

The search for new high-mass resonances in the dilepton final state using the full Run-2 ATLAS dataset of 13 TeV proton-proton collisions

by

Etienne Dreyer

BSc, University of the Fraser Valley, 2015

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Physics
Faculty of Science

© Etienne Dreyer 2021
SIMON FRASER UNIVERSITY
Summer 2021

Copyright in this work rests with the author. Please ensure that any reproduction
or re-use is done in accordance with the relevant national copyright legislation.



Declaration of Committee

Name: Etienne Dreyer

Degree: Doctor of Philosophy (Experimental High Energy Physics)

Thesis title: The search for new high-mass resonances in the dilepton final state using the full Run-2 ATLAS dataset of 13 TeV proton-proton collisions

Examining Committee:

Chair: Karen Kavanagh
Professor

Bernd Stelzer
Supervisor
Professor

Dugan O’Neil
Committee Member
Professor

Michel Vetterli
Committee Member
Professor

Matthias Danninger
Internal Examiner
Assistant Professor

Christopher Hearty
External Examiner
Professor
Department of Physics & Astronomy
University of British Columbia

Date Defended: July 27, 2021

Abstract

There are good reasons to anticipate that the Standard Model (SM), successful as it is, will nonetheless give way to a more complete description of particle interactions. New physics scenarios beyond the SM often address its key shortcomings through extended gauge symmetries with the implied existence of additional massive bosons. The dilepton final state is a particularly clean experimental channel to explore for new heavy resonances, as was the case for SM discoveries of the 1970's and 80's. This thesis presents a search for high-mass dilepton resonances using the full dataset collected by the ATLAS detector during Run 2 of the LHC. The analysis targets the invariant mass spectrum of an inclusive selection of dielectron and dimuon events with some of the highest energies reported by ATLAS. Prompted by the unprecedented statistical precision of the dataset, the results are based on a new approach for estimating the SM backgrounds using a functional form fit. While no significant excess is observed, the resulting upper limits on the cross section for production of a dilepton resonance are used to constrain the parameter space of models that predict new bosons coupled to leptons. For example, the existence of a Sequential Standard Model Z' boson is excluded for pole masses below 5.1 TeV. Updated bounds are also derived for the mass and leptonic couplings of a hypothetical spin-1 dark matter mediator.

Keywords: particle physics; LHC; high-mass resonance; grand unification theory; dark matter

Dedication

In memory of my oupa, Dr. Heinrich “Hein” Dreyer.

Acknowledgements

It is not possible for me here to acknowledge the large number of people who in various ways have encouraged, supported, and influenced me along my PhD journey. But I am glad for this chance to highlight some of them...

Thank you first of all to my supervisor Prof. Bernd Stelzer for convincing me to go in the direction of high energy physics after my BSc. Talking with you has always left me with more clarity and motivation. Without your steady backing and advice I would have missed several opportunities that have had a defining impact on my PhD.

My studies have been enabled in large part by funding from NSERC and I also am grateful for the generous financial support of Dr. Howard Malm.

I thank Heberth Torres for investing in me as a new student. Many of my eventual research directions trace back to these first projects. Former postdocs Kate Pachal and Alyssa Montalbano – you were a big help! To Konstantin Lehman, I’m grateful to have had you as a PhD companion and for all your listening to me. Long research hours were freshened by good conversations with officemates Koos van Nieuwkoop, Sina Bahrasemani, Benjamin Jäger, and “resident theorist” Subrata Mandal.

To fellow members of the Run-2 dilepton analysis teams – I am grateful for the experience to have worked with and learned from you all. Thanks in particular to Daniel Hayden, Noam Tal Hod, and Paolo Mastrandrea for mentoring me and allowing me to present a lot of half-baked results. Thank you Peter Falke for the many technical and coordination skills I was able to learn because of working alongside you. Also thank you to Tetiana Berger-Hryn’ova for sharing your reinterpretation expertise and for supervising my extended work at CERN.

I am thankful for each friendship with fellow graduate students in the SFU Physics Department. The past six years would have been pretty colorless without your camaraderie and the ability to share frustration and mutual interests alike. Thank you to the brilliant classmates who were a vital resource in grad coursework – Alex Zucca, Adrian Yeung, and Matthew Martin in particular.

Finally, to my family, I have been so blessed to study nearby and have your support through the ups and downs. Mom, thank you; you’ve cheered me on so faithfully and you know how much it meant to me. My church family both here and at CERN have been a much needed mainstay. To the Cookmans who made me feel like family during my stays at CERN, thank you.

Statement of Contribution

During my PhD I contributed to three papers associated with the ATLAS dilepton analysis effort. This thesis focuses on the full Run-2 resonant analysis [1] and describes the relevant research, to which I contributed chiefly in the following areas:

1. sample preparation (chapter 6)
 - helping develop and updating the analysis selection framework
 - requesting Monte Carlo samples and fully processing analysis inputs
 - deriving the spin-0 and spin-2 signal efficiencies
2. modelling the fake electron rate and $m_{\ell\ell}$ distribution using the matrix method (ch. 7)
3. statistical analysis (chapter 8)
 - early research of the data-driven background estimate strategy and validation
 - creating a functional form-based statistical analysis framework, `Rooswift`, and using it to produce independent p -value scans and upper limits (appendix C)
 - studies of the look-elsewhere effect (appendix B)
4. reinterpretation of generic cross section limits (chapter 9)
 - calculating benchmark Z' signal cross sections and PDF uncertainty
 - coding the framework to derive exclusion contours on heavy vector triplet couplings, including width-matching via interpolated limits
 - generating dark matter mediator samples with `MADGRAPH` and producing mass and coupling exclusion contours
5. formatting analysis results for HEPData entry (chapter 10)

I also served as one of the two analysis coordinators for the full Run-2 resonant publication in which I shared responsibility to organize the team effort, chair analysis meetings, write supporting documentation, and meet the approval requirements of the ATLAS Collaboration. Following the publication I served as Statistics Committee liaison to the ATLAS Exotics group.

*There once was a detector called ATLAS
Its electron reconstruction was flawless
One pair in nine hundred thousand
It past 4 TeV founded
But no sign of Z' – alas!*

Table of Contents

Declaration of Committee	ii
Abstract	iii
Dedication	iv
Acknowledgements	v
Statement of Contribution	vi
Table of Contents	viii
1 Introduction	1
1.1 Analysis strategy synopsis	3
2 Theoretical backdrop	5
2.1 Introduction to the Standard Model of particle physics	5
2.2 Electroweak theory	9
2.2.1 Quantum Electrodynamics	9
2.2.2 Weak interactions	11
2.2.3 The Brout-Englert-Higgs mechanism	14
2.3 Quantum chromodynamics	18
2.3.1 Asymptotic freedom and quark confinement	19
2.4 High energy physics toolkit	21
2.4.1 Particle interactions and cross sections	21
2.4.2 Particle colliders	22
2.4.3 Invariant mass in terms of transverse and angular observables	24
2.5 Simulation of ATLAS data	26
2.5.1 Parton distribution function	27
2.5.2 Event generation with Monte Carlo	29
2.5.3 Detector simulation	32
3 Scenarios beyond the Standard Model	34

3.1	The unnaturally light Higgs mass	35
3.2	The disparity in strengths of the fundamental forces	37
3.2.1	Grand Unified Theories	37
3.2.2	Extra-dimensional models	39
3.3	The nature of dark matter	42
3.4	Dilepton resonances	45
4	The ATLAS detector at the LHC	47
4.1	Overview of ATLAS and the LHC	47
4.1.1	Coordinate system and kinematic variables used in ATLAS	49
4.2	Magnet system	50
4.3	Muon spectrometer	51
4.3.1	Closer look at monitored drift tubes	53
4.3.2	Muon spectrometer alignment	53
4.4	Inner detector	54
4.4.1	Pixel detector	55
4.4.2	SCT: Silicon microstrip tracker	56
4.4.3	TRT: transition radiation tracker	56
4.5	Calorimeters	57
4.5.1	LAr-based calorimeters	60
4.5.2	Tile calorimeter	62
4.6	From detector to dataset	62
4.6.1	Trigger and data acquisition (TDAQ)	62
4.6.2	The worldwide LHC computing grid (<i>The Grid</i>)	64
4.6.3	Data formats	65
5	Reconstructing electrons and muons in ATLAS	66
5.1	Reconstructing electrons	66
5.1.1	Electron identification	67
5.1.2	Electron isolation	68
5.1.3	Challenges of electrons at high- E_T	69
5.2	Reconstructing muons	70
5.2.1	The high- p_T muon selection working point	72
5.2.2	Muon isolation	73
5.2.3	Challenges of muons, particularly at high- p_T	73
6	Dilepton event selection	76
6.1	Standard Model background processes	76
6.2	Simulation of dilepton processes	80
6.2.1	Higher-order K -factors for correcting Drell-Yan MC	82

6.3	Selection criteria for high-mass dilepton analyses	83
6.3.1	Event-level criteria	83
6.3.2	Lepton-level criteria	85
6.3.3	Differences in event selection between the 36.1 fb^{-1} and 139 fb^{-1} analyses	87
6.4	Corrections applied to MC	88
6.5	Selection results	89
6.6	Dilepton invariant mass response model	97
6.6.1	The transfer function method	97
6.6.2	Invariant mass resolution and selection efficiency	99
6.6.3	Application of the transfer function to background smoothing	101
6.6.4	Application of the transfer function to signal modelling	102
6.7	Fiducial cuts for generic limits	103
6.8	Systematic uncertainties	105
6.8.1	Signal normalization uncertainties	108
6.8.2	Signal shape uncertainties	109
7	Estimating the fake electron background	110
7.1	Origins of fake electrons	111
7.2	The matrix method	112
7.3	Measuring real electron rate	113
7.4	Measuring fake electron rate	114
7.4.1	Fake-enriched data sample	115
7.4.2	Real contamination subtraction	117
7.4.3	Fake rates after subtraction	118
7.5	Obtaining the m_{ee} template for fake electrons	118
8	Statistical analysis	124
8.1	Discovery statistics	124
8.1.1	Hypothesis testing	124
8.1.2	CL_s limit setting	129
8.2	Global fit	132
8.2.1	Motivation	132
8.2.2	Description	134
8.2.3	Validation	136
8.2.4	Uncertainties	140
8.3	Summary of systematic uncertainties and impact	142
8.4	Statistical results	144
8.4.1	Signal significance scans	145
8.4.2	Upper limits on signal cross section	147

9	Model exclusions by reinterpretation	151
9.1	Signal reweighting	151
9.2	Scenario #1: Benchmark Z' models	153
9.2.1	Deriving PDF uncertainty bands on the Z'_{SSM} cross section	155
9.3	Scenario #2: Heavy Vector Triplet model	157
9.3.1	HVT signal width and cross section values	158
9.3.2	HVT exclusion contours	159
9.4	Scenario #3: Dark matter spin-1 mediator	164
9.4.1	Dark matter mediator “mass-mass” exclusion contours	165
9.4.2	Dark matter s-channel summary plots	168
9.4.3	Dark matter mediator “coupling-mass” exclusion contours	172
10	Summary and outlook	173
	Bibliography	175
	Appendix A Additional dilepton modelling plots	188
	Appendix B Additional statistical analysis material	199
B.1	The look-elsewhere effect	201
B.1.1	Asymptotic behaviour of trials factor for 1D search	202
B.1.2	Asymptotic behaviour of trials factor for 2D search	204
B.2	Comparison of q_μ and q_0 distributions for upper limits	209
	Appendix C Sliding window fit crosscheck	212
C.1	Overview	212
C.2	Method	213
C.3	RooSwift	213
C.4	Uncertainties	214
C.5	Optimization	215
C.6	Results based on SWiFt	220
C.6.1	Significance scans	220
C.6.2	Limit curves for 2015 – 17 dataset (80 fb^{-1})	221
C.7	Comparison between SWiFt and the global fit	223
C.7.1	Spurious signal estimate	223
C.7.2	Limit curves without systematic uncertainties	224
C.7.3	Limit curves including systematic uncertainties	224
C.7.4	Conclusions from comparison	225
	Appendix D Additional reinterpretation material	227
D.1	Additional HVT material	227

D.2 Additional DM material	230
List of Tables	233
List of Figures	235

Chapter 1

Introduction

Notwithstanding the astonishing development of subatomic physics over the past 120 years, there is today no shortage of outstanding questions. What is the nature of dark matter? What leads to the fundamental forces having such disparate strengths? Can we explain the small mass of the Higgs boson apart from finely-tuned model parameters? These are among the questions that form a broad motivation to continue investigating new physics scenarios beyond the Standard Model (SM). Several theoretical solutions propose an extended gauge symmetry group beyond the SM, and hence predict the existence of additional massive gauge bosons. This thesis details the latest search for a new neutral boson producing the particularly clean *dilepton* final state – two high energy leptons¹. Notably, the search was conducted using the latest “full Run-2” dataset collected by the ATLAS detector at the world’s highest-energy particle accelerator complex: the Large Hadron Collider (LHC).

The dilepton final state has historically been an important channel for discovering new particles. Its characteristic observable of the dilepton invariant mass $m_{\ell\ell}$ is fully-reconstructable and able to resolve resonances of new particles acting as neutral mediators. The dimuon invariant mass spectrum reconstructed from one of the earliest datasets collected by ATLAS in 2011 [2] reveals in clear succession the resonances first of the J/ψ and Υ mesons discovered in the 1970’s [3, 4, 5] and then the electroweak Z boson discovered in 1983 [6]. This early dataset was used for the first search by the ATLAS Collaboration for new dilepton resonances with masses beyond the Z peak back in 2011 [7]. In the summer of 2015, following significant upgrades, the LHC turned on again with an over 50% increase in collision centre-of-mass energy from $\sqrt{s} = 8$ TeV to 13 TeV. The resulting extended reach in mediator mass prompted new iterations of the search for dilepton resonances using first 3.2 fb⁻¹ of data [8] and then 36.1 fb⁻¹ in 2017 [9]. The analysis strategy for both partial-Run-2 publications was very similar.

Upon finding no significant indications of a new dilepton resonance, the analysis team took some time to restructure instead of simply repeat the same analysis strategy for the

¹In the context of the dilepton analysis, “leptons” refers to electrons and muons.

full Run-2 dataset. Something which was made apparent by the higher data rates in Run 2 was the challenge of producing sufficiently large simulated samples of the dominant dilepton background processes. In the 2017 analysis, the suboptimal statistical precision of simulated samples was seen to directly reduce the search’s sensitivity, particularly at low mass where signals with small cross sections would be partially obscured. Yet with $\sqrt{s}$ more or less fixed in the coming LHC runs, the growing dataset size will be the main driver of increased experimental sensitivity. Hence the full Run-2 analysis was the golden opportunity to deploy for the first time a data-driven strategy for analysing the dilepton invariant mass spectrum. In this approach, the data are used to directly constrain the shape of a functional form-based background estimate after having been validated against the simulated SM background distributions. A good deal of chapter 8 is devoted to this procedure.

Another aspect of the restructuring aimed to streamline the interpretation of analysis results by the wider high energy physics community, ensuring a lasting impact for the full Run-2 legacy paper. This motivated the shift to a statistical analysis based solely on a generic, parametric signal model. To ensure applicability to a wide range of specific models, upper limits were set on the generic signal cross section within a well-defined fiducial volume to effectively sidestep model-dependent features. The procedure for reinterpreting these generic upper limits into specific scenarios was demonstrated for two use cases published together with the analysis and in a follow-up dark matter mediator interpretation. The resulting exclusions obtained for these three scenarios are discussed in chapter 9.

Despite its significant technical and methodological changes, the revamped dilepton analysis proceeded efficiently. So much so that figures from the preprint presented at the 137th LHCC Meeting in February 2019 became the first public results using the full Run-2 ATLAS dataset, less than three months after the LHC shutdown. Just days prior to this release, the results were updated to incorporate the brand new estimate of 139 fb^{-1} as the full Run-2 integrated luminosity (a number which has since been validated). Since then, the analysis results have been featured in many conference talks and also reinterpretations outside of ATLAS – a synopsis is provided in the summary of this thesis. The final paper published in *Physics Letters B* in September 2019 [1] is the work at the heart of this thesis.

This thesis begins by setting the theoretical backdrop of physics at the LHC (chapter 2) and the reasons for expecting physics beyond the SM (chapter 3). Chapter 4 is an overview of the experimental setup, i.e. the ATLAS detector. The process of reconstructing high-momentum electrons and muons in ATLAS is discussed in more detail in chapter 5. The selection criteria behind the $m_{\ell\ell}$ spectrum and the modelling of detector $m_{\ell\ell}$ response is discussed in chapter 6. Chapter 7 focuses on the estimate of background contributions from misidentified electrons. Finally, chapter 8 sets up the statistical analysis approach and shows the results, and chapter 9 draws out the implications of those results for three signal scenarios.

1.1 Analysis strategy synopsis

The following synopsis is put here as an easy reference for the reader while moving through the chapters describing the analysis in detail.

The analysis of the dilepton invariant mass ($m_{\ell\ell}$) spectrum in the ATLAS full Run-2 dataset is illustrated in figure 1.1. The analysis pipeline begins with the code which applies object (i.e. lepton) and event selection criteria to the data and the MC background samples to produce templates for the analysis (section 6.3). Ultimately the data template is used to directly constrain the shape of a functional form used as a background estimate (figure 8.6). This, combined with the generic signal model, is the basis for the statistical analysis resulting in p -value scans and cross section upper limits (section 8.4).

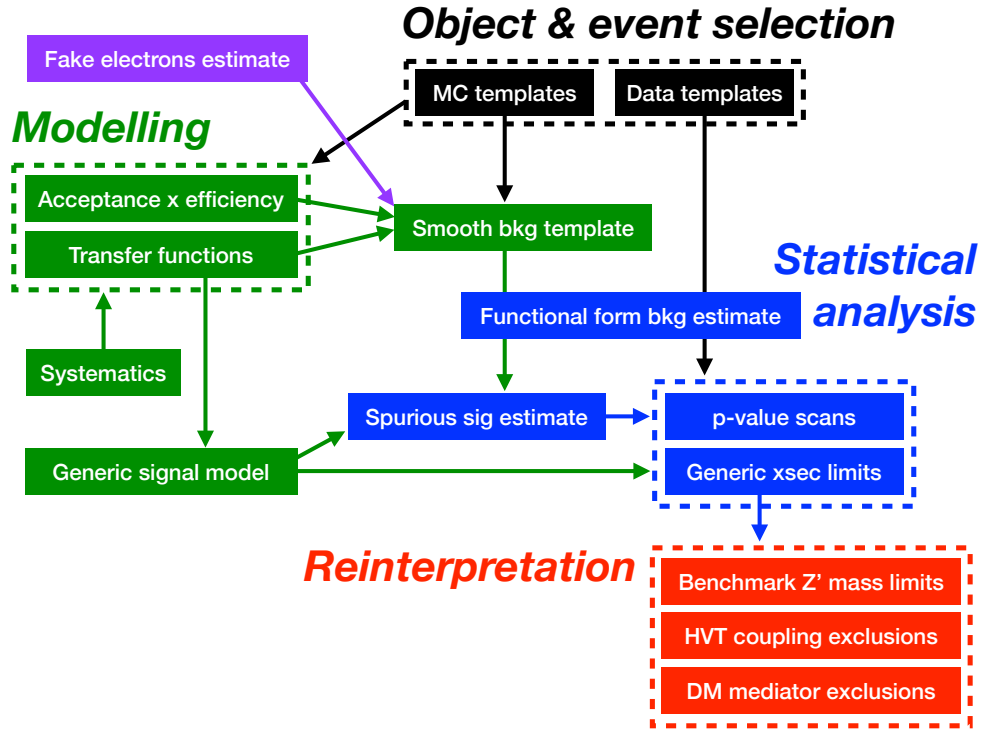


Figure 1.1: Schematic overview of the strategy used to analyse the full Run-2 dilepton invariant mass spectrum. “bkg” and “sig” refer to background and signal.

Modelling of the selected $m_{\ell\ell}$ distribution plays an important role both in validating the functional form background estimate and in deriving the generic signal model. The MC templates of the dominant SM background (Drell-Yan) are the basis for modelling the acceptance times efficiency and the $m_{\ell\ell}$ response via transfer functions (section 6.6). The generic signal model is obtained by convoluting a Breit-Wigner resonance with the transfer functions (section 6.6.4). Similarly, a convolution kernel involving the transfer function is used to smooth a MC background template used to gauge/optimize the functional form

fit performance before using it on data (sections 6.6.3 and 8.2.3). The smooth background template also includes an estimate of the smallest background process coming from fake electrons (chapter 7).

The final stage of the analysis produces bounds on specific signal models by reinterpreting the cross section limits obtained for the generic signal model (chapter 9). This boils down to overlaying predicted and observed cross sections within a common fiducial volume and obtaining a smooth exclusion from the points where they intersect.

Chapter 2

Theoretical backdrop

2.1 Introduction to the Standard Model of particle physics

The most complete and experimentally successful framework to date for describing the properties of fundamental particles and their interactions is known as the Standard Model (SM). The SM is a theory of interacting quantum fields, both bosonic and fermionic. Crucially, the mathematics describing their field dynamics are expressly invariant under a set of transformations. For the SM these include not only Lorentz transformations but also internal transformations generated by elements of the following three symmetry groups:

$$G_{\text{SM}} = \underbrace{SU(3)_c}_{\text{strong}} \times \underbrace{SU(2)_L \times U(1)_Y}_{\text{electroweak}} \quad (2.1)$$

The subscripts identify the groups with their corresponding conserved quantum numbers of color, weak isospin (*Left-chiral*), and hypercharge Y . The threefold symmetry group structure of the SM represents three of the four known fundamental forces¹. What is notably absent is a description of the most familiar fundamental force – gravity. The other fundamental force operative on macroscopic systems – electromagnetism – is formulated in the SM in close connection with the less intuitive weak nuclear force. To appreciate this, it is important to recall how historically, the individual forces of electricity and magnetism had been harmonized through Maxwell’s equations. So the SM further extends this paradigm, demonstrating how in more complicated way, electromagnetism itself emerges from a common interaction shared with the weak nuclear force. This electroweak interaction is described in (2.1) by the symmetry group product $SU(2) \times U(1)$. Finally the strong nuclear force enters independently through the $SU(3)$ group.

There are three types of quantum field contained in different sectors of the SM as shown in table 2.1: fermions (comprising leptons and quarks), gauge bosons, and the Higgs field.

¹It is curious that groups with degrees of “3,2,1” should underlie such a fundamental description of nature. Chapter 3 outlines how the SM symmetry groups could potentially unify at high energies.

Sector	Spin	Field	Species	$Q = Y + T_3$						$SU(2)_L$	$SU(3)_c$
Fermion	1/2	spinor	lepton	chirality	generation			(anti-) particle			
					I	II	III				
				L	$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L$	$\begin{matrix} +1 & -1 \\ (0)? & 0 \end{matrix}$		2	1
			quark	R	e_R	μ_R	τ_R	$\begin{matrix} +1 & -1 \end{matrix}$		1	
				L	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$\begin{matrix} -2/3 & +2/3 \\ +1/3 & -1/3 \end{matrix}$		2	3
				R	$\begin{matrix} u_R \\ d_R \end{matrix}$	$\begin{matrix} c_R \\ s_R \end{matrix}$	$\begin{matrix} t_R \\ b_R \end{matrix}$	$\begin{matrix} -2/3 & +2/3 \\ +1/3 & -1/3 \end{matrix}$		1	
Gauge boson	1	real vector	B					0		1	1
			W					$\begin{pmatrix} +1 \\ 0 \\ -1 \end{pmatrix}$		3	1
			g					0		1	8
Higgs	0	complex scalar	H					$\begin{pmatrix} +1 \\ 0 \end{pmatrix}$		2	1

Table 2.1: Field content of the Standard Model prior to electroweak symmetry breaking. The quantum numbers Q , Y , and T_3 refer to electric charge, hypercharge and weak isospin, respectively. Subscripts L and R label the chirality of a state. The question mark for antineutrinos signifies that the Dirac vs. Majorana nature of neutrinos is not yet known.

The wavefunctions representing these three types of field assign to each point in spacetime a value that is either spinor, vector, or (complex) scalar. Elementary particles can be thought of as quantized excitations of these fields.

In constructing a viable theory of fundamental interactions such as the SM, two concepts are foundational: the principle of least action and the requirement of Lorentz and local gauge invariance. The principle of least action asserts that the predominant behaviour of a physical system corresponds to the trajectory which minimizes the time-integral of its Lagrangian. The full Lagrangian of the system is therefore sufficient to capture its entire dynamics. For

the SM the Lagrangian (density) can be written in the following condensed form²:

$$\begin{aligned}\mathcal{L}_{\text{SM}} = & -\frac{1}{4}G_{\mu\nu}^k G^{k\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ & + \bar{\psi}_f(i\gamma^\mu D_\mu - m_f)\psi_f \\ & + (D^\mu\phi)^\dagger D_\mu\phi + \mu^2\phi^\dagger\phi - \frac{\lambda^2}{2}(\phi^\dagger\phi)^2\end{aligned}\quad (2.2)$$

where the gauge indices a and k run from 1 to 3 and from 1 to 8, respectively. The first line comprises the kinetic and self-interaction terms of the strong and electroweak gauge vector fields which are represented by their field strength tensors $G_{\mu\nu}^k$ and $W_{\mu\nu}^a$, $B_{\mu\nu}$ respectively. The term of the Lagrangian in the second line gives rise to the Dirac equation for fermion spinor fields ψ_f , where the 4-gradient operator ∂_μ has been replaced by the covariant derivative D_μ (the mass term will be discussed in section 2.2.3). The last line includes the kinetic, gauge-interaction (via D_μ), and potential terms of the Higgs complex scalar field doublet ϕ . The parameters μ and λ in the potential are discussed in section 2.2.3. The covariant derivative D_μ represents a local modification of the spacetime gradient by the presence of gauge fields, to preserve local gauge invariance of the Lagrangian. It is through this connection that gauge fields interact with fermions and the Higgs field:

$$D_\mu\{\psi_f, \phi\} = \left[\partial_\mu - \underbrace{ig'YB_\mu}_{\psi \neq \psi_\nu^R} - \underbrace{ig_w T_3 W_\mu^a \sigma^a}_{\psi = \psi^L} - \underbrace{ig_s A_\mu^k \lambda^k}_{f \in \text{quarks}} \right] \{\psi_f, \phi\} \quad (2.3)$$

Here, A_μ^k are the gluon fields, and σ^a and λ^k are generators of the SM gauge groups, to be discussed (table 2.2). The couplings denoted by g 's determine the strength of the gauge interactions. The constants Y and T_3 are the *hypercharge* and *weak isospin* quantum numbers of the fields which are related to the electric charge Q as follows:

$$Q = Y + T_3 \quad (2.4)$$

This relationship is a consequence of electroweak symmetry breaking (section 2.2.3).

The second foundational concept, the requirement of Lorentz and local gauge invariance, determines which terms are allowed in a Lagrangian. According to Noether's theorem, such symmetry requirements are equivalent to statements that certain quantities remain conserved in the system. For example, the symmetry (or invariance) of a Lagrangian under spacetime translations implies that the resulting dynamics will conserve energy-momentum.

²Lagrangian “density” must be integrated in all four spacetime dimensions to get the action. This form admits some major caveats and simplifications: it is shown prior to electroweak symmetry breaking (except for fermion mass term placeholder), Higgs and fermion doublet structures are condensed, fermion species and chiralities are lumped together, and neutrino Dirac mass terms are implicitly included.

Gauge symmetry, on the other hand, ensures that particle interactions will conserve the currents associated with the corresponding gauge field (implying e.g. conservation of electric charge for the photon gauge field). Gauge theories are very successful at describing the fundamental forces in the SM – a realization partly due to 't Hooft and Veltman [10] who demonstrated that they are renormalizable also in the non-Abelian case (1999 Nobel Prize).

The central role of mathematical symmetry groups, then, is to encode the sets of transformations which leave the Lagrangian unchanged. The canonical example is that of a global phase shift applied to a wavefunction, corresponding to a $U(1)$ transformation applied uniformly in spacetime. Since only relative phase differences are measurable (and thus physical), the dynamics of the system should be unaffected by this shift. In the case at hand, the SM Lagrangian is constructed to be invariant under the set of infinitesimal transformations generated by members of the three groups comprising 2.1. And unlike the simple example above of a global phase shift, gauge invariance is required for spacetime-dependent (local) transformations.

In accord with these symmetry demands, the SM fields are constructed to have definite transformation properties (e.g. Lorentz-covariant) such that they can be combined to form the invariant terms in (2.2). The spinor, vector, and scalar fields in table 2.1 reflect this property for Lorentz transformations. Similarly, for gauge transformations fermions come in multiplets of the symmetry group while gauge bosons are identified with the group generators. In terms of their representations of G_{SM} , arranged as $(c, L)_Y$, the SM fermion fields are arranged as follows ([11], §94):

$$\{Q_L, u_R/d_R, L, \ell_R\} = \left\{ (\mathbf{3}, \mathbf{2})_{\frac{1}{3}}, (\bar{\mathbf{3}}, \mathbf{1})_{-\frac{4}{3}}, (\mathbf{1}, \mathbf{2})_{-1}, (\mathbf{1}, \mathbf{1})_{-2} \right\} \quad (2.5)$$

where Q_L are left-handed quark doublets, u_R/d_R are right handed quarks, L is the left-handed lepton doublet and ℓ_R does not include right-handed neutrinos. Note that the quark and lepton fields have different representations depending on their chirality (L or R), as explained in section 2.2.2. The three types of field in the SM experience infinitesimal local gauge transformations differently, as follows:

$$\text{Fermion, Higgs : } \{\psi_f, \phi\} \rightarrow \{\psi_f', \phi'\} = e^{i\frac{q}{2}\omega^a(x)T^a} \{\psi_f, \phi\} \quad (2.6)$$

$$\text{Gauge boson : } A_\mu^a \rightarrow A_\mu^{a'} = A_\mu^a - \partial_\mu \omega^a(x) + f^{abc} \omega^b(x) A_\mu^c \quad (2.7)$$

where the T^a are the set of generators of a symmetry subgroup of G_{SM} (equation 2.1). The variables $\omega^a(x)$ parameterize the infinitesimal rotations in the basis of the generators and are in general spacetime-dependent (i.e. local). The structure constants f^{abc} are traceless antisymmetric tensors which result from the commutation of the generators as shown in table 2.2. In the case of the Abelian group $U(1)$, rotations are generated by a single complex phase (which is a function of spacetime) and hence commute.

Group	Generators	Commutation relation
$U(1)$	$T = 1$	$[T, T] = 0$ (2.8)
$SU(2)$	$T^a = \sigma^a$, $a = 1 \dots 3$	$[\sigma^a, \sigma^b] = 2i\epsilon^{abc}\sigma^c$ (2.9)
$SU(3)$	$T^k = \lambda^k/2$, $k = 1 \dots 8$	$[T^k, T^l] = if^{klm}T^m$ (2.10)

Table 2.2: Generators in the fundamental representation of Standard Model gauge groups and their commutation relations. σ^a are the 2×2 Pauli matrices and λ^k are the 3×3 Gell-Mann matrices. For the non-Abelian groups the structure constants ϵ^{abc} and f^{klm} are arrays of real numbers that are traceless and fully antisymmetric under index permutation.

2.2 Electroweak theory

2.2.1 Quantum Electrodynamics

In the 1920s quantum mechanics came of age through its successful description of electromagnetic bound states in the context of simple atoms. Two decades before, the new theory of special relativity had since established Lorentz invariance as a property of natural law, specifically in the context of electromagnetism. These two developments first converged in Dirac's formulation of a relativistic matter-wave equation for electrons in 1928 [12]. This first step toward the theory of Quantum Electrodynamics (QED) was a milestone in the physics of particle interactions which are both small (quantized) and fast (relativistic). And the theoretical groundwork along the way laid the foundation for quantum field theories of weak and strong nuclear interactions ultimately underpinning the SM.

The QED Lagrangian describing interactions of an electrically-charged fermion field $\psi(x)$ mediated by the photon (4-)vector field $A_\mu(x) = (\Phi, \vec{A})$ is the following:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + q\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (2.11)$$

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.12)$$

where γ^μ are four 4×4 matrices constrained to satisfy the anticommutation relation $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbf{1}_{4 \times 4}$. Though not unique, a popular basis for these *gamma matrices* is the Dirac-Pauli representation where they can be expressed in terms of the three Pauli matrices $\vec{\sigma}$:

$$\gamma^0 = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} \mathbf{0} & \sigma^i \\ -\sigma^i & \mathbf{0} \end{pmatrix} \quad (2.13)$$

The middle term in (2.11) describes interactions between fermions and photons. The last term, quadratic in derivatives of A_μ is the kinetic energy term for the photon field. This form is the gauge- and Lorentz-invariant trademark of a *Yang-Mills* theory. The QED Lagrangian,

with the full gauge-covariant derivative, is directly incorporated into the SM – however in (2.2) the field A_μ is shown in terms of its two basis fields prior to electroweak symmetry breaking (discussed in section 2.2.3). Applying the principle of least action to the QED Lagrangian leads to the following Euler-Lagrange equations of motion for the fields ψ and A_μ :

$$\text{Dirac equation:} \quad (i\gamma^\mu D_\mu - m)\psi = 0, \quad D_\mu \equiv \partial_\mu + iqA_\mu \quad (2.14)$$

$$\text{Gauss's, Ampere's law:} \quad \partial_\mu F^{\mu\nu} = qj^\nu, \quad j^\nu \equiv \bar{\psi}\gamma^\nu\psi \quad (2.15)$$

where j^ν is the 4-current comprised of charged fermion fields. The homogenous Maxwell's equations follow as a consequence of the field strength tensor differential properties. For each fermion species there exists a minimal basis of four plane wave solutions to (2.14) which can be understood as spanning the combinations of left/right chirality and particle/antiparticle states. They are the product of a 4-component spinor and the plane wave:

$$\psi(x) = \begin{bmatrix} u(p) \\ v(p) \end{bmatrix} e^{ip_\mu x^\mu} \quad (2.16)$$

where the bispinors satisfy the linear system

$$\begin{bmatrix} u \\ v \end{bmatrix} = \vec{\sigma} \cdot \vec{p} \begin{bmatrix} 1/(E - m) & 0 \\ 0 & 1/(E + m) \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix} \quad (2.17)$$

This ensures that the solutions satisfy the relativistic energy-momentum relation:

$$E^2 = \vec{p} \cdot \vec{p} + m^2 \quad (2.18)$$

It is important to differentiate the four column indices of ψ in (2.16) from Lorentz indices; spinors contract with their adjoints $\bar{\psi} \equiv \gamma^0 \psi^\dagger$ to form objects with definite Lorentz transformation properties, for example the vector combination $\bar{\psi}\gamma^\mu\psi$.

In the canonical approach, the quantization of (fermion) boson fields results from requiring the fields to (anti)commute with their conjugate fields according to

$$[A_\mu(\vec{x}, t), \dot{A}_\nu(\vec{y}, t)] = -i\hbar\eta_{\mu\nu}\delta^{(3)}(\vec{x} - \vec{y}) \quad (2.19)$$

$$\{\psi_m(\vec{x}, t), \psi_n^\dagger(\vec{y}, t)\} = \delta_{mn}\delta^{(3)}(\vec{x} - \vec{y}) \quad (2.20)$$

with all other combinations vanishing³. Equation 2.19 is analogous to the operator commutation in the Heisenberg uncertainty principle. For fermion fields it turns out the anti-

³Also, $\delta^{(3)}$ and δ_{mn} denote the 3-dimensional Dirac and the Kronecker delta functions, respectively, and $\eta_{\mu\nu}$ is the Minkowski spacetime metric.

commutator is needed for a causal ordering, related to the underlying difference in fermion vs. boson statistics. Expanding the fields in a basis of momentum space eigenstates and imposing these conditions generates an algebra of creation and annihilation operators for both fermion and photon states. This leads to the interpretation of ψ and A_μ as quantized fields and the ability to calculate scattering matrix elements between incoming and outgoing particle states (discussed further in section 2.4.1). Historically the QED calculation of the electron g factor to increasing orders of the fine structure constant α was an early confirmation of the theory. To date g remains the most precise prediction of the SM, with state-of-the-art experimental measurements and QED calculations agreeing to within less than one part per ten billion [13].

2.2.2 Weak interactions

Following its success in QED, quantum field theory offered a framework for describing weak nuclear interactions. Much of the field theory developed for QED can be analogously derived for weak interactions but needs to incorporate the following key differences:

- non-Abelian gauge group
 - three mediators instead of one
 - self-coupling of mediators
- massive mediators
- chiral theory

The latter two features of the theory of weak interaction were required by experimental observations. The earliest weak processes studied (e.g. beta decay and muon decay) were unlike electromagnetic processes in that they altered the particle species. Specifically, weak interactions of leptons produced a species of particle with zero electric charge later identified as the neutrino. Furthermore, the outgoing neutrinos always exhibited the same chirality and hence their production processes violated parity symmetry. Considering their rates compared to electromagnetic ones, weak interactions had evidently smaller cross sections. While the name *weak* interaction is therefore fitting, it is potentially misleading as to the origin of the small interaction cross sections. Indeed the weak force is about ten thousand times smaller than electromagnetism at low energies, yet its intrinsic coupling strength $\alpha_w = g_w^2/4\pi$ is actually nearly five times larger. In fact the relative smallness of weak interactions comes about because it is mediated by massive bosons, often much heavier than the energy scale of the interaction. As a consequence, the characteristic length scale of weak interactions is limited to $\sim 10^{-3}$ fm. Before the weak bosons were discovered, their masses needed to explain the weak force were substantially higher than any fundamental particle yet observed. Historically, then, the gauge theory of weak interactions was predicated on

the existence of particles which were on the one hand too fleeting (neutrinos) and on the other hand too massive (weak bosons) to be experimentally confirmed. Remarkably, in time both particle ingredients were discovered, setting a precedent for the SM that solidified the hope of likewise discovering the Higgs boson.

In the SM, the weak interaction is structured around the group $SU(2)_L$. Its generators, the three Pauli matrices σ^a , are embodied in the fundamental vector (i.e. Lorentz-indexed) gauge fields to form the weak field strength tensor which is contracted in (2.2) *à la* Yang-Mills:

$$W_{\mu\nu}^a = \partial_\mu \mathcal{W}_\nu^a - \partial_\nu \mathcal{W}_\mu^a + g_w \epsilon^{abc} \mathcal{W}_\mu^b \mathcal{W}_\nu^c, \quad \mathcal{W}_\mu^a \equiv \sigma^a \mathcal{W}_\mu \quad (2.21)$$

The weak isospin index a runs over the three generators, and ϵ^{abc} is the Levi-Civita permutation symbol. This is in contrast to the electromagnetic field strength tensor (2.12) arising from the single $U(1)$ generator. Another difference is the presence of a third term arising from the nonzero $SU(2)$ commutation relation (table 2.2) – effectively an isospin cross-product of the $\mathcal{W}_\mu$ fields. What is therefore apparent is that the contraction $W_{\mu\nu}^a W^{a\mu\nu}$ in equation (2.2) will contain terms to the third and fourth power in $\mathcal{W}_\mu$. These correspond to self-interaction processes of the weak vector bosons involving no fermions – something not exhibited by photons.

The “ L ” subscript of the weak symmetry group refers to the fact, deduced solely from empirical observations [14], that weak interactions only take place between left-chiral particle fields and between right-chiral antiparticle fields. Spinor states of definite chirality are defined as the two eigenstates of the γ_5 matrix operator:

$$\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & \mathbf{1}_{2\times 2} \\ \mathbf{1}_{2\times 2} & 0 \end{pmatrix} \quad (2.22)$$

with “right” corresponding to the $+1$ eigenvalue and “left” to the -1 eigenvalue. The weak force only couples to the projection of (anti-)fermion states onto the $(R)L$ chirality eigenstate, which is achieved using the following operator:

$$\mathcal{P}_{L/R} \equiv \frac{1 \mp \gamma_5}{2} \quad (2.23)$$

This implies, for example, that the weak decay of a muon, $\mu^- \rightarrow \nu_\mu + e^- + \bar{\nu}_e$, will take place for a muon in the left-chiral state but not for one in the right-chiral state. This observation for weak interactions is known as *maximal parity violation* – the weak current couples not only preferentially, but exclusively to left-chiral fermion fields (or right-chiral antifermion fields).

Intuition for the physical difference between states of left and right chirality can be gained by considering the massless limit. In this limit, a particle’s *helicity* – the projection

of its spin onto its momentum – becomes equivalent to its chirality. Therefore, a massless⁴ fermion which is fully in the right-chiral eigenstate is one which has its spin and momentum vectors oriented in the same direction (and for left-chiral states in the opposite direction). Spin and momentum are both directional degrees of freedom that are possessed by fermions which explains why they have the property of helicity despite being point-like objects.

The left-chiral fermion fields projected out using $\mathcal{P}_{L/R}$ transform as weak isospin doublets of $SU(2)$ and the right-chiral ones as singlets. The SM quarks and leptons are therefore grouped as follows:

$$L = \frac{1}{2}(1 - \gamma_5) \begin{pmatrix} \psi_{\nu_\ell} \\ \psi_\ell \end{pmatrix} \equiv \begin{pmatrix} \nu_\ell \\ \ell_L \end{pmatrix}, \quad \ell_R \equiv \frac{1}{2}(1 + \gamma_5)\psi_\ell, \quad \ell \in \{e, \mu, \tau\} \quad (2.24)$$

$$Q = \frac{1}{2}(1 - \gamma_5) \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix} \equiv \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad u_R/d_R \equiv \frac{1}{2}(1 + \gamma_5)\psi_{u/d}, \quad \begin{cases} u \in \{u, c, t\} \\ d \in \{d, s, b\} \end{cases} \quad (2.25)$$

Currents $j^{\mu a}$ for weak interactions are therefore sourced exclusively by contractions between the fermion doublets L and Q and their adjoints:

$$j^{\mu a} = \underbrace{\bar{\psi} \frac{1}{2}(1 + \gamma_5) \gamma^\mu \sigma^a}_{\bar{L}, \bar{Q}} \underbrace{\frac{1}{2}(1 - \gamma_5) \psi}_{L, Q} = \bar{\psi} \gamma^\mu \left(\frac{1 - \gamma_5}{2} \right) \sigma^a \psi \quad (2.26)$$

using the gamma matrix anticommutation relations and the identity $\gamma_5 \gamma_5 = 1$. Compared to the vector coupling (γ^μ) in the QED current (2.15) there is in (2.26) the subtraction of the axial-vector component: $(\gamma^\mu - \gamma^\mu \gamma_5)$. This is the so-called $V - A$ theory of weak interactions. In addition, the three possible rotations by the group generators σ^a lead to a collection of weak vertices – both those that are electrically neutral (e.g. $\bar{\nu} - \nu$ and $\bar{u}_L - u_L$) and those that are charged (e.g. $\nu_\ell - \bar{\ell}_L$ and $\bar{u}_L - d_L$).

The Lagrangian for $SU(2)_L \times U(1)_Y$ can now be written in terms of the chiral fermion fields, their minimal couplings to the gauge fields, and the Yang-Mills gauge kinetic terms:

$$\begin{aligned} \mathcal{L} = & \bar{L}(i\partial)L + \bar{Q}(i\partial)Q + \bar{\ell}_R(i\partial)\ell_R + \bar{u}_R(i\partial)u_R + \bar{d}_R(i\partial)d_R \\ & + g_w W_\mu^a j^{\mu a} + g' B_\mu j^{\mu} \\ & - \frac{1}{4} W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \end{aligned} \quad (2.27)$$

⁴For massive spinors, helicity becomes a Lorentz frame-dependent quantity and is therefore different from chirality. On the other hand, the chirality state of a massive fermion propagating freely evolves in time ([15], p. 12).

where $\not{\partial} \equiv \gamma^\mu \partial_\mu$. What is missing here are mass terms for the fermions and gauge bosons. For fermions, mass terms “crossover” between left- and right-chiralities,

$$\text{fermion mass term: } -m\bar{\psi}\psi = -m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) \quad (2.28)$$

and are therefore not invariant under $SU(2)_L$ transformations. But fermion masses clearly exist in nature, so their inclusion in the SM as well as weak gauge boson masses is imperative. The mechanism for doing so in a gauge-invariant way is the subject of the next section.

2.2.3 The Brout-Englert-Higgs mechanism

The SM gauge theory of weak interactions hinges on a field-theoretic mechanism revealed through the overlapping work of Brout together with Englert and of Higgs in a series of 1964 papers [16, 17, 18]. What is often called the *BEH mechanism* explains how gauge bosons can acquire non-zero mass in systems which undergo gauge symmetry breaking. Note that the acronym “BEH” fails to acknowledge the work done in parallel by Guralnik, Hagen, and Kibble [19] also in 1964. Moreover, core ideas behind the BEH mechanism were introduced already in 1963 by Anderson [20] in the context of superconductivity.

In the QED Lagrangian (2.12), the demand of local $U(1)$ gauge invariance is satisfied because of the added gauge field kinetic and interaction terms (in that case, of the photon). This procedure also works in QCD for local $SU(3)$ invariance where the corresponding gluon fields are likewise massless. However, when attempting to couple fermion fields to the weak gauge bosons, the term $\propto m^2 W_\mu W^\mu$ which enters as part of the *Proca Lagrangian* describing a vector field, does not vanish as for the photon or gluon. The problem is that this mass term spoils gauge invariance under the transformation (2.7). In fact, a similar problem is encountered by the fermion mass terms ($\propto \bar{\psi}\psi$) since they are not invariant under local $SU(2)$ transformations. The nonzero masses of leptons and the expected large masses of the weak force mediators therefore presented a dilemma in the early developments of the SM. The solution was realized by incorporating the BEH mechanism in a unified model of electroweak interactions for which Weinberg, Glashow, and Salam (GWS) were later awarded the 1979 Nobel Prize. Their work predicted the existence of a neutral current associated with the weak interaction which was observed at CERN in 1973 [21]. About a decade later the Z^0 boson mediating this neutral current was also discovered [6]. The GWS theory uses the BEH mechanism to explain why the weak gauge bosons have large masses and how the photon ends up being massless while still deriving from the same electroweak symmetry group.

Spontaneous symmetry breaking

The unitarity (and thus validity) of the electroweak gauge theory requires that its Lagrangian exhibit $SU(2) \times U(1)$ symmetry, but it turns out there is an auxiliary way to

realize a breaking of this symmetry in order to accommodate the mass terms. Spontaneous symmetry breaking occurs when a system's ground state breaks a symmetry exhibited by its Lagrangian. The Higgs field ϕ plays precisely this role in the SM where it is introduced as a weak isospin doublet of complex scalar fields:

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \text{Re } \phi^+ + i\text{Im } \phi^+ \\ \text{Re } \phi^0 + i\text{Im } \phi^0 \end{pmatrix} \quad (2.29)$$

where the superscripts denote the electric charge of the components $Q = 1/2 \pm 1/2$. The SM Higgs field therefore comprises four degrees of freedom. It is constrained by a “Mexican hat” potential $V(\phi)$ which, as required, is manifestly $SU(2) \times U(1)$ symmetric:

$$\mathcal{L}_{\text{Higgs}} = (D^\mu \phi)^\dagger D_\mu \phi - V(\phi) , \quad V(\phi) \equiv -\mu^2 \phi^\dagger \phi + \frac{\lambda^2}{2} (\phi^\dagger \phi)^2 \quad (2.30)$$

for some real parameters μ and λ . It is possible to visualize the shape of $V(\phi)$ by plotting it on the complex plane of only one of the fields in the doublet, as shown in figure 2.1. Then the symmetry about the z -axis represents the group $U(1)$ of complex phase rotations.

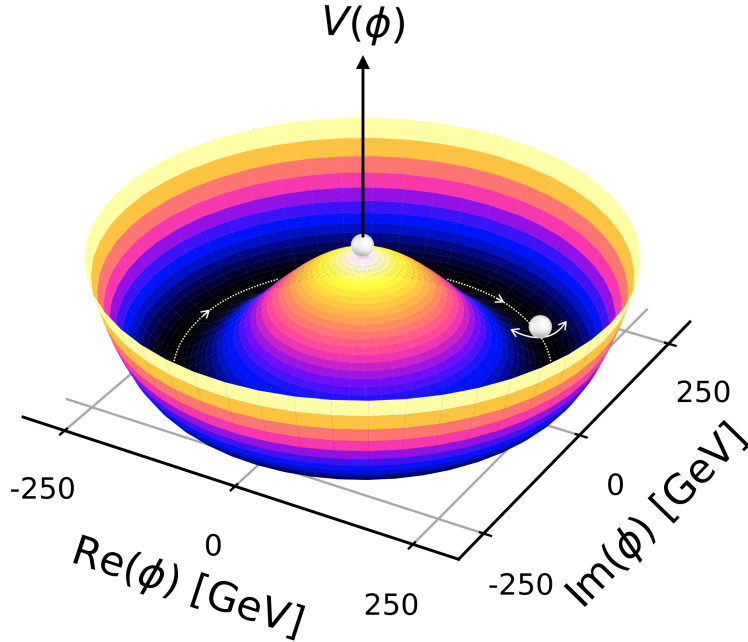


Figure 2.1: Plot of the “Mexican hat” potential for a single complex scalar field ϕ . The local maximum at the origin is clearly unstable against small perturbations and thus gives way to a nonzero vacuum configuration at some point along the circular gutter. For the SM it resides at $v = 246$ GeV as shown.

For a gravitational analogy, the ball placed on the z -axis in the figure shows how $\phi = 0$ is clearly not a stable vacuum configuration. Instead, responding to quantum fluctuations, the field will equilibrate to a nonzero vacuum expectation value (VEV) *somewhere* along

the circular gutter with radius $v \equiv \mu/\lambda$ (shown by the second ball). So while the potential itself is constructed $U(1)$ -symmetric, any particular choice of groundstate spontaneously breaks this symmetry! The two previously identical degrees of freedom now correspond to the two quite different excitations – radial and angular – about the nonzero VEV which are indicated in figure 2.1. The radial excitations engage the potential gradient and therefore give rise to a massive particle, the *Higgs boson*. Angular excitations on the other hand are potential-invisible and correspond to a massless *Nambu-Goldstone boson*. The one generator of the broken $U(1)$ group in this case gives rise to a single degree of freedom which, while massless in itself, now provides the necessary “slot” for a gauge field mass term as we shall see.

Extending this idea to the full SM Higgs doublet, the degenerate set of potential minima define a sphere of radius v in $\mathcal{R}^4$. As in the lower-dimensional illustration, a vacuum configuration anywhere on this boundary can be expressed without loss of generality in terms of a single nonzero (real) field component. Likewise the four degrees of freedom decompose into excitations $h(x)$ in the single radial direction along the VEV and in three angular directions $\xi^a(x)$ that are tangent to the 4-sphere. The Higgs field after spontaneous symmetry breaking can therefore be reparameterized as follows:

$$\phi(x) = \frac{1}{\sqrt{2}} e^{i \frac{\xi^a(x) \sigma^a}{v}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (2.31)$$

where σ^a are the Pauli matrices. When (2.31) is substituted back into the potential (2.30) a term arises that is quadratic in h , indicating that the radial excitations again correspond to a massive particle associated with the Higgs field. No such terms arise for the three ξ^a which cancel out in the norm-squared terms suggesting their interpretation as massless Goldstone modes. In fact they can be fully absorbed at every point in spacetime by specifying the three local gauge functions ω^a in (2.6) representing $SU(2)_L$ invariance. But this specific choice of *unitary gauge* implies that the SM electroweak symmetry has been broken:

$$SU(2)_L \times U(1)_Y \rightarrow U(1)_Q \quad (2.32)$$

where Q is the electric charge which is now conserved by the residual symmetry associated with the electromagnetic interaction (the unitary gauge did not fix this remaining generator). Meanwhile, the three degrees of freedom that were “gauged away” show up in the gauge sector through the covariant derivative in the kinetic term of (2.30):

$$(D^\mu \phi)^\dagger D_\mu \phi = \frac{1}{2} \left| \begin{array}{c} -ig_w T(W_\mu^1 - iW_\mu^2)(v+h) \\ \partial_\mu h - i(g' Y B_\mu - g_w T W_\mu^3)(v+h) \end{array} \right|^2 \quad (T=Y=1/2) \quad (2.33)$$

$$\supset \frac{1}{2} (\partial^\mu h) \partial_\mu h + \frac{1}{8} \left[(g'^2 + g_w^2) Z^\mu Z_\mu + 2g_w^2 W^{+\mu} W^-_\mu \right] (v+h)^2 \quad (2.34)$$

where

$$W_\mu^\pm \equiv \frac{(W_\mu^1 \mp iW_\mu^2)}{\sqrt{2}}, \quad Z_\mu \equiv \frac{g'B_\mu - g_w W_\mu^3}{\sqrt{g'^2 + g_w^2}}, \quad A_\mu \equiv \frac{g'B_\mu + g_w W_\mu^3}{\sqrt{g'^2 + g_w^2}} \quad (2.35)$$

These linear combinations of the electroweak gauge fields correspond to the observable $W^\pm$ and Z_0 weak bosons and the photon. The photon field A_μ accounts for the remaining degree of freedom orthogonal to Z_μ .

Massive bosons

As a result of spontaneous symmetry breaking there are terms in (2.34) quadratic in the weak boson fields and proportional to the constant v – these explain the masses of the weak bosons:

$$m_W = \frac{1}{2}g_w v, \quad m_Z = \frac{m_W}{\cos \theta_W}, \quad \tan \theta_W \equiv \frac{g'}{g_w} \quad (2.36)$$

where the SM electroweak VEV is $v = 246$ GeV and the *Weinberg angle* is $\theta_W = 28.7^\circ$, corresponding to masses of 80.4 GeV and 91.2 GeV for the W and Z bosons, respectively. On the other hand, because of the charge assignment to the Higgs doublet (2.29) the photon does not get a mass term in the expansion (2.34). In summary, the three Goldstone modes from the spontaneous symmetry breaking provide the longitudinal (i.e. massive) polarization states of the three weak bosons while leaving the photon massless.

Equation 2.34 also entails 3-point and 4-point interaction terms between the weak bosons and the massive Higgs field h . The Higgs boson mass $m_H \equiv \sqrt{2}\mu$ is a free parameter in the potential (2.30) and has been experimentally determined at the LHC as 125.10 ± 0.14 GeV [11].

Fermion masses

The BEH mechanism can also introduce mass terms for fermions by coupling them to the Higgs field. An additional scalar $SU(2)_L$ doublet is just what is needed to connect left- and right-chiral fermion fields in a gauge-invariant way. This can be seen in the following adaptation of the mass term (2.28):

$$\text{Yukawa sector: } -y_f \left(\bar{\psi}_R \phi^\dagger \Psi_L + \bar{\Psi}_L \phi \psi_R \right) \quad (2.37)$$

where y_f is the *Yukawa coupling* of the fermion species f ; Ψ and ψ are the doublets and singlets, respectively, defined in equations (2.24) and (2.25) for leptons and quarks. Because the Higgs doublet contracts with the left-handed fermion doublets this mass-type term is $SU(2)$ gauge invariant, as desired. The Higgs vacuum configuration (2.31) then precipitates the following terms:

$$-y_\ell \bar{\ell}_R \phi^\dagger L = -\frac{y_\ell}{\sqrt{2}} \bar{\ell}_R \begin{pmatrix} 0 & v+h \end{pmatrix} \begin{pmatrix} \nu_\ell \\ \ell_L \end{pmatrix} = -\frac{y_\ell}{\sqrt{2}} (v+h) \bar{\ell}_R \ell_L \quad (2.38)$$

When put together with its adjoint the expression on the right produces a mass term $m\bar{\psi}\psi$, where the mass includes a factor of the lepton’s particular Yukawa coupling y_ℓ and the Higgs VEV v . The neighboring term with h in it represents a coupling between fermions and the dynamical Higgs field after spontaneous symmetry breaking. It also is proportional to y_ℓ , leading to the SM prediction that the Higgs boson will couple to fermions in direct proportion to their mass. The values of the Higgs boson and fermion masses are not, however, determined within the SM, suggesting one way in which new theories might be able explain more (chapter 3).

2.3 Quantum chromodynamics

There is an additional fundamental interaction between certain particles which, like the weak nuclear force, only emerges on a small length scale: the strong nuclear force. The quantized charge associated with this interaction is *color charge*, hence the terms *chromodynamics*. In the SM, color charge is carried only by quarks and the gauge boson of the strong force, the gluon. The notion of color is invoked solely as a helpful analogue for the way these charges combine and has its origins in the series of meson and baryon discoveries from 1947 into the 1960s. Gell-Mann and Zweig’s organization of these particles into the *Eightfold Way* formed the basis for a theory of hadrons as bound states of light fermions, ushering in the quark model. An additional quantum number was then needed to explain how baryons could contain more than one otherwise identical quark states while respecting the Pauli exclusion principle. The wavefunctions of quarks were therefore extended by an additional degree of freedom representing its color state in the red, green, and blue eigenvectors:

$$\text{quarks : } r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (2.39)$$

$$\text{antiquarks : } \bar{r} = r^T, \quad \bar{g} = g^T, \quad \bar{b} = b^T \quad (2.40)$$

The need for a separate basis of colors – anticolors – carried by antiquarks is tied to the principle that bound states of QCD are always color-neutral (i.e. singlets): in mesons a color cancels with its anticolor while in baryons three distinct colors combine to form “white”. The absence of qq bound states follows from the property that no two colors can cancel. The quark color triplet is in the $\mathbf{3}$ fundamental representation of the $SU(3)$ group while antiquarks are described by the $\bar{\mathbf{3}}$ representation. Gluons on the other hand correspond to the adjoint representation of $SU(3)$, rotating quark and antiquark states through its 8 of generators λ^k . As shown in table 2.2 they are the complex 3×3 Gell-Mann matrices.

The Lagrangian for QCD is nearly identical in form to that of QED in equation 2.11. However as in the weak gauge theory there are structural differences in both the fermion

fields and the gauge fields. The quark fields are packaged in color triplets corresponding to (2.39) and the matrix-valued generators of $SU(3)$ are implicit within the gluon gauge fields $A_\mu = A_\mu^k \lambda^k$. The latter point again leads to additional commutation terms in the gluon field strength tensor $G_{\mu\nu}^k$ compared to the electromagnetic one. Ultimately the contraction $G_{\mu\nu}^k G^{k\mu\nu}$ therefore describes not only gluon kinetic terms but also self-interaction terms. Though not unlike weak interactions in this regard, QCD produces a very different phenomenology because gluons are massless and couple strongly as discussed in section 2.3.1 below. These same features enable the different quark types to combine in a variety of QCD bound states comprising the mesons and baryons. The most important among them are the nucleons that make up baryonic matter throughout the universe. Also, while the BEH mechanism is responsible for fundamental masses, it is the nuclear binding energy arising from strong force that generates the vast majority of atomic mass, and hence that of ordinary matter (dark matter is another question).

2.3.1 Asymptotic freedom and quark confinement

The fact that the gluon is charged under the same force it mediates is a consequence of the non-Abelian nature of the $SU(3)$ group – that is, its generators do not commute. This leads to self-interaction of gluons in three- and four-point diagrams. While such diagrams also exist for the weak bosons they are suppressed by factors of the weak coupling constant $g_w \sim 0.6$ and more importantly propagators with large mediator masses. To the contrary, the strong coupling constant of QCD rises to values near 1 for low energy (i.e. *soft*) interactions. This means that additional diagrams corresponding to higher-order QCD corrections are not necessarily small compared to the tree-level process. At the lowest energies perturbative methods therefore become unsuitable for QCD calculations. However the strong coupling “constant” α_s weakens for higher energy interactions, enabling perturbative calculations of QCD cross sections at colliders. This dependence on energy scale Q^2 is shown in figure 2.2 and has become known as the “running” of the strong coupling constant. It can equivalently be said that the strong force is strongest at large distance scales but weakens for short-range interactions.

The evolution of α_s with energy scale Q is generated by higher-order QCD corrections which themselves include factors of α_s . In the ultraviolet regime where perturbation theory is applicable this interdependence is modelled by the following renormalization group equation (RGE):

$$\frac{\partial}{\partial(Q^2)} \left(\frac{\alpha_s}{4\pi} \right) = -\frac{1}{Q^2} \left(\frac{\alpha_s}{4\pi} \right)^2 \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{4\pi} \right)^n \quad (2.41)$$

The β_n are constant expansion coefficients of the *Beta function*, corresponding to $(n+1)^{\text{th}}$ -loop order corrections. For example, the 1-loop correction is $\beta_0 = 11 - \frac{2}{3}n_f$, in terms of the number of quark flavours included in the calculation. The expansion has recently been carried to 5-loops by calculation of the β_5 coefficient [22].

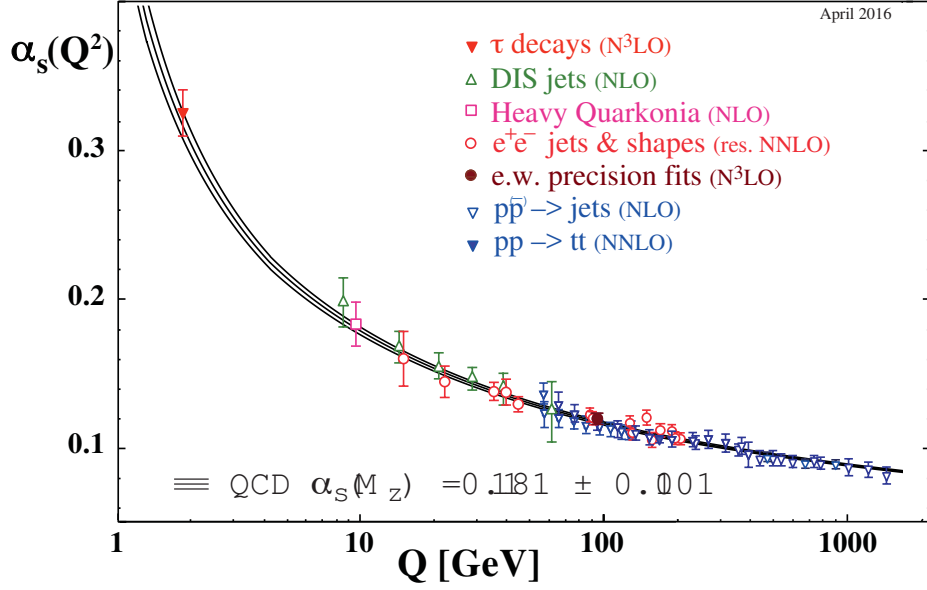


Figure 2.2: Evolution (a.k.a. “running”) of the strong coupling constant with energy scale Q as shown by a fit to experimental results. The precision of the calculations used is written in brackets. This is figure 9.3 of [23].

The running of the strong coupling constant leads to important implications for the behaviour of quarks. In the limit of high energy interactions, quarks exhibit *asymptotic freedom* as they interact with smaller values of the strong coupling constant. Asymptotic freedom essentially follows from the minus sign in the RGE (2.41). The reverse is true for quarks interacting at low energies. These two regimes can be modelled in an approximate form of the color singlet (e.g. $r\text{-}\bar{r}$) potential derived from perturbative calculations at small distances (high energy) and lattice QCD beyond the perturbative regime [24]:

$$V_{\text{QCD}}(r) = -\frac{4}{3} \frac{\alpha_s(r)}{r} + kr + \mathcal{O}(\mu_f^3 r^2) \quad (= V_{\text{pert.}} + V_{\text{latt.}}) \quad (2.42)$$

where μ_f is the fragmentation scale. The potential is roughly Coulombic at small distances (ignoring r -dependence of α_s) but is dominated at larger distances by the term which increases linearly with separation with effective “spring constant” k . This explains why the strong force, despite its massless mediator, is nonetheless confined to distance scales on the order of a proton radius ($\sim \text{fm}^{-1}$). A picture for what actually takes place is that as two quarks are drawn apart, their color potential gradient becomes steep enough to precipitate a $q\bar{q}$ pair out of the quantum vacuum. The new pair then splits off to form two new bound states with the separated quarks. This process may repeat depending on the energy of the

quarks (discussed further in section 2.5.2), but the upshot is that the original quarks are not experimentally accessible apart from their confinement within bound states. The top quark is the only exception because it decays via the weak interaction before it has time to create a bound state with another quark (prompted by the large difference in top-bottom mass).

Asymptotic freedom is confirmed by the observed decrease of α_s with energy scale Q deduced from experimental measurements as shown in figure 2.2. For example, the “DIS jets” points are based on combined measurements of jet cross sections in deep inelastic $e-p$ scattering at high- Q^2 [25] where the partons are treated as free particles. On the other hand, quark confinement is empirically supported by the fact that experiments have never managed to observe quarks in isolation, even at colliders probing energies well above the confinement scale Λ_{QCD} . For instance, quarks produced in collisions at the LHC cannot be detected individually, but only through their final state signature of a collimated stream of hadrons known as a *jet*.

2.4 High energy physics toolkit

2.4.1 Particle interactions and cross sections

In general a particle interaction is a process converting a set of initial state particles into a set of final state particles over a time interval. In Feynman diagrams these *external* particles are represented by the open-ended lines while their interconnecting vertices and internal lines comprise intermediate states which strictly cannot be observed directly. As such the kinematics of the internal lines, called propagators, are not restricted to classical behaviour.

The probability of observing any given particle interaction depends on the external particles’ four-momenta and quantum numbers together with the intrinsic quantum-mechanical probability of the process. The *Feynman amplitude* $\mathcal{M}$ incorporates the external particle states in connection with the coupling strengths of the relevant fundamental interaction (via *vertex factors*) and phase space integrals over propagators. The *Feynman rules* (§A.1 of [26]) provide the recipe for calculating this probability for a given diagram. Considering a process with two initial particles (momenta $p_1 + p_2 \rightarrow \Sigma p_f$) the probability of producing the final state particles in an infinitesimal phase space volume $d^3\mathbf{p}_f$ is calculated as follows:

$$d\sigma = (2\pi)^4 \delta^{(4)}\left(p_1 + p_2 - \sum_f p_f\right) \frac{1}{|v_2 - v_1|} \prod_{i=1,2} \frac{m_i}{E_i} \prod_f \left[\frac{m_f}{E_f} \frac{d^3\mathbf{p}_f}{(2\pi)^3} \right] |\mathcal{M}|^2 \quad (2.43)$$

This is the *differential cross section* of the process. The Dirac delta function simply enforces 4-momentum conservation (so that the E_f are fully determined by $\mathbf{p}_f$). There is a frame-dependent factor in the denominator of the relative velocity of the incoming particles. In

the two most common frames it is:

$$|v_2 - v_1| = \begin{cases} |\mathbf{p}_1 = -\mathbf{p}_2| \cdot (E_1 + E_2)/(E_1 E_2) , & \text{(CoM)} \\ |\mathbf{p}_1 \neq 0|/(E_1) , & \text{(Lab)} \end{cases} \quad (2.44)$$

The Feynman amplitude generally depends on the spin states of the external particles. However, in practice the spin states of the incoming particles (i.e. beam polarization) are often not known and those of the outgoing particles are not always relevant to the experiment. In such cases the quantity of interest is the cross section both averaged over initial spin states and summed over final spin states. Similarly, interactions between quarks and gluons entail sums and averages over possible color states of the particles.

Because internal particle lines are not observable, all possible configurations must be included in order to calculate the full cross section between a given initial and final state configuration (which *is* observable). The lowest-order set of indistinguishable Feynman diagrams associated with a given interaction is commonly referred to as the *Born level*⁵. Diagrams which share an external topology with the Born-level but have n additional pairs of internal vertices are suppressed by n factors of $g^2 = 4\pi\alpha$. Therefore calculations at the Born-level are called *leading-order* (LO) and higher order corrections next-to-leading-order (NLO), next-to-next-to-leading-order (NNLO), and so on.

Although higher-order effects are often small, they can have an important impact in the form of interference. Interference is especially noticeable when diagrams are indistinguishable at leading order (e.g. γ and Z in Drell-Yan). In general, indistinguishable processes, say A and B , must be summed at the amplitude level: $\mathcal{M}_{A+B} = \mathcal{M}_A + \mathcal{M}_B$. Because it goes like the norm-squared amplitude, the corresponding cross section for $A + B$ picks up an additional contribution from the cross-term:

$$\sigma_{A+B} \sim |\mathcal{M}_{A+B}|^2 = |\mathcal{M}_A|^2 + |\mathcal{M}_B|^2 + 2 \operatorname{Re} \{ \mathcal{M}_A^* \mathcal{M}_B \} \quad (2.45)$$

The result is observable interference effects in the differential cross section which are constructive or destructive depending on the sign of the cross term.

2.4.2 Particle colliders

Many parameters are needed to fully describe a high energy particle collider, but the two most defining ones are its beam energy and *luminosity* which will be introduced here. A collider includes the aspect of a particle accelerator as well as the beam manipulation needed to achieve collisions between the accelerated particles.

⁵This popular term apparently combines the ideas of “Born approximation” and “tree level”, but in some cases “Born-level” also refers to diagrams higher than tree level (e.g. $H \rightarrow \gamma\gamma$).

Charged particles can be accelerated by means of the electric potential gradients generated inside resonators, in particular radio-frequency (RF) cavities. Particle accelerators either arrange these RF cavities along open (e.g. linear accelerators) or closed (e.g. circular) paths. The latter option has the advantage of reusing the same set of RF cavities multiple times throughout the acceleration of a single beam. However, maintaining a circular beam trajectory requires in addition a strong magnetic field for acceleration in the transverse direction (bending). There are two main types of circular accelerators: the cyclotron and the synchrotron. In cyclotrons, both the electric and magnetic fields used to accelerate/bend the beam permeate a common volume inside a cylindrical evacuated vessel. The beam is injected at the centre and spirals outward, being accelerated in the gap region between two “D”-shaped electrodes of opposite phase. The energy achievable with a cyclotron is limited by the need to sustain both a strong magnetic field and a high vacuum across the entire extent of the spiraling beam. Synchrotrons on the other hand opt to maintain a constant beam radius while sweeping both the electric field radio frequency and the magnetic field strength⁶. In this case the magnetic field only needs to be present at small bending corridors along an approximately circular beamline. This aspect make it possible for synchrotrons to achieve far higher beam energies than cyclotrons. The angular velocity of a charged (q) particle’s orbit with radius R lying perpendicular to a magnetic field B is

$$\omega = \frac{v}{R} = \frac{Bq}{m\gamma} \iff |\mathbf{p}| = \gamma m v = R B q \quad (2.46)$$

where $\gamma = 1/\sqrt{1-v^2}$. The second equation above implies that beam momentum (and hence beam energy in the ultrarelativistic limit) is proportional to both the magnetic field strength and radius of the accelerator. Even with state-of-the-art superconducting magnet technology, equation 2.46 implies the need for very large radii to probe physics at the TeV scale. For example, the dipole magnets at the LHC are designed to operate at 8.33 TeV, which for a design energy of $|\mathbf{p}| = 7 \text{ TeV}/c$ corresponds to a bending radius of 2.8 km [27].

Once accelerated to high energy, by far the most straightforward way to study particle interactions is to steer the beam into a stationary/fixed target. However this configuration is inefficient in terms of converting beam energy E into centre of mass energy $\sqrt{s}$ of interactions between beam and target particles. The following expressions motivate why, though technically challenging, it is highly advantageous to instead steer two beams into collision course:

$$\sqrt{s} = \begin{cases} \sqrt{2m_t E} , & \text{(fixed-target)} \\ \sqrt{2E_1 E_2 (1 + \cos \theta)} \Big|_{E_1=E_2, \theta=0} = 2E , & \text{(collider)} \end{cases} \quad (2.47)$$

⁶In cyclotrons it is the radio frequency that remains constant to first order, although it also must ramp to account for relativistic effects.

where m_t is the target particle mass for the first case, and θ is the (small) crossing angle of beams 1 and 2 in the second case. While the collision energy for a collider experiment is up to double the collision energy ($2E$), for a fixed-target experiment the collision energy goes like the square-root of $2E$.

Getting a beam of incredibly small particles to collide in an experiment involves not only accelerating the beam but also highly focusing it. Both the cross section σ of the particle interaction (dependent on the collision energy) and the beam luminosity $\mathcal{L}$ factor into the rate dN/dt one can expect this process to occur in the experiment:

$$\frac{dN}{dt} = \sigma \mathcal{L} \quad (2.48)$$

The luminosity measures both the number of colliding particles and their proximity achieved by compression of the beam both in space and time (i.e. frequency of collisions). Two incident bunches⁷ of particles with (transverse) density ρ_A and ρ_B that intersect in a cross section $\mathcal{A}$ produce the following luminosity:

$$\mathcal{L} = f_{\text{col}} \int_{\mathcal{A}} \rho_A \rho_B dx dy \quad (2.49)$$

where f_{col} denotes the bunch collision frequency. This expression assumes a uniform density in the longitudinal direction for each of the two colliding bunches. Considering now the collision of two identical beams with density profiles that are gaussian with width σ_x and σ_y , the integral results in

$$\mathcal{L} = f_{\text{col}} \chi \frac{N^2}{4\pi\sigma_x\sigma_y} \quad (2.50)$$

where N is the number of particles per bunch and $\chi \leq 1$ is a geometric reduction factor accounting for a crossing angle less than 180° .

2.4.3 Invariant mass in terms of transverse and angular observables

The *invariant mass* associated with a two-body decay is defined as the magnitude of the total final state four-momentum:

$$\begin{aligned} m_{\ell\ell}^2 &= P_\mu P^\mu = (P_1 + P_2)^\mu \eta_{\mu\nu} (P_1 + P_2)^\nu \\ &= (E_1 + E_2)^2 - |\vec{p}_1 + \vec{p}_2|^2 \end{aligned} \quad (2.51)$$

where labels 1 and 2 identify the two final state particles (in our case leptons). The invariant mass is most intuitive when both particles are produced with equal and opposite 3-momentum vectors in the lab frame; then the invariant mass becomes simply the sum of

⁷In the case of a collider these are opposite beams but a similar argument holds for fixed-target experiments in the collision center-of-mass frame.

their total energies. A more practical expression can be derived by converting the lepton energies and 3-momenta into the basis of observable quantities at the LHC: transverse momentum p_T , radial angle ϕ , and pseudorapidity η . This proceeds as follows from equation 2.51:

$$\begin{aligned}
m_{\ell\ell}^2 &= (E_1 + E_2)^2 - |\vec{p}_1 + \vec{p}_2|^2 \\
&= E_1^2 + E_2^2 + 2E_1E_2 - |\vec{p}_1|^2 - |\vec{p}_2|^2 - 2\vec{p}_1 \cdot \vec{p}_2 \\
&= m_1^2 + m_2^2 + 2(E_1E_2 - \vec{p}_1 \cdot \vec{p}_2) \\
&= m_1^2 + m_2^2 + 2 \left[E_1E_2 - p_1^z p_2^z - |\vec{p}_1^T| |\vec{p}_2^T| \cos(\phi_1 - \phi_2) \right]
\end{aligned} \tag{2.52}$$

The first two terms in the brackets can be expressed in terms of the leptons' rapidity, defined as:

$$\cosh y \equiv \gamma \implies y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{2.53}$$

The second equation is a deduction for the case where we are considering a frame boosted in the z -direction, e.g. colliding protons. Then,

$$\begin{aligned}
2(E_1E_2 - p_1^z p_2^z) &= (E_1 + p_1^z)(E_2 - p_2^z) + (E_1 - p_1^z)(E_2 + p_2^z) \\
&= \sqrt{(E_1 + p_1^z)(E_2 - p_2^z)(E_1 - p_1^z)(E_2 + p_2^z)}
\end{aligned} \tag{2.54}$$

$$\begin{aligned}
&\times \left[\sqrt{\frac{(E_1 + p_1^z)(E_2 - p_2^z)}{(E_1 - p_1^z)(E_2 + p_2^z)}} + \sqrt{\frac{(E_1 - p_1^z)(E_2 + p_2^z)}{(E_1 + p_1^z)(E_2 - p_2^z)}} \right] \\
&= \sqrt{E_1^2 - (p_1^z)^2} \sqrt{E_2^2 - (p_2^z)^2} \left(\sqrt{e^{2y_1} e^{-2y_2}} + \sqrt{e^{-2y_1} e^{2y_2}} \right) \\
&= \sqrt{m_1^2 + p_1^T^2 - (p_1^z)^2} \sqrt{m_2^2 + p_2^T^2 - (p_2^z)^2} \left(e^{y_1 - y_2} + e^{-y_1 + y_2} \right) \\
&= \sqrt{m_1^2 + (p_1^T)^2} \sqrt{m_2^2 + (p_2^T)^2} \cdot 2 \cosh(y_1 - y_2)
\end{aligned} \tag{2.55}$$

So, substituting (2.55) back into (2.52), the invariant mass is now expressed fully in terms of rapidity, radial angle, and the component of momentum perpendicular to the beam axis:

$$m_{\ell\ell}^2 = m_1^2 + m_2^2 + 2p_1^T p_2^T \left[\sqrt{\left(\frac{m_1}{p_1^T} \right)^2 + 1} \sqrt{\left(\frac{m_2}{p_2^T} \right)^2 + 1} \cdot \cosh(y_1 - y_2) - \cos(\phi_1 - \phi_2) \right] \tag{2.56}$$

As discussed in section 4.1, the pseudorapidity η is used in ATLAS as a proxy for the rapidity y since these two are equal in the ultrarelativistic limit. In the case that both outgoing particles are of the same species, the masses m_1 and m_2 are equal. Further, ultrarelativistic particles satisfy $m/p \ll 1$ (in natural units) such that their mass can be safely neglected. In particular, muon and electron candidates in the high-mass dilepton analysis have $p_T \geq 30 \text{ GeV}$ such that the $(m/p_T)^2$ term under the radical in (2.56) evaluates to 2.9×10^{-10} and 1.2×10^{-5} , respectively. The contributions from nonzero electron and muon rest masses

in (2.56) are therefore far below the invariant mass resolution. With these simplifications the invariant mass becomes

$$\lim_{E \gg m} m_{\ell\ell}^2 = 2p_1^T p_2^T [\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)] \quad (2.57)$$

which is the form used in the dilepton analysis discussed in this thesis.

2.5 Simulation of ATLAS data

There are many ways in which doing physics at the LHC depends on large samples of accurately simulated proton-proton collisions. Comparing real and simulated data is the basis for understanding detector efficiencies, optimizing analysis strategies, and testing theories such as the SM. The fundamental advantage of simulating events is that the underlying process is fully specified and is therefore also known after reconstruction. This is unlike real data where information about the underlying process must be inferred based on an event's properties, and may be limited by the rules of quantum mechanics (e.g. for indistinguishable Feynman diagrams). So, using the detailed mathematical description of particle interactions in quantum field theory, one can calculate the cross sections and hence production rates of processes taking place at the LHC (using equations 2.43 and 2.48). However, to simulate data that matches the form collected by LHC experiments, calculating cross sections is only one link in a chain of a highly involved modelling workflow.

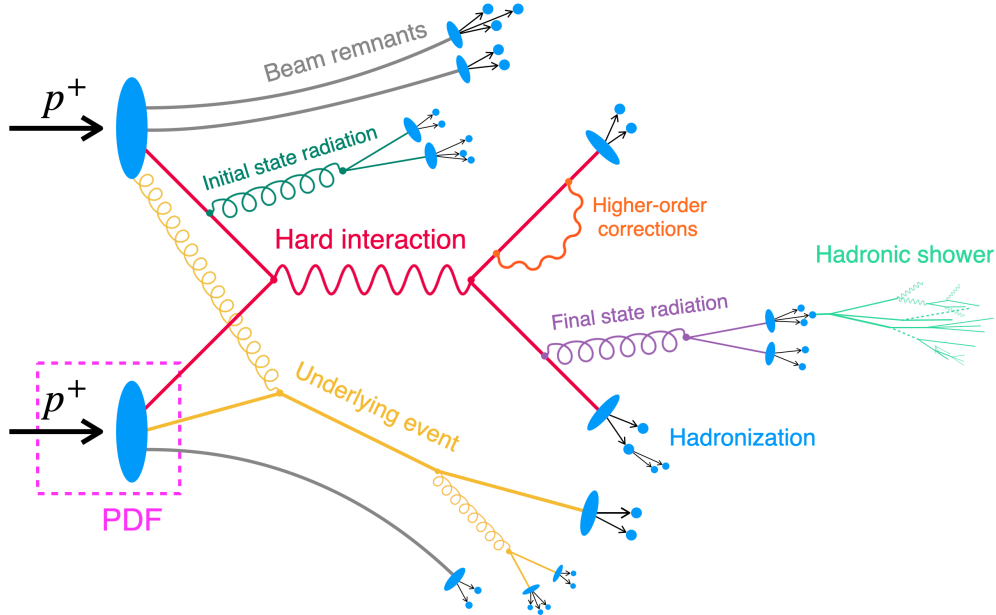


Figure 2.3: Illustration of the different intermediate components of a pp collision.

The so-called *hard interaction* denotes the process at the heart of a pp collision of interest. However, as shown in figure 2.3, the hard interaction is overlaid with a number of interactions which influence its initial and final state kinematics and add to its detector signature. To start with, at a hadron collider the initial state can only be modelled probabilistically through the Parton Distribution Function (PDF), discussed in section 2.5.1. Next, the hard interaction cross section is modified by higher-order corrections, and radiative emissions from its initial and final fermion lines. Then the final state and spectator partons undergo the non-perturbative process of hadronization.

Each of the effects just mentioned is modeled at the stage of *event generation* denoted by the box in figure 2.4. After that stage, all the particle-level information propagates through intermediate stages of interactions with the ATLAS detector before being converted into digital format. The process must also take into account that in the real detector, multiple proton collisions overlap from within the same bunch crossing and also from neighboring ones (pileup). The challenge of simulating this entire picture with sufficient accuracy is the aim of the ATLAS Monte Carlo production scheme. An overview of the different components to be discussed is shown in figure 2.4.

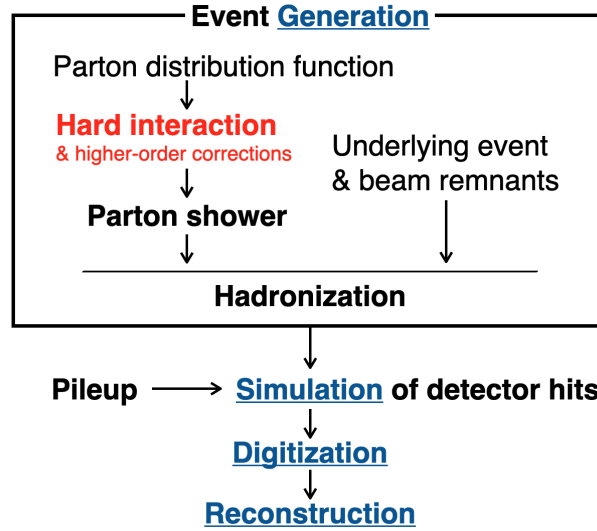


Figure 2.4: Main components of the ATLAS MC workflow, outlined by Generation, Simulation, Digitization, and Reconstruction. See [28] for a more detailed flowchart.

2.5.1 Parton distribution function

The aim in colliding protons is that two of their partons will collide with sufficient energy to achieve some process of interest. Calculating the likelihood of this occurring is complicated by the fact that each of the two partons possesses a different, indeterminate fraction x of the proton's total momentum. In a simplistic picture of the proton's structure composed of three valence quarks, it might be supposed that each parton possesses about one-third

of the proton's total momentum. A closer figure of $\sim 1/6$ could also be intuitively argued to account for momentum-carrying gluons present in the proton. However, deep inelastic scattering experiments support a more complex picture where gluon fields undergo random splitting into quark-antiquark pairs, generating a stochastic “sea” of quarks that accompany the valence uud . The fraction of a proton's momentum which is carried by a given species of quark therefore takes values according to a probability distribution. The probability to find a given species of parton i possessing a fraction of the proton momentum in the interval $[x, x + dx]$ is given by the corresponding moment $x \cdot f_i(x, Q^2)$ of its Parton Distribution Function (PDF). The components carried by all possible quarks, antiquarks, and gluons therefore together satisfy [26]:

$$\sum_i \int_0^1 x f_i(x, Q^2) dx = 1 \quad (2.58)$$

where the sum is over all parton species i , including gluons. The PDFs derived by the NNPDF collaboration are shown in figure 2.5. The second argument of the PDF, Q^2 , denotes dependence of the proton structure on the scale of momentum transfer between the probe particle and the proton, for example in the context of deep inelastic scattering experiments.

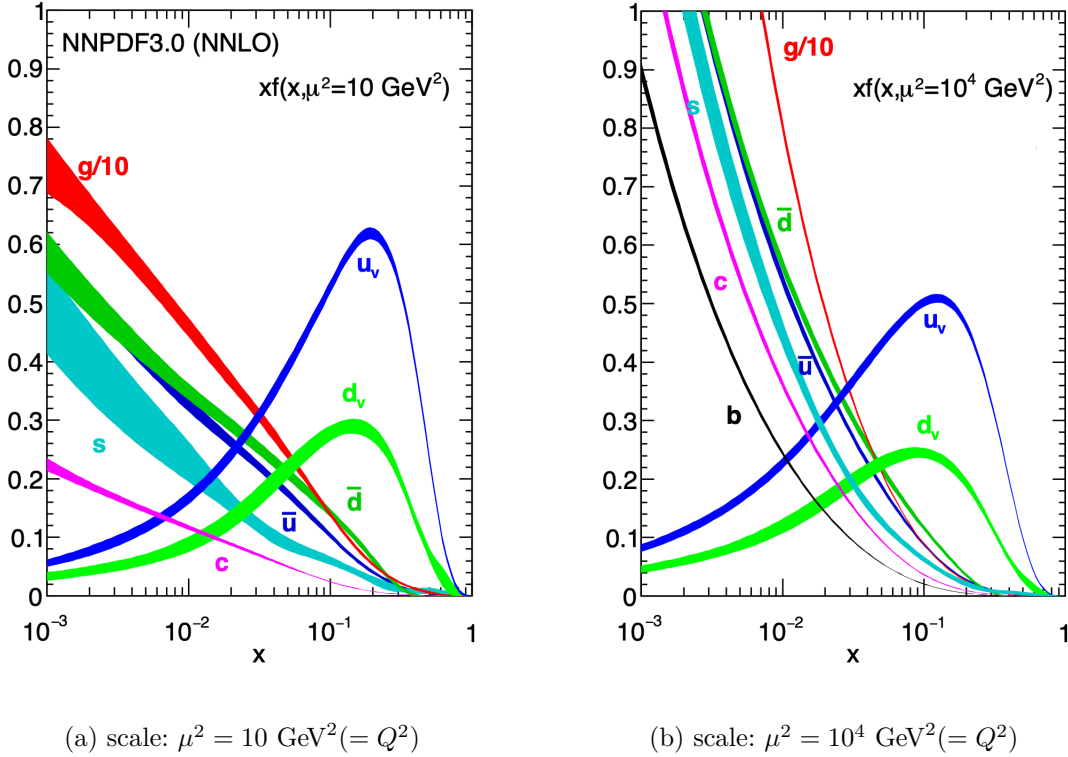


Figure 2.5: Plots of the parton momentum fraction moments $x \cdot f_i(x, Q^2)$ for $Q^2 = 10 \text{ GeV}^2$ using the NNPDF3.0 set [29]. Figure adapted from [23].

The inherently dynamic composition of the proton is both an essential feature and a challenge of pp collisions at the LHC. A proton collider operating at a fixed beam energy has access to multiple initial state combinations involving gluons, quarks, and antiquarks at a range of x values. This is much more varied than the initial states produced by, say, ee collisions within their radiative corrections and beam energy uncertainty. This feature makes the LHC an excellent facility to search for a wide range of processes and has enabled the discovery of the Higgs boson in multiple production modes. The downside is that direct knowledge of the initial state kinematics is limited by the PDF which itself is known with uncertainty. The cross section of a given production process on the partonic level, $ij \rightarrow X$, must therefore be integrated over the range of possible momentum fractions x_A and x_B of the protons A and B , weighted by their respective PDFs:

$$\sigma_{pp \rightarrow X} = \sum_{i,j} \int_0^1 \int_0^1 f_i(x_A, Q^2) f_j(x_B, Q^2) \sigma_{ij \rightarrow X}(x_A p_A^\mu, x_B p_B^\mu) dx_A dx_B \quad (2.59)$$

where $\sigma_{ij \rightarrow X}$ is the partonic cross section involving physics of the hard scatter and, in general, subsequent parton shower and hadronization. This scheme is known as *factorization* [30] because of how it splits the $pp \rightarrow X$ cross section into the soft (long wavelength) interactions distributing the proton momenta among partons and the hard (short wavelength) interaction between partons in the initial state. Factorization is motivated by the different behaviour of QCD interactions at the different characteristic energy scales. The hard interaction correspond to a scale where α_s is small and allows a perturbative treatment. Conversely, soft interactions incorporate a large value of α_s and are therefore treated using non-perturbative techniques.

2.5.2 Event generation with Monte Carlo

Samples of simulated LHC collisions are typically called *Monte Carlo* (MC) after the method used to evaluate phase space integrals for cross section calculations. A form of Monte Carlo is also used to simulate particle interactions with detector components. Generally “Monte Carlo” describes any procedure which performs a calculation by generating random numbers. The canonical example is calculating the area integral of an arbitrary shape by “throwing darts” randomly at the shape within a known area containing it (e.g. a square). The ratio of hits to misses tends to the ratio of areas, as a consequence of the *Law of Large Numbers*. For an n -dimensional shape defined by f and bounded within a region Ω this result can be expressed as

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(\mathbf{X}_i \in \Omega) = \frac{1}{\int_{\Omega} d^n x} \int_{\Omega} f(\mathbf{x}) d^n x \quad (2.60)$$

where $\mathbf{X}_i$ is one of N points drawn randomly in the region. In the case of cross sections, equations 2.43 and 2.59 entail integration over initial momenta from PDF and internal momenta in the underlying Feynman amplitude. Random 3-vectors are drawn for particles inside the $(3n)$ -dimensional phase space restricted by 4 energy-momentum constraints.

The *matrix element* (ME) stage of event generation focuses on sampling the hard interaction cross section. The numeric calculation of the underlying Feynman amplitude is carried out at a prescribed level in perturbation theory. Hence MC generated samples come with labels such as “NLO”, for example, indicating that next-to-leading order corrections have been added to the hard interaction diagram.

A number of mainstream event generators are commonly used for MC samples in ATLAS. Steering files known as “jobOptions” and also run cards are written to tell the generator what hard interaction process to generate and other details such as kinematics. The generators are developed by separate teams, and it is not uncommon to observe significant differences in their predictions for a given physics process. Examples of ME-level generators include POWHEG [31], MADGRAPH [32], and ALPGEN [33]. These ones are typically combined with an additional generator implementing the parton shower (PS) step such as PYTHIA [34]. Other event generators – SHERPA [35] and HERWIG [36] for instance – are commonly used for both ME and PS stages in combination.

A given MC sample will thus typically be labelled according to its physics process, the generator(s) used for ME and PS, and the perturbative order and PDF set used by the generator. Following that are several identifiers related to simulation configuration. However there are additional configurations not usually written. As SM measurements at the LHC increase in precision, they are used to “backpropagate” information about high energy processes for future MC modelling. These data-driven corrections are provided in the form of *tunes* which adjust the PS parameters of a given event generator. In addition to tunes, MC samples may also incorporate “afterburner” generators following event generation to improve the accuracy of process-specific details including photon radiation and higher-order QED corrections.

ISR/FSR and the parton shower

While higher-order corrections augment the hard scatter diagram with additional internal lines, initial and final state radiation (ISR and FSR) affect the kinematics of the diagram by altering the kinematics and multiplicity of its external lines. In cases where ISR and FSR carry off a significant fraction of momentum they produce additional hard objects in the detector signature. Radiation from external fermion lines can be photons, for which case the probability can be calculated perturbatively. The treatment of ISR and FSR for particles which carry QCD color charge is more complicated.

The same properties underlying asymptotic freedom, that gluons are massless and carry color charge, are also the basis for the parton shower phenomenon. Quarks can radiate

gluons, and gluons can split. The real challenge is that for such $a \rightarrow a' + b$ parton splittings which involve low momentum transfer, the strong coupling constant is too large for perturbative calculations. In fact the cross section for a quark radiating a gluon encounters two types of divergences: *collinear divergence* when the gluon is emitted in a direction parallel to the (anti)quark, and *infrared divergence* in the limit of vanishing gluon energy (i.e. “soft” gluon limit). To circumvent these problems, parton shower algorithms rely on non-perturbative approaches to evolve the parton splitting functions.

Parton splitting rates are modeled on the basis of the Altarelli-Parisi⁸ splitting kernels [37], expressed as $P_{a,b}(z)$ for the splitting $a \rightarrow a' + b$. The variable z is the fraction of the initial parton’s energy which is carried by one of the split legs while the other has $(1 - z)$. The *Sudakov form factor* captures the probability for a parton a to not undergo splitting to the sum of possible outgoing partons b when transitioning from scale q to q' :

$$\Delta_a(q^2, q'^2) = e^{-\mathcal{I}_a(q^2, q'^2)} \quad (2.61)$$

$$\text{where } \mathcal{I}_a(q^2, q'^2) = \frac{1}{4\pi} \sum_{b \in \{q, g\}} \int_{q^2}^{q'^2} \int_{z_{\min}}^{z_{\max}} \alpha_s(\tilde{q}^2) P_{a,b}(z) dz \frac{d\tilde{q}^2}{\tilde{q}^2} \quad (2.62)$$

Equation 2.62 resembles the expression $\exp\left(-\int_0^t \lambda dt\right)$ for survival probability until time t of a radioactive nucleus with decay rate λ – except in this case the rate depends on the integration variable. The z integration range is not allowed to go from 0 to 1, but instead is restricted to cases of splitting processes which are deemed *resolvable* – as opposed to a situation where one outgoing parton receives a z fraction below the resolution of the parton shower model. Thus a cutoff scale Q_0 is set so the resolvable energy fraction threshold is $z_{\min} = Q_0^2/q^2$. Conversely, $z_{\max} = 1 - Q_0^2/q^2$.

The evolution of the parton shower is propagated at each iteration with starting scale q^2 by setting $\Delta_a(q^2, q'^2)$ equal to a random value in the interval $[0, 1]$ and solving for q'^2 . A splitting occurs in this iteration if $q'^2 > Q_0^2$ and the process repeats for each outgoing parton until the condition fails. The outgoing momenta and angles of the split partons are distributed according to the splitting kernels $P_{a,b}(z)$.

Hadronization

As shown in figure 2.3, all quarks and gluons produced by the hard scatter plus parton shower subsequently hadronize (introduced in section 2.3.1). During the parton shower, the interaction energy scale lowers, and the strong coupling therefore runs upward, eventually enabling the formation of hadronic bound states that are stable on much larger timescales. Parametric models which have proven successful at describing hadronization offer some

⁸Also attributed to Dokshitzer and also Gribov & Lipatov for their independent derivations (i.e. “DGLAP equations”).

intuition as to what takes place between the quarks and gluons during this process. Two types of hadronization models are commonly used in event generators at the LHC: the *string model* and the *cluster model*.

The string model is based on the picture of quark confinement introduced in section 2.3.1. Flux tubes of the color field act as strings connecting quarks and antiquarks. As they move away from their creation vertex, the string is stretched until it splits into two smaller strings, each ending on the new $q\bar{q}$ pair. Gluons correspond to excitations of the string in the form of kinks.

The cluster model describes hadronization in terms of the formation of color-neutral clusters which then separately dissociate into hadrons. All gluons following the PS are broken up into $q\bar{q}$ pairs which are then grouped with the other quarks from the PS in order to cancel out colors. These clusters are expected to follow a universal mass distribution for singlet states determined by the input scale q and QCD scale λ (derived by Veneziano and Amati [38]). Clusters with invariant mass above 3–4 GeV undergo additional splitting until all clusters are light enough to be converted into hadrons.

The HERWIG event generator, for example, uses the cluster model while PYTHIA uses the string model. Both models have multiple tunable parameters and do better with different aspects of modelling: the cluster model at predicting flavor composition and the string model at energy-momentum modelling.

Underlying event and beam remnants

The hard interaction is also accompanied by fragmentation of the spectator partons (beam remnants) and potentially soft interactions between some of them. This is the likely cause of additional activity observed to accompany events with a hard interaction known as the *underlying event*. As a comparison, the subset of events from minimum-bias triggers that have no identifiable hard scatter exhibit less activity than what is seen to accompany the hard scatter in events which have one (taking into account ISR/FSR) [39]. The detector characteristics of the underlying event are measured in real events with at least one $p_T > 1$ GeV jet [40]. The kinematic and angular distributions are then used to tune MC generators which use parametric models to simulate the underlying event.

2.5.3 Detector simulation

The “Event Generation” box of figure 2.4 has been discussed in the previous section. Next comes the simulation of particle trajectories, interactions, and resulting electrical signals (“hits”) produced by final state hadrons, photons, and leptons throughout the vast array of ATLAS detector components. Afterward, the digitization step emulates the detector readout to convert the detector response into fully digital format. And finally, reconstruction applies the same algorithms as used for data to interpret the digitized information into physics objects.

A virtual model of the ATLAS detector is built in GEANT 4 [41] and incorporated into the ATLAS **Athena** software framework. GEANT 4 simulates the detector response on the level of individual particle interactions with active and dead material. It also accounts for particle decays taking place in the detector such as photon conversions. All this necessary detail makes running the ATLAS simulation on average the largest single use of CPU resources for the ATLAS experiment. The most time-consuming step is the simulation of the ATLAS calorimeter system. This has motivated the development of the ATLAS Fast Calorimeter Simulation (AF2) [42] to shortcut the “fullsim” treatment of particle showers. Many MC samples are produced in AF2, but it is not yet fully suitable for all analyses, including the high-mass dilepton analysis because of the detail needed in high- E_T electron showers.

Pileup

The detector signature of an individual pp collision is complicated by overlapping activity known as *pileup* from particle interactions not coming from the primary vertex. Most pileup is due to additional pp vertices within the same bunch crossing which are a consequence of operating the LHC beam with high luminosity (needed to effectively study most processes of interest; section 2.4.2). In addition, some of the slower detector systems (e.g. LAr calorimeters, TRT) also contend with activity from bunch crossings arriving right before or after (*out-of-time pileup*). Under design conditions, bunch crossings are separated in time by only 25 ns (or roughly 7.5 m). Besides pp collisions, a smaller component of pileup is generated by beam interactions with the beam pipe and residual gas inside it.

Beam conditions in Run 2 produced on average 33.7 interactions per bunch crossing. The vast majority of these interactions fall in the category of elastic, single-diffractive, or double-diffractive processes. The elastic case corresponds to protons scattering but remaining intact while diffractive processes are when one or both protons are broken up by the collision and hadronize. To simulate the effect these pileup vertices have on reconstructing the primary vertex, the processes are first generated in PYTHIA. The generated events are then overlaid with one another according to the expected number of pileup vertices as distributed in data (the pileup “profile”). The whole lot then gets passed to the ATLAS simulation together since that is where their signatures overlap.

Chapter 3

Scenarios beyond the Standard Model

The Run-1 and Run-2 datasets collected at the LHC have provided strong experimental support for the SM. Arguably the most striking confirmation of the SM was the discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations [43]. Furthermore, recent summary plots¹ by the ATLAS experiment demonstrate the general agreement between various predicted and measured SM cross sections across 14 orders of magnitude. However as this section explains, the SM faces an array of theoretical and empirical difficulties as demonstrated by the following list of unsolved questions:

1. Naturalness: (section 3.1)

- *can the small mass of the Higgs boson be explained without excessive fine tuning?*

2. Explanatory power: (section 3.2)

- *why are there three fermion generations and why do they have such disparate masses?*
- *what leads to the fundamental forces having such disparate strengths?*
- *how does gravity work on the subatomic scale and why is it so weak?*

3. Empirical issues:

- *what is the nature of dark matter?* (section 3.3)
- * *where do neutrinos obtain their small but nonzero mass?*
- * *why don't strong interactions violate CP symmetry even though it is not expressly forbidden?*

¹<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/index.html>

where all but the (*) items are shortcomings for which there are theoretical solutions jointly motivating the experimental signature of a high-mass dilepton resonance. This claim will be discussed in this chapter. In this connection, the relevant “answers” to the seemingly unrelated questions listed above take the form of alternative Higgs models, grand unified theories, models with compactified extra dimensions, and dark matter mediator models. Besides the extra-dimensional models, each scenario features massive spin-1 bosons in their scope of phenomenology arising from additional $U(1)'$ or $SU(2)'$ symmetry groups extending the SM gauge structure (2.1). The legitimacy of this type of model building dates back to Georgi and Weinberg [44] who showed that electroweak predictions are not in principle spoiled by introducing $U(1)'$ groups beyond the SM $SU(2)_L \times U(1)_Y$. Because of the breadth of BSM scenarios which overlap in their prediction of heavy gauge bosons, the corresponding search for dilepton resonances is a highly-motivated search channel at the LHC.

3.1 The unnaturally light Higgs mass

Unlike many SM particles, the Higgs boson and several of its key characteristics were foreseen theoretically several decades before its discovery. Following a decade of analysis at the LHC, the Higgs boson’s scalar nature, CP properties, *dominant* production and decay modes, and mass-proportionate Yukawa and gauge couplings have been incrementally confirmed within experimental reach.

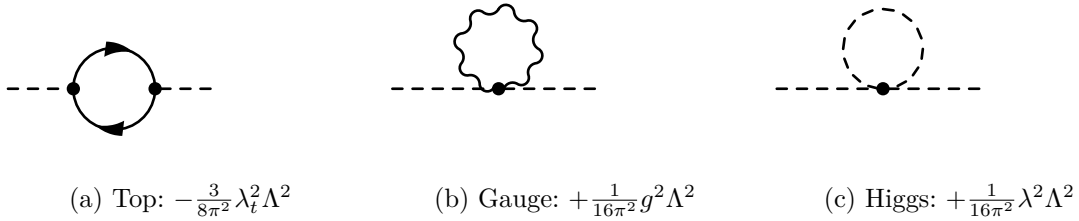


Figure 3.1: Dominant loop corrections to the Higgs mass at scale Λ .

One important property of the Higgs boson that is a free parameter in the SM and therefore *not* predicted is the Higgs mass. Theoretical unitarity bounds and experimental constraints from the LEP and later Tevatron experiments were prior indications of a light Higgs, but the observed mass of 125 GeV confirmed this suspicion and the problem it raises for the SM as a UV-complete theory. Although a light Higgs mass does not contradict the SM it renders incredible the prospect that SM interactions are sufficient to describe physics deep into the multi-TeV scale. The observed Higgs mass is the sum of both a bare mass term and all contributions from the set of loop-order diagrams involving other massive particles. The three dominant diagrams involving top quarks, gauge bosons, and the Higgs self-interaction are shown in figure 3.1. The requirement that these large corrections cancel with the Higgs bare mass to result in a *comparatively* small value of 125 GeV is known as

the *hierarchy problem*. More generally, as will be discussed in the next section, the hierarchy problem refers to the observed vast separation between the electroweak scale and the Planck scale:

$$m_{\text{EW}} \simeq 246 \text{ GeV} \ll M_{\text{Pl}} \equiv \sqrt{\frac{\hbar c}{8\pi G}} \simeq 2.4 \times 10^{18} \text{ GeV} \quad (3.1)$$

Coming to the Higgs mass, the hierarchy problem manifests as an inability to extrapolate the SM dynamics to the Planck scale. For example, assuming the SM holds up to an energy scale $\Lambda = 10 \text{ TeV}$ the Higgs mass is calculated approximately as follows [45]:

$$m_h^2 = (125 \text{ GeV})^2 = m_0^2 - (2000 \text{ GeV})^2 + (700 \text{ GeV})^2 + (500 \text{ GeV})^2 + \dots \quad (3.2)$$

At this scale the difference between the Higgs bare mass term m_0^2 and the corrections must lie within 1% to realize a 125 GeV Higgs. Because of the quadratic dependence of the corrections on cutoff scale Λ this fine tuning only becomes more dramatic the further the SM is extrapolated to high energy. For the SM to hold up to the Planck scale would require fine tuning to one part in 10^{32} . While other SM particle masses also experience quantum corrections, they are limited by symmetry constraints and typically scale with Λ logarithmically so do not entail a problematic degree of tuning. The lightness of the Higgs mass therefore presents a strong case that new physics will either naturally realize the cancellation or otherwise intervene at some high energy scale.

Perhaps there is an additional set of loop diagrams involving BSM particles which cancel the SM contributions by construction. This traditional remedy to the hierarchy problem is a central feature of supersymmetric theories. By now searches for supersymmetry (SUSY) at the LHC have constrained its parameter space enough to cast doubt on its ability to address the hierarchy problem. However instead of invoking a “cancellation by necessity” an alternative possibility is for new physics to interrupt the scaling of the corrections at some energy. The quadratic dependence on energy scale of the Higgs corrections could be assuaged if the Higgs only behaves as a scalar up to some multi-TeV energy scale, for example if the Higgs is a composite particle bound by new strong dynamics. In this case the smallness of its mass can be explained by postulating that it is a pseudo Nambu-Goldstone boson in analogy to the SM pions which are scalar composites of the strong force. The approximate $SU(3)_L \times SU(3)_R$ chiral symmetry exhibited by quarks in the SM is spontaneously broken at the scale of quark condensation ($\sim 1 \text{ GeV}$), generating the spectrum of pseudo (i.e. only approximately massless) Nambu-Goldstone bosons including the relatively light pions (135, 140 MeV). If the Higgs fields are Nambu-Goldstone modes of some broken approximate symmetry an analogous mechanism could explain the smallness of the Higgs mass compared to the compositeness energy scale. This was the basic premise of early composite Higgs models pioneered by Kaplan and Georgi (e.g. [46]). More recent incarnations through variations of *Little Higgs* (e.g. [47, 48]) and *Twin Higgs* models [49]

featured greater compatibility with electroweak precision constraints by naturally elevating the compositeness scale into the tens of TeV.

Models proposing an alternative nature of the Higgs boson typically rely on an expanded symmetry group, for example $SU(5)$, which includes the SM and the TeV-scale new physics particle states prior to symmetry breaking. This generally implies the existence of residual $U(1)'$ or $SU(2)'$ groups at energies below the breaking energy, and hence their corresponding heavy gauge boson(s) (see [50, 51] for reviews). Drell-Yan production through their couplings to SM fermions allows these high-mass states to be directly probed in leptonic final states at the LHC.

3.2 The disparity in strengths of the fundamental forces

3.2.1 Grand Unified Theories

The unification of the electromagnetic and weak nuclear forces is a high-point of the Standard Model’s explanatory power. The SM explains how, despite their substantial differences, electromagnetic and weak interactions are fundamentally connected. Although there is no *a priori* reason why the two forces (or three, recalling Maxwell’s unification) could not have been entirely separate entities, a theory for seeing them as one is compelling. The conceptual appeal of unification schemes correlate with the idea – articulated by Dirac – that mathematical elegance is a core guiding principle in the development of successful theories of fundamental physics. From a practical point of view, Grand Unified Theories (GUTs) can potentially explain otherwise arbitrary properties of nature, in particular (1) the vastly different strengths of the fundamental forces, and (2) the peculiar group-theoretic (see equation 2.5) and family structure of the SM fermions, e.g. Yukawa couplings. While in the SM these properties are simply taken for granted, a successful GUT demonstrates how distinctions between gauge forces and the differences between fermions dissolve at a very high energy scale Λ_{GUT} below the Planck scale. A GUT where this is accomplished could also serve to bridge the large-scale interactions of gravity and the small-scale dynamics of quantum fields.

The founding proposal for gauge unification [52] was based on the smallest symmetry group which is able to embed the SM tri-fold gauge structure G_{SM} : $SU(5)$. With their diagonals connected the 3×3 and 2×2 matrices of the fundamental representation of $SU(3)_c$ and $SU(2)_L$ form traceless 5×5 members of $SU(5)$ while $U(1)_Y$ comprises a single traceful diagonal member. However the minimal $SU(5)$ scheme comes short of grouping all fermion fields in a single multiplet (it needs two per generation) and implies unacceptable rates of nucleon decay. These shortcomings led to the development of GUTs based on larger symmetry groups, first $SO(10)$ and then the larger exceptional group E_6 . These are related

to one another and to $SU(5)$ through the following possible breaking chain:

$$\begin{aligned}
E_6 &\rightarrow SO(10) \times U(1)_\psi \\
&\downarrow \\
&\left\{ \begin{array}{l} SU(5) \times U(1)_\chi \\ SU(3)_c \times SU(2)_L \times SU(2)_R \times U(1)_{B-L} \\ \dots \end{array} \right. \\
&\downarrow \\
&G_{\text{SM}}
\end{aligned} \tag{3.3}$$

where the ellipsis denotes other breaking patterns not shown. Depending on the model, the highest arrow denotes breaking at Λ_{GUT} while the subsequent arrow(s) denotes spontaneous symmetry breaking at an intermediate energy scale. $SO(10)$ is promising in that its **16** spinor representation is large enough for a single SM fermion generation (with room for one additional singlet, potentially a right-handed neutrino). In the larger E_6 then, there are more than enough slots in its fundamental **27** representation such that it risks introducing more exotic fields than what is helpful. Figure 3.2 shows the predicted running of gauge coupling strengths in an $SO(10)$ -based GUT [53] with a grand unification scale $\Lambda_{\text{GUT}} > 10^{17}$ GeV and an intermediate breaking scale of the $SU(5)$ subgroup which can be seen at $10^{15.96}$ GeV.

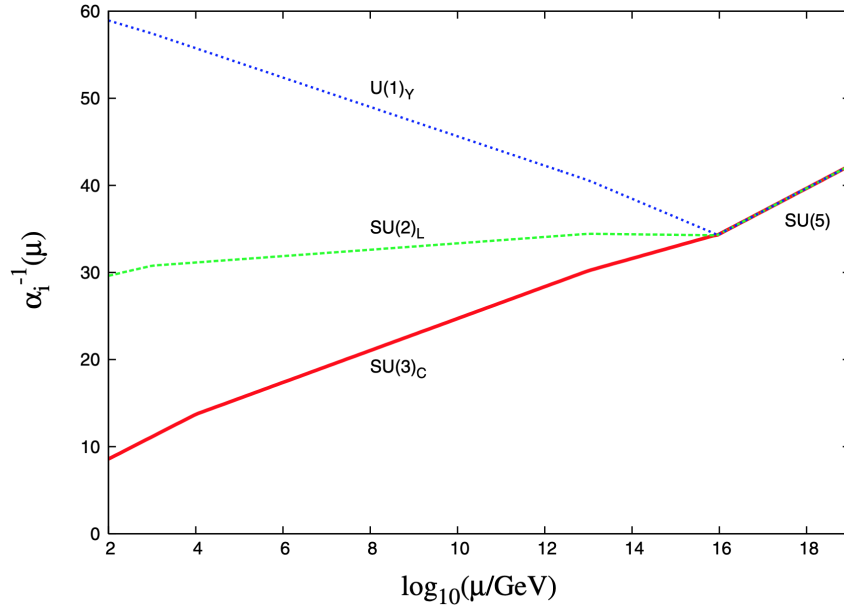


Figure 3.2: Example showing the coalescence of strong and electroweak (reciprocal) gauge couplings as functions of energy scale μ in a non-SUSY $SO(10)$ GUT (Fig. 1 of [53]).

GUTs generally predict residual $U(1)$ symmetry groups beyond the SM. This is clear from the breaking patterns shown in equation 3.3, in particular the groups labelled ψ , χ , and $B - L$ (associated charge is difference between baryon and lepton numbers). The corresponding experimental signature of a heavy neutral resonance has been parameterized in this thesis in terms of the Z'_χ and Z'_ψ resulting from a superstring-inspired E_6 GUT, following [54]. In general, both χ and ψ residual groups may be present in a linear combination at the TeV scale:

$$U(1)' = U(1)_\chi \cos \theta + U(1)_\psi \sin \theta \quad (3.4)$$

Different choices for θ are used in a variety of models which lead to Z' resonances with different couplings and hence widths, cross sections, etc. (table 1 of [55]).

Baryon and lepton number non-conservation (e.g. proton decay) is a general consequence of GUTs because of BSM vertices connecting quarks and leptons via new gauge bosons. As such, viable GUTs require auxiliary mechanisms to make them compatible with increasing lower bounds on proton decay rates. For example the current exclusion on the proton lifetime below 1.6×10^{34} years by Super-Kamiokande [56] is 2-6 orders of magnitude longer than expected by “vanilla” GUTs [57]. Supersymmetric GUT models can make use of symmetries to protect baryon and lepton numbers (e.g. R-parity) but must also contend with additional problematic operators induced by superpartners. On the other hand, in GUTs built on 5 or 6 spacetime dimensions it is possible to strongly suppress the baryon- and lepton-number violating operators.

The idea of GUTs in extra dimensions is also inspired by the extra-dimensional landscape of string theory. String theory is compelling for how naturally it accounts for quantum gravity, yet it has proved difficult to probe experimentally². So the string-GUT intersection both extends the unification paradigm (to include gravity along with the SM gauge forces) and presents a scenario connecting string theory to more testable phenomenology at energies that can potentially be probed in collider experiments.

3.2.2 Extra-dimensional models

Models not reliant on string theory have also shown that extra spatial dimensions can potentially address the hierarchy problem (eq. 3.1), particularly as it manifests in the relative weakness of gravity. The n additional dimensions may have gone unnoticed if they are sufficiently compactified, that is, confined to a small circle of radius R . In that case the intuition for the weakness of gravity is that large-scale interactions only capture a small fraction of the force’s field lines, as illustrated for $n = 1$ in figure 3.3a. If gravity is able

²Superstring theories, which account for elementary fermions in addition to bosons, do so by invoking a form of supersymmetry on bosonic strings which arise naturally as spacetime vibrations. But since the scale of supersymmetry breaking is not predicted, this feature is not easily falsifiable by collider searches for supersymmetry.

to propagate in the $(4 + n)$ -dimensional bulk, Gauss's r^{-1} potential becomes $\sim r^{-(1+n)}$ on distance scales comparable to R . Then it is possible that the separation $m_{\text{Pl}} \gg m_{\text{EW}}$ of the Planck mass from the electroweak scale at $r \gg R$ is due entirely to a subliminal factor of R^{-n} in the gravitational potential, whereas in the full $4 + n$ dimensions the two scales are in fact *one*. In limiting ourselves to Minkowski space we have been calculating an *effective* Planck scale, failing to realize that its size implicitly factors in the geometry of additional dimensions. This argument was the key idea introduced by Arkani-Hamed, Dimopoulos, and Dvali (ADD) in 1998 [58]. For $n = 2$ extra dimensions the compactification radius R to make this work was small enough to duck constraints of sub-mm gravitation at the turn of the century. However subsequent tests of Newtonian gravity – recently down to $52 \mu\text{m}$ [59] – have cast doubt on the $n = 2$ scenario.

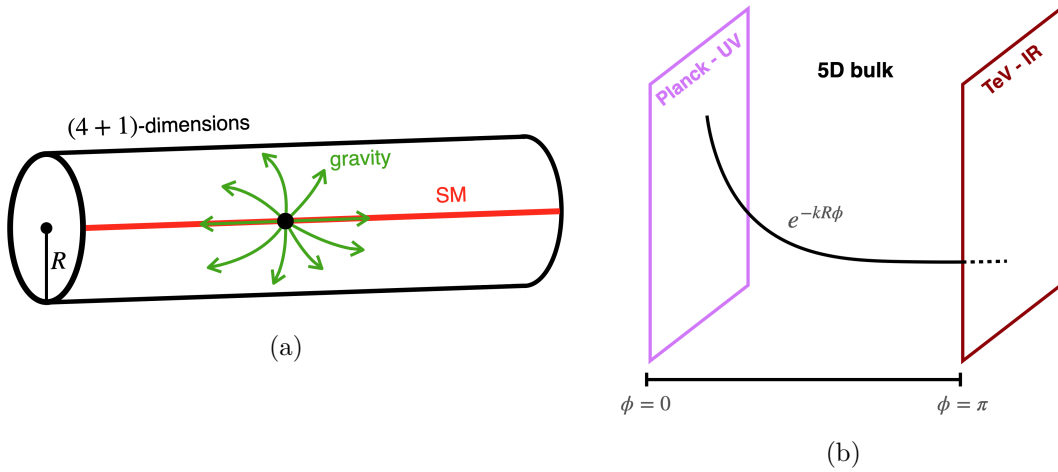


Figure 3.3: Visual aids to the ADD (left) and RS (right) models. In (a) the compactified $n = 1$ extra dimension is shown as a circular extension to the 4-dimensional spacetime depicted as a straight line. The weakness of gravity in this case can be understood as the small fraction of gravitational flux lines (green) which lie along the 4-dimensional SM fields (red). In (b) the separation of two branes at endpoints of the extra ϕ dimension represent the hierarchy between Planck and TeV scales generated by an exponential warp factor of the 5-dimensional metric.

Inspired by the ADD model, Randall and Sundrum (RS) showed that a similar explanation could be obtained with the addition of only one warped dimension ϕ [60]. The extra dimension is likewise compactified on a circle to the segment $\phi \in \{0, \pi\}$. In their 5-dimensional spacetime, distances are measured by the following anti-de Sitter (AdS) metric:

$$ds^2 = w(\phi)^2 \eta_{\mu\nu} dx^\mu dx^\nu + R^2 d\phi^2, \quad w(\phi) = e^{-kR\phi} \quad (3.5)$$

where k is the single fundamental energy scale of nature near the Planck scale. Unlike the flat metric assumed in the ADD, here the exponential $w(\phi)$ warps Minkowski space in the direction of ϕ . The warp factor allows small values of R to generate a large hierarchy

between the strength of gravity at the two boundaries depicted in figure 3.3b: the Planck/UV brane ($\phi = 0$) and the TeV/IR brane ($\phi = \pi$). The TeV brane is the 3-brane world of SM interactions while gravity is again free to propagate in the 5-dimensional bulk. Consequently “we” measure a scale of the gravitational field that is suppressed by factor $\exp(-kR\pi)$ with respect to its value on the Planck brane.

In ADD and RS models and their variations, fields propagating in the higher-dimensional bulk manifest in 4-dimensional spacetime as an infinite series, or “tower” of massive modes in a mechanism known as Kaluza-Klein (KK) reduction³. These modes arise because of the periodic nature of the compactified extra dimension(s); they correspond to standing wave excitations of the bulk field at characteristic frequencies. Since the solution to the hierarchy problem requires that gravity inhabit the bulk the KK modes arising from the graviton are a prime collider signature of compactified extra dimensions. In the ADD model the masses m_n^G of the graviton KK modes are evenly spaced by factors of $1/R$ whereas in the RS model [63],

$$m_n^G = k e^{-kR\pi} x_n \quad , \quad J_1(x_n) = 0 \quad (3.6)$$

that is, spaced according to the roots of the J_1 Bessel function. In both cases the modes get heavier with more compactification (i.e. shrinking R). The first, lightest graviton mode is a target for resonant production at the LHC through $q\bar{q}$ and gg production modes. It corresponds to a narrow (sub-resolution) resonance with angular distribution characteristic of a spin-2 mediator. This distinguishes the graviton signature from a spin-1 resonance as does its ability to decay to $\gamma\gamma$. In minimal models, the $G \rightarrow e^+e^-$ decay channel has a $\sim 2\%$ branching ratio yet drives discovery because of superior mass resolution and SM background modelling at multiple TeV [64].

Subsequent variations on the ADD and (particularly) RS models explore the possibility of SM gauge and fermion fields cohabiting the bulk with the graviton. Strictly speaking, only the Higgs field need be confined to the TeV brane to generate a Planck-TeV hierarchy. This more general scenario is useful for explaining the fermion mass hierarchy [65]. In such flavourful models the coupling of light SM fermions to graviton KK modes can be highly suppressed (related to their Yukawa couplings), essentially squashing the $q\bar{q} \rightarrow G$ and $G \rightarrow \ell^+\ell^-$ collider channels [63]. Decays through bosonic and top quark channels then play a leading role. However even in these cases the fermions may remain weakly coupled to KK modes associated with bulk gauge fields. The Z , W , γ , and g KK towers are highly mass-degenerate, very narrow, and exhibit strong interference effects. The mass of their overlapping lightest modes is of order 60% that of the graviton in simple extensions of the

³The name in fact dates back to the 1920s and an early attempt by Kaluza [61] to unify gravity and electromagnetism in a 5-dimensional general relativity. It was Klein who further contributed the idea – now foundational to string theory etc. – that the extra dimension may be confined to a small circle [62].

RS model [66]. Hence a multi-TeV resonance with a spin-1 angular signature could also arise from a warped extra dimension.

3.3 The nature of dark matter

A major indication for physics beyond the SM stems from the cumulative empirical case for the existence of dark matter (DM). Current estimates suggest that DM comprises over 80% of the total matter in our universe. While the first evidence for DM came from departures in galactic rotation curves, the case has been further strengthened by observations of gravitational lensing and by parameter estimation from acoustic modes of the cosmic microwave background (CMB). A candidate for DM must be able to produce a gravitational effect large enough to account for these astrophysical data while remaining non-luminous, effectively transparent, and cold (i.e. non-relativistic). There are potential sources from macroscopic objects (gas clouds and MACHOS) which can fit this description but these candidates are hard pressed to meet the observed abundance and distribution of DM. On the other hand theoretical solutions of modified Newtonian dynamics (MOND) on large distance scales offer an explanation of galactic rotation curves without DM. However MOND does not account for lensing and CMB data, and the ramifications of replacing general relativity with a modified classical description make the proposal dubious at best. The failure of these alternatives has put the spotlight on candidates of non-baryonic DM as the most credible explanation.

The current distribution of DM derives from its original (*relic*) density which “froze out” as the universe cooled and particle species fell out of thermal equilibrium. The current DM relic density is therefore a function of not only the expansion rate of the universe⁴, but of the thermal average of the DM relative velocity v and its self-annihilation cross section $\sigma_{\chi\chi\rightarrow\text{SM}}$ [67]:

$$\Omega_\chi \propto \frac{T_0^3}{\langle v \cdot \sigma_{\chi\chi\rightarrow\text{SM}} \rangle h^2} \quad (3.7)$$

where T_0 is the temperature of the CMB, about 2.7 K [68]. Note that the value Ω for a given matter type refers to its density relative to the critical density: $\Omega = \rho/\rho_{\text{cr}}$ (weighing in all matter in addition to dark energy yields a total ratio near unity). Accounting for contributions to DM from baryonic sources, then, the non-baryonic DM relic density is extracted from a fit to the CMB anisotropy data, e.g. from *Planck* [69]:

$$\Omega_\chi \simeq 0.26 \quad (\text{or } \Omega_\chi h^2 \simeq 0.12) \quad (3.8)$$

⁴Hubble’s constant enters equation 3.7 through the dimensionless “little h ”, via $H_0 = 100 \cdot h$ (km/s)/Mpc.

In the SM, the neutrino is the only conceivable DM candidate but is simply too light (and hence fast) to fit the bill. Beyond the SM, several proposals exist for non-baryonic DM which is sufficiently stable to persist on cosmological timescales and abundant enough to produce the observed DM relic density. It is important to note that although DM could in principle *only* interact gravitationally, postulating that it does couple weakly to other particles offers the hope of directly detecting DM. This is the basis for the broad scenario of a new weakly-interacting massive particle (WIMP), χ . In general the WIMP mass can be in a large range of values from 2 GeV up until around 100 TeV where it reaches unitarity bounds [70]. Part of why the WIMP is popular is because such a particle can arise in the contexts of several extensions of the SM addressing problems independent of DM. This is particularly true in SUSY where the lightest superparticle (LSP) makes an ideal DM candidate – its stability ensured as a consequence of R-parity. But WIMPs can also be accommodated in non-SUSY scenarios such as Little Higgs [71], Twin Higgs [72], and extra-dimensional GUTs [73] which were mentioned in sections 3.1 and 3.2. The plausibility of WIMP DM was further enhanced by the so-called *WIMP miracle* – the realization that a particle with mass near the electroweak scale and weak-characteristic annihilation cross section leads to a relic density surprisingly near what is expected from equation 3.7. The WIMP is therefore both a suitable DM candidate and is a potential byproduct of various extensions of the SM.

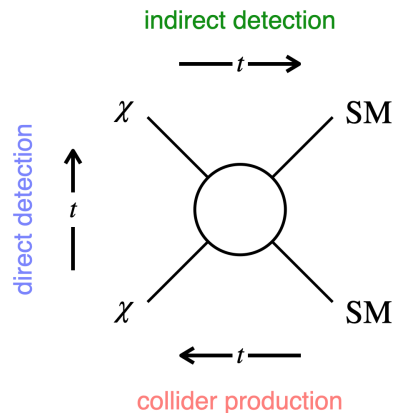


Figure 3.4: Complimentary experimental probes of DM.

So far, the existence of DM is inferred from observations of astrophysical systems, CMB, and the large-scale structure of the universe. To more precisely discern the nature of DM there are three complementary experimental approaches illustrated in figure 3.4. First, DM annihilation could be indirectly observed through its influence on the energy spectra of cosmic rays and neutrinos as measured by ground-based and atmosphere/space-based detectors. So far these measurements remain inconclusive, although a persistent positron excess reported by the AMS02 experiment [74] could be interpreted – though not uniquely – as a cosmic DM annihilation signal. A DM signal would be less ambiguous if reported by direct detection

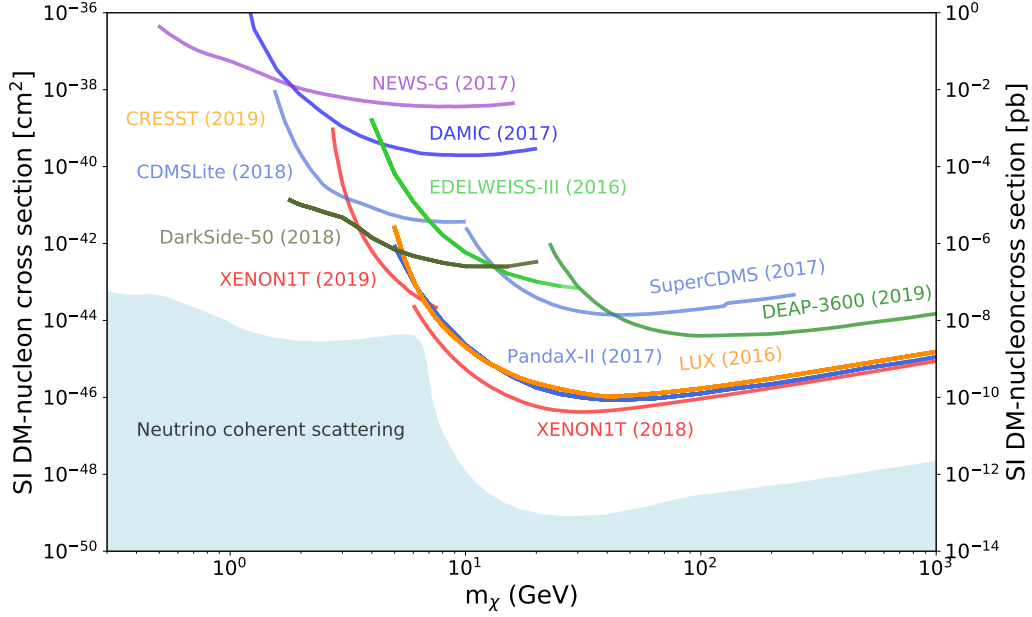


Figure 3.5: Status of spin-independent (SI) DM exclusions from direct detection experiments curated in the 2020 Review of Particle Physics [11].

experiments. Direct detection experiments are designed to observe DM from outer space scattering *in-situ* off of atomic nuclei in the apparatus. Recent experiments have honed a variety of techniques to detect minute scattering signals in the form of photons (LUX, DarkSide, DEAP, PandaX, XENON), phonons/heat (CDMS, CRESST, EDELWEISS), and ionization charge (DAMIC, NEWS-G). Figure 3.5 summarizes current upper bounds on DM-nucleon scattering cross sections from these experiments. Lastly, collider-based experiments are the third complementary approach of directly probing DM by creating it through collisions of SM particles. The ATLAS collaboration probes a diversity of signatures related to DM using LHC collision data. This includes searches for complete models such as SUSY, specific scenarios such as sterile neutrinos, and generic signatures parameterized by effective field theory (EFT) operators or through simplified DM models.

The lack of confirmation from direct-detection and collider experiments has inspired new proposals for the nature of DM. One possibility is that DM does not partake in any SM interactions, but only interacts with a hidden dark sector of particles. Some of the other dark sector particles, such as a new mediator, may have SM couplings. In this case DM would be accessible at the LHC through a dark mediator “portal”. Directly searching for a new mediator with weak couplings to SM fermions is therefore a probe of the dark sector scenario. Recent iterations of relevant searches with ATLAS and CMS have employed simplified models to connect cross-experiment multi-channel results to the range of parameter configurations within this context [75, 76]. The latest bounds on high-mass dilepton

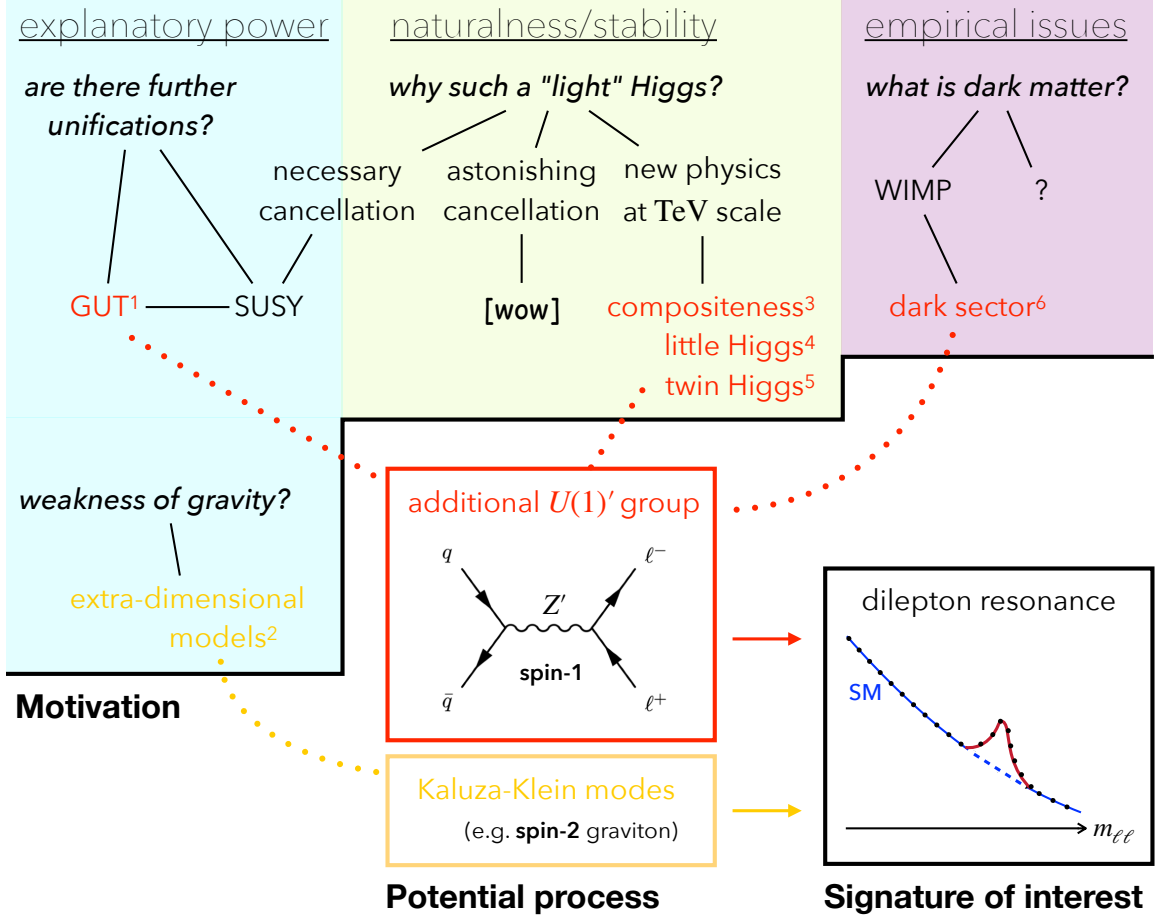


Figure 3.6: Summary of theoretical motivations possibly connected to the experimental signature of a dilepton resonance. Example citations, in order of the numbered superscripts on the scenarios highlighted in color are [54], [60], [77], [78], [79], and [80]

resonances have an important bearing on the DM mediator interpretation as expounded in section 9.4.

3.4 Dilepton resonances

A summary of how the signature of a dilepton resonance is possibly connected to the BSM scenarios discussed above is shown in figure 3.6. Each scenario in general may have specified or preferred coupling structure and mass, width values etc. of the new mediator, but a search for generic dilepton resonances remains broadly sensitive across such differences. The basic parameterization of a new gauge boson Z' coupling to a SM fermion field f is captured in the following Lagrangian [55]:

$$\mathcal{L}' = g' Z'_\mu \bar{f} \gamma^\mu (\epsilon_L^f \mathcal{P}_L + \epsilon_R^f \mathcal{P}_R) f \quad (3.9)$$

$$= g'_V Z'_\mu \bar{f} \gamma^\mu f - g'_A Z'_\mu \bar{f} \gamma^\mu \gamma^5 f \quad (3.10)$$

where $\mathcal{P}_{L/R}$ are the chiral projection operators defined in (2.23) and the $Z' - f$ couplings g' and ϵ^f are related as follows:

$$g'_{V/A} = \frac{g'}{2}(\epsilon_L^f \pm \epsilon_R^f) \quad (3.11)$$

The V and A denote the equivalent parameterization in terms of a new neutral vector or axial-vector current. The differential cross section with respect to the total 4-momentum squared of the s-channel production/decay $q\bar{q} \rightarrow Z' \rightarrow ff$ has the form of a Breit-Wigner distribution:

$$\frac{d\sigma}{ds} \propto \frac{1}{(s - M_{Z'}^2)^2 + M_{Z'}^2 \Gamma_{Z'}^2} \quad (3.12)$$

where $M_{Z'}$ is the mediator mass and $\Gamma_{Z'}$ is its width associated with this decay mode. Note that s in this case is simply the square of invariant mass of the final state fermions. The width of the resonance and its proximity in mass to other (e.g. SM) resonances determines the degree of interference which would augment (3.12), discussed in section 2.4.1. The width, in turn is dependent on the Z' couplings. What this means is that the differential cross section can not exactly be varied independently of the width and vice versa. This presents a technical challenge in fits to data which attempt to extract a cross section (i.e. best-fit signal size) since the resonance width and by extension interference effects vary together with it. However, the interdependence can be shut off by considering resonances with infinitesimal widths. In practice, this *narrow width approximation* (NWA) is made also for resonances with $\Gamma \neq 0$ with an error that goes like $\mathcal{O}(\Gamma/M)$ [81]. The NWA is equivalent to the limit where the s-channel mediator is always on-shell and the Breit-Wigner distribution (3.12) is replaced with a delta function.

In the narrow width approximation, and assuming no threshold effects the width can be determined for a given coupling configuration as follows:

$$\Gamma_{Z'} = \frac{1}{12\pi} N_c \left[(g_V^f)^2 + (g_A^f)^2 \right] M_{Z'} \quad (3.13)$$

where N_c is from the sum over colors of the fermion (i.e. just 1 for leptons). Hence the relative width $\Gamma_{Z'}/M_{Z'}$ is a constant for models with fixed couplings.

Flavour-universal Z' fermion couplings are a frequent assumption although there are many models of generation-dependent couplings. An extended gauge sector which assigns generation-dependent charges necessarily introduces flavour-changing couplings which must contend with precision electroweak constraints [82]. This is more viable for enhanced third-generation couplings (e.g. $Z' \rightarrow t\bar{t}$), whereas for the search presented in this thesis, lepton flavour universality is assumed when combining the electron and muon channels.

Chapter 4

The ATLAS detector at the LHC

What is today the largest particle detector ever built, installed at the world’s most powerful particle collider, is the culmination of several decades of planning and construction by a global community of physicists and engineers. The proportions of the LHC were constrained by the decision to reuse the existing 27 km tunnel of the CERN Large Electron Positron (LEP) accelerator. With the bending radius fixed, the more than two orders of magnitude increase in beam energy is a reflection of advances in superconducting magnet technology (entering through equation 2.46) and decreased synchrotron radiation losses of protons compared to electrons. It took over a decade from initial discussions of the LHC in 1984 until the project was approved by CERN in the end of 1996.

Initial proposals around the late 1980s for general purpose experiments at the LHC converged around two distinct layouts for magnetizing the muon subdetector: toroidal versus solenoidal [83]. A guiding concern was muon momentum resolution on which would hinge the potential to reconstruct a $H \rightarrow 4\mu$ signature. The projected luminosity and centre of mass energy of the LHC called for detectors with unprecedented performance in terms of timing, granularity, radiation hardness, and resolution. ATLAS – A Toroidal LHC ApparatuS – was the merger of two earlier designs both based on a toroidal magnet system. The choice of air-core versus iron-filled toroidal magnets for the muon detector is mainly what led to the enormous size of the ATLAS detector. This is in contrast to the design of what would become the complementary general-purpose experiment, Compact Muon Solenoid (CMS), which is built around a single solenoidal magnet contained in an iron return yoke. In addition to their unique muon and magnet systems, the ATLAS and CMS detectors are distinct on numerous levels, such as their important differences in calorimeter design.

4.1 Overview of ATLAS and the LHC

The LHC and its experiments thus came to be what is now the world’s premiere facility for directly probing fundamental physics at the high energy frontier. The LHC is primarily a proton-proton collider but also dedicates runtime for colliding beams of heavy ions (Pb and

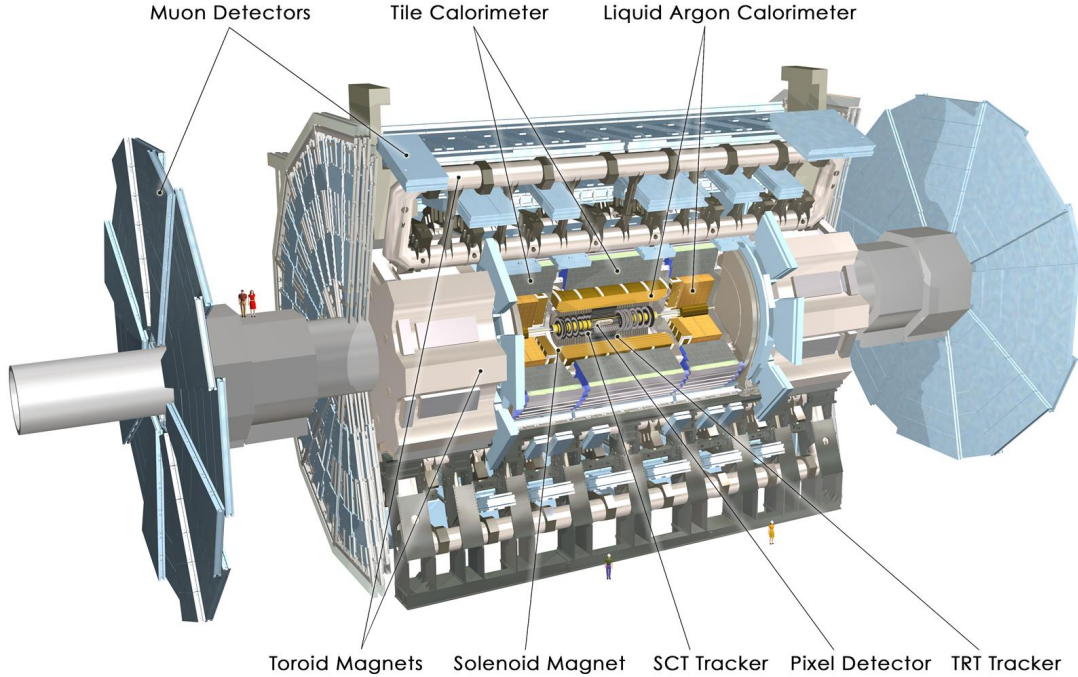


Figure 4.1: Cutaway of the ATLAS detector, with figurines for scale. Some conspicuous parts that are not labeled include: octagonal radiation shields (dark grey) between the large and outer muon wheels; at the very top, L-shaped extensions of the endcap toroid cryostats that connect to flexible liquid helium supply lines (not shown); feet and rail system (dark grey) built from non-magnetic steel. Figure taken from [84].

recently Xe). The first LHC collisions in late 2009 involved proton beams of 3.5 TeV – more than tripling the record of its predecessor the Tevatron. The dataset analysed in this thesis concerns the LHC Run 2 when protons were colliding with $\sqrt{s} = 13$ TeV centre-of-mass energy. Four large experiments operate on the LHC beamline: ATLAS, LHCb, CMS, and ALICE. The LHCb detector is optimized to measure CP violation, decays of b hadrons, and exotic QCD bound states. The ALICE detector is geared to study properties of quark-gluon plasma, for which Pb-Pb collisions play a central role. The ATLAS and CMS experiments were designed to facilitate a wide range of analyses, most importantly searches for the Higgs boson and BSM physics.

The ATLAS detector lies 100m underground at the LHC point 1, just outside Meyrin, Switzerland. It has a cylindrical geometry (figure 4.1) which is 25 m in diameter and 46 m long, weighing in at approximately 7×10^6 kg. Its interface with the LHC is through the 0.8 mm thick wall of a beryllium beam pipe segment centred on the interaction point.

The opposing LHC proton beams circulate in separate pipes side by side¹. Upstream from the interaction point, the beams are bent inward and dramatically focused so as to cross through each other at the centre of the ATLAS detector, swapping sides as they exit. The beams are not continuous but each consist of 2808 proton bunches equally spaced to arrive simultaneously at the collision point every 25 ns. The bunch configuration achieves good energy coherence of the beam particles by constraining their velocity dispersion to sync with the oscillating potential gradient in the RF cavities. To maximize collision rate, the beams are also focused to intersect in a luminous region which in Run 2 was around 34-44 mm in length and 7-14 μm in width on average. When this compression was consistently the highest in 2018 it resulted in 36 proton collisions per bunch crossing (pileup). At this rate, since each bunch contains $\sim 10^{11}$ protons, the bunches can be reused for several hours during a single LHC fill. The corresponding instantaneous luminosity is then highest at the beginning of a fill, when in Run 2 it reached and at times doubled the target of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$.

4.1.1 Coordinate system and kinematic variables used in ATLAS

The ATLAS detector is described using a right-handed coordinate system oriented such that z lies along the beamline, x points to the centre of the LHC, and y points perpendicular to both, nearly vertically². The x - y plane is referred to as the *transverse plane* and is also parameterized in terms of the azimuthal angle ϕ and the radial distance r from the beamline. The polar angle θ is measured from the positive z -axis and used to define the *pseudo-rapidity* variable $\eta \equiv -\ln \tan(\theta/2)$. As its name implies, the pseudo-rapidity is a proxy for the rapidity (for boost in the z -direction) that is used because it can be calculated using θ . For massless particles and for ultrarelativistic particles the two are interchangeable. The variables η and ϕ form an orthogonal basis for describing the set of rays emanating from the origin. Particle tracks, in general, are not straight lines because they undergo bending and multiple scattering in ATLAS, so (η, ϕ) coordinates are assigned according to their initial trajectory from the primary vertex. Distances in $\eta - \phi$ space are measured as follows:

$$\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (4.1)$$

A fixed ΔR defines a cone with its apex at the origin which is typically used to define the neighborhood of a given track (or energy deposit) in $\eta - \phi$ space.

Because it is not known how much of the 6.5 TeV proton energy was carried by any given parton, the net momentum of the final state in the longitudinal direction is generally nonzero. On the contrary, the incoming partons have negligible momentum in the transverse

¹This is one engineering drawback of a pp collider. A $p\bar{p}$ collider, by contrast, is able to circulate beams using the same beam pipe and magnets.

²The ATLAS detector is inclined by 1.24% with respect to the horizontal to match the LHC beamline.

plane³, so it is justified to assume the final state momenta will therefore sum up to zero in this plane. The component of a particle’s momentum that lies in the transverse plane, $\vec{p}_T$, is therefore of particular significance in pp collisions. A nonzero sum of the transverse momenta of all final state particles is known as *missing transverse momentum*, $E_T^{\text{miss}} \equiv -|\Sigma \vec{p}_T|$. The symbol is owing to an unfortunate conflation with the slightly distinct quantity called *missing transverse energy*, i.e. “MET”. To be precise, the transverse energy, (which itself is an awkward term) is defined as $E_T^2 \equiv p_T^2 + m^2$, but for highly relativistic particles the mass can be ignored. If all visible particles were measured accurately, missing transverse momentum in an event implies the presence of an undetected particle, either a neutrino or a BSM particle. Events with large E_T^{miss} are therefore selected by many BSM searches.

4.2 Magnet system

As mentioned above, the gigantic toroidal magnet system is a defining feature of the ATLAS detector. More precisely, ATLAS has four separate magnets: a solenoid enveloping the inner detector, a toroidal magnet in the barrel region of the muon spectrometer (MS), and a separate toroidal magnet for each of the two MS endcaps. The magnets have an air core which space is ultimately occupied by other detector components. In the MS this choice greatly helps minimize dead material in order to reduce multiple scattering of muons. Table 4.1 compares the main specifications of the three magnet systems.

The magnetic field in ATLAS ($\vec{B}$) plays the vital role of bending the trajectories of charged particles emanating from the collision point. Their maximum deflection Δs from straight lines, known as *sagitta*, is then a measure of their inverse momentum:

$$\Delta s \propto \frac{q}{p} \int \vec{B} \cdot d\vec{\ell} \quad (4.2)$$

Accurate measurements of high-momentum particles is therefore achieved using magnetic fields which are large both in field strength (bending power) and spatial extent (lever arm). For the physics goals of ATLAS this would include particle momenta of up to ~ 1 TeV. The strongest of the magnet systems, the solenoid, covers the region surrounding the ATLAS Inner Detector (ID). The resulting axial field serves to reconstruct tracks of charged particles which is crucial for distinguishing track vertices and to measure isolation, identification, and kinematics. Since muons are not absorbed by the calorimeters their momentum measurement relies almost entirely on track sagitta (section 5.2). This measurement is enabled by further bending in the MS which extends the lever arm far beyond what the solenoid provides. The toroidal field in both barrel and endcap regions lies along the ϕ direction.

³Assuming that possible contributions induced by initial state radiation are observed and accounted for.

Hence a muon traveling initially perpendicular to the beamline bends first in the ϕ direction while inside the solenoid and then once inside the toroid bends in the z direction.

magnet	inner radius	outer radius	length	peak field in windings	avg. field in volume
solenoid	1.23 m	1.28 m	5.8 m	2.6 T	2 T
barrel toroid	4.70 m	10.05 m	25.3 m	3.9 T	0.5 T
endcap toroids	0.83 m	5.35 m	5.0 m	4.1 T	1 T

Table 4.1: Important parameters of the three ATLAS magnet systems. Specifications from [85].

The eight individual coils which form the barrel toroid are of size 25.3 m by 5.35 m, making the entire structure the largest superconducting magnet system in the world. Each endcap toroid is also formed by eight smaller coils arranged to interleave with the barrel coils. Table 4.1 lists the approximate field strengths permeating the regions defined by columns 2-4. Superconducting magnets are the only feasible way to achieve such large fields over such large volumes. Strands slightly thicker than 1 mm of the superconductor NbTi are reinforced with copper and aluminum to form the current-carrying coils of the ATLAS magnets. At their nominal operating temperatures of 4.2 K and 4.7 K the solenoid and toroid systems conduct 7.73 kA and 20.5 kA of current, respectively. The low temperature is maintained by a liquid helium cryostat fed by supply lines connecting to the surface hall above ATLAS. An array of probes installed inside the solenoid and MS systems measure a map of the magnetic field in these regions which is then used in the path integral in equation 4.2.

4.3 Muon spectrometer

The ATLAS Muon Spectrometer (MS) detects and tracks muons exiting the calorimeters which are deflected within the toroidal magnet systems. The functions of the MS are to trigger on and identify muons, help determine their charge, and ultimately measure their momentum up to and beyond 1 TeV. Since the toroidal magnetic field runs in the ϕ direction it induces track curvature in a *bending plane* spanned by r and z at constant ϕ . For example, the track of a muon with a transverse momentum of 1 TeV accrues a sagitta of roughly $\Delta s = 500 \mu\text{m}$ before leaving ATLAS. A design goal of the MS was to enable the measurement of such a muon's 1 TeV momentum within 10% uncertainty or better. At lower momenta of 10-200 GeV the MS achieves (on its own) a resolution of $\Delta p_T/p_T = 4 - 5\%$ [86].

The MS comprises the outermost and largest subdetectors, being chiefly responsible for the size of ATLAS. There are three concentric MS barrel layers, three endcap layers per side, and a single annular layer to help cover the η range in transition between them. The

innermost MS endcap layers, known as the “small wheels”, have a diameter of approximately 9.3 m and are sandwiched on each end between the endcap calorimeter and the endcap toroidal magnets. Next, right outside each endcap toroid is a stack of four muon wheels called “big wheels” because of their 25 m diameter. Finally the pair of 23 m diameter “outer wheels” are installed a distance of $z = 21.5$ m away from the interaction point. It is ironic that the name “wheel” is used for a structure designed to be completely stationary during operation. More than one detector technology is used to meet performance goals in the different regions of the MS: 2 types of muon “chambers” are used for precise position measurement and 2 types for fast triggering. All four types are gas-based detectors making them suitable for coverage over a large area.

The projection of muon tracks in the bending plane are reconstructed using hits (i.e. energy deposits) in Monitored Drift Tube (MDT) chambers. This is the one type of chamber where the gas volume is segmented. The MDTs cover $|\eta| < 2.7$ except in the forward region ($2.0 < |\eta| < 2.7$) of the innermost MS layer where muons encounter Cathode Strip Chambers (CSCs) instead. The CSCs also provide good spatial precision but are able to handle rates over six times that of the MDTs, making them suitable for the forward environment. Table 4.2 lists the spatial resolution of both types of precision tracking chamber. In order to precisely measure sagitta within the bending plane, MDT chambers are finely segmented in z in the barrel region and r in the endcap. They provide ϕ only coarsely, through the chamber position. The CSCs, which are arranged in an endcap geometry, are likewise finely segmented in r but also in ϕ .

			resolution (RMS)	
type	# chambers	coverage	bending plane (z or r)	timing
MDT	1150	$ \eta < 2.7^\dagger$	$35 \mu\text{m}$	-
CSC	32	$2.0 < \eta < 2.7$	$40 \mu\text{m}$	7 ns
RPC	606	$ \eta < 1.05$	10 mm	1.5 ns
TGC	3588	$1.05 < \eta < 2.4$	2 – 6 mm	4 ns

Table 4.2: Per-chamber resolution of the tracking (upper rows) and timing (lower rows) muon chambers. [†]CSC chambers replace MDT chambers in the innermost layer forward region. In actual measurements these figures are degraded by alignment (spatial) and electronic (timing) uncertainties. Data taken from table 32 in [85].

The lower two rows of table 4.2 correspond to the two types of fast timing chambers used in the MS. Resistive Place Chambers (RPCs) are used for the barrel region ($|\eta| < 1.05$) and Thin Gap Chambers (TGCs) are used for the endcap wheels ($1.05 < |\eta| < 2.4$). The purpose of these chambers is to time muon hits with a resolution of a few ns for triggering and determining which bunch-crossing they come from. This level of timing is not possible with MDTs because of limitations from drift time. In addition the timing chambers provide

a more precise measurement of the track ϕ coordinate which is then used together with the bending-plane coordinate measured by the precision chamber hits.

The TGCs are multiwire proportional chambers with a small (1.8 mm) inter-wire spacing and large nearly 3 kV potential gradient to ensure signals arrive within 25 ns in 99% of cases. In order to provide a bending plane measurement, the wires in each chamber are tangent to the ϕ direction. In front and behind the wires, the gas volume is bounded by two carbon-coated cathode planes, segmented on the backside with copper “pick up” strips. The strips run in the r direction and therefore provide a measurement of ϕ . Two types of TGC modules are used in the different MS wheels: “doublets” and “triplets” which comprise stacks of two or three wire chambers.

The RPCs, for triggering in the barrel, are the only type of chamber not using wires. Instead tracks ionize a gaseous cavity separating two parallel plates coated in a resistive plastic laminate. The resulting charge avalanches induce a signal on a longitudinal array and a transverse array of metal strips, one on the opposite face of each plate.

4.3.1 Closer look at monitored drift tubes

The MDT technology responsible for tracking muons in the bending plane in most of the MS achieves high spatial resolution in combination with large-area capability, mechanical stability, and reliability. There are a total of 656 rectangular MDT chambers in the barrel region and 494 trapezoidal ones in the endcap. Each contains between 12 and 72 aluminum drift tubes packed side by side. The 3 cm tubes are filled with a 93% Argon and 7% CO₂ gas mixture (with a few hundred ppm of H₂O) which is pressurized at a moderate 3 bar (43.5 PSI). Along the centre of each tube a 50 μ m gold-plated W/Re wire is strung and held at a potential of 3080 V. An incident muon traveling perpendicular to the tube ionizes the gas mixture along a chord of the tube circumference. The line of liberated charges drifts radially along the electric field lines between tube and wire, resulting in a spread of arrival times. In the process the large potential difference amplifies the original signal by a gain of roughly 20000, resulting in a 99.8% efficiency for a tube to register a minimally-ionizing muon hit [87].

4.3.2 Muon spectrometer alignment

Muon momentum resolution is affected by a number of variables, including the energy a muon deposits in the calorimeters and the multiple scattering it experiences in the MS (discussed further in section 5.2). Another source of uncertainty, directly affecting the spatial resolution, is that of the imperfect alignment of the more than 1000 muon spectrometer chambers. For the stiffest tracks at high- p_T , relative alignment thus becomes the dominant source of uncertainty in the momentum measurement. To mitigate this impact a precision goal was set at 30-40 μ m for alignment of the MS. It is not difficult to appreciate the

challenge of determining relative positions of muon chambers to within the width of a human hair across the free-standing 23 m diameter outer wheels of ATLAS.

The alignment of the MS is extracted in real-time using an optical alignment system. The optical alignment system comprises a network of over 12000 CCD and CMOS sensors which each observe an illuminated reference target through a lens. Misalignments on the sensor-target span are measured through the distortion of the observed versus reference image. Optical components within a muon chamber measure its local distortion while those connected between chambers reconstruct relative alignment throughout the MS. The alignment reconstructed by the optical system is combined with constraints from track-fitting of data collected in designated (pp and cosmic ray) runs with the toroidal magnetic field turned off. In this case actual chamber positions can be calibrated by exploiting the fact that muon hits in the MS should form straight tracks (aside from multiple scattering, which is averaged out). This procedure has been done most recently for data collected during a special run in 2017 [88].

4.4 Inner detector

At the heart of the ATLAS detector, the first thing that particles encounter beyond the beam pipe is the Inner Detector (ID). The function of the ID is to precisely measure coordinates of particle hits and thereby reconstruct tracks while minimally obstructing them. For charged particles this translates into a momentum measurement via equation 4.2, given that the ID is immersed in the 2 T field of the solenoid magnet. Besides reconstructing particle 3-momentum vectors, information from the ID is critical to resolve neighboring interaction vertices and displaced secondary vertices.

The entire ID defines a cylinder just over 7 m long and 2.3 m in diameter, covering particle tracks within $|\eta| < 2.5$. It is made up of three detector technologies arranged coaxially in the barrel and sequentially in the endcap regions. Figure 4.2 is a schematic of the ID barrel region. The innermost layers of the ID serve the vital role as the nearest point of reference to the primary interaction vertex. Silicon pixels are therefore chosen for this task. This choice of fine granularity requires a large bandwidth which is the main reason they are used only up until $r = 12.25$ cm ($z = 30.92$ cm in the endcap), beyond which the silicon sensors are segmented into microstrips. Finally, the outermost part of the ID, the transition radiation tracker (TRT), is made of gas-filled straw tubes. The ID, and in particular its innermost layers, faces the challenge of operating in the most radiation-harsh region of the ATLAS detector besides the forward calorimeter [90]. The sensors have been

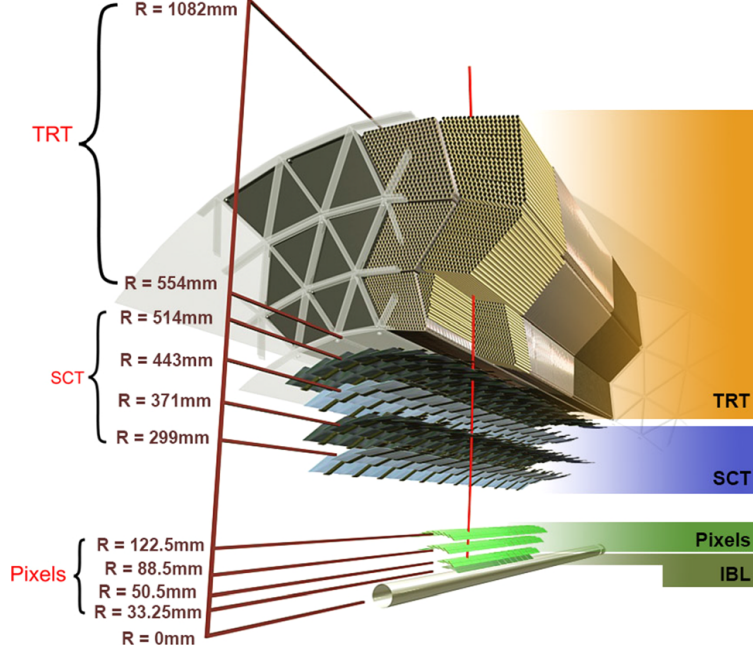


Figure 4.2: Closeup of the ATLAS Inner Detector barrel showing the IBL, Pixels, SCT, and TRT regions. This is figure 1 of [89].

designed to operate up to $2\times$ the anticipated irradiation prior to being replaced by a new inner detector during the LHC Long Shutdown 2⁴.

Both Pixel and SCT detector sensors are based on an n-type silicon bulk. The actual pixels are a pattern of n^+ implants, which deal well with type-inversion from radiation, whereas strips are made of p-type regions. A refrigeration system is integrated into the support structure of the Pixel and SCT subdetectors to cool their Si sensors to around -7°C , with the purpose of reducing noise due to irradiation.

4.4.1 Pixel detector

During the preparations for Run-2 in 2014, the pixel detector was prepended by the *Insertable B-Layer* (IBL) to enhance performance, in particular vertex reconstruction (see [91]). The IBL layout is cylindrical, and with a diameter slightly larger than a pop can (6.65 cm) it lies very close to the LHC beamline. Though it has no endcap modules the IBL contributes to tracking out to $|\eta| < 3$. Fourteen staves (rectangular support structures) run along its roughly 7 m length, each supporting 20 modules which house radiation-hard silicon pixel sensors. The sensors are finely segmented with a pixel size of $50 \times 250\ \mu\text{m}^2$ in $\phi \times z$.

⁴The foreseen degradation of the innermost “B-Layer” of the Pixel detector was partly compensated by the insertion of a smaller layer, the IBL.

Compared to the IBL, the subsequent barrel layers of the Pixel detector are comprised of sensors with a pixel size that is longer in the beam (z) direction: $50 \times 400 \mu\text{m}^2$ in the beam direction. In total, 1744 silicon pixel sensors are distributed over the 3 barrel and 3+3 endcap layers. Table 4.3 compares the number of readout channels for the IBL and Pixel along with the other ID subdetectors.

Subdetector	Sensor	Granularity	Size	Modules	Channels
IBL	Si pixels	$50 \times 250 \mu\text{m}^2$	$0.1 \text{ m}^2 \text{ Si}$	280	12M
Pixel	Si pixels	$50 \times 400 \mu\text{m}^2$	$1.8 \text{ m}^2 \text{ Si}$	1744	80M
SCT	Si strips	$80 \mu\text{m}$ (avg. in endcap)	$60 \text{ m}^2 \text{ Si}$	4088	6M
TRT	drift tubes	7 mm	$12 \text{ m}^3 \text{ vol.}$	(3×32) $+(2 \times 20)$	350k

Table 4.3: Comparison of the ID subdetectors.

4.4.2 SCT: Silicon microstrip tracker

The silicon microstrips in the SCT have a pitch (i.e. inter-segment length) of $80 \mu\text{m}$ in the barrel and ranging between $56.9 \mu\text{m}$ and $94.2 \mu\text{m}$ in the endcap. Moving from silicon pixels to strips dramatically reduces the number of readout channels (from 44M channels per m^2 of Si down to 100k according to table 4.3). In order to allow spatial resolution in the direction parallel with the strips, sensors are housed back-to-back on modules at a small relative stereo angle of 40 mrad. Once installed the strips are oriented along the beam axis in the barrel and roughly radially in the endcap. This arrangement enables each module to have a spatial resolution of $17 \mu\text{m}$ in the bending direction of the solenoid magnetic field (compared to roughly $580 \mu\text{m}$ resolution in the longitudinal coordinate).

4.4.3 TRT: transition radiation tracker

The basic unit of the TRT is similar to that of the muon MDT detectors: gas-filled drift tubes. Because of their smaller 4 mm diameter they are also referred to as “straws”. The essential ingredients for each unit are the Kapton straw coated on the inside with the aluminum cathode (at -1.53 kV), a $31 \mu\text{m}$ gold-plated tungsten central anode wire, and the $70\% \text{ Xe} + 27\% \text{ CO}_2 + 3\% \text{ O}_2$ gas mixture in between. This configuration produces a gain of factor 25000 on the ionization signal. The drift time measurement enables a spatial resolution per straw of $\sim 130 \mu\text{m}$ – much smaller than the inter-straw separation. The barrel region is formed by three sectors each containing 32 modules for a total of 52,544 straws running parallel to the beamline. The endcaps each contain 20 modules in the shape of disks for a total of 122,880 straws aligned with the radial direction.

In addition to large-area precision tracking, the TRT helps identify particles by providing both transition radiation and dE/dx information [92]. When passing through polymer fibres

(foils in the endcap) in the space between the straws, charged particles traverse layers of material with alternating dielectrics which induces the transition radiation. The energy of the radiated photons is dependent on the relativistic factor $\gamma = E/m$. Electrons, with their comparatively small mass, therefore radiate photons with typically higher energy than pions (8–10 keV vs. 2 keV on average [93]), and also radiate more photons on average. The result is a larger ionization signal in the TRT associated with electrons, due to their increased transition radiation losses, which helps distinguish them from pions. The choice of Xenon in the straws is for its high absorption of transition radiation photons which have X-ray wavelengths. A separate high threshold readout is used in addition to a low threshold one to determine which hits include an ionization signal from associated transition radiation.

Information about the rate of energy loss by a particle, dE/dx , is measured using timing information. This quantity is derived from the absolute width of a particle’s ionization signal, which is measured using its “time over threshold” (ToT). In order to measure ToT, the signal from the straw is monitored every 3.12 ns over a period of 3 consecutive bunch crossings to determine the time interval between crossing above and falling below the low threshold. A separation in ToT for electrons and pions reflects their different dE/dx and is most useful for discrimination in the $p_T < 10$ GeV range.

4.5 Calorimeters

Once tracked in the ID, the vast majority of known particles besides muons and neutrinos come to an end by being completely absorbed in the calorimeter system of ATLAS. While tracking detectors attempt to measure particle momenta with minimal energy loss, calorimeters deliberately induce energy deposition of particles in the form of particle *showers*. This is essential for accurately measuring the energy of electrons and, in particular, photons which are invisible to the ID (unless they convert). Calorimeter cells also enable measuring the total energy of jets as a sum over the energy deposits generated by their constituent hadrons. The ATLAS detector uses sampling calorimeters, which means they are composed of layers of absorbing material, which instigate the particle shower, separated by sensitive/active layers which incrementally “sample” the shower energy.

The combination of an electromagnetic (ECAL) and a hadronic (HCAL) calorimeter is used in ATLAS to help distinguish showers produced by objects containing strong color charge (i.e. hadrons) from those which do not (photons and electrons). The ECAL is the next subdetector beyond the ID, and immediately surrounds the solenoid magnet in the barrel region. Then follows the HCAL. These two systems cover up to $|\eta| < 3.2$, having both barrel and endcap regions. Both are segmented in ϕ as well as the longitudinal coordinate (z in barrel, r in endcap) to give information about the direction and spatial distribution of a shower, which is also important for determining whether it was an electromagnetic or hadronic shower. Finally, a separate calorimeter system, the FCAL, is dedicated to covering

the highly active forward region. While the HCAL barrel uses scintillating tiles as the active medium for sampling, the rest of the calorimeter regions of ATLAS employ Liquid Argon (LAr). Figure 4.3 shows the integration of these subdetectors which will be described further in the following sections.

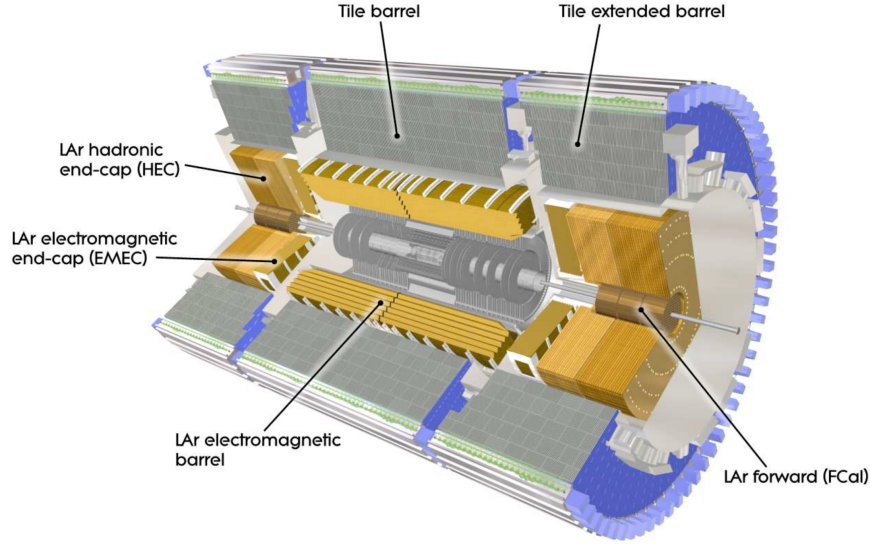


Figure 4.3: The LAr- and tile-based ATLAS calorimeter subdetectors. This is figure 1.3 of [85].

Electromagnetic showers

The energy loss of electrons (and positrons) and photons passing through material is dominated by the following processes, listed in order of increasing energy of the incident particle ([23], section 33.4):

electrons/positrons: positron annihilation, Møller and Bhabha scattering, ionization, bremsstrahlung

photons: photoelectric effect, Compton scattering, e^+e^- pair production

Above 10 MeV, electrons and positrons lose energy predominantly by bremsstrahlung (roughly proportional to energy) while photons by e^+e^- pair production. The interactions are induced by the fields of stationary atoms in the material where they appear as branching events $e^\pm \rightarrow e^\pm \gamma$ and $\gamma \rightarrow e^+e^-$. Successive branches give rise to the tree-like distribution of secondary particles comprising an electromagnetic shower. For a given material, electromagnetic branching can be modelled as occurring stochastically with characteristic length

scales X_0 (*radiation length*) and λ_0 , respectively (following Grupen & Schwartz [94]):

$$\text{bremsstrahlung: } -\frac{dE}{dx} \approx \frac{E}{X_0} \implies E(x) = E_0 e^{-x/X_0} \quad (4.3)$$

$$\text{pair-production: } \frac{dw}{dx} = \frac{e^{-x/\lambda_0}}{\lambda_0} \quad (4.4)$$

where the pair production length turns out to be about $\lambda_0/X_0 \simeq 9/7$, so slightly longer than the radiation length for electrons. But the qualitative picture is clear: (1) after penetrating a distance $\sim X_0$ on average, a high energy electron radiates a photon carrying off a fraction of its energy; (2) the photon travels a distance $\lambda_0 \sim X_0$ before creating an electron-positron pair; the process repeats for each new branch. Because both processes are forward-peaked, the lateral development of electromagnetic showers is mainly due to multiple (Coulomb) scattering, making them characteristically narrow. Each branching epoch represents a successive division of the shower's energy until it drops below some critical energy threshold E_c where absorption processes such as ionization and the photoelectric effect begin to dominate. Since branching drops off rapidly at this stage, it is referred to as the shower maximum. At this stage the number of current particles is approximately $N = 2^{x_{\max}/X_0}$. For $x > x_{\max}$, the electron/positron branches are mostly absorbed while photons will travel for up to an additional nine radiation lengths to deposit 95% of their energy. In the interest of absorbing all the energy of the shower, electromagnetic calorimeters are typically designed to present at least $14 - 16 X_0$ of absorber material (where X_0 is material-specific). It is also important for a calorimeter to be wide enough to enclose the shower's lateral distribution. About 99% of an electromagnetic shower is found to be contained within $3.5 R_M$, where the Moliere radius $R_M \simeq X_0 \cdot (21 \text{ MeV})/E_c$ [23].

Hadronic showers

Because of their relatively larger mass, charged hadrons experience minimal energy losses due to bremsstrahlung. Instead, branching is primarily induced through inelastic hadronic interactions on the nuclei of the absorbing media. The characteristic length scale separating these branching events is the *nuclear interaction length* λ_I which is usually longer than the electromagnetic radiation length for a given material⁵. For example, in lead $X_0 = 0.56 \text{ cm}$ while $\lambda_I = 17.1 \text{ cm}$. As a result, hadronic showers are longer than electromagnetic ones. The nuclear interactions responsible for branching can include large transverse momentum exchange, meaning that hadronic showers are also typically wider than electromagnetic showers.

Pions account for a large fraction of the branches in a hadronic shower, and one third of them on average are neutral pions. They (and other secondary hadrons) complicate the

⁵See table 6.1 of [11]; note the opposite is true for the first four elements.

picture by generating two additional types of branching through their electromagnetic ($\pi^0 \rightarrow \gamma\gamma$) and weak (e.g. $\pi^+ \rightarrow \mu^+ \nu_\mu$) decays. Production of photons above the pair-production critical energy causes a fraction of the hadronic shower to turn into an electromagnetic shower. Weak decays, on the other hand, can cause energy fluctuations that escape the shower in the form of muons and neutrinos. Finally, some of the energy released through hadronic interactions is “invisible” to the calorimeter: low-energy neutrons, delayed gamma radiation, and binding energy in nuclear reactions. These sources are must be accounted for in energy calibration of a hadronic calorimeter⁶.

4.5.1 LAr-based calorimeters

The active medium in the ECAL, the HCAL endcap, and the FCAL is chosen to be LAr for its radiation hardness and charge mobility. All these LAr calorimeters sit in three different toroidal aluminum cryostats, one for the barrel and one for each of the two endcaps. The cryostats are two-walled with a vacuum in-between like a thermos bottle, maintaining a temperature of -183°C .

A calorimeter measures energy by summing over either ionization charge released in the sampling layers or photoelectrons from a scintillating active medium⁷, which have different response but are both proportional to the incident energy E . Since higher energy showers contain more particles ($N \propto E$), they lead to a larger electrical signal and therefore better relative energy measurement. This corresponds to the $\sqrt{E}$ (stochastic) term in the following expression for the relative energy resolution:

$$\frac{\sigma_E}{E} = a \frac{1}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (4.5)$$

where a , b , and c are constants. Electronic noise is a main contributor to the second term. The final term, called the constant term, accounts for things such as imperfections in calorimeter construction and calibration, electronic readout, and losses due to dead material.

ECAL: accordion geometry

The ECAL uses LAr for active material in between absorbing layers of steel-reinforced lead. Instead of lying flat as pancakes, they are folded into an accordion (i.e. triangle wave) geometry which makes for tight coverage in the azimuthal direction. The barrel (endcap) LAr calorimeter can be visualized by taking large planar copper electrode-spacer-lead/steel-spacer sandwiches, folding them in a zigzag pattern with creases along the z (r) direction,

⁶It is also possible to partially compensate for invisible neutron energy by introducing fissionable material such as Uranium to convert it into (visible) photons.

⁷In fact LAr responds in both ways, though only the ionization signal is used in ATLAS.

and then stacking them along the ϕ direction to sweep out a torus. The spacer is a hexagonal grid which the LAr fills and across which runs a 2000 V potential gradient. This is where ionization and charge collection happens during shower sampling. Each sandwich being only several millimetres thick, it takes 1024 layers to wrap around each $\pm z$ half-barrel. The copper electrodes are segmented into 3 layers that run approximately in the direction of the shower depth (except the ECAL endcap has only 2). Furthermore each depth region is segmented in η : layer 1 with 8 times the granularity of layer 2, which has 2 times the granularity of layer 3 (except at edges and for ECAL endcap).

The ECAL barrel is divided in ϕ into 32 modules per $\pm z$ side, each module being instrumented into 3424 readout cells. Each ECAL endcap is made of an inner and outer torus-shaped wheel each divided into eight azimuthal modules. Adding up the modules all around on both sides, the endcap pair contains a total of 1024 folded absorber sheets. The ECAL is at least $22 X_0$ in the barrel and at least $24 X_0$ in the endcap, meaning that in most cases photons and electrons/positrons will be completely absorbed. To accurately measure and account for energy loss of incoming particles prior to showering (i.e. E_0), a thin (11 mm) LAr ionization presampler sits immediately inside the ECAL barrel, with four times the azimuthal granularity. There is also a 5 mm presampler in front of each ECAL endcap which provides needed spatial resolution for distinguishing pairs of collimated photons (from π^0 or possibly boosted Higgs decays) from single photons.

HCAL endcap (HEC): pancake geometry

LAr is used as the active medium in the HCAL endcaps (HECs) because the scintillating tiles used in the HCAL barrel are not suitable for the higher radiation doses there. Unlike the ECAL design, the HEC is based on copper plate absorber layers stacked in a “pancake” geometry. Copper is preferable to lead in the HCAL for its shorter nuclear interaction length λ_I (though in terms of radiation lengths X_0 , lead is about twice as dense). Beginning at $z \simeq \pm 4.25$, a 81.7 cm deep front wheel containing 24 absorber layers is followed by a 96.1 cm deep rear wheel with 16 layers. The layers inside the two wheels are 25 mm and 50 mm thick copper, respectively. The LAr-filled separations of 8.5 mm contain a central electrode used for readout surrounded by outer ones carrying an 1800 V potential. This configuration helps reduce charge buildup. The readout electrodes are segmented into $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ pads for $1.5 < |\eta| < 2.5$ and 0.2×0.2 from there out to $|\eta| = 3.2$. Much of the HEC overlaps in z with the smaller FCAL.

FCAL: honeycomb geometry

Situated 4.7 m from the interaction point and in $3.1 < |\eta| < 4.9$, the FCAL experiences high particle flux which can cause charge over-occupancy in LAr sampling layers. Nonetheless it is important to maintain calorimeter coverage of this region for E_T^{miss} and forward jet (e.g. VBF) performance. Therefore, to shorten ion drift time to about 60 ns, a electrode

rod-in-honeycomb design is used which limits the LAr gap to be 0.27-0.51 mm. An inner layer based on stacked copper plates functions as an electromagnetic calorimeter and two outer layers, based on a tungsten matrix, as hadronic calorimeters. Each layer is 45 cm deep. The three FCAL layers are very dense, weighing 9.6 metric tons per side.

4.5.2 Tile calorimeter

The ATLAS HCAL measures the portion of typically-longer hadronic showers which escape the ECAL. Hadronic showers encounter between $2-4\lambda_I$ nuclear interaction lengths in the ECAL regions and before. The HCAL extends this via the LAr-based HEC by about $9\lambda_I$ in the endcaps and by $7.4\lambda_I$ in the Tile Calorimeter (TileCal) which wraps around the other ATLAS calorimeters as shown in figure 4.3. The TileCal is split into a central barrel region and two extended barrel regions which cover out to $|\eta| < 1.7$. The TileCal is composed of steel absorber plates and scintillating tiles for sampling layers. Stacks of steel and scintillator running in the z -direction are staggered in successive layers in the radial direction, forming a flattened checkered pattern. The TileCal's extent in $z - \phi$ of $(2.6 + 5.8 + 2.6)\text{m} \times 2\pi$ is split into 64 wedge-shaped modules. Steel supports at the back of each wedge are bolted together into a mechanical frame but also provide a flux return for the solenoid magnetic field. Scintillation light is channelled at the $\phi \equiv \text{const.}$ edge of each tile via wavelength-shifting optical fibres ganged into photomultiplier tubes (PMTs) for the readout signal. Since the optical components are subject to radiation damage, their light yield was tested after irradiation with the equivalent of ten years' LHC operation and found to degrade by less than 10%. The TileCal also has independent calibration systems, one based on built-in infrared LEDs and another using a ^{137}Cs source which travels through all its tiles along a network of water tubes [95].

4.6 From detector to dataset

During operation, the ATLAS detector records a dense stream of data which can be likened to that of a 100-megapixel camera capturing several hundred frames per second. This is, however, only a sliver of the information which survives a sophisticated filtering process to select only the most interesting collisions. The design of the ATLAS trigger aims to optimize the usefulness of recorded data with the constraints from finite hardware speed and limited distributed data storage.

4.6.1 Trigger and data acquisition (TDAQ)

Bunches of protons collide inside the ATLAS detector at a rate of 40 MHz. This rate is deliberately far higher than what could be written to permanent storage for subsequent analysis. The reason for this is to facilitate the production of rare processes such as $H \rightarrow 4\ell$ (or potentially $Z' \rightarrow \ell^+\ell^-$). It is therefore the critical role of the ATLAS trigger and

data acquisition (TDAQ) system to ensure that precious bandwidth is prioritized to select and record the small fraction of events which may contain an interesting signature. To qualify, events must pass through first the hardware-based Level-1 (L1) trigger and then the software-based High-Level Trigger HLT, which is implemented in a large dedicated computer farm. The full bunch crossings rate of 40 MHz is reduced by the L1 trigger with an acceptance rate up to 100 kHz⁸ [96], while the inclusive HLT selections reduce this further to about 1.2 kHz on average (in total, 3 bunch crossings per 100,000).

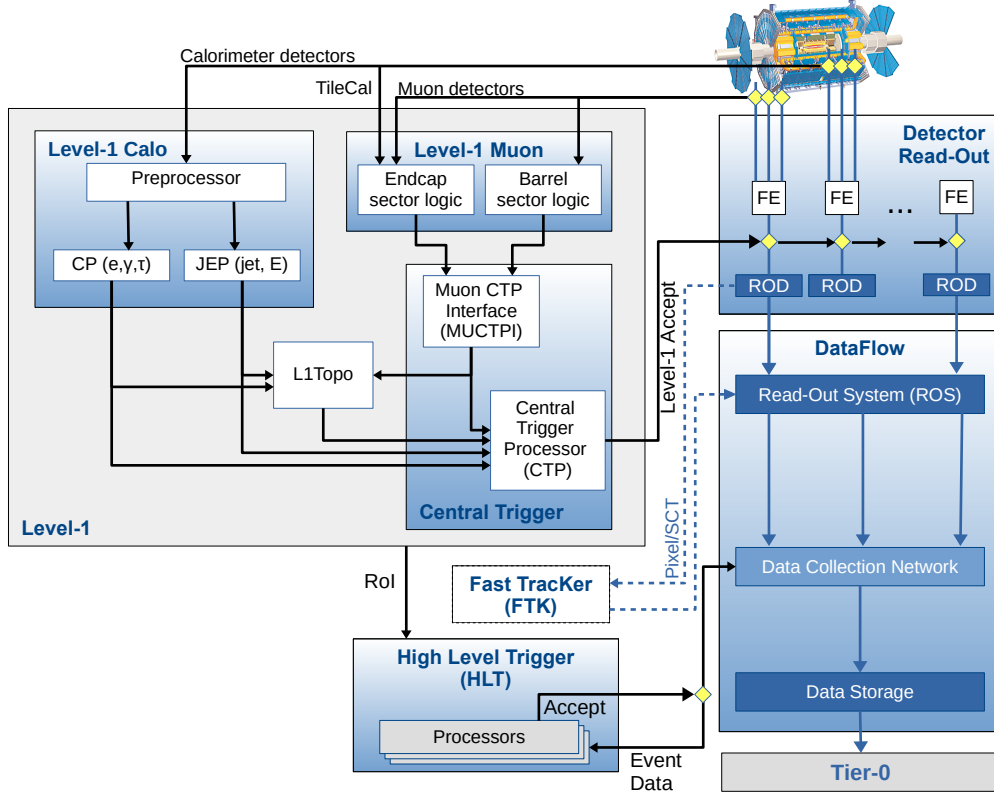


Figure 4.4: Schematic of the ATLAS 2-level trigger system in Run 2. This is figure 1 of [96].

Refer to figure 4.4 for a schematic of the following description. The L1 trigger considers the following information at the full collision rate: the sum of energy deposits in coarse calorimeter trigger towers and hits in RPC (barrel) and TGC (endcaps) muon chambers. Based on only these two sources, the L1 trigger is programmed to identify features of interesting final states: electrons/photons, muons via MS trigger chambers, tau leptons, jets, and $E_T^{\text{miss}}/\Sigma E_T$ in the calorimeter. Analog signals from about 7200 LAr and Tile calorimeter towers are transmitted for digitization in custom electronics implementing the

⁸This improvement on the maximum Run-1 rates of around 65 kHz is thanks to readout upgrades during Long Shutdown 1.

L1Calo algorithm [97]. This hardware is installed underground a short distance from the detector to minimize delays.

The HLT is comprised of two stages: the Level-2 (L2) trigger and the Event Filter. The L2 trigger postpones building a full event by first focusing on select $\eta - \phi$ regions of interest (ROIs) seeded by the findings at L1. Its conclusion based on this partial information determines which events are marked to be fully built for the Event Filter, which makes the final decision whether to keep events for permanent storage. At this point, the full event is checked if it passes the threshold and topology criteria for at least one of multiple HLT trigger “menu” items defined for physics analysis. The two HLT steps are distributed across a computer farm of about 40000 processing units located in building SDX1 above the ATLAS cavern.

Throughout ATLAS subdetectors, Timing Trigger and Control (TTC) information encoding the LHC clock signal and L1 decisions are propagated via interfaces to component readout electronics. If the L1 trigger passes an event, detector signals from the electronics front-end (FE) boards are pushed onto ReadOut Drivers (RODs) before passing through fast ReadOut Links (ROLs) into short-term storage on ReadOut Buffers (ROBs) in the DAQ system. Here, the full detector information waits as the L2 trigger assesses ROIs. If passed, the information is built into an event and handed to the second stage of the HLT, the Event Filter. Events passing at least one branch of the full trigger chain are sent from the HLT farm to the CERN computing centre in four separate streams: a set of physics streams with data passing checks for physics analyses, an express stream containing a subset of physics-quality events for detector monitoring purposes, a calibration stream dedicated to a small fraction events deliberately recorded to study detector/trigger performance, and finally a debug stream storing events with TDAQ errors for diagnostic purposes. The data streams proceed to offline event reconstruction via the CERN computing centre at the Meyrin site. The data rate for the physics stream averages 1.2 GB/s.

4.6.2 The worldwide LHC computing grid (*The Grid*)

The CERN computing centre is Tier 0 of a four-tier global network designed to process and store data from the LHC experiments (and MC samples). At Tier 0, events are passed through a first stage of object reconstruction and copied to backup tape storage. A dedicated network of 10 GB/s optical-fibre links then distributes these data to thirteen Tier-1 facilities around the world, including the Canadian Tier-1 centre hosted on the Simon Fraser University campus. Tier-1 sites serve as backup storage for the raw data and are where reprocessing of data takes place using any updates to reconstruction algorithms. These interface with around 155 Tier-2 sites which provide resources for analysis jobs. Central MC production is also done using Tier 2 resources and supplemented by Tier 1.

The main results shown in this thesis have relied on shared grid resources to parse huge data and MC samples, together several tens of terabytes in size. It would be highly

impractical for the end user to download all these samples, so instead the grid works by sending users' jobs (including full code environment) in parallel to the locations of the samples on Tier 2 sites around the world.

4.6.3 Data formats

Multiple data formats are used at various stages of the data pipeline from raw format to analysis format. The following summarizes some of the main types.

Real events recorded by the ATLAS detector are collected into blocks of definite integrated luminosity (*LumiBlocks*) which define the basic time intervals of data-taking lasting on the order of 1 minute. All the *LumiBlocks* collected during a single fill of the LHC beam (lasting typically a few hours) are grouped in *Runs*. A *Period* contains several tens of runs collected during campaigns of similar data-taking conditions scheduled throughout the long-term LHC runs (e.g. `data17_13TeV.periodK`).

Central MC samples are generated in campaigns with simulated conditions (e.g. pileup profile and detector performance) approximately corresponding to long data-taking intervals within LHC runs. The amount of MC events simulated for each campaign also scales according to the data yields expected for the corresponding intervals. The final MC campaigns for Run-2, are `mc16a`, `mc16d`, and `mc16e`, optimized for comparison with ATLAS data collected during 2015-2016, 2017, and 2018, respectively.

Both real and simulated data exist in a hierarchy of file formats corresponding to stages of the processing chain from raw data to final analysis. The fully-reconstructed format of Analysis Object Data (**xAOD**) is the basic input for physics analyses. For each event, an **xAOD** file stores representations of physical objects (electrons, calorimeter clusters, etc.) and event information (e.g. *LumiBlock*) in classes defined by the central Event Data Model (EDM). The **xAOD** format is designed to be storage-efficient but also versatile and accessible (e.g. readable by ROOT). Central ATLAS **xAOD** samples undergo a final step of selections and corrections (derivation step) before being used for physics analysis. Individual data analysts then use their own code to apply event selection on shared **xAOD** samples to produce flat n-tuples for final results such as histograms.

Chapter 5

Reconstructing electrons and muons in ATLAS

5.1 Reconstructing electrons

The signature of an electron in the ATLAS detector is a track (with potential associated bremsstrahlung) in the ID associated with a cluster of energy deposits in the ECAL.

Tracks are reconstructed starting from 3-dimensional coordinates of hits in silicon layers of the ID. An algorithm (see [98]) then clusters neighboring hits associated with a given seed, accounting for possible significant energy losses due to bremsstrahlung. The track kinematics are extracted from a χ^2 fit to the string of hits. For signal electrons with momenta considered in the dilepton analysis this process is over 98% efficient. Fine-tuning of the electron trajectory is performed in an additional refitting step [99] meant to correct for radiative losses.

The first step in reconstructing the electron calorimeter signature is to identify cells which have a signal-to-noise ratio/significance of ≥ 4 . In this case the noise model includes contributions from both electronic noise and expected pileup. Then their surrounding cells which have significance ≥ 2 are collected recursively to build up a *topo-cluster*. Once no more connected cells can be found which pass the significance threshold, a final layer is added to the topo-cluster using all cells at the perimeter regardless of energy. Additional criteria are also built into the topo-clustering algorithm to handle cases of cell clusters which overlap and the treatment of negative energy fluctuations (see [100] for a detailed description). For a topo-cluster to be considered for electron reconstruction, the fraction of its total energy which is contained in the ECAL is required to be $f_{\text{EM}} \geq 0.5$ since this has very little impact on signal efficiency and eliminates over half of pileup-generated candidates.

The leading topo-clusters in an event which agrees in $\eta - \phi$ with an extrapolated ID track are used to seed the construction of electron *superclusters*. Topo-clusters within a fixed window of the seed cluster are successively amalgamated to form the final electron shower. The final electron's kinematics are determined from the summed transverse energy

of the supercluster following a dedicated calibration of the electron energy scale [101]. Hence electrons' transverse momenta are usually quoted in E_T . On the other hand, an electron's $\eta - \phi$ coordinates are provided by its reconstructed track.

5.1.1 Electron identification

Multiple features of an electron's detector signature can be leveraged to distinguish it from light- and heavy-flavour jets and photon conversions. The ATLAS electron identification algorithm uses 13 parameters (table 1 of [102]) relating to the track and calorimeter signature in a likelihood discriminant. An example is R_{had1} : the ratio of E_T measured in the first layer of the HCAL in the direction of the energy cluster to its total energy, which is typically larger for hadronic showers than for showers from electrons. Track-related parameters include the TRT `eProbabilityHT` based on the deposited transition radiation and the measured relative momentum loss $\Delta p/p$ across the track (larger for electrons on average because of bremsstrahlung). Parameters measuring the matching between the cluster and track are also incorporated, since this distribution is broader for jets and photon conversions. For each parameter a signal distribution S is filled from samples of $J/\psi, Z \rightarrow ee$ -tagged events in data and a background distribution B is filled using a fake-enriched data sample¹. The distributions for each parameter are smoothed, transformed, and finally combined in a joint-likelihood function:

$$L_{S,B} = \prod_{\text{params}} P_i(x_i|S, B) \quad (5.1)$$

The following discriminant is then used to compute a likelihood identification score for a given electron candidate based on its features:

$$\frac{1}{15} \cdot \ln \left(\frac{L_S}{L_B} \right) \quad (5.2)$$

The $1/15$ is just a convenient scaling factor. The three main identification working points supported for ATLAS analyses – `LHLoose`, `LHMedium`, and `LHTight` – correspond to cuts on the likelihood score (5.2) in three different configurations which vary across E_T and η bins.

In addition to the likelihood discriminant, the identification working points require that a candidate electron track be associated with at least seven hits in the silicon layers of the ID. In all except the `VeryLoose` working point, two of the silicon hits are required to be in the Pixel detector. To pass `Medium` and `Tight` one of those two additionally needs to be in the innermost Pixel layer that is available, usually the IBL. This helps reject photon conversions by requiring a charged track back to the interaction point. The silicon hits criteria are summarized in table 5.1.

¹The performance using the data-derived background sample is validated using simulated generic $2 \rightarrow 2$ hadronic samples.

Working point	ϵ_{ID} at 40 GeV	n_{Si}	n_{Pixel}	n_{Blayer}	Note
VeryLoose	-	≥ 7	-	-	For fake electron studies
Loose	93%	≥ 7	≥ 2	-	Used in dilepton analysis Adds E/p , w_{stot} cuts
Medium	88%	≥ 7	≥ 2	1	
Tight	80%	≥ 7	≥ 2	1	

Table 5.1: Example of signal efficiency of electron identification working points and comparison of their requirements on silicon hits.

The **VeryLoose** working point is meant to provide a highly-inclusive baseline selection for studies of fake electron background, such as the analysis described in chapter 7. The signal efficiencies decrease from **Loose** to **Tight** shown in table 5.1 correspond to an increase in background rejection of $\sim 3.5\times$. The working points follow a similar trend in η towards higher signal efficiency in the central region, except for the two most central bins which are somewhat limited due to ID services. Their efficiency also rises in E_T , plateauing around 93% for the **Medium** working point broadly above 80 GeV.

5.1.2 Electron isolation

The presence of other tracks or calorimeter activity in proximity to an electron candidate is one of the characteristics of jets which mimic electrons. Similarly, electrons produced in secondary decays of hadrons or photon conversions typically are not well isolated. The goal of the ATLAS electron isolation algorithm is to reject electron candidates which are suspicious in this regard while avoiding to reject prompt electrons unnecessarily. Isolation is assessed in two ways, track-based and calorimeter-based, both of which are incorporated in the final criteria.

Proximity to the candidate track and calorimeter cluster is measured in terms of ΔR (equation 4.1). A cone defined by $\Delta R = 0.2$ is used to define the “neighborhood” around the candidate wherein additional calorimeter clusters are considered (if their energy-weighted barycentre lies inside). Both ECAL and HCAL cells are included. Of course, the cone includes the candidate itself with most of its energy inside a small core $\eta - \phi$ rectangle which is subtracted. A leakage correction is applied to account for energy of the candidate which falls outside the core rectangle. What the cone also unfortunately includes is activity from pileup and the underlying event (section 2.5) which should not count against the isolation quality. This contribution is therefore estimated and subtracted at the per-event level by constructing a median density out of all other activity within $|\eta| < 5$. The resulting sum of cell energies is the isolation variable $E_{\text{T,cone}}^{\text{isol}}$ which measures additional calorimeter deposits near the electron candidate coming from the same primary vertex.

Because of the better spatial resolution of the inner detector compared to the calorimeters, a smaller cone is defined to capture tracks in the neighborhood of the electron candi-

date. This is especially important to avoid rejecting electrons in boosted event topologies comprised of collimated tracks. A variable cone size is therefore used which shrinks with transverse momentum as $10 \text{ GeV}/p_T$ up to a maximal value² of $\delta R_{\text{max}} = 0.2$ (in place for $p_T < 50 \text{ GeV}$). Within this cone, the transverse momentum of tracks with $p_T > 1 \text{ GeV}$ that meet basic quality criteria (including impact parameter association with the primary vertex) are added up in the isolation variable p_T^{isol} . The candidate track p_T is subtracted together with tracks that appear to originate from bremsstrahlung of the electron candidate.

Finally, a menu of cuts applied to both $E_{T,\text{cone}}^{\text{isol}}$ and p_T^{isol} define nine different electron isolation working points centrally provided within ATLAS [102]. Five “Fix”-type working points cut uniformly on the ratio of the isolation variable to the corresponding candidate variable. On the other hand, “Gradient”-type working points tailor the cut on isolation variables to meet p_T -dependent signal efficiency targets in all η bins. In particular, the **Gradient** isolation working point adopted by the dilepton analysis provides

$$\text{cuts on } p_T^{\text{isol}}, E_{T,\text{cone}}^{\text{isol}} \ni \epsilon_{\text{iso}} = (92.14 + 0.1143 \cdot p_T/\text{GeV})\% \quad (5.3)$$

The efficiency of the combined calorimeter and track isolation cuts are therefore a bit lower: 90% at 25 GeV but rising to 99% at 60 GeV.

5.1.3 Challenges of electrons at high- E_T

Although electron momentum is provided by the ID, the final calibrated value for electrons with large transverse momentum is taken from the calorimeter measurement and therefore called E_T instead of p_T . Equation 4.5 shows how the relative energy resolution of the ECAL improves with energy which leads to the superior $m_{\ell\ell}$ resolution in the electron channel of the dilepton analysis (figure 6.14). In this respect, relative performance improves for high- p_T electrons. Yet electron reconstruction at high- E_T does run into the following challenges:

Shower shape modelling

Part of the discrimination of electrons from jets draws on variables modelling the shape of the calorimeter shower. The electron identification efficiency is therefore partly dependent on how well electron shower shapes are modelled in MC. Imperfections in the modelling can be corrected using centrally-provided scale factors which compare against efficiency measured for electrons tagged in data as coming from dielectron events in a narrow window around the Z resonance peak³. Since this crosscheck is kinematically limited to around $p_T < 150 \text{ GeV}$, efficiency scale factors for higher electron E_T come from extrapolation.

²What is used in the dilepton analysis. Other definitions with $\delta R_{\text{max}} = 0.4$ also exist.

³or J/ψ for lower E_T .

However, for electrons with E_T in the hundreds of GeV the shower shape variables were seen to degrade electron ID signal efficiency. In particular, the simulation of the f_3 variable giving the ratio of ECAL-deposited energy that comes from the third layer was not accurate for high- E_T electrons which have longer showers. Similarly, due to increased electron “punch-through”, the ratio R_{had1} of total E_T coming from leakage into the first HCAL layer was not fully captured in simulation. To account for this, f_3 is removed from the joint likelihood discriminant for $E_T > 80$ GeV for all working points [102]. Motivated by a similar signal efficiency loss, the cut on R_{had1} in the **Tight** working point is also loosened to match that of **Medium** in this E_T range.

Shifts are also derived converting between shower shape variable distributions in MC and the $Z \rightarrow$ data. When comparing the identification efficiency of simulated Drell-Yan samples with and without the shift applied, a roughly flat 1% difference is seen across 100 GeV to 3000 GeV [103]. This is taken as an additional uncertainty on the extrapolated efficiency scale factor for high- E_T electrons.

Charge misidentification

Because tracks stiffen with higher p_T it becomes more difficult to measure curvature in the electron track needed to measure its charge. Muons of the same p_T also face this challenge but benefit from the extended lever arm of combined ID-MS hits. Furthermore, electrons are susceptible to so-called “trident” events. When an electron radiates a bremsstrahlung photon which then pair-produces, it is possible that the resulting positron will be mistakenly associated with the original electron track. If this swap happens within the tracker it can lead to a misidentified charge for the original electron. Since at LHC energies the electron bremsstrahlung cross section increases proportionally with energy, the occurrence of trident events poses a larger problem for high- E_T electrons.

The challenges of electron charge identification at high- E_T makes an opposite-sign requirement in the dielectron analysis inefficient. The absolute impact on signal efficiency worsens from -1.6% for events with $m_{ee} = 150$ GeV to -5.6% at $m_{ee} = 1$ TeV and an additional $\sim -5\%/\text{TeV}$ beyond that [104]. Since the dominant background (Drell-Yan) would be affected no more than the signal, the benefit of requiring opposite sign electrons would only be in reducing sub-dominant backgrounds. Hence, the charge of electron candidates is not checked for the high-mass dilepton analysis.

5.2 Reconstructing muons

As discussed in chapter 4, the MS measures η -like bending of muon trajectories in the ATLAS toroidal magnets. Similarly, muons deflected in the ϕ direction within the ATLAS solenoid are tracked by the ID. Both sources are important for optimal reconstruction of muons. Muon momentum is determined by measuring the reciprocal quantity of the sagitta

$s \propto q/p$ (equation 4.2). Relative q/p resolution of the MS has the following form:

$$\sigma_{\text{rel}}^{\text{exp}}(q/p) = \sqrt{\left(\frac{p_0}{p_T}\right)^2 + p_1^2 + (p_2 p_T)^2} \quad (5.4)$$

The p_0 term is uncertainty on the energy loss experienced by a muon traversing detector material, primarily the calorimeters, which dominates for muons of low p_T . Similarly, changes to the muon trajectory from multiple scattering enter in the p_1 term which overtakes the previous term for intermediate p_T . Finally, the p_2 term which dominates at high- p_T is governed by the MDT single-hit resolution and the alignment accuracy of the MS.

Muon reconstruction relies primarily on ID and MS information, ideally in combination. Calorimeter information also supplements the identification of ID tracks as likely muon candidates, and flag cases of large energy deposits by muons (section 5.2.3). First, segments are built from hits that are spatially-aligned within the segmentation of each muon chamber. Segments which can be at least roughly associated with one another then form the ingredients for the MS component of the muon track. A charged particle trajectory is then χ^2 -fitted to the linked segments using the model of the magnetic field in potentially both ID and MS. Five different “types” are defined corresponding to the algorithms used to reconstruct muon candidates in the ATLAS detector [105]:

CB: “Combined” muons have matching track segments in the ID and MS

IO: “Inside-out” muons are complementary to CB but are seeded by an ID track alone which is then used to search for at least three MS hits lying in a compatible direction

ME: “Extrapolated” muons rely on MS information to extrapolate a muon track to the beamline when an associated ID track could not be found

ST: “Segment-tagged” muons are reconstructed similarly to IO muons but looser, only requiring finding one compatible MS hit

CT: “Calorimeter-tagged” muons rely on energy deposits in the calorimeter to corroborate the ID track

CB and IO muons make use of the most information when it is available and are thus favoured when more than one type is reconstructed. The ID track quality of all but ME muons is improved by the following requirements: $p_T > 2$ GeV (5 GeV for CT), ≥ 1 Pixel detector hits, ≥ 5 SCT hits, ≤ 2 missing silicon hits. The final χ^2 of the track fit is used to gauge the quality of muon reconstruction along with the following two variables which measure tension between the ID and MS momentum measurements (when both are available):

$$\rho' \equiv \frac{|p_{T,\text{ID}} - p_{T,\text{MS}}|}{p_{T,\text{CB}}} , \quad z_{q/p} \equiv \frac{|q/p_{\text{ID}} - q/p_{\text{MS}}|}{\sigma(q/p_{\text{ID}}) \oplus \sigma(q/p_{\text{MS}})} \quad (5.5)$$

Three muon identification working points – *Loose*, *Medium*, and *Tight* – are defined incorporating the ID track cuts and the variables mentioned. While these meet most use cases of muons, two additional selections are recommended for analyses targeting muons with either low or high (>100 GeV) p_T . The high-mass dilepton and lepton+ E_T^{miss} have been the primary use cases in the latter category, so the discussion will focus on the high- p_T muon selection working point and the measurement challenges that motivate it.

The Medium identification working point is suitable for many analyses and also forms a baseline criterion for the high- p_T working point. The q/p significance is required to be $z_{q/p} < 7$ for Medium muons. In the range $0.1 < |\eta| < 2.5$ it restricts muons to be CB or IO with at least two *precision stations*, defined as ≥ 3 clustered hits in a MDT or CSC station. Only one precision station is required for muons in $|\eta| < 0.1$ to make up for a gap in MS coverage. Further out to $\eta = \pm 2.7$, ME muons are also accepted to recover some muons beyond ID coverage. The other two identification working points add to or take away from restrictions of the Medium definition.

5.2.1 The high- p_T muon selection working point

Only muons that pass the Medium identification working point are considered for the high- p_T selection. Then the following conditions are added:

1. muon type must be CB (reconstructed using independent ID and MS segments)
2. muon track must be associated with hits in ≥ 3 precision MS stations (MDT or CSC)
3. MS precision chambers with reduced performance are vetoed (see below)

The three-station requirement improves p_T resolution by approximately 30% for muons with $p_T > 1.5$ TeV at the cost of a 20% decrease in selection efficiency [106]. The main impact of requirement (3) is to exclude chambers in the barrel-endcap overlap region ($1.01 < |\eta| < 1.10$) because of increased uncertainty of their relative alignment. A few chambers that are MDT small sectors of the barrel inner layer (BIS7/8) are also vetoed. An additional 12 chambers installed in the region of the ATLAS feet in 2016/17 (BMG) are also excluded for their larger rate of multiple scattering. Finally, muons traversing BEE (“barrel endcap extra”) chambers are required to have four precision stations because of their increased alignment uncertainty and the unusual magnetic field in this region.

“Good muon” requirement

In addition to the criteria above, high- p_T muons are further refined by a quality standard known as the “good muon” requirement: a muon’s relative q/p uncertainty must be below a p_T -dependent factor times the expected relative resolution:

$$\sigma_{\text{rel}}(q/p) = \frac{\sigma(q/p)}{(q/p)} < C(p_T) \cdot \sigma_{\text{rel}}^{\text{exp}}(q/p) \quad (5.6)$$

where $\sigma_{\text{rel}}^{\text{exp}}(q/p)$ is a model of expected resolution in the form of (5.4). The coefficient $C(p_T)$ is a constant C_0 in p_T until a certain point T after which it decreases linearly (i.e. the cut tightens) down to a minimum value of 1 starting around 5 TeV. The purpose of the good muon requirement is to cut away muons which lie in tails of the p_T response distributions, that is, $(p_T^{\text{reco}}/p_T^{\text{truth}}) - 1$. In particular, large positive tails (i.e. overestimated p_T) are observed for high- p_T muons which pass the criteria 1-3 above but are measured with large relative q/p uncertainty. The coefficient parameters C_0 and T are chosen to reduce the tail fraction of muons with $p_T^{\text{reco}} - p_T^{\text{truth}} > 2$ to negligible levels in simulation.

For the full Run-2 dilepton analysis the cut values of $C_0 = 2.5$ and $T = 1$ TeV were chosen through a reoptimization of the good muon requirement. The updated cut is 1-5% more efficient for signal muons above 2 TeV at the cost of increasing the p_T response tails by fractions of a percent. The selection of high-mass dimuon events gains doubly from this optimization effort.

5.2.2 Muon isolation

Like electrons, *non-prompt* muons can originate from leptonic decays of hadrons. Both calorimeter and tracking information is used to distinguish between prompt and non-prompt muons based on their isolation from surrounding activity. A ΔR cone is used to define the $\eta - \phi$ neighborhood of the muon track wherein the p_T of additional tracks is added up above some lower threshold (either 0.5 or 1 GeV). A similar sum is performed for calorimeter cell clusters in a cone of $\Delta R = 0.2$ around the muon track. The result is corrected for pileup and the underlying event in a similar way as for electrons but there remains some pileup dependence.

A variety of muon isolation working points are defined, all of which use the tracking p_T sum and most also the calorimeter sum. Several working points use a cone size for the track isolation which decreases like $1/p_T$ in specified p_T regions. The choice of `FCTightTrackOnly` for muon isolation criteria in the dilepton analysis was based on an efficiency comparison of the different options, particularly towards high- p_T . As the name suggests, this working point is based only on track isolation and uses a cone with fixed size of $\Delta R = 0.2$.

5.2.3 Challenges of muons, particularly at high- p_T

Muon kinematics cannot be measured based on energy deposits in the calorimeter like electrons. Therefore, muon momentum relies on ID+MS track sagitta and is inherently more challenging to measure at high- p_T because of smaller deflection in the magnetic field. This is again reflected in the dimuon relative $m_{\ell\ell}$ resolution in figure 6.14 which grows with $m_{\ell\ell}$. Other challenges are also discussed below.

Cosmic muons

Muons with momentum at and far above LHC energies are ubiquitous at the earth’s surface because of their creation in the upper atmosphere through cosmic ray conversions. This presents a possibility unique to muons that the ATLAS detector will reconstruct a particle originating from outside, in this case via the cavern shaft. For this to happen, a cosmic muon must traverse the ATLAS detector during the trigger live time (coinciding with a bunch crossing) and pass very near the interaction point. The standard cuts on muon track impact parameter (table 6.4) therefore strongly suppress cosmic muons.

The rate at which cosmic muons are reconstructed has previously been studied [107] by selecting muons in data which have large impact parameter values in a specified range outside the standard cuts. Based on the two selected muon candidates in 236 pb^{-1} of recorded data, and assuming a uniform downward flux, it was estimated that less than $(2.0 \pm 1.4) \times 10^{-4}$ such candidates would happen to fall within the region defined by standard impact parameter cuts. Extending this reasoning to the full Run-2 dataset of 139 fb^{-1} , the cosmic muon background contribution can be estimated at 0.118 events, without taking into account full reconstruction efficiency⁴. This is negligible compared to SM backgrounds from pp collisions but important to keep in mind since a cosmic muon can potentially generate a single candidate event signature with very large mass.

Muon station relative alignment

Slight distortion of muon chambers and relative misalignments between chambers degrade the sagitta (and hence momentum) resolution of the MS. For instance, systematic slight deformation of the MS geometry occurs because of gravitational sag, introducing a ϕ -dependent bias in the track sagitta (and hence p_T) measurement.

The MS design goal of 10% relative p_T resolution for 1 TeV muons translates to a position uncertainty of $\pm 50 \text{ } \mu\text{m}$ on the total $\sim 500 \text{ } \mu\text{m}$ sagitta for such a muon [88]. This means, for example, understanding the relative alignment of MS chambers in the 26m-diameter ATLAS outer wheel to within the width of a human hair.

In practice the performance of the MS varies with η and ϕ because of inhomogeneity in chamber coverage and the toroidal magnetic field. This is also true of the impact of MDT chamber alignment on the sagitta which ranges from $28 \text{ } \mu\text{m}$ in the BA large sector (in barrel between the magnet coils) to $121 \text{ } \mu\text{m}$ in the BEE mentioned above [88]. The overall averaged performance meets the MS design goal but is the main bottleneck for measuring muons with high p_T .

⁴Furthermore, the increased centre of mass energy and beam luminosity of collisions in Run 2 yield a larger muon production rate than in Run 1, while the cosmic muon flux has remained the same. For this reason, the simple estimate here is expected to be conservative.

MS sagitta bias

As discussed in section 4.3.2, the MS alignment is calibrated using a track-fitting analysis of no-magnet data runs. By selecting only $Z \rightarrow \mu^+ \mu^-$ events, alignment-induced bias was observed in the variation of the reconstructed mass between different $\eta - \phi$ regions of the MS. The fact that the bias also showed a clear dependence on the charge of the muon in a given region indicated that the effect was related to weak modes in the alignment calibration. Weak modes are degrees of freedom of the geometry which are poorly constrained by the track-fitting minimization procedure.

Muon catastrophic energy loss

Muons with high- p_T are more affected by energy loss via bremsstrahlung which rises significantly in the multi-TeV momentum range. The radiated photons show up as energy deposits in the calorimeters close to the extrapolated muon track. By potentially creating a sizable difference in the muon momentum measured by the ID and the MS, instances of muon “catastrophic” energy loss can evade the matching criteria checked during muon reconstruction. The reduction in reconstruction efficiency in such cases has been studied and found to be substantial, on the order of 5.5% per 100 GeV of energy lost [108]. However, a simulated sample of a 1 TeV dimuon resonance revealed that energy losses of > 50 GeV were experienced by only 1.4% of the muons. Therefore the actual impact on analysis results is considered negligible given the larger uncertainties on muon selection efficiency which exist in the high-mass tail.

Chapter 6

Dilepton event selection

The new physics scenarios discussed in chapter 3 represent potentially new ways for protons to collide, interact, and produce a pair of high- p_T leptons inside the ATLAS detector. But far from being a smoking-gun signature of new physics, dilepton events arise from multiple processes within the SM itself. Both resonant and non-resonant BSM signatures in the dilepton final state are mediated by a new virtual particle which means that on the event level it is strictly impossible to distinguish from the quantum superposition with processes involving SM virtual particles. It is only on the level of kinematic distributions that BSM signatures in the dilepton final state are generally distinguishable from their SM counterparts. Even here, differences between the dominant SM background and BSM distributions is often limited to a few kinematic variables. For example decays of a virtual new boson with spin 0 or 2 would generate a different angular distribution of the final state leptons than expected for purely SM processes involving spin-1 bosons. However if the BSM particle is an additional massive spin-1 boson, angular information is no longer helpful, and the signature appears only in kinematic distributions. For a neutral mediator this is most clearly revealed in the dilepton invariant mass, defined in equations 2.51 and 2.57:

$$\begin{aligned} m_{\ell\ell}^2 &= P_\mu P^\mu = (P_1 + P_2)^\mu \eta_{\mu\nu} (P_1 + P_2)^\nu \\ &= (E_1 + E_2)^2 - |\vec{P}_1 + \vec{P}_2|^2 \end{aligned} \tag{6.1}$$

$$\underset{E \gg m}{\approx} 2p_1^T p_2^T [\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)] \tag{6.2}$$

6.1 Standard Model background processes

Events containing a pair of same-flavour electrons or muons in the final state originate from several processes at the LHC. For invariant masses in the hundreds of GeV and up, the dominant SM process is that of neutral Drell-Yan (DY) production [109] shown in figure 6.1.

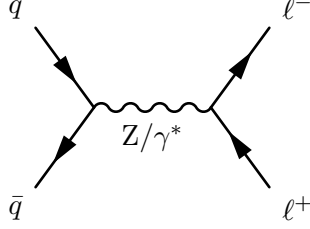


Figure 6.1: Feynman diagram of Drell-Yan process.

The cross section is comprised of two components: a resonant contribution mediated by a Z boson, peaked at its pole mass, and a non-resonant contribution mediated by an off-shell photon. The total DY cross section is shown in figure 6.2 together with its decomposition into Z and photon components. Beyond the Z boson pole, the separate matrix elements for both processes drop steeply for larger values of $m_{\ell\ell}$. Nonetheless its simple s-channel topology allows the DY process to dominate lepton pair production at high mass. Independent from these falling scattering amplitudes, the DY cross section also drops with s because of the decreasing probability to find initial state quarks with large momentum fraction x . This probability is encoded in the Parton Distribution Function described in section 2.5.1.

A small fraction of the DY contribution comes from $Z/\gamma \rightarrow \tau^+\tau^-$ events where each of the two tau leptons decay leptonically to the same flavour. This is a background in both dielectron and dimuon channels which is roughly an order of magnitude smaller than the W -associated top background. In the mixed-flavour $e\mu$ final state this background becomes important relative to the other backgrounds. But for example, in the ee final state it is suppressed by the following factor:

$$\underbrace{(\text{BR}_{\tau \rightarrow e \nu_e})^2}_{\sim 18\%} \cdot \frac{1}{4} \simeq 0.0079 \quad (6.3)$$

That is, this component will be a fraction of 0.8% the size of the $Z/\gamma \rightarrow e^+e^-$ background. The tau lepton’s branching fraction to a muon is closer to 17%, but the outcome is similar. The relatively small ditau DY component is incorporated in the “ Z/γ^* ” histogram for each channel shown in the data-MC comparison in figure 6.9.

Diagrams involving top quarks comprise the second largest SM contribution to the high-mass dilepton final state. Among these, top quark pair production ($t\bar{t}$) as shown in figure 6.3 produces the majority of lepton pairs. Top quark production in association with a W boson (figure 6.4a) can also produce two isolated leptons but at tree level this requires a bottom quark in the initial state. The last and smallest component comes from single top production in the s- and t-channel diagrams shown in figure 6.4. In these diagrams one lepton comes from the top quark decay and the other from the non-prompt decay of the B-meson associated with the b quark in the final state or from gluon splitting leading to another b quark in the initial state. The B-meson decays leptonically with around 34% branching

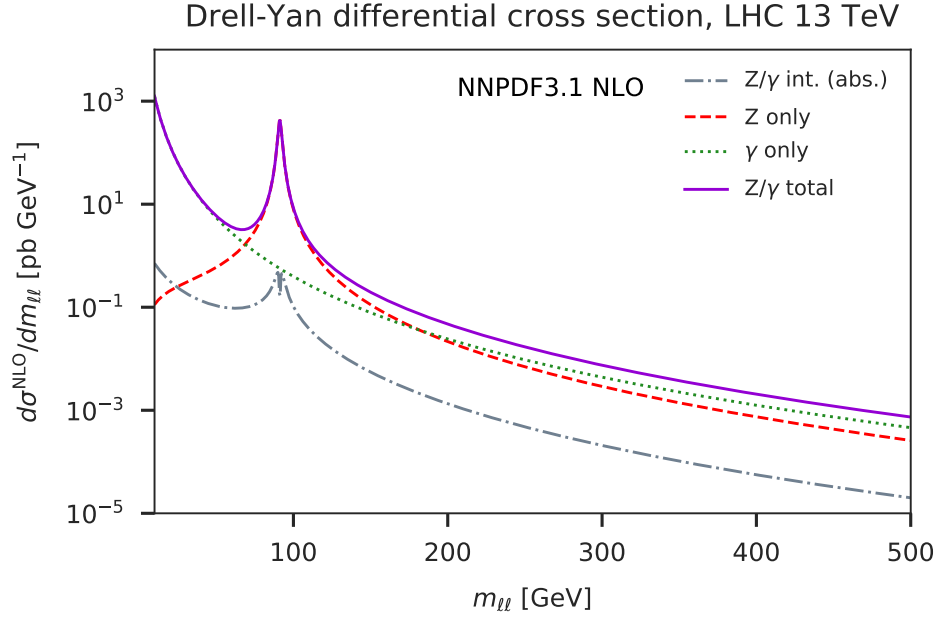


Figure 6.2: Differential cross section for full Drell-Yan and Z , γ components separately, as well as their interference term.

fraction and the resulting lepton is often rejected by isolation requirements. Therefore, single top yields the smallest contribution among the dilepton background processes considered in this analysis.

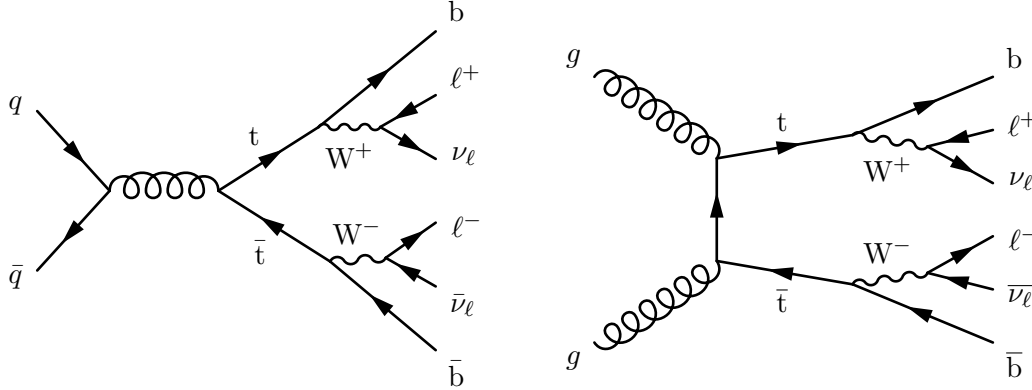


Figure 6.3: The two dominant $t\bar{t}$ diagrams contributing to the dilepton final state. The initial state quarks on the left diagram can be substituted for gluons.

The third-largest category of SM contribution to the high mass dilepton final state is from production of two electroweak bosons with leptonic decays. Representative diagrams of this diboson component are shown in figure 6.5. Each boson can be one of W or Z and decays either hadronically or leptonically. In all, this leads to the fourteen combinations in table 6.1, ignoring diagrams which differ only by exchanged boson labels. The last three

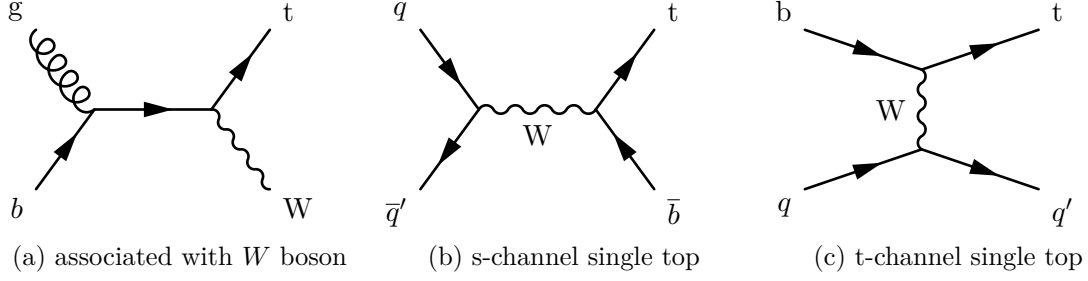


Figure 6.4: Top quark production diagrams contributing to the dilepton final state.

columns in table 6.1 correspond to the six combinations which produce two or more final state leptons (not counting neutrinos).

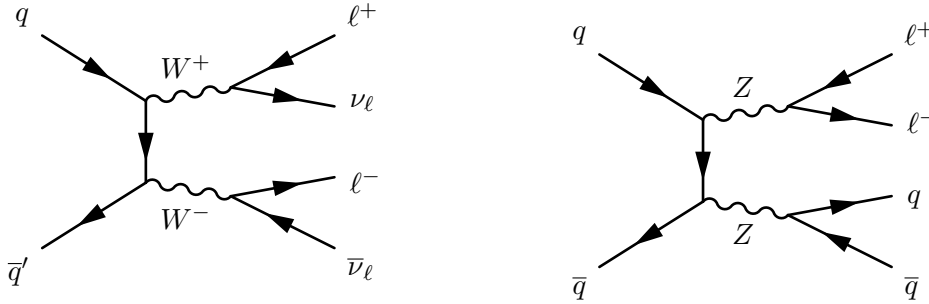


Figure 6.5: Diboson production resulting in two leptons via either a W or Z pair.

bosons	0ℓ	1ℓ	2ℓ	3ℓ	4ℓ
WW	$qqqq$	$\ell\nu qq$	$\ell\nu\ell\nu$		
WZ	$qq\nu\nu, qqqq$	$\ell\nu\nu\nu, \ell\nu qq$	$qq\ell\ell$	$\ell\nu\ell\ell$	
ZZ	$\nu\nu\nu\nu, qq\nu\nu, qqqq$		$\ell\ell\nu\nu, qq\ell\ell$		$\ell\ell\ell\ell$

Table 6.1: Possible decay modes of diboson processes in terms of number of final state (non-neutrino) leptons. Note that in some cases the occurrence of $\ell\ell$ may be $e\mu$ etc. The last three columns contribute to the dilepton final state.

Processes involving not two but three electroweak bosons can also have a final state with two or more leptons. However the cross section of such triboson processes is considered negligible against the larger SM contributions to the dilepton spectrum.

The final SM contribution incorporated into the high-mass dilepton analysis arises from events where one or more objects are mistakenly identified as electrons originating from the hard interaction. The so-called fake electron background is discussed in detail in section 7. The large cross sections of hadronic processes compared to true dilepton processes cause the fake electron background to be comparable in size to the diboson background. It is also possible for muons to enter the selection which are non-prompt or are reconstructed

by mistake from unassociated tracks. However the detector signature of a high- p_T muon is inherently less ambiguous than that of an electron. Impact parameter cuts and isolation criteria further reduce the occurrence of muon “fakes”. They are therefore more rare than electron fakes and are not included in the sum of MC backgrounds compared with data (although a dedicated study was done to validate this choice).

6.2 Simulation of dilepton processes

A variety of MC event generators, parton shower engines, and PDFs were used to model background and signal processes in the dilepton final state as shown in table 6.3. All samples are generated centrally by the ATLAS physics modelling group, with requests and feedback from the analysis teams who share them. Event generation was done in three separate campaigns¹ and reconstructed using separate subversions of the ATLAS software (Athena release 21) as shown in table 6.2. The full ATLAS detector simulation in GEANT 4 was used for all samples instead of the faster AF2 calorimeter parameterization (section 2.5.3). The latter was studied and found to slightly mismodel kinematic distributions for high-mass signal samples, mostly affecting the high- p_T tails.

campaign	corresp. data-taking period	reco. tag	Athena release
mc16a	2015-2016	r9364	21.0.20
mc16d	2017	r10201	21.0.53
mc16e	2018	r10724	21.0.77

Table 6.2: Specifications of the three MC campaigns used to model the full Run-2 dataset.

Background MC samples

Drell-Yan samples used for the background are generated with POWHEG+PYTHIA at NLO in QCD with the CT10 PDF set [110]. To maintain smoothness of the dominant DY MC distribution all the way from the Z peak down 5 orders of magnitude to the high-mass tail, the DY event generation is split into 1+19 slices of $m_{\ell\ell}$ at truth-level. The first bin integrates the Z resonance peak. An inclusive sample, on the other hand, would reduce sensitivity to data excesses in the multi-TeV region because of its falling statistical precision towards high mass. Details of the size and cross sections of the NLO DY samples in each MC campaign are provided in table A.1 in appendix A. The PHOTOS algorithm [111] was used to apply corrections to DY samples due to final state QED radiation.

¹The exception being the SHERPA diboson sample where the mc16d sample was duplicated and normalized to stand in for the mc16e ones which are missing slices.

POWHEG+PYTHIA is also used to generate the top quark backgrounds ($t\bar{t}$, Wt , single top) interfaced with the newer (than CT10) PDF set, NNPDF3.0NLO [29]. A dilepton filter is applied in generation of the $t\bar{t}$ sample which is used. The single top t -channel sample is generated using the 4-flavour (u, d, c, s) PDF scheme [112], and a dedicated treatment of spin correlation between the top and bottom quarks in this case is handled using the MADSPIN plugin [113]. The normalizations of the top quark samples are obtained at NNLO using TOP++2.0 [114].

The contribution from diboson processes are modelled with the SHERPA 2.1.1 event generator [35] and the CT10 PDF set. The three $VV = \{WW, WZ, ZZ\}$ processes are simulated inclusively in four groups of lepton multiplicity: one each for the 3ℓ and 4ℓ columns of table 6.1, and separate ones for $qql\ell$ and $\ell\nu\ell\nu$. Samples are distinct for the dielectron and dimuon channel. Similar to the DY samples, the diboson background samples corresponding to each lepton multiplicity are also generated in 8 bins of truth-level $m_{\ell\ell}$, with the following boundaries (in GeV):

$$[50, 150], [150, 500], [500, 1000], [1000, 2000], [2000, 3000], [3000, 4000], [4000, 5000], [5000, \infty]$$

All samples with PYTHIA parton shower also are interfaced to the EVTGEN afterburner which models decays of c and b hadrons [115].

Signal MC samples

An independent set of DY samples designated for signal reweighting (section 9.1) are generated at LO using PYTHIA v8.186 [116] and the NNPDF23LO PDF set [117]. They are also mass-sliced. Both these LO DY samples and the NLO ones for background modelling are bootstrapped to a common NNLO (in QCD) scheme using higher-order K -factors provided by external packages (section 6.2.1).

Benchmark Z'_χ samples used for illustration in figure 6.9 are likewise generated in PYTHIA 8 (v8.210) interfaced to NNPDF23LO. A similar configuration is used to generate test samples of spin-1 dilepton resonances of negligible width (to validate transfer function results).

Finally, LO signal samples were generated by request for studying selection efficiency in different spin cases. The RS graviton also used the PYTHIA 8 and NNPDF23LO configuration like the other signals, while the spin-0 MSSM Higgs samples were based on POWHEG+PYTHIA and CT10.

Pileup collisions

The presence of multiple pileup collisions (section 2.5.2) leads to important effects in data which need to be accurately captured in simulation. Such collisions are generated with PYTHIA v8.186 and the NNPDF23LO PDF set and superimposed with each hard-scatter

event in the MC samples above. The pileup conditions are generated within the A3 tune of the PYTHIA generator derived from measurements of inelastic scattering events in ATLAS Run-1 and early Run-2 data [118].

Background Process	ME Generator and ME PDFs	PS and non-perturbative effect with PDFs
NLO Drell-Yan	POW-B, CT10, PHOTOS	PYTHIA v8.186, CTEQ6L1 [119, 120], EG1.2.0
$t\bar{t}$	Pow-B [121, 122], NNPDF3.0NLO	PYTHIA v8.230, NNPDF23LO, EG1.6.0
Single top s -channel, Wt	Pow-B, NNPDF3.0NLO	PYTHIA v8.230, NNPDF23LO, EG1.6.0
Single top t -channel	Pow-B, NNPDF3.04fNLO, MADSPIN	PYTHIA v8.230, NNPDF23LO, EG1.6.0
Diboson (WW , WZ , ZZ)	SHERPA 2.1.1, CT10	SHERPA 2.1.1, CT10
Signal Process		
LO Drell-Yan	PYTHIA v8.186, NNPDF23LO	PYTHIA v8.186, NNPDF23LO, EG1.2.0
Randall-Sundrum $G^* \rightarrow \ell\ell$	PYTHIA v8.210, NNPDF23LO	PYTHIA v8.210, NNPDF23LO, EG1.2.0
MSSM $gg \rightarrow H \rightarrow \ell\ell$	Pow-B, CT10	PYTHIA v8.212, CTEQ6L1, EG1.2.0

Table 6.3: Event generators and their matrix element (ME) and parton shower (PS) configurations used for the full Run-2 dilepton analysis. POWH-B stands for POWHEG-BOX and EG stands for EVTGEN. Table modified from [1].

6.2.1 Higher-order K -factors for correcting Drell-Yan MC

Since the DY process dominates the dilepton spectrum, a dedicated effort was made to improve the precision of the DY samples using (truth) mass-dependent K -factors. The K -factors measure the ratio of the nominal DY cross section in MC σ_{MC} to values obtained by cross section calculations σ_{HO} incorporating additional processes and using updated PDF sets:

$$K(m_{\ell\ell}^{\text{tru}}) = \frac{\sigma_{\text{HO}}(m_{\ell\ell}^{\text{tru}})}{\sigma_{\text{MC}}(m_{\ell\ell}^{\text{tru}})} \quad (6.4)$$

The following three effects are treated in separate K -factors as functions of $m_{\ell\ell}^{\text{tru}}$:

Higher-order QCD corrections: Two sets of K -factors are derived to bring both the background and signal-like DY samples up to NNLO in QCD:

- background (POWHEG+PYTHIA): NLO in QCD, CT10 $\rightarrow$ NNLO in QCD, CT14
- signal-like (PYTHIA 8): LO in QCD, NNPDF23LO $\rightarrow$ NNLO in QCD, CT14

The correction also brings the two sample types into alignment under the common G_μ EW parameter scheme of the packages used, FEWZ 3.1.b2 [123] and VRAP v0.9 [124]. Since the LO $\rightarrow$ NNLO entails two orders, the corresponding K -factors are as large as +40% (around 2 TeV) and are positive up until around 5.5 TeV. The NLO $\rightarrow$ NNLO remains within $\pm 20\%$ up to about 6 TeV.

Additional EW corrections: NLO EW diagrams not included in the PHOTOS FSR afterburner are calculated using SANC [125]. Effects include QED ISR, interference between ISR and FSR, and Sudakov corrections at the single-loop level. This contribution comes in the form of an additive correction at each order in QCD which is then converted to a multiplicative K -factor.

Photon-induced contribution: Initial-state photons arising from the protons can pair-produce the dilepton final state. The photon-induced dilepton differential cross section is calculated at LO QCD using the MRST2004QED global parton analysis [126]. The more recent package SANC was also used to cross-check the values. The additive contribution is rolled into the DY nominal samples through a dedicated K -factor.

The `LPXSignalReweightingTool` framework was created to store the K -factors centrally and apply them on an event-by-event basis.

6.3 Selection criteria for high-mass dilepton analyses

The search for new dilepton signatures is performed in a signal region (SR) of data defined by several object- and event-level selection criteria, summarized in table 6.4. The SR is comprised of two categories, interchangeably labeled as the electron/dielectron/ ee channel and the muon/dimuon/ $\mu\mu$ channel.

The event selection is accomplished in two main steps, programmed in two packages: `DileptonST` which outputs primary n-tuples by applying pre-selection criteria to derived xAOD samples, and `DileptonNTR` (“n-tuple reader”) which applies final selection to create smaller, secondary n-tuples, histograms, and data-MC comparison plots. The first package serves as an interface to the `SUSYTools` framework which is shared among many analysis teams and centrally maintained to apply recommendations from ATLAS performance groups.

6.3.1 Event-level criteria

The events of interest contain at least two high- p_T electrons or muons. This basic requirement is met in samples pre-selected under the EXOT0 derivation xAOD format which requires:

EXOT0 upstream criteria

- electrons: $p_T > 20$ GeV and at least LHLoose identification
- muons: $p_T > 20$ GeV, type is {CB, IO, ME}, and “good muon” requirement
- $\#e + \#\mu \geq 2$

Data samples produced in this format are then filtered using the central “Good Run List” (GRL) for each year to include only the LumiBlocks collected with nominal beam and detector conditions. Within otherwise healthy LumiBlocks, individual events may also be flagged due to detector problems, and this is checked at the “event cleaning” step.

	Criterion	Electrons	Muons
track/kinem.	Transverse momentum	$p_T > 30 \text{ GeV}$	
	Pseudorapidity	$(\eta < 1.37)$ $ (1.52 < \eta < 2.47)$	$ \eta < 2.5$
	Transverse IP	$ d_0/\sigma(d_0) < 5\sigma$	$ d_0/\sigma(d_0) < 3\sigma$
	Longitudinal IP	$ z_0 \sin \theta < 0.5 \text{ mm}$	
object	Isolation	Gradient	FCTightTrackOnly
	Identification	LHMedium	high- p_T selection
	Quality	veto dead HV regions (data <i>only</i>)	$\sigma_{\text{rel}}(q/p) < \text{“good muon”}$ threshold
	Calibration	es2017_R21_v1 (ESModel)	Recs2018_05_20, SBC [†] on MC <i>only</i>
		Event	
event	Dataset	pass 2015-2016, 2017, 2018 GRLs and event cleaning	
	Trigger	lowest unprescaled $2e$ HLTs	lowest unprescaled 1μ HLTs
	Multiplicity	$(\geq 2e) (\geq 2\mu)$ (if passes both, select as $\mu\mu$ event)	
	Overlap removal	lepton-favoured WP	
	Charge	not checked	opposite sign
	Pair	2 highest- p_T candidate leptons	
	Mass	$m_{\ell\ell} \geq 130 \text{ GeV}$ (search region is $m_{\ell\ell} \geq 225 \text{ GeV}$)	

Table 6.4: Summary of selection criteria for the signal region of full Run-2 high-mass dilepton analysis – see text for description. [†]SBC stands for “sagitta bias correction”.

The dielectron selection uses **2eX**-type high-level triggers which fire on events containing two electrons with E_T above threshold **X**. The threshold must be lower than the SR requirement $p_T > 30 \text{ GeV}$ to avoid turn-on effects. The chosen triggers listed in table 6.5 are the ones operating for each year with the lowest threshold **X** available for electrons with looser identification and isolation requirements than the SR. For the dimuon selection it is more efficient to use **muX** single-muon triggers since these are available at p_T thresholds below that of the SR requirements (due to the lower production rate of high- p_T muons than candidates for electron triggers). The exception is **mu50** (which also drops the isolation requirement), but in each year it is combined with the **mu26_*** triggers which are efficient to lower p_T . MC events are likewise required to pass the same triggers in simulation.

Once candidate leptons (see section 6.3.2) have been identified within an event, it is first checked that there are at least two electrons or two muons. If two muons are to qualify for a candidate pair their measured charges must have opposite sign since the search only considers neutral mediators. For electrons, due to increased rate of charge misidentification

channel	high-level trigger	2015	2016	2017	2018
ee	2e12_lhloose_L12EM10VH	✓			
	2e17_lhvloose_nod0		✓		
	2e24_lhvloose_nod0			✓	✓
$\mu\mu$	mu26_imedium	✓			
	mu26_ivarmedium		✓	✓	✓
	mu50	✓	✓	✓	✓

Table 6.5: Triggers used in the full Run-2 high-mass dilepton analysis. Triggers which are available for the same year are combined in a logical OR.

at high- E_T , the benefit to purity of the opposite charge requirement is outweighed by the loss of selection efficiency (discussed in section 5). The possible combination which associates the two highest- p_T leptons is then selected as the candidate dilepton pair. The invariant mass of this pair must be greater than 130 GeV to enter the SR. The region above 225 GeV is then used in the statistical analysis, and the first pole mass hypothesis tested is 250 GeV.

6.3.2 Lepton-level criteria

Candidate leptons are required to have p_T (i.e. E_T for electrons) above 30 GeV to sufficiently clear the lowest unprescaled triggers in 2017-2018.

Electrons which shower inside the two transition/“crack” regions ($1.37 \leq |\eta| \leq 1.52$) between the ECAL barrel and endcaps are known to have degraded resolution due to the large amount of material preceding the first LAr sampling layer (5-10 X_0) [100]. The benefit to selection efficiency of including these electrons in this analysis was studied against the accompanying degradation of invariant mass resolution. It was found that although about 7% more events enter the dielectron selection when including the crack region, they broaden the invariant mass response distributions by 4-5%. Furthermore, when considered separately the response distributions of events with one or more crack-region electron candidates exhibit long tails, indicating a larger degree of mismodelling. In the future it will likely be worthwhile to include these events in a separate signal category, but it was decided to exclude them from this analysis since they are not properly modelled within the one SR.

Otherwise, the lepton pseudorapidity acceptance extends out to the precision tracking coverage of $|\eta| < 2.5$ (for electrons slightly less to contain the shower inside the 3-layer ECEC region). Muons are accepted within this range except for the barrel-endcap overlap region ($1.01 < |\eta| < 1.10$) which is precluded by the upstream high- p_T selection working point.

Events are required to have at least one primary vertex, with at least three associated tracks. Among the typically multiple vertices, the one with highest Σp_T^2 of associated tracks with $p_T > 0.5$ GeV is chosen to be the *primary* vertex (PV) assigned to the hard scatter process. Candidate leptons must have tracks that reasonably originate from this PV. This is

gauged using the track’s transverse and longitudinal impact parameters, defined as follows: First, define $\mathcal{C}$ as the point of closest approach of the extrapolated track to the beamline.

- d_0 : distance between $\mathcal{C}$ and V (vertex of interest) in the transverse ($r - \phi$) plane
- z_0 : distance between $\mathcal{C}$ and V in the longitudinal (z) direction

The cuts $|d_0/\sigma(d_0)| < 5\sigma$ ($< 3\sigma$ for muons) and $|z_0 \sin \theta| < 0.5$ mm are standard requirements of lepton-vertex association. The $\sin \theta$ factor accounts for the geometric increase of z differences for tracks inclined away from the transverse plane. These impact parameter cuts are particularly effective for excluding non-prompt and cosmic muons which only rarely associate with the PV.

Central quality criteria are applied to the leptons according to ATLAS-recommended performance standards. As discussed in section 5.1.2, isolation criteria target non-prompt electrons from photon conversions or heavy flavour decays. Electrons must pass the Gradient isolation working point and Muons the FCTightTrackOnly one. These working points were chosen for consistency with the 36.1 fb^{-1} analysis (see section 6.3.3). Both isolation working points maintain an efficiency greater than 99% for signal electrons and muons with $p_T > 60$ GeV [102, 105].

As discussed in section 5.1, the identification quality of electron candidates are ranked according to a joint likelihood discriminant based on multiple feature distributions. The LHMedium identification working point is chosen and is commonly used among analyses that do not have strict requirements on electron purity. As such it is able to select signal electrons with an efficiency above 92% for $E_T > 80$ GeV. An additional standard quality requirement vetoes electrons with clusters in regions affected by dead high-voltage or electronics in known masks (“BADCLUSELECTRON”) and separately in the EMEC during periods of 2016 data-taking.

To enable reliable measurements of muons with momentum in the hundreds of GeV, the high- p_T muon selection working point is used in this analysis. Section 5.2.1 describes the requirement. The selected muons must also pass the “good muon” q/p relative uncertainty threshold also discussed there. These muon selection criteria lead to suboptimal efficiency at low $m_{\mu\mu}$ but are necessary to ensure that events with the highest $m_{\mu\mu}$ of all ATLAS analyses, now in the multi-TeV regime, can be included reasonably safely. The leading and subleading muons in the ten dimuon events with $m_{\mu\mu} > 2$ TeV selected in the full Run-2 dataset were also individually inspected. It was verified that the independent measurements of each muon’s p_T from the MS and the ID were consistent within uncertainty.

During the analysis work a prototype version of the high- p_T working point was under development which relaxes the 3-station requirement to 2-station under certain conditions. For example, candidate 2-station muons must have the missing station in-between the two others because these have a longer lever arm. This optional extension is reflected in the finalized high- p_T working point documented in [105]. When studied using Drell-Yan samples,

the selection efficiency was seen to increase by 5 – 6% (absolute) across the entire $m_{\ell\ell}$ spectrum when adding in the 2-station muons. However, despite additional quality cuts for the 2-station muons, they were seen to lengthen the tails of the q/p distribution. The main concern was the increased probability to select very high-mass events from these tails which have mismeasured momentum. It was therefore decided to allow more time to ensure the quality of these 2-station muons which would likely benefit future iterations of the dilepton (and $\ell + E_{\text{T}}^{\text{miss}}$) analysis.

An overlap removal algorithm is used to ensure the same detector signature is not reused for the reconstruction of two separate objects entering the analysis. The “lepton-favoured” working point is used to disable rejection of lepton candidates overlapping with jets. Of course, such instances are still subject to e.g. isolation requirements downstream which means that in practice the overlap removal has little additional impact. Amongst leptons, if two electrons share track hits the candidate with lower p_{T} is rejected; a muon which overlaps with an electron track is rejected only if it is a CT muon (section 5.2), otherwise the electron is rejected.

6.3.3 Differences in event selection between the 36.1 fb⁻¹ and 139 fb⁻¹ analyses

As mentioned in chapter 1, the full Run-2 analysis [1] being discussed was preceded in 2017 by a publication of analysis results using the 36.1 fb⁻¹ dataset [9]. Overall there is a lot of continuity between the dilepton event selection criteria used to analyse the 36.1 fb⁻¹ and 139 fb⁻¹ datasets. This was intentional to help focus on important changes to other aspects of the dilepton analysis strategy. Some of the main differences were introduced by updates to central ATLAS performance algorithms. These changes were introduced through updates to calibration tools at the analysis level, but also were implemented directly in data because the entire 2015-2016 dataset was reprocessed in Athena release 21 since the 36.1 fb⁻¹ analysis.

- Update to electron reconstruction algorithm (supercluster-based) and “good muon” threshold (see chapter 5)
- In 2018 a new set of four isolation WPs were recommended by ATLAS in place of the old (see section 5), including “Loose” and “LooseTrackOnly” used for the electron and muon channels, respectively. The new electron WP (“Medium”) was chosen because its efficiency was very similar to the previous WP at masses above 200 GeV (being at most 3% less efficient below that, relatively). Similarly, a comparison of the new muon isolation WPs indicated that the “FCTightTrackOnly” efficiency was similar and very slightly better than the previous WP.
- The Z -peak region was previously included (i.e. SR had $m_{\ell\ell} \geq 80$ GeV) in a single very tall bin at the beginning of the analysis templates to help constrain nuisance

parameters in the profile likelihood fit. This region is no longer included because the background estimate is fitted to the spectrum above 225 GeV.

6.4 Corrections applied to MC

Simulated samples do not do a perfect job modelling physics processes and detector response. This is quantified by studies of data-MC agreement in control regions and accounted for in recommended calibrations/corrections by ATLAS performance groups. MC corrections are typically provided in the form of p_T -dependent scale factors (SFs). While scale factors are often close to 1, they are also commonly packaged together with up/down variations to parameterize the effect of systematic uncertainties on MC kinematic distributions. The corrections are briefly described as follows:

***K*-factors:** Event weights applied to DY MC at truth level to account for higher-order effects (see section 6.2.1).

Pileup reweighting: The pileup conditions in MC generated prior to data-taking are simulated based on an estimate of the distribution of the average number of events μ per bunch crossing. Since pileup has an important impact on detector performance, it is desirable that the number of pileup pp vertices μ in MC is similar to that in data. Weights are therefore applied to MC events according to their values of μ in order to morph the overall μ distribution toward that of the data-taking period in comparison.

Lepton scale factors: Individual sets of scale factors are applied as event weights to account for calibrated discrepancies between data and MC reconstruction, trigger, identification, and isolation efficiencies. This applies to both electrons and muons, but for muons there is an additional scale factor recommended to correct the efficiency in simulation of track-to-vertex-association (TTVA).

Electron kinematic calibration: The calorimeter response to electrons from $Z \rightarrow e^+e^-$ events in data is used to calibrate the common ATLAS electron energy scale and resolution models. Both MC and data are calibrated using the `ElectronPhotonFourMomentumCorrection` tool configured to version es2017_R21_v1.

Muon sagitta bias: Correction of muon momentum using η - ϕ sagitta bias calibration maps. The prescription is applied to MC but not to data, as recommended, to avoid affecting individual high-mass events by local inhomogeneities in the maps. The `MuonCalibrationAndSmearingTool` package configured to Recs2018_05_20 recommendations was used to implement this as well as the momentum corrections below.

Muon momentum correction: The momentum of muons in simulation is calibrated according to in-situ studies and also smeared to better reflect the impact on resolution from detector imperfections (e.g. MS misalignment).

6.5 Selection results

The total yields of selected dilepton events in the Run-2 dataset are shown in table 6.6.

Data-taking period	Data events selected per channel	
	Dielectron	Dimuon
2015-2016	235302	201641
2017	280635	239825
2018	367657	313404
Total	883594	754870

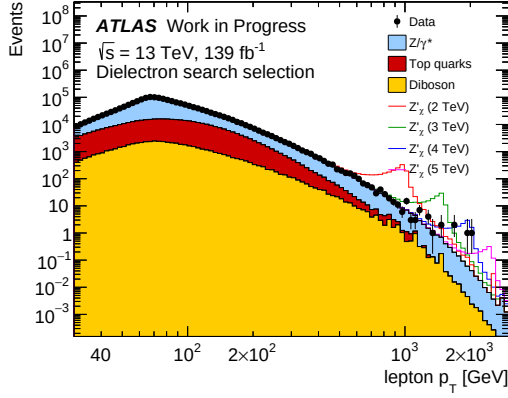
Table 6.6: Number of data events passing selection for each channel in the three data-taking periods, for the SR starting at $m_{\ell\ell} = 130$ GeV.

A comparison of data and MC for the p_T , η , and ϕ distributions of the selected leptons is shown in figures 6.6, 6.7, and 6.8, respectively. In each case, only qualitative agreement is expected because (1) the fake electron background estimate is not included and (2) the background templates do not include any systematic variations.

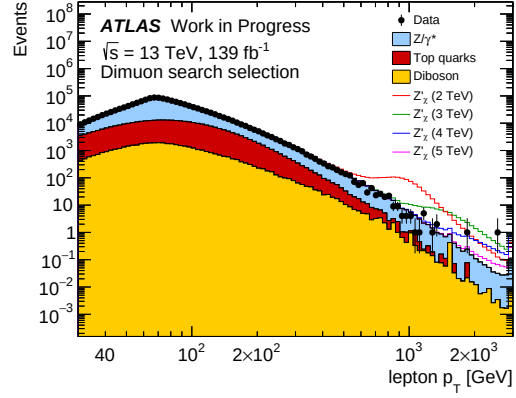
In figure 6.6 the p_T spectra of the leading and subleading leptons are shown for each channel. The broad peak at around 70 GeV is dominated by the effect of the $m_{\ell\ell} \geq 130$ GeV selection cut applied to the Drell-Yan p_T spectrum. The subdominant top quark backgrounds have a wider peak about 10 GeV above that of Drell-Yan. The lineshapes of Z'_χ samples peak at values slightly less than half their pole mass $m_{Z'}$ because their integrated cross section is larger below than above $m_{Z'}$ (see section 6.7). For both figures 6.6a and 6.6b, the single entries in the two highest bins are from the same event which has the highest $m_{\ell\ell}$ in each channel, to be discussed below. Note that bins which “jump” in the diboson MC p_T distribution tail are explained by single events which are assigned a very large generator weight.

The lepton η and ϕ distributions illustrate both the analysis angular criteria as well as the relative efficiency for the reconstruction of electrons and muons in different regions of ATLAS. In figure 6.7a the visible dips in two bins on each side is a result of vetoing the electron crack region which overlaps with these two bins. The muon η distribution (figure 6.7a) exhibits three large dips, one in the middle caused by the gap in MS coverage for $|\eta| < 0.1$, and at the two barrel-endcap transition regions, particularly $1.01 < |\eta| < 1.10$ which is vetoed. The MS performance for high- p_T muons is also more inhomogeneous in ϕ than the ECAL for electrons, as shown in figure 6.8. The two dips at roughly $\phi = 1$ and 2 are caused by additional material present and reduced MDT coverage in the region of the ATLAS feet (including muon chambers vetoed by the high- p_T selection working point).

The invariant mass distribution for events passing the analysis selection are shown in figure 6.9. Starting from 130 GeV, there are 100 log-spaced bins including the final overflow bin at 6 TeV. Note that the statistical analysis is performed with finer binning. Histograms

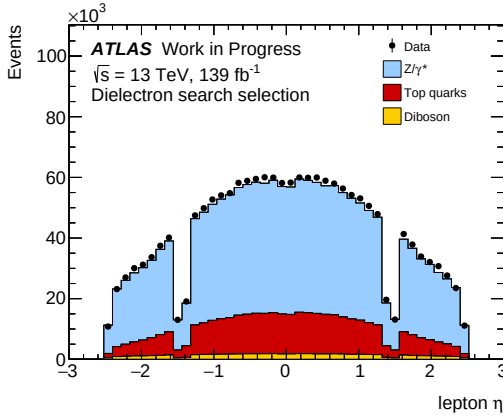


(a) Dielectron channel

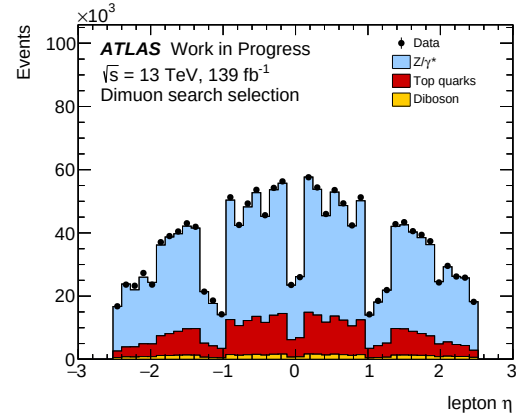


(b) Dimuon channel

Figure 6.6: Comparison of the lepton p_T distributions in data and MC for the full Run-2 dataset. p_T distributions of Z'_χ signal samples with multi-TeV pole masses are also shown as examples. Note that no systematic variations are applied to the MC distributions and the fake electron background is not included.



(a) Dielectron channel



(b) Dimuon channel

Figure 6.7: Comparison of the lepton η distributions in data and MC for the full Run-2 dataset. Note that no systematic variations are applied to the MC distributions and the fake electron background is not included. The features in each channel reflect both analysis η vetoes and performance differences across the layouts of the calorimeter and muon spectrometer.

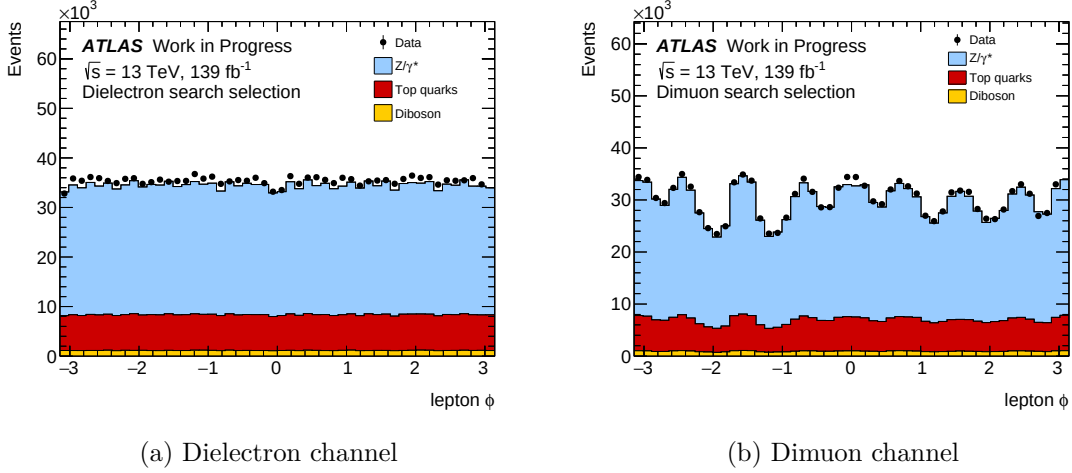


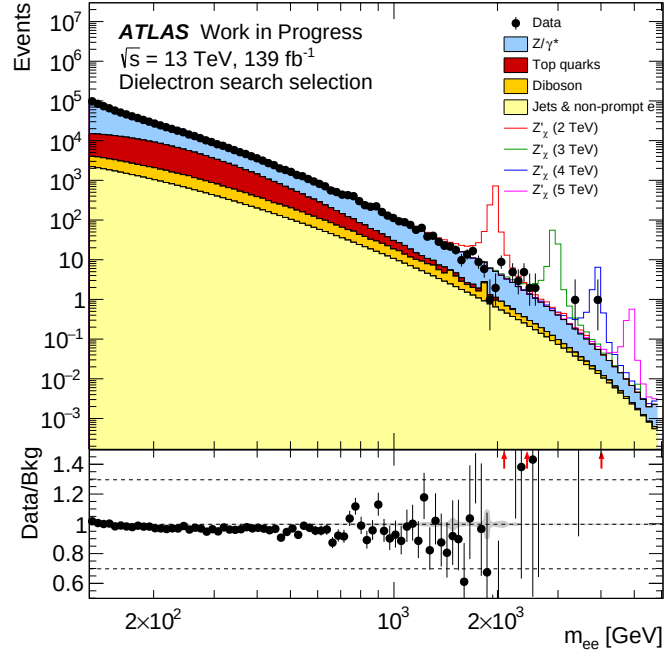
Figure 6.8: Comparison of the candidate lepton ϕ distributions in data and MC for the full Run-2 dataset. Note that no systematic variations are applied to the MC distributions and the fake electron background is not included. The features in the dimuon channel reflect inhomogeneities in the muon spectrometer coverage.

of the dominant SM background processes are stacked in order of dominance. In the electron channel there is the additional contribution from misidentified electrons arising from multi-jet and W +jets events. Visually, the data agree reasonably well with the total MC prediction, even though systematic uncertainties in the MC have not been varied (i.e. these are “pre-fit” distributions). It is important to point out the smoothness of both DY and diboson MC distributions out to 6 TeV thanks to the mass-sliced event generation scheme. Both channels include overlays of example Z'_χ resonances with pole masses of 2, 3, 4, and 5 TeV. The difference in invariant mass resolution between the channels is evidenced by the dramatically wider resonance shapes in the dimuon channel.

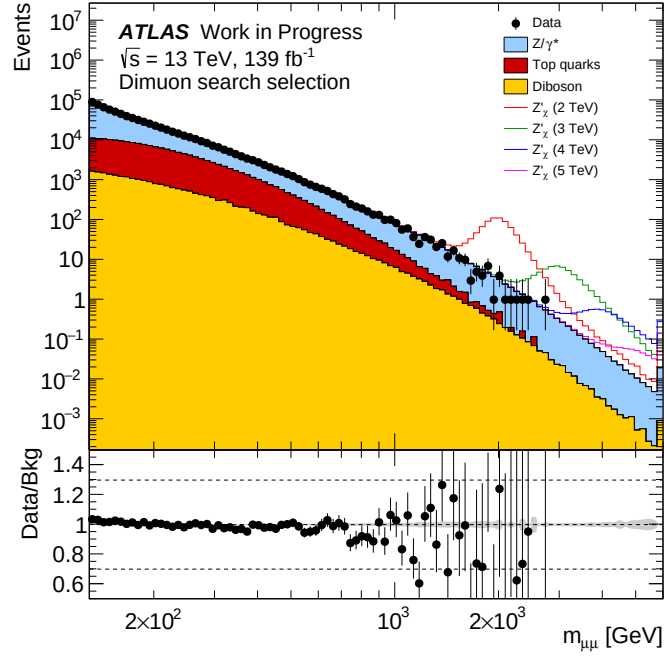
For the electron channel in figure 6.9, the contribution from misidentified electrons is estimated for the full Run-2 dataset using the matrix method described in chapter 7. This is an update with respect to the published version of this plot ([1], aux. fig. 1) which used an appropriately scaled-up version of the fake electron background template derived in the 36.1 fb^{-1} dilepton analysis [9]. That is, the estimate was not redone in that case, simply reused. This is justified only because the data-MC comparison is used in this case only for the purpose of validating the general modelling, whereas the data-driven statistical analysis does not rely directly on this comparison. Rather, potential mismodelling due to reusing the fake electron template would only impact statistical results through the estimate of the spurious signal systematic, as discussed in section 8.2.4. It was therefore checked that removing this contribution entirely does not majorly impact the statistical analysis since it is the smallest background component. For this reason, data-MC comparisons in other variables for the full Run-2 dataset do not include an estimate of the fake electron background. The data-MC

agreement when using the updated fake electron background as shown in the subpanel of figure 6.9a is slightly shifted but remains reasonable.

The final event in each channel’s invariant mass distributions has been highlighted using ATLANTIS [127] event displays, shown in figures 6.10 and 6.11. The dielectron event has $m_{ee} = 4.06$ TeV and the dimuon event has $m_{\mu\mu} = 2.75$ TeV – the two highest dilepton invariant masses reported in an ATLAS publication. Figure 6.12 shows a display of the dimuon event with the highest invariant mass in the 36.1 fb^{-1} dataset (this one from 2016). The visualization is done in PERSINT [128] and makes it easy to see the three-station requirement on the high- p_T muons. The presence of a third muon possibly indicates this is a diboson process. The transverse momenta reported in the caption of figure 6.12 reflect the updated calibration used in the full Run-2 analysis and are slightly different than the values reported in [9] (the invariant mass shifts down by 4.5%).



(a) Dielectron channel



(b) Dimuon channel

Figure 6.9: The invariant mass distribution for the full Run-2 data and MC samples passing full analysis selection. Lineshapes of Z'_χ signal samples with multi-TeV pole masses are also shown as examples. The subpanel shows the ratio of data to MC on top of a shaded MC statistical uncertainty band. Note that no systematic variations have been applied to the MC distributions (i.e. they are at “pre-fit” values), so some tension is to be expected.

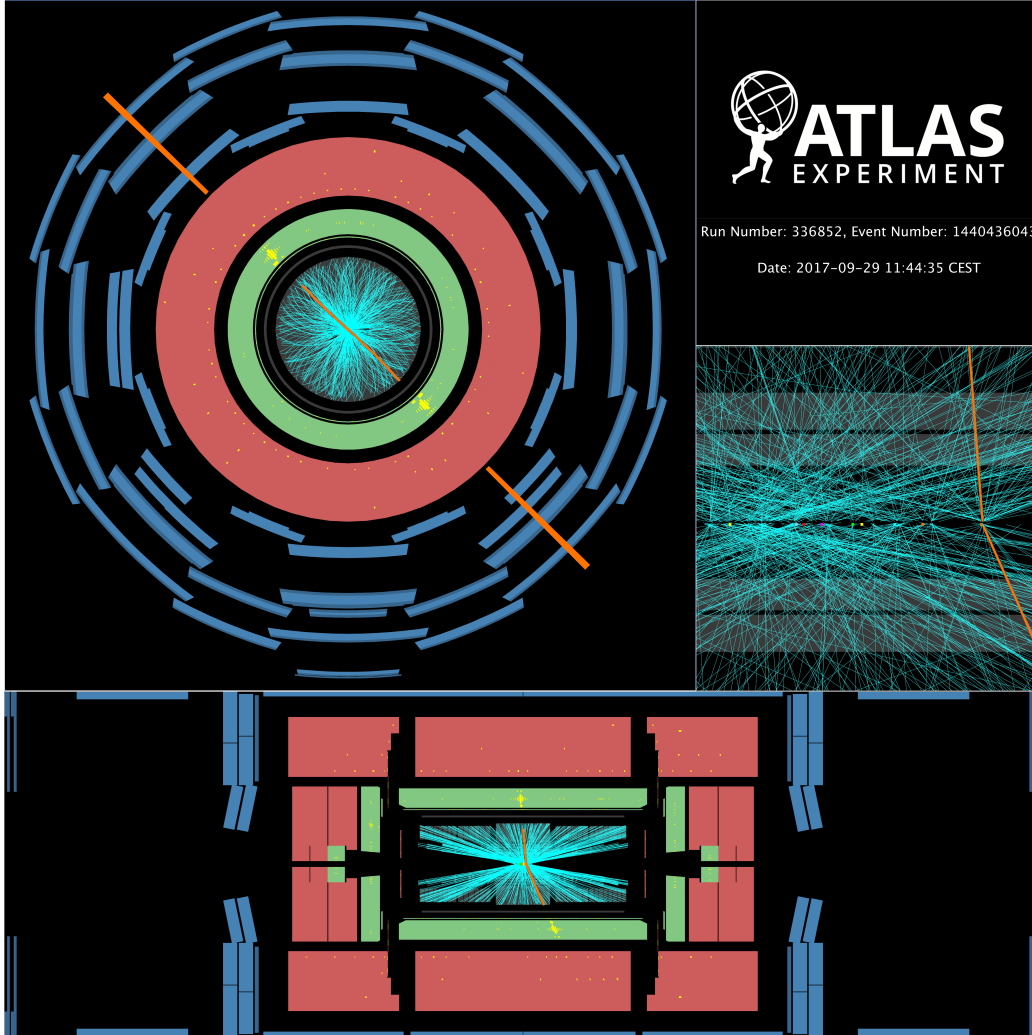


Figure 6.10: Event display of the highest invariant mass candidate event in full Run-2 ATLAS dataset passing the dielectron selection. The two electrons are indicated by the orange tracks, yellow clusters of energy in the ECAL, and orange extensions of the track serving as a visual aid. The dielectron invariant mass is $m_{ee} = 4.06$ TeV. The leading electron, pointing toward the lower half of the display, has $E_T = 2.01$ TeV, $\eta = 0.47$ and $\phi = -0.78$ (measured from the horizontal to the right). The other, subleading electron has $E_T = 1.92$ TeV, $\eta = -0.03$ and $\phi = 2.37$. Since the two electrons match each other in transverse momentum within less than 5%, they are nearly back-to-back in the transverse plane (i.e. their ϕ 's differ by nearly π). Figure taken from [1].

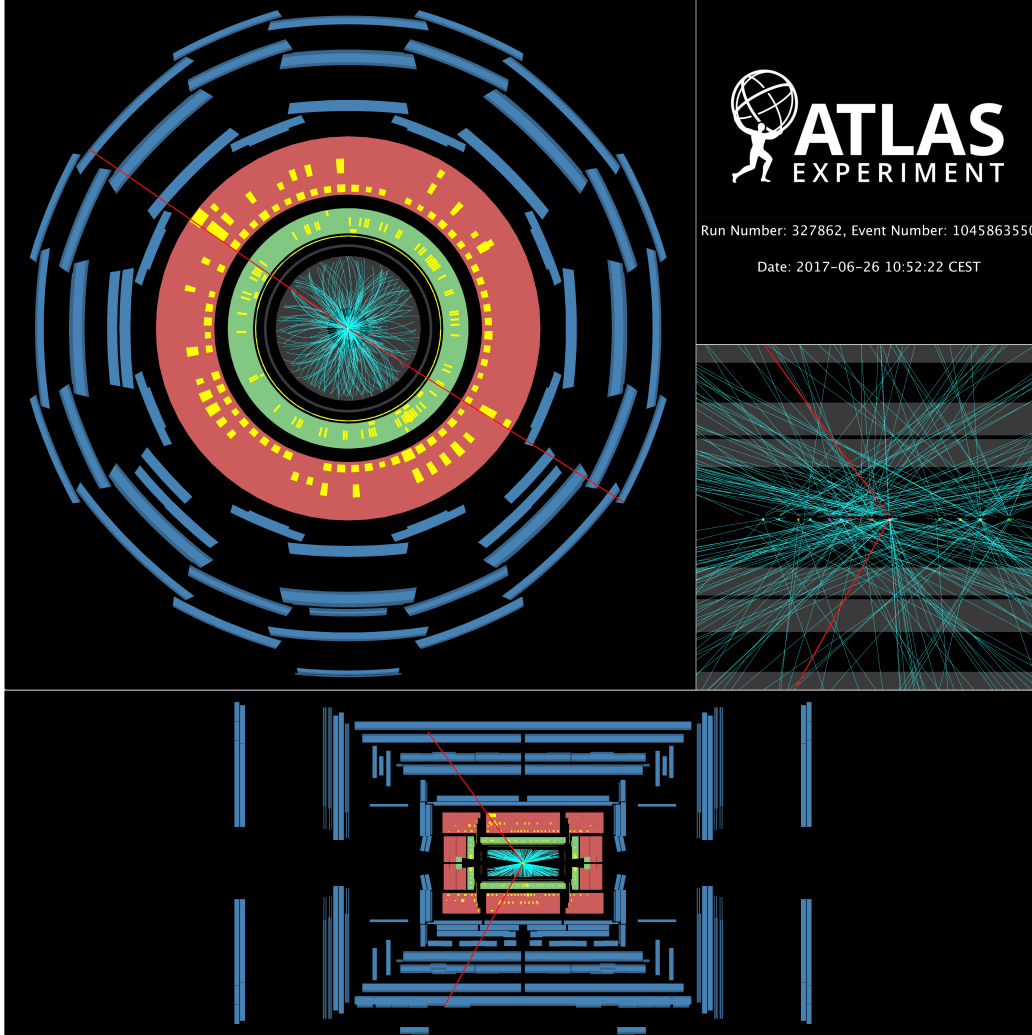


Figure 6.11: Event display of the highest invariant mass candidate event in the full Run-2 ATLAS dataset passing the dimuon selection. The two muon tracks are indicated by red lines. The dimuon invariant mass is $m_{\mu\mu} = 2.75$ TeV. The leading muon, pointing toward the lower half of the display, has $p_T = 1.82$ TeV, $\eta = -0.52$ and $\phi = -0.56$ (measured from the horizontal to the right). The other, subleading muon has $p_T = 1.04$ TeV, $\eta = -0.67$ and $\phi = 2.53$. Nearby each muon are energy deposits in the calorimeters with $p_T < 50$ GeV which is consistent with the energy loss of high- p_T muons (see section 5.2.3). The event has $E_T^{\text{miss}} \simeq 800$ GeV in the direction of the subleading muon. Figure taken from [1].

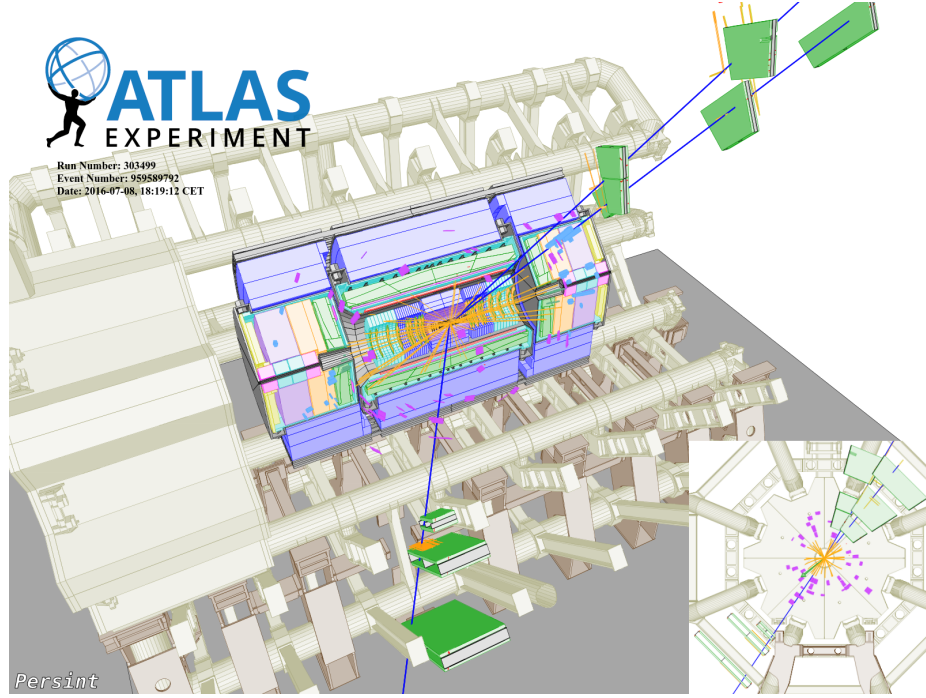


Figure 6.12: Event display of the highest invariant mass candidate event in the 2015-2016 (36.1 fb^{-1}) dataset passing the dimuon selection. Three muon tracks are indicated by blue lines. The three MDT chambers associated with each muon, as required in the high- p_T muon selection, are shown in green. The leading muon, pointing in the lower left quadrant of the display, has $p_T = 564 \text{ GeV}$, $\eta = -0.43$ and $\phi = -2.16$. The subleading muon points nearest to the top right corner and has $p_T = 555 \text{ GeV}$, $\eta = 1.81$ and $\phi = 0.90$. These two are associated with the dimuon pair of combined invariant mass $m_{\mu\mu} = 1.90 \text{ TeV}$. The event contains a third muon of $p_T = 136 \text{ GeV}$ and has $E_T^{\text{miss}} \simeq 179 \text{ GeV}$ in the direction of $\phi = -2.38$, indicated by a green arrow. Figure taken from [9].

6.6 Dilepton invariant mass response model

Since the invariant mass observable is the basis for the statistical analysis of the dilepton final state, it is valuable to model what the ATLAS detector and the selection criteria do to it. That is, given an event producing two leptons with some *truth-level* invariant mass, how does the process of reconstruction affect it within the analysis selection? This section discusses the so-called *transfer-function* (TF) method used to model the detector response, and how this in turn serves in both background and signal modelling for the full Run-2 analysis.

6.6.1 The transfer function method

The invariant mass *response* is defined as the relative difference between the reconstructed ($m_{\ell\ell}$) and truth-level (m_t) invariant mass:

$$\mathcal{R} = \frac{m_{\ell\ell} - m_t}{m_t} = \frac{m_{\ell\ell}}{m_t} - 1 \quad (6.5)$$

where m_t refers to the truth-level mass and $m_{\ell\ell}$ is the ordinary, reconstructed mass. Given a sample of simulated events, the inclusive response distribution can be filled using the mass before and after reconstruction. For a perfect detector the distribution would be a Dirac delta function. However, for a *real* detector, it is a peaking distribution centred at zero (hopefully) with width corresponding to the relative invariant mass resolution. More useful than the inclusive one is the response distribution generated by events within a differential slice of truth-level phase space. Of particular relevance to this analysis is a model of the response as a function of truth mass: $P(\mathcal{R}|m_t)$. Knowing this would make it possible to transform a given invariant mass distribution from truth-level (f_t) to reco-level (f_r) using the following convolution:

$$f_r(m_{\ell\ell}) = f_t(m_t) \otimes [\text{TF}(m_{\ell\ell}|m_t) \cdot A\epsilon(m_t)] \quad (6.6)$$

where the transfer function $\text{TF}(m_{\ell\ell}|m_t)$ is simply the response distribution $P(\mathcal{R}|m_t)$ shifted by +1 and then scaled by m_t . The convolution kernel in (6.6) transforms the truth-level distribution through the following effects:

horizontal: smearing of the invariant mass by the response (packaged in the TF)

vertical: loss of events due to analysis acceptance cuts (A) and reconstruction efficiency (ϵ)

Since the response distribution is a function of the truth-level phase space of its underlying event population, it is therefore physics process-dependent as well. In the dilepton analysis, the most interesting response distribution to model is that of the Drell-Yan process, not

only because it is the dominant background, but because it shares an identical topology with a spin-1 signal diagram.

The response distribution as a function of m_t is obtained by first slicing the NLO Drell Yan MC into 200 log-spaced bins of truth mass. Then, in each bin of roughly constant m_t , the resulting response distribution is fitted using the following model:

$$P(\mathcal{R}|m_t) = \kappa \cdot \text{CB}(\mathcal{R}|\mu_{\text{CB}}, \sigma_{\text{CB}}, \alpha, n) + (1 - \kappa) \cdot \text{Gauss}(\mathcal{R}|\mu_{\text{G}}, \sigma_{\text{G}}) \quad (6.7)$$

where CB denotes a “crystal ball” distribution². This component is needed in addition to a simple Gaussian to model the response tails populated by poorly-measured leptons at high- p_{T} – mainly due to bremsstrahlung of electrons and for muons, chamber alignment or calorimeter energy loss. Example fits of the TF model are shown in figure 6.13.

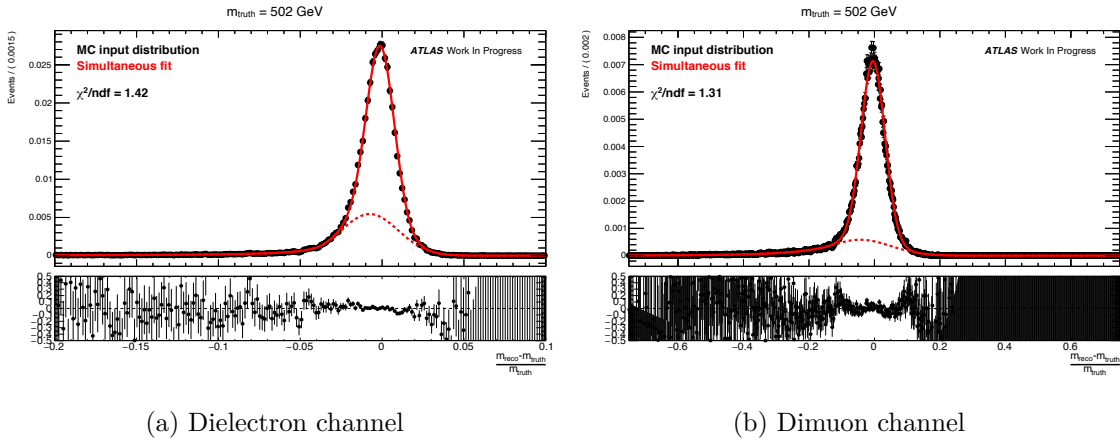


Figure 6.13: Results of the simultaneous transfer function fit shown in red superimposed on the $m_{\ell\ell}$ response distribution for the truth mass bin centred at 502 GeV. The dashed red line shows the Crystal Ball component (equation 6.7). Note that the x -axis range is different for each channel so that their relative width is not immediately evident. Figures adapted from [129].

Though not written explicitly, each parameter of the two distributions in (6.7) are treated as functions of m_t , including the coefficient κ controlling their relative contribution. Therefore, once a functional ansatz has been chosen for each of the 7 parameters by trial and error, the fit in the 200 m_t bins is performed simultaneously to extract the “parameters

²A piecewise distribution comprised of a Gaussian core with one of its tails modified beyond α to fall like power- n :

$$\text{CB}(x; \mu, \sigma, \alpha, n) \propto \begin{cases} e^{-(x-\mu)^2/2\sigma^2} & , \quad \left| \frac{x-\mu}{\sigma} \right| \leq \alpha \\ e^{-\alpha^2/2} \cdot \left[1 - \frac{\alpha^2}{n} + \left(\frac{\alpha}{n} \right) \frac{x-\mu}{\sigma} \right]^{-n} & , \quad \left| \frac{x-\mu}{\sigma} \right| > \alpha \end{cases} \quad (6.8)$$

of the parameters”. Appendix A shows the fitted parameterizations of the response model parameters (figures A.3 and A.4) as well as their functional forms (tables A.3 and A.4).

Finally, the mean and width parameters are shifted and scaled to transform the response model into a TF model $\text{TF}(m_{\ell\ell}|m_t)$.

6.6.2 Invariant mass resolution and selection efficiency

The width of the response distribution is approximated by the following weighted sum³:

$$\frac{\sigma(m_t)}{m_t} = \sqrt{\kappa\sigma_{\text{CB}}^2 + (1 - \kappa)\sigma_{\text{G}}^2} \quad (6.9)$$

The simultaneous fit therefore also provides the relative invariant mass resolution as a function of m_t , which is plotted in figure 6.14. What can immediately be appreciated from this figure is the relative performance impact of electron and muon reconstruction at high- p_{T} . While both channels begin at a comparable relative $m_{\ell\ell}$ resolution of 2 – 3% around 200 GeV, this climbs almost linearly for the dimuon channel. The causes of this degradation are the challenges associated with small muon track sagitta measurements, as discussed in section 5.2.3. On the other hand, the dielectron channel (thanks to calorimeter-based measurements) shows slight improvement of relative resolution towards high- $m_{\ell\ell}$, settling at $\sigma(m_t)/m_t \simeq 1\%$.

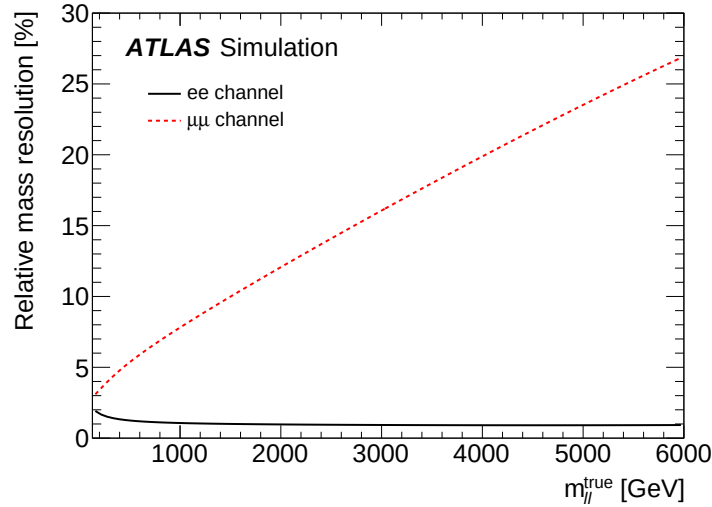


Figure 6.14: Relative invariant mass resolution for the dimuon and dielectron channels as a function of the true invariant mass. The channels’ very different behaviour is driven by the challenges of muon measurement at high- p_{T} .

³This expression neglects an additional $\kappa(1 - \kappa)\delta^2$ term under the radical, where $\delta = \mu_{\text{CB}} - \mu_{\text{G}}$, which has a small impact compared to the σ_i^2 contributions.

Similarly, the 200 slices of NLO Drell Yan MC are used to derive the acceptance and selection efficiency (and their product $A\epsilon$) expected for a spin-1 mediator process:

$$A\epsilon(m_t) = \left(\frac{N_{\text{true}}^i \text{ pass acceptance}}{N_{\text{true}}^i \text{ in sample}} \right) \times \left(\frac{N_{\text{reco}}^i \text{ reconstructed \& selected}}{N_{\text{true}}^i \text{ pass acceptance}} \right) \Big|_{m_t \in \text{bin } i} \quad (6.10)$$

where “reco” labels reconstructed events and “true” events prior to reconstruction (i.e. truth-level). The “fiducial” cuts applied at truth-level and their impact on the acceptance are described in section 6.7. Figure 6.15 shows the resulting values of $\epsilon(m_t)$ for both channels obtained from the Drell Yan slices. More precisely, what is shown by the curves are smooth parameterizations from fitting the 200 ϵ points with 8-parameter polynomials in m_t with powers between -5 and $+2$. The lower efficiency for the dimuon channel is a result of the 3-station and “good muon” requirements of the high- p_T selection working point discussed in section 5.2.1. On the other hand, aside from a mainly isolation-related dip at the low- m_{ee} end, the dielectron channel exhibits slowly increasing efficiency towards a value of about 74%.

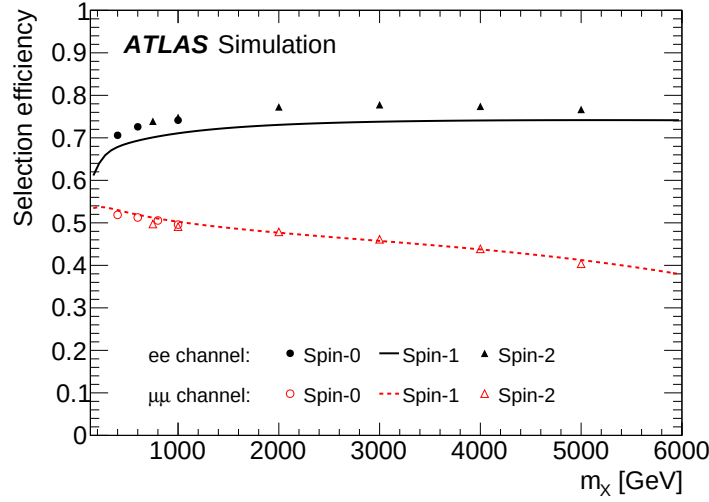


Figure 6.15: Analysis selection efficiency for dielectron (black) and dimuon (red) channels. It represents the fraction of events lying within the acceptance criteria which additionally are reconstructed and satisfy the full signal region selection of the dilepton analysis. Drell-Yan samples are used to measure spin-1 as well as Graviton and MSSM Higgs samples for spin-2 and 0 points, respectively.

Also shown alongside the spin-1 ϵ curves are test masses of spin-2 and spin-0 mediator processes calculated using dedicated samples of high-mass scalar (MSSM⁴ Higgs) and graviton mediators, respectively. Examples of $m_{\ell\ell}$ distributions for these samples are displayed

⁴Minimal Supersymmetric Standard Model

in figures A.5 and A.6 in appendix A. While different spin cases will generally experience different losses at the acceptance level (due to their different angular distributions), these test points demonstrate that selection efficiency is comparable between the different spins. However, the dielectron channel in particular shows that there is some discrepancy; in this case the spin $\neq$ 1 samples exhibit larger efficiency by up to 4% (absolute). This can be explained as follows: spin-0 and spin-2 samples have a larger proportion of their η distribution in the central region where electron identification is more efficient on average. The truth-level lepton η distributions are shown for the three spin cases in figure A.7 in appendix A. For muons, the reconstruction efficiency is more uniform between central and forward regions, so the different spin cases are more similar. This comparison implies that the generic cross section limits for spin-1 mediators can be applied to scalar and graviton resonances to within a few percent (absolute) in the factor ϵ (provided the model-dependent acceptance impact is also incorporated).

6.6.3 Application of the transfer function to background smoothing

As will be discussed in section 8.2.1, the ability to validate a data-driven background estimate is limited by the smoothness of the MC background template available. Generating $\mathcal{O}(100\text{M})$ additional simulated DY events would be the highly CPU-expensive solution. By far the most time is spent fully simulating the ATLAS detector, and in particular the calorimeter response. While there are fast parameterizations [130] available to help, the computational overhead of this route is still prohibitive. Alternatively, considering that for this analysis the background only needs to be modelled in the $m_{\ell\ell}$ observable, the transfer function presents its own application-specific parameterization of the ATLAS detector response. It captures the essential smearing of truth-level invariant mass associated with detector measurements. Thereby, the TF enables very large generator-level events to be converted into reconstruction-level distributions which statistically resemble those of the fully-simulated NLO DY samples. In contrast to fully-simulated samples, generator-level samples with a very large number of events can be produced at far lower cost using a local institute batch computing system.

It is true that statistical fluctuations in the NLO DY also propagate as uncertainties on the parameters of the TF and $A\epsilon$ fits. The impact this has on the smoothed $m_{\ell\ell}$ distribution could in principle be generated by varying the parameters according to the covariance matrix of the simultaneous fit, but this is difficult because of the complexity of the fit and its parameter correlations. Instead, uncertainty bands on the TF and $A\epsilon$ shapes are obtained by fitting an ensemble of 10k pseudo-datasets drawn from the NLO DY samples. The events in each pseudo-dataset are scaled by random weights drawn from a Poisson distribution with mean of 1. The resulting uncertainty bands are negligible in the core region of the TF and only become visible in the tails. For the $A\epsilon$, the impact of MC statistical uncertainty is

of the order $\lesssim 0.1\%$ for masses below 5 TeV. In both cases the statistical uncertainty is therefore considered negligible.

The models of invariant mass response and acceptance times efficiency are combined using equation 6.6 in convolution with the large generator-level sample. The resulting distribution is then reweighted to match the target dataset integrated luminosity. Its statistical variations are small enough to enable its use as a validation template for the data-driven background shape modelling. A modified transfer function is also derived to smooth the $t\bar{t}$ MC $m_{\ell\ell}$ distribution for the same purpose. In this case, the full set of $t\bar{t}$ dilepton events which are accepted at truth-level are used to help recover the full statistical precision available in the sample (a portion of which is lost during reconstruction).

6.6.4 Application of the transfer function to signal modelling

A model of how distributions in the dilepton invariant mass propagate through detector reconstruction can also be applied to the signature of a resonance. In particular, a resonance with negligible intrinsic width (think Dirac delta function) will be smeared into a distribution which purely represents the detector response. In fact, the transformation of such a “zero-width” resonance of a hypothetical particle X of given pole mass m_X is equivalent to the transfer function at $m_t = m_X$.

This connection between the TF method and signal lineshape modelling was the basis for validating the extracted TF model against dedicated zero-width signal samples. The samples of 50k Drell-Yan events were generated in mass windows narrower than 1 GeV and their fully-reconstructed distributions compared to the parameterized distributions predicted by the pre-fitted TF model. Good agreement was achieved across the 11 tested pole masses (i.e. window centres).

The same feature of the TF model is exploited to generate signal lineshapes with nonzero intrinsic resonance widths. The statistical analysis of the $m_{\ell\ell}$ spectrum is done using a generic signal model with the goal of making results applicable to a range of specific signal scenarios. The generic signal is implemented as a semi-analytic function which is the convolution of underlying “physics” lineshape with detector response (TF):

$$f_{\text{sig}}(m_{\ell\ell}|m_X, \Gamma_X) = \text{BW}_X(m_{\ell\ell}) \otimes \text{TF}(m_{\ell\ell}|m_t = m_X) \quad (6.11)$$

where BW is a non-relativistic Breit-Wigner of mass m_X and width Γ_X , and where the form of TF is given in equation 6.7. It is semi-analytic because the convolution is evaluated numerically and cached using the `RooFFTConvPdf` class (using as backend the fast fourier transform library FFTW3 [131]).

Examples of generic signal lineshapes and their underlying components in each channel are shown in figure 6.16. Both subfigures show an identical truth-level Breit-Wigner distribution along with two very different response distributions, as expected. Clearly in

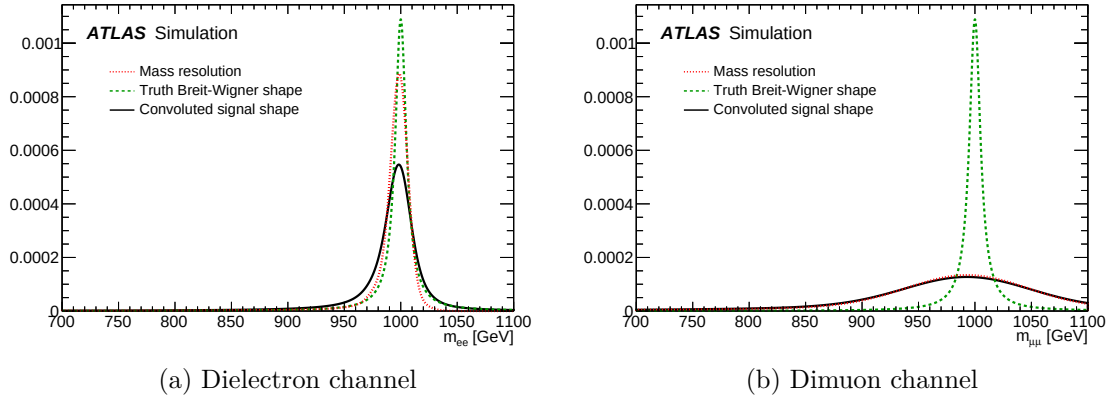


Figure 6.16: The lineshapes of the generic signal model for each channel at $m_X = 1$ TeV and $\Gamma_X/m_X = 1.2\%$ as an example. The two underlying components are shown as dashed lines while the result of their convolution is the solid black line.

the dimuon channel the latter is what dominates the resulting signal lineshape, and this behaviour persists until larger relative widths.

6.7 Fiducial cuts for generic limits

Signal inclusive cross sections tend to fall steeply for larger and larger pole masses into the multiple TeV range. When generating the $m_{\ell\ell}$ distributions expected from proton collision production, two competing effects can come into tension. Although the differential cross section is strongly peaking at the pole mass, the probability to find a quark and antiquark with sufficient momentum fraction x is suppressed according to the PDF. For example, assume that a u quark near the peak of its predicted momentum fraction ($x_1 \sim 0.12$ according to figure 2.5b) took part in s -channel production of an on-shell Z' boson with a mass of 2 TeV. Ignoring ISR, the other parton in this case would have to be a $\bar{u}$ with the following momentum fraction:

$$x_1 x_2 s = m_{\ell\ell}^2 \implies x_2 = (2 \text{ TeV})^2 / [(0.12) \cdot (13 \text{ TeV})^2] = 0.197 \quad (6.12)$$

that is, the antiquark needs to possess nearly 20% of the other proton's momentum which, looking at figure 2.5b, is far out in the tail of the $xf_{\bar{u}}$ distribution. As a result, multi-TeV resonances exhibit sizable *parton luminosity tails* coming from the relative rise of the off-shell production component of the differential cross section due to the PDF. Although the broad profile of parton luminosity tails typically has a small impact on the overall search sensitivity, the tails can integrate a large amount of the differential cross section. Hence, when setting limits on high-mass signal cross sections it is necessary to account for the fraction of the total cross section which lies outside the core of the resonance.

Parton luminosity tails are dependent on the resonance width and production mode (e.g. ggF vs. $q\bar{q}$). Another model-dependent feature is the impact of interference effects on the $m_{\ell\ell}$ spectrum of a signal. Section 2.4.1 mentioned how interference appears between processes with identical external lines. For this reason, s -channel production of BSM mediators in general interfere with the dominant Drell-Yan background in the dilepton channel. The size of the cross-term depends on the degree of kinematic overlap of the resonances (i.e. pole mass separation, width of new resonance) and the new mediator’s coupling to quarks and leptons. Taking both the effects of parton luminosity tail and the SM interference into account would in principle require computing a dedicated set of cross section limits for each model of interest.

To make the results based on the generic signal model applicable to a range of models, the following set of truth-level fiducial cuts is defined with the aim of removing the above model-dependent effects:

$$\text{each lepton: } \begin{cases} p_T^{\text{tru}} > 30 \text{ GeV} \\ |\eta^{\text{tru}}| < 2.5 \end{cases} \quad (6.13)$$

$$\text{lepton pair: } m_{\ell\ell}^{\text{tru}} > m_X - 2\Gamma_X \quad (6.14)$$

While the p_T and η cuts enclose the accepted phase space of the signal region (table 6.4), the one-sided cut (6.14) defines a fiducial “volume” approximately free of model-dependent effects appearing well below m_X . Since it removes the tail of the *truth-level* signal distribution below $2\times$ the intrinsic width, the cut has a negligible impact on the effective shape of the peak region. What it *does* do is scale the cross section of models in proportion to the relative size of their off-shell components. In this way the limits on the generic model can be reinterpreted as applying to a given signal model (of matching m_X and Γ_X) once brought inside the fiducial volume defined by the prescribed cuts.

It is important to note that this procedure assumes that the signal lineshape which gets cut off does not contribute significantly to the signal sensitivity, but only to the overall cross section entering the limits. Models with strong interference effects appearing near the peak region can push this assumption. Section 8.4.2 will make note of the visible impact when cross section upper limits are presented within the defined fiducial volume.

The fraction of events which at truth level satisfy the per-lepton fiducial cuts (6.13) is shown in figure 6.17. Only the electron channel is shown because the muon channel acceptance, in this definition, is identical up to slight statistical variation. This is again a polynomial fit to the acceptance obtained from each of the 200 log-spaced bins of Drell-Yan. The acceptance initially climbs steeply as the Lorentz boost of the dilepton system drops off towards higher masses. Also the p_T cut has an impact mainly at low masses. The acceptance plateaus as the dilepton angular distribution decouples from the mass of the

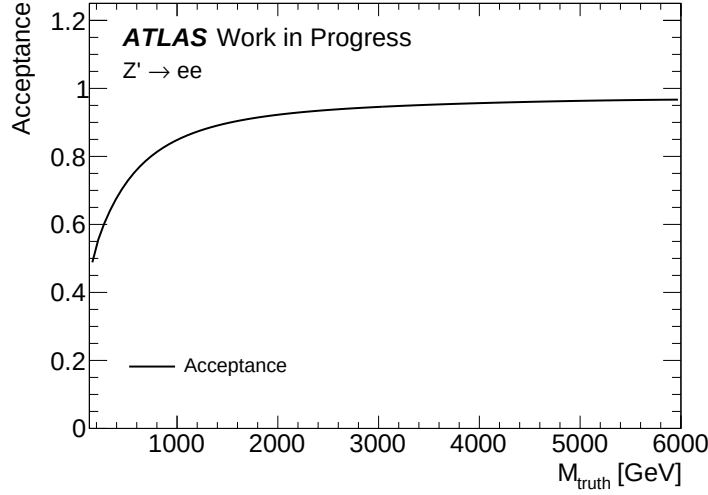


Figure 6.17: Acceptance defined by the fiducial cuts (equations 6.13 and 6.14). It represents the fraction of events present in the original MC samples that lie within the acceptance cuts at truth-level.

(virtual) mediator. For a resonance shape, as opposed to a single narrow bin, the acceptance factor due to the one-sided mass cut (6.14) also applies.

6.8 Systematic uncertainties

Additional sources of uncertainty affect the dilepton analysis which are not related to the finite size of the selected data sample. These systematic uncertainties can be broadly categorized as coming from three sources. First are uncertainties in the auxiliary measurements of theory parameters which affect MC differential cross sections. One of the sources of uncertainty that becomes important at high- p_T is the limited information about the PDF at high- x . Second, many uncertainties are related to the efficiency and accuracy of reconstruction, i.e. the fact that we deal with an imperfect detector. A final source of systematic uncertainty accounts for the dependence that results may have on specific analysis choices or techniques. For example, the spurious signal systematic discussed in section 8.2.4 is meant to account for imperfect modelling of the chosen functional form for the background estimate (along with some other uncertainties as well). Refer to section 8.3 for a summary of how these uncertainties are incorporated into the statistical analysis results.

The full set of systematics considered in one way or another in this analysis is listed in table 6.7. The variations which affect the detector response distribution and hence the signal shape are indicated. In practice these systematics are provided for each MC sample in the form of central variations with different values of lepton kinematics on the event level. The final column indicates whether a systematic was neglected during the final statistical analysis. The only systematic which affects the signal shape that turned out to be negligible

type (channel)	systematic	shape?	negligible?
Thy. (both)	ALPHAS PDF_EV{1...7} PDF CHOICE_HERAPDF20 CHOICE_NNPDF30 REDCHOICE_NNPDF30 SCALE_Z SCALE_EW SCALE_PI TTbarNorm		
Exp. (both)	LUMI PRW_DATASF		✓
Exp. (ee)	EG_SCALE_ALL EG_RESOLUTION_ALL EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR	✓ ✓	✓ ✓
Exp. ($\mu\mu$)	MUON_SCALE MUON_ID MUON_MS MUON_EFF_ISO_STAT MUON_EFF_ISO_SYS MUON_EFF_RECO_STAT MUON_EFF_RECO_SYS MUON_EFF_BADMUON_STAT MUON_EFF_BADMUON_SYS MUON_EFF_TTVA_STAT MUON_EFF_TTVA_SYS MUON_EFF_TrigStatUncertainty MUON_EFF_TrigSystUncertainty MUON_SAGITTA_RESBIAS MUON_SAGITTA_RHO	✓ ✓ ✓	✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

Table 6.7: The set of centrally-provided systematic variations considered per channel. Each are parameterized by corresponding “up” and “down” variations (except for PDF CHOICE). Systematics which have an impact on the signal $m_{\ell\ell}$ shape, through variation of peak or width, are indicated. Systematics which are found to be negligible are also indicated.

is the muon momentum scale, which shifts the signal resonance but significantly less than the $m_{\mu\mu}$ resolution. All other neglected systematics are in the category of variations affecting $A\epsilon$, and do so at levels below 1% throughout the $m_{\ell\ell}$ spectrum. In general systematics affecting $A\epsilon$ are handled centrally through “up” and “down” variations of efficiency scale factors which multiply MC total event weights.

Here follows a brief description of the different uncertainties in table 6.7.

The theoretical uncertainties come from a variety of sources which are reflected in their names. For instance, **ALPHAS** models uncertainty on the value of the strong coupling constant α_s and **TTbarNorm** uncertainty on the $t\bar{t}$ cross section which affects the normalization of this sample. Other quantities which enter cross section calculations include the Z boson mass and the electroweak scale v which likewise have associated uncertainties. The PDF uncertainties comprise both variations coming from the global fits to experimental data and also differences related to the choice of PDF set, owing to differences in their predictions. These two sources are discussed more in section 9.2.1. Theoretical sources of systematic uncertainty affect the analysis results only in an indirect way, as discussed in sections 8.2.4 and 8.3. Ultimately they are not treated as independent, but are overlaid to determine an uncertainty envelope.

Section 8.3 also discusses the uncertainty (**LUMI**) on the estimate of integrated luminosity. The pileup profile measured in data likewise comes with an uncertainty that affects the reweighting of MC via **PRW_DATASF** (which was found to be negligible for this analysis).

Coming to the third group of rows in table 6.7, first listed are the uncertainties in the calibration of ECAL energy scale and resolution which affect the electron E_T measurement. Next are the set of uncertainties on the scale factors used to correct MC efficiencies for passing identification, isolation, reconstruction, and trigger criteria. In these cases, the **_1NPCOR_PLUS_UNCOR** suffix indicates that the variations are provided in a simplified nuisance parameter scheme which does not including the full correlation model.

Finally, in the muon channel, three categories of experimental uncertainties can be identified, affecting the muon p_T measurement, the MC efficiency scale factors, and coming from the sagitta bias correction. The **ID** and **MS** systematic variations parameterize the uncertainty of the position (and hence sagitta) resolution of these two subdetectors. Uncertainties on scale factors are split into statistical and systematic components affecting the efficiency of MC in isolation, reconstruction, track-to-vertex association (**TTVA**), and trigger criteria. Additionally, an uncertainty which is rather large at high- p_T is associated with the efficiency of the “good muon” requirement (which historically was called the “bad muon” veto). The sagitta bias correction described in section 5.2.3 likewise introduces uncertainty through the residual bias remaining after the correction in $\eta - \phi$ maps (**RESBIAS**) and the effect of p_T resolution on a weight parameter ρ used to derive the correction (**RHO**).

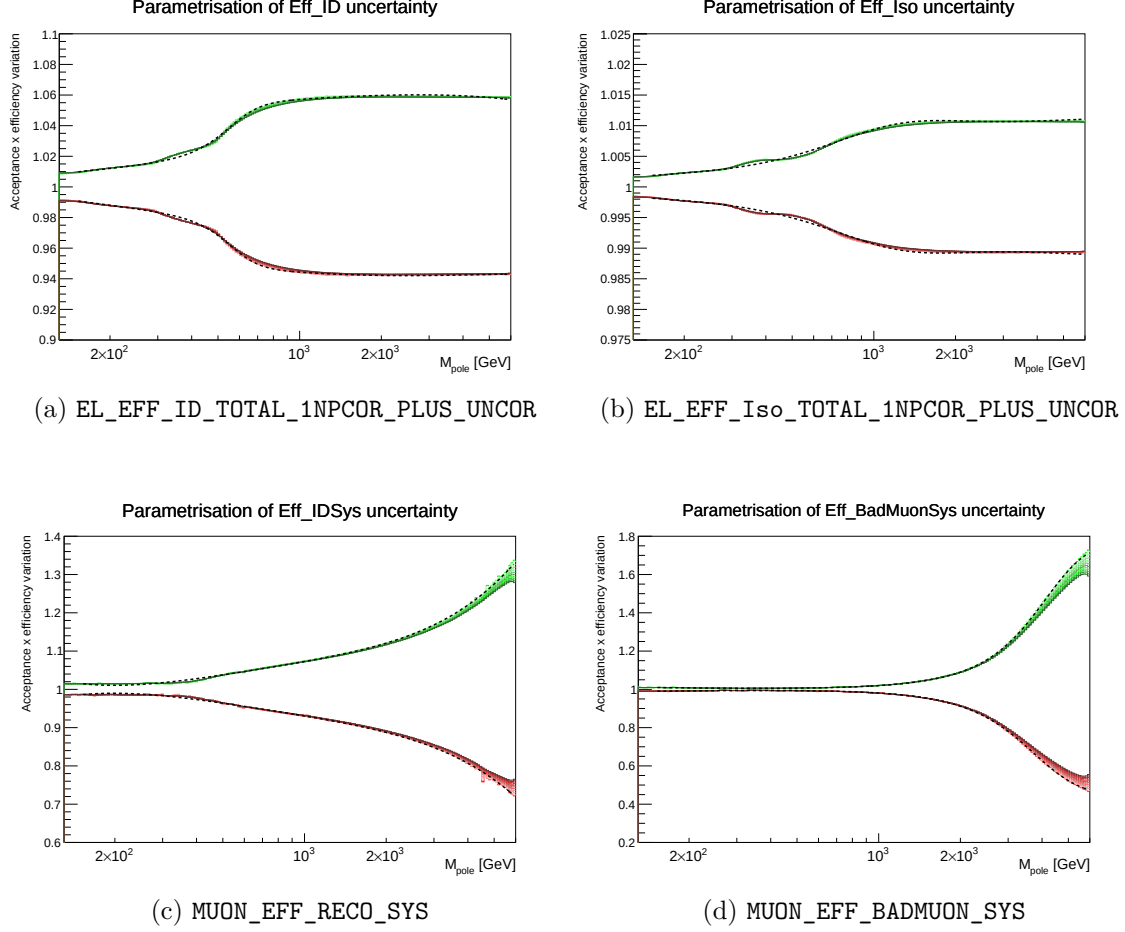


Figure 6.18: Main $m_{\ell\ell}$ -dependent systematic variations of $A\epsilon$ in both channels. Curves in shades of green are the “up” variations for signal widths between 0% and 10%, and curves in shades of red are the “down” variations. The functional parameterizations are shown as dashed black lines. Figures taken from [129].

6.8.1 Signal normalization uncertainties

Systematic uncertainties propagate into the signal normalization/yield through variations of the acceptance times efficiency⁵. For each systematic the “up” and “down”-varied NLO DY templates are used to determine corresponding $A\epsilon$ variations with the procedure of section 6.6.2. The “up” and “down” relative variations – i.e. ratio of varied to nominal – are symmetrized, then averaged, and then fitted using a parametric model. The (mirrored) parameterizations are shown overlaid with the signed variations in figure 6.18 for the main $m_{\ell\ell}$ -dependent systematics in each channel. Also, the impact is shown in different shades of green and red for 20 different signal widths. The similarity of the curves at different widths is the basis for using a common variation for all of them based on the parameterization for

⁵Also through the integrated luminosity; discussed in section 8.3.

the signal with 0% width. The MUON_EFF_RECO_STAT and MUON_EFF_ISO_SYS variations are treated but are not shown in figure 6.18 because they are effectively constant in $m_{\ell\ell}$.

6.8.2 Signal shape uncertainties

The systematics highlighted in table 6.7 as affecting the signal shape do so by either shifting the resonance peak or widening/narrowing it. The former effect is captured in the TF model by a variation in the two distribution means μ_{CB} and μ_G . Similarly, the width variation is captured through changes with respect to the nominal σ_{CB} and σ_G . In this case, a different TF simultaneous fit is extracted for the “up” and “down” variations. The varied response distributions at an example mass bin and the corresponding TF fits are overlaid in figure 6.19 for the four shape systematics. Note that the only case of a systematic shift of signal mass is figure 6.19a while the others are width variations.

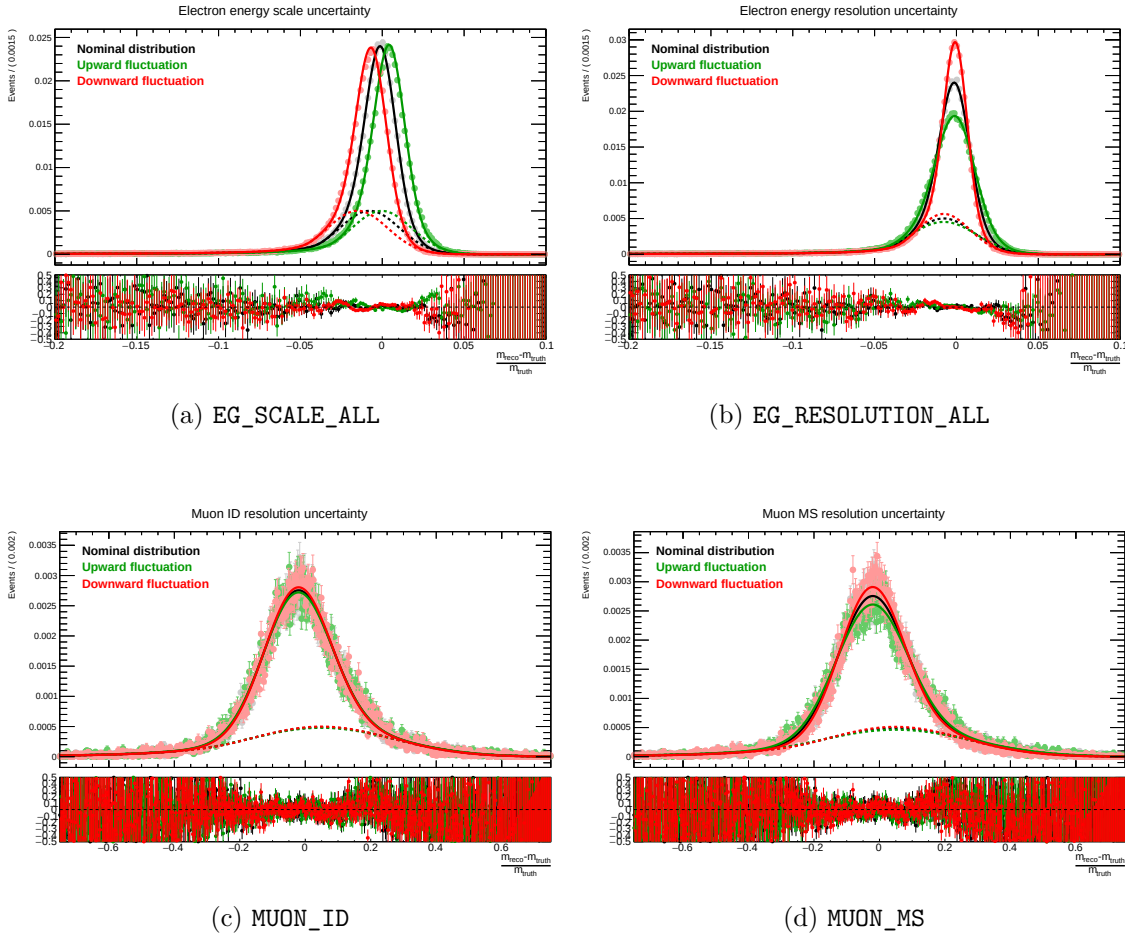


Figure 6.19: “up” and “down” systematic variations affecting an example 0%-width signal shape at an example mass. Figures from [129].

Chapter 7

Estimating the fake electron background

As discussed in section 5.1 the signature of an electron inside the ATLAS detector consists of hits in the inner detector matched to an energy deposit in the electromagnetic calorimeter. Unlike muons which leave a unique signature (when accurately reconstructed), electrons have a signature which by nature can be mimicked by hadronic showers. The likelihood to mistakenly identify a jet as an electron is small, but the effect is amplified by the fact that collisions in the LHC produce $\sim 10^5$ jets for every one prompt electron. The presence of such *fake* electrons in many ATLAS analyses is therefore unavoidable, although this background can be suppressed. Similarly, electrons which are produced in secondary decays downstream from the hard interaction are also an undesirable background.

Quality requirements built into the electron reconstruction algorithm offer baseline rejection of fake electrons prior to the main analysis criteria. Then, at the analysis level, requirements of identification, isolation, and impact parameter cuts add further rejection of fake electrons. The portion of fake electrons which survive all such measures and enter the analysis signal region is, however, non-negligible in the dilepton analysis and is thus modelled as a reducible background contribution.

This chapter describes the procedure used to estimate the “Jets & non-prompt e” electron background entering the full Run-2 analysis shown in figure 6.9a. Note that the corresponding figure in the full Run-2 publication [1] uses the estimate derived for the 36.1 fb^{-1} analysis, renormalized to 139 fb^{-1} (see discussion in section 6.5). While the procedure presented here uses a similar methodology as the earlier estimate, it was derived following the publication using a newly developed framework and additionally estimates the fake electron background properly for the 2017 and 2018 datasets.

7.1 Origins of fake electrons

Fake electrons are objects passing electron reconstruction and potentially quality cuts which are either misidentified jets or actual electrons which do not originate directly from the hard scatter process. The latter are better described by the term *non-prompt* electrons, or *background* electrons. The primary sources of fake electrons are as follows:

Light-flavour jets: comprised of bound states of u , d , and s quarks and also their antiquarks, for example, $\pi^\pm$ and $K^\pm$

Photon conversion: non-prompt electrons pair-produced by a photon interacting with detector material (particularly the ID)

Heavy-flavour decays: non-prompt (or “soft”) electrons produced in leptonic decays of c and b hadrons

The first source results from the partial overlap between the signature of an electron and jet – both produce tracks and calorimeter showers. As discussed in section 4.5, hadronic showers feature lots of pions, and the proportions of charged and neutral can fluctuate to resemble an electromagnetic shower. There are aspects of the ATLAS detector layout that help to differentiate electron from jet signatures. The depths and composition of the calorimeters were chosen so that the bulk of electron and photon showers would be contained within the electromagnetic calorimeter. Jets are therefore characterized by a shower which typically extends beyond the electromagnetic calorimeter, but this does not represent a hard cut in the electron identification (section 5.1.1).

Jets of light flavour hadrons can also produce non-prompt leptons through intermediate decays to photons and subsequent photon conversion before the calorimeter: $\pi^0 \rightarrow \gamma\gamma$; $\gamma \rightarrow e^+e^-$. This segues into the second category above. Prompt photons can also convert into non-prompt electron-positron pairs when encountering the Coulomb fields in detector material. These cases can mimic an electron signature only if some other track is present because otherwise the Pixel hit requirements for electrons will not be satisfied (table 5.1). One case which can produce this unfortunate arrangement is when a high- p_T muon emits a bremsstrahlung photon which leads to an associated calorimeter shower. Thankfully this source is practically eliminated by standard overlap removal which will not permit the same energy deposit to be used in the electron reconstruction that is associated with a muon. In case the photon comes from bremsstrahlung or transition radiation of an electron it can lead to charge misidentification of the electron (section 5.1.3).

The third category is for non-prompt electrons that originate in decays of heavy flavour hadrons. For example, $B^+ \rightarrow e^+\nu_e$. Typically these electrons are not well isolated in $\eta - \phi$ from the constituents of the jet containing the heavy flavour decay products.

7.2 The matrix method

While the other background processes in the dilepton final state are modelled using MC samples, this is not the case for the fake electrons. Samples of simulated dijet and W +jets events could be passed through the selection, but only a tiny fraction of the generated events would enter the analysis signal region. Besides lacking statistical precision, such a MC-based estimate would amplify any potential mismodelling of discriminating features such as shower shape (see section 5.1.3). It is therefore a common strategy to deduce fake backgrounds directly from kinematic distributions in data by manipulating efficiencies (i.e. rates) of electron quality cuts, typically isolation and identification.

The *matrix method* is a robust technique for estimating the proportion of a given data distribution that is expected to comprise fake electrons. Central to the formalism of the matrix method is the relationship between two bases used to describe the same set of data:

$$\begin{array}{cc}
 \text{Actual basis} & \text{Observable basis} \\
 \left\{ \begin{array}{l} R : \text{“Real”} \\ F : \text{“Fake”} \end{array} \right. & \left\{ \begin{array}{l} T : \text{“Tight”} = \text{passBaseline} \ \& \ \text{passID} \ \& \ \text{passIso} \\ L : \text{“Loose”} = \text{passBaseline} \ \& \ \neg(\text{passID} \ \& \ \text{passIso}) \end{array} \right. \quad (7.1)
 \end{array}$$

In data passed through some selection, the numbers of electrons that are *actually* real and fake, N_R and N_F , are not directly accessible (except in simulated data). Instead, they must be inferred by determining the transformation going from the actual basis to the observable basis, and then computing the inverse. In the observable basis, electron candidates are described using two states, *tight* (T) and *loose* (L). The label T is reserved for electron candidates which pass the full selection S . On the other hand, candidates which pass all selection criteria except for the subset of criteria $I \subset S$ targeting fake electrons are labelled L . As defined in equation 7.1, the analysis criteria used for this study take I to be the electron isolation and identification criteria, while the other criteria are lumped into a “baseline” selection.

Presumably there is a high probability r for real electrons to become tight electrons and a low probability f for fake electrons to become tight electrons. That is the goal of the quality criteria I . Indeed, as will be determined in sections 7.3 and 7.4 below, these probabilities or *rates* are in the following ranges:

$$\text{“real rate” } r \equiv P(T|R) \gtrsim 90\% \quad (7.2)$$

$$\text{“fake rate” } f \equiv P(T|F) \sim 16 - 45\% \quad (7.3)$$

The real and fake rates are the key to connecting yields in the observable basis to those in the actual basis. For example, in a sample selected with baseline criteria the number of

electron candidates that are tight N_T will contain a portion f of the total number of fake electrons N_F . The full relationship is

$$\begin{pmatrix} N_T \\ N_L \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_R \\ N_F \end{pmatrix} \quad (7.4)$$

Knowledge of r and f and measurement of the observable yields allows to solve for N_R and N_F . In particular, the matrix inversion can be performed in bins of electron p_T and η to derive a spectrum of fake electrons. This procedure is the basis for extracting the invariant mass spectrum of fake electrons passing the dilepton analysis selection. However, in this case yields are measured for events with both a leading (1) and subleading (2) electron candidate, so the baseline selection includes four categories:

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1(1-r_2) & r_1(1-f_2) & f_1(1-r_2) & f_1(1-f_2) \\ (1-r_1)r_2 & (1-r_1)f_2 & (1-f_1)r_2 & (1-f_1)f_2 \\ (1-r_1)(1-r_2) & (1-r_1)(1-f_2) & (1-f_1)(1-r_2) & (1-f_1)(1-f_2) \end{pmatrix} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} \quad (7.5)$$

where the rate factors are labelled as being evaluated at the (p_T, η) of the leading or subleading electron. As in the single electron case, the rate matrix $\mathbf{M}$ can be inverted to determine the vector of actual yields:

$$\vec{N}_{\text{obs}}(p_{T1,2}, \eta_{1,2}) = \mathbf{M}(r_{1,2}, f_{1,2}) \vec{N}_{\text{act}}(p_{T1,2}, \eta_{1,2}) \quad (7.6)$$

$$\implies \mathbf{M}^{-1}(r_{1,2}, f_{1,2}) \vec{N}_{\text{obs}}(p_{T1,2}, \eta_{1,2}) = \vec{N}_{\text{act}}(p_{T1,2}, \eta_{1,2}) \quad (7.7)$$

Once $\vec{N}_{\text{act}}$ is determined, its lower three entries are of particular interest for back-substituting into (7.5) and evaluating the following sum:

$$N_{TT} \Big|_{\geq 1F} = r_1 f_2 N_{RF} + f_1 r_2 N_{FR} + f_1 f_2 N_{FF} \quad (7.8)$$

which is the matrix method prediction for the amount of events in the signal region containing one or more misidentified (fake) electrons. Contributions from W + jets events will mainly show up in the first two terms while dijet events will be concentrated in the third term.

7.3 Measuring real electron rate

The real rate is easy to measure since well-modelled real electrons are readily available via MC samples. The POWHEG+PYTHIA Drell Yan mass-sliced samples are used for their good coverage out to high electron E_T . Events are passed through the full selection code and each of the two electrons put into tight and loose histograms depending on whether they

pass both isolation and identification (in addition to the baseline selection). The real rate is then calculated as follows:

$$r(\eta, p_T) = \frac{N_T}{N_T + N_L} \bigg|_{DY \rightarrow ee} \quad (7.9)$$

The ratio is calculated in bins of p_T and η for each of the three MC campaigns. Figure 7.1 shows the result of merging all three. The general plateau shape is driven by the isolation and identification signal efficiencies reported in [102]. The real rate is above 90% for all but the first p_T bin where both isolation and identification efficiencies are lowest.

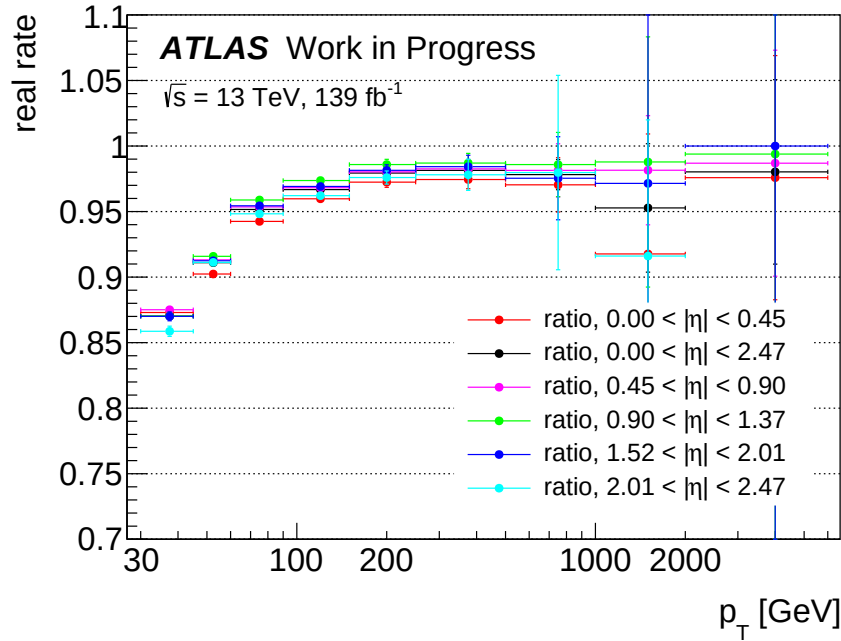


Figure 7.1: The real rate measured with electrons from Drell-Yan MC. Different curves show five different η bins and the black one the η -inclusive case.

7.4 Measuring fake electron rate

Compared to the real rate, determining the fake rate is much more involved. To start, where does one get fake electrons? It was already mentioned that MC simulation is not reliable in this application. In data, on the other hand, fake electrons are reconstructed using the actual ATLAS detector and also are produced in abundance. The majority are rejected by standard criteria upstream from the analysis selection, starting with the trigger object quality requirements. Therefore by deliberately loosening criteria it is possible to recover a sample that includes a sufficiently large number of fake electrons. Processes that produce

real electrons will also contribute to such a sample but can be suppressed with vetoes and also simulated in MC and subtracted off. This strategy is discussed in more detail below.

7.4.1 Fake-enriched data sample

The data sample enriched in fake electrons is obtained in two main steps:

Step 1. Inclusive selection for the fake sample to minimize upstream selection bias by using:

- samples in the EXOT19 derivation format
- prescaled single-electron triggers with minimal quality requirements

Step 2. Purify the fake sample by removing contamination from real electrons

- **NoDY**: veto events with 2 or more lhMedium electrons
- **NoZ**: veto events with invariant mass within 20 GeV of Z peak
- **NoW**: veto events with missing ET > 25 GeV
- **Subtract MC** estimates of processes producing real electrons that enter above selection

The EXOT0 derivation format used for the dilepton analysis is not efficient at collecting fake electrons because of its upstream requirements (see section 6.3.1). On the other hand, EXOT19 is dedicated to fake electron studies and only requires at least 1 electron with $p_T > 20$ GeV. The main quality bottleneck therefore comes in the criteria of the triggers used. To widen this as much as possible, triggers are used which require at least one electron with only LHVLoose (i.e. “very loose”) identification. Also `etcut`-type triggers are used at higher p_T . Different menus of triggers are available in the different data taking periods, and the ones used are listed in table 7.1.

The tradeoff for triggering on electrons with minimal quality requirements is that these appear at very high rates (many of them are actually jets), particularly at low- p_T . Hence it is only possible to store a fraction of such events given the limited bandwidth of the ATLAS high-level trigger. Most of the triggers in table 7.1 are therefore *prescaled*, meaning that only every n^{th} event on average which fires the trigger is accepted (where n often varies within a data-taking period). Triggers with higher p_T threshold typically have smaller prescale factors and are therefore able to keep more events in the p_T spectrum above their turn on compared to using a lower- p_T prescaled trigger in the same region. For this reason, depending on the data-taking period the available prescaled triggers are combined in successive p_T bins listed in table 7.1. For each electron candidate the smallest prescale factor is stored and later applied as a weight so that the p_T spectra from the different triggers are continuous when combined.

period	trigger	prescaled?	p_T bin
2015	e26_lhvloose_nod0_L1EM20VH	Y	[30, 65)
2015	e60_lhvloose_nod0	Y	[65, 315)
2015	e200_etcut	N	[315, ∞)
2016	e26_lhvloose_nod0_L1EM20VH	Y	[30, 65)
2016	e60_lhvloose_nod0	Y	[65, 126)
2016	e120_lhvloose_nod0	Y	[126, 147)
2016	e140_lhvloose_nod0	Y	[147, 315)
2016	e300_etcut	N	[315, ∞)
2017	e28_lhvloose_nod0_L1EM20VH	Y	[30, 65)
2017	e60_lhvloose_nod0	Y	[65, 126)
2017	e120_lhvloose_nod0	Y	[126, 147)
2017	e140_lhvloose_nod0	Y	[147, 168)
2017	e160_lhvloose_nod0	Y	[168, 315)
2017	e300_etcut	N	[315, ∞)
2018	e28_lhvloose_nod0_L1EM22VH	Y	[30, 65)
2018	e60_lhvloose_nod0	Y	[65, 75)
2018	e70_lhvloose_nod0	Y	[75, 85)
2018	e80_lhvloose_nod0	Y	[85, 105)
2018	e100_lhvloose_nod0	Y	[105, 126)
2018	e120_lhvloose_nod0	Y	[126, 147)
2018	e140_lhvloose_nod0	Y	[147, 168)
2018	e160_lhvloose_nod0	Y	[168, 315)
2018	e300_etcut	N	[315, ∞)

Table 7.1: Triggers used for creating fake-enriched samples in the different data-taking periods.

A modified version of the main dilepton analysis code is used to select the fake-enriched sample, so the object-level criteria for electrons in table 6.4 are applied (except isolation and identification are not checked for the baseline selection). The event-level criteria of table 6.4 are not applied, except the GRL and event cleaning. Within a given event, each electron candidate is checked against the fired trigger to see if it matches within $\Delta R < 0.4$.

Once a sample of data is collected using the highly inclusive selection described above, it is purified to remove real electrons. Three vetos are imposed to suppress electrons coming from Drell-Yan, and Z and W decays. As will be seen in the next section, these are the dominant real processes entering the selection. Since events are not required to have more than one electron, $W \rightarrow e\nu_e$ is the main source of real contamination. Events which do have additional electrons beyond the trigger-matched one are required to not have any combination with $|m_{\ell\ell} - 91.1876| < 20$ GeV. Given the steeply falling spectrum around the Z resonance, the p_T range immediately outside this will remain the dominant part.

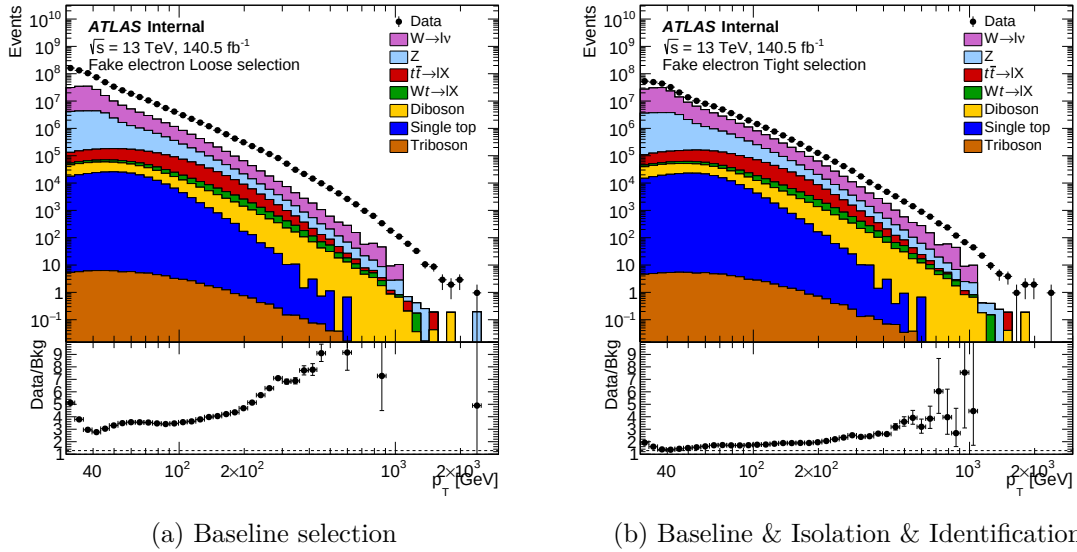


Figure 7.2: Comparison of data and MC electron p_T spectra in the fake-enriched selection for the full Run-2 dataset. Left shows the baseline selection while right shows the baseline in addition to full isolation and identification requirements. The MC samples estimate the real contamination to be subtracted off, therefore agreement is not expected in the data-MC ratio in the subpanel plots.

7.4.2 Real contamination subtraction

Since the purifying vetos are not 100% effective, it is important to take into account the residual contamination from real electrons in the fake-enriched sample. This can be estimated by passing MC samples through the selection and vetos. Inclusive samples of W and Z generated with POWHEG+PYTHIA are used, not the mass-binned Drell-Yan. Otherwise, the same background samples for the dilepton analysis are used for the processes $t\bar{t}$ and single- and W -associated top production. Samples of diboson and triboson processes modelled in SHERPA 2.1 for several lepton multiplicities are also included in the real electron subtraction, although as will be shown this component is small enough to be nearly negligible.

The real subtraction samples are processed for each of the three MC campaigns corresponding to the full Run-2 data-taking period. An important difference compared to processing the data is that the corresponding triggers in simulation are not prescaled – obviously it would be counterproductive to generate events only to discard them. So no prescale weight is applied to MC events.

Figure 7.4.3 shows the “dilution” plots comparing data and MC in the fake-enriched selection. The samples correspond to the full Run-2 integrated luminosity. It is clear that in the selection including vetos, $W \rightarrow e\nu$ dominates the real electron background with $Z \rightarrow ee$ coming in second. The MC estimates of the main processes show that the baseline

selection contains between 12.7% (at 400 GeV) and 36% (at 40 GeV) contamination by real electrons (7.2a subpanel). Applying isolation and identification criteria on top of this selection significantly reduces the data yield while the MC yields are not much affected. As a result, the subpanel of (7.2b) indicates as large as 74% real contamination near 40 GeV which then decreases to 38% at 400 GeV.

7.4.3 Fake rates after subtraction

The sum of MC yields in figure 7.4.3 is subtracted bin-by-bin from the data. While they are shown integrated over η , the subtraction is done in bins of η (as well as p_T bins, but different than the ones shown here). To account for potential differences in the real contamination, the subtraction is also done separately for each of the 2015-2016, 2017, and 2018 periods. The resulting two sets of “post-subtraction” histograms are used to compute the fake rate as follows:

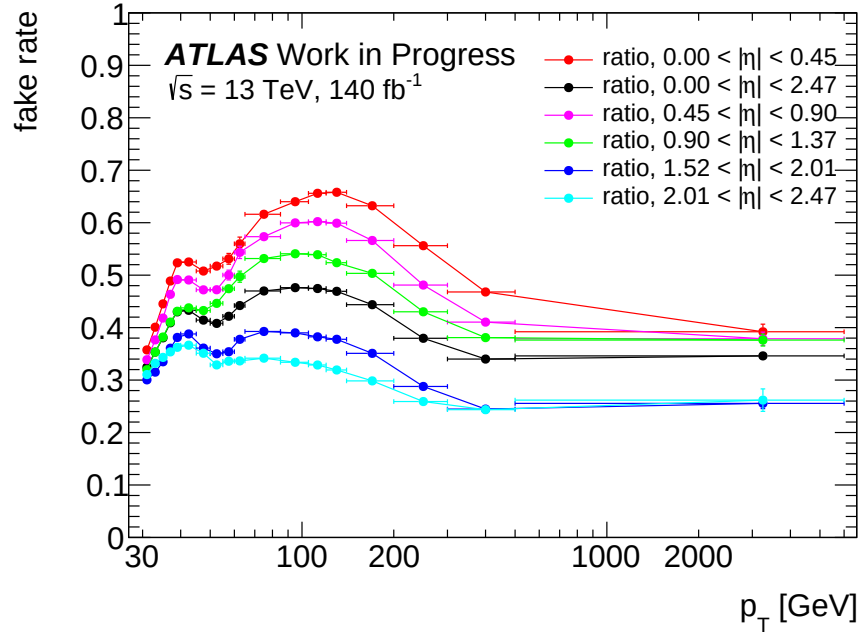
$$f(\eta, p_T) = \frac{N_T}{N_T + N_L} \Big|_{\text{fake-enriched data}} \quad (7.10)$$

Figure 7.3 shows the fake rate inclusive of the three data-taking periods; the fake rate prior to MC subtraction is also shown as a reference. The binning is chosen to line up with the prescaled trigger regions defined in table 7.1 with an additional final bin that integrates all electrons with $E_T > 500$ GeV. There is a trend for lower fake rates for more forward electrons which is probably correlated with the broad decrease in electron identification efficiency towards higher η (see [102] figure 8). This does not necessarily mean that most fake electrons will enter the dilepton selection from more central regions – the yield itself comes from a combination of the fake rate and the production cross section of jets and non-prompt electrons, which also presumably has some η dependence. Prior to subtraction, the fake rate ranges within 25% – 65% while afterwards it is 16% – 45% and the spread between η regions is somewhat reduced.

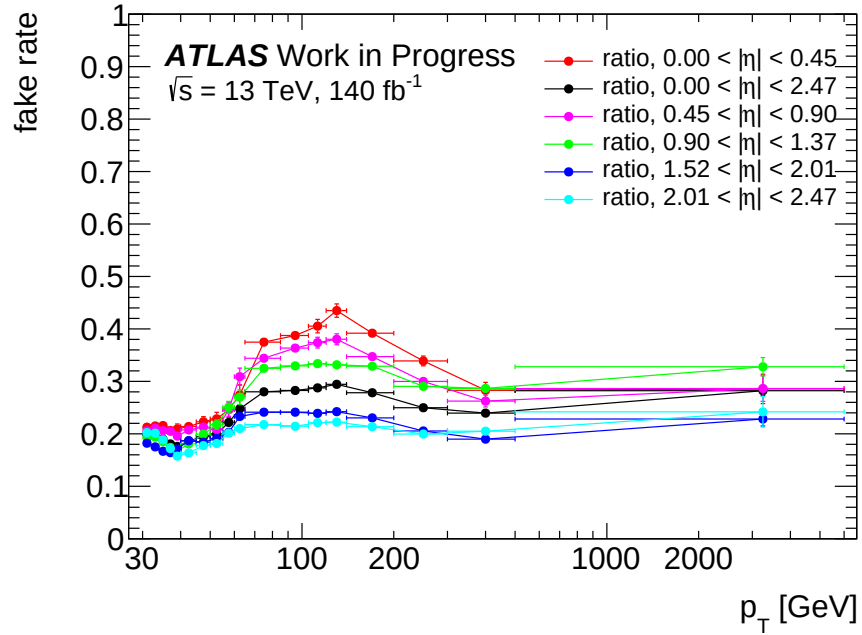
The amount by which the MC subtraction reduces the fake rate depends on the purity of the fake-enriched data sample. To the degree that the cuts used to enrich this sample are aggressive, the real contamination (and hence impact of MC subtraction) will be small, and vice versa. In this particular configuration, it clearly is important to estimate the real contamination (figure) and subtract it prior to evaluating the fake rate.

7.5 Obtaining the m_{ee} template for fake electrons

Having measured the real and fake electron rates, the matrix elements in equation 7.5 can be calculated given the leading and subleading electron p_T and η in a dielectron event. Likewise, the rate matrix can be inverted to predict the actual-basis proportions (mixed RR, RF, FR, FF) underlying this event measured in the observable basis (exclusive TT, TL, LT, LL). This procedure is done on an event-by-event basis to reweight a large data



(a) Before MC subtraction



(b) After MC subtraction

Figure 7.3: Fake rate calculated (a) before and (b) after subtraction of the MC estimate of contamination by real electrons. Different curves show five different η bins and the black one the η -inclusive case.

sample comprised of events passing the full dilepton signal region selection (TT) as well as a loosened selection where one or more electrons fail the tight identification and isolation criteria. Like the sample used to derive the fake electron rate, the EXOT19 derivation is used to avoid the upstream LHLoose requirement built into EXOT0. However this “loosened” data sample is different in that it requires a candidate electron *pair* satisfying the full analysis criteria, such as firing a dielectron trigger. The key differences between these two samples and the signal region used in the main analysis are summarized in table 7.2.

	Data sample		
	fake-enriched	loosened	signal region
Purpose	fake rate	matrix method reweighting	main analysis
Derivation	EXOT19	EXOT19	EXOT0
Trigger	1e (table 7.1)	2e (table 6.5)	2e (table 6.5)
<i>Imposed/required?</i>			
real suppression	Y	N	N
ee candidate pair	N	Y	Y
full ID & isolation	N	N	Y

Table 7.2: Characteristics distinguishing the two data samples used to estimate the fake electron background from the main analysis signal region. The “real suppression” refers to the three vetos and the MC subtraction described in section 7.4.1.

Figure 7.4 shows the elements of the rate matrix in equation 7.5 averaged over events in the loosened data sample and over all three data-taking periods. The largest entry, at row 0 and column 0, indicates that events with two real electrons are 88.4% likely to be selected as events with two tight electrons. This simply expresses the sampling of the real rates shown in figure 7.1 by the kinematics of the electrons in the loosened data sample. The diagonal elements are all at least twice as large as others in their row and column. This essentially encodes the degree to which loose events are correlated with fake events and vice-versa.

Each event in the loosened data sample translates to a $\vec{N}_{\text{obs}}$ unit vector with a single entry filled with “1”. It is then rotated into the actual basis by the inverted rate matrix (equation 7.7) to produce a set of four weights. Single-event yields are frequently negative since the matrix method is properly defined on average yields over an ensemble of events, not individual events. Figure 7.5 shows the proportions of events in the actual basis which are predicted to populate the four observable basis categories. Because of the use of dielectron triggers and the baseline quality criteria, the loosened data sample contains 77% events with both electrons passing the tight identification and isolation criteria. As expected, this TT majority is dominated by events with two real electrons but also contains 3.6% events with at least one fake electron. Conversely, for example, the matrix method also predicts

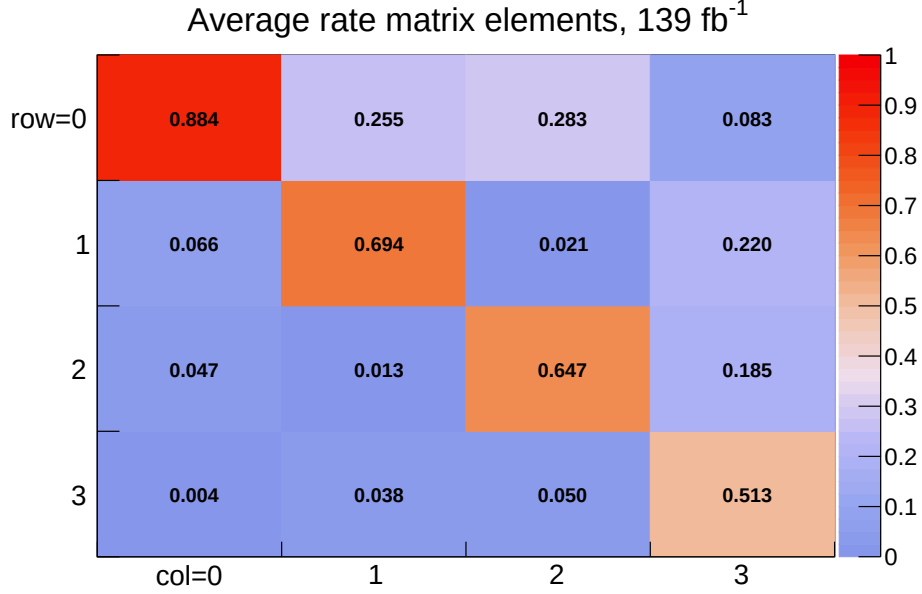


Figure 7.4: Average values of the rate matrix elements.

that only 6.9% of the selected events where both electrons *fail* the tight criteria (LL) are events with two real electrons.

The sum of the three non-RR slivers inside the TT pie chart represent the three terms on the right hand side of equation 7.8. Their m_{ee} spectrum is constructed by applying the RF, FR, and FF weights to each event in the loosened data sample. The $m_{\ell\ell}$ spectrum of events in the loosened data sample is blinded above 1 TeV since it is expected to contain more data events at high mass than the dilepton SR. Note that the latter is approved for unblinding partly because it *does* require the nominal analysis cuts on electron ID and isolation. The resulting weighted $m_{\ell\ell}$ distribution is shown in figure 7.6 up to 1 TeV. A smooth template is obtained by a maximum likelihood fit with the following four-parameter function up to 3010 GeV:

$$f_{\text{fakes}}(x) = (1 - x)^{g_1} \cdot x^{g_2 + g_3 x + g_4 (\ln x)^2} \quad (7.11)$$

where $x \equiv m_{ee}/(13 \text{ TeV})$ and the function is normalized to the distribution. The upper value of the fit range is chosen for fit stability which was sensitive to bins with negative weights showing up in the low-statistics tail. The best-fit values for the parameters are summarized in table 7.3 Since the three fake categories produce different m_{ee} spectrum shapes it was also tested fitting each one separately, but this does not significantly improve the fit quality. Finally, the extrapolation up to 6 TeV is used for the “Jets & non-prompt e” template in figure 6.9a.

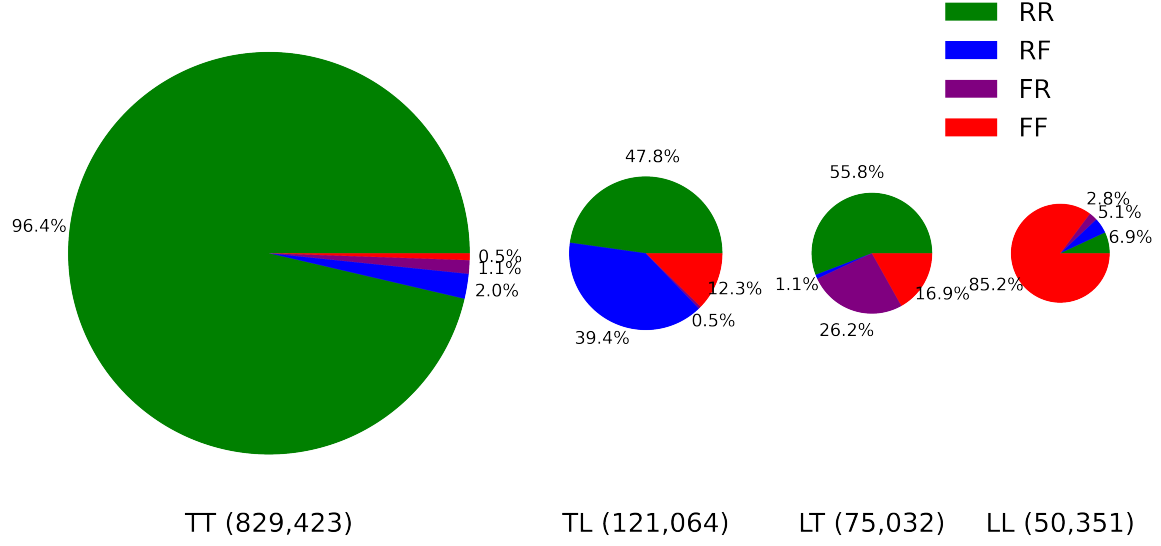


Figure 7.5: Predicted mixing of event yields in observed (TT, TL, LT, LL) and actual bases (RR, RF, FR, FF). The relative sizes of the pie charts are in proportion to the total yields in the four observable-basis categories which are given in the parentheses below them. Note that labels such as “TT” in this case refer not to events which in fact passed the corresponding category but to the *predicted* yields based on back-substituting the answer of equation 7.7 into equation 7.5.

parameter	best-fit value
g_0	9.0716 ± 1.93
g_1	-3.2941 ± 0.161
g_2	0.36776 ± 0.0192
g_3	0.066711 ± 0.00355

Table 7.3: Fitted parameter values of the function in equation 7.11.

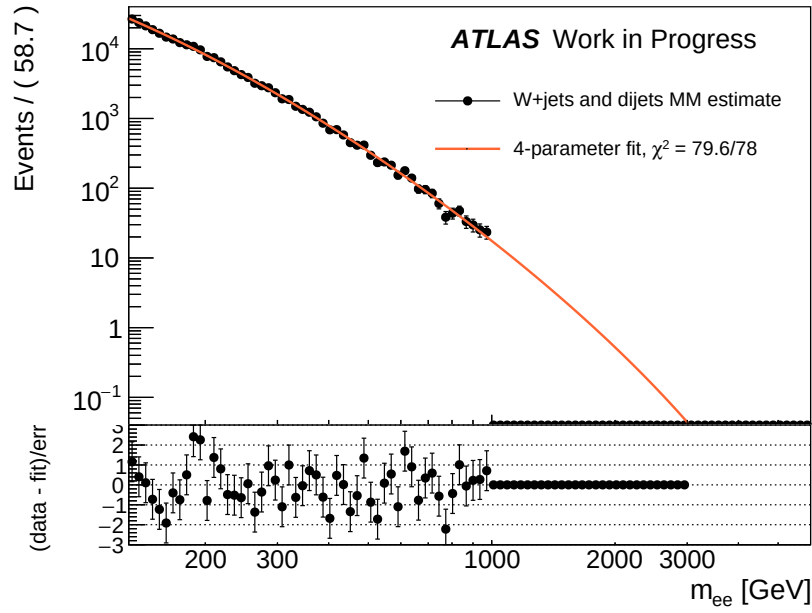


Figure 7.6: Functional fit used to smooth and extrapolate the $m_{\ell\ell}$ estimate of fake electrons from the matrix method. The fit range is between 130 GeV and 3010 GeV.

Chapter 8

Statistical analysis

This chapter discusses the process of drawing statistical conclusions on models of BSM physics from the observed data. For the dilepton analysis, the invariant mass distribution derived in chapter 6 is the main input for the statistical analysis. In the first sections, the general concepts of discovery statistics will be reviewed in the context of LHC searches. Then the statistical analysis strategy for the full Run-2 dilepton analysis (based on a global functional form fit) will be motivated and described. Finally, the results of the search and limit-setting phases of this analysis will be presented.

8.1 Discovery statistics

8.1.1 Hypothesis testing

As explained in chapter 6, various BSM scenarios can show up in the dilepton final state as a distortion (typically enhancement) of the kinematic distributions expected for SM background processes. But in practice, these backgrounds cannot be modelled with 100% accuracy, and also include statistical uncertainty depending on the amount of simulated events that are generated. For this reason there are potentially multiple hypotheses able to explain an observed signal-like feature. In order to establish the hypothesis of interest – namely, that a new signal is responsible – it must be shown that the data are sufficiently unlikely to arise from the background processes and their statistical and systematic variations¹. The latter constitutes the *null* hypothesis which must be rejected in favor of the *alternative*:

Null Hypothesis, $H_0 \Leftrightarrow \{\mu = 0, \theta = \hat{\theta}\}$: The observed data resulted from purely the background distribution and its inherent variations

Alternative Hypothesis, $H_1 \Leftrightarrow \{\mu = \hat{\mu}, \theta = \hat{\theta}\}$: The observed data resulted from a combination of background and signal distributions

¹Note that a frequentist, rather than Bayesian approach is taken throughout this chapter.

The parentheses denote for each hypothesis the final values of the *parameter of interest* (POI) μ controlling the strength of the signal and the vector of *nuisance parameters* $\boldsymbol{\theta}$. The “hats” are explained below. Note that in other cases, such as cross section measurements, some fixed reference value for the POI other than zero may define the null hypothesis, but this discussion is focused on the case of searches for new signals. Nuisance parameters (NPs) are used to parameterize variations of the background distribution caused by systematic effects (energy calibration of electrons, uncertainty in the parton distribution, etc.). Under the null hypothesis, the signal strength is assumed to be zero and the data are explained by some combination of the NPs within their allowed ranges and/or simply statistical fluctuations of the background. On the other hand, the alternative hypothesis H_1 can invoke not only the NPs, but also nonzero values of the signal strength to explain observed deviations from the expected background distribution.

Given a dataset of N bins, the ability of each hypothesis to account for these observations can be quantified using the following likelihood function based on the Poisson probability p :

$$L(\text{data}|H) = \prod_{i=1}^N p(n_i ; m_i(\mu, \boldsymbol{\theta})|H) \cdot \prod_{k \in \text{syst.}} U(\theta_k), \quad p(n ; m) \equiv \frac{m^n e^{-m}}{n!} \quad (8.1)$$

where n_i refers to the observed content of bin i and $m_i(\mu, \boldsymbol{\theta})$ is the bin’s expected content as a function of the POI and NPs and under the assumption that hypothesis H is true. Event yields for a specific process are Poisson-distributed because events are independent from one another and occur with an identical and small probability, determined by the integrated cross section and luminosity. For regions of the measurement where an additive signal component would appear (i.e. the signal region), each bin is expected to contain

$$m_i(\mu, \boldsymbol{\theta}) = b_i(\boldsymbol{\theta}_b) + \mu s_i(\boldsymbol{\theta}_s) \quad (8.2)$$

where s and b are signal and background distributions which are varied by (potentially overlapping) sets of nuisance parameters. In general, however, some subset of the bins in (8.1) may come from a control region used to help constrain systematics in a region devoid of signal, i.e. $m_i = c_i(\boldsymbol{\theta})$.

The second factor in equation 8.1 is a product over the sources of systematic uncertainties which are parameterized by the vector of NPs. Rather than being able to take any real value, each NP is assumed to lie on a probability distribution U . The constraint represents auxiliary knowledge about the size of the uncertainty. In typical cases a normal distribution is a suitable constraint factor.

The hypotheses are most credible when their corresponding likelihood function is maximized. Parameter values which maximize the likelihood (i.e. best-fit values) under the alternative hypothesis are denoted by “hats”: $\{\mu = \hat{\mu}, \theta = \hat{\boldsymbol{\theta}}\}$. When the signal is constrained

to be zero under the null hypothesis, the NPs will generally assume different “double hat” values $\hat{\hat{\theta}}$ at the new likelihood maximum. A ratio of the two likelihood maxima forms a simple test statistic, i.e. a quantity which is the basis for deciding whether to reject the null hypothesis in favor of the alternative:

$$\lambda^{-1}(\mu) = \frac{L(\text{data}|\hat{\mu}, \hat{\theta})}{L(\text{data}|\mu, \hat{\hat{\theta}})} \quad (8.3)$$

The form shown corresponds to a null hypothesis defined by a fixed reference value of the POI μ , which, for a search is $\mu = 0$. The reciprocal $\lambda(\mu)$ is known as the *profile likelihood ratio*. According to the Neyman-Pearson Lemma [132], a related test statistic is optimal for the purpose of constructing a hypothesis test, the so-called *log-likelihood ratio* (LLR):

$$q_0 \equiv -\ln \lambda^2(0) = \ln \left[\frac{L(\text{data}|\hat{\mu}, \hat{\theta})}{L(\text{data}|\mu = 0, \hat{\hat{\theta}})} \right]^2 \quad (8.4)$$

where the subscript 0 indicates the null hypothesis of $\mu = 0$. Saying the LLR is “optimal” means that it provides the largest statistical *power*, defined as the probability to reject the null hypothesis given that the alternative is in fact true. This is equivalent to saying it minimizes the test’s probability β of “false negative” outcomes, or type II errors:

Type I error (“false positive”): Test rejects H_0 although in fact H_1 is not true (occurs with probability α)

Type II error (“false negative”): Test fails to reject H_0 in the case that H_1 is true (occurs with probability β)

But what about type I errors? Any hypothesis test has some nonzero probability α that the null hypothesis will be rejected when the alternative is not in fact true. To clarify, then, the use of the LLR test statistic maximizes the power $(1 - \beta)$ for a fixed value of α . Depending on how much is at stake for reporting a false positive, the test is configured to admit an acceptably small α . For tests of new physics at the LHC, the bar is set high to avoid the risk of reporting a false positive. Two thresholds for α are defined: *evidence*, corresponding to $\alpha \simeq 0.13\%(3\sigma)$ and *discovery/observation* at $\alpha \simeq 3 \times 10^{-7}(5\sigma)$.

To determine whether a given observation justifies rejecting the null hypothesis (passing, say, the threshold for evidence) the corresponding p -value must be calculated. It is defined as follows for a given observed value of the test statistic $\hat{q}_0$:

$$p_0 = p(q_0 \geq \hat{q}_0 | H_0) = \int_{\hat{q}_0}^{\infty} f(q_0 | H_0) dq_0 \quad (8.5)$$

where $f(q_0 | H_0)$ is the underlying probability distribution of the test statistic under the null hypothesis. Thus the p -value can be understood as the odds of drawing a dataset from

purely background processes which is at least as discrepant as the observed value, measured by the test statistic. The outcome of the hypothesis test then depends on whether the p -value of an observation, based on the chosen test statistic, is smaller than the agreed-upon threshold α . A general-purpose method to calculate the p -value is to generate a large set of pseudo-experiments (a.k.a. *toys*) from the known distribution of the observable under the null hypothesis, for example, MC background templates. After fitting each toy with the $s+b$ model, a distribution of the test statistic q_0 can be populated and integrated to get the p -value for a given observation. However, under certain conditions, a much faster method is available based on the expected analytic form of the test statistic distribution (see [133]).

Under the null hypothesis, a sufficiently-large sample is expected to exhibit statistical fluctuations both upward and downward with equal probability. Assuming for now we are only interested in a positive signal, the signal strength can be capped $\mu \geq 0$ such that upward fluctuations will produce $\hat{q}_0 > 0$ and everything else is lumped into $\hat{q} = 0$. Thus the test statistic distribution can be split into two equal-size populations:

$$f(q_0|H_0) = \frac{1}{2}\delta(q_0) + \frac{1}{2}f_+(q_0 > 0|H_0) \quad (8.6)$$

$$\text{where } f_+(q_0 > 0|H_0) = \chi_k^2(q_0) = \frac{q_0^{k/2-1}e^{-q_0/2}}{2^{k/2}\Gamma(k/2)} \quad (8.7)$$

The second line follows from Wilks' theorem [134] in the limit of large sample sizes (a.k.a. the *asymptotic limit*), which states that the positive values of the test statistic follow a chi-square distribution. The k degrees of freedom are equal to the number of POIs that distinguish the alternative hypothesis from the null one. The case being considered is of a single POI, the signal strength μ . Figure 8.1 compares the distribution of q_0 values derived from 100k toys with the asymptotic expression $\chi_{k=1}^2(q_0)$ for a single POI. The first bin integrates the delta function spike at $q_0 = 0$ and thus contains more than half of the toys.

The p -value under the distribution in figure 8.1 can be expressed by substituting (8.7) with $k = 1$ in (8.5):

$$p_0 = \frac{1}{2} \int_{\hat{q}_0}^{\infty} \chi_{k=1}^2(q_0) dq_0 = \frac{1}{2} \int_{\hat{q}_0}^{\infty} \frac{q_0^{k/2-1}e^{-q_0/2}}{2^{k/2}\Gamma(k/2)} \Big|_{k=1} dq_0 = \frac{1}{2\sqrt{2\pi}} \int_{\hat{q}_0}^{\infty} \frac{e^{-q_0/2}}{\sqrt{q_0}} dq_0 \quad (8.8)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{t=\sqrt{\hat{q}_0}}^{\infty} e^{-t^2/2} dt \quad (8.9)$$

$$\therefore p_0 = 1 - \Phi(\sqrt{\hat{q}_0}) \quad (8.10)$$

where Φ is the normal cumulative distribution function, following the substitution $t = \sqrt{q_0}$ in the second line. Thus, in the case of one POI and the asymptotic limit the p -value is simply the Gaussian tail integral to the right of $\sqrt{\hat{q}_0}$. The translation into Gaussian significance is therefore trivial:

$$z_0 = \sqrt{\hat{q}_0} \quad (8.11)$$

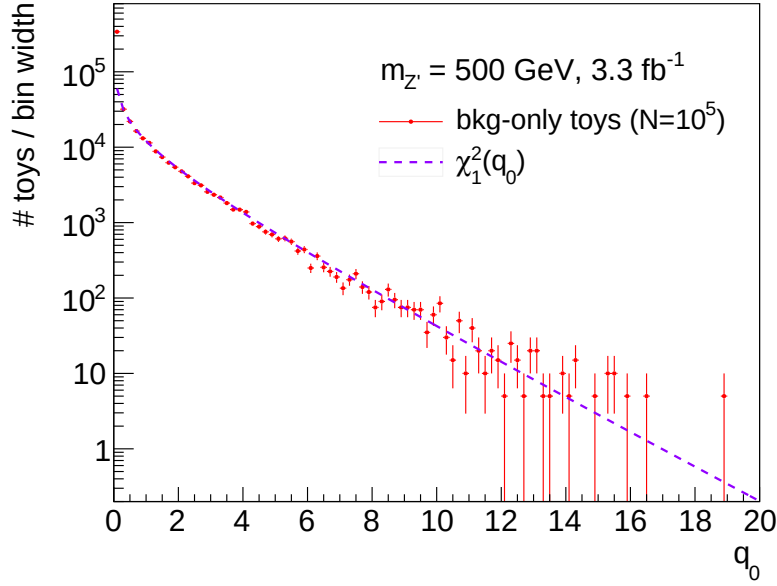


Figure 8.1: Comparison of the LLR test statistic distribution modelled using toys and in the asymptotic limit using the χ^2_1 distribution. Each of the 10^5 toy datasets is drawn from the expected background template normalized to 3.3 fb^{-1} and used for a hypothesis test at pole mass of 500 GeV. This is a region sufficiently populated to justify the asymptotic approximation, as can be seen by the agreement between the two distributions. The bin width is 0.2.

A final simplification can be realized using the result of Wald [135] (as cited in ([133], equation 17)): assuming the observed POI value $\hat{\mu}$ is Gaussian-distributed with standard deviation σ_μ about a central value μ' , the observed LLR tends to

$$q_0 = \frac{\hat{\mu}^2}{\sigma_\mu^2} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad (8.12)$$

where N is the size of the data sample. Up to the correction term, then, (8.12) substituted into (8.11) yields

$$z_0 \simeq \frac{\hat{\mu}}{\sigma_\mu} \quad (8.13)$$

On a practical level, $\hat{\mu}$ is directly the best-fit value returned by the log-likelihood minimizer while σ_μ is its estimated uncertainty. The latter is calculated by propagating the uncertainties of all fit parameters (including NPs) into the POI through the covariance matrix.

Given that the signal strength is constrained to be positive, the LLR test statistic in equation 8.4 assigns to negative fluctuations an equal likelihood of arising under either of the null or alternative hypotheses. The maximum p -value is thus 50%. While resonance-like signatures in the downward direction are not physical (apart from strong interference

effects), it is helpful to quantify them in order to compare with the frequency of positive fluctuations and to identify potential systematic effects. A modified two-sided test statistic is thus defined as follows:

$$\tilde{q}_0 \equiv \begin{cases} + \ln \left[\frac{L(\text{data}|\hat{\mu},\hat{\theta})}{L(\text{data}|\mu=0,\hat{\theta})} \right]^2, & \hat{\mu} \geq 0 \\ - \ln \left[\frac{L(\text{data}|\hat{\mu},\hat{\theta})}{L(\text{data}|\mu=0,\hat{\theta})} \right]^2, & \hat{\mu} < 0 \end{cases} \quad (8.14)$$

Hence $\tilde{q}_0$ will be distributed as a two-sided χ^2_1 – figure 8.1 mirrored about the y -axis. It is shown by the red curve in figure 8.2. In the two-sided definition, downward fluctuations are characterized by p -values greater than 50% and negative Gaussian significance. The latter follows simply by allowing $\hat{\mu}$ to be a negative quantity in equation 8.13.

Note on the look-elsewhere effect

Searches which perform a grid scan of hypothesis tests across one or more signal parameters (e.g. the pole mass) must account for the *look-elsewhere effect*. Because the scanned parameter is not known a priori there is a greater probability for the background to significantly fluctuate *somewhere* in the scan as opposed to at one particular point. Consequently, the largest significance found in a scan must be converted to a *global* significance accounting for the increased likelihood that it could be merely a background fluctuation. The effect is substantial for the dilepton analysis because of the size of the signal hypothesis scan and the well-resolved observable. Refer to appendix B for a more detailed discussion and studies of the look elsewhere effect in the dilepton analysis.

8.1.2 CL_s limit setting

When a search for new physics fails to reject the null hypothesis – that *zero* signal is present – a follow-up set of tests is usually defined to determine which hypotheses of *non-zero* signal strength can be rejected. Given that no signal-like feature is observed, it intuitively becomes doubtful that a large signal could have been veiled by a fluke downward fluctuation in data. Thus, cases of null results in searches can be translated into bounds (a.k.a. constraints) on model parameters wherein they remain statistically compatible with the measurement. For a given model, parameters such as BSM coupling and mediator mass can be constrained through their relation to the signal cross section. In turn, the maximal value of the signal cross section which is deemed compatible with the analysis results is proportional to (in our case equal to) the upper limit on the POI. Setting an upper limit on the POI can be posed as the following problem:

How large can the true POI be and still yield – in $100 \cdot (1 - \alpha)\%$ of trials – observations that are as or less signal-like than the one in question?

where the percentage is referred to as the *confidence level* (CL) of the limit. Typically ATLAS analyses adopt 95% CL (i.e. $\alpha = 0.05$), meaning that any POI value above the upper limit would have produced more significant results with at least 95% probability. As with calculating p -values, a frequentist approach to setting limits can be done in a brute force way, using toy ensembles, and in the asymptotic limit using closed-form expressions.

So far we have discussed the distribution of q_0 (equation 8.7) for the case of background-only data – i.e. true signal strength of $\mu' = 0$ – where its median is 0. If, however, the alternative hypothesis is true, $\mu' > 0$ and the q_0 distribution shifts toward positive values according to equation 8.12. Example distributions of q_0 for different values of underlying μ' are shown in figure 8.2. Also illustrated, in figure 8.2b is the p -value associated with rejecting the null hypothesis, p_b , which has already been discussed. However, the most relevant p -value when deciding whether to reject a hypothesis of non-zero signal strength is the one shown labeled as p_{s+b} . In the example shown, $p_{s+b} = 0.12$ which means that one can expect to observe $\hat{q}_0 \leq 3$ in 12% of trials if an actual signal of strength $\mu' = 16$ is present. Equivalently,

$$\text{CL}_{s+b} : (1 - p_{s+b}) = 88\% \quad (8.15)$$

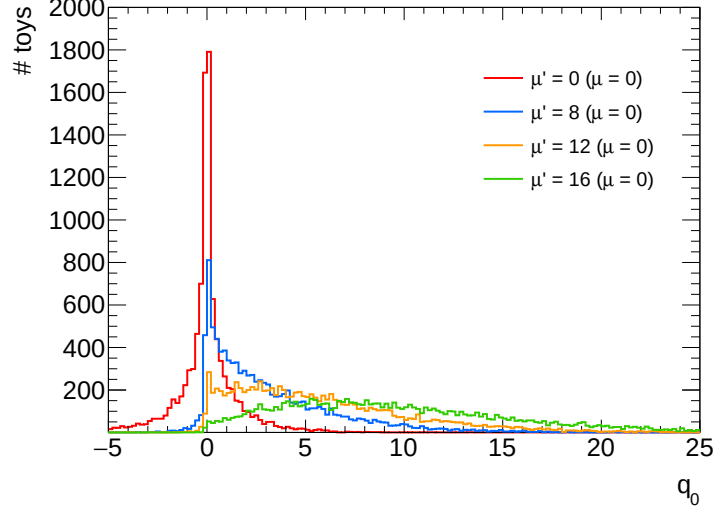
which does not meet the 95% CL threshold for rejecting the $\mu = 16$ hypothesis. The upper limit on μ is thus higher than 16 and can be found by generating test statistic distributions with larger underlying signals until the tail pinches off below $p_{s+b} < \alpha = 0.05$.

The above description has basically explained the procedure used in ATLAS to set upper limits, except for two modifications: (1) a different test statistic, q_μ , is actually used and (2) the α threshold is checked using a modified tail probability, CL_s . For the first point, since it is some $\mu > 0$ hypothesis which an upper limit is meant to reject, it is more fitting to use a test statistic associated with this fixed value of μ (rather than q_0 which tests the $\mu = 0$ hypothesis).

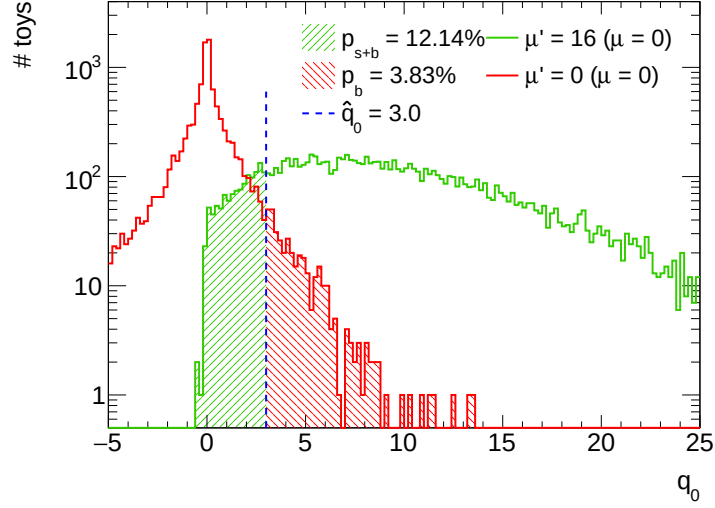
$$q_\mu \equiv \begin{cases} -2 \ln \lambda(\mu), & \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu \end{cases} \quad (8.16)$$

In this definition the null hypothesis is not $\mu = 0$ but rather “ $\mu = \mu > 0$ ”. Observed $\hat{\mu}$ that are lower than the hypothesized μ translate to larger values of q_μ . Conversely, observations that are larger than μ are not considered to count against the null hypothesis, so $q_\mu = 0$ in such cases.

Even though based on the q_0 test statistic, the same result can be obtained using an analogous calculation of p_{s+b} and p_b using q_μ , but it is arguably a bit less intuitive. Appendix B.2 shows distributions of q_μ to validate this statement and the analogous version of figure 8.2b.



(a)



(b)

Figure 8.2: Distributions of the LLR test statistic q_0 for 10k toys generated for different injected signal strengths μ' . (a) shows how the distribution shifts upward with μ' compared to the 2-sided χ^2_1 distribution generated under the background-only hypothesis. (b) illustrates the two p -values used in the CL_s method for the signal strength hypothesis of $\mu' = 16$ and an example observed $\hat{q}_0$ of 3. Each toy dataset (with 200k events) was both generated and then fit at a fixed mass using a simple $s + b$ model based on an exponential background.

Regarding the second modification, ATLAS sets upper limits using the following quantity which incorporates both p_{s+b} and p_b :

$$\text{CL}_s \equiv \frac{p_{s+b}}{1 - p_b} \Big|_{\text{req.}} < \alpha \quad (8.17)$$

The use of CL_s is meant to make the α threshold more difficult to reach for observations which are easily accommodated under the background-only hypothesis (large p_b) [136]. This helps prevent excluding signal hypotheses which the analysis does not sufficiently distinguish from the background-only hypothesis. On the other hand, for observations with a very small probability under the null hypothesis the CL_s method becomes the same as CL_{s+b} . For example, considering again figure 8.2b, the CL_s method is very similar to the p_{s+b} since $p_b = 0.0383$ is small such that the denominator is close to 1.

Fitting a complicated model to thousands of toy datasets per μ' in the upper limit scan to generate the q_μ distributions for upper limits can be very computationally expensive. That is especially true for limit-setting in a scan over signal parameters such as pole mass. Once again an asymptotic form for the distributions of q_μ can be invoked in the large sample limit (hence the following is not valid for regions of low data yields such as tails). Equation 56 of [133] gives the asymptotic q_μ distribution when calculated for datasets that have underlying/injected $\mu' = \mu$. It turns out to be the same as in equation 8.7 for the q_0 case: a (half) chi-square distribution. Figure B.9 in appendix B.2 demonstrates this for q_μ distributions with $\mu' = 8, 12, 16$. Therefore the same shortcut derived in equation 8.10 can be used to calculate p_{s+b} for upper limits:

$$p_{s+b} = 1 - \Phi(\sqrt{\hat{q}_\mu}) \quad (8.18)$$

Note that for a given observed $\hat{\mu}$ the value of $\hat{q}_\mu$ is not the same as $\hat{q}_0$ and needs to be calculated using both a fit to the dataset at the hypothesized μ fixed and at μ free.

8.2 Global fit

Beyond the Z peak, SM processes generate a smoothly-falling dilepton invariant mass spectrum populated in data across a span of over 2 TeV. The least monotonic background is the $t\bar{t}$ process but its first derivative nonetheless remains negative throughout. The breadth and smoothness of the SM background distribution in contrast to a narrow, localized resonance make it possible to construct a search based on a data-driven background estimate, that is, a functional form fit to data. This approach has been adopted for the first time in the full Run-2 iteration of the ATLAS dilepton analysis for reasons which will be motivated in the following section.

8.2.1 Motivation

Exceptional measures have been taken to model the Drell-Yan differential cross section at high mass with high statistical precision (mass-binned samples) and inclusion of higher-order effects through various K -factors (section 6.2.1). These steps made it possible to accurately model the expected dilepton invariant mass spectrum of the full Run-2 ATLAS

dataset, as shown in figure 6.9. However, generating the necessary sample size to keep up with the integrated luminosity of Run 2 raised the question of whether a more sustainable approach could be developed for modelling the SM backgrounds in the dilepton analysis.

During the statistical analysis of the dilepton spectrum at 36.1 fb^{-1} it was noted that the mass-binned Drell-Yan samples provided suboptimal statistical precision in the region below about 1 TeV. In order to discern small departures of the data from expected SM yields, a much larger number of background events must be generated to avoid statistical fluctuations which obscure the signal. Ideally the number of MC events should be $N_{\text{MC}} \gtrsim 100 \times N_{\text{data}}$ to achieve a relative statistical uncertainty of order 10%:

$$\frac{\sigma_{\text{rel}}(\text{MC})}{\sigma_{\text{rel}}(\text{data})} = \frac{\sqrt{N_{\text{MC}}}/N_{\text{MC}}}{\sqrt{N_{\text{data}}}/N_{\text{data}}} = \frac{\sqrt{100N_{\text{data}}}/(100N_{\text{data}})}{\sqrt{N_{\text{data}}}/N_{\text{data}}} = 0.1 \quad (8.19)$$

But MC statistical fluctuations may also be acceptable that are $\sim 20\%$ the size of those in data, corresponding to a MC sample that contains 25 times the amount of events expected in data. Unfortunately, the situation encountered in the low-mass end of the dilepton spectrum did not even reach this 20% figure. Figure 8.3 shows the ratio of relative statistical errors on MC and data as a function of the invariant mass. Clearly, suboptimal MC statistical precision was also a problem for the top backgrounds (mainly $t\bar{t}$).

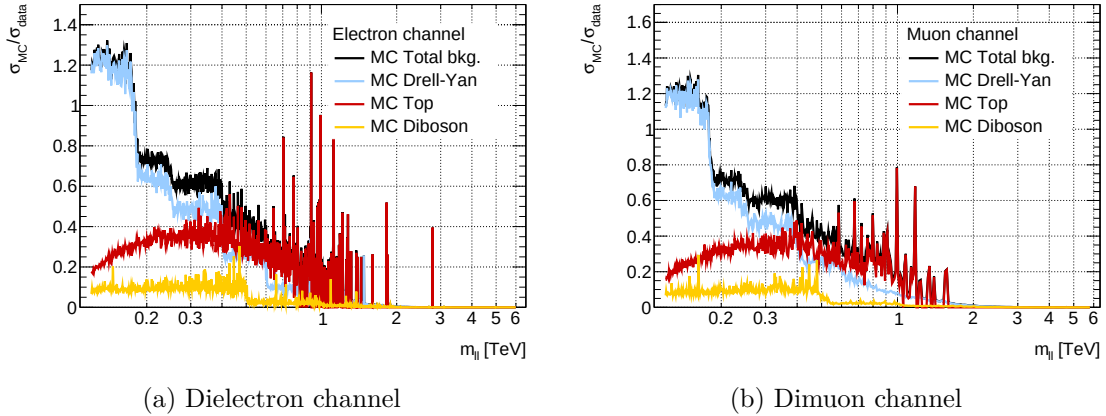


Figure 8.3: Ratio of relative statistical uncertainty of MC to data for the three main MC backgrounds in the 36.1 fb^{-1} analysis. Ideally the ratio should be at most around the 10% level, while up to 20% may also be acceptable. Figures taken from [129].

As a result, the “bumpiness” inherent in the MC-based background estimate reduced the sensitivity of the 36.1 fb^{-1} analysis to small resonances. For example, a real excess present in data would be partially obscured if it coincided with an upward statistical fluctuation in the MC background. The impact was modelled by introducing nuisance parameters in the likelihood which allow the MC bin contents to vary within their statistical uncertainty. The sensitivity loss is shown in table 8.1 for the electron channel. Enlarging the Drell-Yan samples to reach the 10% level would require on the order of 100 million additional MC

Mass [GeV]	Sensitivity loss
150	35%
200	18%
300	13%
500	6%

Table 8.1: Sensitivity loss estimated for different pole masses resulting from statistical uncertainty on the MC background templates in the 36.1 fb^{-1} dilepton analysis. In the end the impact was reduced by a smoothing procedure applied to the templates.

events. The option is clearly computationally very expensive and could become a bottleneck for moving the analysis forward. As a temporary solution in the 36.1 fb^{-1} analysis, a functional form fit was applied to smooth the Drell-Yan MC template between 120 GeV and 1 TeV so that statistical fluctuations could be reasonably neglected.

At the same juncture where a directly MC-based background estimate was becoming problematic, the data-driven alternative became feasible. The successful use of a functional form fit to smooth the MC background opened the door for considering a fit directly to data. In principle, the shape of the background extracted using a functional form fit to data becomes better constrained as a larger amount of data is used. This of course assumes that an appropriate functional form can be chosen which does not introduce mismodelling. In order to check that, an accurate (if not optimally precise) MC background estimate remains highly valuable for validating and optimizing the function choice in an unbiased way. For the functional form used in the full Run-2 analysis, several such checks were done, as described in the following section. Furthermore, a sensitivity comparison was done by overlaying the expected signal cross section limits based on a data-driven analysis and based on MC templates.

Finally, it should be noted that the data-driven strategy includes validation using background MC samples and is therefore not fully immune to uncertainties in those samples. This will be explained in section 8.2.4.

8.2.2 Description

Two complementary frameworks were investigated for the functional background modelling: an iterative technique based on a sliding window fit (SWiFt), and a global fit to the entire signal region. While the SWiFt approach was more robust (i.e. able to model all parts of the spectrum) it turned out to be slightly less sensitive in terms of expected signal cross section limits. In the future, larger datasets will be more challenging to fit and this is likely the situation where the SWiFt approach will have the advantage. The SWiFt methodology and its comparison to the global fit approach is documented in appendix C.

An assortment of functions were tested including those with forms similar to what has been used in past dijet and diphoton resonance searches. In the end the following choice

was found suitable to describe both dielectron and dimuon invariant mass spectra:

$$f_{\text{bkg}}(x) = a \cdot \text{BW}_Z(x) \cdot (1 - x^c)^b \cdot x^{\sum_{i=0}^3 p_i (\ln x)^i} \quad (8.20)$$

where

$$x = \frac{m_{\ell\ell}}{\sqrt{s}}, \quad c = \begin{cases} 1 & , ee \\ 1/3, & \mu\mu \end{cases}, \quad a, b, p_i : \text{free parameters} \quad (8.21)$$

and

$$\text{BW}_Z(x) = \frac{\Gamma_Z}{(m_Z - x \cdot \sqrt{s})^2 + \Gamma_Z^2} \quad (8.22)$$

with PDG values for the Z boson mass and intrinsic resonance width of $m_Z = 91.1876$ GeV and $\Gamma_Z = 2.4952$ GeV, respectively. Firstly, the observable x of the function is a rescaling of the invariant mass by the LHC centre-of-mass energy. Next, notice that the core of the function (8.20) is a 4th-order polynomial in “log-log” space:

$$\underbrace{\ln f_{\text{bkg}}}_Y = \sum_{i=0}^3 p_i \underbrace{(\ln x)}_X^{i+1} \quad (8.23)$$

The space Y vs. X is the usual choice for visualizing the dilepton spectrum in figures such as (6.9), and is also a practical space in which to fit a functional model. An additional factor of $(1 - x^c)$ was found to help provide stability of the fit in the high-mass tail of the spectrum, with the c hand-tuned for goodness of fit in each channel. Note that although the SM $m_{\ell\ell}$ spectra in both channels are essentially identical at truth level, significant differences in their shape and normalization are introduced by reconstruction and selection efficiency as well as detector resolution. Up to this point the form of the function is identical to the one used to model the diphoton invariant mass spectrum in ATLAS [137]. This choice was in turn inspired by earlier searches in the dijet final state by the CDF collaboration [138]. But in application to the dilepton final state, the somewhat *ad hoc* functional form is multiplied by a clearly physics-motivated piece BW_Z describing the Z Breit-Wigner resonance shape.

Fit configuration

The full background plus signal model is fit to the data by minimizing the following full likelihood function:

$$L(m_X, \Gamma_X | \sigma_X, \vec{p}, \boldsymbol{\theta}) = \prod_{i=1}^N P(n_{\text{obs}}^i, n_{\text{exp}}^i) \prod_{k \in \text{syst.}} \text{Gaus}(\theta_k, 0, 1) \quad (8.24)$$

with

$$n_{\text{exp}}^i = f_{\text{bkg}}^i(\vec{p}) + \left[\sigma_X \cdot A \epsilon(m_X, \Gamma_X, \boldsymbol{\kappa}) \mathcal{L} \cdot \kappa_{\mathcal{L}} + \theta_{\text{ss}} \cdot |N_{\text{ss}}(m_X, \Gamma_X)| \right] \cdot f_{\text{sig}}^i(m_X, \Gamma_X, \boldsymbol{\theta}_{\mu, \sigma}) \quad (8.25)$$

where

- f^i refers to the integral of the function f in the i^{th} $m_{\ell\ell}$ bin
- f_{sig} is the generic signal model (equation 6.11)
- σ_X is the POI, the cross section for the signal of mass m_X and width Γ_X
- $A\epsilon$ is the signal acceptance $\times$ efficiency
- $\mathcal{L}$ is the integrated luminosity of the dataset
- $\vec{p}$ is the vector of background function parameters
- θ are the full set of nuisance parameters controlling the systematic variations
- $\theta_\mu, \theta_\sigma$ are the “shape” nuisance parameters which shift the signal resonance mean and vary its width
- κ are the subset of nuisance parameters which enter as factors multiplying the luminosity and acceptance times efficiency²
- $|N_{\text{ss}}|$ is the parameterized symmetrized spurious signal yield (discussed in section 8.2.3 below)

The global fit is performed on a data histogram of $N = 5775$ linear 1 GeV bins running from 225 GeV up to 6000 GeV. The binning is chosen to be fine enough that the experimental mass resolution limits the shape of the distribution, not binning choice. Recall from figure 6.14 that in the electron channel the mass resolution ranges from about 2 GeV at the low end of the spectrum to ~ 10 GeV at 1 TeV. A fit to an unbinned data distribution would be ideal, but this is computationally unfeasible and no substantial gain is expected beyond what is achieved with the 1 GeV binning.

The **Hfitter** framework [139] is used to implement the global fit described above. **Hfitter** is partly standalone and partly a wrapper of **RooStats** libraries which was previously used for the statistical analysis of the high-mass diphoton spectrum [137].

8.2.3 Validation

This section discusses the methods used to guide the choice of functional form for the global fit and to validate its modelling capability. The concerns addressed are that the function adequately models the entire range of the spectrum (across more than 5 orders of magnitude), whether it can potentially induce spurious signal-like discrepancies, and if it will not be so flexible as to “fit away” a real signal excess in data. The function choice

²In this case the Gaussian constraint has mean of 1 (the nominal factor).

and validation is based on the smooth MC background template obtained via the transfer function smoothing technique discussed in section 6.6.

Spurious signal test

The choice of functional form is guided by a measure of its goodness-of-fit (GOF). Various GOF figures of merit can be used to measure the efficacy of a candidate function to describe a dataset, such as the χ^2 test and the Kolmogorov-Smirnov test. The χ^2 is computed as the sum of squared fit residuals:

$$\chi^2 = \sum_{i=1}^N \left[\frac{f(x_i) - n_i}{\sigma(n_i)} \right]^2 = \sum_{i=1}^N \frac{[f(x_i) - n_i]^2}{n_i} \quad \text{if } n_i \text{ are Poisson-distributed} \quad (8.26)$$

However, the difficulty of interpreting the χ^2 measure in this situation stems from the fact that the bin contents in the validation background templates do not follow the “ $\sqrt{N}$ ” prescription. For example, a fit directly to the mass-sliced Drell-Yan sample incorporates bins with statistical uncertainties that are scaled down by a factor $\sqrt{N_{\text{MC}}/N_{\text{data}}}$, and further this varies from slice to slice. It is even less clear how to incorporate bin errors in the case of a validation template obtained using the transfer function smoothing. It is therefore not straightforward to compare an observed χ^2 value with what is expected from a reference distribution based on Poisson-distributed bin counts.

Instead, while χ^2 values are used to gauge the relative performance of different functions, a custom GOF metric is defined: the spurious signal yield N_{ss} . The spurious signal is defined as the signal yield extracted from a background + signal model fit to a background-only (i.e. MC) dataset. As such, it is a function of the signal mass and width hypothesis. Using N_{ss} as a figure of merit assumes that mismodelling of the background estimate can be effectively targeted by minimizing it in terms of fitted signal yield. The absolute size of the spurious signal is required to be at or below the level of half of a standard deviation on the expected signal yield:

$$\frac{|N_{\text{ss}}|}{\sigma_{N_{\text{ss}}}^{\text{exp}}} = \frac{|\hat{\mu}|}{\sigma_{\mu}^{\text{exp}}} \lesssim 0.5 \quad (8.27)$$

where $\hat{\mu}$ is the best-fit (spurious) signal strength and $\sigma_{\mu}^{\text{exp}}$ its uncertainty. The second expression follows from cancelling the common factors of $A\epsilon \cdot \mathcal{L}$ in the numerator and denominator. The condition can therefore be interpreted as limiting the spurious signal “significance” (via equation 8.13) to be below 0.5σ . The performance of the background function 8.20 in terms of both spurious signal yield N_{ss} and the ratio $|N_{\text{ss}}|/\sigma_{N_{\text{ss}}}^{\text{exp}}$ is shown in figure 8.4.

While for the $\Gamma/m_X = 0\%$ width signals the ratio $|N_{\text{ss}}|/\sigma_{N_{\text{ss}}}^{\text{exp}}$ remains essentially below 0.3, larger values were obtained for wider signals. This is driven by the decrease in sensitivity to wide signals because of less distinction from the background. In any case, the values for signals up to $\Gamma/m_X = 10\%$ width remain small – in the range of 0.05 – 0.35 for the dimuon

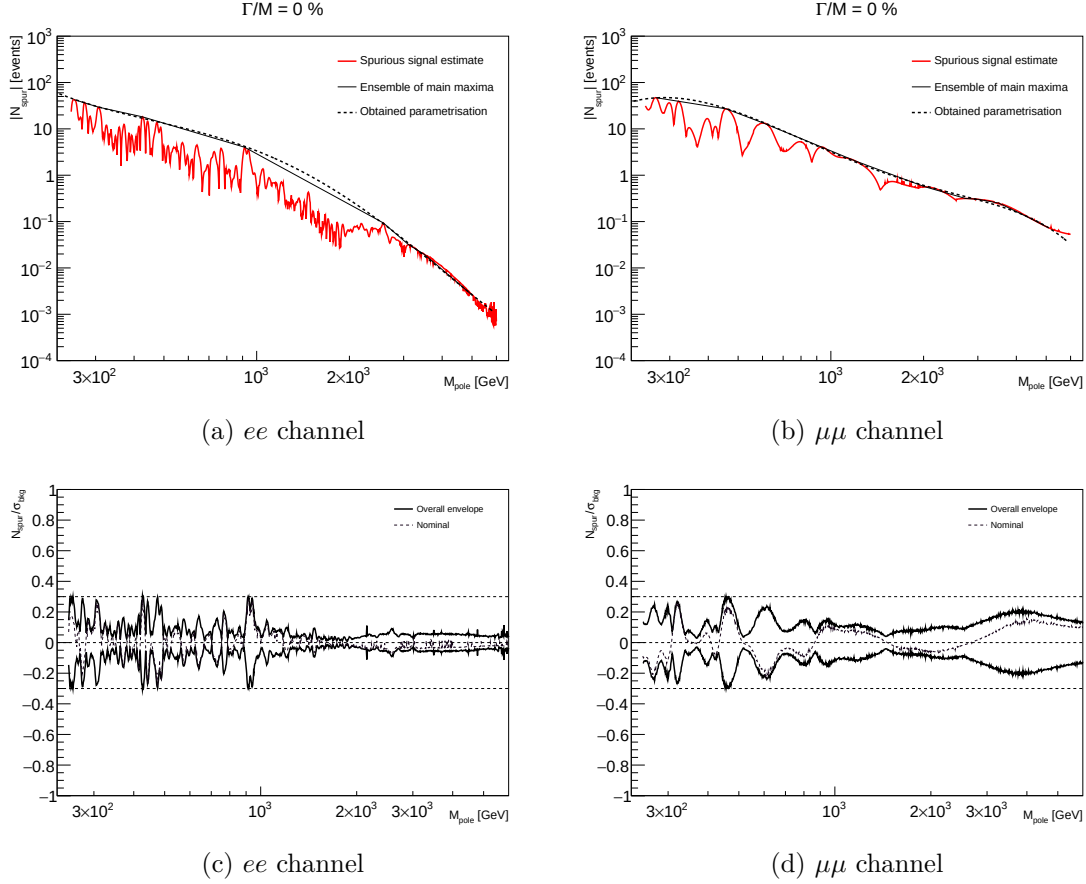


Figure 8.4: Spurious signal absolute yield (top row) and “significance” (bottom row) for the global fit function (equation 8.20) and a $\Gamma/m_X = 0\%$ width signal. The left column is the dielectron channel and the right column is the dimuon channel. Also shown in the plots on the left are a solid line connecting the main maxima and the smooth parameterization of the spurious signal shown as a dashed line. Plots in the bottom row show values obtained from the nominal background template in a dashed line and also those obtained from the envelope of systematic variations in a symmetrized solid line. Figures taken from [129].

channel. For the electron channel this was also mostly true aside from $m_X \sim 950$ GeV where for the widest signals it reached $|N_{\text{ss}}|/\sigma_{N_{\text{ss}}}^{\text{exp}} \simeq 0.5$.

Another assumption underlying the spurious signal test is that the validation template is devoid of signal-like fluctuations. As discussed in section 8.2.1, the presence of statistical fluctuations in the NLO Drell-Yan samples made them less suitable for a smooth background estimate. Similarly, also for the purpose of estimating spurious signal, the MC templates prove inadequately smooth. That is, using them directly to validate the fit will lead to an artificially large spurious signal which points not to background function mismodelling, but to actual statistical fluctuations in the template. This “contamination” can not be fully eliminated but must be reduced to a workable level if possible.

It is therefore the important role of the transfer function smoothing (section 6.6.3) to produce a suitably smooth validation template. However, statistical fluctuations, though considerably smoothed out, are still sizable enough at points that they cannot be ruled out as alternative sources of the measured spurious signal (instead of fit mismodelling). For example a statistical fluctuation in the validation template seems to be the best explanation for the peak in fitted signal in the dielectron channel between 900 – 1000 GeV. For this reason, MC statistical uncertainty still enters the data-driven analysis as a “second-order” effect through its contribution to the size of the spurious signal systematic, which will be discussed in section 8.2.4.

Signal injection test

If minimizing spurious signal was the only criterion for validating a functional form as a background model, it could suggest picking an overly-flexible one. Such a function may have enough parameters to conform to broad features of the smooth validation template, but may also be able to absorb real excesses when fit to data. So it is very important that a data-driven background model be flexible enough to model well the smoothly-falling spectrum but rigid enough that it cannot easily accommodate signal-like features in data.

The expected performance of the global fit function (8.20) in the presence of actual signals was tested by injecting signals into the validation template. A full $b + s$ fit was then done on the resulting mock dataset to extract the underlying signal. The comparison of extracted vs. injected was verified for good linearity across values of the injected signal from zero up to five times the signal cross section of the expected 95% CLs upper limits. Note that the extracted signal for zero injected is simply the spurious signal. The test was done for signal pole masses of 300 GeV and 1000 GeV in the electron, muon, and combined channels.

Additional parameter test

To account for the possibility that the background validation template does not adequately model the full Run-2 dataset after unblinding, a procedure was put in place beforehand to allow for a more flexible background function. In case the nominal function (8.20) turned out to not model the data well, additional parameters multiplying higher powers of $(\ln x)$ in the exponent of x could be introduced if found to substantially improve the background-component fit. The condition is based on a hypothesis test using the LLR of the two fit functions. If the function including the additional parameter improves the fit at a significance of 2σ with respect to the nominal, it is substituted in as the new function choice. Choosing the form of the alternative function prior to looking at the data avoids biasing the analysis. In the end, however, the background fit to the full Run-2 dataset was not found to improve at this level through the addition of a parameter, such that the nominal function (8.20) was therefore used.

8.2.4 Uncertainties

The propagation of systematic and statistical uncertainties into the functional form-based background estimate are discussed below.

Systematic

Since the final background estimate is derived from data, sources of systematic uncertainty which have to do with potential differences between MC and data do not *directly* apply. However the functional form is assumed to have its own modelling limitations which are quantified by the spurious signal. Thus the potential for the background estimate to introduce artificial differences with respect to the data is a source of systematic uncertainty. After vetting the function by the requirement (8.27) the measured spurious signal yield is used to set the size of this systematic. A nuisance parameter θ_{ss} is introduced into the full likelihood function (8.24) as an additional term multiplying the signal model (8.25). θ_{ss} therefore plays an analogous role to the signal strength except for two important differences: it is constrained by a unit-normal penalty term, and it remains free under both the null and alternative hypotheses.

To appreciate how this works, consider an $s + b$ fit to data at the location of a small excess which, for example, happens to be about the size of the spurious signal which was measured at that point on the background template. Under the alternative hypothesis, the signal strength μ (a.k.a. σ_X) can be varied by the minimizer at no cost to the likelihood. However, when $\mu = 0$ is enforced under the null hypothesis, the minimizer can still invoke the same signal component via the θ_{ss} parameter except that it incurs a likelihood penalty from its constraint term, for this example on the order of a 1σ pull. The upshot is that the likelihood fit is partially able to fit using the signal model under the null hypothesis, according to how much spurious signal was measured. As a result, the difference between the numerator and the denominator in the LLR (8.4) is decreased and likewise the significance of the observed excess.

The value N_{ss} varies up and down, as already mentioned, due partly to residual statistical fluctuations in the smoothed MC template. Therefore it needs to be smoothed across the $m_{\ell\ell}$ spectrum to facilitate an NP constraint which does not jump around from mass to mass. A 5-parameter function similar in form to (8.20) – without the Breit-Wigner factor – is used to provide a smooth envelope of the spurious signal as a function of pole mass which is also shown in subfigures 8.4a and 8.4b.

Now, besides statistical fluctuations, the MC used for the smooth background template is subject to a range of systematic variations. Variations of the MC (which may prove more true to the shape of data) might be more difficult for the function to fit. Hence the spurious signal obtained from the nominal template in principle underestimates the actual fit modelling uncertainty. To propagate systematic variations of the MC, the spurious

signal envelope is expanded to bound the $N_{ss}(m_X)$ measured not only on the nominal MC background template, but also its ensemble of systematic variations. Systematic variations on background MC therefore translate into “uncertainties on an uncertainty” – that is, they enter at second order via an inflated spurious signal systematic.

Figure 8.5 shows an example of N_{ss} obtained using templates varied by the systematic uncertainty in α_s . An analogous figure exists for each of the theoretical uncertainties listed in table 6.7. The overall envelope floating on this whole set of varied templates is also shown (which is the same curve shown in figure 8.4).

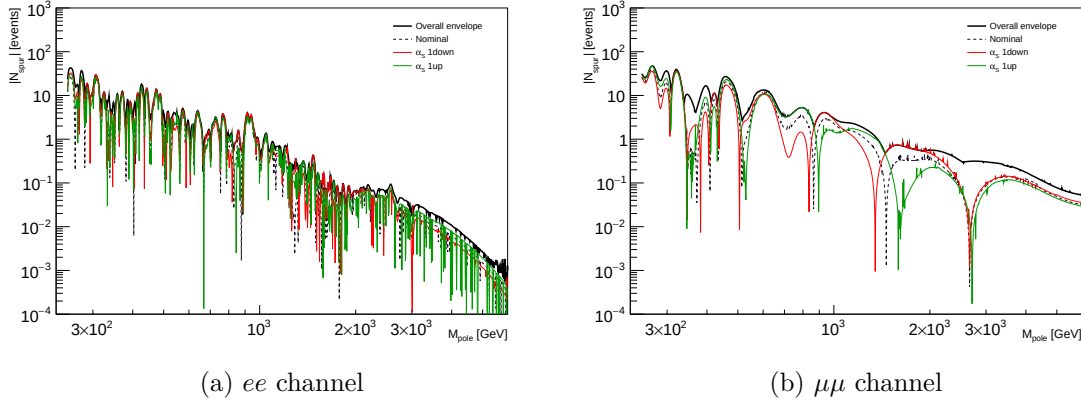


Figure 8.5: Spurious signal absolute yield obtained on the nominal background template compared to “up” (green) and “down” (red) templates of the ALPHAS (α_s) systematic variation. The overall envelope covering all systematic variations is also shown for reference as a solid black line. A signal of width $\Gamma/m_X = 0\%$ is used for this example. Figures taken from [129].

Statistical

The statistical uncertainty of the dataset used to constrain the fit propagates by way of the covariance matrix through to an uncertainty on the function parameters. In a full “background + signal” ($b + s$) model used to extract signal strength, the statistical variations of background parameters are interconnected with uncertainty σ_μ on the best-fit value of the parameter of interest. In this way statistical uncertainty enters directly into the expected signal significance (e.g. sensitivity) via equation 8.13. The impact of data sample size is therefore incorporated automatically through the evaluation of the best-fit model parameters and their correlations in a likelihood minimization where they are allowed to float. The background function parameters and the POI float freely while nuisance parameters vary within their constraints. For the full Run-2 analysis this was implemented in **RooStats** [140] which employs as backend the MINUIT minimization package [141].

8.3 Summary of systematic uncertainties and impact

The treatment of background uncertainties via the spurious signal systematic was discussed in section 8.2.4. Systematic uncertainties introduced through variations of the signal normalization and shape were discussed in sections 6.8.1 and 6.8.2, respectively. To summarize, then, systematic uncertainties are incorporated into the analysis results through the following routes:

Background variation: spurious signal systematic – θ_{ss} in (8.25). Envelope of:

- N_{ss} on the nominal background template
- N_{ss} on the ensemble of bkg. templates varied by theory systematics in table 6.7

Signal variation: experimental systematics

- normalization – κ in (8.25) (enters limits but not p -values)
 - Luminosity (integrated)
 - $A\epsilon$
- shape
 - width – θ_σ in (8.25)
 - shift – θ_μ in (8.25)

Benchmark signal cross section: envelope of PDF (theoretical) systematics applied to Z'_{SSM} cross section for qualitative comparison (figure 8.9a)

The final way in which systematics are treated in the full Run-2 dilepton analysis is to produce the uncertainty band on the predicted Z'_{SSM} cross section. The procedure for applying PDF eigenvector and choice variations in this way is discussed in section 9.2.1. The result is shown in figure 8.9a but not taken into account in the statistical analysis – it merely gives an idea of uncertainty on the obtained Z'_{SSM} mass limit.

The parameterized (relative) variations of the signal normalization described in section 6.8.1 enter as the κ NPs which directly multiply the $A\epsilon$ term in equation 8.25. Systematic variations of the signal shape discussed in section 6.8.2 are more technical to implement. The signal shape is allowed to morph smoothly between the “up” and “down”-varied TF parameterizations through instances of the `FlexibleInterpVar` class in `RooStats`. The minimizer can thereby determine a best-fit signal shape for a given feature in data by varying the θ_μ and θ_σ NPs within the Gaussian constraint.

The impact of systematic uncertainties on the final dilepton cross section limits is relatively small, resulting in a (4–7)% reduction compared to the statistics-only limits. The impact of the dominant sources of systematic uncertainty in each channel are listed in table 8.2. The metric used is the relative change to the signal strength from its expected upper

Systematic	Channel				
	m_X [GeV]	Dielectron [%]		Dimuon [%]	
		300	2000	5000	
					300
					2000
					5000
Common					
Spurious signal		± 12.5 (12.0)	± 4.6 (10.8)	± 0.1 (1.0)	± 11.7 (11.0)
Identification		± 1.6 (1.6)	± 5.6 (5.6)	± 5.6 (5.6)	± 1.8 (1.8)
Isolation		± 0.3 (0.3)	± 1.1 (1.2)	± 1.1 (1.1)	± 0.4 (0.4)
Luminosity		± 1.7 (1.7)	± 1.7 (1.7)	± 1.7 (1.7)	± 1.7 (1.7)
Electrons					
Energy scale		-1.7 $\left(\begin{smallmatrix} +1.0 \\ -1.8 \end{smallmatrix} \right)$	-1.9 $\left(\begin{smallmatrix} +1.7 \\ -2.9 \end{smallmatrix} \right)$	$+0.1$ $\left(\begin{smallmatrix} \pm 0.8 \\ -0.4 \end{smallmatrix} \right)$	-
Energy resolution		$+7.9$ $\left(\begin{smallmatrix} +1.1 \\ -0.9 \end{smallmatrix} \right)$	$+9.0$ $\left(\begin{smallmatrix} +0.7 \\ -0.5 \end{smallmatrix} \right)$	$+0.4$ $\left(\begin{smallmatrix} \pm 0.1 \\ -0.9 \end{smallmatrix} \right)$	-
Muons					
ID resolution		-	-	-	$+0.8$ $\left(\begin{smallmatrix} +0.3 \\ -0.8 \end{smallmatrix} \right)$
MS resolution		-	-	-	$+2.8$ $\left(\begin{smallmatrix} +1.0 \\ -1.3 \end{smallmatrix} \right)$
“Good muon”		-	-	-	± 0.6 (0.6)
					$+0.9$ $\left(\begin{smallmatrix} +0.7 \\ -1.1 \end{smallmatrix} \right)$
					$+3.2$ $\left(\begin{smallmatrix} +2.6 \\ -2.4 \end{smallmatrix} \right)$
					$+9.0$ $\left(\begin{smallmatrix} +9.0 \\ -8.2 \end{smallmatrix} \right)$
					$+55$ $\left(\begin{smallmatrix} +55 \\ -35 \end{smallmatrix} \right)$

Table 8.2: The relative impact of $\pm 1\sigma$ variation of systematic uncertainties on the signal yield in percent for zero (10%) relative width signals at the pole masses of 300 GeV, 2 and 5 TeV for dielectron and dimuon channels. Sources of uncertainties leading to an impact smaller than 0.5% on the signal yield at any point of the mass spectrum are not shown. A signal is injected at the cross-section limit.

limit value at the given mass and width. Systematic uncertainties with impact smaller than $\pm 0.5\%$ are not listed.

At low mass (300 GeV) the spurious signal systematic is the dominant uncertainty in both channels. In the electron channel this gives way first to uncertainty in energy resolution and then identification efficiency. For muons, the rising uncertainty associated with muon reconstruction efficiency (listed as “Identification”) at high- p_T grows to 25% at 6 TeV. However it is surpassed by the efficiency uncertainty of the “good muon” requirement which blows up into the multi-TeV tail, up to $+55\% / -35\%$. Most of the systematics have a similar impact for both the 0% and 10% signal width cases, except for those affecting lepton resolution which have a smaller impact for wider signals.

A flat uncertainty of 1.7% is assigned to the full Run-2 integrated luminosity (i.e. $139.0 \pm 2.4 \text{ fb}^{-1}$) as described in [142]. The LUCID-2 detector [143], installed 17m downstream from ATLAS, provides the primary luminosity measurement and is crosschecked against additional methods.

8.4 Statistical results

The final fits of the function (8.20) to the full Run-2 dilepton invariant mass are shown in figure 8.6. The choice of 100 log-spaced bins in this case is purely for visualization purposes whereas the actual fit is performed on 1 GeV linear binning (see discussion in section 8.2.2). The best-fit values of the background model parameters under the null hypothesis are reported in table 8.3.

parameter	dielectron channel	dimuon channel
a	178000 ± 400	138700 ± 400
b	1.5 ± 1.0	11.8 ± 0.5
c (fixed)	1	1/3
p_0	-12.38 ± 0.09	-7.38 ± 0.12
p_1	-4.295 ± 0.014	-4.132 ± 0.017
p_2	-0.9191 ± 0.0027	-1.0637 ± 0.0029
p_3	-0.0845 ± 0.0005	-0.1022 ± 0.0005

Table 8.3: Final best-fit values and uncertainties of the background function parameters obtained in the global fit to the full Run-2 dataset.

The number of events plotted in figure 8.6 is normalized to the (variable) bin width, which is measured in units of $10 \text{ GeV}^{(3)}$. An additional detail is a transition starting around 1.5 TeV to error bars based on asymmetric Poisson confidence intervals rather than the $\pm\sqrt{N}$

³For instance, the final nonzero bin in (8.6a) – containing the single highest- m_{ee} event – is $\simeq 135 \text{ GeV}$ wide and therefore has a y -value of $0.074 = 1/13.5 \text{ [Events/10 GeV]}$.

treatment. See [144] and [11] §40.4.2.3 for this prescription. The fit residuals plotted in the subpanels show the agreement between the data and the fitted functional form background estimate. Overall, the variations do not suggest any clear trend and seem consistent with a well-modelled background.

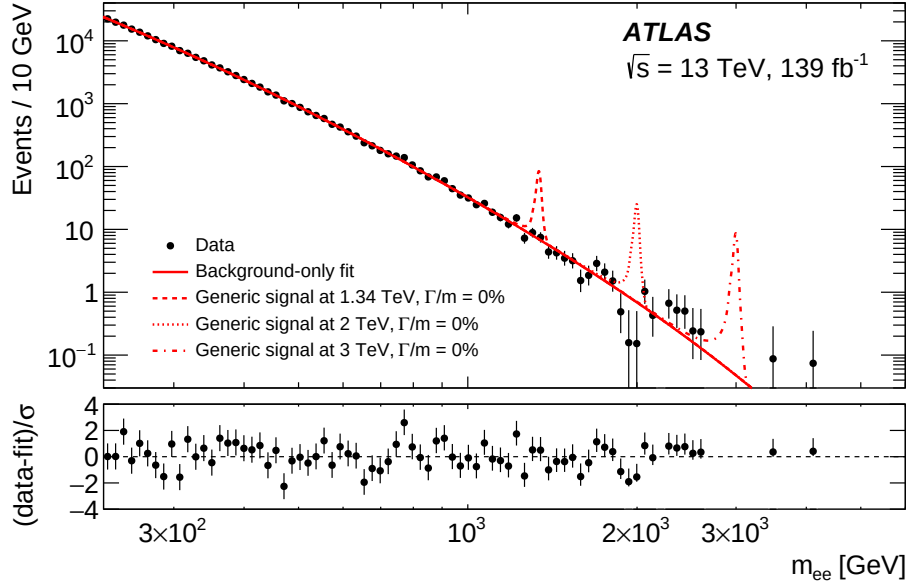
8.4.1 Signal significance scans

The differences between the data and the background estimate shown in the fit residuals of figure 8.6 are quantified by performing $b + s$ fits over a scan of signal hypotheses.

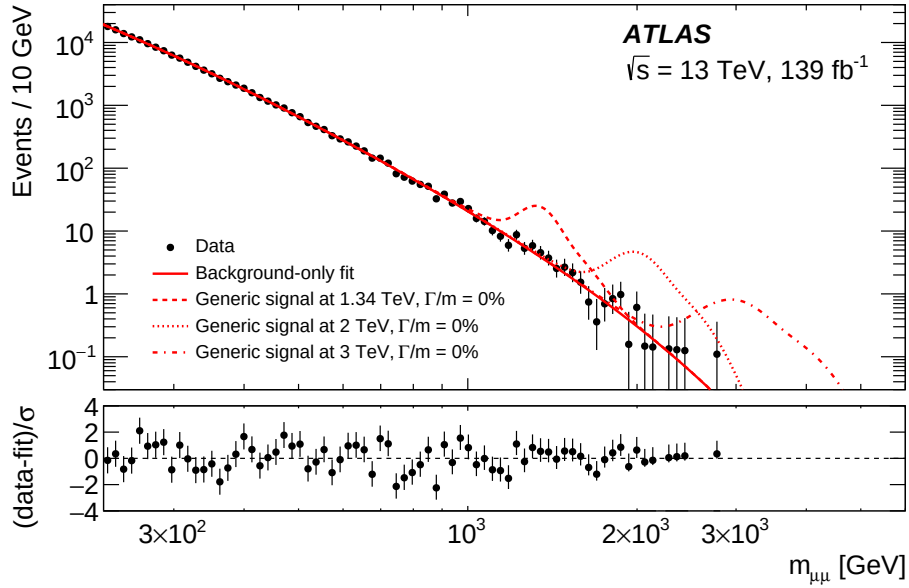
The signal pole mass scan is chosen to coincide with bin centres in the data histogram, except starting only at 250 GeV to allow for a left sideband to the fit. For both channels this scan granularity is finer than the experimental mass resolution. This is simply to avoid the possibility of missing features by scanning in increments that are larger than the resolution. On the other hand, a scan that is finer than the resolution will not be less sensitive to a resonance which happens to lie halfway between two tested pole masses. Similarly, at the resolution threshold the look-elsewhere effect is expected to reach a maximum and not continue to grow beyond that.

The results for signals with zero intrinsic width are shown in figure 8.7. The use of the uncapped LLR test statistic allows for negative fluctuations in data to be quantified. Overall, as noted for the fit residuals, the upward and downward fluctuations appear to be roughly in balance. The electron channel scan clearly exhibits much narrower features than the muon channel owing to its superior mass resolution. The combination of the two channels is therefore driven by the electron channel (this is less so for tests of wider signals). The impact of the two highest- m_{ee} candidates can be clearly seen as narrow peaks at roughly 3500 and 4000 GeV. In the dimuon channel, the highest invariant mass candidate of 2.75 TeV appears as a small bump amidst an otherwise slightly data-deficient region above ~ 2 TeV. The largest excess observed in the combined channel for the 0-width signal is 2.3σ at $m_X = 264$ GeV where the electron and muon channel upward fluctuations overlap.

The pole mass scan was repeated for signals with nonzero intrinsic width up to $\Gamma_X/m_X = 10\%$ in steps of 0.5%. Figure 8.8 shows the “2D” scans of positive local significance across the full set of signal width and mass hypotheses in both channels. In this case the signal hypothesis with largest significance can be seen at $m_X = 773$ GeV and $\Gamma_X/m_X = 2.5\%$. While this feature corresponds to a local significance of 3.0σ it is important to emphasize that the look-elsewhere effect reduces the significance to the level of $\simeq 0.9\sigma$. This caveat is even more important to keep in mind when interpreting the 2D scans since including a scan over signal width introduces an increased look-elsewhere effect. Studies on both the 1D and 2D look-elsewhere effect are documented in appendix B. The signal hypotheses with largest positive and negative local significances are summarized in table 8.4.



(a) Dielectron channel



(b) Dimuon channel

Figure 8.6: The functional form fits to the dielectron and dimuon invariant mass spectra at 139 fb^{-1} . Also shown are example lineshapes of generic resonances with zero intrinsic width in each channel. They are normalized to 20 times the observed 95% CL upper limit on the fiducial cross section at each mass. The subpanel shows the fit residuals. Figures taken from [1].

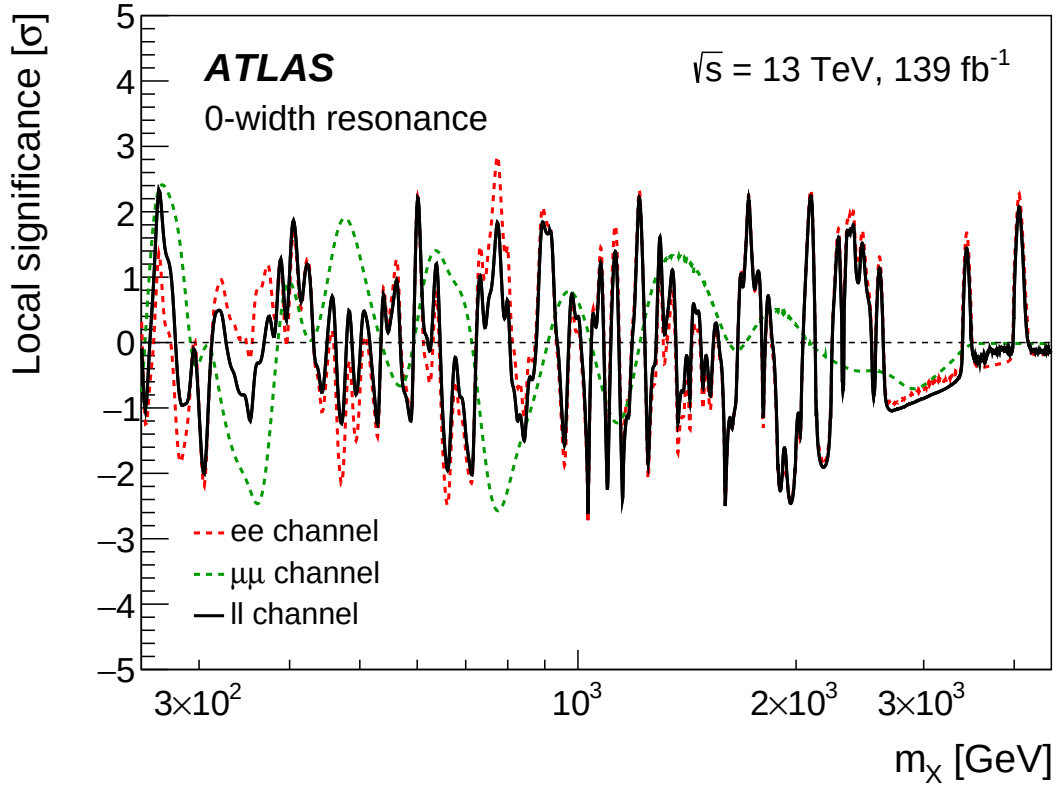


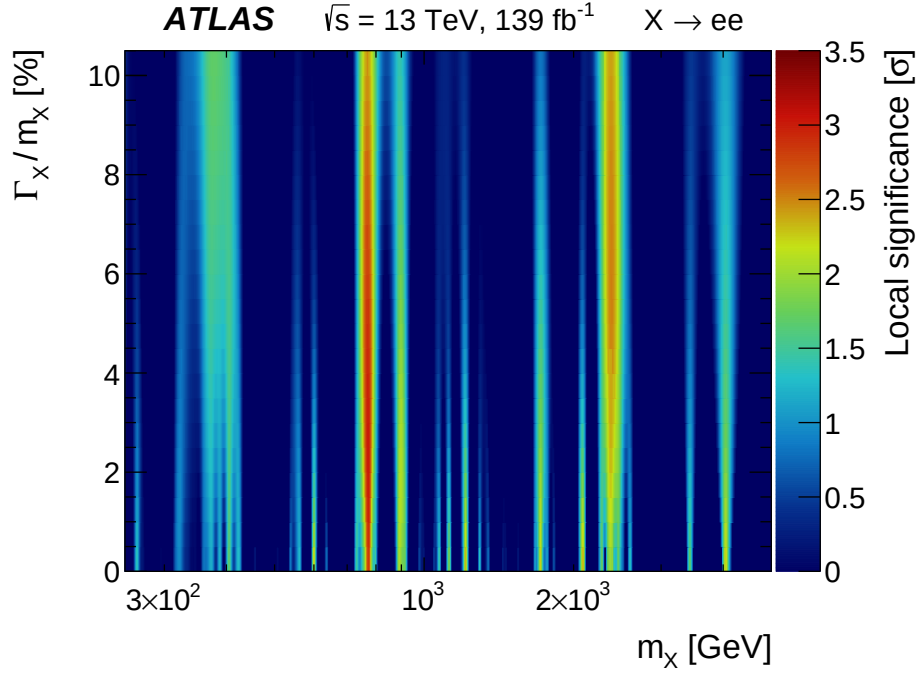
Figure 8.7: Statistical significance of the data as a function of mediator mass m_X , assuming a 0-width resonance. Both dielectron (red) and dimuon (green) channels are overlaid, and their statistical combination is also shown in black. Results are for the full Run-2 dataset and all systematic uncertainties are included. Figure taken from [1].

Channel	Excess			Deficit		
	p_0 (global) [σ]	m_X [GeV]	Γ_X/m_X [%]	p_0 [σ]	m_X [GeV]	Γ_X/m_X [%]
ee	3.0 (0.9 ± 0.2)	773	2.5	-3.2	1957	4.0
$\mu\mu$	2.5 (0.6 ± 0.3)	268	2.5	-2.8	349	8.5
$\ell\ell$	2.3 ($\lesssim 0$)	264	0	-2.9	1958	3.0

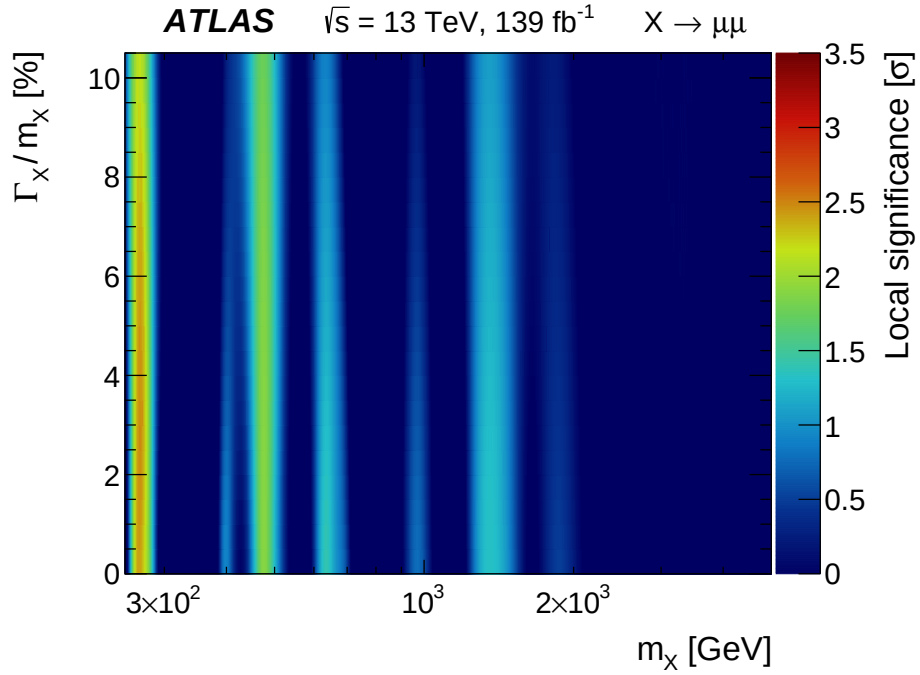
Table 8.4: The largest excesses and deficits observed in the data across the signal width and pole mass scans in the dielectron, dimuon, and combined channels. Global significance values are approximate, based only on upcrossings in the dataset (see appendix B).

8.4.2 Upper limits on signal cross section

Having observed no excess in data meeting the 3σ (global significance) threshold for evidence, the analysis now turns to translate the observations into upper limits on the generic signal cross section. The CLs method is followed as described in section 8.1.2 with an $\alpha = 0.05$ (i.e. 95% confidence level). The hypothesis inversion is carried out for the same set of signal mass points and a subset of six relative widths: $\Gamma_X/m_X \in \{0, 0.5, 1.2, 3, 6, 10\}\%$.



(a) Dielectron channel



(b) Dimuon channel

Figure 8.8: Local significance scan over signal width and pole mass hypotheses. Only positive values are shown. Figures taken from [1].

Both dielectron and dimuon channels are used to produce limits separately and in combination. As discussed in section 6.7, the limits are defined in terms of signal cross sections within the fiducial volume.

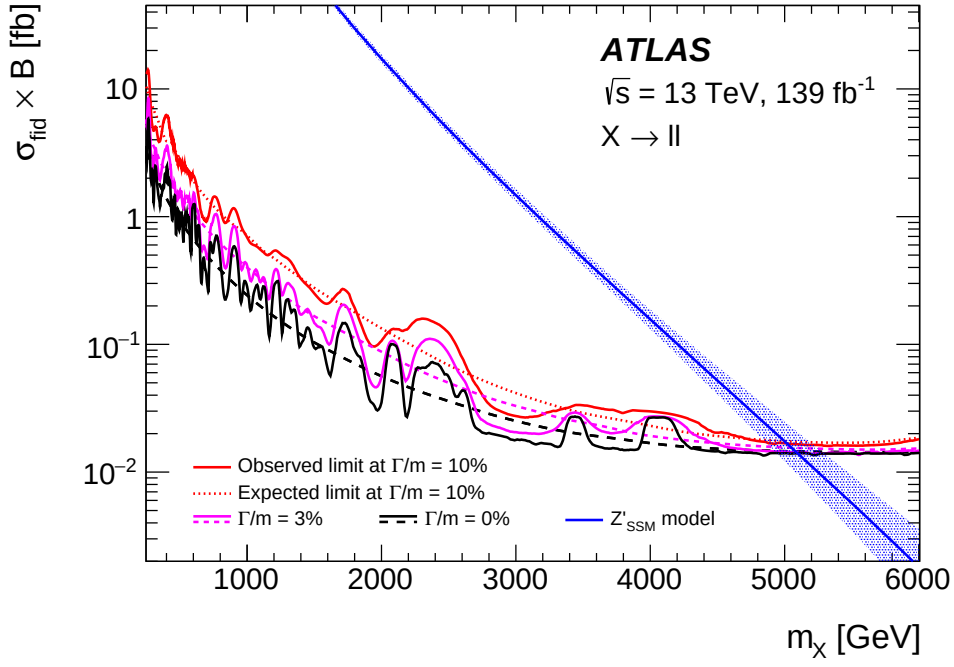
The fiducial cross section upper limits for the combined channel are shown in figure 8.9 in two groups of three curves, each corresponding to a different signal width. The observed limits are obtained by applying the CLs method to data while the expected limits correspond to the smooth background validation template with matching integrated luminosity. Signals with narrow widths are constrained by the results to have lower cross sections than wide signals because of their more conspicuous signature. The corresponding limits in the separate dielectron and dimuon channels are shown in figure B.1 in appendix B.

Because of the shared fiducial cuts applied to both, the overlaid cross sections of benchmark models can be directly compared to the upper limits on the generic signal model. The matching must, however, be done for a common value of assumed resonance width. For example, the Z'_ψ has $\Gamma_X/m_X = 1.2\%$ and therefore can be excluded using the upper limit on the fiducial cross section of the generic signal with 1.2% relative width. This leads to the lower limits on mediator mass for the three benchmark models which are shown in table 8.5.

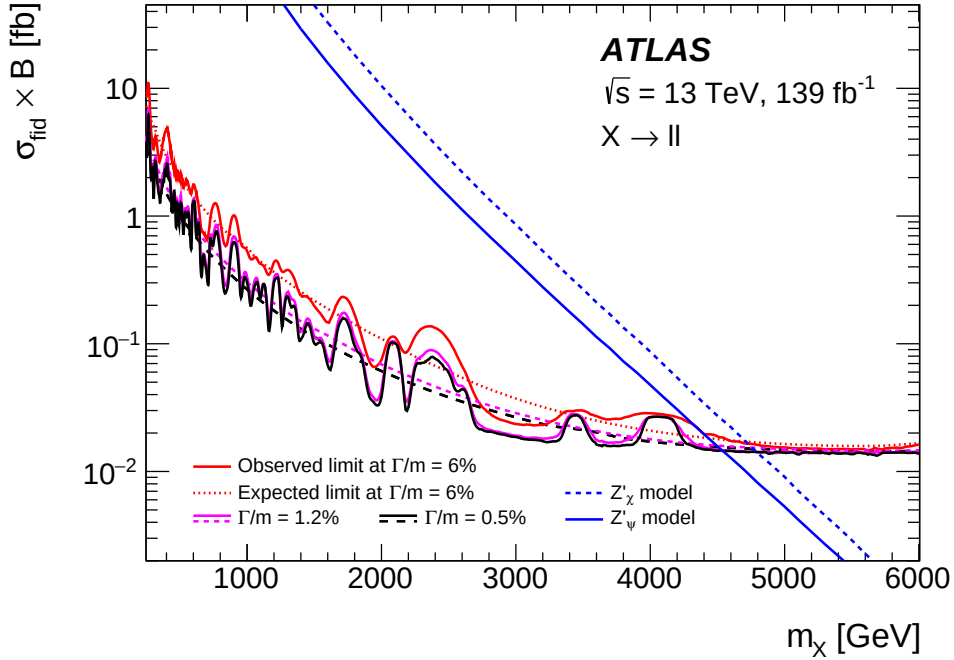
Model	Rel. width Γ_X/m_X	Lower limits on $m_{Z'}$ [TeV]					
		ee		$\mu\mu$		$\ell\ell$	
		obs	exp	obs	exp	obs	exp
Z'_ψ	0.5%	4.1	4.3	4.0	4.0	4.5	4.5
Z'_χ	1.2%	4.6	4.6	4.2	4.2	4.8	4.8
Z'_{SSM}	3.0%	4.9	4.9	4.5	4.5	5.1	5.1

Table 8.5: Lower limits on $m_{Z'}$ obtained from the observed and expected cross section limits at 95% CL in the dielectron, dimuon, and combined channels. Values are listed to the nearest 100 GeV for two E_6 -motivated benchmark models and the SSM model.

Compared to limits set on the full signal cross section, the effect of the fiducial cuts can be seen at the high-mass end of the curves. As the number of expected/observed events drops below 1, the ability of the CLs method to exclude a high-mass peak saturates, revealed by a flattening of the observed limit. On the contrary, for upper limits on signal lineshapes without the fiducial cut (6.14), the curve begins to increase again at this point as the proportion of the cross section contained in the parton luminosity tail grows in comparison to the resonance. This can be seen, for example, by comparing to figure 4 of the 36.1 fb^{-1} dilepton analysis [9]. However, the overall mass limit on the model remains equal in either approach, since in the inclusive case the theoretical cross section also turns upward with respect to its fiducial counterpart.



(a) $\Gamma_X/m_X = 0\%, 3\%, 10\%$



(b) $\Gamma_X/m_X = 0.5\%, 1.2\%, 6\%$

Figure 8.9: Observed and expected upper limits at 95% confidence level on the fiducial cross section of a generic signal in the combined dilepton channel. The curves with different colors correspond to six different resonance width assumptions. Overlaid are predicted cross sections within the fiducial volume for three benchmark models: Z'_χ , Z'_ψ , Z'_{SSM} with relative widths of 0.5%, 1.2%, and 3.0%, respectively. The PDF uncertainty on the Z'_{SSM} cross section curve is indicated by the shaded band. Figures taken from [1].

Chapter 9

Model exclusions by reinterpretation

This chapter discusses the implications of the full Run-2 dilepton cross section limits for three different BSM signal scenarios. In each case, the statistical analysis of section 8.4 is not repeated but rather *reinterpreted* using the appropriate matching of upper limits and predicted cross sections within the fiducial volume. The three signal scenarios considered are as follows:

1. Benchmark Z' models in $E6$ GUT and SSM – section 9.2
2. Heavy vector triplet (HVT) neutral mediator – section 9.3
3. Simplified dark matter (DM) spin-1 mediator – section 9.4

The general scheme for each of these reinterpretations can be summarized in three main steps:

1. Generate/reweight MC signal $m_{\ell\ell}$ distributions
2. Calculate signal resonance width and fiducial cross section
3. Determine bounds on signal parameters by matching to the upper limits on dilepton fiducial cross section

There are some differences in how these steps are carried out for each of the three signal scenarios. The procedure and results are discussed in further detail for each case.

9.1 Signal reweighting

The calculation of cross sections for signals in the benchmark Z' models and the HVT model is achieved through an implementation of the `LPXSignalReweightingTool`, documented in [145, 146]. This tool provides scale factors to be applied at the event level on simulated

Drell-Yan samples to emulate a menu of Z' signal models. The scale factor encodes the ratio of the differential cross-section of the BSM process of interest to that of SM Drell-Yan (in our case, neglecting interference):

$$\text{SF}_{\text{rw}} = \frac{\frac{d\sigma}{d\hat{t}}(q\bar{q} \rightarrow Z' \rightarrow \ell\bar{\ell})}{\frac{d\sigma}{d\hat{t}}(q\bar{q} \rightarrow Z/\gamma \rightarrow \ell\bar{\ell})} \quad (9.1)$$

where the Mandelstam variable $\hat{t}$ is the difference-squared of the initial state quark q and final state lepton ℓ momenta:

$$\hat{t} = |P_{\ell^-}^\mu - P_q^\mu|^2 = |P_{\ell^+}^\mu - P_{\bar{q}}^\mu|^2 \quad (9.2)$$

The BSM differential cross sections are parameterized for a generic form of Lagrangian which admits the benchmark models considered here. The reweighting of Drell-Yan events by means of equation 9.1 is applicable to BSM processes which have an event topology identical to Drell-Yan. Models involving s-channel production of a spin-1 boson therefore fulfill this requirement, whereas models with mediators of different spin do not.

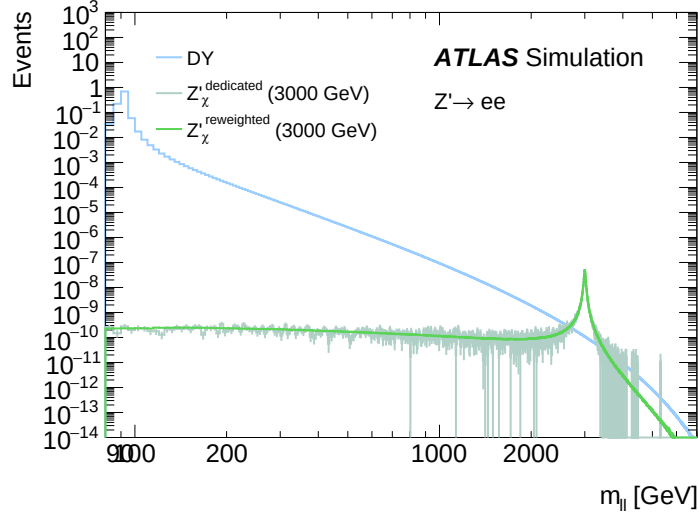


Figure 9.1: The Z'_χ invariant mass distribution at truth level obtained from reweighting dielectron Drell-Yan events, compared to that of a dedicated MC sample and also the original Drell-Yan distribution.

The reweighting approach is used to provide kinematic distributions for the Z'_{SSM} , Z'_χ , Z'_ψ , and Z'_{HVT} signal models in a range of pole masses without the need for dedicated Monte Carlo production. The tool is only configured to deal with Drell-Yan samples at LO, so the list in table A.2 is used for reweighting (the results are based on the available samples in the mc16a and mc16d campaigns). These samples are passed through the full dilepton selection

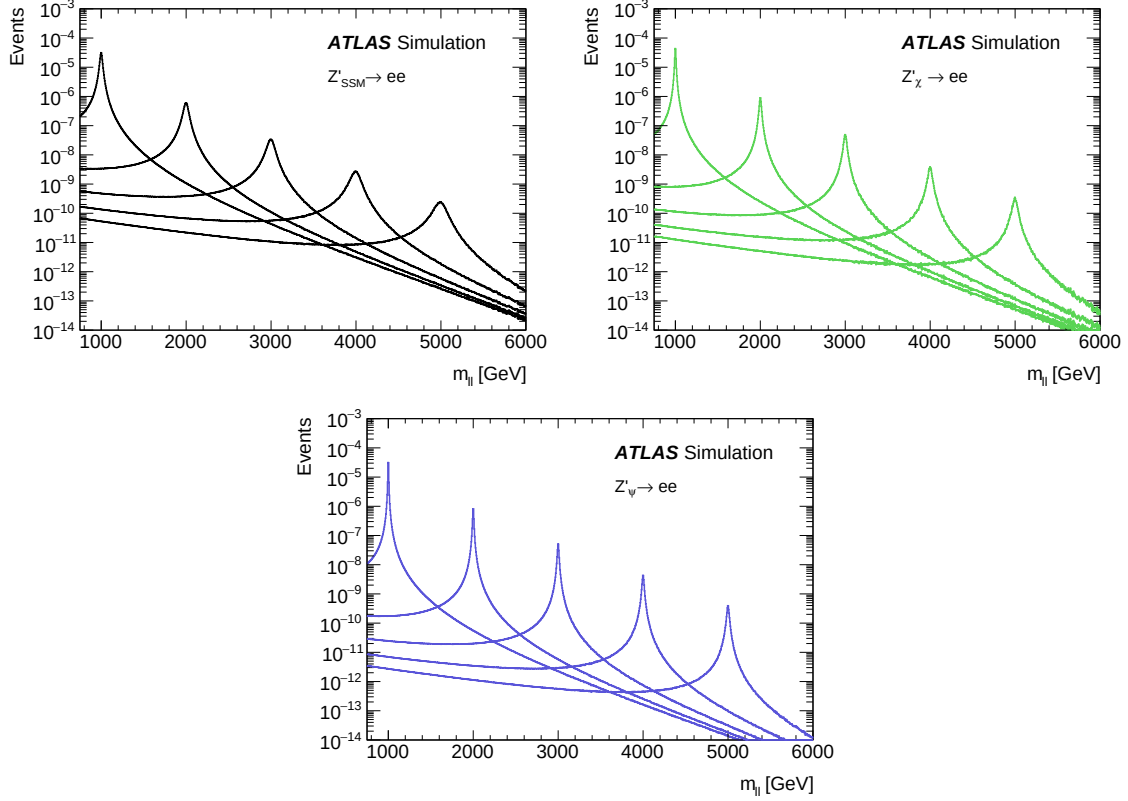


Figure 9.2: Examples of reweighted signal templates at truth-level obtained for the Z'_{SSM} , Z'_{χ} , and Z'_{ψ} benchmark models. The resonances have fixed relative widths of 3.0%, 1.2%, 0.5%, respectively.

framework. They are unskimmed such that truth-level quantities are readily accessible in the final analysis n-tuples.

Figure 9.1 compares the result of the reweighting procedure for the Z'_{χ} model at 3 TeV overlaid with the invariant mass distribution of a dedicated PYTHIA8 MC sample, both at truth-level. Also shown is a binned version of the input DY sample. The agreement between the dedicated and reweighted signal distribution is clearly good across the entire spectrum. The reweighted template benefits from the smoothness of the mass-sliced Drell-Yan input sample.

9.2 Scenario #1: Benchmark Z' models

Examples of dielectron invariant mass distributions obtained for Z'_{SSM} , Z'_{χ} , and Z'_{ψ} by reweighting are shown in figure 9.2. Since these are models with lepton flavour universality, only one channel needs to be chosen to calculate the common fiducial cross section – the results shown use $Z' \rightarrow ee$.

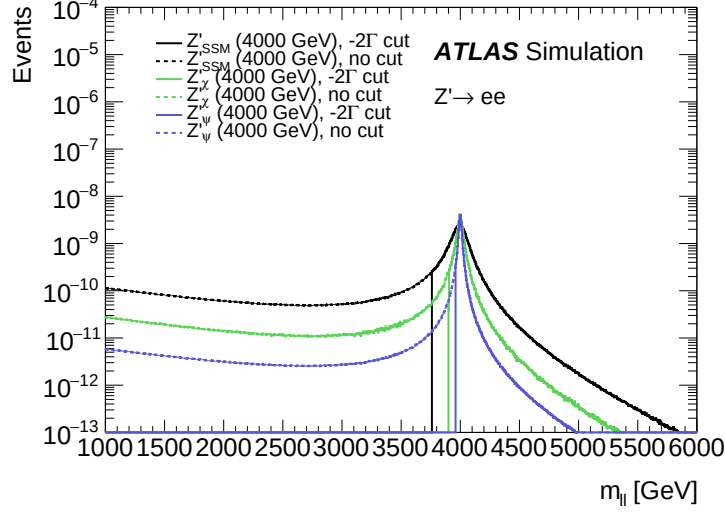


Figure 9.3: Visualisation of the mass cut $m_{\ell\ell} > m_X - 2\Gamma_X$ on the truth-level invariant mass distributions of reweighted Z'_{SSM} , Z'_χ , and Z'_ψ signal distributions with a pole mass of 4TeV.

The integral N of the invariant mass distribution is proportional to the total cross section σ of the new process times its branching ratio B to the dilepton final state:

$$N_{\text{tot}} = \mathcal{L}\sigma \times B \quad (9.3)$$

$$N_{\text{acc}} = \mathcal{L}\sigma A \times B \quad (9.4)$$

$$N_{\text{reco}} = \mathcal{L}\sigma A\epsilon \times B \quad (9.5)$$

where A refers to the acceptance factor introduced by the fiducial cuts defined for the dilepton analysis in equations 6.13 and 6.14. Thus, the fiducial cross section which is suitable to compare directly to the upper limits in figure 8.9 is then calculated from the integral N_{acc} of the intermediate “acceptance-level” distribution:

$$\sigma_{\text{fid}} \times B = \sigma A \times B = N_{\text{acc}}/\mathcal{L} \quad (9.6)$$

The $\mathcal{L}$ factor in this case refers to the effective integrated luminosity of the input Drell-Yan sample.

$$\mathcal{L}_{\text{eff}} = N_{\text{DY,tot}}/\sigma_{\text{DY}} \quad (9.7)$$

In case of mass-sliced samples it is applied as an event weight depending on which sample the event belongs to. The benchmark Z' models have fixed relative widths as listed in table 8.5 which align with chosen signal widths of the cross section upper limit curves. Figure 9.3 shows the impact of the mass cut on the three benchmark models at 4 TeV.

Using instead the integral N_{tot} prior to fiducial cuts yields the full cross section times branching ratio. Since we are focusing on the dilepton final state, the presence of a branching

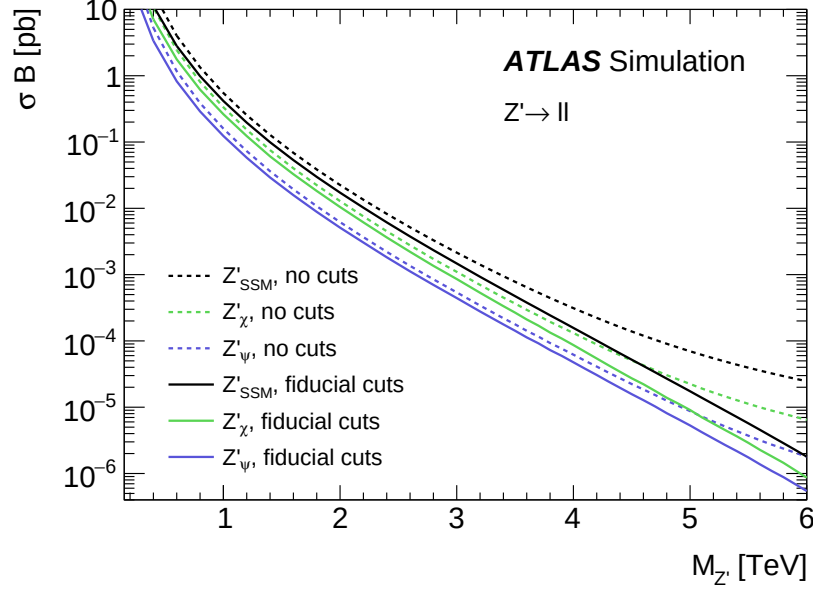


Figure 9.4: Cross sections calculated from reweighted Z'_{SSM} , $Z'_{\chi'}$, and Z'_{ψ} templates both with (solid line) and without (dashed line) the fiducial cuts (6.13)-(6.14) applied.

ratio factor should be inferred where not written. The cross sections as functions of pole mass for the three benchmark Z' models are shown in figure 9.4 both before and after applying fiducial cuts. This illustrates both the exponential fall of the signal cross section with pole mass and how the fiducial cuts have an increasing impact for the largest masses where the parton luminosity tail fraction grows. These are the curves overlaid directly on the generic dilepton upper limits in subfigures 8.9a and 8.9b to produce the benchmark mass limits in table 8.5.

9.2.1 Deriving PDF uncertainty bands on the Z'_{SSM} cross section

The predicted cross section curves obtained from the reweighting are affected by PDF uncertainties, particularly at high mass. For the SSM case, two sources of PDF uncertainty were quantified using the VRAP program [147]. The first is PDF uncertainty caused by imperfect constraint on global fits to experimental data used to extract PDF sets. This has been derived based on the 90% CL error set accompanying the CT14NNLO PDF set following the recommendations of [148]. Normally the variations are provided in 28 eigenvectors which parameterize the independent degrees of freedom within fit uncertainty. For the dilepton and lepton + E_T analyses, these were bundled by the CT14 authors into a new set of seven eigenvectors to avoid overconstraining the PDF at high mass based on data at low mass, because of the very large signal regions of these analyses. The envelope of the eigenvector variations widens a lot in the multi-TeV region because experimental data has a harder time constraining the PDF at high momentum fraction x .

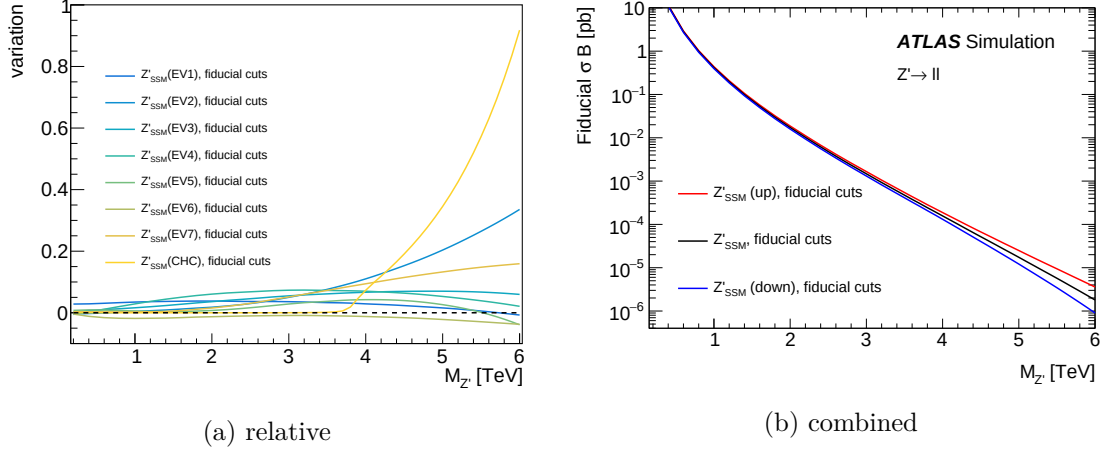


Figure 9.5: Relative impact of PDF eigenvector and reduced choice variations on Z'_{SSM} cross section with respect to nominal (a). Combined impact via equation 9.8 shown in (b).

An additional uncertainty related to the spread in predictions from different PDF sets is also included through the reduced PDF choice variation. It is derived from a comparison of the central values of CT14NNLO [149] with those of MMHT14 [150], NNPDF3.0 [29], ABM12 [151], and JR14 [152]. To avoid double-counting, differences between PDFs are only counted to the degree in which they exceed the sum in quadrature of PDF eigenvector variations (on the CT14NNLO nominal prediction). The resulting envelope of differences coming from each pair of PDFs is then taken as the (reduced) choice variation.

The impact on the SSM cross section from the PDF eigenvector and choice variations is shown in figure 9.5a. The conspicuous rise in the PDF choice variation above 3.8 TeV is driven by the difference between CT14NNLO and NNPDF3.0. The set of PDF eigenvector and choice variations are combined through a sum in quadrature at each pole mass, $\Delta\sigma$, and the overall up (σ^+) and down (σ^-) variations are determined as follows:

$$\sigma^\pm = \sigma \cdot (1 + \Delta\sigma)^{\pm 1} \quad (9.8)$$

Figure 9.5b overlays the overall up and down variations of the Z'_{SSM} fiducial cross section along with the nominal curve. These curves are what define the predicted SSM cross section curve and its uncertainty band in figure 8.9a.

9.3 Scenario #2: Heavy Vector Triplet model

The Heavy Vector Triplet (HVT) framework is a simplified model for parameterizing the results of searches for new high-mass mediators with fermionic and bosonic couplings. The HVT basis enables a joint interpretation of results such as the latest statistical combination of ATLAS heavy resonance searches at 36.1 fb^{-1} [153]. Neutral and charged current searches are both treated as involving a triplet of vector fields $\mathcal{W}_\mu^a$ generated by an $SU(2)'$ extension to the SM:

$$\mathcal{L}_{\mathcal{W}} = -g_q \mathcal{W}_\mu^a \bar{q}_k \gamma^\mu \frac{\sigma_a}{2} q_k - g_l \mathcal{W}_\mu^a \bar{\ell}_k \gamma^\mu \frac{\sigma_a}{2} \ell_k - g_H \left(\mathcal{W}_\mu^a H^\dagger \frac{\sigma_a}{2} i D^\mu H + \text{h.c.} \right) \quad (9.9)$$

The couplings g_q and g_l are assumed to be the same for each fermion generation, k , and implicitly apply only to left-handed fermion fields (q and ℓ). The three gauge bosons associated with the $\mathcal{W}_\mu^a$ field have electrical charges of $-1, 0, +1$ and are nearly mass degenerate. The V' Drell-Yan process shown in figure 9.6a is probed by searches in the $\ell\ell$ and $\ell\nu$ channels. The comparison of constraints from these complementary search channels at 139 fb^{-1} will be shown in section 9.3.2. Figure 9.6b shows the bosonic decay channel probed in searches for diboson resonances in multijet and multilepton channels. Finally figure 9.6c is the vector boson fusion (VBF) V' production mode shown decaying to bosons, but likewise contributes to inclusive selections of the $\ell\ell$ and $\ell\nu$ final states.

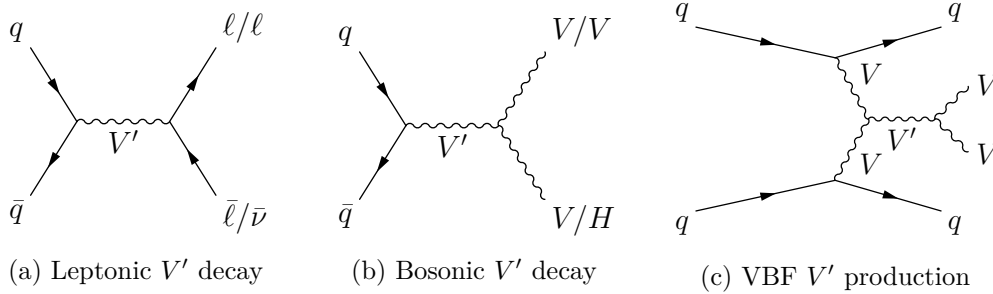


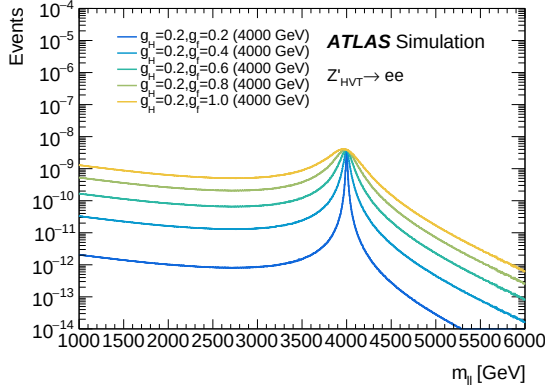
Figure 9.6: Three example diagrams involving the new vector field V' in the HVT model. Taken from figure 1 of [153].

The reweighting procedure is used to obtain a grid of signal templates for the purpose of constraining the HVT model coupling space. Table 9.1 lists the scan configurations used to probe the mass of a Z'_{HVT} boson and its couplings to SM quarks g_q , leptons g_l , and (Higgs and vector) bosons g_h (a.k.a. g_H). Notice $g_f = g_l = f_q$ is used to denote the case of a universal coupling strength to fermions.

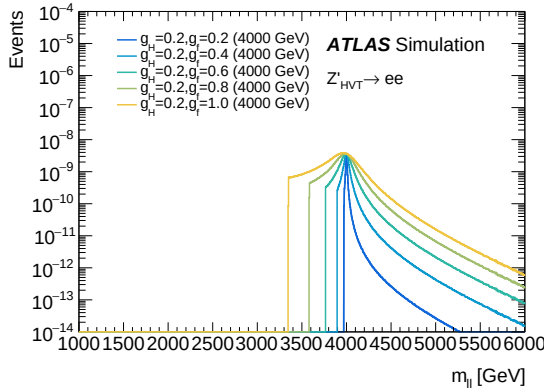
The dilepton generic signal limits are able to constrain the new mediator's coupling to SM bosons mainly through the contribution from the vector boson fusion production mode. The larger g_h , for a fixed g_f , the larger the expected HVT cross section times dilepton branching ratio, but the dependence is smaller than for fermion couplings.

Plane	g_q	g_l	g_h	$M_{Z'} [\text{TeV}]$
$g_l - g_q$ plane	$\{0.00, 0.01, \dots 1.00\}$	$\{0.00, 0.01, \dots 1.00\}$	0	$\{3, 4, 5\}$
$g_f - g_h$ plane	$\{0.00, 0.01, \dots 1.00\}$	$g_l = g_f = g_q$	$\{0.00, 0.03, \dots 3.00\}$	$\{3, 4, 5\}$

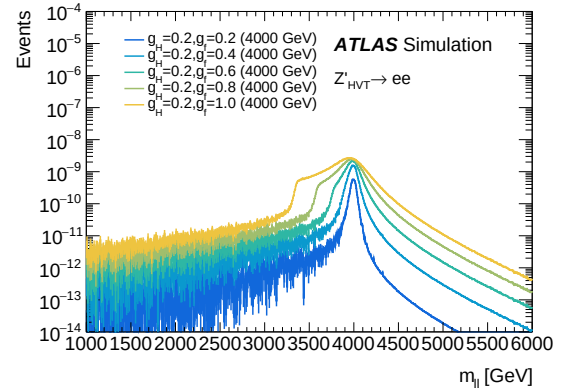
Table 9.1: Configuration of the HVT couplings and pole masses scanned in the two planes considered.



(a) Truth-level



(b) Truth-level, fiducial



(c) Reco-level

Figure 9.7: Examples of reweighted HVT signal templates at truth-level, accepted-level, and reconstructed-level. In this case the Z'_{HVT} pole mass is 4 TeV with couplings of $g_h = 0.2$ and g_f ranging from 0.2 – 1.0.

Some examples of the resulting signal invariant mass templates are shown in figure 9.7. The same fiducial cuts (6.13)-(6.14) are imposed as an intermediate stage between truth- and reco-level.

9.3.1 HVT signal width and cross section values

Figure 9.7 illustrates how both the (relative) width and cross section of the Z'_{HVT} resonance vary for the different choices of couplings. The value of the relative width as a function of the

two scanned couplings in the g_l - g_q and g_f - g_h planes is shown in figure 9.8. For most of this space the relative width remains within the range considered in this analysis. The exception is in the highest coupling values in the g_f - g_h plane, towards the upper right corner of figure 9.8a where the relative width reaches the 10 – 12%, exceeding the width of signals used to set generic upper limits. This corner region is thankfully not near the limit intersection for the pole masses considered.

Figures 9.9b-9.9e similarly show the fiducial cross section obtained by integrating the reweighted HVT invariant mass distributions within the fiducial acceptance criteria at truth level.

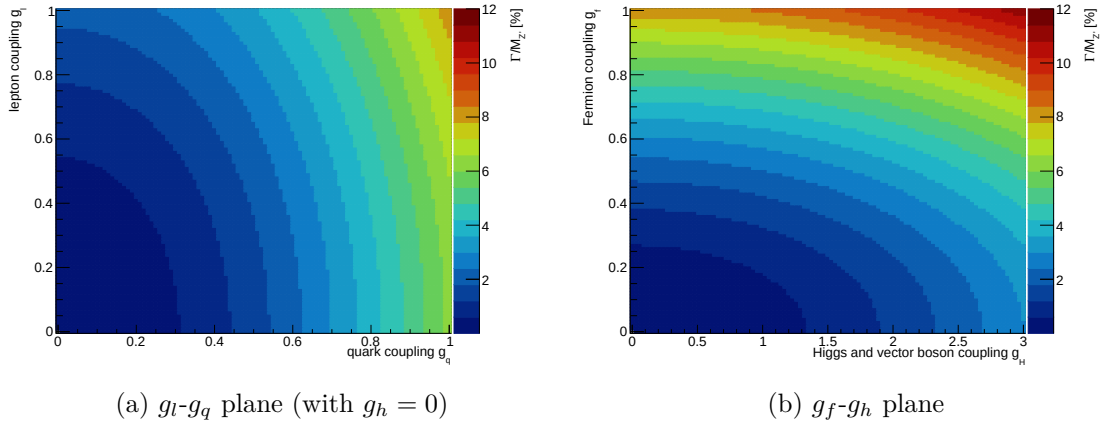


Figure 9.8: The relative width of reweighted Z'_{HVT} dilepton resonances in the two coupling planes considered.

To facilitate an ongoing $VV/VH/\ell\ell/\ell\nu$ combination effort incorporating other HVT decay channels, the cross sections are derived without the application of QCD K -factors. Recall from section 6.2.1 that these account for contributions to the leading-order Drell-Yan cross section from higher-order processes. These K -factors modify the leading-order Drell-Yan cross section by an increase of up to 30% at 5 TeV according to section 4 of [154]. However, higher-order corrections are not readily available for signal samples used in the diboson channels and are therefore omitted to enable comparison.

9.3.2 HVT exclusion contours

Knowing the HVT signal fiducial cross section as a function of the two couplings in each plane enables a direct comparison with the expected and observed limits on generic signal cross sections. As with the benchmark Z' models, the intersection must be done with the upper limit curve for a generic signal of matching width. Although the generic limits are only available for a set of six relative signal widths, these can be smoothly interpolated to provide limits at the widths corresponding to each grid point of the HVT coupling planes. The interpolation relies on Delaunay triangulation [155] (implemented in

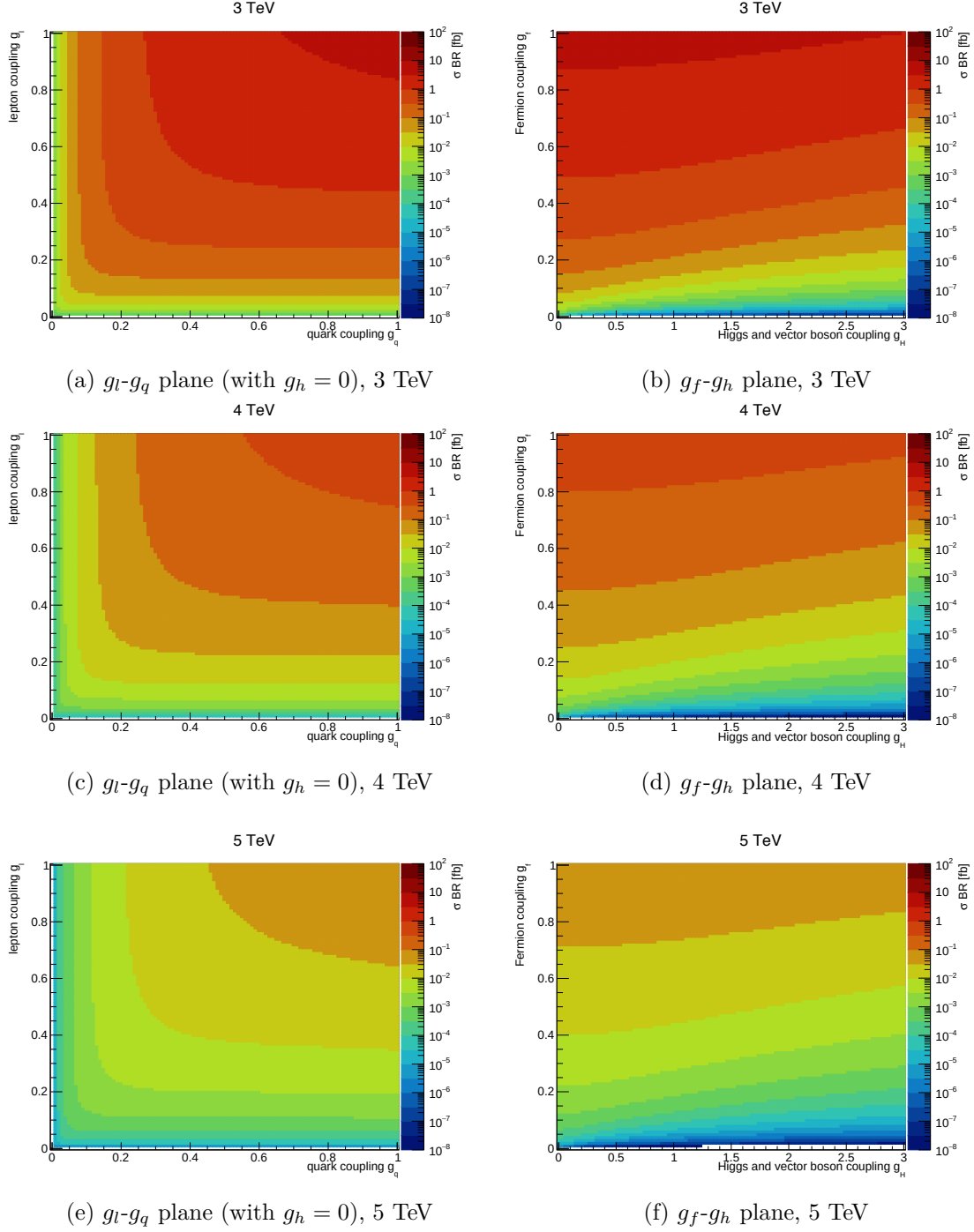


Figure 9.9: The fiducial cross section of reweighted Z'_{HVT} resonances at $M_{Z'} = 3, 4, 5$ TeV in the two coupling planes considered. Values do not include k -factors.

`ROOT::TGraph2D::Interpolate()`) which supports the general case of an arbitrary pole mass and width assumption. Figure 9.10 shows examples of interpolated observed limit curves for the combined $\ell\ell$ channel. The wiggles enabled by the excellent dielectron mass

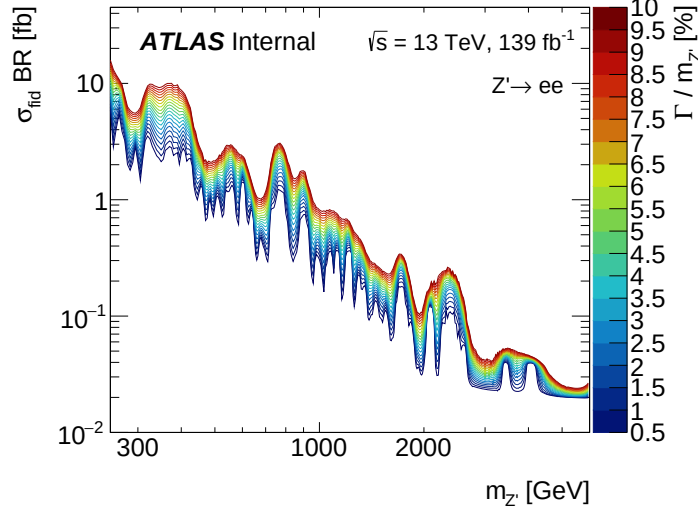


Figure 9.10: Observed upper limits at 95% CL on the cross section of generic resonances decaying to a single-generation lepton pair. The limit values from the original statistical analysis are interpolated to twenty relative width values between 0% and 10%.

resolution are no different from those shown in figure 8.9, but appear somewhat exaggerated due to the aspect ratio of this figure.

Comparing the generic cross section limits to the HVT cross section at a common value of the resonance width is important to match the signal fiducial acceptance which is width-dependent (because of the underlying effect of the width-dependent parton luminosity tail). The width interpolation method is a new development that was not previously used for the HVT contours published in the 36.1 fb^{-1} combination [153]. Figure D.3 in appendix D.1 proves that there is a substantial difference compared to the previous method based on a fixed resonance width throughout the coupling plane.

The generic limits interpolated at each width on the grid of the HVT plane are next divided by the corresponding fiducial HVT cross section values. The contour within the resulting plane where this ratio equals 1 therefore demarks the excluded regions of HVT coupling values. Examples showing the granularity of the contour maps obtained with this method are shown in figure D.1 in appendix D.1.

The HVT couplings to leptons and bosons which are excluded by the ATLAS full Run-2 searches in $\ell\ell$ and $\ell\nu$ final states are shown in figure 9.11. Note that the symmetry of the HVT coupling structure allows the positive quadrants to be simply mirrored as has been done in these plots. The reason for this visualization is to be able to overlay with electroweak precision constraints which are in general sensitive to the sign of the couplings (compare figures 10 and 11 of [153]). The $\ell\nu$ final state generally produces tighter coupling constraints than $\ell\ell$ (the exception being at the “knee” regions of the g_l - g_q contours). Though these observed bounds may be subject to individual fluctuations at high-mass, the same

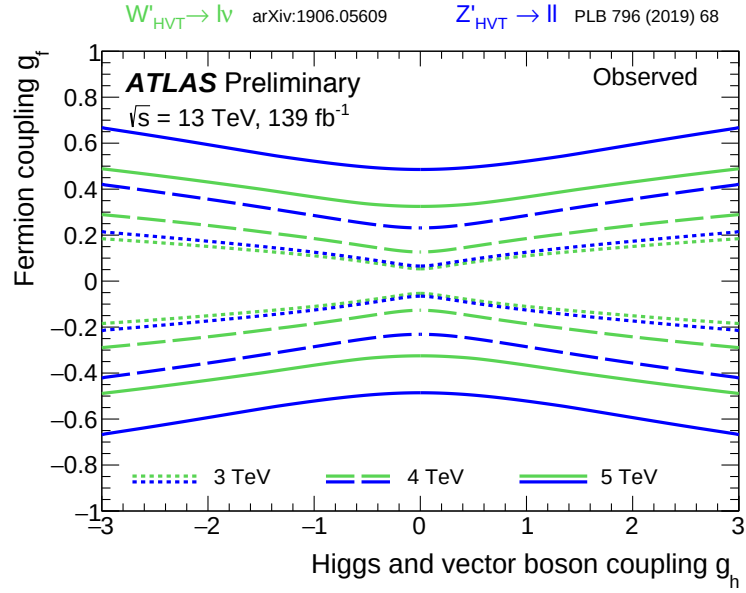
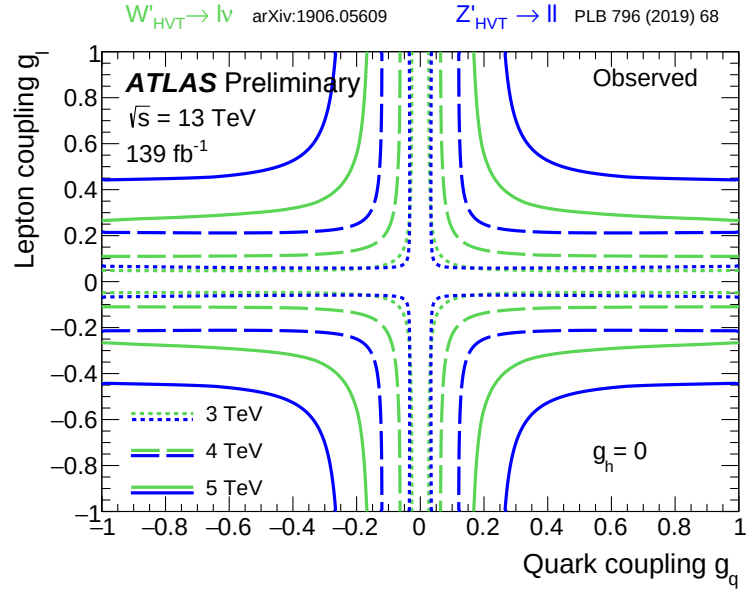


Figure 9.11: Observed 95% CL exclusions on HVT model fermionic and bosonic couplings derived from the ATLAS high-mass dilepton [1] (blue) and lepton + E_T^{miss} [156] (green) searches. The three line styles correspond to three assumptions for the pole mass of the Z'_{HVT} mediator: 3, 4, and 5 TeV. Regions outside the lines are excluded.

behaviour can be verified for the corresponding exclusions based on the expected cross section limits shown in figure D.2 in appendix D.1. The explanation is mainly due to the fact that there are two charged mediator states contributing to the overall $W'^{\pm} \rightarrow \ell\nu$ cross section as opposed to the single neutral mediator for $Z'^0 \rightarrow \ell\ell$.

The observed and expected upper bounds on the fermion coupling $g_f = g_q = g_l$ in the case $g_h = 0$ are presented in table 9.2. Although the full Run-2 combination is still ongoing within ATLAS analysis groups it can already be seen that the updated $\ell\nu$ channel observed bound surpasses the 36.1 fb^{-1} statistical combination of both channels.

int. lumi.	Upper bounds on g_f ($g_h = 0$)				
	139 fb^{-1}		139 fb^{-1}		36.1 fb^{-1}
	$W'_{\text{HVT}} \rightarrow \ell\nu$		$Z'_{\text{HVT}} \rightarrow \ell\ell$		$\ell\ell/\ell\nu$ comb.
pole mass	obs	exp	obs	exp	obs
3 TeV	0.05	0.07	0.07	0.08	0.06
4 TeV	0.13	0.16	0.23	0.18	0.15
5 TeV	0.32	0.38	0.49	0.50	0.42

Table 9.2: 95% CL upper bounds on the HVT triplet field coupling strength to fermions g_f in the case $g_h = 0$ using the full Run-2 ATLAS analyses of the $\ell\ell$ and $\ell\nu$ channels. The final column is a comparison with the leptonic combination result at 36.1 fb^{-1} [153].

9.4 Scenario #3: Dark matter spin-1 mediator

Given non-zero couplings g_ℓ to SM leptons a new neutral mediator associated with a dark sector would produce an excess from $Z'_{\text{DM}} \rightarrow \ell\ell$ in the distribution of the dilepton invariant mass. The upper limits on dilepton signal cross section at 139 fb^{-1} therefore impose lower bounds on the pole mass of a potential Z'_{DM} boson. These bounds are strongest for leptophilic scenarios, with relatively large g_ℓ . The following subsections describe the calculation of $Z'_{\text{DM}} \rightarrow \ell\ell$ cross sections and the resulting exclusion contours within the context of simplified mediator-based DM models.

The effective Lagrangian of the simplified model is as follows [80]:

$$\mathcal{L}_V \supset \frac{1}{2} m_{Z'}^2 V_\mu V^\mu - g_q^{ij} V_\mu \bar{q}_i \gamma^\mu q_j - g_\ell^{ij} V_\mu \bar{\ell}_i \gamma^\mu \ell_j - g_\chi V_\mu \bar{\chi} \gamma^\mu \chi \quad (9.10)$$

where χ is the DM field, and V_μ the vector field associated with the new mediator. A similar expression with the replacement $\gamma^\mu \rightarrow \gamma^\mu \gamma^5$ would describe a mediator with (A)xial-vector couplings to fermion and DM fields. The five main model parameters are the masses of the new mediator ($m_{Z'}$) and of the DM particle (m_χ), and the couplings of the mediator to quarks (g_q^{ij}), leptons (g_ℓ^{ij}), and dark matter (g_χ). Note that the fermion couplings may be flavour-dependent in general (but are assumed to be flavour-universal in the scenario considered). Additionally this set of parameters can be explored for either scenario of a vector or axial-vector mediator.

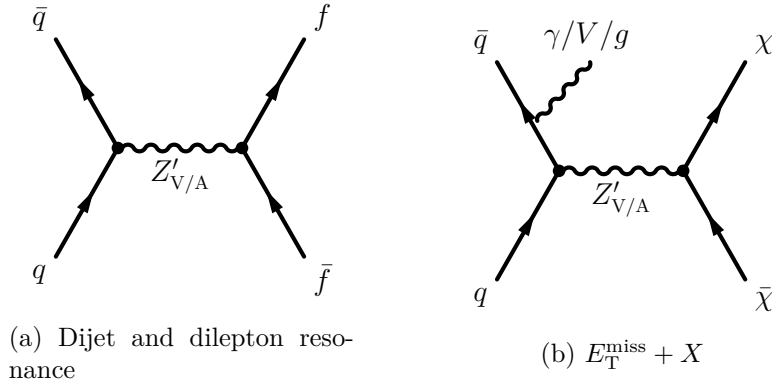


Figure 9.12: Processes probed in s-channel searches for a spin-1 DM mediator. Figures from [76].

The diagrams shown in figure 9.12 are the two processes which are probed by dijet and dilepton resonance searches and $E_T^{\text{miss}} + X$ searches (i.e. mono-jet and mono-photon). A summary of the latest bounds on these processes in the plane of dark matter mass versus mediator mass are shown in section 9.4.2.

As with the HVT case considered in section 9.3, the DM mediator relative width changes in the parameter scan. The same limit interpolation technique illustrated in figure 9.10 is

therefore used to obtain cross section upper limits at matching width. Unlike the HVT case, the DM simplified model is not yet parameterized within the `LPXSignalRewightingTool`, although implementing it would be a good future investment. Instead, a `MADGRAPH` card formatted for $Z'_{\text{DM}} \rightarrow q\bar{q}$ was modified to create samples at generator level. Again, lepton-flavour-universal couplings are assumed for the BSM mediator, so $Z'_{\text{DM}} \rightarrow ee$ events were used without loss of generality.

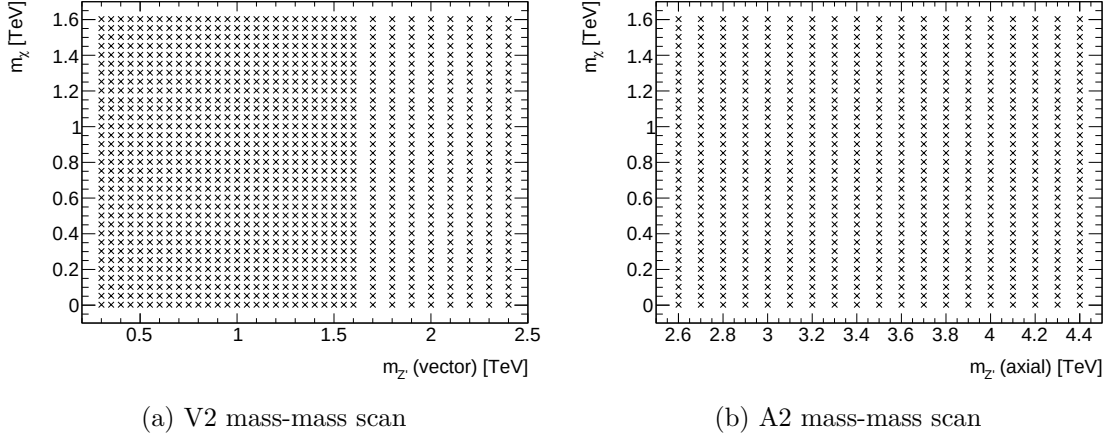


Figure 9.13: Parameter scans of the DM mass m_χ and the DM mediator mass $m_{Z'}$ in the A2 and V2 scenarios. Each point corresponding to a sample of 10k generated events for $Z'_{\text{DM}} \rightarrow \ell\ell$.

For each point in the parameter scan, 10k events are generated. The event generation is done through several batch submission scripts completing the `Generate_tf.py` and `Reco_tf.py` steps to produce `TRUTH0` derivation files. The `MADGRAPH` steering is done through automatic `JobOption` file creation, and the derivations are processed in a simple truth-level event loop macro. The relevant code is contained in the `MadGraph5AtlasVal` project of the `DileptonDM` repository on GitLab [157].

9.4.1 Dark matter mediator “mass-mass” exclusion contours

Parameter scans are defined under the two leptophilic scenarios – “V2” and “A2” – considered in the latest ATLAS DM mediator summary paper [76]. The V2 scenario assumes a spin-1 mediator with vector couplings, in particular a leptonic coupling of $g_\ell = 0.1$. The A2 scenario considers a $10\times$ larger leptonic coupling which is axial-vector in nature. Both scenarios have couplings to quarks and DM particles fixed at $g_q = 0.1$ and $g_\chi = 1.0$, respectively.

Besides the couplings which are fixed in each scenario, the model parameters affecting the Z'_{DM} cross section times dilepton branching ratio are its mass and the DM mass m_χ . These two define the “mass-mass” plane, which is probed in the scan visualized in figure 9.13. The parameter scan was defined to allow coverage around the expected location of the

Scenario	Coupling type	g_q	g_l	g_χ	$m_{Z'} - m_\chi$ scan
V2	vector	0.1	0.01	1	[figure 9.13a]
A2	axial-vector	0.1	0.1	1	[figure 9.13b]

Table 9.3: Configuration of the two parameter scans used for the mass-mass DM mediator exclusions.

new exclusion contours as well as some coverage to crosscheck the location of the contours obtained for the 36.1 fb^{-1} dilepton analysis (which used an entirely different setup).

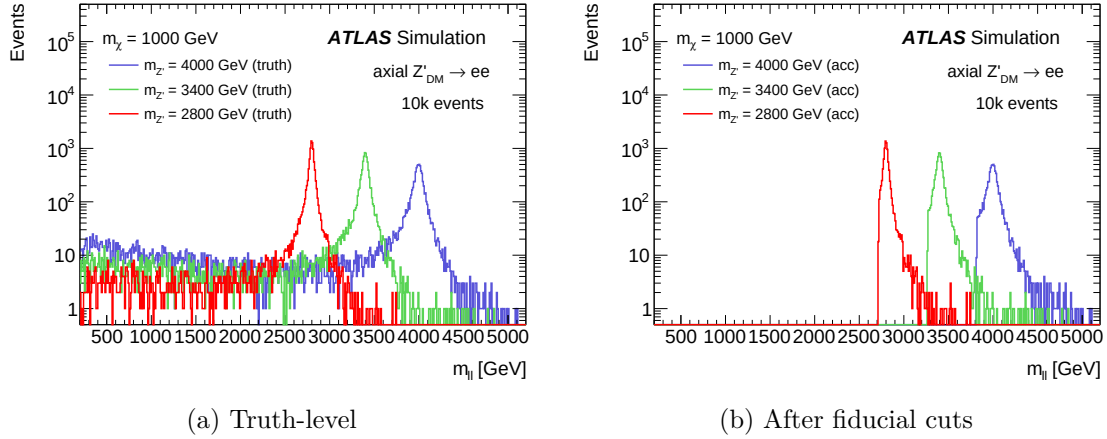
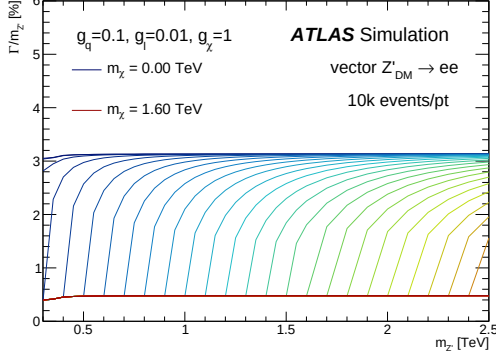


Figure 9.14: Example $m_{\ell\ell}$ distributions for three different pole masses of the generated axial-vector Z'_{DM} signals at truth level and after applying fiducial cuts.

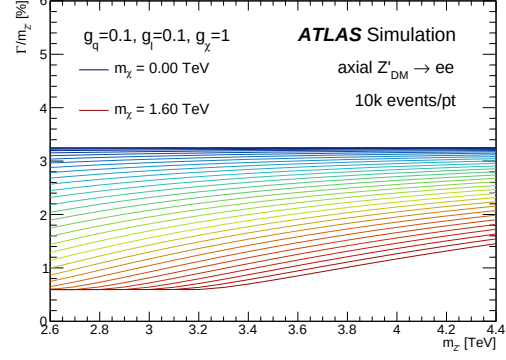
Figure 9.14 shows some example invariant mass distributions of generated $Z'_{\text{DM}} \rightarrow \ell\ell$ events in the A2 scenario before and after applying the width-dependent fiducial cuts of equations 6.13 and 6.14. The corresponding p_T and η distributions for these signal points are also shown in figure D.4 in appendix D.2.

The resonance width $\Gamma_{Z'}$ at each mass point is calculated for each scenario using the `DarkMatterWidthCalculator` utility [158]. Figure 9.15 plots the relative widths $\Gamma/m_{Z'}$ in percent for both scenarios as functions of $m_{Z'}$ for each m_χ . The relative widths are bounded between 0.6% and just over 3.2% for all mass configurations. This behaviour comes about because of two threshold effects. First, the width of the dilepton resonance increases for smaller DM masses since the competing decay $Z' \rightarrow \chi\chi$ becomes more kinematically favourable, but this effect saturates once $m_\chi = 0$ – the upper bound on width. On the other hand, the lower bound on width arises from the opposite effect where the BR to $\chi\chi$ gets smaller as the χ get heavier, but hits the kinematic decay threshold at $2m_\chi \geq m_{Z'}$. Notice how for, say, $m_{Z'} = 4 \text{ TeV}$ this lower limit is not reached.

The acceptance, defined as the ratio of events passing the fiducial cuts to the total number of truth events, is plotted as a function of $m_{Z'}$ for each value of m_χ in figure 9.16.

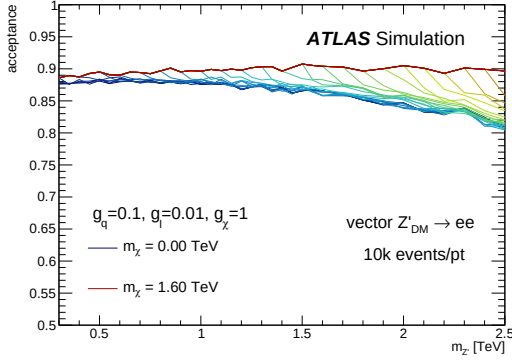


(a) V2 scenario

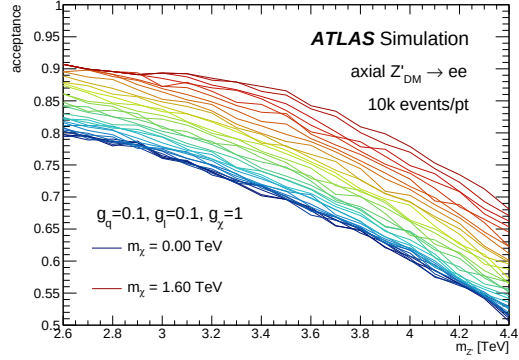


(b) A2 scenario

Figure 9.15: Relative width $\Gamma/m_{Z'}$ in percent of the generated $Z'_{\text{DM}} \rightarrow \ell\ell$ signals for the V2 and A2 scenarios. Each curve is color-coded according to its value of the DM mass m_χ , from blue at small masses to red at large.



(a) V2 scenario

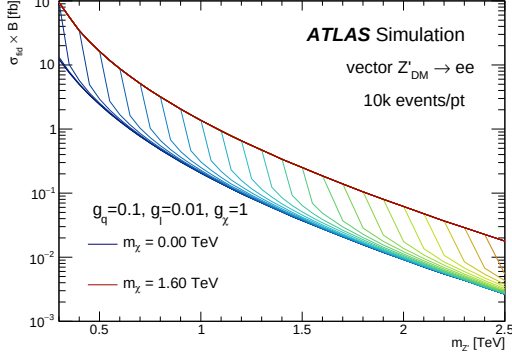


(b) A2 scenario

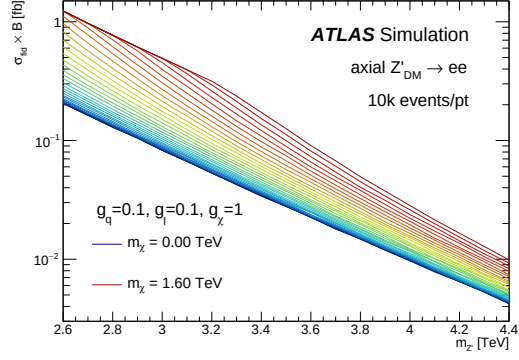
Figure 9.16: Fiducial acceptance of the generated $Z'_{\text{DM}} \rightarrow \ell\ell$ signals for the V2 and A2 scenarios. Each curve is color-coded according to its value of the DM mass m_χ , from blue at small masses to red at large.

The result of multiplying the acceptance with the truth-level cross section is the fiducial cross section, shown in figure 9.17.

Finally, having both the upper limit cross section values from [1] (figure 9.10) and the predicted $Z'_{\text{DM}} \rightarrow \ell\ell$ cross section values (figure 9.17), the location in $m_{Z'}$ of their intersection is found at the matching width for each value of m_χ . The relevant code is contained in the `DileptonDMContours` project of the `DileptonDM` repository on GitLab [157]. The resulting contours in the $(m_{Z'}, m_\chi)$ plane for both vector and axial-vector scenarios are shown in figure 9.18. The exclusion at low mediator mass is taken from the 36.1 fb^{-1} analysis since it set limits down to pole masses of 150 GeV while the full Run-2 analysis started scanning above 250 GeV.

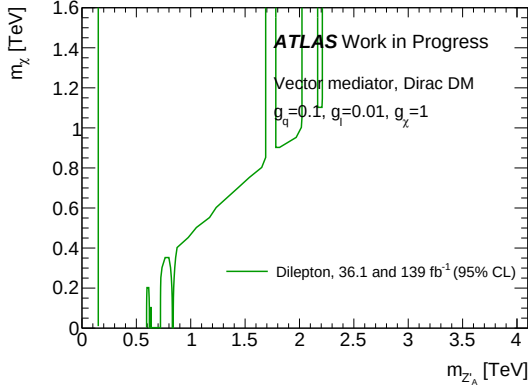


(a) V2 scenario

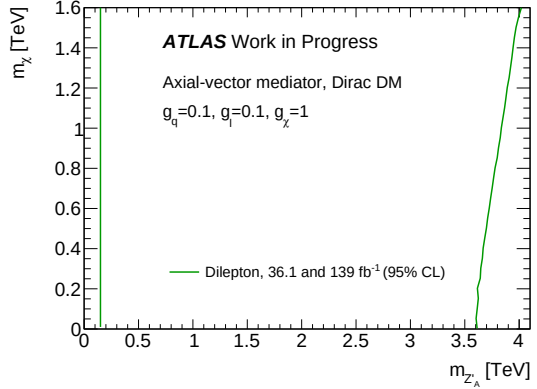


(b) A2 scenario

Figure 9.17: Cross section of the generated $Z'_{\text{DM}} \rightarrow \ell\ell$ signals in the fiducial volume for the V2 and A2 scenarios. Each curve is color-coded according to its value of the DM mass m_χ , from blue at small masses to red at large.



(a) V2 scenario



(b) A2 scenario

Figure 9.18: Dark matter mediator exclusion contours in the mass-mass plane from dilepton cross section upper limits at 95% CL for the V2 and A2 scenarios. The region between the lower- $m_{Z'}$ (36.1 fb^{-1}) and upper- $m_{Z'}$ (139 fb^{-1}) curves is excluded.

A crosscheck of results with that of the 36.1 fb^{-1} contours is provided in appendix D.2, figure D.5. The two curves agree very well despite having been produced by completely independent frameworks.

9.4.2 Dark matter s-channel summary plots

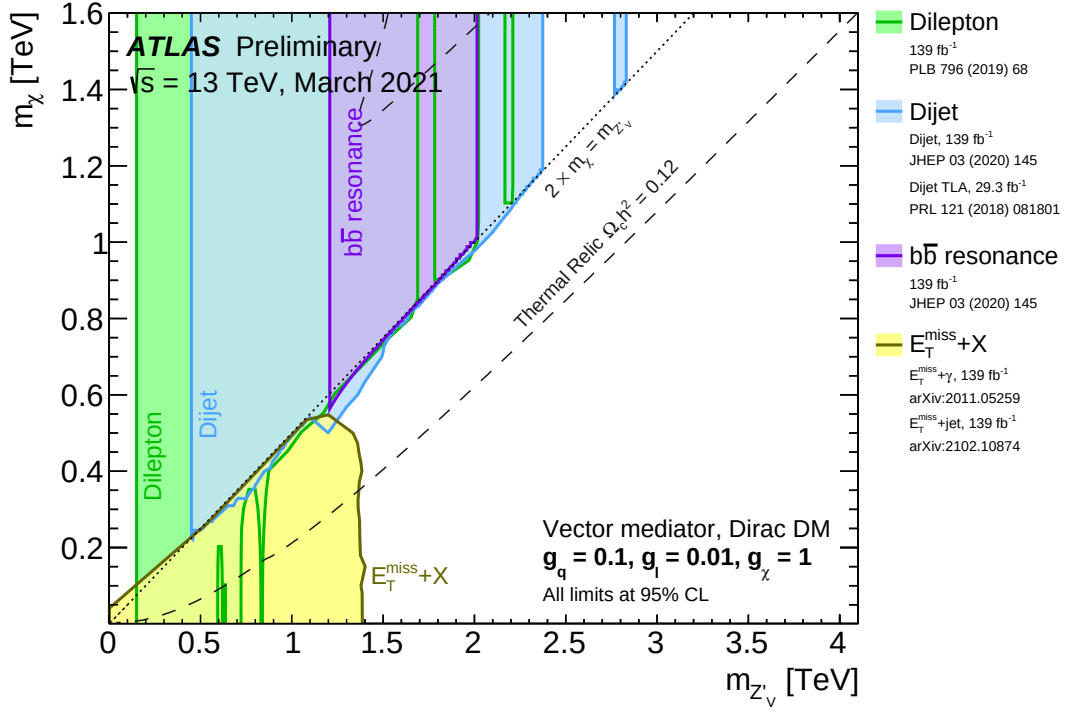
A summary of the latest ATLAS DM mediator exclusions in the mass-mass plane is presented in figure 9.19. The two subfigures correspond to the model scenarios of table 9.3. Clearly the A2 scenario with its ten-times larger leptonic coupling comes under strong exclusion from the full Run-2 dilepton cross section limits. The lower mass limit in this case begins at about 3.6 TeV in the limit of light dark matter and extends out to just beyond 4.0

TeV for $m_\chi = 1.6$ TeV. It gets stronger because of the resonance narrowing effect already discussed. Compared to the analogous contour from the 36.1 fb^{-1} dilepton analysis (figure 12 of [76]) the A2 scenario mass limit has extended by an amount between $400 - 800$ GeV.

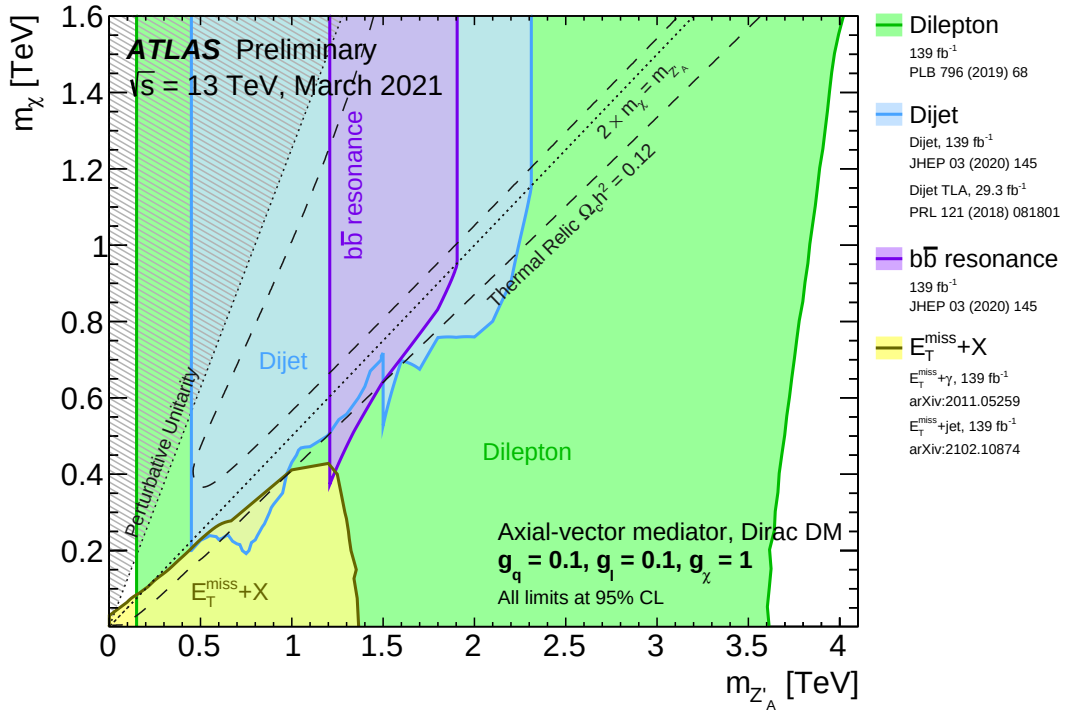
With the lepton coupling in the V2 (vector) scenario dialed back by factor 10, the exclusions from other analyses shown in figure 9.19a are more competitive compared to the dilepton analysis. The di(b)jet signal cross sections are also slightly different in this plane because of the change of coupling structure. The colored regions from the four $Z'_V \rightarrow f\bar{f}$ searches (diagram 9.12a) are therefore complementary up until the kinematic threshold $2 \times m_\chi = m_{Z'_V}$ indicated by a dashed line. As shown in figure 9.17a, this can be understood as coming from the abrupt drop in predicted Z'_V cross section times branching ratio to fermions once the invisible decay opens up kinematically. For the inverse reason, the region excluded by the $Z'_V \rightarrow \chi\chi$ (diagram 9.12b) is bounded above by the same kinematic threshold.

For both scenarios an additional dashed contour indicates which masses in the plane are consistent with a thermal relic density of $\Omega_c h^2 = 0.12$ as suggested by modern cosmological measurements (see equation 3.8 in section 3.3). The curve is obtained by calculating the values of $\Omega_c h^2$ across the plane using the MADDMplugin [159] as demonstrated in figure D.6.

As discussed in section 3.3, direct searches for DM at the LHC are complementary to indirect detection experiments. To compare the ATLAS results in figure 9.19 with latest bounds from other experiments, they are translated into bounds on DM-nucleon scattering cross section under the assumed mediator model. The overlay of the A2 scenario with direct detection bounds from LUX and XENON1T is shown in figure 9.20.



(a) V2 scenario



(b) A2 scenario

Figure 9.19: Exclusion summary overlaying results of six separate ATLAS searches interpreted in the simplified DM mediator model. Shaded areas are excluded. Published in [160].

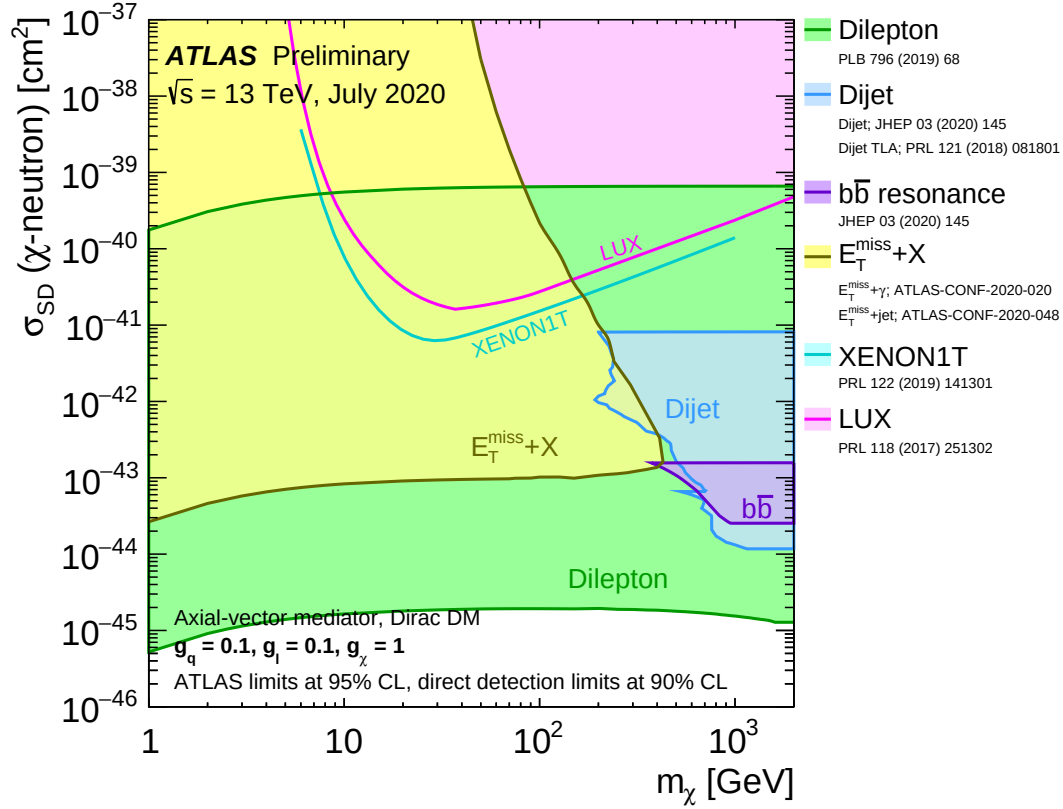


Figure 9.20: Excluded DM-nucleon scattering cross sections inferred using the limits obtained by six different ATLAS searches, assuming a DM axial-vector mediator in the A2 scenario. Overlaid LUX and XENON1T bounds give example comparisons with current spin-dependent (SD) bounds from direct detection experiments. Shaded areas are excluded. Note that the ATLAS limits are at 95% CL whereas the available direct-detection bounds are at 90% CL. Published in [160] (Figure 12).

9.4.3 Dark matter mediator “coupling-mass” exclusion contours

Alongside the previously-released mass-mass exclusion contours, a preliminary version of lepton coupling versus mediator mass exclusions has been prepared for the next iteration of public plots. The result is shown in figure 9.21, where the area above the curves is excluded. This type of plot is analogous to the quark coupling versus mediator mass exclusions usually summarized for analyses with hadronic final states (Figure 15 of [160]). Additional signal mass points were generated for lepton couplings g_ℓ varying between the two fixed values, 0.1 and 0.01, assumed for the axial-vector and vector mass-mass contours respectively, and further down to $g_\ell = 0.001$. These lower coupling values enable bounds on scenarios with lighter DM mediators than are included in the current choice of mass-mass contours. Signal points were generated in the limit of heavy DM ($m_\chi = 10$ TeV in this case) so that contributions to the resonance width from $Z'_A \rightarrow \chi\chi$ are negligible. At each coupling value 10k events were generated for each mediator pole mass at intervals of 50 GeV or 100 GeV as shown in figure D.7. Currently the scan has only been produced for the axial-vector scenario. The corresponding exclusion curves are obtained at 95% CL using the same procedure as for their mass-mass counterparts and .

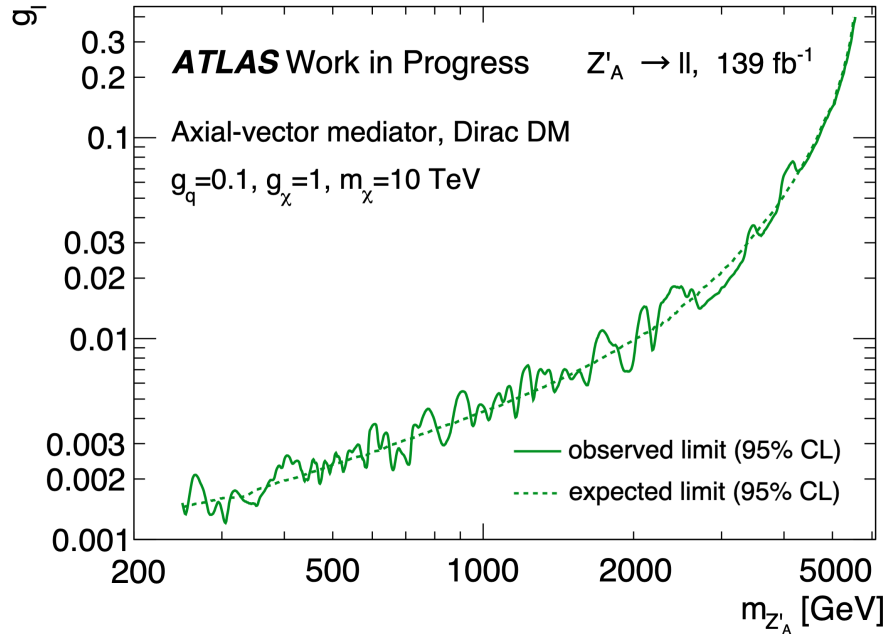


Figure 9.21: Observed and expected upper limits at 95% CL on the lepton coupling of an axial-vector dark matter mediator, as a function of mediator mass. A DM mass of $m_\chi = 10$ TeV is assumed.

Chapter 10

Summary and outlook

This thesis began by reviewing the theoretical and experimental context of the search for high mass dilepton resonances using the ATLAS detector. The full dataset collected in LHC Run 2 from 2015 and 2018 has been filtered through a set of selection criteria to extract the dilepton invariant mass spectrum above 130 GeV. Simulated samples of the main SM backgrounds have likewise been processed to inform the analysis strategy and to model detector $m_{\ell\ell}$ response and acceptance times efficiency. The statistical analysis to search for potential signal resonances is based on a new functional form background estimate fitted to the $m_{\ell\ell}$ spectrum in data.

Since no significant excess was observed, upper limits were derived for a generic signal cross section (figure 8.9). These bounds can be translated to rule out parameter space in a range of BSM scenarios, helping determine which model configurations are still viable. As reviewed in chapter 3, the signature of a dilepton resonance at high mass features in multiple hypothetical extensions beyond the SM addressing a range of its shortcomings. For this reason the search presented in this thesis has had a broad impact in the high energy physics community from publication until now. The exclusions obtained in chapter 9 are examples of this in the context of the HVT and simplified DM mediator models.

To promote similar reinterpretations outside of ATLAS, the data spectra and cross section limits were formatted in a public HEPData repository¹ along with other information. As a result, a number of reinterpretations have appeared, “backpropagating” the ATLAS observed results into updated constraints for a variety of models. The following citations, for example, are representative of the three classes of models motivating a dilepton resonance outlined in figure 3.6:

- **GUT** [161]: E_6 -motivated $U(1)'$ extension in Minimal Supersymmetric SM
- **Alternative Higgs** [162]: Higgs as a composite pseudo-Nambu-Goldstone boson under new strong dynamics

¹<https://www.hepdata.net/record/88425>

- **Dark matter** [163]: simplified DM model with spin-1 mediator
- **Flavour anomalies** [164]: Z' with flavour-non-universal couplings

The fourth example appeared just over a month after the dilepton preprint, recasting ATLAS results to probe Z' explanations of significantly discrepant $\bar{b} \rightarrow \bar{s}\ell^+\ell^-$ transition rates reported by the LHCb Collaboration².

Future versions of this analysis could incorporate additional improvements that were beyond the scope of this one. Mentioned in chapter 6 are the potential to (1) add a separate analysis category for electrons in the calorimeter transition region and (2) include muons which pass the 2-station extension to the high- p_T working point. The studies done in this version of the analysis will likely aid future efforts to “safely” benefit from these two additional categories of leptons. The selection framework developed for this analysis has also paved the way for exclusive dilepton selections such as VBF³ dilepton production, $\ell\ell + E_T^{\text{miss}}$, and $\ell\ell + b$ jets, some of which are already underway in ATLAS.

Looking further ahead, the merits of the data-driven methods developed for this round of the dilepton analysis are set to become more pertinent for datasets of the high-luminosity LHC (HL-LHC). A timeline of experimental progress in the dilepton analysis is shown in figure 10.1 in terms of the benchmark Z'_{SSM} cross section. While the jump in centre of mass energy up to $\sqrt{s} = 13$ TeV led to a clear boost in sensitivity for the Run-2 searches, an increase of this magnitude is not planned for future LHC datasets. Instead, the reach will extend mainly through accumulation of integrated luminosity, especially after Phase-2 upgrades. This is an effective way towards progress on a wide range of physics goals. However, for the Z'_{SSM} prospects in particular, the progress from increased dataset size will be more gradual than in the past. Notice how the collection of 139 fb^{-1} in $\sim$ four years of Run 2 extended the SSM mass limit by 2.2 TeV while the addition of 3000 fb^{-1} over a decade of HL-LHC is foreseen to extend it merely another ~ 1.4 TeV (keep in mind the prospects [167] study is conservative).

Moreover, the high-luminosity gains also come with challenges at both reconstruction and analysis levels. Already in Run 2 the challenge of suboptimal MC statistical precision was encountered in the 36.1 fb^{-1} dilepton analysis. Strategies to fully exploit the upcoming ATLAS datasets will therefore likely include data-driven background estimates, perhaps in sliding window fits or by more exotic means such as Gaussian processes [168]. Novel analysis techniques based on weakly-supervised deep learning algorithms may also find application in dilepton exclusive selections as they have recently in hadronic final states [169]. On the other hand, the rise of faster calorimeter models (e.g. based on generative

²Neutral B meson decays [165], and charged [166] (updated this year)

³Vector boson fusion

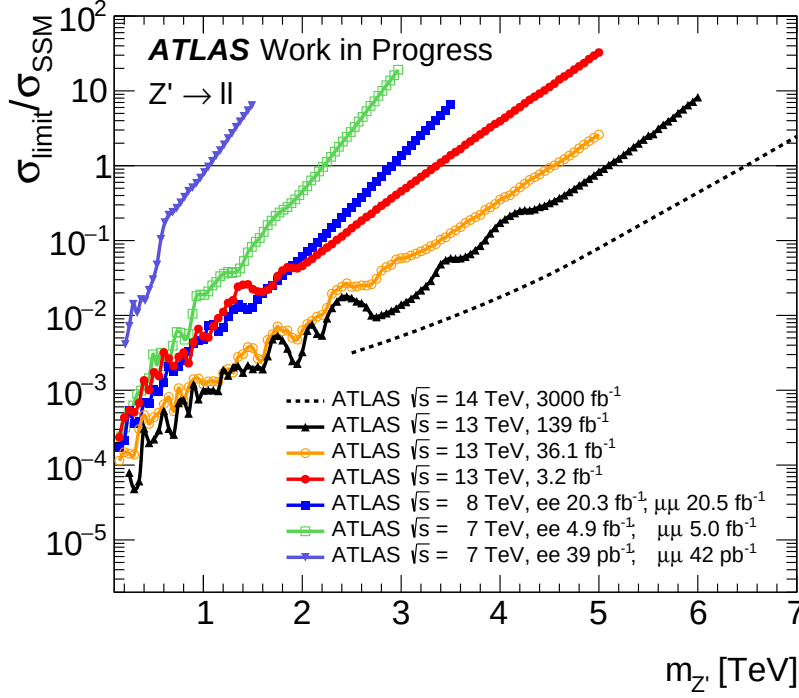


Figure 10.1: Historical progression of the SSM Z' exclusion from published ATLAS dilepton resonance searches using different datasets, including a projection for the full HL-LHC dataset of 3000 fb^{-1} at $\sqrt{s} = 14 \text{ TeV}$ (from [167]). Pole masses are excluded where the ratio of upper limit to predicted cross section crosses is below unity. For the full Run-2 analysis shown in solid black the pole mass lower limit is 5.1 TeV. Figure adapted from [1].

adversarial networks [170]) may dramatically reduce the computation needed to simulate the SM dilepton backgrounds at HL-LHC.

In either case, there are promising improvements in store for the dilepton analysis on top of expected gains through continued excellent operation of the LHC and the ATLAS detector. Given its multiple physics motivations, the signature of a dilepton resonance will remain an important target for future searches at the LHC. Whether such a resonance⁴ hides beyond those of historic discoveries in the dilepton mass spectrum can only be known through continued exploration – as in this thesis – for signs of grand unifications, extra dimensions, alternative Higgs models, dark sectors, or even something entirely unexpected.

⁴besides SM $H \rightarrow \ell\ell$!

Bibliography

- [1] ATLAS Collaboration. “Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector”. In: *Phys. Lett. B* 796 (2019), p. 68. DOI: 10.1016/j.physletb.2019.07.016. arXiv: 1903.06248 [hep-ex].
- [2] ATLAS Collaboration. “Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton–proton collisions at $\sqrt{s} = 7\text{ TeV}$ in 2010”. In: *Eur. Phys. J. C* 74 (2014), p. 3034. DOI: 10.1140/epjc/s10052-014-3034-9. arXiv: 1404.4562 [hep-ex].
- [3] J. J. Aubert et al. “Experimental Observation of a Heavy Particle J ”. In: *Phys. Rev. Lett.* 33 (Dec. 1974), pp. 1404–1406. DOI: 10.1103/PhysRevLett.33.1404.
- [4] J. E. Augustin et al. “Discovery of a Narrow Resonance in e^+e^- Annihilation”. In: *Phys. Rev. Lett.* 33 (Dec. 1974), pp. 1406–1408. DOI: 10.1103/PhysRevLett.33.1406.
- [5] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (Aug. 1977), pp. 252–255. DOI: 10.1103/PhysRevLett.39.252.
- [6] G. Arnison et al. “Experimental observation of lepton pairs of invariant mass around $95\text{ GeV}/c^2$ at the CERN SPS collider”. In: *Phys. Lett. B* 126.5 (1983), pp. 398–410. ISSN: 0370-2693. DOI: 10.1016/0370-2693(83)90188-0.
- [7] ATLAS Collaboration. “Search for high mass dilepton resonances in pp collisions at $\sqrt{s} = 7\text{ TeV}$ with the ATLAS experiment”. In: *Phys. Lett. B* 700 (2011), p. 163. DOI: 10.1016/j.physletb.2011.04.044. arXiv: 1103.6218 [hep-ex].
- [8] ATLAS Collaboration. “Search for high-mass new phenomena in the dilepton final state using proton–proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector”. In: *Phys. Lett. B* 761 (2016), p. 372. DOI: 10.1016/j.physletb.2016.08.055. arXiv: 1607.03669 [hep-ex].
- [9] ATLAS Collaboration. “Search for new high-mass phenomena in the dilepton final state using 36 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector”. In: *JHEP* 10 (2017), p. 182. DOI: 10.1007/JHEP10(2017)182. arXiv: 1707.02424 [hep-ex].
- [10] Gerard ’t Hooft and Martinus Veltman. “Regularization and renormalization of gauge fields”. In: *Nuclear Physics B* 44.1 (1972), pp. 189–213. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(72\)90279-9](https://doi.org/10.1016/0550-3213(72)90279-9). URL: <https://www.sciencedirect.com/science/article/pii/0550321372902799>.

- [11] Particle Data Group et al. “Review of particle physics”. In: *Progress of Theoretical and Experimental Physics* 2020.8 (2020), p. 083C01.
- [12] Paul Adrien Maurice Dirac. “The quantum theory of the electron”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.778 (1928), pp. 610–624.
- [13] Léo Morel, Zhibin Yao, Pierre Cladé, and Saïda Guellati-Khélifa. “Determination of the fine-structure constant with an accuracy of 81 parts per trillion”. In: *Nature* 588.7836 (2020), pp. 61–65.
- [14] Chien-Shiung Wu et al. “Experimental test of parity conservation in beta decay”. In: *Physical Review* 105.4 (1957), p. 1413.
- [15] Alex E Bernardini. “Dynamics of chiral oscillations: a comparative analysis with spin flipping”. In: *Journal of Physics G: Nuclear and Particle Physics* 32.1 (2005), p. 9.
- [16] François Englert and Robert Brout. “Broken symmetry and the mass of gauge vector mesons”. In: *Physical Review Letters* 13.9 (1964), p. 321.
- [17] Peter Ware Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Phys. Lett.* 12 (1964), pp. 132–133.
- [18] Peter W Higgs. “Broken symmetries and the masses of gauge bosons”. In: *Physical Review Letters* 13.16 (1964), p. 508.
- [19] Gerald S Guralnik, Carl R Hagen, and Thomas WB Kibble. “Global conservation laws and massless particles”. In: *Physical Review Letters* 13.20 (1964), p. 585.
- [20] Philip W Anderson. “Plasmons, gauge invariance, and mass”. In: *Physical Review* 130.1 (1963), p. 439.
- [21] FJ Hasert et al. “Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment”. In: *Nuclear Physics B* 73.1 (1974), pp. 1–22.
- [22] PA Baikov, KG Chetyrkin, and JH Kühn. “Five-Loop Running of the QCD coupling constant”. In: *Physical Review Letters* 118.8 (2017), p. 082002.
- [23] Particle Data Group et al. “Review of particle physics”. In: *Chinese Physics C* 40.10 (2016), p. 100001.
- [24] Y Sumino. “‘Coulomb + linear’ form of the static QCD potential in operator-product expansion”. In: *Physics Letters B* 595.1-4 (2004), pp. 387–392.
- [25] Francise D Aaron et al. “Combined measurement and QCD analysis of the inclusive $e \pm p$ scattering cross sections at HERA”. In: *Journal of High Energy Physics* 2010.1 (2010), p. 109.
- [26] Michael E Peskin and Daniel V Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.
- [27] Lyndon Evans and Philip Bryant. “LHC machine”. In: *Journal of instrumentation* 3.08 (2008), S08001.
- [28] ATLAS Collaboration. “The ATLAS simulation infrastructure”. In: *The European Physical Journal C* 70.3 (2010), pp. 823–874.
- [29] Richard D. Ball et al. “Parton distributions for the LHC run II”. In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

- [30] Geoffrey T Bodwin. “Factorization of the Drell-Yan cross section in perturbation theory”. In: *Physical Review D* 31.10 (1985), p. 2616.
- [31] Paolo Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.
- [32] Johan Alwall et al. “MadGraph 5 : Going Beyond”. In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].
- [33] Michelangelo L. Mangano et al. “ALPGEN, a generator for hard multiparton processes in hadronic collisions”. In: *JHEP* 07 (2003), p. 001. DOI: 10.1088/1126-6708/2003/07/001. arXiv: hep-ph/0206293.
- [34] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].
- [35] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].
- [36] Johannes Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].
- [37] Guido Altarelli and Giorgio Parisi. “Asymptotic freedom in parton language”. In: *Nuclear Physics B* 126.2 (1977), pp. 298–318.
- [38] Daniele Amati and Gabriele Veneziano. “Preconfinement as a Property of Perturbative QCD”. In: *Physics Letters B* 83.1 (1979), pp. 87–92.
- [39] Andy Buckley et al. “General-purpose event generators for LHC physics”. In: *Physics Reports* 504.5 (2011), pp. 145–233.
- [40] ATLAS Collaboration. “Measurement of the underlying event in jet events from 7 TeV proton–proton collisions with the ATLAS detector”. In: *Eur. Phys. J. C* 74 (2014), p. 2965. DOI: 10.1140/epjc/s10052-014-2965-5. arXiv: 1406.0392 [hep-ex].
- [41] S. Agostinelli et al. “GEANT4: A Simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- [42] J Schaarschmidt and. “The new ATLAS Fast Calorimeter Simulation”. In: *Journal of Physics: Conference Series* 898 (Oct. 2017), p. 042006. DOI: 10.1088/1742-6596/898/4/042006. URL: <https://doi.org/10.1088/1742-6596/898/4/042006>.
- [43] Georges Aad et al. “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 (2013), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].
- [44] Howard Georgi and Steven Weinberg. “Neutral currents in expanded gauge theories”. In: *Phys. Rev. D* 17 (1 Jan. 1978), pp. 275–279. DOI: 10.1103/PhysRevD.17.275. URL: <https://link.aps.org/doi/10.1103/PhysRevD.17.275>.
- [45] Martin Schmaltz. “Physics beyond the standard model (theory): Introducing the little Higgs”. In: *Nuclear Physics B-Proceedings Supplements* 117 (2003), pp. 40–49.

- [46] David B Kaplan, Howard Georgi, and Savas Dimopoulos. “Composite Higgs scalars”. In: *Physics Letters B* 136.3 (1984), pp. 187–190.
- [47] Nima Arkani-Hamed, Andrew G Cohen, Emanuel Katz, and Ann E Nelson. “The Littlest Higgs”. In: *Journal of High Energy Physics* 2002.07 (2002), p. 034.
- [48] David E Kaplan and Martin Schmaltz. “The Little Higgs from a simple group”. In: *Journal of High Energy Physics* 2003.10 (2003), p. 039.
- [49] Z Chacko, Hock-Seng Goh, and Roni Harnik. “Natural electroweak breaking from a mirror symmetry”. In: *Physical Review Letters* 96.23 (2006), p. 231802.
- [50] Tao Han, Heather E. Logan, Bob McElrath, and Lian-Tao Wang. “Phenomenology of the little Higgs model”. In: *Phys. Rev. D* 67 (9 May 2003), p. 095004. DOI: 10.1103/PhysRevD.67.095004. URL: <https://link.aps.org/doi/10.1103/PhysRevD.67.095004>.
- [51] Hock-Seng Goh and Shufang Su. “Phenomenology of Left-Right Twin Higgs Model”. In: *AIP Conf. Proc.* 903.1 (2007). Ed. by Jonathan L. Feng, pp. 431–434. DOI: 10.1063/1.2735216. arXiv: [hep-ph/0608330](https://arxiv.org/abs/hep-ph/0608330).
- [52] Howard Georgi and Sheldon L Glashow. “Unity of all elementary-particle forces”. In: *Physical Review Letters* 32.8 (1974), p. 438.
- [53] RN Mohapatra and Mina K Parida. “Type-II seesaw dominance in nonsupersymmetric and split-supersymmetry SO(10) models and proton lifetime”. In: *Physical Review D* 84.9 (2011), p. 095021.
- [54] David London and Jonathan L. Rosner. “Extra gauge bosons in E_6 ”. In: *Phys. Rev. D* 34 (5 Sept. 1986), pp. 1530–1546. DOI: 10.1103/PhysRevD.34.1530. URL: <https://link.aps.org/doi/10.1103/PhysRevD.34.1530>.
- [55] Elena Accomando et al. “Z’ physics with early LHC data”. In: *Physical Review D* 83.7 (2011), p. 075012.
- [56] K Abe et al. “Search for proton decay via $p \rightarrow e^+ \pi^0$ and $p \rightarrow \mu^+ \pi^0$ in 0.31 megaton Δ years exposure of the Super-Kamiokande water Cherenkov detector”. In: *Physical Review D* 95.1 (2017), p. 012004.
- [57] Stuart Raby. *Supersymmetric Grand Unified Theories: From Quarks to Strings Via SUSY GUTs*. eng. Cham: Springer International Publishing AG, 2017. ISBN: 3319552538.
- [58] Nima Arkani-Hamed, Savas Dimopoulos, and Gia Dvali. “The hierarchy problem and new dimensions at a millimeter”. In: *Physics Letters B* 429.3 (1998), pp. 263–272. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/S0370-2693\(98\)00466-3](https://doi.org/10.1016/S0370-2693(98)00466-3). URL: <http://www.sciencedirect.com/science/article/pii/S0370269398004663>.
- [59] JG Lee et al. “New Test of the Gravitational $1/r^2$ Law at Separations down to $52 \mu\text{m}$ ”. In: *Physical Review Letters* 124.10 (2020), p. 101101.
- [60] Lisa Randall and Raman Sundrum. “Large mass hierarchy from a small extra dimension”. In: *Physical Review Letters* 83.17 (1999), p. 3370.
- [61] Theodor Kaluza. “Zum unitätsproblem der physik”. In: *Sitzungsber. Preuss. Akad. Wiss. Berlin.(Math. Phys.)* 1921.966972 (1921), pp. 45–46.
- [62] O. Klein. “The Atomicity of Electricity as a Quantum Theory Law”. In: *Nature* 118 (1926), p. 516. DOI: 10.1038/118516a0.

- [63] Hewett Davoudiasl, JL Hewett, and TG Rizzo. “Experimental probes of localized gravity: On and off the wall”. In: *Physical Review D* 63.7 (2001), p. 075004.
- [64] Benjamin C Allanach et al. “Exploring small extra dimensions at the large hadron collider”. In: *Journal of High Energy Physics* 2002.12 (2003), p. 039.
- [65] Tony Gherghetta and Alex Pomarol. “Bulk fields and supersymmetry in a slice of AdS”. In: *Nuclear Physics B* 586.1-2 (2000), pp. 141–162.
- [66] Hooman Davoudiasl, Shrihari Gopalakrishna, Eduardo Ponton, and Jose Santiago. “Warped five-dimensional models: phenomenological status and experimental prospects”. In: *New Journal of Physics* 12.7 (2010), p. 075011.
- [67] Scott Dodelson. *Modern cosmology*. Elsevier, 2003.
- [68] DJ Fixsen. “The temperature of the cosmic microwave background”. In: *The Astrophysical Journal* 707.2 (2009), p. 916.
- [69] Nabila Aghanim et al. “Planck 2018 results-VI. Cosmological parameters”. In: *Astronomy & Astrophysics* 641 (2020), A6.
- [70] Leszek Roszkowski, Enrico Maria Sessolo, and Sebastian Trojanowski. “WIMP dark matter candidates and searches – current status and future prospects”. In: *Reports on Progress in Physics* 81.6 (2018), p. 066201.
- [71] Andreas Birkedal, Andrew Noble, Maxim Perelstein, and Andrew Spray. “Little Higgs dark matter”. In: *Physical Review D* 74.3 (2006), p. 035002.
- [72] Nathaniel Craig, Andrey Katz, Matt Strassler, and Raman Sundrum. “Naturalness in the dark at the LHC”. In: *Journal of High Energy Physics* 2015.7 (2015), pp. 1–49.
- [73] Kaustubh Agashe and Géraldine Servant. “Baryon number in warped grand unified theories: model building and (dark matter related) phenomenology”. In: *Journal of Cosmology and Astroparticle Physics* 2005.02 (2005), p. 002.
- [74] M Aguilar et al. “The Alpha Magnetic Spectrometer (AMS) on the international space station: Part II – Results from the first seven years”. In: *Physics reports* (2020).
- [75] Daniel Abercrombie et al. “Dark matter benchmark models for early LHC run-2 searches: report of the ATLAS/CMS dark matter forum”. In: *Physics of the Dark Universe* 27 (2020), p. 100371.
- [76] ATLAS Collaboration. “Constraints on mediator-based dark matter and scalar dark energy models using $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS detector”. In: *JHEP* 05 (2019), p. 142. DOI: 10.1007/JHEP05(2019)142. arXiv: 1903.01400 [hep-ex].
- [77] Christopher T Hill and Elizabeth H Simmons. “Strong dynamics and electroweak symmetry breaking”. In: *Physics Reports* 381.4-6 (2003), pp. 235–402.
- [78] Nima Arkani-Hamed, Andrew G Cohen, and Howard Georgi. “Electroweak symmetry breaking from dimensional deconstruction”. In: *Physics Letters B* 513.1-2 (2001), pp. 232–240.
- [79] Hock-Seng Goh and Shufang Su. “Phenomenology of the left-right twin Higgs model”. In: *Physical Review D* 75.7 (2007), p. 075010.
- [80] Jalal Abdallah et al. “Simplified models for dark matter searches at the LHC”. In: *Physics of the Dark Universe* 9 (2015), pp. 8–23.

- [81] D Berdine, N Kauer, and D Rainwater. “Breakdown of the narrow width approximation for new physics”. In: *Physical Review Letters* 99.11 (2007), p. 111601.
- [82] RS Chivukula and EH Simmons. “Electroweak limits on nonuniversal Z’ bosons”. In: *Physical Review D* 66.1 (2002), p. 015006.
- [83] ATLAS collaboration. *ATLAS: A 25-year Insider Story of the LHC Experiment*. World Scientific, 2019.
- [84] ATLAS Collaboration. “Computer generated image of the whole ATLAS detector”. Mar. 2008. URL: <https://cds.cern.ch/record/1095924>.
- [85] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.
- [86] ATLAS Collaboration. “Studies of the performance of the ATLAS detector using cosmic-ray muons”. In: *Eur. Phys. J. C* 71 (2011), p. 1593. DOI: 10.1140/epjc/s10052-011-1593-6. arXiv: 1011.6665 [hep-ex].
- [87] ATLAS Collaboration. “Commissioning of the ATLAS Muon Spectrometer with cosmic rays”. In: *Eur. Phys. J. C* 70 (2010), p. 875. DOI: 10.1140/epjc/s10052-010-1415-2. arXiv: 1006.4384 [hep-ex].
- [88] *Performance of the muon spectrometer alignment in 2017 and 2018 data*. Tech. rep. ATL-MUON-PUB-2021-002. Geneva: CERN, Mar. 2021. URL: <https://cds.cern.ch/record/2753329>.
- [89] ATLAS Collaboration. “Alignment of the ATLAS Inner Detector in Run 2”. In: *The European Physical Journal C* 80.12 (2020), pp. 1–41.
- [90] ATLAS Collaboration. *Interactive web pages for radiation environment exploration of ATLAS*. ATL-SOFT-PUB-2020-003. 2020. URL: <https://cds.cern.ch/record/2742378/files/ATL-SOFT-PUB-2020-003.pdf>.
- [91] Malte Backhaus et al. “The upgraded pixel detector of the ATLAS experiment for Run 2 at the Large Hadron Collider”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 831 (2016), pp. 65–70.
- [92] ATLAS Collaboration. *Particle Identification Performance of the ATLAS Transition Radiation Tracker*. ATLAS-CONF-2011-128. 2011. URL: <https://cds.cern.ch/record/1383793>.
- [93] Vasiliki A Mitsou and ATLAS TRT Collaboration. “The ATLAS transition radiation tracker”. In: *Astroparticle, Particle and Space Physics, Detectors and Medical Physics Applications*. World Scientific, 2004, pp. 497–501.
- [94] Claus Grupen et al. *Particle detectors*. Vol. 11. Cambridge University Press, 2008.
- [95] N Shalanda et al. “Radioactive source control and electronics for the ATLAS tile calorimeter cesium calibration system”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 508.3 (2003), pp. 276–286.
- [96] ATLAS collaboration. “Operation of the ATLAS trigger system in Run 2”. In: *Journal of Instrumentation* 15.10 (2020), P10004.
- [97] R Achenbach et al. “The ATLAS level-1 calorimeter trigger”. In: *Journal of Instrumentation* 3.03 (2008), P03001.

- [98] 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. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].
- [99] 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>.
- [100] 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. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].
- [101] ATLAS Collaboration. “Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data”. In: *JINST* (2018). arXiv: 1812.03848 [hep-ex].
- [102] ATLAS Collaboration. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV”. In: (2019). arXiv: 1902.04655 [physics.ins-det].
- [103] Christos Anastopoulos et al. *Support Note for Electron ID: electron reconstruction and identification*. Tech. rep. Geneva: CERN, July 2017. URL: <https://cds.cern.ch/record/2273328>.
- [104] T Addy et al. *Search for high-mass dilepton resonances in 5/fb of pp collisions at $\sqrt{s} = 7$ TeV*. Tech. rep. Support note of the EXOTICS High mass dilepton paper at 7 TeV (arXiv:1209.2535). Geneva: CERN, Feb. 2012. URL: <https://cds.cern.ch/record/1421545>.
- [105] ATLAS Collaboration. *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*. 2020. arXiv: 2012.00578 [hep-ex].
- [106] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 76 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].
- [107] Simon Viel. “Search for new neutral high-mass resonances decaying into muon pairs with the ATLAS detector”. PhD thesis. 2014. URL: <https://cds.cern.ch/record/1748486>.
- [108] AA Abdelalim et al. *Search for high-mass dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV (2011 update for the EPS conference)*. Tech. rep. Geneva: CERN, June 2011. URL: <https://cds.cern.ch/record/1361706>.
- [109] Sidney D. Drell and Tung-Mow Yan. “Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies”. In: *Phys. Rev. Lett.* 25 (5 Aug. 1970), pp. 316–320. DOI: 10.1103/PhysRevLett.25.316. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.25.316>.
- [110] Hung-Liang Lai et al. “New parton distributions for collider physics”. In: *Phys. Rev. D* 82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

- [111] Piotr Golonka and Zbigniew Was. “PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays”. In: *Eur. Phys. J. C* 45 (2006), pp. 97–107. DOI: 10.1140/epjc/s2005-02396-4. arXiv: hep-ph/0506026.
- [112] Rikkert Frederix, Emanuele Re, and Paolo Torrielli. “Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@ NLO”. In: *Journal of High Energy Physics* 2012.9 (2012), p. 130.
- [113] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”. In: *JHEP* 03 (2013), p. 015. DOI: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph].
- [114] Michal Czakon and Alexander Mitov. “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”. In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].
- [115] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth. A* 462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.
- [116] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. “A brief introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), p. 852. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].
- [117] Richard D. Ball et al. “Parton distributions with LHC data”. In: *Nucl. Phys. B* 867 (2013), pp. 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].
- [118] 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>.
- [119] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: <http://cds.cern.ch/record/1966419>.
- [120] Daniel Stump et al. “Inclusive jet production, parton distributions, and the search for new physics”. In: *JHEP* 10 (2003), p. 046. DOI: 10.1088/1126-6708/2003/10/046. arXiv: hep-ph/0303013.
- [121] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *Journal of High Energy Physics* 2010.6 (2010), p. 43.
- [122] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method”. In: *Journal of High Energy Physics* 2007.11 (2007), p. 070.
- [123] Ye Li and Frank Petriello. “Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code”. In: *Physical Review D* 86.9 (2012), p. 094034.
- [124] Charalampos Anastasiou, Lance Dixon, Kirill Melnikov, and Frank Petriello. “High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order”. In: *Physical Review D* 69.9 (2004), p. 094008.
- [125] D Bardin et al. “SANC integrator in the progress: QCD and EW contributions”. In: *JETP letters* 96.5 (2012), pp. 285–289.

- [126] AD Martin, RG Roberts, WJ Stirling, and RS Thorne. “Parton distributions incorporating QED contributions”. In: *The European Physical Journal C-Particles and Fields* 39.2 (2005), pp. 155–161.
- [127] ATLAS Collaboration. “Atlantis Event Display”. In: (2015). URL: <http://atlantis.web.cern.ch/atlantis/>.
- [128] L Chevalier et al. *PERSINT Event Display for ATLAS*. Tech. rep. ATL-SOFT-PUB-2012-001. Geneva: CERN, Dec. 2012. URL: <https://cds.cern.ch/record/1501131>.
- [129] Deshan Kavishka Abhayasinghe et al. *Internal documentation for ATLAS full Run-2 dilepton resonant analysis*. Tech. rep. Geneva: CERN, June 2018. URL: <https://cds.cern.ch/record/2622420>.
- [130] ATLAS Collaboration. *The new Fast Calorimeter Simulation in ATLAS*. ATL-SOFT-PUB-2018-002. 2018. URL: <https://cds.cern.ch/record/2630434>.
- [131] Matteo Frigo and Steven G. Johnson. “The Design and Implementation of FFTW3”. In: *Proceedings of the IEEE* 93.2 (2005). Special issue on “Program Generation, Optimization, and Platform Adaptation”, pp. 216–231.
- [132] Jerzy Neyman and Egon Sharpe Pearson. “IX. On the problem of the most efficient tests of statistical hypotheses”. In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 231.694-706 (1933), pp. 289–337.
- [133] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011). [Erratum: *Eur. Phys. J. C* 73 (2013) 2501], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].
- [134] Samuel S Wilks. “The large-sample distribution of the likelihood ratio for testing composite hypotheses”. In: *The annals of mathematical statistics* 9.1 (1938), pp. 60–62.
- [135] Abraham Wald. “Tests of statistical hypotheses concerning several parameters when the number of observations is large”. In: *Transactions of the American Mathematical society* 54.3 (1943), pp. 426–482.
- [136] Alexander L Read. “Presentation of search results: the CLs technique”. In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), p. 2693.
- [137] ATLAS Collaboration. “Search for resonances in diphoton events at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 09 (2016), p. 001. DOI: 10.1007/JHEP09(2016)001. arXiv: 1606.03833 [hep-ex].
- [138] CDF Collaboration. “Search for new particles decaying into dijets in proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV”. In: *Physical Review D* 79.11 (2009), p. 112002.
- [139] Nicolas Berger, Andreas Hoecker, and Bruno Lenzi. *Hfitter : a Statistical Framework for High-Energy Physics Analysis*. Tech. rep. ATL-COM-PHYS-2016-442. Geneva: CERN, May 2016. URL: <https://cds.cern.ch/record/2151063>.
- [140] Lorenzo Moneta et al. “The RooStats project”. In: *arXiv preprint arXiv:1009.1003* (2010).
- [141] M Hatlo et al. “Developments of mathematical software libraries for the LHC experiments”. In: *IEEE Transactions on Nuclear Science* 52.6 (2005), pp. 2818–2822.

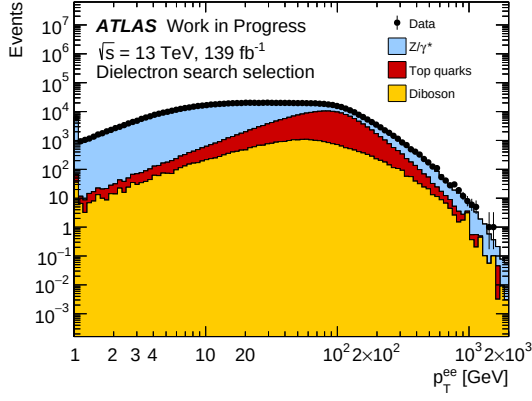
- [142] ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. ATLAS-CONF-2019-021. 2019. URL: <https://cds.cern.ch/record/2677054>.
- [143] G Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *Journal of Instrumentation* 13.07 (2018), P07017.
- [144] ATLAS Statistics Forum. “Recommendations for presentation of error bars”. In: (2011). URL: <http://www.pp.rhul.ac.uk/~cowan/atlas/ErrorBars.pdf>.
- [145] Christopher Willis. *The LPXSignalReweightingTool: A Package for Signal Template Reweighting in High-Energy Physics Analysis*. Tech. rep. ATL-COM-PHYS-2016-912. Geneva: CERN, July 2016. URL: <https://cds.cern.ch/record/2199258>.
- [146] ATLAS collaboration. “LPXSignalReweightingTool code repository”. In: (). URL: <https://svnweb.cern.ch/trac/atlasphys-exo/browser/Physics/Exotic/LPX/CommonTools/LPXSignalReweightingTool/>.
- [147] Charalampos Anastasiou, Lance J. Dixon, Kirill Melnikov, and Frank Petriello. “High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order”. In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266.
- [148] Jon Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].
- [149] Sayipjamal Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Phys. Rev. D* 93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph].
- [150] Patrick Motylinski, Lucian Harland-Lang, Alan D Martin, and Robert S Thorne. “Updates of PDFs for the 2nd LHC run”. In: *Nuclear and particle physics proceedings* 273 (2016), pp. 2136–2141.
- [151] S Alekhin, J Blümlein, and S Moch. “The ABM parton distributions tuned to LHC data”. In: *Physical Review D* 89.5 (2014), p. 054028.
- [152] Pedro Jimenez-Delgado and Ewald Reya. “Delineating parton distributions and the strong coupling”. In: *Physical Review D* 89.7 (2014), p. 074049.
- [153] ATLAS Collaboration. “Combination of searches for heavy resonances decaying into bosonic and leptonic final states using 36 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Rev. D* 98 (2018), p. 052008. DOI: 10.1103/PhysRevD.98.052008. arXiv: 1808.02380 [hep-ex].
- [154] Markus Zinser et al. *Search for new high-mass resonances in the dilepton final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2016-1404. Geneva: CERN, Sept. 2016. URL: <https://cds.cern.ch/record/2217641>.
- [155] Boris Delaunay et al. “Sur la sphere vide”. In: *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk* 7.793-800 (1934), pp. 1–2.

- [156] ATLAS Collaboration. “Search for a heavy charged boson in events with a charged lepton and missing transverse momentum from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Rev. D* 100 (2019), p. 052013. DOI: 10.1103/PhysRevD.100.052013. arXiv: 1906.05609 [hep-ex].
- [157] “Repository containing code for obtaining DM cross sections and exclusion curves for the dilepton analysis.” In: (). URL: <https://gitlab.cern.ch/atlas-phys/exot/lpx/dilepton/dileptondm>.
- [158] Andy Nelson. “DarkMatterWidthCalculator code repository”. In: (). URL: <https://github.com/LHC-DMWG/DarkMatterWidthCalculation>.
- [159] Chiara Arina et al. “Studying dark matter with MadDM 3.1: a short user guide”. In: *PoS TOOLS2020* (2021), p. 009. DOI: 10.22323/1.392.0009. arXiv: 2012.09016 [hep-ph].
- [160] *Dark matter summary plots for s-channel mediators*. ATL-PHYS-PUB-2021-006. Geneva, 2021. URL: <http://cds.cern.ch/record/2758386>.
- [161] Mariana Frank, Yaşar Hiçyılmaz, Stefano Moretti, and Özer Özdal. “E6 motivated UMSSM confronts experimental data”. In: *Journal of High Energy Physics* 2020.2004.01415 (2020), pp. 1–23.
- [162] Ethan Carragher et al. “Convergent Bayesian Global Fits of 4D Composite Higgs Models”. In: *arXiv preprint arXiv:2101.00428* (2021).
- [163] Felix Kahlhoefer, Alexander Mück, Stefan Schulte, and Patrick Tunney. “Interference effects in dilepton resonance searches for Z' bosons and dark matter mediators”. In: *Journal of High Energy Physics* 2020.3 (2020), pp. 1–23.
- [164] BC Allanach, JM Butterworth, and Tyler Corbett. “Collider constraints on Z' models for neutral current B-anomalies”. In: *Journal of High Energy Physics* 2019.8 (2019), pp. 1–23.
- [165] Roel Aaij et al. “Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays”. In: *Journal of High Energy Physics* 2017.8 (2017), pp. 1–31.
- [166] LHCb Collaboration. “Test of lepton universality in beauty-quark decays”. In: *arXiv:2103.11769* (2021).
- [167] ATLAS Collaboration. *Prospects for searches for heavy Z' and W' bosons in fermionic final states with the ATLAS experiment at the HL-LHC*. ATL-PHYS-PUB-2018-044. 2018. URL: <https://cds.cern.ch/record/2650549>.
- [168] Meghan Frate et al. “Modeling smooth backgrounds and generic localized signals with Gaussian processes”. In: *arXiv preprint arXiv:1709.05681* (2017).
- [169] ATLAS Collaboration. “Dijet Resonance Search with Weak Supervision Using $\sqrt{s} = 13$ TeV pp Collisions in the ATLAS Detector”. In: *Physical Review Letters* 125.13 (2020), p. 131801.
- [170] ATLAS Collaboration. *Fast simulation of the ATLAS calorimeter system with Generative Adversarial Networks*. ATL-SOFT-PUB-2020-006. 2020. URL: <https://cds.cern.ch/record/2746032>.
- [171] Robert B Davies. “Hypothesis testing when a nuisance parameter is present only under the alternative”. In: *Biometrika* 74.1 (1987), pp. 33–43.

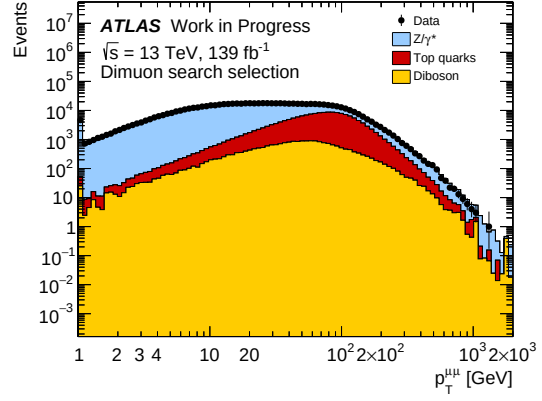
- [172] Eilam Gross and Ofer Vitells. “Trial factors for the look elsewhere effect in high energy physics”. In: *The European Physical Journal C* 70.1 (2010), pp. 525–530. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1470-8. URL: <http://dx.doi.org/10.1140/epjc/s10052-010-1470-8>.
- [173] Ofer Vitells and Eilam Gross. “Estimating the significance of a signal in a multi-dimensional search”. In: *Astroparticle Physics* 35.5 (2011), pp. 230–234. ISSN: 0927-6505. DOI: <https://doi.org/10.1016/j.astropartphys.2011.08.005>. URL: <https://www.sciencedirect.com/science/article/pii/S0927650511001630>.
- [174] Ryan Edgar, Dante Amidei, Christopher Grud, and Karishma Sekhon. “Functional Decomposition: A new method for search and limit setting”. In: *arXiv preprint arXiv:1805.04536* (2018).
- [175] ATLAS Collaboration. “Search for Low-Mass Dijet Resonances Using Trigger-Level Jets with the ATLAS Detector in pp Collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Rev. Lett.* 121 (2018), p. 081801. DOI: 10.1103/PhysRevLett.121.081801. arXiv: 1804.03496 [hep-ex].
- [176] ATLAS Collaboration. “Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 03 (2020), p. 145. DOI: 10.1007/JHEP03(2020)145. arXiv: 1910.08447 [hep-ex].
- [177] “RooSWiFt”. In: (). URL: <https://gitlab.cern.ch/atlas-phys/exot/lpx/dilepton/RooSwift>.

Appendix A

Additional dilepton modelling plots

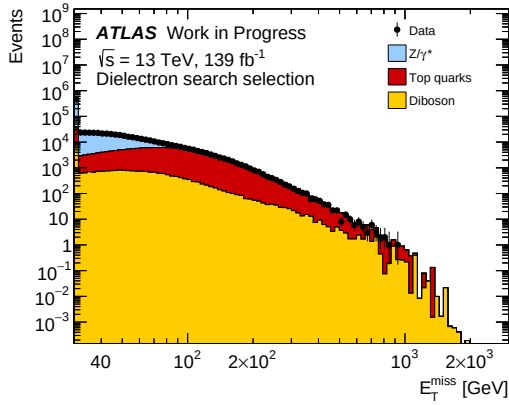


(a) Dielectron channel

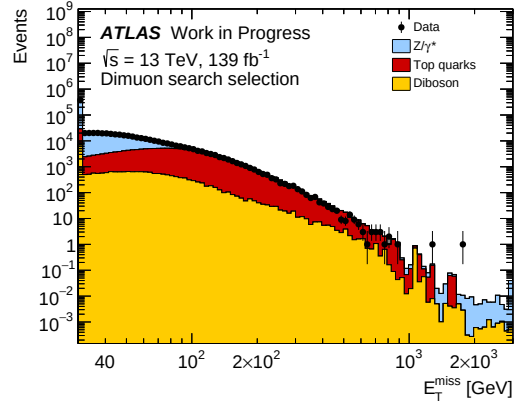


(b) Dimuon channel

Figure A.1: Comparison of the dilepton system transverse momentum $p_T^{\ell\ell}$ distributions in data and MC for the full Run-2 dataset. Note that no systematic variations are applied to the MC distributions and the fake electron background is not included.



(a) Dielectron channel



(b) Dimuon channel

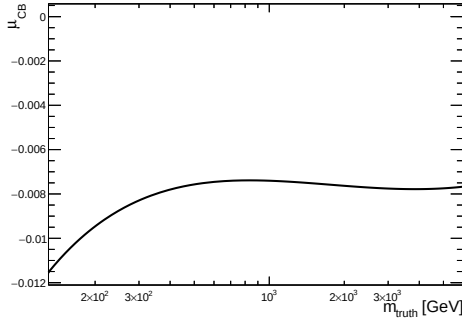
Figure A.2: Comparison of the missing transverse momentum E_T^{miss} distributions in data and MC for the full Run-2 dataset. Note that no systematic variations are applied to the MC distributions and the fake electron background is not included.

Table A.1: POWHEG+PYTHIA8 Drell-Yan Monte Carlo samples used in the dielectron and dimuon channel analysis. The first column gives the process, the second column gives the mass range in which the Drell-Yan process was simulated, the third column gives the internal ATLAS dataset number (DSID). For each sample the cross section times branching ratio with which the POWHEG+PYTHIA8 generator produced the sample, the number of produced events and effective luminosity per campaign are given. Table taken from [129].

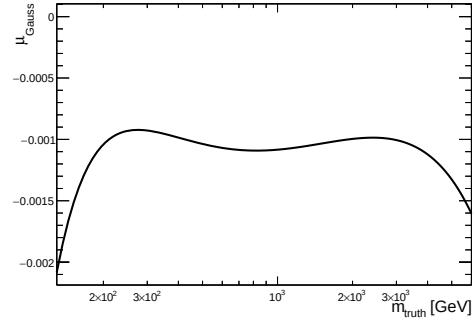
Process	m_{ee} [GeV]	ee: DSID	σ_{Br} [pb]	mc16a N_{evt}	eff lumi [fb $^{-1}$]	mc16d N_{evt}	eff lumi [fb $^{-1}$]	mc16e N_{evt}	eff lumi [fb $^{-1}$]	truth N_{evt}	eff lumi [fb $^{-1}$]
$Z \rightarrow ee$	120 – 180	301000	$1.748 \times 10^{+01}$	1496000	8.56E+01	1615000	9.24E+01	2482000	1.42E+02	139840000	8.00E+03
$Z \rightarrow ee$	180 – 250	301001	$2.921 \times 10^{+00}$	900000	3.08E+02	957000	3.28E+02	1500000	5.13E+02	23370000	8.00E+03
$Z \rightarrow ee$	250 – 400	301002	$1.082 \times 10^{+00}$	1200000	1.11E+03	1240000	1.15E+03	1998000	1.85E+03	8660000	8.00E+03
$Z \rightarrow ee$	400 – 600	301003	1.955×10^{-01}	900000	4.60E+03	925000	4.73E+03	1495000	7.65E+03	1800000	9.21E+03
$Z \rightarrow ee$	600 – 800	301004	3.740×10^{-02}	510000	1.36E+04	536000	1.43E+04	850000	2.27E+04	1300000	3.48E+04
$Z \rightarrow ee$	800 – 1000	301005	1.061×10^{-02}	250000	2.36E+04	262000	2.47E+04	420000	3.96E+04	1000000	9.43E+04
$Z \rightarrow ee$	1000 – 1250	301006	4.258×10^{-03}	100000	2.35E+04	112000	2.63E+04	100000	2.35E+04	1000000	2.35E+05
$Z \rightarrow ee$	1250 – 1500	301007	1.422×10^{-03}	50000	3.52E+04	62000	4.36E+04	90000	6.33E+04	1000000	7.03E+05
$Z \rightarrow ee$	1500 – 1750	301008	5.452×10^{-04}	50000	9.17E+04	62000	1.14E+05	90000	1.65E+05	1000000	1.83E+06
$Z \rightarrow ee$	1750 – 2000	301009	2.299×10^{-04}	100000	4.35E+05	125000	5.43E+05	170000	7.39E+05	1000000	4.35E+06
$Z \rightarrow ee$	2000 – 2250	301010	1.039×10^{-04}	50000	4.81E+05	62000	5.96E+05	90000	8.65E+05	550000	5.29E+06
$Z \rightarrow ee$	2250 – 2500	301011	4.940×10^{-05}	50000	1.02E+06	62000	1.27E+06	90000	1.84E+06	550000	1.12E+07
$Z \rightarrow ee$	2500 – 2750	301012	2.445×10^{-05}	50000	2.08E+06	62000	2.58E+06	90000	3.75E+06	550000	2.29E+07
$Z \rightarrow ee$	2750 – 3000	301013	1.249×10^{-05}	50000	4.17E+06	62000	5.17E+06	90000	7.50E+06	550000	4.58E+07
$Z \rightarrow ee$	3000 – 3500	301014	1.003×10^{-05}	50000	5.00E+06	50000	5.00E+06	90000	9.00E+06	300000	3.00E+07
$Z \rightarrow ee$	3500 – 4000	301015	2.934×10^{-06}	50000	1.67E+07	56000	1.80E+07	90000	3.00E+07	250000	8.33E+07
$Z \rightarrow ee$	4000 – 4500	301016	8.976×10^{-07}	50000	5.00E+07	58000	5.80E+07	90000	9.00E+07	250000	2.50E+08
$Z \rightarrow ee$	4500 – 5000	301017	2.807×10^{-07}	50000	1.78E+08	61000	2.17E+08	90000	3.21E+08	250000	8.91E+08
$Z \rightarrow ee$	5000 – ∞	301018	1.265×10^{-07}	50000	3.95E+08	62000	4.90E+08	90000	7.12E+08	250000	1.98E+09
$Z \rightarrow \mu\mu$	120 – 180	301020	$1.748 \times 10^{+01}$	2000000	1.14E+02	1486000	8.50E+01	3318000	1.90E+02	139840000	8.00E+03
$Z \rightarrow \mu\mu$	180 – 250	301021	$2.921 \times 10^{+00}$	1150000	3.94E+02	898000	3.07E+02	1899000	6.50E+02	23370000	8.00E+03
$Z \rightarrow \mu\mu$	250 – 400	301022	$1.082 \times 10^{+00}$	1345000	1.24E+03	1195000	1.10E+03	2249000	2.08E+03	8660000	8.00E+03
$Z \rightarrow \mu\mu$	400 – 600	301023	1.955×10^{-01}	1000000	5.11E+03	900000	4.60E+03	1660000	8.49E+03	1800000	9.21E+03
$Z \rightarrow \mu\mu$	600 – 800	301024	3.740×10^{-02}	600000	1.60E+04	495000	1.32E+04	1000000	2.67E+04	1300000	3.48E+04
$Z \rightarrow \mu\mu$	800 – 1000	301025	1.061×10^{-02}	300000	2.83E+04	250000	2.36E+04	495000	4.67E+04	1000000	9.43E+04
$Z \rightarrow \mu\mu$	1000 – 1250	301026	4.258×10^{-03}	150000	3.52E+04	100000	2.35E+04	250000	5.87E+04	1000000	2.35E+05
$Z \rightarrow \mu\mu$	1250 – 1500	301027	1.422×10^{-03}	100000	7.03E+04	50000	3.52E+04	170000	1.20E+05	1000000	7.03E+05
$Z \rightarrow \mu\mu$	1500 – 1750	301028	5.452×10^{-04}	100000	1.83E+05	50000	9.17E+04	170000	3.12E+05	1000000	1.83E+06
$Z \rightarrow \mu\mu$	1750 – 2000	301029	2.299×10^{-04}	100000	4.35E+05	50000	2.17E+05	170000	7.39E+05	1000000	4.35E+06
$Z \rightarrow \mu\mu$	2000 – 2250	301030	1.039×10^{-04}	100000	9.62E+05	50000	4.81E+05	170000	1.63E+06	550000	5.29E+06
$Z \rightarrow \mu\mu$	2250 – 2500	301031	4.940×10^{-05}	100000	2.04E+06	50000	1.02E+06	170000	3.47E+06	550000	1.12E+07
$Z \rightarrow \mu\mu$	2500 – 2750	301032	2.445×10^{-05}	100000	4.17E+06	50000	2.08E+06	170000	7.08E+06	550000	2.29E+07
$Z \rightarrow \mu\mu$	2750 – 3000	301033	1.249×10^{-05}	100000	8.33E+06	50000	4.17E+06	170000	1.42E+07	550000	4.58E+07
$Z \rightarrow \mu\mu$	3000 – 3500	301034	1.003×10^{-05}	100000	1.00E+07	50000	5.00E+06	170000	1.70E+07	300000	3.00E+07
$Z \rightarrow \mu\mu$	3500 – 4000	301035	2.934×10^{-06}	100000	3.33E+07	50000	1.67E+07	170000	5.67E+07	250000	8.33E+07
$Z \rightarrow \mu\mu$	4000 – 4500	301036	8.976×10^{-07}	100000	1.00E+08	50000	5.00E+07	170000	1.70E+08	250000	2.50E+08
$Z \rightarrow \mu\mu$	4500 – 5000	301037	2.807×10^{-07}	100000	3.56E+08	50000	1.78E+08	170000	6.06E+08	250000	8.91E+08
$Z \rightarrow \mu\mu$	5000 – ∞	301038	1.265×10^{-07}	100000	7.91E+08	50000	3.95E+08	170000	1.34E+09	250000	1.98E+09

Table A.2: PYTHIA8 Drell-Yan Monte Carlo samples used in the dielectron and dimuon channel analysis mostly to calculate theoretical cross-sections in fiducial volume using sample reweighting technique. The first column gives the process, the second column gives the mass range in which the Drell-Yan process was simulated, the third column gives the internal ATLAS dataset number (DSID). For each sample the cross section times branching ratio with which the POWHEG+PYTHIA8 generator produced the sample, and the number of produced events per MC campaign, are given. Table taken from [129].

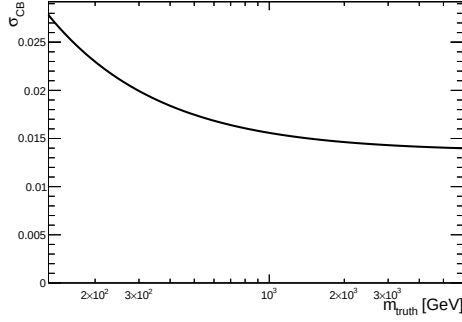
Process	$m_{\ell\ell}$ [GeV]	DSID	σBr [pb]	mc16a N_{evt}	mc16d N_{evt}	mc16e N_{evt}
$Z \rightarrow ee$	70 – 120	301540	$1.426 \times 10^{+03}$	250000	313000	250000
$Z \rightarrow ee$	120 – 180	301541	$1.384 \times 10^{+01}$	1250000	1565000	250000
$Z \rightarrow ee$	180 – 250	301542	$2.333 \times 10^{+00}$	1250000	1565000	250000
$Z \rightarrow ee$	250 – 400	301543	8.675×10^{-01}	1400000	1753000	150000
$Z \rightarrow ee$	400 – 600	301544	1.557×10^{-01}	1349000	1690000	100000
$Z \rightarrow ee$	600 – 800	301545	2.971×10^{-02}	1300000	1630000	50000
$Z \rightarrow ee$	800 – 1000	301546	8.347×10^{-03}	1050000	1315000	50000
$Z \rightarrow ee$	1000 – 1250	301547	3.311×10^{-03}	1050000	1315000	50000
$Z \rightarrow ee$	1250 – 1500	301548	1.098×10^{-03}	1050000	1314000	50000
$Z \rightarrow ee$	1500 – 1750	301549	4.187×10^{-04}	1050000	1305000	50000
$Z \rightarrow ee$	1750 – 2000	301550	1.760×10^{-04}	1050000	1315000	50000
$Z \rightarrow ee$	2000 – 2250	301551	7.996×10^{-05}	550000	690000	50000
$Z \rightarrow ee$	2250 – 2500	301552	3.823×10^{-05}	550000	689000	50000
$Z \rightarrow ee$	2500 – 2750	301553	1.905×10^{-05}	550000	690000	50000
$Z \rightarrow ee$	2750 – 3000	301554	9.854×10^{-06}	550000	690000	50000
$Z \rightarrow ee$	3000 – 3500	301555	8.045×10^{-06}	300000	376000	50000
$Z \rightarrow ee$	3500 – 4000	301556	2.419×10^{-06}	299000	380000	50000
$Z \rightarrow ee$	4000 – 4500	301557	7.577×10^{-07}	300000	350000	50000
$Z \rightarrow ee$	4500 – 5000	301558	2.433×10^{-07}	300000	380000	50000
$Z \rightarrow ee$	5000 – ∞	301559	1.167×10^{-07}	298000	356000	50000
$Z \rightarrow \mu\mu$	70 – 120	301560	$1.426 \times 10^{+03}$	250000	315000	250000
$Z \rightarrow \mu\mu$	120 – 180	301561	$1.384 \times 10^{+01}$	1250000	1565000	250000
$Z \rightarrow \mu\mu$	180 – 250	301562	$2.333 \times 10^{+00}$	1250000	1565000	250000
$Z \rightarrow \mu\mu$	250 – 400	301563	8.675×10^{-01}	1400000	1754000	150000
$Z \rightarrow \mu\mu$	400 – 600	301564	1.557×10^{-01}	1350000	1690000	100000
$Z \rightarrow \mu\mu$	600 – 800	301565	2.971×10^{-02}	1300000	1630000	50000
$Z \rightarrow \mu\mu$	800 – 1000	301566	8.347×10^{-03}	1050000	1315000	50000
$Z \rightarrow \mu\mu$	1000 – 1250	301567	3.311×10^{-03}	1050000	1315000	50000
$Z \rightarrow \mu\mu$	1250 – 1500	301568	1.098×10^{-03}	1050000	1315000	50000
$Z \rightarrow \mu\mu$	1500 – 1750	301569	4.187×10^{-04}	1050000	1315000	50000
$Z \rightarrow \mu\mu$	1750 – 2000	301570	1.760×10^{-04}	1050000	1315000	50000
$Z \rightarrow \mu\mu$	2000 – 2250	301571	7.996×10^{-05}	550000	690000	50000
$Z \rightarrow \mu\mu$	2250 – 2500	301572	3.823×10^{-05}	550000	690000	50000
$Z \rightarrow \mu\mu$	2500 – 2750	301573	1.905×10^{-05}	550000	690000	50000
$Z \rightarrow \mu\mu$	2750 – 3000	301574	9.854×10^{-06}	550000	690000	50000
$Z \rightarrow \mu\mu$	3000 – 3500	301575	8.045×10^{-06}	300000	380000	50000
$Z \rightarrow \mu\mu$	3500 – 4000	301576	2.419×10^{-06}	300000	380000	50000
$Z \rightarrow \mu\mu$	4000 – 4500	301577	7.577×10^{-07}	300000	380000	50000
$Z \rightarrow \mu\mu$	4500 – 5000	301578	2.433×10^{-07}	300000	380000	50000
$Z \rightarrow \mu\mu$	5000 – ∞	301579	1.167×10^{-07}	300000	380000	50000



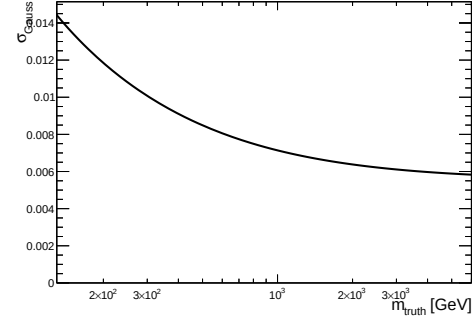
(a) μ_{CB}



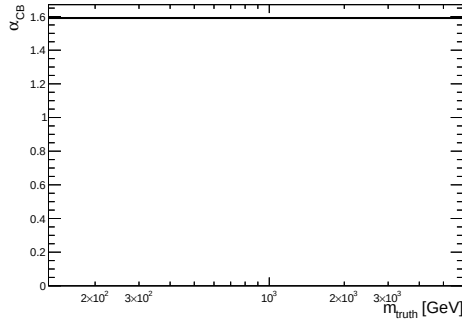
(b) μ_G



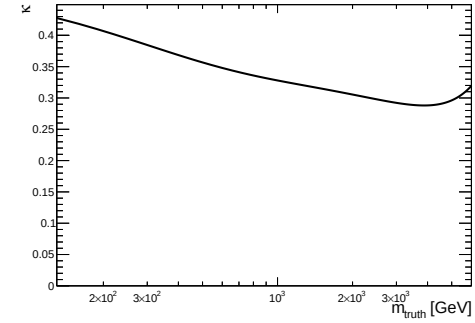
(c) σ_{CB}



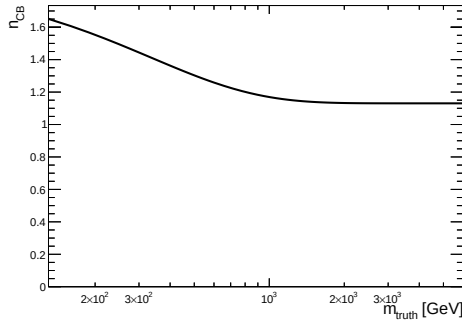
(d) σ_G



(e) α_{CB}

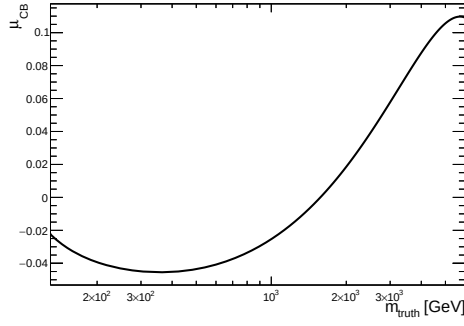


(f) κ

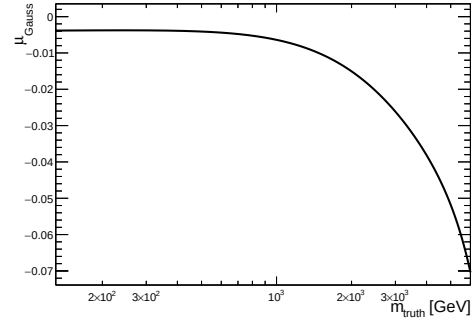


(g) n_{CB}

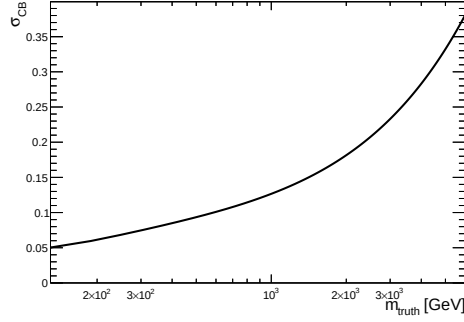
Figure A.3: Fitted parameterizations of the dielectron channel $m_{\ell\ell}$ response model parameters (equation 6.7) as functions of truth-level dilepton mass. Figures taken from [129].



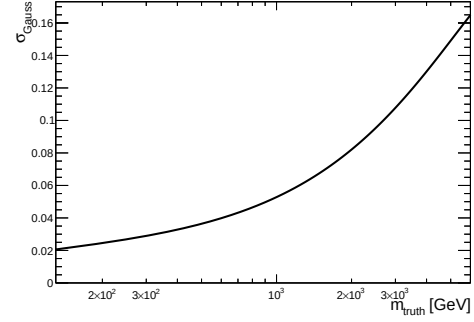
(a) μ_{CB}



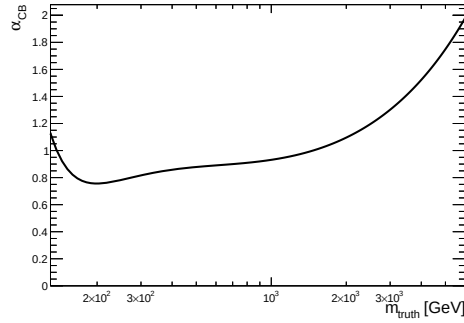
(b) μ_G



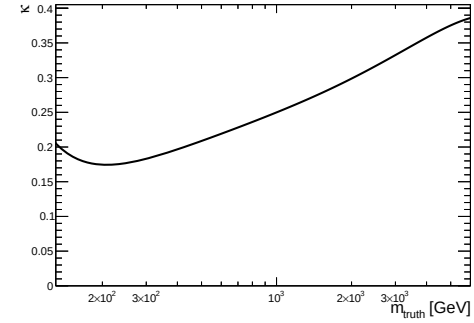
(c) σ_{CB}



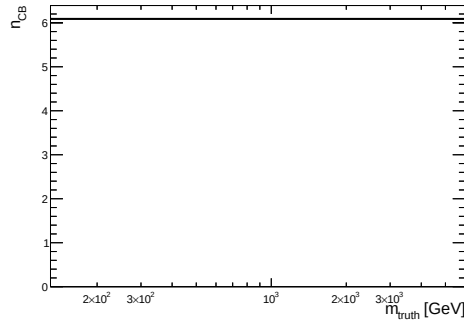
(d) σ_G



(e) α_{CB}



(f) κ



(g) n_{CB}

Figure A.4: Fitted parameterizations of the dimuon channel $m_{\ell\ell}$ response model parameters (equation 6.7) as functions of truth-level dilepton mass. Figures taken from [129].

Parameter M_t -dependence	Fit result
$\mu_{\text{CB}}(M_t) = p_{\mu_{\text{CB}},1} + p_{\mu_{\text{CB}},2}/\ln(M_t) + p_{\mu_{\text{CB}},3} \cdot \ln(M_t) + p_{\mu_{\text{CB}},4} \cdot \ln(M_t)^4$	$p_{\mu_{\text{CB}},1} = 0.13287$ $p_{\mu_{\text{CB}},2} = -0.410663$ $p_{\mu_{\text{CB}},3} = -0.0126743$ $p_{\mu_{\text{CB}},4} = 2.9547 \cdot 10^{-6}$
$\sigma_{\text{CB}}(M_t) = \sqrt{p_{\sigma_{\text{CB}},1}^2 + (p_{\sigma_{\text{CB}},2}/\sqrt{M_t})^2 + (p_{\sigma_{\text{CB}},3}/M_t)^2}$	$p_{\sigma_{\text{CB}},1} = 0.0136624$ $p_{\sigma_{\text{CB}},2} = 0.230678$ $p_{\sigma_{\text{CB}},3} = 1.73254$
$\alpha_{\text{CB}}(M_t) = p_{\alpha_{\text{CB}},1}$	$p_{\alpha_{\text{CB}},1} = 1.59112$
$n_{\text{CB}}(M_t) = p_{n_{\text{CB}},1} + p_{n_{\text{CB}},2} \cdot \exp(-p_{n_{\text{CB}},3} \cdot M_t)$	$p_{n_{\text{CB}},1} = 1.13055$ $p_{n_{\text{CB}},2} = 0.76705$ $p_{n_{\text{CB}},3} = 0.00298312$
$\mu_{\text{G}}(M_t) = p_{\mu_{\text{G}},1} + p_{\mu_{\text{G}},2}/M_t + p_{\mu_{\text{G}},3} \cdot M_t + p_{\mu_{\text{G}},4} \cdot \ln(M_t)^3 + p_{\mu_{\text{G}},5}/M_t^2 + p_{\mu_{\text{G}},6} \cdot M_t^2$	$p_{\mu_{\text{G}},1} = -0.00402708$ $p_{\mu_{\text{G}},2} = 0.814172$ $p_{\mu_{\text{G}},3} = -3.94281 \cdot 10^{-7}$ $p_{\mu_{\text{G}},4} = 7.97076 \cdot 10^{-6}$ $p_{\mu_{\text{G}},5} = -87.6397$ $p_{\mu_{\text{G}},6} = -1.64806 \cdot 10^{-11}$
$\sigma_{\text{G}}(M_t) = \sqrt{p_{\sigma_{\text{G}},1}^2 + (p_{\sigma_{\text{G}},2}/\sqrt{M_t})^2 + (p_{\sigma_{\text{G}},3}/M_t)^2}$	$p_{\sigma_{\text{G}},1} = 0.00553858$ $p_{\sigma_{\text{G}},2} = 0.140909$ $p_{\sigma_{\text{G}},3} = 0.644418$
$\kappa(M_t) = p_{\kappa,1} + p_{\kappa,2} \cdot \exp(-p_{\kappa,3} \cdot M_t) + p_{\kappa,4} \cdot M_t + p_{\kappa,5} \cdot M_t^3$	$p_{\kappa,1} = 0.347003$ $p_{\kappa,2} = 0.135768$ $p_{\kappa,3} = 0.00372497$ $p_{\kappa,4} = -2.2822 \cdot 10^{-5}$ $p_{\kappa,5} = 5.06351 \cdot 10^{-13}$

Table A.3: Functional dependence on the truth mass M_t of the seven $m_{\ell\ell}$ response model parameters (equation 6.7) in the dielectron channel. Note: this is table 28 in the dilepton analysis internal documentation [129].

Parameter M_t -dependence	Fit result
$\mu_{\text{CB}}(M_t) = p_{\mu_{\text{CB}},1} + p_{\mu_{\text{CB}},2}/M_t + p_{\mu_{\text{CB}},3}/M_t^2 + p_{\mu_{\text{CB}},4}/M_t^3 + p_{\mu_{\text{CB}},5} \cdot M_t + p_{\mu_{\text{CB}},6} \cdot M_t^2 + p_{\mu_{\text{CB}},7} \cdot M_t^3$	$p_{\mu_{\text{CB}},1} = -0.0891397$ $p_{\mu_{\text{CB}},2} = 10.6169$ $p_{\mu_{\text{CB}},3} = -951.712$ $p_{\mu_{\text{CB}},4} = 74775.3$ $p_{\mu_{\text{CB}},5} = 5.60192 \cdot 10^{-5}$ $p_{\mu_{\text{CB}},6} = -1.58827 \cdot 10^{-9}$ $p_{\mu_{\text{CB}},7} = -3.81706 \cdot 10^{-13}$
$\sigma_{\text{CB}}(M_t) = p_{\sigma_{\text{CB}},1} + p_{\sigma_{\text{CB}},2}/M_t + p_{\sigma_{\text{CB}},3}/M_t^2 + p_{\sigma_{\text{CB}},4} \cdot M_t + p_{\sigma_{\text{CB}},5} \cdot M_t^2$	$p_{\sigma_{\text{CB}},1} = 0.0836349$ $p_{\sigma_{\text{CB}},2} = -8.98476$ $p_{\sigma_{\text{CB}},3} = 491.19$ $p_{\sigma_{\text{CB}},4} = 5.18068 \cdot 10^{-5}$ $p_{\sigma_{\text{CB}},5} = -3.45042 \cdot 10^{-10}$
$\alpha_{\text{CB}}(M_t) = p_{\alpha_{\text{CB}},1} + p_{\alpha_{\text{CB}},2}/M_t + p_{\alpha_{\text{CB}},3}/M_t^2 + p_{\alpha_{\text{CB}},4}/M_t^3 + p_{\alpha_{\text{CB}},5} \cdot M_t$	$p_{\alpha_{\text{CB}},1} = 0.512577$ $p_{\alpha_{\text{CB}},2} = 252.922$ $p_{\alpha_{\text{CB}},3} = -79337.4$ $p_{\alpha_{\text{CB}},4} = 7.31863 \cdot 10^6$ $p_{\alpha_{\text{CB}},5} = 0.000237883$
$n_{\text{CB}}(M_t) = p_{n_{\text{CB}},1}$	$p_{n_{\text{CB}},1} = 6.08818$
$\mu_{\text{G}}(M_t) = p_{\mu_{\text{G}},1} + p_{\mu_{\text{G}},2} \cdot M_t + p_{\mu_{\text{G}},3} \cdot M_t^2 + p_{\mu_{\text{G}},4} \cdot M_t^3 + p_{\mu_{\text{G}},5} \cdot M_t^4$	$p_{\mu_{\text{G}},1} = -0.00410659$ $p_{\mu_{\text{G}},2} = 2.82352 \cdot 10^{-6}$ $p_{\mu_{\text{G}},3} = -6.264 \cdot 10^{-9}$ $p_{\mu_{\text{G}},4} = 1.25547 \cdot 10^{-12}$ $p_{\mu_{\text{G}},5} = -9.94431 \cdot 10^{-17}$
$\sigma_{\text{G}}(M_t) = p_{\sigma_{\text{G}},1} + p_{\sigma_{\text{G}},2}/M_t + p_{\sigma_{\text{G}},3}/M_t^2 + p_{\sigma_{\text{G}},4} \cdot M_t + p_{\sigma_{\text{G}},5} \cdot M_t^2$	$p_{\sigma_{\text{G}},1} = 0.0214264$ $p_{\sigma_{\text{G}},2} = -0.795058$ $p_{\sigma_{\text{G}},3} = 15.4726$ $p_{\sigma_{\text{G}},4} = 3.38205 \cdot 10^{-5}$ $p_{\sigma_{\text{G}},5} = -1.64984 \cdot 10^{-9}$
$\kappa(M_t) = p_{\kappa,1} + p_{\kappa,2}/M_t + p_{\kappa,3}/M_t^2 + p_{\kappa,4} \cdot M_t + p_{\kappa,5} \cdot M_t^2$	$p_{\kappa,1} = 0.235477$ $p_{\kappa,2} = -31.2227$ $p_{\kappa,3} = 3447.34$ $p_{\kappa,4} = 4.54408 \cdot 10^{-5}$ $p_{\kappa,5} = -3.25374 \cdot 10^{-9}$

Table A.4: Functional dependence on the truth mass M_t of the seven $m_{\ell\ell}$ response model parameters (equation 6.7) in the dimuon channel. Note: this is table 29 in the dilepton analysis internal documentation [129].

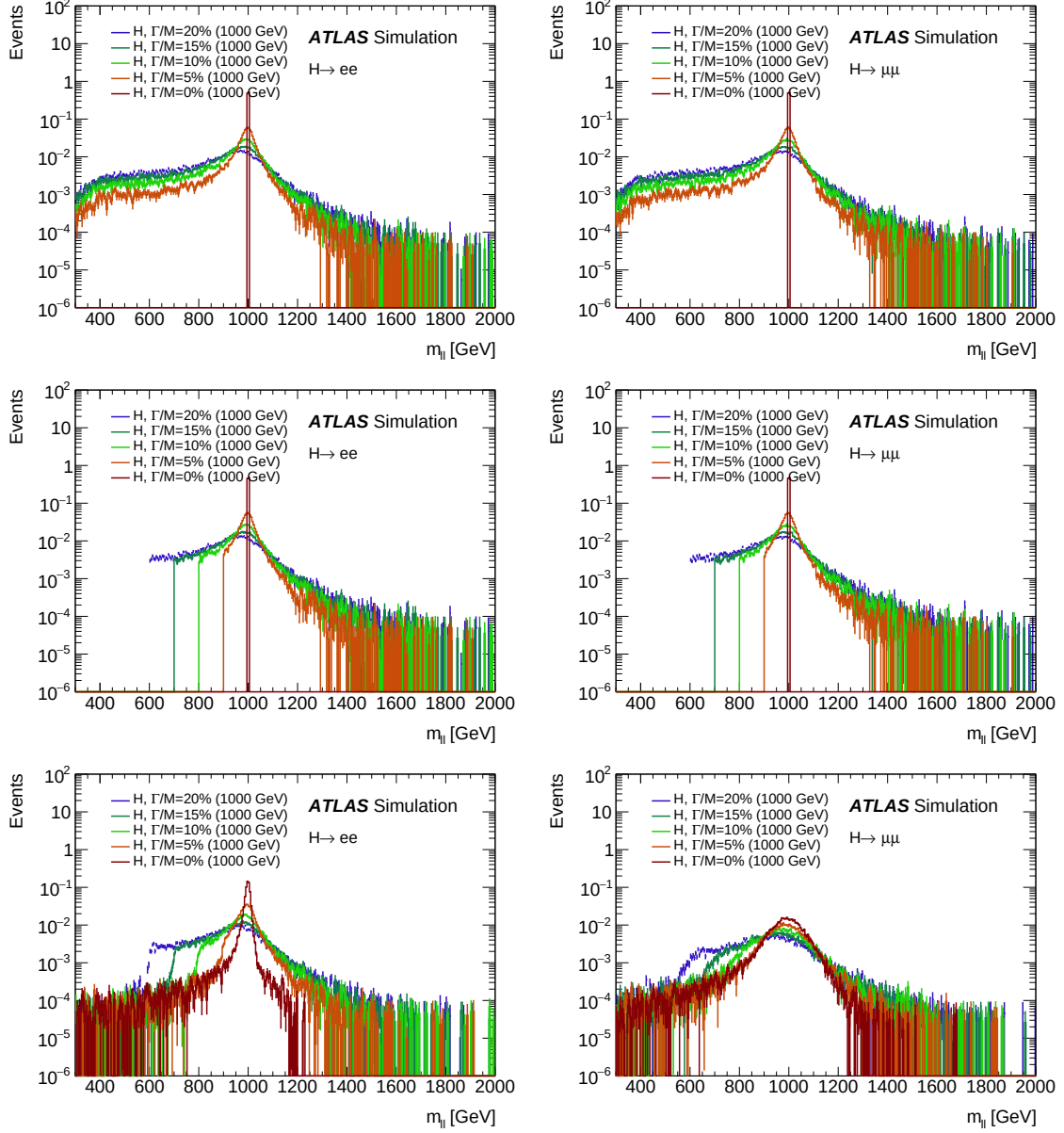


Figure A.5: Invariant mass distributions of 1 TeV MSSM Higgs samples for different relative widths Γ/M_H in the dielectron (left) and dimuon (right) channels at truth-level (top row), after fiducial cuts (middle row), and after reconstruction (bottom row).

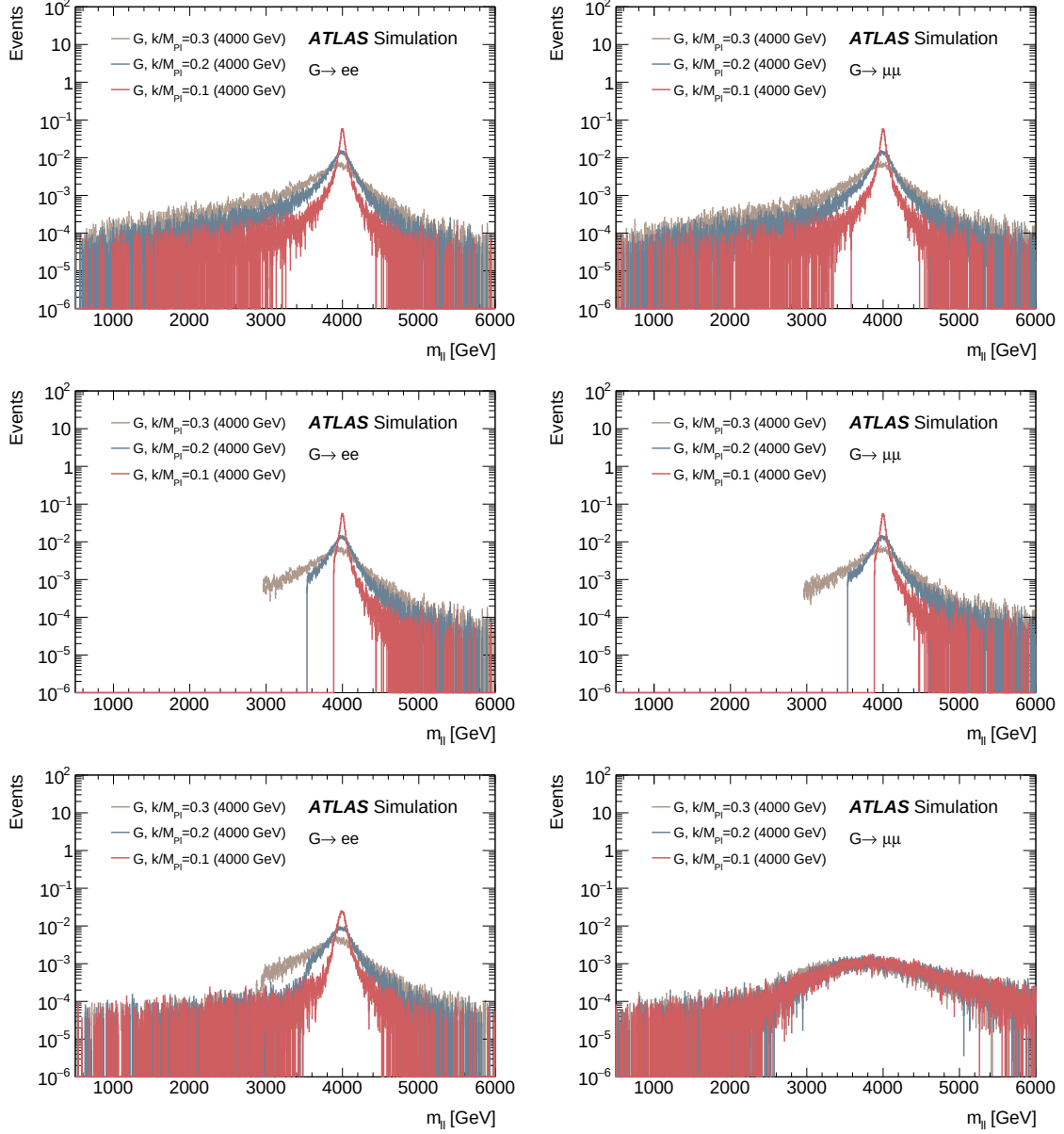


Figure A.6: Invariant mass distributions of 4 TeV Graviton samples for different values of the extra-dimensional curvature parameter k/M_{Pl} in the dielectron (left) and dimuon (right) channels at truth-level (top row), after fiducial cuts (middle row), and after reconstruction (bottom row).

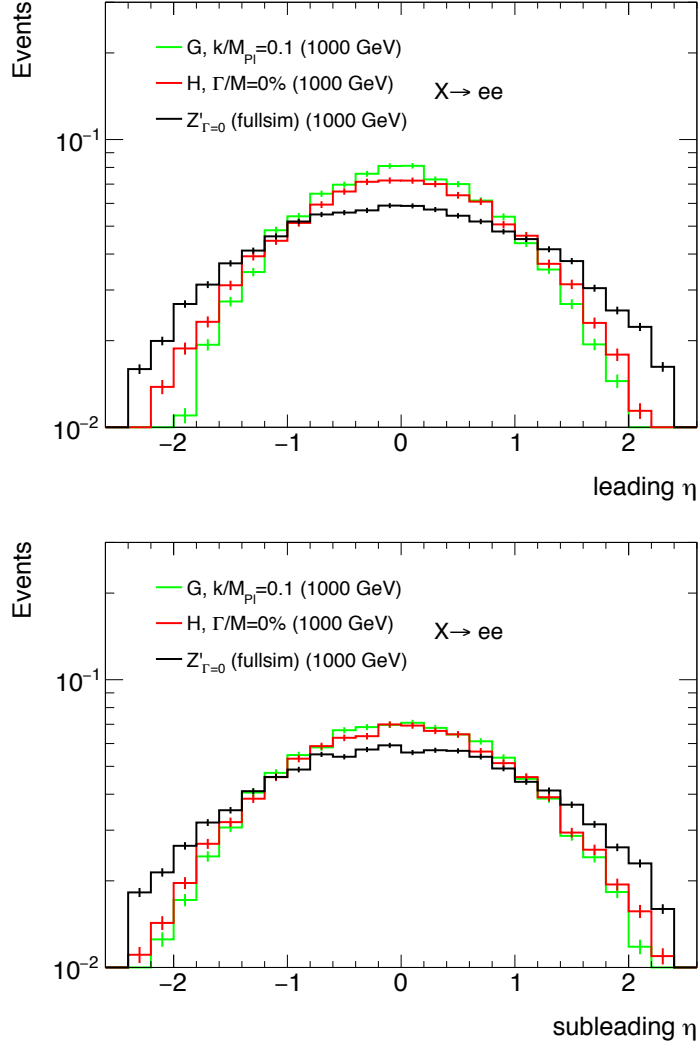


Figure A.7: Comparison of the leading (top) and subleading (bottom) lepton η distributions at truth-level for the 0-width Higgs, 0-width Z' , and a Graviton sample at 1 TeV. The distributions are each normalized to 1. The fiducial cuts (equations 6.13 and 6.14) have been applied. This demonstrates how the spin-1 mediator (Z') produces leptons that are more forward-distributed than in the spin-0 or spin-2 cases.

Appendix B

Additional statistical analysis material

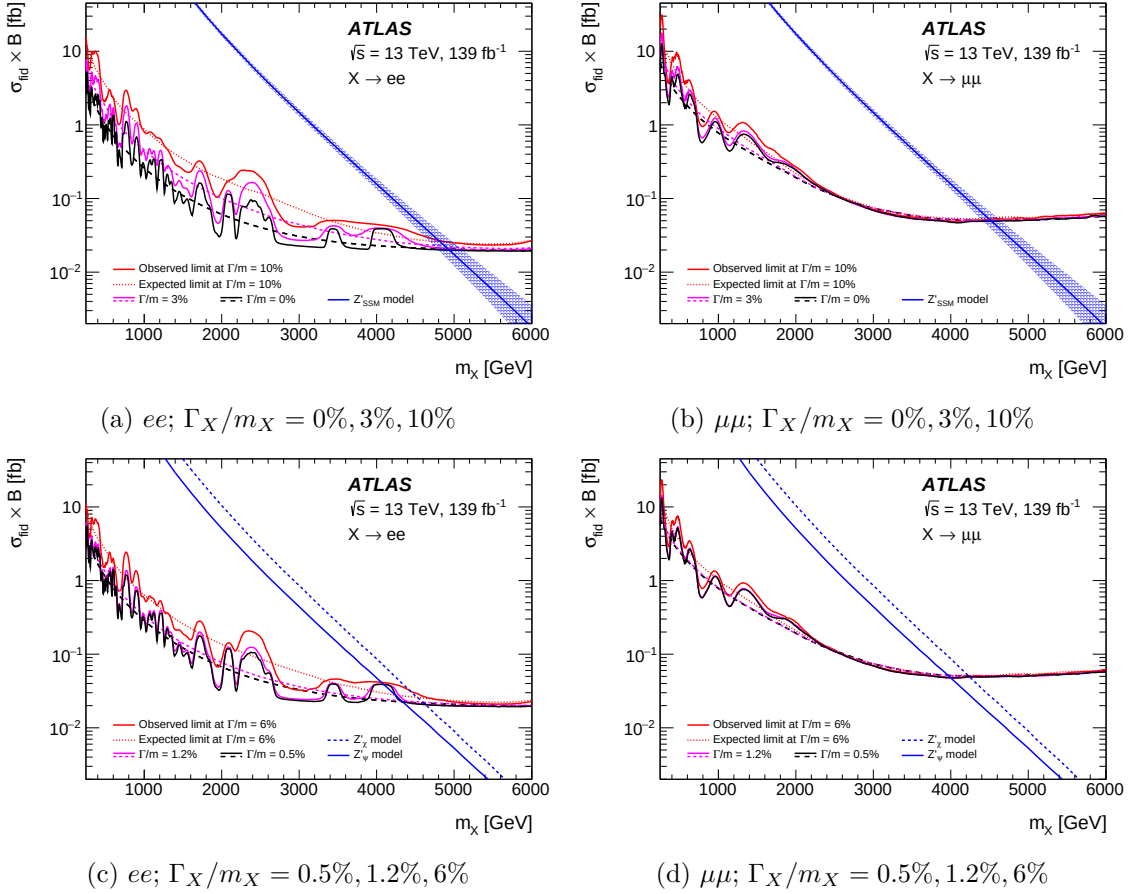


Figure B.1: Observed and expected upper limits at 95% confidence level on the fiducial cross section of a generic signal in the dielectron (left) and dimuon (right) channels. The curves with different colors correspond to six different resonance width assumptions. Overlaid are predicted cross sections within the fiducial volume for three benchmark models: Z'_χ , Z'_ψ , Z'_{SSM} with relative widths of 0.5%, 1.2%, and 3.0%, respectively. The PDF uncertainty on the Z'_{SSM} cross section curve is indicated by the shaded band. Figures taken from [1].

B.1 The look-elsewhere effect

In general, the situation of a mass scan can be categorized as a hypothesis test where one or more parameters exist for the alternate hypothesis but are undefined for the null¹:

$$\mathcal{P}(H_1) = \mathcal{P}(H_0) \cup \{m_X, \Gamma_X, \dots\} \quad (\text{B.1})$$

where $\mathcal{P}$ denotes the parameter set. Note that the parameter of interest is defined under both hypotheses although it is fixed at $\mu = 0$ under H_0 . The additional parameters which are *not* defined under H_0 are, in this case, the signal mass m_X and width Γ_X but could be other ones in general. These can be thought of as nuisance parameters (NPs) which are neither constrained nor free to float in the fit, but fixed during each minimization at a point in an overall scan of values². The search for a new signal is carried out, agnostic toward these additional NPs until the most significant excess is identified at a particular mass and width and “breaks the symmetry”. Similarly, nowhere in the usual p -value calculation of this particular excess does the fact enter that hundreds or even thousands of neighboring mass points did *not* exhibit such a large significance.

A search designed to be sensitive to excesses across a wide signal region irrespective of their mass is more likely to observe significant statistical fluctuations than a search limited to only one signal mass, for example. In the understanding of “significant” as “unlikely to be a statistical fluctuation”, a scan over multiple signal hypotheses simply introduces more probabilistic opportunities for the background-only hypothesis to generate fluctuations. This is called the *look-elsewhere effect* (LEE).

The LEE can be accounted for by reporting the global p -value: the probability of observing a fluctuation *anywhere* across the entire background invariant mass spectrum which has a local p -value at least as small as the one observed in data, generally at a different mass. If f is the probability distribution of the most significant (i.e. minimal) p -value expected from a background-only dataset, then the global p -value can be expressed as follows:

$$\hat{p}_{\text{global}} = P(\min(p_{\text{local}}|_{\text{bkg}}) \leq \hat{p}_{\text{local}}) = \int_0^{\hat{p}_{\text{local}}} f(\min(p_{\text{local}}|_{\text{bkg}})) dp_{\text{local}} \quad (\text{B.2})$$

where hats denote the values associated with the observation in question. Hence the global p -value can be interpreted as a p -value of local p -values. For a given local p -value, the corresponding global p -value increases with the range of the mass scan because the min function in f has more “candidates” to choose from. Similarly, increased granularity of the mass scan drives a larger LEE down to a limit of the experimental mass resolution.

The distribution f can be built up by collecting the maximally-significant fluctuations from an ensemble of pseudo-experiments (toys) generated under the background-only hypothesis. Measuring p -values in Gaussian significance is a one-to-one mapping. The distribution of

¹An early study of this situation as stated is by Davies in [171]

²In principle it is also possible, albeit technically problematic, to allow the mass, width, etc to float if only the most significant excess is of interest. To a small, constrained extent the electron energy scale and resolution systematic variations do this.

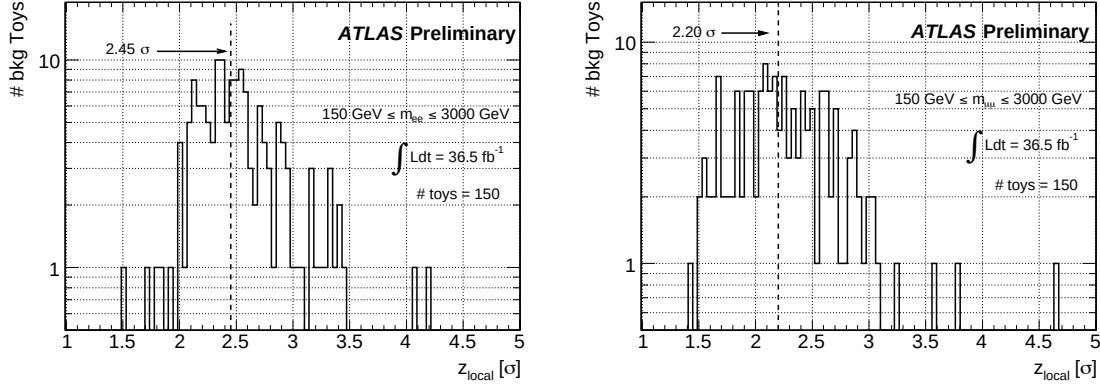


Figure B.2: Distribution of largest local significances observed in (left) dielectron and (right) dimuon channel pseudoexperiments. Median shown by vertical dashed line.

the largest local significance values is shown for 150 toys in the electron and muon channels separately in figure B.2. The statistical analysis configuration used in this figure is from the 36.1 fb^{-1} dilepton publication. In this case the *global significance* can be calculated as the normalized integral on the right tail of this distribution above a given local significance z_{local} :

$$\hat{p}_{\text{global}} = P(\max(p_{\text{local}}|_{\text{bkg}}) \geq \hat{z}_{\text{local}}) = \frac{1}{N} \sum_{i=\hat{i}}^{\infty} n_i, \quad \hat{i} = \text{index}(\hat{z}_{\text{local}}) \quad (\text{B.3})$$

$$\text{and } \hat{z}_{\text{global}} = \text{Erf}^{-1}(1 - 2\hat{p}_{\text{global}}) \quad (\text{B.4})$$

where N is the number of toy datasets scanned and the sum counts up all toy scans whose largest significance was as large or larger than the given $\hat{z}_{\text{local}}$. The median values indicated in figure B.2 suggest that 50% of the background-only processes will produce a fluctuation with significance of at least 2.45σ (2.20σ) in the dielectron (dimuon) channel.

B.1.1 Asymptotic behaviour of trials factor for 1D search

Evidently the sample size of the distribution in figure B.2 is quite limited to $\sim 10^2$ which is a result of the prohibitive computing time required to perform numerous scans (incorporating systematic uncertainties is the expensive part). To calculate global p -values associated with local significance beyond 2.5σ , the 36.1 fb^{-1} dilepton analysis utilized the asymptotic relationship derived by Gross and Vitells [172]³:

$$\hat{p}_{\text{global}} = \underbrace{P(\chi_1^2(q_0) \geq \hat{q}_0)}_{\hat{p}_{\text{local}}} + \langle N(c_0) \rangle e^{-\frac{\hat{q}_0 - c_0}{2}} \quad (\text{B.5})$$

³Shown is the special case for $k = 1$ degrees of freedom, for a single POI.

where $\hat{q}_0$ is the observed value of the LLR test statistic (equation 8.4) and c_0 is some reference threshold. The first term on the right hand side is simply the local p -value in the asymptotic approximation of Wilks' theorem for one degree of freedom. The second term therefore describes the LEE as a term involving the average number of test statistic “upcrossings” $\langle N(c_0) \rangle$, counted when q_0 crosses above the threshold c_0 in a background-only toy. Lastly, the exponential represents a scaling law which allows freedom in the choice of reference value c_0 . Equation B.5 holds in the limit of large significances which conveniently is also where the tails of the toy distributions in figure B.2 become statistically uncertain.

The factor $\langle N(c_0) \rangle$ can be measured to sufficient accuracy by counting upcrossings in a relatively small collection of background-only toys. To illustrate, figure B.3a shows how 12 threshold upcrossings are counted for a q_0 scan on a typical toy dataset. This one example corresponds to a single entry in the bin containing $\# \text{ upcrossings} = 12$ in the distribution shown in figure B.3b. In this case $\langle N(c_0) \rangle$ of equation B.5 equals 11.89.

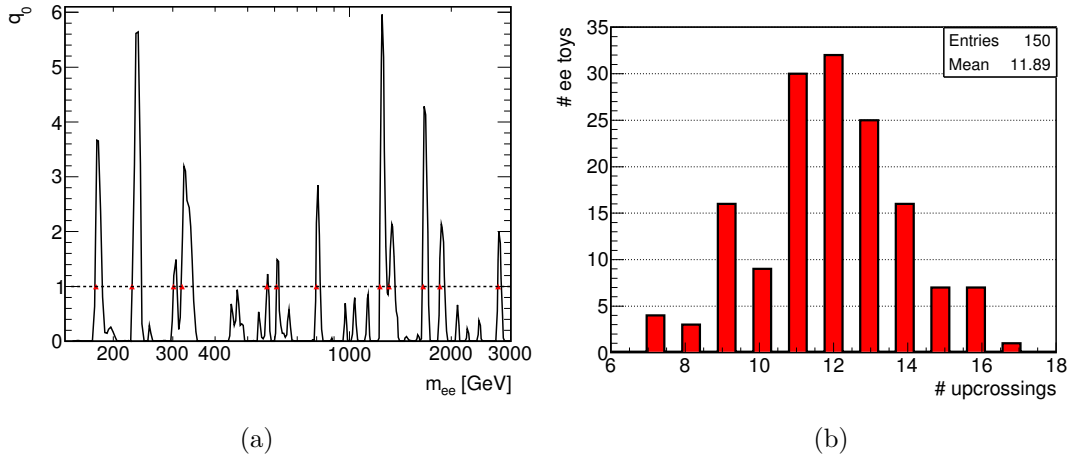


Figure B.3: Illustration of test statistic upcrossings. (a) shows an example test statistic scan on a background-only toy dataset in the electron channel where red triangles locate 12 upcrossings of the test statistic q_0 above a reference threshold of $c_0 = 1.0$. (b) shows the distribution in number of upcrossings of the test statistic for the same c_0 in 150 electron channel toys.

To demonstrate the validity of equation B.5 in the large-significance limit, an ensemble of 50k background-only toys were scanned without systematic uncertainties (to make them quicker). With the increased statistical precision in the large-significance tails, equation B.3 could be used up to $\hat{z}_{\text{local}} = 5\sigma$. The test statistic upcrossings were also counted for each toy for use in equation B.5. The comparison of the two results is shown for the range of $\hat{z}_{\text{local}}$ values on the x -axis of figure B.4. The asymptotic relationship agrees to within less than 0.2σ with the tail integral method beginning at $\hat{z}_{\text{local}} = 3\sigma$ and improves towards higher values, as expected.

Based on the validation of the asymptotic approach, for the 36.1 fb^{-1} analysis a hybrid of equation B.3 and equation B.5 was used. The first method is used as far as the limited statistical precision in figure B.2 allows. Then a linear interpolation is used to interpolate across roughly 0.5σ to where the asymptotic description (B.5) begins. Using the same col-

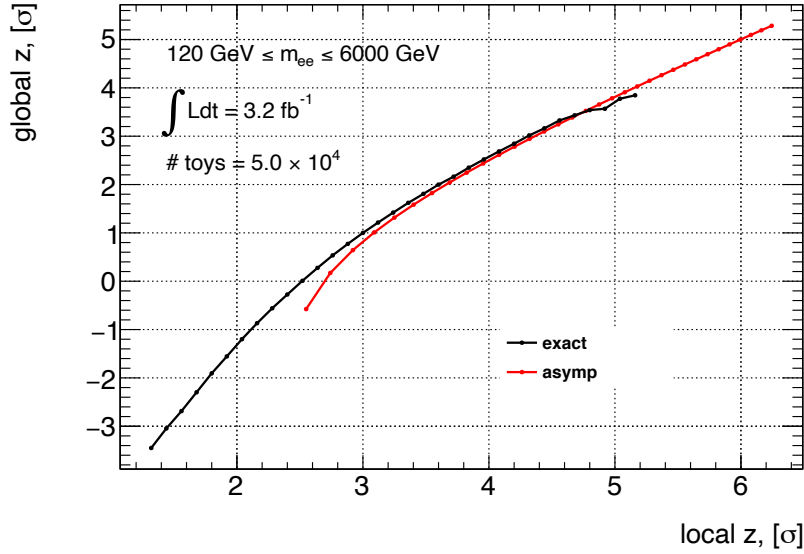


Figure B.4: Comparison of local-to-global significance map derived using the asymptotic relationship in equation B.5 and from pseudoexperiments (equation B.3).

lection of 150 pseudo experiments, the average number of upcrossings above a $c_0 = 1.0$ threshold is measured as the mean of distributions such as the one in figure B.3b.

Finally, figure B.5 shows the global p-value phrased in Gaussian significance plotted for a local significance range. The y -intercepts corresponds to the medians of the distributions in figure B.2. Negative values of global significance indicate data which were less signal-like than is typically expected of the background. Because of the LEE, observations in the combined dilepton channel will need to exceed local significance of 4.4σ to meet the conventional 3σ global significance threshold required for claims of evidence. The effect is essentially the same in the dielectron channel alone since its much better $m_{\ell\ell}$ resolution makes it contribute much more to the combined LEE than the dimuon channel.

In the full Run-2 dilepton analysis the LEE was estimated solely using equation B.5 and using only the observed dataset as a single “pseudoexperiment” for counting upcrossings. This is consistent with a recommendation by the ATLAS Statistics Forum for search results with an observed local significance of $2 - 3\sigma$. However, depending on how large the scan is (with respect to the experimental resolution), the total number of upcrossings may be small, leading to uncertainty in $N(c_0)$. In this case because of the large LEE (i.e. lots of fluctuations) the uncertainty is acceptable, and the resulting estimate of global significances for the largest excesses are listed in table 8.4.

B.1.2 Asymptotic behaviour of trials factor for 2D search

In equation B.1 the signal width Γ_X was also listed for an example of a nuisance parameter which is defined for the alternate hypothesis but not for the null. Thus, performing a scan over Γ_X also introduces a LEE. In the dilepton analysis, a 2D scan is performed over both m_X and Γ_X , hence engaging a two-fold LEE. Quantifying the 2D LEE can be done in

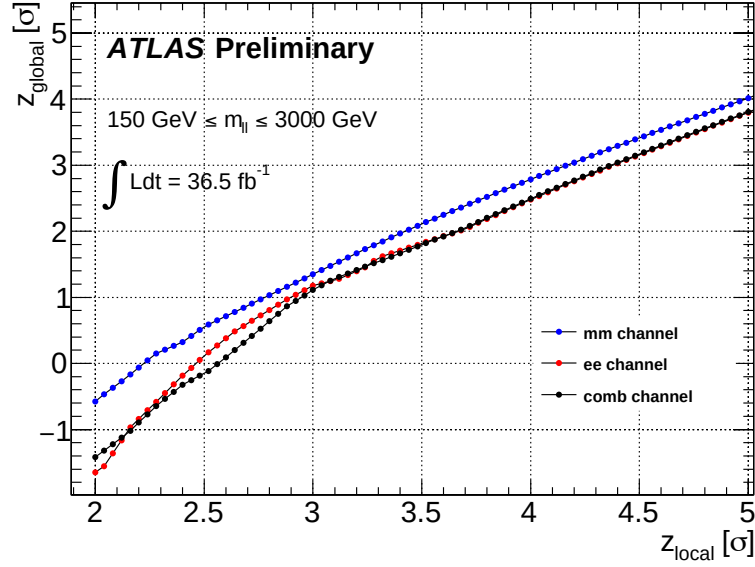


Figure B.5: Distribution of highest local significances observed in (left) dielectron and (right) dimuon channel pseudo experiments. Median shown by vertical dashed line.

exactly the same way as described above by scanning ensembles of background-only toy datasets. But if the computational time for only the mass scan was already prohibitive, running the width scan is slower by a factor of the number of width points. Therefore the follow-up publication by Gross and Vitells [173] treating the multi-dimensional search case is all the more helpful.

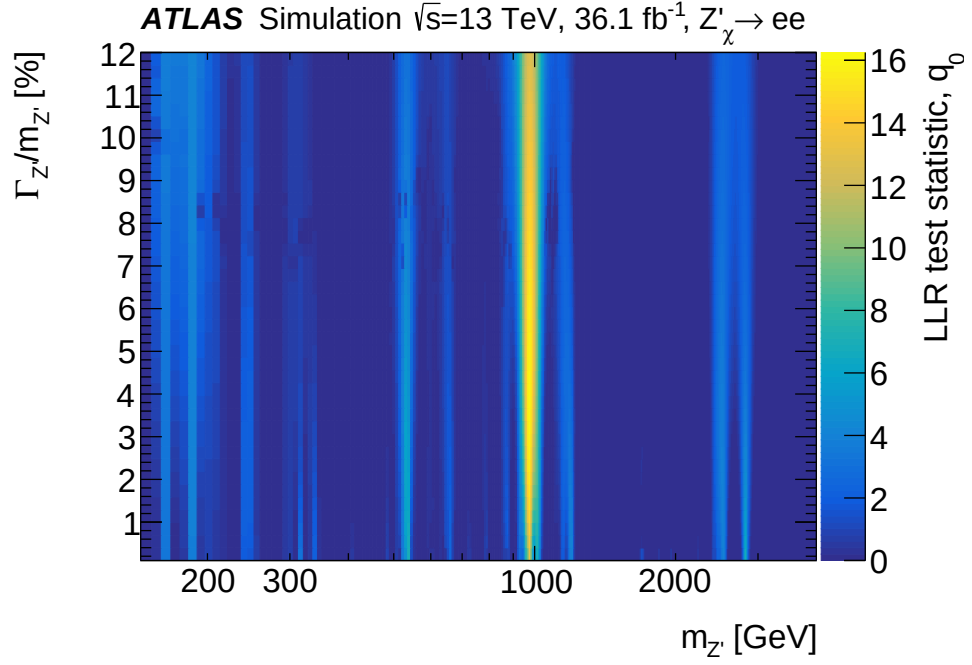
The test statistic in a 2D scan of a background-only toy dataset constitutes a random field in the two search parameters. Upcrossings therefore appear as “islands” instead of the peaks shown in figure B.3a. An example of the 2D analogue is shown in figure B.6. In subfigure B.6b, the shaded regions where the test statistic is above threshold is referred to as the *excursion set*. Since signal shapes are relatively similar across the range of widths considered compared to the range of masses, the excursion set has more correlation along the vertical direction. Again, this study is in the context of the 36.1 fb^{-1} dilepton analysis.

The asymptotic expression for the global p -value in a 2D search is related to a topological property of the excursion set known as the *Euler characteristic* $\phi()$, calculated as shown in figure B.7⁴. In fact, according to [173],

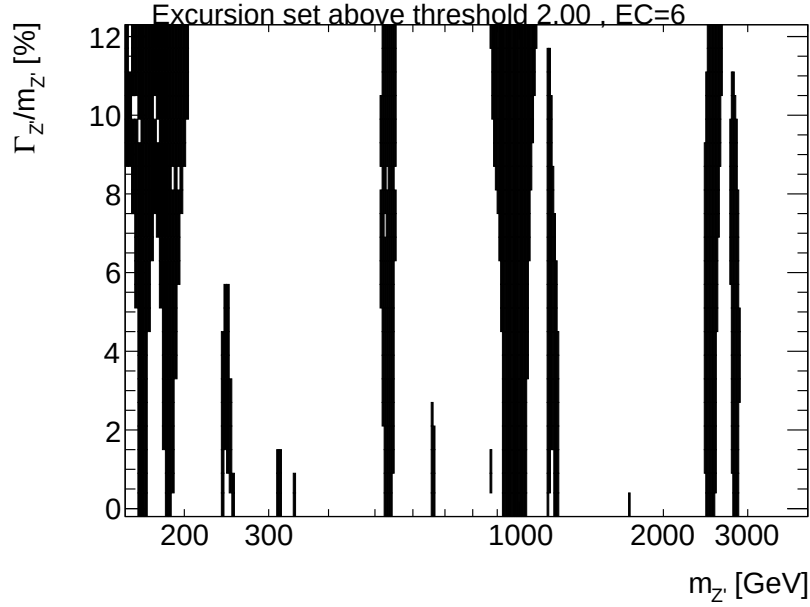
$$\hat{p}_{\text{global}} \simeq \langle \phi(\hat{q}_0) \rangle = \underbrace{P(\chi_1^2(q_0) \geq \hat{q}_0)}_{\hat{p}_{\text{local}}} + (N_1 + \sqrt{\hat{q}_0} N_2) e^{-\hat{q}_0/2} \quad (\text{B.6})$$

where $\langle \rangle$ denotes the average. The first approximate equality is that in the limit of small (i.e. significant) local p -values, the global p -value tends toward the average value of the Euler

⁴The Euler characteristic turns out to be on average the number of “islands” minus the number of “lakes” (i.e. holes in islands), as is nearly the case in B.7.



(a)



(b)

Figure B.6: Example of a 2D scan of the test statistic on a background-only pseudoexperiment. (b) shows the excursion set comprised of regions where the test statistic lies above the reference threshold of $c = 2$ in this case.

characteristic of the corresponding excursion set. This is analogous to the behaviour of the

asymptotic relationship (B.5) in 1D. However, as with the 1D case, this can be determined using the more frequent upcrossings occurring at a lower threshold. In the second equality of (B.6) the average Euler characteristic is related to two unknown coefficients N_1 and N_2 . Since they are assumed to not vary with $\hat{q}_0$, they can be determined by measuring $\langle \phi(\hat{q}_0) \rangle$ for the excursion sets at two different thresholds, say $\hat{q}_0 = c_0$ and c_1 , and solving the linear system of equations.

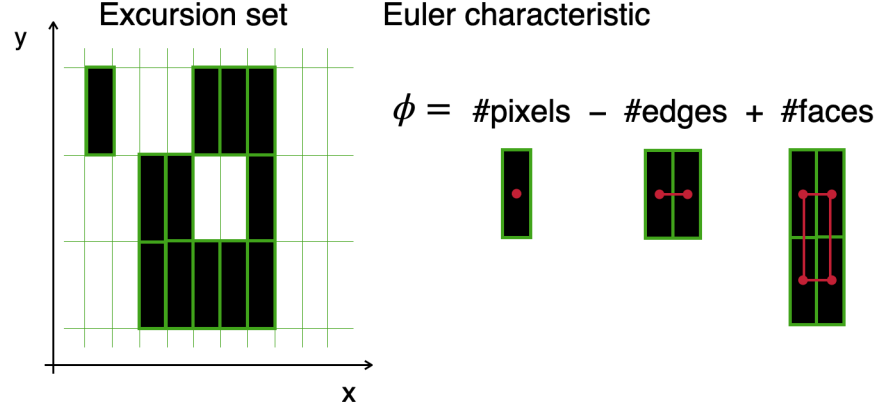


Figure B.7: Illustration showing how the Euler characteristic is calculated from the excursion set. In the example shown $\phi = 12 - (7 + 4) + 1 = 0$.

This procedure was followed using 2D scans on 1000 background-only toy datasets. A variety of choices for c_0 and c_1 were tried, and for each choice equation B.6 used to isolate N_1 and N_2 . It was found that choosing $c_0 = 0$ led to reasonable results. This and seven different choices for c_1 produced the relationships between global and local significance plotted in figure B.8. Also shown is the map resulting from using the “brute force” approach of integrating the tails of the toys’ q_0 distributions (equation B.3). It is clear that the choice of thresholds appears to have an impact on the resulting relationship, whereas this is not indicated in [173]. Also the agreement with the fully toys-based calculation is not as persuasive as in the 1D case shown in figure B.4. This study has been only investigative and was not yet used in any publication. There is clearly more to be understood, but these results already show a degree of feasibility of the asymptotic method for a 2D scan which may be applicable in future analysis iterations.

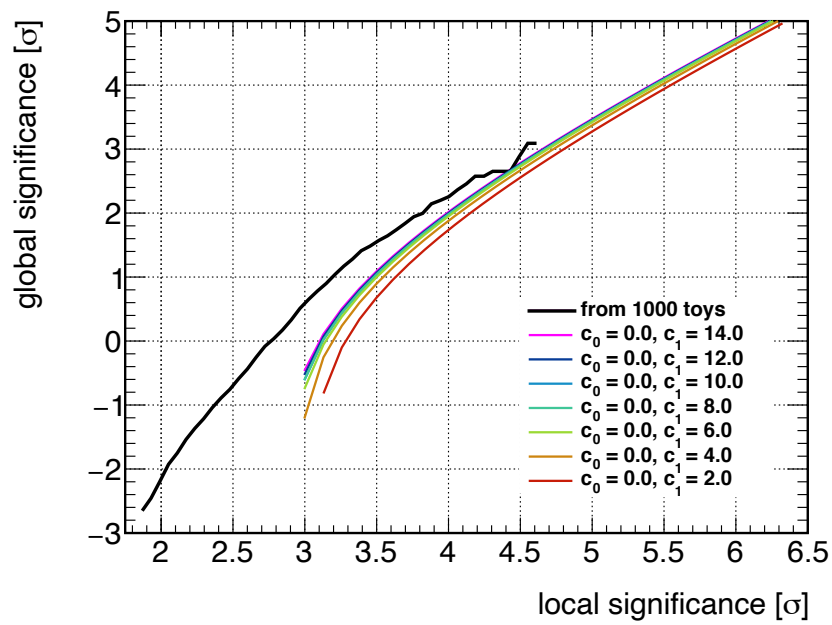


Figure B.8: Comparison of local-to-global significance map derived using the asymptotic relationship in equation B.6 and from pseudoexperiments. A range of choices for the $\hat{q}_0$ thresholds c_0 and c_1 are shown with some variation.

B.2 Comparison of q_μ and q_0 distributions for upper limits

This section shows how using either the q_0 (equation 8.14) and q_μ (equation 8.16) test statistic leads to the same upper limit. I am not aware of this equivalence being shown in previous publications but also have not intentionally looked. In section 8.1.2 the intuition for determining an upper limit on the signal strength μ was built from q_0 distributions which shift progressively to larger values and therefore correspond to shrinking p_{s+b} cut off by a stationary $\hat{q}_0$. When using instead the q_μ test statistic, the dynamics are reversed but preserve the same effect. Figure B.9 shows the q_μ analogues of figure 8.2 as follows:

- figure 8.2a $\leftrightarrow$ figure B.9a, figure B.9b, figure B.9c
- figure 8.2b $\leftrightarrow$ figure B.9d

Note that a two-sided version of q_μ is used for illustrative purposes.

$$q_\mu \equiv \begin{cases} -2 \ln \lambda(\mu), & \hat{\mu} \leq \mu \\ +2 \ln \lambda(\mu), & \hat{\mu} > \mu \end{cases} \quad (\text{B.7})$$

What shifts in this case is the $\hat{q}_\mu$ for each hypothesis in the scan, while the distribution of q_μ actually remains stationary (it is the mirrored chi-square distributions in each subfigure of B.9). Similarly, while in the q_0 case the background-only distribution remained stationary for different μ , in the q_μ case it shifts to larger values. In fact the test statistic distributions for background-only toys under q_μ are the same as the distributions for toys with $\mu' = \mu$ under q_0 .

To check that the p -values needed to compute the CL_s are the same for either choice of test statistic, consider again the example of an observed $\hat{q}_0$ which was shown in figure 8.2b. Using the Wald expression of equation 8.12, such an observation corresponds to

$$\hat{q}_0 = 3 = \frac{\hat{\mu}^2}{(5.278)^2} \implies \hat{\mu} = 9.142 \quad (\text{B.8})$$

where the denominator is the value of $\sigma_{(\mu=0)}$ measured as the standard deviation of the distribution in figure B.10a.

Now, to convert this observation into a q_μ value the corresponding asymptotic expression can be used:

$$q_\mu \simeq \frac{(\mu - \hat{\mu})^2}{\sigma_\mu^2} \quad (\text{B.9})$$

Substituting in the value obtained in (B.8) and the standard deviation of the underlying $\hat{\mu}$ measurement at $\mu' = 16$ shown in figure B.10b,

$$\therefore \hat{q}_{(\mu=16)} = \frac{(16 - 9.142)^2}{(5.58)^2} = 1.511 \quad (\text{B.10})$$

This value of $\hat{q}_\mu$ can now be overlaid with the q_μ distribution as shown in figure B.9d to determine p_{s+b} and p_b . They are shown in table B.1 in comparison to those obtained using

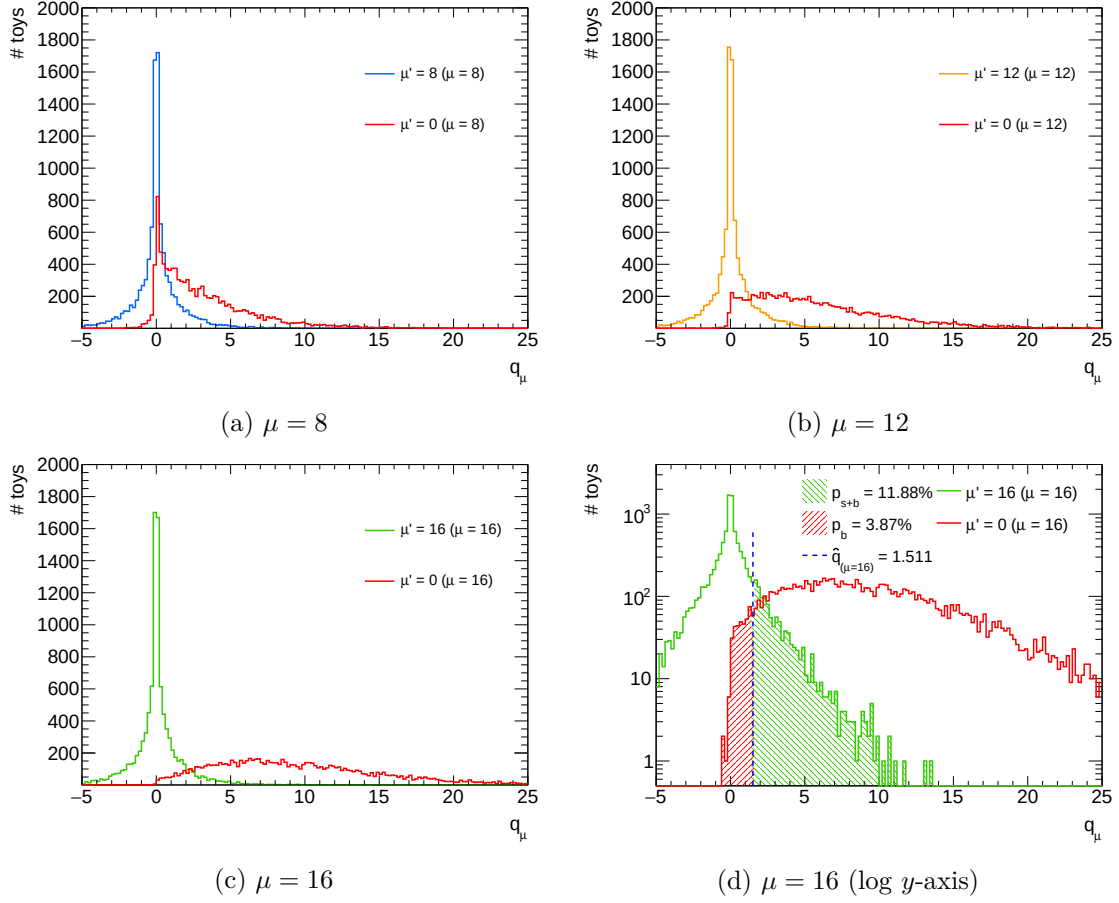


Figure B.9: Distributions of the LLR test statistic q_μ for 10k toys generated for different hypothesized values of the signal strength. Each subfigure shows one distribution for 10k background-only toys ($\mu' = 0$) and another for 10k toys with injected signal matching the hypothesized value ($\mu' = \mu$). In each case the q_μ distribution for toys with the injected signal takes the form of a two-sided χ_1^2 distribution. The background-only distribution is shown in red and shifts upward with μ . Each toy dataset (with 200k events) was both generated and then fit at a fixed mass using a simple $s + b$ model based on an exponential background.

the q_0 test statistic. Dealing with q_μ results in p -values that agree with those obtained using the q_0 test statistic.

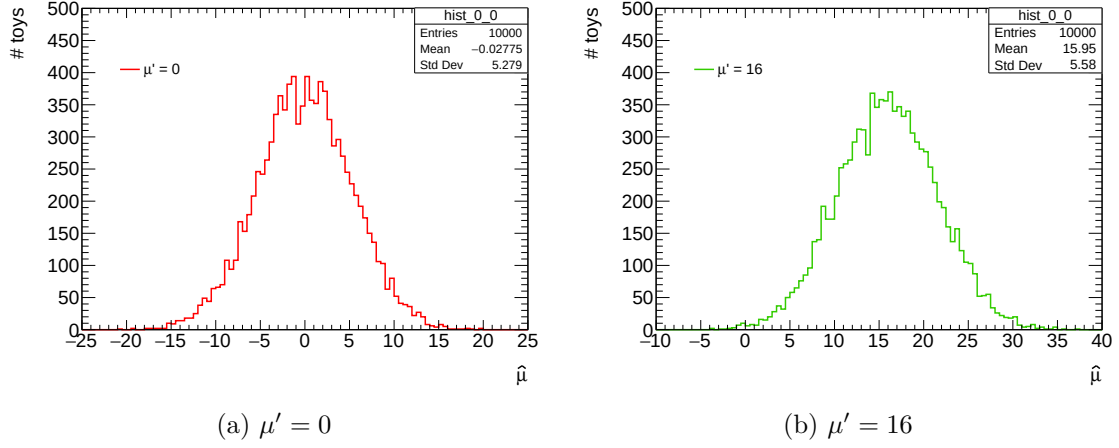


Figure B.10: Distributions of best-fit $\hat{\mu}$ for (a) background-only toys and (b) toys with injected $\mu' = 16$. Mean and standard deviation shown in upper right corner. Each toy dataset (with 200k events) was both generated and then fit at a fixed mass using a simple $s + b$ model based on an exponential background.

test statistic for $\hat{\mu} = 9.142$	p_b	p_{s+b}	CL_s
$\hat{q}_0 = 3.0$	$3.83\% = 383/10000$	$12.14\% = 1214/10000$	0.1262
$\hat{q}_\mu = 1.511$	$3.87\% = 387/10000$	$11.88\% = 1188/10000$	0.1236

Table B.1: Comparison of p -values obtained using $\hat{q}_0$ and $\hat{q}_\mu$ for an example observation of $\hat{\mu} = 9.142$.

Appendix C

Sliding window fit crosscheck

Note: the sections in this appendix have been adapted from appendices N and Q of the dilepton analysis internal documentation [129].

C.1 Overview

The large dataset size which presents a challenge for a fully Monte Carlo based background estimate opens up an opportunity for a data-driven background estimate that is sufficiently constrained. In the context of the dilepton analysis, the data-driven background estimate is derived from a parametric, monotonic background function which is fit to the data in a certain range. Note that in general there are alternatives to using a parametric function such as Gaussian process regression [168] and functional decomposition [174]. If the range of the background fit is too small, it will not be sufficiently constrained in the case of a true signal. This results in a loss of sensitivity as the background function’s parameters are free enough to pull and “fit out” the signal. On the other hand, achieving a good fit to the data in the entire range is potentially challenging given the small statistical uncertainty in the data and the presence of multiple background components within the inclusive spectrum. This possibility is checked using the spurious signal test for the global fit configuration on the full Run-2 dataset (chapter 8.2.3). However it is expected that the same configuration will not be able to properly model larger datasets in the future. For example already in this analysis round it was seen that including the lowest part of the spectrum, down to 130 GeV substantially increased the spurious signal of the global fit, presumably due to fit mismodelling. Therefore, what follows in this chapter is a more general approach which attempts to optimize the sensitivity and robustness of a data-driven background estimate by fitting the data in ranges of the observable called *windows*. Each signal pole mass hypothesis corresponds to a separate window which generally overlaps with its neighbors. In this way the width and position of the window with respect to the signal mass can be optimized instead of arbitrarily choosing to fit the full mass spectrum.

C.2 Method

The simplest incarnation of a “Sliding Window Fit”, or SWiFt involves fitting the background function to the data in a window of fixed width and centred on the pole mass hypothesis. In this set up, the window will “slide” across the invariant mass spectrum in the scan over pole mass hypotheses. This type of approach is quite established in dijet searches within ATLAS [175, 176] where it is used to stitch together a semi-rigid background template from data. However, the method presented here is novel in that it combines the robustness offered by the sliding window fits while completely avoiding the problems associated with stitching them together.

The width and position of the window relative to the signal mass need not be static, and section C.5 describes a procedure for optimizing the window configuration as a function of signal mass. Figure C.1 illustrates the windows obtained from the optimization done separately in the electron and muon channels. When performing a mass scan on data, the sliding window may eventually encounter a region where it does not encompass enough data to be sufficiently constrained thereby. In this case, an “anchor” point can be chosen as the maximum value for the left edge of the sliding window.

The requirement for a monotonic background function implies that a suitable function fitting in a restricted range should have less parameters to provide a rigid estimate in the presence of a signal. In our implementation, the following background function with four parameters (including normalization) produces stable fits across the invariant mass spectrum:

$$f(m_{\ell\ell}) = \frac{1}{(M_Z^2 - m_{\ell\ell}^2)^2 + (m_{\ell\ell}^2 * \Gamma_Z/M_Z)^2} \times p_0 \left(1 - \frac{m_{\ell\ell}}{13}\right)^{p_1} \left(\frac{m_{\ell\ell}}{13}\right)^{p_2 + p_3 * \log(m_{\ell\ell}/13)} \quad (\text{C.1})$$

where the first factor is a relativistic Breit-Wigner shape with parameters $M_Z = 0.0911876$ TeV and $\Gamma_Z = 0.0024952$ fixed to the Particle Data Group values for the Z boson resonance. The choice of function is driven primarily by its ability to model the low-mass range of the invariant mass spectrum where the shoulder of the Z peak is challenging to describe.

The background function C.1 is then added to the generic signal model obtained as described in section 6.6.4 (the optimization was based on the zero-width signal). Fitting the data with this $s + b$ probability density function is the basis for search and limit-setting scans with the same configuration as those in section 8.2.2.

C.3 RooSwift

The SWiFt search, optimization, and limit-setting are implemented in a RooFit/RooStats-based framework called RooSwift [177] developed within the dilepton group. RooSwift was preceded by the original SWiFt package used in JDM analyses but has the advantage of utilizing the RooWorkspace class which is easily portable and recognized within the ATLAS community.

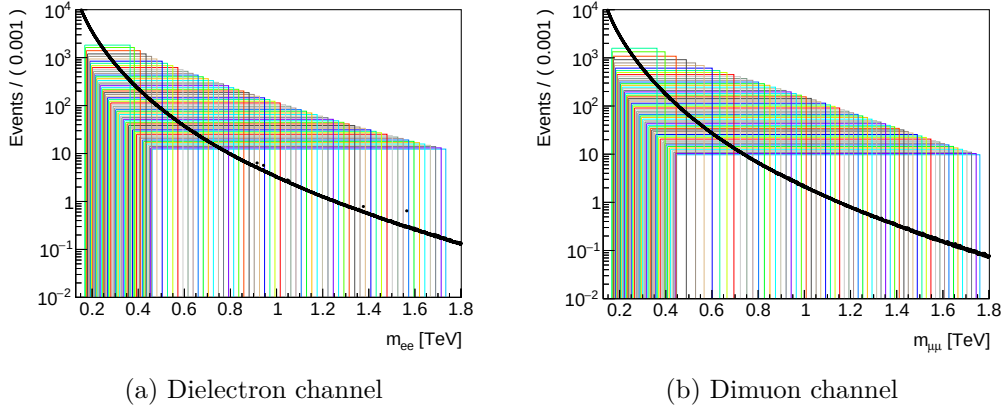


Figure C.1: An illustration of the window ranges used in the SWiFt procedure in the ee (left) and $\mu\mu$ (right) channel shown as colored rectangles overlaying the smoothed MC background template. These are the window configurations produced by the optimization procedure described in section C.5.

C.4 Uncertainties

The finite dataset size introduces a statistical uncertainty into the data-driven results. This uncertainty is taken into account implicitly through the maximum-likelihood fit which allows the background fit parameters to vary within the assumed Poisson uncertainties on each bin of data.

The only systematic uncertainty applied to the data-driven background estimate is the spurious signal systematic. The spurious signal systematic uncertainty quantifies the degree to which background mismodelling distorts the true signal extraction. A poor background fit will deviate from the data so that even in a background-only sample, a signal may be induced by the discrepancy of the fit near the pole mass. The optimal window configuration and function choice will reduce the amount of spurious signal but not eliminate it entirely.

To estimate the size of the spurious signal uncertainty a smooth background $m_{\ell\ell}$ spectrum is constructed by adding the TF-smoothed MC templates for Drell-Yan and $t\bar{t}$ described in section 6.6.3 to the next-dominant templates and the QCD template scaled up from [9]. This background template is normalized to 140fb^{-1} to represent the target Run-2 integrated luminosity. The spurious signal is then measured in a mass scan on the smoothed MC background template as if extracting the presence of a real signal. The absolute value of the resulting signal yield ($|N_{SS}|$) as a function of mass is shown in figure C.2 and is then modelled using the following smooth function to eliminate fluctuations:

$$|N_{ss}|(x) = p_0(1-x)^{p_1}x^{p_2+p_3\log x+p_4(\log x)^2} + p_5x^{p_6} \quad , \quad x = M_{Z'}/13 \quad (\text{C.2})$$

In order to obtain a conservative estimate, only a subset of the $|N_{SS}|$ points are fitted which are selected to represent the envelope of the distribution. The fitted parameter values of Equation C.2 corresponding to the fits shown in figure C.2 have the following values in the ee and $\mu\mu$ channels:

channel	ee	$\mu\mu$
$p_0 =$	0.100209	0.497362
$p_1 =$	22.2081	21.6417
$p_2 =$	-1.76909	-2.09654
$p_3 =$	0.435344	-0.00915778
$p_4 =$	0.117314	0.0448331
$p_5 =$	0.0126385	0.0083518
$p_6 =$	-1.21579	-0.243600

(C.3)

Finally, this model of the spurious signal yield is taken to be the standard deviation of a gaussian constraint term profiling the spurious signal systematic in the likelihood as in equation C.4. Effectively, the background-only fit is allowed to pull the spurious signal strength nuisance parameter (θ_{SS}) to enhance the fit at the cost of a likelihood penalty meted by the constraint term.

$$\mathcal{L} = \prod_{i \in \text{bins}} \text{Prob}(N_i | N_{\text{bkg}} f_{\text{bkg}} + (\mu N_{\text{sig}} + \theta_{\text{ss}}) f_{\text{sig}}) \times \text{Gaus}(\theta_{\text{ss}}, 0, |N_{\text{ss}}|) \quad (\text{C.4})$$

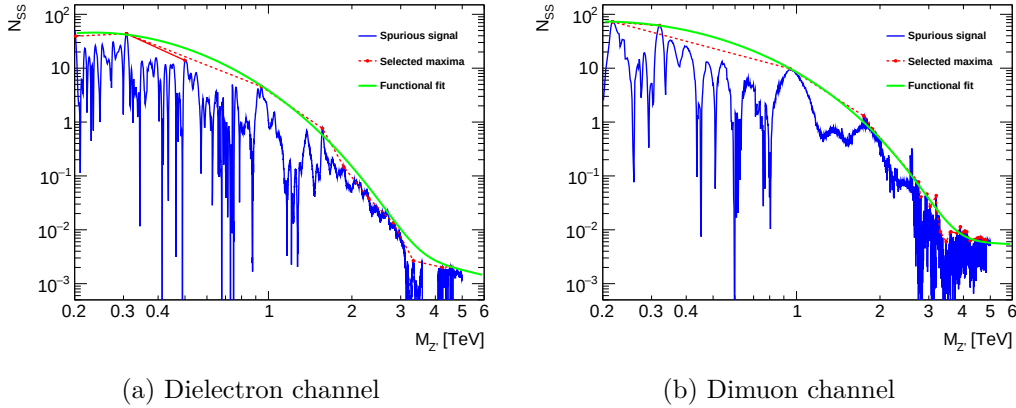


Figure C.2: The measured spurious signal for a zero-width signal extracted from the smoothed MC background template normalized to 140fb^{-1} . The green curve indicates the functional fit of equation C.2 to the selected maxima identified with red markers.

C.5 Optimization

As mentioned in section C.1, the sensitivity of the SWiFt procedure depends on the width of the sliding window. While increasing width will, for an ideal background function lead to increasing sensitivity, for an imperfect function with a fixed number of parameters it also leads to increased background mismodelling. When incorporated through the spurious signal systematic as described in section C.4, the background mismodelling accompanying wide windows is expected to curb the effect of increasing sensitivity. These two effects suggest that in principle an optimal window configuration should exist for a given function as a function of pole mass.

In addition to the window width, only one more parameter is required to form a basis for the space of all window configurations, and an obvious choice is the window *centrality*, defined as the difference between the geometric center of the window and the pole mass being tested, divided by the width of the window. In the width (w) and centrality (c) basis, the left and right edges of windows are parameterized as follows:

$$m_{\ell\ell}^{\text{left}} = M_{Z'} - w \times (1 - c)/2 \quad (\text{C.5})$$

$$m_{\ell\ell}^{\text{right}} = M_{Z'} + w \times (1 + c)/2 \quad (\text{C.6})$$

The output of the optimization procedure is a model of the window configuration parameterized through w and m as functions of pole mass $M_{Z'}$ which optimizes the sensitivity of the SWiFt approach for a given background function. Strictly speaking, the optimization is also limited to a particular signal model (in this case the zero-width shape) such that when considering wide signals, w and m would also be functions of the resonance width $\Gamma_{Z'}$.

The sensitivity of the SWiFt procedure is quantified using the expected significance: the value of the Gaussian significance measured on an Asimov dataset with an injected signal. In order to avoid bias from the background mismodelling, the background used in the Asimov dataset is taken from the smooth MC background template. Furthermore, for the optimization procedure the size of the injected signal is calculated as a multiple of the background integral underneath the signal within a central range four times the absolute invariant mass resolution. The value itself is not relevant, only that it remains constant for each window configuration at a given pole mass.

The background mismodelling is incorporated into the fit producing the expected significance via the spurious signal systematic as described in section C.4. However, within the optimization procedure there is no way to first extract a spurious signal envelope since there is no *a priori* choice of window configuration. Instead, the spurious signal for a given choice of w and c at a given pole mass is estimated as directly the signal extracted from a fit to the smooth MC template at that pole mass. The downside of this approach is that it is sensitive to fluctuations within the smooth MC template which the envelope approach avoids. But because the optimization is performed as a function of pole mass, any given fluctuation should impact all windows configurations roughly the same. Though the absolute value of the spurious signal may be offset by fluctuations, the different window choices will still result in a relative difference that provides the gradient for the optimization.

To set up a comparison of a reasonable set of window configurations, the a few steps are performed within a nested “for loop” over window width w , window centrality c , and pole mass $M_{Z'}$:

for w in steps of 0.020 TeV between 0.020 TeV and 0.420 TeV

for c in steps of 0.1 between -0.8 and $+0.8$

for $M_{Z'}$ in steps of 0.005 TeV between 0.200 TeV and 0.500 TeV
and steps of 0.010 TeV between 0.500 TeV and 1.100 TeV

1. Fit signal normalization from smoothed MC background

2. Create Asimov dataset with injected signal
3. Measure significance of Asimov dataset (2) using the measurement in (1) as the standard deviation on the spurious signal nuisance parameter

Then, for each pole mass, the expected significance values of these windows are “cleaned up” as described below before being plotted on a two-dimensional grid of window width versus centrality as shown in figure C.3. The smoothed MC background template has a lower boundary of 0.130 TeV, so only window configurations whose left edge satisfy $m_{\ell\ell}^{\text{left}} \geq 0.130$ are shown. Further, to narrow the comparison to more sensitive window configurations, at each pole mass only windows with expected significance above 3.0σ and 2.5σ are shown in the electron and muon channels, respectively. Lastly smoothing is applied in two ways: first, statistical fluctuations in the background are smoothed by taking the expected significance at any pole mass using the average of the surrounding ten pole masses in the loop. The second step of smoothing is applied within the $w - c$ plane to identify broader regions of interest under the assumption that similar window configurations should have correspondingly similar sensitivity.

Once the candidate window configurations are arrayed in the $w - c$ plane, the optimal point at each pole mass is taken as the center of the highest contour in expected significance. In the case of degenerate maximum contours, the one with the largest area is selected, and candidates near the edge are disfavored. In the example of figure C.3, the optimal (w, c) coordinate is marked with a star. Figures C.4 and C.5 show the measured optimal width and centrality, respectively, as functions of pole mass along with the functional fits used to model their evolution smoothly:

$$w(M_{Z'}) = \alpha_0 + \alpha_1 M_{Z'} + \alpha_2 / M_{Z'} \quad (\text{C.7})$$

$$c(M_{Z'}) = \beta_0 e^{\beta_1 M_{Z'}} + \beta_2 M_{Z'} \quad (\text{C.8})$$

where the best-fit values of the α_n and β_n in the two channels are

channel	ee	$\mu\mu$
$\alpha_0 =$	0.0296051	0.581191
$\alpha_1 =$	1.14651	0.758327
$\alpha_2 =$	-0.0131862	-0.109202
$\beta_0 =$	2.32409	1.11605
$\beta_1 =$	-6.28307	-2.75518
$\beta_2 =$	-1.3×10^{-7}	-0.0379527

(C.9)

The prescription of Equation C.8 is used to define the windows illustrated in figure C.1 and used in the SWiFt results both in extracting the spurious signal estimate of section C.4 and in the scans on data in section C.6.

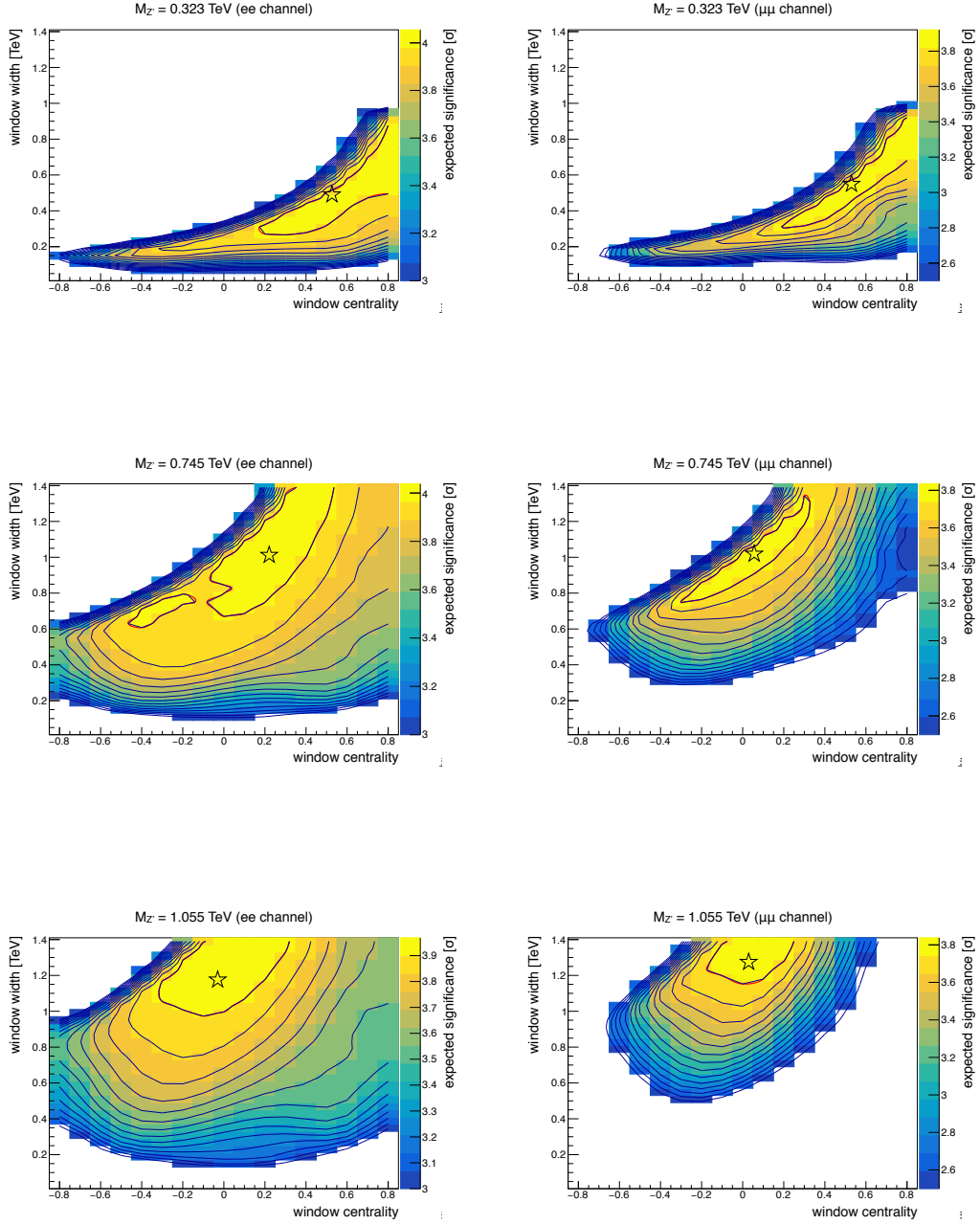
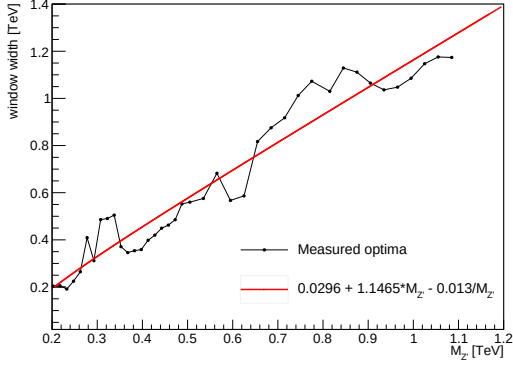
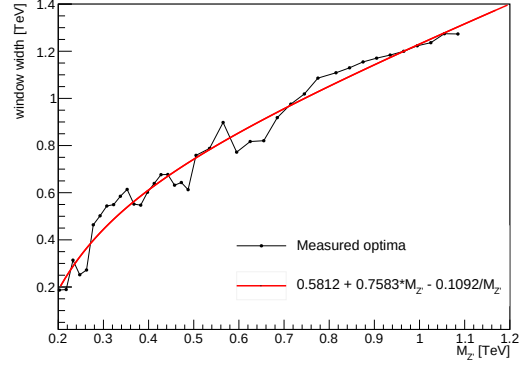


Figure C.3: Expected significance above 3.0σ and 2.5σ levels in the ee (left) and $\mu\mu$ (right) channels, respectively, for windows of different widths and centralities at $M_{Z'} = 0.323$ TeV, 0.745 TeV, 1.055 TeV. The star indicates the selected optimal window configuration in each case.

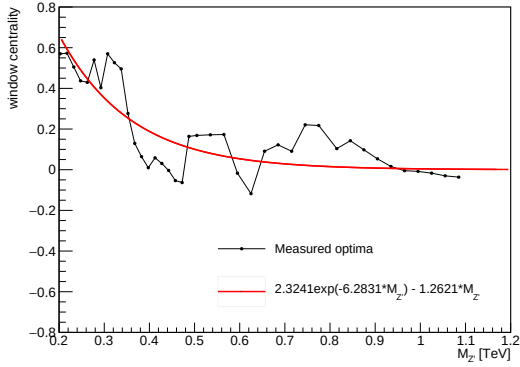


(a) Dielectron channel

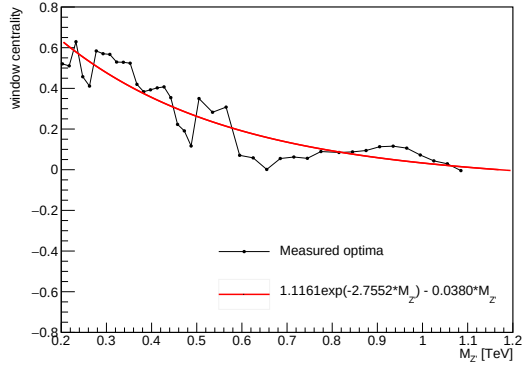


(b) Dimuon channel

Figure C.4: Optimal values of window width w as a function of pole mass in the ee (left) and $\mu\mu$ (right) channels along with their functional fits.



(a) Dielectron channel



(b) Dimuon channel

Figure C.5: Optimal values of window centrality c as a function of pole mass in the ee (left) and $\mu\mu$ (right) channels along with their functional fits.

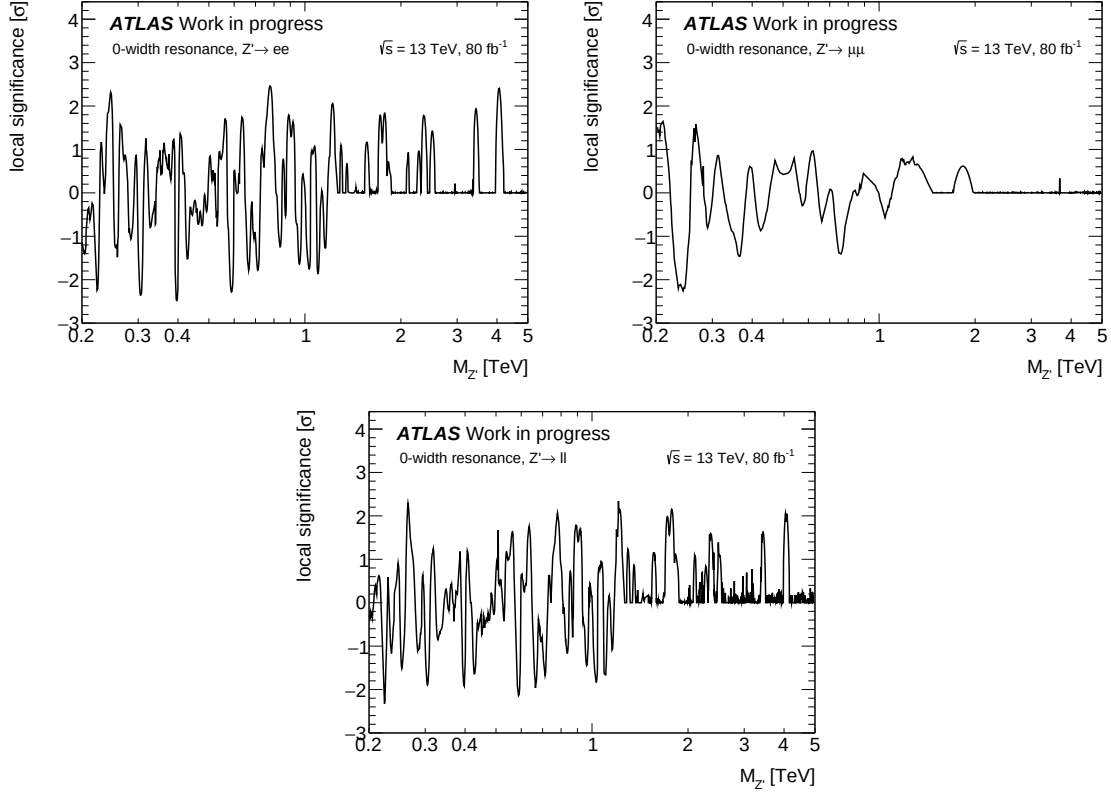


Figure C.6: Observed 1-dimensional significance scans of deviations from the background-only hypothesis, as a function of the Z' pole mass, $M_{Z'}$, for a signal width assumption of $\Gamma/M = 0\%$ on truth-level for the 2015-17 dataset in the ee (upper left), $\mu\mu$ (upper right) and $\ell\ell$ (bottom) channels. Results obtained from SWiFt implementation

C.6 Results based on SWiFt

C.6.1 Significance scans

Figure C.6 shows the significance scans on the 2015-2017 data from the SWiFt implementation using the zero-width signal model. The results incorporate the experimental uncertainties described in section 6.8 and the spurious signal systematic uncertainty described in section C.4. To also probe wide signal hypotheses, the significance scan was repeated using the zero-width signal model convoluted with Breit-Wigner lineshapes at ten different widths up to $\Gamma/M = 10\%$. The results are shown together with the zero-width significance scan in figure C.7 and do not include any systematic uncertainties. Note that the allowed range of the POI (the signal size) has been limited to take only positive values above 1.5 TeV for fit stability.

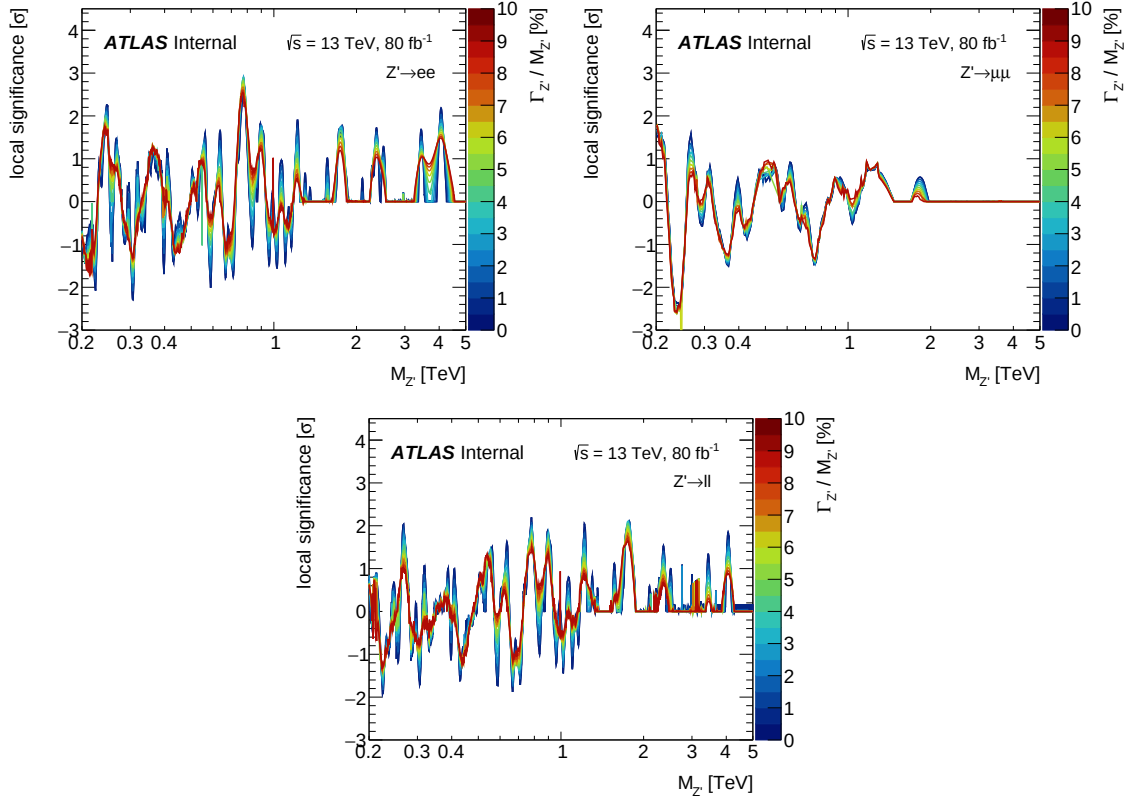


Figure C.7: Observed significance scans of deviations from the background-only hypothesis, as a function of the Z' pole mass, $M_{Z'}$ for the 2015-17 dataset in the ee (upper left), $\mu\mu$ (upper right) and $\ell\ell$ (bottom) channels. A truth-level signal width assumption of $\Gamma/M = 0\%$ is shown overlaid with ten wide signal width assumptions between $\Gamma/M = 1\%$ and $\Gamma/M = 10\%$ identified by the color axis. Results obtained from SWiFt implementation neglecting systematic uncertainties.

C.6.2 Limit curves for 2015 – 17 dataset (80 fb^{-1})

figure C.8 shows the cross section limits on the 2015-2017 data from the SWiFt implementation using the zero-width signal model. The results incorporate the experimental uncertainties described in section ?? and the spurious signal systematic uncertainty described in section C.4.

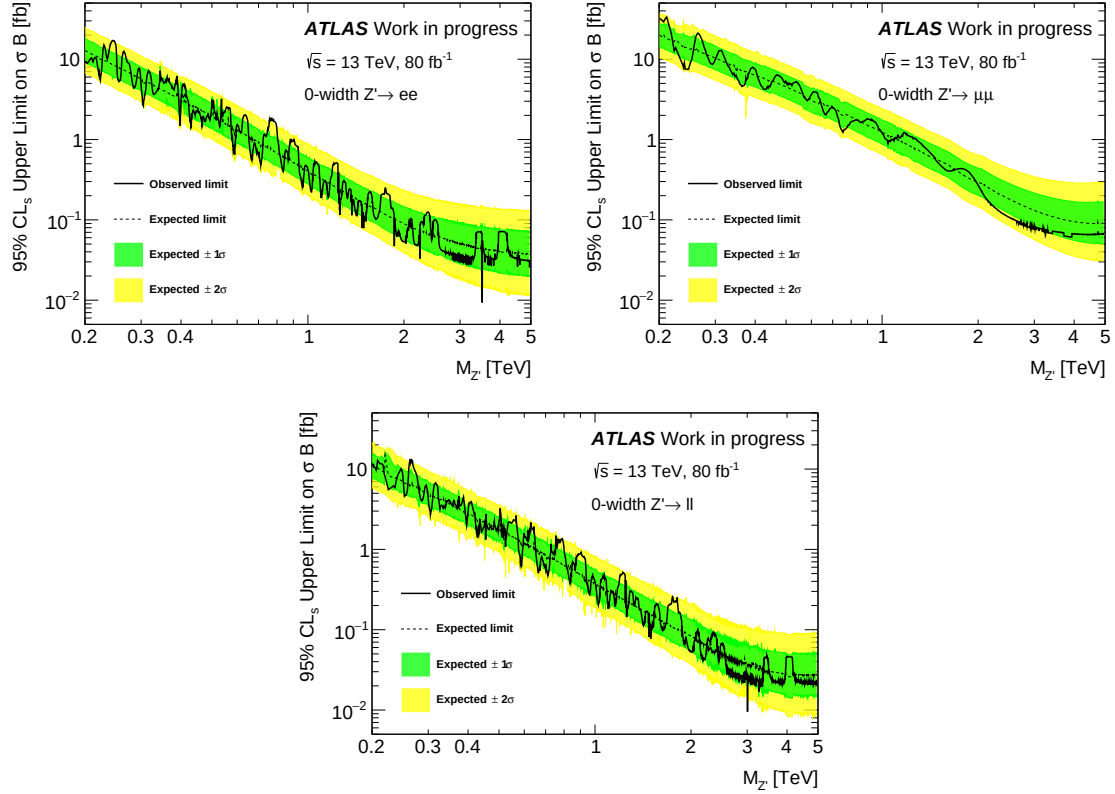


Figure C.8: Expected and observed limits on the Z' production cross-section $\times$ branching ratio into one fermion generation, $\sigma_{Z' \rightarrow \ell\ell} \times \text{BR}$, as a function of the Z' pole mass, $M_{Z'}$ in the ee (upper left), $\mu\mu$ (upper right), and $\ell\ell$ (bottom) channels. Results obtained from SWiFt implementation assuming a signal width of 0%.

C.7 Comparison between SWiFt and the global fit

The data-driven background estimation was originally investigated using two independent frameworks, a SWiFt approach and a global fit approach. Both setups are implemented using RooFit classes and the statistical procedure is as described in chapter 8. The eventual decision to adopt the global fit option was based on the performance comparison documented in this section. The spurious signal and expected upper limits were used as performance metrics at the full Run 2 target luminosity of 140 fb^{-1} along with fit stability considerations.

Both methods were optimized against the same smooth background template with the procedures described in sections C.5 and 8.2.3, and both use the same signal model. The systematic uncertainties on the signal shape, acceptance $\times$ efficiency and luminosity are treated equally between the two methods when incorporated. The effect of background uncertainties, e.g. due to PDF variations, are taken into account through the spurious signal estimate only for the global fit (compare section 8.2.4). The impact of these background variations are mainly given for the high M_{truth} region.

The function choice for the SWiFt option is given by a 3-parameter function (equation C.1), which is fitted inside dedicated windows for each M_{truth} assumption. The fit range in the SWiFt option depends on the M_{truth} assumption and can extend down to the beginning of the background template at $M_{\text{inv}} = 130 \text{ GeV}$. The high-mass cutoff is maximally given by $M_{\text{inv}} = 6 \text{ TeV}$ for the windows in the high M_{truth} region. The SWiFt option allows to scan M_{truth} values starting at 200 GeV .

The global fit option uses a 5-parameter function (equation 8.20), which is fit to the full considered invariant mass range. The fit range is given by $M_{\text{inv}} \in [225 \text{ GeV}, 6 \text{ TeV}]$ and mainly affected by the Z peak region, which starts to modify the spectrum towards lower M_{inv} . The global fit option allows to scan M_{truth} values starting at 250 GeV .

C.7.1 Spurious signal estimate

The comparison of the spurious signal estimates is given in figure C.9. The raw spurious signal values of the two methods do not show considerable differences in the intermediate mass regions, while for the low mass region the global fit option yields up to 20% (30%) higher spurious signal values for the statistical fluctuations at around 300 GeV for the ee ($\mu\mu$) channel. Higher spurious signal values are present in the high mass region for the global fit, which might be due to the inclusion of background uncertainties in the global fit option. The spurious signal in the high mass region is typically below 1 signal event and inconsequential for the statistical results of both methods.

The spurious signal parameterization involves a somewhat *ad hoc* functional choice intended as a reasonable envelope of the raw values, and as such it is different for the two methods. The parameterization could be adapted according to the requirements placed on the functional form chosen and if it needs to capture *all* raw spurious signal values. Currently the SWiFt parameterization is more conservative for intermediate masses, whereas the global fit parameterization slightly underestimates some raw values which are most likely driven by statistical fluctuations within the background template.

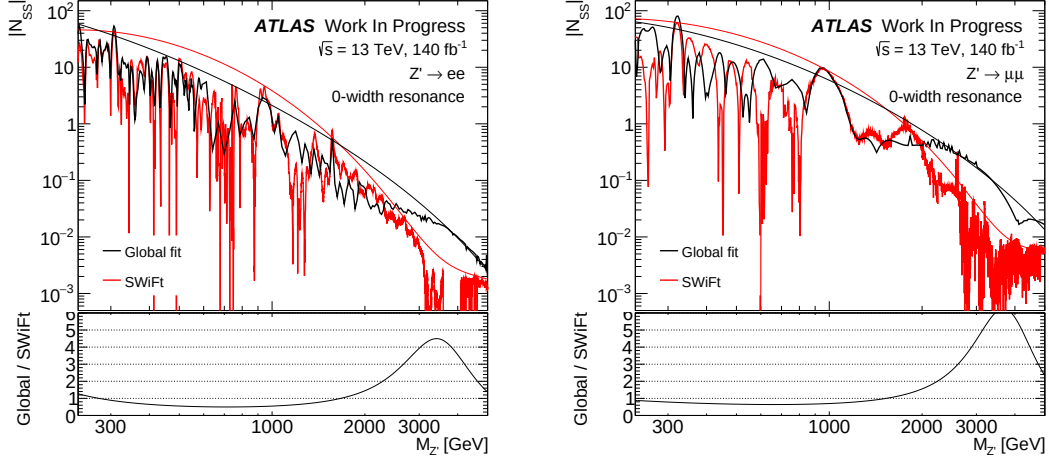


Figure C.9: Comparison of the spurious signal estimations in the SWiFt and global fit options at 140 fb^{-1} . The thick lines give the raw spurious signal values, while the thin lines give the functional parameterization.

C.7.2 Limit curves without systematic uncertainties

The comparison of the SWiFt and global fit stat.-only limits (which neglect all systematic uncertainties) at 140 fb^{-1} for the ee , $\mu\mu$ and $\ell\ell$ channels is given in figure C.10.

In the ee channel no significant difference between the two methods is present. The limits in the $\mu\mu$ channel are up to 20% lower with global fit option in the high-mass region, which is most probably an effect of the large $\mu\mu$ mass resolution and thus the very wide signal distributions. The global fit option might provide a more stable background estimate in this case. In the low mass region, SWiFt provides stronger limits by 5 – 10% in the $\mu\mu$ channel, probably being related to the possibility to cover invariant masses lower than the global fit cut-off at 225 GeV.

The $\ell\ell$ channel limits are affected by the already described differences in the single channels and is up to 5% stronger for the global fit option in the intermediate and high-mass regions.

C.7.3 Limit curves including systematic uncertainties

The comparison of the SWiFt and global fit limits at 140 fb^{-1} including systematic uncertainties for the ee , $\mu\mu$ and $\ell\ell$ channels is given in figure C.11.

Compared to the limit curves obtained without taking into account systematic uncertainties, the global fit provides consistently stronger limits than SWiFt in figure C.11. The increased discrepancy between the limits at intermediate mass is in part due to the difference in choice of spurious signal parameterization shown in figure C.9.

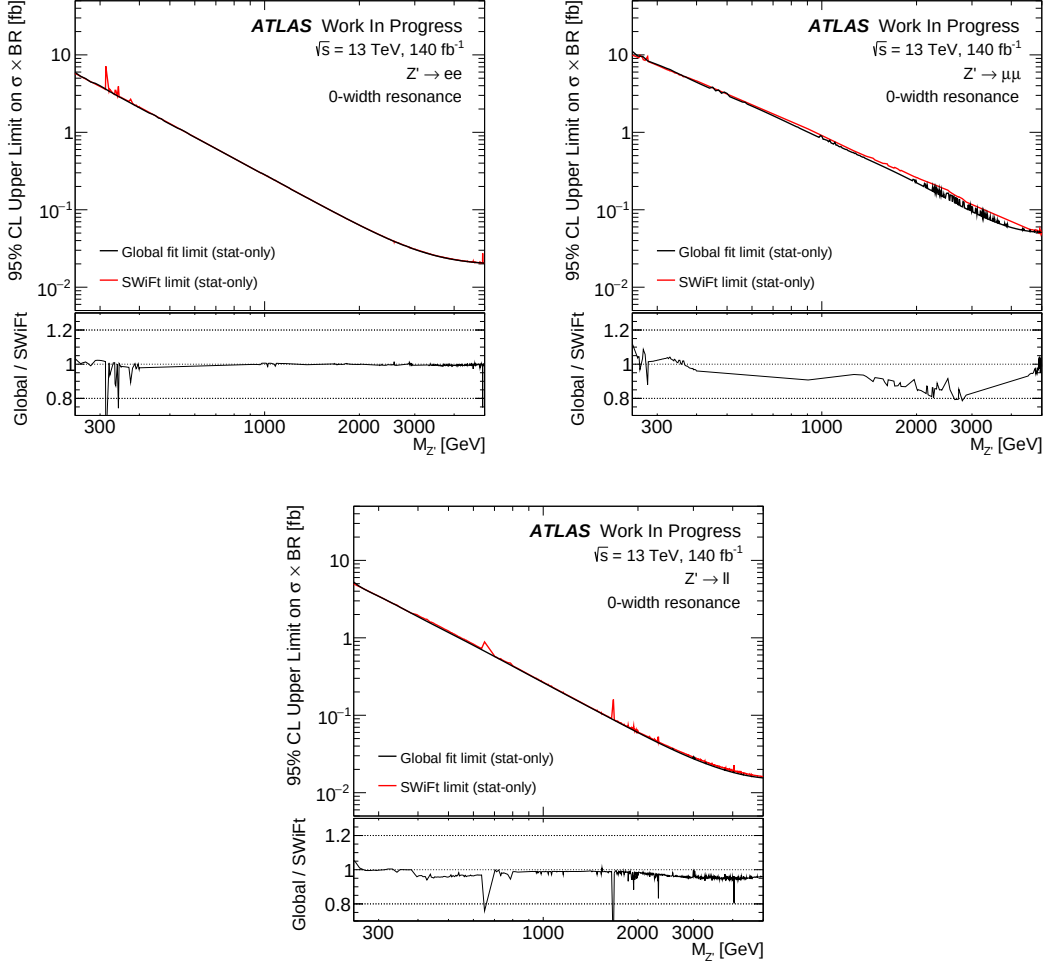


Figure C.10: Comparison of the expected limits on the Z' production cross-section $\times$ branching ratio into one fermion generation, $\sigma_{Z' \rightarrow \ell\ell} \times \text{BR}$, as a function of the Z' pole mass, $M_{Z'}$ in the ee (upper left), $\mu\mu$ (upper right) and $\ell\ell$ channels (bottom) at 140 fb^{-1} for the SWiFt and global fit options. A signal width of 0% is assumed and no systematic uncertainties are taken into account.

C.7.4 Conclusions from comparison

Both methods are comparably stable and sensitive descriptions of the expected ee and $\mu\mu$ background distributions at 140 fb^{-1} . A comparison is done in terms of the spurious signal and the expected limits for SWiFt and the global fit. The difference in the expected stat.-only limits of the two methods is typically below 5% (20%) in the ee ($\mu\mu$) channels with the global fit generally providing better sensitivity for masses above 350 GeV. This discrepancy is increased for the expected limits including systematic uncertainties largely because of the difference in spurious signal parameterization choices. The level of agreement between expected limits indicates that both **RooSwift** and **Hfitter** are viable options and this is useful in case the need for a cross-check arises. Moreover both methods support the same conclusion regarding the competence of a data-driven resonant search strategy.

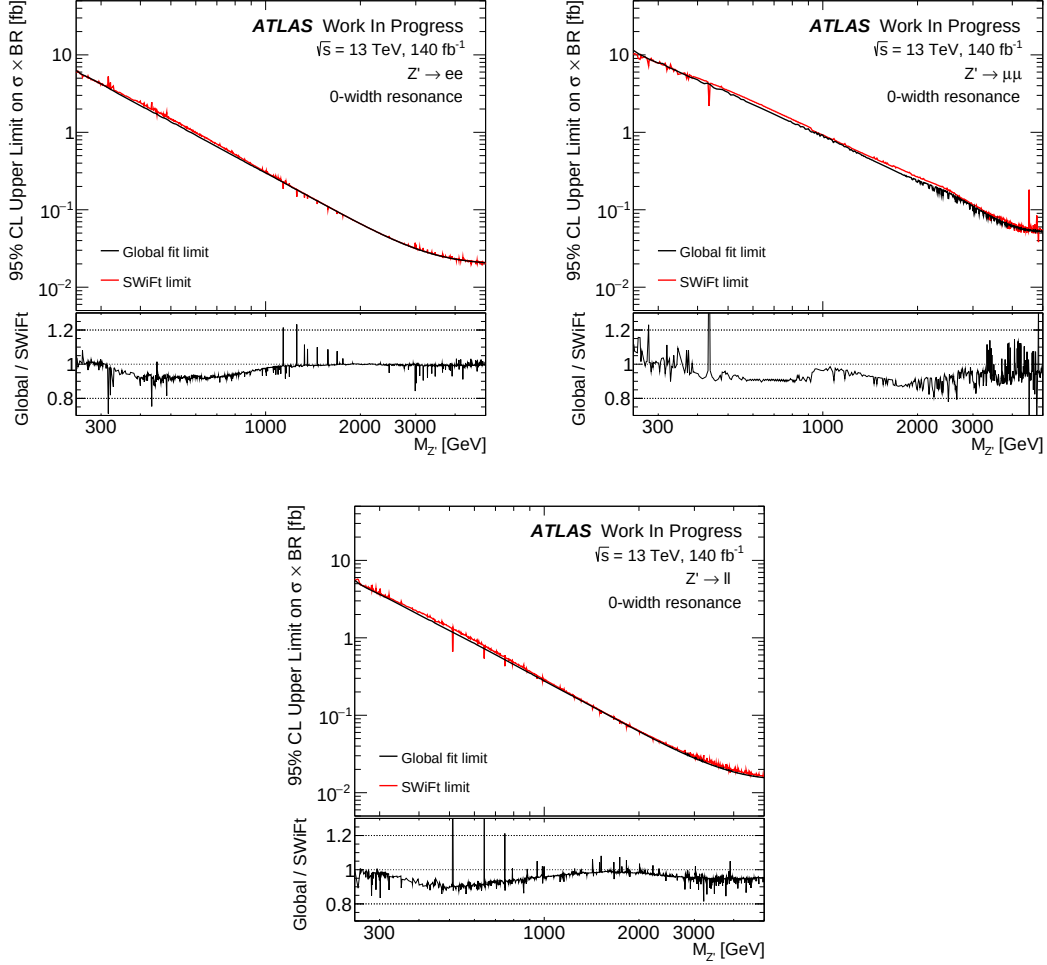


Figure C.11: Comparison of the expected limits on the Z' production cross-section $\times$ branching ratio into one fermion generation, $\sigma_{Z' \rightarrow \ell\ell} \times \text{BR}$, as a function of the Z' pole mass, $M_{Z'}$, in the ee (upper left), $\mu\mu$ (upper right) and $\ell\ell$ channels (bottom) at 140 fb^{-1} for the SWiFt and global fit options. A signal width of 0% is assumed and all experimental uncertainties as well as the spurious signal uncertainty are taken into account. The standalone upper limit plots for SWiFt are given in figure C.8, while figure 8.9 contains the global fit information.

Because of the slightly better limits in the global fit option and the fact that it had been successfully used in past analyses (the diphoton resonance search), it was chosen over the more novel SWiFt approach. The agreement between statistical results both on the full Run-2 expected dataset and on the 80 fb^{-1} dataset (figures C.8, C.6) was helpful to have confidence in the data-driven analysis strategy.

Appendix D

Additional reinterpretation material

D.1 Additional HVT material

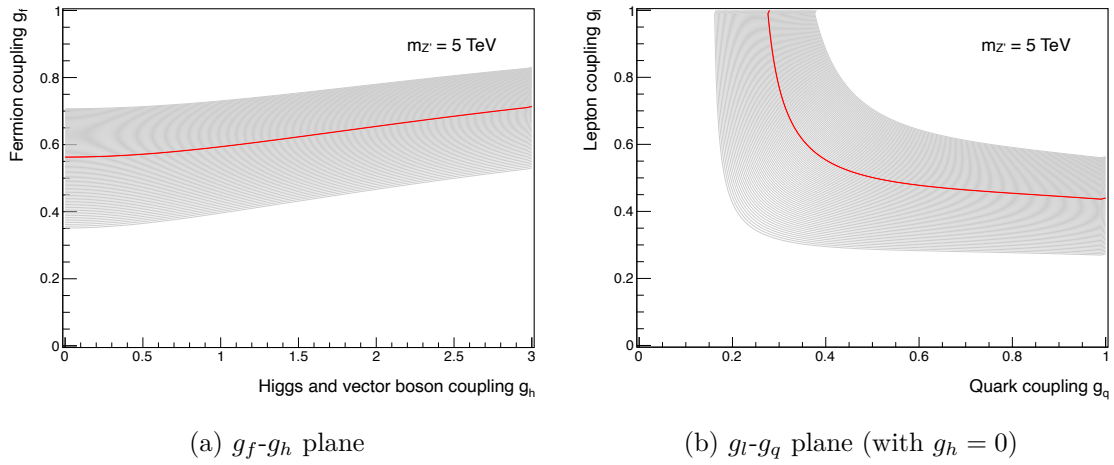
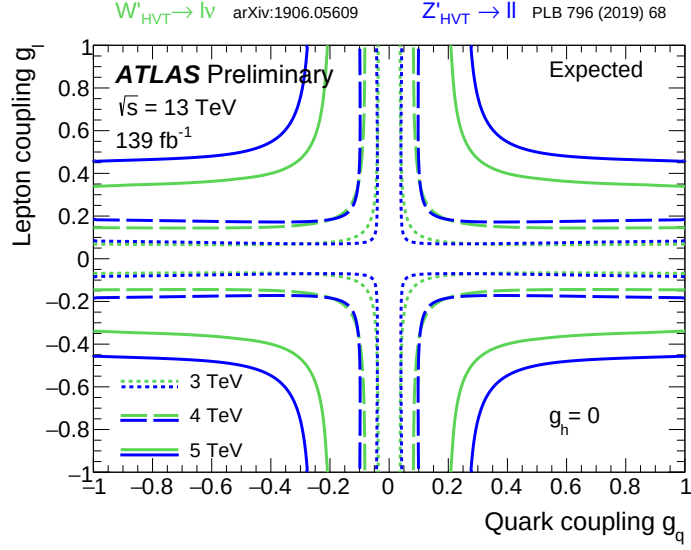
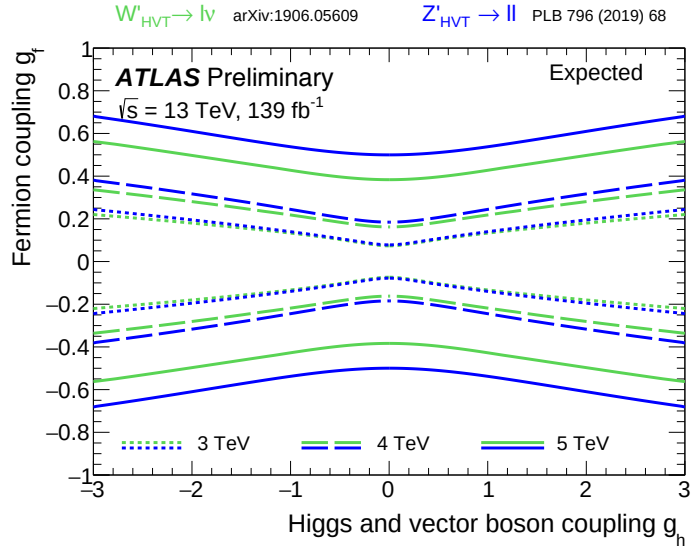


Figure D.1: Example showing array of contours drawn on the ratio of HVT fiducial cross section to the upper limit. Red curve is the contour where ratio is nearest to 1.



(a) g_l - g_q plane (with $g_h = 0$)



(b) g_f - g_h plane

Figure D.2: Expected 95% CL exclusions on HVT model fermionic and bosonic couplings derived from the ATLAS high-mass dilepton [1] (blue) and lepton + $E_{\text{T}}^{\text{miss}}$ [156] (green) searches. Regions outside the the lines are excluded.

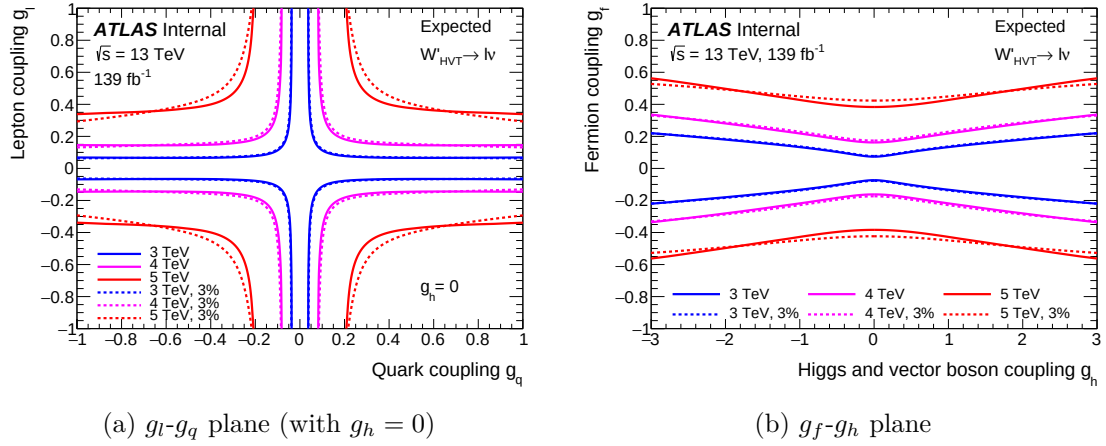


Figure D.3: Comparison of HVT expected exclusions obtained in the $\ell\nu$ final state [156] using cross section limits at a fixed width of $\Gamma_{Z'}/M_{Z'} = 3\%$ (dashed) versus the width-matching method (solid).

D.2 Additional DM material

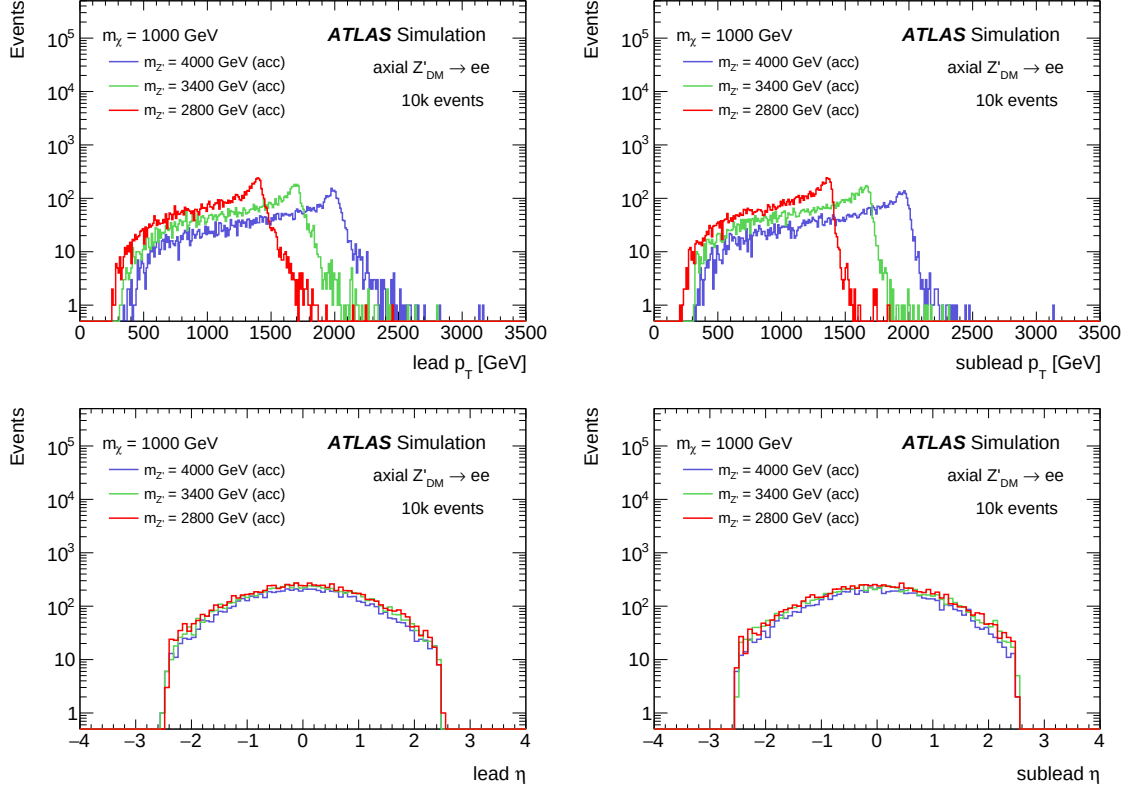


Figure D.4: Distributions of p_T and η for the leading and subleading leptons in each of the 10k generated axial-vector Z'_{DM} events after applying fiducial cuts.

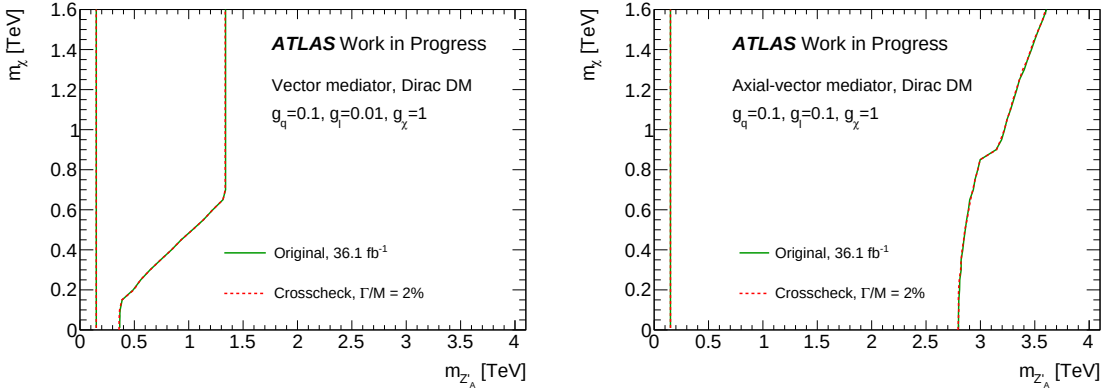


Figure D.5: Exclusion contours from dilepton cross section limits for the vector (left) and axial-vector (right) scenarios at 36.1 fb^{-1} . The solid green curve is taken from the previous DM summary plots repository (behind figures 11b and 12b of [76]) whereas the dashed red curve is a re-production of the 36.1 fb^{-1} exclusion contour using the current framework. The $\Gamma/m_{Z'} = 2\%$ observed limits from [9] were plugged into the new framework (no width-matching), and the fiducial cuts were changed to match those of [9].

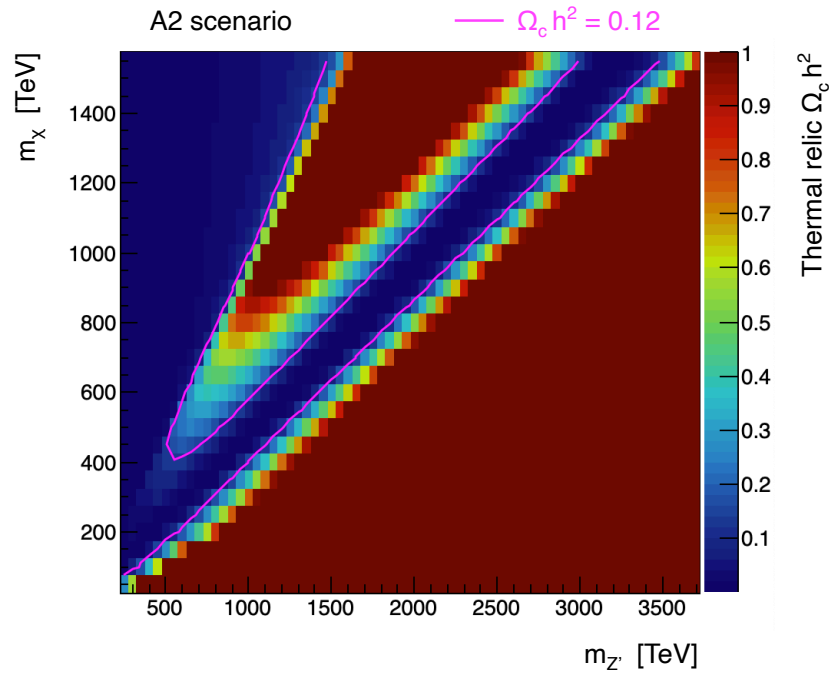


Figure D.6: Plot of thermal relic density Ωh^2 calculated using MADDM 3.1 [159] in the mass-mass plane for the A2 scenario (axial-vector mediator, $g_1 = 0.1$, $g_l = 0.1$, $g_\chi = 1$). The contour corresponding to the expected value of 0.12 by cosmological measurements is indicated by the magenta curve.

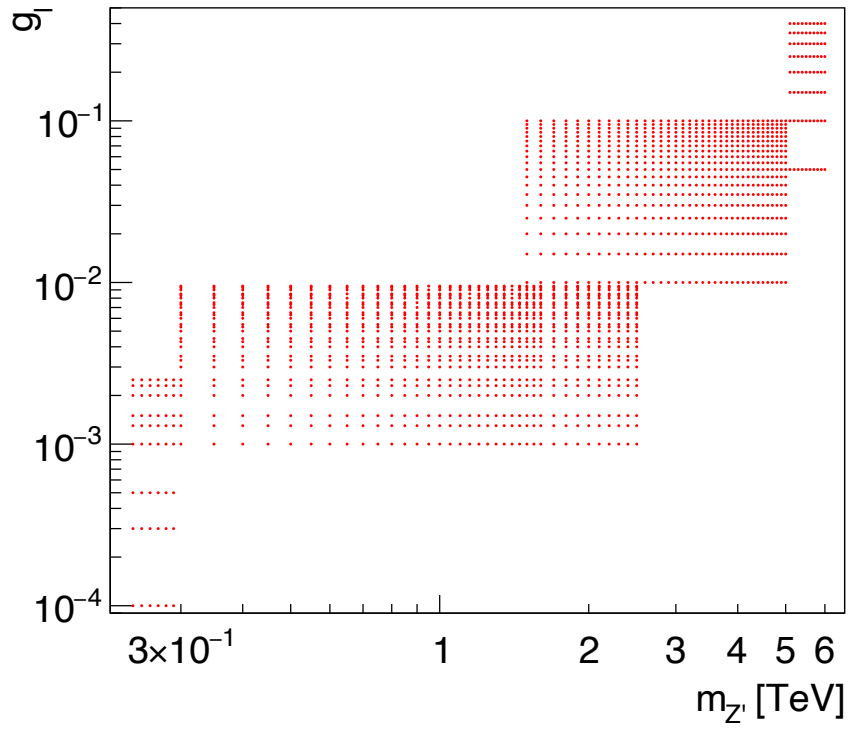


Figure D.7: The total number of 1784 points in the $(m_{Z'}, g_\ell)$ plane each corresponding to 10k generated events for $Z'_{\text{DM}} \rightarrow \ell\ell$ cross section calculation for axial-vector couplings.

List of Tables

Table 2.1	Field content of the SM prior to electroweak symmetry breaking . . .	6
Table 2.2	SM gauge group generators and their commutation relations	9
Table 4.1	Important parameters of the three ATLAS magnet systems	51
Table 4.2	Position and timing resolutions of ATLAS muon chambers	52
Table 4.3	Comparison of the ID subdetectors.	56
Table 5.1	Comparison of electron ID working points	68
Table 6.1	Possible decay modes of diboson process	79
Table 6.2	MC campaign specifications	80
Table 6.3	Event generator configurations for MC samples	82
Table 6.4	Summary of selection criteria for the full Run-2 dilepton analysis . . .	84
Table 6.5	Triggers used in the full Run-2 dilepton analysis	85
Table 6.6	Number of data events passing selection for each channel	89
Table 6.7	Set of systematic uncertainties considered	106
Table 7.1	Triggers for collecting fake-enriched data sample	116
Table 7.3	Fit parameter values for fakes estimate smoothing	122
Table 8.1	Estimated sensitivity loss due to statistical uncertainty on MC templates	134
Table 8.2	Systematic uncertainties relative channel-wise impact	143
Table 8.3	Final global fit parameter values	144
Table 8.4	Most significant observed excesses and deficits	147
Table 8.5	Benchmark scenario lower mass limits	149
Table 9.1	HVT parameter scan configuration	158
Table 9.2	Upper bounds on HVT coupling strength to fermions	163
Table 9.3	DM mediator mass-mass parameter scan configuration	166
Table A.1	Details of NLO POWHEG+PYTHIA8 Drell-Yan samples	190
Table A.2	Details of LO PYTHIA8 Drell-Yan samples	191
Table A.3	Functional forms of the $m_{\ell\ell}$ response model parameters in the dielec- tron channel	194

Table A.4	Functional forms of the $m_{\ell\ell}$ response model parameters in the dimuon channel	195
Table B.1	Comparison of p -values obtained using $\hat{q}_0$ and $\hat{q}_\mu$	211

List of Figures

Figure 1.1	Analysis strategy schematic	3
Figure 2.1	Plot of the “Mexican hat” potential	15
Figure 2.2	Running of the strong coupling constant	20
Figure 2.3	Illustration of the different intermediate components of a pp collision.	26
Figure 2.4	Main components of the ATLAS MC workflow	27
Figure 2.5	Plot of NNPDF3.0 PDFs	28
Figure 3.1	Dominant loop corrections to the Higgs mass at scale Λ	35
Figure 3.2	Example of gauge coupling unification in an $SO(10)$ GUT	38
Figure 3.3	Visual aids to the ADD and RS extra-dimensional models	40
Figure 3.4	Complimentary experimental probes of DM.	43
Figure 3.5	Spin-independent DM exclusions from direct detection experiments	44
Figure 3.6	Summary of BSM scenarios potentially connected to a dilepton resonance	45
Figure 4.1	Cutaway diagram of the ATLAS detector	48
Figure 4.2	Closeup of the ATLAS Inner Detector barrel	55
Figure 4.3	The LAr- and tile-based ATLAS calorimeter subdetectors	58
Figure 4.4	Schematic of the ATLAS 2-level trigger system in Run 2	63
Figure 6.1	Feynman diagram of Drell-Yan process	77
Figure 6.2	Drell-Yan differential cross section	78
Figure 6.3	Dominant $t\bar{t}$ diagrams contributing to the dilepton final state	78
Figure 6.4	Top quark production diagrams contributing to the dilepton final state.	79
Figure 6.5	Diboson production resulting in two leptons via either a W or Z pair.	79
Figure 6.6	Data-MC comparison in lepton p_T spectrum	90
Figure 6.7	Data-MC comparison in lepton η spectrum	90
Figure 6.8	Data-MC comparison in lepton ϕ spectrum	91
Figure 6.9	Data-MC comparison in $m_{\ell\ell}$ spectrum	93
Figure 6.10	Event display for highest- $m_{\ell\ell}$ dielectron candidate	94
Figure 6.11	Event display for highest- $m_{\ell\ell}$ dimuon candidate	95

Figure 6.12	Event display for highest- $m_{\ell\ell}$ dimuon candidate in 36.1fb^{-1} dataset	96
Figure 6.13	Example TF fit results	98
Figure 6.14	Per-channel relative $m_{\ell\ell}$ resolution	99
Figure 6.15	Per-channel selection efficiency	100
Figure 6.16	Example generic signal lineshapes	103
Figure 6.17	Fiducial acceptance of spin-1 signal	105
Figure 6.18	Systematic variations of $A\epsilon$	108
Figure 6.19	Example of signal shape systematic variations	109
Figure 7.1	Real electron rate for 2015-2018 MC campaigns	114
Figure 7.2	Data-MC comparison for fake-enriched selection	117
Figure 7.3	Fake electron rate for 2015-2018 dataset	119
Figure 7.4	Averaged rate matrix elements	121
Figure 7.5	Predicted proportions of observed event yields in actual basis	122
Figure 7.6	Functional smoothing of fakes estimate from the matrix method	123
Figure 8.1	LLR test statistic distribution in asymptotic limit	128
Figure 8.2	Example distributions of q_0 for different μ'	131
Figure 8.3	Ratio of MC to data statistical error	133
Figure 8.4	Spurious signal yields and significance for global fit	138
Figure 8.5	Spurious signal yields with α_s variations	141
Figure 8.6	Functional form fit to $m_{\ell\ell}$ spectra	146
Figure 8.7	Signal significance scan at $\Gamma_X/m_X = 0\%$	147
Figure 8.8	2D local significance scans	148
Figure 8.9	Signal fiducial cross section 95% CL upper limits in combined channel	150
Figure 9.1	Truth-level Z'_χ $m_{\ell\ell}$ distribution from reweighting DY	152
Figure 9.2	Examples of reweighted signal templates for benchmark Z' models	153
Figure 9.3	Visualisation of the fiducial mass cut on benchmark Z' signal $m_{\ell\ell}$ distributions	154
Figure 9.4	Cross sections for benchmark Z' models before and after fiducial cuts	155
Figure 9.5	Impact of PDF uncertainties on SSM cross section	156
Figure 9.6	Example diagrams of HVT model	157
Figure 9.7	Examples of reweighted HVT signal templates	158
Figure 9.8	Relative widths of HVT dilepton resonances	159
Figure 9.9	Fiducial cross sections of HVT dilepton resonances	160
Figure 9.10	Interpolated observed cross section limits	161
Figure 9.11	Observed exclusions on HVT model fermion and boson couplings	162
Figure 9.12	Feynman diagrams of s-channel DM mediator searches	164
Figure 9.13	DM particle mass versus mediator mass scans for V2 and A2 scenarios	165

Figure 9.14	Example signal $m_{\ell\ell}$ templates for axial-vector DM mediator	166
Figure 9.15	Resonance width for generated $Z'_{\text{DM}} \rightarrow \ell\ell$ signals	167
Figure 9.16	Fiducial acceptance for generated $Z'_{\text{DM}} \rightarrow \ell\ell$ signals	167
Figure 9.17	Fiducial cross sections for generated $Z'_{\text{DM}} \rightarrow \ell\ell$ signals	168
Figure 9.18	Dark matter mediator exclusions in the mass-mass plane resulting from dilepton cross section upper limits	168
Figure 9.19	ATLAS dark matter s -channel mediator summary plots	170
Figure 9.20	Comparison of ATLAS DM mediator exclusions and direct-detection bounds	171
Figure 9.21	Dark matter mediator exclusion in lepton coupling vs. mediator mass plane	172
Figure 10.1	Historical progression of the SSM Z' mass exclusion by ATLAS . .	175
Figure A.1	Data-MC comparison in event $p_{\text{T}}^{\ell\ell}$ spectrum	189
Figure A.2	Data-MC comparison in event $E_{\text{T}}^{\text{miss}}$ spectrum	189
Figure A.3	Parameterizations of the dielectron channel $m_{\ell\ell}$ response model parameters	192
Figure A.4	Parameterizations of the dimuon channel $m_{\ell\ell}$ response model parameters	193
Figure A.5	Invariant mass distributions of MSSM Higgs samples for different relative widths	196
Figure A.6	$m_{\ell\ell}$ distributions of Graviton samples for different values of k/M_{Pl}	197
Figure A.7	Comparison of η distributions for different spin cases	198
Figure B.1	Cross section upper limits in the dielectron and dimuon channels .	200
Figure B.2	Distribution of largest local significance in each channel	202
Figure B.3	Test statistic upcrossings in a 1D scan of background-only toys . .	203
Figure B.4	Validation of asymptotic relationship for 1D LEE	204
Figure B.5	Local-to-global significance map for 36.1 fb^{-1} dilepton analysis . .	205
Figure B.6	Example 2D scan of test statistic on background-only pseudoexperiment	206
Figure B.7	How to calculate the Euler characteristic	207
Figure B.8	Validation of asymptotic relationship for 2D LEE	208
Figure B.9	Example distributions of q_{μ} for different μ'	210
Figure B.10	Distributions of best-fit $\hat{\mu}$ for different injected μ'	211
Figure C.1	Illustration of SWiFt window ranges	214
Figure C.2	Measured spurious signal at zero-width using SWiFt	215
Figure C.3	Expected significance maps for SWiFt window optimization	218

Figure C.4	Optimal values of window width w as a function of pole mass . . .	219
Figure C.5	Optimal values of window centrality c as a function of pole mass . .	219
Figure C.6	Significance scans for 0% width signal at 80fb^{-1} using SWiFt . . .	220
Figure C.7	Significance scans for ten different signal width assumptions at 80fb^{-1} using SWiFt	221
Figure C.8	Upper limits on generic dilepton signal cross section at 80fb^{-1} using SWiFt	222
Figure C.9	Comparison of the spurious signal estimations in the SWiFt and global fit options at 140fb^{-1}	224
Figure C.10	Comparison of expected upper limits for SWiFt and global fit options without systematic uncertainties	225
Figure C.11	Comparison of expected upper limits for SWiFt and global fit options with systematic uncertainties included	226
Figure D.1	Example of HVT contours around exclusion	227
Figure D.2	Expected 95% CL exclusions on HVT model fermion and boson cou- plings	228
Figure D.3	Comparison of HVT contours obtained using fixed width versus width-matching	229
Figure D.4	Distributions of p_T and η for the leading and subleading leptons in each of the 10k generated axial-vector Z'_{DM} events after applying fiducial cuts.	230
Figure D.5	Crosscheck of DM mediator exclusion framework with the 36.1fb^{-1} result	230
Figure D.6	Thermal relic density in mass-mass plane	231
Figure D.7	Signal DM mediator lepton coupling versus mediator mass grid scan	232