

Precision measurements of four-lepton production in $\sqrt{s} = 13$ TeV proton-proton collisions
with the ATLAS detector

A Dissertation

Presented to

The Faculty of the Graduate School of Arts and Sciences
Brandeis University

Department of Physics

Gabriella Sciolla, Advisor

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
Zachary Schillaci

February 2022



This dissertation, directed and approved by Zachary Schillaci's Committee, has been accepted and approved by the Faculty of Brandeis University in partial fulfillment of the requirements for the degree of:

DOCTOR OF PHILOSOPHY

Wendy Cadge, Dean
Graduate School of Arts and Sciences

Dissertation Committee:
Gabriella Schilla, Physics
Craig Blocker, Physics
Jonathan Butterworth, Physics, University College London

Copyright by
Zachary Schillaci

2022

Acknowledgements

The work presented here and my progress as a physics graduate student are thanks to the help of many people to whom I am sincerely grateful. Though I cannot name everyone here, those listed below deserve special recognition.

First and foremost, I must thank my doctoral advisor **Gabriella Sciolla** for her guidance, support, and wisdom over the past five years. I would be nowhere without your excellent mentorship. Thank you for providing me with many opportunities to grow as a student and researcher. I have learned much from you in regards to physics, teaching, and research.

Secondly, to the other members of my defense committee **Craig Blocker** and **Jon Butterworth**. I am honored you served on my committee. Thank you for reviewing my dissertation and providing insightful feedback. Through our various research endeavors, I have learned a great deal about particle physics from the both of you.

Next, to my supervising postdoctorate **Max Goblirsch** for teaching me how to do particle physics research from the fundamentals to the advanced. Your technical expertise and software support are unparalleled and were integral to the work of this thesis. Thank you for your patience, kindness, and dedication in mentoring me these past five years.

To all members of the ***Inclusive Four-lepton*** analysis team for their essential contributions to the work presented here. Together, you taught me the intricacies of doing particle physics research in the ATLAS Collaboration and made my first physics analysis such an educational and rewarding experience.

To my undergraduate research advisor **Matt Amthor** for believing in my capabilities and starting me on the path of physics research. Your teaching and mentorship were instrumental in preparing me for a career as a graduate student.

To many other mentors in ATLAS who shaped me along the way - including **Gaetano Barone**, **Noemi Calace**, **Phil Kuczewski**, **Dave Lynn**, **Nora Pettersson**, **Thomas Strebler**, **Stefania Stucci**, and **Alessandro Tricoli**. From each of you I have benefitted

from your generous support and technical expertise.

To all of my Sciolla Lab colleagues for learning with and teaching me along the way - particularly **Prajita Bhattarai**, **Laura Bergsten**, **Francesca Capocasa**, **Jiayi Chen**, and **Hannah Herde**. Through our many years in collaboration, I consider myself incredibly fortunate to have worked alongside others whom I can also call my friends.

To all other Bucknell and Brandeis physics colleagues with whom I spent countless late nights working on problems sets and studying for exams - including **Spasen Chaykov**, **Sagnik Ghosh**, **Gus Gordon**, and **Jonathan Redwine**. I would have never succeeded as a student were it not for you.

To my family - **Dylan**, **Katelin**, and my **Mom** and **Dad** - for all of their love, support, and encouragement throughout my life. To my Mom, you have been central to all of my successes in life, thank you for everything you have done and continue to do. To my Dad, I credit you with inspiring my curiosity in understanding how things work on a fundamental level, without which, I would have never found my love of physics.

Finally, to **Aga** for all of your support and everything that you mean in my life.

To my father, Onofrio Schillaci.

ABSTRACT

Precision measurements of four-lepton production in $\sqrt{s} = 13$ TeV proton-proton collisions with the ATLAS detector

A dissertation presented to the Faculty of the
Graduate School of Arts and Sciences of Brandeis University
Waltham, Massachusetts

By Zachary Schillaci

Measurements of inclusive four-lepton (4ℓ) production in $\sqrt{s} = 13$ TeV proton-proton collisions at the Large Hadron Collider using the ATLAS Run-2 (2015-2018) dataset of 139 fb^{-1} are presented. The 4ℓ final state, consisting of two same-flavor, opposite-charge lepton pairs, exhibits a rich structure with contributions from several modes of production - including Z -boson and Higgs resonant decays and a variety of multi-boson processes - and benefits from a clean detector signature, making it the ideal candidate for both precision studies of the Standard Model and interpretations of new physics beyond-the-Standard Model. Four-lepton production cross-sections are measured inclusively in several invariant mass regions and differentially versus a variety of other kinematic observables. The measurements are corrected for detector effects and compared to state-of-the-art Standard Model predictions, which are found to be consistent with the data. The most precise measurement to date of the $Z \rightarrow 4\ell$ branching fraction is extracted and found to be in agreement with prior measurements and theoretical predictions. Constraints are placed on two scenarios of new physics, while further re-interpretations remain promising given the inclusive nature of the analysis and minimal model dependence in the measurements.

Table of Contents

Acknowledgements	iv
Abstract	vii
List of Tables	xii
List of Figures	xiii
List of Abbreviations	xvii
I Introduction	1
II Theory	2
1 History	2
1.1 Special Relativity	2
1.2 Quantum Mechanics	5
1.3 Relativistic Quantum Mechanics	8
2 The Standard Model	10
2.1 Constituents	10
2.2 Theoretical Underpinnings	13
2.3 Outstanding Problems	21
3 Testing the Standard Model	23
3.1 Collider Principles	23
3.2 Diboson Physics	24
3.3 Four-lepton Processes	25
III Apparatus	30
4 CERN	30
5 The Large Hadron Collider	31

6	The ATLAS Detector	34
6.1	Inner Detector	36
6.2	Electromagnetic and Hadronic Calorimeters	38
6.3	Muon Spectrometer	39
7	Data Processing	41
7.1	Triggering and Data Acquisition	41
7.2	From Signals to Sigmas	42
8	Particle Reconstruction	45
8.1	Techniques	45
8.2	Objects	50
9	Outlook for Future Upgrades	57
9.1	High Luminosity LHC Program	57
9.2	ATLAS Detector Upgrades	57
IV	Analysis Overview	63
10	Analysis Goals	66
11	Fiducial Selection	67
11.1	Signal Definition	67
11.2	Kinematic Selections	69
12	Reconstruction Selection	71
12.1	Physics Objects	71
12.2	Overlap Removal	73
12.3	Event Selection	74
13	Dataset and Monte Carlo Samples	76
13.1	LHC Dataset	76
13.2	Monte Carlo Samples	78
14	Observables	82

V	Analysis Strategy	85
15	Trigger	86
	15.1 Menu	86
	15.2 Efficiency	86
16	Backgrounds	93
	16.1 Sources	94
	16.2 Analysis Regions	96
	16.3 Estimation Strategy	100
	16.4 Calculation of Efficiencies	103
	16.5 Choice of Estimation Strategy	107
	16.6 Smoothing of Prediction	108
	16.7 Validation of Estimate	109
17	Unfolding	115
	17.1 Overview	115
	17.2 Algorithm	117
	17.3 Binning	118
	17.4 Pre-unfolding	121
	17.5 Optimization	127
	17.6 Validation	129
18	Uncertainties	133
	18.1 Systematics	134
	18.2 Statistics	141
	18.3 Impact	143
VI	Results and Interpretations	155
19	Detector Measurements	156
20	Unfolded Measurements	158

21	$Z \rightarrow 4\ell$ Branching Fraction	166
22	Interpretations	172
	22.1 Standard Model Effective Field Theory	174
	22.2 $B - L$ Gauge Symmetry Model	184
23	Re-interpretations	187
VII Conclusion		188
References		189
Appendices		196
A	Primary Contributions	197
	A.1 Inclusive Four-lepton Analysis	197
	A.2 ATLAS Collaboration	198
B	Collider Kinematics	200
C	Selection Optimization	201
	C.1 Loosening of Selections	201
	C.2 Lepton Isolation	203
D	Background Estimation Strategy	206
E	Data-driven Closure Test	207

List of Tables

1	Summary of the Standard Model elementary particles	11
2	Summary of the Standard Model fundamental interactions	11
3	Comparison of the Run-2 and HL-LHC ATLAS tracking detectors	59
4	Fiducial selection	70
5	Baseline lepton selection	72
6	Signal lepton selection	73
7	Baseline and signal jet selections	73
8	Event-level reconstruction selection	74
9	Luminosity conditions during LHC Run-2	78
10	Overview of signal Monte Carlo samples	81
11	Overview of background Monte Carlo samples	81
12	Active lepton triggers by run period	87
13	$Z \rightarrow \ell\ell$ control region selection	98
14	Detailed breakdown of unfolded uncertainties	152
15	Reconstruction-level inclusive yields	156
16	Fiducial-level inclusive cross-sections	158
17	Summary of the inputs to the $\mathcal{B}_{Z \rightarrow 4\ell}$ measurement	170
18	Comparison of recent $\mathcal{B}_{Z \rightarrow 4\ell}$ measurements	171
19	Overview of the dimension-six EFT coefficients considered in this analysis . .	176

List of Figures

1	Summary of the Standard Model elementary particles	12
2	Four-lepton production in the Standard Model	27
3	Four-lepton production beyond-the-Standard Model	28
4	Aerial image of CERN and the surrounding area	33
5	The CERN accelerator complex	33
6	The ATLAS detector	35
7	The ATLAS Inner Detector	37
8	The ATLAS calorimeter systems	38
9	The ATLAS Muon Spectrometer	40
10	The ATLAS trigger and data acquisition system	42
11	Data and Monte Carlo processing chains in ATLAS	44
12	Parameterization of a charged particle trajectory in a magnetic field	46
13	Overview of the track reconstruction chain in ATLAS	48
14	Overview of an electromagnetic calorimeter cell in ATLAS	50
15	Principles of particle reconstruction in ATLAS	51
16	Electron reconstruction in ATLAS	53
17	Muon reconstruction in ATLAS	54
18	Jet reconstruction in ATLAS	56
19	Timeline of the LHC operation	58
20	Latest ITk detector layout	61
21	Track reconstruction efficiency with the ITk detector	61
22	Pile-up scaling of track reconstruction with the ITk detector	62
23	Impact parameter resolution with the ITk detector	62
24	Summary of Standard Model cross-section measurements	64
25	Data versus theory agreement of diboson cross-section measurements	65
26	Dependence of particle-level lepton definitions on mass resolution	68

27	Event display of a candidate $ZZ \rightarrow \mu\mu ee$ decay	75
28	Integrated luminosity of LHC Run-2	77
29	Observed pile-up profiles during LHC Run-2	77
30	Illustration of an example $pp \rightarrow ZZ \rightarrow \mu^+\mu^-e^+e^-$ decay	84
31	Validation of the toy-based trigger efficiency computation	89
32	Evaluation of trigger efficiency scale factors	90
33	Trigger efficiency breakdown by lepton multiplicity	90
34	Combined trigger and matching efficiency	91
35	Validation of the trigger efficiency computation	92
36	Depiction of prompt and non-prompt leptons	93
37	Truth origin of quadruplet leptons	95
38	Analysis signal and validation regions	97
39	Truth origin of non-prompt leptons in signal and validation regions	99
40	Truth origin of non-prompt leptons across analysis regions	99
41	Real efficiencies measured for electrons and muons	104
42	Electron fake efficiency with associated errors	105
43	Muon fake efficiency with associated errors	106
44	Background estimation method comparison	108
45	Overview of the background smoothing procedure	110
46	Validation of the background estimation	112
47	Detailed validation of the background estimation in the VR-DF	113
48	Detailed validation of the background estimation in the VR-SC	114
49	Unfolding correction factors for $m_{4\ell}$	119
50	Unfolding migration matrix for $m_{4\ell}$	119
51	Unfolding migration matrix for inclusive- $m_{4\ell}$	120
52	Unfolding migration matrix for m_{Z_1} - $m_{4\ell}$	120
53	Reconstruction efficiency derivation	123

54	Isolation and track-to-vertex association efficiency derivation	124
55	Combined pre-unfolding efficiency derivation	125
56	Unfolding robustness to a new physics signature	126
57	Unfolding robustness to variations in signal composition	127
58	Investigation of optimal number of unfolding iterations	129
59	Overview of the Monte Carlo closure test	130
60	Overview of the data-driven closure method	132
61	Evaluation of QCD scale uncertainties	137
62	Evaluation of PDF variation uncertainties	137
63	Statistical uncertainty evaluation from bootstrapping	142
64	Detailed breakdown of fiducial-level $m_{4\ell}$ uncertainties	144
65	Detailed breakdown of reconstruction-level $m_{4\ell}$ uncertainties	145
66	Detailed breakdown of unfolded $m_{4\ell}$ uncertainties	147
67	Detailed breakdown of unfolded $m_{4\ell}$ -channel uncertainties	148
68	Detailed breakdown of unfolded m_{Z_1} - $m_{4\ell}$ uncertainties	149
69	Detailed breakdown of unfolded m_{Z_2} - $m_{4\ell}$ uncertainties	150
70	Detailed breakdown of unfolded $ \Delta\phi_{\text{pairs}} $ - $m_{4\ell}$ uncertainties	151
71	Unfolded $m_{4\ell}$ uncertainty correlation matrices	154
72	Detector-level $m_{4\ell}$ measurement	157
73	Unfolded $m_{4\ell}$ cross-section	161
74	Unfolded $m_{4\ell}$ -channel cross-section	162
75	Unfolded m_{Z_1} - $m_{4\ell}$ cross-section	163
76	Unfolded m_{Z_2} - $m_{4\ell}$ cross-section	164
77	Unfolded $ \Delta\phi_{\text{pairs}} $ - $m_{4\ell}$ cross-section	165
78	Representative $Z \rightarrow 4\ell$ signal and background processes	166
79	Comparison of recent $\mathcal{B}_{Z \rightarrow 4\ell}$ branching fraction measurements	171
80	Summary of the constraints derived on EFT coefficient c_{HWB}	180

81	Limits on the dimension-six EFT coefficients under the full model	182
82	Limits on the dimension-six EFT coefficients under the linear-only model . .	183
83	Limits on the $B - L$ gauge symmetry model	186
84	Impact of loosening analysis selections	202
85	Isolation working point event yield comparison	205
86	Isolation working point signal efficiency comparison	205
87	Overview of the data-driven closure method for $m_{4\ell}$ -channel	207
88	Overview of the data-driven closure method for m_{Z_1} - $m_{4\ell}$	208
89	Overview of the data-driven closure method for m_{Z_2} - $m_{4\ell}$	209
90	Overview of the data-driven closure method for $ \Delta\phi_{\text{pairs}} $ - $m_{4\ell}$	210

List of Abbreviations

- **ATLAS**: A Toroidal LHC Apparatus
- **BSM**: Beyond-the-Standard Model
- **CL**: Confidence level
- **CMS**: Compact Muon Solenoid
- **CSC**: Cathode Strip Chambers
- **CR**: Control region
- **ECAL**: Electromagnetic Calorimeter
- **EFT**: Effective field theory
- **EW**: Electroweak
- **FSR**: Final state radiation
- **GRL**: Good Run List
- **HCAL**: Hadronic Calorimeter
- **HGTD**: High Granularity Timing Detector
- **HL-LHC**: High Luminosity Large Hadron Collider
- **HLT**: High-level trigger
- **IBL**: Insertable B-layer
- **ID**: Inner Detector
- **IFF**: Isolation and Fake Forum
- **IP**: Interaction point
- **ITk**: Inner Tracker
- **JVT**: Jet-vertex tagger
- **LAr**: Liquid-argon
- **LEP**: Large Electron Positron Collider
- **LHC**: Large Hadron Collider
- **LO**: Leading order
- **MC**: Monte Carlo
- **MDT**: Monitored Drift Tubes
- **MS**: Muon Spectrometer

- **NLO**: Next-to-leading order
- **NNLO**: Next-next-to-leading order
- **OR**: Overlap removal
- **PDF**: Parton distribution function
- **QCD**: Quantum Chromodynamics
- **QED**: Quantum Electrodynamics
- **RF**: Radio frequency
- **RPC**: Resistive Plate Chamber
- **SCT**: Semiconductor Tracker
- **SF-OC**: Same-flavor, opposite-charge
- **SM**: Standard Model
- **SMEFT**: Standard Model Effective field theory
- **SR**: Signal region
- **SUSY**: Supersymmetry
- **TGC**: Thin Gap Chambers
- **TRT**: Transition Radiation Tracker
- **TTVA**: Track-to-vertex association
- **VR-DF**: Different-flavor validation region
- **VR-SC**: Same-charge validation region
- **WIMP**: Weakly-interacting massive particle
- **WP**: Working point

Chapter I: Introduction

“What we observe is not nature in itself, but nature exposed to our method of questioning.”

—— Werner Heisenberg, in *Physics and Philosophy*

As best we can tell, the Universe is governed by a set of four fundamental forces dictating the interactions of a collection of elementary, irreducible particles. The relative sparseness of this description should not be mistaken for simplicity, for the underlying theories are mathematically complex and, more importantly, experimentally well vetted. This fundamental theory of almost everything¹, known as the Standard Model of particle physics, was developed only fifty years ago, yet has origins stretching back to the dawn of scientific inquiry. The operative word here being almost, for the theory is far from offering a complete description of nature, leaving many questions in fundamental physics unanswered. Theoretical reasoning can help to motivate and guide the search for these open questions, but there can be no answer without experimentation. Through rigorous experimentation, new paths forward may be discovered if unexpected deviations from theoretical predictions are found. Following such clues, new theories may be rejected or substantiated on the basis of experimental evidence. In this way, our scientific understanding may be refined, gradually approaching a more genuine description of nature. This thesis represents an increment, however slight, towards this goal of scientific progress.

¹Inspired by the title of Robert Oerter’s book [1].

Chapter II: Theory

1 History

The foundation of modern particle physics is based upon the union of two revolutionary theories formulated in the early 20th century: special relativity and quantum mechanics. These seemingly unrelated theories were developed out of necessity, shattering long-held scientific assumptions in order to resolve inconsistencies in classical physics. The theory of special relativity, published by Albert Einstein in 1905 [2], overturned the notions of absolute space and time in favor of a unified, malleable space-time, capable of properly describing coordinate transformations of Maxwell's theories of electromagnetism. Quantum mechanics was formulated over the course of two decades, beginning with Max Planck's hypothesis of light quanta in 1900, and followed by the work of many others, to solve a myriad of problems in atomic physics.

1.1 Special Relativity

The theory of special relativity is based upon the postulate of physical invariance, that is, the requirement that the laws of physics are identical in all inertial frames of reference. This was already understood to be the case for much of classical physics, in which inertial frames of reference moving in constant relative motion with respect to one another are trivially related according to the transformations of Galilean relativity. For a reference frame moving

in constant motion v along the x -axis, these transformations are simply

$$\begin{aligned}x' &= (x - vt) \\y' &= y \\z' &= z \\t' &= t.\end{aligned}\tag{1}$$

The problem, however, lay in the treatment of electromagnetism, the unified theory of electricity and magnetism formulated by James Clerk Maxwell some fifty years prior. Maxwell's theories capture a host of electric and magnetic phenomena and, crucially, describe light as a form of electromagnetic radiation which propagates at a fixed speed, c , in vacuum². Under Einstein's invariance postulate, all observers in inertial frames of reference must agree upon these same laws of electromagnetism and, therefore, predict and measure the same speed of light. This is impossible in the framework of Galilean relativity, in which observers in motion relative to one another would necessarily disagree on measurements of the speed of light due to the additive nature of relative speeds³. In order to solve this fundamental inconsistency, Einstein derived a set of space-time coordinate transformations capable of maintaining the invariance of physical laws. As is often the case in science, these transformations had already been derived, unbeknownst to Einstein, by another physicist, Hendrik Lorentz, whose surname they now bear. Once again, considering uniform motion along the

² In SI units, this is defined to be exactly $c = 299,792,458$ m/s.

³ To see this, differentiate Equation 1 with respect to time along the coordinate of motion, $\frac{dx'}{dt} = \frac{dx}{dt} - v$. This implies the unprimed velocity $u := \frac{dx}{dt}$ is shifted according to the motion of the frame, $u' = u - v$.

x -axis, the corresponding Lorentz transformations are

$$\begin{aligned}x' &= \gamma(x - vt) \\y' &= y \\z' &= z \\t' &= \gamma(t - \frac{vx}{c^2}),\end{aligned}\tag{2}$$

where the Lorentz factor is defined $\gamma := (1 - \beta^2)^{-\frac{1}{2}}$, with $\beta := v/c$.

The effect of Lorentz transformations along a specific axis of motion results in time dilation and length contraction along the spatial axis, relative to another inertial frame of reference. This violates the intuitive notions of absolute space and time, as well as simultaneity, since observers in motion with respect to one another no longer agree on measurements of length or duration. However, such effects are practically imperceptible at classical speeds, $\beta \ll 1$, where the Lorentz factor corresponding to the magnitude of the correction is approximately unity, $\gamma \approx 1$.

Mathematically, relativistic observables are best expressed in terms of space-time four-vectors, which themselves transform according to the rules of the Lorentz transformation. For instance, assuming energy and momentum form a union, akin to time and space in the coordinate transformations, their relativistic counterparts can be derived as $E = \gamma mc^2$ and $\vec{p} = \gamma m \vec{\beta}$, which form a contravariant four-vector $p^\mu = (E/c, \vec{p})$. The equation for relativistic energy has the surprising, and now famous, result that even in the rest frame of a massive particle ($\gamma = 1$), there is an intrinsic energy $E_0 = mc^2$ associated with its *rest mass*. Energy and three-momentum are separately conserved, meaning the four-vector is also conserved. In special relativity, the scalar product of any contravariant four-vector with the corresponding covariant four-vector is necessarily a Lorentz invariant quantity. Using the metric tensor $g_{\mu\nu}$

defined as

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (3)$$

the four-momentum scalar product is

$$p^\mu p_\mu = E^2/c^2 - \vec{p}^2. \quad (4)$$

In a massive particle's rest frame ($\vec{p} = 0$), this reduces to $E_0^2 = c^2 p^\mu p_\mu$. Hence, it is natural to define the product in terms of the rest mass m of the particle $p^\mu p_\mu = m^2 c^2$, using $E_0 = mc^2$. This is the origin of the famous mass-energy equivalence of special relativity,

$$E^2 = \vec{p}^2 c^2 + m^2 c^4. \quad (5)$$

1.2 Quantum Mechanics

The theory of quantum mechanics provides a coherent, albeit counterintuitive description of physics at the smallest scales. The strange postulates of quantum mechanics arose out of necessity, to the chagrin of many contemporary physicists, to describe unresolved paradoxes of classical physics in the early 1900s. The origin of the name *quantum*, referring to a discrete physical quantity, comes from Max Planck's 1900 solution to the ultraviolet catastrophe of thermodynamics, in which he assumed that electromagnetic radiation could only come in discrete packets of energy [3] - specifically, an integer multiple of some fundamental constant, h , now known as Planck's constant⁴, multiplied by the frequency of radiation,

$$E = h\nu. \quad (6)$$

⁴ In SI units, this has the exact value of $h = 6.62607015 \times 10^{-34}$ Js.

It took until 1905, however, for Einstein to propose that these light quanta correspond not only to the spectral properties of atoms, but to real physical particles, whose existence could also explain the photoelectric effect [3].

A similar *quantization* was postulated by Niels Bohr in 1913 in his model of the atom in order to solve the atomic lifetime paradox of classical physics [3]. The simple application of Newtonian mechanics and Maxwell’s electromagnetism predicts that electrons orbiting a nucleus should spiral inwards due to the continual emission of electromagnetic radiation, leading to a near-instantaneous collapse of the atom. In order to solve this obvious discrepancy, Bohr posited that mechanical energy, in addition to radiation energy, is quantized. As a result, in the Bohr atomic model, only certain discretized orbitals are allowed. In 1924, Louis de Brogile interpreted this model as a standing wave solution, in which the electron is described by a wave with a whole number of wavelengths confined to the orbital [3]. This directly leads to the notion of wave-particle duality, which asserts that matter particles exhibit wave-like properties, and vice versa, under certain conditions. Generalizing the formula for Planck’s light quanta, the de Brogile wavelength of matter particles is related to the inverse of their momentum,

$$\lambda = \frac{h}{p}. \tag{7}$$

Following upon the work of Bohr and de Brogile, Erwin Schrödinger sought a three-dimensional wave equation to describe the behavior of electron waves in atoms. In 1926, he published a solution in terms of a complex-valued *wavefunction* $\Psi(x, t)$, encoding the information of a particular state, which evolves in time according to the equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = \hat{H} \Psi(x, t), \tag{8}$$

where $\hbar := h/2\pi$ and $\hat{H}$ represents the Hamiltonian of the system [3]. The meaning of the wavefunction, initially perplexing to Schrödinger himself, came to be interpreted as a probability amplitude, whose square-modulus represents a probability density for the outcome of

a particular measurement,

$$\rho = |\Psi|^2 = \Psi^\dagger \Psi. \quad (9)$$

The probabilistic interpretation inherent to quantum theory represents a striking departure from the familiar determinism of classical mechanics, introducing truly unpredictable randomness in nature.

Complementary to Schrödinger's interpretation, the wavefunction can be described by a state vector $|\psi\rangle$ belonging to a Hilbert space. The measurable properties of the system are then captured by Hermitian operators acting on the state vector. More specifically, the measurement of a given observable corresponds to the action of a time-independent operator A , the result of which is the eigenvalue a of the equation

$$\hat{A}|\psi\rangle = a|\psi\rangle, \quad (10)$$

assuming $|\psi\rangle$ is an eigenvector of A . In general, systems can be in a quantum superposition of eigenvectors, in which case the outcome of the corresponding measurement is not a priori known, and instead, is fundamentally probabilistic - depending on the square-modulus of the associated amplitudes. Moreover, given the operator description of observables, certain measurements cannot be made simultaneously if the commutation relation of their operators, given by

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}, \quad (11)$$

is nonzero. This introduces a fundamental barrier in the precision to which one can measure both observables simultaneously. For the particular case of the operators of position and momentum along the same axis, having a commutation relation of $[\hat{x}, \hat{p}_x] = i$, this results in the famous Heisenberg uncertainty principle,

$$\Delta x \Delta p_x \geq \hbar/2. \quad (12)$$

In a similar vein to special relativity, the effects of which only become noticeable at speeds comparable to the speed of light, quantum mechanical effects are practically imperceptible in the natural world owing primarily to the smallness of the constant $\hbar$ with respect to macroscopic energy scales and the effects of decoherence, which generally cause the collapse of quantum superpositions on larger scales.

1.3 Relativistic Quantum Mechanics

The theory of relativistic quantum mechanics describes the physical regime in which special relativity and quantum mechanics both play central roles. Therefore, it is natural to define the two fundamental constants of the underlying theories, the speed of light and the reduced Planck constant, as being equal to one: $c = \hbar = 1$. Using these *natural* units, energy, mass, and momentum are equivalently described in units of energy $[E] = [m] = [p]$, while space and time are placed on an equal footing with units of inverse energy $[d] = [t] = [E]^{-1}$. In this regime, units of energy are frequently expressed in multiples of electron-volt (eV), defined as the amount of energy gained by a single electron accelerating from rest through a potential difference of 1 V - equivalent to $1.602176634 \times 10^{-19}$ J in SI units.

Given the advent of the two groundbreaking theories of special relativity and quantum mechanics at the start of the 20th century, many attempts followed to unify the theories in a consistent framework. Paul Dirac, in seeking a relativistic formulation of the Schrödinger equation, arrived upon the first solution in 1928 [3]. Dirac's solution describes the relativistic quantum mechanics of massive spin-1/2 particles, and naturally describes the intrinsic spin of fermions as an inevitable outcome of unifying the two theories. The theory also provided the first prediction of antiparticles - interpreted as a set of complementary, negative energy solutions to the positive energy states. Validating Dirac's efforts, the antiparticle of the electron, known as the positron, would be discovered a mere four year later in experiments by Carl Anderson [3].

Expressed in its covariant form in terms of the four-component wavefunction ψ , known

as a Dirac spinor, and the γ -matrices, the Dirac equation is

$$(i\gamma^\mu\partial_\mu - m)\psi = 0, \tag{13}$$

where $\gamma^\mu := (\gamma^0, \gamma^1, \gamma^2, \gamma^3)$. The four-vector probability current $j^\mu = (\rho, \vec{j})$ of a spinor as expressed in a Lorentz covariant form is

$$j^\mu = \bar{\psi}\gamma^\mu\psi, \tag{14}$$

where the *adjoint spinor* is defined, $\bar{\psi} := \psi^\dagger\gamma^0$. This satisfies the continuity equation expressing the conservation of particle probability, $\partial_\mu j^\mu = 0$. Though relatively simple in its original form, the Dirac equation describes the basis of much of modern particle physics, as will be shown in the following sections.

2 The Standard Model

2.1 Constituents

The Standard Model (SM) of particle physics describes the properties of and interactions between the elementary constituents of the Universe. The theory consists of two classes of particles: the integer-spin bosons and the half-integer spin fermions - respectively obeying Bose-Einstein and Fermi-Dirac statistics. Generally speaking, bosons are *force-carrying* particles, while fermions are *matter* particles. The seventeen elementary particles are depicted schematically in Figure 1 and summarized in detail in Table 1. The scalar Higgs boson is the singular particle of the Higgs field, a crucial component of the theory necessary to ascribe masses to particles. The four gauge bosons mediate the three fundamental forces of the theory: the strong, electromagnetic, and weak interactions. There are three generations of fermions sharing similar properties, each consisting of two quarks and two leptons. Each lepton generation consists of an electrically charged particle and a neutral partner, known as a neutrino. For each fermion there exists a corresponding antiparticle with identical properties, save for negated gauge charges. Together, these twelve fermionic particles make up all of the observable matter in the Universe. Practically speaking, though, the entirety of atomic matter consists of only the first generation of fermions: the up and down quarks - together forming protons (uud) and neutrons (udd) - and the electron.

Each fundamental force acts upon only those particles which are charged under the corresponding gauge interaction, as outlined in Table 2, and the strength of the interaction is governed by a gauge coupling parameter. The strong force, mediated by gluons, affects particles with color charge - that is, the six quarks and gluons themselves. The electromagnetic force, mediated by the photon, couples to electrically charged particles. Lastly, the weak force, mediated by the charged $W^\pm$ and neutral Z bosons, affects particles charged under weak isospin - that is, all of the fermions, the W -boson itself, and the Higgs boson.

Table 1: Summary of the elementary bosons and fermions of the Standard Model. Masses are taken from the 2020 summary tables provided by the Particle Data Group [4]. Electric charges are given in units of elementary charge for the particle and antiparticle pairs.

Category		Particle	Symbol	Spin	Charge	Mass
Bosons	Scalar	Higgs	H	0	0	125.25 GeV
	Gauge	Gluon	g	1	0	0
		Photon	γ	1	0	0
		Z-boson	Z	1	0	91.19 GeV
		W-boson	$W^\pm$	1	± 1	80.38 GeV
Fermions	Quarks	Up	u	1/2	$\pm 2/3$	2.16 MeV
		Down	d	1/2	$\mp 1/3$	4.67 MeV
		Charm	c	1/2	$\pm 2/3$	1.27 GeV
		Strange	s	1/2	$\mp 1/3$	93 MeV
		Top	t	1/2	$\pm 2/3$	173 GeV
		Bottom	b	1/2	$\mp 1/3$	4.18 GeV
	Leptons	Electron	e	1/2	∓ 1	0.511 MeV
		Muon	μ	1/2	∓ 1	105.66 MeV
		Tau	τ	1/2	∓ 1	1776.86 MeV
		e -neutrino	ν_e	1/2	0	< 1.1 eV
		μ -neutrino	ν_μ	1/2	0	< 0.19 MeV
		τ -neutrino	ν_τ	1/2	0	< 18.2 MeV

Table 2: Summary of fermion interactions under the three fundamental forces of the Standard Model. A checkmark indicates that the particles are charged under the corresponding gauge interaction and therefore participate in the associated force, while a cross denotes otherwise.

Particles		Strong	Electromagnetic	Weak
Quarks	$u, c, t,$ d, s, b	✓	✓	✓
Leptons	e, μ, τ	×	✓	✓
	ν_e, ν_μ, ν_τ	×	×	✓

Standard Model of Elementary Particles

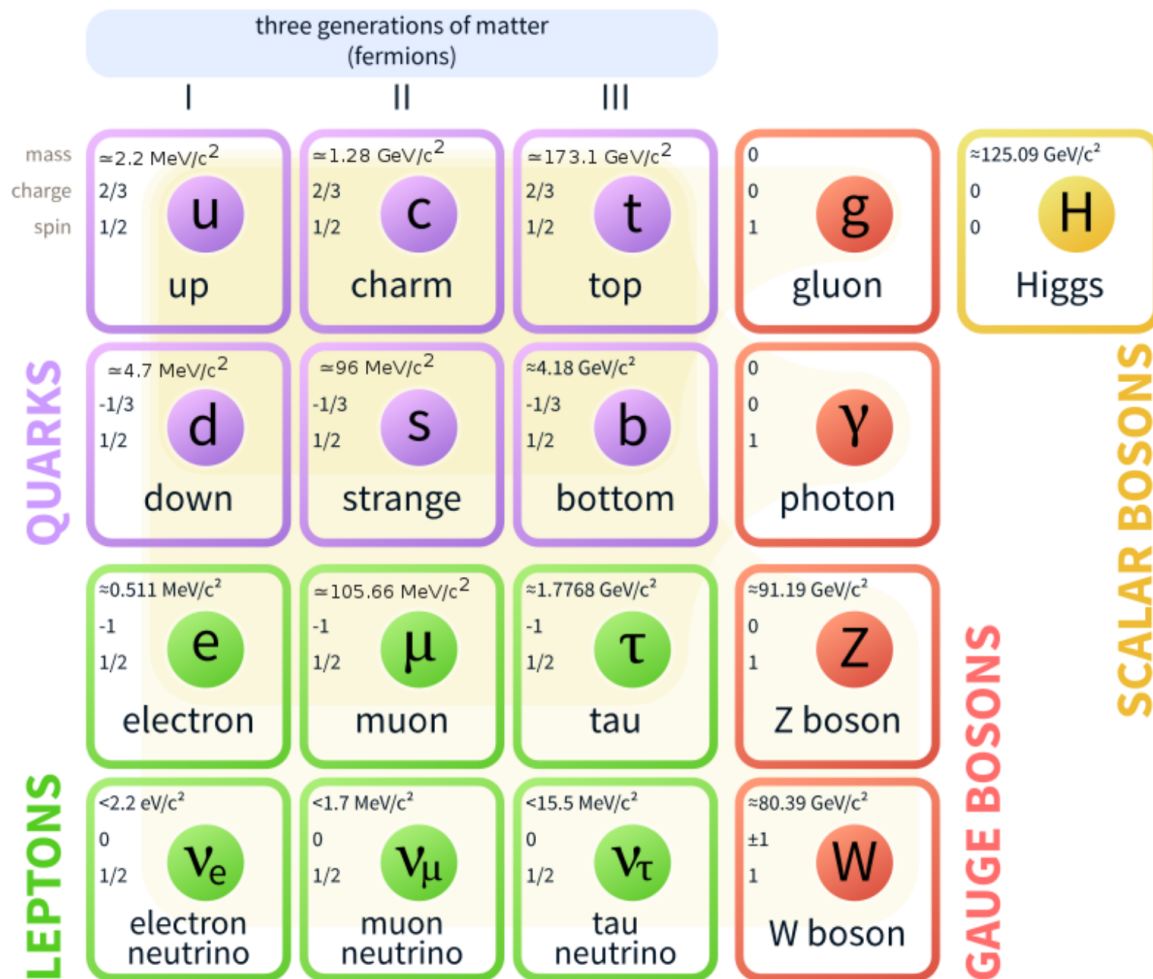


Figure 1: The elementary particles of the Standard Model, including: the twelve fermions, the four gauge bosons, and the scalar Higgs boson. The six quarks and leptons can be arranged in three generations based on similar properties. Taken from *Wikipedia*.

2.2 Theoretical Underpinnings

The underlying physics of the Standard Model is based upon the framework of relativistic quantum field theory. In this picture, each particle is fundamentally an excitation in an underlying quantum field. Following the Lagrangian formalism of classical field theory, the entirety of the dynamics can be encapsulated in the form of a Lagrangian⁵. For the Standard Model, this can be condensed to a single line containing the boson and fermion fields of the theory,

$$\mathcal{L}_{\text{SM}} = \underbrace{-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{gauge boson interactions}} + \underbrace{i\bar{\psi}\gamma^\mu D_\mu\psi}_{\text{fermion interactions}} + \underbrace{\bar{\psi}_i y_{ij} \psi_j \phi + h.c.}_{\text{Higgs-fermion couplings}} + \underbrace{|D_\mu\phi|^2}_{\text{Higgs-gauge couplings}} - \underbrace{V(\phi)}_{\text{Higgs self-couplings}}, \quad (17)$$

where $h.c.$ denotes the Hermitian conjugate of the previous term. The first term, proportional to the scalar product of the field strength tensor, describes the dynamics of the force-carrying gauge bosons. The interaction of fermionic matter particles, as mediated by the gauge bosons, is encapsulated in the second term. The remaining terms describe the Higgs field ϕ and its interaction with other particles of the Standard Model. The third and fourth terms generate mass for the fermions and antifermions, respectively, based on their Yukawa coupling y_{ij} to the Higgs field. The fifth term describes the interaction of the gauge bosons with the Higgs field, which, as will be discussed, provides an explanation for the

⁵ For practical reasons, this is represented as a Lagrangian *density*, denoted with a curly $\mathcal{L}$, which, when integrated over all of space, returns the Lagrangian, denoted by L . Therefore, the integral of the Lagrangian density over all of spacetime yields the familiar action S ,

$$S = \int L dt = \int \mathcal{L} d^4x. \quad (15)$$

As in classical mechanics, the equations of motion may be derived by solving for the extremum of the action, equivalent to applying the Euler-Lagrange equations,

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = 0. \quad (16)$$

non-zero masses of the W and Z bosons. Lastly, the final term describes the Higgs potential itself and its allowed self-interactions.

The three forces described by the Standard Model each correspond to an underlying gauge field, in which the Lagrangian remains invariant under a set of local phase transformations belonging to a Lie group. The quantization of a classical gauge field describes the resulting quanta of the field as gauge bosons, and the Lagrangian is said to obey the symmetry group corresponding to the transformation. According to Noether's theorem, a symmetry of the Lagrangian also implies the existence of a conserved quantity. For reasons to be discussed, the Standard Model as a whole is said to obey a $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry structure. The following discussion of gauge theories in the Standard Model is largely inspired by the treatment in Mark Thomson's *Modern Particle Physics* [5].

Quantum Electrodynamics

Historically, Quantum Electrodynamics (QED) was the first successful relativistic quantum field theory, and describes the simplest of the Standard Model interactions - the electromagnetic force. Fundamentally, QED describes the interaction of spin-1/2 electrons and positrons as mediated by the massless, spin-1 photon. Delving into its gauge structure, QED is said to obey the $U(1)$ symmetry group - the real Lie group of unitary ($U^\dagger U = 1$) matrices of dimension one. The $U(1)$ group consists of all complex numbers with absolute value one, and reflects the arbitrary choice in phase at every point in space. This symmetry necessitates the existence of the massless, spin-1 photon and the conservation of electric charge, Q . Because the generators of this symmetry group commute, the theory of QED is said to be an Abelian gauge theory - which implies that photons carry no electric charge and, hence, cannot directly self-interact.

The Lagrangian of QED can be derived by imposing local $U(1)$ gauge invariance on the Dirac Lagrangian. In its Lagrangian form, the Dirac equation (see Equation 13) can be

written as

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi. \quad (18)$$

This is clearly invariant under a *global* $U(1)$ phase transformation α , given the definition of the adjoint spinor, $\bar{\psi} := \psi^\dagger\gamma^0$. To meet the criteria of a $U(1)$ gauge theory, though, the Lagrangian of QED must remain invariant under *local* phase transformations of the form

$$\psi(x) \rightarrow \psi'(x) = e^{iq\alpha(x)}\psi(x), \quad (19)$$

This is only possible by including an additional gauge-fixing term - found by ansatz to be

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - q\bar{\psi}\gamma^\mu A_\mu\psi, \quad (20)$$

where the newly introduced field A_μ is required to transform by

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\alpha. \quad (21)$$

The new term in the Lagrangian describes the interaction of the A_μ field with the fermions of the theory.

Given its transformation properties, the field of A_μ exactly corresponds to that of the photon, and is related to the electromagnetic field tensor by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (22)$$

The photon kinetic term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ describing Maxwell's free field equations - also invariant under local phase transformations - may also be inserted in Equation 20. Substituting the electric charge ($q = -e$) into the rightmost term, gives the three-point interaction vertex of QED with coupling strength e . For practical purposes, the gauge-fixing interaction term can be absorbed into the definition of the *covariant derivative*, D_μ - as done in Equation 17. In

QED this is simply

$$D_\mu = \partial_\mu + ieA_\mu, \quad (23)$$

allowing for the expression of the Lagrangian in a manifestly gauge invariant form,

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi. \quad (24)$$

The continuity equation of Dirac's theory is still valid, now interpreted as mandating the conservation of electric charge - a direct result of the global $U(1)$ gauge symmetry of QED.

From Equation 24, the entirety of electromagnetic phenomena may be derived by solving for the equations of motion and imposing the Lorenz gauge condition, $\partial_\mu A^\mu = 0$. Given the small dimensionless coupling strength, α , of QED⁶, the probability of particle decay rates and interactions may be calculated as the square-modulus of a perturbative series in powers of α of transition matrix elements. Higher order terms in the perturbative expansion correspond to insertions of additional interaction vertices, and are therefore suppressed over lower order processes according to the size of the dimensionless coupling. The unprecedented accuracy of QED is due to the small size of α , since higher order terms in the expansion quickly become negligible and the infinite series rapidly converges. This is not generally true for other gauge field theories.

Quantum Chromodynamics

The quantum field theoretic description of the strong force, known as Quantum Chromodynamics (QCD), is invariant under local $SU(3)$ transformations - corresponding to the group of special ($\det(U) = 1$), unitary matrices of dimension three. The generators of the $SU(3)$ group are the set of eight traceless, Hermitian 3×3 matrices - known as the Gell-Mann matrices, λ^a . Analogously to QED, the theory of QCD may be derived by imposing local

⁶ Small here refers to the size of the dimensionless gauge coupling constant in relation to unity. For QED, this is the familiar fine structure constant - expressed in natural units as $\alpha = e^2/4\pi \approx 1/137$.

$SU(3)$ gauge invariance on the Dirac Lagrangian,

$$\psi(x) \rightarrow \psi(x)' = e^{ig\alpha^a(x)T^a} \psi(x), \quad (25)$$

where $T^a = \frac{1}{2}\lambda^a$ are the eight generators of the $SU(3)$ group. Imposing this condition leads to eight massless, spin-1 gluons and a trifold, conserved color charge, C . The non-commutation of the group generators gives rise to gluon self-interactions in the form of triple and quartic vertices. For this reason, the theory of QCD is said to be non-Abelian.

Following the procedure demonstrated in QED, the gluon field is completely specified by the mandate of local $SU(3)$ gauge invariance. The gluon field strength tensor $G_{\mu\nu}^a$, analogous to the field strength tensor of electromagnetism, is defined in terms of the gluon field $\mathcal{A}_\mu^a$ by

$$G_{\mu\nu}^a = \partial_\mu \mathcal{A}_\nu^a - \partial_\nu \mathcal{A}_\mu^a + gf^{abc} \mathcal{A}_\mu^b \mathcal{A}_\nu^c, \quad (26)$$

where f^{abc} are the structure constants of $SU(3)$ and the gluon indices (a, b, c) run from one to eight. The dynamics of QCD are much complicated over QED due to the presence of eight force-carrying bosons and the nature of gluon self-interactions - represented by the rightmost term in the field tensor. The interaction terms of the theory, necessary to maintain local invariance under $SU(3)$ phase transformations, are absorbed into the definition of the covariant derivative,

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} - ig(T^a)_{ij} \mathcal{A}_\mu^a. \quad (27)$$

Written in its gauge invariant form in terms of the quark field ψ , the QCD Lagrangian is

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} + \bar{\psi}_i (i\gamma^\mu (D_\mu)_{ij} - m\delta_{ij}) \psi_j. \quad (28)$$

The theory of QCD as outlined here leads to several important consequences. Gluon self-interactions are hypothesized to account for the phenomenon of *color confinement*, in which colored states of quarks and gluons are not observed to exist in nature, instead confined to

colorless bound states. These bound states are known as hadrons - formed either by quark-antiquark pairs (mesons) or combinations of three quarks (baryons) in a color-neutral state. This has the important effect of hadronization, whereby isolated color states of quarks or gluons must quickly coalesce into colorless hadrons. Further complicating the analogy to QED, the dimensionless coupling strength of the strong interaction, α_S , is of order unity at low-energy scales. Consequently, the perturbative expansion employed in QED is no longer valid and alternative calculation techniques must be pursued. This coupling constant is, however, not constant - instead scaling and, crucially, decreasing with the energy scale. Beyond energies of 100 GeV, the coupling is sufficiently small enough, of order $\mathcal{O}(0.1)$, to allow for perturbative calculations - this result is known as *asymptotic freedom* [5].

Electroweak Unification

The quantum field theoretic formulation of the charged-current weak interaction rests upon the $SU(2)$ symmetry group - corresponding to the group of special, unitary matrices of dimension two. The charged weak interaction is the only means by which fermions can change flavor and, interestingly, violates parity symmetry - only coupling to left-handed (L) particles and right-handed (R) antiparticles. Analogously to QCD, the generators of the $SU(2)$ group are the set of three traceless, Hermitian 2×2 matrices - known as the Pauli matrices, σ^a . Here again, the generators do not commute so the theory is said to be non-Abelian, and interactions between the gauge bosons necessarily arise. Imposing local $SU(2)$ gauge invariance on the Dirac Lagrangian, requires the definition of the covariant derivative to be

$$D_\mu = \partial_\mu + igT^a W_\mu^a, \quad (29)$$

where $T^a = \frac{1}{2}\sigma^a$ and W_μ^a are the three gauge fields of the symmetry. This requirement leads to three spin-1 gauge bosons and the conservation of weak isospin, T_3 . As before, the gauge bosons are required to be massless by the mandate of local gauge invariance. While true for QED and QCD, the weak gauge bosons have been experimentally observed to carry non-zero

mass. Ascribing mass to these particles is possible by unifying the electromagnetic and weak interactions, and subsequently *breaking* the symmetry - in a process known as spontaneous symmetry breaking.

As discovered by Stephen Weinberg [6] Abdus Salam [7], and Sheldon Glashow [8], the electromagnetic and weak interactions can be unified in a single *electroweak* (EW) interaction based on a combined $SU(2)_L \times U(1)_Y$ symmetry. This requires replacing the $U(1)_Q$ gauge symmetry of QED with one of weak hypercharge $U(1)_Y$, where the relevant conserved quantity is related to the electric charge and weak isospin by $Y = 2(Q - T_3)$. The three $SU(2)_L$ weak isospin fields are denoted W_μ^1 , W_μ^2 , and W_μ^3 and the single $U(1)_Y$ field of weak hypercharge, B_μ . According to the electroweak theory, the differences in the electromagnetic and weak interactions are only apparent at low energies below the unification threshold.

Following electroweak unification, the Higgs mechanism [9–11] is responsible for spontaneous symmetry breaking of the electroweak symmetry, which ascribes masses to the weak gauge bosons while keeping the photon massless. By acquiring a non-zero vacuum expectation value, the Higgs field must orient itself along some direction in field space. In doing so, rotations about this axis of orientation preserve the vacuum expectation value and the corresponding gauge group remains unbroken, while rotations about the other axes do not preserve the vacuum, meaning the corresponding symmetries are broken. The photon remains massless precisely because the particular linear combination of electric charge does not couple to the Higgs boson, while the masses of the weak bosons arise from their coupling to the Higgs field. The Higgs mechanism also separately generates masses for the charged fermions of the Standard Model based on their Yukawa couplings to the Higgs field. The charged $W_\mu^\pm$ boson fields are formed from linear combinatorics of the W_μ^1 and W_μ^2 weak isospin fields,

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2). \quad (30)$$

Meanwhile, the neutral Z_μ and A_μ bosons fields, respectively governing the weak neutral-current and electromagnetic interactions, are similarly obtained from a linear combination

of the W_μ^3 and B_μ fields. Defining the Weinberg mixing angle, θ_W , by

$$\sin \theta_W = \frac{g_Y}{\sqrt{g_W^2 + g_Y^2}}, \quad (31)$$

in terms of the gauge couplings of the $SU(2)_L$ weak isospin (g_W) and $U(1)_Y$ weak hypercharge (g_Y) interactions, this can be succinctly written as a rotation in the $W_\mu^3 - B_\mu$ plane by θ_W ,

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}. \quad (32)$$

The unification of the electromagnetic and weak interactions has many important consequences for the structure of the theory. The Weinberg mixing angle specifies a mathematical relation between the masses of the W and Z bosons,

$$m_Z = \frac{m_W}{\cos \theta_W}, \quad (33)$$

both of which are in turn specified by the two gauge couplings of the theory and the Higgs vacuum expectation value v ,

$$m_Z = \frac{1}{2}v\sqrt{g_W^2 + g_Y^2}. \quad (34)$$

The $U(1)_Q$ gauge coupling of electromagnetism can be recovered as

$$e = g_W \sin \theta_W. \quad (35)$$

Given their similar field description, the Z -boson and photon participate in many of the same interactions, though, owing to the mass of the Z -boson, its contribution is suppressed at lower energies and dominates for energies near m_Z . The neutral gauge bosons do importantly differ in that the Z -boson is predicted to couple to left- and right-handed fermions with unequal coupling strengths - while the photon does so equally. At short distance scales, the weak

interaction has a comparable strength to the electromagnetic force, though this decreases rapidly on larger scales, allowing for the perturbative calculation of probability amplitudes as in pure QED. The accuracy afforded by electroweak calculations, in contrast to those of QCD, is central to many precision tests of the Standard Model.

2.3 Outstanding Problems

The Standard Model is the world’s best tested and experimentally verified theory ever devised⁷. Over the course of several decades, experiments have repeatedly verified many of its predictions to an extreme accuracy. However, despite its many successes, the theory is almost certainly incomplete. Most notably, gravity is entirely left out of the theory and, furthermore, attempts to reconcile quantum mechanics with Einstein’s theory of gravity, known as general relativity, have continually failed. Additionally, many astronomical observations point to the existence of an unknown, invisible form of matter outnumbering normal matter by a factor of five. The Standard Model provides no such description of this so-called *dark matter*. More recently in the realm of particle physics, experiments have found suggestive hints of new physics in the form of deviations in the muon magnetic moment [13] and possible violations of lepton flavor universality [14].

Outside of unexplained experimental discrepancies, there are also concerns regarding the structure of the theory itself. Hidden in Equation 17, there are 19 free parameters which are completely unspecified by the theory, including the 9 charged fermion masses, 3 gauge couplings, and the Higgs mass and vacuum expectation value. Many of these parameters must be fine-tuned to match experimental observations, which clashes with the theoretical ideal of naturalness. Although the importance of naturalness in constructing theories is contested, there are many features of the Standard Model which suggest a deeper explanation exists. The mass of the charged fermions spans a large, yet unexplained scale of over 10^5 between

⁷ Famously, the measurement of the electron magnetic moment has been measured to an accuracy of better than one part per trillion [12], and agrees with the QED prediction up to 10 significant figures, making it the most accurate verification of a physical theory to date.

the lightest and heaviest particles - respectively, the electron and top quark. Concerning the fundamental forces, there is an enormous and unexplained difference in their effective coupling strengths - an outstanding question known as the *hierarchy problem* - and it is unknown whether the strong and electroweak interactions can be unified.

3 Testing the Standard Model

Considering the outstanding problems, the search for *new physics* beyond-the-Standard Model (BSM) is essential. Doing so requires not only measuring the interactions of elementary particles to a high precision, but, more importantly, the expertise in knowing which processes of the Standard Model offer the best window into the search for new physics.

3.1 Collider Principles

A particle collider is an experimental apparatus designed to probe fundamental physics interactions through the collision of high-energy elementary or composite particles. In simple terms, colliders operate on two fundamental principles borrowed from special relativity and quantum mechanics - respectively, mass-energy equivalence and wave-particle duality. Firstly, according to the mass-energy equivalence of special relativity, the energy of colliding particle beams can be converted into mass-energy and momentum in the form of outgoing particles. Collision energies of several hundred GeV and above allow for the creation of all known Standard Model particles, the most massive of which are very unstable, virtually nonexistent in nature, and decay nearly instantaneously. Measuring the kinematics and properties of the outgoing decay particles, by means of sophisticated detector technologies, provides insight into the underlying physics interaction itself. Secondly, according to the quantum mechanical principle of wave-particle duality encapsulated by the de Broglie wavelength, increasing the energy of colliding beams effectively decreases the distance scale being probed. For instance, an energy of 1 TeV corresponds to a distance scale of 10^{-19} m - roughly ten thousand times smaller than the radius of a proton [4]. At such an energy, the collision of composite hadrons, effectively probes parton interactions of the constituent quarks and gluons, which are otherwise inaccessible at lower energies. High energy particle colliders therefore provide the optimal setting for experimental tests of the Standard Model.

Although a particle collider is a complex apparatus characterized by many properties, two

measures stand out in their importance: the center-of-mass energy, $\sqrt{s}$, and instantaneous luminosity, $\mathcal{L}$. The center-of-mass energy is equivalent to the square root of the Lorentz-invariant Mandelstam variable s , which itself is given by the scalar product of the combined four-momentum of particles p_1 and p_2 , $s = (p_1 + p_2)^2$. This is directly related to the energy reach of the collider, providing an upper limit on the mass of accessible heavy resonances. The instantaneous luminosity is a measure of the density of the colliding beams, having units of events per area per time $[d]^{-2}[t]^{-1}$, and instead determines the statistical reach of the collider - which is to say, the amount of collision data possible to obtain in a fixed time. For a given process, the luminosity is related to the expected event count, N , by

$$N = \sigma \int \mathcal{L} dt, \tag{36}$$

where σ represents the *cross-section* of the particular process. Following dimensional analysis, the cross-section must have units of area $[d]^2$ and relates to the probability of the specific process to occur. In high energy particle physics, cross-sections are typically expressed in units of inverse barns, b^{-1} , where one $1 \text{ b} = 10^{-28} \text{ m}^2$.

3.2 Diboson Physics

Diboson measurements are a crucial pillar of collider physics, complementary to the Higgs and top quark sectors. Measurements of the W and Z bosons are particularly suited for precision tests of the Standard Model for a variety of reasons, discussed in detail in Ref. [15]. The massive bosons are the most strongly coupled particles to electroweak symmetry breaking, and therefore provide insight into one of the more interesting, though less studied aspects of the theory. Moreover, measurements of the electroweak bosons allow for the direct testing of non-Abelian gauge theories, in a way not possible in QCD given the nature of color confinement. The non-Abelian nature of the electroweak theory gives rise to vector boson self-interactions in the form of triple and quartic gauge boson couplings, which offer an

unmatched sensitivity to higher energy physics. Well-known examples of these interactions include the vector boson fusion and vector boson scattering processes. From an experimental standpoint, decays of the electroweak bosons can be reconstructed in a variety of clean final states, particularly those involving leptons.

Diboson interactions, that is, those mediated by two massive gauge bosons, are suppressed over single boson interactions, but provide some additional advantages in the search for new physics. The higher energy scale of diboson processes is generally more sensitive to various models of new physics, particularly heavy resonances and effective field theory interpretations. Additionally, the clear signature of, in particular, leptonic decay channels provides excellent discrimination from background sources, allowing for high precision measurements. However, the production of two or more heavy gauge bosons in a collision requires a high energy collider, in the range of several hundred GeV and above. Modern hadron colliders, which are currently capable of reaching energies beyond 10 TeV, therefore offer the ideal ground for studying processes mediated by the exchange of these heavy particles.

3.3 Four-lepton Processes

Considering multi-boson processes, the four-lepton decay channel is the cleanest signature available, in which all decay products can be fully reconstructed. The specific signature is that of four *prompt* leptons consisting of two same-flavor, opposite-charge (SF-OC) lepton pairs. Here, prompt leptons refer to those emanating directly from the hard scatter interaction, in contrast to non-prompt leptons which originate from secondary decays further from the interaction point. The process is typically mediated through the exchange of two intermediary Z -bosons (or photons⁸), each of which decaying to a pair of SF-OC leptons. In this specific decay, the Z -boson, being a neutral gauge boson, must decay into same-flavor, opposite-charge leptons to conserve lepton-number and electric charge, respectively. Within the Standard Model, the charged leptonic Z -boson decay, $Z \rightarrow \ell^+ \ell^-$, has a branching frac-

⁸In many processes, photons may participate in the same interactions as Z -bosons, though this is often left unstated for the sake of simplicity.

tion of approximately 10%, with equal contributions from each lepton flavor [4]. Therefore, roughly 1 in 100 SM ZZ events will decay to the four-lepton final state.

Standard Model Contributions

In proton-proton colliders, the contribution from various SM production modes varies depending on the invariant mass of the four-lepton system⁹. The representative Feynman diagrams depicting the primary processes are shown in Figure 2. The dominant, QCD-driven production mode, $q\bar{q} \rightarrow 4\ell$, shown in Figure 2 (a), occurs via a quark-induced t -channel interaction, yielding a pair of intermediary Z -bosons, which themselves decay leptonically. Near the mass of the Z -boson, the production is dominated by single Z resonance, with final state radiation from one of the outgoing leptons producing the secondary lepton pair, as shown in Figure 2 (b). Gluon-induced production via an intermediary quark loop, $gg \rightarrow 4\ell$, as in Figure 2 (c), contributes at next-to-next-to-leading order in the strong coupling α_S relative to the quark-induced process. Near the Higgs mass, contributions from resonant Higgs mediated s -channel production, $H \rightarrow 4\ell$, become significant, as shown for the dominant gluon-gluon fusion production mode in Figure 2 (d). Additional SM processes contributing to the four-lepton final state include gauge boson production in association with a pair of top quarks ($t\bar{t}V$), with both tops decaying semileptonically, and various modes of triboson production (VVV). These processes contribute mainly at higher energy given the large masses of the intermediary particles.

⁹ Kinematics of the four-lepton final state are defined in terms of the four-momenta of the individual leptons $p_\ell^\mu = (E_\ell, \vec{p}_\ell)$ forming the quadruplet. The total four-momentum of the four-lepton system is simply the vector sum of the quadruplet lepton four-momenta,

$$p_{4\ell}^\mu = \sum_{\ell=1}^{\ell=4} p_\ell^\mu = \left(\sum_{\ell=1}^{\ell=4} E_\ell, \sum_{\ell=1}^{\ell=4} \vec{p}_\ell \right) = (E_{4\ell}, \vec{p}_{4\ell}). \quad (37)$$

The four-lepton invariant mass is then related to the scalar product of the total four-momentum by

$$m_{4\ell}^2 = (p^\mu p_\mu)_{4\ell} = E_{4\ell}^2 - \vec{p}_{4\ell}^2. \quad (38)$$

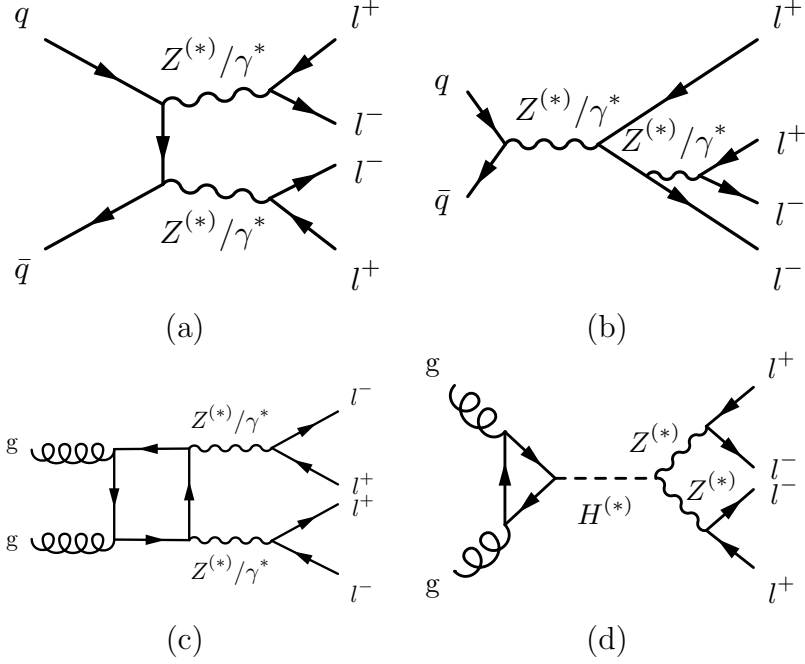


Figure 2: Representative leading order Feynman diagrams for the dominant modes of Standard Model four-lepton production at high energy hadron colliders. The superscript of $*$ denotes a particle that is off-shell, while $(*)$ indicates the particle can be either on-shell or off-shell. In (a), the quark-initiated t -channel process produces a pair of intermediary Z -bosons, which themselves decay leptonically yielding two same-flavor, opposite-charge lepton pairs. In (b), resonant Z production along with internal conversion of a secondary Z in the final state gives the four-lepton signature. In (c), gluon-initiated production via an intermediary quark loop contributes at higher order in QCD. In (d), resonant Higgs production leads to the four-lepton final state, once again through two intermediary Z -bosons.

Beyond-the-Standard Model Contributions

Several scenarios of new physics could either directly, through leptonic decays of BSM particles, or indirectly, by means of BSM-SM interference, affect four-lepton production. Examples of direct contributions are shown in Figure 3. Considering the direct contributions, many BSM models introduce a heavy Z' gauge boson, corresponding to a broken $U(1)'$ symmetry, which decays in a manner similar to the SM Z -boson [16]. This includes perhaps the best-motivated scenario of WIMP-based (weakly-interacting massive particle) thermal dark matter, in which dark matter particles couple to the SM through a Z' -boson [17]. Additionally, various modes of supersymmetry (SUSY) - an extension to the SM predicting the existence of fermion (boson) pairs to the bosons (fermions) of the SM - can result in signatures with high charged lepton multiplicities. This fact has already been used in ATLAS [18] to place constraints on SUSY models using four-lepton events. Furthermore, several extensions of the SM predict the existence of exotic Higgs particles, often in the form of doubly-charged scalar bosons [19]. Pair production of these charged Higgs bosons, which may decay leptonically into a pair of same-charge leptons, would also contribute to the four-lepton final state.

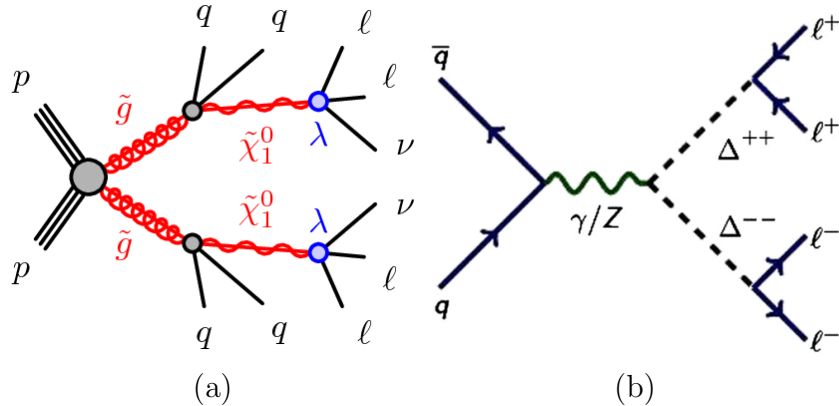


Figure 3: A sample of leading order Feynman diagrams for potential modes of beyond-the-Standard Model four-lepton production at high energy hadron colliders. In (a), SUSY gluino pair production followed by decays to the lightest SUSY particle, $\tilde{\chi}_1^0$ - assumed to always decay by $\tilde{\chi}_1^0 \rightarrow \ell\ell\nu$ - gives a final state with four charged leptons [18]. In (b), the pair production of doubly-charged Higgs bosons, each decaying into same-charge lepton pairs, can also give rise to the four-lepton signature [19].

Considering indirect effects, new physics could enter, more subtly, by means of higher energy processes interfering with SM processes in the production of four-leptons. This interference would manifest itself in the form of slight modifications to SM couplings or deviations in the high energy tails of differential cross-sections. Effective field theory (EFT) interpretations - to be discussed in more detail in Section 22 - provide a robust, model-independent framework for parameterizing such effects, thereby mapping the space of possible new physics contributions. In this regard, the four-lepton signature is an interesting candidate to study because of the large energy scale involved in diboson processes and the precision offered by charged lepton reconstruction. These measurements serve an important role in the pursuit of a global EFT interpretation - where the combination of complementary diboson, Higgs, and top quark measurements across many analyses would improve the search sensitivity over any single analysis. Many hurdles remain towards this goal, though smaller scale combinations, as demonstrated in Ref. [20], have already proved effective.

Chapter III: Apparatus

4 CERN

The nature of experimental high energy particle physics requires probing shorter and shorter distance scales, corresponding to higher and higher collision energies. This requires the construction of ever-larger particle colliders capable of pushing forward in the energy frontier. The feasibility of cutting-edge research is therefore only possible through international collaboration at a massive scale. The European Organization for Nuclear Research, better known by the acronym CERN¹⁰, in Geneva, Switzerland perfectly encapsulates this spirit of collaboration. Proposed by leading European scientists in the aftermath of the Second World War and signed into existence in June of 1953 [21], the CERN Convention called for an international alliance in the burgeoning field of nuclear physics research. A central European laboratory would not only serve the practical purposes of pooling resources and talent in one location, but also provide a means of mending the fraught political landscape. Construction of the laboratory began two years later in the countryside surrounding Geneva. Over the course of several decades, the focus of the research shifted towards higher energies in order to probe interactions between subatomic particles. As the experiments grew in complexity, so too did the collaborative network, now extending far beyond the European continent. Today, the collider and detector facilities at CERN are home to the world-leading technologies in elementary particle physics research.

¹⁰The acronym derives from the name of the provisional body established with founding the laboratory - Conseil Européen pour la Recherche Nucléaire.

5 The Large Hadron Collider

Halfway through the operation of the Large Electron Positron Collider (LEP) at CERN (1989-2000 [22]), the CERN Council approved the more powerful successor in December of 1994 [23]. This new collider, known as the Large Hadron Collider (LHC), would collide hadrons, namely protons or heavy-ions, at instantaneous luminosities up to $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and center-of-mass energies up to $\sqrt{s} = 14 \text{ TeV}$, far exceeding Fermilab's Tevatron record of 0.98 TeV and making it the world's most powerful particle accelerator. Construction of the LHC finished in 2008, made significantly faster by reusing the LEP tunnel system - a 27 km ring located outside of Geneva, Switzerland and buried at an average depth of 100 m underneath the French-Swiss border (see Figure 4). The first official run of data-taking (Run-1) started two years later in 2010 at a center-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$, and continued through 2012 reaching energies up to 8 TeV. Following a series of planned upgrades beginning in 2013, the collider began its second run of data-taking (Run-2) in 2015, which concluded in late 2018. Data collected during this second period consisting of proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ serve as the foundation of this thesis.

Shown schematically in Figure 5, the LHC is the largest of the accelerators in the CERN accelerator complex. A series of smaller accelerators responsible for incrementally increasing the beam energies ultimately feed the primary ring. Protons in the collider begin their journey as hydrogen gas, which is housed in tanks and stripped of orbital electrons by means of ionization. Beams of protons are formed in the linear accelerator LINAC 2, operational since 1978, and injected into the circular PS Booster at an energy of 50 MeV [23]. First accelerated to an energy of 1.4 GeV, beams are then injected into the Proton Synchrotron and further accelerated to energies of 25 GeV [23]. The beams are then fed into the Super Proton Synchrotron and accelerated up to 450 GeV [23] before being bidirectionally transferred to the main LHC ring. Over the course of many passes in the radio frequency (RF) acceleration

cavities, the opposing beams ultimately reach the desired center-of-mass energy and are maintained for several hours of data-taking. During Run-2 of data-taking, the individual beams were accelerated to a final energy 6.5 TeV, yielding the previously quoted center-of-mass energy of $\sqrt{s} = 13$ TeV.

As a result of the radio frequency acceleration scheme, beams are discretized in several thousand evenly-spaced bunches, each consisting of some 10^{11} protons [23]. Each RF cavity is tuned to oscillate at a frequency of 400 MHz, accelerating or decelerating only those protons arriving later or earlier, respectively. Under peak luminosity conditions, beams circulate with a typical inter-bunch spacing of 25 ns, corresponding to a collision frequency of approximately 40 MHz [23] at the four interaction points. Circular bending of the charged hadron beams requires enormous magnetic fields throughout the chain of colliders. In the primary LHC ring, some 1,232 superconducting dipole magnets supply nominal magnetic fields of 8 T to provide this continuous bending [23]. Focusing of the beams in the transverse plane requires an additional 392 higher order main quadrupole magnets [23]. Operation of the magnets is dependent upon liquid helium cooling down to 1.9 K [23] in order to maintain the superconducting state. This requires the operation of two vacuum systems to house the cryomagnets and the helium distribution lines, in addition to the ultrahigh beam-line vacuum of 10^{-13} atm [23] necessary to avoid undesired interactions with air molecules.



Figure 4: An aerial image of CERN and the surrounding Geneva area, featuring Lac Léman and Mont Blanc in the background. The primary LHC ring is superimposed in yellow, while the feeder accelerators are shown in the upper right of the image. The ATLAS detector, situated at Interaction Point 1, is highlighted in the top right (Credit: CERN).

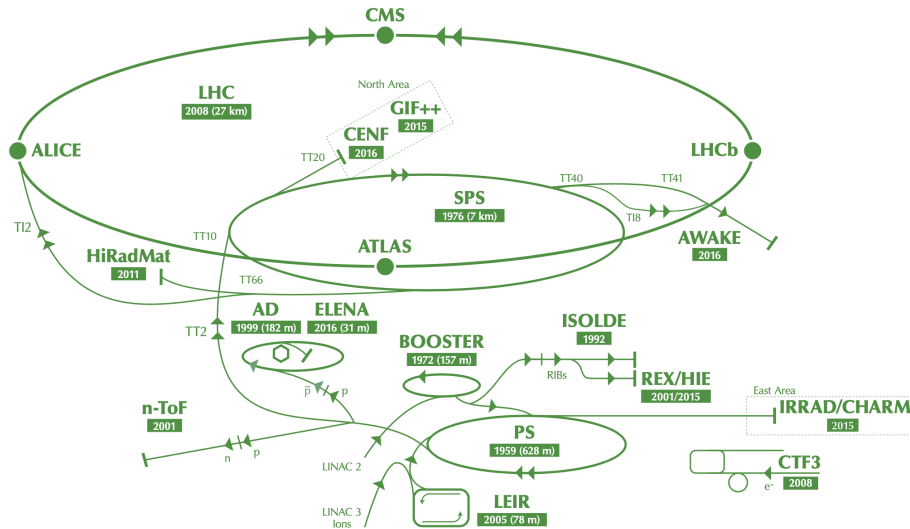


Figure 5: Schematic representation of the CERN accelerator complex. The LHC is the largest of the accelerators, shown at the top of the image. The chain of accelerators which feed proton beams in the LHC begin with LINAC 2, visible at the bottom of the image [23].

6 The ATLAS Detector

Depicted in Figure 6, the ATLAS (**A Toroidal LHC ApparatuS**) detector is a large, general-purpose detector built around LHC Interaction Point 1 in Meyrin, Switzerland - visible in the top right of Figure 4 and the center of Figure 5. A forward-backward symmetric cylindrical geometry provides a nearly hermetic 4π solid angle coverage around the collision point. The 44 m long detector is comprised of concentric cylindrical subdetectors extending to a radial distance of 12.5 m [24] from the beam-line. Specifically, it consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic sampling calorimeters, and an outer muon spectrometer situated on three large superconducting air-core toroidal magnets. The innermost detector, known as the Inner Detector (ID), precisely measures the trajectories of charged particles in an encompassing 2 T axial magnetic field [24]. The electromagnetic (ECAL) and hadronic (HCAL) calorimeters serve as dense barriers, stopping particles while measuring their energy in the form of material-induced showers. The outermost detector, known as the Muon Spectrometer (MS), provides a secondary measurement of muon trajectories within a 2 – 6 Tm toroidal magnetic field [24].

The ATLAS Coordinate System

The beam pipe of the LHC runs along the cylindrical axis of symmetry of the detector. A right-handed coordinate system is defined with the origin at the nominal interaction point. The *longitudinal* z -axis points along the beam direction, while the *transverse* xy -plane is perpendicular to the beam direction. The direction of the x -axis is defined as pointing from the interaction point to the center of the LHC ring, while the direction of the y -axis is defined as pointing upwards towards the surface. The azimuthal angle $\phi \in [0, 2\pi)$ is measured along the beam axis in the xy -plane and the polar angle $\theta \in [0, \pi]$ is measured from the beam axis. The transverse momentum is defined as the momentum in the xy -plane, given by

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin \theta. \quad (39)$$

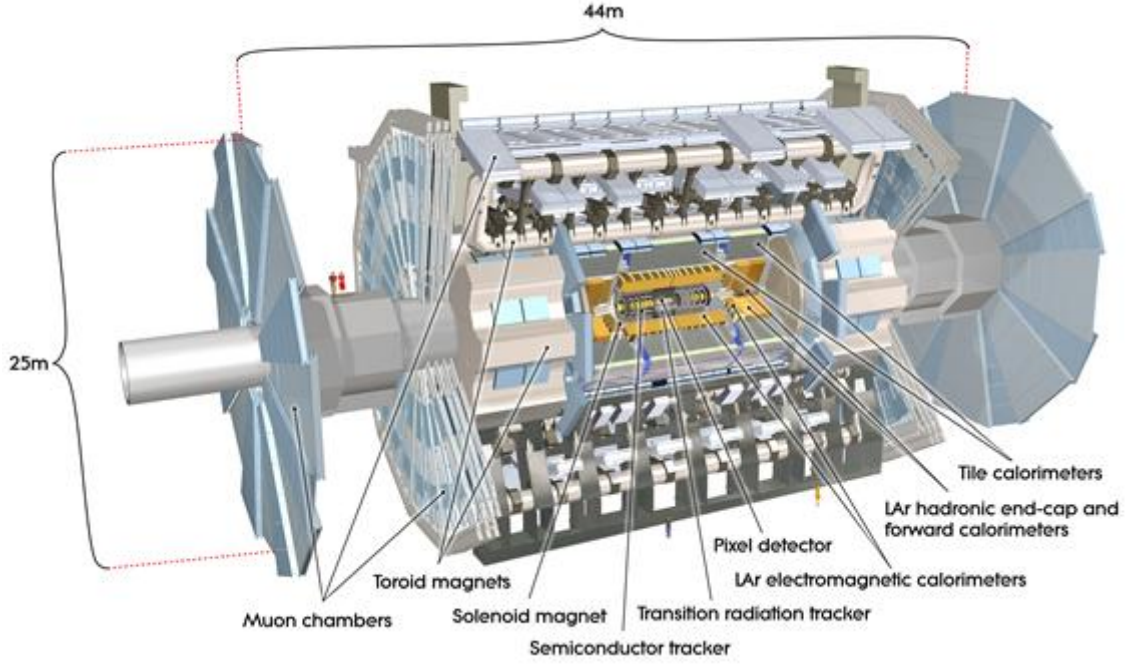


Figure 6: A computer-generated image of the ATLAS detector showing the primary subsystems. Length and width dimensions are shown, with humans added for scale [25].

A commonly used metric in collider physics, known as the rapidity, is defined in terms of a particle's energy and momentum along the beam axis,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \quad (40)$$

The absolute value of rapidity is maximized for particles with significant momentum along the z -axis, and minimized for particles with momentum dominated in the transverse plane. For particles with negligible mass-energy (i.e. $m \ll |\vec{p}|$, hence $E \approx |\vec{p}|$), the rapidity reduces to a purely angular quantity depending only on θ ,

$$\eta = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right) = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right], \quad (41)$$

known as the pseudorapidity¹¹. High absolute values of rapidity and pseudorapidity correspond to forward regions near the beam-axis, while values near zero correspond to the central region in the transverse plane. The ATLAS detector provides coverage of various subsystems within the bounds of $|\eta| < 4.5$. Angular separation is defined as the absolute difference in the ϕ - η plane,

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \quad (42)$$

6.1 Inner Detector

The Inner Detector is the innermost subdetector of ATLAS, responsible for tracking charged particles in an encompassing magnetic field. It consists of an inner silicon-based system composed of pixel and strip sensors, known as the Semiconductor Tracker (SCT), and an outer drift-tube system, known as the Transition Radiation Tracker (TRT). During the first long shutdown period, the Pixel Insertable B-Layer (IBL) was installed in the innermost region, located 3.3 cm from the beam-line [26, 27]. A cross-sectional wedge of the detector is shown schematically in Figure 7. The detector provides tracking capabilities for charged particles above a sub-GeV p_T threshold within the pseudorapidity range $|\eta| < 2.5$ [24]. The 1,744 identical sensors making up the original pixel system are each composed of 47,232 individual pixels with a nominal pitch of $50 \times 400 \mu\text{m}^2$ [24], while the roughly 12 million pixel sensors of the IBL have a smaller size of $50 \times 250 \mu\text{m}^2$ [26]. The 15,912 silicon-microstrip sensors of the SCT have a nominal pitch of $80 \mu\text{m}$ in the square barrel sensors, and a varying pitch between approximately 57 – 94 mm in the trapezoidal end-cap sensors [24]. The 370,000 straw tubes of the TRT are composed of $31 \mu\text{m}$ diameter tungsten wires, with a length of 144 cm in the barrel and 37 cm in the end-caps [24]. Aligned along the beam-axis and surrounding the tracking volume, a thin superconducting solenoid provides a 2 T axial magnetic field throughout the detector volume [24].

¹¹In the realm of collider physics, the combination of (p_T, η, ϕ, m) generally provides a more natural encapsulation of a particle’s four-momentum than the equivalent (E, p_x, p_y, p_z) formulation (see Appendix B).

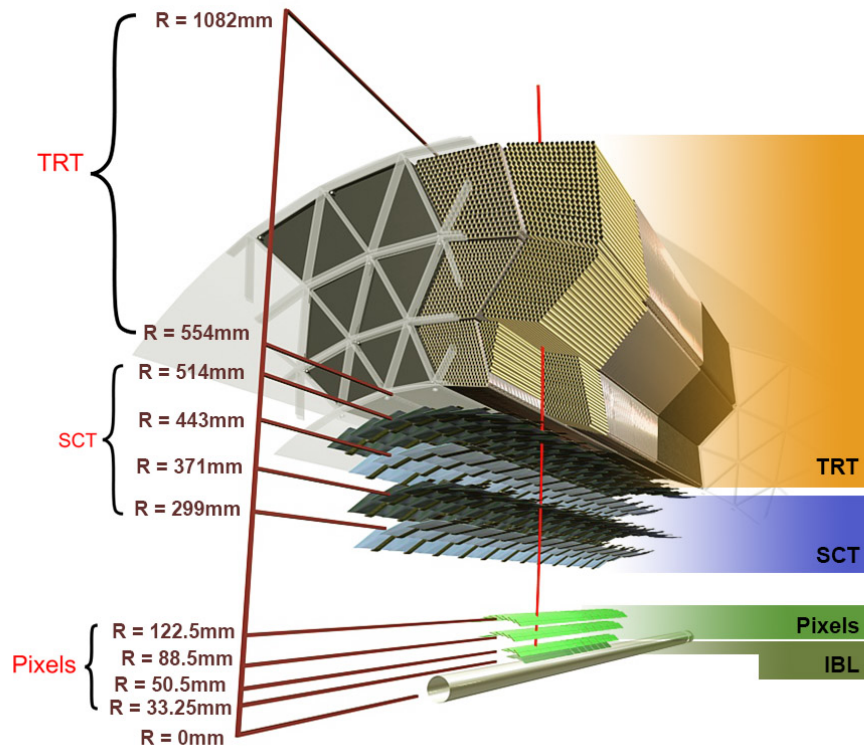


Figure 7: Cross-sectional wedge of the ATLAS Inner Detector, showing the three primary subdetectors: Pixels (incl. IBL), SCT, and TRT. Radial dimensions are measured with respect to the beam-line [28].

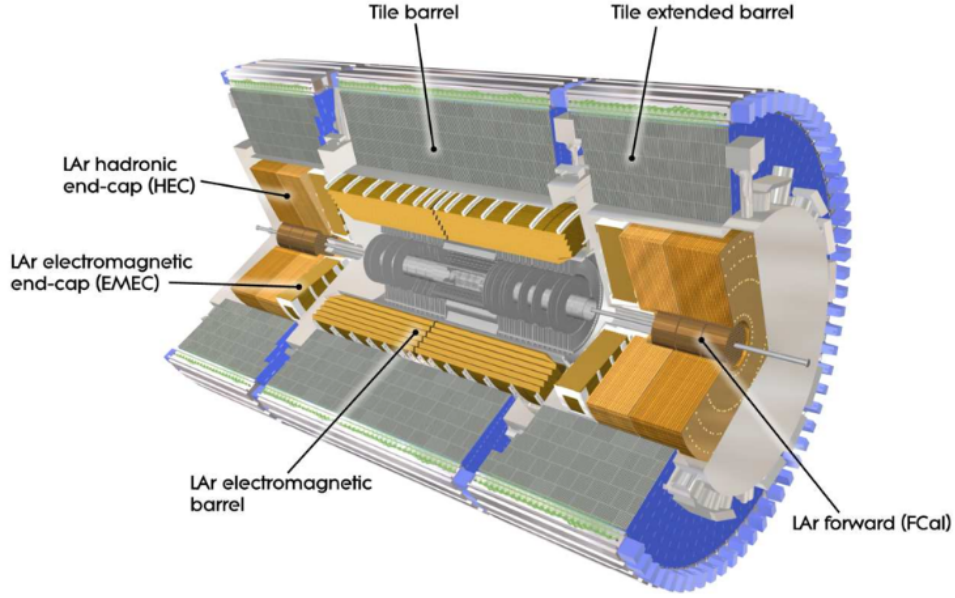


Figure 8: Cut-away image of the ATLAS calorimeter systems, highlighting the electromagnetic and hadronic subsystems [24].

6.2 Electromagnetic and Hadronic Calorimeters

The ATLAS calorimeter system, situated directly outside of the ID region, provides energy measurements of charged and neutral particles. Shown in Figure 8, it is composed of separate electromagnetic and hadronic subsystems covering the pseudorapidity range $|\eta| < 4.9$ [24]. The inner electromagnetic subsystem, consisting of a barrel section ($|\eta| < 1.475$) and two end-cap sections ($1.375 < |\eta| < 3.2$), is built from alternating accordion-shaped layers of lead absorber plates and liquid-argon (LAr) cells [24]. Within the pseudorapidity region corresponding to the ID volume ($|\eta| < 2.5$), finer granularity in the $\Delta\eta \times \Delta\phi$ segmentations provides precision measurements of electrons and photons [24]. The outer hadronic subsystem consists of a central steel/scintillator-tile calorimeter ($|\eta| < 1.7$), two parallel-plate copper/LAr end-caps ($1.5 < |\eta| < 3.2$), and a tungsten/LAr forward calorimeter ($3.2 < |\eta| < 4.9$) [24].

6.3 Muon Spectrometer

The Muon Spectrometer, depicted in Figure 9, forms the outermost layer of the ATLAS detector and extends to a pseudorapidity of $|\eta| = 2.7$ [24]. It provides a complementary measurement of track trajectories for deeply-penetrating charged particles which leave only minimum ionizing deposits in the calorimeters (i.e. muons). The detector consists of the central barrel region up to $|\eta| = 1.05$ and two end-cap sections covering the forward regions, with a 16-fold azimuthal segmentation of chambers arranged in alternating large and small sectors [24]. Four different detector technologies are used throughout the system, each offering a particular specialization in precision tracking or triggering. Three stations of Monitored Drift Tubes (MDT) provide precise momenta measurements for muons within $|\eta| < 2.7$, with Cathode Strip Chambers (CSC) replacing the inner layer of MDTs for $|\eta| > 2$ [24]. Triggering capabilities in the MS are provided by Resistive Plate Chambers (RPC) within the barrel region and Thin Gap Chambers (TGC) in the forward region for $1.05 < |\eta| < 2.4$ [24]. The toroidal magnet system is composed of a barrel toroid, supplying a magnetic field of 0.5 T in the central region, and two end-cap toroids, each producing a 1 T field in the respective end-cap regions [24].

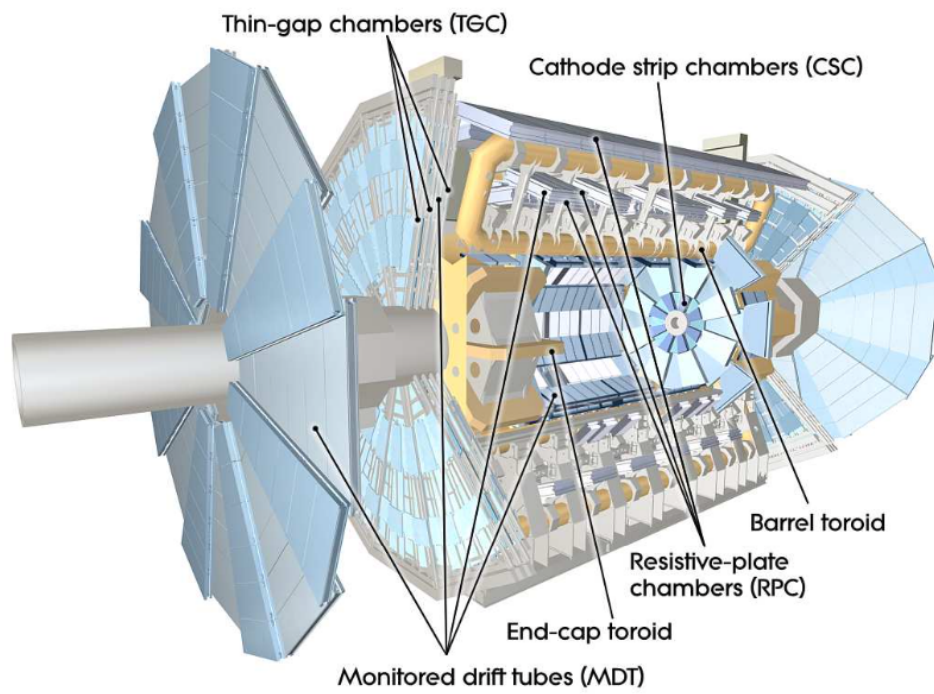


Figure 9: Computer-generated image of the ATLAS Muon Spectrometer, highlighting the four detector subsystems and the toroid magnets [24].

7 Data Processing

7.1 Triggering and Data Acquisition

The high bunch crossing rate of beams in the LHC presents an enormous hurdle of data acquisition and storage. The combination of the 40 MHz frequency, corresponding to a 25 ns latency between successive crossings, and the immense detail captured by ATLAS, make it impossible to store complete event records of all collisions. Additionally, aside from any physical limitations, the majority of collisions involve only minimal scattering of protons and are therefore less important to the physics goals of the experiment. For these reasons, only a fraction of the total events are selected and saved to permanent storage for later analysis. This is the role of the ATLAS trigger and data acquisition system.

As depicted in Figure 10, the ATLAS trigger system consists of two levels: a Level 1 (L1) hardware trigger and a secondary, high-level software trigger (HLT) [29]. The L1 trigger runs on custom-built hardware using information from the calorimeters and Muon Spectrometer in order to quickly determine whether or not to accept an event. The incoming data rate of 40 MHz, driven by the bunch crossing frequency, is reduced by a factor of 400 to 100 kHz before reaching the secondary trigger system [29]. The HLT runs on high-performance CPU farms using data from all detectors in order to run simplified reconstruction algorithms emulating the more complex offline versions. Based on these algorithms, which provide crude information on particle candidates, a decision is made whether or not to save the event. This software-based trigger reduces the L1 rate of 100 kHz to an average output rate of 1.5 kHz, corresponding to a data readout of 1.5 GB/s [29]. In total, the data rate is reduced by a factor of about 40,000, making it possible to write the full record of all remaining events to permanent storage.

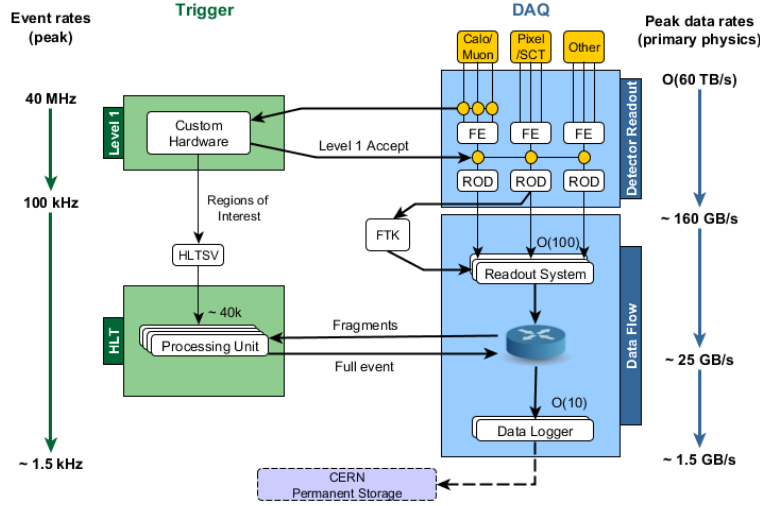


Figure 10: A schematic overview of the ATLAS trigger and data acquisition system, responsible for reducing the incoming data rate by over ten-thousandfold in order to store full event records of potentially interesting collisions. The trigger system is comprised of a Level-1 (L1) hardware trigger and a secondary, high-level software trigger (HLT) [29].

7.2 From Signals to Sigmas

The ATLAS data processing chain describes the multistage transformation of raw detector signals collected by ATLAS to analysis-ready physics objects. The triggering and data acquisition system, discussed previously, represents the first stage in this process - in which events passing the two-level trigger system are saved to permanent storage in a high-detail file format (RAW). This initial stage is necessarily run *online* during data-taking, while subsequent stages are run *offline*, often relying on the distributed LHC Computing Grid infrastructure. The following stage, known as reconstruction, involves the transformation of raw detector signals into physics objects, written in the so-called Analysis Object Data (AOD) format. Given the large $\mathcal{O}(PB)$ size of AODs, analysis groups often perform an additional derivation stage, involving substantial data reduction and minimal augmentation based on the physics needs. The outputs of this stage are known as Derived AODs (DAODs), and are typically on the order of several terabytes in size. The final stage, often particular to

the analysis, typically consists in the transformation of DAODs into flat n-tuples, containing only essential information pertinent to the analysis and typically small enough in size to process locally.

A parallel chain is also run for Monte Carlo (MC) samples to simulate the expected ATLAS response to physics processes of interest. This chain starts with the event generation itself, involving the simulation of the underlying hard-scatter process and the subsequent parton showering and hadronization into stable particles. The following stage simulates the interactions and showering of these particles within the ATLAS material, typically by means of the GEANT4 [30] simulation toolkit, particularly in terms of their charge and energy deposition in detector-sensitive elements. Next, the digitization stage evaluates the expected detector response to these signals, emulating the output of real data generated from the detector. From this point, the reconstruction and derivation stages can be run identically as for the genuine data. A schematic overview of these parallel data and MC processing chains are shown in Figure 11.

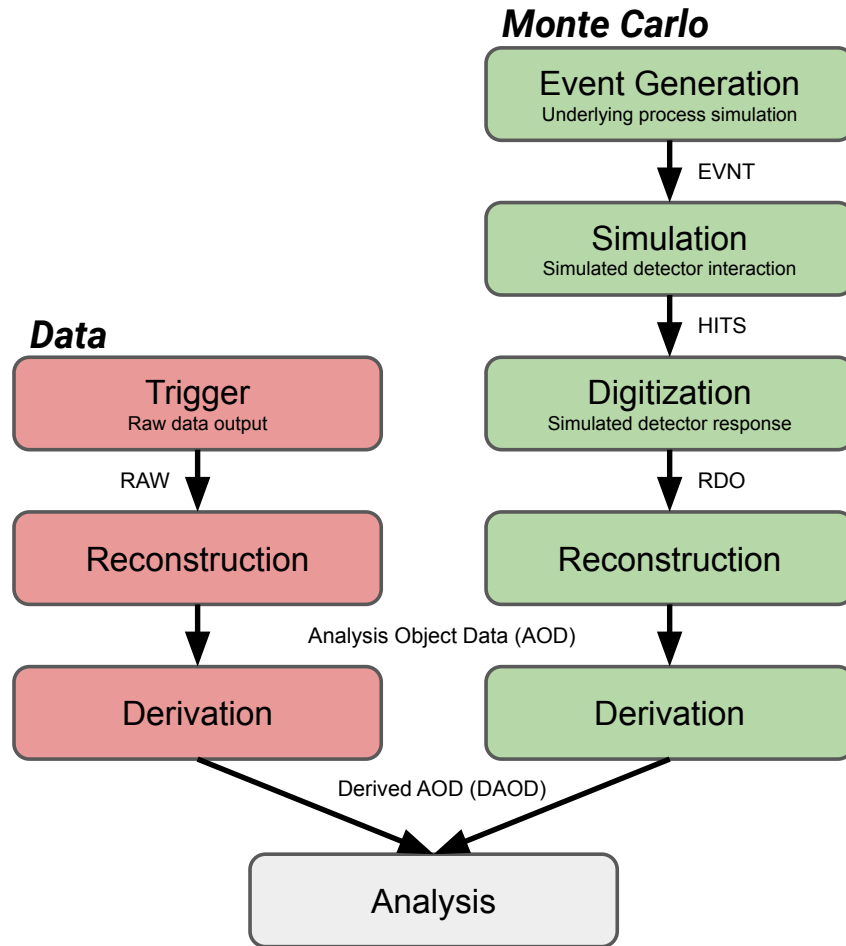


Figure 11: Overviews of the ATLAS data processing chains for data and Monte Carlo. File formats from each stage are shown adjacent to the arrows. Adapted from James Catmore’s talk *The ATLAS data processing chain: from collisions to papers*.

8 Particle Reconstruction

The interaction of particles with the ATLAS detector is the key to their detection and identification. Particles traversing the detector leave a series of idiosyncratic signals through the many detector layers, specific to the particular type and kinematics of the object. Reconstruction in ATLAS describes the multistage process of transforming raw detector signals (e.g. hits, clusters) from all relevant subdetectors into well-defined physics objects (e.g. muons, photons, jets), representing experimental measurements of a candidate particle's identification and kinematics.

8.1 Techniques

Generally speaking, reconstruction in ATLAS depends on two complementary techniques: tracking and calorimetry. Tracking corresponds to the measurement of charged particle trajectories within a magnetic field, while calorimetry involves the energy measurement of neutral and charged particles by the complete absorption of material-induced showers. The fundamental basis of tracking therefore consists in measuring curved tracks, characterized by five helix parameters, by fitting a series of compatible precision hits. The foundation of calorimetry consists in measuring the position and energy of clusters spread across neighboring calorimeter cells. The combination of these detection techniques is crucial to the overall performance goals and each serves to complement one another in terms of resolution - while the relative momentum resolution from tracking worsens with increasing momentum, the relative energy resolution from calorimetry improves with increasing energy.

Tracking

Charged particle tracking is an essential component of collider detectors, particularly for high-luminosity experiments having a significant overlap in simultaneous inelastic collisions (*pile-up*). Tracking provides both a precise measurement of position and momentum, as well

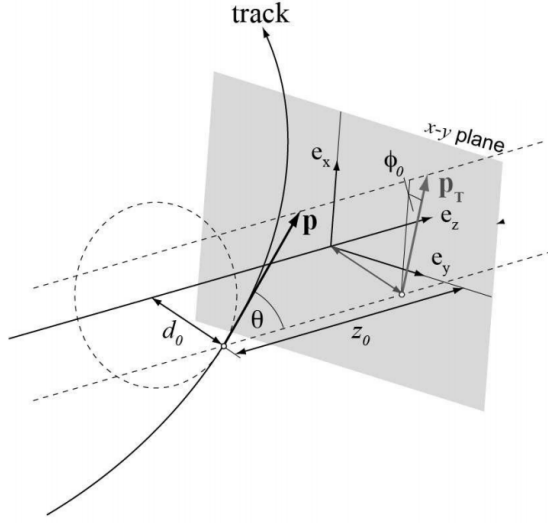


Figure 12: The trajectory of a charged particle in a solenoidal magnetic field can be fully described by five helix parameters. In global detector coordinates, the choice of $(d_0, z_0, \phi_0, \theta, q/p_T)$ as depicted here is frequently used.

as a means to extrapolate particle trajectories inwards to the interaction point (IP). More specifically, it consists in the measurement of five helix parameters completely describing a charged particle's trajectory in a solenoidal magnetic field. In global detector coordinates, the choice of $(d_0, z_0, \phi_0, \theta, q/p_T)$ as depicted in Figure 12 is typically used. The transverse impact parameter d_0 corresponds to the distance of closest approach to the IP in the transverse plane. The location of this perigee point also defines the longitudinal impact parameter z_0 , corresponding to the distance from the IP projected along the z -axis, and the azimuthal angle ϕ_0 . The last parameter q/p_T represents what is actually measured via the Lorentz force - that is, the ratio of a particle's charge to its component of momentum perpendicular to the magnetic field.

Tracking detectors are generally placed as closed to the beam-line as possible. The reasoning is twofold: firstly, this decreases the material traversed by particles in the tracking region, thereby minimizing the deleterious effects of multiple scattering and energy losses, and, secondly, this provides a series of high precision position measurements close to the beam-line, crucial for the extrapolation of tracks to the interaction point. To further mini-

mize energy losses, tracking detectors are typically made using low radiation length materials such as silicon. Given the proximity to the beam-line, tracking is critical for the process of *vertexing* in which tracks are extrapolated inwards and matched to unique interaction points, termed primary vertices. This process allows for the definition of vertex-specific information, as well as the separation from potentially interesting secondary decay-chain interactions and unwanted pile-up vertices. The relative momentum resolution of a tracking detector can roughly be decomposed as

$$\frac{\sigma(p_T)}{p_T} = a \cdot p_T \oplus b, \quad (43)$$

consisting of one component scaling with p_T , owing to the reduced curvature of high momentum tracks, and one fixed component, representing the low- p_T limit driven by multiple scattering. The coefficient of each term is inversely related to the magnetic field and path length of the detector, thereby motivating a strong solenoid field covering a far radial extent of tracking-sensitive layers.

A summary of the primary and secondary back-tracking reconstruction chains in ATLAS are shown in Figure 13. The primary reconstruction chain in the ATLAS ID follows an inside-out approach building track seeds in the innermost region and progressively extending track candidates into the outer layers. Charged particles traversing the inner silicon layers leave charge deposits by means of ionization in the pixel and microstrip sensors, which are interpreted as hits if the collected charge is above the sensor noise. In the outer TRT, electrons produced by ionization in the straw tubes are collected and measured by electrodes, and similarly interpreted as hits, commonly known as drift circles. A pre-processing stage combines these signals in neighboring Pixel and SCT channels into clusters. Clusters are converted into three-dimensional space-points with a corresponding position uncertainty based on the sensor geometry and pitch. Track seeds are formed from a collection of three space-points, passing curvature and purity requirements, in the Pixel or SCT subdetectors. Silicon track candidates are built by successively extrapolating seeds into the outer layers using a combinatorial Kalman filter [31], in which compatible clusters are iteratively included in the

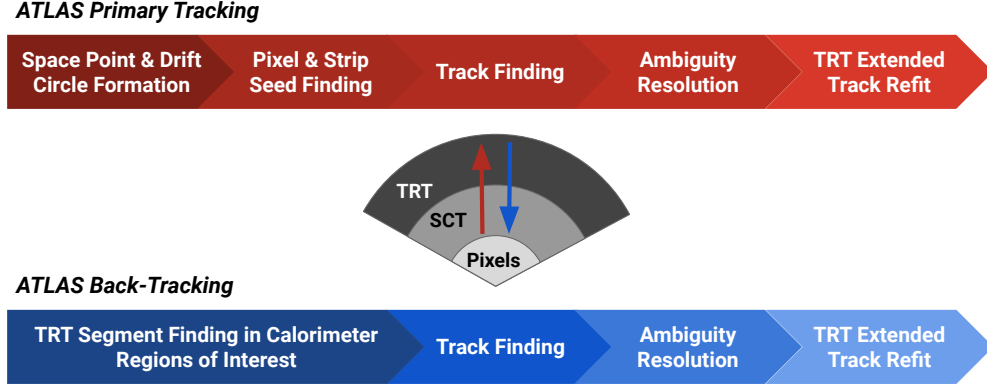


Figure 13: Simplified overview of the primary and secondary back-tracking chains in ATLAS reconstruction. The primary reconstruction runs from inside-out, starting from silicon space-points. The secondary back-tracking chain runs from outside-in, seeded from leftover TRT drift circles within electromagnetic calorimeter regions of interest.

estimated track trajectory. A dedicated ambiguity resolution stage resolves overlaps in track candidates, reassigning shared clusters to tracks based on a competitive scoring procedure. The purified track candidates are then fit using a global χ^2 method. Lastly, if there are compatible drift circles present in the TRT, tracks are extended into the outermost layer and subsequently refitted.

A secondary back-tracking chain is run outside-in to recover photon conversions within the detector material. In this case, reconstruction is only run in regions of interest defined by clusters in the electromagnetic calorimeter. Drift circles in the TRT compatible with calorimeter regions of interest are used to seed an inside-out reconstruction similar to the primary tracking chain. Only leftover drift circles not used in track fits during the initial tracking pass are considered.

Following the track extrapolation and fitting, the vertexing procedure uses the resulting track candidates to identify unique collision points along the beam-line. In standard ATLAS reconstruction, the hard-scatter primary vertex is defined as the vertex with the largest $\sum p_T^2$ of the associated tracks, while all remaining vertices are considered pile-up.

Calorimetry

Calorimetry comprises the second essential component of collider detectors, necessary for the measurement of neutral particles and high-density particle showers. The role of a calorimeter is to both cause the development and measure the properties of electromagnetic or hadronic showers within a large detector volume. Showers are collimated avalanches of secondary particles initiated by the interaction of the incident particle with the high-density material of the detector. By measuring the many individual energy deposits in a particle shower, either by means of scintillation light or ionization energy, the original particle's position and energy can be deduced. Given the larger segmentation of calorimeters, the spatial resolution is generally worse than tracking detectors.

Calorimeter systems are typically segmented into cells, with alternating layers of high-density passive material and active material. For example, the cells of the ATLAS electromagnetic calorimeter, shown in Figure 14, form repeating accordion-shaped layers of lead absorber plates and liquid-argon in the $\eta - \phi$ plane. Passive regions provide the high radiation lengths necessary to stop energetic particle, in the process inducing showers of particles which are measured within the active cells. The relative energy resolution of a calorimeter can roughly be decomposed into three terms as

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c. \quad (44)$$

Concerning the first term, the energy resolution of an ideal calorimeter is fundamentally stochastic, depending on the statistical fluctuations in the number of measured shower particles. On average, the number of shower particles is directly proportional to the energy of the primary particle and, therefore, the relative energy resolution improves with energy by $\sqrt{E} \propto \sqrt{N}$. The additional terms in Equation 44 account for, among other sources, electronic noise and shower leakage and are driven primarily by the detector design and geometry, respectively.

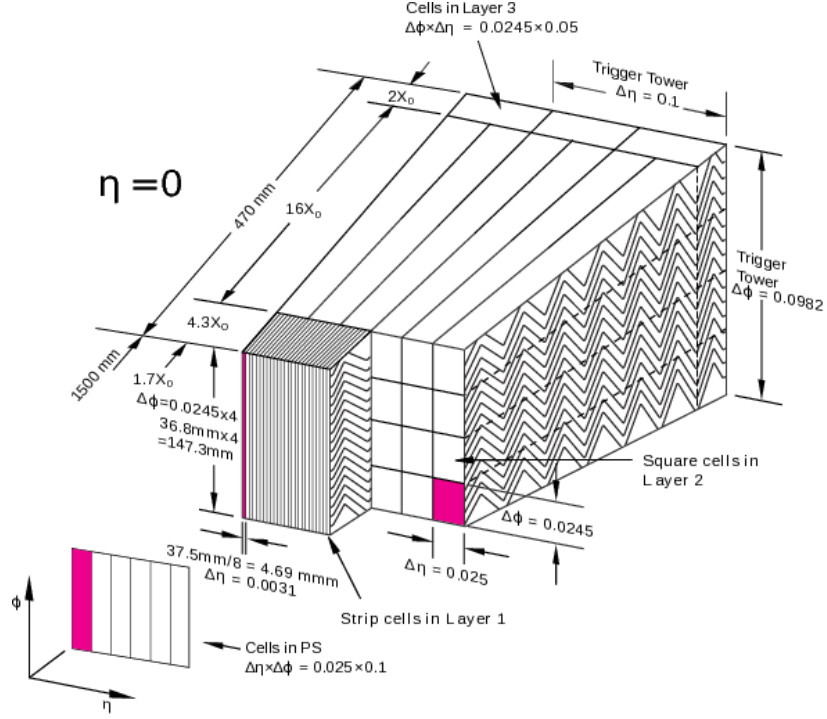


Figure 14: Schematic view of an electromagnetic calorimeter cell in ATLAS [32].

Calorimeter reconstruction in ATLAS starts by forming topological clusters (topo-clusters) from calorimeter seed cells with an energy significance well above noise levels [33]. Neighboring cells with high energy significance are iteratively added to expand the seed clusters [33]. A splitting algorithm then separates clusters based on the local energy maximum, and the resulting purified clusters are calibrated based on the local detector properties and presence of nearby clusters [33]. The calibration stage allows for a classification of the cluster as electromagnetic- or hadronic-initiated. The final energy measurement is obtained by summing the calibrated signals of connected cells within a cluster, with a corresponding position measurement calculated from the individual cell locations [33].

8.2 Objects

The basis of object reconstruction in ATLAS is shown in a simplified means in Figure 15. The figure depicts a cross-sectional wedge of the detector in the transverse plane, superimposed with the characteristic signatures of various particles traversing the several layers of

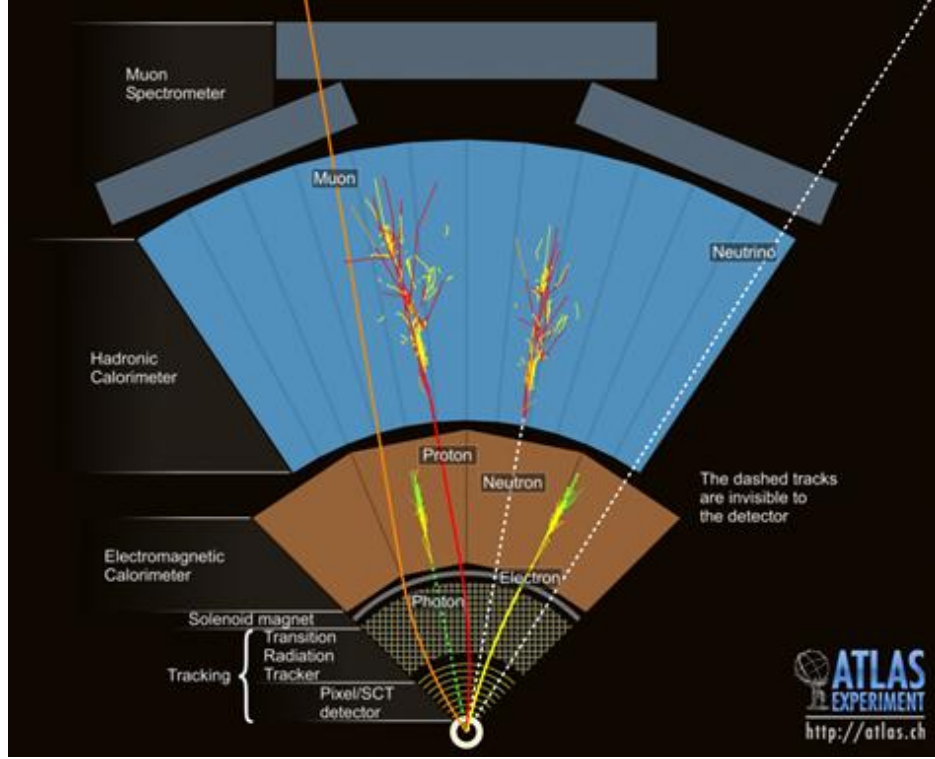


Figure 15: Cross-sectional wedge of ATLAS in the transverse plane. Simplified representations of various particles traversing the detector are shown to highlight the general differences in particle reconstruction [34].

subdetectors. Charged particles (e.g. protons, electrons, muons) leave curved tracks within the Inner Detector, allowing for precise measurements of position and momentum. Photons and (anti-)electrons, by interacting with detector material, induce electromagnetic showers within the first calorimeter system, thereby allowing for a measurement of energy. Similarly, hadronic jets, produced through the showering and hadronization of final state partons, are measured and absorbed within the outer calorimeter as showers of hadronic activity. Muons generally leave a secondary track in the Muon Spectrometer, providing an independent measurement of position and momentum complementing that of the Inner Detector. Weakly-interacting, neutral particles, namely neutrinos, leave no signatures in the detector, except in the indirect form of momentum imbalance in the transverse plane.

Electrons

The challenge of electron reconstruction is a direct result of their low mass, which largely determines their interactions with matter and, therefore, the means of their detection. Electrons in the multi-GeV range and above interact with matter primarily by means of bremsstrahlung radiation [5] - the process by which a charged particle deflected by the electric field of an atomic nucleus emits radiation in the form of a photon - rather than ionization. High-energy photons emitted by bremsstrahlung can in turn pair produce an e^+e^- pair in the field of a nucleus. The original electron and positron-electron pair will each in turn interact with detector material, repeating the chain and producing a cascade of particles known as an electromagnetic shower. These electromagnetic showers, though stochastic in nature, are on average similar for electrons of the same energy, having a characteristic penetration depth in a given material. Calorimeters designed specifically for the detection of electromagnetic activity are therefore capable of providing a precise measurement of energy for electrons, which is complemented by the momentum and position measurements provided by tracking detectors.

Electrons in ATLAS are reconstructed from a combination of track measurements in the ID and energy clusters measured in the electromagnetic calorimeter. The trajectory of a typical electron passing through the many layers of the ATLAS detector is shown schematically in Figure 16. The several tracking layers of the ID provide a series of position measurements along the trajectory, before the electron initiates a shower which is absorbed and measured within the electromagnetic calorimeter. The TRT provides additional information for electron identification by detecting transition radiation X-rays, which are preferentially emitted by electrons. Identification of electrons is performed with a likelihood discriminator, using the track and cluster properties as input.

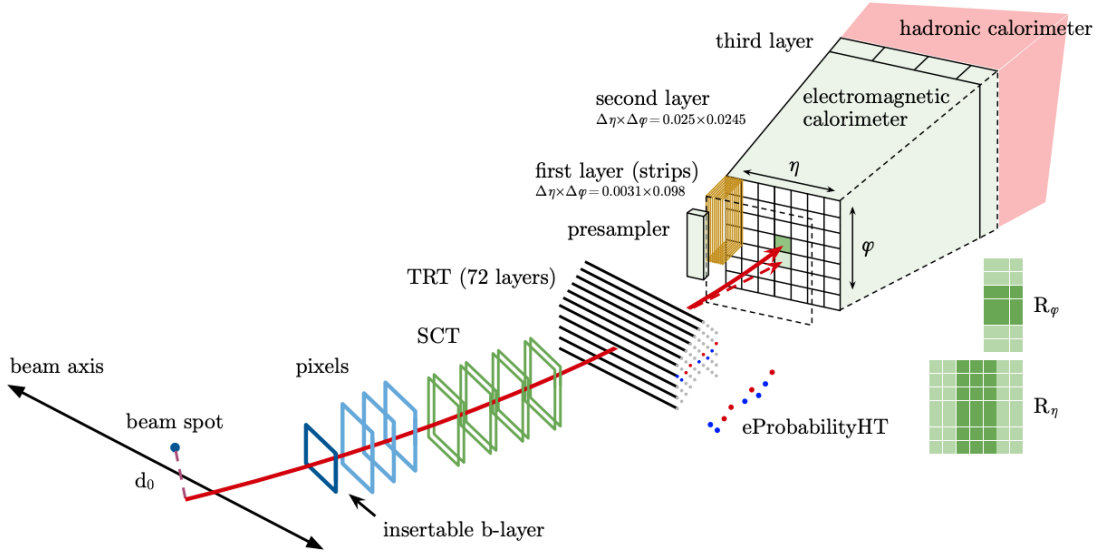


Figure 16: Schematic view of the principles underlying electron reconstruction and identification in ATLAS [35].

Muons

Muons, being a much heavier counterpart of electrons, interact with matter primarily through ionization, rather than bremsstrahlung radiation. This is because the rate of bremsstrahlung radiation is inversely proportional to the square of the particle mass, and is thus $(m_\mu/m_e)^2 \approx 40,000$ times more suppressed for muons than electrons [5]. As a result, muons typically do not initiate electromagnetic showers in the calorimeters of ATLAS, instead leaving only minimum ionizing deposits of energy and continuing through to the MS. Therefore, charged particle tracking is the primary means of muon detection.

Muon reconstruction in ATLAS benefits significantly from the independent track measurements of the ID and MS and, to a lesser extent, from the measurement of energy deposits in the calorimeters. Although the combination of track measurements in the ID and MS generally provides the best measurement of muons in ATLAS, it is not a strict requirement of the reconstruction. As depicted in Figure 17, the variety of detection means provided by ATLAS allows for several complementary muon reconstruction algorithms:

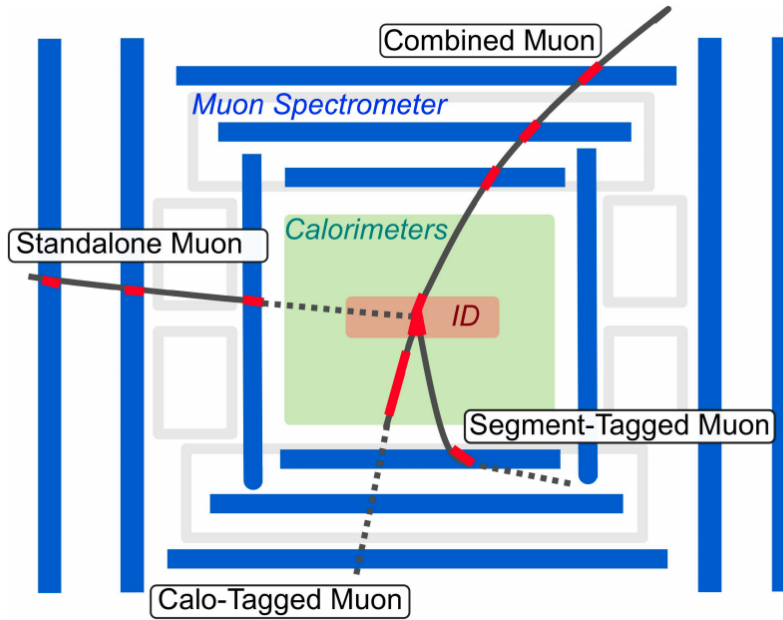


Figure 17: Schematic view of four different muon reconstruction algorithms used in ATLAS, and the various subdetectors involved for each technique [36].

- **Combined** muons are obtained from a global refit of ID and MS tracks
- **Segment-Tagged** muons consist of a fitted ID track and MS segment
- **Calo-Tagged** muons consist of an ID track matched to a minimum ionizing energy deposit in the calorimeters
- **Standalone** muons consist of a lone MS track

Jets

Jets are collimated showers of hadrons and other particles initiated by the hadronization of final state quarks or gluons. Characterized by a large penetration depth, jets typically pass through the electromagnetic calorimeter and require a significant distance to be fully absorbed in the hadronic calorimeter. The measurement of jets is in principle similar to that of electromagnetic showers, based largely on the formation of topo-clusters in the calorimeters, but in practice is much more complicated due to the variety of constituent particles involved.

Many well-known reconstruction algorithms exist to determine the clustering and splitting of energy deposits into jets. One of the broad classes of jet reconstruction algorithms is that of sequential recombination, in which *nearby* particles are iteratively combined to form pseudojets, based on some distance metric defining the separation of pseudojets from other particles. The process is iterated over the set of particles until the distance metric defining the pseudojet itself falls below that of the particle and pseudojet, at which point the object is deemed a jet and removed from further iterations [37]. The anti- k_T algorithm [37] uses a separation metric dependent upon the transverse momentum of particles and pseudojets, the angular separation between the two, and a fixed radius parameter. The radius parameter, common to most jet algorithms, roughly determines the angular size of the resulting jets in the $y - \phi$ plane.

Jets relevant to this analysis are reconstructed using a particle flow implementation of the anti- k_T algorithm with radius parameter 0.4 (ANTI k_T 4EMPFLOW) [37, 38]. Particle flow reconstruction techniques use measurements from both calorimeter and tracking systems, in the form of clusters and tracks, respectively. The inclusion of track measurements is advantageous over purely cluster-based algorithms for a variety of reasons. Most importantly, the momentum resolution of tracking detectors is superior to the energy resolution of calorimeters for low-energy charged particles, and the reconstruction is extended to softer particles whose energy deposits do not exceed the noise thresholds for topo-cluster seeding [38]. Additionally, the finer granularity of tracking detectors provides a significantly better angular resolution for single charged particles [38]. Moreover, unwanted contributions from pile-up collisions can be removed from clusters by using track-to-vertex association. A schematic view of the particle flow jet reconstruction algorithm used in ATLAS is shown in Figure 18.

The process starts with the standard measurements of calorimeter topo-clusters and ID tracks. In cases where a topo-cluster and ID track are matched to one another, the track measurement is used in place of the cluster, and its energy contribution is subtracted from the cluster (modified cluster) [38]. Clusters without a matching track, potentially originating

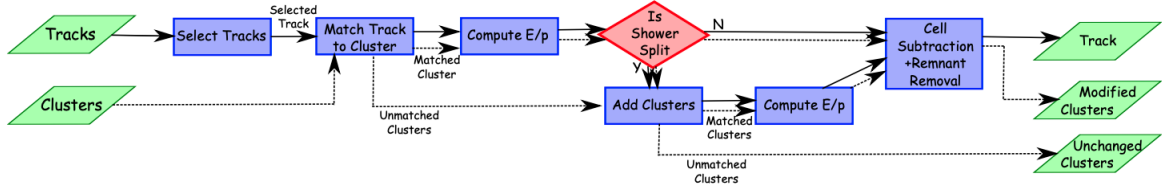


Figure 18: Flowchart detailing the complementary usage of tracks and clusters in the ATLAS particle flow jet reconstruction algorithm [38].

from a neutral particle or outside of the tracking acceptance, are used directly without any cell subtraction (unchanged cluster) [38]. In this way, the performance of particle flow jets is superior to jets formed from topo-clusters alone, particularly in terms of resolution and pile-up robustness [38].

9 Outlook for Future Upgrades

9.1 High Luminosity LHC Program

Over the course of a decade, the LHC has provided an enormous quantity of data to the on-board experiments, facilitating the joint discovery of the Higgs boson by ATLAS and CMS in 2012 [39, 40] and the measurement of many rare Standard Model processes to unprecedented precision. However, many extremely rare process, such as di-Higgs production, have so far eluded detection and will likely remain undiscovered even given the additional Run-3 dataset. Fortunately, the LHC has a long future planned beyond the next run of data-taking, with ultimate high-luminosity runs extending well into the 2030s. The past and future timeline of the LHC, as anticipated at the time of writing, is shown in Figure 19.

The upcoming High Luminosity Large Hadron Collider (HL-LHC), expected to start operation in the second half of 2027, aims to further increase the instantaneous luminosity and center-of-mass energy of proton-proton collisions to $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and $\sqrt{s} = 14 \text{ TeV}$, respectively. These changes will yield an expected $\langle\mu\rangle = 200$ inelastic proton-proton collisions per bunch crossing, several times larger than ever observed in any previous LHC run¹². The high-luminosity conditions will deliver unprecedented quantities of data over the decade-long program, thereby allowing for the study of currently inaccessible rare process and improving the precision of world-leading Standard Model measurements. Prior to the start of the program, the ATLAS detector will undergo major subsystem overhauls during the Phase-II upgrade period. These upgrades are essential to address the high detector occupancies and harsh radiation conditions expected during the high-luminosity program.

9.2 ATLAS Detector Upgrades

Since its installation in 2008, the ATLAS detector has already undergone several upgrades in order to cope with the increasing demands of the challenging LHC environment. Perhaps

¹² The average observed pile-up during Run-2 of the LHC was $\langle\mu\rangle = 33.7$ (see Figure 29), with some high pile-up runs reaching $\langle\mu\rangle = 70$

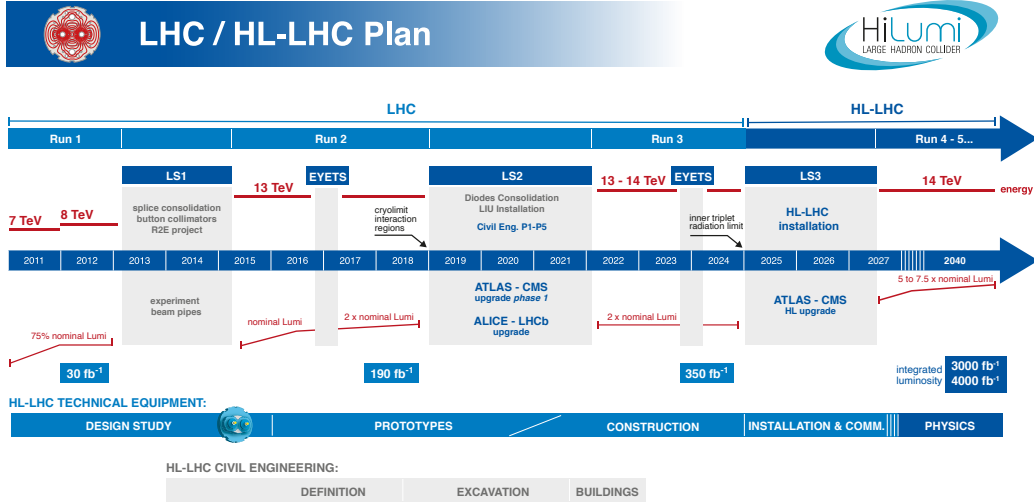


Figure 19: Past and future timeline of the LHC operation. The three long shutdown periods (LS1, LS2, and LS3) correspond to the completed and planned upgrade periods. The ultimate high-luminosity runs of the LHC, expected to begin in the second half of 2027, aim to deliver a total integrated luminosity beyond 3000 fb⁻¹, exceeding the expected 300 fb⁻¹ combination of the first three LHC runs by a factor of ten (Credit: CERN).

most notably, the IBL [26, 27] was installed in 2014 during the first long shutdown period (LS1) to improve tracking performance, particularly for the reconstruction of b -jets and displaced vertices. At the time of writing, the so-called Phase-I upgrades, are currently being developed and installed in the ATLAS cavern during the second long shutdown period (LS2) in preparation for Run-3 of the LHC. The next major series of upgrades to ATLAS, as well as the collider, will begin during the third long shutdown period (LS3) in preparation for the final high-luminosity runs of the LHC.

While many essential upgrades to the subdetector and data acquisition systems of ATLAS are planned during LS3, perhaps most notably are the replacements and upgrades of the tracking system. The all-silicon Inner Tracker (ITk) will replace the entirety of the current ATLAS Inner Detector in order to maintain and improve upon tracking performance, particularly for the reconstruction of very forward jets. Additionally, the High Granularity Timing Detector (HGTD) will be installed outside of the ITk end-caps to supplement the

Table 3: Overview of the Run-2 Inner Detector (ID) and the upcoming replacement Inner Tracker (ITk), planned for the advent of the HL-LHC program. The ITk layout referenced here and throughout this section refers to version 23-00-03, as documented in Ref. [41].

Category	Run-2 ID	HL-LHC ITk
Sensor Technologies	Silicon pixel/strip, gaseous straw tube	Silicon pixel/strip
Pseudorapidity Coverage	$ \eta < 2.5$	$ \eta < 4.0$
Innermost Layer Radius	$r = 33$ mm	$r = 34$ mm
Number of Pixel Layers	3	5
Number of Strip Layers	4	4
Nominal Pixel Pitch	$50 \times 400 \mu\text{m}^2$	$50 \times 50 \mu\text{m}^2$
Innermost Pixel Pitch	$50 \times 250 \mu\text{m}^2$	$25 \times 100 \mu\text{m}^2$
Nominal Strip Pitch	$80 \mu\text{m}$	$75.5 \mu\text{m}$

tracking system with picosecond timing information, crucial for the discrimination of out-of-time pile-up vertices. The combination of timing measurements from the HGTD and spatial measurements from the ITk will facilitate four-dimensional tracking in the reconstruction chain.

In order to increase granularity and improve radiation hardness while minimizing the material footprint, the ITk is based entirely on silicon detector technologies. Shown schematically in Figure 20, it consists of an inner silicon pixel subsystem extending to $|\eta| = 4.0$ and an outer silicon strip subsystem covering $|\eta| < 2.7$. This drastically extends the $|\eta| < 2.5$ coverage of the ATLAS ID, thereby greatly increasing the acceptance for forward track reconstruction. Due to the low radiation length of silicon, the material budget of the ITk is significantly lower than that of the ID in the forward region. A nominal pixel pitch of $50 \times 50 \mu\text{m}^2$ will be used throughout much of the pixel detector, with an alternative pitch of $25 \times 100 \mu\text{m}^2$ used in the innermost barrel layer. A comparative overview of the Run-2 ID and ITk is shown in Table 3 to highlight the key differences.

Maintaining track reconstruction performance in the high pile-up environment of the HL-LHC is central to the design of the ITk. Owing to reconstruction and detector improvements, the ITk reconstruction as optimized for $\langle\mu\rangle = 200$ is already faster than the Run-2 ID reconstruction at $\langle\mu\rangle = 60$, and expected to improve more given further developments.

Importantly, these drastic improvements in processing speed do not significantly impact the physics performance. The track reconstruction efficiency¹³ of the ITk at $\langle\mu\rangle = 200$ is shown alongside a comparison to the Run-2 detector at $\langle\mu\rangle = 38$ in Figure 21. With respect to the the Run-2 detector, the ITk provides a comparable reconstruction efficiency within 5% while operating under a factor five increase in pile-up. Crucially, the presence of *fake* tracks - such as those corresponding to incorrect combinations of hits - is also kept low, with an estimated rate of 10^{-4} at $\langle\mu\rangle = 200$ [41]. Figure 22 shows the number of reconstructed tracks versus the number of interactions for the ITk and Run-2 detectors under their respective nominal conditions. The number of real tracks is expected to scale linearly with pile-up, while fake tracks are expected to scale with a higher power, owing to the dramatic increase in combinatorial possibilities. The near-linear behavior of the ITk detector in Figure 22 therefore illustrates a robust tracking efficiency and a negligible fake rate over a wide range of pile-up conditions [41]. In terms of the resolution of track impact parameters, crucial for vertexing and pile-up mitigation, the current ITk layout provides factor two improvements in the transverse d_0 and factor four improvements in the longitudinal z_0 impact parameters for $p_T = 100$ GeV muons, as shown in Figure 23. The significant improvements in the z_0 resolution are driven by the reduction in pixel pitch in the innermost layer, as outlined in Table 3. Though not shown here, the transverse momentum resolution of the ITk is also expected to outperform the Run-2 detector due to improvements in the spatial resolution along the bending direction [41].

¹³ The track reconstruction efficiency is measured using Monte Carlo truth information. It corresponds to the fraction of truth particles with a matching reconstructed track, and is profiled as a function of truth-level observables - frequently η and p_T owing to the geometric and kinematic dependencies.

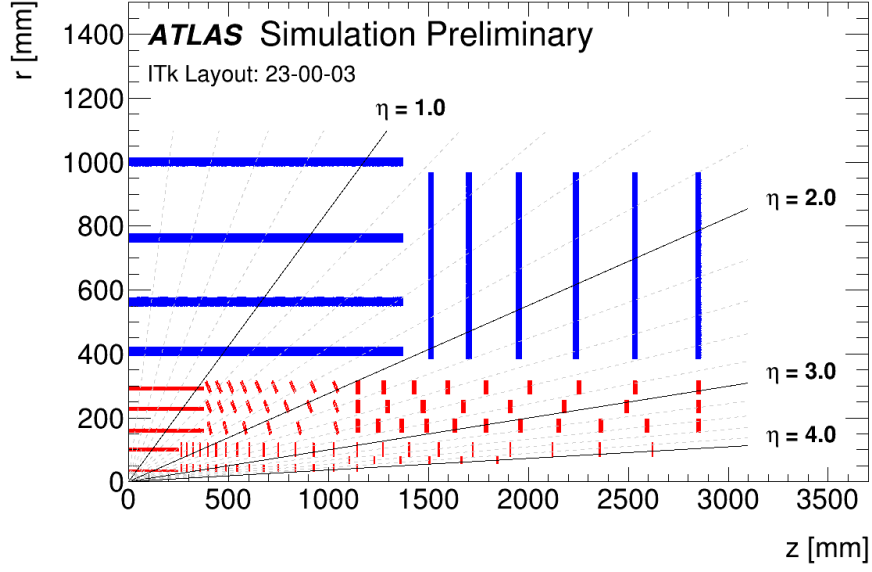


Figure 20: A schematic depiction of the ITk layout showing the active pixel (red) and strip (blue) detector elements in one rz -quadrant [41].

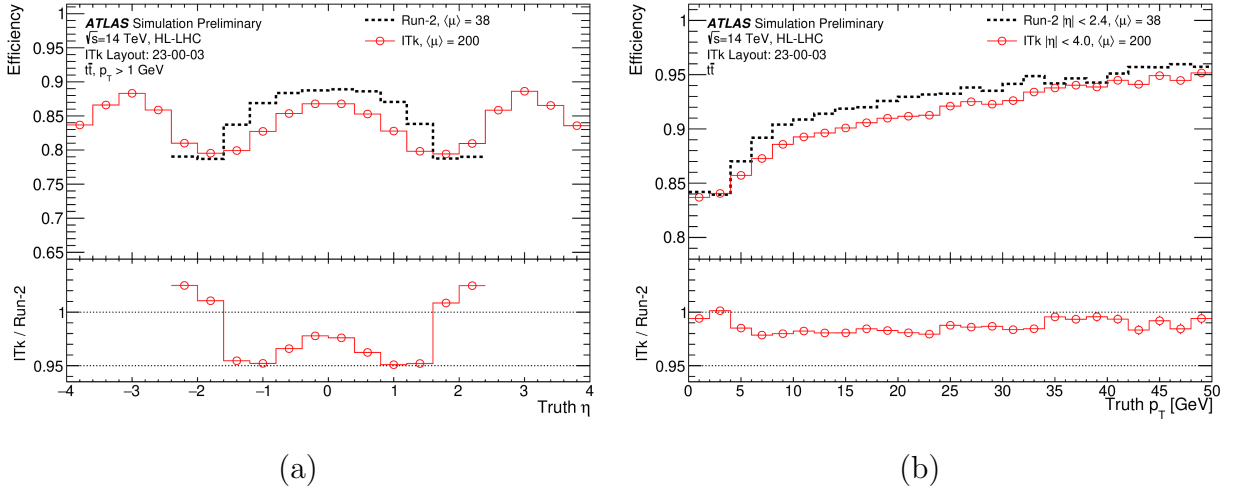


Figure 21: Track reconstruction efficiency versus η (a) and p_T (b) for all tracks from $t\bar{t}$ decays, shown separately for the ITk at $\langle\mu\rangle = 200$ and the Run-2 detector at $\langle\mu\rangle = 38$ [41].

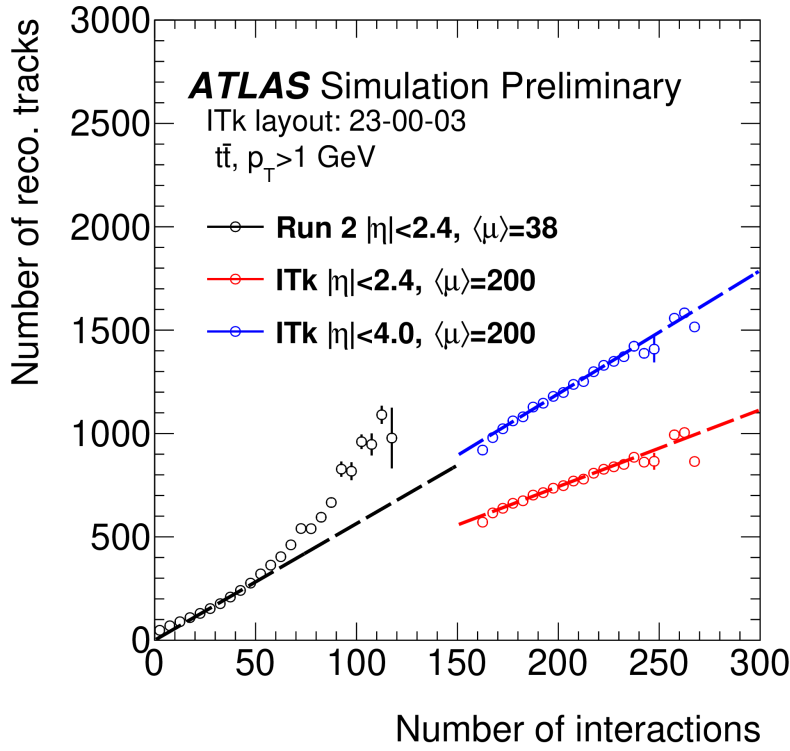


Figure 22: Number of reconstructed tracks with $p_T > 1$ GeV per event versus the number of interactions in $t\bar{t}$ events, shown separately for the ITk at $\langle \mu \rangle = 200$ in two $|\eta|$ windows and the Run-2 detector at $\langle \mu \rangle = 38$ [41].

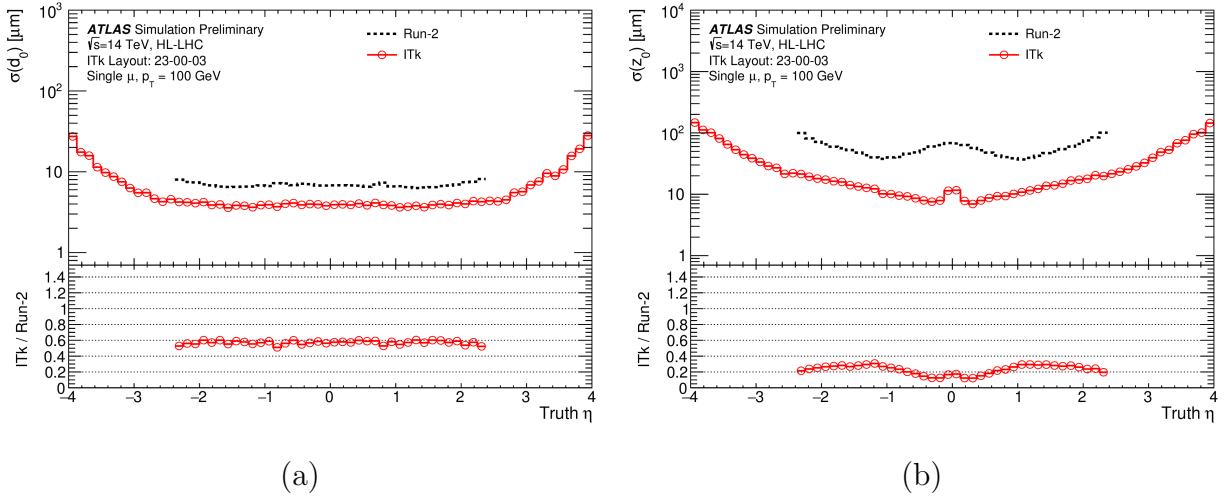


Figure 23: Transverse d_0 (a) and longitudinal z_0 (b) impact parameter resolution as a function of η for $p_T = 100$ GeV single muons without pile-up, shown separately for the latest ITk layout and the Run-2 detector [41].

Chapter IV: Analysis Overview

As discussed in Section 3.3, the four-lepton final state presents an interesting albeit rare process for precision tests of the Standard Model. Benefiting from a clean detector signature and relatively low backgrounds, four-lepton analyses are generally limited by statistics owing to the low production cross-sections involved. The dominant production mode $pp \rightarrow ZZ \rightarrow 4\ell$, mediated by the exchange of heavy gauge bosons, is naturally more suppressed in the Standard Model over lower energy interactions. To better illustrate this, the total production cross-section for a variety of Standard Model processes is shown as a function of center-of-mass energy in Figure 24. The uppermost process $pp \rightarrow X$ represents the total inelastic cross-section for proton-proton collisions at the LHC. This increases with the center-of-mass energy and was measured to be $(78.1 \pm 2.9) \times 10^9$ pb at $\sqrt{s} = 13$ TeV [42]. Inclusive $pp \rightarrow ZZ$ production is visible as the lowermost entry, with a measured cross-section of 17.3 ± 0.9 pb at 13 TeV [43]. Considering the $Z \rightarrow \ell\ell$ branching fraction of approximately 10% [4], this means roughly two in one trillion proton-proton collisions will give the $ZZ \rightarrow 4\ell$ final state.

While several previous LHC analyses have measured ZZ production inclusively, and exclusively in the four-lepton final state, the results have generally been limited by data statistics. The data versus theory agreement of several diboson cross-section measurements, including those aforementioned, is shown in Figure 25. Notably, there are some discrepancies in the ZZ measurements which cannot as of yet be determined statistically significant, owing to the large uncertainties. The analysis discussed here is the first ATLAS measurement of inclusive four-lepton production using the full Run-2 dataset, the final result of which is shown in the $ZZ^* \rightarrow 4\ell$ row of Figure 25. Although this measurement significantly benefits from the nearly fourfold increase in statistics over the previous analysis [44], combinations using data from Run-3 and beyond remain particularly promising in further reducing the statistical uncertainty below the bounds of the theoretical and experimental uncertainties.

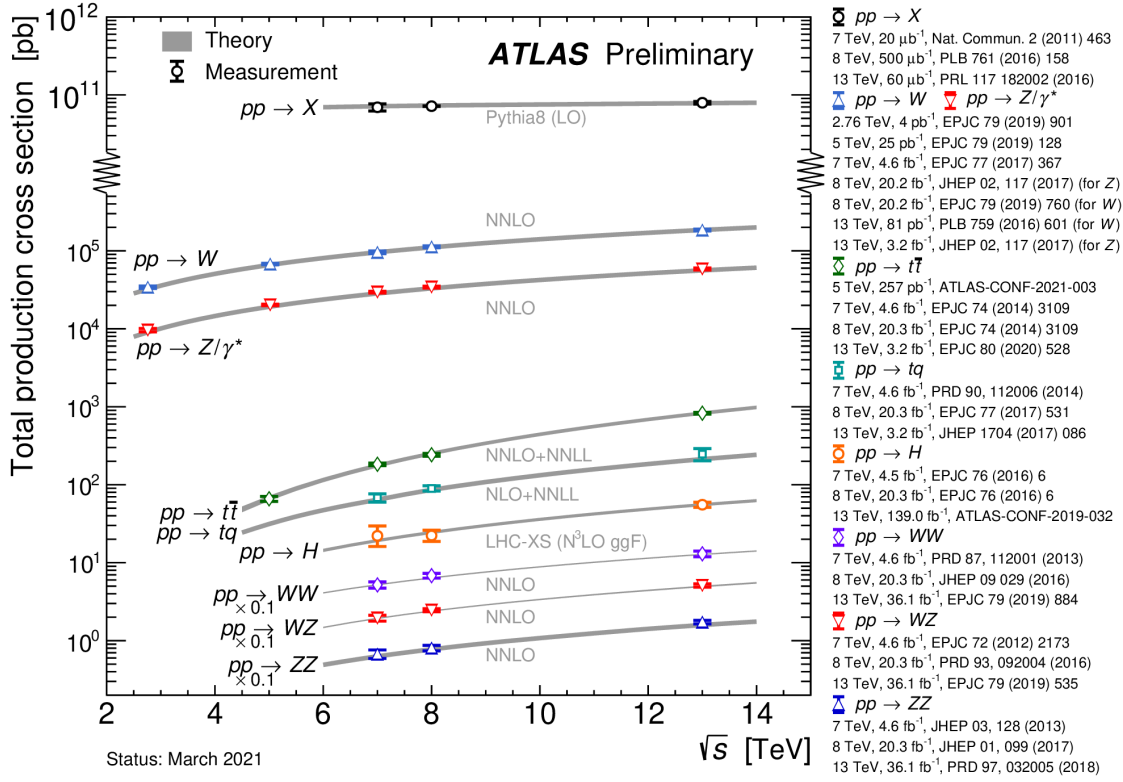


Figure 24: Predictions of Standard Model total production cross-sections for various processes as a function of center-of-mass energy up to 13 TeV. Experimental measurements of the cross-sections from Run-1 and Run-2 of the LHC are superimposed, with journal references listed on the right [45].

Diboson Cross Section Measurements

Status: March 2021

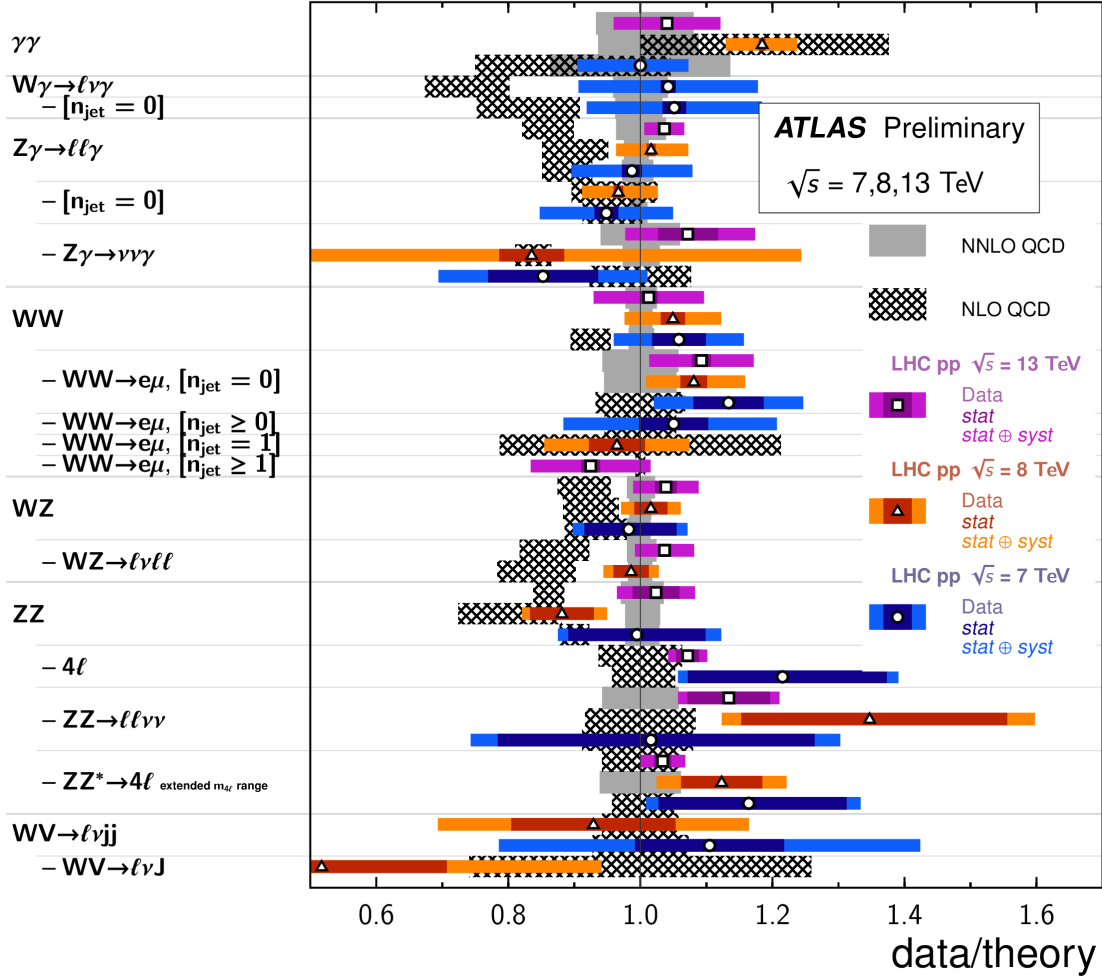


Figure 25: The data versus theory agreement for various diboson fiducial production cross-section measurements. The $\sqrt{s} = 13 \text{ TeV}$ cross-section measurement of this analysis is included in the $ZZ^* \rightarrow 4\ell$ row [45].

10 Analysis Goals

The goals of this analysis are to measure a variety of four-lepton differential cross-sections to use for precision studies of Standard Model processes and constraints on many scenarios of beyond-the-Standard Model physics. In order to exploit the versatility of the measurements, this analysis relies on a technique known as *unfolding* - a process which corrects for detector effects and facilitates a more straightforward usage of the measurements by those outside of the ATLAS Collaboration. As will be discussed in Section 17, unfolding consists in estimating the true, underlying distributions independent of detector effects by extrapolating the *detector-level* event yields obtained from reconstruction to *particle-level* cross-sections in a *fiducial* phase space. This has the advantage of disentangling the measurements from the specifics of ATLAS reconstruction, allowing for direct comparisons to theoretical predictions - based on the Standard Model or otherwise - in the fiducial phase space. The *unfolded* cross-sections are in turn used for precision studies and *re-interpretations* against arbitrary models of new physics affecting the four-lepton final state.

11 Fiducial Selection

The fiducial selection defines the particle-level phase space to which the measured reconstructed-level distributions are extrapolated via the unfolding procedure. This section outlines the selections defining the fiducial region of the analysis, which are summarized in Table 4.

11.1 Signal Definition

The fiducial region is inclusive of all $pp \rightarrow 4\ell$ ($\ell = e, \mu$) processes leading to the simultaneous production of at least four prompt leptons originating from the hard scatter interaction. Unlike electrons and muons, tau leptons are not stable within the volume of ATLAS and, therefore, are only measured via their decay products. This fiducial definition is inclusive of leptonically-decaying tau leptons from the hard scatter, but not of semileptonically-decaying tau leptons, generally referred to as hadronic taus. Of the prompt leptons in an event, there must exist at least one lepton quadruplet composed of two same-flavor, opposite-charge lepton pairs. This results in three possible final state flavor channels, outlined below:

- **4 μ** : $\mu^+\mu^-\mu^+\mu^-$
- **4e**: $e^+e^-e^+e^-$
- **2 μ 2e**: $\mu^+\mu^-e^+e^-$

The fiducial region is defined in terms of truth leptons, following one of three generator-level conventions:

- *Born* leptons are those produced after parton showering, but before the subsequent QED final state radiation (FSR) correction.
- *Bare* leptons correspond to stable, final state leptons after the QED FSR correction.
- *Dressed* leptons correspond to bare leptons after a *dressing* procedure, in which the four-momentum contribution of nearby photons within a cone of size $\Delta R < 0.1$ and excluding those originating from hadrons is added to the lepton four-momentum.

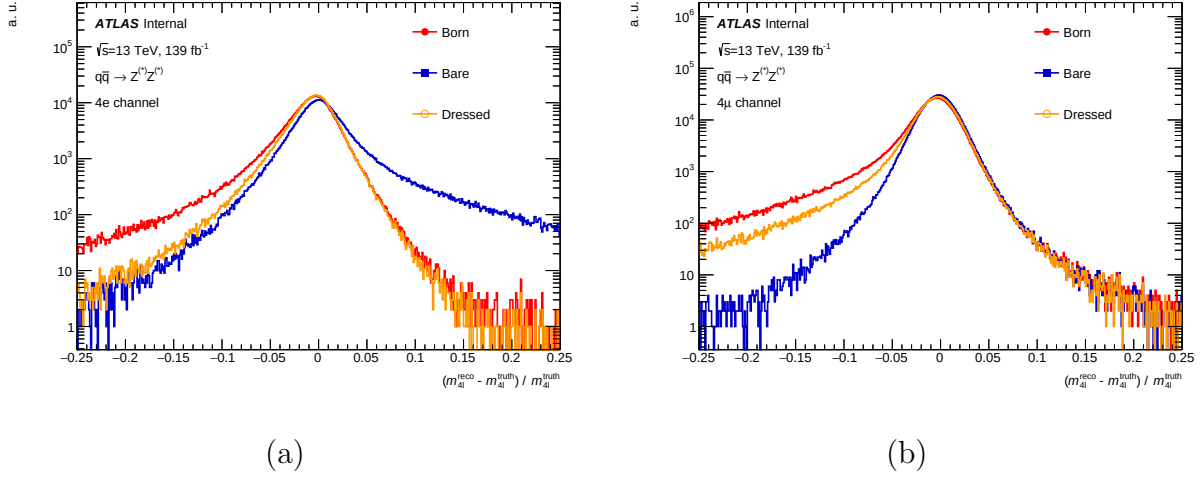


Figure 26: Resolution of the reconstructed four-lepton invariant mass for three definitions of particle-level leptons, shown separately for electrons (a) and muons (b).

The fiducial region is based on a hybrid definition of *dressed* electrons and *bare* muons, reflecting the inherent differences in electron and muon reconstruction in ATLAS.

As discussed in Section 8.2, electrons are primarily measured by adding clusters of energy deposits in the electromagnetic calorimeter, whereas muons are typically reconstructed using charged particle track measurements in the ID and MS. The important difference being that the former method necessarily includes the contribution of collinear radiative photons, while the latter does not. To clearly demonstrate this effect, the resolution of the reconstructed four-lepton invariant mass is shown based on the three definitions of particle-level leptons separately for electrons (a) and muons (b) in Figure 26. Concerning electrons, the dressing procedure best matches with the calorimeter-based reconstruction. In the case of muons, though, the bare definition provides the most accurate description of the track-based reconstruction. Therefore, defining the particle-level leptons in terms of *dressed* electrons and *bare* muons best conforms with what is actually measured in data, thereby minimizing extrapolations from detector- to particle-level in the unfolding procedure.

11.2 Kinematic Selections

Lepton kinematic selections are applied at particle-level to match the detector acceptance of ATLAS and the limits of high efficiency reconstruction. These selections are different for electrons and muons, owing once again to the inherent differences in reconstruction. Electrons are required to have $p_T > 7$ GeV and $|\eta| < 2.47$, while muons must satisfy $p_T > 5$ GeV and $|\eta| < 2.7$.

Additional kinematic selections are applied to match the phase-space limitations of available MC event generators and to suppress the backgrounds originating from misidentified or secondary leptons, including those from the well-known J/ψ resonance:

- The leading (subleading) lepton, as ranked in p_T , must satisfy $p_T > 20$ (10) GeV
- The angular separation between any two leptons in the event must satisfy $\Delta R > 0.05$
- The invariant mass of any SF-OC lepton pair in the event must satisfy $m_{\ell\ell} > 5$ GeV

Lastly, a simplified truth-level isolation criterion is enforced, approximating the more complex detector-level isolation working point used in reconstruction (see Appendix C.2), in order to minimize extrapolation uncertainties in the unfolding procedure. The measure of activity surrounding a lepton is defined as the p_T -sum of all charged particles, excluding any selected leptons, within a cone of size $\Delta R = 0.3$. In order to pass the selection, the resulting sum, referred to as ptcone30, must be less than 0.16 times the lepton p_T ,

$$\text{ptcone30} := \sum_{i, \Delta R_{i,\ell} < 0.3} p_{T,i} < 0.16 \times p_T. \quad (45)$$

Table 4: Definition of the fiducial region.

Lepton Selection	
Electrons	Dressed, $p_T > 7$ GeV, $ \eta < 2.47$
Muons	Bare, $p_T > 5$ GeV, $ \eta < 2.7$
Event Selection	
Final State	≥ 4 leptons, with 2 SF-OC pairs
Lepton Kinematics	$p_T > 20$ (10) GeV for leading (subleading) lepton
Lepton Separation	$\Delta R_{ij} > 0.05$ for any two leptons
J/ψ -Veto	$m_{ij} > 5$ GeV for all SF-OC lepton pairs
Truth Isolation	$\text{ptcone30} < 0.16 \times p_T$

12 Reconstruction Selection

The reconstruction selection defines the detector-level phase space in which the data is measured. As in the fiducial region, events are required to contain a lepton *quadruplet*, formed from two same-flavor, opposite-charge lepton pairs. *Baseline* leptons are required to pass a set of kinematic selections as well as vertex association, loose identification criteria, and a lepton-favored overlap removal (OR) procedure. *Signal* leptons are defined as the subset of baseline leptons also passing, among other selections, isolation and track-to-vertex association requirements. Although jets are not used in the final unfolded measurements, they enter in the overlap removal stage and are an important component of the fake factor background estimation. As for leptons, *baseline* and *signal* jets are defined based on sets of looser and stricter selections, respectively. The final unfolded measurements consist of only signal leptons, while baseline leptons enter in control regions used for the data-driven reducible background estimation. The baseline and signal object selections are summarized in Tables 5, 6, and 7. The final event selection is summarized in Table 8.

12.1 Physics Objects

Electrons

Baseline electrons are required to have $p_T > 7$ GeV and $|\eta| < 2.47$ and pass the loose identification working point (LOOSELH ID). An additional vertex association requirement based on the longitudinal impact parameter of $|z_0 \sin \theta| < 0.5$ mm is applied for all electrons. Baseline electrons enter the overlap removal as described in Section 12.2.

In addition to the baseline selections, signal electrons are required to pass the FIXED-CUTPFLOWLOOSE isolation working point and a transverse impact parameter significance requirement of $|d_0/\sigma_{d_0}| < 5$. In order to reduce backgrounds from photon conversions, a stricter identification working point (LOOSEBLAYERLH ID), requiring a silicon hit in the innermost b-layer, is enforced for signal electrons.

Table 5: Summary of the baseline lepton selection.

Category	Requirement
Kinematics	Electrons: $p_T > 7 \text{ GeV}$ $ \eta < 2.47$
	Muons: $p_T > 5 \text{ GeV}$ If Calo-Tagged: $p_T > 15 \text{ GeV}$ $ \eta < 2.7$
Vertex Association	Both: $ z_0 \sin \theta < 0.5 \text{ mm}$
Identification	Electrons: LOOSELH ID
	Muons: LOOSE ID
Overlap Removal	Both: Lepton-favored OR

Muons

Baseline muons in this analysis are required to have $p_T > 5 \text{ GeV}$, or $p_T > 15 \text{ GeV}$ if they are Calo-Tagged, and $|\eta| < 2.7$ and pass the loose identification working point (LOOSE ID). As for electrons, a vertex association requirement of $|z_0 \sin \theta| < 0.5 \text{ mm}$ is applied for all muons. Baseline muons enter the overlap removal as described in Section 12.2.

In addition to the baseline selections, signal muons are required to pass the FIXED-CUTPFLOWLOOSE isolation working point and a transverse impact parameter significance requirement of $|d_0/\sigma_{d_0}| < 3$. Lastly, to reduce backgrounds from cosmic muons, a transverse impact parameter cut of $|d_0| < 1 \text{ mm}$ is applied.

Jets

Jets are reconstructed following the particle flow anti- k_T jet clustering algorithm using a radius parameter of 0.4 (ANTIKT4EMPFLOW) [38]. Baseline jets are required to have $p_T > 30 \text{ GeV}$ and $|\eta| < 4.5$. In addition to the baseline kinematic selections, signal jets are required to pass the jet-vertex tagger (JVT) after the overlap removal stage to reduce contamination from pile-up.

Table 6: Summary of the signal lepton selection.

Category	Requirement	
Isolation	Both:	FIXEDCUTPFLOWLOOSE working point
Cosmic Muon Veto	Muons:	$ d_0 < 1$ mm
Impact Parameter	Electrons:	$ d_0/\sigma_{d_0} < 5$
	Muons:	$ d_0/\sigma_{d_0} < 3$
Stricter Identification	Electrons:	LOOSEBLAYERLH ID

Table 7: Summary of the baseline and signal jet selections.

Category	Requirement
Collection:	ANTIKT4EMPFLOW
Kinematics:	$p_T > 30$ GeV
	$ \eta < 4.5$
Overlap Removal	Lepton-favored OR
Vertex Association	If signal: pass JVT

12.2 Overlap Removal

Overlap removal is a standard ATLAS procedure designed to remove analysis objects reconstructed from the same detector signals. The configuration used in this analysis is a lepton-favored working point which places a higher priority on leptons than jets in the decision of which object to discard. The process proceeds iteratively through a set of overlap criteria, in which only objects surviving all previous steps are used in the subsequent steps. A summary of the overlap removal steps are as follows, where the ΔR angular separation is calculated using rapidity:

1. if electrons share an ID track, the one with lower p_T is rejected,
2. Calo-Tagged muons sharing an ID track with an electron are rejected,
3. electrons sharing an ID track with a muon are rejected,
4. jets within $\Delta R < 0.2$ of an electron are rejected,
5. ghost-associated jets¹⁴ with $N_{\text{track}} < 3$ and within $\Delta R < 0.2$ of a muon are rejected.

¹⁴A ghost-associated jet refers to a jet associated to the hard scatter vertex by means of (*ghost*) tracks which are clustered within the bounds of the jet but not considered in the energy measurement.

Table 8: Summary of the event-level reconstruction selection.

Category	Requirement
Event Preselection	Fire ≥ 1 lepton trigger ≥ 1 vertex with ≥ 2 tracks
Final State	≥ 4 leptons (e, μ)
Lepton Kinematics	$p_T > 20$ (10) GeV for leading (subleading) lepton
Lepton Separation	$\Delta R_{ij} > 0.05$ for any two leptons
J/ψ -Veto	$m_{ij} > 5$ GeV for all SF-OC lepton pairs
Trigger Matching	Baseline leptons matched to ≥ 1 lepton trigger
Quadruplet Formation	≥ 1 quadruplet, with 2 SF-OC pairs
Quadruplet Categorization	Signal: 4 signal, 0 non-signal Inclusive Control Region: ≤ 3 signal, ≥ 1 non-signal

12.3 Event Selection

Lepton quadruplets are formed from two lepton pairs consisting of baseline leptons. The primary lepton pair is formed from the two same-flavor, opposite-charge leptons with a combined invariant mass closest to the mass of the Z -boson (m_Z), while the secondary pair is chosen from the remaining leptons which form a same-flavor, opposite-charge pair with an invariant mass closest to m_Z . In this way, a single, unique quadruplet is formed in all events containing at least four compatible leptons. Quadruplets entering the signal region must consist entirely of signal leptons, while quadruplets with baseline, non-signal leptons are used in the fake background estimation discussed in Section 16.

For an event to be selected, at least one of the leptons in the quadruplet must be matched to one of the offline triggers which fired for the event. To ensure a high trigger efficiency, a requirement of $p_T > 20$ (10) GeV is required for the leading (subleading) lepton of the quadruplet. Lastly, as for the fiducial region, loose ΔR and $m_{\ell\ell}$ selections are applied to minimize known backgrounds without significantly impacting the acceptance.

To illustrate the typical quadruplet signature, a representative event display, recorded during Run Number 271298 of 2015 data-taking, is shown for a candidate $ZZ \rightarrow \mu\mu ee$ event in Figure 27. The primary lepton pair is formed from two oppositely charged muons, while the secondary pair is formed from two oppositely charged electrons.

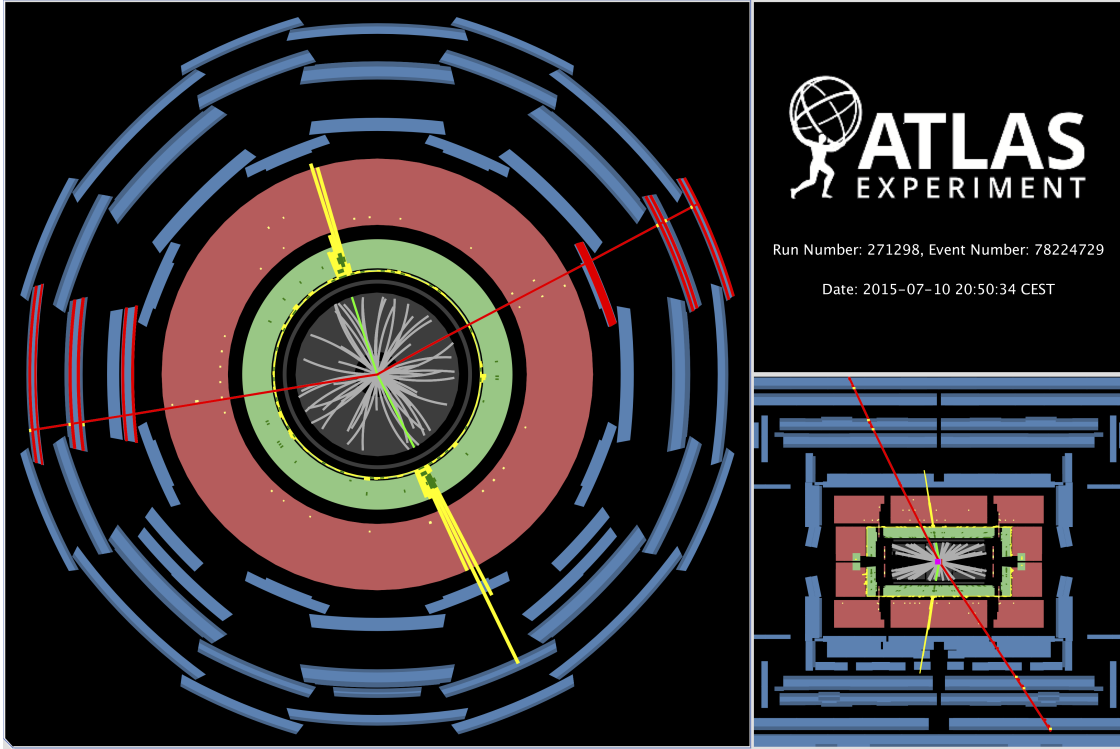


Figure 27: Display of a candidate $ZZ \rightarrow \mu\mu ee$ event from proton-proton collisions recorded by ATLAS, with LHC beams at a collision energy of $\sqrt{s} = 13$ TeV. The primary Z -boson candidate has a dilepton invariant mass of 94 GeV and transverse momentum of 35 GeV, and is reconstructed from two oppositely charged muons, shown as red tracks. The secondary Z -boson candidate has a mass of 86 GeV and transverse momentum of 19 GeV, and is reconstructed from two oppositely charged electrons, shown as green tracks leading to energy deposits in the electromagnetic calorimeter. The four-lepton system has a combined invariant mass of 191 GeV and transverse momentum of 24 GeV. Taken from ATLAS public results <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun2Collisions>.

13 Dataset and Monte Carlo Samples

13.1 LHC Dataset

The LHC collision data used in this analysis, known as the ATLAS Run-2 dataset, correspond to proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV totaling an integrated luminosity of 139 fb^{-1} . This dataset consists of four individual run periods, divided according to the year of operation: 2015, 2016, 2017, and 2018. These periods are further divided into sub-periods, lasting on the order of one to three weeks, with varying beam and detector conditions. For each run period, only those subset of events satisfying the relevant Good Run List (GRL) are considered. Events on the GRL must pass detector condition and data quality thresholds to ensure the data are suitable for physics analysis.

The total integrated luminosity over the course of Run-2 is shown as a function of time in Figure 28. Three measures of the luminosity are shown: the total as delivered by the LHC (“LHC Delivered”), the total as recorded by the ATLAS detector (“ATLAS Recorded”), and the total passing the relevant GRL (“Good for Physics”). The previously quoted Run-2 luminosity of 139 fb^{-1} corresponds to the lattermost. The general trend of increasing instantaneous luminosity over time can be seen in the increasing slope of the three graphs. This is also visible in the average number of inelastic proton-proton collisions observed per bunch crossing, termed pile-up, which is shown for each of the four run periods in Figure 29. Roughly speaking, a higher instantaneous luminosity corresponds to a denser packing of proton bunches within the beams, which in turn leads to more collisions per bunch crossing. The figures are summarized for each of the four run periods in Table 9.

The LHC beam luminosity, as delivered to ATLAS, is primarily measured by the Luminosity Cherenkov Integrating Detector (LUCID-2) [46]. An overall relative uncertainty of $\pm 1.7\%$, equal to $\pm 2.4 \text{ fb}^{-1}$ in absolute terms, has been determined for the full Run-2 operation [47].

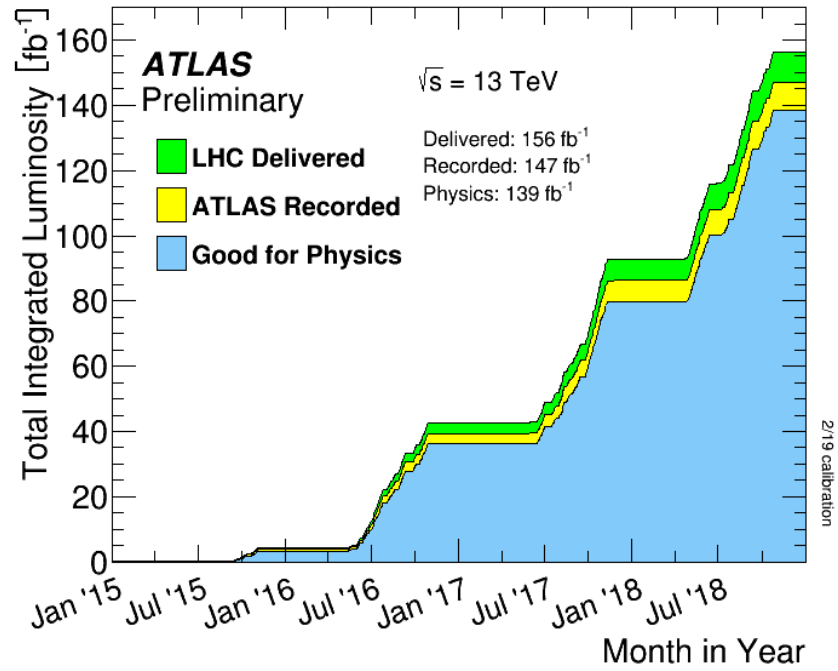


Figure 28: Total integrated luminosity over the course of the Run-2 operation.

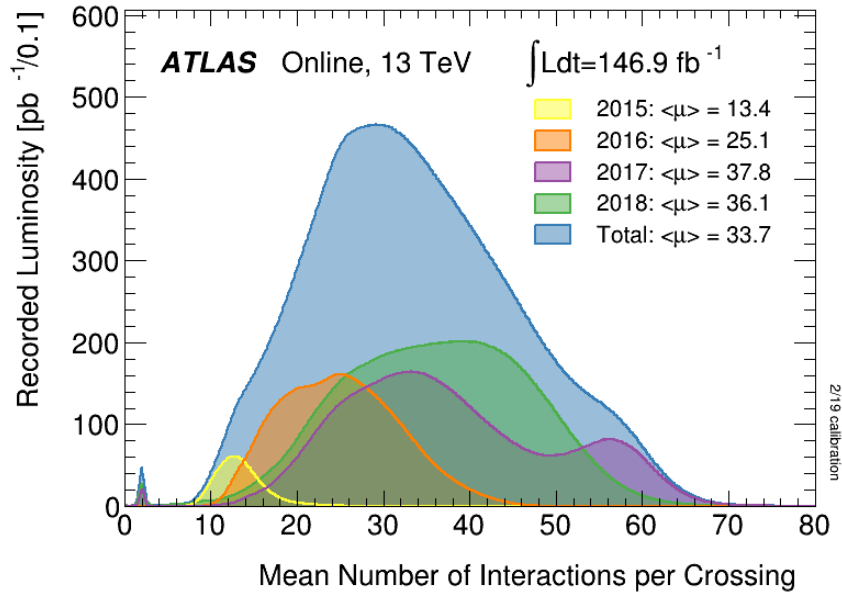


Figure 29: Average number of inelastic proton-proton collisions per bunch crossing (pile-up) for each run period and the combination of Run-2.

Table 9: Summary of luminosity conditions for the four periods of Run-2. The luminosities as delivered by the LHC (“Delivered”), recorded by ATLAS (“Recorded”), and passing the relevant GRL (“Good for Physics”) are shown in units of fb^{-1} . The average number of inelastic proton-proton interactions per bunch crossing (“Pile-up”) is also shown to further highlight the general trend of increasing luminosity.

Run Period (Year)	Delivered (fb^{-1})	Recorded (fb^{-1})	Good for Physics (fb^{-1})	Pile-up
2015	4.0	3.7	3.2	13.4
2016	38.1	35.5	33.0	25.1
2017	49.7	46.6	44.3	37.8
2018	64.2	60.9	58.5	36.1
Run-2	156.0	146.7	139.0	33.7

13.2 Monte Carlo Samples

The signal and background physics processes in the analysis phase space are simulated using Monte Carlo (MC) event generation, following the procedure outlined in Section 7.2. The derivation stage following reconstruction relies on the HIGG2D1 scheme, provided by the Higgs analysis group, which applies a dilepton (e, μ) filter, at least one electron or muon trigger to be fired, and a minimum dilepton invariant mass of $m_{\ell\ell} > 5 \text{ GeV}$. Additionally, the transverse momentum of the leading and subleading leptons are required to exceed 15 GeV and 5 GeV, respectively. A summary of the MC samples considered as signal is shown in Table 10, while samples considered as reducible background are listed in Table 11.

Signal Processes

Details on multi-boson simulation in ATLAS are discussed further in Ref. [48]. Additional details on the specific samples used in this analysis can be found in Ref. [49]. All signal and background samples generated with SHERPA [50] use the NNPDF3.0nnlo [51] parton distribution function (PDF) set¹⁵. The dominant quark-initiated $q\bar{q} \rightarrow Z^{(*)}Z^{(*)} \rightarrow 4\ell$ process, including resonant $Z \rightarrow 4\ell$, is modeled using SHERPA 2.2.2, with next-to-leading order

¹⁵ Given the complexity of composite hadrons, a choice of parton distribution functions is required to accurately model their collisions. These are determined experimentally and estimate the probability of finding a given parton (i.e. u, d, s, g) as a function of its fractional momentum of the hadron, often denoted by x . For a more detailed overview see Ref. [52].

(NLO) accuracy in QCD for up to one additional parton emission, and leading order (LO) accuracy for up to three additional emissions. An alternative sample used for comparison to the nominal Sherpa $q\bar{q} \rightarrow 4\ell$ sample is generated at NLO accuracy in QCD, excluding any parton emissions, with POWHEG-BOX v2 [53] using the CT10nlo [54] PDF set, and interfaced to PYTHIA 8 [55] for the parton shower and hadronization. The higher order electroweak production of four leptons in association with two jets is modeled using a separate sample generated with SHERPA 2.2.2. The loop-induced, gluon-initiated $gg \rightarrow Z^{(*)}Z^{(*)} \rightarrow 4\ell$ process, emerging at next-to-next-to-leading order (NNLO) in α_S , is modeled using SHERPA 2.2.2, with LO precision for up to one additional parton emission. Resonant Higgs production $pp \rightarrow H^{(*)} \rightarrow Z^{(*)}Z^{(*)} \rightarrow 4\ell$ is simulated using POWHEG+PYTHIA8 to at least NLO accuracy in QCD, separately for the four primary Higgs production modes: gluon-gluon fusion (ggF), vector boson fusion (VBF), Higgsstrahlung (VH), and production in association with top quarks ($t\bar{t}H$) - where the ggF sample uses the POWHEG NNLOPS program [56] to achieve NNLO accuracy in QCD. Triboson processes which produce four or more prompt leptons in the final state, including WWZ , WZZ , ZZZ , are modeled with SHERPA 2.2.2 at NLO accuracy in QCD for zero or one additional parton emissions, and LO accuracy for up to two additional parton emissions. Lastly, Z -boson production in association with a pair of top quarks $pp \rightarrow t\bar{t} + (Z \rightarrow)2\ell$ is modeled with SHERPA 2.2.0.

Background Processes

The sources of reducible background in this analysis correspond to events with one or more non-prompt or fake leptons. Although the fake background is estimated using a fully data-driven method, MC samples are used to develop and validate the estimation procedure. The dominant background process is that of Z -boson production in association with jets $pp \rightarrow Z^{(*)} \rightarrow 2\ell + \text{jets}$, which is simulated for both three or more leptons using SHERPA 2.2.0 and inclusively using SHERPA 2.2.1, in order to balance the tradeoffs between statistical precision and inclusivity. The subdominant background process is that of $t\bar{t}$ production in

which both top quarks decay semileptonically $pp \rightarrow t\bar{t} \rightarrow 2\ell$. This process is simulated with POWHEG+PYTHIA8 using the A14 NNPDF23LO [58] PDF set. An additional contribution from WZ production in which both bosons decay leptonically $pp \rightarrow WZ \rightarrow 3\ell 1\nu$ is simulated using SHERPA 2.2.2. Lastly, a very minor contribution from $Z+\Upsilon$ production in multi-parton interactions $pp \rightarrow Z + \Upsilon \rightarrow 4\ell$ is simulated with PYTHIA 8 using the A14 NNPDF23LO PDF set.

Calibrations

A variety of calibrations are applied to the simulations to account for known discrepancies with measured collision data. The MC is split into three campaigns known as mc16a, mc16d, and mc16e - corresponding, respectively, to the 2015-2016, 2017, and 2018 data-taking periods. For each campaign, the MC is reweighted to match the observed pile-up profile in the corresponding data period. Additional detector-level correction weights, known as scale factors, are applied to reconstructed events in MC to account for other differences with data. These scale factors generally fall into one of three categories, each accounting for differences in a specific selection efficiency between data and MC:

- **Electron:** reconstruction, identification, and isolation efficiency of electrons
- **Muon:** reconstruction, isolation, and track-to-vertex association efficiency of muons
- **Trigger:** combined event-level trigger efficiency (see Section 15)

Table 10: A summary of the Standard Model Monte Carlo samples considered as signal. The last column denotes the accuracy in QCD depending on the number of additional parton emissions. The sample indicated with a dagger highlights the alternative $q\bar{q} \rightarrow Z^{(*)}Z^{(*)} \rightarrow 4\ell$ sample used to derive a generator uncertainty with respect to the nominal.

Process	Filter	Generator	PDF	Accuracy
$q\bar{q} \rightarrow Z^{(*)}Z^{(*)} \rightarrow 4\ell$	inclusive	SHERPA 2.2.2 [50]	NNPDF3.0nnlo [51]	0, 1j@NLO+2, 3j@LO
	inclusive [†]	POWHEG+PYTHIA8 [53, 55]	CT10nlo [54]	0j@NLO
	≥ 2 jets (EW)	SHERPA 2.2.2	NNPDF3.0nnlo	0j@LO
$gg \rightarrow Z^{(*)}Z^{(*)} \rightarrow 4\ell$	$m_{4\ell} > 130$ GeV	SHERPA 2.2.2	NNPDF3.0nnlo	0, 1j@LO
	no H , $m_{4\ell} < 130$ GeV	SHERPA 2.2.2	NNPDF3.0nnlo	0, 1j@LO
$pp \rightarrow H^{(*)} \rightarrow Z^{(*)}Z^{(*)} \rightarrow 4\ell$	ggF	POWHEG+PYTHIA8	PDF4LHC15nnlo [57]	NNLO
	VBF	POWHEG+PYTHIA8	PDF4LHC15nlo	NLO
	VH	POWHEG+PYTHIA8	PDF4LHC15nlo	NLO
	ttH	POWHEG+PYTHIA8	PDF4LHC15nlo	NLO
$pp \rightarrow W^{(*)}W^{(*)}Z^{(*)} \rightarrow 4\ell 2\nu$	inclusive	SHERPA 2.2.2	NNPDF3.0nnlo	0j@NLO+1, 2j@LO
$pp \rightarrow W^{(*)}Z^{(*)}Z^{(*)} \rightarrow 5\ell 2\nu$	inclusive	SHERPA 2.2.2	NNPDF3.0nnlo	0j@NLO+1, 2j@LO
$pp \rightarrow Z^{(*)}Z^{(*)}Z^{(*)} \rightarrow 6\ell$	inclusive	SHERPA 2.2.2	NNPDF3.0nnlo	0j@NLO+1, 2j@LO
$pp \rightarrow Z^{(*)}Z^{(*)}Z^{(*)} \rightarrow 4\ell 2\nu$	inclusive	SHERPA 2.2.2	NNPDF3.0nnlo	0j@NLO+1, 2j@LO
$pp \rightarrow t\bar{t} + (Z \rightarrow) 2\ell$	$m_{t\bar{t}} > 5$ GeV	SHERPA 2.2.0 [50]	NNPDF3.0nnlo	LO

Table 11: A summary of the Standard Model Monte Carlo samples considered as background.

Process	Filter	Generator	PDF
$pp \rightarrow Z^{(*)} \rightarrow 2e+\text{jets}$	≥ 4 truth leptons, $p_T^\ell > 4$ GeV	SHERPA 2.2.0 [50]	NNPDF3.0nnlo [51]
	$m_1(\ell\ell) > 40$ GeV		
	$m_2(\ell\ell) > 8$ GeV		
$pp \rightarrow Z^{(*)} \rightarrow 2\mu+\text{jets}$	≥ 4 truth leptons, $p_T^\ell > 4$ GeV	SHERPA 2.2.0	NNPDF3.0nnlo
	$m_1(\ell\ell) > 40$ GeV		
	$m_2(\ell\ell) > 8$ GeV		
$pp \rightarrow Z^{(*)} \rightarrow 2e+\text{jets}$	≥ 3 truth leptons, $p_T^\ell > 4$ GeV vetoing 4ℓ filter	SHERPA 2.2.0	NNPDF3.0nnlo
$pp \rightarrow Z^{(*)} \rightarrow 2\mu+\text{jets}$	≥ 3 truth leptons, $p_T^\ell > 4$ GeV vetoing 4ℓ filter	SHERPA 2.2.0	NNPDF3.0nnlo
$pp \rightarrow Z^{(*)} \rightarrow 2e+\text{jets}$	inclusive	SHERPA 2.2.1 [50]	NNPDF3.0nnlo
$pp \rightarrow Z^{(*)} \rightarrow 2\mu+\text{jets}$	inclusive	SHERPA 2.2.1	NNPDF3.0nnlo
$pp \rightarrow Z^{(*)} \rightarrow 2\tau+\text{jets}$	inclusive	SHERPA 2.2.1	NNPDF3.0nnlo
$pp \rightarrow t\bar{t} \rightarrow 2\ell$	inclusive	POWHEG+PYTHIA8 [53, 55]	A14 NNPDF23LO [58]
$pp \rightarrow WZ \rightarrow 3\ell 1\nu$	inclusive	SHERPA 2.2.2	NNPDF3.0nnlo
$pp \rightarrow Z + \Upsilon \rightarrow 4\ell$	inclusive	PYTHIA 8	A14 NNPDF23LO

14 Observables

The presence of four leptons in the final state gives many interesting kinematic observables to study, particularly those pertaining to angular separations - a sample of which are displayed in Figure 30. While a wide-variety of lepton- and event-level observables are utilized for internal validation studies, this section outlines only those observables considered for the final unfolded measurements.

As in the previous inclusive four-lepton analysis [44], the four-lepton invariant mass ($m_{4\ell}$) is measured inclusively and differentially, i.e. in slices of the transverse momentum ($p_{T,4\ell}$), rapidity ($y_{4\ell}$), and lepton flavor of the quadruplet system. The lepton flavor slices correspond to the three final state channels: 4μ , $4e$, and $2\mu 2e$. New to this analysis, several observables are measured differentially in four slices of the four-lepton invariant mass. These mass slices are chosen to be enriched in four separate 4ℓ production modes:

- **Single resonant Z :** $60 < m_{4\ell} < 100$ GeV
- **Resonant Higgs:** $120 < m_{4\ell} < 130$ GeV
- **Off-shell ZZ :** $20 < m_{4\ell} < 60$ GeV or $100 < m_{4\ell} < 120$ GeV or $130 < m_{4\ell} < 180$ GeV
- **On-shell ZZ :** $180 < m_{4\ell} < 2000$ GeV

Within these mass slices, the invariant mass and transverse momentum of the primary and secondary lepton pairs are measured. Adding to the differential measurements are the absolute rapidity and azimuthal separation of the primary and secondary lepton pairs. The absolute azimuthal separation between the two leading- p_T leptons of the quadruplet is also measured. Completing the differential observables, polarization-sensitive variables are measured for both lepton pairs, defined as the cosine of the angle between the negative lepton in its respective pair rest frame and the same lepton pair in the laboratory frame, as displayed in Figure 30. In addition to these observables, a series of five *inclusive* measurements are performed - one in the entire fiducial phase space and one in each of the four four-lepton

invariant mass regions outlined above. The thirteen differential observables are listed below according to the relevant variables profiled.

- **Single-differential $m_{4\ell}$:**

- Four-lepton invariant mass: $m_{4\ell}$

- **Double-differential, versus $m_{4\ell}$:**

- Four-lepton invariant mass, sliced in transverse momentum: $m_{4\ell} - p_{T,4\ell}$
- Four-lepton invariant mass, sliced in rapidity: $m_{4\ell} - y_{4\ell}$
- Four-lepton invariant mass, sliced in quadruplet flavor: $m_{4\ell}$ –flavor

- **Double-differential, sliced in $m_{4\ell}$:**

- Invariant mass of the primary (secondary) lepton pair: m_{Z_1} (m_{Z_2})
- Transverse momentum of the primary (secondary) lepton pair: p_{T,Z_1} (p_{T,Z_2})
- Absolute rapidity separation between the lepton pairs: $|\Delta y_{\text{pairs}}|$
- Absolute azimuthal separation between the lepton pairs: $|\Delta\phi_{\text{pairs}}|$
- Absolute azimuthal separation between the leading- p_T quadruplet leptons: $|\Delta\phi_{\ell\ell}|$
- Cosine of the angle between the negative lepton in the primary (secondary) pair frame and the primary (secondary) pair in the laboratory frame: $\cos\theta_1^*$ ($\cos\theta_2^*$)

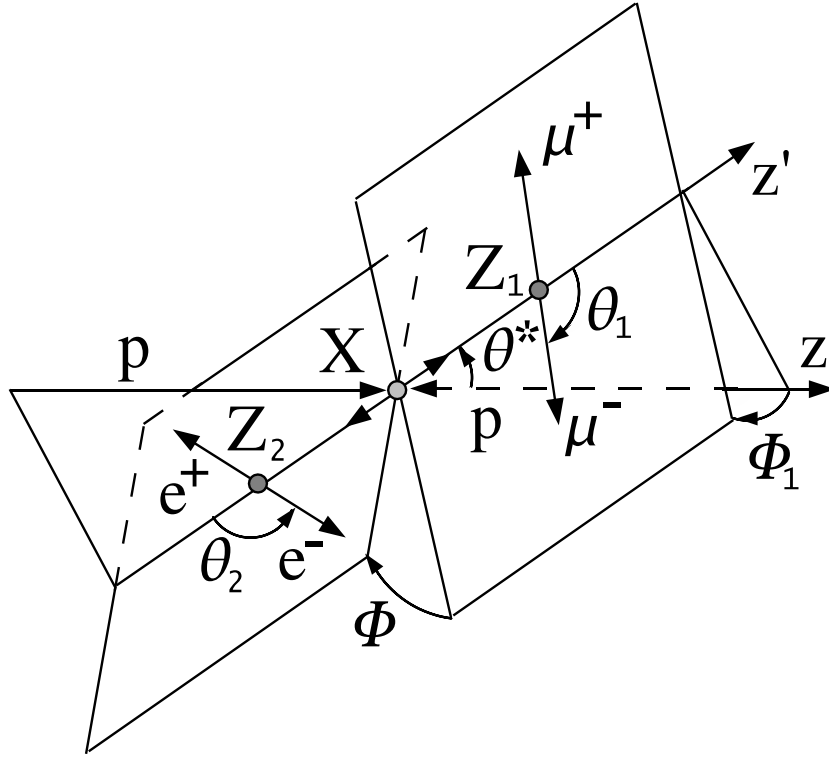


Figure 30: Illustration of an example $pp \rightarrow ZZ \rightarrow \mu^+\mu^-e^+e^-$ decay, showing the six angles (excluding the arbitrary azimuthal offset Φ^*) which fully characterize the decay chain [59].

Chapter V: Analysis Strategy

The overarching aims of this analysis are that of *inclusivity* and *re-interpretability*. The former goal motivates, among other decisions, the loosening of many kinematic selections and the inclusion of leptonic τ -decays in the signal region, as outlined in Section 12. These changes greatly extend the reach of this analysis over previous iterations but pose many challenges in the triggering and background estimation, which will be discussed in Sections 15 and 16, respectively. The effects of loosening selections and updating the lepton isolation working point with respect to the previous analysis are discussed in Appendix C. Re-interpretability is the ability to place strict constraints on many models of new physics using the measurements, which depends crucially on minimizing model dependence. The data-driven background estimation and the unfolding procedure, to be discussed in Sections 16 and 17, respectively, are central to achieving this goal. Lastly, physical measurements mean nothing without associated uncertainties. Uncertainties in this analysis arise from systematic and statistical sources, and are discussed extensively in Section 18.

15 Trigger

15.1 Menu

Given the analysis aim of inclusivity, the trigger menu was chosen to maximize the acceptance in the four-lepton phase space, particularly at low invariant mass. The menu therefore consists of an expansive combination of the lowest prescaled single, double, and triple lepton triggers. The exact menu of triggers varies according to the run period, reflecting changes in the high-level trigger system intended to cope with increasing data rates. Events in data are selected using a logical-OR of the active triggers in the corresponding run period. An additional requirement known as trigger matching selects the subset of events in which at least one lepton of the quadruplet is matched to one of the fired triggers. A breakdown of the trigger menu by electron-only, muon-only, and mixed electron-muon triggers is shown per data period in Table 12.

15.2 Efficiency

Calculation

The trigger efficiency in MC is defined as the fraction of events passing a logical-OR selection of the active triggers with respect to the number of events passing the reconstruction-level pre-trigger selection. The ATLAS-supported tool `TrigGlobalEfficiencyCorrectionTool`¹⁶ is used in order to calculate the trigger efficiency scale factors which are applied to MC in order to account for differences in the trigger efficiency with respect to data. As is generally the case, the trigger efficiency scale factor is defined as the ratio of the measured efficiencies between MC and data,

$$\text{SF}^{\text{trigger}} = \frac{\epsilon_{\text{MC}}^{\text{trigger}}}{\epsilon_{\text{Data}}^{\text{trigger}}}. \quad (46)$$

The ATLAS global trigger tool supports the computation of trigger efficiency scale factors

¹⁶<https://gitlab.cern.ch/atlas/athena/tree/21.2/Trigger/TrigAnalysis/TrigGlobalEfficiencyCorrection>

Table 12: Electron-only, muon-only, and mixed electron-muon triggers used for this analysis list by active data period.

Year	Leptons	Included Triggers
2015	Electron:	HLT_e24_lhmedium_L1EM20VH
		HLT_e60_lhmedium
		HLT_e120_lhloose
		HLT_2e12_lhvloose_L12EM10VH
	Muon:	HLT_mu20_iloose_L1MU15
		HLT_mu50
		HLT_2mu10
		HLT_mu18_mu8noL1
	Mixed:	HLT_e7_lhmedium_mu24
		HLT_e17_lhloose_mu14
2016	Electron:	HLT_e26_lhtight_nod0_ivarloose
		HLT_e60_lhmedium_nod0
		HLT_e140_lhloose_nod0
		HLT_2e17_lhvloose_nod0
	Muon:	HLT_mu26_ivarmedium
		HLT_mu50
		HLT_2mu14
		HLT_mu22_mu8noL1
	Mixed:	HLT_e7_lhmedium_nod0_mu24
		HLT_e17_lhloose_nod0_mu14
2017	Electron:	HLT_e26_lhtight_nod0_ivarloose
		HLT_e60_lhmedium_nod0
		HLT_e140_lhloose_nod0
		HLT_2e24_lhvloose_nod0
	Muon:	HLT_e24_lhvloose_nod0_2e12_lhvloose_nod0_L1EM20VH_3EM10VH
		HLT_mu26_ivarmedium
		HLT_mu50
		HLT_2mu14
	Mixed:	HLT_mu22_mu8noL1
		HLT_e17_lhloose_nod0_mu14
2018	Electron:	HLT_e26_lhtight_nod0_ivarloose
		HLT_e60_lhmedium_nod0
		HLT_e140_lhloose_nod0
		HLT_2e17_lhvloose_nod0_L12EM15VHI
	Muon:	HLT_2e24_lhvloose_nod0
		HLT_e24_lhvloose_nod0_2e12_lhvloose_nod0_L1EM20VH_3EM10VH
		HLT_mu26_ivarmedium
		HLT_mu50
	Mixed:	HLT_2mu14
		HLT_mu22_mu8noL1
	Electron:	HLT_e17_lhloose_nod0_mu14
		HLT_e26_lhmedium_nod0_mu8noL1
	Muon:	HLT_e17_lhloose_nod0_mu14
		HLT_e26_lhmedium_nod0_mu8noL1
	Mixed:	HLT_e17_lhloose_nod0_mu14
		HLT_e26_lhmedium_nod0_mu8noL1

either by explicit calculation, based on the combined probabilities of unions, or by generating an ensemble of pseudo-experiments. In the former case, the tool supports only restricted combinations of triggers, given the increasingly complicated analytic formulae required by complex configurations. In the simplest case of a trigger menu consisting of only one single lepton trigger ($\ell = e, \mu$), the probability of the trigger firing in an event with n leptons, essentially a measure of the trigger efficiency, is calculated by the logical-OR of the probability of firing for each lepton,

$$P(t_n) = P(\ell_1 \vee \ell_2 \dots \vee \ell_n). \quad (47)$$

This can be factorized and, therefore, more easily calculated by considering the probability of all individual leptons not firing the trigger,

$$P(t_n) = 1 - P(\bar{\ell}_1 \wedge \bar{\ell}_2 \dots \wedge \bar{\ell}_n) = 1 - \prod_{i=1}^n (1 - P(\ell_i)). \quad (48)$$

In general, however, this computation cannot be factorized so cleanly and requires considering the respective contributions from both lepton flavors.

Initially, a reduced set of triggers was chosen in order to benefit from an explicit calculation of scale factors. However, the inclusion of mixed triple-lepton triggers, which is analytically supported only in very simple trigger menus, was found to yield a non-negligible acceptance gain at low four-lepton invariant mass. This motivated a detail study of the pseudo-experiment based approach. Using a reduced, analytically-supported trigger menu, the trigger efficiencies in data and MC were calculated separately for both approaches, using 100,000 toys for the pseudo-experiment option. As shown in Figure 31, the calculation by pseudo-experiments gives per-mille agreement in the data (a) and MC (b) trigger efficiencies with the explicit calculation.

Given the results of this validation, the pseudo-experiment option was chosen over the explicit calculation, and the trigger menu was expanded to include all triggers with full scale factor support, as shown in Table 12. The resulting MC and data trigger efficiencies, as

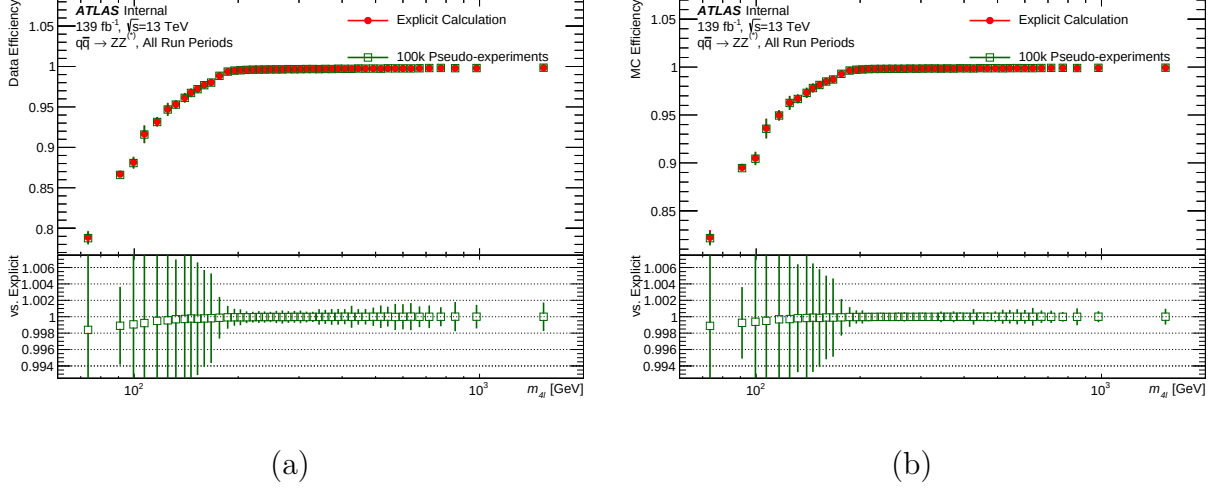


Figure 31: Validation of the trigger efficiency computation by the pseudo-experiment based approach ($n_{toys} = 100,000$) against the explicit calculation using analytic formulae, shown separately for the resulting data (a) and MC (b) trigger efficiencies.

well as the corresponding scale factors, are shown as a function of $m_{4\ell}$ for the dominant $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$ process in Figure 32. The trigger efficiency is above 50% in all bins and saturates to 100% above 200 GeV. The following plots are restricted to an upper limit of 253 GeV to highlight the low-mass region where the differences are noticeable. As expected, the efficiency calculated in MC overestimates the efficiency measured in data, although the scale factor remains within 4% of unity.

Breakdowns

The contributions to the overall trigger efficiency separated by lepton multiplicity are shown in Figure 33 by successively including each category of triggers to the menu. The effect of including triple lepton triggers on top of single and double lepton triggers amounts to a 0.5% acceptance gain for $m_{4\ell} < 253$ GeV. The overall trigger efficiency of the finalized menu is shown in Figure 34, also highlighting the additional impact of requiring trigger matching to quadruplet leptons.

The MC trigger efficiency as evaluated by the global trigger tool is further validated in $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$ MC by comparing the *post-trigger* yield observed after applying a manual

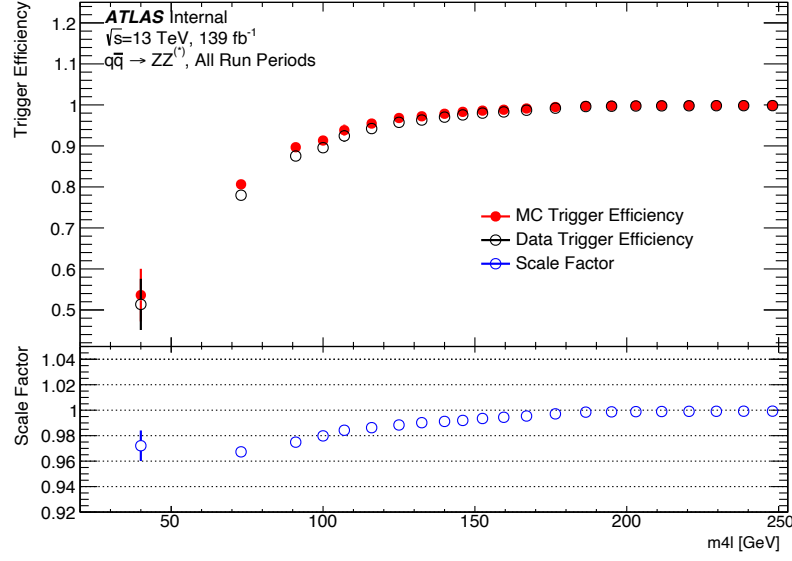


Figure 32: The MC and data trigger efficiencies, and the resulting efficiency scale factor as calculated by the TrigGlobalEfficiencyCorrectionTool using 100,000 pseudo-experiments.

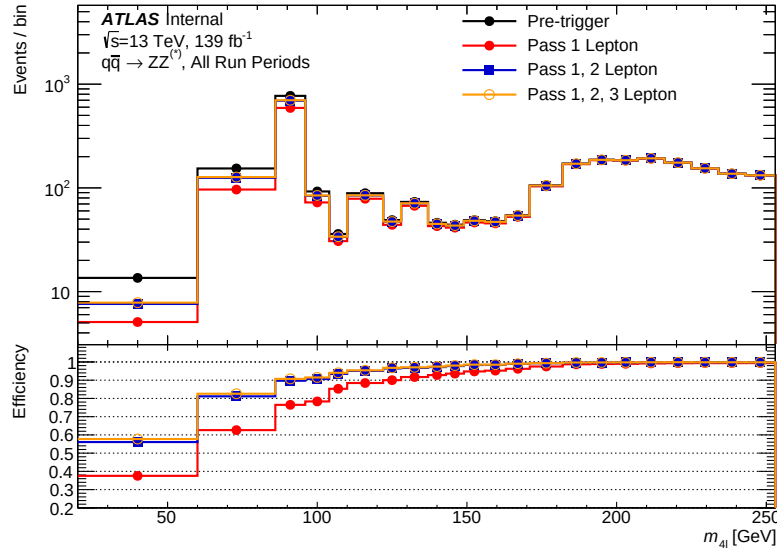


Figure 33: Trigger efficiency breakdown including successive sets of triggers by lepton multiplicity, corresponding to the categories in Table 12. The inclusion of double lepton triggers amounts to a 6.8% acceptance gain for $m_{4\ell} < 253$ GeV, while triple lepton triggers provide an additional 0.5% increase.

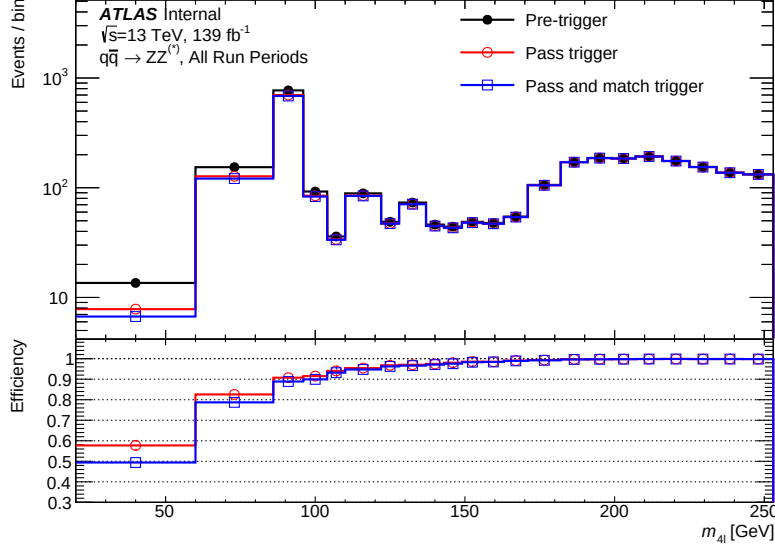


Figure 34: The overall trigger efficiency of the finalized trigger menu, also showing the impact of the additional requirement of trigger matching.

trigger selection against the *pre-trigger* reconstruction-yield weighted by the MC trigger efficiency ϵ_{MC} , as shown in Figure 35. Closure of these two yields indicates that the MC trigger efficiency as evaluated using pseudo-experiments accurately describes the actual trigger efficiency as measured in MC. The results of this closure test show a generally good agreement within 10% percent. The discrepancies at low four-lepton invariant mass arise due to differing isolation requirements between the analysis signal region and the individual trigger efficiencies used as inputs to the tool. Specifically, the isolation working point as used in the analysis is looser for low- p_T leptons than the working point which the individual trigger efficiencies are measured against. Therefore, the trigger efficiency as derived by the tool assuming more-isolated leptons is overestimated with respect to the less-isolated leptons allowed in the analysis signal region. However, given this bias affects the data and MC trigger efficiencies equally, it is cancelled out in the resulting scale factor.

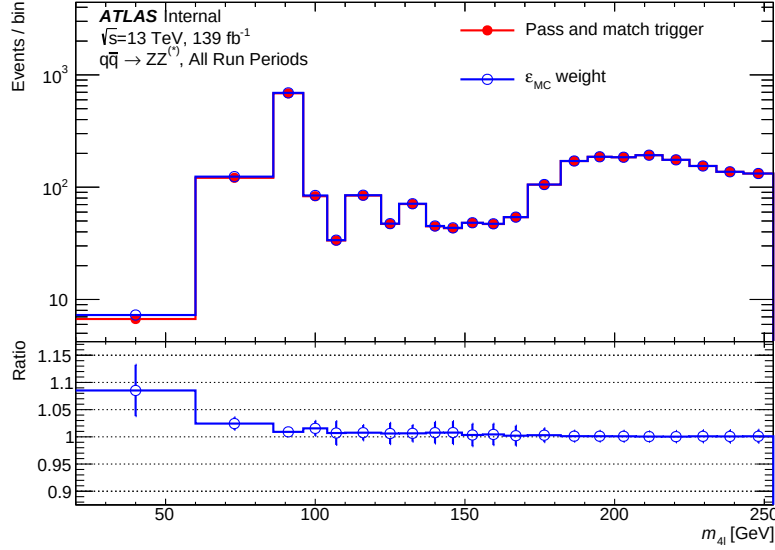


Figure 35: Validation of the TrigGlobalEfficiencyCorrectionTool comparing the *post-trigger* yield observed after applying a manual trigger selection to the reconstruction-yield weighted by the MC trigger efficiency ϵ_{MC} . The generally good closure indicates the tool provides a reliable measure of the MC and, by extension, data trigger efficiencies through the pseudo-experiment based evaluation. The overestimation of the ϵ_{MC} weighting at low-mass is driven by differing isolation working points between the signal region and the measured trigger efficiencies. This bias affects data and MC efficiencies equally and is therefore cancelled in the resulting scale factor.

16 Backgrounds

In the previous iteration of this analysis [44], the signal region was restricted to diboson-mediated four-lepton production, while all other processes were considered as *irreducible* or *reducible* sources of background. The contributions of other genuine four-lepton processes, namely vector boson production in association with a top quark pair ($t\bar{t}V$) and triboson production (VVV), were considered as irreducible sources of background and subtracted from the observed data using the respective predictions. As part of the effort to reduce model dependence in this analysis, it is preferable not to rely on the modeling of these processes in simulations. Therefore, these sources are no longer considered as backgrounds and, instead, included in the inclusive four-lepton signal definition. In other words, all events with four prompt leptons are considered as signal, leaving only reducible sources to consider as background. These backgrounds correspond to events in which at least one non-prompt lepton enters the selected quadruplet. In the terminology used here, *real leptons* are genuine leptons produced directly in the hard scatter, while *fake leptons* refer to all non-prompt objects - including genuine leptons produced from secondary interactions and other charged tracks incorrectly identified as leptons (*misidentified leptons*). This is depicted for a simplified hadronic jet event in Figure 36, where the fake lepton in this example may be a genuine or misidentified lepton produced in secondary decays within the jet activity.

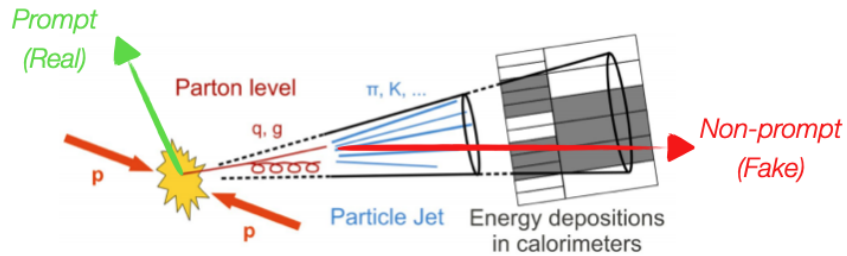


Figure 36: Depiction of prompt (real) and non-prompt (fake) leptons in an event with hadronic jet activity. In this example, the fake lepton may be a genuine or misidentified lepton produced in the secondary interactions. Adapted from Eva Hansen’s talk *Jet Reconstruction and Calibration with ATLAS*.

16.1 Sources

Although the modeling of fake leptons is generally worse than that of real leptons in MC predictions, the simulations provide a rough approximation of the expected background composition. For purposes of the lepton composition studies, the truth origin of leptons is determined by the Isolation and Fake Forum (IFF) truth classification tool¹⁷. The tool classifies leptons into several categories according to the available truth-level information. The classifications relevant to the phase space of this analysis are as follows:

- **Unknown/KnownUnknown**: insufficient truth information to classify particle
- **IsoElectron**: prompt electron from the hard scatter interaction
- **PromptMuon**: prompt muon from the hard scatter interaction
- **TauDecay**: leptons originating from the decay of a prompt tau lepton
- **PromptPhotonConversion**: electrons originating from a prompt photon conversion
- **ChargeFlipIsoElectron**: electrons whose reconstructed charge is mis-measured
- **BHadronDecay**: leptons originating from hadrons containing a b -quark
- **CHadronDecay**: leptons originating from hadrons containing a c -quark
- **LightFlavorDecay**: leptons originating from lighter quark mesons and hadrons

The IFF truth classifier origin of quadruplet leptons is shown for the signal and background samples which compose the four-lepton signal region in Figure 37. The vast majority of leptons are prompt electrons or muons, coming predominantly from the combined $qq \rightarrow 4\ell$ and $gg \rightarrow 4\ell$ signal samples ($llll$). Leptons from prompt tau decays, also considered as signal in this analysis, contribute a minor portion mostly from the $t\bar{t}V$ ($t\bar{t}ll$) process. The **Unknown/KnownUnknown** classifications, resulting from insufficient truth information, are composed mainly of the signal $t\bar{t}V$ and VVV processes. Based on ΔR -based truth-association studies, it is observed that these unclassified leptons appear to be real in the

¹⁷<https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier>

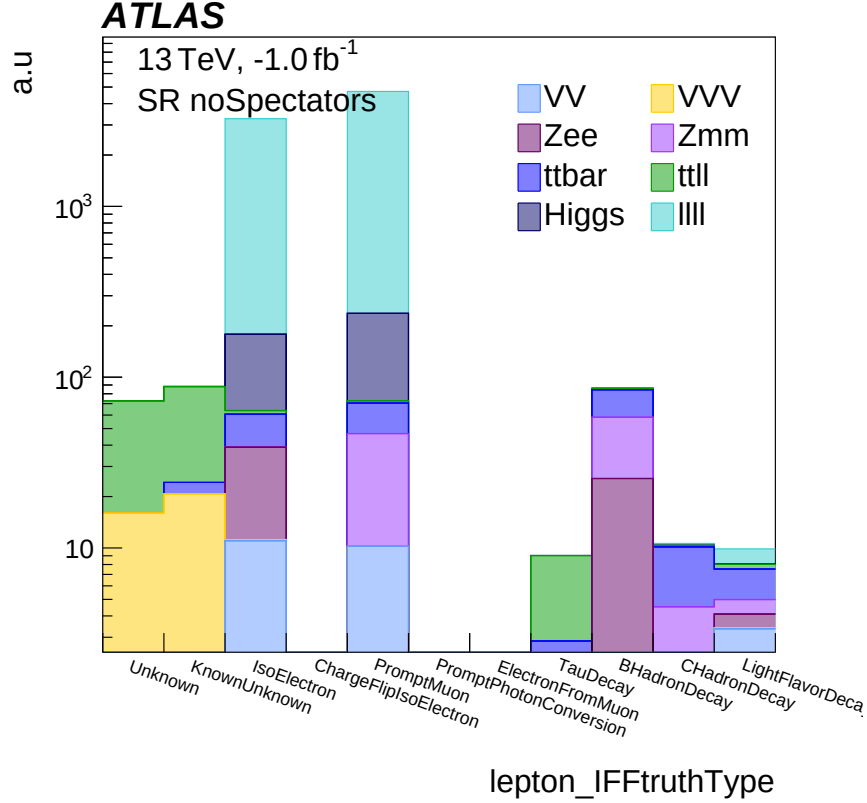


Figure 37: The truth origin of quadruplet leptons in the signal and background samples which enter the four-lepton signal region, weighted by the relevant cross-section. “Spectator” leptons which are not part of the selected quadruplet are excluded.

signal region and fake in other regions - and are therefore assumed as such for the estimation. All other classifications are considered backgrounds for the purposes of this measurement, and together contribute less than 10% of the total composition. The dominant background source is that of leptons from b -hadron decays, while lighter flavor hadron and meson decays comprise the second largest contribution. The former arises mainly from $Z \rightarrow \ell\ell$ ($Z + \text{jets}$) production, while the latter is composed roughly equally of $Z + \text{jets}$ and $t\bar{t}$ production. The contributions of electrons with mis-measured charge and those from prompt photon conversions, both of which are considered as background, are shown to be negligible.

Though the overall composition of the fake background is largely understood from simulations, a data-driven background estimation is preferred over MC predictions for two main reasons. Firstly, the MC simulations are limited in statistics with respect to the data and

therefore suffer from larger statistical uncertainties. Secondly, although simulations generally model genuine fakes (e.g. secondary leptons from b -hadron decays) well, the modeling of reconstruction artifacts (e.g. jets reconstructed as leptons) is much worse. Therefore, a data-driven approach based on the *fake factor* method is used for the background estimation.

16.2 Analysis Regions

Four-lepton Signal and Validation Regions

As outlined in Section 12, the signal selections form a subset of the more inclusive baseline selections. There are three classes of leptons relevant to the background estimation: *baseline* leptons, *signal* leptons, and *baseline-not-signal* leptons. Baseline and signal leptons correspond to leptons passing, respectively, the baseline and signal selections outlined in Section 12. Baseline-not-signal leptons are those which pass the baseline selections, but fail the additional signal requirements. The signal region (SR) of the analysis consists of baseline quadruplets in which all leptons of the quadruplet pass the additional signal selections. The larger region consisting of baseline-not-signal quadruplets, in which at least one quadruplet lepton does not pass signal requirements, forms the inclusive four-lepton control region (CR). Although the events outside of the SR do not enter the final measurement, the inclusive CR is essential for the data-driven estimation of the fake lepton background.

In addition to the primary four-lepton signal region, consisting of two same-flavor, opposite-charge lepton pairs, two additional validation regions are introduced to verify the accuracy of the background prediction. These regions are orthogonal to the primary SF-OC regions, based on the presence of one non SF-OC lepton pair in the quadruplet, but otherwise identical in terms of upstream selections. In the different-flavor validation region (VR-DF) exactly one pair of the quadruplet is required to be different-flavor, conversely, in the same-charge validation region (VR-SC) exactly one pair is required to be same-charge. In each of the validation regions, the subset of events which do not contain a signal quadruplet are again used in the respective background estimations. A summary of the signal and validation

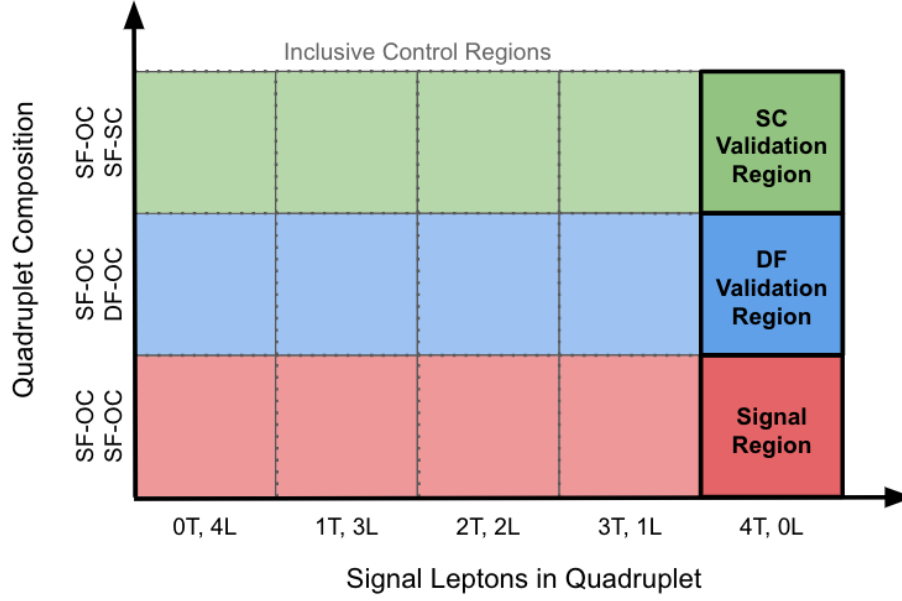


Figure 38: Graphical representation of the four-lepton regions used in the background estimation and final measurement. The horizontal axis shows the number of signal (tight) and, inversely, baseline-not-signal (loose) leptons in the quadruplet. The vertical axis denotes the composition of the two lepton pairs in the quadruplet. The signal region is shown in the bottom right, consisting of four signal leptons arranged in two same-flavor, opposite-charge pairs. The DF (SC) validation region, fully orthogonal to the signal region, consists of one same-flavor, opposite-charge lepton pair and one different-flavor, opposite-charge (same-flavor, same-charge) lepton pair. Along each of the three horizontal slices, all regions having less than four signal leptons, collectively referred to as the inclusive control regions, are used in the respective background estimations.

regions, plotted along the axes of the number of signal leptons in the quadruplet and the lepton pair composition, is shown in Figure 38.

Although the secondary validation regions are orthogonal to the SR, the composition of fake leptons remains largely unchanged, allowing for an independent validation of the background estimation. The compositions of non-prompt leptons in the VR-DF and VR-SC are shown alongside the non-prompt composition in the SR in Figure 39. The relative contribution of fake leptons from b -hadron decays to those from lighter quark decays is comparable across all three regions.

Table 13: Summary of the $Z \rightarrow \ell\ell$ CR reconstruction-level selection.

Category	Requirement
Event Preselection	Fire ≥ 1 lepton trigger
Final State	≥ 2 signal leptons, plus ≥ 1 baseline lepton
J/ψ -Veto	$m_{ij} > 5$ GeV for all SF-OC lepton pairs
Lepton Separation	$\Delta R_{ij} > 0.05$ for any two leptons
Dilepton Formation	≥ 1 SF-OC lepton pair
Z -mass Constraint	$76 < m_{\ell\ell} < 106$ GeV
Pair Kinematics	$p_T > 20$ (15) GeV for leading (subleading) lepton
Trigger Matching	Pair leptons matched to ≥ 1 lepton trigger

$Z \rightarrow \ell\ell$ Control Region

Given the known mis-modeling of fake leptons in simulations, as well as the limited MC statistics, a fake-enriched $Z \rightarrow \ell\ell$ control region is defined in data to more accurately study the behavior of fake leptons. The region is defined based on the selections outlined in Table 13 so as to be enriched in the dominant $Z + \text{jets}$ background. The region uses a combination of single and double lepton triggers, and requires a minimum of two signal leptons plus at least one additional baseline lepton. Of the signal leptons in the event, a Z -compatible pair must be present consisting of two same-flavor, opposite-charge leptons with an invariant mass between $76 < m_{\ell\ell} < 106$ GeV. As in the four-lepton regions, a J/ψ veto of $m_{\ell\ell} > 5$ GeV is enforced for all same-flavor, opposite-charge lepton pairs, and all leptons must be separated by a minimum $\Delta R > 0.05$. The leading and subleading leptons of the pair must respectively satisfy $p_{T,1} > 20$ GeV and $p_{T,2} > 15$ GeV. Based on this selection, the primary lepton pair consists largely of genuine prompt leptons from Z -decays, while any additional baseline leptons in the event are expected to be fake. As shown in Figure 40, the composition of fake leptons in the $Z \rightarrow \ell\ell$ control region is comparable to the signal region and, by extension, the validation regions. Therefore, measurements of fake leptons in this region can be directly applied to the four-lepton signal and validation regions.

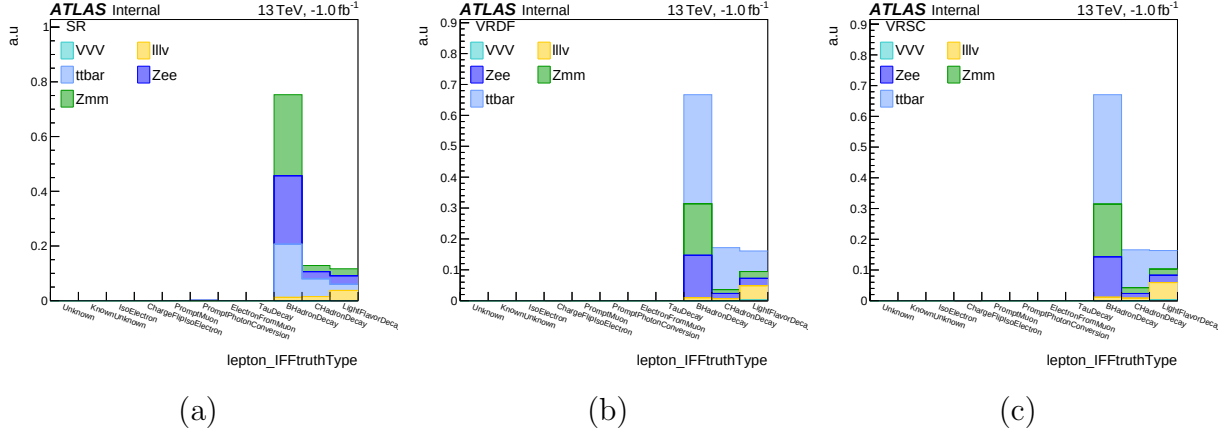


Figure 39: The truth origin of non-prompt leptons in the four-lepton signal region (a), the different-flavor validation region (b), and the same-charge validation region (c).

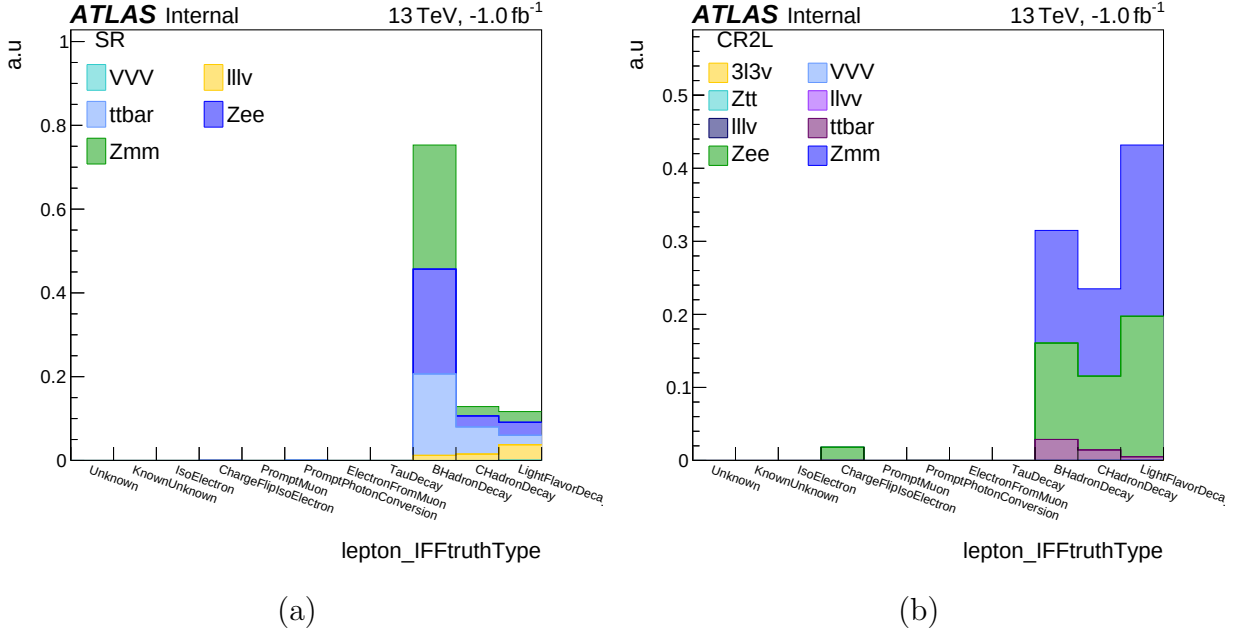


Figure 40: The truth origin of non-prompt leptons in the four-lepton signal region (a) and the $Z \rightarrow \ell\ell$ control region (b).

16.3 Estimation Strategy

The *fake factor* method is a data-driven background estimation technique, closely related to the *matrix method*, both of which are discussed in detail in Ref. [60]. Although both methods were tested for use in this analysis, the former proved more effective for reasons to be discussed. Central to both methods are the so-called real (r) and fake (f) efficiencies, which are typically measured in data control regions enriched in real and fake leptons, respectively. The *real efficiency* is a measure of the fraction of prompt leptons passing baseline selections (N_R^B) which further pass the tighter signal selections (N_R^T),

$$r = \frac{N_R^T}{N_R^B}. \quad (49)$$

Similarly, the *fake efficiency* measures the fraction of non-prompt leptons passing baseline selections (N_F^B) which further pass the tighter signal selections (N_F^T),

$$f = \frac{N_F^T}{N_F^B}. \quad (50)$$

Both techniques operate on the assumption that the fake efficiency is much smaller than the real efficiency ($f \ll r$) and, therefore, non-linear terms in f can be ignored.

For the simplest case of a signal region defined by a single tight lepton, the observed signal (N^T) and baseline-not-signal (N^L) lepton yields can be decomposed in terms of the real (N_R^B) and fake (N_F^B) baseline predictions as

$$N^T = rN_R^B + fN_F^B, \quad (51)$$

and

$$N^L = (1 - r)N_R^B + (1 - f)N_F^B. \quad (52)$$

This system of equations can be arranged in a 2×2 matrix equation,

$$\begin{pmatrix} N^T \\ N^L \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_R^B \\ N_F^B \end{pmatrix}. \quad (53)$$

The matrix method predicts the fake baseline contribution by inversion of the matrix,

$$N_F^B = \frac{1}{r-f} (r(N^T + N^L) - N^T), \quad (54)$$

where higher order corrections dependent on f^2 are ignored. The predicted number of fake leptons passing signal selections is therefore given by

$$N_F^T = f N_F^B = \frac{f}{r-f} (r(N^T + N^L) - N^T). \quad (55)$$

The fake factor method is a simplification of the matrix method technique, essentially equivalent to the limit of $r \rightarrow 1$ in Equation 55. However, since in reality the real efficiency of any analysis is less than unity this will overestimate the fake background composition. Therefore, the genuine baseline-not-signal prompt contribution as measured in MC (N_R^L) must be subtracted from the observed baseline-not-signal yield,

$$N_F^T = \frac{f}{1-f} (N^L - N_R^L). \quad (56)$$

The calculation of N_R^L assumes real leptons are well-modeled in MC, which is typically a safe assumption. The leading coefficient in Equation 56 is called the *fake factor*, denoted by F ,

$$F = \frac{f}{1-f}, \quad (57)$$

which has a lower bound of zero when $f \rightarrow 0$ and is unbounded above when $f \rightarrow 1$, with larger values corresponding to higher fake efficiencies. Unlike the matrix method estimate in

Equation 55, the fake factor estimate in Equation 56 does not depend on the signal region yield in which all leptons are tight. Therefore, the fake factor technique offers the important advantage of deriving a background estimate which is blind to the contents of the signal region [60].

Each method can be extended to an arbitrarily complex multi-lepton final state, defined by the presence of N^B baseline leptons. In a multi-lepton final state, there are 2^{N^B} observations to consider, one for every permutation of the individual leptons being either signal or baseline-not-signal. Hence, a two lepton final state requires measuring N^{TT} , N^{TL} , N^{LT} , and N^{LL} , which are related to the predicted real and fake contributions by a 4×4 matrix equation,

$$\begin{pmatrix} N^{TT} \\ N^{TL} \\ N^{LT} \\ N^{LL} \end{pmatrix} = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1 \bar{r}_2 & r_1 \bar{f}_2 & f_1 \bar{r}_2 & f_1 \bar{f}_2 \\ \bar{r}_1 r_2 & \bar{r}_1 f_2 & \bar{f}_1 r_2 & \bar{f}_1 f_2 \\ \bar{r}_1 \bar{r}_2 & \bar{r}_1 \bar{f}_2 & \bar{f}_1 \bar{r}_2 & \bar{f}_1 \bar{f}_2 \end{pmatrix} \begin{pmatrix} N_{RR}^B \\ N_{RF}^B \\ N_{FR}^B \\ N_{FF}^B \end{pmatrix}, \quad (58)$$

where $\bar{\varepsilon} = (1 - \varepsilon)$.

For the more relevant case of a four-lepton signal region, there are $2^4 = 16$ equations to consider, corresponding to the individual observations of the N^{TTTT} , N^{TTTL} , ..., N^{LLLL} yields. In this analysis, any event with at least one baseline-not-signal lepton in the quadruplet is considered as background, equivalent to $N_{\mathcal{F}}^{TTTT}$ where $\mathcal{F}$ represents all quadruplet permutations of real and fake leptons excluding the four-real combination ($RRRR$). The estimation of the background therefore consists in the prediction of events containing four signal leptons in which at least one lepton is fake. However, owing to a slight complication arising from the loose ΔR selections of this analysis, a modification to the calculation must be made to correct for cases in which the real and fake probabilities of nearby leptons are correlated. This lepton skipping work-around is discussed further in Appendix D. In the final matrix method or fake factor method calculation, an event-level weight is evaluated based on the non-skipped quadruplet leptons in the inclusive CR to obtain an estimate of

the background composition in the SR. The limited data statistics in the inclusive region affects the precision of the background estimate and is therefore considered as a source of systematic uncertainty in the final measurement (see Section 18).

16.4 Calculation of Efficiencies

Real Efficiency

The real efficiency is required as input only for the matrix method estimation, entering in Equation 55. Owing to the excellent modeling of prompt lepton reconstruction, it can be directly measured in simulation using the IFF truth classification. The real efficiency is calculated using the dominant signal sample $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$ and parameterized as a function of lepton p_T and the number of jets, N_{jets} . Only leptons which form a quadruplet in the inclusive four-lepton region are considered in the calculation. The choice of jet multiplicity over the more commonly used lepton η variable owes to an observed greater dependency on the former variable over the latter. The measured real efficiencies are shown separately for electrons and muons in Figure 41. In both cases, the efficiency significantly increases with lepton p_T and slightly decreases with the jet multiplicity. The real efficiency is on average close to unity, ranging from $69 \rightarrow 100\%$, and well above 90% in the high populated regions.

Fake Efficiency

The fake efficiency is a necessary input to both the matrix method and fake factor method, entering in Equations 55 and 56. Unlike the real efficiency, it cannot be accurately measured in simulation for reasons previously discussed and is instead measured in the fake-enriched $Z \rightarrow \ell\ell$ control region. In this region, any baseline leptons present in the event in addition to the selected lepton pair are assumed to be fake. Therefore, the fake efficiency can be directly profiled as a function of the additional baseline lepton kinematics. Based on extensive studies, an accurate estimation is obtained by measuring the efficiency double-differentially, in bins of the additional lepton p_T and the number of jets present in the event. As shown in Figure 42,

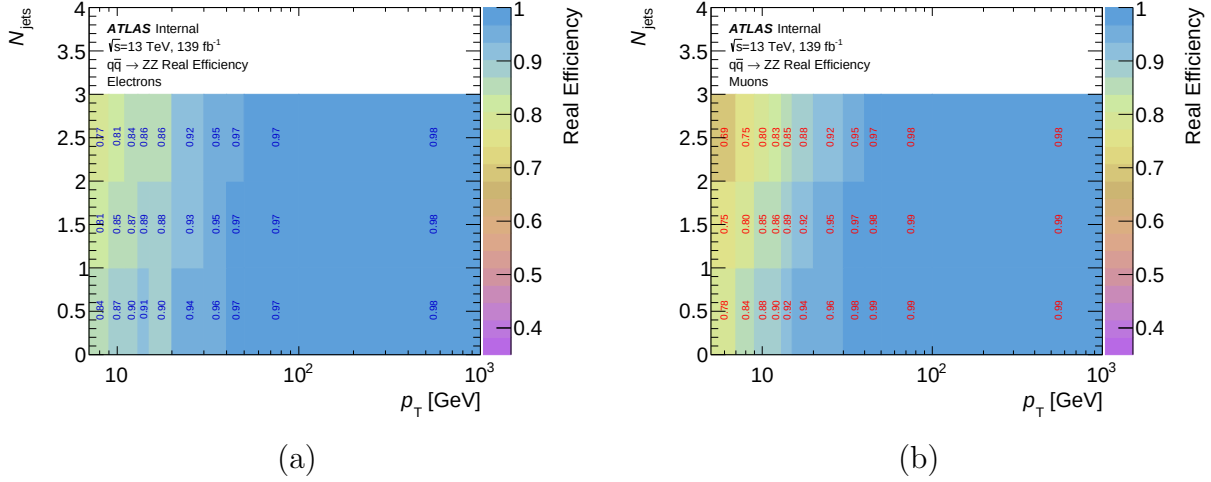


Figure 41: The real efficiency, defined as the fraction of prompt baseline leptons passing the signal requirements, measured in $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$ MC separately for electrons (a) and muons (b) in the inclusive four-lepton region.

the fake efficiency as measured for electrons increases with lepton p_T , and is mostly stable versus the number of jets. Shown in Figure 43, the fake efficiency as measured for muons decreases with both lepton p_T and the jet multiplicity. The opposite dependence on lepton p_T for electrons and muons is understood to be due to differences in the isolation working point, while the dependence on the number of jets, particularly for muons, is less understood. The two-dimensional histograms shown here are used to evaluate and apply the fake factor weight (Equation 57) in the inclusive control region when estimating the reducible background in the signal region. Statistical and systematic uncertainties are evaluated on the measured fake efficiencies, also shown in Figures 42 and 43, where the latter component accounts for QCD renormalization and factorization scale, PDF, and α_s variations in the WZ sample used for the prompt subtraction. These uncertainties are propagated to the final measurement by repeating the fake factor estimation using the symmetric up and down variations on the fake efficiencies, the impacts of which are discussed in more detail in Section 18.

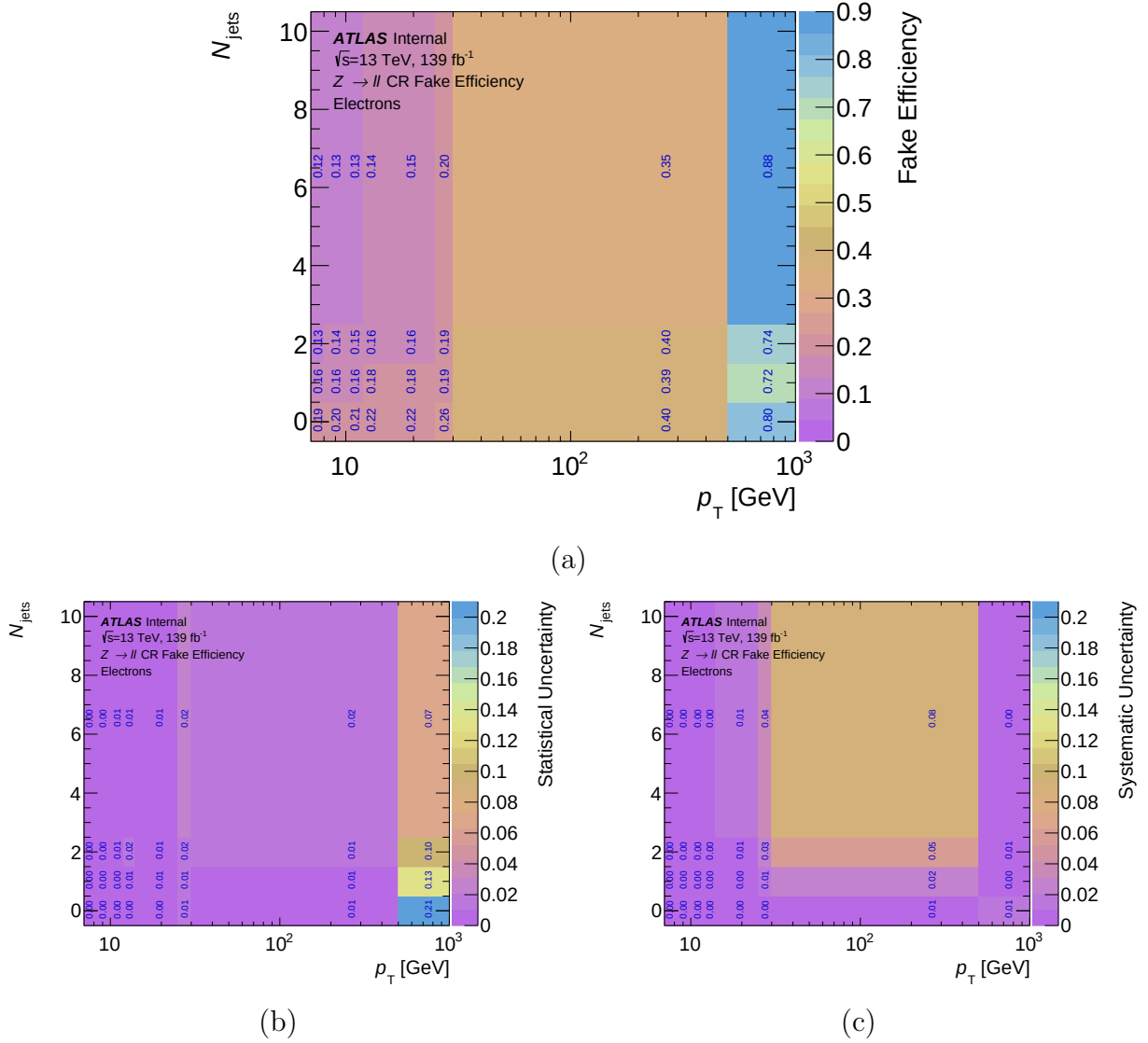


Figure 42: The electron fake efficiency (a), defined as the fraction of non-prompt baseline electrons passing the signal requirements, measured in data in the $Z \rightarrow \ell\ell$ control region, and corrected for genuine prompt contributions from MC. Statistical (b) and systematic (c) uncertainties on the fake efficiencies are also shown, where the latter accounts for renormalization and factorization scale, PDF, and α_S variations in the subtracted WZ sample.

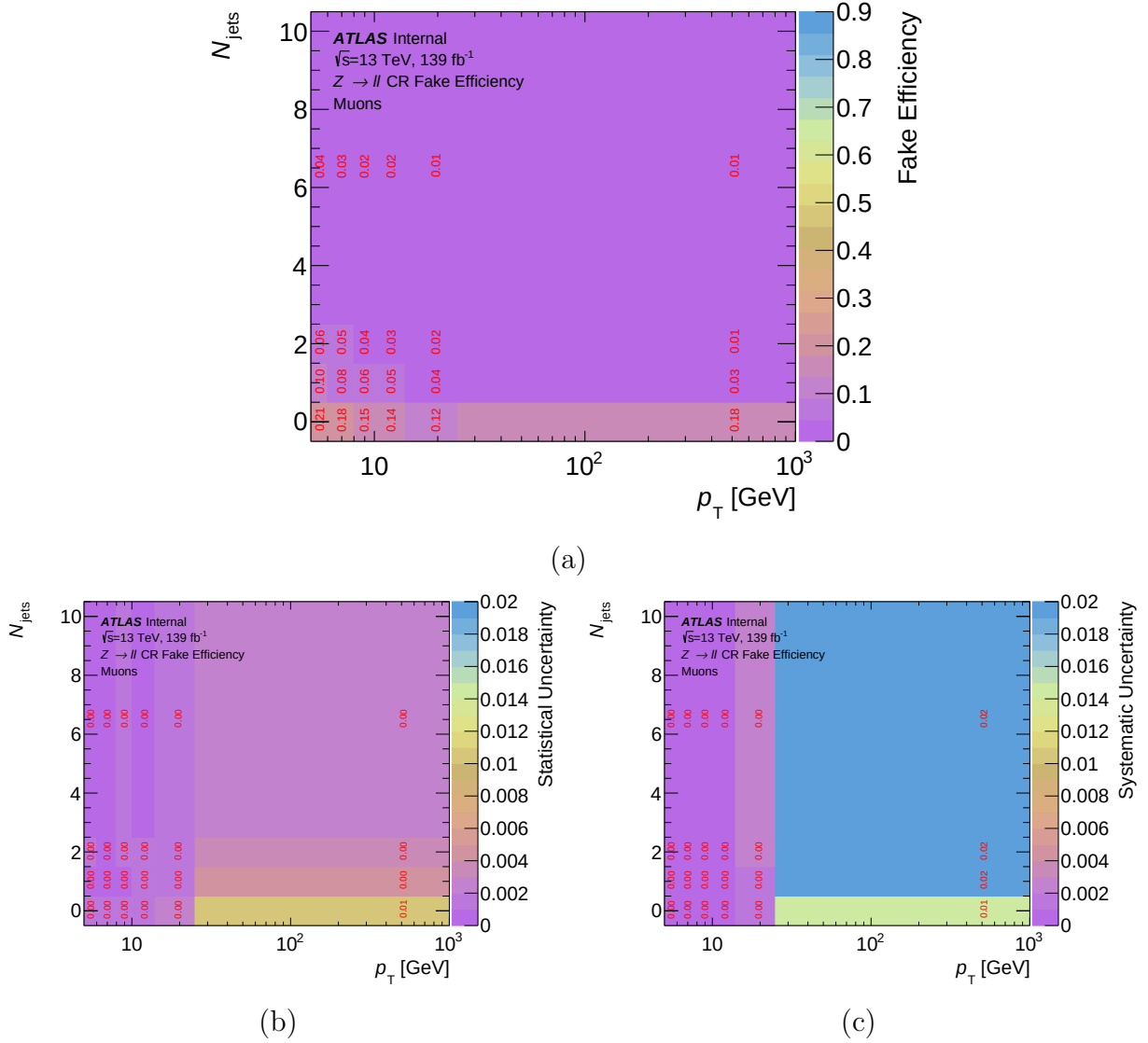


Figure 43: The muon fake efficiency (a), defined as the fraction of non-prompt baseline muons passing the signal requirements, measured in data in the $Z \rightarrow \ell\ell$ control region, and corrected for genuine prompt contributions from MC. Statistical (b) and systematic (c) uncertainties on the fake efficiencies are also shown, where the latter accounts for renormalization and factorization scale, PDF, and α_S variations in the subtracted WZ sample.

16.5 Choice of Estimation Strategy

The primary advantage of the fake factor method over the matrix method is the lack of dependence on the signal region, allowing for a truly signal-blind background estimation. As a result of this simplification, though, the fake factor method requires the genuine baseline-not-signal prompt component to be taken from MC. In practice, however, this is hardly a disadvantage since prompt leptons are well-modeled by MC and, moreover, the real efficiency inputs for the matrix method are also derived from prompt MC. Studies of the two methods found the matrix method generally performs worse than the fake factor method, particularly in regions with a very high signal-to-background ratio. Given that the background contamination in the signal region is relatively low, typically below 10%, the matrix method proved a poor strategy for this analysis. Therefore, the fake factor method was chosen as the nominal background estimation technique.

To ensure the fake factor estimate is not significantly different from that of the matrix method, a validation of the two estimates is performed based on a method asymmetry metric A . This metric is defined as the ratio of the difference to the sum of the two estimates,

$$A = \frac{N^{FF} - N^{MM}}{N^{FF} + N^{MM}}, \quad (59)$$

so as to be zero when the methods are approximately equal and non-zero otherwise. This metric is evaluated with a subset of the leading systematic uncertainties in the validation regions, the results of which are shown versus $m_{4\ell}$ in Figure 44. The uncertainties are applied in a correlated way on the two estimates and are shown separately for the fake efficiency uncertainties in red and the combined uncertainties also including additional theoretical sources in gray. Even excluding the larger inclusive CR statistical uncertainty, the two methods are in good agreement within systematic uncertainties.

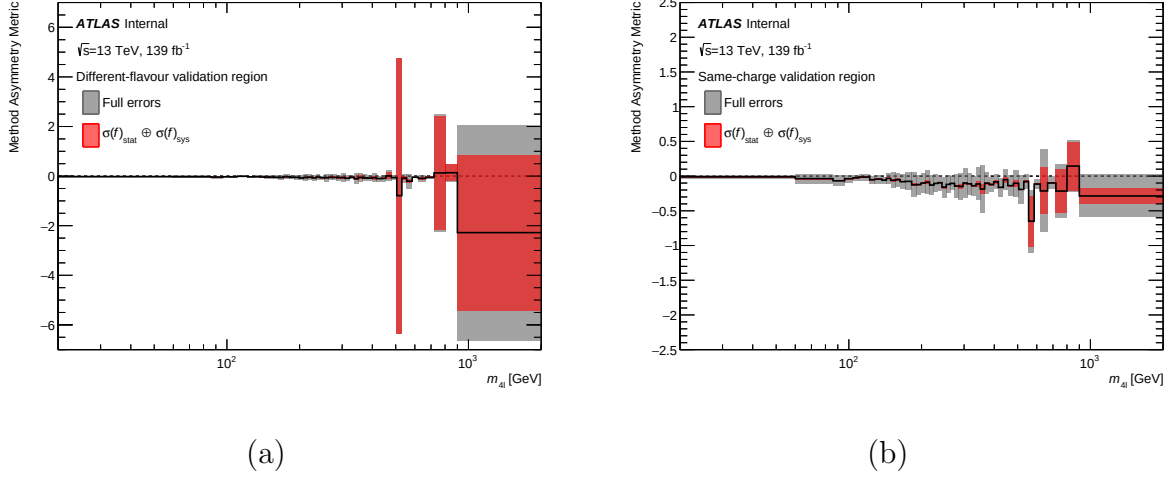


Figure 44: Method asymmetry comparison between the fake factor and matrix methods, as defined in Equation 59, versus $m_{4\ell}$ in both the different-flavour validation region (a) and the same-charge validation region (b).

16.6 Smoothing of Prediction

The background prediction, as evaluated using the fake factor or matrix methods, exhibits bin-to-bin fluctuations owing to statistical fluctuations in the measured efficiencies. This noise is non-physical in nature, instead the result of limited data and MC statistics. To prevent these statistical fluctuations from being propagated to the background-subtracted data, a smoothing procedure is used to reduce noise without losing genuine shape information in the background estimate.

Studies using various smoothing techniques were found to over-smooth the background predictions when applied directly to the estimated differential yields binned according to the unfolded measurements. Due to the relative coarseness of the unfolded binning (discussed in Section 17.3), the smoothing techniques could not accurately resolve the underlying background shape. Therefore, an alternative approach was devised based on the smoothing of observables in an intermediary fine binning. This multistage smoothing technique is shown for the single-differential $m_{4\ell}$ observable in Figure 45 and is carried out per-slice for double-differential observables. In the first step, the background estimate is evaluated in an intermediary fine binning, consisting of 500 fixed-sized bins spaced either linearly or log-

linearly depending on the relative difference of the smallest and largest bins in the observable slice¹⁸. In the second step, Friedman’s Super Smoother technique [61] is used to smooth the finely-binned estimate. In the third step, the finely-binned smoothed estimate is integrated to the nominal binning of the observable. In the fourth step, intended as an internal validation, this final smoothed result is compared to the unsmoothed estimate as derived in the nominal binning. The results for single-differential $m_{4\ell}$ show the background smoothing accurately captures the shape of the prediction, while reducing statistical fluctuations in the raw estimate.

To account for uncertainties arising from the smoothing procedure itself, a source of systematic uncertainty is considered in the final measurement. This uncertainty is derived by following the above smoothing procedure using 250 or 1000 intermediary bins instead and taking the envelope of the predictions from the nominal smoothed result as the error (see Section 18).

16.7 Validation of Estimate

The background estimation is validated in two orthogonal validation regions: VR-DF and VR-SC, as depicted in Figure 38. Although limited in statistics, the regions provide independent validations of the fake factor background estimation strategy as used in this analysis. To quantify the agreement with data, systematic uncertainties are evaluated on the combined yield of the signal MC prediction and reducible background estimate. As done in the SR, systematic variations are applied in a correlated way to the MC samples used for the signal prediction and background estimation. Only a subset of the leading reconstruction-level systematics are considered, while the full set of systematics evaluated in the SR is discussed in Section 18. This includes all background-related systematics - namely the sta-

¹⁸ The choice of log-linear or linear spaced binning depends on the nominal binning of observables. Positive-valued observables which exceed 1,000 in the upper limit and whose widest and smallest bins differ by at least a factor of ten are binned log-linearly. The remaining observables are binned linearly.

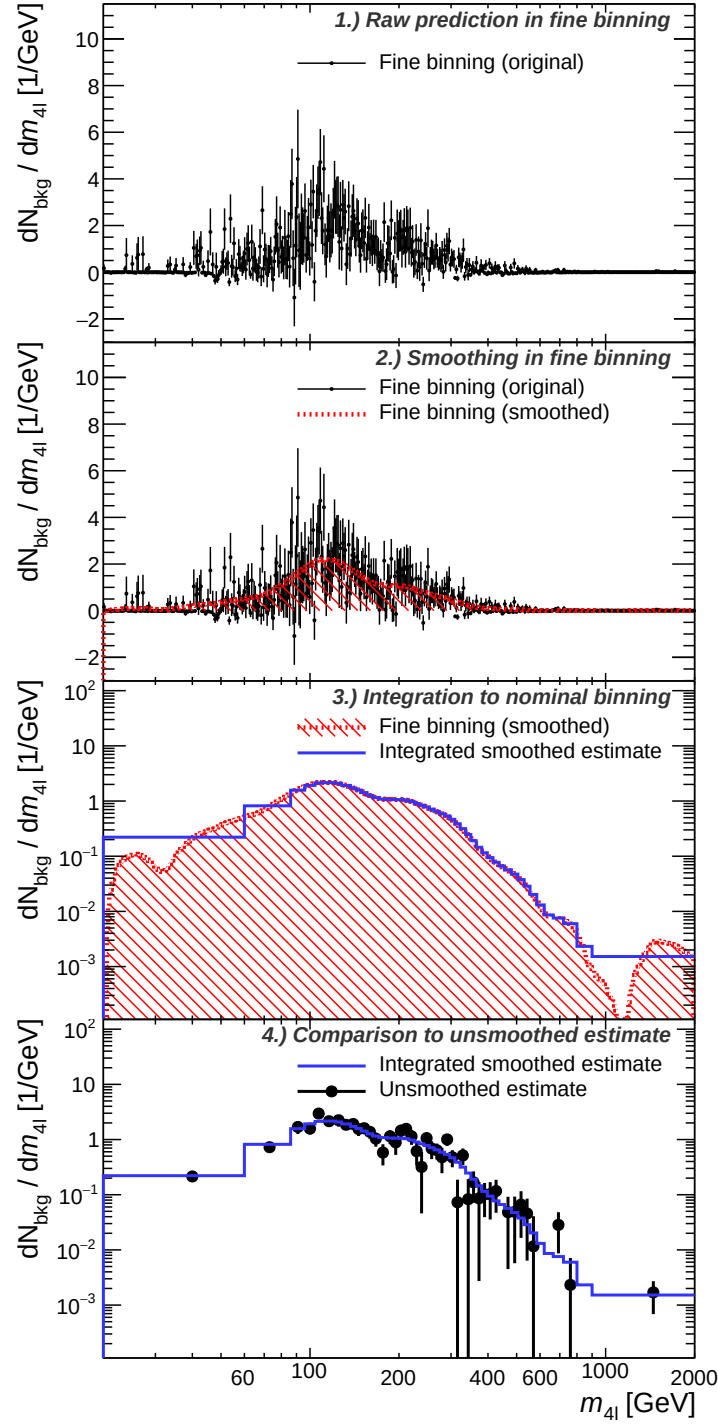
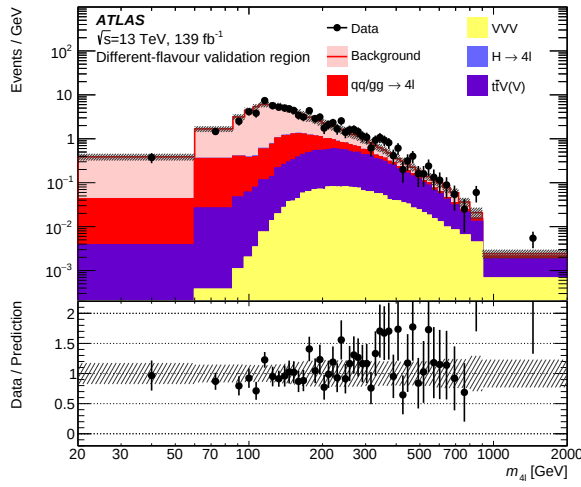


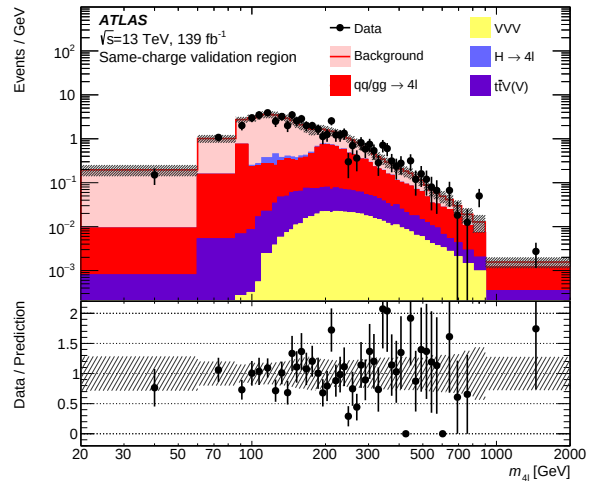
Figure 45: A step-by-step overview of the background smoothing technique for the single-differential $m_{4\ell}$ observable. In (1), the background estimate is evaluated in an intermediary fine binning. In (2), the smoothing routine is run to smooth the finely-binned estimate. In (3), the smoothed result is integrated to the nominal binning. In (4), the integrated result is compared to the unsmoothed estimate in the nominal binning.

tistical and systematic uncertainties on the fake efficiencies, the envelope of the alternative smoothing configurations, and the data statistical uncertainty in the inclusive CR. Flat cross-section variations of $\pm 15\%$ and $\pm 10\%$ on, respectively, the $t\bar{t}V$ and VVV predictions are also included. Lastly, a theoretical uncertainty accounting for QCD renormalization and factorization scale variations in the $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$ sample is also included. Due to framework constraints, this uncertainty is evaluated in the SR and propagated to the VRs by scaling to the expected $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$ admixture in each region.

The observed data is compared against the combined signal plus background prediction versus $m_{4\ell}$ for the two validation regions in Figure 46. The predictions are generally in good agreement with data, falling well within the bounds of the total uncertainty. Plots for additional quadruplet- and event-level observables are shown for the VR-DF in Figure 47 and the VR-SC in Figure 48. For all observables studied across both regions, the predictions are found to be in good agreement with data within errors. As shown by the $\Delta R_{\ell, \min}$ comparisons in Figure 47 (f) and Figure 48 (f), the closure is observed even for minimally-separated leptons, indicating the nearby lepton skipping strategy discussed in Appendix D is valid.



(a)



(b)

Figure 46: Validation of the reducible background estimation in the different-flavor validation region (a) and same-charge validation region (b) for the single-differential $m_{4\ell}$ observable. Black error bars denote the statistical uncertainty on the data, while the gray shading shows the total systematic uncertainty on the combined prediction.

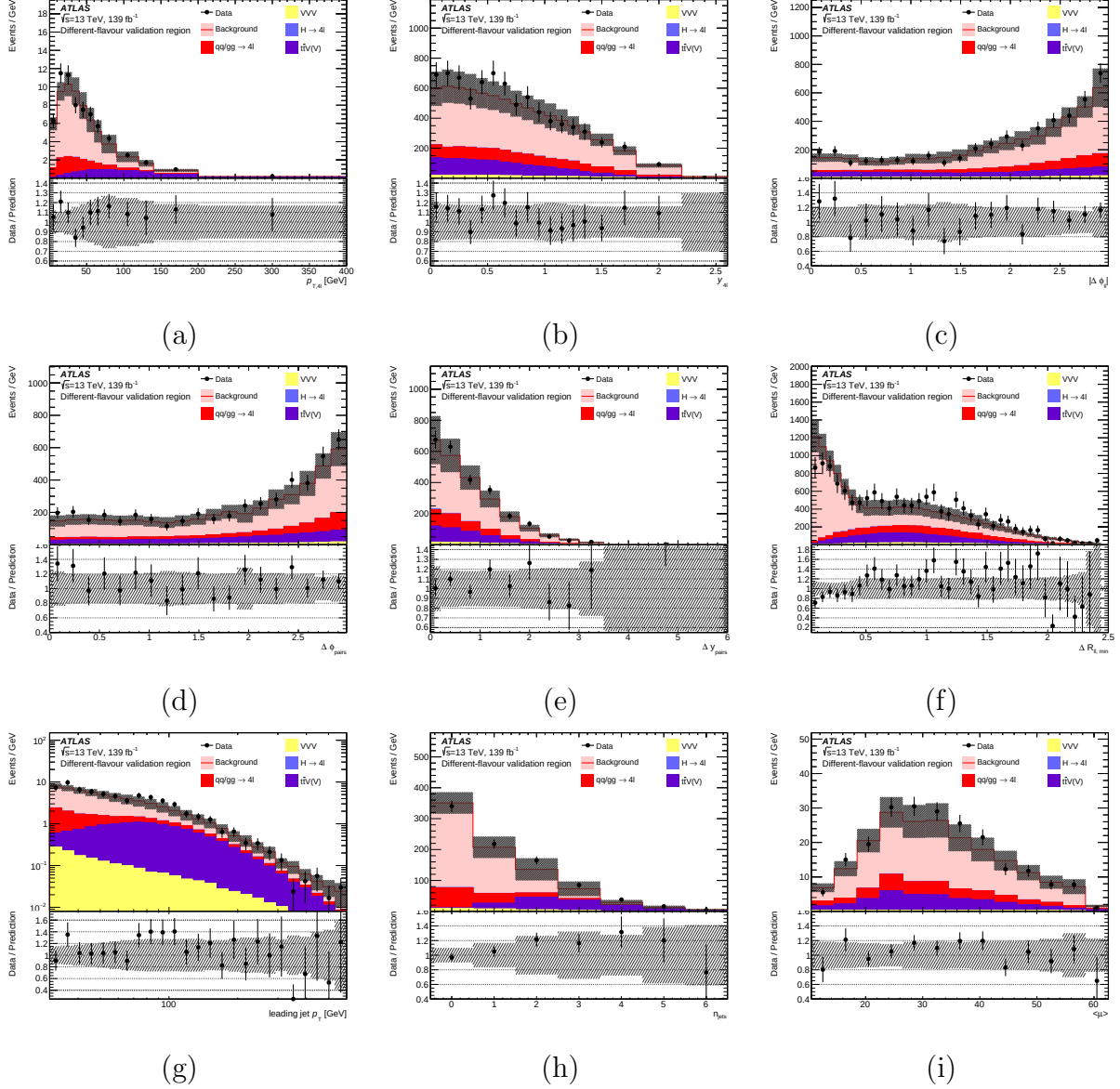


Figure 47: Validation of the reducible background estimation in the different-flavour validation region for a variety of observables. Black error bars denote the statistical uncertainty on the data, while the gray shading shows the total systematic uncertainty on the combined prediction.

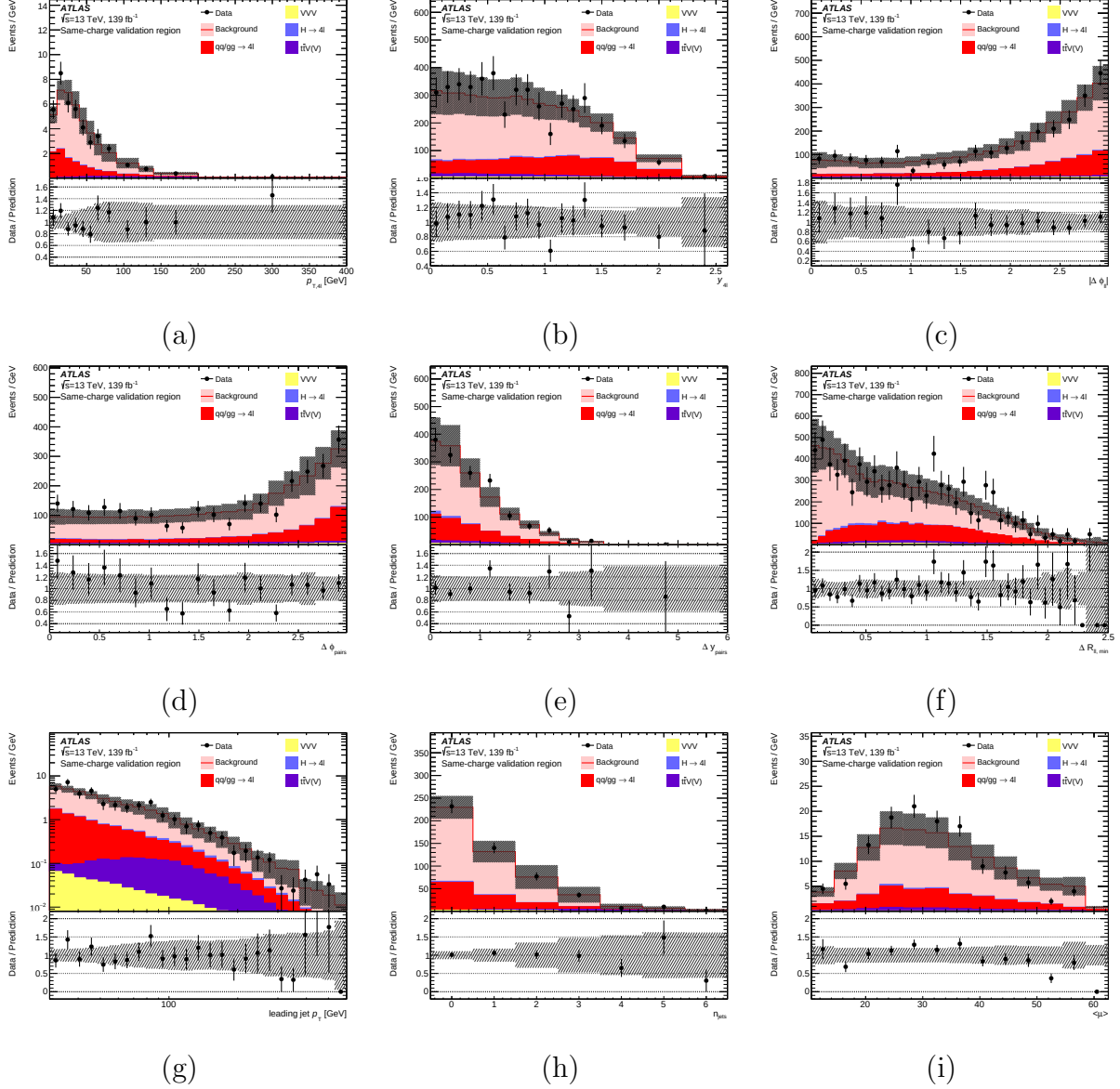


Figure 48: Validation of the reducible background estimation in the same-charge validation region for a variety of observables. Black error bars denote the statistical uncertainty on the data, while the gray shading shows the total systematic uncertainty on the combined prediction.

17 Unfolding

One of the foundational goals of experimental particle physics is the measuring of particle cross-sections of specific processes. The number of observed events for a given process $ab \rightarrow X$ as measured in a collider experiment is essentially given by

$$N_{ab \rightarrow X}^{\text{meas}} = A \times \varepsilon \times \sigma_{ab \rightarrow X}^{\text{true}} \times \int \mathcal{L}(t) dt, \quad (60)$$

where $\sigma_{ab \rightarrow X}^{\text{true}}$ corresponds to the underlying *true* cross-section of the process and $\mathcal{L}(t)$ is the instantaneous luminosity delivered by the collider. The factors of A and ε account, respectively, for the limited acceptance and reconstruction efficiency of the detector. Dividing out the integrated luminosity gives the measured inclusive cross-section, which is trivially related to the true cross-section by

$$\sigma_{ab \rightarrow X}^{\text{meas}} = A \times \varepsilon \times \sigma_{ab \rightarrow X}^{\text{true}}. \quad (61)$$

Thus, by measuring $\sigma_{ab \rightarrow X}^{\text{meas}}$, in addition to the acceptance and reconstruction efficiency of the detector, one can estimate the true value of the cross-section. However, this relatively simple picture is much more complicated in the case of differential measurements, where resolution effects can result in migrations between nearby bins. The process of unfolding offers a full treatment of these experimental effects for the purposes of estimating true cross-sections.

17.1 Overview

Given the nature of experimentation, measured data as reconstructed by the ATLAS detector necessarily suffers from inefficiencies and resolution effects. Inefficiencies particularly affect low-momentum objects and vary depending on the pseudorapidity, corresponding to changes in the active detector region. The effect of these inefficiencies generally decreases the measured yield in certain regions of phase space. Resolution effects play a more complicated

role because the smearing of measurements can result in migrations between neighboring bins of a measured distribution. It is illuminating to think of these experimental effects as transforming the underlying particle-level (true) distributions into a set of detector-level (reconstructed) distributions.

Unfolding corrects for the inefficiencies and resolution effects in data by transforming the measurements from a detector-level phase space back to a particle-level phase space - essentially inverting the reconstruction process. The specific unfolding technique used here relies on first deriving a series of *unfolding inputs*, which encode the expected response of the detector to the signal processes of interest. The necessary unfolding inputs, listed below, are derived using the signal MC samples outlined in Section 13.2.

- ***Reconstruction efficiency:*** The reconstruction efficiency, ε , is a measurement of the fraction of events in MC which pass both reconstruction- and fiducial-level selections to all those passing only the fiducial selection, as a function of the particle-level variable. This accounts for the limited acceptance and efficiency of the detector, which result in a fraction of particle-level events being captured and reconstructed at detector-level.
- ***Fiducial fraction:*** The fiducial fraction, f , is a measurement of the fraction of events in MC which pass both reconstruction- and fiducial-level selections to all those passing only the reconstruction selection, as a function of the reconstruction-level variable. This accounts for events outside of the fiducial region at particle-level which enter the measured distributions at detector-level, due to the smearing effects of limited detector resolution.
- ***Migration matrix:*** The migration matrix is a two-dimensional matrix, M_{dp} , whose entries correspond to the probability that an event in particle-level bin p will be reconstructed in the detector-level bin d . It is measured in MC by comparing the distributions at particle- and detector-level for events which pass both fiducial- and reconstruction-level selections. This accounts for measurements at detector-level which

have migrated with respect to the particle-level distribution - a result of resolution effects smearing the reconstructed distributions. The detector-level normalization of the matrix diagonal, corresponding to the fraction of detector-level events which originate from the same bin at particle-level, is referred to as the *fiducial purity*.

17.2 Algorithm

This analysis uses an iterative unfolding technique based on Bayes' theorem¹⁹, known as *iterative Bayesian unfolding*, in which the fiducial-level prediction serves as the initial prior and the output of each iteration is fed back in as the subsequent prior [62]. The algorithm can be summarized for a single iteration by

$$U_i^n = \frac{1}{\varepsilon_i} \times \sum_j^{\text{reco}} \left((R_j - F_j) * f_j * w_j^{\text{pre-unf}} * \frac{M_{ji} U_i^{n-1}}{\sum_k^{\text{truth}} M_{jk} U_k^{n-1}} \right), \quad (63)$$

where U_i^n represents the unfolded yield in the i^{th} particle-level bin on the n^{th} iteration. The observed reconstruction-level yield in the j^{th} detector-level bin is denoted as R_j , while the data-driven reducible background estimate is labelled as F_j . The factor $w_j^{\text{pre-unf}}$ represents the *pre-unfolding* correction, to be discussed in more detail in Section 17.4.

¹⁹ Bayes' theorem provides a mathematical formulation for how the probability of a certain event A_i should be updated based on observation of event B , based on the likelihood of B given the assumption of A_i and the prior belief in A_i - written as

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}. \quad (62)$$

In the case of unfolding, events are considered to be reconstruction- or particle-level bin counts of a distribution, and the migration matrix serves the role of the likelihood. When iterating the process, the outcome of the previous iteration is used as the prior for the subsequent iteration.

17.3 Binning

The particular binning of measured observables is an important consideration in the unfolding procedure. Given the large increase of statistics over the previous analysis round, differential observables measured in this analysis are binned more finely to better preserve the underlying shape information. However, the granularity of the binning is limited by the available data statistics and the necessity of minimizing large bin migrations in the unfolding procedure. Additionally, a sufficient coarseness in the binning is necessary to ensure that the event yields in bins are approximately Gaussian distributed - an important assumption for the treatment of statistical uncertainties and the subsequent interpretations.

Using the signal predictions as input, an algorithmic procedure is used to set the binning for all measured observables. Bins are required to contain a minimum of 14 events passing the reconstruction-level selection and furthermore satisfy a variable fiducial purity threshold depending on the number of predicted events, summarized as follows:

- $25 \leq N_{\text{reco}}$ and fiducial purity $\geq 60\%$, or
- $20 \leq N_{\text{reco}} < 25$ and fiducial purity $\geq 70\%$, or
- $14 \leq N_{\text{reco}} < 20$ and fiducial purity $\geq 80\%$.

Where applicable, bins are also required to be centered around expected resonant peaks in the distributions. The results of this procedure are shown for the single-differential $m_{4\ell}$ distribution in Figure 49. By design, the fiducial purity is above 60% in all bins, ensuring the migrations in neighboring bins are kept low.

The migration matrix for single-differential $m_{4\ell}$, shown in Figure 50, is observed to be highly diagonal, a direct result of the high purity requirements of the binning procedure. The four differential mass regions outlined in Section 14 are also considered as an observable and treated following the same unfolding procedure. The migration matrix for this four-bin observable is shown in Figure 51 and is also highly diagonal, with a purity of 99.1% in the $Z \rightarrow 4\ell$ enriched bin. In the case of double-differential observables, the migration matrix

is formed using the entire two-dimensional distribution concatenated along a single axis to account for migrations between slices. These inter-slice migrations are by design low, as shown for the m_{Z_1} - $m_{4\ell}$ migration matrix in Figure 52.

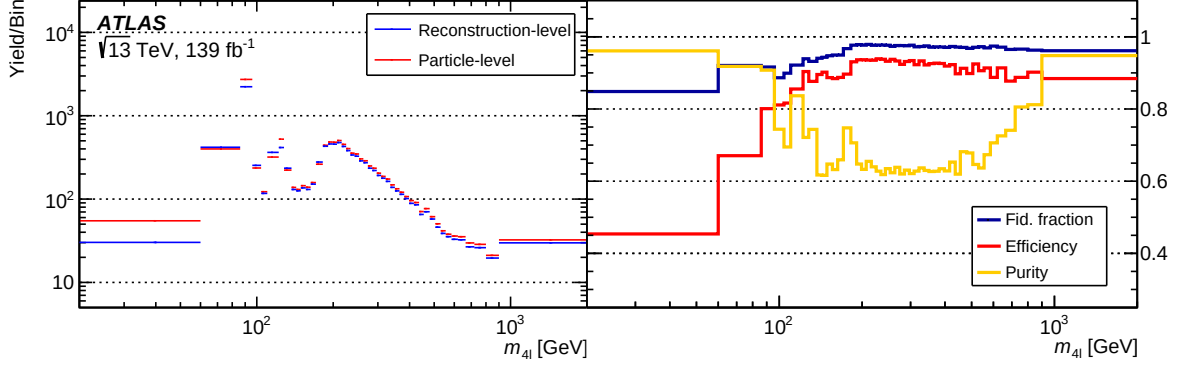


Figure 49: The predicted reconstruction- and particle-level yields (left) and the one-dimensional unfolding inputs (right) used to set the binning for the single-differential $m_{4\ell}$ observable. All yields and inputs are shown after pre-unfolding weights have been applied.

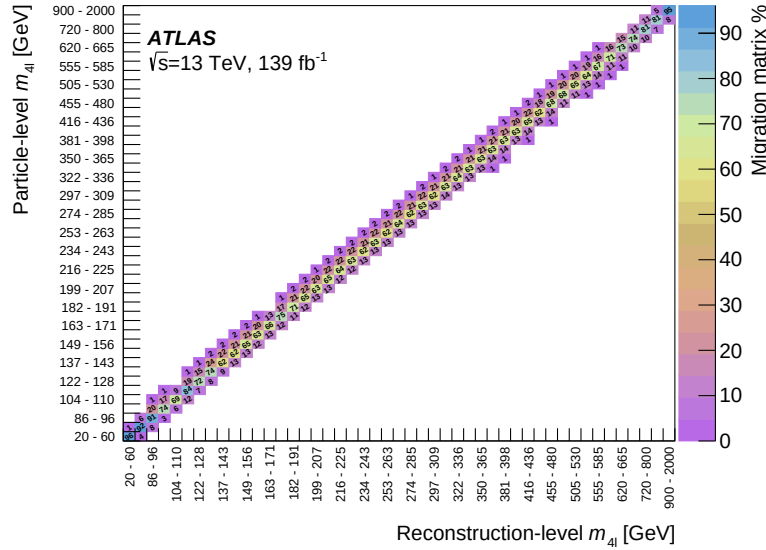


Figure 50: The single-differential $m_{4\ell}$ migration matrix showing the probability that an event in a given particle-level bin (along the vertical axis) will be reconstructed in the same detector-level bin (along the horizontal axis). The bin contents are normalized to the detector-level prediction, making the diagonal equivalent to the fiducial purity. Derived after pre-unfolding weights have been applied.

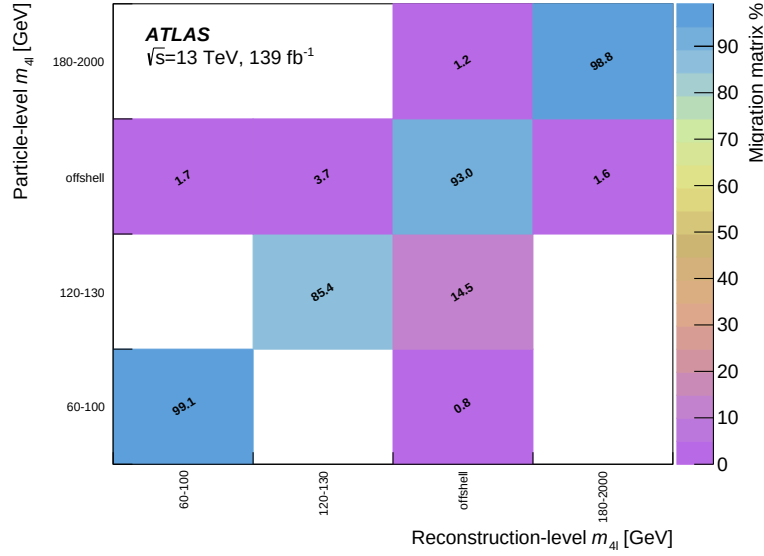


Figure 51: The inclusive- $m_{4\ell}$ migration matrix showing the probability that an event in a given particle-level bin (along the vertical axis) will be reconstructed in the same detector-level bin (along the horizontal axis). The bin contents are normalized to the detector-level prediction, making the diagonal equivalent to the fiducial purity. Derived after pre-unfolding weights have been applied.

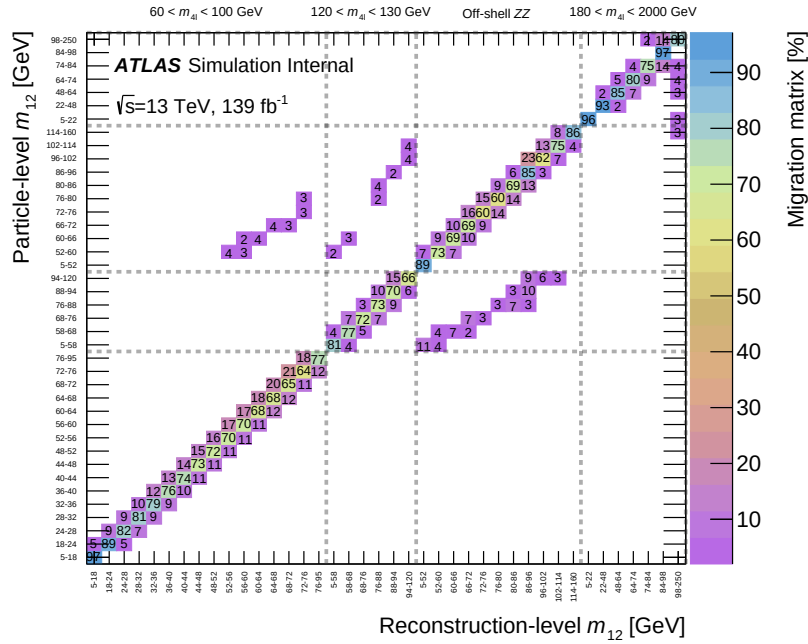


Figure 52: The m_{Z_1} - $m_{4\ell}$ migration matrix showing the probability that an event in a given particle-level bin (along the vertical axis) will be reconstructed in the same detector-level bin (along the horizontal axis). The bin contents are normalized to the detector-level prediction, making the diagonal equivalent to the fiducial purity. Derived after pre-unfolding weights have been applied.

17.4 Pre-unfolding

Definition

In order to further improve the unfolding robustness, a novel technique referred to as pre-unfolding is used in this analysis. The process consists of applying an event-level pre-unfolding weight $w^{\text{pre-unf}}$ on both the signal MC and the background-subtracted data directly before the iterative Bayesian unfolding process. It is derived as the inverse of the pre-unfolding efficiency $\varepsilon^{\text{pre-unf}}$, which itself is defined as the product of the individual lepton efficiencies ε_ℓ of the quadruplet,

$$w^{\text{pre-unf}} = \frac{1}{\varepsilon^{\text{pre-unf}}} = \frac{1}{\prod_{\ell=1}^4 \varepsilon_\ell(p_T^\ell, \eta^\ell)}. \quad (64)$$

The lepton-level efficiencies take into account the combined efficiency of reconstructing, identifying, isolating, and track-to-vertex associating (TTVA) each lepton in the quadruplet. These are obtained in MC by separately measuring the reconstruction efficiency $\varepsilon_{\text{reco}}$, as the fraction of baseline to true leptons profiled in truth kinematics, and the isolation and TTVA efficiency $\varepsilon_{\text{iso \& TTVA}}$, as the fraction of signal to baseline leptons profiled in reconstructed kinematics. The product of these two efficiencies gives a good approximation of the overall efficiency, contingent on the intentional similarity of the fiducial and detector region definitions,

$$\varepsilon_\ell \approx \varepsilon_{\text{reco}} \times \varepsilon_{\text{iso \& TTVA}}. \quad (65)$$

The lepton efficiencies are profiled versus η and p_T independently for electrons and muons to account for the kinematic and flavor dependencies of lepton reconstruction. Additionally, owing to the added difficulty of reconstructing leptons from tau decays, efficiencies are measured separately for leptons originating from tau decays and other sources. The overall pre-unfolding correction is then obtained from the weighted sum of tau and non-tau sources

according to the tau composition f_τ predicted by MC,

$$w^{\text{pre-unf}} = f_\tau \times w^{\text{pre-unf}, \tau} + (1 - f_\tau) \times w^{\text{pre-unf}, \text{non-}\tau}. \quad (66)$$

The measured reconstruction efficiencies for electrons and muons from non-tau and tau sources, profiled versus true lepton η and p_T , are shown in Figure 53. In general, the efficiency for muons is significantly higher, particularly at high lepton p_T , as is the efficiency for leptons not from tau decays. The measured isolation and TTVA efficiencies for electrons and muons from non-tau and tau sources, profiled versus reconstructed lepton η and p_T , are shown in Figure 54. The efficiency of electrons and muons not from tau decays is comparable, owing largely to the robust `FIXEDCUTPFLOWLOOSE` isolation working point. Following Equation 65, the bin-wise product of these efficiencies gives the overall pre-unfolding efficiencies, which are shown in Figure 55. Together with the predicted fraction of leptonic tau decays, these total efficiencies are used to derive the final pre-unfolding weight as per Equations 64 and 66.

Impact

The pre-unfolding correction is designed to minimize bias in the unfolding procedure, particularly in cases where the predicted SM signal composition does not match the composition observed in data. The impact and benefit of including the pre-unfolding correction is therefore best observed by performing an injection test of a genuine BSM signal. The reconstruction-level observation of the combined SM+BSM signal is taken as the pseudodata sample, which is unfolded using the SM-only unfolding response. The unfolded distribution is then compared to the true SM+BSM particle-level distribution, and the difference between the two is taken as the unfolding bias.

The results of such an injection test are shown for a $m_{\text{res}} = 300$ GeV heavy resonance in Figure 56, with and without the pre-unfolding (“pre-UF”) correction applied. In the case of

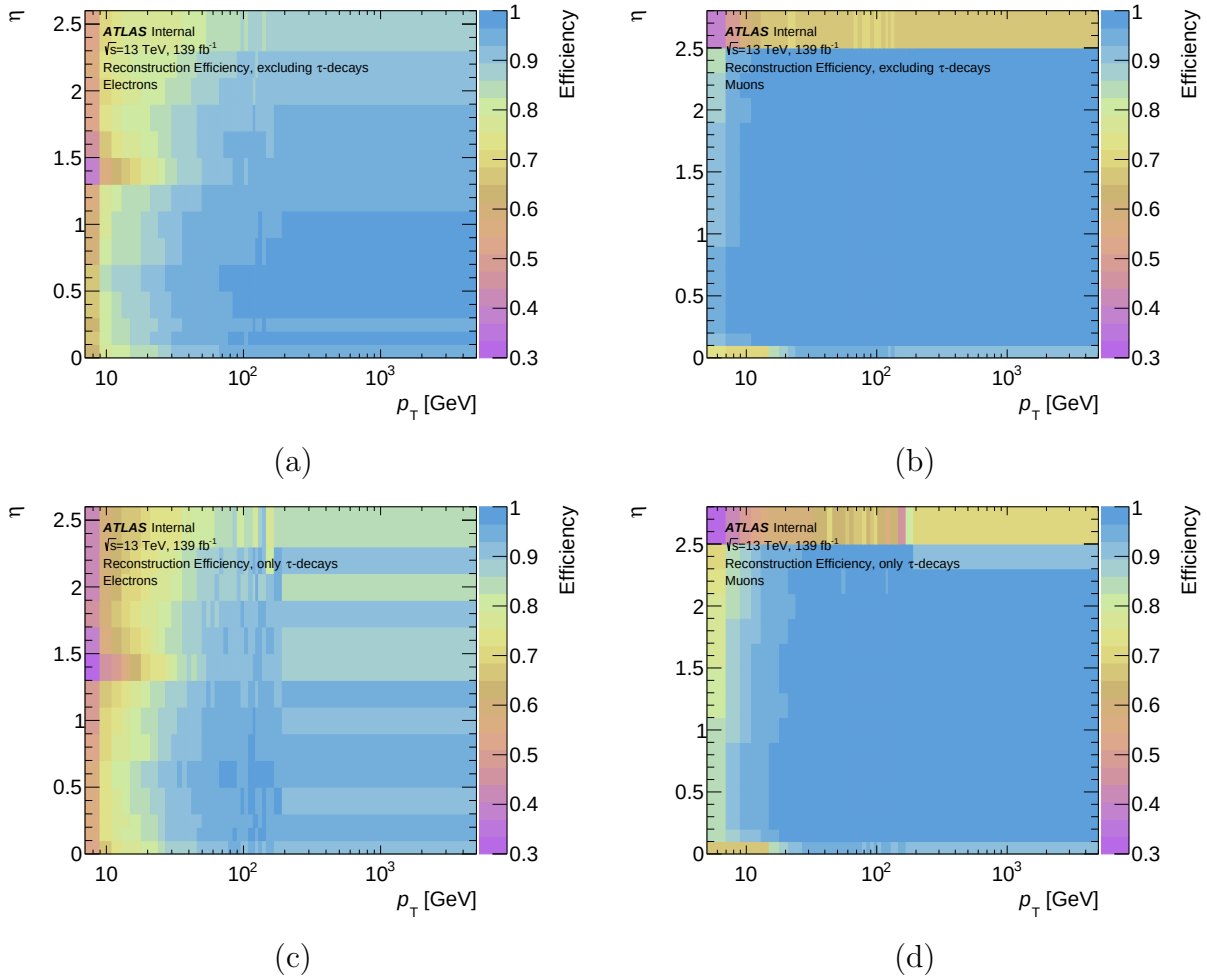


Figure 53: The reconstruction efficiency measured using the signal MC and profiled versus the true lepton η and p_T for electrons (a) and muons (b) not from τ -decays, and electrons (c) and muons (d) only from τ -decays.

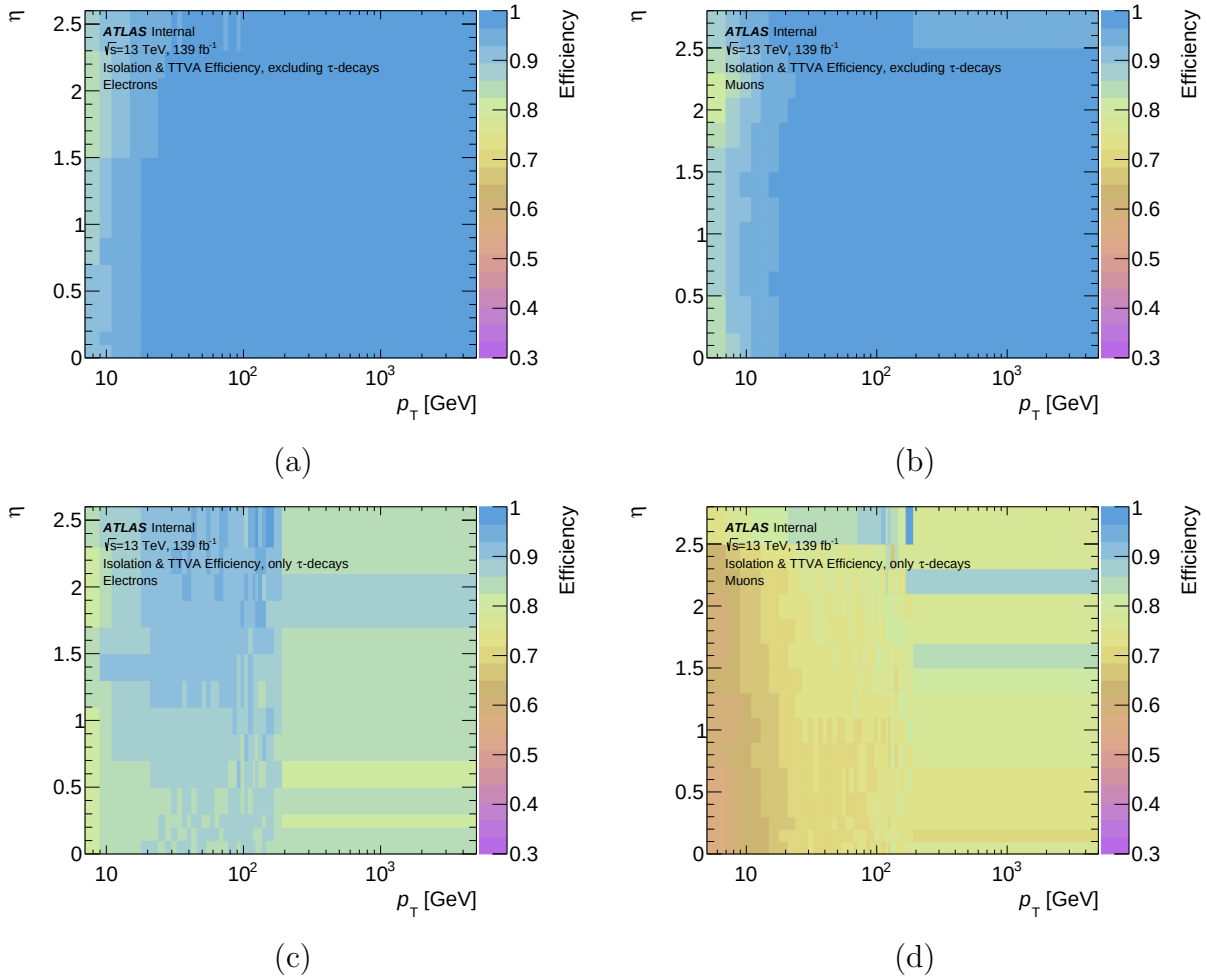


Figure 54: The isolation and track-to-vertex association efficiency measured using the signal MC and profiled versus the reconstructed lepton η and p_T for electrons (a) and muons (b) not from τ -decays, and electrons (c) and muons (d) only from τ -decays.

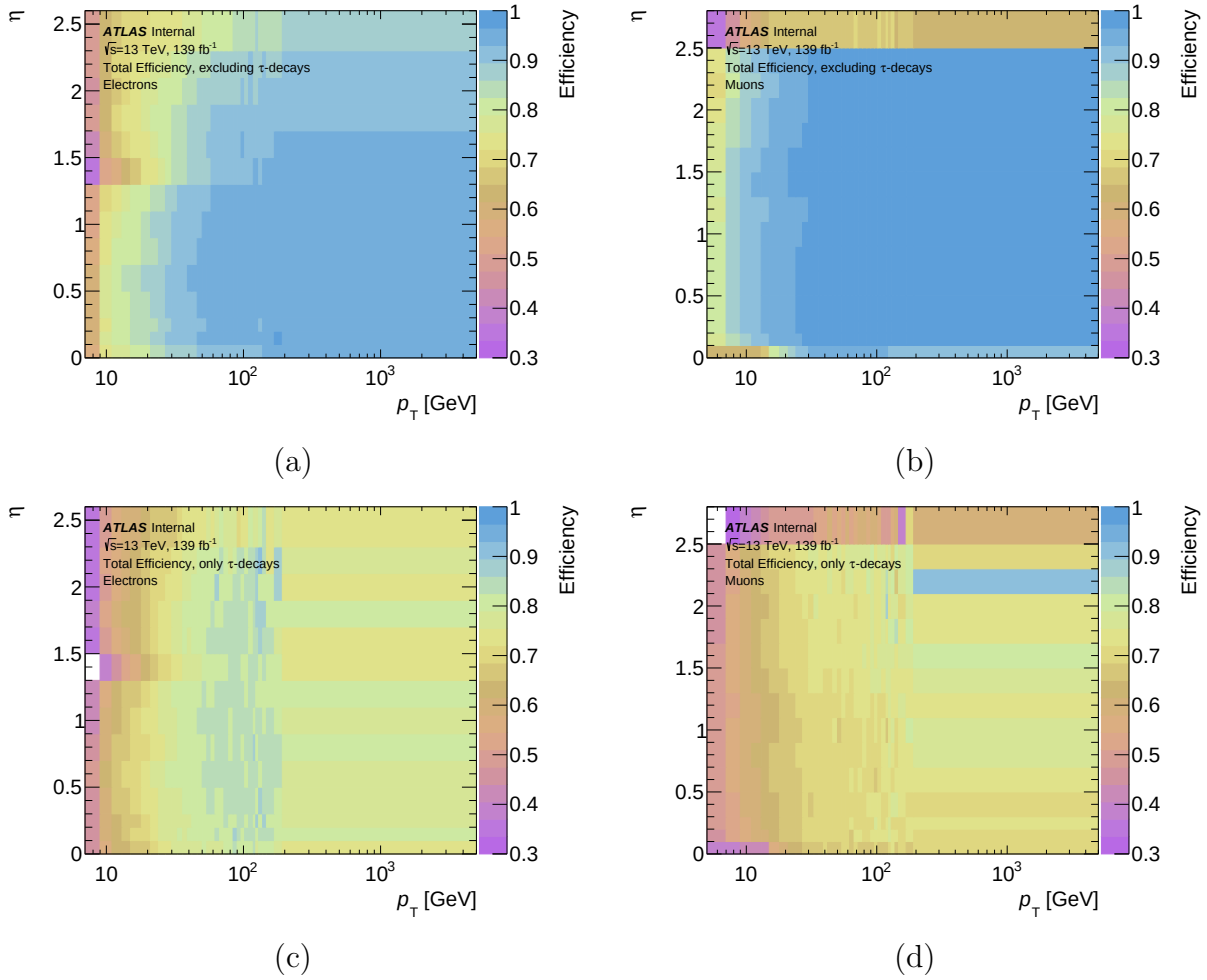


Figure 55: The pre-unfolding efficiency inputs, defined as the total reconstruction, isolation, and track-to-vertex association efficiency, as measured using the signal MC for electrons (a) and muons (b) not from τ -decays, and electrons (c) and muons (d) only from τ -decays.

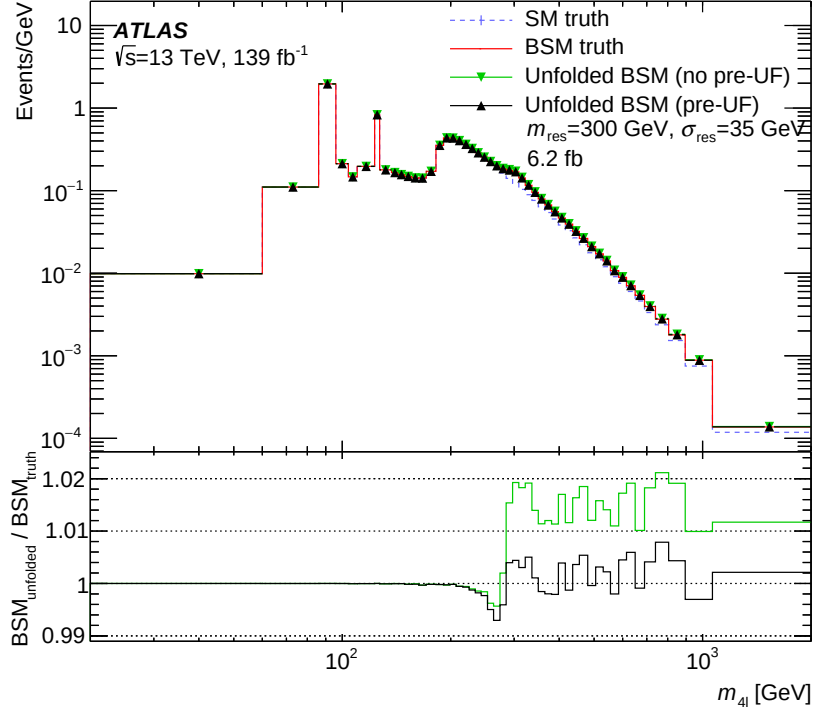


Figure 56: Summary of an injection test demonstrating the unfolding robustness and the importance of the pre-unfolding correction. A $m_{\text{res}} = 300$ GeV heavy resonance with a width of $\sigma_{\text{res}} = 35$ GeV and cross-section of 6.2 fb is injected into the nominal SM prediction. The reconstruction-level SM+BSM signal is unfolded using the nominal SM-only unfolding response, with and without the pre-unfolding correction applied, and compared to the corresponding SM+BSM truth distribution. The pre-unfolding correction is shown to significantly reduce the unfolding bias near the injected resonance.

not applying the pre-unfolding correction, a large bias of over 2% between the unfolded and truth SM+BSM distributions is visible near the injected resonance. Applying pre-unfolding reduces this bias to within half of a percent, demonstrating its effectiveness in recovering genuine BSM signals using SM-based unfolding.

The advantage of including the pre-unfolding correction can also be demonstrated by unfolding a varied-SM pseudodata sample, in which the signal strengths of one or more processes are rescaled from the SM expectation. Similarly to the injection test, the pseudodata sample is unfolded using the nominal SM unfolding response and the relative difference from the corresponding truth distribution is taken as the unfolding bias. The sensitivity

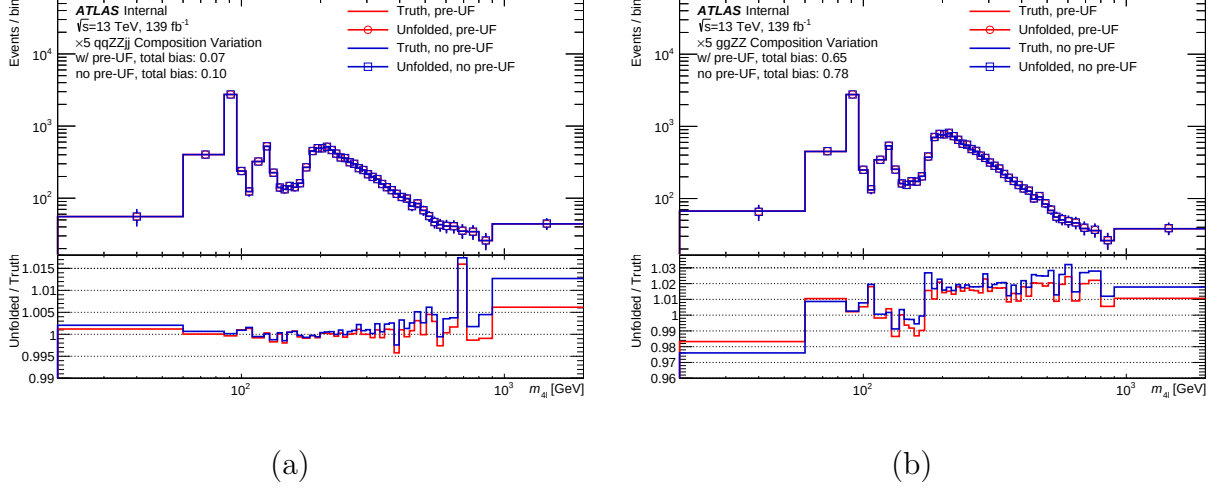


Figure 57: Results of unfolding varied-SM pseudodata samples, in which the signal strength of either the electroweak $qqZZjj$ (a) or $ggZZ$ (b) component is scaled upwards by a factor of five from the SM expectation. The pseudodata samples are unfolded using the nominal SM unfolding response and the relative difference from the corresponding truth distributions are taken as measures of the unfolding bias. The effect of including the pre-unfolding correction reduces the unfolding bias in both cases. The integrated bias is given in the upper left of each plot to better quantify the differences.

of the unfolding procedure on the assumed signal composition should be reduced with pre-unfolding weights applied. The results of such a test are shown with and without applying the pre-unfolding correction in Figure 57, where either the electroweak $qqZZjj$ (a) or $ggZZ$ (b) component is varied by a factor of five. In either case, the bias is reduced with the pre-unfolding correction applied, reflecting a reduction in the model dependence.

17.5 Optimization

The unfolding procedure is designed to minimize the model dependence (bias), in terms of both the simulation of the underlying SM processes as well as the potential presence of BSM physics. An important parameter in optimizing the iterative Bayesian unfolding technique is the number of iterations (see Equation 63). The tradeoff lies between reducing the model dependence by increasing the number of iterations and lowering the statistical uncertainty by decreasing the number of iterations. Model dependence in the procedure arises from using the SM predictions to derive the initial prior and unfolding inputs, and generally decreases

with more iterations. As the number of iterations increases, however, statistical fluctuations in the data become amplified through repeated bin migrations, resulting in a larger statistical uncertainty. The impact of each effect varies depending on the observable being measured; therefore, it is important to determine the optimal number of iterations for each observable.

One approach to determine the optimal number of iterations for a given observable is via the minimization of an uncertainty-based metric. In this first case, the metric is defined as the sum in quadrature of the statistical and systematic (bias) uncertainties, with a relative multiplicative weight of c on the latter component,

$$\sigma_{\text{total}} = \sqrt{\sigma_{\text{stat}}^2 + (c \times \sigma_{\text{sys}})^2}. \quad (67)$$

The bias component is evaluated from the data-driven closure test, discussed in the coming Section 17.6. The results of evaluating this uncertainty metric for several unfolding iterations, as well as an alternative bin-by-bin procedure, are shown as a function of single-differential $m_{4\ell}$ in Figure 58, setting the multiplicative weight to $c = 1$ (a) and $c = 10$ (b). The “Bayesian0” option corresponds to the maximally-biased procedure of simply returning the initial prior without any unfolding, where the statistical uncertainty is assumed to be zero. Without weighting the bias component, this method, which completely disregards the measured data, is preferred over all other unfolding procedures. By instead weighting the systematic component by a factor of ten, the single iterative Bayesian procedure is preferred. This shows that without properly weighting the bias over the statistical component, fewer iterations are generally preferred because of the much lower statistical uncertainty.

The former approach is limited because it relies on the measured data, which is subject to statistical fluctuations. A more statistically-robust approach for determining the optimal number of iterations is therefore preferred. The technique used here relies on drawing an ensemble of 1,000 Gaussian-distributed toys from the SM prediction. The bias is defined on a per-toy basis based on the non-closure between a given unfolded toy and the SM truth

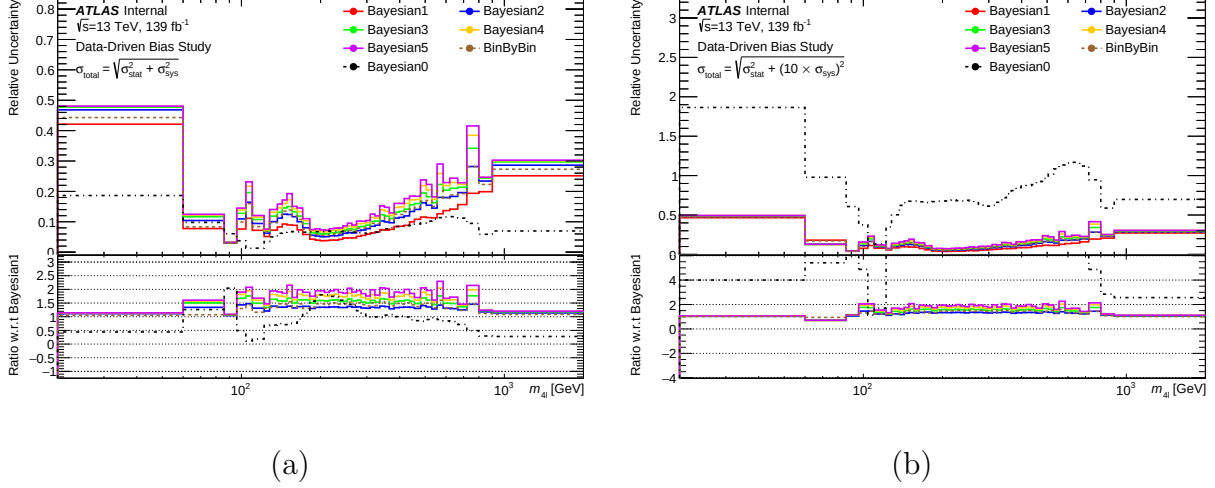


Figure 58: Uncertainty-based approach for choosing the optimal number of unfolding iterations, shown for single-differential $m_{4\ell}$. The total uncertainty is calculated based on the pre-defined metric in Equation 67. In (a), the statistical and systematic sources are weighted equally. In (b), the systematic source is given a relative weight of 10 with respect to the statistical component. The “Bayesian0” option is equivalent to returning the initial prior without unfolding.

distribution. Instead of evaluating a total uncertainty as in the previous method, the relative impact of the bias to the statistical uncertainty is measured for each toy. A statistically-robust estimate of this is obtained by evaluating the root-mean-square over all toys. The choice of iterations is then decided by requiring the bias to be suppressed by a least a factor of two compared to the statistical uncertainty. Based on this metric, three iterations is optimal for the majority of distributions, while m_{Z_1} , $|\Delta\phi_{\ell\ell}|$, $|\Delta y_{\text{pairs}}|$, and the inclusive observables require only two iterations.

17.6 Validation

Monte Carlo Closure Test

The unfolding procedure is first validated using MC, via a half-sample closure test. The full set of signal MC is randomly split into two equally sized sets based on the parity of the event number (i.e. odd and even). The reconstruction-level odd MC is taken as the pseudodata, and unfolded using the even MC. The result of which is compared to the fiducial-level odd

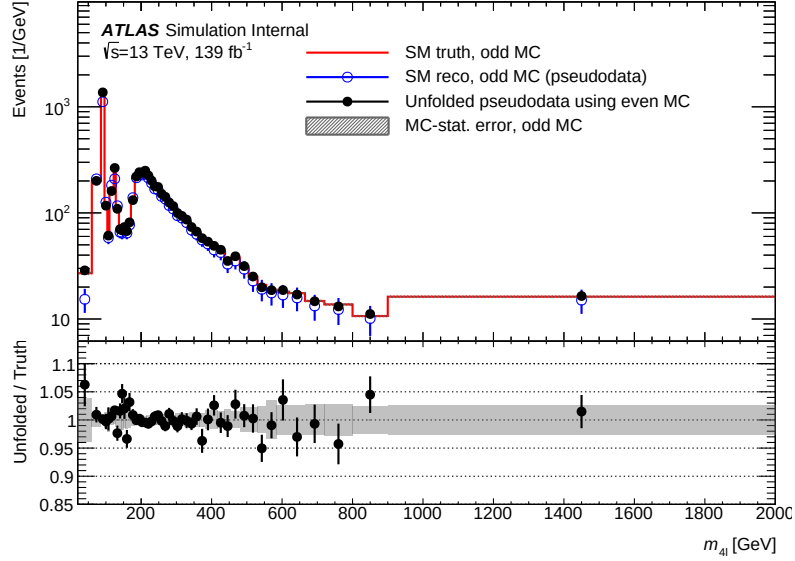


Figure 59: Results of the half-sample MC closure test for single-differential $m_{4\ell}$. In this test, the MC is split randomly into two equally sized sets (odd and even). The reconstruction-level odd MC is then unfolded using the even MC, and compared against the fiducial-level odd MC. Statistical uncertainties on the pseudodata and even MC are evaluated using the bootstrapping method.

MC, where good closure within statistical uncertainties indicates the unfolding procedure is working properly. The statistical uncertainties are evaluated on both sets of the MC using the bootstrapping technique with sets of 100 randomly sampled toys. This technique is discussed in more detail in Section 18, where it is also used for the evaluation of statistical uncertainties on the unfolded data. The statistical uncertainty on the pseudodata is akin to the data statistical uncertainty, while the uncertainty on the even MC used for unfolding accounts for the limited MC statistics. The results of this half-sample closure test are shown for the single-differential $m_{4\ell}$ observable in Figure 59. There is generally a good agreement between the unfolded pseudodata and the corresponding truth distribution within uncertainties.

Data-driven Closure Test

A secondary validation of the unfolding procedure is performed using the measured data. This data-driven closure test is also used to derive a systematic uncertainty estimating the inherent bias in the unfolding procedure. As before, the general idea is to construct a pseu-

data sample where the exact fiducial distribution is known in order to derive an estimate of the unfolding bias. In order to best reflect what is measured in data, the pseudodata is derived by reweighting the MC by the measured data/MC ratio. More specifically, the SM fiducial-level prediction is reweighted at truth-level by the data/MC ratio measured at reconstruction-level, thereby accounting for any overall shape differences in the MC. The data/MC ratio is parametrized in each single-differential observable slice using Friedman's Super Smoother technique [61], fixing the end points to the value of the ratio in the first and last bins. This provides a series of continuous reweighting functions for each one-dimensional observable. By design, the reconstruction-level pseudodata observation provides a closer description of the measured data than the nominal prediction. The pseudodata is then unfolded using the nominal response matrix and the relative difference of this result with respect to the reweighted fiducial-level distribution is taken as a systematic uncertainty. The summary of this procedure is shown step-by-step for the single-differential $m_{4\ell}$ observable in Figure 60. The results of this test for various other observables are shown in Appendix E. This bias is generally well below 1% for most observables, demonstrating the robustness of the unfolding procedure.

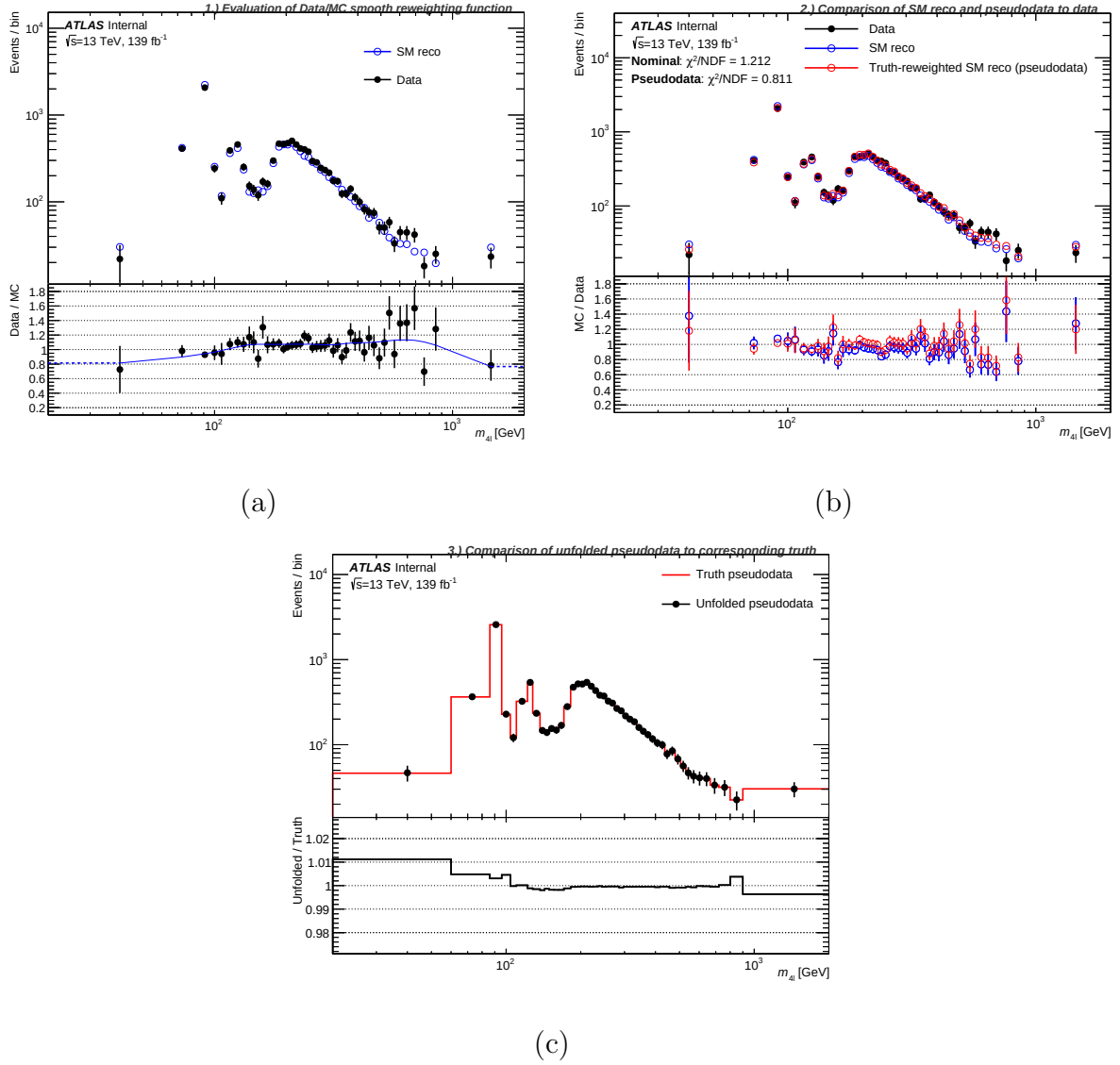


Figure 60: Step-by-step overview of the data-driven closure method for the single-differential $m_{4\ell}$ observable. In (a), the observed data is compared to the nominal reconstruction-level prediction, and a continuous reweighting function is derived by smoothing the observed data/MC ratio while keeping the endpoints fixed, as shown by the dashed lines. In (b), the reconstruction-level truth-reweighted MC (“pseudodata”) and the nominal MC are shown alongside the observed data, showing, as expected, an improved agreement with the data for the reweighted sample. In (c), the result of unfolding the pseudodata using the nominal response matrix is compared to the true prediction of the pseudodata. The relative difference between the two is taken as a source of systematic uncertainty, capturing the inherent bias in the unfolding procedure.

18 Uncertainties

This section details the treatment of uncertainties on the predictions and measurements. Statistical uncertainties, namely those from limited data statistics, are the dominant source of error in most regions, though the effect of systematic uncertainties is non-negligible. Systematic uncertainties, arising from theoretical and experimental sources, directly impact the particle- and detector-level predictions, and are propagated to the unfolded measurements by means of their impact on the unfolding response.

Formalism

Considering the i^{th} bin of nominal distribution N_i , the uncertainty σ_i due to a given variation V_i is generally defined as the relative deviation from the nominal value,

$$\sigma_i = \frac{V_i - N_i}{N_i}. \quad (68)$$

In cases where up ($+1\sigma$) and down (-1σ) variations are available, the corresponding uncertainty band is evaluated, conservatively, as the maximal, symmetric envelope,

$$\sigma_i^{\text{band}} = \max(|\sigma_i^{\text{up}}|, |\sigma_i^{\text{down}}|). \quad (69)$$

The total uncertainty of S uncorrelated uncertainties is equivalent to their sum in quadrature,

$$\sigma_i^{\text{tot.}} = \sqrt{\sum_{s=1}^S (\sigma_i^s)^2}. \quad (70)$$

Concerning the bin-to-bin correlations of a given variation, the ij^{th} elements of the corresponding uncertainty covariance matrix are given by

$$C_{ij} = (V_i - N_i)(V_j - N_j), \quad (71)$$

which have corresponding correlation matrix elements defined by

$$\rho_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \quad (72)$$

Summing S covariance matrices linearly is equivalent to combining the corresponding uncertainties in quadrature,

$$C_{ij}^{\text{tot.}} = \sum_{s=1}^S C_{ij}^s. \quad (73)$$

18.1 Systematics

Systematic uncertainties on the particle- and detector-level predictions are evaluated by calculating the relative impact of a given variation on the corresponding nominal prediction. Fiducial-level uncertainties account for only theoretical sources of uncertainty on the predictions. Reconstruction-level uncertainties include the former theoretical sources of uncertainty, as well as all sources of experimental uncertainty such as those pertaining to the background estimation. Unlike the results at fiducial-level, these reconstruction-level uncertainties are evaluated on the combination of the SM prediction and reducible background estimate. Uncertainties on the unfolded measurements include all former theoretical and experimental uncertainties, as well as additional uncertainties accounting for the particular unfolding process itself. The evaluation of unfolded uncertainties requires repeating the unfolding process using varied MC to derive the unfolding inputs (see Section 17), where each variation is also applied in a correlated manner on the samples used in the fake factor background estimation technique (see Section 16). Once the background-subtracted data is unfolded, the calculation of the uncertainty follows the same procedure as for the predictions. Unfolded-level uncertainties are also evaluated using the SM reconstruction-level prediction in place of the collision data, in order to derive the expected uncertainties. For reasons discussed in Section 22, these expected uncertainties are preferred over the measured uncertainties for use in the interpretations.

Theoretical Sources

To assess theoretical uncertainties on the simulated samples (see Table 10), on-the-fly event weights accounting for QCD scale and parton distribution function (PDF) variations in the matrix elements are evaluated for the dominant $q\bar{q} \rightarrow 4\ell$ and $gg \rightarrow 4\ell$ samples. Uncertainties on the higher order corrections (k -factors) applied to the $q\bar{q} \rightarrow 4\ell$ and $gg \rightarrow 4\ell$ samples are also propagated through the predictions. Lastly, normalization uncertainties based on recent cross-section measurements are considered for the subdominant samples. Unless otherwise noted, the evaluation of uncertainties on the dominant $q\bar{q} \rightarrow 4\ell$ process is identical for the nominal SHERPA and alternative POWHEG+PYTHIA8 samples. The theoretical uncertainties considered in this analysis are outlined in more detail below.

- QCD scale:** To account for QCD factorization (μ_F) and renormalization (μ_R) scale uncertainties, a six-point envelope is constructed by independently scaling μ_F and μ_R by -50% and 100% , excluding the maximally opposing factors [63]. The total uncertainty is evaluated as the absolute envelope of the six variations from the nominal $\{\mu_F, \mu_R = 1.0, 1.0\}$, as shown for single-differential $m_{4\ell}$ in Figure 61. For the $gg \rightarrow 4\ell$ sample, however, which relies on an external cross-section normalization, the scale variations are first normalized to the same inclusive cross-section of the nominal such that only changes in acceptance are considered.
- PDF tune:** Following the NNPDF3.0nnlo prescription [51], the PDF variations include a set of 100 internal variations, reweightings to the alternative MMHT2014nnlo [64] and CT14nnlo [65] PDF sets, and variations of the strong coupling constant of ± 0.001 ($\alpha_S = 0.117, 0.119$). The total uncertainty band is taken as the absolute envelope of the standard deviation of the internal variations and the two alternate PDF variations, combined in quadrature with the envelope of the α_S variations,

$$\sigma_{\text{PDF}}^{\text{NNPDF3.0nnlo}} = \sqrt{\max(\sigma_{\text{internal}}, |\sigma_{\text{MMHT2014nnlo}}|, |\sigma_{\text{CT14nnlo}}|)^2 + \sigma_{\alpha_S}^2}. \quad (74)$$

This procedure is shown for single-differential $m_{4\ell}$ in Figure 62.

- For the alternative POWHEG+PYTHIA8 $q\bar{q} \rightarrow 4\ell$ sample, the CT10nlo uncertainty evaluation prescription [54] is instead followed. This includes 26 pairs of upward and downward internal variations and reweightings to the alternative NNPDF3.0nnlo and MSTW2008 [66] PDF sets. The total internal uncertainty is evaluated, asymmetrically, as the sum in quadrature of the internal variations. The total uncertainty band is then taken as the absolute envelope of the total internal upward and downward uncertainties and the two alternate PDF variations,

$$\sigma_{\text{PDF}}^{\text{CT10nlo}} = \max(\sigma_{\text{internal}}^{\text{up}}, |\sigma_{\text{internal}}^{\text{down}}|, |\sigma_{\text{NNPDF3.0nnlo}}|, |\sigma_{\text{MSTW2008}}|). \quad (75)$$

- **Parton shower modeling:** A symmetric flat uncertainty of $\pm 2\%$ is applied on the $q\bar{q} \rightarrow 4\ell$ sample to account for possible differences in the parton shower modeling.
- **$q\bar{q} \rightarrow 4\ell$ NLO EW corrections:** The uncertainties on the differential- $m_{4\ell}$ virtual NLO EW k -factors [67] are propagated through the $q\bar{q} \rightarrow 4\ell$ prediction.
- **$gg \rightarrow 4\ell$ NLO QCD corrections:** Likewise, the uncertainties on the differential- $m_{4\ell}$ NLO QCD k -factors [68] are propagated through the $gg \rightarrow 4\ell$ prediction.
- **ggH cross-section:** The cross-section uncertainty of on-shell Higgs boson production via the primary gluon-gluon fusion mode is evaluated based on two variations, separately accounting for QCD scale $\left({}^{+4.6\%}_{-6.7\%}\right)$ and PDF ($\pm 3.2\%$) uncertainties, taken from the previous ATLAS publication [69].
- **$t\bar{t}V$ and VVV cross-sections:** Flat cross-section uncertainties are taken, conservatively, from recent ATLAS measurements [70, 71] and applied on the $t\bar{t}V$ ($\pm 15\%$) and VVV ($\pm 10\%$) predictions. An alternative $t\bar{t}V$ sample is used to evaluate a more precise uncertainty on the fiducial-level prediction. This sample is generated with MADGRAPH5_aMC@NLO 2.3.3 [72] using the NNPDF3.0nnlo PDF set, and contains

on-the-fly event weights accounting for QCD scale and PDF variations. The corresponding fiducial-level uncertainty is taken as the sum in quadrature of the scale and PDF components, which are calculated following the same prescription as the SHERPA $q\bar{q} \rightarrow 4\ell$ sample.

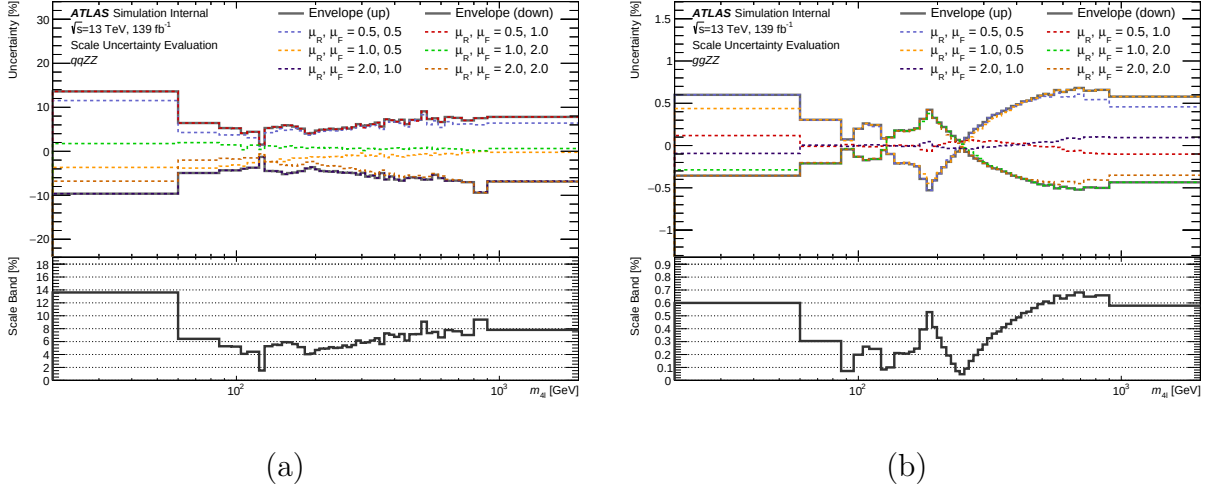


Figure 61: Evaluation of the fiducial-level QCD renormalization and factorization scale uncertainty bands for the $q\bar{q} \rightarrow 4\ell$ (a) and $gg \rightarrow 4\ell$ (b) samples versus single-differential $m_{4\ell}$.

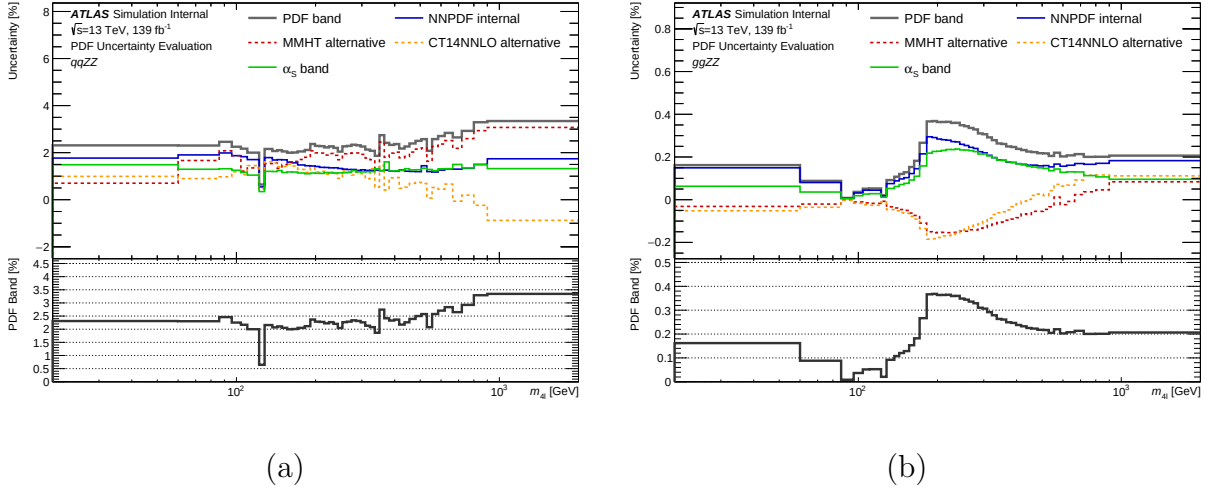


Figure 62: Evaluation of the fiducial-level PDF uncertainty bands for the $q\bar{q} \rightarrow 4\ell$ (a) and $gg \rightarrow 4\ell$ (b) samples versus single-differential $m_{4\ell}$.

Experimental Sources

Experimental uncertainties primarily account for the modeling of lepton efficiencies, resolution, and energy scale in simulations. Individual sources are generally provided by ATLAS performance groups as Gaussian nuisance parameters with $\pm 1\sigma$ upward and downward variations. The corresponding uncertainties are evaluated as the absolute envelope of the upward and downward variations from the nominal. Lepton-level variations are applied in a correlated way to all leptons in the event to obtain a conservative measure of the uncertainty. The experimental uncertainties considered in this analysis are discussed below.

- **Electron efficiencies:** Uncertainties are considered on the trigger, identification, reconstruction, and isolation efficiencies of electrons, which are provided by the ATLAS-supported tool `ElectronEfficiencyCorrection`²⁰. In total, a set of 61 nuisance parameters are included, each with upward and downward components, corresponding to 61×2 variations. A single nuisance parameter with 1×2 variations covers the uncertainties in the trigger efficiency scale factors. A set of 34×2 variations accounts for uncertainties in the identification efficiency, 25×2 variations for the reconstruction efficiency, and a single 1×2 nuisance parameter for the isolation efficiency.
- **Muon efficiencies:** Uncertainties are considered on the trigger, identification and reconstruction, isolation, and track-to-vertex association efficiencies of muons, which are provided by the ATLAS-supported tool `MuonEfficiencyCorrections`²¹. In total, a set of ten nuisance parameters are included, each with upward and downward components, corresponding to 10×2 variations. A set of two nuisance parameter with 2×2 variations covers the uncertainties in the trigger efficiency scale factors. A set of 4×2 variations accounts for uncertainties in the combined identification and reconstruction efficiency,

²⁰<https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/ElectronPhotonID/ElectronEfficiencyCorrection>

²¹<https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/MuonID/MuonIDAnalysis/MuonEfficiencyCorrections>

2×2 variations for the isolation efficiency, and 2×2 variations for the track-to-vertex association efficiency.

- **Electron resolution and scale:** Two nuisance parameters, corresponding to 2×2 variations, account for uncertainties in the electron momentum resolution and scale.
- **Muon resolution and scale:** Two nuisance parameters, corresponding to 2×2 variations, account for uncertainties in the muon momentum resolution, as measured separately by the Inner Detector and Muon Spectrometer. Additionally, two nuisance parameters, again having 2×2 variations, account for uncertainties in the Sagitta correction. Lastly, a single 1×2 nuisance parameter covers the uncertainties in the muon momentum scale.
- **Pile-up reweighting:** A single 1×2 nuisance parameter accounts for upward and downward variations in the pile-up reweighting factor used to re-weight the predictions to match the observed pile-up profiles in data.
- **Luminosity:** The uncertainty in the 139 fb^{-1} integrated luminosity as collected by ATLAS over the full Run-2 data-taking period is $\pm 1.7\%$ [47]. This is applied as a flat uncertainty variation on the fiducial- and reconstruction-level predictions.

Background Sources

Background systematics account for uncertainties in the fake factor estimation procedure, discussed in Section 16. These uncertainties are evaluated on both the reconstruction-level predictions, where the total yield is taken as the signal prediction plus the reducible background estimate, and the unfolded measurements, where the background is first subtracted from the measured data. In total, four background-related uncertainties are included, which are outlined below.

- **Fake factors:** Systematic and statistical uncertainties are derived on the fake factors,

based on the central values and uncertainties of the fake efficiencies as measured in the $Z \rightarrow \ell\ell$ CR (see Section 16.4), and propagated through the estimation.

- **Smoothing:** A single uncertainty accounts for the specifics of the background smoothing procedure, as discussed in Section 16.6.
- **Inclusive CR statistics:** A single uncertainty accounts for the limited data statistics in the inclusive control region where the fake factors are applied, which is evaluated via the bootstrapping method.

Unfolding Sources

There are two systematics specific to the unfolding process itself and, hence, relevant for the unfolded measurements only. These are summarized below.

- **Unfolding bias:** A single uncertainty accounts for the inherent bias of the unfolding method, evaluated based on the data-driven closure test discussed in Section 17.6.
- **Generator choice:** A single uncertainty covers the differences in the response matrix modeling of the dominant $q\bar{q} \rightarrow 4\ell$ process between the SHERPA and POWHEG+PYTHIA8 event generators. This is motivated by known differences in the modeling of QED-driven FSR between the generators. To prevent the double-counting of uncertainties related to the unfolding method itself, a version of the SM prediction including the POWHEG+PYTHIA8 $q\bar{q} \rightarrow 4\ell$ sample in place of the nominal SHERPA sample is reweighted at particle-level to reproduce the SHERPA lineshape. The background-subtracted data is then unfolded using this reweighted POWHEG+PYTHIA8 sample, and the relative difference with respect to the nominal unfolded result is taken as the systematic uncertainty.

18.2 Statistics

In the case of predicted SM fiducial- or reconstruction-level yields, the statistical uncertainty is taken directly as the Poisson error on the simulated event counts. The situation is more complicated for unfolded measurements due to bin-to-bin correlations in the unfolding process. In the latter case, the bootstrap method is used to estimate the statistical uncertainties on the data, as well as the MC-derived unfolding inputs. The technique is also used to evaluate the impact of the limited statistics in the inclusive CR on the reducible background estimate. Concerning the unfolded measurements, the data statistical uncertainty is the dominant source of uncertainty in most bins, reflecting the statistical limit of the analysis.

Bootstrapping

The bootstrap method uses random sampling with replacement to estimate a statistic from a population [73]. The usage of the technique in ATLAS is discussed further in Ref. [74]. For the purposes of this analysis, it is used to estimate the impact of various statistical uncertainties on the unfolded measurements.

Given a set of data consisting of N events, a bootstrap replica (toy) is generated by resampling with replacement N times. Each toy thus consists of a subset of events from the nominal dataset with an integer weight corresponding to their number of occurrences. The uncertainty is estimated as the relative standard deviation of a set of T bootstrap toys, typically on the order of several thousand. The statistical uncertainty is evaluated on a per-bin basis as the relative standard deviation of the distribution of bootstrap toys in each bin. Representing the set of all bootstrap toys evaluated in the i^{th} bin as the vector $\vec{B}_i$, the statistical uncertainty in the i^{th} bin is given by

$$\sigma_i = \frac{\sqrt{\langle \vec{B}_i^2 \rangle - \langle \vec{B}_i \rangle^2}}{\langle \vec{B}_i \rangle}. \quad (76)$$

Likewise, obtaining the covariance of the statistical uncertainty follows

$$C_{ij} = \frac{1}{T} \sum_{t=1}^T (B_i^t - \langle \vec{B}_i \rangle) (B_j^t - \langle \vec{B}_j \rangle). \quad (77)$$

For the case of the uncertainty on the unfolded data, 3,500 bootstrap toys are generated from the data and unfolded using the nominal response matrix. The results of this are shown for single-differential $m_{4\ell}$ in Figure 63, where the shaded error band denotes the relative statistical uncertainty as evaluated by Equation 76. By construction, the mean of the unfolded toys is virtually identical to the nominal unfolded data. In the case of the expected data uncertainty, 3,500 Poisson toys are drawn from the SM reconstruction-level prediction without pre-unfolding weights applied and unfolded using the corresponding response matrix. The uncertainty is then calculated in a similar fashion according to Equation 76 in terms of the unfolded Poisson toys.

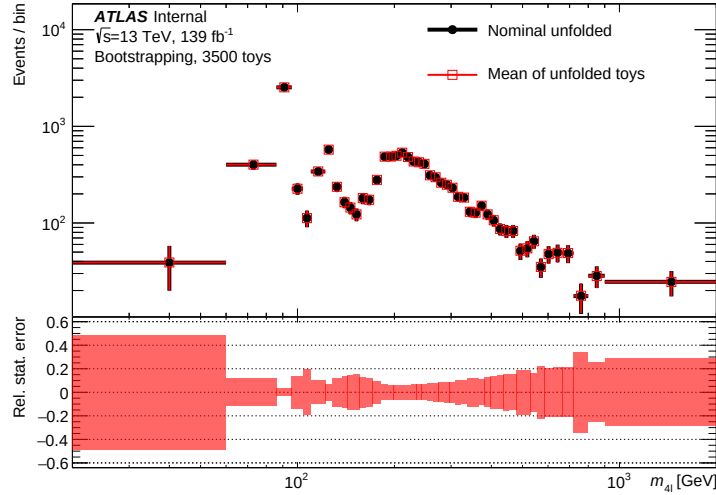
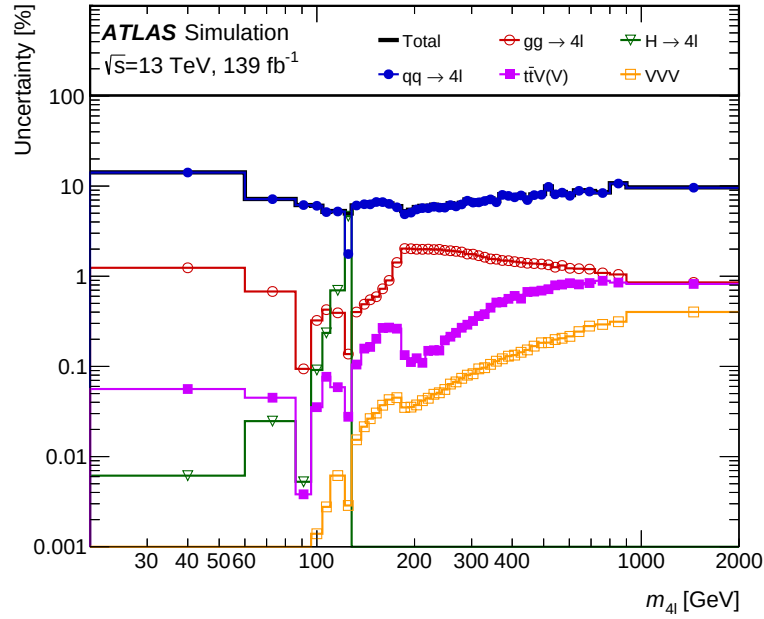


Figure 63: Summary of the bootstrapping technique using 3,500 randomly sampled Poisson toys to evaluate the statistical uncertainty on the unfolded data, as a function of single-differential $m_{4\ell}$.

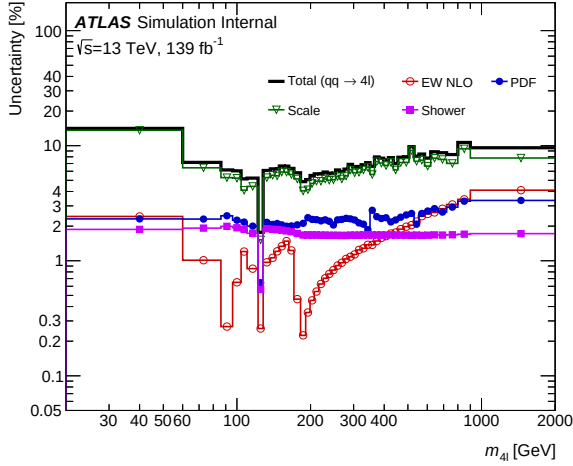
18.3 Impact

Fiducial Prediction

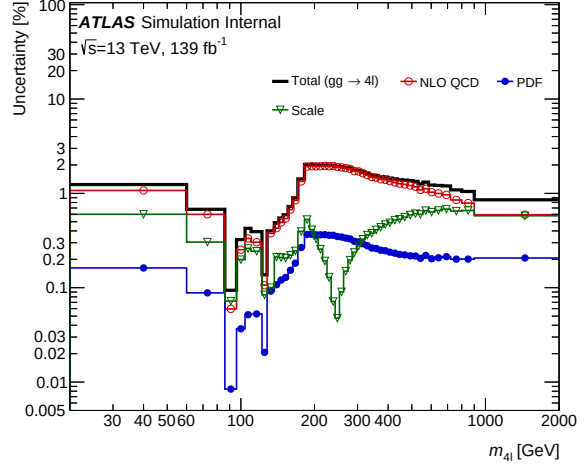
The uncertainties on the predicted fiducial-level $m_{4\ell}$ distribution are shown in Figure 64. In Figure 64 (a), the uncertainties are grouped by physics process, scaled according to the respective fraction of the total SM prediction in each bin. Due to the overall abundance of the quark-initiated process $q\bar{q} \rightarrow 4\ell$ in the analysis phase space, uncertainties on this process are dominant through much of the spectrum. A detailed breakdown of these uncertainties is shown in Figure 64 (b), where by far the dominant source is that of the QCD renormalization and factorization scale. Uncertainties specific to the gluon-initiated process $gg \rightarrow 4\ell$ are similarly shown in Figure 64 (c), where the dominant contribution comes from uncertainties on the NLO QCD correction. Given the suppression of the gluon process over the dominant quark-initiated process, these uncertainties are proportionally smaller on the inclusive prediction. Near the mass of the Higgs boson, uncertainties relating to on-shell Higgs production $H \rightarrow 4\ell$ become significant. The total uncertainty is roughly 10% over the bulk of the spectrum.



(a)



(b)



(c)

Figure 64: In (a), a breakdown of the fiducial-level systematic uncertainties on the total SM prediction, grouped by the six primary physics processes, is shown versus single-differential $m_{4\ell}$. In (b) and (c), the uncertainties on the inclusive prediction are shown separately for those affecting the $q\bar{q} \rightarrow 4\ell$ and $gg \rightarrow 4\ell$ samples, respectively.

Reconstruction Prediction

The uncertainties on the predicted reconstruction-level $m_{4\ell}$ distribution are shown in Figure 65. Theoretical uncertainties dominate in all bins, with a similar impact of approximately 10% as at fiducial-level. Of the experimental uncertainties, background and lepton-level systematics present the largest impacts in the low and high mass ranges, while the luminosity uncertainty is dominant in the intermediate range.

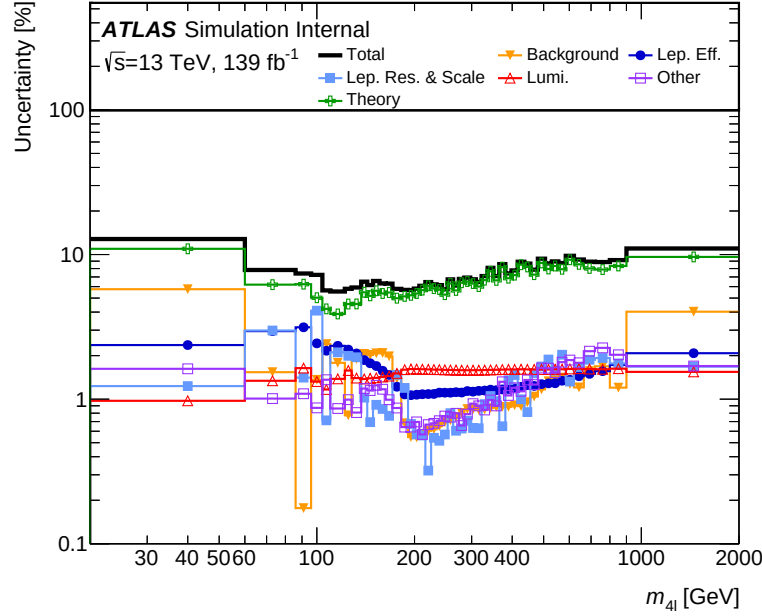


Figure 65: Breakdown of the reconstruction-level systematic uncertainties, grouped by theoretical or experimental categories, on the combined SM plus reducible background prediction versus single-differential $m_{4\ell}$.

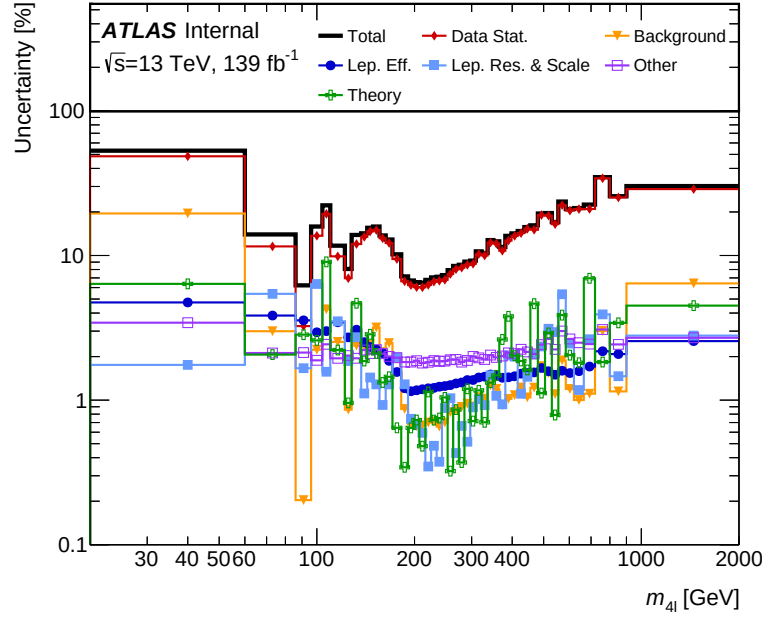
Unfolded Measurements

The uncertainties on the unfolded $m_{4\ell}$ distribution are shown in Figure 66. The statistical uncertainty is dominant in all bins except the one closest to the Z -boson mass, where lepton efficiencies play a larger role. As expected, the theory uncertainties are much smaller on the unfolded cross-section than the fiducial- or reconstruction-level predictions - excluding the $q\bar{q} \rightarrow 4\ell$ generator component which is only evaluated on the unfolded measurements

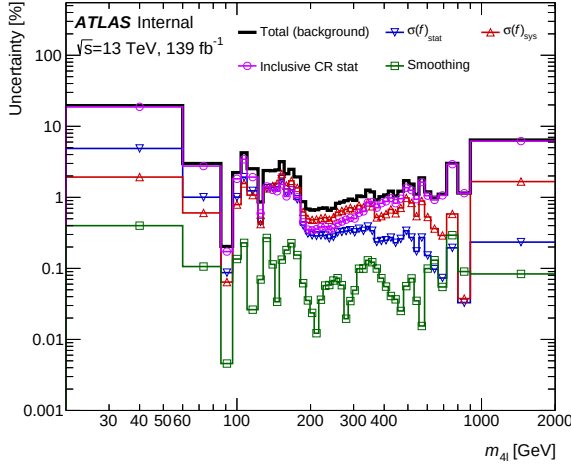
- reflecting the minimal model-dependence of the unfolding procedure. The breakdowns of background- and theory-related uncertainties are shown in Figure 66 (b) and (c), respectively. Concerning the background sources, the inclusive CR statistical uncertainty dominates over the bulk of the spectrum, while the statistical and, in particular, systematic uncertainties on the fake factors contribute significantly in the intermediate mass range. Of the theory uncertainties, the SHERPA versus POWHEG+PYTHIA8 generator comparison is dominant in most regions, followed and, in some bins, exceeded by the combined uncertainty on the $q\bar{q} \rightarrow 4\ell$ process.

The uncertainty breakdowns versus $m_{4\ell}$ are shown across the three quadruplet flavor channels in Figure 67. Notably, the systematic and statistical uncertainties are significantly lower on the fully muonic channel owing to the more efficient and robust reconstruction of muons in ATLAS. The unfolded uncertainty breakdowns are also shown for the double-differential m_{Z_1} , m_{Z_2} , and $|\Delta\phi_{\text{pairs}}|$ distributions across the four $m_{4\ell}$ slices in Figures 68, 69, and 70, respectively. In all distributions, including those not shown here, the statistical uncertainty surpasses the combined systematic uncertainty by a large margin in virtually every bin, except for those corresponding to Z -boson resonances, where uncertainties on lepton efficiencies and resolutions have a larger impact.

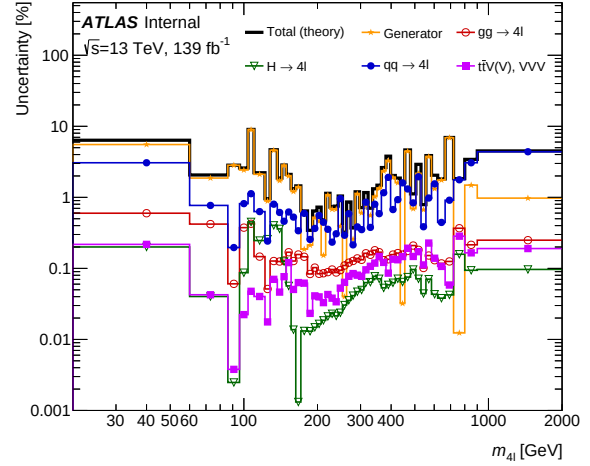
A more detailed breakdown of the systematic and statistical uncertainties is shown for the inclusive measurements in Table 14. Systematic uncertainties are ranked by their impact on the inclusive measurement, and all uncertainties greater than 1% are highlighted in bold. The systematic uncertainty is under 5% in all regions, and below 2% in the on-shell ZZ region. The statistical uncertainty in the inclusive regions is comparably low, excluding the on-shell Higgs measurement where it is by far dominant.



(a)

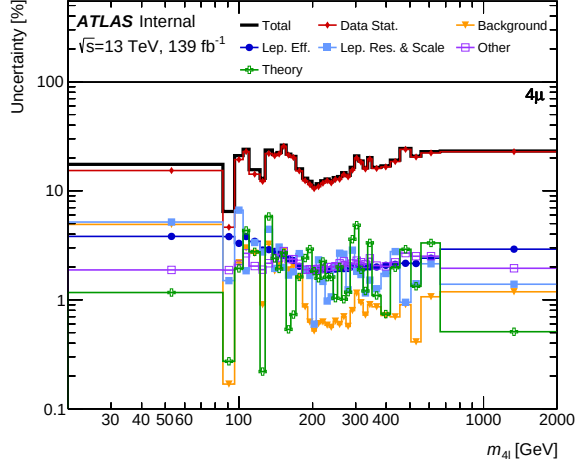


(b)

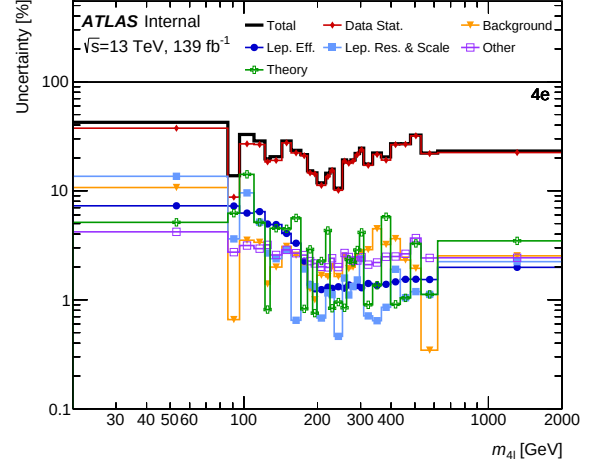


(c)

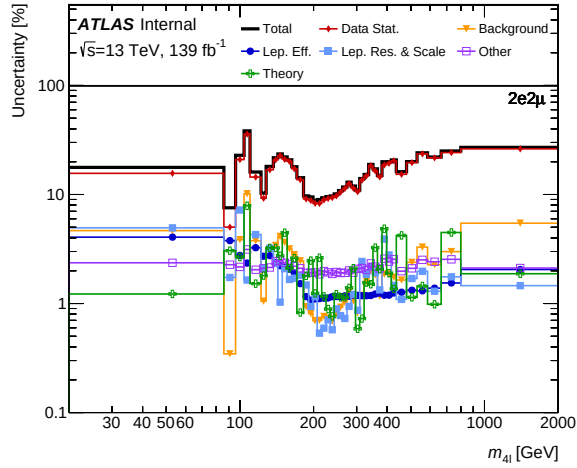
Figure 66: In (a), a breakdown of the unfolded uncertainties, grouped by theoretical or experimental categories, is shown versus single-differential $m_{4\ell}$. The statistical uncertainty on the data is the dominant source of uncertainty in all but the third mass bin where $m_{4\ell} \approx m_Z$. In (b), a detailed breakdown of only those sources related to the background estimation is shown. The total background uncertainty is mostly dominated by the statistical uncertainty in the inclusive CR, followed by the systematic uncertainty on the fake factors. In (c), the breakdown of theory sources is shown, where the generator comparison is observed to be dominant in most bins.



(a)

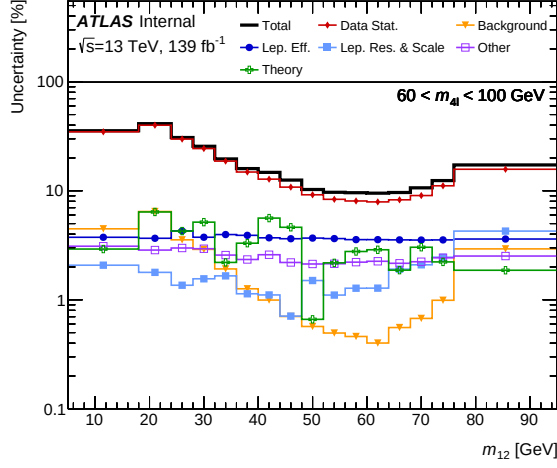


(b)

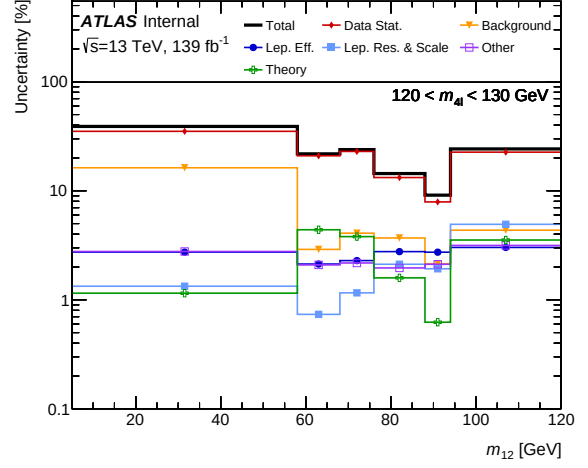


(c)

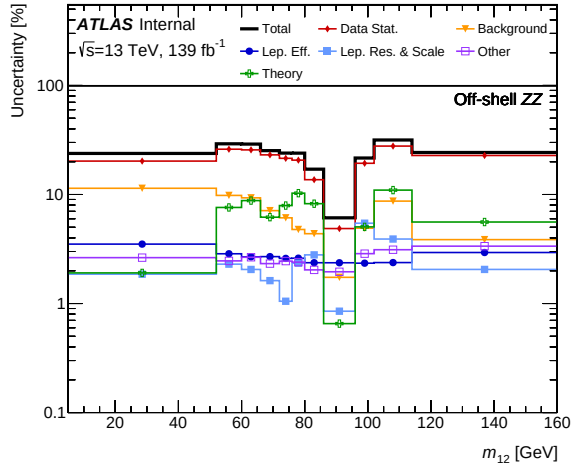
Figure 67: Breakdown of the unfolded uncertainties, grouped by theoretical or experimental categories, as a function of $m_{4\ell}$ across the three flavor channels.



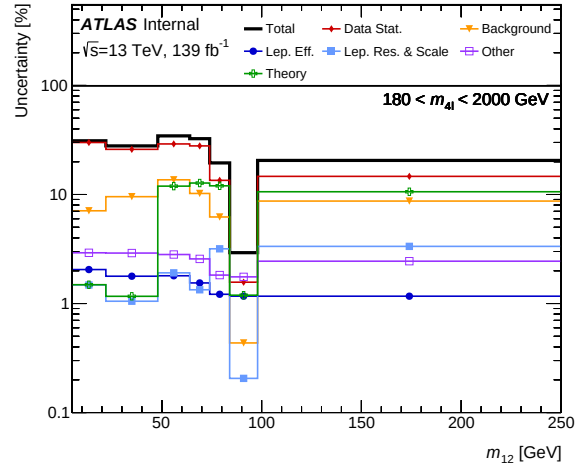
(a)



(b)

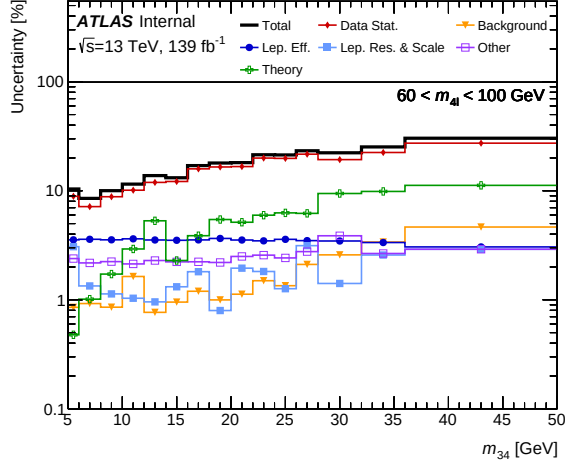


(c)

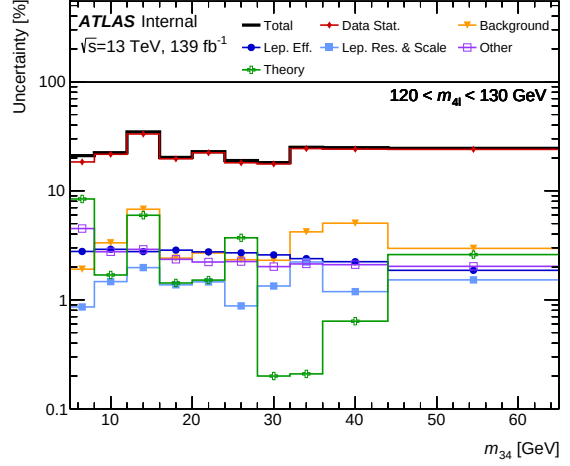


(d)

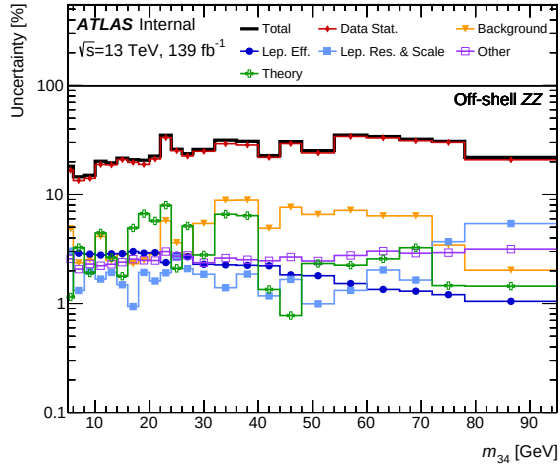
Figure 68: Breakdown of the unfolded uncertainties, grouped by theoretical or experimental categories, as a function of m_{Z_1} across the four $m_{4\ell}$ slices.



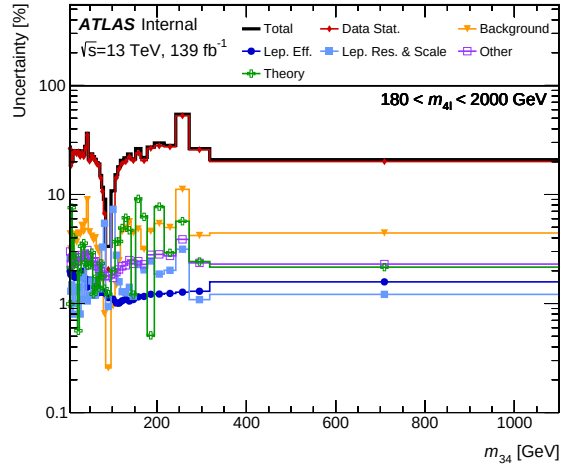
(a)



(b)

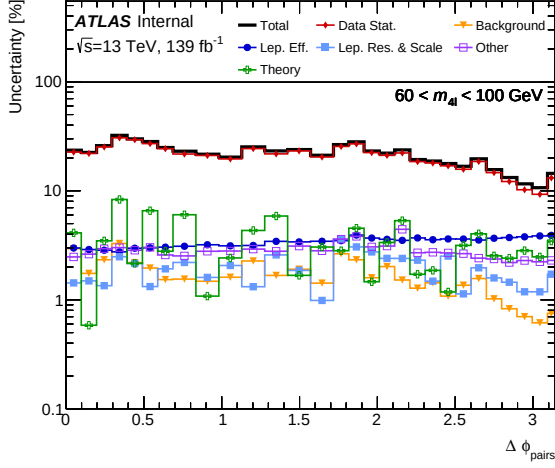


(c)

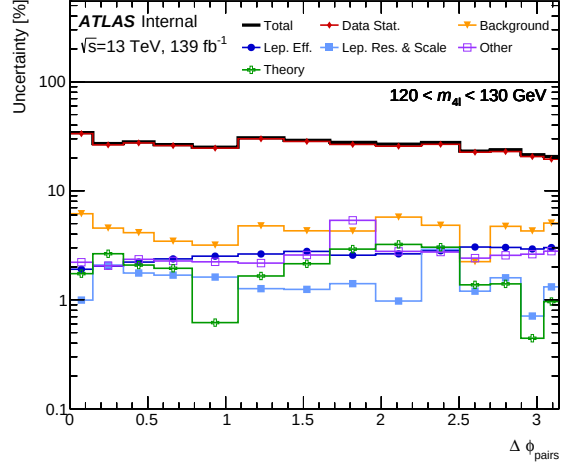


(d)

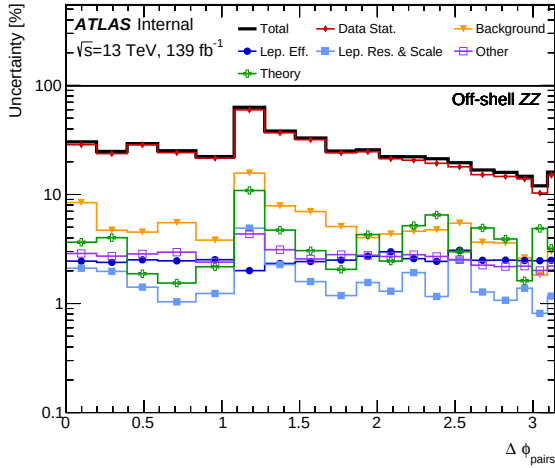
Figure 69: Breakdown of the unfolded uncertainties, grouped by theoretical or experimental categories, as a function of m_{Z_2} across the four $m_{4\ell}$ slices.



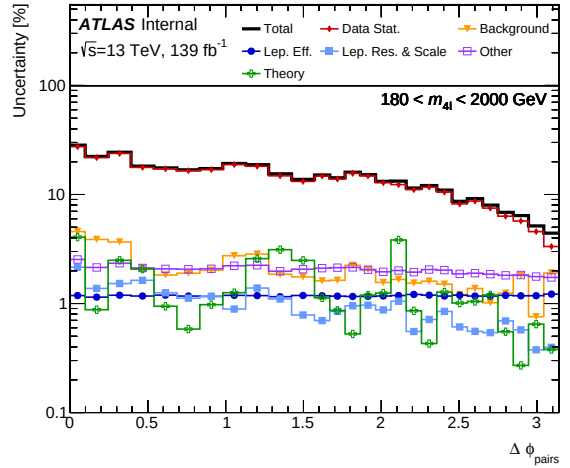
(a)



(b)



(c)



(d)

Figure 70: Breakdown of the unfolded uncertainties, grouped by theoretical or experimental categories, as a function of $|\Delta\phi_{\text{pairs}}|$ across the four $m_{4\ell}$ slices.

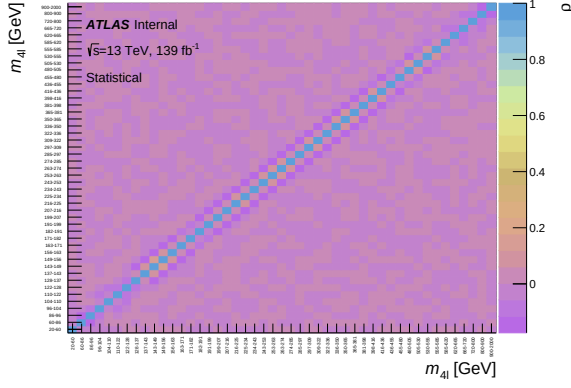
Table 14: Systematic and statistical uncertainty breakdown for the unfolded inclusive cross-sections. Systematic uncertainties are grouped by category and sorted by their impact on the inclusive measurement. Uncertainties greater than one percent are highlighted in bold.

Region $m_{4\ell}$ [GeV]	Inclusive any	$Z \rightarrow 4\ell$ 60-100	On-shell H 120-130	Off-shell ZZ 20-60/100-120/130-180	On-shell ZZ 180-2000
Generator	1.3%	2.6%	1.3%	2.7%	0.13%
Muon Isolation	0.96%	1.6%	1.2%	1.2%	0.58%
Electron ID	0.94%	1.9%	1.5%	1.3%	0.49%
Electron Reco	0.84%	1.7%	1.3%	1.2%	0.31%
Muon ID & Reco	0.83%	1.1%	0.91%	0.89%	0.82%
Reducible	0.8%	0.55%	1.7%	2.5%	0.74%
Pile-up Reweighting	0.73%	1.2%	1%	0.81%	0.47%
Electron Isolation	0.52%	1.1%	0.79%	0.73%	0.18%
Electron Res. & Scale	0.46%	1.1%	0.83%	0.54%	0.12%
Trigger	0.33%	0.8%	0.44%	0.44%	0.084%
Muon Res. & Scale	0.3%	0.65%	0.55%	0.53%	0.13%
Theory	0.27%	0.31%	0.23%	0.45%	0.27%
Muon TTVA	0.21%	0.46%	0.28%	0.31%	0.071%
Unfolding Bias	0.088%	0.35%	0.13%	0.45%	0.035%
MC Stat.	0.087%	0.22%	0.38%	0.26%	0.088%
Total Systematic	2.6%	4.8%	3.7%	4.6%	1.5%
Luminosity	1.7%	1.6%	1.6%	1.6%	1.7%
Data Stat.	1.3%	2.9%	6.2%	4.1%	1.5%
Total	3.3%	5.8%	7.4%	6.4%	2.7%

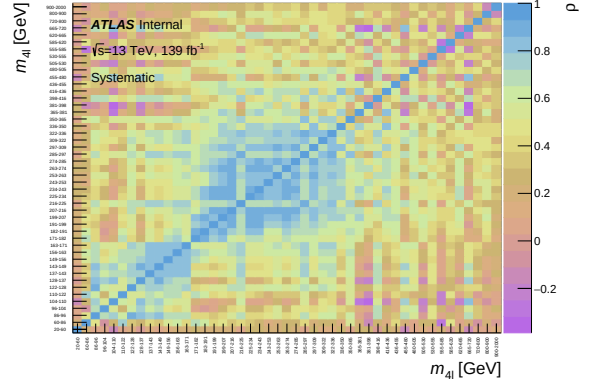
Uncertainty Correlations

The covariance matrices encoding the bin-to-bin correlations of the uncertainties on the unfolded measurements are a necessary input for the statistical interpretations. Individual covariance matrices are derived for each systematic variation, following Equation 71, and summed together linearly, as per Equation 73, to obtain the total systematic matrix. The statistical covariance matrix is derived in a similar way using the bootstrap toys, as outlined in Equation 77. The total covariance matrix is simply the linear summation of the systematic and statistical covariance matrices.

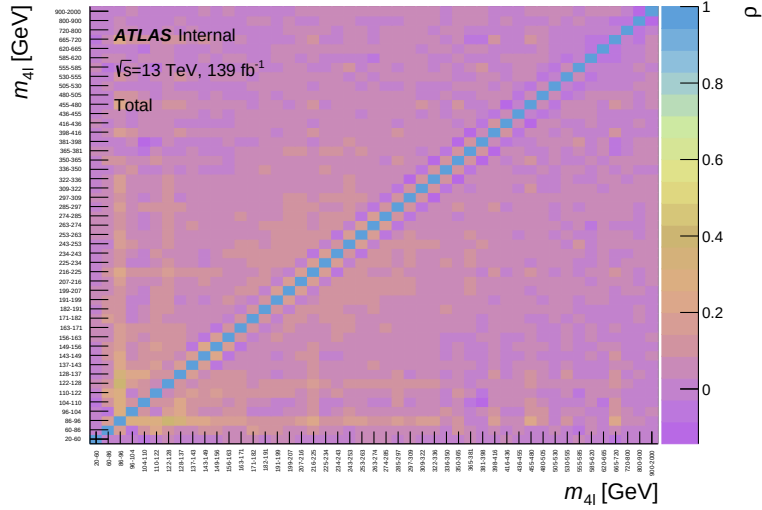
The combined statistical (a), systematic (b), and total (c) correlation matrices are shown in Figure 71. The statistical correlation matrix is largely diagonal, with off-diagonal contributions arising from bin migrations in the unfolding procedure. The systematic correlation matrix exhibits more pronounced correlation structures, given the tendency of particular systematic variations to impact entire regions of the four-lepton phase space. The total correlation matrix in Figure 71 is mostly diagonal, owing to the dominance of the data statistical uncertainty, but exhibits some off-diagonal structures from correlations in the systematic uncertainties.



(a)



(b)



(c)

Figure 71: Total statistical (a), systematic (b), and statistical + systematic (c) correlation matrices for unfolded $m_{4\ell}$.

Chapter VI: Results and Interpretations

Reconstruction-level distributions are measured using the full ATLAS Run-2 dataset for the thirteen observables and five inclusive regions outlined in Section 14. A sample of these detector-level measurements, mainly serving as intermediaries for the final unfolded results, are shown in Section 19. The unfolded differential cross-sections of these observables are compared to state-of-the-art Standard Model predictions in the fiducial phase space in Section 20. Inclusive cross-sections are also measured in the entirety of the fiducial region, as well as the four $m_{4\ell}$ regions, each of which is enriched in a different physics process. In Section 21, a precision measurement of the rare $Z \rightarrow 4\ell$ branching fraction is obtained using the unfolded cross-section as measured in the single resonant Z region. As an example of the re-interpretative power of these measurements, the unfolded cross-sections are used to constrain two scenarios of new physics in Section 22. Lastly, the prospects of future re-interpretations using the publicly available measurements are briefly discussed in Section 23.

Table 15: Predicted and measured detector-level event yields in the inclusive region and each of the four mass regions. Contributions to the total prediction are broken down by signal and background processes with associated systematic uncertainties.

Region $m_{4\ell}$ [GeV]	Inclusive any	$Z \rightarrow 4\ell$ 60-100	On-shell H 120-130	Off-shell ZZ 20-60/100-120/130-180	On-shell ZZ 180-2000
$q\bar{q} \rightarrow 4\ell$	6100 ± 500	1490 ± 120	128 ± 10	800 ± 60	3640 ± 280
$gg \rightarrow 4\ell$	680 ± 90	10.8 ± 2.9	3.9 ± 0.7	49 ± 6	620 ± 80
On-shell H	245 ± 20	2.16 ± 0.18	207 ± 17	33.5 ± 3.1	1.98 ± 0.2
VVV	35 ± 4	0.018 ± 0.005	0.127 ± 0.018	2.05 ± 0.22	32.9 ± 3.4
$t\bar{t}V$	123 ± 19	1.37 ± 0.22	1.24 ± 0.2	15.5 ± 2.4	105 ± 16
$Z + \Upsilon$	13.2 ± 1.3	1.7 ± 0.4	1.2 ± 0.4	8.7 ± 1	1.7 ± 0.4
Reducible	320 ± 50	42 ± 8	25 ± 5	120 ± 19	137 ± 30
Total Pred.	7500 ± 500	1540 ± 110	367 ± 19	1030 ± 60	4530 ± 290
Data	7755	1452	379	1095	4828

19 Detector Measurements

Reconstruction-level measurements represent the raw detector observations before any unfolding corrections are applied. First subtracting the reducible background estimate (see Section 16) and then following the iterative Bayesian unfolding technique (see Section 17), these are extrapolated to particle-level cross-sections in the fiducial phase space. The predicted and measured event yields in the five inclusive regions are shown in Table 15. The reducible background, estimated by the fake factor method, is included in the total prediction along with the predicted $Z + \Upsilon$ contribution, the latter of which contributes a minuscule source of irreducible background. Considering the associated uncertainties, there is generally good agreement between the predictions and measurements. The single-differential $m_{4\ell}$ detector observation is shown in Figure 72, with stacked histograms showing the predicted contributions from each of the signal processes and the combined background estimate. Here too, a good agreement with data is observed, where larger deviations in the high mass tail appear to be driven by statistical fluctuations in the data.

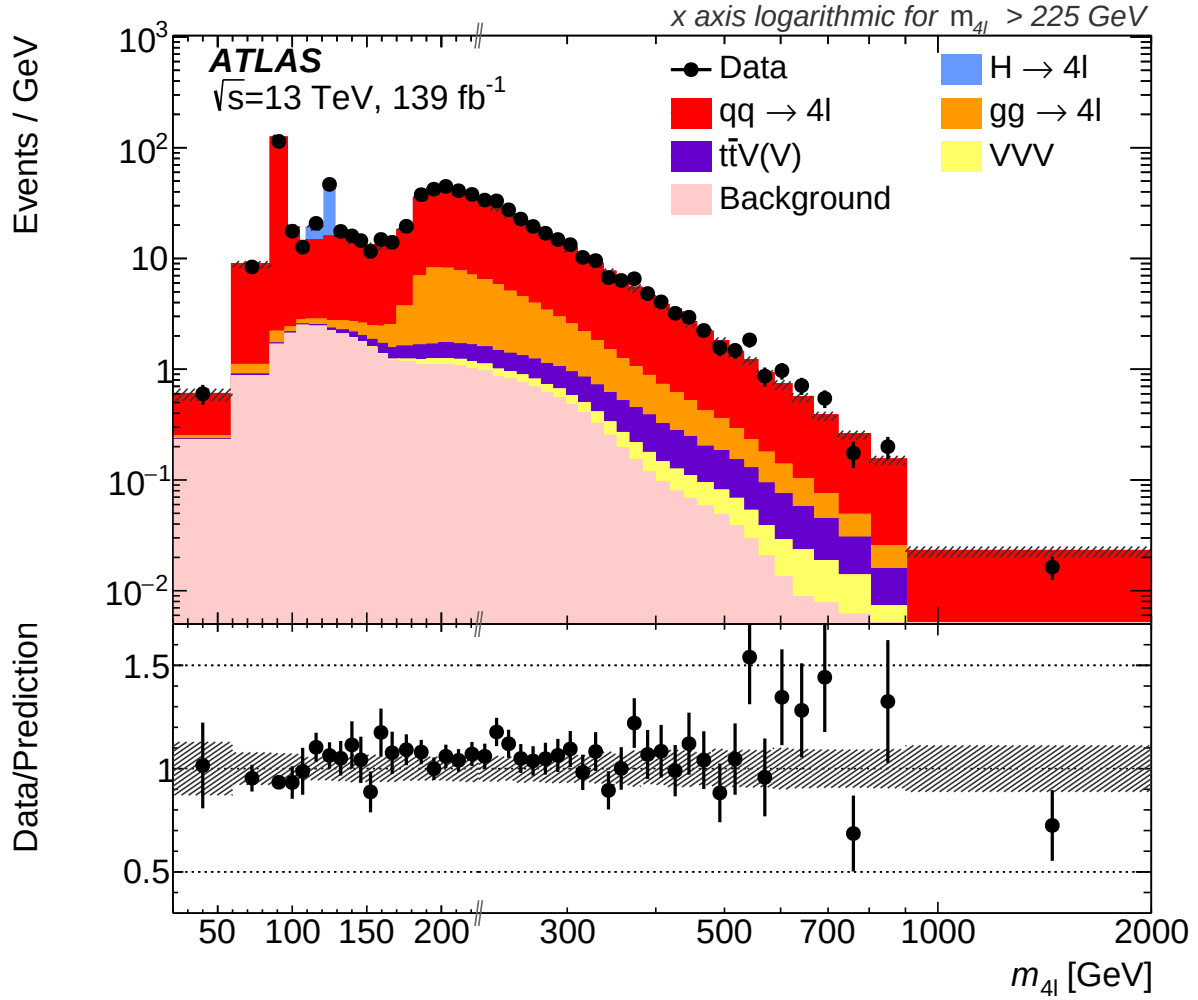


Figure 72: The detector-level m_{4l} distribution as measured in data, along with the predicted breakdown for different processes shown in the form of successively stacked histograms. The shaded band covers the systematic uncertainty on the predictions, while the error bars on the data points show the Poisson statistical uncertainty. The horizontal axis uses a linear scale until $m_{4l} = 225$ GeV, and a logarithmic scale beyond this point.

Table 16: Fiducial cross-sections measured in femtobarns in the inclusive region and each of the four mass regions, compared to the respective Standard Model predictions using either SHERPA or POWHEG+PYTHIA8 to simulate the $q\bar{q} \rightarrow 4\ell$ process. All uncertainties on the measurements and predictions, as discussed in Section 18, are included.

Region $m_{4\ell}$ [GeV]	Inclusive any	$Z \rightarrow 4\ell$ 60-100	On-shell H 120-130	Off-shell ZZ 20-60/100-120/130-180	On-shell ZZ 180-2000
Measured	88.8	22.1	4.76	12.4	49.3
fiducial	$\pm 1.1(\text{stat})$	$\pm 0.6(\text{stat})$	$\pm 0.30(\text{stat})$	$\pm 0.5(\text{stat})$	$\pm 0.8(\text{stat})$
cross-section	$\pm 2.3(\text{syst})$	$\pm 1.1(\text{syst})$	$\pm 0.18(\text{syst})$	$\pm 0.6(\text{syst})$	$\pm 0.7(\text{syst})$
[fb]	$\pm 1.5(\text{lumi})$	$\pm 0.4(\text{lumi})$	$\pm 0.08(\text{lumi})$	$\pm 0.2(\text{lumi})$	$\pm 0.8(\text{lumi})$
	$\pm 3.0(\text{total})$	$\pm 1.3(\text{total})$	$\pm 0.35(\text{total})$	$\pm 0.8(\text{total})$	$\pm 1.3(\text{total})$
Sherpa	85.7 ± 5.2	23.6 ± 1.5	4.56 ± 0.21	11.5 ± 0.7	46.0 ± 2.9
Powheg	82.6 ± 3.2	21.2 ± 1.3	4.36 ± 0.19	10.7 ± 0.5	46.4 ± 1.6

20 Unfolded Measurements

Based on the reconstruction-level observations, the thirteen observables discussed in Section 14, as well as the inclusive measurements, are extrapolated to the fiducial phase space following the iterative Bayesian unfolding procedure described in Section 17. The particle-level cross-sections are then compared to two competing Standard Model predictions, in which the dominant $q\bar{q} \rightarrow 4\ell$ process is simulated at NLO accuracy in QCD with either SHERPA or POWHEG+PYTHIA8, where the latter is normalized to a prediction at NNLO accuracy in QCD. Additional differences between the generators are expected due to the modeling of QED FSR. The measured and predicted cross-sections in the full fiducial phase space and each of the four mass regions are outlined in Table 16. The predictions are systematically lower than the measured data in all regions except $Z \rightarrow 4\ell$, where the SHERPA prediction lies above the measurement. Considering the quoted uncertainties, though, the predictions are generally within 1σ uncertainty of the data.

Statistical Compatibility

For the differential measurements, the compatibility of the theoretical predictions with the unfolded data is evaluated in terms of a p -value for each observable. This represents a measure of the statistical compatibility of the predictions in light of the measured data, and

is defined as follows. For an observable with ν bins, the residual, $\vec{r}$, is a ν -length vector defined as the difference between the unfolded data, $\vec{d}$, and the theoretical prediction, $\vec{p}$,

$$\vec{r} = (\vec{d} - \vec{p}). \quad (78)$$

Based on the total $\nu \times \nu$ covariance matrix $\mathbf{C}$, accounting for the combination of statistical and systematic uncertainties, the observed χ^2 is

$$\chi_{\text{obs}}^2 = \vec{r}^T \mathbf{C}^{-1} \vec{r}. \quad (79)$$

Assuming this follows a χ^2 -PDF (here *probability density function*) with ν degrees of freedom, the p -value corresponds to the integral of the probability distribution beyond χ_{obs}^2 ,

$$p = \int_{\chi_{\text{obs}}^2}^{\infty} \chi_{\text{PDF}}^2(\chi; \nu) d\chi. \quad (80)$$

Therefore, the p -value as defined here corresponds to the probability of observing a value of χ^2 larger than the observed value, given the prediction. For larger p -values closer to the upper bound of one, the measured χ^2 is low, meaning the agreement with data is excellent.

For the purposes of this analysis, the total covariance in Equation 79 includes both the expected SM statistical and systematic covariances on the unfolded measurements, obtained by unfolding the reconstruction-level prediction in place of the data, and the total theory covariance on the fiducial-level prediction. The expected covariance is used in place of the measured to account for possible downward fluctuations in the data below the per-bin asymptotic Gaussian limit of approximately 14 events, where the assumption of Gaussian-distributed statistical errors is no longer valid. The theory covariance includes all theoretical sources of uncertainty, which are conservatively treated as uncorrelated across the four mass regions owing to the different physics processes involved. Following Equation 73, the linear summation of these covariances corresponds to the sum in quadrature of the uncertainties.

Differential Cross-sections

The unfolded cross-section as measured for the single-differential $m_{4\ell}$ observable is shown in Figure 73, alongside the competing SHERPA and POWHEG+PYTHIA8 fiducial cross-section predictions. The data is well modeled by both simulations over the entire mass range from 20 GeV to 2000 GeV, with differences under 50% in most bins and within 2σ uncertainty bands in all bins. The Z -boson and Higgs resonances, clearly visible in the measured spectrum, are in excellent agreement with theory. The differential $m_{4\ell}$ cross-section is also measured separately for the three flavor channels: 4μ , $4e$, and $2\mu 2e$ - the results of which are shown in Figure 74. No significant deviations are observed between the fully muonic or electronic channels. A subset of the additional measurements are shown for those observables which offer particular sensitivity in the interpretations. This includes the double-differential lepton pair mass observables, $m_{Z_1}-m_{4\ell}$ and $m_{Z_2}-m_{4\ell}$, in Figures 75 and 76, respectively, and the absolute azimuthal separation of the lepton pairs, $|\Delta\phi_{\text{pairs}}|-m_{4\ell}$, in Figure 77. Given the fine granularity of the binning, statistical fluctuations are expected and indeed observed in some regions. Both the SHERPA and POWHEG+PYTHIA8 predictions are generally in good agreement with the measured data, although differences in normalization are present. These normalization differences, combined with statistical fluctuations in the data, lead to some low observed p -values, particularly for the POWHEG+PYTHIA8 predictions.

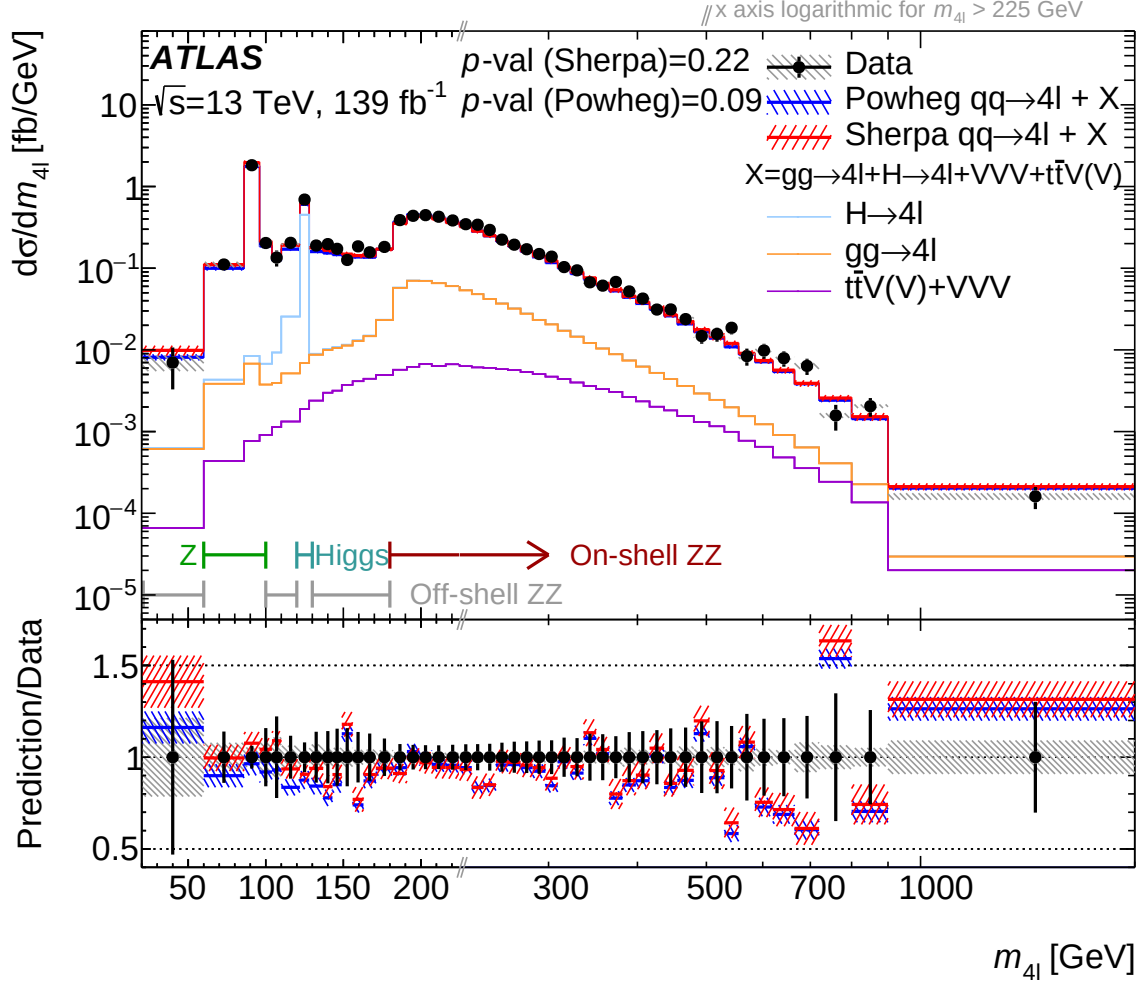


Figure 73: The unfolded single-differential $m_{4\ell}$ cross-section compared against full Standard Model predictions using either SHERPA or POWHEG+PYTHIA8 to model the dominant $q\bar{q} \rightarrow 4\ell$ contribution. The breakdown of the MC contribution is shown for different processes in the form of successively stacked histograms. Shaded bands denote the systematic uncertainties on the unfolded data and the predicted fiducial cross-sections, while the error bars on the data points show the combined statistical and systematic uncertainties. The horizontal axis uses a linear scale until $m_{4\ell} = 225$ GeV, and a logarithmic scale beyond this point.

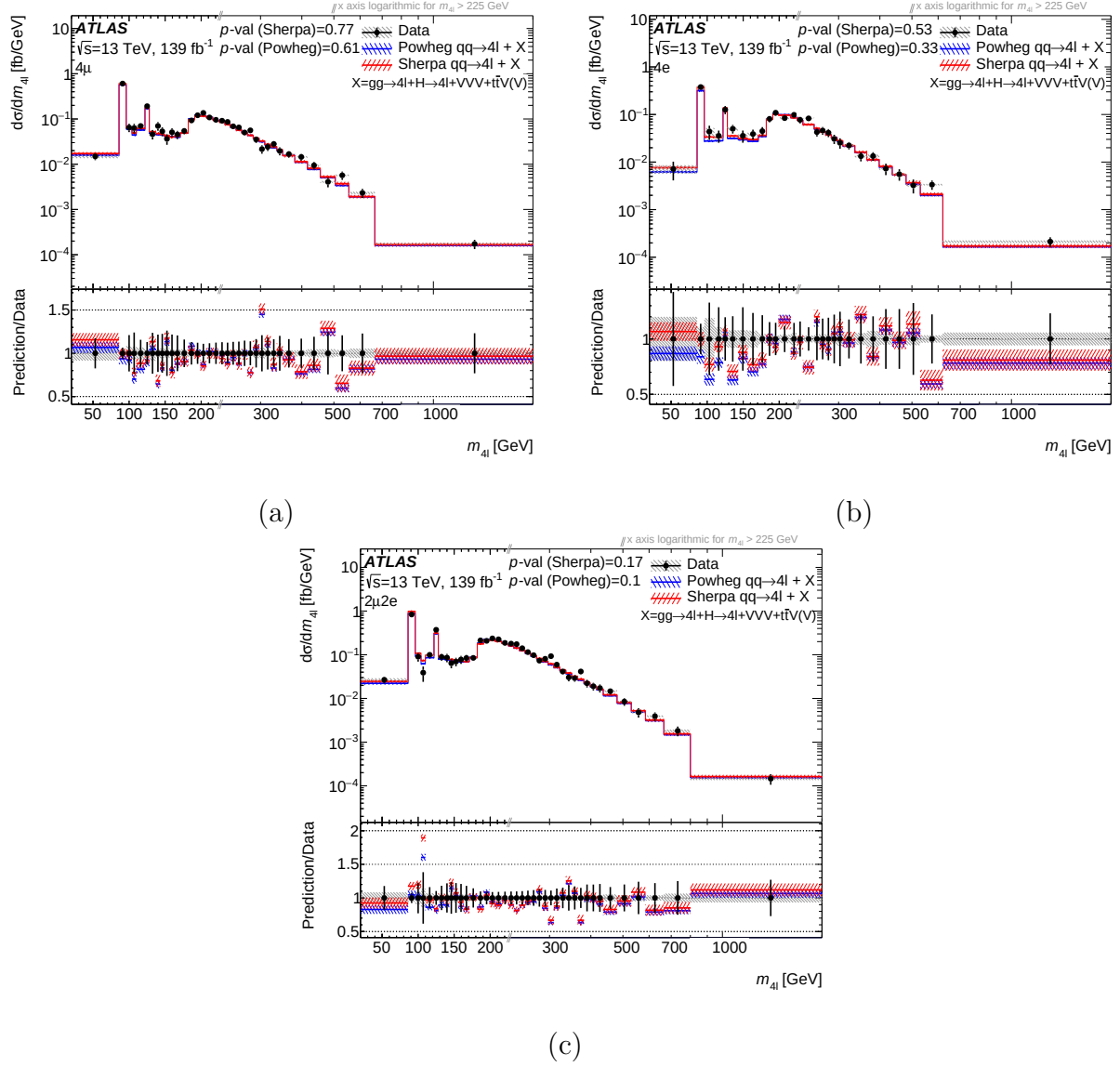
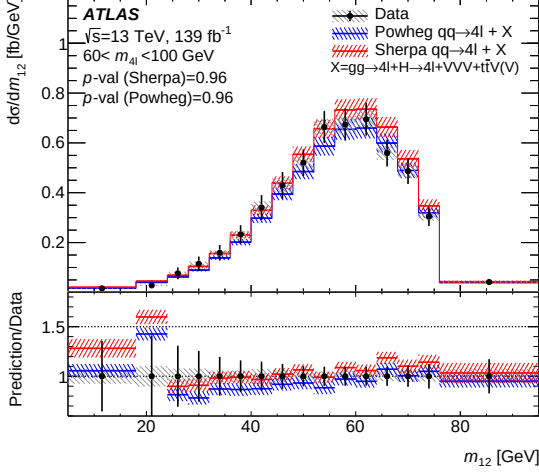
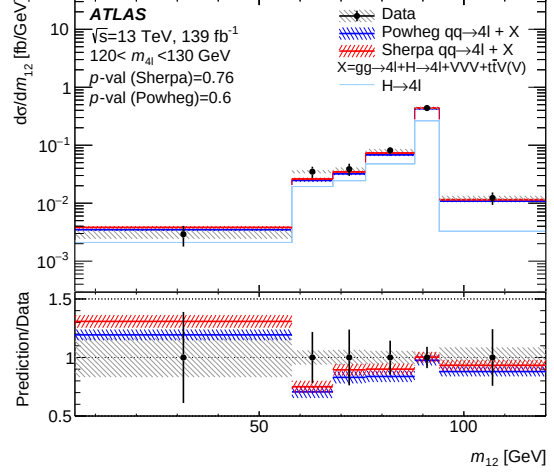


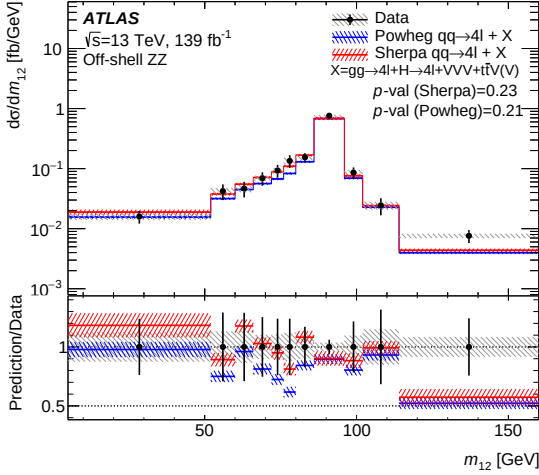
Figure 74: The unfolded differential $m_{4\ell}$ cross-section, split by quadruplet flavor channel, compared against full Standard Model predictions using either SHERPA or POWHEG+PYTHIA8 to model the dominant $q\bar{q} \rightarrow 4\ell$ contribution. Shaded bands denote the systematic uncertainties on the unfolded data and the predicted fiducial cross-sections, while the error bars on the data points show the combined statistical and systematic uncertainties.



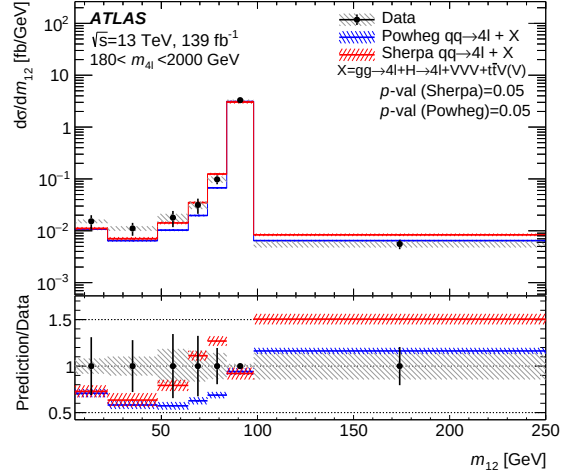
(a)



(b)



(c)



(d)

Figure 75: The unfolded double-differential m_{Z_1} - $m_{4\ell}$ cross-section, compared against full Standard Model predictions using either SHERPA or POWHEG+PYTHIA8 to model the dominant $q\bar{q} \rightarrow 4\ell$ contribution. Shaded bands denote the systematic uncertainties on the unfolded data and the predicted fiducial cross-sections, while the error bars on the data points show the combined statistical and systematic uncertainties.

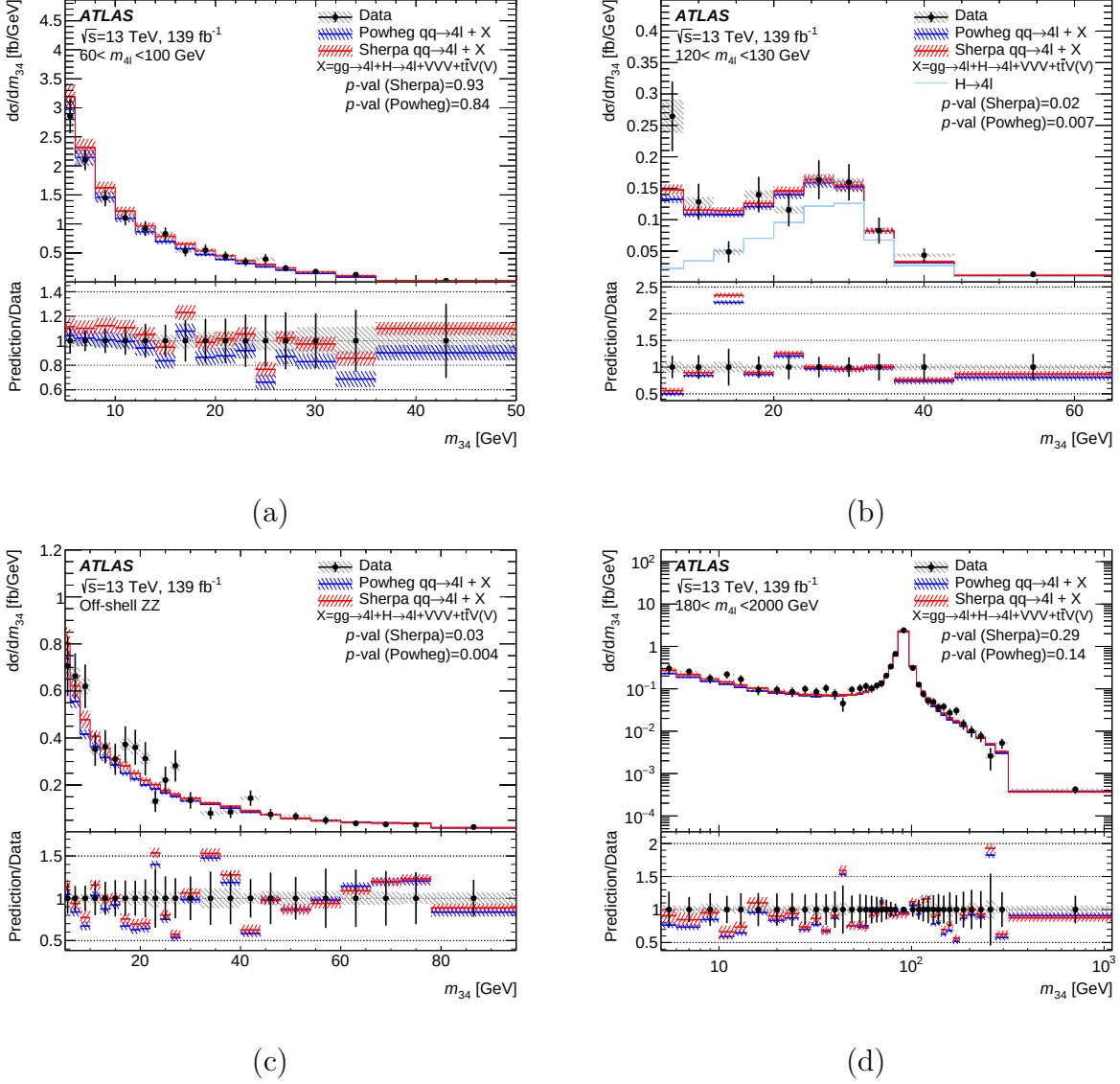
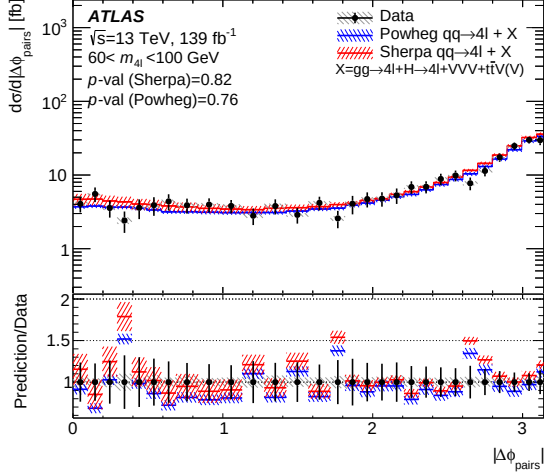
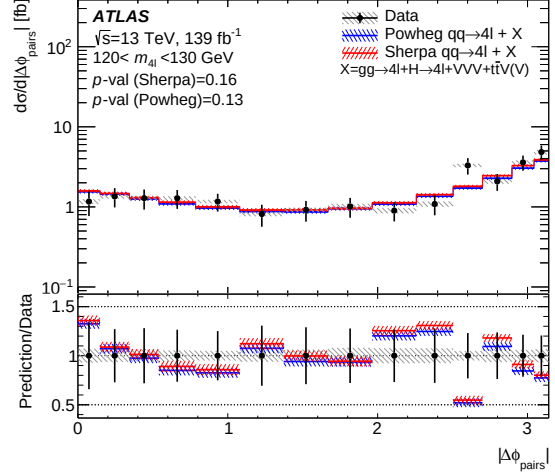


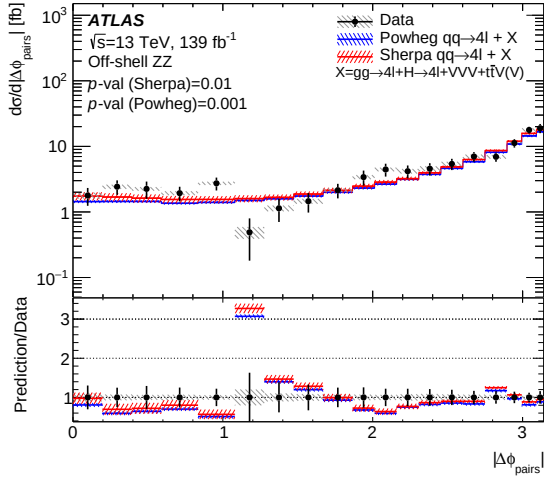
Figure 76: The unfolded double-differential m_{Z_2} - $m_{4\ell}$ cross-section, compared against full Standard Model predictions using either SHERPA or POWHEG+PYTHIA8 to model the dominant $q\bar{q} \rightarrow 4\ell$ contribution. Shaded bands denote the systematic uncertainties on the unfolded data and the predicted fiducial cross-sections, while the error bars on the data points show the combined statistical and systematic uncertainties.



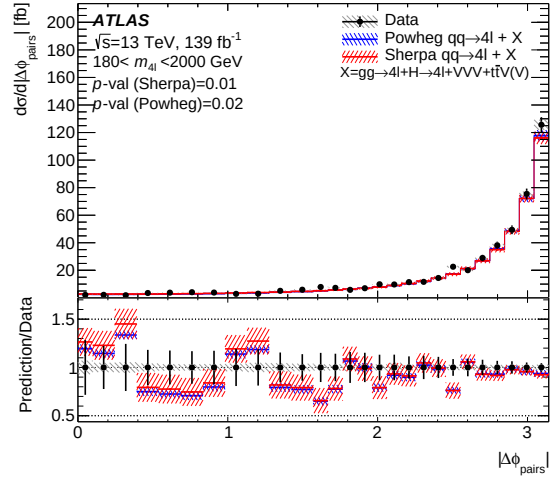
(a)



(b)



(c)



(d)

Figure 77: The unfolded double-differential $|\Delta\phi_{\text{pairs}}|-m_{4\ell}$ cross-section, compared against full Standard Model predictions using either SHERPA or POWHEG+PYTHIA8 to model the dominant $q\bar{q} \rightarrow 4\ell$ contribution. Shaded bands denote the systematic uncertainties on the unfolded data and the predicted fiducial cross-sections, while the error bars on the data points show the combined statistical and systematic uncertainties.

21 $Z \rightarrow 4\ell$ Branching Fraction

Measuring the branching fraction of the $Z \rightarrow 4\ell$ process is an excellent precision test of the Standard Model, in particular because of the relatively low theoretical uncertainties on the prediction. Near the Z resonance, $pp \rightarrow 4\ell$ production is dominated by the s -channel process shown in Figure 78 (a), where internal conversion of a virtual Z or γ FSR emission produces a lepton pair in addition to the primary $Z \rightarrow \ell\ell$ decay. Irreducible backgrounds, including the t -channel process shown in Figure 78 (b), contribute only a minor source in this region [75], allowing for a precise extraction of the resonant cross-section. Although benefitting from low backgrounds and a clean final state, the $Z \rightarrow 4\ell$ decay is an extremely rare process - the ALEPH experiment of LEP, which reported first observation in 1991, found only 24 four-lepton candidates out of 200,000 Z decays while running near the resonance [76]. The LHC Run-2 dataset, containing over one hundred times the total integrated luminosity delivered by LEP [22], is therefore crucial in obtaining a precision measurement of this process.

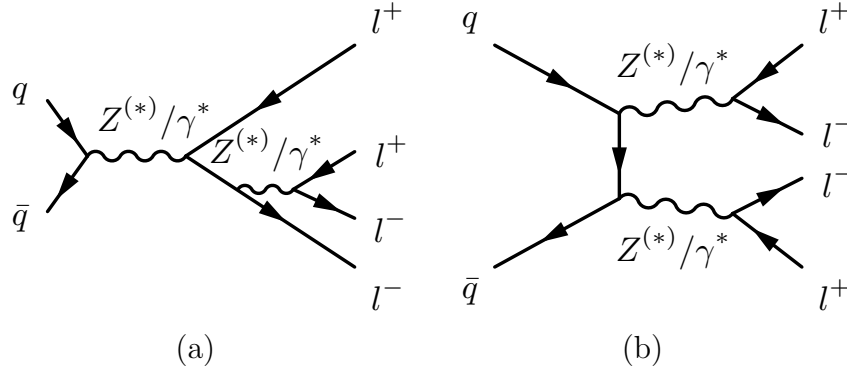


Figure 78: The s -channel $Z \rightarrow 4\ell$ decay (a), which dominates near the Z resonance, alongside the t -channel mode (b), the latter contributing a minor source of irreducible background.

The measurement is performed using the inclusive fiducial cross-section as measured in the single resonant Z region, defined by $60 \text{ GeV} < m_{4\ell} < 100 \text{ GeV}$, quoted in Table 16. In order to be directly comparable to previous ATLAS and CMS results, the measurement is extrapolated to an extended phase space defined by an m_Z -centered four-lepton invariant mass window of $80 \text{ GeV} < m_{4\ell} < 100 \text{ GeV}$ and the requirement that the dilepton invariant

mass of all six possible lepton pair combinations satisfies $m_{\ell\ell} > 4$ GeV. Additionally, this extended phase space excludes leptons originating from τ -decays.

In order to compute the $Z \rightarrow 4\ell$ ($\ell = e, \mu$) branching fraction, the cross-section measurement must be normalized to the total cross-section for single- Z production, σ_Z . This is derived from the ATLAS measurement of the $Z \rightarrow \ell\ell$ cross-section within $66 \text{ GeV} < m_Z < 116 \text{ GeV}$ using 81 pb^{-1} of early 2015 data [77] and the theoretical value of the branching fraction for the same process. A transfer factor is included to align the measurement with the Z -mass window as defined in the extended phase space,

$$\sigma_Z = \frac{\sigma_{Z \rightarrow \ell\ell}^{66-116}}{\mathcal{B}_{Z \rightarrow \ell\ell}} \times T_{66-116}^{80-100}. \quad (81)$$

This transfer factor is itself derived from two previously published m_Z window correction factors: $T_{60-120}^{80-100} = 0.926 \pm 0.001$ as used in Ref. [78] and $T_{66-116}^{60-120} = 1.01$ as used reciprocally in Ref. [79]. The factor used here is therefore calculated by

$$T_{66-116}^{80-100} = T_{60-120}^{80-100} \times T_{66-116}^{60-120}, \quad (82)$$

which gives a value of $T_{66-116}^{80-100} = 0.935 \pm 0.001$.

The predicted contribution of non- $q\bar{q} \rightarrow 4\ell$ processes, a source of irreducible background for this measurement, is subtracted from the measured fiducial cross-section. This cross-section is calculated using the signal MC samples and found to be $\sigma_{\text{non-}q\bar{q} \rightarrow 4\ell}^{\text{pred}} = 0.22 \pm 0.04 \text{ fb}$, including the same sources of uncertainty as discussed in Section 18. The small contribution from non-resonant t -channel $q\bar{q} \rightarrow 4\ell$ production, another source of irreducible background, is removed by multiplying by the predicted fractional composition of resonant production. This is calculated in the extended phase space using the POWHEG-BOX v2 generator, yielding a value of $f_{\text{res.-}q\bar{q} \rightarrow 4\ell} = 0.952 \pm 0.005$. Next, the contribution of leptons from τ -decays, excluded in the extended phase space, are accounted for via a multiplicative factor corresponding to the predicted fraction of non- τ events in the region. This is calculated using the $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$

MC to be $f_{\text{non-}\tau} = 0.9919 \pm 0.0001$, including only MC statistical uncertainties.

Lastly, the extrapolation to the extended phase space is accomplished by normalizing to the fiducial acceptance, defined as the ratio of the fiducial selection with respect to the loose selections of the extended phase space. Leptons originating from τ -decays are excluded in the numerator and denominator because the contribution thereof is already removed using the $f_{\text{non-}\tau}$ correction factor. This acceptance is measured using an extended $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$ MC sample and found to be $A_{\text{fid}} = 0.085 \pm 0.002$, again including the same sources of uncertainty discussed in Section 18.

Including all inputs, the formula used to compute the final measurement is given by

$$\mathcal{B}_{Z \rightarrow 4\ell} = \frac{(\sigma^{\text{meas}} - \sigma_{\text{non-}q\bar{q} \rightarrow 4\ell}^{\text{pred}}) \times f_{\text{res.-}q\bar{q} \rightarrow 4\ell} \times f_{\text{non-}\tau} \times \mathcal{B}_{Z \rightarrow \ell\ell}}{A_{\text{fid}} \times \sigma_{Z \rightarrow \ell\ell}^{66-116} \times T_{66-116}^{80-100}}. \quad (83)$$

The uncertainties on the individual inputs are accounted for by repeating the measurement for each variation and taking the relative difference with respect to the nominal measurement as the corresponding uncertainty. All of the individual uncertainties are treated conservatively to be uncorrelated, including the detector and luminosity uncertainties on the $Z \rightarrow \ell\ell$ cross-section. A detailed breakdown of the inputs, with their associated values and uncertainties, is shown below in Table 17.

Inputting the values outlined in Table 17 to Equation 83 gives a branching fraction measurement of $\mathcal{B}_{Z \rightarrow 4\ell} = (4.41 \pm 0.30) \times 10^{-6}$, with an uncertainty breakdown of

$$\mathcal{B}_{Z \rightarrow 4\ell} = (4.41 \pm 0.13(\text{stat}) \pm 0.23(\text{syst}) \pm 0.08(\text{theo}) \pm 0.12(\text{lumi})) \times 10^{-6}. \quad (84)$$

This is in agreement with previous ATLAS and CMS measurements [80], [75], [78], and [44], which are outlined in Table 18, and is consistent with the Standard Model prediction of $\mathcal{B}_{Z \rightarrow 4\ell}^{\text{pred}} = (4.50 \pm 0.01) \times 10^{-6}$ derived with POWHEG - as taken from Ref. [75]. Following the procedure of Ref. [81], an alternative theoretical prediction is derived by simulating 100,000 $Z \rightarrow 4\ell$ events with MADGRAPH5 [72] and normalizing by the program's default Z decay

width,

$$\frac{\Gamma_{Z \rightarrow 4\ell}}{\Gamma_Z} = \frac{11.48 \pm 0.01 \text{ keV}}{2.441 \text{ GeV}} = (4.70 \pm 0.01) \times 10^{-6}. \quad (85)$$

This, too, is in agreement with the latest measured value. The measurements and predictions discussed here are summarized together in Figure 79. This analysis provides the most precise measurement to date, benefitting from the large dataset and increase in acceptance over previous analyses. The overall error on the measurement of $\pm 6.8\%$ is dominated by sources of systematic uncertainty ($\pm 5.2\%$), followed by statistical ($\pm 2.9\%$), luminosity ($\pm 2.7\%$), and theory ($\pm 1.8\%$) uncertainties. The leading systematic uncertainties affecting this measurement are summarized in the $Z \rightarrow 4\ell$ column of Table 14.

Table 17: Summary of the inputs to Equation 83 used in the calculation of the $Z \rightarrow 4\ell$ branching fraction. The contributing sources of uncertainty are shown below each value in order of their respective size. All measured and predicted values derived in this analysis are obtained within the inclusive single resonant Z region, defined by $60 \text{ GeV} < m_{4\ell} < 100 \text{ GeV}$.

Symbol	Definition	Value
σ^{meas}	Measured fiducial cross-section ATLAS Run-2 139 fb^{-1} [49]	$22.13 \pm 1.29 \text{ fb}$ $\pm 1.06 \text{ fb}$ (syst) $\pm 0.64 \text{ fb}$ (stat) $\pm 0.36 \text{ fb}$ (lumi)
$\sigma_{\text{non-}q\bar{q} \rightarrow 4\ell}^{\text{pred}}$	Predicted cross-section of non- $q\bar{q} \rightarrow 4\ell$ sources SHERPA	$0.22 \pm 0.04 \text{ fb}$ $\pm 0.040 \text{ fb}$ (theo) $\pm 0.004 \text{ fb}$ (lumi) $\pm 0.001 \text{ fb}$ (stat)
$f_{\text{res.-}q\bar{q} \rightarrow 4\ell}$	Predicted fraction of resonant $q\bar{q} \rightarrow 4\ell$ production POWHEG-BOX v2	0.952 ± 0.005 ± 0.005 (theo)
$f_{\text{non-}\tau}$	Predicted fraction of events excluding τ -decays SHERPA $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$	0.9919 ± 0.0001 ± 0.0001 (stat)
A_{fid}	Predicted fiducial acceptance, excluding τ -decays, with respect to the extended phase space SHERPA $q\bar{q} \rightarrow Z^{(*)}Z^{(*)}$	0.085 ± 0.002 ± 0.0014 (scale) ± 0.0006 (pdf) ± 0.0001 (stat) ± 0.0001 (nlo)
$\mathcal{B}_{Z \rightarrow \ell\ell}$	Predicted branching fraction of $Z \rightarrow \ell\ell$ decay	0.03366 ± 0.00002 ± 0.00002 (theo)
$\sigma_{Z \rightarrow \ell\ell}^{66-116}$	Measurement of $Z \rightarrow \ell\ell$ cross-section ATLAS 81 pb^{-1} [77]	$1.98 \pm 0.06 \text{ nb}$ $\pm 0.042 \text{ nb}$ (lumi) $\pm 0.039 \text{ nb}$ (syst)
T_{66-116}^{80-100}	Transfer factor accounting for m_Z window POWHEG	0.935 ± 0.001 ± 0.001 (theo)

Table 18: Comparison of recent measurements of the $Z \rightarrow 4\ell$ ($\ell = e, \mu$) branching fraction in the extended phase defined by $80 \text{ GeV} < m_{4\ell} < 100 \text{ GeV}$ and $m_{\ell\ell} > 4 \text{ GeV}$. Center-of-mass energies and integrated luminosities are shown for each measurement. The last column highlights the total relative error on the measurement. The result of this analysis is shown in the final row.

Measurement	$\mathcal{B}_{Z \rightarrow 4\ell} [10^{-6}]$					Precision
	Stat.	Syst.	Theo.	Lumi.		
CMS, $\sqrt{s} = 7 \text{ TeV}$ [80] $L = 5.02 \text{ fb}^{-1}$	4.2	$^{+0.9}_{-0.8}$	± 0.2	-	-	$^{+22.0\%}_{-19.6\%}$
ATLAS, $\sqrt{s} = 7, 8 \text{ TeV}$ [75] $L = 4.5, 20.3 \text{ fb}^{-1}$	4.31	± 0.34	± 0.17	-	-	$\pm 8.8\%$
CMS, $\sqrt{s} = 13 \text{ TeV}$ [78] $L = 35.9 \text{ fb}^{-1}$	4.83	$^{+0.23}_{-0.22}$	$^{+0.32}_{-0.29}$	± 0.08	± 0.12	$^{+8.7\%}_{-8.1\%}$
ATLAS, $\sqrt{s} = 13 \text{ TeV}$ [44] $L = 36.1 \text{ fb}^{-1}$	4.70	± 0.32	± 0.21	-	± 0.14	$\pm 8.7\%$
ATLAS, $\sqrt{s} = 13 \text{ TeV}$ [49] $L = 139 \text{ fb}^{-1}$	4.41	± 0.13	± 0.23	± 0.08	± 0.12	$\pm 6.8\%$

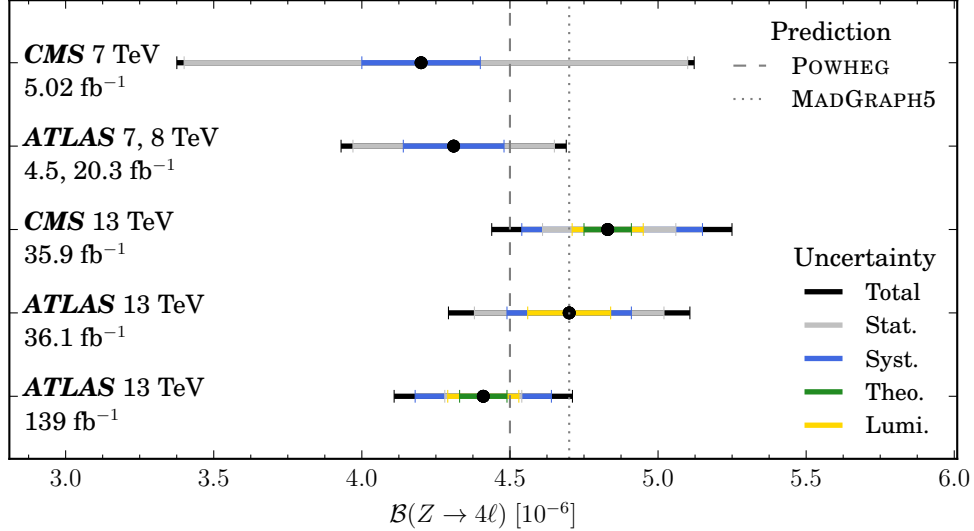


Figure 79: Comparison of recent measurements of the $Z \rightarrow 4\ell$ ($\ell = e, \mu$) branching fraction in the extended phase defined by $80 \text{ GeV} < m_{4\ell} < 100 \text{ GeV}$ and $m_{\ell\ell} > 4 \text{ GeV}$. The result of this analysis is shown by the lowermost entry. The predictions obtained from POWHEG and MADGRAPH5 are shown as vertical lines.

22 Interpretations

Given the inclusivity of the analysis and the wide range of contributing processes, there is serious potential for placing constraints on new physics which either directly or indirectly contribute to the four-lepton final state. Two scenarios of beyond-the-Standard Model physics are considered and constrained using a common profile likelihood statistical framework. The possibility of further interpretations using the available measurements is discussed in the next section.

Statistical Framework

The profile likelihood technique [82] is a commonly used approach for constraining the parameter space of models, relying on a test statistic which is defined in terms of the ratio of likelihood functions²². For the purposes of this analysis, the likelihood function for a particular distribution of k bins is defined by

$$\mathcal{L} = \frac{1}{\sqrt{(2\pi)^k |\mathbf{C}|}} \exp \left\{ -\frac{1}{2} \left[\vec{\sigma}^{\text{meas}} - \vec{\sigma}^{\text{pred}}(\vec{\theta}) \right]^T \mathbf{C}^{-1} \left[\vec{\sigma}^{\text{meas}} - \vec{\sigma}^{\text{pred}}(\vec{\theta}) \right] \right\} \times \prod_{\theta_i \in \vec{\theta}} \mathcal{G}(\theta_i, 0, 1), \quad (86)$$

where $\vec{\sigma}^{\text{meas}}$ is the measured unfolded cross-section, $\vec{\sigma}^{\text{pred}}$ the model prediction, $\mathbf{C}$ the total uncertainty covariance matrix, and $\vec{\theta}$ the nuisance parameters of the model. As for the evaluation of p -values in Section 20, the statistical and systematic covariances are taken from the SM expectation instead of the observed data to account for bins where downward fluctuations in the data below the Gaussian limit lead to an underestimation of the statistical uncertainty. The theoretical uncertainties of the model are not included in the covariance matrix, instead treated as Gaussian-constrained nuisance parameters with a mean of zero and standard deviation of one, denoted by $\mathcal{G}(\theta_i, 0, 1)$. The term in the exponent, ignoring the multiplicative factor, is equivalent to the χ^2 between the observed data and the prediction

²² A likelihood is a conditional probability $P(x|\theta)$, encoding the probability of observing the data x under assumption of the model θ .

evaluated at a specific point in the nuisance parameter space, allowing for the simplification,

$$\mathcal{L} = \frac{1}{\sqrt{(2\pi)^k |\mathbf{C}|}} \exp \left\{ -\frac{1}{2} \chi(\vec{\theta})^2 \right\} \times \prod_{\theta_i \in \vec{\theta}} \mathcal{G}(\theta_i, 0, 1). \quad (87)$$

Letting the single parameter c uniquely specify the model of interest, the test statistic is defined as the natural logarithm of the ratio of likelihoods,

$$q(c) = -2 \ln \frac{\arg\max_{\vec{\theta}} \mathcal{L}(c, \vec{\theta})}{\arg\max_{c', \vec{\theta}'} \mathcal{L}(c', \vec{\theta}')} := -2 \ln \frac{\mathcal{L}(c, \hat{\vec{\theta}}(c))}{\mathcal{L}(\hat{c}, \hat{\vec{\theta}})}. \quad (88)$$

The numerator in Equation 88 represents the likelihood maximized in the space of $\vec{\theta}$, *conditioned* on the value of c . In contrast, the denominator represents the *unconditional* maximum likelihood, in which the values of c and $\vec{\theta}$ are allowed to simultaneously float in order to globally maximize the likelihood. Following the definition of the likelihood in Equation 87, this reduces to a summation of two terms: the $\Delta\chi^2$ between numerator and denominator and a sum in quadrature over the nuisance parameter terms,

$$q(c) = \Delta\chi(c)^2 + \Delta \sum_i \Theta^2. \quad (89)$$

The test statistic is itself a statistical variable drawn from some underlying distribution. Based on the definition of Equation 88, low values of the test statistic close to zero correspond to a good agreement between the hypothesized model and observed data, while higher values indicate an increasing incompatibility of the model in describing the data. Constraints in the parameter space are derived based upon a desired p -value exclusion contour α , equivalent to a confidence level (CL) of $1 - \alpha$. Following Wilk's theorem [83], the distribution of the test statistic is expected to asymptotically approach the χ_ν^2 distribution with $\nu = 1$ degree of freedom. In this case, limits may be easily derived using the $\chi_{\nu=1}^2$ -PDF (here *probability density function*) and inputting the desired statistical significance (e.g. $q = 3.84$ for 95%

CL). If, however, the asymptotic condition is not assumed to be satisfied, the distribution of the test statistic may instead be estimated by throwing Monte Carlo toys and limits may be then derived using the toy-based PDF.

Instead of simultaneously fitting the thirteen differential measurements outlined in Section 14, which requires precise knowledge of the correlations, each observable is considered separately in Equation 88. The observable providing the most stringent limits for a given model is then used to derive the quoted results. This varies depending on the particular new physics scenario, highlighting the advantage of measuring several complementary observables. Expected limits are evaluated in addition to the observed limits by using the Asimov dataset²³ in place of the measured data.

22.1 Standard Model Effective Field Theory

Formalism

The Standard Model Effective Field Theory (SMEFT) provides an approximate description of new physics beyond a cutoff scale Λ in terms of the SM Lagrangian and a set of non-renormalizable operators. The operators describing new physics interactions are built out of the SM fields in higher dimensions of energy ($d > 4$) and obey the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry structure of the low-energy, dimension-four theory [84]. The effective Lagrangian is written as a polynomial expansion in the inverse of the energy cutoff scale Λ ,

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda} \sum_i C_i^{(5)} \mathcal{O}_i^{(5)} + \frac{1}{\Lambda^2} \sum_i C_i^{(6)} \mathcal{O}_i^{(6)} + \dots, \quad (90)$$

where each successive term is suppressed by an additional inverse factor of Λ . The first term in the expansion represents the SM Lagrangian, while the remaining terms are parameter-

²³ The Asimov dataset is simply taken as the SM prediction with Poisson errors derived from the weighted event counts. The name derives from Isaac Asimov's novel *Franchise*, in which the most representative person out of an entire population of people is chosen as the sole voter in an election.

ized by dimension- d operators, $\mathcal{O}_i^{(d)}$ with dimensionless coupling constants $C_i^{(d)}$, known as Wilson coefficients, determining their relative strength. The multiplicative factors of Λ are generally absorbed into the coupling constants $C_i^{(d)}$, defining *effective* coupling constants $c_i^{(d)} = C_i^{(d)}/\Lambda^{d-4}$. The inclusion of higher-dimensional operators to the SM Lagrangian leads to two important consequences: firstly, new vertices forbidden in the dimension-four theory arise, and secondly, the usual vertices and parameters (e.g. masses, Weinberg mixing angle) of the SM are modified depending on the strength of the EFT contributions [85]. As described in Ref. [85], effective field theories are useful in two general cases: firstly, to study the low-energy behavior of a given ultraviolet theory and, secondly, as a model-independent means to parameterize experimental deviations from SM predictions. The latter use case, of more relevance to this analysis, provides a means to systematically map the landscape of potential new physics contributions where deviations from SM predictions are observed [85].

With each dimension term included in the SMEFT expansion, the number of operators grows given the increase in possible field combinations. At dimension-five, there is only one operator, which violates lepton and baryon number conservation [85], and is therefore neglected in this analysis. Higher-dimension operators beyond $d = 6$ are neglected based on their additional suppression in energy. The new physics component constrained in this analysis is therefore described entirely in terms of dimension-six operators, and the energy cutoff scale for new physics is taken to be 1 TeV. The operators are described in terms of SM fields before electroweak symmetry breaking within the Warsaw basis [86]. A subset of those 59 terms which do not violate baryon number conservation are considered and, based on a screening study, 22 operators offering particular sensitivity in the four-lepton final state are constrained in the interpretations.

The 22 Wilson coefficients of relevance to this analysis, along with their corresponding operator expansion in terms of SM fields, are summarized in Table 19. The coefficients are separated into three broad categories depending on the type of SM fields composing the operators, as follows:

Table 19: Summary of the 22 dimension-six Wilson coefficients considered in this analysis. The first column denotes the general category of SM fields included in the corresponding operators, where X represents an arbitrary gauge field and (L, R) denote fermion chirality states. Following the convention of Ref. [86], the operators are written in terms of SM fields before electroweak symmetry breaking. The gauge fields are written as $G_{\mu\nu}^A$, $W_{\mu\nu}^I$, and $B^{\mu\nu}$, while ϕ denotes the scalar Higgs field. The left-handed quark and lepton doublets are represented by q and l , and the right-handed quarks and leptons by u, d , and e , respectively. The subscripts p, r, s , and t denote flavor indices.

Category		Coefficient	Operator
Bosonic	$\phi^2 X^2$	c_{HG}	$\phi^\dagger \phi G_{\mu\nu}^A G^{A\mu\nu}$
		$\tilde{c}_{HG}$	$\phi^\dagger \phi \tilde{G}_{\mu\nu}^A G^{A\mu\nu}$
		c_{HWB}	$\phi^\dagger \tau^I \phi W_{\mu\nu}^I B^{\mu\nu}$
	$\phi^4 D^2$	c_{HD}	$(\phi^\dagger D^\mu \phi)^* (\phi^\dagger D_\mu \phi)$
Mixed	$\phi^3 \psi^2$	c_{He}	$(\phi^\dagger \phi)(\bar{l}_p e_r \phi)$
		c_{Hu}	$(\phi^\dagger \phi)(\bar{q}_p u_r \tilde{\phi})$
		c_{Hd}	$(\phi^\dagger \phi)(\bar{q}_p d_r \phi)$
	$\phi^2 \psi^2 D$	$c_{Hl}^{(1)}$	$(\phi^\dagger i \overleftrightarrow{D}_\mu \phi)(\bar{l}_p \gamma^\mu l_r)$
		$c_{Hl}^{(3)}$	$(\phi^\dagger i \overleftrightarrow{D}_\mu^I \phi)(\bar{l}_p \tau^I \gamma^\mu l_r)$
		$c_{Hq}^{(1)}$	$(\phi^\dagger i \overleftrightarrow{D}_\mu \phi)(\bar{q}_p \gamma^\mu q_r)$
		$c_{Hq}^{(3)}$	$(\phi^\dagger i \overleftrightarrow{D}_\mu^I \phi)(\bar{q}_p \tau^I \gamma^\mu q_r)$
Fermionic	$(\bar{L}L)(\bar{L}L)$	c_{ll}	$(\bar{l}_p \gamma_\mu l_r)(\bar{l}_s \gamma^\mu l_t)$
		$c_{ll}^{(1)}$	$(\bar{l}_p \gamma_\mu l_t)(\bar{l}_r \gamma^\mu l_s)$
		$c_{lq}^{(1)}$	$(\bar{l}_p \gamma_\mu l_r)(\bar{q}_s \gamma^\mu q_t)$
		$c_{lq}^{(3)}$	$(\bar{l}_p \gamma_\mu \tau^I l_r)(\bar{q}_s \gamma^\mu \tau^I q_t)$
	$(\bar{R}R)(\bar{R}R)$	c_{ee}	$(\bar{e}_p \gamma_\mu e_r)(\bar{e}_s \gamma^\mu e_t)$
		c_{eu}	$(\bar{e}_p \gamma_\mu e_r)(\bar{u}_s \gamma^\mu u_t)$
		c_{ed}	$(\bar{e}_p \gamma_\mu e_r)(\bar{d}_s \gamma^\mu d_t)$
	$(\bar{L}L)(\bar{R}R)$	c_{le}	$(\bar{l}_p \gamma_\mu l_r)(\bar{e}_s \gamma^\mu e_t)$
		c_{lu}	$(\bar{l}_p \gamma_\mu l_r)(\bar{u}_s \gamma^\mu u_t)$
		c_{ld}	$(\bar{l}_p \gamma_\mu l_r)(\bar{d}_s \gamma^\mu d_t)$
		c_{qe}	$(\bar{q}_p \gamma_\mu q_r)(\bar{e}_s \gamma^\mu e_t)$

- **Bosonic** operators consist purely of boson fields, namely the Higgs and gauge bosons.
- **Mixed** operators consist of combinations of boson and fermion fields.
- **Fermionic** operators consist purely of fermion fields and describe four-fermion contact terms - including direct four-lepton and lepton-quark interactions.

The bosonic operators affect a host of Higgs interactions, as well as those mediated by the exchange of gauge bosons. The mixed operators primarily enter by means of their effect on the $Z \rightarrow \ell\ell$ vertex. The four-fermion operators clearly give rise to contact terms not present in the SM, though the effect of other individual operators is generally more nuanced. For instance, the pure-bosonic c_{HWB} coefficient directly affects triple and quartic gauge vertices of the SM, but also enters in the Z -boson field normalization, thereby indirectly affecting a range of additional vertices [85]. Hence, the individual effect of each operator is often more complicated than the modification or addition of single vertices [85]. A detailed discussion of the remaining dimension-six operators in terms of their constituent Standard Model fields is provided in Ref. [86].

For the purposes of this analysis, only one dimension-six operator is considered at a time, thereby reducing the EFT parameterization to a single effective coupling constant²⁴. This eliminates cross-terms in the matrix element expansion which depend on the product of two different EFT couplings. Within this framework, the matrix element for a given physics process including the SM and BSM contributions at leading order in the perturbative expansion is given by

$$\mathcal{M}_{\text{SMEFT}} = \mathcal{M}_{\text{SM}} + c \cdot \mathcal{M}_{\text{BSM}}, \quad (91)$$

the square of which equals

$$|\mathcal{M}_{\text{SMEFT}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + 2c \cdot \mathcal{R}(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{BSM}}) + c^2 \cdot |\mathcal{M}_{\text{BSM}}|^2. \quad (92)$$

²⁴From hereon, the abbreviation c refers to an arbitrary dimension-six effective coupling constant.

The predicted cross-section, given by the square of the matrix element, can therefore be decomposed into three components: SM-only, SM-BSM interference, and BSM-only. The interference (or *linear*) term depends linearly on the effective coupling constant c and is therefore the leading contribution of new physics at order Λ^{-2} in the expansion. The BSM-only (or *quadratic*) term depends quadratically on c , therefore contributing at a higher order of Λ^{-4} . Notably, these quadratic terms are at the same order of energy suppression as dimension-eight linear terms, the latter of which are neglected in this analysis. Given the additional suppression of the quadratic term with respect to the linear term, limits are derived for two separate models, either including (*full model*) or excluding (*linear-only model*) the quadratic term. Studying both models allows for the importance of dimension-eight linear terms to be inferred in cases where the inclusion of the quadratic term in the full model significantly improves the observed limits as derived in the linear model. The predicted SMEFT cross-section in the i^{th} bin is evaluated under the full model by

$$\sigma_i^{\text{pred}} = \sigma_i^{\text{SM}}(1 + c \cdot A_i + c^2 \cdot B_i) \quad (93)$$

and the linear-only model by

$$\sigma_i^{\text{pred}} = \sigma_i^{\text{SM}}(1 + c \cdot A_i), \quad (94)$$

where A_i and B_i correspond to the linear and quadratic EFT contributions in bin i relative to the SM cross-section σ_i^{SM} in the same bin.

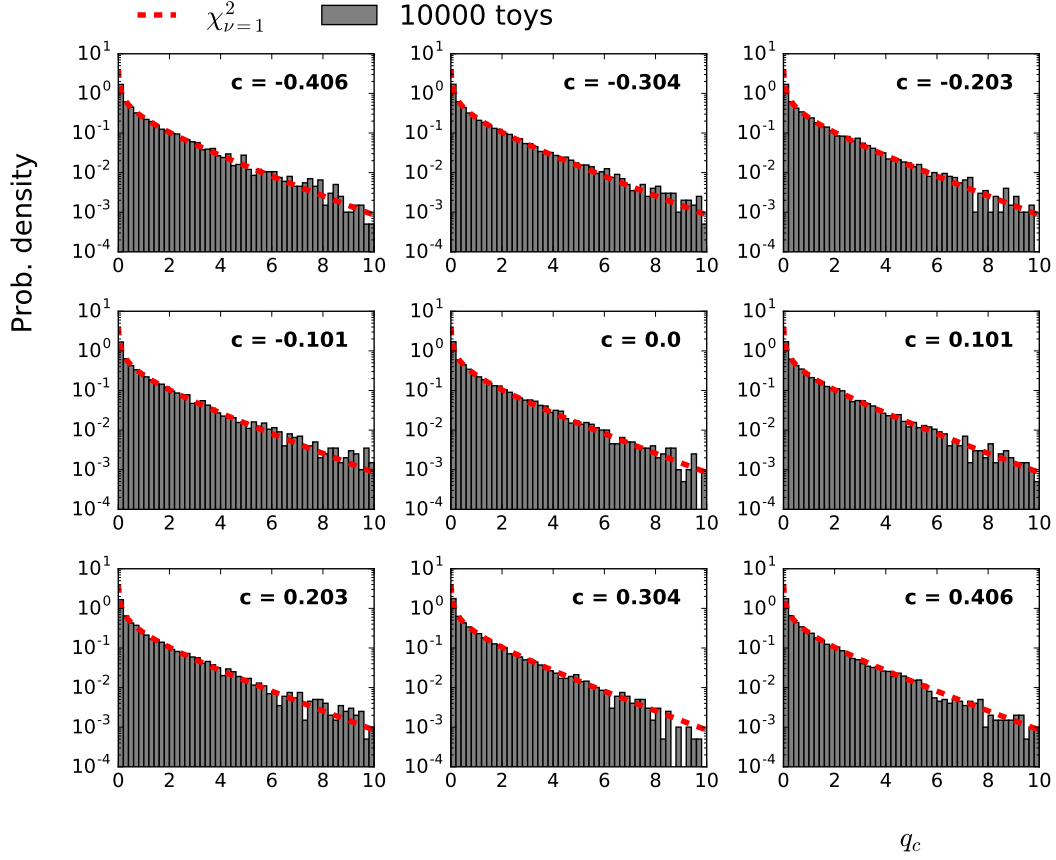
Strategy

Effective field theory samples are generated using the SMEFTsim [84,87] implementation of FEYNRULES [88], running MADGRAPH5_aMC@NLO2.6.5 + PYTHIA8.243 at LO accuracy in QCD with the NNPDF3.0nlo PDF set. Since the signal samples of this analysis are generated beyond LO accuracy, a k -factor is applied on the LO SMEFT cross-sections defined as $k = \bar{\sigma}^{\text{NLO SM}} / \bar{\sigma}^{\text{LO SM}}$, where the numerator is taken from the signal samples discussed

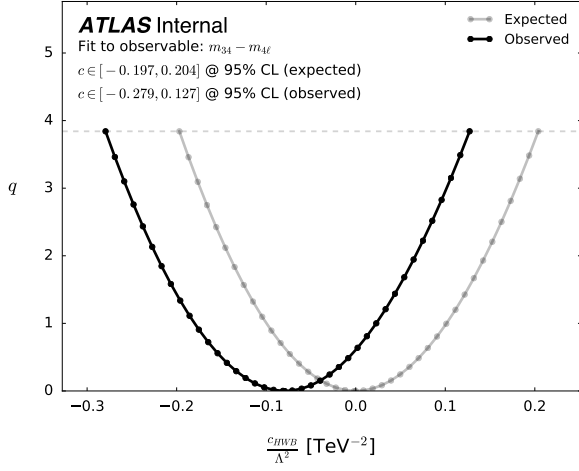
in Section 13.2 and the denominator from the SMEFT setup discussed above. This correction implicitly assumes the higher order effects are the same for the SM and new physics components. Uncertainties accounting for QCD scale variations, PDF reweightings, and α_S variations are evaluated following the same strategy as outlined in Section 18.

Using the cross-section decomposition in either Equation 93 or 94, it is sufficient to generate only the SM prediction and the SMEFT prediction at two arbitrary scan points to produce the prediction for any given value of c . This allows for a non-resource intensive means to perform one-dimensional likelihood scans versus a single parameter of interest. At each scan point, the compatibility of the model with the data is evaluated based upon the test statistic outlined in Equation 88. The uncertainties on the SM prediction, encoded in the covariance matrix of Equation 86, are taken to be fixed at the SM value $c = 0$. Lower and upper limits on the coefficient of interest are derived at 95% CL based on the test statistic distribution. The postulate of Wilk's theorem - that this distribution will asymptotically approach a $\chi^2_{\nu=1}$ -PDF - was observed not to hold in all cases, particularly in the presence of dominant quadratic terms or large EFT enhancements. Therefore, the distribution of the test statistic is instead evaluated using Monte Carlo toys thrown around the SM prediction.

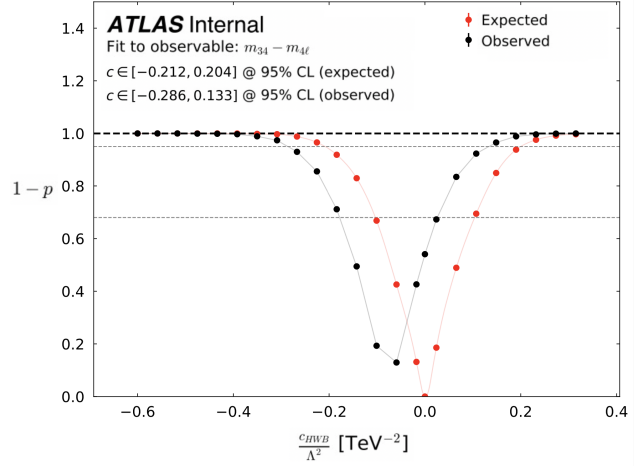
An overview of the fitting procedure is shown in Figure 80, where the bosonic c_{HWB} coefficient is constrained under the full model using the observable offering the best sensitivity, $m_{Z_2}-m_{4\ell}$. In Figure 80 (a), the distributions of 10,000 Monte Carlo toys are compared against the $\chi^2_{\nu=1}$ -PDF for nine scan points of c_{HWB} . Over all scan values, the distributions of test statistics are well-modeled by the $\chi^2_{\nu=1}$ -PDF, reflecting that in this specific case Wilk's theorem is satisfied. In Figure 80 (b), the expected and observed limits at 95% CL are then derived under assumption of Wilk's theorem, by evaluating the $q = 3.84$ crossing points. In Figure 80 (c), limits are instead derived using the $1 - p = 95\%$ crossing of toy distributions. Given the observed validity of Wilk's theorem demonstrated in Figure 80 (a), these two sets of limits are largely comparable. However, as this does not generally hold in other cases, the latter, toy-based approach is used to derive limits for all coefficients.



(a)



(b)



(c)

Figure 80: Overview of the full model likelihood scan for c_{HWB} , as constrained by the m_{Z_2} - $m_{4\ell}$ observable. In (a), the distributions of the test statistic drawn from Monte Carlo toys are compared against the $\chi^2_{\nu=1}$ -PDF. In (b), the expected and observed limits are derived under the asymptotic assumption of Wilk's theorem. In (c), limits are instead derived - as for the final results - based on the distributions of Monte Carlo toys.

Results

The expected and observed 95% confidence limits on the 22 dimension-six Wilson coefficients are shown separately for the full and linear-only EFT models in Figures 81 and 82, respectively. For detailed tables see Ref. [49]. The $\tilde{c}_{HG}$ coefficient has a vanishing interference term and is therefore only constrained in the full model, leaving 21 coefficients constrained under the linear-only model.

The $m_{4\ell}$, $|\Delta\phi_{\text{pairs}}|-m_{4\ell}$, and, particularly, $m_{Z_2}-m_{4\ell}$ measurements provide the majority of the constraining power. The observed limits are generally compatible with the expected limits under both decomposition models. Seven coefficients (c_{HD} , c_{HWB} , c_{He} , $c_{Hl}^{(1)}$, $c_{Hq}^{(3)}$, $c_{Hl}^{(3)}$, $c_{ll}^{(1)}$) are constrained to a similar degree regardless of whether or not the quadratic term is included, suggesting its contribution is negligible. This partially explains the validity of the asymptotic approximation observed for the c_{HWB} test statistic distributions in Figure 80. Two coefficients ($c_{Hq}^{(1)}$, $c_{lq}^{(3)}$) have slightly negative-shifted limits when including the quadratic term, indicating that its contribution is relatively small and, necessarily, positive. The presence of double intervals in two coefficients (c_{HG} , c_{Hu}) under the full model is indicative of non-negligible quadratic terms, which contribute secondary maxima in the likelihood scans due to cancellation with the interference terms. The remaining eleven coefficients ($\tilde{c}_{HG}$, c_{Hd} , c_{ed} , c_{ee} , c_{eu} , c_{ld} , c_{le} , c_{ll} , $c_{lq}^{(1)}$, c_{lu} , c_{qe}) are also observed to have non-negligible quadratic contributions, leading to more stringent limits when including the higher order term. Following the previous discussion, this suggests the neglected dimension-eight linear terms may also play an important role for these coefficients.

The coefficients affecting the $Z \rightarrow \ell\ell$ vertex have already been constrained to a much greater degree than reported here from a global SMEFT fit to LEP and LHC data [89], benefiting significantly from the precision electroweak data provided by LEP. Other coefficients, however, show improvements over recent results published by the ATLAS $H \rightarrow 4\ell$ cross-section analysis [90]. In particular, the limits on c_{HWB} are significantly improved with respect to the Higgs measurement [90].

ATLAS

$\sqrt{s}=13 \text{ TeV}, 139 \text{ fb}^{-1}$

full model

Expected 95% CL

Observed 95% CL

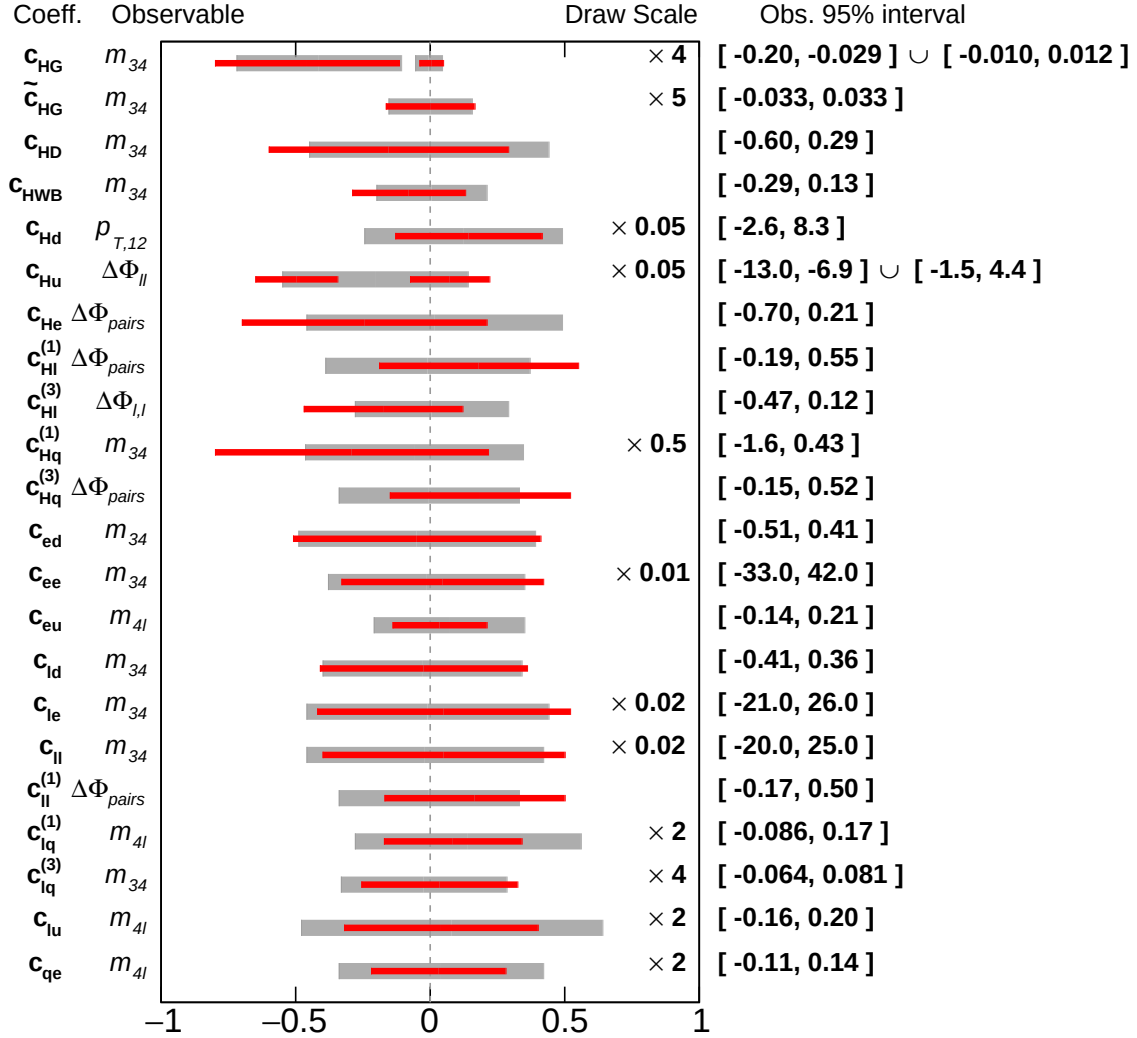


Figure 81: Expected and observed 95% confidence limits on the 22 dimension-six Wilson coefficients constrained in this analysis under the full decomposition model (see Equation 93). The observable column denotes the observable offering the most sensitivity to the particular coefficient, which is used to derive the limits. Multiplicative factors on the right denote the relative scaling of the limits for the purposes of maintaining readability.

ATLAS

$\sqrt{s}=13$ TeV, 139 fb⁻¹

only linear term

Expected 95% CL

Observed 95% CL

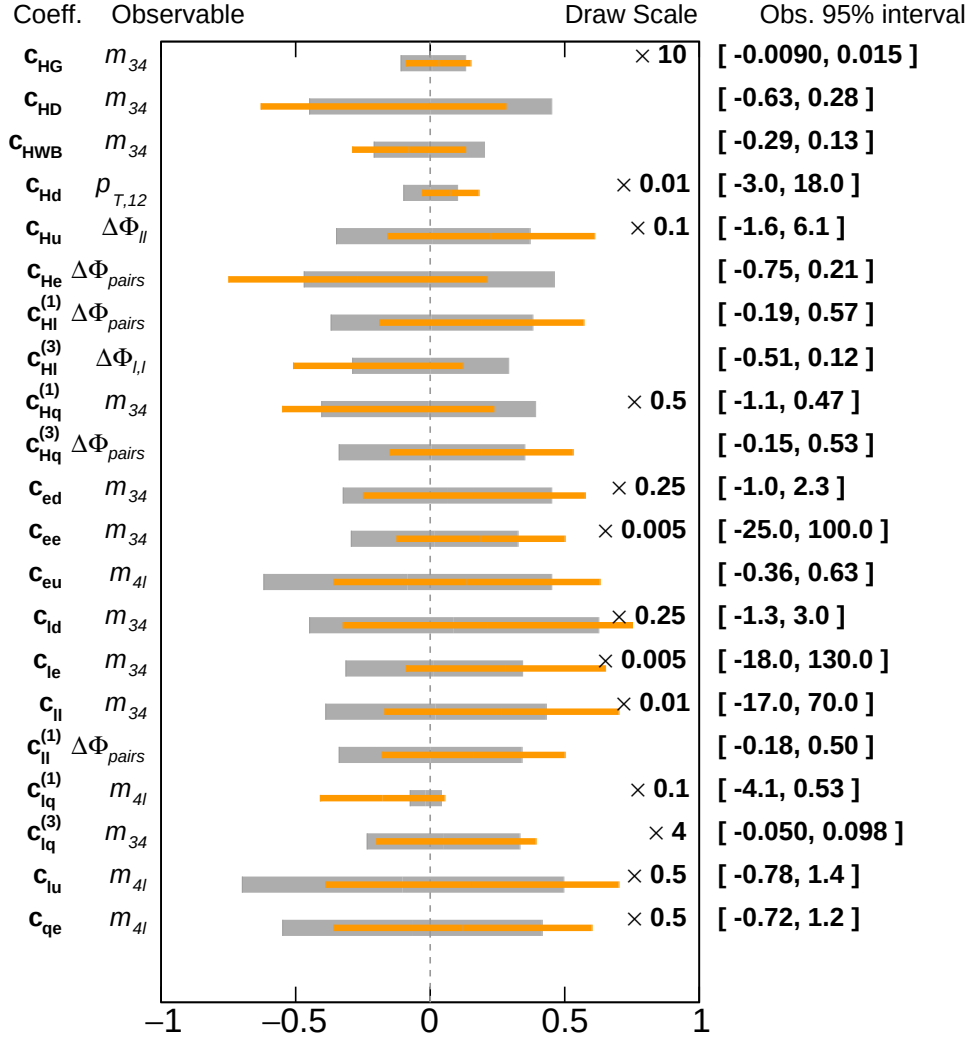


Figure 82: Expected and observed 95% confidence limits on the 21 dimension-six Wilson coefficients constrained in this analysis under the linear-only decomposition model (see Equation 94). The observable column denotes the observable offering the most sensitivity to the particular coefficient, which is used to derive the limits. Multiplicative factors on the right denote the relative scaling of the limits for the purposes of maintaining readability.

22.2 $B - L$ Gauge Symmetry Model

Formalism

The $B - L$ model [91] is a well-motivated new physics scenario which introduces a spontaneously broken global baryon-number minus lepton-number $U(1)_{B-L}$ gauge symmetry to solve the neutrino mass problem in the Standard Model. This model postulates a heavy Z' gauge boson of the new local symmetry, as well as an exotic Higgs h_2 from the symmetry breaking. The hypothetical Higgs particle couples to the SM Higgs with a mixing angle α . Both Higgs particles can decay into a pair of Z or Z' bosons, which themselves can decay leptonically to produce the four-lepton signature. The four-lepton phase space can therefore provide sensitivity in the region where the mass of the exotic Higgs exceeds twice the mass of the Z' -boson.

Strategy

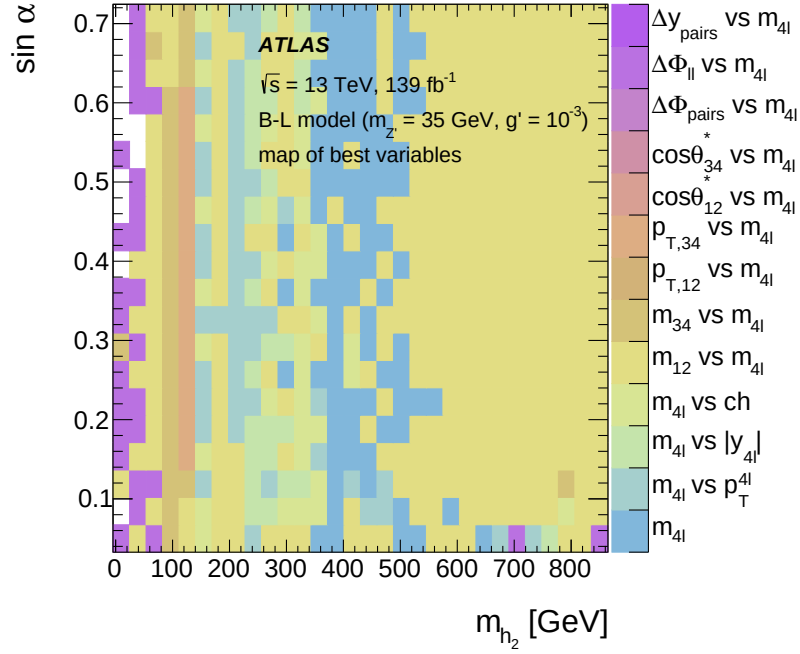
New physics samples are generated using HERWIG 7.2 [92], excluding interference effects with the SM which are expected to be small. To compare with previous LHC measurements [93], the model is constrained to a low-mass Z' ($m_{Z'} = 35$ GeV), which is weakly coupled to the SM ($g' = 10^{-3}$), and exclusion contours are derived in the plane of $\sin \alpha$ versus m_{h_2} . The constraint of $m_{Z'} = 35$ GeV necessarily removes any sensitivity in the four-lepton final state in the region $m_{h_2} < 70$ GeV.

Constraining the model relies on the same likelihood function defined in Equation 86, but differs in the precise definition of the test statistic. As is common in searches [39, 94], the model parameter μ determines the global signal strength of the new physics, such that $\mu = 0$ corresponds to the SM-only hypothesis and $\mu = 1$ to the nominal signal hypothesis. In order to account for downward fluctuations in the data below the SM-only hypothesis, where $\hat{\mu} < 0$, a piecewise function defines the modified test statistic $\tilde{q}$, shown in Ref. [49]. The CL_S [95] method is used to derive the 95% CL exclusion contours in the plane of $\sin \alpha$

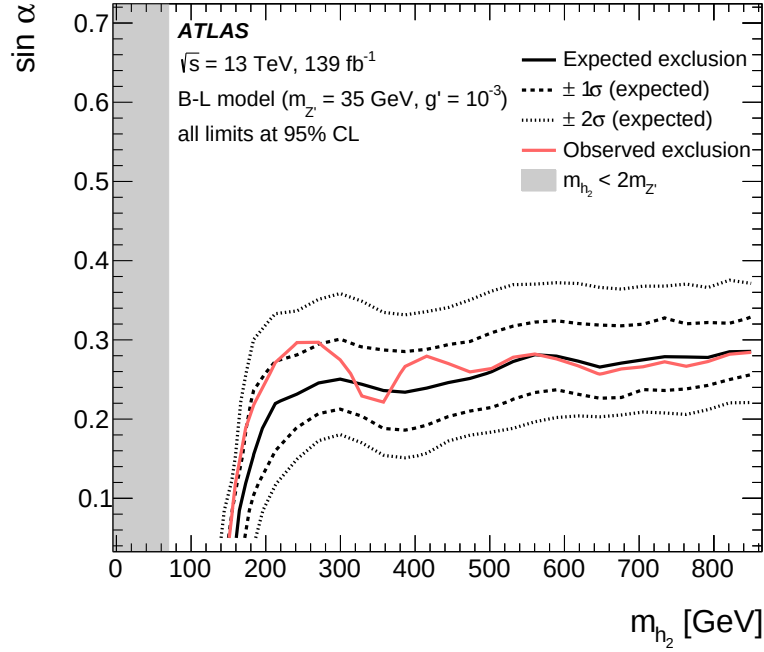
versus m_{h_2} . Unlike the EFT interpretation, the asymptotic approximation of Wilk’s theorem is observed to be valid here, and is therefore used instead of a toy-based approach.

Results

The results are shown in Figure 83, where (a) shows the observable offering the best sensitivity at each scan point, which is used to derive the expected and observed contours shown in (b). The observed limits are compatible with the expected limits within uncertainties. As expected, the dilepton invariant mass observables, in particular $m_{Z_1-m_{4\ell}}$, provide good sensitivity across much of the phase space given the presence of non-SM Z' -bosons. The limits presented here significantly improve upon the previous results of [93], extending the exclusion in $\sin \alpha$ from 0.4 to 0.28 across most of the plane and constraining higher masses than before, beyond $m_{h_2} > 600$ GeV.



(a)



(b)

Figure 83: Constraints on the new physics scenario based on the $B-L$ model, in the plane of $\sin \alpha$ versus m_{h_2} . In (a), a color map of which unfolded observable gives the best fit sensitivity at each point in the scan is shown. In (b), the expected and observed 95% confidence limits are derived with uncertainty bands, based on the most sensitive observable.

23 Re-interpretations

The complexity and effort behind physics analyses motivate maximal use of the measurements in the search for new physics beyond-the-Standard Model. While searches provide an important means to test particular models, precision measurements benefit from being re-interpretable in light of many models. The widespread availability of generic analysis frameworks (e.g. RIVET [96]) and online data repositories provide an excellent means of rapidly testing measurements against a variety of new physics models.

To this end, the particle-level cross-sections and uncertainty correlations measured in this analysis have been made publicly available on HEPData²⁵ to allow for further re-interpretations of the data. Already at the time of writing, the measurements have been used to constrain composite dark matter models in Ref. [97], where they offer particular sensitivity in the 200 GeV mass range of hypothetical dark pions. Further re-interpretations remain promising given the variety of measured observables and the many potentially contributing processes.

²⁵<https://www.hepdata.net/record/ins1849535>

Chapter VII: Conclusion

The four-lepton (4ℓ) final state, characterized by the clean signature of two same-flavor, opposite-charge lepton pairs, offers an excellent avenue for precision tests of the Standard Model at high energy particle colliders. Using the large Run-2 dataset of the Large Hadron Collider as recorded by the ATLAS detector, inclusive 4ℓ production cross-sections have been measured to an unprecedented granularity. State-of-the-art Standard Model predictions are found to be in excellent agreement with the detector-corrected measurements. A world-leading precision measurement of the rare $Z \rightarrow 4\ell$ branching fraction has been measured to be $(4.41 \pm 0.30) \times 10^{-6}$, in agreement with previous measurements and leading theoretical predictions. The complementary 4ℓ cross-section measurements allow for competitive constraints to be placed on two scenarios of new physics - using either a model-independent effective field theory framework or a specific $B - L$ gauge symmetry model - and provide an excellent foundation for further re-interpretations and beyond-the-Standard Model searches.

References

- [1] R. Oerter. *The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics*. Penguin Publishing Group, 2006.
- [2] Albert Einstein. On the electrodynamics of moving bodies. *Annalen Phys.*, 17:891–921, 1905.
- [3] G. Gamow. *Thirty Years that Shook Physics: The Story of Quantum Theory*. Anchor Books, 1966.
- [4] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.
- [5] Mark Thomson. *Modern particle physics*. Cambridge University Press, Cambridge, 2013.
- [6] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [7] A. Salam and J.C. Ward. Electromagnetic and weak interactions. *Physics Letters*, 13(2):168–171, 1964.
- [8] S L Glashow. Partial symmetries of weak interactions. *Nuclear Phys.*, 22, 2 1961.
- [9] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [10] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.
- [12] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse. Cavity control of a single-electron quantum cyclotron: Measuring the electron magnetic moment. *Physical Review A*, 83(5), May 2011.
- [13] B. et al. Abi. Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm. *Physical Review Letters*, 126(14), Apr 2021.
- [14] LHCb Collaboration. Test of lepton universality in beauty-quark decays. Technical report, CERN, Geneva, Mar 2021.
- [15] D. R. Green, P. Meade, and M.A. Pleier. Multiboson interactions at the LHC. *Reviews of Modern Physics*, 89(3), Sep 2017.
- [16] Paul Langacker. The physics of heavy Z' gauge bosons. *Reviews of Modern Physics*, 81(3):1199–1228, Aug 2009.
- [17] J. A. Aguilar-Saavedra, J. A. Casas, J. Quilis, and R. Ruiz de Austri. Multilepton dark matter signals. *Journal of High Energy Physics*, 2020(4), Apr 2020.

- [18] ATLAS Collaboration. Search for supersymmetry in events with four or more leptons in $\sqrt{s} = 13$ TeV pp collisions with ATLAS. *Physical Review D*, 98(3), Aug 2018.
- [19] K. S. Babu and Sudip Jana. Probing doubly charged Higgs bosons at the LHC through photon initiated processes. *Phys. Rev. D*, 95:055020, Mar 2017.
- [20] Jacob J. Ethier, Giacomo Magni, Fabio Maltoni, Luca Mantani, Emanuele R. Nocera, Juan Rojo, Emma Slade, Eleni Vryonidou, and Cen Zhang. Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC, 2021.
- [21] Armin Hermann, Lanfranco Belloni, Gerhard John Krige, Ulrike Mersits, and Dominique Pestre. *History of CERN*. North-Holland, Amsterdam, 1987.
- [22] R W Assmann. LEP Operation and Performance with Electron-Positron Collisions at 209 GeV. 2001.
- [23] CERN. LHC Guide. Brochure, Mar 2017.
- [24] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3, 2008.
- [25] Joao Pequeno. Computer generated image of the whole ATLAS detector. General Photo, Mar 2008.
- [26] Yosuke Takubo. ATLAS IBL operational experience. Technical Report ATL-INDET-PROC-2016-012, CERN, Geneva, Nov 2016.
- [27] ATLAS Collaboration. ATLAS Run-2 status and performance. Technical Report ATL-GEN-PROC-2015-001, CERN, Geneva, Sep 2015.
- [28] ATLAS Collaboration. Alignment of the ATLAS Inner Detector in Run-2. *Eur. Phys. J. C*, 80(arXiv:2007.07624. 12):1194. 41 p, Jul 2020.
- [29] ATLAS Collaboration. DAQ approved plots. Available at <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ>.
- [30] S. Agostinelli et al. GEANT4 – a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [31] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444–450, 1987.
- [32] Nikiforos Nikiforou. Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade. Technical Report arXiv:1306.6756, CERN, Jun 2013.
- [33] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77(CERN-PH-EP-2015-304):490. 87 p, Mar 2016.

- [34] Joao Pequeno. Event Cross Section in a computer generated image of the ATLAS detector. General Photo, Mar 2008.
- [35] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. Technical Report ATLAS-CONF-2016-024, CERN, Geneva, Jun 2016.
- [36] Sebastien Rettie. Muon identification and performance in the ATLAS experiment. Technical report, CERN, Geneva, Jun 2018.
- [37] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [38] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77:466, 2017.
- [39] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716, 2012. <https://arxiv.org/abs/1207.7214>.
- [40] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, Sep 2012.
- [41] ATLAS Collaboration. Expected tracking and related performance with the updated ATLAS Inner Tracker layout at the High-Luminosity LHC. Technical report, CERN, Geneva, Jul 2021. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-024>.
- [42] ATLAS Collaboration. Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC. Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC. *Phys. Rev. Lett.*, 117:182002. 19 p, Jun 2016.
- [43] ATLAS Collaboration. $ZZ \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector. $ZZ \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector. *Phys. Rev. D*, 97:032005. 39 p, Sep 2017.
- [44] ATLAS Collaboration. Measurement of the four-lepton invariant mass spectrum in 13 TeV proton-proton collisions with the ATLAS detector. *Journal of High Energy Physics*, 2019(4):48, Apr 2019.
- [45] ATLAS Collaboration. Standard Model Summary Plots March 2021. Technical Report ATL-PHYS-PUB-2021-005, CERN, Geneva, Mar 2021.
- [46] G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13(07):P07017, 2018.

- [47] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical Report ATLAS-CONF-2019-021, CERN, Geneva, Jun 2019.
- [48] ATLAS Collaboration. Multi-Boson Simulation for 13 TeV ATLAS Analyses. Technical Report ATL-PHYS-PUB-2017-005, CERN, Geneva, May 2017.
- [49] ATLAS Collaboration. Measurements of differential cross-sections in four-lepton events in 13 TeV proton-proton collisions with the ATLAS detector. *Journal of High Energy Physics*, 2021(7):5, July 2021.
- [50] Enrico Bothmann, Gurpreet Singh Chahal, Stefan Höche, Johannes Krause, Frank Krauss, Silvan Kuttimalai, Sebastian Liebschner, Davide Napoletano, Marek Schönherr, Holger Schulz, and et al. Event generation with Sherpa 2.2. *SciPost Physics*, 7(3), Sep 2019.
- [51] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, and et al. Parton distributions for the LHC run II. *Journal of High Energy Physics*, 2015(4), Apr 2015.
- [52] Davison E. Soper. Parton distribution functions. *Nuclear Physics B - Proceedings Supplements*, 53(1-3):69–80, Feb 1997.
- [53] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6), Jun 2010.
- [54] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.P. Yuan. New parton distributions for collider physics. *Physical Review D*, 82(7), Oct 2010.
- [55] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, Jun 2008.
- [56] Keith Hamilton, Paolo Nason, Emanuele Re, and Giulia Zanderighi. NNLOPS simulation of Higgs boson production. *Journal of High Energy Physics*, 2013(10), Oct 2013.
- [57] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys. G*, 43:023001, 2016.
- [58] ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. Technical report, CERN, Geneva, Nov 2014.
- [59] Yanyan Gao, Andrei V Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V Tran. Spin determination of single-produced resonances at hadron colliders. Technical Report arXiv:1001.3396, Jan 2010.

- [60] Fabio Cardillo, Otilia Anamaria Ducu, Julien Maurer, Claudia Merlassino, Joseph Reichert, Erich Varnes, Pepijn Johannes Bakker, and Marcus Morgenstern. Tools for estimating fake/non-prompt lepton backgrounds in ATLAS. Technical report, CERN, Geneva, Sep 1999.
- [61] J H Friedman. A variable span smoother. Technical Report SLAC-PUB-3477. STAN-LCS-005, SLAC, Stanford, CA, Oct 1984.
- [62] G. D’Agostini. A multidimensional unfolding method based on Bayes’ theorem. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 362(2):487 – 498, 1995.
- [63] Enrico Bothmann, Marek Schönherr, and Steffen Schumann. Reweighting QCD matrix-element and parton-shower calculations. *The European Physical Journal C*, 76(11), Oct 2016.
- [64] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *The European Physical Journal C*, 75(5), May 2015.
- [65] Sayipjamal Dulat, Tie-Jiun Hou, Jun Gao, Marco Guzzi, Joey Huston, Pavel Nadolsky, Jon Pumplin, Carl Schmidt, Daniel Stump, and C.-P. Yuan. New parton distribution functions from a global analysis of quantum chromodynamics. *Physical Review D*, 93(3), Feb 2016.
- [66] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *The European Physical Journal C*, 63(2):189–285, Jul 2009.
- [67] Benedikt Biedermann, Ansgar Denner, Stefan Dittmaier, Lars Hofer, and Barbara Jäger. Next-to-leading-order electroweak corrections to the production of four charged leptons at the LHC. *Journal of High Energy Physics*, 2017(1), Jan 2017.
- [68] Fabrizio Caola, Matthew Dowling, Kirill Melnikov, Raoul Röntsch, and Lorenzo Tancredi. QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC. *Journal of High Energy Physics*, 2016(7), Jul 2016.
- [69] ATLAS Collaboration. Measurement of the Higgs boson coupling properties in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2018(3), Mar 2018.
- [70] ATLAS Collaboration. Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 99:072009, Apr 2019.
- [71] ATLAS Collaboration. Evidence for the production of three massive vector bosons with the ATLAS detector. *Physics Letters B*, 798:134913, 2019.

- [72] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), Jul 2014.
- [73] Ifan G Hughes and Thomas P A Hase. *Measurements and their uncertainties: a practical guide to modern error analysis*. Oxford Univ. Press, New York, NY, 2010.
- [74] ATLAS Collaboration. Evaluating statistical uncertainties and correlations using the bootstrap method. Technical report, CERN, Geneva, Apr 2021. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-011>.
- [75] ATLAS Collaboration. Measurement of the 4ℓ Cross Section at the Z Resonance and Determination of the Branching Fraction of $Z4\ell$ in pp collisions at $\sqrt{s} = 7$ and 8 TeV with ATLAS. *Physical Review Letters*, 112(23), Jun 2014.
- [76] ALEPH Collaboration. Charged particle pair production associated with a lepton pair in Z decays: indication of an excess in the tau channel. *Phys. Lett. B*, 263:112–122. 27 p, Apr 1991.
- [77] ATLAS Collaboration. Measurement of $W^\pm$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ with the ATLAS detector. *Physics Letters B*, 759:601 – 621, 2016.
- [78] CMS Collaboration. Measurements of the $ppzz$ production cross section and the $Z4\ell$ branching fraction, and constraints on anomalous triple gauge couplings at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 78(2), Feb 2018.
- [79] CMS Collaboration. Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 7$ and 8 TeV and search for anomalous triple gauge couplings at $\sqrt{s} = 8$ TeV. *The European Physical Journal C*, 77(4), Apr 2017.
- [80] CMS Collaboration. Observation of Z decays to four leptons with the CMS detector at the LHC. *Journal of High Energy Physics*, 2012(12), Dec 2012.
- [81] J. Lovelace Rainbolt and Michael Schmitt. Branching fraction for Z decays to four leptons and constraints on new physics. *Physical Review D*, 99(1), Jan 2019.
- [82] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011.
- [83] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):60 – 62, 1938.
- [84] Ilaria Brivio, Yun Jiang, and Michael Trott. The SMEFTsim package, theory and tools. *Journal of High Energy Physics*, 2017(12), Dec 2017.

- [85] Raquel Gomez-Ambrosio. Studies of dimension-six EFT effects in vector boson scattering. *The European Physical Journal C*, 79(5):389, 2019.
- [86] B. Grzadkowski, M. Iskrzyński, M. Misiak, and J. Rosiek. Dimension-six terms in the Standard Model Lagrangian. *Journal of High Energy Physics*, 2010(10), Oct 2010.
- [87] Ilaria Brivio. SMEFTsim 3.0 — a practical guide. *Journal of High Energy Physics*, 2021(4), Apr 2021.
- [88] Adam Alloul, Neil D. Christensen, Céline Degrande, Claude Duhr, and Benjamin Fuks. FeynRules 2.0 — A complete toolbox for tree-level phenomenology. *Computer Physics Communications*, 185(8):2250–2300, Aug 2014.
- [89] John Ellis, Christopher W. Murphy, Verónica Sanz, and Tevong You. Updated global SMEFT fit to Higgs, diboson and electroweak data. *Journal of High Energy Physics*, 2018(6), Jun 2018.
- [90] ATLAS Collaboration. Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 80:957. 76 p, Apr 2020.
- [91] Frank F. Deppisch, Wei Liu, and Manimala Mitra. Long-lived heavy neutrinos from Higgs decays. *Journal of High Energy Physics*, 2018(8), Aug 2018.
- [92] Johannes Bellm, Stefan Gieseke, David Grellscheid, Simon Plätzer, Michael Rauch, Christian Reuschle, Peter Richardson, Peter Schichtel, Michael H. Seymour, Andrzej Siódmok, and et al. Herwig 7.0/Herwig++ 3.0 release note. *The European Physical Journal C*, 76(4), Apr 2016.
- [93] S. Amrith, J. M. Butterworth, F. F. Deppisch, W. Liu, A. Varma, and D. Yallup. LHC constraints on a $B - L$ gauge model using CONTUR. *Journal of High Energy Physics*, 2019(5), May 2019.
- [94] ATLAS Collaboration. Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s}=7$ TeV with the ATLAS detector. *Phys. Rev. D*, 86:032003, Aug 2012.
- [95] A.L. Read. Presentation of search results: the CL_S technique. *J. Phys. G*, 28(2693), 2002.
- [96] Christian Bierlich, Andy Buckley, Jonathan Butterworth, Christian Holm Christensen, Louie Corpe, David Grellscheid, Jan Fiete Grosse-Oetringhaus, Christian Gutsche, Przemyslaw Karczmarczyk, Jochen Klein, and et al. Robust Independent Validation of Experiment and Theory: Rivet version 3. *SciPost Physics*, 8(2), Feb 2020.
- [97] J. M. Butterworth, L. Corpe, X. Kong, S. Kulkarni, and M. Thomas. New sensitivity of LHC measurements to Composite Dark Matter Models, 2021.

Appendices

The following appendices are intended as supplementary material to the main body of this dissertation. In Appendix A, I outline first my specific contributions to the *Inclusive Four-lepton* analysis central to this thesis, and second my larger involvements within the ATLAS Collaboration. In the remaining appendices, I provide additional information on particle kinematics in collider settings (Appendix B), the loosening and optimization of selections (Appendix C), the background estimation strategy (Appendix D), and the unfolding data-driven closure test (Appendix E).

A Primary Contributions

Over the past five years of my doctoral program, I have conducted cutting-edge research in experimental high energy particle physics within the ATLAS Collaboration, contributing across the three major domains of the experiment – physics analysis, detector hardware, and reconstruction software. Though this thesis focuses on a particular physics analysis, my supplementary hardware and software experiences were essential in writing a coherent dissertation. In the following sections, I outline first my primary contributions to the analysis presented in this thesis, and second my larger involvements within the experiment.

A.1 Inclusive Four-lepton Analysis

My specific area of physics expertise in ATLAS is that of four-lepton processes. Over the previous three years, I have made substantial contributions to three physics analyses in the Standard Model Electroweak subgroup. The work of this thesis focuses entirely on the *Inclusive Four-lepton* analysis [49], to which I contributed from its initial conception in 2018 until its final publication in 2021. This dissertation would not be possible without the combined efforts of the entire analysis team, in particular **Max Goblirsch**, **Zara Grout**, and **Emily Nurse** for their essential leadership efforts as analysis contacts. My specific contributions in this analysis centered on preliminary sensitivity projections, selection and isolation working point optimization, lepton triggering, smoothing and validation of the background estimate, unfolding bias and uncertainty quantification, the $Z \rightarrow 4\ell$ branching fraction measurement, and statistical compatibility studies. In subsequent four-lepton analyses, I have led the efforts in background estimation techniques and effective field theory interpretations of new physics, while also taking a more prominent leadership role as the analysis framework maintainer, responsible for implementing and validating changes in the upstream selections and producing skimmed samples for common usage. Although the results of these ongoing analyses are as of yet unpublished and therefore left out of this thesis, my experiences in these

more recent endeavors have significantly shaped my discussion of the inclusive analysis.

A.2 ATLAS Collaboration

My involvement with ATLAS began in the summer of 2016, when I first moved to CERN and started my authorship qualification task in the Muon Combined Performance group. The task would blend performance and hardware work by studying the impacts of future detector upgrades on muon reconstruction in the expected environment of the upcoming High Luminosity Large Hadron Collider (HL-LHC) program. Through this role, I learned the fundamentals of muon reconstruction and performance in ATLAS and had my first encounter with the ATLAS Inner Tracker (ITk) tracking detector upgrade. The culmination of my work resulted in the development of a more pile-up robust muon track isolation working point for use in the HL-LHC, which exploited the superior impact parameter resolution of the ITk to veto likely pile-up tracks from isolation cones.

Following my author qualification and completion of my Brandeis advanced exam in the spring of 2018, I started my involvements in physics analysis by joining the post-publication efforts of the prior four-lepton analysis [44]. Fittingly, I investigated the potential of a follow-up analysis given the increased statistics of the full Run-2 dataset, including the effects of loosening kinematic selections and re-optimizing the binning - efforts which would pave the way for the subsequent four-lepton analysis forming the bulk of this thesis.

In the summer of 2018, I moved to Brookhaven National Laboratory (BNL) for a one-year residency supported by the US Department of Energy Office of Science Graduate Student Research fellowship, to focus on the design and construction of silicon-based tracking and timing detectors intended for use in the HL-LHC. During this period, I led the development of a semi-automated loading procedure for the precision mounting of silicon strip sensors on stave support structures and developed quality assurance standards for sensors on- and off-stave. In parallel, I conducted early-stage electrical characterization studies of Low Gain Avalanche Detectors to guide their design for use in a future timing detector - known as the

High Granularity Timing Detector. These projects were not only crucial in preparing the BNL site for the upcoming stages of production, but also in teaching me the fundamentals of detector development and operation.

After finishing my residency at BNL in the summer of 2019, I moved back to CERN to focus on analysis work for my dissertation. Around this time, I joined the *Inclusive Four-lepton* analysis, where I would go on to spend the following two years working towards the analysis publication and, subsequently, my dissertation writing and defense.

Beginning later in 2019, I resumed my prior efforts in ATLAS reconstruction software, now focused on generic charged particle tracking, leveraging my hands-on experience with silicon-based tracking detectors. I began contributing to the ongoing development of track reconstruction software for the future ATLAS detector of the HL-LHC era. Towards this goal, I carried out tracking performance studies of the ITk detector through its many layout updates, the most recent results of which are highlighted in Section 9.2. Later on, I shifted focus to the present tracking detector to join the efforts in reconstruction improvements in advance of Run-3, participating in a large-scale software overhaul which achieved algorithmic speed-ups beyond a factor of two and large reductions in disk usage. This work has given me a deep insight into the principles underlying reconstruction, in particular tracking, and its important connection to charged lepton measurements in ATLAS.

Midway through the inclusive analysis in the spring of 2020, I joined the *Differential ZZjj* analysis, with a goal of measuring differential vector boson scattering in the $4\ell jj$ final state. One year later in the spring of 2021, I also joined the *Angular ZZ* analysis, focused on measuring 4ℓ angular observables for their use in polarization and interpretation studies. Much of this thesis, though based on the results of the *Inclusive Four-lepton* analysis, relies heavily on my complementary experiences drawn from these subsequent analyses.

B Collider Kinematics

Following Equation 39, the transverse momentum p_T is defined as the momentum in the xy -plane. Therefore, the linear momentum can be uniquely specified by the transverse momentum, azimuthal angle ϕ , and polar angle θ . Equivalently, the polar angle can be replaced by the pseudorapidity η by inverting Equation 41. The linear momentum written in terms of (p_T, η, ϕ) is then

$$\begin{aligned} p_x &= p \sin \theta \cos \phi = p_T \cos \phi \\ p_y &= p \sin \theta \sin \phi = p_T \sin \phi \\ p_z &= p \cos \theta = p_T \cot \theta = p_T \sinh \eta. \end{aligned} \tag{95}$$

For a particle of mass m , the combination of (p_T, η, ϕ, m) provides a complete description of the four-momentum, equivalent to the standard (E, p_x, p_y, p_z) formulation,

$$(E, p_x, p_y, p_z) \rightarrow \left(\sqrt{m^2 + (1 + \sinh^2 \eta) p_T^2}, p_T \cos \phi, p_T \sin \phi, p_T \sinh \eta \right). \tag{96}$$

The latter is more convenient in collider settings since it is expressed in terms of quantities whose absolute values (i.e. p_T, m) or differences in (i.e. $\Delta\eta, \Delta\phi$) are Lorentz invariant under boosts along the longitudinal axis.

C Selection Optimization

With respect to the previous four-lepton analysis [44], many selections have been loosened to benefit from increased inclusivity and statistics. This section outlines the impact of loosening a subset of those selections, as well as other important changes in the reconstruction strategy.

C.1 Loosening of Selections

The reconstruction-level event selection as used in this analysis is outlined in Table 8, while the selection as used in the previous analysis can be found in Ref. [44]. Four selections in particular have been loosened in this analysis to increase statistics and improve sensitivity to certain models of new physics. The effects of individually (a,b,c,d) and simultaneously (e) reverting each of these selections to the baseline as used in previous analysis are shown in Figure 84. Allowing leptons from τ -decays, which were previously excluded in the former analysis, grants percent-level inclusive acceptance gains. Loosening the angular separation requirement to $\Delta R > 0.05$, previously 0.1 (0.2) for different-flavor (same-flavor) pairs, amounts to only a small percentage gain at low mass, but greatly improves sensitivity to boosted production scenarios. Loosening the secondary and dropping the tertiary lepton p_T selections yields a large acceptance gain, particularly for softer topologies at low four-lepton invariant mass. Lastly, dropping the dilepton invariant mass requirements, previously applied as a fixed selection on the leading pair and a variable selection on the subleading pair [44], massively increases the acceptance for off-shell ZZ topologies and greatly improves the sensitivity to decays of non-SM Z -bosons. The combined effect of dropping or loosening these four selections amounts to an overall increase in acceptance of approximately 37%, corresponding to roughly a factor two gain for $m_{4\ell} < 100$ GeV.

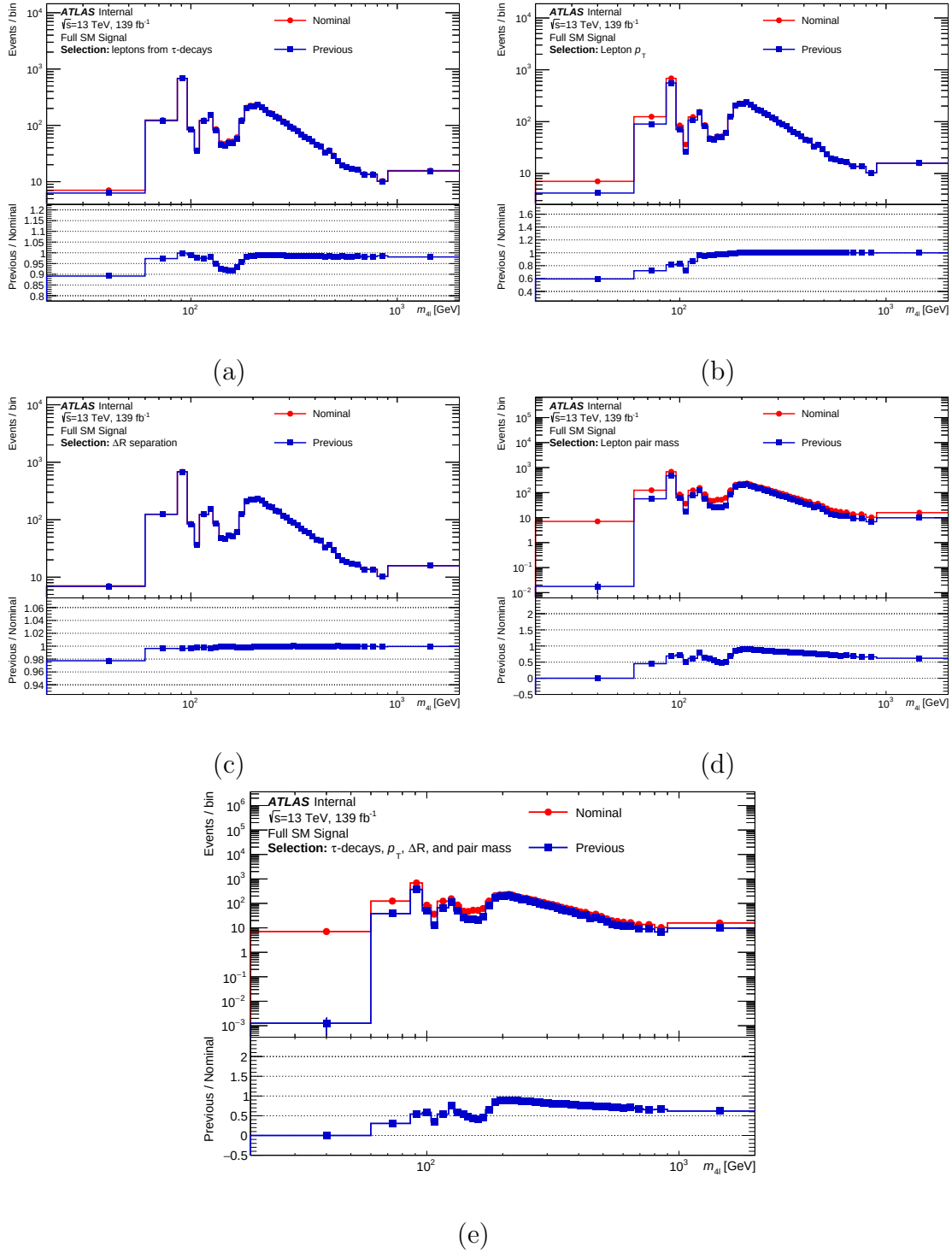


Figure 84: The impact of reverting four selections to the baseline used in the previous four-lepton analysis [44], compared to the nominal selection as used in this analysis. Shown separately for reverting the selections on τ -decays (a), the quadruplet lepton p_T (b), quadruplet lepton ΔR separation (c), and pair mass (d) one at a time. The combination of all reverting all four selections is shown in (e).

C.2 Lepton Isolation

Isolation is a measure of the particle activity surrounding a reconstructed object, and is crucial for the discrimination of prompt particles from non-prompt ones. Prompt leptons originating from the hard-scatter interaction are typically well-isolated, while non-prompt leptons from secondary hadronic decays are usually embedded in additional jet activity. The isolation working point (WP) used in the previous analysis round `FIXEDCUTLOOSE` is no longer supported, therefore the use of a new WP is required. The overarching analysis goal of maximizing acceptance and the relatively low reducible background motivate the adoption of a loose isolation WP, with a higher priority on signal efficiency over background rejection.

The `FIXEDCUTPFLOWLOOSE` WP offers a similarly loose threshold to the previous `FIXEDCUTLOOSE` configuration, as well as improved robustness against pile-up. The WP is based on an isolation measure, here referred to as `pflow30`, which uses pile-up robust versions of the track and calorimeter isolation measurements. It is defined in Equation 97 as the maximum of two measures of the track isolation surrounding the lepton, the first within a fixed cone of size $\Delta R < 0.2$ and the second within a p_T -dependent cone of minimum size $\Delta R < 0.3$, plus an additional calorimeter isolation measurement `newflowisol20` calculated using particle flow reconstruction techniques within a cone of size $\Delta R < 0.2$. To improve pile-up robustness, all tracks entering the track isolation summations are required to pass tight track-to-vertex association selections and satisfy $p_T > 500$ MeV.

$$\text{pflow30} := \max \left(\sum_i^{\Delta R < 0.2} p_T^i ; \sum_i^{\Delta R < \min(0.3, \frac{10 \text{ GeV}}{p_T^\ell})} p_T^i \right) + 0.4 \times \text{newflowisol20} \quad (97)$$

The `FIXEDCUTPFLOWLOOSE` WP corresponds to the requirement that the measurement of `pflow30` is less than 0.16 times the lepton p_T , motivating the similar truth isolation selection outlined in Section 11.2.

The performance in terms of signal acceptance and background rejection obtained by

adopting the `FIXEDCUTPFLOWLOOSE` WP for both electrons and muons (“Reference”) are shown alongside two alternative options of maintaining the previous `FIXEDCUTLOOSE` WP for either electrons and muons (“BothFCLoose”) or electrons only (“ElecFCLoose”) in Figure 85. The signal acceptance of the `FIXEDCUTPFLOWLOOSE`-based selection significantly improves the low-mass acceptance for both electrons and muons, while the background rejection remains largely unchanged within statistical uncertainties. The signal efficiency below $m_{4\ell} < 100$ GeV is shown versus the average pile-up density in Figure 86 for the same configurations to better highlight the performance differences. The `FIXEDCUTPFLOWLOOSE` WP, based on more pile-up robust isolation measurements, is shown to significantly reduce the pile-up degradation, particularly for muons (i.e. “BothFCLoose” $\rightarrow$ “ElecFCLoose”).

Based on these studies, as well as comparisons against a third WP not shown here, the `FIXEDCUTPFLOWLOOSE` WP was chosen for both electrons and muons in this analysis. Centrally provided muon isolation efficiency scale factors derived specifically for this WP are used to correct for differences between data and MC. For electrons, however, no such scale factors were available at the time. Therefore, electron isolation efficiency scale factors derived for the most similar alternative (`FIXEDCUTLOOSE`) are used instead.

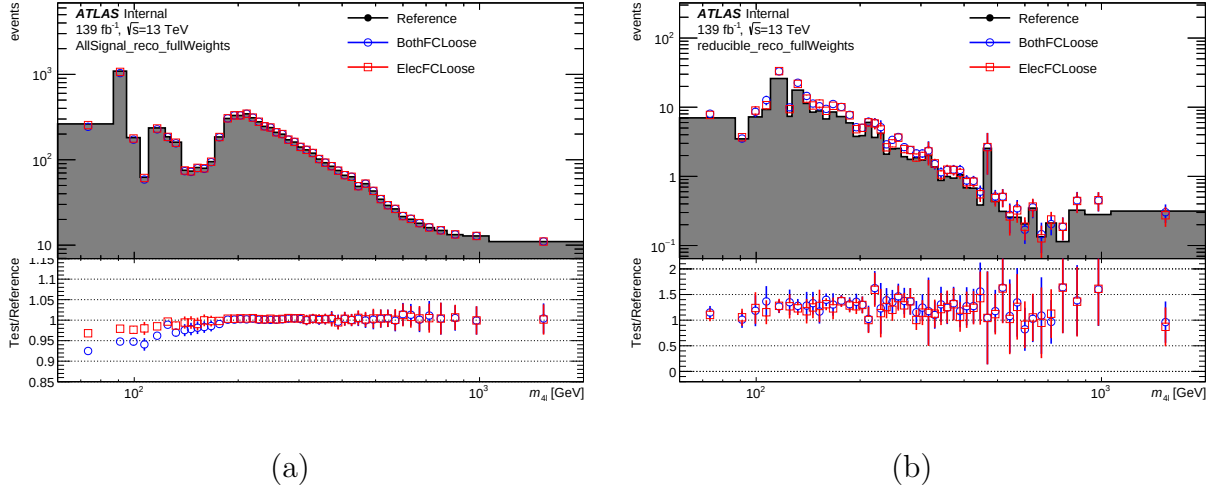


Figure 85: The predicted signal (a) and reducible background (b) reconstruction yields obtained for three isolation WP configurations: (1) adopting FIXEDCUTPFLOWLOOSE for both electrons and muons (“Reference”), (2) maintaining FIXEDCUTLOOSE for both electrons and muons (“BothFCLoose”), and (3) the combination of FIXEDCUTPFLOWLOOSE for muons and FIXEDCUTLOOSE for electrons only (“ElecFCLoose”).

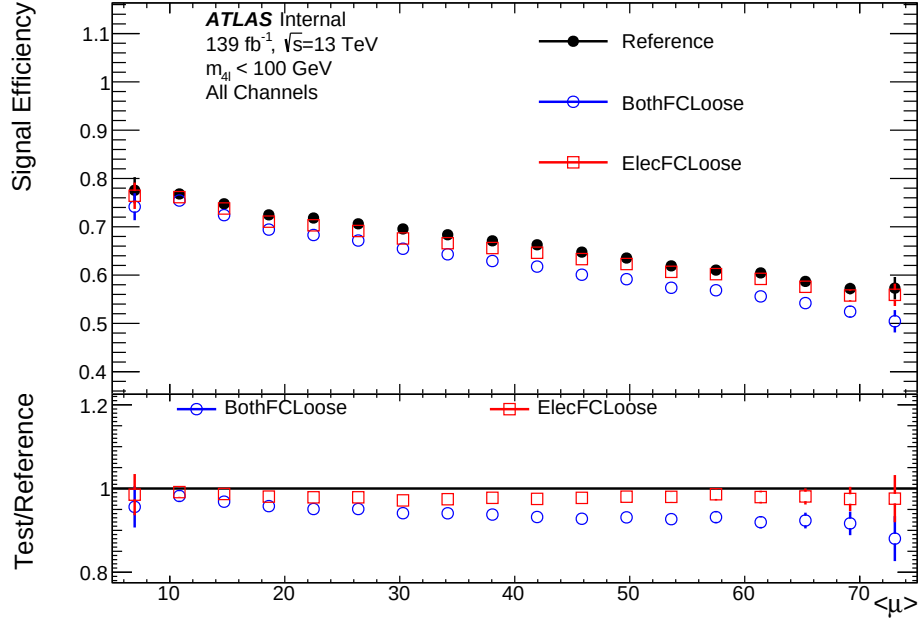


Figure 86: The predicted signal efficiency below $m_{4\ell} < 100$ GeV versus the average pile-up density for three isolation WP configurations: (1) adopting FIXEDCUTPFLOWLOOSE for both electrons and muons (“Reference”), (2) maintaining FIXEDCUTLOOSE for both electrons and muons (“BothFCLoose”), and (3) the combination of FIXEDCUTPFLOWLOOSE for muons and FIXEDCUTLOOSE for electrons only (“ElecFCLoose”).

D Background Estimation Strategy

Although the background estimation used in this analysis in principle operates on a four-lepton selection, a slight complication arises in the precise background estimation for this analysis to account for a core assumption of the matrix method and fake factor method - that the probability of any given lepton to be real or fake is uncorrelated to that of other leptons in the event. While this assumption is normally safe for well-separated leptons, this analysis uses a relatively loose separation requirement of $\Delta R > 0.05$. Studies found that for the particular case of cascade decays of b -hadrons, primarily occurring in $t\bar{t}$ events, hadronic activity can lead to two or more nearby leptons with highly correlated real and fake probabilities. In these sorts of cases, the estimation techniques are invalid and cannot be reliably used. Therefore, a workaround based on omitting nearby leptons in the matrix method or fake factor calculation was devised. Specifically, if two signal leptons are found to be within $\Delta R \leq 0.4$ of one another then either the electron is skipped if the pair is SF-OC or the lowest p_T lepton is skipped if the leptons are of the same flavor. The choice of skipping electrons over muons is based on the greater influence of electrons on nearby muon isolation cones - that is, calorimeter clusters corresponding to electrons can contribute significantly to the calorimeter isolation of nearby muons, while the reciprocal effect of muons is negligible in comparison. Hence, if a muon is observed to be isolated, despite a nearby electron, it is likely the electron is also isolated and, therefore, its probability of passing the signal requirements is correlated to that of the muon. The effect of skipping one or more leptons (N_{skipped}) in the matrix method or fake factor calculation means the background estimation corresponds to a modified $(4 - N_{\text{skipped}})$ selection. This technique is used in the signal region and validation regions alike, where good agreement is observed in MC closure tests.

E Data-driven Closure Test

The results of the data-driven closure test, discussed in Section 17.6, are summarized here for several other observables not included in the main body.

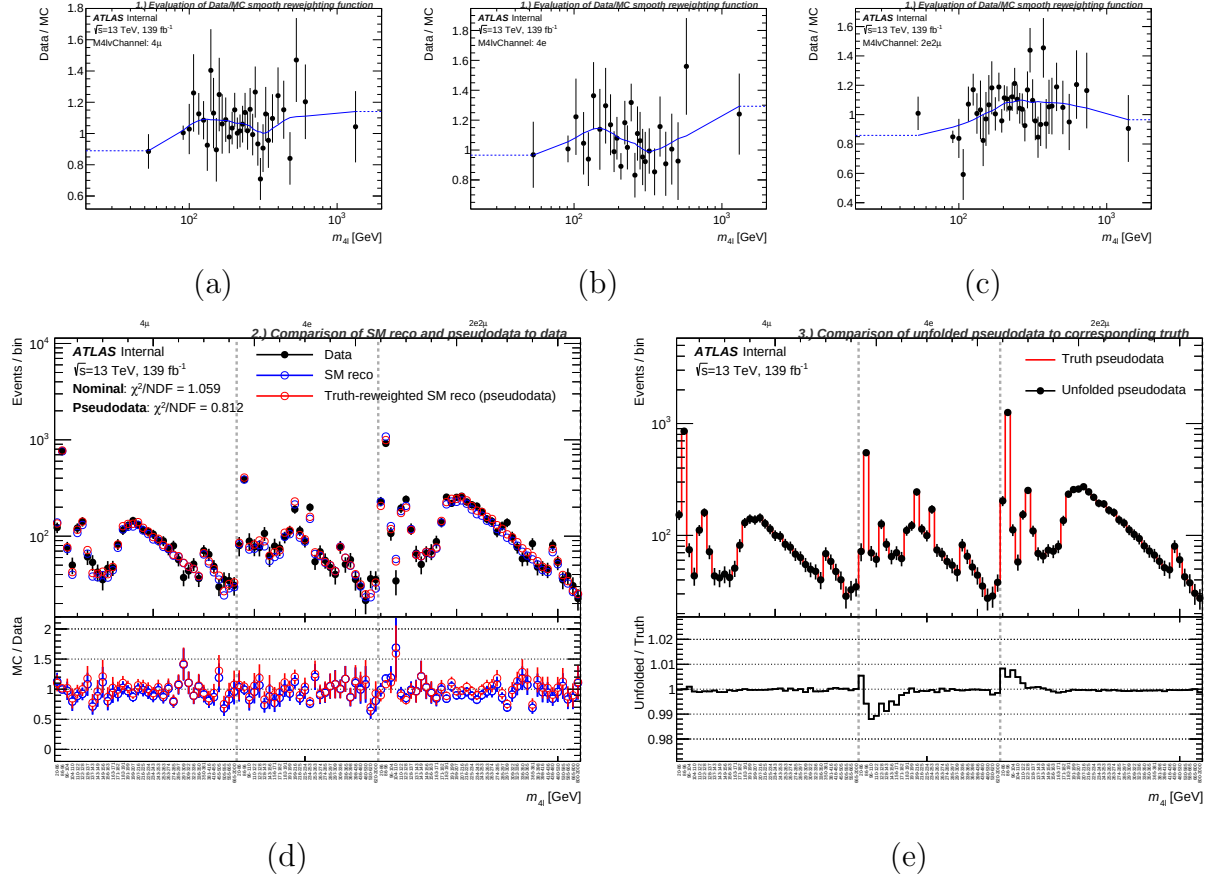


Figure 87: Step-by-step overview of the data-driven closure method for the double-differential $m_{4\ell}$ -channel observable. In (a)-(c), a series of continuous reweighting functions are derived by smoothing the observed data/MC ratios while keeping the endpoints fixed, as shown by the dashed lines. In (d), the reconstruction-level truth-reweighted MC (“pseudodata”) and the nominal MC are shown alongside the observed data, showing, as expected, an improved agreement with the data for the reweighted sample. In (e), the result of unfolding the pseudodata using the nominal response matrix is compared to the truth-reweighted fiducial prediction. The relative difference between the two is taken as a source of systematic uncertainty, capturing the inherent bias in the unfolding procedure.

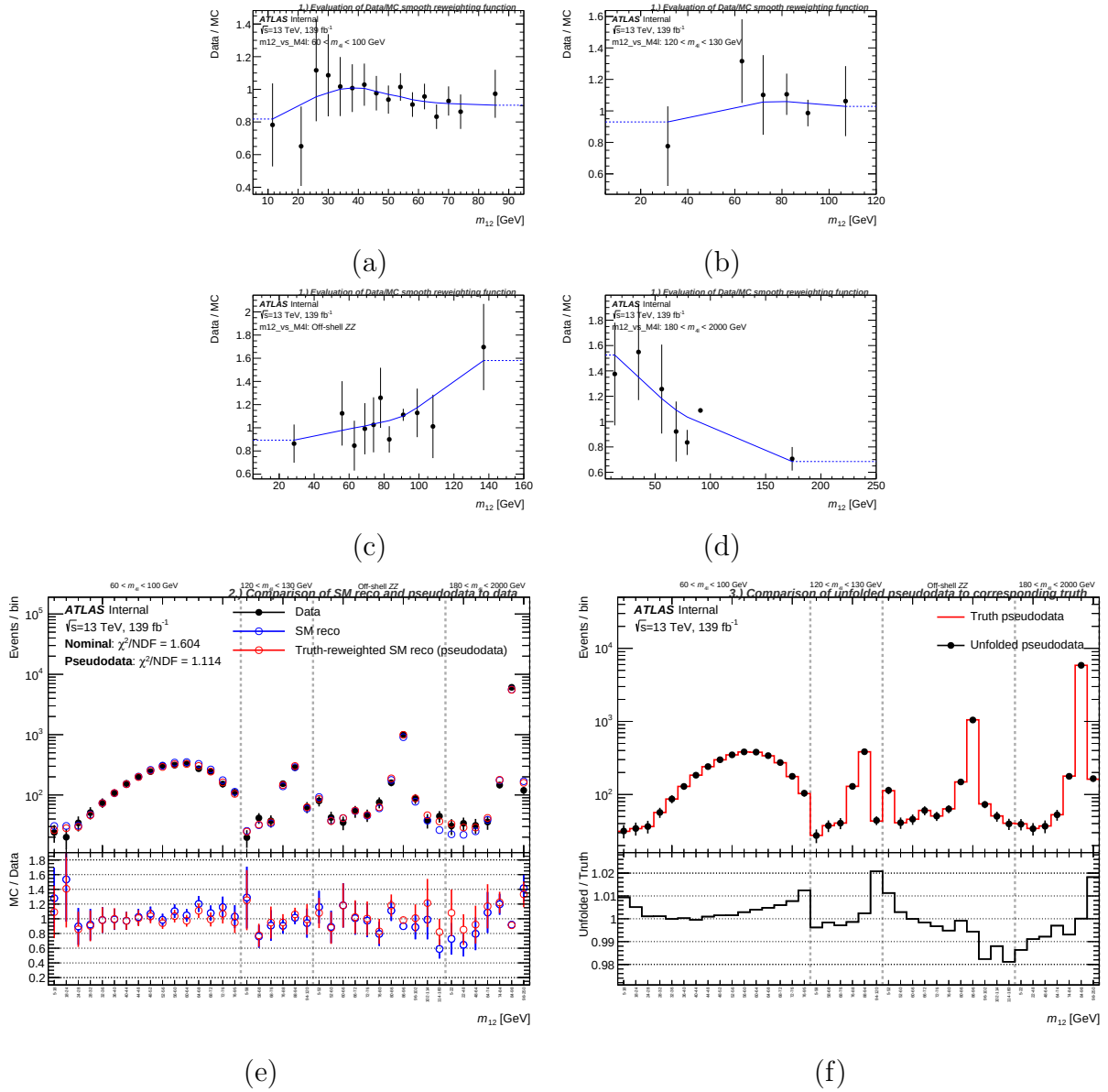


Figure 88: Step-by-step overview of the data-driven closure method for the double-differential $m_{Z1}-m_{4\ell}$ observable. In (a)-(d), a series of continuous reweighting functions are derived by smoothing the observed data/MC ratios while keeping the endpoints fixed, as shown by the dashed lines. In (e), the reconstruction-level truth-reweighted MC (“pseudodata”) and the nominal MC are shown alongside the observed data, showing, as expected, an improved agreement with the data for the reweighted sample. In (f), the result of unfolding the pseudodata using the nominal response matrix is compared to the truth-reweighted fiducial prediction. The relative difference between the two is taken as a source of systematic uncertainty, capturing the inherent bias in the unfolding procedure.

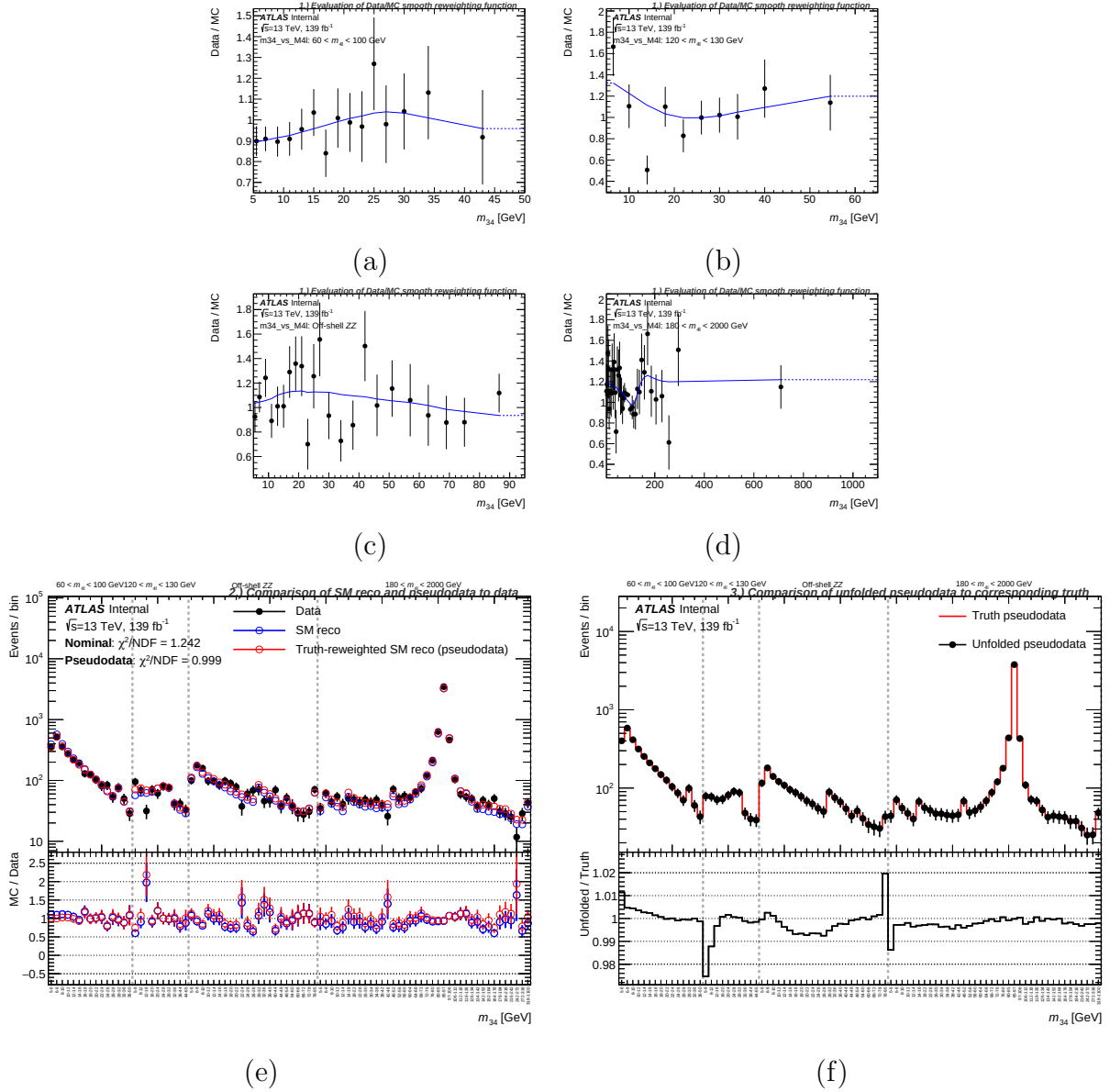


Figure 89: Step-by-step overview of the data-driven closure method for the double-differential m_{Z_2} - $m_{4\ell}$ observable. In (a)-(d), a series of continuous reweighting functions are derived by smoothing the observed data/MC ratios while keeping the endpoints fixed, as shown by the dashed lines. In (e), the reconstruction-level truth-reweighted MC (“pseudodata”) and the nominal MC are shown alongside the observed data, showing, as expected, an improved agreement with the data for the reweighted sample. In (f), the result of unfolding the pseudodata using the nominal response matrix is compared to the truth-reweighted fiducial prediction. The relative difference between the two is taken as a source of systematic uncertainty, capturing the inherent bias in the unfolding procedure.

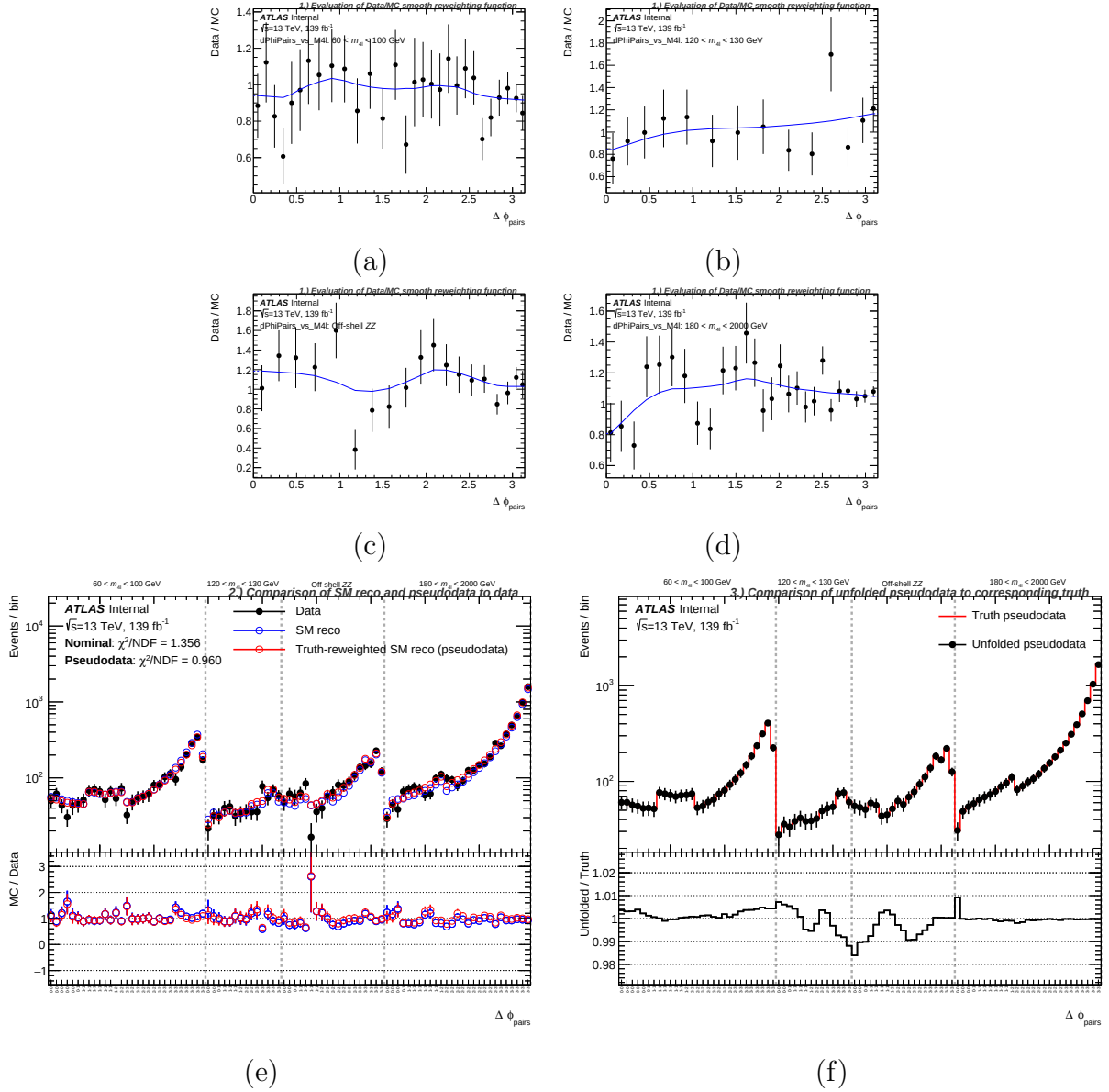


Figure 90: Step-by-step overview of the data-driven closure method for the double-differential $|\Delta\phi_{\text{pairs}}|-m_{4\ell}$ observable. In (a)-(d), a series of continuous reweighting functions are derived by smoothing the observed data/MC ratios while keeping the endpoints fixed, as shown by the dashed lines. In (e), the reconstruction-level truth-reweighted MC (“pseudodata”) and the nominal MC are shown alongside the observed data, showing, as expected, an improved agreement with the data for the reweighted sample. In (f), the result of unfolding the pseudodata using the nominal response matrix is compared to the truth-reweighted fiducial prediction. The relative difference between the two is taken as a source of systematic uncertainty, capturing the inherent bias in the unfolding procedure.