

A SEARCH FOR RESONANT Z' PRODUCTION IN HIGH-MASS DIELECTRON
FINAL STATES WITH THE ATLAS DETECTOR IN RUN-2 OF THE LARGE
HADRON COLLIDER

By

Christopher G. Willis

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics - Doctor of Philosophy

2018

ABSTRACT

A SEARCH FOR RESONANT Z' PRODUCTION IN HIGH-MASS DIELECTRON FINAL STATES WITH THE ATLAS DETECTOR IN RUN-2 OF THE LARGE HADRON COLLIDER

By

Christopher G. Willis

A search is performed for new resonant high-mass phenomena in the dielectron final state. The search uses 36.1 fb^{-1} of proton-proton collision data, collected at $\sqrt{s} = 13 \text{ TeV}$ by the ATLAS experiment at the Large Hadron Collider during its 2015 and 2016 data-taking runs. The dielectron invariant mass is used as the search variable. No significant deviations from the Standard Model prediction are observed. Upper limits at the 95% credibility level are set on the cross section times branching fraction to dielectron pairs for resonant Z' models considered in the search. Lower limits on the resonance pole mass are also presented. For the Z'_{SSM} , masses are excluded up to 4.5 TeV, while masses up to 4.1 TeV are excluded in the E_6 -motivated Z'_χ model. Limits are also derived in the Minimal Z' Model on the relative coupling strength γ' .

In addition, a series of studies are conducted in order to assess and reduce the dominant systematic uncertainty of this analysis, which arises from the imprecise knowledge of the Parton Distribution Functions in regions of very high parton x . While this uncertainty does not limit the discovery potential of the analysis presented here, it has the potential to do so in future searches. A novel approach is developed, and is shown to significantly reduce this systematic uncertainty in the high-mass search region of interest, thereby improving the discovery potential of future analyses.

For my parents, Susan and Gibson Willis, my brother, Matthew Willis, and my full time partner in crime, Jone Brunelle. Thanks for always believing in me.

ACKNOWLEDGMENTS

First, I would like to give a huge thanks to my excellent adviser Chip Brock. It is hard to compress 7 years worth of gratitude into a single paragraph. Chip played a formative role in this experience, providing me with enough space to learn and work on my own, but also the necessary guidance if I got stuck, and to him I'll forever be grateful.

I would also like to thank Dan Hayden, the contact of the dilepton analysis and my surrogate adviser at CERN. Dan always took the time to answer the many questions of a new student, and continued to point me in the right direction as we got up and running with the Run-2 dilepton search. From Dan, I learned how to do a real LHC physics analysis, and by the end I *actually understood* how everything fit together.

Thanks to the MSU HEP group for supporting me during my Ph.D., and for funding me as I went to workshops and conferences to present my work. A special thanks to those who were responsible for maintaining the MSU Tier3, especially during analysis crunch time.

I would like to thank my committee members, Norman Birge, Joey Huston, Chris Wrede, and C.-P. Yuan, for accepting the job and taking the time to read and comment on this thesis. Also, a big thanks to Brenda Wenzlick and Kim Crosslan for helping me to schedule my thesis defense, and taking care of the paperwork that followed.

Finally, I'd like to thank C.-P. Yuan, Carl Schmidt, Josh Isaacson, and T.-J. Hou for sharing their PDF knowledge and theory insights for the completion of the ePump project, which formed the latter half of this thesis.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xv
Part I The Theory	1
Chapter 1 Introduction	2
Chapter 2 The Standard Model	9
2.1 Fundamental Particles and Interactions	9
2.1.1 Matter Particles	10
2.1.2 Forces and Interactions	12
2.1.3 Particle Masses	14
2.2 Mathematical Formulation	14
2.2.1 Lagrangian Dynamics	15
2.2.2 Symmetries of the Standard Model	16
2.2.3 Gauge Symmetries	17
2.2.4 Standard Model Lagrangian	20
2.3 The Unified Electroweak Theory	21
2.3.1 History	21
2.3.2 Fermion Fields	22
2.3.3 The Electroweak Lagrangian	24
2.3.4 The Higgs Mechanism	26
2.3.5 Yukawa Interactions	30
2.3.6 The Electroweak Lagrangian Revisited	33
2.4 Quantum Chromodynamics	36
2.4.1 Asymptotic Freedom and Confinement	38
Chapter 3 Proton-Proton Collisions	40
3.1 Proton Structure	40
3.1.1 Proton-Proton Collisions	42
3.1.2 QCD Factorization Theorem	44
3.1.3 Hadronic Events	45
3.2 Parton Distribution Functions	48
3.2.1 Extracting PDFs from Experimental Data	50
3.2.2 Assessing PDF Uncertainties	53
3.2.3 Summary	56
Chapter 4 Two Body Scattering and The Drell-Yan Process	57
4.1 The S-Matrix and Feynman Rules	57

4.1.1	Scattering Amplitudes	58
4.1.2	Feynman Diagrams	59
4.1.3	Two Body Scattering	60
4.2	The Drell-Yan Process	62
4.2.1	Leading-Order Cross Section	64
4.2.2	PDF Uncertainty	69
4.2.3	Higher-Order Corrections	71
Chapter 5	Physics Beyond the Standard Model	74
5.1	Limitations of the Standard Model	74
5.2	Theories Beyond the Standard Model	78
5.3	Models with Extended Gauge Symmetries	80
5.3.1	Dilepton Phenomenology	81
5.3.2	The Sequential Standard Model	84
5.3.3	The E_6 Model	85
5.3.4	The Minimal Model	88
5.4	Previous Experimental Limits	90
Part II	The Experiment	93
Chapter 6	The Large Hadron Collider	94
6.1	The LHC	94
6.2	CERN Accelerator Complex	96
6.3	LHC Operation	96
6.3.1	Pile-up	98
6.3.2	Luminosity	100
6.4	LHC Performance in Run-II	101
6.5	Experiments at the LHC	102
Chapter 7	The ATLAS Experiment	104
7.1	Detector Overview	104
7.2	Coordinate System and Event Kinematics	107
7.3	Tracking System	109
7.3.1	Pixel Detector	110
7.3.2	Semiconductor Tracker	112
7.3.3	Transition Radiation Tracker	112
7.4	Calorimetry System	113
7.4.1	Electromagnetic Calorimeter	115
7.4.2	Hadronic Calorimeter	117
7.5	Muon System	118
7.6	Trigger and Data-Acquisition System	120
Chapter 8	Electron Reconstruction in ATLAS	124
8.1	Track and Vertex Reconstruction	124

8.1.1	Track Reconstruction	125
8.1.2	Vertex Reconstruction	127
8.2	Electrons	128
8.2.1	Electron Trigger	128
8.2.2	Electron Reconstruction	129
8.2.3	Electron Identification	130
8.2.4	Electron Isolation	134
8.3	Energy Corrections	136
8.3.1	Energy Scale Calibration	136
8.3.2	Energy Resolution Calibration	137
8.4	Efficiency Corrections	138
Chapter 9	Monte Carlo Simulation	141
9.1	Event Generators	141
9.1.1	Parton Shower Generators	142
9.1.2	Matrix Element Generators	143
9.1.3	Generators Used in the Dilepton Analysis	143
9.2	ATLAS Simulation Chain	145
9.2.1	Event Generation	146
9.2.2	Detector Simulation	146
9.2.3	Digitization	147
9.3	Simulation Samples	148
9.3.1	Drell-Yan Production	148
9.3.2	Photon-Induced Production	149
9.3.3	$t\bar{t}$ Production	150
9.3.4	Single-Top Quark Production	151
9.3.5	Diboson Production	152
9.3.6	Z' Production	153
9.3.7	MC Sample Summary	154
9.4	Corrections to Simulation	154
9.4.1	Sample Normalization	155
9.4.2	Efficiency Scale Factor Corrections	155
9.4.3	Energy Scale and Resolution Smearing	157
9.4.4	Pile-Up Profile	159
9.4.5	k-Factor Corrections	160
Part III	The Analysis	162
Chapter 10	Electron Channel Dilepton Analysis	163
10.1	Analysis Strategy	163
10.2	Data Samples	166
10.2.1	ATLAS Data	166
10.2.2	Analysis Data	167
10.2.3	Dataset Derivations	168

10.3	Monte Carlo Samples	170
10.4	Analysis Software	170
10.5	Event Selection	172
10.5.1	Selection Criteria	173
10.5.2	Cutflows	176
10.5.3	Signal Acceptance	178
10.6	Fake Background Estimate	180
10.6.1	The Matrix Method	180
10.6.2	Real- and Fake-Rate Determination	182
10.6.3	Background Extrapolation	183
10.7	Comparison of Data to the Standard Model Background Expectation	184
10.7.1	Event Yields	185
10.7.2	Kinematic Distributions	186
10.8	Systematic Uncertainties	189
10.8.1	Experimental Uncertainties	193
10.8.2	Theoretical Uncertainties	198
10.8.3	Summary	202
10.8.4	Comment on PDF Uncertainties	202
Chapter 11	Statistical Interpretation and Results	206
11.1	Statistical Model	207
11.2	Signal Reweighting	210
11.2.1	Signal Template Generation	211
11.2.2	Interference Templates	215
11.3	Signal Search	217
11.3.1	Log-Likelihood Ratio Test	219
11.3.2	BumpHunter Test	225
11.4	Exclusion Limits	227
11.4.1	Bayesian Inference	228
11.4.2	Limits on SSM and E_6 Model Z' Bosons	229
11.4.3	Limits on Minimal Model Z' Bosons	236
11.5	Discussion of Results	242
Part IV	PDF Uncertainty	244
Chapter 12	The Drell-Yan Process Revisited	245
12.1	The Drell-Yan Process	246
12.1.1	Drell-Yan Kinematics	246
12.1.2	The Collins-Soper Frame	250
12.1.3	Triple-Differential Cross Section	252
12.1.4	Kinematic Distributions	254
12.1.5	Explaining the $\cos\theta^*$ Asymmetry	255
12.1.6	$\cos\theta^*$ as a Quark Flavor Enhancer	260

Chapter 13 PDF Error Reduction	263
13.1 The ePump Package	264
13.2 PDF Update Strategy	266
13.3 Construction of ePump Templates	267
13.3.1 Kinematic Reach at the LHC	268
13.3.2 Template Simulation Samples	269
13.3.3 Template Event Selection	270
13.3.4 Template Binning	271
13.3.5 Pseudo-Data Generation	272
13.3.6 Finalized Templates	273
13.4 Results of the Update	273
13.4.1 Impact on CT14Hera2 PDFs	276
13.4.2 Impact on High-Mass Drell-Yan	282
13.5 Outlook	283
Part V The Conclusion	287
Chapter 14 Conclusion	288
14.1 Summary	288
14.2 Outlook	291
APPENDICES	294
Appendix A Properties of Analysis Simulation and Data	295
Appendix B Additional Dielectron Cutflow Tables	300
Appendix C ATLANTIS Event Displays	302
Appendix D ePump Inputs and Supplimental Results	305
Appendix E Early Warning System	334
BIBLIOGRAPHY	346

LIST OF TABLES

Table 2.1: Matter content of the Standard Model. The three generations of quarks and leptons are shown along with their name, symbol, charges under each interaction, and mass [1].	11
Table 2.2: Force carriers and Higgs boson of the Standard Model. The name, symbol, charges under each interaction, and mass of each particle is listed [1]. . . .	12
Table 2.3: Three generations of fermions in the Electroweak theory with their respective quantum numbers. Left-handed particles are grouped into iso-doublets, while right-handed particles exist as iso-singlets. The right-handed neutrinos are included for reference, but are not formally included as part of the Electroweak theory.	24
Table 2.4: Neutral current couplings of Standard Model fermions. The rows specify the couplings within each respective quark and lepton generation. Alternate definitions often are written in terms of $g_V = \frac{1}{2}(g_L + g_R)$ and $g_A = \frac{1}{2}(g_L - g_R)$, which function as coefficients to the vector (γ^μ) and axial-vector ($\gamma^\mu\gamma^5$) terms of the neutral current Lagrangian.	36
Table 5.1: Standard Model fermion charges under the $U(1)_\psi$ and $U(1)_\chi$ symmetries of the E_6 GUT model. Once the mixing angle θ_{E_6} is specified, the Z' Lagrangian of Eq. 5.2 is uniquely defined, and can be used to make E_6 -related predictions.	86
Table 5.2: E_6 models for several motivated choices of mixing angle θ_{E_6} . From left to right, these are the ψ model, the neutral model, the η model, the inert model, the secluded sector model, and the χ model.	87
Table 5.3: Several Z'_{Min} models for well-motivated choices of θ_{Min} and γ' . From left to right, these are the pure $Z'_{\text{B-L}}$ model, the Z'_χ of the E_6 model, and the Z'_{3R} of the left-right symmetric model.	89
Table 5.4: Quantum numbers associated with Standard Model fermions according to Eq. 5.12. Once the mixing angle θ_{Min} and relative coupling γ' are specified, the Z' Lagrangian of Eq. 5.2 is uniquely defined and can be used to make Minimal model predictions.	89

Table 5.5: Observed limits at 95% C.L. on the mass of a Z' boson from previous experimental searches at hadron colliders. The CMS analyses listed do not provide limits on the Z'_χ , which would otherwise be slightly higher than what was obtained for the Z'_ψ . The integrated luminosity for each analysis is rounded to nearest whole number. The $\ell\ell$ channel refers searches that combine individual electron and muon channels. The analysis presented here is denoted with “*.”	91
Table 6.1: Main LHC beam parameters under typical running conditions. Shown are values for the 2015 run and the 2016 run through June, along with the LHC design values [2].	101
Table 8.1: Run-2 electron identification variables [3].	132
Table 9.1: Mass-dependent NLO-to-NNLO QCD and higher-order EW k-Factor corrections for several values of dielectron invariant mass. These corrections are applied to the enhance the nominal Drell-Yan background estimate with NNLO QCD and higher-order EW effects.	149
Table 9.2: Mass-dependent Photon-Induced k-Factor corrections at several values of dielectron invariant mass. These corrections are applied to the nominal the Drell-Yan background estimate to incorporate Photon-Induced effects.	150
Table 9.3: Mass-dependent LO-to-NNLO QCD k-Factor corrections at several values of dielectron mass m_{ee} . These corrections are applied to enhance the nominal Z' signal estimate with NNLO QCD effects.	154
Table 9.4: A summary of MC samples used for the background and signal modeling in the electron channel of the dilepton analysis. The columns include the process of interest, the MC generation program, the matrix element order, the order of the k-Factor correction, the parton shower program, and the PDF set used [4].	154
Table 10.1: The number of events passing the electron channel event selection in the 36.1 fb^{-1} dataset obtained by ATLAS during the 2015+2016 physics runs. The first column indicates the absolute number of events passing selection, the middle column depicts the relative efficiency of each cut, and the last column shows the absolute efficiency.	177
Table 10.2: The number of events passing the electron channel event selection in a sample of simulated Z'_χ events with pole mass $M_{Z'} = 3 \text{ TeV}$. The sample has all corrections applied that are described in Section 9.4, and is normalized to an integrated luminosity of 36.1 fb^{-1} according to the procedure outline in Section 9.4.1.	177

Table 10.3: Expected and observed event yields in different invariant mass intervals according to the event selection of Section 10.5.1. The errors include both statistical and systematic uncertainties (described in Section 10.8).	186
Table 10.4: A Summary of the relative systematic uncertainties affecting the expected signal and background event yields at dielectron invariant masses of 2 and 4 TeV. Background values correspond to the relative change in the total number of expected events in the 2 and 4 TeV bins of the m_{ee} histogram depicted in Figure 10.6, while signal values are calculated from changes in the FWHM of the Z'_χ signal at pole masses of $M_{Z'} = 2$ TeV and $M_{Z'} = 4$ TeV. Here N/A refers to uncertainties not considered on the signal process [5].	203
Table 11.1: Expected and observed 95% C.L. lower mass limits for the SSM and $E_6 Z'$ bosons considered in the analysis. Electron, muon, and combined channel limits are shown [5].	235
Table 11.2: Observed and expected 95% C.L. lower mass limits on the Z'_{Min} pole mass for the three representative models considered in the text. As a result of including the interference effects and the variation of the decay width $\Gamma_{Z'}$ through the modulation of γ' , a slightly weaker limit for the Z'_{χ} is obtained in the Minimal model approach as compared to the standard approach of Section 11.4.2.	241
Table 11.3: Range in the observed and expected 95% C.L. lower limits on the Z'_{Min} pole mass for several representative values of the relative coupling strength γ' . Both the electron and muon channels are combined.	241
Table 11.4: Range in the observed and expected 95% C.L. upper limits on the relative coupling strength γ' for several representative values of the Z'_{Min} boson pole mass. Both the electron and muon channels are combined.	241
Table 13.1: The ranges in parton- x probed in the Drell-Yan process at the $\sqrt{s} = 13$ TeV LHC for various values of dilepton rapidity and four-momentum transfer $Q = m_{\ell\ell}$. The maximum values assumed for the dilepton rapidity are determined from the $ y_{\ell\ell} $ distributions in Figure 12.4, where the Drell-Yan cross section falls to zero. The minimum and maximum values in x are calculated using Eq. 12.7.	269
Table 13.2: Impact of 3000 fb^{-1} update on the CT14Hera2 $u_v(x)$ and $d_v(x)$ valence and $\bar{u}(x)$ and $\bar{d}(x)$ sea distributions for several values of x using the standard triple-differential templates at $Q = 3$ TeV. To be compared with Figures 13.4 and 13.5.	277

Table 13.3: The estimated PDF uncertainty in several invariant mass bins for the distributions shown in Figure 13.8. The current CT14Hera2 uncertainty estimates are shown in the first column, the result of the 3000 fb^{-1} update is shown next, and the 3000 fb^{-1} update with an additional $\cos \theta^* > 0$ requirement on the dilepton selection is shown last. Notice the pre-update values for the central-central (CC) selection are consistent with those assessed in the dilepton analysis (See Fig. 10.11).	283
Table 13.4: The post-update PDF uncertainty as compared to the experimental and dominant theoretical uncertainties in the electron channel of the dilepton analysis. As the PDF uncertainty will be reduced well below the current experimental uncertainty, attention will be shifted to the reduction of others, such as the “PDFChoice” uncertainty, improving the discovery potential of future iterations of the dilepton analysis.	285
Table A.1: The NLO PowHEG+PYTHIA8 Drell-Yan MC samples used in the modeling of the electron channel [4].	296
Table A.2: The LO PYTHIA8 Z'_χ MC samples used to validate the electron channel reweighting procedure [4].	296
Table A.3: The LO PYTHIA8 Drell-Yan MC samples used to construct reweighted Z' signal templates for the electron channel search and limit setting of the dilepton analysis [4].	297
Table A.4: The NLO PowHEG+PYTHIA6 $t\bar{t}$ and Single-Top MC samples used in the modeling of the electron channel in the dilepton analysis. A generator-level filtering efficiency of 0.10534 is accounted for in the $t\bar{t}$ sample cross section estimate [4].	297
Table A.5: The NLO SHERPA WW , WZ , and ZZ Diboson MC samples used in the modeling of the electron channel in the dilepton analysis [4].	298
Table A.6: LHC Datasets used in the dilepton analysis [4].	299
Table B.1: The number of simulated Drell-Yan events passing the electron channel event selection of the dilepton analysis.	300
Table B.2: The number of simulated Top ($t\bar{t} + Wt$) events passing the electron channel event selection of the dilepton analysis.	301
Table B.3: The number of simulated Diboson ($WW + WZ + ZZ$) events passing the electron channel event selection of the dilepton analysis.	301

Table D.1: Impact of 300 fb^{-1} update on the CT14Hera2 $u_v(x)$ and $d_v(x)$ valence and $\bar{u}(x)$ and $\bar{d}(x)$ sea distributions for several values of x using the standard triple-differential templates at $Q = 3 \text{ TeV}$. To be compared with Fig.s D.25 and D.26.	332
Table D.2: The estimated PDF uncertainty in several invariant mass bins for the distributions shown in Figure D.27. The current CT14Hera2 uncertainty estimates are shown in the first column, the result of the 300 fb^{-1} update is shown next, and the 300 fb^{-1} update with an additional $\cos \theta^* > 0$ requirement on the dilepton selection is shown last.	332

LIST OF FIGURES

<p>Figure 2.1: The QED interaction vertex. The solid lines represent the motion of the charged leptons and the curved line represent the propagation of the photon. The strength of the interaction is determined by the magnitude of electric charge Q in units of fundamental charge e and is labeled at the vertex.</p> <p>Figure 2.2: The shape of the Higgs potential $V(\Phi)$ of Eq. 2.19 for the choice of $\mu^2 > 0$. The potential $V(\Phi)$ obtains a minimum at the non-zero value of $\Phi^\dagger\Phi = \mu^2/2\lambda \equiv v^2/2$ [6].</p> <p>Figure 2.3: Fundamental interaction vertices with the Higgs boson in the EW theory: the Higgs-fermion interaction (a), and the three-point (b) and four-point (c) gauge interactions.</p> <p>Figure 2.4: Fundamental three-point (a) and four-point (b) Higgs boson self-interaction vertices in the EW theory.</p> <p>Figure 2.5: Fundamental three-point (a) and four-point (b)-(c) gauge boson vertices in the EW theory.</p> <p>Figure 2.6: Fundamental interaction vertices with quarks in the EW theory: the charged current interactions (a)-(b) and neutral current interaction (c).</p> <p>Figure 2.7: Fundamental interaction vertices with leptons in the EW theory: the charged current interactions (a)-(b) and neutral current interactions (c).</p> <p>Figure 2.8: Fundamental interaction vertices in QCD: the quark-gluon interaction (a), the three-point gluon interation (b), and the four-point gluon interation (c).</p> <p>Figure 2.9: The running of the α_S coupling constant in QCD as a function of the energy scale Q with a variety of experimental measurements overlaid [1].</p> <p>Figure 3.1: Substructure of proton depicting interactions between its partonic constituents: valence quarks, sea quarks, and gluon [7].</p> <p>Figure 3.2: Diagram of a typical hard scattering interaction between partons at the LHC [8].</p>	<p>19</p> <p>27</p> <p>29</p> <p>30</p> <p>33</p> <p>34</p> <p>35</p> <p>37</p> <p>39</p> <p>41</p> <p>42</p>
--	---

Figure 3.3: An inelastic collision at the LHC. The dark red blob depicts the hard process, with its associated red lines the resulting final state radiation. Beam remnants are light blue, and initial state radiation is depicted by the dark blue lines. The purple blob is a secondary hard scatter along with its own final state radiation in purple. The light green blobs depict hadronization of any color-charged partons emanating from the hard scatter or the ISR/FSR showering. Subsequent dark green blobs corresponds to hadronic decays of unstable particles and yellow lines indicate the presence of electromagnetic radiation [9].	46
Figure 3.4: CT14NNLO PDFs provided by the CTEQ collaboration. Depicted are gluon and all quark flavors as a function of x , evaluated at the scales of $Q = 2 \text{ GeV}$ (a) and $Q = 100 \text{ GeV}$ (b) [10].	49
Figure 3.5: Parton kinematics for the regions in (x, Q^2) probed by the LHC and previous DIS and fixed target experiments. The scale Q^2 of a process is often identified with the mass M of a produced resonance. This is indicated by four discrete values on the plot, for the LHC with $\sqrt{s} = 14 \text{ TeV}$ [8]. . . .	52
Figure 3.6: PDF uncertainties associated with the $g(x)$, $u(x)$, $d(x)$, $s(x)$, $u_v(x)$, $d_v(x)$, $\bar{u}(x)$, and $\bar{d}(x)$ distributions in the CT14NNLO and CT14Hera2 PDF sets evaluated at a scale of $Q = 3 \text{ TeV}$	54
Figure 4.1: Two body scattering as described by Eq. 4.5 [11].	61
Figure 4.2: The Drell-Yan process as observed at the LHC [12].	62
Figure 4.3: The LO Feynman Diagram for Drell-Yan production mediated by a virtual photon γ^* (a), and on-shell Z boson (b). The propagator and vertex factors of each diagram are depicted. The initial- and final-state particles are represented by the Dirac spinors u and v , which label their momenta and spin.	66
Figure 4.4: The invariant mass distribution of dilepton pairs $m_{\ell\ell}$ produced via the Drell-Yan mechanism at $\sqrt{s} = 13 \text{ TeV}$. The MC generator RESBOS, configured with the CT14Hera2 PDF set, was used for the prediction. The PDF error, calculated with the symmetrized form of Eq. 3.10, is indicated in the sub-panel, and steadily increases with dilepton mass.	70
Figure 4.5: Example higher-order contributions to the Drell-Yan process: NLO initial-state gluon radiation $q\bar{q} \rightarrow \gamma^*/Z + g$ (a), NLO initial-state quark radiation $qg \rightarrow \gamma^*/Z + q$ (b), NLO virtual gluon correction $q\bar{q} \rightarrow \gamma^*/Z$ (c), NLO final-state photon radiation $q\bar{q} \rightarrow \gamma^*/Z + \gamma$ (d).	72

Figure 5.1: Leading Order Feynman diagram for Z' production at the LHC. The propagator and vertex factors are depicted, and the initial- and final-state fermions are represented by the Dirac spinors u and v , which label their momenta and spin.	82
Figure 5.2: Predicted Z'_{SSM} event yields for $M_{Z'} = 1$ TeV (a) and $M_{Z'} = 1$ TeV (b). The vertical axis displays the number of predicted Z' events per given unit of integrated luminosity, as would be observed at the $\sqrt{s} = 13$ TeV LHC.	85
Figure 5.3: Predicted Z' event yields for various choices of θ_{E_6} in the E_6 model with $M_{Z'} = 1$ TeV (a) and $M_{Z'} = 1$ TeV (b). Starting from the Z'_ψ value of $\theta_{E_6} = 0$, E_6 models will exhibit increasing cross sections and decay widths, until the Z'_χ value of $\theta_{E_6} = \pi/2$ is reached, at which point a decrease is observed. All E_6 models fall within the range defined by Z'_ψ and Z'_χ . The vertical axis displays the number of predicted Z' events per given unit of integrated luminosity, as would be observed at the $\sqrt{s} = 13$ TeV LHC.	87
Figure 5.4: Predicted Z'_{Min} event yields for $M_{Z'} = 3$ TeV according to Eq. 5.12. In (a), θ_{Min} is held constant for several distinct values of γ' . Similarly, in (b), γ' is held constant while θ_{Min} is varied. The vertical axis displays the number of predicted Z'_{Min} events per given unit of integrated luminosity, as would be observed at the $\sqrt{s} = 13$ TeV LHC.	90
Figure 6.1: Illustration of the 27 km long LHC, its four main experiments (ALICE, ATLAS, CMS, and LHCb), the pre-accelerators used in staging proton acceleration to reach a CM energy of $\sqrt{s} = 13$ TeV, and the region surrounding Geneva, Switzerland [13].	95
Figure 6.2: Illustration of the CERN accelerator complex. Proton acceleration is achieved through the following stages: Linac 2, PSB, SP, SPS, and LHC. Accelerators used in other CERN experiments are shown as well [14].	97
Figure 6.3: Superconducting dipole magnets as seen in the LHC tunnel [15].	98
Figure 6.4: Depiction of a $Z \rightarrow \mu^+ \mu^-$ event as recorded by the ATLAS detector during Run-1 of the LHC, along with the 25 simultaneous proton-proton interactions during this particular bunch crossing. The two muons originate from a common vertex, due to the decay of the Z boson, and are highlighted in yellow [16].	99

Figure 6.5: Cumulative integrated luminosity obtained per week as delivered by the LHC and recorded by the ATLAS detector during the 2015 (a) and 2016 (b) data-taking runs [17, 18].	102
Figure 7.1: Individual particles are identified and measured in the ATLAS detector according to their characteristic signatures left in each detector subsystem. Solid lines represent trajectories of charged particles, while dashed lines indicate neutral particles. Showers are depicted for particles that interact with the dense detector material [19].	105
Figure 7.2: Cut-away view of the ATLAS detector with individual subsystems labeled [20].	106
Figure 7.3: Cut-away view of the ATLAS Inner Detector. Subsystem components are shown in both the barrel and end-cap regions [21].	110
Figure 7.4: Schematic of the barrel region of ATLAS Inner Detector, as crossed by a charged particle [21].	111
Figure 7.5: Cut-away view of the ATLAS calorimetry system. Depicted are the Electromagnetic and Hadronic sampling calorimeters, both of which consist of individual barrel and end-cap modules. The forward components are closest to the beam line [22].	114
Figure 7.6: Depiction of an ECAL barrel module in the precision region of the ATLAS detector. The accordion shape is shown, along with the three independent sampling layers. The granularity in η and ϕ of the cells of each layer is labeled [23].	116
Figure 7.7: Sketch of an ATLAS tile calorimeter module [24].	117
Figure 7.8: Schematic of the ATLAS detector forward calorimeter systems. The three FCAL modules are depicted [25].	119
Figure 7.9: Cut-away view of the ATLAS Muon System [26].	119
Figure 7.10: The production cross sections and corresponding event rates for a variety of physics processes observable at the LHC. BSM processes, such as Z' production, are expected to occur at extremely small rates compared to the nominal rate of proton-proton inelastic collisions indicated by σ_{tot} . In this regard, a trigger is essential to correctly identify and select such events for offline analysis [8].	121
Figure 8.1: Depiction of ATLAS detector systems used to perform electron reconstruction and identification [3].	125

Figure 8.2: Illustration of the transverse impact parameter d_0 (a) and longitudinal impact parameter z_0 (b) [27].	127
Figure 8.3: Illustration comparing the ETCONE and TOPOETCONE isolation methods. The grid corresponds to the $\eta - \phi$ granularity of the middle layer of the EM calorimeter. The electron candidate cluster is mostly contained in the central 5×7 window. An isolation cone with radius $\Delta R = 0.4$ is drawn around the candidate. In the ETCONE method, all cells (yellow) within the radius ΔR are used, while in the TOPOETCONE method, only topoclusters (red), whose barycenters fall within ΔR , are used, effectively reducing noise from pile-up [28].	135
Figure 9.1: The LO Feynman diagrams for Photon-Induced Drell-Yan production. Both t -channel (a) and u -channel (b) modes are shown.	150
Figure 9.2: The LO Feynman diagrams for $t\bar{t}$ production. The gluon initiated s -channel (a) and t -channel (b) modes are depicted, along with quark-antiquark initiated production (c).	150
Figure 9.3: The LO Feynman diagrams (a) and (b) for Wt -channel Single-Top production.	151
Figure 9.4: The LO Feynman diagrams for Diboson production. Depicted are the s -channel (a), t -channel (b), and u -channel (c) production modes. The WZ and ZZ processes can proceed through all modes, while the ZZ process can produce only through the t - and u -channels. In each diagram $V_i = W, Z$ or γ	152
Figure 9.5: Electron efficiency scale factor corrections as a function of dielectron invariant mass m_{ee} . Shown are the scale factors applied to correct the electron reconstruction, identification, isolation, and trigger efficiencies.	156
Figure 9.6: Invariant Mass distribution for the $M_{Z'} = 3$ TeV Z'_χ predicted by the simulation with and without the application of the energy scale correction and resolution smearing. The effect is to both shift and broaden the shape of the Z' peak.	158
Figure 9.7: Mean number of interactions per bunch crossing in the $Z \rightarrow e^+e^-$ MC sample with and without the pile-up correction as compared to data. The effect of reweighting is to match the pile-up conditions in data.	159

Figure 9.8: The four mass-dependent k-Factor corrections applied in the dilepton analysis. The NLO-to-NNLO QCD correction, LO-to-NLO EW correction, and the LO Photon-Induced QED correction are applied to enhance the description of the Drell-Yan spectrum, while the LO-to-NNLO QCD correction is applied to the Z' signal.	160
Figure 10.1: Day-by-day data-taking efficiency of the ATLAS detector during in the 2015 (a) and 2016 (b) runs. The efficiency over the course of each run was above 92%. Empty bins correspond to days during which the LHC did not deliver stable beams [18].	167
Figure 10.2: Average number of proton-proton interactions per bunch crossing during the 2015 and 2016 data-taking periods. The data shown correspond to periods with stable beams delivered to the ATLAS detector [18].	168
Figure 10.3: The Z' signal efficiency, according to the dielectron selection defined in Section 10.5.1, as a function of Z'_χ pole mass. The selection is maximally efficient at high mass, the primary target of the search. The quick “turn-on” at low mass and slow “turn-off” at high mass are due to the electron kinematics at each of these scales.	179
Figure 10.4: Z'_χ signals post-selection for several characteristic mass points. A low-mass tail begins to develop as the kinematic limit is reached. The number of events within specific invariant mass intervals is displayed in Table 10.3 for the 4 and 5 TeV pole mass Z'_χ	179
Figure 10.5: The data-driven fake background estimate used in the dielectron channel of the dilepton analysis. The estimate is obtained with the matrix method, following the technique outlined in Section 10.6 [4].	185
Figure 10.6: Dielectron invariant mass distribution for candidate dielectron events selected by the full event selection. The stacked sum of Standard Model backgrounds is normalized to 36.1 fb^{-1} and compared to data. Three Z'_χ signals are included for reference, overlaid with the Standard Model expectation. No obvious signal-like excess is present. The lower-panel sub-plot shows the “post-fit” result obtained from the statistical analysis.	189
Figure 10.7: Leading (a) and subleading (b) electron E_T distributions in the dielectron invariant mass region $m_{ee} > 120 \text{ GeV}$, after the full event selection. The stacked sum of Standard Model backgrounds is normalized to 36.1 fb^{-1} and compared to data.	190

Figure 10.8: Leading (a) and subleading (b) electron η distributions in the dielectron invariant mass region $m_{ee} > 120$ GeV, after the full event selection. The stacked sum of Standard Model backgrounds normalized to 36.1 fb^{-1} and compared to data.	191
Figure 10.9: Leading (a) and subleading (b) electron ϕ distributions in the dielectron invariant mass region $m_{ee} > 120$ GeV, after the full event selection. The stacked sum of all Standard Model backgrounds normalized to 36.1 fb^{-1} and compared to data.	192
Figure 10.10: Experimental systematic uncertainties in the electron channel of the dilepton analysis. Shown are effects that shift the Standard Model event yield by greater than 1% anywhere in the dielectron invariant mass search spectrum. The largest experimental uncertainty is due to the $W+\text{jets} \& \text{QCD Multi-jet fake background estimate}$, which rises to 10% by $m_{ee} = 2$ TeV. See Section 10.8.1 for explanations of the source of each systematic uncertainty.	194
Figure 10.11: Theoretical systematic uncertainties in the electron channel of the dilepton analysis (a). Shown are effects that shift the Standard Model event yield by greater than 1% anywhere in the dielectron invariant mass search spectrum. The largest uncertainty is due to the PDF Variation uncertainty, which is also shown decomposed into its individual eigenvector variations (b). See Section 10.8.2 for an explanation of the source of each systematic uncertainty.	199
Figure 10.12: Drell-Yan cross section uncertainty as a function of invariant mass as predicted by several modern PDF sets [4].	204
Figure 11.1: Validation of the signal reweighting procedure. Reweighted signal templates are compared to their dedicated PYTHIA counterparts at pole masses of 2 (a) and 3 (b) TeV for the Z'_χ of the E_6 model. The effect of interference is not included in the dedicated generation, and is similarly not included in the signal reweighting procedure for this comparison. Percent level agreement is seen in the peak region, and the signal shape is effectively captured across the entire invariant mass spectrum.	214
Figure 11.2: Example Minimal model signal template with fixed mixing angle $\theta_{\text{Min}} = 0$ and Z' pole mass $M_{Z'} = 3$ TeV. The template is a function of dielectron mass m_{ee} and the fourth power of the Minimal model relative coupling strength $\gamma' = g_{Z'}/g_Z$, and is generated from reweighted PYTHIA Drell-Yan events after passing the full electron channel event selection.	217

Figure 11.3: Sampling distribution $f(q_0 0)$ for the LLR test statistic q_0 . The distribution is generated from background-only pseudo-experiments and is used to extract the p-value associated with individual Z' signal hypotheses. The distribution follows a χ^2 function with one degree of freedom according to Wilk's theorem [5].	221
Figure 11.4: The resulting p_0 scan in the electron channel search region, derived using Z_χ signal templates with pole masses between 0.15 and 3.0 TeV. The local p-values are depicted as the black curve, where the results for individual mass points have been interpolated. The corresponding local and global significance levels are depicted as gray and red dashed lines, respectively. The lack of sufficient statistics at high mass can result in an ill defined PDF and negative probabilities. Therefore, above 1.5 TeV the so-called “capped” p-value is used, while the “un-capped” value is used below. See Ref. [4] for more details.	224
Figure 11.5: Results of BUMPHUNTER test for the search range $m_{ee} \in [0.12, 3.00]$ TeV. Depicted are the local p_0 of each interval searched (a), and the global p_0 (b), which is extracted from an ensemble of background-only psuedo-experiments.	226
Figure 11.6: The dielectron mass distribution, showing the observed data, the Standard Model background prediction, and the corresponding bin-by-bin local significance. The extremely fine search binning is clearly visible in this plot. The interval identified by the pair of blue lines corresponds to the most significant local excess [5].	227
Figure 11.7: The 95% C.L. upper limits on the Z' production cross section times branching fraction for the electron channel (a) and the combined electron plus muon channel (b) as a function of Z' pole mass. The expected limit is depicted as a black dashed line, with yellow and green bands indicating $\pm 1\sigma$ and $\pm 2\sigma$ variations from the median value. The observed limit is shown as a solid red line and the above region is excluded at 95% C.L. Signal theoretical uncertainties are shown on the SSM cross section prediction as a gray band for reference, but are not included as part of the limit calculation [5].	232
Figure 11.8: The LO cross section times branching fraction predictions for Z' production in the SSM and E_6 models considered in the analysis. Values were obtained using PYTHIA8 configured with the NNPDF2.3 PDF set, and are plotted as a function of Z' pole mass for $\sqrt{s} = 13$ TeV [5].	234

Figure 12.4: The absolute dilepton rapidity $ y_{\ell\ell} $ in several slices in dilepton invariant mass $m_{\ell\ell}$ ranging from 40 GeV to 1 TeV ((a)-(f)). The relative fractions of the u -type and d -type Drell-Yan sub-processes are shown to illustrate the partonic flavor decomposition in this observable.	256
Figure 12.5: The lepton angular distribution $\cos \theta^*$ in several slices of dilepton invariant mass $m_{\ell\ell}$ ranging from 40 GeV to 1 TeV ((a)-(f)). The relative fractions of u -type and d -type sub-processes are shown as well, which exhibit a strong angular dependency, especially at high mass.	257
Figure 12.6: Drell-Yan asymmetry coefficients of Eq. 12.12. The parameter C_q^0 ((a)) describes the magnitude of the symmetric component of the Drell-Yan cross section, while C_q^1 ((b)) describes the magnitude of the asymmetric component.	259
Figure 12.7: Ratio of Drell-Yan asymmetry coefficients C_q^0 and C_q^1 of Eq. 12.12.	260
Figure 12.8: The dilepton invariant mass $m_{\ell\ell}$ as selected by forward ($\cos \theta^* > 0$) or backward ($\cos \theta^* < 0$) events in the CS frame ((a)) and the corresponding PDF uncertainty associated with each selection ((b)). Events arising from the forward-forward lepton topology are not included. As with the other kinematic distributions, the CT14Hera2 PDF set is used. The selection has the effect of modulating the PDF systematic uncertainty.	261
Figure 13.1: ePump requires two sets of inputs to perform an update: the first consists of a set of binned signal templates, each of which is comprised of a nominal theory prediction and n eigenvector variations ($n = 56$ in CT14Hera2), and the second consists of an analogous set of binned data (or pseudo-data) templates, along with their statistical uncertainties. The observed reduction in PDF uncertainty associated with the updated PDF set depends on the statistical precision of the data. Datasets with high statistical precision are better able to constrain the PDF uncertainty.	264
Figure 13.2: The triple-differential cross section of Eq. 13.2 in the range of $m_{\ell\ell} \in [381, 525]$ GeV and $0 < y_{\ell\ell} < 1.8$, scaled to 3000 fb^{-1} . The specific kinematic region is labeled in each plot (a)-(i). The RESBos prediction is indicated by the solid line, and pseudo-data are shown as solid black markers. In each plot, the sub-panel shows the ratio between the theoretical prediction and the measurement. Error bars represent the statistical uncertainty on the pseudo-data and the solid band represents the PDF uncertainty on the dilepton event yield.	274

Figure 13.3: The triple-differential cross section of Eq. 13.2 in the range of $m_{\ell\ell} \in [381, 525]$ GeV and $1.8 < |y_{\ell\ell}| < 3.6$, scaled to 3000 fb^{-1} . The specific kinematic region is labeled in each plot (a)-(i). The RESBos prediction is indicated by the solid line, and pseudo-data are shown as solid black markers. In each plot, the sub-panel shows the ratio between the theoretical prediction and the measurement. Error bars represent the statistical uncertainty on the pseudo-data and the solid band represents the PDF uncertainty on the dilepton event yield. 275

Figure 13.4: Impact of the 3000 fb^{-1} update on the CT14Hera2 $u_v(x)$ (a) and $d_v(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos\theta^*$ templates: “Update3D” at $Q = 3$ TeV. The utilization of the $\cos\theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty. 278

Figure 13.5: Impact of the 3000 fb^{-1} update on the CT14Hera2 $\bar{u}(x)$ (a) and $\bar{d}(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos\theta^*$ templates: “Update3D” at $Q = 3$ TeV. The utilization of the $\cos\theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty. 279

Figure 13.6: Impact of the 3000 fb^{-1} update on the CT14Hera2 $u_v(x)$ (a) and $d_v(x)$ (b) distributions using the double-differential mass and rapidity templates: “Update2D_Rapidity,” and double-differential mass and $\cos\theta^*$ templates: “Update2D_Cosine” at $Q = 3$ TeV. 280

Figure 13.7: Impact of the 3000 fb^{-1} update on the CT14Hera2 $\bar{u}(x)$ (a) and $\bar{d}(x)$ (b) distributions using the double-differential mass and rapidity templates: “Update2D_Rapidity,” and double-differential mass and $\cos\theta^*$ templates: “Update2D_Cosine” at $Q = 3$ TeV. 281

Figure 13.8: The dilepton invariant mass distribution for central-central dilepton events (a), as would be selected by the dilepton event selection of Section 10.5, and with an additional $\cos\theta^* > 0$ requirement added to the selection (b). The ratio sub-plot depicts the CT14Hera2 PDF uncertainty before and after the 3000 fb^{-1} update. 284

Figure 14.1: Expected 95% C.L. upper limits on σB with 300 fb^{-1} of data (a) and 3000 fb^{-1} of data (b). Theory curves are shown for the Z'_ψ , Z'_{LR} , and Z'_{SSM} models. Mass limits can be extracted from the intersection between the expected limit and the theory prediction. The region above each curve is excluded at 95% C.L. [32]. 292

Figure C.1: The ATLANTIS event display for the highest dielectron invariant mass event selected in the 2015 analysis. The “leading” electron has an E_T of 373 GeV and an η of -1.03. The “subleading” electron has an E_T of 246 GeV and an η of 2.45. The invariant mass of the dielectron pair is 1.78 TeV [5].	303
Figure C.2: The ATLANTIS event display for the highest dielectron invariant mass event selected in the 2016 analysis. The “leading” electron has an E_T of 841 GeV, an η of -1.26, and a ϕ of -1.15. The “subleading” electron has an E_T of 655 GeV, an η of 1.32, and a ϕ of 1.83. The invariant mass of the dielectron pair is 2.90 TeV [5].	304
Figure D.1: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [40, 66]$ GeV and $0 < y_{\ell\ell} < 1.8$	306
Figure D.2: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [40, 66]$ GeV and $1.8 < y_{\ell\ell} < 3.6$	307
Figure D.3: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [66, 80]$ GeV and $0 < y_{\ell\ell} < 1.8$	308
Figure D.4: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [66, 80]$ GeV and $1.8 < y_{\ell\ell} < 3.6$	309
Figure D.5: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [80, 91]$ GeV and $0 < y_{\ell\ell} < 1.8$	310
Figure D.6: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [80, 91]$ GeV and $1.8 < y_{\ell\ell} < 3.6$	311
Figure D.7: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [91, 102]$ GeV and $0 < y_{\ell\ell} < 1.8$	312
Figure D.8: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [91, 102]$ GeV and $1.8 < y_{\ell\ell} < 3.6$	313
Figure D.9: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [102, 116]$ GeV and $0 < y_{\ell\ell} < 1.8$	314
Figure D.10: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [102, 116]$ GeV and $1.8 < y_{\ell\ell} < 3.6$	315
Figure D.11: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [116, 145]$ GeV and $0 < y_{\ell\ell} < 1.8$	316

Figure D.12: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [116, 145]$ GeV and $1.8 < y_{\ell\ell} < 3.6$.	317
Figure D.13: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [145, 200]$ GeV and $0 < y_{\ell\ell} < 1.8$.	318
Figure D.14: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [145, 200]$ GeV and $1.8 < y_{\ell\ell} < 3.6$.	319
Figure D.15: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [200, 275]$ GeV and $0 < y_{\ell\ell} < 1.8$.	320
Figure D.16: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [200, 275]$ GeV and $1.8 < y_{\ell\ell} < 3.6$.	321
Figure D.17: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [275, 381]$ GeV and $0 < y_{\ell\ell} < 1.8$.	322
Figure D.18: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [275, 381]$ GeV and $1.8 < y_{\ell\ell} < 3.6$.	323
Figure D.19: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [381, 525]$ GeV and $0 < y_{\ell\ell} < 1.8$.	324
Figure D.20: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [381, 525]$ GeV and $1.8 < y_{\ell\ell} < 3.6$.	325
Figure D.21: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [525, 725]$ GeV and $0 < y_{\ell\ell} < 1.8$.	326
Figure D.22: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [525, 725]$ GeV and $1.8 < y_{\ell\ell} < 3.6$.	327
Figure D.23: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [725, 1000]$ GeV and $0 < y_{\ell\ell} < 1.8$.	328
Figure D.24: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [725, 1000]$ GeV and $1.8 < y_{\ell\ell} < 3.6$.	329
Figure D.25: Impact of the 300 fb^{-1} update on the CT14Hera2 $u_v(x)$ (a) and $d_v(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos \theta^*$ templates: “Update3D” at $Q = 3$ TeV. The utilization of the $\cos \theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty.	330

Figure D.26: Impact of the 300 fb^{-1} update on the CT14Hera2 $\bar{u}(x)$ (a) and $\bar{d}(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos\theta^*$ templates: “Update3D” at $Q = 3 \text{ TeV}$. The utilization of the $\cos\theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty.	331
Figure D.27: The dilepton invariant mass distribution for central-central dilepton events (a), as would be selected by the dilepton event selection of Section 10.5, and with an additional $\cos\theta^* > 0$ requirement added to the selection (b). The ratio sub-plot depicts the CT14Hera2 PDF uncertainty before and after the 300 fb^{-1} update.	333
Figure E.1: Run-1 dielectron invariant mass distribution (a) and electron+MET transverse mass distribution (b), the primary search variables in the dilepton and lepton+MET analyses, respectively [33, 34].	336
Figure E.2: Run-1 95% C.L. lower mass limits in the dilepton analysis (a) and lepton+MET analysis (b) [33, 34].	337
Figure E.3: The EWS applied to the full dataset collected during Run-1 of the LHC. The fit is performed in the low p_T region of $[125, 300] \text{ GeV}$ and extrapolated to the monitoring region to assess the high- p_T data. Three signals are overlaid using simulation samples: a $1.5 \text{ TeV } Z'_{\text{SSM}}$, a $2.5 \text{ TeV } Z'_{\text{SSM}}$, and a $3.0 \text{ TeV } W'_{\text{SSM}}$, each of which displays a characteristic Jacobian peak in the p_T spectrum.	340
Figure E.4: Operation of the Run-2 EWS in the electron channel: May to June 2016 (a)-(f).	342
Figure E.5: Operation of the Run-2 EWS in the electron channel: June to July 2016 (a)-(f).	343
Figure E.6: Operation of the Run-2 EWS in the muon channel: May to June 2016 (a)-(f).	344
Figure E.7: Operation of the Run-2 EWS in the muon channel: June to July 2016 (a)-(f).	345

Part I

The Theory

Chapter 1

Introduction

The idea that matter is composed of a finite number of elementary constituents is very old and dates back to at least the 5th century BCE in Ancient Greece. It was during this period that the **atomic hypothesis** was first formulated Democritus and his mentor Leucippus, that all of matter was composed of small, indivisible particles called atoms.¹ However, this was one of many competing ideas about the fundamental nature of matter at the time , and it would take another 2400 years for modern experiments to test for and later establish the existence of atoms.

In the 19th century, John Dalton proposed the modern form of the atomic hypothesis, that each chemical element found in nature was composed of indivisible atoms, which can combine, separate, and rearrange in chemical reactions, and form the true building blocks of matter. Early into the 20th century, however, it was realized that atoms themselves were not indivisible as originally thought; rather they comprised even more fundamental particles: electrons, protons, and neutrons.

J.J. Thompson discovered the electron in 1897 through his study of cathode rays [35] and Ernest Rutherford discovered the atomic nucleus in 1911 and later identified the proton through early α particle scattering experiments [36]. The following discovery of the neutron [37] by James Chadwick in 1932 seemed to complete the picture of atoms being

¹The word “atom” derives from the Ancient Greek $\alpha\tauομος$, *atomos*, i.e., “uncuttable,” or “indivisible.”

composed of three “elementary particles.”

Over the subsequent decades, hundreds of new particles were discovered and each was thought to be elementary in its own right, collectively dubbed the “Particle Zoo.” Attempts at a systematic classification lead to the development of the **quark model** [38, 39], which proposed that the known strongly interacting particles were actually bound states of truly elementary quarks.

The prediction and subsequent discovery of the Ω^- baryon in 1964 at Brookhaven [40] was one of the early successes of the quark model. In the late 1960s, deep inelastic scattering experiments at the Stanford Linear Accelerator (SLAC) [41, 42], where highly energetic electrons were scattered off of nuclei, showed that the proton had an internal structure, providing strong evidence for existence of quarks.

As technology progressed, increasingly powerful particle accelerators were developed, allowing for ever higher energies to be achieved and ever smaller sub-nuclear distances to be probed. This led to the production and discovery of more new and exotic particles, and, driven by the interplay between experimental observations and theoretical predictions, the **Standard Model of Particle Physics** began to take shape.

The gluon, the force mediator of strong interaction, was discovered in 1979 with the Positron-Elektron Tandem Ring Anlage (PETRA) at the Deutsches Elektronen-Synchrotron (DESY) [43], and four years later in 1983, the W^\pm and Z bosons, the mediators of the weak interaction, were discovered with the Super-Proton-Antiproton Synchrotron (Sp \bar{p} S) at the Conseil Européen pour la Recherche Nucléaire (CERN) [44, 45, 46, 47]. The top quark was discovered in 1995 by the DØ and CDF collaborations at the Tevatron [48, 49], the tau neutrino in 2000 by the DONUT collaboration [50], and the Higgs boson in 2012 by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) at CERN [51, 52], all

of which were predicted by the Standard Model.

The Standard Model has been precisely tested over the course of several decades, and is able to describe all current experimental observations in particle physics, making it an extraordinarily successful theory. It explains the structure of matter, classifies the known elementary particles, and accounts for three of four **fundamental forces of nature**.²

The precision and scope of Standard Model are indeed astonishing; however, despite its enormous experimental success, there are several classes of observed phenomena that it cannot obviously explain. These include the existence and nature of dark matter and dark energy, the baryon asymmetry of the universe, and the fundamental nature of neutrinos, to name a few. The Standard Model also suffers from several theoretical, or conceptual, problems, such as how to consistently incorporate gravity into the theory. Taken together, these observations suggest that the Standard Model is incomplete and some kind of extension or modification is required. If the 20th century was about establishing foundational principles, the 21st century is about discovering what lies beyond.

Today, searches are performed for new particles and interactions not predicted by the Standard Model, but by theories beyond that attempt to address the limitations mentioned above by extending it in various ways. Since the Standard Model has been precisely tested for several decades over an enormous range of energies, new physics phenomena are expected to occur very rarely, to be produced only at high energies, or both. However, as the history of particle physics has shown, searches performed at high-energy colliders are a very promising way to discover new particles, and many theories predict new phenomena at the TeV scale, possibly within the reach of the LHC.

The LHC, a proton-proton collider located outside Geneva, Switzerland at CERN, holds

²The strong and weak nuclear forces, electromagnetism, and gravity.

the record as the highest energy particle collider constructed thus far, and as such, offers a unique opportunity to search *directly* for the production of new physics. Four specialized particle detectors, such as ATLAS (A Toroidal LHC ApparatuS), are used to detect, record, and reconstruct the assemblage of particles that originate from the high-energy proton-proton collisions provided by the LHC.

The focus of this dissertation is the search for a new heavy electrically neutral spin-1 boson,³ known colloquially as a Z' , which is produced in the proton-proton collisions at the LHC and decays into a pair of charged leptons. For the search considered here, the final-state consists of an electron and a positron, or a **dielectron pair**; i.e., $pp \rightarrow Z' \rightarrow e^+e^- + X$. The search is performed with the full dataset collected by the ATLAS detector during the 2015 and 2016 runs of the LHC at a center-of-mass (CM) energy of $\sqrt{s} = 13$ TeV, allowing for a previously unexplored energy scale to be probed for the first time.

As a Z' would decay too quickly to be detected directly in the ATLAS detector, events with highly energetic dielectron pairs are searched for instead. The smoking-gun signature of Z' production would be the appearance of a narrow resonant peak in the dielectron invariant mass spectrum, where the position of the peak would depend on the mass of the Z' boson. Its specific properties, such as its decay width $\Gamma_{Z'}$, production cross section $\sigma_{Z'}$, and branching fraction to dielectron pairs $Br(Z' \rightarrow e^+e^-)$, all depend on the particular model under consideration.

While the Z' is not predicted by Standard Model, it is predicted by several well-motivated extensions, such as the E_6 model [53], a **Grand Unified Theory**.⁴ Thus, discovering a Z' would contribute significantly to our fundamental understanding of the elementary particles

³A particle with integer spin.

⁴Grand Unified Theories (GUTs) predict the three gauge interactions of the Standard Model are unified by a single gauge symmetry at very high energy.

and their interactions, and would help to address some of the limitations of the Standard Model outlined above. In the case of non-observation, data are used to set exclusion limits on proposed extensions of the Standard Model, thereby restricting their ranges of validity.

In order to perform a robust search for new physics like the one presented here, precise predictions of the Standard Model processes that result in dielectron pairs are required. This in turn requires a precise understanding of the **Parton Distribution Functions** (PDFs) of the proton, which describe its internal structure, and are therefore instrumental in the modeling of the high-energy proton-proton collisions at the LHC. While the discovery potential of the search presented here is not strictly limited by the current precision of the PDFs, the ability to discriminate between different Z' models in a discovery scenario and to set more powerful exclusion limits in future iterations of the search will be strongly affected unless the current uncertainty of the PDFs can be sufficiently reduced. In the final chapter of this dissertation, an exploratory study is performed with precisely this goal in mind. The dissertation is organized as follows:

Part I: The theoretical foundations and predictions needed for the subsequent chapters are described.

- Chapter 1: Introduction
- Chapter 2: The particles and forces of the Standard Model are introduced, along with its mathematical formulation, including the Electroweak Theory, the Higgs Mechanism, and Quantum Chromodynamics (QCD).
- Chapter 3: The physics of LHC proton-proton collisions are described and PDFs are introduced.
- Chapter 4: The primary Standard Model background process to the dilepton search is

due to Drell-Yan production. Drell-Yan physics is introduced and its production cross section is calculated using the diagrammatic technique developed by Feynman. The effect of the PDF uncertainty on the predicted high-mass Drell-Yan cross section is qualitatively assessed.

- Chapter 5: The limitations of the Standard Model are described. Several theories of new physics that predict high-mass dilepton resonances are introduced. Candidate models searched for in this dissertation are presented along with their potential experimental signatures. Exclusion limits on Z' production are provided for recent experiments.

Part II: The LHC and ATLAS detector are described, along with electron reconstruction and identification, and the Monte Carlo simulation used to model the signal and Standard Model background processes present in the search.

- Chapter 6: The operation and performance of LHC is presented.
- Chapter 7: The ATLAS experiment is introduced with an emphasis on detector subsystems relevant for measuring electrons.
- Chapter 8: The reconstruction and identification of electron candidates from inner detector tracks and calorimeter energy clusters in the ATLAS detector are explained. Corrections applied to electrons in data and simulation are also described.
- Chapter 9: Simulation samples used to model the signal and background processes contributing to the analysis are described, along with their various corrections.

Part III: The search for new physics in the dielectron final-state is performed and the results of the search are interpreted and presented.

- Chapter 10: The search for dielectron resonances is performed. The estimate of the QCD Multijet & $W+jets$ fake background is obtained using a data-driven technique. The dielectron event selection is described and data are compared directly to the Standard Model background estimate. Event yields are quantified and several kinematic distributions are shown, including the dielectron invariant mass. Systematic uncertainties are assessed.
- Chapter 11: The data in the dielectron invariant mass distribution are scrutinized for the presence of a signal. If no significant excess is found, a Bayesian technique is used to set 95% C.L. upper limits on the cross section times branching fraction for Z' production, as well as on several other Z' model parameters of interest.

Part IV: The systematic uncertainty of PDFs is studied and a novel approach is developed that can be used to reduce it for future analyses.

- Chapter 12: The Drell-Yan process is revisited in an attempt to better understand the nature of the high-mass PDF uncertainty that arises in the primary search region of the dielectron invariant mass spectrum. The $\cos \theta^*$ variable is identified as potentially useful in a future PDF global analysis.
- Chapter 13: A novel technique is developed to reduce the high-mass PDF uncertainty. Results are quantified and considered in the context of a future dilepton search.

Part V: The conclusion to the dissertation is given, which includes a summary of the results obtained and an outlook to the future of the dilepton search and LHC program.

- Chapter 14: Conclusion

Chapter 2

The Standard Model

In this chapter, an introduction to the Standard Model of Particle Physics is given. The fundamental particles and interactions are described, its mathematical formulation is presented, and the theories of the Electroweak interaction, the Higgs Mechanism, and Quantum Chromodynamics are explained in detail.

2.1 Fundamental Particles and Interactions

The Standard Model of particle physics is a gauge-invariant, relativistic quantum field theory (QFT). It describes the dynamics of and interactions between all currently known elementary particles, and successfully incorporates three of the four fundamental interactions into its theoretical framework.¹ Since its inception in the early 1970s, the Standard Model has passed every experimental test and precisely predicted a wide range of phenomena, making it today one of the best established and well-tested theories in the history of science.

Broadly speaking, the particles of the Standard Model can be divided two classes: there are the matter particles and the force-carriers, or mediators. The matter particles are **fermions**,² which interact through the exchange of force mediating **gauge bosons**.³ In

¹The 4th, Gravity, is explained by Einstein's General Theory of Relativity as the warping of the geometry of space-time.

²Fermions are particles with half-integer spin.

³Bosons are particles with integer spin. All force mediating particles in the Standard Model are spin-1.

addition, the Standard Model contains a spin-0 scalar boson, known as the Higgs boson, as a consequence of the **Higgs Mechanism** (See Sec. 2.3.4), which is responsible for the generation of all elementary particle masses in the model.

2.1.1 Matter Particles

The matter particles, the basic building blocks of matter, can be further subdivided into two categories according to their allowed interactions: these are the **leptons** and the **quarks**.

Table 2.1 lists the properties of each matter particle in the Standard Model.⁴

There are six leptons: three charged leptons, the electron (e), the muon (μ), and the tau (τ), each of which carries an electric charge of $Q = -1 e$, and three neutral leptons, the electron-neutrino (ν_e), the muon-neutrino (ν_μ), and the tau-neutrino (ν_τ), each of which carries an electric charge of zero. Thus, the charged leptons can interact electromagnetically, while the neutral leptons can not. Furthermore, each leptons carries a weak isospin t_3 and weak hypercharge quantum number Y , and are thus susceptible to the weak interactions as well.

Analogous to the leptons, there are six quarks: three **up-type** quarks, the up (u), the charm (c), and the top (t), each of which carries an electric charge of $Q = +2/3 e$, and three **down-type** quarks, the down (d), the strange (s), and the bottom (b), with $Q = -1/3 e$. Quarks are not found in isolation like the leptons, rather they form composite structures called **hadrons**, which come in two varieties: two-quark state **mesons**, and three-quark state **baryons**, the most well known being the proton (uud), and neutron (ddu). Not only do the quarks have weak isospin t_3 and weak hypercharge Y , they also carry one of three

⁴Natural units will be used throughout this dissertation, such that $\hbar = c = 1$. As a result energy and mass have the same dimension.

Particle Name	Symbol	EM Charge [e] (Q)	Weak Charge (t_3)	Strong Charge (Color)	Mass [MeV]
electron neutrino	ν_e	0	+1/2	No	$< 2 \cdot 10^{-6}$
	e	-1	-1/2	No	0.511
muon neutrino	ν_μ	0	+1/2	No	< 0.19
	μ	-1	-1/2	No	106
tau neutrino	ν_τ	0	+1/2	No	< 18.2
	τ	-1	-1/2	No	1777
up	u	+2/3	+1/2	Yes	2.2
	d	-1/3	-1/2	Yes	4.7
charm	c	+2/3	+1/2	Yes	96
	s	-1/3	-1/2	Yes	1.3
top	t	+2/3	+1/2	Yes	$173.1 \cdot 10^3$
	b	-1/3	-1/2	Yes	$4.2 \cdot 10^3$

Table 2.1: Matter content of the Standard Model. The three generations of quarks and leptons are shown along with their name, symbol, charges under each interaction, and mass [1].

possible color charges, called “red,” “green,” or “blue,” as opposed to the singular charge familiar in electromagnetism, signifying their possible interaction through the strong force.

In addition, pairs of elementary fermions form *three* distinct **generations**. The first generations of leptons is defined by the (ν_e , e) pair, the second by the (ν_μ , μ) pair, and the third by the (ν_τ , τ) pair, with a similar pairing for the quarks. Each particle within a generation has greater mass than the corresponding particle of lower generation, but their possible interactions are identical. Why each generation should be replicated like this in nature is currently unknown.

The first generation of matter particles is stable, forming all ordinary matter in the universe. Atoms, for example, are composed of electrons orbiting atomic nuclei, themselves composed of protons and neutrons, which are ultimately made from bound states of up and down quarks. On the other hand, the second and third generations are only seen in high-energy environments like cosmic rays, particle accelerators, and the early universe [54], and

Particle Name	Symbol	EM Charge [e] (Q)	Weak Charge (t_3)	Strong Charge (Color)	Mass [MeV]
photon	γ	0	0	No	0
W -boson	W^\pm	± 1	± 1	No	$80.3 \cdot 10^3$
Z -boson	Z	0	1	No	$91.2 \cdot 10^3$
gluon	g	0	0	Yes	0
Higgs	H	0	-1/2	No	$125.5 \cdot 10^3$

Table 2.2: Force carriers and Higgs boson of the Standard Model. The name, symbol, charges under each interaction, and mass of each particle is listed [1].

quickly decay into the lower generation particles. Neutrinos of all generations are apparently stable due to their extremely small masses, but they rarely interact with normal matter. Fermions across generations are referred to as **flavors**, such that, e.g., the electron and muon are distinct flavors of charged leptons.

Each quark and lepton has an **antimatter** counterpart with identical values for mass and spin, but an opposite electric charge, weak charge, and color charge. The most well known antimatter particle is the anti-electron, or positron, which was first theorized by Dirac in 1928 [55] and discovered by Anderson in 1932 [56] from his analysis of cosmic ray tracks.

2.1.2 Forces and Interactions

There are four known fundamental forces in nature: gravitation, the electromagnetic force, and the weak, and strong nuclear forces. The Standard Model describes the later three, with the incorporation of gravity into the theory still an open problem.⁵ Each force can be attributed to the exchange of one or more spin-1 gauge bosons: the photon γ , the massive W^\pm and Z bosons, and the gluon g , whose properties are listed in Table 2.2. The photon γ is the mediator of the electromagnetic force. It is responsible for all forms of electromagnetic

⁵The influence of gravity at the currently probed subatomic energy scales is entirely negligible, and, as a consequence, its incorporation into the theory not strictly necessary to make accurate predictions.

radiation, including heat and light, as well as the phenomena of electricity, and magnetism. The photon is massless and its range of influence is infinite, diminishing in influence as $1/r^2$ according to Coulomb’s Law. Only particles with an electric charge couple to the photon, which includes all matter particles except the neutrinos. The photon itself does not have a charge.

The W^+ , W^- , and Z bosons are mediators of the weak interaction, responsible for radioactive β -decay and any process involving neutrinos. Weak interactions can proceed through a “charged current” interaction, mediated by the W^+ or W^- bosons, or though a “neutral current” interaction, mediated by the Z boson. The β^- decay of the neutron $n \rightarrow p e^- \bar{\nu}_e$, for example, is mediated by a W^- boson, in which the down quark d of the neutron is converted into an up quark u , and the W^- subsequently decays into an electron e^- and electron antineutrino $\bar{\nu}_e$. Likewise, the exchange of a Z boson can occur in neutrino elastic scattering reactions such as $\nu_\mu e^- \rightarrow \nu_\mu e^-$.

The weak bosons are massive. The mass of the W^+ and W^- bosons are identical with $M_W \simeq 80.3$ GeV and the mass of the Z boson is slightly larger with $M_Z \simeq 91.2$ GeV, which limits their range of influence to roughly 10^{-18} m, or about 0.1% the diameter of a proton. All matter particles carry weak charge and can therefore interact through the exchange of a weak boson.

The mediators of the strong interaction are a set of eight massless, electrically neutral gluons g . The strong force holds matter together, confining quarks into “color neutral” bound-states of hadrons. As a residual effect, the strong force is responsible for the nuclear force that binds protons and neutrons together and as a result allows for the formation of stable atomic nuclei. Gluons themselves carry color charge, and can therefore couple to quarks and to each other. Despite being massless, the fact that gluons can self-couple leads

to the observed short-ranged influence of the interaction, a distance of roughly 1 fm, or the size of an atomic nucleus.

2.1.3 Particle Masses

In its simplest form, the Standard Model predicts that all particles should be massless, in strong contradiction with experiment. To overcome this obstacle, a new spin-0, or scalar, field is introduced into the theory, the Higgs field. Through the Higgs Mechanism (See Sec. 2.3.4), the masses of the gauge bosons and fermions are generated from **spontaneous symmetry breaking** induced by the Higgs field. The result is the Higgs boson H , and the Standard Model particles that interact with the Higgs field obtain a mass.

The properties of the Higgs boson can be seen in Table 2.2. The Higgs boson has been searched for since its original proposal by Higgs [57], Englert and Brout [58], and Guralnik, Hagen, and Kibble [59] independently in 1964. It was finally discovered in 2012 by the ATLAS and CMS collaborations at the LHC at CERN.

2.2 Mathematical Formulation

The Standard Model is formulated as a gauge-invariant quantum field theory (QFT) [60]. The combination of quantum mechanics and special relativity leads to the introduction of quantum fields and their associated field quanta, or particle [54]. All particles in the Standard Model are described as excitations of underlying quantum fields, which are operator valued and defined at each point in space-time. The fermions in the model are defined by fermion, or Dirac, fields $\psi(x)$, gauge bosons by vector fields $A_\mu(x)$, and the Higgs field by a scalar field $\phi(x)$. The dynamics and evolution of the fields is described using a Lagrangian formalism,

which is introduced in the following section.

In addition, the Standard Model is a gauge theory⁶, meaning the Lagrangian can remain invariant with respect to certain symmetry transformation applied to the fields. The generators of the transformations form a group, and the gauge group of the Standard Model is $SU(3)_C \times SU(2)_L \times U(1)_Y$.

2.2.1 Lagrangian Dynamics

The Standard Model is usually expressed in a Lagrangian formalism. The quantity that encodes the possible dynamics of the fundamental fields is the action \mathcal{S} , with

$$\mathcal{S}[\phi] = \int L dt = \int \mathcal{L}(x^\mu, \phi, \partial_\mu \phi) d^4x, \quad (2.1)$$

where $L = T - V$ is the Lagrangian. In quantum field theory, it is customary to work with the Lagrangian density \mathcal{L} .⁷ Here, the Lagrangian is written $\mathcal{L} = \mathcal{L}(x^\mu, \phi, \partial_\mu \phi)$, which is a function of the space-time dependent quantum field $\phi(x^\mu)$, the derivative $\partial_\mu \phi(x^\mu)$, and the standard space-time four-vector x^μ , with $x^0 = ct$, and $x^i = \mathbf{x}$.⁸

Possible terms in the Lagrangian include kinetic terms, which describe the propagation of free fields, mass terms, which specify particle masses, and interaction terms, which couple different particles together. The Lagrangian is often said to *define* the theory, since, using the rules of quantum field theory, all directly observable quantities, such as scattering cross section and particle decay widths, can be calculated from it.

⁶Additional degrees of freedom exist in the definition of the Lagrangian that do not necessarily correspond to changes in the physical state of the system.

⁷The Lagrangian density is commonly referred to as the Lagrangian, which will be the case here, and throughout the rest of the dissertation.

⁸The field theory definition effectively generalizes that from classical mechanics, where the Lagrangian $L = L(t, q, \dot{q})$ is a function of time, t , the generalized coordinates q , and its derivative \dot{q} .

A physical system evolves so that the action S is minimized. By requiring $\delta\mathcal{S} = 0$, the Euler-Lagrange Equation is derived,

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \quad (2.2)$$

which can be used to obtain the equations of motion for the space-time evolution of the quantum fields. In the case of electromagnetism, these are known as Maxwell's equations, and in the case of the weak and strong interactions, these are known as the Yang-Mills equations.

2.2.2 Symmetries of the Standard Model

The Standard Model is said to exhibit a symmetry if the form of its Lagrangian is left invariant by a transformation applied to its fields. Symmetries are described by symmetry groups, such as the group of translation or group of rotations. If the symmetry is continuous, Noether's Theorem [61] states there must be a conserved current, and hence a conserved charge. For example, the conserved “charge” of a rotationally invariant system is its angular momentum.

The Standard Model broadly exhibits two kinds of symmetries: external and internal. External symmetries are related to space-time transformations. These can be continuous, as in the Lorentz transformations, rotations, and boosts that form the Poincaré group, or discrete. The discrete transformations include **parity-inversion**, \mathcal{P} (changing the sign of the spatial coordinates, $\mathbf{r} \rightarrow -\mathbf{r}$), **charge-conjugation**, \mathcal{C} (changing the sign of charges), and the operation of **time-reversal**, \mathcal{T} (changing the sign of the time coordinate, $t \rightarrow -t$).

The electromagnetic and strong interactions are invariant under, and hence conserve,

both \mathcal{P} and \mathcal{CP} ; the weak interaction, on the other hand, conserves neither. It turns out the Standard Model is not invariant under any one of the discrete transformations, nor any successive two, however it is invariant under the combined operation of \mathcal{CPT} . The internal, or gauge symmetries, are discussed in the next section.

2.2.3 Gauge Symmetries

The Standard Model Lagrangian is invariant under a set of local (i.e. space-time dependent) symmetry transformations of the fields, known as **gauge transformations**. Gauge symmetries are some of the most consequential symmetries for the Standard Model, as they *mandate* the existence of the force carriers.

The following illustrates the consequences of requiring a theory to be locally gauge invariant in the context of Quantum Electrodynamics (QED). The Lagrangian describing the motion of a *free* spin-1/2 fermion field $\psi(x)$ of mass m is

$$\mathcal{L} = \bar{\psi}(x) (i\cancel{\partial} - m) \psi(x). \quad (2.3)$$

Here, $\psi(x)$ refers to the electron, $\bar{\psi}(x) = \psi^\dagger \gamma^0$ to the positron, and $\cancel{\partial} \equiv \gamma_\mu \partial^\mu$ denotes the Lorentz contraction between γ^μ and ∂_μ , where γ^μ are the 4×4 Dirac γ matrices,

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad (2.4)$$

which themselves are composed of the 2×2 Pauli spin matrices σ^i .

The Lagrangian of Eq. 2.3 is left invariant under a rescaling of the fields by a global phase factor α . To see this explicitly, if $\psi(x) \rightarrow \psi'(x) = e^{iQe\alpha} \psi(x)$ is substituted into Eq. 2.3,

then $\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L}$. Thus, the Lagrangian is globally gauge invariant. The phase factor $e^{iQe\alpha}$ is an element of the symmetry group $U(1)_Q$ of rotations in one dimension. The charge operator Q is generator of the group, and returns the magnitude of electric charge associated with the field $\psi(x)$.

Taking this a step further and demanding that *local* gauge invariance holds, i.e., that the Lagrangian of Eq. 2.3 remain invariant under a transformation where the phase $\alpha(x)$ now depends on local space-time coordinates, results in

$$\begin{aligned}\mathcal{L} \rightarrow \mathcal{L}' &= \overline{\psi'} (i\gamma^\mu \partial_\mu - m) \psi' \\ &= e^{+iQe\alpha(x)} \overline{\psi} (i\gamma^\mu \partial_\mu - m) e^{-iQe\alpha(x)} \psi \\ &= \underbrace{\overline{\psi} (i\gamma^\mu \partial_\mu - m) \psi}_{\text{original Lagrangian}} - \underbrace{\overline{\psi} (Qe\gamma^\mu \partial_\mu \alpha(x)) \psi}_{\text{derivative term}}.\end{aligned}\tag{2.5}$$

The Lagrangian of Eq. 2.5 is no longer invariant under the local $U(1)$ transformation due to resulting derivative term associated with $\alpha(x)$. However, gauge invariance can be restored with the following substitution,

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - ieQA_\mu,\tag{2.6}$$

where D_μ is known as the **gauge covariant derivative**, which is constructed by adding a new spin-1 field A_μ to the standard definition of the four-derivative ∂_μ . If the new field A_μ is required to transform like

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \alpha(x),\tag{2.7}$$

during a gauge transformation, then the $\partial_\mu \alpha(x)$ term in Eq. 2.7 exactly compensates for the derivative term in Eq. 2.5 and thus, the requirement of local gauge invariance is satisfied.

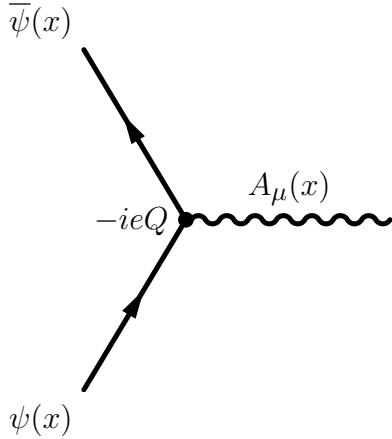


Figure 2.1: The QED interaction vertex. The solid lines represent the motion of the charged leptons and the curved line represent the propagation of the photon. The strength of the interaction is determined by the magnitude of electric charge Q in units of fundamental charge e and is labeled at the vertex.

The “new” field A_μ is none other than the photon field, the mediator of the electromagnetic force. The insertion of Eq. 2.6 into Eq. 2.5 results in the following Lagrangian,

$$\begin{aligned}\mathcal{L} &= \bar{\psi}(x) (i\gamma^\mu D_\mu - m) \psi(x) \\ &= \bar{\psi}(x) (i\gamma^\mu \partial_\mu - m) \psi(x) - Qe\bar{\psi}(x)\gamma^\mu A_\mu \psi(x)\end{aligned}\tag{2.8}$$

where the second term corresponds to the *interaction* between the fermion fields $\psi(x)$ and $\bar{\psi}(x)$ and the photon field A^μ , which is represented pictorially in Figure 2.1. As the photon can propagate in its own right, an additional kinetic term can be added to 2.8. The result is the full QED Lagrangian,

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(x) (i\gamma^\mu \partial_\mu - m) \psi(x) - Qe\bar{\psi}(x)\gamma^\mu A_\mu \psi(x)\tag{2.9}$$

where $F^{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the usual electromagnetic field strength tensor. Incidentally, the photon field A^μ must be massless, since any term of the form $\mathcal{L}_{\text{mass}} = \frac{1}{2}m^2 A^\mu A_\mu$ would violate the local gauge invariance of Eq. 2.9.

Thus, the requirement that the free field theory of Eq. 2.3 be invariant with respect to local gauge transformations, forces the introduction of a new massless field A_μ , understood to be the photon, and recovers the complete theory of quantum electrodynamics. In fact, this procedure of “promoting” a free field theory to an interacting theory is exactly how each of the fundamental interactions of the Standard Model are generated.

The gauge group of the Standard Model is $SU(3)_C \times SU(2)_L \times U(1)_Y$. Invariance under $SU(3)_C$ leads to the introduction of the gluons of the strong interaction, while invariance under $SU(2)_L \times U(1)_Y$ leads to the force mediators of the unified Electroweak interaction.

2.2.4 Standard Model Lagrangian

The Standard Model Lagrangian can be broken down into four sectors,

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{QCD}}. \quad (2.10)$$

The following sections describe the Lagrangian of each sector in detail. The dynamics of the unified Electroweak interaction are encoded in \mathcal{L}_{EW} , while $\mathcal{L}_{\text{Higgs}}$ and $\mathcal{L}_{\text{Yukawa}}$ describe the dynamics of the Higgs field, and how fermion masses are generated, respectively. The last term, \mathcal{L}_{QCD} , describes the dynamics of the strong interaction. The electromagnetic component, \mathcal{L}_{QED} , is included in the unified description of the Electroweak sector, and becomes apparent only after the spontaneous symmetry breaking induced by the Higgs field.

2.3 The Unified Electroweak Theory

The Electroweak (EW) theory unifies the weak and electromagnetic interactions, which historically were treated as two separate theories. The unification of these theories was first described by Weinberg, Glashow, and Salam, earning them the 1979 Nobel Prize in Physics. The Electroweak theory is invariant under simultaneous $SU(2)_L \times U(1)_Y$ gauge transformations, where L refers to “left-handed” particles, with the weak isospin \mathbf{t} as the generator of $SU(2)_L$ transformations, and the weak hypercharge Y as the generator of $U(1)_Y$ transformations. Under the action of the Higgs Mechanism, the Electroweak group is spontaneously broken, such that $SU(2)_L \times U(1)_Y \rightarrow U(1)_Q$, which recovers the theory of electromagnetism in the low energy limit.

2.3.1 History

Enrico Fermi proposed the first theory of the weak interaction in 1934 in order to explain radioactive β decays [62, 63] of nuclear physics. The original theory proposed by Fermi was subsequently generalized over the course of many decades to account the new and varied observations associated with weak interaction physics.

C.S. Wu discovered parity violation in 1956 by measuring the beta decay spectrum of Co-60 [64] and around the same time James Cronin and Val Fitch discovered \mathcal{CP} violation in their analysis of the neutral Kaon system [65]. The discovery of parity violation suggested Fermi’s theory must be modified. In 1958, Feynman and Gell-Mann proposed a V-A (vector minus axial-vector) Lagrangian for the theory, so that the weak interaction acted only on left-handed particles (and right-handed antiparticles) [66, 67].

The mathematical framework for the modern formulation of the theory was laid down

by Yang and Mills in 1956 by extending the classic idea of gauge symmetry to non-abelian field theories [68]. Building on the work of Yang and Mills, in 1961 Glashow linked the weak and electromagnetic forces by formulating a theory with three *massless* gauge bosons, two charged and one neutral, the progenitors of the W^\pm and Z bosons of the Standard Model, and the photon γ [69]. The only massless field that had been observed in nature at the time was the photon, and the short range of the weak interaction indicated its force carriers must be massive.

This problem of massless force carriers was solved when in 1967 Weinberg [70], and in 1968 Salam [71], incorporated the Higgs Mechanism into Glashow's theory, thereby attributing the gauge bosons mass. The result was a theory that predicted one massless particle, the photon γ , and three new massive W^\pm and Z bosons, the carriers of the weak interaction.

The first evidence for this theory came in 1973 when the Gargamelle bubble chamber at CERN provided indirect evidence for the weak neutral current from the elastic scattering of neutrinos in matter [72, 73]. For their work, Glashow, Weinberg, and Salam (GWS) shared the 1979 Nobel Prize in physics. Four year later, direct evidence came in 1983 with the subsequent discoveries of the W^\pm and Z bosons by the UA1 and UA2 collaborations at CERN [44, 45, 46, 47], marking one of the major intellectual achievements of humankind. It would take until 2012 for the Higgs boson, the last particle predicted by the GWS theory, to be discovered at CERN by the ATLAS and CMS collaborations [51, 52].

2.3.2 Fermion Fields

The force mediators of the weak interaction couple with different strengths to left-handed and right-handed fermion fields; and the charged current couples *exclusively* to left-handed fermions. The left- and right-handed components of a fermion fields can be obtained from

the projection operators P_L and P_R ,

$$\psi_L = P_L \psi = \frac{1}{2} (1 - \gamma^5) \psi, \quad \psi_R = P_R \psi = \frac{1}{2} (1 + \gamma^5) \psi, \quad (2.11)$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is a product of the 4×4 Dirac γ matrices. Because of the asymmetric treatment of particles and antiparticles on the basis of their handedness, the Electroweak interaction is said to be chiral.

The Electroweak Theory introduces the weak isospin quantum number \mathbf{t} . This allows all left-handed fermions to be arranged into $SU(2)_L$ iso-doublets of $\mathbf{t} = \mathbf{1/2}$ with an upper component of $t_3 = 1/2$ and lower component of $t_3 = -1/2$. The right-handed fermions are arranged into $SU(2)_L$ iso-singlets with $\mathbf{t} = \mathbf{0}$ and $t_3 = 0$, which is a consequence of their non-participation in the charged-current interaction. The right-handed neutrinos (and left-handed antineutrinos) are not included with the rest of the singlets.⁹ Thus, in the Electroweak theory, the fermions are described as

$$Q^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}, \quad u_R^i, \quad d_R^i, \quad L^i = \begin{pmatrix} \nu_L^i \\ e_L^i \end{pmatrix}, \quad e_R^i, \quad (2.12)$$

where the index i refers to individual fermion generation.

Furthermore, regardless of their handedness, all fermions in the theory are given an intrinsic weak hypercharge Y , which is introduced ultimately to accommodate the electromagnetic interaction. Table 2.3 details the fields and their associated quantum numbers. The standard electric charge Q of a fermion can be expressed in terms of the third component of weak

⁹At the time of the construction of the theory, there was no experimental evidence that neutrinos had mass and they could therefore be neglected from the theory; however the discovery of neutrino oscillations suggests a more complex story. See Section 5.1 for more details.

	Fermion Generation			Quantum Number		
	I	II	III	Q	t_3	Y
Quarks	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	+2/3	+1/2	+1/6
	u_R	c_R	t_R	-1/3	-1/2	+1/6
	d_R	s_R	b_R	+2/3	0	+2/3
				-1/3	0	-1/3
Leptons	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$	0	+1/2	-1/2
	e_R^-	μ_R^-	τ_R^-	-1	-1/2	-1/2
	ν_R	$\nu_{\mu R}$	$\nu_{\tau R}$	-1	0	-1
				0	0	0

Table 2.3: Three generations of fermions in the Electroweak theory with their respective quantum numbers. Left-handed particles are grouped into iso-doublets, while right-handed particles exist as iso-singlets. The right-handed neutrinos are included for reference, but are not formally included as part of the Electroweak theory.

isospin t_3 and the weak hypercharge Y as

$$Q = t_3 + Y, \quad (2.13)$$

which is a consequence of Electroweak unification.

2.3.3 The Electroweak Lagrangian

The symmetry group $SU(2)_L$ contains three generators $t_i = \frac{1}{2}\sigma_i$, where σ_i are the Pauli matrices, and the L in this context means “left-handed only.” The gauge bosons corresponding to the generators of the $SU(2)_L$ symmetry are called the W_1^μ , W_2^μ , and W_3^μ , which by construction only couple to the left-handed fermions of the theory. Furthermore, the symmetry group associated with weak hypercharge is $U(1)_Y$, which contains a single generator Y . The corresponding gauge boson, known as the B^μ , couples equally between fermion chiralities.

The procedure of generating the Electroweak interactions follows the same prescription as was carried out for QED in Section 2.2.3. Thus, the requirement of local gauge invariance

under the $SU(2)_L \times U(1)_Y$ group results in the following Lagrangian,

$$\begin{aligned} \mathcal{L}_{\text{EW}} = & \sum_{i=1}^3 \bar{Q}_L^i (iD) Q_L^i + \sum_{i=1}^3 \bar{L}_L^i (iD) L_L^i + \sum_{i=1}^3 \bar{u}_R^i (iD) u_R^i + \sum_{i=1}^3 \bar{d}_R^i (iD) d_R^i \\ & + \sum_{i=1}^3 \bar{e}_R^i (iD) e_R^i - \frac{1}{4} W_{\mu\nu}^i W^{\mu\nu i} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \end{aligned} \quad (2.14)$$

where i is the generation index, and Q_L^i and L_L^i refer to the left-handed quark and lepton iso-doublets, respectively, and u_R^i , d_R^i , and e_R^i to the right-handed up-type and down-type quark and lepton iso-singlets, respectively, according to Table 2.3.

The covariant derivative is

$$D_\mu = \partial_\mu + igt^i W_\mu^i + ig' Y B_\mu, \quad t^i = \frac{1}{2} \sigma^i. \quad (2.15)$$

The two coupling constants: g associated with $SU(2)_L$, and g' associated with $U(1)_Y$, set the strengths of their corresponding interactions. The field strength tensors are

$$\begin{aligned} W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon^{ijk} W_\mu^j W_\nu^k \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned} \quad (2.16)$$

The $SU(2)_L$ group is **non-abelian**, meaning its generators do not commute, which is codified in the relation $[t_i, t_j] = i\epsilon_{ijk} t_k$, where ϵ_{ijk} is the Levi-Civita symbol. As a result, self-interactions between the W_i^μ gauge fields become possible, as seen in the third term of the field strength tensor of Eq. 2.16. The $U(1)_Y$ group is **abelian**, and as a result no self-interactions are present between the gauge fields.

The Electroweak Lagrangian of Eq. 2.14 is incomplete in one crucial respect: the fermions and gauge fields lack mass terms, and are hence massless fields in direct contradiction with

experiment. In fact, it is impossible to add explicit mass terms since the left- and right-handed fields exhibit different transformation properties; i.e., standard mass terms of the form $\mathcal{L}_{\text{mass}} = m\bar{\psi}\psi = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$ break $SU(2)_L$ gauge invariance. The solution to this puzzle is developed in the next section. Still, the reason nature treats left- and right-handed fields so differently is one of the unresolved mysteries of the Standard Model.

2.3.4 The Higgs Mechanism

The solution to the fermion and gauge boson mass paradox is obtained through the spontaneously breaking of Electroweak symmetry via the Higgs Mechanism [57, 58, 59]. The idea is to introduce a new particle into the Electroweak theory $\Phi(x)$, which is an $SU(2)_L$ iso-doublet of complex scalar fields,

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}. \quad (2.17)$$

The upper $t_3 = 1/2$ component ϕ^+ is electrically charged with $Q = +1$ and lower $t_3 = -1/2$ component ϕ^0 is electrically neutral. Each component carries a weak hypercharge of $Y_\phi = 1/2$. The field $\Phi(x)$, known as the Higgs field, can obtain a non-zero value in its vacuum state, which results in the breaking of the $SU(2)_L \times U(1)_Y$ symmetry of the Electroweak Lagrangian of Eq. 2.14 and in the spontaneous generation of gauge-invariant mass terms.

The Higgs Lagrangian is

$$\mathcal{L}_{\text{Higgs}} = |D_\mu \Phi|^2 - V(\Phi), \quad (2.18)$$

where D_μ is the Electroweak covariant derivative of Eq. 2.15 and the Higgs potential is

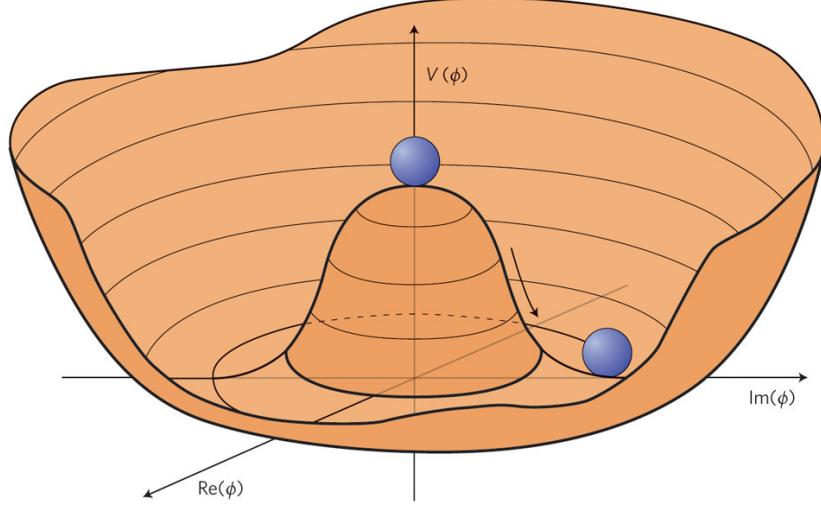


Figure 2.2: The shape of the Higgs potential $V(\Phi)$ of Eq. 2.19 for the choice of $\mu^2 > 0$. The potential $V(\Phi)$ obtains a minimum at the non-zero value of $\Phi^\dagger\Phi = \mu^2/2\lambda \equiv v^2/2$ [6].

defined as

$$V(\Phi) = -\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, \quad (2.19)$$

where μ and λ are real-valued constants.

The Higgs potential of Eq. 2.19 is invariant under $SU(2)_L \times U(1)_Y$ gauge transformations and is constructed in such a way that the vacuum state is *degenerate*. By requiring $\mu^2 > 0$ and $\lambda > 0$, the potential $V(\Phi)$ is minimized when $\Phi^\dagger\Phi = \mu^2/2\lambda \equiv v^2/2$, where v is a non-zero number known as the **vacuum expectation value** (or vev). The situation is depicted in Figure 2.2.

Due to the degeneracy of the $V(\Phi)$ vacuum, there are an infinite number of states that satisfy the minimum requirement. A specific choice amounts to choosing a gauge and the vacuum state of the scalar field becomes $\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$. The Higgs field Φ can now be expressed in terms of an expansion around its vacuum state Φ_0 and as a real valued scalar

field $H(x)$, which has a zero vev, so that Eq. 2.17 becomes

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \quad (2.20)$$

In terms of $H(x)$ the $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken, such that $SU(2)_L \times U(1)_Y \rightarrow U(1)_Q$.¹⁰ The field $H(x)$ describes a neutral scalar, called the Higgs boson. While one degree of freedom of $\Phi(x)$ contributes to $H(x)$, the other three contribute to the longitudinal polarizations of the W^\pm and Z bosons.

As a result, in terms of the redefined Higgs field of Eq. 2.20, the scalar Lagrangian of Eq. 2.18 becomes

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = & -\frac{1}{2} (\partial_\mu H)^\dagger \partial_\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4 \\ & - \frac{g^2}{8} (v + H)^2 \left| W_\mu^1 - i W_\mu^2 \right|^2 - \frac{1}{8} (v + H)^2 (g' B_\mu - g W_3^\mu)^2 \end{aligned} \quad (2.21)$$

which now contains the desired gauge boson mass terms. To identify the mass terms explicitly, it helps to redefine the gauge fields in the *mass basis* in order to eliminate the mixed terms between fields. With that, the physical gauge fields become

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2) \\ Z_\mu &= \frac{1}{\sqrt{g^2 + g'^2}} (g W_\mu^3 - g' B_\mu) \equiv W_\mu^3 \cos \theta_W - B_\mu \sin \theta_W \\ A_\mu &= \frac{1}{\sqrt{g^2 + g'^2}} (g' W_\mu^3 + g B_\mu) \equiv W_\mu^3 \sin \theta_W + B_\mu \cos \theta_W \end{aligned} \quad (2.22)$$

where θ_W is the **Weinberg Angle**, defined by $\tan \theta_W = g'/g$, which parametrizes the

¹⁰A better term might be “hidden,” or no longer manifest, since the Lagrangian is still invariant, but the vacuum is not.

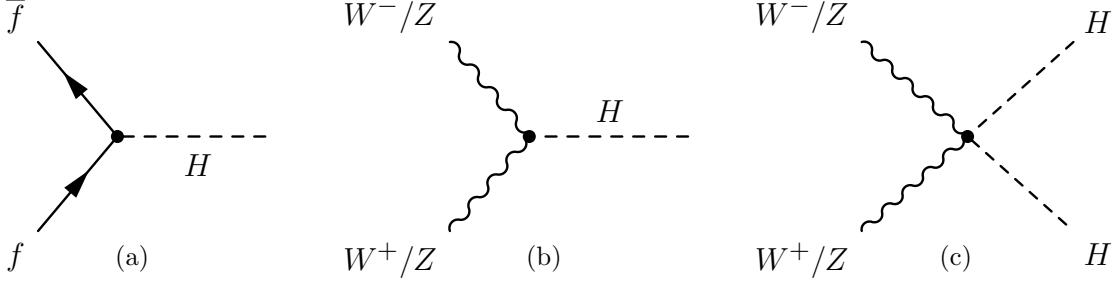


Figure 2.3: Fundamental interaction vertices with the Higgs boson in the EW theory: the Higgs-fermion interaction (a), and the three-point (b) and four-point (c) gauge interactions.

degree of mixing between the gauge fields after spontaneous symmetry breaking.¹¹ The gauge couplings g of $SU(2)_L$ and g' of $U(1)_Y$ are related to the electromagnetic coupling e with

$$e = g \sin \theta_W = g' \cos \theta_W, \quad (2.23)$$

and a coupling for the weak neutral current can be specified by

$$g_Z = \frac{e}{\sin \theta_W \cos \theta_W}. \quad (2.24)$$

With these substitutions, the scalar Lagrangian of Eq. 2.18 takes on a much simpler form,

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = & -\frac{1}{2} (\partial_\mu H)^\dagger \partial_\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4 \\ & - \frac{g^2}{4} (v + H)^2 W_\mu^+ W^{\mu-} - \frac{1}{8} (g^2 + g'^2) (v + H)^2 Z_\mu Z^\mu \end{aligned} \quad (2.25)$$

which highlights the three- and four-point interactions between the Higgs field and the other gauge fields, depicted in Figure 2.3, as well as its own self-interactions, depicted in Figure 2.4.

¹¹Said another way, the Weinberg angle relates the weak eigenstates (W^3, B) to the mass eigenstates (Z, A) .

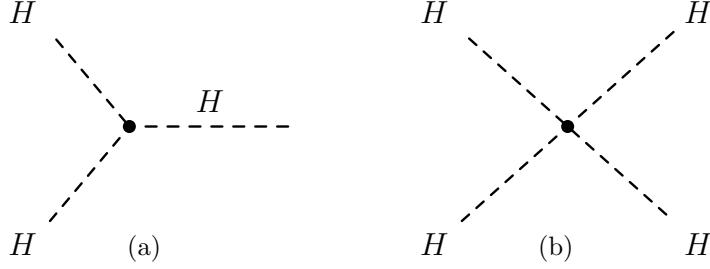


Figure 2.4: Fundamental three-point (a) and four-point (b) Higgs boson self-interaction vertices in the EW theory.

In this form, Eq. 2.25 readily yields the following mass terms,

$$\begin{aligned}
 M_H &= \sqrt{2\lambda}v \\
 M_W &= \frac{1}{2}gv \\
 M_Z &= \frac{1}{2}g_Zv,
 \end{aligned} \tag{2.26}$$

which are all proportional to the vev v . The photon remains massless. From the experimental values of M_W and M_Z , the vev is deduced to be $v \simeq 246$ GeV [1]. It has units of mass, and is the only parameter of the Standard Model that is not a dimensionless number, setting an absolute scale for the Electroweak theory.

2.3.5 Yukawa Interactions

With the addition of the scalar field $\Phi(x)$ into the model, the fermion masses can similarly be generated through spontaneous symmetry breaking according to Yukawa type interactions.¹²

¹²i.e. an interaction between a scalar and fermion fields, reminiscent of Yukawa's original theory that described the nuclear force between nucleons as mediated by Pions.

The general gauge invariant Yukawa Lagrangian is [60]

$$\mathcal{L}_{\text{Yukawa}} = - \sum_{j,k=1}^3 \bar{Q}_{jL} \Phi \lambda_{jk}^{(d)} d_{kR} - \sum_{j,k=1}^3 \bar{Q}_{jL} (-i\sigma_2 \Phi^*) \lambda_{jk}^{(u)} u_{kR} - \sum_{j,k=1}^3 \bar{L}_{jL} \Phi \lambda_{jk}^{(\ell)} e_{kR} + \text{h.c.}, \quad (2.27)$$

where the generational indices j and k allow inter-generation mixing between the fields, and the $\lambda_{jk}^{(f)}$ terms are called **Yukawa Matrices**, each of which depends on 9 complex parameters. After the scalar field $\Phi(x)$ acquires a vev as in Eq. 2.20, the Yukawa Lagrangian becomes

$$\mathcal{L}_{\text{Yukawa}} = -\frac{1}{\sqrt{2}} (v + H) \sum_{j,k=1}^3 \left(\lambda_{jk}^{(d)} \bar{d}_{jL} d_{kR} + \lambda_{jk}^{(u)} \bar{u}_{jL} u_{kR} + \lambda_{jk}^{(\ell)} \bar{e}_{jL} e_{kR} \right) + \text{h.c.} \quad (2.28)$$

The fermion mass terms can be made explicit by diagonalizing the Yukawa Matrices. This can be accomplished with a suitable set of field transformations such that, for the up- and down-type quarks,

$$U_{uL}^\dagger \lambda^{(u)} U_{uR} = \begin{pmatrix} \lambda_u & 0 & 0 \\ 0 & \lambda_c & 0 \\ 0 & 0 & \lambda_t \end{pmatrix}, \quad U_{dL}^\dagger \lambda^{(d)} U_{dR} = \begin{pmatrix} \lambda_d & 0 & 0 \\ 0 & \lambda_s & 0 \\ 0 & 0 & \lambda_b \end{pmatrix}, \quad (2.29)$$

with an identical definition for the leptons of Eq. 2.28. The unitary matrices $U_{fL,R}$ are used to transform the fermion fields from the electroweak basis to the mass basis. As a result, distinct mass eigenstates are formed from linear superpositions of electroweak eigenstates. The price of this procedure is that, like before, it mixes the fundamental fields together, in this case of different generations.

Thus, the Yukawa Lagrangian of Eq. 2.28 can be re-expressed as

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= -\frac{1}{\sqrt{2}}(v + H) \sum_{i=1}^3 \left(\lambda_i^{(d)} \bar{d}_i L d_{iR} + \lambda_i^{(u)} \bar{u}_i L u_{iR} + \lambda_i^{(\ell)} \bar{e}_i L e_{iR} \right) + \text{h.c.} \\ &= -\frac{1}{\sqrt{2}}(v + H) \sum_{i=1}^3 \left(\lambda_i^{(d)} \bar{d}_i d_i + \lambda_i^{(u)} \bar{u}_i u_i + \lambda_i^{(\ell)} \bar{e}_i e_i \right),\end{aligned}\tag{2.30}$$

and the fermion mass terms can now be read off explicitly,

$$m_f = \frac{1}{\sqrt{2}} \lambda_f v,\tag{2.31}$$

which are also proportional to the vev v , just like the mass terms of the Standard Model gauge fields. Effectively, Eq. 2.27 split into two parts: the first of which provided the required fermion mass terms, and the second of which specified the Yukawa interactions between fermions and the Higgs field (as in Fig. 2.3).

The electroweak theory provides no explanation for the empirically determined values of λ_f , and hence provides no fundamental understanding of the fermion masses [60]. Furthermore, as the neutrinos have no right-handed partners ν_{iR} in the Standard Model, they cannot acquire mass terms through the Yukawa interaction described here, nor subsequently can they interact with the Higgs field. Again, the neutrinos are strictly massless in the Standard Model.

The electroweak interactions with fermions defined in Eq. 2.14 were originally specified between identical fermion generations. However, the fermion field redefinitions of Eq. 2.29 mix the generations together. As fermions propagate through the vacuum, they are perceived as maintaining a distinct identity, known as a mass eigenstate, which is a superposition of flavor eigenstates. Fermions oscillate between their flavor eigenstates as they travel, a

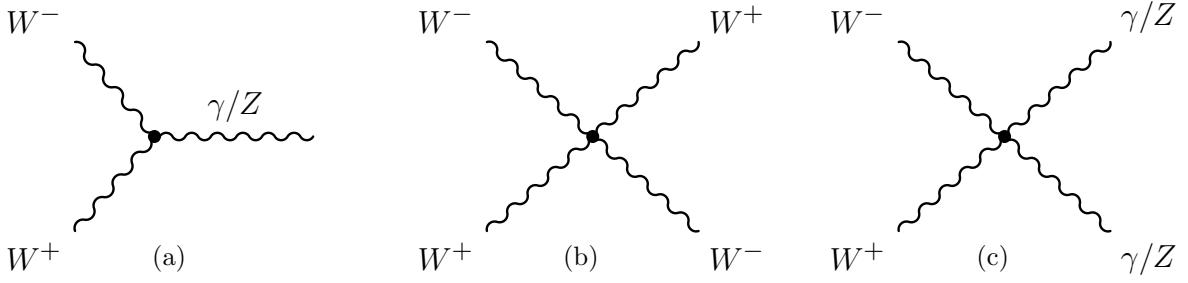


Figure 2.5: Fundamental three-point (a) and four-point (b)-(c) gauge boson vertices in the EW theory.

phenomenon called fermion flavor oscillation.

2.3.6 The Electroweak Lagrangian Revisited

After spontaneous symmetry breaking, the Electroweak Lagrangian of Eq. 2.14 can be re-expressed in terms of the physical fermion and gauge fields.

The substitution of the physical gauge fields of Eq. 2.22 into Eq. 2.14 results in distinct three- and four-point gauge interactions as seen in Figure 2.5. These include WWZ and $WW\gamma$ interactions, as well as $WWWW$, $WW\gamma\gamma$, $WWZZ$, and $WWZ\gamma$ interactions. Furthermore, the covariant derivative of Eq. 2.15 becomes

$$D_\mu = \partial_\mu + ieQA_\mu + i\frac{g}{\sqrt{2}}(t^+W_\mu^+ + t^-W_\mu^-) + ig_Z(t_3 - Q\sin^2\theta_W), \quad (2.32)$$

where the weak isospin raising and lowering operators are defined as $t^\pm = \frac{1}{\sqrt{2}}(t_1 \pm it_2)$.

Upon substitution of Eq. 2.32 into the Electroweak Lagrangian of Eq. 2.14, the standard form of charged and neutral current electroweak interactions with fermions can be written down.

The charged current interaction between the W boson and the quarks (as in Fig. 2.6) is

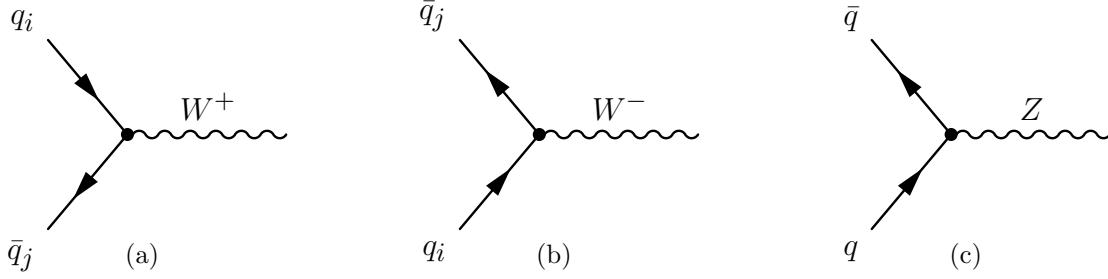


Figure 2.6: Fundamental interaction vertices with quarks in the EW theory: the charged current interactions (a)-(b) and neutral current interaction (c).

given by

$$\mathcal{L}_{Wq} = -i \frac{g}{\sqrt{2}} \sum_{n,m=1}^3 \left(V_{mn} \bar{u}_{Lm} \gamma^\mu d_{Ln} W_\mu^+ + (V_{mn})^\dagger \bar{d}_{Lm} \gamma^\mu u_{Ln} W_\mu^- \right). \quad (2.33)$$

Because all quarks are massive, it is natural to express their fields in the mass basis. Here, the matrix $V = U_u^\dagger U_d$ results from the fermion field redefinitions of Eq. 2.29, and is known as the Cabibbo-Kobayashi-Maskawa (CKM) Matrix [74, 75], defined as

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (2.34)$$

$$= \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4).$$

The CKM matrix parametrizes the degree of mixing between the quark flavor eigenstates of the weak interaction and their corresponding mass eigenstates due to spontaneous symmetry breaking.

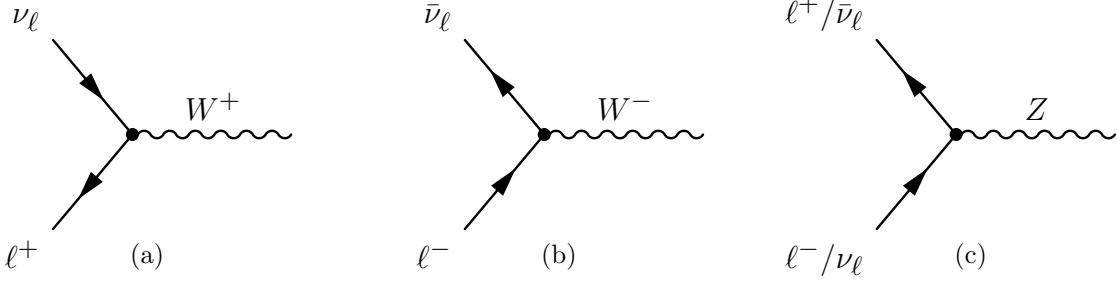


Figure 2.7: Fundamental interaction vertices with leptons in the EW theory: the charged current interactions (a)-(b) and neutral current interaction (c).

The matrix is expressed in terms of the Wolfenstein parametrization [76], which is valid to fourth order in λ with $\lambda \simeq 0.225$, $A \simeq 0.8$, $\rho \simeq 0.13$, and $\eta \simeq 0.35$ [1]. This form nicely highlights the magnitude of each CKM matrix element, and the nearly diagonal structure of the matrix, with off-diagonal terms becoming proportionally smaller. Since V_{CKM} is not strictly diagonal, the W^\pm bosons couple to quark mass eigenstates of different generations.

Likewise, the charged current interaction with leptons (as in Fig. 2.7) is given by

$$\mathcal{L}_{W\ell} = -i \frac{g}{\sqrt{2}} \sum_{i=1}^3 (\bar{\nu}_{Li} \gamma^\mu \ell_{Li} W_\mu^+ + \bar{\ell}_{Li} \gamma^\mu \nu_{Li} W_\mu^-). \quad (2.35)$$

The charged current interaction with leptons does not require an analogous mixing matrix of the form $V = U_\nu^\dagger U_\ell$, since the neutrinos are massless in the Standard Model by definition.¹³

Lastly, the interaction between the neutral electroweak gauge bosons and fermions is given by

$$\begin{aligned} \mathcal{L}_{Z/\gamma\psi} &= -ieQ\bar{\psi}\gamma^\mu\psi A_\mu - ig_Z\bar{\psi}\gamma^\mu \left(t_3 - Q \sin^2 \theta_W \right) \psi Z_\mu \\ &= -ieQ\bar{\psi}\gamma^\mu\psi A_\mu - ig_Z\bar{\psi}\gamma^\mu (g_L P_L + g_R P_R) \psi Z_\mu, \end{aligned} \quad (2.36)$$

¹³In the minimal extension of the Standard Model with three massive neutrinos, $V = V_{PMNS}$ is known as the PMNS matrix, analogous to the CKM matrix of the quark sector.

	g_L	g_R	g_V	g_A
u	$+\frac{1}{2} - \frac{2}{3}x_W$	$-\frac{2}{3}x_W$	$+\frac{1}{4} - \frac{2}{3}x_W$	$-\frac{1}{4}$
d	$-\frac{1}{2} + \frac{1}{3}x_W$	$+\frac{1}{3}x_W$	$-\frac{1}{4} + \frac{1}{3}x_W$	$+\frac{1}{4}$
ν_ℓ	$+\frac{1}{2}$	0	$+\frac{1}{4}$	$-\frac{1}{4}$
ℓ	$-\frac{1}{2} + x_W$	$+x_W$	$-\frac{1}{4} + x_W$	$+\frac{1}{4}$

Table 2.4: Neutral current couplings of Standard Model fermions. The rows specify the couplings within each respective quark and lepton generation. Alternate definitions often are written in terms of $g_V = \frac{1}{2}(g_L + g_R)$ and $g_A = \frac{1}{2}(g_L - g_R)$, which function as coefficients to the vector (γ^μ) and axial-vector ($\gamma^\mu\gamma^5$) terms of the neutral current Lagrangian.

where ψ stands for any quark or lepton of the three fermion generations. The first term in the Lagrangian of Eq. 2.36 is identified with the standard electromagnetic interaction induced by the photon γ and the second term with the neutral current interaction of the Z boson.

In contrast to the charged current interactions, the neutral current interactions have no flavor changing transitions. The left- and right-handed couplings are defined as $g_L = t_3 - Q \sin^2 \theta_W$ and $g_R = -Q \sin^2 \theta_W$, respectively. Therefore, the neutral current interaction is chiral, coupling with different strengths to the left- and right-handed fields. The values of the couplings are listed in Table 2.4.

2.4 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of the strong interaction. It is an unbroken non-abelian gauge theory of gauge group $SU(3)_C$. The quark fields carry one of three possible color charges $N_C = 3$ and are grouped into $SU(3)_C$ color-triplets ψ_j where j refers to “red,” “green,” or “blue.” The leptons are color-singlets and do not participate in the strong interaction. The gluons are the force mediators of the strong interaction and also carry color charge themselves. There are eight color combinations of gluons.

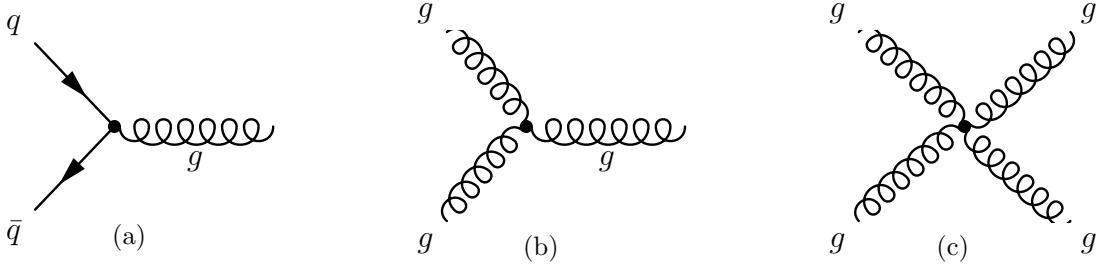


Figure 2.8: Fundamental interaction vertices in QCD: the quark-gluon interaction (a), the three-point gluon interaction (b), and the four-point gluon interaction (c).

The QCD Lagrangian is

$$\mathcal{L}_{\text{QCD}} = \sum_{q=1}^6 \sum_{i,j=1}^3 \bar{\psi}_i^q (iD_\mu - m_q \delta_{ij}) \psi_j^q - \sum_{a=1}^8 \frac{1}{4} G_{\mu\nu}^a G_{\mu\nu}^a, \quad (2.37)$$

where the summations occur over each quark of flavor q with mass m_q and also over the color indices i and j in reference to the three possible color charges. The gauge covariant derivative is

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} + ig_S t_{ij}^a G_\mu^a, \quad t_{ij}^a = \frac{\lambda_{ij}^a}{2}, \quad (2.38)$$

which introduces the eight Gell-Mann matrices λ_{ij}^a , the $SU(3)$ analogs of $SU(2)$ Pauli matrices, as well as the gauge fields for the gluons G_μ^a with a running from 1 to $N_C^2 - 1 = 8$; i.e., there are eight kinds of gluons. The coupling g_S sets the strength of the strong interaction and can be expressed as a “fine-structure constant” for QCD by $\alpha_S = g_S^2/4\pi$, often called the strong coupling constant.

The gluon field strength tensor is

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_S f^{abc} G_\mu^b G_\nu^c, \quad [t_a, t_b] = if_{abc} t_c, \quad (2.39)$$

which, as was also the case in the $SU(2)_L$ Electroweak theory, includes an additional term due to the non-abelian structure of the gauge group. Therefore, the kinetic term for the gluon fields in Eq. 2.37 predicts 3- and 4-point self-interactions, which are depicted in Figure 2.8. The f_{abc} terms are the structure constants of the $SU(3)_C$ gauge group.

2.4.1 Asymptotic Freedom and Confinement

The strong force is appropriately named due to the large value of its coupling constant α_S . By comparison, the electromagnetic coupling α_{EM} is more than an order of magnitude smaller. The values assumed by the coupling constants depend on the energy or momentum-transfer scale μ associated with the given physical process.

This “running” of the coupling can be parametrized by the beta function [60],

$$\beta(\alpha) = \mu^2 \frac{\partial \alpha(\mu^2)}{\partial \mu^2} = \sum_{i=0}^{\infty} \beta_i \alpha(\mu^2)^{i+2} \quad (2.40)$$

whose solution gives the scale dependence of the coupling. When keeping only the leading term in the expansion, the solution is

$$\alpha(\mu^2) = \frac{\alpha(\mu_0^2)}{1 - \alpha(\mu_0^2) \beta_0 \ln(\mu^2/\mu_0^2)}, \quad (2.41)$$

where μ_0 is a reference scale where α is known, and the factor β_0 is the leading-order coefficient in the perturbative expansion of Eq. 2.40, and can be calculated explicitly.

In QED, the coefficient $\beta_0 = 1/(3\pi)$, which results in a running, such that $\alpha_{EM}(\mu) \rightarrow \infty$ as $\mu \rightarrow \infty$; i.e., as the energy scale increases, so does the strength of the electromagnetic interaction. Here, the running is fairly slow, increasing from its low energy value of $\alpha_{EM}(0) \approx$

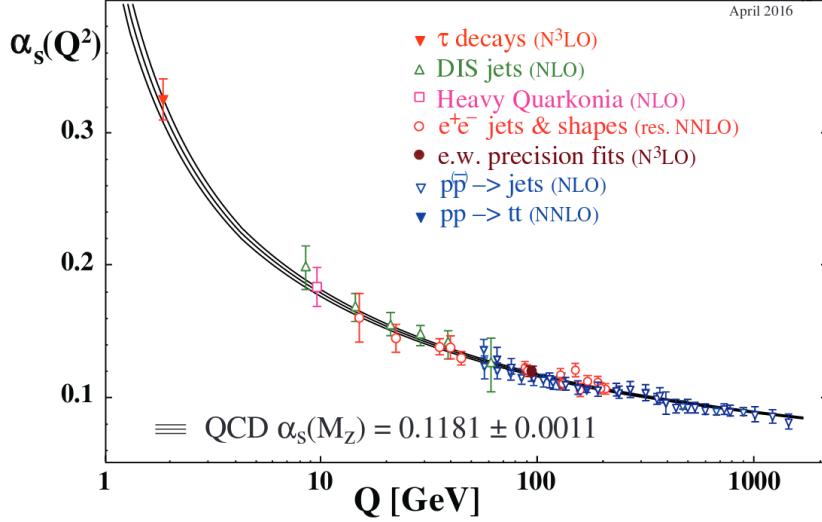


Figure 2.9: The running of the α_S coupling constant in QCD as a function of the energy scale Q with a variety of experimental measurements overlaid [1].

$1/137$ to a value of $\alpha_{\text{EM}}(M_Z) \approx 1/128$ at the scale of the Z boson mass.

However, the behavior of the QCD coupling constant differs significantly from its electromagnetic counterpart due to non-abelian nature of the theory. The corresponding QCD beta function is *negative* with $\beta_0 = -(33 - 2N_F)/12\pi$, where N_F is the number of active quark flavors. Consequently, this leads to a running where $\alpha_S(\mu) \rightarrow 0$ as $\mu \rightarrow \infty$, a phenomenon known as **Asymptotic Freedom** [77, 78]. As a result, the quarks do not feel the strong force at high energies and behave essentially as free particles. On the other hand, at low energies the interactions become strong, leading to the confinement of quarks and gluons into bound colorless hadrons.

As long as the energy scale is sufficiently large, the smallness of α_S makes perturbative calculations of QCD observables, such as hadronic scattering cross sections, possible. However, as α_S approaches unity, the perturbative approach breaks down. The scale at which this happens is known as $\Lambda_{QCD} \approx 220$ MeV. At the mass of the Z boson, the value of $\alpha_S(M_Z) = 0.118$ [1]. Figure 2.9 plots the running of α_S as a function of the energy scale Q .

Chapter 3

Proton-Proton Collisions

This chapter provides a foundation for understanding the phenomenology of the proton-proton collisions at the LHC. This is important because the signature of new physics, if it exists, will be identified by performing a detailed analysis on a large dataset of such collisions. In this chapter, the composite nature of the proton is described in detail, along with the dynamics of proton-proton collisions themselves.

3.1 Proton Structure

The proton is a **composite particle**. Unlike the **point-like** particles of the Standard Model (as in Table 2.1), the proton has an internal structure. By scattering point-like probes off of it, the internal structure of the proton can be investigated. The so-called resolution of the probe can be expressed in terms of the energy scale Q^2 , the squared momentum transfer between the probe and proton during a scattering process.

The proton consists of an ensemble of interacting quarks and gluons, collectively known as **partons**, whose dynamics are described by QCD, the $SU(3)$ gauge theory of the strong interaction (See Sec. 2.4). Due to the non-abelian nature of the theory, the interaction strength between the quarks and gluons increases (decreases) with distance (energy). As a result, individual quarks and gluons are only ever found as bound hadronic states, such as the proton, a process known as confinement. At low values of Q^2 , the proton appears point-

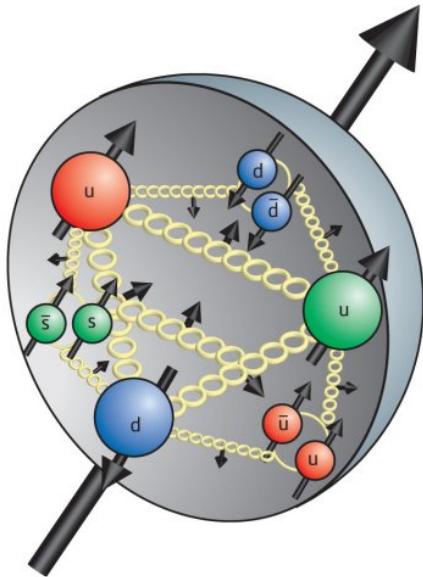


Figure 3.1: Substructure of proton depicting interactions between its partonic constituents: valence quarks, sea quarks, and gluon [7].

like, and its internal structure cannot be resolved. However, as smaller and smaller distances are probed (larger energy scale Q^2), the strength of the QCD interaction decreases, which allows the partons to be treated essentially as a free particles. This is known as asymptotic freedom.

The static properties of the proton (spin, electric charge, isospin, etc.) are fully determined by its three **valence quarks**: two up quarks and one down quark (uud). Along with the three valence quarks, the proton (and all hadrons for that matter) consists of an indefinite number of quark-antiquark pairs known as **sea quarks** (See Fig. 3.1). Sea quarks arise when a gluon splits, or fluctuates, into a $q\bar{q}$ pair (as in Fig. 2.8). The constant flux of gluons splitting and recombining results in what is known as “the quark sea,” the set of all off-shell quarks in the proton consistent with the uncertainty principle.

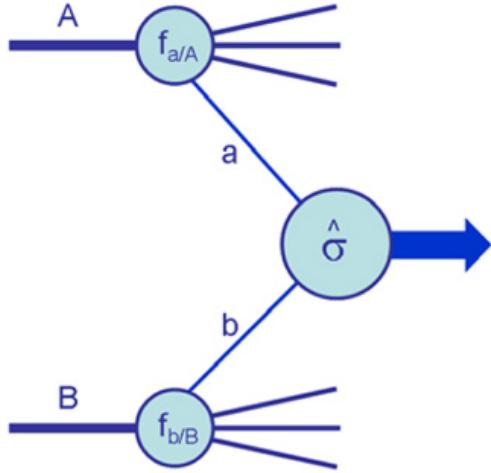


Figure 3.2: Diagram of a typical hard scattering interaction between partons at the LHC [8].

3.1.1 Proton-Proton Collisions

The majority of proton-proton collisions at the LHC are elastic, meaning the colliding protons remain intact leaving their substructure unresolved. On the other hand, in an inelastic collision, or **hard scatter**, the protons break apart. In this case, the momentum transfer squared Q^2 is large, and the proton substructure can be resolved. Here, the interaction is not between the protons themselves, but between two of their constituent partons, one from each proton. Since the partons go on to initiate in the interaction essentially act as free particles, perturbative QCD (pQCD) can be applied to calculate the **partonic cross section** $\hat{\sigma}$ of the two body interaction.

A sketch of a generic hard scatter at the LHC can be seen in Figure 3.2. Here, the incoming protons A and B each carry a momentum $P_A = P_B = \frac{1}{2}\sqrt{s}$, half of the center-of-mass (CM) energy \sqrt{s} of the LHC. The proton momentum is distributed among its many partonic constituents, and as such, the partons a and b that go on to initiate the hard scatter each carry a momentum p_a and p_b , some fraction less than the total proton momentum. The

momentum fraction x of each colliding parton is defined by

$$\begin{aligned} x_a &= \frac{p_a}{P_A}, \quad 0 < x_a < 1 \\ x_b &= \frac{p_b}{P_B}, \quad 0 < x_b < 1 \end{aligned} \tag{3.1}$$

and the parton four-vectors can be written,

$$p_a^\mu = \frac{\sqrt{s}}{2}(x_a, 0, 0, x_a), \quad p_b^\mu = \frac{\sqrt{s}}{2}(x_b, 0, 0, -x_b). \tag{3.2}$$

Since each parton carries a fraction of the total proton momentum, the energy that goes into the collision is only a fraction of the total CM energy of the LHC. The energy scale Q^2 of the collision can be expressed as the squared sum of the parton four-vectors,

$$Q^2 = (p_a^\mu + p_b^\mu)^2 = x_a x_b s, \tag{3.3}$$

which is proportional to the momentum fraction x of each parton and the CM energy of the collider. The two partons that initiate the process scatter with a *process dependent* cross section $\hat{\sigma}(Q^2, x_a, x_b)$, which describes the probability for the event to occur, and can be calculated perturbatively as a series in α_S in the framework of pQCD (to be introduced in Sec. 4.1).

The probability to find a given parton with momentum fraction x at resolution scale Q^2 is given by a **parton distribution function** (PDF) where, for example, the probability to find parton a in proton A with x_a at scale Q^2 is denoted $f_{a/A}(x_a, Q^2)$. These functions are non-perturbative and parton dependent, and describe how the proton momentum is distributed among its constituent partons. A detailed discussion starts in Sec. 3.2.

Event kinematics depend heavily on how much momenta the initial state partons carry. For example, in the production of a resonance particle, such as a W or Z boson, the degree to which a system is boosted longitudinally depends on the difference between the momentum fractions x_a and x_b of the incoming partons. Assuming the partons are massless and that the collision occurs head on, the **rapidity**, a measure of the longitudinal boost can be defined,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) = \frac{1}{2} \ln \frac{x_a}{x_b}. \quad (3.4)$$

The expression simplifies when the energy E and longitudinal momentum p_z are taken to be the summed components of the partonic four-vectors in Eq. 3.3. In general, since $x_a \neq x_b$, Eq. 3.4 predicts a mediating W or Z boson produced in the hard collision will move along the z -axis relative to the CM frame of the proton-proton system.

The inversion of Eq. 3.4, combined with Eq. 3.3, results in an expression for the partonic x fractions in terms of measurable kinematic variables,

$$x_a = \frac{Q}{\sqrt{s}} e^{+y}, \quad x_b = \frac{Q}{\sqrt{s}} e^{-y}. \quad (3.5)$$

Thus, given the fixed CM energy \sqrt{s} of the LHC, different values of Q^2 and y can be used to probe different values of x .

3.1.2 QCD Factorization Theorem

The **QCD Factorization Theorem** greatly simplifies the calculation of hadronic cross sections. It states that the hadronic cross section σ_{AB} can be reduced, or factorized, into two separate components. The first is the short-distance part, defined by the partonic cross

section $\hat{\sigma}$, which is calculable in pQCD, and the second is the non-perturbative long-distance part, codified by the PDFs $f(x, Q^2)$. This factorization is possible for two reasons: first, the partonic cross section is *independent* of the type of colliding hadrons; and second, the PDFs are **universal**, meaning they are a property of the colliding hadrons themselves and are therefore independent of the dynamics of the scattering process.

According to the factorization theorem, the total hadronic cross section can be written [8],

$$\sigma_{AB} = \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times [\hat{\sigma}_{LO} + \alpha_S(\mu_R^2) \hat{\sigma}_{NLO} + \dots]_{ab \rightarrow X} \quad (3.6)$$

where the indices a and b run over all relevant parton flavors, and the partonic cross section $\hat{\sigma}$ is expressed as a perturbative series in terms of α_S .

The parameter μ_F^2 is known as the **factorization scale**, while μ_R^2 , on the other hand, is known as the **renormalization scale**. Each scale is an artifact of the perturbative approach and reflects the uncertainty in the prediction due to unknown higher-order terms. For the Drell-Yan process, each scale is generally characterized by the scale of the collision Q^2 , so that $\mu_R^2 = \mu_F^2 = Q^2$. Different choices for μ_R^2 and μ_F^2 lead to in general different numerical results for the cross section estimate, so up and down variations of the nominal choice are often used as a way to assess the theoretical uncertainty.

3.1.3 Hadronic Events

A complete picture of the collision does not end with the calculation of the hadronic cross section of Eq. 3.6. A diagram of a typical hard scatter, and the subsequent evolution of the event, is depicted in Figure 3.3. In such events, hundreds of particles are produced, with energies and momenta spanning several orders of magnitude, which results in numerous

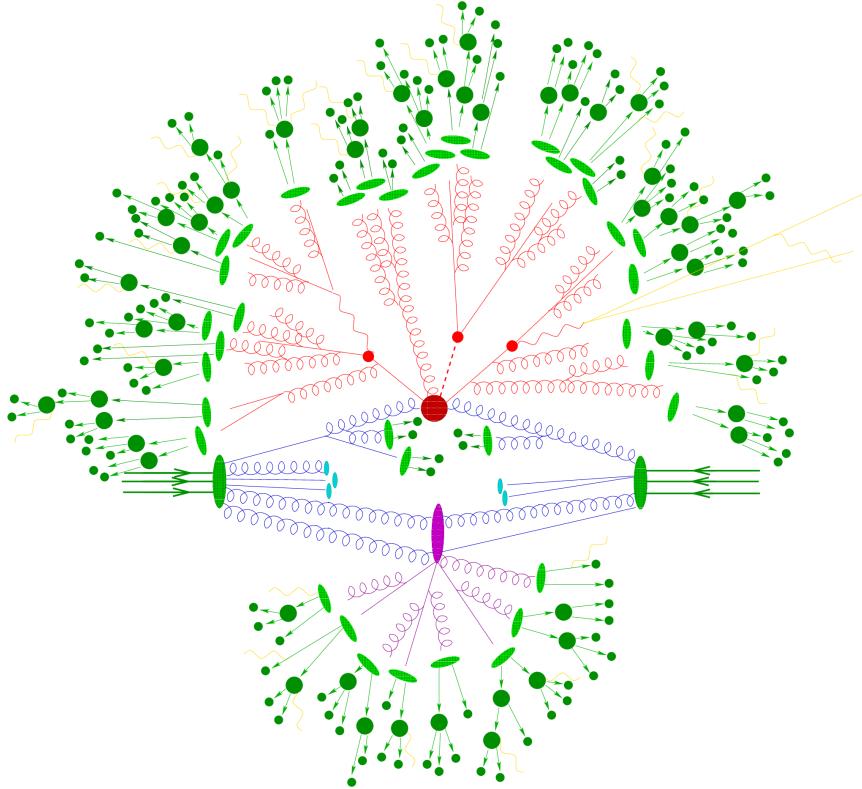


Figure 3.3: An inelastic collision at the LHC. The dark red blob depicts the hard process, with its associated red lines the resulting final state radiation. Beam remnants are light blue, and initial state radiation is depicted by the dark blue lines. The purple blob is a secondary hard scatter along with its own final state radiation in purple. The light green blobs depict hadronization of any color-charged partons emanating from the hard scatter or the ISR/FSR showering. Subsequent dark green blobs corresponds to hadronic decays of unstable particles and yellow lines indicate the presence of electromagnetic radiation [9].

complex topologies and distinct final states.

In such a collision, each proton breaks apart, leaving behind a remnant of the partons that did not participate in the hard scatter. The debris that results is left in a color-charged state from which a particle cascade, or **parton shower**, develops. Subsequent radiation and quark and gluon splittings drive the shower evolution. Eventually, the shower stops around scale of Λ_{QCD} in a process known as **hadronization**, where the partons become confined within hadrons. However most hadrons are unstable and subsequently decay, sometimes multiple times, resulting in additional hadrons and radiation. The complex interplay between

the remaining **beam remnants**, and the additional partons produced in the event from splittings, the shower, and radiation, is referred to as the **underlying event**, which serves as a background to the original hard scatter.

Despite the complexity, the phenomenology of the collision can be factorized into several distinct stages:

Hard Scattering Process - the collision between partons a and b in proton A and B . The partonic cross section can be calculated at some fixed-order in perturbation theory.

Multi-Parton Interactions (MPIs) - because protons are composite objects, more than one hard scatter can occur from the same proton-proton collision.

Underlying Event (UE) - an unavoidable background to the hard scatter. It consists of all event particles that accompany the hard scatter; i.e., beam remnants, MPIs, and those from branchings, as well other in-time proton collisions.

Initial State Radiation (ISR) - partons can radiate after the proton breaks apart but before the hard collision itself. Most ISR is soft; i.e., confined to small transverse momenta.

Final State Radiation (FSR) - final state partons can undergo radiation as well. If the final state is color-neutral, like a pair of leptons, only electromagnetic radiation is possible.

Beam Remnants - the extraction of the partons from the proton to initiate the hard scatter leaves behind a remnant of the proton that is no longer color-neutral.

Hadronization - the non-perturbative process of transforming color-charged partons originating from the hard scatter, ISR and FSR branchings, and elsewhere in the event

into color-neutral hadrons.

Particle Decays - hadrons produced are generally unstable and subsequently decay, sometimes multiple times through a decay-chain, until a stable state is reached. If they are charged, additional electromagnetic radiation can be produced.

Understanding the evolution of the event in stages like this allows for robust **Monte Carlo** (MC) simulation samples to be developed to model the proton-proton collisions at the LHC, the subject of Section 9.1.

3.2 Parton Distribution Functions

The PDFs $f_i(x, Q^2)$ determine how the momentum of the proton is shared among its partonic constituents, where $f_i(x, Q^2)$ is the density of parton species i per unit of momentum fraction. This is a complicated situation because each parton flavor (i.e. valence quarks, sea quarks, and gluons) shares a different fraction x of the total proton momentum. Not only that, but the momentum fraction each parton carries is strongly dependent on the energy scale Q^2 of the hard scatter.

Figure 3.4 illustrates the situation for the valence quarks $u_v(x)$ and $d_v(x)$, the sea quarks $\bar{u}(x)$, $\bar{d}(x)$, $s(x)$, $\bar{s}(x)$, $c(x)$, and $\bar{c}(x)$, and the gluon $g(x)$, at the scales of $Q^2 = (2 \text{ GeV})^2$ and $Q^2 = (100 \text{ GeV})^2$.¹ The evolution of the PDFs from one scale Q^2 to another can be described analytically in the framework of pQCD with the **DGLAP equations** [79].²

Furthermore, it can be seen that the valence quarks dominate at low Q^2 and for $x \gtrsim 0.1$,

¹The total u quark content of the proton is the sum of its valence and sea contributions (i.e. $u(x) = u_v(x) + \bar{u}(x)$). The same holds for the d quark content (i.e. $d(x) = d_v(x) + \bar{d}(x)$). For the heavy quarks, like the strange, $s(x) = \bar{s}(x)$, since there is no valence component.

²Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

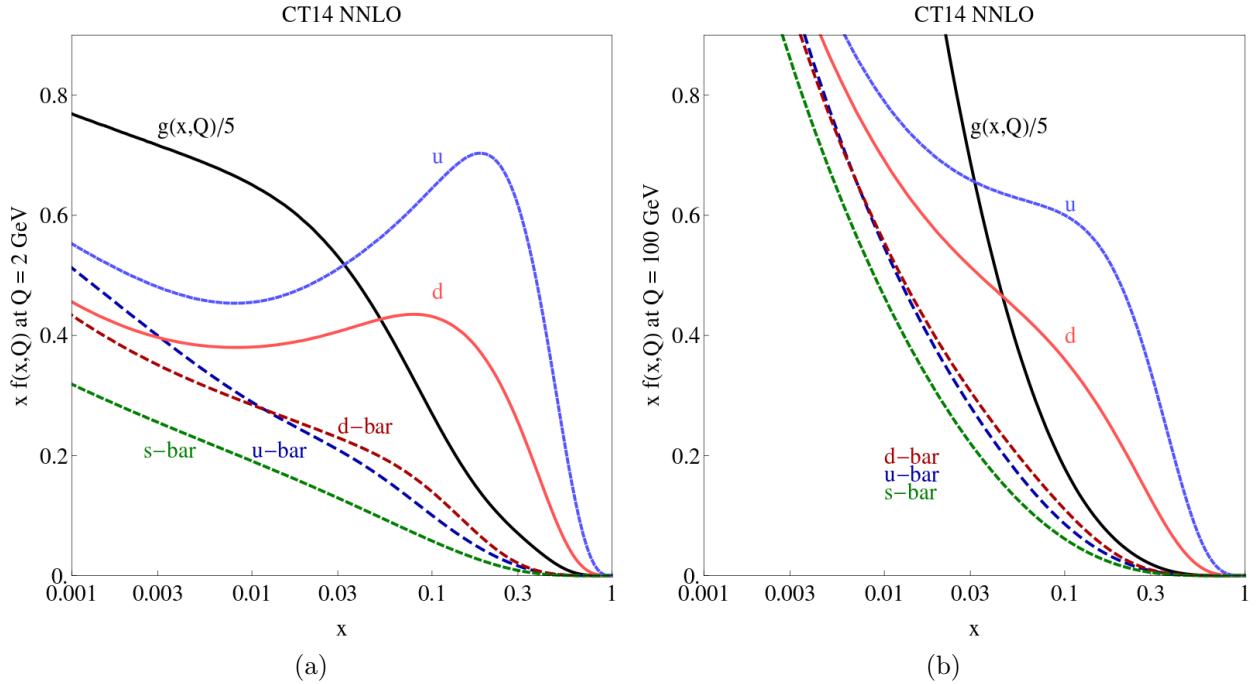


Figure 3.4: CT14NNLO PDFs provided by the CTEQ collaboration. Depicted are gluon and all quark flavors as a function of x , evaluated at the scales of $Q = 2$ GeV (a) and $Q = 100$ GeV (b) [10].

with the u quark more than twice as large as the d quark, and both much larger than the heavy quarks of the sea in this region. At $x \simeq 1/3$, the $u(x)$ and $d(x)$ distributions peak, which corresponds to the valence part of the proton. At low x , the distributions show increasing contributions from the sea and gluon. The flavor symmetry of the quark sea is actively being explored at the LHC, with the goal of understanding how proton momentum is shared by the $s(x)$ and $\bar{s}(x)$ distributions as compared to the $\bar{u}(x)$ and $\bar{d}(x)$ distributions.

As the energy scale Q^2 is increased, the contributions from the sea and the gluon grow in importance, as can be seen in the rapidly changing shapes of their distributions at very low x . Further, the peaks associated with the valence $u(x)$ and $d(x)$ diminish in significance, as now more of the momentum of the proton is carried by the low- x sea. In particular, the gluon distribution dominates in the region $x \lesssim 0.05$. Also, the sea becomes more flavor

symmetric, as additional sea quarks arise from $q\bar{q}$ gluon splittings.

Lastly, an important requirement for the PDFs is that they be properly normalized. This is ensured by two **sum rules**. The first is momentum sum rule where,

$$\sum_i \int_0^1 x f_i(x, Q^2) dx = 1, \quad (3.7)$$

which requires the total momentum fraction of each parton within the proton adds up to unity. The second is the quark-number sum rule where,

$$\int_0^1 f_{uv}(x, Q^2) dx = 2, \quad \int_0^1 f_{dv}(x, Q^2) dx = 1, \quad (3.8)$$

which ensures the number of up and down valence quarks agrees with the *uud* definition of the proton, independent of the scale Q^2 of the scattering.

3.2.1 Extracting PDFs from Experimental Data

The PDFs are non-perturbative by nature and cannot, therefore, be calculated analytically in the framework of pQCD. Nevertheless, the determination of the PDFs is crucial, as they are required for the estimation of LHC cross sections according to Eq. 3.6.

The situation can be overcome by performing a **global analysis**, in which the PDF for each parton flavor is parametrized by an arbitrary functional form, dependent on x and Q^2 , and fit to an ensemble of cross section data points, spanning a large range in x and Q^2 values from many experiments. The factorization theorem allows cross sections to be predicted for each experimental dataset used. The predicted cross sections are then fit simultaneously to the available experimental data, which allows the fit parameters to be extracted, and hence

the PDFs. The entire procedure owes its success to the universal nature of the form of the PDFs, allowing for a wide variety of experimental inputs to be used.

Extracting PDFs is a major theoretical industry with multiple collaborations participating. The most prominent collaborations are CTEQ [10], MSTW [80], and NNPDF [81], with each providing semi-regular updates to their PDF parameterizations, as new experimental data and theoretical techniques become available. However, there is no unique set of PDFs commonly accepted by the HEP community, as each group strives to provide the best parametrization possible. Differences can arise from, for example, the choice of the PDF fit parameterizations, the use and exclusion of certain input datasets, various theoretical assumptions, and the way in which experimental and theoretical errors are handled.

The determination of most global fits relies on data from lepton-nucleon fixed target experiments, Deep Inelastic Scattering (DIS) experiments at HERA, and collider-based Tevatron and LHC experiments [8, 82].³ These sources cover a wide range in the (x, Q^2) kinematic phase space, as seen in Figure 3.5, with somewhat considerable overlap as well.

In general, the PDF global analyses start by choosing suitable parameterizations for the parton distributions. These are generic functional forms that specify the valence quark distributions $u_v(x)$ and $d_v(x)$, the gluon $g(x)$, and the $q\bar{q}$ sea distributions, and can be represented by [8]

$$f(x, Q_0^2) = a_0 x^{a_1} (1 - x)^{a_2} P(x; a_3, \dots). \quad (3.9)$$

The parameters a_1 and a_2 are physically motivated, and the reference value Q_0^2 is chosen in the 1-2 GeV range. The first two terms alone cannot sufficiently describe the quark and gluon

³The lepton-nucleon data arises from experiments at SLAC, FNAL, and CERN, which includes e , μ , and ν DIS measurements on Hydrogen, Deuterium, and nuclear targets [83].

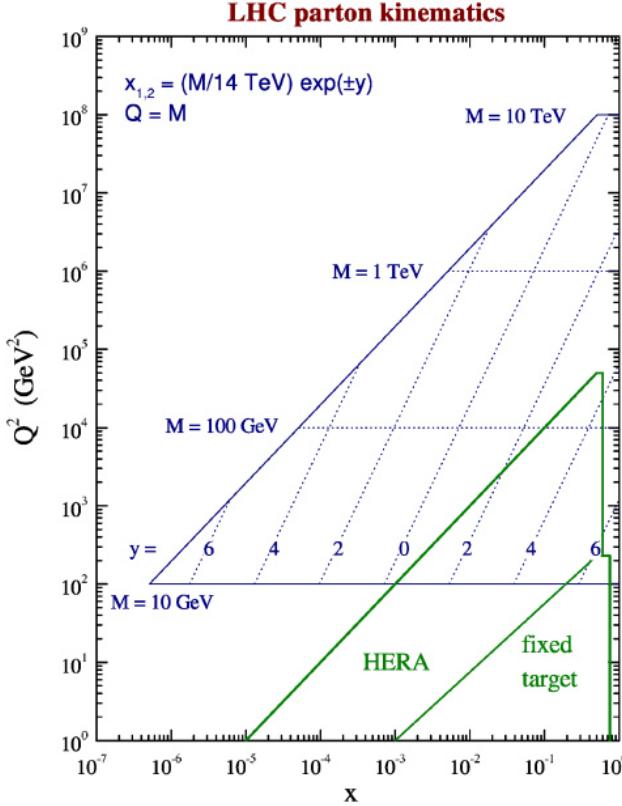


Figure 3.5: Parton kinematics for the regions in (x, Q^2) probed by the LHC and previous DIS and fixed target experiments. The scale Q^2 of a process is often identified with the mass M of a produced resonance. This is indicated by four discrete values on the plot, for the LHC with $\sqrt{s} = 14$ TeV [8].

distributions, so a third term $P(x; a_3, \dots)$, is added, which provides the required flexibility to the PDF parametrization during the fit. This is a smooth function depending on one or more parameters, and separate functions exist for each parton flavor.⁴

The numerical values of the PDF fit parameters are extracted by performing a χ^2 minimization, which compares the measured experimental data to the theoretical expectations obtained with Eq. 3.9. The global minimum of the χ^2 function defines the central, or best-fit, values $\{a_i\}_0$, whose uncertainty is given by the variation around the minimum of the

⁴For example, in the CT14NNLO PDF set, the functions $P(x; a_3, \dots)$ are taken to be a linear combination of specialized orthogonal polynomials. All together, 8 parameters are associated with the valence distributions $u_v(x)$ and $d_v(x)$, 5 with the gluon $g(x)$, and 13 with sea distributions $\bar{u}(x)$, $\bar{d}(x)$, and $\bar{s}(x)$ for a total of $n = 26$ fit parameters.

χ^2 function. DGLAP evolution then allows the post-fit PDFs to be evolved from the initial scale Q_0^2 to higher scales Q^2 , and to be compared against data to evaluate the performance of the fit.

3.2.2 Assessing PDF Uncertainties

In addition to obtaining the best-fit PDFs from the global analysis, it is also necessary assess their associated uncertainties (as in Fig. 3.6). Uncertainties on the PDFs arise from the experimental uncertainties of the input measurements and the theoretical assumptions made in the global fit. These uncertainties can be propagated to the PDFs, which allows for a standard **PDF uncertainty** to be defined. The PDF uncertainty is one of the major contributing factors to the overall theoretical uncertainty on the hadronic cross section calculation of Eq. 3.6, and therefore must be assessed when performing a search or measurement at the LHC.

The PDF uncertainties are estimated using the **Hessian method** [84, 85]. Here, a quadratic expansion of the χ^2 function is made around its global minimum, which results in an $n \times n$ matrix of second-order partial derivatives known as the **Hessian matrix**. The Hessian matrix is then diagonalized by a set of n orthonormal eigenvectors, where n is the number of free parameters in the PDF fit (e.g. $n = 26$ in the CT14NNLO PDF set). The PDFs are varied along the positive and negative direction of each eigenvector, which correspond to unique variations of the nominal PDF parameters. Each variation results in a new PDF, and the set of all $2n$ (one set f_i^+ for up and one set f_i^- for down) eigenvector variations is referred to as a **PDF error set**.⁵

⁵Due to the diagonalization of the Hessian matrix, the eigenvectors are admixtures of the n global PDF fit parameters, so an individual eigenvector, for example, can be composed of contributions from all parton flavors.

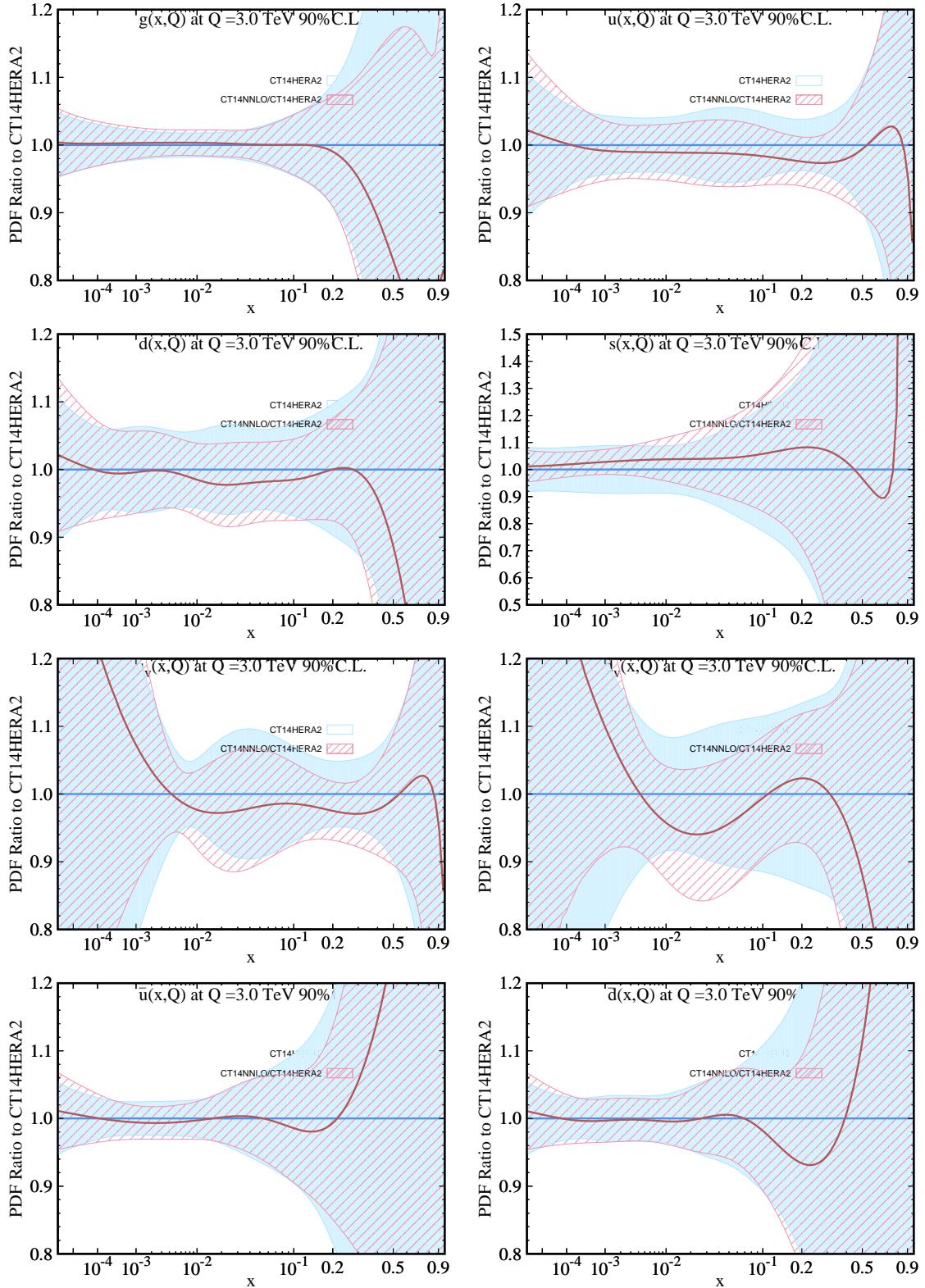


Figure 3.6: PDF uncertainties associated with the $g(x)$, $u(x)$, $d(x)$, $s(x)$, $u_v(x)$, $d_v(x)$, $\bar{u}(x)$, and $\bar{d}(x)$ distributions in the CT14NNLO and CT14Hera2 PDF sets evaluated at a scale of $Q = 3$ TeV.

With the PDF error set, the PDF uncertainty for LHC observables can be obtained. For example, in the prediction of generic hadronic cross section σ , the PDF uncertainty $\Delta\sigma_{\text{PDF}}$, is determined by

$$\begin{aligned}\Delta\sigma_{\text{PDF}}^+ &= \sqrt{\sum_{i=1}^n [\max(\sigma(f_i^+) - \sigma(f_0), \sigma(f_i^-) - \sigma(f_0), 0)]^2} \\ \Delta\sigma_{\text{PDF}}^- &= \sqrt{\sum_{i=1}^n [\max(\sigma(f_0) - \sigma(f_i^+), \sigma(f_0) - \sigma(f_i^-), 0)]^2},\end{aligned}\tag{3.10}$$

where $\sigma(\cdot)$ is evaluated using f_0 , the nominal set of PDFs obtained in the global fit, and f_i^+ and f_i^- , the + and – variations of PDF eigenvector i in the error set of n , respectively. The $\Delta\sigma^+$ term adds the PDF error contributions in quadrature that lead to an increase in the nominal value of the observable σ , while $\Delta\sigma^-$ adds those that lead to a decrease, thus providing a definition for asymmetric errors.

The plot in Figure 3.6 shows the PDF uncertainties for several individual parton flavors in the **CT14NNLO** and **CT14Hera2** PDF sets. Generally speaking, the PDF uncertainties increase at the extreme values of x where input data is scarce. The valence distributions $u_v(x)$ and $d_v(x)$ are least precisely known at very low x , roughly for $x < 10^{-2}$, and the $u_v(x)$ distribution is determined better than $d_v(x)$ at high x , roughly $x > 0.1$. The $\bar{u}(x)$ and $\bar{d}(x)$ distributions are both least precisely known at high x , again where input data is difficult to obtain. Each of these distributions plays an important role in the initial-state quark-antiquark annihilation that results in the Drell-Yan process.

Each PDF set is used to model Standard Model processes in later sections of the dissertation. The **CT14NNLO** set is used to model the Drell-Yan background in the dilepton analysis, presented in the part III, while the **CT14Hera2** set is used to perform exploratory studies in

an attempt to reduce of the high- x PDF uncertainty in part IV.

3.2.3 Summary

PDFs play an essential part in all measurements and searches performed at the LHC. The state-of-the-art PDF sets provided by the PDF collaborations come with an accuracy of a few to several percent over the wide kinematic range probed by the LHC, a significant accomplishment.

However, in the regions where no little or no data are provided for global analyses, such as at very large or small x , the PDF uncertainty can grow considerably and start to impact the accuracy of the theory prediction for the total cross section. This is especially evident in the high-mass dilepton spectrum of the Drell-Yan process.

In the final chapter, the results of several studies are presented on how to reduce the high-mass PDF uncertainty for future iterations of the dilepton analysis. In the mean time, the next chapter introduces the Drell-Yan process in detail.

Chapter 4

Two Body Scattering and The Drell-Yan Process

In this chapter, the Drell-Yan process is introduced and described in detail. The production of a pair of massive leptons through the Drell-Yan mechanism provides for one of the best motivated channels to search for Z' production, and a detailed understanding therefore is essential to performing a robust search at the LHC.

The goal of this chapter is to calculate the Drell-Yan production cross section, which will shed light on the nature of its production mechanism and kinematics. To accomplish this, the S-matrix and Feynman Rules are introduced, which provide the means to calculate hard scatter cross sections using a diagrammatic technique. With these ideas in place, the Drell-Yan production cross section is calculated to the leading-order in perturbation theory, and the results are discussed, with a focus on the effects of the PDFs at high dilepton mass. Lastly, higher-order diagrams associated with Drell-Yan sub-processes important to the analysis are presented.

4.1 The S-Matrix and Feynman Rules

In the last chapter, it was shown that in order to calculate hadronic cross sections, one must first convolute the PDFs with an expression of the fixed-order partonic cross section $\hat{\sigma}$ for

the scattering process of interest. This section introduces the concepts relevant for actually calculating these rates, given a specific interaction Lagrangian \mathcal{L}_{int} of the Standard Model.

4.1.1 Scattering Amplitudes

In quantum field theory, the operator S that connects and evolves an initial-state of particles $|i\rangle$ undergoing a scattering process to an outgoing final-state $|f\rangle$ is known as the **S-matrix**.

The matrix elements of the S-matrix,

$$S_{fi} = \langle f | S | i \rangle, \quad (4.1)$$

are known as **scattering amplitudes**, which encode the probability amplitudes for a given initial-state to scatter into a certain final-state.¹

In general, the S-matrix can be written,

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dx_1^4 \cdots \int_{-\infty}^{\infty} dx_n^4 T [\mathcal{H}_{\text{int}}(x_1) \cdots \mathcal{H}_{\text{int}}(x_n)], \quad (4.2)$$

which is a sum over all possible interactions. The operator T denotes a time-ordered product, ensuring interactions occur in the correct order, and the particle interactions are codified in terms of an interaction Hamiltonian \mathcal{H}_{int} .²

As the S-matrix contains a complete description of the scattering process, what becomes clear from Eq. 4.2 is that it can be used to calculate scattering amplitudes order-by-order according to the series expansion. This forms the starting point to the perturbative approach

¹In QED for example, the initial- or final-state is specified by a definite number of photons, electrons, and positrons with given spins, momenta, and polarizations [86].

²The Hamiltonian density \mathcal{H} is fully equivalent to, and can be derived from, the more standard Lagrangian density \mathcal{L} . Its use here is due to working in the interaction picture.

used in the calculation of all Standard Model scattering processes.

A crucial point of this approach is that each \mathcal{H}_{int} term carries with it one power of the coupling constant $\alpha(Q^2)$ for the Standard Model interaction under consideration.³ The implication is that the perturbative description is only valid when the value of the coupling constant is small, such that $\alpha \ll 1$ for the Q^2 range under study. If this is the case, the approach is highly predictive, and the calculation of only the first few terms in the expansion can result in an accurate approximation of the full amplitude.

For example, the electromagnetic coupling constant is $\alpha_{EM} \approx 1/137$, so higher-order corrections to electromagnetic processes quickly become vanishingly small, making QED an extremely precise theory.

In this way, the infinite number of possible (and increasingly unlikely) particle interactions can be split up into different orders of probability. In the series expansion of S , a calculation truncated at the first non-zero term is called a calculation at the **Leading-Order** (LO). A calculation including the next higher-order term is a calculation at the **Next-to-Leading-Order** (NLO), followed by the **Next-to-Next-to-Leading-Order** (NNLO), etc.

4.1.2 Feynman Diagrams

The calculation of physical cross sections from Eq. 4.2 is a technically challenging task, since it requires the evaluation of complex integrals over a multiple variables.

Feynman's contribution was to recognize regular structures within the integrals and devise a way to represent them graphically as **Feynman diagrams**. Each part of a diagram has a mathematical interpretation, and the **Feynman rules** [87] provide a way to translate them into an expression for the scattering amplitude, thereby circumventing the need for a detailed

³ \mathcal{H}_{int} is related to \mathcal{L}_{int} through a Legendre transformation, which specifies the relevant coupling.

calculation altogether.

The diagrammatic approach has the advantage of being intuitive, depicting the scattering process in terms of the exchange of force particles, and the production and decay of short-lived matter particles. The particle interactions are written in terms of propagators and vertex factors (as seen in Fig. 2.6 and Fig. 2.7 for the electroweak interaction).⁴ Together vertices and propagators can be combined following the Feynman rules to make Feynman diagrams.

The scattering amplitude S of Eq. 4.2 can be expressed in terms of the **Feynman amplitude** \mathcal{M} ,⁵ such that,

$$S_{fi} = \delta_{fi} - i\mathcal{M}_{fi} (2\pi)^4 \delta^4(k_f - p_i), \quad (4.3)$$

where k_i and k_f represent the four-momentum sums of the initial- and final-state particles, respectively. The entire problem boils down to obtaining the amplitude \mathcal{M} , which amounts to drawing all topologically distinct Feynman diagrams for a given order n in the S-matrix expansion according to the Feynman rules.

4.1.3 Two Body Scattering

With knowledge of the Feynman Rules, and the Lagrangian of Eq. 2.36, which specifies the electromagnetic and neutral current interactions, it is enough to write down the Drell-Yan matrix element. The final step is to combine the matrix element with known kinematic factors to obtain an expression for the cross section.

⁴The vertex factors are derived from interaction Lagrangian of a given theory. Likewise, the propagators are derived from free field Lagrangian.

⁵Likewise, the term “matrix element” is often used interchangeably.

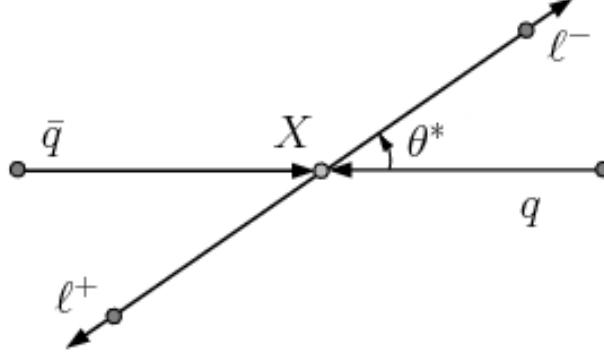


Figure 4.1: Two body scattering as described by Eq. 4.5 [11].

In many body scattering (i.e. $1 + 2 \rightarrow 3 + 4 + \dots + n$), the matrix element \mathcal{M} appears as a factor in the differential cross section as follows [87],

$$d\hat{\sigma} = (2\pi)^4 \delta^4(p_1 + p_2 - k_3 - \dots - k_n) \frac{1}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} |\mathcal{M}|^2 \prod_{j=3}^n \frac{1}{2E_f} \frac{d^3 \mathbf{k}_j}{(2\pi)^3}, \quad (4.4)$$

where the four-momenta of the initial- and final-state particles are denoted p_i and k_j , respectively.

Equation 4.4 reduces significantly in the special case of two body scattering. If the collision takes place in the center-of-mass (CM) frame where $E_1 = E_2 = E$ and the colliding partons are massless and have equal and opposite momenta, then [87],

$$d\hat{\sigma} = \frac{1}{64\pi^2 (E_1 + E_2)^2} \frac{|\mathbf{k}_f|}{|\mathbf{p}_i|} |\mathcal{M}|^2 d\Omega, \quad (4.5)$$

where $|\mathbf{k}_f|$ and $|\mathbf{p}_i|$ are the magnitudes of the outgoing and incoming momentum of either particle, respectively (See Fig 4.1).

The dynamics of the process are entirely encoded in the amplitude \mathcal{M} , with the absolute square proportional to the differential cross section. Once the amplitude \mathcal{M} known, the calculation reduces to pure kinematics, independent of the form of the interaction, making

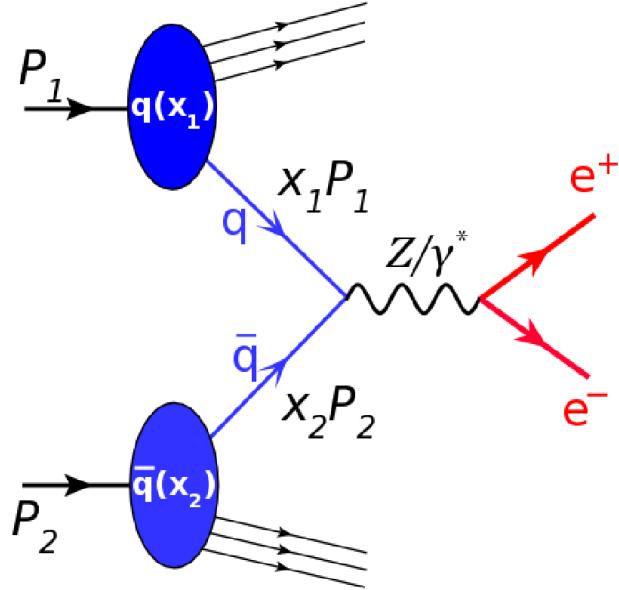


Figure 4.2: The Drell-Yan process as observed at the LHC [12].

two body scattering particularly simple.

4.2 The Drell-Yan Process

The Drell-Yan process refers to the production of a massive pair of charged leptons $\ell^+\ell^-$ from $q\bar{q}$ annihilation (See Fig. 4.2), as predicted by Drell and Yan in 1970 [88].⁶ In the low-energy Drell-Yan process, a quark and antiquark annihilate through the production of a virtual photon γ^* , which then decays into a pair of charged leptons. The angular distributions of the leptons depend on the precise nature of the interaction, which is specified by the matrix element \mathcal{M} .

Since present day colliders operate at sufficiently high CM energies, it is possible for the Drell-Yan process to proceed through the exchange of an on-shell Z boson as well. At the LHC, the γ^* and Z processes interfere, and must, therefore, be considered together in the

⁶The leptons can be electrons, muons, or taus. The focus of this dissertation will be the e^+e^- electron channel.

calculation of Drell-Yan observables, such as the total cross section. In this way, the high-energy LO process is expressed $q\bar{q} \rightarrow \gamma^*/Z \rightarrow \ell^+\ell^- + X$, where X denotes any additional final-state particles (e.g. radiated partons, the underlying event, etc.).

Together, the mass $m_{\ell\ell}$ of the dilepton system and the mass $m_{\gamma^*/Z}$ of the intermediate state γ^*/Z can be related to the hard scale Q^2 of the collision with

$$Q^2 = (p_1^\mu + p_2^\mu)^2 = \left(k_{\ell^+}^\mu + k_{\ell^-}^\mu \right)^2 = m_{\gamma^*/Z}^2 = m_{\ell\ell}^2 = x_1 x_2 s, \quad (4.6)$$

where $k_{\ell^+}^\mu$ and $k_{\ell^-}^\mu$ are the four-vectors of the final-state dilepton pair, p_1^μ and p_2^μ are the four-vectors of the initial-state quark-antiquark pair, and \sqrt{s} is the CM energy of the collision. As a consequence, the dilepton invariant mass will exhibit a spectrum of values consistent with the x values carried by the colliding quark and antiquark in accordance with the PDFs.

The Drell-Yan process itself is important for several reasons. First, it provides a means to compare theoretical calculations to experimental measurements, allowing for stress tests of pQCD.⁷ Second, it provides insight into the nature of the PDFs, allowing for better constraints to be placed on their parametrization.⁸ Third, it is the dominant background to many searches for new physics, and specifically to the one presented later in this dissertation. For these reasons, significant theoretical effort has been committed to its understanding, with the calculation of differential cross section established at up to NNLO [89, 90].

To gain insight into the dynamics of this process, the LO Drell-Yan cross section calculation is now performed using the tools developed thus far in this chapter. Later, the

⁷Consider the following example: at the LO, the transverse momentum $p_T^{\ell\ell}$ of the dilepton system is zero by conservation of momentum, but at higher orders, the initial-state quarks can radiate one or more energetic gluons, which results in a non-zero dilepton $p_T^{\ell\ell}$. Precise measurements of this quantity, then, provide a way to directly test the predictions of pQCD.

⁸Specifically, since an antiquark is needed to initiate the Drell-Yan process, a cross section measurement provides sensitivity to the $\bar{u}(x)$ and $\bar{d}(x)$ sea distributions.

results of the calculation are qualitatively extended by discussing the effects of perturbative higher-order QCD and EW corrections.

4.2.1 Leading-Order Cross Section

To illustrate the Feynman approach, the LO expression for the Drell-Yan cross section is calculated using Feynman diagrams. The approach is to write down the matrix element \mathcal{M} , which encodes the dynamics of the process, and find its absolute square for insertion into Eq. 4.5 to obtain an expression for the total cross section.

The calculation proceeds in the CM frame and in the massless limit. The initial- and final-state particle four-vectors are given by

$$\begin{aligned} p_1^\mu &= E(1, 0, 0, 1) \\ p_2^\mu &= E(1, 0, 0, -1) \\ k_1^\mu &= E(1, \sin\theta, 0, \cos\theta) \\ k_2^\mu &= E(1, -\sin\theta, 0, -\cos\theta), \end{aligned} \tag{4.7}$$

where p_i^μ refer to the quarks and k_i^μ to the leptons (See Fig. 4.1 for example).

For simplicity, it is useful to express the kinematics in terms of three Lorentz invariant quantities, known as the **Mandalstam variables**,

$$\begin{aligned} \hat{s} &= (p_1 + p_2)^2 = (k_1 + k_2)^2 = 4E^2 \\ \hat{t} &= (p_1 - k_1)^2 = (p_2 - k_2)^2 = -\frac{\hat{s}}{2}(1 - \cos\theta) \\ \hat{u} &= (p_2 - k_1)^2 = (p_1 - k_2)^2 = -\frac{\hat{s}}{2}(1 + \cos\theta), \end{aligned} \tag{4.8}$$

where the four-vectors of Eq. 4.7 have been used to simplify each expression. The parameter

\hat{s} can be identified with the squared CM energy of the $q\bar{q}$ initial-state, or equivalently, the $\ell^+\ell^-$ final-state in accordance with Eq. 4.6, and, \hat{t} and \hat{u} introduce angular dependencies.

The Drell-Yan process can proceed through a virtual photon γ^* or an on-shell Z boson at the LHC in accordance with the interaction Lagrangian of Eq. 2.36. The total amplitude \mathcal{M} can be written as a sum, with $\mathcal{M} = \mathcal{M}_{\gamma^*} + \mathcal{M}_Z$. Figure 4.3 depicts the corresponding Feynman diagrams.

The Feynman rules state that each fermion line should be tracked backwards through the diagram, and the associated factors are written down in the direction opposite to the arrows. With this, the matrix element for the virtual photon γ^* is

$$i\mathcal{M}_{\gamma^*} = \left[\bar{u}(k_1)(-iQ^\ell e\gamma_\mu)v(k_2) \right] \frac{-ig^{\mu\nu}}{(p_1 + p_2)^2} [\bar{v}(p_2)(-iQ^q e\gamma_\nu)u(p_1)], \quad (4.9)$$

and similarly, the matrix element for the Z boson is

$$\begin{aligned} i\mathcal{M}_Z &= \left[\bar{u}(k_1) (-ig_Z \gamma_\mu (g_L P_L + g_R P_R)) v(k_2) \right] \\ &\times \frac{-ig^{\mu\nu}}{(p_1 + p_2)^2 - M_Z^2 + iM_Z\Gamma_Z} \\ &\times [\bar{v}(p_2) (-ig_Z \gamma_\nu (g_L P_L + g_R P_R)) u(p_1)]. \end{aligned} \quad (4.10)$$

The initial- and final-state fermions are associated with pairs of Dirac spinors u and v , which label the momenta and spin states of the particles. Writing the spinor indices explicitly allows the total matrix element \mathcal{M} to be expressed concisely as

$$\mathcal{M}_{ij} = \mathcal{G}_{ij} [\bar{u}_j(k_1)\gamma^\mu v_j(k_2)] [\bar{v}_i(p_2)\gamma_\mu u_i(p_1)], \quad (4.11)$$

where the subscripts i and j index the helicity of the initial- and final-state fermions, and

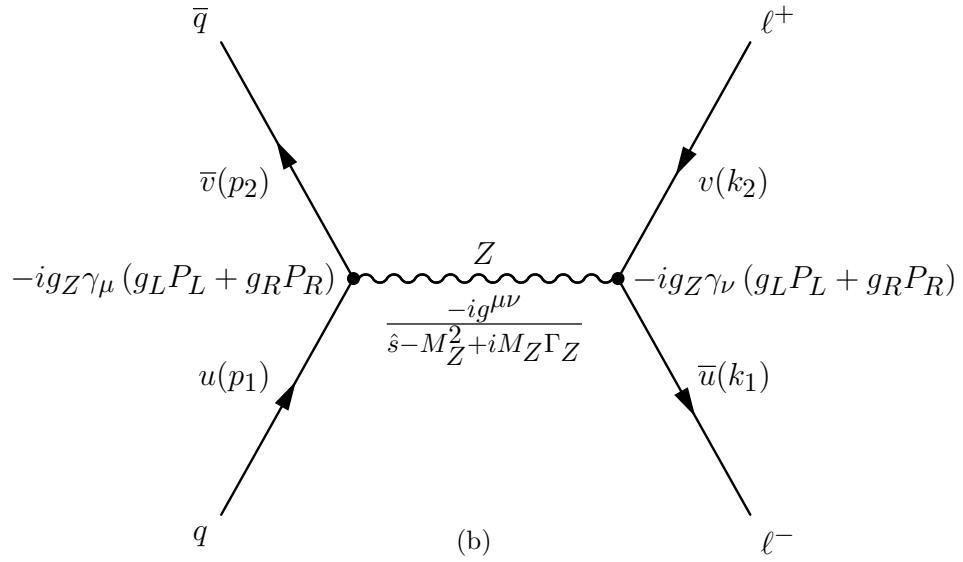
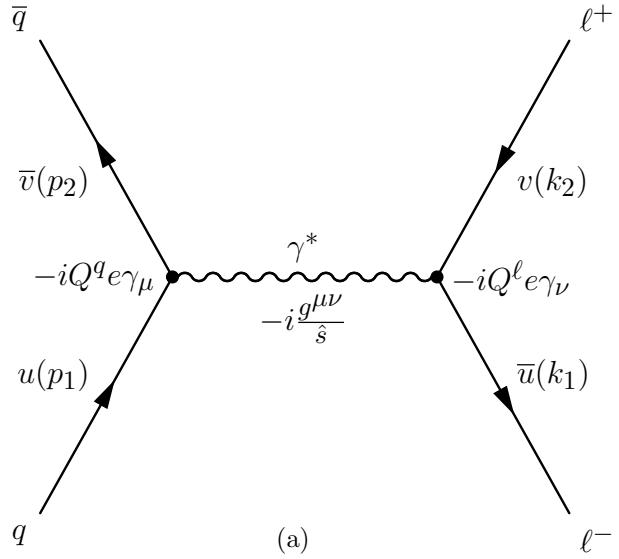


Figure 4.3: The LO Feynman Diagram for Drell-Yan production mediated by a virtual photon γ^* (a), and on-shell Z boson (b). The propagator and vertex factors of each diagram are depicted. The initial- and final-state particles are represented by the Dirac spinors u and v , which label their momenta and spin.

\mathcal{G}_{ij} groups together the couplings, charges, and propagator factors with

$$\mathcal{G}_{ij}(\hat{s}) = e^2 \frac{Q^\ell Q^q}{\hat{s}} + g_Z^2 \frac{g_i^\ell g_j^q}{\hat{s} - M_Z^2 + iM_Z\Gamma_Z}. \quad (4.12)$$

The electroweak charges (Q , g_L , and g_R) for the incoming and outgoing fermions are provided in Table 2.4, and the mass and decay width of the Z Boson are denoted M_Z and Γ_Z , respectively.

To enter the expression for the two body cross section of Eq. 4.5, the modulus squared of the matrix element \mathcal{M} must be evaluated. As the cross sections at the LHC are unpolarized, an average over initial-state helicity modes and sum over final-state modes must be taken. Since there are two particles in the initial-state, a factor of $1/4$ is included for the average, and from angular momentum conservation, the four possible helicity states are \mathcal{M}_{LL} , \mathcal{M}_{LR} , \mathcal{M}_{RL} , \mathcal{M}_{RR} .

Therefore, an expression for the squared total matrix element can be written,

$$|\mathcal{M}|^2 = \sum_{i,j \in \{L,R\}} |\mathcal{M}_{ij}|^2. \quad (4.13)$$

The evaluation of each term in the sum involves calculating traces of products of Dirac γ matrices, and the result is

$$\begin{aligned} |\mathcal{M}_{LL}|^2 &= |\mathcal{G}_{LL}(\hat{s})|^2 4\hat{u}^2, & |\mathcal{M}_{LR}|^2 &= |\mathcal{G}_{LR}(\hat{s})|^2 4\hat{t}^2 \\ |\mathcal{M}_{RL}|^2 &= |\mathcal{G}_{RL}(\hat{s})|^2 4\hat{t}^2, & |\mathcal{M}_{RR}|^2 &= |\mathcal{G}_{RR}(\hat{s})|^2 4\hat{u}^2. \end{aligned} \quad (4.14)$$

Inserting the matrix element \mathcal{M} of Eq. 4.14 into the expression for the differential cross section of Eq. 4.5, accounting for the spin factor of $1/4$, and integrating over solid angle

leads to the following partonic cross section,

$$\hat{\sigma}(q\bar{q} \rightarrow \gamma^*/Z \rightarrow \ell^+\ell^-) = \frac{\hat{s}}{48\pi} \left[|\mathcal{G}_{LL}(\hat{s})|^2 + |\mathcal{G}_{LR}(\hat{s})|^2 + |\mathcal{G}_{RL}(\hat{s})|^2 + |\mathcal{G}_{RR}(\hat{s})|^2 \right], \quad (4.15)$$

which completes the partonic component of the LO calculation.

Equation 4.15 has a strong dependence on the partonic CM energy, and hence the dilepton mass by Eq. 4.6. In the high-mass limit of $\hat{s} \gg M_Z^2$, far from the Z boson mass peak, the cross section is inversely proportional to the partonic CM energy $\hat{\sigma} \sim 1/\hat{s}$, since the amplitudes $|\mathcal{G}_{ij}|^2 \sim 1/\hat{s}^2$. Thus, in this regime, the cross section exhibits a smooth and rapidly falling behavior.⁹

On the other hand, if $\hat{s} \approx M_Z^2$, a dramatic change to the shape of the smoothly falling spectrum occurs, as it is replaced by a Breit-Weigner resonance due to the exchange of an on-shell Z boson, or a possible Z' boson at $\hat{s} \approx M_{Z'}^2$.

Equation 4.15 also makes manifest the interference effect due to the exchange of either the γ^* or the Z boson in the intermediate state. For example, cross terms from each of processes appear when squaring the amplitudes $|\mathcal{G}_{ij}|^2$. Interference effects like this can play an important role if a Z' boson is exchanged as well, which will be discussed further in Section 5.3.

The final step of the calculation is to convolute Eq. 4.15 with the PDFs to obtain the LO estimate of the total cross section. This requires two modifications. First, since the colliding $q\bar{q}$ pair annihilates into a colorless final-state, it must carry one of three possible color combinations ($r\bar{r}, b\bar{b}, g\bar{g}$), so an average over initial-state colors must be taken as well. This is usually expressed in terms of the QCD color factor with $N_C = 1/3$. Second, since

⁹As the high-mass region is the main target of the search, any observed discrepancy that is not smoothly falling might then be an indication of new physics.

the incoming quark-antiquark system exhibits a spectrum of CM energies due to the PDFs, the replacement $\hat{s} \rightarrow x_1 x_2 s$ is made, where \sqrt{s} is the CM energy of the LHC.

Folding in the PDFs, as described by Eq. 3.6, results in an expression for the total hadronic cross section,

$$\begin{aligned}\sigma = & \sum_q \int dx_1 dx_2 [f_q(x_1, \hat{s}) f_{\bar{q}}(x_2, \hat{s}) + f_{\bar{q}}(x_2, \hat{s}) f_q(x_1, \hat{s})] \\ & \times \frac{\hat{s}}{48\pi N_C} \left[|\mathcal{G}_{LL}(\hat{s})|^2 + |\mathcal{G}_{RR}(\hat{s})|^2 + |\mathcal{G}_{LR}(\hat{s})|^2 + |\mathcal{G}_{RL}(\hat{s})|^2 \right],\end{aligned}\quad (4.16)$$

where the sum runs over the quark flavors q , and $f_q(x)$ and $f_{\bar{q}}(x)$ are the associated PDFs, evaluated at the scale $Q^2 = \hat{s} = m_{\ell\ell}^2$ of the event.

4.2.2 PDF Uncertainty

Equation 4.16 is the result of the LO calculation. The expression for the cross section is given by a sum over all quark-antiquark flavors, weighted by various coupling and propagator factors. This is significant because it indicates the Drell-Yan process has direct access to individual parton flavors and their PDFs through its initial-state. Specifically, the Drell-Yan cross section receives significant contributions from the less well measured sea quark PDFs.

Therefore, the Drell-Yan process can be used to get a handle on the **flavor decomposition** of the sea, as well as the valence distributions $u_v(x)$, $d_v(x)$, and the gluon $g(x)$ PDFs.¹⁰ Furthermore, the Drell-Yan process at the LHC is somewhat unique in that it provides complimentary information in (x, Q^2) not well covered by previous experiments, as well as extending this range significantly (See Fig. 3.5).

For the purposes of the dilepton analysis, knowledge of Drell-Yan production is important

¹⁰The Drell-Yan processes at NLO is initiated by quark-gluon (and antiquark-gluon) scattering, providing a probe of the gluon distribution $g(x)$ (see Sec. 4.2.3).

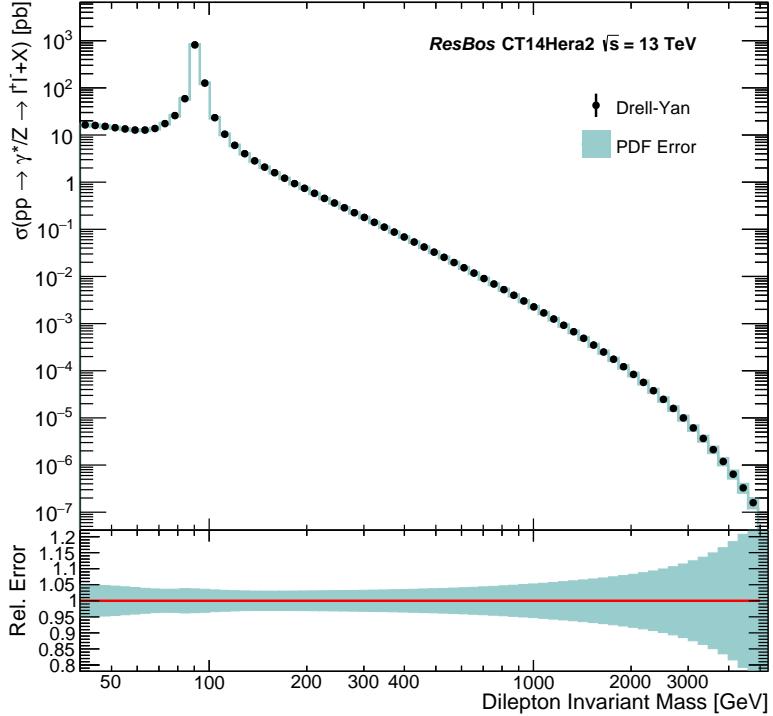


Figure 4.4: The invariant mass distribution of dilepton pairs $m_{\ell\ell}$ produced via the Drell-Yan mechanism at $\sqrt{s} = 13$ TeV. The MC generator RESBOS, configured with the CT14Hera2 PDF set, was used for the prediction. The PDF error, calculated with the symmetrized form of Eq. 3.10, is indicated in the sub-panel, and steadily increases with dilepton mass.

because it is the dominant background to the dilepton search, so understanding it well is paramount to performing a successful analysis. However, as will be discussed in subsequent chapters, the PDF uncertainty is the dominant theoretical uncertainty at high dilepton invariant mass, the primary region of interest for the search.

This uncertainty results from the imprecise knowledge of the valence and sea quark distributions at high x , which are required to produce dilepton states with large invariant masses as indicated by $m_{\ell\ell}^2 = x_1 x_2 s$. However, at the same time, these are the very distributions probed by the Drell-Yan process via Eq. 4.16, which means dedicated measurements of the Drell-Yan cross section may be used to constrain the PDFs affecting the high-mass search region of the dilepton analysis.

The effect of the PDFs on the description of the dilepton spectrum is shown in Figure 4.4. Here, the RESBos [91, 92, 93] MC generator, configured with the CT14Hera2 [94] set of PDFs was used.¹¹ Samples were generated in slices of dilepton invariant mass, weighted, and stitched together. The PDF error is calculated bin-by-bin with the CT14Hera2 error set PDFs using the symmetrized form of Eq. 3.10. At $m_{\ell\ell} = 1$ TeV, the uncertainty is roughly 5%, but by $m_{\ell\ell} = 5$ TeV it stands at 30%.

Accurate knowledge of the PDFs is crucial for direct searches of new physics (like in this dissertation), but it is also crucial for indirect searches through precision measurements. Future iterations of the dilepton analysis could benefit greatly from new global fit measurements, which translate directly into improved discovery potential, since the PDF uncertainty would become further constrained. At the same time, LHC measurements of Drell-Yan might help resolve disagreements between the various global fit datasets offered by the PDF fitting groups, via its enhanced sensitivity to the strange quark and gluon PDFs.

In the final chapter, an exploratory study is presented which aims to reduce the PDF uncertainty associated with Drell-Yan production by incorporating new measurements in the PDF global fits. The achieved error reduction is quantified and the results show future LHC measurements can indeed improve the PDF descriptions.

4.2.3 Higher-Order Corrections

The Feynman diagrams in Figure 4.3 show the LO contributions to the Drell-Yan process, but countless other **sub-processes** contribute to $q\bar{q} \rightarrow \gamma^*/Z \rightarrow \ell^+\ell^- + X$ as well. These sub-processes include additional final-state particles that arise from initial- or final-state

¹¹MC generators are the subject of Section 9.1. The generator RESBOS calculates fully differential cross sections for electroweak processes, such as neutral current Drell-Yan, at up to the NNLO in QCD.

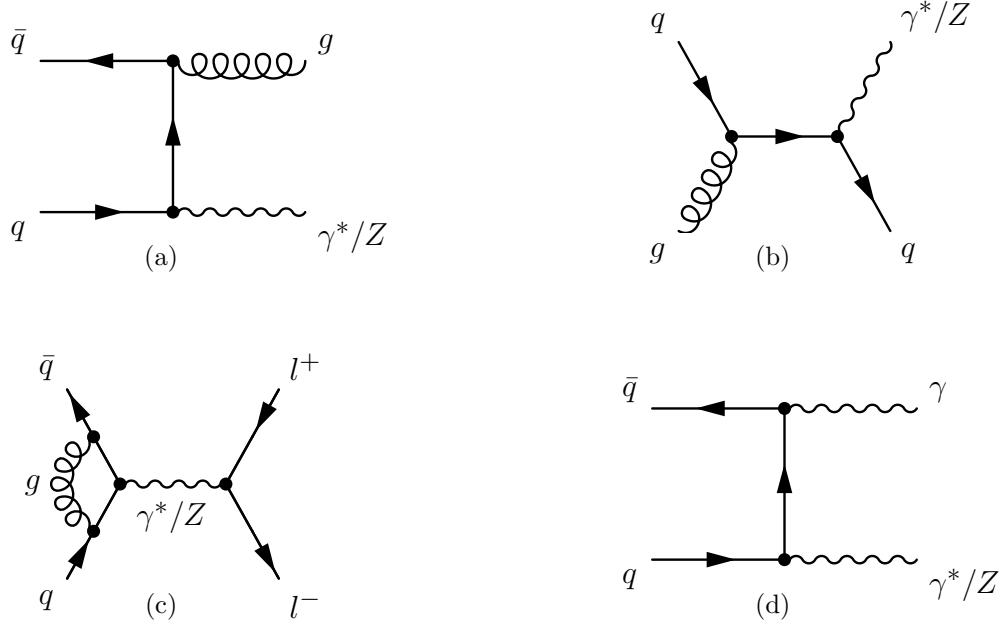


Figure 4.5: Example higher-order contributions to the Drell-Yan process: NLO initial-state gluon radiation $q\bar{q} \rightarrow \gamma^*/Z + g$ (a), NLO initial-state quark radiation $qg \rightarrow \gamma^*/Z + q$ (b), NLO virtual gluon correction $q\bar{q} \rightarrow \gamma^*/Z$ (c), NLO final-state photon radiation $q\bar{q} \rightarrow \gamma^*/Z + \gamma$ (d).

radiation (i.e. from photons or gluons), as well as internal **particle loops**, which arise from the combination of two or more interaction vertices according to the Feynman rules.¹²

Sub-processes like this are known as **higher-order corrections**, since they originate from higher-order terms in the S-matrix expansion. For example, at LO that only quark-antiquark annihilation can occur, while at NLO quark-gluon (and antiquark-gluon) scattering becomes possible, and at NNLO gluon-gluon scattering opens up as well. Some examples of this are shown in Figure 4.5. Each sub-processes acts to correct to the LO estimate obtained for the Drell-Yan cross section.

In general, higher-order corrections due to the strong interaction are larger than those of EW interaction at the same order in perturbation theory, since the strong coupling α_S is

¹²Each order in the series expansion of \mathcal{M} includes one additional vertex factor, and therefore, the number of diagrams at a given order increases like a factorial (due to the possible permutations), resulting in complex multi-loop diagrams.

much larger than the electromagnetic α_{EM} . QCD corrections to the Drell-Yan process can be sizeable, and both types are accounted in the description of the Drell-Yan background for the dilepton analysis (as described in Sec. 9.4.5).

Higher-order corrections also introduce qualitatively new features to the kinematics of the LO process. For example, at LO the dilepton transverse momentum $p_T^{\ell\ell}$ is identically zero, but at NLO and beyond, it is non-zero due to the emission of initial-state radiation, allowing for the γ^*/Z to recoil in the transverse plane of the reaction.

Beyond LO calculations can be extremely challenging. It seems like at this point, all that is required to calculate the cross section to any accuracy is to draw the relevant Feynman diagrams for the process under consideration, and follow feynman rules to obtain an expression for the matrix element. However, a technicality arises beyond the LO in the S-matrix expansion: loop diagrams are encountered, which when evaluated lead to divergent integrals in the calculation of the cross section. Finite results can only be obtained after the **renormalization** of the theory, which is a way to systematically handle the aforementioned divergences.

Chapter 5

Physics Beyond the Standard Model

The Standard Model is an extremely successful theory and one of the best tested in the history of physics. However, despite its outstanding successes, there are multiple experimental and theoretical indications that the Standard Model is an incomplete theory, which suggests it is only a part or an aspect of one more fundamental. In the following chapter, the limitations of the Standard Model are discussed and an overview of several Beyond the Standard Model (BSM) theories is presented. The emphasis is placed on models that predict new high-mass dilepton resonances, which are the main target of the search performed in this dissertation.

5.1 Limitations of the Standard Model

Over the past several decades, experiments in the fields of cosmology, astronomy, neutrino physics, and collider physics have uncovered phenomena that cannot naturally be explained within the framework of the Standard Model. These phenomena include:

Neutrino Masses: Neutrinos are massless in the Standard Model, but recent experiments [95] have established that the three generations of neutrinos mix, or undergo **flavor oscillations**, as they travel, and therefore must have tiny but non-zero masses.¹ The

¹This can only happen if there is a mismatch between the neutrino mass and flavor eigenstates analogous to the situation in the quark sector.

matrix responsible for the mixing is known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, just like the CKM matrix of the quark sector. However, unlike the CKM matrix, the off-diagonal entries of the PMNS matrix are large, indicating there is a much more substantial amount of mixing in the lepton sector. The reason for such asymmetric behavior between the quark and lepton sectors is unknown, as are the origin and nature of the neutrino masses; specifically, why they are so many orders of magnitude smaller than the other fundamental fermions. The current limit on the neutrino masses is $m_{\nu e} + m_{\nu \mu} + m_{\nu \tau} < 0.12$ eV (95% C.L.) [96]. The neutrinos are the only electrically neutral fermions in the Standard Model, and as such, they could be described with **Majorana**,² as opposed to **Dirac** fields, unlike all other Standard Model fermions. If this is the case, it has profound implications for the mechanism responsible for neutrinos mass generation [97].³

Dark Matter and Dark Energy: Cosmological measurements have determined that only 5% of the universe is comprised of ordinary baryonic matter, as described by the Standard Model, while the rest consists of entirely unknown forms, known as **dark matter** ($\approx 23\%$) and **dark energy** ($\approx 72\%$). The first evidence of dark matter came from measurements of the orbital velocities of stars and gas in spiral galaxies [99]. The observed dynamics could only be explained if there was some non-luminous component of matter present. Now the existence of dark matter is supported by many other astro-

²A Majorana fermion is its own antiparticle.

³If neutrinos are Dirac particles, their masses can be generated through the Higgs Mechanism by introducing right-handed neutrinos into the Standard Model (analogous to the generation of the up- and down-type quark masses from Eq. 2.27). However, since the neutrino masses are known to be at or below the eV scale, their Yukawa couplings would be unnaturally small. The alternative is to consider them as Majorana particles, in which case, the See-Saw mechanism would be responsible for their mass generation. Left-Right symmetric models [98] can naturally accommodate these ideas, which predict heavy right-handed versions of the W and Z bosons of the Standard Model, making them a natural target for the search performed in this dissertation.

physical and cosmological observations, including gravitational lensing measurements and features in the cosmic microwave background [100]. Due to its observed properties, a potential dark matter particle cannot interact via the strong or electromagnetic forces. The neutrino is the only particle in the SM for which this applies, but neutrinos can only account for a small fraction of the observed dark matter because they are so light, so a viable candidate is still missing. Likewise, the nature of dark energy is just as mysterious. Astronomers have measured the rate of expansion of the universe, and determined it is accelerating [101]. This can be accommodated in the framework of general relativity with the addition of a small non-zero cosmological constant; however, attempts to derive its value within the Standard Model lead to predictions vastly different from observation.

Matter-Antimatter Asymmetry: A common assumption is that matter and antimatter are created in equal amounts in the Big Bang. However, another set of cosmological observations indicate that the visible fraction of the universe is composed almost entirely of matter with comparatively little antimatter present [102]. This imbalance is known as the matter-antimatter asymmetry. One way to account for this asymmetry is to introduce interactions that treat matter and antimatter differently. Indeed, CP violating interactions are firmly established in the Standard Model, but their effect is too small to account for the observed imbalance, suggesting new BSM processes may be responsible.

Gravity: The Standard Model provides a successful description of the strong, weak, and electromagnetic forces, but leaves gravity out, as it is currently unknown how to describe gravitational interactions quantum mechanically. Likewise, the Standard Model

does not offer an explanation as to why the gravitational force is so many orders of magnitude weaker than the other fundamental forces. Attempts to develop a quantum theory of gravity consistent with the Standard Model are an active area of research.

In addition to the above experimental phenomena, there are several theoretical aspects of the Standard Model that are not well understood. The following list highlights some of the areas of concern:

Fermion Generations: The Standard Model contains exactly three generations of quarks and three generation of leptons. This is not a consequence of the theory, but rather it is an experimental fact. Each quark and lepton in the Standard Model seems to be *replicated* three times, differentiated only by their masses. Why there are three generations, and why each generation differs so much in mass is currently unknown. A more compelling model would explain the seemingly arbitrary nature of three generations and why there are not more or less.

Hierarchy Problem: Quark and charged lepton masses in the Standard Model are generated through the Higgs mechanism (See Sec 2.3.4). In the Standard Model, the Higgs boson mass receives large quantum corrections through its interaction with virtual particles, specifically the top quark. Whereas most other loop corrections in the Standard Model scale by factors of $\ln \Lambda$, the corrections to the Higgs mass scale additively and quadratically like Λ^2 . If the Standard Model is valid up to the Plank scale, $\Lambda = M_{\text{Plank}} \approx 10^{19} \text{ GeV}$, then the size of these corrections is actually much larger than the observed mass of the Higgs boson. The relation between the bare mass and observed mass of the Higgs boson is $M_H^2 = M_{H,0}^2 + \mathcal{O}(\Lambda^2)$, with $M_{H,0}$ denoting the bare mass and $\mathcal{O}(\Lambda^2)$ the size of the corrections. Since the observed value of M_H is

around 125 GeV, an extremely precise cancellation must take place between the bare mass term and the correction term. This **unnatural** fine tuning of the bare Higgs mass due to the discrepant scales involved is known as the hierarchy problem [103].

Free Parameters: The Standard Model has 19 free parameters. With the introduction of neutrino masses, and the phases and mixing angles of the PMNS matrix, the number of free parameters increases to 26. The values of these parameters cannot be explained by the Standard Model, nor do there seem to be any coherent relations between them; rather they have to be determined directly from experiment. With so many free parameters, the Standard Model seems too *ad-hoc* to be the fundamental theory, suggesting there are unresolved conceptual issues. Therefore, the search for new physics is motivated not just by the experimental observations outlined above, but also by the desire to simplify what seems like an overly complicated and arbitrary theory.

Taken together, these experimental and theoretical difficulties suggest the Standard Model is part of a more fundamental theory. Several candidate theories that attempt to extend the Standard Model are presented in the following section.

5.2 Theories Beyond the Standard Model

Many Beyond the Standard Model theories have been developed in an attempt to address one or more of the limitations presented in Section 5.1, and some of the most well-motivated ones are presented here. They include:

Supersymmetry: In Supersymmetry (SUSY), each particle in the Standard Model is assigned a “superpartner” particle, whose spin differs by 1/2 compared to its Standard

Model counterpart; i.e., each boson receives a fermion partner and vice-versa. Since these partners have not been observed, SUSY must be broken at some higher energy scale, making the partners much heavier than their Standard Model counterparts. SUSY provides a solution to the hierarchy problem. The quadratic divergence in the squared Higgs mass is stabilized through a series of cancellations between loop diagrams involving the superpartners and those involving particles of the Standard Model. Likewise, in SUSY models that conserve R-parity,⁴ the lightest SUSY superpartner (LSP) is stable and can function as a viable dark matter candidate.

Extra Dimensions: Extra Dimensional models extend the Standard Model with extra spatial dimensions in an attempt to solve the hierarchy problem. In these models, the Standard Model particles are localized to the usual (3+1)-dimensions of spacetime, but the graviton, the force mediator of gravitation, is allowed to propagate into the extra spatial dimensions, which explains why gravity seems so much weaker than the other fundamental forces. As a result, the true strength of gravity is comparable to the other forces, and the Plank scale is reduced to the electroweak scale, thus solving the hierarchy problem. Two well-known models are the ADD model [104], proposed by Arkani-Hamed, Dimopoulos, and Dvali, and the Randall-Sundrum (RS) model [105].

Grand Unified Theories: Grand Unified Theories (GUTs) propose that the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry of the Standard Model originates from a single, larger symmetry group, which provides a deeper relation between the quarks and leptons. In GUTs, all interactions are described by a local gauge theory with a single running coupling constant, effectively unifying the separate forces of the Standard Model into

⁴The Standard Model particles are given an R-parity of +1 while the supersymmetric particles receive an R-parity of -1. If R-parity is conserved, then the lightest superpartner is necessarily stable.

a single fundamental interaction. The scale of this unification is known as the GUT scale with $M_{\text{GUT}} \approx 10^{16}$ GeV, similar to how the weak and electromagnetic forces unify at the electroweak scale. The original GUT model was based on the gauge group $SU(5)$ and was proposed by Glashow and Georgi in 1974 [106]. This is the simplest group that can contain the Standard Model; examples of other unifying groups include $SO(10)$ and E_6 [107, 53]. GUTs are attractive because their phenomenology touches on almost every problem outlined in Section 5.1. However, a problematic feature shared by many GUTs is the prediction that the proton decay, with a lifetime much larger than 10^{30} years, which has not been observed experimentally [108]. Likewise gravity is not included as part of the unification.

5.3 Models with Extended Gauge Symmetries

Many BSM theories, including each of those outlined in Section 5.2, predict the existence of one or more new high-mass bosons, which can decay into lepton-antilepton, or **dilepton**, pairs of Standard Model particles. These include such objects as a spin-0 sneutrino of R-parity violating SUSY, a spin-2 graviton as in the RS extra dimensional model, or a spin-1 force mediating gauge boson that arises from a new $U(1)$ symmetry, predicted in many GUTs [109]; and these are just to name a few.⁵ Because this feature is common to so many extensions of the SM, searches for new high-mass bosons are highly motivated, and are therefore some of the first to be conducted at newly commissioned hadron colliders.

The focus of this dissertation will be on the case where the Standard Model gauge group

⁵Of course there could also exist particles that no one has yet theorized.

is extended to include at least one additional $U(1)'$ symmetry [107, 110], so that it becomes

$$SU(3)_C \times SU(2)_L \times U(1)_Y \times U(1)'). \quad (5.1)$$

The breaking of this symmetry results in a new massive neutral spin-1 gauge boson, known as the Z' boson, the signature of which would be a resonance in the invariant mass spectrum of its decay products at the pole mass of the new particle.

For the analysis presented in this dissertation, the Z' decay products are taken to be the electron and positron; i.e., the search is performed in the so-called **dielectron** final-state, or the “electron channel.” Moreover, as part of the dilepton analysis, a search is also performed in the **dimuon** final-state, or “muon channel,” but this is not presented here (see Ref. [111] for details).

5.3.1 Dilepton Phenomenology

The general Lagrangian for the interaction between a hypothetical Z' state and pair of fermions is [110]

$$\mathcal{L}_{Z'} = g_{Z'} \bar{f} \gamma^\mu (z_{f_L} P_L + z_{f_R} P_R) f Z'_\mu, \quad (5.2)$$

where $g_{Z'}$ is the new $U(1)'$ gauge coupling constant, z_{f_L} and z_{f_R} are the left- and right-handed fermion charges under the new $U(1)'$ symmetry, and P_L and P_R are the standard projection operators. This interaction is phenomenologically identical to the Standard Model neutral current interaction defined in Eq. 2.36, and is only differentiated by the values taken by the new $U(1)'$ fermion charges and gauge coupling constant.⁶ The associated Feynman

⁶For the search performed in this dissertation, **family universality** is assumed among the $U(1)'$ fermion charges, such that, for example, the values taken by the electron, muon, and tau are identical.

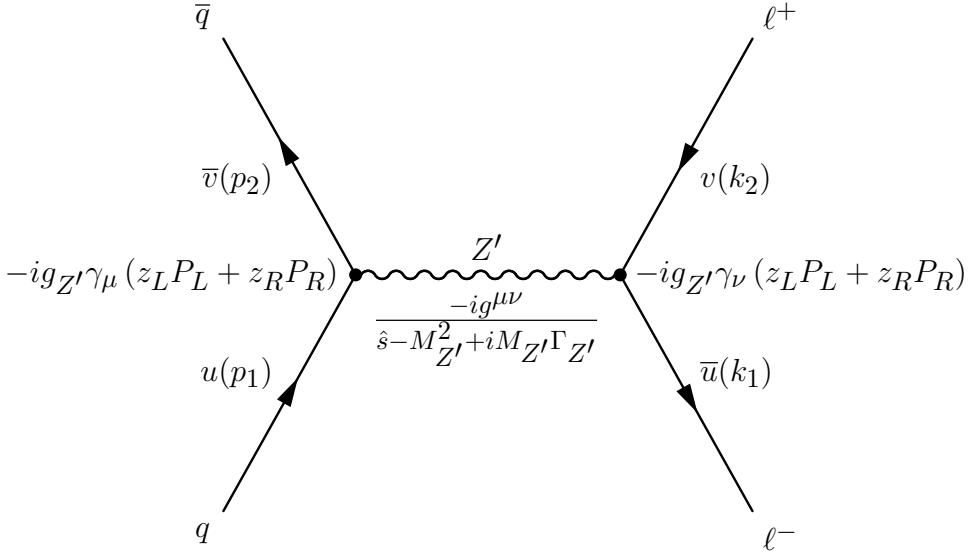


Figure 5.1: Leading Order Feynman diagram for Z' production at the LHC. The propagator and vertex factors are depicted, and the initial- and final-state fermions are represented by the Dirac spinors u and v , which label their momenta and spin.

diagram for Z' production can be seen in Figure 5.1.

The result of extending the electroweak sector with an additional neutral current interaction leads to mass and kinetic mixing between the Z boson of the Standard Model and the BSM Z' boson [109]. This can result in tiny modifications to the nominal mass M_Z and width Γ_Z of the Standard Model Z boson; however, precision measurements have put stringent constraints on such mixing, and in this analysis it can be safely ignored [109].

The matrix element for the Drell-Yan process in Eq. 4.11 of Section 4.2.1 can be extended to include a term for the new Z' boson with $\mathcal{M} = \mathcal{M}_{\gamma*} + \mathcal{M}_Z + \mathcal{M}_{Z'}$. Applying the Feynman rules to the diagram in Figure 5.1 allows the amplitude of Eq. 4.12 to be extended so that,

$$\mathcal{G}_{ij} = e^2 \frac{Q^\ell Q^q}{\hat{s}} + g_Z^2 \frac{g_i^l g_j^q}{\hat{s} - M_Z^2 + iM_Z\Gamma_Z} + g_{Z'}^2 \frac{z_i^l z_j^q}{\hat{s} - M_{Z'}^2 + iM_{Z'}\Gamma_{Z'}}, \quad (5.3)$$

where the Z' mass is $M_{Z'}$ and its width is $\Gamma_{Z'}$.

Despite the addition of this new intermediate state, the general expression for the Drell-

Yan cross section remains identical to what was derived in Eq. 4.16. By simply inserting Eq. 5.3 into the expressions for the cross section of Eq. 4.16, the effect of the new Z' state is fully included.

However, the inclusion of the Z' as a possible exchange particle gives rise to new model-dependent interference effects, which can be traced back to the cross terms present in $|\mathcal{M}|^2$. The size and nature of these effects depend on the values taken by the new gauge coupling constant $g_{Z'}$ and the $U(1)'$ fermion quantum numbers $z_{fL,R}$.

The total Z' decay width to fermion-antifermion pairs can be written,

$$\Gamma_{Z' \rightarrow f\bar{f}} = N_C g_{Z'} \frac{M_{Z'}}{24\pi} \left(z_{fL}^2 + z_{fR}^2 \right), \quad \Gamma_{Z'} = \sum_f \Gamma_{Z' \rightarrow f\bar{f}}, \quad (5.4)$$

where the color factor N_C is 1 for lepton final-states, and 3 for quark final-states, and the Z' mass is assumed to be much heavier than its decay products.

Often models with extended gauge symmetries require the existence of additional exotic fermions in order to be self consistent. If these exotic decay modes are kinematically accessible, then the size of the nominal Z' width in Eq. 5.4 is increased, which has the effect of *reducing the branching fractions* to the conventional Standard Model fermions. In general, these effects are model-dependent in nature, and are therefore neglected in the dilepton analysis.⁷

The branching fraction of the Z' into fermion-antifermion pairs is

$$Br(Z' \rightarrow f\bar{f}) = \frac{\Gamma_{f\bar{f}}}{\Gamma_{Z'}}. \quad (5.5)$$

⁷Likewise, for each of the Z' models considered in the analysis, the couplings that specify the cubic and quartic gauge interactions to other Standard Model bosons are set to zero; i.e., only Standard Model decays to fermions are considered in the expression for the Z' width. This choice is avoids the introduction of model-dependent degrees of freedom into the interpretation of the search results.

Once the Z' mass and fermion couplings are defined for the model under consideration, the relative Z' width $\Gamma_{Z'}/M_{Z'}$ can be calculated, which functions as a useful metric for comparisons between different models. For the models considered in the search, the relative width varies over a wide range, from less than 1% to upwards of 40%. Among these are The Sequential Standard Model (SSM), the E_6 model, and the Minimal Z' model, which are now presented.

5.3.2 The Sequential Standard Model

In the dilepton analysis, the Sequential Standard Model (SSM) Z' [107] is used as a useful benchmark model. Here, the Z'_{SSM} is assigned the same couplings to fermions as the Standard Model Z boson, but has a different mass and width (via Eq. 5.4). Similarly, the gauge coupling constant $g_{Z'}$ is taken to be identical to g_Z . As a result, the Z'_{SSM} has a relatively large width relative to its mass, with a value of 3.0%.

Strictly speaking, aside from the difference in mass, the Z'_{SSM} is identical to the Standard Model Z boson, and can be seen in Figure 5.2 for values of $M_{Z'} = 1$ TeV and $M_{Z'} = 3$ TeV. The fermion charges in the SSM are defined,

$$\begin{aligned} z_{fL} &= t_f^3 - q_f \sin^2 \theta_W \\ z_{fR} &= -q_f \sin^2 \theta_W, \end{aligned} \tag{5.6}$$

where t_3 is the weak isospin, q is the electric charge, and θ_W is the Weinberg angle. The values associated with each of these parameters are presented in Table 2.4 of Section 2.3.6.

The Z'_{SSM} is not a well-motivated model from a theoretical point of view due to its strict replication of Standard Model parameters. However, as an entry point, the Z'_{SSM} serves as

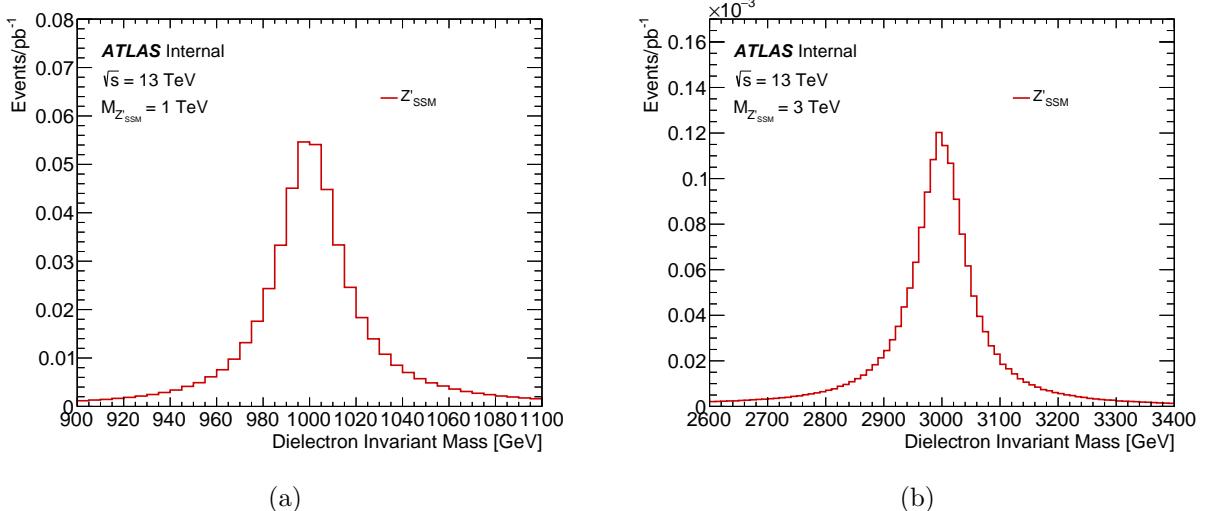


Figure 5.2: Predicted Z'_{SSM} event yields for $M_{Z'} = 1$ TeV (a) and $M_{Z'} = 3$ TeV (b). The vertical axis displays the number of predicted Z' events per given unit of integrated luminosity, as would be observed at the $\sqrt{s} = 13$ TeV LHC.

a useful benchmark for comparisons between experiments, as well as comparisons against other more theoretically motivated models considered in this analysis.

5.3.3 The E_6 Model

The E_6 model [53, 112, 110] is a GUT which predicts the symmetry group at the GUT scale is the gauge group E_6 . The successive decomposition of this group results in the gauge group of the Standard Model plus two additional abelian $U(1)'$ groups, the mixing of which results in a Z' boson that could be potentially observable at the TeV scale.

At the GUT scale, the E_6 symmetry is broken in three stages: first, $E_6 \rightarrow SO(10) \times U(1)_\psi$, followed by $SO(10) \rightarrow SU(5) \times U(1)_\chi$, and finally $SU(5)$ breaks to the Standard Model group defined by Eq. 5.1. Schematically, this can be written,

$$E_6 \rightarrow SO(10) \times U(1)_\psi \rightarrow SU(5) \times U(1)_\chi \times U(1)_\psi \rightarrow SM \times U(1)_{\theta_{E_6}}. \quad (5.7)$$

f	$q_{f_L}^\psi$	$q_{f_L}^\chi$	$q_{f_R}^\psi$	$q_{f_R}^\chi$
u	+1	-1	-1	+1
d	+1	-1	-1	-3
ν_ℓ	+1	+3	-1	+5
ℓ	+1	+3	-1	+1

Table 5.1: Standard Model fermion charges under the $U(1)_\psi$ and $U(1)_\chi$ symmetries of the E_6 GUT model. Once the mixing angle θ_{E_6} is specified, the Z' Lagrangian of Eq. 5.2 is uniquely defined, and can be used to make E_6 -related predictions.

If a linear combination of the two abelian symmetries $U(1)_\chi$ and $U(1)_\psi$ is broken at the TeV scale, a physical Z' state results, where,

$$Z'(\theta_{E_6}) = Z'_\psi \cos \theta_{E_6} + Z'_\chi \sin \theta_{E_6}. \quad (5.8)$$

Here, the degree of mixing is determined by the angle $\theta_{E_6} \in [0, \pi]$, which is a free parameter of the theory. The Z' gauge coupling constant $g_{Z'}$ is normalized in such a way that the Standard Model gauge groups are unified into $SU(5)$ at the GUT scale, where,

$$g_{Z'} = \sqrt{\frac{5}{3}} g_Z \sin \theta_W = \sqrt{\frac{5}{3}} g \tan \theta_W. \quad (5.9)$$

Further, the mixing angle θ_{E_6} determines the values of the $U(1)'$ fermion charges,

$$\begin{aligned} z_{fL} &= \frac{q_{fL}^\psi}{2\sqrt{6}} \cos \theta_{E_6} - \frac{q_{fL}^\chi}{2\sqrt{10}} \sin \theta_{E_6} \\ z_{fR} &= \frac{q_{fR}^\psi}{2\sqrt{6}} \cos \theta_{E_6} - \frac{q_{fR}^\chi}{2\sqrt{10}} \sin \theta_{E_6}, \end{aligned} \quad (5.10)$$

where $q_{fL,R}^\psi$ and $q_{fL,R}^\chi$ are the left- and right-handed fermion charges under the $U(1)_\psi$ and $U(1)_\chi$ symmetries, which are depicted in Table 5.1.

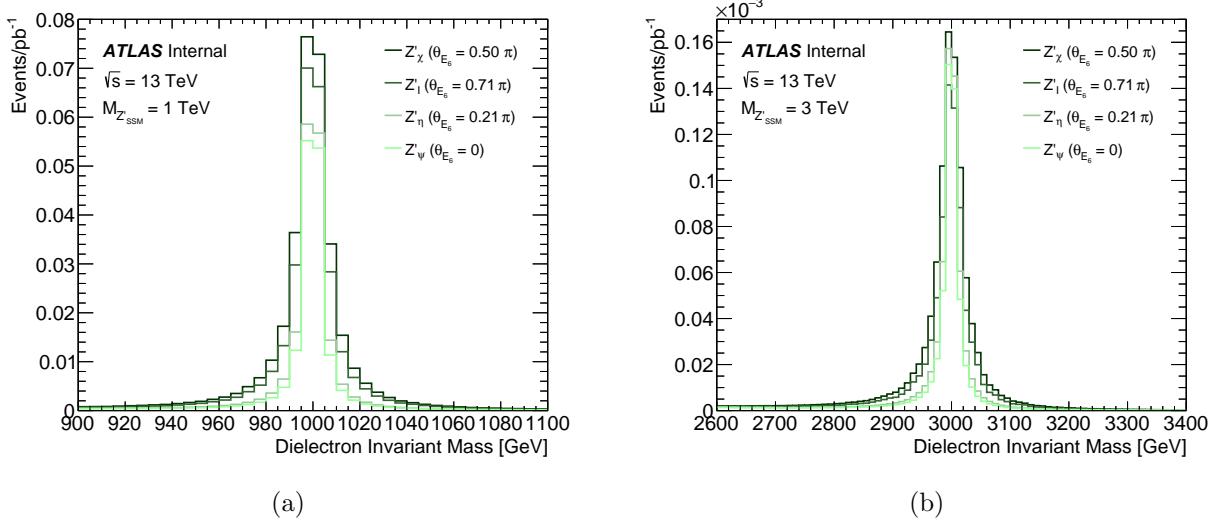


Figure 5.3: Predicted Z' event yields for various choices of θ_{E_6} in the E_6 model with $M_{Z'} = 1$ TeV (a) and $M_{Z'} = 3$ TeV (b). Starting from the Z'_ψ value of $\theta_{E_6} = 0$, E_6 models will exhibit increasing cross sections and decay widths, until the Z'_χ value of $\theta_{E_6} = \pi/2$ is reached, at which point a decrease is observed. All E_6 models fall within the range defined by Z'_ψ and Z'_χ . The vertical axis displays the number of predicted Z' events per given unit of integrated luminosity, as would be observed at the $\sqrt{s} = 13$ TeV LHC.

Several well-motivated choices for θ_{E_6} exist [110], and once its value is specified, associated Z' observables can be determined, such as the decay width via Eq. 5.4 and cross section via Eq. 4.16. The names of these well-motivated E_6 models are listed in Table 5.2.

Several Z' signals predicted in E_6 model can be seen in Figure 5.3 for different values of θ_{E_6} . In the case of the Z'_ψ , the relative width takes on a minimum value of 0.5%, while for the Z'_χ it is maximized at 1.2%. The values for all other E_6 models fall in between.

Model	Z'_ψ	Z'_N	Z'_η	Z'_I	Z'_S	Z'_χ
θ_{E_6}	0.0	0.419π	0.710π	0.210π	0.129π	0.5π

Table 5.2: E_6 models for several motivated choices of mixing angle θ_{E_6} . From left to right, these are the ψ model, the neutral model, the η model, the inert model, the secluded sector model, and the χ model.

5.3.4 The Minimal Model

The Minimal Z' model [113, 114] attempts to describe Z' phenomenology according to a minimal set of assumptions. Unlike the various GUT scenarios, in the Minimal model, no exotic fermions are required to generate an anomaly free theory. Here, only three right-handed neutrinos (one for each fermion generation) are required. The Standard Model gauge group of Eq. 5.1 is extended by a $U(1)_{B-L}$ symmetry, so that it becomes $SU(3)_C \times SU(2)_L \times U(1)_Y \times U(1)_{B-L}$.

The new Z'_{Min} boson results from the mixing between the $U(1)_Y$ symmetry of the Standard Model and the new $U(1)_{B-L}$ symmetry. As a result, the generator of the new $U(1)'$ symmetry is a linear combination of Standard Model weak hypercharge Y and baryon-minus-lepton number $B - L$.

The phenomenology of Z' production and decay in the Minimal model is characterized by two effective coupling constants: g_{B-L} , associated with the $B - L$ current, and g_Y , with the weak hypercharge Y where,

$$\begin{aligned} g_Y &= \gamma' g_Z \sin \theta_{\text{Min}} \\ g_{B-L} &= \gamma' g_Z \cos \theta_{\text{Min}}. \end{aligned} \tag{5.11}$$

Here, the mixing angle θ_{Min} controls the level of mixing between the $U(1)_Y$ and $U(1)_{B-L}$ gauge groups, and the relative coupling γ' measures the strength of the Z' boson gauge coupling relative to that of the Z boson of the Standard Model. Each of these are free parameters in the Minimal model.

Just like in the context of the E_6 model, specific choices for γ' and θ_{Min} characterize well-known Z' states, several of which are listed in Table 5.3. These include the Z'_χ , which belongs

Model	Z'_{B-L}	Z'_χ	Z'_{3R}
γ'	$\sqrt{\frac{5}{8}} \sin \theta_W$	$\sqrt{\frac{41}{24}} \sin \theta_W$	$\sqrt{\frac{25}{12}} \sin \theta_W$
$\sin \theta_{\text{Min}}$	0	$-\sqrt{\frac{16}{41}}$	$-\sqrt{\frac{4}{5}}$
$\cos \theta_{\text{Min}}$	1	$\sqrt{\frac{25}{41}}$	$\sqrt{\frac{1}{5}}$

Table 5.3: Several Z'_{Min} models for well-motivated choices of θ_{Min} and γ' . From left to right, these are the pure Z'_{B-L} model, the Z'_χ of the E_6 model, and the Z'_{3R} of the left-right symmetric model.

f	Y_{fL}	Y_{fR}	$(B - L)_{fL}$	$(B - L)_{fR}$
u	$+\frac{1}{6}$	$+\frac{2}{3}$	$+\frac{1}{3}$	$+\frac{1}{3}$
d	$+\frac{1}{6}$	$-\frac{1}{3}$	$+\frac{1}{3}$	$+\frac{1}{3}$
ν_ℓ	$-\frac{1}{2}$	0	-1	-1
ℓ	$-\frac{1}{2}$	-1	-1	-1

Table 5.4: Quantum numbers associated with Standard Model fermions according to Eq. 5.12. Once the mixing angle θ_{Min} and relative coupling γ' are specified, the Z' Lagrangian of Eq. 5.2 is uniquely defined and can be used to make Minimal model predictions.

to the E_6 model of the previous section, the Z'_{3R} of the left-right symmetric model [115, 116], and the Z'_{B-L} of the pure $B - L$ model [117].⁸

The left-handed and right-handed fermions charges under the new $U(1)'$ symmetry are

$$\begin{aligned} z_{fL} &= g_Y Y_{fL} + g_{B-L} (B - L)_{fL} \\ z_{fR} &= g_Y Y_{fR} + g_{B-L} (B - L)_{fR}, \end{aligned} \tag{5.12}$$

where $Y_f = Q_f - t_f^3$ is the standard weak hypercharge, and $(B - L)_f$ is the fermion Baryon Number minus its Lepton Number, whose values are listed explicitly in Table 5.4. Several Z'_{Min} signal scenarios are depicted in Figure 5.4 as the values of θ_{Min} and γ' are varied. The Z' width is sensitive to these choices, and taking the Z'_{3R} and Z'_{B-L} models as examples, the

⁸A common point in Z' parameter space is shared between both the Minimal model and E_6 model for the Z'_χ boson.

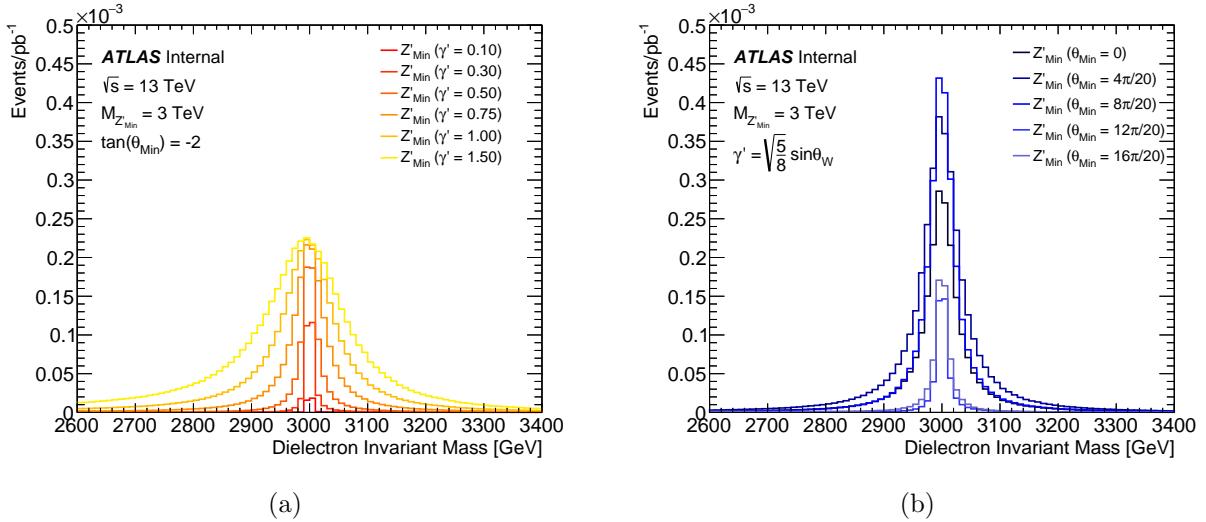


Figure 5.4: Predicted Z'_{Min} event yields for $M_{Z'} = 3 \text{ TeV}$ according to Eq. 5.12. In (a), θ_{Min} is held constant for several distinct values of γ' . Similarly, in (b), γ' is held constant while θ_{Min} is varied. The vertical axis displays the number of predicted Z'_{Min} events per given unit of integrated luminosity, as would be observed at the $\sqrt{s} = 13 \text{ TeV}$ LHC.

width varies from $< 1\%$ to up to 12.8% and 39.5% respectively for the γ' ranges considered in the dilepton search.

5.4 Previous Experimental Limits

Both direct and indirect searches for Z' bosons have been conducted at several previous collider experiments. Early results were obtained from the LEP experiment at CERN, followed by the DØ and CDF experiments at the Tevatron, and more recently, the ATLAS and CMS experiments at the LHC. Examples of previous limits on the Z' mass are shown in Table 5.5.

Some of the earliest limits on Z' production come from indirect measurements taken at the LEP experiment. LEP searched for Z' bosons in several fermionic final-states via $e^+e^- \rightarrow Z' \rightarrow f\bar{f}$. Here, the Z' state would interfere with the intermediate state γ^*/Z of Drell-Yan, modifying the total cross section and angular distributions of the final-state fermions. The

Collaboration	\sqrt{s} [TeV]	\mathcal{L} [fb^{-1}]	Channel	Lower Limit on $M_{Z'} [\text{TeV}]$			Reference
				Z'_{SSM}	Z'_{ψ}	Z'_{χ}	
DØ	1.96	5	ee	1.023	0.891	0.903	[120]
CDF	1.96	5	$\mu\mu$	1.071	0.917	0.930	[121]
CMS	8	20	$\ell\ell$	2.90	2.57	-	[122]
ATLAS	8	20	$\ell\ell$	2.90	2.51	2.62	[123]
CMS	13	3	$\ell\ell$	3.18	2.70	-	[124]
ATLAS	13	3	$\ell\ell$	3.36	2.74	3.05	[111]
CMS	13	36	$\ell\ell$	4.5	3.9	-	[125]
ATLAS*	13	36	$\ell\ell$	4.5	3.8	4.1	Sec. 11.4.2

Table 5.5: Observed limits at 95% C.L. on the mass of a Z' boson from previous experimental searches at hadron colliders. The CMS analyses listed do not provide limits on the Z'_{χ} , which would otherwise be slightly higher than what was obtained for the Z'_{ψ} . The integrated luminosity for each analysis is rounded to nearest whole number. The $\ell\ell$ channel refers searches that combine individual electron and muon channels. The analysis presented here is denoted with “*.”

strongest indirect limits come from a combined analysis using the data taken by the four LEP experiments, which combines differential cross section and angular measurements from multiple final-states (See. Sec. 3.5 of Ref. [118]). The result is a lower mass limit of 1.787 TeV on the Z'_{SSM} , and 0.481 TeV and 0.673 TeV on the Z'_{ψ} and Z'_{χ} , respectively [119].

Following LEP, the DØ and CDF collaborations at the Tevatron set direct limits on Z' production using proton-antiproton collision data at $\sqrt{s} = 1.96$ TeV. These were the strongest direct limits set before the operation of the LHC. The DØ limits [120] were set in the electron channel with 5.3 fb^{-1} of data, while the CDF limits [121] were set with 4.6 fb^{-1} in the muon channel. For DØ, this resulted in lower mass limits of 1.023, 0.891, 0.903 TeV for the Z'_{SSM} , Z'_{ψ} , and Z'_{χ} respectively. Likewise, for CDF the results were 1.071, 0.917, and 0.930 TeV.

The LHC currently holds record as the highest energy particle collider, and as such, its ability to set exclusion limits on Z' production is vastly improved compared to what was

achievable at LEP and the Tevatron. The most stringent direct limits come from the ATLAS and CMS experiments using proton-proton collision data. Searches have been conducted at $\sqrt{s} = 7, 8$ and 13 TeV, with varying amounts of integrated luminosity. Some of the highest direct limits set to date come from the dilepton analysis presented in this dissertation (See Sec. 11.4.2). Results obtained by ATLAS and CMS at $\sqrt{s} = 8$ and 13 TeV are displayed in Table 5.5.

Part II

The Experiment

Chapter 6

The Large Hadron Collider

This chapter introduces the Large Hadron Collider (LHC) and discusses its operation and performance during the data-taking periods used for the dilepton analysis. The LHC is a synchrotron used to accelerate, store, and collide beams of protons and heavy ions at high energies. The analysis of such collisions provides evidence for the laws governing subatomic particle interactions, and allows for theories Beyond the Standard Model (BSM) to be tested by searching for potentially unobserved particles such as the Z' boson.

6.1 The LHC

The LHC is currently the world's largest and most powerful particle accelerator (See Fig. 6.1). It is a hadron collider that accelerates beams of protons (and heavy ions) up to an energy of 6.5 TeV, resulting in a total **center-of-mass** (CM) energy of $\sqrt{s} = 13$ TeV. Operated by The European Organization for Nuclear Research (CERN), the LHC is housed in an excavated tunnel under the Franco-Swiss border outside of Geneva, Switzerland [126].

The tunnel is 26.7 km in circumference, with a depth of 45 m (toward lake Geneva) and 170 m (at the foot of the Jura mountains). It was originally excavated in the 1980s for a different machine, the Large Electron Positron Collider (LEP) [126]. LEP was decommissioned in 2000, so that the existing infrastructure could be reused, with adequate modification, for the construction of the LHC and its new experiments [126].

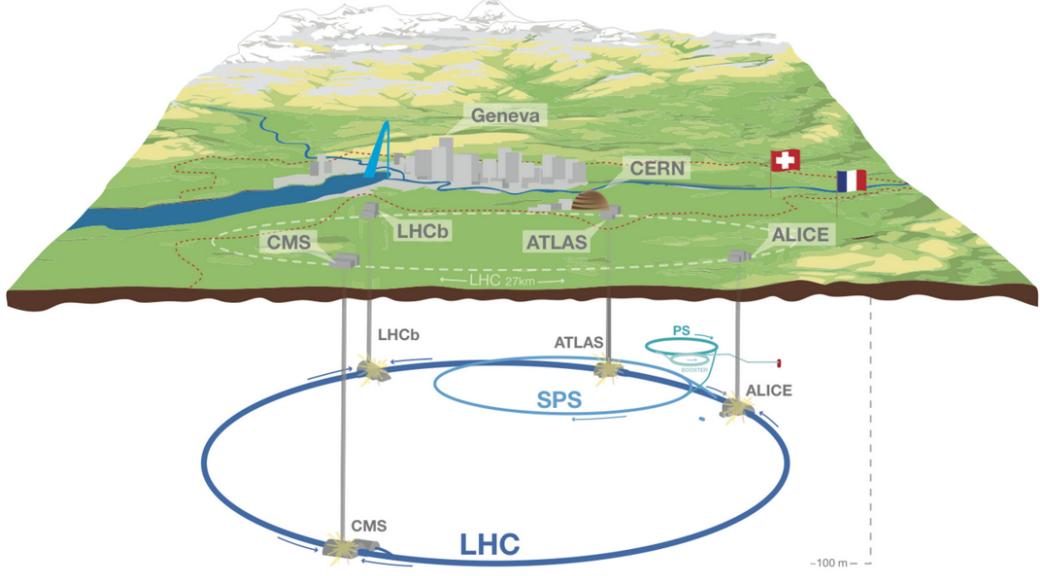


Figure 6.1: Illustration of the 27 km long LHC, its four main experiments (ALICE, ATLAS, CMS, and LHCb), the pre-accelerators used in staging proton acceleration to reach a CM energy of $\sqrt{s} = 13$ TeV, and the region surrounding Geneva, Switzerland [13].

The LHC was designed to search physics Beyond the Standard Model (as outlined in Sec. 5.1), and test the predictions of the Standard Model in an energy range inaccessible to previous experiments. Its first research run was conducted from 2010 to 2013 at $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV. Run-1 saw the discovery of the Higgs Boson, as well as a host of searches and measurements consistent with Standard Model predictions. From 2013 to 2015 the machine was shut down for repairs and upgrades, in preparation the second run when the CM energy was increased to $\sqrt{s} = 13$ TeV.

6.2 CERN Accelerator Complex

In order to reach the maximum beam energy of 6.5 TeV in the main ring of the LHC, protons are accelerated through successive stages by smaller accelerators at CERN’s accelerator complex [127] (See Fig. 6.2).

Proton acceleration starts with the Linac2 linear accelerator. The proton source is a single bottle of hydrogen gas. The hydrogen is passed through an electric field, ionizing it, and the resulting protons are accelerated to 50 MeV. The second stage occurs when the beam is injected into the Proton Synchrotron Booster (PSB), a circular accelerator made of four superimposed synchrotron rings. The PSB accelerates protons up to 1.4 GeV and injects them into the 628 m ring of the Proton Synchrotron (PS). In the third stage, the PS pushes the beam up to 25 GeV and transfers it to the nearly 7 km ring of the Super Proton Synchrotron (SPS). Finally, in stage five, the SPS further accelerates the beam to 450 GeV and injects it into the LHC [127].

6.3 LHC Operation

Once injected, the LHC is designed to accelerate and store two counter-rotating beams of high-energy particles. The beams travel in opposite directions, and are contained in separate beam pipes held at ultra-high vacuum (10^{-13} atm) to prevent collisions with extraneous gas molecules [128].

Protons are accelerated by 16 Radio-Frequency (RF) cavities (8 per beam) situated along the straight sections of the LHC main ring [129]. The RF cavities are designed to provide 400 MHz oscillating electric fields, so that a charged particle, as it passes through, will feel an overall force in its direction of motion. For each pass, the proton receives an extra 16 MeV

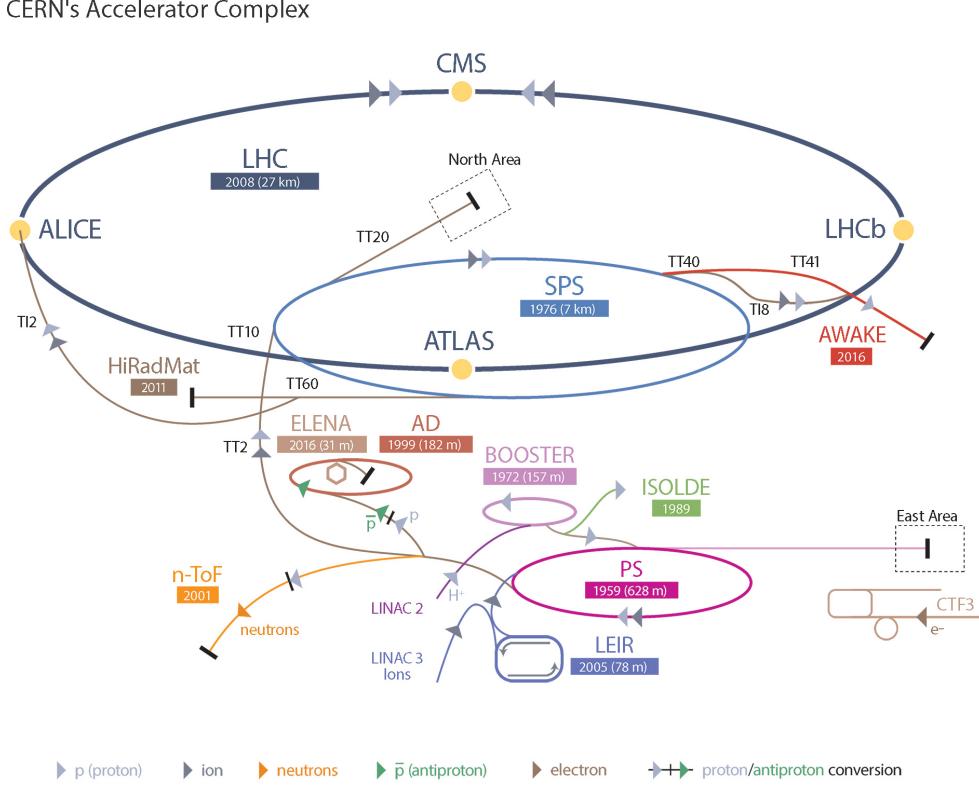


Figure 6.2: Illustration of the CERN accelerator complex. Proton acceleration is achieved through the following stages: Linac 2, PSB, SP, SPS, and LHC. Accelerators used in other CERN experiments are shown as well [14].

of energy, and as a result, it takes about 20 minutes to go from 0.45 TeV at injection to 6.5 TeV for collisions [130]. After this, the main role of the RF cavities is to keep the protons tightly bunched together to ensure the highest possible luminosity.

The proton beams are guided through the main ring of the LHC by powerful magnetic fields. The LHC utilizes 1232 superconducting dipole magnets, measuring 15 m in length and weighing 35 tons (See Fig. 6.3). The electromagnet's coils are threaded with a niobium-titanium (NbTi) alloy and cooled to 1.9 K. With an applied current of 11,080 A, the resulting field strength is 8.3 T. The beams spread out as they travel, so it is also necessary to align and focus them. This is achieved with 392 quadrupole magnets, which alternate between

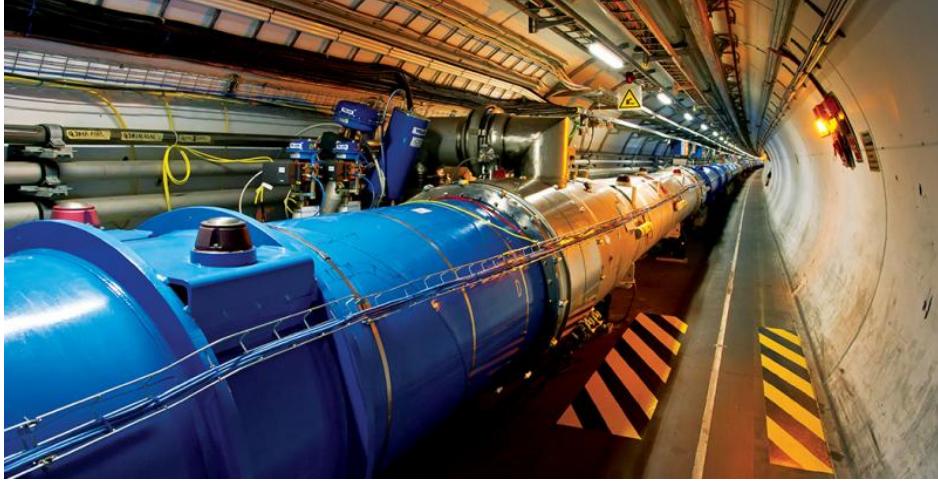


Figure 6.3: Superconducting dipole magnets as seen in the LHC tunnel [15].

squeezing the beams horizontally and vertically as they circulate. Additional higher-order magnets help as well by correcting for small imperfections in the extreme regions of the dipole fields [131].

6.3.1 Pile-up

LHC beams circulate with a frequency of 11.245 kHz,¹ and can do so for many hours under normal operating conditions. As a result of the RF acceleration scheme, the beams fall into discrete packets called “bunches.”² At the LHC design value, each proton beam contains 2808 bunches, with 1.2×10^{11} protons per bunch at the start of data-taking [128]. The bunch-spacing of the LHC is 25 ns, which corresponds to a bunch crossing rate of 40 MHz, the number of times per second bunches cross at each interaction point (IP). The bunch crossing rate is related to the frequency of the RF cavities through $f_{RF} = h f_{rev}$, where h is an integer, known as the harmonic number. This ensures the proton always receives a

¹The proton revolution frequency is derived from $f_{rev} = \beta c / 2\pi R_{LHC}$.

²A “perfectly timed” proton will have a revolution frequency f_{rev} that is a multiple of the 400 MHz RF oscillation frequency, and as a result will experience zero accelerating voltage; however, any proton that exhibits a small variation in f_{rev} will be slightly accelerated or decelerated, and will therefore oscillate around the perfectly timed proton. The aggregate of such particles forms a bunch [132].

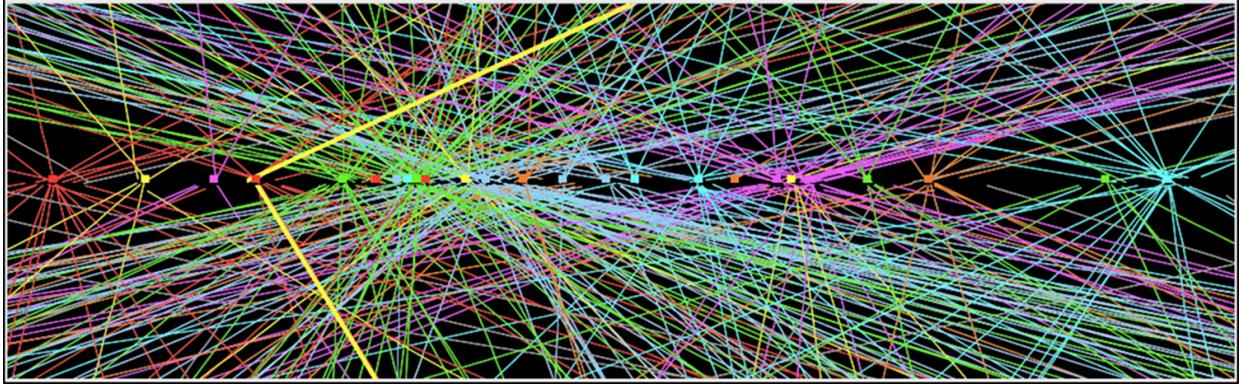


Figure 6.4: Depiction of a $Z \rightarrow \mu^+ \mu^-$ event as recorded by the ATLAS detector during Run-1 of the LHC, along with the 25 simultaneous proton-proton interactions during this particular bunch crossing. The two muons originate from a common vertex, due to the decay of the Z boson, and are highlighted in yellow [16].

positive accelerating voltage from each RF cavity.

During a single bunch crossing, there can be multiple collisions between the roughly 2×10^{11} particles (as in Fig. 6.4). These additional interactions are known as **pile-up**. The collisions in a single bunch crossing, in addition to the one of interest, are known as in-time pile-up. This serves as an additional background to identifying and reconstructing various physics objects in the event. Further, out-of-time pile-up results from collisions that occur in bunch crossing just before or after the collision of interest. This can affect the detector response to the in-time collision, since various detector electronics integrate for more time than the 25 ns between individual bunch crossings.

The amount of pile-up present in events can be quantified by the average number of interactions per bunch crossing. In Run-1 for example, this value was $\langle \mu \rangle = 20.7$ at $\sqrt{s} = 8$ TeV [17]. The reconstruction of every physics object in an event is affected by pile-up in some way, so it must be accounted for when performing an LHC physics analysis (as in Sec. 9.4.4).

6.3.2 Luminosity

Aside from its CM energy, the most important design feature of an accelerator is its **instantaneous luminosity** (or just the luminosity), the number of protons passing through a given unit area per unit time. In fact, the LHC was chosen to be a pp collider (as opposed to $p\bar{p}$ like the Tevatron) as a way to maximize its luminosity.

Many physics processes of interest at the LHC are only produced rarely, so a large number of collisions is required. The instantaneous luminosity links the **event rate** of a physics process to its production cross section σ where,

$$\frac{dN}{dt} = \sigma \mathcal{L}, \quad (6.1)$$

so a higher instantaneous luminosity increases the rate at which rare events are produced, and hence improves the potential for the discovery of new physics. Likewise, the **integrated luminosity**, $\mathcal{L}_{\text{int}} = \int \mathcal{L} dt$, can be used to determine how many collisions events occurred for a process of interest over a particular data-taking period or run of the LHC.³

The instantaneous luminosity can be expressed in terms of the following accelerator parameters [133],

$$\mathcal{L} = \frac{N^2 n_b f \gamma}{4\pi \epsilon_n \beta^*} F, \quad (6.2)$$

where N is the number of particles per bunch, n_b is the number of circulating bunches, f_{rev} is the proton revolution frequency, and γ is the Lorentz factor. The cross-sectional area of each beam is related to ϵ_n , the normalized emittance, and β^* the value of the beta function

³At the LHC, cross sections are measured in barns ($1 \text{ barn} = 10^{-24} \text{ cm}^2$). The integrated luminosity is similarly expressed in terms of inverse barns, or more commonly inverse femtobarns, fb^{-1} .

Year	E_{beam} [TeV]	N_p	n_b	ϵ_n [μm]	β^* [m]	Bunch Spacing [ns]	Peak lumi. [cm $^{-2}$ s $^{-1}$]
2015	6.5	1.10×10^{11}	2244	3.5	0.8	50/25	5.1×10^{33}
2016	6.5	1.10×10^{11}	2076	3.4	0.4	25	1.4×10^{34}
Design	7.0	1.15×10^{11}	2808	3.75	0.55	25	1.0×10^{34}

Table 6.1: Main LHC beam parameters under typical running conditions. Shown are values for the 2015 run and the 2016 run through June, along with the LHC design values [2].

at the interaction point.⁴ The parameter F is a geometrical factor, expressing the fact that the bunches do not necessarily collide head-on, resulting in a small luminosity reduction, with a design value of around 0.8. Table 6.1 lists the main LHC beam parameters during the 2015 and 2016 data-taking runs, as well as the design values for comparison.

6.4 LHC Performance in Run-II

The LHC started its second run in 2015, after its first long shutdown (LS1). During this two year shutdown, the magnet systems were overhauled to prepare for the increase in the CM energy of the collider from $\sqrt{s} = 8$ TeV in Run-1 to $\sqrt{s} = 13$ TeV in Run-2. After several months of commissioning, the first stable beams were delivered to the experiments on June 3, 2015, and the intensity was ramped up slowly throughout the summer [134].

The bunch spacing was initially set to 50 ns, then was increased to 25 ns, making it possible to fill up to 2244 bunches into the machine. By the end of the proton physics running period, the peak instantaneous luminosity was $\mathcal{L} \approx 5 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$, about 50% its design value. The integrated luminosity delivered over the course of the run was just above 4 fb^{-1} , falling below the initial projections, due to various machine related difficulties interrupting the run. The bulk of the data arrived toward the end, reaching production rates

⁴The beta function is $\pi\sigma^2/\epsilon$, the width of the beam squared divided by the emittance. If beta is small, the beam is narrow, like at the interaction point.

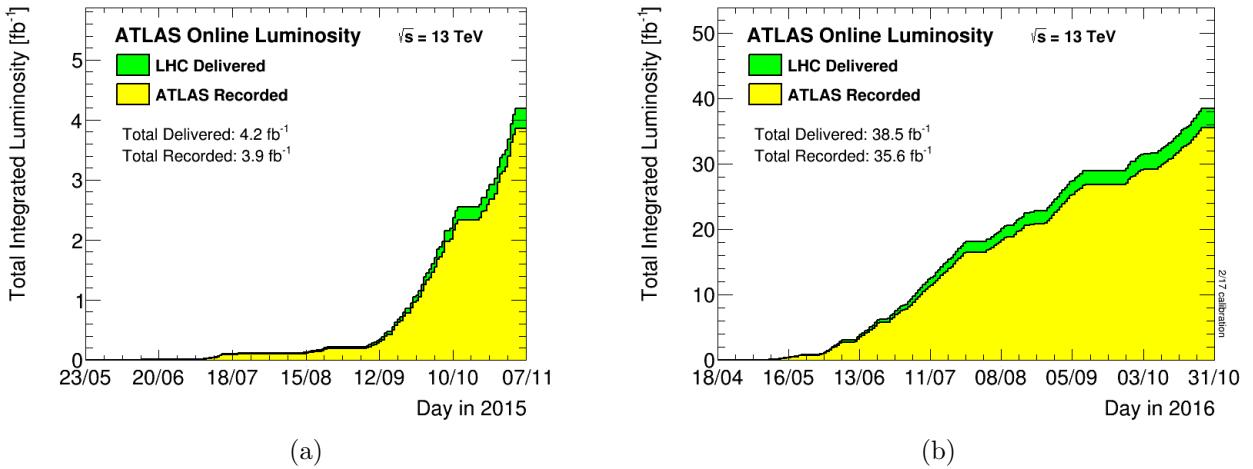


Figure 6.5: Cumulative integrated luminosity obtained per week as delivered by the LHC and recorded by the ATLAS detector during the 2015 (a) and 2016 (b) data-taking runs [17, 18].

of 200-250 pb⁻¹/day [134].

On the other hand, the 2016 run was extremely successful. The machine reached its design luminosity in June, for the first time, and then regularly operated at 30% above that [135]. The target integrated luminosity for the year was 25 fb^{-1} , and almost 40 fb^{-1} was recorded, with the machine spending a record 60% of its time delivering stable beams to the experiments [136].

During the 2015 + 2016 data-taking periods, the LHC delivered a total of 42.7 fb^{-1} , of which 36.1 fb^{-1} was recorded by the ATLAS detector, and made usable for analysis. The average number of interactions per bunch crossing for this period was $\langle\mu\rangle = 23.7$. Figure 6.5 shows the integrated luminosity for the 2015 and 2016 runs as a function of time.

6.5 Experiments at the LHC

The four main experiments at the LHC are ATLAS (A Toroidal LHC Apparatus), CMS (the Compact Muon Solenoid), LHCb (the Large Hadron Collider beauty), and ALICE (A

Large Ion Collider Experiment). ATLAS and CMS are general purpose detectors, designed to investigate the largest possible range of physics phenomena. LHCb, on the other hand, is optimized to study the production and decay of B hadrons as a way to investigate CP violation⁵. ALICE specializes in the study of heavy-ion collisions and the physics of the quark-gluon plasma. Each detector is installed in a large underground cavern built around the four collision points of the LHC beam (See Fig. 6.1).

In addition, the main experiments are complemented by three specialized smaller ones: TOTEM (the TOTal Elastic and diffractive cross section Measurement), LHCf (the Large Hadron Collider forward), and MoEDAL (Monopole and Exotics Detector at the LHC). TOTEM is installed close to the CMS interaction point, LHCf near ATLAS, and MoEDAL near LHCb. Each of the smaller experiments is designed with enhanced sensitivity to study specific aspects of proton collisions [128].

⁵B hadrons contain a bottom quark

Chapter 7

The ATLAS Experiment

This chapter introduces the ATLAS detector, the experimental apparatus used to record LHC proton-proton collisions and search for Z' boson in the dilepton analysis.

7.1 Detector Overview

The ATLAS (A Toroidal LHC ApparatuS) detector is one of two **general purpose** high-energy physics detectors at the LHC. It is the largest particle detector ever constructed, measuring 46 m long, 26 m high, 26 m wide, and weighing over 7000 tons.

As a general purpose detector, it is designed to cover the broadest possible range of physics phenomena, rather than focusing on a particular signature or type of process. Indeed, if new physics exists, ATLAS will have a good chance of detecting it and performing dedicated measurements to understand its properties. At the same time, the ATLAS detector is designed to be able to perform a host of Standard Model measurements over a wide range of final-states and complex decay topologies.

The experiment is designed to operate at a luminosity of $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, one of the highest proton-proton collision rates yet achieved in a collider experiment, while simultaneously maintaining sensitivity to a wide variety of physics signatures [19]. This puts definite requirements on the design of the detector and helps to establish its physics goals. Specifically, the ATLAS experiment needs to be able to detect, measure, and reconstruct

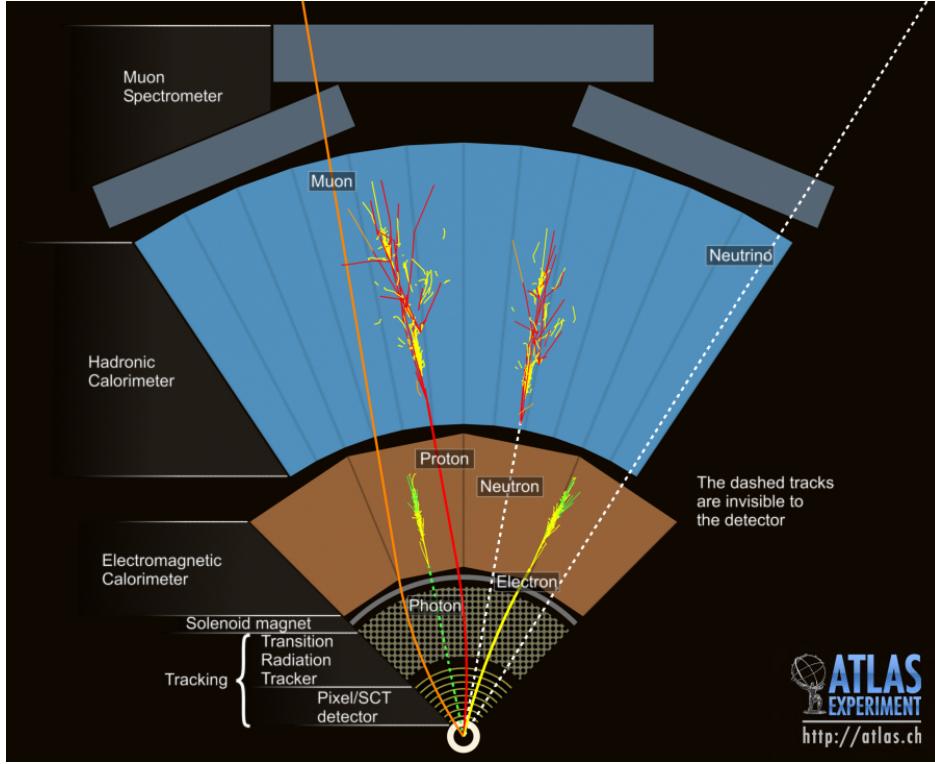


Figure 7.1: Individual particles are identified and measured in the ATLAS detector according to their characteristic signatures left in each detector subsystem. Solid lines represent trajectories of charged particles, while dashed lines indicate neutral particles. Showers are depicted for particles that interact with the dense detector material [19].

final-states defined by electrons, photons, jets, B hadrons, muons, taus, and/or neutrinos, all over a wide kinematic range.

The ATLAS detector is comprised of a series of dedicated subsystems to take advantage of the distinct signatures left by each particle as it interacts with the detector material (See Fig. 7.1). Charged particles are identified by ionization tracks, as their trajectory bends as they pass through a magnetic field. Tracks are formed from a series of ionizing hits in the inner detector, which can be used to estimate charge and momentum. Neutral particles leave no tracks; rather, they interact with dense detector material, creating a shower of secondary particles. The resulting shower profile can be used to estimate the incident particle energy. Muons are highly penetrating, and their path is determined with additional tracking

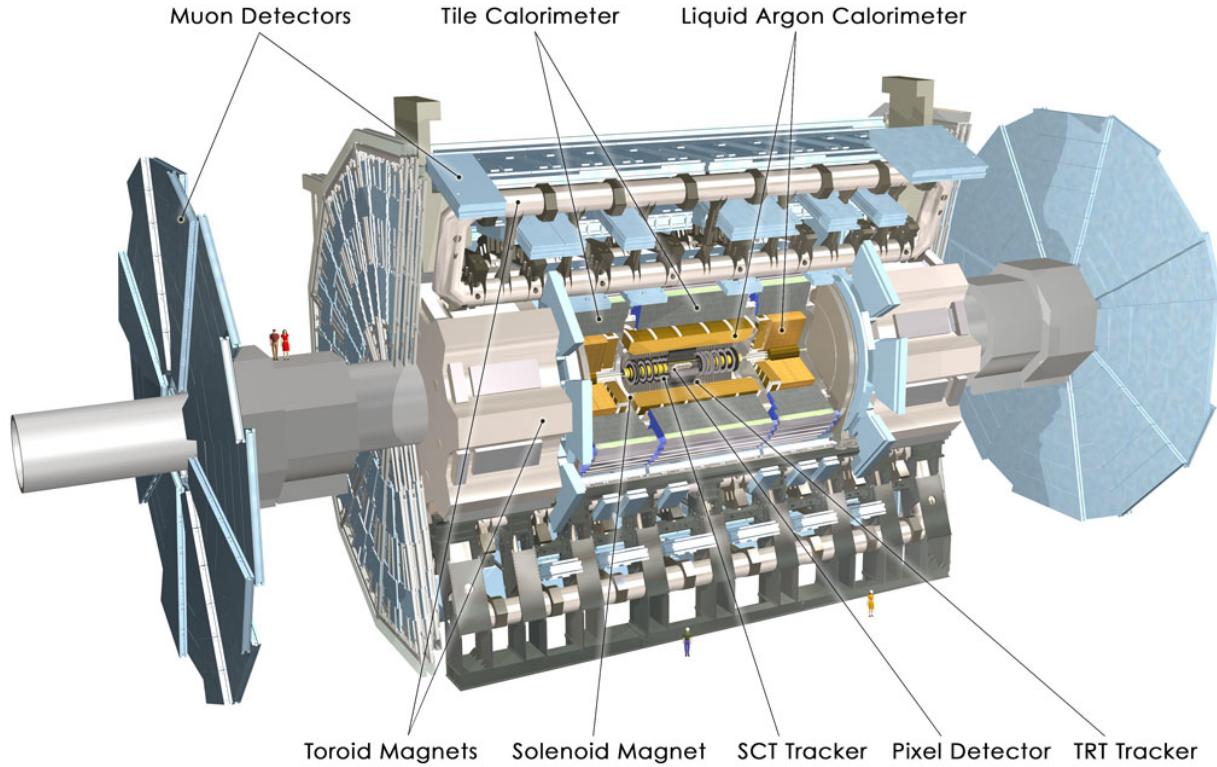


Figure 7.2: Cut-away view of the ATLAS detector with individual subsystems labeled [20].

information provided in the outer detector. Neutrinos are undetectable, but their presence can be inferred by calculating the amount of missing energy in the event. Unstable particles like B hadrons have long lifetimes and decay in flight producing a **secondary decay vertex**. A secondary (or displaced) vertex like this can be used to identify and reconstruct unstable particles in the event.

The ATLAS detector layout can be seen in Figure 7.2. For particle detection, three main subsystems are used: the tracking subsystem, the calorimetry subsystem, and the muon subsystem. The **tracking system** resides in the inner detector (ID), which is immersed in a 2 T solenoidal magnetic field, and is used to measure the charge and momentum of charged particles. The ID itself is comprised of three subsystems: the Pixel Detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT), which together are

used to achieve high momentum resolution and reconstruction efficiency. The **calorimetry system** surrounds the ID, and is comprised of electromagnetic and hadronic components, which measures the energy of electrons, photons, and hadrons, and contributes to their identification. The **muon system** sits outside the hadronic calorimeter and provides stand-alone precision muon-momentum measurements that supplement the tracking measurements of the ID. Lastly, the **trigger and data-acquisition (TDAQ) system** is responsible for triggering on interesting hard interaction events, and saving them to permanent storage for later offline analysis.

Each subsystem consists of a barrel detector, centered around the interaction point (IP), and two end-caps, for detecting particles in forward regions of the detector. This maximizes the detector **acceptance**, and is the origin of its cylindrical shape.

7.2 Coordinate System and Event Kinematics

The ATLAS detector uses a right-handed coordinate system. The origin is set at the nominal IP at the center of the detector. The z -axis coincides with the beam line, the x -axis points to the center of the LHC ring, and the y -axis points upward to complete the right-handed system.

Spherical coordinates are used as well, where R is the radius in the $x - y$ (transverse) plane of the detector, and ϕ is the azimuthal angle, measured clockwise around the z -axis from the positive x -axis. Instead of the polar angle θ , measured from the positive z -axis, the **pseudorapidity** η is used, where

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \quad (7.1)$$

The advantage is that η is a Lorentz invariant quantity for boosts along the z -axis. In the massless limit, the rapidity y of a particle defined by Eq. 3.4 reduces to the pseudorapidity of Eq. 7.1, which is a good approximation for the final-state electrons of dielectron events at the LHC.

It is also useful to define several kinematic variables in the transverse plane of the detector due to conservation of momentum in the collision. The **transverse momentum** p_T and **transverse energy** E_T of a particle are given by

$$\begin{aligned} p_T &= \sqrt{p_x^2 + p_y^2}, \\ E_T &= \sqrt{p_T^2 + m^2}. \end{aligned} \tag{7.2}$$

Again, in the massless limit, the transverse energy E_T reduces to the transverse momentum p_T . For particles like neutrinos, it is useful to define the missing transverse momentum $\mathbf{p}_T^{\text{miss}}$, the negative vector sum of all visible reconstructed transverse momenta in the event,

$$\mathbf{p}_T^{\text{miss}} = - \sum_i \mathbf{p}_{T,i}. \tag{7.3}$$

A large value of $\mathbf{p}_T^{\text{miss}}$ indicates that one or more particles in the event escaped detection, the primary candidate being a neutrino. The **missing transverse energy** (MET) is the magnitude, defined by $E_T^{\text{miss}} = |\mathbf{p}_T^{\text{miss}}|$.

Lastly, the distance ΔR between two objects in the $\eta - \phi$ space of detector can be calculated from

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{7.4}$$

7.3 Tracking System

The ATLAS Inner Detector (ID) [23] covers the region $|\eta| \leq 2.5$, also known as the **precision region** of the ATLAS detector, and is designed to measure charged particle trajectories and locate interaction vertices (See Fig. 7.3). A precise measurement of the trajectory can be used to extract the charge and momentum of a charged particle. The ID is located within a 2 T axial magnetic field provided by the external superconducting central solenoid magnet [23].¹ The curvature of the particle's trajectory in the magnetic field is due to the Lorentz Force and is proportional to its momentum. The achieved momentum resolution is $\sigma/p_T = 0.05\%p_T \oplus 1\%$ [23].

In addition to tracking, the ID helps with vertex reconstruction. Pattern recognition software is used to identify tracks that originate from the primary vertex, as well as those from secondary vertices due to in-flight decays of B hadrons, b -quark jets, or τ leptons. Measuring 7.0 m long and 2.3 m wide, the ID is the detector closest to the beam line.

The ID consists of three subsystems: the Silicon Pixel Detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). The Pixel Detector and SCT contribute to the determination of primary and secondary vertices and precision tracking, while the TRT supplements these momenta measurements and aids in electron identification.

A charged particle entering the ID passes through 3 layers of silicon pixels in the Pixel Detector, 8 layers of silicon strips in SCT, and finally through the straw-tubes of the TRT, which provide 36 hits on average (See Fig. 7.4). Pattern recognition software is then used to fit a high precision track through the series of hits provided by each subsystem. In addition, a fourth subsystem, the Insertable B-Layer (IBL), was added in Run-2 to maintain a high

¹The solenoid uses a single winding of aluminum-stabilized niobium-titanium (NbTi) coil, operating at a nominal current of 7.73 kA, and temperature of 4.6 K. The casing is designed to be as thin as possible to minimize particle scatterings before they reach the calorimetry system [23].

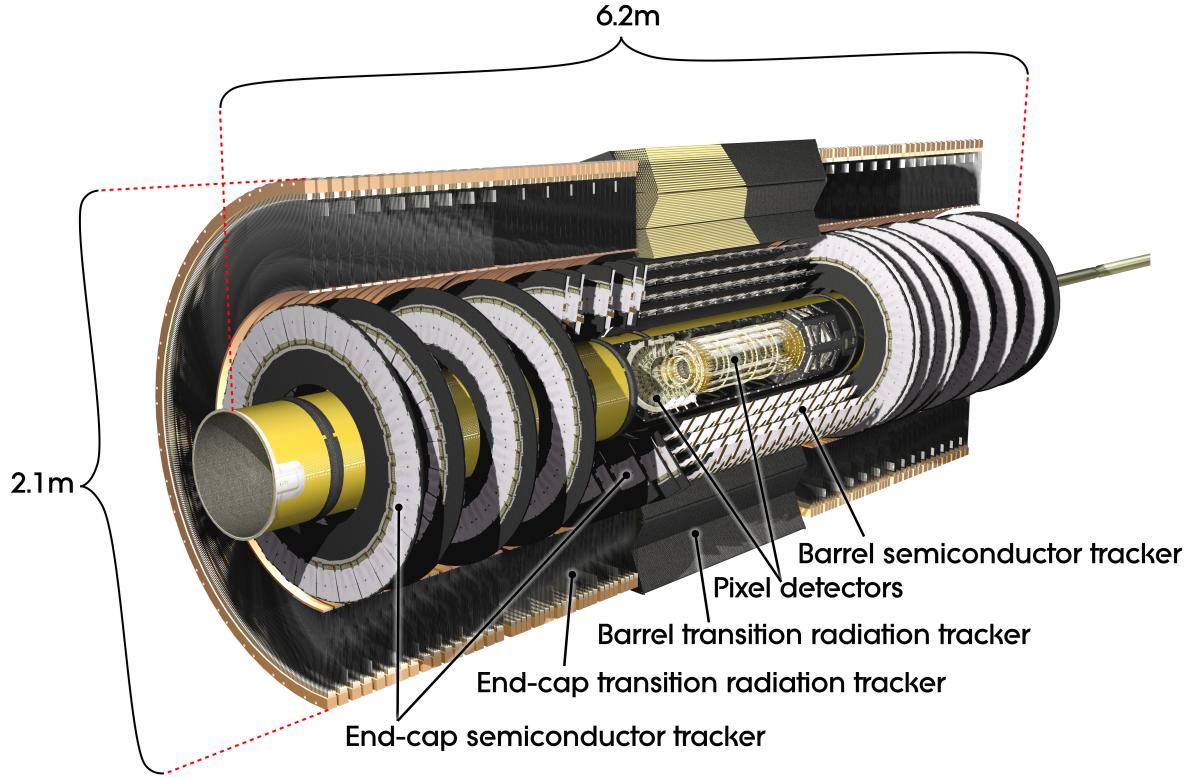


Figure 7.3: Cut-away view of the ATLAS Inner Detector. Subsystem components are shown in both the barrel and end-cap regions [21].

level of tracking performance in the high luminosity environment of Run-2, where increased pile-up was expected.

7.3.1 Pixel Detector

The high granularity silicon Pixel Detector [23] is the innermost part of the ID system and closest to the IP, covering $|\eta| < 2.5$. It is designed for the measurement of charged particle tracks to high resolution, and identify and reconstruct primary and secondary decay vertices. When a high energy charged particle passes through the silicon layer of each module, it ionizes the silicon molecules. The resulting current produced from the movement of electron-hole pairs is detected by readout electronics.

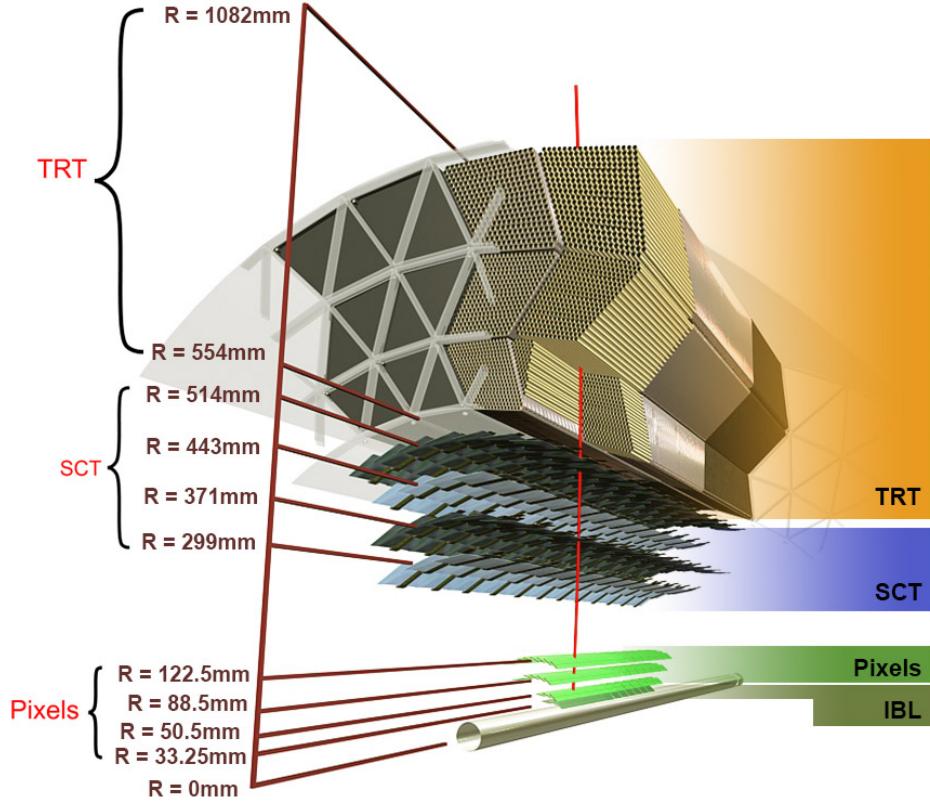


Figure 7.4: Schematic of the barrel region of ATLAS Inner Detector, as crossed by a charged particle [21].

The pixel detector consists of three concentric layers of silicon pixel modules in the barrel region, measuring 50.5 mm, 88.5 mm, 122.5 mm from the beam axis. In the end-caps, the modules are located on perpendicular disks measuring 495 mm, 580 mm, and 650 mm from the IP. There are 1456 pixel modules in the barrel and 288 in the end-caps. The high-granularity design results in approximately 80 million pixels, each with individual readout channels, about half of the number of channels in ATLAS.

The modules measure $2 \times 6 \text{ cm}^2$ in area, have a thickness of $250 \mu\text{m}$, and are composed of roughly 47,000 silicon pixels. Likewise, the silicon pixels measure $50 \times 400 \mu\text{m}^2$ in area, which is how precision tracking and vertexing is obtained. The position resolution is $10 \mu\text{m}$ in the $R - \phi$ plane and $115 \mu\text{m}$ in the z (R) direction of the barrel (end-caps).

7.3.2 Semiconductor Tracker

The SCT [23] surrounds the Pixel Detector and forms the next layer of the ID. It covers the same range as the Pixel Detector with $|\eta| \leq 2.5$, and works in a similar way, making use of the ionization of the silicon for charged particle detection. Like the Pixel Detector, the high granularity of the SCT results in approximately 6.3 million readout channels.

In the barrel region, the SCT is comprised of four cylindrical layers of silicon microstrip sensors, and in the forward region, nine perpendicular disks are located in each end-cap. The 285 μm thick silicon microstrips measure 80 $\mu\text{m} \times 12 \text{ cm}$, allowing the SCT to maintain coverage over a larger surface area, as it is further from the beam line.

The silicon microstrips are two sided, allowing for two hits per layer. Each layer is arranged stereoscopically; i.e., angled with respect to the other by 40 mrad, allowing for a position measurement in all three coordinates. In the end-caps, the strips run radially and utilize the two-layer stereo structure as well. The position resolution of each module is 17 μm in the $R - \phi$ plane, and 580 μm in the z (R) direction of the barrel (end-caps).

7.3.3 Transition Radiation Tracker

The outermost subsystem of the ID is the TRT [23], which covers a range $|\eta| \leq 2.0$. It provides additional tracking, utilizing a large number of hits, up to 36 on average per track. Like the Pixel Detector and SCT, the TRT is composed of separate barrel and end-cap modules. The barrel consists of many layers of gaseous straw drift-tubes, which run parallel to the beam axis, and are interleaved with transition radiation material. The construction of the end-caps is similar, with the straws arranged radially into 18 wheels.

The drift tubes are 4 mm in diameter, 144 cm (37 cm) long in the barrel (end-cap)

region, and are filled with a Xe-based gas mixture. The interior of each tube is coated with a conducting layer of aluminum, and a gold plated tungsten wire runs along the center. This cathode-anode pair is held at a potential of a few kV. When a charged particle passes through the transition radiation material, X-ray photons are produced. The particle and the photons pass through the drift tube and ionize of the Xe-based gas. The gas electrons are collected at the anode, and the resulting current is stereoscopically read out.

The TRT is also used to perform particle identification. The amount of transition radiation produced by a charged particle is proportional to the Lorentz factor $\gamma = E/m$. The electron, with its small mass, typically produces significant amounts of transition radiation, while the pion, almost 300 times heavier, produces much less. This results in different amounts of electric current collected from the ionized gas. Typically, seven to ten high-threshold hits from transition radiation are expected for electrons with energies above 2 GeV [23].

The TRT has a position resolution of $130 \mu\text{m}$ per drift tube in the $R-\phi$ plane. While the position resolution of the TRT worse than the Pixel or SCT subdetectors, its large number of drift tubes ensure many measurements are taken, especially far from the beam line, which contribute significantly to the momentum reconstruction of the charged particle. The use of the lower resolution drift tubes results in approximately 351,000 TRT readout channels.

7.4 Calorimetry System

The ATLAS calorimetry system [23] surrounds the inner detector, and covers the region $|\eta| < 4.9$. It is designed to provide dedicated energy measurements for incident electrons, photons, jets, and hadrons. The calorimetry system consists of three independent subsystems: the Electromagnetic Calorimeter (ECAL), the Hadronic Calorimeter (HCAL), both of which

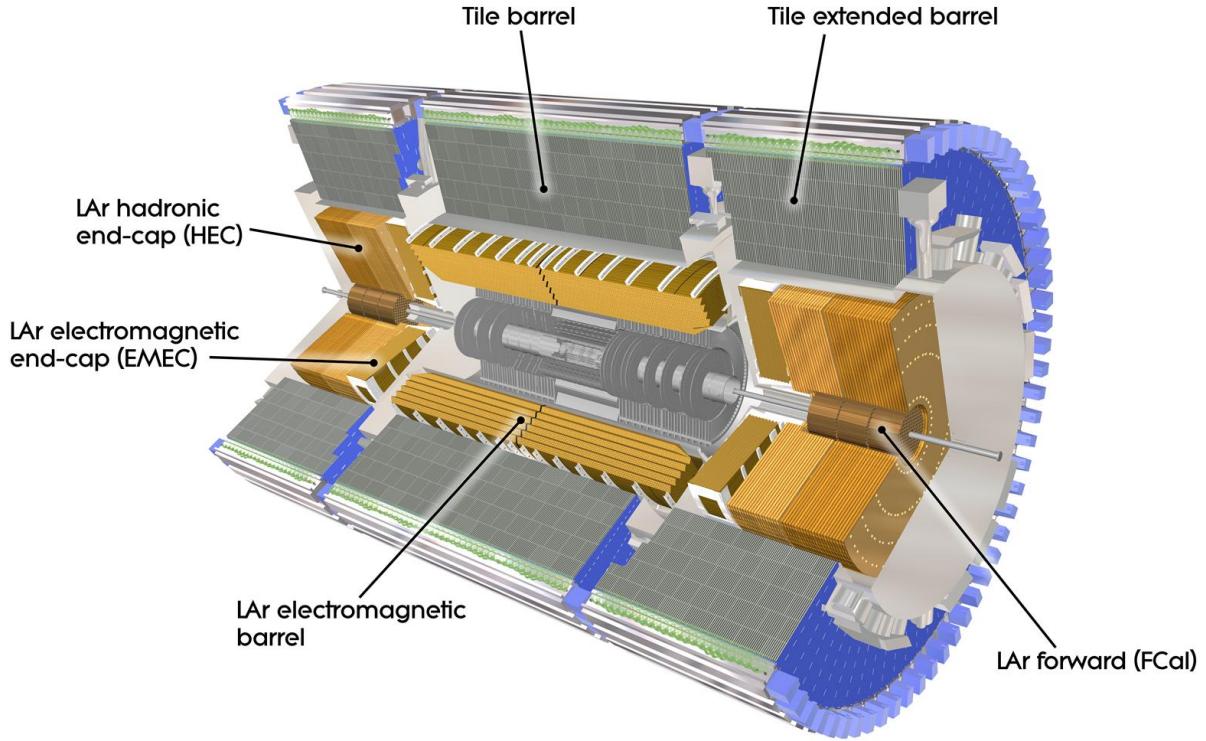


Figure 7.5: Cut-away view of the ATLAS calorimetry system. Depicted are the Electromagnetic and Hadronic sampling calorimeters, both of which consist of individual barrel and end-cap modules. The forward components are closest to the beam line [22].

cover up to $|\eta| < 3.2$, and the Forward Calorimeter (FCAL), which covers the forward region $3.2 < |\eta| < 4.9$ (See Fig. 7.5). The ECAL and HCAL consist of both barrel and end-cap components, while the FCAL is entirely forward by definition.

Each calorimetry subsystem relies on sampling calorimetry, which makes use of two distinct materials sandwiched together in alternating layers: a dense absorber and an active medium. An incident particle passing through the dense layers of absorbing material initiates a particle cascade, or **shower**, of lower energy secondary particles. The secondary particles in the shower interact with the active layers, which are connected to readout electronics for measurement. The shower continues to propagate until a minimal energy threshold is

reached, at which new particle production stops. In this way, the energy of the incident particle is obtained by measuring the energy deposited by its shower over the multiple active layers of the calorimeter. An energy calibration is required using particles of known energies.

7.4.1 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [23] measures the energies of electrons and photons. The EM calorimeter is divided into a central barrel, covering the region $|\eta| < 1.475$, and two symmetric end-caps for $1.375 < |\eta| < 3.2$. The EM barrel (EMB) takes the form of two identical half-barrels, separated by a small gap (4 mm) at $z = 0$, while the EM end-caps (EMEC) are themselves divided into two coaxial wheels (See Fig. 7.5). The outer wheel runs from $1.375 < |\eta| < 2.5$, and the inner one from $2.5 < |\eta| < 3.2$. The region between the barrel and end-caps, $1.37 < |\eta| < 1.52$, is known colloquially as the **crack region**, which consists of cables and cooling for the ID.² The ECAL uses lead plates as the absorbing medium and liquid Argon (LAr) as the active medium, which fills the gaps between the absorber layers, to produce a signal, while copper electrodes are used for readout. The alternating layers of active and passive material are arranged in an accordion style geometry, which leads to very uniform performance in terms of linearity and resolution as a function of ϕ . The design resolution is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ [23].

When a charged particle passes through the dense lead absorbers, it decelerates, resulting in the emission of high energy photons due to bremsstrahlung. These photons convert into electron-positron pairs, initiating an electromagnetic cascade, and forming a shower of lower energy particles. The shower particles ionize the LAr, and the ions drift toward electrodes

²Due to the poor energy resolution, particles from this region are usually excluded in ATLAS physics analyses.

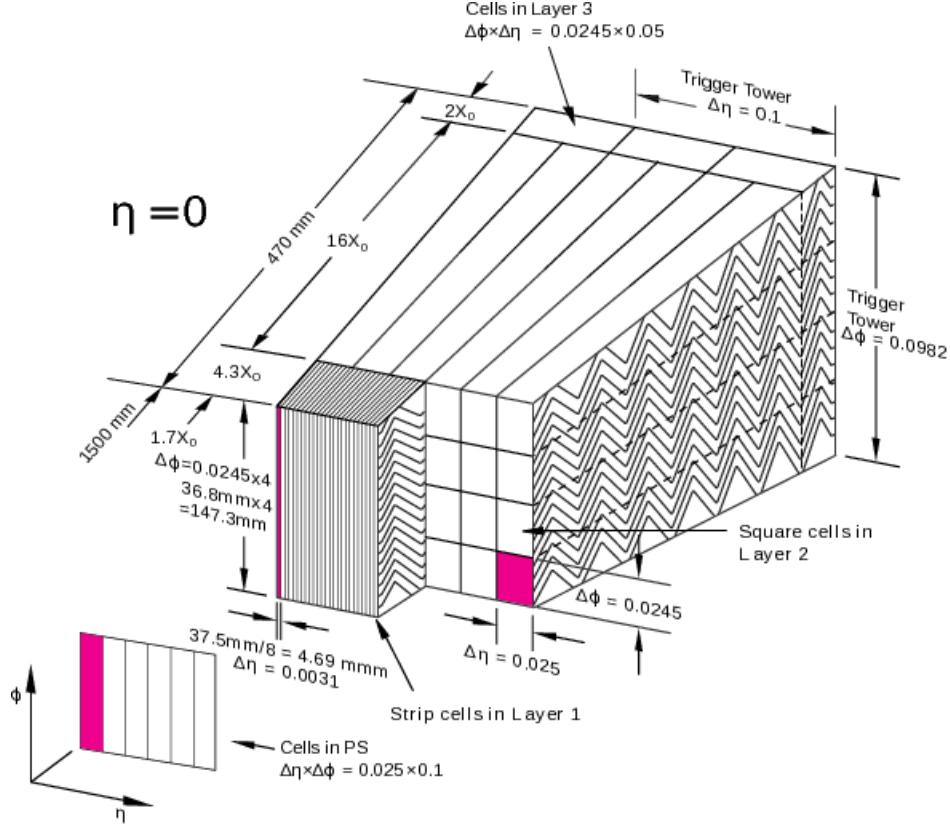


Figure 7.6: Depiction of an ECAL barrel module in the precision region of the ATLAS detector. The accordion shape is shown, along with the three independent sampling layers. The granularity in η and ϕ of the cells of each layer is labeled [23].

for readout, which results in a signal proportional to the deposited energy.

The precision region of the detector covers up to $|\eta| < 2.5$, matching the acceptance of the ID. Here, the ECAL is subdivided into three sampling layers of varying granularities in η and ϕ (See Fig. 7.6). The first sampling layer has a depth of 4.3 radiation lengths X_0 ,³ and is composed of cells finely segmented in η , with a granularity of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$. The fine granularity allows for precision position measurements to be made, and can be used to distinguish photons from converted pions.⁴ The second sampling layer has a depth of $16 X_0$, and collects the majority of the energy of the electromagnetic shower. It is also

³The width of the absorber is measured in radiation lengths X_0 , the distance over which an incident electron loses $1/e$ of its initial energy.

⁴The in-flight decay $\pi^0 \rightarrow \gamma\gamma$ can result in two highly collimated photons.

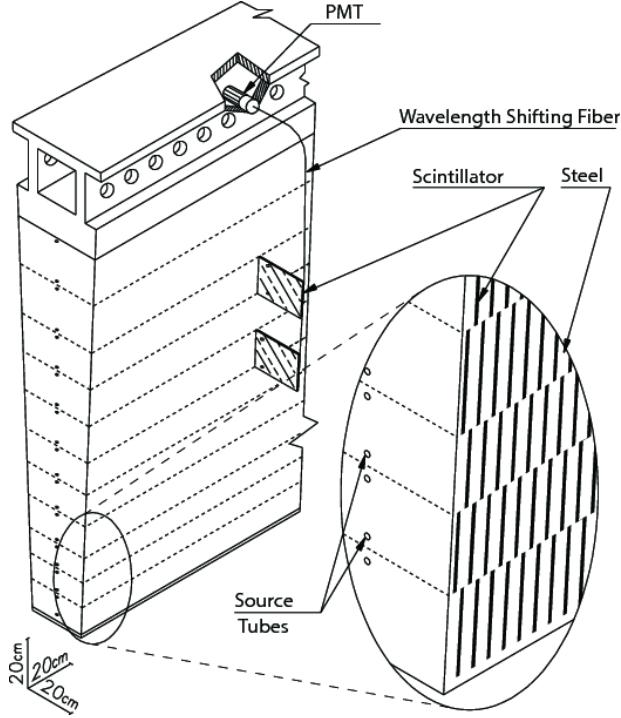


Figure 7.7: Sketch of an ATLAS tile calorimeter module [24].

coarser, with a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$. The third layer is the coarsest at $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$, and is only reached by the highest energy particles, collecting the tail of the shower.

In addition, within the region $|\eta| < 1.8$ and situated before the first sampling layer, is a layer of LAr, known as the presampler, which helps correct for energy losses of incident electrons or photons due to uninstrumented material within the inner detector. Outside the precision region, there is less material blockage, so a presampler is not necessary.

7.4.2 Hadronic Calorimeter

The hadronic calorimeter (HCAL) [23] forms the next detector layer of the ATLAS detector, covering a region up to $|\eta| < 3.2$. Like the ECAL, it is a sampling calorimeter, and consists of two subdetectors: the tile calorimeter and the hadronic end-cap calorimeter (HEC).

The tile calorimeter consists of a barrel, covering $|\eta| < 1.0$, and two extended barrels for $0.8 < |\eta| < 1.7$, each segmented in three layers. Compared to the ECAL, the granularity $\Delta\eta \times \Delta\phi$ is much coarser, measuring 0.1×0.1 in the first two layers, and 0.2×0.1 in the third. As a sampling calorimeter, it uses scintillating tiles as its active material, interleaved with layers of passive steel absorbers (See Fig. 7.7). Hadronic interactions with the scintillator material causes it to luminesce. Light signals are read out by wavelength shifting fibers connected to photomultiplier tubes, which convert the light signals into an electric current to be read out by electronics.

The HEC covers the region $1.5 < |\eta| < 3.2$ and is located in front of the EMEC. It is a LAr sampling calorimeter with copper plate absorbers, which are more robust for the high radiation environment of the forward region of the detector. It consists of two wheels per end-cap where each wheel is divided into two segments for a total of four segmented layers.

The forward calorimeter (FCAL) covers the range $3.1 < |\eta| < 4.9$. It is located in front of the EMEC and below the HEC (See Fig. 7.8) and is closest to the beam line. It consists of three modules which are used to make electromagnetic and hadronic measurements. The design energy resolution for the tile and HEC is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$, and for the FCAL it is $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$.

7.5 Muon System

The Muon Spectrometer (MS) [23] forms the outermost layer of the ATLAS detector (See Fig. 7.9). It is a tracking detector designed to measure the position and momentum of muons in the range $|\eta| < 2.7$, as well as trigger on muons in the range $|\eta| < 2.4$, based on their deflection in an applied magnetic field.

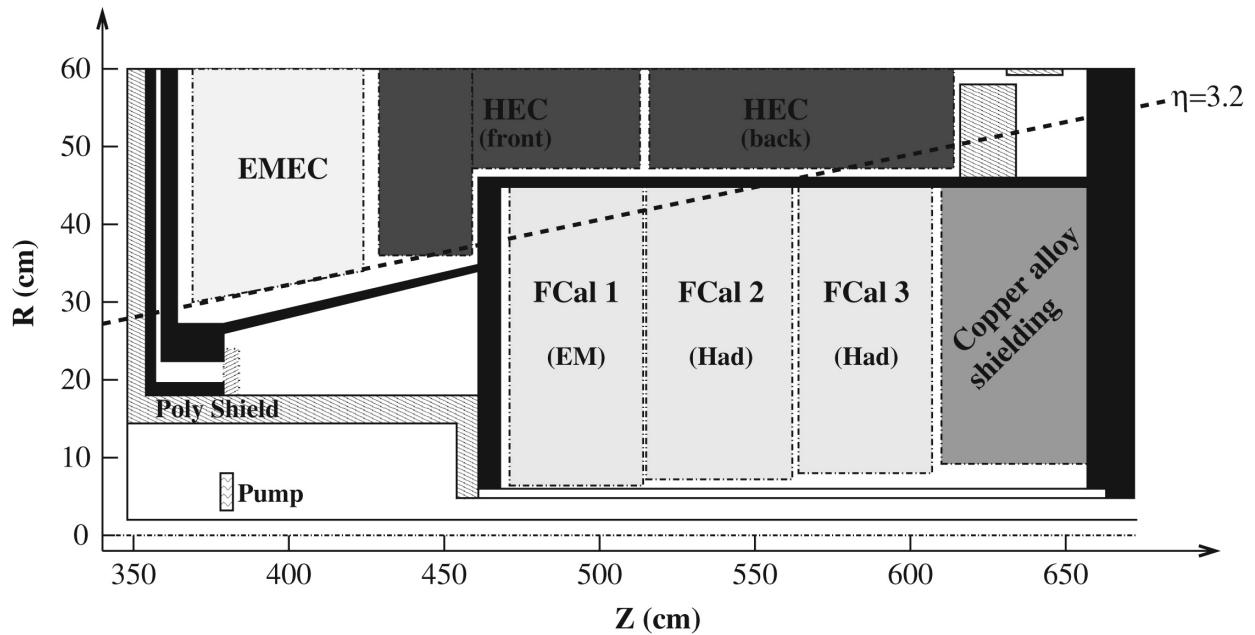


Figure 7.8: Schematic of the ATLAS detector forward calorimeter systems. The three FCAL modules are depicted [25].

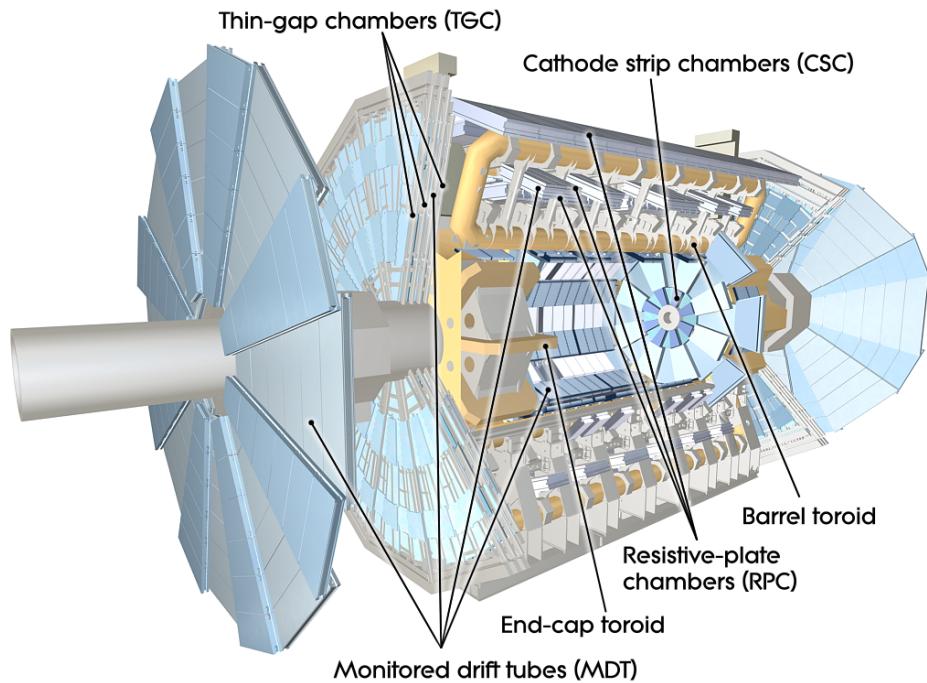


Figure 7.9: Cut-away view of the ATLAS Muon System [26].

Three toroid magnets, one in the barrel and two in the end-caps, produce the necessary magnetic field. The superconducting air-core toroid magnet is formed by an eight-fold symmetric set of 25.3 m long flat race-track coils, which operate at a current of 20.5 kA and temperature of 4.6 K to generate an approximately 0.5 T magnetic field for the region $|\eta| < 1.4$ of the MS. The two additional end-cap toroids are positioned at each end of the central solenoid magnet (for the ID) and provide a 1.0 T magnetic field for additional bending power in the MS end-caps, $1.6 < |\eta| < 2.7$. The **transition region** is defined by the gap in coverage between the two η regions, which relies on the fields of each magnet.

Muons are highly penetrating particles. They leave energetic hits in the ID, but pass almost uninterrupted through the calorimeters before reaching the MS, which is comprised of many gas-filled chambers. As they pass through, the muons ionize the gas, producing a series of hits, which can then be used to reconstruct its track. Four subdetectors are used: the Monitored Drift Tube chambers (MDT) and the Cathode-Strip Chambers (CSC) perform precision tracking measurements, while the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC) provide triggering capabilities and secondary track measurements.

The strong fields and large size of MS allow for the trajectories of energetic muons to be measured with a high level of accuracy; something not possible with the use of the ID subsystems alone. The MS can measure muon momenta up to 3 TeV, and provides a transverse momentum resolution of 10% at 1 TeV.

7.6 Trigger and Data-Acquisition System

An essential component of any collider experiment is an efficient trigger and data-acquisition (TDAQ) system. The trigger selects collision events that contain physics processes of interest,

proton - (anti)proton cross sections

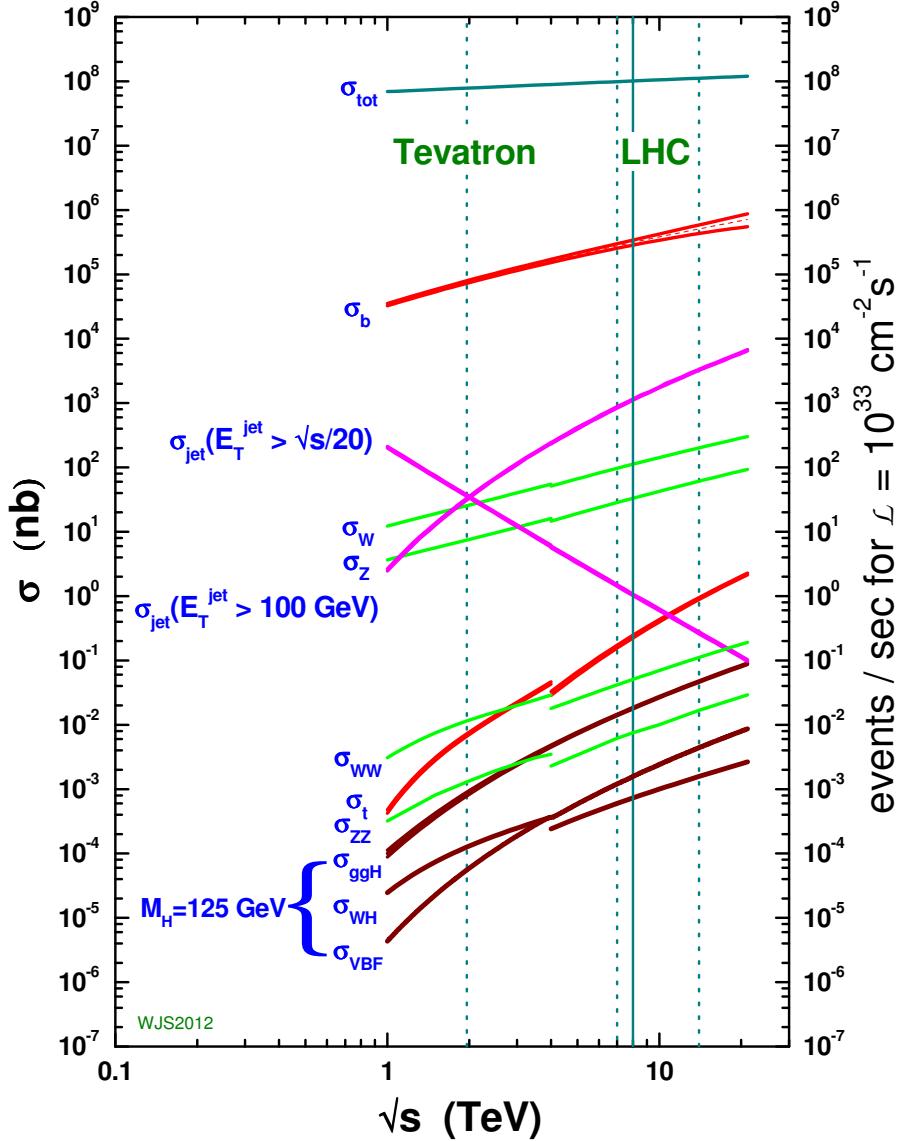


Figure 7.10: The production cross sections and corresponding event rates for a variety of physics processes observable at the LHC. BSM processes, such as Z' production, are expected to occur at extremely small rates compared to the nominal rate of proton-proton inelastic collisions indicated by σ_{tot} . In this regard, a trigger is essential to correctly identify and select such events for offline analysis [8].

and in so doing, reduces the enormous event rate provided by the high luminosity of the LHC to something manageable that can be recorded and permanently stored for subsequent offline analysis (See Fig. 7.10). The ATLAS trigger system consists of two tiers, a hardware-based first-level (L1) trigger, and a single software-based high-level trigger (HLT). With this setup, the 40 MHz bunch crossing rate is reduced to a rate of 100 kHz at the L1 trigger, and then to an average recording rate of about 1 kHz at the HLT, which is compatible with offline computing and storage requirements [137]. The average size of an event accepted by the HLT is on the order of 1 MB. The types of signatures accepted by the trigger are encoded in the trigger menu, which contains several hundred predefined items.⁵

The L1 trigger is implemented with custom electronics, and is designed to search for regions of interest (RoIs) of high-energy activity within the detector using coarse-grained information from the L1 calorimeter trigger (L1Calo) and L1 muon (L1Muon) trigger systems. The window to make a positive identification is about $2.5 \mu\text{s}$, during which time the event is stored in temporary pipeline memories [137]. The L1 trigger decision to keep, and pass the event to the HLT, or to discard it, is made by the Central Trigger Processor (CPT). In the case of an L1 trigger acceptance, the event is buffered in the Read-Out System (ROS), and sent to the HLT for subsequent processing.

The HLT is purely software based and reduces the L1 trigger output rate of 100 kHz to approximately 1 kHz on average within a processing time of about 200 ms. Software algorithms run over the RoIs identified by the L1 trigger, or if needed over the full detector, in order to reconstruct the event. In addition, finer-granularity calorimeter information, precision measurements from the MS, and tracking information from the ID, are made use

⁵For example, the `HLT_2e17_lhloose` trigger, used to select events in the dilepton analysis, requires an event to have at least two electrons with $E_T > 17 \text{ GeV}$ that pass a loose level of likelihood based identification at the HLT level.

of, which are not available at the L1 trigger [138].

In order to minimize processing time, most HLT triggers follow a two-stage approach, where a fast first-pass reconstruction is performed to reject events early on, and a slower precision reconstruction is applied in the second-stage for any remaining events. The reconstruction algorithms run on a unique PC farm of about 40,000 processor cores [138]. Events accepted by the HLT are transferred to the CERN computing center, and stored in the RAW (bytestream) data format on magnetic tapes for later reprocessing. With an average of 25 interactions per bunch crossing at 40 MHz in 2016, of all the proton-proton collisions observed by the ATLAS detector, only about one in a million is saved for later offline analysis.

Chapter 8

Electron Reconstruction in ATLAS

The search conducted in this dissertation is focused on a final-state defined by an electron-positron pair. This chapter describes the reconstruction and identification of **central** ($|\eta| < 2.47$) electron candidates in the ATLAS experiment. Electrons are reconstructed from the hits they leave in the Inner Detector (ID) and from the energy clusters they deposit in the EM calorimeter.¹ Dedicated algorithms are used to reconstruct these detector level signals into meaningful physics objects usable for analysis. Moreover, identification criteria are applied to the reconstructed objects in order to separate true electron candidates from so-called **fakes**, or non-electron objects like jets which have been misidentified.

8.1 Track and Vertex Reconstruction

As a charged particle, an electron that passes through the central region of the ATLAS detector leaves a series of hits in the ID and an energy deposit in the EM calorimeter as its signature (See Fig. 8.1).² The reconstruction of central electrons starts with **track reconstruction**, which aims to reconstruct the trajectories of charged particles according to the hits left in the three ID subsystems. Here, sophisticated track reconstruction algorithms

¹ Aside from the curvature of their tracks, electrons and positrons share the same signature in the ATLAS detector, so for the remainder of the chapter, unless otherwise stated, positrons are also denoted as “electrons.”

²Muons do as well, but leave only a negligible amount of energy in the calorimeter, since they are highly penetrating.

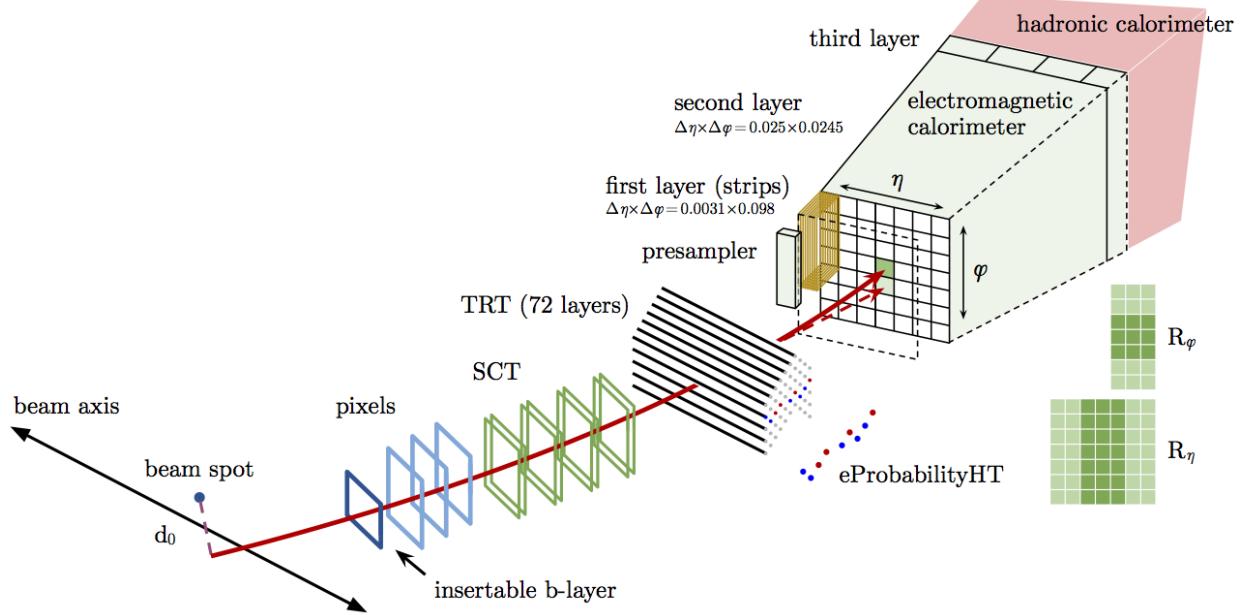


Figure 8.1: Depiction of ATLAS detector systems used to perform electron reconstruction and identification [3].

are employed and are discussed below.

After each event track has been reconstructed, **vertex reconstruction** occurs, where the tracks are extrapolated toward the beam line to reconstruct the positions of the event vertices, which originate from the multiple proton-proton interactions during a bunch crossing. Likewise, sophisticated vertex finder algorithms are used for the reconstruction.

8.1.1 Track Reconstruction

Several track finding strategies exist within the ATLAS experiment [139]. The primary strategy is known as the “inside-out” strategy [140], which is also the baseline for the reconstruction of **prompt** charged particles.³ Here, the reconstruction follows the natural particle trajectory, starting from the silicon hits present in the pixel and SCT detectors, and

³Prompt, or signal-like, electrons originate from an event primary vertex, such as from the decay of a W or Z boson.

gradually extends outward to the TRT.

The inside-out strategy works in two steps: first pattern recognition is employed, followed by a track fit. In the first step, pixel and SCT hits are transformed into three dimensional “space-points.” Pattern recognition software is used to form track candidates (or track seeds) out of 3 silicon space-points, providing an initial track direction and orientation. This is followed by a track-building process where the directional information is propagated through the detector layers, away from the interaction point, and silicon hits that match the assumed trajectory are included in the updated track candidate definition.

The track-building procedure is performed using a Kalman filter algorithm [139], which updates the track fit iteratively with each successive hit, accounting for the material effects of the detector. The tracks are formed according to the “pion hypothesis,” where energy losses due to radiation are assumed to be negligible. However if a track candidate is coincident with an EM cluster and cannot be fully extended using the pion hypothesis, the “electron hypothesis” is used instead. Under this hypothesis, the trajectory is allowed to be affected by energy losses, such as from bremsstrahlung, as the particle can interact with detector material at each material surface [3].

In the second step, track candidates are extended to the TRT, and refit using both the silicon and TRT information again assuming either the electron or pion hypothesis. The refitted track is compared to its silicon-only segment and retained only if the two are consistent, and the track contains at least seven silicon hits. Tracks successfully reconstructed with the inside-out algorithm must have a transverse momentum $p_T > 400$ MeV.

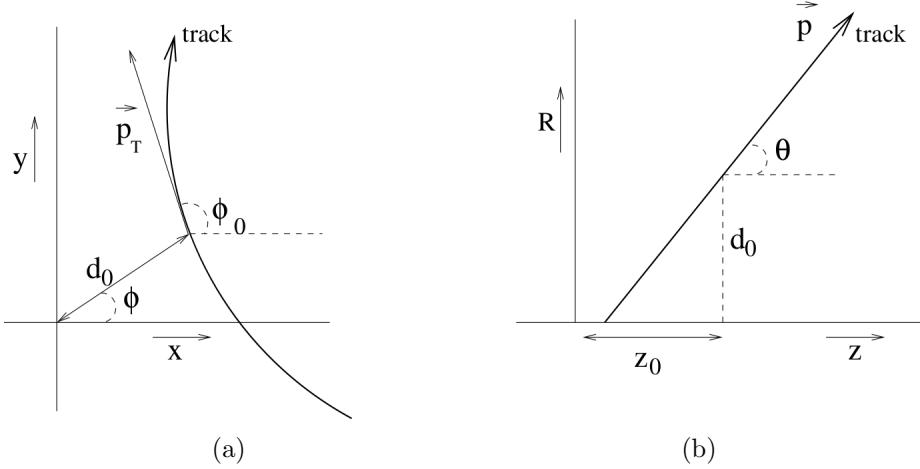


Figure 8.2: Illustration of the transverse impact parameter d_0 (a) and longitudinal impact parameter z_0 [27].

8.1.2 Vertex Reconstruction

After all tracks have been successfully reconstructed, vertex finder algorithms [140] are used to assign them to their respective vertices.⁴ Vertex reconstruction starts by forming vertex seeds out of bunches of closely related tracks. Tracks incompatible with the initial seed go on to form new seeds, and the process repeats until a set of consistent vertices has been identified and all tracks are uniquely matched [140]. Reconstructed vertices are required to have at least two associated tracks, and the primary vertex is defined as the one with the highest $\sum p_T^2$ of its associated tracks.

The position of the primary vertex can be described by the transverse and longitudinal impact parameters of its tracks (See Fig. 8.2). The transverse impact parameter d_0 measures the distance of closest approach between the track and the primary vertex in the transverse plane of the detector, while the longitudinal impact parameter z_0 measures the distance of the track from the vertex along the beam line (the z -axis). The d_0 significance, the value

⁴Vertices can originate as pile-up vertices from soft-interactions during the bunch crossing, as secondary vertices from in-flight particle decays, and as the primary vertices from event hard-scatters.

of d_0 divided by its uncertainty, can be used to separate prompt from non-prompt electron candidates, as non-prompt candidates usually have a larger value.

8.2 Electrons

Electrons in the central region of the ATLAS detector are reconstructed from energy deposits in the EM calorimeter that are matched to charged particle tracks in the inner detector. The procedure starts with the triggering of a high energy cluster in the EM calorimeter, and is followed by the reconstruction and calibration of the measured cluster energy. Discrimination between signal-like (or true) and background-like (or fake) electrons can be achieved with specific electron identification and isolation requirements.

8.2.1 Electron Trigger

Events with prompt electrons are selected by the ATLAS triggering system. All electron triggers are seeded at L1. Here, L1 trigger algorithms search for one or more EM calorimeter trigger objects above a certain energy threshold, and identify RoIs where clusters are built [141].

Events passing the L1 trigger are sent to the HLT, which employs two further stages of selection. The first stage algorithms are less precise, but are fast to reject events early on and reduce the overall trigger rate. The second stage algorithms utilize more precise offline-like requirements, such as isolation and likelihood-based identification, which function to further reduce the event rate. These requirements are looser than the actual offline algorithms to avoid losses in efficiency at this early stage in selection.

Electrons used to perform the electron channel search in the dilepton analysis are selected

with the `HLT_2e17_lhloose` trigger, which requires events to have at least two $E_T > 17$ GeV electrons that satisfy loose likelihood-based identification at the HLT level.

8.2.2 Electron Reconstruction

The reconstruction of central electrons relies on two detector subsystems: the inner detector, which provides tracking information, and the EM calorimeter, which provides energy measurements. In a successful reconstruction, energy clusters are built from cluster seeds in the EM calorimeter, and are subsequently matched to at least one reconstructed track in the inner detector.

The reconstruction starts by building clusters from the energy deposits in the EM calorimeter. Here, the $\eta - \phi$ space of the calorimeter is partitioned into a grid with $N_\eta \times N_\phi = 200 \times 256$ cells. Each cell in the grid has a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, the same as the middle layer of the EM calorimeter (as in Fig. 7.6). Within each grid cell, an “energy tower” is formed by summing together the energy of all calorimeter depth layers. Next, a **sliding window algorithm** [142] scans the grid of energy towers with a fixed 3×5 window. A seed cluster is formed from the energy towers if the transverse energy within the window is at least $E_T > 2.5$ GeV, and if it is a local maximum within the scanned region. Finally, a smaller 3×3 window, less sensitive to detector noise, is used to determine the position of the seed cluster. The energy weighted barycenter of the cells within the resulting seed cluster is used to define its position.

In the next step, a search is performed to match reconstructed tracks to seed clusters. Reconstructed tracks are extrapolated to the middle layer of the EM calorimeter, and are considered loosely matched to clusters if they pass through the cluster barycenter within a cone of $\Delta R < 0.3$. If multiple tracks match the cluster, then the one with the most silicon

hits and the smallest distance to the cluster barycenter is chosen as the primary. A cluster without any matched tracks is discarded as an electron candidate, and reclassified as an unconverted photon. With a successful match, the cluster is rebuilt using a 3×7 window of energy towers in the EM calorimeter barrel, and a 5×5 window in the end-caps. As a final step, the cluster energy is calibrated to the original electron energy using multivariate techniques [143], which are based on simulation.

The four-momentum of the electron candidate is determined from the calibrated cluster energy and the corresponding track parameters. The energy component E comes from the cluster energy, and the directional components θ and ϕ are taken from the matched track. The finalized four-momentum of the reconstructed electron is

$$p^\mu = (E, p_x, p_y, p_z) = (E, p \sin \theta \cos \phi, p \sin \theta \sin \phi, p \cos \theta), \quad (8.1)$$

making it ready for use in physics analysis.

8.2.3 Electron Identification

A sample of reconstructed electrons consists of both prompt electrons and objects that are misidentified as electrons, such as hadronic jets or electrons from in-flight photon conversions, which leave similar signatures in the ID and EM calorimeter.⁵ Therefore, a set of **electron identification** criteria must be applied to distinguish the real electrons from the fakes. This procedure is known as electron identification. For central electrons in Run-2, electron identification relies on a likelihood-based identification technique.

Electron identification algorithms are based on the discriminating variables in Table 8.1.

⁵In the following discussion, unless otherwise stated, prompt electrons are considered “signal,” while all other objects are considered “background.”

These variables describe the lateral and longitudinal shape of EM showers in the calorimeter,⁶ the properties of reconstructed ID tracks, information from the TRT, and the degree to which tracks and energy clusters are well matched [144]. Together, these variables can be combined into a single menu to obtain the desired level of background rejection.

In Run-2, the electron identification algorithms were reoptimized in accordance with the ATLAS detector upgrades. With the commissioning of the IBL, additional hits in the innermost pixel layer could be used to further discriminate between electrons and converted photons. Likewise, modifications to the TRT gas mixture led to an improved likelihood-based method to characterize high-threshold TRT hits.

In Run-1, two different electron identification methods were developed within ATLAS. The first method [3] utilized a simple cut-based approach, where selection requirements are applied sequentially to the track and shower shape variables of the electron candidate in order to make a decision. The second method [3] relies on a likelihood function to evaluate multiple shower shape and tracking variables simultaneously in order to make a selection decision. This multivariate analysis (MVA) technique results in an improved background rejection and increased signal acceptance relative to what could be obtained with the Run-1 cut-based approach. Therefore, in Run-2, the cut-based approach was retired, and the likelihood-based algorithm was made the baseline for electron identification.

The likelihood function is obtained from data, and is constructed out of the signal and background probability distributions of the discriminating variables of Table 8.1. The distributions are two dimensional, and are divided into 9 $|\eta|$ bins and 6 E_T bins to balance the available statistics with their evolution in $|\eta|$ and E_T . The probability for a reconstructed electron to be signal or background is calculated, and is then combined into a single

⁶These often go by the name of “shower shape” variables.

Type	Description	Name
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta < 0.8$ or $ \eta > 1.37$)	$R_{\text{had}1}$
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)	R_{had}
Back layer of EM Calorimeter	Ratio of the energy in the back layer to the total energy in the EM accordion calorimeter. This variable is only used below 100 GeV because it is known to be inefficient at high energies.	f_3
Middle layer of EM calorimeter	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	$w_{\eta 2}$
	Ratio of the energy in 3×3 cells over the energy in 3×7 cells centered at the electron cluster position	R_ϕ
	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position	R_η
Strip layer of EM Calorimeter	Shower width, $\sqrt{(\sum E_i(i - i_{\max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η , and i_{\max} is the index of the highest-energy strip	w_{stot}
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	E_{ratio}
	Ratio of the energy in the strip layer to the total energy in the EM accordion calorimeter	f_1
Track conditions	Number of hits in the innermost pixel layer; discriminates against photon conversions	n_{Blayer}
	Number of hits in the pixel detector	n_{Pixel}
	Number of total hits in the pixel and SCT detectors	n_{Si}
	Transverse impact parameter with respect to the beam line	d_0
	Significance of transverse impact parameter defined as the ratio of d_0 and its uncertainty	d_0/σ_{d_0}
	Momentum lost by the track between the perigee and the last measurement point divided by the original momentum	$\Delta p/p$
TRT	Likelihood probability based on transition radiation in the TRT	ePrHT
Track-cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track	$\Delta\eta_1$
	$\Delta\phi$ between the cluster position in the middle layer and the track extrapolated from the perigee	$\Delta\phi_2$
	Defined as $\Delta\phi_2$, but the track momentum is rescaled to the cluster energy before extrapolating the track from the perigee to the middle layer of the calorimeter	$\Delta\phi_{\text{res}}$
	Ratio of the cluster energy to the track momentum	E/p

Table 8.1: Run-2 electron identification variables [3].

discriminant $d_{\mathcal{L}}$, on which the final selection decision is made, where,

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \quad \mathcal{L}_{S(B)}(\mathbf{x}) = \prod_{i=1}^n P_{S(B),i}(x_i), \quad (8.2)$$

where $P_{S,i}(x_i)$ is the signal distribution and $P_{B,i}(x_i)$ is the background distribution for discriminating variable i evaluated at a value of x_i .

In this approach, three distinct electron identification operating points are defined: LOOSE, MEDIUM, and TIGHT. Each operating point offers an increased level of background rejection at the cost of a decreased signal efficiency. The levels are defined in such a way that a sample of electrons obtained from a tighter operating point will be a subset of looser one. For example, electrons that pass TIGHT will be a subset of those passing MEDIUM, etc. The LOOSE criterion offers a selection that specializes in rejecting light flavor jets, while the MEDIUM and TIGHT criteria include additional discriminating variables in their definitions and are better suited at rejecting electrons from heavy flavor decays and photon conversions [3].

The signal and background likelihood functions are optimized using Monte Carlo (MC) simulation. For the signal, a sample of $Z \rightarrow e^+e^-$ events is used, while the background uses a sample of dijet events. Disagreements are observed between the data and MC in the positions or widths of several variables due to incomplete knowledge of the detector description and the modeling of shower shapes in the detector simulation. Therefore, data-driven corrections are applied to the simulation to match the data during the optimization procedure to ensure similar performance.

8.2.4 Electron Isolation

In addition to the electron identification requirement, it is possible to further separate signal from background with an **electron isolation** requirement. Isolation quantifies the amount of detector activity surrounding the electron candidate, either as additional calorimeter energy or as transverse momenta from ID tracks.⁷ As such, both calorimeter-based and track-based definitions of isolation exist, and both are used to select electrons in the dilepton analysis.

In the calorimeter-based definition, the isolation energy $E_T^{\text{cone}\Delta R}$ is obtained by summing over the transverse energy deposited in each EM calorimeter cell within a certain cone radius ΔR around the center of the electron candidate cluster,

$$E_T^{\text{cone}0.2} = \sum_{\Delta R < 0.2} E_T^{\text{cell}}, \quad \Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (8.3)$$

The transverse energy E_T of the electron cluster itself is not included in the sum. The core of the electron cluster is defined to be within the 5×7 window, measuring $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$, centered on the reconstructed electron cluster barycenter. The isolation is corrected on an event-by-event basis for electron energy leakage into the cone, and for pile-up effects.

In Run-1 this ETCONE method suffered from a high sensitivity to pile-up noise, which decreased its ability to discriminate between signal and background. The situation was improved in Run-2 with the TOPOETCONE method, which naturally applies some noise suppression by using topological clusters, or **topoclusters**, instead of using individual calorimeter cells in the isolation calculation [28]. Topoclusters are built from contiguous calorimeter

⁷Electrons that originate from non-prompt decays or from misidentified hadronic jets exhibit broader energy clusters in the calorimeter and feature multiple ID tracks.

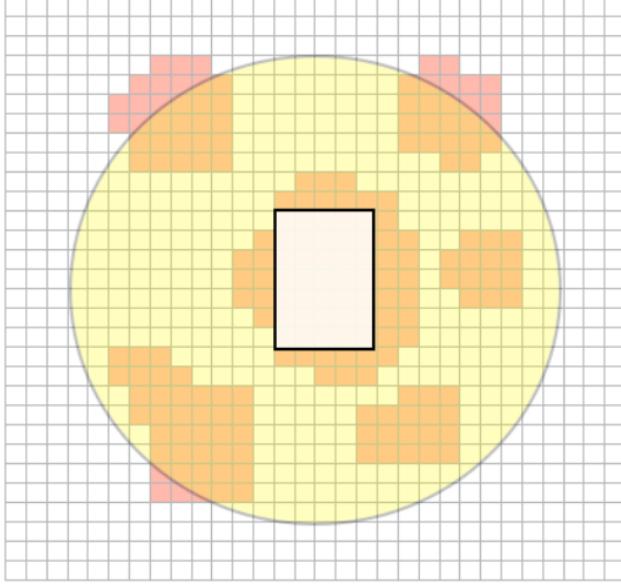


Figure 8.3: Illustration comparing the ETCONE and TOPOETCONE isolation methods. The grid corresponds to the $\eta - \phi$ granularity of the middle layer of the EM calorimeter. The electron candidate cluster is mostly contained in the central 5×7 window. An isolation cone with radius $\Delta R = 0.4$ is drawn around the candidate. In the ETCONE method, all cells (yellow) within the radius ΔR are used, while in the TOPOETCONE method, only topoclusters (red), whose barycenters fall within ΔR , are used, effectively reducing noise from pile-up [28].

cells whose energy falls above a certain threshold. These are then clustered together and the TOPOETCONE isolation variable is constructed by summing up the transverse energy of all topoclusters whose barycenters fall into the isolation cone ΔR (See Fig. 8.3).

Likewise, the track-based isolation variable $p_T^{\text{varcone}\Delta R}$ is obtained by summing over the transverse momenta p_T of all ID tracks within a cone ΔR centered on the track that was matched to the electron candidate cluster,

$$p_T^{\text{varcone}0.2} = \sum_{\Delta R < X} p_T^{\text{track}}, \quad X = \min(0.2, 10 \text{ GeV}/p_T). \quad (8.4)$$

Here, the cone size varies according to the transverse momentum p_T of the track matched to the electron cluster. Tracks must pass certain quality requirements to be included in the

sum: they must have a sufficient number of silicon hits in the ID and must originate from the event primary vertex. Analogously, the transverse momentum p_T of the track matched to the electron cluster is excluded from the isolation sum.

8.3 Energy Corrections

Several effects cause the reconstructed energy of the electron cluster to differ from the true energy of the electron produced in the collision. In order to minimize these differences, energy scale corrections are applied to data, and energy resolution corrections are applied to simulation.

8.3.1 Energy Scale Calibration

An electron candidate is built from its cluster energy during reconstruction (as in Sec. 8.2.2). After a track is successfully matched, the electron energy is calibrated using data-driven corrections to account for any non-uniformities in the detector response.⁸ In addition, a simulation-based calibration using a MVA algorithm is applied to calibrate the electron energy, correcting for any energy deposited in front of the calorimeter or outside the cluster.⁹

Despite this, there can still exist a residual disagreement in the electron energy scale and energy resolution between data and simulation [145]. In order to correct for this residual mismatch, a third set corrections is applied in-situ to recalibrate the absolute energy scale in data. These energy scale corrections are small, at most at the level of 1%, depending on

⁸These corrections are applied only to electrons in data.

⁹These corrections are applied both to electrons in data and simulation.

the particular values of electron η and E_T , and are parametrized as,

$$E_i^{\text{data}} = E_i^{\text{true}}(1 + \alpha_i) \quad (8.5)$$

where E_i^{data} is the measured energy and E_i^{true} is the prediction from simulation. The parameters α_i are η -dependent corrections for each η region i , and are extracted using a sample of $Z \rightarrow e^+e^-$ events by comparing the fitted Z boson lineshape between data and simulation [143].

The ATLAS electron/photon combined performance group (EGP) provides a specialized tool, the `ElectronPhotonFourMomentumCorrection` tool [146], that can be used to apply these corrections to selected data events during analysis.

8.3.2 Energy Resolution Calibration

In addition to the energy scale calibration applied to data, the energy resolution is calibrated in simulation. The relative energy resolution can be expressed [143],

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (8.6)$$

where a , b , and c are η -dependent parameters, known respectively as the “sampling” term, the “noise” term, and the “constant” term, and \oplus refers to addition in quadrature. The sampling term a contributes predominantly at low energy, and is related to the shower fluctuations and sampling frequency in the EM sampling calorimeter. Its design value is around 10%. The noise term b is due to pile-up and electronic noise. The constant term c dominates the resolution at higher energies and has a design value around 0.7%. It parametrizes effects

due to the calorimeter geometry and amount of upstream material, for example, and is not those related to the intrinsic energy of the particle.

The calibration of the energy resolution is achieved by “smearing” the electron energy in simulation by a random Gaussian factor. The width of the Gaussian distribution determines the relative amount of smearing allowed, and depends on the specific η -region of the detector. The difference in resolution between data and simulation can be modeled as

$$\left(\frac{\sigma}{E}\right)_i^{\text{data}} = \left(\frac{\sigma}{E}\right)_i^{\text{true}} \oplus c'_i, \quad (8.7)$$

where c'_i is an effective constant term in a given pseudorapidity region i [145]. These energy smearing factors are derived in-situ by comparing the difference in width of the Z boson lineshape between data and simulation, obtained from a sample of $Z \rightarrow e^+e^-$ events [143]. Likewise, the ATLAS EGP group provides the same dual purpose tool, the `Electron-PhotonFourMomentumCorrection` tool, which can be used in physics analysis to perform the resolution smearing on simulation samples.

8.4 Efficiency Corrections

The **total efficiency** to detect an electron in the ATLAS detector is the product of four individual terms: the reconstruction, identification, trigger, and isolation efficiencies [144], where,

$$\epsilon_{\text{Total}} = \epsilon_{\text{Reco}} \times \epsilon_{\text{ID}} \times \epsilon_{\text{Trigger}} \times \epsilon_{\text{Iso}}. \quad (8.8)$$

The reconstruction efficiency ϵ_{Reco} is the probability that an energy cluster identified in the EM calorimeter is successfully reconstructed as an electron candidate. The identification efficiency ϵ_{ID} is the probability that the reconstructed electron candidate satisfies a particular level of electron identification. The trigger efficiency $\epsilon_{\text{Trigger}}$ is the probability that the reconstructed electron further satisfies an electron trigger. Finally, the isolation efficiency ϵ_{Iso} is the probability that, in addition to being satisfying reconstruction, identification, and trigger requirements, the electron candidate also satisfies a certain isolation threshold.

The efficiencies themselves are calculated from ratios of selected events,

$$\epsilon_{\text{Reco}} = \frac{N_{\text{Reco}}}{N_{\text{Clusters}}}, \quad \epsilon_{\text{ID}} = \frac{N_{\text{ID}}}{N_{\text{Reco}}}, \quad \epsilon_{\text{Trigger}} = \frac{N_{\text{Trigger}}}{N_{\text{ID}}}, \quad \epsilon_{\text{Iso}} = \frac{N_{\text{Iso}}}{N_{\text{Trigger}}}, \quad (8.9)$$

where, for example, the reconstruction efficiency ϵ_{Reco} is obtained from the number of successfully reconstructed electron candidates N_{Reco} in a clean sample of electrons, relative to the total number of electron clusters N_{Clusters} in that same sample. Similar definitions apply to the evaluation of the other efficiencies.

The efficiencies can be estimated using the **tag-and-probe** [3] method. This can be performed with a clean sample of electrons obtained from $Z \rightarrow e^+e^-$ or $J/\psi \rightarrow e^+e^-$ decays. Here, one electron, the “tag,” is required to pass very strict requirements, ensuring it is in fact a true signal electron, while the other electron, the “probe,” is selected only with very loose requirements. The invariant mass of the pair must be within a window centered around M_Z or $M_{J/\psi}$ to reject spurious background combinations. With a clean set of probes, the denominators in Eq. 8.9 can be determined, and the electron efficiencies calculated.

When tested on simulation, the same tag-and-probe technique results in estimates for the reconstruction, identification, trigger, and isolation efficiencies that differ from those

measured in data. The discrepancy arises from the known mismodeling of electromagnetic showers and tracks in the simulation. To account for this, efficiency corrections known as **scale factors** are derived and applied to the simulation during analysis in order to correct for the mismodeling. For a complete discussion, see Section 9.4.2.

Chapter 9

Monte Carlo Simulation

In the electron channel of the dilepton search, the characteristic signature for the production of a Z' resonance in LHC proton-proton collisions can be identified from a pair of prompt, high- E_T electrons. However, the same signature can arise from several unrelated Standard Model processes, which present a non-negligible background to the Z' signal process under investigation. Monte Carlo (MC) simulation is used to estimate the Standard Model contribution from these background processes, and hence allows for a direct comparison to be made against the observed LHC data. These backgrounds include the Drell-Yan process, the $t\bar{t}$ and Wt Single-Top quark processes, and the WW , WZ , and ZZ Diboson processes.¹ This chapter introduces MC event generators, the ATLAS simulation chain, and the MC samples used to model the electron channel of the dilepton analysis.

9.1 Event Generators

Event generators [147] are tools that use **MC sampling** to simulate high energy physics scattering processes.² Combined with a robust detector simulation, event generators can be used to make reliable estimates of kinematic distributions in accordance with Standard

¹An additional “fake” background arises from QCD Multi-jet and $W+jets$ production due to the misidentification of hadronic jets as electrons. This component is estimated using a data-driven technique (See Sec. 10.6).

²Event generators are multi-purpose: they can simulate a wide variety of scattering processes for multiple collider scenarios (e.g. pp , $p\bar{p}$, e^+e^- , $e^\pm p$, etc.).

Model predictions. This allows for direct comparisons to be made between simulation and LHC data, making event generators extremely powerful tools.

Broadly speaking, event generators fall into two categories: the first, parton shower (PS) generators, aims to provide a comprehensive description of the entire high-energy event (as in Fig. 3.3), while the second, Matrix Element (ME) generators, are designed to evaluate specific higher-order (NLO, NNLO, etc.) matrix elements, usually for a single process or a class of similar processes.

9.1.1 Parton Shower Generators

PS event generation is subdivided into several stages [1], and follows closely those outlined in Section 3.1.3. In the first stage, a matrix element is calculated to some fixed order in α_{EM} or α_S , and is evaluated using MC sampling. Here, the user selects a physics process from an available list, and specifies the set of Parton Distribution Functions (PDFs) to be used in the generation.

The matrix element is calculated and inserted into an expression for the partonic cross section, which is then convoluted with the PDFs to produce an expression for the hadronic cross section (as in Eq. 3.6). Events are generated from MC integration, in which the expression for the hadronic cross section is sampled with random numbers [147]. Each sampling specifies the momentum fractions x_1 and x_2 of the colliding partons, and defines the event four-vectors for the initial- and final-state hard process particles.

In the next stage, parton showers are simulated with dedicated parton showering algorithms [147]. Soft radiation is “attached” iteratively, either as $g \rightarrow gg$ or $g \rightarrow q\bar{q}$, to the initial- and final-state partons of the hard process. Each emission or splitting is generated at an energy scale lower than the previous one, and the developing shower is evolved until some

predefined cutoff is reached. Here, a phenomenological hadronization model takes over.

In the third stage, hadronization is simulated, along with the underlying event, and subsequent hadronic decays.

9.1.2 Matrix Element Generators

Whereas PS generators are designed to simulate a broad range of physics processes, and model the subsequent development of the hadronic event, ME generators are instead designed to precisely evaluate higher-order matrix elements.³ The higher-order result obtained from the ME calculation can be applied on top of the corresponding lower-order PS calculation as a way to obtain a more accurate description of the Standard Model process, while simultaneously retaining the parton shower and hadronization effects. This strategy is known as **reweighting** and is employed in the dilepton analysis. Here, mass-dependent **k-Factor** corrections, or scaling factors applied on an event by event basis, are derived and applied to enhance the accuracy of several Standard Model background estimates (See Sec. 9.4.5).⁴

9.1.3 Generators Used in the Dilepton Analysis

In order to simulate the background processes present in the dilepton analysis, several different generators are used. Each generator has its own strengths and weaknesses according to its design philosophy. Event generators used include the following:

PYTHIA [148] is a general purpose PS generator, with a focus on $2 \rightarrow 1$ and $2 \rightarrow 2$ type scattering processes. Hard process matrix elements are calculated at the LO.

³For this reason, they are also known as cross section integrators or calculators.

⁴A k-Factor is obtained from the ratio of a higher-order cross section prediction to a lower order one for a given process, such as, for example, $w_{kF} = \sigma_{NNLO}/\sigma_{NLO}$.

It contains a library of over 200 predefined SM and BSM processes, several models for parton showering and hadronization, built in PDF sets, and a wide variety of phenomenological tunes, making it extremely modular and extensively used within the ATLAS collaboration. If higher-order precision is required, PYTHIA can interface with an external ME generators. Here, the external generator calculates the hard process, and PYTHIA is used to subsequently perform the showering, hadronization, and particle decays.

SHERPA [149] is another general purpose PS generator. Matrix elements calculations are performed at the NLO. Many matrix elements in SHERPA are hard-coded as in PYTHIA, but those not included can be automatically generated according to a user specified final-state. There is no upper limit on the particle multiplicity, making SHERPA a powerful tool for the realistic description of multi-particle final-states. SHERPA provides a comprehensive description of the event, including its own implementation of parton showering, multi-parton interactions, hadronization, and the resummation of soft photons.

POWHEG [150] is a NLO ME generator. It supports a library of standard processes, and is based on the POWHEG method, which specifies how to merge NLO calculations with parton showers. Therefore, POWHEG can be interfaced with PYTHIA, for example, to obtain an improved description of the event, which is often the approach taken by ATLAS physics analyses.

TOP++ [151] is a ME generator designed to calculate the total inclusive cross section for the production of top quark pairs ($\bar{t}t + X$) at hadron colliders. The calculation takes place at the NNLO, with the option of including the effect of soft-gluon resummation

at next-to-next-to-leading logarithmic (NNLL) accuracy.

VRAP [152] is another specialized ME generator. It computes the doubly-differential Drell-Yan cross section at the NNLO in QCD with respect to dilepton invariant mass and rapidity. Higher-order EW corrections are not included. Variations of the factorization and renormalization scales result in a theoretical uncertainty at the less than one percent level.

MCSANC [153] is a ME generator that computes the NLO QCD and NLO EW corrections to the cross section for several Standard Model processes, such as inclusive Drell-Yan production. Photon-Induced contributions of the form $\gamma\gamma \rightarrow \ell^+\ell^-$, including t - and u -channel modes, are included in the EW corrections, which serve as a small, but non-negligible, irreducible background to the high-mass search.

PHOTOS [154] is an “afterburner,” which is used to include QED radiative corrections. PHOTOS interfaces with a “host” generator, which does not account for QED radiation in the parton shower, and proceeds to add final-state radiated photons to the already generated events, accounting for QED interference and multiple photon emissions.

9.2 ATLAS Simulation Chain

In order to be able to compare the Standard Model background simulation to LHC data, the response of the ATLAS detector must be accounted for in the generation as well. This is accomplished with a dedicated simulation of the ATLAS detector [155], which is integrated into the **ATLAS software framework** ATHENA [156] and relies on the use of the GEANT4 [157] simulation toolkit.

The ATLAS simulation chain proceeds in three steps: first, a sample of events for a physics process of interest is generated; then, event particles are passed through a full-scale simulation of the ATLAS detector, accounting for its detailed response; finally, the recorded detector hits and energy deposits are digitized and saved to output.

Aside from MC event generation, the flow of the simulation chain proceeds identically between data and simulation, which allows for the same trigger and reconstruction packages to be used on each, and ensures direct comparisons can be made.

9.2.1 Event Generation

The first step of the ATLAS simulation chain is event generation. A host of event generators are configured to run in ATHENA and the output of the generation is written in the standard HepMC format [158]. The step-by-step procedure for generating events follows the description in Section 9.1.1. Event filtering can be applied according to the specific needs of the analysis.⁵ All prompt particles are written to the event record, as well as all outgoing particles expected to propagate through the ATLAS detector. Just as with LHC data events, simulated events are given unique event numbers, and the simulated dataset receives a corresponding run number.

9.2.2 Detector Simulation

In the next step of the ATLAS simulation chain, the generated events are passed through the ATLAS detector simulation. Each final state particle is propagated through a virtual model of the ATLAS detector described with GEANT4, a software package for the accurate

⁵Event filtering discards certain events from the generation unless they meet some user-defined requirement. For example, prompt electrons must have an $E_T > 20$ GeV.

description of the geometry and material composition of particle detectors. The detailed interactions between the particles and the detector material is simulated, and additional particles that result from bremstrahlung, or from electromagnetic or hadronic showers, are propagated through the detector as well. An accurate description of the detector geometry is crucial for the simulation. Experimental results are continually compared with previous versions of the ATLAS geometric model to help further refine its description [159]. The result is a record of hits and energy deposits for each simulated event. This information is written to an output hit file.

9.2.3 Digitization

In the final stage of the simulation chain, the responses of the detector readout electronics are simulated and digitized using the hit files. The effect of pile-up is included in this stage as well. Pile-up is not strictly simulated as part of the event generation; rather, it is included at this stage by reading in various types of simulated events. Hits from these events are “overlaid” with the hits from the generated event, effectively reducing the amount of computation required.

Furthermore, the hardware-based L1 trigger is configured to run in a “pass-through” mode on the digitized detector information, and its L1 decisions are emulated. No event is discarded, and the set of L1 trigger hypotheses for each event is evaluated and saved. The digitized events are written in the Raw Data Object (RDO) file format, just like real data. Additional “truth level” information is retained, which includes a detailed record of each particle in the event from the generation stage.

ATLAS reconstruction software (as in Sec. 8.2) can be run on the RDO files. The output of reconstruction is an Analysis Object Dataset (AOD) file, which contains both “truth

level” and “reconstruction level” information usable for physics analysis. These files contain all the necessary information to interface with the tools provided by the ATLAS combined performance groups, and apply, for example, energy scale calibration to the data (as in Sec. 8.3.1), and resolution smearing (as in Sec. 9.4.3), and efficiency scale factor corrections (as in Sec. 9.4.2) to the simulation. Likewise, they provide the ability to assess the systematic uncertainties associated with these operations (as in Sec. 10.8).

9.3 Simulation Samples

Each Standard Model background to the electron channel search is defined by a final-state with at least two prompt electrons. An **irreducible** component of this arises from the Drell-Yan and Photon-Induced processes, and non-negligible **reducible** components, in order of largest to smallest, arise from $t\bar{t}$ and Single-Top quark production, Multi-jet & $W+jets$ production, and Diboson production.⁶ The Multi-jet & $W+jets$ backgrounds are the result of the misidentification of one or more non-prompt electrons or hadronic jets as a prompt electrons, and is estimated using a data-driven approach (See Sec. 10.6). All other backgrounds are described by dedicated MC simulation.

9.3.1 Drell-Yan Production

The Drell-Yan process (depicted in Fig. 4.2) is generated at the NLO with the POWHEG event generator, configured with the CT10 [160] PDF set. Showering and hadronization is performed with PYTHIA8 using the ATLAS AZNLO tune [161].

⁶Irreducible meaning the $Z' \rightarrow e^+e^-$ signal has an identical final-state to the $\gamma^*/Z \rightarrow e^+e^-$ background. Reducible backgrounds are identified by other final-state objects in addition to the two electrons. For example, the $t\bar{t}$ and Wt processes include at least one additional b -jet in the final-state, and Diboson processes include neutrinos or additional jets.

m_{ee} [GeV]	500	750	1000	1500	2000	3000	4000	5000
$w_{QCD} = \sigma_{NNLO}/\sigma_{NLO}$	1.044	1.044	1.043	1.035	1.018	0.985	0.955	0.844
w_{EW}	0.985	0.972	0.953	0.927	0.896	0.857	0.834	0.784

Table 9.1: Mass-dependent NLO-to-NNLO QCD and higher-order EW k-Factor corrections for several values of dielectron invariant mass. These corrections are applied to the enhance the nominal Drell-Yan background estimate with NNLO QCD and higher-order EW effects.

POWHEG is used to generate 19 individual mass-binned samples ranging from 120 GeV to 5000+ GeV in slices of true dielectron invariant mass m_{ee} (See Appendix A for details). Generating multiple mass-binned samples ensures a small statistical uncertainty is achieved and results in an adequate number of events to model the kinematic distributions of interest. Below 120 GeV, a dedicated sample of $Z \rightarrow e^+e^-$ events is used to model the Z boson mass peak (See Appendix A for details). This sample has a lower mass threshold of 60 GeV, but no upper threshold, so an upper limit of 120 GeV is applied manually to smoothly match it with the lowest mass-binned sample.

In order to improve the accuracy of this prediction, Drell-Yan event yields are enhanced by several mass-dependent k-Factors corrections. These include higher-order QCD corrections, derived with VRAP at the NNLO, and higher-order EW corrections derived with McSANC at the NLO. Table 9.1 lists the characteristic values of the corrections as a function of dielectron invariant mass. Further detail on the derivation and application of k-Factor corrections is provided in Section 9.4.5.

9.3.2 Photon-Induced Production

An additional irreducible component of the background comes from the Photon-Induced process $\gamma\gamma \rightarrow \ell^+\ell^-$ through t - and u -channel modes, as depicted in Figure 9.1. This process generally makes a small contribution to the overall Standard Model background estimate,



Figure 9.1: The LO Feynman diagrams for Photon-Induced Drell-Yan production. Both t -channel (a) and u -channel (b) modes are shown.

but becomes sizeable at high mass, as depicted in Table 9.2. The Photon-Induced process is not included in the dedicated generation of Drell-Yan; instead, it is accounted for with a mass-dependent k-Factor correction, derived using the MRST2004QED [162] PDF set. The photon PDF $\gamma(x, Q^2)$ is included in this specialized PDF set, which enables the initial-state production of QED related processes.

m_{ee} [GeV]	500	750	1000	1500	2000	3000	4000	5000
w_{PI}	1.039	1.045	1.051	1.060	1.075	1.111	1.139	1.263

Table 9.2: Mass-dependent Photon-Induced k-Factor corrections at several values of dielectron invariant mass. These corrections are applied to the nominal the Drell-Yan background estimate to incorporate Photon-Induced effects.

9.3.3 $t\bar{t}$ Production

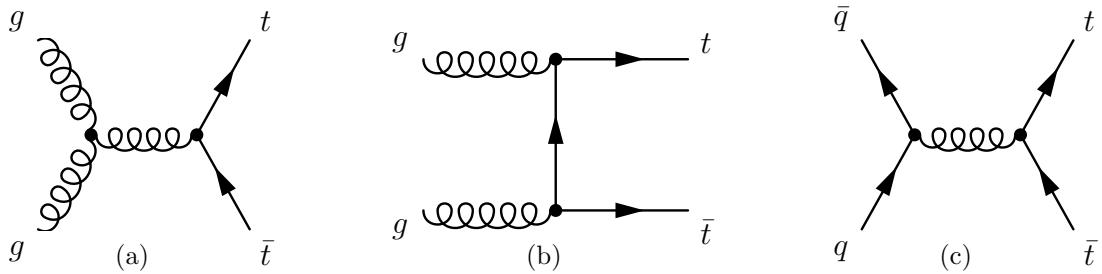


Figure 9.2: The LO Feynman diagrams for $t\bar{t}$ production. The gluon initiated s -channel (a) and t -channel (b) modes are depicted, along with quark-antiquark initiated production (c).

The $t\bar{t}$ process is generated at NLO using PowHEG with the CT10 PDF set. The top quark mass m_{top} is set to 172.5 GeV. A single sample is produced, and no mass-threshold requirements are made in the generation. The PYTHIA6 generator, configured with the Perugia 2012 tune [163], is used to model the underlying event and parton showering.

A global k-Factor is derived using the TOP++ 2.0 generator, which brings the cross section from the NLO to the NNLO in QCD, and includes Next-to-Next-to-Leading-Log (NNLL) terms from soft-gluon resummation. The total $t\bar{t}$ cross section comes to 831.76 pb after the application of the global k-Factor correction, which has a value of 1.1949. The relevant $\bar{t}t$ processes are shown in Figure 9.2. The electron channel background originates from $t \rightarrow bW$, where each W boson decays leptonically through $W \rightarrow e\nu$.

9.3.4 Single-Top Quark Production

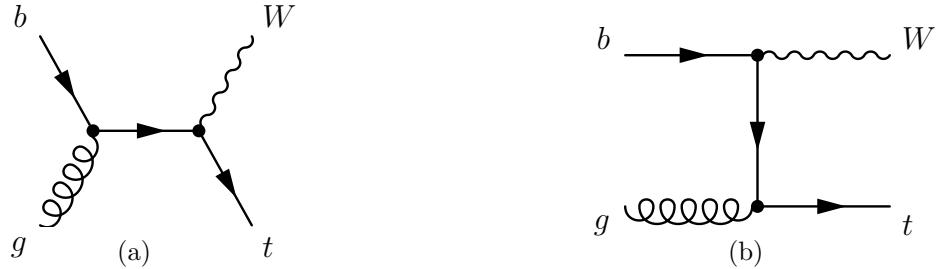


Figure 9.3: The LO Feynman diagrams (a) and (b) for Wt -channel Single-Top production.

The majority of top quarks at the LHC are produced in pairs via the strong interaction (See Fig. 9.2). However, the weak interaction allows for the production of a *single* top quark in association with another particle (See Fig. 9.3. The Single-Top background arises predominantly from so-called Wt -channel Single-Top production, where a top quark is produced in association with a W boson.⁷

⁷Single-Top production also proceeds through s - and t -channel modes.

This background is generated at NLO in QCD with Powheg with the CT10 PDF, and showed in PYTHIA6 using the Perugia 2012 tune. Again, TOP++2.0 is used to bring the calculation from NLO to NNLO in QCD, where the total cross section measures 71.7 pb after the application of the global k-Factor correction, whose value is 1.054.

9.3.5 Diboson Production

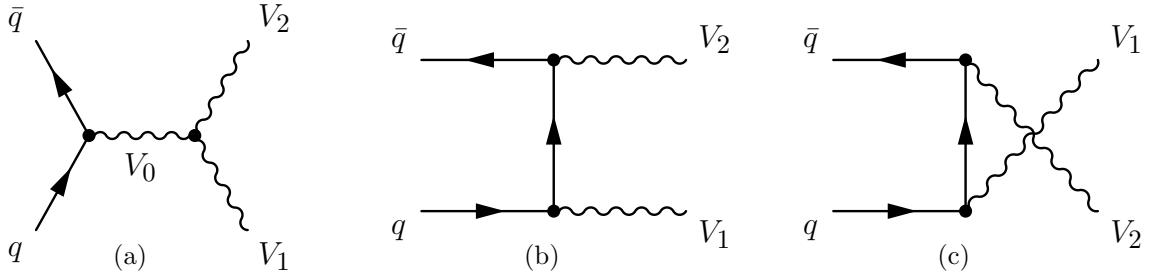


Figure 9.4: The LO Feynman diagrams for Diboson production. Depicted are the s -channel (a), t -channel (b), and u -channel (c) production modes. The WZ and ZZ processes can proceed through all modes, while the $Z\gamma$ process can produced only through the t - and u -channels. In each diagram $V_i = W, Z$ or γ .

The smallest background component arises from Diboson production, or the production of pairs of vector bosons. These pairs form three distinct final-states: WW , WZ , and ZZ , as seen in Figure 9.4. Each process is generated in SHERPA 2.1.1 at NLO with the CT10 PDF, which also includes the parton showering. Because of its small contribution to the overall background, no attempt is made to include a k-Factor correction. As with the Drell-Yan process, a series of mass-thresholds are required in the generation (See Appendix A for details).

In the WW channel, the dielectron background originates from the decay $WW \rightarrow e\nu ee\nu$; in the WZ channel, though the decays $WZ \rightarrow \ell\nu ee$ or $WZ \rightarrow qqee$; and in the ZZ channel, though the decays $ZZ \rightarrow \ell\ell ee$, $ZZ \rightarrow qqee$, or $ZZ \rightarrow \nu\nu ee$. Each set of mass-binned samples covers the entire search region, and provides an adequate statistical uncertainty. The largest

contribution to the Diboson background arises from the $ZZ \rightarrow qqee$ process with a cross section of 16.59 pb.

9.3.6 Z' Production

Dedicated Z' signal samples are generated with PYTHIA8 at LO using the NNPDF23LO [164] PDF set and the A14 tune [165] for showering and hadronization. Here, the several “benchmark” samples are generated: four at Z' pole masses of 2, 3, 4, and 5 TeV for the Z'_χ , and one at a pole mass of 3 TeV for the Z'_{SSM} . The effect of interference with the Standard Model Drell-Yan process is not included in the Z' generation.

These benchmark mass-points are used to validate the **signal reweighting** technique used later for the signal search (See Sec. 11.3) and limit setting (See Sec. 11.4) phases of the analysis.⁸ Signal reweighting is applied to a set of LO Drell-Yan samples, generated with PYTHIA8 using the NNPDF23LO PDF set. These samples are generated with an identical range in truth dielectron invariant mass m_{ee} as the corresponding NLO PowHEG Drell-Yan samples described in Section 9.3.1.

In addition, higher-order QCD corrections are applied to the reweighted Z' signals. A LO-to-NNLO QCD k-Factor, derived using VRAP, takes the LO PYTHIA8 Z' signal estimates to the NNLO, the same order as the PowHEG Drell-Yan background estimate. While the higher-order QCD effects are expected to be the same between signal and background, the higher-order EW corrections are model dependent, and are therefore not applied to the signal. Several values of the QCD k-Factor correction are shown in Table 9.3.

⁸A Signal reweighting tool was developed to generate Z' signal samples for user-defined pole masses and model parameters using the set of LO Drell-Yan samples as input. This avoids the otherwise computationally intensive process of generating new samples for each mass point and model under investigation. See Section 11.2 for more details.

m_{ee} [GeV]	500	750	1000	1500	2000	3000	4000	5000
$w_{QCD} = \sigma_{NNLO}/\sigma_{LO}$	1.356	1.380	1.411	1.435	1.423	1.339	1.264	1.025

Table 9.3: Mass-dependent LO-to-NNLO QCD k-Factor corrections at several values of dielectron mass m_{ee} . These corrections are applied to enhance the nominal Z' signal estimate with NNLO QCD effects.

9.3.7 MC Sample Summary

A summary of the generators and PDF sets used for the electron channel analysis is displayed in Table 9.4. For a comprehensive list of MC samples and their properties, see Appendix A.

Process	Generator	Order	Corrected	Parton Shower	PDF
$q\bar{q} \rightarrow Z/\gamma^* \rightarrow e^+e^-$	POWHEG v2	NLO	NNLO	PYTHIA 8.186	CT10
$t\bar{t}$	POWHEG v2	NLO	NNLO	PYTHIA 6.428	CT10
Wt	POWHEG v2	NLO	NNLO	PYTHIA 6.428	CT10
WW, WZ, ZZ	SHERPA 2.1.1	NLO	NLO	SHERPA 2.1.1	CT10
$q\bar{q} \rightarrow Z/\gamma^* \rightarrow e^+e^-$	PYTHIA 8.186	LO	NNLO	PYTHIA 8.186	NNPDF23LO
$q\bar{q} \rightarrow Z' \rightarrow e^+e^-$	PYTHIA 8.186	LO	NNLO	PYTHIA 8.186	NNPDF23LO

Table 9.4: A summary of MC samples used for the background and signal modeling in the electron channel of the dilepton analysis. The columns include the process of interest, the MC generation program, the matrix element order, the order of the k-Factor correction, the parton shower program, and the PDF set used [4].

9.4 Corrections to Simulation

A series of corrections are applied to the simulation to achieve the closest possible level of agreement between data and MC. In the ATLAS detector simulation, all major components and material interactions are simulated within the detector volume, and the resulting electron distributions are generally well matched to data. However, the modeling is by no means perfect, and discrepancies can arise, for example, from assumptions made about the exact

material distribution within the ATLAS detector. To account for these effects, efficiency scale factor and energy resolution corrections are applied to enhance the simulation. Likewise, corrections to the pile-up distribution assumed in the MC generation are performed, and k-Factor corrections are used to enhance the description of the simulation with the inclusion of higher-order effects absent in the original generation.

9.4.1 Sample Normalization

As the amount of integrated luminosity in the final LHC dataset is not known a-priori, the MC samples used in the analysis were generated with integrated luminosities that were in general different from the dataset value of $\mathcal{L}_{\text{int}} = 36.1 \text{ fb}^{-1}$. Therefore, to match the observed integrated luminosity in data, each background must be correctly normalized. The following **event weight** is used,

$$w_{\text{MC}} = \frac{\sigma_{\text{bkg}} \epsilon_{\text{filter}}}{N_{\text{evts}}} \mathcal{L}_{\text{int}}, \quad (9.1)$$

where σ_{bkg} is the associated background sample cross section, ϵ_{filter} is its filtering efficiency, N_{evts} is the number of events in the sample, and \mathcal{L}_{int} is the integrated luminosity of the dataset. The values used are obtained from the generator outputs listed in Appendix A.

9.4.2 Efficiency Scale Factor Corrections

The total efficiency to identify an electron in the ATLAS detector is the product of the electron reconstruction, ID, trigger, and isolation efficiencies (as outlined in Sec. 8.4). When extracted from simulation, these efficiencies show discrepancies, at the level of a few percent, to those measured in data, due to the imperfect modeling of the ATLAS detector. Therefore,

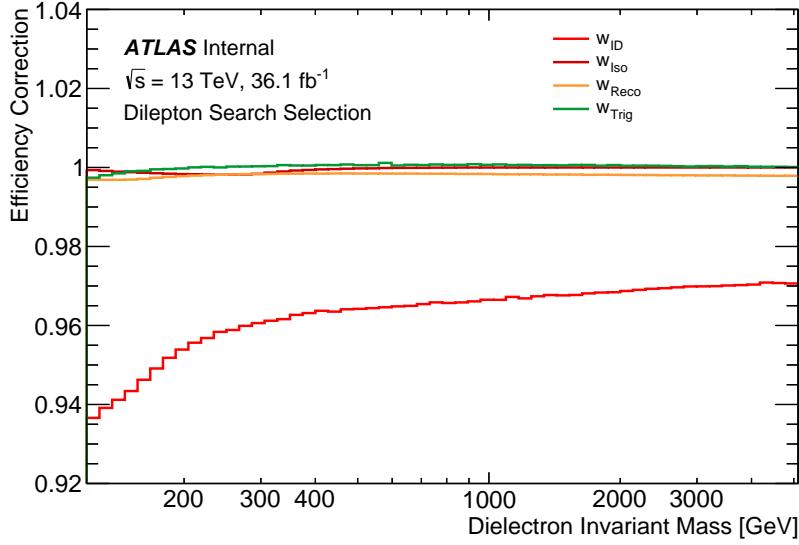


Figure 9.5: Electron efficiency scale factor corrections as a function of dielectron invariant mass m_{ee} . Shown are the scale factors applied to correct the electron reconstruction, identification, isolation, and trigger efficiencies.

to mitigate these effects and bring the simulation into the best possible agreement with the data, efficiency scale factor corrections are applied to the simulation.

These corrections are obtained from the following efficiency measurements,

$$w_{\text{Reco}} = \frac{\epsilon_{\text{Reco}}^{\text{data}}}{\epsilon_{\text{Reco}}^{\text{MC}}}, \quad w_{\text{ID}} = \frac{\epsilon_{\text{ID}}^{\text{data}}}{\epsilon_{\text{ID}}^{\text{MC}}}, \quad w_{\text{Trigger}} = \frac{\epsilon_{\text{Trigger}}^{\text{data}}}{\epsilon_{\text{Trigger}}^{\text{MC}}}, \quad w_{\text{Iso}} = \frac{\epsilon_{\text{Iso}}^{\text{data}}}{\epsilon_{\text{Iso}}^{\text{MC}}} \quad (9.2)$$

The scale factor corrections of Eq. 9.2 are applied on an event-by-event basis as multiplicative event weights. The size of the corrections are sensitive to the event electron E_T and η , but in general they are close to unity. In the case of Drell-Yan, with two electrons in the final-state, two sets of scale factor corrections are applied, one for each electron. Figure 9.5 shows the effect of each correction as a function of dielectron invariant mass m_{ee} .

The ATLAS electron/photon combined performance group (EGP) [166] provides the efficiency scale factors used in the electron channel of the dilepton analysis. Here, the

EGP group identifies the Z boson mass peak as a very clean control region and uses the tag-and-probe [3] method, for example, to estimate electron efficiency scale factors, using increasingly coarse bins in E_T . The scale factors are extracted from measurements performed in two-dimensional bins of E_T and η , which account for changes in efficiency according to the electron energy and ATLAS detector region. Scale factor uncertainties arise due to assumptions about the stability of the measurements out to high- E_T , where electrons cannot be measured precisely (See Sec. 10.8.1). The corrections are applied with the `Electron-EfficiencyCorrection` [167] tool, which is called during the main analysis workflow (See Sec. 10.5.1).

The electron ID efficiency correction is larger than the others depicted in Figure 9.5. This is due to the different behaviors of the “leading,” or higher E_T , and “subleading,” or lower E_T , electrons in the dielectron event, which are used in the construction of the dielectron invariant mass, and due to the assumptions regarding the evolution of the electron shower shapes in E_T and η . As a likelihood-based electron identification is used (See Sec. 8.2.3), the measurements performed to extract the electron ID scale factors are particularly sensitive to any lack of knowledge in the E_T and η kinematic variables.

9.4.3 Energy Scale and Resolution Smearing

The electron energy resolution observed in data differs from what is predicted in simulation, so energy smearing is performed on the simulation to match what is observed in data (See Sec. 8.3.2). This is achieved by smearing the simulated electron energy with Gaussian correction factors, dependent on electron E_T and η . These smearing factors are extracted from dedicated measurements comparing the Z boson width between simulation and data using a sample of $Z \rightarrow e^+e^-$ events.

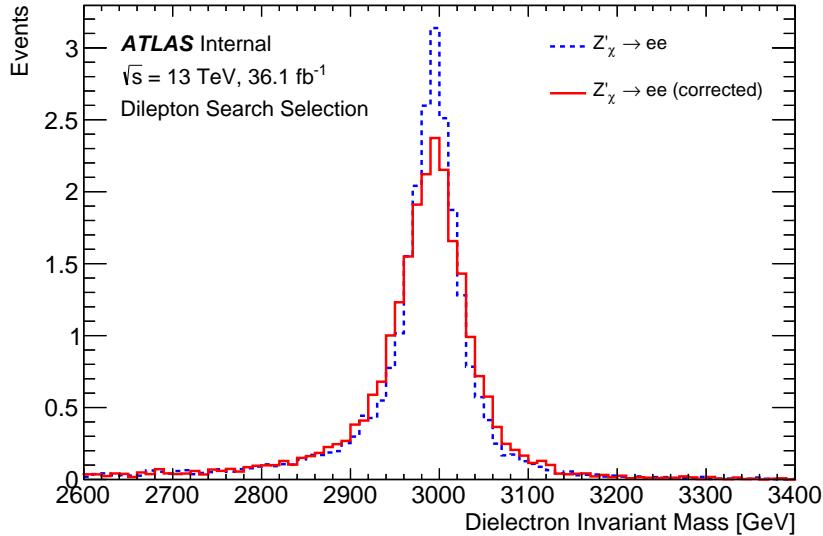


Figure 9.6: Invariant Mass distribution for the $M_{Z'} = 3$ TeV Z'_χ predicted by the simulation with and without the application of the energy scale correction and resolution smearing. The effect is to both shift and broaden the shape of the Z' peak.

In addition, the electron energy scale is calibrated in order to improve the level of agreement between data and simulation. The Z boson mass peak is used as a way to estimate the measured electron energy, since it provides a constraint on the invariant mass of dielectron pairs, and hence the transverse energy E_T of each electron. From this, scale factors can be extracted and used to scale the electron energy. Disagreements between data and simulation can arise due to a number of reasons (See Sec. 8.3.1). The ATLAS EGP group derives these corrections, and includes them in an analysis tool, the `ElectronPhotonFour-MomentumCorrection` tool [146], which is used to smear event-by-event the electron energy in simulation during analysis (See Sec. 10.5.1). Further, it is used to calibrate the electron energy scale in both simulation and data. The invariant mass distribution of the $M_{Z'} = 3$ TeV Z'_χ with and without the application of energy resolution smearing and energy scale corrections can be seen in Figure 9.6. The effect is to shift the Z' peak from its nominal value and broaden its shape.

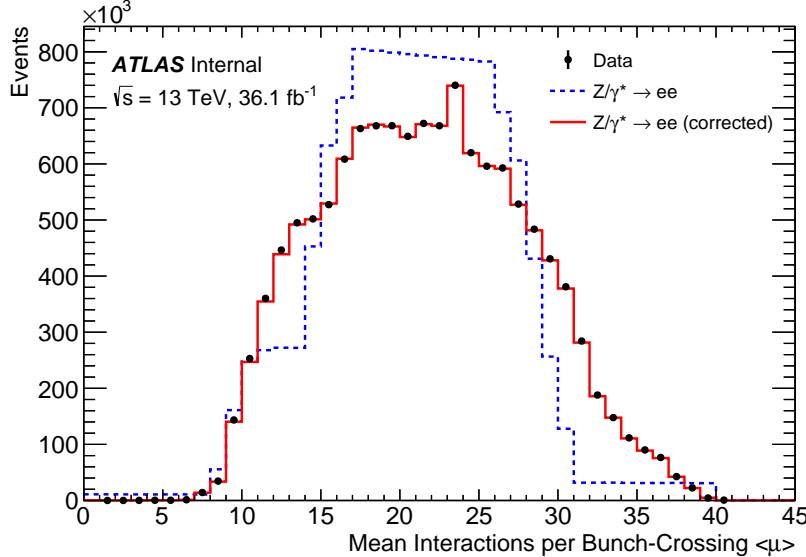


Figure 9.7: Mean number of interactions per bunch crossing in the $Z \rightarrow e^+e^-$ MC sample with and without the pile-up correction as compared to data. The effect of reweighting is to match the pile-up conditions in data.

9.4.4 Pile-Up Profile

The multiple simultaneous proton-proton collisions that occur during a single bunch crossing at the LHC are known as pile-up, as depicted in Figure 6.4. The simulation samples are usually produced before or during a given data-taking period, and while the effect of pile-up is included, the exact conditions that need to be matched to emulate the data are not known *a-priori*. Therefore, only a reasonable estimate of the pile-up conditions in data can be folded into the simulation, which must be corrected for later, once the run conditions have been realized.

Each simulation sample is generated with a pile-up profile around an average number of proton-proton collisions per bunch crossing $\langle\mu\rangle$. The $\langle\mu\rangle$ profile in simulation is reweighted to match the one observed in data using event weights, such that the evolving LHC running conditions are accounted for.

Reweighting is performed using the `PileupReweighting` tool [168], provided by the `Py`

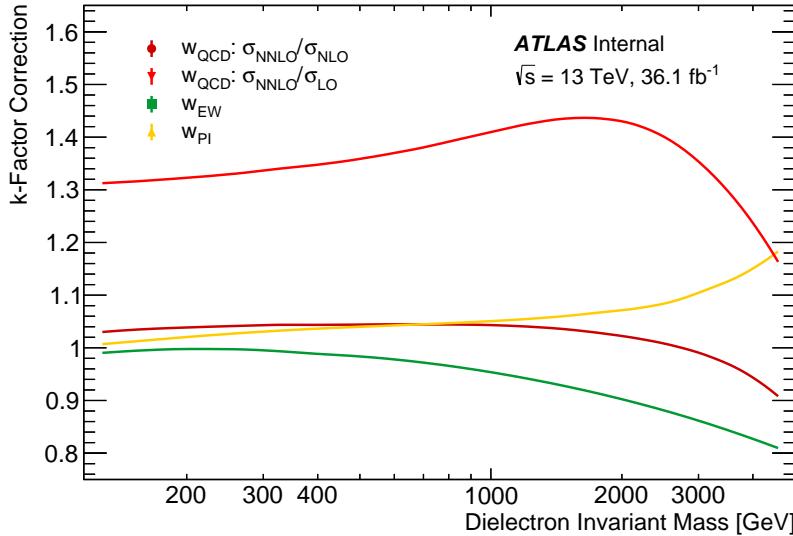


Figure 9.8: The four mass-dependent k-Factor corrections applied in the dilepton analysis. The NLO-to-NNLO QCD correction, LO-to-NLO EW correction, and the LO Photon-Induced QED correction are applied to enhance the description of the Drell-Yan spectrum, while the LO-to-NNLO QCD correction is applied to the Z' signal.

the ATLAS Analysis Software Group (ASG). The result of the reweighting procedure can be seen in Figure 9.7, where the $\langle \mu \rangle$ profile in the dedicated $Z/\gamma^* \rightarrow e^+e^-$ sample is reweighted and compared to data. The overall effect of pile-up on the dilepton search is expected to be small, since the signal is characterized by two high E_T objects, effectively suppressing pile-up interactions.

9.4.5 k-Factor Corrections

Several k-Factor corrections are applied to the simulation in order to improve the nominal descriptions of the Drell-Yan, $t\bar{t}$, and Wt background estimates, as well as the Z' signal estimate. These corrections include higher-order effects, which are absent from the original event generation, and are derived by taking the ratio between a higher-order cross section calculation and a lower-order one.

Both mass-dependent and mass-independent, or global, k-Factor corrections are used in the dilepton analysis. For the NLO Drell-Yan background, the following mass-dependent k-Factor corrections include the NLO-to-NNLO QCD correction, the LO-to-NLO EW correction, and the Photon-Induced QED correction. For the LO Z' signal, only the LO-to-NNLO QCD correction is applied.⁹

The mass-dependent NNLO QCD correction is derived with VRAP 0.9, while the NLO EW correction comes from McSANC. Final-state photon radiation is simulated with PHOTOS, so the EW corrections derived with McSANC include only the contributions from initial-state photon radiation and electroweak loops [4]. The set of mass-dependent corrections is depicted in Figure 9.8.

Similarly, global k-Factor corrections are applied to the $\bar{t}t$ and Wt Single-Top processes, each derived with the TOP++ generator, providing a description at the NNLO in QCD. Corrections to the Diboson processes are not required, due to their relatively small contribution to the overall background estimate.

The LPXKFactorTool [169] incorporates the functional forms of these corrections, and is used to apply them as event weights during the main analysis workflow (See Sec. 10.5.1). Several example weights are listed in Table 9.1 for several values of dielectron invariant mass. Each k-Factor correction comes with an associated systematic uncertainty; these are assessed in Section 10.8.

⁹For the signal, the precise form of EW k-Factor correction depends on the model-dependent couplings between the Z' and the Standard Model W and Z bosons; hence this correction is left out.

Part III

The Analysis

Chapter 10

Electron Channel Dilepton Analysis

In this chapter, the dilepton analysis is introduced, and the search for new physics in dielectron final-states is performed using data taken during the 2015 and 2016 runs of the LHC at $\sqrt{s} = 13$ TeV. This analysis constitutes the contribution to the ATLAS published result by the author.

10.1 Analysis Strategy

The limitations of the Standard Model outlined in Section 5.1 all point indirectly to the existence of physics Beyond the Standard Model (BSM), suggesting some kind of modification or extension to the Standard Model must take place. Many BSM theories extend the Standard Model with additional $U(1)'$ gauge symmetries (as outlined in Sec. 5.3), the breaking of which lead to one or more massive Z' bosons. New Z' states can potentially decay into pairs of charged leptons through the $Z' \rightarrow \ell^+ \ell^-$ decay channel, the signature of which would appear as a narrow resonance, or “bump,” in the dilepton invariant mass spectrum.¹

To that end, a model-independent analysis is developed and conducted in order to search for signatures of new physics, such as the production of a heavy Z' boson, in final-states defined by pairs of same-flavor, opposite-charge leptons, which are produced as a result of proton-proton collisions at the LHC. The analysis is performed in two distinct leptonic

¹The decay width $\Gamma_{Z'}$ is model-dependent and can be calculated with Eq. 5.4.

channels: the **electron channel**, in which the final-state consists of an electron-positron pair, and the **muon channel**, in which the final-state consists of a muon-antimuon pair. Each channel is as its own independent analysis, and later statistically combined to obtain a single dilepton result. The focus of this dissertation, and the following chapter, will be on the search for resonant new physics in the electron channel.

In the electron channel, the dielectron invariant mass is used to discriminate the signal from the Standard Model background processes. This is the most direct way to search for new gauge bosons, as the invariant mass spectrum will display a resonance peak at the pole mass of the new particle. The dielectron invariant mass is built according to

$$m_{e^+e^-} = \sqrt{2E_T^{e^+} E_T^{e^-} (\cosh \Delta\eta - \cos \Delta\phi)}, \quad (10.1)$$

where $E_T^{e^+}$ and $E_T^{e^-}$ are the electron transverse energies, $\Delta\eta$ is the difference in pseudorapidities, and $\Delta\phi$ is the difference in azimuthal angles in the transverse plane of the ATLAS detector.

The general strategy of the analysis is to select events recorded in LHC proton-proton collisions with two good quality, high E_T electron candidates, in order to reconstruct the dielectron invariant mass. The invariant mass distribution in data is then compared to the background expectation provided by the Standard Model, and any observed deviations are statistically quantified.

All signal and background processes are described by Monte Carlo (MC) simulation (See Sec. 9.3), except for the W +jets & QCD Multi-jet fake background, which is obtained through the data-driven **matrix method**. The leading Standard Model background contribution arises from the Drell-Yan process, while subleading backgrounds arise from the Top

$(t\bar{t}, Wt)$, Diboson (WW , WZ , ZZ), and fake ($W+$ jets & QCD Multi-jet) processes.

The **signal region** of the search is defined by the region in the dielectron invariant mass beyond the Z boson mass peak. Two separate analyses took place in the years 2015 and 2016: the first was performed on the complete 2015 dataset [170], while the second, without any major modifications, was performed on the full 2015+2016 combination [5]. The 2015 analysis was initially “blinded,” such that any event with an invariant mass $m_{ee} > 500$ GeV was rejected.²

After it had been shown that the search strategy was sound and a sufficient level of agreement had been obtained between the data and the MC simulation below the 500 GeV threshold, the analysis was “unblinded” with the approval of the ATLAS collaboration. During this time, this author developed an Early-Warning-System (EWS) [171] (See Appendix E), which ran on slimmed-down datasets produced at the Tier-0 computing facility at CERN, in order to alert the analysis team of the development of any anomalous signatures as data were being taken, which could facilitate a faster unblinding if necessary.³

The dilepton analysis provides excellent sensitivity to signatures of new physics owing to its clean, fully reconstructed final-state, high signal-efficiency, and set of small well-understood backgrounds [5]. For these reasons, it is one of the earliest searches to be conducted at newly commissioned hadron colliders. This chapter now presents the 2015+2016 analysis.

²In a blind analysis, the final result is kept hidden (in this case, the observed invariant mass spectrum above a certain mass threshold) until the analysis is essentially complete. The idea is to avoid any “subconscious” experimental bias, such as modifying the event selection to enhance a potential signal.

³The search sensitivity of the Run-1 dilepton analysis would be quickly surpassed with only 1-2 fb^{-1} of Run-2 $\sqrt{s} = 13$ TeV data. The rate of data collection in early Run-2 was so fast that the integrated luminosity would double every few weeks, so a signal really could “jump out” from one week to the next. This is the importance of the EWS, especially since the performance of the main analysis and subsequent unblinding could take months.

10.2 Data Samples

The dilepton analysis uses proton-proton collision data collected at the LHC during the 2015 and 2016 data-taking periods at $\sqrt{s} = 13$ TeV. The total integrated luminosity corresponds to 36.1 fb^{-1} , as recorded by the ATLAS detector after excluding data-taking periods where one or more subdetectors were not fully operational.

10.2.1 ATLAS Data

ATLAS data-taking is divided into **periods**. Periods usually last several weeks, and are classified according to consistent configurations of the detector and trigger. Significant changes either to the configuration of the detector or to the trigger usually result in a new period. Periods are listed alphabetically with a single letter, and consist of specific data-taking **runs**.

Each run is assigned a unique run number, and usually corresponds to a single fill of the LHC, which lasts for multiple hours.⁴ Runs are further broken down into **luminosity blocks**, collections of data recorded over the course of a few minutes, during which time the instantaneous luminosity, and detector conditions, are approximately constant.

Luminosity blocks are tabulated in a file called the Good Runs List (GRL) [172]. A luminosity block is flagged as “good” in the GRL if certain data quality requirements are met, such as the stability of LHC beams, the absence of any hardware malfunctions, and that relevant ATLAS subsystems are operational. Datasets usable for LHC physics analysis consist only of runs and luminosity blocks found within a GRL.

⁴Some of the longer runs have lasted for slightly more than a day.

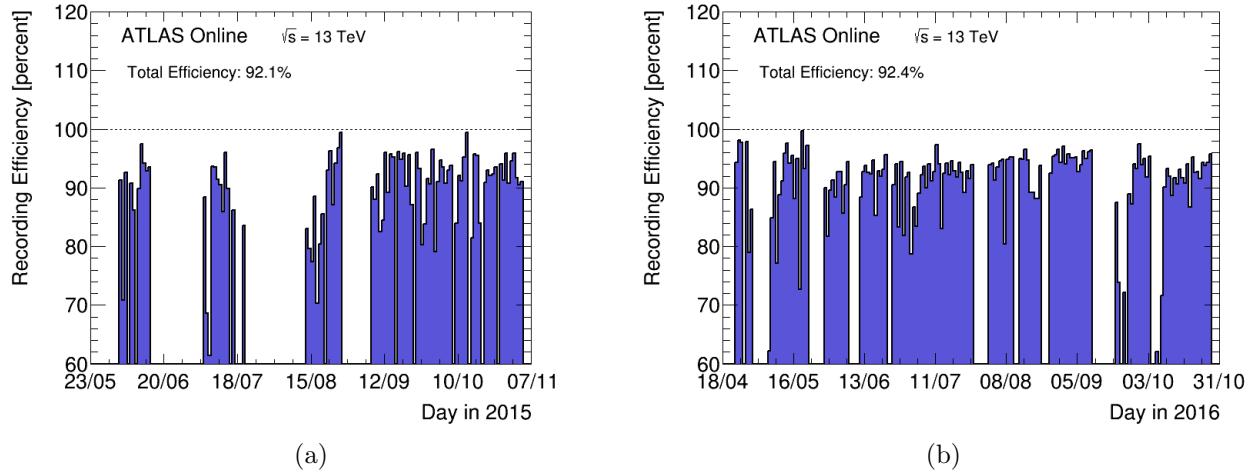


Figure 10.1: Day-by-day data-taking efficiency of the ATLAS detector during the 2015 (a) and 2016 (b) runs. The efficiency over the course of each run was above 92%. Empty bins correspond to days during which the LHC did not deliver stable beams [18].

10.2.2 Analysis Data

The search uses LHC proton-proton collisions at $\sqrt{s} = 13$ TeV collected in the years 2015 and 2016, as recorded by the ATLAS detector. The 2015-2016 dataset is comprised of events recorded from August through October 2015,⁵ and from April through October 2016 [173]. During these periods, the LHC delivered a total integrated luminosity of 42.7 fb^{-1} . For the 2015 run alone, this corresponded to 4.2 fb^{-1} , with a total of 3.9 fb^{-1} recorded by the ATLAS detector (See Fig. 6.5). Likewise, in the following year, the LHC delivered 38.5 fb^{-1} , and of that, the ATLAS detector recorded 36.5 fb^{-1} (See Fig. 6.5).

Among all luminosity blocks recorded, only those identified in a GRL as “good” were analyzed. Two GRLs were used in the analysis, one for each year of data-taking. Thus, the final dataset consisted of an integrated luminosity of 36.1 fb^{-1} usable for analysis, of which 3.2 fb^{-1} belonged to 2015, and 32.9 fb^{-1} to 2016. A 3.2% luminosity uncertainty was

⁵In June and July of 2015, the LHC operated with a bunch-spacing of 50 ns. In August, it was subsequently reduced to its design value of 25 ns. Only data taken with the 25 ns bunch spacing is analyzed, since the integrated luminosity of the 50 ns run is small and would otherwise require entirely new MC samples.

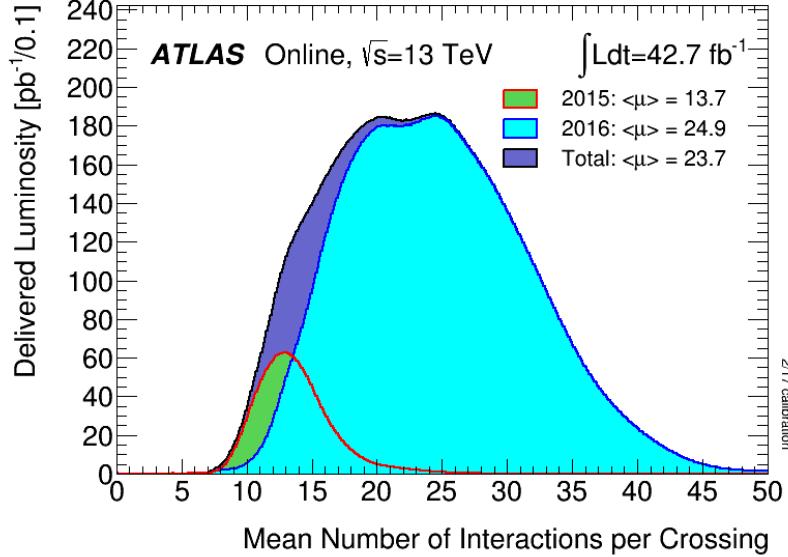


Figure 10.2: Average number of proton-proton interactions per bunch crossing during the 2015 and 2016 data-taking periods. The data shown correspond to periods with stable beams delivered to the ATLAS detector [18].

assessed on the final dataset, as recommended by ATLAS luminosity group [174], by taking a weighted average of the luminosity uncertainty of each year. A complete list of data-taking periods used for analysis can be found in Appendix A.

The data-taking efficiency of the ATLAS detector can be seen in Figure 10.1. For the 2015 run, the average efficiency was 92.1%, which was roughly the same in 2016 at 92.4%, despite the average number of pile-up interactions almost doubling (See Sec. 6.3.1). The average number of pile-up interactions was $\langle \mu \rangle = 13.7$ in 2015, which increased to $\langle \mu \rangle = 24.0$ in 2016. The average over both years was $\langle \mu \rangle = 23.7$, and the associated pile-up distributions can be seen in Figure 10.2.

10.2.3 Dataset Derivations

Many physics analyses, including this one, produce slimmed down datasets associated with the particular final-state under study. These **dataset derivations** remove otherwise ex-

traneous event or particle level information from the AOD event record, which eliminates overhead, streamlines the analysis, and speeds up the turnaround time for common analysis tasks. This can be especially useful, since the data and simulation need to be run over frequently.⁶

In Run-1, the situation regarding derivations was untenable. The AOD files (See Sec. 9.2.3) were not ROOT [175] compatible, and therefore had to be converted into a series of intermediate ROOT-readable formats.⁷ The production and formatting of these “n-tuples” was left to each analysis team, and without any quality control, a large number of incompatible formats resulted, making it difficult to perform cross-checks, or validate results obtained by other analysis teams working on similar channels.

The Run-2 event data model [176] (EDM) solves this by uniformizing all group data-formats. The AOD is replaced by a ROOT-readable format from the start, called the xAOD, and derived datasets, DxAOEs, are produced centrally at CERN from xAOEs during event reconstruction (See Sec. 9.2.3) [177].

The DxAOEs produced for the dilepton analysis are designed to retain events with two or more reconstructed electrons with $E_T > 20$ GeV or muons with $p_T > 20$ GeV. In addition, the electron has to satisfy the LOOSE likelihood-based identification criterion (See Sec. 8.2.3). All data and simulation samples analyzed in the analysis are produced in this derived format.

As a result, the physical size of the DxAOEs is reduced to approximately 1% the size of the input xAOEs. This reduction in size was large enough to allow for both the data and simulation samples to be stored locally on the MSU Tier-3 [178] computing cluster, on which

⁶This occurs for a variety of reasons: to conduct performance studies, to study new kinematic variables, to check the effect of changes to the selection, if bugs are found, if updated CP recommendations are released, if running systematic studies, after a bulk-reprocessing of the data, etc.

⁷ROOT is statistical analysis framework commonly used in HEP to process and visualize data, usually in the form of histograms.

the analysis was performed by this author.

10.3 Monte Carlo Samples

As outlined in the previous chapter, simulation samples provide a way to estimate the Standard Model background expectation, and therefore allows for a direct comparison to be made against events selected in data. The Standard Model backgrounds in the electron channel arise from processes with two real final-state electrons, and are outlined in Section 9.3. Background samples are used to model events arising from the Drell-Yan process, Photon-Induced production, Top (Wt -channel Single-Top and $t\bar{t}$) production, and Diboson (WW , WZ , and ZZ) production.

For each sample, the ATLAS detector response is simulated with GEANT4, and all events are reconstructed with the same reconstruction software that is used for data (See Sec. 9.2); and like the data, all MC samples are built as DxAOD derivations. The simulation is corrected for a variety of effects. The electron energy resolution is smeared (See Sec. 9.4.3), event level reweighting is performed to account for pile-up effects (See Sec. 9.4.4), and scale factors are applied to correct for differences observed in the reconstruction, identification, trigger, and isolation efficiencies between the MC samples and data (See Sec. 9.4.2).

10.4 Analysis Software

In order to conduct the analysis, a dedicated framework is used to “run over”⁸ the proton-proton collision events stored in the DxAOD derivations. The analysis framework houses the event selection (to be defined in Sec. 10.5.1), and interfaces with the various CP tools, which

⁸i.e. load, analyze, and accept or reject

are used to apply the required event and object level corrections to data and simulation, and to assess systematic uncertainties.

The dilepton analysis uses the `ROOTCORE` framework [179], which is provided and maintained by the ATLAS collaboration. Using `ROOTCORE`, data and MC samples are loaded with the `SampleHandler` [180] package, and processed using `EventLoop` [181]. The analysis framework is built within a particular **Analysis Base Release** (ABR) [182]. The ABR contains specific releases of general purpose physics analysis tools, which facilitate coordination between analysis groups, and ensures projects are efficiently organized and up to date. The development and deployment of ABRs is handled by the ATLAS Software Group [183].

The results of the dilepton analysis were obtained under ABR version 2.4.27, the most up-to-date version at the time of publication, which contains the following physics analysis tool versions:

- `ElectronPhotonFourMomentumCorrection-02-03-00`: package used to perform electron energy calibration (See Sec. 8.3.1) and resolution smearing (See Sec. 9.4.3)
- `ElectronEfficiencyCorrection-00-01-88`: package used to apply efficiency scale factor corrections to simulation (See Sec. 9.4.2)
- `PileupReweighting-00-04-01`: package used to correct the pile-up profile in simulation (See Sec. 9.4.4)
- `IsolationCorrection-00-01-27`: package used to correct electron candidate isolation energy
- `ElectronPhotonSelectorTools-00-02-92-18`: package used to perform likelihood-based electron identification (See Sec. 8.2.3)

- `IsolationSelection-00-06-05`: package used to calculate the isolation energy of electron candidates (See Sec. 8.2.4)

In addition, packages specific to the dilepton analysis include:

- `LPXKfactorTool-00-01-05` [169]: package used to apply higher-order k-Factor corrections to simulation (See Sec. 9.4.5), as well as to normalize signal and background samples with the appropriate cross sections (See Sec. 9.4.1)
- `LPXSignalReweightingTool-00-00-07` [184]: package used to create Z' signal templates from LO Drell-Yan simulation for use in the statistical analysis (See Sec. 11.2)

10.5 Event Selection

The electron channel search targets LHC proton-proton events with two or more high- E_T electron candidates. For an event to be considered a *high-mass dielectron candidate event*, each event electron must pass certain quality requirements, defined by the dielectron **event selection** in the proceeding section. These requirements enhance the probability of selecting a real electron, rather than a misidentified jet or an electron from a photon conversion. Events that retain at least two electrons passing selection are then used to calculate the dielectron invariant mass according to Eq. 10.1.

The same set of selection criteria are applied equally to events in data and simulation, including both the signal and background MC. Background events passing the selection are subsequently scaled to the integrated luminosity observed in data (as in Sec. 9.4.1), and combined to provide an estimate for the Standard Model background expectation, which is then used to model the observed event yield in data. This allows for a direct comparison to

be made between the data, the overall Standard Model expectation, and the various BSM predictions, allowing for a conclusion to be drawn regarding the presence of a signal.

10.5.1 Selection Criteria

A set of selection criteria is applied to obtain a high quality sample of high-mass dielectron events from data and simulation. Events passing the selection are retained for the subsequent statistical analysis (See Sec. 11.1). The selection criteria are applied first at the “event level,” then at the “object level,” and finally at the dielectron “pair level,” moving from broad to narrow in scope. As a preliminary step, the DxAOD preselection is applied to skim the data and MC samples used in the analysis (See Sec. 10.2.3). The event selection starts at the event level with the following requirements:

GRL: Data events must originate from luminosity blocks identified as good by the GRL, ensuring the LHC and ATLAS detector are operating under stable conditions (See Sec. 10.2.1). A GRL requirement is not placed on the simulation, since it is generated with nominal detector and beam conditions.

Trigger: Events must pass the `HLT_2e17_lhloose` trigger, which requires at least two electrons to satisfy the LOOSE likelihood-based identification requirement, each with a minimum transverse energy of $E_T > 17$ GeV.

Event Cleaning: Event cleaning removes incomplete or corrupt events. The GRL removes entire luminosity blocks affected by detector related problems. To provide additional discrimination, event level detector flags exist to reject events that are incomplete, contain corrupt data, or are inconsistent with LAr, Tile, and SCT error requirements.

Two Electrons: At least two electron candidates must originate from the event primary vertex (See Sec. 8.1.2). All electrons considered are reconstructed according to the standard ATLAS algorithm (See Sec. 8.2.2), which reconstructs energy depositions in the EM calorimeter, and attempts to match each with an single inner detector track.

Next, the object level selection is applied to select events with at least two good electron candidates. A set of selection criteria is applied to each electron in the event. If an electron fails a cut at any stage in the proceeding selection, then it is excluded from the list of candidate good electrons. If the number of candidates falls below two, then the event is rejected from the analysis. At this stage, the energy scale calibrations (See Sec. 8.3.1) and energy resolution smearing (See Sec. 8.3.2) are applied to the individual electron candidates in data and MC, respectively. The object level selection is:

Pseudorapidity: The electron candidate must be reconstructed from an energy cluster in the central region ($|\eta| < 2.47$) of the ATLAS detector. Additionally, electrons in the transition region, $1.37 < |\eta| < 1.52$, between the barrel and end-cap EM calorimeters are rejected, due to a degraded energy resolution in this region.

Object Quality: The quality of the cluster associated with an electron candidate is checked. The object quality requirement rejects candidates with bad quality or fake clusters resulting from calorimeter problems, such as malfunctioning hardware, LAr noise bursts, or data integrity errors [185].

Transverse Energy: An electron candidate must have a transverse energy $E_T > 30$ GeV. The E_T cut is chosen to ensure that the `HLT_2e17_lhloose` trigger is fully efficient.

d_0 Significance: The track associated with an electron candidate must have a d_0 significance σ_{d_0} that satisfies $d_0/\sigma_{d_0} < 5$ (See Sec. 8.1.2). This requirement helps to select

prompt electron candidates that originate from the event primary vertex, and not from secondary vertices, or photons conversions.

Longitudinal Impact Parameter: Similarly, the track associated with an electron candidate must have a longitudinal impact parameter z_0 with respect to the reconstructed primary vertex, satisfying $|\Delta z_0 \sin \theta| < 0.5$ mm (See Sec. 8.1.2).

Electron Identification: Each electron candidate must satisfy the MEDIUM likelihood-based electron identification criterion (See Sec. 8.2.3), which provides discrimination based on longitudinal and lateral shower shape variables, tracking variables, and cluster-track matching quantities (See Table 8.1). The electron identification requirement serves to reduce background processes that fake the electron signature.

Isolation: Each electron candidate must satisfy the LOOSE isolation criterion, which requires that the energy deposition in the EM calorimeter and inner detector track matched to the electron candidate both be well isolated (See Sec. 8.2.4). The isolation requirement further suppresses backgrounds originating from non-prompt decays and misidentified hadronic jets.

The final stage of selection focuses on the construction of the dielectron object:

Pair Formation: If two or more electron candidates are retained after the object level selection, then the two with the highest transverse momentum E_T are selected to form a dielectron pair.

Dielectron Invariant Mass: The invariant mass of the pair is calculated with Eq. 10.1 and must satisfy $m_{ee} > 80$ GeV.

Events that satisfy all of the preceding requirements are retained for analysis.⁹

10.5.2 Cutflows

Tables 10.1 and 10.2 list the number of events in the full 36.1 fb^{-1} dataset and the simulated Z'_χ signal sample, respectively, that pass the dielectron event selection of Section 10.5.1. These **cutflow tables** summarize the number of events that pass each stage of selection, along with their relative and absolute selection efficiencies.

There are initially 82,684,764 events in data, and after applying the selection, 12,716,634 candidate events remain in use for analysis.¹⁰ An identical set of selection criteria is applied to the signal and background simulation, in order to make a meaningful comparison with the data.¹¹ Event and object level corrections are applied to the MC to enhance its accuracy. All corrections described in Section 9.4 are included, and their combined effect on the resulting number of Z'_χ events can be seen in the final row of Table 10.2, denoted as “Event Weights.”

The MC simulation is normalized to an integrated luminosity of 36.1 fb^{-1} to match the observed value in data. The resulting signal acceptance is approximately 71% for a 3 TeV Z'_χ signal. Additional cutflow tables for the Drell-Yan, Top, and Diboson background processes can be found in Appendix B. The total background rejection for the Top and Diboson backgrounds is approximately 98% and 88%, respectively.

⁹In the muon channel, an opposite-sign charge requirement is placed on dimuon candidates. This is not required in the electron channel, however, due to the possible charge misidentification of high- E_T electrons, resulting from either bremsstrahlung, or from the limited momentum resolution associated with high- p_T tracks in the Inner Detector.

¹⁰In fact, considering the dataset was initially skimmed, the number of events *before* preselection totaled 6,327,792,063. Consequently, out of the entire 36.1 fb^{-1} dataset, only approximately 0.2% of LHC events were selected for use in the electron channel analysis.

¹¹The GRL requirement is not included, however, since the MC is generated under nominal beam and detector conditions.

Selection Cut	Number Events	Relative Eff. [%]	Absolute Eff. [%]
Post-selection	82 684 764	100.00	100.00
GRL	80 525 303	97.39	97.39
Trigger	24 639 061	30.60	29.80
Event Cleaning	24 601 393	99.85	29.75
2 Electrons	24 601 392	100.00	29.75
η	23 784 390	96.68	28.77
Object Quality	23 722 173	99.74	28.69
E_T	16 035 333	67.60	19.39
d_0 Sig.	15 774 118	98.37	19.08
z_0	15 597 169	98.88	18.86
Identification	13 756 703	88.20	16.64
Isolation	13 241 876	96.26	16.01
Invariant Mass	12 716 634	96.03	15.38

Table 10.1: The number of events passing the electron channel event selection in the 36.1 fb^{-1} dataset obtained by ATLAS during the 2015+2016 physics runs. The first column indicates the absolute number of events passing selection, the middle column depicts the relative efficiency of each cut, and the last column shows the absolute efficiency.

Selection Cut	Number Events	Relative Eff. [%]	Absolute Eff. [%]
Post-selection	38.48	100.00	100.00
GRL	38.48	100.00	100.00
Trigger	33.12	86.08	86.08
Event Cleaning	33.12	100.00	86.08
2 Electrons	33.12	100.00	86.08
η	32.54	98.23	84.56
Object Quality	32.50	99.87	84.45
E_T	30.55	94.02	79.40
d_0 Sig.	30.48	99.75	79.20
z_0	30.40	99.74	78.99
Identification	28.64	94.21	74.42
Isolation	28.17	98.36	73.20
Invariant Mass	28.16	99.99	73.19
Event Weights	27.27	96.82	70.86

Table 10.2: The number of events passing the electron channel event selection in a sample of simulated Z'_χ events with pole mass $M_{Z'} = 3 \text{ TeV}$. The sample has all corrections applied that are described in Section 9.4, and is normalized to an integrated luminosity of 36.1 fb^{-1} according to the procedure outline in Section 9.4.1.

10.5.3 Signal Acceptance

The pole mass $M_{Z'}$ of a hypothetical Z' boson is generally a free parameter of the new physics model, as is the case for the models considered in this analysis. Therefore, it is important to quantify the performance of the dielectron event selection as a function of the *a-priori* unknown Z' pole mass, as the kinematics and reconstruction properties of the decay electrons are sensitive to the precise value of $M_{Z'}$.

The signal **acceptance times efficiency** is plotted as a function of $M_{Z'}$ in Figure 10.3.¹² Here, the dielectron event selection is applied to individual Z' signal templates, ranging from $M_{Z'} = 150$ GeV to $M_{Z'} = 5000$ GeV in 50 GeV intervals, and the resulting total efficiency for each is quantified. A smooth polynomial curve is fit through each resulting point, providing an estimate of the total efficiency within the 50 GeV intervals.

For a Z'_χ with $M_{Z'} = 150$ GeV, the total efficiency is approximately 32%, which quickly rises to 71% for the 3 TeV Z'_χ (as can also be seen in Table 10.2), before slowing dropping off again at high mass. The quick “turn-on” at low mass is due to the high- E_T and central- η requirements in the dielectron selection. At higher pole masses the Z' is more likely to be produced at rest, and as a result, its decay electrons tend to be more central, leading to additional events passing the η requirement.

However, above about 2.5 TeV in pole mass, the efficiency starts to drop again due to the strong decrease in parton luminosity at high momentum transfer. The colliding quarks have insufficient energy to produce on-shell Z' events, and as a result, the relative fraction of events in the low-mass tail of the spectrum becomes enhanced (See Fig. 10.4). The electrons in the tail are again less central, causing a drop in acceptance.

¹²The acceptance times efficiency $\mathcal{A}\epsilon$ is the number of selected dielectron candidate events divided by the total number of generated events; sometimes also called the total efficiency, or just the efficiency.

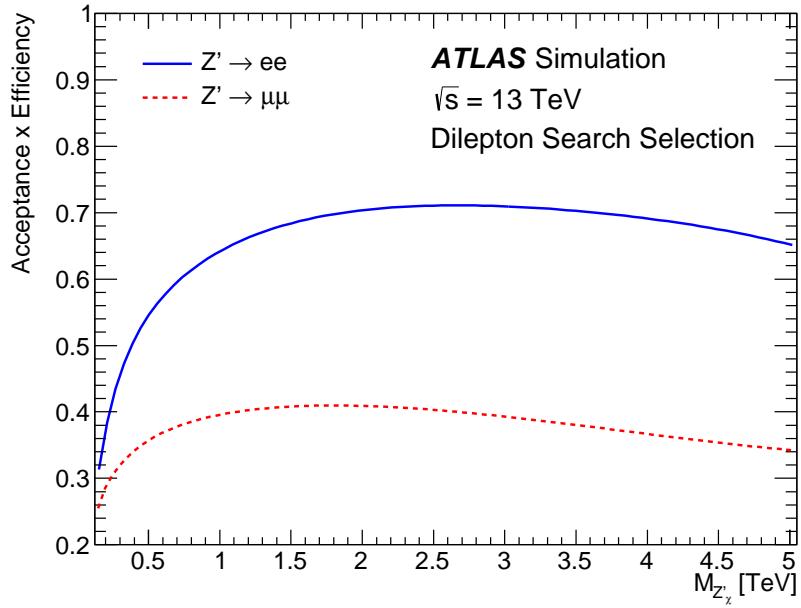


Figure 10.3: The Z' signal efficiency, according to the dielectron selection defined in Section 10.5.1, as a function of Z'_χ pole mass. The selection is maximally efficient at high mass, the primary target of the search. The quick “turn-on” at low mass and slow “turn-off” at high mass are due to the electron kinematics at each of these scales.

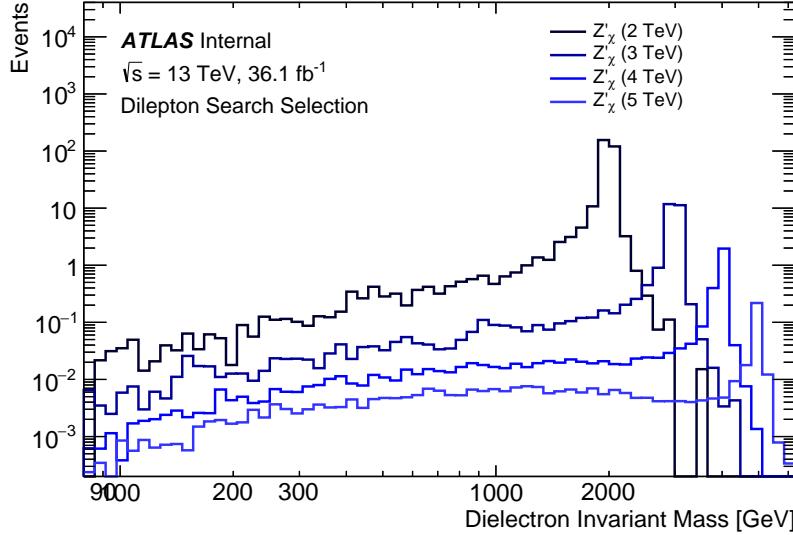


Figure 10.4: Z'_χ signals post-selection for several characteristic mass points. A low-mass tail begins to develop as the kinematic limit is reached. The number of events within specific invariant mass intervals is displayed in Table 10.3 for the 4 and 5 TeV pole mass Z'_χ .

10.6 Fake Background Estimate

The dielectron event selection is robust, greatly reducing the probability of selecting events with non-prompt, or fake, electrons. However, because of the overwhelming production rate of Multi-jet events at the LHC, a small but non-negligible number of fake events pass the dielectron event selection. Therefore, this “fake” background exists as a component of the selected data, and must be properly accounted for in the modeling the overall dielectron background estimate.

The fake background consists of events that originate primarily from two sources: $W + \text{jets}$ production, in which a misidentified jet and electron pass selection, and QCD Multi-jet production, in which two misidentified jets pass selection. Due to the relatively low probability of this happening, it is prohibitive to generate MC simulation samples with sufficient statistics to model this background accurately. Instead, a fully data-driven approach is taken, known as the matrix method [186, 187]. Here, the fake background is fully estimated from a sample of events obtained directly from the 36.1 fb^{-1} dataset used in the analysis.

10.6.1 The Matrix Method

In the matrix method, two comparable event selections are defined: “loose” and “tight.” The tight selection coincides with the nominal dielectron event selection of Section 10.5.1. The loose selection is defined relative to the tight selection, where the electron identification requirement is relaxed from the MEDIUM to the LOOSE operating point, and the isolation requirement removed entirely. Therefore, the set of objects passing the tight selection forms a subset of those passing the loose selection. The idea is to measure the electron efficiency in each selection in order to identify the signal contamination from misidentified jets.

Accordingly, with two unique selections and two final-state electrons, four distinct observables can be defined that measure the number of dielectron pairs passing the tight and loose selections, respectively. These observables are N_{TT} , N_{TL} , N_{LT} , N_{LL} , where, the first (second) index refers to the leading (subleading) electron,¹³ and the L subscript refers to an electron passing the loose selection, but failing the tight selection, while the T subscript refers to an electron that passes both.

Similarly, for any dielectron kinematic distribution, the true background expectation can be decomposed into four components, which are unmeasurable and consist of distinct combinations of real and fake electrons. These components are N_{RR} , N_{RF} , N_{FR} , N_{FF} , where, again, the first and second indices refer to the leading and subleading electron candidates, and the R and F subscripts define whether the electron candidate is real or fake. One of the four component is due entirely to real electrons, while the other three account for the fake combinations.

It is not possible to estimate these true quantities directly, but it is possible to construct a system of equations that relates the experimentally measurable quantities to the unmeasurable ones:

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} r^2 & rf & fr & f^2 \\ r(1-r) & r(1-f) & f(1-r) & f(1-f) \\ (1-r)r & (1-r)f & (1-f)r & (1-f)f \\ (1-r)^2 & (1-r)(1-f) & (1-f)(1-r) & (1-f)^2 \end{pmatrix} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix}. \quad (10.2)$$

Here, the **real-rate** r and **fake-rate** f provide the connection between the measurable and

¹³Of the selected electron pair, “leading” denotes the one with the highest E_T , while “subleading” refers to the one with the second highest E_T .

unmeasurable terms. The real-rate r (fake-rate f) is the probability that a real (fake) electron selected by the loose selection will also be selected by the tight selection.

If r and f can be estimated, then the fake background contribution to the electron channel can be identified, reading off the matrix entries, as

$$N_{TT}^{\text{Multi-jet \& W+jets}} = rfN_{RF} + frN_{FR} + f^2N_{FF} \quad (10.3)$$

which, as required, is in terms of dielectron pairs with at least one fake electron. The final step of the method is to re-express the unmeasurable terms on the right hand side of Eq. 10.3 in terms of the measurable ones, which can be obtained simply by inverting the matrix of Eq. 10.2.

10.6.2 Real- and Fake-Rate Determination

The real-rate and fake-rate are defined as

$$r = \frac{N_{real}^{tight}}{N_{real}^{loose}}, \quad f = \frac{N_{fake}^{tight}}{N_{fake}^{loose}}, \quad (10.4)$$

where the number of electrons passing the tight selection N_{tight} forms a subset of those passing the loose selection N_{loose} . In general, the real-rate and fake-rate of Eq. 10.4 are functions of electron E_T and $|\eta|$. In order to obtain an estimate of these rates, **enriched samples** of real and fake electrons are needed.

The enriched sample of real electrons can be easily obtained from the Drell-Yan MC simulation, since by definition its events contain two real electrons. The real-rate r is estimated to be greater than 90% for $E_T > 30$ GeV and increases with E_T . When binned as a function

of $|\eta|$, the rate drops by approximately 2%, moving from low to high $|\eta|$.

An enriched sample of fake electrons is obtained directly from data. Here, a special selection is defined, which combines several single electron triggers that incorporate the likelihood-based LOOSE or vLOOSE electron identification requirements.

To mitigate contamination from real electrons in the fake enriched sample, three further selection criteria are imposed. First, an $\cancel{E}_T > 25$ GeV requirement (See Sec. 7.2) is used to veto events that originate from $W \rightarrow e\nu$ decays. Second, in order to suppress events that originate from the Drell-Yan process, a veto is placed on events with two electrons that satisfy likelihood-based MEDIUM identification criterion. Lastly, the requirement $|M_Z - m_{ee}| > 20$ GeV is used to veto events within 20 GeV of the Z boson mass peak.

Any remaining contamination from real electrons is estimated using the Drell-Yan, Top, and Diboson MC samples. Whatever remains of the MC post-selection is subtracted from the data to finally obtain the fake enriched sample, which is used to determine N_{fake}^{tight} and N_{fake}^{loose} . For $E_T > 30$ GeV, the fake rate is estimated to be between approximately 16% and 34% depending on the $|\eta|$ and E_T regions considered.

10.6.3 Background Extrapolation

The fake background estimate is obtained from the sample of events defined by Eq. 10.3. The invariant mass of each dielectron pair m_{ee} in the sample is calculated and binned in a histogram, providing the fake background estimate to be used in conjunction with the MC for the analysis.¹⁴ However, at this stage, the resulting background estimate only applies above the region of the Z boson mass peak, due to veto applied in the fake-rate derivation,

¹⁴With a sample of fake background events, kinematic variables in addition to the invariant mass can be extracted, the results of which can be seen in Figures 10.7, 10.8, and 10.9.

and runs out of statistics at high mass.

In order to obtain coverage over the entire dielectron search region, the estimate is extrapolated both down to low mass, below the Z boson mass peak, and out to high mass using a series of fits to the m_{ee} distribution. The functional form of the fit is not known *a priori*, but the so-called “dijet function” has been used successfully in previous ATLAS dijet searches [188, 189], and is therefore similarly used here. The form is given by

$$f(x) = p_1(1-x)^{p_2}x^{p_3+p_4\log(x)+p_5\log^2(x)}, \quad x = \frac{m_{ee}}{\sqrt{s}} \quad (10.5)$$

where p_1 through p_5 are free parameters, allowing for flexibility in the resulting fit.

To account for any bias in the choice of fit range, the starting point is varied from 125 GeV to 195 GeV in steps of 10 GeV. This results in eight individual fits, and for each, the χ^2/ndf falls between 1.2 and 1.6. The resulting central estimate is taken to be the χ^2 weighted mean of the set of fits, and the fit uncertainty is characterized by the resulting envelope.

The final result can be seen in Figure 10.5. Systematic uncertainties related to the matrix method technique, and to the fit uncertainty, are included as part of the error estimate.

10.7 Comparison of Data to the Standard Model Background Expectation

In the following section, several direct comparisons are made between the observed data, selected from the 36.1 fb^{-1} LHC dataset, and the full Standard Model expectation. All non-negligible backgrounds are taken into account for the comparison.

The Standard Model expectation is obtained by applying the dielectron event selection

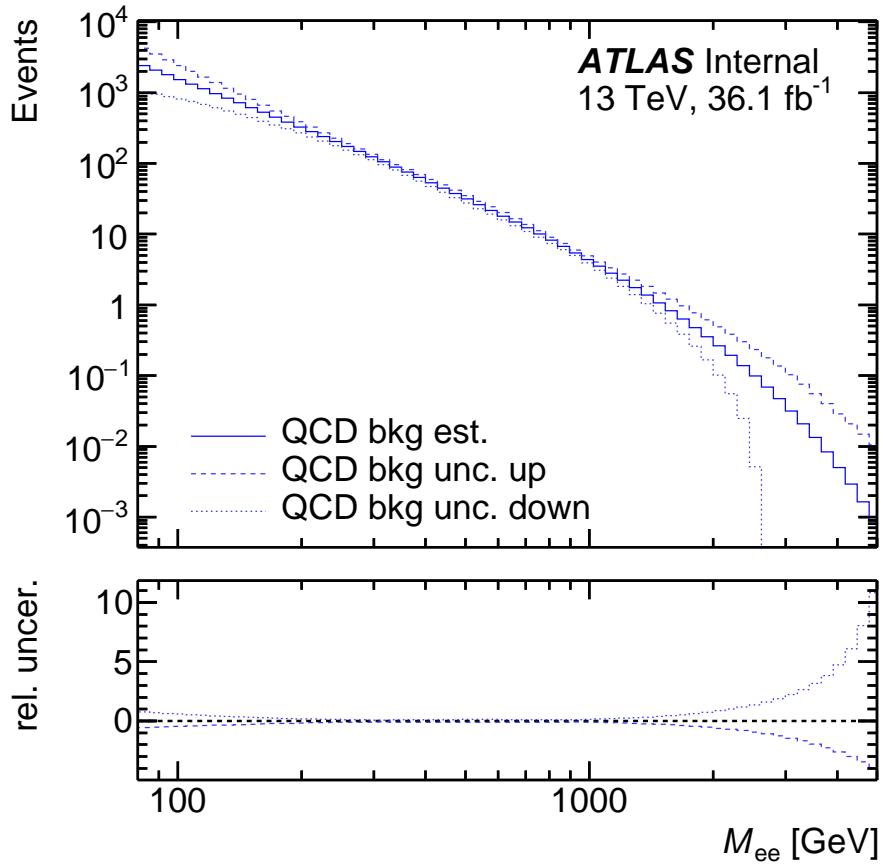


Figure 10.5: The data-driven fake background estimate used in the dielectron channel of the dilepton analysis. The estimate is obtained with the matrix method, following the technique outlined in Section 10.6 [4].

(See Sec. 10.5.1) to the Drell-Yan, Top, and Diboson MC samples. The resulting kinematic distributions are normalized (See Sec. 9.4.1) and summed together. The inclusion of the fake background in the sum completes the estimate.

10.7.1 Event Yields

The expected and observed **event yields** are presented in Table 10.3 in several bins of dielectron invariant mass. The expected event yields are subdivided into the various background processes, and the Z'_χ signal process, with a pole mass of 3 and 4 TeV, is presented as well

m_{ee} [GeV]	80–120	120–250	250–400	400–500	500–700
Drell–Yan	$11\,800\,000 \pm 700\,000$	$216\,000 \pm 11\,000$	$17\,230 \pm 1000$	2640 ± 180	1620 ± 120
Top quarks	$28\,600 \pm 1800$	$44\,600 \pm 2900$	8300 ± 600	1130 ± 80	560 ± 40
Dibosons	$31\,400 \pm 3300$	7000 ± 700	1300 ± 140	228 ± 25	146 ± 16
Fakes	$11\,000 \pm 9000$	5600 ± 2000	780 ± 80	151 ± 2	113 ± 17
Total SM	$11\,900\,000 \pm 700\,000$	$273\,000 \pm 12\,000$	$27\,600 \pm 1100$	4150 ± 200	2440 ± 130
Data	12 415 434	275 711	27 538	4140	2390
Z'_χ (4 TeV)	0.00635 ± 0.00021	0.0390 ± 0.0015	0.0564 ± 0.0025	0.0334 ± 0.0027	0.064 ± 0.004
Z'_χ (5 TeV)	0.00305 ± 0.00012	0.0165 ± 0.0006	0.0225 ± 0.0010	0.0139 ± 0.0007	0.0275 ± 0.0015
m_{ee} [GeV]	700–900	900–1200	1200–1800	1800–3000	3000–6000
Drell–Yan	421 ± 34	176 ± 17	62 ± 7	8.7 ± 1.3	0.34 ± 0.07
Top quarks	94 ± 8	27.9 ± 2.8	5.1 ± 0.7	< 0.001	< 0.001
Dibosons	39 ± 4	16.9 ± 2.1	5.8 ± 0.8	0.74 ± 0.11	0.028 ± 0.004
Fakes	39 ± 6	16.1 ± 2.0	7.9 ± 2.3	1.6 ± 1.2	0.08 ± 0.27
Total SM	590 ± 40	237 ± 17	81 ± 7	11.0 ± 1.8	0.45 ± 0.28
Data	589	209	61	10	0
Z'_χ (4 TeV)	0.0585 ± 0.0035	0.074 ± 0.005	0.121 ± 0.011	0.172 ± 0.017	2.57 ± 0.27
Z'_χ (5 TeV)	0.0218 ± 0.0013	0.0295 ± 0.0021	0.040 ± 0.004	0.040 ± 0.004	0.280 ± 0.030

Table 10.3: Expected and observed event yields in different invariant mass intervals according to the event selection of Section 10.5.1. The errors include both statistical and systematic uncertainties (described in Section 10.8).

for comparison. Coarser binning is used at high mass to account for the lower number of expected events.

For each mass bin considered, the observed event yield in data agrees well with the Standard Model prediction, generally falling within $\pm 1\sigma$ according to the quoted uncertainties.¹⁵ The errors include the combined statistical, experimental, and theoretical systematic uncertainties, which are assessed in Section 10.8.

10.7.2 Kinematic Distributions

In this section, the kinematic properties of electrons in both data and MC simulation passing the full event selection are compared. The MC includes the Drell-Yan, Top, and Diboson processes, each stacked, summed, and normalized to 36.1 fb^{-1} . Together with the fake

¹⁵A deficit exists in the 900–1200 GeV and 1200–1800 GeV bins. Potential sources for this discrepancy were investigated, and no problems were found.

background estimate, the overall Standard Model expectation is obtained, and compared directly to data for several kinematic distribution of interest.

Figure 10.6 shows the dielectron invariant mass distribution, the main discriminating variable of the search, while Figures 10.7, 10.8, and 10.9 show distributions for the leading and subleading electron transverse energy E_T , pseudorapidity η , and azimuthal angle ϕ , respectively. The electron E_T , η , and ϕ distributions are presented in the signal region region beyond the Z boson mass peak $m_{ee} > 120$ GeV.

For each distribution presented, the data, shown with their associated Poisson statistical uncertainty, are compared to the stacked Standard Model expectation, and the ratio between the two can be seen in the sub-panel. Here, the turquoise band quantifies total systematic uncertainty on the Standard Model expectation, and it can generally be seen that the background modeling is consistent with the statistical uncertainty of the data.

The dielectron invariant mass serves as the primary search distribution for resonant new physics (as outlined in Sec. 10.1). The bin widths are constant in $\log(m_{ee})$, and the spectrum is seen to be quickly falling, covering multiple orders of magnitude by the time the TeV scale is reached. Clearly visible is the Z boson mass peak at $M_Z \simeq 91$ GeV. Also visible are three resonant Z'_χ signals from MC with pole masses of $M_{Z'} = 3$ TeV, $M_{Z'} = 4$ TeV, and $M_{Z'} = 5$ TeV, overlaid with the Standard Model expectation.

The observed data agree well with the Standard Model expectation over the entirety of the invariant mass spectrum, with no obvious signal-like excesses present. As quantified in the yield table, a slight deficit can be seen in the 1-2 TeV range. The highest dielectron invariant mass event has $M_{ee} = 2.90$ TeV. A display for this event, as well as the highest mass event from the 2015 run can be seen in Appendix C, along with their associated kinematic information.

Like the invariant mass, the electron E_T spectrum, coincident with the $E_T > 30$ GeV requirement, falls rapidly over several orders of magnitude with increasing electron E_T . The contribution from jets, which are generally “harder” than electrons, increases as high E_T is reached. In the ratio sub-panel, a good level of agreement between the data and the Standard Model expectation can be seen, as the statistical error of the data overlaps with the systematic band. The overlaid Z' signals exhibit a peak at roughly half their mass value, due to the two body nature of the decay.

The electron η distribution is symmetric relative to the beam axis, displaying a maximum at $\eta = 0$, and slowly falling off at higher values, until $|\eta| = 2.47$ is reached. The interval $1.37 < |\eta| < 1.52$ defines the transition region between the barrel and end-caps of the calorimeter. Events still populate this region, since the η plotted is the so-called “best-estimate,” which utilizes the tracking information (See Sec. 8.2.2). The event selection requirement is placed on the η position of the energy deposition in the EM calorimeter second layer.¹⁶ The systematic uncertainty, displayed in the sub-panel, is an approximately constant 5%, and a good level of agreement is seen between data and background expectation.

The electron ϕ distribution exhibits azimuthal symmetry in the $x - y$ transverse plane of the detector, with electrons equally distributed over $[-\pi, \pi]$. The overall level of agreement between the Standard Model expectation and data is good, with data accounted for by the flat 5% systematic band.

While no obvious signature of resonant new physics can be seen in the m_{ee} distribution, a rigorous statistical analysis is needed to conclude to what degree the data agrees with background expectation, which is the subject of the next chapter.

¹⁶Indeed, plotting this shows the transition region devoid of events.

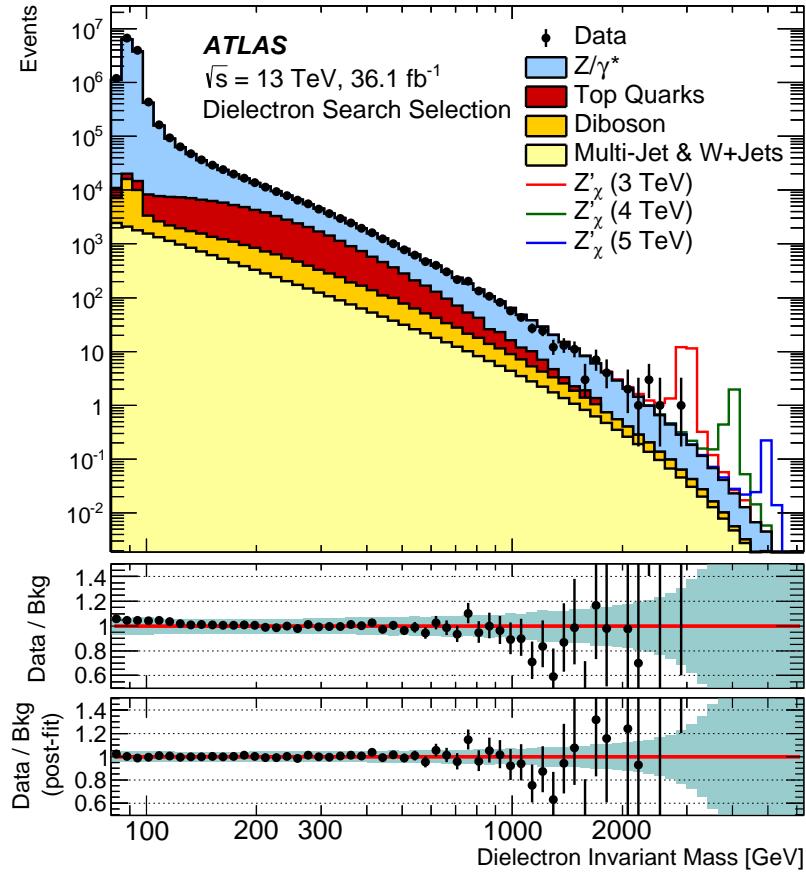


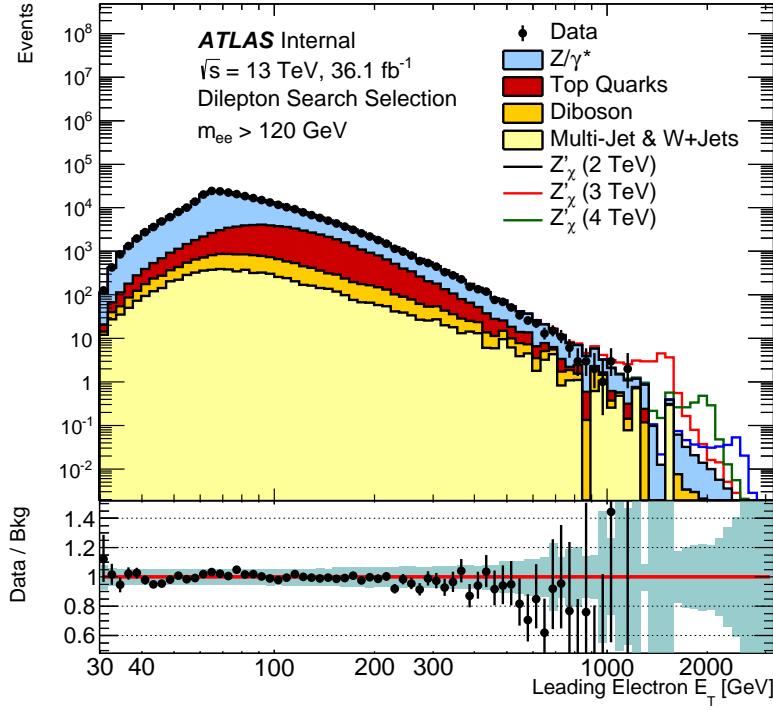
Figure 10.6: Dielectron invariant mass distribution for candidate dielectron events selected by the full event selection. The stacked sum of Standard Model backgrounds is normalized to 36.1 fb^{-1} and compared to data. Three Z'_χ signals are included for reference, overlaid with the Standard Model expectation. No obvious signal-like excess is present. The lower-panel sub-plot shows the “post-fit” result obtained from the statistical analysis.

10.8 Systematic Uncertainties

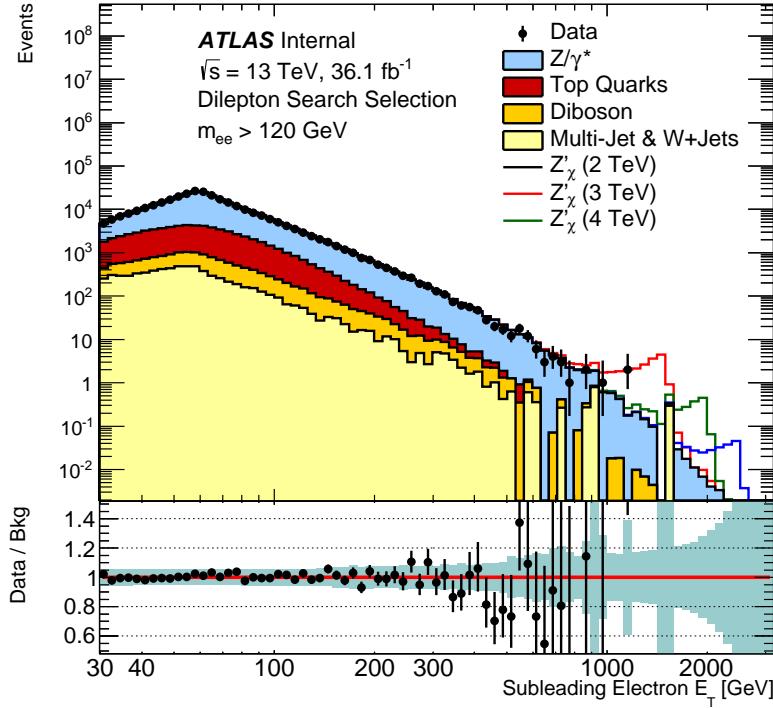
This final section describes the systematic uncertainties accounted for in the electron channel of the dilepton analysis. Statistical uncertainties affecting the Standard Model background expectation are negligible ($< 1\%$), owing to the high statistics simulation samples used in the analysis, and can therefore be safely neglected.¹⁷

The systematic uncertainties affecting the analysis can be divided into two general cat-

¹⁷Statistical uncertainties arise from the inherent random fluctuations in a measurement. In a large number of measurements, statistical uncertainties average to zero.

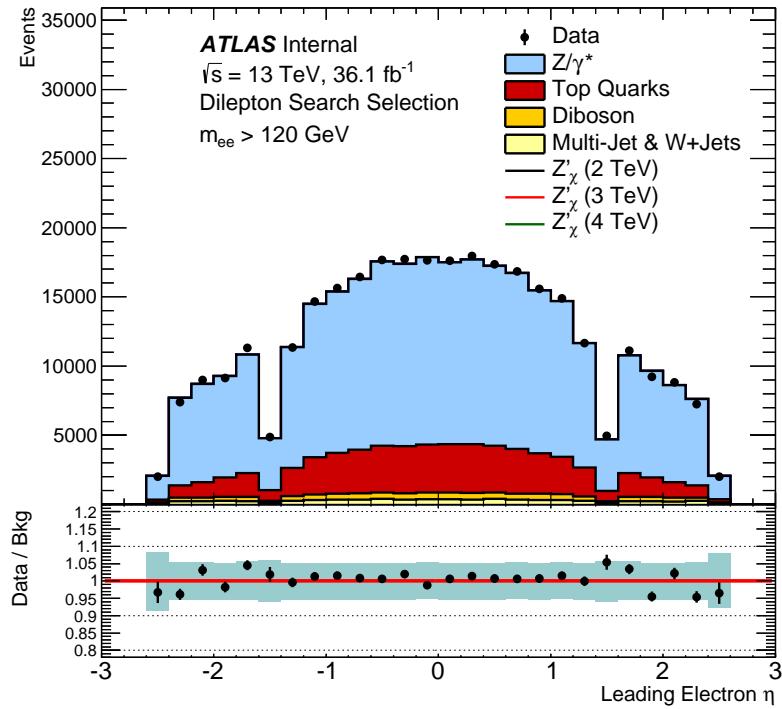


(a)

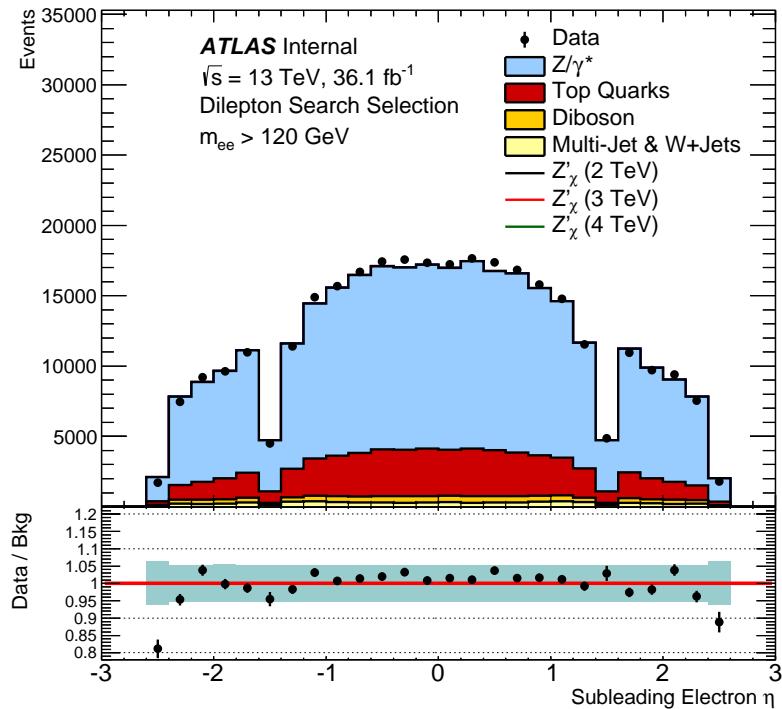


(b)

Figure 10.7: Leading (a) and subleading (b) electron E_T distributions in the dielectron invariant mass region $m_{ee} > 120 \text{ GeV}$, after the full event selection. The stacked sum of Standard Model backgrounds is normalized to 36.1 fb^{-1} and compared to data.

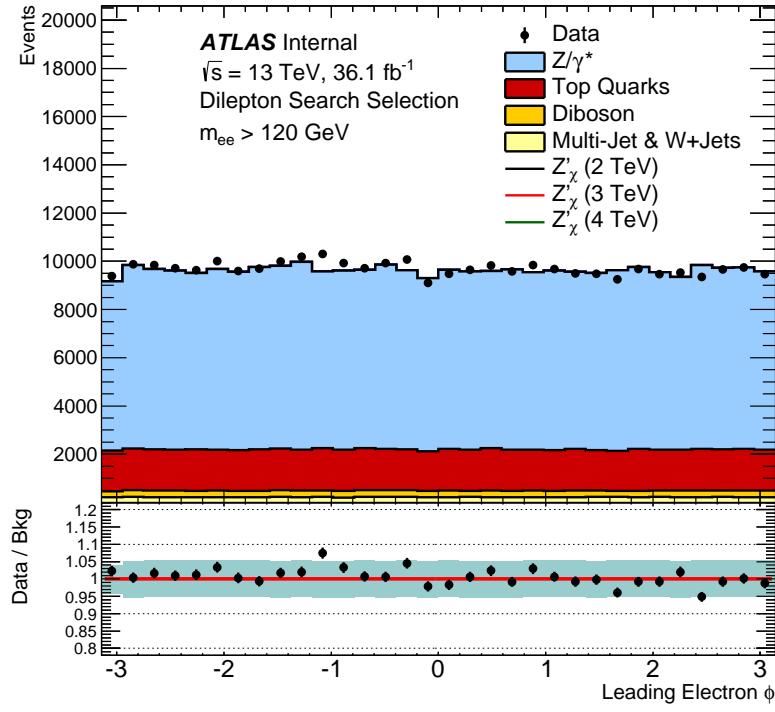


(a)

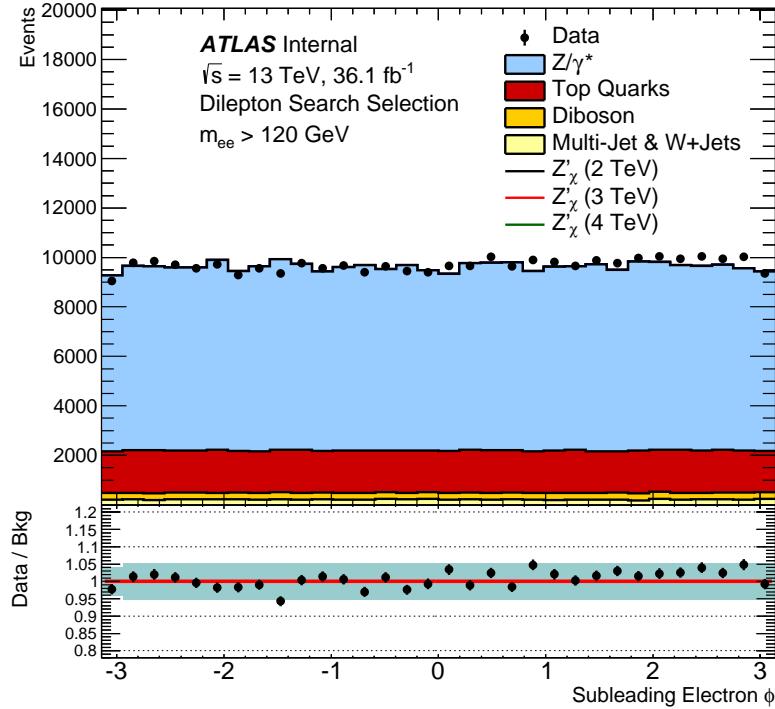


(b)

Figure 10.8: Leading (a) and subleading (b) electron η distributions in the dielectron invariant mass region $m_{ee} > 120 \text{ GeV}$, after the full event selection. The stacked sum of Standard Model backgrounds normalized to 36.1 fb^{-1} and compared to data.



(a)



(b)

Figure 10.9: Leading (a) and subleading (b) electron ϕ distributions in the dielectron invariant mass region $m_{ee} > 120 \text{ GeV}$, after the full event selection. The stacked sum of all Standard Model backgrounds normalized to 36.1 fb^{-1} and compared to data.

egories: experimental (to be described in Sec. 10.8.1) and theoretical (to be described in Sec. 10.8.2).¹⁸ All systematic uncertainties are assessed as functions of dielectron invariant mass m_{ee} , and quantify the $\pm 1\sigma$ variations in the expected dielectron event yield, according to the Standard Model expectation (See Table 10.3).

For each systematic source, the larger of two “upward” or “downward” **systematic variations** is taken to construct a symmetrized uncertainty estimate.¹⁹ The result of this procedure can be seen in Figure 10.10 for the experimental sources and Figure 10.11 for the theoretical sources. These symmetric variations are then accounted for as **nuisance parameters** in the search and limit setting procedures of the statistical analysis performed in the next chapter (See Sec. 11.1).

10.8.1 Experimental Uncertainties

The experimental uncertainties affecting the description of the Standard Model background estimate in the electron channel from several sources. Those relating to the measurement of electrons include: the simulation of the detector response [190] (i.e. the electron efficiencies), and the precision of the electron energy measurement [191] (i.e. the energy scale calibration, and electron energy resolution). Additional experimental uncertainties include those arising from the luminosity estimate, the LHC beam energy, the pile-up profile, and the Multi-jet & $W+jets$ fake background estimate. The magnitude of each can be seen in Figure 10.10 as a function of the dielectron invariant mass m_{ee} . There are a total of 11 experimental uncertainties considered in the electron channel. They include:

¹⁸Systematic uncertainties arise from consistent systematic effects introduced during the measurement, and by the assumptions, or choices, made in the analysis that affect the interpretation of the data. Systematic uncertainties are the result of experimental bias, and cannot simply be reduced by taking more data.

¹⁹An upward or downward systematic variation leads to an increased or decreased estimate on the Standard Model background yield.

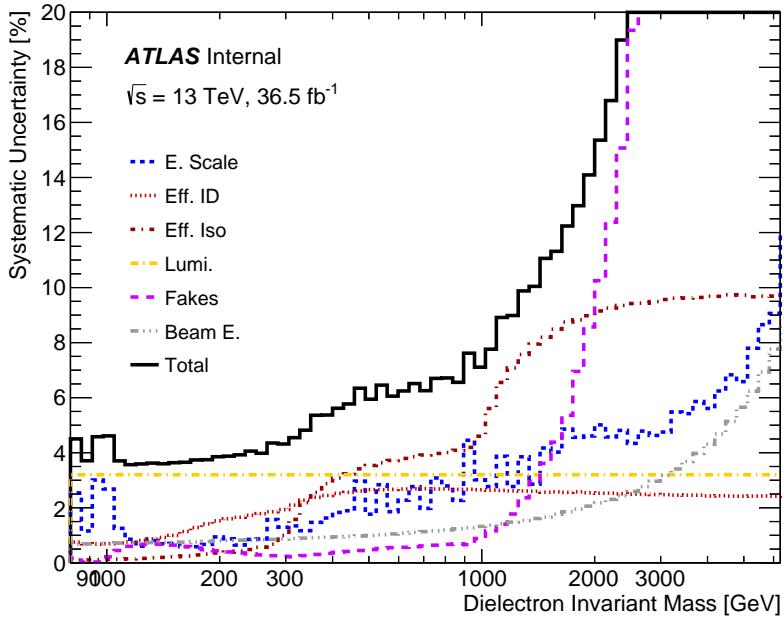


Figure 10.10: Experimental systematic uncertainties in the electron channel of the dilepton analysis. Shown are effects that shift the Standard Model event yield by greater than 1% anywhere in the dielectron invariant mass search spectrum. The largest experimental uncertainty is due to the $W+jets$ & QCD Multi-jet fake background estimate, which rises to 10% by $m_{ee} = 2$ TeV. See Section 10.8.1 for explanations of the source of each systematic uncertainty.

Electron Efficiencies: The electron reconstruction, identification, isolation, and trigger efficiency scale factors account for the difference between the efficiency observed in data and the one predicted by the simulation (See Sec. 9.4.2). Each efficiency correction comes with an associated systematic uncertainty, provided by the ATLAS electron/photon combined performance group (EGP), as an “upward” and “downward” variation from the nominal scale factor estimate. The variations are applied as event weights with the `ElectronEfficiencyCorrectionTool` when running over the simulation samples for analysis. The systematic uncertainties are determined by varying the tag-and-probe methodology used to extract the scale factor corrections [3]. For electrons with an $E_T > 200$ GeV the tag-and-probe technique can no longer be used

due to lack of statistics, so scale factors and their uncertainties are extrapolated to high E_T . The uncertainty in the identification scale factor is determined to be 2.0% for $E_T > 150$ [3]. Similarly, the uncertainty in the isolation scale factor was determined to be 2.0% for $E_T > 150$ GeV and 5.0% for $E_T > 500$ [190]. At an invariant mass of $m_{ee} = 2$ TeV, the propagation of the electron reconstruction and trigger efficiency uncertainties each result in a negligible uncertainty ($< 1\%$) on the Standard Model background estimate, while the ID and isolation efficiencies result in uncertainties of 2.6%, and 9.1%, respectively (See the “Eff. ID” and “Eff. Iso” curves in Fig. 10.10).²⁰

Electron Energy Scale: The measured electron energies are calibrated using a multivariate analysis (MVA) and in-situ techniques. Systematic uncertainties arise from the calibration procedure, and are provided by the ATLAS EGP group (See Sec. 8.3.1). As with the nominal correction, the `ElectronPhotonFourMomentumCorrection` tool is used to apply the “upward” and “downward” systematic variations in the energy scale correction on an event-by-event basis to the associated event electrons. While the calibration is strictly applied to data, in order to assess the systematic uncertainty, it is instead applied to the simulation, which has a negligible statistical uncertainty. There are over 60 systematic variations associated with the energy scale calibration [145]. The dilepton analysis uses a simplified calibration scheme in which all systematic variations are summed together in quadrature, reducing this number down to one, which greatly simplifies the complexity of the statistical analysis, presented in the following chapter. At an invariant mass of $m_{ee} = 2$ TeV, the uncertainty in the electron energy scale leads to a 4.1% uncertainty on the Standard Model background estimate (See the

²⁰Indeed, two “kinks” can be seen in isolation scale factor uncertainty at $m_{ee} \simeq 300$ GeV and $m_{ee} \simeq 1$ TeV in Figure 10.10, which arise as the scale factor uncertainty estimate abruptly changes at $E_T = 150$ GeV and $E_T = 500$ GeV.

“E. Scale” curve in Fig. 10.10). The energy scale in general does not greatly impact the sensitivity of the search, as the main effect would be to shift the precise location of a potential Z' peak, while still preserving its shape.

Electron Energy Resolution: The electron energy resolution in simulation differs from the one observed in data. The simulation is corrected by smearing the individual electron energies according to data-driven energy smearing factors extracted by the ATLAS EGP group (as explained in Sec. 8.3.2). As was the case with the energy scale calibration, a single systematic uncertainty is used to account for the energy resolution uncertainty [145]. The “upward” and “downward” variations are applied on an event-by-event basis using the `ElectronPhotonFourMomentumCorrection` tool. For high E_T electrons, the energy resolution is dominated by the so-called calorimeter “constant term” (See Sec. 8.3.2), which has an uncertainty at the per-mille level [143]. Therefore, only small shifts are introduced to the width of the energy smearing distribution, and the resultant uncertainty on the Standard Model expectation is negligible.

$W+jets$ & Multi-jet Fake Background Estimate: The matrix method is used to estimate the background contribution that arises from misidentified jets that pass the dielectron selection. The uncertainty on the fake background estimate is ascertained by comparing variations in the methodology of the matrix method. This includes varying the extrapolation methodology, how the electron fake-rate is determined, and comparing the resulting estimate from events with same-charge and opposite-charge electrons. Each uncertainty is summed together in quadrature to produce an overall uncertainty estimate. The extrapolation uncertainty can become very large at high dielectron masses m_{ee} (See the “Fakes” curve in Fig. 10.10), but it does not greatly

affect the interpretation of the final results of the analysis.

Luminosity: A flat 3.2% uncertainty [192] is assessed on the integrated luminosity of the 36.1 fb^{-1} dataset according to the ATLAS luminosity group [174] (See the “Lumi.” curve in Fig. 10.10). The luminosity uncertainty affects the normalization of all signal and background simulation samples.

LHC Beam Energy: The LHC beam energy is estimated to have an uncertainty of 0.65%, which results in an 2.0% uncertainty on the Standard Model background estimate at an invariant mass of $m_{ee} = 2 \text{ TeV}$ (See the “Beam E.” curve in Fig. 10.10)

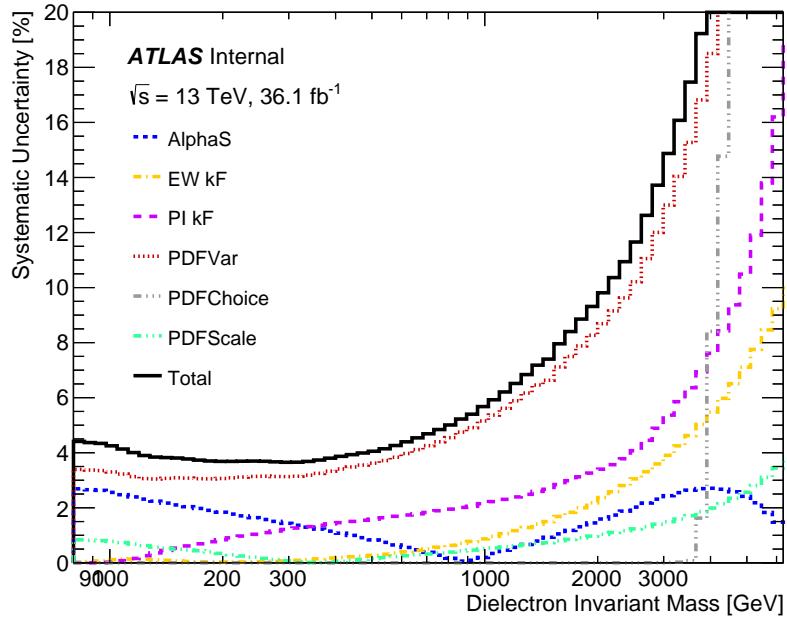
Monte Carlo statistics: While not strictly a systematic uncertainty, the MC statistical uncertainty is assessed and found to have a negligible effect (< 1%) on the Standard Model estimate. Each MC sample used in the electron channel analysis is generated with an adequate number of statistics to ensure accurate modeling of the dielectron invariant mass distribution.

Pile-up Reweighting: The PileupReweighting tool is used to reweight the MC simulation to account for the changing pile-up conditions in data. The tool provides “upward” and “downward” variations to account for the systematic uncertainty associated with the extraction of the pile-up scale factors from data [193]. Systematic variations are applied on an event-by-event basis as event weights during the analysis. The pile-up uncertainty was determined to have a negligible effect (< 1%) on uncertainty of the overall Standard Model background estimate.

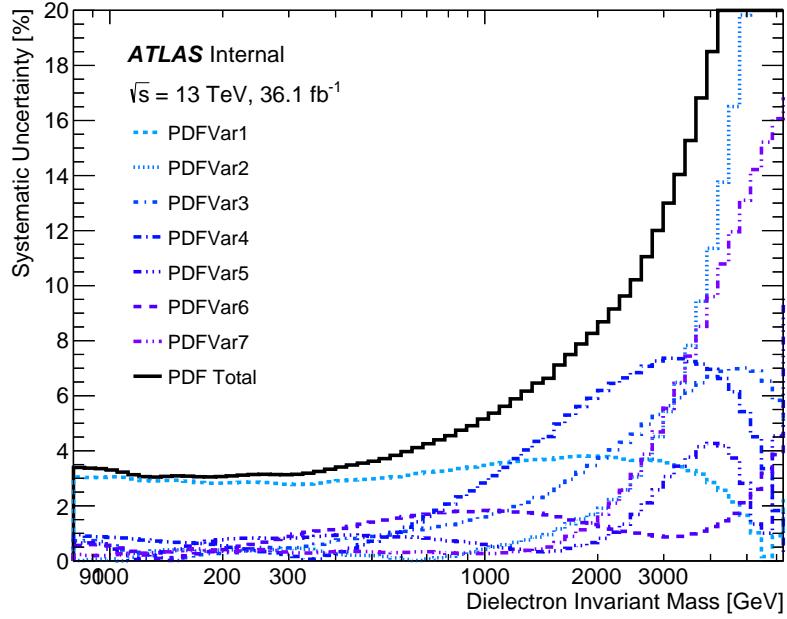
10.8.2 Theoretical Uncertainties

Similarly, multiple theoretical uncertainties affect the description of the Standard Model background estimate in the electron channel, and can be assessed on each individual background. Theoretical uncertainties affecting the Drell-Yan background estimate, the dominant background of the analysis, include those arising from the Parton Distribution Functions (PDFs), the specific choice of PDF set used to perform the uncertainty estimate, the PDF scale uncertainty, the high-mass running of α_S , and assumptions regarding the calculations of the Electroweak (EW) and Photon-Induced (PI) k-Factor corrections. Uncertainties affecting the Top and Diboson background estimates arise from their respective cross section estimates, and hence are treated as normalization uncertainties. The magnitude of each uncertainty can be seen in Figure 10.11 as a function of the dielectron invariant mass m_{ee} . There are a total of 8 theoretical uncertainties considered in the electron channel. They include:

PDF Variation: The Drell-Yan cross section is corrected from NLO to NNLO in QCD with the mass-dependent QCD k-Factor corrections (See Sec. 9.4.5). The NNLO cross section is calculated with VRAP configured with the CT14NNLO PDF set. The PDF uncertainties are obtained using the 90% C.L. CT14NNLO PDF error set, and propagated to the uncertainty on the Drell-Yan cross section estimate, which can be interpreted as the uncertainty on the QCD k-Factor correction. At an invariant mass of $m_{ee} = 2$ TeV, the PDF variation uncertainty is 8.7% (See the “PDFVar” curve in Fig. 10.11). The CT14NNLO error set contains 26 eigenvector variations (See Sec. 3.2.2), which are used to obtain an estimate of the overall PDF uncertainty with Eq. 3.10. As a simplification to the dilepton search, instead of utilizing the full set of 26 eigenvector variations for the



(a)



(b)

Figure 10.11: Theoretical systematic uncertainties in the electron channel of the dilepton analysis (a). Shown are effects that shift the Standard Model event yield by greater than 1% anywhere in the dielectron invariant mass search spectrum. The largest uncertainty is due to the PDF Variation uncertainty, which is also shown decomposed into its individual eigen-vector variations (b). See Section 10.8.2 for an explanation of the source of each systematic uncertainty.

PDF uncertainty calculation, a re-diagonalized set of 7 was used, which was provided by the CTEQ collaboration [194] and obtained with MP4LHC [195].²¹ This is the minimal number of variations that still captures the necessary mass-dependent correlations, and maintains an accurate description of the PDF error, as it pertains to the Drell-Yan process.²² The forms of each of the 7 eigenvector variations are incorporated into the `LPXKFactorTool` and applied on an event-by-event basis as individual event weights to account for the PDF uncertainty on the QCD k-Factor correction. The PDF variation uncertainty is the largest theoretical systematic uncertainty in the dilepton search.

PDF Choice: In addition to the PDF variation uncertainty on the NNLO QCD calculation, a PDF choice uncertainty is also considered, according to the recommendations of the PDF4LHC forum [195]. The PDF choice uncertainty arises from the specific choice of PDF set used to calculate the NNLO QCD k-Factor correction. At an invariant mass of $m_{ee} = 2$ TeV, the PDF choice uncertainty is $< 1.0\%$, but becomes significant at high mass (See the “PDFChoice” curve in Fig. 10.11). It is derived by comparing the central cross section prediction of several other PDF parameterizations, namely `MMHT14` [196] and `NNPDF3.0` [81], to the 90% C.L. error envelope prediction of `CT14NNLO` (See Fig. 10.12). If the central value of an alternative PDF set lies outside the `CT14NNLO` error envelope, it is taken as an additional uncertainty. Specifically, this occurs roughly at an invariant mass of 3.5 TeV, at which point the nominal `NNPDF` prediction diverges from the `CT14NNLO` envelope.

PDF Scale: Owing to the finite precision of the fixed-order NNLO QCD calculation, the

²¹The reduction in the number of nuisance parameters prevents the statistical analysis from becoming computationally intractable.

²²This is to say, the quadrature sum of the 7 eigenvector variations matches the `CT14NNLO` error envelope of the full 26.

theoretical uncertainties arising from the renormalization and factorization scale choices are considered (See Sec. 3.1.2). Also derived with VRAP, these uncertainties are obtained by simultaneously varying the renormalization scale μ_R and factorization scale μ_F of the nominal CT14NNLO prediction up and down by a factor of two (See the “PDFScale” curve in Fig. 10.11).

α_S Value: The final PDF related uncertainty arises from the value of α_S chosen for the VRAP calculations (See Fig. 2.9). The uncertainty on α_S is computed by varying the nominal value of 0.118 by ± 0.003 , and propagating it to the NNLO Drell-Yan cross section prediction (See the “AlphaS” curve in Fig. 10.11).

EW k-Factor: A systematic uncertainty on the EW k-Factor correction is considered. It is currently not known how to precisely combine calculation involving higher-order EW effects with those involving higher-order QCD effects. Both additive and multiplicative approaches are considered, and the difference between the two approaches is taken as a systematic uncertainty (See the “EW kF” curve in Fig. 10.11).

PI k-Factor: A systematic uncertainty on the PI k-Factor correction is considered and arises from the choice of quark mass values and uncertainty on the photon PDF $\gamma(x, Q^2)$ in the calculation of the PI cross section (See the “PI kF” curve in Fig. 10.11).

Top Background Normalization: Like the Drell-Yan background, theoretical uncertainties affecting the normalization of the Top background are considered. The systematic effects are accounted for with variations on the renormalization and factorization scales, the PDF variations, and variations on α_S . Additionally, variations in association with several top quark specific effects, such as the value of top quark mass, are included, as recommended by the Top working group [197].

Diboson Background Normalization: Like the Drell-Yan and Top backgrounds, systematic effects for the Diboson background are accounted for, which included variations on the renormalization and factorization scales, the PDFs, and α_S .

10.8.3 Summary

This final section summarizes the effect each systematic has on the expected Z' signal and Standard Model background yields. Table 10.4 presents the relative impact each uncertainty has on the nominal event yield, according to the prescribed $\pm 1\sigma$ variations. Values presented are obtained directly from the estimates shown in Figures 10.10 and 10.11.

Any systematic uncertainty that changes the event yield by less than 1% over the entirety of the dielectron search region is neglected in the statistical analysis, as the final result is not affected. These neglected uncertainties include the uncertainty on the electron trigger and reconstruction efficiencies, the electron energy resolution, the MC statistical uncertainty, the pile-up uncertainty, and the Diboson background normalization.

Furthermore, theoretical uncertainties are not considered on the signal processes, since the analysis aims to be as model-independent as possible, and systematics associated with the signal can strongly depend on the particular theory of new physics. This decision does not have a significant effect on the final statistical results. The following chapter discusses the statistical analysis in detail.

10.8.4 Comment on PDF Uncertainties

As described in Section 10.8.2, the systematic uncertainty associated with the PDFs gives rise to the dominant source of theoretical uncertainty in the dilepton analysis. While this

Source	m_{ee} at 2 TeV [%]		m_{ee} at 4 TeV [%]	
	Signal	Background	Signal	Background
Luminosity	3.2	3.2	3.2	3.2
MC statistical	<1.0	<1.0	<1.0	<1.0
Beam energy	2.0	2.0	4.1	4.1
Pile-up effects	<1.0	<1.0	<1.0	<1.0
DY PDF choice	N/A	<1.0	N/A	8.4
DY PDF variation	N/A	8.7	N/A	19
DY PDF scales	N/A	1.0	N/A	2.0
DY α_S	N/A	1.6	N/A	2.7
DY EW corrections	N/A	2.4	N/A	5.5
DY PI corrections	N/A	3.4	N/A	7.6
Top quarks theoretical	N/A	<1.0	N/A	<1.0
Dibosons theoretical	N/A	<1.0	N/A	<1.0
Reconstruction efficiency	<1.0	<1.0	<1.0	<1.0
Isolation efficiency	9.1	9.1	9.7	9.7
Trigger efficiency	<1.0	<1.0	<1.0	<1.0
Identification efficiency	2.6	2.6	2.4	2.4
Electron energy scale	<1.0	4.1	<1.0	6.1
Electron energy resolution	<1.0	<1.0	<1.0	<1.0
$W+\text{jets} \& \text{Multi-jet}$	N/A	10	N/A	129
Total	10	18	11	132

Table 10.4: A Summary of the relative systematic uncertainties affecting the expected signal and background event yields at dielectron invariant masses of 2 and 4 TeV. Background values correspond to the relative change in the total number of expected events in the 2 and 4 TeV bins of the m_{ee} histogram depicted in Figure 10.6, while signal values are calculated from changes in the FWHM of the Z'_χ signal at pole masses of $M_{Z'} = 2$ TeV and $M_{Z'} = 4$ TeV. Here N/A refers to uncertainties not considered on the signal process [5].

uncertainty does not necessarily prohibit the discovery of a high-mass resonance, such as a Z' boson, it can greatly affect the *characterization* of such a discovery, as the ability to distinguish between different possible models may be diminished. Furthermore, the size of the PDF uncertainty can severely *reduce* the discovery potential of other theoretical interpretations, such as those arising from contact-interaction type models [62], or any kind of new physics that features a broad excess, rather than a distinct resonant peak, over the Standard

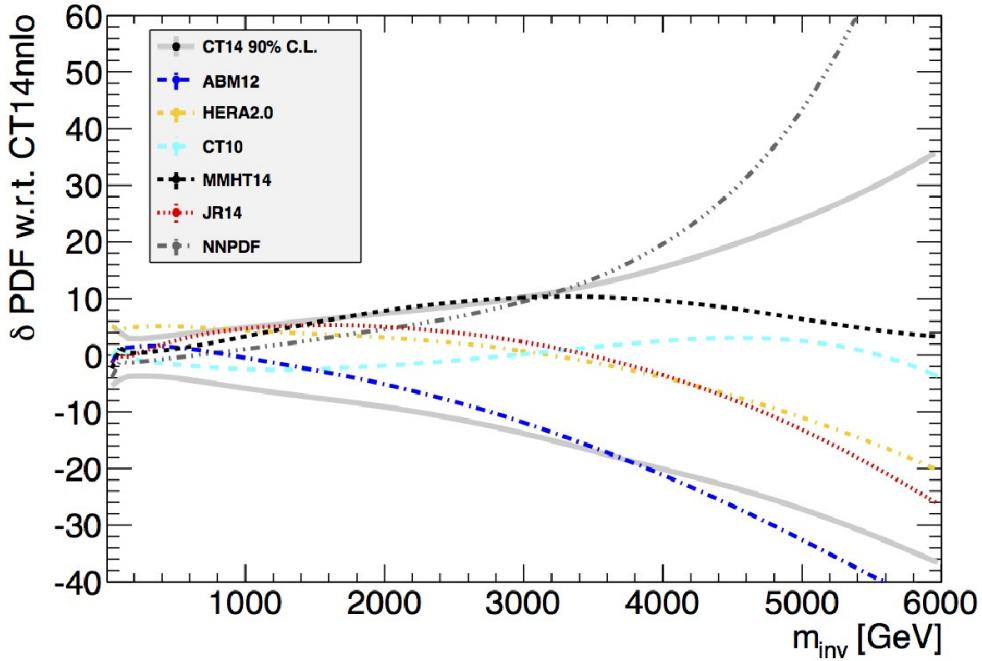


Figure 10.12: Drell-Yan cross section uncertainty as a function of invariant mass as predicted by several modern PDF sets [4].

Model expectation.

As the dilepton final-state is sensitive to a host of new phenomena (See Sec. 5.2), it is crucial to find a way to reduce this theoretical uncertainty; otherwise, despite the increasing size of future LHC datasets, it will become increasingly difficult to exclude non-resonant models of new physics or characterize the Z' in case of discovery.

Moreover, in accordance with the PDF4LHC recommendation [195], the dilepton analysis includes an ad-hoc ‘‘PDF Choice’’ uncertainty (See Sec. 10.8.2) to account for the fact that the nominal **MMHT14** and **NNPDF3.0** predictions depart from the **CT14NNLO** 90% C.L. error band, which is the baseline PDF used in the dilepton analysis (See Fig. 10.12).

Folding in the PDF uncertainties from multiple sets results in an overestimate and a loss in discovery potential. If the ‘‘PDF Variation’’ uncertainty can be reduced, such that the modern PDF sets are brought into closer agreement, then the ‘‘PDF Choice’’ uncertainty

becomes irrelevant.²³ Therefore, attacking the PDF uncertainty is crucial to the future success of the dilepton search. A strategy devised to further reduce this uncertainty is presented in Section 10.1.

²³The inclusion of additional high-quality LHC data into future PDF global fits could have the effect of resolving disagreements between the current estimates provided by the PDF global fitting groups, especially at high x where the PDFs are least constrained.

Chapter 11

Statistical Interpretation and Results

In this chapter, the statistical methods used to search for new high-mass resonances in the electron channel of the dilepton analysis are presented. The search for a signal is performed by scanning the dielectron invariant mass distribution of Figure 10.6 for deviations in the observed event yield relative to the Standard Model expectation; a significant excess might indicate the presence of a signal. The search is performed with a dedicated statistical model, which is constructed in Section 11.1.

To facilitate the desired interpretations of the data, the statistical model utilizes Z' signal templates, which parametrize the expected Z' event yields for different theoretical models and mass hypotheses $M_{Z'}$ under investigation. These templates are produced with a dedicated signal reweighting technique, described in Section 11.2. The search for a signal and the results thereof are presented in Section 11.3.

Two complimentary methods are used to quantify the significance of any potentially observed excess. In case of a null result, exclusion limits are set at 95% C.L. on several Z' model parameters of interest using a Bayesian technique outlined in Section 11.4. The results are interpreted in the context of the SSM, the E_6 model, and the Minimal model.¹ The chapter closes with a final discussion of results in Section 11.5.

¹Additional resonant and non-resonant interpretations are possible as well, but are beyond the scope of this dissertation. See Section 5.2 for a brief discussion.

11.1 Statistical Model

The following section describes the construction of the statistical model used to perform the statistical analysis in the electron channel of the dilepton analysis. The model is based on the dielectron invariant mass m_{ee} , due to its excellent discriminating power between the signal-like excesses of a Z' boson and the smoothly falling background predicted by the Standard Model (as discussed in Sec. 10.1). To ensure consistency between the search and limit setting, a common statistical model is developed, which is then used to perform statistical hypothesis tests for the search (as in Sec. 11.3.1), and set upper limits on Z' model parameters of interest (as in Sec. 11.4).

The model is formulated in terms of a binned likelihood function $\mathcal{L}(\mu, \boldsymbol{\theta})$, where μ is the **parameter of interest**, which for this analysis parametrizes either the signal strength, the signal cross section, or a coupling constant, and $\boldsymbol{\theta}$ is a set of **nuisance parameters**, which account for the electron systematic uncertainties (as described in Sec. 10.8).

The likelihood itself is constructed from a product of Poisson² probabilities from the $N_{\text{bins}} = 66$ invariant mass bins of the dielectron invariant mass distribution of Figure 10.6. The multi-bin approach allows for signal shape information, such as the decay width $\Gamma_{Z'}$ of a Z' boson, to be distributed over multiple m_{ee} mass bins. Incorporating this additional information into the likelihood function improves the overall sensitivity of the search.

The likelihood function is

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{k=1}^{N_{\text{bins}}} \frac{\lambda_k(\mu, \boldsymbol{\theta}) e^{-\lambda_k(\mu, \boldsymbol{\theta})}}{n_k!} \prod_{i=1}^{N_{\text{sys}}} \mathcal{N}(\theta_i, 0, 1) \quad (11.1)$$

²The Poisson distribution $P(n|\lambda) = \lambda^n e^{-\lambda} / n!$ quantifies the probability that exactly n events are observed when a mean number of events λ is expected.

where the first product is over each m_{ee} mass bin k and the second is over the systematic uncertainties. For a given mass bin k , the number of observed events n_k is described by a Poisson distribution with expectation value $\lambda_k(\mu, \boldsymbol{\theta})$.

The total number of expected dielectron events $\lambda_k(\mu, \boldsymbol{\theta})$ in a given m_{ee} mass bin k depends on the values of μ and $\boldsymbol{\theta}$. The parameter $\mu \geq 0$ is known as the signal strength, which acts a multiplicative normalization factor on the expected signal yield, while the parameter $\boldsymbol{\theta}$ specifies a set of nuisance parameters, which encode the effect of the systematic uncertainties (defined in Sec. 10.8) on the nominal signal and background expectations. The nominal range of each variation can be seen in Figure 10.10 and Figure 10.11. In order to aid in constraining the best-fit values of the nuisance parameters in the statistical analysis, the likelihood is augmented by a product of unit normal distributions $\mathcal{N}(\theta_i, 0, 1)$, one for each nuisance parameter θ_i .

The expected number of events $\lambda_k(\mu, \boldsymbol{\theta})$ can be decomposed into a sum of expected Z' signal events $s_k(\boldsymbol{\theta})$ and Standard Model background events $b_k(\boldsymbol{\theta})$, where

$$\lambda_k(\mu, \boldsymbol{\theta}) = \mu s_k(\boldsymbol{\theta}) + b_k(\boldsymbol{\theta}). \quad (11.2)$$

Again, the signal strength μ serves as a floating normalization constant on the signal expectation $s_k(\boldsymbol{\theta})$. The background expectation $b_k(\boldsymbol{\theta})$ is the sum of dielectron events that pass the dielectron event selection (See Sec. 10.5) and go on to populate each m_{ee} mass bin k .³ The expected number of signal events $s_k(\boldsymbol{\theta})$ can be further expressed as

$$s_k(\boldsymbol{\theta}) = s_k \times \left(1 + \sum_{i=1}^{N_{\text{sys}}} \theta_i \frac{\delta s_{ki}}{s_k} \right). \quad (11.3)$$

³These include the standard Drell-Yan, Top, Diboson, and $W+\text{jets}$ & QCD Multi-jet processes.

The central value s_k is obtained from Eq. 6.1, where

$$s_k = \sigma_{Z'} Br(Z' \rightarrow e^+ e^-) \cdot \mathcal{A}\epsilon(Z' \rightarrow e^+ e^-) \cdot \mathcal{L}_{\text{int}} \quad (11.4)$$

where $\sigma_{Z'}$ is the Z' production cross section calculated with Eq. 4.16, $Br(Z' \rightarrow e^+ e^-)$ is the Z' branching fraction to electron pairs, $\mathcal{A}\epsilon(Z' \rightarrow e^+ e^-)$ is the total efficiency to select a Z' signal event, quantified in Figure 10.3 as a function of pole mass, and \mathcal{L}_{int} is the integrated luminosity of the dataset.

The ratio $\delta s_{ki}/s_{ki}$ in Eq. 11.3 expresses the relative shift of the expected signal yield s_k in m_{ee} mass bin k according to the $\pm 1\sigma$ variation induced by systematic uncertainty i .⁴ These shifts are parametrized by the mass-dependent relative systematic uncertainties of Figure 10.10 and Figure 10.11, whose explicit values are summarized in Table 10.4.

Likewise, the expected number of background events can be expressed

$$b_k(\boldsymbol{\theta}) = b_k \times \left(1 + \sum_{i=1}^{N_{\text{sys}}} \theta_i \frac{\delta b_{ki}}{b_k} \right) \quad (11.5)$$

where the central value b_k is the Standard Model expectation in m_{ee} mass bin k and includes the contributions from each background process. It similarly follows from Eq. 6.1, accounting for the respective cross sections, branching fractions, and total efficiencies of each individual background process. The relative shift in the background expectation from systematic i is accounted for by $\delta b_{ki}/b_{ki}$, and is due to the same systematics uncertainties in Figure 10.10 and Figure 10.11 as the signal process. Correlations between signal and background systematic uncertainties are accounted for across all mass bins, since both $s_k(\boldsymbol{\theta})$ and $b_k(\boldsymbol{\theta})$ are each

⁴The nuisance parameters θ_i , associated with each systematic uncertainty i , act as free normalizations, which have the effect of smearing out the nominal Poisson prediction.

affected by the same set of systematics.

Equations 11.3, 11.4, and 11.5 specify all the required inputs to perform the statistical analysis. However, in order to perform a comprehensive search, multiple Z' models and mass hypotheses $M_{Z'}$ must be tested. Therefore, the dilepton analysis makes use of a **signal reweighting** technique, which is used to generate Z' signal samples on-the-fly using the Leading-Order (LO) Drell-Yan Monte Carlo (MC) samples (See Sec. 9.3.6).

The power in this technique lies in its ability to *reweight* the Drell-Yan MC to any Z' model or pole mass $M_{Z'}$ of interest. A reweighted Z' signal sample can then be used to construct a **Z' signal template**, a dielectron invariant mass histogram that contains the precise signal expectation $s_k(\boldsymbol{\theta})$ for the specific Z' model and pole mass $M_{Z'}$ under investigation. This formulation allows for the parameter $s_k(\boldsymbol{\theta})$ of Eq. 11.3 to be fully specified for a wide variety of Z' models or model parameters under consideration, thus enabling the possibility of a comprehensive search.

11.2 Signal Reweighting

This is a contribution to the ATLAS dilepton analysis by the author and is described in detail in Reference [198]. The dilepton analysis uses signal templates in order to interpret the level of disagreement between the observed data and the Standard Model expectation as part of the statistical analysis. Signal templates are produced from the LO Drell-Yan MC samples with the `LPXSignalReweightingTool`. Event-by-event weights are calculated and applied to Drell-Yan events passing the event selection, thereby “reweighting,” or rescaling, them into corresponding Z' signal events. Effectively this produces a unique sample of Z' signal events, which, for all intents and purposes, matches what would otherwise be produced

from a dedicated MC generation, but with orders of magnitude less computing required.

As a result, signal templates can be generated for a diverse range of Z' models and model parameter choices using only the standard analysis Drell-Yan simulation. This allows for a much more comprehensive statistical analysis to be performed, which in the case of the Minimal model, would otherwise be impossible due to the large amount of signal MC required to effectively scan the model parameter space.

The signal templates include the same set of experimental and theoretical corrections as applied to the other simulation samples, including the standard electron efficiency, energy resolution, and pile-up corrections (described in Sec. 9.4). In addition, higher-order QCD corrections are applied to the signal templates, utilizing the same LO-to-NNLO QCD mass-dependent k-Factor as the background MC (See Table 9.3); however, higher-order EW corrections are left out due to the model-dependent nature of the interaction between the Z' and the other Standard Model gauge bosons.⁵

Two distinct reweighting strategies are employed in the dilepton analysis. The SSM and E_6 model signal templates neglect the effect of interference (as described in Sec. 5.3.1) between the γ^*/Z and Z' processes, while the Minimal model templates include this effect fully. The following sections describe each of these approaches in detail.

11.2.1 Signal Template Generation

Signal templates are obtained from the reweighting of the PYTHIA8 LO Drell-Yan simulation (See Sec. 9.3.6). The idea is to replace the Drell-Yan differential cross section with the

⁵Indeed, in the calculation of the Z' decay width $\Gamma_{Z'}$, only the Standard Model fermion decay channels are considered; i.e., right-handed neutrinos, additional gauge bosons, and exotic fermions are neglected as possible decays.

corresponding cross section for Z' production, where

$$\frac{d\sigma}{dt}(q\bar{q} \rightarrow \gamma^*/Z \rightarrow e^+e^-) \rightarrow \frac{d\sigma}{dt}(q\bar{q} \rightarrow Z' \rightarrow e^+e^-). \quad (11.6)$$

This can be accomplished if analytic expressions for the Drell-Yan and Z' matrix elements are known. As these expressions were calculated in Section 4.2.1 and Section 5.3.1, they can now be used here. The ratio between these two matrix elements, each evaluated with the truth level kinematic information⁶ of the dielectron event, defines the associated event weight w_{RW} . This weight can then be used to reweight the specified Drell-Yan event to the corresponding Z' event.

The event weight, a function of the Z' pole mass, takes the form

$$w_{RW}(M_{Z'}) = \frac{\sum_{i,j \in \{L,R\}} |\mathcal{M}_{ij}(Z')|^2}{\sum_{i,j \in \{L,R\}} |\mathcal{M}_{ij}(\gamma^*/Z)|^2}, \quad (11.7)$$

where the numerator and denominator define the Z' and Drell-Yan matrix elements, respectively, and the sum is performed over all helicity states. The matrix element for each process share a common structure and can be expressed as the product of an angular term, and term that factorizes the couplings, such that

$$|\mathcal{M}_{ij}|^2 = 4 |\mathcal{G}_{ij}|^2 (1 + \lambda_{ij} \cos \theta)^2, \quad \lambda_{ij} = \begin{cases} +1, & \text{if } i = j \\ -1, & \text{if } i \neq j, \end{cases} \quad (11.8)$$

⁶i.e. with the four-momenta of the initial-state quarks and final-state electrons

where the couplings are

$$\mathcal{G}_{ij}(\hat{s}) = \begin{cases} e^2 \frac{Q^l Q^q}{\hat{s}} + g_Z^2 \frac{g_i^l g_j^q}{\hat{s} - M_Z^2 + iM_Z \Gamma_Z}, & \text{for } \gamma^*/Z \\ g_{Z'}^2 \frac{z_i^l z_j^q}{\hat{s} - M_{Z'}^2 + iM_{Z'} \Gamma_{Z'}}, & \text{for } Z'. \end{cases} \quad (11.9)$$

Thus, to evaluate the event weight $w_{RW}(M_Z)$ of Eq. 11.7, a set of factors associated with the Drell-Yan process, with the Z' model under investigation, and with the kinematics of the event are required. The Drell-Yan matrix element is defined by the EW fermion quantum numbers Q , g_L , and g_R , which are listed in Table 2.4, by the mass M_Z and decay width Γ_Z of the Standard Model Z boson, and by its the gauge coupling constant g_Z .

Likewise, the Z' matrix element is defined by the $U(1)'$ gauge coupling $g_{Z'}$, by the mass $M_{Z'}$ and decay width $\Gamma_{Z'}$ of the Z' boson, and by the fermion $U(1)'$ quantum numbers z_L and z_R . The analytic expression for decay width of the Z' boson is obtained from Eq. 5.4. The associated parameters for the SSM, the E_6 model, and the Minimal model can be found in Section 5.3.2, Section 5.3.3, and Section 5.3.4, respectively, although this formulation is fully general, and any set of Z' model parameters can suffice.

Each matrix element shares several common kinematic factors as well. Both depend on the scale of event $\hat{s} = m_{ee}^2$, obtained from the dielectron invariant mass, and both share a dependence on the initial-state quark flavors and final-state electron angular terms $\cos \theta$. It is necessary to identify the flavor of the initial-state quark-antiquark pair on an event-by-event, so that the quark quantum numbers can be correctly specified in Eq. 11.9; similarly, it is necessary to identify the four-momenta of the electrons that originate from the Z' decay vertex in order to extract $\cos \theta$.

The reweighting technique itself is codified in a the `LPXSignalReweightingTool` [184], a

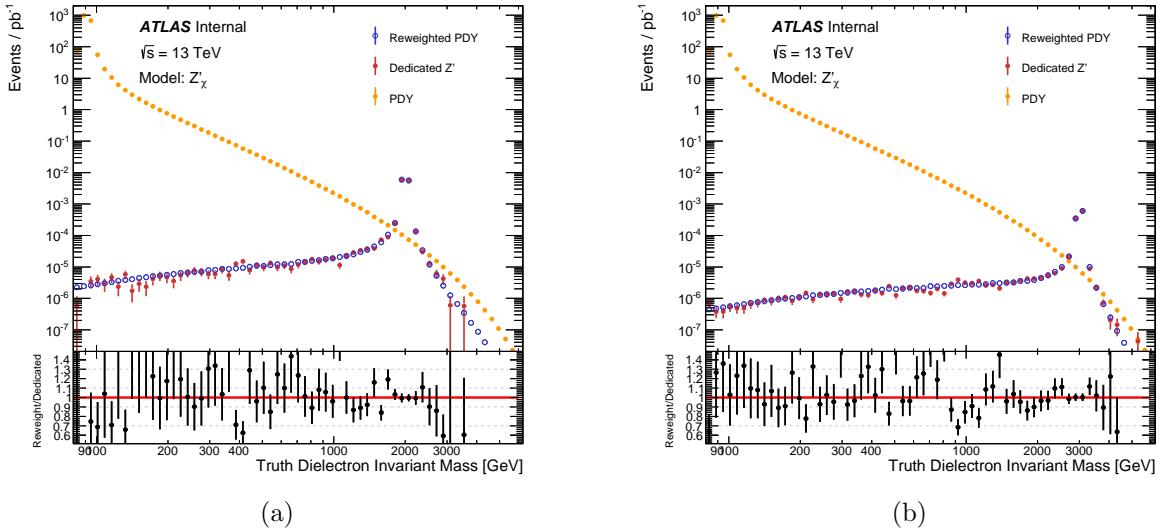


Figure 11.1: Validation of the signal reweighting procedure. Reweighted signal templates are compared to their dedicated PYTHIA counterparts at pole masses of 2 (a) and 3 (b) TeV for the Z'_χ of the E_6 model. The effect of interference is not included in the dedicated generation, and is similarly not included in the signal reweighting procedure for this comparison. Percent level agreement is seen in the peak region, and the signal shape is effectively captured across the entire invariant mass spectrum.

package which interfaces with the ROOTCORE framework, and is designed to be plugged into the main dilepton analysis workflow in the same way as the other standard analysis tools (See Sec. 10.4). The user specifies a Z' model of interest, which determines gauge coupling constant $g_{Z'}$ and $U(1)'$ fermion quantum numbers, and chooses a Z' pole mass in order to obtain the associated event weights.

The Standard Model parameters are predefined, and the tool selects the required incoming quarks and outgoing leptons from the event, so that their four-momenta and associated quantum numbers can be used to evaluate the matrix elements of Eq. 11.8, and hence return the weight of Eq. 11.7. The operation of the tool can be seen in Figure 5.2, 5.3, and 5.4 in which reweighted Z' signal templates for the SSM, the E_6 model, and the Minimal model were generated for several model parameters of interest.

The reweighting procedure was extensively validated by comparing sets of reweighted

signal templates from the `LPXSignalReweightingTool` to their dedicated MC counterparts.

Figure 11.1 shows the electron channel validation plots for E_6 model Z'_χ at pole masses of 2 and 3 TeV.⁷ An excellent level of agreement can be seen between the dedicated and reweighted signal templates, especially in the region surrounding the peak.

11.2.2 Interference Templates

The previous strategy can be extended by taking into account the full quantum mechanical interference between the γ^*/Z Drell-Yan and Z' signal processes, which results in an **interference template**. Like before, this amounts to modifying the differential cross section, such that

$$\frac{d\sigma}{d\hat{t}}(q\bar{q} \rightarrow \gamma^*/Z \rightarrow e^+e^-) \rightarrow \frac{d\sigma}{d\hat{t}}(q\bar{q} \rightarrow \gamma^*/Z/Z' \rightarrow e^+e^-). \quad (11.10)$$

Here, the interference modification applies only to the numerator of the event weight of Eq. 11.7, so that now,

$$w_{RW}(M_{Z'}) = \frac{\sum_{i,j \in \{L,R\}} |\mathcal{M}_{ij}(\gamma^*/Z/Z')|^2}{\sum_{i,j \in L,R} |\mathcal{M}_{ij}(\gamma^*/Z)|^2}. \quad (11.11)$$

The substitution does not affect the expressions for the matrix elements of Eq. 11.8, rather the form of the factorized amplitude is extended, where

$$\mathcal{G}_{ij}(\hat{s}) = \begin{cases} e^{2\frac{Q^l Q^q}{\hat{s}}} + g_Z^2 \frac{g_i^l g_j^q}{\hat{s} - M_Z^2 + iM_Z\Gamma_Z}, & \text{for } \gamma^*/Z \\ e^{2\frac{Q^l Q^q}{\hat{s}}} + g_Z^2 \frac{g_i^l g_j^q}{\hat{s} - M_Z^2 + iM_Z\Gamma_Z} + g_{Z'}^2 \frac{z_i^l z_j^q}{\hat{s} - M_{Z'}^2 + iM_{Z'}\Gamma_{Z'}}, & \text{for } \gamma^*/Z/Z'. \end{cases} \quad (11.12)$$

⁷Similar validation plots were generated for the muon channel, as the reweighting technique is identical.

The dilepton analysis applies this extended method to the generation of signal templates for the Minimal model, which provides for a more general interpretation of the search results than the SSM or E_6 models. In the later cases, σB is taken as the parameter of interest, while for the Minimal model, the parameters of interest is the relative coupling strength γ' for a fixed mixing angle θ_{Min} (See Sec. 5.3.4). This means limits on these model parameters can be directly translated into limits on σB for entire classes of Z' models owing to the more general phenomenology.

The Minimal model Z' gauge coupling and fermion quantum numbers are determined by

$$\begin{aligned} g_{Z'} &= \gamma' g_Z \\ z_f &= Y_f \sin \theta_{\text{Min}} + (B - L)_f \cos \theta_{\text{Min}} \end{aligned} \tag{11.13}$$

where z_f is the fermion $U(1)'$ charge with its left- or right-handed chiral index is suppressed, and Y and $B - L$ are the Standard Model fermion weak hypercharge and baryon number minus lepton number, respectively.

With two independent model parameters $(\gamma', \theta_{\text{Min}})$, it is not sufficient to generate signal templates as a function of the invariant mass m_{ee} alone for each Z' pole mass of interest. Instead, the template definition is extended to two dimensions, so that the Minimal model templates are generated as a function of both m_{ee} and γ' for fixed values of Z' pole mass $M_{Z'}$ and mixing angle θ_{Min} .

Figure 11.2 shows an example of these combined two-dimensional templates, which account for the interference between the γ^*/Z and Z' processes and assume a mixing angle of $\theta_{\text{Min}} = 0$ and pole mass of $M_{Z'} = 3$ TeV. For each Minimal model template, the decay width $\Gamma_{Z'}$ is recalculated for each value of γ' considered during the template construction.

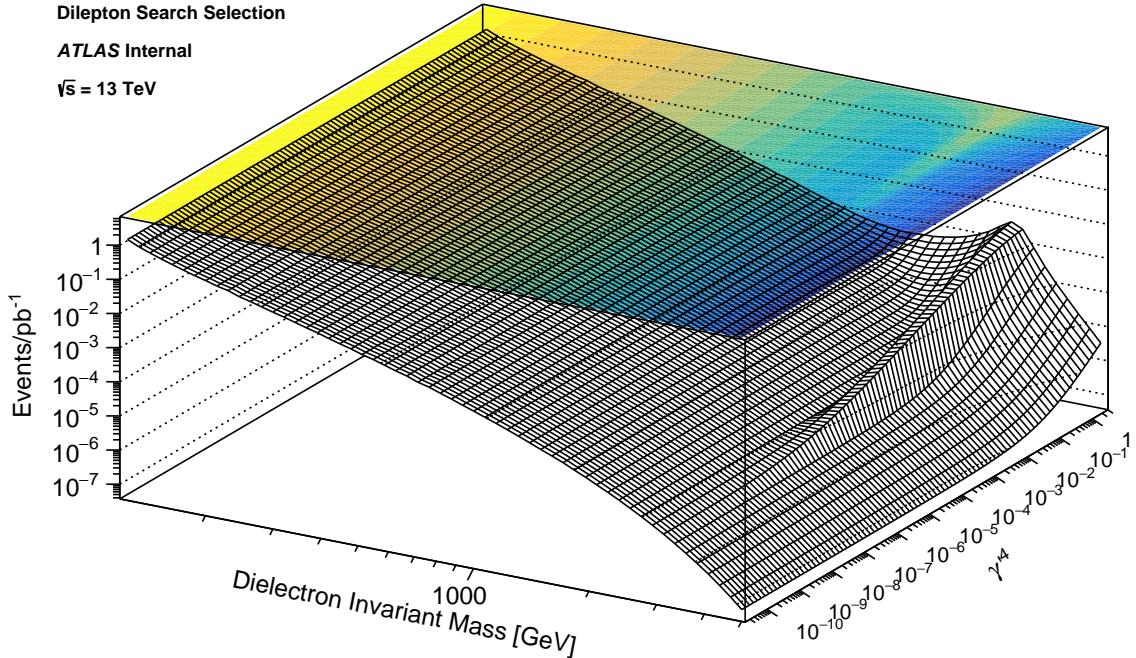


Figure 11.2: Example Minimal model signal template with fixed mixing angle $\theta_{\text{Min}} = 0$ and Z' pole mass $M_{Z'} = 3 \text{ TeV}$. The template is a function of dielectron mass m_{ee} and the fourth power of the Minimal model relative coupling strength $\gamma' = g_{Z'}/g_Z$, and is generated from reweighted PYTHIA Drell-Yan events after passing the full electron channel event selection.

This ensures the signal cross section evolves with the value taken by the decay width $\Gamma_{Z'}$ in Eq. 11.12, as required when adjusting the coupling strength between the fermions and Z' boson.

11.3 Signal Search

The search for a signal-like excess is conducted by scanning the dielectron invariant mass spectrum for local deviations in the observed event yield relative to the Standard Model background expectation. Two complementary search methods are used to interpret the significance of any observed excesses in the data. For each method, the significance of

an observed excess is extracted with a **p-value** p_0 , the probability of observing a result equal to or more extreme than the one actually observed, according to the background-only hypothesis.⁸

In the first method, a log-likelihood ratio (LLR) test is performed, which assesses the compatibility of the observed dataset under a so-called **null hypothesis** H_0 as compared to a specific **alternative hypothesis** H_1 . In the dilepton analysis, H_0 refers to the prediction of the Standard Model, while H_1 refers to the Beyond the Standard Model (BSM) prediction of the SM+ Z' , evaluated for a given Z' model and pole mass $M_{Z'}$ as described by Eq. 11.2.

The LLR test is performed multiple times, once for each Z' pole mass under investigation, since the pole mass of a hypothetical Z' boson is unknown *a-priori*. This requires the use of individual Z' signal templates (See Sec. 11.2). The LLR test results in a **local** p_0 , or a p_0 associated with a specific mass hypothesis $M_{Z'}$ under investigation.

The second search method relies on the BUMPHUNTER (BH) test [199], which scans all possible contiguous intervals in the dielectron invariant mass spectrum, and identifies the one where the data maximally deviates from the Standard Model prediction; i.e., it locates the most significant “bump.” Based on this interval, it extracts the corresponding p-value according to Poisson statistics. Unlike the LLR test, the BH test does not require an explicit signal model to be assumed for the test to be carried out, making it a fully model-independent approach.

⁸A p-value is the probability of obtaining a test statistic t equal to or more extreme than the one observed in the actual data t_0 , assuming the null hypothesis H_0 is true: $p_0 = p(t \geq t_0 | H_0)$.

11.3.1 Log-Likelihood Ratio Test

The LLR test is a statistical test which compares how well two different statistical models fit a set of observations. For the dilepton analysis, the comparison is made between the Standard Model (i.e. the null hypothesis H_0) to a possible BSM signal model (i.e. the alternative hypothesis H_1). The signal hypothesis H_1 consists of the combined background plus signal expectation SM+ Z' according to Eq. 11.2 for a given pole mass of interest, while H_0 consists of the Standard Model expectation only. By construction, the Standard Model hypothesis is a special, or **nested**, case of the BSM hypothesis, since setting μ equal to zero in Eq. 11.2 fully recovers the Standard Model hypothesis H_0 .

The LLR test is performed as a function of Z' pole mass; i.e., multiple times in discrete steps, which requires the use of Z' signal templates. A set of over 500 Z'_χ templates is used, which are generated without interference and range in pole mass from 120 GeV to 5 TeV. The Z'_χ model has a decay width $\Gamma_{Z'}$ that is 1.2% the value of its pole mass $M_{Z'}$.

The templates are constructed with a specialized dielectron invariant mass distribution, which utilizes an optimized histogram binning based on the dielectron invariant mass resolution of the ATLAS detector. In this case, each signal hypothesis, indexed by its pole mass $M_{Z'}$, is tested in steps that correspond to the spacing between histogram bins. The fine spacing between $M_{Z'}$ mass-points ensures an adequate scan is taken over dielectron invariant mass spectrum for potential signal-like excesses, and is what gives rise to the over 500 signal templates used in the scan.

The use of the LLR test is based on the Neyman-Person lemma [200], which states that this is most powerful test to reject H_0 in favor of H_1 . To test for excesses in the data, the LLR test is carried out using the HISTFACTORY [201] package, along with the RooStats [202]

and RooFit [203] packages. The LLR **test statistic** q_0 is used to extract the probability p_0 that the Standard Model background will fluctuate up to its observed value in data or greater.⁹ The p-value p_0 for each mass hypothesis is computed using the following test statistic, the logarithm of the profile likelihood ratio,

$$q_0 = \begin{cases} 0, & \text{for } \hat{\mu} < 0 \\ -2 \ln \left[\frac{\mathcal{L}(0, \hat{\theta}_0)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \right], & \text{for } \hat{\mu} \geq 0, \end{cases} \quad (11.14)$$

where $\mathcal{L}(\mu, \boldsymbol{\theta})$ is the likelihood function of Eq. 11.1, and μ and $\boldsymbol{\theta}$ are the signal strength and nuisance parameters, respectively.

The numerator defines the Standard Model hypothesis, where $\hat{\theta}_0$ is the value of $\boldsymbol{\theta}$ that maximizes \mathcal{L} under the condition that $\mu = 0$; i.e., in accordance with the background-only hypothesis H_0 . The best-fit values, or most probable values for a given dataset, are obtained with the **Maximum Likelihood** (ML) method.¹⁰ Likewise, the denominator defines the alternative hypothesis H_1 , where the ML estimates obtained by the likelihood are $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$.

As a consequence, higher values of q_0 correspond to increasing incompatibility between the data and the null hypothesis H_0 . For a given dataset, the numerator is always smaller than or equal to the denominator, so the ratio is always less than or equal to one. Thus, the more significant an excess, the larger the difference between the respective likelihoods in Eq. 11.14.

In the case of a downward fluctuation in data, the ML estimate for the signal strength $\hat{\mu}$

⁹The test statistic is a scalar number that quantifies the difference between hypotheses H_0 and H_1 for a given dataset.

¹⁰This method finds the values of the parameter of interest μ and nuisance parameters $\boldsymbol{\theta}$ that provide the likelihood function with a global maximum. These values are known as maximum likelihood estimators and are denoted with a hat; i.e. $\hat{\mu}$ is the maximum likelihood estimate for μ .

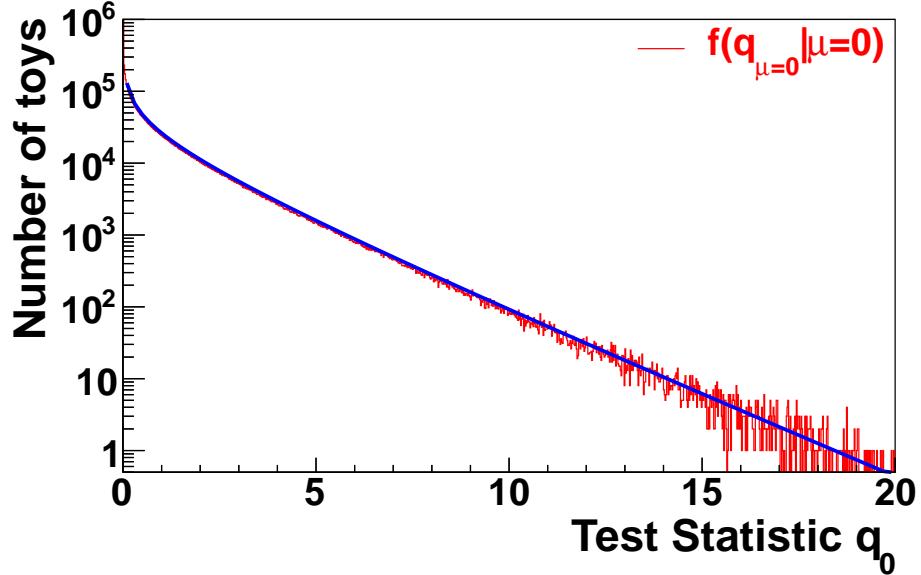


Figure 11.3: Sampling distribution $f(q_0|0)$ for the LLR test statistic q_0 . The distribution is generated from background-only pseudo-experiments and is used to extract the p-value associated with individual Z' signal hypotheses. The distribution follows a χ^2 function with one degree of freedom according to Wilk's theorem [5].

can become negative. In this case the test statistic q_0 of Eq. 11.14 is set to zero by definition, associated with the fact that the number of signal events is a strictly positive quantity. This definition leads to a peak at $q_0 = 0$ in the **sampling distribution** $f(q_0|0)$ of the test statistic, which represents the probability that a background-only dataset will produce a downward fluctuation, a situation that occurs 50% of the time.

The sampling distribution $f(q_0|0)$ is used to extract the p-value associated with each signal hypothesis under investigation. The sampling distribution $f(q_0|0)$ is built by generating background-only **pseudo-experiments**, or toys, collections of simulated event counts drawn from the Standard Model expectation that corresponds to the integrated luminosity of the actual dataset. An individual pseudo-experiment is obtained by sampling the number of expected background events in each mass bin k of the dielectron invariant mass distribution according to Poisson statistics. The test statistic q_0 is recalculated and binned in a

histogram, thereby building up $f(q_0|0)$ one sample at a time, as shown in Figure 11.3.

In the large sample limit, the distribution $f(q_0|0)$ follows a χ^2 distribution for one degree of freedom. This is the result of **Wilk's theorem** [204], which states the distribution of the LLR test statistic follows a χ_k^2 distribution, as the sample size approaches infinity, where k is the difference in the number of degrees of freedom between H_1 and H_0 . Since the two models are nested by construction, the difference in free parameters k is equal to one. The resulting χ^2 fit can be seen in Figure 11.3 as the solid line. Thus, the p-value can be extracted directly without the need to generate an ensemble of pseudo-experiments, greatly reducing the required amount of computing [205].

The probability p_0 associated with the actual experimental observation q_0^{obs} is calculated from

$$p_0 = p(q_0 \geq q_0^{obs} | H_0) = \int_{q_0^{obs}}^{\infty} f(q_0|0) dq_0, \quad (11.15)$$

using the ensemble. Further, the p-value p_0 is usually translated to significance scale z based on the number of Gaussian standard deviations,

$$z = \Phi^{-1}(1 - p_0). \quad (11.16)$$

The function $\Phi^{-1}(x)$ is the inverse of the cumulative Gaussian distribution. If the observed p-value falls below a predefined critical threshold, the data are unlikely to have been realized under H_0 , and therefore H_0 is rejected in favor of H_1 . In High Energy Physics, a p_0 smaller than $1.35 \cdot 10^{-3}$ constitutes evidence for a signal, while a p_0 less than $2.87 \cdot 10^{-7}$ formally constitutes a discovery. These two thresholds are derived from the one sided integral of a unit Gaussian distribution, where **evidence** corresponds to the value of the integral beyond $+3\sigma$ and **discovery** to beyond $+5\sigma$. In this sense, the p-value can be interpreted as the

false discovery probability [199].

The above machinery is used to extract **local p-values**, those that correspond to a specific signal hypotheses, localized in Z' pole mass. However, as many Z' mass hypotheses are tested, it becomes increasingly likely that at least one should produce a significant result. This is known as the **look-elsewhere-effect**, and can be accounted for by calculating the so-called **global p-value**. The global p-value can be extracted from an ensemble of background-only pseudo-experiments, where, for each pseudo-experiment, the value of the most significant local fluctuation z_0 in the dielectron spectrum is used to build up the sampling distribution $n(z_0|0)$ of largest local significances.

Thus, the global p-value can be obtained from

$$p_{\text{global}} = p(z_0 \geq z_0^{\text{obs}} | H_0) = \int_{z_0^{\text{obs}}}^{\infty} n(z_0|0) dz_0. \quad (11.17)$$

where z_0^{obs} is the value of the largest observed local significance in the actual dataset.

The data were scrutinized for a signal, but no significant excess was found. Figure 11.4 shows the extracted local and global p-values as a function of Z' pole mass for the electron channel search. The observed global p-value is $p_{\text{global}} = 58\%$, meaning more than half of background-only pseudo-experiments produce an excess as significant as the one observed in data.

The smallest local p-value is 0.6%, which translates to an excess of 2.5σ , and is observed for a Z' pole mass of 2.37 TeV. When compared to the global significance scale, however, this excess falls below the value obtained by the majority of background-only experiments, and is therefore not significant. As such, the observed data are consistent with Standard Model hypothesis, and, at least in this analysis, the Standard Model cannot be rejected in

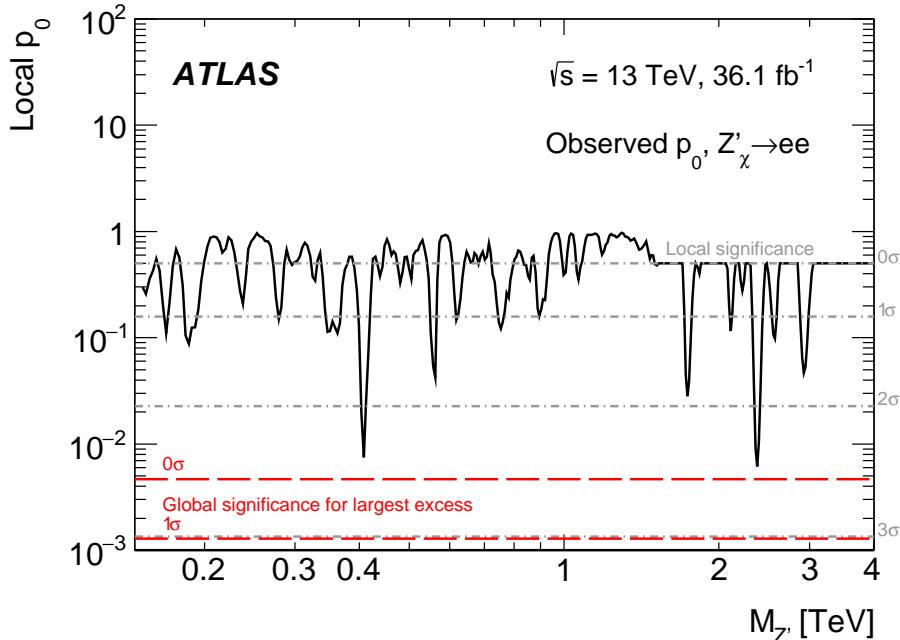


Figure 11.4: The resulting p_0 scan in the electron channel search region, derived using Z_χ signal templates with pole masses between 0.15 and 3.0 TeV. The local p-values are depicted as the black curve, where the results for individual mass points have been interpolated. The corresponding local and global significance levels are depicted as gray and red dashed lines, respectively. The lack of sufficient statistics at high mass can result in an ill defined PDF and negative probabilities. Therefore, above 1.5 TeV the so-called “capped” p-value is used, while the “un-capped” value is used below. See Ref. [4] for more details.

favor of any Z' signal hypothesis.

Lastly, it should be noted that a small unavoidable level of model-dependence exists in the definition of p_0 , since a particular signal scenario must be assumed for the calculation of q_0 in Eq. 11.14. In the dilepton search, the theoretically well-motivated Z'_χ is assumed, as a benchmark, and therefore, the observed p-values are extracted in the context of this model. To account for this model-dependence, the complimentary model-independent BUM-PHUNTER test is also performed.

11.3.2 BumpHunter Test

The BUMPHUNTER [199] (BH) is a tool that can be used to find the largest localized excess in data, without the need to assume a particular signal model. Therefore, the BUMPHUNTER is sensitive to a broader range of new physics models, as compared to the LLR test, which is designed to test for resonant Z' production.

The BUMPHUNTER tool makes use of the BH test, which scans all possible contiguous bins in the dielectron invariant mass distribution and identifies the one with the most significant disagreement between the data and the Standard Model background expectation. For each window searched, the significance is identified with Poisson p-value by computing the number of events equal to or greater than the observed event count as compared to the expected number under the null hypothesis H_0 of the Standard Model, such that

$$p_0 = p(n \geq n_{obs} | H_0) = \sum_{k=n_{obs}}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!}, \quad (11.18)$$

where n is the number of observed events and λ is the Standard Model expectation in the window being tested. In the dilepton search, the smallest window scanned ranges from two bins to half the total number of bins in the invariant mass distribution. The results of the electron channel scan can be seen in Figure 11.5a, where the search intervals range from 120 GeV to 3 TeV.

The window with the smallest p-value, or most significant excess, is used to construct the BH test statistic, defined as the negative logarithm of the smallest p-value identified in the scan,

$$t = -\log \left[p_0^{\min} \right]. \quad (11.19)$$

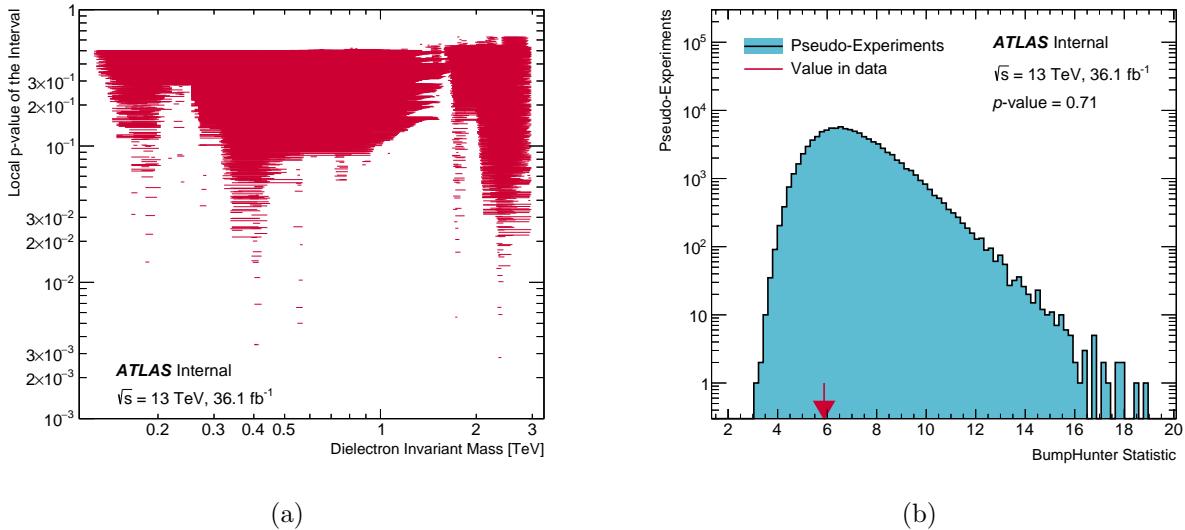


Figure 11.5: Results of BUMPHUNTER test for the search range $m_{ee} \in [0.12, 3.00]$ TeV. Depicted are the local p_0 of each interval searched (a), and the global p_0 (b), which is extracted from an ensemble of background-only psuedo-experiments.

This statistic is used to extract a global p-value of the search. As with the LLR test, the sampling distribution for the BH test statistic is built up using pseudo-experiments, where pseudo-data is sampled from the H_0 background estimate of the Standard Model, and the observed BH statistic is then compared to the ensemble. The global p-value is the fraction of background-only pseudo-experiments where the BH test statistic is greater than or equal to the value observed in data, whose value is denoted by the arrow in Figure 11.5b. As can be seen, the majority of pseudo-experiments have a BH statistic larger than the value observed in data, indicating the observed data are consistent with the Standard Model.

As a result of the scan, the global p-value is found to be 71% in the electron channel. In Figure 11.6, the observed data with accompanying Standard Model background prediction is depicted, along with the bin-by-bin local significance in ratio sub-panel. The interval identified as having the most discrepant upward fluctuation is indicated by the pair of blue lines, which corresponds to a set of mass bins around 2.4 TeV, the same region identified by

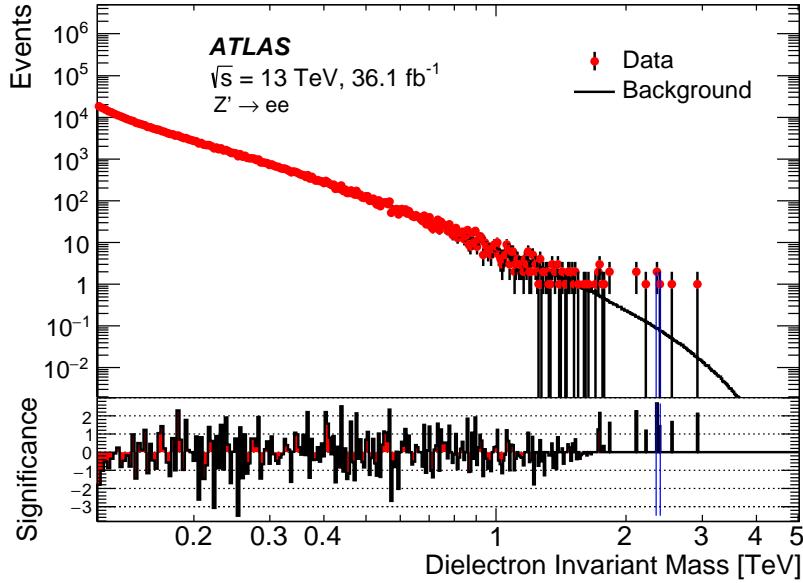


Figure 11.6: The dielectron mass distribution, showing the observed data, the Standard Model background prediction, and the corresponding bin-by-bin local significance. The extremely fine search binning is clearly visible in this plot. The interval identified by the pair of blue lines corresponds to the most significant local excess [5].

the LLR test.

As with the LLR test, the mass spectrum scrutinized according to the BH test exhibits no characteristic signs of new physics. In this case, exclusion limits are placed on various model parameters of interest, such as production cross sections and coupling scenarios for the Z' models considered in this analysis, the subject of the following section.

11.4 Exclusion Limits

The signal search determined the data were consistent with the Standard Model, which rules out any alternative explanations in terms of new physics. In the absence of a significant excess, **exclusion limits**¹¹ are set at 95% C.L. using a Bayesian approach on several Z'

¹¹An exclusion limit identifies a region in the model parameter space that is still consistent with the data.

model parameters of interest.

For the SSM and E_6 models, limits are placed on $\sigma B(Z' \rightarrow e^+e^-)$, the Z' cross section times branching fraction to electron pairs. Similarly for the Minimal model, limits are placed on the relative coupling strength γ' , which controls the overall strength of the Z'_{Min} interaction.

The limit setting procedure, like the signal search, relies on the binned likelihood function of Eq. 11.1. For the limit setting, however, the nominal parameter of interest μ is re-expressed as the cross section times branching fraction σB for the extraction of the SSM and E_6 model limits, and as the relative coupling strength γ' for the extraction of the Minimal model limits.

Furthermore, like the search, the use of Z' signal templates is continued here as well. A total of 118 signal templates is considered for the limit setting. Each is generated without interference, and range in pole mass from 150 GeV to 5 TeV in 50 GeV increments. The templates are a function of the Z' pole mass, and are defined over the standard log-binned dielectron invariant mass distribution of Figure 10.6, and include the Z boson mass peak in their definition as well. This allows for the nuisance parameters associated with each mass hypothesis to be maximally constrained in the limit setting.

11.4.1 Bayesian Inference

In the dilepton analysis, exclusion limits are calculated using **Bayes' theorem** [206],

$$p(H|D) = \frac{p(D|H)p(H)}{p(D)}, \quad (11.20)$$

which relates $p(H|D)$, the probability that a hypothesis H is true given the observed data D , to $p(D|H)$, the probability of observing the data given that the hypothesis is true, and

to the individual probabilities for the hypothesis $p(H)$ and data $p(D)$ separately.¹²

Each term in Eq. 11.20 has an interpretation. The **likelihood** $p(D|H)$ is the evidence for the hypothesis provided by the observed data, which is identified with the likelihood function of Eq. 11.1. The **prior probability**, or just the prior, $p(H)$ is the probability that the hypothesis is true before the data are taken. The **evidence** $p(D)$ is the total probability of the data, accounting for all possible hypotheses, and acts as a normalization constant. The **posterior probability**, or just posterior, $p(H|D)$ is the probability that the hypothesis is true after taking the data. The posterior quantifies the impact the data has on the prior, so Bayes' theorem specifies how to update the probability of a certain hypothesis as new evidence becomes available.

Bayesian probability requires the specification of a prior. In the dilepton analysis, the prior is assumed to be **flat**, or uniform and positive, such that, for example, the prior on the signal cross section is $p(\sigma B) = 1$ for $\sigma B \geq 0$ and 0 otherwise. This choice leaves the statistical interpretation as general as possible. With a well-defined prior and likelihood, a posterior can be constructed for each signal hypothesis under investigation using Bayes' theorem of Eq. 11.20. Then, the posterior can be *integrated* to set exclusion limits at 95% C.L. on the Z' model parameter of interest; i.e., σB or γ' .¹³

11.4.2 Limits on SSM and E_6 Model Z' Bosons

Since the data were found to be consistent with the Standard Model, exclusion limits are set on the existence of Z' bosons. The following section presents limits for the SSM and E_6

¹²In the Bayesian statistics, probability is interpreted as a measure of degree of belief in a proposition or hypothesis, such as the existence of a Z' boson with a certain pole mass. Frequentist statistics, on the other hand, interprets probability as the relative frequency of a certain event occurring in a large number of trials.

¹³In the Bayesian approach, a credibility interval is the range of values in the posterior that includes 95% of the total probability.

models, where upper limits are placed on the signal cross section times branching fraction σB for Z' production.¹⁴

Exclusion limits are extracted as a function of Z' pole mass, meaning individual Z' signal templates are used to model each mass point under investigation. A total of 118 Z' mass points are tested using signal templates for the individual SSM and E_6 models considered. The templates themselves are functions of the dielectron invariant mass m_{ee} (as in Fig. 11.1) and have a Z' pole mass $M_{Z'}$ that ranges from 150 GeV to 5 TeV in 50 GeV steps.

As the limit calculation relies on the binned likelihood function of Eq. 11.1, which is nominally in terms of the parameter of interest μ , a change of variables must be made in order to set limits directly on σB . Therefore, μ is replaced by $\mu = (\sigma B)/(\sigma B)_{Z'}$, the signal cross section times branching fraction relative to its theoretical expectation, effectively trading one floating signal normalization μ for another (σB) .

With this change, the posterior is obtained via Bayes' theorem,

$$p(\sigma B, \boldsymbol{\theta} | \mathbf{n}) = \frac{\mathcal{L}(\mathbf{n} | \sigma B, \boldsymbol{\theta}) p(\sigma B)}{p(\mathbf{n})}. \quad (11.21)$$

The likelihood $\mathcal{L}(\mathbf{n} | \sigma B, \boldsymbol{\theta})$ is a function of σB and $\boldsymbol{\theta}$, the vector of electron channel nuisance parameters.¹⁵ The prior on the signal is $p(\sigma B)$, which is uniform for $\sigma B \geq 0$ and is equal to 0 otherwise and the prior on the data is $p(\mathbf{n})$, which is independent of σB . The vector \mathbf{n} parametrizes the number of observed data events in each mass bin k of the dielectron invariant mass distribution.

In order to formulate the posterior strictly in terms of the parameter of interest σB , the

¹⁴Upper limits define the region in Z' parameter space that is still consistent with the observed dataset; or, equivalently, they exclude the region that is now inconsistent.

¹⁵Systematics that contribute at the level of 1% or less over the dielectron invariant mass search region are not included, since it is computationally expensive and does not affect the final result.

nuisance parameters $\boldsymbol{\theta}$ of the likelihood function must be integrated over. This dependence can be removed numerically with Markov Chain Monte Carlo (MCMC), which is performed with the Bayesian Analysis Toolkit (BAT) [207]. This process of integrating out the nuisance parameters is known as **marginalization**, and the marginalized likelihood is

$$\mathcal{L}'(\mathbf{n}|\sigma B) = \int \mathcal{L}(\mathbf{n}|\sigma B, \boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (11.22)$$

With that, the posterior on σB alone becomes

$$p(\sigma B|\mathbf{n}) = \frac{\mathcal{L}'(\mathbf{n}|\sigma B)p(\sigma B)}{p(\mathbf{n})}. \quad (11.23)$$

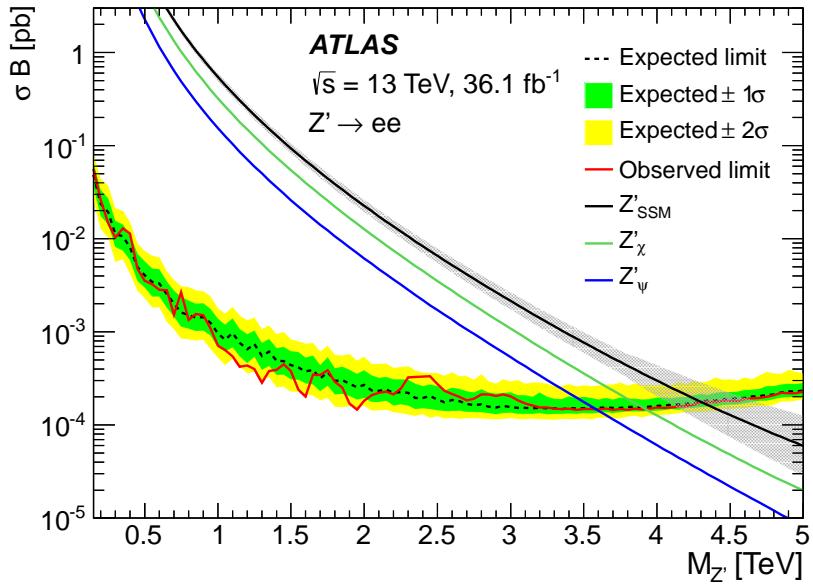
The 95% C.L. upper limits on σB can be obtained by integrating posterior $p(\sigma B|\mathbf{n})$; the maximum value corresponds to most likely signal excess in data. Exclusion limits on σB are obtained such that,

$$0.95 = \frac{\int_0^{(\sigma B)_{95}} p(\sigma B|\mathbf{n}) d(\sigma B)}{\int_0^{\infty} p(\sigma B|\mathbf{n}) d(\sigma B)} \quad (11.24)$$

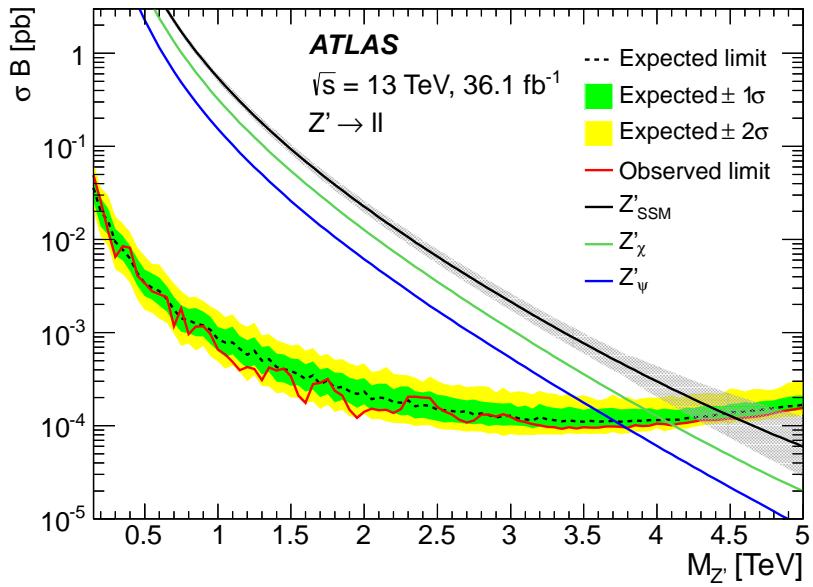
where the value $(\sigma B)_{95}$ is the value of the parameter of interest for which the 95% C.L. upper limit is satisfied. This integration is also performed using BAT.

The upper limits on σB as a function of Z' pole mass are depicted in Figure 11.7, along with the Z' production cross sections for the Z'_{SSM} , and Z'_{ψ} and Z'_{χ} of the E_6 model. The **observed limits**, or those extracted with the observed 36.1 fb^{-1} dataset, are represented by the solid red line. These tend to fluctuate due to the local excesses and deficits in the observed data of the dielectron invariant mass distribution. The parameter space of models that predict a σB above the red line are excluded at 95% C.L.

The gray band on the Z'_{SSM} theory line depicts the total theoretical uncertainty on the



(a)



(b)

Figure 11.7: The 95% C.L. upper limits on the Z' production cross section times branching fraction for the electron channel (a) and the combined electron plus muon channel (b) as a function of Z' pole mass. The expected limit is depicted as a black dashed line, with yellow and green bands indicating $\pm 1\sigma$ and $\pm 2\sigma$ variations from the median value. The observed limit is shown as a solid red line and the above region is excluded at 95% C.L. Signal theoretical uncertainties are shown on the SSM cross section prediction as a gray band for reference, but are not included as part of the limit calculation [5].

signal, but it is not included as part of the limit calculation. This uncertainty can become significant at high mass, where it is due almost exclusively to the imprecise knowledge of the Parton Distribution Functions (PDFs). A strategy to reduce this uncertainty in future iterations of the analysis is presented in Section 10.1.

The **expected limits**, or those extracted under the assumption that the Standard Model is true, are also shown, which are marked by a dotted-dashed black line.¹⁶ The expected limits are obtained by generating 10,000 background-only pseudo-experiments (pseudo-data invariant mass distributions) for each Z' mass point tested. The pseudo-data are sampled according to the Poisson expectation in each m_{ee} mass bin k and the Gaussian parameterizations of the systematic uncertainties. For each pseudo-experiment, the 95% C.L. upper limit on $(\sigma B)_{95}$ is recalculated via Eq. 11.22, and the resulting set of values is used build a distribution for $(\sigma B)_{95}$. The median value of the distribution defines the expected limit, and the green and yellow bands define the $\pm 1\sigma$ (68%) and $\pm 2\sigma$ (95%) quantiles. This procedure is carried out for each mass point tested and the limit plot is built up point-by-point.

Just as the Z'_χ was used in p_0 scan for the LLR test, it is similarly used here as the benchmark model to extract the limits, meaning it is the signal model assumed in the likelihood definition. In addition, the same signal acceptance (as in Fig. 10.3) is assumed for all Z' models tested. This approximation works well since interference is neglected and the models here have similar decay widths.

Above a mass of roughly 3.5 TeV, the upper limits start to weaken. This is due primarily to the quickly falling production cross section, as the kinematic limit is reached: signal production is occurring increasingly off-shell in the low-mass tail, as opposed to in the width of

¹⁶“Expected” limit is somewhat of a misnomer as it means nothing new or interesting is expected. A better term might be “Standard Model” limit.

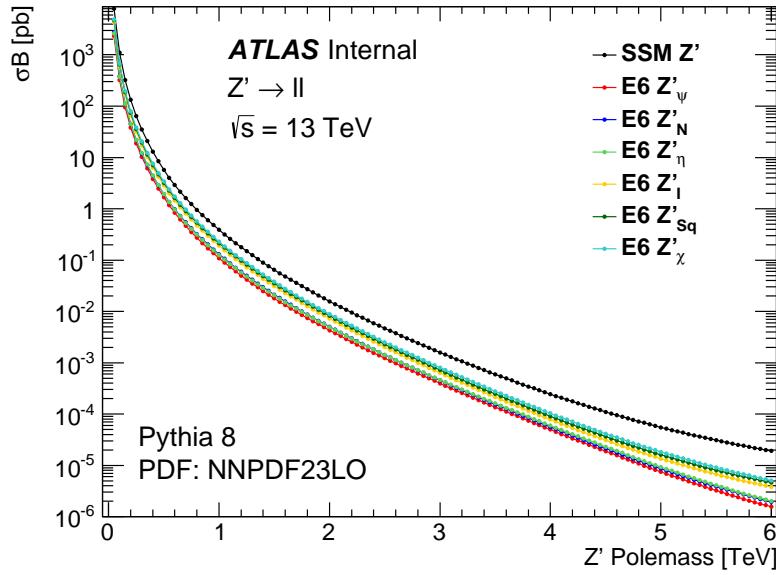


Figure 11.8: The LO cross section times branching fraction predictions for Z' production in the SSM and E_6 models considered in the analysis. Values were obtained using PYTHIA8 configured with the NNPDF2.3 PDF set, and are plotted as a function of Z' pole mass for $\sqrt{s} = 13$ TeV [5].

the resonance (compare Fig. 10.4), which subsequently leads to a drop in selection efficiency (compare Fig. 10.3).

The Z' pole mass at which the theoretical value of σB intersects the observed limit defines a lower limit on mass. The theoretical cross section for each model as a function of mass can be seen in Figure 11.8. In the electron channel, the lower mass limits range from 3.6 TeV for the Z'_ψ to 4.3 TeV for the Z'_{SSM} . Table 11.1 lists the lower mass limits for the SSM and each of the E_6 models considered in the analysis.

Finally, in order to set the strongest exclusion limits possible, the individual electron and muon channels are statistically combined using BAT.¹⁷ For this, the likelihood of Eq. 11.1 is extended to include the muon channel, incorporating the observed dimuon data and signal and background expectations in all dimuon invariant mass bins, as well as the muon chan-

¹⁷This is done assuming lepton universality, which applies for the signal models considered in this analysis.

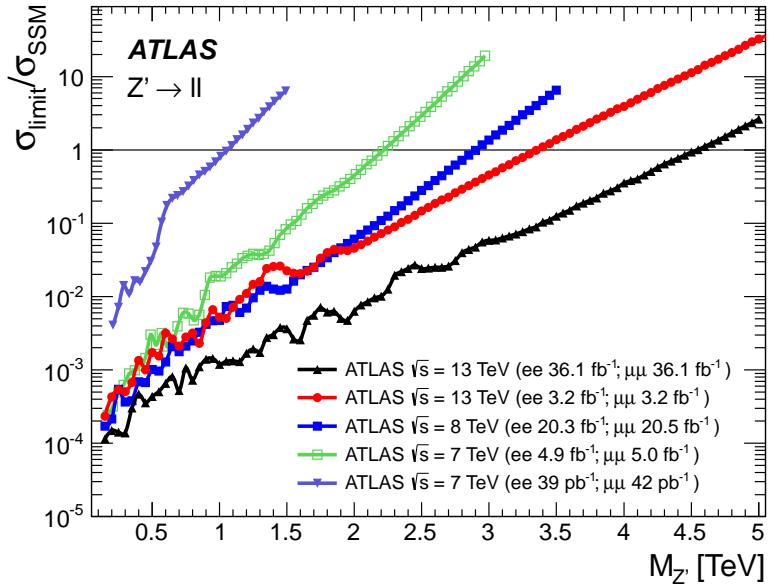


Figure 11.9: Ratio of the observed 95% C.L. upper limit to the theoretical prediction for the Z'_{SSM} cross section times branching fraction as a function of Z' pole mass for the combined electron and muon channels. Previous iterations of the ATLAS dilepton search, performed at $\sqrt{s} = 7, 8$, and 13 TeV, are compared to the current result obtained here, which is shown in black. The region above each line is excluded at 95% C.L. [5].

Model	Width [%]	θ_{E_6} [rad]	Lower limits on $M_{Z'}$ [TeV]					
			ee		$\mu\mu$		$\ell\ell$	
			Obs	Exp	Obs	Exp	Obs	Exp
Z'_{SSM}	3.0	-	4.3	4.3	4.0	3.9	4.5	4.5
Z'_χ	1.2	0.50π	3.9	3.9	3.6	3.6	4.1	4.0
Z'_S	1.2	0.63π	3.9	3.8	3.6	3.5	4.0	4.0
Z'_I	1.1	0.71π	3.8	3.8	3.5	3.4	4.0	3.9
Z'_η	0.6	0.21π	3.7	3.7	3.4	3.3	3.9	3.8
Z'_N	0.6	-0.08π	3.6	3.6	3.4	3.3	3.8	3.8
Z'_ψ	0.5	0π	3.6	3.6	3.3	3.2	3.8	3.7

Table 11.1: Expected and observed 95% C.L. lower mass limits for the SSM and E_6 Z' bosons considered in the analysis. Electron, muon, and combined channel limits are shown [5].

nel systematic uncertainties (See Ref. [5] for a complete list). The multichannel likelihood becomes

$$\mathcal{L}(\mathbf{n}|\sigma B, \boldsymbol{\theta}) = \prod_{l=1}^{N_{\text{channel}}} \prod_{k=1}^{N_{\text{bins}}} \frac{\lambda_{lk}(\sigma B, \boldsymbol{\theta}) e^{-\lambda_{lk}(\mu, \boldsymbol{\theta})}}{n_{lk}!} \prod_{i=1}^{N_{\text{sys}}} \mathcal{N}(\theta_i, 0, 1), \quad (11.25)$$

which, following the same procedure as before, is marginalized, and used to construct a posterior from which the associated combined channel limits are extracted at 95% C.L. The combined upper limits on σB are similarly depicted in Figure 11.7, and the combined lower limits on mass can be seen in Table 11.1. In addition, Figure 11.9 compares the exclusion limits obtained for the Z'_{SSM} in this analysis to several previous iterations of the ATLAS dilepton search [111, 123, 208, 209].

11.4.3 Limits on Minimal Model Z' Bosons

For the Minimal model, exclusion limits are set directly on the relative coupling strength γ' of the Z'_{Min} boson relative to the Standard Model Z boson, as a function of both the Z'_{Min} pole mass and mixing angle θ_{Min} .

This requires the use of two-dimensional signal templates (See Sec. 11.2.2), which are generated as a function of both the dielectron invariant mass m_{ee} and γ' for fixed values for θ_{Min} and $M_{Z'}$ (See Fig. 11.2). As with the SSM and E_6 limit setting of the previous section, 118 signal templates are utilized, corresponding to the individual Z' pole masses tested.

The use of two-dimensional templates requires a modification to the nominal definition of the likelihood function of Eq. 11.1, since the expected number of events and systematic variations are now extended to depend on both m_{ee} and γ' . Otherwise, the limit setting procedure is identical to the approach of the previous section.

In order to modify the likelihood of Eq. 11.1, the signal expectation $s_k(\boldsymbol{\theta})$ in each m_{ee} mass bin k is replaced by a non-linear function $f_k(\gamma', \boldsymbol{\theta})$, which encodes the Minimal model signal expectation as a function of the relative coupling γ' . As such, the expected number of dielectron events becomes

$$\lambda_k(\gamma', \boldsymbol{\theta}) = f_k(\gamma', \boldsymbol{\theta}) + b_k(\boldsymbol{\theta}). \quad (11.26)$$

In this formalism, each function f_k is defined for a fixed pole mass $M_{Z'}$ and mixing angle θ_{Min} . Furthermore, this replacement fully incorporates the $\gamma^*/Z/Z'$ interference effect, as well as the functional dependence of the Z' decay width $\Gamma_{Z'}$ on the relative coupling γ' .

The functions f_k are obtained from a series of interpolations, performed in each m_{ee} mass bin k of the two-dimensional signal templates (as in Fig. 11.2). Since the Standard Model expectation b_k in Eq. 11.26 already contains the Drell-Yan background estimate, it is isolated and subtracted from each function f_k to avoid double counting. Thus, each function f_k consists of a pure Z' signal component and the $\gamma^*/Z/Z'$ interference component. For specific choices of θ_{Min} and γ' , the signal expectation can be negative, which results from the $\gamma^*/Z/Z'$ interference. In the limiting case $\gamma' \rightarrow 0$, the Standard Model expectation is exactly recovered as the function $f_k \rightarrow 0$.¹⁸

With these changes, the marginalized likelihood becomes,

$$\mathcal{L}'(\mathbf{n}|\gamma') = \int \mathcal{L}(\mathbf{n}|\gamma', \boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (11.27)$$

Bayes' theorem is used with the prior $p(\gamma')$, which is constant as a function of $\gamma'^4 > 0$ and 0 otherwise. This choice of γ'^4 relies on the fact that the pure Z'_{Min} component of the total

¹⁸Analogously, the Standard Model expectation is exactly recovered in the limit of $\mu \rightarrow 0$ in Eq. 11.2.

production cross section is proportional to γ'^4 in the Minimal model. Thus, the posterior takes the form,

$$p(\gamma'|\mathbf{n}) = \frac{\mathcal{L}'(\mathbf{n}|\gamma')p(\gamma')}{p(\mathbf{n})}. \quad (11.28)$$

The 95% C.L. upper limits on γ' are obtained by integrating posterior $p(\gamma'|\mathbf{n})$, as was done in the previous section, to find γ'_{95} such that

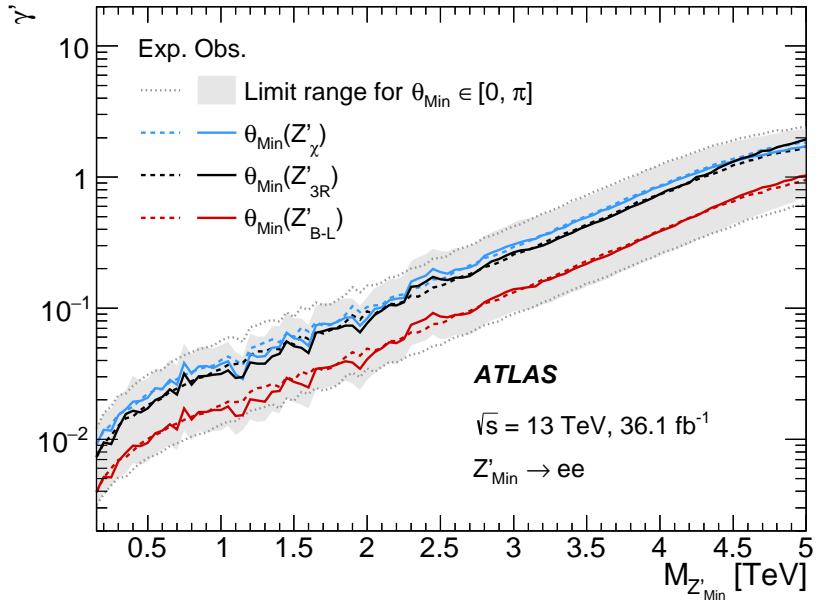
$$0.95 = \frac{\int_0^{\gamma'_{95}} p(\gamma'|\mathbf{n})d\gamma'}{\int_0^{\infty} p(\gamma'|\mathbf{n})d\gamma'}. \quad (11.29)$$

As with the SSM and E_6 model limits, these calculations are performed with BAT.

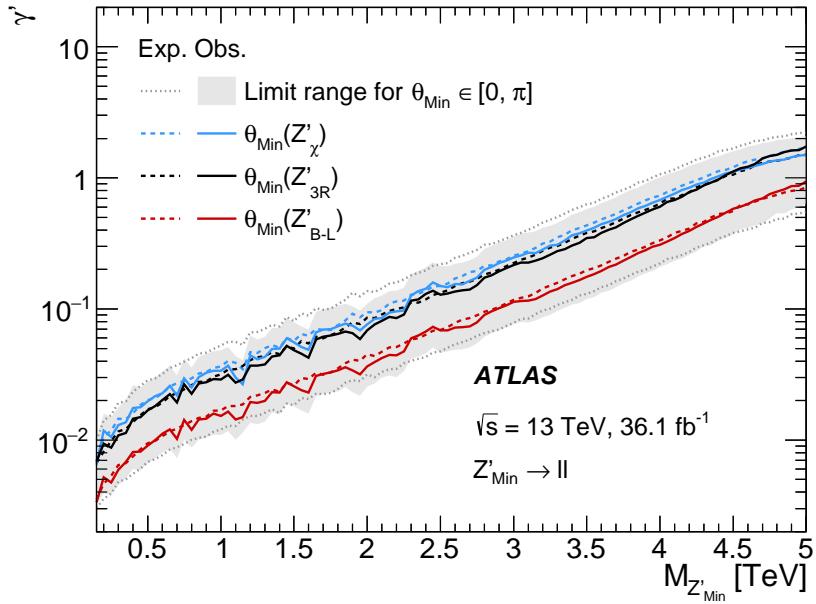
The resulting expected and observed upper limits can be seen in Figure 11.10. The collection $\theta_{\text{Min}} \in [0, \pi]$ of all Minimal model limit curves fits within the gray band. The values of θ_{Min} that correspond to the minimum and maximum Z'_{Min} cross sections are used to specify the edges of the band. Limit curves for the three representative models are shown, which for specific values of γ' and θ_{Min} correspond to the Z'_χ , Z'_{3R} , and Z'_{B-L} , respectively (See Table 5.3 for details). The region above each curve is excluded at 95% C.L.

Exclusion limits as a function of θ_{Min} are also depicted in Figure 11.11, which provide a complimentary interpretation to the mass limits where θ_{Min} takes a fixed value. For each case, following the strategy of extending the likelihood definition to include the muon channel, the combined limits are shown as well.

Lower limits on $M_{Z'_{\text{Min}}}$ for specific Z'_{Min} models can be obtained as well, following an approach similar to the one in the previous section. The intersection between the upper limit on γ' and the nominal value of γ' provided by theory, such as those in Table 5.3, defines a lower limit on Z'_{Min} pole mass. The lower mass limits are given in Table 11.2 for the electron,

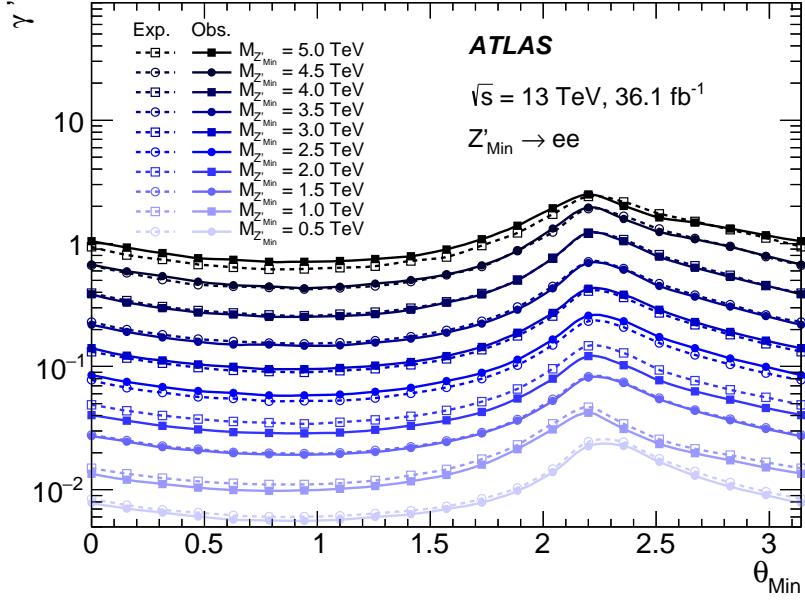


(a)

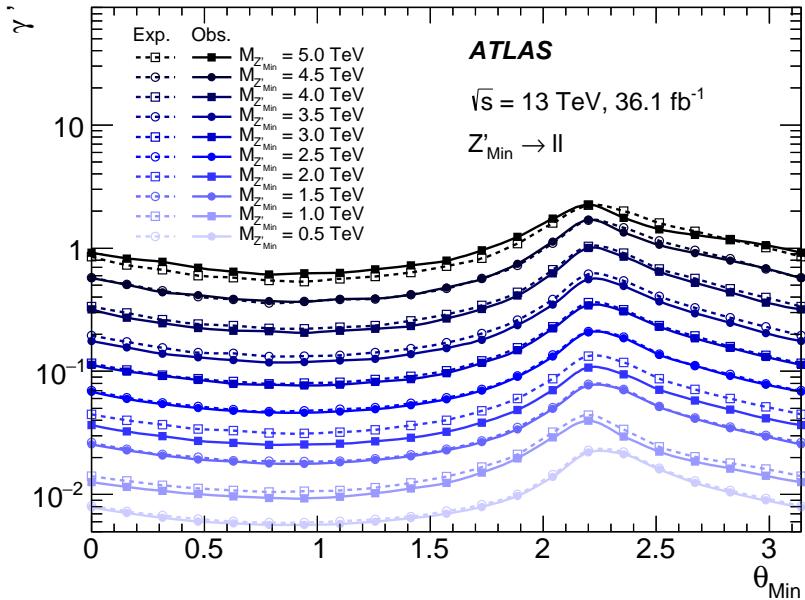


(b)

Figure 11.10: Observed and expected 95% C.L. upper limits on the relative coupling strength γ' as a function of the Z'_Min pole mass for the electron channel (a) and the combined electron plus muon channel (b). Three representative limit curves are shown, for specific values of the mixing angle θ_{Min} , which correspond to the Z'_χ , Z'_{3R} , and Z'_{B-L} models at specific values of γ' . All Minimal model limits fit within the gray band, which depend on the choice of $\theta_{\text{Min}} \in [0, \pi]$. The edges of the band are defined by the two models with the minimum and maximum Z'_Min cross sections. The region above each curve is excluded at 95% C.L.



(a)



(b)

Figure 11.11: Observed and expected 95% C.L. upper limits on the relative coupling strength γ' as a function of the mixing angle θ_{Min} for the electron channel (a) and the combined electron and muon channels (b). Limit curves are set for ten representative values of the Z'_{Min} pole mass $M_{Z'}$. The structure observed in the limit curves for a given pole mass as a function of θ_{Min} originates from the changing shape of the resonance. The resonance with the smallest (largest) width is obtained at $\theta_{\text{Min}} \simeq 1.0$ ($\theta_{\text{Min}} \simeq 2.2$), which define the edges of the gray band in Figure 11.10. The region above each curve is excluded at 95% C.L.

Model	γ'	$\tan \theta_{\text{Min}}$	Lower limits on $M_{Z'_{\text{Min}}}$ [TeV]					
			ee		$\mu\mu$		$\ell\ell$	
			Obs	Exp	Obs	Exp	Obs	Exp
Z'_χ	$\sqrt{\frac{41}{24}} \sin \theta_{\text{Min}}$	$-\frac{4}{5}$	3.7	3.7	3.4	3.3	3.9	3.8
Z'_{3R}	$\sqrt{\frac{5}{8}} \sin \theta_{\text{Min}}$	-2	4.0	3.9	3.6	3.6	4.1	4.1
Z'_{B-L}	$\sqrt{\frac{25}{12}} \sin \theta_{\text{Min}}$	0	4.0	4.0	3.6	3.6	4.2	4.1

Table 11.2: Observed and expected 95% C.L. lower mass limits on the Z'_{Min} pole mass for the three representative models considered in the text. As a result of including the interference effects and the variation of the decay width $\Gamma_{Z'}$ through the modulation of γ' , a slightly weaker limit for the Z'_χ is obtained in the Minimal model approach as compared to the standard approach of Section 11.4.2.

γ'	Range in limits on $M_{Z'_{\text{Min}}}$ [TeV]	
	Obs.	Exp.
0.01	0.2-0.8	0.2-0.8
0.02	0.4-1.6	0.4-1.5
0.04	0.8-2.3	0.8-2.3
0.10	1.7-3.3	1.6-3.2
0.20	2.4-4.0	2.4-3.9
0.40	3.2-4.5	3.1-4.6

Table 11.3: Range in the observed and expected 95% C.L. lower limits on the Z'_{Min} pole mass for several representative values of the relative coupling strength γ' . Both the electron and muon channels are combined.

$M_{Z'_{\text{Min}}}$ [TeV]	Range in limits on γ'	
	Obs.	Exp.
1	0.011-0.049	0.012-0.053
2	0.026-0.11	0.032-0.14
3	0.077-0.36	0.079-0.37
4	0.21-1.0	0.22-1.1
5	0.63-2.1	0.55-2.2

Table 11.4: Range in the observed and expected 95% C.L. upper limits on the relative coupling strength γ' for several representative values of the Z'_{Min} boson pole mass. Both the electron and muon channels are combined.

muon, and combined dilepton channels.

Furthermore, the ranges in the observed and expected lower limits on $M_{Z'_{\text{Min}}}$ for several characteristic values of γ' are shown in Table 11.3, and likewise, the ranges of observed and expected limits on γ' for several characteristic values of $M_{Z'_{\text{Min}}}$ are shown in Table 11.4.

11.5 Discussion of Results

This chapter presented the dilepton statistical analysis, which consisted of distinct search and limit setting phases. The dataset used to conduct these analyses consisted of 36.1 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$, as collected at the LHC and recorded by the ATLAS detector in the 2015 and 2016 data-taking runs. The record collision energy of the LHC allowed for a search for new physics to be conducted in a new as yet unexplored high-energy regime.

The main result of the dilepton analysis is the non-observation of any statistically significant resonant-like excesses over the predicted Standard Model background in either the dielectron or dimuon invariant mass distributions. As such, exclusion limits on Z' production were set in a variety of models.

The models considered were the SSM, the E_6 model, and the Minimal model, all of which predict the existence of a heavy spin-1 resonance that couples and decays to pairs of Standard Model fermions. As the search confirmed there was no excess in the data, upper limits were placed on the cross section times branching fraction σB for Z' production in the SSM and the E_6 model for well-motivated choices of θ_{E_6} , and on the relative coupling strength γ' in the Minimal model. In each case, these results were converted into lower limits on the Z' pole mass, which are some of the most stringent such limits to date. The resultant limits do

not strictly exclude any particular model, but they set tight constraints on Z' production over the invariant mass range searched at the LHC and on the allowed regions of model parameter space.

While this dissertation presents exclusion limits on models predicting resonant Z' production, other interpretations of the dilepton results are possible as well, and are included in detail in the most recent ATLAS dilepton search paper (See Ref. [5]). In particular, limits are set on several non-resonant Contact-Interaction (CI) models, whose signature would appear as a broad excess, rather than a resonant peak, over the Standard Model expectation. In addition, so-called **generic limits** are also provided to facilitate the reinterpretation of the dilepton results in the context of other models that predict narrow dilepton resonances, such as the Randall-Sundrum (RS) model [105], briefly mentioned in Section 5.2.

Part IV

PDF Uncertainty

Chapter 12

The Drell-Yan Process Revisited

As described in Section 10.8.2, the dominant systematic uncertainty in the dilepton analysis is purely theoretical and originates with the PDFs. The primary contributions to this theory uncertainty arise from the “PDF Variation” and “PDF Choice” systematic uncertainties, where the former is due to the modeling of the PDFs, and the later is due to a lack of agreement among the modern PDF sets.

To see this explicitly, consider the Standard Model prediction at a dielectron mass of $m_{ee} = 4$ TeV, which resides precisely in high-mass search region of interest, and is the primary target of the dilepton analysis. Here, the total theoretical uncertainty is 23%, which is roughly twice as large as the experimental uncertainty of 13%.¹ The individual “PDF Variation” and “PDF Choice” uncertainties are 19% and 8.4%, respectively. As $m_{ee} = 5$ TeV is reached, these increase to 29% and 47%, respectively, so that the total theory uncertainty becomes 57%; the total experimental uncertainty, on the other hand, remains largely the same, increasing only to 15%.

A precise understanding of the PDF uncertainty is therefore crucial to future Standard Model measurements, as well as high-mass searches for new physics. Consequently, in this chapter, the physics of the Drell-Yan process is further investigated in order to better understand the nature of the PDF uncertainty with the hope of further constraining it.

¹The contribution from the fake background is neglected in this estimate.

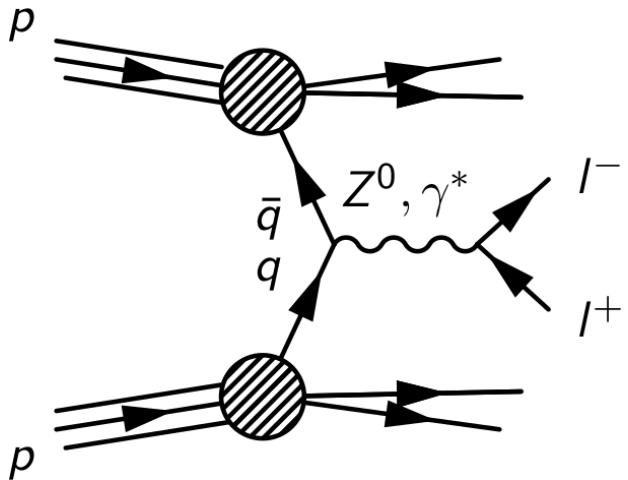


Figure 12.1: The Drell-Yan process initiated by a quark-antiquark pair as observed at the LHC [29].

12.1 The Drell-Yan Process

The Drell-Yan process was introduced in Section 4.2. At the LO in QCD, the process is initiated by the annihilation of quark and antiquark pair via the *s*-channel exchange of an electroweak boson, which then decays into a pair of same-flavor, oppositely-charged leptons as in Fig. 12.1. The intermediate boson can be produced either through the annihilation of a valance quark with a sea quark (e.g. $u\bar{u}$, $d\bar{d}$, etc.), or through the annihilation of two sea quarks (e.g. $s\bar{s}$, $c\bar{c}$, etc.) in the collision.

12.1.1 Drell-Yan Kinematics

The kinematics of the Drell-Yan process can be fully described in terms of three kinematic variables, which are defined below. Although some of the definitions have already been provided in Section 4.2, they are repeated here for convenience.

In the Drell-Yan process, the quark q carries a fraction x_1 of total proton momentum P_1 . Likewise, the antiquark \bar{q} carries a fraction x_2 of total proton momentum P_2 . Thus, the

parton four-momenta can be written,

$$p_1 = x_1 P_1, \quad p_2 = x_2 P_2 \quad (12.1)$$

Using Eq. 12.1, the relativistic invariant \hat{s} (as defined in Sec. 4.2.1) can be calculated from

$$\hat{s} = (p_1 + p_2)^2 = x_1 x_2 s, \quad (12.2)$$

where \sqrt{s} is the CM energy of the colliding proton-proton (pp) system, and $\sqrt{\hat{s}}$ is the CM energy of the quark-antiquark ($q\bar{q}$) system. The squared CM energy \hat{s} equals the scale of the collision Q^2 and invariant mass $m_{\ell\ell}$ of the dilepton system,

$$\hat{s} = Q^2 = m_{\ell\ell}^2. \quad (12.3)$$

The dilepton invariant mass $m_{\ell\ell}$ defines the first of the three Drell-Yan kinematic observables of interest.

The second variable to be considered is the dilepton rapidity,

$$y_{\ell\ell} = \frac{1}{2} \ln \left(\frac{E + P_z}{E - P_z} \right), \quad (12.4)$$

where E and P_z are the energy and longitudinal momentum of the dilepton system. The four-momenta of the colliding quark and antiquark can be written,

$$p_1 = \frac{\sqrt{s}}{2} (x_1, 0, 0, x_1), \quad p_2 = \frac{\sqrt{s}}{2} (x_2, 0, 0, -x_2), \quad (12.5)$$

and upon substitution of Eq.12.5 into Eq. 12.4, the rapidity can be re-expressed as

$$y_{\ell\ell} = \frac{1}{2} \ln \frac{x_1}{x_2}, \quad (12.6)$$

which is exclusively in terms of the parton momentum fractions x_1 and x_2 . Furthermore, the substitution of Eq. 12.2 and Eq. 12.3 into Eq. 12.6 results in an expression for the partonic x fractions in terms of Drell-Yan kinematic observables,

$$x_1 = \frac{m_{\ell\ell}}{\sqrt{s}} e^{+y_{\ell\ell}}, \quad x_2 = \frac{m_{\ell\ell}}{\sqrt{s}} e^{-y_{\ell\ell}}. \quad (12.7)$$

With this, the PDFs $f_{q/P_1}(x_1, Q^2)$ and $f_{\bar{q}/P_2}(x_2, Q^2)$ (as in Sec. 3.2) can be extracted from differential measurements of the Drell-Yan cross section in the dimensions of dilepton mass $m_{\ell\ell}$ and rapidity $y_{\ell\ell}$.

In the high-mass region of the dilepton analysis, the x values relevant for Drell-Yan production can be calculated from Eq. 12.7. For a central dilepton event with an invariant mass of $m_{\ell\ell} = 3$ TeV and rapidity of $y_{\ell\ell} = 0$, the momentum fractions involved are $x \simeq 0.23$. For higher masses, the relevant average x values are of course larger, so any future dilepton search would sample fairly high values of x . These values correspond roughly to where the valence distributions peak (See Fig. 3.4); for larger dilepton masses or rapidities, the x values that result are placed firmly beyond, in the extreme region of the PDF tails.

As depicted in Figure 3.6, the uncertainties associated with the valence and sea PDFs in the moderate- x region of $10^{-3} \lesssim x \lesssim 0.1$ are fairly well constrained; however, as the high- x region of $x \gtrsim 0.1$ is reached, these uncertainties become significant. Therefore, this suggests the PDF systematic uncertainty that dominates the high-mass dilepton spectrum

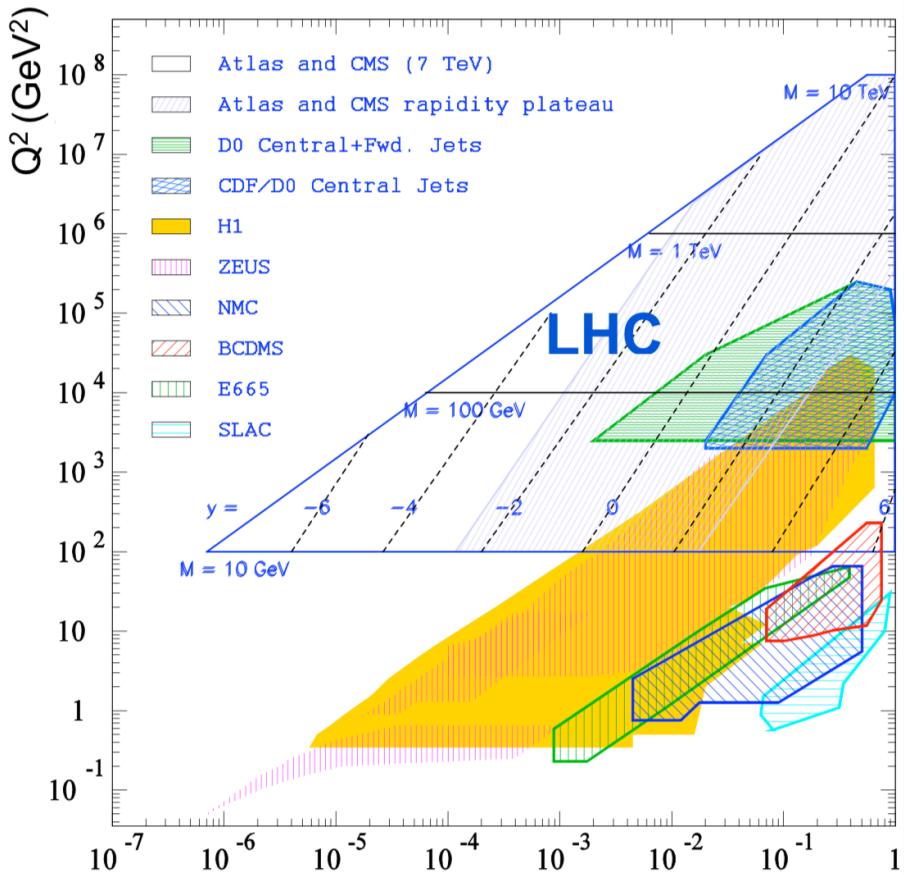


Figure 12.2: The transferred momentum squared Q^2 versus the parton momentum fraction x at $\sqrt{s} = 7$ TeV. The regions probed by previous DIS, fixed-target, and collider-based experiments are labeled. The accessible regions to the ATLAS and CMS experiments are indicated by the solid blue line, and the shaded region within denotes measurements using the central region of each detector. The high-mass search region of the dilepton analysis can be considered as beyond $m_{ee} = 1$ TeV, where two high- x partons ($x \gtrsim 0.1$) are required to initiate the Drell-Yan process [30].

is precisely due to the unconstrained high- x behavior of the PDFs.

The reason for this inherent high- x uncertainty in the modern PDF sets is due to a lack of experimental data, as previous experiments have yet to probe the PDFs in these extreme regions (x, Q^2) of kinematic phase space (See Fig. 12.2). At the same time, measurements in these regions can be challenging due to the steeply falling nature of the PDFs at high x (See Fig. 3.4).

The valence quark PDFs in the high- x region have been measured in previous experiments [30], so the largest uncertainties in these distributions are found at low x and at very high x . For example, in the CT14Hera2 error set, the $u_v(x)$ and $d_v(x)$ distributions have an uncertainty that reaches 20% at $x \gtrsim 0.5$ and $x \lesssim 10^{-3}$ for $Q^2 = (3 \text{ TeV})^2$ (See Fig. 3.6).

The sea quark and gluon PDFs, on the other hand, are the least constrained at high x due to a lack of experimental data in this region, and due to the much more rapid fall-off that takes place as compared to the valence PDFs. The gluon $g(x)$ uncertainty reaches 20% at $x \gtrsim 0.3$ at $Q^2 = (3 \text{ TeV})^2$, and likewise for the $\bar{u}(x)$ and $\bar{d}(x)$ distributions, this occurs at $x \gtrsim 0.2$ (See Fig. 3.6). From this, it can be seen that the high-mass Drell-Yan PDF uncertainty arises precisely from this lack of knowledge in the high- x sea PDFs, and to a lesser extend the high- x valence PDFs, as both a quark and antiquark are required to initiate the LO Drell-Yan process.

That a difficult region of x causes a large uncertainty for high-mass Drell-Yan perhaps it is also an opportunity: the LHC may allow for the PDFs to be probed in this previously inaccessible (x, Q^2) kinematic regime. The Drell-Yan process is crucial in this regard, as contemporary cross section measurements may help to validate the less-precise datasets used in the PDF global fits and thereby improve their overall quality. Thus, dedicated measurements targeting these precise x regions at the LHC could help to reduce the PDF uncertainty and lead to an overall better understanding of proton substructure.

The third observable of interest is defined in the following section.

12.1.2 The Collins-Soper Frame

The third observable of interest is the polar angle θ^* of the outgoing *lepton* relative to the incoming *quark*, measured in a special rest frame of the dilepton system known as the

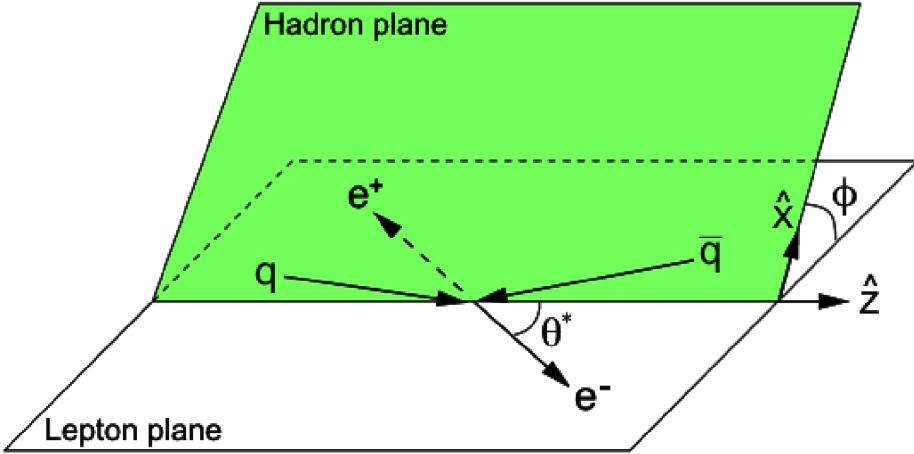


Figure 12.3: The Collins-Soper reference frame. The \hat{z} axis bisects the angle between the quark momentum p_1 and the negative of the antiquark momentum $-p_2$. The \hat{y} axis is defined perpendicular to the plane spanned by the parton momenta, and \hat{x} axis completes the right handed coordinate system. The polar angle θ^* is measured from the \hat{z} axis in this frame [31].

Collins-Soper (CS) frame [210] (See Fig. 12.1.2).²

Starting from the rest frame of the dilepton system, the CS frame can be obtained from the following set of coordinate transformations:

1. The \hat{z} axis is defined such that it bisects the angle between the momentum vector of the incoming quark q and the negative momentum vector of the incoming antiquark \bar{q} .³
2. The \hat{y} axis is defined as the vector normal to the plane spanned by the momenta of the quark-antiquark pair.
3. The \hat{x} axis is chosen such that a right handed coordinate system is obtained.

The polar angle θ^* , measuring the direction of the outgoing lepton ℓ^- relative to \hat{z} in the

²The CS frame accounts for the non-zero p_T of the dilepton system. When the dilepton system has zero transverse momenta, the CS frame reduces to the dilepton rest frame.

³This definition introduces an inherent sign ambiguity, since it is not clear at the LHC which proton carries the quark. However, the quark direction can be inferred on an event-by-event basis according to the longitudinal boost of the dilepton system $y_{\ell\ell}$, assuming this is in the same direction as that of the initial-state quark.

CS frame, can be calculated directly from lab frame quantities,

$$\cos \theta^* = \frac{P_z}{|P_z|} \frac{2(p_1^+ p_2^- - p_1^- p_2^+)}{M_{\ell\ell} \sqrt{M_{\ell\ell}^2 + P_T^2}}, \quad (12.8)$$

where P_T and P_z are the transverse and longitudinal momentum of the dilepton system, respectively and,

$$p_i^\pm = \frac{1}{\sqrt{2}} (E_i \pm p_{z,i}), \quad i = 1, 2, \quad (12.9)$$

where the lepton (antilepton) energy and longitudinal momentum are specified by E_1 and $p_{z,1}$ (E_2 and $p_{z,2}$), respectively. Notice this definition requires the electric charge of each lepton to be correctly identified.

Within the CS frame, Drell-Yan events can be classified as **forward** or **backward** according to the lepton decay angle. Events where the outgoing *lepton* travels in the same direction as the incoming *quark* (i.e. $\cos \theta^* > 0$) are classified as forward events, while those where it travels in the opposite direction (i.e. $\cos \theta^* < 0$) are classified as backward events. For forward events, the helicity of the final-state lepton generally matches that of the initial-state quark, while for backwards events, the helicities are opposite (See Eq. 4.14).

12.1.3 Triple-Differential Cross Section

The Drell-Yan triple-differential cross section can be represented as a function of the three variables outlined in the previous two sections: the dilepton invariant mass $m_{\ell\ell}$, the dilepton rapidity $y_{\ell\ell}$, and the cosine of the polar angle θ^* of the incoming quark and outgoing lepton measured in the CS frame $\cos \theta^*$. At the LO, the Drell-Yan triple-differential cross section

can be written,

$$\frac{d^3\sigma}{dm_{\ell\ell}dy_{\ell\ell}d\cos\theta^*} = \frac{\pi\alpha^2}{3m_{\ell\ell}s} \sum_q P_q \left[f_{q/P_1}(x_1, Q^2) f_{\bar{q}/P_2}(x_2, Q^2) + (q \leftrightarrow \bar{q}) \right], \quad (12.10)$$

where \sqrt{s} is the CM energy of the LHC, x_1 and x_2 are the incoming parton momentum fractions, and $f_{q/P_1}(x_1, Q^2)$ and $f_{\bar{q}/P_2}(x_2, Q^2)$ are the PDFs for partons of flavor q and \bar{q} , respectively. The scale of the collision Q^2 is set to the dilepton invariant mass $m_{\ell\ell}$, which is equal to the partonic CM energy as in Eq. 12.3. The term $(q \leftrightarrow \bar{q})$ accounts for the fact that either proton can carry a sea quark, as the LHC is a proton-proton collider.

The function P_q encodes the dynamics with

$$\begin{aligned} P_q &= Q_\ell^2 Q_q^2 \left(1 + \cos^2 \theta^* \right) \\ &+ Q_\ell Q_q \frac{2m_{\ell\ell}^2 (m_{\ell\ell}^2 - M_Z^2)}{\sin^2 \theta_W \cos^2 \theta_W \left[(m_{\ell\ell}^2 - M_Z^2)^2 + \Gamma_Z^2 M_Z^2 \right]} \left[g_V^\ell g_V^q (1 + \cos^2 \theta^*) + 2g_A^\ell g_A^q \cos \theta^* \right] \\ &+ \frac{m_{\ell\ell}^4}{\sin^4 \theta_W \cos^4 \theta_W \left[(m_{\ell\ell}^2 - M_Z^2)^2 + \Gamma_Z^2 M_Z^2 \right]} \\ &\times \left[\left(g_A^\ell + g_V^\ell \right) \left(g_A^q + g_V^q \right) \left(1 + \cos^2 \theta^* \right) + 8g_A^\ell g_V^\ell g_A^q g_V^q \cos \theta^* \right], \end{aligned} \quad (12.11)$$

where M_Z and Γ_Z are the mass and width of the Z boson, Q_f is the fermion electric charge, and g_V^f and g_A^f are the electroweak vector and axial-vector fermion couplings, respectively, the values of which are listed in Table 2.4.

Each term in Eq. 12.11 has an associated interpretation: the first term corresponds to the exchange of a virtual photon γ^* , the second results from γ^*/Z interference, and the third from pure Z boson exchange. The cross section of Eq. 12.10 exhibits a strong mass

dependence, falling continuously, and smoothly like $1/m_{\ell\ell}^2$, except at the mass of the Z boson, where the smoothly falling behavior is replaced by a Breit-Wigner resonance.

The dependence on rapidity $y_{\ell\ell}$ provides direct sensitivity to the PDFs, as Eq. 12.6 is expressed in terms of the parton momentum fractions x_1 and x_2 . The dilepton mass $m_{\ell\ell}$ provides further sensitivity, since the relative contribution of u -type and d -type quarks to the total cross section varies with mass.⁴ Finally, the dependence on $\cos\theta^*$ provides additional sensitivity to the terms in Eq. 12.11 that contain axial- and axial-vector prefactors.

12.1.4 Kinematic Distributions

The kinematic distributions for the absolute dilepton rapidity $|y_{\ell\ell}|$ and cosine of the polar angle $\cos\theta^*$ are presented according to the prediction of the triple-differential cross section of Eq. 12.10. The RESBos MC generator (See Sec. 9.1.2), configured with the CT14Hera2 PDF set, is used for the event generation. The simulation consists of a set of mass-binned samples, covering an invariant mass range of $m_{\ell\ell} = 40$ GeV to $m_{\ell\ell} = 1$ TeV at $\sqrt{s} = 13$ TeV.

For each sample, the lepton pseudorapidity is required to be within $|\eta| < 4.9$, which covers the forward region of the ATLAS detector. A simple selection is applied to the events of each sample: first, leptons must have a transverse momentum of $E_T > 30$ GeV, and second, events must consist of either a central-central ($|\eta| < 2.47$) or central-forward ($2.5 < |\eta| < 4.9$) lepton pair (See Sec. 13.3.1), in order to be detectable within the ATLAS detector.

The $|y_{\ell\ell}|$ (See Fig. 12.4) and $\cos\theta^*$ (See Fig. 12.5) distributions are depicted in slices of dilepton invariant mass $m_{\ell\ell}$ to illustrate how they evolve as a function of mass. Further, the **flavor decomposition** of each observable is presented, which shows how the overall Drell-

⁴At the Z peak, for example, the partons couple almost exclusively through the weak interaction; off-peak, however, the electromagnetic coupling to γ^* becomes increasingly dominant.

Yan event yield decomposes into its individual contributions from the u -type and d -type Drell-Yan sub-processes.⁵

The $\cos \theta^*$ distributions show that the rates associated with each quark-type sub-process are in fact highly dependent on the polar angle θ^* . For example, above the Z boson mass peak and in the forward region of $\cos \theta^* > 0$, an increasing degree of separation can be seen between the rates associated with the u -type and d -type sub-processes as compared to the backward region of $\cos \theta^* < 0$. In this region, and especially at high masses, it can be seen that the contribution to the total cross section is due predominantly to u -type sub-process.

The fact that $\cos \theta^*$ exhibits such a strong sensitivity to the distinct parton sub-processes suggests it might be useful to include as an additional dimension in future PDF global fits. Along with the more traditional observables of dilepton invariant mass and rapidity, as established in Section 12.1.1, $\cos \theta^*$ might provide additional discriminating power between the individual parton flavors, and thereby function to constrain the PDF uncertainty further than what could be achieved otherwise. Indeed, this idea is explored in detail in Section 12.1.6.

The $|y_{\ell\ell}|$ distributions show the contributions of the u -type and d -type sub-processes to the total cross section as well. Below the Z boson mass peak, the maximum rapidity attained is $|y_{\ell\ell}| \lesssim 3.5$. By the time $m_{\ell\ell} = 1$ TeV is reached, the maximum rapidity reduces to $|y_{\ell\ell}| \lesssim 2.5$. The relative fraction of the d -type rate is enhanced at higher rapidities.

12.1.5 Explaining the $\cos\theta^*$ Asymmetry

In order to better understand the asymmetric behavior of the u - and d -type Drell-Yan sub-processes as a function of mass and $\cos \theta^*$ (as depicted in Fig. 12.5), the function P_q of

⁵The u -type sub-processes include initial-states defined by $u\bar{u}$, ug , and $\bar{u}g$, where u is any up-type quark and g is the gluon. An identical definition applies to the d -type sub-processes, with u replaced by d .

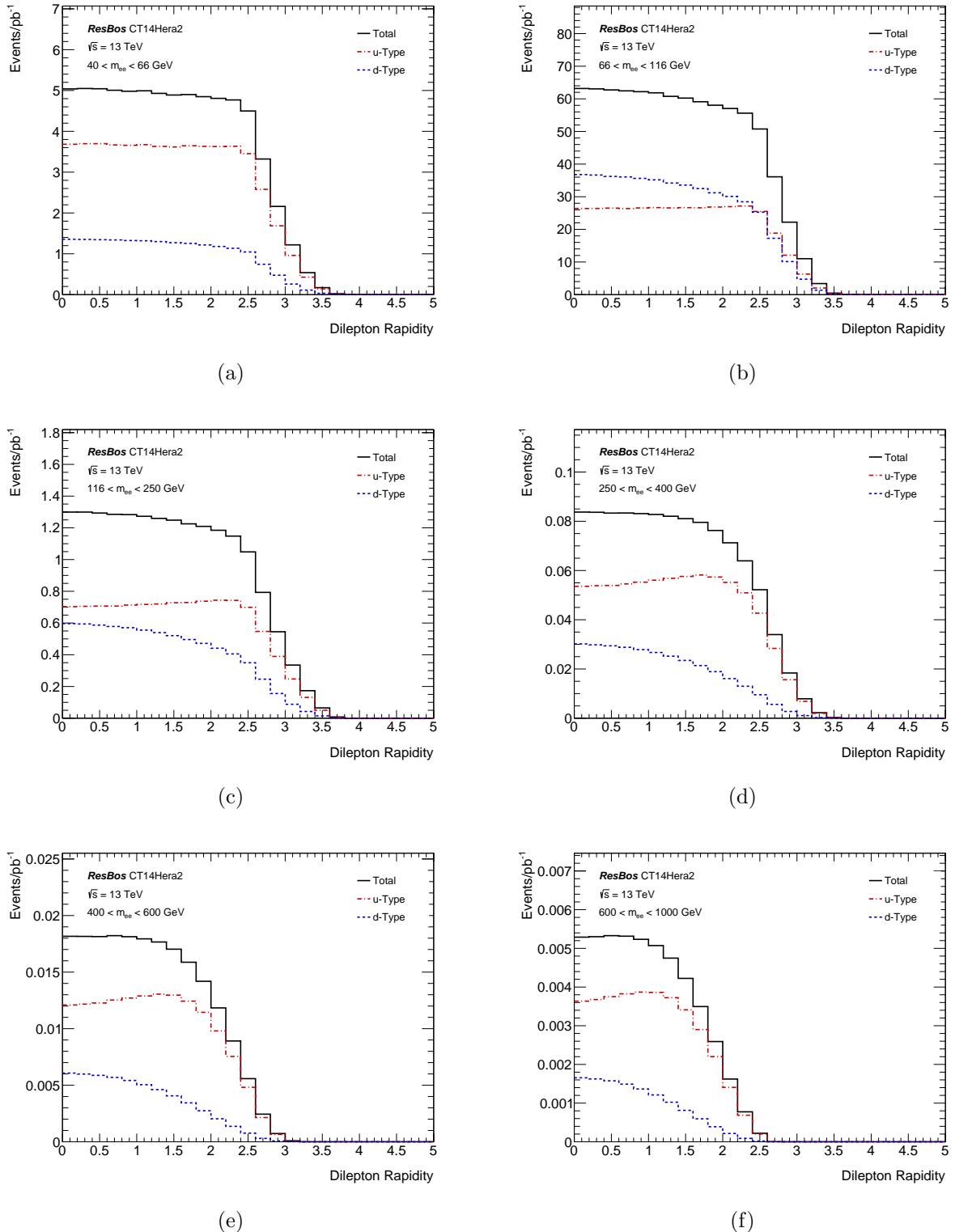


Figure 12.4: The absolute dilepton rapidity $|y_{\ell\ell}|$ in several slices in dilepton invariant mass $m_{\ell\ell}$ ranging from 40 GeV to 1 TeV ((a)-(f)). The relative fractions of the u -type and d -type Drell-Yan sub-processes are shown to illustrate the partonic flavor decomposition in this observable.

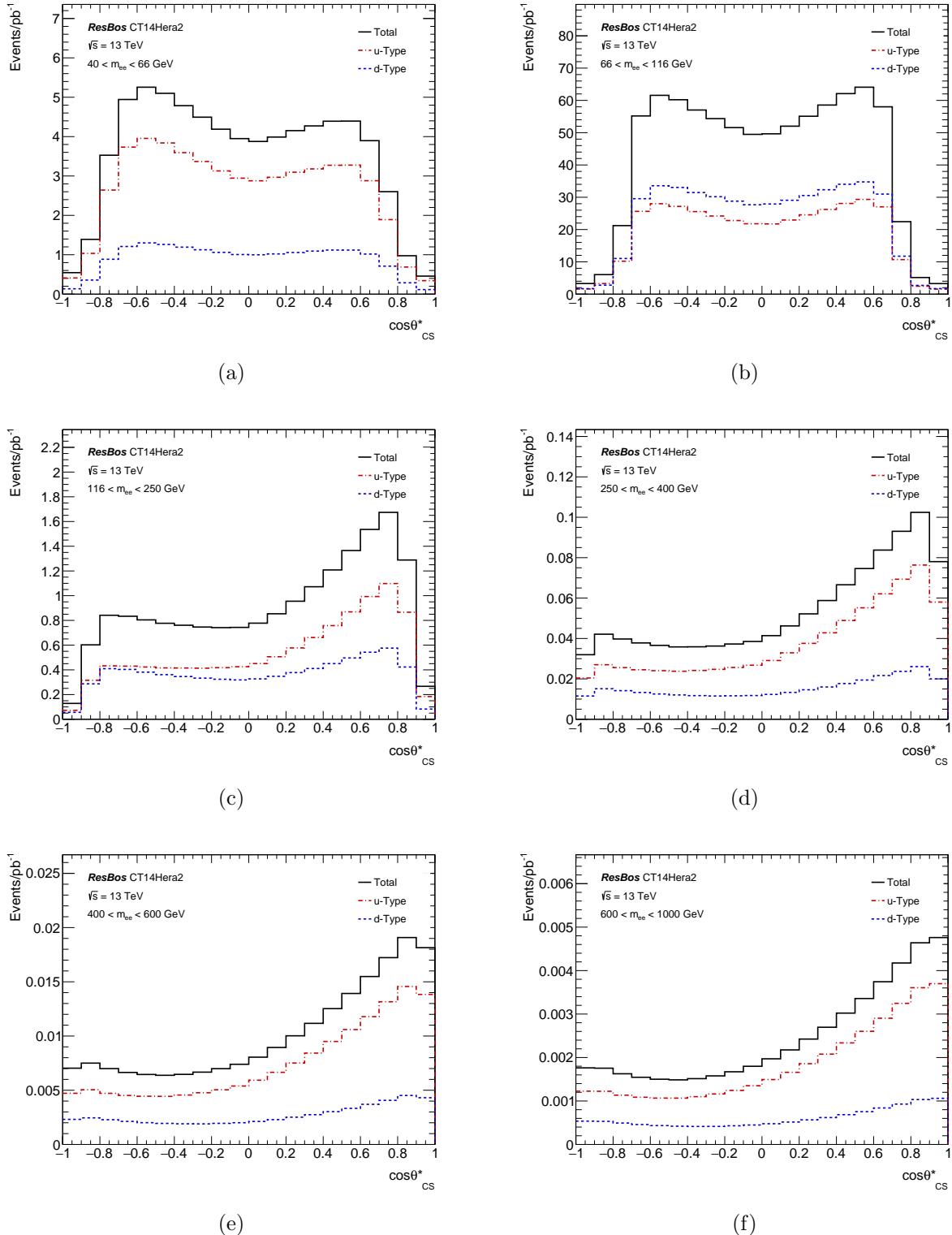


Figure 12.5: The lepton angular distribution $\cos\theta^*$ in several slices of dilepton invariant mass m_{ll} ranging from 40 GeV to 1 TeV ((a)-(f)). The relative fractions of u -type and d -type sub-processes are shown as well, which exhibit a strong angular dependency, especially at high mass.

Eq. 12.11 can be recast in a more suitable form,

$$P_q = C_q^0 \left(1 + \cos^2 \theta^* \right) + C_q^1 \cos \theta^*, \quad (12.12)$$

where,

$$\begin{aligned} C_q^0 &= \left[Q_\ell^2 Q_q^2 + 2Q_\ell Q_q g_V^\ell g_V^q \chi_1(m_{\ell\ell}) + \left(g_A^{\ell 2} + g_V^{\ell 2} \right) \left(g_A^{q 2} + g_V^{q 2} \right) \chi_2(m_{\ell\ell}) \right] \\ C_q^1 &= \left[4Q_\ell Q_q g_A^\ell g_A^q \chi_1(m_{\ell\ell}) + 8g_A^\ell g_V^\ell g_A^q g_V^q \chi_2(m_{\ell\ell}) \right] \\ \chi_1(m_{\ell\ell}) &= \kappa \frac{m_{\ell\ell}^2 (m_{\ell\ell}^2 - M_Z^2)}{(m_{\ell\ell}^2 - M_Z^2)^2 + \Gamma_Z^2 M_Z^2} \\ \chi_2(m_{\ell\ell}) &= \kappa^2 \frac{m_{\ell\ell}^4}{(m_{\ell\ell}^2 - M_Z^2)^2 + \Gamma_Z^2 M_Z^2} \\ \kappa &= \frac{1}{\sin \theta_W \cos \theta_W}. \end{aligned} \quad (12.13)$$

Here, χ_1 is due to γ^*/Z interference and χ_2 is due to Z boson exchange. In this form, P_q is described by an even function⁶ ($1 + \cos^2 \theta^*$) and an odd function⁷ ($\cos \theta^*$). The odd part is responsible for inducing the well-known forward-backward asymmetry A_{FB} .

At low mass, below the Z boson mass peak, the Drell-Yan process proceeds primarily through γ^* photon exchange. In this limit, χ_1 and $\chi_2 \rightarrow 0$ as $m_{\ell\ell} \rightarrow 0$. Therefore, the C_q^1 term of Eq. 12.12 vanishes and C_q^0 is described purely by the electromagnetic component of the cross section. As a result, P_q of Eq. 12.12 is described almost exclusively by an even function, and little to no asymmetry is present. This accounts for the behavior of the $\cos \theta^*$ distribution of Figure 12.5a.

At the Z boson mass peak, the Drell-Yan proceeds almost exclusively though on-shell

⁶An even function satisfies $f(x) = f(-x)$

⁷An odd function satisfies $f(-x) = -f(x)$

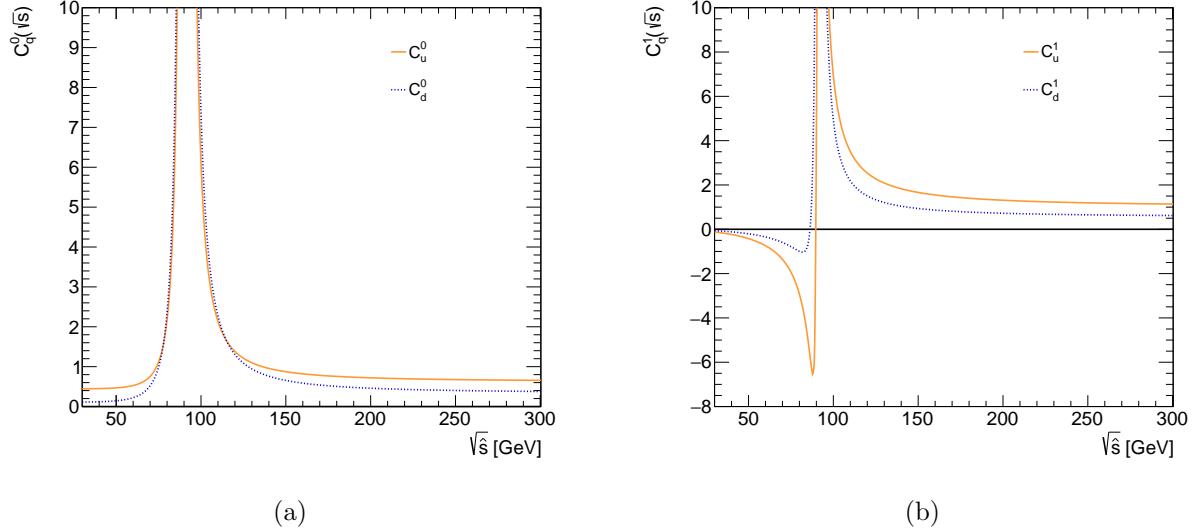


Figure 12.6: Drell-Yan asymmetry coefficients of Eq. 12.12. The parameter C_q^0 ((a)) describes the magnitude of the symmetric component of the Drell-Yan cross section, while C_q^1 ((b)) describes the magnitude of the asymmetric component.

Z boson exchange. Here, $\chi_1(M_Z) = 0$ and $\chi_2(M_Z) = \kappa^2 \frac{M_Z^2}{\Gamma_Z^2}$. While the term $C_q^1 = 8g_A^\ell g_V^\ell g_A^q g_V^q \chi_2(M_Z)$ is indeed non-zero, ensuring P_q has an odd function component, the value of the coefficient $8g_A^\ell g_V^\ell g_A^q g_V^q$ is too small to induce any asymmetric effects. Therefore, the forward-backward asymmetry is appreciably small at the Z peak, as in the $\cos\theta^*$ distribution of Figure 12.5b.

At masses beyond the Z peak, the contributions from γ^*/Z interference and pure Z boson exchange are non-negligible, and as a result, both χ_1 and χ_2 functions are non-zero, and are *roughly comparable* in magnitude. In this limit, both the C_q^0 and C_q^1 terms contribute to the total cross section. The dominant contribution to C_q^1 comes from the term $4Q_\ell Q_q g_A^\ell g_A^q \chi_1(m_{\ell\ell})$, and it is precisely this term that induces the forward-backward asymmetry observed in the Drell-Yan process. The size of this effect differs for the u - and d -type Drell-Yan sub-processes due to the values of their electroweak couplings and the dilepton mass scale of interest, which results in $C_u^1 > C_d^1$.

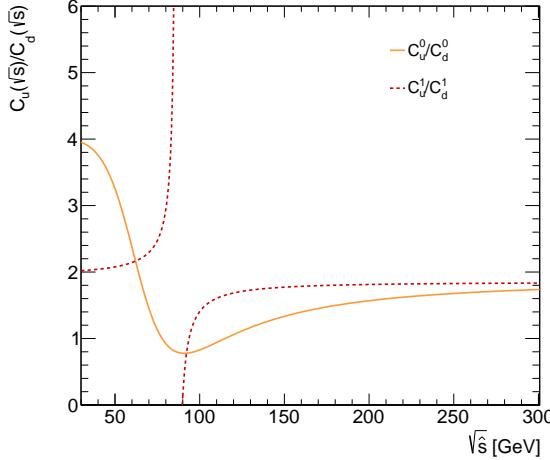


Figure 12.7: Ratio of Drell-Yan asymmetry coefficients C_q^0 and C_q^1 of Eq. 12.12.

Consequently, this explains the increasingly asymmetric behavior of the remaining $\cos \theta^*$ distributions in Figures 12.5c through 12.5f. Likewise, the fact that $C_u^0 > C_d^0$ explains the increasing separation between the rates of the u - and d -type sub-processes as a function of dilepton mass $m_{\ell\ell}$.

To drive this point home, consider the dilepton behavior at $m_{\ell\ell} = 3$ TeV. Here, the u -type coefficients are $C_u^0 = 0.44 + 0.015 + 0.16 = 0.62$ and $C_u^1 = 0.99 + 0.033 = 1.03$, while the d -type are $C_d^0 = 0.11 + 0.013 + 0.21 = 0.33$ and $C_d^1 = 0.49 + 0.059 = 0.56$. The d -type components are each roughly a factor of two smaller than their u -type counterparts. For each quark flavor q , $C_q^1 > C_q^0$, indicating a sizeable asymmetry is present, as the magnitude of the odd function in Eq. 12.12 dominates. Figure 12.6 displays the C_q terms explicitly as a function of $\sqrt{\hat{s}} = m_{\ell\ell}$, and Figure 12.7 plots their ratio.

12.1.6 $\cos\theta^*$ as a Quark Flavor Enhancer

The cosine of the polar angle $\cos\theta^*$ is a highly discriminating variable, sensitive to the kinematics of the different partonic sub-processes in the Drell-Yan process. This fact can

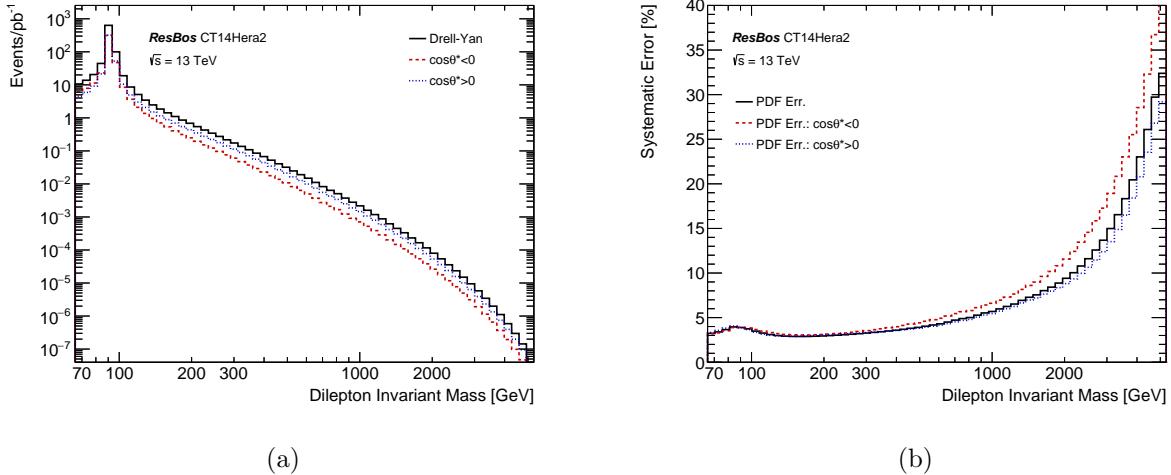


Figure 12.8: The dilepton invariant mass $m_{\ell\ell}$ as selected by forward ($\cos \theta^* > 0$) or backward ($\cos \theta^* < 0$) events in the CS frame ((a)) and the corresponding PDF uncertainty associated with each selection ((b)). Events arising from the forward-forward lepton topology are not included. As with the other kinematic distributions, the CT14Hera2 PDF set is used. The selection has the effect of modulating the PDF systematic uncertainty.

be exploited to “sculpt” the PDFs so as to reduce their overall uncertainty. For example, with a forward $\cos \theta^* > 0$ requirement, the d -type contribution to the total cross section will be suppressed, and, as the $d_v(x)$ PDFs are less precisely determined than their $u_v(x)$ counterparts (as in Fig. 3.6), the overall PDF uncertainty will be reduced.⁸ This effect can be used *even today* to modulate the PDF systematic uncertainty as in Figure 12.8. Here, just as with the other kinematic distributions, the state-of-the-art CT14Hera2 PDF set is used in the ResBos event generation.

The invariant mass $m_{\ell\ell}$ is shown along with the uncertainty on the event yield due to the PDFs according to forward and backward selections applied to events in the CS frame. In the nominal selection, the PDF uncertainty at $m_{\ell\ell} = 3$ TeV is 15% and at $m_{\ell\ell} = 5$ TeV it is 31% (black curves). The forward selection (blue-dashed curves) reduces the PDF uncertainty

⁸The $u_v(x)$ distribution is better constrained than the $d_v(x)$ distribution precisely because the dominant contribution to the valence structure of the proton is from the up quark. As a result, the $u_v(x)$ distribution is more likely to initiate the Drell-Yan process in high-energy scattering experiments.

to 14% at 3 TeV and to 29% at 5 TeV, while the backward (red-dashed) selection has the opposite effect; it actually *increases the overall uncertainty*. Therefore, by imposing the requirement of $\cos \theta^* > 0$, the u -type component of the Drell-Yan cross section is enhanced (and the d -type component is suppressed); as such, the overall PDF uncertainty is reduced, which, when propagated to the invariant mass, has the effect of reducing the uncertainty on the high-mass Drell-Yan event yield.

This strategy of using $\cos \theta^*$ to modulate the flavor content of the Drell-Yan cross section to reduce the overall PDF uncertainty is novel and could, for example, be applied to the dilepton analysis, where, in addition to the nominal selection defined in Section 10.5, an additional requirement could be placed to select only forward events. The advantage of such a selection is that this additional requirement can be applied to the search now; there is no need to wait for future PDF sets or additional LHC data. Furthermore, in the case of a non-resonant excess, this method could provide a useful handle in diagnosing the nature of the excess, which would otherwise be stymied by the large PDF uncertainty at high-mass.

To reiterate, the fact that a simple selection requirement on $\cos \theta^*$ can increase or decrease the PDF systematic uncertainty suggests, again, it can supplement the invariant mass and rapidity distributions as constraints in the PDF global fits, and therefore motivates the followup study presented in the next chapter.

Chapter 13

PDF Error Reduction

The goal of this chapter is to assess how future LHC measurements of the Drell-Yan cross section can help to constrain the valence and sea PDF uncertainties at high x , and thereby reduce the PDF systematic uncertainty on the Drell-Yan event yield at high dilepton invariant mass. This is a work proposed by this author.

To accomplish this, the `ePump` [211] package is used, which is able to update an existing PDF set with new experimental data (or pseudo-data) and consequently produce new best-fit and error PDFs. In the previous chapter, $\cos \theta^*$ was established as a potentially sensitive discriminating variable, since it separates u -type and d -type parton flavors. In this chapter, `ePump` is used to update an existing PDF set with Drell-Yan pseudo-data, which takes advantage of $\cos \theta^*$ explicitly. Figure 13.1 sketches the planned use of `ePump`.

The pseudo-data are generated for the Drell-Yan process at several characteristic integrated luminosities and are incorporated into the `CT14Hera2` PDF set using `ePump`. The updated PDFs are subsequently used to assess any potential improvement in the high-mass dilepton spectrum.¹

This chapter is formatted as follows: first the `ePump` package and its functionality is introduced (Sec. 13.1), then the PDF update analysis is performed (Sec. 13.3), and its effects are assessed (Sec. 13.4). Finally, a brief outlook is presented (Sec. 13.5).

¹The `CT14Hera2` set incorporates data from fixed target experiments, the Tevatron, LHC Run-1, and HERA [94].

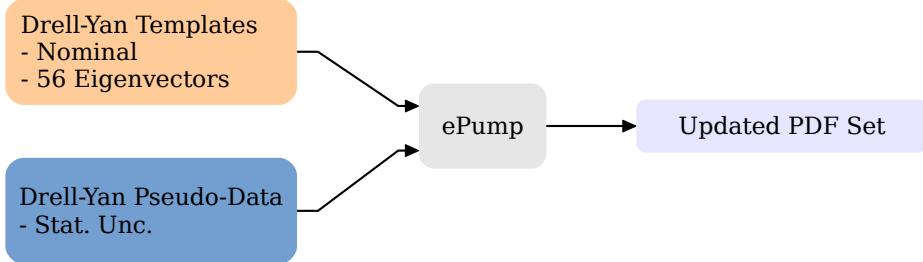


Figure 13.1: `ePump` requires two sets of inputs to perform an update: the first consists of a set of binned signal templates, each of which is comprised of a nominal theory prediction and n eigenvector variations ($n = 56$ in CT14Hera2), and the second consists of an analogous set of binned data (or pseudo-data) templates, along with their statistical uncertainties. The observed reduction in PDF uncertainty associated with the updated PDF set depends on the statistical precision of the data. Datasets with high statistical precision are better able to constrain the PDF uncertainty.

13.1 The `ePump` Package

The `ePump` (Error PDF Updating Method Package) package can be used to **update** or **optimize** a PDF set for analysis.² The package itself package contains two executables: `UpdatePDFs` and `OptimizePDFs`.

The first executable `UpdatePDFs` can be used to predict the impact of new experimental data on an already existing PDF set (e.g. the CT14Hera2 set). This is achieved by producing updated best-fit and updated error PDFs given new data by essentially re-performing the χ^2 minimization of the global PDF analysis (See Sec. 3.2.1) and Hessian error analysis (See Sec. 3.2.2). However, a full global re-analysis of the data is not required, as the same χ^2 minimum found during global analysis is assumed for the update, allowing for results to be obtained extremely quickly.

As the global χ^2 function is a sum over discrete measurements, the requisite `ePump` inputs, or **templates**, are formatted as histograms with a finite binning. The **data templates** in-

²A PDF set in the context of this discussion includes both the nominal (or best-fit) and hessian error PDFs (see Sec. 3.2.2 for details).

clude new experimental data, each bin of which is treated a possible input to the PDF fit (i.e. the “pseudo-data”), and their associated statistical uncertainties. The **signal templates**³ include theoretical predictions for a given process, and are comprised of a nominal prediction and n eigenvector variations (See Fig. 13.1). In the previous chapter, the importance of $m_{\ell\ell}$, $y_{\ell\ell}$, and $\cos\theta^*$ was described, and in **ePump**, these variables are specifically used as parameters.

The individual entries of these input templates are appended to the original global χ^2 function, in order to extend it with the new data (or pseudo-data). Since the global χ^2 minimum is already known from the global fit, the new best-fit parameters can be found algebraically, along with the new eigenvectors of the Hessian matrix, leading to updated best-fit PDFs and updated error PDFs. The updated PDFs can then be used directly to assess the impact of the new data on the PDF uncertainty for various observables of interest, such as the dilepton invariant mass $m_{\ell\ell}$.

The uncertainty reduction observed in the PDFs after the update depends on the statistical precision of the “old” experimental inputs, as well as the “new” inputs from the pseudo-data used in the update. All things being equal, a dataset with a larger integrated luminosity (and hence a smaller statistical uncertainty) will be able to better constrain the PDF uncertainty, than one with a smaller integrated luminosity (and hence a larger statistical uncertainty). In the results that follow, the pseudo-data used in the update are generated with $\mathcal{L}_{\text{int}} = 3000 \text{ fb}^{-1}$ to coincide with the lifetime dataset collected at the LHC.

The second executable **OptimizePDFs** can be used to transform an input set of error PDFs into a “simplified” set, optimized for the description of a specific observable. The

³Here, the **ePump** “signal templates” are not to be confused with those used for the dilepton signal reweighting procedure developed in Section 11.2.

error PDFs of the optimized set are ranked according to their overall contribution to the total PDF error for the observable under study. As a result, this procedure simplifies the calculation of the PDF error, since the sum in Eq. 3.10 can be truncated after the first few optimized eigenvectors. A reduced PDF error set like this was used to estimate the PDF uncertainty in the dilepton analysis (See Sec. 10.8.2 and Fig. 10.11).

Because of its versatility, `ePump` offers a way to quickly and accurately assess how new datasets might impact the best-fit and error PDFs of contemporary PDF sets. This makes `ePump` the ideal tool to assess the impact of future Drell-Yan measurements at the LHC on the high-mass dilepton PDF uncertainty, the results of which are presented in Section 13.4.

13.2 PDF Update Strategy

The proceeding sections of this chapter detail the “PDF Update Analysis,” in which pseudo-data and signal templates for the Drell-Yan process are constructed and used for the PDF update with `ePump`.

The overall strategy is this: two regions in the dilepton invariant mass $m_{\ell\ell}$ are defined and treated differently; the first, a “control region,” encompasses well-measured masses and excludes the potential effects of new physics; the second, a “signal region,” extends beyond the excluded masses and is not used to predict the updated PDF set. The PDF update is performed using templates constructed in the control region, and the updated PDFs are then used to assess the reduction in PDF uncertainty in the signal region, which coincides with the high-mass search region of the dilepton analysis.

The analysis presented here is performed at “truth level,” such that the acceptance and efficiency effects associated with the reconstruction and identification of prompt leptons in

the ATLAS detector are neglected. However, leptons are very well measured, so this is an acceptable first look at this technique.

13.3 Construction of ePump Templates

In order to probe the PDFs in the three dimensions of interest: $m_{\ell\ell}$, $y_{\ell\ell}$, and $\cos\theta^*$, the ePump templates are produced in slices of the triple-differential Drell-Yan cross section of Eq. 12.10. The templates cover a finite region of kinematic phase space, consistent with the **fiducial volume** of the ATLAS detector, in order to simulate an actual future measurement. The control region considered for the present update is designed explicitly to probe the PDFs at high x and is defined by

$$40 < m_{\ell\ell} < 1000 \text{ GeV}, \quad |y_{\ell\ell}| < 3.6, \quad -1 < \cos\theta^* < 1. \quad (13.1)$$

Here, the range in dilepton mass is limited to a maximum value of 1 TeV. This “Standard Model” control region is defined in order to prevent possible contamination from high-mass new physics processes that would otherwise bias the outcome of a measurement. The final-state leptons must have a transverse momentum $p_T^\ell > 30$ GeV, and be produced in either the central-central or central-forward decay configurations (discussed further in Sec. 13.3.1).⁴ The inclusion of the $\cos\theta^*$ observable in the definition is crucial, as it provides the additional discriminating power between the u -type and d -type parton level Drell-Yan sub-processes (as in Fig. 12.5).

⁴Central leptons are defined by $|\eta| < 2.47$ while forward leptons are $2.5 < |\eta| < 4.9$.

13.3.1 Kinematic Reach at the LHC

The extent to which the PDFs can be probed at high x via the Drell-Yan mechanism is largely dependent on the fiducial region considered in the ATLAS detector. In the Drell-Yan process, for example, three distinct final-state topologies are possible within the ATLAS detector acceptance, depending on the rapidity of the intermediate γ^*/Z boson:⁵

- The **central-central** topology results when $y_{\gamma^*/Z}$ is small, so both decay leptons enter the central region ($|\eta| < 2.47$) of the ATLAS detector.
- The **central-forward** topology results when $y_{\gamma^*/Z}$ is larger, such that one lepton enters the central region ($|\eta| < 2.47$), and the other is boosted into the forward region ($2.5 < |\eta| < 4.9$) of the ATLAS detector.
- The **forward-forward** topology results when the $y_{\gamma^*/Z}$ is large, so both leptons are boosted into the forward region ($2.5 < |\eta| < 4.9$) of the ATLAS detector.

The central, or precision, region of the ATLAS detector is covered by both the tracking (See Sec. 7.3) and calorimetry (See Sec. 7.4) systems, providing for a robust electron measurement. The forward region beyond is covered only by the calorimetry system, making the reconstruction and identification of forward electrons much more challenging.

However, the use of forward leptons allows for higher values of dilepton rapidity to be measured, and hence more extreme values of x to be probed. For example, the dilepton rapidity in the central-central configuration is limited to $|y_{\gamma^*/Z}| < 2.47$. On the other hand, the use of central-forward leptons extends this range up to $|y_{\gamma^*/Z}| < 3.6$ (See Fig. 12.4).

⁵The γ^*/Z boson receives a longitudinal boost when there is a difference between x_1 and x_2 of the colliding quark and antiquark as in Eq. 12.6. As it decays, its longitudinal momentum is transferred to the final-state leptons.

$m_{\ell\ell}$ [GeV]	Max. $ y_{\ell\ell} $	x_{Min}	x_{Max}
40 - 66	3.5	9.3×10^{-5}	0.17
66 - 116	3.5	1.5×10^{-4}	0.30
116 - 250	3.5	2.7×10^{-4}	0.64
250 - 400	3.3	7.1×10^{-4}	0.83
400 - 600	2.9	1.7×10^{-3}	0.84
600 - 1000	2.5	3.8×10^{-3}	0.94

Table 13.1: The ranges in parton- x probed in the Drell-Yan process at the $\sqrt{s} = 13$ TeV LHC for various values of dilepton rapidity and four-momentum transfer $Q = m_{\ell\ell}$. The maximum values assumed for the dilepton rapidity are determined from the $|y_{\ell\ell}|$ distributions in Figure 12.4, where the Drell-Yan cross section falls to zero. The minimum and maximum values in x are calculated using Eq. 12.7.

Previous ATLAS measurements have been performed in the central region of the detector, but to obtain the highest possible sensitivity to the x values relevant at high mass, the forward region must be considered as well. Thus, the forward region motivates the choice of the fiducial volume considered in Eq. 13.1 for the generation of ePump templates.

Table 13.1 lists the values of x_{Min} and x_{Max} calculated for the $\sqrt{s} = 13$ TeV LHC in several invariant mass ranges using y_{Min} and y_{Max} for central-forward leptons. Here, it can be seen that the measurement of the Drell-Yan cross section starting at low invariant mass, such as $40 < m_{\ell\ell} < 60$ GeV, already probe values of $x \gtrsim 0.1$, and by $m_{\ell\ell} = 1$ TeV the values of x probed approach even unity.

13.3.2 Template Simulation Samples

Simulated samples for the Drell-Yan process are generated at NNLO in QCD using the RESBos MC generator and the CT14Hera2 PDF set. This provides a precise theoretical prediction of the Drell-Yan signal expectation. Additional background contributions are

neglected, but would be similar to those assessed in the dilepton analysis.⁶

As a technique to deal with the steeply falling cross section, Drell-Yan samples are generated in slices of dilepton invariant mass $m_{\ell\ell}$ to maintain high statistical precision. At the generator level, the final-state leptons are required to have a transverse momentum of $p_T > 30$ GeV and pseudorapidity of $|\eta| < 4.9$, and the dilepton system is required to be within a rapidity of $|y_{\gamma^*/Z}| < 4.9$. All MC samples are stitched together, and are generated at a CM energy of $\sqrt{s} = 13$ TeV to coincide with Run-2 of the LHC. A description of the parton shower and underlying event (See Sec. 3.1.3) that normally accompany the high energy event are not simulated, nor is the effect of QED FSR radiation.⁷

For the error analysis, RESBos provides a set of $n = 57$ cross section weights for each generated MC event, which correspond to the $n = 1$ nominal cross section prediction and the $n = 56$ eigenvector variations associated with the CT14Hera2 error PDFs (as described in Sec. 3.2.2). This is a crucial feature, since these event weights are used to assess the PDF uncertainties (via Eq. 3.10 in Sec. 3.2.2) in the fiducial volume considered.

13.3.3 Template Event Selection

ePump signal templates are constructed using events from simulation that pass a loose event selection. For each dilepton event, the final-state leptons are required to pass a minimum transverse momentum threshold and be produced in either the central-central or central-forward configurations (See Sec. 13.3.1).

For events with an invariant mass of $m_{\ell\ell} > 80$ GeV, the lepton transverse momentum requirement is $p_T > 30$ GeV, while for low-mass events that fall between $40 < m_{\ell\ell} < 80$,

⁶These include $t\bar{t}$ production, Wt Single-Top production, Diboson production, and $W+jets$ & Multi-jet production if considering the electron channel.

⁷These effects are negligible for the present study.

the requirement is $p_T > 15$ GeV. In the central-central configuration, both leptons must be within $|\eta| < 2.47$. In the central-forward configuration, one lepton must be within $|\eta| < 2.47$ while the other must be within $2.5 < |\eta| < 4.9$.

13.3.4 Template Binning

Drell-Yan events passing selection are binned in histograms and used to construct ePump templates for the update. The templates are defined in terms of the triple-differential cross section of Eq. 12.10, which is parametrized as

$$\mathcal{L}_{\text{int}} \left(\frac{d^3\sigma}{dm_{\ell\ell} dy_{\ell\ell} d\cos\theta^*} \right)_{ijk} = \frac{N_{\text{pseudo-data}}^{ijk}}{(\Delta m_{\ell\ell})_i (2\Delta|y_{\ell\ell}|)_j (\Delta \cos\theta^*)_k}, \quad (13.2)$$

where i , j , and k correspond to the bin indices of the distributions of interest: the invariant mass $m_{\ell\ell}$, the absolute rapidity $|y_{\ell\ell}|$, and the cosine of the polar angle $\cos\theta^*$, respectively.

The number of pseudo-data events are given by $N_{\text{pseudo-data}}^{ijk}$, the integrated luminosity of the pseudo-dataset is \mathcal{L}_{int} , and $(\Delta m_{\ell\ell})_i$, $(2\Delta|y_{\ell\ell}|)_j$, and $(\Delta \cos\theta^*)_k$ are the bin widths for each of the distributions of interest.⁸ The factor of two in the denominator accounts for the modulus in the rapidity bin width. The binning used to parametrize Eq. 13.2 is

- $40 < m_{\ell\ell} < 1000 : \{40, 66, 80, 91, 102, 116, 145, 200, 275, 381, 525, 725, 1000\}$
- $0 < |y_{\ell\ell}| < 2.4 : \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$
- $2.4 < |y_{\ell\ell}| < 3.6 : \{2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 3.6\}$
- $-1 < \cos\theta^* < 1 : \{-1.0, -0.7, -0.4, 0.0, 0.4, 0.7, 1.0\}$.

⁸In a realistic measurement, the numerator of Eq. 13.2 would be replaced by $N_{\text{data}}^{ijk} - N_{\text{bkg}}^{ijk}$, where the background component arises from the standard set of dilepton backgrounds (as described in Sec. 9.3.)

There are 12 bins in mass $m_{\ell\ell}$, 18 equidistant bins in rapidity $|y_{\ell\ell}|$, and 6 bins in the cosine of the polar angle $\cos \theta^*$. The total number of measurement bins is $N_{\text{bins}} = 12 \times 18 \times 6 = 1296$ over the fiducial region considered. Each bin is assigned a unique bin number from 1 to N_{bins} .

The binning scheme is motivated by several previous ATLAS cross section measurements [212, 213]. The fine-granularity binning in $|y_{\ell\ell}|$ is crucial, since the parton momentum fraction x is highly dependent on $y_{\ell\ell}$, as in Eq. 12.7. A large number of rapidity bins, therefore, allows for as much PDF shape information to be extracted as possible. Likewise, the $\cos \theta^*$ binning separates forward and backward events, providing improved sensitivity to u -type and d -type parton sub-processes. The range considered in dilepton rapidity covers both central-central and central-forward topologies, and the cosine of the polar angle $\cos \theta^*$ covers the entire acceptance.

13.3.5 Pseudo-Data Generation

The next step is to use the binned signal templates to generate pseudo-data using the Drell-Yan event samples of finite statistics. This step is sufficient to create pseudo-data templates now more sensitive to the high- x PDFs relevant to a dilepton search.

Pseudo-data are generated according to the Drell-Yan event yield in each of the $N_{\text{bins}} = 1296$ template bins defined by Eq. 13.2. For each bin considered, the Drell-Yan cross section estimate $\sigma_{\text{Drell-Yan}}^{ijk}$ is scaled by a characteristic integrated luminosity \mathcal{L}_{int} to arrive at a definite Drell-Yan event yield $N_{\text{Drell-Yan}}^{ijk}$. The resulting yield is assumed to be the mean of a Poisson distribution, which is then used to throw a random number according to Poisson statistics, thereby populating the bin with $N_{\text{pseudo-data}}^{ijk}$ events.⁹ The characteristic lumi-

⁹The probability to observe $N_{\text{pseudo-data}}$ events in a given bin is $P(N_{\text{pseudo-data}} | \sigma_{\text{Drell-Yan}} \mathcal{L}_{\text{int}})$,

nosity considered is $\mathcal{L}_{\text{int}} = 3000 \text{ fb}^{-1}$, coincident with the final dataset of the LHC. Thus, this procedure results in a set of $N_{\text{bins}} = 1296$ pseudo-data entries, each associated with the Drell-Yan theory prediction over the fiducial region defined by Eq. 13.1. This constitutes a set of pseudo-data templates ready to be used for the update.

13.3.6 Finalized Templates

With the requisite `ePump` signal and pseudo-data templates, the update can be performed as in Figure 13.1). The templates can be represented visually as single dimensional $\cos \theta^*$ histograms in slices of $m_{\ell\ell}$ and $|y_{\ell\ell}|$ according to the binning scheme outlined above in Section 13.3.4. As a concrete example, the $\cos \theta^*$ distributions in the 381-525 GeV mass slice are depicted in Figure 13.2, and Figure 13.3 for $\mathcal{L}_{\text{int}} = 3000 \text{ fb}^{-1}$ with both the Drell-Yan and pseudo-data event yields overlaid. The 3000 fb^{-1} templates constructed in the other mass slices are presented in Appendix D.

The MC simulation provides sufficient statistics over the majority of the fiducial volume considered, but there are isolated regions of phase space where the statistical uncertainty on the MC can become large. In order to not bias the update, bins in the MC estimate where the statistical uncertainty is greater than 3% are removed.

13.4 Results of the Update

With a prepared set of templates, the `UpdatePDFs` executable of `ePump` is used to assess the impact of the 3000 fb^{-1} pseudo-dataset on the `CT14Hera2` PDFs, as well as assess the reduction in the PDF systematic uncertainty in the high-mass dilepton spectrum.

where $P(n|\nu)$ is the Poisson distribution.

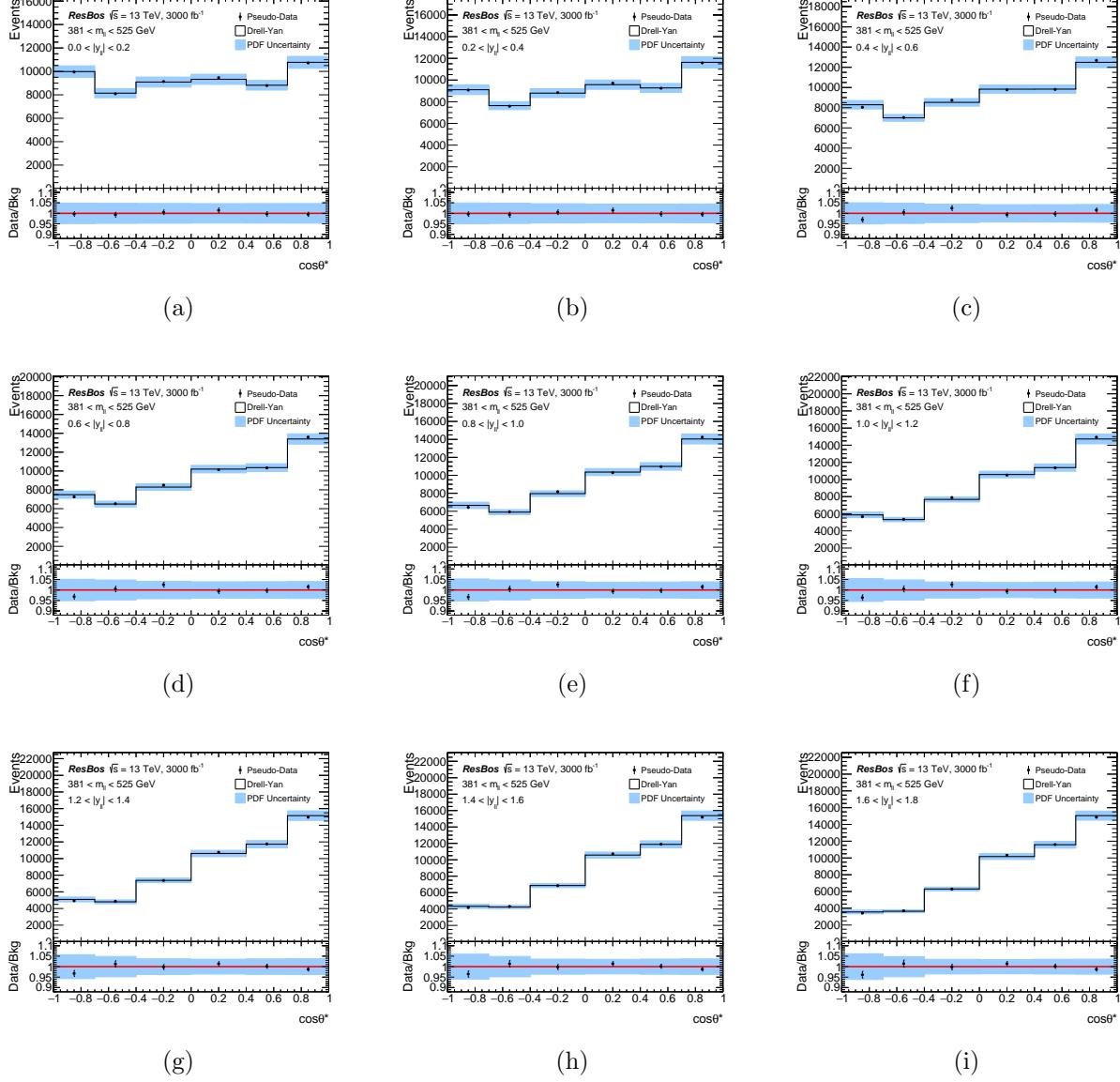


Figure 13.2: The triple-differential cross section of Eq. 13.2 in the range of $m_{\ell\ell} \in [381, 525]$ GeV and $0 < |y_{\ell\ell}| < 1.8$, scaled to 3000 fb^{-1} . The specific kinematic region is labeled in each plot (a)-(i). The RESBOS prediction is indicated by the solid line, and pseudo-data are shown as solid black markers. In each plot, the sub-panel shows the ratio between the theoretical prediction and the measurement. Error bars represent the statistical uncertainty on the pseudo-data and the solid band represents the PDF uncertainty on the dilepton event yield.

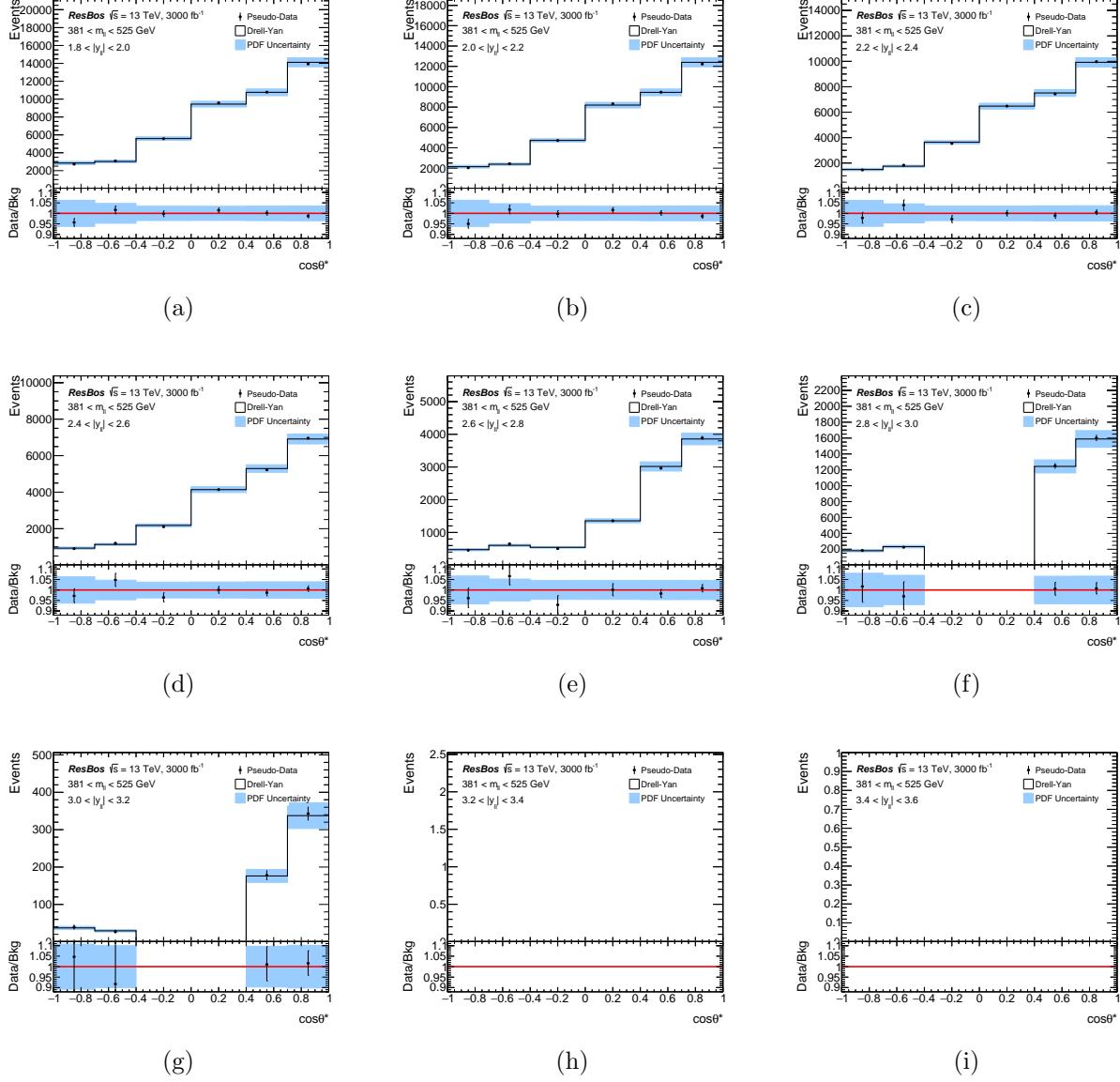


Figure 13.3: The triple-differential cross section of Eq. 13.2 in the range of $m_{\ell\ell} \in [381, 525] \text{ GeV}$ and $1.8 < |y_{\ell\ell}| < 3.6$, scaled to 3000 fb^{-1} . The specific kinematic region is labeled in each plot (a)-(i). The RESBOS prediction is indicated by the solid line, and pseudo-data are shown as solid black markers. In each plot, the sub-panel shows the ratio between the theoretical prediction and the measurement. Error bars represent the statistical uncertainty on the pseudo-data and the solid band represents the PDF uncertainty on the dilepton event yield.

13.4.1 Impact on CT14Hera2 PDFs

The results of the `ePump` update with the 3000 fb^{-1} pseudo-dataset can be seen in Figures 13.4 and 13.5, which were generated with the `PDFFigure` [214] package and depict the PDF error bands for several parton flavors (the impact of a 300 fb^{-1} pseudo-dataset can be seen in Appendix D). Of particular importance to description of high-mass Drell-Yan are the $\bar{u}(x)$ and $\bar{d}(x)$ sea and $u_v(x)$ and $d_v(x)$ valence distributions, which are the focus of the following discussion.

The sea distributions show a considerable reduction in uncertainty at high x . For example, the sea uncertainty is reduced to $\simeq 20\%$ at $x \simeq 0.5$ from the pre-update value of $\simeq 70\%$. On the other hand, the improvement in valence distributions at high x is not as dramatic; however, substantial improvement is observed in the ranges of low and intermediate x . The $u_v(x)$ distribution remains better constrained than $d_v(x)$ at high x after the update, comparing 2.6% to 11%, respectively. Table 13.2 lists the pre- and post-update uncertainties for several parton flavors and values of x explicitly.

Results are also interpreted in terms of templates with reduced dimensionality; i.e., with single-differential $\left(\frac{d\sigma}{dm_{\ell\ell}}\right)$ or double-differential $\left(\frac{d^2\sigma}{dm_{\ell\ell}dy_{\ell\ell}}\right)$ templates, as compared to the fully triple-differential templates constructed in Section 13.3. In this case, the single-differential templates are produced by integrating over the rapidity and $\cos\theta^*$ bins of Eq. 13.2, leaving the invariant mass as the only remaining dimension; likewise, the double-differential templates are produced by integrating over $\cos\theta^*$, leaving the both invariant mass and rapidity as the remainders for the update. The results obtained from each set are presented and overlaid in Figures 13.4 and 13.5, so that the relative impact of the PDF update can be assessed as each new dimension of Drell-Yan cross section is added. The addition

x	$u_v(x)$		$d_v(x)$		$\bar{u}(x)$		$\bar{d}(x)$	
	δ_{pre} [%]	δ_{post} [%]						
0.1	3.4	0.7	5.8	1.5	9.8	2.2	11	3.8
0.3	2.6	0.9	7.5	3.6	30	8.3	32	11
0.5	4.8	2.6	16	11	71	20	69	20
0.7	12	7.0	45	30	280	77	250	67

Table 13.2: Impact of 3000 fb^{-1} update on the **CT14Hera2** $u_v(x)$ and $d_v(x)$ valence and $\bar{u}(x)$ and $\bar{d}(x)$ sea distributions for several values of x using the standard triple-differential templates at $Q = 3 \text{ TeV}$. To be compared with Figures 13.4 and 13.5.

of the $\cos \theta^*$ dimension provides for a dramatic reduction in PDF uncertainty as compared to what could otherwise be achieved in a single- or double-differential measurement alone.

Furthermore, the effect of $\cos \theta^*$ *itself* on the performance of the PDF update is studied. The results are presented in Figures 13.6 and 13.7, which compare the reduction in PDF uncertainty using templates based on the double-differential cross sections $\left(\frac{d^2\sigma}{dm_{\ell\ell} d\cos \theta^*}\right)$ and $\left(\frac{d^2\sigma}{dm_{\ell\ell} d|y_{\ell\ell}|}\right)$. Just as before, the triple-differential templates of Section 13.3.4 are used to produce the reduced dimensionality templates here by integrating out the rapidity $y_{\ell\ell}$ or the $\cos \theta^*$ dimensions explicitly from Eq. 13.2. Each template retains its binning in the mass dimension to accommodate the steeply falling Drell-Yan cross section.

The results show the uncertainty on the $u_v(x)$ and $d_v(x)$ valence distribution is reduced in the low and intermediate x ranges in the $\cos \theta^*$ dimension, while the rapidity $y_{\ell\ell}$ dimension provides for a slightly better reduction at high x , especially on the $u_v(x)$ uncertainty. On the other hand, the uncertainties on the $\bar{u}(x)$ and $\bar{d}(x)$ sea distributions each show a similar level of reduction at both low and high x , independent of the update performed. Taken together, these results indicate that $\cos \theta^*$ by itself is able to constrain the PDF uncertainty to a degree comparable with $y_{\ell\ell}$, and in some cases better, confirming its discriminating power that was suspected in the preceding chapter (See Sec. 12.1.6)

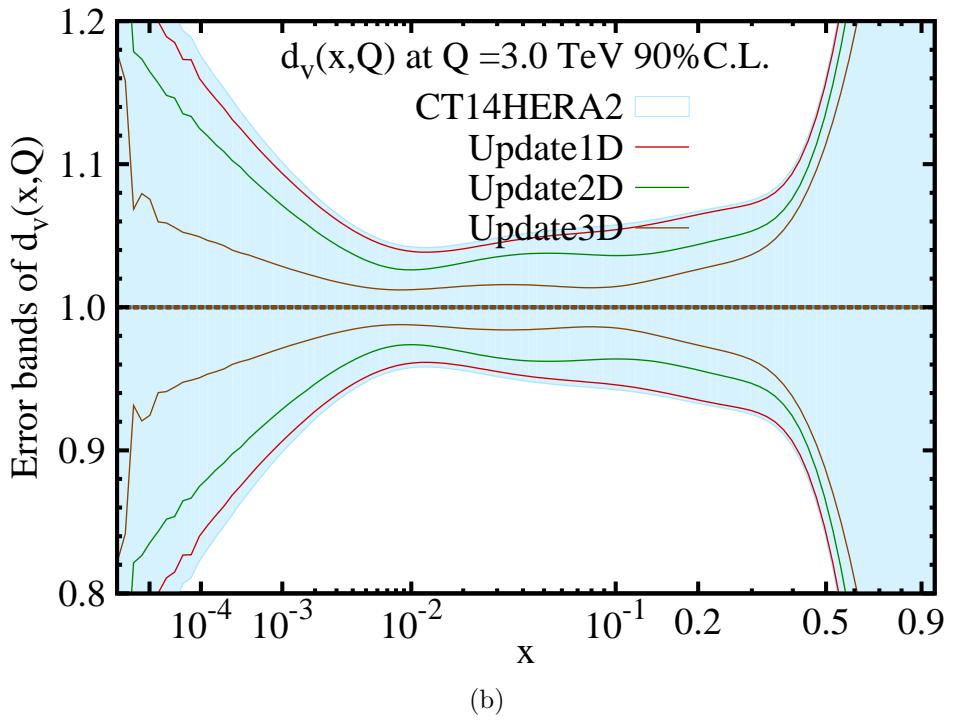
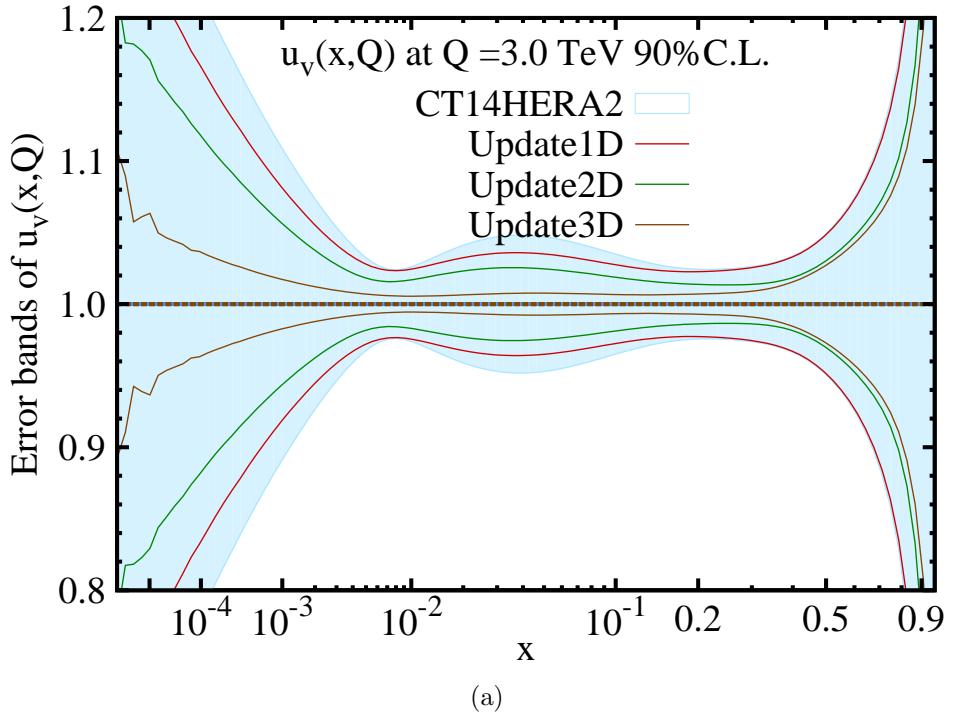


Figure 13.4: Impact of the 3000 fb^{-1} update on the CT14Hera2 $u_v(x)$ (a) and $d_v(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos\theta^*$ templates: “Update3D” at $Q = 3 \text{ TeV}$. The utilization of the $\cos\theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty.

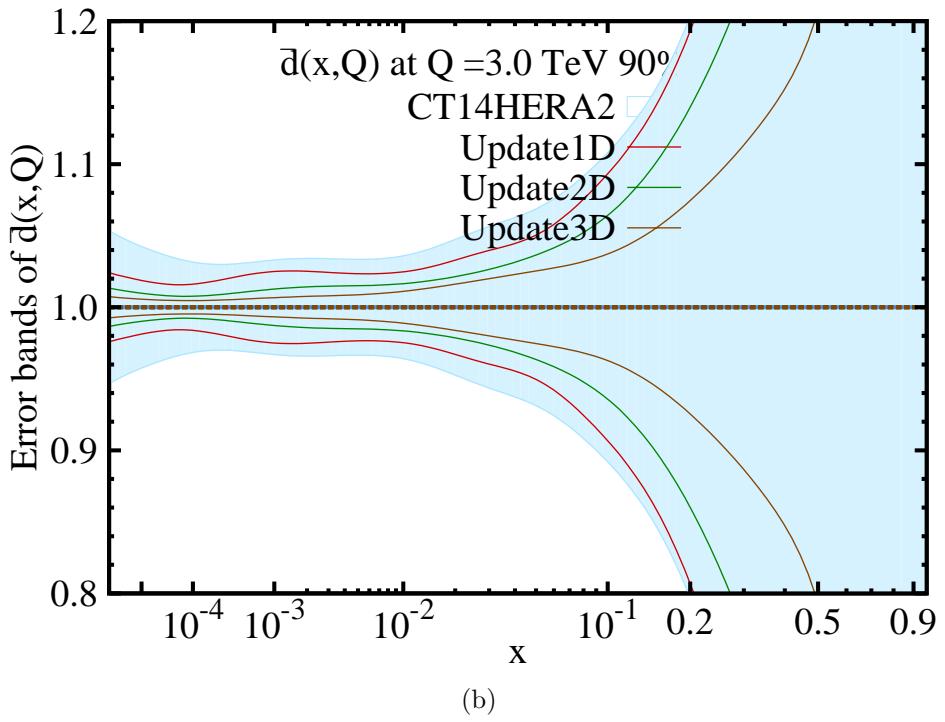
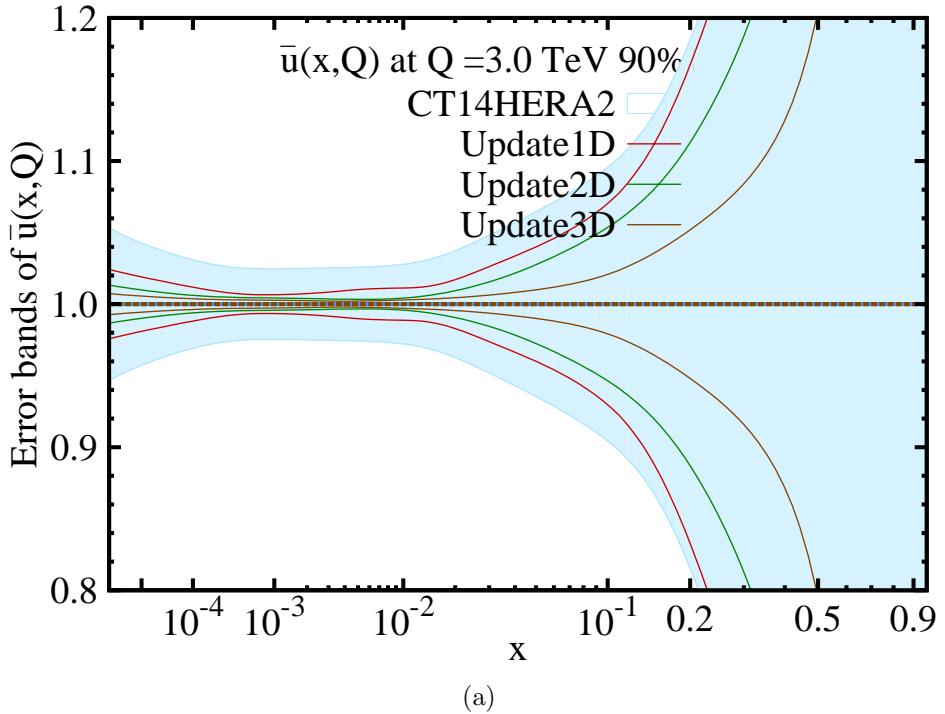
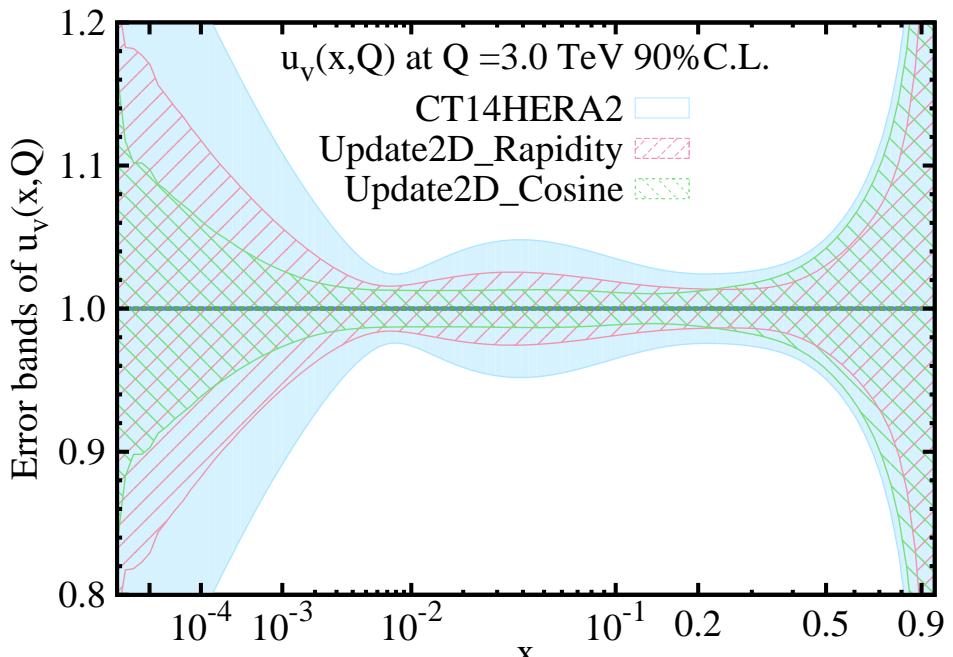
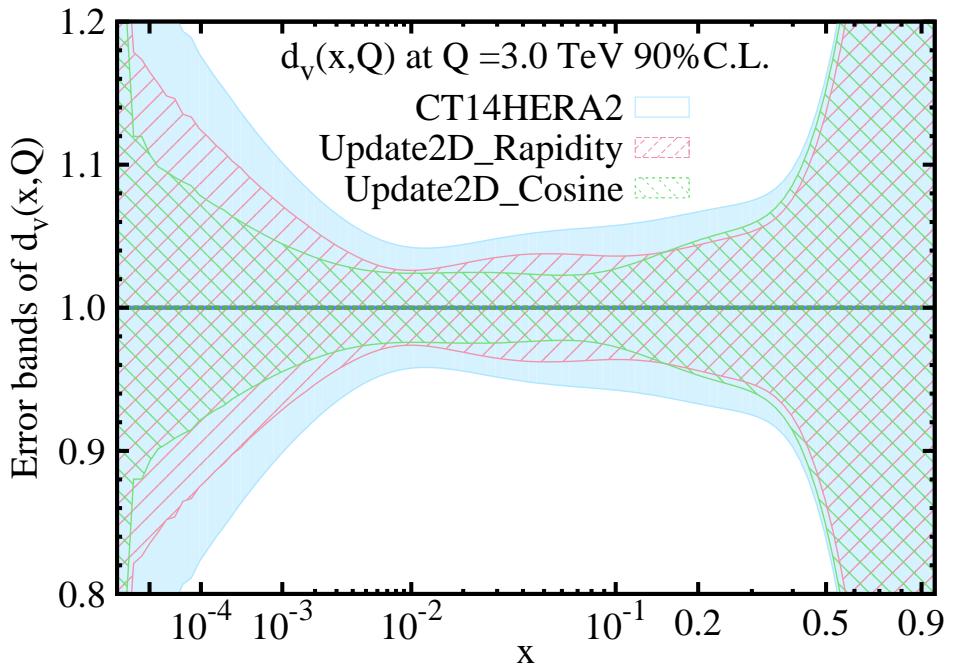


Figure 13.5: Impact of the 3000 fb^{-1} update on the CT14Hera2 $\bar{u}(x)$ (a) and $\bar{d}(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos\theta^*$ templates: “Update3D” at $Q = 3$ TeV. The utilization of the $\cos\theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty.



(a)



(b)

Figure 13.6: Impact of the 3000 fb^{-1} update on the CT14Hera2 $u_v(x)$ (a) and $d_v(x)$ (b) distributions using the double-differential mass and rapidity templates: “Update2D_Rapidity,” and double-differential mass and $\cos\theta^*$ templates: “Update2D_Cosine” at $Q = 3 \text{ TeV}$.

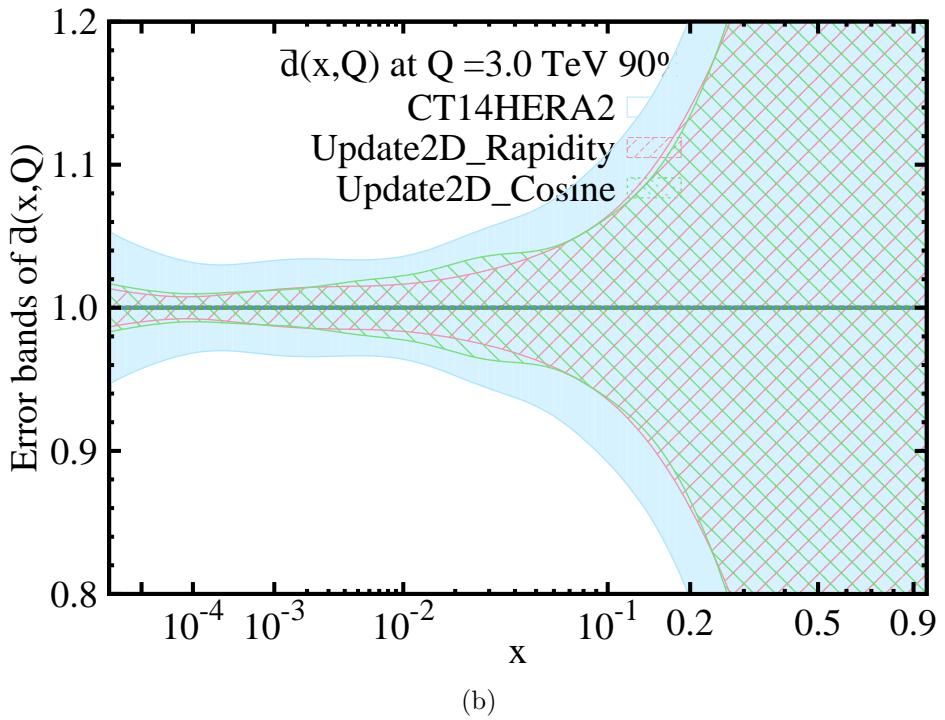
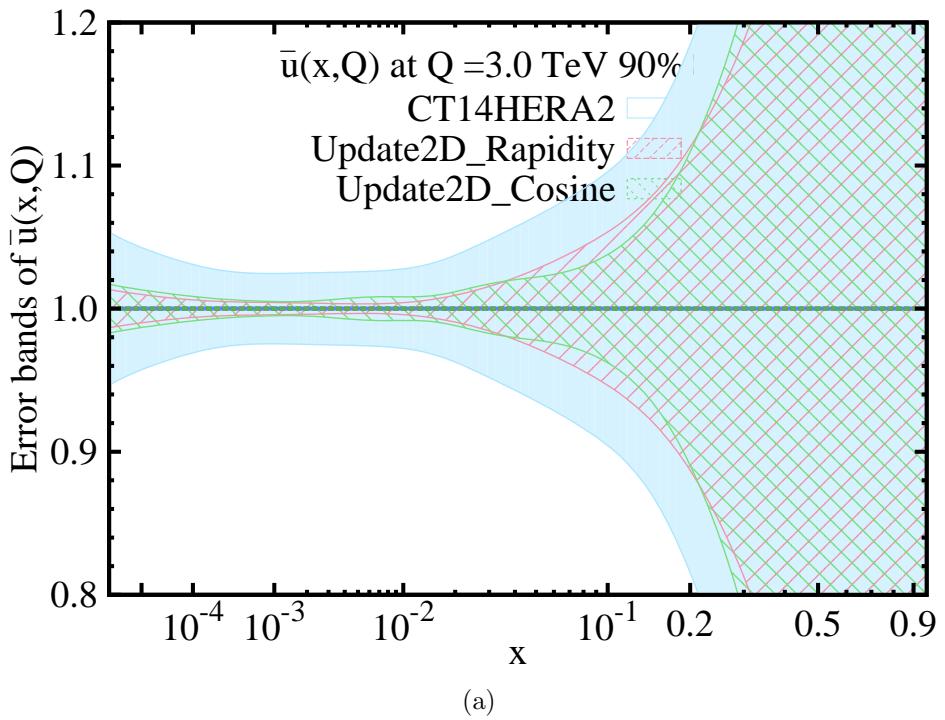


Figure 13.7: Impact of the 3000 fb^{-1} update on the CT14Hera2 $\bar{u}(x)$ (a) and $\bar{d}(x)$ (b) distributions using the double-differential mass and rapidity templates: “Update2D_Rapidity,” and double-differential mass and $\cos\theta^*$ templates: “Update2D_Cosine” at $Q = 3 \text{ TeV}$.

The bottom line is this: a global fit of PDFs which includes medium mass Drell-Yan LHC data, with rapidity and $\cos \theta^*$ sensitivity, would dramatically improve the precision of the up and down PDFs. No other data inputs are capable of achieving this improvement.

13.4.2 Impact on High-Mass Drell-Yan

With an updated set of PDFs, the systematic uncertainty on the high-mass Drell-Yan cross section prediction can now be re-assessed. As part of the update, `ePump` is able to propagate the reduced PDF uncertainty directly to an observable of interest, in this case the dilepton invariant mass. In order to make contact with the dilepton analysis, the invariant mass distribution assessed here utilizes leptons that originate in the central-central decay configuration, as these are precisely the ones selected by the dilepton event selection (See Sec. 10.5).

The results are presented in Figure 13.8, which shows the impact of the 3000 fb^{-1} pseudo-dataset on the high-mass PDF systematic uncertainty (the impact of the 300 fb^{-1} pseudo-dataset can be seen in Appendix D). The PDF uncertainty is evaluated at several characteristic values of dilepton mass, which are listed in Table 13.3. At $m_{\ell\ell} = 5 \text{ TeV}$, the PDF systematic uncertainty is reduced from 31% to 8.9%, a reduction of roughly a factor of 3.5. Similarly, at $m_{\ell\ell} = 3 \text{ TeV}$, the uncertainty is reduced from 15% to 3.7%, roughly a factor of 4. In each case, a substantial improvement is obtained compared to the current state-of-the-art predictions (as depicted in Fig 12.8). The PDF uncertainty assessed in the dilepton analysis is, for example, 13% and 29% at 3 and 5 TeV, respectively (See Fig. 10.11).

Furthermore, Figure 13.8 displays the dilepton invariant mass with the analysis strategy of imposing a $\cos \theta^* > 0$ requirement, where the error is assessed with the updated PDFs (as first explored in Sec. 12.1.6). Such a combination provides for a marginal, but non-trivial

$m_{\ell\ell}$ [TeV]	CC Selection		CC+ $\cos \theta^*$ Selection	
	δ_{pre}^{PDF} [%]	δ_{post}^{PDF} [%]	δ_{pre}^{PDF} [%]	δ_{post}^{PDF} [%]
1	5.9	1.0	5.6	0.9
2	9.6	2.0	8.9	1.7
3	15	3.7	13	3.2
4	22	6.0	20	5.3
5	31	8.9	28	8.0

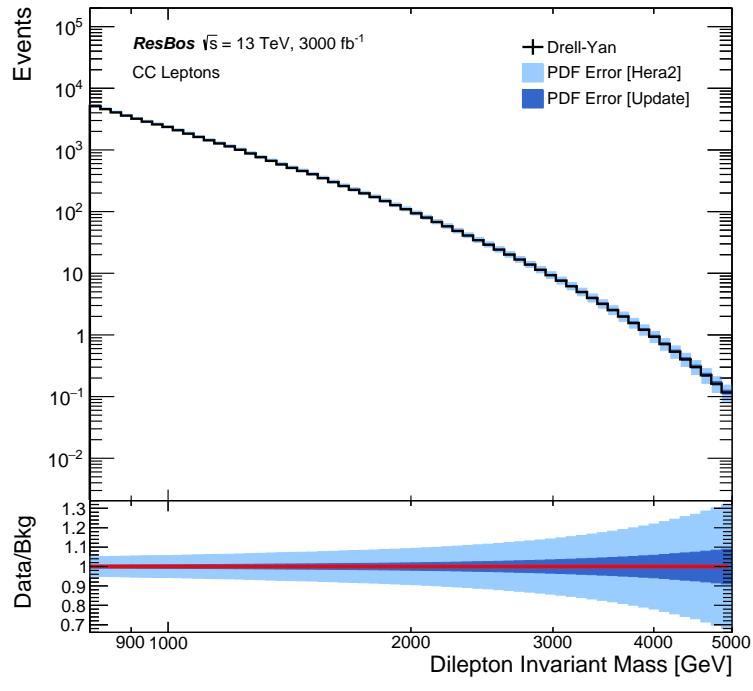
Table 13.3: The estimated PDF uncertainty in several invariant mass bins for the distributions shown in Figure 13.8. The current CT14Hera2 uncertainty estimates are shown in the first column, the result of the 3000 fb^{-1} update is shown next, and the 3000 fb^{-1} update with an additional $\cos \theta^* > 0$ requirement on the dilepton selection is shown last. Notice the pre-update values for the central-central (CC) selection are consistent with those assessed in the dilepton analysis (See Fig. 10.11).

further reduction in the PDF uncertainty at high mass at the expense of a reduction in statistics. Table 13.3 quantifies the PDF uncertainty before and after the update in several high-mass bins of interest for each of these selections.

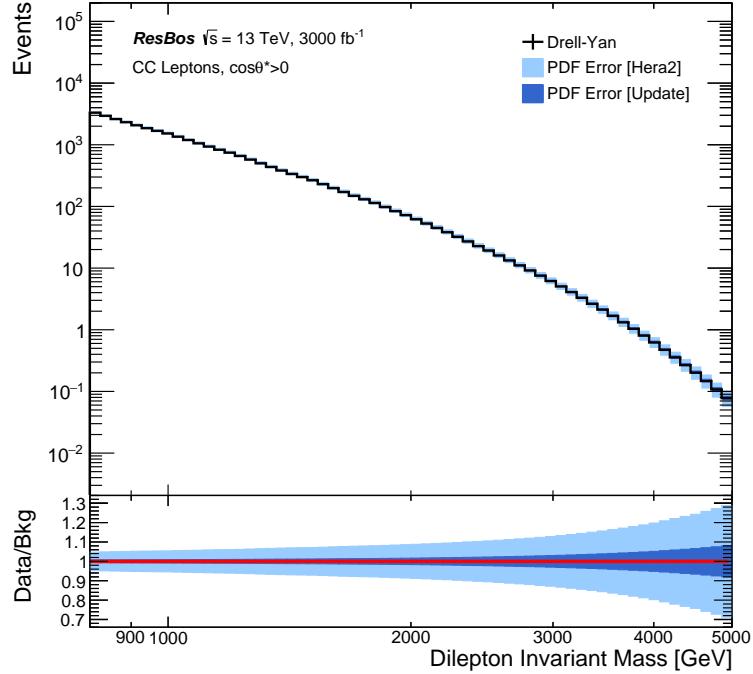
Finally, further improvement at high mass could be obtained by extending the nominal 1 TeV control region if, for example, the presence of new physics is definitely ruled out. The bottom line is this: The heretofore enormous PDF uncertainties in any BSM search at high mass can be reduced *below* the current experimental uncertainties with the use of LHC Drell-Yan inputs from a possible future PDF global fit.

13.5 Outlook

The impact of a future Drell-Yan cross section measurement on the CT14Hera2 PDF uncertainty was assessed using the `ePump` package at the $\sqrt{s} = 13 \text{ TeV}$ LHC with 3000 fb^{-1} of Drell-Yan pseudo-data. The phase space region considered for the PDF update was based on three variables: the dilepton mass $m_{\ell\ell}$, the dilepton rapidity $y_{\ell\ell}$, and the cosine of the polar angle $\cos \theta^*$. This region was divided into 1296 histogram bins and used to construct `ePump`



(a)



(b)

Figure 13.8: The dilepton invariant mass distribution for central-central dilepton events (a), as would be selected by the dilepton event selection of Section 10.5, and with an additional $\cos \theta^* > 0$ requirement added to the selection (b). The ratio sub-plot depicts the CT14Hera2 PDF uncertainty before and after the 3000 fb^{-1} update.

$m_{\ell\ell}$ [TeV]	CC Selection δ_{post}^{PDF} [%]	Dilepton Analysis				
		δ^{PDF} [%]	δ^{Choice} [%]	δ^{Theory} [%]	δ^{Exp} [%]	δ^{Total} [%]
1	1.0	5.4	0.0	5.9	7.7	9.7
2	2.0	8.7	0.0	9.8	11	15
3	3.7	13	0.0	15	12	19
4	6.0	19	8.4	23	13	26
5	8.9	29	47	57	14	59

Table 13.4: The post-update PDF uncertainty as compared to the experimental and dominant theoretical uncertainties in the electron channel of the dilepton analysis. As the PDF uncertainty will be reduced well below the current experimental uncertainty, attention will be shifted to the reduction of others, such as the “PDFChoice” uncertainty, improving the discovery potential of future iterations of the dilepton analysis.

pseudo-data and signal templates, which were designed to probe the PDFs in the extreme kinematic regions of (x, Q^2) only accessible at the LHC.

The CT14Hera2 PDF set was used for the update. The results showed a significant reduction in the uncertainties associated with all parton flavors, especially $\bar{u}(x)$ and $\bar{d}(x)$ sea at high x . Likewise, these reduced PDF uncertainties, when propagated to the high-mass dilepton spectrum, lead to a significantly improved estimate of the event yield.

These proof-of-concept results indicate a great deal of improvement can still be obtained from precision PDFs measurements at LHC. The use of $\cos \theta^*$ as an additional dimension in future PDF global fits is absolutely crucial, as it supplements the more standard double-differential measurements in invariant mass and rapidity; when used in conjunction, as was done here, the reduction in uncertainty can be dramatic.

For these reasons, Drell-Yan cross section measurements could be vital to the success of future searches and measurements at the LHC. Not only will the PDF uncertainty that affects the high-mass dilepton analysis be reduced, improving the discovery potential of many non-resonant new physics models, but the inclusion of new and robust data into the modern PDF global fits will bring the uncertainty estimates of the various global fitting groups into

better agreement (See Fig. 10.12).

Such an opportunity would call for a re-assessment of the “PDFChoice” uncertainty, which may further improve the dilepton discovery potential, as the largest uncertainty would no longer be due to the PDFs. Table 13.4 compares these uncertainties explicitly.¹⁰ Therefore, for the reasons outlined in this chapter, experiments at the LHC should strongly consider performing precision measurements of the Drell-Yan triple-differential cross section in order to further constrain the PDF uncertainties in future PDF global fits.

¹⁰The uncertainty on the QCD background estimate is not included in the above calculations.

Part V

The Conclusion

Chapter 14

Conclusion

In this final chapter, a conclusion is given by summarizing the results obtained in the dilepton analysis, and a brief outlook to the future of the analysis and the broader LHC physics program is provided.

14.1 Summary

The Standard Model is an extremely successful theory, but we know it cannot be the *final* theory of elementary particles and interactions. As a consequence of addressing the experimental and/or conceptual limitations of the Standard Model, Beyond the Standard Model (BSM) theories often predict striking new phenomena, such as new particles or interactions, many of which are potentially observable at the LHC.

This search performed in this dissertation focused on a class of BSM theories that extend the $SU(3) \times SU(2) \times U(1)$ Standard Model gauge group with one or more $U(1)'$ gauge symmetries. These theories predict new heavy gauge bosons, known as Z' bosons, which potentially interact with and decay into pairs of charged leptons. The smoking-gun signature of Z' production at the LHC would be the appearance of a resonant peak in the dilepton invariant mass spectrum.

A search for a new high-mass resonance was performed in dielectron final-states using 36.1 fb^{-1} of proton-proton collision data. The dataset was recorded with the ATLAS detector

during the 2015 and 2016 data-taking runs of the LHC at a CM energy of $\sqrt{s} = 13$ TeV at CERN in Geneva, Switzerland.

The dielectron invariant mass spectrum was measured and the highest-mass event observed was 2.9 TeV. The Standard Model background was estimated with MC simulation and data-driven techniques. The dominant irreducible background originates with the production of high- E_T dielectron pairs via the Drell-Yan process. Additional backgrounds arise from $t\bar{t}$ production, Single-Top production, Diboson production, and QCD Multi-jet and $W+jets$ production, in which one or more jets of the event are misidentified as electrons.

The dielectron invariant mass spectrum was compared to data, and possible deviations were quantified using a log-likelihood ratio (LLR) test and the **BUMPHUNTER** package. According to the LLR test, the largest local excess occurred at a mass of $M_{Z'} = 2.37$ TeV with a p-value of 0.6% or a significance of 2.5σ . The global significance of the excess was negligible, and hence the data were found to be consistent with the Standard Model. A similar result was obtained with the **BUMPHUNTER**.

As no significant excess was observed, 95% C.L. upper limits were set on several Z' model parameters of interest. These included the cross section times branching fraction to electrons σB in the SSM and E_6 models and the relative coupling strength γ' in the Minimal Z' model. Limits were set using the electron channel dataset and later statistically combined with the muon channel for maximal sensitivity.

For the SSM, masses below 4.5 TeV were excluded, and for the E_6 models, lower mass limits were set between 3.8 and 4.1 TeV for the Z'_ψ and Z'_χ , respectively. Other E_6 motivated models are constrained between these values. For the Minimal model, a range of mass limits were set depending on the assumed value of the relative coupling strength γ' . In particular, for the Z'_{3R} and Z'_{B-L} models, masses below 3.9 and 4.0 TeV were excluded. While no model

was definitely ruled out, these limits help to strongly restrict the allowed Z' parameter space and are currently some of the most stringent limits set to date (See Table 5.5).

The sensitivity of the dilepton search will continue to improve as additional data are collected in subsequent LHC runs. However, as events increasingly populate the high-mass dilepton tail, knowledge of the Parton Distribution Functions (PDFs) will become increasingly important. While the current analysis is relatively insensitive to the theoretical uncertainty of the PDFs, the sensitivity of future analyses will eventually weaken and plateau, unless the PDF uncertainty can be sufficiently reduced.

To this end, `ePump`, which can emulate a full PDF global fit, was used to produce an updated set of `CT14Hera2` PDFs that incorporate 3000 fb^{-1} of Drell-Yan pseudo-data in the dimensions of dilepton mass m_{ee} , rapidity y_{ee} , and $\cos\theta^*$. The updated PDFs were then used to assess the reduction in PDF uncertainty in the high-mass search region of interest of the dilepton analysis.

The results indicate a significant reduction is possible, such that, for example, at an invariant mass of $m_{ee} = 5 \text{ TeV}$, the current `CT14Hera2` uncertainty of 31% is reduced to 8.9%, a roughly factor of 4 improvement. As a result, the PDFs would no longer be the dominant theoretical uncertainty at high mass, thereby shifting the focus to reduction of the others. Thus, a greatly improved understanding of the PDFs may likely be obtained from future Drell-Yan measurements at the LHC, ensuring that future iterations of the dilepton search are able to maintain a high discovery potential.

14.2 Outlook

Run-2 of the LHC saw an increase of CM energy from $\sqrt{s} = 8$ to 13 TeV. This upgrade dramatically enhanced the sensitivity of the dilepton search to new high-mass phenomena. By the end of 2016, a total of 36.1 fb^{-1} had been recorded by the ATLAS detector. This allowed some of the strongest exclusion limits on the existence of Z' bosons to be set to date. The 2017 run is now complete, which collected an additional 50 fb^{-1} [215], and by the end of 2018, a total of $\simeq 150 \text{ fb}^{-1}$ is expected, which will conclude Run-2 at the LHC.

As more data are collected, more comprehensive searches can be performed. With increased statistics, the cosine of the polar angle $\cos \theta^*$ could be used, for example, as a discriminating variable in conjunction with the dilepton invariant mass $m_{\ell\ell}$ to enhance the sensitivity to a potential high-mass signals via interference effects [216]. Likewise, performing the search in exclusive channels, such a $Z' + n$ jets also becomes possible, as well as the application of b -jet vetoes or \cancel{E}_T cuts to further reduce the Top and Diboson Standard Model backgrounds.

Alternatively, to achieve even more powerful exclusion limits using the existing data, **multi-channel combinations** are being considered for the first time. Just as the electron and muon channels are statistically combined within the dilepton analysis, search channels can also be combined *across* analyses. One of the first attempts at this come from the lepton+MET and dilepton searches [217].

This is the first in a series of more ambitious and ongoing grand combinations, which aim to combine not only the $\ell\nu$ and $\ell\ell$ leptonic channels, but also the VV and VH Diboson channels¹ for maximal sensitivity [218]. Future grand combinations might benefit from

¹Here, V denotes a W or Z boson.

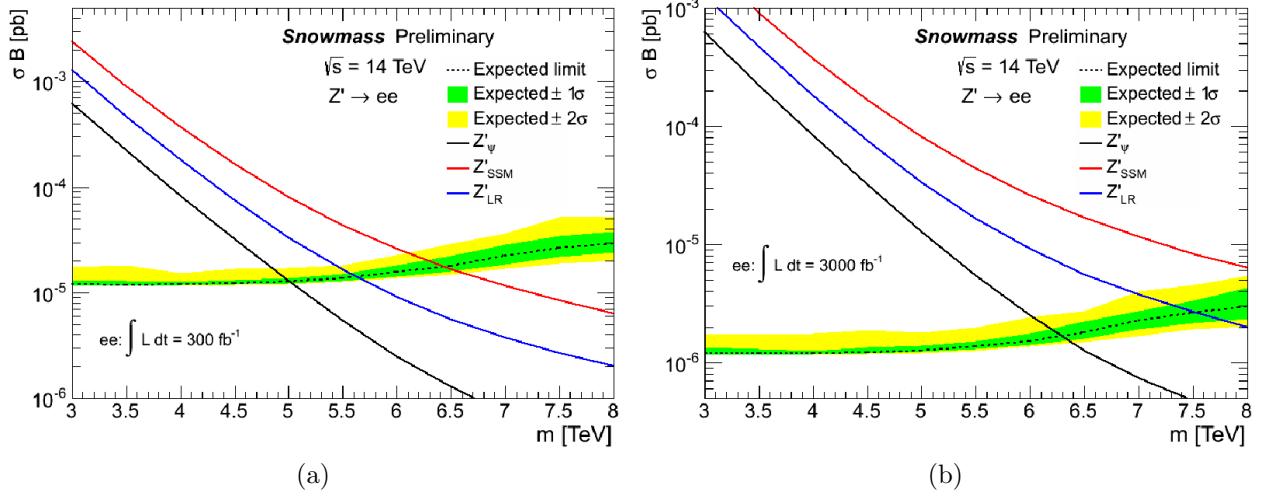


Figure 14.1: Expected 95% C.L. upper limits on σB with 300 fb^{-1} of data (a) and 3000 fb^{-1} of data (b). Theory curves are shown for the Z'_ψ , Z'_{LR} , and Z'_{SSM} models. Mass limits can be extracted from the intersection between the expected limit and the theory prediction. The region above each curve is excluded at 95% C.L. [32].

folding in additional search channels as well (e.g. $\tau\tau$, $t\bar{t}$, tb , bb , jj , etc.). Combinations like this provide a more comprehensive picture of the excluded parameter space, and will greatly help with the interpretation of future data.

Looking further ahead, Run-3 starts in 2021, after a two year long shutdown for LHC and detector related upgrades, where for the first time it is anticipated that the machine will run with a CM energy of $\sqrt{s} = 14$ TeV. The Run-3 goal is to collect an additional 150 fb^{-1} for a total of 300 fb^{-1} . After Run-3, the LHC and detectors will be upgraded once again for the **High-Luminosity LHC** (HL-LHC) program, which will increase the LHC luminosity by a factor of 10 beyond its original design value, with a goal of collecting 3000 fb^{-1} by the mid-2030s [219]. These Phase-2 upgrades will greatly increase the potential for discovery and the possible study of new exotic phenomena.

On the other hand, if no new physics is to be found in the 300 fb^{-1} (3000 fb^{-1}) dataset, the LHC is expected to exclude the Z'_ψ below a mass of 5.01 (6.29) TeV, the Z'_{LR} below 5.62

(7.52) TeV, and the Z'_{SSM} below 6.44 (8.55) TeV in the electron channel alone [32]. The inclusion of the muon channel would improve these limits further still. Figure 14.1 depicts the upper cross section limits for these three models given 300 fb^{-1} and 3000 fb^{-1} of data collected at the $\sqrt{s} = 14 \text{ TeV}$ LHC.

Of course the Phase-2 improvements will require the heroic efforts of the LHC collaborations, as the tremendous gain in instantaneous luminosity comes with its own challenges, such as an extremely high event pile-up. In Run-3 alone, pile-up is expected to occur at an average rate of 70 interactions per bunch crossing, and 140 is foreseen at the HL-LHC. These high pile-up scenarios call for extensive modifications to almost every aspect of the ATLAS detector, especially in the areas of triggering, tracking, and reconstruction, in order to accommodate the higher number of charged tracks and collision vertices per event. Research and development in these areas is currently ongoing to ensure the successful operation of the LHC now and into the foreseeable future.

APPENDICES

Appendix A

Properties of Analysis Simulation and Data

Monte Carlo Samples

This sections lists the properties of the Monte Carlo (MC) samples used in the modeling of the electron channel of the dilepton analysis. The NLO Drell-Yan samples are listed in Table A.1, the Z' samples in Table A.2, the LO Drell-Yan samples in Table A.3, the Top samples in Table A.4, and the Diboson samples in Table A.5. Each table lists the process, the generated range in invariant mass, the ATLAS dataset ID (DSID), the sample cross section σB , and the total number of events associated with each sample. The integrated luminosity of a sample can be calculated from $\mathcal{L}_{\text{int}} = N_{\text{evt}} / (\sigma B)$.

Data Samples

This sections lists the datasets used in the dilepton analysis, as recorded at $\sqrt{s} = 13$ TeV by the ATLAS detector during the 2015 and 2016 runs of the LHC. The datasets, along with their period, associated runs, total integrated luminosity, and year recorded, are listed in Table A.6.

It should be noted that the central value of 33257 pb^{-1} for the 2016 dataset was updated

Process	m_{ee} [GeV]	DSID	σB [pb]	N_{evt} [10^3]
$Z \rightarrow ee$	-	361106	$1.901 \times 10^{+03}$	2000
$\gamma^*/Z \rightarrow ee$	120-180	301020	$1.748 \times 10^{+01}$	500
$\gamma^*/Z \rightarrow ee$	180-250	301021	$2.921 \times 10^{+00}$	250
$\gamma^*/Z \rightarrow ee$	250-400	301022	$1.082 \times 10^{+00}$	150
$\gamma^*/Z \rightarrow ee$	400-600	301023	1.955×10^{-01}	100
$\gamma^*/Z \rightarrow ee$	600-800	301024	3.740×10^{-02}	150
$\gamma^*/Z \rightarrow ee$	800-1000	301025	1.061×10^{-02}	50
$\gamma^*/Z \rightarrow ee$	1000-1250	301026	4.258×10^{-03}	50
$\gamma^*/Z \rightarrow ee$	1250-1500	301027	4.422×10^{-03}	50
$\gamma^*/Z \rightarrow ee$	1500-1750	301028	5.452×10^{-04}	50
$\gamma^*/Z \rightarrow ee$	1750-2000	301029	2.299×10^{-04}	50
$\gamma^*/Z \rightarrow ee$	2000-2250	301030	1.039×10^{-04}	50
$\gamma^*/Z \rightarrow ee$	2250-2500	301031	4.940×10^{-05}	50
$\gamma^*/Z \rightarrow ee$	2500-2750	301032	2.445×10^{-05}	50
$\gamma^*/Z \rightarrow ee$	2750-3000	301033	1.249×10^{-05}	50
$\gamma^*/Z \rightarrow ee$	3000-3500	301034	1.003×10^{-05}	50
$\gamma^*/Z \rightarrow ee$	3500-4000	301035	2.934×10^{-06}	50
$\gamma^*/Z \rightarrow ee$	4000-4500	301036	8.976×10^{-07}	50
$\gamma^*/Z \rightarrow ee$	4500-5000	301037	2.807×10^{-07}	50
$\gamma^*/Z \rightarrow ee$	5000+	301038	1.265×10^{-07}	50

Table A.1: The NLO PowHEG+PYTHIA8 Drell-Yan MC samples used in the modeling of the electron channel [4].

Process	$M_{Z'}$ [GeV]	DSID	σB [pb]	N_{evt} [10^3]
$Z'_\chi \rightarrow ee$	2000	301215	8.8463×10^{-03}	20
$Z'_\chi \rightarrow ee$	3000	301246	8.1018×10^{-04}	20
$Z'_\chi \rightarrow ee$	4000	301217	1.0375×10^{-04}	20
$Z'_\chi \rightarrow ee$	5000	301218	1.8298×10^{-05}	20

Table A.2: The LO PYTHIA8 Z'_χ MC samples used to validate the electron channel reweighting procedure [4].

Process	m_{ee} [GeV]	DSID	σB [pb]	N_{evt} [10^3]
$\gamma^*/Z \rightarrow ee$	70-120	301540	$1.4260 \times 10^{+03}$	250
$\gamma^*/Z \rightarrow ee$	120-180	301541	$1.3836 \times 10^{+01}$	250
$\gamma^*/Z \rightarrow ee$	180-250	301542	$2.3332 \times 10^{+00}$	250
$\gamma^*/Z \rightarrow ee$	250-400	301543	8.6745×10^{-01}	150
$\gamma^*/Z \rightarrow ee$	400-600	301544	1.5568×10^{-01}	100
$\gamma^*/Z \rightarrow ee$	600-800	301545	2.9710×10^{-02}	50
$\gamma^*/Z \rightarrow ee$	800-1000	301546	8.3465×10^{-03}	50
$\gamma^*/Z \rightarrow ee$	1000-1250	301547	3.3105×10^{-03}	50
$\gamma^*/Z \rightarrow ee$	1250-1500	301548	1.0982×10^{-03}	50
$\gamma^*/Z \rightarrow ee$	1500-1750	301549	4.1872×10^{-04}	50
$\gamma^*/Z \rightarrow ee$	1750-2000	301550	1.7595×10^{-04}	50
$\gamma^*/Z \rightarrow ee$	2000-2250	301551	7.9961×10^{-05}	50
$\gamma^*/Z \rightarrow ee$	2250-2500	301552	3.8234×10^{-05}	50
$\gamma^*/Z \rightarrow ee$	2500-2750	301553	1.9048×10^{-05}	50
$\gamma^*/Z \rightarrow ee$	2750-3000	301554	9.8535×10^{-06}	50
$\gamma^*/Z \rightarrow ee$	3000-3500	301555	8.0449×10^{-06}	50
$\gamma^*/Z \rightarrow ee$	3500-4000	301556	2.4190×10^{-06}	50
$\gamma^*/Z \rightarrow ee$	4000-4500	301557	7.5769×10^{-07}	50
$\gamma^*/Z \rightarrow ee$	4500-5000	301558	2.4326×10^{-07}	50
$\gamma^*/Z \rightarrow ee$	5000+	301559	1.1667×10^{-07}	50

Table A.3: The LO PYTHIA8 Drell-Yan MC samples used to construct reweighted Z' signal templates for the electron channel search and limit setting of the dilepton analysis [4].

Process	m_{ee} [GeV]	DSID	σB [pb]	N_{evt} [10^3]
$t\bar{t} \rightarrow \ell X$	-	410009	$7.333 \times 10^{+01}$	20000
$Wt \rightarrow \ell X$	-	410015	$3.584 \times 10^{+00}$	1000
$W\bar{t} \rightarrow \ell X$	-	410016	$3.581 \times 10^{+00}$	1000

Table A.4: The NLO Powheg+PYTHIA6 $t\bar{t}$ and Single-Top MC samples used in the modeling of the electron channel in the dilepton analysis. A generator-level filtering efficiency of 0.10534 is accounted for in the $t\bar{t}$ sample cross section estimate [4].

Process	m_{ee} [GeV]	DSID	σB [pb]	N_{evt} [10^3]
$VV \rightarrow e\nu e\nu$	50-150	303014	8.2795×10^{-01}	200
$VV \rightarrow e\nu e\nu$	150-500	303015	2.3217×10^{-01}	50
$VV \rightarrow e\nu e\nu$	500-1000	303016	9.3993×10^{-03}	50
$VV \rightarrow e\nu e\nu$	1000-2000	303017	1.1795×10^{-03}	50
$VV \rightarrow e\nu e\nu$	2000-3000	303018	1.2414×10^{-04}	50
$VV \rightarrow e\nu e\nu$	3000-4000	303019	2.7399×10^{-05}	50
$VV \rightarrow e\nu e\nu$	4000-5000	303020	6.4562×10^{-06}	50
$VV \rightarrow e\nu e\nu$	5000+	303021	1.3890×10^{-06}	50
$VV \rightarrow \ell\nu ee$	50-150	303022	$1.3096 \times 10^{+00}$	200
$VV \rightarrow \ell\nu ee$	150-500	303023	2.0000×10^{-02}	50
$VV \rightarrow \ell\nu ee$	500-1000	303024	6.4948×10^{-04}	50
$VV \rightarrow \ell\nu ee$	1000-2000	303025	4.9253×10^{-05}	50
$VV \rightarrow \ell\nu ee$	2000-3000	303026	1.3996×10^{-06}	50
$VV \rightarrow \ell\nu ee$	3000-4000	303027	8.2389×10^{-08}	50
$VV \rightarrow \ell\nu ee$	4000-5000	303028	6.3062×10^{-09}	50
$VV \rightarrow \ell\nu ee$	5000+	303029	5.4441×10^{-10}	50
$VV \rightarrow \ell\nu ee$	50-150	303030	6.8373×10^{-01}	200
$VV \rightarrow \ell\nu ee$	150-500	303031	8.6295×10^{-02}	50
$VV \rightarrow \ell\nu ee$	500-1000	303032	3.4398×10^{-03}	50
$VV \rightarrow \ell\nu ee$	1000-2000	303033	2.7244×10^{-04}	50
$VV \rightarrow \ell\nu ee$	2000-3000	303034	7.6162×10^{-06}	50
$VV \rightarrow \ell\nu ee$	3000-4000	303035	4.0674×10^{-07}	50
$VV \rightarrow \ell\nu ee$	4000-5000	303036	2.3892×10^{-08}	50
$VV \rightarrow \ell\nu ee$	5000+	303037	1.8118×10^{-09}	50
$VV \rightarrow q\nu ee$	50-150	303038	$5.7027 \times 10^{+00}$	200
$VV \rightarrow q\nu ee$	150-500	303039	1.7567×10^{-01}	50
$VV \rightarrow q\nu ee$	500-1000	303040	4.0750×10^{-03}	50
$VV \rightarrow q\nu ee$	1000-2000	303041	3.1586×10^{-04}	50
$VV \rightarrow q\nu ee$	2000-3000	303042	1.0039×10^{-05}	50
$VV \rightarrow q\nu ee$	3000-4000	303043	6.6015×10^{-07}	50
$VV \rightarrow q\nu ee$	4000-5000	303044	5.3459×10^{-08}	50
$VV \rightarrow q\nu ee$	5000+	303045	4.8468×10^{-09}	50

Table A.5: The NLO SHERPA WW , WZ , and ZZ Diboson MC samples used in the modeling of the electron channel in the dilepton analysis [4].

to 32862 pb^{-1} in February 2017 for the Moriond conference, according to the ATLAS luminosity group recommendations [172]. As a result, the total integrated luminosity of the 2015 + 2016 dataset combination becomes $3213 \text{ pb}^{-1} + 32862 \text{ pb}^{-1} = 36075 \text{ pb}^{-1}$, or 36.1 fb^{-1} as quoted in each of the kinematic plots of the analysis chapter. The resulting luminosity uncertainty was updated as well, with an estimate of 3.2% for the 2015+2016 combination.

Period	Runs	$\mathcal{L}_{\text{int}} [\text{pb}^{-1}]$	Year
D	276262 - 276954	51.646	2015
E	278880 - 279928	435.61	2015
F	279932 - 280368	301.69	2015
G	280423 - 281075	724.40	2015
H	281317 - 281411	264.36	2015
J	282625 - 284484	1435.2	2015
Total	276262 - 284484	3213.0	2015
A	297730 - 300279	546.57	2016
B	300345 - 300908	1949.5	2016
C	301912 - 302393	2884.2	2016
D	302737 - 303560	4684.1	2016
E	303638 - 303892	1507.1	2016
F	303943 - 304494	3449.0	2016
G	305380 - 306451	3886.4	2016
I	307126 - 309759	8124.4	2016
L	310015 - 311481	6325.7	2016
Total	297730 - 311481	33257	2016
Total	276262 - 311481	36470	2015+2016

Table A.6: LHC Datasets used in the dilepton analysis [4].

Appendix B

Additional Dielectron Cutflow Tables

The following section presents the electron channel cutflow tables for the Drell-Yan, Top, and Diboson background processes according to the analysis event selection criteria of Section 10.5.1. Table B.1 presents the Drell-Yan cutflow, Table B.2 presents the combined Top ($t\bar{t} + Wt$) cutflow, and Table B.3 presents the combined Diboson ($WW + WZ + ZZ$) cutflow. Each table is generated using the standard set of simulation samples of Appendix A.

Selection Cut	Number Events	Relative Eff. [%]	Absolute Eff. [%]
Post-selection	27 792 145.29	100.00	100.00
GRL	27 792 145.29	100.00	100.00
Trigger	23 084 040.92	83.06	83.06
Event Cleaning	23 084 040.92	100.00	83.06
2 Electrons	23 083 701.84	100.00	83.06
η	22 378 782.21	96.95	80.52
Object Quality	22 325 785.14	99.76	80.33
E_T	15 435 590.20	69.14	55.54
d_0 Sig.	15 316 004.67	99.23	55.11
z_0	15 166 910.86	99.03	54.57
Identification	14 107 638.84	93.02	50.76
Isolation	13 586 918.28	96.31	48.89
Invariant Mass	13 133 770.28	96.66	47.26
Event Weights	12 093 108.99	92.08	43.51

Table B.1: The number of simulated Drell-Yan events passing the electron channel event selection of the dilepton analysis.

Selection Cut	Number Events	Relative Eff. [%]	Absolute Eff. [%]
Post-selection	3 432 740.84	100.00	100.00
GRL	3 432 740.84	100.00	100.00
Trigger	237 117.65	6.91	6.91
Event	237 117.65	100.00	6.91
2 Electrons	237 107.39	100.00	6.91
η	236 688.49	99.82	6.90
Object Quality	236 650.55	99.98	6.89
E_T	159 981.32	67.60	4.66
d_0 Sig.	150 297.89	93.95	4.38
z_0	149 239.60	99.30	4.35
Identification	120 975.13	81.06	3.52
Isolation	112 549.93	93.04	3.28
Invariant Mass	87 752.50	77.97	2.56
Event Weights	82 520.12	94.04	2.40

Table B.2: The number of simulated Top ($t\bar{t} + Wt$) events passing the electron channel event selection of the dilepton analysis.

Selection Cut	Number Events	Relative Eff. [%]	Absolute Eff. [%]
Post-selection	315 370.19	100.00	100.00
GRL	315 370.19	100.00	100.00
Trigger	98 956.35	31.38	31.38
Event Cleaning	98 956.35	100.00	31.38
2 Electrons	98 863.77	99.91	31.35
η	96 974.46	98.09	30.75
Object Quality	96 814.36	99.83	30.70
E_T	56 665.20	58.53	17.97
d_0 Sig.	55 965.28	98.76	17.75
z_0	55 287.77	98.79	17.53
Identification	49 027.21	88.68	15.55
Isolation	46 306.62	94.45	14.68
Invariant Mass	40 228.69	86.87	12.76
Event Weights	37 328.22	92.79	11.84

Table B.3: The number of simulated Diboson ($WW + WZ + ZZ$) events passing the electron channel event selection of the dilepton analysis.

Appendix C

ATLANTIS Event Displays

This section presents the event display for the highest invariant mass dielectron events selected in the 2015+2016 dilepton analysis, according to the 3.2 fb^{-1} and 36.1 fb^{-1} datasets that were used, respectively. The event displays here are produced with ATLANTIS [220], as part of the ATLAS software framework [156].

Figure C.1 depicts the highest-mass dilepton event in the 2015 search, and Figure C.2 depicts the same in the 2016 search. Energy depositions in the cells of the ECAL and HCAL are indicated by green and red towers, respectively. Reconstructed tracks associated with the selected electron candidates are indicated by red lines. Other tracks in the event with $E_T > 2 \text{ GeV}$, including those from pile-up, are indicated by cyan lines.

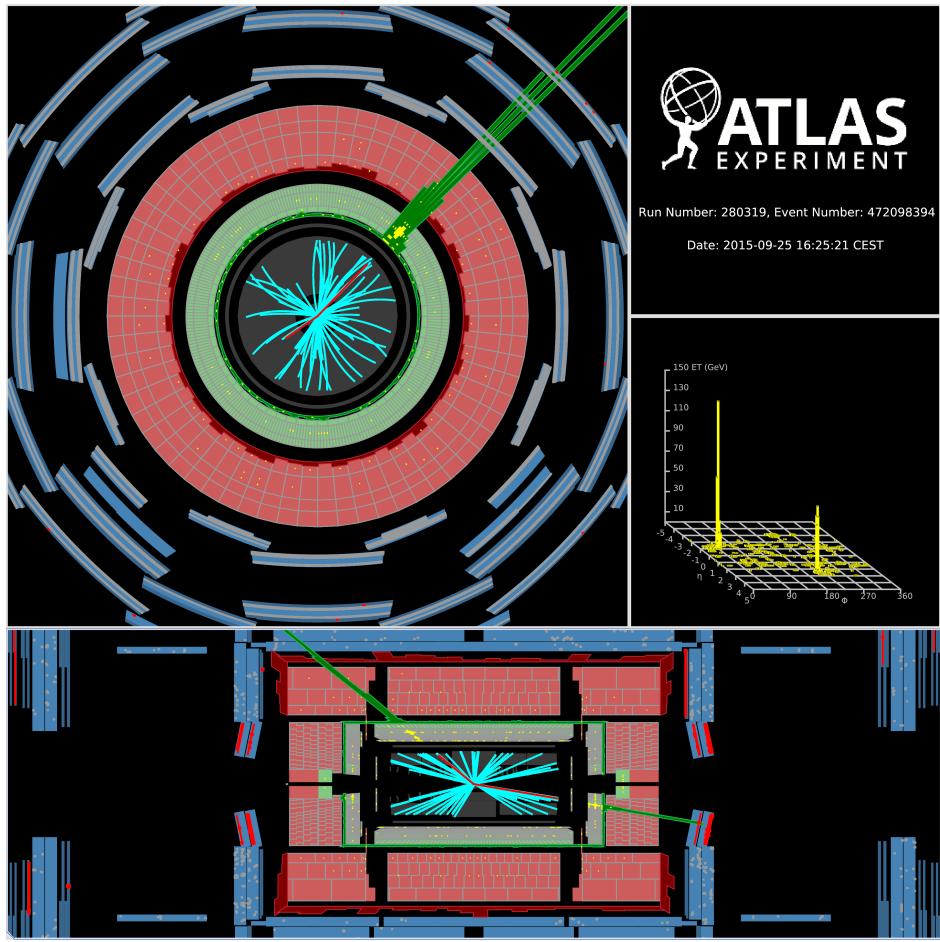


Figure C.1: The ATLANTIS event display for the highest dielectron invariant mass event selected in the 2015 analysis. The “leading” electron has an E_T of 373 GeV and an η of -1.03. The “subleading” electron has an E_T of 246 GeV and an η of 2.45. The invariant mass of the dielectron pair is 1.78 TeV [5].

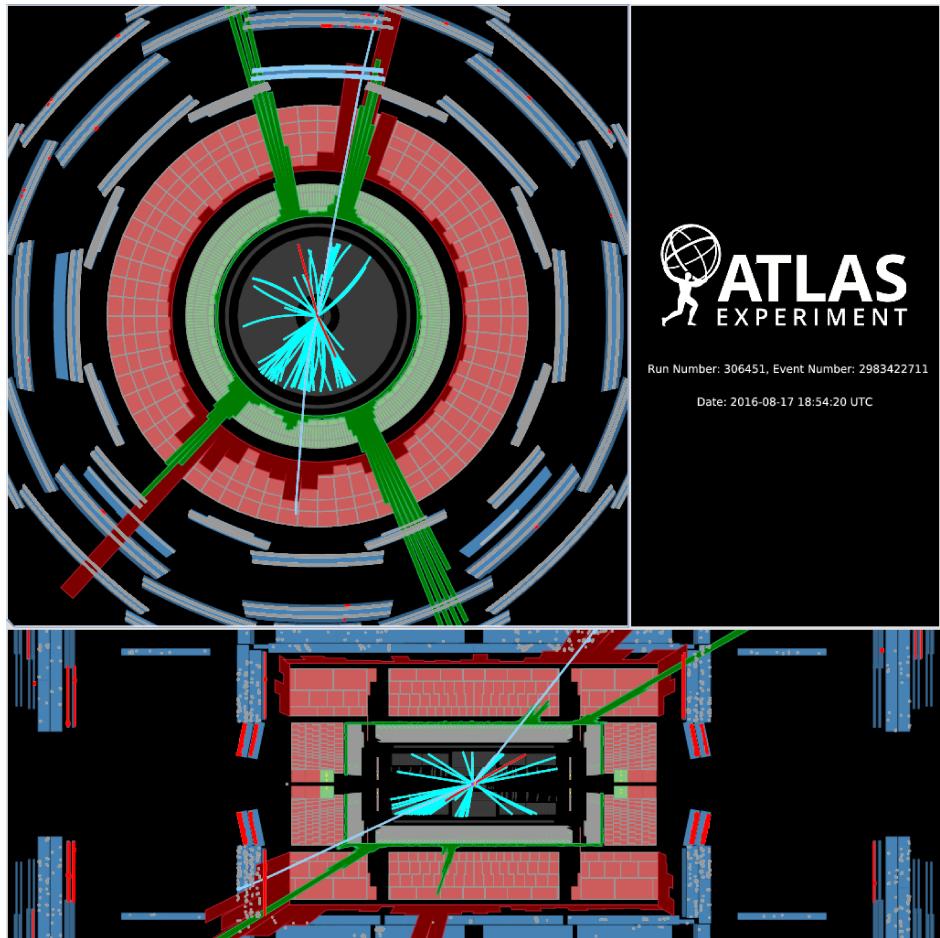


Figure C.2: The ATLANTIS event display for the highest dielectron invariant mass event selected in the 2016 analysis. The “leading” electron has an E_T of 841 GeV, an η of -1.26, and a ϕ of -1.15. The “subleading” electron has an E_T of 655 GeV, an η of 1.32, and a ϕ of 1.83. The invariant mass of the dielectron pair is 2.90 TeV [5].

Appendix D

ePump Inputs and Supplemental Results

3000 fb⁻¹ Templates

This section presents the complete set of finalized ePump templates used to model the triple-differential Drell-Yan cross section of Eq. 13.2 that were used to perform the update analysis of the final chapter of the dissertation. The cross section is presented in slices of invariant mass according to histogram binning of Section 13.3.4 and is scaled to 3000 fb⁻¹. Pseudo-data are generated according to the procedure of Section 13.3.5.

The specific kinematic region is labeled in each plot. The RESBos prediction is indicated by the solid black line and pseudo-data are shown as solid black markers. The sub-panel shows the ratio between the theoretical prediction and the resulting pseudo-data. Error bars represent the statistical uncertainty on the pseudo-data and the solid band represents the PDF uncertainty on the dilepton event yield.

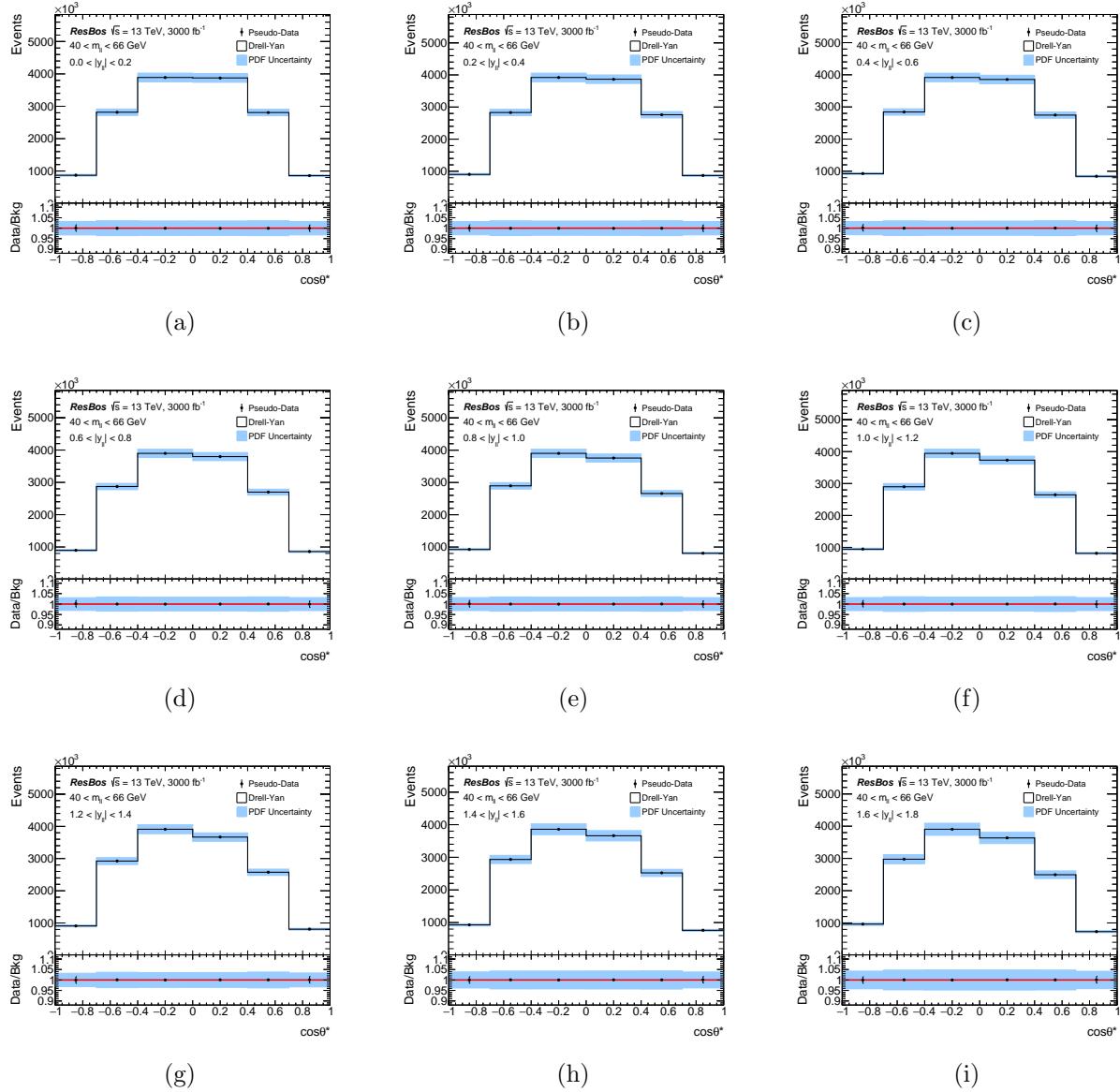


Figure D.1: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [40, 66]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

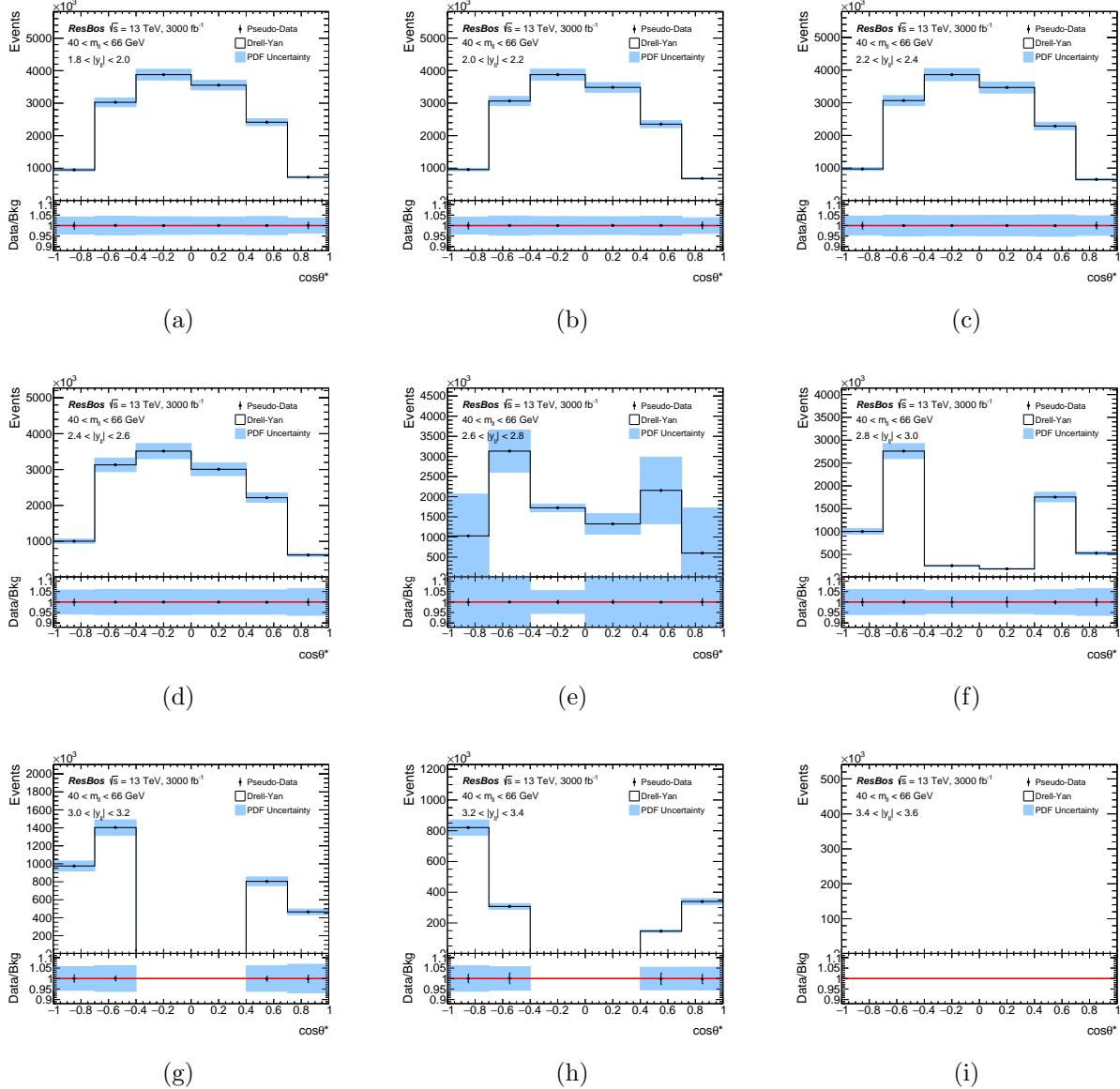


Figure D.2: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [40, 66] \text{ GeV}$ and $1.8 < |y_{\ell\ell}| < 3.6$.

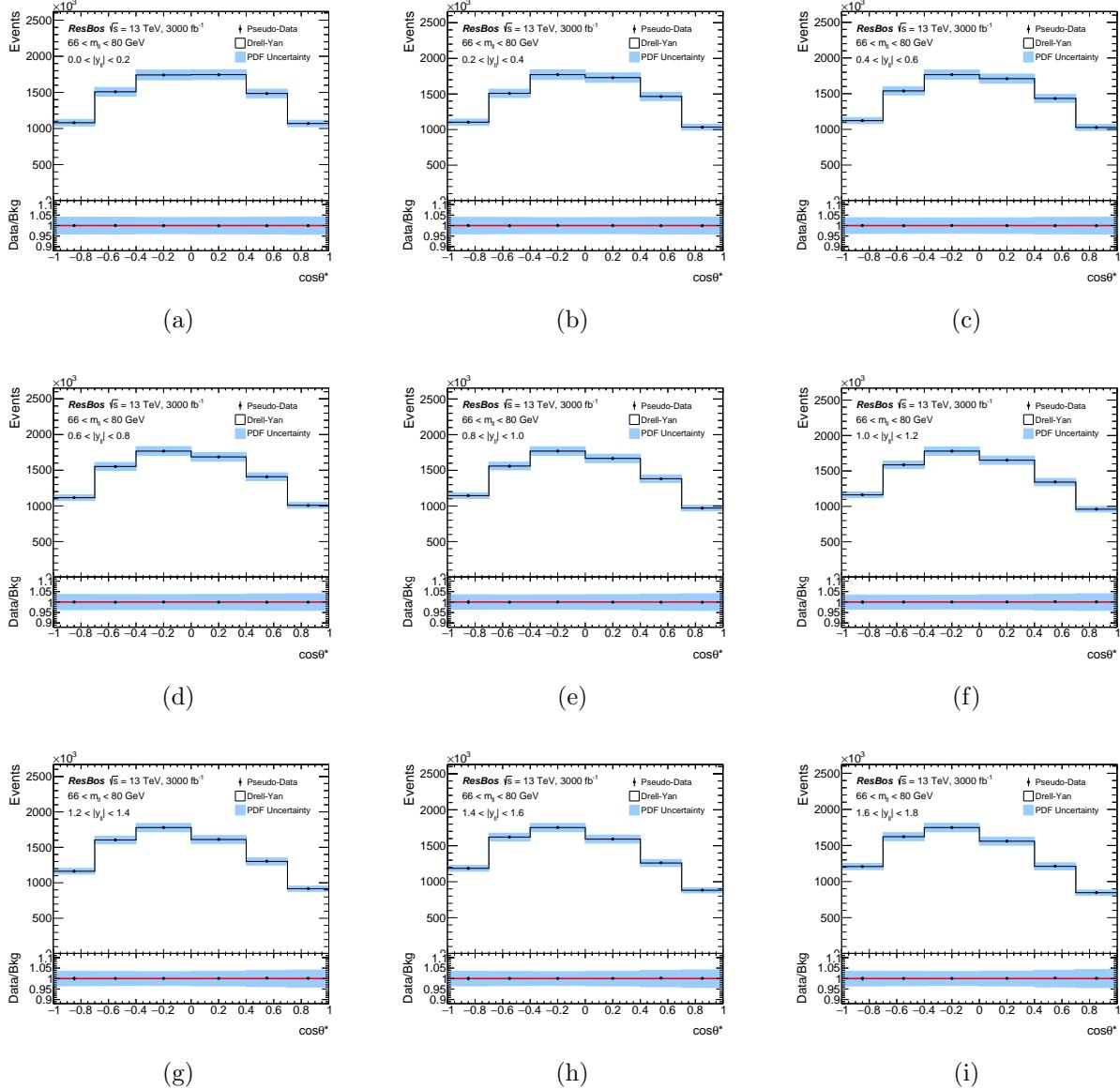


Figure D.3: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [66, 80]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

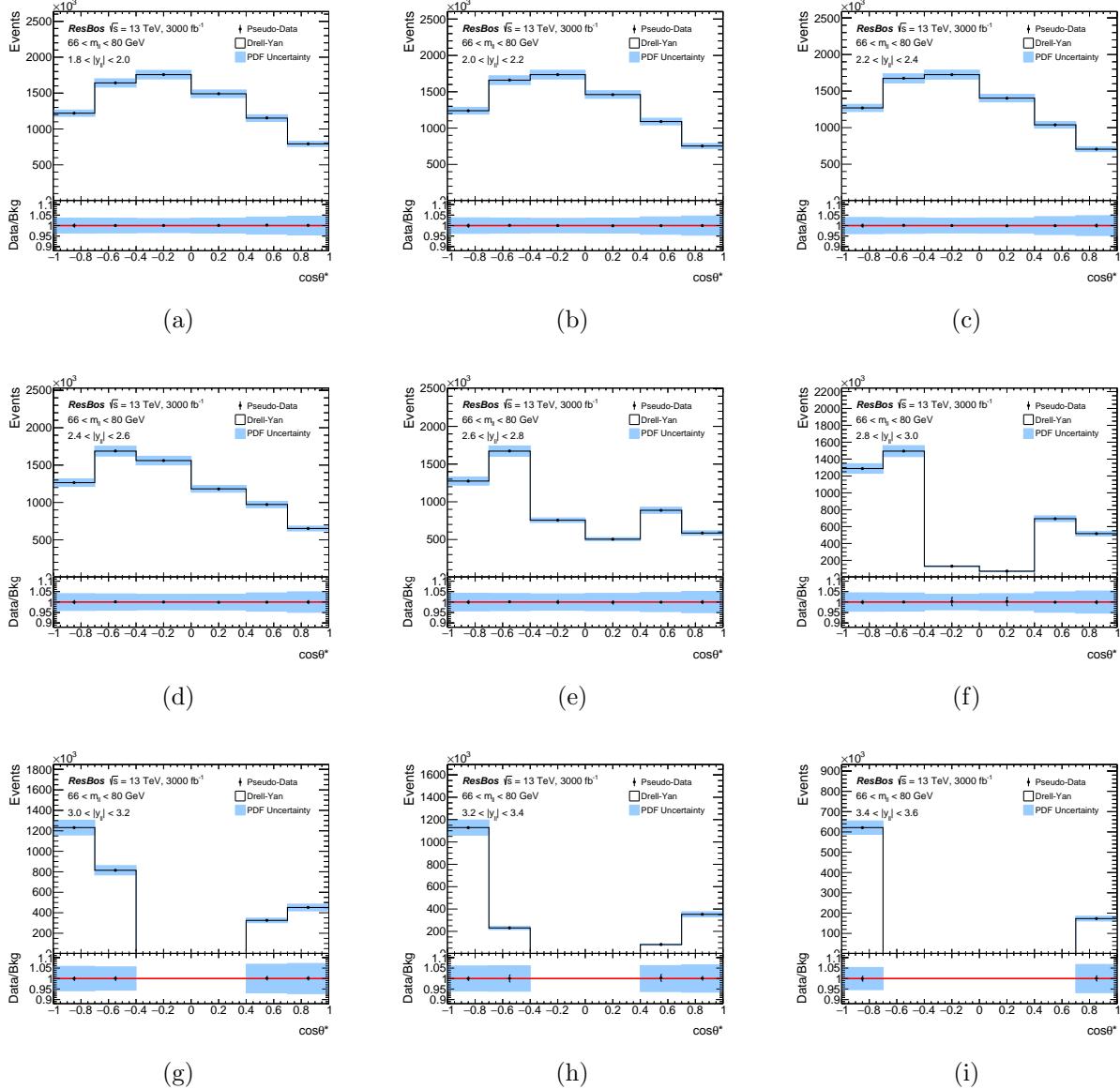


Figure D.4: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [66, 80]$ GeV and $1.8 < |y_{\ell\ell}| < 3.6$.

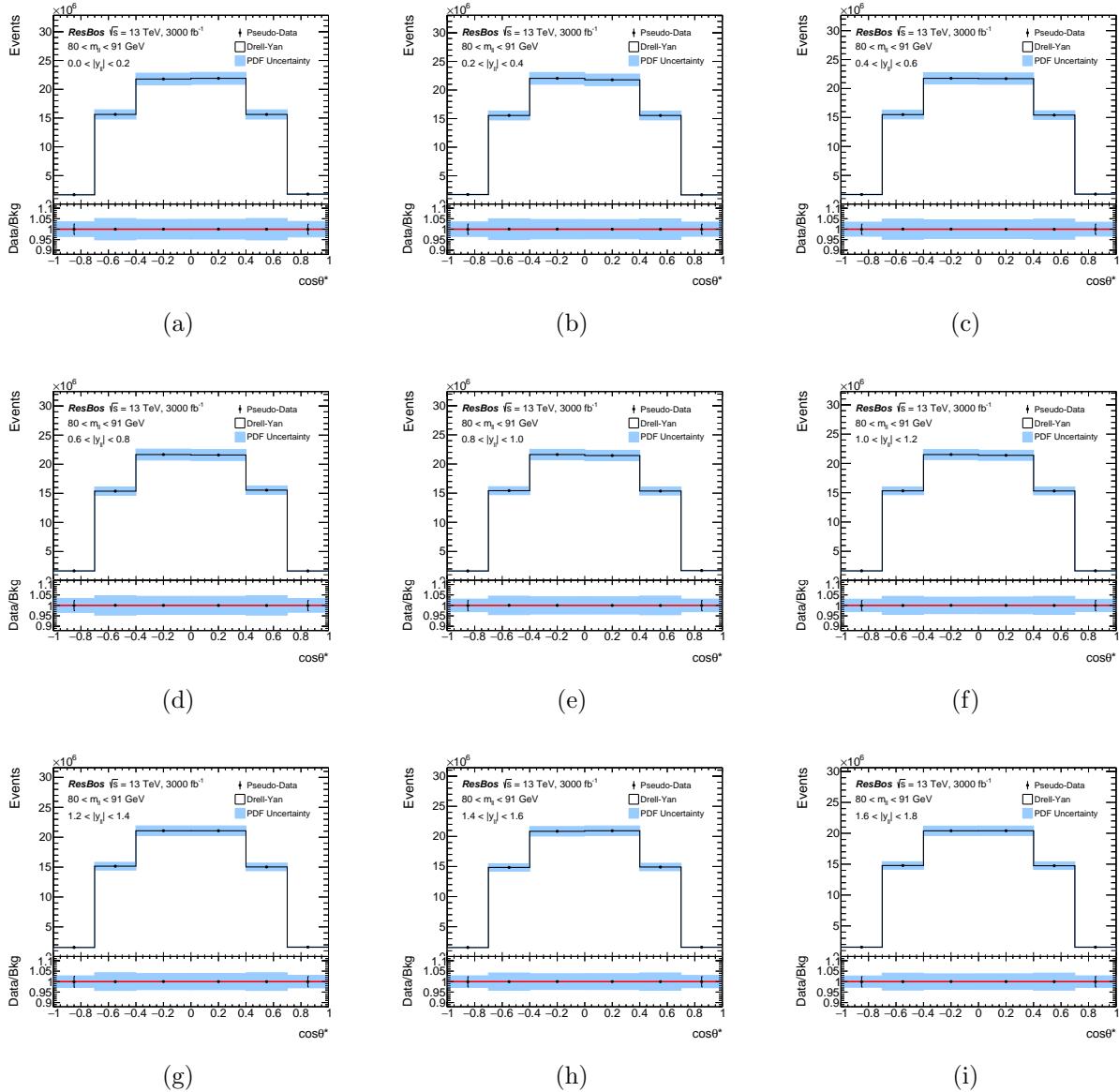


Figure D.5: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [80, 91]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

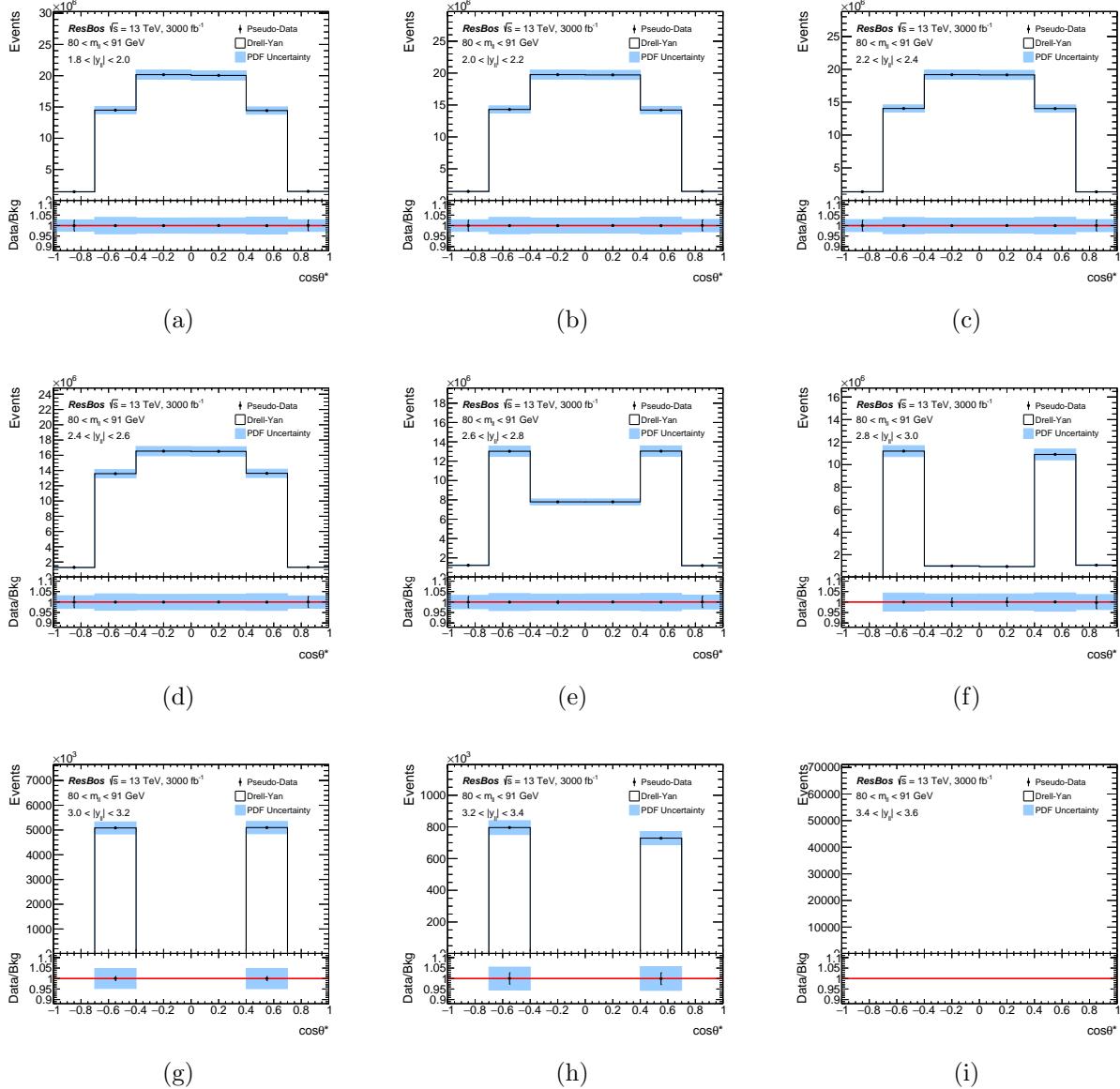


Figure D.6: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [80, 91] \text{ GeV}$ and $1.8 < |y_{\ell\ell}| < 3.6$.

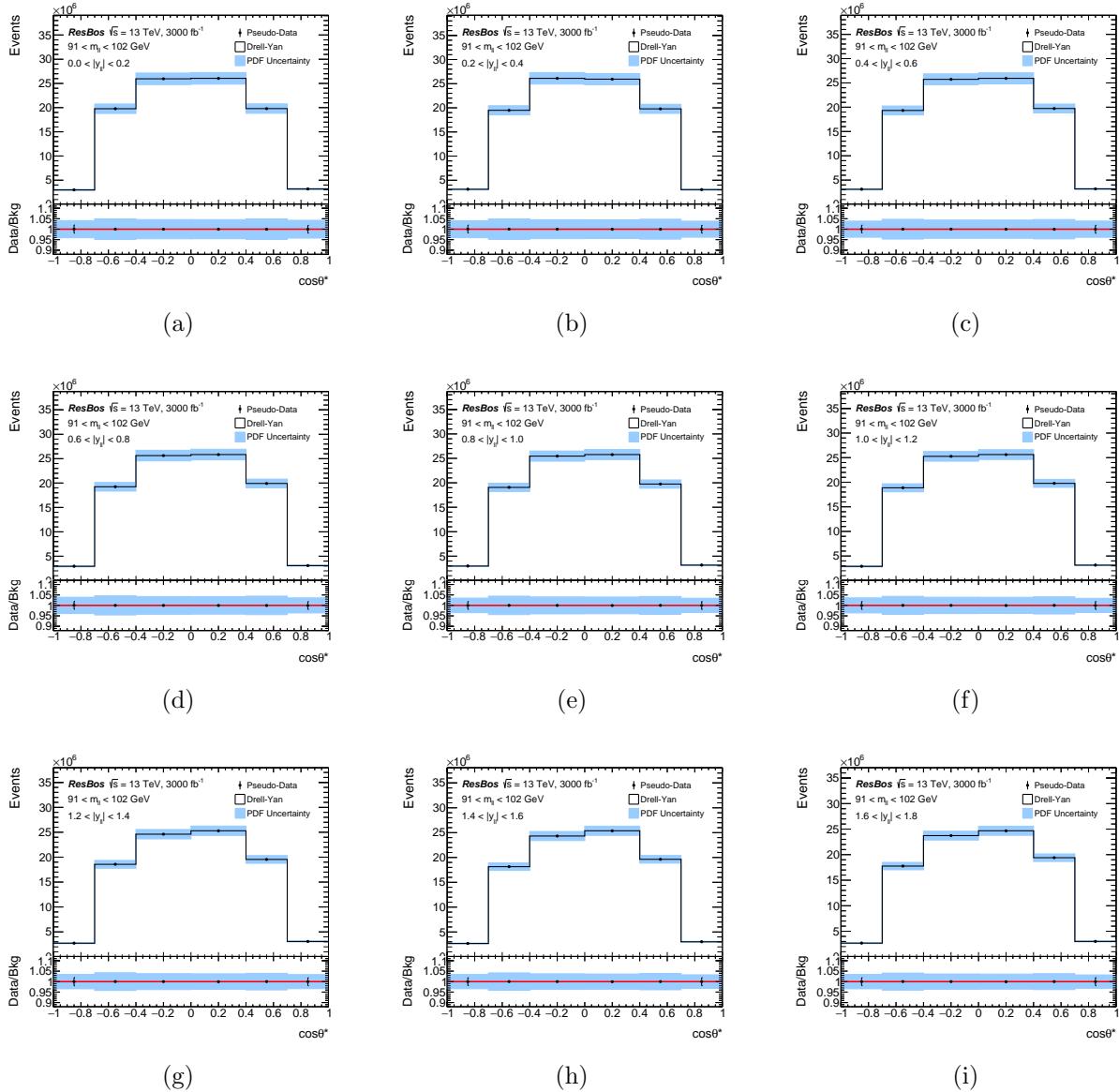


Figure D.7: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [91, 102]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

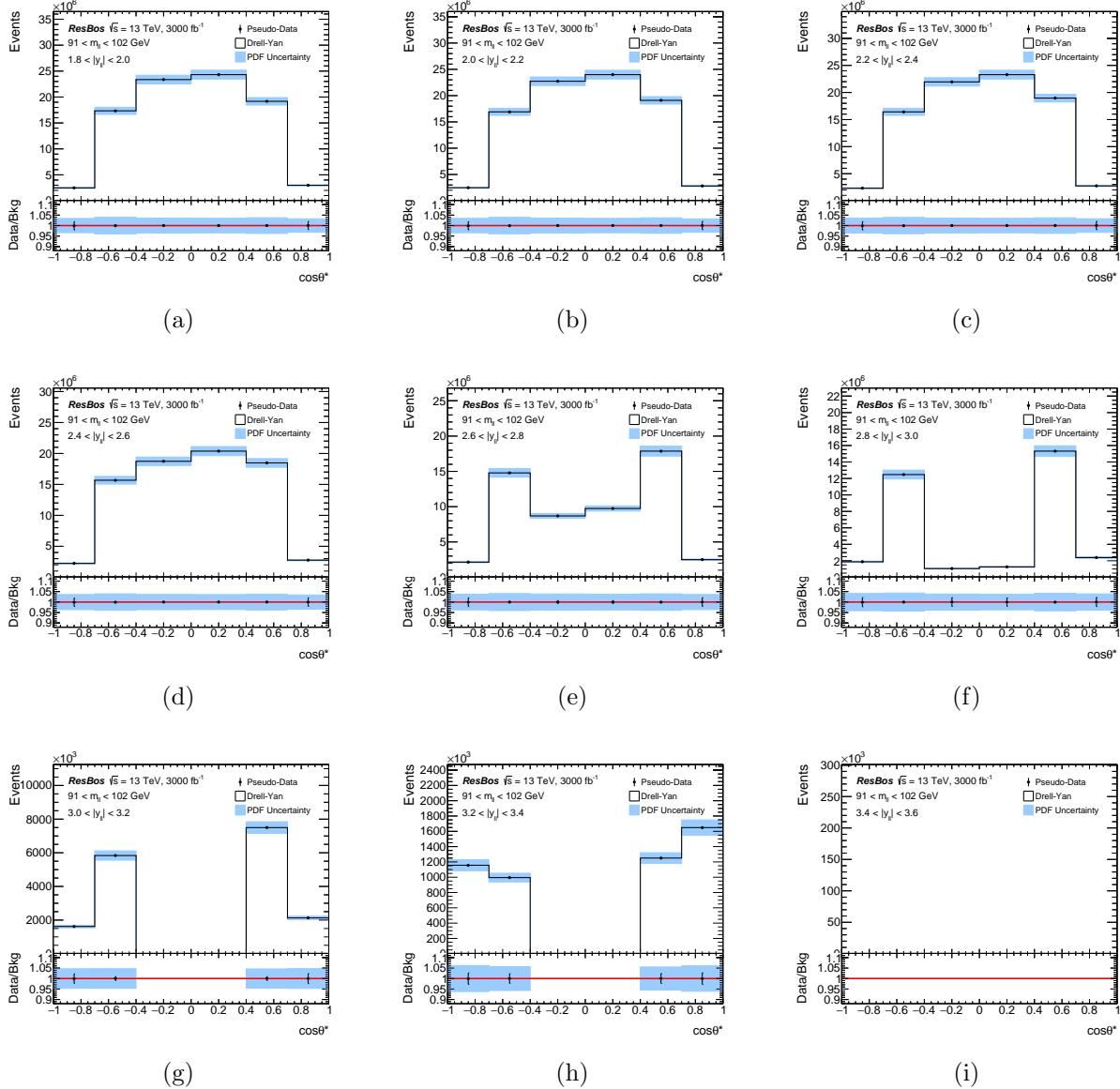


Figure D.8: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [91, 102] \text{ GeV}$ and $1.8 < |y_{\ell\ell}| < 3.6$.

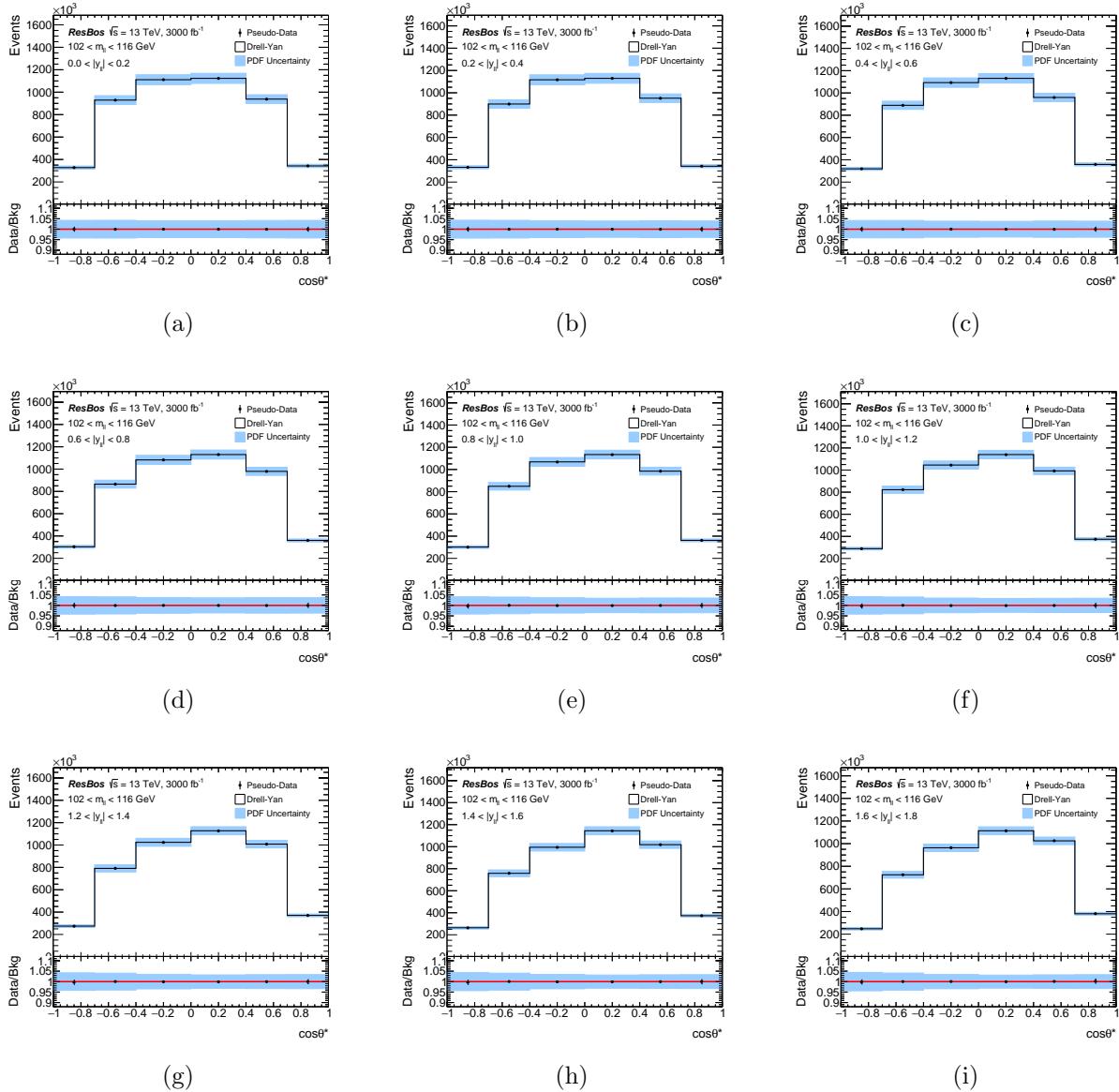


Figure D.9: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [102, 116]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

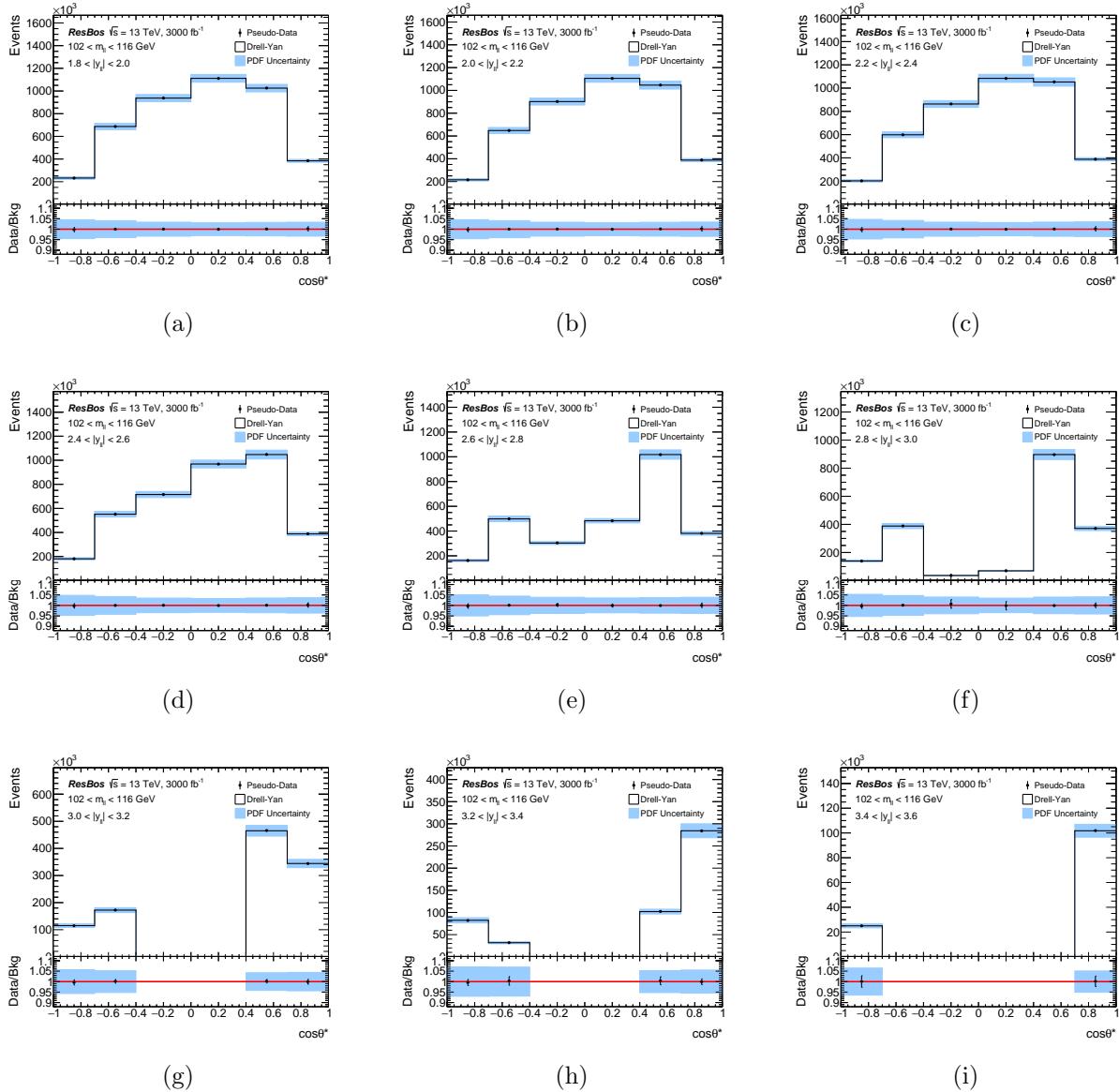


Figure D.10: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [102, 116]$ GeV and $1.8 < |y_{\ell\ell}| < 3.6$.

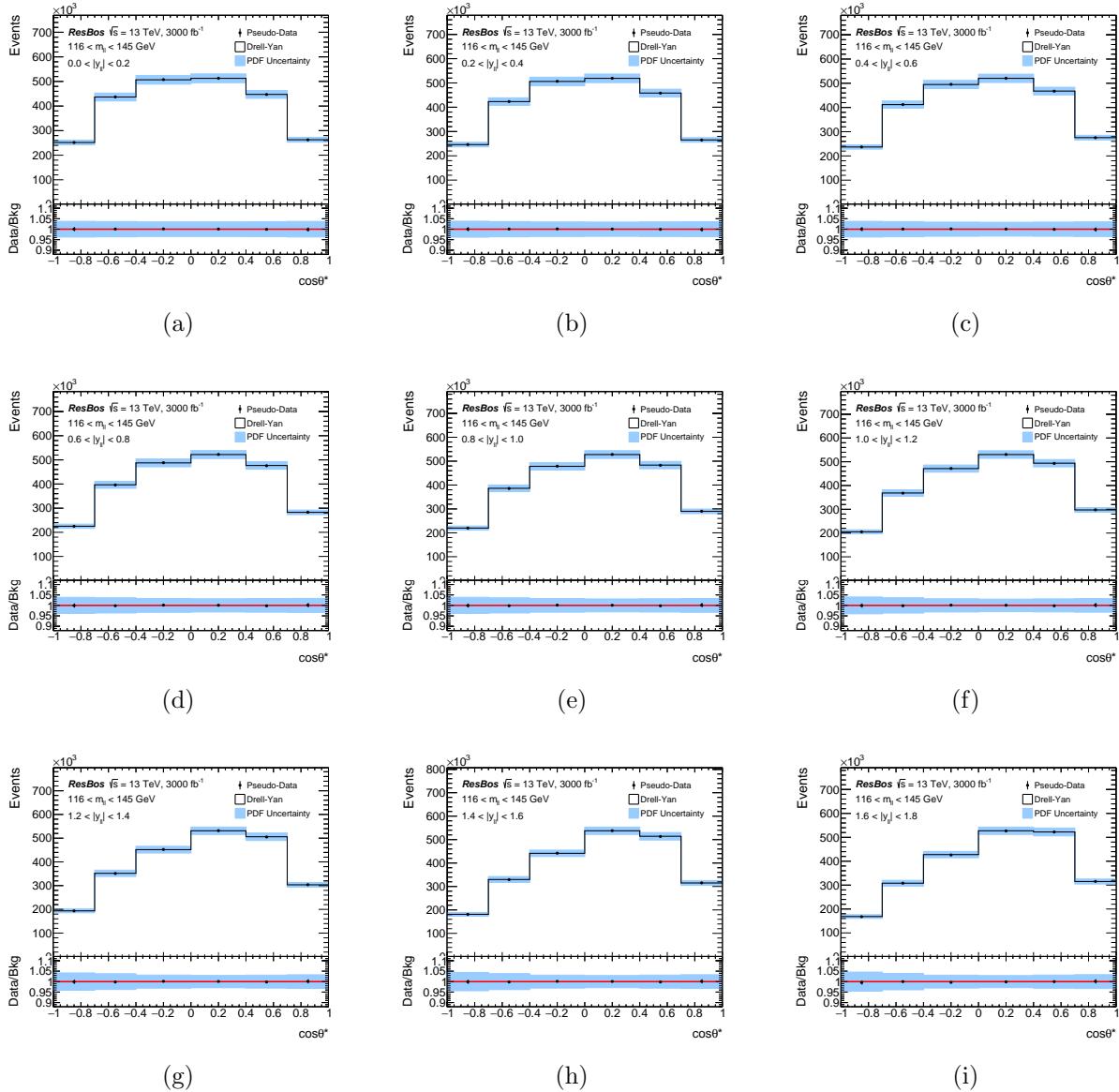


Figure D.11: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [116, 145]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

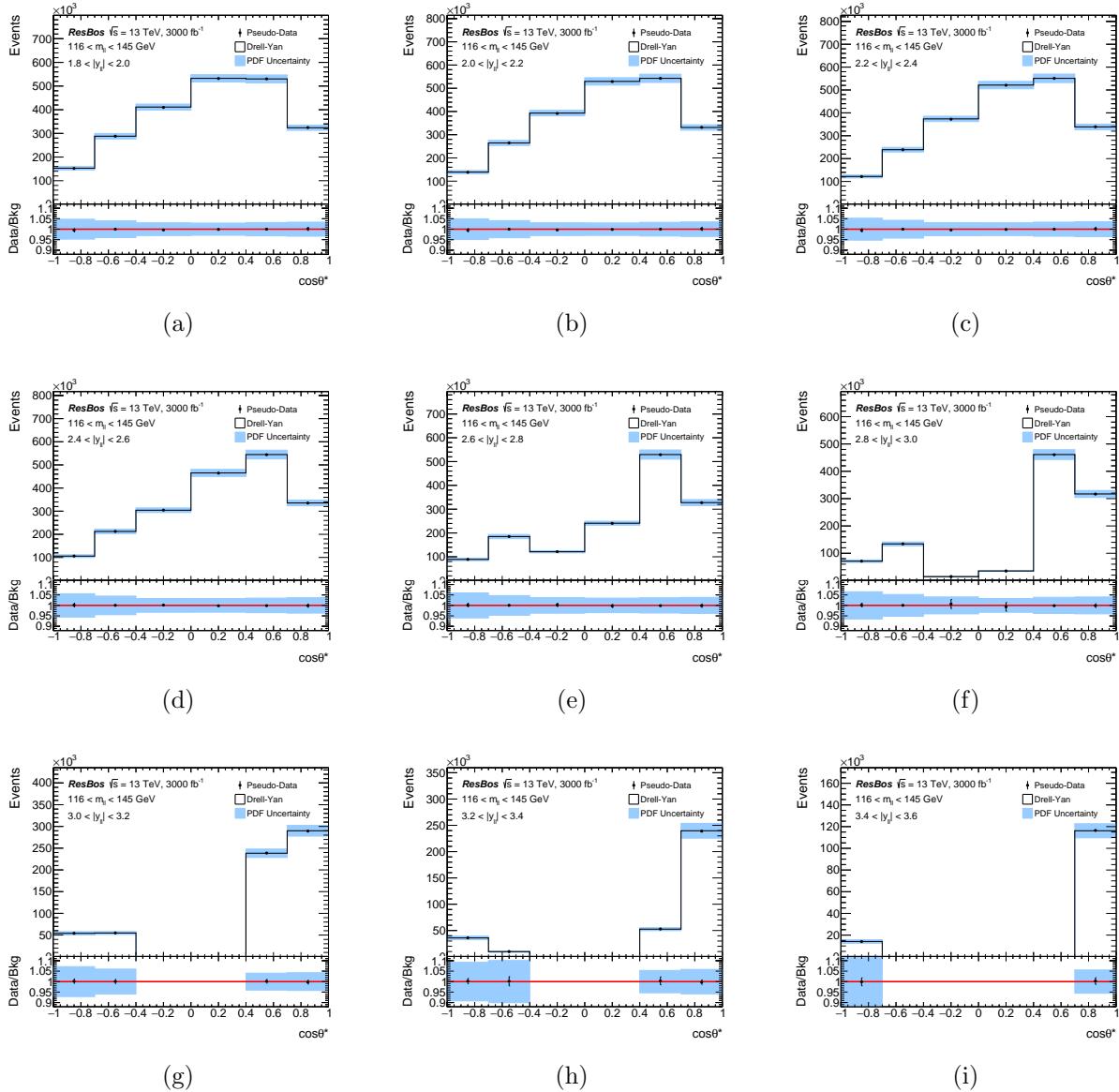


Figure D.12: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [116, 145] \text{ GeV}$ and $1.8 < |y_{\ell\ell}| < 3.6$.

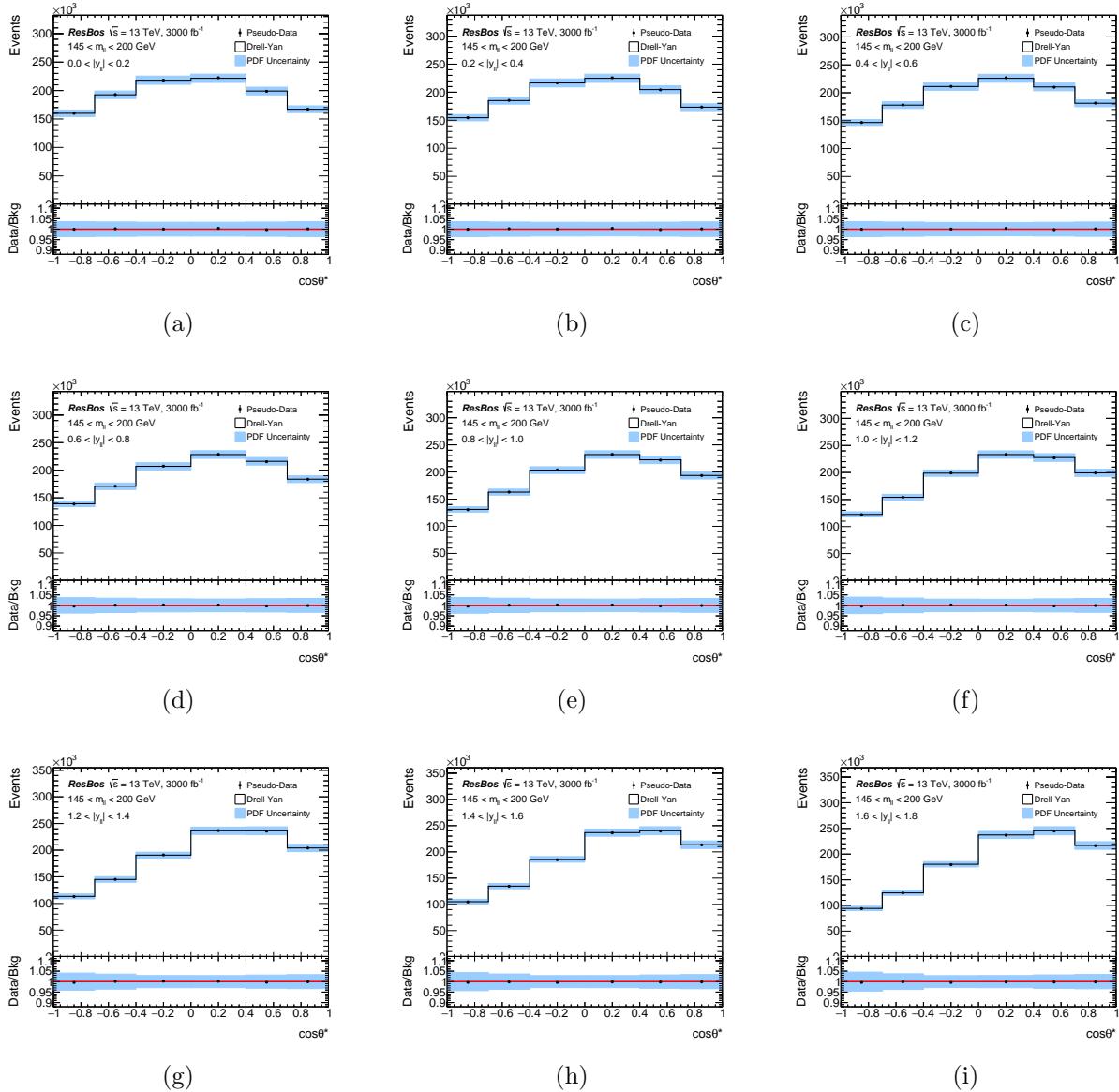


Figure D.13: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [145, 200]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

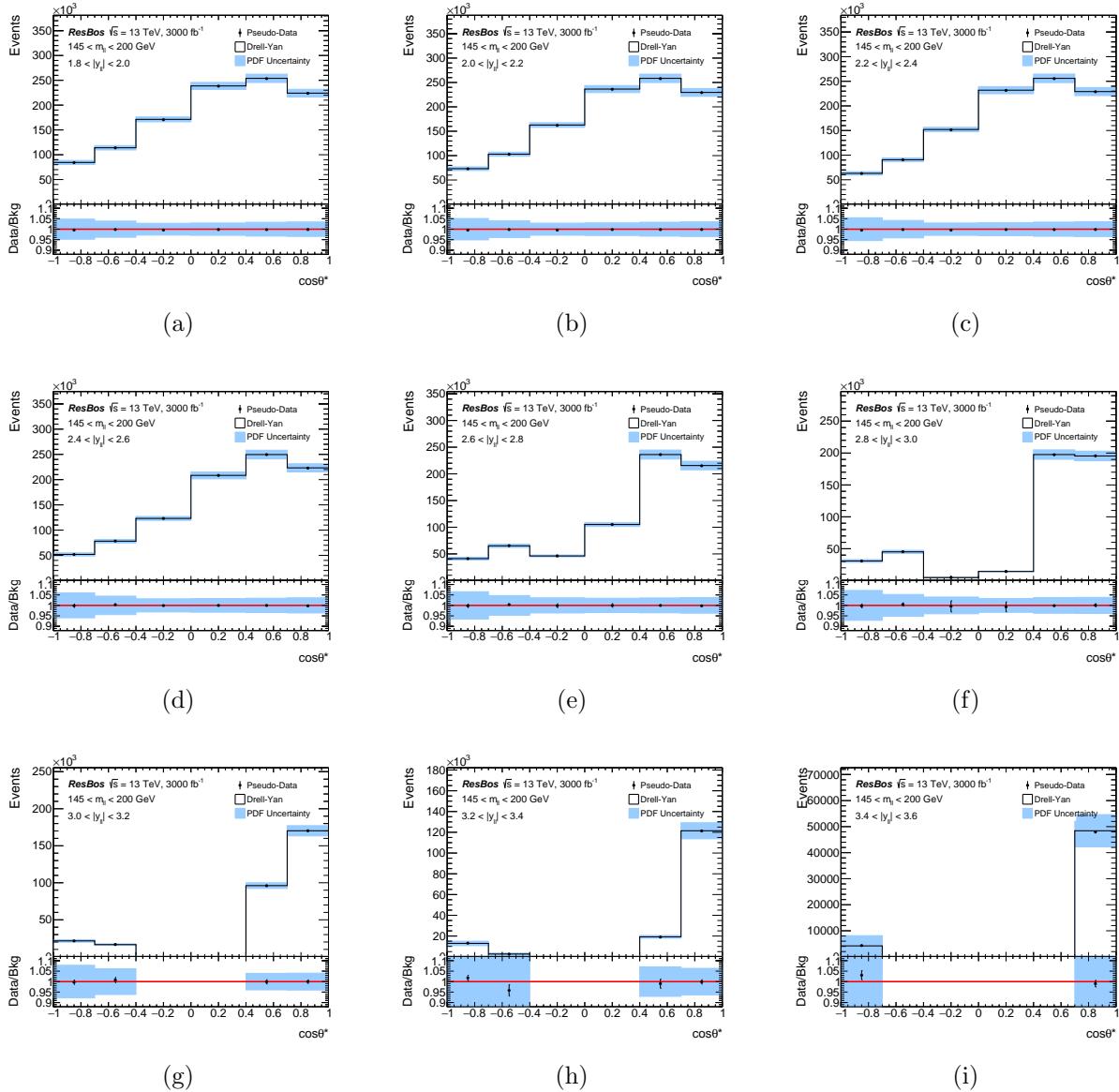


Figure D.14: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [145, 200] \text{ GeV}$ and $1.8 < |y_{\ell\ell}| < 3.6$.

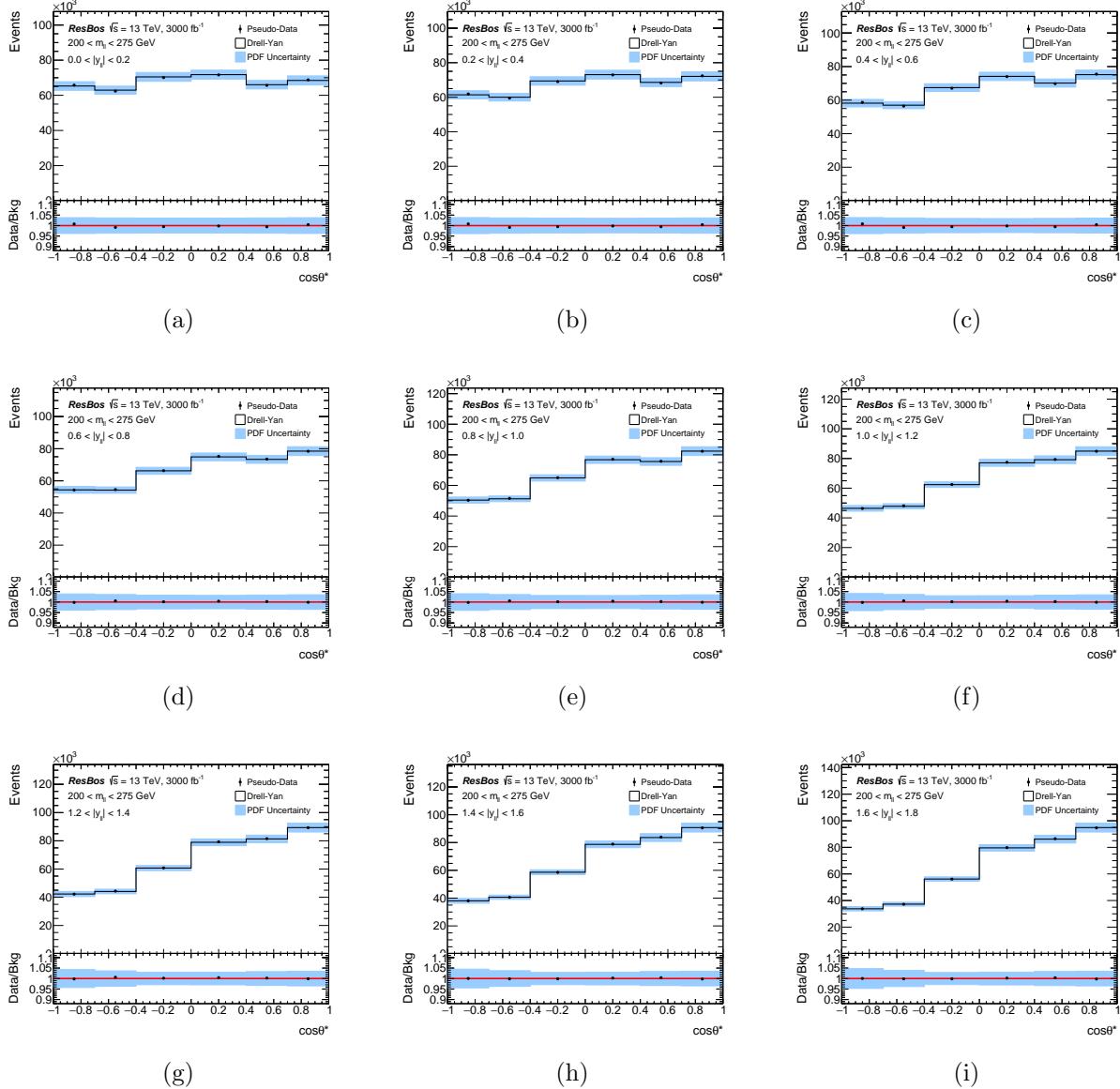


Figure D.15: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [200, 275]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

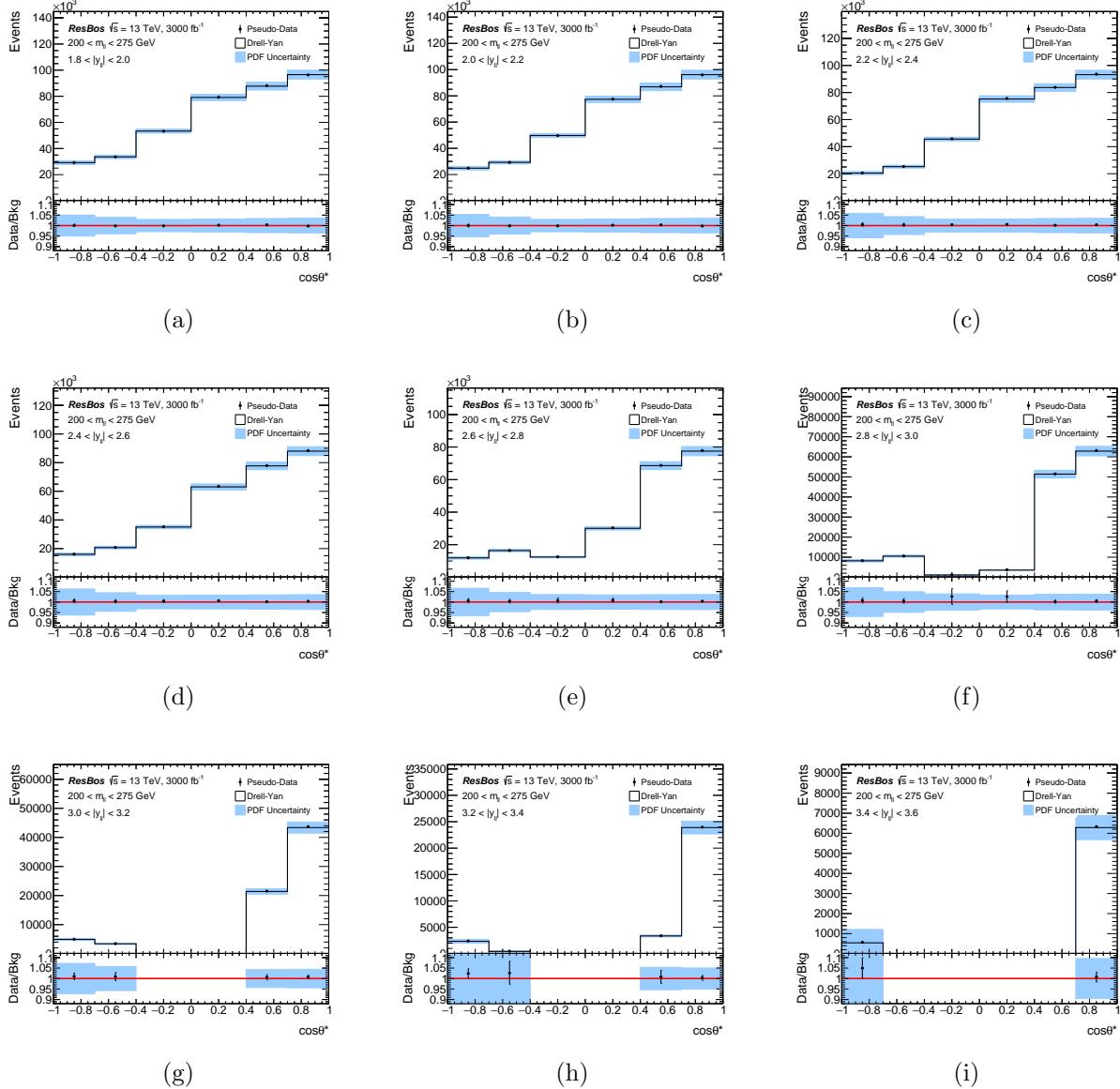


Figure D.16: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [200, 275]$ GeV and $1.8 < |y_{\ell\ell}| < 3.6$.

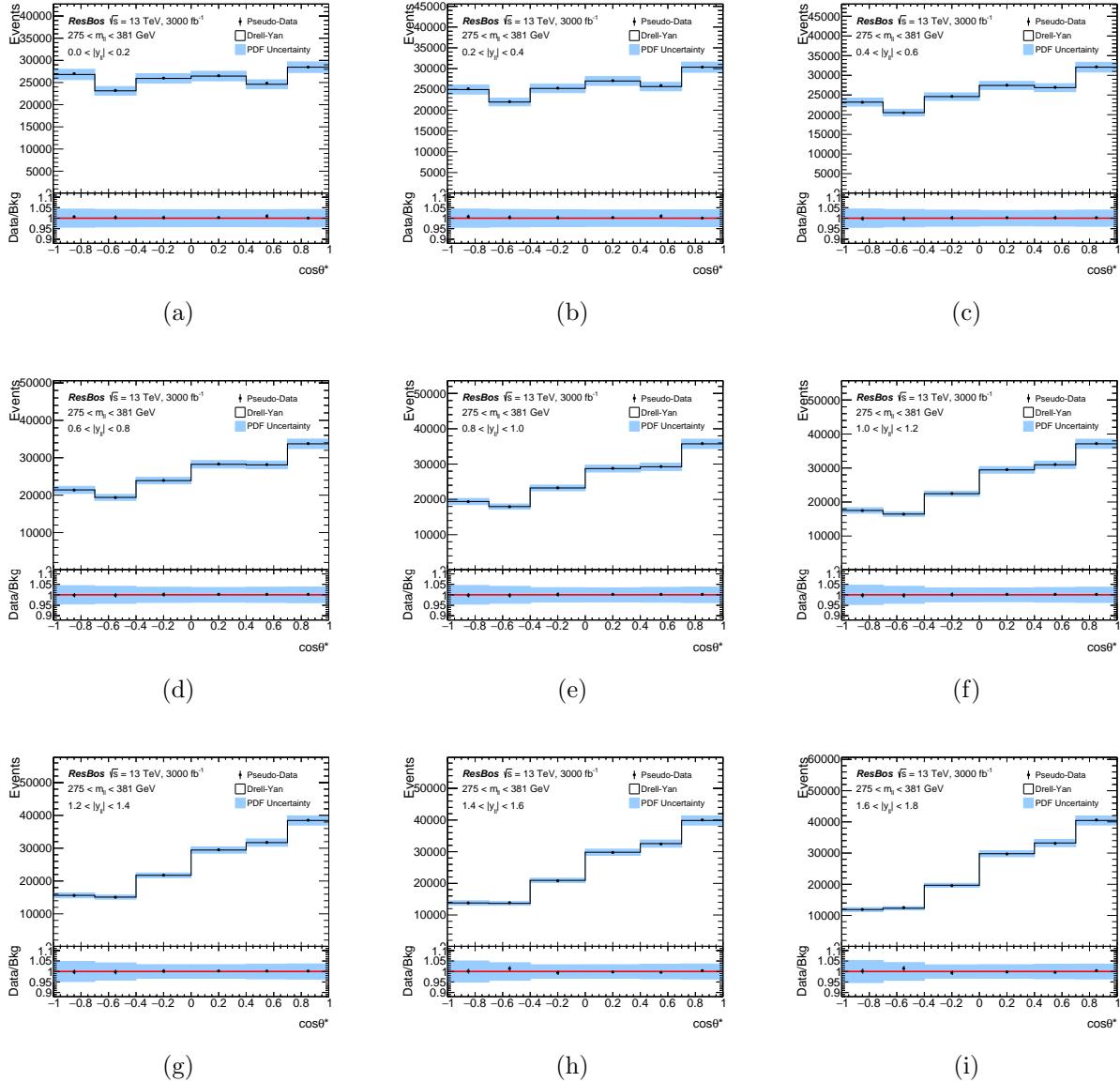


Figure D.17: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [275, 381]$ GeV and $0 < |y_{\ell\ell}| < 1.8$.

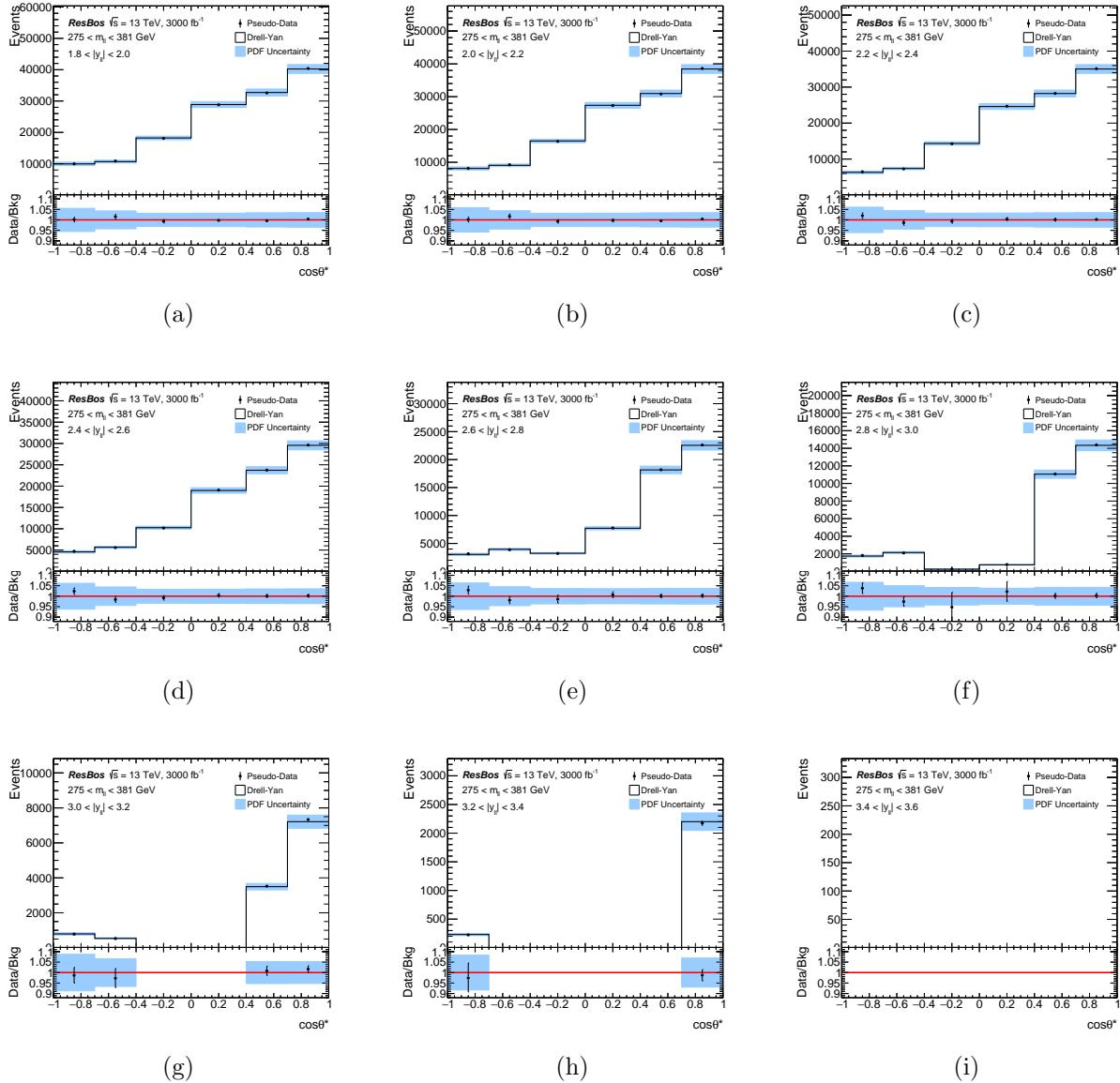


Figure D.18: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [275, 381] \text{ GeV}$ and $1.8 < |y_{\ell\ell}| < 3.6$.

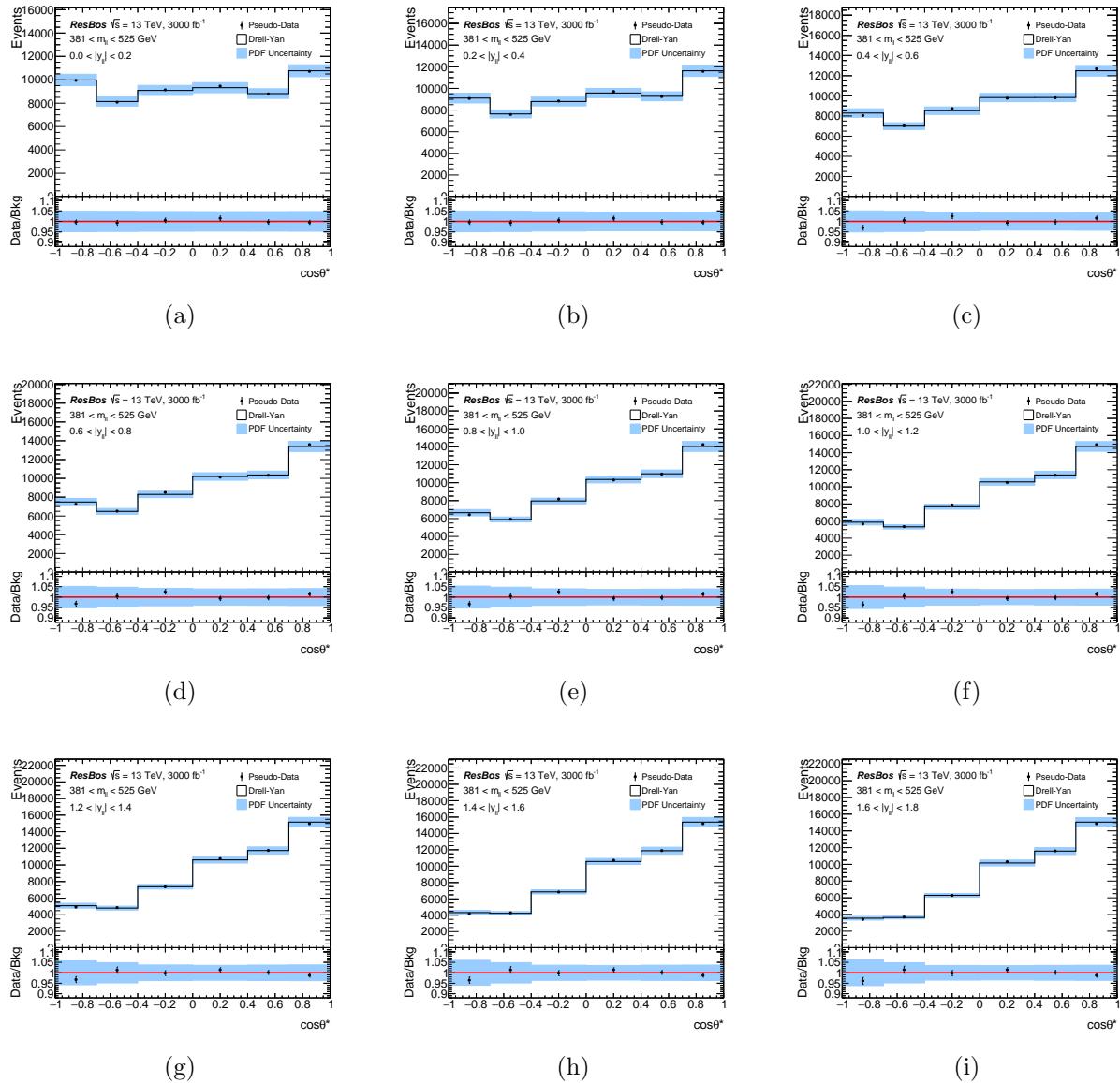


Figure D.19: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [381, 525] \text{ GeV}$ and $0 < |y_{\ell\ell}| < 1.8$.

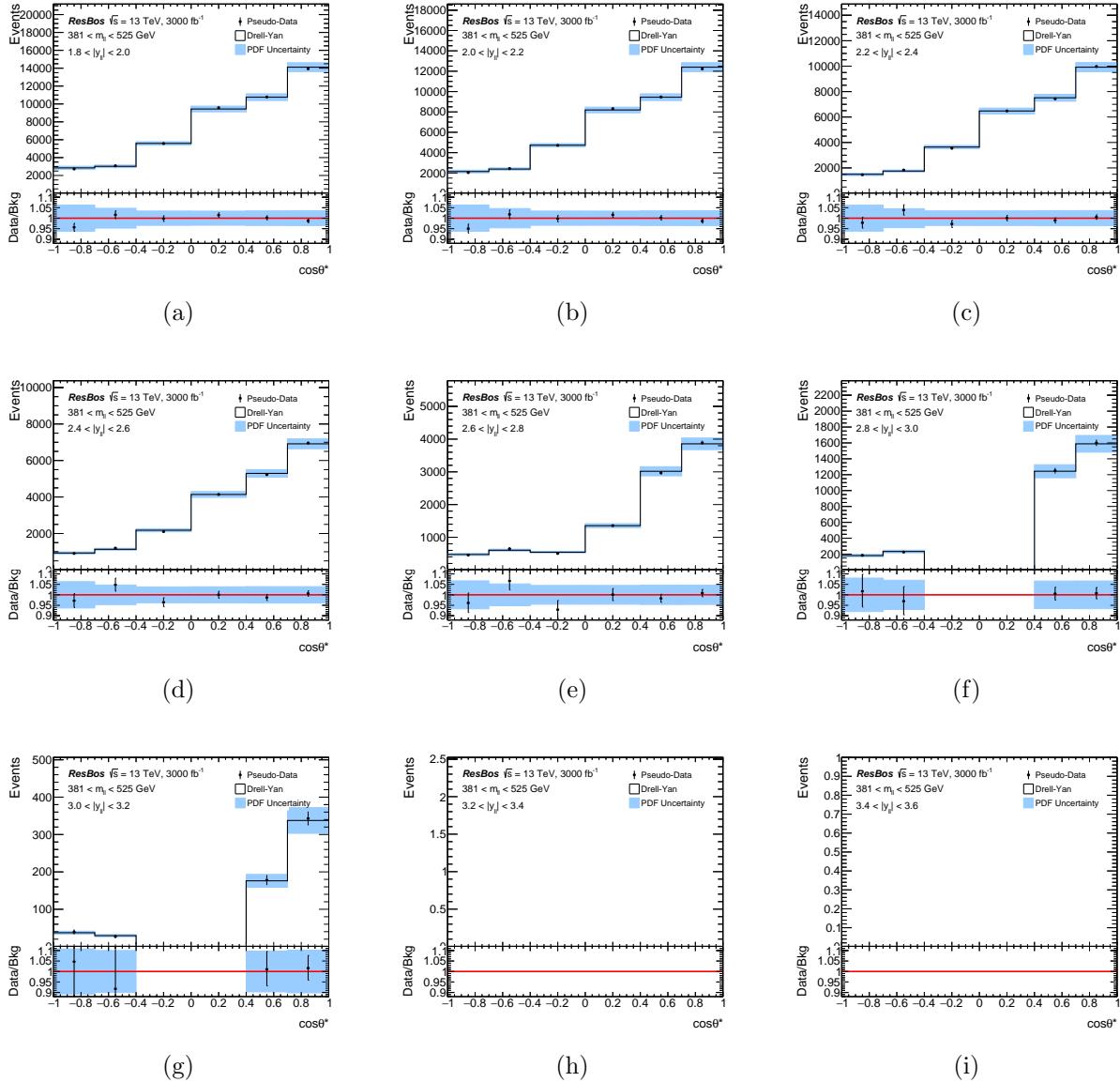


Figure D.20: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [381, 525]$ GeV and $1.8 < |y_{\ell\ell}| < 3.6$.

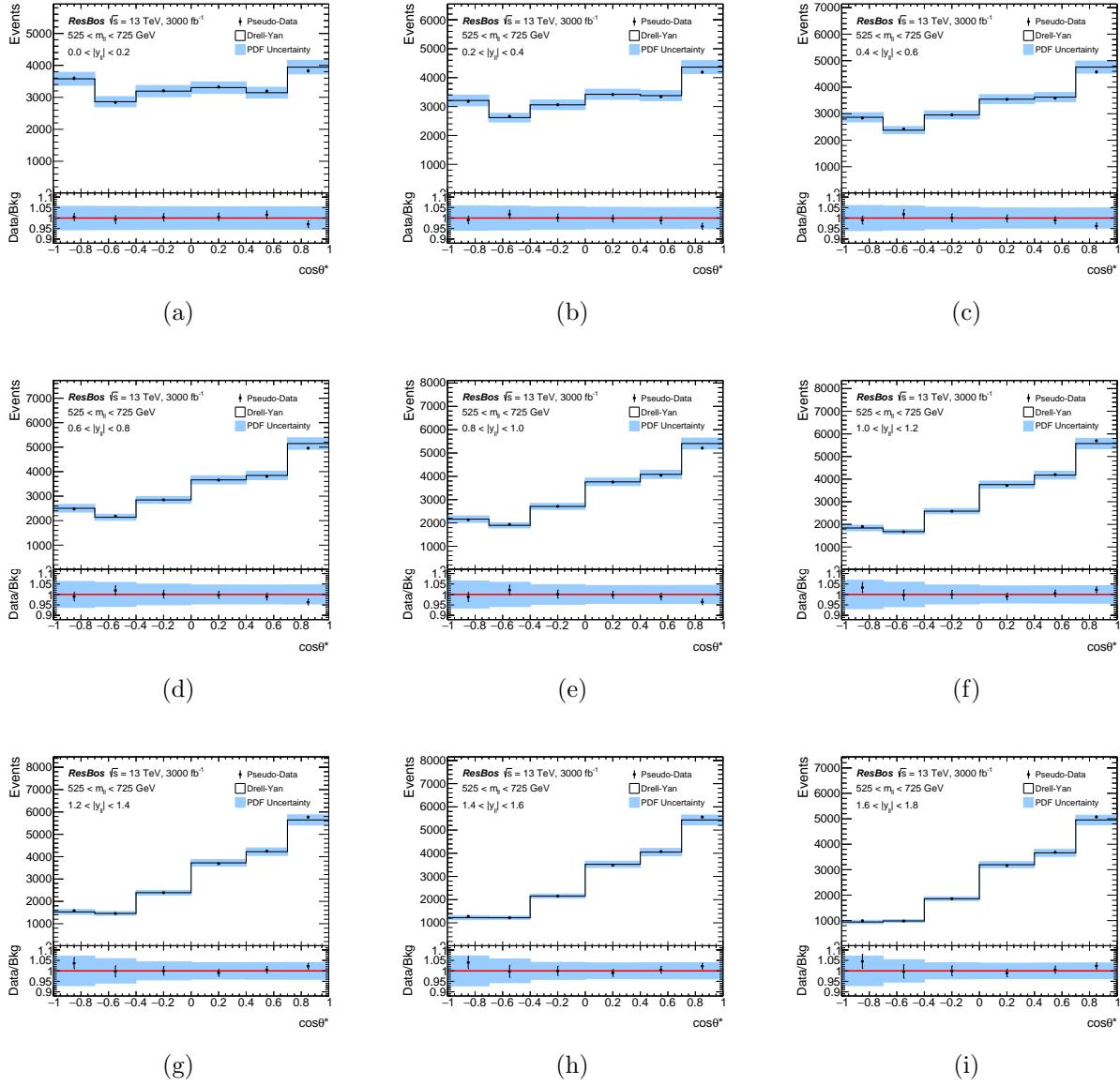


Figure D.21: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [525, 725] \text{ GeV}$ and $0 < |y_{\ell\ell}| < 1.8$.

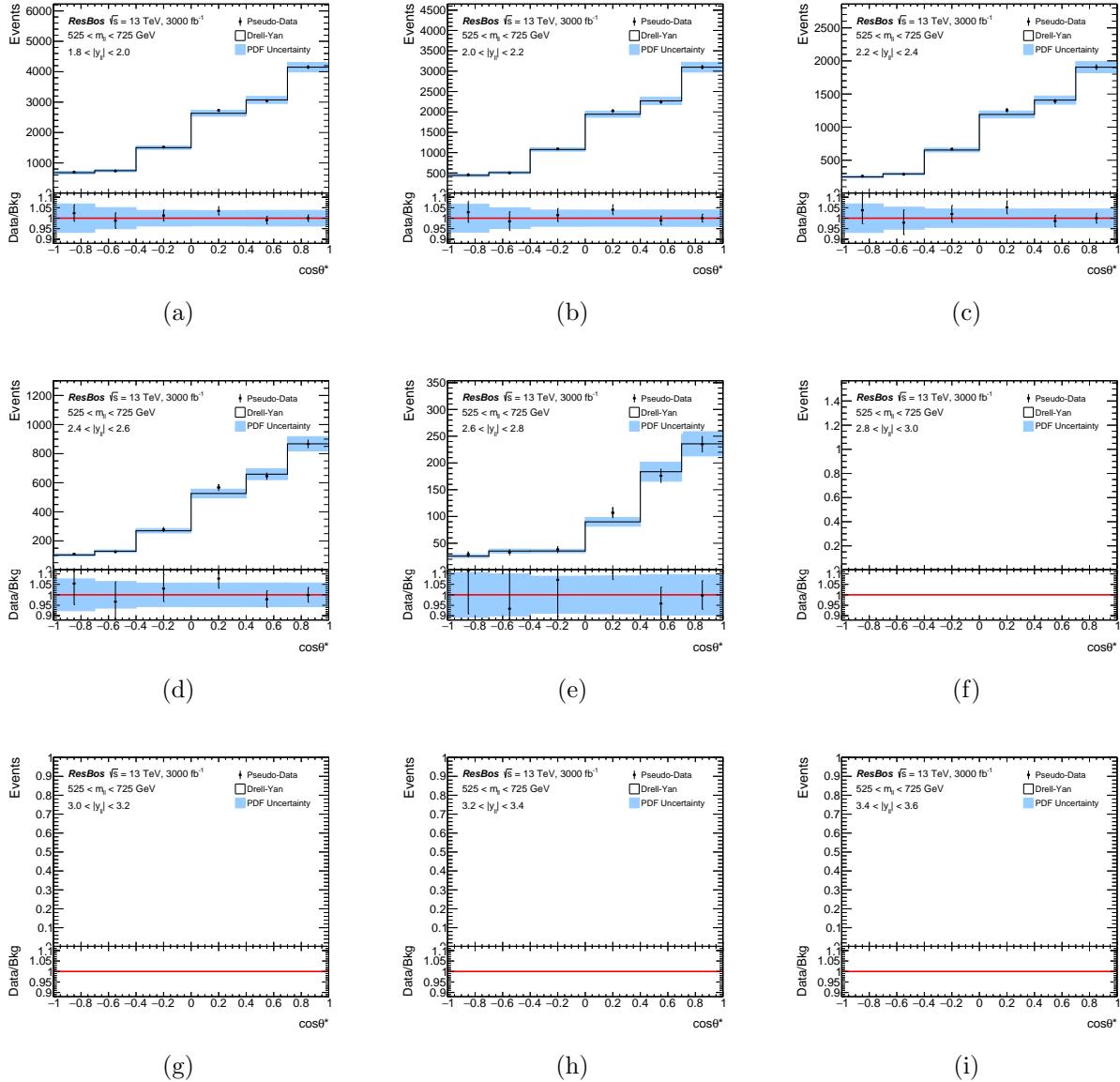


Figure D.22: $e\mu\mu$ templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{ll} \in [525, 725] \text{ GeV}$ and $1.8 < |y_{ll}| < 3.6$.

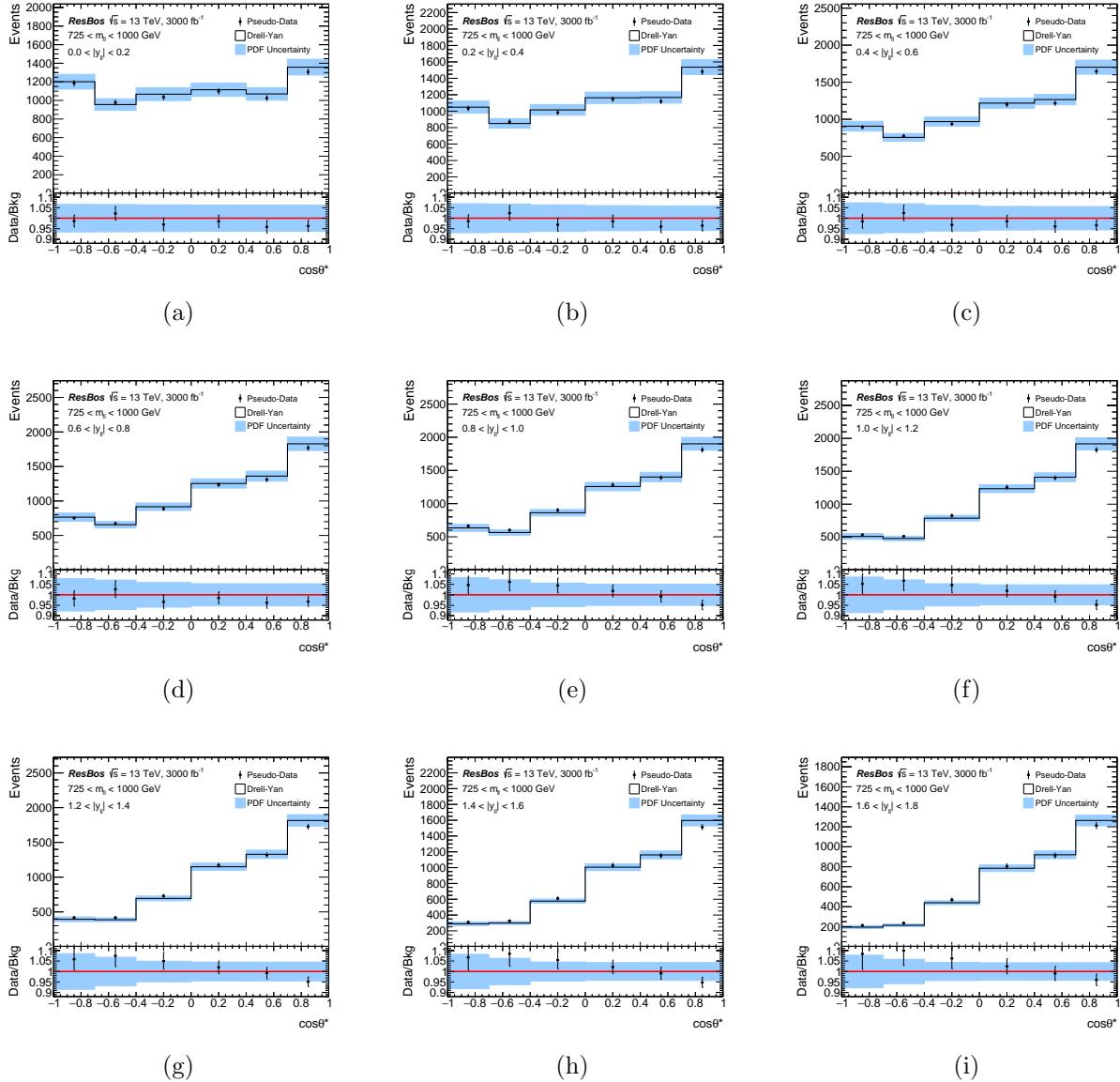


Figure D.23: $e\text{Pump}$ templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [725, 1000] \text{ GeV}$ and $0 < |y_{\ell\ell}| < 1.8$.

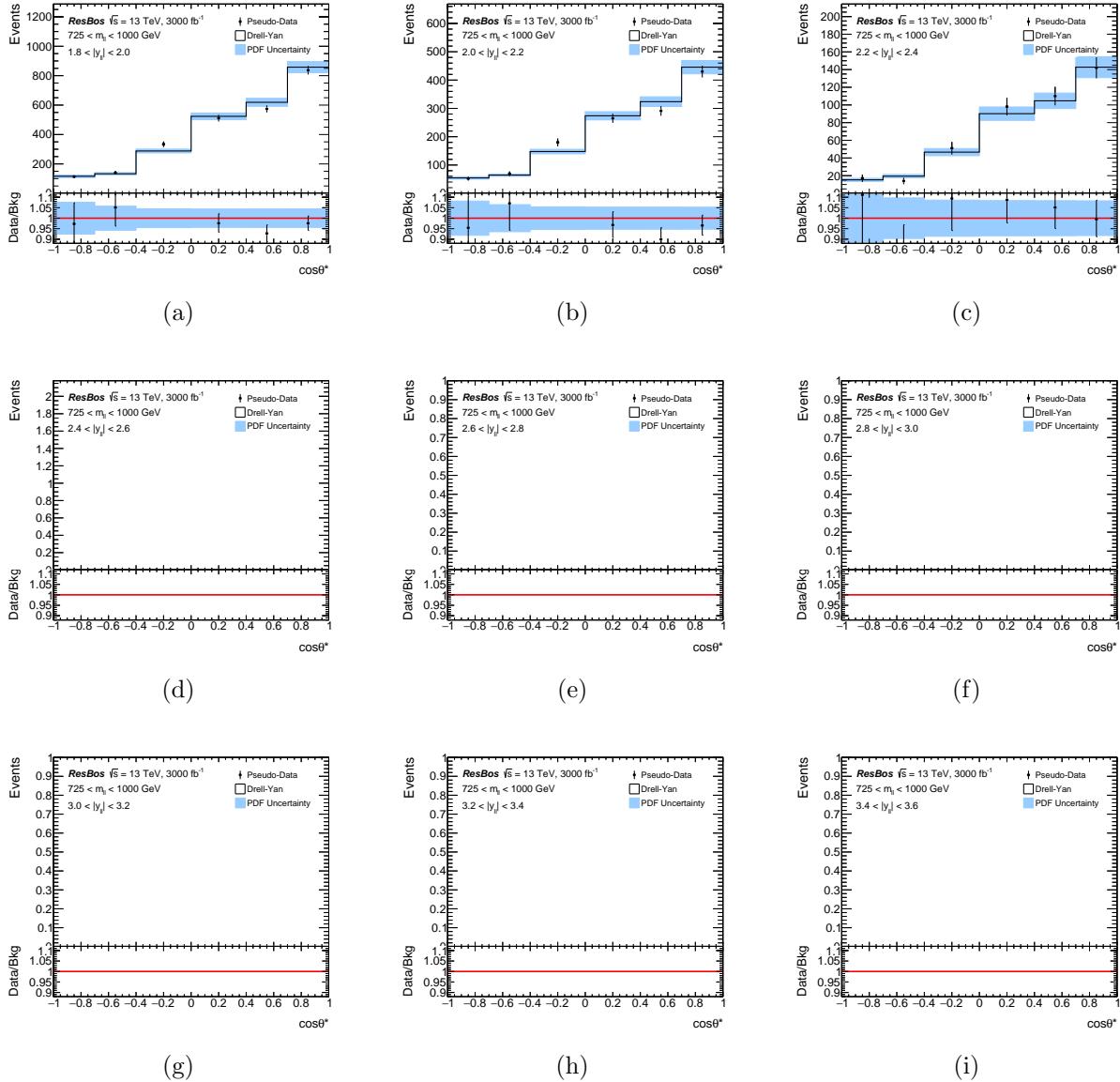


Figure D.24: ePump templates (a)-(i) for the triple-differential cross section of Eq. 13.2 in $m_{\ell\ell} \in [725, 1000]$ GeV and $1.8 < |y_{\ell\ell}| < 3.6$.

Impact on CT14Hera2 PDFs with 300 fb^{-1}

The results of the ePump update on the CT14Hera2 PDFs with the 300 fb^{-1} pseudo-dataset can be seen in Figures D.25 and D.26, and in Table D.1.

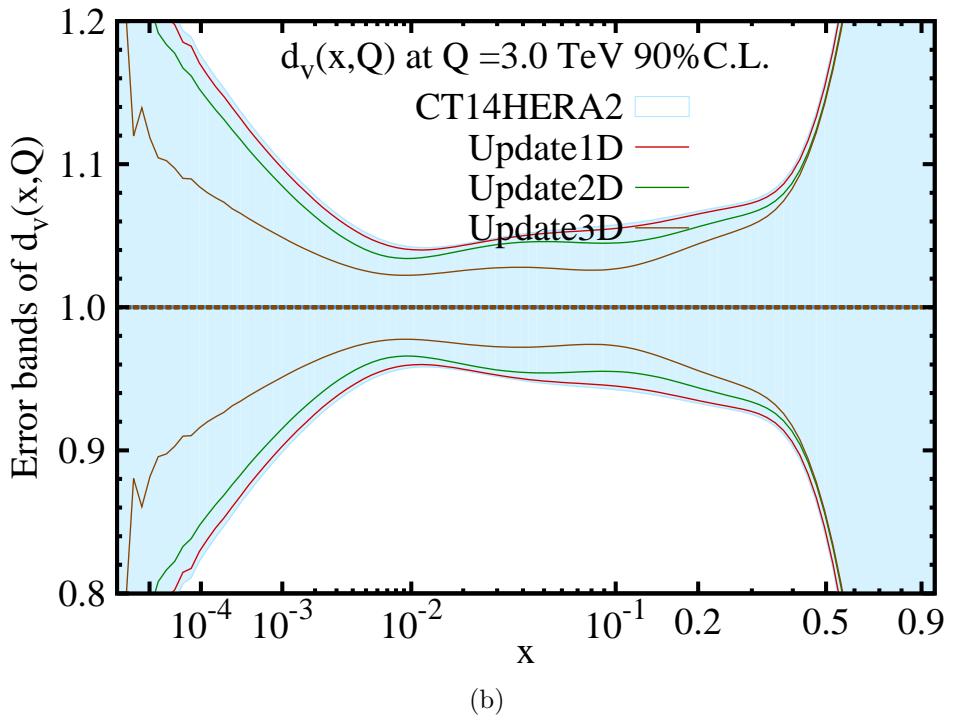
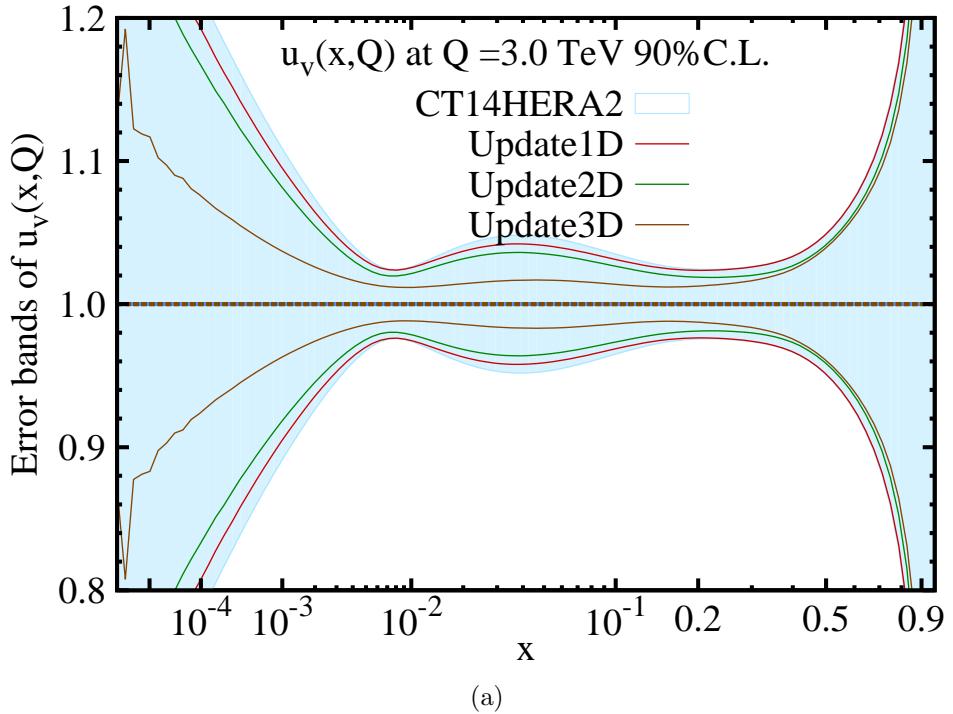
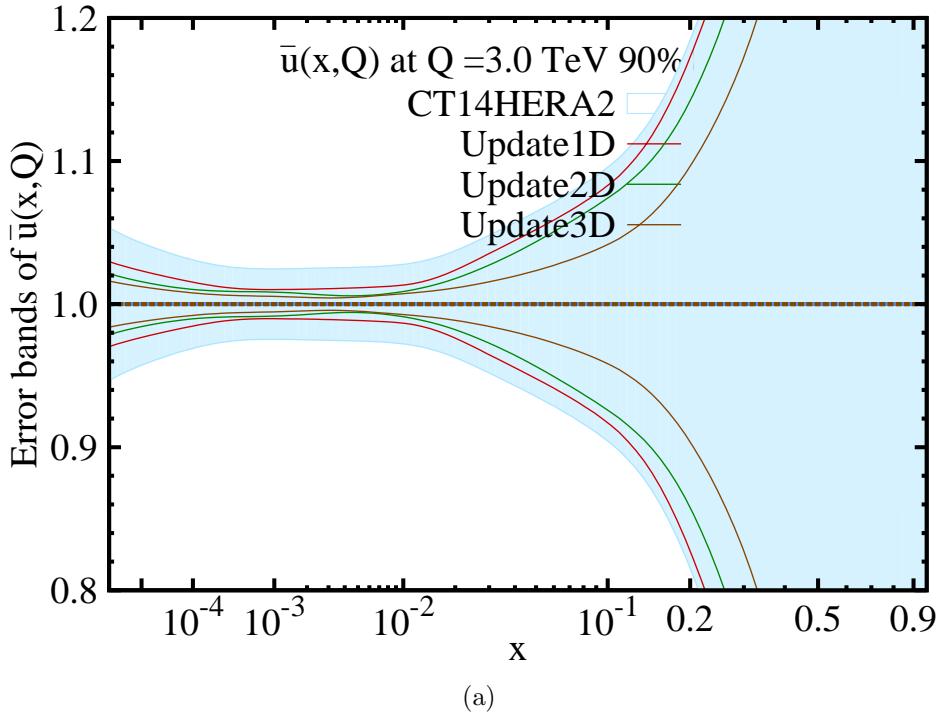
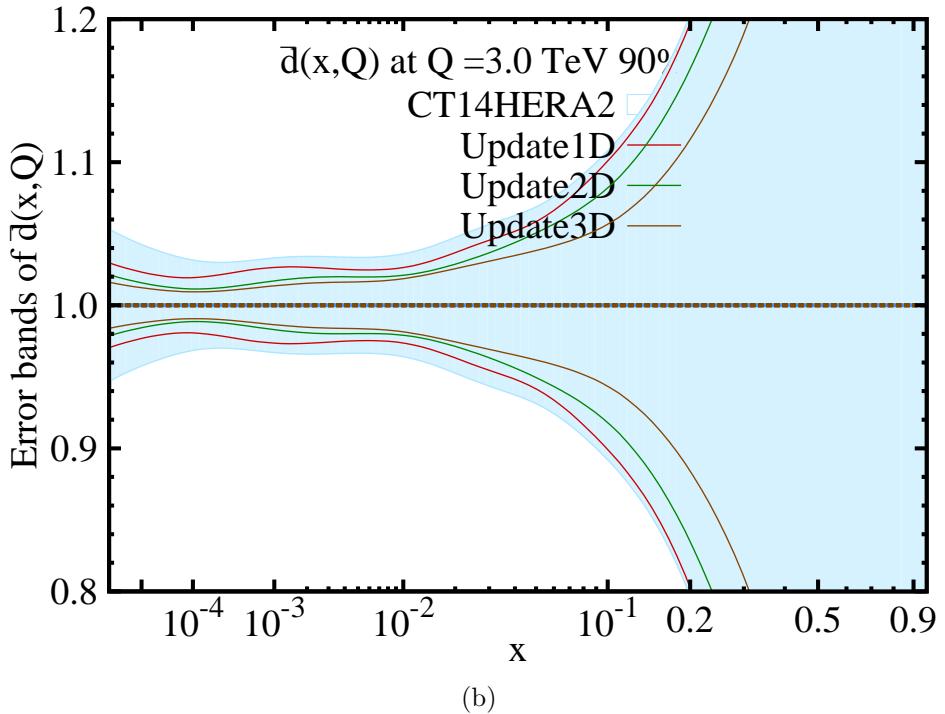


Figure D.25: Impact of the 300 fb^{-1} update on the CT14Hera2 $u_v(x)$ (a) and $d_v(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos \theta^*$ templates: “Update3D” at $Q = 3$ TeV. The utilization of the $\cos \theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty.



(a)



(b)

Figure D.26: Impact of the 300 fb^{-1} update on the CT14Hera2 $\bar{u}(x)$ (a) and $\bar{d}(x)$ (b) PDF uncertainties assuming single-differential mass-only templates: “Update1D,” double-differential mass and rapidity templates: “Update2D,” and the standard triple-differential mass, rapidity, and $\cos \theta^*$ templates: “Update3D” at $Q = 3$ TeV. The utilization of the $\cos \theta^*$ dimension in the update provides for a substantial reduction in the PDF uncertainty.

x	$u_v(x)$		$d_v(x)$		$\bar{u}(x)$		$\bar{d}(x)$	
	δ_{pre} [%]	δ_{post} [%]						
0.1	3.4	1.4	5.8	2.7	9.8	4.3	11	6.0
0.3	2.6	1.6	7.5	5.7	30	17	32	19
0.5	4.8	3.9	16	14	71	43	69	41
0.7	12	9.7	45	41	280	180	250	160

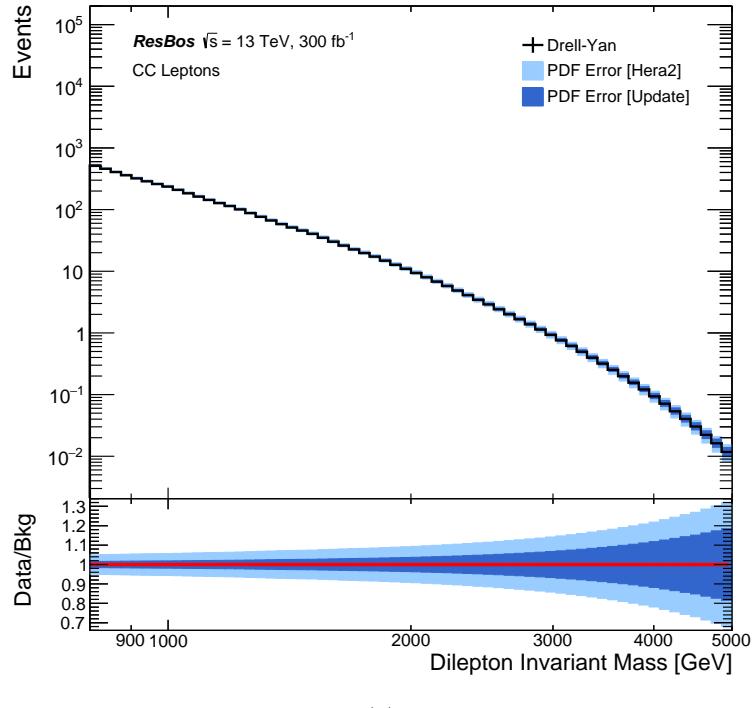
Table D.1: Impact of 300 fb^{-1} update on the **CT14Hera2** $u_v(x)$ and $d_v(x)$ valence and $\bar{u}(x)$ and $\bar{d}(x)$ sea distributions for several values of x using the standard triple-differential templates at $Q = 3 \text{ TeV}$. To be compared with Fig.s D.25 and D.26.

Impact on High-Mass Drell-Yan with 300 fb^{-1}

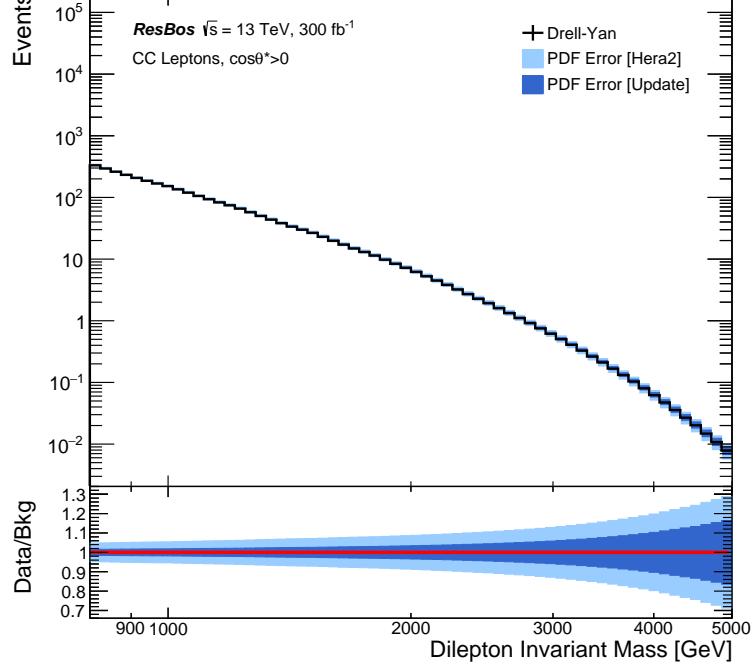
The results of the **ePump** update on the high-mass Drell-Yan spectrum with the 300 fb^{-1} pseudo-dataset can be seen in Figure D.27. Table D.2 depicts the PDF uncertainty before and after the update in several high-mass bins of interest.

$m_{\ell\ell}$ [TeV]	CC Selection		CC+cos θ^* Selection	
	δ_{pre}^{PDF} [%]	δ_{post}^{PDF} [%]	δ_{pre}^{PDF} [%]	δ_{post}^{PDF} [%]
1	5.8	1.9	5.5	1.7
2	9.4	3.8	8.8	3.5
3	15	7.1	13	6.1
4	23	13	21	11
5	32	19	29	16

Table D.2: The estimated PDF uncertainty in several invariant mass bins for the distributions shown in Figure D.27. The current **CT14Hera2** uncertainty estimates are shown in the first column, the result of the 300 fb^{-1} update is shown next, and the 300 fb^{-1} update with an additional $\cos \theta^* > 0$ requirement on the dilepton selection is shown last.



(a)



(b)

Figure D.27: The dilepton invariant mass distribution for central-central dilepton events (a), as would be selected by the dilepton event selection of Section 10.5, and with an additional $\cos\theta^* > 0$ requirement added to the selection (b). The ratio sub-plot depicts the CT14Hera2 PDF uncertainty before and after the 300 fb^{-1} update.

Appendix E

Early Warning System

This Appendix details the Early Warning System (EWS) used in the monitoring of signatures of new physics in the dilepton and lepton+MET analyses during the early data-taking in Run-2 of the LHC.

Physics Motivation

Many BSM physics scenarios predict the existence of new force mediating gauge bosons, such as the electrically neutral Z' or electrically charged W' bosons (See Sec. 5.3). The Run-2 dilepton and lepton+MET Exotics analyses, which search for Z' and W' production respectively, will surpass their Run-1 of the LHC mass limits [33, 34] with as little as a $1\text{-}2 \text{ fb}^{-1}$ of 13 TeV data. As such, if new physics is indeed present, it could start to become apparent very quickly as the data are collected.

The electron transverse energy E_T and muon transverse momentum p_T ¹ spectra offer a unique way to search for new model-independent resonant physics, as the signature of a two body decay would appear as a **Jacobian peak**² over the smoothly falling Standard Model background.

¹For the rest of the section, the “lepton p_T ” will be used in reference to the electron E_T and muon p_T for simplicity.

²A visible feature in the p_T spectrum at roughly half the mass of the resonance due to the back-to-back nature of a two body decay.

To be prepared for potential new physics scenarios, an EWS was developed to monitor the *inclusive* lepton p_T spectrum for anomalies as the data are recorded by the ATLAS detector. This system runs in parallel with the offline dilepton and lepton+MET analyses, and provides crucial feedback to the question “Are we in discovery or exclusion mode?”, thereby allowing the analyses teams to develop a strategy for quicker unblinding if necessary.

Resonant Searches in Run-1

With the onset of Run-2 of the LHC, high energy proton-proton collisions will provide new opportunities to search for BSM physics. The traditional method to search for new resonant phenomena is to reconstruct the dilepton invariant mass $m_{\ell\ell}$ as in the Z' search and the transverse mass m_T as in the W' search. The invariant mass of dilepton pairs is

$$m_{\ell^+\ell^-} = \sqrt{2p_T^{\ell^+} p_T^{\ell^-} (\cosh \Delta\eta - \cos \Delta\phi)}, \quad (\text{E.1})$$

where $p_T^{\ell^+}$ and $p_T^{\ell^-}$ are the lepton transverse momenta, $\Delta\eta$ is the difference in their pseudo-rapidities, and $\Delta\phi$ is the difference in their azimuthal angles in the transverse plane of the ATLAS detector. Likewise, the transverse mass is

$$m_T = \sqrt{2E_T^{\text{miss}} p_T^\ell (1 - \cos \phi_{\ell\nu})}, \quad (\text{E.2})$$

where p_T^ℓ is the lepton transverse momentum, E_T^{miss} is the magnitude of the missing transverse momentum vector, and $\phi_{\ell\nu}$ is the angle between the lepton p_T and neutrino \cancel{E}_T in the transverse plane of the ATLAS detector.

Figure E.1 shows the dielectron invariant Mass and electron+MET transverse mass spec-

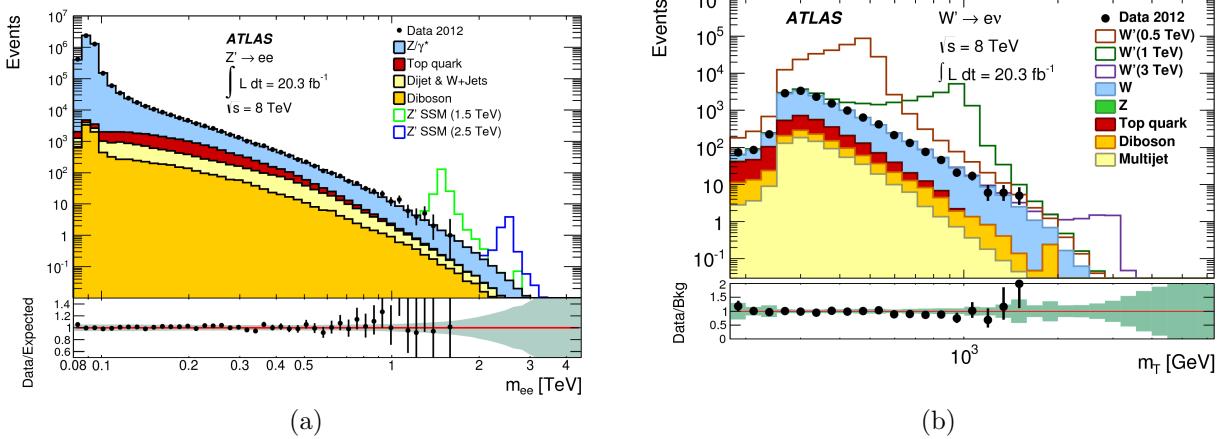


Figure E.1: Run-1 dielectron invariant mass distribution (a) and electron+MET transverse mass distribution (b), the primary search variables in the dilepton and lepton+MET analyses, respectively [33, 34].

tra, the primary search variables in each respective analysis. Because the signature of a new heavy resonance would appear as a resonant peak against the smoothly falling Standard Model background, the above searches have traditionally relied on these two search variables for their respective analyses; however, at the same time, because of the two body nature of each decay, the presence of such a resonance would also be indicated by a Jacobian peak in the lepton p_T spectrum. In this way, the inclusive lepton p_T spectrum could be used to perform a model-independent search for both the Z' and W' bosons *simultaneously*, as well as other resonances decaying into leptons (e.g. a lepton-flavor violating Z' boson [221]).

Furthermore, the search sensitivity of the Exotics Run-1 dilepton and lepton+MET analyses will be quickly surpassed with only 1-2 fb^{-1} of Run-2 data. The Run-1 combined limits for each analysis are shown in Figure E.2. The dilepton analysis interprets the search results in terms of the Sequential Standard Model (SSM)³ (See Sec. 5.3.2), setting a combined channel limit of 2.90 TeV. Similarly, the lepton+MET analysis interprets their results in terms

³The Z'_{SSM} is assigned the same couplings to fermions as the Standard Model Z boson.

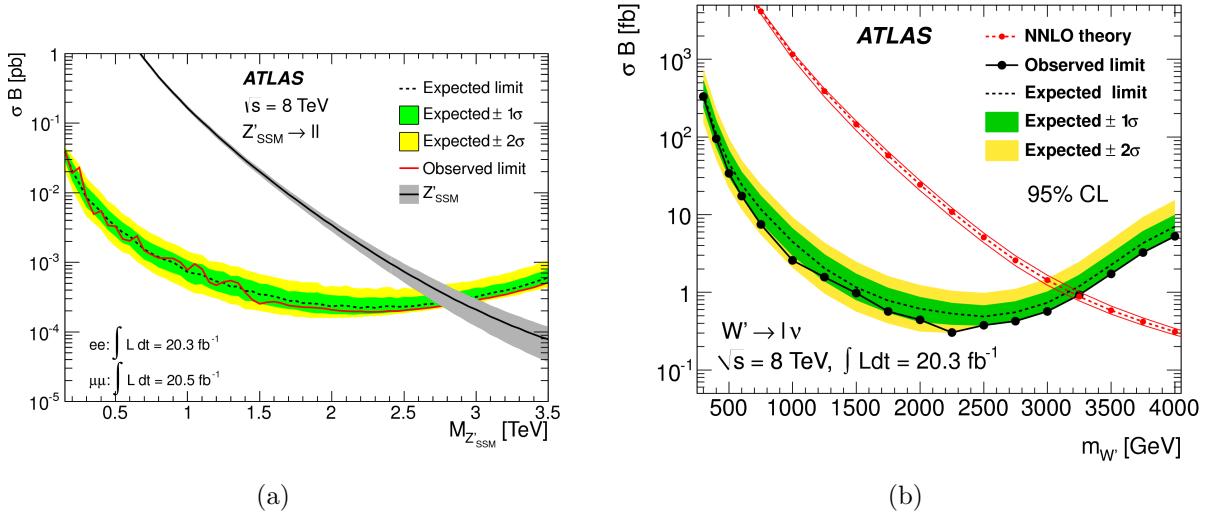


Figure E.2: Run-1 95% C.L. lower mass limits in the dilepton analysis (a) and lepton+MET analysis (b) [33, 34].

of the SSM W' Boson⁴, setting a combined channel limit of 3.24 TeV.

After the first 1-2 fb^{-1} of Run-2 data is collected, the Run-1 search sensitivity will be exceeded. As the amount of data collected will double every few weeks, and then every few months, a new signal really could “jump out” from one week to the next. This demonstrates the need for a robust EWS, especially since the official analysis and subsequent unblinding procedure could take several months, thereby delaying a possible early discovery.

EWS Construction

The inclusive lepton p_T spectrum is used as the primary variable to discriminate between signal and background in a search for Z' and W' bosons. The choice of lepton p_T is based on two important considerations: first, it naturally accommodates many new physics scenarios in a model-independent way, and second, it mitigates any experimental bias, since it is a

⁴The W'_{SSM} is assigned the same couplings to fermions as the Standard Model W boson.

fundamental observable at the LHC. For a two body decay, the signal would appear as a Jacobian peak on the smoothly falling Standard Model background, which would grow with increasing integrated luminosity.

The EWS procedure is as follows: the inclusive lepton p_T distribution is built and updated as new data are recorded, a functional form is fit on a small subrange of the data at low p_T , the form of the fit is extrapolated to high lepton p_T (i.e. the “monitoring” region), and the data are compared to the extrapolation, refitting as new data become available. The monitoring will occur in both the electron and muon channels, such that a peak observed in the electron channel could be verified in the muon channel, which would function as an independent cross check.

Any sort of persistent shape difference between the extrapolated fit and the data in the monitoring region could indicate some kind of anomalous behavior in the high *invariant mass* or high *transverse mass* search regions; in fact it would not be possible to tell *a-priori* as to which is the source. One possible way to identify the source would be to introduce a progressively stronger missing transverse energy \cancel{E}_T requirement into the EWS event selection. A potential W' signal would be insensitive to the \cancel{E}_T requirement, while the Z' signal would be strongly diminished. Alternatively, a two lepton requirement could be used to isolate a potential Z' signal.

EWS Tests on Run-1 Data

In order to test the EWS idea, the full Run-1 LHC dataset is used.⁵. An event selection is developed for the EWS in order to select events from LHC proton-proton collisions that

⁵At the time of its construction, Run-2 data-taking had not yet started.

contain at least one electron or muon. The EWS event selection is modified according to the 8 TeV Run-1 dilepton analysis. The modifications include: an introduction of single electron triggers, a relaxation of the requirement that every event must contain at least 2 electrons, and lastly, the removal of a minimal invariant mass requirement. Furthermore, the EWS selection requires event electrons to pass the same electron identification and isolation criteria as applied to the “leading” electron of the 8 TeV dilepton analysis.

The event level selection criteria used for this study are:

1. Event is in the GoodRunsList (GRL)
2. Event has at least one primary vertex, with number of tracks > 2
3. Event satisfies the trigger configuration:

```
EF_e24vhi_medium1||EF_e60_medium1||EF_g35_loose_g25_loose
```

4. Event fulfills **LArError** < 2 for protection against noise bursts and data corruption

The following object level selection criteria are used as well:

1. Each electron has author 1 or 3
2. Each electron must be within $|\eta| < 2.47$ and exclude the crack region $1.37 < |\eta| < 1.52$
3. Each cluster must pass calorimeter quality requirements
4. Each electron must have $p_T > 40$ GeV
5. Each electron must have at least **isEM MEDIUM++** identification
6. Each electron must be isolated using an E_T dependent cut: **EtCone20_pT_NPV_corrected**
 $< 0.007 \times E_T + 5.0$ GeV

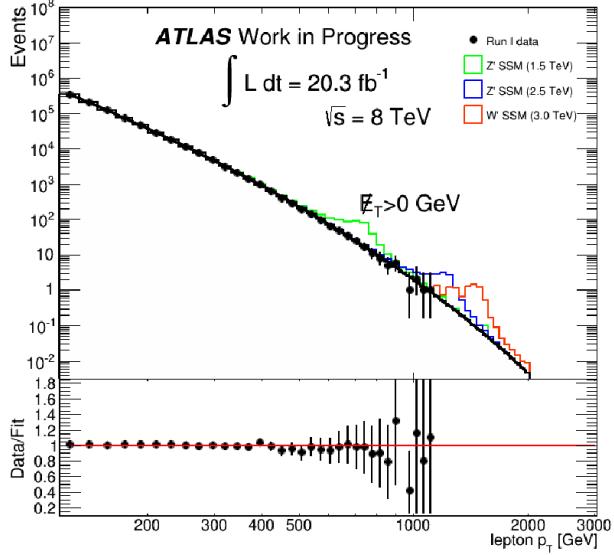


Figure E.3: The EWS applied to the full dataset collected during Run-1 of the LHC. The fit is performed in the low p_T region of [125, 300] GeV and extrapolated to the monitoring region to assess the high- p_T data. Three signals are overlaid using simulation samples: a 1.5 TeV Z'_SSM , a 2.5 TeV Z'_SSM , and a 3.0 TeV W'_SSM , each of which displays a characteristic Jacobian peak in the p_T spectrum.

In the case that multiple electron candidates satisfy the above EWS requirements, each is selected and retained for analysis, in order to build the fully inclusive lepton p_T distribution.

With the requisite lepton p_T distribution, the fit and extrapolation can be performed out to high lepton p_T . Since the shape of the lepton p_T spectrum is unknown *a priori*, a functional form for the fit must be carefully chosen. The so-called **dijet function** [188, 189],

$$f(x) = a(1-x)^b x^{c+d \log(x)}, \quad x = \frac{p_T}{\sqrt{s}}, \quad (\text{E.3})$$

where a , b , and c are parameters to be fitted, is used, as it demonstrated the best reliability among several alternative fits. Follow-up studies showed it was robust against the chosen low- p_T interval of the fit, and the luminosity scale set by pseudo-data.⁶ The fit is performed

⁶Pseudo-data were generated at different characteristic integrated luminosities, and the performance of the EWS fitting was evaluated.

on the nominal low- p_T interval of [125, 300] GeV and is extrapolated to $p_T = 2$ TeV, where the 20.3 fb^{-1} of Run-1 data runs out. The result of the event selection and fitting strategy can be seen in Figure E.3. With an extrapolated background estimate, the need to compare data to simulation at high lepton p_T , where MC statistics start to run out, is relieved.

EWS Operation during Run-2 Data-Taking

The formulation of the EWS for Run-2 followed directly from its successful test-phase in Run-1. As before, the EWS utilized a modified event selection based on the 13 TeV dilepton analysis (See Sec. 10.5). The dilepton requirements were loosened to select events with at least 1 electron, and the invariant mass requirement is removed entirely. All other selection requirements were identical. If multiple electrons passed selection, each was used to construct the electron E_T spectrum, making it fully inclusive. The same strategy was employed in the muon channel, resulting in an inclusive muon p_T spectrum.

Furthermore, the Run-2 EWS ran on TAG dataset derivations, slimmed-down DxAODs which were produced centrally at the CERN Tier-0 computing center [222], and consequently became available several days before the DxAOD derivations used for offline analysis (See Sec.10.2.3). The use of TAG datasets allowed for some level of data-scouting, as the EWS was consistently about a week “ahead” of the offline analysis.

Figures E.4 and E.5, and E.6 and E.7 show several runs of the EWS in the electron and muon channels, respectively, during LHC data-taking from May to July 2016. No signal-like high- p_T excesses were observed in either channel during the 2015 or 2016 LHC runs. However, the EWS allowed for the identification and study of several candidate high- p_T leptons before subsequent observation in the offline dilepton and lepton+MET analyses.

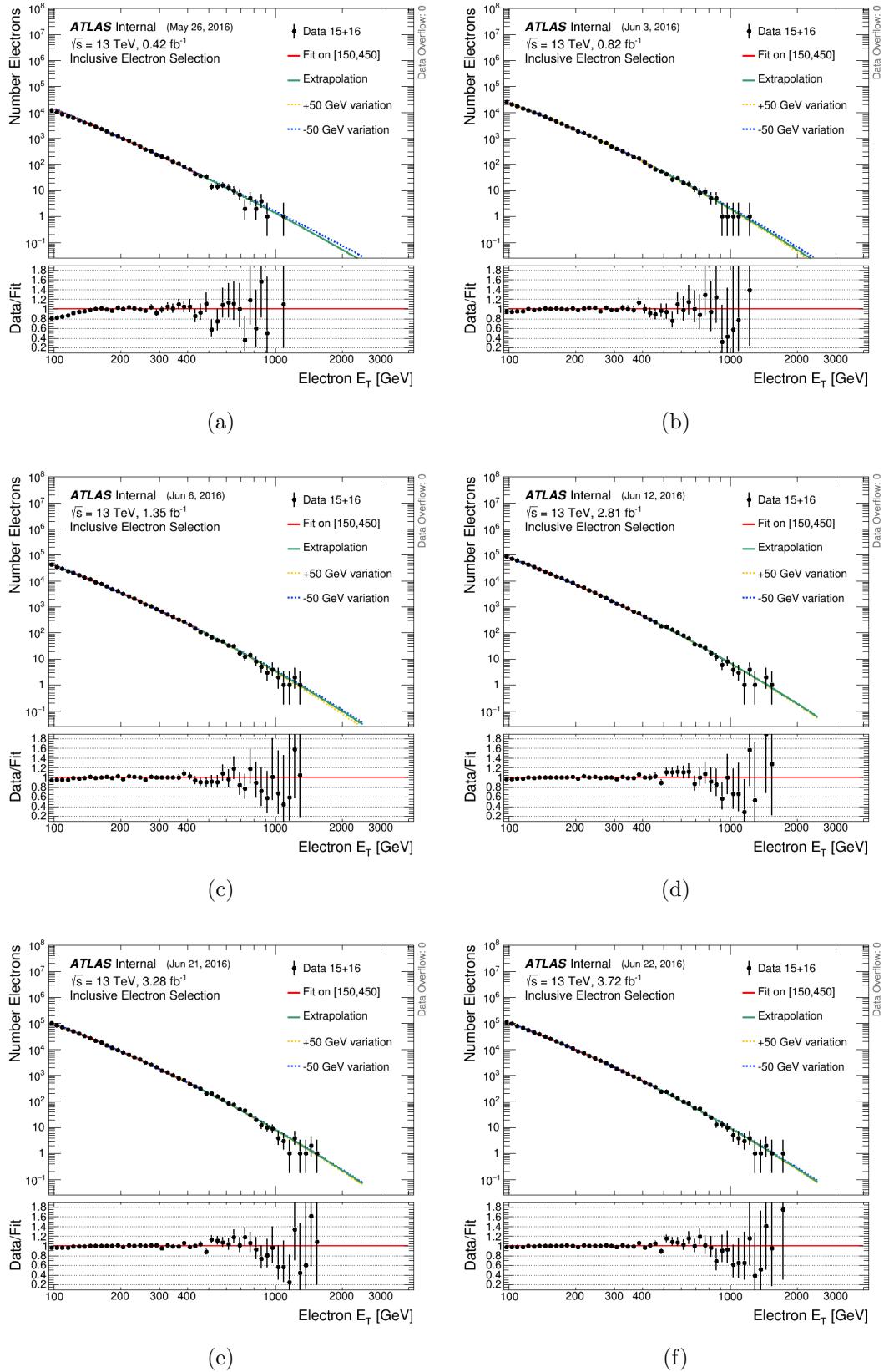


Figure E.4: Operation of the Run-2 EWS in the electron channel: May to June 2016 (a)-(f).

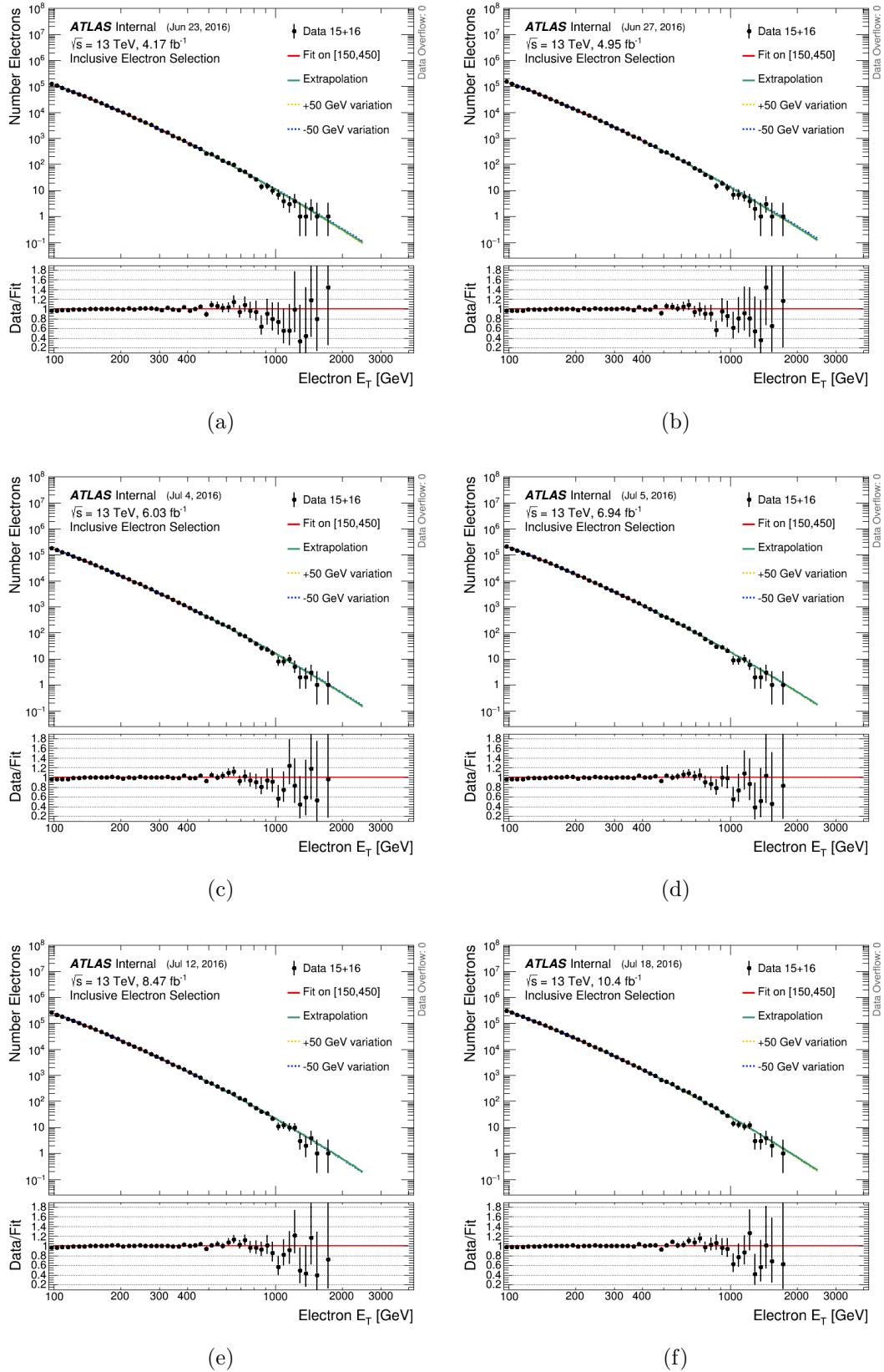


Figure E.5: Operation of the Run-2 EWS in the electron channel: June to July 2016 (a)-(f).

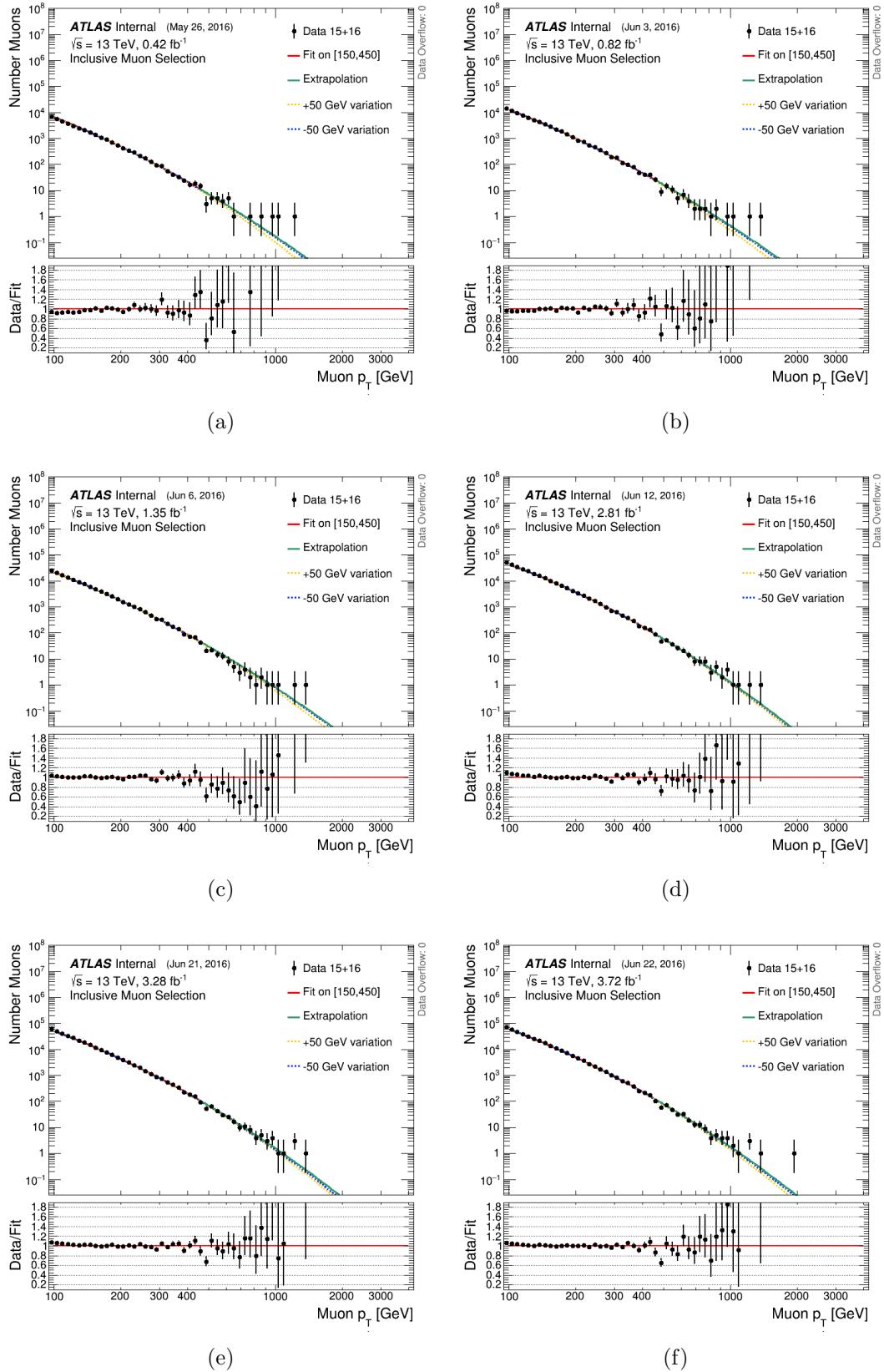


Figure E.6: Operation of the Run-2 EWS in the muon channel: May to June 2016 (a)-(f).

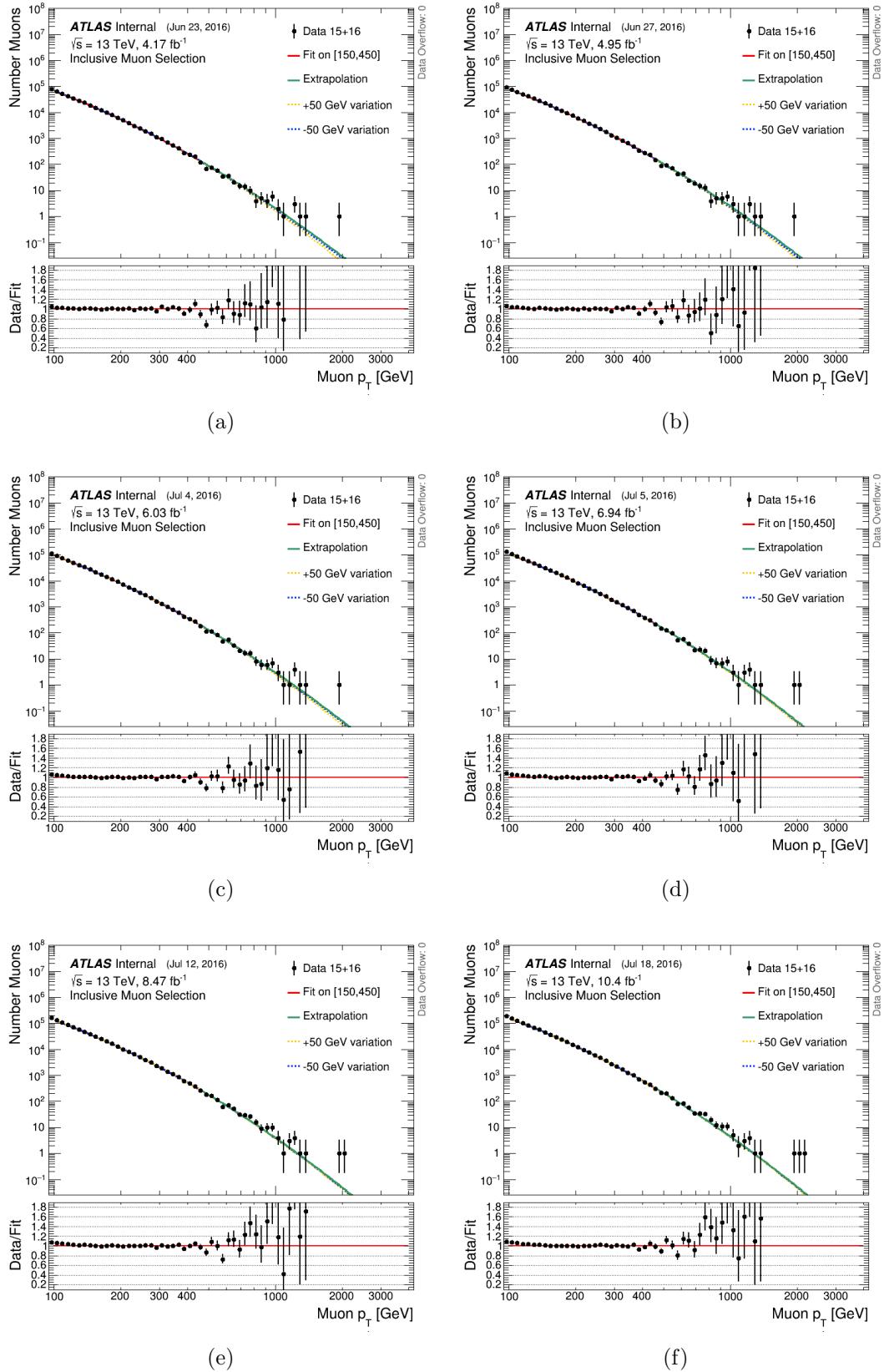


Figure E.7: Operation of the Run-2 EWS in the muon channel: June to July 2016 (a)-(f).

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] Particle Data Group Collaboration, C. Patrignani et al., *Review of Particle Physics*, Chin. Phys. **C40** (2016) no. 10, 100001.
- [2] R. Bruce, G. Arduini, H. Bartosik, R. De Maria, M. Giovannozzi, G. Iadarola, J. Jowett, K. S. B. Li, M. Lamont, A. Lechner, E. Metral, D. Mirarchi, T. Pieloni, S. Redaelli, G. Rumolo, B. Salvant, R. Tomas Garcia, and J. Wenninger, *LHC Run 2: Results and Challenges*, Tech. Rep. CERN-ACC-2016-0103, CERN, Geneva, Jul, 2016. <https://cds.cern.ch/record/2201447>.
- [3] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016.
<https://cds.cern.ch/record/2157687>.
- [4] M. Zinser, D. K. Abhayasinghe, G. Artoni, T. Berry, R. Brock, M. K. Bugge, M. Cano Bret, J. Catmore, M. Chizhov, T. Dai, E. Dreyer, P. J. Falke, O. Fedin, W. Fedorko, W. Fedorko, F. Giuli, D. Hayden, S. Hellesund, N. Hod, T. Hryna'ova, E. E. Khoda, U. Klein, M. Levchenko, Y. Liu, V. Maleev, P. Mastrandrea, V. Morisbak, F. Ould-Saada, V. Pleskot, S. Rettie, S. Schmitz, B. Stelzer, O. Stelzer-Chilton, H. Torres, C. Willis, S. Willocq, I. Yeletskikh, Z. Zhao, and B. Zhou, *Search for new high-mass resonances in the dilepton final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2016-1404, CERN, Geneva, Sep, 2016.
<https://cds.cern.ch/record/2217641>.
- [5] ATLAS Collaboration, M. Aaboud et al., *Search for new high-mass phenomena in the dilepton final state using 36.1 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. CERN-EP-2017-119, Jul, 2017.
<https://cds.cern.ch/record/2273892>.
- [6] CMS Collaboration, W. de Boer, *The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology*, in *Time and Matter 2013 (TAM2013) Venice, Italy*. 2013. arXiv:1309.0721 [hep-ph].
<http://inspirehep.net/record/1252561/files/arXiv:1309.0721.pdf>.
- [7] R. Nathvani, *Spun out of proportion: The Proton Spin Crisis*, 2016.
<https://www.quantumdiaries.org/2016/02/01/spun-out-of-proportion-the-proton-spin-crisis/>.

- [8] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard interactions of quarks and gluons: a primer for LHC physics*, Reports on Progress in Physics **70** (2007) no. 1, 89. <http://stacks.iop.org/0034-4885/70/i=1/a=R02>.
- [9] S. Höche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*. 2014. arXiv:1411.4085 [hep-ph].
<https://inspirehep.net/record/1328513/files/arXiv:1411.4085.pdf>.
- [10] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C.-P. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (Feb, 2016) 033006.
<https://link.aps.org/doi/10.1103/PhysRevD.93.033006>.
- [11] CMS Collaboration, S. Chatrchyan et al., *Measurement of the weak mixing angle with the Drell-Yan process in proton-proton collisions at the LHC*, Phys. Rev. **D84** (2011) 112002, arXiv:1110.2682 [hep-ex].
- [12] E. Yatsenko et al., *Measurement of Neutral Current Drell-Yan production using ATLAS data collected at $\sqrt{s} = 8$ TeV*, Nov, 2014.
<https://cds.cern.ch/record/2041829>.
- [13] O. Brüning, *Reparatur und Umbau des LHC*, 2015.
<https://www.weltderphysik.de/gebiet/teilchen/experimente/teilchenbeschleuniger/lhc/reparatur-und-umbau-des-lhc>. Last Accessed Feb 9, 2018.
- [14] F. Marcastel, *CERN's Accelerator Complex. La chaîne des accélérateurs du CERN*, Tech. Rep. OPEN-PHO-CHART-2013-001, CERN, Geneva, Oct, 2013.
<https://cds.cern.ch/record/1621583>.
- [15] CERN, *The Large Hadron Collider*, 2014. <https://cds.cern.ch/record/1998498>. Last Accessed Mar 2018.
- [16] L. R. F. Castillo, *The Large Hadron Collider*, in *The Search and Discovery of the Higgs Boson*, 2053-2571, pp. 3–1 to 3–9. Morgan & Claypool Publishers, 2015.
<http://dx.doi.org/10.1088/978-1-6817-4078-2ch3>.
- [17] E. Torrence, *LuminosityPublicResults*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults/>. Last Accessed Jan 2018.

- [18] E. Torrence, *LuminosityPublicResultsRun2*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>. Last Accessed Jan 2018.
- [19] M. L. Dunford and P. Jenni, *The ATLAS experiment*, Scholarpedia **9** (2014) no. 10, 32147. revision #184627.
- [20] ATLAS Collaboration, K. Sliwa, *ATLAS Overview and Main Results*, in *Proceedings, International School on High Energy Physics: Workshop on High Energy Physics in the near Future. (LISHEP 2013): Rio de Janeiro, Brazil, March 17-24, 2013.* 2013. arXiv:1305.4551 [hep-ex].
<http://inspirehep.net/record/1234234/files/arXiv:1305.4551.pdf>.
- [21] J. Pequenao, *Computer generated image of the ATLAS inner detector*, Tech. Rep. CERN-GE-0803014, CERN, Geneva, Mar, 2008.
<http://cds.cern.ch/record/1095926>.
- [22] J. Pequenao, *Computer generated image of the ATLAS calorimeter*, Tech. Rep. CERN-GE-0803015, CERN, Geneva, Mar, 2008.
<http://cds.cern.ch/record/1095927>.
- [23] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08003.
<http://stacks.iop.org/1748-0221/3/i=08/a=S08003>.
- [24] B. Sotto-Maior Peralva, *Calibration and Performance of the ATLAS Tile Calorimeter*, in *Proceedings, International School on High Energy Physics : Workshop on High Energy Physics in the near Future. (LISHEP 2013): Rio de Janeiro, Brazil, March 17-24, 2013.* 2013. arXiv:1305.0550 [physics.ins-det].
<http://inspirehep.net/record/1231740/files/arXiv:1305.0550.pdf>.
- [25] J. Pinfold et al., *Evaluation of the local hadronic calibration with combined beam-test data for the endcap and forward calorimeters of ATLAS in the pseudorapidity region $2.5 < |\eta| < 4.0$* , Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **693** (2012) 74 – 97. <http://www.sciencedirect.com/science/article/pii/S016890021200705X>.
- [26] ATLAS Collaboration, E. Diehl, *Calibration and Performance of the ATLAS Muon Spectrometer*, in *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2011, Providence, USA, August 9-13, 2011.* 2011. arXiv:1109.6933 [physics.ins-det].
<http://inspirehep.net/record/930228/files/arXiv:1109.6933.pdf>.
- [27] M. Limper, S. Bentvelsen, and A. P. Colijn, *Track and vertex reconstruction in the ATLAS inner detector*, 2009. <https://cds.cern.ch/record/1202457>.

- [28] J.-B. De Vivie De Regie, J. Guo, N. Kondrashova, S. Laplace, G. R. Lee, A. Lesage, J. Liu, S. M. Mazza, J. Morris, R. Nicolaidou, J. Poveda, N. Proklova, Y.-T. Shen, J. G. Vasquez, Z. Wang, H. Yang, M. Zgubic, D. Zhang, and L. Zhang, *ATLAS electron, photon and muon isolation in Run 2*, Tech. Rep. ATL-COM-PHYS-2017-290, CERN, Geneva, Mar, 2017.
<https://cds.cern.ch/record/2256658>.
- [29] Graipher (<https://physics.stackexchange.com/users/74720/graipher>), *Sea quark parton annihilation?*, 2017. <https://physics.stackexchange.com/q/312600>. Last Accessed Jun 4 2018.
- [30] M. Boonekamp, F. Chevallier, C. Royon, and L. Schoeffel, *Understanding the Structure of the Proton: From HERA and Tevatron to LHC*, Acta Phys. Polon. **B40** (2009) 2239–2321, [arXiv:0902.1678 \[hep-ph\]](https://arxiv.org/abs/0902.1678).
- [31] DØ Collaboration, V. M. Abazov et al., *Measurement of $\sin^2 \theta_{\text{eff}}^\ell$ and Z-light quark couplings using the forward-backward charge asymmetry in $p\bar{p} \rightarrow Z/\gamma^* \rightarrow e^+e^-$ events with $\mathcal{L} = 5.0 \text{ fb}^{-1}$ at $\sqrt{s} = 1.96 \text{ TeV}$* , Phys. Rev. D **84** (Jul, 2011) 012007. <https://link.aps.org/doi/10.1103/PhysRevD.84.012007>.
- [32] D. Hayden, R. Brock, and C. Willis, *Z Prime: A Story*, in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*. 2013. [arXiv:1308.5874 \[hep-ex\]](https://arxiv.org/abs/1308.5874).
<https://inspirehep.net/record/1251060/files/arXiv:1308.5874.pdf>.
- [33] ATLAS Collaboration, G. Aad et al., *Search for high-mass dilepton resonances in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector*, Phys. Rev. **D90** (2014) no. 5, 052005, [arXiv:1405.4123 \[hep-ex\]](https://arxiv.org/abs/1405.4123).
- [34] ATLAS Collaboration, G. Aad et al., *Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector*, JHEP **09** (2014) 037, [arXiv:1407.7494 \[hep-ex\]](https://arxiv.org/abs/1407.7494).
- [35] J. J. Thomson M.A. F.R.S., *XL. Cathode Rays*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **44** (1897) no. 269, 293–316.
<https://doi.org/10.1080/14786449708621070>.
- [36] E. Rutherford F.R.S., *LXXIX. The scattering of α and β particles by matter and the structure of the atom*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **21** (1911) no. 125, 669–688.
<https://doi.org/10.1080/14786440508637080>.

- [37] J. Chadwick, F.R.S., *The existence of a neutron*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **136** (1932) no. 830, 692–708. <http://rspa.royalsocietypublishing.org/content/136/830/692>.
- [38] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964) no. 3, 214 – 215.
<http://www.sciencedirect.com/science/article/pii/S0031916364920013>.
- [39] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 2*, . <https://cds.cern.ch/record/570209>.
- [40] V. E. Barnes et al., *Observation of a Hyperon with Strangeness Minus Three*, Phys. Rev. Lett. **12** (Feb, 1964) 204–206.
<https://link.aps.org/doi/10.1103/PhysRevLett.12.204>.
- [41] E. D. Bloom, D. H. Coward, H. DeStaeler, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenbach, J. I. Friedman, G. C. Hartmann, and H. W. Kendall, *High-Energy Inelastic $e - p$ Scattering at 6° and 10°* , Phys. Rev. Lett. **23** (Oct, 1969) 930–934. <https://link.aps.org/doi/10.1103/PhysRevLett.23.930>.
- [42] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. DeStaeler, J. Drees, L. W. Mo, and R. E. Taylor, *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (Oct, 1969) 935–939.
<https://link.aps.org/doi/10.1103/PhysRevLett.23.935>.
- [43] D. P. Barber et al., *Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA*, Phys. Rev. Lett. **43** (Sep, 1979) 830–833.
<https://link.aps.org/doi/10.1103/PhysRevLett.43.830>.
- [44] UA2 Collaboration, M. Banner et al., *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider*, Physics Letters B **122** (1983) no. 5, 476 – 485.
<http://www.sciencedirect.com/science/article/pii/0370269383916052>.
- [45] UA1 Collaboration, G. Arnison et al., *Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}=540$ GeV*, Physics Letters B **122** (1983) no. 1, 103 – 116.
<http://www.sciencedirect.com/science/article/pii/0370269383911772>.
- [46] UA2 Collaboration, P. Bagnaia et al., *Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ collider*, Physics Letters B **129** (1983) no. 1, 130 – 140.
<http://www.sciencedirect.com/science/article/pii/037026938390744X>.

- [47] UA1 Collaboration, G. Arnison et al., *Experimental observation of lepton pairs of invariant mass around 95 GeV/c² at the CERN SPS collider*, Physics Letters B **126** (1983) no. 5, 398 – 410.
<http://www.sciencedirect.com/science/article/pii/0370269383901880>.
- [48] CDF Collaboration, F. Abe et al., *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (Apr, 1995) 2626–2631. <https://link.aps.org/doi/10.1103/PhysRevLett.74.2626>.
- [49] DØ Collaboration, S. Abachi et al., *Observation of the Top Quark*, Phys. Rev. Lett. **74** (Apr, 1995) 2632–2637.
<https://link.aps.org/doi/10.1103/PhysRevLett.74.2632>.
- [50] DONUT Collaboration, K. Kodama et al., *Observation of tau neutrino interactions*, Physics Letters B **504** (2001) no. 3, 218 – 224.
<http://www.sciencedirect.com/science/article/pii/S0370269301003070>.
- [51] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) no. 1, 30 – 61.
<http://www.sciencedirect.com/science/article/pii/S0370269312008581>.
- [52] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) no. 1, 1 – 29.
<http://www.sciencedirect.com/science/article/pii/S037026931200857X>.
- [53] D. London and J. L. Rosner, *Extra gauge bosons in E₆*, Phys. Rev. D **34** (Sep, 1986) 1530–1546. <https://link.aps.org/doi/10.1103/PhysRevD.34.1530>.
- [54] G. Kane, *Modern elementary particle physics: the fundamental particles and forces?* Addison-Wesley Publishing Group, 1993.
- [55] P. A. M. Dirac, *The Quantum Theory of the Electron*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **117** (1928) no. 778, 610–624.
<http://rspa.royalsocietypublishing.org/content/117/778/610>.
- [56] C. D. Anderson, *The Positive Electron*, Phys. Rev. **43** (Mar, 1933) 491–494.
<https://link.aps.org/doi/10.1103/PhysRev.43.491>.
- [57] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (Oct, 1964) 508–509.
<https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.

- [58] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323.
<https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [59] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (Nov, 1964) 585–587.
<https://link.aps.org/doi/10.1103/PhysRevLett.13.585>.
- [60] V. Barger and R. Phillips, *Collider Physics*. Westview Press, 1991.
- [61] E. Noether, *Invariant variation problems*, Transport Theory and Statistical Physics **1** (1971) no. 3, 186–207. <https://doi.org/10.1080/00411457108231446>.
- [62] E. Fermi, *Versuch einer Theorie der β -Strahlen. I*, Zeitschrift fr Physik **88** (1934) no. 3, 161–177. <http://dx.doi.org/10.1007/BF01351864>.
- [63] F. L. Wilson, *Fermi's Theory of Beta Decay*, American Journal of Physics **36** (1968) no. 12, 1150–1160. <http://dx.doi.org/10.1119/1.1974382>.
- [64] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (Feb, 1957) 1413–1415. <https://link.aps.org/doi/10.1103/PhysRev.105.1413>.
- [65] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (Jul, 1964) 138–140.
<https://link.aps.org/doi/10.1103/PhysRevLett.13.138>.
- [66] R. P. Feynman and M. Gell-Mann, *Theory of the Fermi Interaction*, Phys. Rev. **109** (Jan, 1958) 193–198. <https://link.aps.org/doi/10.1103/PhysRev.109.193>.
- [67] E. C. G. Sudarshan and R. E. Marshak, *Chirality Invariance and the Universal Fermi Interaction*, Phys. Rev. **109** (Mar, 1958) 1860–1862.
<https://link.aps.org/doi/10.1103/PhysRev.109.1860.2>.
- [68] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (Oct, 1954) 191–195.
<https://link.aps.org/doi/10.1103/PhysRev.96.191>.
- [69] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) no. 4, 579 – 588.
<http://www.sciencedirect.com/science/article/pii/0029558261904692>.
- [70] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (Nov, 1967) 1264–1266.
<https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.

- [71] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377. <http://inspirehep.net/record/53083>.
- [72] Gargamelle Collaboration, F. Hasert et al., *Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment*, Physics Letters B **46** (1973) no. 1, 138 – 140.
<http://www.sciencedirect.com/science/article/pii/0370269373904991>.
- [73] Gargamelle Collaboration, F. Hasert et al., *Search for elastic muon-neutrino electron scattering*, Physics Letters B **46** (1973) no. 1, 121 – 124.
<http://www.sciencedirect.com/science/article/pii/0370269373904942>.
- [74] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (Jun, 1963) 531–533. <https://link.aps.org/doi/10.1103/PhysRevLett.10.531>.
- [75] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657.
<http://dx.doi.org/10.1143/PTP.49.652>.
- [76] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (Nov, 1983) 1945–1947.
<https://link.aps.org/doi/10.1103/PhysRevLett.51.1945>.
- [77] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (Jun, 1973) 1343–1346.
<https://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [78] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (Jun, 1973) 1346–1349.
<https://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [79] G. Altarelli, *QCD evolution equations for parton densities*, Scholarpedia **4** (2009) no. 1, 7124. revision #91681.
- [80] A. Martin, R. Roberts, W. Stirling, and R. Thorne, *Physical gluons and high-ET jets*, Physics Letters B **604** (2004) no. 1, 61 – 68.
<http://www.sciencedirect.com/science/article/pii/S0370269304015059>.
- [81] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, *Parton distributions for the LHC run II*, Journal of High Energy Physics **2015** (Apr, 2015) 40. [https://doi.org/10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).

- [82] R. Placakyte, *Parton Distribution Functions*, in *Proceedings, 31st International Conference on Physics in collisions (PIC 2011): Vancouver, Canada, August 28-September 1, 2011.* 2011. arXiv:1111.5452 [hep-ph].
<https://inspirehep.net/record/954990/files/arXiv:1111.5452.pdf>.
- [83] J. Feltesse, *Introduction to Parton Distribution Functions*, Scholarpedia **5** (2010) no. 11, 10160. revision #91386.
- [84] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, *Uncertainties of predictions from parton distribution functions. II. The Hessian method*, Phys. Rev. D **65** (Dec, 2001) 014013.
<https://link.aps.org/doi/10.1103/PhysRevD.65.014013>.
- [85] J. Pumplin, D. R. Stump, and W. K. Tung, *Multivariate fitting and the error matrix in global analysis of data*, Phys. Rev. D **65** (Dec, 2001) 014011.
<https://link.aps.org/doi/10.1103/PhysRevD.65.014011>.
- [86] F. Mandl and G. Shaw, *Quantum Field Theory*. John Wiley and Sons, 2010.
- [87] D. Griffiths, *Introduction to Elementary Particles*. Wiley-VCH, 2008.
- [88] S. D. Drell and T.-M. Yan, *Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies*, Phys. Rev. Lett. **25** (Aug, 1970) 316–320.
<https://link.aps.org/doi/10.1103/PhysRevLett.25.316>.
- [89] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$* , Phys. Rev. D **74** (Dec, 2006) 114017.
<https://link.aps.org/doi/10.1103/PhysRevD.74.114017>.
- [90] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, *Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order*, Phys. Rev. Lett. **103** (Aug, 2009) 082001.
<https://link.aps.org/doi/10.1103/PhysRevLett.103.082001>.
- [91] G. A. Ladinsky and C.-P. Yuan, *Nonperturbative regime in QCD resummation for gauge boson production at hadron colliders*, Phys. Rev. D **50** (Oct, 1994) R4239–R4243. <https://link.aps.org/doi/10.1103/PhysRevD.50.R4239>.
- [92] C. Balázs and C.-P. Yuan, *Soft gluon effects on lepton pairs at hadron colliders*, Phys. Rev. D **56** (Nov, 1997) 5558–5583.
<https://link.aps.org/doi/10.1103/PhysRevD.56.5558>.
- [93] F. Landry, R. Brock, P. M. Nadolsky, and C.-P. Yuan, *Fermilab Tevatron run-1 Z boson data and the Collins-Soper-Sterman resummation formalism*, Phys. Rev. D **67** (Apr, 2003) 073016. <https://link.aps.org/doi/10.1103/PhysRevD.67.073016>.

- [94] T.-J. Hou, S. Dulat, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C.-P. Yuan, *CTEQ-TEA parton distribution functions and HERA Run I and II combined data*, Phys. Rev. D **95** (Feb, 2017) 034003.
<https://link.aps.org/doi/10.1103/PhysRevD.95.034003>.
- [95] M. C. Gonzalez-Garcia and Y. Nir, *Neutrino masses and mixing: evidence and implications*, Rev. Mod. Phys. **75** (Mar, 2003) 345–402.
<https://link.aps.org/doi/10.1103/RevModPhys.75.345>.
- [96] N. Palanque-Delabrouille, C. Yèche, J. Baur, C. Magneville, G. Rossi, J. Lesgourgues, A. Borde, E. Burtin, J.-M. LeGoff, J. Rich, M. Viel, and D. Weinberg, *Neutrino masses and cosmology with Lyman-alpha forest power spectrum*, Journal of Cosmology and Astroparticle Physics **2015** (2015) no. 11, 011.
<http://stacks.iop.org/1475-7516/2015/i=11/a=011>.
- [97] R. N. Mohapatra and G. Senjanović, *Neutrino Mass and Spontaneous Parity Nonconservation*, Phys. Rev. Lett. **44** (Apr, 1980) 912–915.
<https://link.aps.org/doi/10.1103/PhysRevLett.44.912>.
- [98] R. N. Mohapatra and G. Senjanović, *Neutrino masses and mixings in gauge models with spontaneous parity violation*, Phys. Rev. D **23** (Jan, 1981) 165–180.
<https://link.aps.org/doi/10.1103/PhysRevD.23.165>.
- [99] V. C. Rubin and W. K. Ford, Jr., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, Astrophys. J. **159** (1970) 379–403.
- [100] K. Freese, *Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars*, EAS Publications Series **36** (2009) 113–126.
<https://doi.org/10.1051/eas/0936016>.
- [101] A. G. Riess et al., *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, The Astronomical Journal **116** (1998) no. 3, 1009. <http://stacks.iop.org/1538-3881/116/i=3/a=1009>.
- [102] M. Trodden, *Electroweak baryogenesis*, Rev. Mod. Phys. **71** (Oct, 1999) 1463–1500.
<https://link.aps.org/doi/10.1103/RevModPhys.71.1463>.
- [103] G. F. Giudice, *Naturally Speaking: The Naturalness Criterion and Physics at the LHC*, pp. 155–178. 2012.
https://www.worldscientific.com/doi/abs/10.1142/9789812779762_0010.
- [104] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *The hierarchy problem and new dimensions at a millimeter*, Physics Letters B **429** (1998) no. 3, 263 – 272.
<http://www.sciencedirect.com/science/article/pii/S0370269398004663>.

- [105] L. Randall and R. Sundrum, *Large Mass Hierarchy from a Small Extra Dimension*, Phys. Rev. Lett. **83** (Oct, 1999) 3370–3373.
<https://link.aps.org/doi/10.1103/PhysRevLett.83.3370>.
- [106] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*, Phys. Rev. Lett. **32** (Feb, 1974) 438–441.
<https://link.aps.org/doi/10.1103/PhysRevLett.32.438>.
- [107] P. Langacker, *The physics of heavy Z' gauge bosons*, Rev. Mod. Phys. **81** (Aug, 2009) 1199–1228. <https://link.aps.org/doi/10.1103/RevModPhys.81.1199>.
- [108] Super-Kamiokande Collaboration, K. Abe et al., *Search for proton decay via $p \rightarrow e^+ \pi^0$ and $p \rightarrow \mu^+ \pi^0$ in 0.31 megaton·years exposure of the Super-Kamiokande water Cherenkov detector*, Phys. Rev. D **95** (Jan, 2017) 012004.
<https://link.aps.org/doi/10.1103/PhysRevD.95.012004>.
- [109] T. G. Rizzo, *Z' phenomenology and the LHC*, in *Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics : Exploring New Frontiers Using Colliders and Neutrinos (TASI 2006): Boulder, Colorado, June 4-30, 2006*, pp. 537–575. 2006. arXiv:hep-ph/0610104 [hep-ph]. <http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-pub-12129>.
- [110] E. Accomando, A. Belyaev, L. Fedeli, S. F. King, and C. Shepherd-Themistocleous, *Z' physics with early LHC data*, Phys. Rev. D **83** (Apr, 2011) 075012.
<https://link.aps.org/doi/10.1103/PhysRevD.83.075012>.
- [111] ATLAS Collaboration, M. Aaboud et al., *Search for high-mass new phenomena in the dilepton final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Physics Letters B **761** (2016) no. Supplement C, 372 – 392.
<http://www.sciencedirect.com/science/article/pii/S0370269316304750>.
- [112] M. Dittmar, A.-S. Nicollerat, and A. Djouadi, *Z' studies at the LHC: an update*, Physics Letters B **583** (2004) no. 12, 111 – 120.
<http://www.sciencedirect.com/science/article/pii/S0370269304000760>.
- [113] E. Salvioni, G. Villadoro, and F. Zwirner, *Minimal Z' models: present bounds and early LHC reach*, Journal of High Energy Physics **2009** (2009) no. 11, 068.
<http://stacks.iop.org/1126-6708/2009/i=11/a=068>.
- [114] L. Basso, *Minimal Z' models and the 125 GeV Higgs boson*, Physics Letters B **725** (2013) no. 4, 322 – 326.
<http://www.sciencedirect.com/science/article/pii/S037026931300587X>.

- [115] G. Senjanović and R. N. Mohapatra, *Exact left-right symmetry and spontaneous violation of parity*, Phys. Rev. D **12** (Sep, 1975) 1502–1505.
<https://link.aps.org/doi/10.1103/PhysRevD.12.1502>.
- [116] R. N. Mohapatra and J. C. Pati, *Left-right gauge symmetry and an “isoconjugate” model of CP violation*, Phys. Rev. D **11** (Feb, 1975) 566–571.
<https://link.aps.org/doi/10.1103/PhysRevD.11.566>.
- [117] L. Basso, A. Belyaev, S. Moretti, and C. H. Shepherd-Themistocleous, *Phenomenology of the minimal $B - L$ extension of the standard model: Z' and neutrinos*, Phys. Rev. D **80** (Sep, 2009) 055030.
<https://link.aps.org/doi/10.1103/PhysRevD.80.055030>.
- [118] J. Erler, P. Langacker, S. Munir, and E. Rojas, *Improved constraints on Z' bosons from electroweak precision data*, Journal of High Energy Physics **2009** (2009) no. 08, 017. <http://stacks.iop.org/1126-6708/2009/i=08/a=017>.
- [119] ALEPH, DELPHI, L3, and OPAL Collaboration, LEP Electroweak Working Group, *Electroweak measurements in electron–positron collisions at W -boson-pair energies at LEP*, Physics Reports **532** (2013) no. 4, 119 – 244.
<http://www.sciencedirect.com/science/article/pii/S0370157313002706>.
- [120] DØ Collaboration, V. M. Abazovai et al., *Search for a heavy neutral gauge boson in the dielectron channel with 5.4 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$* , Physics Letters B **695** (2011) no. 1, 88 – 94.
<http://www.sciencedirect.com/science/article/pii/S0370269310012700>.
- [121] CDF Collaboration, T. Aaltonen et al., *Search for High Mass Resonances Decaying to Muon Pairs in $\sqrt{s} = 1.96 \text{ TeV}$ $p\bar{p}$ Collisions*, Phys. Rev. Lett. **106** (Mar, 2011) 121801. <https://link.aps.org/doi/10.1103/PhysRevLett.106.121801>.
- [122] CMS Collaboration, V. Khachatryan et al., *Search for physics beyond the standard model in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 8 \text{ TeV}$* , Journal of High Energy Physics **2015** (Apr, 2015) 25.
[https://doi.org/10.1007/JHEP04\(2015\)025](https://doi.org/10.1007/JHEP04(2015)025).
- [123] ATLAS Collaboration, G. Aad et al., *Search for high-mass dilepton resonances in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector*, Phys. Rev. D **90** (Sep, 2014) 052005. <https://link.aps.org/doi/10.1103/PhysRevD.90.052005>.
- [124] CMS Collaboration, V. Khachatryan et al., *Search for narrow resonances in dilepton mass spectra in protonproton collisions at $\sqrt{s} = 13 \text{ TeV}$ and combination with 8 TeV data*, Physics Letters B **768** (2017) no. Supplement C, 57 – 80.
<http://www.sciencedirect.com/science/article/pii/S0370269317301004>.

- [125] CMS Collaboration, A. M. Sirunyan et al., *Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, arXiv:1803.06292 [hep-ex].
- [126] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostoja, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. <https://cds.cern.ch/record/815187>.
- [127] CERN, *The accelerator complex*, Jan, 2012.
<https://cds.cern.ch/record/1997193>. Last Accessed Mar 2018.
- [128] *LHC Guide*, Mar, 2017. <https://cds.cern.ch/record/2255762>.
- [129] CERN, *Radiofrequency cavities*, Sep, 2012. <https://cds.cern.ch/record/1997424>.
Last Accessed Mar 2018.
- [130] Xabier Cid Vidal and Ramon Cid Manzano, *RF Cavities*, 2017.
https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.rf_cavities.
Last Accessed Mar 2018.
- [131] CERN, *Pulling together: Superconducting electromagnets*, Aug, 2012.
<https://cds.cern.ch/record/1997395>. Last Accessed Mar 2018.
- [132] X. C. Vidal and R. C. Manzano, *Buckets and bunches*, 2017. https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.buckets_and_bunches.
Last Accessed Mar 2018.
- [133] M. Lamont, *Status of the LHC*, Journal of Physics: Conference Series **455** (2013) no. 1, 012001. <http://stacks.iop.org/1742-6596/455/i=1/a=012001>.
- [134] G. Papotti et al., *Operation of the LHC with Protons at High Luminosity and High Energy*, in *Proceedings, 7th International Particle Accelerator Conference (IPAC 2016): Busan, Korea, May 8-13, 2016*, p. WEOCA01. 2016.
<http://inspirehep.net/record/1470111/files/weoca01.pdf>.
- [135] Corinne Pralavorio, *LHC performance reaches new highs*, Jul, 2016.
<https://cds.cern.ch/record/2203203>. Last Accessed Mar 2018.
- [136] Frédéric Bordry and Corinne Pralavorio, *LHC smashes targets for 2016 run*, Nov, 2016. <https://cds.cern.ch/record/2235979>. Last Accessed Mar 2018.
- [137] Y. Nakahama, *The ATLAS Trigger System: Ready for Run-2*, Journal of Physics: Conference Series **664** (2015) no. 8, 082037.
<http://stacks.iop.org/1742-6596/664/i=8/a=082037>.

- [138] M. zur Nedden, *The LHC Run 2 ATLAS trigger system: design, performance and plans*, Journal of Instrumentation **12** (2017) no. 03, C03024.
<http://stacks.iop.org/1748-0221/12/i=03/a=C03024>.
- [139] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007. <https://cds.cern.ch/record/1020106>.
- [140] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, Tech. Rep. ATLAS-CONF-2012-042, CERN, Geneva, Mar, 2012.
<https://cds.cern.ch/record/1435196>.
- [141] R. Achenbach et al., *The ATLAS Level-1 Calorimeter Trigger*, Journal of Instrumentation **3** (2008) no. 03, P03001.
<http://stacks.iop.org/1748-0221/3/i=03/a=P03001>.
- [142] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008.
<https://cds.cern.ch/record/1099735>.
- [143] ATLAS Collaboration, G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, The European Physical Journal C **74** (Oct, 2014) 3071. <https://doi.org/10.1140/epjc/s10052-014-3071-4>.
- [144] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014.
<https://cds.cern.ch/record/1706245>.
- [145] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2016-015, CERN, Geneva, Aug, 2016.
<https://cds.cern.ch/record/2203514>.
- [146] R. Turra, *ElectronPhotonFourMomentumCorrection*, 2017.
<https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronPhotonFourMomentumCorrection>. Last Accessed Mar 2018.

- [147] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Höche, H. Hoeth, F. Krauss, L. Lönnblad, E. Nurse, P. Richardson, S. Schumann, M. H. Seymour, T. Sjöstrand, P. Skands, and B. Webber, *General-purpose event generators for LHC physics*, Physics Reports **504** (2011) no. 5, 145 – 233.
<http://www.sciencedirect.com/science/article/pii/S0370157311000846>.
- [148] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, Computer Physics Communications **191** (2015) 159 – 177.
<http://www.sciencedirect.com/science/article/pii/S0010465515000442>.
- [149] T. Gleisberg, S. Höche, F. Krauss, M. Schönher, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, Journal of High Energy Physics **2009** (2009) no. 02, 007. <http://stacks.iop.org/1126-6708/2009/i=02/a=007>.
- [150] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, Journal of High Energy Physics **2010** (2010) no. 6, 43.
[http://dx.doi.org/10.1007/JHEP06\(2010\)043](http://dx.doi.org/10.1007/JHEP06(2010)043).
- [151] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Computer Physics Communications **185** (2014) no. 11, 2930 – 2938.
<http://www.sciencedirect.com/science/article/pii/S0010465514002264>.
- [152] C. Anastasiou, L. Dixon, K. Melnikov, and F. Petriello, *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to-leading order*, Phys. Rev. D **69** (May, 2004) 094008.
<https://link.aps.org/doi/10.1103/PhysRevD.69.094008>.
- [153] S. G. Bondarenko and A. A. Sapronov, *NLO EW and QCD proton–proton cross section calculations with mcsanc-v1.01*, Computer Physics Communications **184** (2013) no. 10, 2343 – 2350.
<http://www.sciencedirect.com/science/article/pii/S0010465513001707>.
- [154] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, The European Physical Journal C - Particles and Fields **45** (Jan, 2006) 97–107. <http://dx.doi.org/10.1140/epjc/s2005-02396-4>.
- [155] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, The European Physical Journal C **70** (Dec, 2010) 823–874.
<https://doi.org/10.1140/epjc/s10052-010-1429-9>.

- [156] ATLAS Collaboration, G. Duckeck, D. Barberis, R. Hawkings, R. Jones, N. McCubbin, G. Poulard, D. Quarrie, T. Wenaus, and E. Obreshkov, *ATLAS Computing: technical design report*. Technical Design Report ATLAS. CERN, Geneva, 2005. <https://cds.cern.ch/record/837738>.
- [157] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) no. 3, 250 – 303.
<http://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [158] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo Event Record for High Energy Physics*, Tech. Rep. ATL-SOFT-2000-001, CERN, Geneva, Jun, 2000.
<https://cds.cern.ch/record/684090>.
- [159] ATLAS Collaboration, *Comparison of the response of the ATLAS detector to electromagnetic processes in data at 8 TeV and simulation using different G4 setups*, Tech. Rep. ATL-PHYS-PUB-2014-003, CERN, Geneva, Mar, 2014.
<https://cds.cern.ch/record/1669338>.
- [160] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, *New parton distributions for collider physics*, Phys. Rev. D **82** (Oct, 2010) 074024.
<https://link.aps.org/doi/10.1103/PhysRevD.82.074024>.
- [161] ATLAS Collaboration, G. Aad et al., *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Journal of High Energy Physics **2014** (Sep, 2014) 145.
[https://doi.org/10.1007/JHEP09\(2014\)145](https://doi.org/10.1007/JHEP09(2014)145).
- [162] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, *Parton distributions incorporating QED contributions*, The European Physical Journal C - Particles and Fields **39** (Feb, 2005) 155–161. <http://dx.doi.org/10.1140/epjc/s2004-02088-7>.
- [163] P. Z. Skands, *Tuning Monte Carlo generators: The Perugia tunes*, Phys. Rev. D **82** (Oct, 2010) 074018. <https://link.aps.org/doi/10.1103/PhysRevD.82.074018>.
- [164] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. D. Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, *Parton distributions with LHC data*, Nuclear Physics B **867** (2013) no. 2, 244 – 289.
<http://www.sciencedirect.com/science/article/pii/S0550321312005500>.
- [165] ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes*, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014.
<https://cds.cern.ch/record/1966419>.

- [166] S. Laplace, *ElectronGamma*, 2017.
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronGamma>. Last Accessed Mar 2018.
- [167] K. Lohwasser, *XAODElectronEfficiencyCorrectionTool*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/XAODElectronEfficiencyCorrectionTool>. Last Accessed Mar 2018.
- [168] K. Koeneke, *PileupReweighting*, 2017. <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PileupReweighting>. Last Accessed Mar 2018.
- [169] M. Zinser, *LPXKfactorTool*, 2017.
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/LPXKfactorTool>. Last Accessed Mar 2018.
- [170] ATLAS Collaboration, M. Aaboud et al., *Search for high-mass new phenomena in the dilepton final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **761** (Jul, 2016) 372–392. 35 p.
<https://cds.cern.ch/record/2198982>.
- [171] C. Willis, D. Hayden, R. Brock, and W. Fedorko, *Constructing an Inclusive Lepton Early Warning System*, Tech. Rep. ATL-COM-PHYS-2015-576, CERN, Geneva, Jun, 2015. <https://cds.cern.ch/record/2027501>.
- [172] R. Nikolaïdou, *GoodRunListsForAnalysisRun2*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/GoodRunListsForAnalysisRun2>. Accessed Feb, 2018.
- [173] E. Gallas, *ConditionsMetadata*, 2017. <https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/ConditionsMetadata>. Accessed Feb, 2018.
- [174] L. W. Group, *LuminosityForPhysics*, 2017.
<https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics>. Accessed Feb, 2018.
- [175] R. Brun and F. Rademakers, *ROOT – An object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **389** (1997) no. 1, 81 – 86.
<http://www.sciencedirect.com/science/article/pii/S016890029700048X>. New Computing Techniques in Physics Research V.

- [176] J. Catmore, J. Cranshaw, T. Gillam, E. Gramstad, P. Laycock, N. Ozturk, and G. A. Stewart, *A new petabyte-scale data derivation framework for ATLAS*, Journal of Physics: Conference Series **664** (2015) no. 7, 072007.
<http://stacks.iop.org/1742-6596/664/i=7/a=072007>.
- [177] T. J. Khoo, *DerivationFramework*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/DerivationFramework>. Accessed Mar, 2018.
- [178] F. Phillips, *Welcome to the ATLAS Tier3 web*, 2017.
https://hep.pa.msu.edu/wiki/bin/view/ATLAS_Tier3/WebHome. Accessed Mar, 2018.
- [179] A. Krasznahorkay, *RootCore*, 2017.
<https://twiki.cern.ch/twiki/bin/view/AtlasComputing/RootCore>. Accessed Mar, 2018.
- [180] B. Carlson, *SampleHandler*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisSoftwareGroup>. Accessed Mar, 2018.
- [181] A. Madsen, *EventLoop*, 2017.
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EventLoop>. Accessed Mar, 2018.
- [182] J. Nielsen, *Analysis Releases*, 2017.
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisRelease>. Accessed Mar, 2018.
- [183] L. Heelan, *ATLAS Analysis Software Group (ASG)*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisSoftwareGroup>. Accessed Mar, 2018.
- [184] C. Willis, *LPXSignalReweightingTool*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/LPXSignalReweightingTool>. Accessed Mar, 2018.
- [185] N. Ilic, *Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade*, Journal of Instrumentation **9** (2014) no. 03, C03049. <http://stacks.iop.org/1748-0221/9/i=03/a=C03049>.
- [186] ATLAS Collaboration, G. Aad et al., *Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector*, JHEP **09** (Jul, 2014) 037. 28 p.
<http://cds.cern.ch/record/1746306>.

- [187] ATLAS Collaboration, M. Aaboud et al., *Search for a new heavy gauge boson resonance decaying into a lepton and missing transverse momentum in 36 fb^{-1} of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS experiment*, Tech. Rep. CERN-EP-2017-082, CERN, Geneva, Jun, 2017.
<http://cds.cern.ch/record/2269048>.
- [188] ATLAS Collaboration, G. Aad et al., *Search for new phenomena in the dijet mass distribution using pp collision data at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector*, Phys. Rev. D **91** (Mar, 2015) 052007.
<https://link.aps.org/doi/10.1103/PhysRevD.91.052007>.
- [189] ATLAS Collaboration, G. Aad et al., *Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector*, Physics Letters B **754** (2016) no. Supplement C, 302 – 322.
<http://www.sciencedirect.com/science/article/pii/S0370269316000447>.
- [190] E. Benhar, *Electron Efficiencies for Run2 (scale factors and uncertainties)*, 2017.
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronEfficiencyRun2>. Accessed Feb, 2018.
- [191] R. Turra, *Electron and Photon Calibration for Run2 (energy scale and resolution)*, 2017. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaCalibrationRun2>. Accessed Feb, 2018.
- [192] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** (Feb, 2013) 2518. 40 p. [https://cds.cern.ch/record/1517411](http://cds.cern.ch/record/1517411).
- [193] ATLAS Collaboration, M. Aaboud et al., *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (Oct, 2016) 182002.
<https://link.aps.org/doi/10.1103/PhysRevLett.117.182002>.
- [194] J. Gao and P. Nadolsky, *A meta-analysis of parton distribution functions*, Journal of High Energy Physics **2014** (Jul, 2014) 35.
[https://doi.org/10.1007/JHEP07\(2014\)035](https://doi.org/10.1007/JHEP07(2014)035).
- [195] J. Butterworth, S. Carrazza, A. Cooper-Sarkar, A. D. Roeck, J. Feltesse, S. Forte, J. Gao, S. Glazov, J. Huston, Z. Kassabov, R. McNulty, A. Morsch, P. Nadolsky, V. Radescu, J. Rojo, and R. Thorne, *PDF4LHC recommendations for LHC Run II*, Journal of Physics G: Nuclear and Particle Physics **43** (2016) no. 2, 023001.
<http://stacks.iop.org/0954-3899/43/i=2/a=023001>.

- [196] P. Motylinski, L. Harland-Lang, A. D. Martin, and R. S. Thorne, *Updates of PDFs for the 2nd LHC run*, Nuclear and Particle Physics Proceedings **273** (2016) no. Supplement C, 2136 – 2141.
<http://www.sciencedirect.com/science/article/pii/S2405601415008366>.
- [197] I. Connelly, *Systematics affecting Top MC for 13 TeV 2015 analyses (rel19)*, 2017.
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopSystematics2015>. Accessed Mar, 2018.
- [198] C. Willis, *The LPXSignalReweightingTool: A Package for Signal Template Reweighting in High-Energy Physics Analysis*, Tech. Rep. ATL-COM-PHYS-2016-912, CERN, Geneva, Jul, 2016.
<https://cds.cern.ch/record/2199258>.
- [199] G. Choudalakis, *On hypothesis testing, trials factor, hypertests and the BumpHunter*, in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva, Switzerland 17-20 January 2011*. 2011. arXiv:1101.0390 [physics.data-an].
<https://inspirehep.net/record/883244/files/arXiv:1101.0390.pdf>.
- [200] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character **231** (1933) 289–337. <http://www.jstor.org/stable/91247>.
- [201] ROOT Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, 2012.
- [202] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv:1009.1003 [physics.data-an].
- [203] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116 [physics].
- [204] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statist. **9** (03, 1938) 60–62.
<https://doi.org/10.1214/aoms/1177732360>.
- [205] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an].

- [206] T. Bayes, *LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton*, A. M. F. R. S, Philosophical Transactions **53** (1763) 370–418.
<http://rstl.royalsocietypublishing.org/content/53/370.short>.
- [207] A. Caldwell, D. Kollár, and K. Kröninger, *BAT – The Bayesian analysis toolkit*, Computer Physics Communications **180** (2009) no. 11, 2197 – 2209.
<http://www.sciencedirect.com/science/article/pii/S0010465509002045>.
- [208] ATLAS Collaboration, G. Aad et al., *Search for high-mass resonances decaying to dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Journal of High Energy Physics **2012** (Nov, 2012) 138.
[https://doi.org/10.1007/JHEP11\(2012\)138](https://doi.org/10.1007/JHEP11(2012)138).
- [209] ATLAS Collaboration, G. Aad et al., *Search for high mass dilepton resonances in pp collisions at $s=7$ TeV with the ATLAS experiment*, Physics Letters B **700** (2011) no. 3, 163 – 180.
<http://www.sciencedirect.com/science/article/pii/S0370269311004436>.
- [210] J. C. Collins and D. E. Soper, *Angular distribution of dileptons in high-energy hadron collisions*, Phys. Rev. D **16** (Oct, 1977) 2219–2225.
<https://link.aps.org/doi/10.1103/PhysRevD.16.2219>.
- [211] C. Schmidt. Personal communication, Apr, 2018.
- [212] ATLAS Collaboration, G. Aad et al., *Measurement of the high-mass Drell-Yan differential cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Physics Letters B **725** (2013) no. 4, 223 – 242.
<http://www.sciencedirect.com/science/article/pii/S0370269313006138>.
- [213] ATLAS Collaboration, G. Aad et al., *Precision measurement and interpretation of inclusive W^+ , W^- , and Z/γ^* production cross sections with the ATLAS detector*, The European Physical Journal C **77** (Jun, 2017) 367.
<https://doi.org/10.1140/epjc/s10052-017-4911-9>.
- [214] H. Tie-Jiun. Personal communication, Apr, 2018.
- [215] C. Pralavorio, *Record luminosity: well done LHC*, 2017.
<https://cds.cern.ch/record/2295027>. Last Accessed Apr 2018.
- [216] ATLAS Collaboration, G. Aad et al., *Search for contact interactions and large extra dimensions in the dilepton channel using proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Eur. Phys. J. **C74** (2014) no. 12, 3134, arXiv:1407.2410 [hep-ex].

- [217] ATLAS Collaboration, T. Hrynev'ova, D. Hayden, T. Berry, P. J. Falke, Y. Takubo, M. M. Perez-Victoria Moreno De Barreda, E. Kay, U. Klein, C. Gwilliam, and M. K. Bugge, *Combination of $Z' \rightarrow \ell\ell$ and $W' \rightarrow \ell\nu$ in Heavy Vector Triplet model with 36 fb^{-1} of 13 TeV data*, Tech. Rep. ATL-COM-PHYS-2015-021, CERN, Geneva, Jan, 2015. <https://cds.cern.ch/record/1982245>. Other Authors: Manuel Perez-Victoria and Javier Lizana.
- [218] ATLAS Collaboration, O. AbouZeid, V. Cavaliere, F. De Lorenzi, W. C. Fisher, G. Halladjian, V. Pascuzzi, A. Picazio, D. Denysiuk, K. Iordanidou, I. Ochoa, S. Hassani, and L. L. Ma, *VV/VH combination*, Tech. Rep. ATL-COM-PHYS-2017-986, CERN, Geneva, Jun, 2017. <https://cds.cern.ch/record/2272170>.
- [219] CERN, *The High-Luminosity LHC*, Nov, 2015. <https://cds.cern.ch/record/2114693>. Last Accessed Apr 2018.
- [220] A. Martyniuk, *Atlantis Event Display*, 2018. <https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/Atlantis>. Last Accessed Apr 2018.
- [221] ATLAS Collaboration, G. Aad et al., *Search for a Heavy Neutral Particle Decaying to $e\mu$, $e\tau$, or $\mu\tau$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, Phys. Rev. Lett. **115** (2015) no. 3, 031801, arXiv:1503.04430 [hep-ex].
- [222] CERN, *The Grid: A system of tiers*, 2012. <https://cds.cern.ch/record/1997396>. Last Accessed Apr 2018.