UNIVERSITY
OF OSLO

Master's thesis

Dark matter searches in two-lepton final states with missing transverse energy using ATLAS open data

Emily Ania Puchala

Physics: Nuclear and Particle Physics
60 ECTS study points

Department of Physics
Faculty of Mathematics and Natural Sciences



Spring 2026

Emily Ania Puchala

Dark matter searches in two-lepton final states with missing transverse energy using ATLAS open data

Supervisors:
Professor Farid Ould-Saada
Dr. Eirik Gramstad
Dr. Even Simonsen Håland

Abstract

While the Standard Model of particle physics encompasses our current understanding of fundamental particles and interactions, it leaves several open questions, including the nature of dark matter. This thesis presents searches for dark matter candidates using ATLAS Open Data for education with two theoretical frameworks: supersymmetric chargino-neutralino production, and semi-dark Higgs boson decays, both focusing on two-lepton final states and missing transverse energy. The analyses are performed using proton-proton collision data collected by the ATLAS detector at 13 TeV corresponding to 36 fb^{-1}. Event selection criteria and statistical methods are applied to distinguish potential signals from the Standard Model background. No statistically significant excess above the expected background is observed, and the results are consistent with Standard Model expectations. The supersymmetric search sets 95% confidence level upper limit on chargino-neutralino production, excluding masses up to approximately 580 GeV for a massless neutralino and a 95% confidence level upper limit of 0.12 fb on the visible cross section. For the semi-dark Higgs boson decay search, a 95% confidence level upper limit of 3.26 fb is set on the visible cross section. Additionally, this work develops complementary educational Jupyter notebooks designed to bridge the gap between particle physics research and outreach. These notebooks guide users through key aspects of the analyses, enabling students to reproduce select results. By combining research analyses with educational materials, this thesis contributes to both ongoing dark matter searches and outreach and education in particle physics.

Contents

I Introduction **1**

1 Introduction and Motivation . . . 3
- 1.1 Outreach and Education . . . 4
- 1.2 ATLAS Open Data . . . 4
- 1.3 Combining Research with Outreach . . . 5
- 1.4 Thesis Outline . . . 5

II Background **7**

2 Dark Matter and the Standard Model . . . 9
- 2.1 Dark Matter . . . 9
- 2.2 The Standard Model . . . 9
  - 2.2.1 Gauge Theory and Spontaneous Symmetry Breaking . . . 10
  - 2.2.2 Neutrinos . . . 11
- 2.3 Missing Energy . . . 12
- 2.4 Simplified Models . . . 12

3 Beyond the Standard Model . . . 15
- 3.1 Supersymmetry . . . 15
  - 3.1.1 Weakly Interacting Massive Particles . . . 15
  - 3.1.2 Spin and R-Parity . . . 16
  - 3.1.3 Minimal Supersymmetric Standard Model . . . 16
  - 3.1.4 Chargino-Neutralino Production and Decay . . . 18
- 3.2 Exotic Higgs Decay . . . 19
  - 3.2.1 Semi-Dark Higgs Models . . . 20

4 Experimental Applications and Statistics . . . 23
- 4.1 Collider Experiments . . . 23
  - 4.1.1 Circular Colliders . . . 23
  - 4.1.2 The ATLAS Detector . . . 24
  - 4.1.3 Cross Section and Luminosity . . . 24
- 4.2 Statistics . . . 25
  - 4.2.1 Poisson Distribution . . . 25
  - 4.2.2 Approximating Significance . . . 26
  - 4.2.3 Likelihood-based Counting Experiments . . . 26
  - 4.2.4 Exclusion Limits . . . 28

III Methods **29**
5 Physics Analysis . . . 31
5.1 Software and Tools . . . 31
5.2 Data, Signals, and Backgrounds . . . 32
5.2.1 Data . . . 32
5.2.2 MC Simulated Events . . . 32
5.3 Variables . . . 36
5.3.1 Invariant Mass . . . 36
5.3.2 Transverse Mass . . . 37
5.3.3 Stransverse Mass . . . 37
5.3.4 Angular Separation . . . 38
5.4 Event Selection Criteria . . . 38
5.4.1 Lepton Identification Selection Criteria . . . 38
5.4.2 Chargino-Neutralino Production with W and Z Bosons . . . 38
5.4.3 Semi-dark Higgs Decay . . . 41
5.5 Statistical Analysis . . . 42
5.5.1 Uncertainties . . . 42
5.5.2 Bayesian Method . . . 43
5.5.3 Frequentist Method . . . 43
5.6 Limitations . . . 44
6 Educational Materials . . . 45
6.1 Intended Audience . . . 45
6.2 Notebook Structure . . . 45
6.3 Notebook Content . . . 47
6.4 Reproducibility . . . 47

IV Results and Discussion **49**
7 Analysis Results . . . 51
7.1 Standard Model Background Validation . . . 51
7.2 Chargino-Neutralino Production with W and Z Bosons . . . 53
7.2.1 Statistical Interpretation . . . 58
7.3 Semi-Dark Higgs Decay . . . 60
7.3.1 Statistical Interpretation . . . 63
8 Educational Notebooks . . . 65
8.1 Supersymmetry Analysis . . . 65
8.2 Statistical Analysis . . . 66
8.3 Applications . . . 66

V Conclusion **69**
9 Conclusion . . . 71

Appendices **73**
A Frequentist and Bayesian Statistics 75
B Additional Validation Region 79
C Comparison to a Published Study 81
D MC Simulated Event Samples 85
D.1 Background Samples 85
D.2 Signal Samples 86

List of Figures

2.1 The Standard Model of Particle Physics . . . 10
2.2 Missing transverse energy . . . 12
2.3 Simplified dark matter model. . . . 13
2.4 Feynman diagram for mono-X dark matter production . . . 14

3.1 Minimal Supersymmetric Standard Model . . . 17
3.2 Chargino-neutralino production through an s-channel and t-channel . . . 18
3.3 Chargino-neutralino production with W and Z boson decays . . . 19
3.4 Semi-invisible Higgs boson decay . . . 20

5.1 Supersymmetric particle masses for considered chargino-neutralino production signal models . . . 33
5.2 Feynman diagrams for diboson processes . . . 34
5.3 Feynman diagrams for $t\bar{t}$ processes . . . 34
5.4 Feynman diagram for the Drell-Yan process . . . 35
5.5 Feynman diagram for single top quark production process . . . 35
5.6 Distributions of progressive selection requirements for the chargino-neutralino production process . . . 40
5.7 Distributions of progressive selection requirements for the semi-dark Higgs process . . . 42

6.1 Image of the education notebook “Try it yourself!” section . . . 46
6.2 Image of the education notebook “What do you think?” section . . . 46

7.1 Standard Model validation distributions . . . 52
7.2 Comparison of intermediate and high signal region signal efficiencies for chargino-neutralino production . . . 54
7.3 Baseline region distributions of the missing transverse energy and stransverse mass for chargino-neutralino production . . . 56
7.4 Intermediate signal region distributions of the missing transverse energy and stransverse mass for chargino-neutralino production . . . 57
7.5 High signal region distributions of the missing transverse energy and stransverse mass for chargino-neutralino production . . . 57
7.6 Grid of the observed CL_s value for each signal mass hypothesis in the combined signal region for chargino-neutralino production . . . 59
7.7 Observed and expected exclusion contour for chargino-neutralino production 59
7.8 Distributions of missing transverse energy and dilepton transverse mass for the semi-dark Higgs decay process in the baseline region . . . 62

7.9 Distributions of missing transverse energy and dilepton transverse mass for the semi-dark Higgs decay process in the signal region 62

A.1 Comparison of observed and expected 95% confidence limits on the signal strength parameter for chargino-neutralino production using frequentist and Bayesian methods 76
A.2 Comparison of frequentist and Bayesian statistical methods for observed and expected exclusion limits on chargino-neutralino production 76

B.1 Distributions of missing transverse energy and stransverse mass for the validation region for chargino-neutralino production 80

C.1 Missing transverse energy distribution validation region compared to results of a published analysis 81
C.2 Missing transverse energy distribution in the signal region compared to results of a published analysis 82
C.3 Exclusion contour compared to results of a published analysis 82

List of Tables

5.1 Python packages and libraries 31
5.2 Event selection criteria for intermediate and high signal regions for chargino-neutralino production 39
5.3 Event selection criteria for the semi-dark Higgs decay signal region. 41

7.1 Summary of event yields in signal regions for chargino-neutralino production 55
7.2 Representative signal model event yields in signal regions for chargino-neutralino production 58
7.3 Observed and expected event yields and model-independent limits for chargino-neutralino production 60
7.4 Event yields in the signal regions for the semi-dark Higgs decay 61
7.5 Observed and expected limits for the semi-dark Higgs decay process . . . 63

A.1 Comparison of model-independent 95% upper limits using frequentist and Bayesian methods 75
A.2 Comparison of model-independent 95% upper limits with and without uncertainties using frequentist and Bayesian methods 77
A.3 Comparison of 95% upper limits for the semi-dark Higgs process using frequentist and Bayesian methods for chargino-neutralino production . . . 78
A.4 Comparison of model-independent 95% upper limits with and without uncertainties using frequentist and Bayesian methods for semi-dark Higgs decay 78

B.1 Summary of event selection criteria for the validation region targeting chargino-neutralino production 79
B.2 Summary of event yields in the validation region for chargino-neutralino production 80

D.1 MC sample information for the Z+jets processes 85
D.2 MC sample information for the W+jets processes 85
D.3 MC sample information for the diboson processes 86
D.4 MC sample information for the $t\bar{t}$ process 86
D.5 MC sample information for the single top quark processes 86
D.6 MC sample information for the semi-dark Higgs signal sample 86
D.7 MC sample information for negatively charged chargino signal samples . . 87
D.8 MC sample information for positively charged chargino signal samples . . . 88

Preface

Acknowledgments

I would like to begin by thanking my supervisors, Farid, Eirik, and Even, for their support, encouragement, and guidance throughout this project. Each of you has shared your unique perspective and expertise, helping to provide clarity and direction, and I have learned so much over these past two years. Your thoughtful feedback and suggestions at every stage have been extremely helpful, and it has been a pleasure to work with all of you.

I would also like to thank everyone in the high-energy physics section at the University of Oslo, and especially my fellow master's students, for contributing to such a welcoming and enjoyable environment.

Last but not least, I would like to thank my parents and my brother for their constant support. Even though we have lived in different countries, I have always been able to rely on your support and encouragement. Thank you for always believing in me and inspiring me to pursue my dreams.

Declaration of the use of generative artificial intelligence

In this scientific work, generative artificial intelligence (AI) has been used. All data and personal information have been processed in accordance with the University of Oslo's regulations, and I, as the author of the document, take full responsibility for its content, claims, and references. An overview of the use of generative AI is provided below. The service ChatGPT-5 developed by OpenAI [1] has been used in the following ways:

- Systematic language editing
  - Draft text was submitted to the model to improve clarity, flow, and structure. Suggestions for wording and phrasing were reviewed and edited by the author to ensure the final text reflects the author's own voice. Only sentence and paragraph level revisions were considered; full sections of text were not generated for direct use.
- Programming
  - Programming errors and code snippets were submitted to the model to assist with debugging, improving code structure, simplifying implementation, and enhancing readability. Suggested alternative programming approaches were reviewed and implemented where appropriate.

Part I

Introduction

Chapter 1

Introduction and Motivation

Science outreach in physics serves to bridge the gap between advanced research and public understanding, especially for students and educators. Science outreach as a concept may differ from person to person, but it generally aims to promote public awareness and inspire an interest in the subject through activities such as public talks, visiting schools, workshops, and online resources. Within the field of physics, many institutions have outreach initiatives aimed at increasing public awareness and inspiring future scientists.

Physics research, ranging from fundamental studies of particles to applied technologies, has many real-world applications. For example, in medicine, developed technologies have been used for imaging, diagnostic tools, and cancer therapies [2]. Another area that benefited from physics research is computing, where the World Wide Web was originally developed as a way for physicists to communicate and share information quickly across the globe [3]. These few examples highlight the wide range of topics physics research contributes to.

The field of particle physics studies fundamental particles and forces, whose interactions form the Standard Model of particle physics. Theoretical work develops models and frameworks to better understand the Standard Model, and also goes beyond the Standard Model, looking for physics that is not explained by our current understanding. Experimental work, such as collider experiments, is central to validating these theories. The European Organisation for Nuclear Research (CERN) is home to a diverse range of experiments, where the experiments at the Large Hadron Collider (LHC) use detectors to analyse particle collisions [4]. A Toroidal LHC AppartuS (ATLAS) is one of the detectors and is particularly well suited for both physics analyses and outreach due to the availability of the ATLAS Open Data, which provides publicly available collision data for both research and education [5].

Within the ATLAS experiment, there is a range of physics that can be studied, from precision measurements of the Standard Model to new physics searches. This thesis focuses on new physics searches relating to the search for dark matter. Despite making up 27% of the matter in the Universe, it remains one of the mysteries in modern physics research [6]. Since we do not know what makes up dark matter, many theories and models have been proposed that aim to describe what it is and how it may be produced in collider experiments. Among these are supersymmetric extensions of the Standard Model, and models that focus on non-Standard Model signatures in Higgs boson decays, each motivating complementary dark matter search strategies. Supersymmetry theories were proposed in the early 1970s [7] and have been explored for over 50 years. With the discovery of the Higgs boson in 2012 [8, 9], studying exotic Higgs decays offers a slightly

more recent approach in dark matter searches.

This thesis connects physics research with outreach. In addition to conducting physics analyses on supersymmetry and exotic Higgs-motivated dark matter searches, the development of educational materials is a key component. Specifically, educational materials are developed in the form of interactive Jupyter [10] notebooks, which can be used by students to reproduce and explore the analyses on their own. Through the combination of physics analyses and educational material development, I aim to support accessible, hands-on particle physics outreach.

1.1 Outreach and Education

Contributing to physics education and outreach bridges the gap between modern physics research and a wide audience, particularly students and educators. Globally, a variety of institutions have adopted outreach efforts. For example, the American Physical Society [11], a non-profit membership organisation, has an outreach initiative, the Joint Network for Informal Physics Education and Research, focused on engaging with the public through informal physics education and events [12]. Many university physics departments have also developed outreach initiatives and programs, including, but not limited to, workshops, science camps, and public talks or lecture series [13–15]. Although outreach efforts often target the broader field of physics, initiatives within sub-fields like particle physics have also been adopted. The International Particle Physics Outreach Group (IPPOG) [16], a network of scientists, science educators, and communication specialists, works to strengthen the understanding and support of particle physics and "raise the standards" of outreach and informal science education efforts [17].

Different tools can be used to deliver information in outreach initiatives and projects. As highlighted in the examples above, the method of delivery can be tailored to fit the needs of the initiative's specific goals. One such tool is the Jupyter notebook, which has been widely used in research settings and has been adopted in educational environments [18–20]. The development of web-based software and platforms, such as JupyterHub [10], Binder [21], and Google Colab [22] has further improved accessibility. Lowering the barrier for student involvement, the notebook interface eliminates the need for access to specialised software or a powerful personal computer, only requiring an internet connection. The overall appeal of these notebooks in education is their ability to allow students to learn the theoretical concepts, coding practices, and data analysis techniques through a single file.

1.2 ATLAS Open Data

Physicists at the ATLAS experiment have been involved in education and outreach initiatives since the 1990s [23]. The impacts of some of these initiatives are highlighted in a report by Barnett and Johansson in 2023, where they summarise ATLAS outreach projects from 1996 to 2012 [23]. The ATLAS Open Data project is a key resource for making particle physics research more accessible, with the first significant release of collision data made publicly available in 2016 [24]. Their mission "is committed to the progressive release of its data for public use, aiming to support educational initiatives, foster citizen science, and enhance the potential for scientific research" [24].

The project offers data for both research and education, including real collision data as well as Monte Carlo (MC) simulations. As of 2025, the latest educational data release includes the 13 TeV 2025 Data-Beta corresponding to collision data from Run 2 of the

LHC, where data was collected between 2015-2016 [25]. As part of the initiative, the collaboration provides analysis notebooks on a variety of topics, which establish an excellent foundation for students to build upon for their own analyses and interests [26].

One application of using ATLAS Open Data in education is Galaxy Z-Path [27], a workflow portal with some accompanying tools developed at the University of Oslo. Based upon the open-source system, Galaxy [28], Z-path was originally developed in the late 2000s/early 2010s to allow high school students to work with data. The goal of Z-Path is to follow the "heartbeats" of the LHC and allow students to engage with recent discoveries [29]. Used in some physics courses at the University of Oslo, this portal provides an easy-to-use web interface allowing students to run analyses without the need to install or configure software locally.

1.3 Combining Research with Outreach

Focusing on dark matter searches provides a useful starting point for introducing one of the open questions in modern physics. Supersymmetric models and exotic Higgs models have been explored in ATLAS analyses [30, 31], and can serve as examples of typical analyses. This foundation is then expanded to include more advanced topics such as statistical analysis techniques, providing an overview of both the theory needed for a conceptual understanding of the physics analysis and applications to research techniques.

From the physics analyses, the development of interactive Jupyter notebooks forms the educational and outreach aspect. The Jupyter notebooks' dual-purpose approach of performing an analysis and later adapting it to education highlights that these notebooks reflect real research methodologies, while presenting them in a format that is better suited to varying skill levels, thereby providing an introduction to particle physics research.

1.4 Thesis Outline

This thesis is divided into five main parts. Part I introduces the motivation for the physics research and how it is combined with educational outreach, focusing on the use of ATLAS Open Data for education.

Part II provides the theoretical background, including an overview of dark matter, the Standard Model, and physics beyond the Standard Model. It also introduces the experimental applications of collider searches and statistical methods used to interpret results.

The methodologies used in this thesis are presented in Part III, and include the physics analyses using ATLAS Open Data for education, as well as descriptions of the data, Standard Model background processes, selection criteria, and statistical framework. Methods related to the development of the accompanying educational materials are also described.

Part IV presents the results of the analysis, including searches for dark matter candidate particles predicted by supersymmetric and exotic Higgs models, together with their statistical interpretations. It also discusses the developed educational notebooks.

Part V summarises the results and places them in a broader context. It also discusses limitations, possible improvements, as well as pointing to possibilities for future work.

The appendices provide supplementary materials that support the main chapters. For example, Appendix A compares Bayesian and frequentist statistical techniques, while other appendices contain additional results.

Part II

Background

Chapter 2

Dark Matter and the Standard Model

This chapter introduces the theoretical background related to dark matter and the Standard Model. It presents key concepts such as missing transverse energy and simplified models, which form the foundation necessary for physics beyond the Standard Model in Chapter 3.

2.1 Dark Matter

One of the longstanding open questions in physics is the nature of dark matter. While ordinary matter, which can be observed and studied, makes up around 5% of the universe, the remainder is largely unknown. Although dark matter had been referenced in scientific literature prior to the 1930s, Zwicky is often cited as one of the leading pioneers of dark matter research. In 1933, he published work on galaxy clusters suggesting some form of invisible matter contributes to their behaviour [32, 33]. In the 1970s, further evidence emerged from observations of spiral galaxies, providing strong evidence for the existence of dark matter [32].

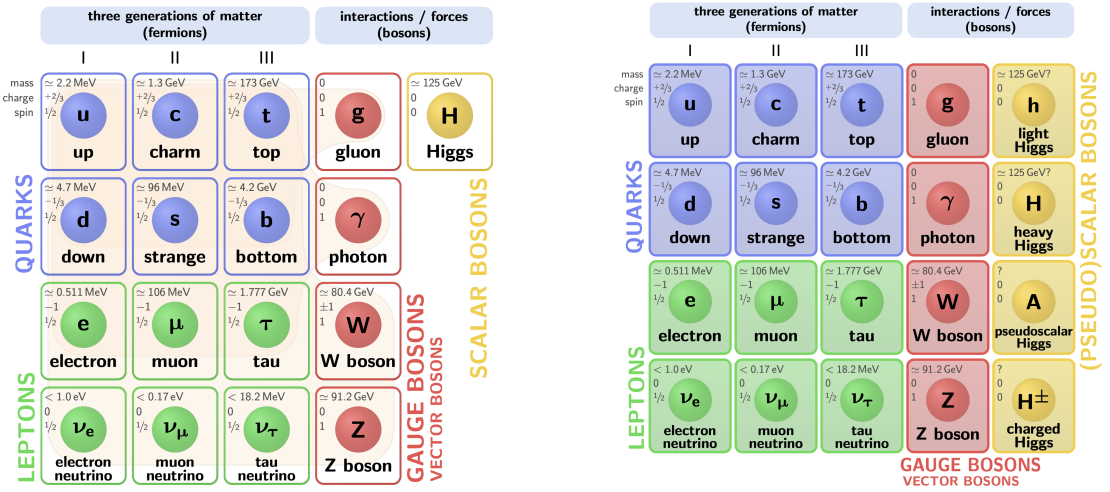
Astrophysical observations provide insight into the key properties of dark matter. Gravitational lensing, where gravity bends light and distorts the appearance of the galaxies, offers strong evidence for its existence [34]. These observations indicate that dark matter interacts via gravity and has been present since the early universe. Comparisons between cosmological observations and models of structure formation, the process by which large-scale structures such as galaxies form, suggest that dark matter is cold. Cold dark matter is non-relativistic and therefore moves slowly compared to the speed of light [35]. In addition, dark matter must be stable, or at least long-lived, to account for its present-day abundance.

These constraints have led to numerous theoretical models, many of which suggest that dark matter may be a particle. As a result, particle physics-based frameworks have been developed to explain its nature. “By the late 1980s, the hypothesis that the missing mass consists of one or more yet-unknown subatomic particle species had gained enough support to become established as the leading paradigm for dark matter” [32]. These properties motivate particle candidates such as weakly interacting massive particles and axions, discussed in later sections.

2.2 The Standard Model

The Standard Model, as shown in Figure 2.1, encompasses our current understanding of three of the four fundamental forces: electromagnetic, weak, and strong forces,

excluding gravity. It classifies the known elementary particles, including fermions, which make up matter, and bosons, which facilitate the interactions. All fundamental fermions participate in weak interactions, while only quarks and gluons participate in the strong interaction, as they carry colour charge in quantum chromodynamics. Leptons, in contrast, do not carry colour charge and therefore do not experience the strong interaction, interacting instead via the electroweak force [36]. The structure of these particle interactions is mathematically described within a gauge theory framework.



(a) The Standard Model

(b) Higgs Extension

Figure 2.1: Figures of (a) the Standard Model of Particle Physics and (b) expanded to include additional Higgs bosons which are needed in theories requiring two Higgs doublets, such as those described in Chapter 3. Figures are adapted from Reference [37], which are based on an original diagram licensed under the Creative Commons Attribution-ShareAlike 4.0 International License [38].

2.2.1 Gauge Theory and Spontaneous Symmetry Breaking

The Standard Model is a gauge theory based on the symmetry group

$$\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$$

which describes the strong, weak, and electromagnetic interactions through the exchange of spin-1 gauge fields [39]. In this formalism, U denotes a unitary group and SU a special unitary group, where the number in parentheses specifies the dimension of the matrices (i.e. an SU(n) group is represented by unitary $n \times n$ matrices with determinant 1). Here $\mathrm{SU}(3)_\mathrm{C}$ describes the quantum chromodynamics and colour charge, while $\mathrm{SU}(2)_\mathrm{L}$ and $\mathrm{U}(1)_\mathrm{Y}$ describe the electroweak interaction, acting on left-handed fermions and weak hypercharge, respectively.

Gauge invariance refers to the requirement that the Lagrangian remains invariant (meaning unchanging) under local transformations of these symmetry groups. Imposing this local symmetry introduces gauge fields, which mediate interactions. A consequence of this is that mass terms for the gauge bosons and fermions are not allowed, implying that the corresponding particles should be massless. While this is consistent with the massless photon and gluons, it does not reflect the experimentally observed masses of the $\mathrm{W}^\pm$ and Z bosons. This indicates that electroweak symmetry is not preserved by

the vacuum, but is instead spontaneously broken. This is achieved through the Brout-Englert-Higgs mechanism, which describes spontaneous symmetry breaking and explains how elementary particles acquire mass through the introduction of a scalar field [36].

In the Standard Model, this is implemented through the introduction of a scalar Higgs doublet, which transforms under the $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ symmetry group. This doublet contains four real degrees of freedom, and is described by a potential that leads to a non-zero vacuum expectation value [39]. In this framework, the Higgs field has a non-zero vacuum expectation value, meaning it remains non-zero even in the lowest energy state. As a result, the electroweak symmetry,

$$\mathrm{SU(2)_L} \times \mathrm{U(1)_Y} \rightarrow \mathrm{U(1)_{QED}}$$

is spontaneously broken, where $\mathrm{U(1)_{QED}}$ corresponds to the unbroken symmetry of the massless photon. The $\mathrm{SU(3)_C}$ symmetry remains unchanged, as it describes the strong interaction and is not affected by electroweak symmetry breaking.

Following symmetry breaking, three of the four degrees of freedom are absorbed by the weak gauge bosons, becoming their longitudinal polarisation states [39]. Through this process, the weak gauge bosons (W^+, W^-, Z) acquire mass, while the photon remains massless. The remaining degree of freedom results in a physical scalar particle, the Higgs boson. This mechanism, therefore, generates masses for the weak gauge bosons while introducing a new scalar particle.

The masses of the gauge bosons are determined by the strength of their coupling to the Higgs field and the value of the vacuum expectation value. The Z and $\mathrm{W^{\pm}}$ bosons acquire masses of approximately 91 GeV and 80 GeV, respectively. These electroweak interactions, together with the Higgs mechanism, play a key role in particle production processes studied at collider experiments at the LHC. In particular, they can produce weakly interacting particles that escape detection.

2.2.2 Neutrinos

Neutrinos provide an example of weakly interacting particles within the Standard Model that escape direct detection in collider experiments. They are neutral leptons that do not experience electromagnetic or strong interactions, properties similar to those of expected dark matter candidates [32]. Although not directly detected at the LHC, their presence can be inferred through the missing momentum and energy measurements. While neutrinos are not considered viable dark matter candidates because of their low mass and high velocities, they provide a useful analogy for studying weakly interacting massive particles. Both are neutral and interact only via the weak force. Neutrinos also demonstrate how weakly interacting particles can escape direct detection in collider experiments, leading to missing energy signatures.

Although the Standard Model has been successful in explaining known particles and interactions, it leaves several physical phenomena unexplained. One notable example is the lack of a dark matter candidate. Despite these limitations, it provides a foundation for extensions beyond the Standard Model. By predicting new weakly interacting particles, these extensions introduce the concept of undetected particles contributing to missing momentum and energy, key features in dark matter searches in collider experiments, as discussed in the next section.

2.3 Missing Energy

Due to the nature of the particles being dark, meaning they cannot be directly observed in collider experiments, we turn to the momentum of the particles. In proton-proton collisions, the protons travel along the beamline and therefore have no net momentum in directions perpendicular to the beam (the transverse plane). As a result, the total momentum in the transverse plane is expected to be zero before the collision and to remain conserved after the collision. If the measured transverse momenta of the visible particles are not balanced, it suggests the presence of additional momentum carried by invisible particles, such as neutrinos. Known as the missing transverse momentum, $\boldsymbol{p}_T^{\text{miss}}$, this imbalance is a key signature in dark matter searches.

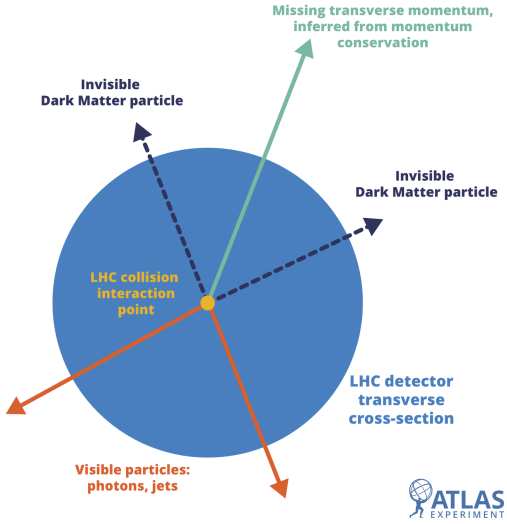


Figure 2.2: Diagram of the cross section showing how missing transverse momentum can be inferred. Image © 2019 CERN for the ATLAS Experiment. Used in accordance with CERN's Conditions of Use for non-commercial purposes [40].

Missing transverse energy, E_T^{miss}, is a key component of dark matter searches in collider experiments, defined as:

$$E_T^{\text{miss}} = |\boldsymbol{p}_T^{\text{miss}}| \tag{2.1}$$

where,

$$\boldsymbol{p}_T^{\text{miss}} = -\left(\sum_{\text{objects}} \boldsymbol{p}_T + \boldsymbol{p}_T^{\text{soft}} \right). \tag{2.2}$$

Here, the $\sum_{\text{objects}} \boldsymbol{p}_T$ includes the transverse momentum of the reconstructed objects, such as leptons, photons, and jets, while $\boldsymbol{p}_T^{\text{soft}}$ is the momentum associated with tracks not associated with these objects. Visualised in Figure 2.2, the missing transverse momentum is inferred from conservation laws.

2.4 Simplified Models

Simplified models provide a starting point for constructing more sophisticated models in the search for dark matter candidate particles. When constructing such models, several

key components are defined: the type of interaction, initial state particles, mediator, and final state particles. The mediator is the force carrier that facilitates the interaction connecting the Standard Model sector to the hypothetical dark matter sector. The initial state consists of Standard Model particles, while the final state may contain either Standard Model particles, dark matter candidates, or a combination of both, depending on the process under consideration.

The three main classes of dark matter searches are indirect detection, direct detection, and collider searches. These are associated with the following processes:

$$\chi\chi \rightarrow \mathrm{SM} + \mathrm{SM}, \tag{2.3}$$

$$\chi + \mathrm{SM} \rightarrow \chi + \mathrm{SM}, \tag{2.4}$$

$$\mathrm{SM} + \mathrm{SM} \rightarrow \chi\chi, \tag{2.5}$$

where SM denotes Standard Model particles, and χ denotes a dark matter candidate. These correspond to annihilation, scattering, and production processes, respectively, and are directly related to the main dark matter search strategies. Indirect detection searches target annihilation products such as gamma rays, cosmic rays, or neutrinos, while direct detection experiments aim to observe dark matter scattering off atomic nuclei [41]. Collider searches, such as those at the LHC, focus primarily on production processes in high-energy collisions. Each detection method explores different mass ranges and interaction strengths, providing complementary approaches.

These processes may proceed through different interaction channels depending on the model and mediator properties. A general example of a simplified model is shown in Figure 2.3, where two Standard Model particles interact via a mediator that can couple either to Standard Model particles or to dark matter particles.

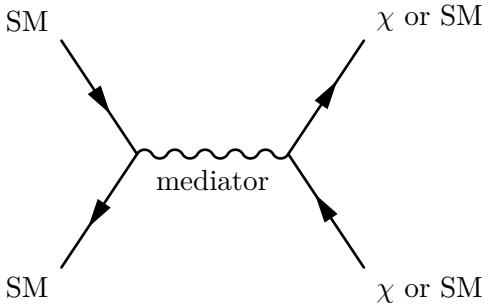


Figure 2.3: Feynman diagram of a simplified dark matter model, where two Standard Model particles (SM) interact through a mediator and result in final states of either Standard Model particle or dark matter candidate particle (χ).

To interpret such signatures, simplified dark matter models are introduced, as minimal extensions of the Standard Model introducing a dark matter candidate particle and a mediator that connects the visible Standard Model and dark sectors. These models are defined by a small number of free parameters: the mass of the dark matter candidate, the mass of the mediator, and the coupling constants at each interaction vertex. The coupling constants describe the strength of the interaction of the mediator with Standard Model particles and dark matter particles.

As a concrete example, in scenarios where the mediator is a new vector boson such as Z', the model can be specified by four parameters: the mediator mass, the dark matter mass, and the two coupling constants [42]. This class of models is commonly used in mono-X searches, where dark matter particles are produced in association with a visible Standard Model particle [42]. An example of such a process is shown in Figure 2.4.

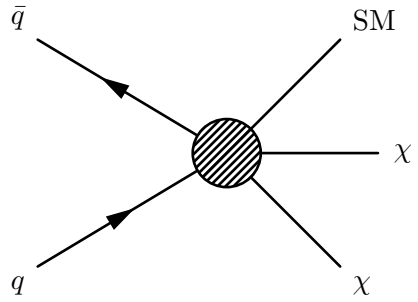


Figure 2.4: Example Feynman diagram for a mono-X dark matter production process, where dark matter particles (χ) recoil against a visible Standard Model (SM) particle.

In collider experiments, the dark matter particles escape detection, resulting in missing transverse energy. When produced in association with visible Standard Model particles, they give rise to characteristic mono-X signatures [43]. This makes missing transverse energy a key signature in searches for physics beyond the Standard Model.

While the simplified model shown in Figure 2.3 does not distinguish between specific particles or mediators, it provides a useful starting point for exploring possible interactions. The framework of simplified models can be extended to study more complex scenarios in beyond the Standard Model theories, discussed in the next chapter.

Chapter 3

Beyond the Standard Model

In this chapter, some of the theoretical frameworks and models used in beyond the Standard Model searches are introduced. These theories aim to describe processes and phenomena that are not explained by the Standard Model, one notable example being dark matter, for which no candidate particle exists within the Standard Model. In this thesis, supersymmetry and exotic Higgs models are introduced as the primary frameworks of interest, as both offer potential explanations for dark matter. Supersymmetry extends the Standard Model by introducing supersymmetric partner particles, while exotic Higgs models explore extensions of the Higgs sector and its decay channels.

3.1 Supersymmetry

To address unresolved phenomena in the Standard Model, several theoretical extensions have been proposed. Although there is currently no experimental evidence for supersymmetry, it remains an attractive framework for addressing limitations of the Standard Model. These include providing a dark matter candidate, addressing the hierarchy problem,[1] and facilitating grand unification of the electroweak and strong forces, among others.

Supersymmetry introduces a symmetry between fermions and bosons, predicting a range of new particles known as superpartners of the Standard Model particles. Among these, the lightest supersymmetric particles, such as the neutralino or sneutrino, are often considered dark matter candidates, as they would fit the expected characteristics of dark matter. This makes supersymmetry an interesting field of research, both for extending the Standard Model and guiding dark matter searches.

3.1.1 Weakly Interacting Massive Particles

Weakly interacting massive particles are hypothesised to be elementary particles that interact only via gravity and the weak force, with an interaction strength comparable to the weak scale. In the early universe, when temperatures were high, these particles would have been in thermal equilibrium and continuously annihilate into lighter particles [44]. As the universe cooled, the number density decreased, and the annihilation rate of these

[1] The hierarchy problem refers to the large discrepancy between the weak force and gravity and the question of why the Higgs boson mass is much lighter than the Planck mass despite quantum corrections. In supersymmetric theories, these corrections are cancelled by contributions from supersymmetric partner particles.

particles fell below the expansion rate of the universe, a process known as thermal freeze-out [44, 45]. This freeze-out left a relic abundance.

Particles with a larger interaction cross section continue to annihilate for a longer time, causing fewer to remain today [46]. Particles with a smaller cross section decouple earlier and leave behind a larger relic abundance. Based on the current abundance of dark matter, if dark matter consists of one of these relic particles, their annihilation cross section is expected to be of the order of the weak interaction scale [46], particularly for particles with masses in the approximate range $10^2 - 10^3$ GeV.

The relic density or abundance describes the amount of particles that remain in the universe today, and determines how many dark matter particles of a given mass would be needed to match the cosmological observations. An approximate solution to the Boltzmann equation gives the abundance as,

$$\Omega_X h^2 \approx \frac{3 \times 10^{-27} \mathrm{cm}^2 \mathrm{s}^{-1}}{\langle \sigma v \rangle} \tag{3.1}$$

where h is the Hubble constant [46, 47]. Using the abundance gives an annihilation cross section of $\langle \sigma v \rangle \approx 3 \times 10^{-26} \mathrm{cm}^3/\mathrm{s}$ [45] which is of the same order of magnitude as the cross sections characteristic of electroweak interactions.

This coincidence is often referred to as the "WIMP Miracle." It suggests that a particle with weak-scale interactions and a mass in the order of $10^2 - 10^3$GeV would naturally result in the dark matter abundance, providing the motivation for such dark matter models [46]. Within supersymmetric frameworks, neutralinos satisfy these properties and are therefore classified as weakly interacting massive particles. This connection between the electroweak scale and cosmological observations is one of the key reasons why weakly interacting particles are considered strong dark matter candidates.

3.1.2 Spin and R-Parity

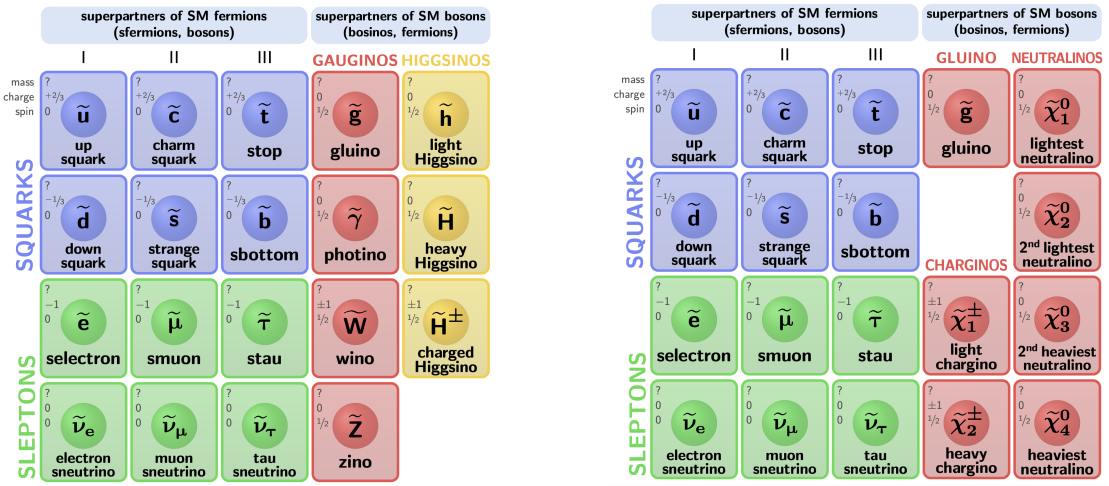
Spin is an intrinsic property of particles, similar to angular momentum. Standard Model particles either have half-integer (e.g., 1/2) or integer (e.g., 0 or 1) spins. In supersymmetric models, each superpartner differs in spin by 1/2 from its corresponding Standard Model particle. For example, leptons with spin 1/2 have superpartners with spin 0, while gauge bosons with spin 1 have superpartners with spin 1/2 [36].

R-parity is defined as $R = (-1)^{3(B-L)+2s}$ where B is baryon number, L is lepton number[2], and s is the spin of the particles. Standard Model particles have an R-parity of +1 (even R-parity) while the supersymmetric particles have an R-parity of -1 (odd R-parity) [48]. If R-parity is conserved, supersymmetric particles would be produced in pairs, and the lightest supersymmetric particle cannot decay into lighter Standard Model particles. This stability makes the lightest supersymmetric particle a great dark matter candidate as it would be neutral, and weakly interacting, with a mass typically in the range of order $10^2 - 10^3$GeV.

3.1.3 Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model is a realisation of supersymmetry that introduces superpartners for all Standard Model particles [48], shown in Figure 3.1.

[2]These quantities refer to the quantum numbers which characterise particle properties. The baryon number is defined as $B = \frac{1}{3}(n_q - n_{\bar{q}})$, where n_q and $n_{\bar{q}}$ are the numbers of quarks and antiquarks, respectively. The lepton number is defined as the difference between the number of leptons and antileptons.



(a) Minimal Supersymmetric Standard Model (b) Chargino-Neutralino Extension

Figure 3.1: The (a) Minimal Supersymmetric Standard Model and its (b) chargino and neutralino extension, adapted from Reference [37]. These figures are based on original diagrams licensed under the Creative Commons Attribution-ShareAlike 4.0 International License [38].

Despite its name, it significantly extends the Standard Model by introducing a large number of additional free parameters, increasing its complexity.

In its most general form, the model contains more than one hundred free parameters, including masses, phases, and mixing angles [49]. This large parameter space makes it difficult to study systematically or constrain experimentally. To address this, the Phenomenological Minimal Supersymmetric Standard Model reduces to 19 parameters, while preserving the key phenomenological features [50]. This reduction enables more focused studies of relevant features. Among the reduced number of parameters, the gaugino mass parameters are especially important for the phenomenology considered here, as they determine the mixing between the gauginos and higgsinos, as well as the gaugino masses.

In the Minimal Supersymmetric Standard Model, the fermionic superpartners of the electroweak gauge bosons and Higgs bosons mix to form electrically neutral states known as neutralinos and charged states known as charginos, as shown in Figure 3.1.

The neutral gauginos consist of the superpartners of the gauge bosons, known as the bino and the neutral wino. After electroweak symmetry breaking, these states can be expressed in terms of the photino and zino, the superpartners of the photon and Z boson, respectively. The neutral gauginos mix with the neutral higgsinos[3] forming neutralinos. The four neutralinos are denoted as $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_3^0$, $\tilde{\chi}_4^0$, where the subscript indicates their increasing mass. Each neutralino is a linear combination of bino, wino, and higgsino. Similarly, the charged wino and charged higgsinos mix to form two charged states, the charginos, denoted as $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^\pm$. The mixing in both the neutralino and chargino states plays a key role in determining how these particles are produced and how they decay.

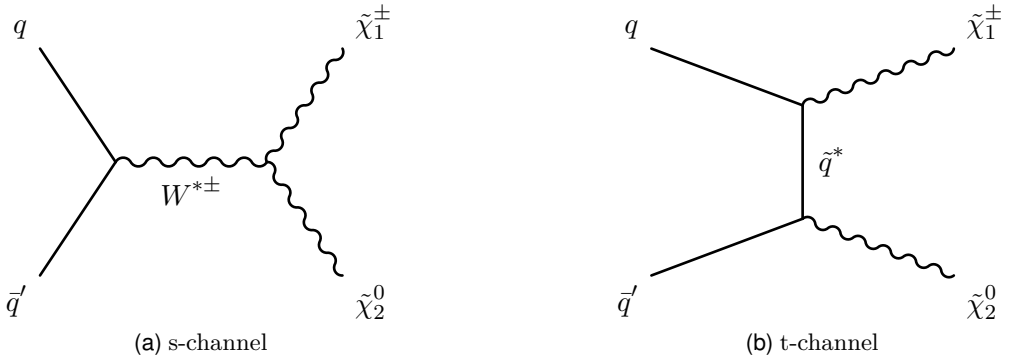
The lightest neutralino is often the lightest supersymmetric particle and one of the leading dark matter candidates. If R-parity is conserved, it is stable and weakly interacting, making it consistent with the expected properties of dark matter [48].

[3] In the Minimal Supersymmetric Standard Model, electroweak symmetry breaking requires two Higgs doublets, giving eight degrees of freedom and resulting in five Higgs bosons: the light Higgs (h), heavy Higgs (H), pseudo-scalar Higgs (A), and two charged Higgs bosons (H^+, H^-).

3.1.4 Chargino-Neutralino Production and Decay

Supersymmetry introduces new particles and extends the interaction structure by allowing additional couplings between Standard Model particles and their superpartners. Their production follows the same underlying principles as Standard Model processes. Chargino and neutralino production, as illustrated in Figure 3.2, is closely related to the electroweak interactions.

In the model considered here, the masses and composition of the charginos and neutralinos are primarily determined by the gaugino and higgsino mass parameters. These parameters determine both the masses and the couplings to the electroweak gauge bosons, which in turn affect the production and decay patterns.



(a) s-channel (b) t-channel

Figure 3.2: Chargino-neutralino production through an (a) s-channel and (b) t-channel. Figures adapted from Reference [51].

Chargino-neutralino production proceeds through both s-channel and t-channel processes. In the s-channel, a quark and an antiquark annihilate to produce an off-shell W boson, which decays into a chargino-neutralino pair. This channel is analogous to the Drell-Yan process, where a lepton pair is produced in hadron-hadron collisions from quark-antiquark annihilation [36]. In the t-channel, a squark is exchanged between two incoming quarks, producing the same final state. The masses of these supersymmetric particles play an important role, where if the squark is very massive, the t-channel is suppressed, making the s-channel the dominant production mechanism [51].

Once produced, charginos and neutralinos decay into Standard Model particles and the lightest neutralino, which escapes detection. This results in missing transverse momentum, a key signature in supersymmetry searches. Depending on the masses and mixing, these decays can produce a range of final states, including combinations of leptons, jets, and missing energy.

In the model considered in Figure 3.3, the mass splitting between the chargino, the next-to-lightest neutralino, and the lightest neutralino allows for decays via W and Z bosons. The positively or negatively charged chargino decays into a W boson and the lightest neutralino, while the heavier neutralino decays into a Z boson and the lightest neutralino. The Z boson decays leptonically, producing two leptons, while the W boson decays hadronically into jets. The lightest neutralinos contribute to missing transverse momentum.

This decay chain leads to final states with two leptons, missing transverse energy, and at least two jets, which defines the signature targeted in this analysis. The kinematic properties of these final states are influenced by the underlying supersymmetric parameters, particularly the gaugino-higgsino mixing and the mass spectrum of the charginos and neutralinos.

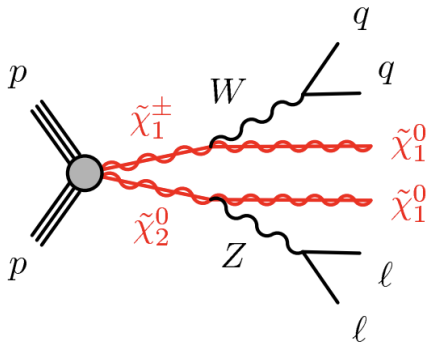


Figure 3.3: Diagram of chargino and neutralino production with W and Z boson decays. Image sourced from Reference [30].

Such signatures are also relevant in other searches for physics beyond the Standard Model. While supersymmetry offers a well-motivated framework for dark matter searches, it is not the only approach.

3.2 Exotic Higgs Decay

These final-state signatures also appear in other extensions of the Standard Model, including models with an extended Higgs sector that introduce new particles and decay channels. These exotic Higgs models provide complementary search strategies and can also be useful in dark matter searches, with decay modes not present in the Standard Model.

The Higgs boson decays can be categorised based on the visibility of their final states [52] as fully visible, fully invisible, and semi-dark:

$$H \to ZZ^* \to 4\ell, \tag{3.2}$$
$$H \to ZZ^* \to 4\nu, \tag{3.3}$$
$$H \to \chi\bar{\chi}, \tag{3.4}$$
$$H \to Z + \chi. \tag{3.5}$$

The process in equation (3.2) is one example of a fully visible decay of a Standard Model Higgs boson. Fully invisible decays can arise within the Standard Model, for example, in equation (3.3), where the final-state neutrinos escape detection. In addition, fully invisible decays may also occur in beyond the Standard Model scenarios, such as in equation (3.4). In contrast, semi-dark decays, equation (3.5), contain both visible and invisible components. In these cases, the Higgs boson decays into a Standard Model particle, such as the Z boson, and an invisible particle denoted by χ. The visible decay enables detection, while missing transverse momentum and energy associated with χ may indicate the presence of new physics.

The identity of the invisible particle depends on the model. Possible candidates include dark photons and axion-like particles, which interact weakly with Standard Model particles. Dark photons are hypothetical spin-1 gauge bosons, while axion-like particles are spin-0 pseudoscalars [35]. While both can produce similar experimental signatures, their underlying properties differ.

Although the Higgs boson has spin 0, conservation of angular momentum is maintained through contributions from orbital angular momentum [36]. This ensures

overall conservation and allows a scalar particle to decay into final states containing particles with non-zero spin, such as the Z boson and dark photon. In this thesis, the focus is on models where the Higgs boson decays into a Z boson and a dark photon, discussed in the next section.

3.2.1 Semi-Dark Higgs Models

Semi-dark decays combine visible and invisible signatures. One example of this is a process where a Higgs boson decays into a Z boson and a dark photon (γ_d). The Z boson produces two leptons, forming the visible final state, while the dark photon contributes to the missing transverse energy. This process is represented by,

$$H \to Z + \gamma_d \to \ell^+\ell^- + E_T^{\text{miss}}, \tag{3.6}$$

and shown in Figure 3.4.

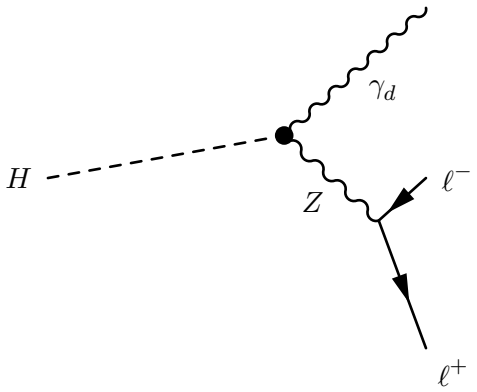


Figure 3.4: Feynman diagram showing an exotic Higgs boson (H) decay where a Z boson (Z) and dark photon (γ_d) are produced. The Z boson decays into two leptons ($\ell^+\ell^-$), providing a visible final state, while the dark photon is assumed to be invisible, leading to missing transverse energy.

This type of interaction arises in models where the Standard Model, as introduced in Section 2.2.1, is extended by an additional abelian gauge symmetry associated with a hidden sector. In this framework, the Standard Model gauge group is extended by a new symmetry, $\mathrm{U(1)_d}$, where the subscript refers to the dark sector. The gauge symmetry is therefore extended to

$$\mathrm{SU(3)_C} \times \mathrm{SU(2)_L} \times \mathrm{U(1)_Y} \times \mathrm{U(1)_d}. \tag{3.7}$$

This extension introduces the dark photon as a new gauge boson.

The dark photon is a hypothetical gauge boson that can interact with Standard Model particles through kinetic mixing with a visible photon [53]. Kinetic mixing describes an interaction between gauge bosons of two different gauge groups and is common in beyond the Standard Model frameworks [54]. This mechanism provides a portal between the visible and hidden sectors, allowing weak coupling between them. The interaction strength is determined by the kinetic mixing parameter, while the mass determines the kinematic properties of the decay.

The specific model considered in this analysis is characterised by two free parameters: the dark photon mass and the kinetic mixing parameter. The kinetic mixing parameter

determines the coupling strength between the dark photon and the Standard Model particles, while the dark photon mass determines the kinematic behaviour of the decay.

The coupling between the Higgs boson, Z boson, and dark photon arises after electroweak symmetry breaking through this mixing mechanism [55]. This allows semi-dark Higgs decays, such as the one described previously. Compared to supersymmetric models, described in Section 3.1, dark photon models are relatively minimal, typically characterised by two parameters [53]. This simplicity makes them an attractive alternative framework for dark matter searches.

With the theoretical frameworks established, the next step is to examine how these models can be tested using collider experiments and statistical methods discussed in Chapter 4.

Chapter 4

Experimental Applications and Statistics

This chapter introduces collider experiments and how particle accelerators and detectors can be used to test theoretical models. It also presents statistical techniques that quantify the significance of the results obtained from experimental searches.

4.1 Collider Experiments

High-energy particle accelerators play a key role in testing theoretical models in particle physics. The largest and most powerful of these is the LHC, a circular proton-proton collider with a circumference of 27 kilometres, situated approximately 100 metres underground at CERN on the France-Switzerland border. It accelerates particles to very high energies before colliding them at interaction points, enabling studies of particle interactions. These collisions can produce both known particles and potentially new particles predicted by theoretical models.

There are several experiments at the LHC, including four major detectors: ATLAS, CMS, ALICE, LHCb, each designed to study different physics processes. In this work, we focus on the ATLAS detector, which is a general-purpose detector designed to explore a wide range of physics processes. The descriptions of circular colliders and the ATLAS detector presented in the following sections are primarily based on References [56–58].

4.1.1 Circular Colliders

The main components of a circular particle accelerator include: vacuum systems, magnets, and radio frequency cavities. The vacuum systems prevent particles from colliding with gas molecules, maintaining a controlled environment for the collisions. Magnets are used to direct the trajectory of particles around the accelerator tracks and focus the beam to collide particles at collision points. Radio frequency cavities with oscillating fields accelerate the protons to the desired energy and ensure they are tightly grouped. These components work together to keep particles moving around the accelerator, building up their energy.

The beams of particles travel in opposite directions around the tracks before being directed towards colliding at a specific interaction point, where detectors are positioned to record the resulting events. These collisions can produce a range of particles, whose properties are measured by the detector placed around the interaction point.

4.1.2 The ATLAS Detector

The ATLAS detector is a general-purpose detector composed of layers of specialised sub-detectors arranged around one of the LHC interaction points. Each subsystem is designed to measure properties of particles, such as trajectory, energy, or identity. The main components of the detector include: tracking systems, magnet systems, calorimeters, and a muon spectrometer.

Starting with the innermost region, the tracking systems consist of a pixel detector, a semiconductor tracker, and a transition radiation tracker. The pixel detector and semiconductor tracker are based on silicon technology, which provides measurements of particle trajectories close to the interaction point. The transition radiation tracker uses drift tubes filled with gas, where charged particles passing through ionise the gas, producing a detectable signal. This provides tracking information together with the pixel and semiconductor tracker detectors, and also helps to distinguish between particle types. The region is surrounded by a solenoidal magnetic field, which bends the charged particle trajectories. From these curved trajectories, the momentum and charge of the particles can be determined.

Located outside the solenoid magnet are the calorimeters, which measure the energy deposited by particles as they pass through and interact with the material. The electromagnetic calorimeter, composed primarily of liquid argon, measures the energy of electrons and photons via electromagnetic showers. Next is the hadronic calorimeter, which is designed to absorb and measure the energy of hadrons like protons and pions. These particles typically pass through the electromagnetic calorimeter and deposit their energy through hadronic interactions.

The outermost layer is the muon spectrometer, which detects muons. This system includes toroidal magnets which provide the magnetic field for momentum measurements, along with monitored drift tubes, which are filled with a gas mixture, allowing their trajectories to be determined.

Given the high rate of collisions, not all events contain interesting characteristics for further study. Trigger systems, implemented as both hardware and software, are used to identify and retain only relevant events for further analysis.

By combining the information from all detector layers, particles can be identified and their properties, such as energy and momentum, can be reconstructed. Particles that do not interact directly with the detector, such as neutrinos or hypothetical weakly interacting particles, cannot be observed directly. Instead, their presence is inferred through missing transverse momentum, calculated from the imbalance in the measured momentum of visible particles, as outlined earlier in Section 2.3.

4.1.3 Cross Section and Luminosity

Two key quantities used to characterise how likely and how frequently particle interactions take place are the cross section and luminosity.

The cross section σ measures the probability of a given particle interaction to take place. It is defined as,

$$\sigma = \frac{\text{number of interactions per unit time per target particle}}{\text{incident flux}}, \tag{4.1}$$

where flux accounts for the relative motion of the particles [36]. Simply, a larger cross section means an interaction is more likely to occur.

The differential cross section describes how the scattering probability is distributed with respect to a chosen variable. For a distribution given in terms of the solid angle, the differential cross section is defined as,

$$\frac{d\sigma}{d\Omega} = \frac{\text{number of particles scattered into } d\Omega \text{ per unit time per target particle}}{\text{incident flux}}, \quad (4.2)$$

where $d\sigma$ represents the cross section, and $d\Omega$ represents the solid angle, describing the angular distribution of particles after a collision [36].

The instantaneous luminosity $\mathcal{L}$ is defined as the number of interactions per unit area per unit time [59]. The integrated luminosity, which is often used when analysing data, is obtained by integrating the instantaneous luminosity over the full data-taking period,

$$\mathcal{L}_{\text{int}} = \int \mathcal{L} dt. \quad (4.3)$$

The luminosity is related to the interaction rate through the cross section,

$$N_{\text{interactions/sec}} = \mathcal{L}_{\text{int}} \cdot \sigma. \quad (4.4)$$

where σ is the total scattering cross section. These values help determine the performance of the accelerator, where a higher luminosity implies there is a greater chance of particle interactions occurring, while a larger integrated luminosity corresponds to a greater total dataset.

When discussing cross sections in particle physics, the most common unit of area is the barn (b), where 1 b = 10^{-28}m^2. Since typically cross sections are much smaller than this, femtobarn (fb), where 1 fb = 10^{-15}b or 10^{-43}m^2, is often used. Integrated luminosity is commonly expressed in units of inverse femtobarn fb^{-1}.

4.2 Statistics

Statistical methods are introduced to test experimental results by quantifying the agreement between observed data and theoretical predictions. These methods provide a framework to determine the significance of deviations between predicted and observed event yields, and set exclusion limits on the considered signal models. The statistical techniques presented in this section are primarily based on the approach of Cowan et al. [60], supplemented with concepts from References [61–63].

4.2.1 Poisson Distribution

In counting experiments, the number of observed events is assumed to follow a Poisson distribution with a mean,

$$\lambda = s + b, \quad (4.5)$$

where s is the expected signal yield and b is the expected background yield. The probability of observing n events is,

$$P(n;\lambda) = \frac{\lambda^n e^{-\lambda}}{n!}. \quad (4.6)$$

This probability distribution forms the basis of the statistical methods described in the following sections.

4.2.2 Approximating Significance

At high event counts, the Poisson distribution can be approximated by a Gaussian distribution, allowing for a simple approximation of the significance. The significance, Z, compares the observed data and the background-only hypothesis. In high statistics, meaning the background yield is large, the significance is approximated as:

$$Z = \frac{n_{\text{obs}} - b}{\sqrt{b}}, \tag{4.7}$$

where n_{obs} is the number of observed events. To account for systematic uncertainties in the background estimate, the uncertainty, δ, is introduced:

$$Z = \frac{n_{\text{obs}} - b}{\sqrt{b + (\delta b)^2}}. \tag{4.8}$$

For low event counts, a more accurate expression better suited for low-statistics analyses is used, where the significance is given by

$$Z = \sqrt{2\left[n_{\text{obs}} \ln\left(\frac{n_{\text{obs}}}{b}\right) - n_{\text{obs}} + b\right]}. \tag{4.9}$$

These expressions define the observed significance, which compares the observed event counts to the background-only hypothesis. The significance corresponds to the number of standard deviations the observed results deviate from the expected background.

While these approximations provide a starting point to determine significance, a more robust statistical analysis is employed for hypothesis testing and limit setting, relying on likelihood-based statistical techniques.

4.2.3 Likelihood-based Counting Experiments

Likelihood

The likelihood function quantifies how compatible a signal hypothesis is with observed data under a given model. For a single counting experiment, the likelihood under the signal hypothesis s is,

$$L(s) = P(n|s + b). \tag{4.10}$$

where b is the expected background, and n is the observed number of events.

In a binned analysis, the number of observed events in each bin i follows a Poisson distribution,

$$P(n_i|\lambda_i) = \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}, \tag{4.11}$$

where n_i is the observed number of events, and λ_i is the expected number of events in bin i.

The signal strength parameter μ scales the expected signal prediction to maximise the likelihood. It is defined as the ratio of the observed signal to the predicted signal. Introducing this parameter, we find,

$$\lambda_i = \mu s_i + b_i. \tag{4.12}$$

and the full likelihood over N bins is,

$$L(\mu) = \prod_{i=1}^{N} P(n_i|\mu s_i + b_i). \tag{4.13}$$

This model is expanded to incorporate systematic uncertainties from nuisance parameters, $\boldsymbol{\theta}$. Including the nuisance parameters gives the expected yield in each bin,

$$\lambda_i(\mu, \boldsymbol{\theta}) = s_i(\mu, \boldsymbol{\theta}) + b_i(\boldsymbol{\theta}). \tag{4.14}$$

The extended likelihood becomes,

$$L(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} P(n_i|\lambda_i(\mu, \boldsymbol{\theta}))\, C(\boldsymbol{\theta}), \tag{4.15}$$

where $C(\boldsymbol{\theta})$ represents terms that constrain the systematic uncertainties.

Hypothesis Testing

Hypothesis testing is performed using the general likelihood ratio test statistic,

$$q_\mu = -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{4.16}$$

where $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ maximises the likelihood for a fixed μ, and $\hat{\mu}, \hat{\boldsymbol{\theta}}$ maximises the likelihood over both signal strength and nuisance parameters. The denominator represents the best-fit description of the data, while the numerator evaluates a specific signal hypothesis.

For discovery, the background only hypothesis ($\mu = 0$) is compared to the signal plus background hypothesis ($\mu > 0$) using

$$q_0 = -2 \ln \frac{L(\mu = 0, \hat{\boldsymbol{\theta}})}{L(\hat{\mu}, \hat{\hat{\boldsymbol{\theta}}})}. \tag{4.17}$$

A one-sided constraint on the test statistic is applied, ensuring $\mu \geq 0$,

$$q_0 = \begin{cases} -2 \ln \lambda(0) & \text{if } \hat{\mu} \geq 0 \\ 0 & \text{if } \hat{\mu} < 0. \end{cases} \tag{4.18}$$

A small q_0 indicates data compatible with the tested hypothesis, while a large q_0 suggests incompatibility.

The p-value quantifies the probability of observing data at least as extreme as what is measured under the null hypothesis (H_0),

$$p_0 = P(q_0 \geq q_0^{\text{obs}}|H_0). \tag{4.19}$$

P-values are often expressed as a significance in units of Gaussian standard deviations,

$$Z = \Phi^{-1}(1 - p_0), \tag{4.20}$$

where Φ^{-1} is the inverse cumulative distribution of a standard Gaussian function. For large datasets, Wilks' theorem [64] allows the asymptotic approximation:

$$Z_0 \approx \sqrt{q_0}, \tag{4.21}$$

$$p_0 \approx 1 - \Phi(Z_0). \tag{4.22}$$

4.2.4 Exclusion Limits

Exclusion tests determine which signal models are incompatible with observed data. Two statistical approaches are evaluated, both based on the likelihood function described in Section 4.2.3.

Frequentist Method

In a frequentist framework, exclusion limits are derived using the modified confidence level CL_s, which accounts for both the signal-plus-background and background-only hypotheses:

$$CL_{s+b} = P(q_\mu \geq q_\mu^{\text{obs}}|\mu), \tag{4.23}$$

$$CL_b = P(q_\mu \geq q_\mu^{\text{obs}}|0), \tag{4.24}$$

$$CL_s = \frac{CL_{s+b}}{CL_b}. \tag{4.25}$$

A signal hypothesis is excluded at the 95% confidence level if $CL_s < 0.05$. Using the asymptotic approximations, this can be expressed in terms of p-values and significance, where,

$$Z_\mu \approx \sqrt{q_\mu}, \tag{4.26}$$

$$p_\mu \approx 1 - \Phi(Z_\mu). \tag{4.27}$$

Under asymptotic approximations to compute the expected behaviour under the background-only hypothesis, we use the Asimov dataset, which represents the expected data under the background-only model. Using this dataset, we compute the corresponding test statistic, $q_{\mu,A}$. The modified confidence level becomes

$$CL_s = \frac{1 - \Phi(\sqrt{q_\mu})}{\Phi(\sqrt{q_{\mu,A}} - \sqrt{q_\mu})}, \tag{4.28}$$

where the numerator corresponds to CL_{s+b} and the denominator corresponds to CL_b, as expressed in equation (4.25).

Bayesian Method

Compared to profiling used in the frequentist approach, the Bayesian approach marginalises over nuisance parameters. This approach constructs a posterior probability and treats the signal strength as a probability distribution. The 95% upper limit s_{up} is defined by,

$$\int_0^{s_{up}} p(s|n_{\text{obs}})ds = 0.95, \tag{4.29}$$

meaning there is a 95% probability that s is below s_{up} and a 5% probability that s is above s_{up}.

The theory and experimental strategies presented in this part are applied to the search for dark matter. The corresponding methodologies are described in Part III and form the foundation of both the research itself and the development of the educational materials.

Part III

Methods

Chapter 5

Physics Analysis

This chapter presents the methodologies used to perform the analyses presented in this thesis. It describes the software and tools employed, processing of data, derivation of analysis variables, event selection criteria, and statistical techniques applied, as well as limitations inherit in the analyses. These methods not only provide the foundation of the physics analyses themselves, but also inform the development of the educational notebooks described in Chapter 6, which adapt these methodologies into suitable workflows for education and outreach.

5.1 Software and Tools

A Jupyter notebook in a python-based environment is chosen as the primary platform used throughout the analysis. This environment is chosen for a simple transition from analysis code to educational materials. The relevant libraries and packages with their purposes are described in Table 5.1.

In the context of ATLAS Open Data, the `atlasopenmagic` Python package simplifies accessing and handling the datasets. It provides utilities and metadata for tasks such as selecting the appropriate data release, caching files locally, choosing skims, and retrieving dataset URLs [65]. Since the data and simulated MC samples are stored in ROOT ntuples, the `uproot` library is used to read these files and convert them into Python-friendly structures for analysis [66].

Category	**Package/Library**	**Purpose**
Data Processing	`atlasopenmagic` [65]	Retrieving ATLAS Open Data files
	`uproot` [66]	Reading ROOT files
	`awkward` [67]	Representing columnar data structures
Calculations	`vector` [68]	Four-vector and momentum calculations
	`mt2` [69]	Computation of the stransverse mass m_{T2}
	`numpy` [70]	Numerical functions
	`scipy` [71]	Statistical calculations
Visualisation	`matplotlib` [72]	Plot generation and customisation
	`pandas` [73]	Data export and organisation

Table 5.1: Python packages and libraries with their purposes.

5.2 Data, Signals, and Backgrounds

5.2.1 Data

The data [74] used throughout this analysis were collected by the ATLAS detector during the 2015-2016 collection periods at $\sqrt{s} = 13$ TeV. The integrated luminosity for the 2015 data period is 3.24 ± 0.04 fb^{-1}, and for the 2016 data period it is 33.40 ± 0.30 fb^{-1} [75]. The combined integrated luminosity for these datasets is 36.0 fb^{-1}.

This analysis uses the 13 TeV 2025 Data-Beta release of the ATLAS Open Data for education [76], which includes collections, or analysis skims, which are subsets of the data that apply a predefined selection corresponding to specific final states. The "2lep" analysis skim is used, containing events with at least two leptons, each with transverse momentum, $p_T > 7$ GeV.

No additional requirements are imposed, for this 2lep skim, on observables such as jet multiplicity or missing transverse momentum. However, reconstructed physics objects satisfy general kinematic criteria: electrons are required to have $p_T > 7$ GeV and pseudorapidity, $|\eta| < 2.47$, muons must have $p_T > 7$ GeV and $|\eta| < 2.5$, and small-R jets are required to have $p_T > 20$ GeV and $|\eta| < 2.5$ [76].

These criteria highlight the pre-selection requirements defined by the chosen skim, while the final event selection requirements used in this analysis are described in subsequent sections.

5.2.2 MC Simulated Events

MC simulated events are used to model both background and signal processes. Tables with the simulated MC sample information, for the background and signal processes, are included in Appendix D. These simulations are produced using a variety of event generators, each suited for the process they are simulating. To ensure that the MC simulations accurately reflect the data conditions, scale factors are applied to correct and normalise the weights of the simulated events.

The cross section weights, normalised by the sum of weights, are multiplied by the MC weights and corresponding scale factors. This gives the total event weight, defined as

$$w_{event} = \frac{(\sigma \cdot \varepsilon_{\text{filter}} \cdot k_{\text{fac}}) \cdot \mathcal{L}_{\text{int}} \cdot w_{\text{MC}}}{\sum w} \times \text{SF}, \tag{5.1}$$

where σ is the cross section, $\varepsilon_{\text{filter}}$ is the filter efficiency, k_{fac} is the k-factor which applies higher order corrections to the cross section, $\mathcal{L}_{\text{int}}$ is the integrated luminosity, w_{MC} is the individual MC event weight, $\sum w$ is the sum of MC weights, and SF is the product of all the applied scale factors. The applied scale factors account for pile-up reweighting, electron and muon identification and isolation efficiency, lepton trigger efficiency, flavour-tagging efficiency for identifying jets originating from specific quark flavours, such as b-jets, and jet vertex tagger efficiency for suppressing jets associated with pile-up interactions.

Signal Processes

The signal models considered in this analysis include two-lepton final states with missing transverse energy. The supersymmetric signal model considered is chargino-neutralino production with W and Z bosons, shown in Figure 3.3. The chargino and neutralino each decay into the lightest neutralino, contributing to missing transverse energy in the

final state. The Z boson decays leptonically, producing the two-lepton final state, while jets arise from the hadronically decaying W boson.

Although the chargino-neutralino production process remains the same, multiple supersymmetric particle mass hypotheses are considered in this analysis. The individual simulated signal model samples correspond to different mass splittings between the chargino/second-lightest neutralino and corresponding lightest neutralino. The resulting grid of mass points is shown in Figure 5.1.

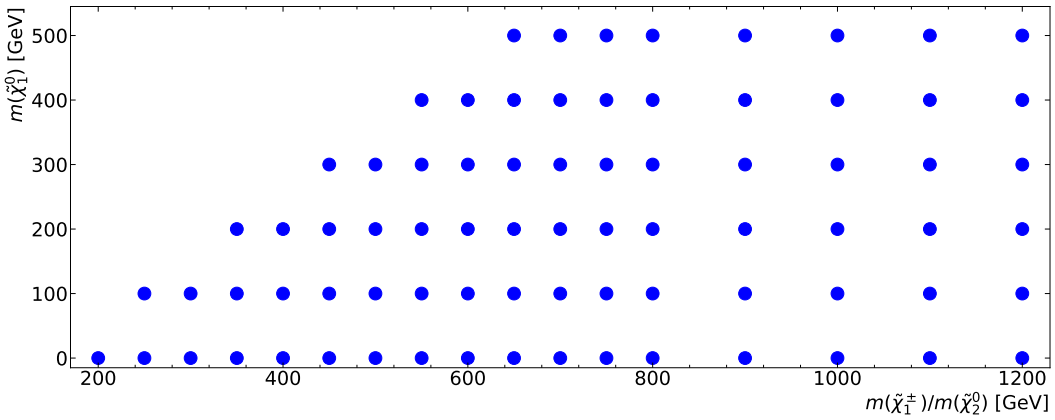


Figure 5.1: Supersymmetric particle masses for each considered chargino-neutralino production signal model. The chargino/neutralino mass is denoted by $m(\tilde{\chi}_1^\pm/\tilde{\chi}_2^0)$ while the corresponding lightest neutralino mass is denoted by $m(\tilde{\chi}_1^0)$.

The semi-dark Higgs decay signal model differs from the supersymmetric scenario due to the production of a Higgs boson through gluon-gluon fusion, resulting in a similar final state, as shown in Figure 3.4. In this model, the Higgs boson decays into a Z boson, contributing to a two-lepton final state, and a dark photon, contributing to missing transverse energy signatures. The final states of the signal models determine which background processes are relevant to this analysis.

Background Processes

Standard Model processes producing final states with two leptons and missing energy signatures closely mimic the signal models we aim to identify, making background reduction critical in the analysis.

The main background processes considered in this analysis include: Z+jets, W+jets, diboson, top quark pair production ($t\bar{t}$), and single top production, specifically in decay channels leading to final states with two leptons and missing transverse energy. Among these processes, the largest backgrounds typically arise from the diboson, $t\bar{t}$, and Z+jets processes.

Diboson production involves the production of two electroweak bosons, which can mimic the signal by producing final states with two leptons and missing transverse energy. In this analysis, only the diboson processes that mimic the signal signature are considered. These processes include final states with two Z bosons, two W bosons, or a W and a Z boson, with specific decay channels relevant to this analysis. For the ZZ processes, one Z boson decays to charged leptons while the other decays to neutrinos, leading to missing transverse energy,

$$q\bar{q} \to ZZ \to \ell^+\ell^-\nu\bar{\nu}. \tag{5.2}$$

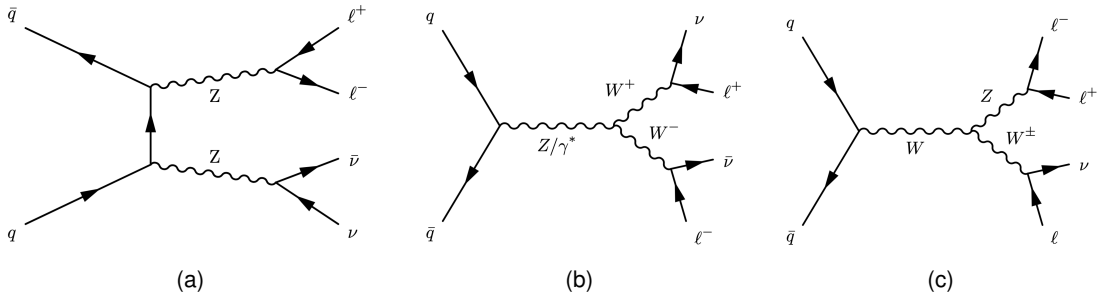


Figure 5.2: Feynman diagrams for the (a) ZZ, (b) WW, and (c) WZ diboson processes.

For the WW processes, both W bosons decay leptonically, resulting in two leptons and missing transverse energy,

$$q\bar{q} \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}. \tag{5.3}$$

In WZ processes, the W boson decays leptonically into a lepton and a neutrino, producing missing transverse energy, while the Z boson decays into a lepton pair. This typically leads to a three-lepton final state,

$$q\bar{q} \to W^\pm Z \to \ell^+\ell^-\ell\nu, \tag{5.4}$$

but can also contribute to the two-lepton final state if one lepton is not reconstructed. These diboson processes are visualised by the Feynman diagrams in Figure 5.2.

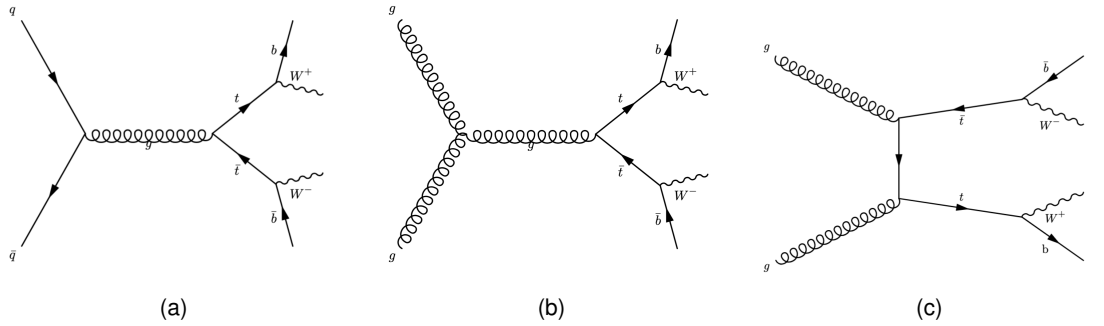


Figure 5.3: Feynman diagrams for production of $t\bar{t}$ through (a) quark-antiquark annihilation, (b) gluon fusion through an s-channel, and (c) gluon fusion through a t-channel.

In $t\bar{t}$ processes, each top quark decays almost exclusively to a W boson and a b-jet. In this analysis, dileptonic $t\bar{t}$ decays are relevant since they produce the two-lepton plus missing transverse energy signature considered. When both W bosons decay leptonically, the final state contains two leptons, missing transverse energy from neutrinos, and multiple jets. This process is typically dominant in high missing transverse energy regions, but can be reduced by vetoing events with b-tagged jets. Two primary production mechanisms contribute to $t\bar{t}$ production: quark-antiquark annihilation,

$$q\bar{q} \to t\bar{t} \to W^+bW^-\bar{b} \to \ell^+\nu b\ell^-\bar{\nu}\bar{b}, \tag{5.5}$$

and gluon fusion,

$$gg \to t\bar{t} \to W^+bW^-\bar{b} \to \ell^+\nu b\ell^-\bar{\nu}\bar{b}. \tag{5.6}$$

These processes are visualised by the Feynman diagrams in Figure 5.3.

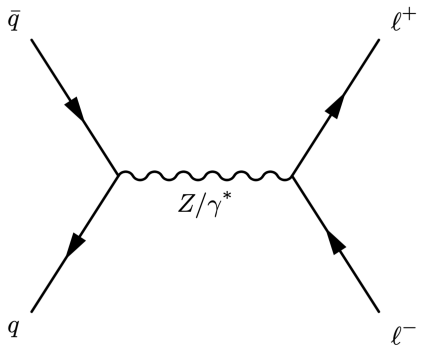


Figure 5.4: Feynman diagram for the Drell-Yan process.

In this analysis, the Drell-Yan process contributes to the background when it satisfies the opposite-sign same-flavour criteria and acquires apparent missing transverse energy through detector effects or neutrinos. The Drell-Yan process is given by,

$$q\bar{q} \to Z/\gamma^* \to \ell^+\ell^-, \tag{5.7}$$

with the corresponding Feynman diagram shown in Figure 5.4. Z bosons can also be produced in association with jets (Z+jets), such as

$$q\bar{q} \to Zg, \tag{5.8}$$

$$q\bar{q} \to Zq. \tag{5.9}$$

In these processes, jets from quantum chromo-dynamic interactions can lead to missing transverse energy signatures due to jet mismeasurement, as well as neutrinos from weak decays of hadrons originating from b-quarks and c-quarks. This background is typically dominant in low missing transverse energy regions.

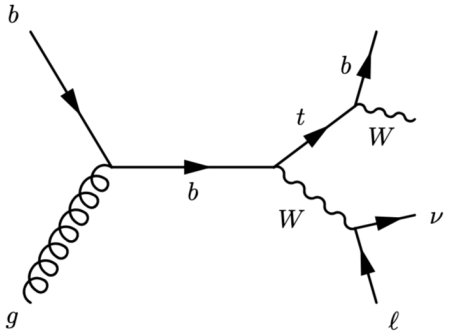


Figure 5.5: Feynman diagram for a single top quark production process.

Standard Model processes, such as single top production and W+jets production, have smaller contributions to the overall backgrounds. Single top processes contribute to the background when their decay products include leptons and neutrinos, leading to final states compatible with the two-lepton plus missing transverse energy selection. Single top quark production results in a final state containing a top quark and is often accompanied by W bosons and b-jets. Additional leptons can arise from semi-leptonic decays of b-hadrons or from misidentified (fake) leptons. These processes are described by,

$$qb \to q't \to q'W^{\pm}b \to \ell^{\pm}\nu b, \tag{5.10}$$

and

$$gb \to tW^{\pm} \to bW^{+}W^{-} \to b\ell^{+}\nu\ell^{-}\bar{\nu}. \tag{5.11}$$

A Feynman diagram for the processes described in equation (5.11) is shown in Figure 5.5. W+jets production contributes to the background through events where the W boson decays leptonically, producing a charged lepton and a neutrino, which can lead to missing transverse energy. An additional reconstructed lepton can arise from jet misidentification. Higgs boson production in processes such as ZH or WH can also lead to backgrounds with final states containing two leptons and missing transverse energy.

As both the background and signal processes can lead to the same final state, distinguishing between them is a key aspect of the analysis. This is achieved by exploiting kinematic differences between the event topologies.

5.3 Variables

To discriminate signal and background processes, kinematic variables are constructed to characterise the events. These variables highlight differences in the kinematic properties of the final states and are used when defining the event selection criteria.

5.3.1 Invariant Mass

The invariant mass describes the mass of an object or system independent of its overall motion. It is particularly useful in reconstructing resonances that decay into visible final state particles. In this analysis, it is beneficial in discriminating processes such as Z boson decays, which produce a characteristic peak in the two-lepton invariant mass distribution. The invariant mass of the two-lepton system is defined as,

$$m_{\ell\ell}^2 = m_1^2 + m_2^2 + 2(E_1E_2 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_2), \tag{5.12}$$

where m_1 and m_2 are the masses, E_1 and E_2 are the energies, and $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are the momenta of the first and second lepton, respectively.

To express the invariant mass in terms of angles, we define pseudorapidity as,

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right], \tag{5.13}$$

where θ is the polar angle. Pseudorapidity describes the angle of a particle relative to the beam axis. The azimuthal angle ϕ describes the angle around the beam axis. Expressing the invariant mass in terms of azimuthal angle, pseudorapidity, and transverse momentum (p_T), in the limit $E \gg m$ gives,

$$m_{\ell\ell}^2 = 2p_{T1}p_{T2}(\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)). \tag{5.14}$$

The dijet invariant mass follows the same formula as in the lepton invariant mass in equation (5.12) but is constructed from the energy and momentum of the jets, rather than that of the leptons.

5.3.2 Transverse Mass

The transverse mass is useful for discriminating events containing one visible particle and one invisible particle inferred from missing transverse energy. For example, when a W boson decays into a lepton and a neutrino, it would produce a signature with one visible lepton and missing transverse energy arising from the neutrino. The transverse mass is defined as

$$m_T^2 = m_\ell^2 + 2\left(E_T^\ell E_T^{\text{miss}} - \boldsymbol{p}_T^\ell \cdot \boldsymbol{p}_T^{\text{miss}}\right), \tag{5.15}$$

where m_ℓ is the lepton mass, E_T^ℓ is the transverse energy of the lepton, $\boldsymbol{p}_T^\ell$ is the transverse momentum vector, and E_T^{miss} is missing transverse energy.

By neglecting the lepton's mass ($m_\ell \approx 0$), which implies $E_T^\ell = p_T^\ell$, this expression reduces to

$$m_T^2 = 2\left(p_T^\ell E_T^{\text{miss}} - \boldsymbol{p}_T^\ell \cdot \boldsymbol{p}_T^{\text{miss}}\right). \tag{5.16}$$

This can be expressed in terms of the azimuthal angle difference $\Delta\phi$ between the lepton and missing transverse momentum, where the transverse mass becomes,

$$m_T^2 = 2p_T^\ell E_T^{\text{miss}}(1 - \cos\Delta\phi). \tag{5.17}$$

This equation is used in the case of a single lepton [77].

The formalism above can be generalised to systems consisting of more than one visible particle. For example, for a dilepton system, the two leptons are combined to form, in this scenario, a visible Z boson using its invariant mass m_Z, transverse momentum p_T^Z, and transverse energy $E_T^Z = \sqrt{(p_T^Z)^2 + m_Z^2}$. In this case, the full expression in equation (5.15) is used to include the non-zero invariant mass of the dilepton system.

5.3.3 Stransverse Mass

Similar to the transverse mass, the stransverse mass is useful for discriminating events with visible and invisible particles, but extends the concept to systems with two visible and two invisible particles. The stransverse mass is a key variable used in supersymmetric analyses, where invisible particles are produced in pairs. It is defined as,

$$m_{T2} = \min_{\boldsymbol{q}_T}\Big[\max\Big(m_T(\boldsymbol{p}_T^{\ell 1}, \boldsymbol{q}_T),\, m_T(\boldsymbol{p}_T^{\ell 2}, \boldsymbol{p}_T^{\text{miss}} - \boldsymbol{q}_T)\Big)\Big], \tag{5.18}$$

where $\boldsymbol{p}_T^{\ell 1}$ and $\boldsymbol{p}_T^{\ell 2}$ are the transverse momentum vectors of the two leptons, and $\boldsymbol{q}_T$ minimises the larger of the two transverse masses defined in equation (5.16) [77, 78].

Implementation

In this analysis, the stransverse mass [77] was implemented using the `mt2` Python package [69]. The function uses the mass and transverse momentum components for each visible particle, the missing transverse momentum vector, and the assumed masses of the invisible particles. The mass of the invisible particle is set to 0 GeV, assuming a massless particle.

Some research suggests setting the mass of the invisible particle, $m_{\text{invis}} = 100$ GeV, to increase sensitivity [79]. When running a test analysis for the chargino-neutralino production models, a slightly higher sensitivity was found when using $m_{\text{invis}} = 0$ GeV, compared to $m_{\text{invis}} = 100$ GeV. This is consistent with studies using $m_{\text{invis}} = 0$ GeV for chargino models [79] and is not further investigated here.

5.3.4 Angular Separation

The angular separation between two objects is defined as,

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \tag{5.19}$$

where $\Delta\eta$ is the difference in pseudorapidity and $\Delta\phi$ is the azimuthal angle difference. The same equation is used for both lepton pairs and jet pairs.

5.4 Event Selection Criteria

The event selection criteria for chargino-neutralino production with W and Z boson decays closely follow the methodologies outlined in Reference [30]. Since the final state remains unchanged across the models, many of the selection requirements are directly applicable to other signal models. The semi-dark Higgs decay is loosely based on the existing framework in the supersymmetric chargino-neutralino analysis, but modified to account for characteristics of the semi-dark Higgs decay. In addition, fully invisible Higgs models focusing on ZH production [31, 80] were used as a reference to guide and motivate the selection criteria.

5.4.1 Lepton Identification Selection Criteria

This analysis focuses on final states containing two same-flavour, opposite-sign leptons, e^+e^- or $\mu^+\mu^-$, where e denotes an electron and μ denotes a muon. Events are required to satisfy a lepton trigger to identify them as either electrons or muons. Electrons are required to satisfy $|\eta| < 2.47$, while muons satisfy $|\eta| < 2.7$. Both the leading and sub-leading leptons are required to have transverse momentum $p_T^{\ell} > 25$ GeV.

To ensure the leptons originate from the primary interaction vertex, identification requirements must satisfy constraints on the track impact parameters. For electrons, $|d_0|/\sigma(d_0) < 5$ and $|z_0 \sin\theta| < 0.5$ millimetres are required. For muons, the corresponding requirements are $|d_0|/\sigma(d_0) < 3$ with $|z_0 \sin\theta| < 0.5$ millimetres and $|\eta| < 2.4$.

5.4.2 Chargino-Neutralino Production with W and Z Bosons

In addition to the lepton identification selection criteria described in Section 5.4.1, further requirements are applied to define signal regions related to the chargino-neutralino production with W and Z bosons.

The invariant mass of the dilepton system satisfies a requirement where the invariant mass, $m_{\ell\ell}$, is between $81 - 101$ GeV to select events associated with the Z boson decay. The invariant mass of the dijet system, m_{jj}, is selected to be within $70 - 100$ GeV corresponding to the mass of the W boson decay. The stransverse mass is required to satisfy $m_{T2} > 100$ GeV to suppress $t\bar{t}$ and WW diboson background contributions. In events where both W bosons decay leptonically, producing a pair of leptons and missing transverse momentum from neutrinos, the value of m_{T2} is bounded by the mass of the W boson, so this requirement rejects such events. The transverse momentum of the Z boson, p_T^Z, from the two leptons must be greater than 80 GeV. The transverse momentum of the W boson, from the jets, is required to satisfy $p_T^W > 100$ GeV.

Events are required to have at least two jets. The two leading jets require $p_T^{\text{jet}} > 30$ GeV, while non-leading jets satisfy $p_T^{\text{jet}} > 20$ GeV, with all jets required to be within $|\eta| < 2.5$. Events containing any b-tagged jets with $p_T^{\text{jet}} > 20$ GeV are rejected. Jets are

identified as b-tagged jets using the DL1dv01 algorithm at a working point with 77% b-tagging efficiency [81, 82].

Jets with $p_T^{\text{jet}} < 60$ GeV and $|\eta| < 2.4$ are additionally required to satisfy the jet vertex tagger [83] requirement to suppress pile-up jets. The angular separation between the two leptons satisfies $\Delta R_{\ell\ell} < 1.8$. Similarly, the two leading jets satisfy $\Delta R_{jj} < 1.5$. The azimuthal angle between the missing transverse momentum and the W boson, $\Delta\phi_{p_T^{\text{miss}},W}$, is required to be within the range $0.5 - 3.0$.

Two signal regions are defined, differing in the missing transverse energy requirement. The intermediate signal region requires $E_T^{\text{miss}} > 150$ GeV, while the high signal region requires $E_T^{\text{miss}} > 250$ GeV. The high region is a subset of the intermediate region, and the two regions are therefore not mutually exclusive at this selection level. A summary of the selection criteria is provided in Table 5.2.

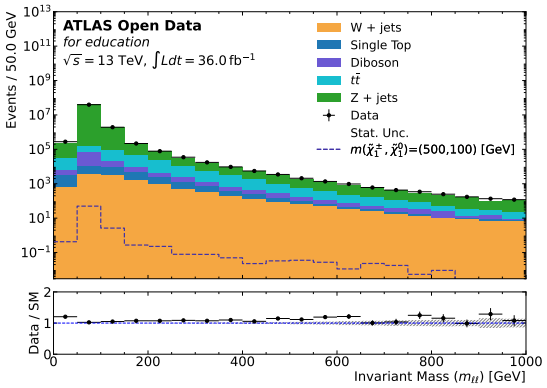
	Intermediate SR	**High SR**
E_T^{miss} [GeV]	> 150	> 250
$m_{\ell\ell}$ [GeV]	$[81, 101]$	$[81, 101]$
m_{jj} [GeV]	$[70, 100]$	$[70, 100]$
m_{T2} [GeV]	> 100	> 100
p_T^Z [GeV]	> 80	> 80
p_T^W [GeV]	> 100	> 100
ΔR_{jj}	< 1.5	< 1.5
$\Delta R_{\ell\ell}$	< 1.8	< 1.8
$\Delta\phi_{(p_T^{\text{miss}},W)}$	0.5 - 3.0	0.5 - 3.0
n_{jets}	≥ 2	≥ 2

Table 5.2: Summary of event selection criteria for intermediate and high signal regions (SR) for chargino-neutralino production with W and Z bosons. The intermediate signal region requires a missing transverse energy greater than 150 GeV, while the high signal region requires a missing transverse energy greater than 250 GeV.

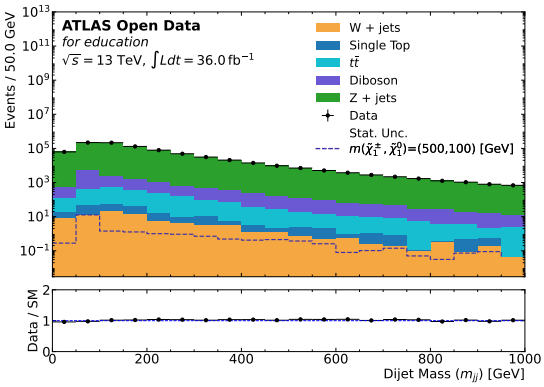
Kinematic Distributions and Selection Progression

The progression of the event selection criteria for chargino-neutralino production with W and Z bosons is illustrated through a series of kinematic distributions shown in Figure 5.6. These distributions show the impact of successive selection requirements on the data, background contributions, and a representative signal model. Presenting the distributions at intermediate stages of the selection provides additional insights into how the chosen kinematic variables discriminate between signal and background processes, motivating the selection requirements summarised in Table 5.2. Some distributions are shown after multiple requirements have been applied.

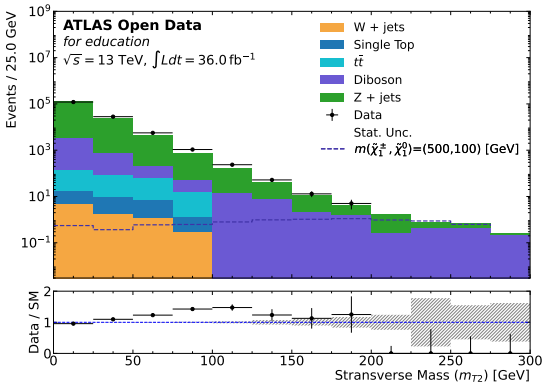
Figure 5.6a shows the dilepton invariant mass distribution after basic lepton selection criteria are applied. The requirement $81 < m_{\ell\ell} < 101$ GeV selects events consistent with the decay of a Z boson. After applying the invariant mass and jet-related requirements,



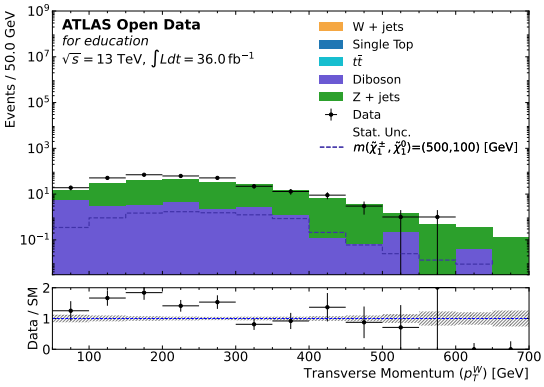
(a) Invariant Mass



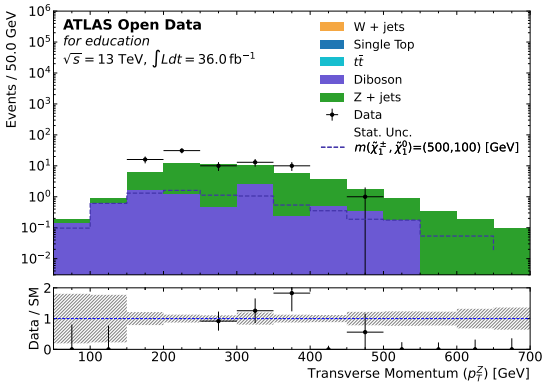
(b) Dijet Mass



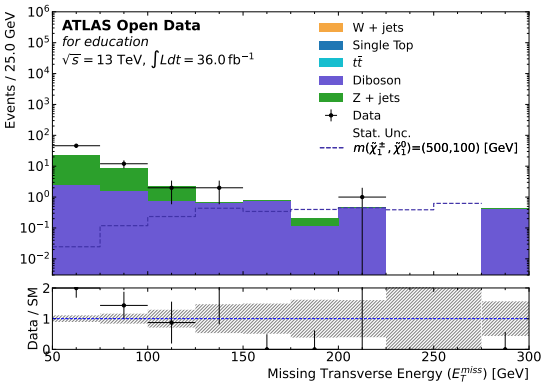
(c) Stransverse Mass



(d) Transverse Momentum of W Boson



(e) Transverse Momentum of Z Boson



(f) Missing Transverse Energy

Figure 5.6: Distributions of the (a) invariant mass, (b) dijet mass, (c) stransverse mass, (d) transverse momentum of W boson, (e) transverse momentum of Z boson, and (f) missing transverse energy at successive stages of the event selection. A representative signal model is included for comparison with the data and Standard Model background contribution.

the resulting dijet mass distribution is shown in Figure 5.6b, where the requirement on m_{jj} selects events corresponding to a hadronically decaying W boson.

The resulting stransverse mass distribution after the dilepton and dijet mass selections is shown in Figure 5.6c. The requirement $m_{T2} > 100$ GeV suppresses contributions from $t\bar{t}$ and WW production. After applying the stransverse mass requirement, the transverse momentum distribution of the reconstructed W boson is

shown in Figure 5.6d. Similarly, Figure 5.6e shows the transverse momentum distribution of the reconstructed Z boson after the additional requirements on the W boson transverse momentum, jet angular separation, and the azimuthal angle between the missing transverse momentum and the W boson are applied.

Finally, Figure 5.6f shows the missing transverse energy distribution after all remaining selection criteria, excluding the requirement on the missing transverse energy. This distribution forms the baseline selection used in the results presented in Section 7.2. Together, these distributions illustrate the effect of the selection requirements and demonstrate how the chosen kinematic variables improve the discrimination between signal and background processes.

5.4.3 Semi-dark Higgs Decay

The lepton identification selection criteria described in Section 5.4.1 are applied to the semi-dark Higgs decay model, with the exception of the momentum of the sub-leading lepton, which satisfies $p_T^{\ell} > 20$ GeV. Requirements for additional selection criteria are as follows. The invariant mass of the dilepton system is required to satisfy a requirement where the invariant mass is between $81 - 101$ GeV to select events associated with the Z boson decay. The transverse momentum of the dilepton system must satisfy $p_T^Z > 80$ GeV. The angular separation between the two leptons must satisfy $\Delta R_{\ell\ell} < 1.8$.

Events with two or more jets with $p_T^{\text{jet}} > 30$ GeV and $|\eta| < 2.5$ are rejected to reduce backgrounds where missing transverse energy may arise from jet mismeasurements and enhance signal contribution from gluon-gluon fusion production of the Higgs boson. The veto of events with b-tagged jets and the jet vertex tagger requirement are identical to the selection requirements defined for the chargino-neutralino production process in Section 5.4.2.

A signal region is defined by the requirements on the missing transverse energy and the dilepton transverse mass, where events in this region require $E_T^{\text{miss}} > 100$ GeV, and $91 \leq m_T^{\ell\ell,E_T^{\text{miss}}} \leq 200$ GeV. The lower bound of the dilepton transverse mass is motivated by the kinematic features of the Z boson decay, where the dilepton system originates, and the upper bound corresponds to the region where the signal rapidly decreases. These selection requirements are summarised in Table 5.3.

	Signal Region
E_T^{miss} [GeV]	> 100
$m_{\ell\ell}$ [GeV]	$[81, 101]$
$m_T^{\ell\ell,E_T^{\text{miss}}}$ [GeV]	$[91, 200]$
p_T^Z [GeV]	> 80
$\Delta R_{\ell\ell}$	< 1.8
n_{jets}	≤ 1

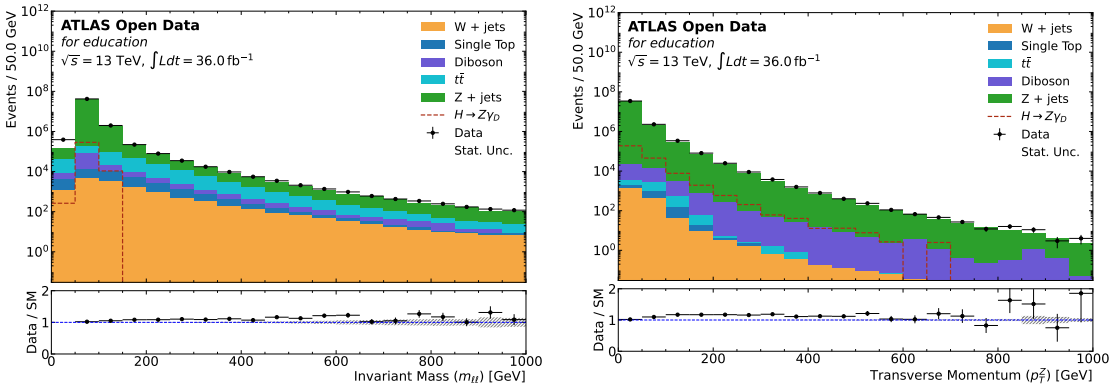
Table 5.3: Summary of event selection criteria for the semi-dark Higgs decay signal region.

Kinematic Distributions and Selection Progression

The progression of the event selection criteria for the semi-dark Higgs decay model is illustrated through the kinematic distributions shown in Figure 5.7. Similar to the chargino-neutralino production case discussed in Section 5.4.2, these distributions demonstrate the effect of the successive selection requirements on the data, background contributions, and signal model. Presenting the distributions at intermediate stages of the selection highlights how well the kinematic variables discriminate between background and signal, motivating the event selection requirements used in the analysis.

Figure 5.7a shows the dilepton invariant mass distribution after basic lepton selection criteria are applied. After applying the invariant mass and jet-related selection requirements, the resulting transverse momentum distribution of the reconstructed Z boson is shown in 5.7b.

The remaining selection requirements are subsequently applied to define the corresponding baseline and signal regions discussed further in Section 7.3. Together, these distributions illustrate the effect of the event selection criteria and demonstrate how the chosen kinematic variables can be used to discriminate between signal and background processes.



(a) Invariant Mass

(b) Transverse Momentum of Z Boson

Figure 5.7: Distributions of the (a) invariant mass, and (b) transverse momentum of the Z boson with progressive selection requirements applied in each distribution.

5.5 Statistical Analysis

The statistical framework described in section 4.2 is applied to quantify the agreement between the observed data and the Standard Model prediction and to determine exclusion limits for each signal hypothesis.

5.5.1 Uncertainties

The statistical uncertainty for the observed data in each bin is,

$$\sigma_{\text{data}} = \sqrt{N_{\text{events}}}. \tag{5.20}$$

For MC simulated backgrounds, the statistical uncertainty per bin is

$$\sigma_{\text{background}} = \sqrt{\sum_{i=1}^{N_{\text{events}}} w_i^2}, \tag{5.21}$$

where w_i corresponds to the event weight. Additionally, a flat 30% systematic uncertainty is applied to the background and signal yields to account for experimental and theoretical uncertainties not explicitly evaluated in this analysis[1].

5.5.2 Bayesian Method

For each signal mass hypothesis, a Bayesian statistical method is employed to calculate upper limits on the visible cross section using the luminosity, signal efficiency, background yields, and observed events. The limits on the visible cross section are given in terms of the cross section times branching fraction, σB. This reflects that the analysis measures a final state where a Z boson decays to a pair of leptons, and accounts for the fraction of events corresponding to this decay channel. This limit is used to convert to upper limits on the expected event yields given by

$$N_{\text{exp}} = \epsilon_{\text{sig}} \sigma B \mathcal{L}_{\text{int}} + N_{\text{bkg}}, \tag{5.22}$$

where $\mathcal{L}_{\text{int}}$ is the integrated luminosity and N_{bkg} is the expected number of background events [84]. The signal efficiency is the fraction of signal events that satisfy the event selection criteria given by

$$\epsilon_{\text{sig}} = \frac{N_{\text{after selection}}}{N_{\text{initial generated}}}. \tag{5.23}$$

The upper limits for the expected number of signal events are compared to the signal event yields for each signal hypothesis. Expected limits and uncertainty bands are obtained from pseudo-experiments under the background-only hypothesis, providing an estimate of the expected distribution without relying on asymptotic approximations, particularly suitable in low-statistics regimes.

This Bayesian limit calculation was implemented using a computational framework developed and provided as part of the course materials [85] for the Research-based Particle Physics course at The University of Oslo. The underlying statistical methodology of the framework remains unchanged, while custom code was developed for the extraction and visualisation of limits.

5.5.3 Frequentist Method

The 95% upper confidence limits on the signal strength parameter, μ_{95}, are derived for each signal hypothesis using the CL_s method (Section 4.2.4). The corresponding upper limit production cross section is

$$\sigma_{95} = \mu_{95} \cdot \sigma_{\text{theory}}, \tag{5.24}$$

where σ_{theory} is the theoretical cross section. Similarly, 95% confidence upper limit on the expected number of signal events is given by,

$$S_{95} = \mu_{95} \cdot S_{\text{exp}}. \tag{5.25}$$

Significance and p-values are calculated using asymptotic approximations from Wilks' theorem. The CL_s values are similarly computed for each signal model under the asymptotic approximation.

[1] This uncertainty is intended to account for contributions from calibration and modelling uncertainties which are not individually considered in this analysis.

5.6 Limitations

Due to the nature of the ATLAS Open Data for education being designed for educational purposes, the dataset faces limitations due to simplicity of use being prioritised over precision [25]. These simplifications limit the precision with which published analyses can be reproduced using open data. While similar methodologies are adopted to mirror existing search strategies and motivate the selection criteria, exact agreement would not be expected.

Reproducing signal regions from published analyses enables a comparison between the behaviour and sensitivity of the models, especially when both datasets correspond to the same luminosity. This approach provides a well-motivated foundation, both in comparing the ATLAS Open Data for education to published results and exploring additional signal models without a direct reproduction.

Systematic uncertainties are treated in a simplified manner throughout this analysis. The use of a flat 30% systematic uncertainty for both signal and background yields does not account for process-dependent effects arising from variations in detector performance or theoretical modelling. As a result, individual uncertainty sources and their correlations between processes are not explicitly modelled.

Despite these limitations, this analysis demonstrates the methodologies and characteristics of searches involving two-lepton final states and missing transverse energy using ATLAS Open Data for education.

Chapter 6

Educational Materials

This chapter describes the methodologies used to develop educational Jupyter notebooks based on the physics analyses previously described. It focuses on design decisions, coding practices, and pedagogical features aimed at supporting active learning in physics.

The educational notebooks are adapted from the code originally developed for the physics analysis. This adaptation ensures that each notebook contains all the necessary elements needed for a student[1] to conduct their own analysis. In addition to including the necessary code segments, the notebooks integrate explanations of physics concepts alongside well-commented code that clarifies the purpose and functionality of each section.

The structure and presentation style of the notebooks are inspired by the publicly available analysis notebooks provided alongside the ATLAS Open Data for education [26, 86]. Selected code segments from these reference notebooks were adapted to fit the context of this analysis. While certain segments of this code remain unchanged where appropriate, the overall implementation of the analysis was developed by the author while adopting a similar stylistic and pedagogical framework.

6.1 Intended Audience

The notebooks are designed for students with some background in physics, ranging from advanced high school to university levels. They accommodate varying levels of prior knowledge in both physics and programming. At a basic level, students can follow the code to reproduce examples and understand the workflow. At an advanced level, students can modify and extend the code, promoting a deeper understanding of both the physics concepts and programming practices.

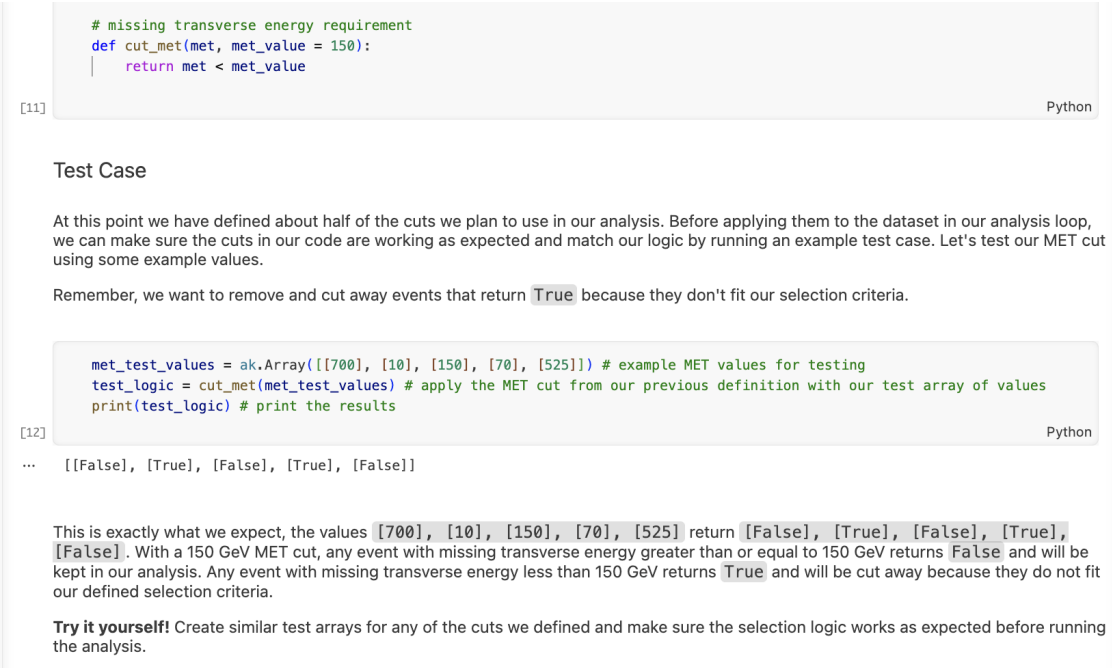
6.2 Notebook Structure

Each notebook contains two main types of cells: markdown and Python. Markdown cells are text cells, which explain concepts more in-depth than an inline comment can convey. Python cells contain executable code with inline comments for additional clarity.

To encourage active learning, the notebooks include sections such as “Try it yourself!” and “What do you think?” The “Try it yourself!” sections, as shown in Figure 6.1, prompt students to interact with the code by modifying examples. The “What do you think?” sections, as shown in Figure 6.2, provide questions for students to consider and

[1] Here, “student” refers to anyone using the notebook.

reflect on when interpreting the results to deepen their understanding. These features are designed to guide students through all aspects of the analysis, from foundational understanding of the physics and programming practices to interpenetrating and engaging with results.



```python
# missing transverse energy requirement
def cut_met(met, met_value = 150):
    return met < met_value
```

[11] Python

Test Case

At this point we have defined about half of the cuts we plan to use in our analysis. Before applying them to the dataset in our analysis loop, we can make sure the cuts in our code are working as expected and match our logic by running an example test case. Let's test our MET cut using some example values.

Remember, we want to remove and cut away events that return `True` because they don't fit our selection criteria.

```python
met_test_values = ak.Array([[700], [10], [150], [70], [525]]) # example MET values for testing
test_logic = cut_met(met_test_values) # apply the MET cut from our previous definition with our test array of values
print(test_logic) # print the results
```

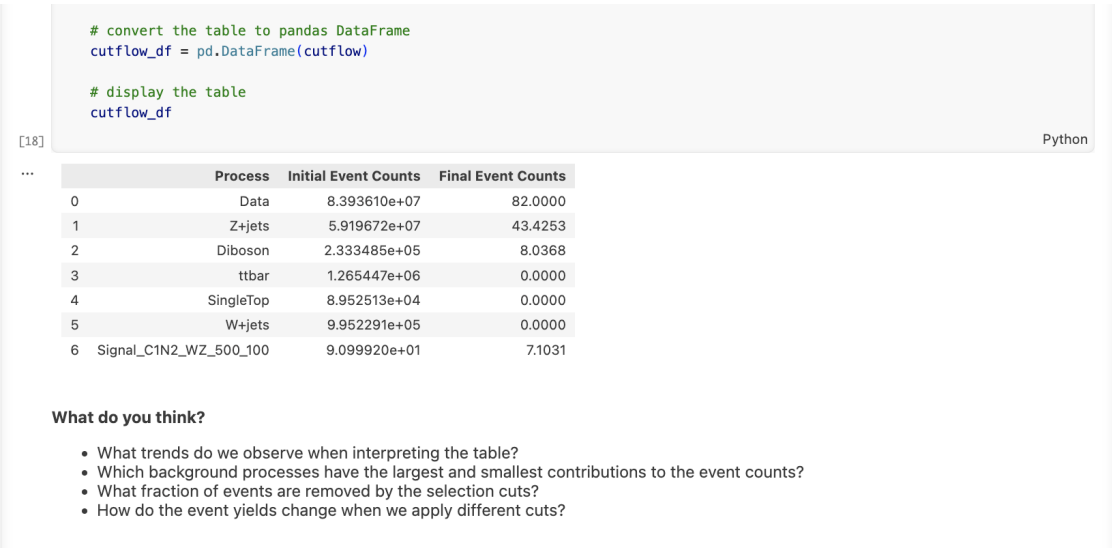
[12] Python

```
[[False], [True], [False], [True], [False]]
```

This is exactly what we expect, the values `[700], [10], [150], [70], [525]` return `[False], [True], [False], [True], [False]`. With a 150 GeV MET cut, any event with missing transverse energy greater than or equal to 150 GeV returns `False` and will be kept in our analysis. Any event with missing transverse energy less than 150 GeV returns `True` and will be cut away because they do not fit our defined selection criteria.

Try it yourself! Create similar test arrays for any of the cuts we defined and make sure the selection logic works as expected before running the analysis.

Figure 6.1: Image of a “Try it yourself” section. A selection cut is defined as code, the purpose of a test case is described, the code to run the test case is included with inline comments, the output is displayed, and a suggestion is provided for how students can interact with the code using similar functionality.



```python
# convert the table to pandas DataFrame
cutflow_df = pd.DataFrame(cutflow)

# display the table
cutflow_df
```

[18] Python

	Process	Initial Event Counts	Final Event Counts
0	Data	8.393610e+07	82.0000
1	Z+jets	5.919672e+07	43.4253
2	Diboson	2.333485e+05	8.0368
3	ttbar	1.265447e+06	0.0000
4	SingleTop	8.952513e+04	0.0000
5	W+jets	9.952291e+05	0.0000
6	Signal_C1N2_WZ_500_100	9.099920e+01	7.1031

What do you think?

- What trends do we observe when interpreting the table?
- Which background processes have the largest and smallest contributions to the event counts?
- What fraction of events are removed by the selection cuts?
- How do the event yields change when we apply different cuts?

Figure 6.2: Image of a "What do you think?" section. The output of a cutflow table is included after the code block, and the section provides questions to guide the student in interpreting the results.

6.3 Notebook Content

Two complementary notebooks support the full physics analysis workflow. The first notebook conducts a supersymmetry analysis, while the second notebook provides an overview of statistical techniques used to interpret the results obtained in the first notebook.

The supersymmetry analysis notebook introduces students to the analysis workflow and guides them in producing results, including cutflow[2] tables, histograms of key kinematic variables, and a simple statistical interpretation. It demonstrates loading data and necessary libraries, defining necessary variables, applying the selection requirements, and visualising results. Each section is structured to follow the analysis steps in a reproducible format.

The statistical analysis notebook introduces the statistical concepts needed to interpret analysis results. It demonstrates how to define statistical models and likelihood functions, incorporate nuisance parameters and uncertainties, perform hypothesis testing, and set exclusion limits. While this notebook builds on the results of the supersymmetry analysis, it presents the statistical methodologies in a general way, allowing flexibility in the student's choice of physics process.

Both analysis notebooks end with suggestions for further exploration. For example, the supersymmetry analysis points students toward extending the interpretation of the results through a statistical analysis, which is covered in the statistics notebook. Together, these notebooks provide a complete foundation for performing a beyond the Standard Model search, from loading data to the statistical interpretation of results.

The semi-dark Higgs analysis shares substantial overlap with the supersymmetry analysis in terms of workflow, analysis framework, event selection strategy, and implementation. For this reason, a dedicated educational notebook for the semi-dark Higgs analysis was not produced. The supersymmetry analysis notebook provides the core workflow and analysis structure, and only minimal modifications would be required to adapt it for the semi-dark Higgs analysis.

6.4 Reproducibility

At the time of writing, the most recent publicly available education release corresponds to the 13 TeV 2025 Data-Beta release of ATLAS Open Data for education [74, 76]. This release serves as the primary openly available dataset used in the development of the analysis workflow and accompanying educational notebooks.

However, several components used in the complete analysis presented in this thesis are not included in this public release, notably the 2lep skim and the MC signal samples used in the supersymmetry analysis. At the time of writing, the full supersymmetry analysis cannot be reproduced using only publicly available resources, although the core workflow remains reproducible using the available data and background MC samples.

To support transparency and future reproducibility, multiple versions of the supersymmetry analysis notebook are provided in the accompanying GitHub repository [87]. When the signal samples become publicly available with the next ATLAS Open Data for education release (expected spring 2026), the notebooks can be updated accordingly to enable a fully reproducible workflow mirroring the analysis presented in this thesis.

[2] A cutflow table is used in high-energy particle physics analyses to visualise event yields and the impact of applying selection requirements.

The statistics notebook is unaffected and remains fully reproducible, as it operates on the produced analysis outputs rather than requiring direct access to the datasets.

With the methodologies established for both the physics analysis and the development of the educational notebooks, the following chapters present the corresponding results and discussion.

Part IV

Results and Discussion

Chapter 7

Analysis Results

This chapter presents the results of the analysis, beginning with a validation of the Standard Model predictions against observed data for key kinematic distributions, followed by the discussion of event yields and distributions in the signal model regions, concluding with a statistical interpretation of the results. Section 7.1 establishes the agreement between observed data and Standard Model background predictions. The results for the chargino-neutralino production with W and Z bosons and the semi-dark Higgs decay processes are presented in Section 7.2 and Section 7.3, respectively, highlighting the effects of the selection requirements to enhance sensitivity to the signal models.

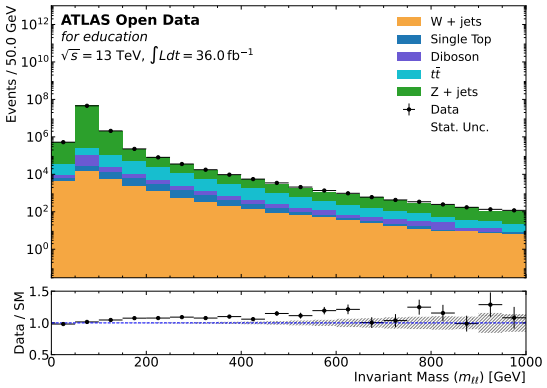
7.1 Standard Model Background Validation

This section validates the analysis by comparing key kinematic distributions between observed data and Standard Model background predictions. A basic event selection is applied as defined in Section 5.4.1, requiring events with exactly two leptons, and suppressing low mass contributions by imposing a dilepton invariant mass greater than 40 GeV. No additional signal-specific requirements as defined in Section 5.4.2, or Section 5.4.3 are applied at this stage.

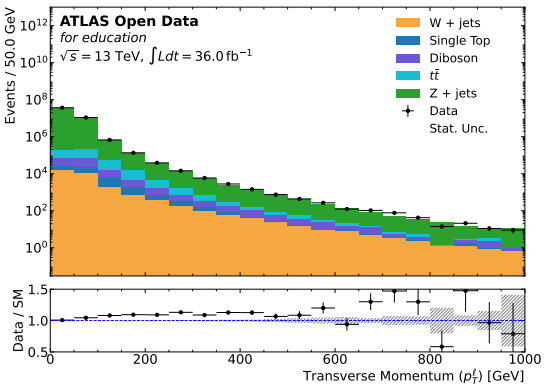
Overall, sufficient agreement is observed between the observed data and Standard Model predictions across the variables as shown in Figure 7.1. The ratio of data to prediction remains close to 1.0, particularly in regions with high statistics. Due to the reduced number of events at higher energies, larger fluctuations are observed, although they generally remain in agreement when considering greater statistical uncertainties. This agreement provides confidence in the modelling of the Standard Model background processes, and forms the basis for the subsequent application of progressively tighter selection requirements.

The invariant mass (Figure 7.1a) distribution shows a clear peak around 91 GeV, consistent with Z boson decays to dilepton pairs, such as Z+jets and diboson production. Low mass events are suppressed, as required by the $m_{\ell\ell} > 40$ GeV selection criteria. The transverse momentum distribution of the leading lepton (Figure 7.1b) shows an expected falling spectrum, where most events are concentrated at low transverse momentum.

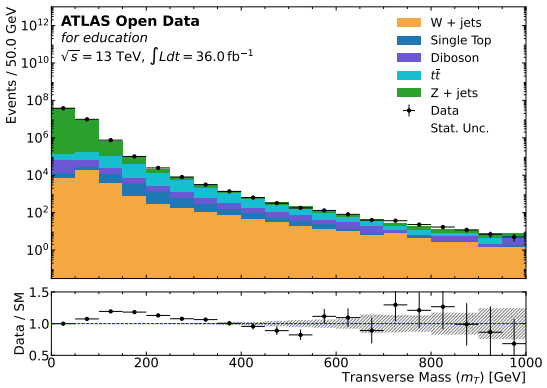
The transverse mass (Figure 7.1c) and the dilepton transverse mass (Figure 7.1d) distributions represent the kinematics of systems involving visible objects and missing transverse momentum. The transverse mass is constructed under the assumption of a single massless visible particle recoiling against an invisible particle. The resulting distribution exhibits the expected falling behaviour, representing the kinematics of the



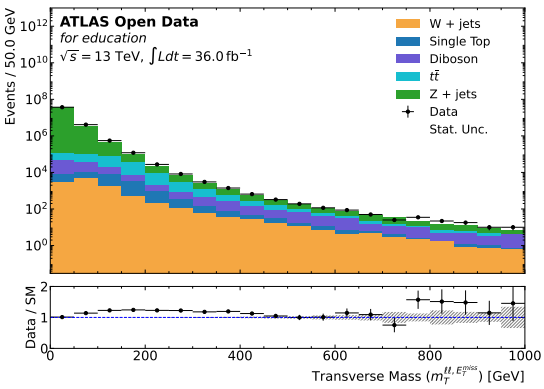
(a) Invariant Mass



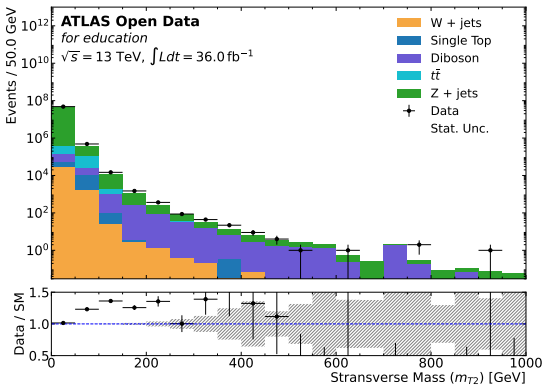
(b) Transverse Momentum Leading Lepton



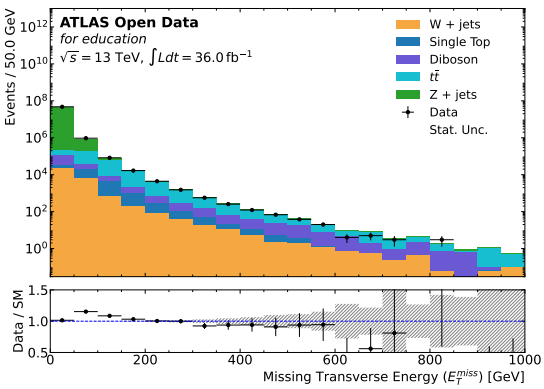
(c) Transverse Mass



(d) Dilepton Transverse Mass



(e) Stransverse Mass



(f) Missing Transverse Energy

Figure 7.1: Distributions of the (a) invariant mass, (b) transverse momentum of the leading lepton, (c) transverse mass, (d) transverse mass of the dilepton system, (e) stransverse mass, and (f) missing transverse energy with basic event selection criteria applied.

Standard Model backgrounds, such as in processes involving neutrinos. In contrast, the dilepton transverse mass extends this definition by treating the visible system as a single object with non-zero mass, corresponding to, in this scenario, a dilepton system consistent with a Z boson. This makes the variable better suited to processes where the visible final state is not massless, and therefore sensitive to processes involving massive visible objects in addition to the missing transverse momentum. This is particularly relevant for scenarios such as the semi-dark Higgs decay considered in this analysis,

where a dark photon recoils against a Z boson. The dilepton transverse mass distribution exhibits a similar falling behaviour as seen in the transverse mass distribution.

The stransverse mass (Figure 7.1e) distribution further extends this sensitivity to final states with two invisible particles, as expected in supersymmetric scenarios involving pair-produced particles. This makes it particularly useful for searches with multiple sources of missing transverse momentum. The distribution also shows a steeply falling behaviour, with a more rapid decrease compared to the transverse mass variables, as expected due to the assumption of an additional invisible particle.

The missing transverse energy (Figure 7.1f) distribution shows the expected falling behaviour where missing energy arises primarily from neutrinos in Standard Model processes. As the signal models predict additional sources of high missing transverse energy from invisible particles, this is an especially important variable for the analysis.

A discrepancy between data and background predictions is observed in several distributions. These differences are most notably seen in the low missing transverse energy, transverse mass, and stransverse mass regions. These differences may be related to limitations in the MC modelling of Z+jets and diboson processes, particularly related to higher-order corrections. In typical analyses, such effects are constrained using dedicated control regions, which allow data-driven corrections to be derived for the backgrounds [30, 88]. However, no control regions are used in this analysis, and this limitation should be taken into account when interpreting the level of agreement.

Further studies of the data and background agreement are presented in Appendix B, where an additional validation region is introduced. The selection requirements in this additional validation region are tighter than basic validation, while still remaining looser than those of the signal regions requirements targeting chargino-neutralino production with W and Z bosons. This region further confirms the behaviour of the modelling in a region closer to the signal selections. The modelling of Standard Model processes is therefore considered adequate for the purpose of this analysis, and we proceed to the definition and interpretation of the signal regions.

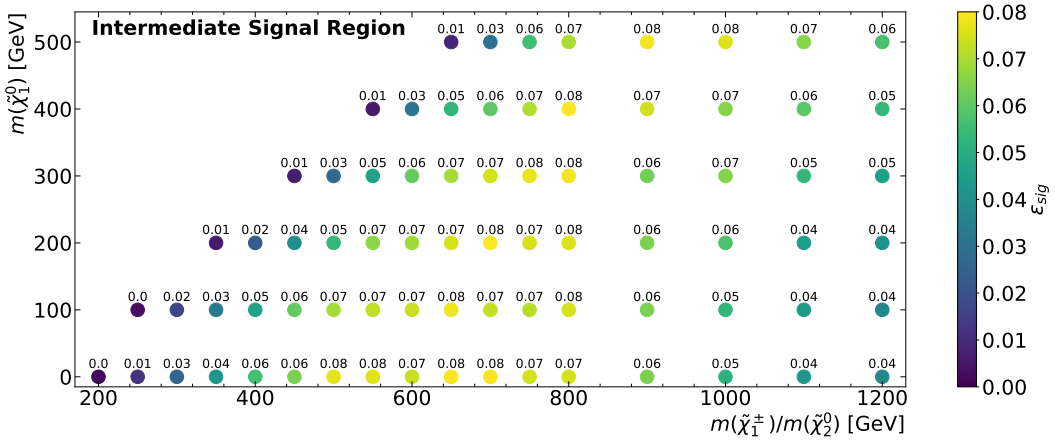
7.2 Chargino-Neutralino Production with W and Z Bosons

This section presents the results of the search for chargino-neutralino production with W and Z bosons in a two-lepton final state using the selection criteria described in Section 5.4.2. This analysis strategy uses large missing transverse energy as the primary variable to differentiate signal from Standard Model background processes. Signal events are expected to contain two leptons from a Z boson decay, jets from the W boson decay, and large missing transverse momentum from the lightest neutralinos. Contributions from Standard Model backgrounds, such as Z+jets, typically exhibit lower missing transverse energy, as seen in Figure 7.3a, motivating the progressively tighter requirements on missing transverse energy to enhance signal sensitivity.

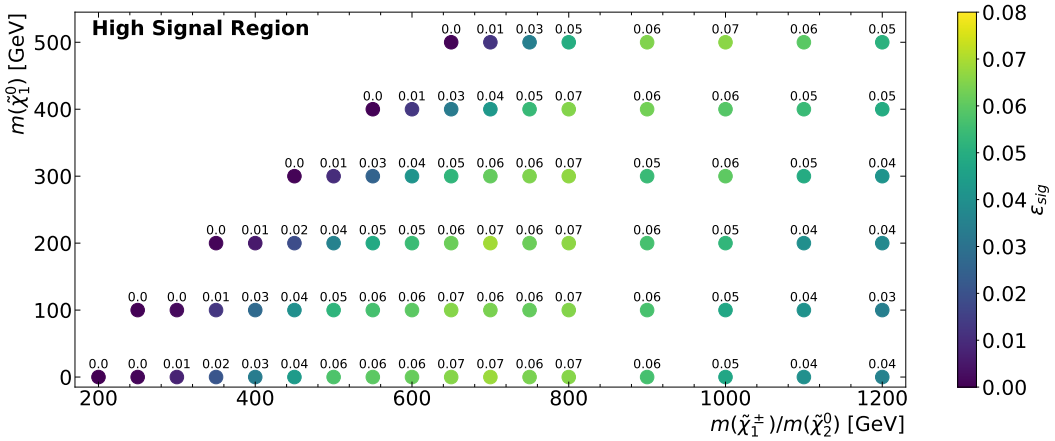
A baseline region is defined by applying all selection criteria except the missing transverse energy requirement. Two signal regions are then constructed: an intermediate signal region with missing transverse energy greater than 150 GeV, and a high signal region with missing transverse energy greater than 250 GeV. These regions are cumulative and are designed to retain signal efficiency while suppressing Standard Model background contributions.

The performance of the selection is quantified by the signal efficiency, as defined by equation (5.23) in Section 5.5.2. Since the kinematics of the process depend on the masses of the chargino/neutralino and lightest neutralino, the signal efficiency

varies across the two-dimensional mass grid. Figure 7.2 shows the signal efficiencies in both the intermediate and high signal regions. In both regions, signal models with chargino/neutralino masses approximately in the range 400 - 800 GeV tend to exhibit the highest efficiencies across the grid.



(a) Intermediate Signal Region ($E_T^{\text{miss}} > 150$ GeV)



(b) High Signal Region ($E_T^{\text{miss}} > 250$ GeV)

Figure 7.2: Comparison of signal efficiencies for the signal models corresponding to chargino-neutralino production with W and Z bosons in the (a) intermediate and (b) high signal regions. Each point on the grid corresponds to a specific combination of particle masses. The heat map displays the signal efficiency (ϵ_{sig}), where lighter (yellow) colours indicate higher efficiencies and darker (purple) colours indicate lower efficiencies. The numerical value of ϵ_{sig} is shown above each point.

Table 7.1 summarises the impact of the selection requirements on both data and background processes in the different regions. The uncertainties listed in Table 7.1 include statistical uncertainties together with a 30% flat systematic uncertainty applied to the background. At the pre-selection level, event yields correspond to the number of events generated after the initial analysis skim and the application of minimal selection criteria designed to prioritise events containing two leptons in the final state. At this stage, the total event sample is dominated by Z+jets production, as expected.

Applying the baseline selection significantly reduces the total event yield, while maintaining Z+jets as the dominant background, followed by diboson processes. Contributions from $t\bar{t}$, W+jets, and single top production become negligible. The

	Pre-selection($\times 10^6$)	Baseline	Intermediate	High
Data	83.9	82	1	0
Total SM	61.6 ± 18.5	51.5 ± 15.8	2.05 ± 0.95	0.65 ± 0.59
Z+jets	59.2 ± 17.8	43.4 ± 13.3	0.15 ± 0.07	0.02 ± 0.02
Diboson	0.23 ± 0.07	8.04 ± 3.01	1.90 ± 0.92	0.63 ± 0.58
$t\bar{t}$	1.27 ± 0.38	-	-	-
W+jets	0.99 ± 0.30	-	-	-
Single Top	0.12 ± 0.04	-	-	-

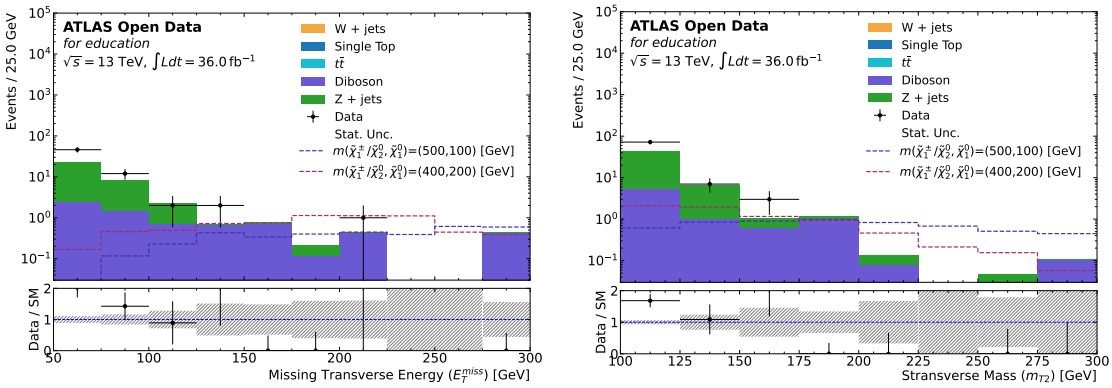
Table 7.1: Summary of event yields in pre-selection (with basic lepton selection requirements applied), baseline region (before the E_T^{miss} requirement is applied), intermediate signal region ($E_T^{\text{miss}} > 150$ GeV) , and high signal region ($E_T^{\text{miss}} > 250$ GeV) for chargino-neutralino production with W and Z bosons. Each of these regions corresponds to progressively tighter requirements on missing transverse energy and represents cumulative yields. The data events, total Standard Model (SM) expectation, and the expectations for the individual Standard Model processes are shown. The statistical and systematic uncertainties are reported for each Standard Model process. Processes with negligible contributions are denoted with a "-" symbol.

background rejection rate is defined as $(1 - \frac{N_{\text{f}}}{N_{\text{i}}})$, where N_{f} is the remaining event yield after the selection criteria are applied, and N_{i} is the initial event yield. The baseline selection reduces the total Standard Model background by 99% relative to pre-selection yields.

In the baseline region, 82 events are observed compared to the total predicted Standard Model yield of 51.5 ± 15.8. While the observed number of events exceeds the predicted background, this region is not used for statistical interpretation of the signal models as it is not optimised for signal sensitivity. Rather, this region serves as a reference before the missing transverse energy requirements are applied.

In the intermediate signal region, the background is strongly suppressed, with one event observed compared to an expected background of 2.05 ± 0.95 events. This corresponds to a background rejection of 99% relative to the pre-selection yields, and an additional 96% reduction relative to the baseline region. The dominant contribution shifts to the diboson processes in this region. The high signal region further reduces the background contributions to 0.65 ± 0.59 events, with zero data events observed. This corresponds to the same overall rejection relative to pre-selection, and a further 68% reduction relative to the intermediate signal region. Both intermediate and high regions show good agreement within uncertainties between the Standard Model expectations and observed data, and no significant excess is observed in either region.

The kinematic distributions further illustrate the effectiveness of the selection strategy. Distributions of the missing transverse energy and stransverse are mass shown in Figure 7.3 for the baseline region, where Z+jets provide the largest contribution in this region, followed by diboson processes. The observed data generally follow the predicted background distributions, with some discrepancies in the lowest bins, consistent with the behaviour observed in the Standard Model background validation (Section 7.1). The missing transverse energy distribution shown in Figure 7.3a indicates that the signal models maintain a significant contribution across the distribution, with relatively larger yields in the high missing transverse energy tail. In contrast, the observed data and predicted background yields decrease towards larger missing transverse energy values. A similar trend is observed in the stransverse mass distribution (Figure 7.3b), where the



(a) Missing Transverse Energy

(b) Stransverse Mass

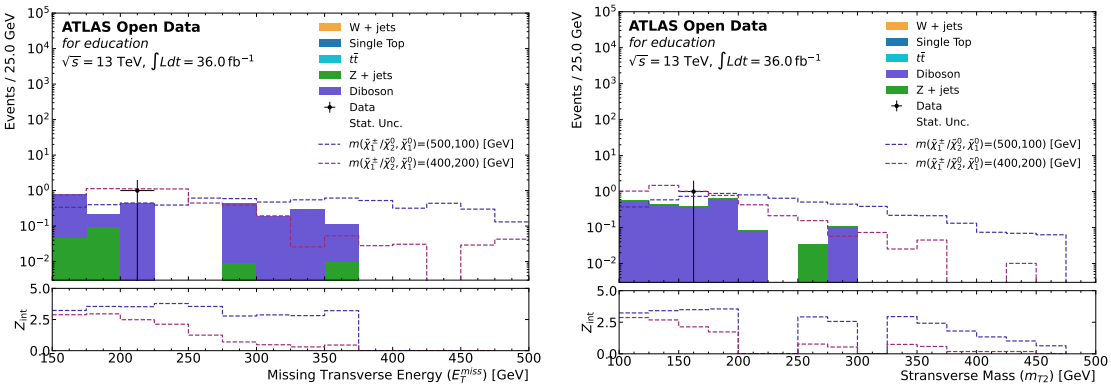
Figure 7.3: Distributions of the (a) missing transverse energy and (b) stransverse mass relating to chargino-neutralino production with W and Z bosons. These distributions represent the baseline region where missing transverse energy selection requirements are not applied. The uncertainty bands include only the statistical uncertainty (Stat. Unc.) on the background processes. Two representative signal models for chargino-neutralino production are shown. The ratio of observed data to Standard Model background is shown in the subplot with corresponding statistical uncertainties.

signal contribution remains visible while the Standard Model background decreases at higher mass values.

As the missing transverse energy requirement is applied, the background contributions decrease, while the signal distributions remain significant, particularly in the high missing transverse energy and high stransverse mass regions. This behaviour reflects the presence of the invisible particles in the signal models. The intermediate signal region and corresponding distributions in the high signal region are shown in Figure 7.4 and Figure 7.5, respectively. In the high signal region, the background contribution is almost entirely suppressed, leaving only a small number of events dominated by diboson processes. Although the stricter selection requirements slightly reduce the signal efficiency, as seen in Figure 7.2, the remaining signal event contributions continue to remain visible in the distribution, demonstrating that the selection requirements retain sensitivity to the signal models.

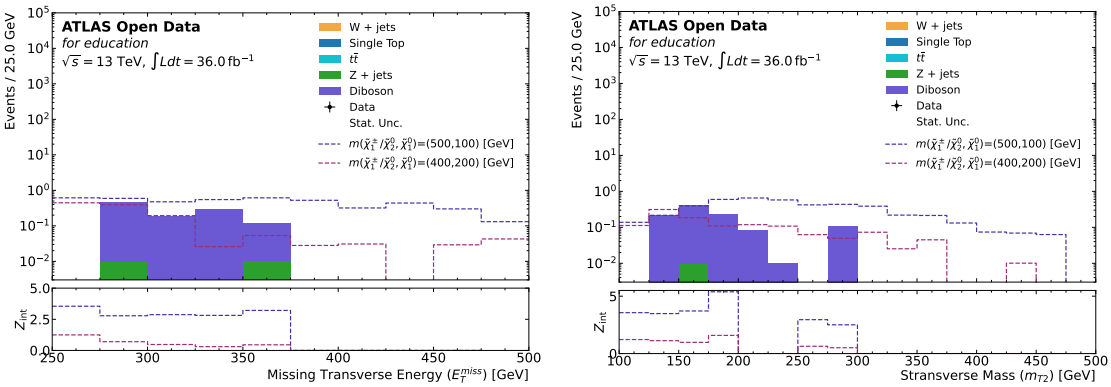
Two representative signal models are shown in the distributions and examined to illustrate the behaviour of the signal for different mass splittings between the chargino/neutralino and the lightest neutralino. The first of these chosen models has a chargino/neutralino mass of 500 GeV and a lightest neutralino mass of 100 GeV, corresponding to a 400 GeV mass splitting, denoted as $\mathrm{m}(\tilde{\chi}_1^\pm/\tilde{\chi}_2^0, \tilde{\chi}_1^0) = (500, 100)$ GeV. The second chosen model has a chargino/neutralino mass of 400 GeV and a lightest neutralino mass of 200 GeV, corresponding to a 200 GeV mass splitting.

In the pre-selection and baseline regions, the (400, 200) GeV model yields a larger number of signal events than the (500, 100) GeV model, as shown in Table 7.2. This behaviour is driven by the larger production cross section of the lower mass model. When requirements on missing transverse energy are applied in the intermediate and high regions, the (500, 100) GeV model begins to exceed the (400, 200) GeV model in event yields. This becomes particularly apparent in the high missing transverse energy tail of the distribution, as seen in Figure 7.4a, where the (500, 100) GeV model surpasses the (400, 200) GeV model at approximately 250 GeV. The larger mass splitting in the (500, 100) GeV model results in greater available energy in the decay, therefore producing



(a) Missing Transverse Energy (b) Stransverse Mass

Figure 7.4: Distributions of (a) missing transverse energy and (b) stransverse mass corresponding to chargino-neutralino production with W and Z bosons. A selection criterion in the intermediate signal region with missing transverse energy greater than 150 GeV is applied. The uncertainty bands include the statistical uncertainty (Stat. Unc.) on the background processes. Two representative signal models for chargino-neutralino production are shown. The integrated significance (Z_{int}) is shown in the subplot for the representative signal models.



(a) Missing Transverse Energy (b) Stransverse Mass

Figure 7.5: Distributions of (a) missing transverse energy and (b) stransverse mass corresponding to the chargino-neutralino production with W and Z bosons. A selection criterion in the high signal region with missing transverse energy greater than 250 GeV is applied. The uncertainty bands include the statistical uncertainty (Stat. Unc.) on the background processes. Two representative signal models for chargino-neutralino production are shown. The integrated significance (Z_{int}) is shown in the subplot for the representative signal models.

larger missing transverse momentum.

This behaviour is further illustrated by the integrated significance, calculated using the low-statistics approximation similar to equation (4.9). The integrated significance is obtained by cumulatively summing the signal and background yields from each bin to the upper end of the distribution, so that each bin corresponds to all events above that value. Overall, the (500, 100) GeV model exhibits a greater significance. In Figure 7.5b, one bin exceeds an integrated significance of $Z_{\text{int}} = 5$. However, this corresponds to an expected significance obtained from a simplified estimate that does not account for systematic uncertainties or correlations.

These results demonstrate how the missing transverse energy requirements enhance the sensitivity to different regions of the supersymmetric parameter space. No clear

m($\tilde{\chi}_1^\pm/\tilde{\chi}_2^0, \tilde{\chi}_1^0$)	Pre-selection	Baseline	Intermediate	High
(500, 100)	91.0 ± 27.3	7.10 ± 2.14	6.30 ± 1.91	4.73 ± 1.44
(400, 200)	233 ± 69.9	7.20 ± 2.20	5.26 ± 1.65	1.24 ± 0.41

Table 7.2: Signal event yields for the representative signal models are shown for the baseline region (before the E_T^{miss} requirement is applied), intermediate signal region ($E_T^{\text{miss}} > 150$ GeV), and high signal region ($E_T^{\text{miss}} > 250$ GeV) for chargino-neutralino production with W and Z bosons and compared to the pre-selection event yields. The signal models are defined by the mass (m) in units of GeV of the chargino/neutralino ($\tilde{\chi}_1^\pm/\tilde{\chi}_2^0$), and the corresponding mass of the lightest neutralino ($\tilde{\chi}_1^0$). Each of these regions corresponds to progressively tighter requirements on missing transverse energy. The statistical and systematic uncertainties are reported for each signal model.

excess over the Standard Model expectation is observed, and a statistical interpretation is therefore performed in the next section to quantify the level of agreement and set constraints on the signal models.

7.2.1 Statistical Interpretation

The following statistical analysis is performed to interpret the results obtained in the signal regions, and exclusion limits are derived to place upper bounds on the considered signal models. Two complementary approaches are evaluated, frequentist and Bayesian. In the frequentist framework, the CL_s, p-value, and significance are computed using an asymptotic approximation, while the Bayesian method is used to determine the upper limits on the event yields and visible cross section. A detailed comparison of these two approaches is provided in Appendix A.

Since the previously defined signal regions are not mutually exclusive, the statistical interpretation is performed using orthogonal regions to avoid double-counting of events. The intermediate signal region is redefined as $150 < E_T^{\text{miss}} \leq 250$, while the high signal region remains $E_T^{\text{miss}} > 250$. These regions are evaluated individually and combined to form a single intermediate plus high region, labelled as "combined." The combined region takes the sum of the yields from both regions to obtain the observed and expected yield values as presented in Table 7.3. Both statistical and 30% flat systematic uncertainties are incorporated as nuisance parameters in the model.

The observed event yields are consistent with the Standard Model expectations within uncertainties. This is supported by the p-value, $p_0 = 0.5$ and corresponding significance, $Z = 0$, indicating no statistically significant deviation from the background-only hypothesis. Within the frequentist framework, exclusion limits are then derived using the CL_s method.

Confidence levels for each signal mass hypothesis are determined using the CL_s method introduced in Sections 4.2.4 and 5.5.3. Figure 7.6 shows the resulting grid of mass hypotheses, where points with $CL_s^{\text{obs}} < 0.05$ are excluded at 95% confidence level. This method highlights the compatibility of each signal hypothesis with the observed data. These CL_s results provide the frequentist exclusion limits for the signal hypotheses.

For the final interpretation, the Bayesian limits are used to quantify the exclusion in terms of the signal strength parameter μ, defined as the ratio of the fitted signal yield to the predicted signal yield. A signal hypothesis is excluded at 95% confidence level if the signal strength parameter $\mu_{95} < 1.0$. The median expected limits are obtained from pseudo-experiments, with $\pm 1\sigma$ uncertainty bands. The Bayesian approach is chosen as the primary result, as it does not rely on asymptotic approximations and is better suited

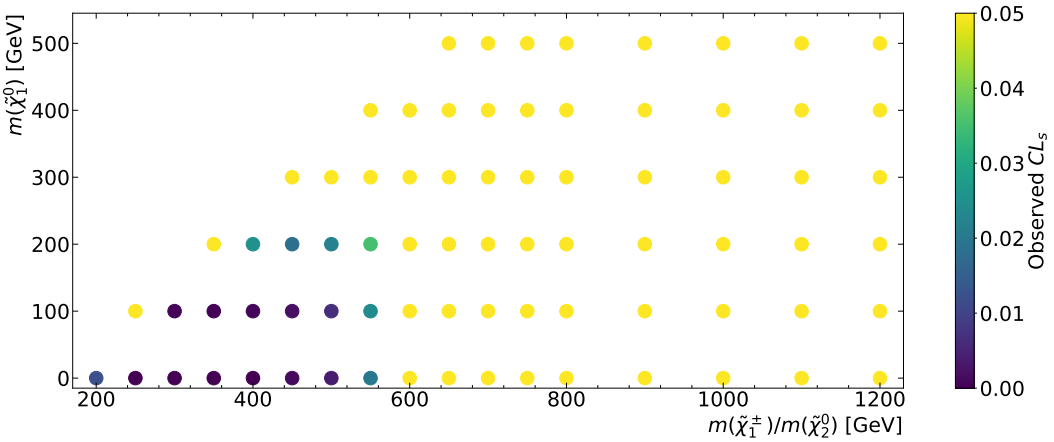


Figure 7.6: Grid of the observed CL_s value for each signal mass hypothesis in the combined region. Points with $CL_s < 0.05$ (dark green-purple) are excluded at 95% confidence level, while points with $CL_s \geq 0.05$ (light yellow) are not considered excluded. The colour scale ranges from 0.00 to 0.05.

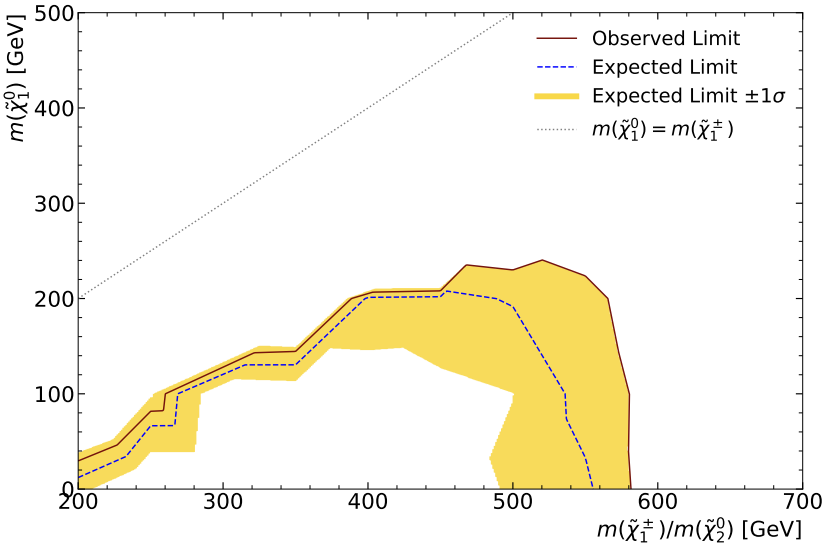


Figure 7.7: Observed (solid red) and expected (dashed blue) exclusion limits on chargino-neutralino production with W and Z bosons. The shaded (yellow) band represents $\pm 1\sigma$ variation on the expected limit. The grey dotted line indicates the region where the chargino mass equals the neutralino mass.

to the low-statistics regime in this analysis. A comparison between the frequentist and the Bayesian results is provided in Appendix A.

Figure 7.7 shows the resulting exclusion contour. The observed limit closely follows that of the expected $+1\sigma$ band, especially in intermediate to high mass regions. This behaviour is consistent with statistical fluctuations in the low-statistics regime, where small variations in the observed event yields with relatively large uncertainties can lead to shifts in the exclusion sensitivity. Differences between the signal regions are accounted for in the combined likelihood fit, which determines the final exclusion contour. When individually considered, the intermediate signal region excludes chargino/neutralino masses up to approximately 420 GeV, while the high signal region excludes chargino/neutralino masses between 290 - 570 GeV, for a massless neutralino

at a 95% confidence level. These individual contributions reflect the combined region, where the observed limits exclude chargino/neutralino masses up to approximately 580 GeV for a massless neutralino at a 95% confidence level, in agreement with previous searches [30]. A comparison of the results of this analysis and those of a published ATLAS analysis is presented in Appendix C.

Model-independent limits are derived from the model-dependent results for the intermediate, high, and combined regions. These correspond to the maximum value obtained across all model-dependent limits. Table 7.3 summarises the observed and expected 95% upper limits on the signal yield using the Bayesian method, which is well-suited for the low-statistics regime. The corresponding limits on the visible cross section are given by $\sigma^{95}_{\text{obs}} = S^{95}_{\text{obs}}/\mathcal{L}$, yielding 0.14 fb and 0.13 fb for the intermediate region and high region, respectively.

Signal Region	N_{obs}	N_{exp}	σ^{95}_{obs} [fb]	S^{95}_{obs}	S^{95}_{exp}	p_0	Z
Intermediate	1	1.4 ± 0.6	0.14	4.9	$4.9^{+3.2}_{-1.3}$	0.5	0
High	0	0.6 ± 0.6	0.13	4.6	$6.5^{+2.3}_{-2.1}$	0.5	0
Combined	1	2.0 ± 1.6	0.12	4.1	$5.3^{+2.2}_{-1.7}$	0.5	0

Table 7.3: Summary of observed and expected event yields and model-independent limits. N_{exp} corresponds to the Standard Model expectations, while N_{obs} shows the observed data events. 95% upper limits on the signal yield (S^{95}_{obs} and S^{95}_{exp}) and the visible cross section (σ^{95}_{obs}) are shown. The combined region corresponds to the intermediate plus high regions. The p-value (p_0) and significance (Z) assume the background only-hypothesis.

Although this analysis targets supersymmetric chargino-neutralino production, the final state with two leptons and missing transverse energy is not exclusive to this process. Similar final states may arise in other beyond the Standard Model scenarios, such as a semi-dark Higgs decay, where missing transverse energy may originate from a dark photon.

Motivated by this, a complementary analysis targets a semi-dark Higgs decay with a dark photon. Although the final state is generally consistent, the kinematic properties differ due to the decay of the Higgs boson. The following section presents the results of this search, highlighting how selection requirements are adapted to enhance sensitivity to the alternative signal model, providing a complementary perspective on dark matter searches in two-lepton final states.

7.3 Semi-Dark Higgs Decay

This section presents the results of a search for semi-dark Higgs decays in a similar two-lepton and missing transverse energy state considered in the previous analysis. The signal model assumes the production of a Standard Model-like Higgs boson with a mass of 125 GeV, which decays via a Z boson and a dark photon. The two leptons originate from the Z boson decay, while in this scenario, the missing transverse energy arises from a dark photon, which is assumed to be massless in this model. Although the experimental signature of the final state is similar to the chargino-neutralino production model, the underlying kinematics differ. Therefore, this analysis follows similar event selection criteria, with modifications to enhance sensitivity to the semi-dark Higgs decay signal, as described in Section 5.4.3.

A baseline region is defined by applying all selection criteria defined in Section 5.4.3 except the requirements on missing transverse energy and dilepton transverse mass. These requirements are then introduced in the signal region where missing transverse energy is greater than 100 GeV, and the dilepton transverse mass is within 91-200 GeV.

	Pre-selection($\times 10^6$)	Baseline($\times 10^3$)	Signal Region
Data	83.9	619	193
Total SM	61.6 ± 18.5	528 ± 158	142 ± 43.3
Z+jets	59.2 ± 17.8	523 ± 157	67.2 ± 20.5
Diboson	0.23 ± 0.07	4.90 ± 1.47	46.3 ± 15.2
$t\bar{t}$	1.27 ± 0.38	0.69 ± 0.02	23.2 ± 7.74
W+jets	0.99 ± 0.30	0.07 ± 0.01	2.02 ± 0.91
Single Top	0.12 ± 0.04	0.17 ± 0.01	3.07 ± 1.34
Signal: $H \to Z\gamma_D$	0.38 ± 0.12	14.1 ± 4.24	459 ± 138

Table 7.4: Summary of event yields in baseline region and signal region for the semi-dark Higgs decay compared to the pre-selection event yields. Each of these regions corresponds to progressively tighter requirements and represents cumulative yields. The data events, total Standard Model (SM) expectation, and the expectations for the individual Standard Model processes are shown. The generated expected signal yield for the process, $H \to Z\gamma_D$, where γ_D is a dark photon, is also shown for each region. Statistical and systematic uncertainties are included for each process.

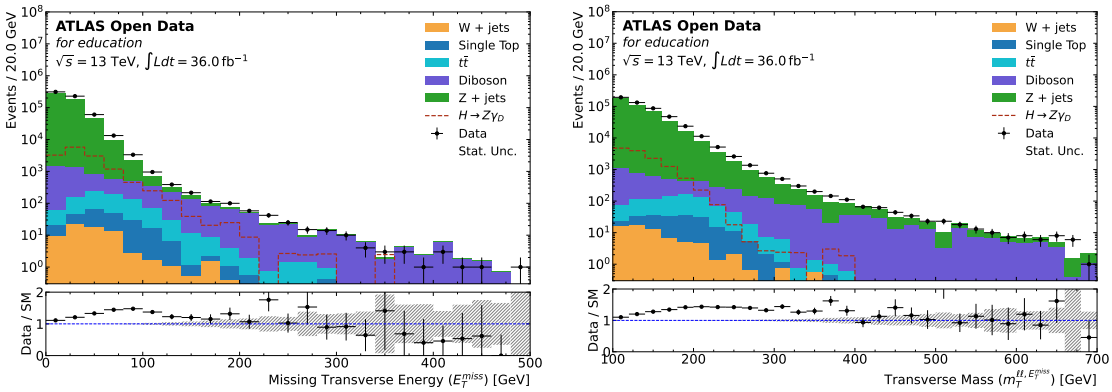
The event yields for the data and Standard Model backgrounds in each region are summarised in Table 7.4 including both statistical and systematic uncertainties. The selection criteria applied in the baseline region and signal region significantly reduce the Standard Model background contributions, with a corresponding decrease in the observed data. The baseline selection reduces the total Standard Model background by 99% relative to pre-selection yields. In both regions, the dominant background arises from Z+jets, followed by diboson and $t\bar{t}$ processes. The W+jets and single top contributions are suppressed once the missing transverse energy requirement is applied.

In the signal region, a total of 193 data events are observed compared to the total predicted Standard Model yield of 142 ± 43.3. This results in a background rejection of 99% relative to pre-selection, and a further 99% reduction relative to the baseline region. These results are consistent within uncertainties, and no significant deviations are observed. The expected yields for the considered signal model are also shown in Table 7.4, where the signal model has a signal efficiency of $\epsilon_{sig} = 0.037$ in the baseline region, and $\epsilon_{sig} = 0.001$ in the signal region.

The kinematic distributions provide further insight into the behaviour of the signal and background processes. Figure 7.8 shows the missing transverse energy and dilepton transverse mass in the baseline region. In this region, Z+jets dominate, and the data are generally consistent with the predicted background.

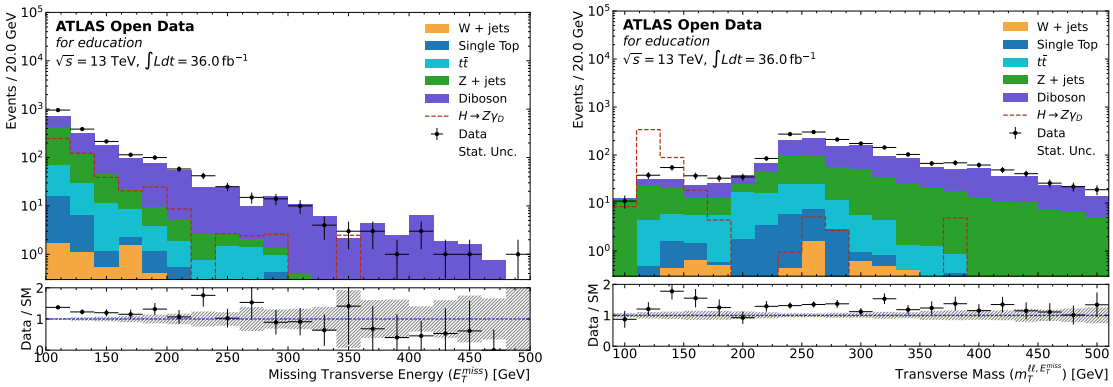
After applying the missing transverse energy requirement as shown in Figure 7.9, the behaviour of the background changes significantly. The Z+jets is suppressed, and diboson processes become the dominant background. This behaviour is consistent with expectations as Z+jets processes typically contribute in the low missing transverse energy regions.

The missing transverse energy distribution in Figure 7.9a indicates that the signal model contributes primarily at lower values and decreases more rapidly than the



(a) Missing Transverse Energy (b) Dilepton Transverse Mass

Figure 7.8: Distributions of (a) missing transverse energy and (b) dilepton transverse mass corresponding to the semi-dark Higgs decay process. A selection criterion in the baseline region without the missing transverse energy or dilepton transverse mass is applied. The uncertainty bands include the statistical uncertainty (Stat. Unc.) on the background processes. The signal model is shown as a red dashed line. The ratio of observed data to Standard Model background is shown in the subplot with corresponding statistical uncertainties.



(a) Missing Transverse Energy (b) Dilepton Transverse Mass

Figure 7.9: Distributions of (a) missing transverse energy and (b) dilepton transverse mass corresponding to the semi-dark Higgs decay process. A selection criterion in the signal region with missing transverse energy greater than 100 GeV, excluding the dilepton transverse mass requirement, is applied. The uncertainty bands include the statistical uncertainty (Stat. Unc.) on the background processes. The signal model is shown as a red dashed line. The ratio of observed data to Standard Model background is shown in the subplot with corresponding statistical uncertainties.

dominant Standard Model background. Some differences between data and Standard Model background are observed, which are consistent with what was observed in the Standard Model background validation discussed in Section 7.1, where the use of data-driven corrections may improve data to Standard Model background agreement.

In contrast, the dilepton transverse mass distribution in Figure 7.9b exhibits a pronounced peak for the signal. The signal peaks near 125 GeV, consistent with the Higgs boson mass. This behaviour reflects the kinematics of the signal decay, where the Z boson is reconstructed from the leptons and the dark photon contributes to the missing transverse energy. The distribution of the background processes for the dilepton transverse mass also changes significantly after applying the missing transverse energy

requirement. While the baseline region shows the expected falling distribution, this region produces a flatter distribution at low transverse mass, followed by a rise and peak, before falling again. This behaviour is driven by the kinematics between the reconstructed Z boson and the recoiling system contributing to missing transverse energy. Since the missing transverse energy is required to be greater than 100 GeV, paired with the Z boson mass of approximately 91 GeV, this corresponds well to the background behaviour in the low mass region with a flat distribution and rise at approximately 200 GeV. Applying a tighter or looser requirement on the missing transverse energy shifts the behaviour of this distribution, respectively.

Although this signal model produces a distinct peak in the dilepton transverse mass region around the Higgs boson mass, the observed data remain consistent with the Standard Model background expectation. A statistical analysis is performed in the subsequent section to quantify any potential deviation.

7.3.1 Statistical Interpretation

Using the previously defined signal region, a statistical analysis is performed. The p-value and significance under the background-only hypothesis are computed using the frequentist asymptotic approximation of the likelihood test statistic. A value of $p_0 = 0.15$ and $Z = 1.05$ are obtained. These values differ slightly from those obtained in the supersymmetric chargino-neutralino analysis (Section 7.2.1), reflecting the slight excess of observed events relative to the background prediction. A total of 193 events are observed, compared to the expected yield of 141 ± 43.3, placing the observation just above the 1σ uncertainty. This is reflected in the significance being slightly larger than 1, and the corresponding lower p-value. These values indicate that the observed data are consistent with the Standard Model expectations, and no statistically significant deviation is observed.

As no statistically significant deviation is observed, exclusion limits are set on the model as summarised in Table 7.5. The observed and expected 95% confidence upper limits on the signal yield are derived using the Bayesian methods and are found to be 117 and $70.3^{+47.2}_{-29.1}$, respectively. The upper limit on the visible cross section σ^{95}_{obs} is 3.26 fb.

Signal Process	N_{obs}	N_{exp}	σ^{95}_{obs} [fb]	S^{95}_{obs}	S^{95}_{exp}	p_0	Z
$H \to Z\gamma_D$	193	141 ± 43.3	3.26	117	$70.3^{+47.2}_{-29.1}$	0.15	1.05

Table 7.5: Summary of observed and expected event yield limits. N_{exp} corresponds to the Standard Model expectations, while N_{obs} shows the observed data events. 95% upper limits on the signal yield (S^{95}_{obs} and S^{95}_{exp}) and the visible cross section (σ^{95}_{obs}) are shown. The combined region corresponds to the intermediate plus high regions. The p-value (p_0) and significance (Z) assume the background only-hypothesis.

Both these analyses show agreement between the data and Standard Model expectations. With no statistically significant excess observed, exclusion limits are determined for each of the models, with resulting 95% upper confidence limits on the signal yields and visible cross sections. The methodologies used to arrive at the results provide the necessary elements motivating the educational materials described in Chapter 8.

Chapter 8

Educational Notebooks

This chapter presents the Jupyter notebooks developed for outreach and education. Each notebook follows the methodologies employed during the physics analysis sections, providing a framework that mimics how experimental particle physics analyses are conducted in practice.

Two complementary notebooks have been developed. The first presents a dark matter motivated supersymmetry analysis, while the second provides a statistical interpretation of the results. The statistics notebook builds directly on the outputs of the supersymmetry analysis but is also structured to introduce statistical techniques in a general context, making it applicable to other analyses.

The notebooks are available in a GitHub repository [87]. They can be downloaded and executed locally, or accessed via web-based platforms such as Binder and Google Colab. Instructions and links to the notebooks through these platforms are included in the repository's `README.md`[1] file.

8.1 Supersymmetry Analysis

The supersymmetry analysis notebook guides students through a complete workflow, following the methodologies outlined earlier in Section 5.4.2, and reproducing select results presented in Section 7.2.

The notebook combines conceptual explanations with executable code. This structure allows students to follow the analysis workflow step by step. The notebook begins by loading the ATLAS Open Data for education and defining detector data, and MC simulated samples for both background and signal processes. Derived variables that are used in the analysis, such as the invariant mass and stransverse mass, among others, are computed within the notebook. Event selection criteria are implemented through selection cuts on the variables. The main analysis loop processes all samples using these selections, enabling the full workflow to be executed. Analysis outputs are produced, such as cutflow tables and histograms. Following the analysis outputs, a simple statistical interpretation of the results is provided where the observed and expected significance are calculated.

Executing the notebook without modifications reproduces key results from the physics analysis. In particular, the cutflow table corresponding to the baseline selection column in Table 7.2, and the missing transverse energy distribution matching the one

[1] https://github.com/emilypuchala/MasterThesis/blob/main/README.md

shown in Figure 7.3. This demonstrates that the notebook provides a reproducible implementation of the analysis.

Interactive elements are integrated throughout the notebook to encourage active learning and independent exploration of the analysis workflow. "Try it yourself" sections prompt students to modify the code and explore the impact on the results, while "What do you think?" sections encourage interpretation of the outputs. Programming tips are included to highlight relevant programming practices, supporting both physics understanding and computational skills. Upon completion of the analysis, students are presented with a "What's next?" section, which suggests further extensions as well as directs students to continue the analysis interpretation with the statistical analysis notebook.

8.2 Statistical Analysis

The statistical analysis notebook follows a similar structure, focusing on the interpretation of analysis results through statistical methods. It introduces the statistics framework alongside executable code.

The notebook presents a simplified frequentist-based likelihood approach using a modified CL_s method with asymptotic approximations. It includes implementations of the Poisson distribution, likelihood function construction with uncertainties and nuisance parameters, test statistics, p-values, significance calculations, and exclusion limits. As in the supersymmetry analysis, interactive elements are included to promote active learning. The notebook concludes with suggestions for further exploration and extensions.

Running the notebook without modifications produces similar results to those shown in the physics analysis chapter. Specifically, Figure 7.6 is recreated, as well as an approximated version of Figure 7.7 is produced. These outputs demonstrate that the statistical implementations presented in the notebook are directly applicable to the results of the analysis.

The use of asymptotic approximations prioritises computational efficiency and conceptual clarity over precision. While more computationally intensive approaches, such as toy MC methods, could be implemented, they are intentionally omitted to maintain accessibility within an educational context, resulting in approximated results. The use of these methods and how they are validated against the Bayesian methods primarily used in the physics analysis section is discussed in Appendix A.

8.3 Applications

Although developed for a supersymmetry analysis, the notebooks are adaptable to other analyses, such as the semi-dark Higgs decay discussed in Section 5.4.3 and Section 7.3. Due to the similar analysis structure, minimal modifications are required to apply the workflow to other physics models with two-lepton final states.

The notebooks demonstrate how analysis code can be translated into an educational format, thereby bridging the gap between research and education. As the implementation is directly adapted from the code used to produce the physics results in Chapter 7 and those results are consistent with established studies, it highlights the value of ATLAS Open Data as a resource for teaching experimental particle physics through realistic analysis workflows.

Together, the physics analysis results and the development of educational materials illustrate the study of dark matter searches in two-lepton final states with missing transverse energy. A summary of the main conclusions, along with reflections on the limitations, improvements, and directions for future work, is presented in the following chapter.

Part V

Conclusion

Chapter 9

Conclusion

Searches for physics beyond the Standard Model remain central in particle physics, especially in efforts to understand the nature of dark matter. This thesis presents analyses of supersymmetric chargino-neutralino production and semi-dark Higgs boson decays in two-lepton final states with missing transverse energy using ATLAS Open Data for education. In parallel, it develops complementary educational materials to make such analyses more accessible.

The analyses use the 13 TeV 2025 Data-Beta release of ATLAS Open Data for education, corresponding to an integrated luminosity of 36 fb^{-1}. In the supersymmetric search, no statistically significant excess above the expected background is observed, consistent with Standard Model expectations. A 95% confidence level upper limit of 0.12 fb is set on the visible cross section, excluding chargino/neutralino masses up to approximately 580 GeV for a massless lightest neutralino. Similarly, the semi-dark Higgs search shows no significant deviation from the background-only hypothesis. A 95% confidence level upper limit of 3.26 fb is set on the visible cross section. Although these results show no evidence for new physics, they contribute to efforts to constrain theoretical models and refine the search for physics beyond the Standard Model.

A comparison with a published ATLAS analysis [30], presented in Appendix C, shows that the overall distribution shapes and exclusion trends for the supersymmetric scenario are broadly consistent. However, differences are observed in the event yields and in the limits obtained in this analysis compared to those reported in the publication. These deviations are likely influenced by the simplifications inherent in this analysis.

Several limitations contribute to these differences. In particular, the treatment of systematic uncertainties is significantly simplified, where a flat 30% systematic uncertainty is applied to the Standard Model backgrounds, rather than accounting for individual sources of uncertainty, which vary between processes. In addition, the absence of control regions and data-driven background estimates may affect the level of agreement between data and Standard Model background prediction, leading to discrepancies in the distributions and event yields. Furthermore, the use of a cut-and-count strategy, as opposed to more advanced analysis techniques, contributes to a reduction in the ability to distinguish signal from background.

Together, these limitations affect the normalisation of the distributions, particularly in the missing transverse energy and stransverse mass distributions. They may also explain the differences between the event yields and limits obtained in this analysis and those reported in the ATLAS publication in Reference [30]. Despite these limitations, this analysis demonstrates the feasibility of performing particle physics searches using

open data. It highlights the strengths and constraints of such datasets, especially in educational contexts where simplicity is often prioritised.

To bridge research and education, two Jupyter notebooks were developed to reproduce the supersymmetric analysis and its statistical interpretation. These notebooks provide an accessible and reproducible framework for engaging with real experimental data and can be extended to other beyond the Standard Model analyses, including the semi-dark Higgs decay, with minimal modifications.

Future improvements could include not using open data by incorporating the larger datasets, such as the full Run 2 or Run 3 data, to increase sensitivity. Additionally, implementing machine learning techniques such as neural networks or boosted decision trees to enhance signal selection could improve signal discrimination. These methods would improve analysis performance while reflecting modern research techniques in experimental particle physics.

The exotic Higgs sector also presents multiple avenues for exploring new physics, where semi-dark Higgs decay models provide alternatives to fully invisible Higgs decay scenarios. These approaches can target not only dark photon models, as discussed previously, but also other mediators between the Standard Model particles and a dark matter candidate.

In summary, through the combination of physics research and the development of educational materials, this thesis contributes both to the search for new physics and the advancement of interactive, hands-on learning in particle physics. Together, these elements highlight the broader impact of modern physics research and physics outreach.

Appendices

Appendix A

Frequentist and Bayesian Statistics

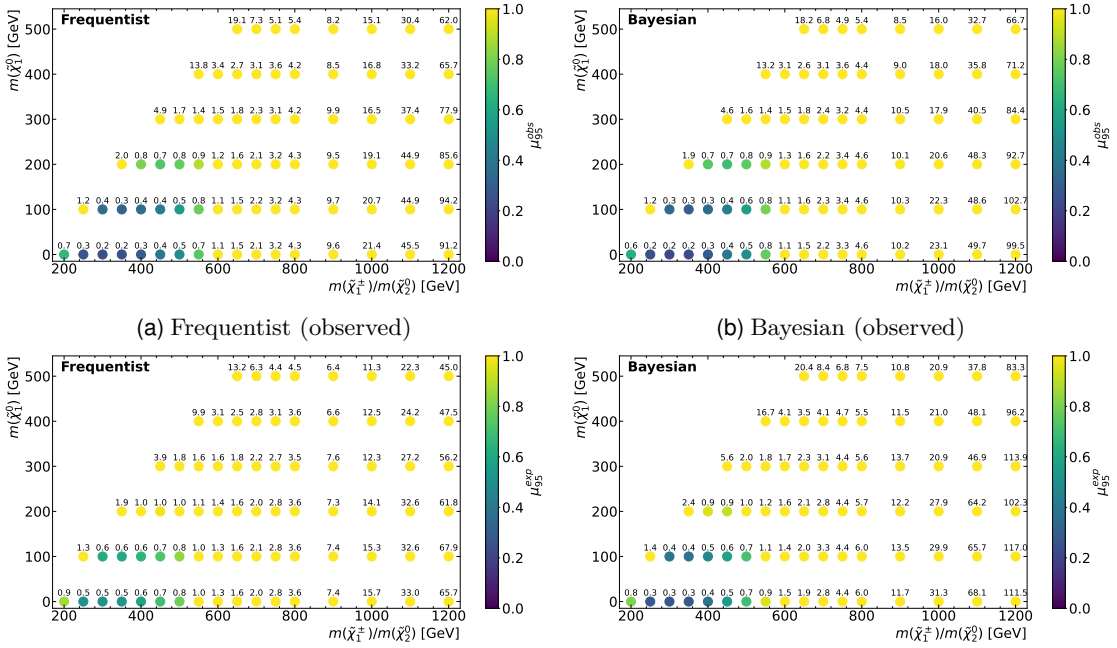
In this work, two statistical methods are used to derive exclusion limits on the models. The methods are compared here, and the differences between the implementations used in this analysis are discussed. The frequentist method employs an asymptotic approximation of the CL_s method, while the Bayesian method constructs a posterior probability. Both methods evaluate the limits at a 95% confidence level. In the following, these will be referred to as frequentist or Bayesian approaches, noting that these labels correspond to the specific implementations used in this analysis, and may differ in detail across studies.

Both approaches are based on a Poisson likelihood distribution, but differ in their statistical interpretation and treatment of nuisance parameters. In the frequentist framework, limits are obtained from a profile likelihood ratio test statistic, where nuisance parameters are fitted to maximise the likelihood for each assumed signal strength parameter. The Bayesian approach constructs a posterior probability by integrating over nuisance parameters and averaging over their possible value.

Although both approaches constrain the signal strength parameter μ, the computational methods used to arrive at this parameter differ. The frequentist approach directly determines an upper limit on μ, while the Bayesian approach first derives limits on the visible cross section using the signal efficiency, luminosity, and expected event yields, which are then converted into constraints on μ. The signal strength parameter is the key quantity used to derive both model-dependent and model-independent limits for the cross section and signal event yields in both methods. These values are compared in Table A.1, with the observed and expected values of μ_{95} shown in Figure A.1.

			Frequentist			Bayesian		
Signal Region	N_{obs}	N_{exp}	σ^{95}_{obs}[fb]	S^{95}_{obs}	S^{95}_{exp}	σ^{95}_{obs}[fb]	S^{95}_{obs}	S^{95}_{exp}
Intermediate	1	1.4 ± 0.6	0.11	4.1	$8.0^{+3.1}_{-2.2}$	0.14	4.9	$4.9^{+3.2}_{-1.3}$
High	0	0.6 ± 0.6	0.08	2.7	$5.3^{+2.4}_{-1.5}$	0.13	4.6	$6.5^{+2.3}_{-2.1}$
Combined	1	2.0 ± 1.6	0.12	4.3	$9.4^{+3.6}_{-2.6}$	0.12	4.1	$5.3^{+2.2}_{-1.7}$

Table A.1: Comparison of model-independent event yield results at a 95% upper confidence level using frequentist and Bayesian methods for the chargino-neutralino production process. The number of expected events (N_{exp}) is the Standard Model prediction, and N_{obs} are observed data. Limits on observed and expected signal yields ($S^{95}_{\text{obs}}, S^{95}_{\text{exp}}$) and visible cross section (σ^{95}_{obs}) are shown for both statistical approaches.



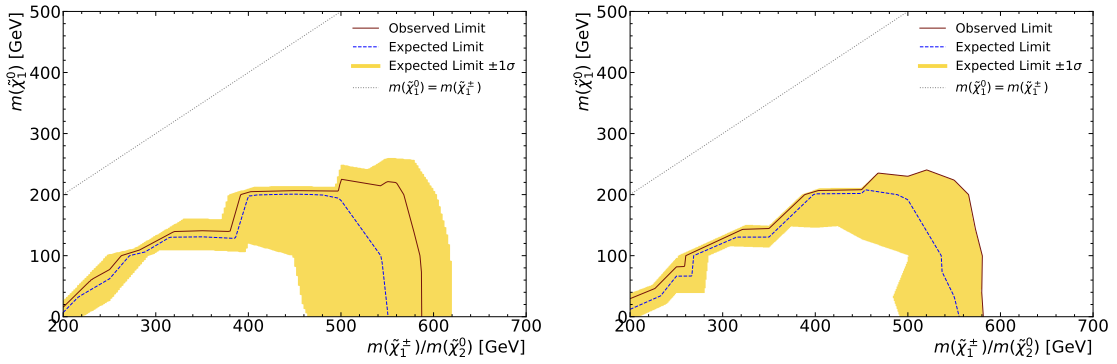
(a) Frequentist (observed)

(b) Bayesian (observed)

(c) Frequentist (expected)

(d) Bayesian (expected)

Figure A.1: Observed (μ_{95}^{obs}) and expected (μ_{95}^{exp}) 95% confidence limits on the signal strength parameter for chargino-neutralino production with W and Z bosons. A grid of mass points for the chargino/neutralino ($\tilde{\chi}_1^\pm/\tilde{\chi}_2^0$) and corresponding lightest neutralino ($\tilde{\chi}_1^0$) where the limits for the (a, c) frequentist and (b, d) Bayesian statistical methods are compared. The values of μ_{95} are listed for each point. The colour-bar maps if the mass point is excluded, where lighter (yellow) points are not excluded, and darker (green-purple) points are excluded.



(a) Frequentist

(b) Bayesian

Figure A.2: Observed and expected exclusion limits on chargino-neutralino production with W and Z bosons. The observed limit (red solid line) and expected limit (blue dashed line) are indicated, and the shaded (yellow) band corresponds to the $\pm 1\sigma$ variation in the expected limit, including both statistical and systematic uncertainties. A grey dotted line shows the bounds of the data where the mass of the chargino ($\tilde{\chi}_1^\pm$) equals the mass of the neutralino ($\tilde{\chi}_1^0$). The limits are computed at a 95% confidence level, where the exclusion limits for the (a) frequentist and the (b) Bayesian statistical methods are compared.

In the frequentist approach, the expected limits are derived using an Asimov dataset, representing the expected data under the background-only model, where the observed data events are set equal to the expected background events ($N_{\text{obs}} = N_{\text{exp}}$). The

exclusion contours for both methods are shown in Figure A.2. The observed limits allow for a slightly more direct comparison between the two methods as they are derived from the actual observed event counts rather than the Asimov dataset inputs. Table A.1 summarises the model-independent limits for the intermediate, high, and combined regions.

As seen in Figure A.1, the model-dependent values of μ_{95} for the observed limits have similar agreement across the parameter space, with small deviations. In contrast, the expected limits in regions with relatively higher statistics corresponding to low to mid mass points show a similar exclusion limit for both methods. In regions with relatively lower statistics corresponding to higher mass points, where the signal efficiency is decreased, larger deviations occur. This highlights how asymptotic approximations may be less reliable in low-statistics regimes. These differences arise from a combination of the treatment of nuisance parameters and the statistical regime of the analysis. The impact of lower statistics in the frequentist method is especially apparent in Table A.1, where the value of S^{95}_{obs} in the high signal region deviates from the values obtained in the intermediate and combined signal regions.

To assess the impact of uncertainties, the analysis was repeated with statistical and systematic uncertainties set to zero for the background yields and signal efficiencies. While these uncertainties are modelled as nuisance parameters, statistical fluctuations still arise from the inherent nature of the Poisson distribution in the construction of the likelihood. The resulting limits are shown in Table A.2. In both scenarios, removing these uncertainties results in stronger (lower) limits, while the trend between statistical methods remains similar. However, the degree of agreement between the Bayesian and frequentist results does not fully converge when the uncertainties are removed. This indicates that nuisance parameters contribute to the differences between the methods, but the statistical fluctuations and event counts may also play a role.

	Frequentist			Bayesian		
Uncertainty Treatment	σ^{95}_{obs} [fb]	S^{95}_{obs}	S^{95}_{exp}	σ^{95}_{obs} [fb]	S^{95}_{obs}	S^{95}_{exp}
With uncertainties	0.12	4.3	$9.4^{+3.2}_{-6.4}$	0.12	4.1	$5.3^{+2.2}_{-1.7}$
Without uncertainties	0.08	3.0	$6.9^{+2.7}_{-1.9}$	0.11	3.9	$5.0^{+2.1}_{-1.6}$

Table A.2: Comparison of model-independent 95% upper limits for the combined region for the chargino-neutralino production process, with and without statistical and systematic uncertainties, using frequentist and Bayesian methods.

This effect is more apparent in the frequentist results, while the Bayesian result shows a smaller relative change between the inclusion or exclusion of uncertainties. This may be a result of the different treatment of nuisance parameters and how the Bayesian marginalises over uncertainty rather than profiling them as in the frequentist approach, which reduces the overall effect on the model-independent limits themselves.

The treatment of uncertainties as nuisance parameters, and whether they are treated as correlated or uncorrelated, can introduce an additional difference between the methods. In the frequentist implementation, the statistical uncertainties are treated as uncorrelated, while systematic uncertainties are correlated across signal regions. In the Bayesian implementation, both systematic and statistical uncertainties are treated as uncorrelated nuisance parameters, reflecting a simplified modelling choice, rather than a property of the Bayesian approach.

A similar comparison can be made in a higher-statistics regime using the semi-dark Higgs model, where the role of nuisance parameters becomes more apparent compared to statistical fluctuations. This model is evaluated using both methods, and a summary of event yield upper limits is shown in Table A.3. In contrast to the chargino-neutralino analysis, this model yields a larger number of expected and observed events.

Method	N_{obs}	N_{exp}	σ^{95}_{obs} [fb]	S^{95}_{obs}	S^{95}_{exp}
Frequentist	193	141 ± 43.3	5.22	188	$194^{+76.0}_{-54.3}$
Bayesian	193	141 ± 43.3	3.26	117	$70.3^{+47.2}_{-29.1}$

Table A.3: Comparison of 95% upper limits for the semi-dark Higgs model using Bayesian and frequentist methods.

In this higher-statistics regime, the frequentist approach yields more conservative upper limits on both visible cross section and signal event yields compared to the Bayesian approach. To investigate the role of uncertainties, the comparison was repeated with uncertainties removed, as shown in Table A.4. In this regime, removing uncertainties leads to a closer agreement between the two methods compared to the supersymmetric scenario. This difference suggests that in higher-statistics scenarios, the handling of nuisance parameters and statistical treatment becomes a key factor between the two methods, rather than statistical fluctuations.

Both the supersymmetric and semi-dark Higgs scenarios confirm that the inclusion of uncertainties leads to more conservative (weaker, larger) upper limits. In addition, it shows that the level of agreement between the Bayesian and frequentist methods depends on the statistical regime of the analysis.

	Frequentist			Bayesian		
Uncertainty Treatment	σ^{95}_{obs} [fb]	S^{95}_{obs}	S^{95}_{exp}	σ^{95}_{obs} [fb]	S^{95}_{obs}	S^{95}_{exp}
With uncertainties	5.22	188	$194^{+76.0}_{-54.3}$	3.26	117	$70.3^{+47.2}_{-29.1}$
Without uncertainties	2.08	75.0	$39.9^{+15.6}_{-11.2}$	2.23	80.5	$26.4^{+10.5}_{-7.4}$

Table A.4: Comparison of model-independent 95% upper limits for the semi-dark Higgs signal region, with and without statistical and systematic uncertainties, using frequentist and Bayesian methods.

Overall, the comparison between these two methods provides a useful validation of the statistical interpretation. The Bayesian approach implemented here offers flexibility through marginalisation over nuisance parameters, but is computationally more expensive. The frequentist method offers a faster alternative, though its behaviour depends on the statistical regime of the analysis. This highlights that the differences between Bayesian and frequentist methods depend not only on the characteristics of the data but also on the statistical regime and implementation details of the analysis.

Appendix B

Additional Validation Region

In addition to validating the distributions of various kinematic variables using basic lepton selection criteria, a dedicated validation region targeting chargino-neutralino production with W and Z bosons is defined in this section.

The basic lepton selection requirements remain unchanged from those defined in Section 5.4.1. In addition, the invariant mass of the dilepton system is required to be within $81 - 101$ GeV, while the invariant mass of the dijet system, m_{jj} is not within $60 - 100$ GeV. The transverse momentum of the reconstructed bosons is constrained where $p_T^Z > 80$ GeV and $p_T^W > 100$ GeV. Events are required to have at least two non-b-tagged jets, where the two leading jets require $p_T^{\text{jet}} > 30$ GeV, non-leading jets satisfy $p_T^{\text{jet}} > 20$ GeV, and all jets are constrained to $|\eta| < 2.5$. Events containing any b-tagged jets with $p_T^{\text{jet}} > 20$ GeV are rejected. Jets are identified as b-tagged jets using the DL1dv01 algorithm at a working point with 77% b-tagging efficiency [81, 82].

In contrast to the signal regions, no explicit requirements are imposed on the stransverse mass, or angular variables such as ΔR and $\Delta\phi$. The key selection criteria defining this validation region are summarised in Table B.1.

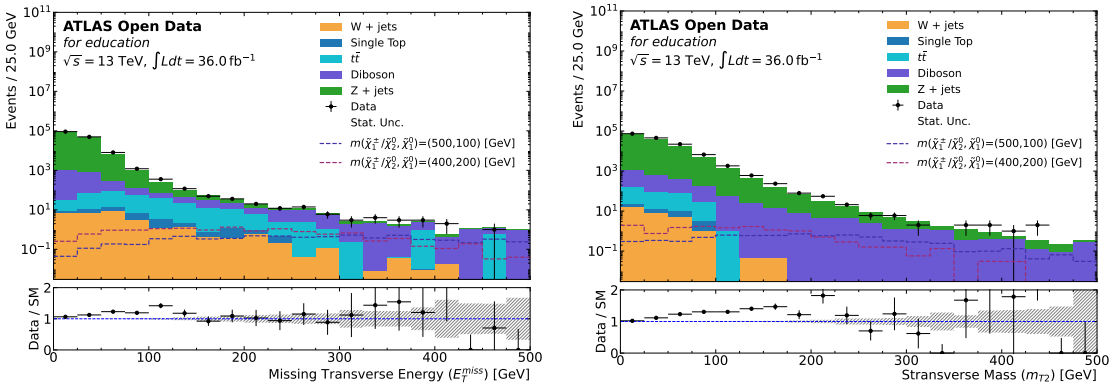
	Validation Region
$m_{\ell\ell}$ [GeV]	$[81, 101]$
m_{jj} [GeV]	$\notin [60, 100]$
p_T^Z [GeV]	> 80
p_T^W [GeV]	> 100
n_{jets}	≥ 2

Table B.1: Summary of event selection criteria for the validation region targeting chargino-neutralino production with W and Z bosons.

Table B.2 summarises the event yields in the validation region, compared to the pre-selection yields. Figure B.1 shows the distributions of missing transverse energy and stransverse mass. In contrast to the initial validation discussed in Section 7.1, this validation region targets a region similar to the signal regions, while maintaining sufficient background event yields for comparison with data. The Z+jets process is the dominant background in this region, followed by diboson and $t\bar{t}$ contributions. A discrepancy between Standard Model background and data is observed in the stransverse mass distribution; however, similar differences were also noted in Section 7.1. As

	Pre-selection($\times 10^6$)	Validation($\times 10^3$)
Data	83.9	151
Total SM	61.6 ± 18.5	138 ± 41.5
Z+jets	59.2 ± 17.8	136 ± 40.7
Diboson	0.23 ± 0.07	2.22 ± 0.67
$t\bar{t}$	1.27 ± 0.38	0.31 ± 0.09
W+jets	0.99 ± 0.30	0.03 ± 0.01
Single Top	0.12 ± 0.04	0.02 ± 0.01

Table B.2: Summary of event yields in the validation region targeting chargino-neutralino production with W and Z bosons are compared to the pre-selection event yields. The data events, total Standard Model (SM) expectation, and the expectations for the individual Standard Model processes are shown. The statistical and systematic uncertainties are reported for each Standard Model process.



(a) Missing Transverse Energy

(b) Stransverse Mass

Figure B.1: Distributions of (a) missing transverse energy and (b) stransverse mass for the validation region targeting chargino-neutralino production with W and Z bosons.

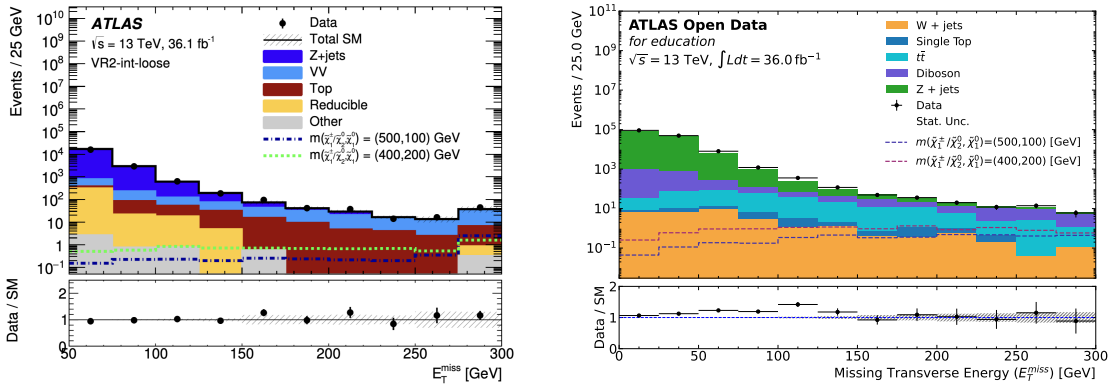
discussed there, the usage of dedicated control regions and data-driven estimates may reduce these discrepancies in the ratio between observed data and Standard Model background. Overall, sufficient agreement is observed within uncertainties, supporting the application of tighter requirements in the signal regions.

Appendix C

Comparison to a Published Study

The results of this analysis are compared with those reported in the published study "Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13$ TeV with the ATLAS detector" [30]. The methodology used for the chargino-neutralino production process follows the approach described in the "2ℓ+jets channel" in that publication. Any references to "the publication" in this appendix refer to the methods, definitions, and results presented in that study. Differences between the two analyses are examined, including both methodological variations and deviations in the results.

Both analyses use datasets corresponding to an integrated luminosity of 36 fb^{-1}. The same event selection criteria were applied in order to reproduce the "VR2-int-loose" validation region. As shown in Figure C.1, a larger number of events is observed for both data and Standard Model background processes in this analysis compared to the publication. The final bin of the distribution is not compared directly, as the publication includes an overflow bin, whereas this analysis does not. However, the overall shape of the distribution appears to be consistent, and the dominant contributions from Z+jets and diboson processes are reproduced.



(a) Publication Figure 2c [30]

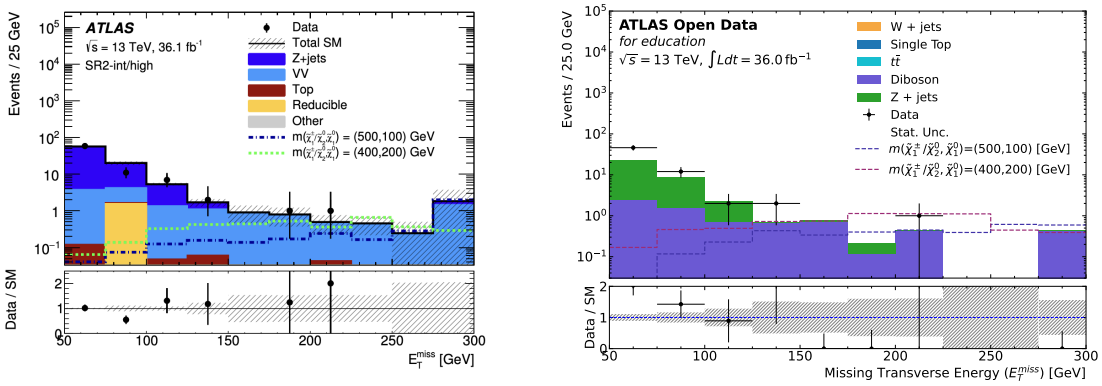
(b) Figure B.1a

Figure C.1: Missing transverse energy distribution for the validation region in (a) the publication compared to (b) the results of this analysis.

For the signal regions, the "SR2 int/high" regions from the publication are compared to the baseline region, intermediate signal region, and high signal region defined in this analysis. The missing transverse energy distributions in Figure C.2, exhibit similar overall behaviour in both analyses. Differences are observed in the event yields per

bin, as well as in the level of agreement between observed data and Standard Model background predictions. The final bin of the distribution is not compared directly, as the publication includes an overflow bin, whereas this analysis does not.

Quantitatively, differences between event yields are evident. In the intermediate region, this analysis observes one event, while the publication reports two events. The total Standard Model background yield is 2.05 ± 0.95 events in this analysis, compared to $4.1^{+2.6}_{-1.8}$ in the publication. In the high region, both analyses observe zero events. However, the predicted total Standard Model background differs: 0.65 ± 0.59 events in this analysis compared to $1.6^{+1.6}_{-1.1}$ events in the publication.



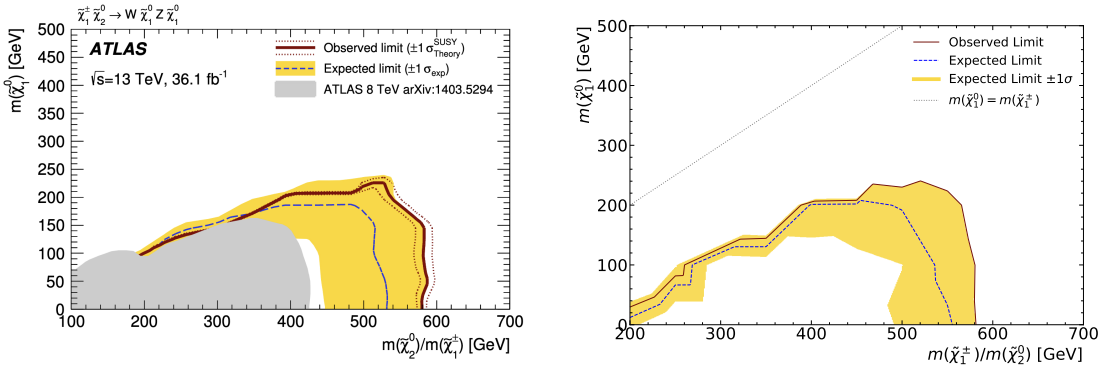
(a) Publication Figure 5a [30]

(b) Figure 7.3a

Figure C.2: Missing transverse energy distribution of the signal region in (a) the publication compared to (b) the results of this analysis.

A significant methodological difference arises in the statistical interpretation. The publication derives exclusion limits by combining both the three-lepton and two-lepton-plus-jets channels. These signal regions are not mutually exclusive, and the best CL_s value is selected for each mass point. In contrast, this analysis defines orthogonal signal regions and considers only the two-lepton channel.

Despite these differences, both analyses excluded chargino/neutralino masses up to approximately 580 GeV for a massless lightest neutralino. However, discrepancies remain in the upper limits on the visible cross sections, as well as in the observed and expected signal yields.



(a) Publication Figure 8d [30]

(b) Figure 7.7

Figure C.3: 95% confidence level exclusion contour result for chargino-neutralino production with W and Z bosons in (a) the publication compared to (b) the results of this analysis.

The observed differences are likely influenced by the simplified treatment of systematic uncertainties in this analysis, as well as the absence of control regions and data-driven background estimates, discussed in previous chapters. These limitations are expected to affect the modelling of dominant backgrounds, particularly the Z+jets and diboson processes. Additionally, the use of open data intended for educational purposes, where simplicity is prioritised over precision, may further contribute to observed deviations.

While exact agreement is not expected, the general consistency in distribution shapes and exclusion reach suggests that the overall behaviour of this analysis is in line with that of the publication, within the limitations described above.

Appendix D

MC Simulated Event Samples

The simulated Standard Model background and signal samples [74] used in this analysis are grouped by physics process, where the background samples are listed in Appendix D.1, while the signal samples are listed in Appendix D.2. Each sample includes its dataset identification (DSID), which is a unique number to identify the dataset. The physics short name contains information for the simulated process, including event generator information. The cross section, filter efficiency, and k-factor for each sample are shown, and are used in the calculation of the event weight as described in Section 5.2.2.

D.1 Background Samples

The background samples are grouped by physics process and listed in Table D.1, Table D.2, Table D.3, Table D.4, and Table D.5, for Z+jets, W+jets, diboson, $t\bar{t}$, and single top quark processes, respectively.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
700320	Sh_2211_Zee_maxHTpTV2_BFilter	2221.20	0.02	1.00
700321	Sh_2211_Zee_maxHTpTV2_CFilterBVeto	2221.20	0.13	1.00
700322	Sh_2211_Zee_maxHTpTV2_CVetoBVeto	2221.20	0.85	1.00
700323	Sh_2211_Zmumu_maxHTpTV2_BFilter	2221.30	0.02	1.00
700324	Sh_2211_Zmumu_maxHTpTV2_CFilterBVeto	2221.30	0.13	1.00
700325	Sh_2211_Zmumu_maxHTpTV2_CVetoBVeto	2221.30	0.85	1.00

Table D.1: MC sample information for the Z+jets processes. DSID identifies the dataset, while the physics short name contains information for the process and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
700338	Sh_2211_Wenu_maxHTpTV2_BFilter	21742.00	0.01	1.00
700339	Sh_2211_Wenu_maxHTpTV2_CFilterBVeto	21742.00	0.15	1.00
700340	Sh_2211_Wenu_maxHTpTV2_CVetoBVeto	21742.00	0.84	1.00
700341	Sh_2211_Wmunu_maxHTpTV2_BFilter	21806.00	0.01	1.00
700342	Sh_2211_Wmunu_maxHTpTV2_CFilterBVeto	21806.00	0.15	1.00
700343	Sh_2211_Wmunu_maxHTpTV2_CVetoBVeto	21806.00	0.84	1.00

Table D.2: MC sample information for the W+jets processes. DSID identifies the dataset, while the physics short name contains information for the process and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
700492	Sh_2211_WqqZll	3.30	1.00	1.00
700493	Sh_2211_ZqqZll	6.30	0.26	1.00
700494	Sh_2211_ZbbZll	0.97	0.48	1.00
700600	Sh_2212_llll	1.30	1.00	1.00
700601	Sh_2212_lllv	4.66	1.00	1.00
700602	Sh_2212_llvv_os	12.08	1.00	1.00
700603	Sh_2212_llvv_ss	0.02	1.00	1.00
700604	Sh_2212_lvvv	3.11	1.00	1.00

Table D.3: MC sample information for the diboson processes. DSID identifies the dataset, while the physics short name contains information for the process and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
410470	PhPy8EG_A14_ttbar_hdamp258p75_nonallhad	729.77	0.54	1.14

Table D.4: MC sample information for the $t\bar{t}$ process. DSID identifies the dataset, while the physics short name contains information for the process and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
410644	PowhegPythia8EvtGen_A14_singletop_schan_lept_top	2.03	1.00	1.02
410645	PowhegPythia8EvtGen_A14_singletop_schan_lept_antitop	1.27	1.00	1.02
410658	PhPy8EG_A14_tchan_BW50_lept_top	37.00	1.00	1.19
410659	PhPy8EG_A14_tchan_BW50_lept_antitop	22.17	1.00	1.18
601624	PhPy8EG_tW_DS_dyn_dil_antitop	3.78	1.00	1.00
601628	PhPy8EG_tW_DS_dyn_dil_top	3.71	1.00	1.00

Table D.5: MC sample information for the single top quark processes. DSID identifies the dataset, while the physics short name contains information for the process and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

D.2 Signal Samples

The signal samples are grouped by physics process, where the semi-dark Higgs decay sample is listed in Table D.6. Table D.8 and Table D.8 include the samples for the chargino-neutralino production with W and Z bosons process.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
603891	PhPy8EG_ggH_HZyd_Zll	28.31	0.52	1.00

Table D.6: MC sample information for the semi-dark Higgs signal sample. DSID identifies the dataset, while the physics short name contains information for the process and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
563381	MGPy8EG_C1mN2_WZ_200_0_2L7_MadSpin	0.6371	0.09	1.00
563383	MGPy8EG_C1mN2_WZ_250_0_2L7_MadSpin	0.2653	0.10	1.00
563385	MGPy8EG_C1mN2_WZ_250_100_2L7_MadSpin	0.2653	0.09	1.00
563386	MGPy8EG_C1mN2_WZ_300_0_2L7_MadSpin	0.1266	0.10	1.00
563388	MGPy8EG_C1mN2_WZ_300_100_2L7_MadSpin	0.1266	0.10	1.00
563390	MGPy8EG_C1mN2_WZ_350_0_2L7_MadSpin	0.0663	0.11	1.00
563392	MGPy8EG_C1mN2_WZ_350_100_2L7_MadSpin	0.0663	0.11	1.00
563394	MGPy8EG_C1mN2_WZ_350_200_2L7_MadSpin	0.0663	0.10	1.00
563395	MGPy8EG_C1mN2_WZ_400_0_2L7_MadSpin	0.0371	0.11	1.00
563397	MGPy8EG_C1mN2_WZ_400_100_2L7_MadSpin	0.0371	0.11	1.00
563399	MGPy8EG_C1mN2_WZ_400_200_2L7_MadSpin	0.0371	0.10	1.00
563401	MGPy8EG_C1mN2_WZ_450_0_2L7_MadSpin	0.0219	0.12	1.00
563403	MGPy8EG_C1mN2_WZ_450_100_2L7_MadSpin	0.0219	0.11	1.00
563405	MGPy8EG_C1mN2_WZ_450_200_2L7_MadSpin	0.0219	0.11	1.00
563407	MGPy8EG_C1mN2_WZ_450_300_2L7_MadSpin	0.0219	0.10	1.00
563408	MGPy8EG_C1mN2_WZ_500_0_2L7_MadSpin	0.0134	0.12	1.00
563410	MGPy8EG_C1mN2_WZ_500_100_2L7_MadSpin	0.0134	0.12	1.00
563412	MGPy8EG_C1mN2_WZ_500_200_2L7_MadSpin	0.0134	0.11	1.00
563414	MGPy8EG_C1mN2_WZ_500_300_2L7_MadSpin	0.0134	0.11	1.00
563416	MGPy8EG_C1mN2_WZ_550_0_2L7_MadSpin	0.0085	0.12	1.00
563418	MGPy8EG_C1mN2_WZ_550_100_2L7_MadSpin	0.0085	0.12	1.00
563420	MGPy8EG_C1mN2_WZ_550_200_2L7_MadSpin	0.0085	0.12	1.00
563422	MGPy8EG_C1mN2_WZ_550_300_2L7_MadSpin	0.0085	0.11	1.00
563424	MGPy8EG_C1mN2_WZ_550_400_2L7_MadSpin	0.0085	0.10	1.00
563425	MGPy8EG_C1mN2_WZ_600_0_2L7_MadSpin	0.0056	0.13	1.00
563427	MGPy8EG_C1mN2_WZ_600_100_2L7_MadSpin	0.0056	0.13	1.00
563429	MGPy8EG_C1mN2_WZ_600_200_2L7_MadSpin	0.0056	0.12	1.00
563431	MGPy8EG_C1mN2_WZ_600_300_2L7_MadSpin	0.0056	0.12	1.00
563433	MGPy8EG_C1mN2_WZ_600_400_2L7_MadSpin	0.0056	0.11	1.00
563435	MGPy8EG_C1mN2_WZ_650_0_2L7_MadSpin	0.0037	0.13	1.00
563437	MGPy8EG_C1mN2_WZ_650_100_2L7_MadSpin	0.0037	0.13	1.00
563439	MGPy8EG_C1mN2_WZ_650_200_2L7_MadSpin	0.0037	0.13	1.00
563441	MGPy8EG_C1mN2_WZ_650_300_2L7_MadSpin	0.0037	0.12	1.00
563443	MGPy8EG_C1mN2_WZ_650_400_2L7_MadSpin	0.0037	0.11	1.00
563445	MGPy8EG_C1mN2_WZ_650_500_2L7_MadSpin	0.0037	0.10	1.00
563446	MGPy8EG_C1mN2_WZ_700_0_2L7_MadSpin	0.0025	0.13	1.00
563448	MGPy8EG_C1mN2_WZ_700_100_2L7_MadSpin	0.0025	0.13	1.00
563450	MGPy8EG_C1mN2_WZ_700_200_2L7_MadSpin	0.0025	0.13	1.00
563452	MGPy8EG_C1mN2_WZ_700_300_2L7_MadSpin	0.0025	0.12	1.00
563454	MGPy8EG_C1mN2_WZ_700_400_2L7_MadSpin	0.0025	0.12	1.00
563456	MGPy8EG_C1mN2_WZ_700_500_2L7_MadSpin	0.0025	0.11	1.00
563457	MGPy8EG_C1mN2_WZ_750_0_2L7_MadSpin	0.0017	0.13	1.00
563459	MGPy8EG_C1mN2_WZ_750_100_2L7_MadSpin	0.0017	0.13	1.00
563461	MGPy8EG_C1mN2_WZ_750_200_2L7_MadSpin	0.0017	0.13	1.00
563463	MGPy8EG_C1mN2_WZ_750_300_2L7_MadSpin	0.0017	0.13	1.00
563465	MGPy8EG_C1mN2_WZ_750_400_2L7_MadSpin	0.0017	0.12	1.00
563467	MGPy8EG_C1mN2_WZ_750_500_2L7_MadSpin	0.0017	0.11	1.00
563468	MGPy8EG_C1mN2_WZ_800_0_2L7_MadSpin	0.0012	0.14	1.00
563470	MGPy8EG_C1mN2_WZ_800_100_2L7_MadSpin	0.0012	0.13	1.00
563472	MGPy8EG_C1mN2_WZ_800_200_2L7_MadSpin	0.0012	0.13	1.00
563474	MGPy8EG_C1mN2_WZ_800_300_2L7_MadSpin	0.0012	0.13	1.00
563476	MGPy8EG_C1mN2_WZ_800_400_2L7_MadSpin	0.0012	0.13	1.00
563478	MGPy8EG_C1mN2_WZ_800_500_2L7_MadSpin	0.0012	0.12	1.00
563479	MGPy8EG_C1mN2_WZ_900_0_2L7_MadSpin	0.0006	0.14	1.00
563481	MGPy8EG_C1mN2_WZ_900_100_2L7_MadSpin	0.0006	0.14	1.00
563483	MGPy8EG_C1mN2_WZ_900_200_2L7_MadSpin	0.0006	0.14	1.00
563485	MGPy8EG_C1mN2_WZ_900_300_2L7_MadSpin	0.0006	0.14	1.00
563487	MGPy8EG_C1mN2_WZ_900_400_2L7_MadSpin	0.0006	0.13	1.00
563489	MGPy8EG_C1mN2_WZ_900_500_2L7_MadSpin	0.0006	0.13	1.00
563490	MGPy8EG_C1mN2_WZ_1000_0_2L7_MadSpin	0.0003	0.14	1.00
563492	MGPy8EG_C1mN2_WZ_1000_100_2L7_MadSpin	0.0003	0.14	1.00
563494	MGPy8EG_C1mN2_WZ_1000_200_2L7_MadSpin	0.0003	0.14	1.00
563496	MGPy8EG_C1mN2_WZ_1000_300_2L7_MadSpin	0.0003	0.14	1.00
563498	MGPy8EG_C1mN2_WZ_1000_400_2L7_MadSpin	0.0003	0.14	1.00
563500	MGPy8EG_C1mN2_WZ_1000_500_2L7_MadSpin	0.0003	0.13	1.00
563501	MGPy8EG_C1mN2_WZ_1100_0_2L7_MadSpin	0.0002	0.15	1.00
563503	MGPy8EG_C1mN2_WZ_1100_100_2L7_MadSpin	0.0002	0.15	1.00
563505	MGPy8EG_C1mN2_WZ_1100_200_2L7_MadSpin	0.0002	0.15	1.00
563507	MGPy8EG_C1mN2_WZ_1100_300_2L7_MadSpin	0.0002	0.14	1.00
563509	MGPy8EG_C1mN2_WZ_1100_400_2L7_MadSpin	0.0002	0.14	1.00
563511	MGPy8EG_C1mN2_WZ_1100_500_2L7_MadSpin	0.0002	0.14	1.00
563512	MGPy8EG_C1mN2_WZ_1200_0_2L7_MadSpin	0.0001	0.15	1.00
563514	MGPy8EG_C1mN2_WZ_1200_100_2L7_MadSpin	0.0001	0.15	1.00
563516	MGPy8EG_C1mN2_WZ_1200_200_2L7_MadSpin	0.0001	0.15	1.00
563518	MGPy8EG_C1mN2_WZ_1200_300_2L7_MadSpin	0.0001	0.15	1.00
563520	MGPy8EG_C1mN2_WZ_1200_400_2L7_MadSpin	0.0001	0.15	1.00
563522	MGPy8EG_C1mN2_WZ_1200_500_2L7_MadSpin	0.0001	0.14	1.00

Table D.7: MC sample information for negatively charged chargino signal samples, grouped by chargino mass. DSID identifies the dataset, while the physics short name contains information for the process, masses, and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

DSID	Physics Short	σ [pb]	ϵ_{filter}	k-factor
563239	MGPy8EG_C1pN2_WZ_200_0_2L7_MadSpin	1.1703	0.09	1.00
563241	MGPy8EG_C1pN2_WZ_250_0_2L7_MadSpin	0.5173	0.10	1.00
563243	MGPy8EG_C1pN2_WZ_250_100_2L7_MadSpin	0.5173	0.09	1.00
563244	MGPy8EG_C1pN2_WZ_300_0_2L7_MadSpin	0.2603	0.10	1.00
563246	MGPy8EG_C1pN2_WZ_300_100_2L7_MadSpin	0.2603	0.10	1.00
563248	MGPy8EG_C1pN2_WZ_350_0_2L7_MadSpin	0.1431	0.11	1.00
563250	MGPy8EG_C1pN2_WZ_350_100_2L7_MadSpin	0.1431	0.10	1.00
563252	MGPy8EG_C1pN2_WZ_350_200_2L7_MadSpin	0.1431	0.10	1.00
563253	MGPy8EG_C1pN2_WZ_400_0_2L7_MadSpin	0.0839	0.11	1.00
563255	MGPy8EG_C1pN2_WZ_400_100_2L7_MadSpin	0.0839	0.11	1.00
563257	MGPy8EG_C1pN2_WZ_400_200_2L7_MadSpin	0.0839	0.10	1.00
563259	MGPy8EG_C1pN2_WZ_450_0_2L7_MadSpin	0.0516	0.12	1.00
563261	MGPy8EG_C1pN2_WZ_450_100_2L7_MadSpin	0.0516	0.11	1.00
563263	MGPy8EG_C1pN2_WZ_450_200_2L7_MadSpin	0.0516	0.11	1.00
563265	MGPy8EG_C1pN2_WZ_450_300_2L7_MadSpin	0.0516	0.10	1.00
563266	MGPy8EG_C1pN2_WZ_500_0_2L7_MadSpin	0.0329	0.12	1.00
563268	MGPy8EG_C1pN2_WZ_500_100_2L7_MadSpin	0.0329	0.12	1.00
563270	MGPy8EG_C1pN2_WZ_500_200_2L7_MadSpin	0.0329	0.11	1.00
563272	MGPy8EG_C1pN2_WZ_500_300_2L7_MadSpin	0.0329	0.11	1.00
563274	MGPy8EG_C1pN2_WZ_550_0_2L7_MadSpin	0.0216	0.12	1.00
563276	MGPy8EG_C1pN2_WZ_550_100_2L7_MadSpin	0.0216	0.12	1.00
563278	MGPy8EG_C1pN2_WZ_550_200_2L7_MadSpin	0.0216	0.12	1.00
563280	MGPy8EG_C1pN2_WZ_550_300_2L7_MadSpin	0.0216	0.11	1.00
563282	MGPy8EG_C1pN2_WZ_550_400_2L7_MadSpin	0.0216	0.10	1.00
563283	MGPy8EG_C1pN2_WZ_600_0_2L7_MadSpin	0.0146	0.12	1.00
563285	MGPy8EG_C1pN2_WZ_600_100_2L7_MadSpin	0.0146	0.12	1.00
563287	MGPy8EG_C1pN2_WZ_600_200_2L7_MadSpin	0.0146	0.12	1.00
563289	MGPy8EG_C1pN2_WZ_600_300_2L7_MadSpin	0.0146	0.12	1.00
563291	MGPy8EG_C1pN2_WZ_600_400_2L7_MadSpin	0.0146	0.11	1.00
563293	MGPy8EG_C1pN2_WZ_650_0_2L7_MadSpin	0.0100	0.13	1.00
563295	MGPy8EG_C1pN2_WZ_650_100_2L7_MadSpin	0.0100	0.13	1.00
563297	MGPy8EG_C1pN2_WZ_650_200_2L7_MadSpin	0.0100	0.13	1.00
563299	MGPy8EG_C1pN2_WZ_650_300_2L7_MadSpin	0.0100	0.12	1.00
563301	MGPy8EG_C1pN2_WZ_650_400_2L7_MadSpin	0.0100	0.11	1.00
563303	MGPy8EG_C1pN2_WZ_650_500_2L7_MadSpin	0.0100	0.10	1.00
563304	MGPy8EG_C1pN2_WZ_700_0_2L7_MadSpin	0.0070	0.13	1.00
563306	MGPy8EG_C1pN2_WZ_700_100_2L7_MadSpin	0.0070	0.13	1.00
563308	MGPy8EG_C1pN2_WZ_700_200_2L7_MadSpin	0.0070	0.13	1.00
563310	MGPy8EG_C1pN2_WZ_700_300_2L7_MadSpin	0.0070	0.12	1.00
563312	MGPy8EG_C1pN2_WZ_700_400_2L7_MadSpin	0.0070	0.12	1.00
563314	MGPy8EG_C1pN2_WZ_700_500_2L7_MadSpin	0.0070	0.11	1.00
563315	MGPy8EG_C1pN2_WZ_750_0_2L7_MadSpin	0.0050	0.13	1.00
563317	MGPy8EG_C1pN2_WZ_750_100_2L7_MadSpin	0.0050	0.13	1.00
563319	MGPy8EG_C1pN2_WZ_750_200_2L7_MadSpin	0.0050	0.13	1.00
563321	MGPy8EG_C1pN2_WZ_750_300_2L7_MadSpin	0.0050	0.13	1.00
563323	MGPy8EG_C1pN2_WZ_750_400_2L7_MadSpin	0.0050	0.12	1.00
563325	MGPy8EG_C1pN2_WZ_750_500_2L7_MadSpin	0.0050	0.11	1.00
563326	MGPy8EG_C1pN2_WZ_800_0_2L7_MadSpin	0.0035	0.13	1.00
563328	MGPy8EG_C1pN2_WZ_800_100_2L7_MadSpin	0.0035	0.13	1.00
563330	MGPy8EG_C1pN2_WZ_800_200_2L7_MadSpin	0.0035	0.13	1.00
563332	MGPy8EG_C1pN2_WZ_800_300_2L7_MadSpin	0.0035	0.13	1.00
563334	MGPy8EG_C1pN2_WZ_800_400_2L7_MadSpin	0.0035	0.12	1.00
563336	MGPy8EG_C1pN2_WZ_800_500_2L7_MadSpin	0.0035	0.12	1.00
563337	MGPy8EG_C1pN2_WZ_900_0_2L7_MadSpin	0.0019	0.14	1.00
563339	MGPy8EG_C1pN2_WZ_900_100_2L7_MadSpin	0.0019	0.14	1.00
563341	MGPy8EG_C1pN2_WZ_900_200_2L7_MadSpin	0.0019	0.14	1.00
563343	MGPy8EG_C1pN2_WZ_900_300_2L7_MadSpin	0.0019	0.13	1.00
563345	MGPy8EG_C1pN2_WZ_900_400_2L7_MadSpin	0.0019	0.13	1.00
563347	MGPy8EG_C1pN2_WZ_900_500_2L7_MadSpin	0.0019	0.13	1.00
563348	MGPy8EG_C1pN2_WZ_1000_0_2L7_MadSpin	0.0010	0.14	1.00
563350	MGPy8EG_C1pN2_WZ_1000_100_2L7_MadSpin	0.0010	0.14	1.00
563352	MGPy8EG_C1pN2_WZ_1000_200_2L7_MadSpin	0.0010	0.14	1.00
563354	MGPy8EG_C1pN2_WZ_1000_300_2L7_MadSpin	0.0010	0.14	1.00
563356	MGPy8EG_C1pN2_WZ_1000_400_2L7_MadSpin	0.0010	0.14	1.00
563358	MGPy8EG_C1pN2_WZ_1000_500_2L7_MadSpin	0.0010	0.13	1.00
563359	MGPy8EG_C1pN2_WZ_1100_0_2L7_MadSpin	0.0006	0.15	1.00
563361	MGPy8EG_C1pN2_WZ_1100_100_2L7_MadSpin	0.0006	0.15	1.00
563363	MGPy8EG_C1pN2_WZ_1100_200_2L7_MadSpin	0.0006	0.15	1.00
563365	MGPy8EG_C1pN2_WZ_1100_300_2L7_MadSpin	0.0006	0.14	1.00
563367	MGPy8EG_C1pN2_WZ_1100_400_2L7_MadSpin	0.0006	0.14	1.00
563369	MGPy8EG_C1pN2_WZ_1100_500_2L7_MadSpin	0.0006	0.14	1.00
563370	MGPy8EG_C1pN2_WZ_1200_0_2L7_MadSpin	0.0003	0.15	1.00
563372	MGPy8EG_C1pN2_WZ_1200_100_2L7_MadSpin	0.0003	0.15	1.00
563374	MGPy8EG_C1pN2_WZ_1200_200_2L7_MadSpin	0.0003	0.15	1.00
563376	MGPy8EG_C1pN2_WZ_1200_300_2L7_MadSpin	0.0003	0.15	1.00
563378	MGPy8EG_C1pN2_WZ_1200_400_2L7_MadSpin	0.0003	0.14	1.00
563380	MGPy8EG_C1pN2_WZ_1200_500_2L7_MadSpin	0.0003	0.14	1.00

Table D.8: MC sample information for positively charged chargino signal samples, grouped by chargino mass. DSID identifies the dataset, while the physics short name contains information for the process, masses, and generator. The cross section (σ), filter efficiency (ϵ_{filter}), and k-factor are given.

Bibliography

[1] OpenAI. *ChatGPT*. Large language model. Accessed: 2026-04-06. 2025. URL: https://chatgpt.com/.

[2] Manjit Dosanjh. *From Particle Physics to Medical Applications*. IOP Publishing, 2017. ISBN: 978-0-7503-1444-2. DOI: 10.1088/978-0-7503-1444-2. URL: http://iopscience.iop.org/book/978-0-7503-1444-2.

[3] CERN. *The birth of the Web*. Accessed: 2026-04-07. URL: https://home.cern/science/computing/birth-web.

[4] CERN. *Experiments*. Accessed: 2026-04-07. URL: https://home.cern/science/experiments.

[5] *ATLAS Open Data*. Accessed: 2026-05-04. URL: https://opendata.atlas.cern/.

[6] Dipak Nath. 'The Darkness of Dark Matter and Dark Energy'. In: *International Journal of Engineering and Applied Sciences* 5.6 (2018), p. 257209. ISSN: 2394-3661. URL: https://www.neliti.com/publications/257209/.

[7] CERN Courier. 'Glimpses of superhistory'. In: *CERN Courier* (Feb. 2001). Accessed: 2026-04-07. URL: https://cerncourier.com/a/glimpses-of-superhistory/.

[8] ATLAS Collaboration. 'Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC'. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.020. URL: http://arxiv.org/abs/1207.7214.

[9] CMS Collaboration. 'Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC'. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://arxiv.org/abs/1207.7235.

[10] Brian E. Granger and Fernando Pérez. 'Jupyter: Thinking and Storytelling With Code and Data'. In: *Computing in Science & Engineering* 23.2 (Mar. 2021), pp. 7–14. ISSN: 1558-366X. DOI: 10.1109/MCSE.2021.3059263. URL: https://ieeexplore.ieee.org/document/9387490.

[11] American Physical Society. *American Physical Society*. Accessed: 2026-04-07. URL: https://www.aps.org/.

[12] American Physical Society. *JNIPER: Joint Network for Informal Physics Education and Research*. Accessed: 2026-04-07. URL: https://www.aps.org/initiatives/advocate-amplify/public-engagement/jniper.

[13] University of Massachusetts Amherst Physics Department. *Outreach and Public Engagement*. Accessed: 2026-04-07. URL: https://www.umass.edu/physics/community/outreach.

[14] University of Cambridge Department of Physics. *Department of Physics Educational Outreach*. Accessed: 2026-04-07. URL: https://outreach.phy.cam.ac.uk/.

[15] University of Melbourne School of Physics. *Engagement and outreach in the School of Physics*. Accessed: 2026-04-07. Jan. 2026. URL: https://physics.unimelb.edu.au/engage.

[16] IPPOG Collaboration. *IPPOG*. Accessed: 2026-04-07. URL: https://ippog.web.cern.ch.

[17] IPPOG Collaboration. *About | Goals*. Accessed: 2026-04-07. URL: https://ippog.org/goals.

[18] Joao Felipe Pimentel et al. 'A Large-Scale Study About Quality and Reproducibility of Jupyter Notebooks'. In: *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)*. IEEE, May 2019, pp. 507–517. ISBN: 978-1-7281-3412-3. DOI: 10.1109/MSR.2019.00077. URL: https://ieeexplore.ieee.org/document/8816763/.

[19] Jeremiah W. Johnson. 'Benefits and Pitfalls of Jupyter Notebooks in the Classroom'. In: *Proceedings of the 21st Annual Conference on Information Technology Education*. ACM, Oct. 2020, pp. 32–37. ISBN: 978-1-4503-7045-5. DOI: 10.1145/3368308.3415397. URL: https://dl.acm.org/doi/10.1145/3368308.3415397.

[20] J. Bascuñana et al. 'Impact of Jupyter Notebook as a tool to enhance the learning process in chemical engineering modules'. In: *Education for Chemical Engineers* 44 (July 2023), pp. 155–163. ISSN: 1749-7728. DOI: 10.1016/j.ece.2023.06.001. URL: https://www.sciencedirect.com/science/article/pii/S174977282300026X.

[21] Project Jupyter et al. 'Binder 2.0 - Reproducible, interactive, sharable environments for science at scale'. In: *scipy* (June 2018). DOI: 10.25080/Majora-4af1f417-011. URL: https://proceedings.scipy.org/articles/Majora-4af1f417-011.

[22] Google Colaboratory. *Google Colab*. Accessed: 2026-05-10. URL: https://colab.research.google.com/.

[23] R. Michael Barnett and K. Erik Johansson. *The Impact of an Innovative Education and Outreach Project by a Physics Experiment*. Oct. 2023. DOI: 10.48550/arXiv.2306.13207. URL: http://arxiv.org/abs/2306.13207.

[24] *The ATLAS Open Data project*. Accessed: 2026-05-04. URL: https://opendata.atlas.cern/docs/aboutus.

[25] ATLAS Collaboration. *Review of the 13 TeV ATLAS Open Data release*. Apr. 2020. URL: https://cds.cern.ch/record/2707171/files/ANA-OTRC-2019-01-PUB-updated.pdf.

[26] ATLAS Collaboration. *Analysis notebooks | ATLAS Open Data*. Accessed: 2026-04-07. URL: https://opendata.atlas.cern/docs/13TeV25Doc.

[27] *Galaxy Zpath*. Accessed: 2026-04-07. URL: https://galaxy-hepp.hpc.uio.no/.

[28] The Galaxy Community. 'The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update'. In: *Nucleic Acids Research* 52.W1 (July 2024), W83–W94. ISSN: 0305-1048. DOI: 10.1093/nar/gkae410. URL: https://doi.org/10.1093/nar/gkae410.

[29] University of Oslo Department of Physics. *ZPATH - Department of Physics*. Accessed: 2026-04-07. URL: https://www.mn.uio.no/fysikk/english/research/projects/zpath/index.html.

[30] ATLAS Collaboration. 'Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13$ TeV with the ATLAS detector'. In: *The European Physical Journal C* 78.12 (Dec. 2018), p. 995. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-018-6423-7. URL: http://arxiv.org/abs/1803.02762.

[31] ATLAS Collaboration. 'Search for dark photons from Higgs boson decays via ZH production with a photon plus missing transverse momentum signature from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector'. In: *Journal of High Energy Physics* 2023.7 (July 2023), p. 133. ISSN: 1029-8479. DOI: 10.1007/JHEP07(2023)133. URL: http://arxiv.org/abs/2212.09649.

[32] Gianfranco Bertone and Dan Hooper. 'A History of Dark Matter'. In: *Reviews of Modern Physics* 90.4 (Oct. 2018), p. 045002. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.90.045002. URL: http://arxiv.org/abs/1605.04909.

[33] F. Zwicky. 'On the Masses of Nebulae and of Clusters of Nebulae'. In: *Astrophys. J.* 86 (1937), pp. 217–246. DOI: 10.1086/143864.

[34] Henk Hoekstra, Howard Yee and Mike Gladders. 'Current status of weak gravitational lensing'. In: *New Astronomy Reviews* 46.12 (Nov. 2002), pp. 767–781. ISSN: 13876473. DOI: 10.1016/S1387-6473(02)00245-2. URL: http://arxiv.org/abs/astro-ph/0205205.

[35] Francesca Chadha-Day, John Ellis and David J. E. Marsh. 'Axion Dark Matter: What is it and Why Now?' In: (Oct. 2022). DOI: 10.48550/arXiv.2105.01406. URL: http://arxiv.org/abs/2105.01406.

[36] Mark Thompson. *Modern Particle Physics*. Cambridge University Press, 2013. ISBN: 978-1-107-03426-6.

[37] Izaak Neutelings. *Standard Model*. Accessed: 2025-10-07. Sept. 2023. URL: https://tikz.net/sm_particles/.

[38] *Deed - Attribution-ShareAlike 4.0 International - Creative Commons*. URL: https://creativecommons.org/licenses/by-sa/4.0/.

[39] Antonio Pich. *The Standard Model of Electroweak Interactions*. Jan. 2012. DOI: 10.48550/arXiv.1201.0537. URL: http://arxiv.org/abs/1201.0537.

[40] Emma Ward. *LHC detector transverse cross-section*. Mar. 2019. URL: https://cds.cern.ch/images/ATLAS-PHOTO-2019-011-3.

[41] Michael Klasen, Martin Pohl and Günter Sigl. 'Indirect and direct search for dark matter'. In: *Progress in Particle and Nuclear Physics* 85 (Nov. 2015), pp. 1–32. ISSN: 01466410. DOI: 10.1016/j.ppnp.2015.07.001. URL: http://arxiv.org/abs/1507.03800.

[42] 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'. In: *Journal of High Energy Physics* 2018.10 (Oct. 2018), p. 180. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2018)180. URL: http://arxiv.org/abs/1807.11471.

[43] Nicole F. Bell et al. 'Searching for Dark Matter at the LHC with a Mono-Z'. In: *Physical Review D* 86.9 (Nov. 2012), p. 096011. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.86.096011. URL: http://arxiv.org/abs/1209.0231.

[44] Mariangela Lisanti. *Lectures on Dark Matter Physics*. Jan. 2017. DOI: 10.1142/9789813149441_0007. URL: http://arxiv.org/abs/1603.03797.

[45] Debarshi Mukherjee. *Dark Matter Freeze-Out and Relic Density in an Expanding Universe*. Tech. rep. Zenodo, Aug. 2025. DOI: 10.5281/zenodo.16735390. URL: https://zenodo.org/records/16735390.

[46] Marc Kamionkowski. *WIMP and Axion Dark Matter*. 1997. URL: https://arxiv.org/abs/hep-ph/9710467.

[47] Gianfranco Bertone, Dan Hooper and Joseph Silk. 'Particle Dark Matter: Evidence, Candidates and Constraints'. In: *Physics Reports* 405.5-6 (Jan. 2005), pp. 279–390. ISSN: 03701573. DOI: 10.1016/j.physrep.2004.08.031. URL: http://arxiv.org/abs/hep-ph/0404175.

[48] Stephen P. Martin. *A Supersymmetry Primer*. July 1998. DOI: 10.1142/9789812839657_0001. URL: http://arxiv.org/abs/hep-ph/9709356.

[49] Howard E. Haber. 'The Status of the Minimal Supersymmetric Standard Model and Beyond'. In: *Nuclear Physics B - Proceedings Supplements* 62.1-3 (Mar. 1998), pp. 469–484. ISSN: 09205632. DOI: 10.1016/S0920-5632(97)00688-9. URL: http://arxiv.org/abs/hep-ph/9709450.

[50] A. Djouadi et al. *The Minimal Supersymmetric Standard Model: Group Summary Report*. Jan. 1999. DOI: 10.48550/arXiv.hep-ph/9901246. URL: http://arxiv.org/abs/hep-ph/9901246.

[51] The CDF Collaboration. 'Search for new physics in the mu+mu+e/mu+met channel with a low-pT lepton threshold at the Collider Detector at Fermilab'. In: *Physical Review D* 79.5 (Mar. 2009), p. 052004. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.79.052004. URL: http://arxiv.org/abs/0810.3522.

[52] J. A. Aguilar-Saavedra et al. 'Semi-dark Higgs decays: sweeping the Higgs neutrino floor'. In: *Physical Review D* 106.11 (Dec. 2022), p. 115023. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.106.115023. URL: http://arxiv.org/abs/2206.01214.

[53] Marco Fabbrichesi, Emidio Gabrielli and Gaia Lanfranchi. *The Dark Photon*. 2021. DOI: 10.1007/978-3-030-62519-1. URL: http://arxiv.org/abs/2005.01515.

[54] G. Barello, Spencer Chang and Christopher A. Newby. 'Correlated Signals at the Energy and Intensity Frontiers from Nonabelian Kinetic Mixing'. In: *Physical Review D* 94.5 (Sept. 2016), p. 055018. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.94.055018. URL: http://arxiv.org/abs/1511.02865.

[55] Maria Cepeda et al. 'Exotic Higgs Decays'. In: *Annual Review of Nuclear and Particle Science* 72.1 (Sept. 2022), pp. 119–149. ISSN: 0163-8998, 1545-4134. DOI: 10.1146/annurev-nucl-102319-024147. URL: http://arxiv.org/abs/2111.12751.

[56] Carmine Elvezio Pagliarone. 'Collider Physics an Experimental Introduction'. In: *Journal of Physics: Conference Series* 287.1 (Apr. 2011), p. 012005. ISSN: 1742-6596. DOI: 10.1088/1742-6596/287/1/012005. URL: https://dx.doi.org/10.1088/1742-6596/287/1/012005.

[57] Hermann Kolanoski and Norbert Wermes. *Particle Detectors: Fundamentals and Applications*. Oxford University Press, 2020. ISBN: 978-0-19-885836-2.

[58] ATLAS Collaboration. *ATLAS Fact Sheets*. Accessed: 2026-04-26. URL: https://atlas.cern/Resources/Fact-sheets.

[59] CERN. Geneva. CERN Accelerator School. *CAS - CERN Accelerator School: Intermediate Accelerator Physics*. Tech. rep. CERN, 2006. DOI: 10.5170/CERN-2006-002. URL: https://cds.cern.ch/record/603056.

[60] Glen Cowan et al. 'Asymptotic formulae for likelihood-based tests of new physics'. In: *arXiv.org* (July 2010). DOI: 10.1140/epjc/s10052-011-1554-0. URL: https://arxiv.org/abs/1007.1727v3.

[61] Particle Data Group. *Reviews, Tables, and Plots*. Accessed: 2026-04-07. URL: https://pdg.lbl.gov/2025/reviews/contents_sports.html.

[62] Alexander L. Read. 'Presentation of search results: The CLs technique'. In: *J. Phys. G* 28 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[63] Bernhard Mistlberger and Falko Dulat. *Limit setting procedures and theoretical uncertainties in Higgs boson searches*. May 2012. DOI: 10.48550/arXiv.1204.3851. URL: http://arxiv.org/abs/1204.3851.

[64] S. S. Wilks. 'The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses'. In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.

[65] ATLAS Collaboration. *ATLAS Open Magic | ATLAS Open Data*. Accessed: 2026-04-07. URL: https://opendata.atlas.cern/docs/atlasopenmagic.

[66] Jim Pivarski et al. *scikit-hep/uproot: 3.12.0*. July 2020. DOI: 10.5281/zenodo.3952728. URL: https://zenodo.org/records/3952728.

[67] Jim Pivarski et al. *scikit-hep/awkward-array: 0.13.0*. July 2020. DOI: 10.5281/zenodo.3952674. URL: https://zenodo.org/records/3952674.

[68] Saransh Chopra et al. 'Vector: JIT-compilable mathematical manipulations of ragged Lorentz vectors'. In: *Journal of Open Source Software* 10.109 (May 2025), p. 7791. ISSN: 2475-9066. DOI: 10.21105/joss.07791. URL: https://joss.theoj.org/papers/10.21105/joss.07791.

[69] Christopher G. Lester and Benjamin Nachman. 'Bisection-based asymmetric MT2 computation: a higher precision calculator than existing symmetric methods'. In: *Journal of High Energy Physics* 2015.3 (Mar. 2015), p. 100. ISSN: 1029-8479. DOI: 10.1007/JHEP03(2015)100. URL: http://arxiv.org/abs/1411.4312.

[70] Charles R. Harris et al. 'Array programming with NumPy'. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2649-2. URL: https://www.nature.com/articles/s41586-020-2649-2.

[71] Pauli Virtanen et al. 'SciPy 1.0: fundamental algorithms for scientific computing in Python'. In: *Nature Methods* 17.3 (Mar. 2020), pp. 261–272. ISSN: 1548-7105. DOI: 10.1038/s41592-019-0686-2. URL: https://www.nature.com/articles/s41592-019-0686-2.

[72] John D. Hunter. 'Matplotlib: A 2D Graphics Environment'. In: *Computing in Science & Engineering* 9.3 (May 2007), pp. 90–95. ISSN: 1558-366X. DOI: 10.1109/MCSE.2007.55. URL: https://ieeexplore.ieee.org/document/4160265/.

[73] The pandas development team. *pandas-dev/pandas: Pandas*. Sept. 2025. DOI: 10.5281/zenodo.17229934. URL: https://zenodo.org/records/17229934.

[74] ATLAS Collaboration. *ROOT ntuple format 2015-2016 proton-proton Open Data for Education and Outreach beta release from the ATLAS experiment*. CERN Open Data Portal. 2025. DOI: 10.7483/OPENDATA.ATLAS.B5M9.44TN.

[75] ATLAS Collaboration. 'Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC'. In: *The European Physical Journal C* 83.10 (Oct. 2023), p. 982. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11747-w. URL: http://arxiv.org/abs/2212.09379.

[76] ATLAS Collaboration. *13 TeV 2025 Data - Beta*. Accessed: 2026-04-26. 2025. URL: https://opendata.atlas.cern/docs/data/for_education/13TeV25_details.

[77] C. G. Lester and D. J. Summers. 'Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders'. In: *Physics Letters B* 463.1 (Sept. 1999), pp. 99–103. ISSN: 03702693. DOI: 10.1016/S0370-2693(99)00945-4. URL: http://arxiv.org/abs/hep-ph/9906349.

[78] ATLAS Collaboration. 'Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector'. In: *The European Physical Journal C* 80.2 (Feb. 2020), p. 123. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-7594-6. URL: http://arxiv.org/abs/1908.08215.

[79] ATLAS Collaboration. 'Search for direct pair production of sleptons and charginos decaying to two leptons and neutralinos with mass splittings near the W-boson mass in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector'. In: *Journal of High Energy Physics* 2023.6 (June 2023), p. 31. ISSN: 1029-8479. DOI: 10.1007/JHEP06(2023)031. URL: http://arxiv.org/abs/2209.13935.

[80] CMS Collaboration. 'Search for invisible decays of Higgs bosons in the vector boson fusion and associated ZH production modes'. In: *The European Physical Journal C* 74.8 (Aug. 2014), p. 2980. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-014-2980-6. URL: http://arxiv.org/abs/1404.1344.

[81] ATLAS Collaboration. 'ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV'. In: *The European Physical Journal C* 79.11 (Nov. 2019), p. 970. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-7450-8. URL: http://arxiv.org/abs/1907.05120.

[82] ATLAS Collaboration. *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*. Geneva, May 2020. URL: https://cds.cern.ch/record/2718948.

[83] ATLAS Collaboration. 'Tagging and suppression of pileup jets with the ATLAS detector'. In: (May 2014). URL: https://cds.cern.ch/record/1700870.

[84] ATLAS Collaboration. 'Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector'. In: *Journal of High Energy Physics* 2014.9 (Sept. 2014), p. 37. ISSN: 1029-8479. DOI: 10.1007/JHEP09(2014)037. URL: http://arxiv.org/abs/1407.7494.

[85] Magnar K. Bugge. *FYS5555 Statistics Notebook*. Course Materials. Accessed: 2025-12-04. 2022. URL: https://galaxy-hepp.hpc.uio.no/.

[86] ATLAS Collaboration. *Notebooks Collection Opendata*. Accessed: 2026-04-07. URL: https://github.com/atlas-outreach-data-tools/notebooks-collection-opendata/tree/master/13-TeV-examples/uproot_python.

[87] Emily Puchala. *MasterThesis*. Github. 2026. URL: https://github.com/emilypuchala/MasterThesis.

[88] ATLAS Collaboration. 'Search for new phenomena in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector'. In: *The European Physical Journal C* 77.3 (Mar. 2017), p. 144. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-017-4700-5. URL: http://arxiv.org/abs/1611.05791.