A step towards Supersymmetry? Searching for dark matter at the LHC with the ATLAS detector, using missing transverse energy and leptons

Alessandro Ruggiero

Magdalen College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2025

Abstract

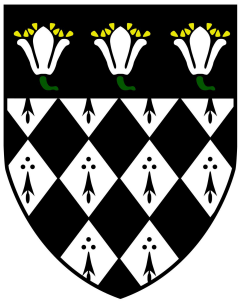
This thesis focuses on the search for new physics using the ATLAS detector. The nature of dark matter remains one of the fundamental open questions in modern physics, and the absence of a viable dark matter candidate within the Standard Model (SM) is a significant limitation of the theory. Supersymmetry (SUSY) offers a compelling extension of the SM, providing a natural dark matter candidate while also addressing several other shortcomings. The central content of this thesis covers searches for the direct pair production of charged light-flavour supersymmetric sleptons ($\tilde{e}$ or $\tilde{\mu}$), each decaying into a stable neutralino ($\tilde{\chi}_1^0$) and an associated Standard Model lepton. The thesis concludes with a study targeting improvements for Run 3, in the trigger used to collect the data for the searches.

The analysis targets the challenging "corridor" region where the mass difference, Δm, between the slepton and the lightest neutralino is small. This region has historically been difficult to probe, and the analysis aims to close the persistent gap in sensitivity to models with $\Delta m \lesssim m(W)$. Events are selected by requiring a high-energy jet, significant missing transverse momentum, and two same-flavour opposite-sign leptons (e or μ). The study uses proton–proton collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector, corresponding to an integrated luminosity of 140 fb^{-1}. A series of kinematic selections is applied, including multiple boosted decision trees optimised for different values of Δm. This approach delivers expected sensitivity across the full Δm corridor for the first time.

The results are generally consistent with Standard Model expectations, where the largest deviations are observed with local significances of $2.0\,\sigma$ in the selectron search and $2.4\,\sigma$ in the smuon search. Discovery significances are calculated for specific SUSY signal models. For the selectron search, a maximum discovery significance of 1.9σ is obtained for a model with selectron mass of 375 GeV, with a neutralino mass of 335 GeV. For the smuon search, a maximum discovery significance of 2.1σ is obtained for a model with smuon mass of 150 GeV, with a neutralino mass of 140 GeV. Although these excesses weaken the observed exclusion reach in some regions of parameter space, the previously unaddressed gap in sensitivity in this corridor region is substantially reduced. At 95% confidence level, constraints are set on simplified models of selectron and smuon pair production, with selectron (smuon) masses up to 300 (350) GeV excluded for Δm values between 2 GeV and 100 GeV.

A step towards Supersymmetry? Searching for dark matter at the LHC with the ATLAS detector, using missing transverse energy and leptons





Alessandro Ruggiero

Magdalen College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Abstract

This thesis focuses on the search for new physics using the ATLAS detector. The nature of dark matter remains one of the fundamental open questions in modern physics, and the absence of a viable dark matter candidate within the Standard Model (SM) is a significant limitation of the theory. Supersymmetry (SUSY) offers a compelling extension of the SM, providing a natural dark matter candidate while also addressing several other shortcomings. The central content of this thesis covers searches for the direct pair production of charged light-flavour supersymmetric sleptons ($\tilde{e}$ or $\tilde{\mu}$), each decaying into a stable neutralino ($\tilde{\chi}_1^0$) and an associated Standard Model lepton. The thesis concludes with a study targeting improvements for Run 3, in the trigger used to collect the data for the searches.

The analysis targets the challenging "corridor" region where the mass difference, Δm, between the slepton and the lightest neutralino is small. This region has historically been difficult to probe, and the analysis aims to close the persistent gap in sensitivity to models with $\Delta m \lesssim m(W)$. Events are selected by requiring a high-energy jet, significant missing transverse momentum, and two same-flavour opposite-sign leptons (e or μ). The study uses proton–proton collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector, corresponding to an integrated luminosity of 140 fb^{-1}. A series of kinematic selections is applied, including multiple boosted decision trees optimised for different values of Δm. This approach delivers expected sensitivity across the full Δm corridor for the first time.

The results are generally consistent with Standard Model expectations, where the largest deviations are observed with local significances of $2.0\,\sigma$ in the selectron search and $2.4\,\sigma$ in the smuon search. Discovery significances are calculated for specific SUSY signal models. For the selectron search, a maximum discovery significance of 1.9σ is obtained for a model with selectron mass of 375 GeV, with a neutralino mass of 335 GeV. For the smuon search, a maximum discovery significance of 2.1σ is obtained for a model with smuon mass of 150 GeV, with a neutralino mass of 140 GeV. Although these excesses weaken the observed exclusion reach in some regions of parameter space, the previously unaddressed gap in sensitivity in this corridor region is substantially reduced. At 95% confidence level, constraints are set on simplified models of selectron and smuon pair production, with selectron (smuon) masses up to 300 (350) GeV excluded for Δm values between 2 GeV and 100 GeV.

Acknowledgements

When I began this journey, I don't think I truly realised how big a challenge I was taking on, and if you had told me back when I left Downend in 2017 what I'd end up achieving, I'm not sure I would have believed you. At the start of my DPhil, my knowledge of experimental particle physics could probably have filled a single page. To have gone from that starting point, to guiding an ATLAS analysis from beginning to end, has been the journey of a lifetime. I could never have completed this alone and there are so many people to whom I owe a deep and genuine gratitude, in both work and life.

First and foremost, I have to thank my supervisor Alan, for guiding me through the last 3.5 years and always making time to answer my questions – even when a 'quick' chat turns into an hour. It has been a real pleasure to work with you, but I also want to thank you for looking out for me, and most importantly for believing in me.

Next I have to thank Claudia, I am not sure that I could have achieved this without you. Thank you for having such endless patience through all my 'Alessandro problems', for sharing your endless knowledge, and for your invaluable friendship. I hope our paths cross again one day.

Maggie, we began this journey side by side and faced it all together, from the very first day and through all the highs and lows, you have been there every step of the way. *We did it.*

Sophie, I don't know if I can ever thank you enough, you were one of the first true friends I made at Magdalen. You are one of the strongest people I know – brighter times are just around the bend.

Anna, I am so glad to have met you along the course of this journey, you are one of the best people I know, and thank you for everything, you deserve nothing but the best in whatever the future holds for you. I hope whatever comes next for both of us we remain friends and are able to meet again at the 'sunset kebab' bench one day.

Ynyr, thank you for listening to me complain, reading all my nonsense Skype messages, and the many (many) beers. I hope we reunite very soon – in a pub of course.

Seb, hermano you might be the funniest person I know, and you made this experience

so much more enjoyable. I cannot wait to learn more Venezuelan sayings from you, just a shame about the football team you support...

I also need to thank the whole Oxford SM+ group, across my time here, for being such a kind and welcoming group of people. In particular, Claire, Eimear and Koichi for being the ever-presents, and to Claire and Eimear for being the driving force behind the Tuesday pub quizzes. An additional thank you to the rest of the wider Oxford ATLAS group for the interesting conversation and pub trips, and to Kim and Sue for being such incredible help with any and all admin questions. To my friends from Oxford: Beth, Sparshita, Gregor, Tom, Radek, Baris, Siyuan, Laura, Brian, Chris P, Jon, Holly, Niamh, Josh, Federico, Rob, Seraphim, Ollie, Ben and Ruben, thank you all. A special thank you to Simon, for all the stimulating conversations and listening to me complain all the time – I wonder how many times we have got a coffee from Horsebox?

Thank you to everyone in the ATLAS SUSY group who made it such a welcoming place, specifically Sara, Dani, Simran, Ben and Jonas. In addition, a special thank you to Jeff and Michael for showing me the ropes as an analysis contact, and listening to all my rambling in the many hours of meetings we sat through.

To the first group of friends I met in college: Pascaline, Joe, Will C, Frankie, Grace, Will S, and Sophie (again), thank you so much for helping me to settle into Oxford life.

To everyone I met and became friends with on LTA that I have not already mentioned, we had some great times and I hope we bump into each other again soon. An additional thank you to the old heads of LTA, Iza, Mei, Chris, Julia, Tim, Mark and Giovanna (and Ynyr again), thank you for taking me under your wings and only slightly traumatising me on some memorable skiing and hiking trips.

Thank you to the third (and fourth) year college crew, Nina, Lindsey, Tom, Grace, and James (and Joe and Sophie again). We have had some great times, and I am so glad that we all ended up coming together after I returned from LTA.

To all my Kings friends, especially Riley, Jodie, Ben and April: thank you for all the good times we had in London and beyond, I hope we are able to reunite soon.

To everyone from the Downend days, especially Joseph and Samuel, I still can't believe where my life has taken me, but thank you for keeping me humble.

Finally, thank you to my parents for making me the person I am today.

This is somewhere to be. This is all you have, but it's still something. Streets and sodium lights. The sky, the world. You're alive.

Preface

This thesis contains both the work of the author and of collaborators. This section states the original contributions of the author to the work presented in this thesis. All figures from ATLAS conference notes are labelled 'ATLAS preliminary', and figures from ATLAS papers are labelled 'ATLAS', with appropriate references provided in both cases. All unpublished figures do not contain any ATLAS labelling, as per the ATLAS thesis guidelines. Where figures have been produced by collaborators but are not published, this is indicated in the caption.

A Run 2 search for compressed sleptons

Chapter 5 concerns work related to a search for compressed sleptons. I developed everything described in the chapter, other than the cut-and-count approach and the data-driven fakes estimate. Specifically, I worked on:

- Developing the BDT approach, from start to finish.
- Training the BDTs, deciding to use multiple BDTs across the target region
- The background estimate strategy.
- Deriving theory uncertainties.
- Running the fits, extracting the exclusion limits.
- Investigating the excesses, and producing the signal model significances.
- Maintaining the Histfitter statistical fit framework for the analysis team.
- Additionally, I was one of the analysis contacts responsible for directing analysis strategy and navigating through the ATLAS approval process.

The ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger

Chapter 6 concerns work carried out as part of my qualification project to optimising the ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger for Run 3. I carried out the optimisation of the pileup suppression algorithm 'pufit' covered in this chapter. The performance evaluation with Run 3 data was carried out by a collaborator, and this is indicated in the relevant figures.

Contents

1 Introduction **1**

2 The LHC and the ATLAS Detector **5**
2.1 The Large Hadron Collider 6
2.1.1 Proton–proton Collisions, Luminosity and Pileup 7
2.2 The ATLAS Detector 10
2.2.1 The ATLAS Coordinate System 10
2.2.2 The Magnets 12
2.2.3 The Inner Detector 13
2.2.4 The Calorimeters 15
2.2.5 The Muon Spectrometer 17
2.3 Trigger System 19
2.3.1 Level 1 Trigger 20
2.3.2 High Level Trigger 21

3 Event Simulation, Reconstruction and Object Identification **23**
3.1 Monte-Carlo Generation 24
3.1.1 Event Generation 24
3.1.2 Detector Simulation 25
3.2 Low Level Objects 26
3.2.1 Tracks 26
3.2.2 Primary Vertex 28
3.2.3 Calorimeter Clusters 28
3.2.4 Particle Flow Objects 29
3.3 Object Reconstruction 30
3.3.1 Jets 31
3.3.2 Jet Flavour Tagging 33
3.3.3 Electrons 33
3.3.4 Muons 36
3.3.5 Lepton Isolation 38
3.3.6 Overlap Removal 40
3.3.7 Missing Transverse Energy 40

4 The Standard Model and Supersymmetry 43
4.1 Supersymmetry . . . 44
4.2 Problems with The Standard Model . . . 46
4.2.1 Dark Matter and R-Parity . . . 46
4.2.2 Hierarchy Problem and the Higgs Mass . . . 49
4.2.3 The Anomalous Muon Gyro-magnetic Ratio . . . 51
4.2.4 Grand Unification . . . 52
4.3 The Minimal Supersymmetric Standard Model . . . 53
4.3.1 Symmetry Breaking . . . 54
4.3.2 Sparticles and Mixing . . . 55
4.3.3 Simplified Models . . . 57
4.4 Experimental Searches for SUSY . . . 60
4.4.1 Strong SUSY . . . 61
4.4.2 Electroweak SUSY . . . 63
4.5 Summary . . . 69

5 An ATLAS Run 2 Search for Sleptons 71
5.1 Analysis Strategy . . . 72
5.2 Datasets, MC samples and Objects . . . 74
5.2.1 Dataset and Triggers . . . 74
5.2.2 Background MC Samples . . . 75
5.2.3 Signal MC Samples . . . 75
5.2.4 Object Definitions . . . 76
5.3 Event Selection . . . 78
5.3.1 Kinematic Variables . . . 78
5.3.2 Preselection Requirements . . . 82
5.3.3 BDT Training . . . 85
5.3.4 Signal Region Definitions . . . 96
5.4 Background Estimation . . . 98
5.4.1 Irreducible Backgrounds . . . 98
5.4.2 Reducible Backgrounds . . . 99
5.4.3 Control and Validation Regions . . . 102
5.5 Statistical Methods . . . 106
5.5.1 Fit types . . . 106
5.5.2 Toy Pseudo-experiments . . . 108
5.6 Systematic Uncertainties . . . 109
5.6.1 Experimental Uncertainties . . . 109
5.6.2 Theoretical Uncertainties . . . 111
5.7 Results . . . 117

5.7.1 Control Regions . . . 117
5.7.2 Validation Regions . . . 118
5.7.3 Signal Regions . . . 119
5.7.4 Model Independent SR Significances . . . 122
5.7.5 SR-BDT$^{\mu\mu}_{5+10}$ Investigations . . . 123
5.7.6 SR-BDT$^{\mathrm{ee}}_{40+50}$ Investigations . . . 131
5.8 Interpretations . . . 132
5.8.1 Exclusion Limits . . . 133
5.8.2 Signal Model Discovery Significances . . . 141
5.9 The Wider Context . . . 143
5.10 Summary . . . 151

6 $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger Optimisation **153**
6.1 The ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger . . . 154
6.1.1 Level 1 $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger . . . 155
6.1.2 HLT $E_{\mathrm{T}}^{\mathrm{miss}}$ Algorithms . . . 155
6.1.3 Run 2 Primary Triggers . . . 157
6.2 Pufit Algorithm Optimisation . . . 158
6.2.1 Datasets and MC Samples . . . 160
6.2.2 Optimisation . . . 162
6.3 Performance in Run 3 data . . . 166

7 Conclusions **169**

Appendices

A MC Simulation **173**
A.1 Event Generation . . . 173
A.1.1 Parton Distribution Functions . . . 174
A.1.2 Matrix Element . . . 175
A.1.3 Parton Shower . . . 175
A.1.4 Hadronisation . . . 177
A.1.5 Multi-parton Interactions . . . 177

B Flavour Dependency of m_{T2} **179**
B.1 The graphical approach . . . 180
B.2 The mathematical approach . . . 185

C The Statistical Fits **189**
C.1 CR-only fit . . . 190
C.2 Exclusion fit . . . 191
C.3 Model-dependent discovery fit . . . 193
C.4 Model-independent discovery fit . . . 194
C.5 Toys . . . 194

D Toys and Systematic Symmetrisation **197**
D.1 Unstable toy fit results . . . 197
D.2 Number of toys . . . 200

E Top Generator Systematics **201**

F Additional Information for the BDT Results **205**
F.1 Breakdown of diboson background contributions in analysis regions 205
F.2 Signal Region Post-fit Yield Tables . . . 207
F.3 Additional Excess Information . . . 208
F.4 Post-fit NP pulls . . . 212
F.5 CL_s Values For Each Interpretation . . . 217

G The Pileup Fit **225**
G.1 Particle Flow Object implementation of pufit . . . 227

References **229**

If you focus on what you left behind, you will never see what lies ahead!

Ratatouille

1 Introduction

The study of particle physics strives to understand the fundamental nature of the universe and the origins of human existence. This pursuit has deep roots in humanity's long-standing desire to understand the building blocks of matter. From the concept of indivisible atoms first proposed by Ancient Greek philosophers, to the discovery of the electron in the late 19th century, the field has evolved dramatically over time. The 20th century witnessed revolutionary developments in the field, including the formulation of quantum mechanics and the identification of a growing number of subatomic particles. These discoveries culminated in the development of the Standard Model, a unified theoretical framework describing the electromagnetic, weak, and strong interactions among fundamental particles. The Standard Model of particle physics provides an extremely successful framework for understanding the fundamental forces and constituents of matter and is one of the most thoroughly tested theories in the history of science.

The Large Hadron Collider (LHC) is the most powerful particle accelerator ever constructed, capable of colliding protons at unprecedented centre-of-mass energies. The 'A Toroidal LHC Apparatus' (ATLAS) detector is one of two general purpose particle detectors built at the LHC, along with Compact Muon Solenoid (CMS) experiment. The LHC and its two flagship detectors were primarily designed to discover the Higgs boson, a goal successfully achieved by both experiments in 2012.

An overview of the LHC and the ATLAS detector is provided in Chapter 2, and Chapter 3 covers the prescriptions undertaken in order to turn detector signals into objects that can be used for data analysis. The discovery of the Higgs Boson stands as one of the major scientific milestones of modern physics, but the work of these experiments has continued far beyond this discovery.

Despite its successes, the Standard Model cannot account for several key phenomena observed in the universe. Among the most compelling shortcomings of the Standard Model is its inability to account for dark matter, a non-baryonic form of matter that interacts primarily via gravity. Astrophysical and cosmological observations, indicate that dark matter constitutes approximately 85% of the total matter content of the universe. However, none of the particles within the Standard Model possess the properties required to serve as a viable dark matter candidate. As a result, theories of physics beyond the Standard Model (BSM) aim to address these limitations by proposing new mechanisms and particles. Since the discovery of the Higgs boson, one of the primary objectives of the ATLAS experiment has been the search for BSM physics, including potential dark matter candidates. Chapter 4 discusses the limitations of the Standard Model and introduces Supersymmetry as a promising BSM theory capable of addressing these issues – most notably by providing a natural dark matter candidate.

ATLAS searches for dark matter through indirect detection, looking for characteristic signatures such as events with large missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$), which can arise from particles that escape the detector without interacting. Chapter 5 presents a search for electroweakly produced supersymmetric scalar particles, known as sleptons, using on the 140fb^{-1} 13 TeV Run 2 ATLAS dataset with boosted decision trees to discriminate between signal and background events. The analysis is based on the work presented in Ref. [1], which has been submitted to the Journal of High Energy Physics (JHEP). This study achieves a new level of expected exclusion sensitivity in a region of supersymmetric parameter space where LHC experiments have made limited progress since the era of the Large Electron-Positron Collider (LEP). In addition, two signal regions have observed excesses

with statistical significances of 2.0 σ and 2.4 σ, in the electron and muon channels respectively. These results offer a promising avenue for further investigation with the new Run 3 dataset to clarify the origin of these discrepancies.

It is important not only to analyse the data that has already been recorded, but to also facilitate the collection of new data to analyse in the future. In Run 3, the increased number of simultaneous interactions in the detector (pileup), meant that there were significant challenges in trigger implementation and operation. To this end, Chapter 6 describes efforts to improve the ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger for Run 3 of the LHC. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is particularly sensitive to these pileup effects and so it was important to re-optimise the pileup suppression algorithm used by the trigger.

Finally, in Chapter 7 the thesis is summarised and conclusions are presented.

To see the world, things dangerous to come to, to see behind walls, draw closer, to find each other, and to feel. That is the purpose of life.

The Secret Life of Walter Mitty

2

The LHC and the ATLAS Detector

Contents

2.1 The Large Hadron Collider 6
2.1.1 Proton–proton Collisions, Luminosity and Pileup 7
2.2 The ATLAS Detector 10
2.2.1 The ATLAS Coordinate System 10
2.2.2 The Magnets 12
2.2.3 The Inner Detector 13
2.2.4 The Calorimeters 15
2.2.5 The Muon Spectrometer 17
2.3 Trigger System 19
2.3.1 Level 1 Trigger 20
2.3.2 High Level Trigger 21

This chapter will briefly describe the LHC [2], the largest, most complex machine ever built, and protons are accelerated and collided. The ATLAS detector [3], the largest multi-purpose particle detector ever built, will be discussed, focusing on how it records the vast amount of proton-proton (pp) collision data produced by the LHC. Information will focus on the ATLAS detector as it existed in Run 2, as this is relevant to the work in this thesis. Since Run 2 (2015–2018), upgrades have been carried out on the detector, for Run 3 (2022–2026) and beyond, but these will not be covered.

2.1 The Large Hadron Collider

The LHC is a 26.7 km circumference particle accelerator, located approximately 100 m underground, under the French-Swiss border, near Geneva. In Run 2, the LHC collided protons [1] at centre-of-mass energies of $\sqrt{s} = 13$ TeV, using two oppositely circulating 6.5 TeV beams. This energy has never been reached due to the extreme difficulty in training the LHC magnets to reach the required strength. The full CERN accelerator and detector complex, consisting of many particle accelerators and experiments, is shown in Figure 2.1.

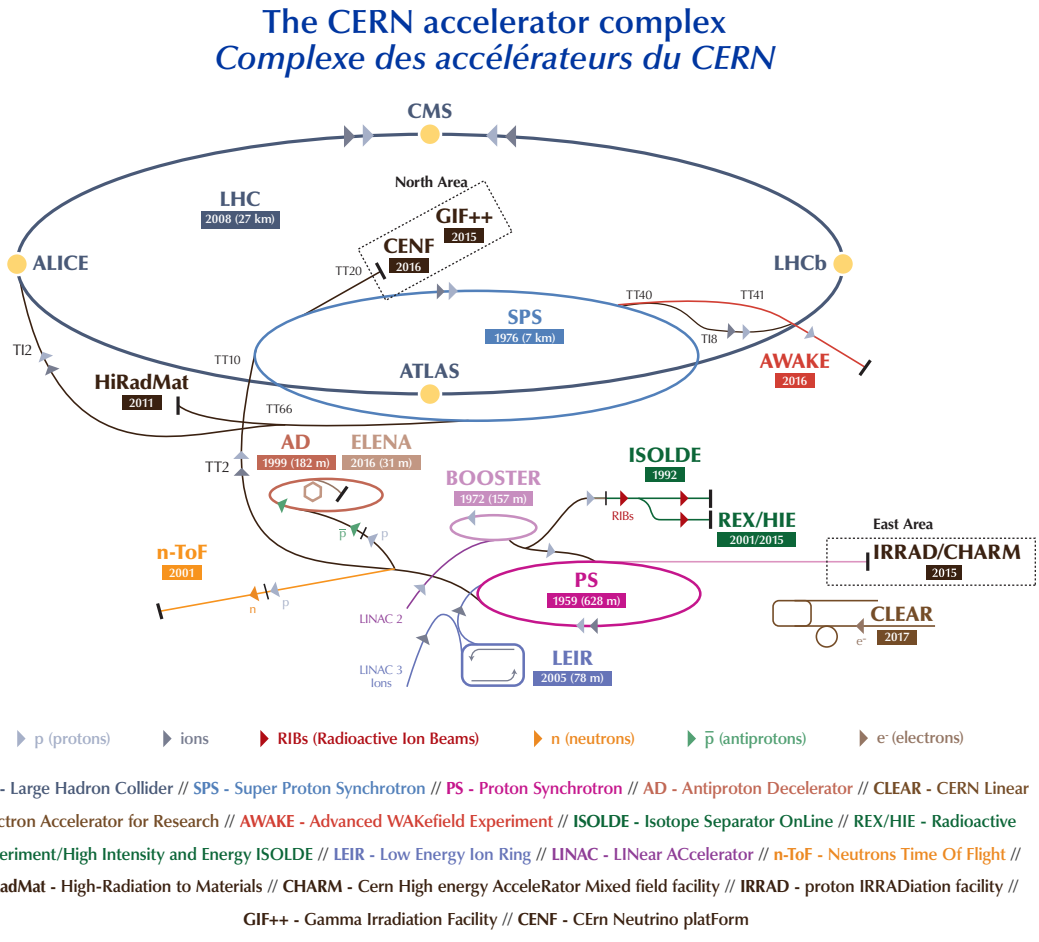


Figure 2.1: Diagram of the LHC accelerator and detector complex, from Ref. [4].

The protons are obtained by ionising H_2 from a bottle of hydrogen gas, by applying an electric field. A linear accelerator (Linac 2, in Run 2) then takes these

[1] Ions, like lead, are also collided, but will not be covered here.

protons and accelerates them to 50 MeV. This beam is then injected into the Proton Synchrotron Booster (PSB), which consists of four vertically stacked synchrotron rings. The PSB accelerates the protons to 1.4 GeV and also groups the protons in the beam into ≈ 2400 bunches of around 10^{11} protons each. Then the beam is injected into the Proton Synchrotron (PS), which increases the beam energy to 25 GeV. The penultimate acceleration is performed by the Super Proton Synchrotron (SPS), which accelerates the protons to 450 GeV. Finally, the beams are injected into the LHC and once the LHC is at capacity, the fill is complete. After the beams enters the LHC, they are accelerated until an energy of 6.5 TeV is reached, with a spacing of 25 ns between the bunches. The protons are bent in a nearly circular orbit by 1232 superconducting dipole magnets, which are cooled to 1.9 K and made out of niobium-titanium. The beam is focused by 858 quadrupole magnets that, depending on the positioning of the poles, either focus the width or height of the beam.

2.1.1 Proton–proton Collisions, Luminosity and Pileup

The performance of a collider can be measured through the instantaneous luminosity ($\mathcal{L}_{\text{inst}}$), which is defined as the ratio between the cross-section of a process and the expected rate at which events from that process would be expected to be generated in LHC collisions:

$$\frac{dN}{dt} = \mathcal{L}_{\text{inst}} \times \sigma, \tag{2.1}$$

where $\frac{dN}{dt}$ is the event rate, and σ is the target cross-section. Once the parameters characterising the beam are known the instantaneous luminosity can be calculated:

$$\mathcal{L}_{\text{inst}} = \frac{N_1 N_2 N_b R}{4\pi\sigma_x\sigma_y} F, \tag{2.2}$$

where, N_1, N_2 are the number of protons in the two beams, N_b is the number of bunches, R is the revolution frequency of the beams around the circular path of the accelerator. Additionally, σ_x σ_y are the widths of the beam density distributions in the x and y direction, respectively. The additional factor 'F' accounts for the crossing angle of the colliding beams. As the beams are brought together into the

same beam pipe before the IP, it is necessary to collide the beams at an angle, to avoid large beam induced energy losses. This crossing angle is reduced throughout a fill to maintain luminosity as the beam decays – a processes called levelling. The colliding beams produce interactions in ATLAS at a rate of one every 25 ns, giving a collision frequency of 40 MHz.

Luminosity in ATLAS is nominally measured by the LUCID–2 detector [5], that detects Cherenkov light produced in the quartz windows of photomultipliers located close to the beam. Other methods are also used to cross-check the luminosity, using ATLAS sub-detectors. LUCID is able to measure beam parameters in Equation 2.2 using special runs called Van der Meer Scans [6, 7]. By fixing one beam and moving the other in discrete increments in x and then y, σ_x and σ_y can be very precisely measured from the changing interaction rates.

In Run 2, the maximum $\mathcal{L}_{\text{inst}}$ was 2.1×10^{34} cm^{-2}s^{-1} [8]. The integrated luminosity, $\int \mathcal{L}_{\text{inst}} dt$, is the instantaneous luminosity integrated over time. Figure 2.2 shows the integrated luminosity as a function of time for ATLAS in Run 2. The physics luminosity is the final luminosity available for analyses after events where there are issues with the detector are rejected, and corresponds to the 'good run list'.

The beams contain bunches of $\approx 10^{11}$ protons, and therefore it is possible to have multiple collisions per event. This is referred to as in-time pileup, while out-of-time pile-up refers to background collisions from a previous bunch crossing that still affect the detector during the collision of interest. Pileup can be parameterised as the mean number of interactions per bunch crossing:

$$\langle \mu \rangle = \frac{L_b \sigma_{\text{inel}}}{f}, \tag{2.3}$$

where $L_b = L/n_b$ is the bunch luminosity, and σ_{inel} is the inelastic pp collilsion cross-section, and 'f' is the collision frequency (40 MHz). Pileup can have a significant impact on the detector trigger efficiency and precision reconstructed objects due to increased background hadronic activity.

In Run 1 of the LHC (2009 – 2013), the collision energy was nominally 7 and then 8 TeV. A total of 23.3 fb^{-1} of pp data was collected at 8 TeV with an average

pileup of 20. Generally, in current ATLAS searches, this dataset is not used due to the much lower energy and comparatively small luminosity when compared to the Run 2 dataset. In Run 2, the energy was increased to 13 TeV, and a total of 140 fb^{-1} of pp data was collected, with an average pileup of 33.7. Figure 2.3 shows the mean number of interactions per bunch crossing as a function of the recorded luminosity for Run 2. In the ongoing Run 3, the energy was again increased, to the maximum 13.6 TeV, at time of writing 183 fb^{-1} of pp data has been recorded, with an average pileup of $\approx$ 60.

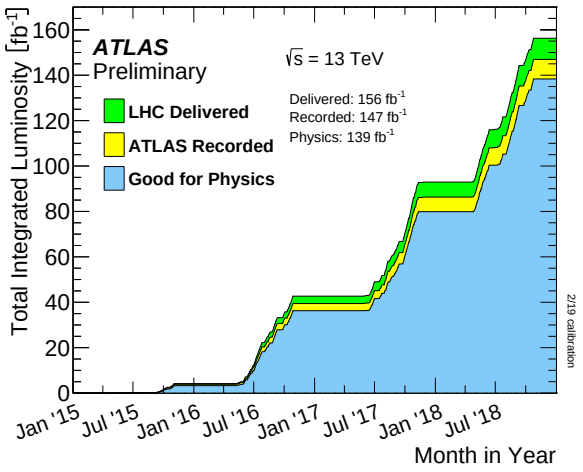


Figure 2.2: The integrated, delivered and recorded luminosity as a function of time for the ATLAS experiment in Run 2, taken from Ref. [8]

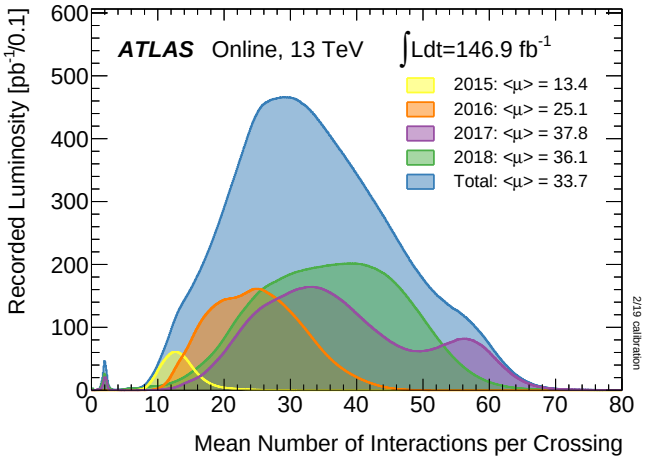


Figure 2.3: The mean number of interactions per bunch crossing as a function of the recorded luminosity, for each year of Run 2 and the entirety of Run 2, taken from Ref. [8].

2.2 The ATLAS Detector

ATLAS is a general purpose particle detector, one of two at the LHC along with CMS. A sketch of the ATLAS detector is shown in Figure 2.4. The detector consists of three major subsystems, the inner detector (ID), the calorimeters and the muon spectrometers (MS), which in turn consist of multiple subdetectors.

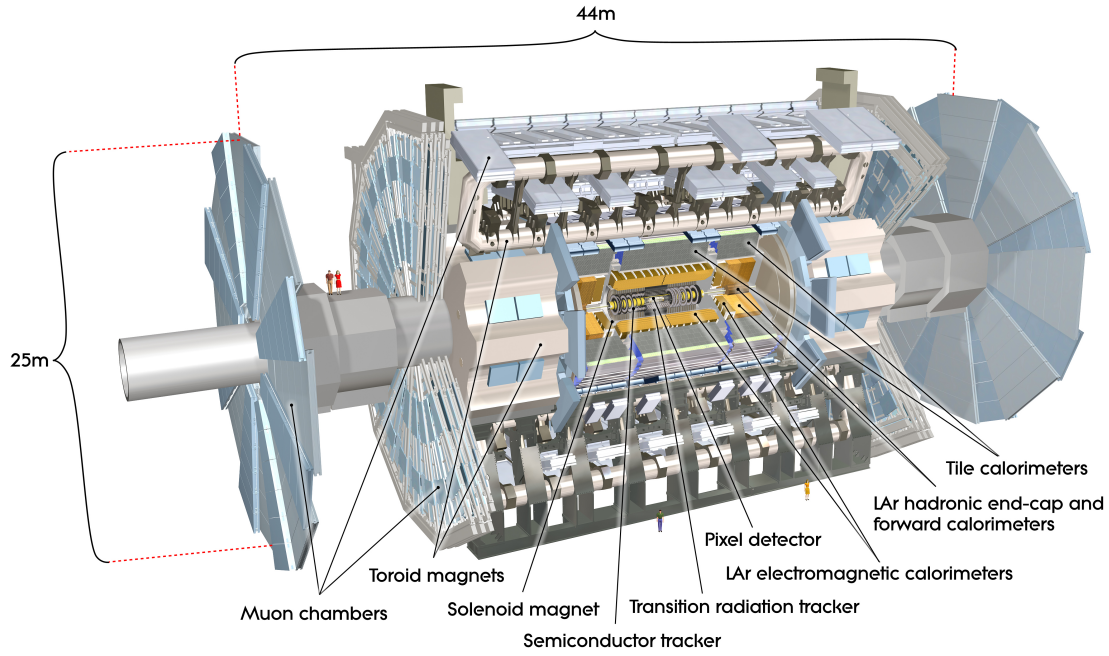


Figure 2.4: A rendering of the ATLAS detector in Run 2, taken from Ref. [9].

2.2.1 The ATLAS Coordinate System

ATLAS uses a right-handed coordinate system with its origin at the IP in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. A graphical representation of this coordinate system is given in Figure 2.5a

The angles (θ, ϕ) are used to define the ATLAS coordinate system. It is the convention to use pseudorapidity, η, instead of θ to measure the angle of a particle pointing away from the beam axis:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{2.4}$$

Pseudorapidity is utilised as an approximation of the rapidity [2], and therefore is assumed to be independent of the frame of reference taken.

Figure 2.5b illustrates the relationship between the angular coordinates, while Figure 2.5a shows how these relate to the LHC's Cartesian coordinate system. It can be noticed that there is a pole in the value of η as a particle tends toward the beam pipe ($z = 0$). Particles produced at large values of $|\eta|$ (> 3.2) are called forward objects, and their momenta is typically less well measured, due to energy losses from particles which are outside of the detector acceptance, and escape undetected down the beam pipe. An additional spatial quantity of interest is ΔR, which is the angular distance between two points, calculated as:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{2.5}$$

In addition, impact parameters, shown in Figure 2.6, are defined and are commonly used in object reconstruction. The transverse impact parameter, d_0, is defined as the distance of closest approach of the track to the primary vertex (PV) in the transverse plane:

$$d_0 = \sqrt{(x_{\text{track}} - x_{\text{PV}})^2 + (y_{\text{track}} - y_{\text{PV}})^2}. \tag{2.6}$$

The longitudinal impact parameter, z_0, is defined as the distance along the beam axis between the point of closest approach and the PV:

$$z_0 = z_{\text{track}} - z_{\text{PV}}. \tag{2.7}$$

Convention is that the region of the detector in the positive z-direction is denoted as the 'A side' and the negative direction the 'C side'. The detector is constructed to have near full 4π coverage, and hence exploit momentum conservation in the transverse plane. To achieve this, ATLAS roughly forms a cylinder around the LHC beampipe. The central area is called the barrel, and each end is called an endcap.

[2] Defined as: $\frac{1}{2}\ln\frac{E+p_z}{E-p_z}$, where E is the energy of the particle and p_z the z component of its momentum.

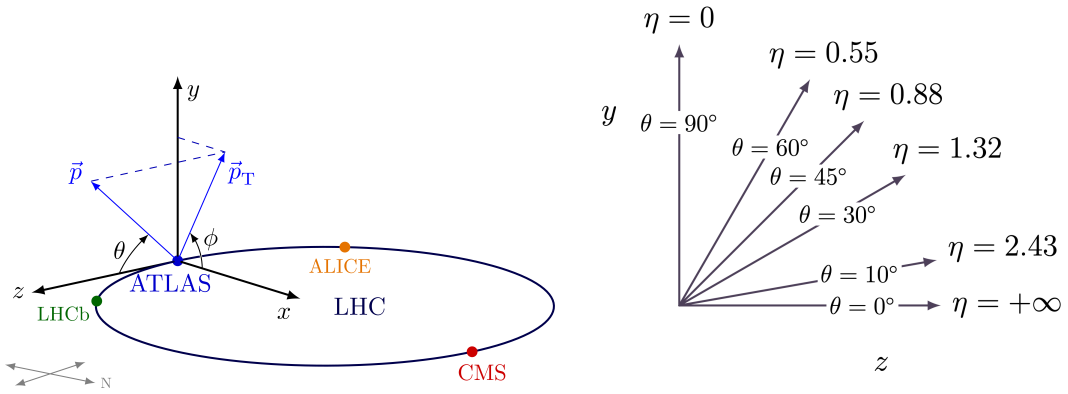


Figure 2.5: In (a), the relationship between the LHC collisions cartesian coordinates and the ATLAS (θ, ϕ) coordinates is shown. Also shown is the momentum vector $\vec{p}$ and the transverse momentum vector $\vec{p}_T$. In (b), the relationship between η and θ in the y–z plane is shown. Both figures are taken from Ref. [10].

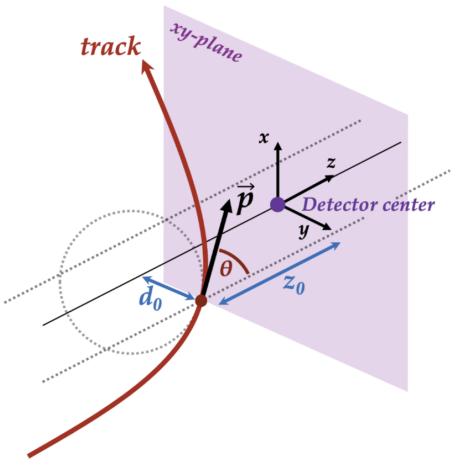


Figure 2.6: Transverse (d_0) and longitudinal (z_0) impact parameters of a track with respect to the interaction point. Figure taken from Ref. [11].

2.2.2 The Magnets

ATLAS has two separate types of magnets. A 2T superconducting solenoid magnet applies a uniform magnetic field in the z-direction, such that charged particles are deflected in ϕ. This enables measurements of charged particle momenta through the curvature of tracks they leave behind in the ID. In addition, toroidal magnets, at the barrel and end cap, apply between a field of 0.5T–2.0T to deflect muons in the η direction and obtain an accurate measurement of their momenta [12].

2.2.3 The Inner Detector

The ATLAS inner detector consists of three sub-detectors, shown in Figure 2.7. The inner detector covers the full ϕ range and η values up to $|\eta| = 2.5$. Two of these detectors use silicon chip technology to measure energy deposition (hits), from charged particles produced in collisions, using these to reconstruct tracks and interaction vertices. Silicon tracking detectors are based on semiconductor technology that exploit the properties of reverse-biased p–n junctions. Applying a reverse-bias voltage widens the p–n junction's depletion region, as electrons and holes are pulled away from the interface in opposite directions. When a charged particle traverses this region, it ionises the silicon, generating electron-hole pairs. The applied electric field drives these charge carriers towards the electrodes, producing a measurable current that is read out by the detector electronics.

The Pixel Detector

The first sub-detector in the ID is the pixel detector, which consists of three barrel layers and three end-cap layers of silicon pixel modules. The silicon pixel modules mostly have dimensions $50\mu\text{m} \times 400\mu\text{m}$. In addition to the three layers in the barrel, there is an addition layer of sensors, the Insertable B-Layer (IBL), an ultra-thin silicon detector that was added to ATLAS at the start of Run 2 [13, 14]. The IBL is located as close to the beam-pipe as possible to improve tracking performance, particularly for reconstruction of secondary collisions.

The Semiconductor Tracker

The next sub-detector, after the pixel detector, is the semiconductor tracker (SCT). In the barrel, each SCT module consists of two 64 mm $\times$ 128 mm silicon strip sensors, mounted back-to-back with a 40 mrad stereo angle between them, to form a single module. The stereo angle allows for two-dimensional track reconstruction, which is a key feature of the SCT. The modules have a pitch of 80 μm, and are arranged into four concentric barrel layers, providing high-resolution tracking capabilities with increasing radial distances from the beam axis.

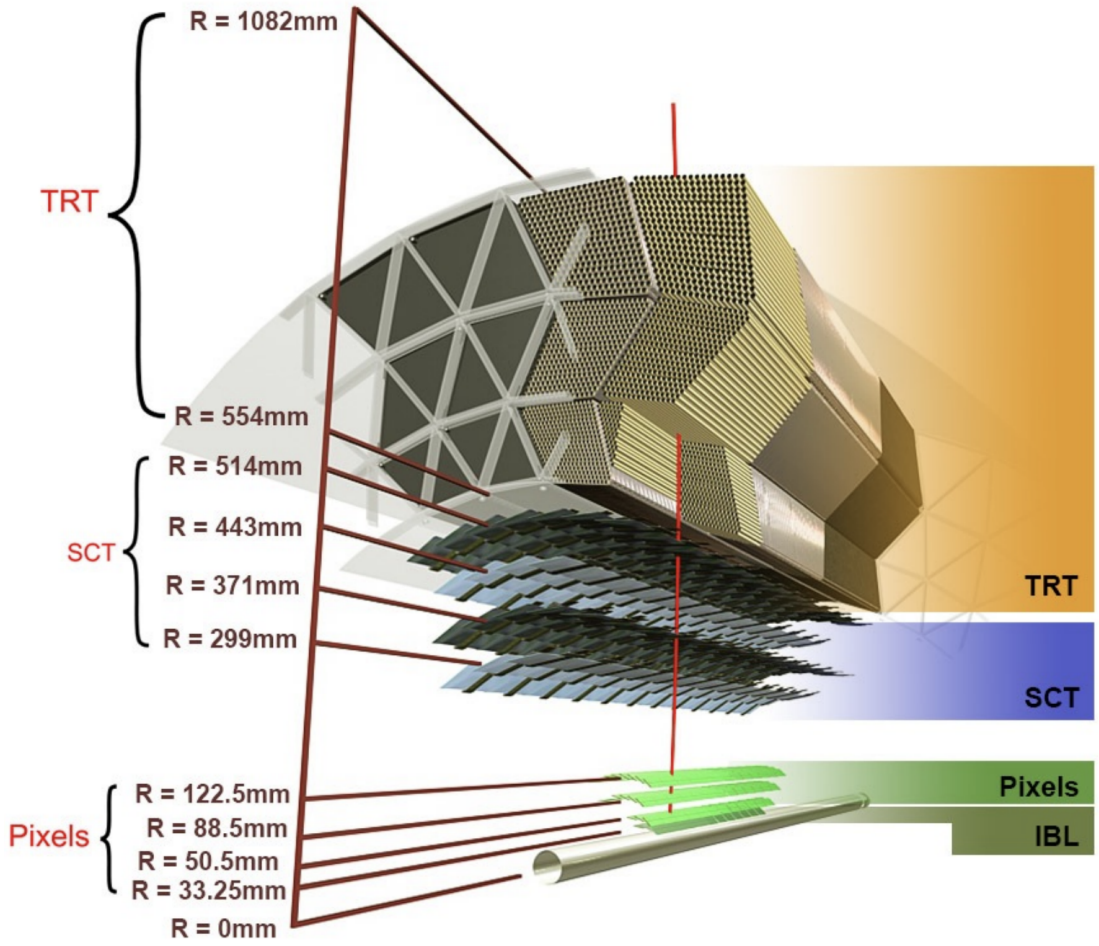


Figure 2.7: Rendering of the ATLAS inner detector showing all its sub-detectors, taken from Ref. [13]. The distances to the beam pipe are also shown.

The Transition Radiation Tracker

The furthest ID sub-detector from the beam-pipe is the transition radiation tracker (TRT), which only covers up to $|\eta| = 2$, and uses different technology to the silicon detectors discussed above. The TRT is made up of 300,000 drift tubes (straws), consisting of a gold-plated anode wire and filled with a Xenon gas mixture. [3] When a charged particle passes through the tube, it ionises the gas, and the resulting electrons drift towards the central wire, where they initiate an avalanche in the electric field, producing a measurable signal that is then read out using optical link cables.

[3]The TRT has numerous leaky straws, and so where there are leaks Argon gas is used instead as it is cheaper.

The TRT is able to not only measure the momentum for charged particles, but also identify them. In the work in this thesis, the TRT is primarily used to distinguish charged pions from electrons with very high efficiency, due to the sizeable mass difference. The space between the straws is filled with polyethylene, and when a relativistic particle traverses the boundary between the polyethylene and straws, it can emit transition radiation, which depends on the Lorentz boost of the particle $\gamma \approx E/m$. As a result, the resulting transition radiation can be used to determine the particle mass given two particles of the same energy.

Resolution

Each sub-detector has different resolutions, shown in Table 2.1. When hits are observed in multiple sub-detectors, a track can be reconstructed with resolution:

$$\frac{\sigma_{p_T}}{p_T} = 0.05\%\, p_T\,[\text{GeV}] \oplus 1\%, \tag{2.8}$$

where $\oplus$ indicates the sum in quadrature. The relative momentum resolution for high p_{T} tracks in Run 2 is: $\frac{\sigma_p}{p} = (4.83 \pm 0.16) \times 10^{-4}\ \text{GeV}^{-1} \times p_{\text{T}}$ [15].

Table 2.1: Spatial resolution of the ATLAS ID sub-detectors, resolution obtained from Refs. [14, 16].

Sub-detector	**Resolution in $R-\phi$ [μm]**	**Resolution in z [μm]**
IBL	10	60
Pixel	10	115
SCT	17	–
TRT	130	–

2.2.4 The Calorimeters

ATLAS has two calorimeters, covering the full ϕ range and $|\eta| < 4.9$. These are the electromagnetic calorimeter (ECal) [17], and the hadronic calorimeter (HCal) [18], which are both shown in Figure 2.8. These are both sampling calorimeters, so consist of multiple layers of active and passive material. Typically, the passive material is dense, like lead, in order to cause particle cascades. Active layers

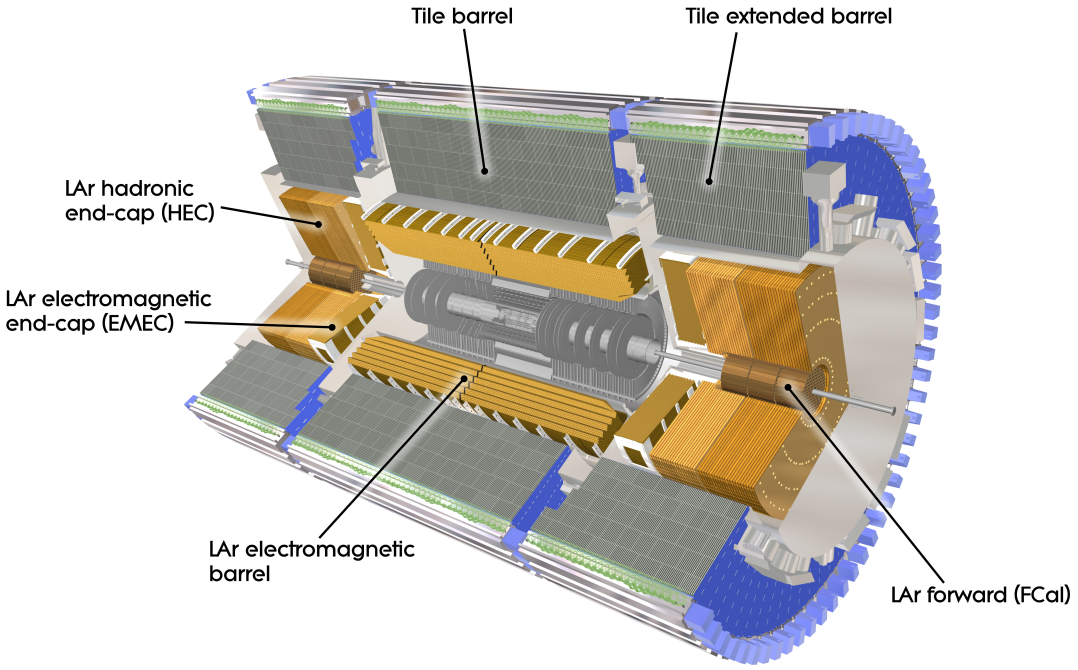


Figure 2.8: Rendering of the ATLAS calorimeters, taken from Ref. [19].

measure the energy deposited by these showers, by measuring either the produced ionisation or scintillation light. The calorimeters also aim to stop all hadronic radiation before it reaches the muon spectrometer, to ensure clean muon signals. The calorimeters are read out in segments, called cells.

The resolution of the calorimeter energy measurements is given by:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{2.9}$$

where 'a' is the stochastic term, which describes statistical fluctuations due to Poisson statistics, for example in the sampling process, 'b' is the noise term, which accounts for electronic noise in the detector readout. The constant term 'c', describes resolution effects that do not depend on E_T, like energy deposition outside the active material and the non-uniform calorimeter response.

The Electromagnetic Calorimeter

The ECal covers $|\eta| < 1.475$ in the barrel and $1.375 < |\eta| < 3.2$ for the end caps, and is used to measure the energy of electrons, photons and partially the jet

energy. The ECal uses lead as the passive material, and liquid argon as the active material. Electromagnetic showers caused by the passive material are parametrised by the radiation length X_0, which is the average distance over which the energy of a high-energy electron will be reduced to $1/e$ of its original value. The energy of the produced showers is measured by recording the current induced via the ionisation of the liquid argon.

The Hadronic Calorimeter

After the ECal, comes the HCal which covers $|\eta| < 1.0$ in the barrel, $0.6 < |\eta| < 1.7$ in the extended barrel, the end caps covering $1.5 < |\eta| < 3.2$ [20]. In the barrel region, steel is used as the passive material and plastic scintillator tiles are used as the active material. The energy of the showers produced by the steel layers is measured by recording the scintillation light produced by the tiles, using photomultipliers. For the end-cap and forward HCals, the active material used is liquid argon. In the end cap, copper is used as the passive material.

The Forward Calorimeters

The forward calorimeters cover $3.1 < |\eta| < 4.9$ [21] and aims to provide accurate measurements of particle energies with high pseudorapidities. The forward calorimeters carry out both hadronic and EM measurements of particle energies. For the forward hadronic measurements, the active material used is liquid argon and for the forward EM measurements, both copper and tungsten are used.

2.2.5 The Muon Spectrometer

The outermost sub-detector of ATLAS is the muon spectrometer (MS), which targets the weakly interacting muons that pass through the ID and calorimeters largely unimpeded [22]. The MS covers the range $|\eta| < 2.7$ and lies in the magnetic field generated by the toroidal magnets, which bend the muon tracks and allow the momentum to be measured from the curvature. There are four different components to the MS, shown in Figure 2.9, each of which utilises slightly different techniques and serves a different role. The sub-detectors used for triggering muon events is

done by the resistive plate chambers (RPC) in the barrel region and the thin gap chambers (TGC) in the end-caps. For precision measurements of muon kinematics, the monitored drift tubes (MDT) and cathode strip chambers (CSC) are used. The precision of the measured muon momentum is around 3% for low p_{T} muons ($2 < p_{\mathrm{T}} < 250$ GeV) and 10% for high p_{T} muons.

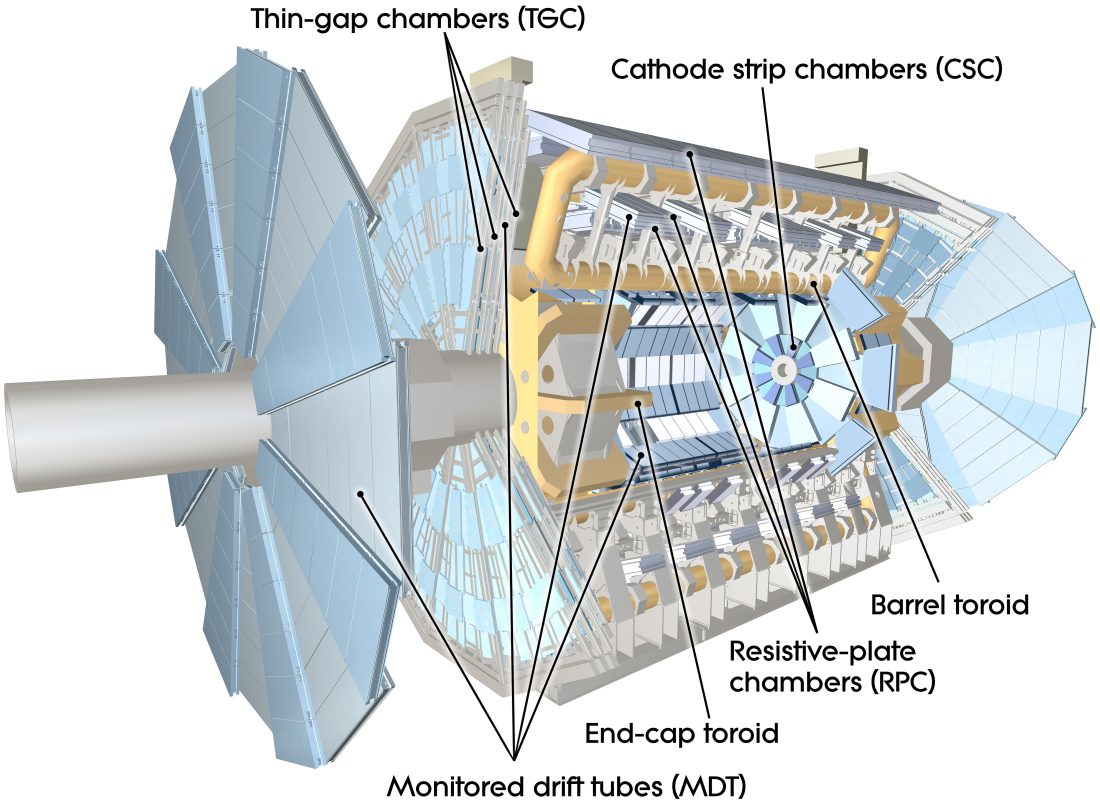


Figure 2.9: Rendering of the ATLAS muon spectrometer, taken from Ref. [23].

Monitored Drift Tubes

The MDTs contain an Argon/CO_2 gas mixture surrounding a central wire. When a high energy muon passes through this gas mixture it causes the gas to ionise, leading to a signal current. This technology has a long response time, as the ionised electrons drift relatively slowly. As a result, it is necessary to halt the detector readout (impose dead time) after the first charge is produced, and different, faster, detector technologies are needed for triggering.

Cathode Strip Chambers

The precision measurements from the MDTs are supported by the CSCs in the range $2.0 < |\eta| < 2.7$. Each chamber is filled with an Argon/CO_2 gas mixture and consists of four planes of wires positioned perpendicular to each other, which act as anodes. The corresponding cathodes are split into strips, allowing position information to be measured. By measuring the relative induced charges on adjacent strips, a signal is extracted.

Resistive Plate Chambers

The RPCs consist of two oppositely charged plates with a spacing of 2mm, filled with a predominantly $C_2H_2F_4$ gas mixture [4] When a muon which passes through the chamber, it causes avalanches of electrons which get drawn to the anode, giving a signal. There are four layers of RPCs, with two sets of chambers in each layer.

Thin Gap Chambers

The TGCs are used to trigger in the end-cap regions and to provide additional information to complement the MDTs. They use similar multiwire proportional chambers to the CSCs, but with a shorter distance between the wires and the cathode strips. The shorter distance between the strips and the wires gives a shorter read-out time than the CSCs, making them ideal for triggering events.

2.3 Trigger System

The LHC collides protons at a frequency of 40 MHz, which, if every event were recorded, would correspond to over 50 TB/s of data – far exceeding the capacity of the ATLAS read-out and storage systems. To reduce this output rate, ATLAS utilises triggers to select events of interest. In Run 2, a two-stage trigger system was used: the hardware-based Level-1 (L1) Trigger and software-based High-Level Trigger (HLT) [25]. The ATLAS L1 trigger rate reduces the 40 MHz collision rate

[4]This gas has a very high Global Warming Potential, and leaks from the ATLAS, CMS and ALICE RPC sub-detectors were responsible for 87% of the greenhouse gas emission by LHC detectors in Run 2. [24]

of the LHC, down to 100kHz, where this value is limited by the maximum read-out rate for the components of the sub-detectors. Then, the HLT trigger runs on events which pass the L1 trigger, and further reduces the output rate to approximately 1 kHz by carrying out more complex selections. Dedicated triggers have been developed for the main signatures: electrons, photons, muons, taus, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, as well as more specialised triggers designed to select events containing b-quarks or potential new particles with long lifetimes. This section will focus on the physical implementation of the trigger, and operation of the HLT. A more in-depth look at trigger software and individual ATLAS HLT triggers will be given in Chapter 6.

2.3.1 Level 1 Trigger

The L1 trigger is split into two main systems: the calorimeter trigger (L1Calo) and the L1 muon trigger. The trigger decisions from these two systems, as well as from several other systems, such as LUCID, are combined in the Central Trigger Processor (CTP) which then makes the final L1 trigger decision. The CTP applies dead-time to prevent two L1 accepts from occurring too close together, and preventing overlap in the detector readouts. It is also responsible for applying complex dead-time, which restricts the number of L1 accepts allowed in a given period, to avoid too many events being passed to the HLT. Both the L1Calo and L1 muon triggers are used to select events to pass onto the HLT, along with Regions-of-Interest (ROIs) in the detector.

Level 1 Calorimeter Trigger

The L1Calo system receives information from the entire calorimeter, which is divided into 'towers'. These towers are a combination of the individual cells that make up the calorimeter cells, considering the EM and hadronic calorimeters separately. The signals from the towers are processed by an FPGA-based preprocessor system, which digitises and calibrates them to obtain the E_T values used in the trigger decision. A dynamic, bunch-by-bunch pedestal subtraction is applied to correct for variations in conditions throughout each bunch train. Detector noise filters are also applied which mean that a tower is only considered if its E_T falls above a threshold. These

requirements significantly reduce the impact of pile-up on the trigger rates. Events of interest and ROIs which are passed onto the HLT are identified corresponding to potential physics objects of interest like electrons or jets.

Level 1 Muon Trigger

In the L1 muon trigger, the RPCs are used to trigger candidates with $|\eta| < 1.05$. For the range $1.05 < |\eta| < 2.4$ the trigger is provided by the TGCs. Different requirements are imposed on hits in the muon detectors to define ROIs, and to give acceptance across the full spectrum of muon p_{T}.

2.3.2 High Level Trigger

The HLT trigger uses the ROIs that pass the L1 trigger and uses information from the entire detector to carry out more complex event reconstruction. The HLT 'farm' consists of approximately 40,000 CPUs and GPUs, and makes the final decision as to whether to save an event for offline analysis in less than 230μs. Events are saved for analysis at an approximate rate of about 1.1 kHz. The HLT uses software-based triggers to run algorithms that can (if needed) require the full event information to make a decision.

Each physics signature of interest has its own set of trigger 'chains', which define the required thresholds reconstructed physics objects need to meet to save the event. Further information on the ATLAS missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) HLT triggers is given in Chapter 6. The full list of trigger chains, called the trigger 'menu', can vary from run to run and depend on the beam conditions and physics goals. In addition, to ensure the HLT rate remains acceptable, triggers are often 'prescaled'. For a prescaled trigger, only every n-th event that satisfies the trigger requirements, is saved for offline analysis.

La vita non è come l'hai vista al cinema. La vita... è più difficile.

(Life isn't like in the movies. Life... is much harder.)

Nuovo Cinema Paradiso

3

Event Simulation, Reconstruction and Object Identification

Contents

3.1 Monte-Carlo Generation 24
3.1.1 Event Generation 24
3.1.2 Detector Simulation 25
3.2 Low Level Objects 26
3.2.1 Tracks 26
3.2.2 Primary Vertex 28
3.2.3 Calorimeter Clusters 28
3.2.4 Particle Flow Objects 29
3.3 Object Reconstruction 30
3.3.1 Jets 31
3.3.2 Jet Flavour Tagging 33
3.3.3 Electrons 33
3.3.4 Muons 36
3.3.5 Lepton Isolation 38
3.3.6 Overlap Removal 40
3.3.7 Missing Transverse Energy 40

This chapter focuses on the reproducing the *pp* collisions at the LHC, using Monte Carlo (MC) simulation to provide an accurate representation of the data collected by ATLAS. In addition, this chapter will cover how ATLAS takes detector level information in data and MC simulation, and turns these signals into physics objects.

3.1 Monte-Carlo Generation

Monte Carlo simulations are used to estimate event yields and evaluate systematic uncertainties for both signal processes and Standard Model (SM) backgrounds. In proton–proton collisions, this task is particularly challenging due to the non-perturbative nature of quantum chromodynamics (QCD) at low energy scales, where the strong coupling (α_s) becomes large, and perturbation theory is no longer reliable. After the event generation, the output must be passed through a detector simulation, which simulates the detector response. This step is essential to recover accurate simulation of the data observed in ATLAS, and to account for various detector effects and inefficiencies. MC simulations are also used to derive theoretical systematic uncertainties, by generating alternative MC samples. This is achieved by varying the choices made in generation of the nominal MC defined in the following sections, and evaluating the effect on the predicted yields in analysis regions. The simulation of proton–proton collisions is a complex task, and this section is only a brief introduction to the topic, heavily based on Ref. [26]. Additional information is given in Appendix A, to provide context for the theoretical uncertainties implemented in the analysis in Chapter 5.

3.1.1 Event Generation

The proton is a composite subatomic particle, consisting of three valence quarks: two up quarks and one down quark. However, this description is incomplete. Over 98% of the proton's rest mass originates, not from the valence quarks, but from complex QCD effects, involving a 'sea' of virtual quark-antiquark pairs and gluons [27]. In high-energy proton-proton collisions, it is these constituents, collectively referred to as partons, that initiate the hard scattering processes of interest. Due to the huge number of possible initial conditions and resulting final states, it is computationally impossible to simulate all possible interactions. Instead, targeted final states are modelled using MC simulations, which combine theoretical predictions with statistical sampling to provide approximations of the processes.

3.1.2 Detector Simulation

After the events have been generated, they need to be propagated through a simulation of the ATLAS detector to accurately reproduce observed detector signatures and data distributions. To simulate the detector response, the MC samples are processed through the ATLAS simulation framework [28] in GEANT4 [29]. The interaction of particles with the active material of the detector are saved as hits (size of energy deposition) in the relevant sub-detector(s). When the hits reach a threshold, they are converted into the expected electrical signal, and are digitized into the output format of observed collision data. In the digitisation step, run-dependent conditions related to the varying responses of channels, detector readout noise levels, and inactive sub-detector modules is taken into account and the hardware trigger decision is recorded.

Full detector simulation is very computationally expensive and generally reserved for MC background processes used by the entire collaboration. Often signal samples for specific analyses use the ATLAS fast simulation (specifically AF2), which parameterises the response of the calorimeters, to speed up processing time [30]. The simulation of the calorimeter showers in the detector are by far the largest consumer of CPU resources in the MC simulation chain, and hence this AF2 simulation is widely employed to reduce CPU time.

Pileup

In ATLAS, the effect of multiple interactions in the same bunch crossing (pileup) in MC simulation, is modelled by overlaying the simulated hard-scatter event with inelastic proton–proton events generated with PYTHIA 8.186. All MC events are then re-weighted to match the pile-up distribution observed in the data.

Truth-matching

After event generation, the truth origin of final state particles can be tracked. However, the detector simulation introduces uncertainties in the measurement of particle kinematics. Consequently, soft hadronic activity and leptons originating

from heavy flavour decays may be misidentified as 'fake' hard scatter lepton signatures. Often the fake lepton background is estimated through data-driven methods – see Section 5.4.2. To ensure that this fake lepton background is estimated entirely with a data-driven method (without any double counting), prompt leptons are required to satisfy a truth-matching requirement. For reconstructed leptons, the reconstructed track is associated with the track hits in the GEANT4 truth record, caused by the true object from the MC simulation. Then this information can be used to assign truth information to the reconstructed lepton, including the type of truth object and its origin. This information is then used to classify each reconstructed lepton.

3.2 Low Level Objects

Detector level information is reconstructed into what are called 'low level' objects. Inner detector hits and muon detector hits are reconstructed into tracks, based on an identified primary collision vertex. Calorimeter cells meeting thresholds are combined to form calorimeter clusters. A set of basic quality criteria is also applied to ensure a fully operating detector and to reject events with large detector noise or non-collision backgrounds [31].

3.2.1 Tracks

Trajectories of charged particles, called tracks, are reconstructed mainly using from hits the ATLAS ID, and where relevant, the muon detectors. These provide precision momentum measurements and charge information due to the bending from the 2T axial magnetic field. Track reconstruction utilises the inside-out method [32], starting from groups of three hits in the Pixel detector or SCT. Additional hits are then added to the track by using a Kalman filter [33] to identify the most likely additional hit based on the preliminary trajectory. Tracks can also be extended to included TRT hits. Through this process, it is possible for hits to be assigned to multiple tracks. To aid the reconstruction of tracks with shared hits, if a track has more than two shared hits, an artificial neural network (NN) is used to identify

shared hits [34]. The measured charge, which is proportional to the deposited energy, and relative position of pixels in the hit can be used to identify merged hits.

Tracks are ranked based on quality, with each one assigned a score and high p_{T} tracks being preferred. A χ^2-fit is performed to assess the quality of the track. This information, along with the number of hits, number of holes (layers in the ID where there are gaps in the hits of the predicted track), number of shared tracks and track p_{T}, is used to calculate the score. Then potential tracks are ranked, with tracks below a threshold rejected. Tracks can also be rejected if they fail any of the following criteria:

- A minimum of seven hits in the pixel detector and SCT, in order to reduce track uncertainties.
- A maximum of one hole.
- A maximum of two shared hits.
- Kinematic requirements: $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$.

For all remaining track candidates fulfilling these requirements, a high-resolution fit is performed using all available information. For these high-resolution fits, the position and uncertainty of each cluster is determined using additional NNs. They predict the positions where charged particles intersect the pixel sensors, and are based on the same architecture as the NN used to resolve shared hits.

Muon Tracks

In the muon spectrometer, hits in the muon chambers are collected to form segments in each layer of the detector. These segments are then used as the starting point to construct candidate tracks, combining multiple compatible hits. Similar to the inner detector tracks, multiple track candidates can be formed, which use the same hits, and so a similar overlap procedure must be performed. In addition, muon spectrometer tracks can be combined with ID tracks to give 'combined' muon candidates.

3.2.2 Primary Vertex

Every event is required to have at least one vertex built from at least two associated tracks with $p_{\mathrm{T}} > 0.5$ GeV. If there are multiple such vertices, the vertex with the highest sum of squared transverse momenta ($\sum p_{\mathrm{T}}^2$) of associated tracks [35] is selected as the primary vertex.

3.2.3 Calorimeter Clusters

Clusters are formed from measurements of the individual calorimeter cell energies [36]. Two types of cluster are used in ATLAS: 'egamma' clusters formed using a sliding window algorithm, and 'topoclusters' formed using a topological clustering algorithm [37]. The egamma clusters only use information from the EM calorimeters, whereas the topoclusters use data from both EM and hadronic calorimeters.

For the egamma clusters, the sliding window algorithm is used to cluster EM calorimeter cells. First, the EM calorimeter is divided into a grid in $(\eta \times \phi)$, and the energy of all longitudinal cells are summed to give a combined 'tower' energy. For cells that are part of more than one tower, the transverse energy (E_T) is partitioned according to the factional area of the cells in each tower. The sliding window algorithm considers fixed windows in $(\Delta\eta \times \Delta\phi)$. All possible windows that contain a total E_T above a threshold are considered as seeds, if two seeds overlap then the one with the higher E_T is chosen. The creation of the egamma clusters then proceeds from these seeds.

For the topoclusters, an algorithm groups together nearby cells into clusters and proceeds in three main steps. The energy requirement on the cells in each step is expressed in terms of a ratio to the cell's expected noise value, $\gamma = \frac{E_{\mathrm{cell}}}{\sigma_{\mathrm{noise}}}$. First, all cells with $|E_{\mathrm{cell}}| > 4\gamma$ are used as seeds and labelled as 'proto-clusters'. Next, cells with $|E_{\mathrm{cell}}| > 2\gamma$ that are adjacent to a proto-cluster are added to that proto-cluster. If a given cell is in-between two proto-clusters then the proto-clusters are merged. This continues until there are no more cells that meet the energy requirement, adjacent to any proto-clusters. Finally, one layer of cells surrounding proto-clusters, with $|E_{\mathrm{cell}}| > 0$, are added to the proto-clusters, and the clusters

are formed. To avoid loss of sensitivity where particles are close to each-other, clusters can be split if they contain multiple maxima.

For topoclusters there is an (optional), calibration called the local cluster weighting (LCW) calibration. This calibration corrects for the energy loss from neutrons, which are invisible in the ECAL as they are neutral particles. Clusters which use this calibration are described as being measured at the 'hadronic scale', and ones which do not use it are measured at the 'EM scale'.

3.2.4 Particle Flow Objects

Track and calorimeter cluster information can be combined to define particle flow objects (PFOs). The PFlow algorithm [38] is widely used in ATLAS, due to the significant improvement energy resolution it provides. First a set of object requirements are applied, $p_{\mathrm{T}} > 500$ MeV, $|\eta| < 2.5$, $|d_0| < 1.5\ mm$ and $|z_0 \sin\theta| < 1.5\ mm$. Then the PFO alorithm matches ID tracks and topoclusters to form PFOs as follows:

- ID tracks which fulfil a set of quality requirements are p_{T} ordered.
- Starting with the highest p_{T} track, this is matched to a topo-cluster forming a 'topo-track' system.
- The probability that the particle the track originated from, deposited energy in more than one topo-cluster is calculated. If the probability passes a certain threshold, additional topoclusters are added to the system.
- Once the topo-track system is complete, the track energy is removed from the topocluster, giving a PFO.
- This is repeated for the next highest p_{T} track and so on, until all tracks and topoclusters have been evaluated.

At low p_{T}, ID tracks provide precision measurements for low-energy charged particles, and can provide a measurement the energy for particles which do not exceed the calorimeter cell energy threshold. Alternatively, the calorimeter has better energy resolution for high-energy particles. By combining the tracking and

calorimeter information the energy and angular resolution of the objects can be improved, as the ID and calorimeter provide complementary information. Only charged particles leave measurable tracks in the ID. As a result, only the calorimeter is used to provide measurements for the PFOs of neutral particles. Additionally, due to the lack of ID coverage in the forward region, only the calorimeter is used for objects with $|\eta| > 2.5$.

3.3 Object Reconstruction

The low level objects are then used to construct high level physics objects:

- Electrons
- Muons
- Taus [1]
- Jets – a collimated group of particles resulting from the hadronisation of quarks or gluons
- Photons [2]
- Missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) – the imbalance in the total transverse momentum of all reconstructed objects in an event

These reconstructed objects are the fundamental building blocks in most ATLAS analyses. There is a distinction between objects produced in the primary collision (prompt) and produced through secondary decays (non-prompt). Prompt objects are essential for reconstructing decay chains and for probing both Standard Model processes, and potential signatures of new physics. Table 3.1 shows the full range of Standard Model particles and the associated signature used to reconstruct them in the ATLAS detector.

[1] Taus will not be covered in detail in this thesis.
[2] Neither will photons.

Table 3.1: Summary of SM particle properties and detector signatures. The masses are taken from Ref. [39]. Metastable indicates particles which have a finite lifetime in ATLAS and give displaced vertex signatures. The metastable quarks also hadronise. "Stable" indicates particles, which are stable within the ATLAS, though not fundamentally so.

Particle	Mass [GeV]	Stability	Detector signature
Quarks			
Up u	2×10^{-3}	Hadronise	ECal, HCal (jets)
Down d	5×10^{-3}	Hadronise	ECal, HCal (jets)
Strange s	0.1	Hadronise	ECal, HCal (jets)
Charm c	1.3	Metastable	ECal, HCal (jets), displaced vertex
Bottom b	4.2	Metastable	ECal, HCal (jets), displaced vertex
Top t	173	Unstable	m_t from e.g. $t \to bW \to$ jets
Leptons			
Neutrinos $\nu_{e,\mu,\tau}$	$< 10^{-9}$	Stable	Missing transverse momentum
Electron e	5×10^{-4}	Stable	ID hits, ECal
Muon μ	0.106	Stable	ID hits, muon spectrometer
Tau τ	1.78	Metastable	Jets or leptons, displaced vertex
Bosons			
Gluon g	0	Hadronise	ECal, HCal (jets)
Photon γ	0	Stable	No tracks, ECal deposit
$W^{\pm}$ boson	80.3	Unstable	m_W from e.g. $W \to \mu\nu_\mu$
Z boson	90.1	Unstable	m_Z from e.g. $Z \to \mu\mu$
Higgs boson h	125.1	Unstable	m_h from e.g. $h \to \gamma\gamma$

3.3.1 Jets

Since quarks [3] and gluons cannot be observed directly due to colour confinement, they are reconstructed in the detector as jets. The quarks and gluons hadronise, radiating showers that form cones of hadronic activity. Jet reconstruction in ATLAS is carried using the anti-k_t algorithm [40], implemented in `FastJet` [41], with a standard jet radius parameter of $R = 0.4$. Considering a set of all PFO objects in an event, the distance between two clusters (d_{ij}), and the distance between each cluster and the beam (d_{iB}), are defined as follows:

$$\begin{aligned} d_{iB} &= p_{\mathrm{T}i}^{2n}, \\ d_{ij} &= \min\left(p_{\mathrm{T}i}^{2n}, p_{\mathrm{T}j}^{2n}\right) \frac{\Delta R_{ij}^2}{R^2}. \end{aligned} \tag{3.1}$$

[3] Other than the top quark.

where $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, is the radial distance between clusters. Additionally, y_i and ϕ_i are the rapidity and azimuthal angle of the cluster i, and $p_{\mathrm{T}i}$ is the transverse momentum. For anti-k_t, n is fixed to -1. The value of R defines the width of the reconstructed jets, where the default value used in ATLAS is 0.4, with alternative values utilised for particular analyses. These defined values are calculated for all possible distances, then:

- If the smallest element is a d_{iB} the related cluster becomes a jet, and it is removed from consideration,
- If the smallest element is a d_{ij}, the objects are combined, and a new object replaces i and j in the list, with $p_{\mathrm{T\ new}} = p_{\mathrm{T}i} + p_{\mathrm{T}j}$.

This process is repeated until all clusters are classified as jets.

The jet energy scale and resolution (JES and JER) are calibrated using simulations and the corrections are made by applying calibration factors. The relative uncertainties are derived as a function of the p_{T} and η of the jet, as well as the pile-up conditions and the jet flavour composition. They are determined using a combination of data and MC simulated samples, comparing the jet response in di-jet, Z+jet and γ+jet events [42, 43]. Certain aspects of this calibration are not available at the trigger level as they depend on quantities like track multiplicity. This forms the distinction between online (without this calibration) and offline (with this calibration) objects utilised in Chapter 6.

Jets from pileup interactions are also taken into account, by utilising tracking information. The jet vertex fraction (JVF) [44] helps remove jets that are not associated to the primary vertex. JVF is defined as the scalar sum of the p_{T} of tracks associated with the jet and that originated from the primary vertex, divided by the scalar sum of the p_{T} of all the tracks associated to the jet. Applying a cut on JVF allows the majority of pile-up jets to be rejected, but leads to a selection that depends on the number of reconstructed primary vertices in the event ($\mathrm{N_V}$). To solve this issue, a two-dimensional likelihood combination is carried out, using

JVF and N_V, called the Jet Vertex Tagger (JVT) [45], is constructed, so that the resulting jet selection efficiency is not dependent on N_V.

3.3.2 Jet Flavour Tagging

In ATLAS, jets originating from bottom quarks can be 'tagged' as b-jets, and similarly for charm quarks (c-jets), or strange, up, down quarks or gluons (light jets). This is possible due to the finite lifetime of b-quarks, and to a lesser extent c-quarks, which lead to displaced vertices and distinctive decay kinematics. The flavour tagging (f-tag) efforts in ATLAS are split into low-level and high-level taggers. The low-level taggers utilise a whole host of track information, impact parameter values, and secondary vertex information.

A variety of high level taggers were developed over the course of Run 2 [46], and the relevant tagger to this thesis is DL1r [45]. DL1r is a deep neural network which combines the low-level information with the jet p_T, η information, to provide the probability a jet is from a b-, c- or light-jet (p_b, p_c, and p_l). The probabilities are combined into a log-likelihood discriminant defined as:

$$D_{\mathrm{DL1r}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_l}\right), \tag{3.2}$$

where f_c represents the effective c-jet fraction in the background training sample, in addition to the probabilities defined earlier. A jet is b-tagged, if its DL1r score exceeds a pre-defined threshold, called a working point (WP), defined using the b-tagging efficiency measured in top, anti-top ($t\bar{t}$) events [47]. The default WP for DL1r is the 77% WP, which can be interpreted as the tagger having a 77% b-tag efficiency on $t\bar{t}$ MC.

3.3.3 Electrons

Electrons are identified by combining tracks from the ID and EM calorimeter clusters [48]. In the forward region ($2.5 < |\eta| < 4.9$) only calorimeter information is used, as there is no ID coverage. Additionally, there is a crack in the calorimeter $1.37 < |\eta| < 1.52$, which has a relatively large amount of inactive material. This

leads to large uncertainties on the momenta of the reconstructed electrons in the region, and this η range can be vetoed if desired.

Tracks matched to a cluster are refitted using a Gaussian Sum Filter [49], to take into account electron energy loss via bremsstrahlung. In the case where multiple tracks are matched to a cluster, the track closest to the cluster is chosen. Then, the tracks are refitted to clusters with more stringent requirements. If an electron candidate has no hits after the refitting, they are removed and designated photons. Additionally, electron candidate tracks must be compatible with the primary vertex and satisfy impact parameter requirements.

Then a likelihood discriminant is constructed from: track information, track-cluster matching information, and variables which describe the shape of the electromagnetic shower [50]. Identification working points (WPs) are then defined for the reconstructed electrons:

- VeryLooseLLH – 97% reconstruction efficiency
- LooseLLH – 95% reconstruction efficiency
- LooseAndBLayerLLH – 95% the same as loose, but including IBL hit information to further improve reconstruction efficiency.
- MediumLLH – 90% reconstruction efficiency
- TightLLH – 80% reconstruction efficiency

The reconstruction efficiencies given are approximate and for electrons with $p_{\mathrm{T}} \approx$ 40 GeV. Figure 3.1, shows the efficiency of a selection of WPs on $Z \to ee$ events.

The electrons are then calibrated using well understood processes, $Z \to ee$ and $J/\psi \to ee$. Calibration is necessary to correct for effects like energy loss in passive detector material, electromagnetic shower leakages, and mismeasurement in the deposited calorimeter energy. The electron energy scale and resolution are corrected using scale factors derived from fits to the invariant mass distributions, ensuring an accurate reconstruction of the resonance peak. Uncertainties on this calibration, along with reconstruction and identification efficiency, are computed

using the error on the scale factors applied to match the MC with the data. Figure 3.2, shows the di-electron invariant mass $Z \to ee$ events, in data compared with calibrated and uncalibrated MC.

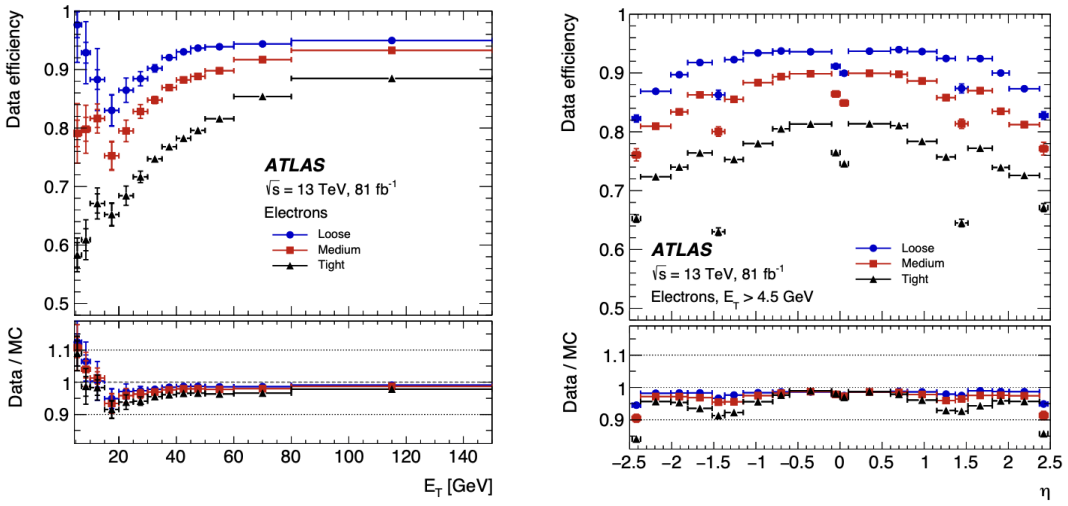


Figure 3.1: Electron distributions in (a) E_{T} and (b) η for the loose, medium and tight WPs. The subplot shows the data/MC ratio. Figures taken from Ref. [50].

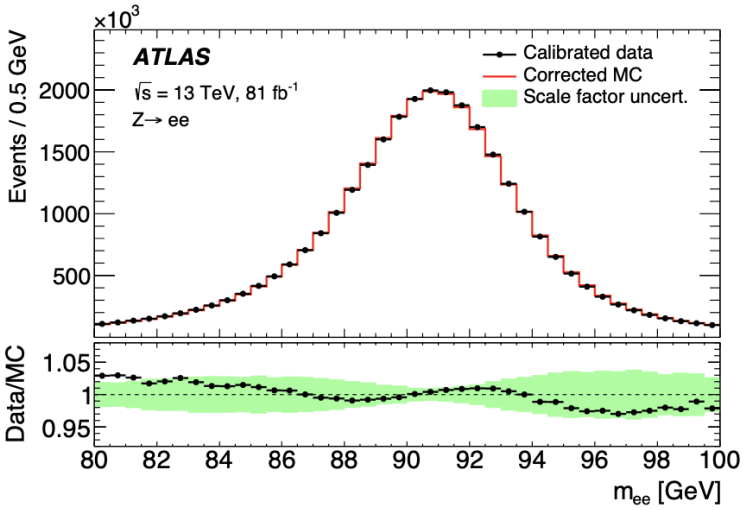


Figure 3.2: Di-electron invariant mass distributions for $Z \to ee$ events. The subplot shows the data/MC ratio and the uncertainty band illustrates the impact of the uncertainties in the calibration and resolution correction factors. Figure taken from Ref. [50].

3.3.4 Muons

Muon reconstruction [51] primarily utilises information provided by the Inner Detector and the Muon Spectrometer. There are four different types of muon, depending on the sub-detectors used to reconstruct them:

- Combined muons – identified by matching a MS track with an ID track. These muons are the most precisely measured, as the momentum is recalculated with a global fit using both sets of track information.
- Segment tagged muons – identified using a combination of tracks in the ID with at least one partially reconstructed track in the MDT or CSC.
- Calorimeter tagged muons – identified from ID track(s) matched to a calorimeter cluster, where the cluster is compatible with a minimum ionising particle. This is a particle that loses energy due to ionisation at or near the minimum rate as it passes through matter.
- Extrapolated muons – which are reconstructed only using an MS track, with a relatively loose association requirement with the primary vertex.

Working points similar to the electron case are then defined:

- Loose – aim to maximise reconstruction efficiency. All muon categories are used, with only kinematic requirements to veto η regions where the MS has low efficiency.
- Medium – the default selection used in ATLAS. These are built in order to reduce the systematic uncertainties of the reconstructed muons. Only combined and extrapolated muons are used, in different $|\eta|$ regions. In addition, a loose selection is applied related to the compatibility between ID and MS momentum, to reduce contamination from misidentified hadrons.
- Tight – these are selected to maximise the purity of the muon selection. Only combined muon tracks are used, with a strict requirement on the reconstruction quality and compatibility between the ID and MS momenta measurement.

In addition to these WPs, there are also low- and high-p_{T} WPs, which will not be covered here, but are specialised to increase efficiency for muons with $p_{\mathrm{T}} < 5$ GeV and > 100 GeV respectively. These WPs have a strong η dependence, and the efficiency as a function of η is shown in Figure 3.3. Not shown in the figure, the loose WP is able to increase efficiency for forward muons, due to the use of calorimeter and segment tagged muons.

In order to evaluate the reconstruction efficiency of these different muons, well understood final states are used: $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ The momentum scale and reconstruction are compared between data and MC, to derive scale factors to correct for differences arising detector effects like energy losses in the detector, which are not accurately described in simulation. The invariant mass distributions for the targeted final states, for the uncorrected and corrected MC are shown in Figure 3.4. An uncertainty associated to the momentum resolution is obtained by varying the fit parameters between the ID and MS tracks used to define the muons. Additional uncertainties are derived on the reconstruction efficiency by considering the error on the scale factors applied to match the MC to the data [52].

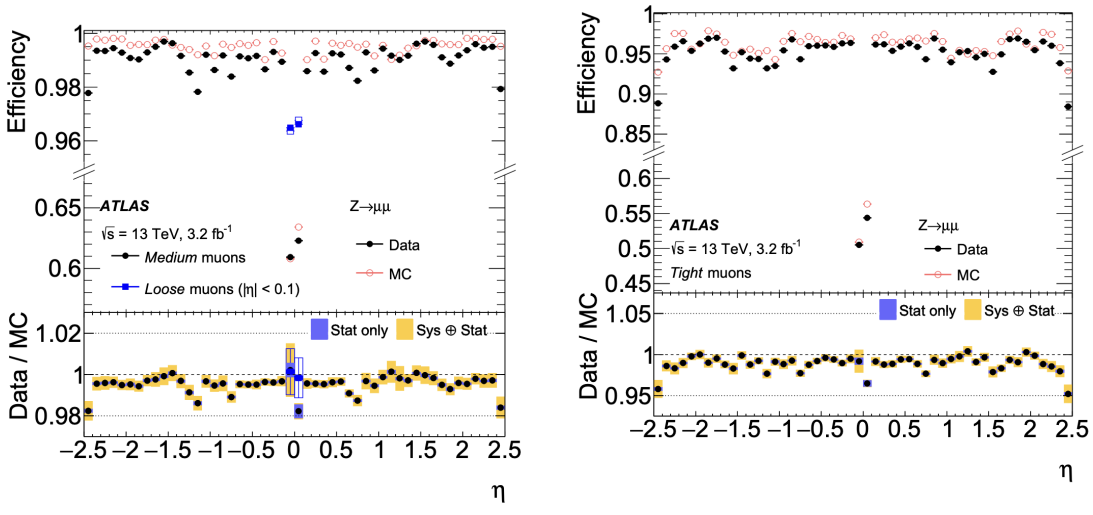


Figure 3.3: Muon efficiency distributions as a function of η for (a) the loose and medium WPs and (b) the tight WP. The subplot shows the data/MC ratio. Figures taken from Ref. [51].

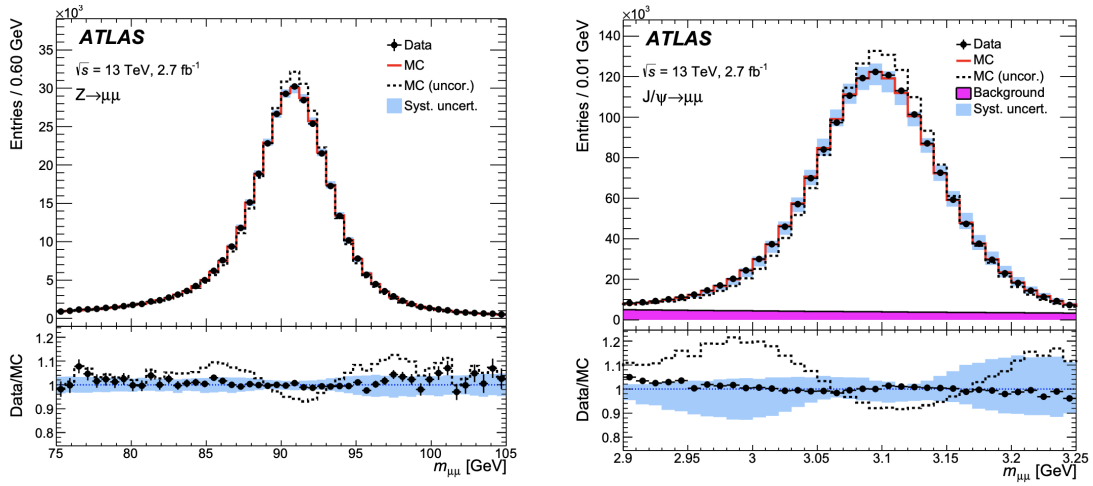


Figure 3.4: Di-muon invariant mass distributions for (a) $Z \to \mu\mu$ and (b) $J/\psi \to \mu\mu$, using combined muons. In both cases, a comparison is shown for data, with the uncorrected and corrected MC distributions. The subplot shows the data/MC ratio and the uncertainty band illustrates the impact of the uncertainties in the calibration and resolution correction factors. Figures taken from Ref. [51].

3.3.5 Lepton Isolation

Fake/non-prompt (FNP) leptons are generally produced near a jet, surrounded by high detector activity. This contrasts with prompt leptons which are generally well separated from other particles in the event. The measure of detector activity around a reconstructed lepton is called isolation, and requirements are implemented reduce the contamination from the FNP events in reconstructed leptons.

An isolation-based discriminant, similar to the one described in Ref. [53], is employed for both electrons and muons. Two boosted decision trees (BDTs) are trained, one general purpose BDT and one specifically targeting low-p_{T} leptons (< 10 GeV). For electrons, calorimeter energy deposits and charged-particle tracks within a cone around the electron candidate are used as inputs. Similar information is used for muons, with the addition of MS track information. Different working points are defined using requirements on the BDT output scores, as a function of lepton p_{T}:

- PLVLoose – loose requirement on the BDT scores, high selection efficiency but rejects less FNP events.

- PLVTight – higher requirement on the BDT scores, lower selection efficiency but rejects more FNP events.
- PLImprovedTight – similar to PLVTight but using an improved BDT training. [4]
- PLImprovedVeryTight – more stringent requirement on the PLImprovedTight BDT score, largest FNP rejection but lowest selection efficiency.

Figure 3.5, shows the efficiency to select well-identified prompt leptons for electrons and muons with the PLVTight WP. Scale factors are derived for each WP and lepton flavour, to account for differences between data and MC, and can be interpreted as the ratio in the subplot of the figure. In the analysis covered in Chapter 5, the PLVLoose isolation WP was chosen to maximise signal selection efficiency. For electrons, the PLVLoose isolation WP gives an efficiency between 80% and 95% for prompt electrons with p_{T} of 10 GeV and 100 GeV, respectively. For muons, the PLVLoose isolation WP gives an efficiency between 86% and 98% for prompt muons with p_{T} of 10 GeV and 100 GeV, respectively.

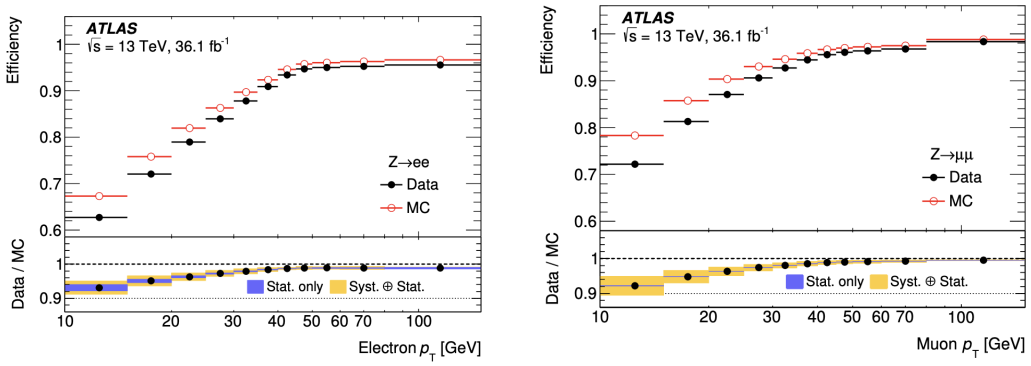


Figure 3.5: The efficiency to select well-identified prompt leptons in data and MC for (a) electrons requiring tight ID and the PLVTight isolation WP, and (b) muons with loose ID and the PLVTight isolation WP. The subplot shows the data/MC ratio and can be interpreted as the correction scale factors applied to each lepton. The uncertainty band illustrates the impact of the systematic uncertainties. Figures taken from Ref. [53].

[4] This includes introducing new, more sensitive variables and using secondary vertex information among other innovations.

3.3.6 Overlap Removal

To prevent the use of the same reconstructed detector signals in multiple objects, an overlap removal (OR) procedure is applied to the reconstructed objects. The OR procedure considers: of electrons, muons and jets, and removes overlapping objects based on ΔR matching or shared inner detector tracks. [5] In addition, the 'boosted leptons' prescription of the OR procedure is used, which implements a sliding ΔR cone that shrinks with the jet p_{T} in the lepton-jet OR step. The order of overlaps removed between different objects is as follows:

1. Muon-electron – reject electron if ID track is shared with muon
2. Jet-electron – reject jet if $\Delta R < 0.2$
3. Electron-jet – reject electron if $0.2 < \Delta R < \min(0.4, 0.04 + 10\mathrm{GeV}/p_{\mathrm{T}}(e))$
4. Jet-muons – reject jet if $\Delta R < 0.2$
5. Muon-jet – reject muon if $0.2 < \Delta R < \min(0.4, 0.04 + 10\mathrm{GeV}/p_{\mathrm{T}}(\mu))$

3.3.7 Missing Transverse Energy

Due to the conservation of momentum, the transverse momentum measured a detector in a collider experiment like ATLAS, should sum to zero in each event. Any deviation from zero is called missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) [54]. Large values of $E_{\mathrm{T}}^{\mathrm{miss}}$ can indicate the presence of weakly-interacting particles in an event. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as the magnitude of the negative vector sum of the transverse momenta of all reconstructed objects. This is calculated before overlap removal, as the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation carries out its own procedure where calorimeter clusters associated to multiple objects are assigned with the following priority: electrons, photons, hadronically decaying taus, jets, then muons.

The $\overrightarrow{\boldsymbol{E}}_{\mathrm{T}}^{\mathrm{miss}}$ is defined as:

$$\overrightarrow{\boldsymbol{E}}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{e} \overrightarrow{\boldsymbol{p}_{\mathbf{T}}} - \sum_{\gamma} \overrightarrow{\boldsymbol{p}_{\mathbf{T}}} - \sum_{\mu} \overrightarrow{\boldsymbol{p}_{\mathbf{T}}} - \sum_{\tau} \overrightarrow{\boldsymbol{p}_{\mathbf{T}}} - \sum_{\text{jets}} \overrightarrow{\boldsymbol{p}_{\mathbf{T}}} - \sum_{\text{soft term}} \overrightarrow{\boldsymbol{p}_{\mathbf{T}}}. \tag{3.3}$$

[5] In the analysis in Chapter 5, photons and taus are not considered in the OR procedure, and so these will be left out of the following description.

In the first four sum terms the subscript indicates the sum over all of a given object, for example the first term indicates the negative vector sum of the transverse momentum of all electrons. For the sum over jets, the object list is modified to reduce the impact of jets from pileup: central jets ($|\eta| < 2.4$) are required to pass the JVT selection, and forward jets are required to have $p_{\mathrm{T}} > 30$ GeV. The soft term is constructed from all ID tracks matched to the primary vertex but not associated with any of the reconstructed objects.

The $\overrightarrow{\boldsymbol{E}}_{\mathrm{T}}^{\mathrm{miss}}$ is composed of x and y components, and the magnitude and associated angle can be calculated using:

$$\begin{aligned} \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} &= \left(E_x^{\mathrm{miss}}, E_y^{\mathrm{miss}}\right), \\ E_{\mathrm{T}}^{\mathrm{miss}} &= \left|\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}\right| = \sqrt{(E_x^{\mathrm{miss}})^2 + \left(E_y^{\mathrm{miss}}\right)^2}, \\ \phi^{\mathrm{miss}} &= \tan^{-1}\left(E_y^{\mathrm{miss}}/E_x^{\mathrm{miss}}\right). \end{aligned} \tag{3.4}$$

The resolution of the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ can be calculated using $Z \to \mu\mu$ events which contain very little real $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 3.6 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the $E_{\mathrm{x}}^{\mathrm{miss}}$ and $E_{\mathrm{y}}^{\mathrm{miss}}$ planes, for data and MC. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is calculated using the root-mean-square of the combined $E_{\mathrm{x}}^{\mathrm{miss}}$ and $E_{\mathrm{y}}^{\mathrm{miss}}$, and is parametrised as a function of the scalar sum of the transverse energy ($\sum E_T$), and of the number of vertices. The resolution varies with the magnitude of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the targeted final state, ranging from $\approx 10--15$ GeV at lower values of $E_{\mathrm{T}}^{\mathrm{miss}}$ (< 200 GeV), increasing with $E_{\mathrm{T}}^{\mathrm{miss}}$ to above 30 GeV in the tails of the distribution. An object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ significance can be used to identify events where the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ comes from undetected particles rather than mismeasured energy deposits by dividing the $E_{\mathrm{T}}^{\mathrm{miss}}$ by an estimate of the resolution [55].

The uncertainty related to the $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated by propagating the uncertainties of all the reconstructed objects, in addition to a dedicated uncertainty on the soft term. The soft term uncertainty is decomposed into components which are parallel and perpendicular to the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction when removing the soft term and comparing MC to data to derive the uncertainty.

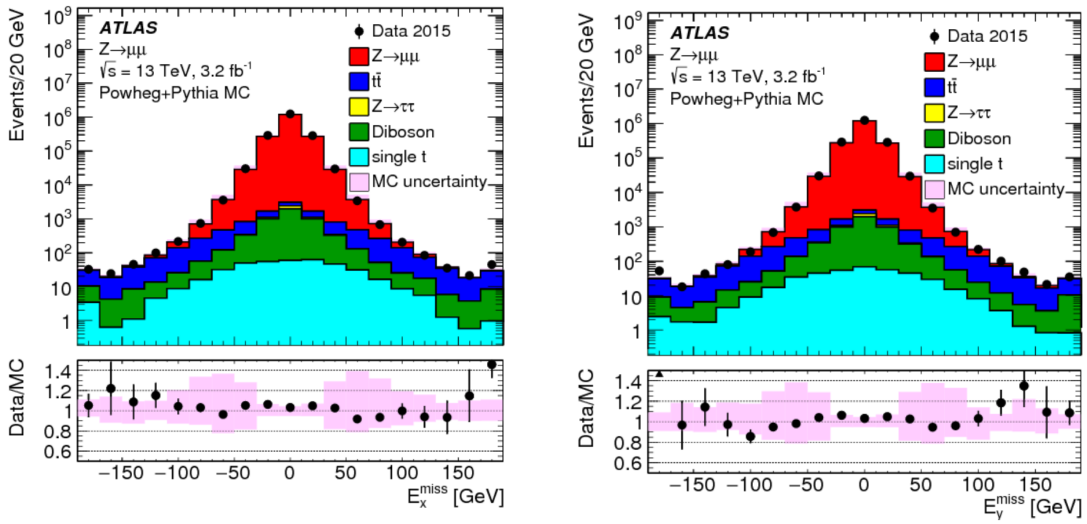


Figure 3.6: Distributions of (a) $E_{\mathrm{x}}^{\mathrm{miss}}$ and (b) $E_{\mathrm{y}}^{\mathrm{miss}}$, for data and MC backgrounds using $Z \to \mu\mu$ events. The subplot shows the data/MC ratio. The uncertainty band illustrates the impact of the systematic uncertainties. Figures taken from Ref. [56].

Sorry, guys; I gotta see about a girl.

Good Will Hunting

4

The Standard Model and Supersymmetry

Contents

4.1 Supersymmetry 44
4.2 Problems with The Standard Model 46
4.2.1 Dark Matter and R-Parity 46
4.2.2 Hierarchy Problem and the Higgs Mass 49
4.2.3 The Anomalous Muon Gyro-magnetic Ratio 51
4.2.4 Grand Unification 52
4.3 The Minimal Supersymmetric Standard Model 53
4.3.1 Symmetry Breaking 54
4.3.2 Sparticles and Mixing 55
4.3.3 Simplified Models 57
4.4 Experimental Searches for SUSY 60
4.4.1 Strong SUSY 61
4.4.2 Electroweak SUSY 63
4.5 Summary 69

The Standard Model (SM) of particle physics is the theoretical framework that describes fundamental particles and their interactions. It is a quantum field theory based on the gauge symmetry: $SU(3)_C \times SU(2)_L \times U(1)_Y$, which governs the interactions of quarks, leptons, and gauge boson interactions of the strong, weak, and electromagnetic forces. The Standard Model is one of the most rigorously tested theories in physics, and its theoretical predictions have consistently shown strong agreement with experimental observations to date [39]. However, it is not

a complete theory and there are clear problems the SM, which have led to the proposal of beyond the Standard Model (BSM) theories.

4.1 Supersymmetry

One of the more well motivated of these BSM theories is supersymmetry (SUSY) [57–62]. This chapter will only be a brief introduction to SUSY, enough to provide motivation for the theory, and to build an understanding of predictions, as related to the analysis described in Chapter 5. A much deeper review of SUSY can be found in Ref. [63]. SUSY imposes an additional symmetry not present in the SM, between fermions and scalars, such that each fermion degree of freedom has a scalar partner and vice versa.

Initially, SUSY arose from the proof, the exception to the Coleman-Mandula theorem, that the form of symmetry defined in Eq. 4.1, is the only possible extension of the Poincaré group, which forms the foundation the space-time invariance of the SM [64]. In a sense, all the problems with the SM solved by SUSY are accidental consequences of imposing this fundamental symmetry.

A representation of the complex operator (Q) which defines the SUSY is:

$$\begin{cases} Q|\text{boson}\rangle = |\text{fermion}\rangle \\ Q|\text{fermion}\rangle = |\text{boson}\rangle. \end{cases} \tag{4.1}$$

The operator Q satisfies the following relations:

$$\left\{Q_\alpha, Q^\dagger_{\dot\alpha}\right\} = -2\sigma^\mu_{\alpha\dot\alpha}P_\mu, \tag{4.2}$$

$$\{Q_\alpha, Q_\beta\} = \left\{Q^\dagger_{\dot\alpha}, Q^\dagger_{\dot\beta}\right\} = 0, \tag{4.3}$$

$$[Q_\alpha, P_\mu] = \left[P_\mu, Q^\dagger_{\dot\alpha}\right] = 0. \tag{4.4}$$

Here, σ^μ is the Pauli four-vector, with $\sigma^0 = \mathbb{I}_2$, σ^i are the three Pauli matrices, and P^μ, is the momentum operator. The first relation shows that if we perform two successive SUSY transformations on a field (eg: boson $\rightarrow$ fermion $\rightarrow$ boson), the net effect is to move it in spacetime, in the direction of a four-momentum given by P^μ. In other words, SUSY connects quantum properties of particles, with motion in spacetime. The second relation shows that SUSY transformations do not generate

any new bosonic charges, other than spacetime translations. This ensures that SUSY does not introduce additional unwanted internal symmetries, and instead only introduces a relation between fermions and bosons. Finally, the third relation tell us that the SUSY generators conserve momentum. So the transformations act on the internal degrees of freedom, but do not change the momentum of a particle. This ensures that supersymmetry respects spacetime translation symmetry and acts independently, as an extension of the Poincaré group generators.

The major consequence of these relations is that for each degree of freedom of each particle in the SM, SUSY predicts the existence of another partner particle with the same quantum numbers, but with its spin quantum number differing by $\frac{1}{2}$. These are named superpartners or sparticles. The naming convention is to add an 's' to the SM fermion name to give the corresponding bosonic superpartner field, and the suffix 'ino' to the SM bosons corresponding fermionic superpartner. The SM particles and their corresponding superpartners exist in irreducible representations of the SUSY algebra as generated by Q and $Q^\dagger$, called supermultiplets.

Supermultiplets must contain the same number of bosonic and fermionic degrees of freedom. Additionally, Q and $Q^\dagger$ also commute with the generators of the gauge symmetries of the Standard Model, and so particles in the same supermultiplet necessarily have the same colour, electric charge and weak isospin. Most importantly, particles in a given supermultiplet must have the same mass, if SUSY is an exact symmetry. This is due to another important consequence of Eq. 4.4, which can be used to show that P^2 commutes with the SUSY generators. Therefore, all the particles belonging to the same supermultiplet must have the same eigenvalue of P^2, and hence the same mass. However, as we have not observed superpartners of SM particles, this symmetry must be broken in our universe. A mechanism to generate this symmetry breaking, and how to define an experimentally testable version of SUSY will be introduced in Section 4.3.

4.2 Problems with The Standard Model

This section will discuss the deficiencies of the SM, and explore how SUSY could provide a mechanism to explain them. The discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations in 2012 [65, 66] was the discovery of the final particle predicted by the SM. However, there remains several significant questions about key experimental observations, which cannot be answered by the SM. These are wide-ranging, from the lack of a mechanism to explain the small, but non-zero, neutrino masses, to the inability to describe gravity in the framework of the SM. Only problems with the SM that could be solved by SUSY will be discussed in the following section.

4.2.1 Dark Matter and R-Parity

Perhaps the most striking gap in the particle content of the SM, is the lack of a possible candidate for dark matter (DM) [67–69]. There is a host of evidence for dark matter, with the first hints in the 1970's, coming from galactic rotation curves [70]. Prior to these measurements, the rotational velocity of a galaxy was expected to decrease with the inverse square root of the distance from the galactic centre, as predicted by Newtonian gravity. However, the rotation curves obtained by Rubin and Ford demonstrated that the velocities tend to remain approximately constant at large distances. This unexpected flattening in the rotational velocity implies the presence of an extended and roughly uniform mass distribution, inconsistent with the distribution of visible matter alone. Visible matter is concentrated at the centre of a galaxy, and so there must be a large amount of invisible matter present [1] even at the edges.

Observations of phenomena such as the Bullet Cluster, a collision between two clusters of galaxies, provide additional evidence for dark matter [71]. In this collision, most of the visible matter (primarily hot gas) interacts and slows down, becoming separated from the galaxies, which continue moving largely undisturbed. Measurements of the gravitational potential, inferred from gravitational lensing,

[1] Or we have understood gravity incorrectly, but we now think that is not the case.

show that it is centred around the galaxies themselves rather than the visible gas. This indicates that most of the mass is not in the visible matter but in some category of invisible matter, consistent with dark matter and very difficult to explain through theories of modified gravity.

In addition to this evidence, observations of the cosmic microwave background (CMB) by the WMAP and Planck missions have provided independent support for the existence of dark matter [72, 73]. These measurements, which map temperature fluctuations in the CMB across the sky, are sensitive to the composition of the early Universe. By analysing the amplitude and spatial distribution of these fluctuations, scientists have inferred that dark matter makes up roughly five times more mass than ordinary, visible matter. When considering the large amount of independent evidence, the search for dark matter is one of the most compelling questions in modern Particle Physics.

The MSSM, as described in Section 4.3, can lead to the non-conservation of baryon and lepton quantum numbers. As a consequence, the proton is not stable, and additionally, no SUSY particle could be a dark matter candidate. However, proton decay has not yet been observed, with the current upper limit on the lifetime set at 1.6×10^{34} years, set by the Super-Kamiokande experiment [74]. Consequently, it is well motivated to introduce a new conservation rule ensure that the proton remains stable. This new conserved quantum number is known as R-parity, and is often introduced as an assumption in supersymmetric theories for experimental searches [75]. The definition of R-parity is:

$$R = (-1)^{3(B-L)+2S}, \tag{4.5}$$

where for a given particle, B is the baryon number, L is the lepton number, and S is the spin. All SM particles have $R = 1$, and all sparticles have $R = -1$.

Enforcing this new conservation rule has major consequences on SUSY phenomenology. The conservation of baryon and lepton quantum numbers is restored, the proton is stable, as is the lightest supersymmetric particle (LSP), and therefore the LSP is a strong candidate for dark matter [76, 77]. The specific LSP candidate

investigated in this thesis is the lightest neutralino ($\widetilde{\chi}_1^0$), which will be discussed in Section 4.3.

Another consequence is that the decay chain of all non-stable sparticles must end with a decay into an odd number of LSP(s), plus SM particle(s). Finally, in collider experiments, sparticles must be produced in pairs to ensure R-parity is conserved.

The mass of the lightest neutralino ($\widetilde{\chi}_1^0$) is unknown, but is constrained by other measurements like the dark matter relic density. One of the favoured categories of dark matter particles are so called 'weakly interacting massive particles' (WIMPs). WIMP dark matter (often denoted χ) is based on the hypothesis that WIMPs are thermal relics, and were in thermal equilibrium with other SM particles during the early universe. As the universe cooled, WIMPs decoupled from other matter through a process called 'freeze-out', meaning their annihilation rate dropped below the Hubble expansion rate, and they could no longer efficiently annihilate.

Dark matter can come in two main categories [2]. Non-relativistic (cold) dark matter is essential to reproduce the observed structure formation and to maintain consistency with measurements of the CMB. In contrast, relativistic (hot) dark matter, such as neutrinos, would suppress the formation of small-scale structures in the universe and lead to deviations from the observations of the CMB. One of the strongest constraints on the potential WIMP mass is the observed relic density ($\Omega_\chi h^2$). [3]

So, assuming non-relativistic (cold) WIMP dark matter, the predicted relic density can be approximated using:

$$\Omega_\chi h^2 \simeq \frac{0.1\text{pb}}{\langle \sigma v \rangle} \simeq 0.1 \left(\frac{0.01}{\alpha}\right)^2 \left(\frac{m_\chi}{100\text{GeV}}\right)^2, \tag{4.6}$$

where 'm_χ' is the WIMP (in our case $\widetilde{\chi}_1^0$) mass, 'α' is the effective coupling strength between the WIMP and SM particles, governing the annihilation interaction, and $< \sigma \text{v} >$ is the thermally-averaged self-annihilation ($\chi + \chi \rightarrow SM + SM$) cross-section [78]. If m_χ is around the weak scale – $\mathcal{O}(100$ GeV) – then the value observed

[2] With 'warm' dark matter being a mix of the two

[3] Where in this context $h = H_0/(100\text{kms}^{-1}\text{megaparsec}^{-1})$, the reduced Hubble constant, not the Higgs boson.

by Planck, $\Omega_\chi h^2 = 0.1200 \pm 0.0012$ [72], is obtained. Therefore, if $\widetilde{\chi}_1^0$ is dark matter and thermally produced in the early universe, we can constrain the allowed mass through the measurement of dark matter relic density by WMAP/Planck. Through the introduction of R-parity, SUSY becomes a strong contender for a theory which could provide a dark matter candidate.

This equation does not take into account any co-annihilation between the dark matter candidate and other particles close in mass. If a particle close in mass to the dark matter candidate exists, then this particle can remain in equilibrium and modifies the annihilation rate. A possible example of this in SUSY would be a chargino, which is close in mass to the $\widetilde{\chi}_1^0$.

4.2.2 Hierarchy Problem and the Higgs Mass

The most precise measurement of the Higgs boson mass is 125.11 ± 0.11 GeV, by the ATLAS experiment [79]. To recover this mass from the SM, extremely fine-tuned cancellations between higher order corrections to the Higgs mass are necessary. This arises because in the SM, the bare Higgs mass (m_{h0}) is a free parameter of the Lagrangian that describes the theory, and higher order corrections are proportional the Higgs mass. In quantum field theories, renormalisation is used to deal with the ultraviolet divergences, which arise from these higher order contributions [80, 81]. Renormalisation of a theory, imposes a cut-off (Λ) defining the energy scale beyond which the theory breaks down. The leading-order quantum correction

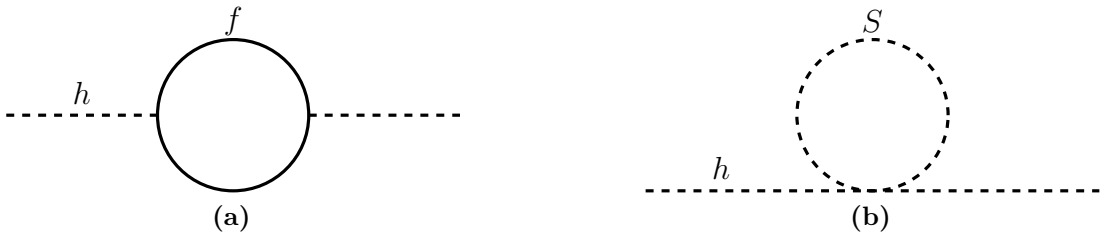


Figure 4.1: Diagrams showing (a) fermionic and (b) scalar first order corrections to the Higgs propagator.

from the fermion loop, depicted in Figure 4.1, arises from the interaction term

$-\lambda_f H f \bar{f}$ in the Lagrangian, where the fermion couples to the Higgs field. This correction to the physical Higgs is given by:

$$m_h^2 \approx m_{h0}^2 - \frac{|\lambda_f|^2}{8\pi^2}\Lambda^2 + \ldots, \tag{4.7}$$

where the omitted terms grow at most logarithmically with Λ.

The presence of the terms proportional to Λ are not an issue alone, but rather the relative magnitude of the terms compared with the observed Higgs boson mass. Comparing the energy scales involved in the calculation of the physical Higgs mass shows the problem. We do not know the scale of Λ, in other words we do not know at what energy scale the SM becomes invalid, but typically this is taken to be the Planck scale, $\Lambda \approx 10^{19}$ GeV. There must necessarily be a large amount of fine-tuning such that the very large correction terms proportional to the Plank scale, cancel exactly to give the observed Higgs mass of 125.11 GeV. This goes against the principles of naturalness, which expects all relevant parameters of a theory to be the same order of magnitude [82, 83]. The larger problem is that this means that the SM has quadratic sensitivity to high mass scales, so small differences in the parameters that parameterise the corrections, would lead to huge changes in the observed Higgs mass.

This high mass scale fine-tuning can be resolved through SUSY. Figure 4.1 also shows a possible scalar first order correction to the Higgs mass, and first order scalar contribution is then:

$$\Delta m_h^2 \approx +\frac{\lambda_s}{16\pi^2}\Lambda^2 + \ldots, \tag{4.8}$$

where the interaction term between the scalar and the Higgs is: $-\lambda_s |h|^2 |S|^2$. SUSY predicts two scalars for each SM fermion, one superpartner each for the left- and right-handed fermions. For example the muon has two scalar superpartners, the left- and right-handed smuon. In SUSY, the fundamental relation: $\lambda_s = |\lambda_f|^2$ defines the relation between the couplings of the scalars and fermions. Therefore, in versions of SUSY where the symmetry is unbroken, the higher order corrections to the Higgs mass obtained from Eq. 4.7 and 4.8, cancel exactly. This ensures

that the divergences cancel exactly and resolves the quadratic sensitivity of the SM to high mass scales. In versions of SUSY where the symmetry is broken – like the MSSM – log terms remain but the size of the fine-tuning is more reasonable so long as the SUSY symmetry breaking scale is around the TeV scale. Though an aesthetic argument, the resolution of the hierarchy problem remains a key motivator in the search for SUSY.

4.2.3 The Anomalous Muon Gyro-magnetic Ratio

A more recent anomalous result which could be explained by SUSY, is the measurement of the muon anomalous gyro-magnetic ratio ($g-2$), at Fermilab [84]. This result is in tension with the 2020 predicted theory value at the 4.2 σ level [85]. More recent investigations by theorists, using lattice QCD, have indicated that the source of this discrepancy may instead be due to problems with the theoretical prediction of the SM value [86]. However, while this discrepancy is not conclusively unresolved, it is worth evaluating possible theoretical solutions to this problem.

SUSY can provide an explanation to this discrepancy through additional loop corrections to the muon magnetic moment, shown in Figure 4.2, involving the smuon - the scalar superpartner of the muon. Efforts by theorists to explain the observed value for $g-2$ through SUSY are numerous [87–89], and a few important conclusions can be drawn. The first is that relatively light smuons (< 1 TeV) could explain the observation. Additionally, many SUSY models which are compatible the observed $g-2$ value, cannot yet be ruled out by ATLAS, CMS or direct detection experiments. Finally, these models allow for non-mass-degenerate smuons and selectrons [90, 91], so long as the mass difference is not too large. This adds significant motivation to search for relatively light smuons.

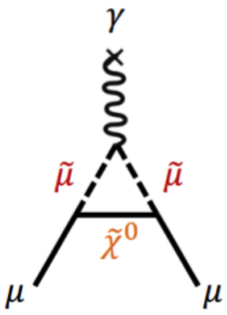


Figure 4.2: Diagram showing a possible correction to the muon g-2 value from SUSY, involving the smuon ($\tilde{\mu}$) and lightest neutralino ($\tilde{\chi}_1^0$).

4.2.4 Grand Unification

Electroweak unification was proposed by Weinberg, Glashow, and Salam provided a mechanism to unify the weak force with electromagnetism [92, 93]. Grand unification theories (GUT) proposed that at some high energy scale, the electroweak forces and strong force become unified through a similar mechanism. However, a problem arises when the couplings of the three fundamental forces are extrapolated to higher energy scales, as shown in Figure 4.3. Assuming the Standard Model holds, they fail to unify at a single point. Figure 4.3 also shows that the addition of supersymmetric particles with mass around 1 TeV leads to unification of the three forces, at an energy scale of 10^{16} TeV [94–96]. This again motivates the search for SUSY as a plausible explanation for physics beyond the Standard Model and grand unification.

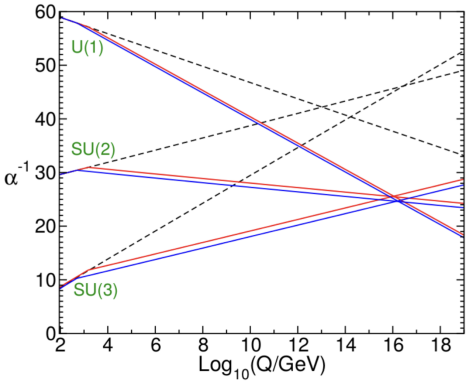


Figure 4.3: The evolution of the inverse gauge couplings of $SU(3)$, $SU(2)$, $U(1)$ as a function of the energy scale for (dashed black lines) the SM and (solid red, blue lines) for SUSY, with the sparticle masses varied between 750 GeV and 2.5 TeV, from Ref. [63].

4.3 The Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model (MSSM), is a version of Supersymmetry whereby the additional supersymmetric fields are restricted to give the smallest number of new additional particles [97, 98]. When constructing the MSSM, the first key consideration is that all observed SM particles must be included in the supermultiplets of the theory. The simplest way to achieve this is by grouping a spin 1/2 fermion field (both left- and right-handed components), with two real scalar fields, which are arranged as a single complex scalar. This specific arrangement is called the 'chiral supermultiplet', and as a result, the right and the left components of the fermion field still behave independently under the gauge transformations, and so all observed leptons and quarks can fit in this new supersymmetric structuring. The scalar superpartners of the leptons and quarks are called 'sleptons' and 'squarks' respectively, and are denoted with the same symbols as their SM partners but with a tilde. As these are scalars, the L, R notation denotes the chirality of the corresponding SM partner.

In addition, two complex Higgs doublets are needed in order to correctly give mass to all particles,

$$\left(H_u^+, H_u^0\right), \left(\tilde{H}_u^+, \tilde{H}_u^0\right), \tag{4.9}$$

and have weak isospin $+1/2$ and $-1/2$. These two new complex doublets have a total of eight real degrees of freedom, corresponding to three Goldstone bosons – which give mass to the electroweak vector bosons, and five observable fields defined in Section 4.3.2. In order to define the two chiral supermultiplets of these doublets, four new sfermion fields are needed, which are again defined with a tilde with respect to the fields in Eq. 4.9. The full set of chiral supermultiplets and associated fields of the MSSM are noted in Table 4.1, along with the SM counterparts and the related charges.

The final piece of the MSSM puzzle is the superpartners of the spin-1 SM vector bosons. The superpartners, known as gauginos, are massless spin-1/2 fermions that exist in two chiral states, so the number of fermions and bosons in SUSY is

the same. The gauge supermultiplets for the gluons and electroweak vector boson fields and their superpartners, are noted in Table 4.2, along with associated charges. It is important to remember that at this stage, before symmetry breaking, these particles are massless and so only have two degrees of freedom.

Table 4.1: Chiral supermultiplets of the MSSM and their associated charges. The right-handed components of the fields are given in the conjugate form, as is convention when defining a chiral supermultiplet.

	spin 0	**spin 1/2**	$\mathbf{SU(3)_c, SU(2)_L, U(1)_Y}$
quarks & squarks	$\left(\tilde{u}_L, \tilde{d}_L\right)$	(u_L, d_L)	$\left(\mathbf{3}, \mathbf{2}, \frac{1}{6}\right)$
	$\tilde{u}_R^*$	$u_R^\dagger$	$\left(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\right)$
$\times 3$ generations	$\tilde{d}_R^*$	$d_R^\dagger$	$\left(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3}\right)$
leptons & sleptons	$(\tilde{\nu}, \tilde{e}_L)$	(ν, e_L)	$\left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}\right)$
$\times 3$ generations	$\tilde{e}_R^*$	$e_R^\dagger$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs & Higgsinos	(H_u^+, H_u^0)	$\left(\tilde{H}_u^+, \tilde{H}_u^0\right)$	$\left(\mathbf{1}, \mathbf{2}, +\frac{1}{2}\right)$
	$\left(H_d^0, H_d^-\right)$	$\left(\tilde{H}_d^0, \tilde{H}_d^-\right)$	$\left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}\right)$

Table 4.2: Gauge supermultiplets of the MSSM and their associated charges.

	spin 1	**spin 1/2**	$\mathbf{SU(3)_c, SU(2)_L, U(1)_Y}$
gluon & gluinos	g	$\tilde{g}$	$(\mathbf{8}, \mathbf{1}, 0)$
W bosons & winos	$W^\pm \quad W^0$	$\tilde{W}^\pm \quad \tilde{W}^0$	$(\mathbf{1}, \mathbf{3}, 0)$
B boson & bino	B^0	$\tilde{B}^0$	$(\mathbf{1}, \mathbf{1}, 0)$

4.3.1 Symmetry Breaking

In the MSSM, all particles and their superpartners would be exactly degenerate in mass if supersymmetry were unbroken, as outlined in the previous section. However, since no superpartners have been observed, so if SUSY exists in nature, must be a broken symmetry. Although the exact mechanism of supersymmetry breaking is unknown, the most general way to describe it is through the formalism introduced in this section. By extending the SUSY Lagrangian to include soft

supersymmetry-breaking terms, it is possible to introduce spontaneous SUSY breaking in a generic way:

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}} , \tag{4.10}$$

where $\mathcal{L}_{\text{SUSY}}$ contains all the gauge and Yukawa interactions that preserve supersymmetric invariance. The additional term $\mathcal{L}_{\text{soft}}$ explicitly breaks supersymmetry, but only through terms with positive mass dimension, such as scalar masses, gaugino masses, and trilinear couplings. This ensures that the soft terms do not reintroduce the quadratic divergences that SUSY naturally cancels, thereby maintaining the ability to solve the Hierarchy problem. The main consequence of soft SUSY breaking is that superpartners can have different masses from their SM counterparts. However, it introduces over 100 free parameters, controlling masses, mixing, and phases controlling the SUSY spectrum, making the theory difficult to test experimentally due to the large number of tunable parameters.

4.3.2 Sparticles and Mixing

Sparticles do not map directly to the fields in Tables 4.1 and 4.2. Mixing between eigenstates can occur for sparticles with the same quantum numbers, as occurs for the neutrinos in the SM. Gluinos have the same quantum numbers, but cannot mix as they carry colour charge. The observable sparticle mass eigenstates of the MSSM are listed in Table 4.3.

The Higgs Sector

For the Higgs sector of the MSSM, the eight Higgs scalar fields mix to form the following five observable bosons:

- Two charged Higgs bosons: H^+, H^-
- A neutral, CP odd, Higgs boson: A_0
- Two CP even Higgs bosons: H_0 and h_0

Table 4.3: The gauge eigenstates and mass eigenstates of the MSSM after symmetry breaking. Mixing between the first and second generations of both squarks and sfermions is experimentally constrained to be negligible.

Names	**Spin**	**Gauge Eigenstates**	**Mass Eigenstates**
Higgs sector	0	$H_u^0 \quad H_d^0 \quad H_u^+ \quad H_d^-$	$h^0 \quad H^0 \quad A^0 \quad H^\pm$
Squarks	0	$\tilde{u}_L \quad \tilde{u}_R \quad \tilde{d}_L \quad \tilde{d}_R$	(same)
		$\tilde{s}_L \quad \tilde{s}_R \quad \tilde{c}_L \quad \tilde{c}_R$	(same)
		$\tilde{t}_L \quad \tilde{t}_R \quad \tilde{b}_L \quad \tilde{b}_R$	$\tilde{t}_1 \quad \tilde{t}_2 \quad \tilde{b}_1 \quad \tilde{b}_2$
Sleptons	0	$\tilde{e}_L \quad \tilde{e}_R \quad \tilde{\nu}_e$	(same)
		$\tilde{\mu}_L \quad \tilde{\mu}_R \quad \tilde{\nu}_\mu$	(same)
		$\tilde{\tau}_L \quad \tilde{\tau}_R \quad \tilde{\nu}_\tau$	$\tilde{\tau}_1 \quad \tilde{\tau}_2 \quad \tilde{\nu}_\tau$
Neutralinos	1/2	$\tilde{B}^0 \quad \tilde{W}^0 \quad \tilde{H}_u^0 \quad \tilde{H}_d^0$	$\tilde{\chi}_1^0 \quad \tilde{\chi}_2^0 \quad \tilde{\chi}_3^0 \quad \tilde{\chi}_4^0$
Charginos	1/2	$\tilde{W}^\pm \quad \tilde{H}_u^+ \quad \tilde{H}_d^-$	$\tilde{\chi}_1^\pm \quad \tilde{\chi}_2^\pm$
Gluino	1/2	$\tilde{g}$	(same)

The h_0, corresponds to the observed SM Higgs boson, and so its mass is constrained to be ≈ 125 GeV. The masses of the other Higgs bosons are necessarily expected to be more massive than the SM Higgs boson. Generic searches for charged Higgs bosons have been carried out within ATLAS [99, 100], with no significant results observed.

Charginos and Neutralinos

The gauginos and higgsinos are able to mix into mass eigenstates as they have the same quantum numbers. Two separate mixings are allowed, for the charged and neutral gauginos and Higgsinos. After mixing of the charged states, there are four observable spin 1/2 states called charginos, denoted as $\tilde{\chi}_i^\pm$. From the neutral states, four observable states called neutralinos, denoted as $\tilde{\chi}_i^0$. The first neutralino state corresponds to the $\tilde{\chi}_1^0$, the stable, neutral, dark matter candidate, as mentioned in Section 4.2.1.

Squarks and sleptons

The left and right handed sleptons and squarks could, in theory mix between generations. However, flavour physics measurements strongly constrain mixing between for first two generations [101, 102]. As a result, this mixing is necessarily assumed to be very small in the MSSM. The third generation is much less constained by other experiments, and flavour mixing is able to be larger.

The SUSY Mass Spectrum

The SUSY mass spectrum can be rather flexible, with a few constraints on mass and ordering, detailed in the previous sections, and from naturalness considerations. A simple example is shown in Figure 4.4, where two possible SUSY mass spectra are shown, assuming degeneracy for the first and second generation sleptons and similary for the first two generations of squarks [4]. In this example, it can be seen that the masses and ordering of, for example, the sleptons versus the squarks can vary significantly. This highlights the difficulty in targeting specific SUSY models in an analysis, without making further simplicifications.

4.3.3 Simplified Models

In addition to the assumptions underlying the MSSM, further simplifications are required to create experimentally testable signal models for optimizing analyses. These 'simplified models', add further assumptions to provide signal models with minimal new particle content, that can still capture the physics of more complex models [103–105]. Results from these simplified models can then be reinterpreted using models generated with the full spectrum of SUSY particles, like the phenomenological minimal supersymmetric Standard Model (pMSSM) [106, 107].

A key assmuption in these simplified models is that only the targeted final state particles can be produced, with all other SUSY particles decoupled at high mass. Sleptons must decay into the same flavour of lepton and mixing is forbidden. Squarks must decay into quarks of the same generation, with no mixing allowed for

[4] This is done for simplicity rather than due to a physical motivation.

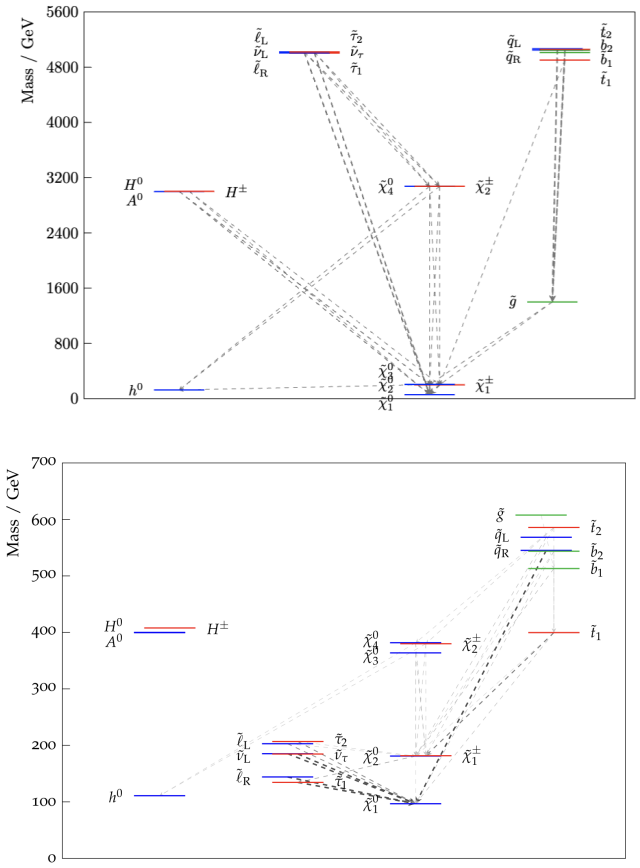


Figure 4.4: Two example MSSM mass spectra, highlighting the large range of possible masses.

the first and second generation squarks. In the case of the charginos and neutralinos, various assumptions can be made on the mass degeneracy of these states, and this determines the targeted SUSY model spectra. Additionally, we assume R-parity conservation and so the $\widetilde{\chi}_1^0$ is stable, and is assumed to be 100% Bino-like.

Different simplified models can be created to constrain a specific sparticle decay mode. Usually these target a specific detector signature, often $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the necessary precence of $\widetilde{\chi}_1^0$'s, in conjunction with lepton or jet requirements. The main process of interest in this thesis is slepton pair production, Figure 4.5, where the sleptons both decay the same flavour SM lepton and the $\widetilde{\chi}_1^0$. In the detector, the signature of this process is two same flavour, oppositely charged leptons, with large $E_{\mathrm{T}}^{\mathrm{miss}}$.

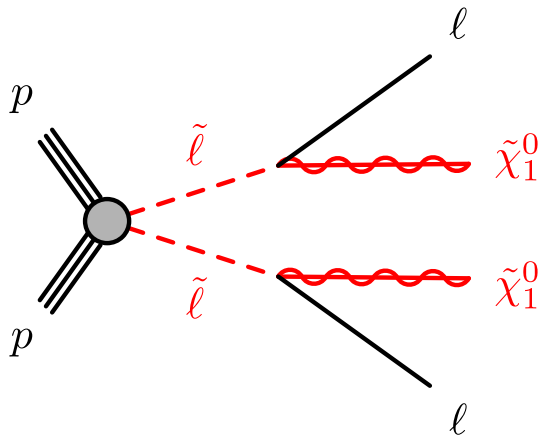


Figure 4.5: Diagram for direct slepton production at the the LHC.

Slepton Assumptions

The analysis in Chapter 5 targets direct slepton pair production, in Figure 4.5. Throughout the rest of this thesis, and generally in ATLAS SUSY searches, 'slepton' refers to either selectron or smuon. While stau's could also be produced and decay via the same processes, the comparatively tiny branching ratio for fully leptonic di-stau decays means that there is currently no sensitivity when targeting the di-leptonic final state. [5] For sleptons, various choices are available regarding the assumed degeneracy of smuons and selectrons [6]. Seven different interpretations are available:

- Four-fold slepton mass degenerate, where we have one interpretation:
  - $m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$
- Two-fold slepton mass degenerate, where we have two interpretations:
  - $m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}})$
  - $m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$
- No slepton mass degeneracy assumed, where we have four interpretations:
  - $m(\tilde{e}_{\mathrm{L}})$ only

[5] Potential sensitivity to stau's was evaluated in the analysis covered in Chapter 5, and seen to be negligible.

[6] It is worth noting, the first of these is not particularly physically motivated.

- $m(\tilde{\mu}_{\mathrm{L}})$ only
- $m(\tilde{e}_{\mathrm{R}})$ only
- $m(\tilde{\mu}_{\mathrm{R}})$ only

Four-fold slepton mass degeneracy is often assumed for simplicity, as the reported sensitivity is the largest, but all interprations are relevant for setting limits on various slepton interprations. Details on how the signal models are generated in MC simulation is given in Section 5.2.3.

4.4 Experimental Searches for SUSY

Extensive effort has been expended to search for SUSY particles in collider experiments like LEP and the LHC. This section will briefly cover the current status of R-parity conserving SUSY searches in ATLAS and CMS. A full overview from the Particle Data Group can be found in Ref. [39].

The cross-section of SUSY particles depends on the mass and coupling, with the size of cross-sections varying significantly across the SUSY spectrum. Figure 4.6, shows the cross-sections as a function of mass, for a range of sparticles. The sparticles produced through strong processes, like the stop and gluino, have larger cross-sections due to the larger coupling. As a result, strong constraints, above one TeV, have been placed on the possible masses of these 'strong SUSY' particles like the stop, at collider experiments. Electroweak sparticles, like the sleptons and the electroweakino's, have smaller couplings and are less constrained by experimental results, making them an interesting target to search for at the LHC.

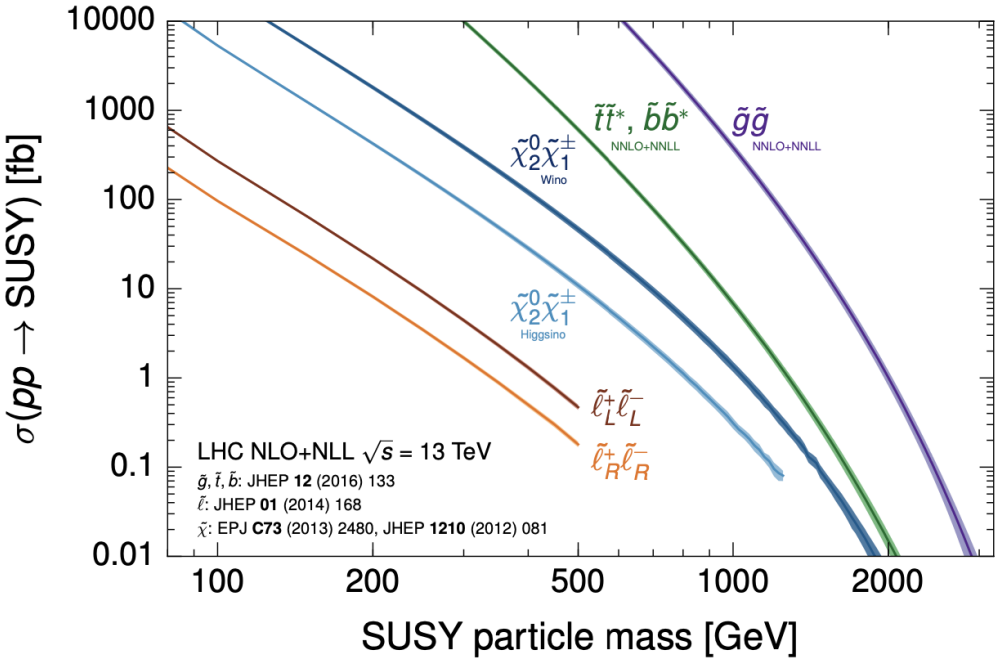


Figure 4.6: Cross-sections from simplified models for pair production of the: gluino, stop, sbottom, and separately for left and right handed sleptons. The wino only and higgsino only cross-sections for neutralino plus chargino production are also shown. The cross-sections are calculated for pp collisions at $\sqrt{s} = 13$ TeV. The width of lines indicates the uncertainty. The figure is taken from Ref. [108], and interpolates cross-sections calculated using Refs. [109–113].

4.4.1 Strong SUSY

The strong SUSY sector has been significantly constrained by ATLAS and CMS due to the relatively high cross-sections at the LHC. Naturalness arguments mean that the $\tilde{t}_1$ should be around 1 TeV, hence there is strong motivation to search for this particle as it could be in reach of the LHC. Figure 4.7, shows the existing stop and gluino limits from the ATLAS experiment. In the figure, the assumptions made in the various strong SUSY signal models varys. For example, the yellow signal model has the Gravitino as the LSP, whereas the majority of other models have the $\tilde{\chi}_1^0$. The $\tilde{t}_1$ is excluded up to 1.2 TeV, for massless neutralinos, and mostly excluded up to 600 GeV for a range of neutralino masses. Similarly gluinos up to 2.4 TeV are excluded for massless neutralinos, and almost entirely excluded up to 2.2 TeV.

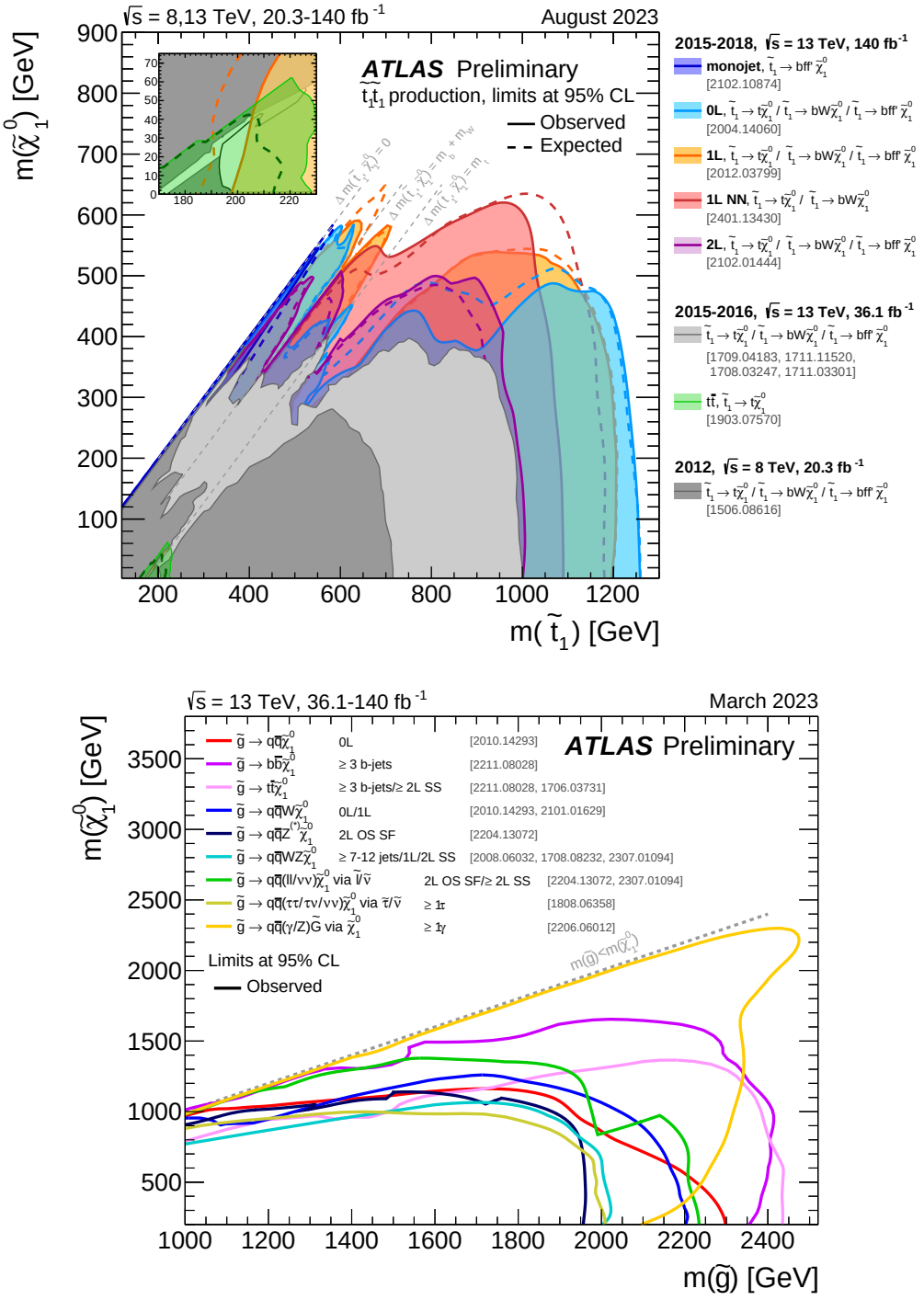


Figure 4.7: Exclusion limits from ATLAS, for simplified SUSY models of (a) stop production and (b) gluino prodction. In (a) the dotted lines indicate the expected limits and the filled areas indicate the observed. In (b) the lines indicate the observed limits. All limits are computed at 95% CL, and the grey dotted line indicates the kinematically forbidden region where $m(\tilde{t}_1)/m(\tilde{g}) < m(\tilde{\chi}_1^0)$. Figures taken from [114].

4.4.2 Electroweak SUSY

Moving towards the lower cross-section electroweak sector, Figure 4.8 shows the ATLAS exclusion limits for Higgsinos ($\tilde{\chi}^{\pm,0}_{1,2}$ and $\tilde{\chi}^0_2$). The $\tilde{\chi}^{\pm,0}_{1,2}$ and $\tilde{\chi}^0_2$ are constrained by ATLAS up to 1 TeV for a massless $\tilde{\chi}^0_1$, and largely excluded for $\tilde{\chi}^0_1$ up to 200 GeV, assuming $m_{\tilde{\chi}^{\pm,0}_{1,2}} = m_{\tilde{\chi}^0_2}$. An important feature to note is that signal models with a small mass splitting, $\Delta m(\tilde{\chi}^\pm_1, \tilde{\chi}^0_1)$, commonly referred to as models with 'compressed' mass spectra, are not well covered by current searches. This is because the final states produced have very low energy, making this region very hard to tackle. These compressed scenarios are of increasing interest in ATLAS SUSY searches as they are well motivated by naturalness and have large uncovered regions for electroweakly produced sparticles.

Sleptons are less strongly constrained than the strong SUSY particles and the Higgsinos, as they have a relatively small cross-section in comparison. The LEP experiments placed constraints on the $\tilde{e}_\mathrm{R}$ and $\tilde{\mu}_\mathrm{R}$ masses, excluding $m(\tilde{e}_\mathrm{R}) < 99.9\,\mathrm{GeV}$ and $m(\tilde{\mu}_\mathrm{R}) < 95\,\mathrm{GeV}$, provided that $\Delta m(\tilde{\ell}, \tilde{\chi}^0_1) \gtrsim 2$ GeV [115]. A summary of the ATLAS results, preceeding the work in this thesis, are shown in Figure 4.9, assuming four-fold slepton mass degeneracy. The results are typically presented in both the $m(\tilde{\ell})$ – $m(\tilde{\chi}^0_1)$ and $m(\tilde{\ell})$ – $\Delta m(\tilde{\ell}, \tilde{\chi}^0_1)$ planes. No significant deviations from the SM are seen in any of these ATLAS slepton interpretations.

ATLAS first performed searches for slepton production using lepton triggers [116, 117] (the purple and pink regions in Figure 4.9 respectively) that achieved lower limits on $m(\tilde{\ell})$ of up to 700 GeV when assuming four-fold mass degeneracy. These limits fall to the LEP limit at $\Delta m(\tilde{\ell}, \tilde{\chi}^0_1) \approx 80$ GeV, and so to gain sensitivity in the regions with smaller $\Delta m(\tilde{\ell}, \tilde{\chi}^0_1)$, a dedicated ATLAS search targeting $\Delta m(\tilde{\ell}, \tilde{\chi}^0_1) \approx m(W)$ was performed [118], (the green region in Figure 4.9). However, the limits from this search decrease at lower $\Delta m(\tilde{\ell}, \tilde{\chi}^0_1)$, and in some regions are unable to improve on LEP.

To gain sensitivity to this region with smaller $\Delta m(\tilde{\ell}, \tilde{\chi}^0_1)$, other searches were performed at ATLAS triggering on the $E^\mathrm{miss}_\mathrm{T}$ signature, allowing the selection of lower-p_T leptons [119], (the blue region in Figure 4.9). Limits were placed on $m(\tilde{\ell})$ in the range of $0.5\,\mathrm{GeV} \lesssim \Delta m(\tilde{\ell}, \tilde{\chi}^0_1) \lesssim 30\,\mathrm{GeV}$, with a Δm-dependent bound on

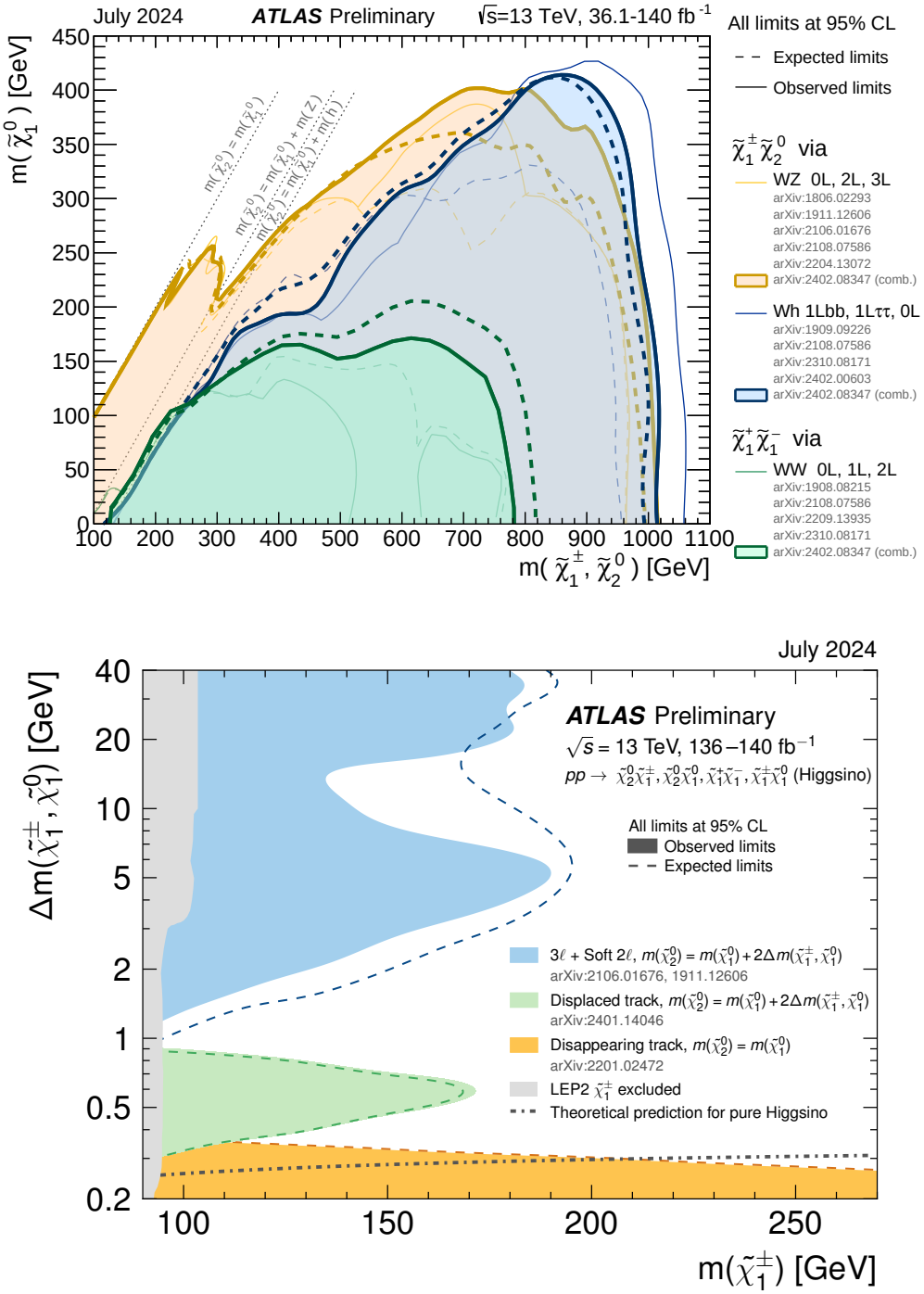


Figure 4.8: Exclusion limits from ATLAS, for simplified SUSY models of Higgsino production. In (a), the $\tilde{\chi}_{1,2}^{\pm,0}$, $\tilde{\chi}_2^0$ are assumed to have the same mass, and results are shown in the $m_{\tilde{\chi}_1^\pm,\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$ plane. In (b), the compressed Higgsino region is shown, and results are shown in the $m_{\tilde{\chi}_1^\pm} - \Delta m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$ plane. Varying assumptions on Higgsino mass ordering are in this figure and these are indicated in the corresponding legend entry. The dotted lines indicate the expected limits and the filled areas indicate the observed. All limits are computed at 95% CL, and in (a) the grey dotted line indicates the kinematically forbidden regions. Figures taken from [114].

individual slepton masses that ranged from just larger than the LEP limits, up to a maximum of $\approx 240\,\text{GeV}$, for $m(\tilde{\ell})$ at $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \approx 10\text{GeV}$.

The CMS Collaboration has also published results evaluating sensitivity to sleptons, and a summary of these is in Figure 4.10, assuming four-fold degeneracy. For $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \gtrsim m(W)$, bounds were set for values of $m(\tilde{\ell}) \lesssim 700\,\text{GeV}$ [120]. To explore the more difficult region with $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \lesssim m(W)$, the CMS Collaboration reinterpreted their soft di-lepton electroweakino search [121] as a slepton search [122], and obtained limits to exclude up to $m(\tilde{\ell}) \lesssim 215$ GeV for $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \approx 5$ GeV, with limits for larger mass differences reducing to that of LEP at $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \approx 20$ GeV. No significant deviations from the SM are seen in the slepton interpretations for any of these CMS results either.

In addition, it is well motivated to relax the assumption of four-fold degeneracy, and look at individual selectron and smuon exclusion. Figure 4.11, shows the existing ATLAS limits for selectrons and smuons individually, assuming two-fold mass degeneracy in the $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$ plane. In the region with $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \lesssim m(W)$, in the case of individual two-fold mass degenerate selectron and smuon exclusion, the limits drop significantly compared to the four-fold limits. For the intermediate region, approximately 20 GeV $\lesssim \Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \lesssim 40$ GeV, improvements in sensitivity have not been obtained since LEP. For $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \gtrsim m(W)$, beyond the bounds of these figures, the lower limit reduces to $550\,\text{GeV}$ for the $m(\tilde{\mu}_\text{L}) = m(\tilde{\mu}_\text{R})$ or $m(\tilde{e}_\text{L}) = m(\tilde{e}_\text{R})$ interpretations, well below the four-fold degenerate limits.

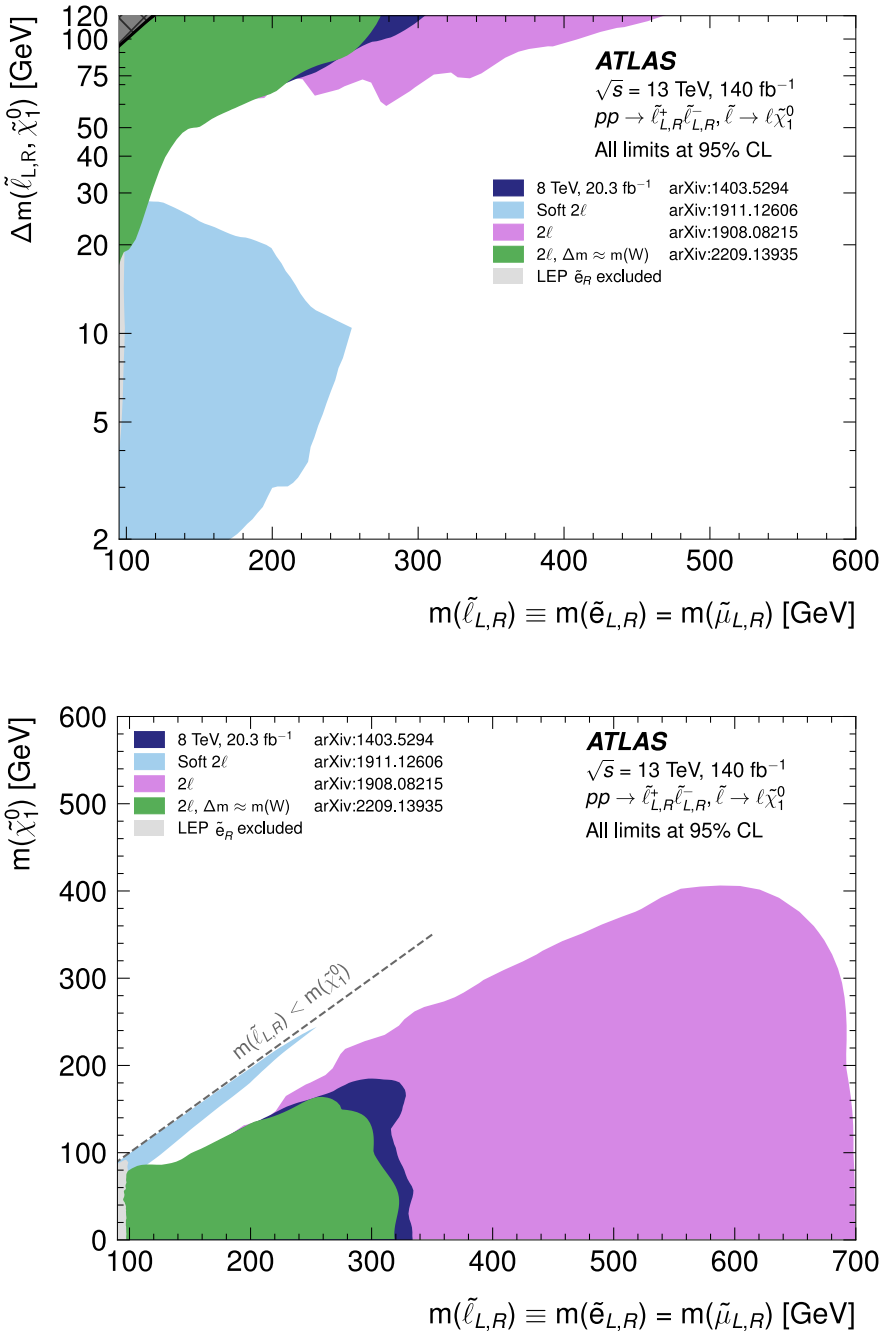


Figure 4.9: Exclusion limits from ATLAS, for simplified SUSY models of direct slepton production assuming four-fold mass degeneracy: $m(\tilde{\ell}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$. The limits are shown in (a) the $m(\tilde{\ell})$ – $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ plane and (b) the $m(\tilde{\ell})$ – $m(\tilde{\chi}_1^0)$ plane. All limits are computed at 95% CL, and the black delineated grey region indicates the kinematically forbidden region where $m(\tilde{\chi}_1^0) < 0$. The observed limits obtained at LEP [115], along with previous searches at the ATLAS experiment [116–119] are shown.

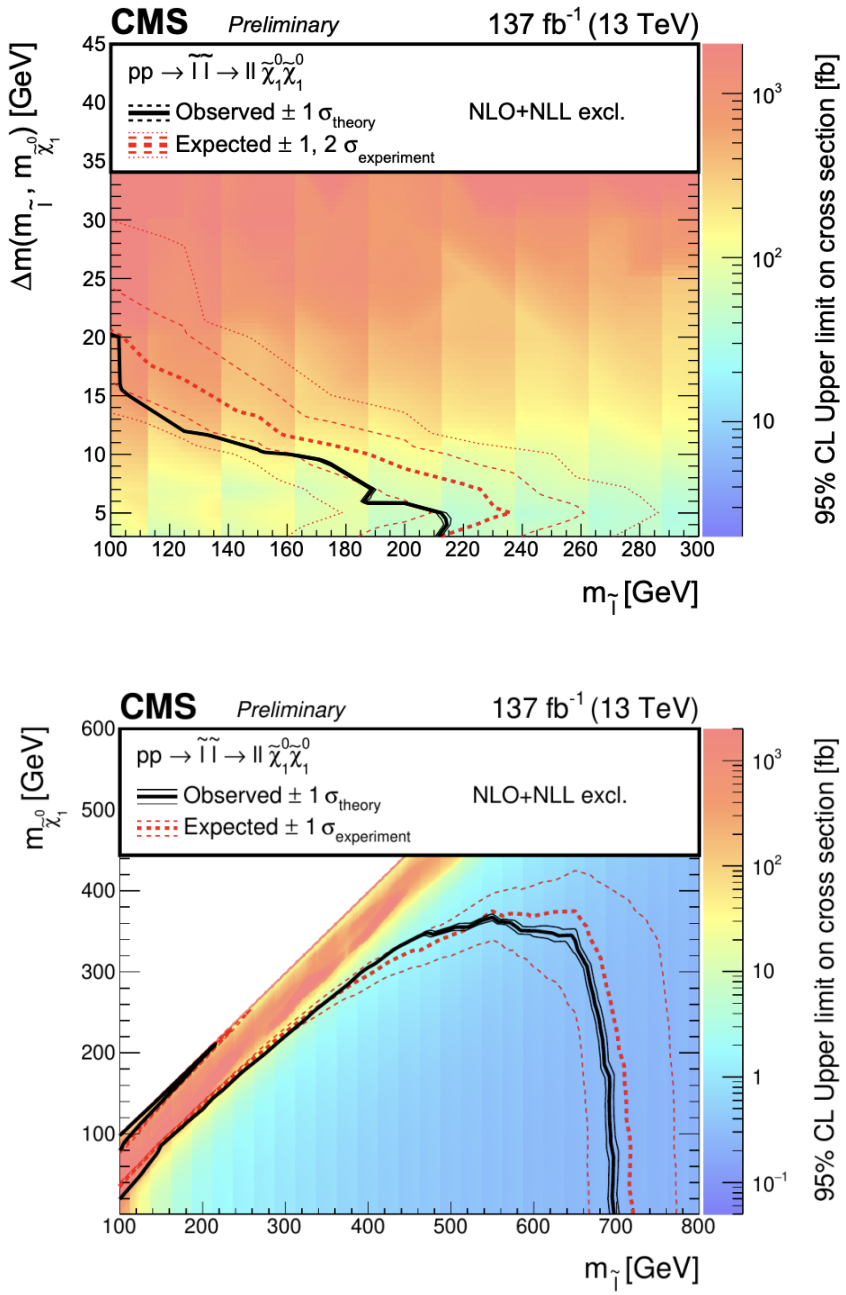


Figure 4.10: The observed and expected exclusion limits from CMS, obtained from Refs. [120, 122], are shown for simplified SUSY models of direct slepton production, assuming four-fold mass degeneracy: $m(\tilde{\ell}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$. The limits are shown in (a) the $m(\tilde{\ell})$ – $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ plane and (b) the $m(\tilde{\ell})$ – $m(\tilde{\chi}_1^0)$ plane. All limits are computed at 95% CL, and in (b) the white filled region indicates the kinematically forbidden region where $m(\tilde{\ell}_{L,R}) < m(\tilde{\chi}_1^0)$. The thickest red dotted lines are the expected limits, and the slightly thinner dotted lines corresponds to the $\pm 1\,\sigma$ variations of the expected limit, accounting for all uncertainties aside from the theoretical uncertainties in the signal cross-section. The thinnest dotted red lines are the $\pm 2\,\sigma$ variations of the expected limit. The thick black solid lines are the observed exclusion limits, and the thinner black lines indicate the variation as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. The colour scale indicates the upper limit on the slepton cross section across the range signal space.

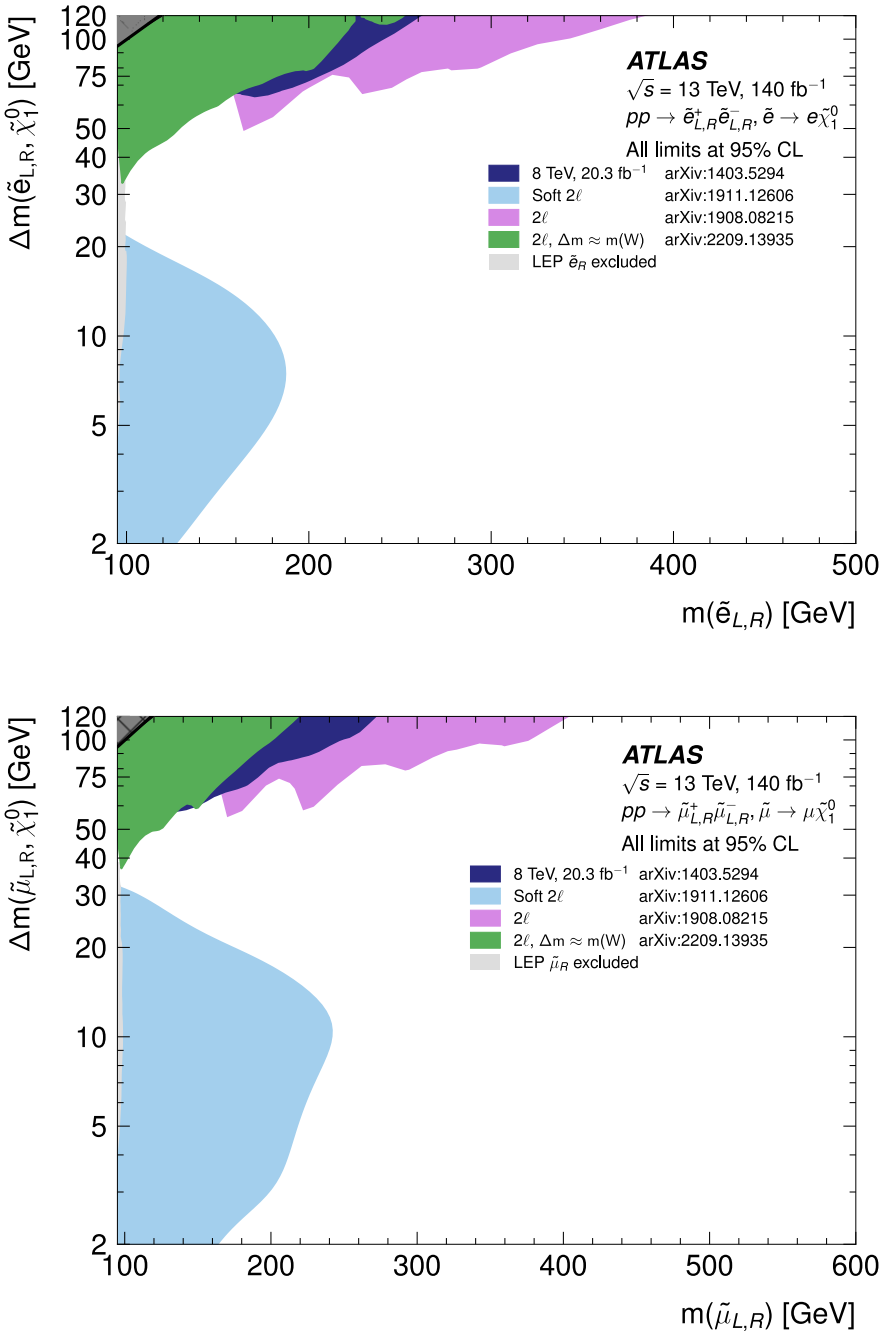


Figure 4.11: Exclusion limits from ATLAS, for simplified SUSY models of direct slepton production assuming two-fold mass degeneracy, (a) $m(\tilde{\ell}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}})$ or (b) $(\tilde{\ell}) \equiv m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$. The limits are shown in (a) the $m(\tilde{e}_{\mathrm{L,R}})$ – $\Delta m(\tilde{e}_{\mathrm{L,R}}, \tilde{\chi}^{0}_{1})$ plane and (b) the $m(\tilde{\mu}_{\mathrm{L,R}})$ – $\Delta m(\tilde{\mu}_{\mathrm{L,R}}, \tilde{\chi}^{0}_{1})$ plane. All limits are computed at 95% CL, and the black delineated grey region indicates the kinematically forbidden region where $m(\tilde{\ell}_{L,R}) < m(\tilde{\chi}^{0}_{1})$. The observed limits obtained at LEP [115], along with previous searches at the ATLAS experiment [116–119] are shown.

4.5 Summary

From these previous ATLAS and CMS results, it is clear that there exists an important gap in sensitivity for $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \leq m(W)$, in the approximate range of mass differences between 20 GeV to 60 GeV. This region has proved sufficiently difficult that, for a substantial range of $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$, the LHC experiments have not yet gained sensitivity beyond LEP for the two-fold mass degenerate interpretations. This region where the slepton mass spectrum is compressed, is challenging to probe due to the similarity of the signal model kinematics to the large W^+W^- background, as well as the presence of low-p_{T} final-state leptons. The next chapter details efforts to target this region of compressed mass spectrum slepton models, using the Run 2 ATLAS dataset and machine learning techniques.

Carpe diem. Seize the day, boys. Make your lives extraordinary.

Dead Poet's Society

5 An ATLAS Run 2 Search for Sleptons

Contents

5.1 Analysis Strategy 72
5.2 Datasets, MC samples and Objects 74
5.2.1 Dataset and Triggers 74
5.2.2 Background MC Samples 75
5.2.3 Signal MC Samples 75
5.2.4 Object Definitions 76
5.3 Event Selection 78
5.3.1 Kinematic Variables 78
5.3.2 Preselection Requirements 82
5.3.3 BDT Training 85
5.3.4 Signal Region Definitions 96
5.4 Background Estimation 98
5.4.1 Irreducible Backgrounds 98
5.4.2 Reducible Backgrounds 99
5.4.3 Control and Validation Regions 102
5.5 Statistical Methods 106
5.5.1 Fit types 106
5.5.2 Toy Pseudo-experiments 108
5.6 Systematic Uncertainties 109
5.6.1 Experimental Uncertainties 109
5.6.2 Theoretical Uncertainties 111
5.7 Results 117
5.7.1 Control Regions 117
5.7.2 Validation Regions 118
5.7.3 Signal Regions 119
5.7.4 Model Independent SR Significances 122
5.7.5 SR-BDT$^{\mu\mu}_{5+10}$ Investigations 123
5.7.6 SR-BDT$^{\mathrm{ee}}_{40+50}$ Investigations 131

5.8 Interpretations 132
5.8.1 Exclusion Limits 133
5.8.2 Signal Model Discovery Significances 141
5.9 The Wider Context 143
5.10 Summary 151

This chapter presents the main analysis of this thesis: a search for sleptons using the ATLAS Run 2 dataset. Previous results from ATLAS and CMS have shown a gap in sensitivity for $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \leq m(W)$, roughly in the range of 20 GeV to 60 GeV. The analysis specifically targets the 'compressed' region, also refered to as the 'corridor', where $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$ is small. The analysis provides sensitivity to large areas of this tricky to target region for the first time. In addition, some interesting deviations between the observed yields and Standard Model predictions were seen, which merit further investigation with the Run 3 dataset.

Two different approaches are taken in the ATLAS analysis: the 'cut-and-count' based approach, and the Boosted Decision Tree (BDT) approach. The focus of this chapter will be the BDT approach, as this is the approach I worked on and provides the best sensitivity to the simplified slepton models. A comparison of the limits achieved using the two approaches will be presented in Section 5.8, but the alternative approach will not be covered in detail [1]. At the end of the chapter, the results of this analysis will be placed in the context of wider ATLAS and CMS electroweak SUSY searches and other existing excesses.

5.1 Analysis Strategy

The searches reported in this chapter select events with a pair of same-flavour electrons or muons with opposite electric charge. They require the presence of a high-p_{T} jet from initial-state radiation (ISR) to boost the neutralinos in the transverse plane. This results in enough missing transverse momentum to trigger the events, removing the need to trigger on leptons, hence extending sensitivity to smaller $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$. Such ISR configurations also boost the parent particles, which

[1] Further information on the cut-and-count approach can be found in Ref. [1].

helps distinguish the kinematics from those of the W^+W^- background, as the sleptons are more massive than the bosons [123]. Figure 5.1 sketches the effect of the ISR boost on the decay products. This increases the acceptance and efficiency for leptons from slepton decays with $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0) \lesssim m(W)$. A diagram showing the slepton pair production mode targeted in this analysis is shown in Figure 5.2.

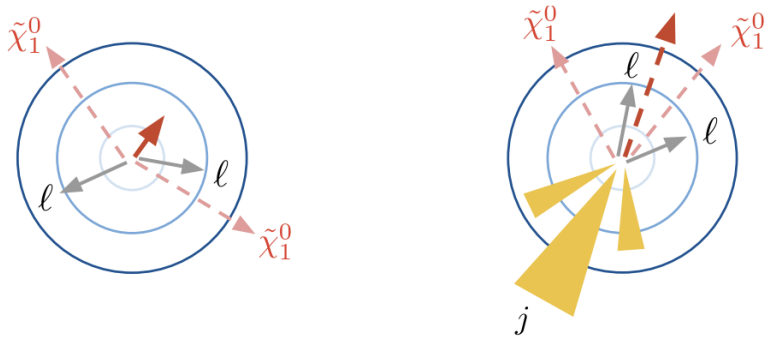


Figure 5.1: Sketch illustrating the signal event topologies (left) before and (right) after requiring an ISR jet (j). The yellow triangles are jets. The dark red arrow indicates the reconstructed missing transverse momentum vector and the relative length indicates its magnitude. Taken from [108].

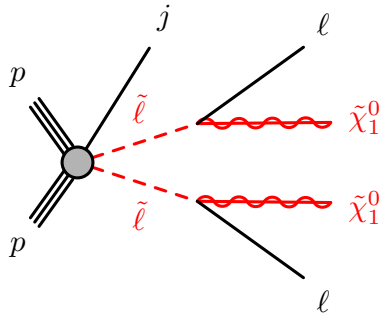


Figure 5.2: Diagram representing the direct production of sleptons ($\tilde{\ell} \in \tilde{e}, \tilde{\mu}$) decaying into their corresponding SM lepton partner ($\ell \in e, \mu$) and the lightest neutralino ($\tilde{\chi}_1^0$), in association with a jet (j) from initial-state radiation.

BDTs are trained on slepton signal samples and on the SM backgrounds, with each BDT targeting signals within a relatively narrow range of $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$, and able to exploit a larger number of kinematic variables and their correlations, improving sensitivity when compared with the cut-and-count-based analysis. In combination, training multiple BDTs accounts for differences in the signal kinematics that are

driven by $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$, thereby providing wide-ranging discrimination between the signal and the backgrounds across the gap region.

Signal regions (SRs) are defined using the BDT score, and optimised based on the expected sensitivity for the simplified supersymmetric models. The background estimate in the SRs constrained using statistical fits between MC predictions and data in orthogonal control regions (CRs). Then, only after extensive validation of this background estimate in validation regions (VRs), the data in the SRs is checked (unblinded), and the statistical significance of any deviation between the data and MC prediction can be evaluated. Exclusion limits are set using these SRs, at 95% CL using the CL_s method [124], on the seven different slepton interpretations detailed in the previous chapter. Sensitivity to supersymmetric models across this corridor region is found.

5.2 Datasets, MC samples and Objects

This search uses $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS experiment across Run 2 of the LHC. The SM backgrounds are modelled using MC simulation and all di-leptonic SM processes are considered. Simplified SUSY signal models are also simulated, and these are used to optimise the analysis regions.

5.2.1 Dataset and Triggers

After the application of data quality requirements [125] results, the ATLAS Run 2 dataset corresponds to an integrated luminosity of 140 fb^{-1} with an uncertainty of 0.83% [7]. The data events used in this search were collected using the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers [126] with varying thresholds depending on the data-taking periods. The triggers are found to be $> 95\%$ efficient in all data-taking periods for $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV, with further details on this efficiency measurement given in Chapter 6. Details of the primary Run 2 $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used in this analysis can be found in Chapter 6.

5.2.2 Background MC Samples

The SM backgrounds involving the production of prompt leptons are modelled with a variety of different MC generators. Table 5.1 provides an overview of all SM background samples utilised in the analysis. It details the generators, the PDF sets, the parameter tunes used for the parton shower, and the perturbative order in the strong coupling constant, α_s, used for matrix element calculations.

In the analysis, the $t\bar{t}$ and single top processes are grouped together to in the 'top' category. Diboson is the other major background category and all other minor background MC samples are grouped into the 'other' category.

The fake/non-prompt (FNP) lepton background is estimated using data-driven methods, as these are not well modelled through MC simulation. FNP leptons originate from hadronic activity in the detector, misidentified as leptons, and non-prompt leptons originate from secondary decays of, for example B hadrons. Further details on the use of the data-driven methods used in this analysis will be given in Section 5.4.2.

Table 5.1: Simulated background event samples with the corresponding matrix element and parton shower (PS) generators, cross-section order in α_s used to normalise the event yield, underlying-event tune and the generator PDF sets used. For diboson, triboson and $t\bar{t}+V$ samples, $V \in \{W, Z\}$. Diboson samples also include Higgs boson contributions. 'Default' refers to the default tune of the Sherpa generator. Abbreviations used are defined as: leading-order (LO), next-to-leading-order (NLO), next-to-next-to-leading-order (NNLO), next-to-leading-logarithmic (NLL), next-to-next-to-leading-logarithmic (NNLL).

Physics process	Generator	Parton shower	Normalisation	Tune	PDF (generator)	PDF (PS)
$t\bar{t}$	Powheg Box v2	Pythia 8.230	NNLO+NNLL	A14	NNPDF3.0nlo	NNPDF2.3lo
Single-top	Powheg Box v2	Pythia 8.230	NLO+NNLL	A14	NNPDF3.0nlo	NNPDF2.3lo
Diboson VV	Sherpa 2.2.2, 2.2.11, 2.2.12	Sherpa 2.2.2, 2.2.11, 2.2.12	LO–NLO	Default	NNPDF3.0nnlo	NNPDF3.0nnlo
Triboson VVV	Sherpa 2.2.2	Sherpa 2.2.2	NLO	Default	NNPDF3.0nnlo	NNPDF3.0nnlo
$t\bar{t}+V$	MadGraph5_aMC@NLO 2.3.3	Pythia 8.210	NLO	A14	NNPDF3.0nlo	NNPDF2.3lo
$t\bar{t}+\gamma$	MadGraph5_aMC@NLO 2.3.3	Pythia 8.210	NLO	A14	NNPDF2.3lo	NNPDF2.3lo
$t\bar{t}+H$	Powheg Box v2	Pythia 8.230	NLO	A14	NNPDF3.0nlo	NNPDF2.3lo
$t\bar{t}+WW$	MadGraph5_aMC@NLO 2.2.2	Pythia 8.186	NLO	A14	NNPDF2.3lo	NNPDF2.3lo
$t\bar{t}+WZ$	MadGraph5_aMC@NLO 2.3.3	Pythia 8.212	NLO	A14	NNPDF2.3lo	NNPDF2.3lo
$tZ, tWZ, t\bar{t}t\bar{t}, t\bar{t}t$	MadGraph5_aMC@NLO 2.3.3	Pythia 8.230	NLO	A14	NNPDF3.0nlo	NNPDF2.3lo
$Z/\gamma^*(\to \ell\ell, \tau\tau)$+jets	Sherpa 2.2.11, 2.2.14	Sherpa 2.2.11, 2.2.14	NNLO	Default	NNPDF3.0nnlo	NNPDF3.0nnlo
ggF $H \to \tau\tau, \mu\mu$	Powheg Box v2	Pythia 8.2	NNLO	AZNLO	PDF4LHC15nnlo	PDF4LHC15nnlo
VBF $H \to \tau\tau$	Powheg Box v2	Pythia 8.2	NNLO	AZNLO	PDF4LHC15nlo	PDF4LHC15nlo
ZH, WH	Powheg Box v2	Pythia 8.2	NNLO	AZNLO	PDF4LHC15nlo	PDF4LHC15nlo

5.2.3 Signal MC Samples

Simulated signal samples consisting of direct selectron – ($\tilde{e}_{\mathrm{L,R}}$), and smuon – ($\tilde{\mu}_{\mathrm{L,R}}$) pair production are used to optimise the event selection and interpret the results.

During event generation, simplified models are generated assuming that selectrons and smuons are four-fold mass-degenerate: $m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$. The sleptons are required to decay into their corresponding SM lepton partner and a purely bino-like neutralino LSP, $\tilde{\chi}_1^0$, with a branching ratio of 100%. For interpretations where the selectron-smuon mass degeneracy is lifted, no dependence of the event kinematics on the slepton flavour and chirality is expected, and so the $m(\tilde{e}_{\mathrm{L}})$, $m(\tilde{\mu}_{\mathrm{L}})$ and $m(\tilde{e}_{\mathrm{R}})$, $m(\tilde{\mu}_{\mathrm{R}})$ cross-sections are simply rescaled according to the kinematically accessible SUSY states. The events are produced at LO in α_{s} using MadGraph5_aMC@NLO 2.7.3 [127] with the NNPDF2.3lo PDF set and include up to two additional partons in the matrix element. Pythia 8.244 [128] is used to model the slepton decays, parton shower, hadronisation, and underlying event with the A14 set of tuned parameters [129]. The decays of bottom and charm hadrons are performed by EvtGen 1.7.0. The signal production cross-sections and uncertainties are calculated using Resummino 2.0.1 at NLO+NLL precision [109, 110, 112, 113, 130–133]. As an example, the combined production cross-sections for mass-degenerate selectrons and smuons with a mass of 200 GeV are $\sigma(\tilde{\ell}_{\mathrm{L}}) = 43.9 \pm 0.8$ fb and $\sigma(\tilde{\ell}_{\mathrm{R}}) = 16.7 \pm 0.4$ fb.

5.2.4 Object Definitions

The reconstruction of physics objects for this analysis is carried out according to the ATLAS recommendations, as described in Chapter 3. For the leptons and jets, distinctions are made between objects with looser (baseline), and tighter selection (signal). The baseline objects are used in the overlap removal procedure and Fake/Non-prompt lepton estimate. The signal objects are the definitions used in the final analysis regions.

Table 5.2 shows the object definitions for the leptons and jets. The leptons have impact parameter (d_0, $z_0 \sin\theta$) requirements in order to reduce contamination from fake leptons and leptons from secondary decays. The lepton p_{T} requirements are set to the lowest values that ATLAS reconstructs electrons and muons down to as the signal models targeted produce soft (low p_{T}) leptons.

In addition, the standard ATLAS overlap removal prescription and $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is used. Only jets, electrons, muons and photons [2] passing the baseline selection criteria are considered for the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The jets used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation have additional requirements, as detailed in Section 3.3.7 to reduce the contamination from pileup activity.

Table 5.2: Summary of the definitions for baseline and signal objects used in the analysis.

Property	Signal	Baseline
Electrons		
Kinematic	$p_{\mathrm{T}} > 4.5$ GeV, $\lvert\eta\rvert < 2.47$	
Identification	MediumLLH	LooseAndBLayerLLH
Isolation	PLVLoose	–
Impact parameter	$\lvert d_0/\sigma(d_0)\rvert < 5$, $\lvert z_0 \sin\theta\rvert < 0.5$ mm	$\lvert z_0 \sin\theta\rvert < 0.5$ mm
Muons		
Kinematic	$p_{\mathrm{T}} > 3$ GeV, $\lvert\eta\rvert < 2.7$	
Identification	Medium	
Isolation	PLVLoose	–
Impact parameter	$\lvert d_0/\sigma(d_0)\rvert < 3$ & $\lvert z_0 \sin\theta\rvert < 0.5$ mm	$\lvert z_0 \sin\theta\rvert < 0.5$ mm
Jets (Anti-k_t $R = 0.4$ PFlow)		
Kinematic	$p_{\mathrm{T}} > 30$ GeV, $\lvert\eta\rvert < 2.8$	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 4.5$
Pileup mitigation	JVT Tight for $p_{\mathrm{T}} < 60$ GeV, $\lvert\eta\rvert < 2.4$	
b-Jets (Anti-k_t $R = 0.4$ PFlow)		
Kinematic	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$	
Pileup mitigation	JVT Tight for $p_{\mathrm{T}} < 60$ GeV, $\lvert\eta\rvert < 2.4$	
b-tagging	DL1r FixedCutBeff 77%	

[2] There is no particular motivation for using photons in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, as the photons would anyway be considered as jets in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation (like tau's are). It just makes this analysis consistent with previous efforts.

5.3 Event Selection

In the analysis, there are different stages in the event selection, that will be covered in this section. Some acronyms are frequently used to describe the different lepton selections:

- SFOS – two same flavour opposite sign leptons (e^+e^- or $\mu^+\mu^-$)
- DFOS – two different flavour opposite sign leptons ($e^+\mu^-$ or μ^+e^-)
- SS – two (any flavour) same sign leptons (e^+e^+, e^-e^- or $\mu^+\mu^+$, $\mu^-\mu^-$)

A training preselection is defined based on the targeted final state, is used to train five separate BDTs, carrying out binary classification to discriminate between signal and background events. The BDTs are trained using different combinations of signal models, grouped by $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$. The output BDT scores are then used to define three opposite-charge di-electron and three opposite-charge di-muon signal regions (SRs).

5.3.1 Kinematic Variables

A number of kinematic variables are utilised as inputs to the trained BDTs, and in defining the analysis regions. A $^{(\dagger)}$ superscript indicates a variable which is used in the BDT training. The definitions of these are given in this section. The leading lepton (jet) is defined as the lepton (jet) with the largest p_{T} in the event. Similarly, in cases where two lepton flavours are used to described region selections, eg: $e\mu$, the leptons are given in order of p_{T}.

- The leading ($p_{\mathrm{T}}^{\ell_1\,(\dagger)}$) and subleading ($p_{\mathrm{T}}^{\ell_2\,(\dagger)}$) lepton transverse momenta.
- The invariant mass of the two leptons, $m_{\ell\ell}{}^{(\dagger)}$.
- The angular distance between the two leptons, $\Delta R_{\ell\ell}{}^{(\dagger)}$.
- The modulus of the vector sum of the $\mathbf{p}_{\mathrm{T}}$ of the two leptons, $p_{\mathrm{T}}^{\ell\ell\,(\dagger)}$.
- The p_{T} of the leading ($p_{\mathrm{T}}^{j_1(\dagger)}$) and subleading ($p_{\mathrm{T}}^{j_2(\dagger)}$) jet. For events without a second jet, the latter is set to -1.

- The number of signal jets with $p_{\mathrm{T}} > 30$ GeV, N_{jet}^{30}.
- The number of b-tagged jets with $p_{\mathrm{T}} > 20$ GeV, $N_{b-\mathrm{jet}}^{20}$.
- The magnitude of the missing transverse momentum vector, $E_{\mathrm{T}}^{\mathrm{miss}\,(\dagger)}$.
- The azimuthal angle between the leading jet and the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(j_1, E_{\mathrm{T}}^{\mathrm{miss}})^{\,(\dagger)}$.
- The minimum azimuthal angle between any signal jet and the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, $\min(\Delta\phi(\mathrm{jets}, E_{\mathrm{T}}^{\mathrm{miss}}))$.
- The azimuthal angle between $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ and the leading lepton $\Delta\phi(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})^{\,(\dagger)}$ and subleading lepton $\Delta\phi(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})^{\,(\dagger)}$.
- The object-based significance of the $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance$^{(\dagger)}$, which provides a measure of the likelihood that the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ originated from real invisible particles and not from detector effects [55].
- The angular variable $\cos\theta_{\ell\ell}^{*(\dagger)}$, which is a measure of the polar angle of one of the two leptons with the beam axis in the laboratory frame and is sensitive to the spin of the particles produced in the collision [134]. Similarly, $\cos\theta_{\ell\ell}^{V(\dagger)}$ measures the polar angle of one of the two leptons with the beam axis, evaluated in the dilepton centre of mass frame [135]. As sleptons are scalar particles, these variables exploit shape differences between the signal models and the SM backgrounds.
- An approximation of the invariant mass of a pair of leptonically decaying τ-leptons, denoted $M_{\tau\tau}{}^{(\dagger)}$, assuming that the neutrinos are collinear with their parent τ-lepton [136, 137]. This variable utilises the lepton p_{T} and the $E_{\mathrm{T}}^{\mathrm{miss}}$ to discriminate between backgrounds containing boosted $Z \to \tau\tau$ decays, for which $M_{\tau\tau}$ peaks near the Z boson mass, and the signal processes that do not feature such a peak.
- The transverse mass of the leading, $m_{\mathrm{T}}^{\ell_1\,(\dagger)}$, and subleading, $m_{\mathrm{T}}^{\ell_2\,(\dagger)}$, lepton, defined as:

$$m_{\mathrm{T}}(p_{\mathrm{T}}, q_{\mathrm{T}}, m_\chi) = \sqrt{m_\ell^2 + m_\chi^2 + 2\left(E_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{q} - \mathbf{p}_{\mathrm{T}} \cdot \mathbf{q}_{\mathrm{T}}\right)}, \tag{5.1}$$

where m_ℓ is the mass of the lepton, m_χ is the invisible particle mass (taken to be zero for the calculation of $m_{\mathrm{T}}^{\ell_1}$ and $m_{\mathrm{T}}^{\ell_2}$), and E_{T}^{ℓ} and E_{T}^{q} are defined as: $E_{\mathrm{T}} = \sqrt{m^2 + \mathbf{p}_{\mathrm{T}}^2}$, for the lepton and invisible particle respectively. The two vectors $\mathbf{p}_{\mathrm{T}}$ and $\mathbf{q}_{\mathrm{T}}$ are the transverse momentum vector of the lepton and the invisible particle $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ vector, respectively.

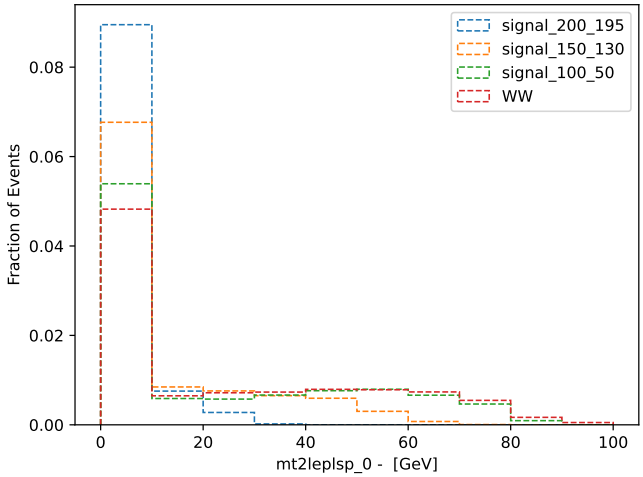
- The stransverse mass, m_{T2}, a kinematic variable that is analogous to m_{T}, but in the case of a pair of particles that have both decayed into a visible and an invisible particle [138, 139]. It is defined as:

$$m_{\mathrm{T2}}^{\chi}(\mathbf{p}_{\mathrm{T},1}, \mathbf{p}_{\mathrm{T},2}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}) = \min_{\mathbf{q}_{\mathrm{T},1}+\mathbf{q}_{\mathrm{T},2}=\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}} \left\{ \max \left[m_{\mathrm{T}}(\mathbf{p}_{\mathrm{T},1}, \mathbf{q}_{\mathrm{T},1}; m_\chi), m_{\mathrm{T}}(\mathbf{p}_{\mathrm{T},2}, \mathbf{q}_{\mathrm{T},2}; m_\chi) \right] \right\},$$

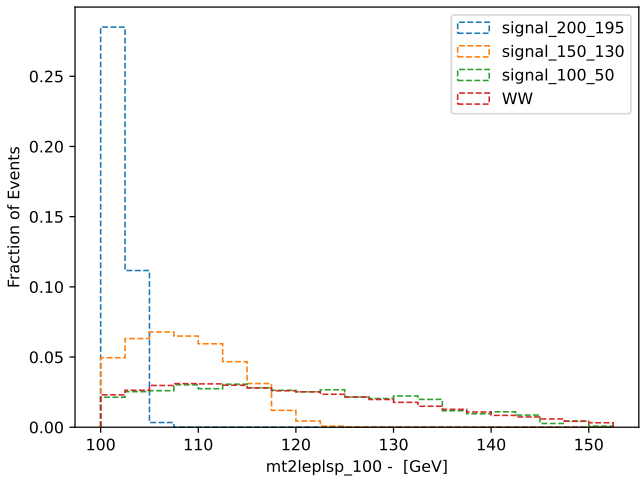
where m_{T} indicates the transverse mass, as defined above. The two vectors $\mathbf{p}_{\mathrm{T},1}$ and $\mathbf{p}_{\mathrm{T},2}$ are the transverse momentum vectors of the two leptons, and $\mathbf{q}_{\mathrm{T},1}$, $\mathbf{q}_{\mathrm{T},2}$ are vectors such that $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} = \mathbf{q}_{\mathrm{T},1} + \mathbf{q}_{\mathrm{T},2}$. The mass of the invisible particle, m_χ, is a free parameter of the equation. A χ mass of zero corresponds to the SM neutrino and this variant is useful to reject backgrounds from W^+W^- and $t\bar{t}$ decays. Non-zero values of m_χ are also chosen to reflect the signal model(s) with that invisible particle mass. A key feature of this variable is a kinematic endpoint, dependent on the true invisible particle mass of the event, shown in Figure 5.3. Since the signal and SM background have different kinematic endpoints, this provides significant discriminating power. In total, six different values of m_χ in the range 0 to 300 GeV are used: $m_{\mathrm{T2}}^{0\,(\dagger)}, m_{\mathrm{T2}}^{50(\dagger)}, m_{\mathrm{T2}}^{100(\dagger)}, m_{\mathrm{T2}}^{150(\dagger)}, m_{\mathrm{T2}}^{200(\dagger)}$ and $m_{\mathrm{T2}}^{300(\dagger)}$.

A seemingly trivial feature of m_{T2} is that the visible lepton masses are inputs to the calculation. Though it may seem like these are tiny in comparison to the energies involved at the LHC, in certain situations, m_{T2} distribution can become proportional to the lepton mass. This becomes problematic when the trained BDTs must be flavour-blind in order to use different-flavour lepton events for validating the background estimate. During the validation of the BDT training, a significant difference was observed between the electron and muon score distributions. This is because in certain kinematic configurations,

in this case when one of the leptons is parallel to the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, the m_{T2}^{0} variant returns a value directly proportional to the lepton masses [140]. As a result, all lepton masses are set to zero in the m_{T2} calculation used in the BDTs. A detailed explanation of this effect is given in Appendix B.



(a)



(b)

Figure 5.3: The normalised distributions of m_{T2}^{0} and m_{T2}^{100} for three example slepton signals and the major WW background. The endpoint of the distribution for each signal, minus the assumed LSP mass in the m_{T2} calculation (0 or 100 GeV) is largely driven by the mass-splitting between the slepton and LSP.

5.3.2 Preselection Requirements

An initial common preselection, shown in Table 5.3, is applied to all samples, and these baseline requirements are used in all analysis regions. This common preselection selects for two leptons and enforces an ISR topology by selecting for a high-p_{T} jet at large angles to the $E_{\mathrm{T}}^{\mathrm{miss}}$. A $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV requirement is implemented to ensure the trigger is fully efficient. Events containing leptons with $p_{\mathrm{T}} < 6$ GeV are found to contribute little to the sensitivity, due to the large FNP lepton background and so are rejected. Finally, $\Delta R_{\ell\ell} > 0.75$ is required to reduce the FNP lepton background. Additional requirements, shown in Table 5.3, are made in addition to these baseline requirements, to target the signal process in the training of the BDTs. For this training preselection, events are required to contain a SFOS lepton pair, no b-tagged jets, and no on-shell Z boson decays, enforced via a veto condition $|m_{\ell\ell} - m_Z| > 10$ GeV. The training selection is inclusive of lepton flavour, as the di-electron and di-muon selections are approximately kinematically equivalent.

Normalised, pre-fit distributions of kinematic variables with the training preselection applied are shown in Figure 5.4. These show the potential of a selection of variables to discriminate between signal and SM backgrounds. Overlaid are the kinematic distributions of three example SUSY models, corresponding to $m(\tilde{\ell}, \tilde{\chi}_1^0) = (200, 180)$ GeV, $(200, 170)$ GeV, and $(200, 150)$ GeV.

Table 5.3: Summary of the requirements applied all regions of the analysis.

Common preselection requirements	
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 200 GeV
$N_{\mathrm{lep}}^{\mathrm{base}}$	$= 2$
$N_{\mathrm{lep}}^{\mathrm{sig}}$	$= 2$
$\min(\Delta\phi(\mathrm{jets}, E_{\mathrm{T}}^{\mathrm{miss}}))$	> 0.4
$\Delta\phi(j_1, E_{\mathrm{T}}^{\mathrm{miss}})$	> 2.0
$\Delta R_{\ell\ell}$	> 0.75
N_{jet}^{30}	> 0
$p_{\mathrm{T}}^{j_1}$	> 100 GeV
$p_{\mathrm{T}}^{\ell_1}$ and $p_{\mathrm{T}}^{\ell_2}$	> 6 GeV

Table 5.4: Requirements for the BDT training selection, applied across all trained BDTs.

Variable	Requirement
Common preselection cuts	See Table 5.3
Lepton charge and flavour	SFOS
$N^{20}_{b-\text{jet}}$	$= 0$
$\lvert m_{\ell\ell} - m_Z \rvert$	> 10 GeV

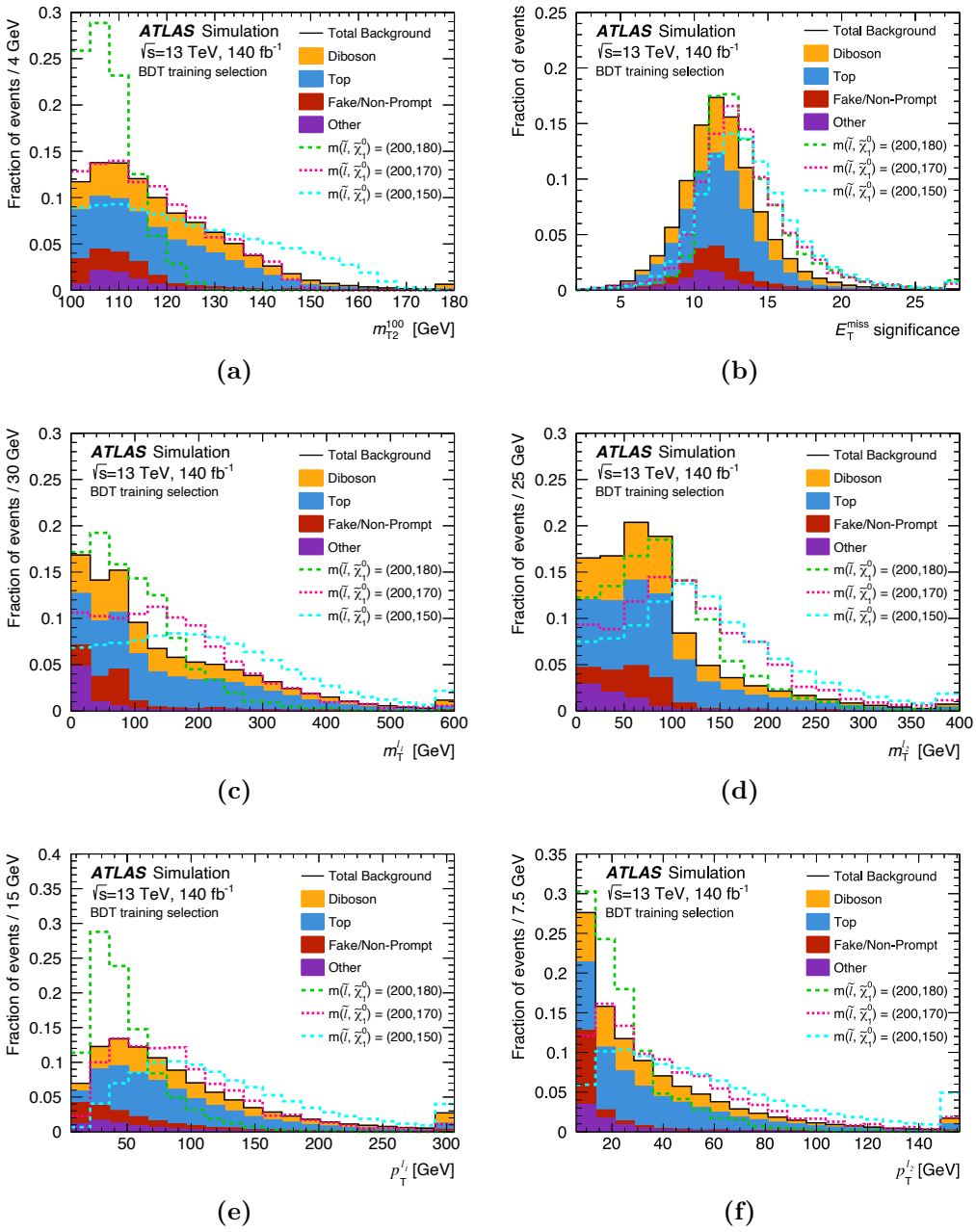


Figure 5.4: Pre-fit SM background (histograms) distributions normalised to unity, for (a) m_{T2}^{100}, (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, (c) $m_{\mathrm{T}}^{\ell_1}$, (d) $m_{\mathrm{T}}^{\ell_2}$, (e) $p_{\mathrm{T}}^{\ell_1}$, and (f) $p_{\mathrm{T}}^{\ell_2}$, using the BDT training preselection, Table 5.7. Three example SUSY scenarios are overlaid for illustration that correspond to $m(\tilde{\ell}, \tilde{\chi}_1^0) = (200, 180)$ GeV, $(200, 170)$ GeV, and $(200, 150)$ GeV, assuming $m(\tilde{\ell}_{\mathrm{L}}) = m(\tilde{\ell}_{\mathrm{R}})$. The last bin includes the overflow.

5.3.3 BDT Training

Using the training preselection requirements, five different BDTs are trained using the XGBoost package [141]. A decision tree is a type of machine learning model that makes decisions by recursively splitting data into subsets based on feature values, creating a tree of 'leaves' (also called nodes) with multiple branches, for classification or regression [142]. Boosted decision trees are an ensemble of decision trees that are built sequentially, where each tree corrects the errors of the previous one (boosting), improving the accuracy of the final predictions [143]. The BDTs trained in this analysis carry out binary classification to discriminate between signal and background processes by minimising the log loss in each sequentially trained decision tree:

$$L_{\log}(y_i, p_i) = -(y_i \log(p_i) + (1 - y_i) \log(1 - p_i)), \tag{5.2}$$

where 'y' is the truth signal/background label for the event and 'p' is the probabilistic prediction for an event between 0 (background-like) and 1 (signal-like). To evaluate the full ensemble of all events, Eq. 5.2 is computed for all events, and the result is then normalised by dividing by the total number of events. In XGBoost, the minimisation is carried out using first (gradient) and second (Hessian) derivative of Eq. 5.2 with respect to p:

$$\frac{dL_{\log}}{dp_i} = g_i = \frac{p_i - y_i}{p_i(1 - p_i)}, \tag{5.3}$$

$$\frac{d^2L_{\log}}{dp_i^2} = h_i = \frac{1}{p_i(1 - p_i)}. \tag{5.4}$$

The gradient is just the prediction minus the truth label, and the Hessian gives information on the local curvature of the loss function.

The initial leaf contains all events, and the leaf is split using a requirement on the variable that leads to the biggest decrease in loss. Each BDT is trained using the same set of 23 kinematic variables, indicated in Section 5.3.1, representing a mix of low- and high-level variables. These were chosen by examining variables important to the cut-and-count-based approach, and variables used in previous analyses targeting similar final states [117–119].

After each split, each leaf is then assigned a weight, calculated following:

$$w_j = -\frac{\sum_{i \in \text{leaf } j} g_i}{\sum_{i \in \text{leaf } j} h_i + \lambda}, \tag{5.5}$$

where, 'w_j' is the weight for an individual leaf, the '$\sum_{i \in \text{leaf } j}$' is the sum over all events in the leaf, and 'λ' is an optimisable parameter that penalises a large number of leaves. The terms g_i and h_i are the gradient and Hessian respectively, defined in Eq. 5.4. The weight of a leaf gives the predicted value (p_i) for all points that fall within that leaf, and so with a full set of weights the final decision tree can be entirely constructed. After each existing leaf in a tree has been split, the procedure is repeated until a limit on the number of leaves is reached, or overtraining protections halt the training. Figure 5.5, shows a schematic of how the decision tree is built using this successive requirements on input variables.

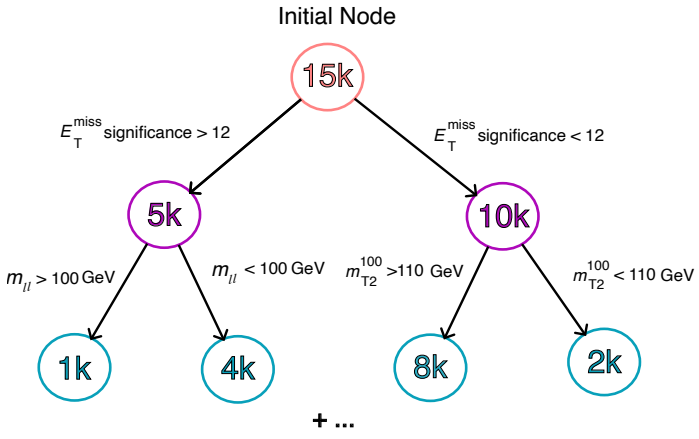


Figure 5.5: Schematic of how a decision tree iteratively splits an initial leaf into multiple nodes through requirements on kinematic input variables. The numbers in each are an example of the change in events in a leaf.

Once a tree is trained, it is compared to the previous tree:

$$z_{\text{new},i} = z_{\text{old},i} - \eta \cdot \frac{p_i - y_i}{p_i(1 - p_i)}, \tag{5.6}$$

where 'z' is the logit ($= \log\left(\frac{p}{1-p}\right)$) calculated for the current and previous tree, which maps a probability $p \in (0, 1)$ to a real number $z \in \mathbb{R}$, and η is the learning rate.

In words, if $p > y$, the tree overestimated the probability, so the logit is reduced, and if $p < y$, the tree underestimated the probability, so the logit is increased. Once the maximum number of trees is reached (N_{Max}) or the training is halted, the final logit for a given input is given by the sum of the individual tree contributions:

$$z_{\text{final},i} = z_{\text{initial},i} + \sum_{t=1}^{T} \eta \cdot f_t(x_i), \tag{5.7}$$

where $f_t(x_i)$ is the output of the t-th tree for a given event x_i, and η is the learning rate. Finally, the predicted probability is obtained by applying the logistic function to the final logit:

$$p_{\text{final},i} = \frac{1}{1 + \exp\left(-z_{\text{final},i}\right)}. \tag{5.8}$$

A BDT left without any restrictions on the training, will perfectly fit the training dataset, but not extrapolate well to an unseen dataset, which is called 'overfitting' or 'overtraining'. To evaluate the level of overtraining, a validation sample is used in the training, and evaluated using the trained BDT after each iteration, with the log loss calculated. To constrain the amount of overtraining, early stopping is employed in this analysis, where training is halted if log loss of the validation set does not improve after 20 additional trees [3]. In addition, in the training itself the overtraining can be controlled through the introduction of regularisation terms to the loss function:

$$L_{\text{Reg}}(y_i, p_i) = L_{\log}(y_i, p_i) + \Omega(f_t), \tag{5.9}$$

where $\Omega(f_t)$ is the regularisation term for the t-th tree, given by:

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \tag{5.10}$$

where γ is a parameter that constrains the number of leaves, T is the total number of leaves in the tree, and λ and w_j are defined in Eq. 5.5.

The training of the BDTs is tweaked through hyperparameters. Hyperparameter optimisation was carried out independently for all five BDTs. This is expected as the different training sets may not necessarily converge to the same minimum.

[3] This number was found to be optimal through varying the parameter, following the same procedure as the hyperparameter optimisation.

The following hyperparameters were varied:

- The maximum number of trees in the BDT (N_{Max}) – this does not get reached by design due to early stopping.
- The learning rate (η) – a scale factor determining how much a given tree contributes to the overall ensemble of trees.
- The minimum number of events per leaf – total $\sum_{j\in\text{leaf } j} w_j$ before splitting a leaf.
- The gamma parameter (γ) – constrains the number of tree to control complexity and prevent overtraining.
- The alpha regularisation parameter (α) – a penalty term added to the gradient to prevent overtraining.
- The lambda regularisation parameter (λ) – a penalty term added to Hessian to prevent overtraining.

First the hyperparameters were varied across the same range for all BDTs. Then a second round of optimisation was carried out where the values were varied more finely around the values that gave the best performance for each BDT. The final hyperparameter configurations for each of the five BDTs are detailed in Table 5.5.

Table 5.5: Hyperparameter choices for each BDT. The final number of trees varies across the different folds of the same BDT and so the average is taken.

Hyperparameter	BDT_{5+10}	BDT_{20}	BDT_{30}	BDT_{40+50}	BDT_{60+75}
(Average) final number of trees	372	489	668	789	924
Learning rate	0.025	0.025	0.04	0.025	0.025
Minimum number of events per leaf	5	2	5	2	15
Alpha	6	1	1	1	1
Gamma	1.1	0.4	0.3	0.25	0.25
Lambda	12	5	6	20	15

Training Samples

Each BDT is trained using a different selection of signal samples, grouped according to the mass-splitting of the targeted signal models. Each of these signal sample groups is trained with the same background samples, the MC samples for truth-matched leptons and the data-driven fakes estimate for the FNP lepton contribution. Signal kinematics change significantly with the mass-splitting, and so multiple BDTs were used, as these give strong sensitivity across the entire range of targeted $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$. The signal samples are grouped as follows, where the subscript denotes the grouping of the mass-splittings in that BDT: Δm_{5+10}, Δm_{20}, Δm_{30}, Δm_{40+50}, Δm_{60+75}. The groupings of the signal samples for the training of each BDT are shown in Figure 5.6. Studies were carried out on grouping samples with neighbouring mass-splittings, and an optimal grouping of mass-splittings was determined through comparison of the classification performance of the trained BDTs. Some signals were not included in the training, as they were not in the region which the analysis initially set out to target.

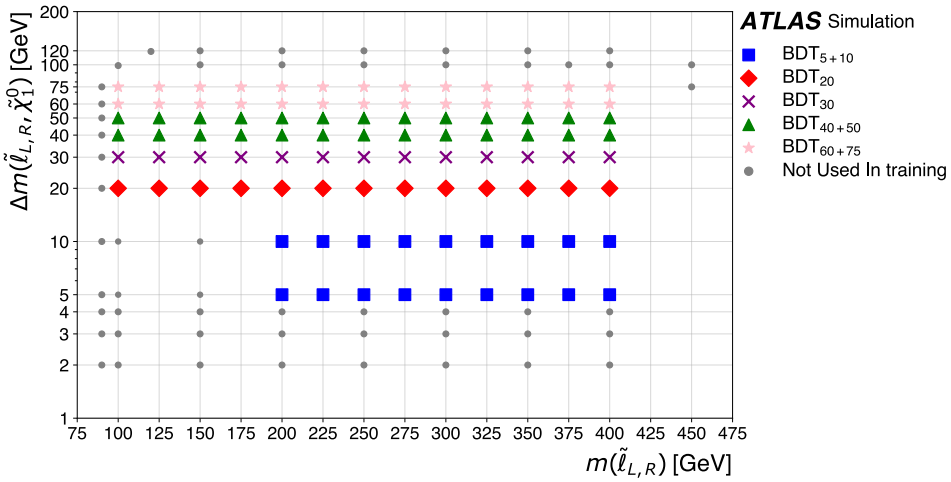


Figure 5.6: Overview in the $m(\tilde{\ell}_{\mathrm{L,R}})$ – $\Delta m(\tilde{\ell}_{\mathrm{L,R}}, \tilde{\chi}_1^0)$ plane of how the simulated benchmark signal models are grouped for the BDT training. The samples entering into each BDT are colour and symbol coded.

For a given BDT, a five-fold cross-validation strategy is employed, where the MC samples for each background process and signal grouping, are divided into fifths. The k-fold cross validation method allows all data to be evaluated using a

BDT where the training was 'blind' to those specific events, increasing robustness of evaluation. Five BDTs are trained for each mass splitting, with each using a different three fifths of the data for training, one fifth for validation (to monitor overtraining), and one fifth for evaluation. Figure 5.7 shows a schematic of the procedure.

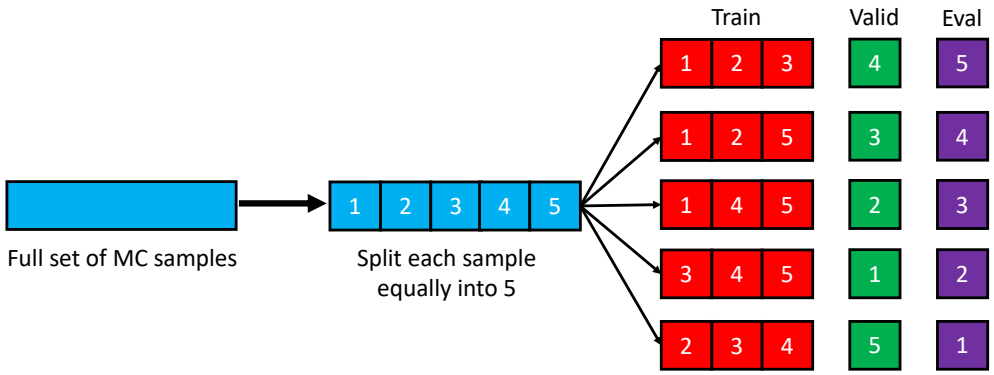


Figure 5.7: Schematic of the KFold procedure. This is carried out for each signal sample and SM background individually.

The same hyperparameter configurations are used for all five different k-fold trainings, with the final number of trees varying by 1–2% between the different k-fold trainings. For each mass splitting, the final BDT score for each event is taken from the BDT for which the event was in the evaluation fifth, ensuring that each BDT is only used to evaluated events not used in its training. This ensures no overlap between training and evaluation data across the five BDTs. The split is achieved by taking the modulus of the event number relative to five, and the same strategy is applied when determining which of the five BDTs used to evaluate the samples (e.g., data) not used in the training.

The background and signal samples have class weights applied so that the signal samples are weighted up to match the number of background events. In addition, MC event weights are used in the BDTs as sample weights. For the background these are simply the absolute value of the event weights [4]. This is done because negative weights lead to the value of the log loss becoming undefined, as the logarithm of a negative number is undefined.

[4]The signal samples are generated at NLO, so have no negative weights.

For the MC backgrounds, these only form a very small percentage of the overall events, however for the FNP lepton background, the contribution is larger. To preserve the shape of the weighted kinematic distributions of the FNP lepton background, the absolute values are re-weighted so that the overall sum of weights is equal to the initial sum of weights before modification.

For the signal events, the sample weights used are also the event weights, but are rescaled to ensure equal representation of each signal sample in a Δm BDT training. For a given signal sample, k, and for a given event weight, w_j, in the sample, the BDT training weight w_{BDT} is assigned using the following equation:

$$w_{BDT} = w_j \times \frac{\sum_{i=0}^{N_{\text{Tot}}} w_i}{\sum_{k=0}^{N_{\text{sig}}} w_k}, \tag{5.11}$$

where the sum over 'i' is the sum over the event weights of all signal samples in a given training, and the sum over 'k' is the sum over the event weights of the signal sample the event belongs to.

Diagnostics and Validation

Four main metrics are used to evaluate the performance of the BDTs, in order to pick the optimal hyperparameter combination and to verify that all 5 folds per Δm have similar performance. The log loss, Equation 5.2, is calculated for an orthogonal validation sample to check for overtraining.

Next, there is a comparison of the training and validation sample scores, normalised to unit area for both signal (averaging over all signals in the training) and background samples. This shows performance in signal versus background classification of the BDT but also used to check that the training and validation distribution shapes are similar. A Kolmogorov-Smirnov (KS) test is implemented to compare training and validation scores [144]. This is calculated for both signal and background, and the result of this is shown in the legend. The closer to one, the greater the similarity in the distributions, and the smaller the overtraining. This is shown for the first fold of the BDT_{40+50} in Figure 5.8.

Another diagnostic plot is the Receiver Operating Characteristic (ROC) Curve. In this section the TP, FP are used as shorthand for true and false positives respectively and TN, FN are used for true and false negatives respectively. The true positive rate: TPR = $\frac{\text{TP}}{\text{TP+FN}}$ is plotted against the false positive rate: $\frac{\text{FP}}{\text{FP+TN}}$. This allows classifier performance to be evaluated, with the aim of maximising the TPR while minimising the FPR, in other words, aiming for the curves to be in the top left corner of the axis. The area under the ROC Curve (AUC) is also calculated, and can be used to measure of BDT performance - a higher value indicates better performance. Comparing the training versus validation AUC values shows that there is little difference in performance on the training and validation samples. This is shown for the first fold of the BDT_{40+50} in Figure 5.8

The number of events in the signal and background samples is unbalanced, with significantly more background than signal. As a result, the number of easily classified true negatives can be large, and this can give misleading results, when using a ROC curve evaluating the performance. A precision-recall curve (PRC) is a measure of how well a BDT finds the true result (precision) without missing too many (recall). Precision is defined as $\frac{\text{TP}}{\text{TP+FP}}$, and recall is another name for the true positive rate. The PRC is shown for the first fold of the BDT_{40+50} in Figure 5.8. The performance can be evaluated by comparing the value of the precision at high recall, and allows for better optimisation of performance regardless of class imbalance. Also shown on the PRC is the average precision which is analogous to the AUC but for PRC Curves.

Variable Importance

The importance of the variables in each of the trained BDTs was evaluated using the SHAP package [145]. The SHAP score is a measure of variable importance in the training. The larger the positive value, the greater the importance of the variable for classifying events as signal-like. The larger the negative value, the greater the importance of the variable for classifying events as background-like. A summary plot called a 'beeswarm plot' is provided by the SHAP package and shown for the first fold of the BDT_{40+50} in Figure 5.9. On the x-axis is the SHAP

score, the width of the 'swarms' in the vertical, y direction represents the relative density of events, and the colour represents the relative value of the variable (high or low). The variables are ordered by the mean absolute value of the SHAP score (x-axis), the value of which is also shown in Figure 5.10.

Figure 5.9 shows that low $m_{\ell\ell}$ values and events without a second jet are used to reject background events, and high $E_{\mathrm{T}}^{\mathrm{miss}}$ significance and $m_{\mathrm{T}}^{\ell_1}$ values are used to accept signal events. The importance rankings were seen to be consistent across the five k-fold BDTs of the same mass-splitting. An overview of the top three most important variables in each BDT is given in Table 5.6.

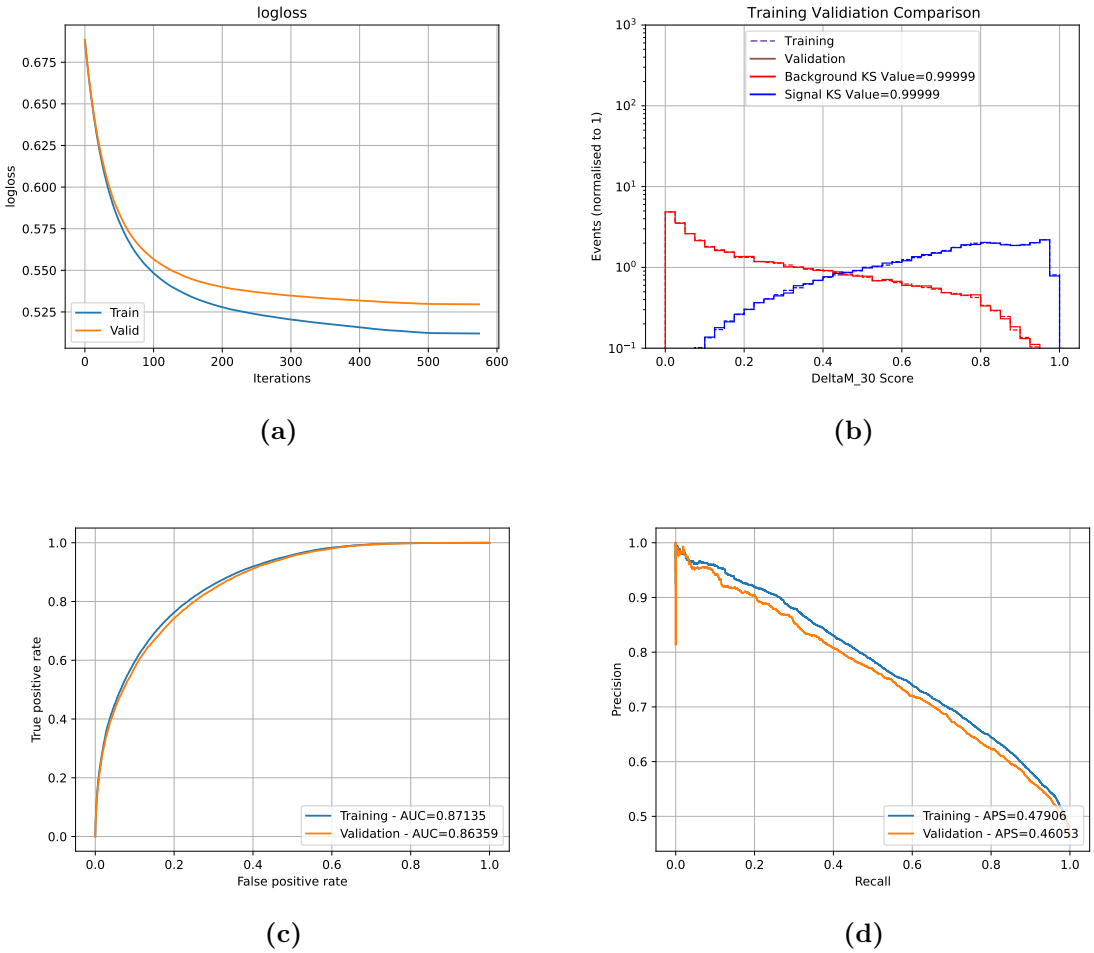


Figure 5.8: BDT Diagnostic plots for the first fold of the BDT_{40+50}. In (a) the log loss is shown, for the training and validation samples. In (b), a comparison between the BDT score distributions for the training and validation sets, for background and signal, are shown. In (c) the ROC curve is shown, for the training and validation samples. In (d) the PRC curve is shown, for the training and validation samples.

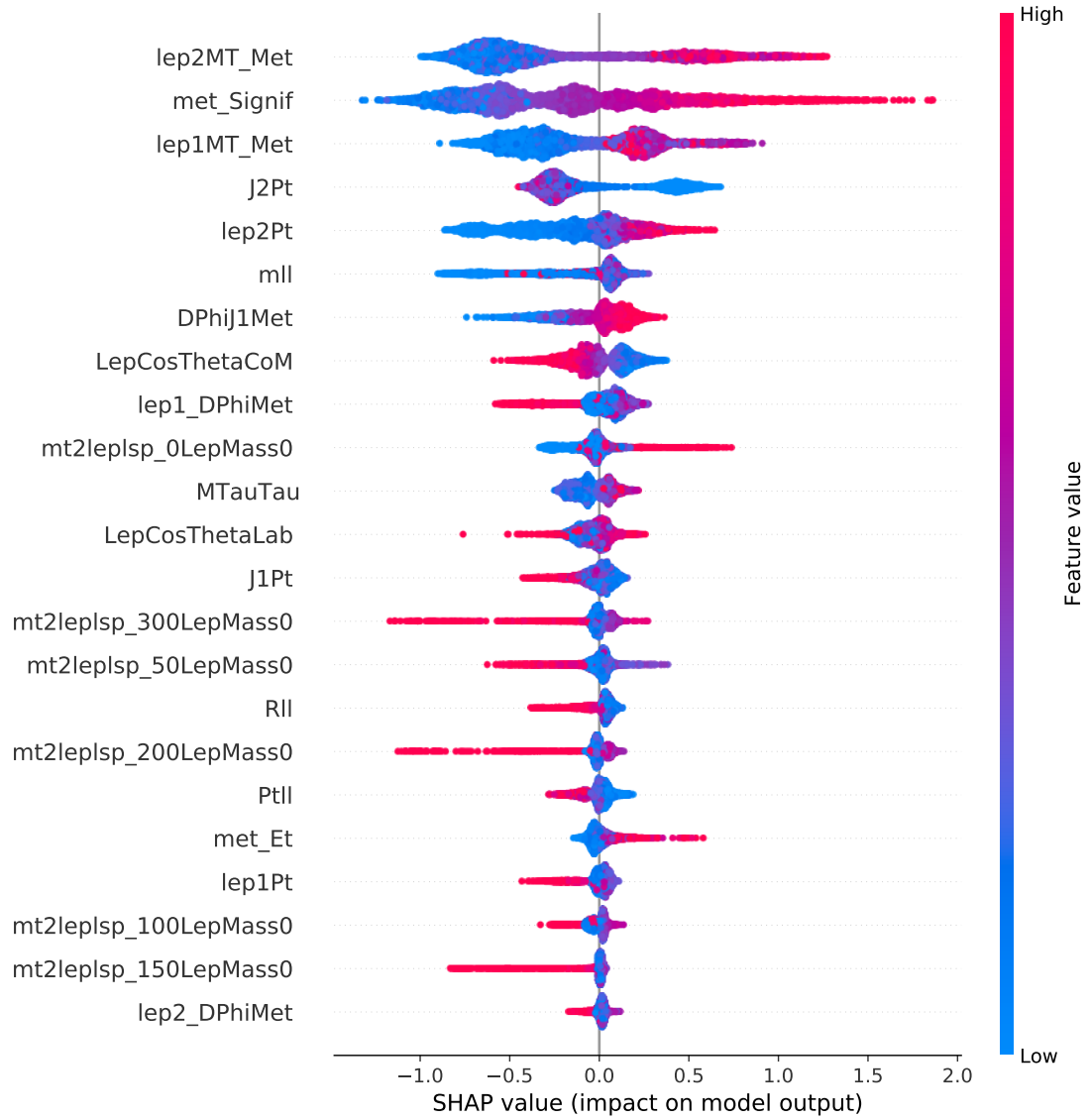


Figure 5.9: The beeswarm plot for the first fold of BDT_{40+50}.

Table 5.6: The top three ranked variables for each BDT.

BDT	Feature 1	Feature 2	Feature 3
BDT_{5+10}	High m_{T2}^{200}, m_{T2}^{300}	Low $p_T^{\ell\ell}$ combined with low m_{T2}^{0}	High E_T^{miss} and E_T^{miss} significance
BDT_{30}	Low $p_T^{\ell_1}$	High m_{T2}^{100}, m_{T2}^{150}	High E_T^{miss} significance
BDT_{30}	High m_{T2}^{200}	High E_T^{miss} significance	High $m_T^{\ell_2}$
BDT_{40+50}	High E_T^{miss} significance	High $m_T^{\ell_1}$, $m_T^{\ell_2}$	High $p_T^{\ell_2}$
BDT_{60+75}	High E_T^{miss} significance	High $m_T^{\ell_1}$, $m_T^{\ell_2}$	High $p_T^{\ell_2}$

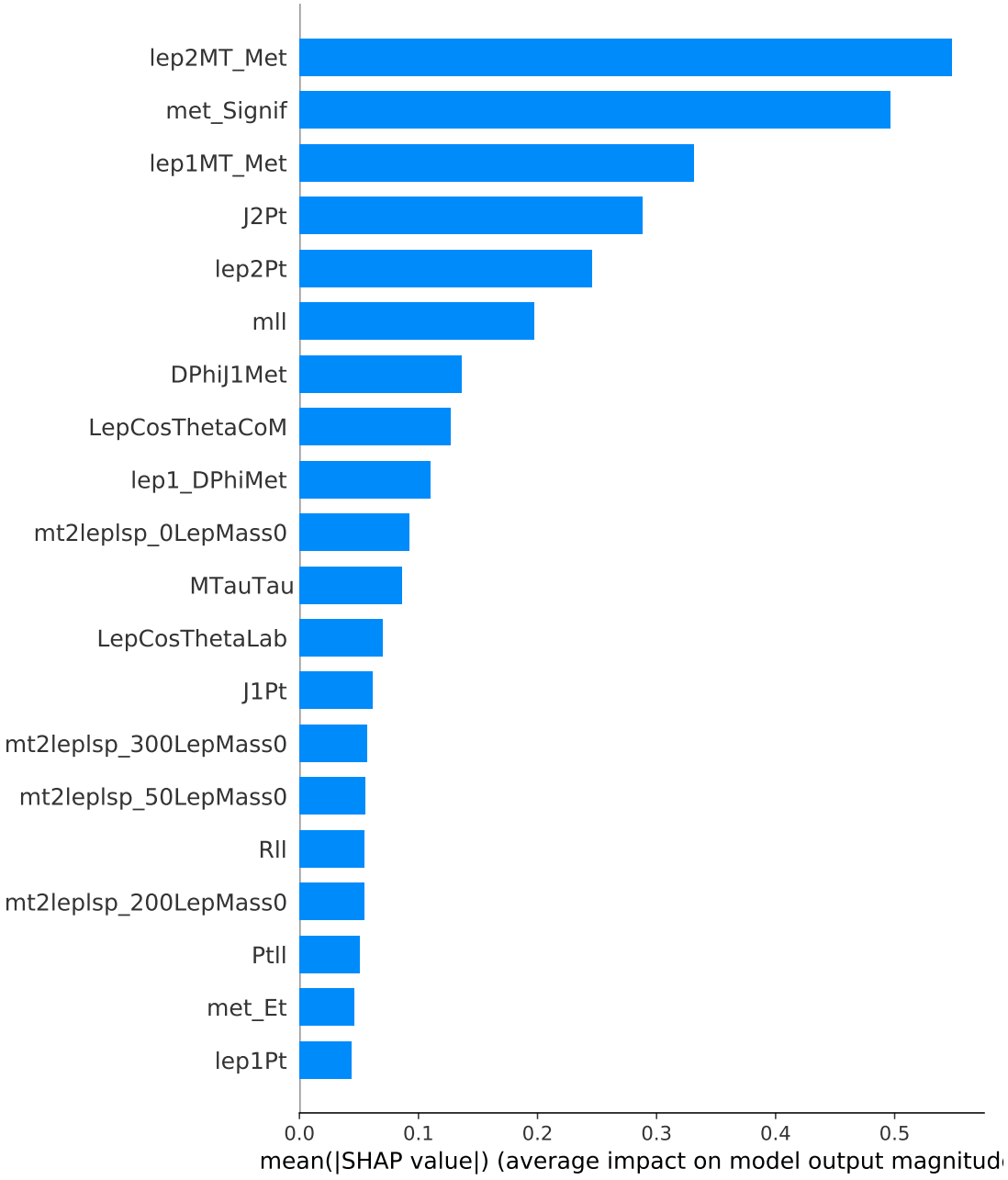


Figure 5.10: The variable importance for the first fold of BDT_{40+50}, where the importance is calculated as the mean absolute SHAP value for the variable.

5.3.4 Signal Region Definitions

The BDT output, a classification score between 0 (background-like) and 1 (signal-like), is used to distinguish signal and background processes. The normalised distributions for the signal and background BDT score are shown in Figure 5.11. The signal distribution shown is an average over each signal model that BDT was trained on.

These resulting BDT score distributions are used to define the SRs for the analysis. The definitions for the SRs are shown in Table 5.7, and are applied in addition to the training preselection described in Tables 5.3 and 5.4 above. The tail of the BDT score is binned into three non-overlapping SR bins to ensure sensitivity to the full range of signals, as the BDT score distributions peak at different values depending on the slepton mass. The binning was optimised individually for each set of BDTs to account for the different shape of the score distributions. Each SR bin is further split via the lepton flavour into ee and $\mu\mu$ events yielding a total of six SRs for each BDT. The SR naming convention follows the convention: $\mathrm{SRX-BDT}^{\mathrm{Z}}_{\mathrm{Y}}$, where X is the SR number $(1, 2, 3)$, Y is the BDT used to define the region, and Z is the lepton-flavour requirement. For example, the first di-electron SR, defined using the smallest mass-splitting BDT is denoted by $\mathrm{SR1\text{-}BDT}^{\mathrm{ee}}_{5+10}$.

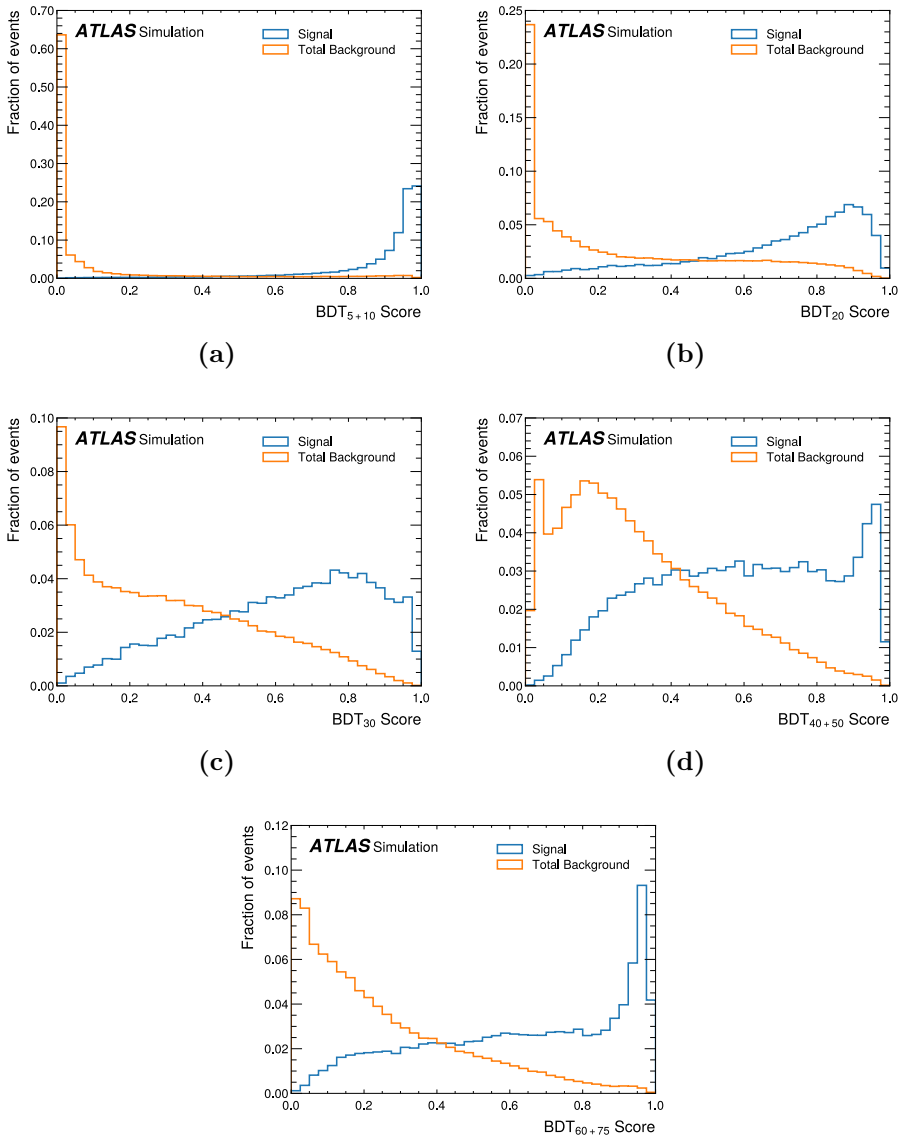


Figure 5.11: Normalised score distributions of (a) BDT_{5+10}, (b) BDT_{20}, (c) BDT_{30}, (d) BDT_{40+50} and (e) BDT_{60+75} for the SUSY signal and the inclusive SM background using all events considered for the BDT training. The SUSY signal distribution represents an average over all signal samples used in the associated training.

Table 5.7: Definition of the BDT-based SR requirements that are applied in addition to the preselection requirements shown in Table 5.3. A BDT score requirement that is exclusive of the lower bound is indicated by the use of "(" in defining the BDT score requirements. Along with the BDT-score requirement, the SRs are further split into three opposite-sign di-electron SRs, and three opposite-sign di-muon SRs.

	BDT Score requirements		
BDT set	$SR1^{ee/\mu\mu}$	$SR2^{ee/\mu\mu}$	$SR3^{ee/\mu\mu}$
BDT_{5+10}	$[0.93, 0.953]$	$(0.953, 0.976]$	$(0.976, 1.0]$
BDT_{20}	$[0.90, 0.928]$	$(0.928, 0.956]$	$(0.956, 1.0]$
BDT_{30}	$[0.88, 0.915]$	$(0.915, 0.95]$	$(0.95, 1.0]$
BDT_{40+50}	$[0.90, 0.928]$	$(0.928, 0.956]$	$(0.956, 1.0]$
BDT_{60+75}	$[0.90, 0.935]$	$(0.935, 0.97]$	$(0.97, 1.0]$

5.4 Background Estimation

Each BDT has its own separate background estimate, giving five individual sets of regions in total, defined following a similar overall strategy. Two control regions (CRs) are used to constrain the major irreducible backgrounds (top and diboson) in a simultaneous fit, to extract two normalisation factors. In addition, a set of validation regions (VRs) are used to validate the background estimate, by extrapolating the normalisation factors extracted from the CRs. The major reducible background originates from FNP leptons, and is estimated through a data-driven method, detailed in Section 5.4.2.

5.4.1 Irreducible Backgrounds

Two control regions (CRs) are used to constrain the two major sources of irreducible background. CR-VV is enriched in diboson processes, CR-Top is used to target processes involving top quarks. The composition of the $t\bar{t}$ and single-top processes were found to be similar across the phase space of the analysis regions, so these are estimated in conjunction, using a common normalisation factor. Subdominant irreducible backgrounds such as $Z \to \tau\tau$ + jets and ttV events were found to contribute typically less than 5% in the SRs, and do not have an associated normalisation factor. These are grouped into the "other" backgrounds category in

the analysis, which consists of all non-VV, $t\bar{t}$ and single-top background processes listed in Table 5.1.

A dedicated set of CRs is constructed for each of the BDTs, which ensures that the SM backgrounds are constrained in a phase space similar to, but statistically independent of the corresponding SRs. The background normalisation factors derived from these CRs, are then verified in a set of validation regions (VRs) that are also individually defined for each BDT. By changing the same-flavour to a different-flavour requirement on the leptons in the SRs, orthogonal VR-DF regions that are otherwise identical to the SRs are defined. This takes advantage of the dominant, flavour symmetric W^+W^- and $t\bar{t}$ background processes, which produce SFOS and DFOS leptons at equal rates, whereas the signal process only produces SFOS leptons.

5.4.2 Reducible Backgrounds

Reducible backgrounds attributed to FNP leptons can arise from several sources: jets misidentified as leptons, photon conversions into e^+e^- pairs, and real, non-prompt leptons produced in the semileptonic decays of hadrons containing bottom or charm quarks. In this analysis, the FNP background is dominated by $W \to \ell\nu + \text{jets}$ events, where one real lepton originates from the W decay and the second lepton from a misidentified jet. As lepton misidentification rates are typically small it would be computational expensive to model this FNP lepton with MC simulations with sufficiently large statistics. Hence, a data-driven method referred to as fake-factor method [146] is employed to estimate the FNP background. A brief introduction to the fake-factor method will be given here – for a more detailed description, see Ref. [146]. Although this estimate was not performed by me, an explanation is provided to give the reader the necessary context to interpret the results of the analysis.

In the fake-factor method, the FNP contribution to a selection requiring tight, signal ('ID') leptons such as the SRs is estimated from an orthogonal control sample failing to meet the tight criteria but satisfying looser ('anti-ID') requirements on the leptons. In this analysis, ID leptons are identical to signal leptons, while anti-ID

leptons are preselected leptons that fail to satisfy at least one of the criteria for signal leptons on identification, isolation and impact parameters, as defined in Table 5.2.

The fake factors are measured in a data sample collected with prescaled low-p_{T} single-lepton triggers. This sample is referred to as the measurement sample and is dominated by multijet events with minor contributions from top and W + jets processes. The sample is required to contain one preselected lepton [5] and is enriched in FNP leptons similar to those contaminating the target phase space by requiring the leading jet p_{T} to be greater than 100 GeV. The fake factor is defined as the ratio of the number of ID leptons to the number of anti-ID leptons in the measurement sample:

$$F(i) = \frac{N_{\mathrm{ID}}(i)}{N_{\mathrm{anti-ID}}(i)}. \tag{5.12}$$

The quantities $N_{\mathrm{ID}}(i)$ and $N_{\mathrm{anti-ID}}(i)$ are the number of ID and and anti-ID leptons respectively, paramatrised by a chosen variable 'i'. In this analysis, i is chosen as the lepton p_{T} to capture the strong kinematic dependence of the FNP leptons on this variable, and the fake factors are measured individually for electrons and muons. Furthermore, there is a dependence of the fake factors on the absence or presence of a b-tagged jet in the event, which is attributed to the increased proportion of heavy flavour hadron decays in the latter. Hence, the fake factors are also measured separately for events with no b-tagged jets, and with at least one.

To take into account the contamination from real leptons, their contribution is estimated from $t\bar{t}$ and W + jet MC simulation and subtracted from the ID and anti-ID samples. As a result, the fake factor definition is then modified:

$$F(i) = \frac{N_{\mathrm{ID}}(i) - N_{\mathrm{ID}}^{\mathrm{MC}}(i)}{N_{\mathrm{anti-ID}}(i) - N_{\mathrm{anti-ID}}^{\mathrm{MC}}(i)}, \tag{5.13}$$

where $N_{\mathrm{ID}}^{MC}(i)$ and $N_{\mathrm{anti-ID}}^{MC}(i)$ are the number of ID and anti-ID prompt leptons in MC, for the ith p_{T} bin, respectively. The measured fake factors in events without b-jets range from approximately 0.1 (0.15) to 0.65 (0.1) for electrons (muons) with a p_{T} of around 10 GeV and 50 GeV, respectively and are shown in Figure 5.12.

[5] Evaluation was carried out on using Z+ jets events, with two preselected leptons however the FF's from this selection were seen to agree less well when extrapolated to SS lepton regions.

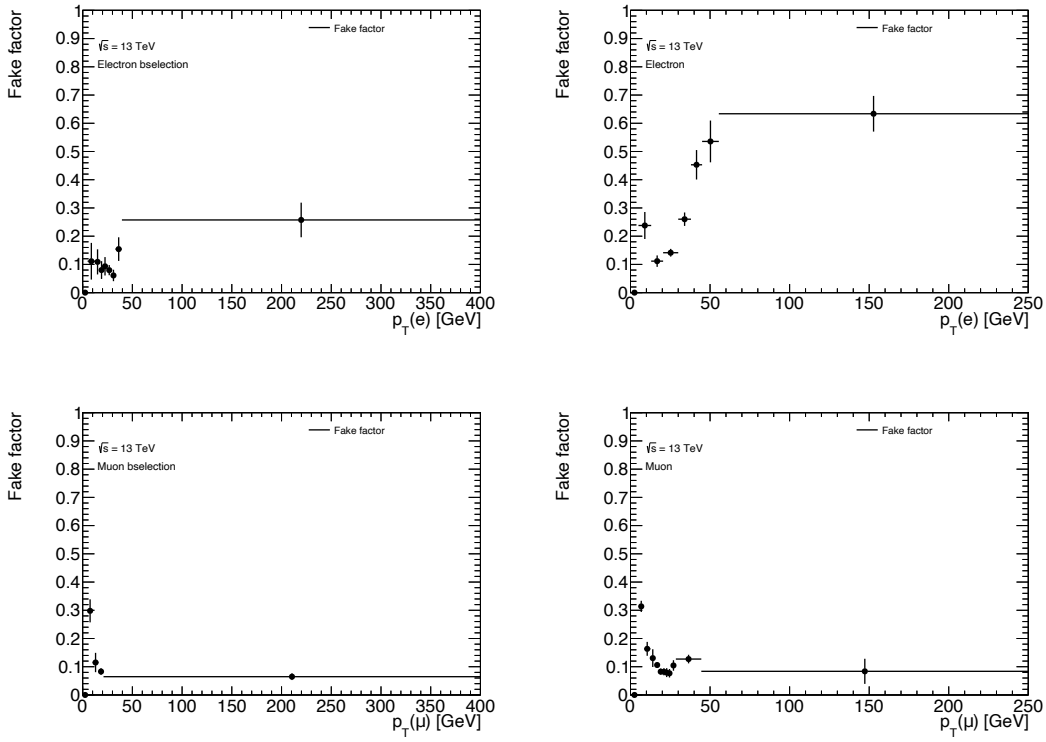


Figure 5.12: Measured fake factors as a function of lepton p_{T}, for electrons (top) and muons (bottom) in b-jet (left) and b-veto (right) selections.

The fake factors are then used to obtain the reducible FNP background in the signal region, N_{SR}. A control sample is constructed by applying the same kinematic criteria as the associated target region (which can be any CR, VR or SR), but requires one or both leptons to be anti-ID. The contribution from irreducible backgrounds featuring two prompt leptons in the control sample is estimated from MC simulation and subtracted before the extrapolation, as is done in the measurement region. The number of leptons in the ith bin is denoted $N_{\mathrm{anti-ID\,CR}}(i)$. From here, N_{SR} for the case where there is only one fake lepton can be computed:

$$N_{\mathrm{SR}} = \sum_i [N_{\mathrm{anti-ID\,CR}}(i) - N^{MC}_{\mathrm{anti-ID\,CR}}(i)] \cdot F(i). \tag{5.14}$$

In this analysis events may contain more than one fake lepton, so the equation above needs to be extended to allow for combinatorics arising from cases where there are two fake leptons. To account for these double fake lepton events, the

leading and subleading leptons are considered separately when estimating N_{SR}:

$$N_{\mathrm{SR}} = F_1 \left(N_{\mathrm{AI}} - N_{\mathrm{AI}}^{\mathrm{MC}}\right) + F_2 \left(N_{\mathrm{IA}} - N_{\mathrm{IA}}^{\mathrm{MC}}\right) - F_1 F_2 \left(N_{\mathrm{AA}} - N_{\mathrm{AA}}^{\mathrm{MC}}\right). \qquad (5.15)$$

Here, N indicates the number of events in the anti-ID control region and the subscripts indicate whether the leptons pass the ID (I) or anti-ID (A) criteria. The subscripts are also ordered such that the first subscript corresponds to the lepton with the largest p_{T}. For example, N_{AI} is the number of events in the anti-ID CR where the highest-p_{T} lepton is an anti-ID lepton and the lower-p_{T} lepton is an ID lepton. Terms with an 'MC' superscript indicate the contributions from prompt lepton processes that are estimated using MC and subtracted. The F_1 and F_2 correspond to the fake factors of the higher and lower p_{T} lepton in the event, respectively. The minus sign in the last term serves to correct for the double-counting of events with two FNP leptons, since these can contribute to both the N_{IA} and N_{AI} terms. The magnitude of this effect is small, at most 5% of the overall FNP lepton yield in any region.

To validate the data-driven FNP lepton estimate in the SRs, dedicated VRs are defined. These VRs impose a SS lepton requirement when defining the VR, enriching the FNP background, and are consequently referred to as VR-SS. The kinematic requirements applied to each VR-SS are the same as the ones used in the corresponding SR, ensuring the FNP lepton processes are similar in the two regions. The only difference being that, in some cases, the BDT score requirements are relaxed to increase the event yields. The contribution of FNP background in the VR-SS regions is typically above 90%, with the remaining background originating from diboson processes with two prompt leptons of the same charge (WZ, where one lepton from the Z is not reconstructed.).

5.4.3 Control and Validation Regions

Two CRs and four VRs per BDT, statistically orthogonal to one another for each BDT, are defined as shown in Table 5.8. Kinematic selections, along with BDT score requirements, are applied to avoid overlap with the SRs, and to target specific

background processes. CR-Top-BDT targets background events involving top quarks and is orthogonal to the SRs due to a $N^{20}_{b-\text{jet}} > 0$ requirement. CR-VV-BDT targets VV background events and is orthogonal to the SRs through requirements on the BDT score and lepton flavour. The naming convention of the regions follows: Region-BDT$^{\text{Y}}_{\text{X}}$, where X indicates the BDT used to define the region and Y corresponds to an additional lepton flavour requirement, if any. For example, the CR-VV defined using the Δm_{5+10} BDT is denoted by CR-VV-BDT$_{5+10}$.

The VV component in the phase space targeted by this analysis is $\approx 85\%$ W^+W^-, in all regions other than VR-ZZ-BDT. Appendix **??** shows the distribution in the BDT score of the Δm_{60+75} regions, using an alternative VV MC where the individual contributions are split.

For CR-VV-BDT$_{5+10}$ and CR-VV-BDT$_{20}$, the requirement $p^{\ell_2}_{\text{T}} > 7.5$ GeV is introduced to reduce the FNP-lepton contamination and increase the purity of the VV background.

To validate the background normalisation factors extracted from these CRs, a set of validation regions is defined. VR-DF-BDT is orthogonal to the SRs through a DF lepton requirement while keeping all other selections of the corresponding SRs. VR-DF-Top-BDT is equivalent to the VR-DF but with a $N^{20}_{b-\text{jet}} > 0$ requirement to validate background processes involving top quarks. VR-ZZ-BDT is used to validate the modelling of the ZZ/WZ backgrounds. It is defined by requiring $|m_{\ell\ell} - m_Z| < 10$ GeV in order to increase the contribution from events containing on-shell Z bosons. The FNP lepton estimate is validated in VR-SS-BDT, which is defined by requiring a pair of SS leptons.

In the case of the Δm_{5+10} BDT VRs, the definitions are slightly modified. The VR-DF is split VR-DF-BDT$^{\text{High}}_{5+10}$, which selects events with $p^{\ell_2}_{\text{T}} > 10$ GeV, and VR-DF-BDT$^{\text{Low}}_{5+10}$, which selects events with $p^{\ell_2}_{\text{T}} < 10$ GeV. The former VR is largely top and diboson events, while the latter VR is dominated by FNP background, hence these VRs verify the modelling of these background components individually. VR-SS is also split according to the subleading lepton flavour into two VRs, VR-SS-BDT$^{e}_{5+10}$ and VR-SS-BDT$^{\mu}_{5+10}$. This split independently validates the

FNP background estimate for subleading electrons and muons in the lowest mass-splitting BDT regions, where the SRs contain a significant FNP lepton component. Figure 5.13 summarises the CRs, VRs and SRs, showing graphically how they are orthogonal to one another. The VR-ZZ is not shown, but is identical to the SRs, just with the on-shell Z veto reversed.

Table 5.8: Summary of region definitions for all CRs and VRs in the BDT approach. Only BDT score requirements and selections that enforce orthogonality between regions are shown. Preselection requirements from Table 5.3 are also applied. A BDT score requirement that is exclusive of the upper bound is indicated by the use of ")" in defining the BDT score requirements.

Region	BDT score	Additional requirements	Lepton flavour and charge
CR-VV-BDT	$\text{BDT}_{5+10} = [0.5, 0.93)$	$p_{\text{T}}^{\ell_2} > 7.5$ GeV	OS leptons
	$\text{BDT}_{20} = [0.7, 0.9)$		
	$\text{BDT}_{30} = [0.7, 0.88)$	—	
	$\text{BDT}_{40+50} = [0.7, 0.9)$		
	$\text{BDT}_{60+75} = [0.7, 0.9)$		
CR-Top-BDT	$\text{BDT}_{5+10} = [0.7, 1.0]$	$N_{b-\text{jet}}^{20} > 0$	SFOS leptons
	$\text{BDT}_{20} = [0.7, 1.0]$		
	$\text{BDT}_{30} = [0.7, 1.0]$		
	$\text{BDT}_{40+50} = [0.7, 1.0]$		
	$\text{BDT}_{60+75} = [0.7, 1.0]$		
VR-DF-BDT	$\text{BDT}_{5+10} = [0.93, 1.0]$	—	DFOS leptons
	$\text{BDT}_{20} = [0.9, 1.0]$		
	$\text{BDT}_{30} = [0.88, 1.0]$		
	$\text{BDT}_{40+50} = [0.9, 1.0]$		
	$\text{BDT}_{60+75} = [0.9, 1.0]$		
VR-DF-Top-BDT	$\text{BDT}_{5+10} = [0.93, 1.0]$	$N_{b-\text{jet}}^{20} > 0$	DFOS leptons
	$\text{BDT}_{20} = [0.9, 1.0]$		
	$\text{BDT}_{30} = [0.88, 1.0]$		
	$\text{BDT}_{40+50} = [0.9, 1.0]$		
	$\text{BDT}_{60+75} = [0.9, 1.0]$		
VR-ZZ-BDT	$\text{BDT}_{5+10} = [0.5, 1.0]$	$\lvert m_{\ell\ell} - m_Z \rvert < 10$ GeV	SFOS leptons
	$\text{BDT}_{20} = [0.7, 1.0]$		
	$\text{BDT}_{30} = [0.7, 1.0]$		
	$\text{BDT}_{40+50} = [0.7, 1.0]$		
	$\text{BDT}_{60+75} = [0.9, 1.0]$		
VR-SS-BDT	$\text{BDT}_{5+10} = [0.93, 1.0]$	Split by subleading e or μ	SS leptons
	$\text{BDT}_{20} = [0.7, 1.0]$	—	
	$\text{BDT}_{30} = [0.7, 1.0]$		
	$\text{BDT}_{40+50} = [0.7, 1.0]$		
	$\text{BDT}_{60+75} = [0.7, 1.0]$		

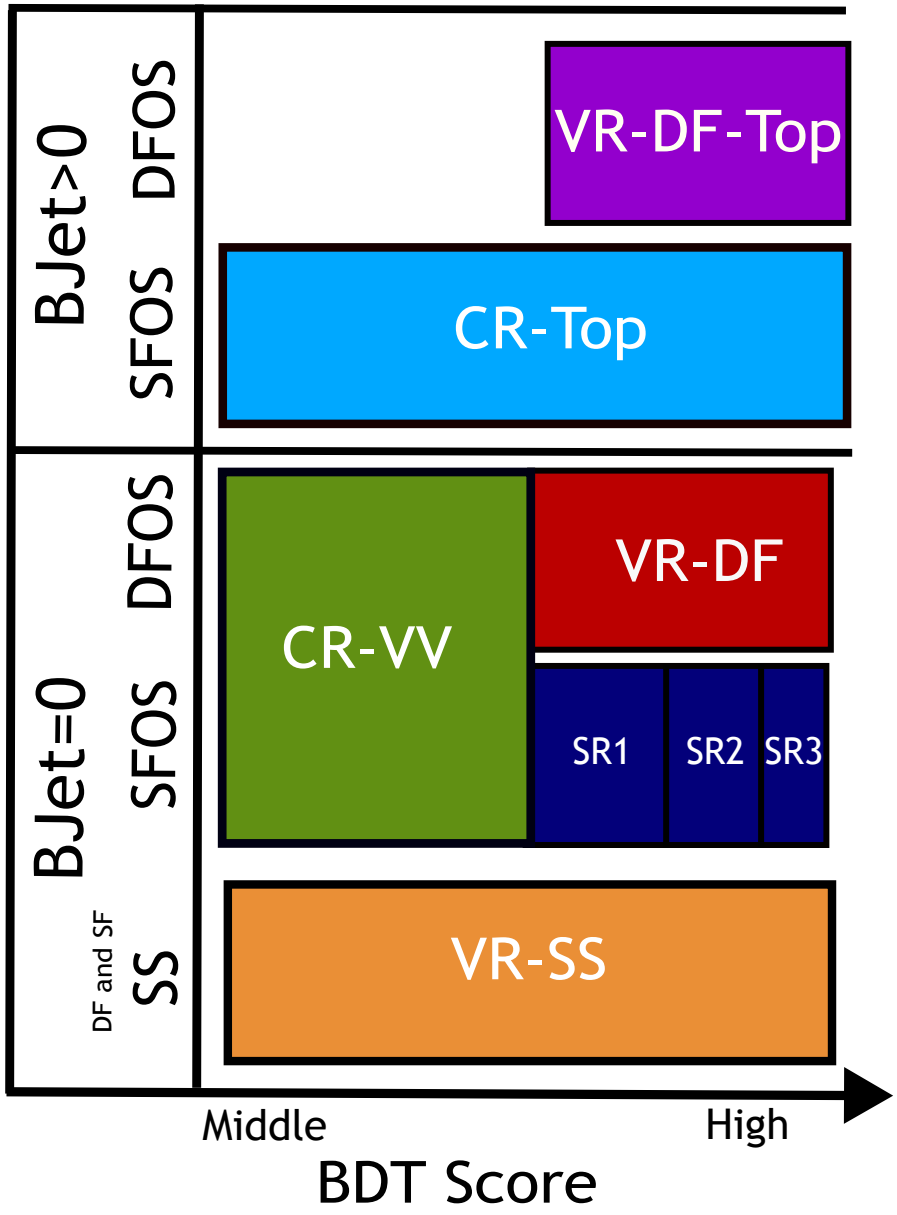


Figure 5.13: CR, VR and SR orthogonality. VR-ZZ is not shown, but is identical to the SR with the on-shell Z veto reversed.

5.5 Statistical Methods

The regions defined for the background estimate are used in a range of different statistical fits to interpret the results. Observed data in the CRs, VRs and SRs are compared with the SM predictions using a profile likelihood method [124] implemented in the HistFitter package [147]. Systematic uncertainties are incorporated as nuisance parameters with Gaussian constraints within the fit. The experimental systematic uncertainties are correlated between signal and background across all regions. Further details are given in Section 5.6 regarding the difference sources of these uncertainties. This section will be a brief introduction into the different statistical fits used in the analysis, and a more in depth look at the mathematics is given in Appendix C.

5.5.1 Fit types

The MC background predictions are constrained using statistical fits to data. After this, hypothesis tests are performed to evaluate the presence of potential signals by comparing the different hypotheses (e.g., background-only vs. signal-plus-background). These tests provide interpretations of the observations and set limits on new physics processes. In each fit region, the probability of observing X events, given a predicted yield Y, is modelled with a Poisson distribution. The overall likelihood is constructed as the product of these Poisson distributions across all regions. In addition, nuisance parameters are introduced to account for the many systematic uncertainties included in the analysis. These are allowed to vary from their nominal value (be pulled) during the fit, but large deviations from the nominal values are penalised. Correlations between nuisance parameters across different regions are fully accounted for. To discriminate between hypotheses, a one-sided profile log-likelihood test statistic is constructed, with the null and alternative hypotheses defined according to the test being performed. Unless otherwise stated, all fits are evaluated using the asymptotic approximation [124], which provides an analytical estimate of the test statistic distribution under the assumption of sufficiently large expected event yields in the fit regions.

This analysis utilises four different fits, each with its own purpose:

- CR-only fit – constrain uncertainties by extracting normalisation factors for major backgrounds.
- Exclusion fit – reject a signal hypothesis using CL_s method [148].
- Model-dependent discovery fit – determine the discovery significance of a signal hypothesis.
- Model independent discovery fit – determine the significance of an individual SR, without assuming a signal hypothesis.

A comparison between the four fit types is given in Figure 5.14. In each case, background normalisation factors (μ_{VV} and μ_{Top}) are extracted, and in all fits other than the CR-only, a signal normalisation factor (μ_{Sig}) is introduced.

In Section 5.7, the results of the CR-only fit are used to present the post-fit VR results, allowing an evaluation of the extrapolation of these extracted normalisation factors. The same fit is also used to obtain the post-fit SR distributions, providing the final comparison between the observed data and MC background predictions given for the SRs.

In Section 5.8, exclusion fits are used to provide exclusion limits across the targeted range of signal models. To exclude a signal model, ATLAS and CMS adopt a threshold of $CL_s < 0.05$, which corresponds to a significance of 1.64 σ.

In Section 5.8.2, model-dependent discovery fits are used to determine the statistical significance of the evaluated signal models, for SRs with excesses. The significance threshold for observation is 3 σ and for discovery the threshold is 5 σ

In Section 5.7.4, model-independent discovery fits are used to provide the statistical significance of the deviation between data and post-fit predicted background, individually for each SR.

Fit	Regions	Purpose	Signal model?	Null hypothesis	Alternative hypothesis	Outcome
CR Only	CRs	Constrain uncertainties of major backgrounds	No	—	—	Normalisation factors for major backgrounds
Exclusion	CRs + SRs	Determine whether a signal can be excluded	Yes	Background plus signal ($\mu_{Sig} = 1$)	Background only ($\mu_{Sig} < 1$)	CLs value for each signal
Discovery significance	CRs + SRs	Determine the significance for a signal model	Yes	Background only ($\mu_{Sig} = 0$)	Background plus signal ($\mu_{Sig} > 0$)	Discovery significance for each signal
Model independent significance	CRs + **One** SR	Determine the significance of a single signal region	No	Background only ($\mu_{Sig} = 0$)	Background plus signal ($\mu_{Sig} > 0$)	Single significance for each SR

Figure 5.14: A comparison between features of the four main fit types used in the analysis described in this chapter.

5.5.2 Toy Pseudo-experiments

In some cases, the SR3's defined in the analysis have a low number of predicted MC events, with $N_{\text{exp}} < 3$. In this regime, the asymptotic approximation may no longer be valid, and the fit results can become unreliable. To obtain accurate fit results, the test statistic is instead constructed using ensemble tests, performed via 'toy' MC simulations that sample from the full likelihood. Pseudo-datasets (toys) are generated for each SR by randomly sampling from a Poisson distribution, where the mean is defined as the expected event yield predicted by the nominal fit. For each toy dataset, the test statistic is recalculated, and this procedure is repeated many times to build up a full distribution of the test statistic under the hypothesis being tested. The p-value for a given hypothesis is then determined by comparing the observed value of the test statistic in data to the distribution obtained from the toys.

In this analysis, toys are utilised in the model independent discovery fits and the signal model discovery fits. The number of toys needs to be optimised to ensure sufficient statistics to accurately sample the full probability distribution. These tests are detailed in Appendix D, and led to the selection of 15,000 toys as the optimal number to guarantee convergence in all fits. Similar tests were carried for the exclusion fits, but only a small difference in CL_s was observed - also given in Appendix D, and so the asymptotic fit results are used for these.

5.6 Systematic Uncertainties

This section details the experimental and theoretical uncertainties implemented for the signal models and background estimate. Background and signal predictions from MC samples are subject to a set of uncertainties in the physics objects used in the analysis. The data-driven FNP lepton estimate has its own derived uncertainties. Theoretical uncertainties are also applied to the signals, as well as the top and diboson backgrounds. These arise from MC generation variations, and uncertainties related to the signal cross-section. The post-fit uncertainties for each of the SRs are summarised in the results section – Section 5.7.3.

The uncertainties are incorporated in the fits as nuisance parameters, each described by a gaussian centred on zero and with width of one. Zero corresponds to the nominal yield in all regions, while the ± 1 correspond to the 'up' and 'down' systematic variations of the nominal yield.

5.6.1 Experimental Uncertainties

The objects used in the analysis come with associated experimental uncertainties for both the signal and background MC. Uncertainties in electrons [149] and muons [51] arise from uncertainties in the corrections applied in MC simulation of their reconstruction, identification, and isolation efficiencies, as well as in the corrections to their momentum resolution and energy scale.

The jet energy resolution and scale uncertainties (JER and JES) are applied as a function of the p_{T} and η of the jet [150, 151].

The JES uncertainties consist of multiple components that account for different sources of systematic variation, such as detector response, pile-up, flavour composition, and jet topology. These components are derived from both simulation and in-situ calibration studies, using dijet, $Z+$ jet and $\gamma+$ jet events. To make the number of NPs more manageable for analyses, an eigenvector decomposition is performed, reducing the total number of NPs to 30.

For the JER uncertainties, the calibration fit to di-jet events results in 117 NPs that parameterise the uncertainty on the resolution. As with the JES uncertainties,

eigenvector decomposition is applied to reduce the number of NPs, and this analysis uses the reduced $6 + 1$ NP configuration. An additional NP is included to take into account the uncertainty in the residual smearing applied to correct for observed differences between data and MC after the JER calibration.

Further jet-related uncertainties originate from efficiency corrections for tagging pile-up jets [44] and the identification of b-jets [47, 152]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties are estimated by propagating the uncertainties in the energy and momentum scale of each of the objects entering the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$, and include the uncertainties in the soft-term resolution and scale [56].

These same set of experimental uncertainties are also implemented for the signals, with the only exception being a few dedicated jet related AF2 uncertainties, due to the different treatment of calorimeter simulation in the signal generation.

Additionally, an uncertainty on the integrated luminosity is included: $\pm 0.83\%$ for the combined Run 2 dataset [7]. An uncertainty is also included related to the average pileup profile, $\langle\mu\rangle$, used in the MC simulation. The $\langle\mu\rangle$ distribution in MC does not necessarily agree with the one observed in data. To account for this, an improved pile-up description is achieved by re-weighting the MC $\langle\mu\rangle$ distribution, to the distribution in data. An uncertainty on this re-weighting procedure is considered through $\pm 1\sigma$ variations on this pile-up weight.

Fake/Non-prompt Uncertainties

The uncertainty in the FNP background estimate based on the fake-factor method has several components. Uncertainties in the fake factors originate from the size of the data sample in which these are derived and further kinematic dependencies than the ones already considered. Uncertainties in the subtraction of the prompt-lepton contamination were estimated by varying their size and are found to be small compared with the other uncertainties. Additional statistical uncertainties arise from the limited size of the loose-lepton control samples. If the control sample does not contain any events, an additional one-sided uncertainty is applied based on the largest fake factor applicable to provide an upper limit on the FNP contribution.

Finally, a non-closure uncertainty is derived to account for remaining discrepancies observed in VR-SS, which are not fully covered by the uncertainties described above, using the difference in the data/MC ratio. This uncertainty is evaluated separately for electrons and muon and is parameterised as a function of the lepton p_{T}. The largest non closure observed across any of the VR-SS's (BDT or cut-and-count) is taken as the uncertainty for a given subleading lepton p_{T} bin.

Experimental Systematic Symmetrisation

Some experimental uncertainties were seen to have 'up' and 'down' variations that led to one-sided variations in the SR yields, due to statistical fluctuations in the MC statistics of the variation samples. This led to the fit results becoming unstable, especially when using toys, and so these uncertainties were manually symmetrised. This was done in all regions, for all MC backgrounds and signals, by taking the root mean squared (RMS) of any two same sided variations, then using this to define symmetric up, down variations. The overall effect of this change on the SR post-fit uncertainties was negligible, and the significances/CL_s values obtained from this new setup were seen to be stable. A more in-depth discussion regarding this fit behaviour and the fix is in Appendix D.

5.6.2 Theoretical Uncertainties

Theory uncertainties are evaluated on the two dominant backgrounds: top and diboson, and on the signal samples. For each of these, the scale, PDF and α_s variations are considered by using the internal on-the-fly weights provided by the MC generator. Information regardinging the role of these different in MC generation is given in Appendix A.

In the case of the top backgrounds, additional generator uncertainties are considered, and the signal models have additional cross-section uncertainties. The 'other', minor backgrounds, have a conservative flat 50% uncertainty assigned to cover any cross-section mis-modelling.

The uncertainties are derived by taking the relative variation with respect to the nominal samples as the value of the uncertainty. These are implemented in the fit as two-sided uncertainties for both signal and the major backgrounds. The variations for the $t\bar{t}$ and single-top processes are treated as correlated within the overall 'top' background.

Background Theory Uncertainties

Uncertainties related to MC generator modelling of the diboson and top backgrounds related to the choice of the QCD factorisation (μ_F) and renormalisation (μ_R) scales are evaluated. The scale uncertainties are assessed by varying the associated generator parameters up and down by a factor of two around their nominal values, avoiding combinations that differ by a factor of 4. Permutations for the factorisation and renormalisation scales are shown in Table 5.9. The PDF variations are assessed by generating the MC using a different PDF to the nominal. The α_s uncertainties are taken as a two-point uncertainty from the nominal PDF set. The envelope defined by the largest upward and downward deviations from the nominal prediction is taken as a two-sided systematic uncertainty on the given background.

As an example, Figure 5.15 shows the top and diboson scale variations in CRs and SR1 only, for BDT_{40+50}. Generally, the scale uncertainty ranges from $10\% - 25\%$, with an average of about 20%, and are consistent across BDT SRs. The top scale uncertainty is larger than the diboson across all regions. The PDF and α_s uncertainties are not included in the final fit for the backgrounds, as they are significantly smaller than the scale uncertainties, $< 1\%$.

Table 5.9: 7-point rescaling scheme. Events are weighted with rescaled μ_R, μ_F in pairs shown above, where $(1, 1)$ corresponds to the nominal choice

μ_R	0.5	1	0.5	1	2	1	2
μ_F	0.5	0.5	1	1	1	2	2

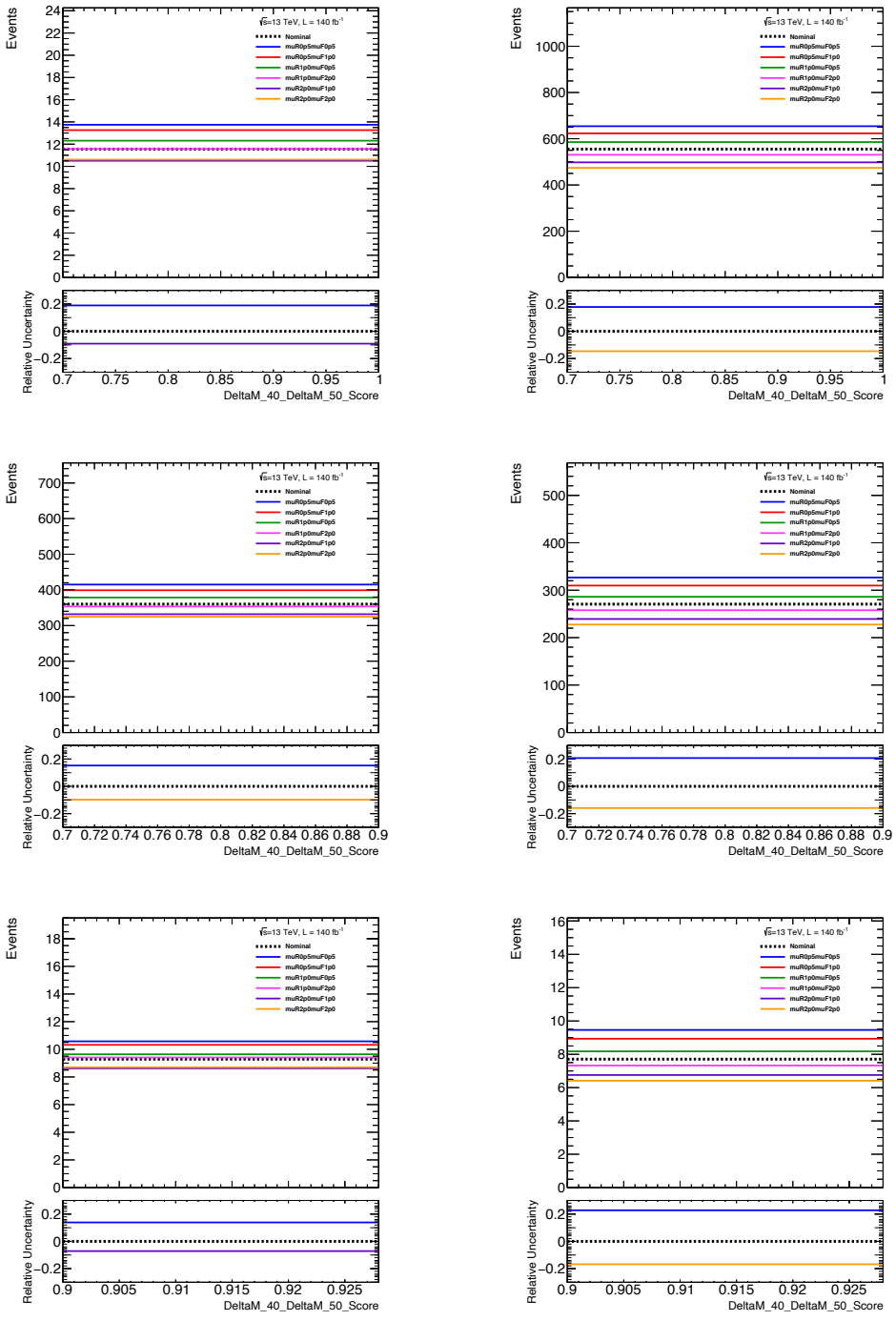


Figure 5.15: The scale uncertainty variations, with the relative uncertainty ($r^{\text{up/down}} - 1$) as a subplot, for the BDT_{40+50} regions. Shown for the (left) diboson and (right) top backgrounds, in (top) CR-Top, (middle) CR-VV, and (bottom) SR1. The legend can be interpreted as for example: muR0p5 – the renormalisation scale set to 1/2× the nominal, and muF2p0 – the factorisation scale set to 2× the nominal. The nominal is equivalent to muR1p0muF1p0.

.

Additional Top Generator Systematics

Additional uncertainties are evaluated for the top background, by evaluating variations of MC generator parameters. Aspects of the generation of these samples are varied from the nominal method and are calculated using all six of a given BDTs SRs together to reduce the impact statistical uncertainties of the generated MC. The one-sided variation in top background yields from this procedure is then symmetrised to give the final two-sided systematic uncertainty.

The variations are as follows:

- Variation of the parton shower (PS) – Pythia as the nominal, versus Herwig as the variation PS [153, 154].
- Variation in the matching scheme between the matrix element and the parton shower (p_{T}-Hard) – Powheg Box as the nominal, versus Pythia as the variation [155].
  - p_{T}-Hard = 0 is the default – where Pythia takes the p_{T} matching scale as the p_{T} of the parton, as generated by Powheg Box.
  - p_{T}-Hard = 1 is the variation – where Pythia takes the p_{T} matching scale as the minimum p_{T} that the parton could have been generated in Powheg Box - so the showering begins at a lower scale than the nominal.
- Treatment of $t\bar{t}$ and Wt interference, – comparison between the diagram removal (DR) method as the nominal [156], and the diagram subtraction (DS) method as the variation [157].

In some cases, the top p_{T}-Hard variation direction flips direction between CR and VRs/SRs, as shown in Figure 5.16 for BDT_{40+50}. This was determined to be due to migration of b-jets below the 20 GeV jet reconstruction threshold, leading to events migrating into the zero b-jet bin, under the p_{T}-Hard variation. A more in-depth look at this behaviour is given in Appendix E. For these systematics,

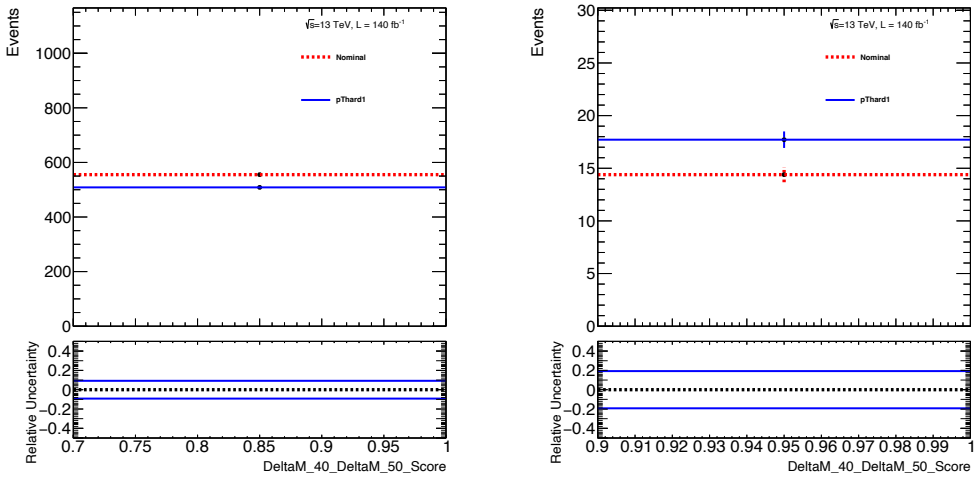


Figure 5.16: Top p_{T}-Hard variation (blue line), and the nominal (red dotted line) in (a) CR-Top and (b) combined SRs for BDT_{40+50}. The subplot shows the symmetrised relative uncertainty with respect to the nominal.

the 'sign-flip' information was preserved in the fit by fixing the direction of the up/down variations in the fit. So for example, in the above example, the (up, down) variation for CR-Top would be (1.1,0.9), whereas the SR variation would be (0.8,1.2). This introduces an anti-correlation between the two regions in the fit, and leads to a small increase post-fit uncertainties in the SR for the top background. The impact on the final region uncertainties was not significant, as the top background is subdominant in most SR bins.

Generally, the $t\bar{t}$ and Wt interference is negligible as single top is only a small proportion of the overall top uncertainty. The parton shower variation is largest in regions that require b-jets, ranging from 20–40% across those analysis regions. The p_{T}-Hard variation is largest in regions that veto b-jets, ranging from 10–20% across those analysis regions.

Signal Theory Uncertainties

The predictions of the signal samples are also subject to a set of theoretical uncertainties. The scale, PDF and α_S, are assessed following the same procedure that is used for the diboson and top backgrounds, and also implemented in the fit as symmetric two-sided uncertainties.

Average values were between 20–30% for the scale variations, and only a mild dependence on $m(\tilde{\ell})$ and $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$ was found. As a result, for each mass splitting, a representative signal model was chosen to define the scale uncertainty for all signals in that mass splitting. Table 5.10, shows the representative signal chosen, and the final scale uncertainties that apply to all signal samples that $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$. Very little variation across signal samples was seen for the PDF uncertainty, and so this was set at 3.5% for all signals. Uncertainties in the value of α_s were found to be negligible for the signal processes and not included in the final set of uncertainties.

In addition, the uncertainties on the slepton production cross-sections are taken into account by rerunning the exclusion fits, with $\pm 1\sigma$ variations on the nominal cross-section. In figures showing the final exclusion limits, the results of these fits are then used to draw (dotted red) lines corresponding to the theory uncertainty on the observed exclusion limits.

Table 5.10: Representative signal taken for each mass splitting and the associated scale uncertainty, given as a percentage.

Mass splitting	Signal $(m(\tilde{\ell}, \widetilde{\chi}_1^0))$	Scale uncertainty (%)
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 2$	$(300, 298)$ GeV	22.4%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 3$	$(300, 297)$ GeV	22.1%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 4$	$(300, 296)$ GeV	21.3%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 5$	$(300, 295)$ GeV	21.1%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 10$	$(300, 290)$ GeV	20.9%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 20$	$(250, 230)$ GeV	20.7%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 30$	$(250, 220)$ GeV	20.6%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 40$	$(250, 210)$ GeV	21.2%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 50$	$(250, 200)$ GeV	20.3%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 60$	$(250, 190)$ GeV	20.5%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 75$	$(250, 175)$ GeV	19.8%
$\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 100$	$(300, 200)$ GeV	19.0%

5.7 Results

Observed data yields in the CRs, VRs and SRs are compared with the SM predictions using the statistical fit methods described in Section 5.5. This section summarises the results of the CR-only fit, and the subsequent extrapolation of the extracted normalisation factors to the relevant VRs and SRs. In addition, the calculation of model-independent individual SR significances is described in Section 5.7.4. Investigations into the observed excesses are discussed in Section 5.7.5.

5.7.1 Control Regions

Each BDT has a dedicated background estimate, and therefore its own set of control regions and normalisation factors. The normalisation factors extracted from each set of CRs are summarised in Table 5.11. The variation in the diboson and top normalisation factors across the various BDTs, originates from the different phase spaces covered by each BDT selection. The pre-fit CR yields, and normalisation factors are summarised in Figure 5.17. Each CR-VV was checked for signal contamination and this was calculated to be $< 10\%$ in all cases.

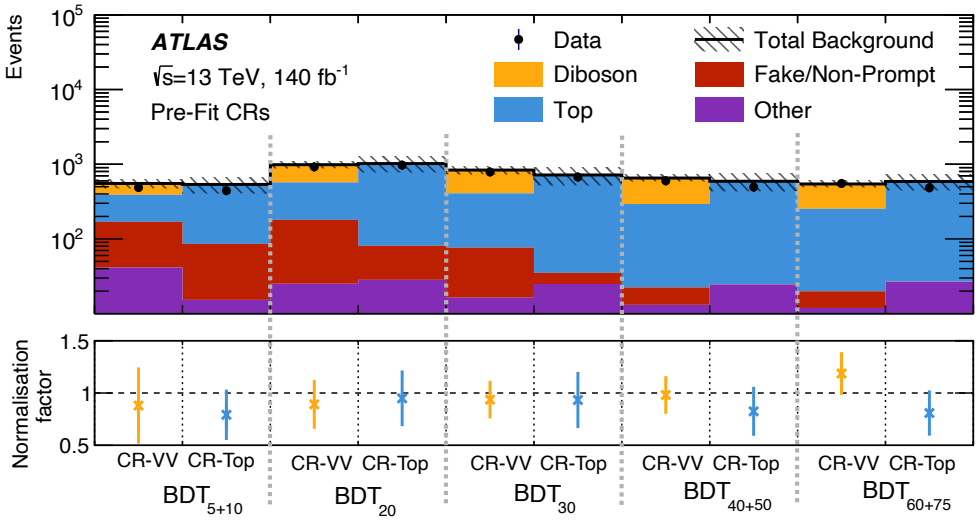


Figure 5.17: Data (dots) and pre-fit expected (histograms) event yields in the CRs. Shown in the lower panel are the normalisation factors (crosses), for μ_{VV} (orange) and μ_{Top} (blue) derived from a CR-only fit, along with the associated uncertainties. Uncertainties in the background estimates include both statistical and systematic uncertainties.

Table 5.11: Normalisation factors μ_{VV} and μ_{Top} for the VV and top backgrounds, respectively, extracted from CR-only fits. The associated uncertainties include all statistical and systematic contributions.

Normalisation factor	μ_{VV}	μ_{Top}
BDT_{5+10}	0.88 ± 0.36	0.79 ± 0.24
BDT_{20}	0.89 ± 0.23	0.95 ± 0.27
BDT_{30}	0.94 ± 0.18	0.93 ± 0.27
BDT_{40+50}	0.98 ± 0.18	0.82 ± 0.23
BDT_{60+75}	1.19 ± 0.20	0.82 ± 0.22

5.7.2 Validation Regions

Figure 5.18 shows the data and the post-fit SM predictions in the BDT VRs, from the CR-only fits. Good agreement is found between observed data and post-fit background in all VRs, where all deviations are found to be below two standard deviations as approximated with the prescription from Ref. [158]. The largest discrepancy is in VR-ZZ-BDT_{60+75}, due to the VV normalisation factor from CR-VV-BDT_{60+75}, which is dominated by W^+W^- and does not extrapolate well to the higher ZZ fraction in the VR.

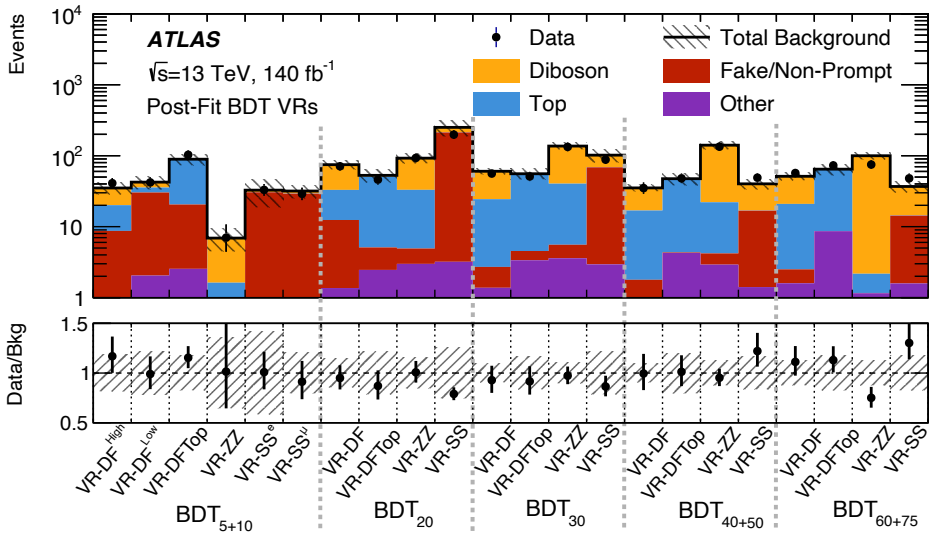


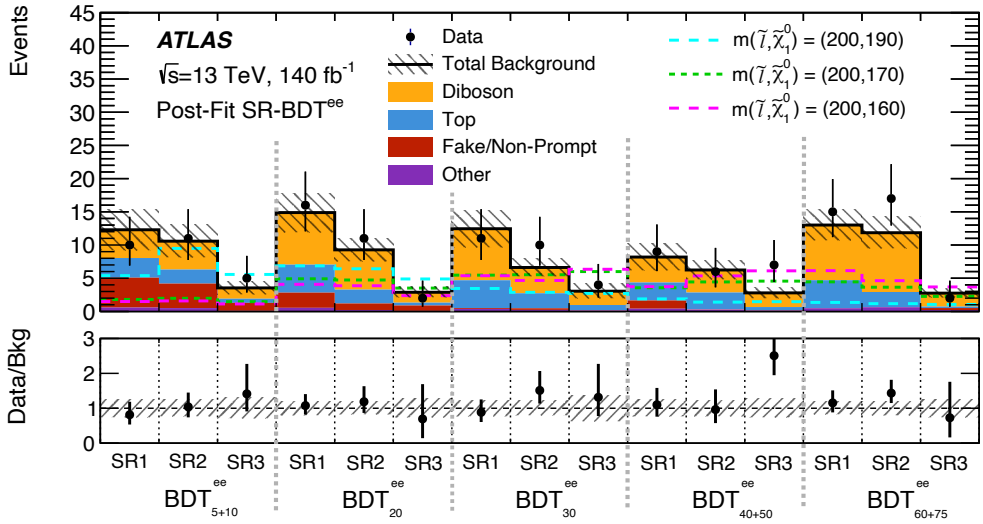
Figure 5.18: Data (dots) and post-fit SM predictions (histograms) in the VRs using a CR-only fit. The lower panel shows the ratio of observed data to the total post-fit SM prediction. Uncertainties in the background estimates include both statistical and systematic uncertainties.

5.7.3 Signal Regions

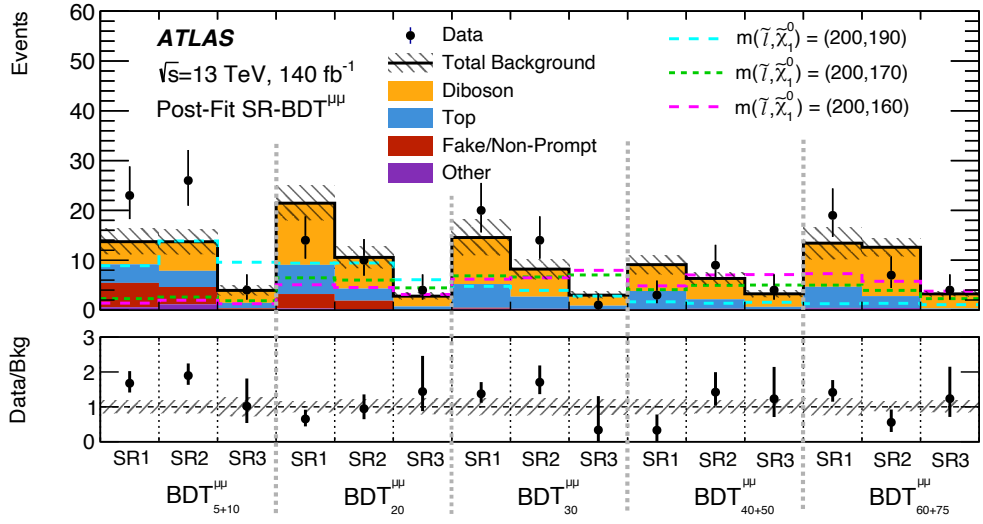
The observed data and predicted SM event yields for the SRs, from the CR-only fits are shown in Figure 5.19, along with the ratio of the observed data to the post-fit SM expectation. While in most SRs the SM predictions agree well with the data, a few bins show excesses of around two standard deviations. These are further discussed in Section 5.7.5.

Figure 5.20, contains a breakdown of the SR uncertainty, following a CR-only fit. The breakdown into individual categories is derived by iteratively performing fits where only the parameters associated with a particular category are allowed to float with all other parameters set constant. The uncertainties are then propagated to the background predictions using the covariance matrix of the nominal fit where all parameters are allowed to float [147]. The individual uncertainties are correlated and do not necessarily add up in quadrature to the total uncertainty.

The leading uncertainties vary across the BDTs as the background composition changes, but their overall magnitude remains similar. Overall, the leading experimental uncertainties in the background estimates were found to originate from the jet energy resolution, as well as the resolution and scale of the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term. The dominant MC-modelling uncertainties are the QCD scale variations and those due to the matching between matrix element and parton shower for the top background. For SR-BDT$_{5+10}$, the uncertainties related to the extrapolation of normalisation factors dominate, due to the relatively large FNP lepton contamination in the CRs. The contribution of these uncertainties decreases as $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ increases, as the FNP background contribution reduces. The overall SR uncertainty ranges from 20–35% in the di-electron SRs and 15–25% in the di-muon SRs. Generally the electron SRs have higher uncertainties, as electron reconstruction is more difficult than muon reconstruction especially at low p_{T}. In addition, there is a larger FNP lepton contribution to the electron CRs, and hence higher post-fit uncertainties.

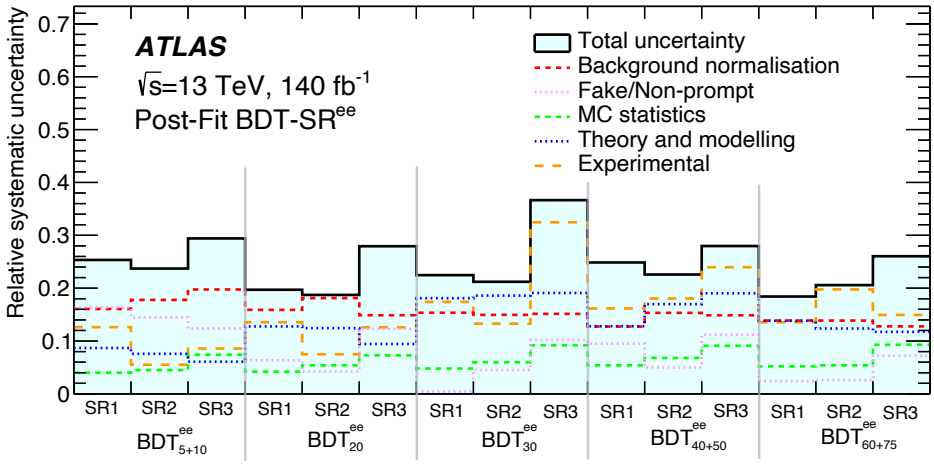


(a)

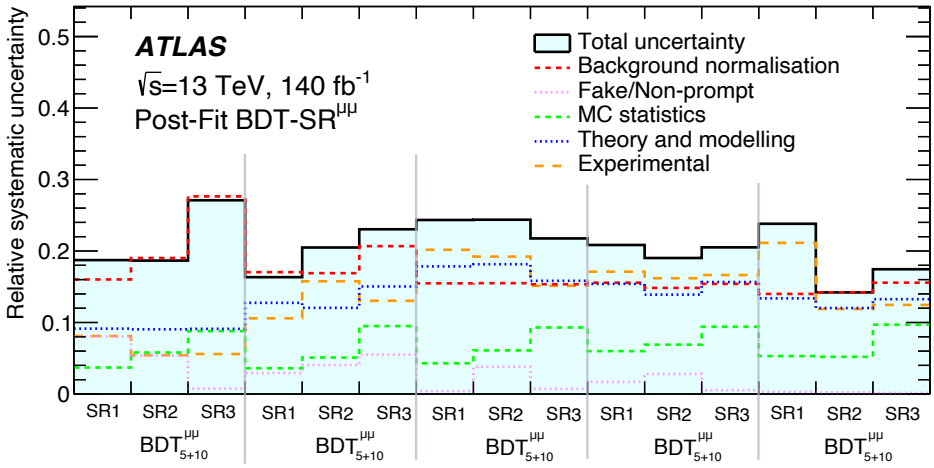


(b)

Figure 5.19: Data (dots) and post-fit SM predictions (histograms) in (a) SR-BDT^{ee} and (b) SR-BDT$^{\mu\mu}$. The SM predictions are obtained from a CR-only fit. Uncertainties in the background estimates include both statistical and systematic uncertainties. The bottom panel shows the ratio of observed data to the total post-fit SM prediction. The dashed lines show predicted yields for three benchmark signal models corresponding to $m(\tilde{\ell}, \tilde{\chi}_1^0) = (200, 190)$ GeV, $(200, 170)$ GeV and $(200, 160)$ GeV, assuming $m(\tilde{\ell}_L) = m(\tilde{\ell}_R)$.



(a)



(b)

Figure 5.20: Relative systematic uncertainties in the post-fit SM background estimates in the (a) di-electron and (b) di-muon SRs obtained from a background-only fit to the CRs. The "Fake/Non-Prompt" category reflects uncertainties impacting the FNP background estimate. Uncertainties originating from the limited size of the MC samples used to model the irreducible background contributions are contained in the "MC statistics" category. The "Background normalisation" category reflects uncertainty in normalisation factors for VV and top backgrounds extracted from CR-Top-BDT and CR-VV-BDT. The "Theory and modelling" category includes the different sources of theoretical modelling uncertainties for the top and diboson backgrounds. The "Experimental" category covers detector related uncertainties from the reconstruction and selection of objects in the analysis.

5.7.4 Model Independent SR Significances

Table 5.12 summarises the observed and total post-fit SM expectation from the CR only fit for each BDT SR. Also shown is the statistical significance of the difference between observed data and the post-fit predicted yield from the CR-only fit, for each SR. The significances are determined using a discovery fit, and using 15,000 toy pseudo-experiments. The CRs are fit simultaneously with one SR at a time, allowing for the extraction of an individual single-bin significance for each SR.

In the di-electron channel the most discrepant region is SR3-BDT$^{ee}_{40+50}$, with an excess of events that corresponds to a local significance of $2.0\,\sigma$. Excesses of smaller statistical significance are observed in SR2-BDT$^{ee}_{30}$ and SR2-BDT$^{ee}_{60+75}$. As these SRs are not statistically independent, the latter excesses are caused to a large extent by the same events that are also found in SR3-BDT$^{ee}_{40+50}$.

In the di-muon channel, the most discrepant SRs are SR1-BDT$^{\mu\mu}_{5+10}$ and SR2-BDT$^{\mu\mu}_{5+10}$, where the excesses have statistical significances of $1.9\,\sigma$ and $2.4\,\sigma$, respectively. By design, there is no overlap between these two SRs. There is also minimal overlap between events in these SR-BDT$^{\mu\mu}_{5+10}$ regions and those in other BDT SRs. This is because, the SR-BDT$_{5+10}$ contains a large proportion of events with low-p_{T} leptons, which is a feature not present in any other SRs. The largest data event overlap occurs with the signal regions of the neighbouring BDT, SR-BDT$^{\mu\mu}_{20}$, which shares $\approx$ 15% of events with SR-BDT$^{\mu\mu}_{5+10}$. Investigations into these excesses follow in the next section.

Table 5.12: Observed and the total post-fit SM expectation, along with the statistical significance of their deviation, shown separately for (left) SR-BDT^{ee}and (right) SR-BDT$^{\mu\mu}$. The expected total SM background event yields, N_{exp}, and the observed event yields, N_{obs}, are shown in the second and third columns of the table. The SM expectation is calculated from a simultaneous fit of only the CRs from the relevant BDT. The fourth column shows the statistical significance of the deviation between the expected total SM prediction and observed data, calculated using individual discovery fits with 15,000 pseudo-experiments. Cases where the calculated significance is negative are indicated by a dash, '–', and no significance is reported.

Region (ee)	N_{exp}	N_{obs}	σ
SR1-BDT$^{ee}_{5+10}$	12.3 ± 3.1	10	–
SR2-BDT$^{ee}_{5+10}$	10.6 ± 2.5	11	0.17
SR3-BDT$^{ee}_{5+10}$	3.5 ± 1.0	5	0.59
SR1-BDT$^{ee}_{20}$	14.9 ± 2.9	16	0.28
SR2-BDT$^{ee}_{20}$	9.3 ± 1.7	11	0.44
SR3-BDT$^{ee}_{20}$	2.9 ± 0.8	2	–
SR1-BDT$^{ee}_{30}$	12.5 ± 2.8	11	–
SR2-BDT$^{ee}_{30}$	6.6 ± 1.4	10	1.0
SR3-BDT$^{ee}_{30}$	3.0 ± 1.1	4	0.53
SR1-BDT$^{ee}_{40+50}$	8.2 ± 2.0	9	0.26
SR2-BDT$^{ee}_{40+50}$	6.2 ± 1.4	6	–
SR3-BDT$^{ee}_{40+50}$	2.8 ± 0.8	7	2.0
SR1-BDT$^{ee}_{60+75}$	13.0 ± 2.4	15	0.69
SR2-BDT$^{ee}_{60+75}$	11.9 ± 2.4	17	1.2
SR3-BDT$^{ee}_{60+75}$	2.8 ± 0.7	2	–

Region ($\mu\mu$)	N_{exp}	N_{obs}	σ
SR1-BDT$^{\mu\mu}_{5+10}$	13.7 ± 2.6	23	1.9
SR2-BDT$^{\mu\mu}_{5+10}$	12.9 ± 2.4	26	2.4
SR3-BDT$^{\mu\mu}_{5+10}$	3.9 ± 1.1	4	–
SR1-BDT$^{\mu\mu}_{20}$	21.5 ± 3.5	14	–
SR2-BDT$^{\mu\mu}_{20}$	10.6 ± 2.2	10	–
SR3-BDT$^{\mu\mu}_{20}$	2.8 ± 0.6	4	0.68
SR1-BDT$^{\mu\mu}_{30}$	14.6 ± 3.5	20	1.1
SR2-BDT$^{\mu\mu}_{30}$	8.2 ± 2.0	14	1.4
SR3-BDT$^{\mu\mu}_{30}$	3.0 ± 0.6	1	–
SR1-BDT$^{\mu\mu}_{40+50}$	9.1 ± 1.9	3	–
SR2-BDT$^{\mu\mu}_{40+50}$	6.3 ± 1.2	9	1.0
SR3-BDT$^{\mu\mu}_{40+50}$	3.2 ± 0.7	4	0.46
SR1-BDT$^{\mu\mu}_{60+75}$	13.4 ± 3.2	19	1.1
SR2-BDT$^{\mu\mu}_{60+75}$	12.6 ± 1.8	7	–
SR3-BDT$^{\mu\mu}_{60+75}$	3.2 ± 0.6	4	0.07

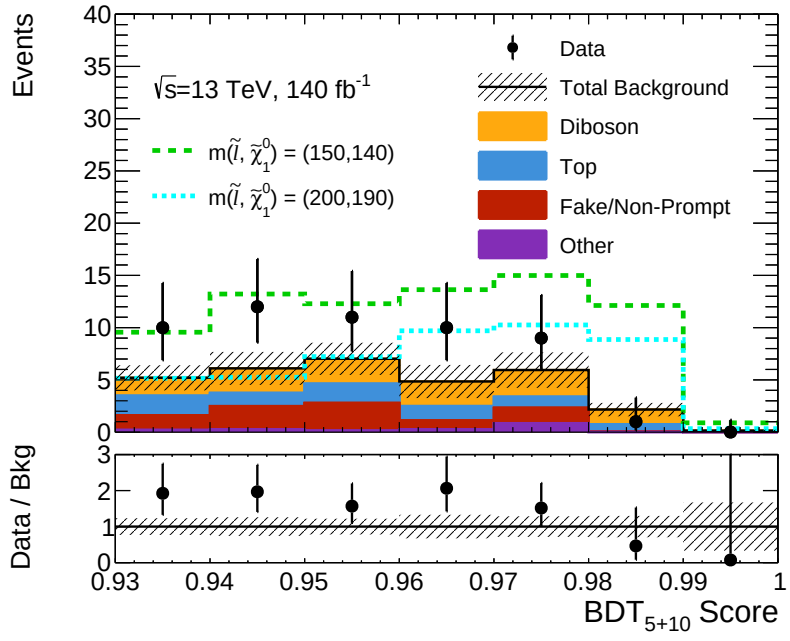
5.7.5 SR-BDT$^{\mu\mu}_{5+10}$ Investigations

For the di-muon analysis, an excess is observed in two adjacent SRs for SR-BDT$^{\mu\mu}_{5+10}$. The BDT score distribution across the three SRs is shown in Figure 5.21. For the BDT score range below the SRs, good agreement is seen in all bins. Kinematic distributions of variables important in the training of this BDT are shown in Figure 5.22, inclusively for all three bins of SR-BDT$^{\mu\mu}_{5+10}$. Overlaid are kinematic distributions of two example signal models: $(m_{\tilde{\mu}}, m_{\tilde{\chi}^0_1}) = (150\ \text{GeV}, 140\ \text{GeV})$ and $(200\ \text{GeV}, 190\ \text{GeV})$.

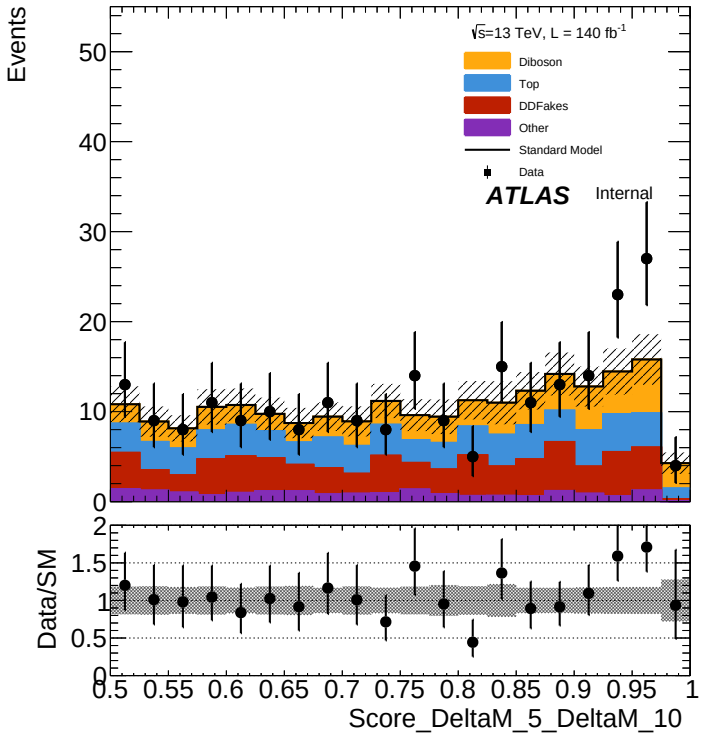
An excess of events appear to accumulate at $\Delta R_{\ell\ell}$ values around 3.0. These are back-to-back in ϕ, and for $20 < m_{\ell\ell} < 30$ GeV. No similar features are observed in

SR-BDT$^{ee}_{5+10}$, VR-DF-BDT$^{Low}_{5+10}$, or VR-SS-BDT$^{\mu}_{5+10}$, as shown in Figures 5.23, 5.24 and 5.25 respectively. Extensive checks were carried out the excess SRs:

- No yearly trend was seen in the data events – see Appendix F.3.
- Events were checked for downward-going muon tracks to test for cosmic muons.
- The η, ϕ maps of the jets and leptons showed no pattern to detector occupancy – see Appendix F.3.
- Muon isolation and reconstruction information was checked and seen to be consistent with correctly reconstructed leptons. Additionally, the majority of muons are combined muons.
- Minimal hadronic activity is observed near the reconstructed muons, and the impact parameters are small, and so muons produced from the decay of b-hadrons are unlikely.
- The angular relationship between leptons and jets was checked, to check for muons from hadron decays – see Appendix F.3.
- When the $p_{\mathrm{T}}^{\ell_2}$ requirement is lowered to 5 GeV, the 5−6 GeV $p_{\mathrm{T}}^{\ell_2}$ bin shows good agreement between data and SM prediction. This indicates the discrepancy is unlikely to be due to an issue with the FNP lepton estimate.
- Contributions from Drell-Yan and low-$m_{\ell\ell}$ Z+jets processes were checked and found to be negligible.



(a)



(b)

Figure 5.21: Data (dots) and post-fit SM background (histograms) distribution in (a) for the BDT score across all three regions of SR-BDT$^{\mu\mu}_{5+10}$ using the CR-only fit, with example signal models overlayed. In (b), the BDT score distribution of the SRs, and the CR-VV is shown, again using the CR-Only fit. The uncertainty bands include statistical and systematic uncertainties. The SRs are defined: SR1-BDT$^{ee}_{40+50}$ from 0.9 – 0.928, SR2-BDT$^{ee}_{40+50}$ from 0.928 – 0.956, and SR3-BDT$^{ee}_{40+50}$ from 0.956 – 1.0

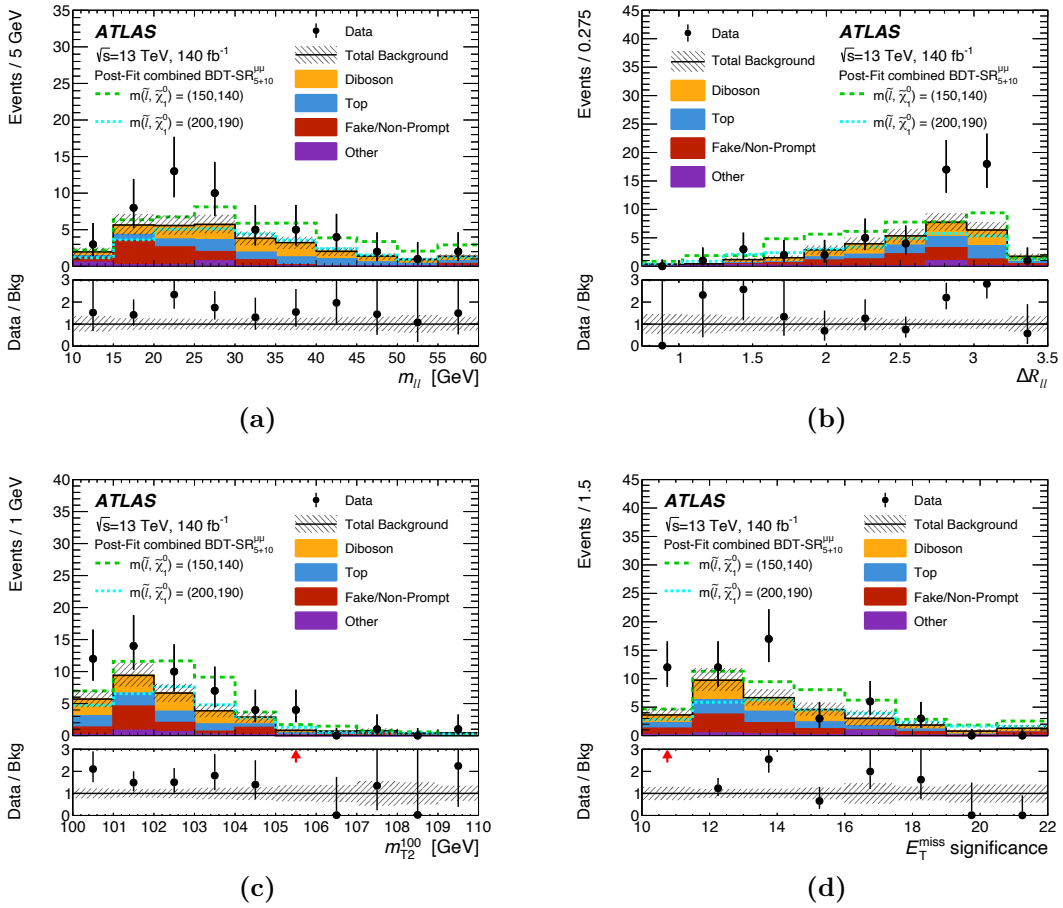


Figure 5.22: Data (dots) and post-fit SM (histograms) distributions for (a) $m_{\ell\ell}$, (b) $\Delta R_{\ell\ell}$, (c) m_{T2}^{100} and (d) the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance in the combined SR-BDT$_{5+10}^{\mu\mu}$. The uncertainty bands include statistical and systematic uncertainties. The first and last bin include the under- and overflow (if applicable), respectively. The lower panel shows the ratio of observed data to the total post-fit SM prediction. The dashed lines show predicted yields for two benchmark signal models corresponding to $m(\tilde{\ell}, \tilde{\chi}_1^0) = (150, 140)$ GeV and $(200, 190)$ GeV, assuming $m(\tilde{\ell}_{\mathrm{L}}) = m(\tilde{\ell}_{\mathrm{R}})$. A red arrow in the lower panel indicates data points outside the vertical range shown.

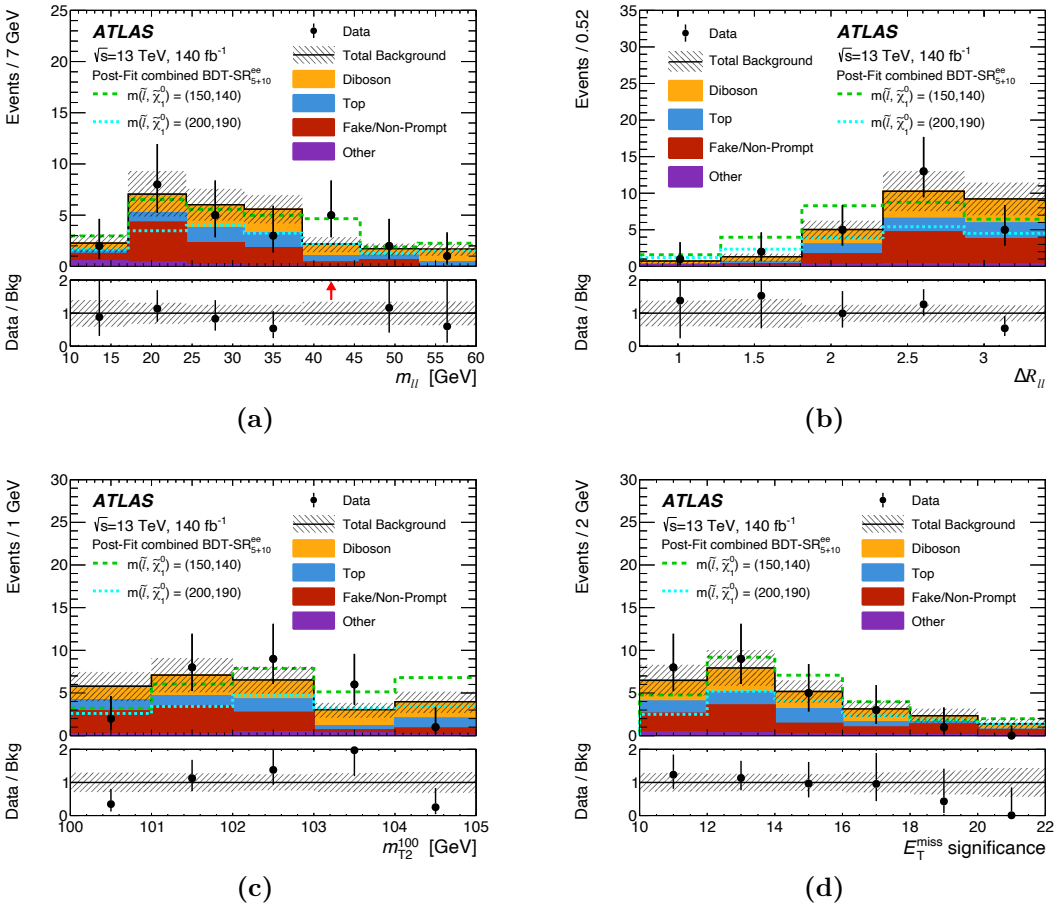


Figure 5.23: Data (dots) and post-fit SM background (histograms) distributions for (a) $m_{\ell\ell}$, (b) $\Delta R_{\ell\ell}$, (c) m_{T2}^{100} and (d) the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance in the combined SR-BDT$_{5+10}^{\mathrm{ee}}$ using the CR-only fit. Two example SUSY scenarios are overlaid for illustration that correspond to $m(\tilde{\ell}, \tilde{\chi}_1^0) = (150, 140)$ GeV and $(200, 190)$ GeV, assuming $m(\tilde{\ell}_{\mathrm{L}}) = m(\tilde{\ell}_{\mathrm{R}})$. The uncertainty bands include statistical and systematic uncertainties. The first and last bin include the under- and overflow (if applicable), respectively. A red arrow in the lower panel indicates data points outside the vertical range shown.

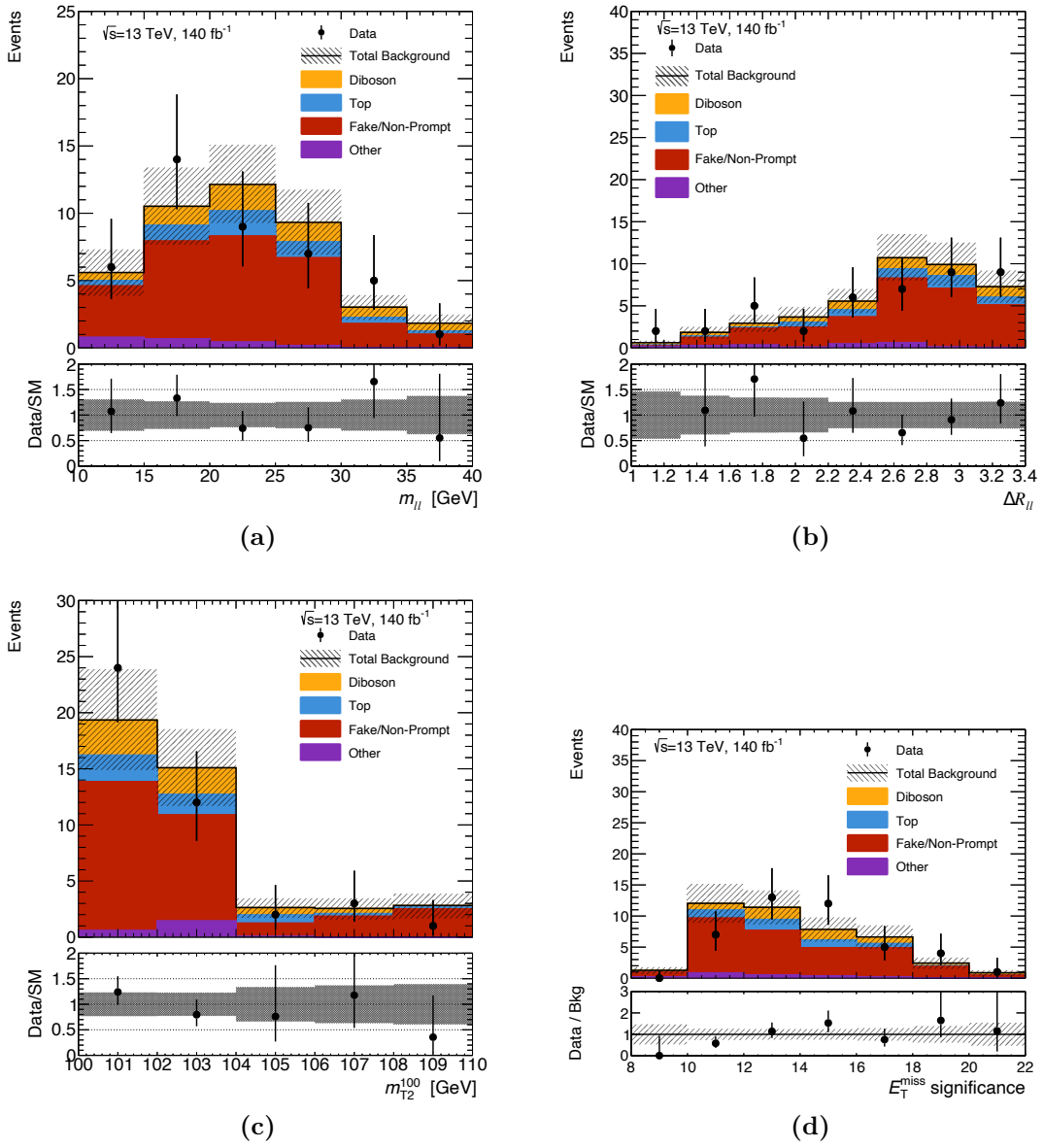


Figure 5.24: Data (dots) and post-fit SM background (histograms) distributions for (a) $m_{\ell\ell}$, (b) $\Delta R_{\ell\ell}$, (c) m_{T2}^{100} and (d) the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance in VR-DF-BDT$_{5+10}^{\mathrm{Low}}$ the using the CR-only fit. The uncertainty bands include statistical and systematic uncertainties. The first and last bin includes the under- and overflow (if applicable), respectively. A red arrow in the lower panel indicates data points outside the vertical range shown.

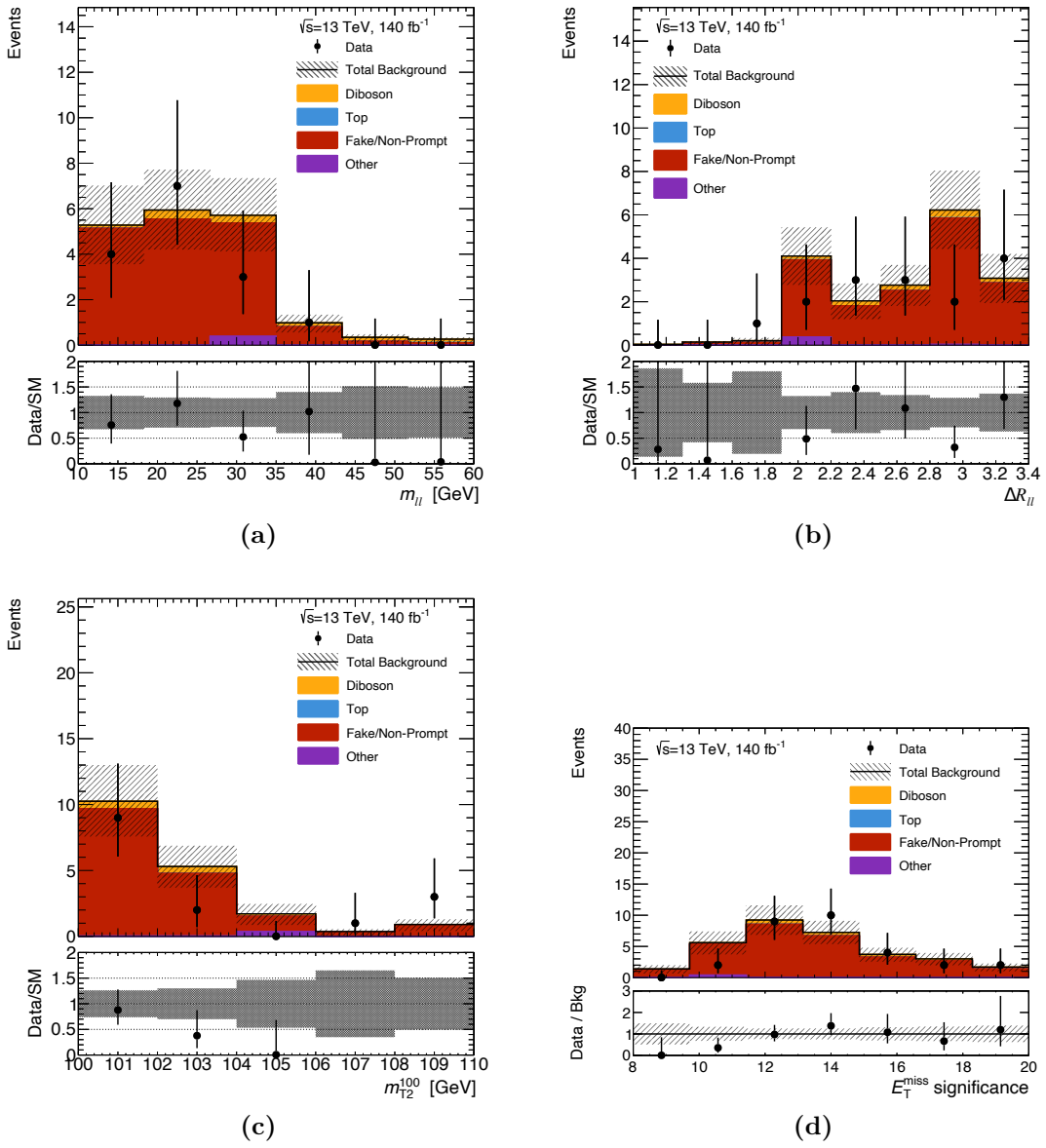


Figure 5.25: Data (dots) and post-fit SM background (histograms) distributions for (a) $m_{\ell\ell}$, (b) $\Delta R_{\ell\ell}$, (c) m_{T2}^{100} and (d) the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance in VR-SS-BDT$^{\mu}_{5+10}$ using the CR-only fit. The uncertainty bands include statistical and systematic uncertainties. The first and last bin includes the under- and overflow (if applicable), respectively. A red arrow in the lower panel indicates data points outside the vertical range shown.

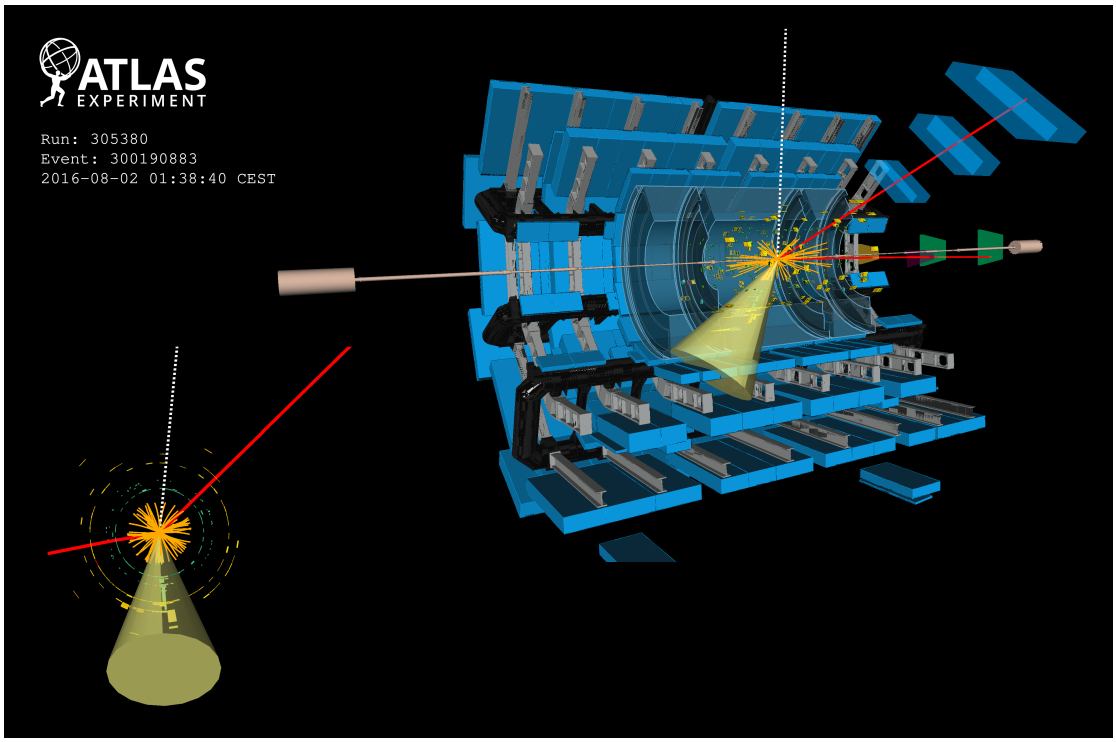


Figure 5.26: Visualisations of a data event recorded in 2016 that enters into SR-BDT$^{\mu\mu}_{5+10}$. The display on the right shows the events within the detector volume along with the LHC beam pipe. Energy deposits in the electromagnetic and hadronic calorimeters are displayed as green and yellow bars, respectively, while a subset of reconstructed inner-detector tracks with $p_{\mathrm{T}} > 1.5$ GeV are shown in orange. An identified jet with $p_{\mathrm{T}} = 297$ GeV is indicated by the yellow cone, and the two identified muons with $p_{\mathrm{T}} = 27.5$ GeV and 24.9 GeV are shown in red. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is indicated by the dashed white line and was measured to be 310 GeV. A BDT$_{5+10}$ score of 0.93 was assigned to the event. The display on the left shows the same event in the $x - y$ plane.

5.7.6 SR-BDT$^{ee}_{40+50}$ Investigations

The excess in SR-BDT$^{ee}_{40+50}$ is localised to the highest BDT scores, as seen in Figure 5.27. Figure 5.28 shows kinematic distributions in SR3-BDT$^{ee}_{40+50}$ only. These seven events were checked in similar detail to the events in the previous section and nothing concerning was spotted. Only limited conclusions can be drawn due to the low statistics in this region. The small excess of events at high $m_{\ell\ell}$ is the cause of the discrepancy, and is not seen in the corresponding muon SR.

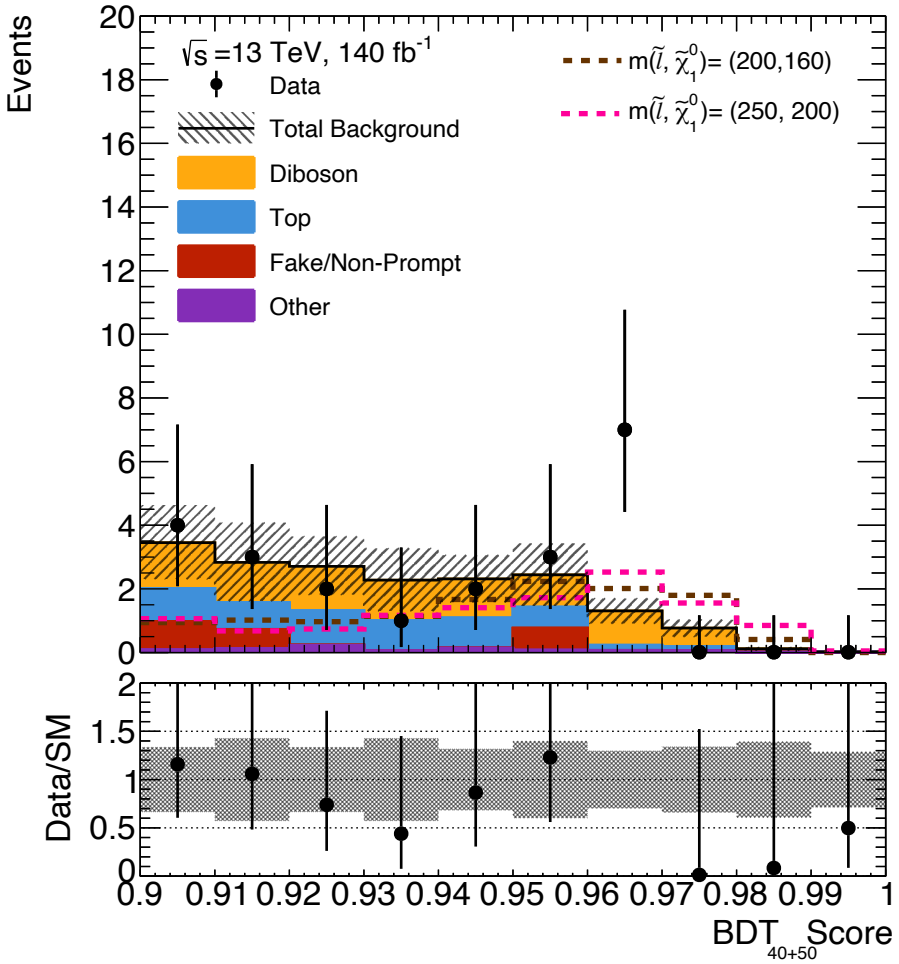


Figure 5.27: Data (dots) and post-fit SM background (histograms) distribution for the BDT score across all three regions of SR-BDT$^{ee}_{40+50}$ using the CR-only fit. Two example SUSY scenarios are overlaid for illustration that correspond to $m(\tilde{\ell}, \tilde{\chi}^0_1) = (200, 160)$ GeV and $(250, 200)$ GeV, assuming $m(\tilde{\ell}_{\mathrm{L}}) = m(\tilde{\ell}_{\mathrm{R}})$. The uncertainty bands include statistical and systematic uncertainties. The SRs are defined: SR1-BDT$^{ee}_{40+50}$ from 0.9 – 0.928, SR2-BDT$^{ee}_{40+50}$ from 0.928 – 0.956, and SR3-BDT$^{ee}_{40+50}$ from 0.956 – 1.0

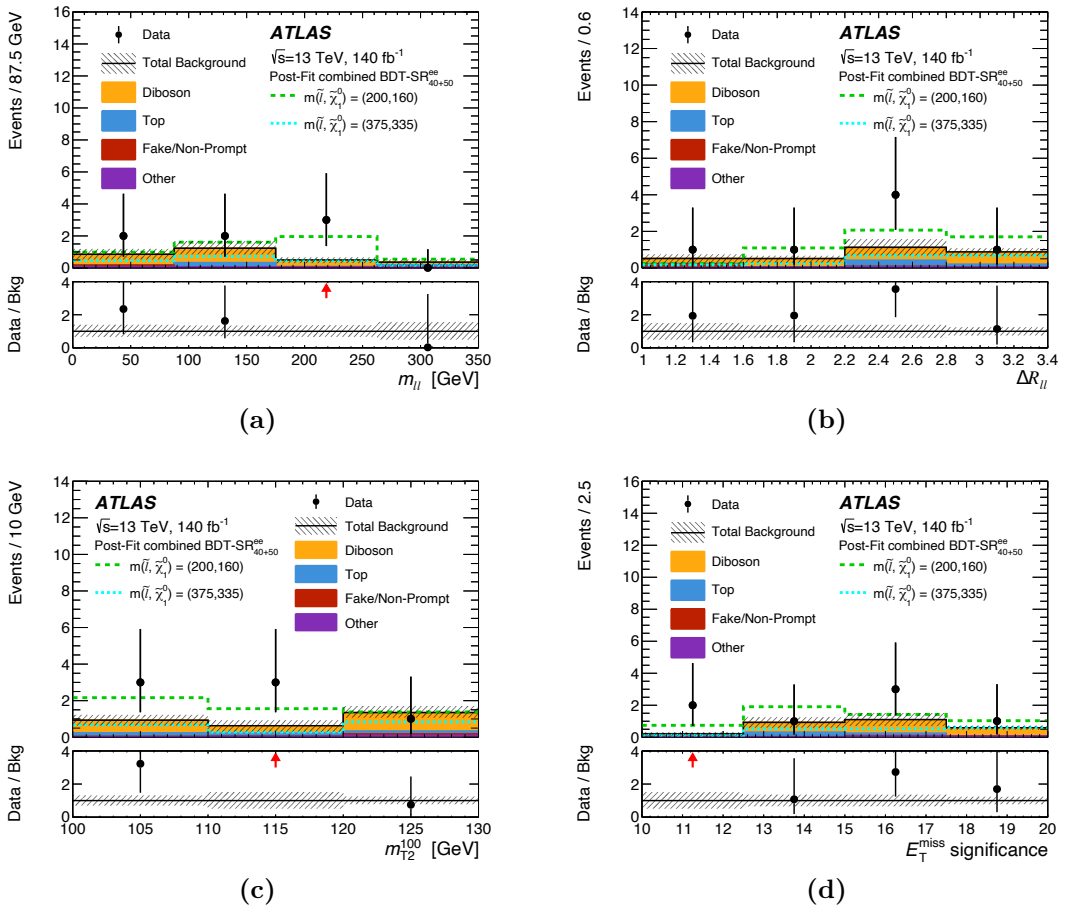


Figure 5.28: Data (dots) and post-fit SM background (histograms) distributions for (a) $m_{\ell\ell}$, (b) $\Delta R_{\ell\ell}$, (c) m_{T2}^{100} and (d) the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance in SR3-BDT$_{40+50}^{\mathrm{ee}}$ using the CR-only fit. Two example SUSY scenarios are overlaid for illustration that correspond to $m(\tilde{\ell}, \tilde{\chi}_1^0) = (200, 160)$ GeV and $(375, 335)$ GeV, assuming $m(\tilde{\ell}_{\mathrm{L}}) = m(\tilde{\ell}_{\mathrm{R}})$. The uncertainty bands include statistical and systematic uncertainties. The first and last bin includes the under- and overflow (if applicable), respectively. A red arrow in the lower panel indicates data points outside the vertical range shown.

5.8 Interpretations

This section details further interpretations of the results using the statistical fit methods described in Section 5.5. These focus on fits which utilise the SRs in a simultaneous fit with the CRs, to derive exclusion limits and discovery significances for signal models.

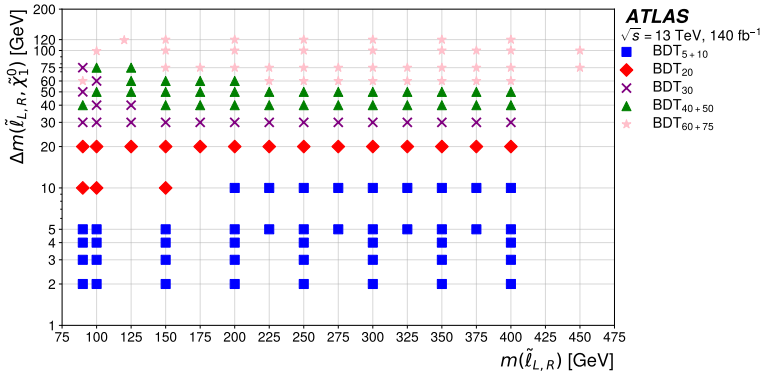
The observations in the SRs can be translated into constraints on the simplified models of slepton pair production. These interpretations are performed for the seven scenarios discussed in Section 4.3.3. The first scenario assumes, for the sake

of simplicity, the four-fold mass degeneracy $m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$ that also served as benchmark model for optimising the searches. The next two scenarios consider either selectron- or smuon-only production with $m(\tilde{e}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}})$ and $m(\tilde{\mu}) \equiv m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$, respectively. Finally, the remaining four scenarios consider either selectron- or smuon-only individual LH/RH production with $m(\tilde{e}) \equiv m(\tilde{e}_{\mathrm{L}})$, $m(\tilde{e}) \equiv m(\tilde{e}_{\mathrm{R}})$, $m(\tilde{\mu}) \equiv m(\tilde{\mu}_{\mathrm{L}})$ and $m(\tilde{\mu}) \equiv m(\tilde{\mu}_{\mathrm{R}})$, respectively. In the selectron-only scenario, for example, the smuons are assumed to be decoupled from the rest of the SUSY sparticle spectrum, and vice versa. Hence, only the ee ($\mu\mu$) SRs are considered when constraining the selectron-only (smuon-only) scenario.

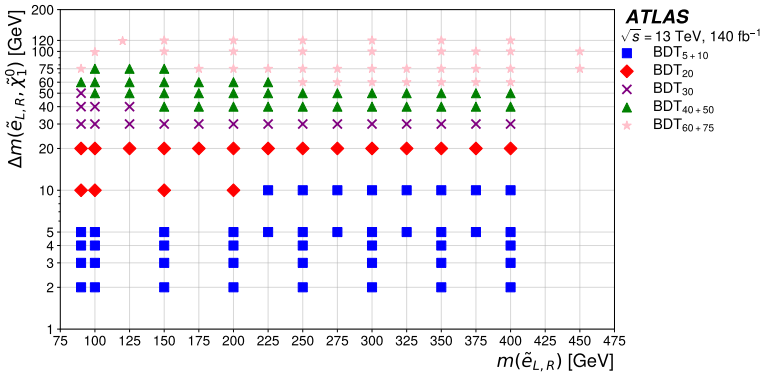
To determine the probability that the observed data is compatible with a particular signal model, the exclusion fit configuration is used in which the SRs are fitted simultaneously with the CRs. The CL_s prescription is then used to perform hypothesis tests and set exclusion limits at 95% CL. The exclusion limits are derived individually for each of the BDT SRs, as these are not statistically independent. To present the final limits, the limits are combined into one set of constraints using the following prescription. For each signal model, the BDT with the lowest expected CL_s value is used, and then this set of CL_s values is used to interpolate the final BDT exclusion contour. The BDT with the best expected CL_s for the four-fold degenerate and two-fold degenerate selectron, smuon only exclusion fits are shown in Figure 5.29. Individual LH/RH interpretations use the same BDT choice as the corresponding two-fold degenerate case. The numeric values for the expected and observed CL_s for each interpretation are presented in Appendix F.

5.8.1 Exclusion Limits

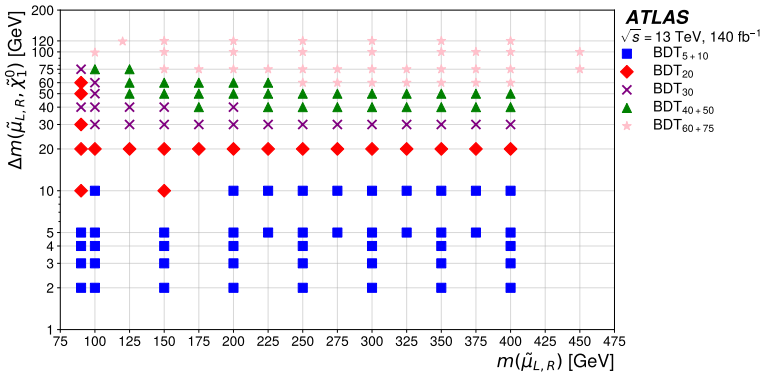
The slepton-mass exclusion contours are restricted to the range 2 GeV $\leq \Delta m(\tilde{\ell}, \tilde{\chi}_1^0) \leq$ 120 GeV due to the presence of strong pre-existing ATLAS constraints at larger [116, 117], and smaller [119] mass splittings. Scenarios with $m(\tilde{\ell}) < 95$ GeV are not considered, as the LEP experiments generally sets strong exclusion for these small slepton masses [115, 159]. The region excluded by the different interpretations is that bounded by the two exclusion contours, where two contours exist.



(a)



(b)



(c)

Figure 5.29: Overview of which of the five BDT-based selections provides the best expected CL_s value across the $m(\tilde{\ell}) - \Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ parameter plane for each signal benchmark model for (a) the $m(\tilde{\ell}) \equiv m(\tilde{e}_\mathrm{L}) = m(\tilde{e}_\mathrm{R}) = m(\tilde{\mu}_\mathrm{L}) = m(\tilde{\mu}_\mathrm{R})$ interpretation, (b) the $m(\tilde{e}) \equiv m(\tilde{e}_\mathrm{L}) = m(\tilde{e}_\mathrm{R})$ interpretation and (c) the $m(\tilde{\mu}) \equiv m(\tilde{\mu}_\mathrm{L}) = m(\tilde{\mu}_\mathrm{R})$ interpretation.

Considering first the signal models in which a four-fold slepton mass degeneracy is assumed, Figure 5.30 shows the observed and expected BDT-based exclusion and cut-and-count limits, as a comparison between the two methods. Further details on the cut-and-count method are given in Ref. [1].

Both the cut-and-count and the BDT-based exclusion reaches are reduced at around $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = 40$ GeV. This originates from the excesses observed in different lepton SRs, a 1.1 σ excess in the muon channel for the cut-and-count approach, and a 2 σ excess in the electron channel for the BDT approach. The increased sensitivity of the BDTs over the cut-and-count-based approach originates from training multiple BDTs to optimise individually for a certain $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$, collectively providing high sensitivity across the full $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$ corridor. [6] In addition, the BDTs are able to exploit complex correlations between input variables.

The results can also be interpreted within models where $m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}})$ and $m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$, while the mass degeneracy between the slepton flavours is broken. Specifically, the $\tilde{e}_{\mathrm{L,R}}$ ($\tilde{\mu}_{\mathrm{L,R}}$) particles are assumed to be within the reach of the search, while the $\tilde{\mu}_{\mathrm{L,R}}$ ($\tilde{e}_{\mathrm{L,R}}$) particles are decoupled at inaccessibly large masses. The expected and observed exclusions for the BDT-based analysis within these models are shown in Figure 5.31. It can be observed that for both the selectron and smuon cases, the expected exclusion for the BDTs collectively spans the full range of $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$, across the challenging "corridor" region in which no previous experiment had sensitivity. Again, the observed exclusions are weaker compared with the predictions of the background-only model, where excesses were observed in the data. In these interpretations, the di-electron excesses affect only the observed selectron exclusion, and analogous for the di-muon excesses and smuon exclusion.

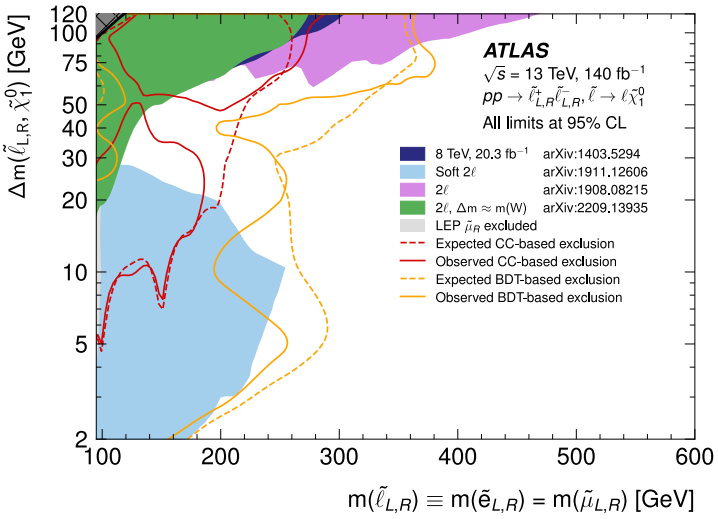
The data excess in SR3-BDT$^{\mathrm{ee}}_{40+50}$ results in a reduction in the observed exclusion limit for selectrons with $\Delta m(\tilde{e}, \tilde{\chi}_1^0) = 30$–$50$ GeV, preventing any exclusion of signal models within this region. This results in a split of the exclusion contours into two "islands" of sensitivity for mass splittings below and above this region. Stringent constraints are placed on selectrons in scenarios with $\Delta m(\tilde{e}, \tilde{\chi}_1^0)$ between 2 GeV

[6] Training a BDT over the complete spectrum of mass splittings was found to yield sensitivity much closer to the cut-and-count-based analysis strategy.

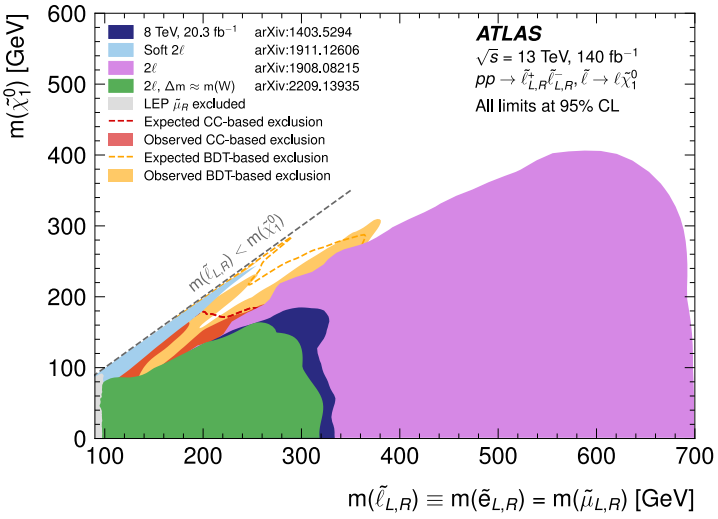
and 20 GeV, representing significant improvements over previous ATLAS results. Notably, sensitivity is achieved at $\Delta m(\tilde{e}, \tilde{\chi}_1^0) = 20$ GeV, extending the exclusion reach up to $m(\tilde{e}) = 200$ GeV, where the strongest previous constraints were still from the LEP experiments [115]. Enhanced exclusion limits are observed for mass splittings between 5 GeV and 10 GeV, up to a similar selectron mass. Larger mass splittings with $\Delta m(\tilde{e}, \tilde{\chi}_1^0)$ between 60 GeV and 75 GeV are constrained up to $m(\tilde{e}) = 300$ GeV, again improving substantially over previous results.

For the smuon exclusion limits, a sizeable reduction in the observed excluded region compared with the expected one is seen for $\Delta m(\tilde{\mu}, \tilde{\chi}_1^0) = 3$–$10$ GeV, driven by excesses in SR1-BDT$^{\mu\mu}_{5+10}$ and SR2-BDT$^{\mu\mu}_{5+10}$. Except where excesses in the numbers of observed events reduce the ability to exclude models, significant improvements in the exclusion reach are observed for $\Delta m(\tilde{\mu}, \tilde{\chi}_1^0)$ between 20 GeV and 100 GeV, extending from $\tilde{\mu}_{\mathrm{L,R}} = 250$ GeV up to $\tilde{\mu}_{\mathrm{L,R}} = 350$ GeV.

Finally, the individual LH and RH selectron and smuon exclusion limits are shown in Figures 5.32 and 5.33. These are obtained by running the same fit configuration as in the two-fold cases, but with the signal cross-sections rescaled to reflect the individual LH, RH selectron/smuon pair production cross-sections. For LH selectrons, the sensitivity is limited to the highest and lowest $\Delta m(\tilde{e}, \tilde{\chi}_1^0)$ values, with modest improvements over previous results. For LH smuons, the observed sensitivity is greater, with sensitivity for $\Delta m(\tilde{\mu}, \tilde{\chi}_1^0)$, reaching down to 40 GeV, but with still some more work to be carried out to gain sensitivity across the full corridor region. The exclusion limits for RH selectrons and smuons are significantly weaker than those for their LH counterparts, primarily due to the reduced production cross-section arising from the chiral structure of the weak interaction. For RH selectrons, a small improvement over existing limits is seen at low $\Delta m(\tilde{e}, \tilde{\chi}_1^0)$, presenting new sensitivity for the first time since LEP. For RH smuons, there are two islands of observed sensitivity at high and low $\Delta m(\tilde{\mu}, \tilde{\chi}_1^0)$, slightly improving on existing limits. Of particular interest is the large increase in expected significance for $\Delta m(\tilde{\mu}, \tilde{\chi}_1^0)$ between 2 GeV and 10 GeV, highlighting the potential for future analyses to set limits, or make discoveries, in these interpretations.

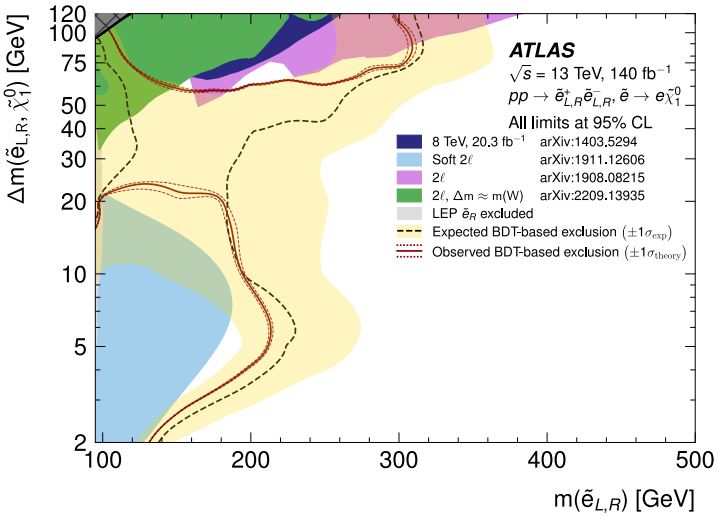


(a)

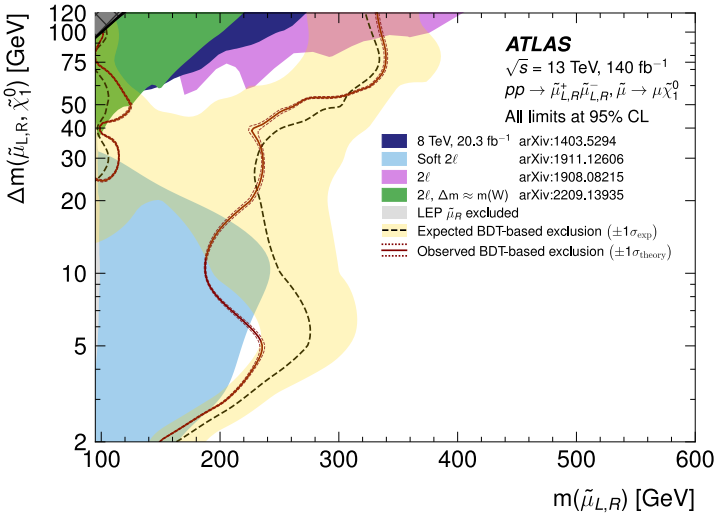


(b)

Figure 5.30: Exclusion limits for simplified SUSY models of direct slepton production assuming four-fold mass degeneracy: $m(\tilde{\ell}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$. The limits are shown in (a) the $m(\tilde{\ell})$ – $\Delta m(\tilde{\ell}, \tilde{\chi}^{0}_{1})$ plane and (b) the $m(\tilde{\ell})$ – $m(\tilde{\chi}^{0}_{1})$ plane. In (a), a comparison is shown between the observed (orange solid lines) and expected (orange dashed lines) exclusion limits from the BDT approach, and the observed (solid red) and expected (dashed red) exclusion limits obtained from the cut-and-count-based approach. In (b), a comparison is shown between the observed (orange fill) and expected (orange dashed lines) exclusion limits from the BDT approach, and the observed (red fill) and expected (dashed red) exclusion limits obtained from the cut-and-count-based approach. All limits are computed at 95% CL, and the black delineated grey region indicates the kinematically forbidden region where $m(\tilde{\ell}_{L,R}) < m(\tilde{\chi}^{0}_{1})$. The observed limits obtained at LEP [115], along with previous searches at the ATLAS experiment [116–119] are shown.

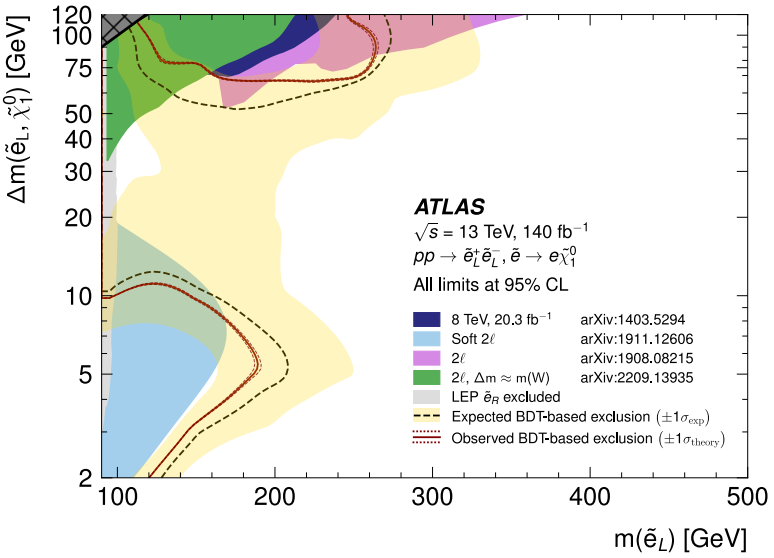


(a)

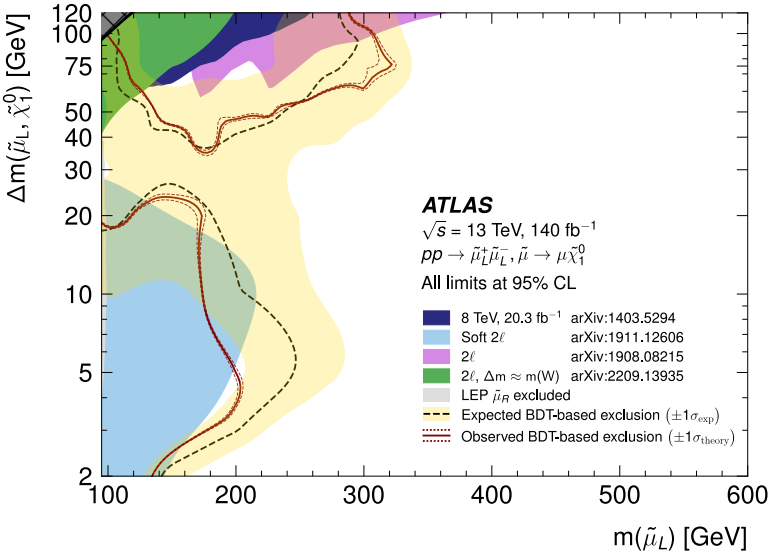


(b)

Figure 5.31: Exclusion limits for simplified SUSY models assuming two-fold mass degeneracy. Separate limits are shown for (a) direct selectron production, where $m(\tilde{e}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}})$, and (b) direct smuon production, where $m(\tilde{\mu}) \equiv m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$. All limits are computed at 95% CL and the observed (red solid lines) and expected (black dashed lines) exclusion limits from the BDT regions are shown in the $m(\tilde{e})$ – $\Delta m(\tilde{e}, \tilde{\chi}^{0}_{1})$ and $m(\tilde{\mu})$ – $\Delta m(\tilde{\mu}, \tilde{\chi}^{0}_{1})$ planes, respectively. In (a), the observed exclusion should be interpreted as the region enclosed by each separate observed curve and the axis. For the expected exclusion in (a) and the expected and observed exclusion in (b), the excluded region should be interpreted as the region enclosed by the two observed curves. The black delineated grey region indicates the kinematically forbidden region where $m(\tilde{e}_{\mathrm{L,R}}/\tilde{\mu}_{\mathrm{L,R}}) < m(\tilde{\chi}^{0}_{1})$. The yellow shaded band around the expected limits corresponds to the $\pm 1\,\sigma$ variations of the expected limit, accounting for all uncertainties aside from the theoretical uncertainties in the signal cross-section. The dotted lines around the observed exclusion limits indicate the variation as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. The observed limits obtained at LEP [115, 159], along with previous searches at the ATLAS experiment [116–119] are shown.

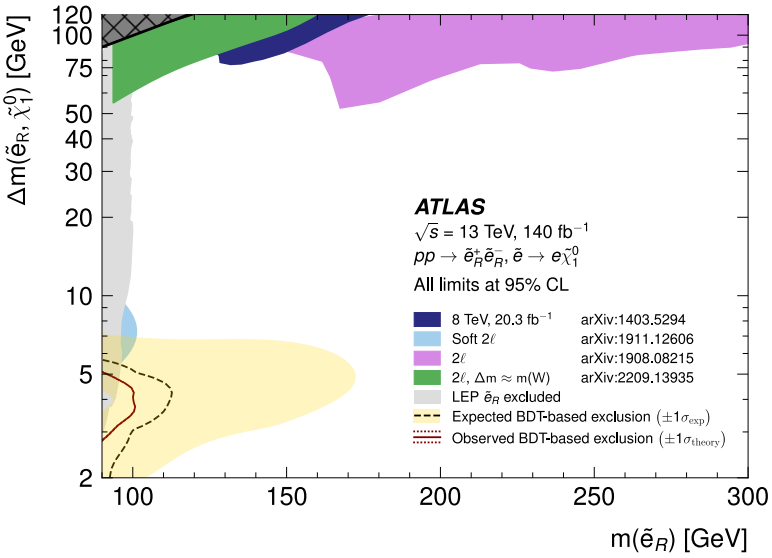


(a)

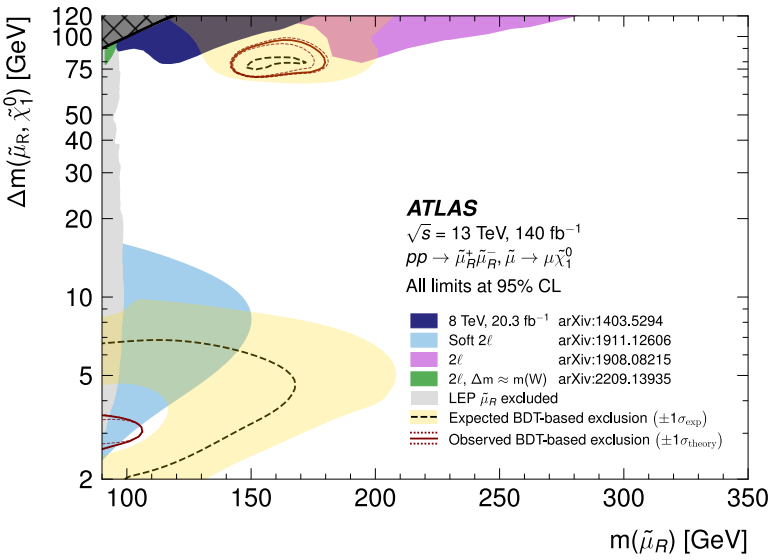


(b)

Figure 5.32: Exclusion limits for simplified SUSY models assuming no mass degeneracy within the sleptons. Separate limits are shown for (a) direct selectron production, where $m(\tilde{e}) \equiv m(\tilde{e}_{\mathrm{L}})$, and (b) direct smuon production, where $m(\tilde{\mu}) \equiv m(\tilde{\mu}_{\mathrm{L}})$. All limits are computed at 95% CL and the observed (red solid lines) and expected (black dashed lines) exclusion limits from the BDT regions are shown in the $m(\tilde{e}_{\mathrm{L}})$ – $\Delta m(\tilde{e}_{\mathrm{L}}, \tilde{\chi}_1^0)$ and $m(\tilde{\mu}_{\mathrm{L}})$ – $\Delta m(\tilde{\mu}_{\mathrm{L}}, \tilde{\chi}_1^0)$ planes, respectively. The black delineated grey region indicates the kinematically forbidden region where $m(\tilde{e}_{\mathrm{L}}/\tilde{\mu}_{\mathrm{L}}) < m(\tilde{\chi}_1^0)$. The yellow shaded band around the expected limits corresponds to the $\pm1\,\sigma$ variations of the expected limit, accounting for all uncertainties aside from the theoretical uncertainties in the signal cross-section. The red dotted lines around the observed exclusion limits indicate the variation as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. The observed limits obtained at LEP [115, 159], along with previous searches at the ATLAS experiment [116–119] are shown.



(a)



(b)

Figure 5.33: Exclusion limits for simplified SUSY models assuming no mass degeneracy within the sleptons. Separate limits are shown for (a) direct selectron production, where $m(\tilde{e}) \equiv m(\tilde{e}_{\mathrm{R}})$, and (b) direct smuon production, where $m(\tilde{\mu}) \equiv m(\tilde{\mu}_{\mathrm{R}})$. All limits are computed at 95% CL and the observed (red solid lines) and expected (black dashed lines) exclusion limits from the BDT regions are shown in the $m(\tilde{e}_{\mathrm{R}})$ – $\Delta m(\tilde{e}_{\mathrm{R}}, \tilde{\chi}_1^0)$ and $m(\tilde{\mu}_{\mathrm{R}})$ – $\Delta m(\tilde{\mu}_{\mathrm{R}}, \tilde{\chi}_1^0)$ planes, respectively. The black line below the dark grey box indicates the region where $m(\tilde{e}_{\mathrm{R}}/\tilde{\mu}_{\mathrm{R}}) < m(\tilde{\chi}_1^0)$. The yellow shaded band around the expected limits corresponds to the $\pm1\,\sigma$ variations of the expected limit, accounting for all uncertainties aside from the theoretical uncertainties in the signal cross-section. The red dotted lines around the observed exclusion limits indicate the variation as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. The observed limits obtained at LEP [115, 159], along with previous searches at the ATLAS experiment [116–119] are shown.

5.8.2 Signal Model Discovery Significances

To identify which signal model fits the observed excesses best, the p_0-value of the background-only hypothesis is evaluated for each signal model and translated into a 'discovery' significance. Similar to the exclusion fits, the BDT with the best expected discovery significance for each signal model is selected, and the observed significance from this set of SRs is used. The BDT which provides the best expected significance for each signal point in the selectron and smuon interpretations is shown in Figure 5.34. The discovery significances are evaluated using 15,000 toy pseudo-experiments. Figure 5.35, shows the discovery significances in the two-fold degenerate interpretations. No significances are reported for the four-fold and individual L/R selectron/smuon interpretations.

For the two-fold degenerate selectron signal models, the highest discovery significance is found to be 1.9σ and corresponds to the scenario with $\Delta m(\tilde{e}, \tilde{\chi}_1^0) = (375, 335)$ GeV evaluated with SR-BDT$^{ee}_{40+50}$. The smuon model with the same masses, $\Delta m(\tilde{\mu}, \tilde{\chi}_1^0) = (375, 335)$ GeV, has a discovery significance of 0.74σ. These discovery significances, evaluated relative to a particular signal model, are lower than the single-bin significances reported above due to differences between the shapes of the signal predictions and the excesses observed in the SRs.

For the two-fold degenerate smuon signal models, the highest discovery significance is provided by the $\Delta m(\tilde{\mu}, \tilde{\chi}_1^0) = (150, 140)$ GeV scenario [7] using SR-BDT$^{\mu\mu}_{5+10}$ and amounts to 2.1σ. The selectron model with the same masses, $\Delta m(\tilde{e}, \tilde{\chi}_1^0) = (150, 140)$ GeV, yields a discovery significance of 0.18σ.

[7] When discounting signal models excluded by the LEP experiment.

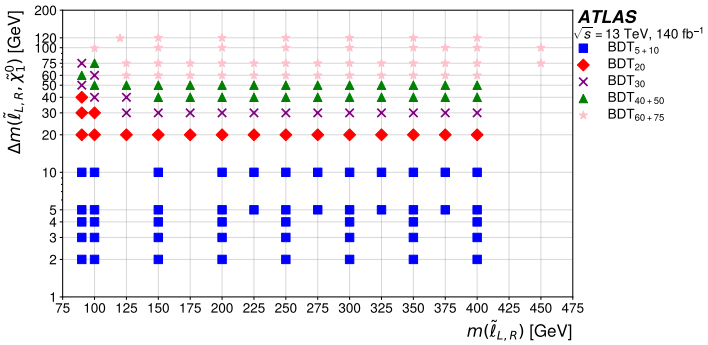
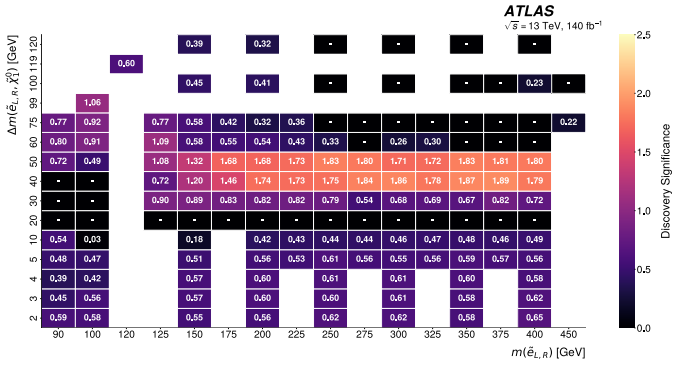
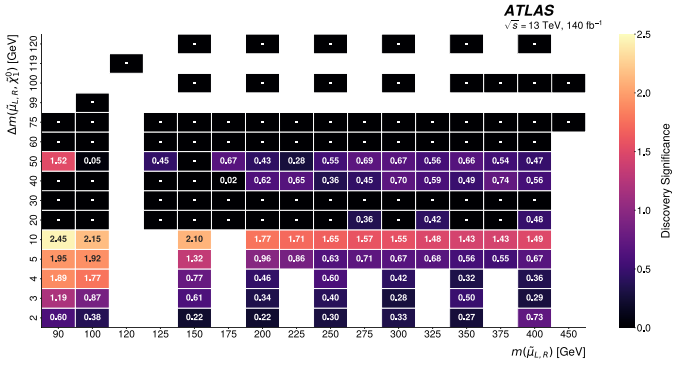


Figure 5.34: Overview of the BDT which provides the best expected discovery significance across the $m(\tilde{\ell}_{\mathrm{L,R}})$ – $\Delta m(\tilde{\ell}_{\mathrm{L,R}}, \tilde{\chi}_1^0)$ plane assuming the four-fold mass-degenerate sleptons scenarios $m(\tilde{\ell}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$. This differs from the BDTs shown in Figure 5.29, as the test statistic now evaluates the statistical significance of the deviation for a signal with the background only model, not exclusion via CL_s.



(a)



(b)

Figure 5.35: Observed discovery significance in the $m(\tilde{\ell})$– $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ plane for each benchmark signal model assuming (a) $m(\tilde{\ell}) \equiv m(\tilde{e}_{\mathrm{L,R}})$ and (b) $m(\tilde{\ell}) \equiv m(\tilde{\mu}_{\mathrm{L,R}})$. Each significance is evaluated using 15,000 pseudo-experiments and the relevant set of BDT SRs; for the selectron-only scenario only di-electron SRs, for the smuon-only scenario only di-muon SRs. The BDT SRs chosen for each signal correspond to the set of SRs that give the highest expected significance, from Figure 5.34. Dashes (-) indicate where the discovery significance is negative, and not reported.

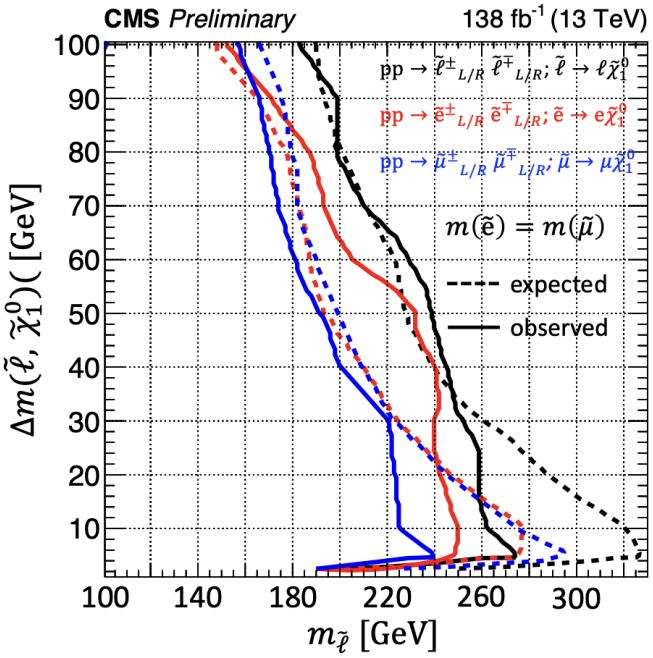
5.9 The Wider Context

Supersymmetry has been a leading theory to search for BSM physics in collider physics for many years, with no breakthrough as of yet. The current picture is an exciting one, with several results in compressed signatures having observed excesses in both ATLAS and CMS. The work described in this chapter is the largest and perhaps the most intriguing of these excesses, due to the specific kinematic features of the observed data. In this section the analysis described earlier in this chapter will be referred to as the 'slepton corridor' analysis.

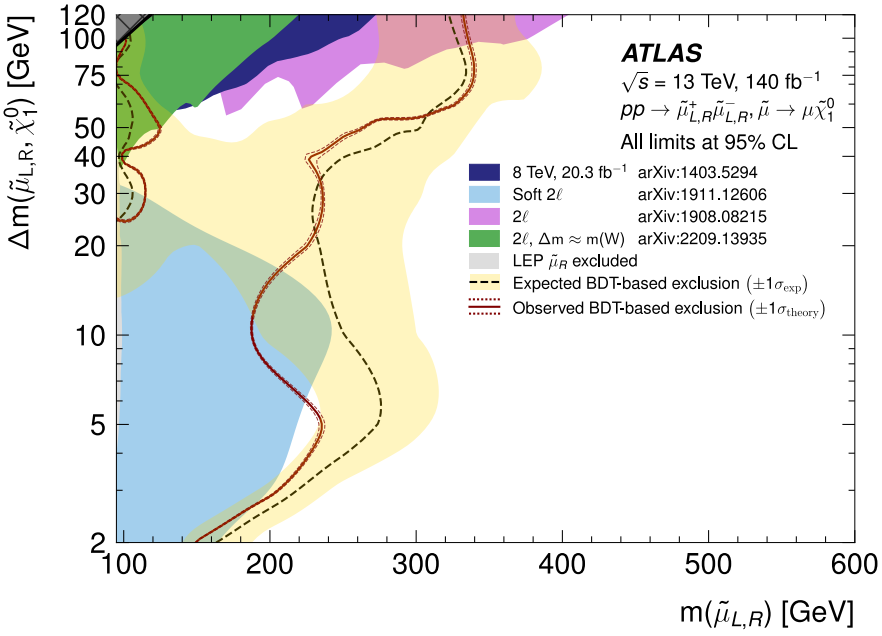
To add further interest to the result shown in this chapter, the CMS collaboration released a preliminary result targeting a general search for supersymmetric particles in compressed mass spectrum scenarios, where limits are set on sleptons, stops and higgsinos [160]. The analysis utilises Recursive Jigsaw Reconstruction to discriminate between signal and background events [161, 162], similar to the ATLAS result in Ref. [117] The results are obtained using a maximum likelihood fit over 2443 bins in 392 categories, with the full Run 2 data set. Figure 5.36, shows the main slepton exclusion contour from the result, where a drop in observed exclusion is seen at $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \approx 10$ GeV for smuons. The results of the ATLAS analysis are also shown in the figure for comparison. The preliminary result does not provide any further kinematic information, and gives no statistical significances so it is hard to draw any conclusions about compatiblity with the ATLAS result. As the result is not yet published, hopefully the final paper will include more information to indicate whether the CMS excess follows the kinematics of the ATLAS analysis.

Sleptons are not the only electroweak SUSY signatures with compressed mass spectra that show intriguing excesses in both ATLAS and CMS. Both ATLAS and CMS have compatible $\approx 2\ \sigma$ excesses observed in Run 2 data for analyses targeting compressed mass spectrum higgsinos, [119], [121]. Figure 5.37 shows the targeted processes in these analyses [8]. Figure 5.38, shows both of the ATLAS 'soft $2l$' higgsino result and the corresponding CMS result. There is an excess in the

[8] in addition to the same ISR slepton pair-production targeted by the slepton corridor analysis.



(a)



(b)

Figure 5.36: Exclusion limits for simplified SUSY models. All limits are computed at 95% CL. The CMS result (a) shows, where the four-fold (solid lines) observed and (dotted lines) expected exclusion limits are shown alongside the two-fold limits from Ref. [160]. In (b), the smuon limits from Figure 5.31 are repeated for comparison.

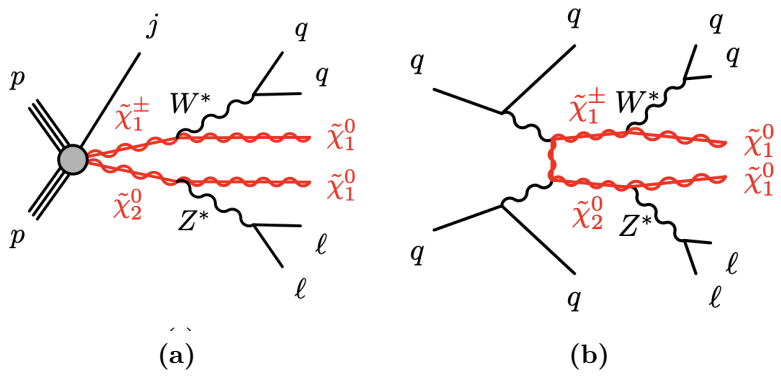


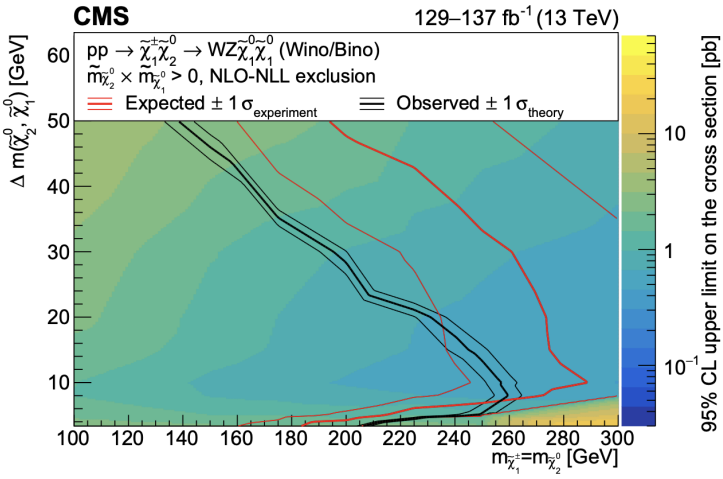
Figure 5.37: Diagrams targeted by the 'soft $2l$' higgsino analysis. The two-lepton final state of (a) production of electroweakinos $\tilde{\chi}_2^0\tilde{\chi}_1^{\pm}$ with initial-state radiation (j), (b) vector boson fusion (VBF) production of electroweakinos $\tilde{\chi}_2^0\tilde{\chi}_1^{\pm}$.

Higgsino search for both results for $\Delta m(\tilde{\chi}_2^0, \tilde{\chi}_1^{\pm}) = 20$ GeV. The $\tilde{\chi}_1^0$ favoured by this excess is also seen to be roughly compatible with the dark matter relic density [163].

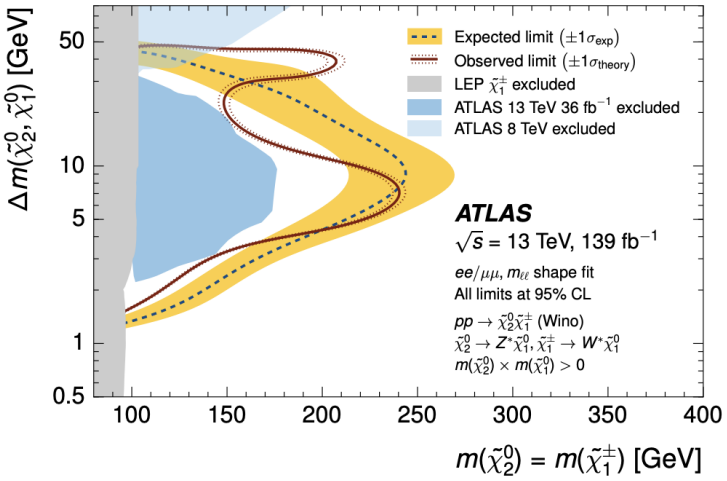
The soft $2l$ analysis also has a slepton search, targeting the same process as our slepton corridor analyis. Investigation was carried out into the direct event overlap between our slepton corridor analysis, and both the slepton and higgsino SRs from soft $2l$. The excess in the soft $2l$ higgsino search occurs in the electron channel, for $5 < m_{\ell\ell} < 10$, and so is not kinematically similar to our slepton corridor electron excess. The compatiblity between the slepton corridor excess and this higgsino excess is not trivial to assess as the higgsino/slepton are kinematically decoupled in the respective benchmark simplifed models. However, broadly the $\tilde{\chi}_1^0$ values that may account for the excesses, are both compatible with values of $m_{\tilde{\chi}_1^0} \approx 150$–$200$ GeV.

Table 5.13 shows the overlap in data events between SR-BDT$^{\mu\mu}_{5+10}$ and the di-muon SRs of the soft $2l$ analysis – without and with a $\Delta R_{\ell\ell}$ requirement to select for the excess region. There is generally a small overlap between the SRs of the two analyses, with the largest overlap occuring with the slepton SRs of soft $2l$. Figure 5.39 shows the m_{T2} distributions for the slepton corridor analysis, and the soft $2l$ di-muon slepton SR. The 101–102 m_{T2}^{100} bin has a Data/MC ratio $\approx$ 1.5, and is the largest discrepancy in the muon channel for the soft $2l$ analysis.

[9] One of these events also appears in the electroweakino SRs.

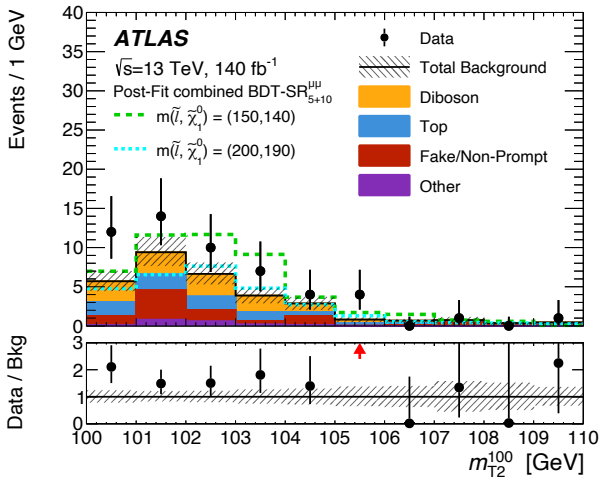


(a)

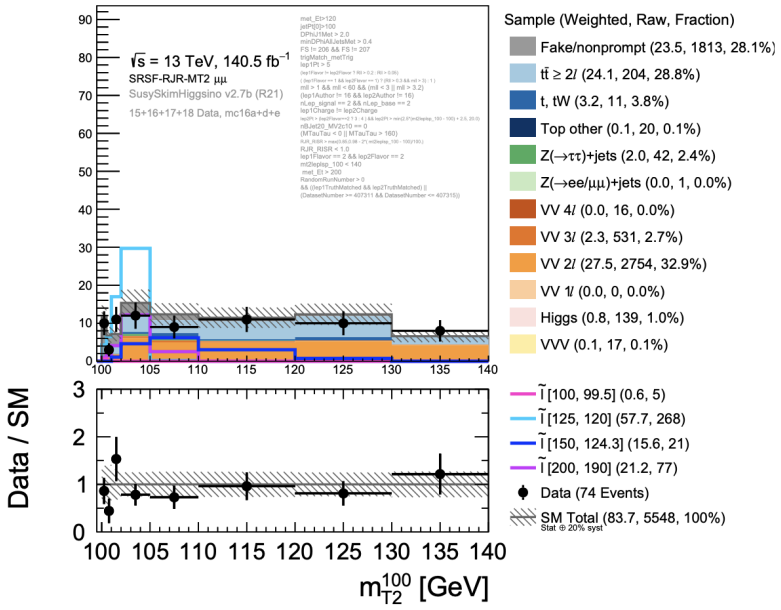


(b)

Figure 5.38: Exclusion limits are set for simplified SUSY models, with all limits are computed at 95% CL. The results assume: $m_{\tilde{\chi}_2^0} = m_{\tilde{\chi}_1^\pm}$ and $m_{\tilde{\chi}_2^0} \times m_{\tilde{\chi}_1^\pm} > 0$, and are Wino/Bino like. The limits are shown in the $m(\tilde{\chi}_2^0)$ – $\Delta m(\tilde{\chi}_2^0, \tilde{\chi}_1^0)$ plane. The CMS result is shown in (a). The red solid lines are the expected limits, and the slightly thinner red solid lines corresponds to the $\pm 1\,\sigma$ variations of the expected limit, accounting for all uncertainties aside from the theoretical uncertainties in the signal cross-section. The thick black solid lines are the observed exclusion limits, and the thinner black lines indicate the variation as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. The colour scale indicates the upper limit on the slepton cross section across the range signal space. In (b), the ATLAS limits are shown. The yellow shaded band around the black dashed expected limits corresponds to the $\pm 1\,\sigma$ variations of the expected limit, accounting for all uncertainties aside from the theoretical uncertainties in the signal cross-section. The dotted lines around the red solid observed exclusion limits indicate the variation as the nominal signal cross-section is scaled up and down by the theoretical uncertainty.



(a)



(b)

Figure 5.39: The post-fit m_{T2}^{100} distributions with background yields from the CR-Only fit. In (a) the distribution for SR-BDT$_{5+10}^{\mu\mu}$ in the slepton corridor analysis is shown. The uncertainties in (a) cover all experimental and theoretical uncertainties. In (b) the distribution for SR$\mu\mu - m_{\mathrm{T2}}^{100}$ in the soft-2l analysis is shown. The uncertainties in (b) include statistical uncertainty and a flat 20% uncertainty on all backgrounds.

Table 5.13: The overlap in the number of observed data events from all three SR-BDT$^{\mu\mu}_{5+10}$ SRs with the di-muon soft 2ℓ SRs. Left shows the overlap in data events from the SR-BDT$^{\mu\mu}_{5+10}$ SRs with both the slepton and electroweakino SRs, and individually. Right shows the overlap when applying a further selection, $2.8 < \Delta R_{\ell\ell} < 3.1$, to select for events in the slepton corridor excess. For example, the entry 5/24 indicates that of the 24 events in the SR-BDT$^{\mu\mu}_{5+10}$ – when requiring $2.8 < \Delta R_{\ell\ell} < 3.1$, 5 of these events overlap with the soft $2l$ slepton and electroweakino SRs.

Considered SRs	**Overlap**	**Details**
All SR-BDT$^{\mu\mu}_{5+10}$ vs. all soft 2ℓ SRs	13/53	Includes both electroweakino and slepton SRs
All SR-BDT$^{\mu\mu}_{5+10}$ vs. Electroweakino SRs	6/53	Binned in $m_{\ell\ell}$
All SR-BDT$^{\mu\mu}_{5+10}$ vs. Slepton SRs	7/53	Binned in m^{100}_{T2}

Considered SRs with $2.8 < \Delta R_{\ell\ell} < 3.1$	**Overlap**	**Details**
All SR-BDT$^{\mu\mu}_{5+10}$ vs. all soft 2ℓ SRs	5/24	Indicates substantial orthogonality with the excess
All SR-BDT$^{\mu\mu}_{5+10}$ vs. Electroweakino SRs	2/24	One each in $m_{\ell\ell} = 30$–40 and 40–60 bins
All SR-BDT$^{\mu\mu}_{5+10}$ vs. Slepton SRs	4/24	One event in 100.5–101 m^{100}_{T2} bin, rest in the 101–102 bin[9]

There is another related excess seen in both ATLAS and CMS, targeting three-lepton final states, similar to Figure 5.37 but requiring a leptonic W decay [160, 164, 165] These processes are shown in Figure 5.40. Figure 5.41 shows an interpretation of the results from the analyses. The ATLAS analysis targeting these final states has $\approx 2\sigma$ excesses in both the WZ and the Wh channels. For CMS, no significance is provided, but there is a sizable excess in two bins of the WZ SRs from Ref. [160] and in the Wh channel from Ref. [165]. The WZ results follow the same broad features as the other compressed spectra excesses discussed in this section. The CMS result is from the same preliminary CMS analysis that has the smuon excess, but different SR(s). These excesses are explicitly orthogonal to the other results discussed in this section due to the lepton multiplicity requirement, but could still originate from decays of similar SUSY particles.

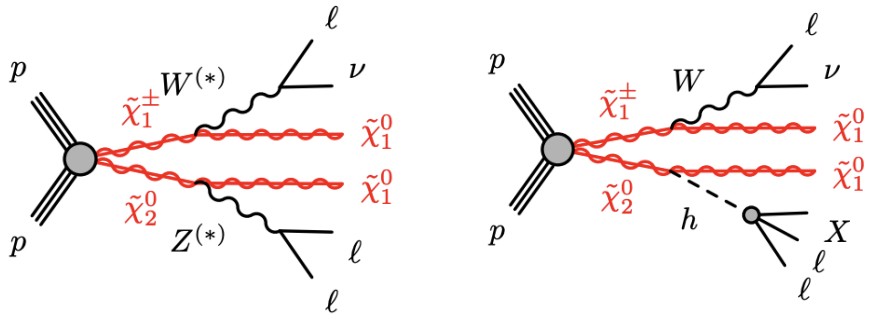
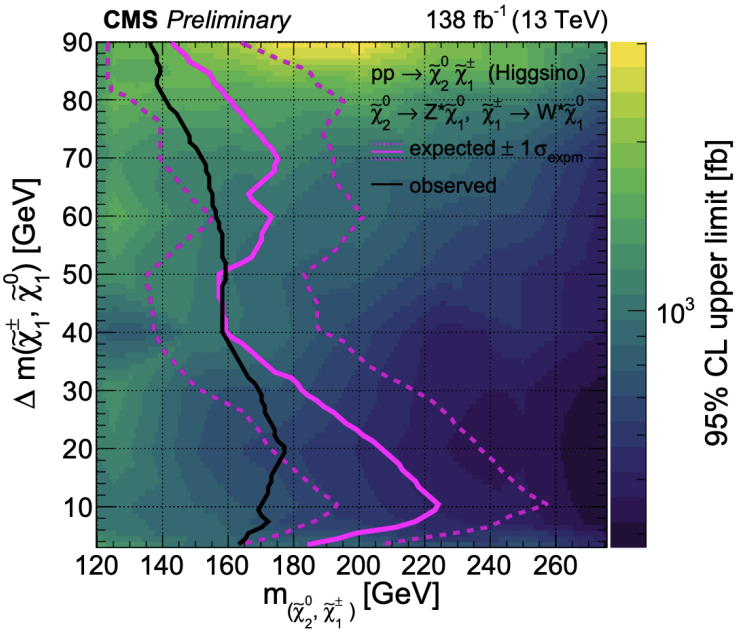
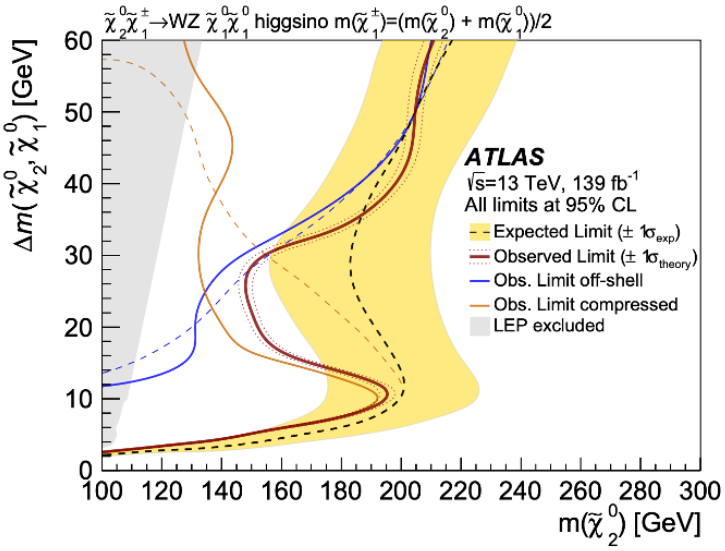


Figure 5.40: Diagrams of simplified models targeted by the 3L analyses for $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0$ pair production. In (a) intermediate WZ (W^*Z^*) production is shown, and in (b) intermediate Wh, where X indicates additional decay products from WW, ZZ, or $\tau\tau$.



(a)



(b)

Figure 5.41: Exclusion limits for simplified SUSY models, with all limits are computed at 95% CL. In both cases the limits are shown in the $m(\tilde{\chi}_2^0)$ – $\Delta m(\tilde{\chi}_2^0, \tilde{\chi}_1^0)$ plane assuming $m_{\tilde{\chi}_2^0} = m_{\tilde{\chi}_1^\pm}$. For both results, the $\tilde{\chi}_2^0, \tilde{\chi}_1^\pm$ are Higgsino-like. In the CMS result (a), the pink solid lines are the expected limits, and the dotted pink lines correspond to the $\pm 1\,\sigma$ variations of the expected limit, accounting for all uncertainties aside from the theoretical uncertainties in the signal cross-section. The thick black solid lines are the observed exclusion limits, no uncertainty is provided on these. The colour scale indicates the upper limit on the cross section across the range signal space. In (b), for the ATLAS result, the $\tilde{\chi}_2^0, \tilde{\chi}_1^\pm$ are Wino/Bino-like and the elements should be interpreted as in Figure 5.38.

Given the multiple electroweak SUSY excesses observed in these searches targeting simplified models, efforts are underway to reinterpret these results both within the pMSSM framework and using another approach inspired by Ref. [163]. These studies aim to assess the consistency of the observed excesses by employing signal models that have the full SUSY particle spectrum. This includes examining the compatibility of ATLAS and CMS results in overlapping final states, as well as evaluating how well the excesses align across the different final states described in this section. Furthermore, these interpretations allow the compatibility of the excesses with dark matter relic density constraints to be evaluated. In particular, these results are compatible with potential coannihilation scenario, where a Bino-like $\widetilde{\chi}_1^0$ has a small mass difference with the $\widetilde{\chi}_1^\pm$ and a larger, but still small, mass difference with the $\widetilde{\chi}_2^0$. The impacts of this on the dark matter relic density can only be evaluated using more comprehensive SUSY models with the full spectrum.

5.10 Summary

This chapter detailed a search for selectrons and smuons, targeting simplified models with a small mass splitting between slepton and neutralino – similar or less than the mass of the W boson, $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) \lesssim m(W)$ – where no previous searches have had sufficient sensitivity. The searches use the complete LHC Run 2 data sample from the ATLAS experiment, equivalent to 140 fb^{-1} of integrated luminosity. The BDT based approach was detailed, where a set of five BDTs, each trained for a different region of $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$ in the range of 5 GeV to 75 GeV, are utilised. Expected sensitivity to $\tilde{\mu}$ or $\tilde{e}$ masses ranging from 200 GeV to 350 GeV for the simplified SUSY models with $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$ larger than 2 GeV, providing expected sensitivity across the full range of targeted $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$, even for scenarios in which four-fold mass degeneracy of the sleptons is not assumed. Limits are set at the 95% CL on $\tilde{\mu}$ masses up to 350 GeV for particular values of $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$, assuming $m(\tilde{\mu}_\text{L}) = m(\tilde{\mu}_\text{R})$ and similarly on $\tilde{e}$ masses up to 300 GeV assuming $m(\tilde{e}_\text{L}) = m(\tilde{e}_\text{R})$. This analysis significantly improves upon the constraints compared to previous searches and largely addresses a long-standing sensitivity gap in the $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0)$

corridor at moderately compressed mass splittings. Some regions of slepton and neutralino mass that are excluded are smaller than expected due to the observed excesses. The most significant excesses are observed with a local significance of $2.0\,\sigma$ in the search for selectrons and $2.4\,\sigma$ in the search for smuons. The largest effect of the excesses are on smuons models with $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0) \approx 10\,\text{GeV}$ and selectrons models with $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0) \approx 40\,\text{GeV}$. The maximum discovery significance found is $1.9\,\sigma$ for the selectron signal model with $m(\tilde{e}, \tilde{\chi}_1^0) = (375, 335)$ GeV, and $2.1\,\sigma$ for the smuon signal model with $m(\tilde{\mu}, \tilde{\chi}_1^0) = (150, 140)$ GeV.

We are in a exciting position: the Run 2 analysis is statistically limited, yet the Run 3 dataset is already larger than the entire Run 2 dataset, and is ready for analysis. An analysis using the combined Run 2 plus 3 dataset, is already underway, and should help to resolve these various excesses, and maybe even discover SUSY!

Seeking something missing, missing something left behind.

The French Dispatch

6 $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger Optimisation

Contents

6.1 The ATLAS $E_{\mathbf{T}}^{\mathbf{miss}}$ Trigger 154
6.1.1 Level 1 $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger 155
6.1.2 HLT $E_{\mathrm{T}}^{\mathrm{miss}}$ Algorithms 155
6.1.3 Run 2 Primary Triggers 157
6.2 Pufit Algorithm Optimisation 158
6.2.1 Datasets and MC Samples 160
6.2.2 Optimisation 162
6.3 Performance in Run 3 data 166

Triggers determine what events get recorded by ATLAS in real-time collisions, and so it is key to ensure that these are well optimised and efficient. The ATLAS trigger group is always involved in ongoing checks on data in order to the triggers are performing as desired. Additionally, revisiting the existing triggers to evaluate areas of improvement, is a large effort across all trigger signatures in ATLAS.

ATLAS has a wide menu of triggers, targeting different event topologies, for use in a wide range of analysis projects. In many ATLAS analyses, the final states of interest include particles that only interact very weakly. Missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) is a key trigger signature used to target these 'invisible' particles, which escape the detector without leaving any energy behind. Many BSM theories predict large $E_{\mathrm{T}}^{\mathrm{miss}}$, as a necessary consequence of predicting dark matter candidates.

The analysis described in Chapter 5 is an example of one which utilises the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, where the targeted SUSY signal model provides large $E_{\mathrm{T}}^{\mathrm{miss}}$ from the lightest neutralino ($\widetilde{\chi}_1^0$). The focus of the chapter will be on the ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ Triggers in Run 2, and on efforts to improve their performance for the beginning of Run 3.

6.1 The ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger

The ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger targets invisible particles that do not interact with the sub-detectors of ATLAS. At the LHC, the protons collide in the z direction, and so the momentum in the transverse $x-y$ plane must sum to zero. The $E_{\mathrm{T}}^{\mathrm{miss}}$ of an event can therefore be evaluated by calculating the negative vector sum of the transverse momentum of all the visible particles in the event,

$$E_T^{\mathrm{Miss}} = |\vec{p}_T^{\mathrm{Miss}}| = -|\sum_i \vec{p}_{T,i}^{\mathrm{Miss}}|. \tag{6.1}$$

The calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$ relies on the accurate reconstruction of the energy of all the objects from the primary collision (also called hard scatter objects), making it a complex task to carry out quickly. We note that the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated at trigger level is necessarily different from the $E_{\mathrm{T}}^{\mathrm{miss}}$ used in ATLAS analyses. In order to distinguish between these, in this section they are assigned different labels. "Online" $E_{\mathrm{T}}^{\mathrm{miss}}$ is used to refer to the trigger level $E_{\mathrm{T}}^{\mathrm{miss}}$, and "offline" $E_{\mathrm{T}}^{\mathrm{miss}}$ is used to refer to the $E_{\mathrm{T}}^{\mathrm{miss}}$ as described in Chapter 3, and used in the analysis in Chapter 5. This difference is because the online, real time data taking conditions that the triggers operate in do not allow for the time intensive reconstruction of all the necessary physics objects, especially tracks. Instead, detector level information, with some limited reconstruction, is used to calculate the online $E_{\mathrm{T}}^{\mathrm{miss}}$ of an event. Detector resolution, noise and pileup objects all impact the measurement of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and need to be accounted for in the calculation. The exact details of the online $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation varies depending on the algorithm used, and the different methods are discussed in Section 6.1.2. Another important distinction is that the online $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation treats muons as invisible in the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$. This is in part due to historical reasons, as the Run 1 and 2 $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers only utilised calorimeter

information, and muons leave little energy in the calorimeters so appear invisible. Additionally, this avoids a large "fake" $E_{\mathrm{T}}^{\mathrm{miss}}$ rate in the trigger, from high p_{T} non isolated muons. This is due to the fact that high p_{T} muons have straight tracks and so have large uncertainty associated with the measured p_{T}. Including these in the visible E_T throws off the kinematic balance in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation and artificially increases the $E_{\mathrm{T}}^{\mathrm{miss}}$, leading to higher trigger rates.

6.1.1 Level 1 $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger

All HLT triggers in ATLAS are built upon an L1 Trigger. The primary L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger sums over the energy of the L1 calorimeter trigger towers, see Section 2.3.1, to calculate the $E_{\mathrm{T}}^{\mathrm{miss}}$. The L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger was not the subject of optimisation efforts and the primary L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, `L1XE50`, is applied throughout the work described in this Chapter.

6.1.2 HLT $E_{\mathrm{T}}^{\mathrm{miss}}$ Algorithms

In addition to an L1 trigger, the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers also have requirements on one or more HLT $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithms. The trigger needs to be able to accept/reject events in a 25 ns time window, so needs to be both fast and efficient. The threshold applied to a given algorithm is chosen based on these considerations. Additionally, the calculation of online $E_{\mathrm{T}}^{\mathrm{miss}}$ is particularly computationally intensive, as it requires information from the entire detector. Running full detector tracking on an event is the single most CPU intensive part of the ATLAS trigger [166], and so choosing an appropriate threshold is key.

Cell based algorithm

In order to reduce the CPU time used by the trigger, an intermediate preselection algorithm called 'cell' was developed. A cell $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is used as preselection before a requirement on another more computationally intensive algorithm. Cell calculates the $E_{\mathrm{T}}^{\mathrm{miss}}$ by summing over the individual calorimeter cells, see Section 2.2.4, with requirements on the cell energy to reduce the impact of detector noise.

The noise suppression is applied such that cells are only included if they pass the energy requirements: $|E_i| > 2\sigma_i$, $E_i > -5\sigma_i$, where σ_i is the pileup dependent expected energy-equivalent noise in each cell [36].

The MHT algorithm

The 'mht' algorithm was the primary algorithm at the start of Run 2. Triggers using this algorithm were not looked at in the optimisation studies in this thesis, but are relevant for analysis in Chapter 5. The inputs to the sum are trigger level jets, using the anti-kt algorithm with R = 0.4, and topoclusters as inputs. The jets are calibrated and have an area-based pile-up subtraction applied [43]. This algorithm was used in the primary trigger during the early years of Run 2, but as pileup increased the rates proved too high, and the pileup-fit-based algorithms were developed to supersede it [167].

Pileup Fit based algorithms

The pileup fit (Pufit) algorithm was developed in order to reduce the impact of pileup on the trigger. MET trigger rates increase exponentially with pileup, as the rate is dominated by exponential resolution effects [167]. This was confirmed during early Run 2 (2016) and as a result an algorithm employing pileup suppression, 'Pufit', was deployed in order to avoid applying stringent $E_{\mathrm{T}}^{\mathrm{miss}}$ thresholds that would lead to significant losses in signal efficiency. The two algorithms using Pufit were optimised, TCPufit and PFOPufit, which follow the same basic principles to account for pileup, but utilise different inputs. For TCPufit, topoclusters are used as inputs (see Section 3.2.3), and for PFOPufit, PFOs (see Section 3.2.4) are used. The main difference between the two inputs is that PFOs contain information based on tracks, which can explicitly determine if charged objects originate from hard scatter or pileup. Any charged particle objects with tracks associated with the primary vertex are hard scatter, and any not associated with the primary vertex, as well as any neutral track objects, are treated as pileup.

6.1.3 Run 2 Primary Triggers

A HLT trigger chain is defined by combining an L1 trigger with one or more thresholds on the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated by a given algorithm. An example trigger name is: `HLT_xe110_tcpufit_xe65_L1XE50`, where `L1XE50` is the primary L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger with a L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold of 50 GeV, `xe65` corresponds to a 65 GeV threshold on the Cell algorithm $E_{\mathrm{T}}^{\mathrm{miss}}$ and `xe110_tcpufit` is a 110 GeV threshold on the TCPufit algorithm $E_{\mathrm{T}}^{\mathrm{miss}}$. The Run 2 primary $E_{\mathrm{T}}^{\mathrm{miss}}$ Triggers are shown in Table 6.1. The rate of the primary $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers throughout Run 2 is shown in Figure 6.1. The improvement brought by Pufit is clear, with the rate at the highest pileup values much lower than the rate of other algorithms. It was key to continue improving in this area, as ATLAS prepared for the high pileup environment of Run 3 (consistently $\langle\mu\rangle > 60$), and beyond.

For the start of Run 3, the primary trigger became `HLT_xe90_pfopufit_xe65_L1XE50`, which uses PFOs as inputs. This is used alongside the Run 2 primary trigger from the end of 2018 (see Table 6.1), with the latter retained in order to capture events where the online primary vertex differs from the offline, and therefore the PFO's are not useful. In Run 3, the Run 2 Pufit trigger was relabelled to TCPufit to differentiate it from the new PFOPufit algorithm. The optimisation detailed in Section 6.2 investigated both the TCPufit and PFOPufit based triggers.

Table 6.1: A summary of the primary $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger chains across Run 2.

Period		Chain
2015		`HLT_xe70_mht`
2016	A - D3	`HLT_xe90_mht_L1XE50`
	D4 - F1	`HLT_xe100_mht_L1XE50`
	F2 -	`HLT_xe110_mht_L1XE50`
2017		`HLT_xe110_pufit_L1XE55`
2018	B - C5	`HLT_xe110_pufit_xe70_L1XE50`
	C6 -	`HLT_xe110_pufit_xe65_L1XE50`

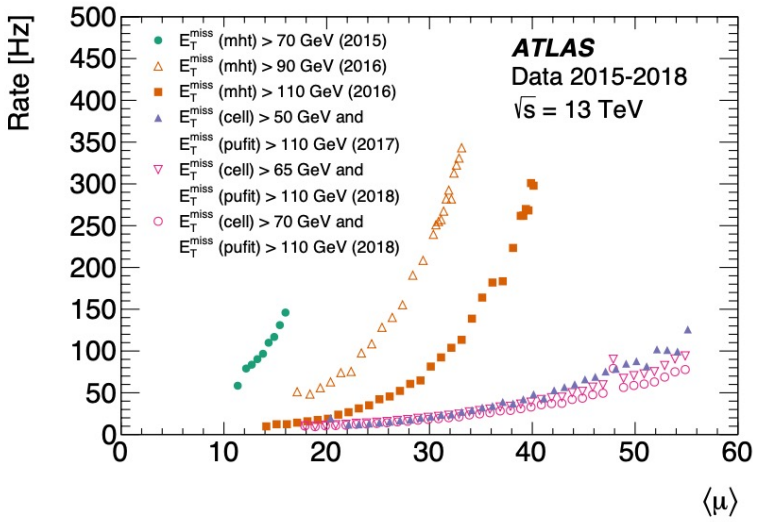


Figure 6.1: Rate as a function of average pileup $\langle\mu\rangle$, for the different primary $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers across Run 2 [167].

6.2 Pufit Algorithm Optimisation

Pufit uses topoclusters binned into "towers" to suppresses the impact of pileup on the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation by carrying out a χ^2 minimisation, with two key constraints:

1. The $E_{\mathrm{T}}^{\mathrm{miss}}$ from pileup collisions sums to zero.

2. The transverse energy from pileup is evenly uniformly distributed across the detector.

Figure 6.2 shows a sketch of the tower E_T and contributions from pileup and hard scatter before and after applying Pufit. A description of the algorithm follows, with more detailed discussion in Appendix G. First, the calorimeter is divided into towers in η, ϕ, with the dimensions currently set at 14×8. So for an event, the value of a tower is the sum of the E_T of all the topoclusters within that tower. Then a trimmed mean transverse energy is calculated;

$$\langle E_T \rangle = \frac{1}{N} \sum_{j=0.05N}^{0.95N} E_{T,j}, \tag{6.2}$$

with towers ordered by E_T and rejecting the top and bottom 5% of towers. Here N is the number of towers before trimming (112 for the above η, ϕ values). The top % of towers are rejected as these are dominated by high energy hard scatter jets, and

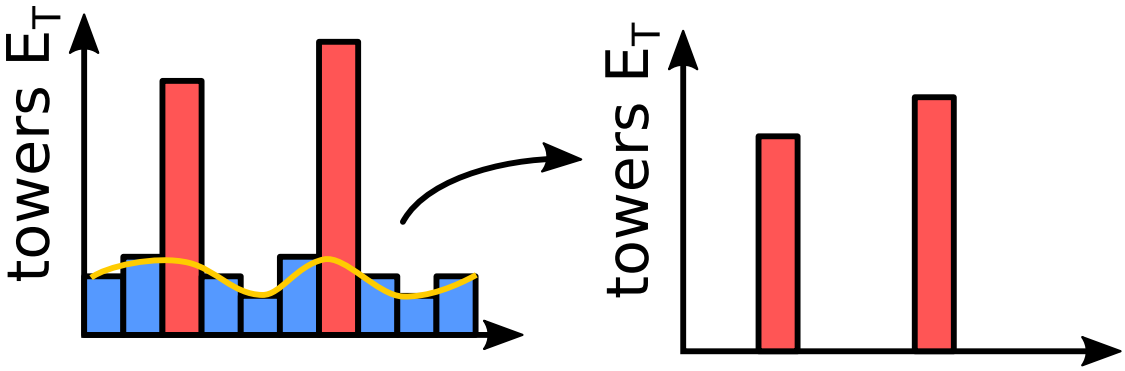


Figure 6.2: Sketch of the tower E_T before and after applying Pufit. The red towers are hard scatter, the blue pileup and the yellow is the estimated pileup contribution to each tower.

the same fraction of bottom towers rejected to avoid biasing the algorithm. This mean is then used to calculate an energy threshold requirement on the towers,

$$E_T^{\text{Th}} = \langle E_T \rangle + n_\sigma \sqrt{V}. \tag{6.3}$$

Below the threshold, a given tower is assumed to be pileup, above the threshold a tower is designated hard scatter. For a typical Run 2 event, this takes a value of around 30 GeV. Here, n_σ (also referred to as "NSigma") is a free parameter found through optimisation, currently set at 5, and V is the variance of the transverse energy in all the towers, with the variance of the top 5% replaced by the bottom 5,

$$V = \frac{1}{N}\left\{ \sum_{j=0.05N}^{0.95N} (E_{T,j} - \langle E_T \rangle)^2 + \sum_{j=0}^{0.05N} (E_{T,j} - \langle E_T \rangle)^2 \right\}. \tag{6.4}$$

The sum over $\sum_{j=0.05N}^{0.95N}$ indicates the sum over the middle 90% of towers, and $\sum_{j=0}^{0.05N}$ indicates the sum over the bottom 5% of towers. Eq. 6.3 and Eq. 6.4 form part of the input into the χ^2 fit. Also included is the calorimeter resolution estimates (see Table 6.2), which are used in the uncertainty on the measured tower E_T's. The χ^2 is then minimised to estimate the pileup contribution in each tower. This contribution is then removed from the measured E_T of each tower, and the $E_{\text{T}}^{\text{miss}}$ is calculated. Further details on the fit are given in Appendix G, and a detailed account can be found in Appendix A of Ref [167]. In Table 6.2 a summary of the configurable parameters of Pufit is shown with some discussion of each parameter.

The effect of varying the NSigma, TrimFraction and Constraint Importance terms are investigated, and reported in this chapter.

Table 6.2: Table of the parameters of the Pufit algorithms, current values, and discussion on their relevance to the work presented here.

Parameter	Purpose	Value	Varied?
NSigma (n_σ)	Sets the Hard scatter vs pileup threshold (Eq.6.3)	5	Yes from 0-6
MaxEta (η_{Max})	Maximum range of eta	5	No - Fixed by detector geometry
NEtaBins (N_η)	Number of towers in eta	14	Previously seen to have minimal effect and limited by processing time
NPhiBins (N_ϕ)	Number of towers in phi	8	Previously seen to have minimal effect and limited by processing time
TrimFraction	Fraction of towers to retain in calculation of the mean (Eq.6.2)	0.9	Yes - from $0.8 - 1.0$
CaloNoise	Coefficient of the Noise term in the calorimeter resolution estimate	50	No - Fixed externally
CaloStochastic	Coefficient of the stochastic term in the calorimeter resolution estimate	15.81	No - Fixed externally
Constraint Importance (s)	Weighing of the $\sum E_T^{\text{Miss}} = 0$ constraint of the χ^2 fit	1	Yes - from $0 - 2$

6.2.1 Datasets and MC Samples

In order to carry out the optimisation a background and target signal need to be decided. For the background pileup is difficult to simulate accurately and so data is used. In this case the choice was made to use an Enhanced Bias (EB) dataset [168]. This consists of collision data taken with a specific selection of triggers that are purposely chosen to evaluate trigger efficiencies and rates, and can be re-weighted to reproduce a background that accurately represents pileup collisions. Using MC for

the pileup background is not done, due to the complex nature of pileup collisions. Detector effects are dominant and very hard to take into account in simulation, as are the soft QCD processes that model pileup collisions. From this dataset we then require the events for the background pass the L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, `L1_XE50`, and come from 'good' luminosity blocks - so the conditions at the time of data taking are consistent and problem free. As this study was carried before the start of Run 3, the specific EB dataset used in the optimisation was taken in Run 2, with average pileup $\langle\mu\rangle = 47$. However, the event reconstruction used in the processing of the dataset reflects running of Run 3 algorithms and reconstruction, making this the best possible proxy to use in lieu of actual Run 3 data.

The signal models used to evaluate performance, target processes with large $E_{\mathrm{T}}^{\mathrm{miss}}$. Two processes were chosen, the primary signal used for optimisation is $t\bar{t}$ MC, with at least one non-hadronic decay. The $t\bar{t}$ process was selected as this provides large $E_{\mathrm{T}}^{\mathrm{miss}}$ from neutrinos. Performance was also evaluated using $Z \to \mu\mu$ MC, which provides $E_{\mathrm{T}}^{\mathrm{miss}}$ from the muons being invisible to the trigger. The results of the optimisation using this alternative MC were similar to the $t\bar{t}$ results, but suffered from low MC statistics and so are not reported here. These MC samples were generated using projected LHC and ATLAS conditions at the beginning of Run 3, and so have a higher average pileup than the background, $\langle\mu\rangle = 55$.

For the signal and background samples, two different types of inputs to the $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithms are used: online and offline. The online inputs run the same reconstruction that is used by the trigger, during data-taking. The offline inputs have more in-depth calibration and event cleaning carried out, giving a more accurate measurement of the object energies, and are more similar to the topoclusters used in analysis object reconstruction. The online PFOs were not available in the samples used in the optimisation, so the PFOPufit algorithm only the offline results are shown. The implementation of Pufit differs slightly for PFOPufit, and these differences are detailed in Appendix G, however, the general concepts and optimisable parameters remain the same.

6.2.2 Optimisation

All ATLAS triggers need to take into account two major factors: the rate of the trigger and the efficiency. The rate is the number of times per second a given trigger is triggered, in Hz. The efficiency is the fraction of signal events which pass the trigger, given a chosen selection,

$$\varepsilon = \frac{\text{N(Trigger \&\& Selection)}}{\text{N(Selection)}}, \tag{6.5}$$

where, N(Selection) is the number of events which pass some event selection, and N(Trigger && Selection) is the number of events which pass the event selection and the trigger requirement. Often the efficiency of a trigger is paramatrised as a function of a relevant variable, and plotted in a 'turn on curve'. Figure 6.3 shows the turn on curve for the trigger `HLT_xe110_tcpufit_L1XE50`, as a function of the offline $E_{\mathrm{T}}^{\mathrm{miss}}$, with muons treated as invisible in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The dataset used to produce this figure is $t\bar{t}$ MC, and so no explicit event selection is applied as the sample is already pure in the targeted event topology. From this plot an offline $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement can be defined, to select a sample of signal events with large $E_{\mathrm{T}}^{\mathrm{miss}}$ for further efficiency studies. This is achieved by determining the region where the trigger efficiency starts to plateau, at about 175 GeV - where it has an efficiency of about 80%. Another key figure of metric to use in optimisation, is the receiver operating characteristic (ROC) curve. This is a measure of the signal acceptance efficiency vs background acceptance, while varying the requirement on the algorithm $E_{\mathrm{T}}^{\mathrm{miss}}$. The aim is to maximise the signal efficiency, while minimising the background acceptance.

To create a ROC curve events which pass the baseline `L1_XE50` trigger are selected, as this is the baseline trigger that the primary $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are built from. Then, the aforementioned offline $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of 175 GeV is implemented in the signal sample only. The purpose of this requirement is to ensure that the trigger efficiency is evaluated on events that have large $E_{\mathrm{T}}^{\mathrm{miss}}$, as it would not be optimal to optimise on signal events with low $E_{\mathrm{T}}^{\mathrm{miss}}$. Next we find the fraction of signal and background events accepted for a given requirement on the HLT algorithm $E_{\mathrm{T}}^{\mathrm{miss}}$

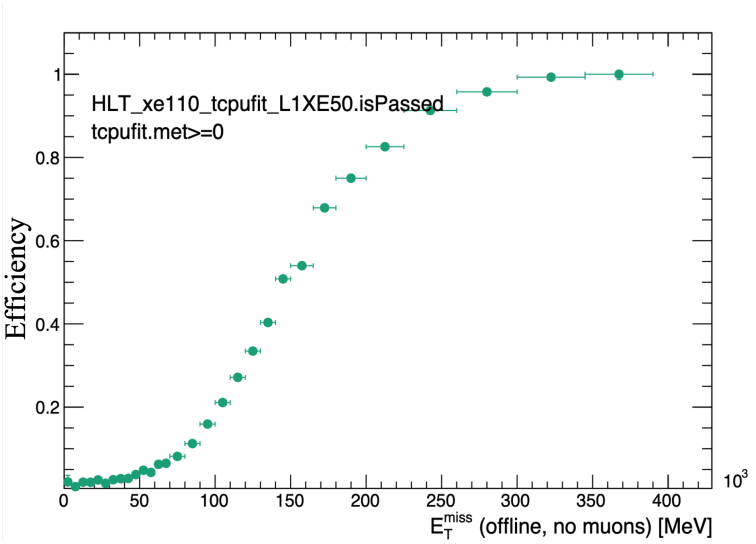


Figure 6.3: Turn on curve showing the efficiency of `HLT_xe110_tcpufit_L1XE50` as a function of offline $E_{\mathrm{T}}^{\mathrm{miss}}$ (with muons included as $E_{\mathrm{T}}^{\mathrm{miss}}$), using $t\bar{t}$ MC.

threshold, then vary this requirement and plot the results to give the ROC curve. In order to evaluate performance for variations of the NSigma parameter, the parameter was varied from $0-6$ with a step of 0.5, then narrowed down to a range from $2.5-5$ using a step of 0.1. The results are shown in Figure 6.4 for both online and offline inputs to TCPufit, for a selection of trial NSigma values. The best performance is achieved when the curve tends towards the bottom right, maximising signal efficiency while minimising background acceptance [1]. The pink line 'tcpufit' represents the current performance, for NSigma = 5. The diamond represents the point of the curve where the Run 2 primary trigger `HLT_xe110_tcpufit_L1XE50` lies and can be used to compare the performance. It can be seen that the NSigma = 3 results show a factor of 2 reduction in background rate, for the same signal acceptance as the Run 2 primary trigger. In the case of TCPufit, these results indicate that changing value of NSigma from 5 to 3 will provide a significant increase in the performance of the algorithm.

The background pass rate and signal efficiency as a function of a cut on the algorithm $E_{\mathrm{T}}^{\mathrm{miss}}$ are also shown in Figure 6.4. It can be seen that the signal efficiency for NSigma = 3 is significantly higher in both the offline and online cases, whereas

[1] Note the difference with respect to the ROC curve definition used in the BDT optimisation in Chapter 5.

the background acceptance remains roughly the same. The abrupt decrease at 0 MeV is due to a number of events in which the algorithm is unable to calculate the $E_{\mathrm{T}}^{\mathrm{miss}}$, usually when the event has no towers above E_T^{Th}.

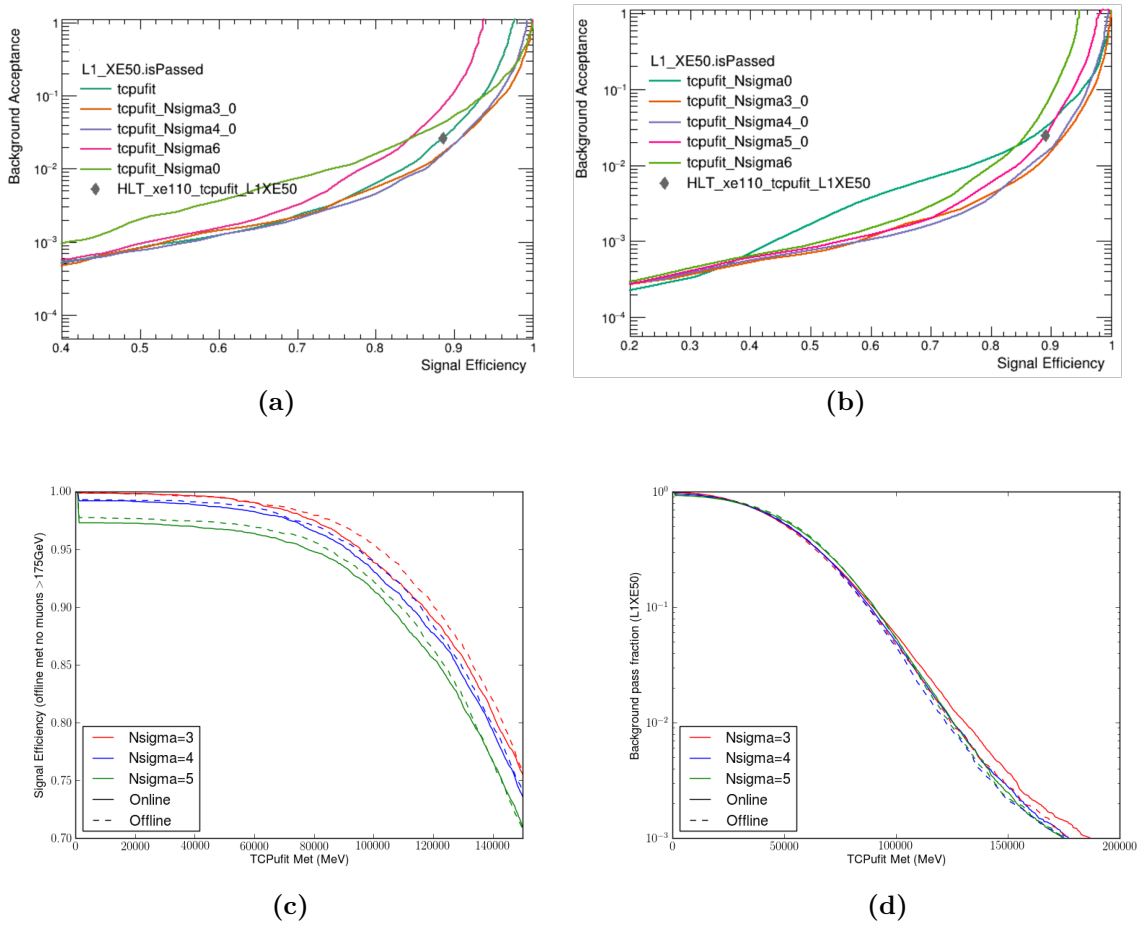


Figure 6.4: The (a) online and (b) offline ROC curves for TCPufit using $t\bar{t}$ MC as signal, and EB dataset as background. The diamond indicates the performance of the existing `HLT_xe110_tcpufit_L1XE50` trigger and NSigma = 5 is the default value of the parameter. Also shown is the (c) signal efficiency, and (d) background pass fraction, for online and offline TCPufit, for variations on NSigma as a function of the algorithm $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold.

A similar optimisation process was repeated for PFOPufit in order to verify the improvement was consistent. The enhanced bias dataset used did not have the online containers necessary to run PFOPufit online, so only the offline results are shown in Figure 6.5. A slightly smaller improvement in efficiency is seen, but this is expected as the PFO's inherently reduce the impact of pileup due to the use of track information. Additionally, results from the variation of the trim fraction and significance terms are shown in Figure 6.6, for TCPufit using the online inputs. Varying these parameters leads to only a small effect on the overall performance.

It was decided to implement four new triggers, for TCPufit and PFOPufit, with NSigma = 3 and NSigma = 4, for the start of Run 3 data taking. These triggers were not immediately set as the primary triggers and were prescaled in order to test the efficiency and acceptance.

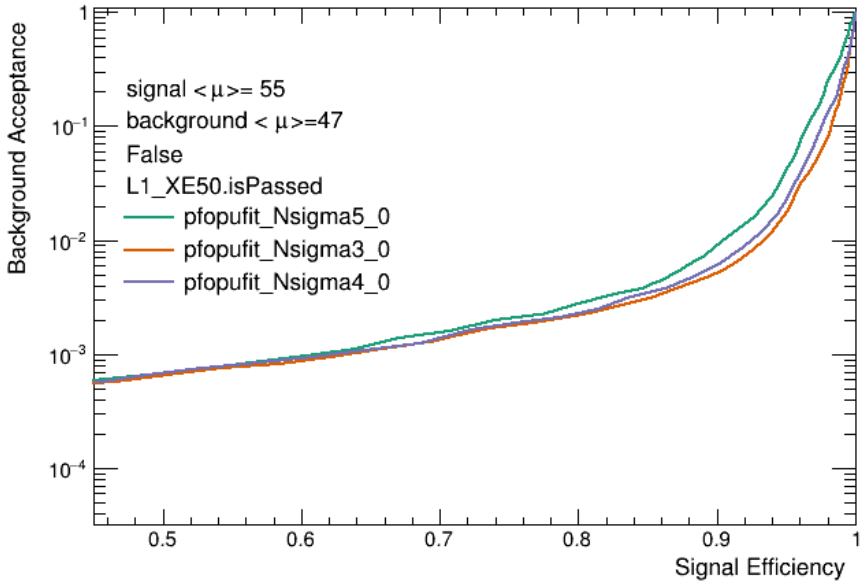
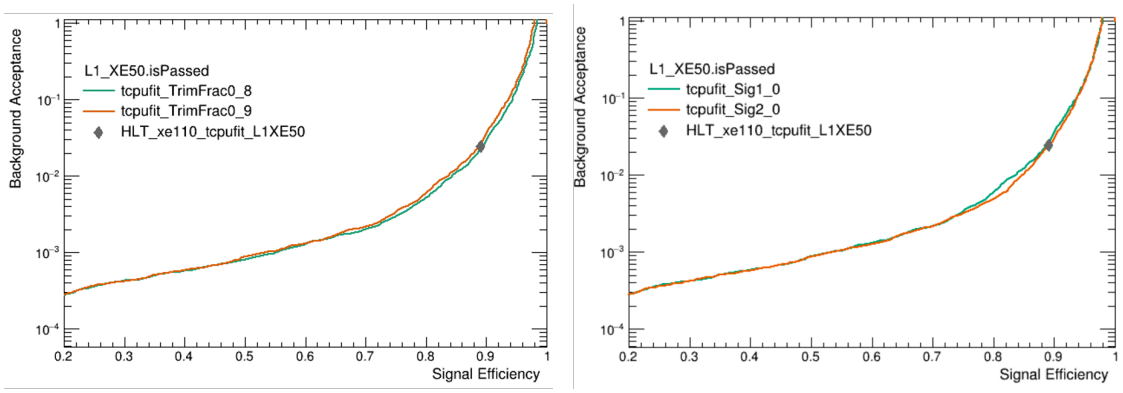


Figure 6.5: Offline ROC curves for variations of NSigma for PFOPufit. Where NSigma = 5 is the default.



(a) **(b)**

Figure 6.6: ROC curves for online and offline TCPufit, for variations on (a) the TrimFraction parameter and (b) The constraint importance parameter. TrimFraction = 1.0 and Constraint Importance = 0.9 are the default values of the parameters.

6.3 Performance in Run 3 data

The performance of the new triggers was also evaluated using Run 3 data. This evaluation was carried out using an EB dataset collected in Run 3 (2022), which has $\langle\mu\rangle = 58$, for the background. For the target signal, Run 3 data is used, again collected in 2022, with the same average pileup as the EB dataset. To evaluate performance using data, selections are applied to target events which are high in $E_{\mathrm{T}}^{\mathrm{miss}}$, in this case $Z \to \mu\mu$ decays. This is achieved by requiring events to pass at least one primary single muon trigger, and restricting the invariant di-muon mass [2], to target the Z resonance: 76 GeV $< m(\mu\mu) <$ 106 GeV. Additionally, all events in both the signal and background samples are required to pass the `L1_XE50` trigger, and we require $\mu \geq 40$ to veto events with low pileup. The ROC curve for this performance evaluation is shown in Figure 6.7, along with the background acceptance rate and signal efficiency.

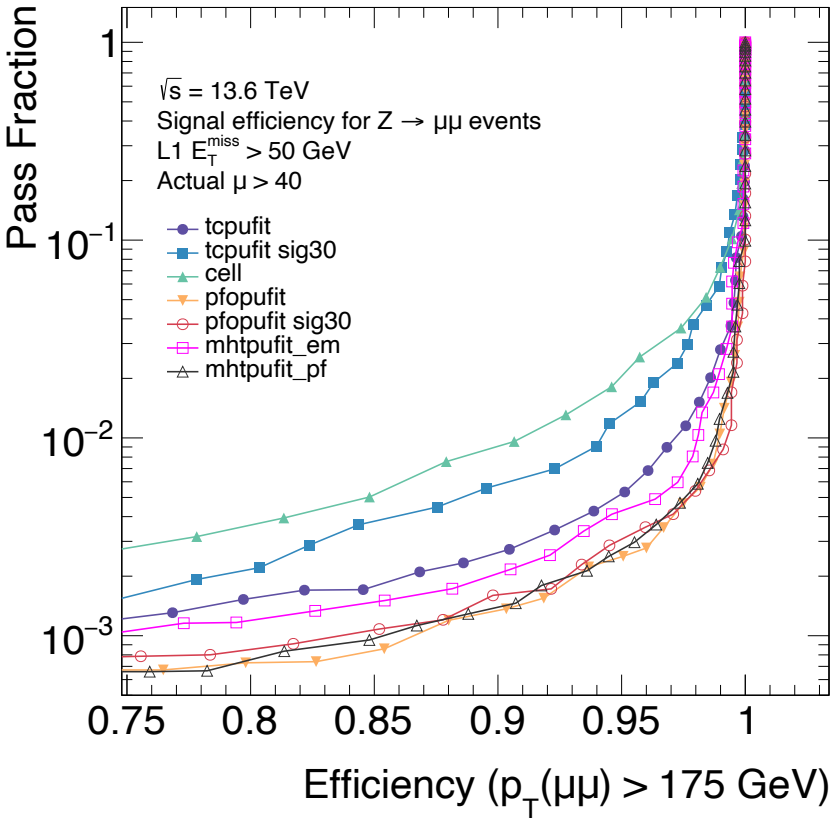
In the Run 3 data, the performance of the new triggers (labelled 'sig_30') was observed to be worse than the existing triggers. The signal acceptance of the nominal triggers compared with the 'sig_30' triggers remains roughly the same, however, the background pass rate increases substantially for the 'sig_30' chains,

[2] Given by $m_{\mu\mu} = \sqrt{(E_{\mu^+} + E_{\mu^-})^2 - (\mathbf{p}_{\mu^+} + \mathbf{p}_{\mu^-})^2}$, where $E_{\mu^\pm}$ and $\mathbf{p}_{\mu^\pm}$ are the energy and momentum vector of the muons, respectively.

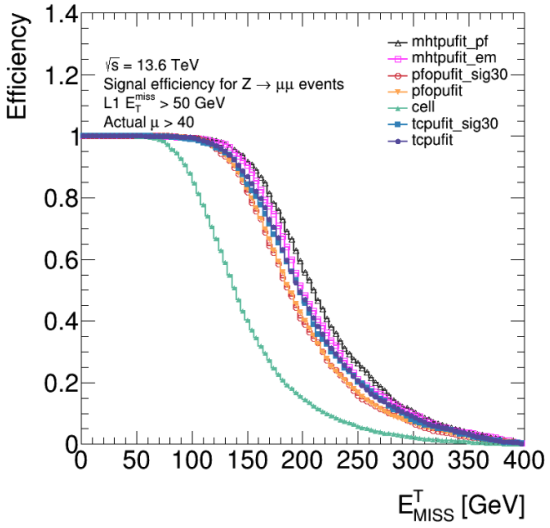
which was not seen in the optimisation studies. This was found to be because the simulated topocluster reconstruction used in the optimisation samples did not accurately match the topoclusters in the Run 3 data.

Further re-optimisation of the parameters of interest here, with an accurate Run 3 reconstruction, was seen to have minimal impact on the performance of the trigger. Though this work did not lead to a new trigger, a lot of information can be inferred from the results. Throughout the years, many efforts have been made to improve the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger beyond the performance of the current Pufit, and all have failed - highlighting the difficult nature of this trigger signature. This indicates that the Pufit algorithm is performing near optimally, and re-optimisation of the parameters for higher pileup conditions will not improve the performance.

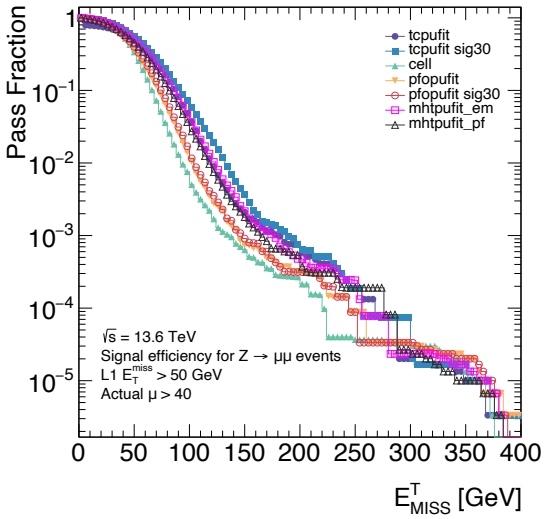
As a result, for the remainder of Run 3, it was decided to refocus efforts onto new ML based triggers, and onto operation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger. The ML triggers utilise the values of the $E_{\mathrm{T}}^{\mathrm{miss}}$ from the various algorithms detailed in this chapter, so maintaining and evaluating the performance of these still remains an important task. Improving the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger remains a key goal for the ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger group, as the trigger will need to be reworked significantly before Run 4, to cope with the large increase in pileup, up to $\langle\mu\rangle = 200$ [169].



(a)



(b)



(c)

Figure 6.7: The (a) ROC Curve, (b) the signal efficiency, and (c) background pass fraction, for various $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers, evaluated using Run 3 data from 'physics main' as the signal and EB data as the background. The 'sig_30' suffix indicates the triggers which use the optimisation detailed in this chapter. The pfopufit and tcpufit lines show the performance of the original NSigma = 5 triggers. These plots were produced by a collaborator.

I wanted to say, in another life, I would have really liked just doing laundry and taxes.

Everything Everywhere All At Once

7 Conclusions

The search for dark matter is one of the most compelling questions in modern particle physics. This thesis presents a comprehensive review of the steps I have undertaken, with the help of many collaborators, to attempt to achieve this goal over my DPhil. It began with an introduction to the LHC, the ATLAS detector, and the processes by which detector signals are reconstructed as particles. Then the focus shifts, providing an overview of the limitations of the Standard Model, and proposing Supersymmetry as a potential solution to some of these. The thesis then examines how data recorded by the ATLAS detector can be used to evaluate sensitivity to simplified Supersymmetric models of sleptons. Finally, the thesis concludes with an investigation into possible improvements to the ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, as we look to improve data-taking efficiency for the future.

The main result of this thesis is the search for simplified supersymmetric models of selectrons and smuons. This analysis targeted a region characterised by small mass splittings between the slepton and neutralino, with $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0) \lesssim m(W)$, where historically there has been a lack of sufficient experimental sensitivity. A machine learning based strategy was employed, involving five BDTs, each optimised for a different region of $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ in the range from 5 to 75 GeV. The search presents results with sensitivity to $\tilde{\mu}$ and $\tilde{e}$ masses between 200 and 350 GeV for models with $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0) > 2\,\mathrm{GeV}$, achieving improvements across in exclusion the most

of the targeted $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ range, even in scenarios where four-fold slepton mass degeneracy is not assumed. These results significantly improve upon previous constraints, effectively closing much of the long-standing gap in sensitivity within the moderately compressed $\Delta m(\tilde{\ell}, \tilde{\chi}_1^0)$ region, especially in the smuon interpretation.

Some SRs show deviations from post-fit background predictions, reducing the excluded regions compared to expectations. The most significant excesses correspond to local single SR significances of $2.0\,\sigma$ in the selectron search and $2.4\,\sigma$ in the smuon search. The highest discovery significances are $1.9\,\sigma$ for the selectron model with $m(\tilde{e}, \tilde{\chi}_1^0) = (375, 335)$ GeV and $2.1\,\sigma$ for the smuon model with $m(\tilde{\mu}, \tilde{\chi}_1^0) = (150, 140)$ GeV.

This search was carried out using the 140 fb^{-1} ATLAS Run 2 dataset, which was recorded between 2015–2018. In order to maintain the success of ATLAS and the LHC, it is important to consistently improve our data-taking techniques as the operation of the LHC continues. The last topic covered in this thesis was an investigation into improving the efficiency of the ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger for the currently ongoing Run 3 data-taking period.

This Run 2 analysis serves as a reminder that the primary objective of LHC searches is the discovery of new particles, not just the exclusion of increasingly large regions of signal space. Should a comparable excess be observed in the Run 3 dataset, it could result in a statistical significance surpassing the observation threshold, potentially even reaching the level required for a discovery. With the excesses from other electroweak supersymmetry searches, Run 3 may yield exciting results in the search for physics beyond the Standard Model, and insight into the nature of dark matter.

The pursuit of answers to the most fundamental questions about the nature of our universe is what has driven me throughout this journey, and it will continue to motivate me in the future. As particle physicists, we can often be sceptical, sometimes even cynical, about the prospect of discovering something new. But every so often in life, we just have to let go and allow ourselves to dream. Let's see what's out there to discover!

Appendices

A
MC Simulation

Contents

A.1 Event Generation 173
A.1.1 Parton Distribution Functions 174
A.1.2 Matrix Element 175
A.1.3 Parton Shower 175
A.1.4 Hadronisation 177
A.1.5 Multi-parton Interactions 177

A.1 Event Generation

The aim of MC simulation is to simulate the production of a target process initiated by the interaction between two partons from the collided protons. This can be interpreted as an estimate of the inclusive cross-section of the target hard scatter process:

$$\sigma_X = \sum_{a,b} \int d\boldsymbol{x} \int PDF\left(x_a, \mu_F\right) PDF\left(x_b, \mu_F\right) d\hat{\sigma}_{(ab\to X)}\left(\mu_F, \mu_R\right), \tag{A.1}$$

for a targeted process $a + b \to X$, where a and b are two partons, and X is the final state particle. In this equation the terms are defined as follows:

- x is the fraction of proton momentum carried by the parton.

- μ_F, μ_R are the factorisation and renormalisation scales respectively, typically chosen to be near the scale of the targeted final state. For example, for $t\bar{t}$ this is chosen as: $\mu_R = \mu_F = \sqrt{m_{\text{top}}^2 + p_{\text{T}}^2}$, where p_{T} is the transverse momentum of the top before any radiation. The factorisation scale controls the separation between long-distance (non-perturbative) and short-distance (perturbative) physics in QCD, affecting the parton distribution functions, while the renormalisation scale sets the scale at which α_s is evaluated up to in the matrix element.

- $PDF\,(x, \mu_R)$ are the parton distribution functions (PDFs) of the two protons, which model the fraction of momentum carried by the partons in the proton at a chosen energy scale controlled by μ_F.

- $d\hat{\sigma}_{(ab\to X)}$ is the differential cross-section of the process, which depends on the matrix element (ME) and the final state phase space.

Various MC generators are available to perform event generation. While they follow broadly similar principles, each employs different techniques with its own advantages and limitations.

A.1.1 Parton Distribution Functions

The PDF is a key element of MC generation, and estimates the probability of finding a parton carrying a given fraction of the proton's momentum at a specific energy scale. By taking experimental data from various sources, such as the LHC experiments, deep inelastic scattering experiments at HERA [170], and precision electroweak experiments at LEP [171], complex multidimensional fits are carried out to determine PDF parameters. This data is typically taken at a different energy scale (Q^2) than we want to apply it to, and so the PDF must be evolved to higher Q^2 using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [172]. This evolution introduces a dependency to the PDF on the chosen factorisation scale (μ_F). The chosen PDF should match the order of the matrix element

calculation. Some examples of PDFs used in Chapter 5 are: NNPDF2.3LO [173], NNPDF3.0NLO [174] and PDF4LHC [175].

A.1.2 Matrix Element

The matrix element (ME) parameterises the parton–parton interaction with the largest momentum transfer in the proton–proton collision – the hard scatter interaction. The matrix element is calculated using perturbative QCD, as the strong coupling constant becomes sufficiently small in high energy proton–proton collisions, perturbation theory can be used to calculate cross-sections. This enables the interaction to be expressed as a series of terms, with each successive term in the expansion corresponding to a higher order in α_s, and so a higher level of precision. The calculation of the ME is typically carried out at one of the following orders:

- Leading-order (LO) – partons only generated up to leading order in α_s at ME level
- Next-to-leading-order (NLO) – one additional order in α_s at ME level
- Next-to-next-to-leading-order (NNLO) – two additional orders in α_s at ME level.

Examples of commonly used generators to obtain the matrix elements are: POWHEG BOX v2 [176–179], SHERPA [180, 181], and MADGRAPH5_AMC@NLO [127].

A.1.3 Parton Shower

After generation of the matrix element hard scatter event, the partons undergo parton showering. The parton shower (PS) simulates the complex QCD dynamics associated with the emission of additional soft (low–energy) and collinear (small–angle) gluons from outgoing partons. These corrections can be large as they give rise to logarithmic terms in the perturbative expansion of the cross-section. Parton showering renormalises large groups of these logarithmic corrections, giving a more accurate simulation of the shower. Similar to the matrix element, the perturbative expansion can occur up to any number of terms, typically the PS is evaluated at:

- Leading-logarithmic (LL) – only the leading log (i.e. $\log^N$) terms are resummed
- Next-to-leading-logarithmic (NLL) – the leading and subleading log power ($\log^N$ and $\log^{N-1}$) terms are resummed
- Next-to-next-to-leading-logarithmic (NNLL) – the first three log power ($\log^N$, $\log^{N-1}$ and $\log^{N-2}$) terms are resummed.

Common parton shower generators used are: SHERPA which uses a method based on dipole radiation, where each emission is modelled as coming from a pair (dipole) of partons. This approach provides a detailed and accurate way to simulate how partons radiate. Another commonly used shower generator is PYTHIA [128], which uses the p_{T}-ordered parton shower method, meaning it orders emissions by their transverse momentum. This method is based on the DGLAP equations and also describes how colour charge affects radiation (colour coherence).

The parameters that control the behaviour of the parton shower need to be tuned to experimental data, typically using minimum-bias and underlying event measurements from LHC experiments. Tunes typically used by ATLAS for PYTHIA are the A3 set of tuned parameters [182] or the A14 set of tuned parameters [129]. For SHERPA the default tune is used [183], and for POWHEG BOX the AZNLO tune is used [184].

Often different MC generators are used to carry out the parton shower and generate the matrix element. This introduces an ambiguity as to what energy scale to switch from using one to the other to describe the parton evolution. For example, in ATLAS $t\bar{t}$ is simulated using POWHEG BOX for the matrix element, and PYTHIA for the shower. The partons generated by POWHEG BOX need to be passed to PYTHIA, through a process called matching. Matching between the matrix element and parton shower is typically performed following the CKKW-L scheme [185]

A.1.4 Hadronisation

Once the produced partons lose enough energy through parton showering, hadronisation can occur. Hadronisation is the process of forming colour-neutral hadrons from partons and is only possible at low Q^2, where α_s is large and QCD becomes non-perturbative. The formation of hadrons is described by phenomenological models, the most commonly used being the Lund string model [186, 187], used by Pythia, and the cluster model [188, 189], used by Herwig and Sherpa. In the Lund string model, the QCD field between two partons is modelled as a string with a kink. The kink represents a gluon at the interaction vertex, which produced the colour connected $q\bar{q}$ pair. As the quarks move further apart and the interaction becomes non perturbative, the strength of the QCD field grows until it becomes possible for $q\bar{q}$ pairs to tunnel out of the QCD vacuum. Then the string starts to break, from the kink outwards, until only colourless states are left.

The cluster hadronisation model utilises the colour pre-confinement of the parton shower [190]. The states that the colour structure of the shower is independent of the energy of the HS and only determined by the colour composition of the partons involved in the HS. So once the non-perturbative regime begins, colour singlet clusters begin to form from the partons, where the mass distribution depends on the energy scale of the parton shower. A cluster splits (fissions) if the mass and potential phase space is sufficient. Otherwise, clusters must colour reconnect with neighbouring clusters before hadrons can be formed.

In both cases, if the created hadrons are unstable, they subsequently decay into lighter particles and hadrons. Additional simulation software has been developed for the decays of bottom and charm hadrons, in order to include complex oscillation and CP violation effects. In ATLAS this is typically carried out using EvtGen [191].

A.1.5 Multi-parton Interactions

In a proton-proton collision at the LHC, the hard scatter process is often accompanied by multiple parton interactions (MPIs), which are additional parton-parton scatterings within the same proton-proton collision. Most MPIs are low-p_{T}

(soft) and contribute to the underlying event. These interactions, along with non-perturbative effects like colour reconnection, can increase the multiplicity and alter the distribution of final-state particles. Colour reconnection is the process through which the colour connections between partons are rearranged prior to hadronisation in order to minimise the total string length. Additional interactions can also be initiated by the remnants of the protons involved in the HS event. These can be relatively high-p_{T} (hard) and are possible sources of MPIs, and are called the underlying event background. MPIs are challenging to simulate because they involve both perturbative and non-perturbative interactions and the difficulty in accurately modelling the interface between the two regimes. As a result, MC generator tuning extensively uses minimum bias and underlying event data to evaluate the modelling of MPIs.

B

Flavour Dependency of m_{T2}

Contents

B.1 The graphical approach 180
B.2 The mathematical approach 185

After the initial work on the BDT_{5+10} training, further investigation into the training was needed, due to the large difference between the SFOS and DFOS lepton distributions, shown in Figure B.1. At high BDT score, the DFOS FNP background does not decrease, whereas the SFOS background does. As the BDT is trained on SFOS leptons the BDT must be learning a feature of the inputs that introduces lepton flavour dependency. This appendix details the investigation into how m^0_{T2} leads to lepton mass dependency in the BDT score.

The m_{T2} of a two-body decay is an upper bound on the mass of a pair-produced parent particle ($m(\tilde{\ell})$), under the assumption of a common invisible particle mass ($m_{\widetilde{\chi}^0_1}$). Importantly, the calculation of m_{T2} requires the minimisation of the partitioning of $\mathbf{p}^{\mathrm{miss}}_{\mathrm{T}}$, between the two visible lepton directions. This minimisation is usually carried out numerically using four vector information, however in some kinematic scenarios, it is possible to prove that this minimisation problem simplifies to a true minimum, as described in [140]. Using the results of the minimisation,

the transverse masses are calculated, and the larger one selected as the value for m_{T2}. A full definition of m_{T2} is given in Section 5.3.1.

As a result of this investigation, the lepton mass was set to 0 in all variants of m_{T2} in the analysis. To provide the whole picture, two approaches are described, a graphical approach - Section B.1, and a mathematical approach - in Section B.2.

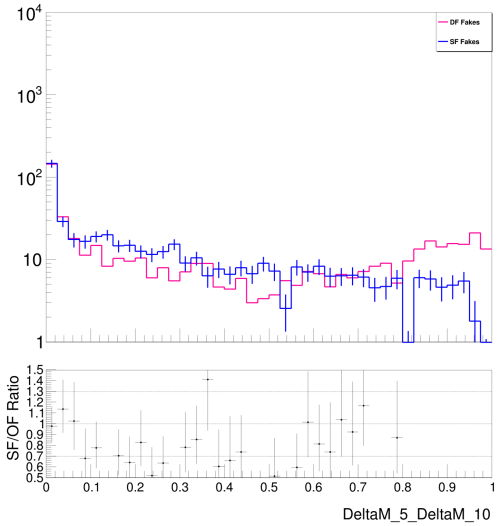
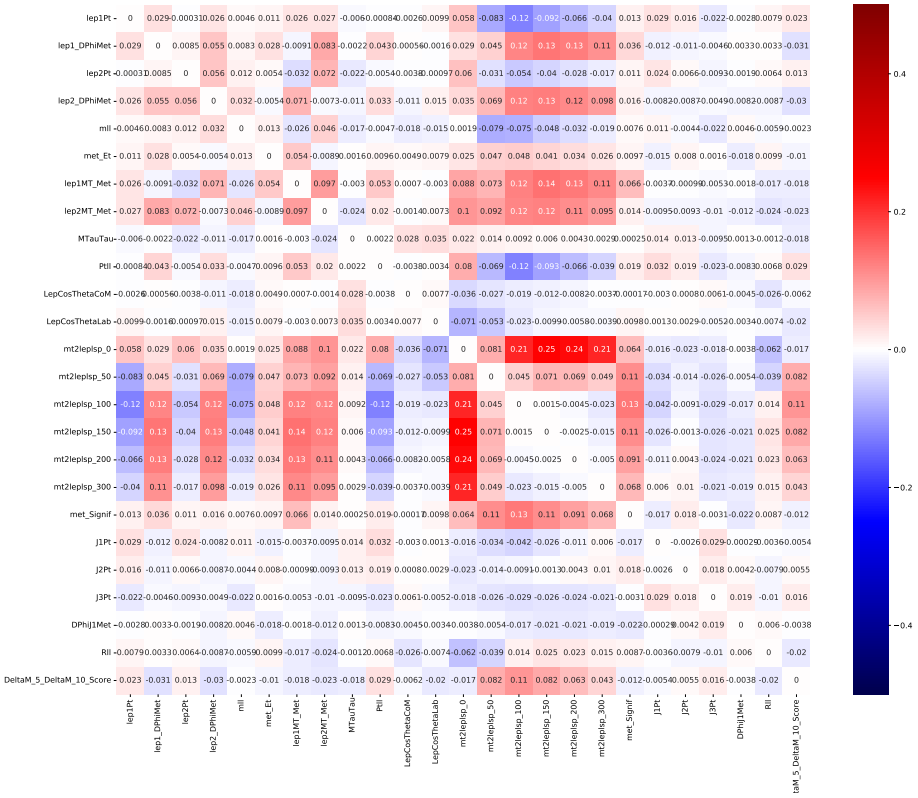


Figure B.1: The SFOS and DFOS lepton BDT score distribution for the data driven fakes background, with the training preselection applied.

B.1 The graphical approach

By comparing the correlation between input variables, in the SFOS and DFOS lepton selections, it is possible to check differences between the variables in the two selections. This is shown in Figure B.2, with the correlation calculated using the Pearson correlation coefficient [192]. It can clearly be seen that there are significant differences in the correlation of the m_{T2} variables between the SF and DF Fakes, in particular for the m^0_{T2} variant.

When looking at the m^0_{T2} distribution, the immediate feature is the large peak in events in the first bin, as seen in Figure B.3a for VR-DF-BDT$_{5+10}$. Previously, it was assumed that there was no structure in this bin, but after the results of the correlation matrix this was investigated closer. Upon plotting the first bin with finer binning, Figure B.3b it is clear there is structure at values m^0_{T2} near zero.



(a)

Figure B.2: The difference between the correlation of the input variables in the SF and DF selections, for the FNP events only.

The FNP lepton background is dominant at low-p_{T}, which is why this effect is largest when looking at this background, but all other major backgrounds, like diboson and $t\bar{t}$ also exhibit this lepton mass dependency. The $t\bar{t}$ background was chosen to carry out the studies which follow. The final states are split into 4 possibilities, ee, $e\mu$, μe, $\mu\mu$, where the first lepton is leading, the second is subleading. This is shown in Figure B.4a, which shows low values of m^0_{T2}, with lepton $p_{\mathrm{T}} <$ 12 GeV. Here it can be seen that there is clear separation between the distributions of ee, and μe, with $e\mu$ and $\mu\mu$. Upon reversing the lepton p_{T} cut, Figure B.4b, the separation reduces significantly but still remains, so this behaviour is a feature at any p_{T} but enhanced at low p_{T}.

The effect is also specific to m_{T2} with an invisible particle mass of zero. Other

invisible mass variants of m_{T2} are shown in Figure B.5 and show that the difference in lepton selections decreases as the invisible particle mass increases. This explains why the mass dependency is largest in m^0_{T2}.

Figure B.6 shows m^0_{T2} recalculated with the lepton masses set to zero. This removes the difference in the distributions, proving that this effect is due to lepton mass dependence in the calculation of m_{T2}.

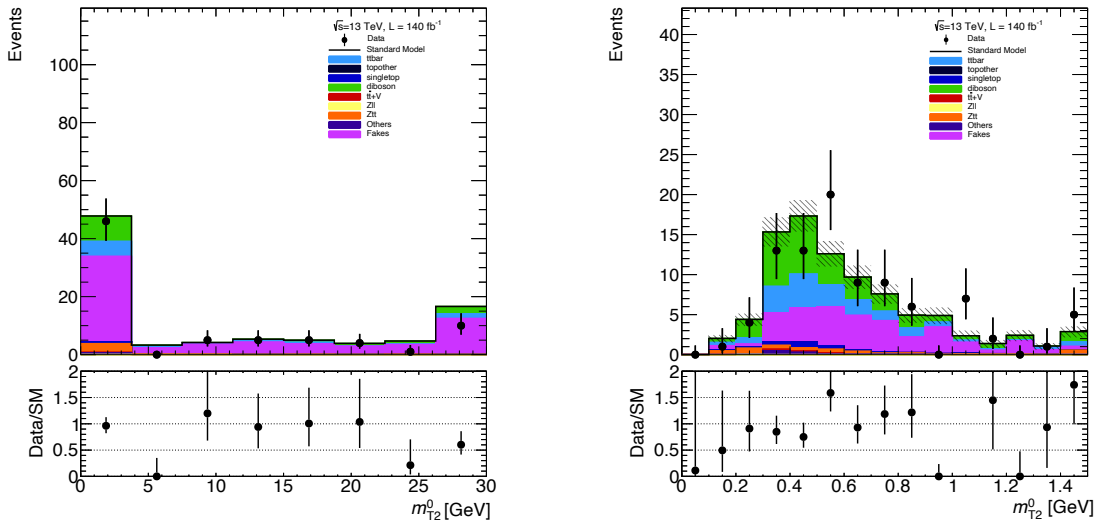
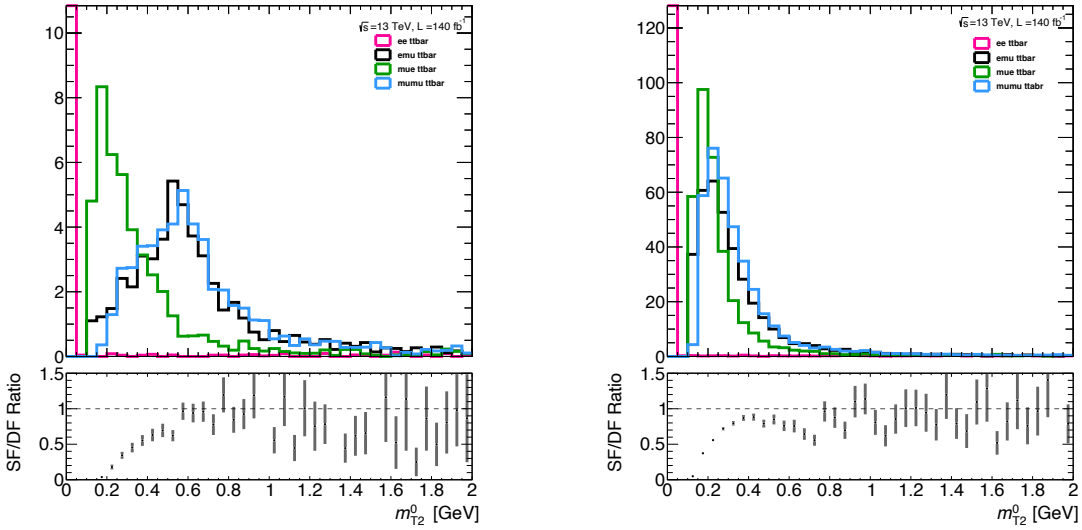
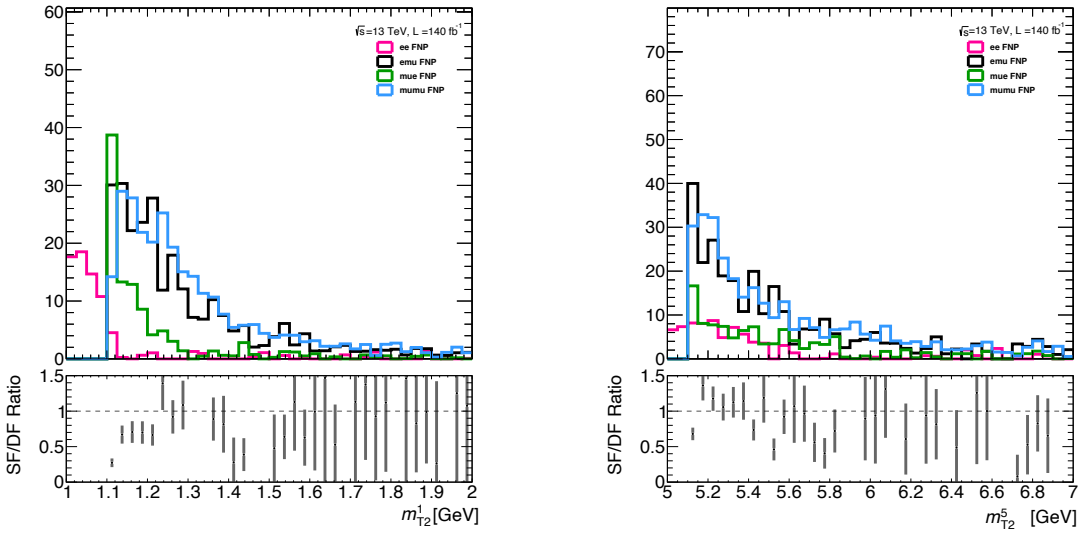


Figure B.3: The m^0_{T2} for VR-DF-BDT$_{5+10}$ in (a) the full range of the distribution and in (b), for just the first bin of (a). The low- and high-p_{T} VR-DFs are combined to improve statistics.



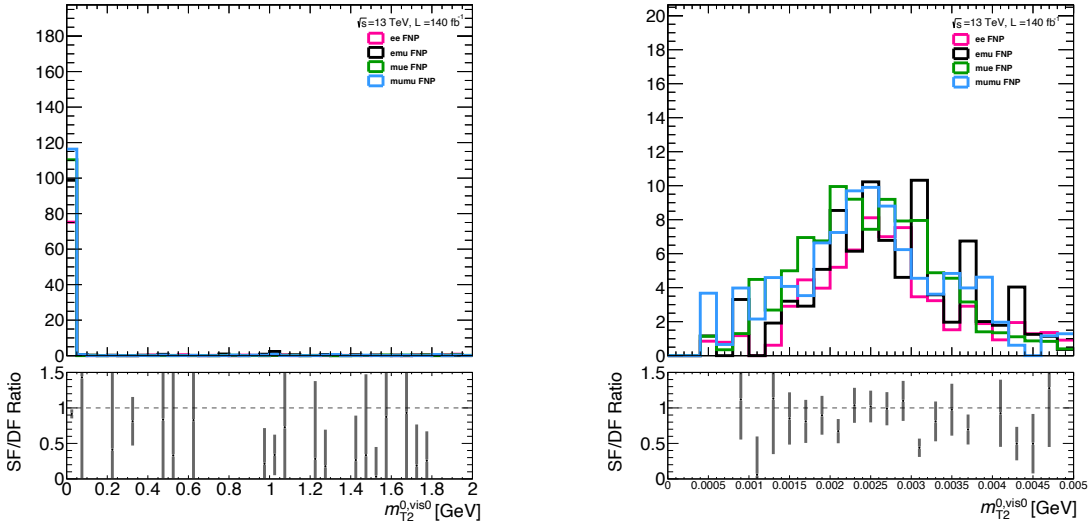
(a) **(b)**

Figure B.4: The m^0_{T2} of $t\bar{t}$, (a) requiring lepton $p_{\mathrm{T}} < 12$ GeV and (b) requiring lepton $p_{\mathrm{T}} > 12$ GeV.



(a) **(b)**

Figure B.5: Distributions obtained for (a) m^1_{T2} and (b) m^5_{T2} - recalculating m_{T2} with different invisible particle masses. Both distributions are for the $t\bar{t}$ background and require lepton $p_{\mathrm{T}} < 12$ GeV.



(a) **(b)**

Figure B.6: The m_{T2}^{0} distribution for $t\bar{t}$, recalculated with all lepton masses equal to zero. In (b), the first bin of (a) is shown with finer bin.

Lepton mass dependency of m_{T2} is not unexpected, but often not noticed in traditional cut-and-count analyses. The BDT is able to exploit these differences at small m_{T2}^{0}, and hence this effect needs to be understood and taken into account in the training. There are known edge cases of m_{T2}, where specific kinematic configurations give rise to specific results [140]. Of the results in this paper, Lemma 5 is most relevant to this analysis. It states that if the total invisible transverse momentum vector ($\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$) of a system is parallel to one of the leptons visible lepton transverse momentum vector $\mathbf{v}_{\mathrm{T}}$, then there exists a partition of the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ such that it is entirely partitioned in the direction of this lepton. In this configuration the transverse masses of both particles take their minimum values, which is proportional to the lepton mass. This is a common kinematic configuration in the BDT_{5+10}, due to the high-p_{T} ISR jet which leads to highly boosted but low-p_{T} leptons. A graphic showing an example of this common kinematic scenario is in Figure B.7.

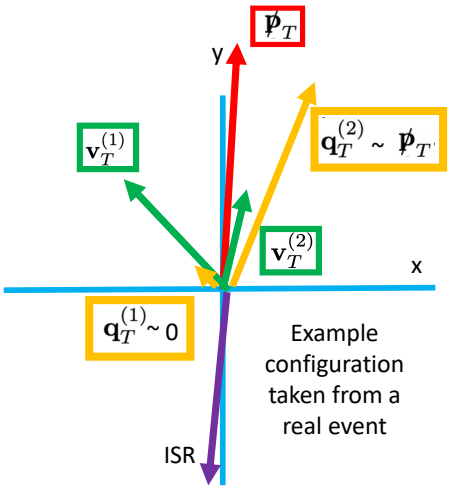


Figure B.7: Kinematic configuration for an example event, taken from data

B.2 The mathematical approach

In this section, general equations for m_{T2} in the scenario described by Lemma 5, will be derived. This will explicitly show that lepton mass dependence occurs regardless of if the leading or subleading transverse mass larger, and occurs regardless of background process or lepton p_{T}. It is worth revisiting the definition of the transverse mass in notation consistent with the m_{T2} calculation:

$$m_T^2 = m_v^2 + m_i^2 + 2\left(e_v e_i - \mathbf{v}_T \cdot \mathbf{q}_T\right), \tag{B.1}$$

where m_v and $m_i(=0)$ are the visible and invisible particle masses respectively. Then, $\mathbf{v}_T$ is the visible lepton momentum and $\mathbf{q}_T$ is the partition of the total $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ assigned to the invisible particle, through the m_{T2} minimisation procedure, such that $\mathbf{q}_T^{(1)} + \mathbf{q}_T^{(2)} = \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. Finally, e_v and e_i are the visible and invisible energies respectively, defined as:

$$e_v^2 = m_v^2 + \mathbf{v}_T^2 \quad \text{and} \quad e_i^2 = m_i^2 + \mathbf{q}_T^2. \tag{B.2}$$

By substituting Equation B.2 to B.1 and simplifying we obtain,

$$m_T^2 = m_v^2 + 2\left(\sqrt{m_v^2 + \mathbf{v}_T^2}|\mathbf{q}_T| - \mathbf{v}_T \cdot \mathbf{q}_T\right). \tag{B.3}$$

From here we can apply the result of Lemma 5, and assume for the leading lepton: $\mathbf{q}_T^{(1)} \to 0$. Then we arrive at an equation for the m_T of the leading lepton in the case where it is not parallel to the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. Where it is seen the m_T directly equals the leading lepton mass:

$$m_T^{(1)} = m_v^{(1)}. \tag{B.4}$$

Next for the subleading lepton, using $\mathbf{q}_T^{(2)} \to \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$,

$$m_T^2 = m_v^2 + 2|\mathbf{q}_T| \left(\sqrt{m_v^2 + \mathbf{v}_T^2} - \mathbf{v}_T \right), \tag{B.5}$$

$$m_T^2 = m_v^2 + 2|\mathbf{q}_T||\mathbf{v}_T| \left(\sqrt{\frac{m_v^2}{\mathbf{v}_T^2} + 1} - 1 \right), \tag{B.6}$$

$$m_T^2 = m_v^2 + 2|\mathbf{q}_T||\mathbf{v}_T| \left(1 + \frac{m_v^2}{2|\mathbf{v}_T|^2} - 1 \right), \tag{B.7}$$

$$m_T^2 = m_v^2 + \frac{|\mathbf{q}_T|}{|\mathbf{v}_T|} m_v^2, \tag{B.8}$$

$$m_T^{(2)} = \sqrt{1 + \frac{|\mathbf{q}_T^{(2)}|}{|\mathbf{v}_T^{(2)}|}} m_v^{(2)}. \tag{B.9}$$

There is also clear mass dependency in the m_T of the subleading lepton. These results show that the lepton mass becomes the dominant contibution to the m_T of both leptons, so regardless of which is the maximum and picked by m_{T2}, lepton mass dependancy enters into m_{T2}^0. The results are general for any case where the subleading lepton is approximately parallel to the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. For the less likely case, where the leading lepton is parallel instead, the same equations hold but with the lepton labels swapped.

From here, we can find which m_T is generically larger, and therefore is the value chosen by m_{T2}. Table B.1 summarises the leading and subleading m_T forms for each flavour combination.

From here, we can use our knowledge of the kinematics of the events to explain Figure B.4a. The $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of the analysis is $\geq$ 200 GeV, and so for low-p_{T} leptons we can assume: $\frac{|\mathbf{q}_T|}{|\mathbf{v}_T|} \gg 1$. So the equations from Table B.1 can be simplified to arrive at the following:

Table B.1: The m_T for the leading and subleading lepton, for the specific kinematic case where the subleading lepton is parallel to the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$.

Flavour Combination	$m_T^{(1)}$	$m_T^{(2)}$
ee	m_e	$\sqrt{1+\frac{\lvert\mathbf{q}_T^{(2)}\rvert}{\lvert\mathbf{v}_T^{(2)}\rvert}}m_e$
$e\mu$	m_e	$\sqrt{1+\frac{\lvert\mathbf{q}_T^{(2)}\rvert}{\lvert\mathbf{v}_T^{(2)}\rvert}}m_\mu$
μe	m_μ	$\sqrt{1+\frac{\lvert\mathbf{q}_T^{(2)}\rvert}{\lvert\mathbf{v}_T^{(2)}\rvert}}m_e$
$\mu\mu$	m_μ	$\sqrt{1+\frac{\lvert\mathbf{q}_T^{(2)}\rvert}{\lvert\mathbf{v}_T^{(2)}\rvert}}m_\mu$

- For ee: $m_{\mathrm{T2}} = m_T^{(2)} \approx \sqrt{\frac{|\mathbf{q}_T^{(2)}|}{|\mathbf{v}_T^{(2)}|}}m_e$ Subleading m_T is always larger due to the extra factor, gives very small values of m_{T2} due to electron mass dependence.
- For $e\mu$: $m_{\mathrm{T2}} = m_T^{(2)} \approx \sqrt{\frac{|\mathbf{q}_T^{(2)}|}{|\mathbf{v}_T^{(2)}|}}m_\mu$ Subleading m_T is always larger due to the extra factor, gives the same value as $\mu\mu$.
- For μe: $m_{\mathrm{T2}} = m_T^{(1)} \approx m_\mu$ - majority of the time. No extra factor hence the smaller value compared to μe and $\mu\mu$. This is because as $m_\mu \gg m_e$, rarely will $m_T^{(2)} > m_T^{(1)}$, unless $q_T^{(2)}$ is very large. The events where $m_{\mathrm{T2}} = m_T^{(1)}$ can be seen by the tail in m_{T2} for this lepton combination that extends to cover the $e\mu$ and $\mu\mu$ distributions.
- For $\mu\mu$: $m_{\mathrm{T2}} = m_T^{(2)} \approx \sqrt{\frac{|\mathbf{q}_T^{(2)}|}{|\mathbf{v}_T^{(2)}|}}m_\mu$ Subleading m_T is always larger due to the extra factor, gives the same value as $e\mu$.

Finally, these results can be extended to cover the higher-p_{T} lepton scenario. If we repeat the above process, but change the assumptions about the size of $\mathbf{v}_T$, it can be shown that the flavour dependency in m_{T2} is not purely a low-p_{T} effect. Directly taking $\sqrt{\frac{|\mathbf{q}_T|}{|\mathbf{v}_T|}} \to 1$, in other words when the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ and lepton $\mathbf{p}_{\mathrm{T}}$ are of similar magnitude, the value of m_T chosen by m_{T2} becomes:

- For ee: $m_{\mathrm{T2}} = m_T^{(2)} \approx \sqrt{2}m_e$
- For $e\mu$: $m_{\mathrm{T2}} = m_T^{(2)} \approx \sqrt{2}m_\mu$

- For μe: $m_{\mathrm{T2}} = m_T^{(1)} \approx m_\mu$
- For $\mu\mu$: $m_{\mathrm{T2}} = m_T^{(2)} \approx \sqrt{2} m_\mu$.

At last explaining the distribution in Figure B.4, where ee events are distinct from events with at least one μ, at higher-p_{T}. As a result, the BDTs trained using variants of m_{T2} where all the lepton masses are set to 0, to remove this dependency.

There are three kinds of lies: lies, damned lies, and statistics

Mark Twain

C
The Statistical Fits

Contents

C.1 CR-only fit 190
C.2 Exclusion fit 191
C.3 Model-dependent discovery fit 193
C.4 Model-independent discovery fit 194
C.5 Toys 194

The regions defined for the background estimate, are used in a range of different statistical fits to interpret the results. Observed data in the CRs, VRs and SRs are compared with the SM predictions using a profile likelihood method [124] implemented in the HistFitter package [147]. Systematic uncertainties are incorporated as nuisance parameters with Gaussian constraints within the fit. The experimental systematic uncertainties are correlated between signal and background across all regions. Further details are given in Section 5.6 regarding the difference sources of these uncertainties. This section will be a brief introduction into the different statistical fits used in the analysis, and a more in depth look at the mathematics is given in Appendix C.

C.1 CR-only fit

Starting first with the CR-only fit, where the two CRs, CR-Top-BDT, CR-VV-BDT, are used together in a simultaneous fit, for each BDT. No signal is considered in these fits. For each control region, the probability of observing n events given a prediction λ is given by a Poisson distribution:

$$P\left(n|\lambda\right) = \frac{\lambda^n e^{-\lambda}}{n!}. \tag{C.1}$$

There are many systematic uncertainties that are accounted for in the analysis, and these uncertainties are included in the fit as nuisance parameters. Generally these take the form of asymmetric Gaussian constraints centred on one, with a width determined by the relative variations derived for a given systematic variation. These parameters are able to vary (be pulled) in the fit but any large deviations with respect to the nominal value are penalised.

Using this construction, a likelihood function is defined as the product of the Poisson distributions of the control regions and the nuisance parameters:

$$\mathcal{L}\left(\boldsymbol{n}, \boldsymbol{\theta}^0 \mid \mu_{\mathrm{NF}}, \boldsymbol{b}, \boldsymbol{\theta}\right) = P_{\mathrm{CR}} \times S_{\mathrm{syst}} = \prod_{i \in \mathrm{CR}} P\left(n_i \mid \lambda_i\left(\mu_{\mathrm{i}}, \boldsymbol{b}, \boldsymbol{\theta}\right)\right) \times S_{\mathrm{syst}}\left(\boldsymbol{\theta}^0, \boldsymbol{\theta}\right). \tag{C.2}$$

The Poisson expectation values λ_{SR} and λ_i depend upon the prediction of the background $\boldsymbol{b}$, the nuisance parameters $\boldsymbol{\theta}$, and the normalisation factors for background processes μ_{NF}. S_{syst} is a product of each nuisance parameter, where $\boldsymbol{\theta}_0$ is the mean value around which a given $\boldsymbol{\theta}$ can be varied. The likelihood is then maximised under the simultaneous variation of normalisation factors associated with each control region, to fit the MC background to the observed data. These normalisation factors are then used to scale the MC predictions other regions. In Section 5.7, the results of the CR-only fit are used to present the post-fit VRs, allowing an evaluation of the extrapolation of these extracted normalisation factors. The same fit is also used to obtain the post-fit SR distributions, providing the final data/MC comparison in the signal regions.

C.2 Exclusion fit

To perform the exclusion and model-dependent discovery significance fits, it is necessary to extend the background-only likelihood formalism to include both the SRs and a specific signal model. The likelihood defined in Equation C.2 is therefore modified as follows:

$$\begin{aligned}\mathcal{L}(\mathbf{n}, \boldsymbol{\theta}^0 \mid \mu_{\text{sig}}, \mu_{\text{SF}}, \mathbf{b}, \boldsymbol{\theta}) = & \prod_{j \in \text{SR}} P\left(n_{\text{SR}_j} \mid \mu_{\text{sig}} \cdot s_{\text{SR}_j}(\boldsymbol{\theta}) + \mu_i \cdot b_{\text{SR}_j}(\boldsymbol{\theta})\right) \\ & \times \prod_{i \in \text{CR}} P\left(n_i \mid \mu_{\text{sig}} \cdot s_i(\boldsymbol{\theta}) + \mu_i \cdot b_i(\boldsymbol{\theta})\right) \\ & \times S_{\text{syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta}),\end{aligned} \tag{C.3}$$

where the product over the SRs is combined with the CR terms to allow for potential signal contamination in control regions and to preserve correlations of systematic uncertainties across all regions. The background expectation parameter ($\mathbf{b}$) is modified to split the signal (s_i) and background contributions (b_i) in the fit regions. A signal strength parameter, μ_{sig}, is introduced to scale the predicted yield from the hypothesis signal model, and is varied in the fit along with the background normalisation factors. In this analysis, these fits are carried out with either three SRs – electron, muon SRs individually, or six SRs – electron and muon SRs simultaneously.

The CL_s method builds upon the likelihood definition above to quantify the degree to which a given signal hypothesis can be excluded by the data. The CL_s value is defined as:

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{1 - \text{CL}_b} = \frac{p_{s+b}}{1 - p_b} = \frac{\int_{q_{\text{obs}}}^{+\infty} f\left(q \mid H_{s+b}\right) \, \mathrm{d}q}{1 - \int_0^{q_{\text{obs}}} f\left(q \mid H_b\right) \, \mathrm{d}q}, \tag{C.4}$$

where p_{s+b} is the p-value associated with the signal plus background hypothesis (H_{s+b}), and p_b is the p-value under the background-only hypothesis (H_b). The variable q represents the test statistic constructed from the profile log-likelihood ratio, and q_{obs} denotes its observed value in data. The probability density functions $f(q \mid H_{s+b})$ and $f(q \mid H_b)$ describe the distributions of the test statistic under the signal-plus-background and background-only hypotheses, respectively. By construction, CL_s avoids the exclusion of signal models based solely on statistical fluctuations of the background, which would otherwise lead to overly aggressive

exclusions. For the exclusion of a signal hypothesis, ATLAS and CMS use a threshold CL_s value of 0.05 – which corresponds to statistical significance of 1.64.

The choice of test statistic is the profile log likelihood ratio, which a measure of how well the data is explained by the best-fit model compared to a model where the signal strength is fixed at some test value. The log likelihood test statistic is defined:

$$q_\mu = -2\ln\left(\frac{\mathcal{L}\left(\hat{\mu}_{\mathrm{sig}}, \overline{\boldsymbol{\theta}}\right)}{\mathcal{L}\left(\mu_{\mathrm{sig}}, \boldsymbol{\theta}\right)}\right), \tag{C.5}$$

where $\mathcal{L}$ is the likelihood defined in Equation C.3. The numerator tests the null hypothesis, $\hat{\mu}_{\mathrm{sig}}$ denotes a chosen fixed value of μ_{sig}, $\overline{\boldsymbol{\theta}}$ are the values of the nuisance parameters which maximise the likelihood for the fixed $\hat{\mu_{\mathrm{sig}}}$. The denominator is the maximum likelihood, where the alternative hypothesis is tested and μ_{sig} and $\boldsymbol{\theta}$ are both allowed to vary.

For the exclusion fit, the null hypothesis is the background plus signal hypothesis and the alternative hypothesis is the background-only hypothesis. As we are only interested in cases where the data disfavours the signal hypothesis, in other words, where the best-fit signal strength is less than the fixed value, the one-sided version of the profile log-likelihood ratio is used. Therefore, the test statistic becomes:

$$q_\mu = \begin{cases} -2\ln\frac{\mathcal{L}(1,\overline{\boldsymbol{\theta}})}{\mathcal{L}(\mu_{\mathrm{sig}},\boldsymbol{\theta})} & \mu_{\mathrm{sig}} \leq 1, \\ 0 & \mu_{\mathrm{sig}} > 1, \end{cases} \tag{C.6}$$

where $\hat{\mu}_{\mathrm{sig}}$ is fixed to one. In cases where the number of predicted events is large, the asymptotic approximation can be used to calculate q_μ by approximating as a χ^2 distribution [124]. Then the full CL_s can be calculated following Eq. C.4.

The observed values are computed by evaluating the test statistic using the real observed data. The expected values are derived using the Asimov dataset, which is a representative dataset constructed such that all statistical fluctuations are set to their expected values. Two Asimov datasets are constructed corresponding to the null and alternative hypotheses. For the null hypothesis, the background plus signal hypothesis, the Asimov dataset is generated using the expected yields from the signal model under a fixed signal strength parameter $\mu = 1$, together with the expected background contributions, and without applying statistical fluctuations.

For the alternative hypothesis, which corresponds to the background-only scenario, the Asimov dataset is constructed using only the expected background yields, again without fluctuations. These Asimov datasets are then used to compute the expected distributions of the test statistic under each hypothesis in the asymptotic approximation In this analysis, this fit configuration is used for the signal model exclusion, shown in Section 5.8.2.

C.3 Model-dependent discovery fit

For the discovery fit, the null hypothesis is the background-only hypothesis and the alternative hypothesis is the background plus signal case. Rather than providing the probability that a signal does not exist, as in the CL_s method, the aim is to quantify the probability of obtaining a test statistic at least as large as the one observed, assuming the background-only hypothesis is true. This probability is given by the denominator in Eq. C.4, $1 - CL_b$, and expresses the degree to which the observed data is incompatible with the background-only hypothesis. As we are only interested in cases where the data favours the presence of a signal, in other words, where the best-fit signal strength is greater than zero, the one-sided version of the profile log-likelihood ratio is used. Therefore, the test statistic for the discovery fit becomes:

$$q_\mu = \begin{cases} -2\ln\frac{\mathcal{L}(0,\overline{\boldsymbol{\theta}})}{\mathcal{L}(\mu_{\text{sig}},\boldsymbol{\theta})} & \mu_{\text{sig}} > 0, \\ 0 & \mu_{\text{sig}} < 0, \end{cases} \tag{C.7}$$

where $\hat{\mu_{\text{sig}}}$ has been fixed to zero. This test statistic is then used to calculate the CL_b p-value. Again, in cases where the number of predicted events is large, the asymptotic approximation can be used to calculate q_μ by approximating as a χ^2 distribution [124].

The obtained p-value is then converted into a significance (Z) by using the fundamental relation:

$$Z = \Phi^{-1}(1-p), \tag{C.8}$$

where p is the p-value ($1 - CL_b$) and Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution. This can be used to calculate

the observed and expected discovery significance for a given signal model. In this analysis, this fit configuration is used for the signal model significances in Section 5.8.2.

C.4 Model-independent discovery fit

Finally, the model-independent discovery fit also employs the one-sided profile log-likelihood ratio, but with a few key differences compared to the signal model discovery fit. In this case, while the form of the test statistic remains the same, no specific signal model is assumed in its construction. Instead, the likelihood function is simplified to include only a single SR, avoiding assumptions about the signal shape or distribution across multiple regions. In addition, the signal prediction is set to a 'dummy' value of one, and free to be scaled by μ_{sig}. The resulting value, after fitting the profile likelihood function, is then converted to a significance using Eq. C.8. This gives a model independent significance for the deviation between data and predicted background, individually for each SR. In this analysis, this fit configuration is used for the individual SR significances shown in Section 5.7.4.

C.5 Toys

In some cases, the SR3's defined in the analysis have a low number of predicted MC events, with $N_{\text{exp}} < 3$. In this regime, the asymptotic approximation may no longer be valid, and the fit results can become unreliable. Instead, the test statistic must be constructed using ensemble tests, by performing toy MC simulations to sample the full likelihood distribution. Many pseudo-datasets ('toys') are constructed by randomly sampling from the probability distribution defined by the likelihood, for each SR. This is done by sampling from a Poisson distribution, parameterised by the number of expected events in a region. Then the p-value for a given hypothesis is given by comparing the value of the toy generated test statistics to the observed one:

$$p = \frac{N(q_\mu^{\text{toy}} \geq q_\mu^{\text{obs}})}{N_{\text{total}}}. \tag{C.9}$$

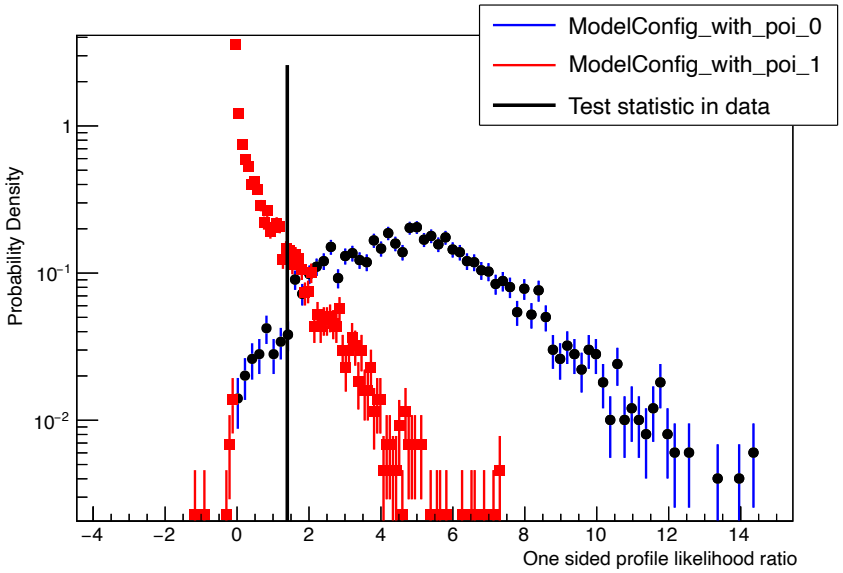


Figure C.1: An example profile likelihood distribution for the profile likelihood ratio test statistic, (red) assuming the null hypothesis (signal plus background: $\mu_{\text{sig}} = 1$) and (blue) using the alternative hypothesis (background only: $\mu_{\text{sig}} = 0$). The distributions are generated using 1000 toys for the null hypothesis and 500 for the alternative, and are normalised to the respective number of toys used in each case to give the probability density. The test statistic in data (black line) is also shown.

For the exclusion fit, this is carried out twice, for the background only and signal plus background hypothesis. Figure C.1, shows an example distribution using $1,000$ toys, in the null and alternative hypothesis for an example exclusion fit with $\mu_{\text{sig}} = 1$. [1] The distribution shows the value of the test statistic in data, is larger than most of the background plus signal tests, but most of the background only test lie above it. This indicates that, and indeed the calculated CL_s for this point is 0.2. The CL_s can be calculated by integrating from the value of the test statistic in data to the right for both cases, using Eq. C.9 to calculate the CL_b and CL_{s+b}, and then the CL_s. The $\hat{\mu}_{\text{sig}} = 1$ distribution goes negative, despite the use of the one-sided log-likelihood ratio as the test statistic. In HistFitter (which uses RooStats [193]), the normalisation factors are typically constrained to be non-negative ($\mu_{\text{sig}} \geq 0$). However, this constraint applies to the fit itself, not to the resulting test statistic. When generating toys, particularly from the signal plus background hypothesis

[1] The value $\mu_{\text{sig}} = 1$ was chosen as an example but in principle this can be done for any value of μ_{sig}.

(where: $\hat{\mu}_{\text{sig}} = 1$), statistical fluctuations can occur, especially in the tails of the distribution. In some cases, the best-fit signal strength (μ_{sig}) from a fit using toys can lead to a lower likelihood than the fixed signal hypothesis ($\hat{\mu}_{\text{sig}}$), such that the test statistic likelihoods follow the relation:

$$\mathcal{L}(\hat{\mu}_{\text{sig}} = 1, \overline{\boldsymbol{\theta}}) > \mathcal{L}(\mu_{\text{sig}}, \boldsymbol{\theta}), \tag{C.10}$$

where the likelihood evaluated with $\hat{\mu}_{\text{sig}} = 1$, is larger than the best fit case where μ_{sig} is free floating. These negative values do not violate the definition of q_μ, and arise from fluctuations in the likelihood when profiling. Values can be clipped to zero in order to enforce the one-sided nature of the test statistic, but this is not done in RooStats/Histfitter. Regardless, the overall effect on the final result is negligible.

In this analysis, toys are utilised in the model independent discovery fits and the signal model discovery fits. The number of toys needs to be optimised to ensure sufficient statistics to accurately sample the full likelihood distribution. These tests are detailed in Appendix D, and led to the selection of 15,000 toys as the optimal number to guarantee convergence in all fits. Similar tests were carried for the exclusion fits, but only a small difference in CL_s was observed - also recorded in Appendix D, and so the asymptotic fit results are used for these.

D

Toys and Systematic Symmetrisation

Contents

D.1 Unstable toy fit results 197
D.2 Number of toys 200

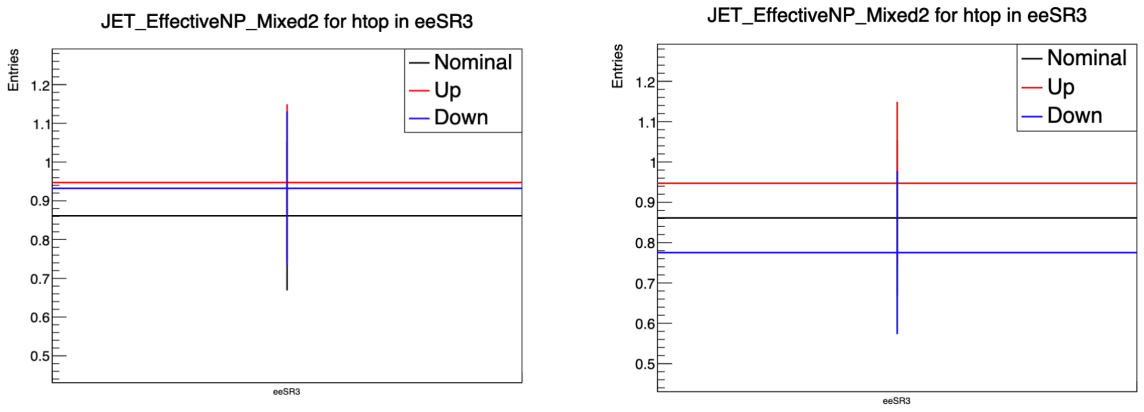
This appendix covers the use of toys in both the exclusion and discovery fits. The appendix first details instability in the fits when generating toys, and the solution. This investigation led to the decision to use toys for the significance fits (discovery and signal model), and asymptotics for the multi-bin exclusion fits. In addition, the results detailed here determine the number of toys required for the fits to converge.

D.1 Unstable toy fit results

Upon running exclusion fits with toys, a large drop in sensitivity was seen, and neighbouring signal points varied by large amounts in observed CL_s value. Additionally, a large difference in the discovery significance was seen when comparing between asymptotic fit results and the toy fit results. The cause of this behaviour was discovered to be unstable toy generation, as a result of systematic variations where the up and down variations were both in the same direction, due to statistical fluctuations. The solution was to symmetrise the problematic systematic

uncertainties. Two methods were trialled: method one takes the larger of the variations and symmetrises using this value, and method two calculates the RMS of the variations and symmetrises using this value.

Figure D.1 shows an example for a jet energy scale systematic and the top background in SR3-BDT$^{ee}_{30}$. This symmetrisation leads to a small loss in sensitivity across the signal grid, as shown in Figure D.2. Figure D.2 shows a comparison between the limits obtained using the unsymmetrised systematics and method one, using the asymptotic approach in both cases. This shows a small drop in sensitivity after symmetrising the systematics in the asymptotic approach, due to the slightly larger overall SR uncertainty. Additionally, a comparison of the limits found using toys and asymptotic approaches using method one, are shown. This shows only a small difference between the toys and asymptotic approaches now that the systematics are symmetrised, indicating that the problem with toy generation is fixed. Finally, a comparison between the exclusion limits from method one (labelled full sym) and method two (labelled half sym) also shown. These results motivated the choice of the asymptotic approach for the exclusion fits, using method two to symmetrise the problematic uncertainties, as the difference in exclusion is small.



(a) **(b)**

Figure D.1: An example jet energy scale systematic variation and the nominal yield for the top background only, in SR3-BDT$^{ee}_{30}$. In (a) the unsymmetrised systematic is shown, and in (b) the systematic variation after symmetrising is shown.

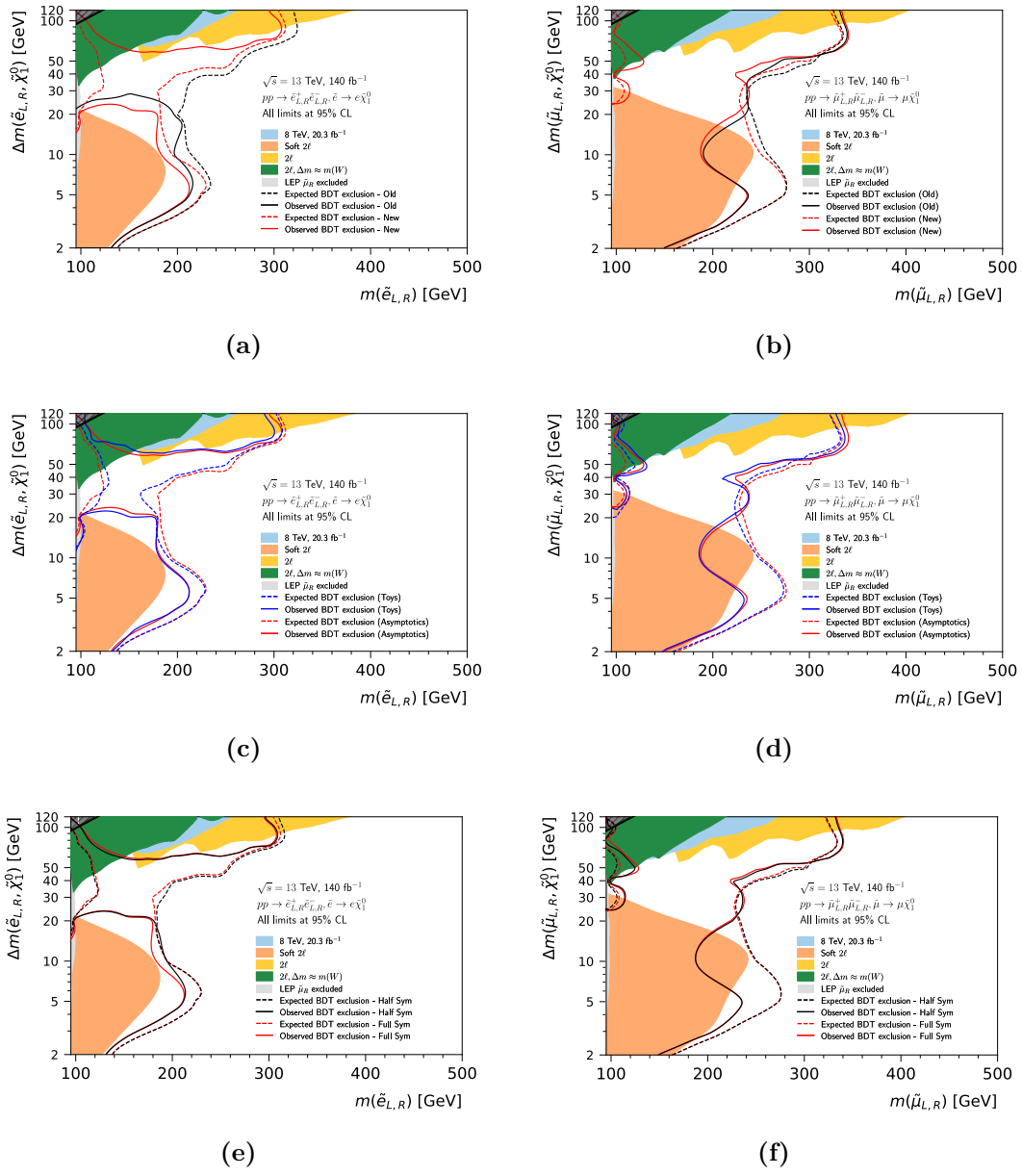
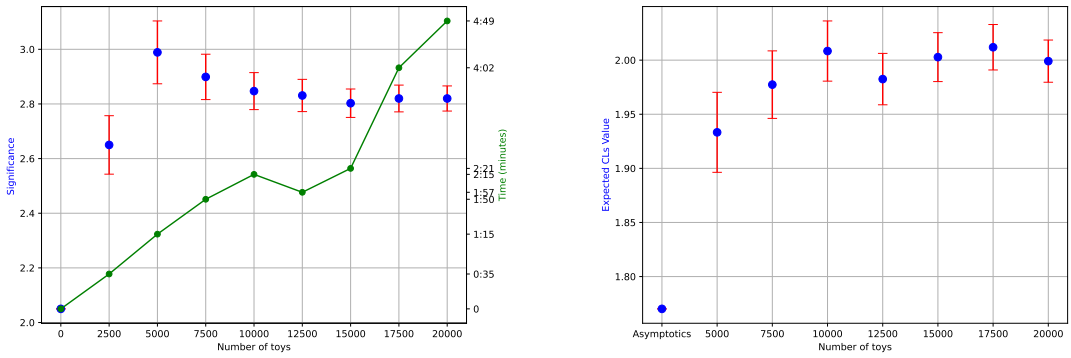


Figure D.2: Exclusion limit comparisons for the (left) selectron and (selection) smuon interpretations. In (a) and (b) results are shown before (labeled old) and after (labeled new) systematic symmetrisation, using method one and asymptotics. In (c) and (d) results are compared for asymptotics and toys, using the systematic symmetrisation following method one, using 15,000 toys. In (e) and (f) results are shown using the systematic symmetrisation following method one (labelled full sym) and method two (labelled half sym), using asymptotics.

D.2 Number of toys

In order to ensure that the number of toys chosen led to convergence in the statistical tests, an optimisation was carried out using a discovery fit in one of the BDT SRs. Figure D.3 shows the results of this optimisation, and it can be seen that the significance stablises at 15,000 toys. This number was chosen for the running of all BDT fits with toys. In addition, tests were performed on each BDT using various signals and exclusion fit configurations, and it was found that 15,000 events provided good convergence for these too.



(a) **(b)**

Figure D.3: Significance as a function of the number of toys for SR3-BDT$^{ee}_{40+50}$. Left shows the significance using the unsymmetrised systematics and right shows the significance after symmetrising all problematic systematics using method two.

E

Top Generator Systematics

This appendix details investigation into the variation direction flip between the SR and CR-Top, seen when evaluating the p_{T}-Hard top generator systematic. Only the results for BDT_{40+50} are shown here, but the same behaviour is seen for the other BDTs, aside from BDT_{5+10}. The p_{T}-Hard variation across the full distributions of CR-Top-BDT and the SFOS lepton preselection, are shown in Figure E.1. At high scores in the SFOS lepton selection, it can be seen that the variation flips direction with respect to CR-Top-BDT.

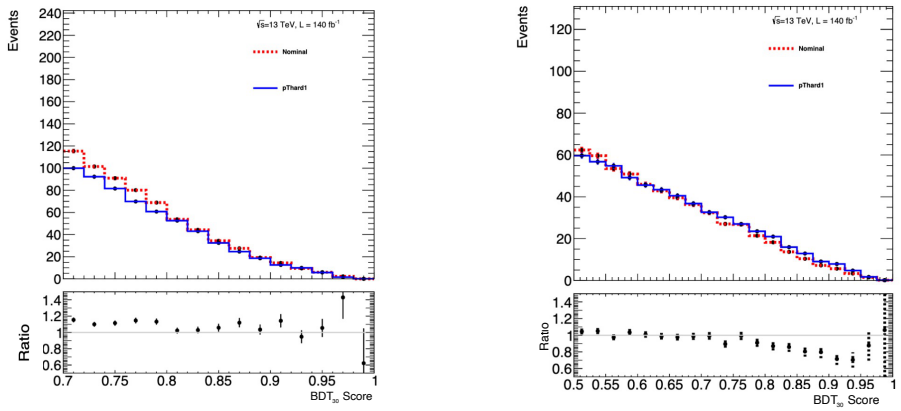


Figure E.1: Pre-fit BDT score distributions for the top background with the p_{T}-Hard variations for CRTop (left) and SFOS lepton selection (right). In the SFOS selection, CR-VV-BDT runs from 0.7–0.9, and the SR from 0.9. The subplot shows the ratio of nominal/variation.

By removing any b-jet selection and implementing a BDT score > 0.8 requirement, the variation of p_{T}-Hard systematic the across b-jet multiplicity bins can be checked, in Figure E.2. The b-jet $= 0$ bin corresponds to the SFOS lepton selection with BDT score > 0.8 region, a proxy for the SRs, and b-jet > 0 bins correspond to CR-Top with BDT score > 0.8. The score cuts were changed to increase statistics in the SFOS lepton selection. This shows a clear flip in direction of the variation between the b-jet $= 0$ bin and b-jet > 0 bins, indicating that the variation differs depending on if b-jets are required or not.

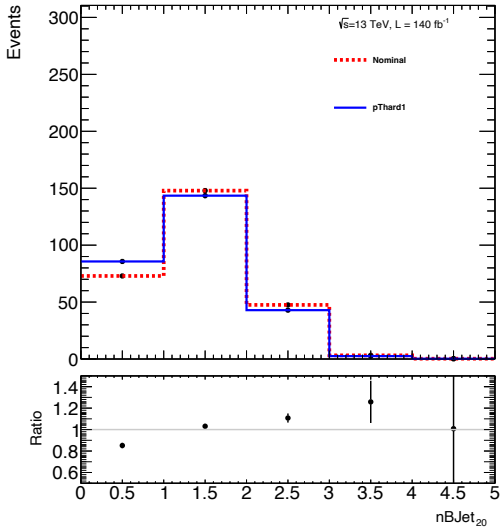


Figure E.2: The $N^{20}_{b-\mathrm{jet}}$ distribution for the top background with the p_{T}-Hard variation and the SFOS lepton selection without a b-jet cut and with a BDT score > 0.8 requirement. The subplot shows the ratio of nominal/variation yields.

The systematic is calculated using only the top background processes, which is majority $t\bar{t}$, and so there will always be two (or more) b-jets produced in the event generation. This means that the events in the b-jet $= 0$ bin arise due to missed b-tags, and the switch from nominal to variation p_{T}-Hard must increase the mis-tag probability of the b-jet. The p_{T}-Hard setting determines the energy scale at which the MC generation switches from using the matrix element (generated by Powheg Box) to parton shower (using Pythia), where the scale depends on the generator p_{T} – which is not the same as detector/reconstructed p_{T}. The nominal p_{T}-Hard setting (0), takes the matching scale to be the p_{T} of the parton as generated

by POWHEG BOX, providing an upper limit to the scale. The variation p_{T}-Hard setting (1), takes the matching scale to be the minimum p_{T} that POWHEG BOX could possibly have generated the parton. As a result, the variation p_{T}-Hard increases the phase space across which the matrix element is used, and so the matrix element is used down to a lower p_{T}. This means the probability to generate a soft b-quark is increased, soft b quarks are more likely to be mis-tagged, and therefore events migrating from b-jet > 0 bins to b-jet $= 0$ as a result. Figure E.3 shows a graphical representation of this explanation. An interesting question to investigate in the future may be why these events are particularly signal-like in BDT score.

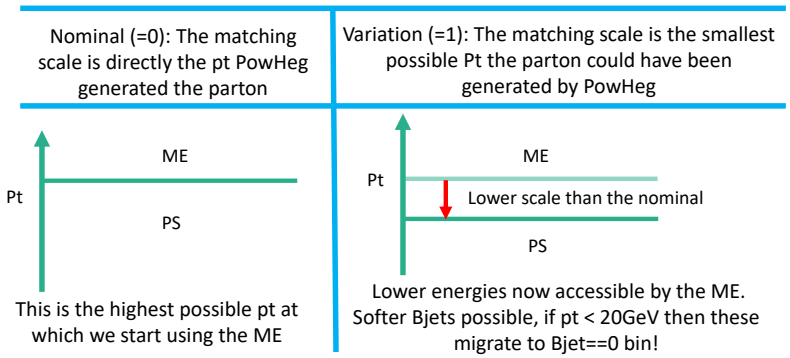


Figure E.3: A sketch explaining the difference in the two p_{T}-Hard schemes, and how this difference causes the observed effect in the SRs.

F

Additional Information for the BDT Results

Contents

F.1 Breakdown of diboson background contributions in analysis regions 205
F.2 Signal Region Post-fit Yield Tables 207
F.3 Additional Excess Information 208
F.4 Post-fit NP pulls 212
F.5 CL_s Values For Each Interpretation 217

This appendix contains a selection of additional material related to the analysis described in Chapter 5.

F.1 Breakdown of diboson background contributions in analysis regions

The BDT with the largest proportion of WZ/ZZ events in the SR is the Δm_{60+75}. This is expected as the signal mass splitting is approaching the Z mass, and hence the signal models become increasingly similar to the background and less like the WW background. Figure F.1 shows a breakdown of the diboson background in the SRs (all 6 combined), the CR-VV, VR-DF and VR-ZZ-BDT for the Δm_{60+75} BDT.

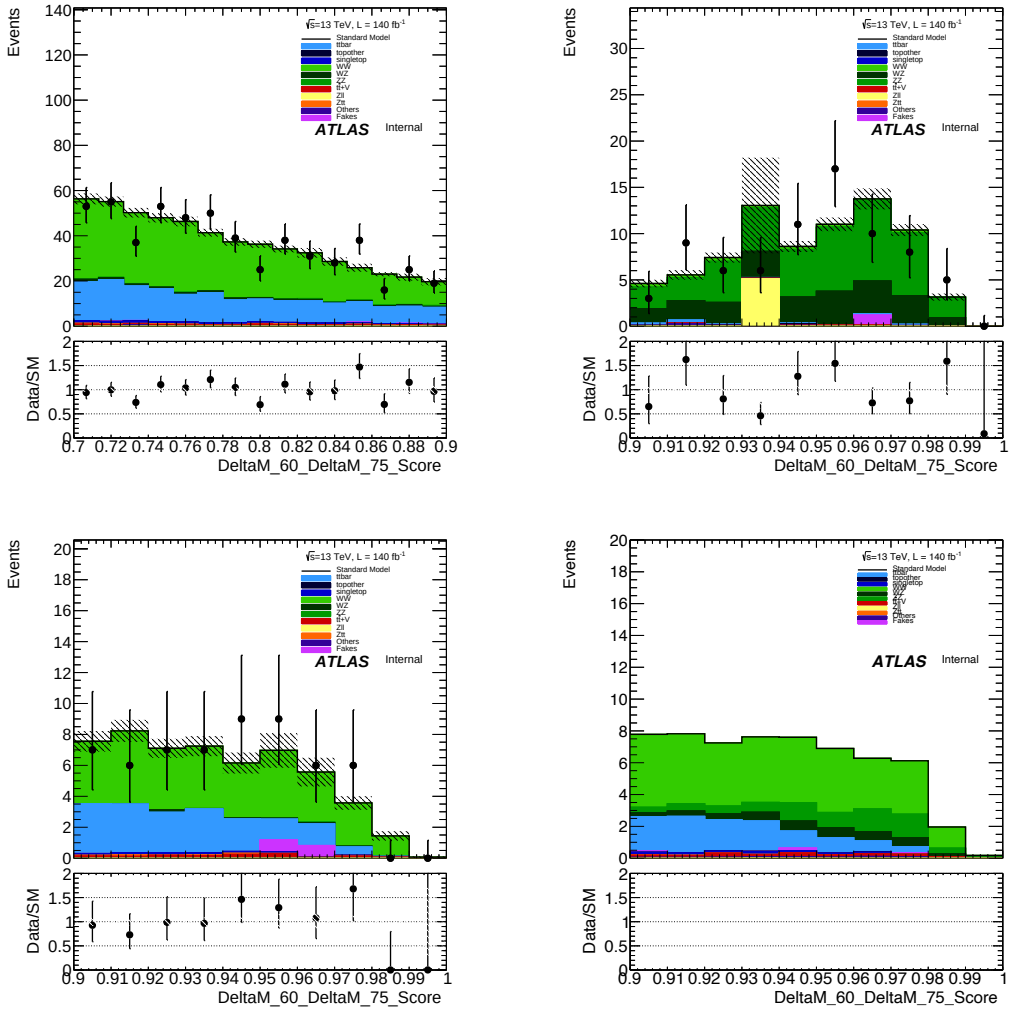


Figure F.1: A post-fit breakdown of the diboson background into WW, ZZ and WZ processes for the CR-VV (top left), VR-ZZ-BDT (top right), VR-DF (bottom left) and combined SRs (bottom right), for the Δm_{60+75} BDT. Note the background grouping is different to the majority of the figures in this thesis as these were made with an older version of the files. Additionally, the large single $Z+$ jets event is later vetoed in the final analysis, and the signal regions are blinded in this figure.

F.2 Signal Region Post-fit Yield Tables

SR	SR1-BDT$^{ee}_{5+10}$	SR2-BDT$^{ee}_{5+10}$	SR3-BDT$^{ee}_{5+10}$	SR1-BDT$^{\mu\mu}_{5+10}$	SR2-BDT$^{\mu\mu}_{5+10}$	SR3-BDT$^{\mu\mu}_{5+10}$
Observed	10	11	5	23	26	4
Fitted SM events	12.3 ± 3.1	10.6 ± 2.5	3.5 ± 1.0	13.7 ± 2.6	13.7 ± 2.6	3.9 ± 1.1
Top	3.0 ± 0.9	2.1 ± 0.5	0.62 ± 0.19	3.7 ± 0.8	3.3 ± 0.7	1.10 ± 0.27
Diboson	4.3 ± 1.9	4.3 ± 1.7	1.6 ± 0.7	4.5 ± 1.9	5.8 ± 2.3	2.5 ± 1.0
Fake/Non-Prompt	4.5 ± 2.3	3.7 ± 1.9	1.2 ± 0.8	4.9 ± 1.6	3.5 ± 1.3	$0.3^{+0.4}_{-0.3}$
Other	0.6 ± 0.4	0.50 ± 0.35	0.09 ± 0.05	0.6 ± 0.5	1.1 ± 0.6	0.05 ± 0.04

Table F.1: Summary of observed and predicted yields in SR-BDT$_{5+10}$ using the CR-only fit. The category 'Other' contains rare backgrounds from processes such as $t\bar{t}V$, multi-top and triboson production. Uncertainties in the fitted background estimates combine statistical and systematic uncertainties.

SR	SR1-BDT$^{ee}_{20}$	SR2-BDT$^{ee}_{20}$	SR3-BDT$^{ee}_{20}$	SR1-BDT$^{\mu\mu}_{20}$	SR2-BDT$^{\mu\mu}_{20}$	SR3-BDT$^{\mu\mu}_{20}$
Observed	16	11	2	14	10	4
Fitted SM events	14.9 ± 2.9	9.3 ± 1.7	2.9 ± 0.8	21.5 ± 3.5	10.6 ± 2.2	2.8 ± 0.6
Top	4.2 ± 1.7	2.1 ± 0.9	0.43 ± 0.22	5.9 ± 2.2	2.4 ± 1.1	0.68 ± 0.31
Diboson	7.8 ± 2.0	6.0 ± 1.6	1.6 ± 0.5	12.4 ± 3.1	6.3 ± 1.7	2.1 ± 0.6
Fake/Non-Prompt	2.3 ± 1.4	1.0 ± 0.8	0.8 ± 0.6	2.9 ± 1.1	1.7 ± 0.8	$0.01^{+0.15}_{-0.01}$
Other	0.5 ± 0.4	0.21 ± 0.15	0.08 ± 0.07	0.31 ± 0.20	0.17 ± 0.16	$0.02^{+0.09}_{-0.02}$

Table F.2: Summary of observed and predicted yields in SR-BDT$_{20}$ using the CR-only fit. The category 'Other' contains rare backgrounds from processes such as $t\bar{t}V$, multi-top and triboson production. Uncertainties in the fitted background estimates combine statistical and systematic uncertainties.

SR	SR1-BDT$^{ee}_{30}$	SR2-BDT$^{ee}_{30}$	SR3-BDT$^{ee}_{30}$	SR1-BDT$^{\mu\mu}_{30}$	SR2-BDT$^{\mu\mu}_{30}$	SR3-BDT$^{\mu\mu}_{30}$
Observed	11	10	4	20	14	1
Fitted SM events	12.5 ± 2.8	6.6 ± 1.4	3.0 ± 1.1	14.6 ± 3.5	8.2 ± 2.0	3.0 ± 0.6
Top	4.2 ± 1.6	2.3 ± 0.8	0.8 ± 0.7	4.8 ± 2.2	2.4 ± 1.2	0.76 ± 0.31
Diboson	7.8 ± 1.8	3.9 ± 0.7	2.1 ± 0.5	9.3 ± 2.0	5.5 ± 1.2	2.1 ± 0.5
Fake/Non-Prompt	$0.1^{+0.5}_{-0.1}$	$0.2^{+0.5}_{-0.2}$	$0.01^{+0.31}_{-0.01}$	$0.18^{+0.25}_{-0.18}$	$0.01^{+0.31}_{-0.01}$	$0.05^{+0.09}_{-0.05}$
Other	0.38 ± 0.30	$0.2^{+0.5}_{-0.2}$	0.15 ± 0.14	$0.3^{+0.4}_{-0.3}$	$0.22^{+0.30}_{-0.22}$	0.07 ± 0.06

Table F.3: Summary of observed and predicted yields in SR-BDT$_{30}$ using the CR-only fit. The category 'Other' contains rare backgrounds from processes such as $t\bar{t}V$, multi-top and triboson production. Uncertainties in the fitted background estimates combine statistical and systematic uncertainties.

SR	SR1-BDT$^{ee}_{40+50}$	SR2-BDT$^{ee}_{40+50}$	SR3-BDT$^{ee}_{40+50}$	SR1-BDT$^{\mu\mu}_{40+50}$	SR2-BDT$^{\mu\mu}_{40+50}$	SR3-BDT$^{\mu\mu}_{40+50}$
Observed	9	6	7	3	9	4
Fitted SM events	8.2 ± 2.0	6.2 ± 1.4	2.8 ± 0.8	9.1 ± 1.9	6.3 ± 1.2	3.2 ± 0.7
Top	2.7 ± 1.2	2.6 ± 1.2	0.5 ± 0.4	3.7 ± 1.6	1.9 ± 1.0	0.53 ± 0.28
Diboson	3.8 ± 1.2	3.4 ± 0.8	2.1 ± 0.5	5.3 ± 1.1	4.2 ± 0.8	2.6 ± 0.5
Fake/Non-Prompt	$1.2^{+1.3}_{-1.2}$	$0.01^{+0.31}_{-0.01}$	$0.01^{+0.31}_{-0.01}$	$0.00^{+0.15}_{-0.00}$	$0.01^{+0.18}_{-0.01}$	$0.04^{+0.09}_{-0.04}$
Other	0.41 ± 0.25	0.28 ± 0.24	0.15 ± 0.08	$0.16^{+0.22}_{-0.16}$	$0.22^{+0.23}_{-0.22}$	0.07 ± 0.05

Table F.4: Summary of observed and predicted yields in SR-BDT$_{40+50}$ using the CR-only fit. The category 'Other' contains rare backgrounds from processes such as $t\bar{t}V$, multi-top and triboson production. Uncertainties in the fitted background estimates combine statistical and systematic uncertainties.

SR	SR1-BDT$^{ee}_{60+75}$	SR2-BDT$^{ee}_{60+75}$	SR3-BDT$^{ee}_{60+75}$	SR1-BDT$^{\mu\mu}_{60+75}$	SR2-BDT$^{\mu\mu}_{60+75}$	SR3-BDT$^{\mu\mu}_{60+75}$
Observed	15	17	2	19	7	4
Fitted SM events	13.0 ± 2.4	11.9 ± 2.4	2.8 ± 0.7	13.4 ± 3.2	12.6 ± 1.8	3.2 ± 0.6
Top	4.2 ± 2.0	2.3 ± 1.5	0.22 ± 0.13	4.2 ± 2.1	2.4 ± 1.2	$0.18^{+0.19}_{-0.18}$
Diboson	8.3 ± 1.8	8.9 ± 1.6	2.0 ± 0.4	8.8 ± 2.4	9.8 ± 1.5	2.9 ± 0.5
Fake/Non-Prompt	$0.01^{+0.31}_{-0.01}$	$0.01^{+0.31}_{-0.01}$	$0.3^{+0.5}_{-0.3}$	$0.07^{+0.14}_{-0.07}$	$0.02^{+0.11}_{-0.02}$	$0.03^{+0.06}_{-0.03}$
Other	0.48 ± 0.32	0.6 ± 0.4	0.21 ± 0.12	0.35 ± 0.22	0.38 ± 0.22	0.09 ± 0.05

Table F.5: Summary of observed and predicted yields in SR-BDT$_{60+75}$ using the CR-only fit. The category 'Other' contains rare backgrounds from processes such as $t\bar{t}V$, multi-top and triboson production. Uncertainties in the fitted background estimates combine statistical and systematic uncertainties.

F.3 Additional Excess Information

This section contains some additional information relating to the investigation of the observed excesses in SR-BDT$^{\mu\mu}_{5+10}$ SRs. The checks in this appendix form only a part of the checks done on the excess events.

Figure F.2 contains a breakdown of the recording year for data events in SR-BDT$^{\mu\mu}_{5+10}$. These checks were done to ensure that the events do not mostly come from the same year, and have roughly a linear distribution across the years when normalised to luminosity.

Figure F.3 contains (η, ϕ) maps for the leading and subleading leptons and jets in data events for SR-BDT$^{\mu\mu}_{5+10}$. These checks were done to verify that the leptons and jets in the events, do not all group up in the same location of the detector.

Figure F.4 contains angular variables calculated between the leptons and the jets for SR-BDT$^{\mu\mu}_{5+10}$. In this case, the checks were carried out to evaluate any potential correlations between lepton and jet angular position. For example, a

secondary muon from a b-jet would be produced at a small angle to the jet. No notable angular dependence was seen between the leptons and the jets. The most notable angular relationship is observed between the two leptons, where they are seen to be very back-to-back, with $\Delta\phi(\ell_1,\ \ell_2) \approx \pi$.

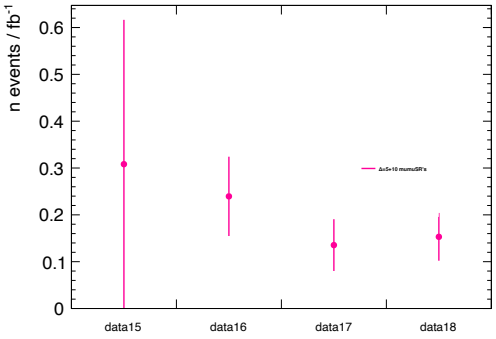


Figure F.2: Data yields per year of Run 2, for events observed in the three SR-BDT$^{\mu\mu}_{5+10}$ regions, normalised to the luminosity recorded in that year.

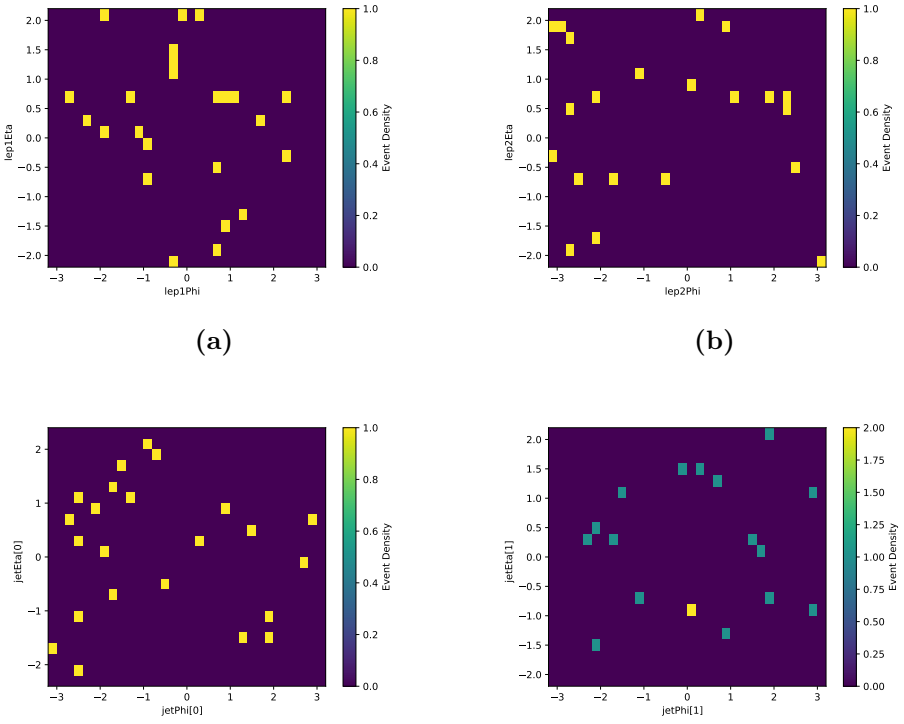


Figure F.3: The (η, ϕ) maps for the (a) leading and (b) subleading leptons, and the (c) leading and (d) subleading jets, for all data events in the three SR-BDT$^{\mu\mu}_{5+10}$ regions. Not all events have a subleading jet.

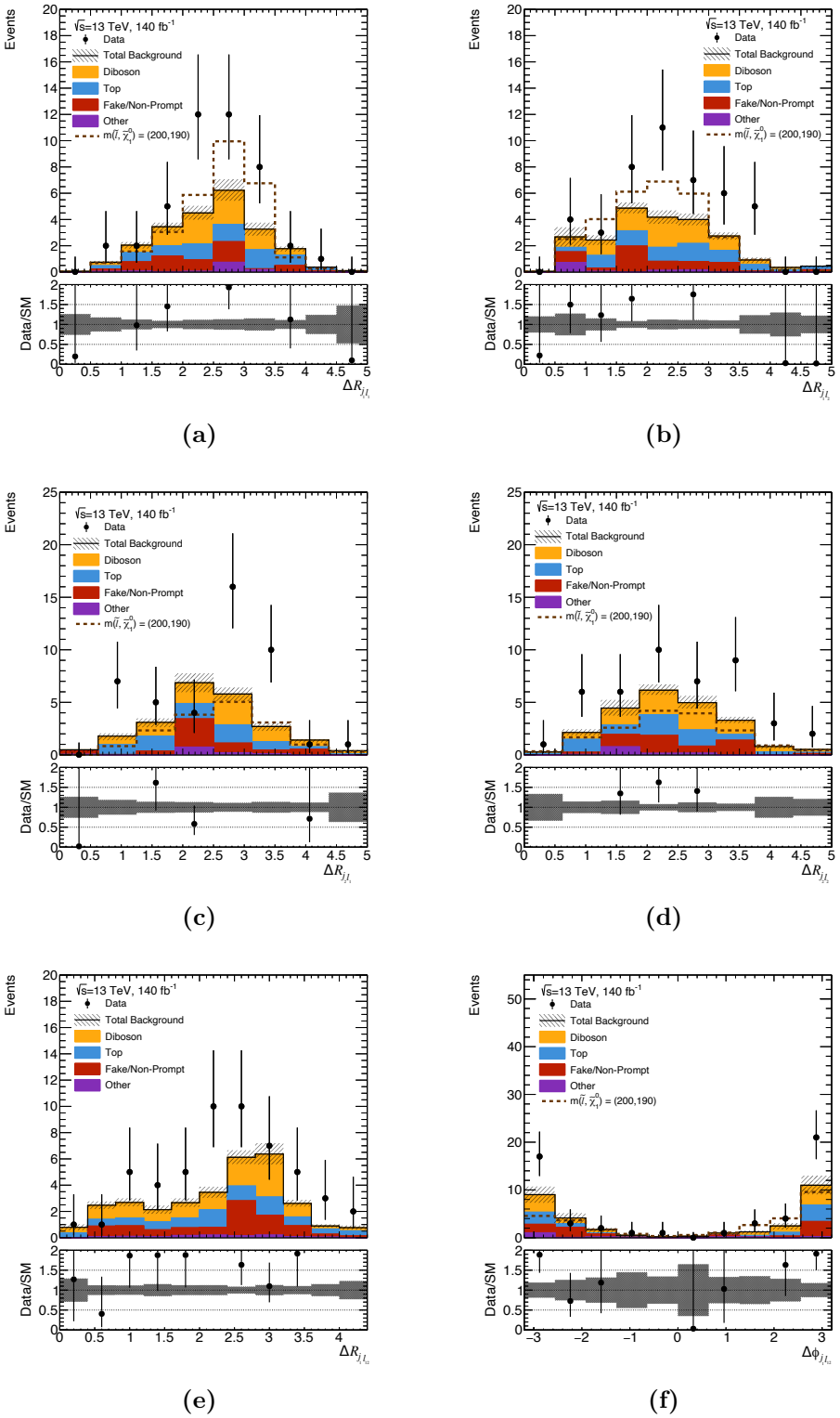


Figure F.4: Data and pre-fit MC background distributions for the combined SR-BDT$^{\mu\mu}_{5+10}$ regions. The angular distance between (a) the leading jet and leading lepton, (b) the leading jet and subleading lepton, (c) the subleading jet and leading lepton, (d) the subleading jet and subleading lepton and (e) the leading jet and di-lepton system. In (f), the difference between the phi angles of the leading and subleading lepton is shown. Only statistical uncertainties are accounted for.

F.4 Post-fit NP pulls

This section contains the post-fit nuisance parameters pulls and normalisation factors for an example exclusion fit, for each BDT. The fits are carried out using all six SRs of a given BDT and a different relevant signal model in each case. In all fits, no significant constraints on uncertainties were seen, and pulls are generally small.

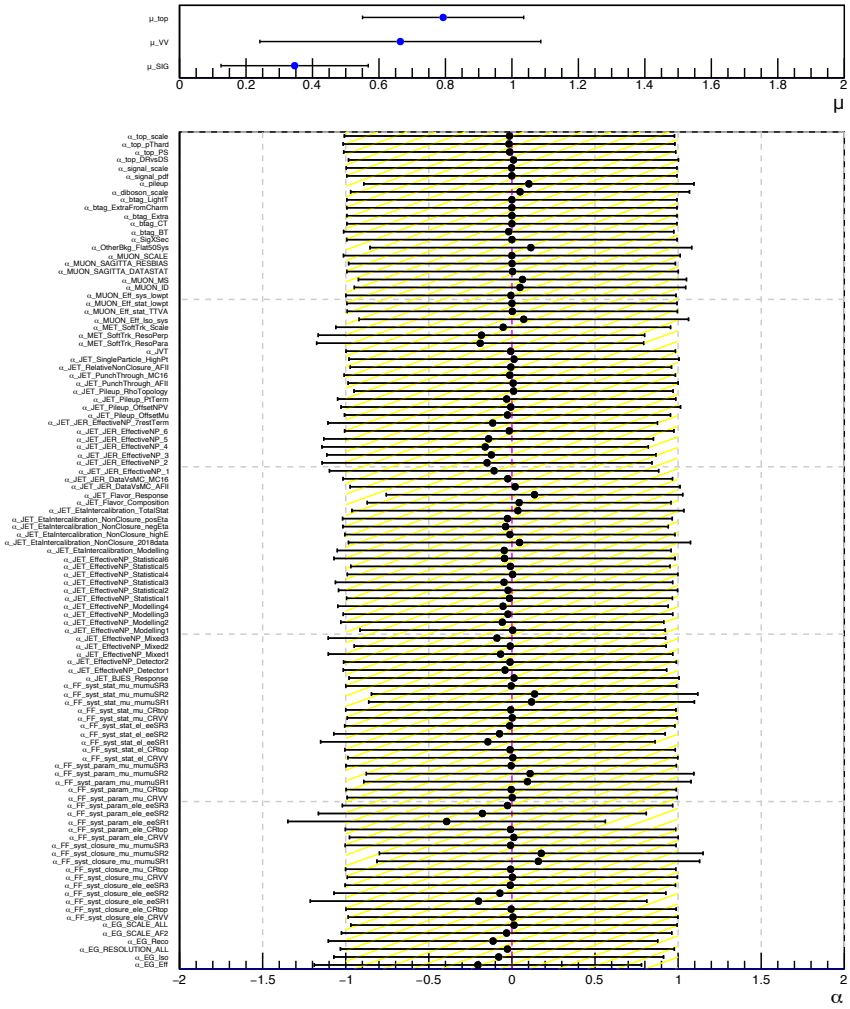


Figure F.5: Normalisation factors and nuisance parameters pulls for an exclusion fit of the signal $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = (150, 140)$ GeV, using all SRs of SR-BDT$_{5+10}$.

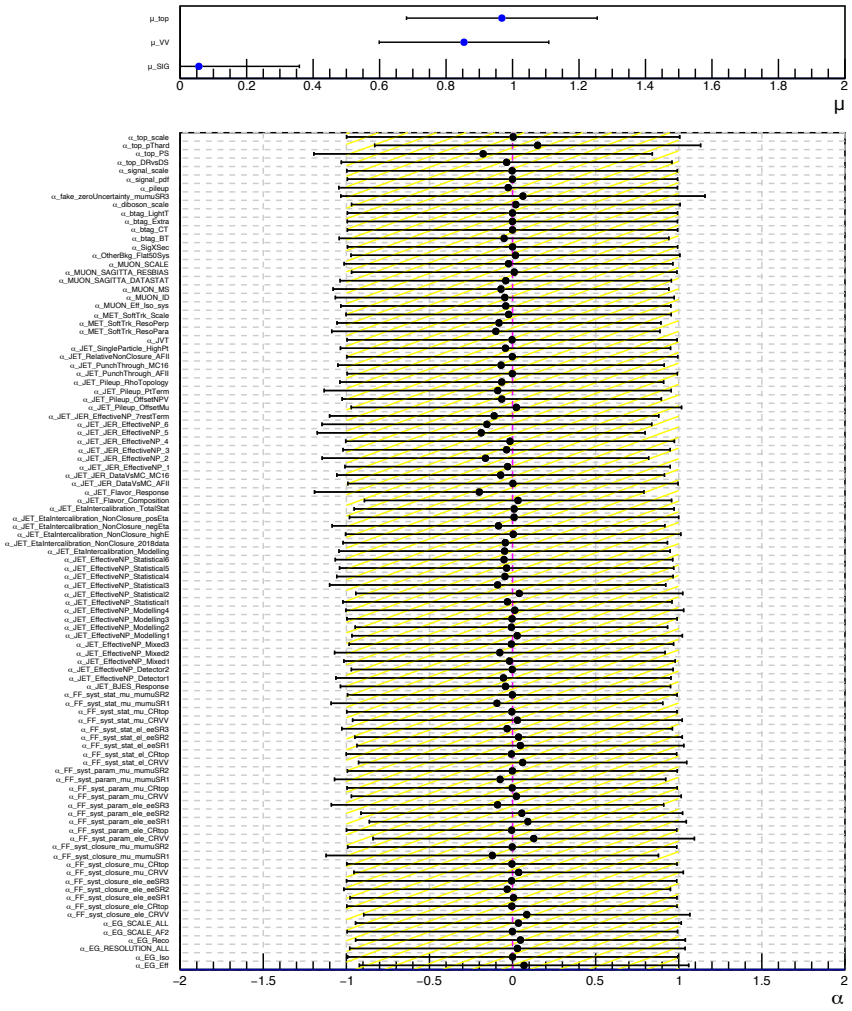


Figure F.6: Normalisation factors and nuisance parameters pulls for an exclusion fit of the signal $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = (200, 180)$ GeV, using all SRs of SR-BDT$_{20}$.

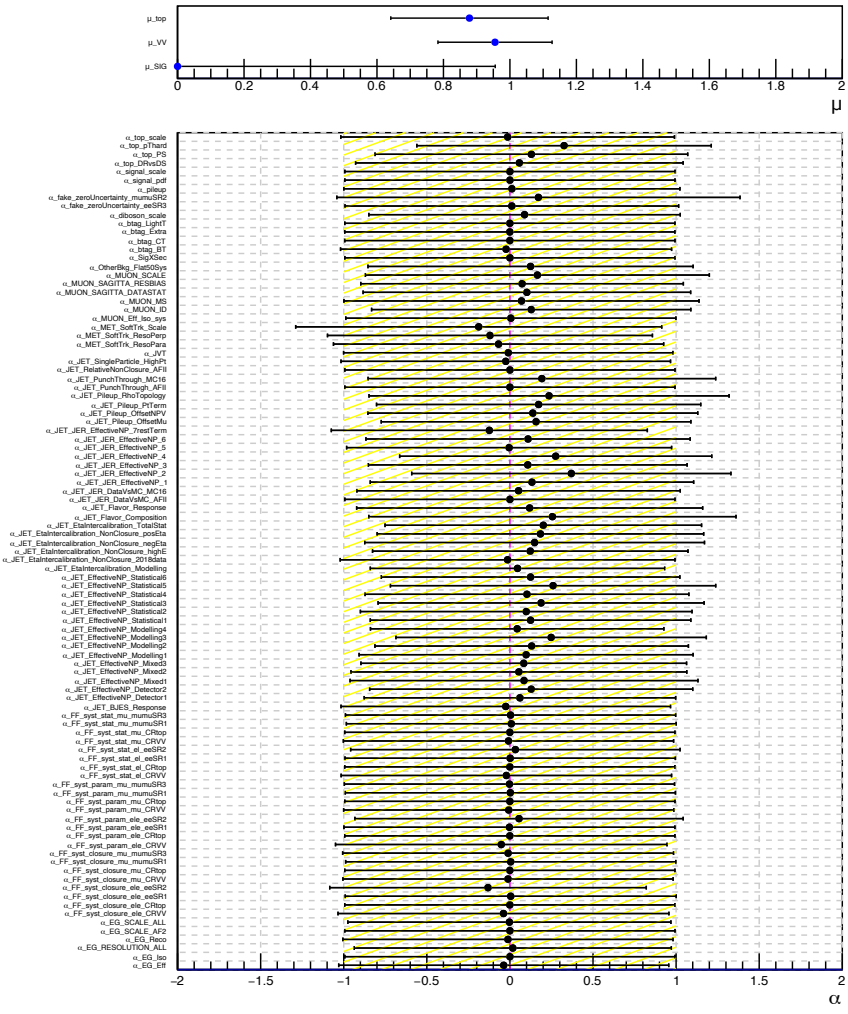


Figure F.7: Normalisation factors and nuisance parameters pulls for an exclusion fit of the signal $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = (200, 170)$ GeV, using all SRs of SR-BDT$_{30}$.

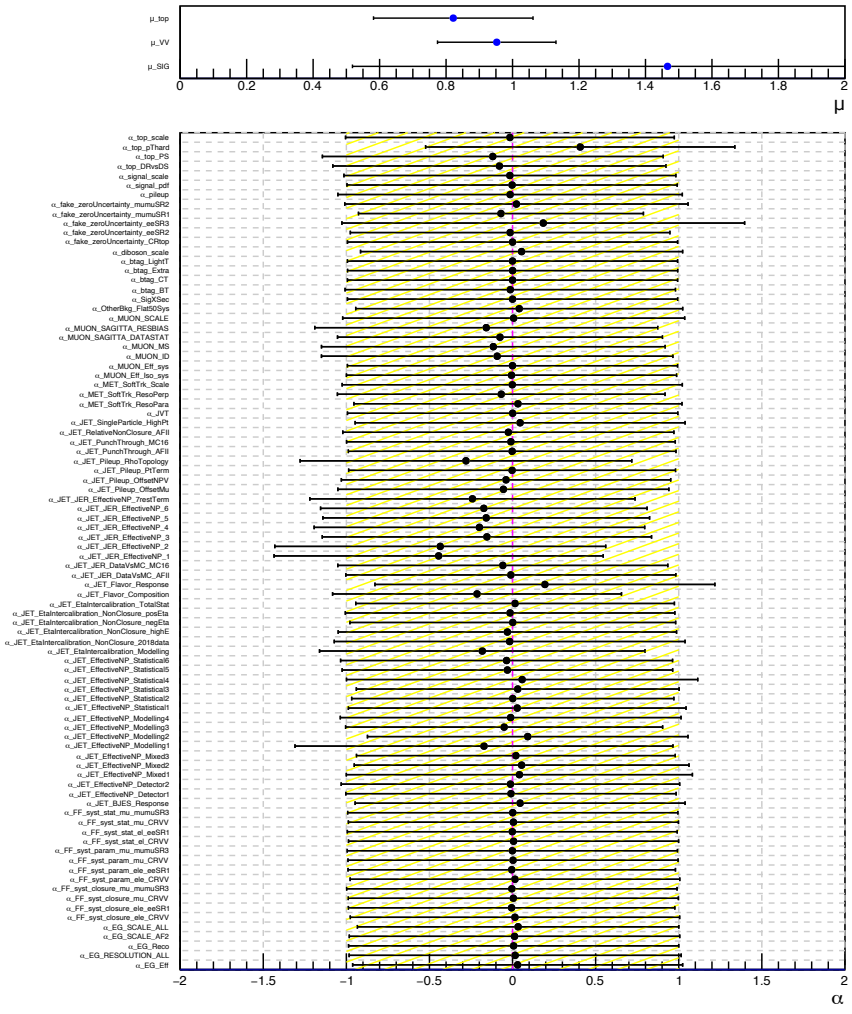


Figure F.8: Normalisation factors and nuisance parameters pulls for an exclusion fit of the signal $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = (375, 335)$ GeV, using all SRs of SR-BDT$_{40+50}$.

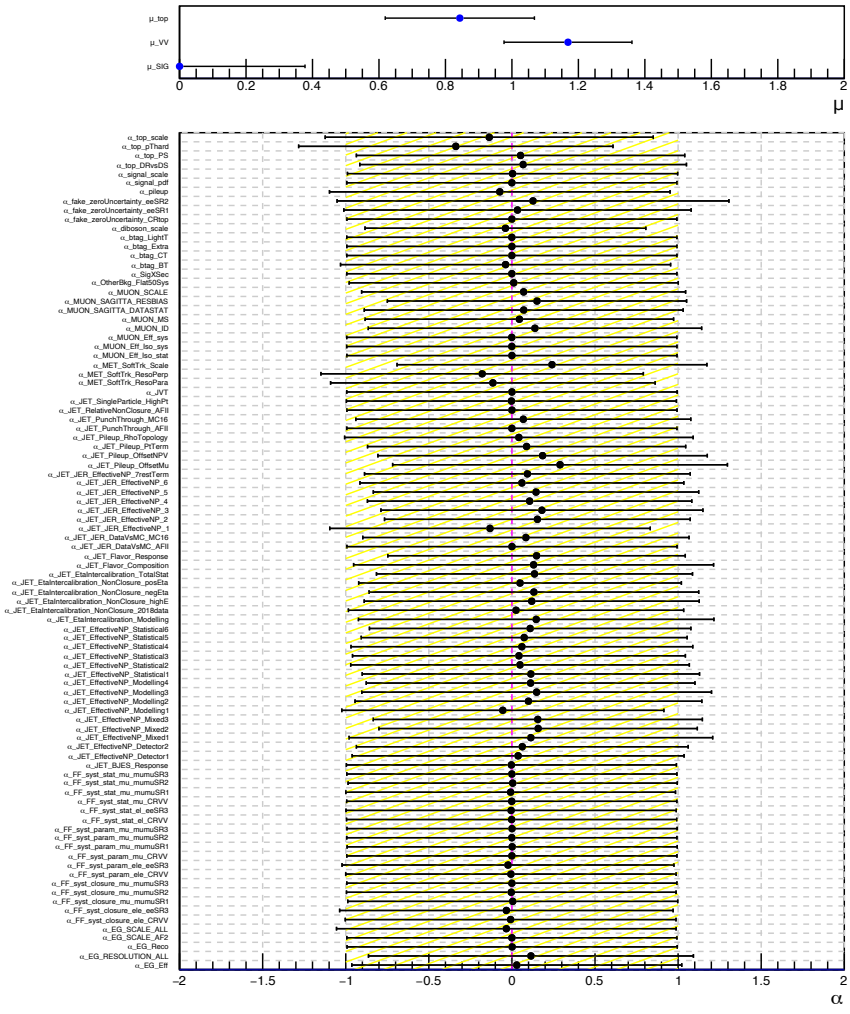
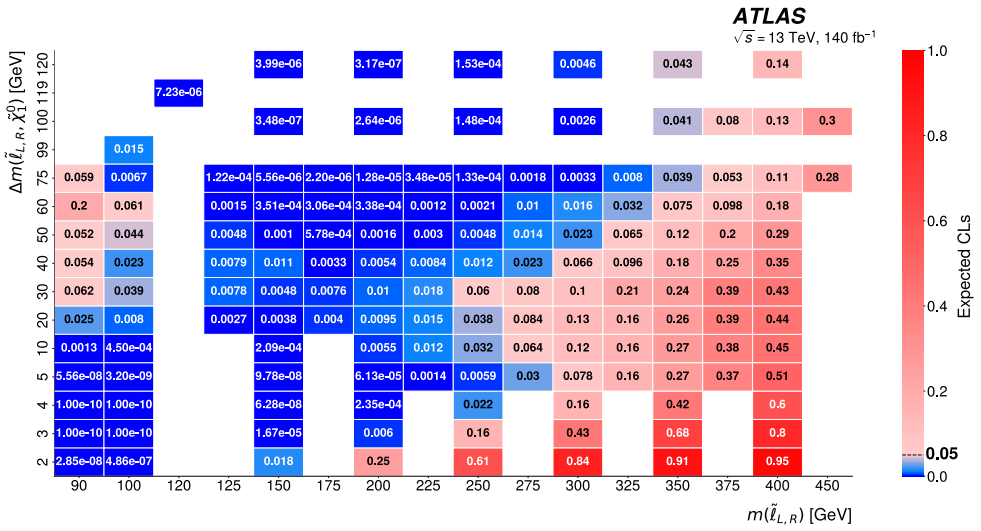
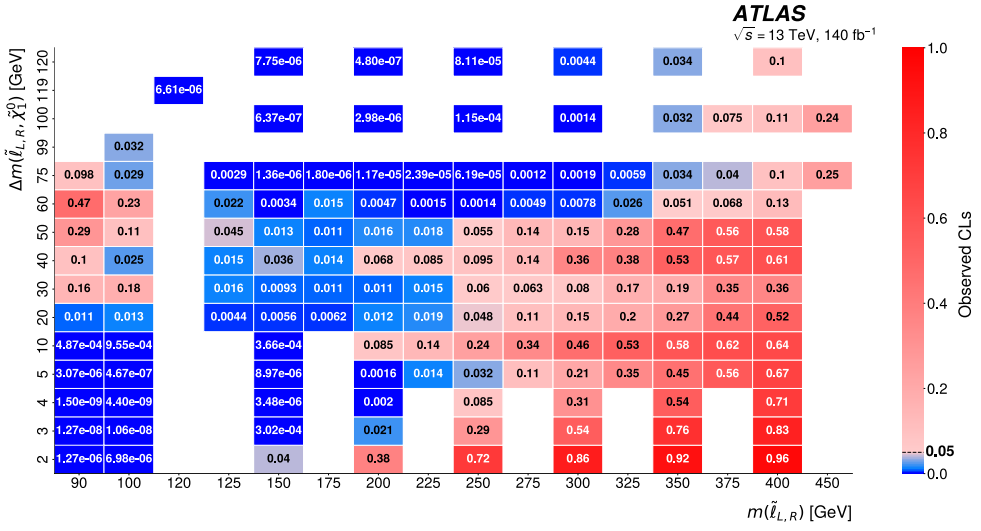


Figure F.9: Normalisation factors and nuisance parameters pulls for an exclusion fit of the signal $\Delta m(\tilde{\ell}, \widetilde{\chi}_1^0) = (300, 225)$ GeV, using all SRs of SR-BDT$_{60+75}$.

F.5 CL_s Values For Each Interpretation

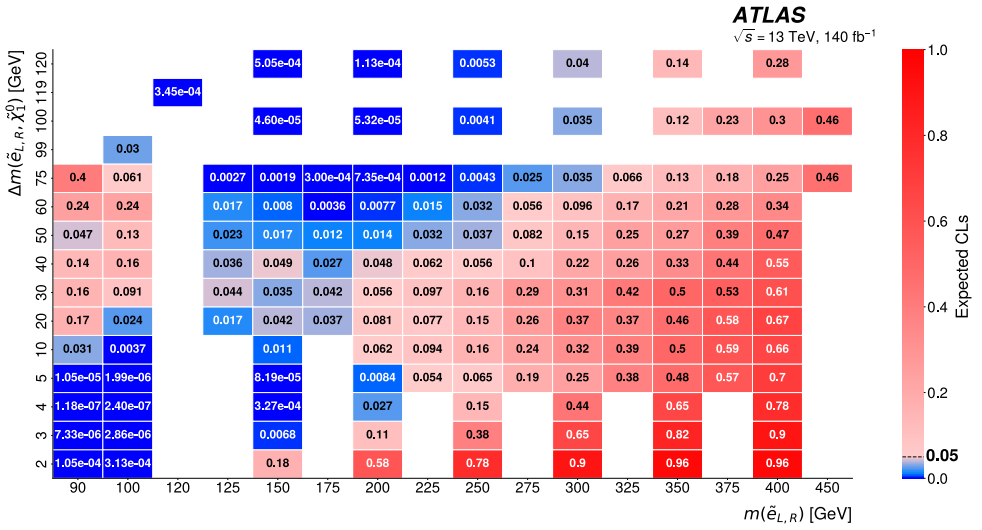


(a)

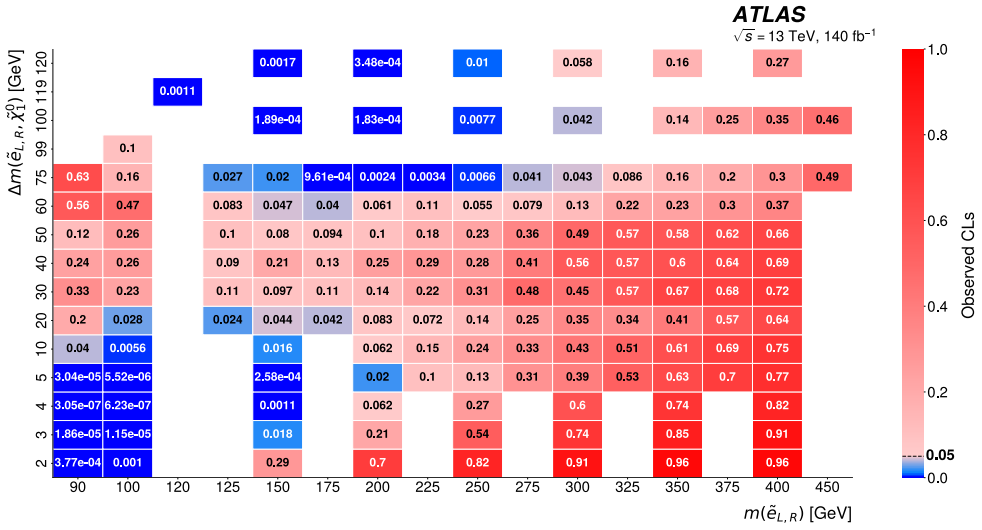


(b)

Figure F.10: Summary of the (a) expected and (b) observed CL_s values across the $m(\tilde{\ell}_{\mathrm{L,R}})$ – $\Delta m(\tilde{\ell}_{\mathrm{L,R}}, \tilde{\chi}_1^0)$ plane in the $m(\tilde{\ell}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}}) = m(\tilde{\mu}_{\mathrm{L}}) = m(\tilde{\mu}_{\mathrm{R}})$ interpretation using the BDT-based selection. The boxes are colour-coded to mark benchmark models for which the CL_s value is found to be below 0.05 (blue) or above (red).

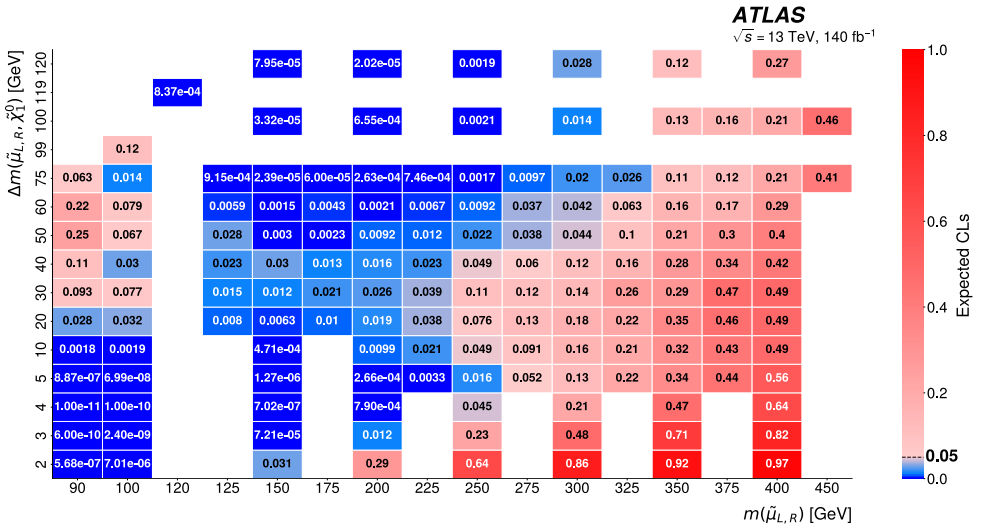


(a)

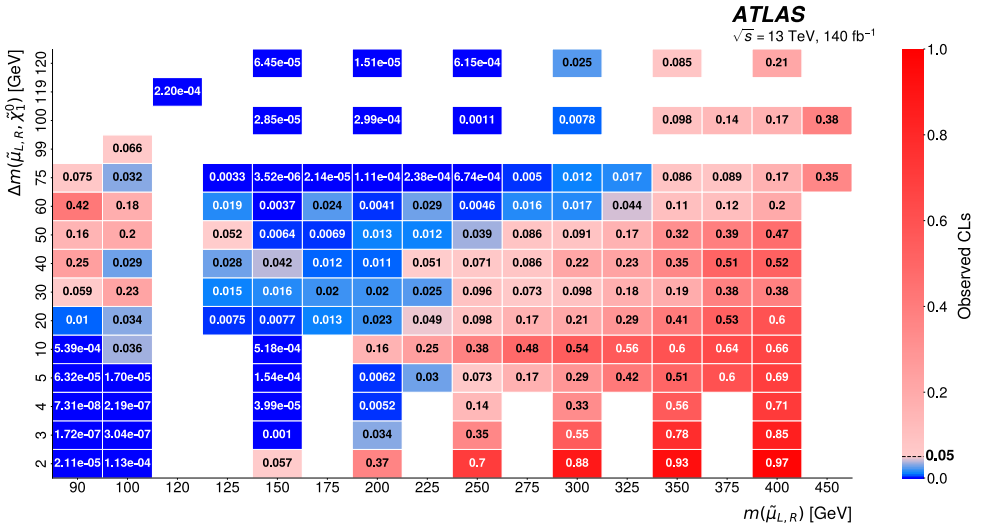


(b)

Figure F.11: Summary of the (a) expected and (b) observed CL_s values across the $m(\tilde{e}_{\mathrm{L,R}}) - \Delta m(\tilde{e}_{\mathrm{L,R}}, \tilde{\chi}_1^0)$ plane in the $m(\tilde{e}) \equiv m(\tilde{e}_{\mathrm{L}}) = m(\tilde{e}_{\mathrm{R}})$ interpretation using the BDT-based selection. The boxes are colour-coded to mark benchmark models for which the CL_s value is found to be below 0.05 (blue) or above (red).

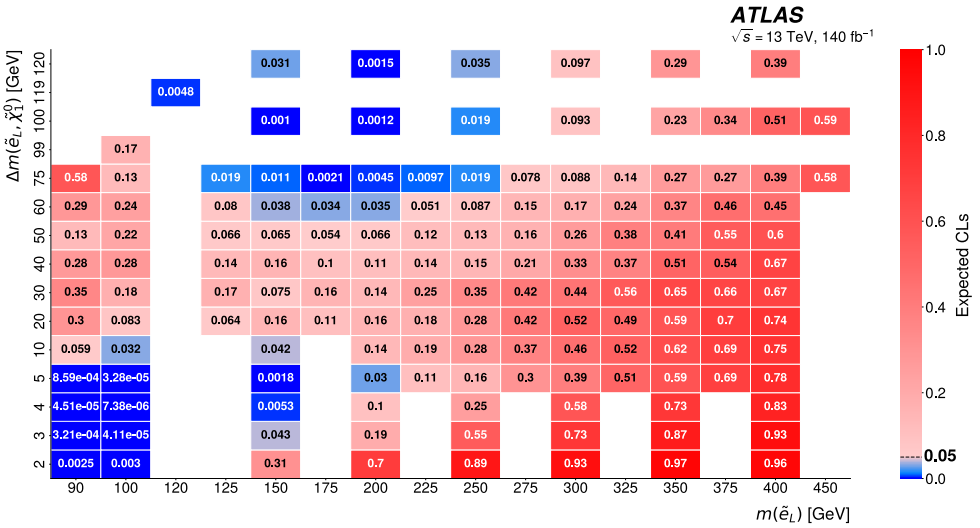


(a)

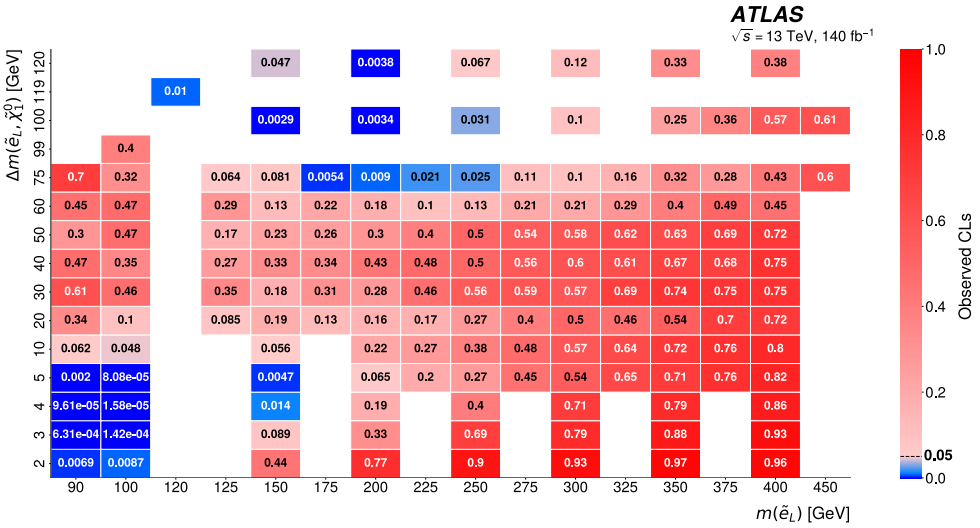


(b)

Figure F.12: Summary of the (a) expected and (b) observed CL_s values across the $m(\tilde{\mu}_{L,R}) - \Delta m(\tilde{\mu}_{L,R}, \tilde{\chi}_1^0)$ plane in the $m(\tilde{\mu}) \equiv m(\tilde{\mu}_L) = m(\tilde{\mu}_R)$ interpretation using the BDT-based selection. The boxes are colour-coded to mark benchmark models for which the CL_s value is found to be below 0.05 (blue) or above (red).

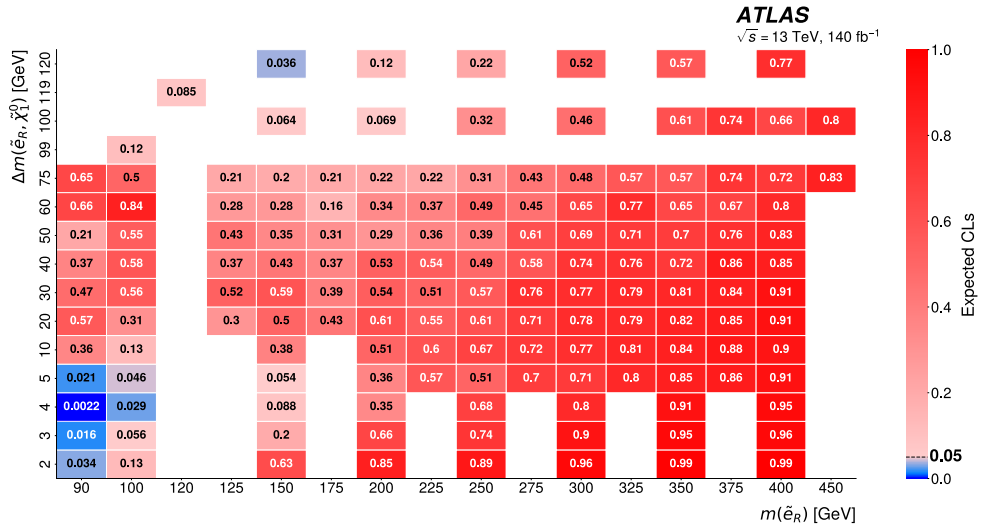


(a)

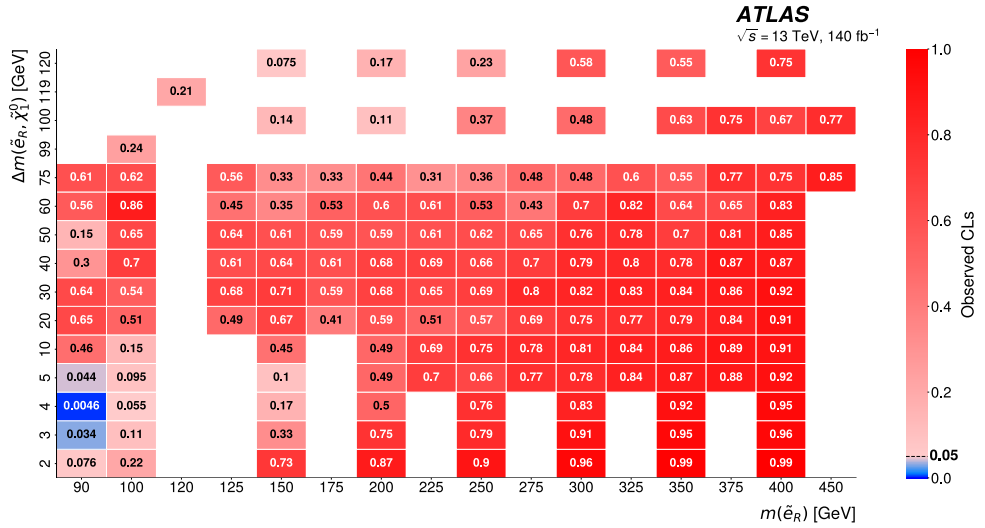


(b)

Figure F.13: Summary of the (a) expected and (b) observed CL_s values across the $m(\tilde{e}_L)$ – $\Delta m(\tilde{e}_L, \tilde{\chi}_1^0)$ plane in the $m(\tilde{e}) \equiv m(\tilde{e}_L)$ interpretation using the BDT-based selection. The boxes are colour-coded to mark benchmark models for which the CL_s value is found to be below 0.05 (blue) or above (red).

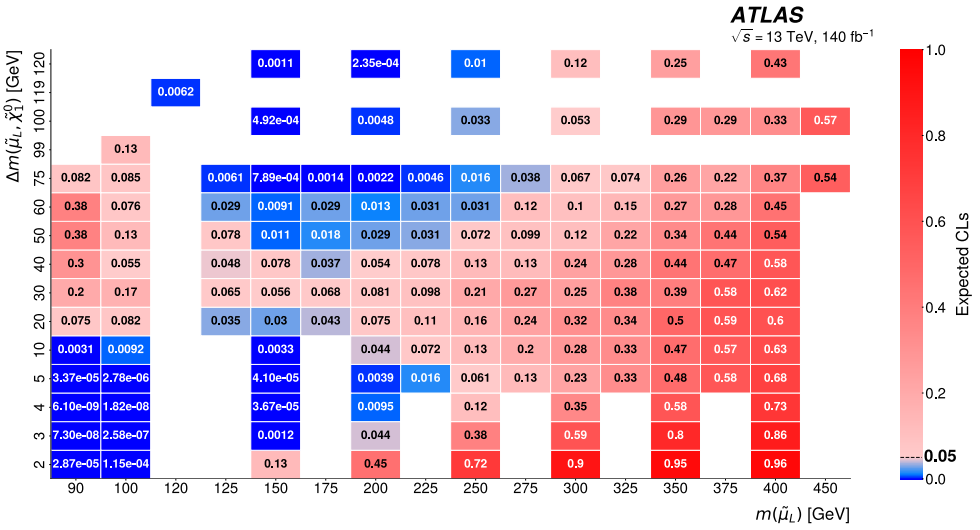


(a)

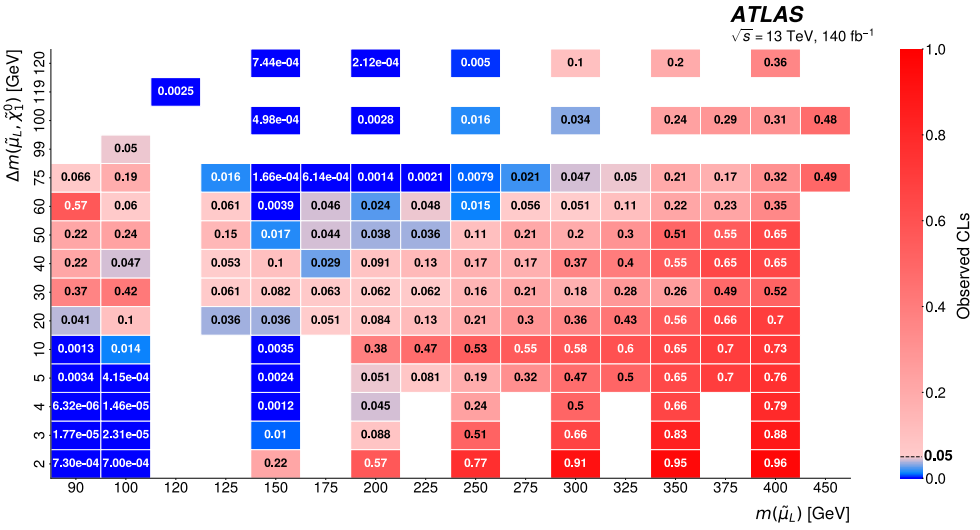


(b)

Figure F.14: Summary of the (a) expected and (b) observed CL_s values across the $m(\tilde{e})$ – $\Delta m(\tilde{e}_R, \tilde{\chi}_1^0)$ plane in the $m(\tilde{\mu}) \equiv m(\tilde{\mu}_L)$ interpretation using the BDT-based selection. The boxes are colour-coded to mark benchmark models for which the CL_s value is found to be below 0.05 (blue) or above (red).

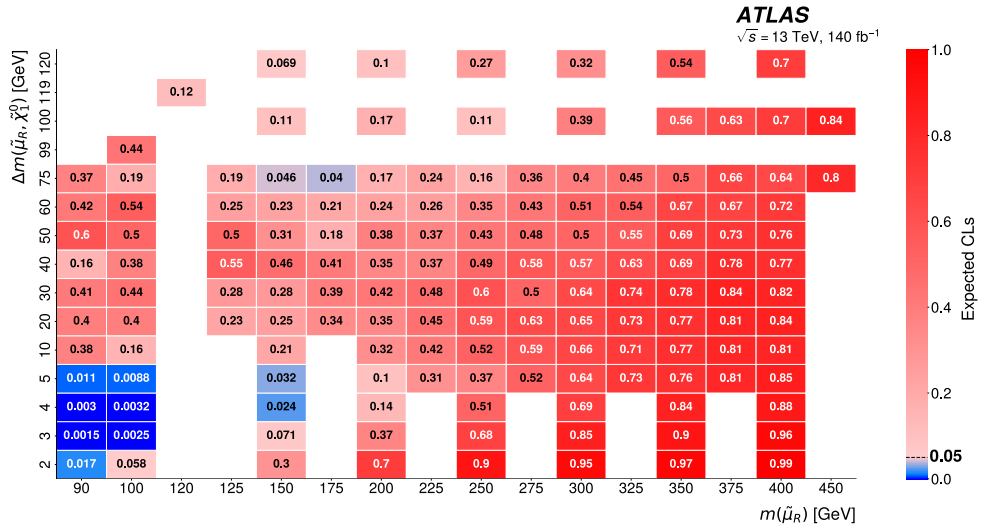


(a)

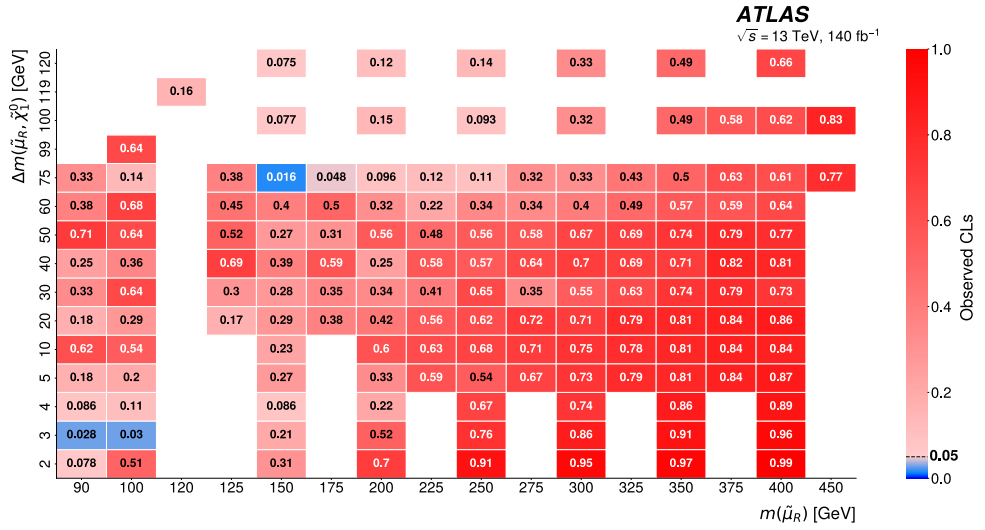


(b)

Figure F.15: Summary of the (a) expected and (b) observed CL_s values across the $m(\tilde{\mu}_L)$ – $\Delta m(\tilde{\mu}_L, \tilde{\chi}_1^0)$ plane in the $m(\tilde{e}) \equiv m(\tilde{e}_R)$ interpretation using the BDT-based selection. The boxes are colour-coded to mark benchmark models for which the CL_s value is found to be below 0.05 (blue) or above (red).



(a)



(b)

Figure F.16: Summary of the (a) expected and (b) observed CL_s values across the $m(\tilde{\mu}_R) - \Delta m(\tilde{\mu}_R, \tilde{\chi}_1^0)$ plane in the $m(\tilde{\mu}) \equiv m(\tilde{\mu}_R)$ interpretation using the BDT-based selection. The boxes are colour-coded to mark benchmark models for which the CL_s value is found to be below 0.05 (blue) or above (red).

G
The Pileup Fit

Contents

G.1 Particle Flow Object implementation of pufit 227

This appendix gives further details on the Pufit calculation. The implementation differs slightly for TCPufit and PFOPufit, and the differences for PFOPufit are detailed in Section G.1.

First, the detector is divided into uniform towers (bins) in η, ϕ. Specifically, there are 14 bins between $-5 < \eta < 5$ and 8 bins in ϕ, $-\pi < \phi < \pi$. A threshold on the transverse energy is defined to group the towers into pileup and hard scatter towers, as defined in Eq. 6.3 and repeated here for clarity,

$$E_T^{\mathrm{Th}} = \langle E_T \rangle + n_\sigma \sqrt{V}. \tag{G.1}$$

The Pufit algorithm uses a χ^2 fit of the hard scatter towers, to derive corrections ($\mathcal{E}_i$) to the E_T of the hard scatter towers, using information from the pileup towers as constraints. The χ^2 fit over the hard scatter towers is constructed using information from the pileup towers,

$$\chi^2 = s\boldsymbol{\Delta}_\mu^{\mathrm{T}} V_\mu^{-1} \boldsymbol{\Delta}_\mu + \sum_{i=1}^{N} \frac{\Delta_i^2}{V_{\mathrm{Tower}}}. \tag{G.2}$$

Here first term implements the constraint: $E_{\mathrm{T}}^{\mathrm{miss}}$ from pileup collisions sums to zero. The second term constraint implements the constraint: E_T from pileup is approximately linearly distributed across the detector, and the sum over 'i' is the sum over all hard scatter towers. The constraint importance, 's', controls the weighting of the two constraints against each other. In the second term, Δ_j is defined,

$$\Delta_j = \mathcal{E}_j - \langle E_T \rangle \,, \tag{G.3}$$

where the weighting depends on $\langle E_T \rangle$, the mean energy of the towers, as defined in Eq. 6.2, and V_{Tower}, the variance of the towers, as defined in Eq. 6.4. In the first term, $\mathbf{\Delta}_\mu$ is defined,

$$\mathbf{\Delta}_\mu = \mathbf{E}_{T_\mu} + \sum_{j=1}^{N} \mathcal{E}_j \mathbf{\Phi}_j. \tag{G.4}$$

The sum over 'j' is again the sum over all hard scatter towers. The corrections, $\mathcal{E}_j$, are derived from the χ^2 fit, and correct the E_T of the j$^{\mathrm{th}}$ hard scatter tower for pileup effects. Additionally, $\mathbf{\Phi}_j = \binom{\cos\phi_j}{\sin\phi_j}$, encodes the angular position of each tower. In the first term $\mathbf{E}_{T_\mu}$ and V_μ terms are the mean and variance of the pileup towers only. Here, the weights of the constraint are given by the covariance matrix of the sum over pileup objects,

$$\begin{aligned} \mathbf{E}_{\mathbf{T}\mu} &= \sum_{i \in \mu} \mathbf{p}_{\mathbf{T}_i}, \\ V_\mu &= \sum_{i \in \mu}^{N} \sigma_i \mathbf{\Phi}_i \mathbf{\Phi}_i^{\mathrm{T}}. \end{aligned} \tag{G.5}$$

The sum over '$i \in \mu$' indicates a sum over pileup towers only. The term $\sigma = a + b\sqrt{E_T}$ is an estimate of the resolution of the pileup towers. The 'a' is the stochastic coefficient, and 'b' is the noise coefficient, as defined in Table 6.2, with the values determined from calorimeter cell noise studies. The corrections ($\mathcal{E}$) are derived by minimising the χ^2,

$$\begin{aligned} \frac{\partial}{\partial \mathcal{E}_k} \chi^2 &= 0, \\ = 2s \frac{\partial \mathbf{\Delta}_\mu^{\mathrm{T}}}{\partial \mathcal{E}_k} V_\mu^{-1} \mathbf{\Delta}_\mu &+ 2 \sum_{j=1}^{N} \frac{\partial \Delta_j}{\partial \mathcal{E}_k} \frac{\Delta_j}{V_{\mathrm{tower}}}. \end{aligned} \tag{G.6}$$

After substitution of the quantities defined above, and subsequent simplification, this gives,

$$0 = \left(s\boldsymbol{\Phi}_j{}^{\mathrm{T}} V_\mu{}^{-1} \boldsymbol{\Phi}_k + \frac{\delta_{jk}}{V_{\mathrm{tower}}} \right) \mathcal{E}_k + s\boldsymbol{\Phi}_j{}^{\mathrm{T}} V_\mu{}^{-1} \mathbf{E}_{T_\mu} - \frac{\langle \boldsymbol{E}_T \rangle_{\mathrm{tower}}}{V_{\mathrm{tower}}}. \tag{G.7}$$

This can be written as a matrix equation which is partically solved analytically:

$$\begin{aligned} \mathcal{E}_j &= X_{ij}^{-1} c_i, \\ X_{ij} &= s\boldsymbol{\Phi}_i{}^{\mathrm{T}} V_\mu{}^{-1} \boldsymbol{\Phi}_j + \frac{\delta_{ij}}{V_{\mathrm{tower}}}, \\ c_i &= \frac{\langle E_T \rangle_{\mathrm{tower}}}{V_{\mathrm{tower}}} - s\boldsymbol{\Phi}_j{}^{\mathrm{T}} V_\mu{}^{-1}. \end{aligned} \tag{G.8}$$

The fit produces an estimate of the pileup contribution to the E_T in each hard scatter tower, which is then subtracted and the $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated,

$$\begin{aligned} E_T^{\mathrm{miss}} &= \sqrt{(-\hat{E}_x)^2 + (-\hat{E}_y)^2}, \\ \hat{E}_x &= \sum_{k=1}^{m} \left(E_{x_k} - \mathcal{E}_{T_k} \cos\phi_k \right), \\ \hat{E}_y &= \sum_{k=1}^{m} \left(E_{y_k} - \mathcal{E}_{T_k} \sin\phi_k \right). \end{aligned} \tag{G.9}$$

Where the sum over 'k' is the sum over all hard scatter towers. As the fit is only performed on the hard scatter towers, the linear matrix equation is typically $4-5$ dimensional for an event, and able to be solved rapidly by the HLT.

G.1 Particle Flow Object implementation of pufit

The procedure is slightly modified in PFOPufit, due to the construction of the PFOs. The PFOs are split into Charged PFlow Objects (cPFOs) and neutral PFlow Objects (nPFOs). Due to the track information in cPFOs, it is possible to determine whether an object is from hard scatter (cPFOHS) or pileup (cPFOPU). The cPFOs can only be reconstructed in the central region covered by the tracking detectors.

For nPFOs these represent only neutral objects in the central region, and both neutral and charged objects in the forward region, as forward objects fall out of the acceptance of the tracking detectors. Given the permutations on which combinations of PFOs to utilise at various points in the pufit calculation, there

are 432 unique combinations when implementing pufit using PFOs. A detailed discussion of the final combination was chosen is detailed in Ref [194]. The cPFOPU are not used at all in the optimal calculation, as these provide well measured information, linked to the primary vertex. Only this final choice will be discussed here, highlighting how it differs from TCPufit.

The trimmed mean and variance are calculated using only the nPFOs and the energy threshold determined seperately for the central and forward nPFOs. In the fit therefore, there are two sets of terms, one for central nPFOs and one for forward nPFOs. So, the final corrections are applied to the nPFOs that are identified as hard scatter. Then the $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated using the vector sum of all cPFOHS and the corrected hard scatter towers from the nPFOs. Finally, the PFOs from muons are manually included as $E_{\mathrm{T}}^{\mathrm{miss}}$ in the calculation, to match the behaviour in TCPufit.

All those moments will be lost in time... like tears in rain...

Blade Runner

References

[1] ATLAS Collaboration. *Searches for direct slepton production in the compressed-mass corridor in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*. 2025. arXiv: 2503.17186 [hep-ex]. URL: https://arxiv.org/abs/2503.17186.

[2] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08001.

[3] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003.

[4] Esma Mobs. "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018". In: (2018). General Photo. URL: https://cds.cern.ch/record/2636343.

[5] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13 (2018), P07017.

[6] S van der Meer. *Calibration of the effective beam height in the ISR*. Tech. rep. Geneva: CERN, 1968. URL: https://cds.cern.ch/record/296752.

[7] ATLAS Collaboration. "Luminosity determination in *pp* collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 83 (2023), p. 982. arXiv: 2212.09379 [hep-ex].

[8] ATLAS Collaboration. *ATLAS Luminosity Public Results for Run-2*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Accessed: 2025-04-24. 2025.

[9] Joao Pequenao. "Computer generated image of the whole ATLAS detector". 2008. URL: https://cds.cern.ch/record/1095924.

[10] Neutelings, Izaak. *CMS coordinate system – TikZ.net*. (Retrieved April 25, 2025). 2021. URL: https://tikz.net/axis3d_cms/.

[11] Maria Mironova. "Experimental Particle Physics with the ATLAS Detector". In: *Search for Higgs Boson Decays to Charm Quarks with the ATLAS Experiment and Development of Novel Silicon Pixel Detectors*. Cham: Springer Nature Switzerland, 2023, pp. 23–53. URL: https://doi.org/10.1007/978-3-031-36220-0_3.

[12] *ATLAS magnet system: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/338080.

[13] ATLAS Collaboration. "Alignment of the ATLAS Inner Detector in Run 2". In: *Eur. Phys. J. C* 80 (2020), p. 1194. arXiv: 2007.07624 [hep-ex].

[14] B. Abbott et al. "Production and integration of the ATLAS Insertable B-Layer". In: *JINST* 13 (2018), T05008. arXiv: 1803.00844 [physics.ins-det].

[15] ATLAS Collaboration. "The ATLAS Inner Detector commissioning and calibration". In: *Eur. Phys. J. C* 70 (2010), p. 787. arXiv: 1004.5293 [physics.ins-det].

[16] ATLAS Collaboration. *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-018. 2015. URL: https://cds.cern.ch/record/2037683.

[17] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061.

[18] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331062.

[19] Joao Pequenao. "Computer Generated image of the ATLAS calorimeter". 2008. URL: https://cds.cern.ch/record/1095927.

[20] The ATLAS collaboration. *Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter*. Tech. rep. Geneva: CERN, 2007. URL: https://cds.cern.ch/record/1026994.

[21] A Artamonov et al. "The ATLAS Forward Calorimeter". In: *JINST* 3 (2008), P02010. URL: https://cds.cern.ch/record/1094547.

[22] *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068.

[23] Joao Pequenao. "Computer generated image of the ATLAS Muons subsystem". 2008. URL: https://cds.cern.ch/record/1095929.

[24] Gianluca Rigoletti, Roberto Guida, and Beatrice Mandelli. "Studies on RPC detectors operated with environmentally friendly gas mixtures in LHC-like conditions". In: *Eur. Phys. J. Plus* 138.9 (2023), p. 841. URL: https://cds.cern.ch/record/2875180.

[25] ATLAS Collaboration. "Operation of the ATLAS trigger system in Run 2". In: *JINST* 15 (2020), P10004. arXiv: 2007.12539 [physics.ins-det].

[26] Andy Buckley et al. "General-purpose event generators for LHC physics". In: *Physics Reports* 504.5 (2011), pp. 145–233. URL: https://www.sciencedirect.com/science/article/pii/S0370157311000846.

[27] Xiangdong Ji. "QCD Analysis of the Mass Structure of the Nucleon". In: *Physical Review Letters* 74.7 (Feb. 1995), pp. 1071–1074. URL: http://dx.doi.org/10.1103/PhysRevLett.74.1071.

[28] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), p. 823. arXiv: 1005.4568 [physics.ins-det].

[29] GEANT4 Collaboration, S. Agostinelli, et al. "GEANT4 – a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250.

[30] ATLAS Collaboration. *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*. ATL-PHYS-PUB-2010-013. 2010. URL: https://cds.cern.ch/record/1300517.

[31] ATLAS Collaboration. "Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton–proton run". In: *JINST* 8 (2013), P07004. arXiv: 1303.0223 [hep-ex].

[32] ATLAS Collaboration. "Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2". In: *Eur. Phys. J. C* 77 (2017), p. 673. arXiv: 1704.07983 [hep-ex].

[33] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45. eprint: https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/82/1/35/5518977/35_1.pdf. URL: https://doi.org/10.1115/1.3662552.

[34] ATLAS Collaboration. "A neural network clustering algorithm for the ATLAS silicon pixel detector". In: *JINST* 9 (2014), P09009. arXiv: 1406.7690 [hep-ex].

[35] ATLAS Collaboration. *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-026. 2015. URL: https://cds.cern.ch/record/2037717.

[36] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77 (2017), p. 490. arXiv: 1603.02934 [hep-ex].

[37] Walter Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. ATL-LARG-PUB-2008-002. 2008. URL: https://cds.cern.ch/record/1099735.

[38] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77 (2017), p. 466. arXiv: 1703.10485 [hep-ex].

[39] S. Navas et al. "Review of particle physics". In: *Phys. Rev. D* 110.3 (2024), p. 030001.

[40] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. arXiv: 0802.1189 [hep-ph].

[41] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. arXiv: 1111.6097 [hep-ph].

[42] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-015. 2015. URL: https://cds.cern.ch/record/2037613.

[43] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 96 (2017), p. 072002. arXiv: 1703.09665 [hep-ex].

[44] ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 76 (2016), p. 581. arXiv: 1510.03823 [hep-ex].

[45] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. ATLAS-CONF-2014-018. 2014. URL: https://cds.cern.ch/record/1700870.

[46] ATLAS Collaboration. "ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset". In: *Eur. Phys. J. C* 83 (2023), p. 681. arXiv: 2211.16345 [physics.data-an].

[47] ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 79 (2019), p. 970. arXiv: 1907.05120 [hep-ex].

[48] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data". In: *JINST* 19 (2024), P02009. arXiv: 2309.05471 [hep-ex].

[49] ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. ATLAS-CONF-2012-047. 2012. URL: https://cds.cern.ch/record/1449796.

[50] ATLAS Collaboration. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 79 (2019), p. 639. arXiv: 1902.04655 [physics.ins-det].

[51] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 76 (2016), p. 292. arXiv: 1603.05598 [hep-ex].

[52] ATLAS Collaboration. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 81 (2021), p. 578. arXiv: 2012.00578 [hep-ex].

[53] ATLAS Collaboration. "Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector". In: *Phys. Rev. D* 97 (2018), p. 072003. arXiv: 1712.08891 [hep-ex].

[54] ATLAS Collaboration. "The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using $140\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ pp collisions". In: (2024). arXiv: 2402.05858 [hep-ex].

[55] ATLAS Collaboration. *Object-based missing transverse momentum significance in the ATLAS Detector*. ATLAS-CONF-2018-038. 2018. URL: https://cds.cern.ch/record/2630948.

[56] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 78 (2018), p. 903. arXiv: 1802.08168 [hep-ex].

[57] Y.A. Golfand and E.P. Likhtman. "Extension of the Algebra of Poincare Group Generators and Violation of P Invariance". In: *JETP Lett.* 13 (1971). [Pisma Zh. Eksp. Teor. Fiz. **13** (1971) 452], p. 323.

[58] D.V. Volkov and V.P. Akulov. "Is the neutrino a goldstone particle?" In: *Phys. Lett. B* 46 (1973), p. 109.

[59] J. Wess and B. Zumino. "Supergauge transformations in four dimensions". In: *Nucl. Phys. B* 70 (1974), p. 39.

[60] Abdus Salam and J. Strathdee. "Super-symmetry and non-Abelian gauges". In: *Phys. Lett. B* 51 (1974), p. 353.

[61] J. Wess and B. Zumino. "Supergauge invariant extension of quantum electrodynamics". In: *Nucl. Phys. B* 78 (1974), p. 1.

[62] S. Ferrara and B. Zumino. "Supergauge invariant Yang-Mills theories". In: *Nucl. Phys. B* 79 (1974), p. 413.

[63] Stephen P. Martin. "A Supersymmetry Primer". In: *Adv. Ser. Direct. High Energy Phys.* 18 (1998), p. 1. arXiv: hep-ph/9709356.

[64] Sidney Coleman and Jeffrey Mandula. "All Possible Symmetries of the S Matrix". In: *Phys. Rev.* 159 (1967), pp. 1251–1256.

[65] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), p. 1. arXiv: 1207.7214 [hep-ex].

[66] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett. B* 716 (2012), p. 30. arXiv: 1207.7235 [hep-ex].

[67] Jonathan L. Feng. "Dark Matter Candidates from Particle Physics and Methods of Detection". In: *Annual Review of Astronomy and Astrophysics* 48.1 (Aug. 2010), pp. 495–545. URL: http://dx.doi.org/10.1146/annurev-astro-082708-101659.

[68] Dan Hooper. *TASI 2008 Lectures on Dark Matter*. 2009. arXiv: 0901.4090 [hep-ph]. URL: https://arxiv.org/abs/0901.4090.

[69] 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. URL: http://dx.doi.org/10.1016/j.physrep.2004.08.031.

[70] V. C. Rubin, W. K. Ford Jr., and N. Thonnard. "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc)." In: *The Astrophysical Journal* 238 (June 1980), pp. 471–487.

[71] Douglas Clowe et al. "A Direct Empirical Proof of the Existence of Dark Matter". In: *The Astrophysical Journal* 648.2 (Aug. 2006), pp. L109–L113. URL: http://dx.doi.org/10.1086/508162.

[72] Planck collaboration. "Planck2018 results: VI. Cosmological parameters". In: *Astronomy and Astrophysics* 641 (Sept. 2020), A6. URL: http://dx.doi.org/10.1051/0004-6361/201833910.

[73] WMAP collaboration. "Cosmological parameters from SDSS and WMAP". In: *Physical Review D* 69.10 (May 2004). URL: http://dx.doi.org/10.1103/PhysRevD.69.103501.

[74] Super-Kamiokande collaboration. "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in 0.31 megaton $\cdot$ years exposure of the Super-Kamiokande water Cherenkov detector". In: *Physical Review D* 95.1 (Jan. 2017). URL: http://dx.doi.org/10.1103/PhysRevD.95.012004.

[75] Glennys R. Farrar and Pierre Fayet. "Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry". In: *Phys. Lett. B* 76 (1978), p. 575.

[76] H. Goldberg. "Constraint on the Photino Mass from Cosmology". In: *Phys. Rev. Lett.* 50 (1983), p. 1419. Erratum: in: *Phys. Rev. Lett.* 103 (2009), p. 099905.

[77] John Ellis et al. “Supersymmetric relics from the big bang”. In: *Nucl. Phys. B* 238 (1984), p. 453.

[78] Mariangela Lisanti. “Lectures on Dark Matter Physics”. In: *New Frontiers in Fields and Strings*. WORLD SCIENTIFIC, Nov. 2016, pp. 399–446. URL: http://dx.doi.org/10.1142/9789813149441_0007.

[79] ATLAS Collaboration. “Combined Measurement of the Higgs Boson Mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ Decay Channels with the ATLAS Detector Using $\sqrt{s} = 7, 8,$ and 13 TeV pp Collision Data”. In: *Phys. Rev. Lett.* 131 (2023), p. 251802. arXiv: 2308.04775 [hep-ex].

[80] Freeman J. Dyson. “The S Matrix in Quantum Electrodynamics”. In: *Physical Review* 75 (1949), pp. 1736–1755.

[81] Freeman J. Dyson. “The Radiation Theories of Tomonaga, Schwinger, and Feynman”. In: *Physical Review* 75 (1949), pp. 486–502.

[82] Riccardo Barbieri and G.F. Giudice. “Upper bounds on supersymmetric particle masses”. In: *Nucl. Phys. B* 306 (1988), p. 63.

[83] B. de Carlos and J.A. Casas. “One-loop analysis of the electroweak breaking in supersymmetric models and the fine-tuning problem”. In: *Phys. Lett. B* 309 (1993), p. 320. arXiv: hep-ph/9303291.

[84] D. P. Aguillard et al. “Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20 ppm”. In: *Phys. Rev. Lett.* 131 (16 2023), p. 161802. arXiv: 2308.06230 [hep-ex].

[85] T. Aoyama et al. “The anomalous magnetic moment of the muon in the Standard Model”. In: *Phys. Rept.* 887 (2020), pp. 1–166.

[86] S. Borsanyi, Z. Fodor, J.N. Guenther, et al. “Leading hadronic contribution to the muon magnetic moment from lattice QCD”. In: *Nature* 593 (2021), pp. 51–55.

[87] Manimala Chakraborti, Sven Heinemeyer, and Ipsita Saha. “The new “MUON G-2” result and supersymmetry”. In: *The European Physical Journal C* 81.12 (Dec. 2021). arXiv: 2104.03287 [hep-ph].

[88] Motoi Endo et al. “Supersymmetric interpretation of the muon g – 2 anomaly”. In: *Journal of High Energy Physics* 2021.7 (July 2021). arXiv: 2104.03217 [hep-ph].

[89] Kaoru Hagiwara et al. “$(g-2)_\mu$ and $\alpha(M_Z^2)$ re-evaluated using new precise data”. In: *J. Phys. G* 38 (2011), p. 085003. arXiv: 1105.3149 [hep-ph].

[90] Marcin Badziak and Kazuki Sakurai. “Explanation of electron and muon g-2 anomalies in the MSSM”. In: *Journal of High Energy Physics* 2019.10 (Oct. 2019). URL: http://dx.doi.org/10.1007/JHEP10(2019)024.

[91] Motoi Endo and Wen Yin. “Explaining electron and muon g-2 anomaly in SUSY without lepton-flavor mixings”. In: *Journal of High Energy Physics* 2019.8 (Aug. 2019). URL: http://dx.doi.org/10.1007/JHEP08(2019)122.

[92] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[93] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579–588. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692.

[94] S. Dimopoulos, S. Raby, and Frank Wilczek. "Supersymmetry and the scale of unification". In: *Phys. Rev. D* 24 (6 Sept. 1981), pp. 1681–1683. URL: https://link.aps.org/doi/10.1103/PhysRevD.24.1681.

[95] L.E. Ibáñez and G.G. Ross. "Low-energy predictions in supersymmetric grand unified theories". In: *Physics Letters B* 105.6 (1981), pp. 439–442. URL: https://www.sciencedirect.com/science/article/pii/0370269381912004.

[96] N. Sakai. "Naturalness in Supersymmetric Guts". In: *Z. Phys. C* 11 (1981), p. 153.

[97] Pierre Fayet. "Supersymmetry and weak, electromagnetic and strong interactions". In: *Phys. Lett. B* 64 (1976), p. 159.

[98] Pierre Fayet. "Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions". In: *Phys. Lett. B* 69 (1977), p. 489.

[99] ATLAS Collaboration. "Search for a heavy charged Higgs boson decaying into a W boson and a Higgs boson in final states with leptons and b-jets in $\sqrt{s} = 13\,\text{TeV}$ pp collisions with the ATLAS detector". In: (2024). arXiv: 2411.03969 [hep-ex].

[100] ATLAS Collaboration. "Search for charged Higgs bosons produced in top-quark decays or in association with top quarks and decaying via $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ in 13 TeV pp collisions with the ATLAS detector". In: (2024). arXiv: 2412.17584 [hep-ex].

[101] LHCb collaboration. *Test of lepton flavour universality with $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$ decays*. 2024. arXiv: 2412.11645 [hep-ex]. URL: https://arxiv.org/abs/2412.11645.

[102] ATLAS Collaboration. "Precise test of lepton flavour universality in W-boson decays into muons and electrons in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Eur. Phys. J. C* 84 (2024), p. 993. arXiv: 2403.02133 [hep-ex].

[103] Daniele Alves et al. "Simplified models for LHC new physics searches". In: *J. Phys. G* 39 (2012), p. 105005. arXiv: 1105.2838 [hep-ph].

[104] Johan Alwall, Philip Schuster, and Natalia Toro. "Simplified models for a first characterization of new physics at the LHC". In: *Phys. Rev. D* 79 (2009), p. 075020. arXiv: 0810.3921 [hep-ph].

[105] Johan Alwall et al. "Searching for directly decaying gluinos at the Tevatron". In: *Phys. Lett. B* 666 (2008), p. 34. arXiv: 0803.0019 [hep-ph].

[106] A. Djouadi et al. "The Minimal supersymmetric standard model: Group summary report". In: *GDR (Groupement De Recherche) - Supersymetrie*. Dec. 1998. arXiv: hep-ph/9901246.

[107] ATLAS Collaboration. "ATLAS Run 2 searches for electroweak production of supersymmetric particles interpreted within the pMSSM". In: *JHEP* 05 (2024), p. 106. arXiv: 2402.01392 [hep-ex].

[108] Jesse Liu. "Discovery strategies for dark matter and Higgsinos at the LHC". Presented 29 May 2019. Oxford U., 2019. URL: https://cds.cern.ch/record/2683899.

[109] Benjamin Fuks et al. "Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV". In: *JHEP* 10 (2012), p. 081. arXiv: 1207.2159 [hep-ph].

[110] W. Beenakker et al. "Production of Charginos, Neutralinos, and Sleptons at Hadron Colliders". In: *Phys. Rev. Lett.* 83 (1999), p. 3780. arXiv: hep-ph/9906298. Erratum: in: *Phys. Rev. Lett.* 100 (2008), p. 029901.

[111] W. Beenakker et al. "Squark and gluino hadroproduction". In: *Int. J. Mod. Phys. A* 26 (2011), p. 2637. arXiv: 1105.1110 [hep-ph].

[112] Benjamin Fuks et al. "Precision predictions for electroweak superpartner production at hadron colliders with RESUMMINO". In: *Eur. Phys. J. C* 73 (2013), p. 2480. arXiv: 1304.0790 [hep-ph].

[113] Benjamin Fuks et al. "Revisiting slepton pair production at the Large Hadron Collider". In: *JHEP* 01 (2014), p. 168. arXiv: 1310.2621 [hep-ph].

[114] *SUSY July 2024 Summary Plot Update*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2904978.

[115] LEPSUSYWG, ALEPH, DELPHI, L3, OPAL Collaborations. "Combined LEP Chargino Results, up to 208 GeV for large m0". In: *LEPSUSYWG/04-01.1* (2004). URL: http://lepsusy.web.cern.ch/lepsusy/www/sleptons_summer04/slep_final.html.

[116] ATLAS Collaboration. "Search for direct production of charginos, neutralinos and sleptons in final states with two leptons and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 05 (2014), p. 071. arXiv: 1403.5294 [hep-ex].

[117] 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: *Eur. Phys. J. C* 80 (2020), p. 123. arXiv: 1908.08215 [hep-ex].

[118] 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: *JHEP* 06 (2023), p. 031. arXiv: 2209.13935 [hep-ex].

[119] ATLAS Collaboration. "Searches for electroweak production of supersymmetric particles with compressed mass spectra in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: *Phys. Rev. D* 101 (2020), p. 052005. arXiv: 1911.12606 [hep-ex].

[120] CMS Collaboration. "Search for supersymmetry in final states with two oppositely charged same-flavor leptons and missing transverse momentum in proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 04 (2021), p. 123. arXiv: 2012.08600 [hep-ex].

[121] CMS Collaboration. "Search for supersymmetry in final states with two or three soft leptons and missing transverse momentum in proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 04 (2022), p. 091. arXiv: 2111.06296 [hep-ex].

[122] CMS Collaboration. "Combined search for electroweak production of winos, binos, higgsinos, and sleptons in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Phys. Rev. D* 109 (2024), p. 112001. arXiv: 2402.01888 [hep-ex].

[123] Alan J. Barr, Ben Gripaios, and Christopher G. Lester. "Weighing Wimps with Kinks at Colliders: Invisible Particle Mass Measurements from Endpoints". In: *JHEP* 02 (2008), p. 014. arXiv: 0711.4008 [hep-ph].

[124] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011), p. 1554. arXiv: 1007.1727 [physics.data-an]. Erratum: in: *Eur. Phys. J. C* 73 (2013), p. 2501.

[125] ATLAS Collaboration. "ATLAS data quality operations and performance for 2015–2018 data-taking". In: *JINST* 15 (2020), P04003. arXiv: 1911.04632 [physics.ins-det].

[126] ATLAS Collaboration. "Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking". In: *JHEP* 08 (2020), p. 080. arXiv: 2005.09554 [hep-ex].

[127] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. arXiv: 1405.0301 [hep-ph].

[128] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. arXiv: 1410.3012 [hep-ph].

[129] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: https://cds.cern.ch/record/1966419.

[130] Jonathan Debove, Benjamin Fuks, and Michael Klasen. "Threshold resummation for gaugino pair production at hadron colliders". In: *Nucl. Phys. B* 842 (2011), p. 51. arXiv: 1005.2909 [hep-ph].

[131] Juri Fiaschi and Michael Klasen. "Neutralino-chargino pair production at NLO+NLL with resummation-improved parton density functions for LHC Run II". In: *Phys. Rev. D* 98.5 (2018), p. 055014. arXiv: 1805.11322 [hep-ph].

[132] Giuseppe Bozzi, Benjamin Fuks, and Michael Klasen. "Threshold resummation for slepton-pair production at hadron colliders". In: *Nucl. Phys. B* 777 (2007), p. 157. arXiv: hep-ph/0701202.

[133] Juri Fiaschi and Michael Klasen. "Slepton pair production at the LHC in NLO+NLL with resummation-improved parton densities". In: *JHEP* 03 (2018), p. 094. arXiv: 1801.10357 [hep-ph].

[134] A.J. Barr. "Measuring slepton spin at the LHC". In: *JHEP* 02 (2006), p. 042. arXiv: hep-ph/0511115.

[135] Tom Melia. "Spin before mass at the LHC". In: *Journal of High Energy Physics* 2012.1 (Jan. 2012). URL: http://dx.doi.org/10.1007/JHEP01(2012)143.

[136] Alan Barr and James Scoville. "A boost for the EW SUSY hunt: monojet like search for compressed sleptons at LHC14 with 100 fb-1". In: *Journal of High Energy Physics* 2015.4 (Apr. 2015). URL: http://dx.doi.org/10.1007/JHEP04(2015)147.

[137] Howard Baer, Azar Mustafayev, and Xerxes Tata. "Monojet plus soft dilepton signal from light higgsino pair production at LHC14". In: *Physical Review D* 90.11 (Dec. 2014). URL: http://dx.doi.org/10.1103/PhysRevD.90.115007.

[138] C. G. Lester and D. J. Summers. "Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders". In: *Phys. Lett. B* 463 (1999), pp. 99–103. arXiv: hep-ph/9906349.

[139] A. Barr, C. G. Lester, and P. Stephens. "A variable for measuring masses at hadron colliders when missing energy is expected; m_{T2}: the truth behind the glamour". In: *J. Phys. G* 29 (2003), pp. 2343–2363. arXiv: hep-ph/0304226.

[140] Alan J. Barr and Claire Gwenlan. "The race for supersymmetry: Using m_{T2} for discovery". In: *Physical Review D* 80.7 (Oct. 2009). URL: http://dx.doi.org/10.1103/PhysRevD.80.074007.

[141] "XGBoost: A Scalable Tree Boosting System". In: *CoRR* abs/1603.02754 (2016). arXiv: 1603.02754. URL: http://arxiv.org/abs/1603.02754.

[142] Leo Breiman et al. *Classification and Regression Trees*. Wadsworth & Brooks/Cole, 1986.

[143] Jerome H. Friedman. "Greedy Function Approximation: A Gradient Boosting Machine". In: *The Annals of Statistics* 29.5 (2001), pp. 1189–1232.

[144] Frank J. Massey. "The Kolmogorov-Smirnov Test for Goodness of Fit". In: *Journal of the American Statistical Association* 46.253 (1951), pp. 68–78.

[145] Scott M Lundberg and Su-In Lee. "A Unified Approach to Interpreting Model Predictions". In: (2017). Ed. by I. Guyon et al., pp. 4765–4774. URL: http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.

[146] ATLAS Collaboration. "Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC". In: *JINST* 18 (2023), T11004. arXiv: 2211.16178 [hep-ex].

[147] M. Baak et al. "HistFitter software framework for statistical data analysis". In: *Eur. Phys. J. C* 75 (2015), p. 153. arXiv: 1410.1280 [hep-ex].

[148] Alexander L. Read. "Presentation of search results: the CL_S technique". In: *J. Phys. G* 28 (2002), p. 2693.

[149] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data". In: *JINST* 14 (2019), P12006. arXiv: 1908.00005 [hep-ex].

[150] ATLAS Collaboration. "Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Eur. Phys. J. C* 81 (2021), p. 689. arXiv: 2007.02645 [hep-ex].

[151] ATLAS Collaboration. "Determination of jet calibration and energy resolution in proton–proton collisions at $\sqrt{s} = 8\,\text{TeV}$ using the ATLAS detector". In: *Eur. Phys. J. C* 80 (2020), p. 1104. arXiv: 1910.04482 [hep-ex].

[152] ATLAS Collaboration. "Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13\,\text{TeV}$ collected with the ATLAS detector". In: *Eur. Phys. J. C* 82 (2022), p. 95. arXiv: 2109.10627 [hep-ex].

[153] M. Bähr et al. "Herwig++ physics and manual". In: *Eur. Phys. J. C* 58 (2008), p. 639. arXiv: 0803.0883 [hep-ph].

[154] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. arXiv: 1512.01178 [hep-ph].

[155] Stefan Höche et al. "A Study of QCD Radiation in VBF Higgs Production with Vincia and Pythia". In: *SciPost Phys.* 12.1 (2022), p. 010. arXiv: 2106.10987 [hep-ph].

[156] Stefano Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. arXiv: 0805.3067 [hep-ph].

[157] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. ATL-PHYS-PUB-2016-020. 2016. URL: https://cds.cern.ch/record/2216168.

[158] ATLAS Collaboration. *Formulae for Estimating Significance*. ATL-PHYS-PUB-2020-025. 2020. URL: https://cds.cern.ch/record/2736148.

[159] ALEPH Collaboration. "Absolute lower limits on the masses of selectrons and sneutrinos in the MSSM". In: *Phys. Lett. B* 544 (2002), pp. 73–88. arXiv: hep-ex/0207056.

[160] *General search for supersymmetric particles in scenarios with compressed mass spectra using proton-proton collisions at* $\sqrt{s} = 13$ TeV. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2904926.

[161] Paul Jackson and Christopher Rogan. "Recursive jigsaw reconstruction: HEP event analysis in the presence of kinematic and combinatoric ambiguities". In: *Physical Review D* 96.11 (Dec. 2017). URL: http://dx.doi.org/10.1103/PhysRevD.96.112007.

[162] Matthew R. Buckley et al. "Super-razor and searches for sleptons and charginos at the LHC". In: *Physical Review D* 89.5 (Mar. 2014). URL: http://dx.doi.org/10.1103/PhysRevD.89.055020.

[163] Diyar Agin et al. "Seeking a coherent explanation of LHC excesses for compressed spectra". In: *The European Physical Journal C* 84.11 (Nov. 2024). URL: http://dx.doi.org/10.1140/epjc/s10052-024-13594-9.

[164] ATLAS Collaboration. "Search for chargino–neutralino pair production in final states with three leptons and missing transverse momentum in $\sqrt{s} = 13\,\text{TeV}$ pp collisions with the ATLAS detector". In: *Eur. Phys. J. C* 81 (2021), p. 1118. arXiv: 2106.01676 [hep-ex].

[165] CMS Collaboration. "Search for electroweak production of charginos and neutralinos in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *JHEP* 04 (2022), p. 147. arXiv: 2106.14246 [hep-ex].

[166] The ATLAS Collaboration. "The ATLAS trigger system for LHC Run 3 and trigger performance in 2022". In: *Journal of Instrumentation* 19.06 (June 2024), P06029. URL: http://dx.doi.org/10.1088/1748-0221/19/06/P06029.

[167] The ATLAS Collaboration. "Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking". In: *Journal of High Energy Physics* 2020.8 (Aug. 2020). URL: http://dx.doi.org/10.1007/JHEP08(2020)080.

[168] *Trigger monitoring and rate predictions using Enhanced Bias data from the ATLAS Detector at the LHC*. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2016-002. Geneva, Oct. 2016. URL: https://cds.cern.ch/record/2223498.

[169] "Technical Design Report for the ATLAS Inner Tracker Pixel Detector". In: (2017). URL: https://cds.cern.ch/record/2285585.

[170] ZEUS Collaborations The H1. "Combined measurement and QCD analysis of the inclusive $e^{\pm}$ – p scattering cross sections at HERA". In: *Journal of High Energy Physics* 2010.1 (Jan. 2010). URL: http://dx.doi.org/10.1007/JHEP01(2010)109.

[171] In: *Physics Reports* 427.5–6 (May 2006), pp. 257–454. URL: http://dx.doi.org/10.1016/j.physrep.2005.12.006.

[172] Torbjörn Sjöstrand. "Jet fragmentation of multiparton configurations in a string framework". In: *Nuclear Physics B* 248.2 (1984), pp. 469–502. URL: https://www.sciencedirect.com/science/article/pii/0550321384906072.

[173] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys. B* 867 (2013), p. 244. arXiv: 1207.1303 [hep-ph].

[174] Richard D. Ball et al. "Parton distributions for the LHC run II". In: *JHEP* 04 (2015), p. 040. arXiv: 1410.8849 [hep-ph].

[175] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys. G* 43 (2016), p. 023001. arXiv: 1510.03865 [hep-ph].

[176] Stefano Frixione, Giovanni Ridolfi, and Paolo Nason. "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: *JHEP* 09 (2007), p. 126. arXiv: 0707.3088 [hep-ph].

[177] Paolo Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. arXiv: hep-ph/0409146.

[178] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. arXiv: 0709.2092 [hep-ph].

[179] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. arXiv: 1002.2581 [hep-ph].

[180] Tanju Gleisberg and Stefan Höche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. arXiv: 0808.3674 [hep-ph].

[181] Steffen Schumann and Frank Krauss. "A parton shower algorithm based on Catani–Seymour dipole factorisation". In: *JHEP* 03 (2008), p. 038. arXiv: 0709.1027 [hep-ph].

[182] ATLAS Collaboration. *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*. ATL-PHYS-PUB-2016-017. 2016. URL: https://cds.cern.ch/record/2206965.

[183] ATLAS Collaboration. *Multi-Boson Simulation for* 13 *TeV ATLAS Analyses*. ATL-PHYS-PUB-2017-005. 2017. URL: https://cds.cern.ch/record/2261933.

[184] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS detector". In: *JHEP* 09 (2014), p. 145. arXiv: 1406.3660 [hep-ex].

[185] Leif Lönnblad and Stefan Prestel. "Matching tree-level matrix elements with interleaved showers". In: *JHEP* 03 (2012), p. 019. arXiv: 1109.4829 [hep-ph].

[186] Torbjorn Sjöstrand. "Jet fragmentation of multiparton configurations in a string framework". In: *Nucl. Phys. B* 248 (1984), pp. 469–502.

[187] Bo Andersson et al. "Parton fragmentation and string dynamics". In: *Phys. Rept.* 97 (1983), pp. 31–145.

[188] B.R. Webber. "A QCD model for jet fragmentation including soft gluon interference". In: *Nuclear Physics B* 238.3 (1984), pp. 492–528. URL: https://www.sciencedirect.com/science/article/pii/055032138490333X.

[189] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff. "A modified cluster-hadronization model". In: *Eur. Phys. J. C* 36 (2004), pp. 381–395. arXiv: hep-ph/0311085.

[190] D. Amati and G. Veneziano. "Preconfinement as a property of perturbative QCD". In: *Physics Letters B* 83.1 (1979), pp. 87–92. URL: https://www.sciencedirect.com/science/article/pii/0370269379908967.

[191] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152.

[192] Karl Pearson. "Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia". In: *Philosophical Transactions of the Royal Society of London. Series A* 187 (1896), pp. 253–318.

[193] Lorenzo Moneta et al. *The RooStats Project*. 2011. arXiv: 1009.1003 [physics.data-an]. URL: https://arxiv.org/abs/1009.1003.

[194] Kenji Hamano. *Pufit Note*. ATL-COM-DAQ-2019-140. Geneva, Aug. 2019. URL: https://cds.cern.ch/record/2687693.