

**MEASURING W BOSON DRELL–YAN ANGULAR COEFFICIENTS AND
DIFFERENTIAL CROSS-SECTION WITH THE ATLAS DETECTOR USING A
LOW–PILEUP DATA SET**

By

Alexander W. Bachiu, BSc., MSc.

A thesis submitted to the Office of Graduate Studies
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy,
Particle Physics

Ottawa–Carleton Institute for Physics
Department of Physics
Carleton University
Ottawa, Ontario

© 2025
Alexander W. Bachiu, BSc., MSc.

ABSTRACT

The W and Z bosons, mediating the fundamental weak force, are some of the particles produced from the proton–proton collisions in the Large Hadron Collider (LHC), located at CERN. A proton–proton interaction producing a W boson that decays immediately into leptonic particles, such as an electron and a neutrino, is known as the Drell–Yan process. The differential cross-section of this process expresses how likely it is to occur depending on the kinematics of the W boson – such as the transverse momentum, p_T^W – and the kinematics of the leptons. This differential cross-section can be decomposed into eight spin-related ratios and an unpolarized term. The eight spin-related ratios are known as the Drell–Yan angular coefficients. The angular coefficients and the unpolarized term are coupled to angular polynomials of $\cos \theta$ and ϕ , which describe the angles between the leptons and the W boson polarization. By measuring the angular distributions of the decay products, the angular coefficients that describe the production of the W boson can be extracted. These angular coefficients are important as they are a direct probe of perturbative Quantum Chromodynamics production of W bosons from proton–proton interactions, making these coefficients an important input to other precision measurements, such as the W boson mass. This thesis presents a measurement of the four dominant angular coefficients – A_0 , A_2 , A_3 , and A_4 – and the production differential cross-section as a function of p_T^W by reconstructing W boson events to measure the $\cos \theta$ and ϕ distributions. The measurements were made possible by utilizing a data set of special runs with fewer interactions per bunch crossing, measured with the ATLAS detector at a centre-of-mass energy $\sqrt{s} = 13$ TeV. Included are the first measurements at the LHC of the W boson coefficients A_2 and A_3 as a function of p_T^W and the first measurements ever of A_0 and A_4 as function of p_T^W .

ACKNOWLEDGEMENTS

This thesis was made possible through the incredible support of all the people in my life. I'd first like to thank my wife, Miranda, and my dog, Potato, for their continued support throughout my entire PhD. program. Their emotional and motivational support made it possible for me to finish this thesis and my degree. I'd also like to thank my parents Chris, Lynda, Pat, Brad, and Francine for all of their support and help throughout my life to get me to this point. To my friends, Shayne, Ben, Coulter, Dylan, Amanda, Jordan, and many others, thank you for enriching my life and being such good friends to get me where I am today. I'd also like to thank all my friends throughout the years in the ATLAS office at Carleton University for making the office such an enjoyable environment.

I'd also like to thank everyone in the ATLAS Collaboration, past and present, that has made this analysis possible, as physics on this scale is only possible through the great effort of many people. Within ATLAS, my supervisor, Manuella, has been the biggest contributor in making these measurements possible by pushing and helping me to become the physicist I am today. Her support through the process, beginning with bringing the idea of making the measurements for my thesis to her, is invaluable. I could not make this measurement alone. My thanks go out to Ludovica, Aaron, Daniil, Ruth, Stefano, Fabrice, and Grigorii, who have worked alongside me to make this possible. Lastly, I'd also like to thank my ATLAS colleagues, Nataliia, Maarten, and Stefan, for their help and guidance through my authorship qualification task.

TABLE OF CONTENTS

Chapter 1: Introduction	1
1.1 Personal Contributions	8
1.2 PhD. Program Experiences	9
Chapter 2: The Standard Model	10
2.1 The Standard Model	10
2.2 Quantum Field Theory	14
2.2.1 Gauge Theories	18
2.2.2 Electroweak Theory and Spontaneous Symmetry Breaking	19
2.2.3 Electroweak Interactions	25
2.2.4 Strong Interactions	27
2.2.5 Summary	29
Chapter 3: LHC and the ATLAS Detector	30
3.1 LHC	30
3.2 Counting Interactions	33
3.3 ATLAS Detector	34
3.3.1 Inner Detector	37
3.3.2 Calorimeters	39

3.3.3	Muon Spectrometer	43
3.3.4	Forward Detectors	45
3.3.5	Triggering	46
3.4	Summary	47
Chapter 4: Particle Reconstruction		48
4.1	Efficiencies and Scale Factors	48
4.2	Tag and Probe Method	50
4.3	Reconstruction	51
4.3.1	Track Reconstruction	52
4.3.2	Electron Reconstruction	53
4.3.3	Photon Reconstruction	54
4.3.4	Muon Reconstruction	55
4.3.5	Hadronic Reconstruction	55
4.4	Particle Identification	56
4.4.1	Electron Identification	57
4.4.2	Muon Identification	64
4.5	Particle Isolation	65
4.6	Particle Trigger	68
Chapter 5: Electron Identification Variable Correction		69
5.1	Variable Correction Factors	70
5.2	Extraction of Variable Correction Factors	74
5.3	Shower Shape Reweighting	76

5.4	eProbabilityHT Transformation	78
5.5	Pear Shaped Correction	80
5.6	Sample Background Consideration	82
5.7	Efficiency and Scale Factor Calculation	85
Chapter 6: Drell–Yan Process		91
6.1	Drell–Yan Angular Coefficients	95
6.2	Collins–Soper Frame	99
6.3	Coefficients in CS Frame	101
6.4	Summary	103
Chapter 7: Measurement Methodology		104
7.1	Moments Method	105
7.2	Building Templates	106
7.3	Template Folding	111
7.4	Likelihood Method	113
Chapter 8: Data, Simulation, and Event Reconstruction		116
8.1	Data Set	116
8.2	MC Sets	118
8.2.1	POWHEG+PYTHIA8	119
8.2.2	SHERPA	119
8.3	Event Selection	120
8.4	Neutrino Reconstruction	121

8.4.1	Solution Choice	127
8.5	W Boson Reconstruction	129
8.6	Data Driven Background Estimation	132
8.7	Kinematic Distributions	133
Chapter 9: Measurement of W Boson Drell–Yan Angular Coefficients and Differential Cross-Section		145
9.1	Single Channel Results	146
9.2	Channel Compatibility	154
9.3	Channel Combination	156
9.4	Comparison to PDF Sets	161
9.5	Summary	164
Chapter 10: Statistical and Systematic Uncertainties on the Measurement of the Drell-Yan Angular Coefficients and Differential Cross-Section		165
10.1	Data Statistical Uncertainties	166
10.1.1	Migration Uncertainty	168
10.1.2	Cutoff Uncertainty	169
10.2	Systematic Uncertainties	171
10.2.1	Experimental Uncertainties	172
10.2.2	Theoretical Uncertainties	182
10.3	Summary of Uncertainties	184
10.4	Constraint of Uncertainties	186
Chapter 11: Combined Channel Measurement of the W Boson Drell–Yan Angular Coefficients and Differential Cross-Section		189

11.1	Measurements of Drell–Yan Angular Coefficients and Differential Cross-Section	189
11.2	Uncertainty Decomposition of the Drell–Yan Angular Coefficients and Differential Cross-Section	193
11.3	Potential Impact on W Mass	196
Chapter 12: Conclusions and Outlook		198
Appendix A: List of Acronyms, Abbreviations, and Common Variables		219
A.1	Acronyms and Abbreviations	219
A.2	Common Variables	222
Appendix B: Spherical Harmonics		224
B.1	Associated Quantum Numbers	224
Appendix C: Moment Derivation		226
Appendix D: Folding		229
Appendix E: Monte Carlo Samples		231
Appendix F: Sign Ambiguity		233
Appendix G: Additional Figures		237
G.1	Additional W^- Electron Channel Figures	237
G.2	Additional W^+ Electron Channel Figures	242
G.3	Additional W^- Muon Channel Figures	255
G.4	Additional W^+ Muon Channel Figures	268

G.5	Additional W^- Combined Channels Figures	281
G.6	Additional W^+ Combined Channels Figures	286

LIST OF TABLES

2.1	Table summarizing the charge and mass of known fermions, grouped by leptons and quarks, and further sub-grouped by generation based on mass. Mass values are in natural units, where $c = 1$, and charges in units of the elementary charge, e . The neutrinos' masses are much smaller than the other fermions and are not included in this table, as they can be treated as massless in most cases. Values taken from Particle Data Group (PDG) [25].	12
2.2	Table of known bosons where mass values are in natural units, where $c = 1$, and charges in units of the elementary charge, e . Values from Particle Data Group (PDG) [25]. The graviton has not been observed and would be the boson of a field theory that describes gravity with a spin ($J = 2$).	13
4.1	List of tracking variables with a brief description that are used in the electron identification LH to determine if events are electrons or background. A more detailed use of each variable is described in [66] and a depiction of some can be found in Figure 4.7.	59
4.2	List of shower shape variables with a brief description that are used in the electron identification LH to determine if events are electrons or background. A more detailed use of each variable is described in [66] and a depiction can be seen in Figure 4.8.	60
6.1	Summary of Drell–Yan angular coefficients, polynomials coupled to each coefficient, how the polynomials relate to the spherical harmonics, the vector and axial–vector couplings – which for the W boson is equal to the third component of the weak isospin $T^3 = \pm 1/2$ –, and the order of α_S when they become non-zero.	99
8.1	A comparison between the high– and low– $\langle\mu\rangle$ requirements for an electron or muon to pass the L1 hardware and HLT software triggers, which shows the lower thresholds used for the low– $\langle\mu\rangle$ run conditions.	118

8.2	List of signal and dominant background processes that are generated using POWHEG+PYTHIA8. A more complete list can be found in Appendix E and contains more information about each process.	120
8.3	List of identification, kinematic selection, and isolation criteria to ensure proper data quality and to select a high purity sample of candidate signal events to be used in this analysis.	121
9.1	Summary of χ^2 , ndf , and p-values for compatibility between the number of observed events and the number of expected events after the fit for all channels as a function of $p_T^{l\nu}$	149
9.2	Summary of χ^2 , ndf , and p-values for compatibility between measured Drell–Yan angular coefficients and the different PDF set predictions in Figure 9.9 and CT10 – the Asimov data set without any PDF reweighting applied. The p-values demonstrate that the measured values are in good compatibility with all PDF set predictions.	163
10.1	Description of the channel independent hadronic recoil systematics NPs used in this analysis. Each NP will have an up and down variation with the exception of SET_Sys which a one-sided up variation.	174
10.2	Description of the background related systematics NPs used in this analysis. Each NP will have an up and down variation except for the $t\bar{t}$ hdamp variations, which are one-sided.	177
10.3	Description of the electron channel only systematics NPs used in this analysis. Each NP has an up and down variation.	179
10.4	Description of the muon channel only systematics NPs used in this analysis. Each NP has an up and down variation.	182
10.5	Description of the PDF systematic uncertainty nuisance parameters used in this analysis. Each NP will have an up and down variation.	183
11.1	Summary of χ^2/ndf and p-values measuring the compatibility between the number of observed events and the number of expected events after the fit for the measurements using combined channels as a function of $p_T^{l\nu}$	190
B.1	The number of angular eigenfunctions needed to describe the angular decay angles for boson of spin $J = 0, 1, 2$ and the associated quantum numbers. The quantum numbers ℓ and m are used for the angular polynomials.	225

E.1 The Monte Carlo (MC) samples at $\sqrt{s} = 13$ TeV for the low- $\langle\mu\rangle$ data set. Given is: a short description of the process, the names and version numbers of the MC generator(s), the used value of the higher order cross section times any branching and filter efficiencies ($\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$) – with the theoretical uncertainty in percent (“unc.”) –, the number of events analyzed that pass the trigger threshold – $N_{\text{evt}}^{\text{skim}}$ –, and the number of originally processed and simulated events – $N_{\text{evt}}^{\text{unskim}}$. In the case of $Z \rightarrow \ell\ell$ samples, there is $\epsilon_{\text{filter}} > 1$ due to the fact the cross sections were calculated for $66 < m^{\ell\ell} < 116$ GeV, but the generated mass range is larger. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}}/(\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}})$ is generally above 3 fb^{-1} for signal and significant backgrounds. The exception are POWHEG+PYTHIA8 $W \rightarrow \tau\nu$ and $Z \rightarrow \tau^+\tau^-$ samples, that have about 0.45 fb^{-1} only. 232

LIST OF FIGURES

1.1	The measured values for the W boson Drell–Yan angular coefficients at $\sqrt{s} = 1.8$ TeV as a function of p_T^W by the CDF Collaboration [10]. Here, A_2 and A_3 were measured while A_0 and A_4 are consistent with zero within their uncertainty. The dashed line is the SM prediction of the angular coefficients.	3
1.2	Schematic of the ATLAS detector with labeled subdetector and magnet locations [13]. The Cartesian coordinate system used by ATLAS includes: the z -axis is along the beamline, the x -axis points towards the centre of the LHC ring, and the y -axis consequentially points upwards. Here p_T , is the momentum in the transverse plane. A more detailed coordinate system and description of the ATLAS detector will be discussed in Chapter 3.	5
1.3	Comparison of the more recent m^W measurements compiled by the CMS Collaboration [17]. The measurement group and citation are listed on the left and correspond to each measurement point listed on an MeVscale. . . .	6
2.1	An example of a Feynman diagram for a quartic coupling, where time flows left to right. This diagram has four fields meeting at a vertex and shows the time order of physics processes but does not contain any space-time information.	16
2.2	Example Feynman diagrams for cubic couplings, where time flows left to right. a) Shows three fields meeting at a vertex where two particles annihilate to produce the third field, which decays into two particles at another three field vertex. b) Shows two particles scattering, which is mediated by another particle. Each scattering particle has their incoming and outgoing field lines forming a cubic vertex with the mediating particle.	17
2.3	A depiction of the Higgs potential taken from [31]. The Higgs potential takes the famous sombrero shape and the symmetry is broken when a vacuum is chosen, which results in a minimum at the vacuum expectation value.	23

2.4	Feynman diagram of nuclear beta decay, where a down quark in a neutron emits a W^- boson to change flavour into an up quark. The W^- boson then promptly decays into an electron and an anti-neutrino. This W^- boson emission is an example of one of the weak charged currents.	26
2.5	Feynman diagram of the Drell–Yan process of quark-antiquark pair annihilation into a Z boson before decaying into an electron-positron pair.	26
2.6	Feynman diagram of a quark-antiquark pair annihilation into a photon before decaying into an electron-positron pair.	26
2.7	Running strong coupling “constant”, α_s , which decreases with increased momentum transfer, Q , and separation distance. Plot taken from Particle Data Group (PDG) [25].	28
2.8	An example of a Feynman diagram where a gluon is emitted within a proton to form sea quarks before being reabsorbed.	28
3.1	Schematic overview of the CERN complex [36], including the LHC. Proton–proton collisions are produced by ionizing hydrogen to isolate the protons before being accelerated through the path LINAC4, BOOSTER, PS, SPS, and LHC. Beams are then collided in experiments, such as ATLAS.	31
3.2	The recorded luminosity vs the mean number of interactions per bunch crossing, μ , taken with the ATLAS detector for Run 1 in orange, Run 2 in green, and the start of Run 3 in purple [46]. For each run, the average of $\mu - \langle \mu \rangle$ – and the most probable value – μ_{MPV} – are listed, rounding to the closest whole value.	35
3.3	Schematic of the ATLAS detector with labeled subdetector and magnet locations [13]. From inside out: the inner detector, the calorimeters, and the muon spectrometer. The Cartesian coordinate system used by ATLAS includes: the z –axis is along the beamline, the x –axis points towards the centre of the LHC ring, and the y –axis consequentially points upwards. The origin is at the centre of the detector but is displaced here for readability. Other coordinates used by ATLAS are r , θ , η , and ϕ , where r is the radius in the xy –plane, θ is the polar angle from the z –axis, $\eta = -\ln [\tan (\theta/2)]$ is the pseudorapidity, and ϕ is the azimuthal angle measured from the x –axis. The transverse momentum, p_T , is in the transverse plane along r	36

3.4	Cut away view of the ATLAS Inner Detector [50]. Closest to the beamline are the Insertable B-Layer (IBL) [51] and the three layers of Pixel detector. Outside of that are the four layers of Semiconductor Tracker (SCT). Lastly, is the Transition Radiation Tracker (TRT), which is constructed out of straws with polypropylene fibers between them.	37
3.5	The schematic diagram of the calorimeter systems in the ATLAS detector [58]. On the inside are the electromagnetic calorimeter (ECAL) is electromagnetic barrel (EMB), electromagnetic endcap (EMEC), and the first layer of the Forward Calorimeter (FCal), which all use liquid argon (LAr). Outside of the ECAL is the hadronic calorimeter (HCAL) comprised of the Tile barrel and extended barrel calorimeter, LAr hadronic endcap (HEC), and the last two layers of the FCal.	40
3.6	Sketch of the cells within the LAr barrel, which lists the $\eta \times \phi$ granularity of each layer. The side cutaway view shows the accordion shape of the layers [1].	42
3.7	Schematic of the muon spectrometer (MS) [61]. The barrel is comprised of Monitored Drift Tubes (MDT), Resistive Plate Chambers (RPC), and the barrel toroid magnets. For each endcap, there is the smaller wheel of Cathode Strip Chambers (CSC) and Thin Gap Chambers (TGC), the endcap toroid magnets, and the larger two wheels, comprised of MDTs and TGCs.	44
4.1	Cartoon example of the Tag&Probe method in the ATLAS detector. The tag lepton meets stringent criteria, while the probe lepton is more loosely selected and is used to determine efficiencies.	51
4.2	A schematic slice of the ATLAS detector in the main barrel region to show the signatures of various particles in the detector [69]. The charged particles, such as electrons, protons, and muons, will leave tracks in the ID. Electrons and photons deposit their energy in the ECAL, while hadronic particles, such as neutrons and protons, will cause showers in the HCAL and some energy is deposited in the ECAL. Muons deposit some energy in the calorimeters but are mostly measured by the MS. Neutrinos do not leave any signature within the ATLAS detector.	52
4.3	Clusters around the seed cluster in an area of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ are added to form a supercluster. Clusters with the same best-matched track to the seed cluster in an area of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ are also added to the supercluster. Cropped image from [68].	54

4.4	From an interaction at a parton level, quarks and gluons are produced. Instead of bare partons, hadrons are formed, which deposit energy in the calorimeter cells in a conical shape, called a jet.	56
4.5	Example of the path traveled by an electron passing through the ATLAS detector from [66]. The solid red line shows the path through each layer of the ID before entering the ECAL. The dashed red line shows the path of a photon if it were produced from the electron interacting with the ID.	57
4.6	Example probability distribution function for R_ϕ for both the signal and background in the region of $20 < E_T < 30$ GeV and $0.6 < \eta < 0.8$ from [66]. There is a discernible difference between signal and background shape, so it is used as part of a LH to discriminate the difference, where it would be inefficient to use a selection-based discrimination.	58
4.7	Depiction of some of the tracking variables used to create probability distribution functions for the electron identification LH. The left image depicts the track geometry near the beam spot to define the transverse impact parameter – d_0 , the perigee, which is used for $\Delta\phi_{\text{rescaled}}$ and $\Delta P/P$, and the longitudinal impact parameter – z_0 , which is not used in the electron identification LH. The image on the right is based on an image from [66], shown previously in Figure 4.5, and shows $\Delta P/P$, eProbabilityHT, $\Delta\eta_1$, and $\Delta\phi_{\text{rescaled}}$	61
4.8	Depiction of the shower shape variables, listed in Table 4.2, used to create probability distribution functions for the electron identification LH. On the left are the ratios of energy between different regions of the calorimeter. On the right are the variables used for the shape of the shower in the calorimeter. Figure inspired by a similar figure in [75].	62
4.9	The measured LH electron identification efficiencies using $Z \rightarrow ee$ events as a function of a) E_T and b) η from [66]. The three working points: <i>Loose</i> (blue circles), <i>Medium</i> (red squares), and <i>Tight</i> (black triangles) correspond to different levels of background rejection. The <i>Tight</i> working point for identification of electrons is used in this thesis, as discussed in Chapter 8.	64
4.10	The measured muon and identification efficiencies using $Z \rightarrow \mu^+\mu^-$ events as a function of η for muons with $p_T > 10$ GeV from [76]. The three working points: <i>Loose</i> (yellow triangles), <i>Medium</i> (red squares), and <i>Tight</i> (blue circles) are shown for data (filled) and MC (open). Each working point corresponds to more background rejection. The <i>Medium</i> working point is used in this thesis, as discussed in Chapter 8.	65

4.11	Example of data isolation efficiencies for electrons from [66] as a function of E_T . In a) <i>Loose</i> and <i>Loose (Track Only)</i> target a uniform efficiency in E_T and η , while the later does not use the calorimeter based isolation. <i>Gradient</i> and <i>Gradient (Loose)</i> target an efficiency uniform in η , while depending on E_T . In b) efficiencies for different fixed selection requirements are shown. The ratio plot on the bottom shows the agreement between efficiencies from data and MC. A selection most similar to <i>Fix (Tight, Track Only)</i> is used in this thesis, as discussed in Chapter 8.	67
4.12	Example of a) the efficiency for the single electron trigger for electrons with $E_T > 7$ GeV from [77] and b) the muon trigger efficiencies from [78]. The plateau of $\sim 80\%$ for the muon trigger efficiency is driven by the requirement of the number of coincidence hits in the TGC [78]. These trigger efficiencies are similar to the ones used in this analysis.	68
5.1	Example of the steps to correct a variable by shifting and stretching, according to Equation 5.2. a) The variables, mean and standard deviation, do not agree between data and MC. b) Shifting the variable by the difference in the means brings both distributions into better agreement. c) Correcting just the width of the distribution makes the shape in better agreement. d) Correcting the mean and the width of the variables improve the agreement. .	71
5.2	The variables used to create a two-dimensional phase space for the VCFs with a combination of $ \eta^l $ and E_T or $ \eta^l $ and μ . a) The signed η^l distribution, b) the transverse energy of the electron, and c) the mean number of interactions per bunch crossing, μ . In the upper plots, data and its statistical error are represented by black points and MC is the blue histograms. The ratio of data/MC can be seen in the lower plots, which shows the relative agreement.	73
5.3	Example corrections are shown for a) one tracking variable, d_0 , in the region $1.81 < \eta^l < 2.01$ and $30 < E_T < 35$ GeV and for b) a shower shape variable, R_η , in the region $0 < \eta^l < 0.1$ and $35 < E_T < 40$ GeV. In the top plot the black points are the uncorrected MC, the red points are the corrected MC, and the blue histograms are the data. The ratio plot on the bottom shows the ratio of data and MC before correction in black and after in red.	75
5.4	Examples of a variable that shows μ dependency in a)-c) and one that does not in d)-f). The dependency for d_0/σ_{d_0} can be seen in μ_{Data} and $\sigma_{\text{Data}}/\sigma_{\text{MC}}$ in the barrel region of the ECAL. There is not a strong dependency for some variables, such as w_{η_2} , which can be seen in all three variables in the endcap region of the ECAL.	76

5.5	An example of the shower shape reweighting (SSR) applied to R_η in the region $0.40 < \eta^l < 0.60$ from [80]. Data is denoted by the black points, MC uncorrected by the red points, and MC after SSR reweighting denoted by the blue points. The ratio of MC over data is shown on the bottom for both the uncorrected MC in red and corrected in blue and shows improvement in the relative agreement, especially in the peak region.	78
5.6	An example of the extra steps need to correct eProbabilityHT for the region $1.15 < \eta^l < 1.37$ and $25 < E_T < 30$ GeV. a) Is eProbabilityHT before being transformed using Equation 5.6. b) The transformed distribution, T(ePHT), is used to calculate the VCFs before it is transformed back using Equation 5.7. In the top plot the black points are the uncorrected MC, the red points are the corrected MC, and the blue histograms are the data. The ratio plot on the bottom shows the ratio of data and MC before correction in black and after in red.	79
5.7	The ECAL is supported by two rails at the top of the subdetector; gravity causes the detector to deform from the intended circular shape of the ECAL, in black, into the distorted shape, in red. This distortion is on the order of a few mm, causing a difference in the spacing between layers in the LAr calorimeter, which has an outer diameter of 4 m.	80
5.8	The phase space of the δ correction to $\Delta\eta_1$ from Equation 5.8 as a function of ϕ and η	81
5.9	The original MC modeled Gaussian shape in black doesn't match with the data shape in the blue histograms. a) The correction without applying a pear shape correction and the VCF corrections are not able to correct the difference. b) The pear shape correction is applied and allows the VCF to bring the agreement within 10%.	82
5.10	Example of the difference between values derived from probe electrons selected with <i>VeryLoose</i> and the <i>Tight</i> selection criteria for VCF variables μ_{MC} and μ_{Data} for the variable d_0/σ_{d_0} . On the left is the barrel region of the ECAL and on the right is the endcap region. This comparison found, that for most variables, the difference is equal to zero within one standard deviation for most $ \eta^l $ and E_T bins.	84
5.11	Examples of the electron efficiencies with different correction configurations derived from Tag&Probe for the <i>Loose</i> identification selection in three E_T slices. They are compared to the measured efficiency in data in black. The configurations are applying just SSR and pear shape correction in red, adding the VCF shift in blue, and adding the stretch in pink. The ratio plots show the relative agreement between data and the configurations.	86

5.12	Comparison of corrections for different $ \eta^l $ bins to demonstrate that the corrections do not improve the data and MC agreement evenly for all $ \eta^l $ bins. a) R_ϕ and c) R_η for $ \eta^l $ bins that the corrections improve the data and MC agreement. b) R_ϕ and d) R_η but in neighbouring bins that do not see the same level of improvement for the data and MC agreement. This commonly occurs in regions where the detector geometry changes significantly. The top shows this for R_ϕ near the transition region, $1.37 < \eta^l < 1.52$, of the ECAL and the bottom is R_η near the outside edge of the ECAL.	88
5.13	Examples of the probe electron SFs using the Tag&Probe framework for the <i>Loose</i> selection in three E_T slices. Here, the baseline SSR and pear shape correction in red, adding the VCF shift in blue, adding the stretch in pink, and to the current recommendations [84] in black.	90
6.1	The cross-sections for different processes produced from proton–proton interactions, measured by ATLAS [85] for centre-of-mass energies $\sqrt{s} = 5, 7, 8, 13, 13.6$ TeV. These cross-section measurements span many orders of magnitude and depend on the probability of the interaction to occur. . . .	91
6.2	Example of a Drell–Yan process, where protons of momenta P_i , interact via partons with momenta, $x_i P_i$, annihilate into a W^- boson, which decays into two leptons.	92
6.3	Examples of parton distribution functions (PDFs) multiplied by their momentum fraction at the scale $Q^2 = 1.9 \text{ GeV}^2$ as a function of their momentum fraction from [87]. a) The PDF for the valence up quark and b) the PDF for the sea up and anti-up quarks obtained from fitting ATLAS and HERA [86] data.	93
6.4	Examples of a) LO, b) NLO, and c) NNLO Feynman diagrams for the Drell–Yan process. The order is determined by the number of QCD vertices. With increased order, the number of Feynman diagrams increases. . . .	95
6.5	The second order angular polynomials associated with each helicity dependent cross-section, plotted as a function of $\cos \theta$ and ϕ	97

6.6	On the left is the lab frame, where protons come from opposite directions producing a W in an arbitrary direction. The Collins–Soper (CS) frame [89] is a rest frame of the W boson, which has the z -axis in the direction of the W boson in the lab frame producing a bisecting angle, α , between the protons. The y -axis is normal to the proton–proton plane and the x -axis is chosen to complete a right-handed Cartesian coordinate system. The CS angles are defined to have the polar angle θ_{CS} as the angle between the “negative” lepton and the z -axis while the ϕ_{CS} as the azimuthal angle between the xz -plane and the lepton plane. For W^+ events, the “negative” lepton is the neutrino.	100
6.7	Examples of different Feynman diagrams for QCD production of W boson, which contribute to the Drell–Yan angular coefficients. a) quark–antiquark annihilation at $\mathcal{O}(\alpha_S^0)$, b) quark–antiquark annihilation at $\mathcal{O}(\alpha_S^1)$, c) Compton scattering of gluon and quark at $\mathcal{O}(\alpha_S^1)$, d) Compton scattering of gluon and antiquark at $\mathcal{O}(\alpha_S^1)$, and e) gluon–gluon interaction at $\mathcal{O}(\alpha_S^2)$	102
6.8	Drell–Yan angular coefficients A_0 to A_4 as a function of p_T^W for $W^- \rightarrow e^- \bar{\nu}$. Produced using POWHEG+PYTHIA8 at generator level with the same binning as the $\sqrt{s} = 8$ TeV Z Drell–Yan angular coefficients analysis [6]. . . .	103
7.1	Examples of how the truth phase space templates that describe the angular distributions evolve as a function of $p_T^{W,\text{Truth}}$ for $W^- \rightarrow e^- \bar{\nu}$ events. Along the x -axis is $\cos \theta_{\text{CS}}$, along the y -axis is ϕ_{CS} , and along the z -axis is the number of events.	108
7.2	Examples of how the reconstructed phase space templates that describe the angular distributions evolve as a function of reconstructed $p_T^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events. Along the x -axis is $\cos \theta_{\text{CS}}$, along the y -axis is ϕ_{CS} , and along the z -axis is the number of events. Events with $\cos \theta_{\text{CS}} > 1$ are from events where the neutrino cannot be reconstructed and are put in the additional $\cos \theta_{\text{CS}}$ bin to not distort the shape of the other bins.	110
7.3	Polynomial templates $P_{i,1}$ for the i th coefficient in the p_T^W bin 1 before folding, which ranges from $8 < p_T^{l\nu} < 17$ GeV for $W^- \rightarrow e^- \bar{\nu}$ events. Along the x -axis is reconstructed $\cos \theta_{\text{CS}}$, along the y -axis is reconstructed ϕ_{CS} , and along the z -axis is the number of events. Each event is weighted by the i th polynomial which is why events are allowed to be negative.	112

7.4	Polynomial templates of $P_{3,j}$ for A_3 in the j th p_T^W bin before folding for $W^- \rightarrow e^- \bar{\nu}$ events to illustrate how the polynomial changes with respect to $p_T^{l\nu}$. Examples are shown for low, middle, and high $p_T^{l\nu}$. Along the x-axis is reconstructed $\cos \theta_{CS}$, along the y-axis is reconstructed ϕ_{CS} , and along the z-axis is the number of events. Each event is weighted by the i th polynomial which is why events are allowed to be negative.	113
8.1	The integrated luminosity vs. mean number of interactions per bunch crossing taken with the ATLAS detector during Run 2 at $\sqrt{s} = 13$ TeV from [45]. Note, the small peak on the left side corresponds to the integrated luminosity of the low- $\langle\mu\rangle$ data set used in this analysis.	117
8.2	The neutrino transverse momentum, p_T^ν , is approximated using the E_T^{miss} , which is calculated using the calibrated hadronic recoil. a) The hadronic recoil, u_T , is calculated by the vectorial sum of all p_T from hadronic activity produced from initial state gluon and quark radiation associated to the event. The hadronic recoil and the charged lepton transverse momentum are used to solve for the missing energy associated the neutrino p_T . b) A comparison of the transverse momentum and hadronic recoil for the Z boson and the bottom panel shows they are in good agreement [12]. c) The hadronic recoil is calibrated by removing a lepton from $Z \rightarrow \ell\ell$ events to calculate E_T^{miss} using u_T and comparing to the measured lepton transverse momentum, $p_T^{l,2}$	123
8.3	Distributions of a) $W^- \rightarrow e^- \bar{\nu}$ hadronic recoil, u_T , b) $Z \rightarrow ee$ hadronic recoil, u_T , c) $W^- \rightarrow e^- \bar{\nu}$ missing transverse energy, E_T^{miss} , and d) $Z \rightarrow ee$ dilepton transverse momentum, $p_T^{\ell\ell}$, for the $\sqrt{s} = 13$ TeV low- $\langle\mu\rangle$ data set from [12]. The data in black points are compared to the sum of the signal and expected backgrounds. The lower panels show the agreement of individual bins between data and the expected signal and background in black points and the prediction uncertainty as a dark grey band. These distributions are shown pre-fit where the expected agreement between data and MC post-fit.	124
8.4	The hadronic recoil detector response for five p_T^Z truth intervals for $Z \rightarrow \mu\mu$ events at $\sqrt{s} = 13$ TeV from [12]. This figure illustrates that, with lower hadronic recoil resolution, it is possible to differentiate between lower p_T^Z values. Likewise, it would also allow for better resolution for lower p_T^W values.	125

8.5	Comparison of the different quadratic equation root choices for p_z^ν reconstruction to the one generated in MC. The solid black points are the values of p_z^ν generated, the red upward triangles are the p_z^ν values when the positive root is chosen, the blue downward triangles are the p_z^ν values when the negative root is chosen, and the open black points are randomly choosing between the two different roots.	128
8.6	Fraction of events that have a real solution as a function of $p_T^{l\nu}$. Events will not have a solution when $m_T > m^W$, which happens more frequently at large $p_T^{l\nu}$	129
8.7	An event display of a $W^- \rightarrow \mu^- \bar{\nu}$ candidate event from proton–proton collisions recorded by ATLAS at $\sqrt{s} = 13$ TeV, from the low- $\langle\mu\rangle$ data set [12]. The event has a missing transverse energy of 49 GeV and the right display shows the transverse energy of the muon is 35 GeV. The muon and missing transverse energy are used to reconstruct a W boson candidate with reconstructed transverse momentum, $p_T^{l\nu} = 16$ GeV, and transverse mass, $m_T = 77$ GeV. In the left display, the path of the muon is shown in orange and the missing transverse energy is denoted by the red arrow. There are two additional reconstructed vertices, as seen in the bottom panel, which shows charged particle tracks from other particles in blue and the track of the muon in orange.	131
8.8	Examples of the track isolation variable for $W^+ \rightarrow \mu^+ \nu$ that is used to estimate the MJ background. Control regions (CR) that are anti-isolated are used to fit the MJ shape and yield, then extrapolated back to the signal region (SR). The amount of MJ in the SR is not shown here because this is pre-extrapolation.	133
8.9	The kinematic distributions for a) p_T^l , b) η^l , c) E_T^{miss} , and d) $\Delta\phi(l, E_T^{\text{miss}})$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.	136

- 8.10 The kinematic distributions for a) m_T , b) $p_T^{l\nu}$ in a fine binning, c) $p_T^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 138
- 8.11 The integrated CS angular distributions of a) $\cos \theta_{CS}^{\text{Reco}}$ and b) ϕ_{CS}^{Reco} for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 140
- 8.12 The two-dimensional $(\cos \theta_{CS}^{\text{Reco}}, \phi_{CS}^{\text{Reco}})$ distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for two bins: a) $0 < p_T^{l\nu} < 8$ GeV and b) $27 < p_T^{l\nu} < 40$ GeV. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 141
- 8.13 The two-dimensional $(\cos \theta_{CS}^{\text{Reco}}, \phi_{CS}^{\text{Reco}})$ distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_T^{l\nu} < 110$ GeV and b) $210 < p_T^{l\nu} < 600$ GeV. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 142

- 9.1 The measured first five coefficients, A_0 to A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for the $W^- \rightarrow e^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are shifted horizontally to avoid point and uncertainty bar overlap but have the same uncertainty and position in $p_T^{l\nu}$ 148
- 9.2 The unfolded three-dimensional ($\cos \theta_{CS}$, ϕ_{CS} , $p_T^{l\nu}$) distribution unwrapped into a one-dimensional distribution for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for each $p_T^{l\nu}$ bin separated by vertical dashed lines and labelled underneath the graph. The measured data, N_{obs} , is denoted by black points, the data-driven MJ background estimate, MC signal, and MC background are combined into N_{exp} and normalized to the data luminosity are denoted by the red points pre-fit and post-fit by the blue points. The ratio of N_{obs} over N_{exp} for both pre- and post-fit is displayed on the bottom plot showing the agreement improves through the NLL minimization. 150
- 9.3 Comparison of the expected coefficients from Asimov data for the four major coefficients A_0 , A_2 , A_3 , A_4 and the differential cross-section as a function of $p_T^{l\nu}$ to show the expected differences between $W^+ \rightarrow e^+ \nu$ (black) and $W^- \rightarrow e^- \bar{\nu}$ (blue). The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. 152
- 9.4 Comparison of the measured coefficients A_3 , A_4 , and the differential cross-section between $W^+ \rightarrow e^+ \nu$ (black) and $W^- \rightarrow e^- \bar{\nu}$ (blue) as a function of $p_T^{l\nu}$ for data. These expected differences for a) A_3 and b) A_4 do not have enough statistical sensitivity to measure, while the difference c) for the differential cross-section can be seen. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. . 153
- 9.5 A test of the compatibility of the $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels as a function of $p_T^{l\nu}$ by taking the difference of the final fit results, denoted by the blue squares. The expected difference is zero and this shows there is good agreement between the channels. Here, only the data statistical uncertainties are used because of unknown correlations of systematic uncertainties between channels. 155
- 9.6 The measured values of four major coefficients A_0 , A_2 , A_3 , A_4 and the differential cross-section for the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels as a function of $p_T^{l\nu}$ can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. . 157

- 9.7 The integrated CS angular distributions of $\cos \theta_{\text{CS}}^{\text{Reco}}$ for a) $W^- \rightarrow e^- \bar{\nu}$ and b) $W^+ \rightarrow e^+ \nu$ and $\phi_{\text{CS}}^{\text{Reco}}$ for c) $W^- \rightarrow e^- \bar{\nu}$ and d) $W^+ \rightarrow e^+ \nu$ for events that pass selection criteria listed in Table 8.3 shown on a linear scale to show the difference in shape. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 159
- 9.8 The difference between W^- and W^+ of the measured coefficients A_3 , A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for data is denoted by blue squares. The expected differences for a) A_3 and b) A_4 are not observed, while the difference for c) the differential cross-section can be observed. . . 160
- 9.9 The values of the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels as a function of $p_T^{l\nu}$ can be seen in black, are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. 162
- 10.1 Depiction of the sources of data statistical uncertainty for A_0 in $p_T^{l\nu}$ bin 0, denoted by $A_{0,0}$, corresponding to $0 < p_T^{l\nu} < 8$ GeV. The uncertainty on the parameter $A_{0,0}$ is the uncorrelated statistical uncertainty. The shape uncertainty is the non-migrational correlated uncertainty between different angular coefficients in the same $p_T^{l\nu}$ bin. The correlated migration uncertainty is comprised of three types: self, shape, and norm. Self-migration is for the migration between $p_T^{l\nu}$ bins for the same coefficient, shape-migration is the migration between angular coefficients and $p_T^{l\nu}$ bins, and norm-migration is the self-migration for the differential cross-section, denoted by A_9 . Based on an image from [6]. 167
- 10.2 Decomposition of the total data statistical uncertainty as a function of $p_T^{l\nu}$ for the absolute uncertainty of A_0 . The black circles are the total data statistical uncertainty, which is the addition in quadrature of the other components. The blue squares correspond to the uncertainty due to the parameter of interest. The yellow upright triangles are the migrational uncertainty correlated between $p_T^{l\nu}$ bins. The red downward triangles are the uncorrelated shape uncertainty for different coefficients within the same bin. Lastly, the grey open circles are the cutoff uncertainty based on the choice of number of migration terms. 167

10.3	Decomposition of the total migrational uncertainty as a function of $p_T^{l\nu}$ for the absolute uncertainty of A_0 . The black circles correspond to the total migrational uncertainty, which is the addition in quadrature of the other components and is the same as the yellow upright triangles in Figure 10.2. The blue squares are the migration between $p_T^{l\nu}$ bins for the same coefficient, the yellow upright triangles are the migration between coefficients and $p_T^{l\nu}$ bins, and the red downwards triangles are the migration between $p_T^{l\nu}$ bins for the norm.	169
10.4	A breakdown of all sources of measured uncertainty for the angular coefficient A_0 as a function of $p_T^{l\nu}$ for Figure 9.1 for the $W^- \rightarrow e^- \bar{\nu}$ channel. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	172
10.5	A breakdown of all sources of measured uncertainty for the angular coefficients, A_0 to A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for Figure 9.1 for the $W^- \rightarrow e^- \bar{\nu}$ channel. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	185
10.6	The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of $p_T^{l\nu}$ for the $W^- \rightarrow e^- \bar{\nu}$ channel. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modelled as most points are within one standard deviation of the pre-fit uncertainty.	188
11.1	The measured four major coefficients – A_0, A_2, A_3, A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for the combination of the $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.	191
11.2	The measured four major coefficients – A_0, A_2, A_3, A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for the combination of the $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.	192

11.3	A breakdown of all sources of measured uncertainty for the measured angular coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for Figure 11.1 for the combination of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The data statistical uncertainty outlined in Section 10.1 is denoted by the black points and all sources of systematic uncertainty are outlined in Section 10.2.	194
11.4	A breakdown of all sources of measured uncertainty for the measured angular coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for Figure 11.2 for the combination of $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$ channels. The data statistical uncertainty outlined in Section 10.1 is denoted by the black points and all sources of systematic uncertainty are outlined in Section 10.2.	195
12.1	A prediction from simulation for the uncertainty on the hadronic recoil, u_T , as a function of the average of the mean number of interactions per bunch crossing, $\langle \mu \rangle$, in the ATLAS detector taken from [94]. This is shown for the normal high- $\langle \mu \rangle$ settings for the calorimeter in the open circles and for the special calorimeter settings used with lower energy thresholds used for the low- $\langle \mu \rangle$ data set.	200
F.1	Choice combination matrix for choosing the correct p_z^ν solution and whether the right sign of $\cos \theta_{CS}$ is obtained. The right sign of $\cos \theta_{CS}$ is what increases the sensitivity to A_4 , while the wrong sign decreases it.	235
F.2	Comparison for A_4 for the inclusion of events dependent on the ambiguity in the sign of $\cos \theta_{CS}$ and y^W in $p_T^{l\nu}$ binning for $W^- \rightarrow e^- \bar{\nu}$ and $W^+ \rightarrow e^+ \nu$. a) and c) shows the values are not biased by the ambiguity. b) and d) show the statistical uncertainty is reduced by including events that only have ambiguity in y^W	236
G.1	The measured last three coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^- \rightarrow e^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	238

- G.2 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $8 < p_{\text{T}}^{l\nu} < 17$ GeV and b) $17 < p_{\text{T}}^{l\nu} < 27$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 239
- G.3 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $40 < p_{\text{T}}^{l\nu} < 55$ GeV and b) $55 < p_{\text{T}}^{l\nu} < 75$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 240
- G.4 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $110 < p_{\text{T}}^{l\nu} < 150$ GeV and b) $150 < p_{\text{T}}^{l\nu} < 210$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 241
- G.5 The integrated CS angular distributions of a) $\cos \theta_{\text{CS}}^{\text{Reco}}$ and b) $\phi_{\text{CS}}^{\text{Reco}}$ for $W^+ \rightarrow e^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 242

- G.6 The kinematic distributions for a) p_T^l , b) η^l , c) E_T^{miss} , and d) $\Delta\phi(l, E_T^{\text{miss}})$ for $W^+ \rightarrow e^+\nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 243
- G.7 The kinematic distributions for a) m_T , b) $p_T^{l\nu}$ in a fine binning, c) $p_T^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^+ \rightarrow e^+\nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 244
- G.8 The two-dimensional $(\cos\theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$ distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^+ \rightarrow e^+\nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $0 < p_T^{l\nu} < 8$ GeV and b) $8 < p_T^{l\nu} < 17$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 245
- G.9 The two-dimensional $(\cos\theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$ distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^+ \rightarrow e^+\nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $17 < p_T^{l\nu} < 27$ GeV and b) $27 < p_T^{l\nu} < 40$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 246

- G.10 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^+ \rightarrow e^+\nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $40 < p_{\text{T}}^{l\nu} < 55$ GeV and b) $55 < p_{\text{T}}^{l\nu} < 75$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 247
- G.11 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^+ \rightarrow e^+\nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_{\text{T}}^{l\nu} < 110$ GeV and b) $110 < p_{\text{T}}^{l\nu} < 150$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 248
- G.12 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^+ \rightarrow e^+\nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $150 < p_{\text{T}}^{l\nu} < 210$ GeV and b) $210 < p_{\text{T}}^{l\nu} < 600$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 249
- G.13 The measured coefficients – A_0 , A_1 , A_2 – and their uncertainties as a function of $p_{\text{T}}^{l\nu}$ for the $W^+ \rightarrow e^+\nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2. 250

G.14	The measured coefficients – A_3 and A_4 –, the differential cross-section, and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow e^+\nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	251
G.15	The measured coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow e^+\nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	252
G.16	The unfolded three-dimensional ($\cos \theta_{CS}$, ϕ_{CS} , $p_T^{l\nu}$) distribution unwrapped into a one-dimensional distribution for $W^+ \rightarrow e^+\nu$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for each $p_T^{l\nu}$ bin separated by vertical dashed lines and labelled underneath the graph. The measured data, N_{obs} , is denoted by black points, the data-driven MJ background estimate, MC signal, and MC background are combined into N_{exp} and normalized to the data luminosity are denoted by the red points pre-fit and post-fit by the blue points. The ratio of N_{obs} over N_{exp} for both pre- and post-fit is displayed on the bottom plot showing the agreement improves through the NLL minimization.	253
G.17	The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow e^+\nu$ channel. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty.	254

- G.18 The integrated CS angular distributions of a) $\cos \theta_{\text{CS}}^{\text{Reco}}$ and b) $\phi_{\text{CS}}^{\text{Reco}}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 255
- G.19 The kinematic distributions for a) p_{T}^l , b) η^l , c) $E_{\text{T}}^{\text{miss}}$, and d) $\Delta\phi(l, E_{\text{T}}^{\text{miss}})$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 256
- G.20 The kinematic distributions for a) m_{T} , b) $p_{\text{T}}^{l\nu}$ in a fine binning, c) $p_{\text{T}}^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 257
- G.21 The two-dimensional $(\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$ distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $0 < p_{\text{T}}^{l\nu} < 8$ GeV and b) $8 < p_{\text{T}}^{l\nu} < 17$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 258

- G.22 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $17 < p_{\text{T}}^{l\nu} < 27$ GeV and b) $27 < p_{\text{T}}^{l\nu} < 40$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 259
- G.23 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $40 < p_{\text{T}}^{l\nu} < 55$ GeV and b) $55 < p_{\text{T}}^{l\nu} < 75$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 260
- G.24 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_{\text{T}}^{l\nu} < 110$ GeV and b) $110 < p_{\text{T}}^{l\nu} < 150$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 261

- G.25 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $150 < p_{\text{T}}^{l\nu} < 210$ GeV and b) $210 < p_{\text{T}}^{l\nu} < 600$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 262
- G.26 The measured coefficients – A_0 , A_1 , A_2 – and their uncertainties as a function of $p_{\text{T}}^{l\nu}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2. 263
- G.27 The measured coefficients – A_3 and A_4 –, the differential cross-section, and their uncertainties as a function of $p_{\text{T}}^{l\nu}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2. 264
- G.28 The measured coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_{\text{T}}^{l\nu}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2. 265

- G.29 The unfolded three-dimensional ($\cos \theta_{\text{CS}}, \phi_{\text{CS}}, p_{\text{T}}^{l\nu}$) distribution unwrapped into a one-dimensional distribution for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for each $p_{\text{T}}^{l\nu}$ bin separated by vertical dashed lines and labelled underneath the graph. The measured data, N_{obs} , is denoted by black points, the data-driven MJ background estimate, MC signal, and MC background are combined into N_{exp} and normalized to the data luminosity are denoted by the red points pre-fit and post-fit by the blue points. The ratio of N_{obs} over N_{exp} for both pre- and post-fit is displayed on the bottom plot showing the agreement improves through the NLL minimization. 266
- G.30 The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of $p_{\text{T}}^{l\nu}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty. 267
- G.31 The integrated CS angular distributions of a) $\cos \theta_{\text{CS}}^{\text{Reco}}$ and b) $\phi_{\text{CS}}^{\text{Reco}}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 268
- G.32 The kinematic distributions for a) p_{T}^l , b) η^l , c) $E_{\text{T}}^{\text{miss}}$, and d) $\Delta\phi(l, E_{\text{T}}^{\text{miss}})$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 269

- G.33 The kinematic distributions for a) m_T , b) $p_T^{l\nu}$ in a fine binning, c) $p_T^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 270
- G.34 The two-dimensional $(\cos \theta_{CS}^{\text{Reco}}, \phi_{CS}^{\text{Reco}})$ distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $0 < p_T^{l\nu} < 8$ GeV and b) $8 < p_T^{l\nu} < 17$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 271
- G.35 The two-dimensional $(\cos \theta_{CS}^{\text{Reco}}, \phi_{CS}^{\text{Reco}})$ distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $17 < p_T^{l\nu} < 27$ GeV and b) $27 < p_T^{l\nu} < 40$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 272

- G.36 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $40 < p_{\text{T}}^{l\nu} < 55$ GeV and b) $55 < p_{\text{T}}^{l\nu} < 75$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 273
- G.37 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_{\text{T}}^{l\nu} < 110$ GeV and b) $110 < p_{\text{T}}^{l\nu} < 150$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 274
- G.38 The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $150 < p_{\text{T}}^{l\nu} < 210$ GeV and b) $210 < p_{\text{T}}^{l\nu} < 600$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ. 275
- G.39 The measured coefficients – A_0 , A_1 , A_2 – and their uncertainties as a function of $p_{\text{T}}^{l\nu}$ for the $W^+ \rightarrow \mu^+ \nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2. 276

G.40	The measured coefficients – A_3 and A_4 –, the differential cross-section, and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow \mu^+ \nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	277
G.41	The measured coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow \mu^+ \nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	278
G.42	The unfolded three-dimensional ($\cos \theta_{CS}$, ϕ_{CS} , $p_T^{l\nu}$) distribution unwrapped into a one-dimensional distribution for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for each $p_T^{l\nu}$ bin separated by vertical dashed lines and labelled underneath the graph. The measured data, N_{obs} , is denoted by black points, the data-driven MJ background estimate, MC signal, and MC background are combined into N_{exp} and normalized to the data luminosity are denoted by the red points pre-fit and post-fit by the blue points. The ratio of N_{obs} over N_{exp} for both pre- and post-fit is displayed on the bottom plot showing the agreement improves through the NLL minimization.	279
G.43	The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow \mu^+ \nu$ channel. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty.	280
G.44	The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The combined channel measurements can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.	281

- G.45 The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The combined channel measurements can be seen in black and are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. 282
- G.46 The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combination of the $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. 283
- G.47 A breakdown of all sources of measured uncertainty for the measured angular coefficients, A_1 , A_5 , A_6 , and A_7 , as a function of $p_T^{l\nu}$ for Figure G.46 for the combination of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2. . . . 284
- G.48 The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of $p_T^{l\nu}$ for the combination of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty. 285
- G.49 The measured values of the four major coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$ channels. The combined channel measurement can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. 287
- G.50 The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$ channels. The combined channel measurement can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. . 288

G.51	The measured values of coefficients A_0 , A_2 , A_3 , A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The combined channel measurements can be seen in black and are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.	289
G.52	The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The combined channel measurements can be seen in black and are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.	290
G.53	The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combination of the $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.	291
G.54	A breakdown of all sources of measured uncertainty for the measured angular coefficients, A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for Figure G.53 for the combination of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.	292
G.55	The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of $p_T^{l\nu}$ for the combination of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty.	293

CHAPTER 1

INTRODUCTION

The Large Hadron Collider (LHC), located on the France–Switzerland border at Conseil Européen pour la Recherche Nucléaire (CERN), accelerates bunches of either protons or heavy ions to high energies that are collided, producing many different particles. Detectors, such as ATLAS [1] (A Toroidal LHC ApparatuS), are built around the interaction points of the collider and are used to measure the signatures of particles produced during the collisions interacting with the detector. Measuring the interactions and decays of the produced particles allows us to improve our understanding of particle interactions, described by the Standard Model (SM), as well as search for possible new physics beyond the SM. The SM is known to be incomplete since there are many observed phenomena that are not described within the model. By measuring deviations in known particle interactions processes, one can make improvements in the current model. It will also be possible to impose constraints in the search for new physics if discrepancies are found between results predicted by the SM and what is experimentally observed.

Proton–proton interactions producing W/Z bosons that decay immediately into leptonic particles, such as electrons, are known as the Drell–Yan process [2, 3]. The differential cross-section for this process expresses the probability to occur depending on the W/Z bosons’ kinematic variables – such as the transverse momentum, mass, and rapidity – and the angles between the leptons and the boson polarization [4]. The differential cross-section can be decomposed into eight spin–related ratios and an unpolarized term. These eight coefficients, are known as the Drell–Yan angular coefficients and are denoted by A_0 to A_7 . They correlate to the different combinations of the particle states, where the initial particles – the quarks from the proton–proton interactions – are spin 1/2, the W/Z bosons are spin 1, and the decay products are spin 1/2. The coefficients and unpolarized

term are coupled to trigonometric polynomials of $\cos \theta$ and ϕ , which describe the angle between the leptons and the W/Z boson polarization. It is possible to isolate each coefficient by using the property of the polynomials being orthogonal to each other because they are related to the spherical harmonics.

The most recent measurement of the Z boson angular coefficients was performed by the LHCb Collaboration [5] with $Z \rightarrow \mu^+ \mu^-$ events at the centre-of-mass energy $\sqrt{s} = 13$ TeV and extracted four of the eight coefficients as a function of the transverse momentum, p_T^Z . All eight Z boson angular coefficients were extracted using $Z \rightarrow e^+ e^-$ and $Z \rightarrow \mu^+ \mu^-$ events by the ATLAS Collaboration [6] at $\sqrt{s} = 8$ TeV as a function of p_T^Z and rapidity, y^Z . With the measurement of the fourth angular coefficient, denoted A_4 , the ATLAS Collaboration was also able to extract the weak mixing angle, $\sin^2 \theta_W$ [7], which can be used to relate the W and Z masses through

$$\sin^2 \theta_W = 1 - \left(\frac{m^W}{m^Z} \right)^2, \quad (1.1)$$

where m^W and m^Z are the masses of the W and Z , respectively. The weak mixing angle is one of the free parameters of the SM, which means the theory does not provide values for them and measuring it allows for constraining the SM. There have been two other measurements of the Z boson coefficients. The first is by the CMS Collaboration [8], who extracted five of the eight coefficients with $Z \rightarrow \mu^+ \mu^-$ events at $\sqrt{s} = 8$ TeV as a function of p_T^Z . And the second is by the CDF Collaboration [9], who measured four of the eight coefficients with $Z \rightarrow e^+ e^-$ events at $\sqrt{s} = 1.96$ TeV as a function of p_T^Z . However, only the coefficients A_2 and A_3 have been measured for the W boson by the CDF Collaboration [10] at $\sqrt{s} = 1.8$ TeV as a function of p_T^W . This measurement can be seen in Figure 1.1, which had limited precision. For this measurement, A_2 and A_3 , were measured in four bins of p_T^W with the bounds $p_T^W = [15, 25, 35, 65, 105]$ GeV and the other six coefficients were consistent with zero within their uncertainty, as seen for A_0 and A_4 in

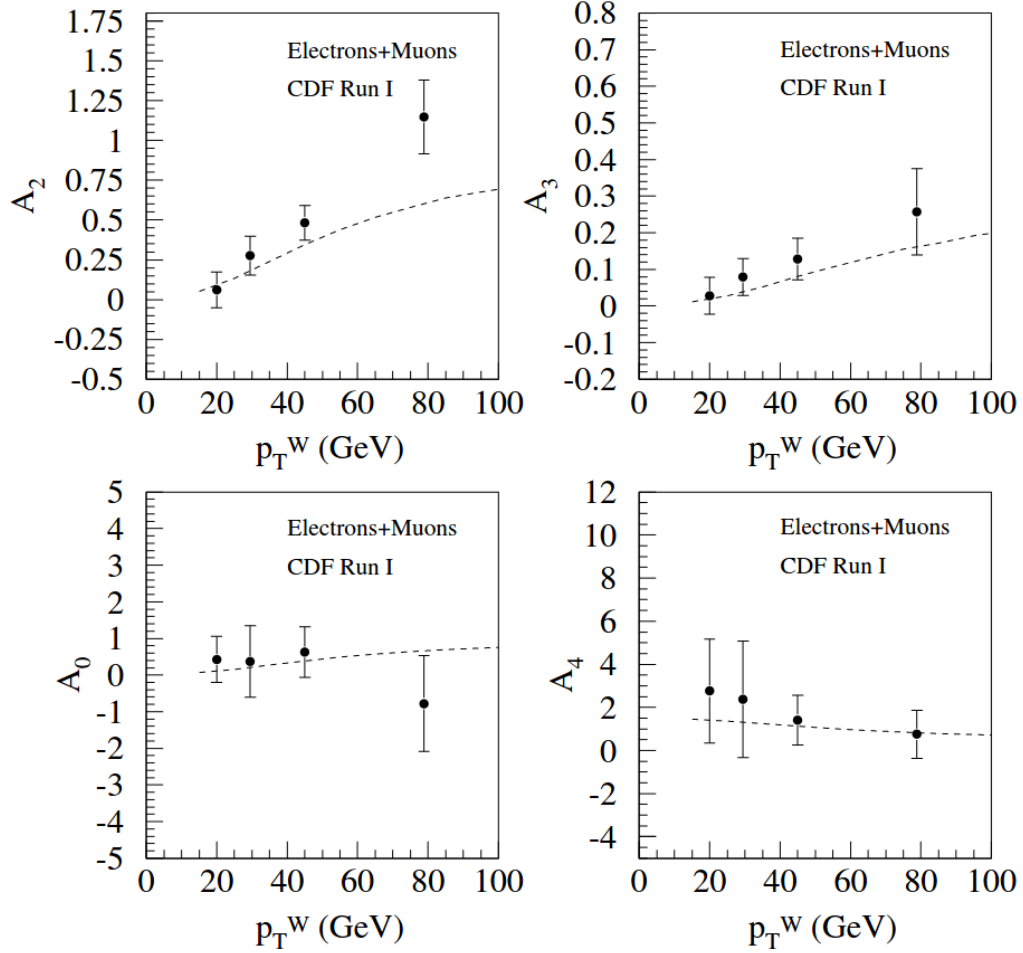


Figure 1.1: The measured values for the W boson Drell-Yan angular coefficients at $\sqrt{s} = 1.8$ TeV as a function of p_T^W by the CDF Collaboration [10]. Here, A_2 and A_3 were measured while A_0 and A_4 are consistent with zero within their uncertainty. The dashed line is the SM prediction of the angular coefficients.

the bottom of Figure 1.1.

The goal of this thesis is to improve on the measurement of the W boson Drell-Yan angular coefficients by measuring the four largest coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of p_T^W in ten bins of p_T^W with bounds:

$$p_T^W = [0, 8, 17, 27, 40, 55, 75, 110, 150, 210, 600] \text{ GeV}. \quad (1.2)$$

This measurement uses a special data set with fewer interactions per proton bunch crossing at $\sqrt{s} = 13$ TeV, taken with the ATLAS detector. During 2017 and 2018, special runs in the

LHC were performed with a lower average number of collisions per proton bunch crossing at $\sqrt{s} = 13$ TeV [11]. During these special data taking periods, the average number of interactions per bunch crossing was ~ 2 ; whereas, the average number of interactions per bunch crossing during the 2015-18 Run 2 was 33.7, reaching as high as 70 at some periods. Some particles, such as neutrinos, produced within interactions are unable to be measured by the detectors at the LHC. They can be estimated using missing energy within the detector, which is easier to estimate when there are fewer simultaneous interactions. More accurately reconstructed missing energy allows for improved estimation of neutrinos from W boson decay to make precision physics measurements. However, with a lower average number of collisions per proton bunch crossing, there is the caveat of a lower number of W bosons being produced.

The measurement of the W boson coefficients has the challenge of requiring a full reconstruction of the W boson. For the Drell–Yan process, the W boson will decay into a detectable lepton, like an electron, and a neutrino, which goes undetected in the ATLAS detector. Despite not being detected, the expected neutrino from the interaction can be approximated by calculating the missing energy in the xy -plane within the ATLAS detector, called the transverse plane. This transverse plane is perpendicular to the z -axis taken to be along the direction of the beam. This coordinate system can be seen in Figure 1.2, which is overlaid a schematic of the ATLAS detector and will be discussed further in Chapter 3. Even though the missing energy can be calculated in the transverse plane, information about the momentum of the neutrino along the beamline is lost. Estimating the neutrino in this way requires precisely measured deposited energy in the ATLAS detector, which can be done more easily with the special low pileup data set. Despite the small integrated luminosity of the special low pileup data set, it can be used for W boson measurements. Because W bosons are produced approximately ten times as frequently compared to Z bosons [12], the W bosons’ coefficients can have more statistical sensitivity than for the Z boson using the same special data set.

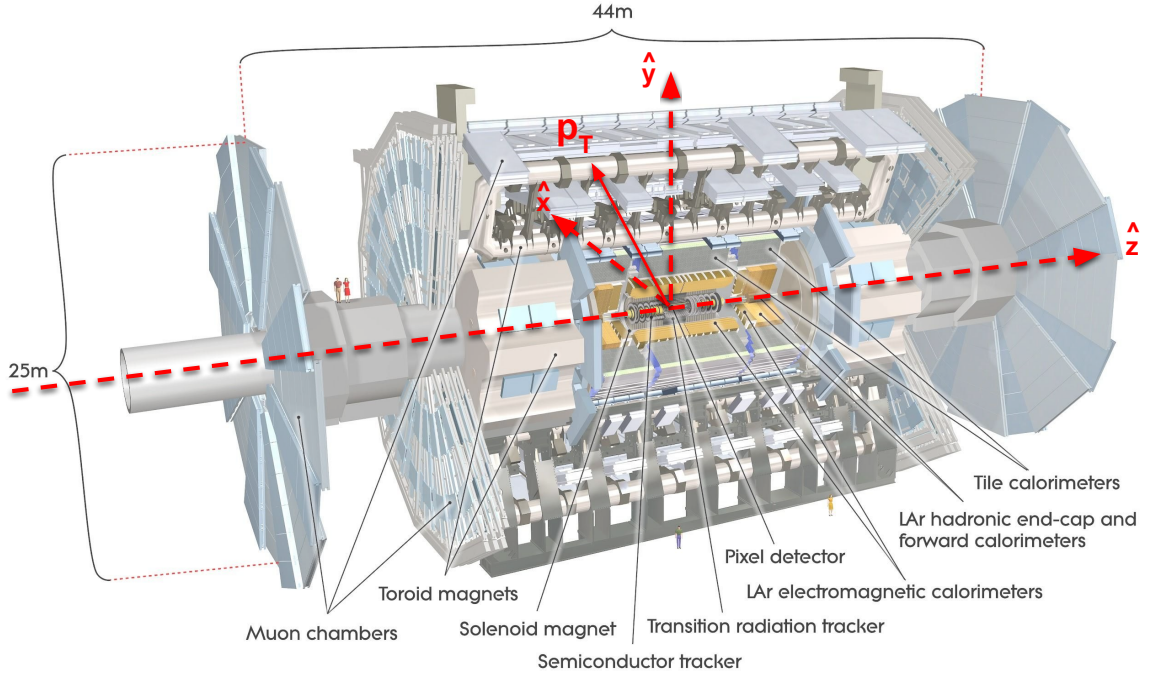


Figure 1.2: Schematic of the ATLAS detector with labeled subdetector and magnet locations [13]. The Cartesian coordinate system used by ATLAS includes: the z -axis is along the beamline, the x -axis points towards the centre of the LHC ring, and the y -axis consequentially points upwards. Here p_T , is the momentum in the transverse plane. A more detailed coordinate system and description of the ATLAS detector will be discussed in Chapter 3.

Quantum Chromodynamics (QCD) is the theory that describes the strong force for the interactions between quarks and gluons. The angular coefficients are a direct probe of the perturbative QCD production of the W boson during proton-proton interactions. This makes them an important input to other precision measurements of the W boson, such as the W boson mass, m^W . Currently, these coefficients are implemented in W boson production modelling by extrapolating the W boson coefficients from the Z boson coefficients, which increases the uncertainty on m^W measurements.

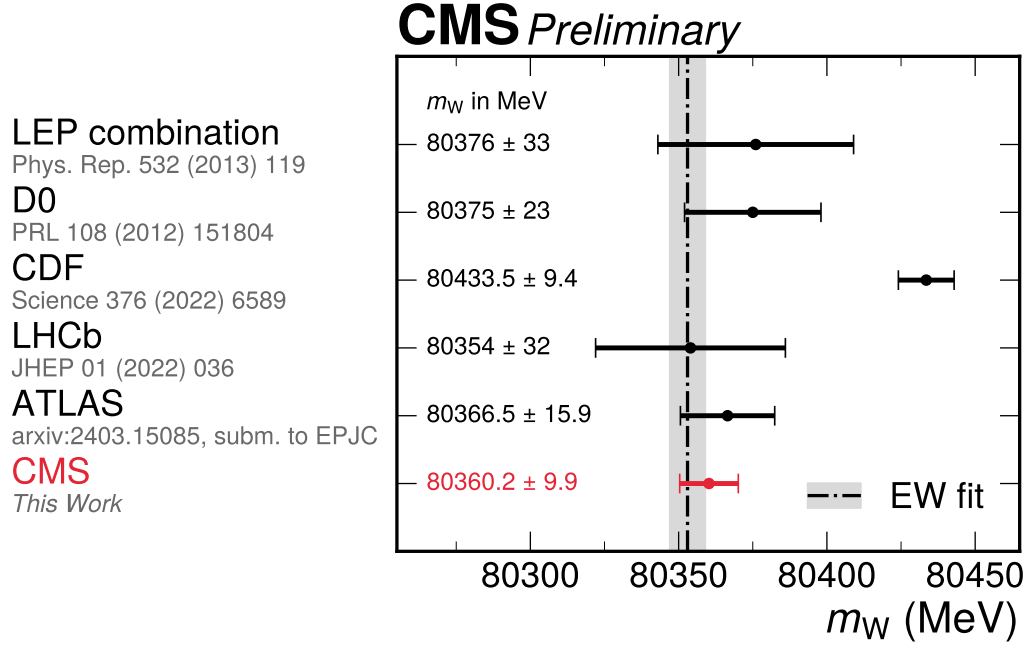


Figure 1.3: Comparison of the more recent m^W measurements compiled by the CMS Collaboration [17]. The measurement group and citation are listed on the left and correspond to each measurement point listed on an MeV scale.

Four measurements of the W boson mass by the CDF [14], LHCb [15], ATLAS [16], and CMS [17] Collaborations have been released in the past two years, making this an important and relevant area of research in precision measurements at the LHC. The aforementioned measurements are summarized in Figure 1.3, taken from the recent CMS Collaboration measurement [17]. This figure highlights a disagreement between the measurement by the CDF Collaboration and all other measurements, including the global electroweak (EW) fit, denoted by the dashed vertical line. A combination of measurements [18], which precludes the CMS Collaboration measurement, has also been published recently. The measurements of the W boson angular coefficients and differential cross-section in this thesis will allow for better perturbative QCD modelling for future m^W measurements.

The first three chapters of this thesis will provide the necessary information to understand and contextualize the remainder of the thesis. This starts with an overview of the SM and Quantum Field Theory (QFT) that are used to mathematically describe particles and their interactions in Chapter 2. Then, in Chapter 3, information is provided about the

LHC, used to accelerate the proton beams, and the ATLAS detector, used to collect the data, in this thesis. Next, Chapter 4, describes the steps used for particle reconstruction in the ATLAS detector to be used for analyses. The information in these first three chapters will allow for a better understanding of the chapters that follow.

Chapter 5 will cover my work towards corrections derived for the variables used to identify electrons in the ATLAS detector. Following this, in Chapter 6, is a detailed explanation of the Drell–Yan process and its angular coefficients. Next, the measurement method for extracting the Drell–Yan angular coefficients using a likelihood fit is outlined in Chapter 7. Chapter 8 describes the special data set used for this analysis, the Monte Carlo (MC) simulation used to model the data, and how the W boson events are reconstructed. This chapter also includes a comparison between the data and the model for kinematic distributions, to ensure that the data and MC are properly modeled. Chapter 9 gives a first look of the measured coefficients for the individual lepton flavour channels and the method for combining them to increase the statistical sensitivity. All sources of statistical and systematic uncertainty for the measurement are discussed in Chapter 10 before the final results are summarized in Chapter 11. The thesis is concluded in Chapter 12, where the results, possible improvements, and potential directions this research could go in the future are discussed.

In addition to the main body of the thesis, there is supplementary material that can be found in the Appendices, ordered by reference in the main body. Due to the large number of abbreviations, acronyms, and variables used in this thesis, a list of terminology with brief explanations can be found in Appendix A. Additional mathematical background for the Drell–Yan angular coefficients can be found in Appendices B and C, which cover the relation to the spherical harmonics and an example derivation of a moment for one polynomial. Extra details for the folding process used in the measurement method from Chapter 7 is outlined in Appendix D. A list of the specific details for the MC samples used for this analysis can be found in Appendix E. Additional information regarding the

ambiguity of the sign of $\cos \theta_{\text{CS}}$ and how it is able to be statistically resolved, allowing for the measurement of A_0 and A_4 , is discussed in Appendix F. Finally, in Appendix G, are additional plots for the channels not shown in Chapters 8, 9, and 11.

1.1 Personal Contributions

Working within a collaboration, such as ATLAS, allows for doing research that benefits from the work and expertise of my colleagues. Within a thesis, it is often difficult to understand what is original work and what information is presented to provide context and explanations. For my thesis, the majority of original work is contained within Chapters 5 and 7-11. Chapter 5 outlines the work I did as part of my authorship qualification project to become an ATLAS author, working alongside my colleague, Nataliia Zakharchuk. The work presented are only my contributions, unless otherwise mentioned.

Chapter 7 covers the measurement methodology used to measure the Drell–Yan angular coefficients, which was previously established within ATLAS by the $\sqrt{s} = 8$ TeV Z Drell–Yan angular coefficients analysis [6]. My work involved expanding the current methodology to work with W bosons, including having a method to incorporate events with no real solution for the neutrino p_z^ν solution. In Chapter 8, I contributed to testing different event selections to determine what selections would increase the statistical sensitivity to measure the Drell–Yan angular coefficients for the W boson. This chapter also includes some of my work towards determining how to appropriately treat the two neutrino p_z^ν solutions. These sensitivity studies were done alongside my colleagues, Ludovica Aperio Bella and Ruth Jacobs, who worked with me to establish how to properly reconstruct W boson events to measure the Drell–Yan angular coefficients. The work on the data-driven background estimation and building the framework to plot the kinematic distributions was performed by my colleague, Daniil Ponomarenko.

The remaining work in Chapters 9-11 is the measurement and uncertainties of the Drell–Yan angular coefficients and the differential cross-section, which was the majority

of my work during my PhD. program. This started with the original studies to ensure that there would be enough statistical sensitivity to measure the W boson Drell–Yan angular coefficients. It then expanded to incorporating the systematic uncertainties so they would be included within the fit and profiling the systematic variations to ensure they were properly modelled. The only exception is the original incorporation and profiling of the PDF uncertainties, which was done by my colleague, Grigori Tolachev. After ensuring that there was expected sensitivity using MC in place of data and that all of the systematic uncertainties were properly modelled, my work became extracting the measurements with real data to get the results presented in Chapters 9-11.

1.2 PhD. Program Experiences

PhD. program experiences and work extend past what is normally contained within a thesis. I will give a brief description of some of my experiences, such as presenting at conferences and working in the ATLAS control room. My PhD. program started with the opportunity of spending the first year doing research at CERN and allowed me to work alongside some of my colleagues, which I would not normally be able to do. This helped with my qualification task as well as allowed me to attend and present at different workshops at CERN. During my PhD. program, I have been able to make two other trips to CERN in 2022 and 2023. This provided me the opportunity to work alongside colleagues and contribute to data taking for ATLAS by working in the control room, where I did shifts at the Run Control and Trigger desks. I have also had the opportunity to present my work at many conferences and workshops, such as Winter Nuclear and Particle Physics Conference (WNPPC) in 2021 and Canadian Association of Physicists (CAP) Congress in 2022. I was also able to present about recent electroweak precision measurements on behalf of the ATLAS Collaboration at the Large Hadron Collider Physics (LHCP) Conference in June 2024.

CHAPTER 2

THE STANDARD MODEL

*Since all models are wrong the scientist must be alert to
what is importantly wrong.*

George E. P. Box, 1976 [19]

This thesis focuses on the measurement of the Drell–Yan angular coefficients for the W boson using a special data set taken during 2017-18 with the ATLAS detector. To be able to physically interpret the Drell–Yan angular coefficients, it is important to have an understanding of the particle interactions that contribute to the Drell–Yan process. This chapter will provide a brief explanation of the Standard Model (SM), which serves as the theoretical framework for elementary particles and their interactions. Using the information provided in this chapter, it is possible to understand the Drell–Yan process and by extension the Drell–Yan angular coefficients, later explained in Chapter 6 and measured in Chapters 9 and 11 of this thesis.

2.1 The Standard Model

There are four fundamental forces in nature: gravity, electromagnetism, the weak force, and the strong force. Gravity is described by general relativity through the curvature of space-time; whereas, the interactions between elementary particles via the latter three are described by the SM. At high enough energies, electromagnetism and the weak force are unified, which was first experimentally demonstrated in 1973 at CERN by measuring neutrino interactions mediated by the Z boson [20]. The strong force is described by quantum chromodynamics (QCD), but has not been experimentally shown to be unified with the electroweak theory. A Grand Unified Theory (GUT), encompassing the three forces other

than gravity, could occur at higher energies ($\sim 10^{16}$ GeV) [21] and would result in observable phenomena, such as the decay of the proton. Many observed cosmological phenomena are not described by the SM, such as the matter–antimatter asymmetry in the universe, dark matter, dark energy, and gravity.

Within the SM, there are 19 free parameters¹, such as the masses of some particles. These parameters do not have theoretically prescribed values and have to be experimentally measured, which constrains the SM. Some of these parameters can be measured via multiple methods and any inconsistencies in these measurements can shed light on how to improve the SM or point towards physics beyond the SM. This section will cover a brief overview of the description of elementary particles and their interactions based upon sources [21, 22, 23, 24], where more extensive explanations can be found.

Elementary particles observed in nature are categorized in two groups, fermions and bosons, based upon their spin. Most particles (x) will have antiparticles ($\bar{x}$) that have the same mass and lifetime. Charged antiparticles will have opposite electric charges, while uncharged particles are their own antiparticle². Matter is comprised of fermions, which have half-integer spin, $J = 1/2$, and interact via the fundamental forces mediated by bosons, which have integer spin, $J = 0, 1, 2$.

Fermions are subdivided into two groups, leptons – ℓ – and quarks – q . A summary of all known SM fermions and some of their properties can be found in Table 2.1. Fermions also have two different chiral states, left and right-handed. For massless particles chirality is equal to its helicity, which relates the particles spin and momentum. The particles produced from high energy interactions, as is the case in this thesis, have energy much greater than their mass, so they can be approximated as massless. Therefore, if a particle’s spin is in the same direction as the momentum, it can be considered as right-handed and if they are in an opposite direction, it can be considered left-handed.

¹This number is for the case of massless neutrinos. If you extend the SM to have massive neutrinos, additional free parameters are needed to describe their masses and the mixing of states. This falls outside of the scope of this thesis.

²This depends on the open question whether neutrinos follow Majorana or Dirac statistics.

Fermions	Generation	Flavour	Charge (e)	Mass
Leptons	1st	e	-1	$0.51099895000 \pm 0.00000000015$ MeV
		ν_e	0	-
	2nd	μ	-1	$105.6583755 \pm 0.0000023$ MeV
		ν_μ	0	-
	3rd	τ	-1	1776.86 ± 0.12 MeV
		ν_τ	0	-
Quarks	1st	u (up)	2/3	$2.16^{+0.49}_{-0.26}$ MeV
		d (down)	-1/3	$4.67^{+0.48}_{-0.17}$ MeV
	2nd	s (strange)	-1/3	$93.4^{+8.6}_{-3.4}$ MeV
		c (charm)	2/3	1.27 ± 0.02 GeV
	3rd	b (bottom)	-1/3	$4.18^{+0.03}_{-0.02}$ GeV
		t (top)	2/3	172.69 ± 0.30 GeV

Table 2.1: Table summarizing the charge and mass of known fermions, grouped by leptons and quarks, and further sub-grouped by generation based on mass. Mass values are in natural units, where $c = 1$, and charges in units of the elementary charge, e . The neutrinos' masses are much smaller than the other fermions and are not included in this table, as they can be treated as massless in most cases. Values taken from Particle Data Group (PDG) [25].

There are six observed lepton flavours: electron e , muon μ , and tau τ , which have charge $Q = -1e$ and their corresponding neutral neutrinos ν_e , ν_μ , and ν_τ . Here, e is the elementary charge and taken to be positive. All six are able to interact via the weak force and only the charged leptons can interact via the electromagnetic force. The electromagnetic and weak interactions are described in more detail in Section 2.2.2.

There are six observed quark flavours: up, down, strange, charm, bottom, and top. The quarks carry fractional elementary charge of either $Q = 2/3e$ or $Q = -1/3e$, allowing them to interact electromagnetically. Quarks also contain an additional quantum number, called colour, which has three states: red, green, and blue. This colour charge is what allows quarks to not violate the Pauli Exclusion Principle when in bound states and is exchanged between quarks via gluons.

Elementary bosons are observed to come in two varieties, vector ($J = 1$) and scalar ($J = 0$). Vector bosons are the mediator particles required for interactions between fermions and are able to mediate interactions for other bosons. The Higgs boson is the only observed elementary scalar boson and was first observed in 2012 by ATLAS and CMS Collabora-

Force	Boson	Charge (e)	Mass
Gravity	Graviton	0	0
Electromagnetic	Photon (γ)	0	0
Weak	W^+	+1	$80.377 \pm 0.012 \text{ GeV}$
	W^-	-1	
	Z	0	$91.1876 \pm 0.0021 \text{ GeV}$
Strong	Gluon	0	0
-	Higgs	0	$125.25 \pm 0.17 \text{ GeV}$

Table 2.2: Table of known bosons where mass values are in natural units, where $c = 1$, and charges in units of the elementary charge, e . Values from Particle Data Group (PDG) [25]. The graviton has not been observed and would be the boson of a field theory that describes gravity with a spin ($J = 2$).

tions [26, 27]. Unlike vector bosons, the Higgs boson does not mediate a force, instead it is an excitation of the Higgs field, which gives particles their mass. A summary of the known bosons can be found in Table 2.2.

The photon, γ , mediates the electromagnetic interaction between charged particles, as described by quantum electrodynamics (QED). The weak interaction is mediated by the $W^\pm$ and Z bosons, and interact with all fermions and bosons, except for the gluon. Lastly, the gluon, g , is the mediator for the strong force for that describes interactions between coloured particles: quarks, antiquarks, and itself, as described by QCD. A hypothetical boson, called the graviton, would be the mediator boson for a field theory that describes gravity in a quantized approach. The photon, gluon, and Higgs boson interact with both fermion chiral states evenly, which is not the case for the weak bosons. Instead, the Z boson has a different coupling strength for left and right-handed fermions, while W bosons only couple to left-handed fermions and right-handed anti-fermions.

2.2 Quantum Field Theory

The SM is built upon the basis of quantum field theory (QFT), which combines the ideas of quantum mechanics and special relativity [28, 29]. QFT is approached by extending classical Lagrangian mechanics to describe fields. Classical Lagrangian mechanics describes the mechanics of a system described by the Lagrangian function, $L(q, \dot{q})$, where q is a set of coordinates and $\dot{q}$ is their first time derivative. A simple case of this is the kinetic energy minus a potential term: $L(q, \dot{q}) = m\dot{q}^2/2 - V(q)$. The dynamics of the system are described by the action, S , which is the time integral

$$S = \int dt L(q, \dot{q}), \quad (2.1)$$

and should follow an extremum so that the change in the action, $\delta S = 0$, is minimized and is known as the principle of least action. This leads to the Euler-Lagrange equations for each $q_i \in q$

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0. \quad (2.2)$$

The Euler–Lagrange equations for each coordinate provides the equations of motion for the Lagrangian.

Field theories extend the coordinates of time, $q(t)$, to be fields of space-time $\phi(x)$. For a local field theory, the Lagrangian can be written in terms of a Lagrangian density, denoted by $\mathcal{L}$. The general form of the action becomes

$$S = \int dt L = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi). \quad (2.3)$$

Similarly, Euler–Lagrange equations of motion can be found for the field

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0. \quad (2.4)$$

Equation 2.4 shows that, with a given Lagrangian density, it is possible to find the equations of motion for the field, similar to how it is possible in classical mechanics with the Lagrangian in Equation 2.2. Because of this, the Lagrangian density is often referred to as the Lagrangian in field theory.

An example of a Lagrangian for a real massive scalar field, ϕ , is of the form

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2, \quad (2.5)$$

where $\partial_\mu\phi$ describes the dynamics of the field, similar to the momentum, and m is mass of the field. The Euler–Lagrange equations of motion results in the Klein–Gordon equation for a free particle

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\phi = 0. \quad (2.6)$$

This example is not realistic but is useful to explain the anatomy of a Lagrangian to understand the form of terms. There are two important types of terms, mass and interaction. For a field X , the mass terms will have the form: CXX for a real scalar field, $CX^\dagger X$ for a complex scalar field, $C\bar{X}X$ for a spinor field, or $CX_\mu X^\mu$ for a vector field. In the previous lists of mass terms, C , is the mass of the fields, $\bar{X} = X^\dagger\gamma^0$ is the adjoint of the field, $\dagger$ is the Hermitian conjugate (complex conjugate and transpose), and γ^0 is a Dirac matrix (which will be discussed in Section 2.2.2). The adjoint or the Hermitian conjugate are required for the fields to remain Lorentz invariant, but specifics of these mathematical operations are not important for this thesis.

Interaction terms in a Lagrangian have the general form $cX_1X_2..X_N$, where N is the number of fields needed, which is 3 or 4 for most interactions. In the Lagrangian example in Equation 2.5, there is no term for the field to interact as the Lagrangian does not have terms with three or more fields. This results in the Klein–Gordon equation for a free particle in Equation 2.6. The fields in interaction terms can be different combinations of scalar, spinor, vector, or their adjoints and Hermitian conjugates. Here, c , is a coupling constant, which

provides the strength of the interaction for each vertex of the field interaction.

Interaction terms can be better understood with two examples of a quartic and cubic coupling. The first example of field interactions is to consider the real scalar field interacting with itself via a quartic coupling. This interaction term is of the form

$$\mathcal{L}_{\text{int}} = -\frac{\lambda}{4!}\phi^4, \quad (2.7)$$

where λ is a coupling constant and $4!$ is a normalization factor. A simple example of this is two fields interacting in $2 \rightarrow 2$ scattering, as seen in the Feynman diagram in Figure 2.1 where fields $\phi(1)$ and $\phi(2)$ interact to produce fields $\phi(3)$ and $\phi(4)$. Feynman diagrams are used to visualize interactions and calculate their rates, the former will be the focus of this section. The diagrams do not contain any space-time information, so any angle between lines can not be used to calculate any vector quantities, such as momentum. Instead they are used to show the time order that interactions happen in, where the convention used in this thesis is time flows from left to right. In Feynman diagrams, each line correlates to a field in the Lagrangian and the vertices of lines correspond to a Lagrangian interaction term. For this example, each ϕ in Equation 2.7 represents one line in the vertex.

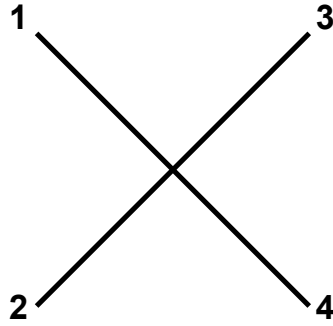


Figure 2.1: An example of a Feynman diagram for a quartic coupling, where time flows left to right. This diagram has four fields meeting at a vertex and shows the time order of physics processes but does not contain any space-time information.

Interactions do not just happen between fields of the same type, instead they will often come in the form

$$\mathcal{L}_{\text{int}} = gX\phi^2, \quad (2.8)$$

where g is the coupling constant and the two real scalar fields form a vertex with the field X . An example $2 \rightarrow 2$ interaction using cubic terms produces multiple Feynman diagrams, such as Figure 2.2a) for annihilation and Figure 2.2b) for scattering. For the first, $\phi(1)$ and $\phi(2)$ annihilate to produce X in one vertex, then X decays into $\phi(1)$ and $\phi(2)$ in the second vertex. In the second case, $\phi(1)$ and $\phi(2)$ interact by exchanging X has one vertex at the top and one at the bottom. These examples serve as a basis to understand Lagrangian equations and Feynman diagrams for actual physical processes when they are presented in future sections.

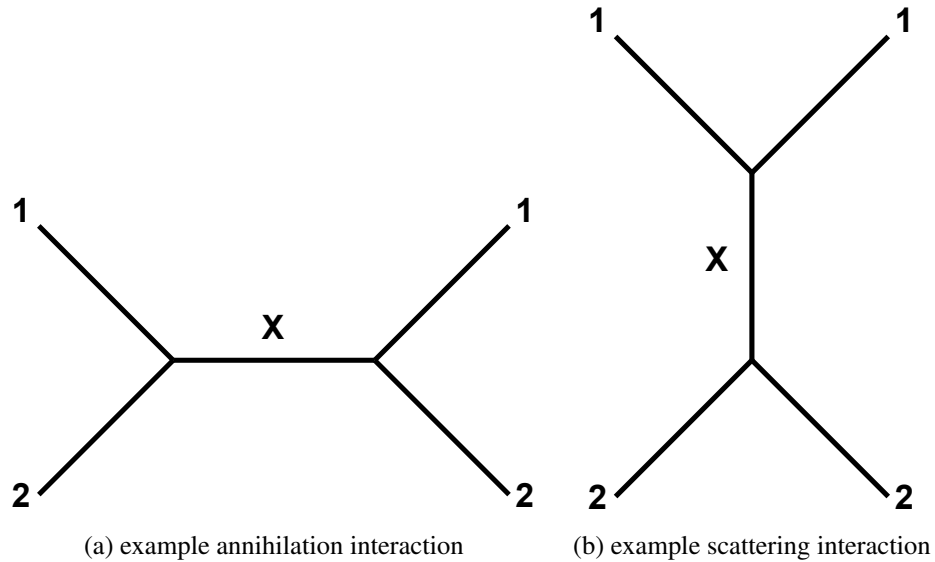


Figure 2.2: Example Feynman diagrams for cubic couplings, where time flows left to right. a) Shows three fields meeting at a vertex where two particles annihilate to produce the third field, which decays into two particles at another three field vertex. b) Shows two particles scattering, which is mediated by another particle. Each scattering particle has their incoming and outgoing field lines forming a cubic vertex with the mediating particle.

2.2.1 Gauge Theories

The SM contains gauge fields, which require the Lagrangian to not change under local transformations (a transformation at every point in space-time). Consider the Dirac Lagrangian for a free fermion ψ , such as an electron, of mass m :

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\not{D} - m)\psi, \quad (2.9)$$

where the slash notation of $\not{a} \equiv a_\mu \gamma^\mu$ is used³. A general gauge transformation can be performed on ψ

$$\psi \rightarrow e^{iQ\lambda(x)}\psi, \quad (2.10)$$

where Q is the quantization of the field. If $\lambda(x)$ is a constant real value, then it is a global transformation and the Lagrangian is invariant. However, this is not the case for a local transformation. To achieve invariance for a local transformation, the Lagrangian in Equation 2.9 needs to be modified with two changes. First, the derivative ∂_μ is replaced with the covariant derivative

$$\partial_\mu \rightarrow \partial_\mu + iQA_\mu = \mathcal{D}_\mu, \quad (2.11)$$

where A is a vector field and transforms as $A_\mu \rightarrow A_\mu + \partial_\mu \lambda$. The second change is to add another kinematic term for the gauge field

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \quad (2.12)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength. The addition of the field strength kinematic term and changing to the covariant derivative allows this local transformation to be Lorentz invariant. This example describes the electron interacting with the photon gauge field, A_μ , and can be used to understand the method when applied for other gauge fields.

³The Einstein summation convention is used where contracted indices are summed over in the range $a = 1, 2, 3$ for Latin characters and $\mu = 0, 1, 2, 3$ for Greek characters. γ^μ is the Dirac gamma matrices, which will be defined later in Equation 2.16.

The photon gauge field is a simpler example as it corresponds to the single generator for the $U(1)$ group, while other Lie groups, such as $SU(2)$ or $SO(3)$, will have more generators. With more generators, there will be more gauge fields and corresponding gauge bosons. Each of these gauge fields will also add to the covariant derivative and the field strength in order to keep it Lagrangian invariant.

Lastly, from Noether's theorem [30] if there is a continuous global symmetry for the action, seen in Equation 2.1, then there is a conserved current and conserved charge for the solutions of the equations of motion. In this example of an electron interacting with the photon gauge field, the transformation leaves the Lagrangian invariant and there is a conserved charge, in this case the electric charge. The interaction term of this Lagrangian from the covariant derivative can be written in terms of the conserved electromagnetic (EM) current

$$\mathcal{L}_{\text{EM}}^{\text{int}} = J_{\mu}^{\text{EM}} A^{\mu}, \quad (2.13)$$

where $J_{\mu}^{\text{EM}} = Q\bar{\psi}\gamma_{\mu}\psi$. Writing in terms of the current is useful for separating the interaction into the gauge boson and the conserved quantity.

2.2.2 Electroweak Theory and Spontaneous Symmetry Breaking

This section will cover the unification of QED and the weak force based on sources [24, 28, 29]. The electroweak interaction is a product of the weak, $SU(2)$, and the electromagnetic, $U(1)$, gauge groups mixing to form the electroweak $SU(2) \times U(1)$ interaction. For the electroweak interaction, there are three weak isospin bosons $\vec{W}^{\mu} = \{W_1^{\mu}, W_2^{\mu}, W_3^{\mu}\}$ from the $SU(2)$ group and one weak hypercharge boson B^{μ} from the $U(1)$ group. From Noether's theorem in Section 2.2.1, there is a conserved charge for each of these bosons. The weak group has the quantum number weak isospin, T , which has three components (one for each boson) and the hypercharge boson has the quantum number hypercharge, Y . The third component of weak isospin and hypercharge are related to the electric charge of a particle through the relation $Q = T^3 + \frac{Y}{2}$. The electroweak interaction Lagrangian

defines how the fermions in the weak isospin and weak hypercharge currents interact with the associated gauge bosons:

$$\mathcal{L}_{EW}^{int} = g \vec{J}_\mu \cdot \vec{W}^\mu + g' J_\mu^Y B^\mu, \quad (2.14)$$

where g and g' are coupling constants that give the strength of the interaction for each current. The isospin currents, $\vec{J}_\mu$, are defined as⁴:

$$\begin{aligned} J_\mu^+ &= \frac{1}{2} \bar{\nu}_e \gamma_\mu (1 - \gamma^5) e, \\ J_\mu^- &= \frac{1}{2} \bar{e} \gamma_\mu (1 - \gamma^5) \nu_e, \\ J_\mu^3 &= \frac{1}{2} \bar{e} \gamma_\mu (v - a \gamma^5) e. \end{aligned} \quad (2.15)$$

The weak hypercharge current J_μ^Y is defined in terms of isospin current J_μ^3 and the electromagnetic current J_μ^{EM} , discussed in Section 2.2.1. They are related in the same proportionality as the electric charge to the third component of weak isospin and hypercharge, discussed earlier, resulting in $J_\mu^{EM} = J_\mu^3 + \frac{J_\mu^Y}{2}$. In Equation 2.15, v and a are the vector and axial–vector couplings, respectively. These provide the coupling strength of the W_3^μ weak isospin boson to the vector current ($\bar{e} \gamma_\mu e$) and axial–vector current ($\bar{e} \gamma_\mu \gamma^5 e$), respectively. The e operator corresponds to creating a positron or destroying an electron. The opposite is true for the $\bar{e}$ operator. Likewise, $\bar{\nu}_e$ creates a neutrino or destroys an antineutrino. Lastly, γ^μ and γ^5 are the Dirac matrices defined in the Bjorken and Drell basis [21] as:

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & -\mathbb{1} \end{pmatrix}; \quad \gamma^i = \begin{pmatrix} \mathbb{0} & \tau_i \\ -\tau_i & \mathbb{0} \end{pmatrix}; \quad \gamma^5 = \begin{pmatrix} \mathbb{0} & \mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix}. \quad (2.16)$$

In these instances, $i = \{1, 2, 3\}$, $\mathbb{1}$ is the 2×2 unit matrix, $\mathbb{0}$ is a 2×2 matrix of zeros, and

⁴All currents presented in this section are for the electron, e , and the electron neutrino, ν_e , interactions for simplicity. The currents are similar for other fermions by replacement of particles with the exception being quarks and antiquarks. They gain a mixing factor but do not change the functional form of the current.

τ_i are the Pauli matrices:

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.17)$$

The physical bosons, W^+ , W^- , Z , and the photon, are not the weak isospin and the hypercharge bosons, instead, they are mixed states of them:

$$\begin{aligned} W^{\pm,\mu} &= \frac{1}{\sqrt{2}} (W_1^\mu \mp iW_2^\mu), \\ Z^\mu &= W_3^\mu \cos \theta_W - B^\mu \sin \theta_W, \\ A^\mu &= W_3^\mu \sin \theta_W + B^\mu \cos \theta_W. \end{aligned} \quad (2.18)$$

As discussed in Section 2.2.1, A^μ is the photon field. The amount the weak isospin and hypercharge mix is parameterized by the mixing angle, θ_W . This mixing angle is a free parameter of the SM and is used to relate the elementary electric charge, e , to the weak isospin and hypercharge coupling constants, g and g' , respectively, through:

$$g \sin \theta_W = g' \cos \theta_W = e. \quad (2.19)$$

The mixing of states in Equation 2.18 is a result of the spontaneous symmetry breaking of the electroweak theory caused by the Higgs field. This symmetry breaking is also the reason the W and Z bosons are not massless like the photon and gluon. To better understand spontaneous symmetry breaking, consider the common physical example of a pencil balanced on its tip. The pencil is rotationally symmetric as it appears the same regardless of which direction you spin it and it has the potential to fall in any direction in a circle. When the pencil falls, a direction is chosen and the symmetry is spontaneously broken. Similarly, the Higgs field, Φ , can first be expressed as a general complex scalar doublet

before symmetry breaking:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (2.20)$$

where ϕ^+ and ϕ^0 are a charged and neutral complex scalar field, respectively. Each has a real, ϕ_1, ϕ_3 , and imaginary, ϕ_2, ϕ_4 , field component in order to keep the theory general. The Lagrangian for this field can be expressed with a kinetic term and a potential term:

$$\mathcal{L} = (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}^\mu \Phi) - V(\Phi^\dagger \Phi), \quad (2.21)$$

where the covariant derivative, $\mathcal{D}_\mu$, is responsible for the kinematics, as described in Section 2.2. The covariant derivative also needs to be expressed so that the Lagrangian is invariant under the $SU(2) \times U(1)$ gauge transformation, as discussed in Section 2.2.1. It is defined as:

$$\mathcal{D}_\mu \Phi = \left(\partial_\mu + \frac{ig}{2} \tau_a W_\mu^a + \frac{ig'}{2} B_\mu \right) \Phi. \quad (2.22)$$

As discussed earlier, W_μ^a are the fields for $SU(2)$ and B_μ is the field for $U(1)$.

The potential term, $V(\Phi^\dagger \Phi)$, can be expressed in powers of $(\Phi^\dagger \Phi)$ so that it is gauge invariant. To keep the theory renormalizable, orders above two are not allowed. The potential is then defined as:

$$V(\Phi^\dagger \Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad (2.23)$$

where λ and μ^2 are constants. To keep the vacuum stable, it is required that the potential is always positive, which requires $\lambda > 0$. There are two possible signs of μ^2 , when $\mu^2 > 0$ or $\mu^2 < 0$. When $\mu^2 > 0$, the potential is parabolic with the minimum at $\Phi = 0$. The second case of $\mu^2 < 0$, the potential is sombrero shaped, as seen in Figure 2.3, and the minimum can be found by taking the derivative of the potential with respect to $(\Phi^\dagger \Phi)$ and setting it

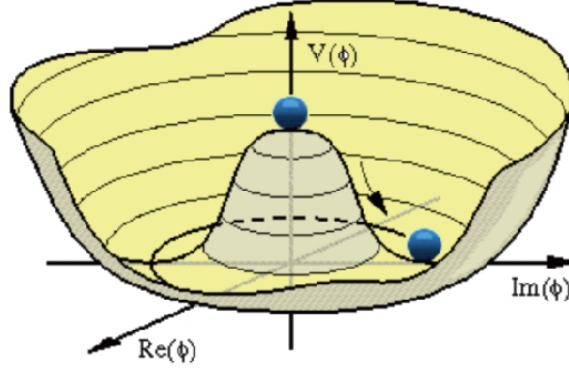


Figure 2.3: A depiction of the Higgs potential taken from [31]. The Higgs potential takes the famous sombrero shape and the symmetry is broken when a vacuum is chosen, which results in a minimum at the vacuum expectation value.

equal to zero. The minimum is a circle along the collection of points

$$\Phi^\dagger \Phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda}. \quad (2.24)$$

This minimum value, v , is the state of the lowest energy and is known as the vacuum expectation value (VEV),

$$\langle \Phi \rangle = v \equiv \sqrt{-\frac{\mu^2}{\lambda}}. \quad (2.25)$$

The potential and the minimum are rotationally symmetric, so a direction can be chosen to simplify Equation 2.24 without a loss of generality. Choosing $\phi_1 = \phi_2 = \phi_4 = 0$, results in $\phi_3^2 = v^2$ and the field Φ can be expressed in terms of the VEV,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.26)$$

Like the example of the pencil falling, which chooses a direction and breaks the symmetry, a direction is chosen by setting all but ϕ_3 to zero. Choosing a specific vacuum breaks three of the four symmetries of the $SU(2) \times U(1)$ group, which results in three massive gauge bosons. This can be seen by substituting Φ in terms of the VEV in Equation 2.26 into Equation 2.21 and rearranging to find just the mass terms. Recall from Section 2.2 that

mass terms will be proportional to only two fields so the relevant part of the Lagrangian is then:

$$\mathcal{L}_M = \frac{1}{8}v^2g^2 [(W_\mu^1)^2 + (W_\mu^2)^2] + \frac{1}{8}v^2 [g^2(W_\mu^3)^2 - 2gg'W_\mu^3B^\mu + g'^2(B_\mu)^2]. \quad (2.27)$$

Looking at the collection of terms in the first set of square brackets, it can be rewritten after substituting the mixing relations for the W bosons from Equation 2.18

$$\frac{1}{8}v^2g^2 [(W_\mu^1)^2 + (W_\mu^2)^2] = \left(\frac{1}{2}vg\right)^2 W^+W^-. \quad (2.28)$$

Here, the constant, $vg/2$, in front of $W^+ W^-$ is $(m^W)^2$, resulting in the W boson mass:

$$m^W = \frac{1}{2}vg. \quad (2.29)$$

The collection of terms in the second set of square brackets in Equation 2.27 is associated to the mass terms for the Z boson and the photon, A_μ . To find the mass terms, it can be cleverly rearranged to be in the form:

$$\frac{1}{8}v^2 [g^2(W_\mu^3)^2 - 2gg'W_\mu^3B^\mu + g'^2(B_\mu)^2] = \frac{1}{8}v^2 [gW_\mu^3 - g'B_\mu]^2 + 0[g'W_\mu^3 - gB_\mu]^2. \quad (2.30)$$

Like in the case of the W boson, Equation 2.18 can be used to express Equation 2.30 in terms of Z_μ and A_μ , with the help of Equation 2.19, to arrive at:

$$\begin{aligned} \frac{1}{8}v^2 [gW_\mu^3 - g'B_\mu]^2 &= \frac{1}{2}(m^Z)^2 Z_\mu^2, \\ 0 [g'W_\mu^3 - gB_\mu]^2 &= \frac{1}{2}(m^A)^2 A_\mu^2. \end{aligned} \quad (2.31)$$

The second term shows the mass of the photon field to be zero, as expected from observations, and the mass of the Z boson is:

$$m^Z = \frac{1}{2}v\sqrt{g^2 + g'^2}. \quad (2.32)$$

Using Equation 2.19, the W and Z boson masses are related by $m^W = m^Z \cos \theta_W$. By choosing a vacuum, the spontaneous symmetry is broken resulting in the physical bosons being a mixture of the four gauge fields and making the W and Z bosons massive, as seen in nature. The masses of the W and Z bosons come naturally out of the theory and depend on the VEV, which has been determined experimentally to be $v \approx 246$ GeV [25].

2.2.3 Electroweak Interactions

Understanding where the physical bosons come from, the electroweak Lagrangian, from Equation 2.14, for interactions between fermions can be expressed in terms of the physical bosons and currents:

$$\mathcal{L}_{EW}^{int} = \frac{g}{\sqrt{2}} (J_\mu^- W^{+\mu} + J_\mu^+ W^{-\mu}) + \frac{g}{\cos \theta_W} (J_\mu^3 - \sin^2 \theta_W J_\mu^{EM}) Z^\mu + e J_\mu^{EM} A^\mu. \quad (2.33)$$

To understand how the interactions in Equation 2.33 physically materialize, examples of an interaction mediated by each boson will be given. First, the weak charged current, J_μ^- in Equation 2.15, is responsible for nuclear beta-minus decay, as seen in Figure 2.4. This occurs when a down quark in a neutron changes flavour to an up quark by emitting a W^- boson, which immediately decays into an electron and an anti-neutrino. The coupling of the W^- , electron, and anti-neutrino is the $\frac{g}{\sqrt{2}} J_\mu^+ W^{-\mu}$ term in Equation 2.33. Similarly, beta-plus decay would be mediated by the second term in Equation 2.33. The neutral weak current is mediated by the Z boson for processes like the Drell-Yan process of $q\bar{q} \rightarrow Z \rightarrow e^+e^-$, as seen in Figure 2.5. Here, the coupling of the Z boson to the electron-positron pair is the middle J_μ^3 term in Equation 2.33. Lastly, the electromagnetic

current, J_μ^{EM} , results in a very similar diagram as Figure 2.5, but the Z boson is replaced by the photon, as seen in Figure 2.6. The coupling of the photon to the electron-positron pair comes from the last term of Equation 2.33.

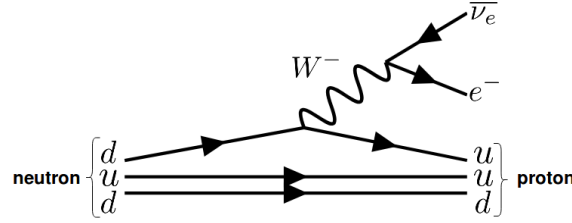


Figure 2.4: Feynman diagram of nuclear beta decay, where a down quark in a neutron emits a W^- boson to change flavour into an up quark. The W^- boson then promptly decays into an electron and an anti-neutrino. This W^- boson emission is an example of one of the weak charged currents.

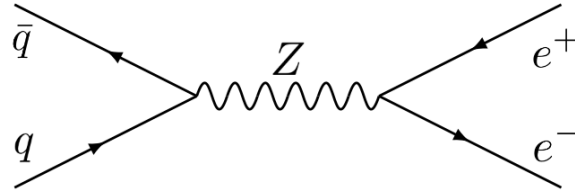


Figure 2.5: Feynman diagram of the Drell–Yan process of quark-antiquark pair annihilation into a Z boson before decaying into an electron-positron pair.

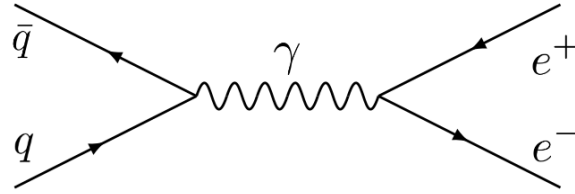


Figure 2.6: Feynman diagram of a quark-antiquark pair annihilation into a photon before decaying into an electron-positron pair.

The mixing of the weak and electromagnetic currents in the middle term for the Z boson in Equation 2.33 allows for coupling both left and right-handed fermions, while the W bosons can only couple to the left-handed fermions.

2.2.4 Strong Interactions

Quarks carry the same quantum numbers as other fermions but have an additional quantum number, called colour. Colour comes in three types: red, green, and blue. This colour charge is exchanged between quarks through gluons, which simultaneously carry colour and anticolour, as described by QCD. So far, everything containing quarks and gluons have only been observed in bound states, called hadrons. This notion of no free quarks or gluons in nature is known as confinement. The most common bound state is a baryon, which has three quarks with different colours, such as a proton (uud) or a neutron (udd). The other common bound state is a meson, which is comprised of a quark–antiquark pair carrying colour and anticolour, such as a π^+ ($u\bar{d}$). More exotic hadronic bound systems, called tetraquarks ($qq\bar{q}\bar{q}$) and pentaquarks ($qqqq\bar{q}$), have been observed [32, 33, 34] by the LHCb Collaboration at CERN. The quarks and antiquarks that form bound states, through gluon exchange, are called valence quarks.

The bound states are a consequence of the strong coupling “constant”, α_S , which defines how strongly quarks and gluons interact. This coupling “constant” depends on the momentum transfer⁵, Q , between quarks and gluons. When quarks are at very short distances, the exchanged gluons have a large momentum transfer and the quarks behave approximately as free particles, known as asymptotic freedom. At larger distances, the momentum transfer decreases and α_S increases to try and pull the quarks back together, like a spring. With a large enough α_S , it becomes energetically favourable to produce a quark-antiquark pair to lower the energy in the system. This process is called hadronization. The strong coupling “constant” being a function of Q^2 , $\alpha_S(Q^2)$, has a running behaviour, where values are experimentally determined from various sources, as seen in Figure 2.7.

⁵Not to be confused with the electric charge, Q .

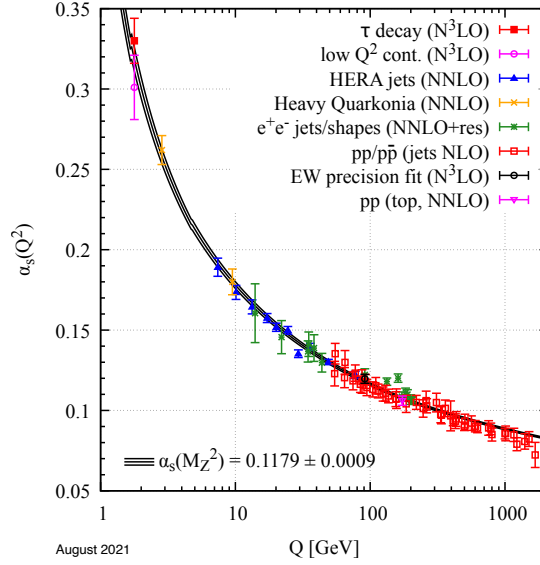


Figure 2.7: Running strong coupling “constant”, α_s , which decreases with increased momentum transfer, Q , and separation distance. Plot taken from Particle Data Group (PDG) [25].

Hadrons are comprised of more than just their valence quarks and the gluons that bind them together. Gluons can be emitted from valence quarks and produce a quark antiquark pair, called sea quarks, before being reabsorbed, as seen in in Figure 2.8. These sea quarks, valence quarks, and gluons are all part of a larger group, called partons, which make up hadrons. When hadrons interact, any of the partons are able to interact to produce other particles.

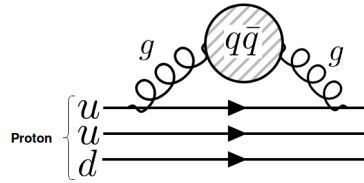


Figure 2.8: An example of a Feynman diagram where a gluon is emitted within a proton to form sea quarks before being reabsorbed.

2.2.5 Summary

This chapter summarizes observed elementary particles and their interactions as described by the SM. With the information provided it will be possible to understand the remainder of this thesis. Next, how these particles are measured in the ATLAS detector will be discussed in Chapter 3. Then how these particles are reconstructed for physics analysis will be discussed in Chapter 4. The information in this chapter will also make it possible to understand the Drell–Yan process described in Chapter 6.

CHAPTER 3

LHC AND THE ATLAS DETECTOR

The measured properties of particles depend on their energy and the method with which they are detected. This chapter will cover the acceleration of bunches of protons with the Large Hadron Collider (LHC), which are subsequently collided at four interaction points in detectors, such as the ATLAS (**A** Toroidal **L**HC **A**pparatu**S**) detector. This thesis utilizes data taken with the ATLAS detector, therefore, this chapter will explain the different subdetectors within the ATLAS detector and how its triggering system works.

3.1 LHC

The LHC [35], as seen in Figure 3.1, is located at Conseil Européen pour la Recherche Nucléaire (CERN) on the border of France and Switzerland, outside of Geneva. It is a 26.7 km circumference ring that was designed to accelerate and collide protons up to a centre-of-mass energy $\sqrt{s} = 14$ TeV. Every 25 ns, bunches of approximately 10^{11} protons are collided within four main detectors along the ring. The LHC started collecting data in 2009 for beam and detector calibration at $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 2.36$ TeV. During Run 1, from 2010-12, data sets were taken at $\sqrt{s} = 7, 8$ TeV before a long shut down, from 2013-15, to upgrade the accelerator and detectors. This upgrade of the accelerator allowed the LHC to operate at a higher centre-of-mass energy of $\sqrt{s} = 13$ TeV during Run 2, from 2015-18. From 2019-21, another long shutdown occurred to perform more accelerator and detector upgrades. These allowed the machine to operate closer to the designed centre-of-mass energy of $\sqrt{s} = 14$ TeV. From 2022 to present, the LHC and the experiments are in Run 3, operating at $\sqrt{s} = 13.6$ TeV. Another long shut down is planned for the start of 2026 for upgrades to operate with a higher number of interactions per proton bunch crossing.

In addition to the proton–proton program, the LHC has a heavy ion program. This

The CERN accelerator complex *Complexe des accélérateurs du CERN*

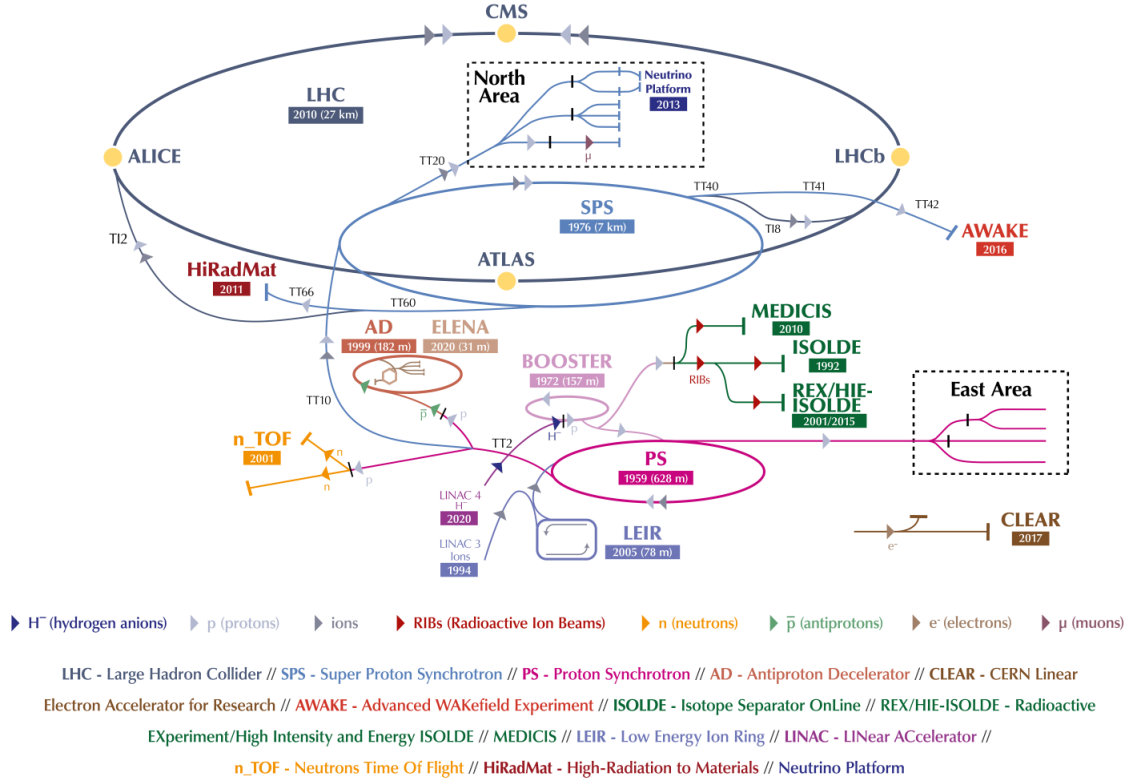


Figure 3.1: Schematic overview of the CERN complex [36], including the LHC. Proton–proton collisions are produced by ionizing hydrogen to isolate the protons before being accelerated through the path LINAC4, BOOSTER, PS, SPS, and LHC. Beams are then collided in experiments, such as ATLAS.

LHC program accelerates heavy ions, such as Pb and Xe, at centre-of-mass energies of $\sqrt{s} = 2.76, 5.02, 5.36, 5.44$, and 8.16 TeV per nucleon, where a nucleon is either a proton or neutron. The heavy ion program is tangentially related and not further considered in this thesis, apart from the fact that the heavy ion program requires corresponding proton collision data at the same energies that can be used for electroweak physics, such as W and Z boson measurements.

The experiments at the interaction points of the LHC include ALICE [37], ATLAS [1], CMS [38], and LHCb [39] as seen in Figure 3.1. ATLAS and CMS are the largest collaborations at the LHC and are multipurpose detectors, located on opposite sides of the ring.

Both experiments aim to analyze a wide range of particle physics processes, which can be used for precision measurements of the SM and searches for new processes beyond the SM. To achieve this, these detectors record as much useful information possible to measure a wide scope of interesting physics events. The experiment LHCb is a more specialized detector that focuses its study on hadrons comprised of b quarks, such as $b\bar{b}$ mesons, that are produced at low angles from the beamline. LHCb is a single-arm spectrometer with an interaction point displaced by 11.25 m from the centre, allowing for better positional tracking close to the beamline. ALICE is designed for precision measurements of thousands of tracks, which are discussed later, per event from heavy ion collisions. Other experiments, such as LHCf [40], FASER [41], and TOTEM [42], are not positioned on interaction points of the detector, but are further away to detect particles in regions with very small angles from the beamline or very long-lived particles. This thesis will focus on ATLAS as it utilizes data measured with the ATLAS detector.

Before bunches of protons collide in a detector, they must go through multiple accelerations steps to achieve higher energies, as seen in Figure 3.1. First, hydrogen gas is ionized to separate protons from the bound electrons. The protons are formed into a beam comprised of the bunches and are then accelerated using radiofrequency cavities up to an energy of 50 MeV. This happens before being sent to the proton synchrotron booster to accelerate them up to an energy of 1.4 GeV. The proton beam is then accelerated by the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) up to energies of 25 GeV and 450 GeV, respectively. In addition to accelerating the protons, the SPS is also responsible for spacing the protons into bunches. The 450 GeV proton bunches are then injected into the LHC ring to undergo its last acceleration up to a maximum energy of 7 TeV per beam. When injecting the proton bunches into the LHC, they are injected into two beamlines, which accelerate the bunches in opposite directions to maximize the centre-of-mass energy per collision. There is a designed 25 ns spacing between bunches per beam, but there are additional spaces added for other purposes.

3.2 Counting Interactions

Searching for rare processes and measuring more common processes, like the production of the W and Z bosons, benefit from having a large data set to increase the number of interactions. This increases the likelihood and statistical significance of observing the desired process. For a process x , the rate of production dN_x/dt , is

$$\frac{dN_x}{dt} = \sigma_x L, \quad (3.1)$$

where σ_x is the production cross-section and L is the instantaneous luminosity, which acts as a proportionality factor. The calculation for instantaneous luminosity is non-trivial and involves knowing the number of protons in the bunch, LHC revolution frequency, beam crossing angle, and the beam density [43]. Characteristics of the beams are found through various methods, such as performing van der Meer (vdM) scans to measure the beam profiles [44]. The LHC was originally designed to have an instantaneous luminosity of $1.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, but this was surpassed during Run 2 by a factor of two [45] and has been consistently at that new value during Run 3 [46]. This higher instantaneous luminosity allows for more data to be collected. It is more useful to report the size of data sets in terms of the integrated luminosity, which is the time integrated instantaneous luminosity. During Run 2, ATLAS recorded an integrated luminosity of approximately 140 fb^{-1} at $\sqrt{s} = 13 \text{ TeV}$, which is useful for physics measurements [45]. To be useful for physics measurements, the delivered luminosity relies on the efficiency of the experiment to record good data. A subset of this Run 2 data set was taken under special running conditions and is used for this analysis. More details for this special data set will be discussed in Section 8.1.

Proton bunches are used to increase the chances of interactions to occur, as only a small fraction of the protons interact per bunch crossing with the majority continuing to circulate. This slowly depopulates the beams and it becomes favourable to dump the beams and restart the process with new beams. A higher instantaneous luminosity leads to an increased

mean number of interactions per bunch crossing, μ . For each data set, the average of the mean number of interactions per bunch crossing, $\langle\mu\rangle$, is also recorded. The instantaneous luminosity increased during Run 2, which changed the μ that ATLAS recorded as they are related through the expression [11]:

$$L = \frac{\mu f_r}{\sigma_{\text{vis}}}, \quad (3.2)$$

where f_r is the LHC bunch revolution frequency, and σ_{inel} is the reference inelastic cross-section. Figure 3.2 shows the instantaneous luminosity as a function of μ for Runs 1,2, and the start of Run 3. An increased μ creates more events in the detector, which makes event reconstruction more challenging because the amount of pileup. These pileup events come in two forms, in-time and out-of-time pileup. When multiple interactions from the same bunch crossing pass through the same region of the detector they are referred to as in-time pileup. Out-of-time pileup occurs when an event from a later bunch crossing passes through the same region of the detector and going through undetected. Here there is an overlap of signal between bunch crossings because some subdetectors take up to several hundred nanoseconds to record the signal [47].

3.3 ATLAS Detector

The ATLAS detector [1] is one of the main experiments at CERN using the LHC ring. The detector utilizes a cylindrical geometry concentric to the beamline, as seen in Figure 3.3, which is subdivided into three regions: the barrel and two endcaps. Separate detection methods are used for the different particles measured by the ATLAS detector. To do this, the ATLAS detector contains subdetectors: the inner detector – closest to the beamline used to identify charged particle trajectory, the calorimeters – surround the inner detector to measure the energy of the particles, and the muon spectrometer – furthest outside to detect muons. Within the ATLAS detector, solenoid and toroid magnets allow for charged

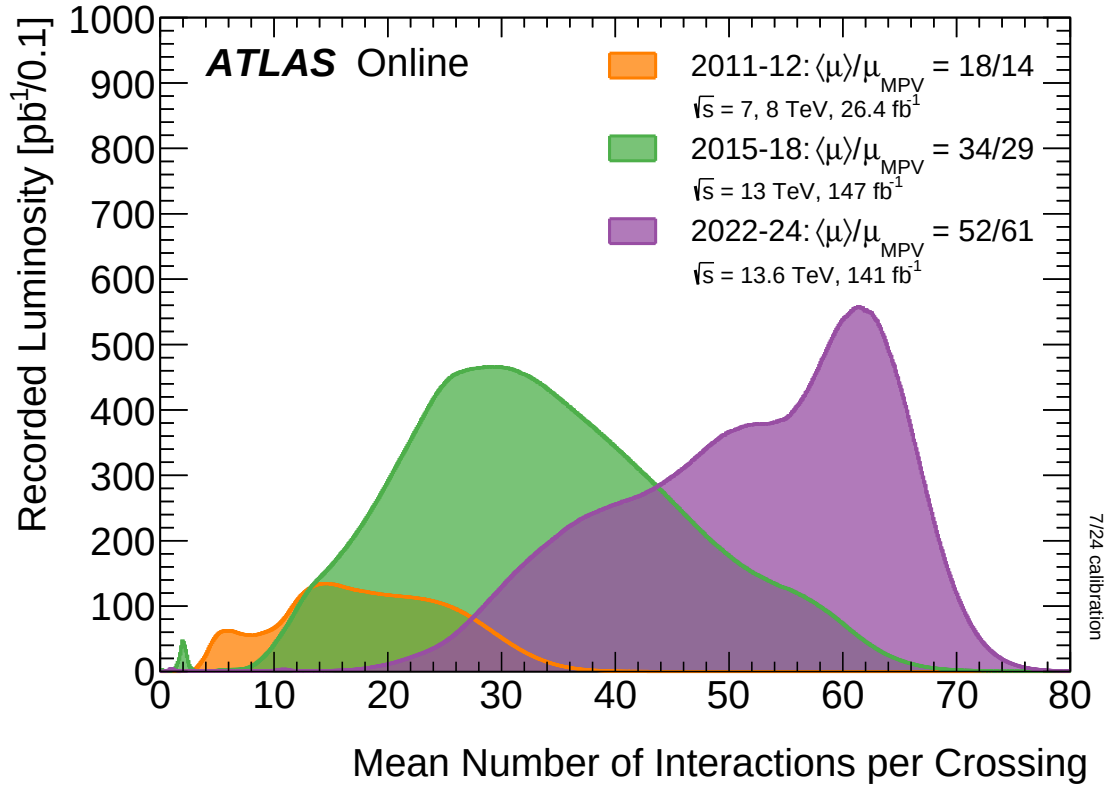


Figure 3.2: The recorded luminosity vs the mean number of interactions per bunch crossing, μ , taken with the ATLAS detector for Run 1 in orange, Run 2 in green, and the start of Run 3 in purple [46]. For each run, the average of $\mu - \langle \mu \rangle$ – and the most probable value – μ_{MPV} – are listed, rounding to the closest whole value.

particles to be deflected and measured by the subdetectors. There are also three forward detectors that are separate from the main ATLAS detector, which are positioned further down the beam line for luminosity and other measurements.

The ATLAS detector uses a Cartesian coordinate system for particle position with the origin at the centre of the detector. The z -axis points along the beamline, the x -axis is toward the centre of the LHC ring, and the y -axis points upwards to be orthogonal. A cylindrical coordinate system (r, ϕ) for the transverse xy -plane is used, where r is the radius and ϕ is the angle from the x -axis around the z -axis. The pseudorapidity, $\eta = -\ln [\tan (\theta / 2)]$, where θ is the polar angle between the trajectory and the positive z -axis is used in conjunction with the cylindrical coordinates. When the mass of a particle is

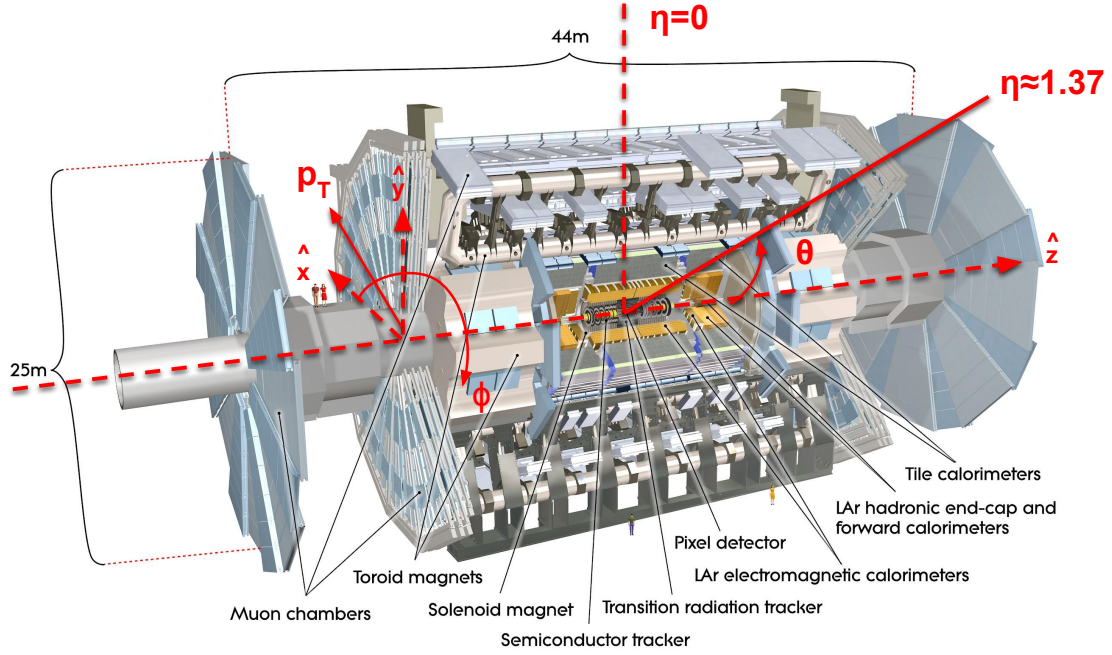


Figure 3.3: Schematic of the ATLAS detector with labeled subdetector and magnet locations [13]. From inside out: the inner detector, the calorimeters, and the muon spectrometer. The Cartesian coordinate system used by ATLAS includes: the z -axis is along the beamline, the x -axis points towards the centre of the LHC ring, and the y -axis consequently points upwards. The origin is at the centre of the detector but is displaced here for readability. Other coordinates used by ATLAS are r , θ , η , and ϕ , where r is the radius in the xy -plane, θ is the polar angle from the z -axis, $\eta = -\ln [\tan (\theta / 2)]$ is the pseudorapidity, and ϕ is the azimuthal angle measured from the x -axis. The transverse momentum, p_T , is in the transverse plane along r .

small compared to its energy, which is the case for electrons, the pseudorapidity is equal to the Lorentz invariant rapidity, y^1 . Both Cartesian and cylindrical coordinate systems are overlayed on the ATLAS detector in Figure 3.3, which also includes the transverse momentum, p_T , in the transverse plane along r .

The data used within this thesis is a subset of the data set taken during Run 2, so this section will go through the various subdetectors within ATLAS as they were at that time.

¹Not to be confused with the y coordinate. In most instances, rapidity will have a superscript y^α to denote that it is the rapidity of particle α .

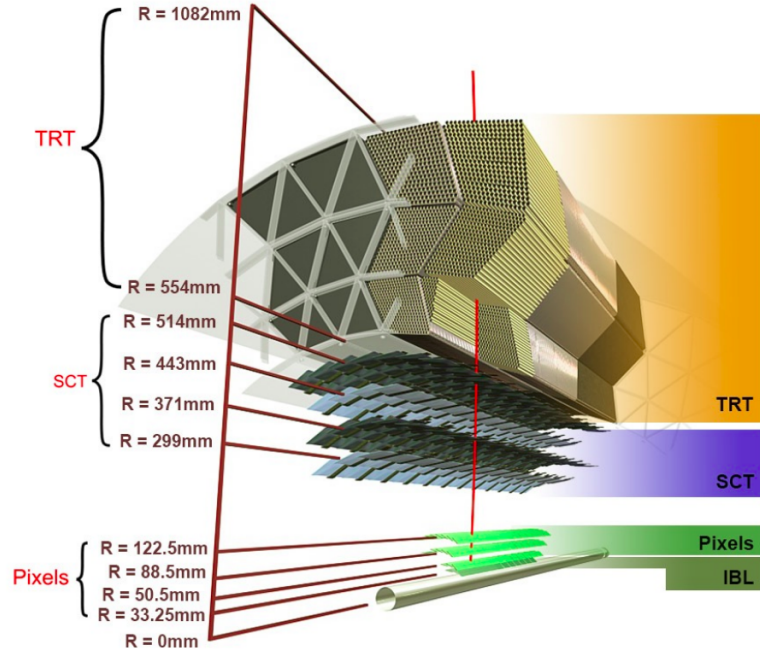


Figure 3.4: Cut away view of the ATLAS Inner Detector [50]. Closest to the beamline are the Insertable B-Layer (IBL) [51] and the three layers of Pixel detector. Outside of that are the four layers of Semiconductor Tracker (SCT). Lastly, is the Transition Radiation Tracker (TRT), which is constructed out of straws with polypropylene fibers between them.

3.3.1 Inner Detector

The Inner Detector (ID) [48, 49] is the subdetector that is at the centre of the ATLAS detector, closest to the beam pipe. This close proximity allows for measuring of charged particle trajectories and reconstruction of interaction vertices for the best resolution and high precision. There are two types of vertices: primary vertices – the original interaction location – and secondary vertices – the location of longer-lived particle decay, such as τ leptons and some hadrons. The ID is comprised of three subdetectors: the Pixel Detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT), as seen in Figure 3.4, which shows a cross-section of the detector. The proximity to the beamline provides a large angular range of $\sim 10^\circ \leq \theta \leq 170^\circ$ or a pseudorapidity range of $|\eta| < 2.5$.

Charged particles interact with multiple layers of the subdetector as they pass through them, allowing the trajectories to be measured. Between the ID and the calorimeters is

a solenoid magnet providing a magnetic field strength of 2 T, which causes the charged particles to curve with the relation:

$$p_T = qBr, \quad (3.3)$$

where B is the magnetic field, p_T is the transverse momentum, q is the electric charge, and r is the radius of curvature. The radius of curvature and the electric charge can be used to help identify charged particles, as heavier particles will have a larger radius and the charge of the particle will determine which direction it is deflected. All three subdetectors combined have a transverse momentum resolution [52] as a function of p_T :²

$$\frac{\sigma_{p_T}}{p_T} \sim 0.05\% \frac{p_T}{\text{GeV}} \oplus 1\%. \quad (3.4)$$

Pixel Detector

The innermost subdetector is the Pixel Detector, which is comprised of four barrel and three endcap layers. It has the finest position resolution of all ID subdetectors and makes up $\sim 80\%$ of the readout channels in the ATLAS detector. It uses silicon semiconductors to detect whether charged particles pass through its layers. The innermost barrel layer, called the Insertable B-Layer (IBL) [51], was added during the upgrade period between Run 1 and 2 and is located at 33.25 mm from the beamline. The IBL has granularity $50 \mu\text{m} \times 250 \mu\text{m}$ with resolution of $10 \mu\text{m}$ in the $r\phi$ -plane and $75 \mu\text{m}$ in the z direction. All the other layers of the Pixel Detector have granularity $50 \mu\text{m} \times 400 \mu\text{m}$ with resolution of $10 \mu\text{m}$ in the $r\phi$ -plane and $40 \mu\text{m}$ in the z direction [52].

²In this thesis, $\oplus$ denotes two quantities being added in quadrature, ie. $a \oplus b = \sqrt{a^2 + b^2}$

Semiconductor Tracker

The Semiconductor Tracker (SCT) is also a silicon semiconductor but has larger detection elements because it is further from the beamline. It is comprised of strips with pitch $80\text{ }\mu\text{m}$ and length 6 or 13 cm, providing a resolution of $17\text{ }\mu\text{m}$ in the $r\phi$ -plane and $580\text{ }\mu\text{m}$ in the z direction for the barrel layers [1]. The layers of strips are organized back-to-back with an angle of 40 mrad between them. The barrel has four layers of strips while each endcap has nine layers.

Transition Radiation Tracker

The outermost subdetector of the ID is the Transition Radiation Tracker (TRT) [53, 54, 55], which is comprised of straws that are 4 mm in diameter. Within the straw an anode wire runs directly through the centre while the outside of the straw acts as the cathode. Each straw is filled with a gaseous mixture of xenon (70%), carbon dioxide (27%), and oxygen (3%) [53]. When a charged particle passes through the gas, the gas atoms/molecules become ionized and are collected on the anode in the straw. Between the straws is polypropylene, where charged particles pass through this material and photons are emitted creating transition radiation. This radiation also accumulates on the anode wire within the straws. This transition radiation is useful for identifying electrons. Within the barrel, there are over 100,000 straws and approximately 250,000 in the endcaps, providing a combined spatial resolution of $130\text{ }\mu\text{m}$ in the $r\phi$ -plane.

3.3.2 Calorimeters

From inside out, the next layer of subdetectors in the ATLAS detector are the calorimeters, as seen in Figure 3.5. There are two calorimeter systems in ATLAS: the electromagnetic (ECAL) [56] and hadronic (HCAL) [56, 57]. Both are sampling calorimeters, which have alternating layers of absorber and active material. The absorber is a denser material that produces showers of particles within the detector, while the active material measures the

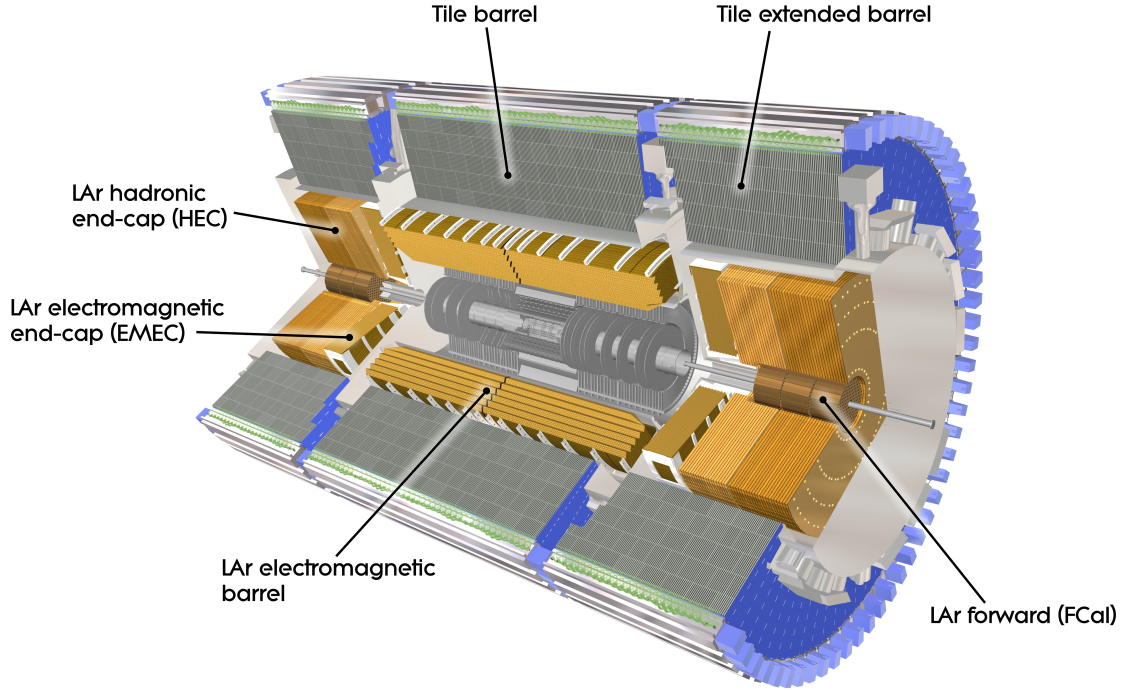


Figure 3.5: The schematic diagram of the calorimeter systems in the ATLAS detector [58]. On the inside are the electromagnetic calorimeter (ECAL) is electromagnetic barrel (EMB), electromagnetic endcap (EMEC), and the first layer of the Forward Calorimeter (FCal), which all use liquid argon (LAr). Outside of the ECAL is the hadronic calorimeter (HCAL) comprised of the Tile barrel and extended barrel calorimeter, LAr hadronic endcap (HEC), and the last two layers of the FCal.

energy deposited by the showering particles. The alternating layers of the two materials allow particles to come to a stop in a more compact space, while also having the particles deposit all of the energy within the calorimeter. The active material measures most of the energy deposited there, however, energy is also deposited in the absorber. To reconstruct the position and shower shape information of the particles, the calorimeters are segmented, allowing identification of particles to be performed later.

In general, the energy resolution of calorimeters is dependent on the energy of the particle and follows the quadratic sum of terms of the form [59]:

$$\frac{\sigma_E}{E} = \frac{a}{E} \oplus \frac{b}{\sqrt{E}} \oplus c. \quad (3.5)$$

In Equation 3.5, a is the term dependent on the electronic noise and pileup in the detector, b is the term dependent on stochastic sampling uncertainty, and c is a constant term from the response of the detector and reconstruction of particles.

ECAL

The innermost calorimeter is the ECAL, which is designed to detect particles that shower electromagnetically. The electromagnetic showers occur in the ECAL as a result of emission of photons from motion through matter – called bremsstrahlung radiation – and electron–positron pair creation from the emitted photon, which continues in a chain until they come to a stop in the detector. The energy deposited follows the form $E(x) = E_0 e^{-x/X_0}$, where x is the number of radiation lengths traveled and X_0 is the material dependent radiation length. The ECAL is designed so that particles travel through at least $24X_0$ of material to ensure that particles, such as electrons and photons, come to a stop within the detector. Other particles, such as muons and hadrons, deposit some energy within the ECAL but are not stopped within it and pass through the system to the rest of the detector.

The ECAL is comprised of three subdetector sections: the Electromagnetic Barrel (EMB), the Electromagnetic Endcap (EMEC), and the first layer of the Forward Calorimeter (FCal), as seen in Figure 3.5. The EMB and EMEC share a similar construction, utilizing lead absorbers in an accordion shape and liquid argon (LAr) as the active material. There are three layers within the EMB and EMEC of varying size, as seen in Figure 3.6. The different sizes allow for specialization of how energy is deposited and their resolution. The first layer discriminates between charged particles by using 4.69 mm wide strips in the η direction, providing a granularity of $\eta \times \phi = 0.0031 \times 0.1$. Most of the electromagnetic showering occurs in the second layer, which has a granularity of $\eta \times \phi = 0.025 \times 0.0245$. The last layer is designed to measure any remaining energy that was not deposited in the first two layers and has a granularity of $\eta \times \phi = 0.05 \times 0.025$. Before these three layers is a presampler to correct for the energy lost by particles in the detector before it reached

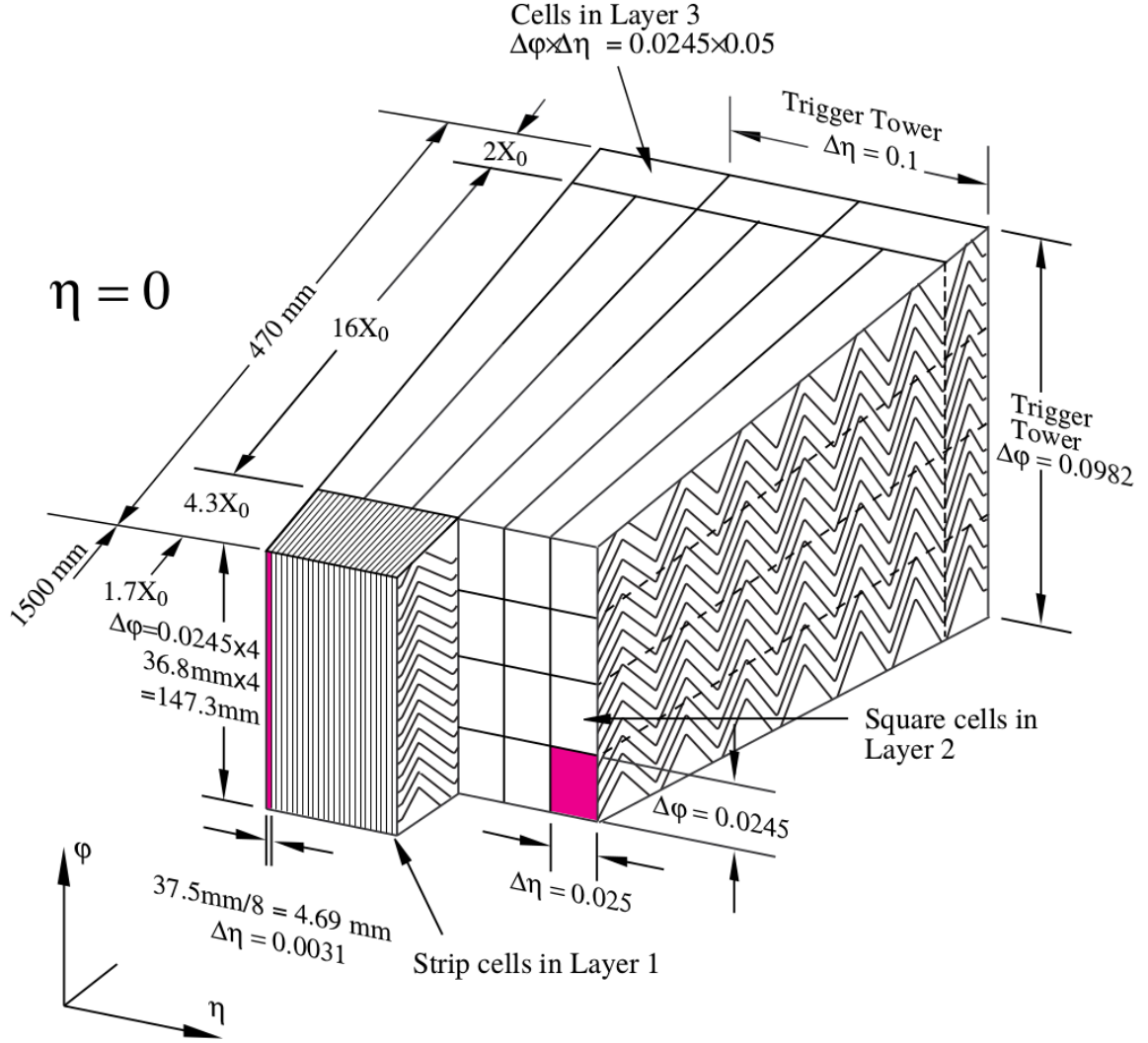


Figure 3.6: Sketch of the cells within the LAr barrel, which lists the $\eta \times \phi$ granularity of each layer. The side cutaway view shows the accordion shape of the layers [1].

the calorimeter, which has a granularity of $\eta \times \phi = 0.025 \times 0.1$ and covers a range of $|\eta| < 1.8$. The EMB and EMEC cover a range of $|\eta| < 1.475$ and $1.375 < |\eta| < 3.2$, respectively. The calorimeter sections coverage overlap in the region of $1.37 < |\eta| < 1.52$ and, due to not as many layers in this region, there is a loss of efficiency. For Equation 3.5, the EMB and EMEC have a resolution of $a \approx 0.4 \text{ GeV}$ [56], $b \approx 10\% \sqrt{\text{GeV}}$, and $c = 0.7\%$ [1].

The ECAL portion of the FCal is constructed with copper as the absorber, rather than the lead used EMB and EMEC. It follows a cylindrical geometry covering a pseudorapidity

range of $3.0 < |\eta| < 4.9$. Due to its geometry, the FCal has worse resolution, where $b \approx 29\% \sqrt{\text{GeV}}$ and $c = 3.5\%$ [1], while a remains $\approx 0.4 \text{ GeV}$.

HCAL

Heavier particles, like hadrons, escape the ECAL and their energy is measured in the HCAL [56, 57]. It is comprised of three subdetectors: the Tile Calorimeter, the Hadronic Endcap (HEC), and the last two layers of the FCal, as seen in Figure 3.5. The Tile calorimeter covers a region of $|\eta| < 1.7$ and is constructed of alternating layers of steel absorber plates and polystyrene scintillator tiles as an active material. There are three layers, which have granularity of $\eta \times \phi = 0.1 \times 0.1$ for the two inner layers and $\eta \times \phi = 0.2 \times 0.1$ for the last layer. For Equation 3.5, the Tile calorimeter has a resolution of $a \approx 0.5 \text{ GeV}$, $b \approx 50\% \sqrt{\text{GeV}}$, and $c = 3\%$ [57].

The HEC is comprised of two wheels and uses copper as the absorber while LAr is the active material. Positioned behind the EMEC, it covers a region of $1.5 < |\eta| < 3.2$ and has a granularity of $\eta \times \phi = 0.1 \times 0.1$ for $|\eta| < 2.5$, while the remainder has a granularity of $\eta \times \phi = 0.2 \times 0.2$. Compared to the Tile calorimeter, the HEC has worse resolution, with a ranging between $0.5 - 1.5 \text{ GeV}$ as a function of η , $b \approx 60\% \sqrt{\text{GeV}}$, and $c = 2\%$ [1]. This worse resolution is due to the HEC being closer to the beam and needing to be robust against radiation damage. The HCAL portion of the FCal uses tungsten as an absorber and covers a region of $3.1 < |\eta| < 4.9$.

3.3.3 Muon Spectrometer

The outermost subdetector of the ATLAS detector is the Muon Spectrometer (MS) [60]. Muons are a minimally ionizing particle, have a long lifetime of $\tau \sim 2.2 \times 10^{-6} \text{ s}$, and a large mass of $m^\mu \sim 105.7 \text{ MeV}$ [25], so they usually escape past the calorimeters and need their own designated detector to properly measure them. During Runs 1 and 2, the MS was comprised of four subdetectors: the Monitored Drift Tube (MDT), Cathode Strip Chamber

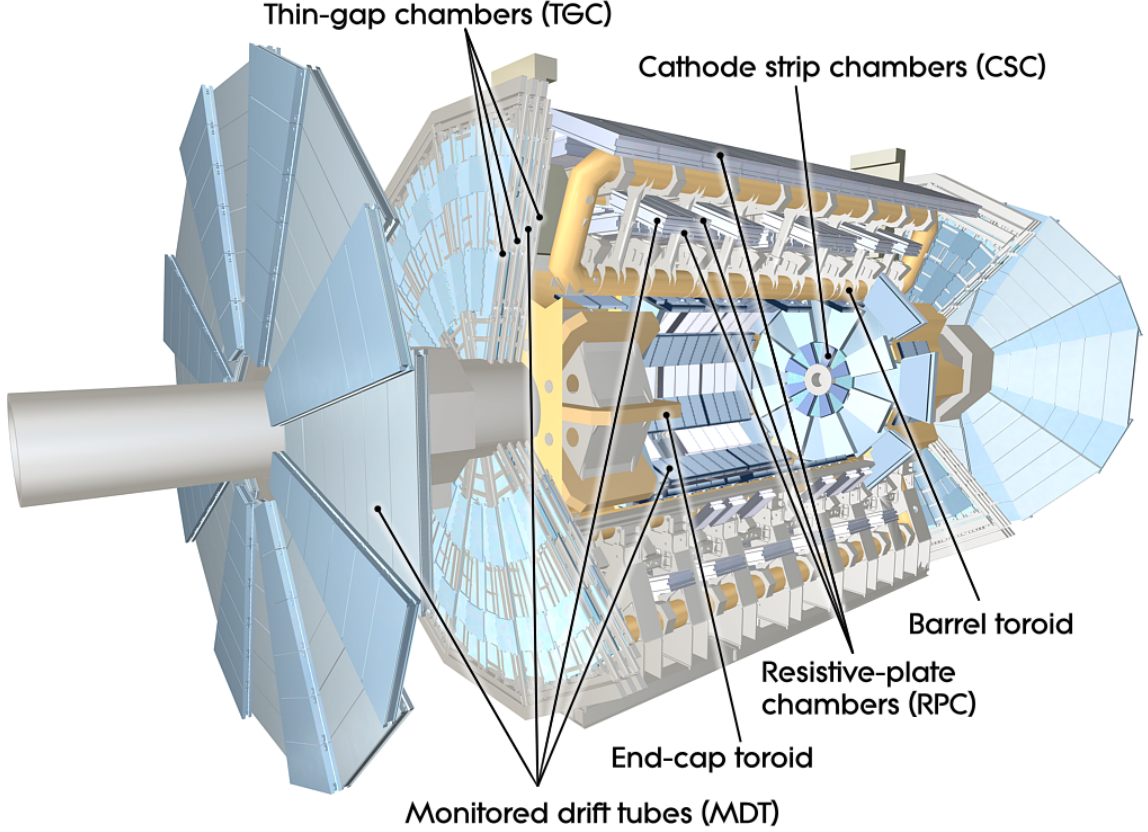


Figure 3.7: Schematic of the muon spectrometer (MS) [61]. The barrel is comprised of Monitored Drift Tubes (MDT), Resistive Plate Chambers (RPC), and the barrel toroid magnets. For each endcap, there is the smaller wheel of Cathode Strip Chambers (CSC) and Thin Gap Chambers (TGC), the endcap toroid magnets, and the larger two wheels, comprised of MDTs and TGCs.

(CSC), Resistive Plate Chamber (RPC), and the Thin Gap Chamber (TGC), as seen in Figure 3.7. Like the other subdetectors, it is also split into a barrel and endcap geometry. The barrel of the MS covers a range of $|\eta| < 1.05$ and is comprised of three layers. The end caps are constructed using three separate wheels and cover a range of $1.05 < |\eta| < 2.7$ but triggering only happens up to $|\eta| < 2.4$. The MS measures the charge and momentum of the muons, which have their trajectory curved from the magnetic fields produced from dedicated toroid magnets in the barrel and the endcaps; the separate magnets produce a non-uniform magnetic field. The MS magnetic field is split into three regions: the barrel $|\eta| < 1.4$, the transition region $1.4 < |\eta| < 1.6$, and the endcap $1.6 < |\eta| < 2.7$.

Subdetectors of the MS utilize ionization of gas to measure when charged particles pass through. Within the barrel and the two outermost wheels, the MDTs are tubes filled with a gas mixture of Ar (91%), N₂ (4%), and CH₄ (5%). The CSCs, located only in the small inner endcap wheel, are constructed using strips with a mesh anode, allowing for a two dimensional position measurement. Located only in the barrel, the RPCs contain two plastic resistive plates with gas between the 2 mm gap and are used as part of the muon triggering system. The TGCs act as a trigger in the endcaps and use strips and a mesh anode, similar to the CSCs. Of the four subdetectors, the MDTs and CSCs give precision measurements for position and transverse momentum, while RPCs and TGCs are used for triggering. The MS has momentum resolution [60]:

$$\frac{\sigma_{p_T}}{p_T} \sim \begin{cases} 2 - 3\% \frac{p_T}{\text{GeV}} & \text{most of the kinematic range,} \\ 10\% \frac{p_T}{\text{GeV}} & p_T \geq 1 \text{ TeV.} \end{cases} \quad (3.6)$$

3.3.4 Forward Detectors

Located separate from the main detector are three subdetectors: LUMinosity measurement using Cerenkov Integrating Detector (LUCID) [62], Zero Degree Calorimeter (ZDC), and Absolute Luminosity For ATLAS (ALFA) [1]. The ZDC and ALFA are located ± 140 m and approximately ± 240 m from the interaction point, respectively. Both are not relevant to this thesis and will not be discussed further. LUCID is located ± 17 m from the interaction point and is the main relative luminosity monitor for ATLAS. It is a dedicated detector for measuring the integrated luminosity independent of the physics data. Each side of LUCID is built with 20 aluminum tubes surrounding the beam pipe and measures Cerenkov radiation produced by particles interacting with the quartz window of the photomultiplier tubes that point in the direction of the interaction point.

Since LUCID provides an independent luminosity measurement from data, it needs an absolute calibration using dedicated vdM scan runs. The vdM scans position the beams to

be head on and then the beam separation is changed in steps of 50 μm , first in the horizontal and then the vertical direction. Then, the beams are positioned with an offset and the 50 μm steps are performed again. These steps change the visible interaction rate and provides a measurement of the beam widths in the x and y direction, σ_x and σ_y respectively. With the beam widths, the luminosity can be calculated using [43]

$$L = \frac{f_{\text{LHC}} N_b N_1 N_2}{4\pi\sigma_x\sigma_y}, \quad (3.7)$$

where f_{LHC} is the revolution frequency of the LHC, N_b the number of bunches, and $N_{1,2}$ are the number of protons in bunch 1 and 2.

3.3.5 Triggering

During Run 2, there were an average of 33.7 interactions occurring every 25 ns, which corresponds to an event rate of ~ 1.3 GHz. A single event in the ATLAS detector produces approximately 1.5 MB of data, thus, not every interaction can be saved and recorded. Many of the interactions that occur are not of much physical interest, so selection criteria can be applied to reduce the number of events that are saved. This selection first happens on a hardware level, called the Level-1 trigger, where custom electronic hardware [63] takes information from the calorimeters and MS and determines if the event could potentially contain physical interest. If the event passes the hardware level trigger, then all the relevant subdetector information is read out in the regions of interest. These selected events then undergo reconstruction algorithms and have further software level event selection, called the high level trigger (HLT), applied to determine if they should be kept. This reduces the event rate to approximately 1 kHz, which is then recorded for further analysis.

3.4 Summary

This chapter covered how bunches of protons are accelerated with the LHC, which are subsequently collided at four interaction points, such as the ATLAS detector. It also covered how the ATLAS detector and its subdetectors operate to measure the signatures from particle interactions within it. With an understanding of the ATLAS detector, the next chapter will discuss how the particle interactions with the ATLAS detector can be used to reconstruct particles and their energies to make measurements.

CHAPTER 4

PARTICLE RECONSTRUCTION

The previous chapter introduced the broad idea of how particles are reconstructed and identified using the subdetectors within the ATLAS detector. This chapter will discuss how the signals within the subdetectors become reconstructed objects that can be measured in analyses. This thesis focuses on measuring W bosons immediately decaying into an electron/muon and a neutrino. The electrons and muons are the only visible part of the process measured in the ATLAS detector, so they need to be measured well. Neutrinos are indirectly detected by reconstructing the full event and looking for missing energy. To understand how these various particles are measured, the concepts of measurement efficiencies, and how they are corrected in Monte Carlo (MC) simulation, will be introduced. Next, the steps needed to reconstruct most detected particles are explained, before describing how electrons and muons are identified, isolated, and triggered on.

4.1 Efficiencies and Scale Factors

When trying to measure a process, such as a W boson decaying to either an electron/muon and a neutrino, the goal is to properly identify particles, such as electrons and muons. Other particles that are misidentified as the desired, prompt electron/muon are referred to as background, which comes in two types. The first is when a different particle results in a trace in the detector that is similar to the electron/muon and the second is when an electron/muon originate from another source other than the desired W boson. Electrons and muons from other sources are put into two different categories, prompt and non-prompt. An example of a prompt background is a Z boson decaying immediately into two muons, where only one muon leaves a signature in the detector. This appears as a single muon and missing energy, which can be mistaken for a W boson decaying into an muon and

neutrino. Non-prompt backgrounds are when the electron/muon do not come from the original interaction, but are a secondary decay product. For example, a B hadron decays semileptonically, producing another hadron and a W boson, when then decays into an electron neutrino pair. This appears as the desired signal, but does not originate from the desired process.

To fully reconstruct the desired particles, there are four major steps: reconstruction, identification, isolation, and triggering; these steps separate the desired particles from background particles to have higher signal purity. Each step has its own efficiency, ϵ , which follows a binomial probability distribution [64]. This probability is estimated by counting the number of events that pass selection criteria out of all events, $N^{\text{pass}}/N^{\text{total}}$. With the efficiency of an individual step the total efficiency of multiple steps can be expressed as the product of the efficiencies of each step. For example, in the case of large N an electron has a total efficiency, ϵ^{Total} :

$$\begin{aligned}\epsilon^{\text{Total}} &= \epsilon^{\text{reco}} \times \epsilon^{\text{id}} \times \epsilon^{\text{iso}} \times \epsilon^{\text{trig}} \\ &= \frac{N^{\text{reco}}}{N^{\text{total}}} \times \frac{N^{\text{id}}}{N^{\text{reco}}} \times \frac{N^{\text{iso}}}{N^{\text{id}}} \times \frac{N^{\text{trig}}}{N^{\text{iso}}}.\end{aligned}\tag{4.1}$$

In this case, ϵ^{reco} is the reconstruction efficiency calculated by the number of reconstructed electrons, N^{reco} , out of the total number of electron candidates, N^{total} . Similarly, ϵ^{id} is the identification efficiency from the number of identified electrons, N^{id} , out of the reconstructed electrons. ϵ^{iso} is the isolation efficiency from the number of isolated electrons, N^{iso} , out of the identified electrons. Lastly, ϵ^{trig} is the trigger efficiency from the number of electrons triggered on, N^{trig} , out of the isolated electrons.

The reconstruction of electrons and muons will be discussed separately in Sections 4.3.2 and 4.3.4, respectively. Similarly, Section 4.4.1 discusses identification for electrons while Section 4.4.2 discusses the same for muons. Both leptons use the same method for isolation discussed in Section 4.5 and the concepts for triggering are discussed, in general, in Section 4.6 and would apply to both.

To get an expectation for the measured efficiencies, the ATLAS detector and its configuration are simulated MC using GEANT4 [65], which take the alignment of subdetector components and the expected detector efficiencies into account. The data and MC efficiency ratio is a scale factor (SF):

$$SF = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}, \quad (4.2)$$

which, when applied to MC, allows the expected number of events to more closely matches what is seen in data. The SFs are calculated for each step individually, as each step will calculate efficiencies with multiple selection requirement options, allowing individual analyses to chose which level of background rejection is best suited.

4.2 Tag and Probe Method

To measure efficiency, a common method called Tag&Probe is used [66, 67]. This method uses a data set containing likely real, prompt leptons from events of likely Z boson decay. For an event to meet the criteria to be a $Z \rightarrow l^+l^-$ candidate it must contain: two oppositely charged leptons, have the dilepton mass lie within a selected window for the Z boson mass spectrum, and one lepton needs to meet stringent selection criteria but the other needs to meet less stringent criteria. The lepton that meets the stringent selection criteria is known as the “tag” lepton, while the more loosely selected is called the “probe” lepton, as seen in Figure 4.1.

The efficiency is then calculated by counting the number of events that pass the criteria of interest for that probe lepton, $N_{\text{Probes}}^{\text{Selected}}$, and dividing by the total number of probes, N_{Probes} ,

$$\epsilon = \frac{N_{\text{Probes}}^{\text{Selected}}}{N_{\text{Probes}}}. \quad (4.3)$$

These efficiencies are usually measured as a function of η and transverse energy, E_T , of the probe lepton. Within data, some background will pass the selection criteria applied and can cause a discrepancy in efficiency, as MC does not include background. Typically,

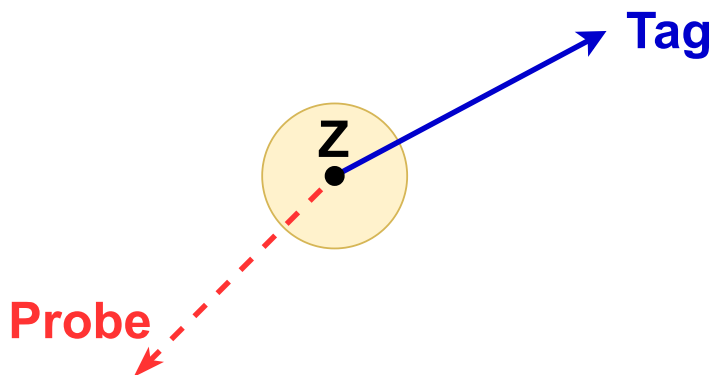


Figure 4.1: Cartoon example of the Tag&Probe method in the ATLAS detector. The tag lepton meets stringent criteria, while the probe lepton is more loosely selected and is used to determine efficiencies.

a mass range of $80 < m^Z < 100$ GeV is used. In this region the expected amount of background is $\sim 2\%$ for electrons [68] and $\sim 0.1\%$ for muons [67].

4.3 Reconstruction

Being a multipurpose detector, ATLAS is designed to reconstruct a wide variety of particles. Depending on the analysis, a combination of particles are used to reconstruct the desired signal to make a measurement. The kinematic properties of the particle detected can be determined by measuring combinations of charged particle tracks in the ID, the energy deposited in the calorimeters, or the trajectory in the MS. Examples can be found in Figure 4.2, which shows the various signatures that particles will leave within the ATLAS detector. Each of these cases will be described in the remainder of this section. First, the method for track reconstruction is explained in Section 4.3.1, as it is common to different particle cases later. Then, electron reconstruction will be covered in Section 4.3.2, photons in Section 4.3.3, muons in Section 4.3.4, and hadronic particles, like neutrons and protons, will be covered in Section 4.3.5.

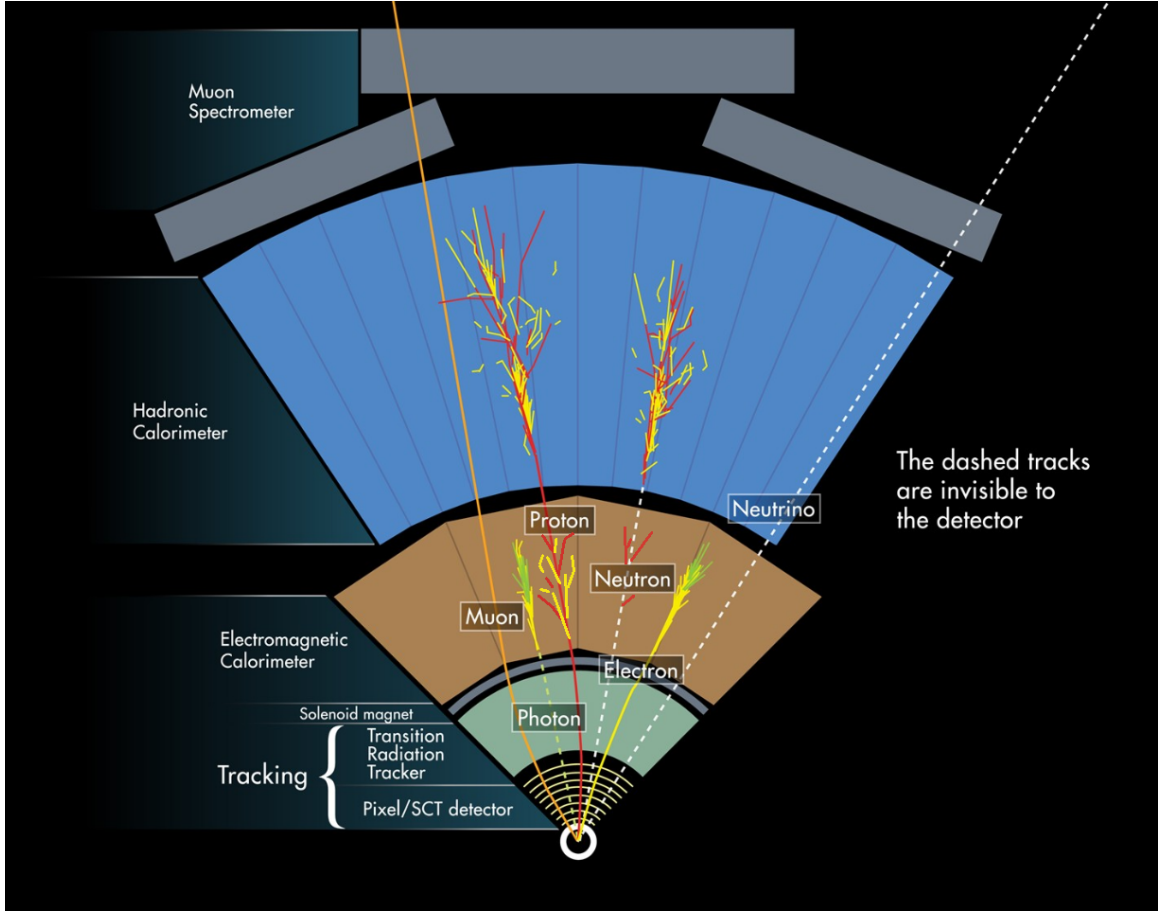


Figure 4.2: A schematic slice of the ATLAS detector in the main barrel region to show the signatures of various particles in the detector [69]. The charged particles, such as electrons, protons, and muons, will leave tracks in the ID. Electrons and photons deposit their energy in the ECAL, while hadronic particles, such as neutrons and protons, will cause showers in the HCAL and some energy is deposited in the ECAL. Muons deposit some energy in the calorimeters but are mostly measured by the MS. Neutrinos do not leave any signature within the ATLAS detector.

4.3.1 Track Reconstruction

Tracks in the ID are reconstructed by using the charged particle interactions within the silicon layers, “hits”, in the Pixel and SCT detectors [66, 70, 71] and the straws in the TRT detector [53, 54, 55]. Pattern recognition algorithms [66, 72] are used to find track candidates, starting with sets of three points in the silicon layers, called the track seeds. Full track reconstruction starts with the track seeds and fits potential paths to find tracks requiring at least seven hits for $p_T > 400$ MeV. If no tracks can be found from the seed,

then a second attempt is performed, where each hit is allowed up to 30% energy loss from bremsstrahlung. These tracks from second attempts appear as an increased curvature in the track. With many events happening concurrently, there are cases where a hit can be associated to more than one track, so ambiguity resolving algorithms are used to find the most likely tracks [66].

4.3.2 Electron Reconstruction

Electron reconstruction uses superclusters of energy in the electromagnetic calorimeter, where there is a matching track from the ID [68]. To find superclusters, clusters of energy in the calorimeters are first identified using the significance of energy in the cells:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}, \quad (4.4)$$

where $E_{\text{cell}}^{\text{EM}}$ is the energy in the cell and $\sigma_{\text{noise,cell}}^{\text{EM}}$ is the expected cell noise from electronics or pileup. Building clusters involves a series of steps. The first step is identifying any cell where $|\zeta_{\text{cell}}^{\text{EM}}| \geq 4$. Then, neighbour cells are checked to see if they have a significance greater than or equal to two, if so, they are added to the cluster. After a cell is added to the cluster, their neighbour cells are checked to potentially add to the cluster. During this cluster reconstruction, nearby clusters can be merged if they share common cells above the threshold, which are separated by finding local maxima. To qualify as a local maximum, the cell must pass the threshold of $E_{\text{cell}}^{\text{EM}} > 500$ MeV and have four neighbour cells with less energy. These energy clusters are then combined to form superclusters.

To form a supercluster, each cluster with $E_{\text{T}} > 1$ GeV is tested in decreasing order of E_{T} . Each cluster is used as a seed and must be matched to a track with at least four hits in the silicon tracking detectors. These seed clusters are used to find nearby clusters of energy within the window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster, as seen in Figure 4.3. Additional clusters in an area $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ are included if their

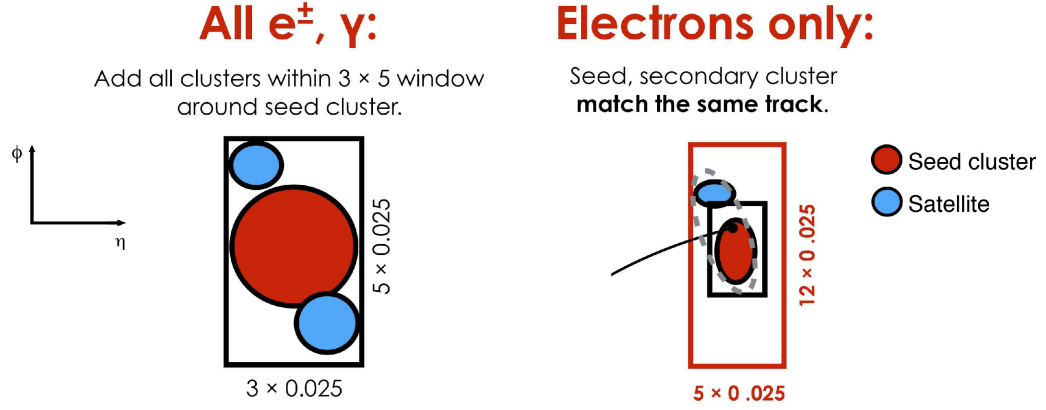


Figure 4.3: Clusters around the seed cluster in an area of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ are added to form a supercluster. Clusters with the same best-matched track to the seed cluster in an area of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ are also added to the supercluster. Cropped image from [68].

best-matched track is the same matched track as the seed cluster. The seed cluster and the associated nearby clusters form a supercluster.

To match to a track, the cluster must satisfy separation requirements in η and ϕ . The η separation is $|\eta_{\text{cluster}} - \eta_{\text{track}}| < 0.05$ and the ϕ separation must satisfy one of:

$$\begin{aligned}
 -0.10 &< -q \times (\phi_{\text{cluster}} - \phi_{\text{track}}) < 0.05 \\
 -0.10 &< -q \times (\phi_{\text{cluster}} - \phi_{\text{track}}^{\text{rescaled}}) < 0.05.
 \end{aligned}
 \tag{4.5}$$

In these requirements q is the sign of the electric charge, ϕ is the azimuthal angle of the cluster or track, and $\phi_{\text{track}}^{\text{rescaled}}$ is where the track is rescaled to the energy from the cluster. The bounds are asymmetric to account for the bending of the track because energy is lost due to bremsstrahlung.

4.3.3 Photon Reconstruction

Photons are measured by their deposited energy in the electromagnetic calorimeter in the same way as electrons. They usually have a different shower shape and do not have a track associated to them. The exception to this is the case of photon conversion, where there are partial tracks originating from a secondary vertex. The reconstruction efficiency of

photons is reduced from two main factors: π^0 decay and electron-positron conversion. The π^0 meson predominately decays into two photons, which can be reconstructed to originate from separate events or overlap to appear as a photon. Before reaching the ECAL, the photon can be converted to an electron-positron pair and be reconstructed as such if the conversion vertex cannot be recovered. Photons are not explicitly used in this analysis and only contribute when being emitted from an initial particle and the energy loss needs to be accounted for.

4.3.4 Muon Reconstruction

Muons used in this analysis are reconstructed using a combination of information from the ID and MS [67], called combined muons. The reconstruction starts in the MS, where tracks are reconstructed by finding a pattern of hits to form track candidates, similar to tracks in the ID, as discussed in Section 4.3.1. Starting from these track candidates, the search is expanded to find full tracks, where middle layers of the MS are prioritized and then expanded to the inner and outer layers. The hits in the RPC or TGC layers give coordinates orthogonal to the bending plane from the magnets, while the CSC gives the coordinates in the $\eta\phi$ -plane. Overlap removal algorithms are then performed to find the most likely tracks when more than one track overlaps. The reconstructed tracks in the MS are extrapolated inward to find associated tracks in the ID, where hits in the MS may be added or removed to improve the fit.

4.3.5 Hadronic Reconstruction

Particles produced from strong interactions, such as gluons or quarks, form hadrons and have most of their energy deposited in the HCAL. These partons will undergo hadronization, as described in Section 2.2.4, and are aggregated in a conical shape, called a jet. These jets are reconstructed with either information from just the calorimeter or in a combination with charged particle tracks from the ID [73]. An example of this can be seen in Figure 4.4,

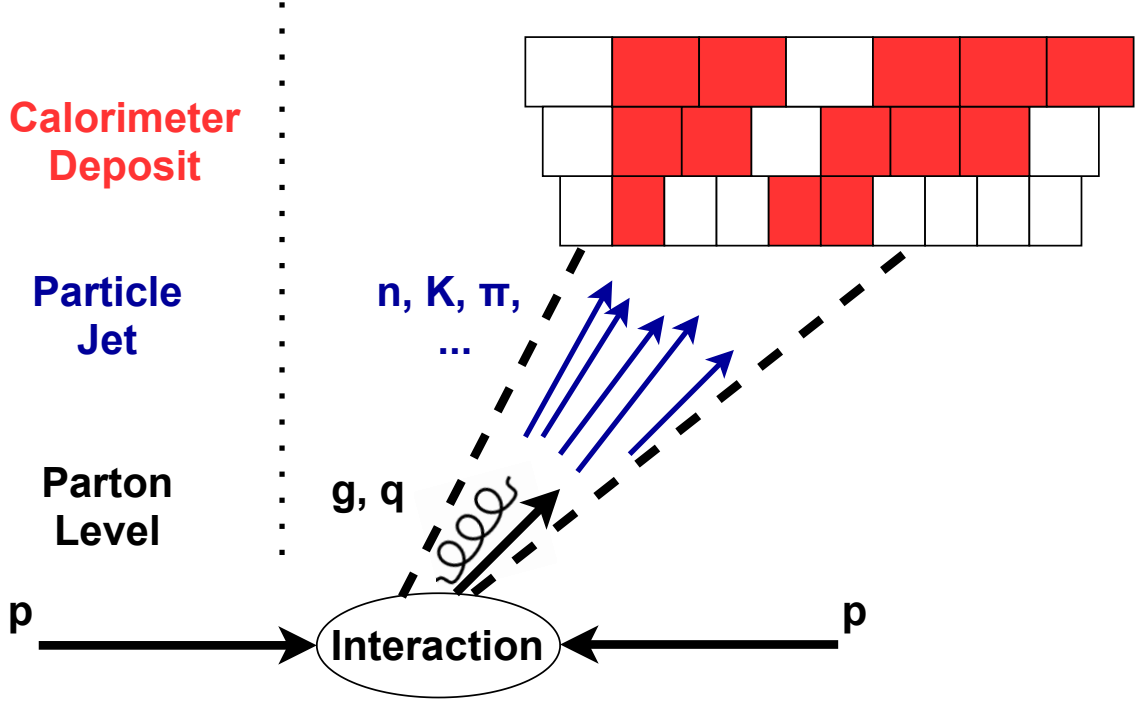


Figure 4.4: From an interaction at a parton level, quarks and gluons are produced. Instead of bare partons, hadrons are formed, which deposit energy in the calorimeter cells in a conical shape, called a jet.

which shows the quarks or gluons produced at parton level which appear as jets of particles and deposit energy in calorimeters. The parton level is more of a theoretical approximation and the jets are the physical objects that are measured. The energy from the Tile calorimeter and HEC is used in a similar way to electrons and photons in the ECAL, where topological clusters of energy are found. These clusters are separated into multiple hadronic particles using the anti- k_t algorithm [74] with a radius parameter of $R = 0.4$. The energy and resolution of jets are calibrated through both simulation-based and in-situ corrections. Jets are not used explicitly in this analysis but the energy they deposit in calorimeters is used to calculate the missing energy, as discussed in Section 8.4.

4.4 Particle Identification

Particles are identified using information from the various ATLAS subdetectors. Section 4.4.1 will cover the identification of electrons in the central region of the detec-

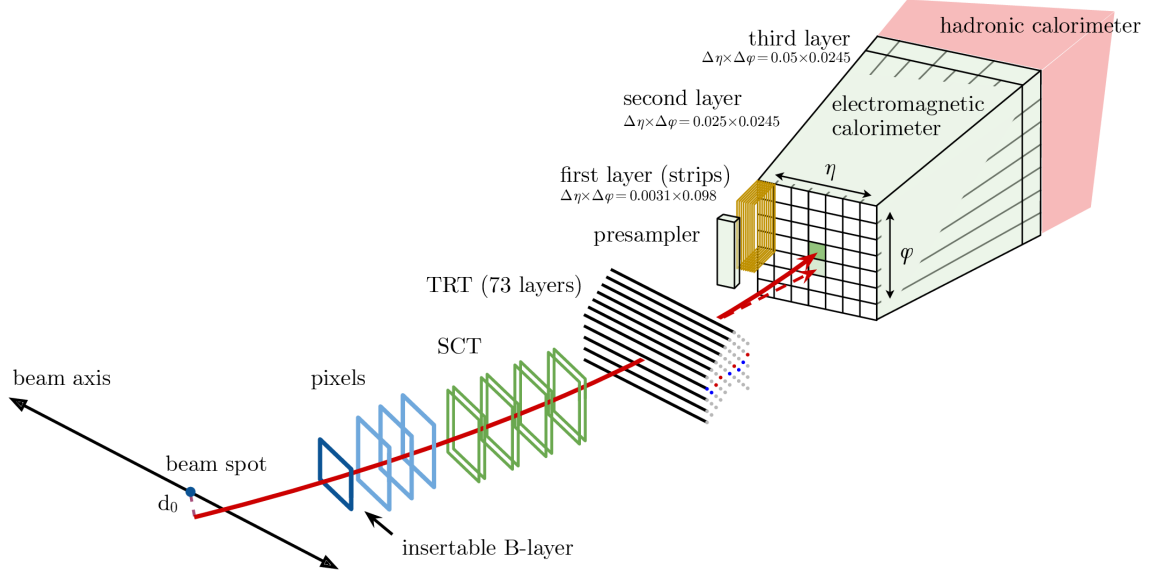


Figure 4.5: Example of the path traveled by an electron passing through the ATLAS detector from [66]. The solid red line shows the path through each layer of the ID before entering the ECAL. The dashed red line shows the path of a photon if it were produced from the electron interacting with the ID.

tor ($|\eta| < 2.47$). An example path of the electron can be seen in Figure 4.5, where the electron interacts with the subdetectors in the ID, creating a track, before depositing its energy in the ECAL. Muon identification will be covered in Section 4.4.2, which uses information from the ID and MS.

4.4.1 Electron Identification

More than the desired, prompt electrons are reconstructed using the methods described in Section 4.3.2. Many sources of background are also reconstructed as electrons, as discussed in Section 4.1. These backgrounds are not easily separable from the signal, reducing the identification efficiency. Instead of selection criteria, a likelihood (LH) for signal, $\mathcal{L}_S$, and background, $\mathcal{L}_B$, are created and used in a discriminating variable. The LHs are calculated using N signal (background) probability distribution functions, $P_{S(B),i}$:

$$\mathcal{L}_{S(B)}(\vec{x}) = \prod_{i=1}^N P_{S(B),i}(x_i). \quad (4.6)$$

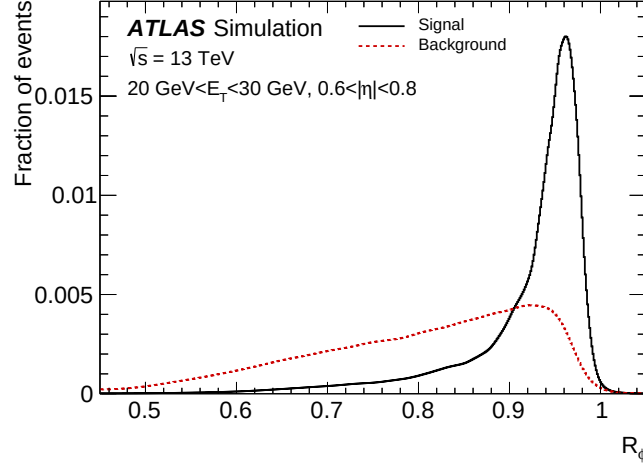


Figure 4.6: Example probability distribution function for R_ϕ for both the signal and background in the region of $20 < E_T < 30$ GeV and $0.6 < |\eta| < 0.8$ from [66]. There is a discernible difference between signal and background shape, so it is used as part of a LH to discriminate the difference, where it would be inefficient to use a selection-based discrimination.

In this equation $\vec{x}$ is the vector of various tracking variables – listed in Table 4.1 and shower shape variables listed in Table 4.2, and $P_{S(B),i}(x_i)$ is the probability that the quantity i at value x_i is signal (background). An example of the signal and background probability distribution function for a shower shape variable, R_ϕ , can be seen in Figure 4.6, where a shape difference between the probability distribution functions. However, there is significant overlap and an efficient selection criteria could not be applied.

Depictions of most of the tracking and shower shape variables can be seen in Figures 4.7 and 4.8, respectively. Figure 4.7 also depicts the track geometry relative to the beam spot to define the transverse impact parameter – d_0 , the longitudinal impact parameter – z_0 , and the perigee, which is used for $\Delta\phi_{\text{rescaled}}$ and $\Delta P/P$. All of the variables in Tables 4.1 and 4.2 are used to create the probability distribution functions for the LH.

Tracking Variable	Description
n_{Blayer}	Number of hits in the IBL.
n_{Pixel}	Number of hits in the Pixel detector.
n_{Si}	Number of hits in the Pixel and SCT detectors.
d_0	Transverse impact parameter relative to the beam-line.
d_0/σ_{d_0}	Significance of d_0 , defined as the ratio of d_0 to its uncertainty.
$\Delta\eta_1$	$\Delta\eta$ between the cluster position in the first layer and the extrapolated track.
$\Delta\phi_{\text{rescaled}}$	$\Delta\phi$ between the cluster position in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge.
$\Delta P/P$	Difference between momentum of the track at the perigee and the last measurement, divided by the momentum at perigee.
eProbabilityHT	Likelihood probability, based on transition radiation in the TRT.
E/p	Ratio of the cluster energy and track momentum.

Table 4.1: List of tracking variables with a brief description that are used in the electron identification LH to determine if events are electrons or background. A more detailed use of each variable is described in [66] and a depiction of some can be found in Figure 4.7.

Shower Shape Variable	Description
f_1	Ratio of energy deposited in first layer to the total energy in EM calorimeter.
f_3	Ratio of energy deposited in third layer to the total energy in EM calorimeter. Only used in the LH for $E_T < 80$ GeV and $ \eta^l < 2.37$.
w_{η_2}	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 within 3×5 cells window.
w_{stot}	Shower width, $\sqrt{\sum E_i (i - i_{\text{max}})^2 / \sum E_i}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \sim 0.0625 \times 0.2$, where i_{max} is the highest-energy strip index. Only used in LH for $E_T > 150$ GeV.
R_η	Ratio of the energy in 3×7 cells over the energy in 7×7 cells, centred at the electron cluster position.
R_ϕ	Ratio of the energy in 3×3 cells over the energy in 3×7 cells, centred at the electron cluster position.
R_{had}	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster. Only used over the range $0.8 < \eta^l < 1.37$.
R_{had1}	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster. Only used over the range $0.8 < \eta^l < 1.37$.
e_{ratio}	Ratio of the energy difference in the cluster between the maximum energy deposit and the energy deposit in a secondary maximum to the sum of these energies.

Table 4.2: List of shower shape variables with a brief description that are used in the electron identification LH to determine if events are electrons or background. A more detailed use of each variable is described in [66] and a depiction can be seen in Figure 4.8.

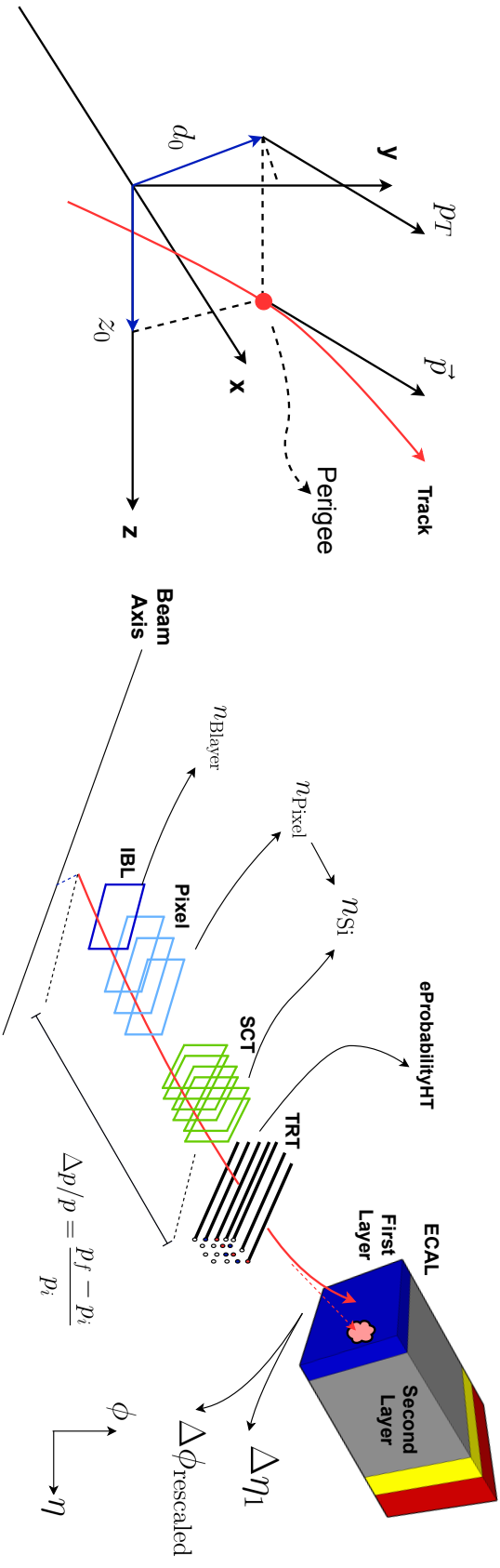


Figure 4.7: Depiction of some of the tracking variables used to create probability distribution functions for the electron identification LH. The left image depicts the track geometry near the beam spot to define the transverse impact parameter $-d_0$, the perigee, which is used for $\Delta\phi_{\text{rescaled}}$ and $\Delta P/P$, and the longitudinal impact parameter $-z_0$, which is not used in the electron identification LH. The image on the right is based on an image from [66], shown previously in Figure 4.5, and shows $\Delta P/P$, eProbabilityHT, $\Delta\eta_1$, and $\Delta\phi_{\text{rescaled}}$.

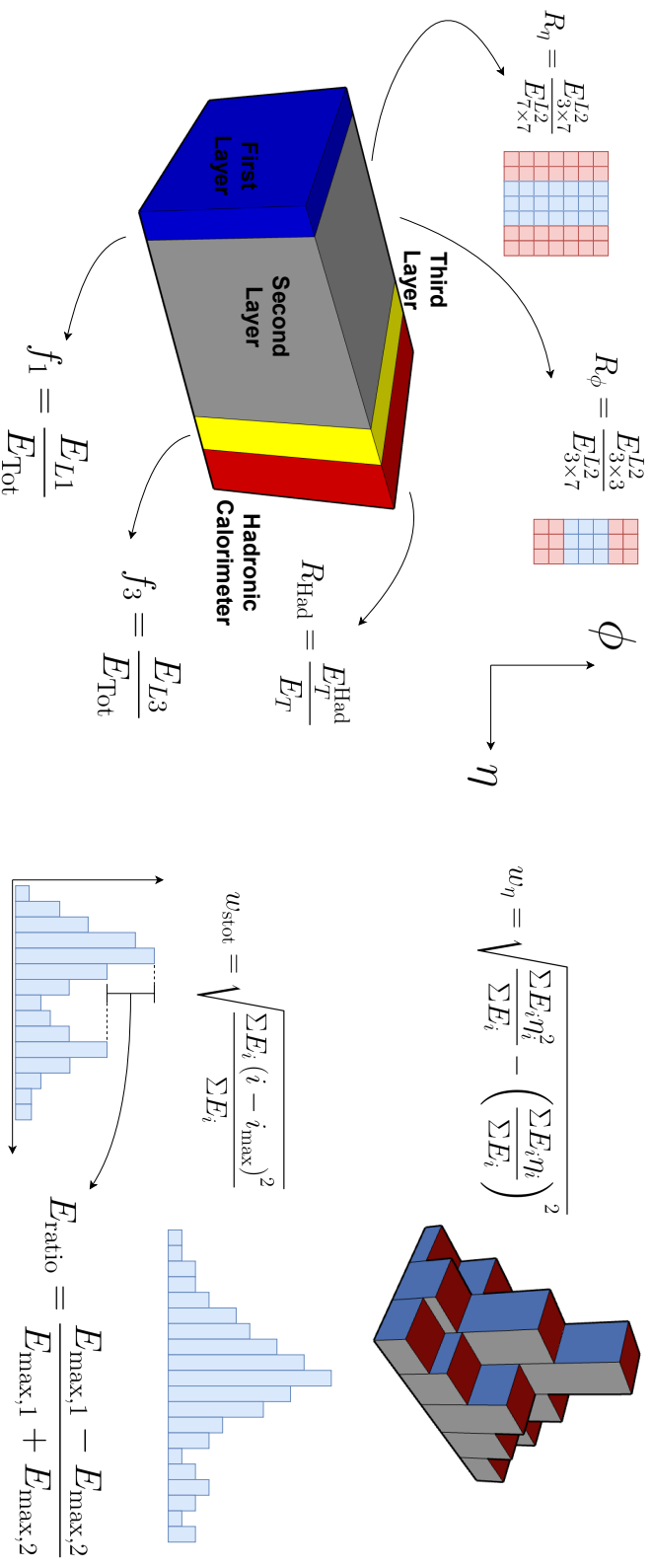


Figure 4.8: Depiction of the shower shape variables, listed in Table 4.2, used to create probability distribution functions for the electron identification LH. On the left are the ratios of energy between different regions of the calorimeter. On the right are the variables used for the shape of the shower in the calorimeter. Figure inspired by a similar figure in [75].

With the LHs for signal and background, a discriminating value is calculated

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \quad (4.7)$$

which peaks at one for signal and zero for background. To make it easier to apply selection criteria to the discriminating variable, the discriminant is transformed using an inverse sigmoid function,

$$\tilde{d}_{\mathcal{L}} = \frac{-1}{15} \ln(d_{\mathcal{L}} - 1), \quad (4.8)$$

where the factor of $1/15$ is found empirically. Selection criteria are then applied to this transformed discriminating variable $\tilde{d}_{\mathcal{L}}$, with different levels of background rejection and efficiency.

There are four fixed values of the LH used within the ATLAS Collaboration, which establishes four main working points for electron identification: *VeryLoose*, *Loose*, *Medium*, and *Tight*. Each subsequent working point has an increased threshold for the LH discriminant, leading to more background rejection but less overall efficiency. Examples of the efficiencies for three of the working points can be seen in Figure 4.9 as a function of a) E_T and b) η . As seen in the bottom ratio plot of Figure 4.9a), there is good agreement between data and MC for $E_T > 20$ GeV. Corrections to a portion of the shower shape variables listed in Table 4.2 improve the modelling at higher E_T but leads to an unintended larger efficiency at lower E_T in data [66] as seen in Figure 4.9a). This thesis uses the *Tight* LH working point for identification of electrons, as discussed in Chapter 8.

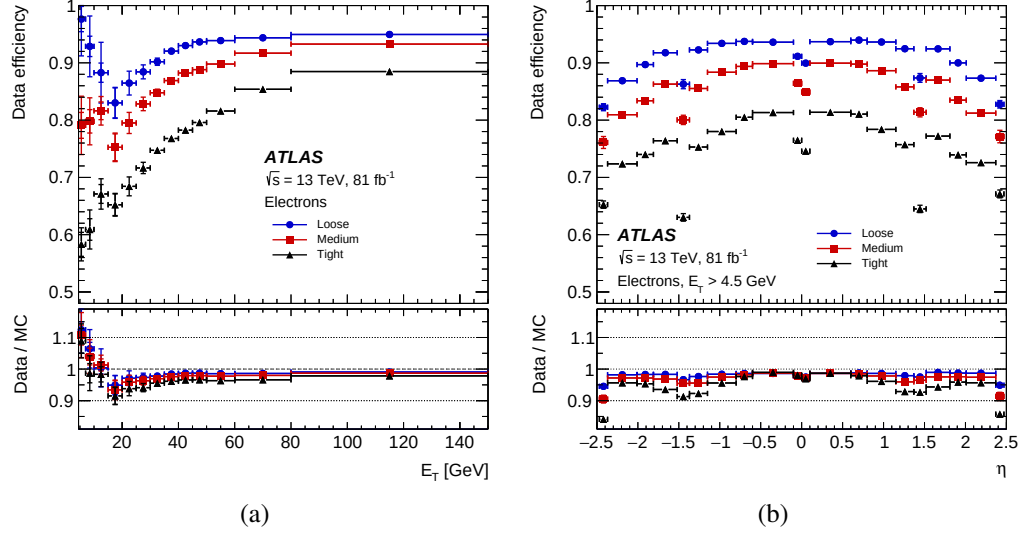


Figure 4.9: The measured LH electron identification efficiencies using $Z \rightarrow ee$ events as a function of a) E_T and b) η from [66]. The three working points: *Loose* (blue circles), *Medium* (red squares), and *Tight* (black triangles) correspond to different levels of background rejection. The *Tight* working point for identification of electrons is used in this thesis, as discussed in Chapter 8.

4.4.2 Muon Identification

As mentioned in Section 4.3.4, this analysis uses what is referred to as combined muons, which use ID and MS information. Identifying combined muons uses three discriminating variables to reduce the background of muons that originate from pion or kaon decay, which produce a non-prompt muon. These muons from hadronic decay will often have tracks with a kink structure, which result in the track from the ID and the track from the MS not having a good fit agreement. The bad fit agreement is used to choose the three discriminating variables [67]:

- q/p significance - The absolute difference of charge over momentum, measured in the ID and MS, divided by the sum in quadrature of the associated uncertainties.
- ρ' - The absolute difference between the p_T measurements in the ID and MS, divided by the p_T of the combined track.
- normalized χ^2 - From combined track fit.

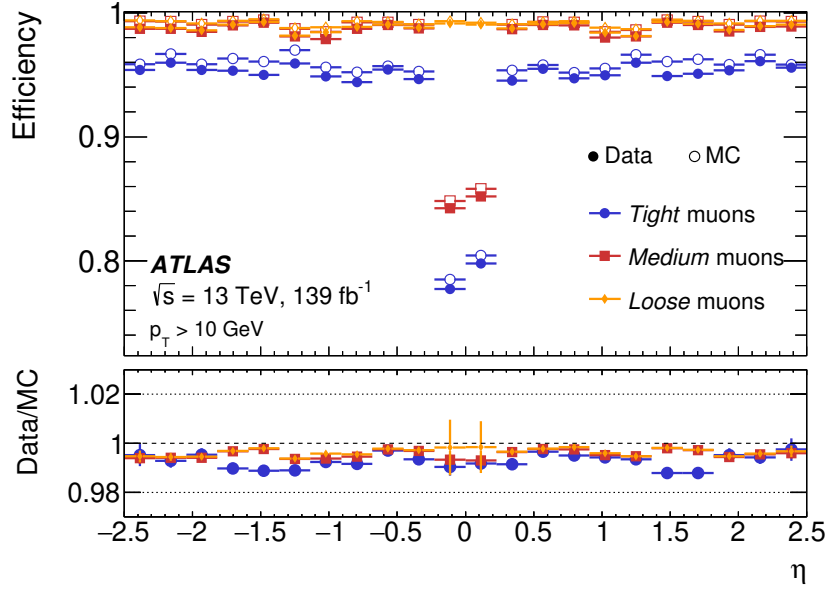


Figure 4.10: The measured muon and identification efficiencies using $Z \rightarrow \mu^+ \mu^-$ events as a function of η for muons with $p_T > 10 \text{ GeV}$ from [76]. The three working points: *Loose* (yellow triangles), *Medium* (red squares), and *Tight* (blue circles) are shown for data (filled) and MC (open). Each working point corresponds to more background rejection. The *Medium* working point is used in this thesis, as discussed in Chapter 8.

Similar to electrons, there are three working points for muon identification: *Loose*, *Medium*, and *Tight*, where the more stringent selection leads to more background rejection and lower efficiency. Examples of the efficiencies can be seen in Figure 4.10, which have a lower efficiency near 0 because of the coverage of the detector acceptance is reduced in this region. This thesis uses the *Medium* working point, as discussed in Chapter 8.

4.5 Particle Isolation

One way to reduce non-prompt background is by requiring the particle be well isolated. This is because non-prompt particles are often surrounded by the other particles that originate from the original prompt particles. There are two types of isolation: track based isolation and calorimeter based isolation. Both use the same principles but are applied to

different objects. Around the particle candidate, a cone of radius, ΔR , is used:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}, \quad (4.9)$$

where $\Delta\phi$ and $\Delta\eta$ are the distance of separation between the centre and the border of the cone in the azimuthal angle and pseudorapidity, respectively. Isolation for both electrons and muons are discussed in general, referring to them as leptons, with any differences noted.

Track based isolation takes the scalar sum of $p_T > 1$ GeV objects within the cone around the lepton candidate, excluding the p_T of the lepton candidate. The tracks are also required to point to the same primary vertex as the lepton candidate and satisfy basic quality requirements. When there is a busy detector environment, a variable sized cone dependent on the p_T of the lepton candidate, ΔR_{var} :

$$\Delta R_{\text{var}} = \min \left(\frac{10 \text{ GeV}}{p_T [\text{GeV}]}, \Delta R_{\text{max}} \right). \quad (4.10)$$

In this equation ΔR_{max} is the max allowable cone size, which is typically 0.2 for electrons [66] and 0.3 for muons [67].

Calorimeter based isolation can be more difficult, when compared to track based isolation, as parts of an energy cluster can be in cells outside of the cone. Instead, the energy in cells is added up in the same way as in Section 4.3.2 and the total energy is added up for clusters, where the centre is within the cone. The energy of the lepton candidate to be subtracted is approximated by adding the energy within a rectangle of $\Delta\eta \times \Delta\phi = 5 \times 7$ cells, centred on the lepton energy cluster. Typically, a cone of size $\Delta R < 0.2$ is used for both electrons [66] and muons [67].

To select lepton candidates, the amount of energy in the cone for both track and calorimeter based isolation is normalized by the transverse momentum of the lepton. An example of the isolation efficiency for electrons with different selection criteria can be seen in Fig-

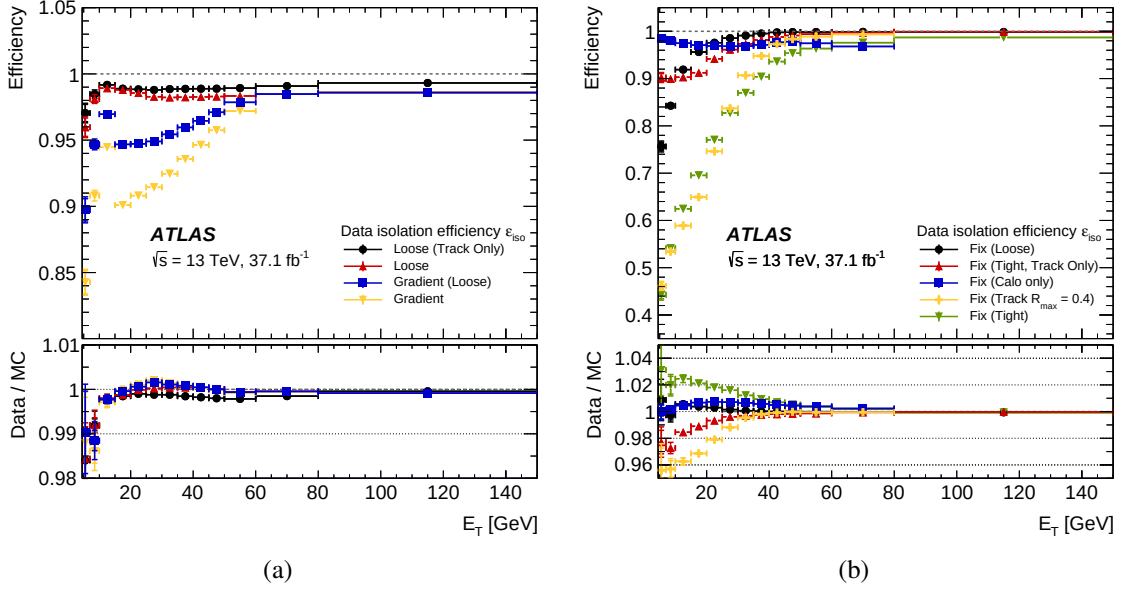


Figure 4.11: Example of data isolation efficiencies for electrons from [66] as a function of E_T . In a) *Loose* and *Loose (Track Only)* target a uniform efficiency in E_T and η , while the latter does not use the calorimeter based isolation. *Gradient* and *Gradient (Loose)* target an efficiency uniform in η , while depending on E_T . In b) efficiencies for different fixed selection requirements are shown. The ratio plot on the bottom shows the agreement between efficiencies from data and MC. A selection most similar to *Fix (Tight, Track Only)* is used in this thesis, as discussed in Chapter 8.

Figure 4.11. Here, *Loose* and *Loose (Track Only)* target a fixed efficiency of $\epsilon^{\text{iso}} = 0.99$ for each isolation category to be uniform in p_T and η , where the latter does not use calorimeter based isolation. The *Gradient* and *Gradient (Loose)* are designed to be uniform in η while depending on E_T to form a gradient. Lastly, fixed selection requirements on the isolation variables vary as a function of E_T and η . A fixed selection requirement of $\text{ptcone20}/\min(p_T^l, 50 \text{ GeV}) < 0.1$ is used in this thesis for electrons and muons, as discussed in Chapter 8. A fixed cone of size $\Delta R < 0.2$ is used and is normalized by whatever is smaller between p_T^l or 50 GeV. This selection is similar to the *Fix (Tight, Track Only)* selection in Figure 4.11, which requires $\text{ptcone20}/p_T^l < 0.06$.

4.6 Particle Trigger

The last efficiency to consider from Equation 4.1 is the trigger efficiency, ϵ^{trig} . For the analysis in this thesis, events with at least one electron with $E_T > 15$ GeV and passing *Loose* identification requirements or at least muon with $p_T > 14$ GeV are triggered. The trigger efficiencies for these working points were calculated for the ATLAS W and Z transverse momentum measurement [12], which uses the same data set as the analysis in this thesis. Similar efficiencies can be seen in Figure 4.12, which shows the trigger efficiencies for electrons and muons also have a turn-on curve, where the efficiencies increase as a function E_T .

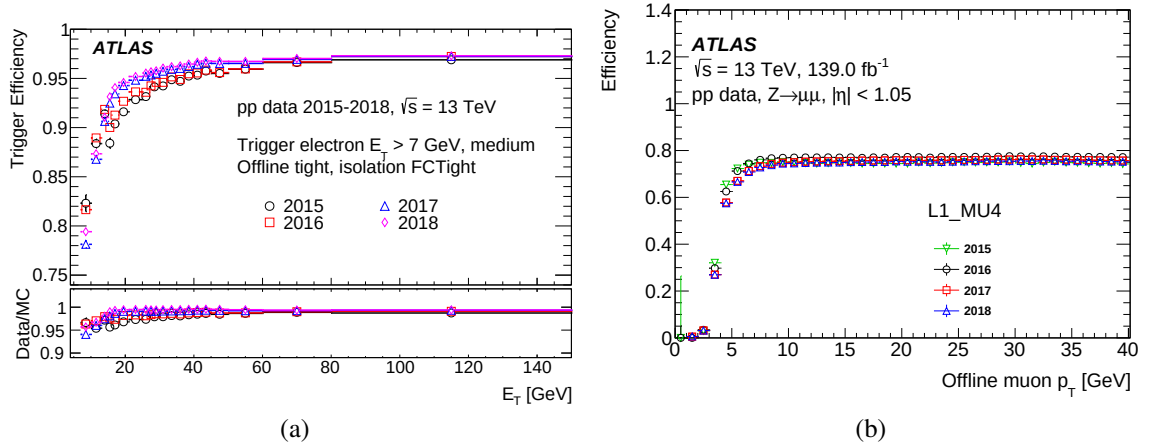


Figure 4.12: Example of a) the efficiency for the single electron trigger for electrons with $E_T > 7$ GeV from [77] and b) the muon trigger efficiencies from [78]. The plateau of $\sim 80\%$ for the muon trigger efficiency is driven by the requirement of the number of coincidence hits in the TGC [78]. These trigger efficiencies are similar to the ones used in this analysis.

CHAPTER 5

ELECTRON IDENTIFICATION VARIABLE CORRECTION

Section 4.4.1 described how a likelihood (LH) is used to identify electrons using the variables from the tracking detectors and calorimeters to create probability distribution functions as input for the LH. To find the expected efficiency of particle identification, $Z \rightarrow e^+e^-$ events are simulated using MC and are compared to the measured identification efficiencies calculated using data. How well the distributions of the variables used in the LH agree between data and simulation will affect how well these efficiencies compare. For the tracking and calorimeter variable distributions, listed in Tables 4.1 and 4.2, there are discrepancies between data and MC due to the complex geometries of the detector, which are difficult to model. This chapter will focus on the attempt to improve the agreement of these electron identification distributions. Throughout this chapter, “electron” will be used to refer to both electrons and positrons as the method is applicable to both, any differences will be noted.

The ATLAS detector was built and installed over approximately a decade, spanning the mid 1990s to mid 2000s, and has been operating since 2008. The ATLAS Collaboration has expanded the number of members since 2008. New members need to qualify as an author by completing an authorship qualification project that is of general benefit to ATLAS to appear on all ATLAS publications. The work within this chapter was part of my authorship qualification project. In addition to meeting the qualification expectations, it was also chosen to better understand the discrepancies in the tracking and shower shapes used for electron identification, which is important for precision W boson measurements. This methodology was not used for the main measurements that is the focus of this thesis, as is not fully implemented within ATLAS; however, it would ultimately improve future precision W boson measurements. With better electron identification the need for correction factors would be reduced, which would mean smaller associated uncertainties.

5.1 Variable Correction Factors

The electron identification efficiencies measured with data and MC differ, where one of the contributing factors is the difference in the shapes of some of the tracking and calorimeter variables, described in Tables 4.1 and 4.2. This mismodelling of the variables used in the LH is complex and is not always fully understood, which prevents correcting the simulation from first principles. The study in this chapter was done to determine if these variables could be empirically corrected in MC so they could better represent the actual distributions seen in data. By correcting the distributions directly, it would be expected that the MC and data efficiencies would be in better agreement. All of the variables in Tables 4.1 and 4.2 are corrected, with an exception of five variables. The number of hits in different detectors n_{Blayer} , n_{Pixel} , and n_{Si} are not included, as they are discrete variables and would not benefit from this method. The ratio of cluster energy over track momentum, E/p , is corrected using a different method while the variable R_{had1} is similar to variable R_{had} , so the corrections were expected to be transferable.

Each variable, x , is corrected using the difference in the mean, μ , of its distribution to shift the values using the relation:

$$x' = x + (\mu_{\text{Data}} - \mu_{\text{MC}}). \quad (5.1)$$

To stretch the distribution, the ratio of the standard deviation, σ , of the distributions can be applied to the variable x and to the MC mean, allowing Equation 5.1 to be rearranged to the relation:

$$x' = (x - \mu_{\text{MC}}) \frac{\sigma_{\text{Data}}}{\sigma_{\text{MC}}} + \mu_{\text{Data}}. \quad (5.2)$$

Examples of the effects of shifting, stretching, or a combination of the two can be seen in Figure 5.1. Shifting the distribution only moves the values by the difference of the data and MC mean, as in Equation 5.1. Stretching the distribution changes the overall width of

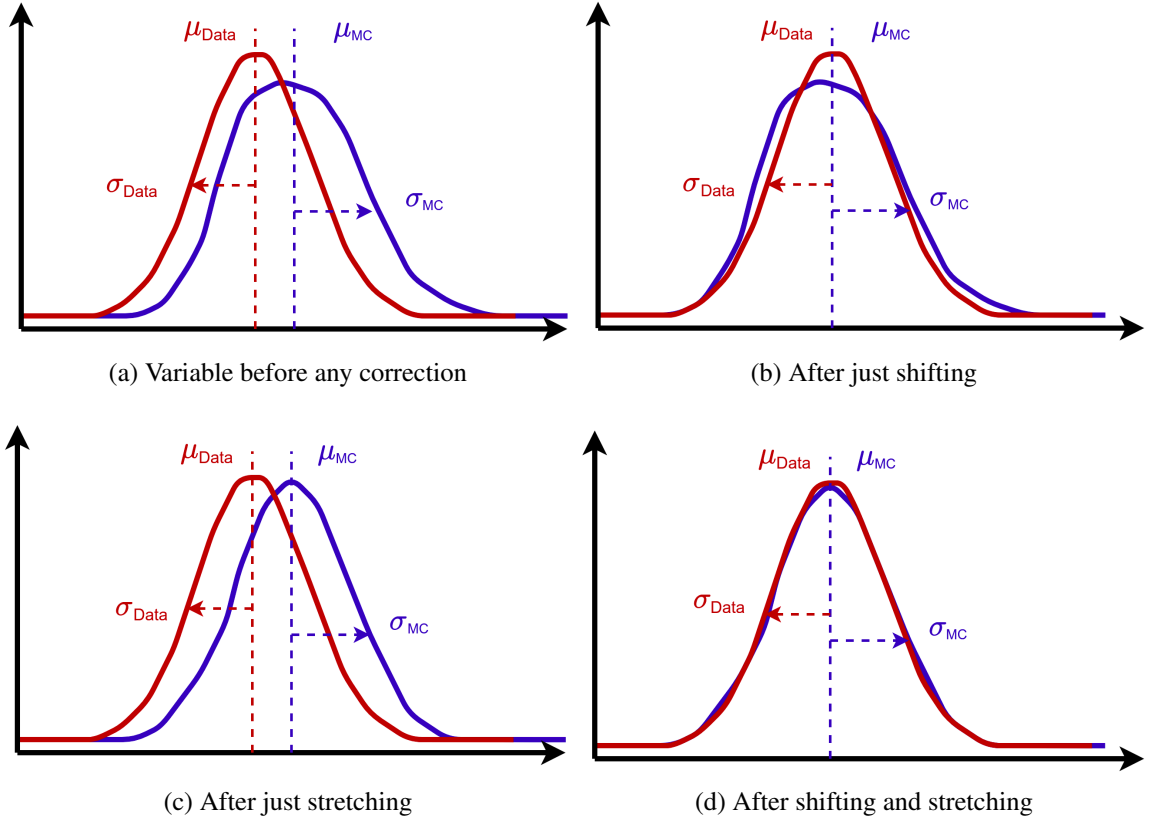


Figure 5.1: Example of the steps to correct a variable by shifting and stretching, according to Equation 5.2. a) The variables, mean and standard deviation, do not agree between data and MC. b) Shifting the variable by the difference in the means brings both distributions into better agreement. c) Correcting just the width of the distribution makes the shape in better agreement. d) Correcting the mean and the width of the variables improve the agreement.

the distribution by either making the distribution narrower – if $\sigma_{\text{Data}}/\sigma_{\text{MC}} < 1$ – or wider – if $\sigma_{\text{Data}}/\sigma_{\text{MC}} > 1$. By stretching the MC distribution, the points stay in the same position relative to its mean $x - \mu_{\text{MC}}$, while being scaled by the ratio of the standard deviation. Shifting and stretching changes the position of the mean of the MC distribution to be closer to the mean of the data distribution, while also stretching the distribution by scaling events by the ratio of the standard deviation.

The stretch and shift variables, known as variable correction factors (VCF), were developed in a two-dimensional phase space of either $(|\eta^l|, E_T)$ or $(|\eta^l|, \mu)$, where μ is the mean number of interactions per bunch crossing. To avoid confusion with the mean of the

data or MC distributions it should be noted that the mean number of interactions per bunch crossing will not contain a subscript. The distributions for these variables for both MC and data can be seen in Figure 5.2. The variables $|\eta^l|$ and E_T are commonly used to allow for granularity for efficiencies based on detector geometry and response. Variations in the detector geometry, amount of material in the ID, and the detector acceptance all change with η^l and will affect the efficiency. There are slight differences between efficiencies for $\eta^l < 0$ and $\eta^l > 0$ but $|\eta^l|$ is used for this study to increase statistics and the difference for signed η^l to be less than what would be gained by the correction.

The response of the detector and the distribution of the tracking and shower shape variables change with particle energy, so E_T is used for the energy of the electron. When more interactions occur within the ATLAS detector, there are two related phenomena that occur that are hard to model appropriately. First, there is more overlap in the showers in the calorimeter, making them more challenging to separate. Second, there are more interactions with the layers in the ID, which makes track reconstruction more difficult. The variable μ was also tested for variable correction as a portion of the variables are affected by the increase in interactions per bunch crossing. Figure 5.2c) does not show good agreement between data and MC as corrections to μ have not been applied here. However, this is not a cause for concern for two reasons. The first, is that they occupy the same phase space and there is not a region in μ that is not modeled in MC. And second, during the extraction of VCF – discussed in Section 5.2 – the MC is normalized to data within each $(|\eta^l|, E_T)$ or $(|\eta^l|, \mu)$ bin so any differences are minimized.

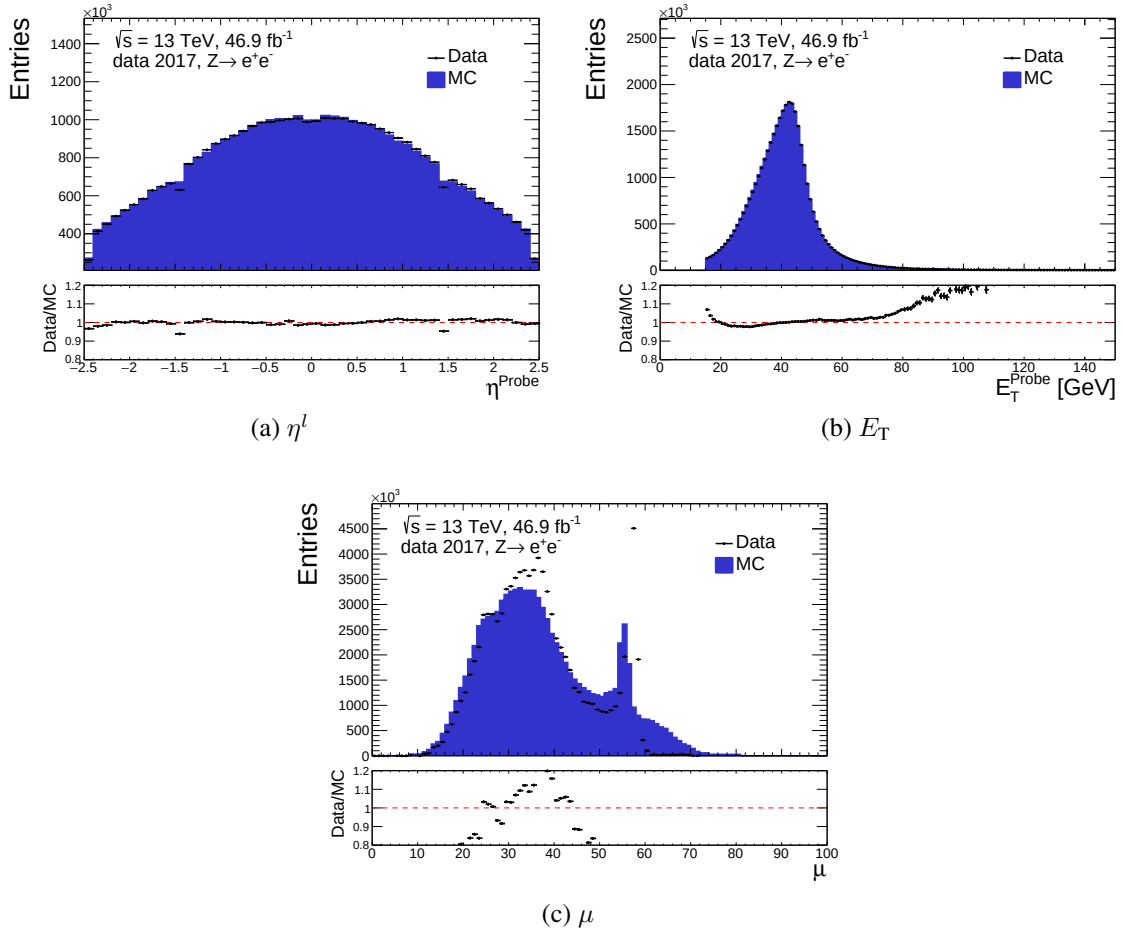


Figure 5.2: The variables used to create a two-dimensional phase space for the VCFs with a combination of $|\eta^l|$ and E_T or $|\eta^l|$ and μ . a) The signed η^l distribution, b) the transverse energy of the electron, and c) the mean number of interactions per bunch crossing, μ . In the upper plots, data and its statistical error are represented by black points and MC is the blue histograms. The ratio of data/MC can be seen in the lower plots, which shows the relative agreement.

5.2 Extraction of Variable Correction Factors

The VCF extraction was performed with a sample of probe electrons and the nominal tag-and-probe method described in Section 4.2. Both electrons are required to be of opposite charge and that the dilepton mass be in the mass window of $80 < m^{ll} < 100$ GeV. The tag electron needs to pass *Tight* identification selection [66], meanwhile, very little selection is applied to the probe electrons to avoid biasing the sample. These probe electrons are only required to pass *VeryLoose* identification selection [66]. Events were stored in two-dimensional histograms of the variables to correct and E_T or μ sliced in $|\eta^l|$ with the bin boundaries:

$$\begin{aligned} E_T &= [15, 20, 25, 30, 35, 40, 45, 80, 150] \text{ GeV}, \\ |\eta^l| &= [0, 0.1, 0.6, 0.8, 1.15, 1.37, 1.52, 1.81, 2.01, 2.37, 2.47], \\ \mu &= [0, 20, 25, 30, 35, 40, 45, 50, 55, 60, 100]. \end{aligned} \tag{5.3}$$

The bin boundaries for $|\eta^l|$ are motivated by the ATLAS detector geometry.

After the selection, these variables are stored in histograms that were then sliced in E_T or μ . From the histograms, the mean and standard deviation are calculated for the binned distributions to extract μ_{Data} , μ_{MC} , and $\sigma_{\text{Data}}/\sigma_{\text{MC}}$. The VCF corrections were then applied to MC on an event-by-event level, using Equation 5.2, to get corrected versions of the same histograms. An example of the MC before and after corrections for a tracking and shower shape variable can be seen in Figure 5.3.

During this extraction of VCFs, it was found that $\Delta P/P$ did not need to be corrected because the original distributions were generally in good agreement between data and MC. It was also found that only a few variables (d_0 , d_0/σ_{d_0} , eProbabilityHT, and R_{had}) had the VCF dependent on μ but all VCF were dependent on E_T . The μ -dependent variables behave in a similar way to d_0/σ_{d_0} in Figure 5.4a)-c) when plotted in slices of $|\eta^l|$. Here, the μ_{Data} and the $\sigma_{\text{Data}}/\sigma_{\text{MC}}$ change as a function of μ , while μ_{MC} does not. This means that

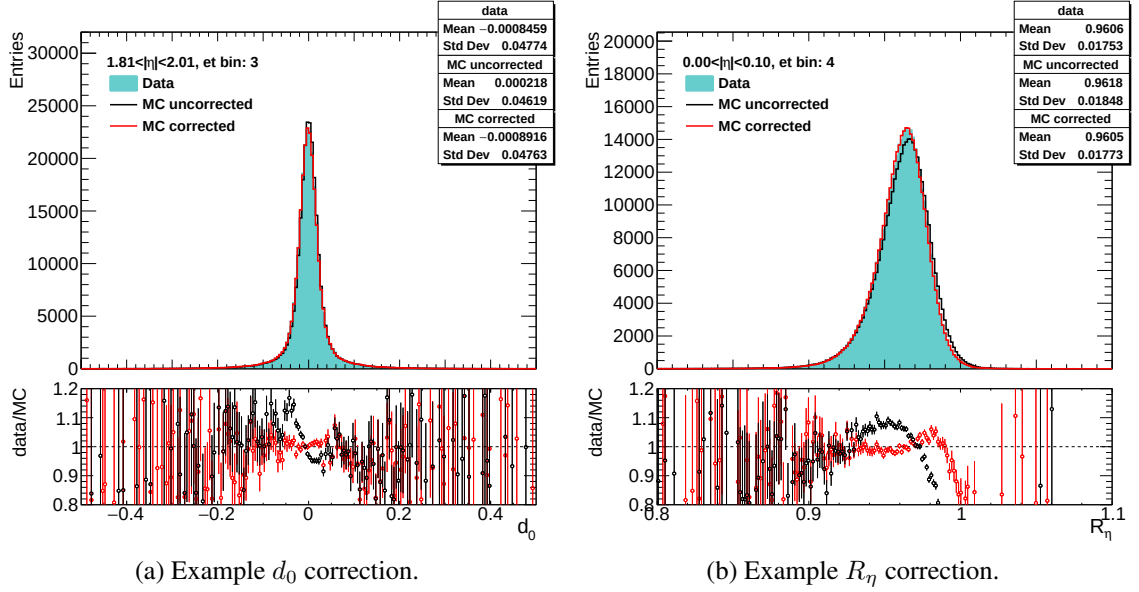


Figure 5.3: Example corrections are shown for a) one tracking variable, d_0 , in the region $1.81 < |\eta^l| < 2.01$ and $30 < E_T < 35$ GeV and for b) a shower shape variable, R_η , in the region $0 < |\eta^l| < 0.1$ and $35 < E_T < 40$ GeV. In the top plot the black points are the uncorrected MC, the red points are the corrected MC, and the blue histograms are the data. The ratio plot on the bottom shows the ratio of data and MC before correction in black and after in red.

both the shift and stretch will change with μ . However, most of the variables look similar to w_{η_2} in Figure 5.4d-f), where little to no variation, with respect to μ , is observed or the variation is within the statistical error of each other. Variables with no μ dependence are corrected only as a function of $|\eta^l|$ and are better corrected as a function of E_T and $|\eta^l|$.

Other studies were performed to look for different possible improvements, such as using the median instead of the mean, only calculating VCF with events within $4 \times \sigma$ of the mean to reduce the effect from the tails, and doing multiple VCF corrections iteratively. These studies were found to not improve results but application of other corrections before extracting the VCFs did improve results. These corrections modify the profile of the distributions by correcting for mismodelling in a different way than what is achievable by the VCF. The first additional correction is a reweighting of shower shape variables on a cell level, as described in Section 5.3. To improve the correction of eProbabilityHT, a trans-

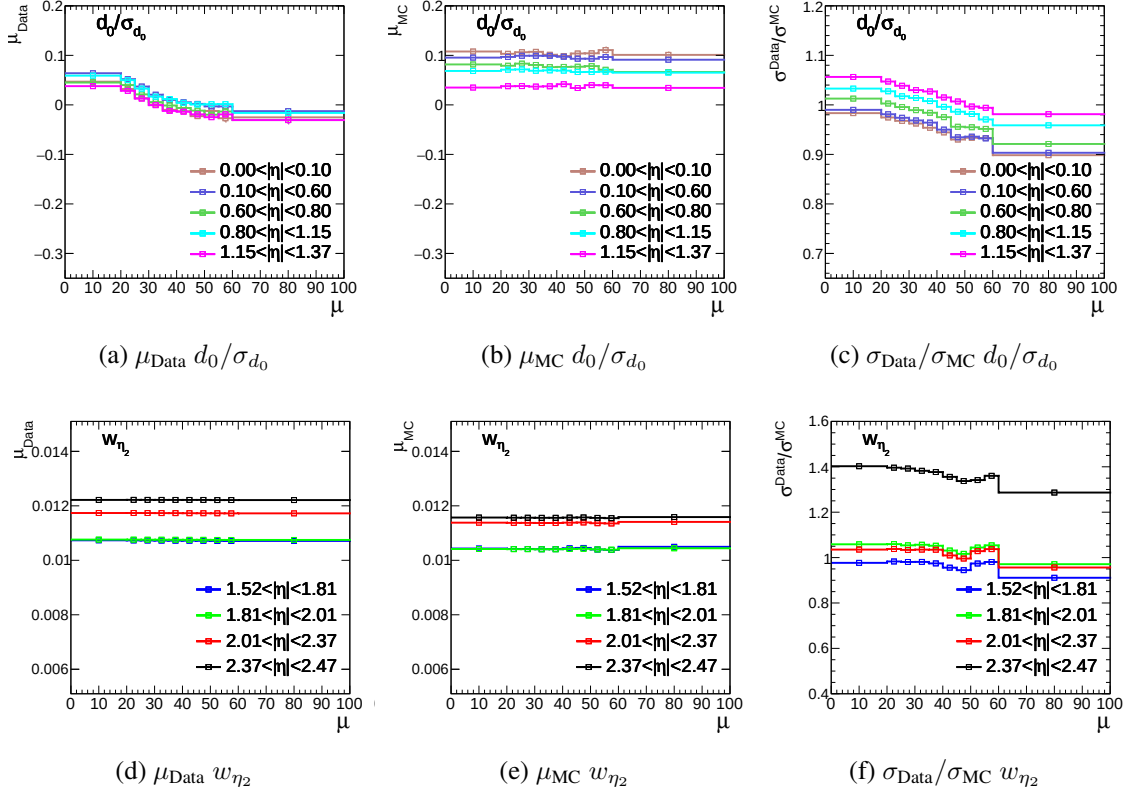


Figure 5.4: Examples of a variable that shows μ dependency in a)-c) and one that does not in d)-f). The dependency for d_0/σ_{d_0} can be seen in μ_{Data} and $\sigma_{\text{Data}}/\sigma_{\text{MC}}$ in the barrel region of the ECAL. There is not a strong dependency for some variables, such as w_{η_2} , which can be seen in all three variables in the endcap region of the ECAL.

formation was used to make the variable more Gaussian before being corrected and then transformed back, as described in Section 5.4. Lastly, a correction to $\Delta\eta_1$ was applied in order to correct for distortion of the ECAL alignment, as described in Section 5.5.

5.3 Shower Shape Reweighting

Previously [79, 80], studies were performed to correct shower shape variables for the discrepancy between MC and data distribution shapes with the same goal to improve compatibility between MC and data electron identification efficiencies. These studies used a shower shape reweighting (SSR), which was calculated using a data driven approach and applied a correction to improve the energy distribution at cell level in the ATLAS detector.

The VCF corrections shown in this chapter are looking at reconstructed variables that use the cell level information as input. This correction is needed as the modelling of energy deposited in the calorimeter does not always match what is seen in data. A two-dimensional reweighting in η and ϕ is done and only affects variables that are explicitly dependent on the energy deposited in the calorimeter: w_{η_2} , w_{stot} , R_η , and R_ϕ . The SSR is done by calculating a two-dimensional correction matrix that contains cluster cell corrections for each η slice calculated by

$$M_i^{\text{Correction}} = \frac{E_i^{\text{Data}}}{\sum E^{\text{Data}}} - \frac{E_i^{\text{MC}}}{\sum E^{\text{MC}}}, \quad (5.4)$$

where E_i^{Data} and E_i^{MC} are the matrix elements of the averaged energy profiles for data and MC, respectively. Because this is a redistribution of energy, the sum over $M_i^{\text{Correction}}$ in all bins should be zero. The matrix, $M_i^{\text{Correction}}$, is then used to reweight the electron cell clusters' energy for all MC events using

$$E'_i = E_i (1 + M_i^{\text{Correction}}). \quad (5.5)$$

The SSR improves the agreement between MC and data and allows the VCF corrections to further correct the variables w_{η_2} , w_{stot} , R_η , and R_ϕ . An example from [80] can be seen in Figure 5.5, which shows the improvement for R_η in the region $0.40 < \eta^l < 0.60$ before correction in red and after correction in blue.

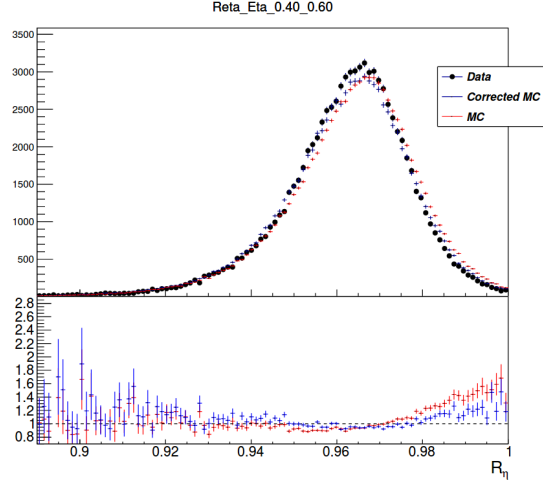


Figure 5.5: An example of the shower shape reweighting (SSR) applied to R_η in the region $0.40 < \eta^l < 0.60$ from [80]. Data is denoted by the black points, MC uncorrected by the red points, and MC after SSR reweighting denoted by the blue points. The ratio of MC over data is shown on the bottom for both the uncorrected MC in red and corrected in blue and shows improvement in the relative agreement, especially in the peak region.

5.4 eProbabilityHT Transformation

The variable eProbabilityHT has a unique shape because it is the LH probability based on transition radiation in the TRT, as seen in Figure 5.6. It has a maximum in the last bin before 1 and approximately exponentially decays as it approaches a value of 0. Because of this non-Gaussian shape, the mean and standard deviation are not the best descriptors for this distribution. To improve corrections for eProbabilityHT, events are transformed with the inverse sigmoid function:

$$x' = \frac{-1}{15} \ln(x^{-1} - 1), \quad (5.6)$$

where x is the value of eProbabilityHT. This function is also used to transform the discriminating variable in Section 4.4.1. The transformed eProbabilityHT, T(ePHT), allows for the distribution to appear more Gaussian, which allows the mean and standard deviation to be a better representation of the distribution. The VCFs are derived for the transformed distribution and also applied to T(ePHT). After being corrected, T(ePHT) is transformed

back to eProbabilityHT using the inverse of Equation 5.6, defined as

$$x = \frac{1}{1 + e^{-15x'}}. \quad (5.7)$$

An example of this process can be seen in Figure 5.6, where on the left is eProbabilityHT before it is transformed and after the VCF correction applied to T(ePHT) on the right. Doing a similar correction to other variables is either not viable, as it requires the value to be between 0 and 1, or the variables are already approximately Gaussian. A similar transformation was attempted for the variables that are bound between 0 and 1, but the results were not improved with this method.

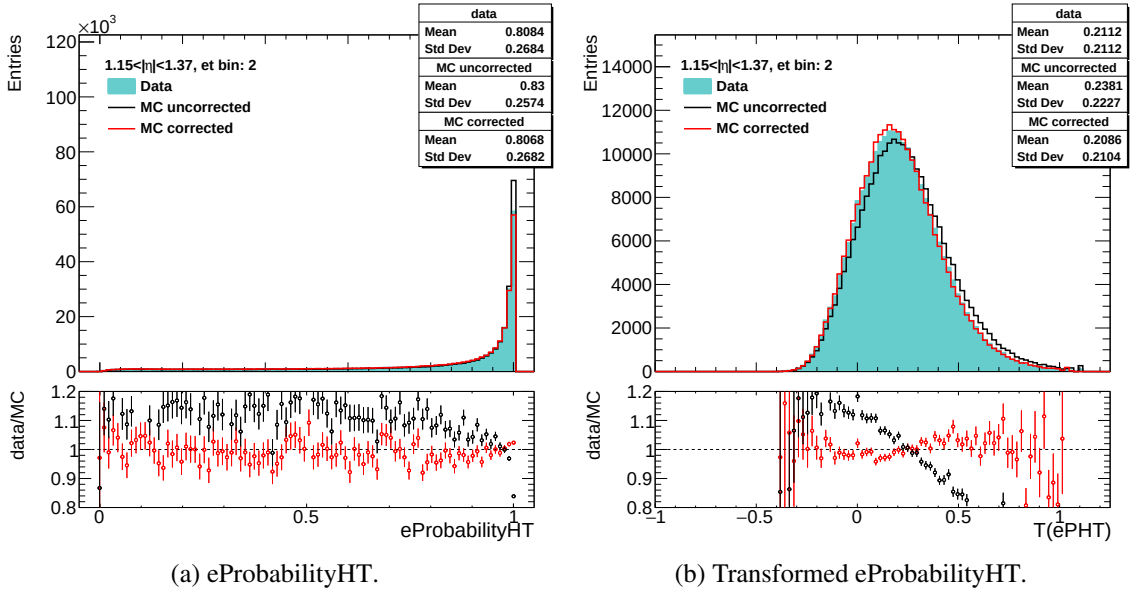


Figure 5.6: An example of the extra steps need to correct eProbabilityHT for the region $1.15 < |\eta^l| < 1.37$ and $25 < E_T < 30$ GeV. a) Is eProbabilityHT before being transformed using Equation 5.6. b) The transformed distribution, T(ePHT), is used to calculate the VCFs before it is transformed back using Equation 5.7. In the top plot the black points are the uncorrected MC, the red points are the corrected MC, and the blue histograms are the data. The ratio plot on the bottom shows the ratio of data and MC before correction in black and after in red.

5.5 Pear Shaped Correction

The barrel of the ECAL is held by two support rails at the top of the subdetector and gravity causes it to sag from its own weight. The level of distortion has been previously measured [81, 82] to be on the order of a few mm downwards, which is a small deviation compared to the overall diameter of the ECAL, 4 m. An over-exaggerated depiction of what this would look like can be seen in Figure 5.7.

This distortion has a measurable effect on the drift time because it changes the gap spacing between layers in the LAr calorimeter [83]. Like the shower shapes discussed in Section 5.3, this distortion in the shape is not properly modeled in MC, so a correction to the now “pear” shape calorimeter needs to be applied. This correction, $\delta(\eta, \phi)$, is applied as a linear shift to the variable $\Delta\eta_1$, $\Delta\eta'_1 = \Delta\eta_1 + \delta(\eta, \phi)$, as a function of position

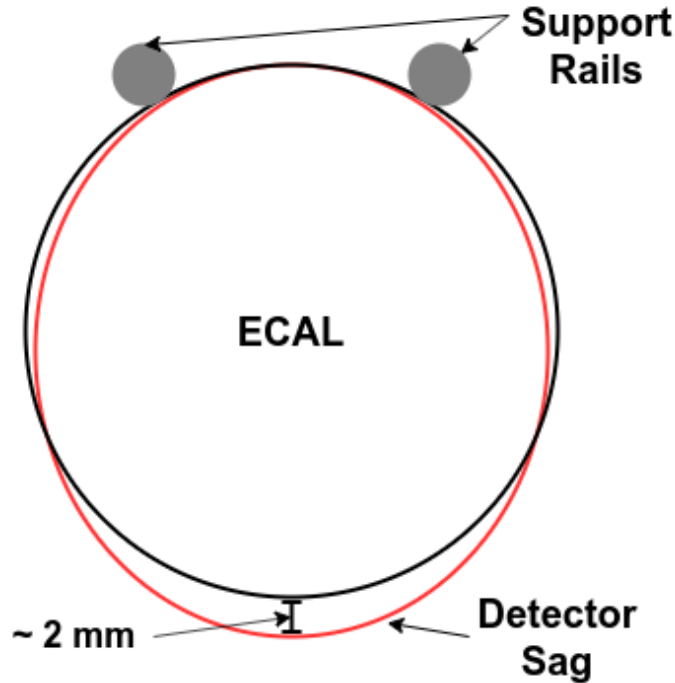


Figure 5.7: The ECAL is supported by two rails at the top of the subdetector; gravity causes the detector to deform from the intended circular shape of the ECAL, in black, into the distorted shape, in red. This distortion is on the order of a few mm, causing a difference in the spacing between layers in the LAr calorimeter, which has an outer diameter of 4 m.

parameterized by

$$\begin{aligned} \delta(\eta, \phi) = & \alpha_0(\eta) + \alpha_1(\eta) \cos 2\phi + \alpha_2(\eta) \cos 3\phi + \alpha_3(\eta) \cos 4\phi \\ & + \alpha_4(\eta) \sin \phi + \alpha_5(\eta) \sin 2\phi + \alpha_6(\eta) \sin 3\phi. \end{aligned} \quad (5.8)$$

The coefficients, α_i , are parameterized as a function of η , which are coupled to an expansion in $\cos n\phi$ and $\sin n\phi$. The η, ϕ phase space for this correction can be seen in Figure 5.8.

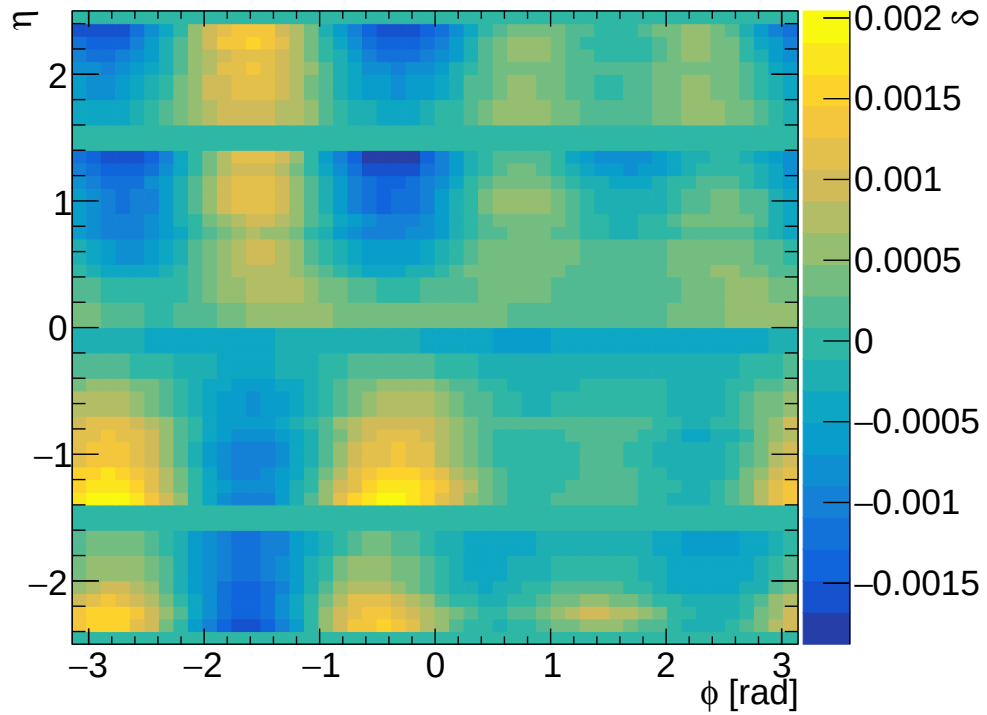


Figure 5.8: The phase space of the δ correction to $\Delta\eta_1$ from Equation 5.8 as a function of ϕ and η .

This pear shaped correction to $\Delta\eta_1$ was performed to correct the difference in shapes before VCF extraction and correction. This allowed the corrected MC to better match the data, as seen in Figure 5.9.

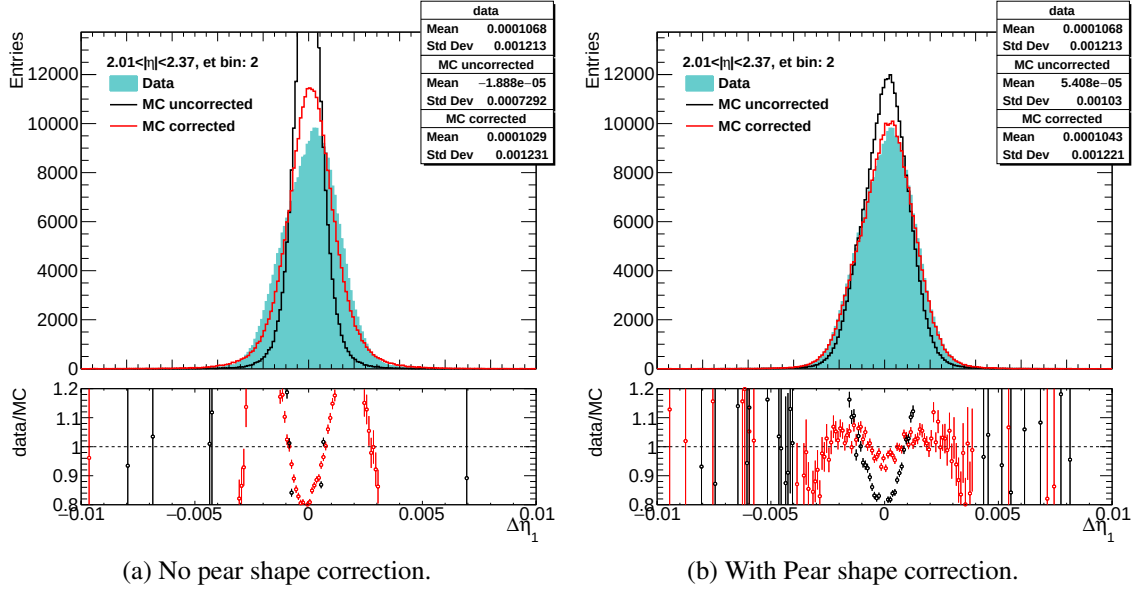


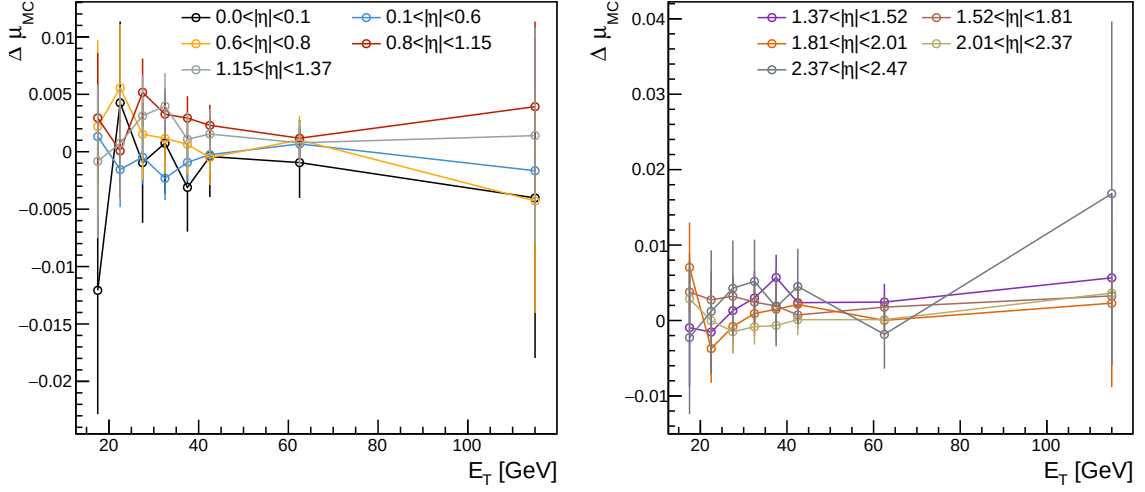
Figure 5.9: The original MC modeled Gaussian shape in black doesn't match with the data shape in the blue histograms. a) The correction without applying a pear shape correction and the VCF corrections are not able to correct the difference. b) The pear shape correction is applied and allows the VCF to bring the agreement within 10%.

5.6 Sample Background Consideration

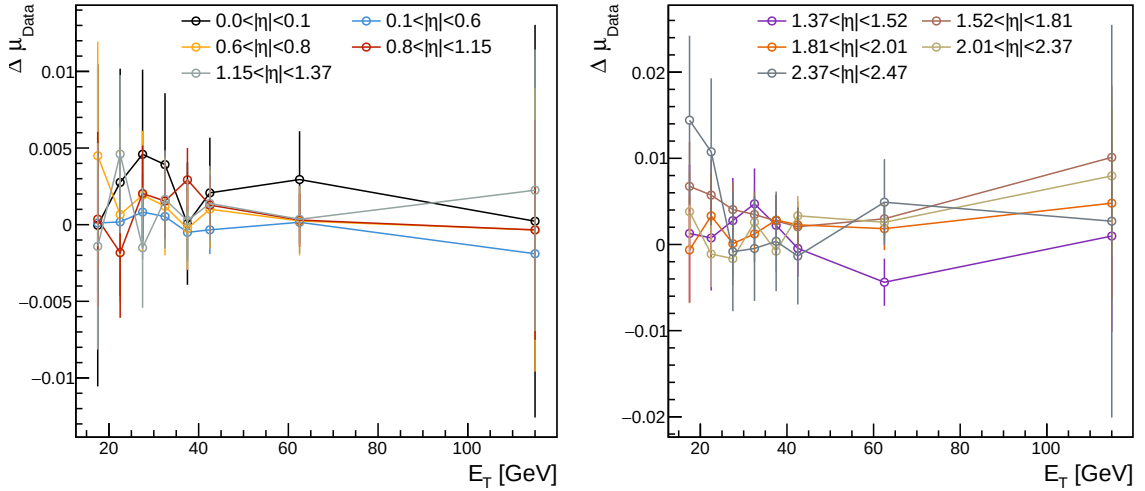
In the data samples used, there are events that pass the selection criteria that mimic the characteristics of the desired signal but originate from events that are not $Z \rightarrow ee$ and are known as background. The VCFs were originally derived from events that pass a *VeryLoose* identification selection criteria, which increases the statistical significance but will have more background contamination. This background is not included in the MC samples, so there is a possibility that the VCFs are correcting signal MC to look more similar to background in data, resulting in lower efficiencies. To test if this is the case, a more restrictive type of event selection was used to derive the VCFs by using probes that pass *Tight* selection [66]. The more restrictive event selection removes both signal and background events, lowering the statistical significance. This removes most of the background allowing a lower relative abundance of background in the data.

Both *VeryLoose* and *Tight* selection were applied to the probe electron. The tracking and

shower shape variables were plotted for each $|\eta^l|$ and E_T bin, which were found to not differ by a noticeable amount. This suggests that background does not contribute significantly to the shapes of the distributions. This could be because there is not enough background to have an effect on the shape or the background has a similar shape to the desired signal of $Z \rightarrow e^+e^-$. The VCF variables were compared and the difference was equal to zero within one standard deviation, as shown for d_0/σ_{d_0} in Figure 5.10. This suggests that the VCFs derived from the *VeryLoose* selection are not changed by difference levels of background contamination and should work for varying level of applied selection criteria.



(a) Difference of μ_{MC} between *VeryLoose* and *Tight* identification selection



(b) Difference of μ_{Data} between *VeryLoose* and *Tight* identification selection

Figure 5.10: Example of the difference between values derived from probe electrons selected with *VeryLoose* and the *Tight* selection criteria for VCF variables μ_{MC} and μ_{Data} for the variable d_0/σ_{d_0} . On the left is the barrel region of the ECAL and on the right is the endcap region. This comparison found, that for most variables, the difference is equal to zero within one standard deviation for most $|\eta^l|$ and E_T bins.

5.7 Efficiency and Scale Factor Calculation

With the VCF derived, the next step is to test if the VCF gives the desired outcome of correcting the MC efficiency to better reflect what is measured with data. The corrections were applied during the Tag&Probe selection of events to see their effect on the efficiencies and the subsequent SFs. The corrections were applied to tracking variables $\Delta\eta_1$, $\Delta\phi_{\text{rescaled}}$, and all of the shower shape variables listed in Table 4.2. The rest of the tracking variables are recalculated during the Tag&Probe selection of events and, therefore, were unable to have the VCFs applied during this study. By correcting the listed variables, it can be seen whether the VCF corrections worked as intended.

Different correction configurations were applied to separate the individual effects of each correction. First, a baseline was established by applying the SSR and pear shape corrections. Then, the application of VCFs were split in two: just applying the shift, and applying both the stretch and the shift. The efficiencies were calculated for the four different levels of electron identification selection: *VeryLoose*, *Loose*, *Medium*, and *Tight*, previously described in Section 4.4.1. This verifies that the amount of background, as discussed in Section 5.6, does not affect the results. Some examples of the derived efficiencies in E_T slices and binned in η for *Loose* selection [66] can be found in Figure 5.11, where they are compared to data.

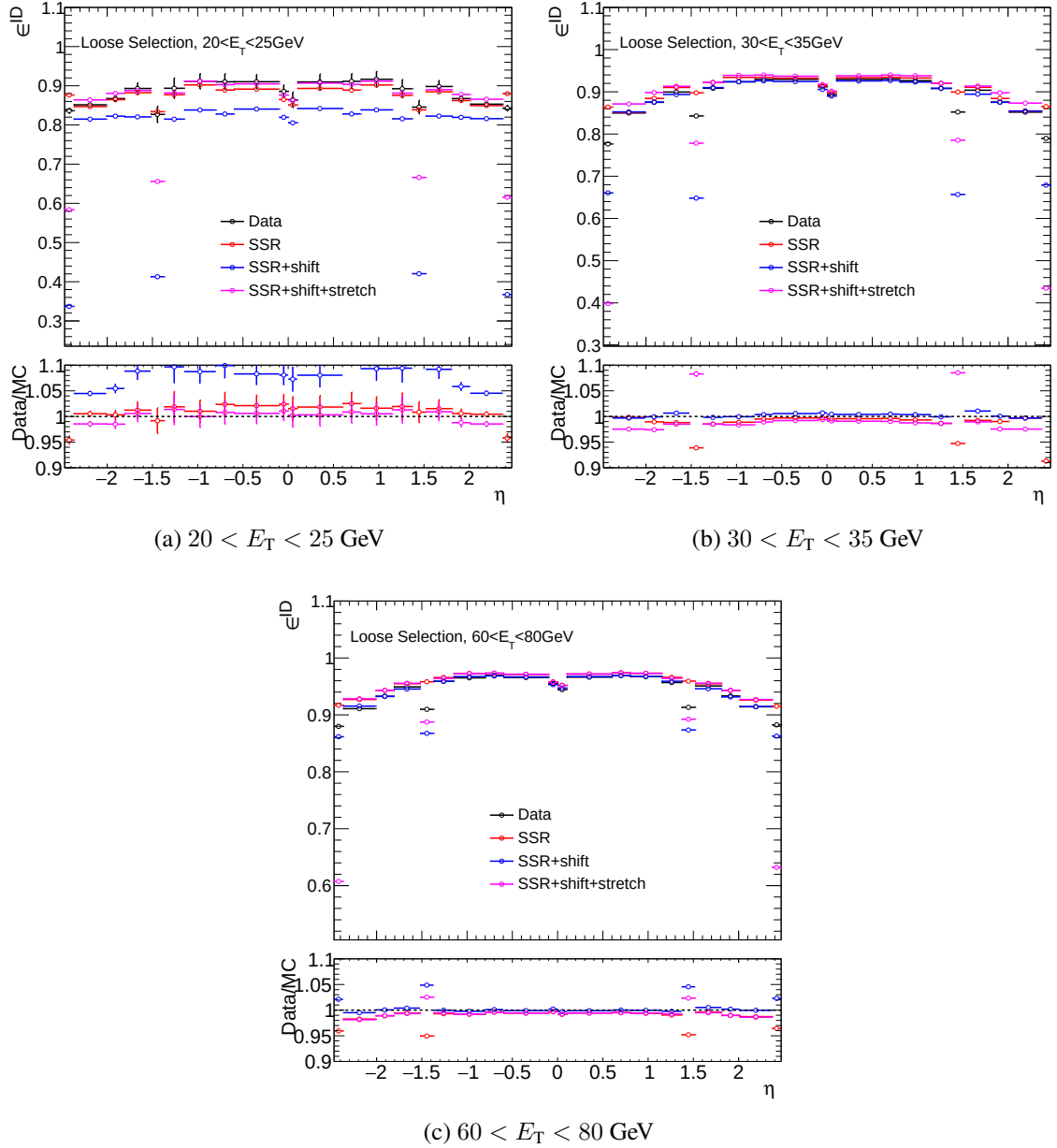


Figure 5.11: Examples of the electron efficiencies with different correction configurations derived from Tag&Probe for the *Loose* identification selection in three E_T slices. They are compared to the measured efficiency in data in black. The configurations are applying just SSR and pear shape correction in red, adding the VCF shift in blue, and adding the stretch in pink. The ratio plots show the relative agreement between data and the configurations.

By correcting the variables, the MC efficiencies should have closer agreement to the measured data efficiencies. The level of agreement can be seen in the ratio plots in the bottom panels of Figure 5.11. For all levels of identification selection, it was found that for $E_T < 30$ GeV the stretch and shift brought the efficiencies closest to that of data. However, using just a shift resulted in better agreement for the rest of the E_T range. This suggests that the central value of the distributions is the main cause for the efficiency difference, but is not the case at lower energies. Both VCF configurations (applying only the shift and applying the shift and the stretch) did not improve the efficiency in the transition region of the calorimeter and the outermost η bins, where the agreement was not improved. This is not surprising as the VCF often did not improve the agreement to the same extent as other bins, as seen in Figure 5.12. Both of these regions have different detector geometry compared to their neighbouring bins and this appears to have a notable effect.

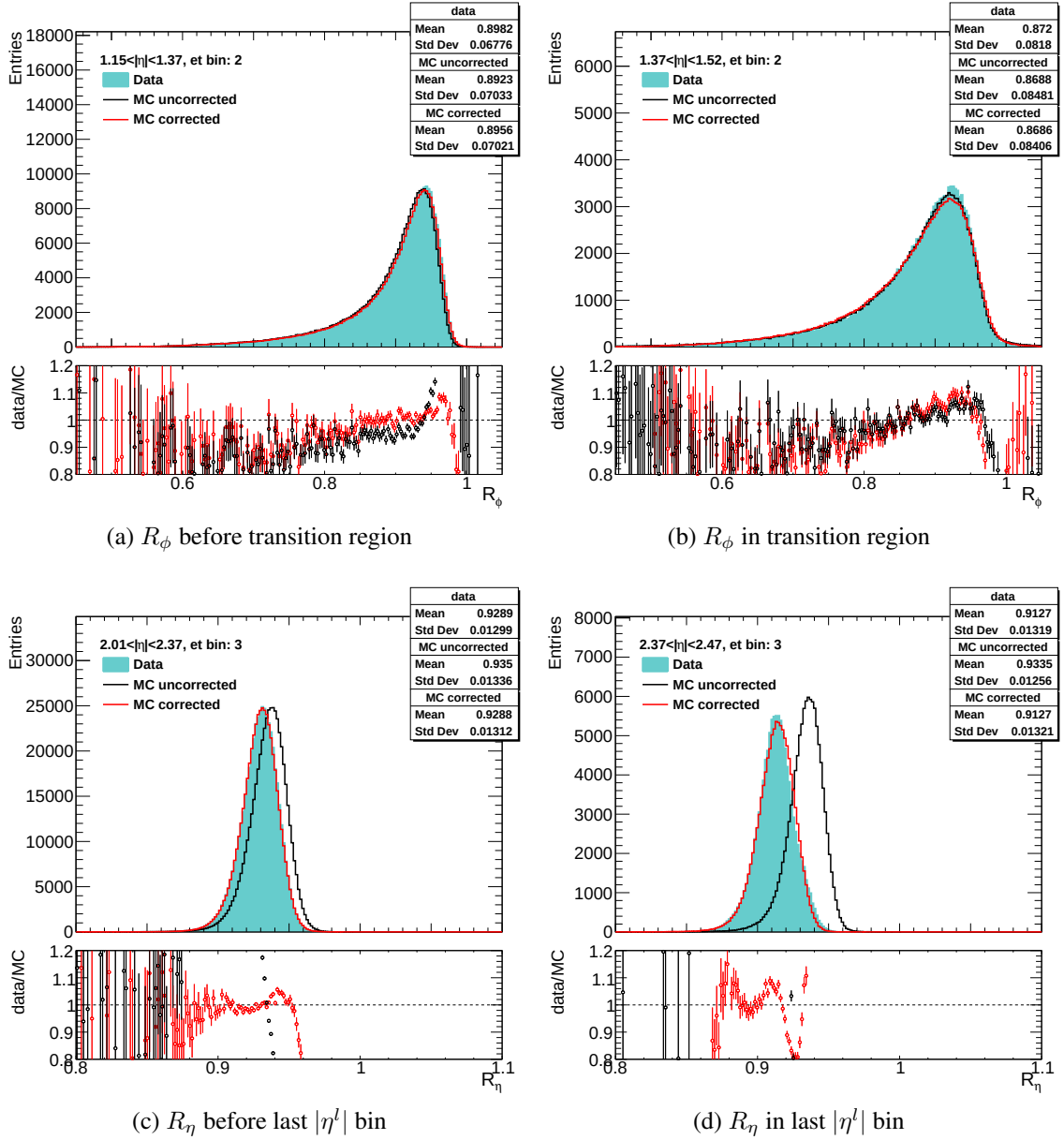
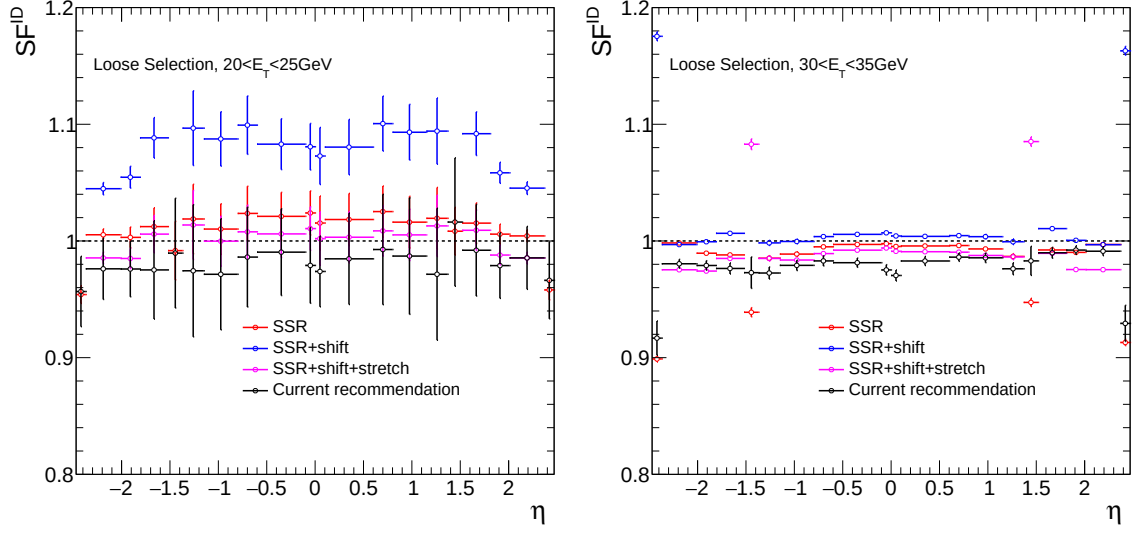


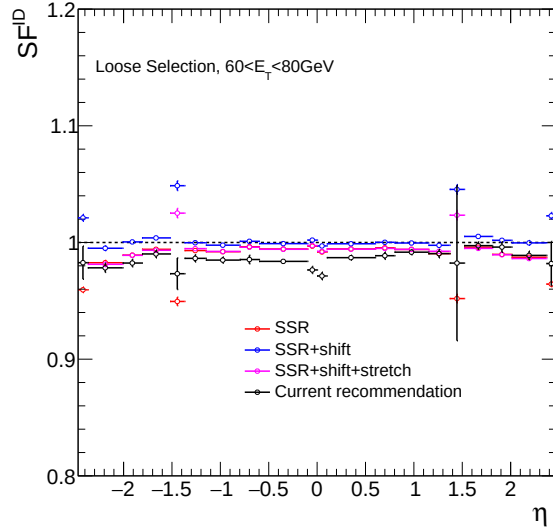
Figure 5.12: Comparison of corrections for different $|\eta^l|$ bins to demonstrate that the corrections do not improve the data and MC agreement evenly for all $|\eta^l|$ bins. a) R_ϕ and c) R_η for $|\eta^l|$ bins that the corrections improve the data and MC agreement. b) R_ϕ and d) R_η but in neighbouring bins that do not see the same level of improvement for the data and MC agreement. This commonly occurs in regions where the detector geometry changes significantly. The top shows this for R_ϕ near the transition region, $1.37 < |\eta^l| < 1.52$, of the ECAL and the bottom is R_η near the outside edge of the ECAL.

With the efficiencies calculated for MC and data, the SFs can be calculated and compared to the previously uncorrected SFs, denoted by the black points labeled “current recommendations” [84] in Figure 5.13. The SFs derived from the corrected electrons are closer to unity than the recommended SFs, showing that the correction method is working as intended. The efficiencies and SFs from the corrected electrons only contain statistical uncertainties, as they were calculated as a proof of principle to determine their viability. For the methodology to be fully implemented and be properly compared to the current recommendations would require evaluating systematic uncertainties related to the methodology. These SFs are not used for the final measurements in this thesis because all sources of uncertainties are not fully implemented. Another reason these SFs were not used in the main measurements that are the focus of this thesis is they were not calculated for the special data set taken with special run conditions, that will be outlined in Chapter 8, where these SFs do not apply. This proof of principle, delivered to ATLAS, shows that it is viable and makes a clear improvement to the kinematic modelling for some variables used in the electron identification. With a full implementation of all sources of uncertainties, this methodology could improve future precision measurements of the W boson in the electron channel. This chapter also gives insight into how SFs, such as the set used in the final measurements, are calculated. In addition it serves as an example of improvements that the ATLAS Collaboration continue to develop to improve their measurements and description of the detector.



(a) $20 < E_T < 25\text{GeV}$

(b) $30 < E_T < 35\text{GeV}$



(c) $60 < E_T < 80\text{GeV}$

Figure 5.13: Examples of the probe electron SFs using the Tag&Probe framework for the *Loose* selection in three E_T slices. Here, the baseline SSR and pear shape correction in red, adding the VCF shift in blue, adding the stretch in pink, and to the current recommendations [84] in black.

CHAPTER 6

DRELL–YAN PROCESS

This chapter will discuss the Drell–Yan process, $q\bar{q}' \rightarrow W^\pm \rightarrow \ell^\pm \bar{\nu}$, and the associated angular coefficients. The background information in this chapter is based on sources [2, 3, 4], where more extensive explanations can be found. Calculated for this thesis, the expected values of the Drell–Yan angular coefficients, their characteristics, and their physical interpretation will be discussed.

Proton–proton interactions can produce a wide variety of particles, each with their own probability of occurring at different centre-of-mass energies. This probability is related to a measured quantity, known as the cross-section denoted by σ . Examples of the cross-section for some of these different processes can be seen in Figure 6.1, which shows the many orders of magnitude that measured cross-section values can have at the LHC.

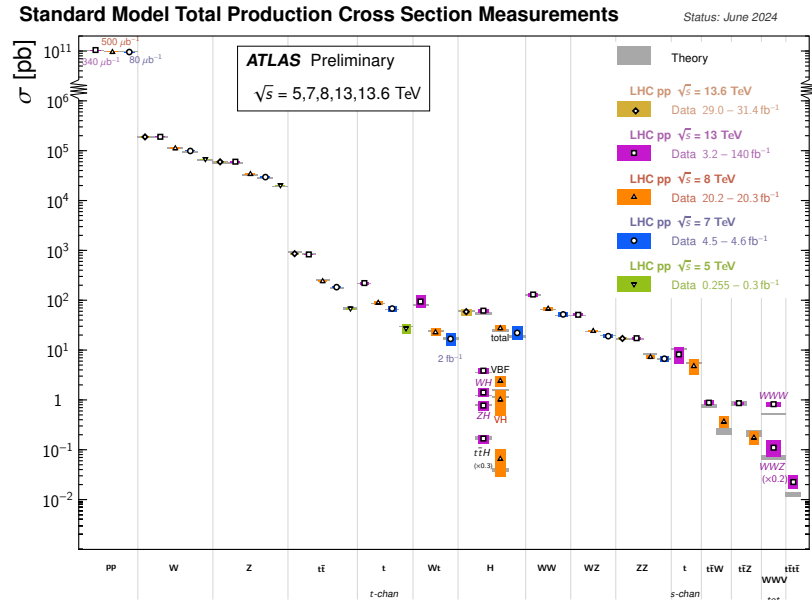


Figure 6.1: The cross-sections for different processes produced from proton–proton interactions, measured by ATLAS [85] for centre-of-mass energies $\sqrt{s} = 5, 7, 8, 13, 13.6 \text{ TeV}$. These cross-section measurements span many orders of magnitude and depend on the probability of the interaction to occur.

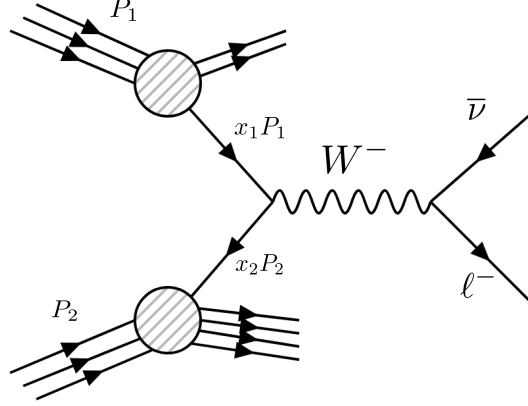


Figure 6.2: Example of a Drell–Yan process, where protons of momenta P_i , interact via partons with momenta, $x_i P_i$, annihilate into a W^- boson, which decays into two leptons.

An important production process at the LHC is the Drell–Yan process [2, 3], where proton–proton interactions produce gauge bosons, $V = \{\gamma, W, Z\}$, before decaying into a lepton pair final state. The γ and Z bosons are produced via $q\bar{q}$ annihilation and the W bosons are produced by $q\bar{q}'$ interactions. Figure 6.2 shows an example of this with $q\bar{q}' \rightarrow W^- \rightarrow \ell^- \bar{\nu}$, where protons with momentum P_i , interact via partons that carry a fraction, $x_i < 1$, of the protons momentum, $x_i P_i$.

Within hadrons, the probability of having a parton i with longitudinal momentum fraction x_i is described by the Parton Distribution Function (PDF), $f_i(x)$. The combined momentum fraction of all partons is equal the total momentum of the hadron:

$$\int_0^1 x dx \left[\sum_q (f_q(x) + f_{\bar{q}}(x)) + f_g(x) \right] = 1, \quad q = \{u, d, s, c, b, t\}. \quad (6.1)$$

The PDFs are obtained experimentally using data from both deep inelastic scattering in lepton–hadron collisions and hadron–hadron collisions. An example of deep inelastic scattering is the HERA data [86] measured by H1 and ZEUS Collaborations. Whereas, proton–proton collisions at the LHC are an example of hadron–hadron collisions used. PDFs depend on the momentum transferred between interacting particles, Q , which is commonly chosen to be $Q^2 = 1.9 \text{ GeV}^2$ to avoid including charm and heavier quarks that do not

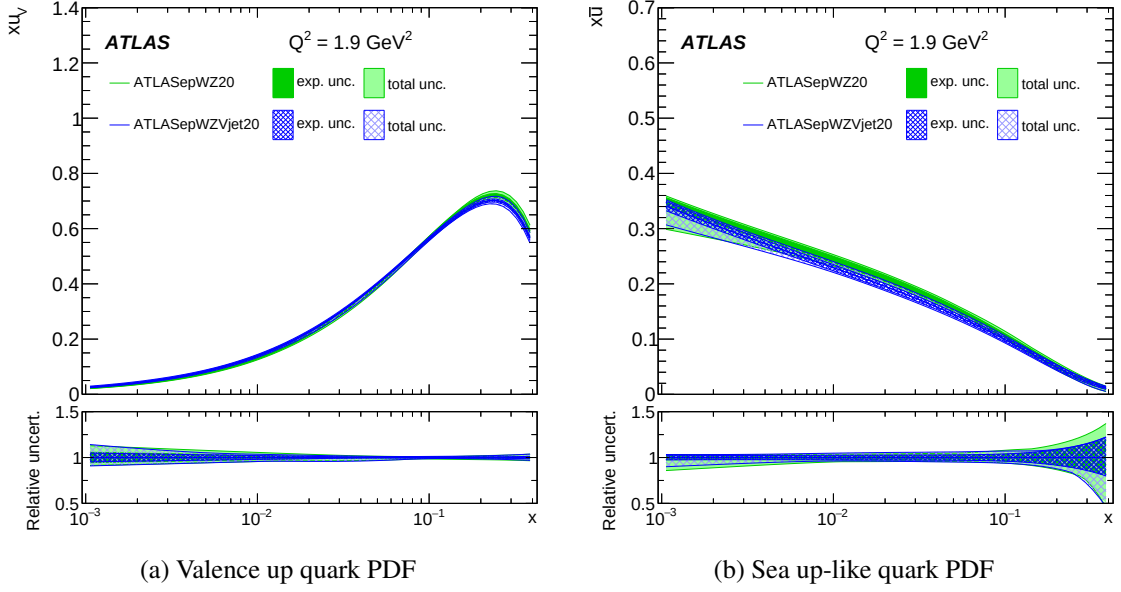


Figure 6.3: Examples of parton distribution functions (PDFs) multiplied by their momentum fraction at the scale $Q^2 = 1.9 \text{ GeV}^2$ as a function of their momentum fraction from [87]. a) The PDF for the valence up quark and b) the PDF for the sea up and anti-up quarks obtained from fitting ATLAS and HERA [86] data.

contribute significantly in proton–proton interactions. Examples of up quark PDFs multiplied by their momentum fraction as a function of their momentum fraction can be found in Figure 6.3. In Figure 6.3a), the valence up quark has a larger probability when x is large because the valence quark is more likely to contribute a large fraction of the proton’s momentum. Whereas, in Figure 6.3b), the sea up (or anti–up) quarks usually only contribute a small fraction of the total momentum and mainly at low x .

To calculate the cross-section of a process, all of the individual sub-processes from each partons’ contribution, $i, j = \{u, d, s, c, b, t\}$, are summed. In most cases, the heavier quarks, c, b, t , are not included in the calculation because of their negligible contribution to the total momentum fraction of the proton. The individual sub-process, $a + b \rightarrow X$, is approximated through the expansion [88]:

$$\sigma = \sum_{i,j} \int dx_a dx_b f_i(x_a; \mu_F^2) f_j(x_b; \mu_F^2) [\hat{\sigma}_0 + \alpha_S(\mu_R^2) \hat{\sigma}_1 + \alpha_S^2(\mu_R^2) \hat{\sigma}_2 \cdots]_{ab \rightarrow X}. \quad (6.2)$$

Here, μ_F is the factorization scale – which sets when the hard interactions of the process can be factorized from the non-perturbative dynamics – and μ_R is the normalization scale – which determines the strength of the strong coupling constant. For Drell–Yan processes, both scales are often set to the dilepton resonance mass, $\mu_F = \mu_R = m^Z$, because this kinematic region is where W/Z are predominately produced. The number of terms in the expansion, $\alpha_S^i \hat{\sigma}_i$, is the order of perturbative QCD used and sets the level of precision of the approximation. Here, $\hat{\sigma}_0$ is the Leading Order (LO), $\alpha_S \hat{\sigma}_1$ is Next to Leading Order (NLO), $\alpha_S^2 \hat{\sigma}_2$ is Next to Next to Leading Order (NNLO), and etc. The naming provides the number of vertices in the production of the W boson, where LO there is only the electroweak $q\bar{q}'W$ vertex, as seen in Figure 6.4a). Each additional order includes a QCD vertex and an additional factor of α_S , which can be seen in Figure 6.4b)-c) showing initial state radiation (ISR) vertices added to the process. Each additional order added to the calculation improves the precision at the cost of computation time and complexity as each introduces more processes: LO is one Feynman diagram, NLO is four additional diagrams, and NNLO is 62 additional diagrams. These additional diagrams can come from ISR vertices, loops in diagrams, or even new processes that could not exist at lower orders. How the different orders of diagrams affect the Drell–Yan process will be discussed in Section 6.3.

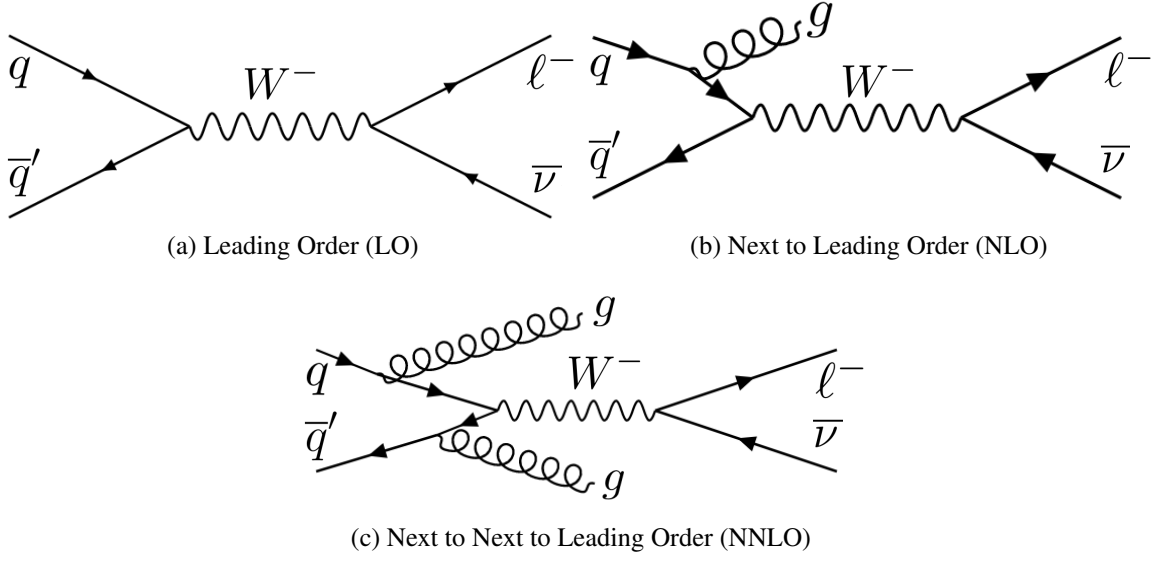


Figure 6.4: Examples of a) LO, b) NLO, and c) NNLO Feynman diagrams for the Drell–Yan process. The order is determined by the number of QCD vertices. With increased order, the number of Feynman diagrams increases.

6.1 Drell–Yan Angular Coefficients

The total cross-section provides the probability of how often a process occurs, such as the Drell–Yan process. These processes can be divided into more discrete parts by measuring the probability that the process occurs as a function of the process kinematics, known as a differential cross-section. For the Drell–Yan process, a differential cross-section can be measured as a function of p_T^V , m^V , y^V , $\cos \theta$, and ϕ , where p_T^V , m^V , and y^V are the transverse momentum, mass, and rapidity of a massive vector boson $V = \{W, Z\}$, respectively. Rapidity is defined as $y = \ln [(E + p_z)/(E - p_z)]$, where E and p_z are the energy and longitudinal momentum of the vector boson, respectively. The angles $\cos \theta$ and ϕ are the decay angles of the leptons, relative to the vector boson.

Differential cross-sections as a function of decay angles can be decomposed into angular polynomials, where the number of polynomials required depends on the spin of the intermediate massive boson. The number of required polynomials is dependent on the allowed quantum numbers and is discussed in more detail in Appendix B. For a spin 0 boson

– such as the Higgs boson – there is only one polynomial, spin 1 bosons – such as the W or Z bosons – require nine, and a hypothetical spin 2 boson would require 25.

The Drell–Yan process needing nine angular polynomials arises from having two spin 1/2 particles producing a spin 1 particle and decaying into two spin 1/2 particles. The different helicity combinations of the quarks, W boson, and the decay leptons produces nine helicity dependent cross-sections that contribute to the differential cross-section. These helicity dependent cross-sections are coupled to the nine angular polynomials and are summed over to produce the differential cross-section [4]:

$$\frac{d^5\sigma}{dp_T^W dm^W dy^W d\cos\theta d\phi} = \frac{3}{16\pi} \sum_{\alpha \in M} g_\alpha(\theta, \phi) \frac{d\sigma^\alpha}{dp_T^W dm^W dy^W}. \quad (6.3)$$

Here, $M = \{U+L, L, T, I, P, A, 7, 8, 9\}$ is the set of nine spin dependent cross-sections, $d\sigma^\alpha$, and g_α are the second order angular polynomials. Examples of $d\sigma^\alpha$ are $d\sigma^L$, $d\sigma^T$, and $d\sigma^I$ which are the longitudinal–polarized, transverse–polarized, and the longitudinal–transverse interference cross-sections, respectively. These cross-sections are related to the vector, v_f , and axial–vector, a_f , couplings between the W boson and the fermions [4], where $f = \{\ell, q\}$:

$$\begin{aligned} \sigma^{U+L, L, T, I} &\sim (v_\ell^2 + a_\ell^2) (v_q^2 + a_q^2), \\ \sigma^{P, A} &\sim v_\ell a_\ell v_q a_q, \\ \sigma^{7, 8} &\sim (v_\ell^2 + a_\ell^2) (v_q a_q), \\ \sigma^9 &\sim (v_\ell a_\ell) (v_q^2 + a_q^2). \end{aligned} \quad (6.4)$$

For the W boson, the vector and axial vector couplings to fermions are the third component of the weak isospin $T^3 = \pm 1/2$, as described in Section 2.2.2. The polynomials, $g_\alpha(\theta, \phi)$

found in Figure 6.5, are [4]:

$$\begin{aligned}
g_{U+L} &= 1 + \cos^2 \theta; & g_L &= 1 - 3 \cos^2 \theta; & g_T &= 2 \sin^2 \theta \cos 2\phi; \\
g_I &= 2\sqrt{2} \sin 2\theta \cos \phi; & g_P &= 2 \cos \theta; & g_A &= 4\sqrt{2} \sin \theta \cos \phi; \\
g_7 &= 2 \sin^2 \theta \sin 2\phi; & g_8 &= 2\sqrt{2} \sin 2\theta \sin \phi; & g_9 &= 4\sqrt{2} \sin \theta \sin \phi.
\end{aligned} \tag{6.5}$$

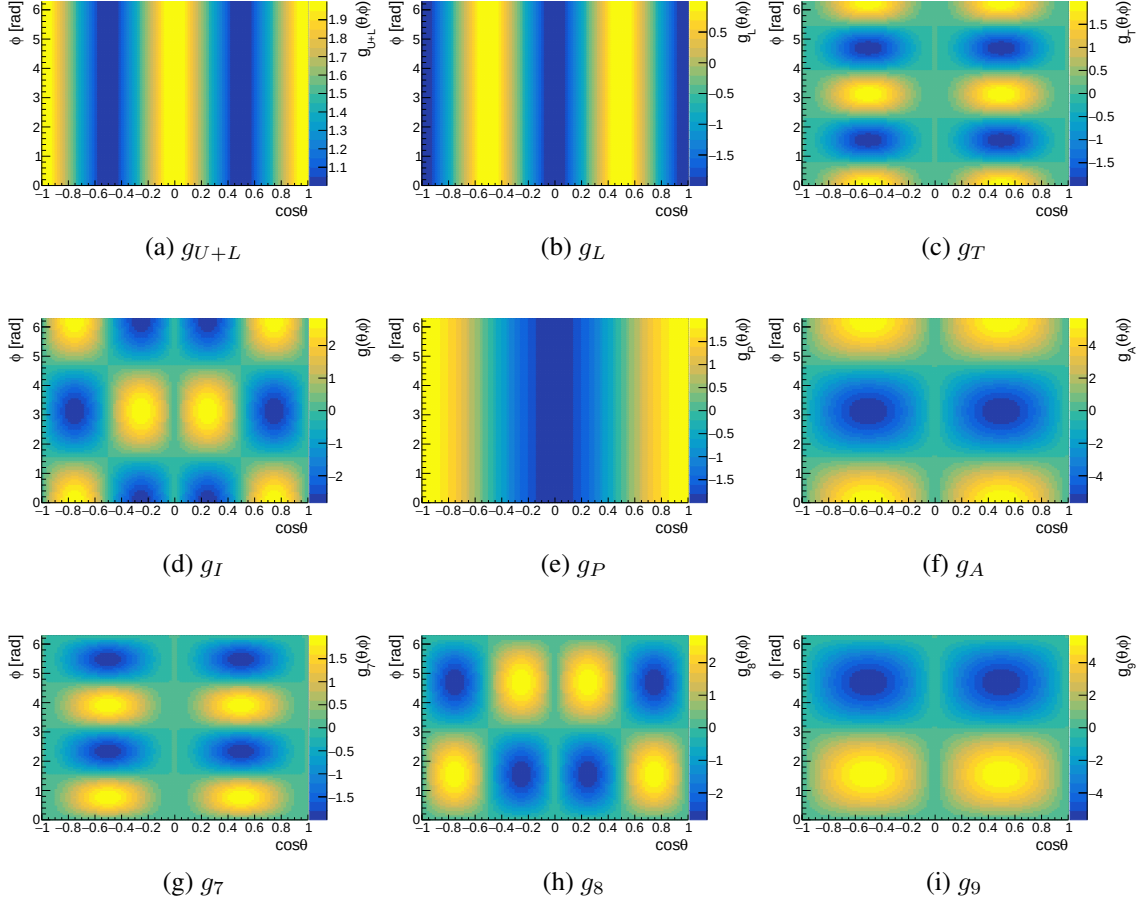


Figure 6.5: The second order angular polynomials associated with each helicity dependent cross-section, plotted as a function of $\cos \theta$ and ϕ .

With unpolarized production of particles, only the unpolarized cross-section can be directly measured. It is then beneficial to put Equation 6.3 in terms of the unpolarized cross-section, $d\sigma^{U+L}$, by factoring it out. This allows for the differential cross-section to be written in terms of dimensionless ratios of helicity dependent cross-sections that act as

angular coefficients:

$$\begin{aligned}
A_0 &= 2d\sigma^L/d\sigma^{U+L}; & A_1 &= 2\sqrt{2}d\sigma^I/d\sigma^{U+L}; & A_2 &= 4d\sigma^T/d\sigma^{U+L}; \\
A_3 &= 4\sqrt{2}d\sigma^A/d\sigma^{U+L}; & A_4 &= 2d\sigma^P/d\sigma^{U+L}; & A_5 &= 2d\sigma^7/d\sigma^{U+L}; \\
A_6 &= 2\sqrt{2}d\sigma^8/d\sigma^{U+L}; & A_7 &= 4\sqrt{2}d\sigma^9/d\sigma^{U+L}.
\end{aligned} \tag{6.6}$$

These are referred to as the Drell–Yan angular coefficients, denoted by A_i , and are coupled to second order polynomials, P_i listed in Table 6.1. The polynomials, P_i , are the same as the g_α polynomials except the normalization factors were absorbed by the Drell–Yan angular coefficients. Expressing the differential cross-section in terms of the Drell–Yan angular coefficients is then:

$$\frac{d^5\sigma}{dp_T^W dm^W dy^W d\cos\theta d\phi} = \frac{3}{16\pi} \frac{d^3\sigma^{U+L}}{dp_T^W dm^W dy^W} \left\{ (1 + \cos^2\theta) + \sum_{i=0}^7 A_i(p_T^W, y^W, m^W) P_i(\cos\theta, \phi) \right\}. \tag{6.7}$$

The Drell–Yan angular coefficients, A_i , are dependent only on the W boson kinematics and encode the information of how the W boson is produced, while the polynomials, P_i , depend only on the lepton kinematics. These polynomials are related to the spherical harmonics, Y_m^ℓ , as described in Table 6.1. Equation 6.7 is correct to all orders of perturbative QCD but not all A_i contribute at each order. A summary of when each coefficient starts contributing can be found in Table 6.1.

By measuring the Drell–Yan angular coefficients, this provides direct access to the vector and axial–vector couplings, which are free parameters of the EW interactions.

$\mathbf{A}_i$	$\mathbf{P}_i$	$\mathbf{Y}_\ell^m(\theta, \phi)$	Coupling	Non-Zero
A_0	$1/2(1 - 3\cos^2\theta)$	Y_2^0	$(v_\ell^2 + a_\ell^2)(v_q^2 + a_q^2)$	$\mathcal{O}(\alpha_S^1)$
A_1	$\sin 2\theta \cos \phi$	$(Y_2^{-1} - Y_2^1)$		$\mathcal{O}(\alpha_S^1)$
A_2	$1/2 \sin^2\theta \cos 2\phi$	$(Y_2^{-2} + Y_2^2)$		$\mathcal{O}(\alpha_S^1)$
A_3	$\sin \theta \cos \phi$	$(Y_1^{-1} - Y_1^1)$	$v_\ell a_\ell v_q a_q$	$\mathcal{O}(\alpha_S^1)$
A_4	$\cos \theta$	Y_1^0		$\mathcal{O}(\alpha_S^0)$
A_5	$\sin^2\theta \sin 2\phi$	$(Y_2^{-2} - Y_2^2)$	$(v_\ell^2 + a_\ell^2)(v_q a_q)$	$\mathcal{O}(\alpha_S^2)$
A_6	$\sin 2\theta \sin \phi$	$(Y_2^{-1} + Y_2^1)$		$\mathcal{O}(\alpha_S^2)$
A_7	$\sin \theta \sin \phi$	$(Y_1^{-1} + Y_1^1)$	$(v_\ell a_\ell)(v_q^2 + a_q^2)$	$\mathcal{O}(\alpha_S^2)$

Table 6.1: Summary of Drell–Yan angular coefficients, polynomials coupled to each coefficient, how the polynomials relate to the spherical harmonics, the vector and axial–vector couplings – which for the W boson is equal to the third component of the weak isospin $T^3 = \pm 1/2$ –, and the order of α_S when they become non-zero.

6.2 Collins–Soper Frame

The Drell–Yan angular coefficients depend on the W boson kinematics, so they will change depending on the choice of reference frame. For this analysis, the Collins–Soper (CS) frame [89] is chosen to make comparisons to previous measured values of Drell–Yan angular coefficients for the Z boson [5, 6] easier. A comparison of the CS frame and the laboratory frame can be seen in Figure 6.6. The CS frame is defined using the rest frame of the W boson and the momentum of the protons in the W boson rest frame. First, the z -axis is determined by taking the momentum of the protons to find a bisecting angle, α , as seen in Figure 6.6, and the sign is chosen to be in the same direction of positive longitudinal W boson momentum in the lab frame. Then, the y -axis is taken to be normal to the proton–proton plane and the x -axis is chosen to be orthogonal to the yz -plane to have a right-handed Cartesian coordinate system.

The direction of the leptons, with respect to the W boson, can be expressed using two angles, θ_{CS} and ϕ_{CS} . The first angle, θ_{CS} , is defined as the angle between the negatively charged lepton and the z -axis. For W^+ events, the neutrino is utilized because there is no negatively charged lepton. Using the kinematics of the leptons in the laboratory frame, it is

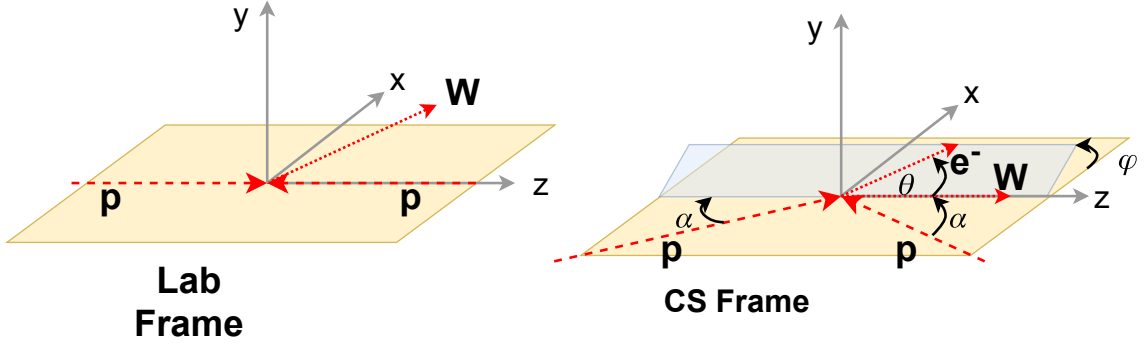


Figure 6.6: On the left is the lab frame, where protons come from opposite directions producing a W in an arbitrary direction. The Collins–Soper (CS) frame [89] is a rest frame of the W boson, which has the z -axis in the direction of the W boson in the lab frame producing a bisecting angle, α , between the protons. The y -axis is normal to the proton–proton plane and the x -axis is chosen to complete a right-handed Cartesian coordinate system. The CS angles are defined to have the polar angle θ_{CS} as the angle between the “negative” lepton and the z -axis while the ϕ_{CS} as the azimuthal angle between the xz -plane and the lepton plane. For W^+ events, the “negative” lepton is the neutrino.

possible to calculate $\cos \theta_{\text{CS}}$ [90] with:

$$\cos \theta_{\text{CS}} = \frac{p_z^{\ell\nu}}{|p_z^{\ell\nu}|} \frac{2}{m^W \sqrt{m^{W2} + p_T^{\ell\nu 2}}} (P_1^+ P_2^- - P_1^- P_2^+), \quad (6.8)$$

where

$$P_i^\pm = \frac{1}{\sqrt{2}} (E_i \pm p_{z,i}). \quad (6.9)$$

For Equation 6.8, $p_z^{\ell\nu}$ is the longitudinal momentum of the dilepton system and m^W is the mass of the W . In Equation 6.9, E_i and $p_{z,i}$ are the energy and longitudinal momentum of the lepton, respectively, and the index i is 1 for the lepton and 2 for the anti-lepton.

To calculate ϕ_{CS} the lepton momentum, p , must be found in the CS frame, which can be expressed as [4]:

$$\begin{aligned} p_x^{\text{CS}} &= \frac{1}{2} \frac{m^W}{m^{W2} + p_T^2} (2p_x - p_T^{W2}), \\ p_y^{\text{CS}} &= p_y, \\ p_z^{\text{CS}} &= \pm \frac{m^W}{2} \sqrt{1 - \frac{(p_x^{\text{CS}})^2 + (p_y^{\text{CS}})^2}{m^{W2}/4}}. \end{aligned} \quad (6.10)$$

With the transverse components, the azimuthal angle can be calculated using:

$$\phi_{\text{CS}} = \arctan p_y^{\text{CS}} / p_x^{\text{CS}}. \quad (6.11)$$

A different way to express Equation 6.8 is by calculating θ_{CS} using the lepton momentum in the CS frame using:

$$\theta_{\text{CS}} = \arctan \sqrt{(p_x^{\text{CS}})^2 + (p_y^{\text{CS}})^2} / p_z^{\text{CS}}. \quad (6.12)$$

There is an ambiguity in the sign of p_z^{CS} because of the unmeasured longitudinal momentum of the neutrino. This ambiguity also exists in Equation 6.8 since the neutrino longitudinal momentum is used to find $P_i^\pm$ and $p_z^{\ell\nu}$. This ambiguity is discussed in much further detail in Section 8.4 and Appendix F.

6.3 Coefficients in CS Frame

Before measuring the Drell–Yan angular coefficients, it is important to understand their expected behaviour as a function of p_{T}^W in the CS frame. Part of understanding this behaviour comes from knowing when the different orders of QCD start contributing to the production of the W boson. As p_{T}^W approaches zero, only $q\bar{q}$ annihilation Feynman diagrams with no QCD vertices contribute, as seen in Figure 6.7a). In this case, A_4 is at its maximum as p_{T}^W approaches zero and is the only non-zero coefficient. At LO in QCD, $q\bar{q}$, gq , and $g\bar{q}$ Feynman diagrams contribute, as seen in Figure 6.7b)-d), respectively. These diagrams increase the transverse momentum of the W boson as there are now particles for the W boson to recoil away from, as a result, coefficients A_0 to A_4 now all contribute. At NLO in QCD, gg Feynman diagrams, as seen in Figure 6.7e), start to contribute as well as higher order versions of previously mentioned production mechanisms, resulting in all coefficients contributing.

The coefficients A_0 , A_2 , and A_3 all increase as a function of p_{T}^W , as seen in Figure 6.8,

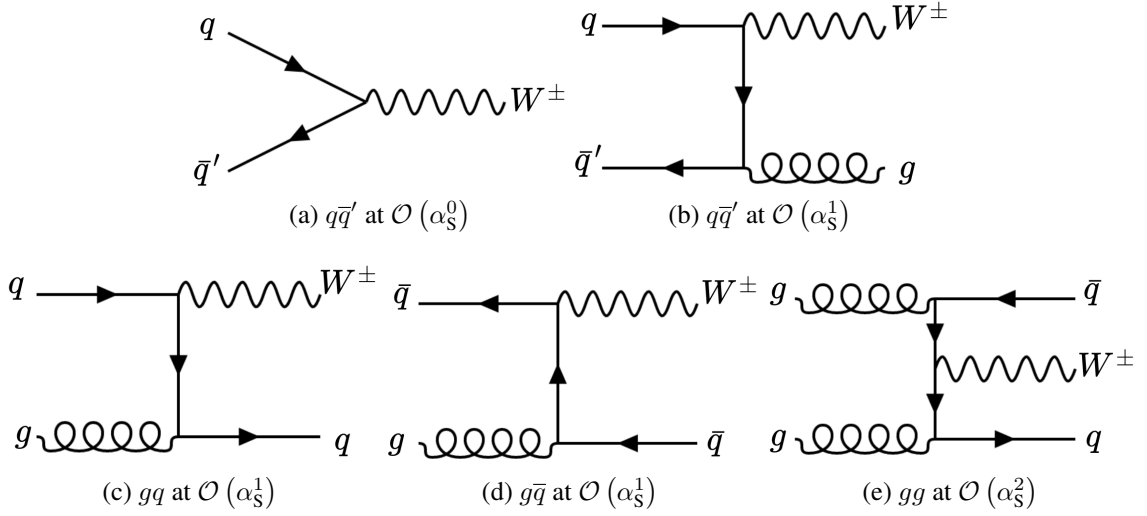


Figure 6.7: Examples of different Feynman diagrams for QCD production of W boson, which contribute to the Drell–Yan angular coefficients. a) quark–antiquark annihilation at $\mathcal{O}(\alpha_S^0)$, b) quark–antiquark annihilation at $\mathcal{O}(\alpha_S^1)$, c) Compton scattering of gluon and quark at $\mathcal{O}(\alpha_S^1)$, d) Compton scattering of gluon and antiquark at $\mathcal{O}(\alpha_S^1)$, and e) gluon–gluon interaction at $\mathcal{O}(\alpha_S^2)$.

where deviations from zero start around $p_T^W = 20 - 50$ GeV. The coefficient A_1 is different than the other coefficients that contribute at NLO, A_0 to A_4 , because it is the interference of the transverse–longitudinal cross-sections. Because of this, it is an order of magnitude smaller, as seen in Figure 6.8b), and it starts deviating from zero more quickly but it starts decreasing when $p_T^W > 50$ GeV. As mentioned earlier, A_4 is at its maximum at low p_T^W and decreases as other coefficients start to contribute, as seen in Figure 6.8. At NNLO and higher orders, gluon loop diagrams start contributing, allowing for non-zero values of A_5 to A_7 , at higher values of p_T^W but are orders of magnitude smaller than the other coefficients and are not shown.

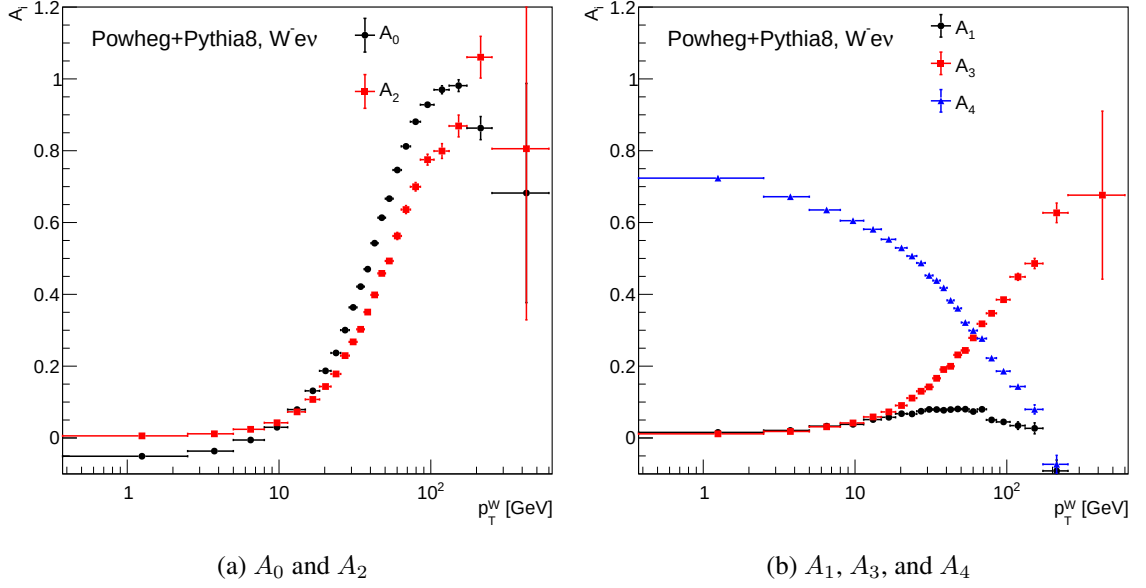


Figure 6.8: Drell–Yan angular coefficients A_0 to A_4 as a function of p_T^W for $W^- \rightarrow e^- \bar{\nu}$. Produced using POWHEG+PYTHIA8 at generator level with the same binning as the $\sqrt{s} = 8$ TeV Z Drell–Yan angular coefficients analysis [6].

6.4 Summary

This chapter provided an understanding of the theoretical background for the Drell–Yan process and the Drell–Yan angular coefficients. Armed with this knowledge, the next chapter will discuss the methodology to measure the angular coefficients.

CHAPTER 7

MEASUREMENT METHODOLOGY

This chapter will cover the measurement methodology used to extract the Drell–Yan angular coefficients, first introduced in Section 6.1, from the differential cross-section in Equation 7.1:

$$\frac{d^5\sigma}{dp_T^W dm^W dy^W d\cos\theta d\phi} = \frac{3}{16\pi} \frac{d^3\sigma^{U+L}}{dp_T^W dm^W dy^W} \left\{ 1 + \cos^2\theta + \sum_{i=0}^7 A_i(p_T^W, y^W, m^W) P_i(\cos\theta, \phi) \right\}. \quad (7.1)$$

The angular coefficients, A_i , are dependent on the transverse momentum, mass, and rapidity of the W boson. They are coupled to angular polynomials, P_i which depend on the kinematics of the charged lepton and neutrino that the W boson decays into. The methodology allows for the angular coefficients to be measured simultaneously as a function of p_T^W , m^W , and y^W , and will be discussed as such in general. However, for the measurements in this thesis, it is only performed as a function of p_T^W . It cannot be measured as a function of m^W because of the methodology used for the reconstruction of the W boson, discussed in more detail in Section 8.4. In the data set used in this thesis, which will be discussed later in Section 8.1, there is a limited number of statistics. Due to the smaller data set size there are not sufficient statistics in the data set used to measure the coefficients as a function of y^W , which is also affected by the reconstruction method of the W boson discussed in Section 8.4. Throughout this chapter, equations will either be left in general terms of a three-dimensional measurement or will be given in terms of p_T^W . The measurement method was used in the ATLAS measurement of the Z angular coefficients [6], which measured the angular coefficients as a two-dimensional measurement of p_T^Z and y^Z . With enough statistics this method is robust enough to measure the Drell–Yan angular coefficients as a one, two, or three-dimensional measurement.

One of the main motivations for measuring the Drell–Yan angular coefficients is to

improve the model used to describe W boson production at the LHC. The W boson is produced in its full phase space, which is when all events are selected regardless of the kinematics of the W boson and its decay products. It is useful to measure the coefficients in the full phase space. However, this is unachievable as the measurement is performed using events that are measured within the ATLAS detector; therefore, they have selection requirements applied or do not meet detection thresholds, which forms a reconstructed phase space. The full phase space can be recovered from the reconstructed phase space using a statistical method, called folding, allowing a likelihood (LH) fit to be performed and the angular coefficients extracted. This chapter will go through all the steps needed to achieve the extraction of the coefficients.

7.1 Moments Method

In the full phase space of the variables p_T^W , m^W , y^W , θ , and ϕ , extraction of each Drell–Yan angular coefficient can be performed individually using a technique called the moments method [91]. Similar to taking the first moment of a probability distribution function, the moment of the angular distributions can be measured. By taking the moment of the trigonometric polynomials coupled to the coefficients, each coefficient can be measured separately. This uses the fact that the trigonometric polynomials, $P_i(\theta, \phi)$ as listed in Table 6.1, are related to spherical harmonics and are, therefore, orthogonal to each other. The moment can be defined as:

$$\langle P_i(\theta, \phi) \rangle = \frac{\int d\sigma(p_T^W, m^W, y^W, \theta, \phi) P_i(\theta, \phi) d\cos\theta d\phi}{\int d\sigma(p_T^W, m^W, y^W, \theta, \phi) d\cos\theta d\phi}, \quad (7.2)$$

where, P_i is a polynomial and $d\sigma$ is the differential cross-section, as seen in Equation 7.1. In the denominator of Equation 7.2, the moment is normalized by the total unpolarized

cross-section, $d^3\sigma/dp_T^W dm^W dy^W$. For example, when $i = 4$ this results in:

$$\langle \cos \theta \rangle = \frac{1}{4} A_4. \quad (7.3)$$

Similarly, the rest of the angular coefficients from Equation 7.1 can be obtained by taking the moments of their respective polynomials, resulting in:

$$\begin{aligned} \langle \frac{1}{2} (1 - 3 \cos^2 \theta) \rangle &= \frac{3}{20} \left(A_0 - \frac{2}{3} \right); & \langle \sin 2\theta \cos \phi \rangle &= \frac{1}{5} A_1; \\ \langle \sin^2 \theta \cos 2\phi \rangle &= \frac{1}{10} A_2; & \langle \sin \theta \cos \phi \rangle &= \frac{1}{4} A_3; \\ \langle \cos \theta \rangle &= \frac{1}{4} A_4; & \langle \sin^2 \theta \sin 2\phi \rangle &= \frac{1}{5} A_5; \\ \langle \sin 2\theta \sin \phi \rangle &= \frac{1}{5} A_6; & \langle \sin \theta \sin \phi \rangle &= \frac{1}{4} A_7. \end{aligned} \quad (7.4)$$

Using this method allows measuring all eight coefficients individually in the full phase space of the MC. A more in-depth example of taking the moment to extract A_4 can be found in Appendix C.

7.2 Building Templates

The events measured with the ATLAS detector contain only a portion of the full phase, called the reconstructed phase space. This reconstructed phase space is comprised of reconstructed events that are in the acceptance range of the detector, pass the selection requirements, and meet proper data quality requirements. This means, at a minimum, that the reconstructed phase space should be a subset of the full phase space. However, this is not the case as events are not reconstructed perfectly with the ATLAS detector. One of the reasons for this is that not all energy of an event passes the threshold to be measured and energy is lost in the inactive portions of the detector. Another reason is that the number of events selected in data differs from what is expected from simulation, which results in different efficiencies that are compensated for, as described in Chapter 5. All of this culmi-

nates to the idea that the reconstructed phase space can be thought of as a distorted subset of the full phase space. To measure the coefficients, templates are made with MC to map from the reconstructed into the full phase space.

The first step is to construct templates of the full phase space using “truth” events from MC simulation. These truth events are simulated events of proton–proton interactions producing a W boson and decaying into a charged lepton and a neutrino. This simulation uses the current understanding of theoretical modelling, up to a certain precision, and does not contain any detector effects. With the truth level decay kinematics of the charged lepton and the neutrino, the W boson is reconstructed. This reconstructed W boson and leptons can be used to calculate $\cos \theta_{CS}$ and ϕ_{CS} , according to Equations 6.8 and 6.11, respectively. The $\cos \theta_{CS}$ and ϕ_{CS} values are used to fill 8×8 two-dimensional histograms, where $\cos \theta_{CS}$ ranges from $[-1, 1]$ and ϕ_{CS} ranges from $[0, 2\pi]$. These two-dimensional phase space histograms are created for each $p_T^{W, \text{Truth}}$ bin, listed in Equation 7.5. Examples of these phase space templates for $W^- \rightarrow e^- \bar{\nu}$ can be found in Figure 7.1, which shows the evolution of the angular distributions with respect to $p_T^{W, \text{Truth}}$.

$$\begin{aligned} p_T^W &= [0, 8, 17, 27, 40, 55, 75, 110, 150, 210, 600] \text{ GeV}, \\ p_T^W &= \{p_T^{W, \text{Truth}}, p_T^{l\nu}\}. \end{aligned} \tag{7.5}$$

The second step is to construct the same type of phase space templates using reconstructed events from MC simulation. The reconstructed events take the truth events and simulate the effects of secondary interactions within the detector, radiative energy loss, and interaction with the detector. This is to approximately measure how the events would appear when measured by the ATLAS detector. Selection requirements are applied to the simulated reconstructed events to model the events as closely as possible to what is expected with reconstructed events in data. These selection requirements and the reconstruction process affect the distributions compared to the same templates in the truth phase space. By applying selection requirements, events in the truth phase space are removed from the

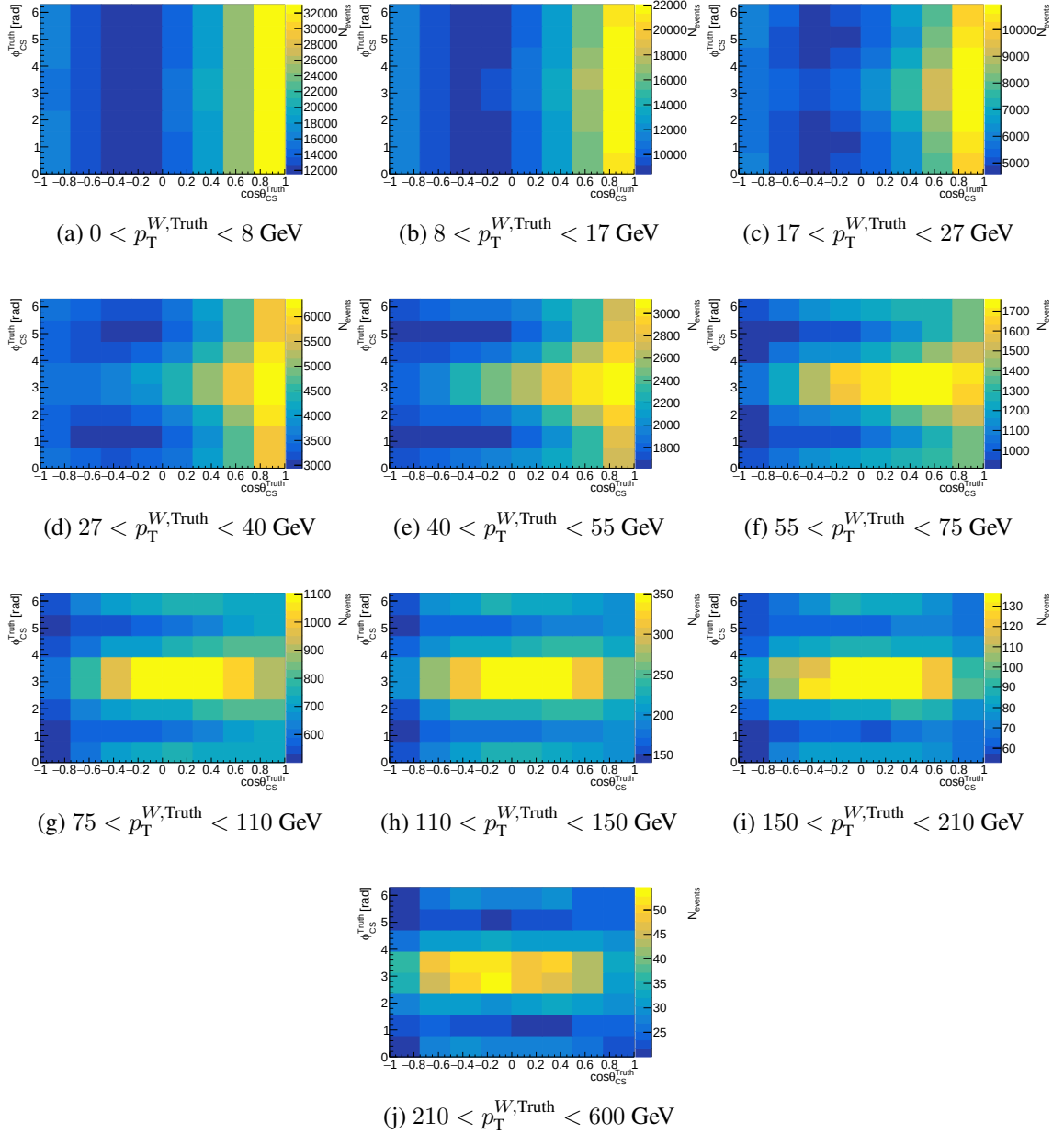


Figure 7.1: Examples of how the truth phase space templates that describe the angular distributions evolve as a function of $p_T^{W, \text{Truth}}$ for $W^- \rightarrow e^- \bar{\nu}$ events. Along the x-axis is $\cos \theta_{CS}$, along the y-axis is ϕ_{CS} , and along the z-axis is the number of events.

reconstructed phase space. Events that are near reconstruction or selection boundaries – for example, close to a bin boundary – can be reconstructed across these boundaries and change the overall distributions.

Similar to the truth events, the W boson is reconstructed to calculate $\cos \theta_{CS}$ and ϕ_{CS} to

fill two-dimensional histograms in the same p_T^W binning as the truth phase space templates listed in Equation 7.5. The reconstructed phase space templates are labeled with $p_T^{l\nu}$ to denote that it is the reconstructed transverse momentum of the lepton-neutrino system ($l\nu$). These reconstructed template histograms differ from the truth templates by expanding the $\cos\theta_{CS}$ range by two additional bins so the total range is from $[-1, 1.5]$. The $\cos\theta_{CS} > 1$ events are where the neutrino can not be reconstructed properly, as will be described in Section 8.4, and are purposefully separated from events where the neutrino can be properly reconstructed as to not distort the angular distributions. The first bin $[1, 1.25]$ is intentionally left empty to provide separation, while events where the neutrino cannot be properly reconstructed are stored in the bin $[1.25, 1.5]$. Examples of these phase space templates for $W^- \rightarrow e^- \bar{\nu}$ can be found in Figure 7.2, which shows the evolution of the angular distributions with respect to $p_T^{l\nu}$.

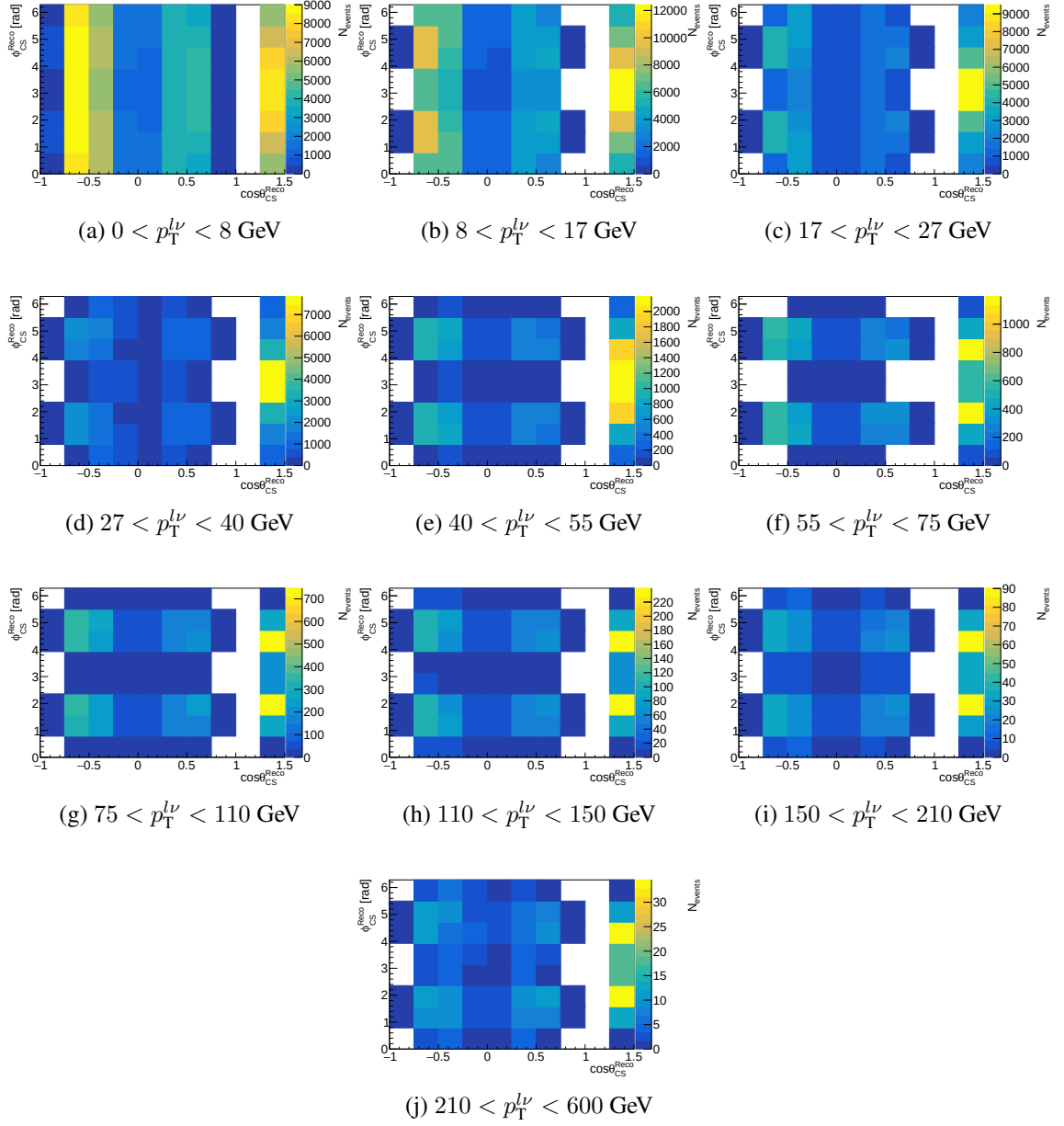


Figure 7.2: Examples of how the reconstructed phase space templates that describe the angular distributions evolve as a function of reconstructed p_T^{ν} for $W^- \rightarrow e^- \bar{\nu}$ events. Along the x-axis is $\cos \theta_{CS}$, along the y-axis is ϕ_{CS} , and along the z-axis is the number of events. Events with $\cos \theta_{CS} > 1$ are from events where the neutrino cannot be reconstructed and are put in the additional $\cos \theta_{CS}$ bin to not distort the shape of the other bins.

7.3 Template Folding

Most analyses in particle physics use a technique called unfolding, where the reconstructed phase space is unfolded to recover the truth phase space [92, 93]. Instead, the inverse of this technique, called folding, is used in this analysis. Folding is where the truth phase space is folded into the reconstructed phase space through a LH fit, which will be discussed more in Section 7.4. Folding and unfolding should give the same result given a sufficient understanding of the relation between the truth and reconstructed phase spaces, as they are just different statistical methods to relate the two. This section will provide a brief overview of folding and more details to the specifics of this analysis can be found in Appendix D.

A set of truth observables – t – that are distributed in a function – $f(t)$ – can be folded using a probability distribution – $p(r|t)$ – into a set of reconstructed observables – r – distributed in a function – $g(r)$. There is no analytical form for the probability distribution $p(r|t)$ and must be derived through simulation by finding the probability of measuring the reconstructed observables given the truth observables. For this analysis, the truth and reconstructed observables, $t = (\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}, p_{\text{T}}^{W, \text{Truth}})$ and $r = (\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}}, p_{\text{T}}^{l\nu})$, respectively, are found as discussed in Section 7.2. The resulting binned probability distribution, described in more detail in Appendix D, is:

$$T_{ij} = \sum_{\text{event} \in \Delta_j} \frac{P_i(\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}}) w_{\text{event}}(r, t)}{\sigma_j \left\{ P_8(\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}) + \sum_{i=0}^7 A_{ij}^{\text{ref}} P_i(\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}) \right\}}, \quad (7.6)$$

where $\Delta_j = (\Delta p_{\text{T}}^{W, \text{Truth}})_j, (\Delta p_{\text{T}}^{l\nu})_j, (\Delta \cos \theta_{\text{CS}}^{\text{Reco}})_j$, and $(\Delta \phi_{\text{CS}}^{\text{Reco}})_j$, where Δ is the width of the j th bin, and A_{ij}^{ref} is the i th reference coefficient in the j th bin from the moments method. In Equation 7.6, σ_j is the differential cross-section in the j th p_{T}^W bin and w_{event} is the product of event scale factors used to compensate for inefficiencies in the detector. To model effects like how $\cos \theta_{\text{CS}}^{\text{Reco}}$ and $\phi_{\text{CS}}^{\text{Reco}}$ change as function of $p_{\text{T}}^{l\nu}$ in the reconstructed phase space the polynomial templates, $P_i(\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$, are constructed as a function

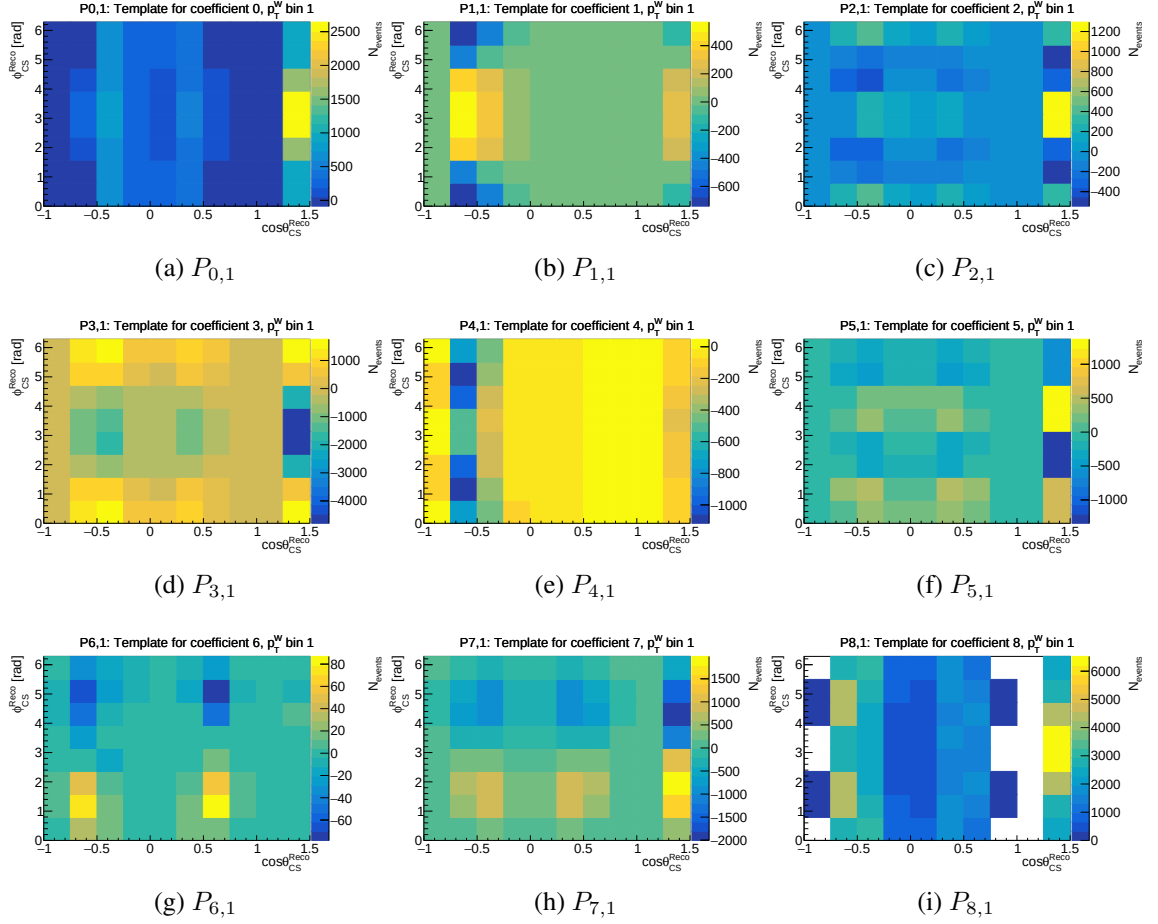


Figure 7.3: Polynomial templates $P_{i,1}$ for the i th coefficient in the p_T^W bin 1 before folding, which ranges from $8 < p_T^{\nu} < 17$ GeV for $W^- \rightarrow e^- \bar{\nu}$ events. Along the x-axis is reconstructed $\cos\theta_{\text{CS}}$, along the y-axis is reconstructed ϕ_{CS} , and along the z-axis is the number of events. Each event is weighted by the i th polynomial which is why events are allowed to be negative.

p_T^{ν} . In Equation 7.6 this means $P_i(\cos\theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$ is replaced by $P_{ij}(\cos\theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$ to account for it being constructed in the j th p_T^{ν} bin. To demonstrate this examples of each polynomial template weighted by w_{event} before folding, which is the numerator in Equation 7.6 will be shown. These can be seen for $8 < p_T^{\nu} < 17$ GeV, $P_{i,2}$, and how $P_{3,j}$ evolves with respect to p_T^{ν} in select bins in Figures 7.3 and 7.4, respectively.

One advantage that folding has when compared to unfolding is that the statistical uncertainties are easier to obtain. When performing unfolding, a covariance matrix is applied to the reconstructed data to get the proper statistical uncertainty, since it does not naturally

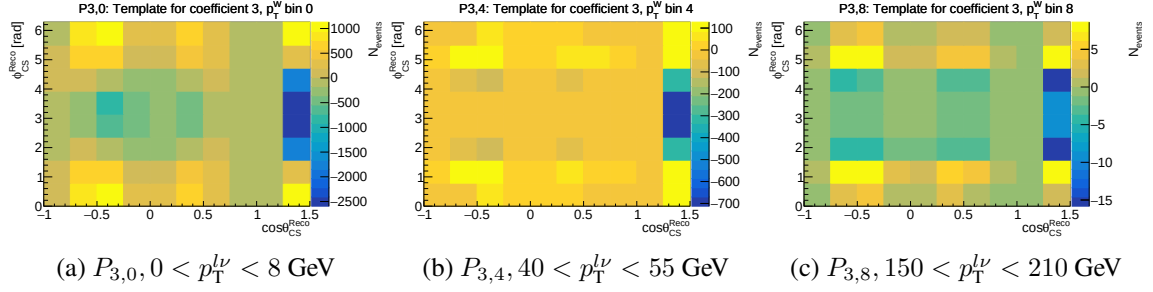


Figure 7.4: Polynomial templates of $P_{3,j}$ for A_3 in the j th p_T^W bin before folding for $W^- \rightarrow e^- \bar{\nu}$ events to illustrate how the polynomial changes with respect to $p_T^{l\nu}$. Examples are shown for low, middle, and high $p_T^{l\nu}$. Along the x-axis is reconstructed $\cos \theta_{CS}$, along the y-axis is reconstructed ϕ_{CS} , and along the z-axis is the number of events. Each event is weighted by the i th polynomial which is why events are allowed to be negative.

follow Poisson statistics. This is not needed for folding, as it follows Poisson statistics, because the proper statistical uncertainty is built into the LH and is easier to determine. Another major advantage is that systematic uncertainties are easy to propagate since they produce varied versions of the templates in the LH function, as described in Section 10.2.

7.4 Likelihood Method

A LH fit is performed to find the most likely value of the coefficients given the templates. The LH is built from both the nominal templates and systematically varied templates. The nominal templates are from the MC sample and is what is expected given the current physics modelling of W boson production at the LHC. The alternative, systematically varied templates are built based on the uncertainties of the physics modelling that is used when constructing the nominal templates. For example, the normalization scale, in Equation 6.2, is varied within its uncertainty and the events with this new normalization scale are stored in systematically varied templates. These systematic variations can affect the shapes of the angular distributions, the number of expected events in each bin, or both the shape and the yield. These systematic variations are described by a set of nuisance parameters (NPs), $\theta = \{\beta, \gamma\}$, which are interpolated to best fit the data and are constrained by auxiliary probability density functions.

The first category, β , contains the NPs representing experimental and theoretical uncertainties. Each β^m in the set $\beta = \{\beta^1, \dots, \beta^M\}$ is constrained by unit Gaussian probability density function, $G(0|\beta^m, 1)$, and interpolate between nominal and varied templates. They are defined by having a nominal value of zero and $\beta^m = \pm 1$ corresponding to $\pm 1\sigma$ for the considered systematic uncertainty.

The second category, γ , contains the NPs that handle systematic uncertainties from the limited size of the MC samples. For each $(\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$ bin n there is a γ^n in the set $\gamma = \{\gamma^1, \dots, \gamma^{N_{\text{bins}}}\}$, where N_{bins} is the number of $(\cos \theta_{\text{CS}}, \phi_{\text{CS}})$ bins in each p_{T}^W bin. The number of bins, N_{bins} , is not constant for each p_{T}^W bin. Instead, the number of bins starts from $(8 \times 8) + (2 \times 8)$, where the first 64 bins are for events with a real neutrino solution and the remaining 16 bins are events that the neutrino cannot be reconstructed properly, as will be discussed in Section 8.4. Bins with zero expected or observed events are pruned to improve fit stability, resulting in N_{bins} ranging between 40 – 72 bins for an individual p_{T}^W bin. γ^n has a nominal value of one and normalizes the expected events in bin n of the templates. They are constrained by Poisson probability density functions, $P(N_{\text{eff}}^n | \gamma^n N_{\text{eff}}^n)$, where N_{eff}^n is the effective number of MC events in bin n . The effective number of events is used due to the non-uniform event weights applied to properly scale the MC to account for effects, such as detector inefficiencies, explained in Section 8.3.

The nominal and systematic varied templates are used to find the expected number of events in bin n and can be written as:

$$N_{\text{exp}}^n(A, \sigma, \theta) = \left\{ \sum_{j=0}^{N_{\text{bins}}^W} \sigma_j \times L \left[T_{8,j}^n(\beta) + \sum_{i=0}^7 A_{ij} \times T_{ij}^n(\beta) \right] \right\} \times \gamma^n + \sum_B^{\text{bkgds}} T_B^n(\beta). \quad (7.7)$$

Here:

- A : Set of all A_{ij}
- A_{ij} : Coefficient parameter A_i for observation bin j

- σ : Set of all σ_j
- σ_j : Signal cross-section parameter
- θ : Set of all NPs
- N_{bins}^W : Number of p_T^W bins
- L : Integrated luminosity
- T_{ij} : Polynomial template P_i for observation bin j
- β : Set of all Gaussian-constrained NPs
- γ^n : Poisson-constrained NP
- T_B : Background estimation templates.

The number of p_T^W bins are summed over to account for the contribution from the events migrating between bins. This migration is when an event would have $p_T^{W,\text{Truth}}$ and $p_T^{l\nu}$ not be in the same bin and may end up in a neighbouring bin. A more in-depth discussion of migration is in Section 10.1.

The LH is the product of Poisson probabilities across all N_{bins} and of auxiliary constraints for each NP β_m in the set M :

$$\mathcal{L}(A, \sigma, \theta | N_{\text{obs}}) = \prod_n^{N_{\text{bins}}} \{ P(N_{\text{obs}} | N_{\text{exp}}^n(A, \sigma, \theta)) P(N_{\text{eff}}^n | \gamma^n N_{\text{eff}}^n) \} \times \prod_m^M G(0 | \beta^m, 1). \quad (7.8)$$

To simplify computation, the maximum of the LH is not found and instead the negative logarithm of the LH is minimized. This is done because, instead of computing the product of the probability distributions, the sum of negative logarithms is easier to compute and still allows for extraction of the coefficients.

CHAPTER 8

DATA, SIMULATION, AND EVENT RECONSTRUCTION

This chapter will focus on the conditions for the measured data set and the simulated Monte Carlo (MC) sets used in this analysis. The MC sets are used to estimate what is expected to be seen in data by modelling the expected distributions for signal and background events. This chapter will also focus heavily on how information about the neutrino can be obtained, despite the neutrino not being measurable with the ATLAS detector, allowing for full $W \rightarrow \ell \nu$ event reconstruction. The selection criteria used throughout this analysis, and a comparison of the kinematic distributions for the expected and measured distributions will also be discussed.

8.1 Data Set

During the Run 2 proton–proton collision running period from 2015-18 with the ATLAS detector, a total integrated luminosity of 146.9 fb^{-1} was recorded. Of this recorded data an integrated luminosity of $140.1 \pm 1.2 \text{ fb}^{-1}$ [11] is usable for physics analysis because it passes data quality requirements. The data sets from this running period were taken with an average of the mean number of interactions per bunch crossing, $\langle \mu \rangle = 33.7$, at $\sqrt{s} = 13 \text{ TeV}$. For the analysis in this thesis, a special data set is used that was taken during Run 2, where this mean was reduced to $\langle \mu \rangle \sim 2$ [11]. To reduce the number of interactions in the detector, the beams were separated in the transverse plane and continuously adjusted to maintain an average $\langle \mu \rangle \sim 2$. This low- $\langle \mu \rangle$ data set was taken with the physics purpose of precision measurements, such as this analysis and other W/Z boson measurements [94].

The low- $\langle \mu \rangle$ data set at $\sqrt{s} = 13 \text{ TeV}$ was taken over two periods, during Fall 2017 with integrated luminosity $147.1 \pm 1.7 \text{ pb}^{-1}$ and in Summer 2018 with integrated luminosity $191.1 \pm 2.1 \text{ pb}^{-1}$, resulting in a combined data set of $338.1 \pm 3.1 \text{ pb}^{-1}$ [11]. This

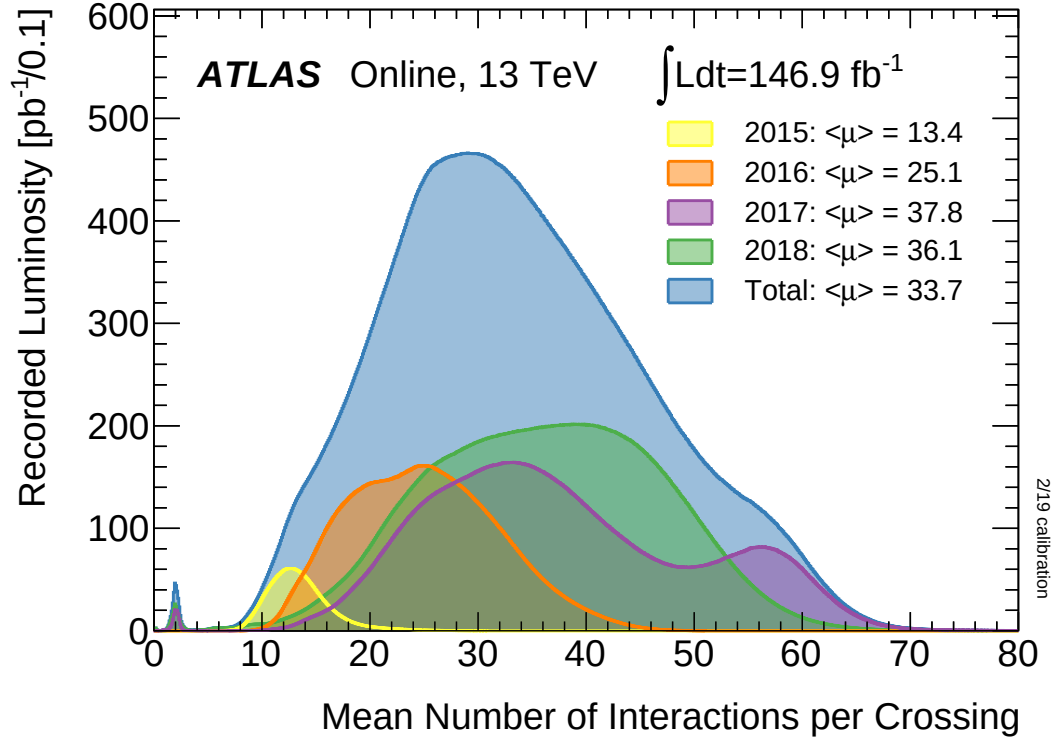


Figure 8.1: The integrated luminosity vs. mean number of interactions per bunch crossing taken with the ATLAS detector during Run 2 at $\sqrt{s} = 13$ TeV from [45]. Note, the small peak on the left side corresponds to the integrated luminosity of the low- $\langle\mu\rangle$ data set used in this analysis.

low- $\langle\mu\rangle$ data set makes up less than 1% of the total proton–proton collision data taken during Run 2 by the ATLAS detector and corresponds to the relatively small peak on the bottom–left in Figure 8.1.

With fewer interactions per bunch crossing in the low- $\langle\mu\rangle$ data set, some run conditions were different compared to the rest of Run 2 with high- $\langle\mu\rangle$ data set. The low- $\langle\mu\rangle$ run conditions allowed for lower energy thresholds for hardware and software triggers, which improves hadronic recoil reconstruction, as the average ΣE_T threshold is lower. In addition, there were less strict identification requirements and no isolation requirements for electrons and muons, which allows for more recorded events, and a lower p_T^l threshold. A comparison of the requirements between high- and low- $\langle\mu\rangle$ for Level-1 (L1) hardware and the software high level trigger (HLT) for a single electron or muon can be seen in Table 8.1.

	Electron		Muon	
	L1	HLT	L1	HLT
high- $\langle\mu\rangle$	$E_T > 26 \text{ GeV}$	$E_T > 26 \text{ GeV}$	$p_T > 14 \text{ GeV}$	$p_T > 24 \text{ GeV}$
		<i>Tight</i> identification		<i>Medium</i> identification
		variable loose isolation		variable medium isolation
low- $\langle\mu\rangle$	$E_T > 13 \text{ GeV}$	$E_T > 15 \text{ GeV}$	$p_T > 10 \text{ GeV}$	$p_T > 14 \text{ GeV}$
		<i>VeryLoose</i> identification		no identification
	no isolation			

Table 8.1: A comparison between the high- and low- $\langle\mu\rangle$ requirements for an electron or muon to pass the L1 hardware and HLT software triggers, which shows the lower thresholds used for the low- $\langle\mu\rangle$ run conditions.

8.2 MC Sets

Signal and background processes that are expected to be measured in data are simulated creating MC samples, where the detailed list of the samples used in this analysis are listed in Appendix E. The simulations are made by the ATLAS Collaboration to match the special run conditions of the data taking and are tuned to model how the particles interact with the ATLAS detector. For this analysis, the simulation of particle production from proton–proton interactions is primarily performed using POWHEG [95, 96, 97, 98], while how the particles shower in the ATLAS detector is performed using PYTHIA8 [99]. A second MC generator, SHERPA [100, 101], is also used to simulate some of the expected background processes. For most signal and significant backgrounds, an equivalent of $\sim 3 \text{ fb}^{-1}$ integrated luminosity of data is simulated. This is approximately 10 times the amount of low- $\langle\mu\rangle$ $\sqrt{s} = 13 \text{ TeV}$ data taken. The events are then reweighted to match the integrated luminosity of the low- $\langle\mu\rangle$ data set to get the number of expected events. By simulating more than the expected number of events, the distributions are less prone to statistical fluctuations. The events are also reweighted with scale factors to account for the difference in measured efficiency for MC and data, first introduced in Section 4.1. This is done by taking

the product of the individual event weights,

$$w_{\text{event}} = \prod_{i \in \mathcal{W}} w_i, \quad (8.1)$$

where $\mathcal{W} = \{w_{\text{MC,gen}}, \text{SF}^{\text{reco}}, \text{SF}^{\text{id}}, \text{SF}^{\text{iso}}, \text{SF}^{\text{trig}}, \text{SF}^{\text{TTPVA}}\}$ is the set of event weights and SFs applied to correct MC to better match data. In order of the list, the set includes the MC generator event weight, the reconstruction SF, identification SF, isolation SF, trigger SF, and track-to-vertex association SF (muons only).

8.2.1 POWHEG+PYTHIA8

The main signal – $W \rightarrow \ell\nu$ – event samples, single boson background, $t\bar{t}$, and single-top production – Wt , t-channel, s-channel – backgrounds use the POWHEG event generator [95, 96, 97, 98] at NLO in QCD using the CT10 [102] PDF set. The events are interfaced to PYTHIA8 [99] to simulate the particles showering in the ATLAS detector, which is tuned to specific requirements that were first introduced in the Z boson p_T distribution measurement at $\sqrt{s} = 7$ TeV [103]. To properly simulate final state QED radiation, POWHEG+PYTHIA8 is interfaced to Photos++ [104]. To improve accuracy the samples are renormalized to NNLO matrix element calculations using the DYTURBO program [105], which is an optimized version of DYNNLO [106, 107] using the MMHT2014nnlo PDF set [108]. A list of some of the samples that are generated using POWHEG+PYTHIA8 are listed in Table 8.2, to provide examples of what is considered signal, single boson backgrounds, and top backgrounds. A more detailed list can be found in Appendix E, which has all sources of background included.

8.2.2 SHERPA

SHERPA [100, 101] is used to simulate various diboson backgrounds used in this analysis (VV , where $V = W, Z$), and there is at least one lepton in the final state. These sam-

Signal	Boson Bkgd.	Top Bkgd.
$W^+ \rightarrow e^+ \nu$	$Z \rightarrow e^+ e^-, W^+ \rightarrow \tau^+ \nu, \dots$	$t\bar{t}, Wt, \dots$
$W^- \rightarrow e^- \bar{\nu}$	$Z \rightarrow e^+ e^-, W^- \rightarrow \tau^- \bar{\nu}, \dots$	
$W^+ \rightarrow \mu^+ \nu$	$Z \rightarrow \mu^+ \mu^-, W^+ \rightarrow \tau^+ \nu, \dots$	
$W^- \rightarrow \mu^- \bar{\nu}$	$Z \rightarrow \mu^+ \mu^-, W^- \rightarrow \tau^- \bar{\nu}, \dots$	

Table 8.2: List of signal and dominant background processes that are generated using POWHEG+PYTHIA8. A more complete list can be found in Appendix E and contains more information about each process.

ples are generated using the NNPDF3.0NNLO [109] PDF set and are also renormalized to NNLO matrix element calculations from DYTURBO in a similar way as what is used for POWHEG+PYTHIA8. SHERPA merges $V + 0, 1, 2$ partons at NLO and $V + 3, 4$ partons at LO using the MEPS@NLO scheme [100].

8.3 Event Selection

Requirements on the kinematics of particles are used to remove regions of inefficiency in the detector, as efficiencies are often E_T and η dependent, as discussed in Section 4.1. Other requirements are usually applied to reduce the amount of background contamination so there is a higher purity of signal. A list of the final event selection criteria used throughout this analysis and the reason for them can be found in Table 8.3. Most selection criteria are only applied to the charged leptons in the event but there are also selection criteria on the missing transverse energy in the detector, E_T^{miss} , and the transverse mass of the W boson, m_T . Electrons and muons are not reconstructed using the same method, as discussed in Section 4.3, so there are some selection differences between them; most notably is that electrons have an additional selection on its pseudorapidity, $|\eta^l|$, to remove the overlap region between the EMB and EMEC.

Electrons	Muons	Reason
Single Lepton		Only events that look most like $W \rightarrow \ell\nu$.
Tight ID	Medium ID	Reduce background contamination.
$p_T^l > 25 \text{ GeV}$		Efficiently reconstructed leptons.
$E_T^{\text{miss}} > 25 \text{ GeV}$		Reduce background contamination.
$m_T > 25 \text{ GeV}$		Reduce background contamination.
$ \eta^l < 2.4$		In the trigger acceptance region for the MS.
$1.37 < \eta^l < 1.52$		Remove overlap region between EMB and EMEC.
$\text{ptcone20}/\min(p_T^l, 50 \text{ GeV}) < 0.1$		Tracks in ID that are well isolated using the cone described in Section 4.5 with $\Delta R_{\text{max}} = 0.2$. Leptons with $p_T^l > 50 \text{ GeV}$ use $p_T^l = 50 \text{ GeV}$ to calculate the isolation requirement.
$ d_0/\sigma_{d_0} < 5$	$ d_0/\sigma_{d_0} < 3$	Transverse distance between the track and the interaction point should be relatively close.
$\Delta z_0 \times \sin \theta < 0.5 \text{ mm}$		The longitudinal distance between the track and the interaction point should be close.

Table 8.3: List of identification, kinematic selection, and isolation criteria to ensure proper data quality and to select a high purity sample of candidate signal events to be used in this analysis.

8.4 Neutrino Reconstruction

Unlike the particles from proton–proton interactions described in Section 4.3, neutrinos are not directly detected by the ATLAS detector. Instead, they are approximated using the hadronic recoil. When vector bosons, $V = (W, Z)$, are produced from proton–proton interactions, they have a non-zero transverse momentum originating from initial state gluon and quark radiation:

$$\vec{p}_T^V = \vec{p}_T^{\text{lep1}} + \vec{p}_T^{\text{lep2}} = - \sum_{\text{PFO}} \vec{p}_T^{\text{ISR,gluon}} = -\vec{u}_T. \quad (8.2)$$

Here, p_T^V is the transverse momentum of the vector boson, p_T^{lep} is the transverse momenta of the decay leptons (charged leptons or neutrinos), and $\sum_{\text{PFO}} \vec{p}_T^{\text{ISR,gluons}}$ is the hadronic recoil denoted by $\vec{u}_T$. The hadronic recoil accounts for the momenta of any charged and neutral partons that recoil away from the vector boson using Particle Flow Objects (PFO) [73, 110] in the transverse plane, as seen in Figure 8.2a). A depiction of the hadronic recoil for W boson events can be seen in Figure 8.2a), which shows the balance between u_T and p_T^W . The calculation of hadronic recoil from W boson events is calibrated from the decay of Z boson events into leptons, as both leptons can be reconstructed. The lepton resolution is close to an order of magnitude better than the hadronic recoil. This calibration is done by using one of the leptons from the $Z \rightarrow \ell\ell$ event and treating it as missing energy to mimic a neutrino in a $W \rightarrow \ell\nu$ event, as seen in Figure 8.2c). Using the hadronic recoil, the missing energy is then calculated for these Z events and compared to the actual transverse momentum of the removed lepton [12], as seen in Figure 8.2b) which shows good agreement. Uncertainties related to this calibration are discussed in Section 10.2.1. With the hadronic recoil, it is possible to extract the neutrino transverse momentum, p_T^ν , by approximating it with the missing transverse energy, $\vec{E}_T^{\text{miss}}$, which is the vectoral sum of the hadronic recoil and the lepton transverse momentum:

$$\vec{E}_T^{\text{miss}} = \vec{p}_T^\nu = -(\vec{u}_T + \vec{p}_T^\ell). \quad (8.3)$$

Examples of the u_T , E_T^{miss} , and the reconstructed transverse momentum of the Z boson, $p_T^{\ell\ell}$, distributions for $W^- \rightarrow e^- \bar{\nu}$ and $Z \rightarrow e^+ e^-$ events from [12] can be seen in Figure 8.3. Here, the hadronic recoil for the Z boson in Figure 8.3b) is similar in shape to the transverse momentum shown in Figure 8.3d). This calibrates the hadronic recoil reconstructed for the W boson in Figure 8.3a) in order to extract the missing transverse energy shown in Figure 8.3c). These distributions are shown pre-fit; the agreement between data and MC is expected to improve post-fit.

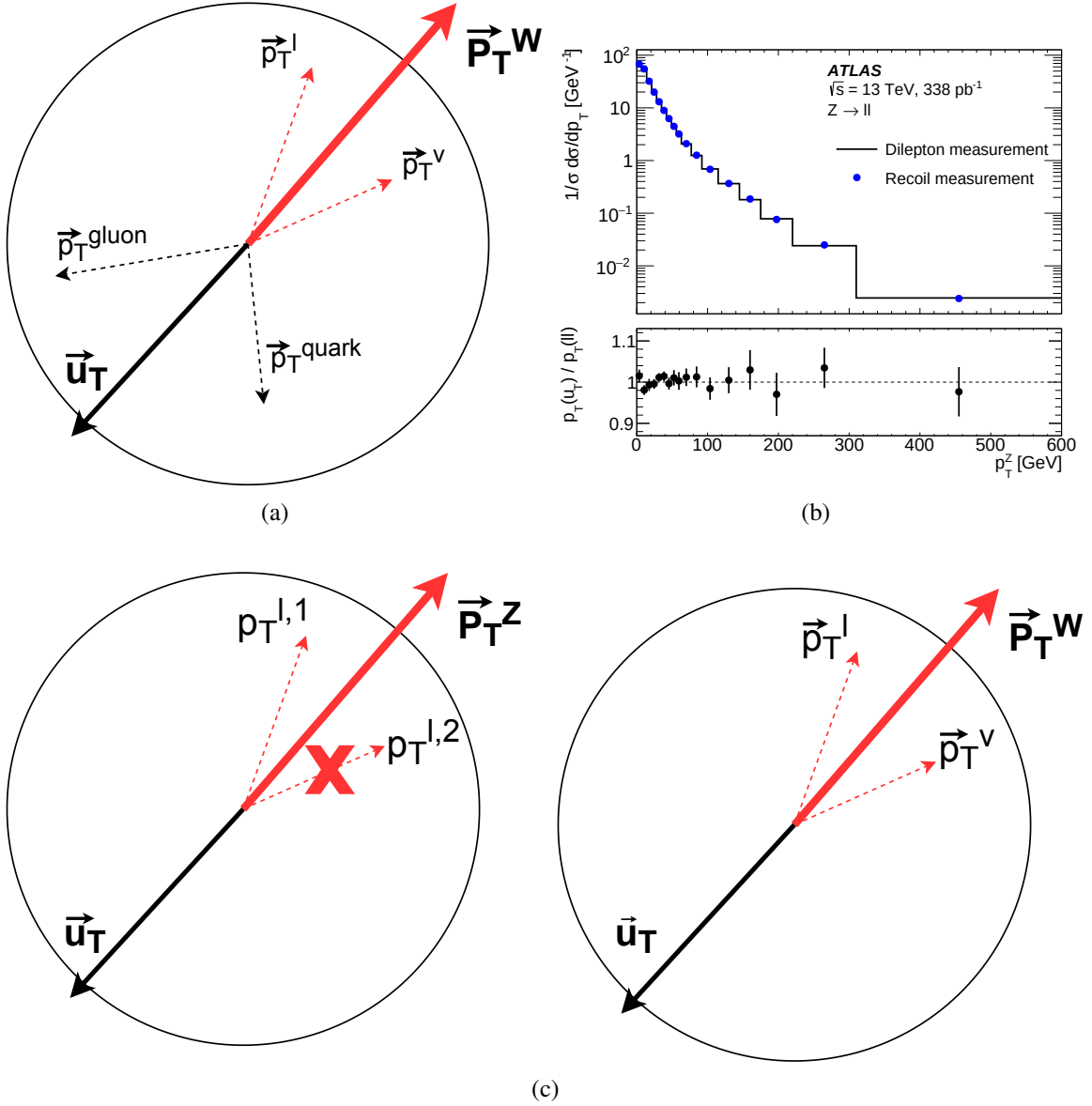
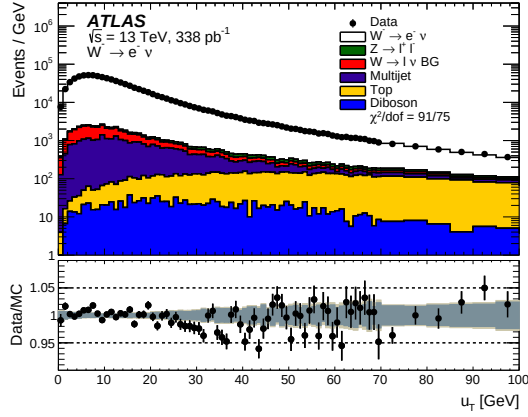
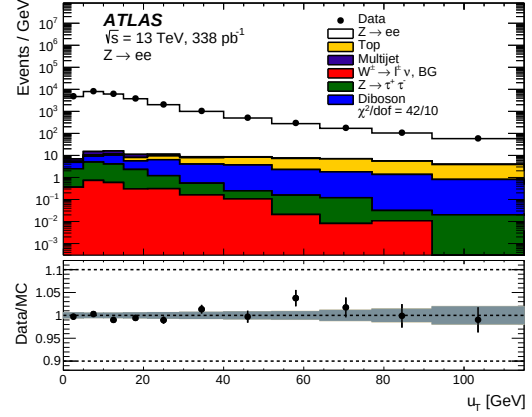


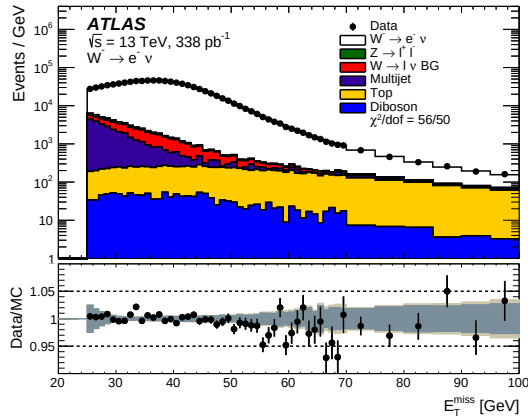
Figure 8.2: The neutrino transverse momentum, p_T^v , is approximated using the E_T^{miss} , which is calculated using the calibrated hadronic recoil, u_T . a) The hadronic recoil, u_T , is calculated by the vectorial sum of all p_T from hadronic activity produced from initial state gluon and quark radiation associated to the event. The hadronic recoil and the charged lepton transverse momentum are used to solve for the missing energy associated the neutrino p_T . b) A comparison of the transverse momentum and hadronic recoil for the Z boson and the bottom panel shows they are in good agreement [12]. c) The hadronic recoil is calibrated by removing a lepton from $Z \rightarrow \ell\ell$ events to calculate E_T^{miss} using u_T and comparing to the measured lepton transverse momentum, $p_T^{l,2}$.



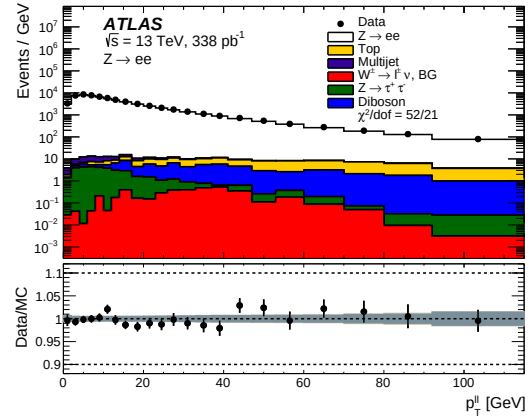
(a) $W^- \rightarrow e^- \bar{\nu} u_T$



(b) $Z \rightarrow ee u_T$



(c) $W^- \rightarrow e^- \bar{\nu} E_T^{\text{miss}}$



(d) $Z \rightarrow ee p_T^{\text{ll}}$

Figure 8.3: Distributions of a) $W^- \rightarrow e^- \bar{\nu}$ hadronic recoil, u_T , b) $Z \rightarrow ee$ hadronic recoil, u_T , c) $W^- \rightarrow e^- \bar{\nu}$ missing transverse energy, E_T^{miss} , and d) $Z \rightarrow ee$ dilepton transverse momentum, p_T^{ll} , for the $\sqrt{s} = 13$ TeV low- $\langle\mu\rangle$ data set from [12]. The data in black points are compared to the sum of the signal and expected backgrounds. The lower panels show the agreement of individual bins between data and the expected signal and background in black points and the prediction uncertainty as a dark grey band. These distributions are shown pre-fit where the expected agreement between data and MC post-fit.

The low pileup data set used in this analysis allows for good resolution of the hadronic recoil because of the lower background in the detector, making it easier to assign the energy associated to each interaction. From Equation 8.3, a better hadronic recoil resolution directly correlates to a better resolution of the missing transverse energy. Equation 8.2 also shows that a better hadronic recoil resolution will improve the resolution of p_T^V , which can also be seen in Figure 8.4. Here, the effects of detector resolution and event migration is illustrated for hadronic recoil for five p_T^Z intervals. An improved resolution would result in narrower distributions, such that most events fall within a smaller range.

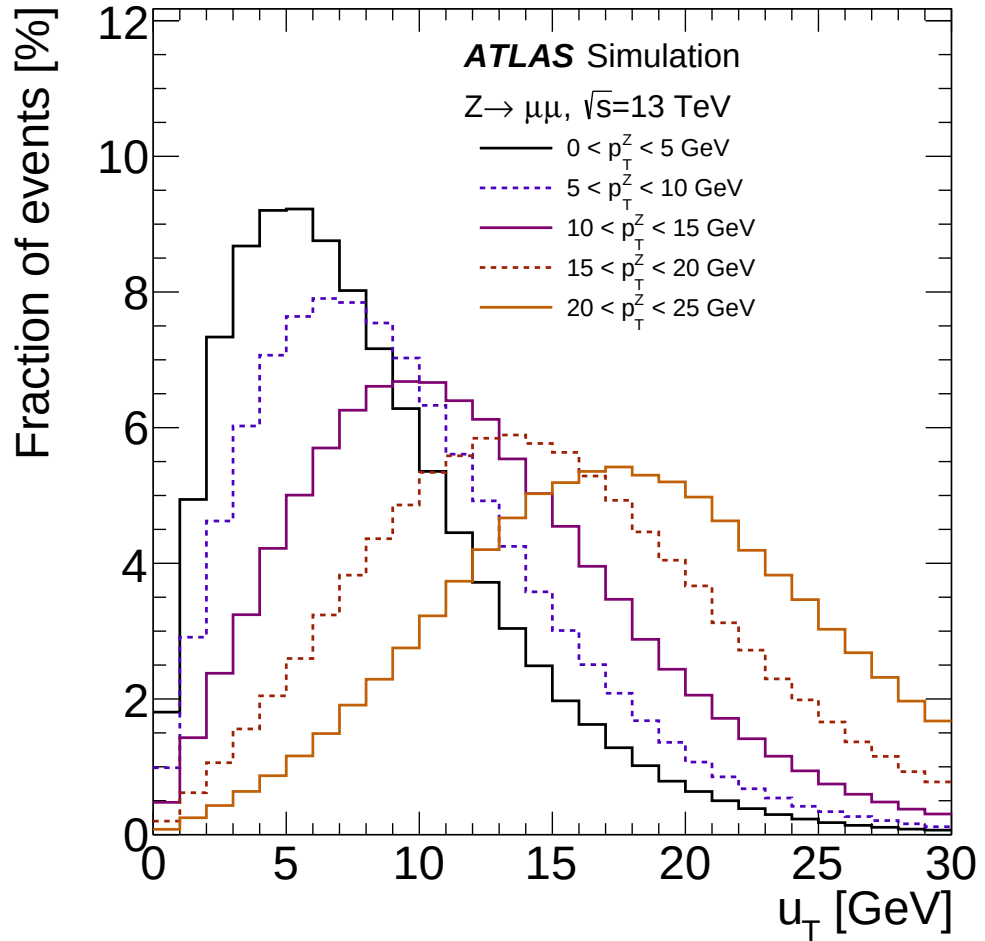


Figure 8.4: The hadronic recoil detector response for five p_T^Z truth intervals for $Z \rightarrow \mu\mu$ events at $\sqrt{s} = 13$ TeV from [12]. This figure illustrates that, with lower hadronic recoil resolution, it is possible to differentiate between lower p_T^Z values. Likewise, it would also allow for better resolution for lower p_T^W values.

Approximating the neutrino's momentum with the missing transverse energy provides information in two dimensions, p_x and p_y in the transverse plane, but the neutrino needs to be fully reconstructed in three dimensions to extract the Drell–Yan angular coefficients. To get the longitudinal momentum of the neutrino, it is assumed that the charged lepton with neutrino decay products of the Drell–Yan process has invariant mass¹ equal to m^W . A charged lepton with four-momentum p^μ and a massless neutrino with four-momentum q^μ are considered in the laboratory frame:

$$\begin{aligned} p^\mu &= \begin{pmatrix} E_\ell \\ \vec{p} \end{pmatrix} & q^\mu &= \begin{pmatrix} |q| \\ \vec{q} \end{pmatrix} \\ \vec{p} &= \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} & \vec{q} &= \begin{pmatrix} q_x \\ q_y \\ q_z \end{pmatrix}, \end{aligned} \tag{8.4}$$

where

$$|p| = \sqrt{p_x^2 + p_y^2 + p_z^2} \quad E_\ell = \sqrt{|p|^2 + m_\ell^2} \quad |q| = \sqrt{q_x^2 + q_y^2 + q_z^2}. \tag{8.5}$$

To determine the longitudinal momentum of the neutrino the invariant mass of the charged lepton and neutrino system is set to the invariant mass of the W boson, $m^W = 80379$ MeV:

$$(p + q)^\mu (p + q)_\mu = (m^W)^2. \tag{8.6}$$

This expression can be expanded and rearranged into powers of the neutrino longitudinal momentum, q_z , allowing it to be expressed in the form of a quadratic equation:

$$aq_z^2 + bq_z + c = 0, \tag{8.7}$$

¹ Throughout this thesis, the combined measured value of $m^W = 80379 \pm 12$ MeV [111], which does not include the 2022 measurement of m^W by the CDF Collaboration [14].

where:

$$\begin{aligned}
a &= 4(p_x^2 + p_y^2 - m_\ell^2) \\
b &= 4p_z\alpha \\
c &= \alpha^2 - (2E_\ell q_T)^2 \\
\text{and } \alpha &= m^{W2} + m_\ell^2 + 2(p_x q_x + p_y q_y).
\end{aligned} \tag{8.8}$$

Finally, it is then possible to get two solutions for the longitudinal momentum of the neutrino using the quadratic formula:

$$q_z(\pm) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{8.9}$$

There is also the possibility for one solution if $\Delta = b^2 - 4ac = 0$, or no real solutions if $\Delta < 0$.

8.4.1 Solution Choice

The rare and simplest case is when $\Delta = 0$ and there is only one solution, where $p_z' = -b/2a$. Most frequently, there are two solutions that are equally valid. One of the solutions is chosen at random and, with enough statistics, the correct p_z' distribution is able to be resolved on average. This can be seen in Figure 8.5, where the positive and negative root choices are denoted by triangular points in red and blue, respectively. In this figure, the random choice of solutions are denoted by the open black points and the generated p_z' denoted by solid black points. Choosing only one solution would bias the distribution, while choosing one at random allows the distribution to get close to what was generated by the MC. Extensive tests were performed to try and improve the rate at which the correct solution was chosen but no better method was found. In a small subset of events, one of the solutions is not physically possible because it would require p_z' or $p_z' + p_z^\ell$ to have more energy than the beam energy, $\sqrt{s}/2$, allowing for one solution to be discarded. The choice of p_z' most importantly affects the value of $\cos \theta_{CS}$, which is discussed in Appendix F.

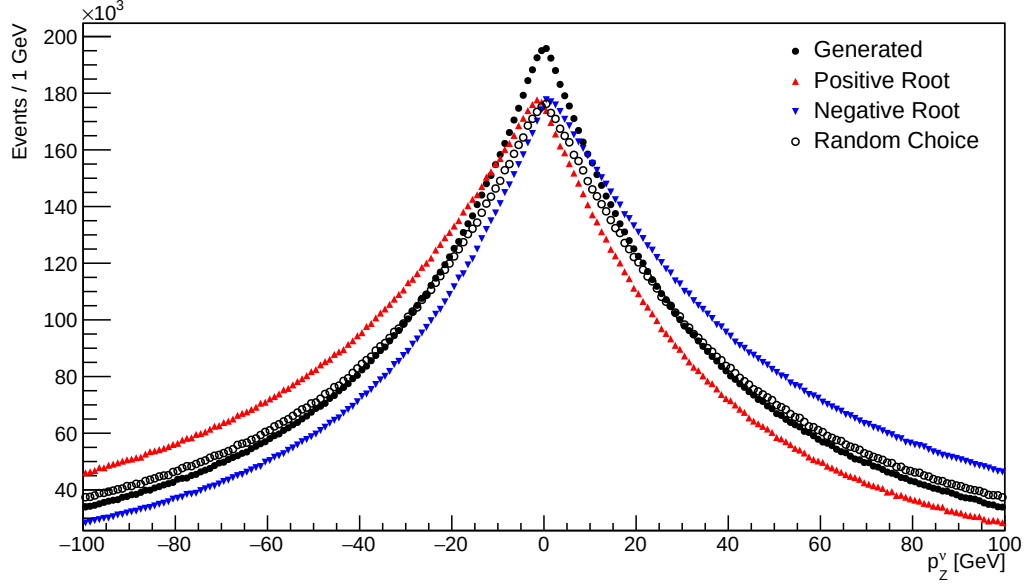


Figure 8.5: Comparison of the different quadratic equation root choices for p_z^ν reconstruction to the one generated in MC. The solid black points are the values of p_z^ν generated, the red upward triangles are the p_z^ν values when the positive root is chosen, the blue downward triangles are the p_z^ν values when the negative root is chosen, and the open black points are randomly choosing between the two different roots.

The last case is when there are no real solutions, which occurs when there is a negative under the square root and is a consequence of the method by setting the total energy of the system equal to m^W . In real events, W bosons are not always produced with the “on-shell” mass of $m^W = 80379 \pm 12$ MeV, and can be produced “off-shell”, where the mass can be less than or greater than the on-shell mass. When this happens, the transverse mass, m_T , which is defined as:

$$m_T = \sqrt{2p_T^\nu p_T^l (1 - \cos \Delta\phi^\nu)}, \quad (8.10)$$

can be greater than the on-shell mass that is used to solve for p_z^ν in Equation 8.6. When $m_T > m^W$ there is no real solution, which happens more frequently with increasing $p_T^{l\nu}$, as seen in Figure 8.6 which shows the percentage of events that have a real solution. At low $p_T^{l\nu}$, approximately 80% of events have a real solution and this rapidly decreases to $\sim 50\%$ of events for $p_T^{l\nu} > 100$ GeV.

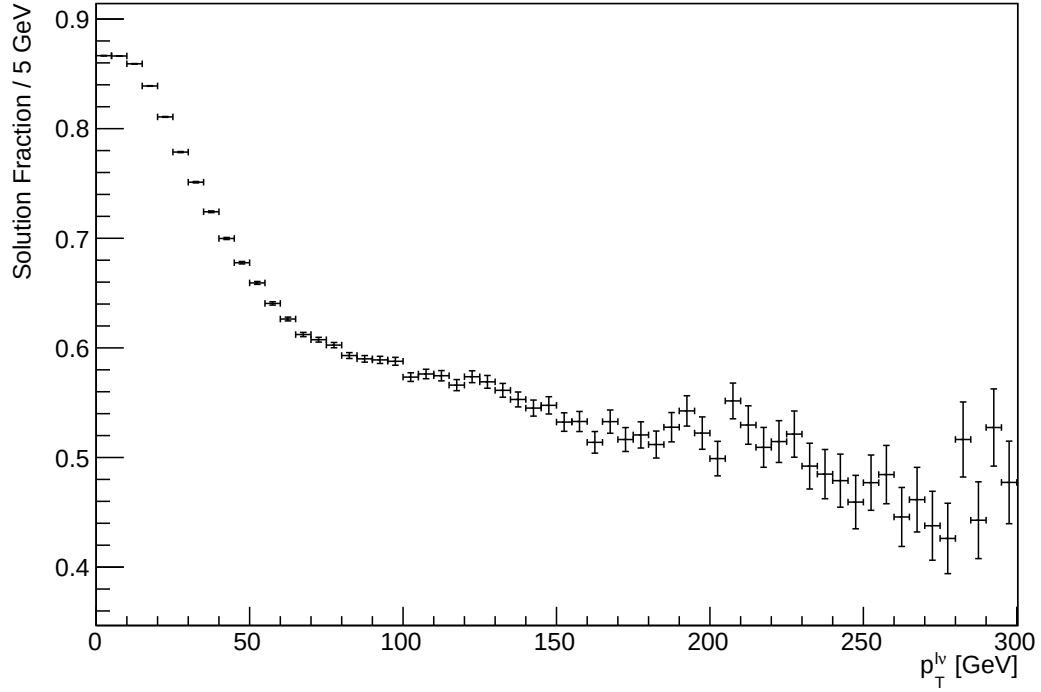


Figure 8.6: Fraction of events that have a real solution as a function of p_T^{ν} . Events will not have a solution when $m_T > m^W$, which happens more frequently at large p_T^{ν} .

These events with no solution are still useful as the information in the transverse plane, which is described by ϕ_{CS} , does not need a solution. In order to use these events, the p_T^{ν} is rotated by the minimum amount needed to arrive at a real solution. This rotation finds the four vector of the neutrino needed for $\cos \theta_{CS} = 0$ and adjusts the ϕ_{CS} to match, which brings it closer to the expected value. When filling the reconstructed event templates, described in Section 7.2, these events are put in a $1.25 < \cos \theta_{CS}^{\text{Reco}} < 1.5$ bin. This separates them from events with a real solution to avoid distortion of the shape for the angular distributions but still allowing for them to be used.

8.5 W Boson Reconstruction

To fully reconstruct a W boson from a Drell–Yan interaction, both a charged lepton and a neutrino are required. The electrons are reconstructed as described in Section 4.3.2,

the muons as described in Section 4.3.4, and the neutrinos are reconstructed as described in Section 8.4. If the event passes the selection criteria outlined in Section 8.3, then the four vector of the W boson is equal to adding the four vectors of the charged lepton and the neutrino. With the W boson, charged lepton, and neutrino four vectors, the variables used for the transverse momentum, p_T^W , and the Collin–Soper angles can be calculated according to Section 6.2. An example of a $W^- \rightarrow \mu^- \bar{\nu}$ candidate event from proton–proton collisions recorded by ATLAS at $\sqrt{s} = 13$ TeV, from the low- $\langle\mu\rangle$ data set [12] can be seen in Figure 8.7.

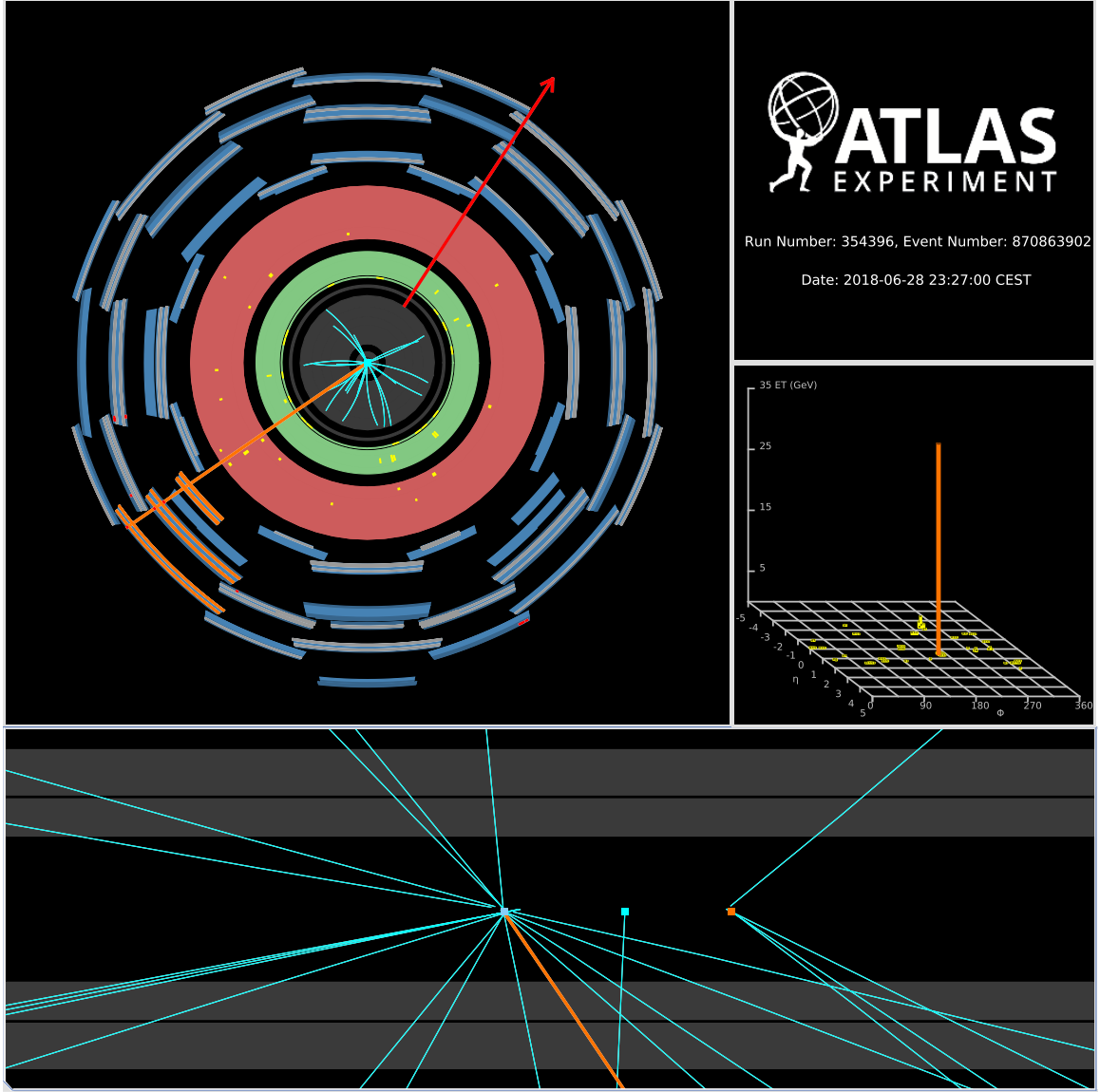


Figure 8.7: An event display of a $W^- \rightarrow \mu^- \bar{\nu}$ candidate event from proton–proton collisions recorded by ATLAS at $\sqrt{s} = 13$ TeV, from the low- $\langle\mu\rangle$ data set [12]. The event has a missing transverse energy of 49 GeV and the right display shows the transverse energy of the muon is 35 GeV. The muon and missing transverse energy are used to reconstruct a W boson candidate with reconstructed transverse momentum, $p_T^{l\nu} = 16$ GeV, and transverse mass, $m_T = 77$ GeV. In the left display, the path of the muon is shown in orange and the missing transverse energy is denoted by the red arrow. There are two additional reconstructed vertices, as seen in the bottom panel, which shows charged particle tracks from other particles in blue and the track of the muon in orange.

8.6 Data Driven Background Estimation

Not all background processes can be properly simulated using MC, such as hadronic events that mimic electrons. This type of background is commonly referred to as multi-jet (MJ) and has to be estimated using data driven methods. For this analysis, a data driven estimate was constructed by fitting the CS angular distributions in a MJ enriched control region and the shape is extrapolated back to the analysis region. This MJ background is one of the dominant sources of background in the electron channel so some of the selection criteria in Section 8.3 are applied to reduce the amount of MJ background because the multiple physics processes encompassing such a background are difficult to model. This includes requirements on identification, E_T^{miss} , and m_T .

The MJ enriched control region is found by reversing the isolation requirements so that the region has primarily MJ and not much signal. For electrons and muons, the track based isolation from Table 8.3 was reversed so that $\text{ptcone20}/\min(p_T^l, 50 \text{ GeV}) > 0.1$. Within the control region, the signal and the background processes from MC simulation are subtracted from data to provide the distribution of MJ. The control region (CR) is subdivided into four control regions – CR1, CR2, CR3, CR4 –, where the CS angular distributions are fit over four variables – p_T^l , E_T^{miss} , m_T , and $\Delta\phi(\ell, E_T^{\text{miss}})$ – in each region. Examples of the control regions and signal region (SR) for $W^+ \rightarrow \mu^+ \nu$ can be seen in Figure 8.8. Here, the comparison between the desired signal of $W^+ \rightarrow \mu^+ \nu$ events from MC in red are compared to the MJ background distribution in blue, which dominates the control regions. Each control region fit estimates the shape and normalization of the MJ distribution in the individual control regions. With the four control region fits for each variable, it can be extrapolated back to estimate the MJ in the signal rich region. The difference between the four variables provides an uncertainty on the MJ estimate.

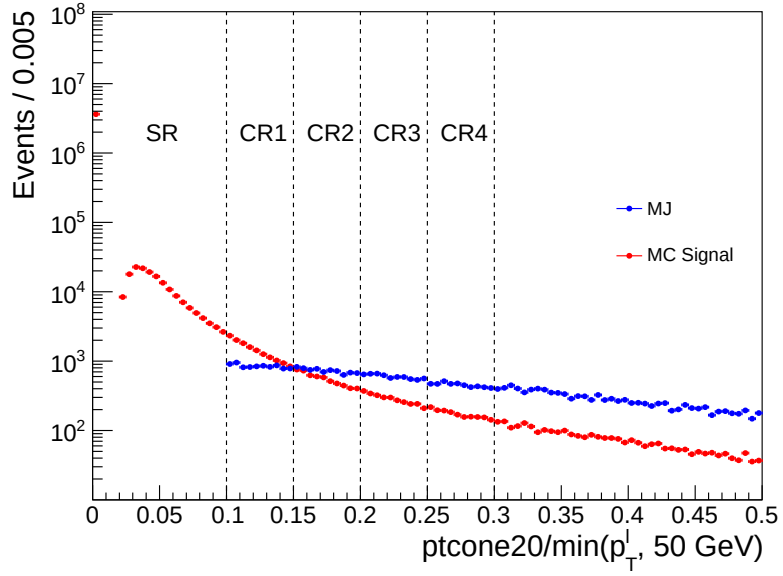


Figure 8.8: Examples of the track isolation variable for $W^+ \rightarrow \mu^+ \nu$ that is used to estimate the MJ background. Control regions (CR) that are anti-isolated are used to fit the MJ shape and yield, then extrapolated back to the signal region (SR). The amount of MJ in the SR is not shown here because this is pre-extrapolation.

8.7 Kinematic Distributions

To ensure that the modelling of MC and that the data-driven MJ background estimation are correct, various kinematic distributions are compared to those from data. This allows for testing different aspects of the analysis – such as making sure the appropriate SFs from particle trigger, identification, and reconstruction – are applied correctly. This section will show the kinematic distributions for the $W^- \rightarrow e^- \bar{\nu}$ channel but the other channels are available in Appendix G. The kinematic distributions presented are for events that pass the selection criteria – outlined in Table 8.3 –, for the low pileup data set at $\sqrt{s} = 13$ TeV – described in Section 8.1 –, the MC signal and background – described in Section 8.2 –, and the data-driven MJ background – described in Section 8.6.

All the kinematic distributions shown are presented on a logarithmic scale as many distributions, such as $p_T^{l\nu}$, cover many orders of magnitude. This allows for the distributions of the background processes to be visible. The data are denoted by black points, while the

modelling of the signal and background MC and data-driven MJ distributions are denoted by histograms of various colours. These histograms are stacked successively and the combination into one model should be what is expected with data. This stacking is performed in order of increasing contribution, with the smallest contribution at the bottom up to the largest contribution. The contribution order is visible in the legend in the top right corner of each plot, which also contains the number of events and the percentage of the total number of data events. Within the plots, the single boson backgrounds, such as $W^- \rightarrow \tau^- \bar{\nu}$ and $Z \rightarrow ee$, are separated to observe the differences between their kinematic distributions, while the different sources of diboson and top related backgrounds are summarized into one category.

There is some difference in the number of events in each plot, as some plots only show a particular range and some events fall outside of this range. Lastly, below the main plot is the ratio of the data over the total model to show the pre-fit agreement between the distributions, where a perfect agreement would be $\text{Data/Model} = 1$. For the ratio plots, the error bars on the points are the statistical uncertainty of the data, while the green band is the systematic uncertainty on the data-driven MJ background. All of the kinematic distributions in this chapter are pre-fit so the agreement between data and the model, characterized by the χ^2 , are often describing that there is not good agreement. During the fitting process, the model is varied within its statistical uncertainty and within the systematic variations to find the best agreement to data. The best fit is not back propagated to the kinematic distributions in this section, to show that there is better agreement post-fit. In the next chapter, the post-fit distributions of the CS angles as a function of $p_T^{l\nu}$ will be compared to those from data and the pre-fit distribution. In the context of the figures shown in this section for pre-fit distributions, an agreement within $\sim 5\%$ or within uncertainty is considered to be acceptable as the simulation is known to not be perfect and this will be brought into better agreement during the fitting process. This $\sim 5\%$ is based on the pre-fit modeling uncertainties from sources like the luminosity, background modeling, and hadronic recoil.

First, the kinematic distributions for the charged lepton and the neutrino, modelled through the missing transverse energy, E_T^{miss} , are shown in Figure 8.9. These figures show the primary distributions of the detected particles before the W boson is reconstructed. In Figures 8.9a) and b), the kinematic distributions for the p_T^l and η^l for the charged lepton are shown. These distributions show there is an agreement between the data and the modelling on the order of 5% as most points fall within this range. For the p_T^l distribution in Figure 8.9a), there is the expected Jacobian peak at 40 GeV, which is half the mass of the W boson. The other feature of note within the plot is the little drop off around 65 GeV, which can be attributed to the isolation selection criteria, $\text{ptcone20}/\min(p_T^l, 50 \text{ GeV}) > 0.1$, where the shape is slightly changed but allows for more background rejection. The η^l distribution in Figure 8.9b) appears as expected since there is agreement on the level of 2% for most points and the selection criterion that removes the crack region of $1.37 < |\eta^l| < 1.52$ is quite visible. In Figure 8.9c), the E_T^{miss} distribution also has a peak at 40 GeV, which is slightly broader than what is seen for p_T^l , as expected, because E_T^{miss} has worse resolution. It also has most points within 5%, making the agreement on par with the p_T^l distribution. Lastly, Figure 8.9d) shows the angular separation in the transverse plane between the charged lepton and the E_T^{miss} from the neutrino. Here, there is good agreement for the majority of the distribution as most points fall within 3% and the only disagreement is at small $|\Delta\phi|$ values. This plot also gives insight into what is happening during the decay of the W boson into a charged lepton and a neutrino. In the rest frame of the W boson, the charged lepton and the neutrino are in opposite directions. Thus, large $|\Delta\phi|$ correlates to the W boson being produced mostly in the transverse direction with a small y^W , where it is possible to observe the decay products in opposite directions. In contrast, a smaller $|\Delta\phi|$ occurs when the charged lepton and the neutrino are traveling in opposite directions with a larger η contribution, correlating to the W boson being produced with a larger y^W .

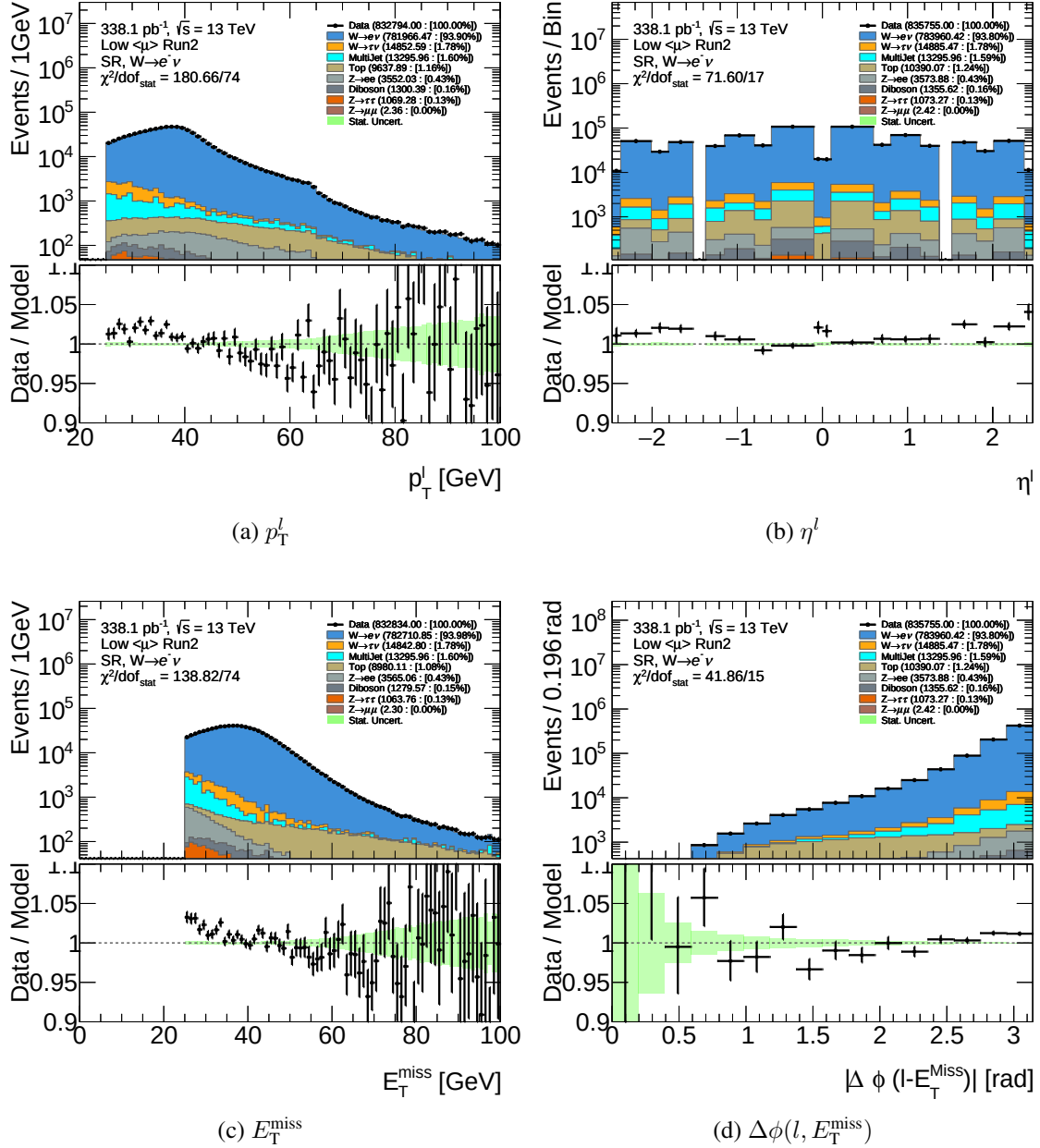


Figure 8.9: The kinematic distributions for a) p_T^l , b) η^l , c) E_T^{miss} , and d) $\Delta\phi(l, E_T^{\text{miss}})$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

With the knowledge that the pre-fit leptons' kinematic distributions are modelled within 5% or agree within uncertainty, the reconstructed leptons can be used to reconstruct the W bosons as, outlined in Section 8.5. This allows plotting of the kinematic distributions of the W boson, as seen in Figure 8.10. Figure 8.10a) is the transverse mass of the W boson, m_T , where there is agreement on the level of 5% for most bins. This figure also shows a peak at the W boson mass of ~ 80 GeV, where there is very good agreement. Figures 8.10b) and c) show $p_T^{l\nu}$ in a fine binning of 1 GeV increments and in the binning used for the measurement listed in Equation 7.5, respectively. The fine binning allows for a more granular view of any potential mismodelling, while the measurement binning displays the pre-fit agreement between the data and the modelling. Both figures show an agreement on the level of 5% or agree within uncertainty for most bins. There is some mismodelling that can be seen in Figures 8.10b) and c), where below 40 GeV there is an excess of data and above 40 GeV there is an excess in the model for most points. Lastly, Figure 8.10d) shows $|y^{l\nu}|$ and the agreement on the level of 3% for all bins. Recall from Chapter 7, the final measurement is three-dimensional ($\cos \theta_{CS}$, ϕ_{CS} , $p_T^{l\nu}$) so Figure 8.10c) allows for verifying the agreement when $\cos \theta_{CS}$ and ϕ_{CS} are integrated over. This can help identify any discrepancies that are not apparent when looking at a more divided phase space.

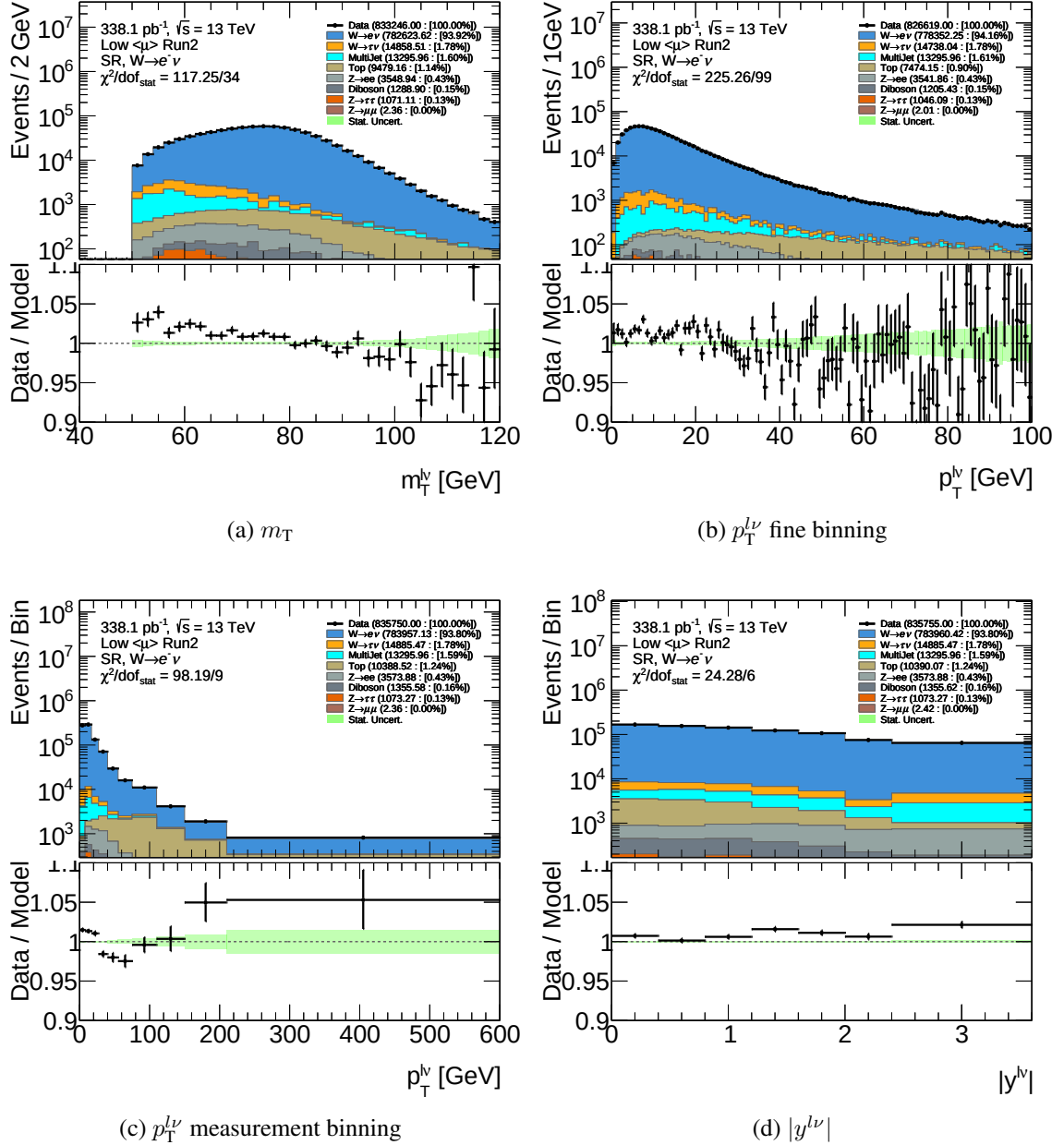


Figure 8.10: The kinematic distributions for a) m_T , b) $p_T^{l\nu}$ in a fine binning, c) $p_T^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

Similar to looking at $p_T^{l\nu}$, where $\cos \theta_{\text{CS}}$ and ϕ_{CS} were integrated over, it is also important to look at each CS angle when integrating over the other two variables. This can be seen in Figure 8.11, which show $\cos \theta_{\text{CS}}^{\text{Reco}}$ and $\phi_{\text{CS}}^{\text{Reco}}$ in an integrated phase space. For $\cos \theta_{\text{CS}}^{\text{Reco}}$ in Figure 8.11a), there is agreement on the level of 2% for most bins, with the exception of the bin for $0.75 < \cos \theta_{\text{CS}}^{\text{Reco}} < 1$, which has an agreement of $\sim 5\%$. Recall from Section 8.4, the $\cos \theta_{\text{CS}}^{\text{Reco}}$ bin of $1.25 < \cos \theta_{\text{CS}}^{\text{Reco}} < 1.5$ is filled with events where there is no real solution for the neutrino longitudinal momentum. These are separated from the events with a real solution to avoid distorting the distribution around $\cos \theta_{\text{CS}}^{\text{Reco}} \sim 0$ and has good agreement. Figure 8.11b) shows $\phi_{\text{CS}}^{\text{Reco}}$ is in good agreement with the model as all points are within 2%. With the integrated phase spaces showing reasonable agreement, the next step is to look at the distributions in the three-dimensional phase space, which will be used to extract the angular coefficients.

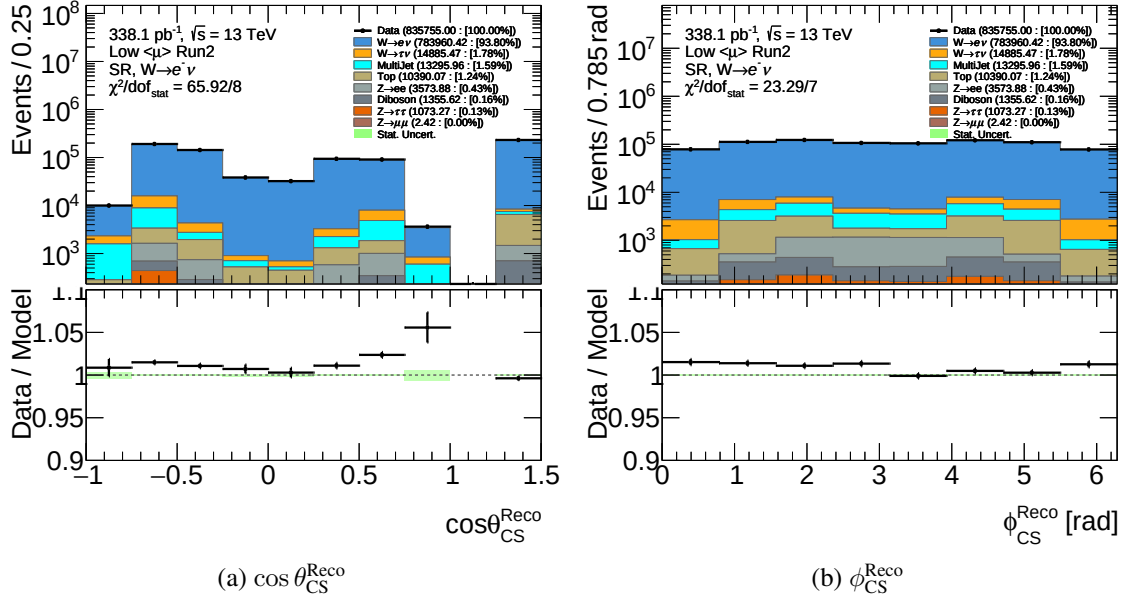


Figure 8.11: The integrated CS angular distributions of a) $\cos \theta_{\text{CS}}^{\text{Reco}}$ and b) $\phi_{\text{CS}}^{\text{Reco}}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

The final kinematic distributions to analyze are the non-integrated three-dimensional phase space $(\cos \theta_{\text{CS}}, \phi_{\text{CS}}, p_{\text{T}}^{\nu})$, as shown in Figures 8.12 and 8.13. These figures show some of the bins as examples of how the angular distributions change as a function of p_{T}^{ν} . The remaining bins, not shown in these Figures, can be found in Appendix G. To represent the three-dimensional phase space in a one-dimensional plot, each figure is of a particular p_{T}^{ν} bin where the $\phi_{\text{CS}}^{\text{Reco}}$ distribution is plotted for each successive $\cos \theta_{\text{CS}}^{\text{Reco}}$ bin. For example, Figure 8.12a) is for the first p_{T}^{ν} bin for the range $0 < p_{\text{T}}^{\nu} < 8 \text{ GeV}$, where the first eight bins are the $\phi_{\text{CS}}^{\text{Reco}}$ distribution for $-1 < \cos \theta_{\text{CS}}^{\text{Reco}} < -0.75$, the next eight bins are for $-0.75 < \cos \theta_{\text{CS}}^{\text{Reco}} < -0.5$, and this pattern continues for each $\cos \theta_{\text{CS}}^{\text{Reco}}$ bin. Unwrapping the distribution in this manner displays how well the data and model agree in a simplified manner, while still retaining all three-dimensional information.

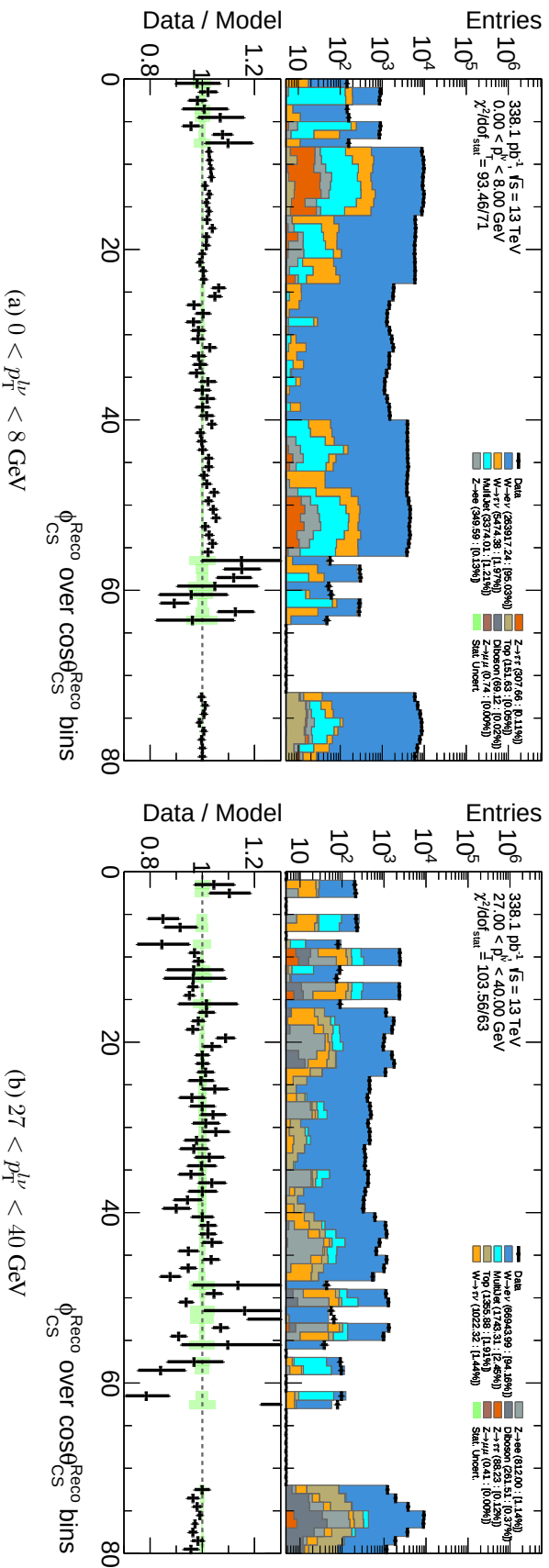


Figure 8.12: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of p_T^{μ} for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for two bins: a) $0 < p_T^{\mu} < 8$ GeV and b) $27 < p_T^{\mu} < 40$ GeV. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

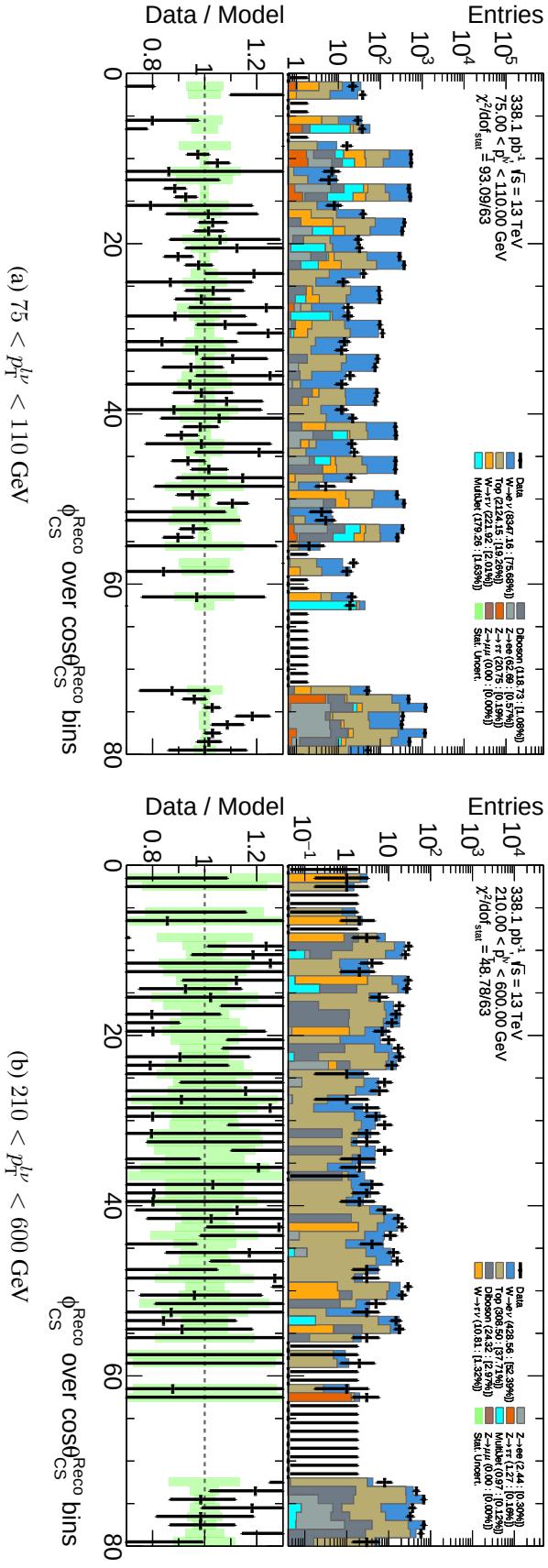


Figure 8.13: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_{\text{T}}^{l\nu} < 110$ GeV and b) $210 < p_{\text{T}}^{l\nu} < 600$ GeV. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

Figures 8.12 and 8.13 show how the angular distributions change as a function of $p_T^{l\nu}$. The first thing to notice is how the ϕ_{CS}^{Reco} distribution changes as a function of $p_T^{l\nu}$, where the first $p_T^{l\nu}$ bin is shown in Figure 8.12a). Here, the ϕ_{CS}^{Reco} distribution in the outer $\cos \theta_{CS}^{\text{Reco}}$ bins of $0.75 < |\cos \theta_{CS}^{\text{Reco}}| < 1$ has two peaks around $\phi_{CS}^{\text{Reco}} \sim \pi/2$ or $\phi_{CS}^{\text{Reco}} \sim 3\pi/2$, while the remaining $\cos \theta_{CS}^{\text{Reco}}$ bins have a flatter shape. As $p_T^{l\nu}$ is increased through Figures 8.12b) and 8.13, more of the $\cos \theta_{CS}^{\text{Reco}}$ bins start to have a similar shape in ϕ_{CS}^{Reco} . To understand what this means physically, it is beneficial to recall the definition of the CS angles from Figure 6.6. In this special W boson rest frame, ϕ_{CS} is the azimuthal angle between the negatively charged lepton and the proton–proton plane, while $\cos \theta_{CS}$ is the polar angle between the negatively charged lepton and the z -axis that points in the direction the W boson travels in the lab frame. This means $\phi_{CS}^{\text{Reco}} \sim \pi/2, 3\pi/2$ is when the leptons are perpendicular to the W boson, which at low $p_T^{l\nu}$ only happens when $|\cos \theta_{CS}^{\text{Reco}}| \sim 1$ or when the negatively charged lepton is (anti)parallel to the direction of the W boson in the rest frame. As $p_T^{l\nu}$ increases, the leptons tend toward being perpendicular to the W boson for both angles.

The other thing to notice, which can also be seen in Figure 8.10c), is the change in the background contributions as a function of $p_T^{l\nu}$. At low $p_T^{l\nu}$, the dominant sources of background are single boson processes, such as $W^- \rightarrow \tau^- \bar{\nu}$, and the data-driven MJ. For high $p_T^{l\nu}$ shown in Figure 8.13, the main sources of background are top related processes and contribute a more significant fraction of events to the total number of data events. This means that the expected statistical sensitivity will be decreased at high $p_T^{l\nu}$, due to an increased background contamination.

These three-dimensional distributions are fit and the resulting post-fit distributions are used to extract the angular coefficients and differential cross-section, as will be discussed in Chapter 9. From Figures 8.12 and 8.13, it can be seen that more bins experience statistical fluctuation compared to the integrated distributions previously shown. However, most of this fluctuation occurs in the outer $\cos \theta_{CS}^{\text{Reco}}$ bins of $0.75 < |\cos \theta_{CS}^{\text{Reco}}| < 1$ where this

behaviour is to be expected, as seen previously in Figure 8.11a). The other thing to note is that these are the distributions before they are fit and constrained. Therefore, the agreement between the model and data will be improved as statistical and systematic variations are used to fit the model to better describe what is measured with data. In Chapter 9, the pre-fit distributions, post-fit distributions, and their compatibility with the measured data will be shown.

CHAPTER 9

MEASUREMENT OF W BOSON DRELL–YAN ANGULAR COEFFICIENTS AND DIFFERENTIAL CROSS-SECTION

This chapter discusses the extraction of the Drell–Yan angular coefficients and differential cross-section for the W boson at $\sqrt{s} = 13$ TeV with the low pileup data set outlined in Section 8.1 using the methodology outlined in Chapter 7. First, the results for $W^- \rightarrow e^- \bar{\nu}$ will be presented in Section 9.1 as a function of $p_T^{l\nu}$ and compared to the predicted values. Next, the compatibility between the coefficients measured separately with electron and muon channels will be shown in Section 9.2. The channels are compatible to each other, allowing them to be combined to decrease the data statistical uncertainties and will be discussed in Section 9.3. Lastly, the measurements will be compared in Section 9.4 to the theoretical predictions reweighted to different modern modelling of parton distribution function (PDF) sets to evaluate the impact of the lack of knowledge of PDFs on the measurement.

To avoid repetition of plots for all four channels ($W^+ \rightarrow e^+ \nu$, $W^+ \rightarrow \mu^+ \nu$, $W^- \rightarrow e^- \bar{\nu}$, and $W^- \rightarrow \mu^- \bar{\nu}$), only examples of one or two channels will be shown. The remaining channels are available in Appendix G. All of the results are plotted with the total uncertainty which is dominated by the data statistical uncertainty, unless noted otherwise. A deeper look into the sources of uncertainty will be discussed in Chapter 10. Some of the coefficients presented in this chapter are being measured for the first time, namely A_0 and A_4 . However, A_2 and A_3 , are being improved upon when compared to the results from the CDF Collaboration [10]. These two coefficients are measured as function of an increased number of $p_T^{l\nu}$ bins, providing better resolution while maintaining a reduced uncertainty.

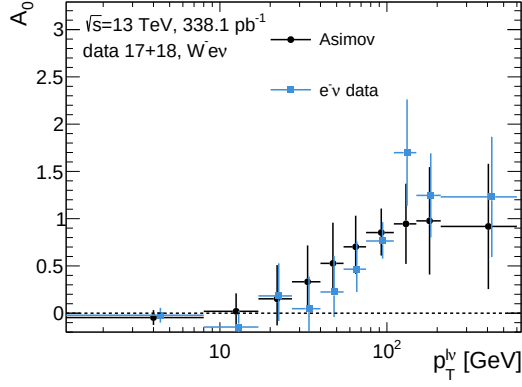
9.1 Single Channel Results

In this section, the results for the $W^- \rightarrow e^- \bar{\nu}$ channel as a function of $p_T^{l\nu}$ will be presented and discussed. These results will be presented for the first five coefficients, A_0 to A_4 , and for the total differential cross-section. The remaining coefficients, A_5 to A_7 , can be found in Appendix G since there is not enough statistical sensitivity to have meaningful results. The results are compared to the predicted values from using an Asimov data set [112]. This Asimov data set is a pseudo-data set using MC simulation and the data-driven MJ background. It is generated by setting the number of observed events equal to the expected events, $N_{\text{obs}} = N_{\text{exp}}$, in the likelihood (LH) function in Equation 7.8. With the Asimov data set the LH fit returns the reference coefficients that are used to build the templates described in Section 7.2.

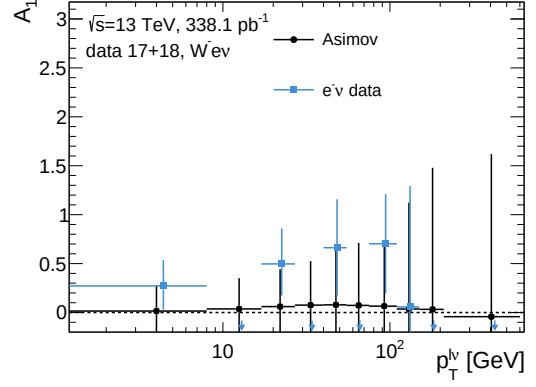
The comparison between N_{obs} , N_{exp} pre-fit, and N_{exp} post-fit will be presented to show how the agreement improves post-fit with the minimization of the negative log-likelihood (NLL). Comparisons between W^- and W^+ will also be shown for A_3 , A_4 , and the differential cross-section as these are expected to depend on the charge of the W boson. This difference is because these coefficients are linearly dependent on the vector and axial-vector couplings to the leptons and quarks, which can be seen in Equation 6.4. These couplings are different for W^+ and W^- as they depend on the third component of the weak isospin $T^3 = \pm 1/2$.

The coefficients are extracted from the LH fit for each channel individually and compared with the predicted values. Fitting channels individually allows testing of the modelling of each individual channel, evaluating the compatibility between channels, and ensuring the variations for the systematic uncertainties are properly propagated and not over-constrained, as will be discussed in Chapter 10. An example of the first five coefficients, A_0 to A_4 , and differential cross-section extracted with data as a function of $p_T^{l\nu}$ can be seen in Figure 9.1. Here, the extracted coefficients with data are denoted by the blue squares and

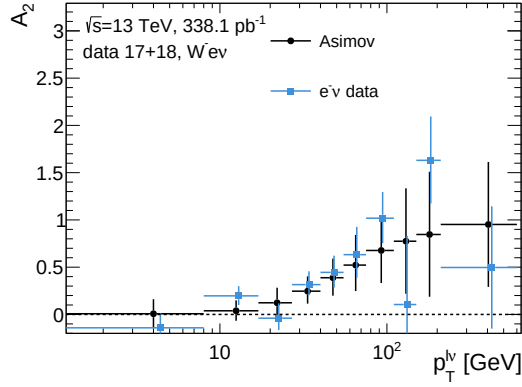
the expected coefficients derived using Asimov data denoted by the black circles. The coefficients are ratios and are normalized by the differential cross-section, which means they are heavily dependent on how well the differential cross-section agrees with the expected value. This is especially the case for coefficients that are dependent on $\cos \theta_{\text{CS}}$, namely A_0 and A_4 , where the fluctuations in Figure 9.1a) and e) are often in the opposite direction of the fluctuations in the differential cross-section plot in Figure 9.1f). These fluctuations in the differential cross-section are a result of the normalization floating within its statistical uncertainty by the amount needed to minimize the difference between N_{obs} and N_{exp} . The coefficients A_2 and A_3 are ϕ_{CS} dependent and can be seen to be in much better agreement in Figure 9.1c) and d) to the expected value. Lastly, it can be seen in Figure 9.1b) that there is no sensitivity to A_1 as the expected values are consistent with zero due to the large uncertainties. This is because A_1 is proportional to the interference between the transverse and longitudinal polarization cross-sections of the W boson, so it is an order of magnitude smaller than the other A_0 to A_4 coefficients.



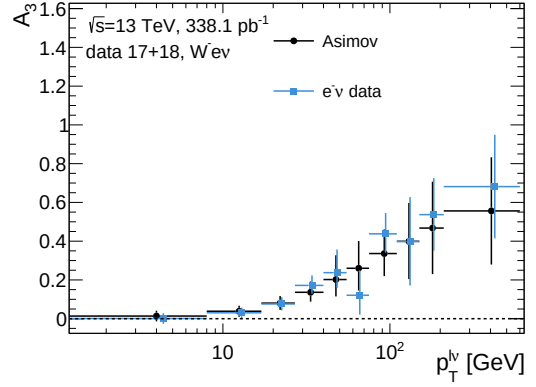
(a) A_0



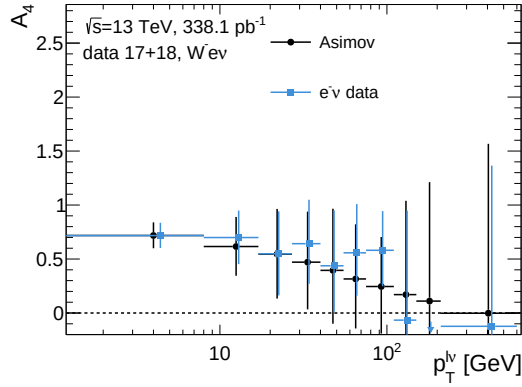
(b) A_1



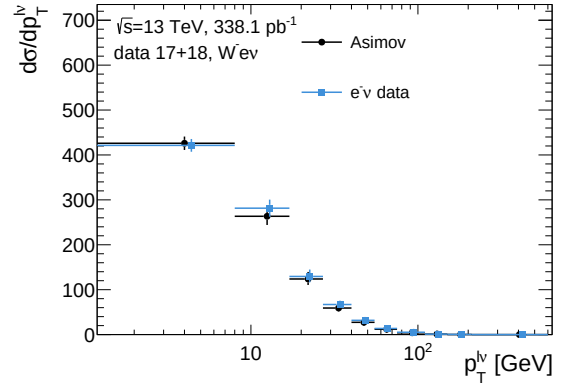
(c) A_2



(d) A_3



(e) A_4



(f) Differential cross-section

Figure 9.1: The measured first five coefficients, A_0 to A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for the $W^- \rightarrow e^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are shifted horizontally to avoid point and uncertainty bar overlap but have the same uncertainty and position in $p_T^{l\nu}$.

The extraction of the angular coefficients come from the minimization of the NLL for the expected and observed number of events in the three-dimensional space ($\cos \theta_{\text{CS}}$, ϕ_{CS} , p_{T}^{ν}). The pre- and post-fit kinematic distributions for this three-dimensional phase space can be seen in Figure 9.2 for the $W^- \rightarrow e^- \bar{\nu}$ channel, while the remaining channels can be found in Appendix G. A summary of the goodness of fit, χ^2 , numbers of degrees of freedom, and the associated p-values - which is the probability under the assumption of the null hypothesis [93] – for each channel are summarized in Table 9.1. The NLL is related to the goodness of fit, χ^2 , through:

$$\chi^2 = -2 \ln \mathcal{L}. \quad (9.1)$$

Because it is the three-dimensional space ($\cos \theta_{\text{CS}}$, ϕ_{CS} , p_{T}^{ν}) being fit to extract the coefficients, the number of degrees of freedom is:

$$ndf = N_{\cos \theta_{\text{CS}}}^{\text{bins}} \times N_{\phi_{\text{CS}}}^{\text{bins}} \times N_{p_{\text{T}}^{\nu}}^{\text{bins}} - N_{A_{ij}}. \quad (9.2)$$

Channel	χ^2/ndf	p-value
$W^- \rightarrow e^- \bar{\nu}$	541.6/543	0.509
$W^+ \rightarrow e^+ \nu$	416.3/546	0.999
$W^- \rightarrow \mu^- \bar{\nu}$	545.4/542	0.451
$W^+ \rightarrow \mu^+ \nu$	543.5/548	0.546

Table 9.1: Summary of χ^2 , ndf , and p-values for compatibility between the number of observed events and the number of expected events after the fit for all channels as a function of p_{T}^{ν} .

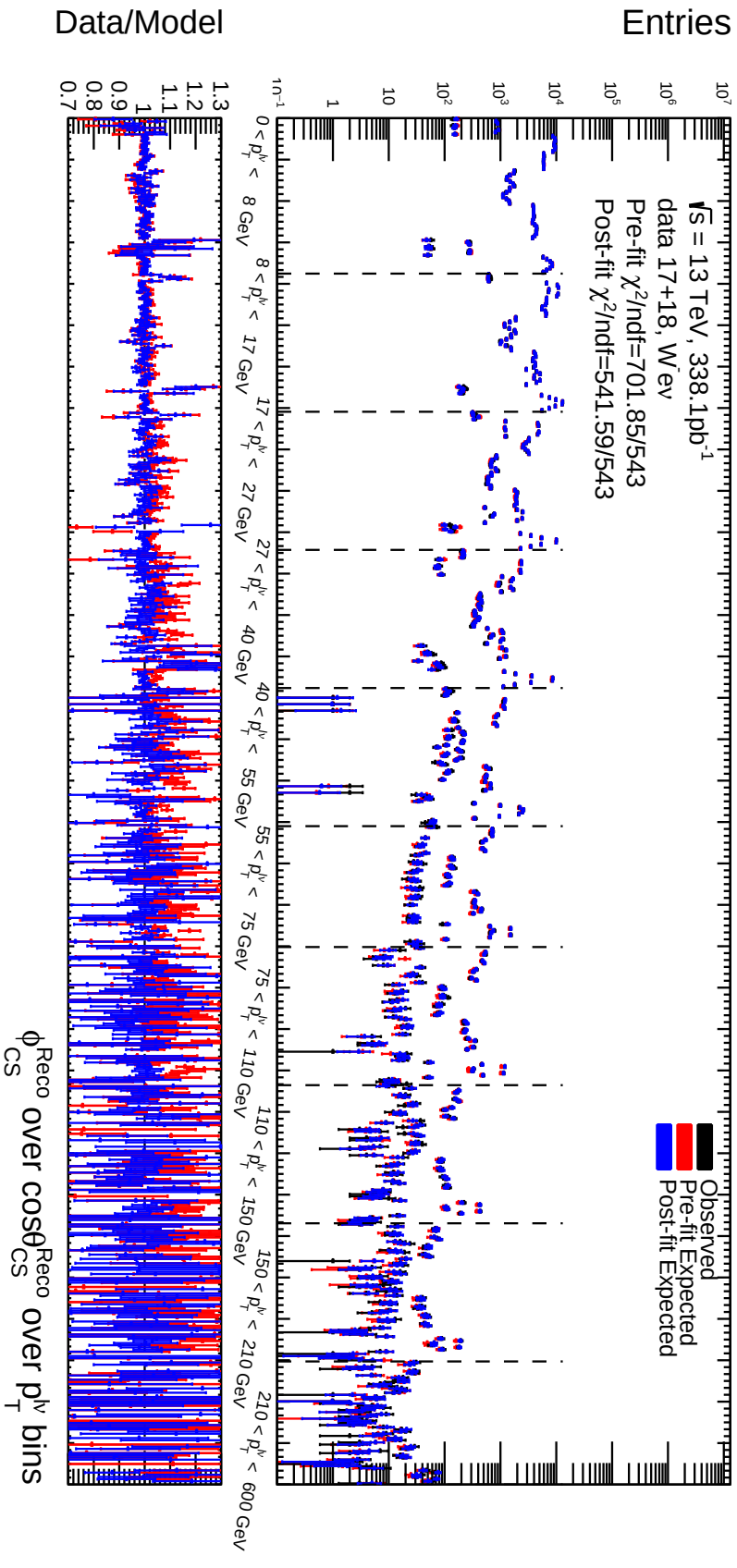
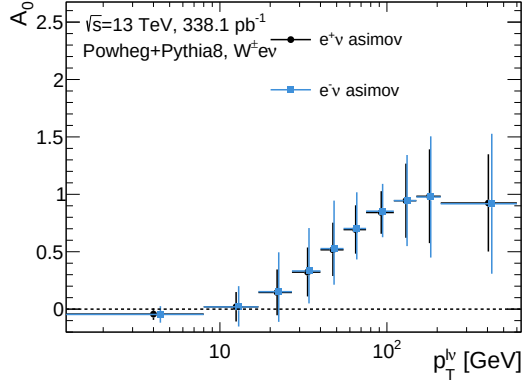


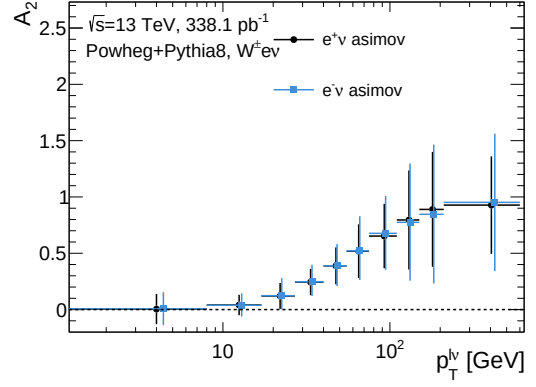
Figure 9.2: The unfolded three-dimensional ($\cos \theta_{CS}$, ϕ_{CS} , $p_T^{l\nu}$) distribution unwrapped into a one-dimensional distribution for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for each $p_T^{l\nu}$ bin separated by vertical dashed lines and labelled underneath the graph. The measured data, N_{obs} , is denoted by black points, the data-driven MJ background estimate, MC signal, and MC background are combined into N_{exp} and normalized to the data luminosity are denoted by the red points pre-fit and post-fit by the blue points. The ratio of N_{obs} over N_{exp} for both pre- and post-fit is displayed on the bottom plot showing the agreement improves through the NLL minimization.

Figure 9.2 is similar to the kinematic distribution plots of the three-dimensional phase space in Figure 8.12, where the background processes are combined with the signal process to form N_{exp} . In this figure the measured data – N_{obs} – is denoted by the black points, the combined model pre-fit is denoted by red points, and the combined model post-fit is denoted by the blue points. This is the three-dimensional phase space unfolded into a one-dimensional distribution, as previously described in Section 8.7, where each $p_T^{l\nu}$ bin is shown separated by vertical dashed lines and labelled under the main graph to denote the range. The ratio of the N_{obs} over N_{exp} for pre- and post-fit is shown on the bottom, which shows the agreement is better post-fit. Individual $(\cos \theta_{\text{CS}}, \phi_{\text{CS}}, p_T^{l\nu})$ bin values are difficult to parse in the ratio plot due to the larger number of them but the important aspect is the better agreement post-fit.

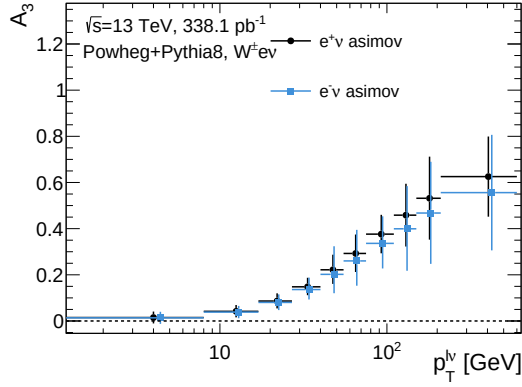
In addition to testing each channel to ensure that it is properly modelled by comparing it with the expected values, it is possible to see if there is sensitivity to the difference between W^+ and W^- . The coefficients A_3 and A_4 should theoretically differ for W^+ and W^- , as seen in Figure 9.3c) and d) which shows the expected values for A_3 and A_4 using Asimov data. Meanwhile, Figure 9.3a) and b) show that there is not an expected difference for A_0 and A_2 . This W boson charge difference for the extracted coefficients was tested using data and was found to have enough statistical sensitivity to measure the differences, as seen in Figure 9.4a) and b). This will be tested again after channels are combined, which will reduce the dominant statistical uncertainty compared to an individual measured channel. Because protons have two valence up quarks and only one valence down quark, the W^+ is produced in a larger abundance than the W^- . This is because the W^+ is primarily produced from a valence up quark and a sea quark, while the W^- is primarily produced from a valence down quark and a sea quark. The difference in the expected production differential cross-section for W^+ and W^- bosons as a function of $p_T^{l\nu}$ can be seen in the Figure 9.3e). The measured differential cross-section, shown in Figure 9.4c), also displays the difference in the production of W^+ and W^- bosons as a function of $p_T^{l\nu}$.



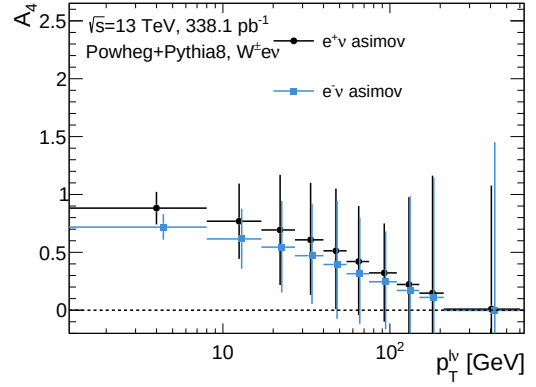
(a) A_0



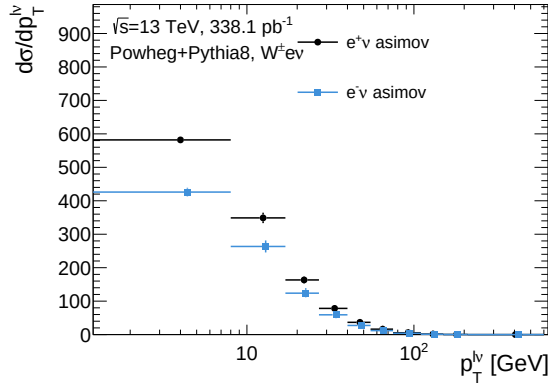
(b) A_2



(c) A_3

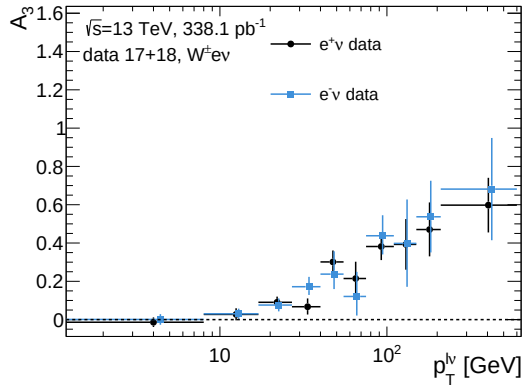


(d) A_4

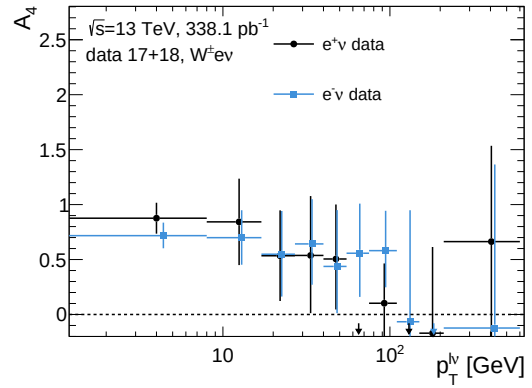


(e) Differential cross-section

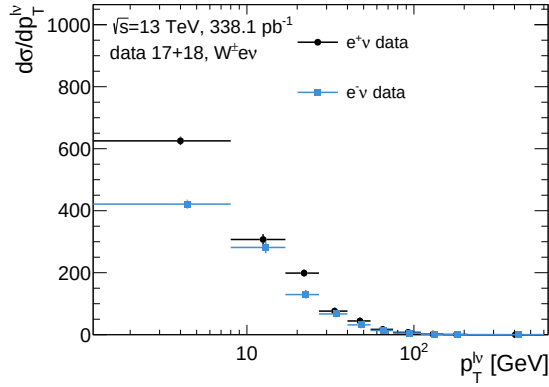
Figure 9.3: Comparison of the expected coefficients from Asimov data for the four major coefficients A_0 , A_2 , A_3 , A_4 and the differential cross-section as a function of $p_T^{l\nu}$ to show the expected differences between $W^+ \rightarrow e^+ \nu$ (black) and $W^- \rightarrow e^- \bar{\nu}$ (blue). The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



(a) A_3



(b) A_4



(c) Differential cross-section

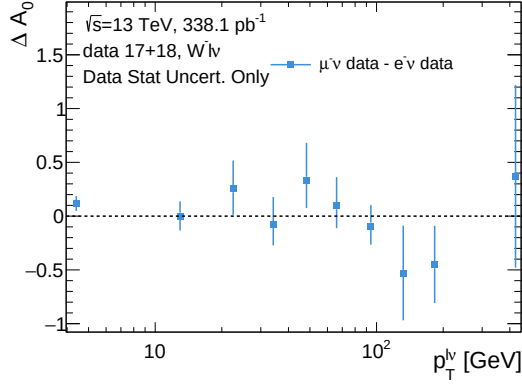
Figure 9.4: Comparison of the measured coefficients A_3 , A_4 , and the differential cross-section between $W^+ \rightarrow e^+\nu$ (black) and $W^- \rightarrow e^-\bar{\nu}$ (blue) as a function of $p_T^{l\nu}$ for data. These expected differences for a) A_3 and b) A_4 do not have enough statistical sensitivity to measure, while the difference c) for the differential cross-section can be seen. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.

9.2 Channel Compatibility

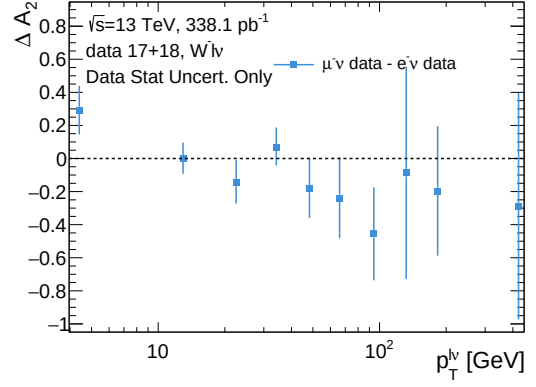
Before the individual channels can be combined, it is important to test the channel compatibility in terms of the angular coefficients and differential cross-section as a function of $p_T^{\ell\nu}$. This is done by taking the fit results of the individual channels and calculating the difference $\Delta A_{ij} = A_{ij}^{\mu\nu} - A_{ij}^{e\nu}$ for the i th coefficient in the j th observable bin. The expectation is that the differences are compatible with zero within their uncertainty. The uncertainty, $\sigma_{\Delta A_{ij}}$, is the sum in quadrature of just the data statistical uncertainties of the individual channels. The statistical uncertainty is the dominant source by an order of magnitude, as will be discussed in Chapter 10, so systematic uncertainties are ignored here. This is because the correlation of systematic uncertainties between channels is not known when fitting individual channels and cannot be expected to be completely decorrelated. Examples between $W^- \rightarrow \mu^- \bar{\nu}$ and $W^- \rightarrow e^- \bar{\nu}$ can be seen in Figure 9.5, which shows ΔA_{ij} for the differential cross-section and coefficients A_0 , A_2 , A_3 , and A_4 . This shows that there is good compatibility between channels in most cases and can be combined. A χ^2 can be used to statistically test if they are compatible, and approximated with:

$$\chi^2 = \sum_{i,j} \left(\frac{\Delta A_{ij}}{\sigma_{\Delta A_{ij}}} \right)^2. \quad (9.3)$$

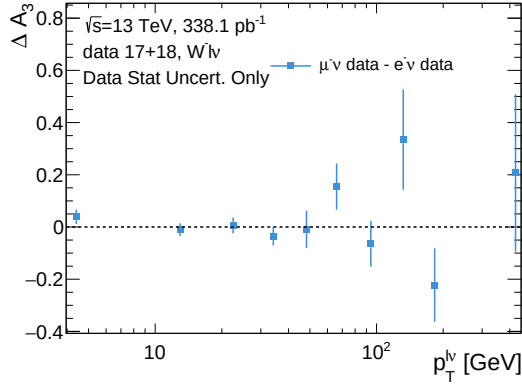
The χ^2 calculation also includes the coefficients where there is no sensitivity to measure, as they are also included in the LH fit for the individual channels. Because there are no fit parameters in this χ^2 to penalize the number of degrees of freedom here it will be penalized by nine for the number of A_i and one for the differential cross-section. So, the number of degrees of freedom will be calculated by $ndf = N_{A_{ij}} - N_{A_i} - 1$, which is $ndf = 81$. The χ^2 for the compatibility between $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels is 64.67, resulting in a p-value of 0.908. The χ^2 between $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$ channels is 60.99 with a p-value of 0.953. Showing good agreement for both W^- and W^+ .



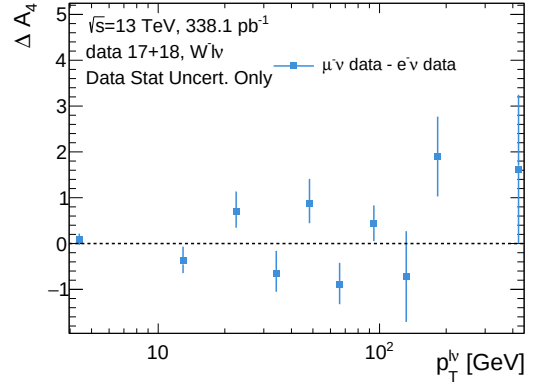
(a) A_0



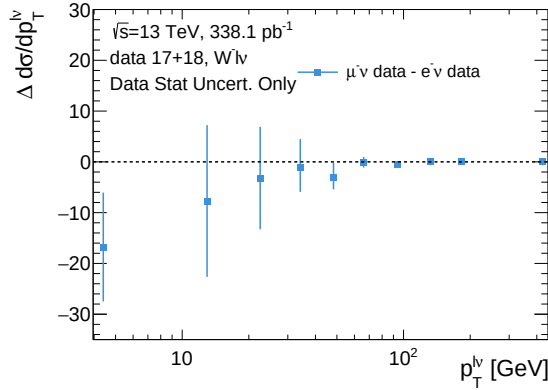
(b) A_2



(c) A_3



(d) A_4



(e) Differential cross-section

Figure 9.5: A test of the compatibility of the $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels as a function of $p_T^{l\nu}$ by taking the difference of the final fit results, denoted by the blue squares. The expected difference is zero and this shows there is good agreement between the channels. Here, only the data statistical uncertainties are used because of unknown correlations of systematic uncertainties between channels.

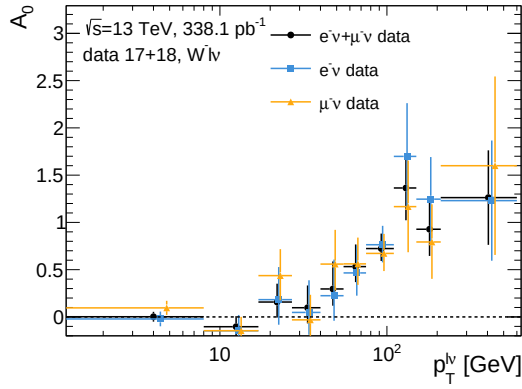
9.3 Channel Combination

In the previous section, it was shown that the electron and muon channels are statistically compatible, which means they can be combined. Combining compatible channels reduces the data statistical uncertainty by a factor of $\sim \sqrt{2}$ compared to measuring individual channels. The individual LHs are combined through a simple LH multiplication, where only one copy of common nuisance parameter (NP) terms are kept for each source of systematic uncertainty, as described in Section 7.4. Each LH can be decomposed into two types of terms: terms containing the observed data, denoted $\mathcal{L}_{\text{data}}$, and the auxiliary terms that constrain the NPs, θ , denoted $\mathcal{A}_i(\theta_i)$. An example of these NPs is the variation on the normalization of the boson background, which is allowed to vary the boson cross-section by 5% either up or down. All NPs are discussed in more detail in Chapter 10, covering the sources of systematic uncertainties considered in these measurements. The combined LH, $\mathcal{L}_{\text{cb}}$, can be written as:

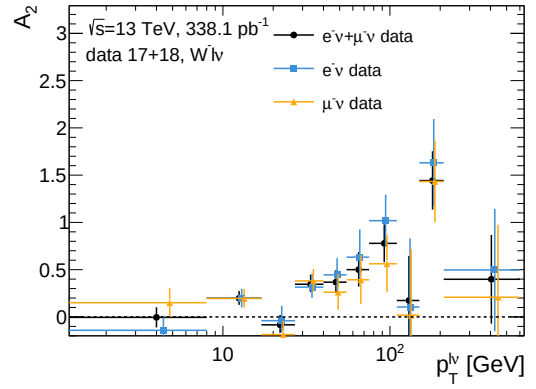
$$\mathcal{L}_{\text{cb}} = \left\{ \prod_i^{\text{channels}} \mathcal{L}_{\text{data}}(A, \theta | N_i) \right\} \left\{ \prod_i^M \mathcal{A}_i(\theta_i) \right\}. \quad (9.4)$$

The combined results for “ $e^- \nu + \mu^- \nu$ ” as a function of $p_{\text{T}}^{l\nu}$ can be found in black in Figure 9.6. In the same figure, the individual $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels are presented in blue and yellow, respectively. The overall agreement between the individual measurements can be seen by comparing the measured values, which are used to calculate the ΔA_{ij} measured values in Section 9.2.

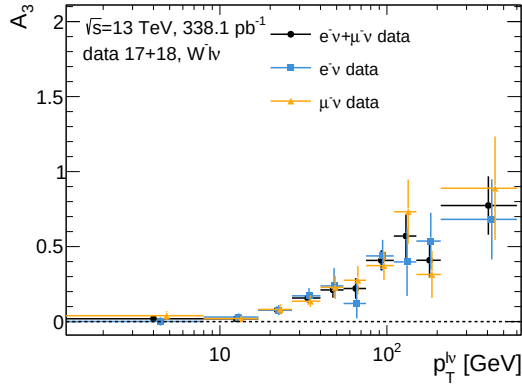
In most cases, the combined value is between the two individual channels, which highlights that most fluctuations are statistical in nature. For some bins, such as those seen at low $p_{\text{T}}^{l\nu}$ for A_0 in Figure 9.6a), the increased statistical power of combining two channels allows for there to be enough events in individual bins to not be as susceptible to statistical fluctuations. Thus, the observed distribution is closer to the expected distribution. As a reminder from Chapter 7, this is because the coefficients are not what are measured directly. Instead, they are extracted from the measured distributions in Figure 9.2.



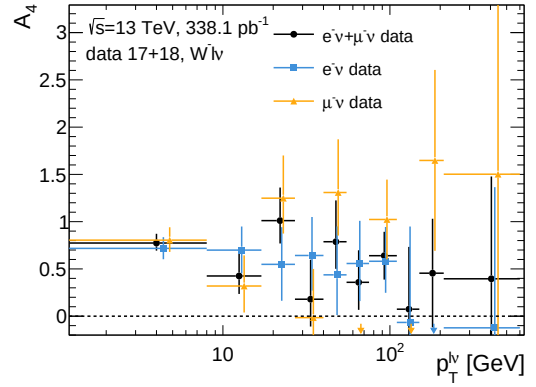
(a) A_0



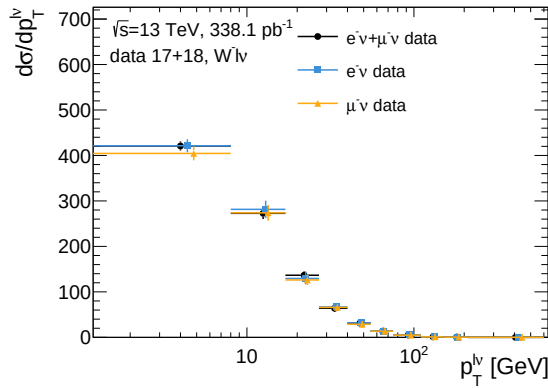
(b) A_2



(c) A_3



(d) A_4



(e) Differential cross-section

Figure 9.6: The measured values of four major coefficients A_0 , A_2 , A_3 , A_4 and the differential cross-section for the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels as a function of $p_T^{l\nu}$ can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.

These measured distributions are used with the folding functions, described in Section 7.3, to extract the “truth” level angular coefficients required to result in the angular distribution in Figure 9.2. By increasing the statistical power, the expected and observed distributions will have less significant statistical fluctuations in their $(\cos \theta_{\text{CS}}, \phi_{\text{CS}})$ distributions and the extracted coefficients will be closer to the expected values. This is more obvious for A_0 and A_4 , which shows that the $\cos \theta_{\text{CS}}$ distribution benefits the most from this increase in statistical sensitivity. From the combined values, it can be seen that they are more smoothly distributed and that the total uncertainties are reduced as expected.

Despite the coefficients for both W^+ and W^- bosons being in agreement for individual channels, it is not possible to combine them. This is primarily driven by the lack of theoretical motivation because A_3 and A_4 are expected to be different, as seen in Figure 9.3c) and d). In contrast to combining the electron and muons channels, where there is not an expected difference between lepton channels because of lepton flavour universality. The second reason is less apparent and has to do with the difference in $\cos \theta_{\text{CS}}$ and ϕ_{CS} shapes for W^+ and W^- . These angular distributions have a different shape for W^+ and W^- , as seen in Figure 9.7, which shows the pre-fit $\cos \theta_{\text{CS}}$ and ϕ_{CS} distributions for $W^+ \rightarrow e^+ \nu$ and $W^- \rightarrow e^- \bar{\nu}$. By combining W^+ and W^- , most of the shape information would be lost. This shape information is critical for extracting the Drell–Yan angular coefficients. For the coefficients that are theoretically expected to be the same, such as A_0 and A_2 , as seen in Figure 9.3a) and b), these could potentially be combined post-fit by averaging them. However, this would require properly correlating uncertainties and is beyond the scope for this thesis.

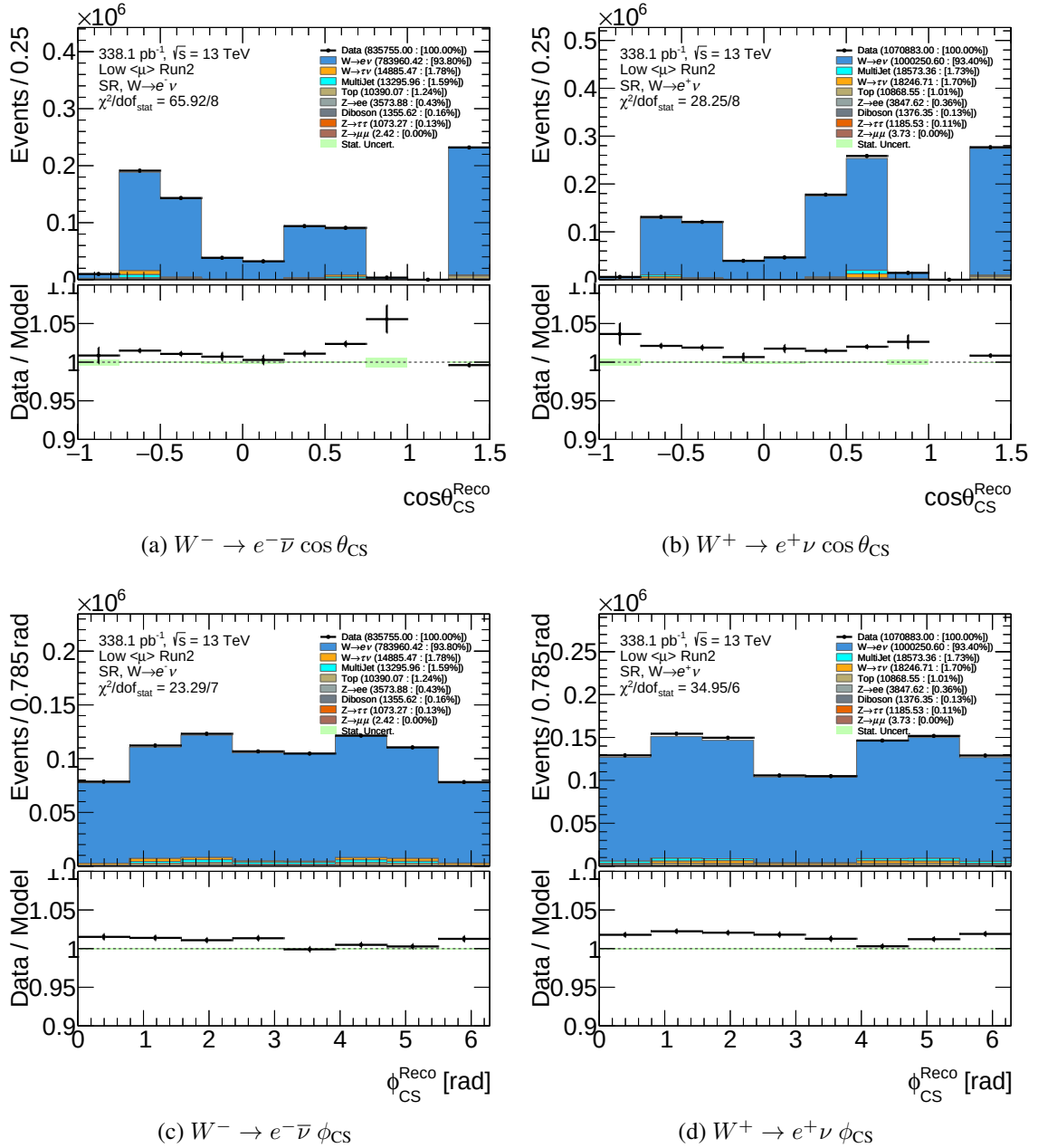


Figure 9.7: The integrated CS angular distributions of $\cos \theta_{\text{CS}}^{\text{Reco}}$ for a) $W^- \rightarrow e^- \bar{\nu}$ and b) $W^+ \rightarrow e^+ \nu$ and $\phi_{\text{CS}}^{\text{Reco}}$ for c) $W^- \rightarrow e^- \bar{\nu}$ and d) $W^+ \rightarrow e^+ \nu$ for events that pass selection criteria listed in Table 8.3 shown on a linear scale to show the difference in shape. The black points represent the measured data and statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

Again, the difference between W^+ and W^- is shown with the combined channels for A_3 and A_4 , as seen in Figure 9.8a) and b). Even with the improved statistical sensitivity of combining lepton flavour channels, there is not sufficient statistical sensitivity to observe this expected difference. For A_3 this difference is expected to increase as a function of $p_T^{l\nu}$, where A_4 should decrease as a function of $p_T^{l\nu}$. A difference in the measured W^+ and W^- differential cross-sections, as seen in Figure 9.8c), is still observed as expected.

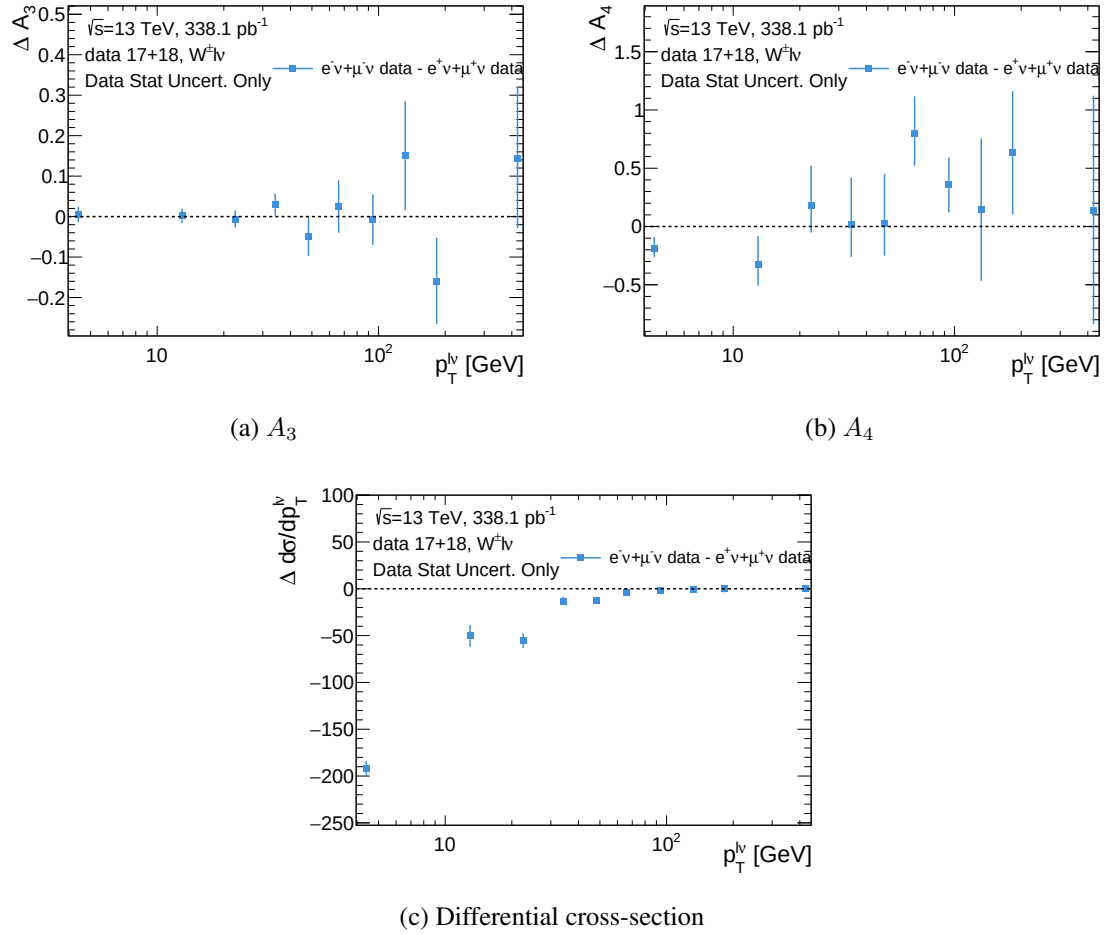
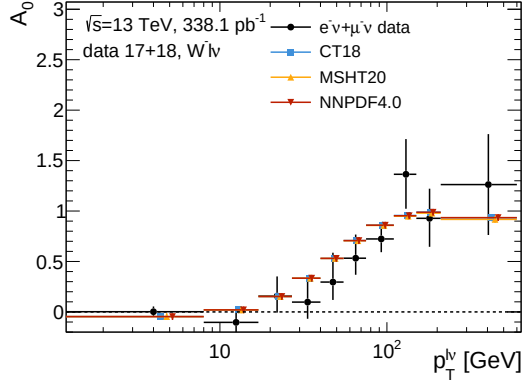


Figure 9.8: The difference between W^- and W^+ of the measured coefficients A_3 , A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for data is denoted by blue squares. The expected differences for a) A_3 and b) A_4 are not observed, while the difference for c) the differential cross-section can be observed.

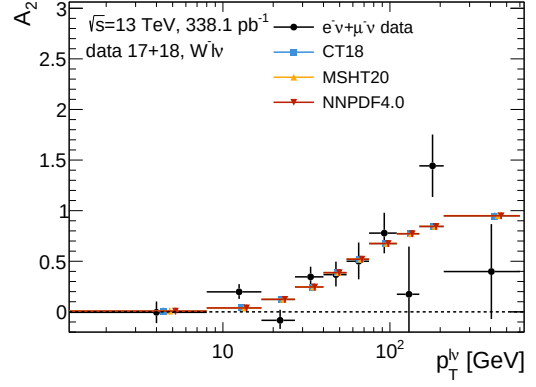
9.4 Comparison to PDF Sets

The final consideration presented is comparing the measurements of the combined values to different PDF set predictions. This allows evaluating the impact of the lack of knowledge of PDFs on the combined measurements. These PDF set predictions are calculated by reweighting the MC simulation using the LHAPDF tool [113] to three different modern PDF sets: CT18 [114], MSHT20 [115], and NNPDF4.0 [116]. With the reweighted MC, Asimov data sets are used to get the theoretical values. Only the uncertainties on the PDF sets are used in the comparison to the measured values and are calculated using variations on the reweighting to the PDF sets. On the scale of the measured values, there is no significant difference between the PDF set predictions and the measured values to preferentially agree.

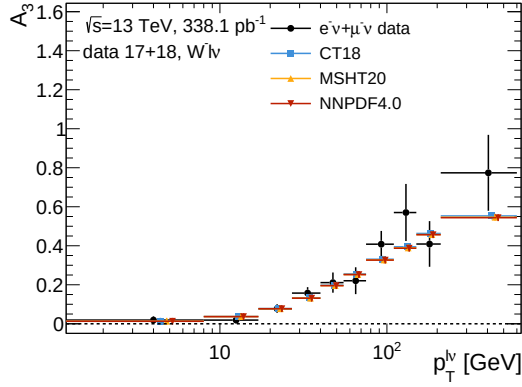
The combined values for W^- as a function of $p_T^{l\nu}$ compared to the different PDF set predictions can be seen in Figure 9.9. Here, there is good agreement with the measured values in black to the PDF set predictions CT18, MSHT20, and NNPDF4.0 in blue, yellow, and red, respectively. The compatibility between the measured values and the different PDF set predictions is summarized in Table 9.2, where a χ^2 is calculated using Equation 9.3. This χ^2 is not the goodness of fit describing how well the observed number of events for the data set matches the expected number of events to determine the Drell–Yan angular coefficients and the differential cross-section. Instead, it is just used as an expression for describing how compatible the measured coefficients are to the expected. The goodness of fit and the final measurements for the combined channels, for both W^+ and W^- , will be presented in Chapter 11.



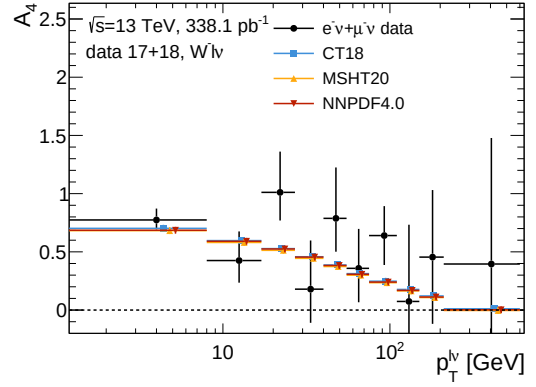
(a) A_0



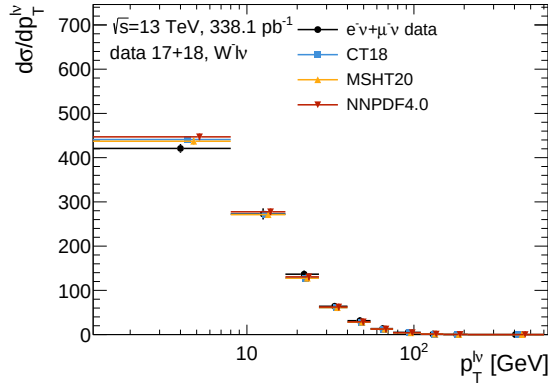
(b) A_2



(c) A_3



(d) A_4



(e) Differential cross-section

Figure 9.9: The values of the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels as a function of $p_T^{l\nu}$ can be seen in black, are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.

Channel	$W^-e^-\nu + \mu^-\nu$		$W^+e^+\nu + \mu^+\nu$	
	χ^2/ndf	p-value	χ^2/ndf	p-value
data-CT10	37.72/81	0.999990	39.69/81	0.999969
data-CT18	37.67/81	0.999990	39.80/81	0.999967
data-MSHT20	38.18/81	0.999987	39.81/81	0.999967
data-NNPDF4.0	38.16/81	0.999987	39.95/81	0.999964

Table 9.2: Summary of χ^2 , ndf , and p-values for compatibility between measured Drell–Yan angular coefficients and the different PDF set predictions in Figure 9.9 and CT10 – the Asimov data set without any PDF reweighting applied. The p-values demonstrate that the measured values are in good compatibility with all PDF set predictions.

In Table 9.2 there is comparison to the three PDF set predictions in Figure 9.9 and CT10, which is the Asimov data set without any PDF reweighting applied. The measured values are in good agreement with all four predictions with little difference in the χ^2 values between them, meaning there is not much preference. The agreement for all of the coefficients is good in the low $p_T^{l\nu}$ range as seen in Figure 9.9, which is the most important region for input to a future W boson mass measurement. For the W mass at $\sqrt{s} = 7$ TeV [16], there was a selection on the hadronic recoil of $u_T < 30$ GeV, where the hadronic recoil is approximately equal to $p_T^{l\nu}$. In particular, the differential cross-section and A_4 have the most impact on future W boson mass measurements, which are in good agreement to the PDF set values and will be discussed more in Section 11.3.

There is good agreement between the measured Drell–Yan angular coefficients and the differential cross-section to the different PDF set predictions. Despite not being sensitive to the different PDF set predictions directly, these measurements can still provide valuable input to future PDF sets that use many different experimental measurements as input. With these being the first measurements of A_2 and A_3 as a function of $p_T^{l\nu}$ at the LHC and the first measurements of A_0 and A_4 as a function of $p_T^{l\nu}$ ever, these will provide additional information used to create the PDF sets and improve the knowledge of PDFs, including their uncertainties.

9.5 Summary

This chapter described the path to measure the channels in individual channels, measure the compatibility between channels, and the combination of channels. The next chapter will focus on the sources of uncertainty that are considered for these measurements. Then, Chapter 11 will present the final measurements, a decomposition of their uncertainties, and the potential impact on future W boson mass measurements.

CHAPTER 10

STATISTICAL AND SYSTEMATIC UNCERTAINTIES ON THE MEASUREMENT OF THE DRELL-YAN ANGULAR COEFFICIENTS AND DIFFERENTIAL CROSS-SECTION

This chapter will focus on the different sources of uncertainty that are considered for the measurement of the Drell–Yan angular coefficients and the differential cross-section presented in Chapter 9. The uncertainties for these measurements are dominated by the sources of statistical uncertainties in the data, which are described in Section 10.1. The other sources of uncertainties that need to be considered are the experimental and theoretical uncertainties discussed in Sections 10.2.1 and 10.2.2, respectively. These systematic uncertainties are an order of magnitude smaller than the data statistical uncertainties. A summary of all sources of uncertainties for the measured values as a function of $p_T^{l\nu}$ will be shown in Section 10.3 for the $W^- \rightarrow e^- \bar{\nu}$ channel. The remaining channels can be found in Appendix G. These systematic uncertainties are partially constrained during the NLL fit, which will be discussed Section 10.4.

Most sources of systematic uncertainty are taken from sources within ATLAS, such as the uncertainties on the electron energy calibration, and are then implemented in these measurements and referenced accordingly. Other sources of uncertainty, such as the data-driven background modelling, were evaluated for the purpose of this analysis and their implementation will be discussed. For the discussion of data statistical uncertainties, only one example will be shown for illustrative purposes. For this thesis, the absolute uncertainty of coefficients A_0 to A_4 will be given, while the relative uncertainty will be given for the differential cross-section. This is because the angular coefficients often approach zero, so a relative uncertainty does not provide a meaningful insight into the uncertainty and an absolute uncertainty allows for a more legible scale.

10.1 Data Statistical Uncertainties

This section will cover the data statistical uncertainties that need to be considered for these measurements. For a coefficient i in the j th $p_T^{l\nu}$ bin, A_{ij} , the total data statistical uncertainty of the coefficient is $\sigma_{A_{ij}}$. It is the addition in quadrature of four major categories of uncertainty: uncorrelated statistical uncertainty, correlated migrational uncertainty between $p_T^{l\nu}$ bins, uncorrelated shape uncertainty, and cutoff uncertainty. The first three are depicted in Figure 10.1 for $A_{0,0}$. The uncorrelated statistical uncertainty is the uncertainty due to the parameter of interest alone and depends on the available statistics of the data. This uncertainty is subdominant for most $p_T^{l\nu}$ bins; however, it contributes more at large $p_T^{l\nu}$ when there are more limited data statistics. The migration uncertainty is the dominant source of statistical uncertainty and is discussed in more detail in Section 10.1.1. The shape uncertainty is the correlated statistical uncertainty between different coefficients within the same $p_T^{l\nu}$ bin. Lastly, the cutoff uncertainty is associated to a choice in the number of migration terms kept and will be discussed in more detail in Section 10.1.2. An example of the absolute data statistical uncertainty and its components can be seen in Figure 10.2 for A_0 as a function of $p_T^{l\nu}$.

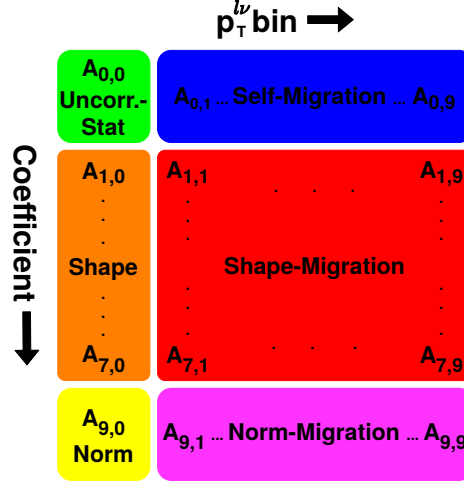


Figure 10.1: Depiction of the sources of data statistical uncertainty for A_0 in $p_T^{l\nu}$ bin 0, denoted by $A_{0,0}$, corresponding to $0 < p_T^{l\nu} < 8$ GeV. The uncertainty on the parameter $A_{0,0}$ is the uncorrelated statistical uncertainty. The shape uncertainty is the non-migrational correlated uncertainty between different angular coefficients in the same $p_T^{l\nu}$ bin. The correlated migration uncertainty is comprised of three types: self, shape, and norm. Self-migration is for the migration between $p_T^{l\nu}$ bins for the same coefficient, shape-migration is the migration between angular coefficients and $p_T^{l\nu}$ bins, and norm-migration is the self-migration for the differential cross-section, denoted by A_9 . Based on an image from [6].

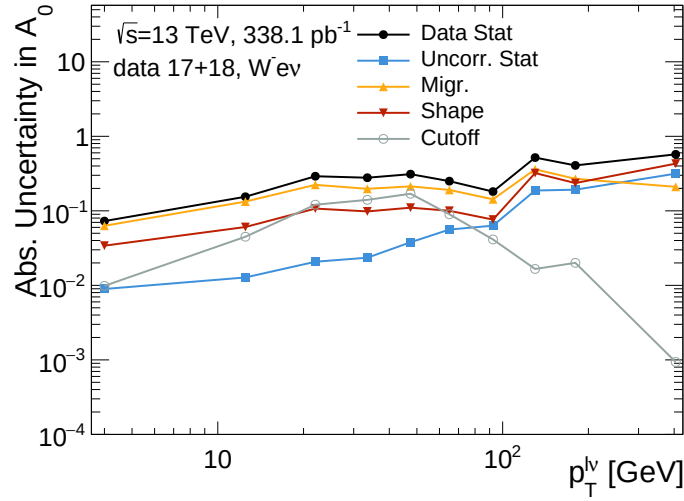


Figure 10.2: Decomposition of the total data statistical uncertainty as a function of $p_T^{l\nu}$ for the absolute uncertainty of A_0 . The black circles are the total data statistical uncertainty, which is the addition in quadrature of the other components. The blue squares correspond to the uncertainty due to the parameter of interest. The yellow upright triangles are the migrational uncertainty correlated between $p_T^{l\nu}$ bins. The red downward triangles are the uncorrelated shape uncertainty for different coefficients within the same bin. Lastly, the grey open circles are the cutoff uncertainty based on the choice of number of migration terms.

10.1.1 Migration Uncertainty

The main source of statistical uncertainty for a coefficient is the correlated migrational uncertainty, which has three components: self, shape, and norm, as seen in Figure 10.1. This figure shows an example for A_0 in $p_T^{l\nu}$ bin 0, corresponding to $0 < p_T^{l\nu} < 8$ GeV. The columns correspond to each coefficient, while the rows correspond to each $p_T^{l\nu}$ bin. The parameters in the first row and second column are the same coefficient but in different $p_T^{l\nu}$ bins. The migration of events between $p_T^{l\nu}$ bins for the same coefficient is called self-migration and is modelled by MC when building the folding templates. These events are when the reconstructed $p_T^{l\nu}$ is in a different bin than the expected $p_T^{W,\text{Truth}}$ bin but remain in the same $(\cos \theta_{CS}, \phi_{CS})$ bin. When the event is reconstructed in a different $p_T^{l\nu}$ bin and a different $(\cos \theta_{CS}, \phi_{CS})$ bin, it is associated to a different coefficient and called shape-migration. Lastly, the differential cross-section, acts as the normalization for the Drell–Yan angular coefficients in the form of the denominator in Equation 7.2. The self-migration for the normalization between $p_T^{l\nu}$ bins, is labeled the norm-migration.

An example of the migration uncertainties for A_0 can be seen in Figure 10.3. For all angular coefficients, the dominant source of uncertainty is the self-migration. Meanwhile, the dominant source for the differential cross-section is the shape-migration for W^- channels but shape and norm-migration are similar for W^+ channels. The main source of migration is from the choice of neutrino p_z^ν solution from the quadratic equation, which affects the measured $p_T^{l\nu}$ and $\cos \theta_{CS}$. This neutrino solution choice is discussed in more detail in Section 8.4.1 and Appendix F. The coefficients as a function of $p_T^{l\nu}$ are mostly sensitive to events where $p_T^{l\nu}$ falls outside of the $p_T^{W,\text{Truth}}$ bin, causing it to have larger self-migration.

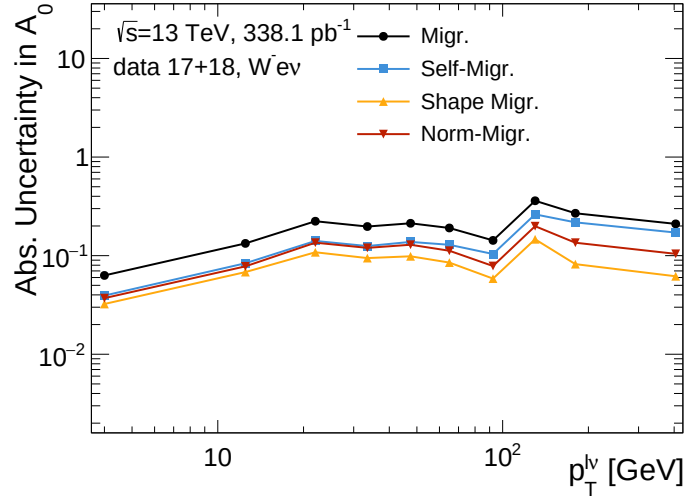


Figure 10.3: Decomposition of the total migrational uncertainty as a function of $p_T^{l\nu}$ for the absolute uncertainty of A_0 . The black circles correspond to the total migrational uncertainty, which is the addition in quadrature of the other components and is the same as the yellow upright triangles in Figure 10.2. The blue squares are the migration between $p_T^{l\nu}$ bins for the same coefficient, the yellow upright triangles are the migration between coefficients and $p_T^{l\nu}$ bins, and the red downwards triangles are the migration between $p_T^{l\nu}$ bins for the norm.

10.1.2 Cutoff Uncertainty

The cutoff uncertainty is associated to the choice in the number of migration terms used in the template folding procedure. MC simulation is used to construct the templates used in the LH fit and allows for modelling the amount of expected migration. For a reconstructed $p_T^{l\nu}$ event, it will have an expected $p_T^{W, \text{Truth}}$ bin and is most likely to be in a neighbouring bin. However, the measured $p_T^{l\nu}$ can be quite different than the expected $p_T^{W, \text{Truth}}$ and these are still recorded. For a folding template to be considered in the LH fit, it needs to have a minimum number of events modelled to pass a threshold, called the cutoff. This uncertainty is to enumerate any potential bias in the choice of cutoff.

To determine an appropriate cutoff threshold, a randomly generated data set using MC was created by varying the expected number of events within its uncertainty. The randomly generated data set was used in a series of LH fits with varying cutoff levels applied to the templates. Each LH fit has an associated χ^2 value describing how well the randomly

generated data set matches the expected. The χ^2 value changes as a function of cutoff and a point was determined when the changes between χ^2 values for different choices of cutoff became minimal. This point was found to be with a cutoff value of 0.02, which means that folding templates for a $p_T^{l\nu}$ bin need to contain at least 2% of the expected $p_T^{W,\text{Truth}}$ events. To assign an uncertainty to this choice, the results from the randomly generated data set were used. Here, the cutoff was varied by 0.01 and the results from those LH fits were used. This was done in two parts, by considering the effect on the value and the effect on the migrational uncertainty. For the first component, the absolute difference between the nominal cutoff value of 0.02 and when varied up or down by 0.01 is taken as the uncertainty on the value for each coefficient and the differential cross-section:

$$\sigma_{\text{cutoff, value}}^{\pm} = |A_{ij}^{0.02} - A_{ij}^{\pm 0.01}|. \quad (10.1)$$

Here, $A_{ij}^{0.02}$ is the nominal value of coefficient i in the j th bin for a cutoff value of 0.02 and $A_{ij}^{\pm 0.01}$ is the value of the same coefficient with a cutoff value of 0.01 or 0.03. The uncertainty is asymmetric depending on whether the varied cutoff was 0.01 or 0.03. Similarly, an uncertainty is assigned for how migrational uncertainty changes. This is done by taking the absolute difference between the migrational uncertainty for the choice in cutoff values:

$$\sigma_{\text{cutoff, migr}}^{\pm} = |\sigma_{\text{migr}}^{0.02} - \sigma_{\text{migr}}^{\pm 0.01}|. \quad (10.2)$$

Here, $\sigma_{\text{migr}}^{0.02}$ is the total migrational uncertainty of coefficient i in the j th bin for a cutoff value of 0.02 and $\sigma_{\text{migr}}^{\pm 0.01}$ is the total migrational uncertainty of the same coefficient with a cutoff value of 0.01 or 0.03.

Both of these sources of cutoff uncertainty are added in quadrature and expected to be uncorrelated, resulting in the total cutoff uncertainty denoted by the grey open circles in Figure 10.2. This uncertainty is derived individually for each lepton flavour channel – $e^-\bar{\nu}$, $e^+\nu$, $\mu^-\bar{\nu}$, $\mu^+\nu$ – with their own randomly generated data sets and added to the total

data statistical uncertainty for the measured data post-fit. It is a subdominant source of statistical uncertainty compared to the total migrational uncertainty for all $p_T^{l\nu}$ bins.

10.2 Systematic Uncertainties

As discussed in Section 7.4 with the LH function, there are two groups of systematics uncertainties constrained by either a Gaussian or a Poisson probability distribution function. The nuisance parameters (NPs) constrained by the Poisson probability distribution function are the statistical uncertainties on the effective number of entries in each $\cos \theta_{CS}$ and ϕ_{CS} bin for the MC used to generate the templates in Equation 7.6. This is one of the dominant sources of systematic uncertainty for this measurement, as seen in the total uncertainty decomposition in Section 10.3. It is labeled “MC Stat” and is denoted by the blue squares. The NPs constrained by Gaussian probability distribution functions are comprised of two groups based on the source of uncertainty: experimental and theoretical – discussed in Sections 10.2.1 and 10.2.2, respectively. All of the systematic uncertainties, shown in Figure 10.4, will be discussed and then will be shown in Section 10.3 for the remaining larger coefficients and the differential cross-section. In addition to the sources of systematic uncertainty a comparison to the total statistical uncertainty will be included. Section 10.3 will show the uncertainties for the larger coefficients A_0 to A_4 and for the differential cross-section as a function of $p_T^{l\nu}$ for the $W^- \rightarrow e^- \bar{\nu}$ channel are shown.

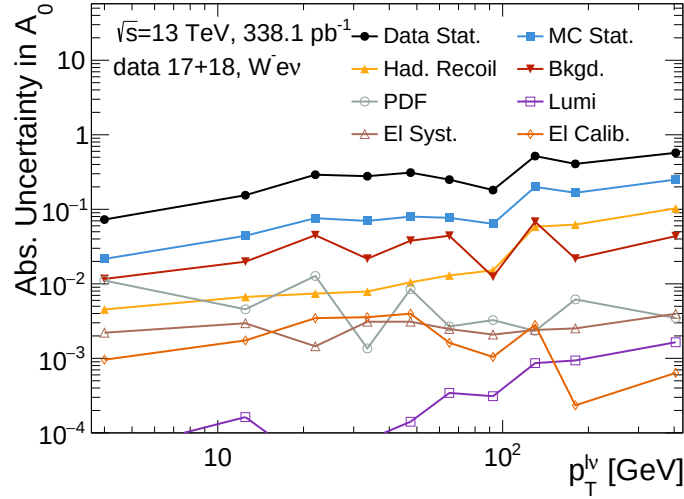


Figure 10.4: A breakdown of all sources of measured uncertainty for the angular coefficient A_0 as a function of $p_T^{l\nu}$ for Figure 9.1 for the $W^- \rightarrow e^- \bar{\nu}$ channel. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.

10.2.1 Experimental Uncertainties

The majority of the systematic uncertainties considered are experimental in nature. They include the uncertainties on hadronic recoil resolution and calibration, the background modelling, the luminosity, and for the calibration and efficiency of leptons. Other sources of systematic uncertainty have been considered [12] but are expected to be smaller and are not included. Each category of uncertainties with more than one NP will contain a table summarizing the NPs considered and what is varied. The experimental uncertainties affect only the reconstructed information for each event and will only impact the numerator of the templates in Equation 7.6.

Hadronic Recoil Systematics

The hadronic recoil calibration is one of the dominant sources of experimental systematic uncertainty in this analysis. This sensitivity is due to uncertainty in the reconstruction of the W boson, which is dependent on the neutrino reconstructed with the hadronic recoil, as discussed in Section 8.4. The uncertainty from the measurement of the hadronic recoil is decomposed into three major categories corresponding to the resolution, response, and the ΣE_T reweighting. A summary of the NPs used and their descriptions can be seen in Table 10.1, which is based on the hadronic recoil uncertainties measured by the ATLAS Collaboration in [12]. The combined uncertainty of these categories is labeled as “Had. Recoil”, denoted by the yellow upright triangles and can be seen in Figure 10.4 for the measure A_0 as a function of $p_T^{l\nu}$.

Nuisance Parameter Name	Number of NPs	Description
Response_Sys	1	Systematic uncertainty on the response correction.
Response_ExtSys	1	Systematic uncertainty on the response correction due to the extrapolation from Z to W .
Response_Stat0	15	Statistical uncertainty on eigenvector 1 of the covariance matrix for the fitted function to model the response correction.
Response_Stat1	15	Statistical uncertainty on eigenvector 2 of the covariance matrix for the fitted function to model the response correction.
Resolution_ExtSys	1	Systematic uncertainty on the resolution correction due to the extrapolation from Z to W .
Resolution_Stat0	18	Statistical uncertainty on eigenvector 1 of the covariance matrix for the fitted function to model the resolution correction.
Resolution_Stat1	18	Statistical uncertainty on eigenvector 2 of the covariance matrix for the fitted function to model the resolution correction.
SET_Sys	1	Systematic uncertainty on the ΣE_T reweighting.
SET_Stat	48	Statistical uncertainty on the ΣE_T reweighting for high values of u_T .

Table 10.1: Description of the channel independent hadronic recoil systematics NPs used in this analysis. Each NP will have an up and down variation with the exception of SET_Sys which a one-sided up variation.

Background Processes Systematics

The other dominant sources of experimental systematic uncertainty are from background modelling for MC and data-driven estimated sources. Each MC simulated background sample is grouped into one of three categories: boson, diboson, and top for a total of three nuisance parameters. The uncertainties for MC estimated samples are taken as a flat percentage based on the order of the calculation used to estimate their cross-sections. The uncertainties used from these calculations are conservative compared to the measured cross-sections cited with each sample. For single boson samples [12], such as $Z \rightarrow l^+l^-$, a 5% uncertainty is taken. Diboson samples [117, 118, 119], such as $WW \rightarrow \ell\nu\ell\nu$, have an uncertainty of 10% and top related processes [120, 121, 122, 123] range from 7 – 10%. Specifics for each sample can be found in Table E.1 in Appendix E. In addition to the uncertainty on the cross-section for the top background contribution, other MC samples are used as variation on $t\bar{t}$ modelling. These samples vary a modelling parameter, hdamp , that affects the high p_T radiation generated by POWHEG. In turn, this affects both the normalization and shape of the $t\bar{t}$ background contribution.

The data-driven multi-jet (MJ) background from Section 8.6 is the larger source of background modelling uncertainty. Multiple sources of uncertainty are considered to quantify the uncertainty of the extraction process. The first category is a flat 5% yield uncertainty analogous to the uncertainty on the cross-sections for MC estimated backgrounds. Next, is the uncertainty on the shape of the MJ templates, which comes from the difference between the shape in the different isolation slices discussed in Section 8.6. Lastly, are the uncertainties from the acceptance function used to correct the samples. Both MC and data-driven background uncertainties have their NPs summarized in Table 10.2. All sources of background uncertainty are combined into one category labeled “Bkgd.” and are denoted by the red downward triangles in Figure 10.4.

Nuisance Parameter Name	Number of NPs	Description
XSec_Boson	1	Flat 5% variation on the cross-section for all single boson samples.
XSec_Diboson	1	Flat 10% variation on the cross-section for all diboson samples.
XSec_Top	1	Flat variation between 7-10% on the cross-section for all top associated samples.
ttbar_PHer7_hdampp258p75_NORM	1	Variation on the number of events from $t\bar{t}$ POWHEG+HERWIG7 generated sample.
ttbar_PHer7_hdampp258p75_SHAPE	1	Variation on the shape of $t\bar{t}$ background from POWHEG+HERWIG7 generated sample.
ttbar_PP8_hdampp517p5_NORM	1	Variation on the number of events from $t\bar{t}$ POWHEG+PYTHIA8 generated sample with modelling parameter, hdamp, is varied to 517.5.
ttbar_PP8_hdampp517p5_SHAPE	1	Variation on the shape of $t\bar{t}$ background from POWHEG+PYTHIA8 generated sample with modelling parameter, hdamp, is varied to 517.5.

MJ_Acc_Closure	3	Difference between templates used to create acceptance correction functions. Divided into three kinematic regions with bounds $0 < p_T^{l\nu} < 17$ GeV, $17 < p_T^{l\nu} < 55$ GeV, and $55 < p_T^{l\nu} < 600$ GeV.
MJ_Shape	1	Uncertainty on MJ template shape correlated across all $p_T^{l\nu}$ bins.
MJ_Yield	3	Uncertainty on yield from intersection point. Divided into three kinematic regions with bounds $0 < p_T^{l\nu} < 17$ GeV, $17 < p_T^{l\nu} < 55$ GeV, and $55 < p_T^{l\nu} < 600$ GeV.

Table 10.2: Description of the background related systematics NPs used in this analysis. Each NP will have an up and down variation except for the $t\bar{t}$ hdamp variations, which are one-sided.

Luminosity Systematics

One of the smaller sources of systematic uncertainties is from the uncertainty on the measured luminosity for the data set [11]. This is an uncertainty of 0.92% that is applied as one NP correlated for all $p_T^{l\nu}$ bins. For the measured coefficients, this uncertainty is very small due the nature of the coefficients being ratios. The differential cross-section luminosity uncertainty has a larger contribution compared to the coefficients because it is not a ratio but it is still not a dominant source of uncertainty. The luminosity uncertainty is labeled “Lumi” and is denoted by the purple open squares in Figure 10.4.

Lepton Systematics

Electrons and muons are measured differently in the ATLAS detector, which means there are different methods of calibrating the momentum of the lepton. For electron channels, there are calibration uncertainties on the resolution and energy scale correction for the electron momentum [124]. Whereas, for muon channels, the calibration uncertainties are on the measurements made in the Inner Detector (ID) and Muon Spectrometer (MS) detectors and on the energy scale corrections [125]. The energy calibration uncertainties are labeled with “Electron Calib.” and denoted by the orange open diamonds in Figure 10.4. For similar uncertainty plots of muons channels, shown in Appendix G, these labels have “Electron” substituted with “Muon”. The NPs for the electron and muon energy calibrations are summarized in Tables 10.3 and 10.4, respectively.

The different leptons have different scale factors (SFs) that are applied and will have different associated uncertainties. Electrons have efficiency uncertainties for identification, isolation, reconstruction, and trigger [72]. The NPs for electron efficiencies are decomposed into 14 eigenvectors and lose some of their physical intuition of what is systematically varied. Muons have efficiency uncertainties for isolation, reconstruction, trigger, and track-to-vertex-association (TTVA) [67]. The efficiency uncertainties are combined and labeled with “Electron Syst.”, as denoted by brown open upright triangles in Figure 10.4. Similar to the calibration uncertainties, uncertainties for the muon channels shown in Appendix G, these labels have “Electron” substituted with “Muon”.

Nuisance Parameter Name	Number of NPs	Description
El_Reso	48	Systematic uncertainty on the resolution of electron momentum.
El_Scale	48	Systematic uncertainty on energy scale correction.
El_ID_Corr	14	Full correlation model systematic variation from [72].
El_ID_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.
El_Iso_Corr	14	Full correlation model systematic variation from [72].
El_Iso_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.
El_Reco_Corr	30	Full correlation model systematic variation from [72].
El_Reco_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.
El_Trig_Corr	14	Full correlation model systematic variation from [72].
El_Trig_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.

Table 10.3: Description of the electron channel only systematics NPs used in this analysis. Each NP has an up and down variation.

Nuisance Parameter Name	Number of NPs	Description
Mu_ID	1	Systematic uncertainty on calibration for muons using just Inner Detector information.
Mu_MS	1	Systematic uncertainty on calibration for muons using just Muon Spectrometer information.
Mu_Scale	1	Systematic uncertainty on energy scale correction.
Sagitta_Bias_Stat	48	Systematic uncertainty on charge dependency of momentum.
Mu_Iso_dRmuj	1	Variation on requirements for angular separation between probe muon and closest jet.
MU_Iso_eta	1	Variation of η requirements for the tag muon.
Mu_Iso_isoBkg	1	Variation of isolation requirements effect on background subtraction.
Mu_Iso_mass	1	Variation of invariant mass window required for tag and probe pair.
Mu_Iso_probeQ	1	Variation on identification requirements for the probe muon.
Mu_Iso_tagIso	1	Variation of isolation requirements for the tag muon.
Mu_Iso_tpdR	1	Variation on tag and probe angular separation.
Mu_Iso_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.

mu_Reco_BkgFraction	1	Systematic uncertainty on background fraction subtracted.
Mu_Reco_HighToLowExtrap	1	Systematic uncertainty of extrapolating SFs from high pileup events to low pileup.
Mu_Reco_LumiUncert	1	Systematic uncertainty on luminosity used to scale MC.
Mu_Reco_MatchingSym	1	Systematic uncertainty on if reconstructed muon is in the cone used to reconstruct it.
Mu_Reco_MCxsec	1	Systematic uncertainty on cross-section used to scale MC.
Mu_Reco_PtDependency	1	Difference in efficiency for $p_T > 200$ GeV, where statistics is more limited.
Mu_Reco_QCDTemplate	1	Systematic uncertainty on QCD background fit templates that mimic probe muons.
Mu_Reco_SuppressionScale	1	Variation on parameter used to estimate energy required to produce di-muon pair.
Mu_Reco_TruthSym	1	Systematic uncertainty on tag and probe efficiencies using truth MC compared to reconstructed MC.
Mu_Reco_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.
Mu_Trig_dPhi	1	Requirements on $\Delta\phi_{\mu\mu}$ to determine detector symmetry.
Mu_Trig_isoLooseAvg	1	Variation of looser isolation requirements.
Mu_Trig_isoTightAvg	1	Variation of tighter isolation requirements.

Mu_Trig_mlltight	1	Variation of invariant mass window required for tag and probe pair.
Mu_Trig_noIP	1	Systematic uncertainty of not requiring a matching interaction point.
Mu_Trig_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.
Mu_TTVA_BkgFraction	1	Systematic uncertainty on background fraction subtracted.
Mu_TTVA_LumiUncert	1	Systematic uncertainty on luminosity used to scale MC.
Mu_TTVA_MCxsec	1	Systematic uncertainty on cross-section used to scale MC.
Mu_TTVA_QCDTemplate	1	Systematic uncertainty on QCD background fit templates that mimic probe muons.
Mu_TTVA_SuppressionScale	1	Variation on parameter used to estimate energy required to produce di-muon pair.
Mu_TTVA_Uncorr	1	Statistical uncertainty of the efficiency in each $p_T \times \eta$ bin.

Table 10.4: Description of the muon channel only systematics NPs used in this analysis. Each NP has an up and down variation.

10.2.2 Theoretical Uncertainties

The last source of systematic uncertainty considered is related to the theoretical modelling. Only one source of theoretical uncertainty is used for this measurement because other sources of theoretical uncertainty are expected to be negligible. These potential uncertainty sources, such as parton shower, generator modelling, and QED/electroweak corrections, are not considered as they were found to be negligible when profiled for the Z angular coef-

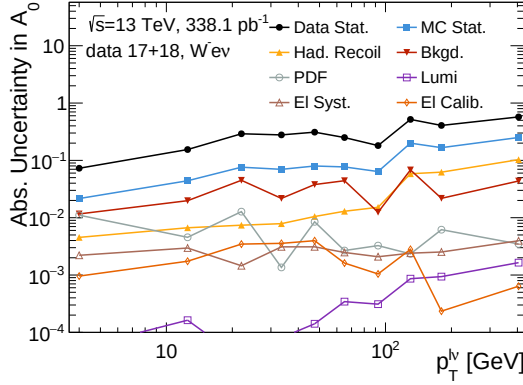
ficient measurement at $\sqrt{s} = 8$ TeV [6]. This measurement had more statistical precision than the data set used in this thesis, so would be more sensitive to these uncertainties. The only theoretical uncertainties considered are associated to PDF. These are expected to be the largest theoretical uncertainty based on the studies for the ATLAS Collaboration Z angular coefficient measurement at $\sqrt{s} = 8$ TeV [6]. The PDF variations reweigh the eigenvectors used to generate events in MC from the CT10 [102] PDF set as well as vary the scale of the strong coupling, α_s . The scale is nominally at $\alpha_s(m^Z) = 0.117982$ and is varied between 0.116 and 0.120. A summary of the PDF NPs are described in Table 10.5. The total PDF uncertainties are labeled with “PDF” and are denoted by the grey open circles in Figure 10.4.

Nuisance Parameter Name	Number of NPs	Description
PDF	26	Systematic uncertainty on the parton distribution function eigenvectors used to weight each MC event for the CT10 [102] PDF set.
PDF α_s	1	Systematic uncertainty on α_s used. Varied between 0.116 and 0.120.

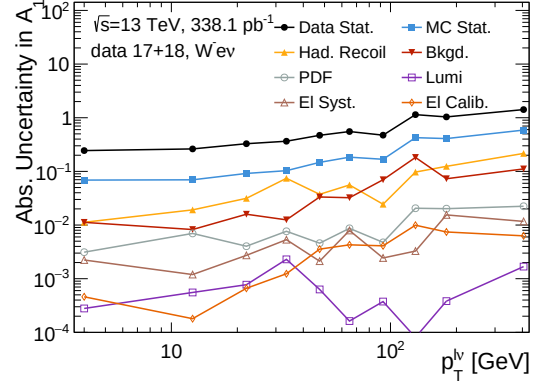
Table 10.5: Description of the PDF systematic uncertainty nuisance parameters used in this analysis. Each NP will have an up and down variation.

10.3 Summary of Uncertainties

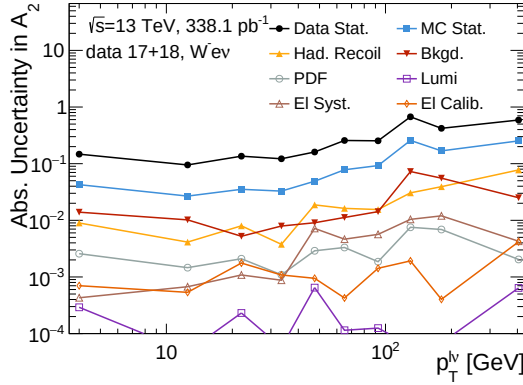
Examples of all sources of uncertainty for the results of the $W^- \rightarrow e^- \bar{\nu}$ channel from Section 9.1 can be found in Figure 10.5. This figure shows the uncertainties for the coefficients A_0 to A_4 and the differential cross-section as function of $p_T^{l\nu}$. A summary of uncertainties for other channels and for A_5 to A_7 can be found in Appendix G. The systematic PDF and luminosity uncertainties on the differential cross-section in Figure 10.5f) are larger in comparison to the angular coefficients A_0 to A_4 in Figure 10.5a)-e). This is because the differential cross-section is not a ratio and does not benefit from cancellation for most sources of uncertainty, as in the case for A_0 to A_4 . The data statistical uncertainty is dominant compared to all other sources of uncertainty and is generally at least one order of magnitude larger compared to any source of systematic uncertainty.



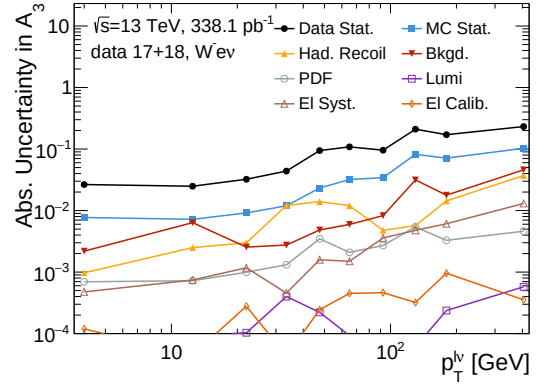
(a) Uncertainties for A_0



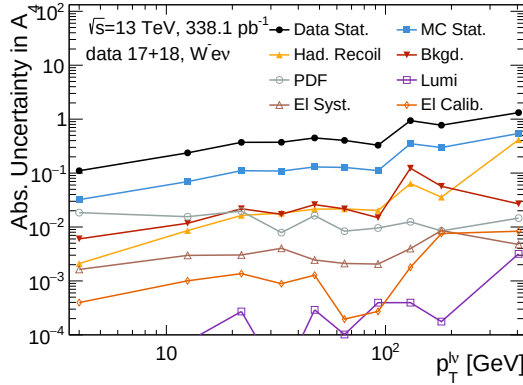
(b) Uncertainties for A_1



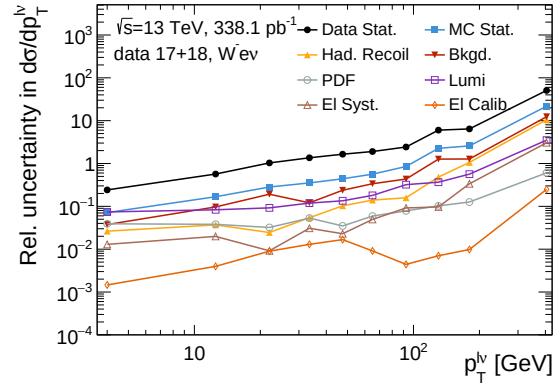
(c) Uncertainties for A_2



(d) Uncertainties for A_3



(e) Uncertainties for A_4



(f) Uncertainties for the differential cross-section

Figure 10.5: A breakdown of all sources of measured uncertainty for the angular coefficients, A_0 to A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for Figure 9.1 for the $W^- \rightarrow e^- \bar{\nu}$ channel. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.

10.4 Constraint of Uncertainties

As part of the NLL fit, the systematic uncertainties are parameterized by a Gaussian probability density function, as described in Section 7.4. They include variations relative to the nominal templates. The systematic variations correlate to being one standard deviation for the Gaussian probability density function used to describe the systematic uncertainty. By varying the templates within their uncertainties during the fitting procedure the NLL is minimized so that the expected number of events best describes the observed number of events. During this minimization, the uncertainty describing the systematic variations can be constrained to have a smaller uncertainty. This constraint is a result of the fit reducing the uncertainty to a level that still describes the observed number of events. It is usually constrained in two ways, through the data statistical uncertainty or using information about the shape of the distribution. If the pre-fit systematic variation is larger than the data statistical uncertainty, then the systematic variation can be constrained by the data statistical uncertainty. This is because the additional information from the data statistical uncertainty will limit the systematic variation within the bounds on the statistical uncertainty.

As previously mentioned in Chapter 9, part of ensuring a good fit is looking at the systematic variations post-fit to make sure they are not over-constrained and are properly propagated. This is done by plotting the pull, g :

$$g = \frac{x - \mu}{\sigma}, \quad (10.3)$$

where x is the randomly distributed variable, described by a Gaussian distribution with mean, μ , and standard deviation, σ . Before the fit, the mean is zero and the standard deviation is one, as described by the dashed line and the green band seen in Figure 10.6. The pulls and constraints of each systematic NP post-fit are plotted in alphanumerical order using the black points and error bars. The used NPs and their descriptions can be found in Section 10.2. For a NP, the pull is the distance between the point and zero denoted

by the dashed line, which is considered good if it is within one standard deviation of the pre-fit uncertainty – described by the green band – and acceptable if within two standard deviations – described by the yellow band. The relative size of the error bars compared to bands provides the amount of constraint on the uncertainty. Figure 10.6 is the pulls on the fit of $W^- \rightarrow e^- \bar{\nu}$ as a function of $p_T^{l\nu}$ and is shown as an example. The plots for the other channels can be found in Appendix G.

Looking at individual NPs allows for understanding what is happening during the fit in order to minimize the NLL. For example, the first NP, “BKGD_XSEC_BOSON”, is the variation on the boson background cross-sections. It has a pull of $g \sim -0.9$, which means that the fit prefers to have -0.9 less of the variation. This particular NP is a 5% variation, to best describe the data, the fit is minimized when the boson background contribution is 4.5% smaller. Within Figure 10.6, most of the points are within one standard deviation of the pre-fit uncertainty and most of the NPs do not have too large of a constraint, which indicates the systematics are appropriately modelled.

With an understanding of how the results are extracted and combined from Chapter 9 and understanding the sources of uncertainty from this chapter, the final results for the combined measurements and uncertainties will be presented in Chapter 11. This upcoming chapter will discuss the potential impact on future measurements of the W boson mass with the measurements in this thesis.

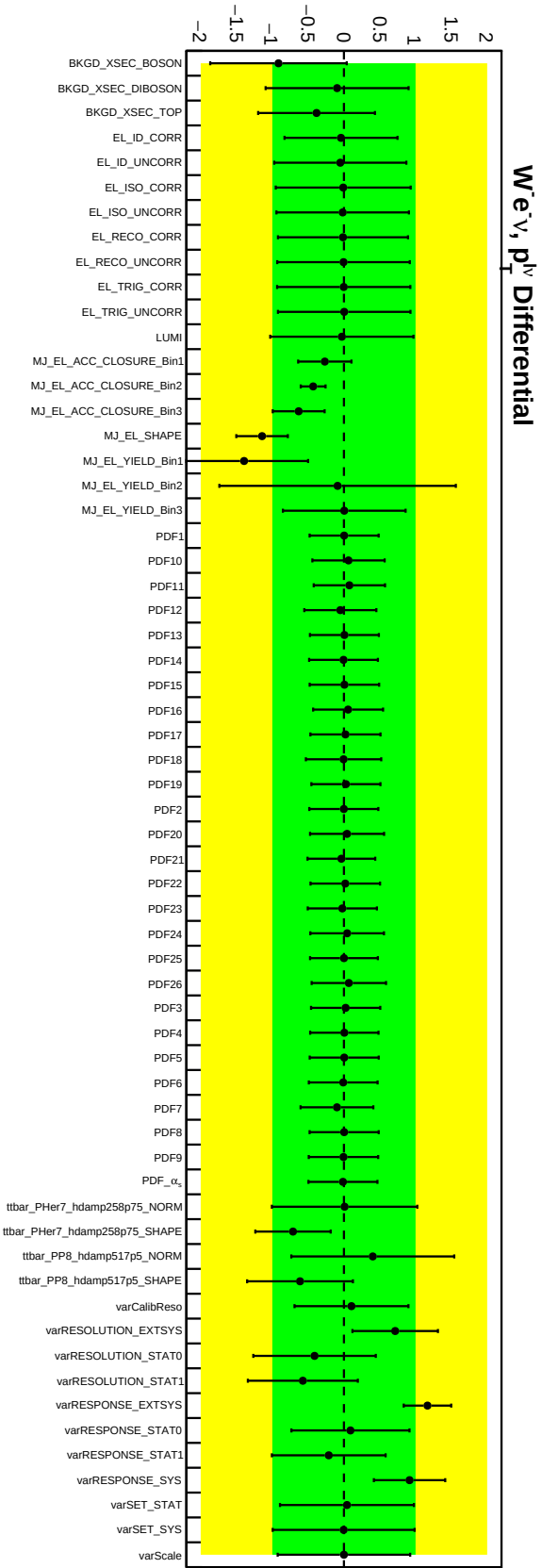


Figure 10.6: The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell-Yan angular coefficients and differential cross-section as a function of $p_T^{l\nu}$ for the $W^- \rightarrow e^- \bar{\nu}$ channel. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modelled as most points are within one standard deviation of the pre-fit uncertainty.

CHAPTER 11

COMBINED CHANNEL MEASUREMENT OF THE W BOSON DRELL–YAN ANGULAR COEFFICIENTS AND DIFFERENTIAL CROSS-SECTION

This chapter will summarize the final measurements of the coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_1^{l\nu}$ for W^+ and W^- . These measurements are the combined electron and muon channels using the method described in Section 9.3. Section 11.2 summarizes the decomposition of all sources of systematic uncertainty and the data statistical uncertainty for both W^+ and W^- . The other coefficients that are not statistically different than zero – A_1 , A_5 , A_6 , and A_7 – are included in Appendix G. These four coefficients are extracted during the fit but do not have sufficient statistical sensitivity to measure them. Lastly, the potential impact of these measurements on future W boson mass measurements will be discussed in Section 11.3.

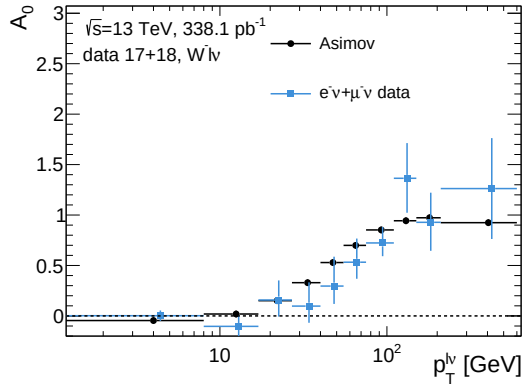
11.1 Measurements of Drell–Yan Angular Coefficients and Differential Cross-Section

The final results of the measured Drell–Yan angular coefficients and the differential cross-section can be seen in Figures 11.1 and 11.2 for W^- and W^+ , respectively. Here, the measured values are denoted by blue squares and are compared to the predicted coefficients using Asimov data, which are denoted by the black points. The measured values are plotted with both data statistical and systematic uncertainties, while Asimov data sets are plotted with only the uncertainty on the CT10 PDF set used to generate the MC, as this is the only theoretical uncertainty source. The compatibility between the predictions using Asimov data and the measured values can be found in Table 9.2 listed as the difference between data and CT10. This compatibility is calculated using all eight Drell–Yan angular coefficients and the differential cross-section. The compatibility between the expected and observed number of events from the NLL fit to extract the coefficients are found to be compatible

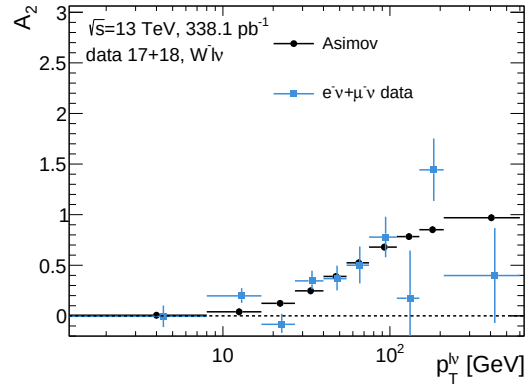
and are listed in Table 11.1.

Channel	χ^2/ndf	p-value
$e^-\nu + \mu^-\nu$ Combination	1117.24/1175	0.885
$e^+\nu + \mu^+\nu$ Combination	1254.22/1184	0.076

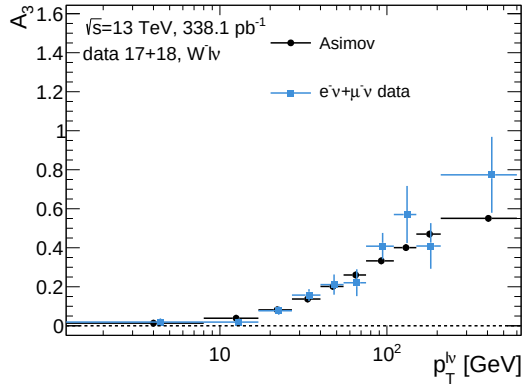
Table 11.1: Summary of χ^2/ndf and p-values measuring the compatibility between the number of observed events and the number of expected events after the fit for the measurements using combined channels as a function of $p_T^{l\nu}$.



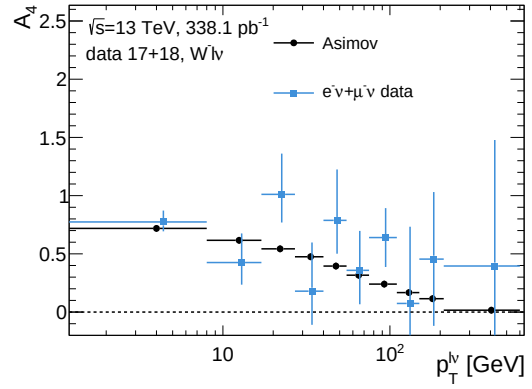
(a) A_0



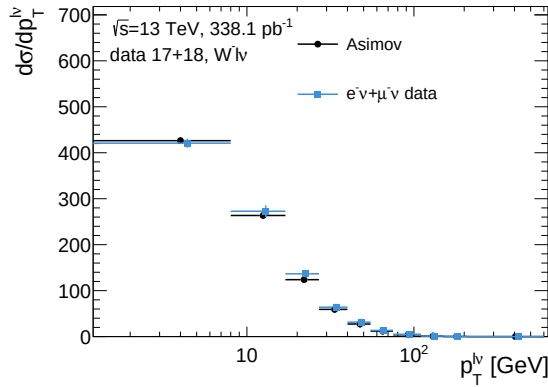
(b) A_2



(c) A_3

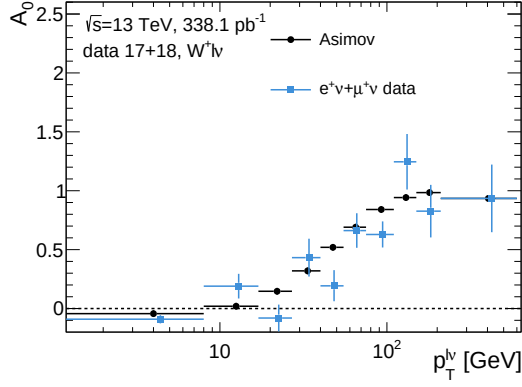


(d) A_4

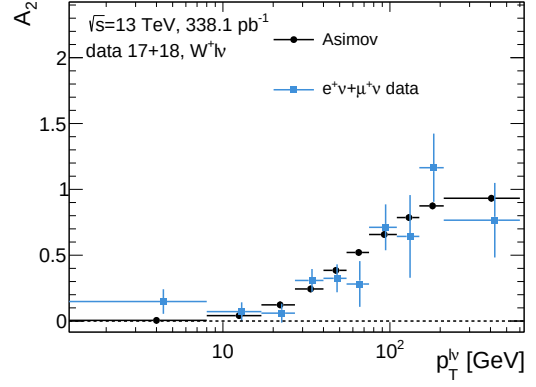


(e) Differential cross-section

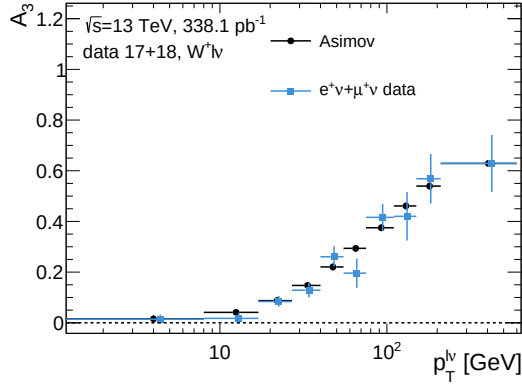
Figure 11.1: The measured four major coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for the combination of the $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



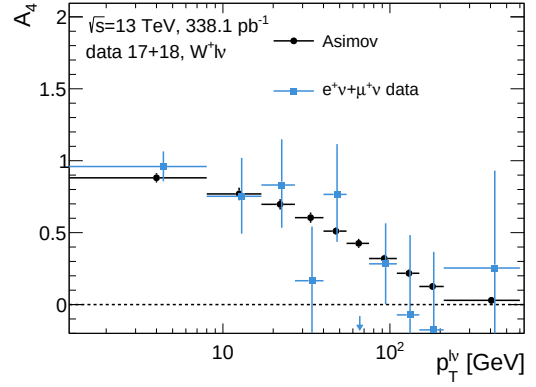
(a) A_0



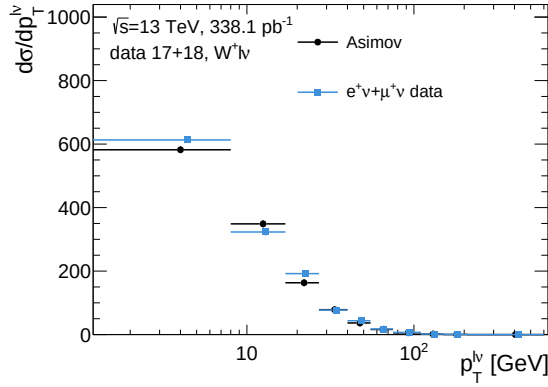
(b) A_2



(c) A_3



(d) A_4

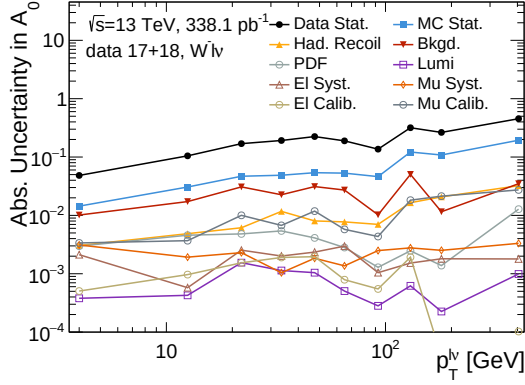


(e) Differential cross-section

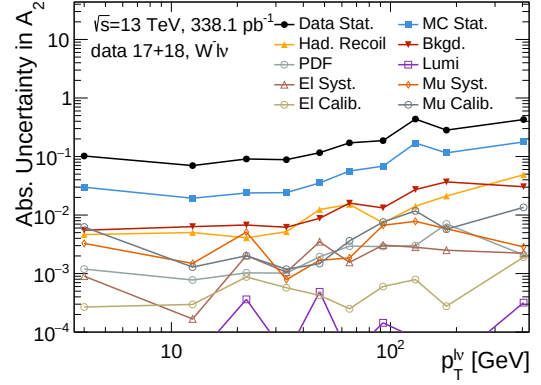
Figure 11.2: The measured four major coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for the combination of the $W^+ \rightarrow e^+ \nu$ and $W^+ \rightarrow \mu^+ \nu$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.

11.2 Uncertainty Decomposition of the Drell–Yan Angular Coefficients and Differential Cross-Section

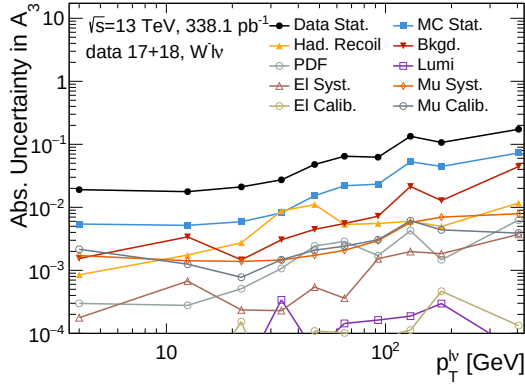
The total uncertainties for the measurements of the Drell–Yan angular coefficients and the differential cross-section for the combined channel measurements can be decomposed to show the contribution of each source of uncertainty. This is shown for W^- and W^+ in Figures 11.3 and 11.4, respectively. The absolute uncertainties are shown for the measured angular coefficients and the relative uncertainties are shown for the differential cross-section. How the data statistical uncertainty is determined is outlined in Section 10.1 and the sources of systematic uncertainty are outlined in Section 10.2. The measurements’ total uncertainties are still data statistically dominated, even with the statistical power of combining channels. The relative uncertainties for the differential cross-section in the first $p_T^{l\nu}$ bin are 2.3% for W^- and 1.7% for W^+ , which are comparable to an ATLAS measurement using the same data set [12].



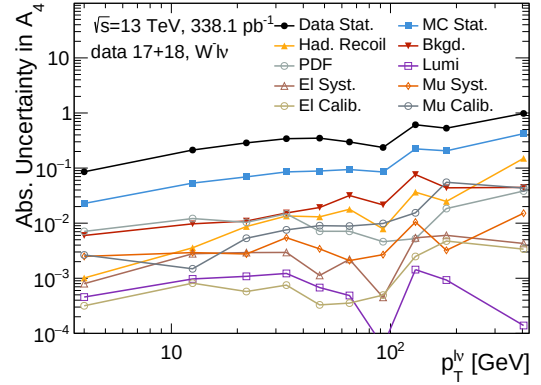
(a) Uncertainties for A_0



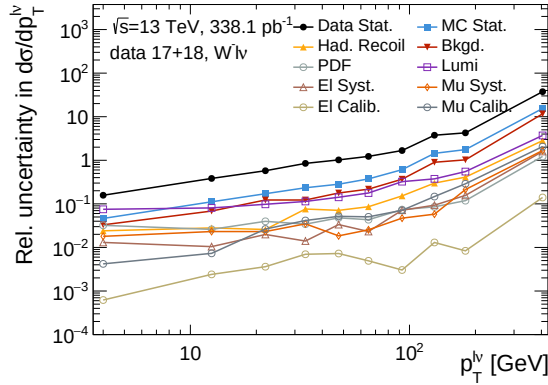
(b) Uncertainties for A_2



(c) Uncertainties for A_3

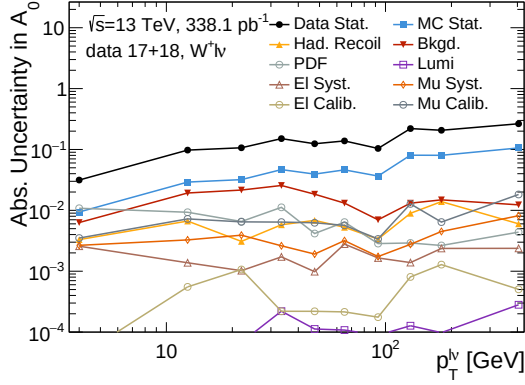


(d) Uncertainties for A_4

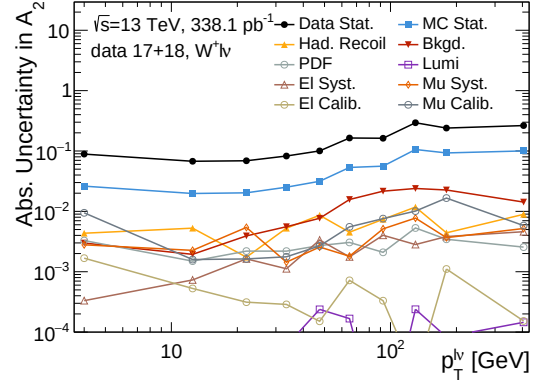


(e) Uncertainties for the differential cross-section

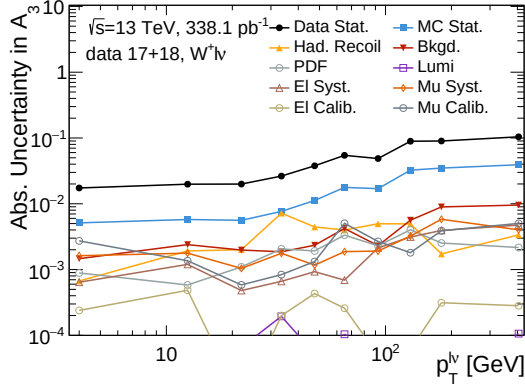
Figure 11.3: A breakdown of all sources of measured uncertainty for the measured angular coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for Figure 11.1 for the combination of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The data statistical uncertainty outlined in Section 10.1 is denoted by the black points and all sources of systematic uncertainty are outlined in Section 10.2.



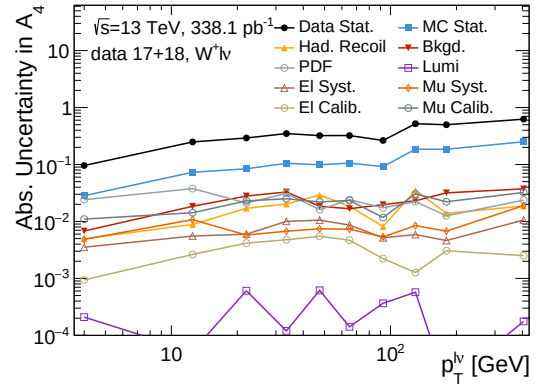
(a) Uncertainties for A_0



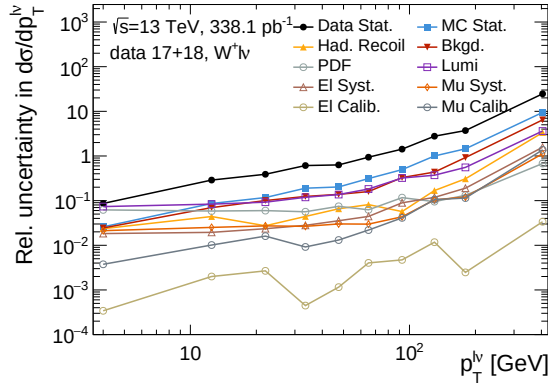
(b) Uncertainties for A_2



(c) Uncertainties for A_3



(d) Uncertainties for A_4



(e) Uncertainties for the differential cross-section

Figure 11.4: A breakdown of all sources of measured uncertainty for the measured angular coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for Figure 11.2 for the combination of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The data statistical uncertainty outlined in Section 10.1 is denoted by the black points and all sources of systematic uncertainty are outlined in Section 10.2.

11.3 Potential Impact on W Mass

To understand how these measurements might improve the uncertainty on the W boson mass, m^W , it is crucial to understand how the measurement is performed. The measurement of the W boson mass utilizes a profile LH fit of p_T^l and m_T to extract m^W [16]. This involves varying the shape of the p_T^l and m_T distributions within the bounds of different uncertainties and trying to find a model that best describes the distributions measured with data. The measured value of m^W is found in two ways. The first, is by taking the peak value of m_T , which should peak at the mass of the W boson. The second, is by taking twice the peak value of p_T^l – known as the Jacobian peak – which should peak at $m^W/2$ [126]. This means that, to find the impact on the W boson mass, it is necessary to understand the impact of the Drell–Yan angular coefficients and differential cross-section on the p_T^l and m_T distributions. An estimate of the impact of the Drell–Yan angular coefficients and differential cross-section on the PDF sets used in Section 9.4 was previously done in [127], which also tested the impact on the uncertainties if the data set were 10 times larger for potential future measurements.

To get the impact, predictions of the differential cross-section and A_4 values were used to profile the PDF sets, which constrains the uncertainty on the PDF sets based on the new information. These profiled PDF sets could then be used to estimate the relative uncertainty for the p_T^l and m_T distributions. From [127], Figures 6.5.1, 6.5.5, and 6.5.7 display the impact for the PDF sets CT18, MSHT20, and NNPDF4.0, respectively. For CT18, the original uncertainty is approximately 3%, it is constrained to 2% after profiling. Profiling 10 times the data constrains it further to 1.6% and 1.3% for p_T^l and m_T , respectively. For MSHT20, the original uncertainty is 1.5%, and is constrained to 1.2% after profiling. Profiling 10 times the data constrains it further to 0.6% and 0.7% for p_T^l and m_T , respectively. Lastly, for NNPDF4.0, the original uncertainty is 0.52%, and is constrained marginally to 0.49% after profiling. Profiling 10 times the data constrains the uncertainty down to 0.39%. The

potential reduction of uncertainty for a future m^W measurement using the measurements of the Drell–Yan angular coefficients and differential cross-section from this chapter would be on par with these predictions. This highlights the usefulness of these measurements and potentially larger data sets.

CHAPTER 12

CONCLUSIONS AND OUTLOOK

The results in Chapter 11 show good agreement to the expected values for the Drell–Yan angular coefficients and the differential cross-section. This highlights the ability to reconstruct the W boson fully and use it for precision measurements despite the neutrino going undetected. This full reconstruction of the W bosons allows for the measurement of the coefficients A_0 and A_4 as a function of $p_T^{\ell\nu}$, which has never been done previously. The reconstruction of the neutrino was made possible by the improved hadronic recoil resolution for the low pileup data set, which was large enough to have sufficient statistical sensitivity to properly resolve the sign of $\cos \theta_{CS}$. The measurements of the A_2 and A_3 as a function of $p_T^{\ell\nu}$ were measured for the first time at the LHC. Finally, the differential cross-section was measured with relative uncertainties -2.3% for W^- and 1.7% for W^+ in the first $p_T^{\ell\nu}$ bin – comparable to an ATLAS measurement using the same data set [12].

These measurements are statistically limited, as expected due to the small data set, but the measurements are still useful as a direct probe of perturbative QCD production modelling. They can be used for profiling the PDF sets used as inputs for the W boson mass measurement, as discussed in Section 11.3. By replacing the expected values used in the profiling in Section 11.3 with the measured values from this thesis. These measurements will provide a more accurate estimate of the uncertainties of the angular coefficients and potentially reduce the uncertainty of PDF profiling, which would be a valuable input to reduce uncertainty to future m^W measurements. These measurements can also be used for global PDF fits that take many measurements as inputs to constrain PDFs and their uncertainty, such as for the up quark in Figure 6.3.

The largest benefit to reduce the uncertainty and expand the reach of these measurements would be to expand the size of the data set to reduce the data statistical uncertainties.

This would require the LHC and the associated experimental collaborations agree to perform more special runs with low pileup. More special runs with low pileup would benefit other precision measurements at the cost of reduced data set size with high pileup. These high pileup data sets are often used for searches of new physics or measuring very rare processes and rely on large data sets. The Drell–Yan angular coefficients and differential cross-section measurements in this thesis highlight the benefit to running in these special low pileup conditions for short periods of time. There are ongoing discussions for another special low pileup data taking period before the end of Run 3 in 2026.

A larger data set would improve more than just the precision of the results shown in this thesis. It would allow for a meaningful measurement of the coefficients A_1 , A_5 , A_6 , and A_7 that are not currently statistically different from zero. This would be important as it would allow for all of the coefficients to be measured, as previously done for the Z boson. It would allow for testing the difference between W^- and W^+ for A_5 , which for the W boson is CP-odd. Testing the difference would allow for assessing for any unexpected CP violation. A larger data set would also allow for measuring the Drell–Yan angular coefficients and differential cross-section as a function of $|y^W|$, which has larger migration between rapidity bins and requires more statistics to constrain. With a data set on the order of 10 fb^{-1} , the Drell–Yan angular coefficients and differential cross-section could also be measured as a two-dimensional function of both p_T^W and y^W , as previously done for the Z boson at $\sqrt{s} = 8 \text{ TeV}$ [6]. This would not require the same data set size, 21 fb^{-1} , as what was used for the Z boson when measured at $\sqrt{s} = 8 \text{ TeV}$. This is because the W boson is produced at a higher rate, due to its cross-section being approximately a factor of 10 larger [128]. It would require a significantly larger data set than what is used for this thesis to achieve the same precision due to the extra migration for W boson event reconstruction compared to the Z boson.

A different potential direction to increase the size of the data set would be to not use the special low pileup data set, but use the data sets from the normal high pileup running con-

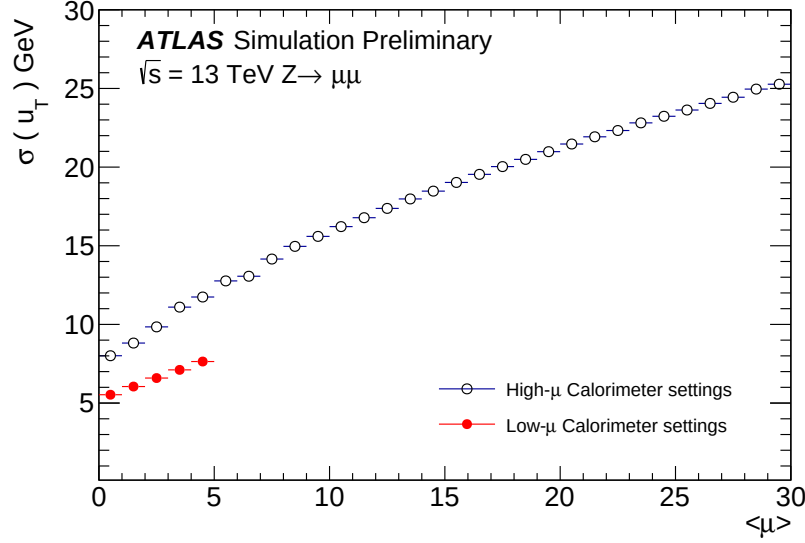


Figure 12.1: A prediction from simulation for the uncertainty on the hadronic recoil, u_T , as a function of the average of the mean number of interactions per bunch crossing, $\langle \mu \rangle$, in the ATLAS detector taken from [94]. This is shown for the normal high- $\langle \mu \rangle$ settings for the calorimeter in the open circles and for the special calorimeter settings used with lower energy thresholds used for the low- $\langle \mu \rangle$ data set.

ditions instead. The most likely candidate would be to use the data set taken in 2015 during Run 2, which corresponds to an integrated luminosity of 3.2 fb^{-1} recorded by ATLAS and had $\langle \mu \rangle = 13.4$ [45]. This would increase the uncertainty on the hadronic recoil, u_T , and the uncertainty for p_T^W by a significant amount, as seen in Figure 12.1. This figure shows the hadronic recoil uncertainty as a function of $\langle \mu \rangle$ from [94]. For $\langle \mu \rangle = 13.4$, this corresponds to an uncertainty of $\sigma(u_T) \sim 17 \text{ GeV}$, which is more than twice the uncertainty on the hadronic recoil for the low pileup data set of $\sigma(u_T) \sim 8 \text{ GeV}$ used for this thesis. This means a measurement with the 2015 Run 2 data set would need larger p_T^W bins to avoid large migration between p_T^W bins because of the increased uncertainty on the hadronic recoil. The m^W measurements selects events with $p_T^W < 30 \text{ GeV}$ and depend on having smaller bins at low p_T^W , so a larger binning would not be as useful to reduce the uncertainty for future m^W measurements. However, this would not negatively affect measurements of the Drell–Yan angular coefficients and differential cross-section at high p_T^W , as most coefficients peak at high p_T^W in the CS frame. It would allow for measuring the Lam–Tung

relation [129, 130, 131], which is the difference of A_0 and A_2 . At NLO in $q\bar{q} \rightarrow Wg$ and $qg \rightarrow Wq$ diagrams, the difference between A_0 and A_2 is expected to be zero if the gluon has spin $J = 1$. Starting at NNLO, the Lam–Tung relation is expected to be non-zero at high p_T^W , so a measurement of this difference would help constrain the gluon PDF.

There are other ways to potentially improve this analysis, besides a larger data set. These improvements would require further testing to determine their viability and are not covered under the scope of this thesis. An example is testing with machine learning models, such as a neural network, for determining if the neutrino p_z^ν solution choice from the quadratic equation is more separable instead of choosing randomly. Another possible improvement would be improving the mass constraint methodology used for the neutrino p_z^ν solution. Instead of using the on-shell mass value from PDG [111], the transverse mass – m_{T-} plus some factor could be used to expand the number of real solutions. Using just m_T in place of the mass to find the p_z^ν solution results in $p_z^\nu \sim 0$. Thus, a method for how much offset is required to be included to not bias the p_z^ν distribution would need to be determined.

All of these things considered, the measurements from this thesis are important to better understand W bosons at the LHC. This is done through measuring A_2 and A_3 as a function of more $p_T^{l\nu}$ bins while maintaining better precision and the first measurements of A_0 and A_4 as a function of $p_T^{l\nu}$ for the W boson. The measurements of these coefficients can be used to give insight into the modelling of perturbative QCD W boson production and will benefit future W boson mass measurements to help constrain the SM. Hopefully, they will inspire future measurements for the W boson Drell–Yan angular coefficients, where improvements can be made to reduce the uncertainty on the measurement to further the collective knowledge on the W boson and its place in the SM.

BIBLIOGRAPHY

- [1] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008). URL: <http://doi.org/10.1088/1748-0221/3/08/s08003>.
- [2] S. D. Drell and T.M. Yan. “Partons and their applications at high energies”. In: *Annals of Physics* 66.2 (1971), pp. 578–623. URL: <http://www.sciencedirect.com/science/article/pii/0003491671900716>.
- [3] S. D. Drell and T. M. Yan. “Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies”. In: *Phys. Rev. Lett.* 25 (5 1970), pp. 316–320. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.25.316>.
- [4] E. Mirkes. “Angular decay distribution of leptons from W-bosons at NLO in hadronic collisions”. In: *Nuclear Physics B* 387.1 (1992), pp. 3–85. URL: <http://www.sciencedirect.com/science/article/pii/055032139290046E>.
- [5] LHCb Collaboration. “First measurement of the $Z \rightarrow \mu^+ \mu^-$ angular coefficients in the forward region of pp collisions at $\sqrt{s} = 13$ TeV”. In: (Mar. 2022). arXiv: [2203.01602](https://arxiv.org/abs/2203.01602) [hep-ex].
- [6] ATLAS Collaboration. “Measurement of the angular coefficients in Z -boson events using electron and muon pairs from data taken at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *JHEP* 08 (2016), p. 159. arXiv: [1606.00689](https://arxiv.org/abs/1606.00689) [hep-ex].
- [7] ATLAS Collaboration. *Measurement of the effective leptonic weak mixing angle using electron and muon pairs from Z -boson decay in the ATLAS experiment at $\sqrt{s} = 8$ TeV*. Tech. rep. ATLAS-CONF-2018-037. Geneva: CERN, 2018. URL: <http://cds.cern.ch/record/2630340>.

- [8] CMS Collaboration. “Angular coefficients of Z bosons produced in pp collisions at $\sqrt{s} = 8$ TeV and decaying to $\mu^+\mu^-$ as a function of transverse momentum and rapidity”. In: *Phys. Lett. B* 750 (2015), p. 154. arXiv: [1504.03512 \[hep-ex\]](#).
- [9] T. Aaltonen et al. “First Measurement of the Angular Coefficients of Drell-Yan e^+e^- Pairs in the Z Mass Region from $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Rev. Lett.* 106 (24 June 2011), p. 241801. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.106.241801>.
- [10] D. Acosta et al. “Measurement of the azimuthal angle distribution of leptons from W boson decays as a function of the W transverse momentum in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV”. In: *Phys. Rev. D* 73 (2006), p. 052002. arXiv: [hep-ex/0504020](#).
- [11] ATLAS Collaboration. “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”. In: *Eur. Phys. J. C* 83.10 (2023), p. 982. arXiv: [2212.09379 \[hep-ex\]](#).
- [12] ATLAS Collaboration. “Precise measurements of W - and Z -boson transverse momentum spectra with the ATLAS detector using pp collisions at $\sqrt{s} = 5.02$ TeV and 13 TeV”. In: (Apr. 2024). arXiv: [2404.06204 \[hep-ex\]](#).
- [13] J. Pequeno. “Computer generated image of the whole ATLAS detector”. 2008. URL: <http://cds.cern.ch/record/1095924>.
- [14] T. Aaltonen et al. “High-precision measurement of the W boson mass with the CDF II detector”. In: *Science* 376.6589 (2022), pp. 170–176. URL: <https://www.science.org/doi/abs/10.1126/science.abk1781>.
- [15] LHCb Collaboration. “Measurement of the W boson mass”. In: *JHEP* 01 (2022), p. 036. arXiv: [2109.01113 \[hep-ex\]](#).

- [16] ATLAS Collaboration. “Measurement of the W-boson mass and width with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 7$ TeV”. In: (Mar. 2024). arXiv: [2403.15085 \[hep-ex\]](#).
- [17] CMS Collaboration. *Measurement of the W boson mass in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2024. URL: <https://cds.cern.ch/record/2910372>.
- [18] S. Amoroso et al. “Compatibility and combination of world W-boson mass measurements”. In: *Eur. Phys. J. C* 84.5 (2024), p. 451. arXiv: [2308.09417 \[hep-ex\]](#).
- [19] G. E. P. Box. “Science and Statistics”. In: *Journal of the American Statistical Association* 71.356 (1976), pp. 791–799.
- [20] F. J. Hasert et al. “Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment”. In: *Phys. Lett. B* 46 (1973), pp. 138–140.
- [21] D. J. Griffiths. *Introduction to elementary particles; 2nd rev. version*. New York, NY: Wiley, 2008. ISBN: 9783527406012.
- [22] Y. Nagashima. *Elementary Particle Physics: Quantum Field Theory and Particles*. Elementary Particle Physics v. 1. Wiley, 2010. ISBN: 9783527630103.
- [23] Y. Nagashima. *Elementary Particle Physics: Foundations of the Standard Model*. Elementary Particle Physics v. 2. Wiley, 2013. ISBN: 9783527409662.
- [24] F. Halzen and A. D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984. ISBN: 0471887412.
- [25] R. L. Workman et al. “Review of Particle Physics”. In: *PTEP* 2022 (2022), p. 083C01.
- [26] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29. arXiv: [1207.7214 \[hep-ex\]](#).

- [27] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 30–61. arXiv: [1207.7235 \[hep-ex\]](#).
- [28] M. Maggiore. *A Modern Introduction to Quantum Field Theory*. EBSCO ebook academic collection. Oxford University Press, 2005. ISBN: 9780198520733.
- [29] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Boulder, USA: Westview Press, 2016. ISBN: 9780813350196. URL: <http://www.slac.stanford.edu/~mpeskin/QFT.html>.
- [30] E. Noether. “Invariant variation problems”. In: *Transport Theory and Statistical Physics* 1.3 (Jan. 1971), pp. 186–207. URL: <https://doi.org/10.1080/2F00411457108231446>.
- [31] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos. “A Historical Profile of the Higgs Boson”. In: (Jan. 2012). arXiv: [1201.6045 \[hep-ph\]](#).
- [32] LHCb Collaboration. “Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$ ”. In: *Phys. Rev. Lett.* 122.22 (2019), p. 222001. arXiv: [1904.03947 \[hep-ex\]](#).
- [33] LHCb Collaboration. “Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$ ”. In: *Submitted to: Phys. Rev. Lett.* (2019). arXiv: [1904.03947 \[hep-ex\]](#).
- [34] LHCb Collaboration. “Amplitude analysis of the $B^+ \rightarrow D^+ D^- K^+$ decay”. In: *Phys. Rev. D* 102 (11 Dec. 2020), p. 112003. URL: <https://link.aps.org/doi/10.1103/PhysRevD.102.112003>.
- [35] L. Evans and P. Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008). URL: <http://doi.org/10.1088/1748-0221/3/08/s08001>.

- [36] E. Lopienska. “The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022”. In: (2022). General Photo. URL: <https://cds.cern.ch/record/2800984>.
- [37] ALICE Collaboration. “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008). URL: <https://doi.org/10.1088/1748-0221/3/08/s08002>.
- [38] CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008). URL: <http://doi.org/10.1088/1748-0221/3/08/s08004>.
- [39] LHCb Collaboration. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008). URL: <https://doi.org/10.1088/1748-0221/3/08/s08005>.
- [40] O. Adriani et al. “The LHCf detector at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008). URL: <http://doi.org/10.1088/1748-0221/3/08/s08006>.
- [41] A. Ariga et al. “Technical Proposal for FASER: ForwArd Search ExpeRiment at the LHC”. In: (Dec. 2018). arXiv: [1812.09139](https://arxiv.org/abs/1812.09139) [physics.ins-det].
- [42] G. Anelli et al. “The TOTEM Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008). URL: <http://doi.org/10.1088/1748-0221/3/08/s08007>.
- [43] W. Herr and B. Muratori. “Concept of luminosity”. In: (2006). URL: <http://cds.cern.ch/record/941318>.
- [44] V. Balagura. “Van der Meer scan luminosity measurement and beam–beam correction”. In: *Eur. Phys. J. C* 81.1 (2021), p. 26. arXiv: [2012.07752](https://arxiv.org/abs/2012.07752) [hep-ex].

- [45] ATLAS Collaboration. *ATLAS Experiment – Public Results, Luminosity Public Results Run 2*. Accessed Mar, 2023. URL: <http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [46] ATLAS Collaboration. *ATLAS Experiment – Public Results, Luminosity Public Results Run 3*. Accessed Mar, 2023. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3>.
- [47] Z. Marshall et al. “Simulation of Pile-up in the ATLAS Experiment”. In: *Journal of Physics: Conference Series* 513.2 (2014), p. 022024. URL: <http://doi.org/10.1088/1742-6596/513/2/022024>.
- [48] ATLAS Collaboration. *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: <http://cds.cern.ch/record/331063>.
- [49] ATLAS Collaboration. *ATLAS inner detector: Technical Design Report, 2*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: <http://cds.cern.ch/record/331064>.
- [50] ATLAS Collaboration. *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*. Tech. rep. ATL-PHYS-PUB-2015-018. Geneva: CERN, 2015. URL: <http://cds.cern.ch/record/2037683>.
- [51] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. 2010. URL: <https://cds.cern.ch/record/1291633>.
- [52] ATLAS Collaboration. *ATLAS Inner Tracker Performance at the Beginning of the LHC Run 2*. Tech. rep. ATL-INDET-PROC-2016-002. Geneva: CERN, 2016. URL: <http://cds.cern.ch/record/2147407>.

- [53] ATLAS Collaboration. “The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance”. In: *Journal of Instrumentation* 3.02 (2008). URL: <http://doi.org/10.1088/1748-0221/3/02/p02013>.
- [54] ATLAS Collaboration. “The ATLAS TRT Barrel Detector”. In: *Journal of Instrumentation* 3.02 (2008). URL: <http://doi.org/10.108/1748-0221/3/02/p02014>.
- [55] ATLAS Collaboration. “The ATLAS TRT end-cap detectors”. In: *Journal of Instrumentation* 3.10 (2008). URL: <http://doi.org/10.1088/1748-0221/3/10/p10003>.
- [56] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: <http://cds.cern.ch/record/331061>.
- [57] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: <http://cds.cern.ch/record/331062>.
- [58] J. Pequenaio. “Computer Generated image of the ATLAS calorimeter”. 2008. URL: <http://cds.cern.ch/record/1095927>.
- [59] R. C. Fernow. *Introduction to Experimental Particle Physics*. Cambridge University Press, 1989. ISBN: 9780521379403.
- [60] ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: <http://cds.cern.ch/record/331068>.
- [61] ATLAS Collaboration. “ATLAS Muon Detector Commissioning”. In: *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2009, Detroit, USA, July 26-31, 2009*. arXiv: 0910.2767 [physics.ins-det].

- [62] F. Lasagni Manghi. *LUCID: The ATLAS Luminosity Detector*. Tech. rep. Geneva: CERN, 2019. URL: <https://cds.cern.ch/record/2646366>.
- [63] ATLAS Collaboration. “Performance of the ATLAS trigger system in 2015”. In: *The European Physical Journal C* 77.5 (2017), p. 317. URL: <https://doi.org/10.1140/epjc/s10052-017-4852-3>.
- [64] Diego Casadei. “Estimating the selection efficiency”. In: *JINST* 7 (2012), P08021. arXiv: [0908.0130 \[physics.data-an\]](https://arxiv.org/abs/0908.0130).
- [65] GEANT4 Collaboration, S. Agostinelli, et al. “GEANT4 – a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250.
- [66] ATLAS Collaboration. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV”. In: *EPJC* 79.8 (2019), p. 639. arXiv: [1902.04655 \[physics.ins-det\]](https://arxiv.org/abs/1902.04655).
- [67] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 76.5 (2016), p. 292. arXiv: [1603.05598 \[hep-ex\]](https://arxiv.org/abs/1603.05598).
- [68] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *JINST* 14.12 (2019), P12006. arXiv: [1908.00005 \[hep-ex\]](https://arxiv.org/abs/1908.00005).
- [69] S. Mehlhase. “ATLAS detector slice (and particle visualisations)”. In: (2021). URL: <https://cds.cern.ch/record/2770815>.
- [70] T. Cornelissen et al. *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*. Tech. rep. Geneva: CERN, 2007. URL: <https://cds.cern.ch/record/1020106>.
- [71] T. Cornelissen et al. “The new ATLAS track reconstruction (NEWT)”. In: *J. Phys.: Conf. Ser.* 119 (2008), p. 032014. URL: <https://cds.cern.ch/record/1176900>.

- [72] ATLAS Collaboration. “Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment”. In: *JHEP* 05 (2024), p. 162. arXiv: [2308.13362 \[hep-ex\]](#).
- [73] ATLAS Collaboration. “Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 81.8 (2021), p. 689. arXiv: [2007.02645 \[hep-ex\]](#).
- [74] M. Cacciari, G. P. Salam, and G. Soyez. “The anti- k_t jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063. arXiv: [0802.1189 \[hep-ph\]](#).
- [75] D. Rauch. “Aspects of Rare Production and Decay Processes of the Higgs Boson”. Presented on 28 Sep 2020. Hamburg U., 2020. URL: <https://cds.cern.ch/record/2749541>.
- [76] ATLAS Collaboration. “Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 81.7 (2021), p. 578. arXiv: [2012.00578 \[hep-ex\]](#).
- [77] ATLAS Collaboration. “Performance of electron and photon triggers in ATLAS during LHC Run 2”. In: *Eur. Phys. J. C* 80.1 (2020), p. 47. arXiv: [1909.00761](#). URL: <https://cds.cern.ch/record/2688248>.
- [78] ATLAS Collaboration. “Performance of the ATLAS muon triggers in Run 2”. In: *Journal of Instrumentation* 15.09 (Sept. 2020), P09015. URL: <https://dx.doi.org/10.1088/1748-0221/15/09/P09015>.
- [79] T. Xu. “Measurements of W boson properties at $\sqrt{s} = 5$ and 13 TeV with the ATLAS detector at the LHC”. Presented 25 Sep 2019. 2019. URL: <http://cds.cern.ch/record/2689846>.
- [80] M. Khandoga. “Calibration of electron shower shapes, hadronic recoil reconstruction using deep learning algorithms and the measurement of W boson transverse momentum distribution with the ATLAS detector”. PhD thesis. IRFU, SPP, Saclay, 2020. URL: <https://cds.cern.ch/record/2752635>.

- [81] B. Aubert et al. “Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 558 (2006), pp. 388–418. URL: <https://www.sciencedirect.com/science/article/pii/S016890020502382X>.
- [82] ATLAS Electromagnetic Liquid Argon Group. “Construction, assembly and tests of the ATLAS electromagnetic end-cap calorimeters”. In: *Journal of Instrumentation* 3.06 (June 2008), P06002–P06002. URL: <https://doi.org/10.1088/1748-0221/3/06/p06002>.
- [83] ATLAS Collaboration. “Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons”. In: *Eur. Phys. J. C* 70 (2010), pp. 755–785. arXiv: [1002.4189](https://arxiv.org/abs/1002.4189) [physics.ins-det].
- [84] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.12 (Dec. 2019), P12006–P12006. URL: <https://doi.org/10.1088/1748-0221/14/12/p12006>.
- [85] *Standard Model Summary Plots June 2024*. Tech. rep. Geneva: CERN, 2024. URL: <https://cds.cern.ch/record/2903866>.
- [86] H. Abramowicz et al. “Combination of measurements of inclusive deep inelastic $e^\pm p$ scattering cross sections and QCD analysis of HERA data”. In: *Eur. Phys. J. C* 75.12 (2015), p. 580. arXiv: [1506.06042](https://arxiv.org/abs/1506.06042) [hep-ex].
- [87] ATLAS Collaboration. “Determination of the parton distribution functions of the proton from ATLAS measurements of differential W and Z boson production in association with jets”. In: *JHEP* 07 (2021), p. 223. arXiv: [2101.05095](https://arxiv.org/abs/2101.05095) [hep-ex].

- [88] J. M. Campbell, J. W. Huston, and W. J. Stirling. “Hard Interactions of Quarks and Gluons: A Primer for LHC Physics”. In: *Rept. Prog. Phys.* 70 (2007), p. 89. arXiv: [hep-ph/0611148](#) [hep-ph].
- [89] J. C. Collins and D. E. Soper. “Angular distribution of dileptons in high-energy hadron collisions”. In: *Phys. Rev. D* 16 (1977), pp. 2219–2225. URL: <http://link.aps.org/doi/10.1103/PhysRevD.16.2219>.
- [90] C. M. Carloni Calame et al. “Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders”. In: *JHEP* 10 (2007), p. 109. arXiv: [0710.1722](#) [hep-ph].
- [91] E. Mirkes and J. Ohnemus. “W and Z polarization effects in hadronic collisions”. In: *Phys. Rev. D* 50 (9 1994), pp. 5692–5703. URL: <http://link.aps.org/doi/10.1103/PhysRevD.50.5692>.
- [92] O. Behnke et al., eds. *Data analysis in high energy physics*. Wiley-VCH, 2013. ISBN: 9783527410583.
- [93] G. Cowan. *Statistical Data Analysis*. Oxford: Clarendon (Oxford), 1998. ISBN: 0-19-850155-2.
- [94] ATLAS Collaboration. *Prospects for the measurement of the W-boson transverse momentum with a low pileup data sample at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2298152>.
- [95] P. Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP* 11 (2004), p. 040. arXiv: [hep-ph/0409146](#).
- [96] S. Frixione, P. Nason, and C. Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method”. In: *JHEP* 11 (2007), p. 070. arXiv: [0709.2092](#) [hep-ph].

- [97] S. Alioli et al. “NLO vector-boson production matched with shower in POWHEG”. In: *JHEP* 07 (2008), p. 060. arXiv: [0805.4802 \[hep-ph\]](#).
- [98] S. Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. arXiv: [1002.2581 \[hep-ph\]](#).
- [99] T. Sjöstrand, S. Mrenna, and P. Skands. “A brief introduction to PYTHIA 8.1”. In: *Computer Physics Communications* 178.11 (2008), pp. 852–867. URL: <https://sciencedirect.com/science/article/pii/S0010465508000441>.
- [100] S. Hoche et al. “NLO matrix elements and truncated showers”. In: *JHEP* 08 (2011), p. 123. arXiv: [1009.1127 \[hep-ph\]](#).
- [101] E. Bothmann et al. “Event generation with Sherpa 2.2”. In: *SciPost Phys.* 7.3 (2019), p. 034. arXiv: [1905.09127 \[hep-ph\]](#).
- [102] M. Guzzi et al. “CT10 parton distributions and other developments in the global QCD analysis”. In: (Jan. 2011). arXiv: [1101.0561 \[hep-ph\]](#).
- [103] ATLAS Collaboration. “Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”. In: *JHEP* 09 (2014), p. 145. arXiv: [1406.3660 \[hep-ex\]](#).
- [104] P. Golonka and Z. Was. “PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays”. In: *Eur. Phys. J. C* 45 (2006), pp. 97–107. arXiv: [hep-ph/0506026](#).
- [105] S. Camarda et al. “DYTurbo: fast predictions for Drell–Yan processes”. In: *The European Physical Journal C* 80.3 (Mar. 2020). URL: <https://doi.org/10.1140%2Fepjc%2Fs10052-020-7757-5>.

- [106] S. Catani and M. Grazzini. “Next-to-Next-to-Leading-Order Subtraction Formalism in Hadron Collisions and its Application to Higgs-boson Production at the Large Hadron Collider”. In: *Phys. Rev. Lett.* 98 (2007), p. 222002. arXiv: [hep-ph/0703012](#) [[hep-ph](#)].
- [107] S. Catani et al. “Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order”. In: *Phys. Rev. Lett.* 103 (8 Aug. 2009), p. 082001. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.103.082001>.
- [108] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *Eur. Phys. J. C* 75.5 (2015), p. 204. arXiv: [1412.3989](#) [[hep-ph](#)].
- [109] R. D. Ball et al. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. arXiv: [1410.8849](#) [[hep-ph](#)].
- [110] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77.7 (2017), p. 466. arXiv: [1703.10485](#) [[hep-ex](#)].
- [111] P. A. Zyla et al. “Review of Particle Physics”. In: *PTEP* 2020.8 (2020). and 2021 update, p. 083C01.
- [112] G. Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011). [Erratum: *Eur.Phys.J.C* 73, 2501 (2013)], p. 1554. arXiv: [1007.1727](#) [[physics.data-an](#)].
- [113] A. Buckley et al. “LHAPDF6: parton density access in the LHC precision era”. In: *Eur. Phys. J. C* 75 (2015), p. 132. arXiv: [1412.7420](#) [[hep-ph](#)].
- [114] T. J. Hou et al. “New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC”. In: *Phys. Rev. D* 103.1 (2021), p. 014013. arXiv: [1912.10053](#) [[hep-ph](#)].

- [115] S. Bailey et al. “Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs”. In: *Eur. Phys. J. C* 81.4 (2021), p. 341. arXiv: [2012.04684 \[hep-ph\]](#).
- [116] R. D. Ball et al. “The path to proton structure at 1% accuracy”. In: *Eur. Phys. J. C* 82.5 (2022), p. 428. arXiv: [2109.02653 \[hep-ph\]](#).
- [117] ATLAS Collaboration. “Measurement of the W^+W^- production cross section in pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV with the ATLAS experiment”. In: *Phys. Lett. B* 773 (2017), pp. 354–374. arXiv: [1702.04519 \[hep-ex\]](#).
- [118] ATLAS Collaboration. “ $ZZ \rightarrow \ell^+\ell^-\ell'^+\ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector”. In: *Phys. Rev. D* 97.3 (2018), p. 032005. arXiv: [1709.07703 \[hep-ex\]](#).
- [119] ATLAS Collaboration. “Measurement of $W^\pm Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 79.6 (2019), p. 535. arXiv: [1902.05759 \[hep-ex\]](#).
- [120] ATLAS Collaboration. “Measurements of top-quark pair to Z -boson cross-section ratios at $\sqrt{s} = 13, 8, 7$ TeV with the ATLAS detector”. In: *JHEP* 02 (2017), p. 117. arXiv: [1612.03636 \[hep-ex\]](#).
- [121] ATLAS Collaboration. “Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS”. In: *JHEP* 01 (2018), p. 063. arXiv: [1612.07231 \[hep-ex\]](#).
- [122] ATLAS Collaboration. “Measurement of single top-quark production in the s-channel in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 06 (2023), p. 191. arXiv: [2209.08990 \[hep-ex\]](#).
- [123] ATLAS Collaboration. “Measurement of t-channel production of single top quarks and antiquarks in pp collisions at 13 TeV using the full ATLAS Run 2 data sample”. In: *JHEP* 05 (2024), p. 305. arXiv: [2403.02126 \[hep-ex\]](#).

- [124] ATLAS Collaboration. “Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data”. In: *JINST* 19.02 (2024), P02009. arXiv: [2309.05471 \[hep-ex\]](#).
- [125] ATLAS Collaboration. “Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 83.8 (2023), p. 686. arXiv: [2212.07338 \[hep-ex\]](#).
- [126] J. Smith, W. L. van Neerven, and J. A. M. Vermaseren. “Transverse Mass and Width of the W Boson”. In: *Phys. Rev. Lett.* 50 (22 May 1983), pp. 1738–1740. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.50.1738>.
- [127] G. Tolkachev. “Measurements of Polarization Angular Coefficients in W-Boson Lepton Decay Processes in the ATLAS Experiment”. 2023. URL: https://indico.particle.mephi.ru/event/348/contributions/3602/attachments/2196/4006/-_29.2023_-_...pdf.
- [128] ATLAS Collaboration. “Measurement of $W^\pm$ and Z -boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 759 (2016), pp. 601–621. arXiv: [1603.09222 \[hep-ex\]](#).
- [129] C. S. Lam and W. K. Tung. “A Systematic Approach to Inclusive Lepton Pair Production in Hadronic Collisions”. In: *Phys. Rev. D* 18 (1978), p. 2447.
- [130] C. S. Lam and W. K. Tung. “Structure Function Relations at Large Transverse Momenta in Lepton Pair Production Processes”. In: *Phys. Lett. B* 80 (1979), pp. 228–231.
- [131] C. S. Lam and W. K. Tung. “A Parton Model Relation Sans QCD Modifications in Lepton Pair Productions”. In: *Phys. Rev. D* 21 (1980), p. 2712.
- [132] B. H. Bransden and C. J. Joachain. *Quantum Mechanics*. Prentice Hall, 2000. ISBN: 9780582356917.

- [133] R. M. Eisberg and R. Resnick. *Quantum physics of atoms, molecules, solids, nuclei and particles; 2nd ed.* New York, NY: Wiley, 1985. URL: <https://cds.cern.ch/record/105889>.

Appendices

APPENDIX A

LIST OF ACRONYMS, ABBREVIATIONS, AND COMMON VARIABLES

A.1 Acronyms and Abbreviations

List of abbreviations and acronyms used throughout the main text in alphabetical order:

ALFA Absolute Luminosity For ATLAS

ATLAS A Toroidal LHC Apparatus

Bkgd Background

CERN European Organization for Nuclear Research

CS Collins–Soper

CSC Cathode Strip Chamber

ECAL Electromagnetic Calorimeter

EM Electromagnetic

EMB Electromagnetic Barrel

EMEC Electromagnetic Endcap

EW Electroweak

FCal Forward Calorimeter

GUT Grand Unified Theory

HCAL Hadronic Calorimeter

HEC Hadronic Endcap

HLT High Level Trigger

IBL Insertable B-Layer

ID Inner Detector

ISR Initial State Radiation

L1 Level-1

LAr Liquid Argon

LH Likelihood

LHC Large Hadron Collider

LO Leading Order

LUCID Luminosity measurement using Cerenkov Integrating Detector

MC Monte Carlo

MDT Monitored Drift Tube

MJ Multi-jet

MS Muon Spectrometer

ndf Number of Degrees of Freedom

NLL Negative Log-Likelihood

NLO Next to Leading Order

NNLO Next to Next to Leading Order

NP Nuisance Parameter

PDF Parton Distribution Function

PFO Particle Flow Objects

PS Proton Synchrotron

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

QFT Quantum Field Theory

RPC Resistive Plate Chamber

SCT Semiconductor Tracker

SF Scale Factor

SM Standard Model

SPS Super Proton Synchrotron

SSR Shower Shape Reweighting

TGC Thin Gap Chamber

TRT Transition Radiation Tracker

VCF Variable Correction Factors

vdM van der Meer

VEV Vacuum Expectation Value

ZDC Zero Degree Calorimeter

A.2 Common Variables

List of common variables and a brief description in no particular order.

m^W - W boson mass

p_T - Transverse momentum

y^W - Rapidity

η - Pseudorapidity

m_T - Transverse mass

E_T - Transverse energy in the xy-plane within the ATLAS detector

E_T^{miss} - Missing transverse energy

p_T^ν - Neutrino transverse momentum

$|\Delta\phi|$ - Angular separation

$\sin^2 \theta_W$ - Weak mixing angle

ΣE_T - Sum of transverse energy

u_T - Hadronic recoil

α_S - Strong force coupling constant

$\langle\mu\rangle$ - Average pileup

$\sqrt{s}$ - Centre of mass energy

σ - Cross-section

$\mathcal{L}$ - Integrated luminosity

d_0 - Transverse impact parameter

d_0/σ_{d_0} - Significance of d_0

$\Delta P/P$ - Difference of momentum of track at perigee from last measurement divided by momentum at perigee

eProbabilityHT - Probability based on transition radiation in TRT

$T(ePHT)$ - Transformed eProbabilityHT with inverse sigmoid function

f_1 - Ratio of energy in first layer to total energy in EM calorimeter

f_3 - Ratio of energy in third layer to total energy in EM calorimeter

$w_{\eta 2}$ - Lateral shower width

w_{stot} - Shower width

R_η - Ratio of the energy in 3×7 cells over the energy in 7×7 cells centred at the electron cluster position

R_{had} - Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster.

R_ϕ - Ratio of the energy in 3×3 cells over the energy in 3×7 cells centred at the electron cluster position

e_{ratio} - Energy ratio between maximum and second maximum in EM calorimeter

APPENDIX B

SPHERICAL HARMONICS

The angular polynomials are related to the spherical harmonics, which are defined by:

$$Y_\ell^m(\theta, \phi) = (-1)^m \left[\frac{(2\ell + 1)(\ell - m)!}{4\pi(\ell + m)!} \right]^{1/2} P_\ell^m(\cos \theta) e^{im\phi}, \quad (\text{B.1})$$

where $P_\ell^m(\cos \theta)$ is the associated Legendre polynomials. They satisfy the orthonormality relation:

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_{\ell'}^{m'}(\theta, \phi) Y_\ell^m(\theta, \phi) = \delta_{\ell\ell'} \delta_{mm'}, \quad (\text{B.2})$$

where δ_{ab} is the Kronecker delta:

$$\delta_{ab} = \begin{cases} 0 & a \neq b \\ 1 & a = b. \end{cases} \quad (\text{B.3})$$

This relation, that the angular polynomials are orthogonal to each other, is used to extract the angular coefficients. An example of this is shown in Appendix C. The angular polynomials are related to the spherical harmonics, as listed in Table 6.1 in Section 6.1.

B.1 Associated Quantum Numbers

The angular polynomials are eigenfunctions and the number required depends on the allowed quantum numbers for each spin. The total spin of a particle – J – is the vectorial sum of the orbital angular momentum – L – and the spin angular momentum – S –, $\vec{J} = \vec{L} + \vec{S}$ [132]. The associated quantum numbers are denoted by lowercase j, ℓ, s , respectively. The quantum numbers ℓ and s are dependent on the value of the quantum number j of the particle. For the Higgs boson $j = 0$, for the W and Z bosons

$j = 1$, and a hypothetical graviton (if it had mass) would require $j = 2$. The values of s are allowed to range between negative and positive values of j in integer increments: $s \in \{-j, -j + 1, \dots, 0, j - 1, j\}$. The quantum number ℓ must then take values to satisfy $j = \ell + s$. The angular polynomials and spherical harmonics use two angles – θ and ϕ – and depend on the quantum number ℓ . Because there are two degrees of freedom with two angles, an additional quantum number is needed to describe the angular polynomials, which is where the additional magnetic quantum number, m . It can take values $m \in \{-\ell, -\ell + 1, \dots, \ell - 1, \ell\}$, is needed [133]. A summary of the allowed quantum numbers for each spin J can be found in Table B.1.

j	0	1			2				
s	0	1	0	−1	2	1	0	−1	−2
ℓ	0	0	1	2	0	1	2	3	4
m	0	0	0, ±1	0, ±1, ±2	0	0, ±1	0, ±1, ±2	0, ±1, ±2, ±3	0, ±1, ±2, ±3, ±4
Number of eigenfunctions for each ℓ	1	1	3	5	1	3	5	7	9
Number of eigenfunctions for each j	1	9			25				

Table B.1: The number of angular eigenfunctions needed to describe the angular decay angles for boson of spin $J = 0, 1, 2$ and the associated quantum numbers. The quantum numbers ℓ and m are used for the angular polynomials.

Thus, the number of polynomials required to describe the angular phase space depends on the possible values ℓ and m can take. This is why the differential cross-section requires nine angular polynomials in its decomposition in Section 6.1.

APPENDIX C

MOMENT DERIVATION

A more in-depth example of calculating the moment of $\cos \theta$ to extract A_4 will be done to represent how it is possible to measure each coefficient individually. The methodology is the same for all of the coefficients, with varying levels of integral complexity.

Recall from Section 7.1, the moment definition for a polynomial in Equation 7.2 is:

$$\langle P_i(\theta, \phi) \rangle = \frac{\int d\sigma P_i(\theta, \phi) d\cos\theta d\phi}{\int d\sigma d\cos\theta d\phi}, \quad (\text{C.1})$$

where $d\sigma$ is the differential portion of Equation 6.7:

$$d\sigma = \frac{d^3\sigma^{U+L}}{dp_T^W dm^W dy^W} \left[1 + \cos^2\theta + \sum_{i=0}^7 A_i(p_T^W, m^W, y^W) P_i(\theta, \phi) \right].$$

For the example polynomial, $\cos \theta$, it is independent of ϕ . Other polynomials are independent of $\cos \theta$ and a similar method is used. Integrating over ϕ allows for a simplification of the expression to be:

$$\int_0^{2\pi} d\sigma d\phi = \frac{3}{8} \frac{d^3\sigma^{U+L}}{dp_T^W dm^W dy^W} \left[1 + \cos^2\theta + \frac{1}{2} A_0(1 - 3\cos^2\theta) + A_4 \cos\theta \right].$$

The numerator and denominator from Equation C.1 are independent of each other so solving them separately simplifies the calculation. The denominator is then:

$$\begin{aligned} \int_{-1}^1 d\sigma d\cos\theta d\phi &= \frac{3}{8} \frac{d^3\sigma^{U+L}}{dp_T^W dm^W dy^W} \int_{-1}^1 \left[1 + \cos^2\theta + \frac{1}{2} A_0(1 - 3\cos^2\theta) + A_4 \cos\theta \right] d\cos\theta. \\ &= \frac{3}{8} \frac{d^3\sigma^{U+L}}{dp_T^W dm^W dy^W} \left[\int_{-1}^1 d\cos\theta + \int_{-1}^1 \cos^2\theta d\cos\theta + \frac{A_0}{2} \int_{-1}^1 d\cos\theta \right. \\ &\quad \left. - \frac{3A_0}{2} \int_{-1}^1 \cos^2\theta d\cos\theta + A_4 \int_{-1}^1 \cos\theta d\cos\theta \right]. \end{aligned}$$

There are three common integrals in this expression:

$$\begin{aligned}\int_{-1}^1 d \cos \theta &= 2, \\ \int_{-1}^1 \cos \theta d \cos \theta &= 0, \\ \int_{-1}^1 \cos^2 \theta d \cos \theta &= \frac{2}{3}.\end{aligned}\tag{C.2}$$

Putting these into the expression for the denominator simplifies to the expected triple differential cross-section which will be defined as $d^3\sigma$:

$$\int d\sigma d \cos \theta d\phi = \frac{d^3\sigma^{U+L}}{dp_T^W dm^W dy^W} \equiv d^3\sigma.\tag{C.3}$$

Now, looking at the numerator, ϕ can be integrated over to give:

$$\begin{aligned}\int d\sigma \cos \theta d \cos \theta d\phi &= \frac{3}{8}d^3\sigma \int_{-1}^1 \left[1 + \cos^2 \theta + \frac{1}{2}A_0(1 - 3 \cos^2 \theta) + A_4 \cos \theta \right] \cos \theta d \cos \theta. \\ &= \frac{3}{8}d^3\sigma \left[\int_{-1}^1 \cos \theta d \cos \theta + \int_{-1}^1 \cos^3 \theta d \cos \theta + \frac{A_0}{2} \int_{-1}^1 \cos \theta d \cos \theta \right. \\ &\quad \left. - \frac{3A_0}{2} \int_{-1}^1 \cos^3 \theta d \cos \theta + A_4 \int_{-1}^1 \cos^2 \theta d \cos \theta \right].\end{aligned}$$

The integrals, again, are in the form of Equation C.2, as well as an additional integral:

$$\int_{-1}^1 \cos^3 \theta d \cos \theta = 0.$$

The numerator can then be simplified to

$$\int d\sigma \cos \theta d \cos \theta d\phi = \frac{A_4}{4}d^3\sigma.\tag{C.4}$$

Putting Equation C.4 into the numerator and Equation C.3 into the denominator for the

definition of the moment in Equation 7.2 results in:

$$\begin{aligned}
 \langle \cos \theta \rangle &= \frac{\int d\sigma \cos \theta d \cos \theta d\phi}{\int d\sigma d \cos \theta d\phi} \\
 &= \frac{\frac{A_4}{4} d^3 \sigma}{d^3 \sigma} \\
 &= \frac{A_4}{4}.
 \end{aligned}
 \tag{C.5}$$

A similar method can be used for the moments of the other polynomials, where the results are listed in Chapter 7.1.

APPENDIX D

FOLDING

In general, truth observables – t – distributed by function – $f(t)$ – are folded through a probability distribution – $p(r|t)$ – into the reconstructed observables – r – and are distributed by function $g(r)$ given by Equation D.1:

$$g(r) = \int f(t) p(r|t) dt. \quad (\text{D.1})$$

The convolution, $p(r|t)$, is the probability to observe truth, t , from reconstructed, r . Similarly, the trigonometric polynomials can be folded. In this case, the truth observables are $t = (\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}, p_{\text{T}}^{W, \text{Truth}})$ and are folded into the reconstructed observables – which are $r = (\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}}, p_{\text{T}}^{l\nu})$. Using these elements, an expression for the i th folded polynomial in the j th p_{T}^W bin $\tilde{P}_{ij}$ can be obtained:

$$\tilde{P}_{ij} \left(r | p_{\text{T}}^{W, \text{Truth}} \in \left(\Delta p_{\text{T}}^{W, \text{Truth}} \right)_j \right) = \int_{\left(\Delta p_{\text{T}}^{W, \text{Truth}} \right)_j} \int_{\cos \theta_{\text{CS}}^{\text{Truth}}} \int_{\phi_{\text{CS}}^{\text{Truth}}} P_i (\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}) p(r|t) dt \quad (\text{D.2})$$

The folded polynomials, $\tilde{P}_{ij}$, are discretely binned over a finite range of $p_{\text{T}}^{W, \text{Truth}}$, expressed as $\left(\Delta p_{\text{T}}^{W, \text{Truth}} \right)_j$. To get an expression for the templates, MC is used to estimate $p(r|t)$. The moments derived in Section 7.1 are taken to build the two dimensional phase space of $\cos \theta_{\text{CS}}$ and ϕ_{CS} in the j th bin of p_{T}^W . This is similar to the truth phase space in Figure 7.1 – denoted by $f_j (\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}})$, which has the form:

$$f_j (\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}) = \sigma_j \left\{ P_8 (\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}) + \sum_{i=0}^7 A_{ij}^{\text{ref}} P_i (\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}}) \right\}. \quad (\text{D.3})$$

In Equation D.3, σ_j is the differential cross section in the j th p_{T}^W bin and $P_8 = 1 + \cos^2 \theta$.

Using $f_j(\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}})$, an estimate of $p(r|t)$ can be constructed:

$$p^{MC}(r|t) = \frac{w_{\text{event}}(r, t)}{f_j(\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}})}, \quad (\text{D.4})$$

where an event based weight, $w_{\text{event}}(r, t)$, is the product of scale factors to compensate for the inefficiencies of the detector. The definition can be found in Equation 8.1, which includes: generator event weight, isolation, identification, trigger, and reconstruction efficiency.

The final template is then summed over all events in the discrete ranges of $(\Delta p_{\text{T}}^{W, \text{Truth}})_j$, $(\Delta p_{\text{T}}^{l\nu})_j$, $(\Delta \cos \theta_{\text{CS}}^{\text{Reco}})_j$, and $(\Delta \phi_{\text{CS}}^{\text{Reco}})_j$ denoted by Δ_j . The templates are expressed as:

$$T_{ij} = \sum_{\text{event} \in \Delta_j} P_i(\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}}) \frac{w_{\text{event}}(r, t)}{f_j(\cos \theta_{\text{CS}}^{\text{Truth}}, \phi_{\text{CS}}^{\text{Truth}})}. \quad (\text{D.5})$$

This is done for the templates described in Section 7.3 that are used in the likelihood function for Equation 7.8 in Section 7.4.

APPENDIX E

MONTE CARLO SAMPLES

The Monte Carlo (MC) samples generated by the ATLAS Collaboration used in this analysis are listed in Table E.1. The production methods for the MC samples are discussed in Section 8.2. A double horizontal rule line is used to separate samples into types in the order: signal, single boson background, diboson background, and top related background. For a given signal process, the other signal processes are included as a single boson background but largely do not contribute. The cross-sections quoted in the table are used to normalize predicted event counts for the various samples and the respective uncertainties are listed, as well.

Process	Generator	$\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$ [nb] (unc.)	$N_{\text{evt}}^{\text{skim}}$	$N_{\text{evt}}^{\text{unskim}}$
$W^+ \rightarrow e^+ \nu$	POWHEG+PYTHIA8	11.61 (5%)	40×10^6	40×10^6
$W^+ \rightarrow \mu^+ \nu$	POWHEG+PYTHIA8	11.61 (5%)	40×10^6	40×10^6
$W^- \rightarrow e^- \bar{\nu}$	POWHEG+PYTHIA8	8.630 (5%)	30×10^6	30×10^6
$W^- \rightarrow \mu^- \bar{\nu}$	POWHEG+PYTHIA8	8.630 (5%)	29×10^6	29×10^6
$W^+ \rightarrow \tau^+ \nu$	POWHEG+PYTHIA8	11.61 (5%)	0.28×10^6	5.0×10^6
$W^- \rightarrow \tau^- \bar{\nu}$	POWHEG+PYTHIA8	8.630 (5%)	0.24×10^6	4.0×10^6
$Z \rightarrow e^+ e^-$	POWHEG+PYTHIA8	1.958 (5%)	10×10^6	10×10^6
$Z \rightarrow \mu^+ \mu^-$	POWHEG+PYTHIA8	1.958 (5%)	10×10^6	10×10^6
$Z \rightarrow \tau^+ \tau^-$	POWHEG+PYTHIA8	1.958 (5%)	0.12×10^6	1.0×10^6
$ZZ(q\bar{q}\ell\ell)$	SHERPA 2.2.1	0.002194 (10%)	6.4×10^3	10×10^3
$WZ(q\bar{q}\ell\ell)$	SHERPA 2.2.1	0.003433 (10%)	6.3×10^3	10×10^3
$WW(q\bar{q}\ell\nu)$	SHERPA 2.2.1	0.02472 (10%)	9.3×10^3	20×10^3
$WW(\ell\nu q\bar{q})$	SHERPA 2.2.1	0.02472 (10%)	9.3×10^3	20×10^3
$WZ(\ell\nu q\bar{q})$	SHERPA 2.2.1	0.01142 (10%)	4.7×10^3	10×10^3
$ZZ(4\ell)$	SHERPA 2.2.2	0.001252 (10%)	5.7×10^3	10×10^3
$WZ(3\ell\nu)$	SHERPA 2.2.2	0.004583 (10%)	6.2×10^3	10×10^3
$WW(2\ell 2\nu)$	SHERPA 2.2.2	0.01250 (10%)	7.3×10^3	10×10^3
$WZ(\ell 3\nu)$	SHERPA 2.2.2	0.003235 (10%)	5.0×10^3	10×10^3
Wt	POWHEG+PYTHIA8	0.03582 (10%)	3.7×10^3	10×10^3
$W\bar{t}$	POWHEG+PYTHIA8	0.03399 (10%)	3.7×10^3	10×10^3
$t\bar{t}$ (nominal)	POWHEG+PYTHIA8	0.4525 (7%)	1.2×10^6	2.0×10^6
$t(t - \text{chan.}t)$	POWHEG+PYTHIA8	0.03699 (10%)	16×10^3	30×10^3
$t(t - \text{chan.}\bar{t})$	POWHEG+PYTHIA8	0.02217 (10%)	11×10^3	20×10^3
$t(s - \text{chan.}t)$	POWHEG+PYTHIA8	0.002027 (10%)	5.0×10^3	10×10^3
$t(s - \text{chan.}\bar{t})$	POWHEG+PYTHIA8	0.001268 (10%)	5.2×10^3	10×10^3

Table E.1: The Monte Carlo (MC) samples at $\sqrt{s} = 13$ TeV for the low- $\langle\mu\rangle$ data set. Given is: a short description of the process, the names and version numbers of the MC generator(s), the used value of the higher order cross section times any branching and filter efficiencies ($\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$) – with the theoretical uncertainty in percent (“unc.”) –, the number of events analyzed that pass the trigger threshold – $N_{\text{evt}}^{\text{skim}}$ –, and the number of originally processed and simulated events – $N_{\text{evt}}^{\text{unskim}}$. In the case of $Z \rightarrow \ell\ell$ samples, there is $\epsilon_{\text{filter}} > 1$ due to the fact the cross sections were calculated for $66 < m^{\ell\ell} < 116$ GeV, but the generated mass range is larger. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}}/(\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}})$ is generally above 3 fb^{-1} for signal and significant backgrounds. The exception are POWHEG+PYTHIA8 $W \rightarrow \tau\nu$ and $Z \rightarrow \tau^+\tau^-$ samples, that have about 0.45 fb^{-1} only.

APPENDIX F

SIGN AMBIGUITY

As discussed in Section 8.4, taking the mass constraint in order get p_z^ν results in a quadratic equation with a two equally viable solutions. The two real solutions can be used to calculate different values of $\cos \theta_{CS}$ and cause a sign ambiguity in $\cos \theta_{CS}$. This appendix will discuss how choosing a solution at random can allow for statistically resolving the angular distributions and allow statistical sensitivity to measure A_4 , which depends on having the correct sign of $\cos \theta_{CS}$.

In Equation 6.8, for the definition of $\cos \theta_{CS}$, there are two parts that contribute to the sign ambiguity. The first is $p_z^{\ell\nu}/|p_z^{\ell\nu}|$, which can be rewritten as the sign of $y^{l\nu}$, and the second contribution is $\Delta P_\ell = (P_1^+ P_2^- - P_1^- P_2^+)$. The main source of solving the ambiguity in $\cos \theta_{CS}$, which happens $\sim 25\%$ of the time, is when the two solutions have opposite signs in $y^{l\nu}$. This causes a sign flip in ΔP_ℓ and both $\cos \theta_{CS}$ solutions are equal. This happens predominately when the W boson has low rapidity values of $y^W < 2$.

This sign flip can be understood by looking at the $W^- \rightarrow e^- \bar{\nu}$ case. In this case P_1 is fixed because the lepton is measured so ΔP_ℓ can be simplified to

$$\Delta P_\ell = P_1^+ (E^\nu - p_z^\nu) - P_1^- (E^\nu + p_z^\nu). \quad (\text{F.1})$$

If ΔP_ℓ flips sign then:

$$\begin{aligned} E^\nu - p_z^\nu > E^\nu + p_z^\nu &\Rightarrow p_z^\nu < 0 \quad \text{for solution 1,} \\ E^\nu - p_z^\nu < E^\nu + p_z^\nu &\Rightarrow p_z^\nu > 0 \quad \text{for solution 2.} \end{aligned} \quad (\text{F.2})$$

The sign flip of $y^{l\nu}$ can be understood by looking at the definition of y :

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (\text{F.3})$$

where if $y^{l\nu}$ flips sign then:

$$\begin{aligned} E^{\ell\nu} + p_z^{\ell\nu} &> E^{\ell\nu} - p_z^{\ell\nu} \text{ for solution 1,} \\ E^{\ell\nu} + p_z^{\ell\nu} &< E^{\ell\nu} - p_z^{\ell\nu} \text{ for solution 2.} \end{aligned} \quad (\text{F.4})$$

The sign flip happens mostly for central $y^{l\nu}$ due to fact the p_z^ν sign flip is more likely to occur when $p_z^\ell + p_z^\nu$ is close to zero. This can be seen by using $p_z^{\ell\nu} = p_z^\ell + p_z^\nu$ in Equation F.4, which can be rearranged into:

$$\begin{aligned} p_z^\ell + p_z^\nu &> 0 \text{ for solution 1,} \\ p_z^\ell + p_z^\nu &< 0 \text{ for solution 2.} \end{aligned} \quad (\text{F.5})$$

The other time that the sign ambiguity can be resolved is when one of the solutions violates the Bjorken condition of $x < 1$ and is unphysical. Here, x is the fraction of momentum the interacting parton of the total proton momentum. A breaking of this Bjorken condition would require either $|p_z^\nu| > \sqrt{s}/2$ or $|p_z^\nu + p_z^\ell| > \sqrt{s}/2$, which is when the solution causes the longitudinal momenta of the neutrino or the W boson to exceed the beam energy. This condition breaking only occurs for $< 1\%$ of all events.

To further understand the sensitivity, it is important to look at what happens when the right and wrong p_z^ν solution are chosen, where the right p_z^ν is when the solution is closest to the truth p_z^ν . Both right and wrong solution choices happen an approximately equal amount. There are four possible combinations of whether p_z^ν is correct and if the sign of $\cos \theta_{CS}$ is correct, as summarized in Figure F.1. The first is when the correct p_z^ν is chosen and get the correct sign of $\cos \theta_{CS}$ as a result. These events are most important for sensitivity at large rapidity, while the correct sign of $\cos \theta_{CS}$ for more central rapidity is determined, as

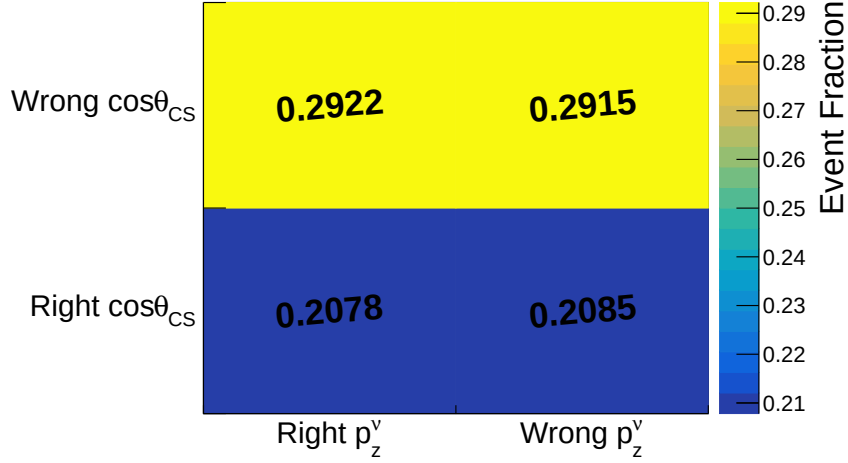


Figure F.1: Choice combination matrix for choosing the correct p_z^ν solution and whether the right sign of $\cos\theta_{CS}$ is obtained. The right sign of $\cos\theta_{CS}$ is what increases the sensitivity to A_4 , while the wrong sign decreases it.

discussed earlier. Second are events where the correct p_z^ν is chosen but due to the limited resolution the sign of $\cos\theta_{CS}$ is flipped. This does not occur often and only when y^W is close to zero. The events that cause the most sensitivity gain to A_4 is when the wrong p_z^ν is chosen but the sign of y^{ν} also flipped, allowing for the right sign of $\cos\theta_{CS}$. This mostly happens for $|y^{\nu}| < 1$. Lastly, are events where the wrong p_z^ν is chosen resulting in the wrong sign of $\cos\theta_{CS}$, which only dilute the sensitivity and occur mostly at high y^{ν} .

To show the statistical sensitivity gained, A_4 was extracted dependent on the ambiguity level of events. First are events that have both an ambiguity in the sign of $\cos\theta_{CS}$ and y^{ν} , where there is only a 50% chance to get the correct sign. These can be seen in Figure F.2 in black circles that correspond to approximately 100% error. Next, the previous events are combined with the events that only have a sign ambiguity in y^{ν} , where the right sign of $\cos\theta_{CS}$ is chosen every time. This is illustrated in blue squares in Figure F.2, which shows the uncertainty reduced by a significant amount. Lastly, the events where one solution violates the Bjorken $x < 1$ condition there is a very small gain in sensitivity increase, as seen in red triangles in Figure F.2. Figure F.2a) and c) also show that there is no bias introduced from this ambiguity.

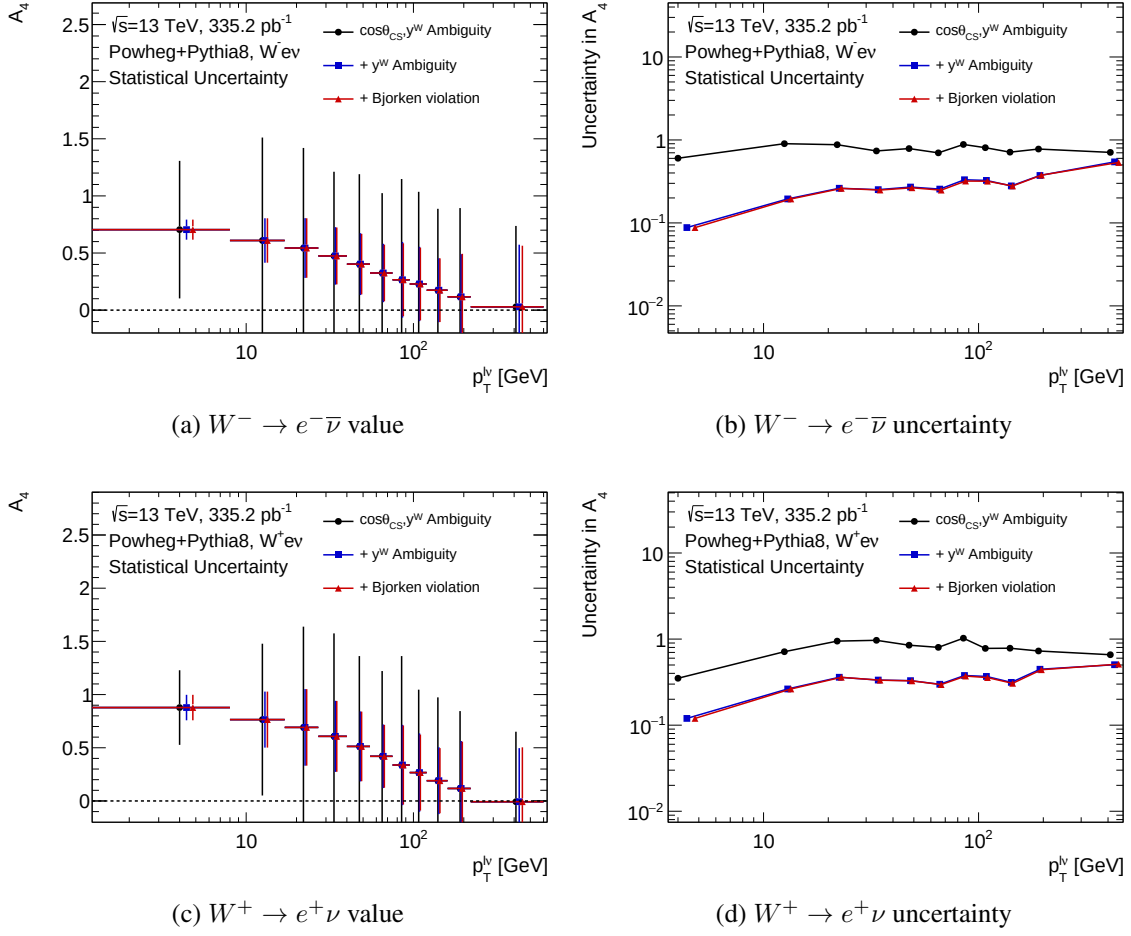


Figure F.2: Comparison for A_4 for the inclusion of events dependent on the ambiguity in the sign of $\cos \theta_{CS}$ and y^W in $p_T^{l\nu}$ binning for $W^- \rightarrow e^- \bar{\nu}$ and $W^+ \rightarrow e^+ \nu$. a) and c) shows the values are not biased by the ambiguity. b) and d) show the statistical uncertainty is reduced by including events that only have ambiguity in y^W .

APPENDIX G

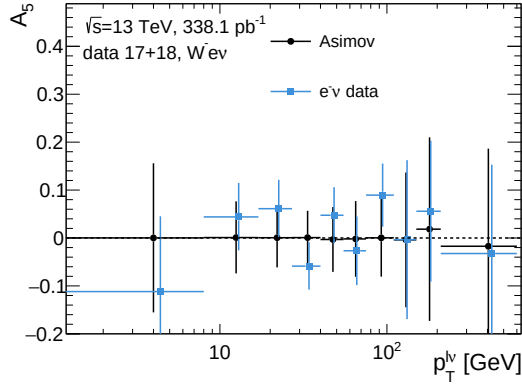
ADDITIONAL FIGURES

This Appendix will include figures that were not included in Chapters 8-11 to avoid repetition, by showing similar distributions for all four channels. Included are the pre-fit kinematic distributions not shown in Section 8.7. For the $W^- \rightarrow e^- \bar{\nu}$ channel, this corresponds to just the two-dimensional $(\cos \theta_{\text{CS}}^{\text{Reco}}, \phi_{\text{CS}}^{\text{Reco}})$ distributions as a function of $p_{\text{T}}^{l\nu}$ for the remaining $p_{\text{T}}^{l\nu}$ bins not shown in Figures 8.12 and 8.13. For the remaining three channels: $W^+ \rightarrow e^+ \nu$, $W^- \rightarrow \mu^- \bar{\nu}$, $W^+ \rightarrow \mu^+ \nu$, this includes all of the pre-fit kinematic distributions. The post-fit unfolded three-dimensional $(\cos \theta_{\text{CS}}, \phi_{\text{CS}}, p_{\text{T}}^{l\nu})$ distribution unwrapped into a one-dimensional distribution from Figure 9.2 is shown for the remaining three channels.

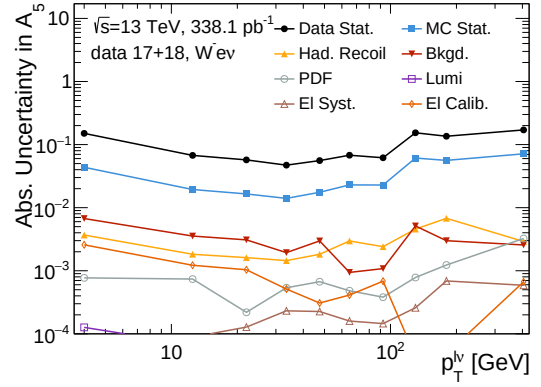
This appendix shows distributions and uncertainties for the measured coefficients and differential cross-section with comparisons to Asimov data. These figures include the coefficients A_1 , and A_5 to A_7 , which do not have the statistical sensitivity to measure but are still extracted during the fitting process and are shown for all four channels as well as the measurements of the combined channels. The final types of plots shown are the post-fit systematic uncertainties NP pull for the other three channels as well as for the fit of the combined channels. All of the figures are organized by channel in the order of $W^- \rightarrow e^- \bar{\nu}$ in Section G.1, $W^+ \rightarrow e^+ \nu$ in Section G.2, $W^- \rightarrow \mu^- \bar{\nu}$ in Section G.3, $W^+ \rightarrow \mu^+ \nu$ in Section G.4, combined W^- in Section G.5, and combined W^+ in Section G.6.

G.1 Additional W^- Electron Channel Figures

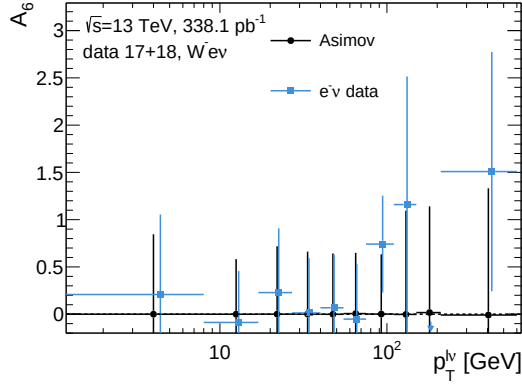
The $W^- \rightarrow e^- \bar{\nu}$ channel was used as the main example throughout most of this thesis, so only a few figures are included in this section.



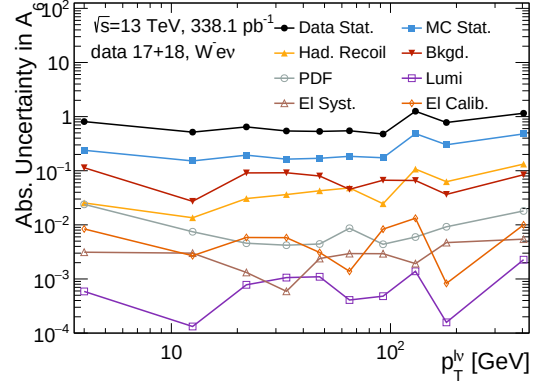
(a) A_5



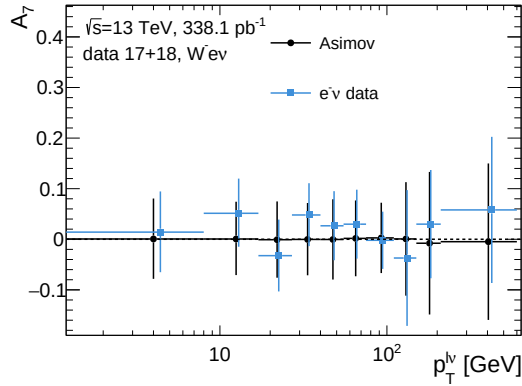
(b) Uncertainties for A_5



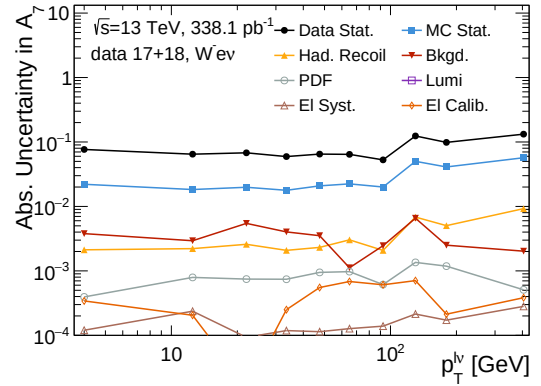
(c) A_6



(d) Uncertainties for A_6

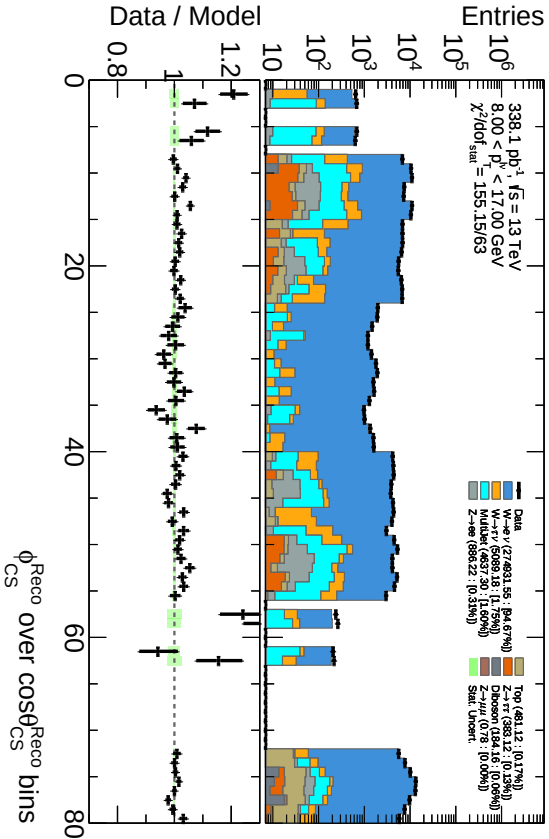


(e) A_7

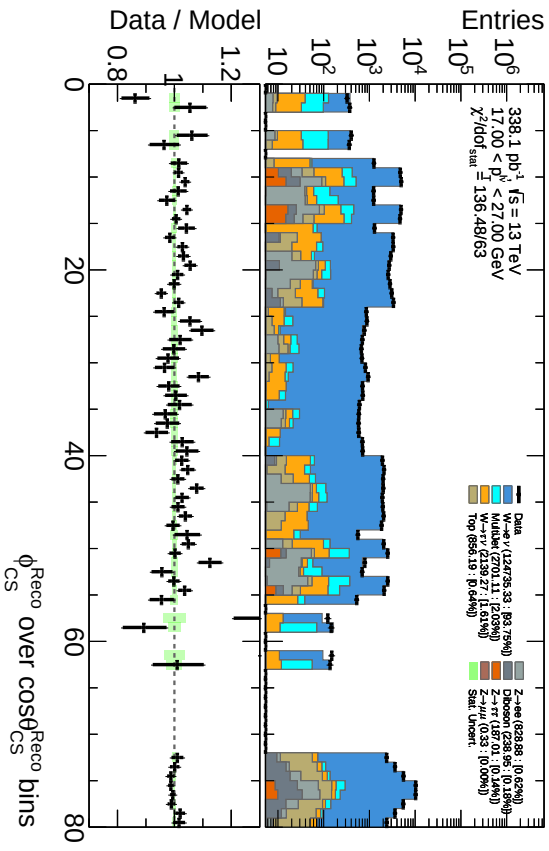


(f) Uncertainties for A_7

Figure G.1: The measured last three coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^- \rightarrow e^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



(a) $8 < p_T^{l\nu} < 17$ GeV



(b) $17 < p_T^{l\nu} < 27$ GeV

Figure G.2: The two-dimensional ($\cos\theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $8 < p_T^{l\nu} < 17$ GeV and b) $17 < p_T^{l\nu} < 27$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

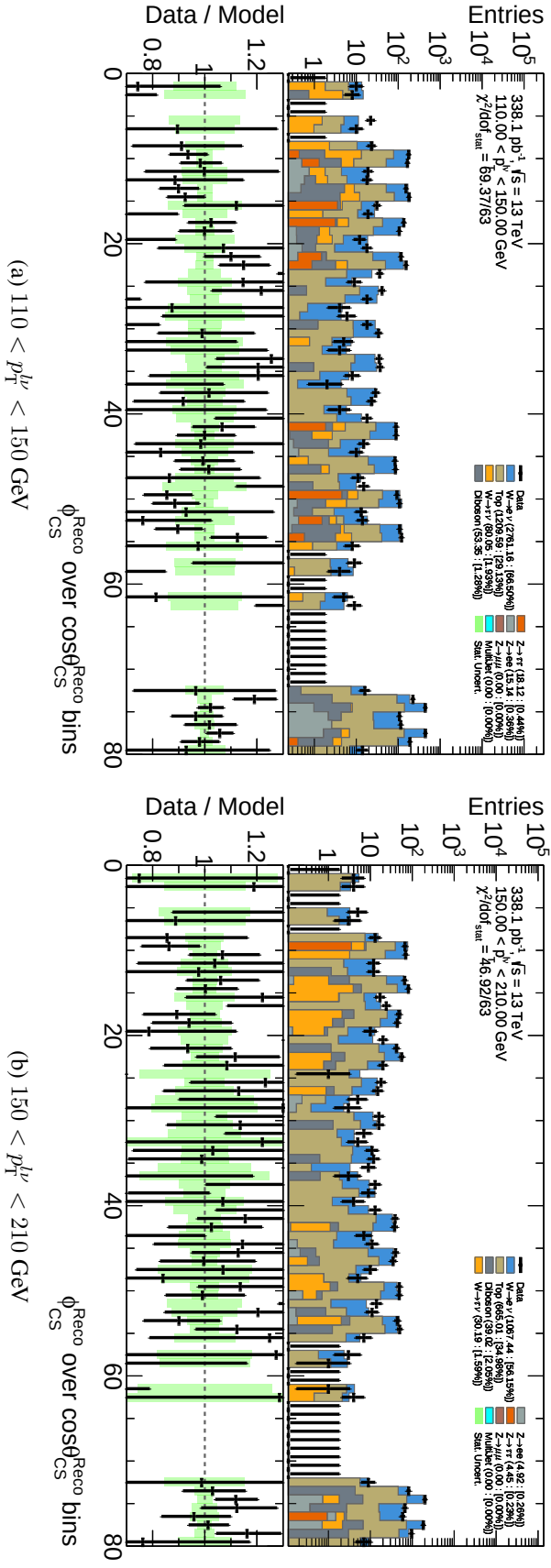


Figure G.4: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^- \rightarrow e^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $110 < p_T^{l\nu} < 150 \text{ GeV}$ and b) $150 < p_T^{l\nu} < 210 \text{ GeV}$. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

G.2 Additional W^+ Electron Channel Figures

Chapters 8-11 usually only included figures for the $W^- \rightarrow e^- \bar{\nu}$ channel. This section will contain the same pre-fit kinematic distributions for the $W^+ \rightarrow e^+ \nu$ channel, measured values of the Drell–Yan angular coefficients and differential cross-section compared to Asimov data, the post-fit unfolded three-dimensional $(\cos \theta_{CS}, \phi_{CS}, p_T^{\nu})$ distribution unwrapped into a one-dimensional distribution, and the plot of the NP pulls.

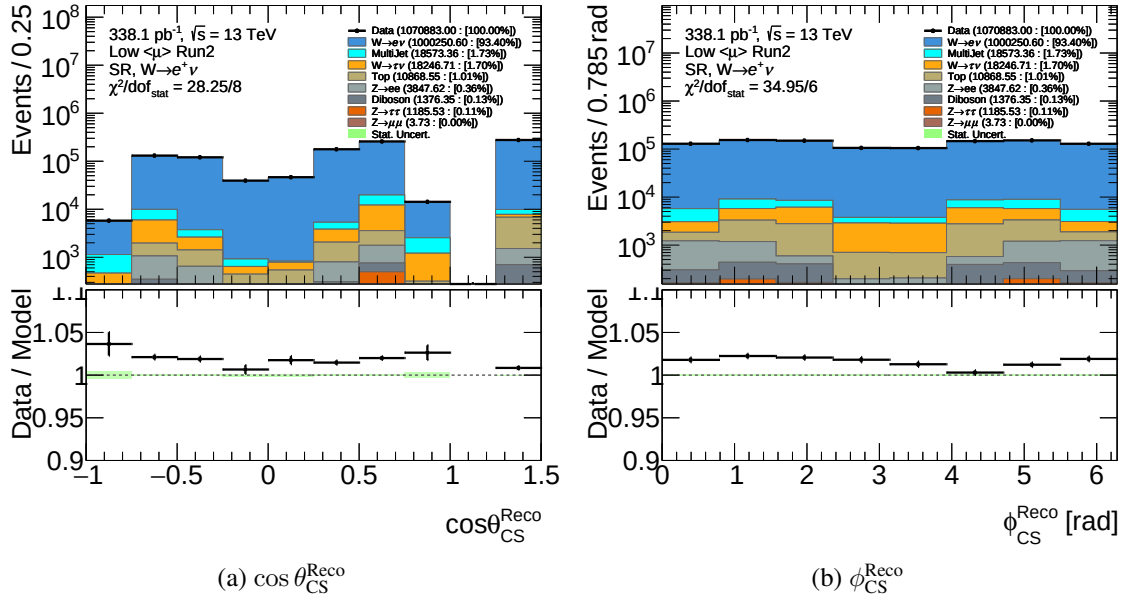


Figure G.5: The integrated CS angular distributions of a) $\cos \theta_{CS}^{\text{Reco}}$ and b) ϕ_{CS}^{Reco} for $W^+ \rightarrow e^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

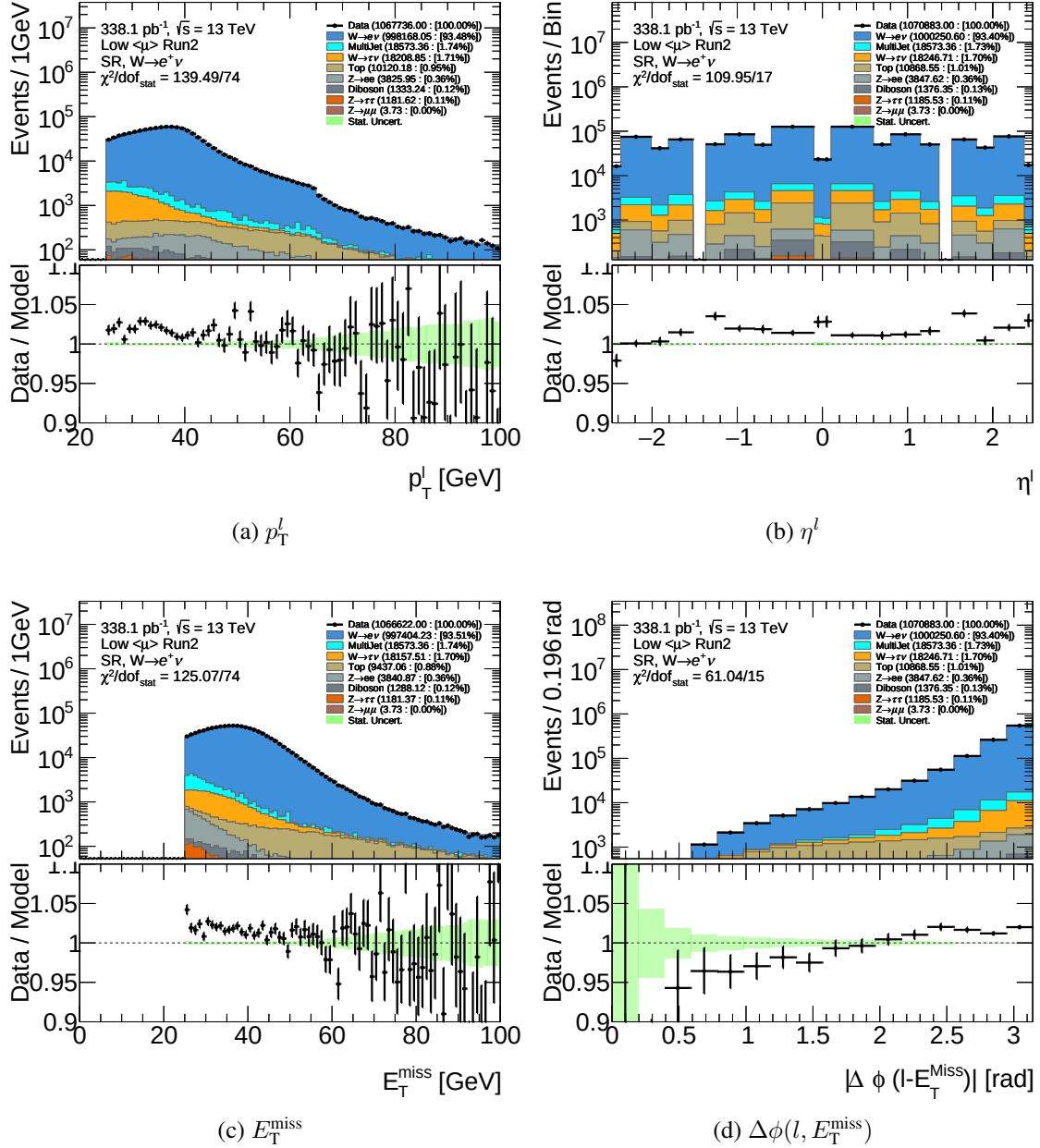


Figure G.6: The kinematic distributions for a) p_T^l , b) η^l , c) E_T^{miss} , and d) $\Delta\phi(l, E_T^{\text{miss}})$ for $W^+ \rightarrow e^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

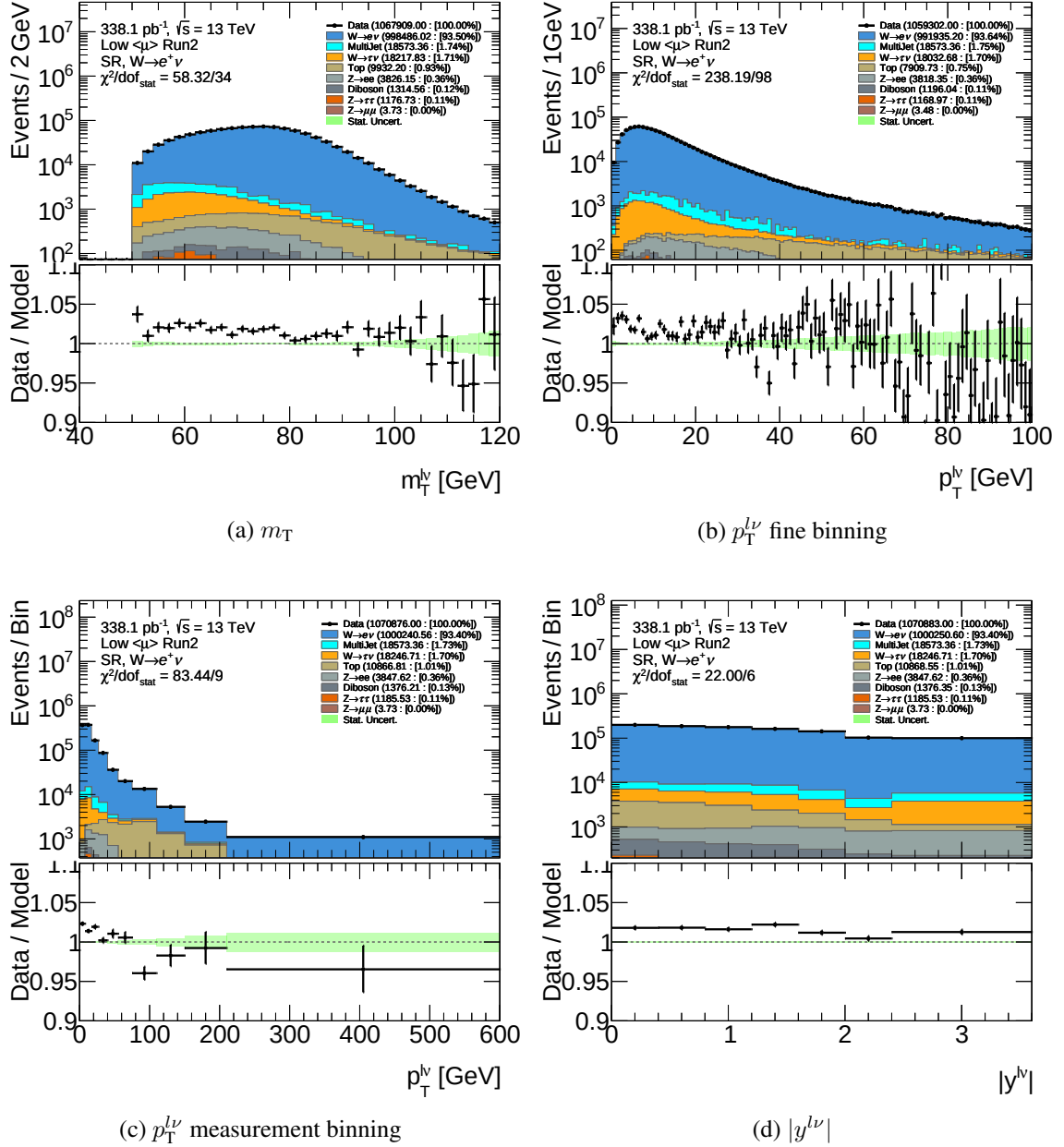


Figure G.7: The kinematic distributions for a) m_T , b) $p_T^{l\nu}$ in a fine binning, c) $p_T^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^+ \rightarrow e^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

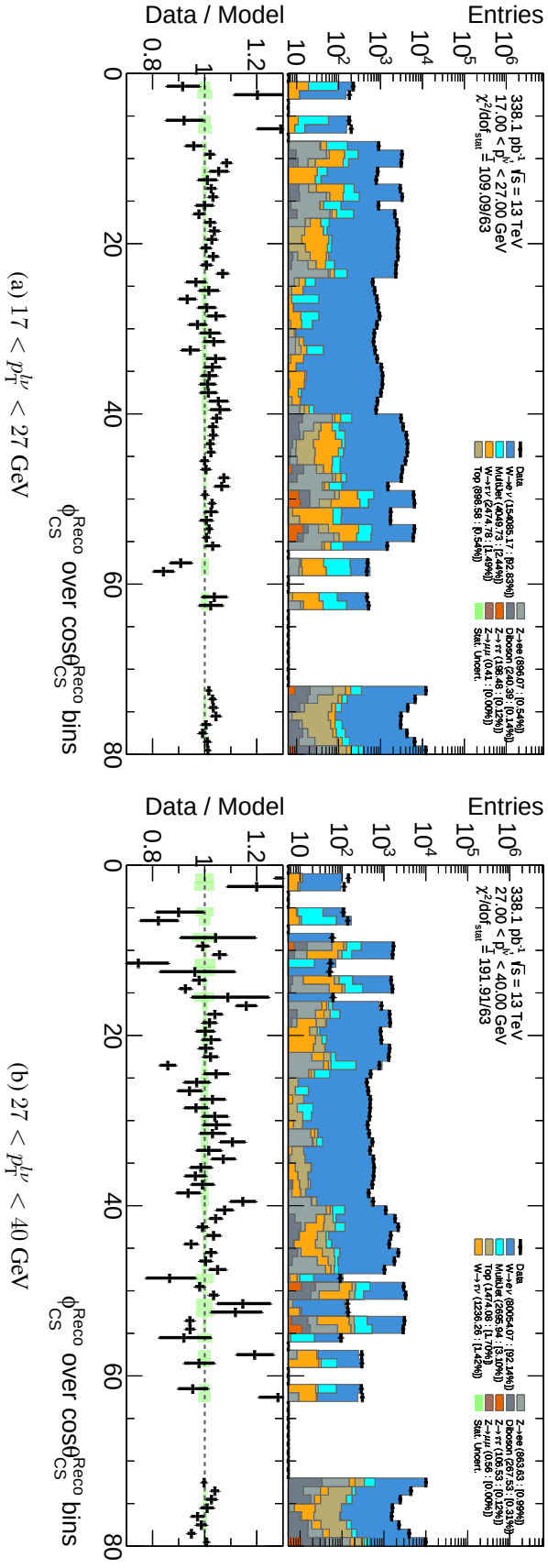


Figure G.9: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of p_T^{ν} for $W^+ \rightarrow e^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $17 < p_T^{\nu} < 27 \text{ GeV}$ and b) $27 < p_T^{\nu} < 40 \text{ GeV}$. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

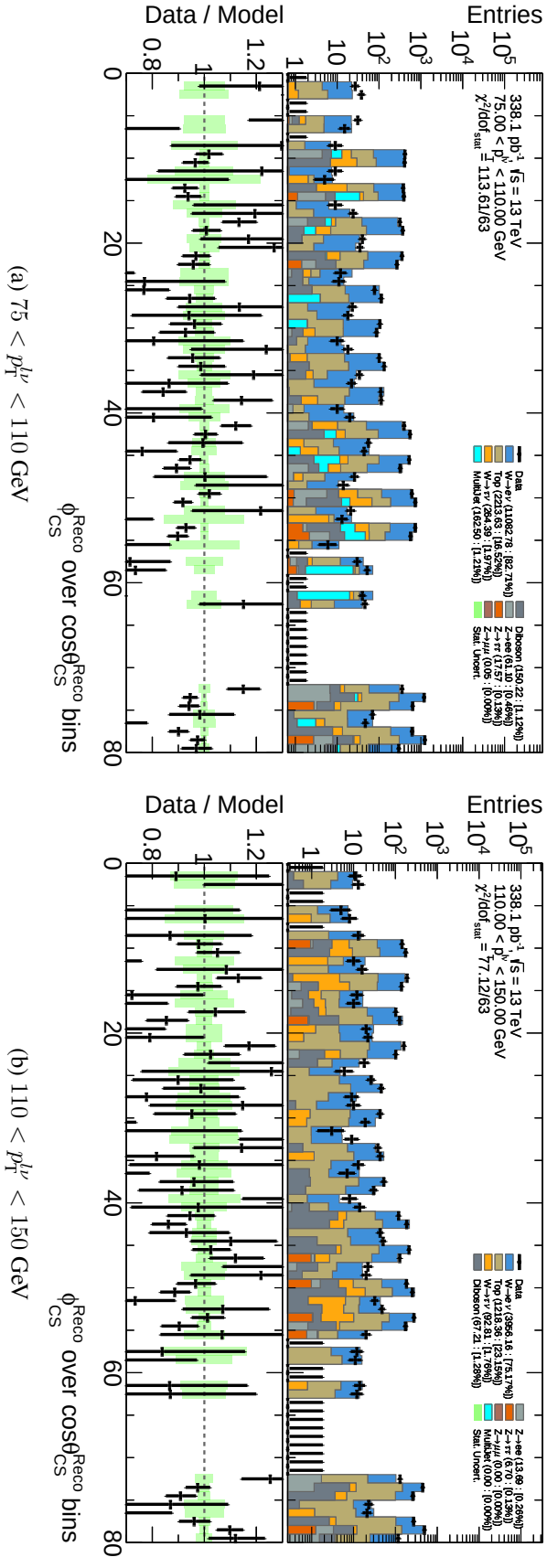
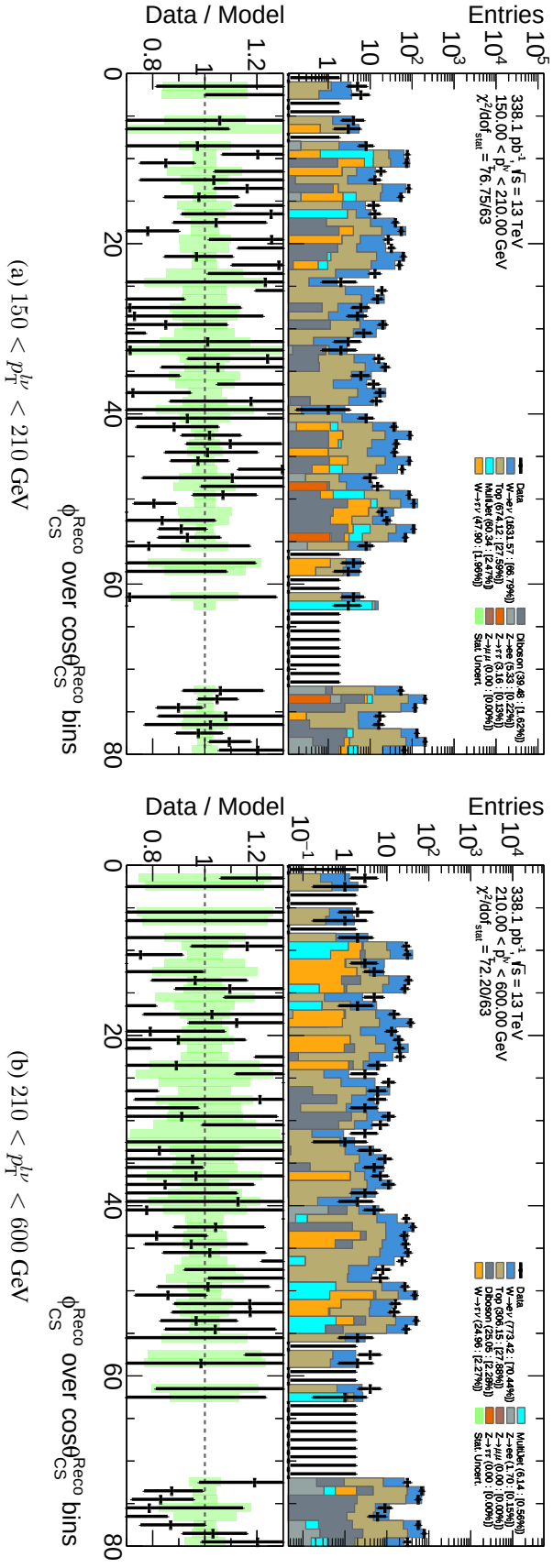
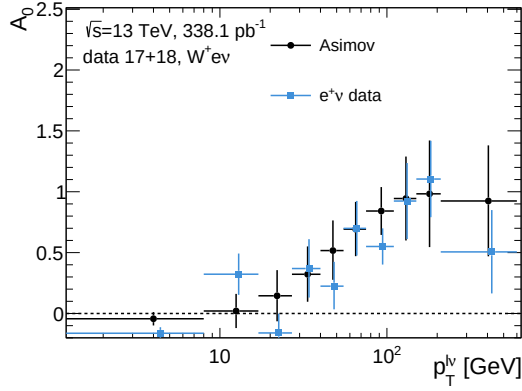
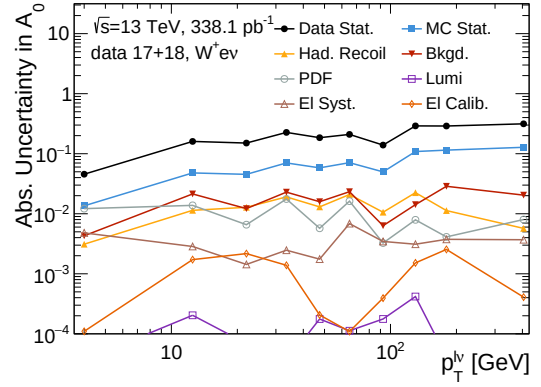


Figure G.11: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of p_T^{le} for $W^+ \rightarrow e^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_T^{\text{le}} < 110 \text{ GeV}$ and b) $110 < p_T^{\text{le}} < 150 \text{ GeV}$. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

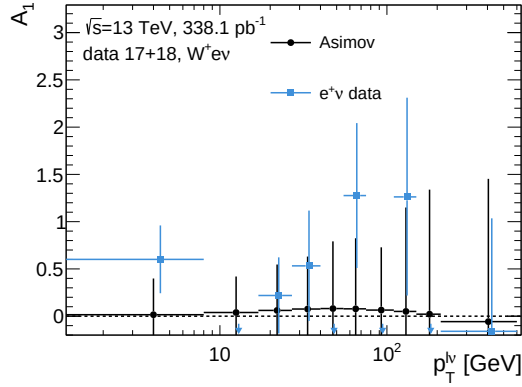




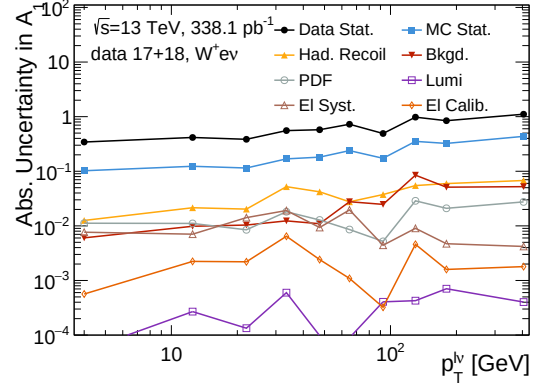
(a) A_0



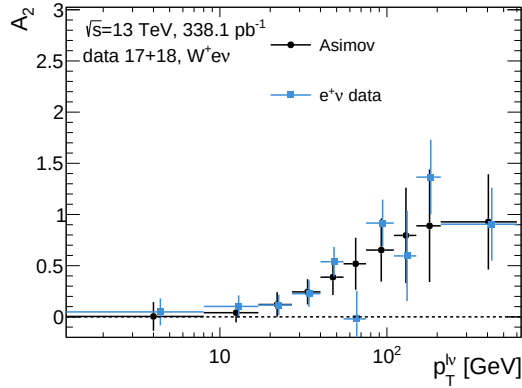
(b) Uncertainties for A_0



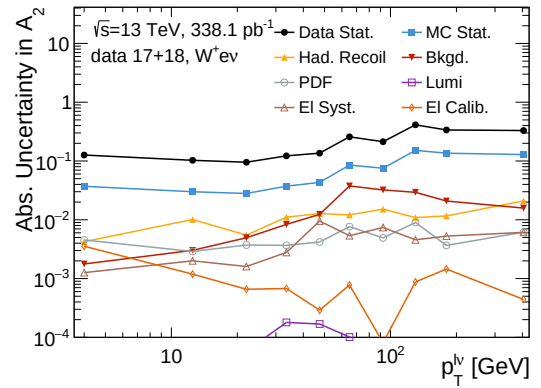
(c) A_1



(d) Uncertainties for A_1

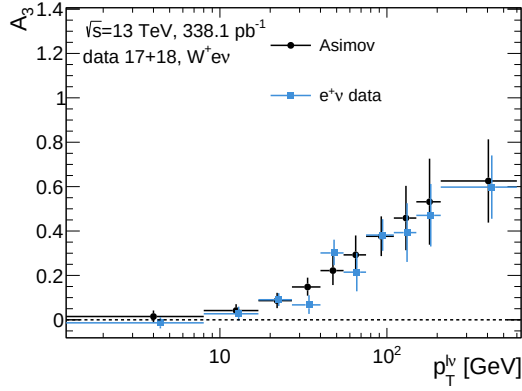


(e) A_2

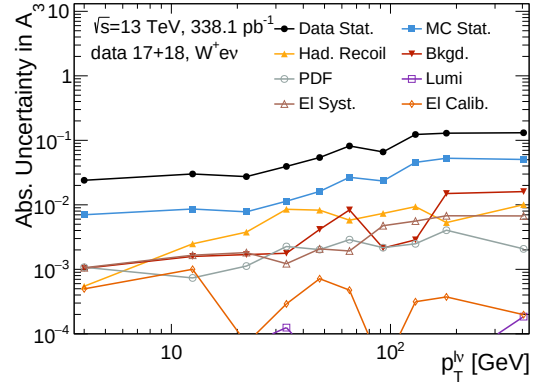


(f) Uncertainties for A_2

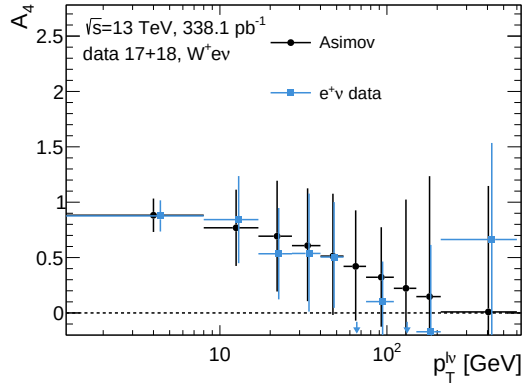
Figure G.13: The measured coefficients – A_0 , A_1 , A_2 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow e^+\nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



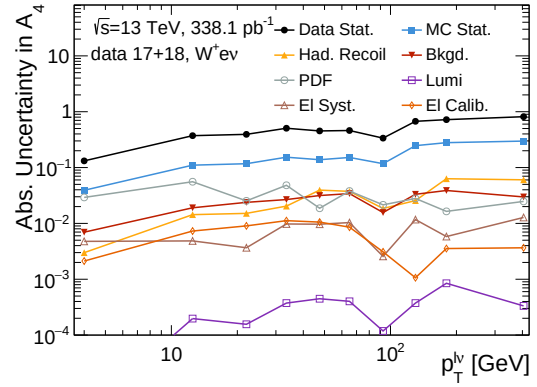
(a) A_3



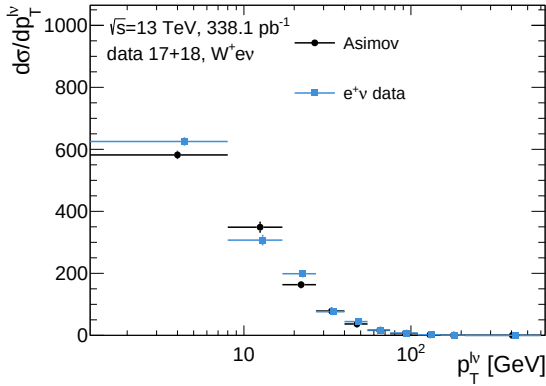
(b) Uncertainties for A_3



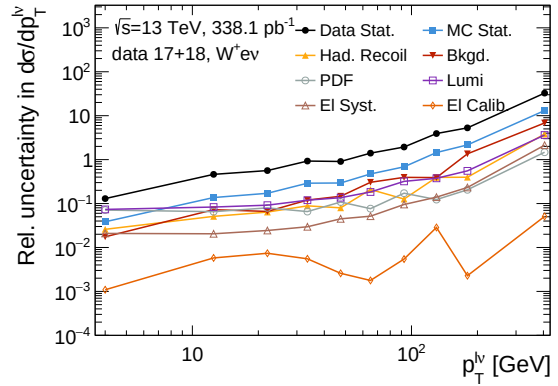
(c) A_4



(d) Uncertainties for A_4

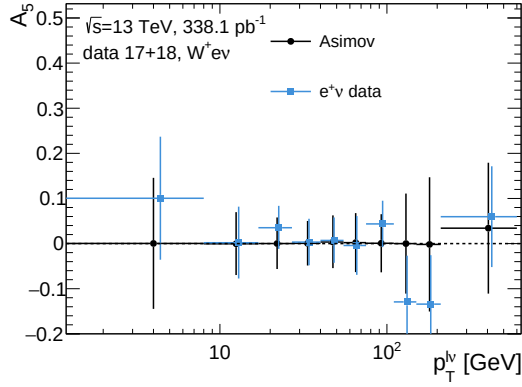


(e) Differential cross-section

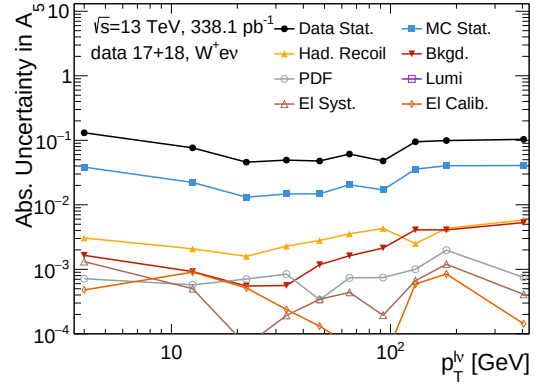


(f) Uncertainties for the differential cross-section

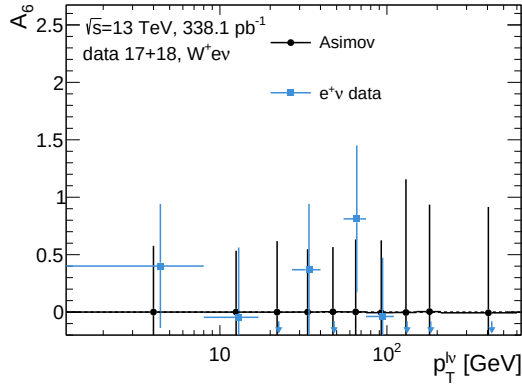
Figure G.14: The measured coefficients – A_3 and A_4 –, the differential cross-section, and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow e^+\nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



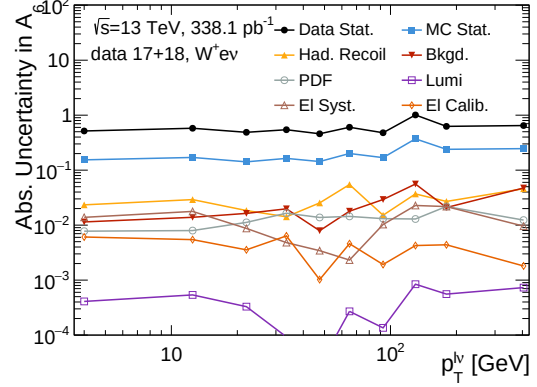
(a) A_5



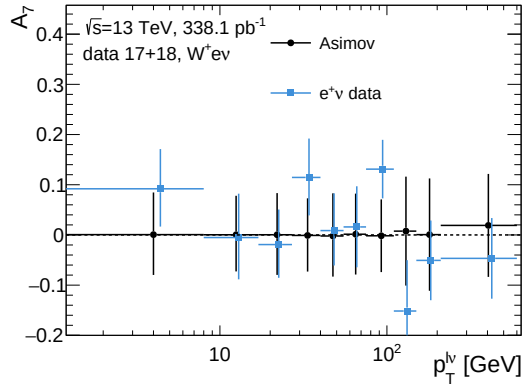
(b) Uncertainties for A_5



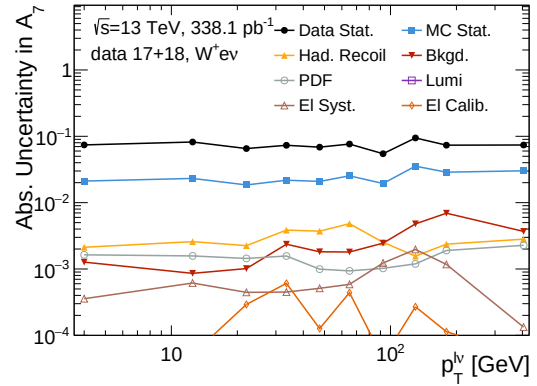
(c) A_6



(d) Uncertainties for A_6

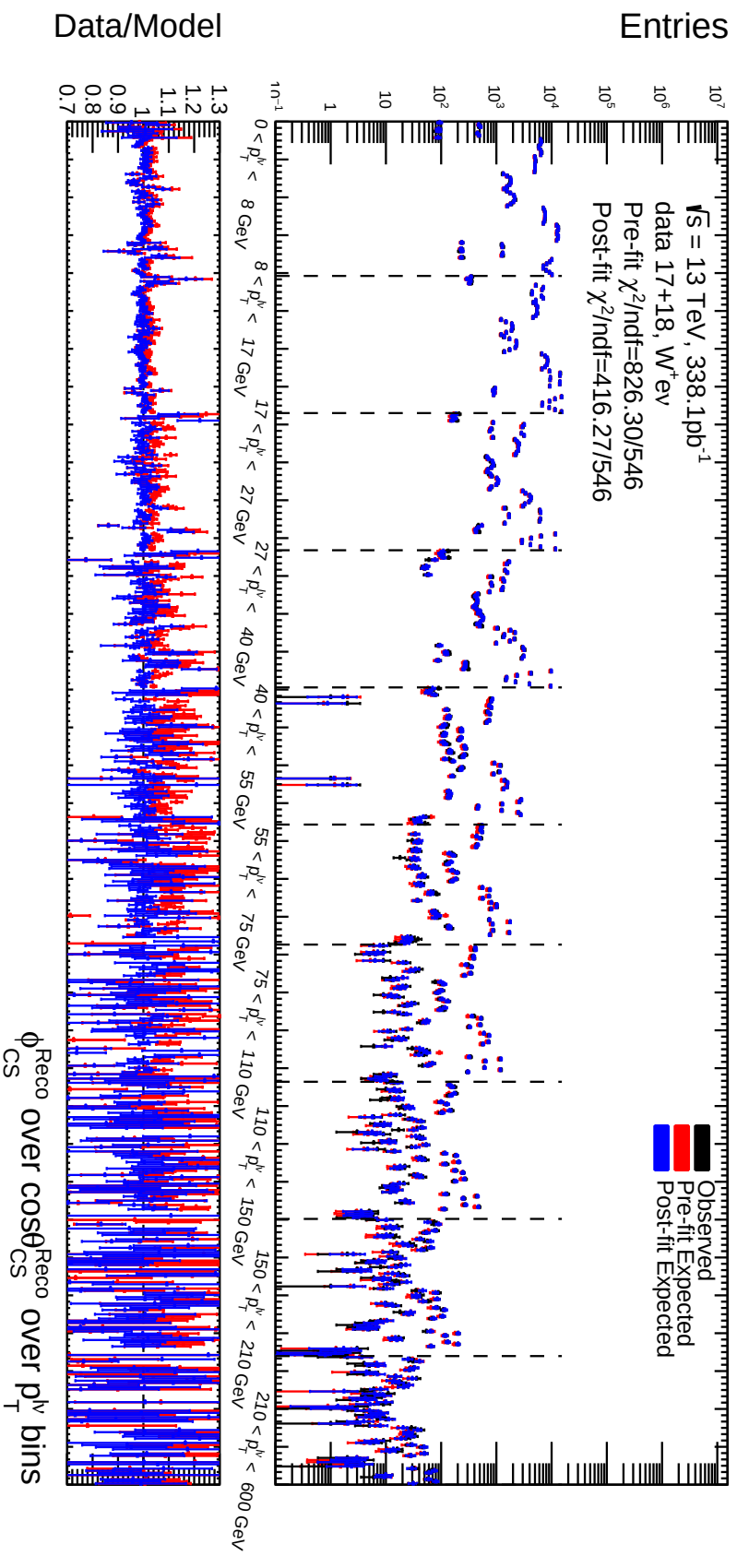


(e) A_7



(f) Uncertainties for A_7

Figure G.15: The measured coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow e^+ \nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



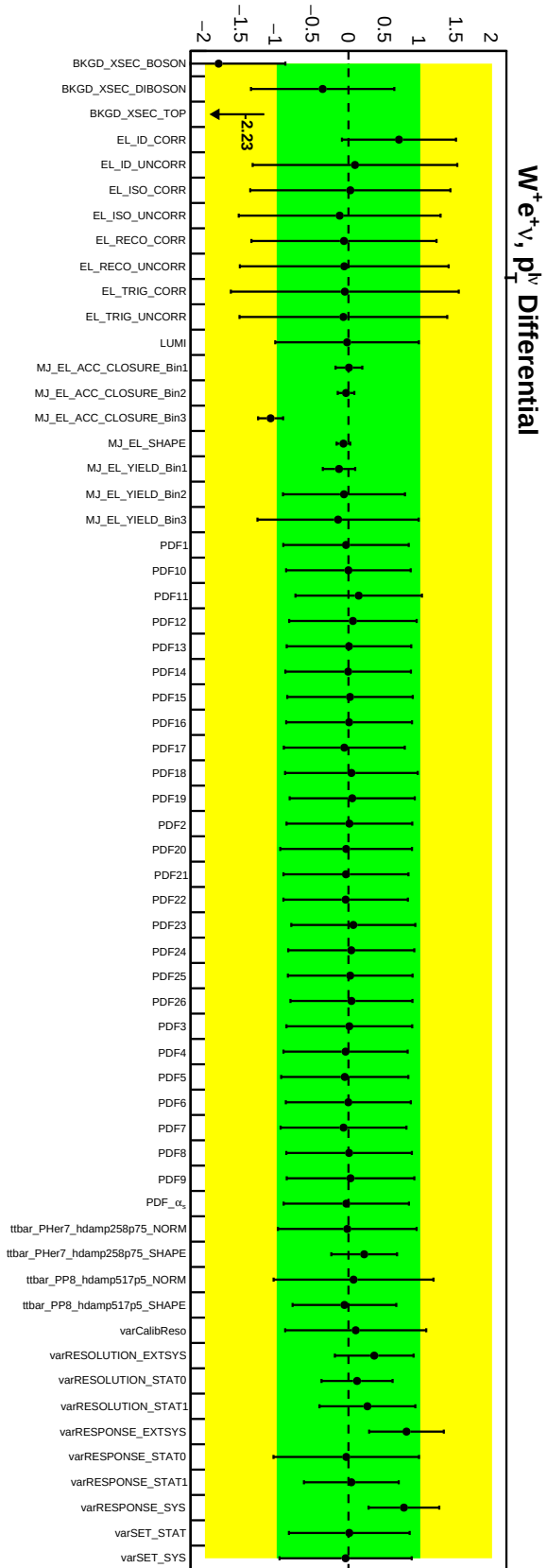


Figure G.17: The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell-Yan angular coefficients and differential cross-section as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow e^+ \nu$ channel. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty.

G.3 Additional W^- Muon Channel Figures

Chapters 8-11 usually only included figures for the $W^- \rightarrow e^- \bar{\nu}$ channel. This section will contain the same pre-fit kinematic distributions for the $W^- \rightarrow \mu^- \bar{\nu}$ channel, measured values of the Drell–Yan angular coefficients and differential cross-section compared to Asimov data, the post-fit unfolded three-dimensional $(\cos \theta_{CS}, \phi_{CS}, p_T^{\mu\nu})$ distribution unwrapped into a one-dimensional distribution, and the plot of the NP pulls.

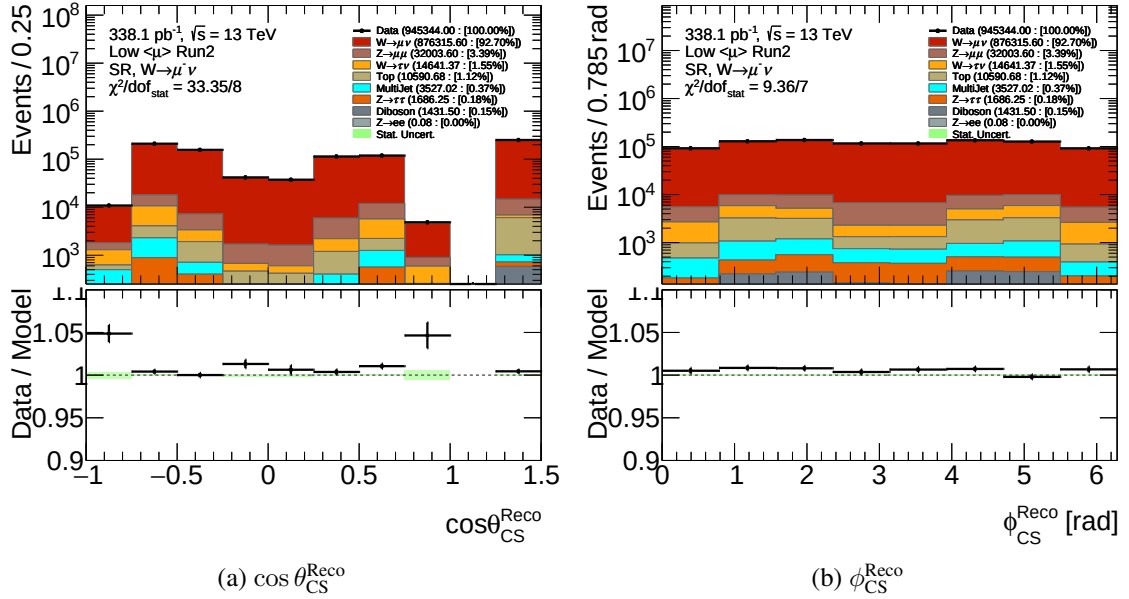


Figure G.18: The integrated CS angular distributions of a) $\cos \theta_{CS}^{\text{Reco}}$ and b) ϕ_{CS}^{Reco} for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

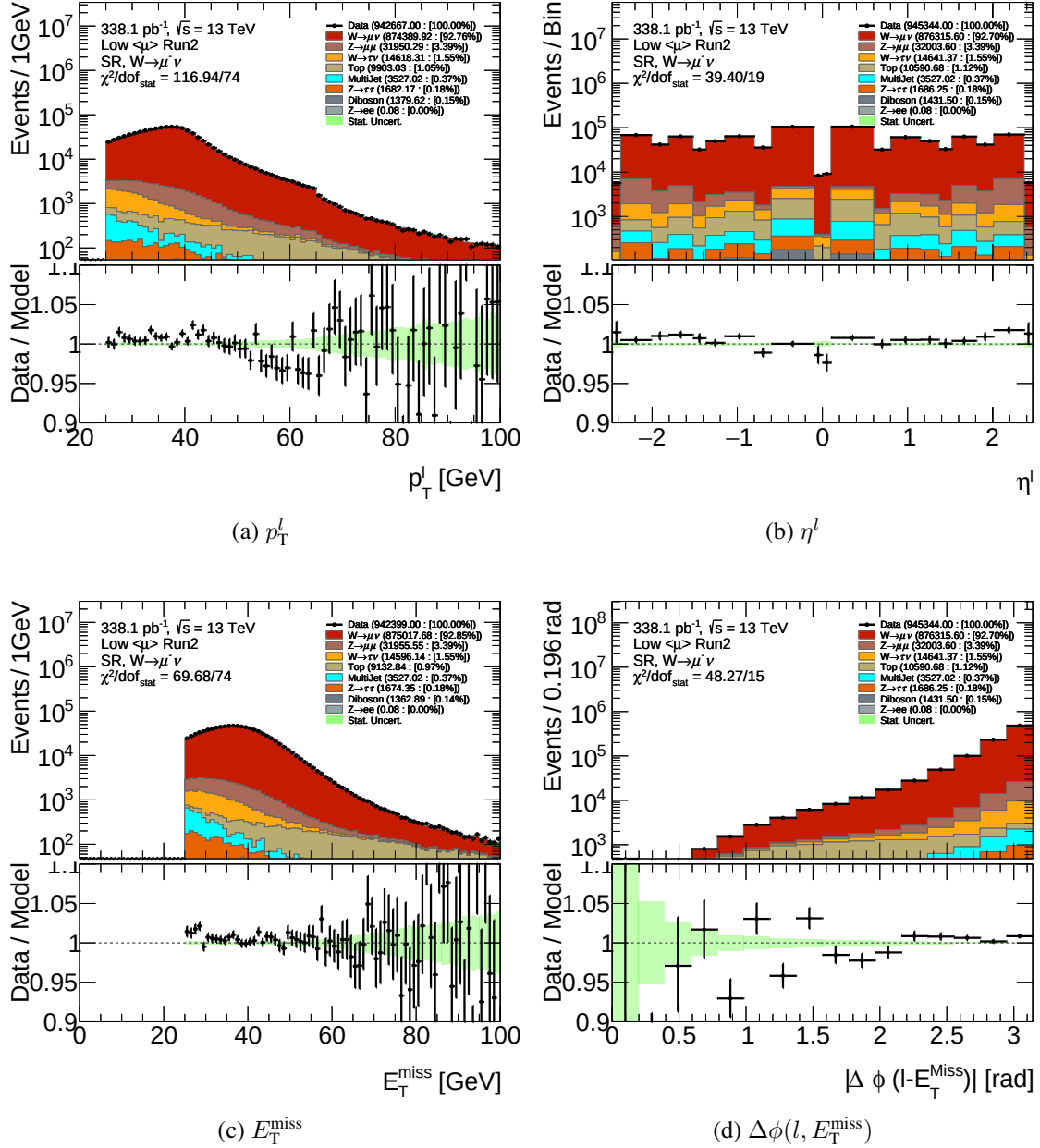


Figure G.19: The kinematic distributions for a) p_T^l , b) η^l , c) E_T^{miss} , and d) $\Delta\phi(l, E_T^{\text{miss}})$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

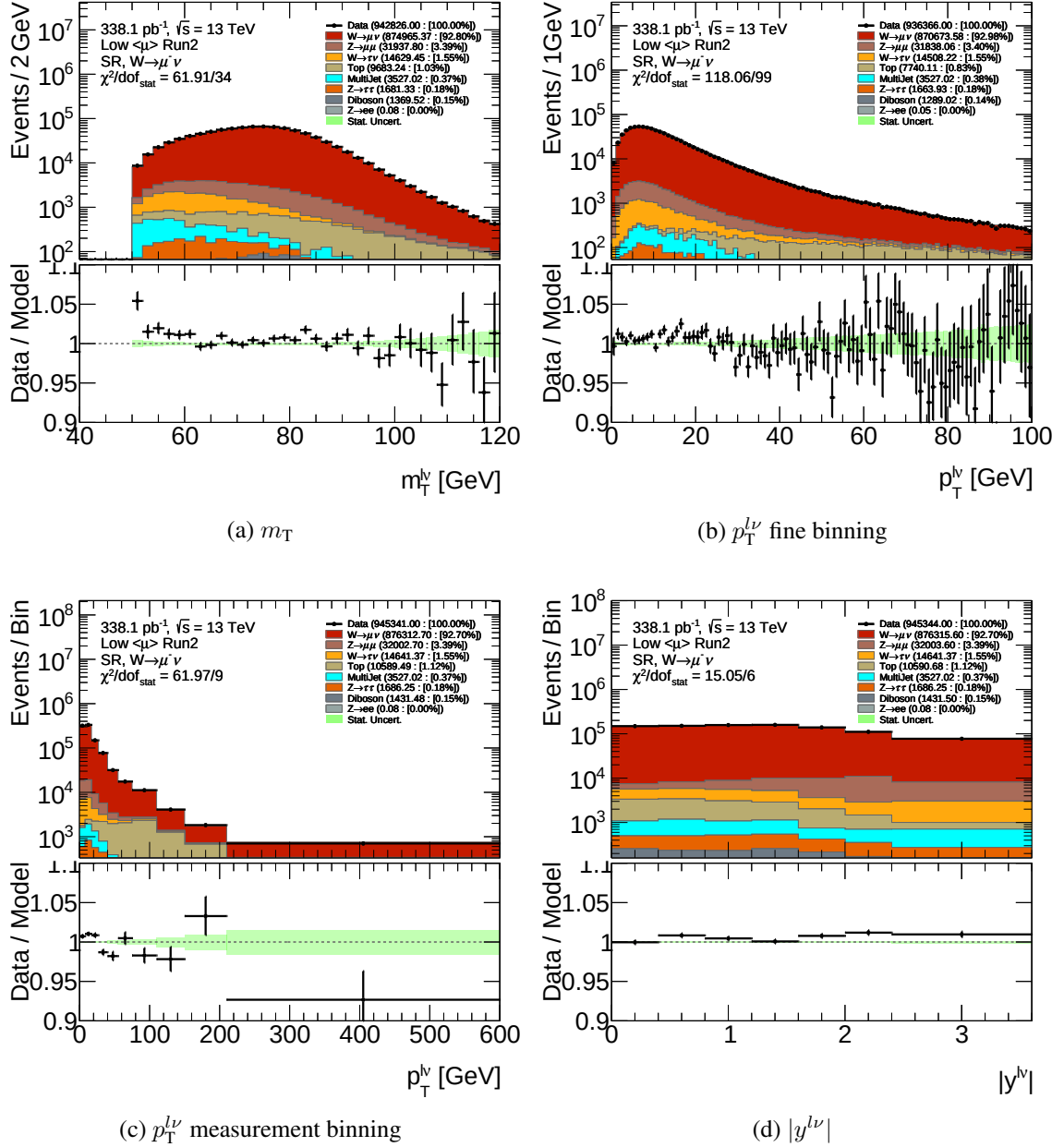
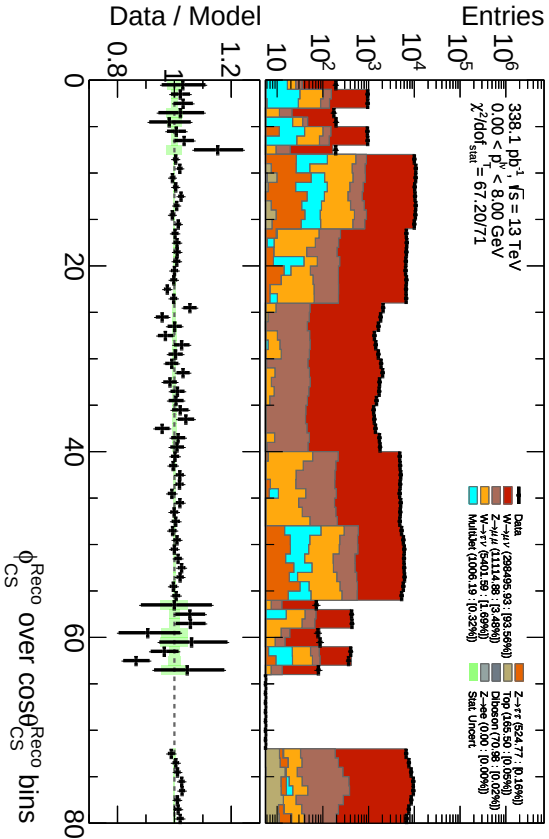
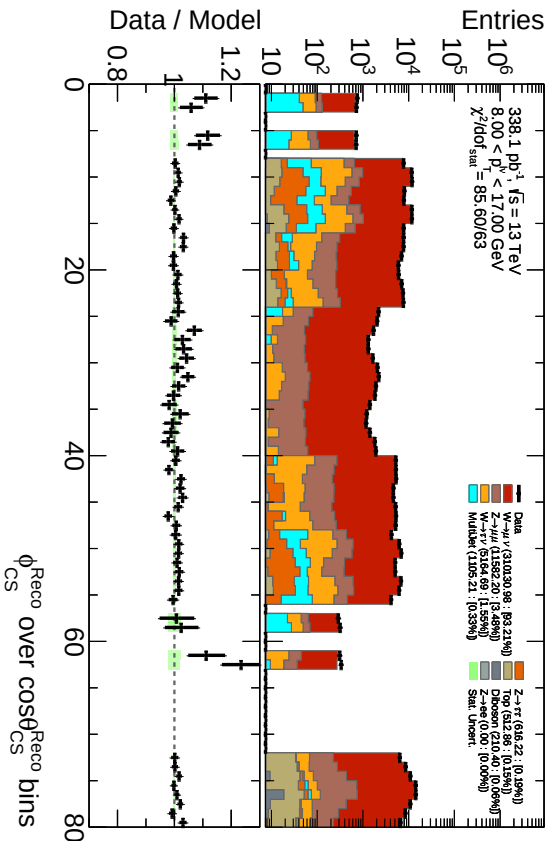


Figure G.20: The kinematic distributions for a) m_T , b) $p_T^{l\nu}$ in a fine binning, c) $p_T^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.



(a) $0 < p_{\text{T}}^{l\nu} < 8 \text{ GeV}$



(b) $8 < p_{\text{T}}^{l\nu} < 17 \text{ GeV}$

Figure G.21: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_{\text{T}}^{l\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $0 < p_{\text{T}}^{l\nu} < 8 \text{ GeV}$ and b) $8 < p_{\text{T}}^{l\nu} < 17 \text{ GeV}$. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

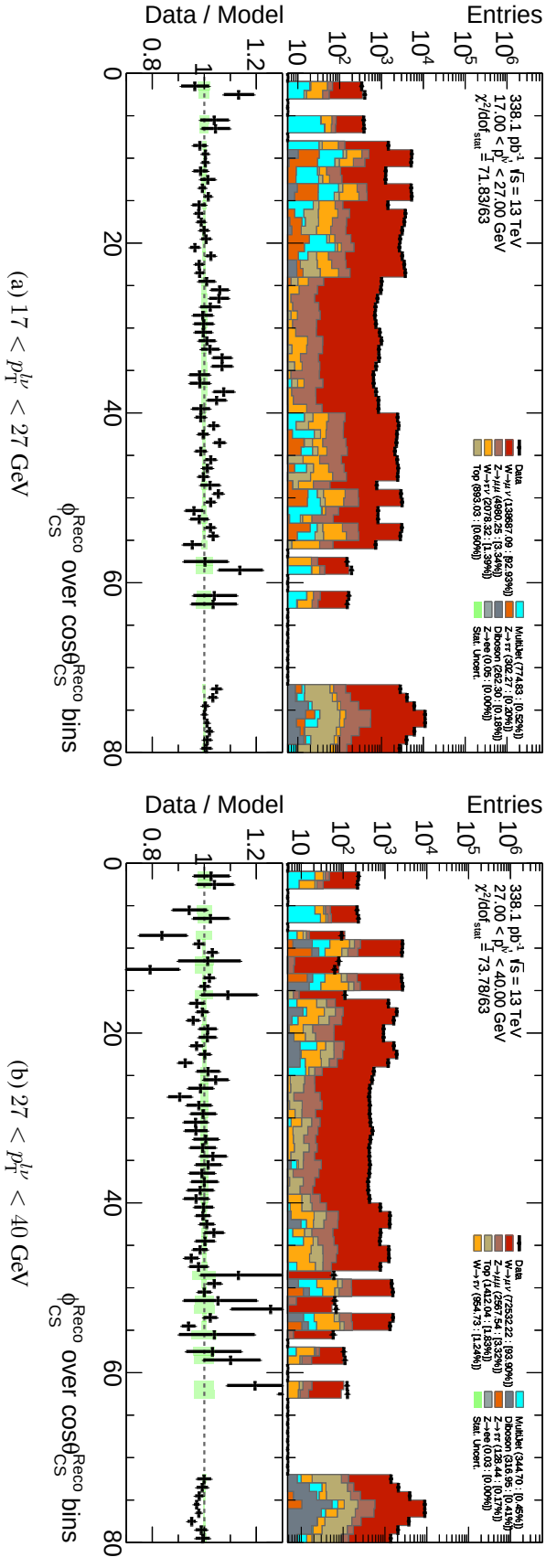
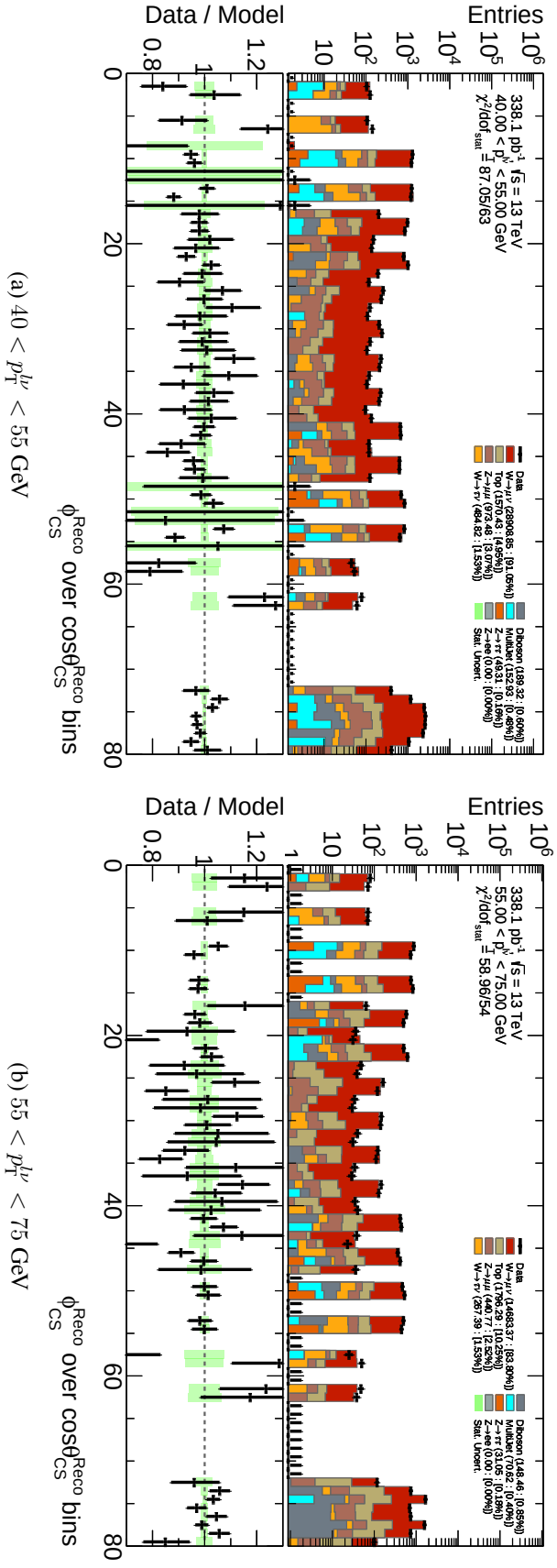


Figure G.22: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{\ell\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $17 < p_T^{\ell\nu} < 27$ GeV and b) $27 < p_T^{\ell\nu} < 40$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.



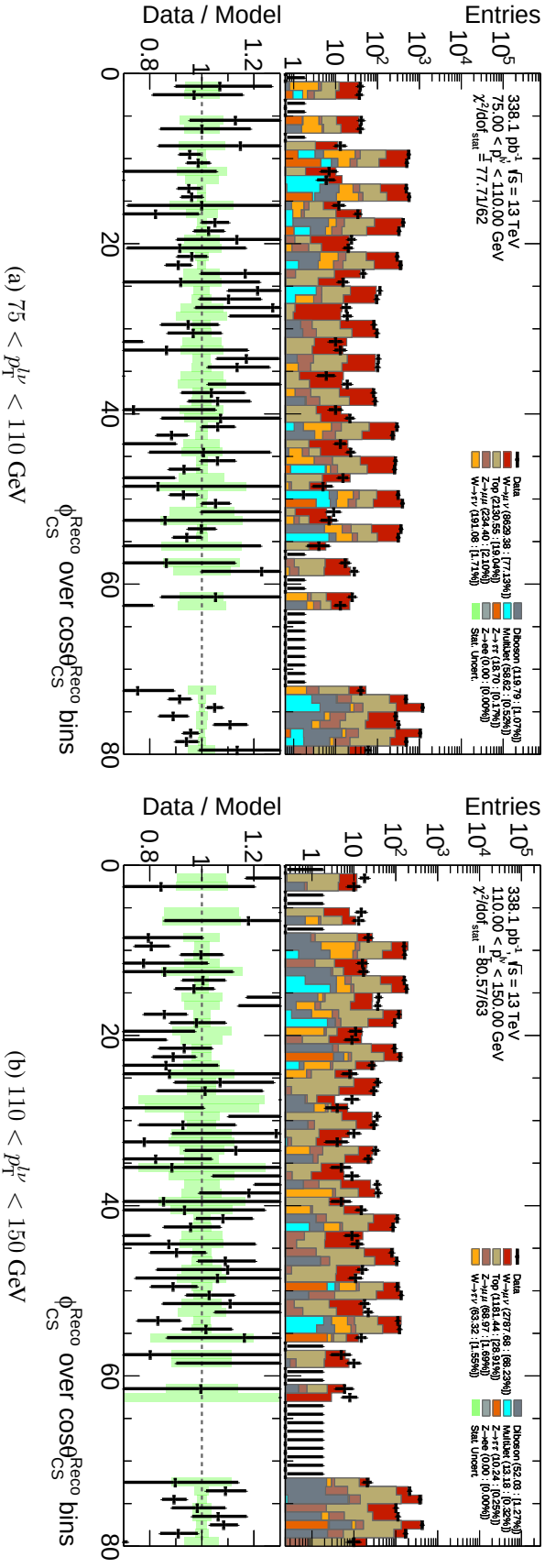
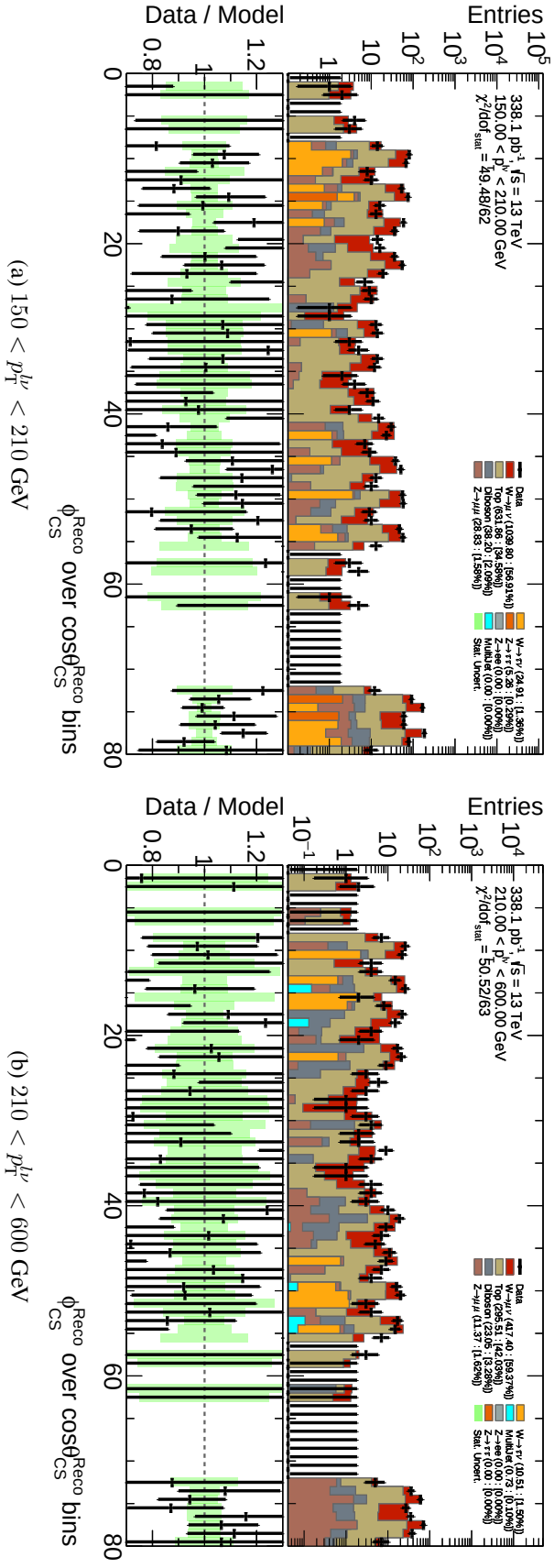
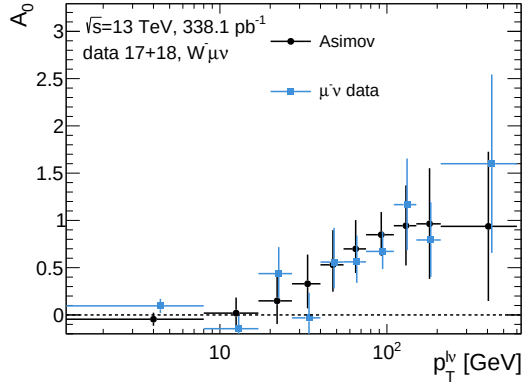
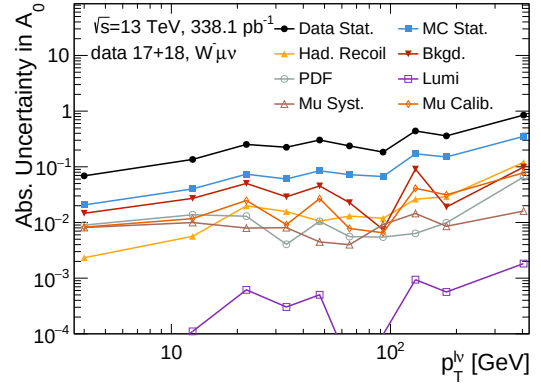


Figure G.24: The two-dimensional ($\cos\theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{l\nu}$ for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_T^{l\nu} < 110$ GeV and b) $110 < p_T^{l\nu} < 150$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

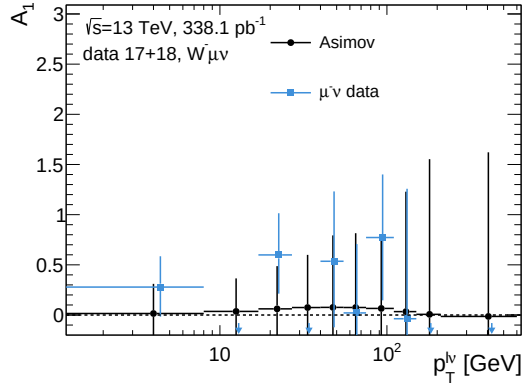




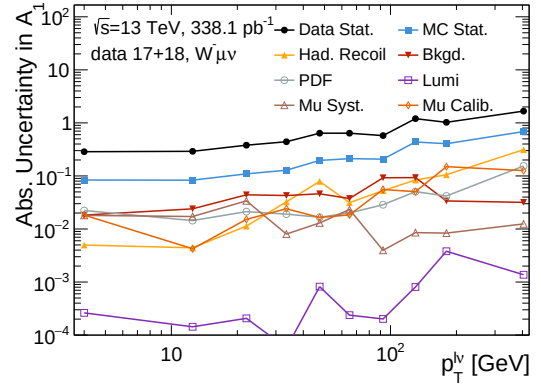
(a) A_0



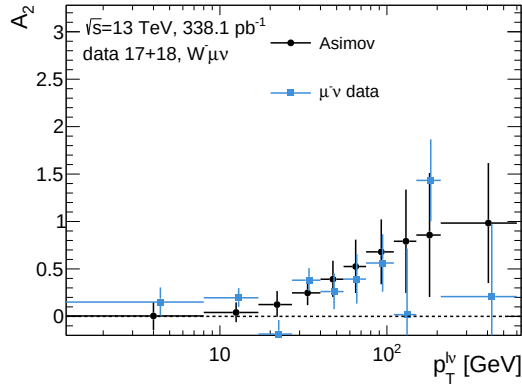
(b) Uncertainties for A_0



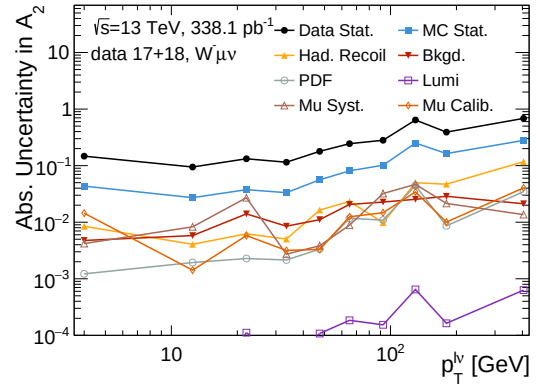
(c) A_1



(d) Uncertainties for A_1

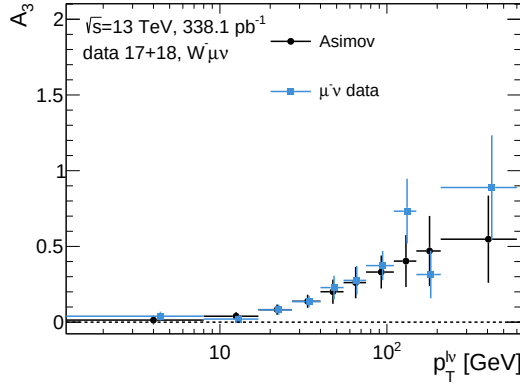


(e) A_2

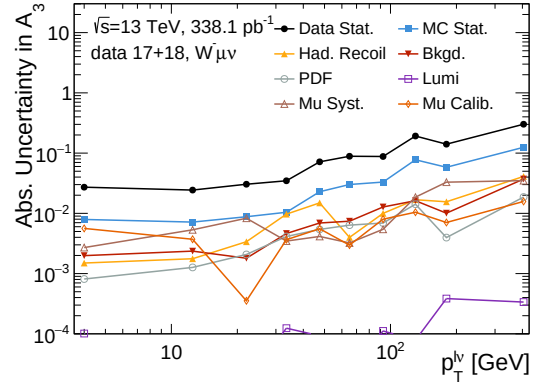


(f) Uncertainties for A_2

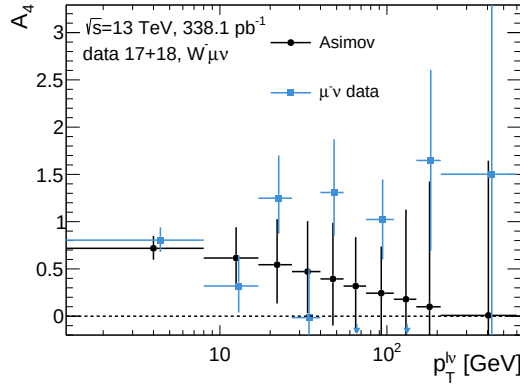
Figure G.26: The measured coefficients – A_0 , A_1 , A_2 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



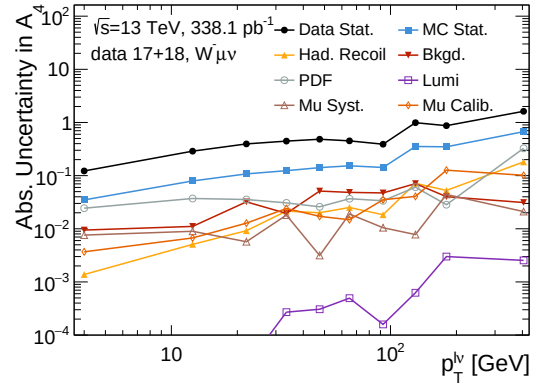
(a) A_3



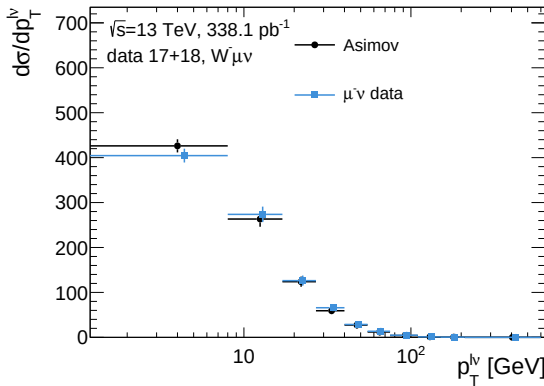
(b) Uncertainties for A_3



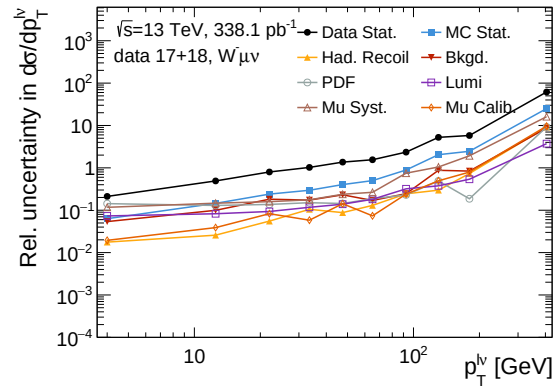
(c) A_4



(d) Uncertainties for A_4

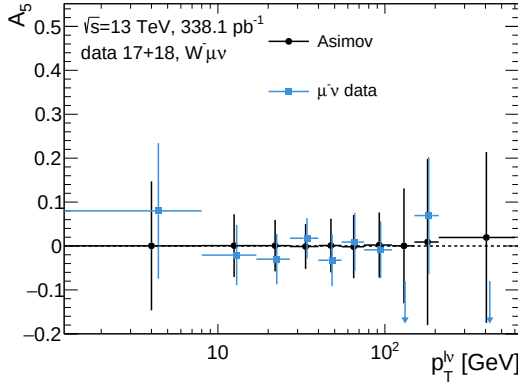


(e) Differential cross-section

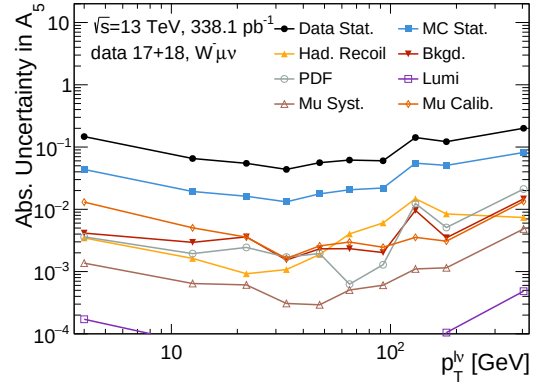


(f) Uncertainties for the differential cross-section

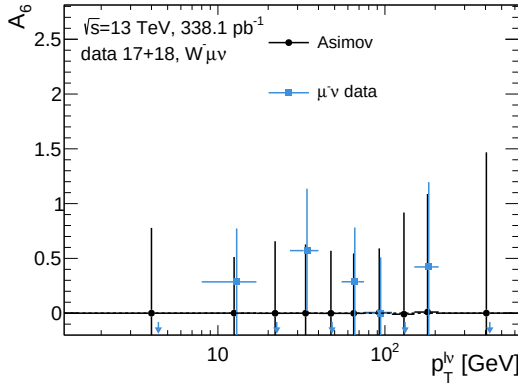
Figure G.27: The measured coefficients – A_3 and A_4 –, the differential cross-section, and their uncertainties as a function of $p_T^{l\nu}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



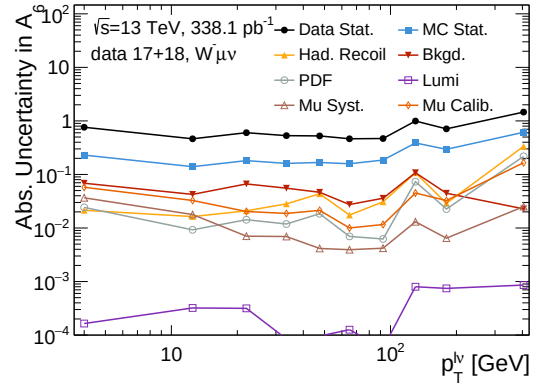
(a) A_5



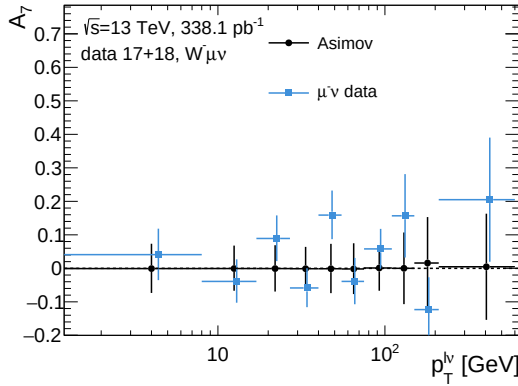
(b) Uncertainties for A_5



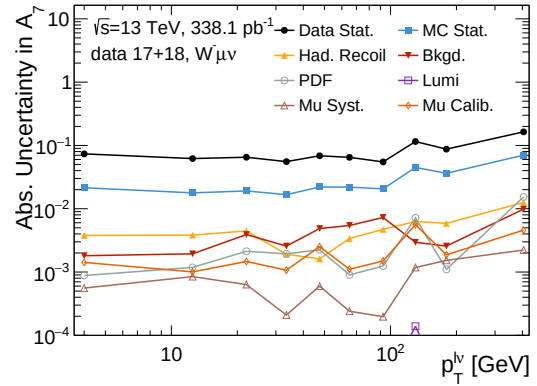
(c) A_6



(d) Uncertainties for A_6



(e) A_7



(f) Uncertainties for A_7

Figure G.28: The measured coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.

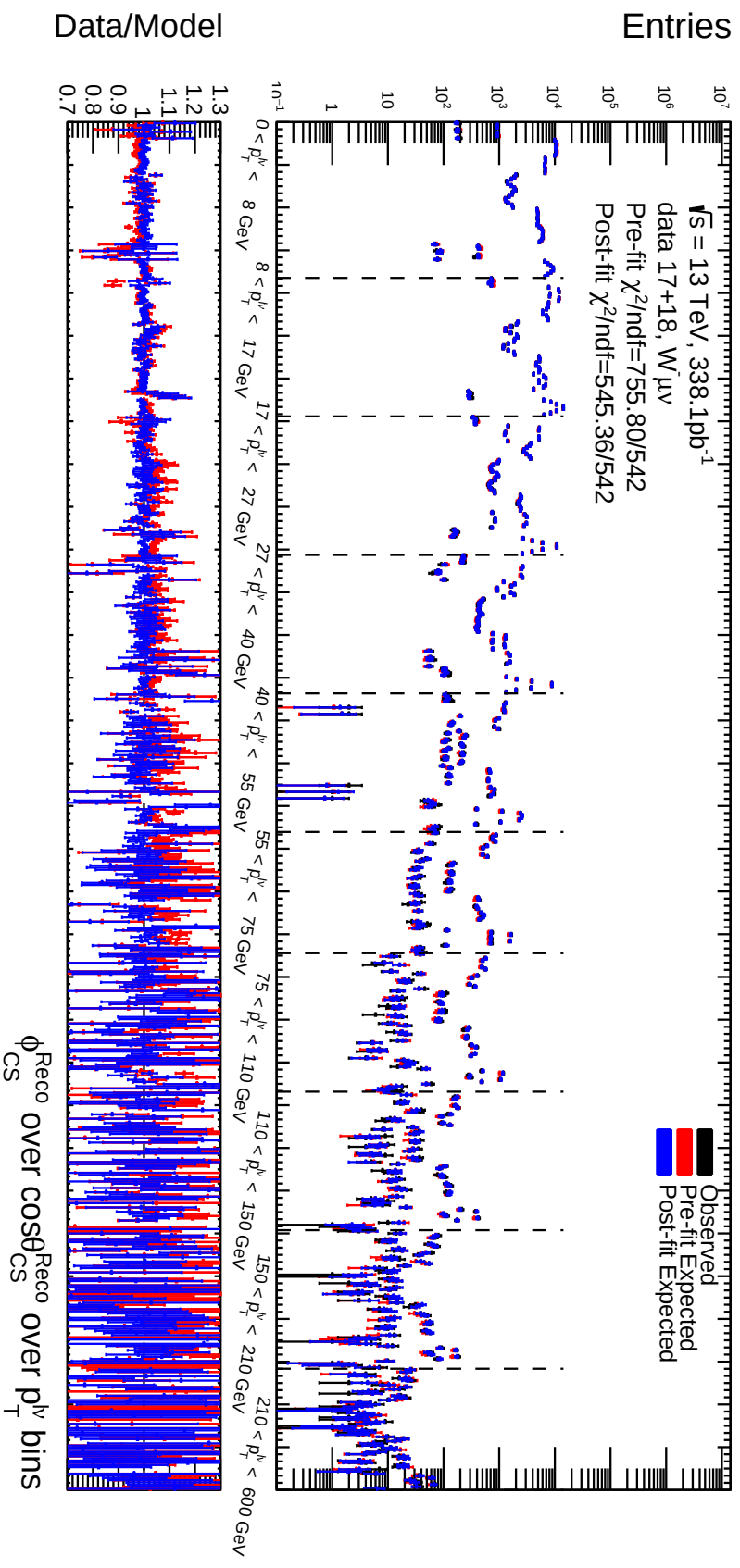


Figure G.29: The unfolded three-dimensional ($\cos\theta_{CS}$, ϕ_{CS} , p_T^{ν}) distribution unwrapped into a one-dimensional distribution for $W^- \rightarrow \mu^- \bar{\nu}$ events that pass selection criteria listed in Table 8.3 shown on a logarithmic scale for each p_T^{ν} bin separated by vertical dashed lines and labelled underneath the graph. The measured data, N_{obs} , is denoted by black points, the data-driven MJ background estimate, MC signal, and MC background are combined into N_{exp} and normalized to the data luminosity are denoted by the red points pre-fit and post-fit by the blue points. The ratio of N_{obs} over N_{exp} for both pre- and post-fit is displayed on the bottom plot showing the agreement improves through the NLL minimization.

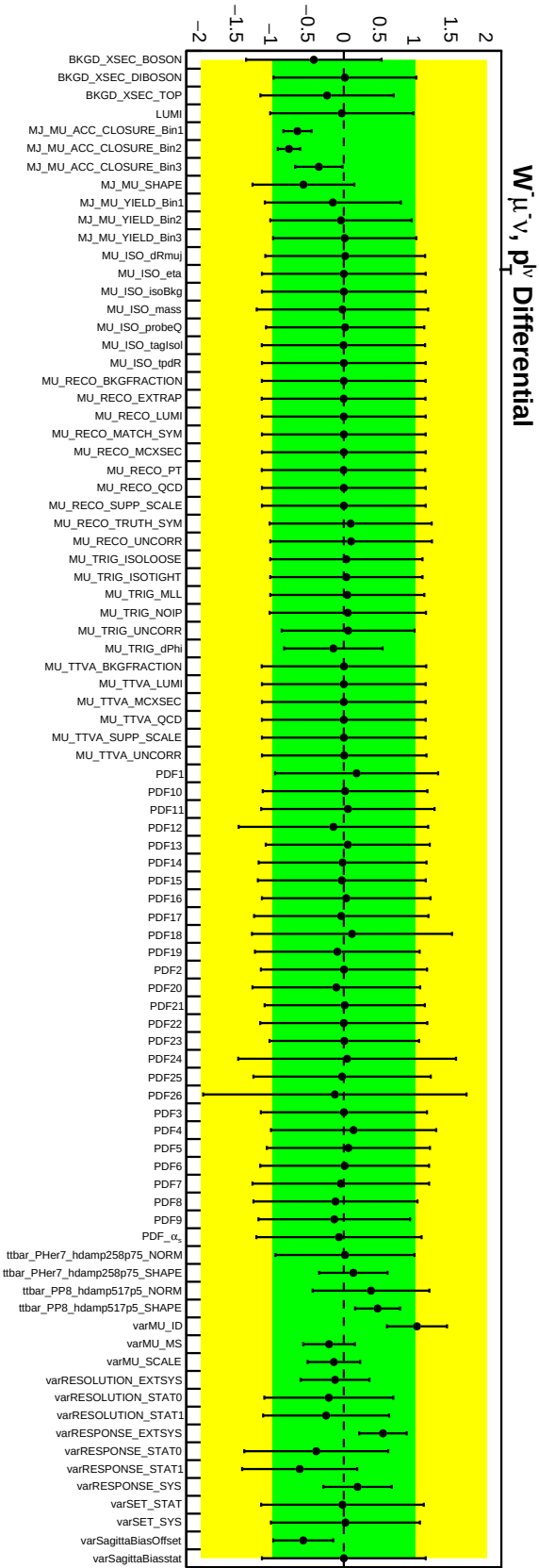


Figure G.30: The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell-Yan angular coefficients and differential cross-section as a function of $p_T^{l\bar{l}}$ for the $W^- \rightarrow \mu^- \bar{\nu}$ channel. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty.

G.4 Additional W^+ Muon Channel Figures

Chapters 8-11 usually only included figures for the $W^- \rightarrow e^- \bar{\nu}$ channel. This section will contain the same pre-fit kinematic distributions for the $W^+ \rightarrow \mu^+ \nu$ channel, measured values of the Drell–Yan angular coefficients and differential cross-section compared to Asimov data, the post-fit unfolded three-dimensional $(\cos \theta_{\text{CS}}, \phi_{\text{CS}}, p_{\text{T}}^{\mu\nu})$ distribution unwrapped into a one-dimensional distribution, and the plot of the NP pulls.

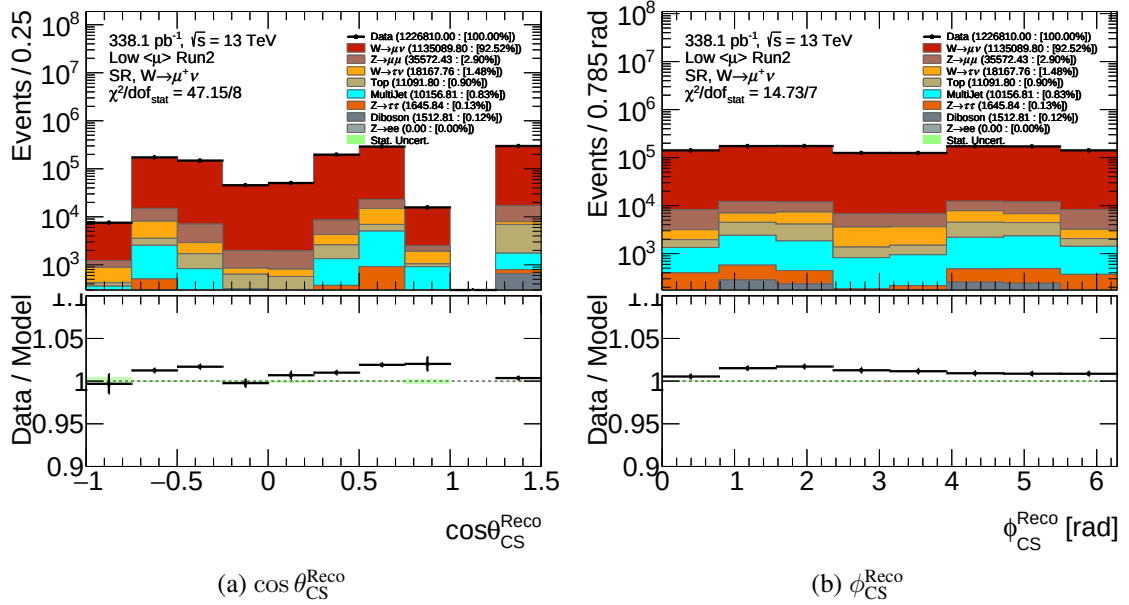


Figure G.31: The integrated CS angular distributions of a) $\cos \theta_{\text{CS}}^{\text{Reco}}$ and b) $\phi_{\text{CS}}^{\text{Reco}}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

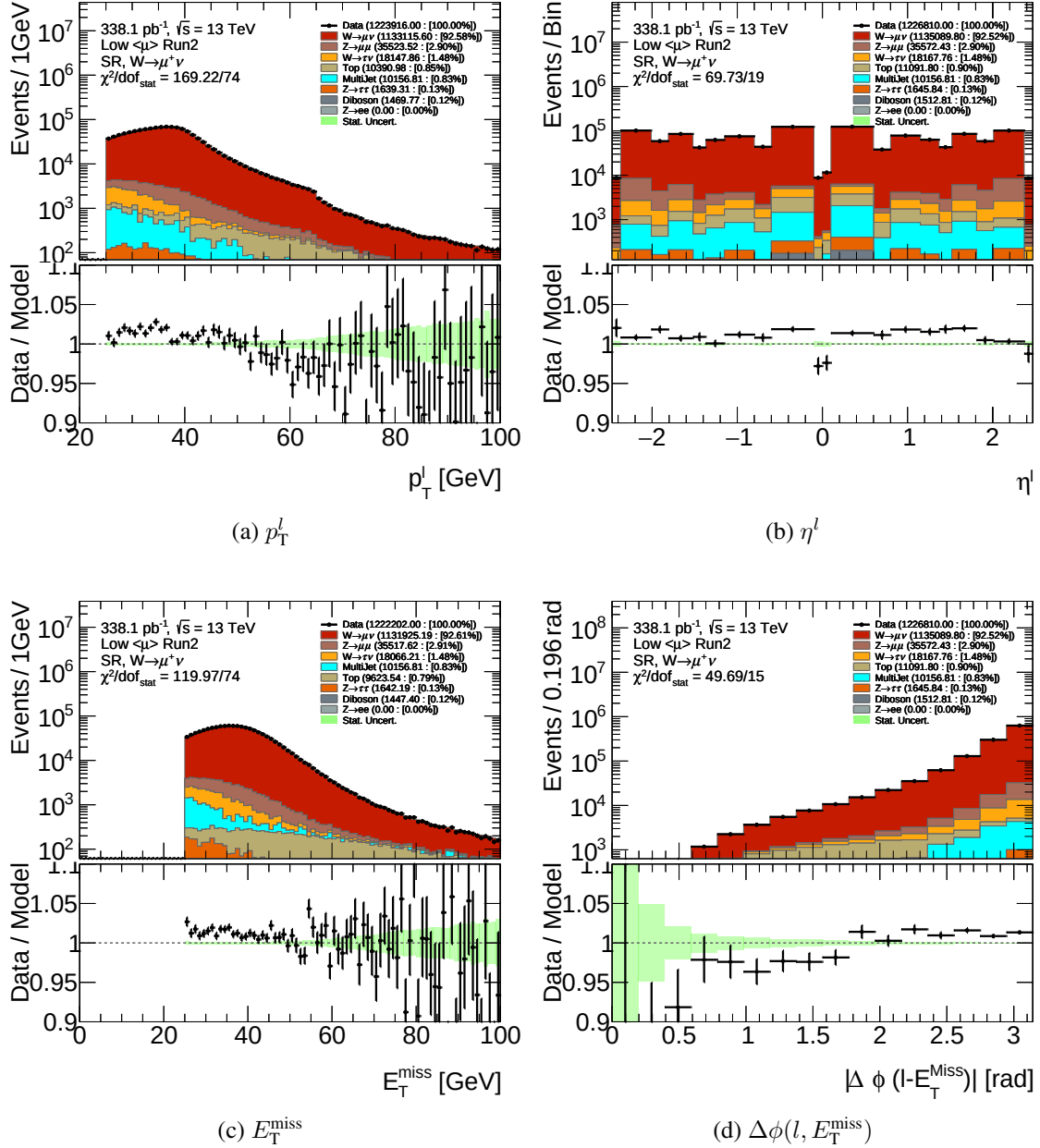


Figure G.32: The kinematic distributions for a) p_T^l , b) η^l , c) E_T^{miss} , and d) $\Delta\phi(l, E_T^{\text{miss}})$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

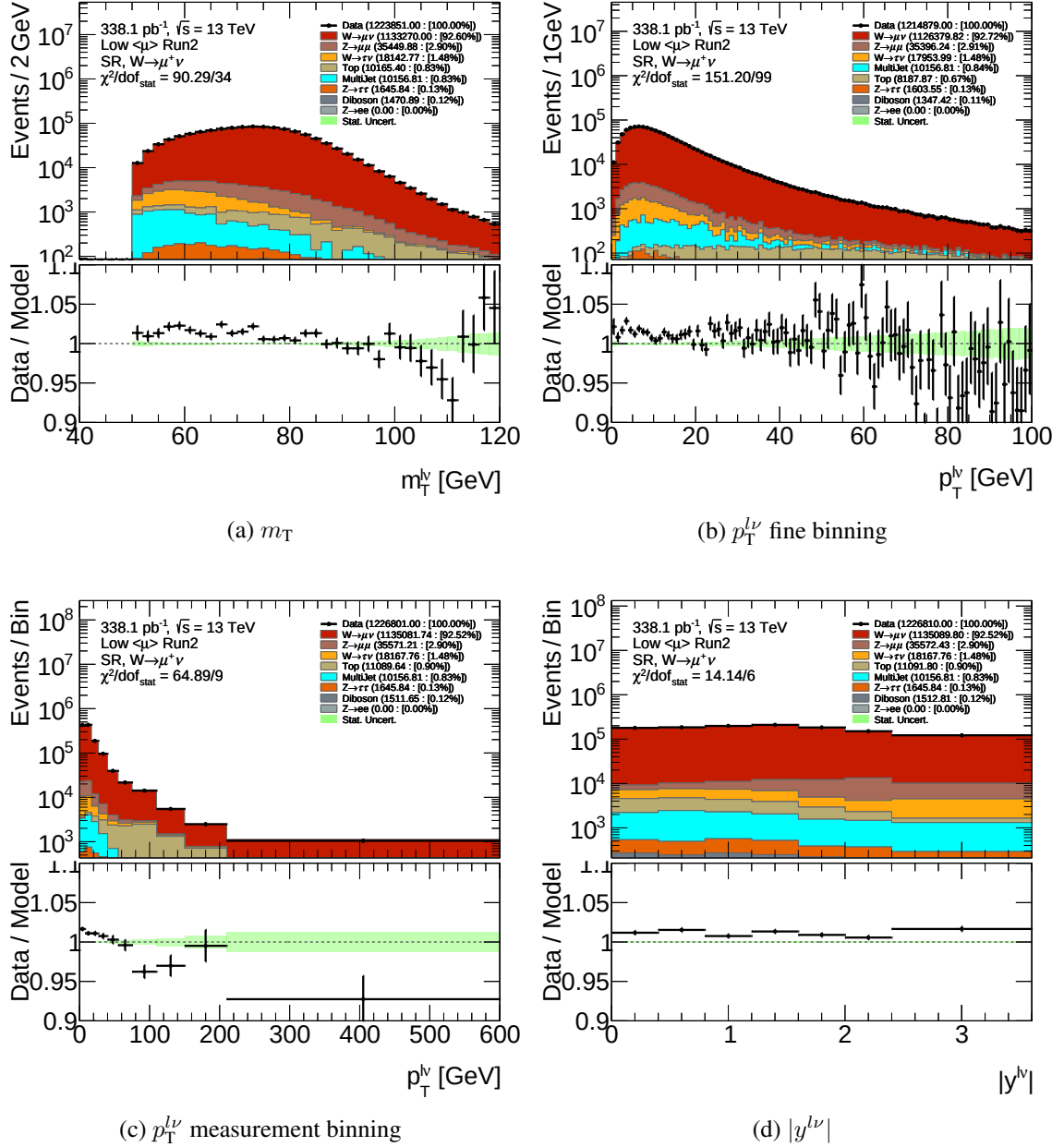
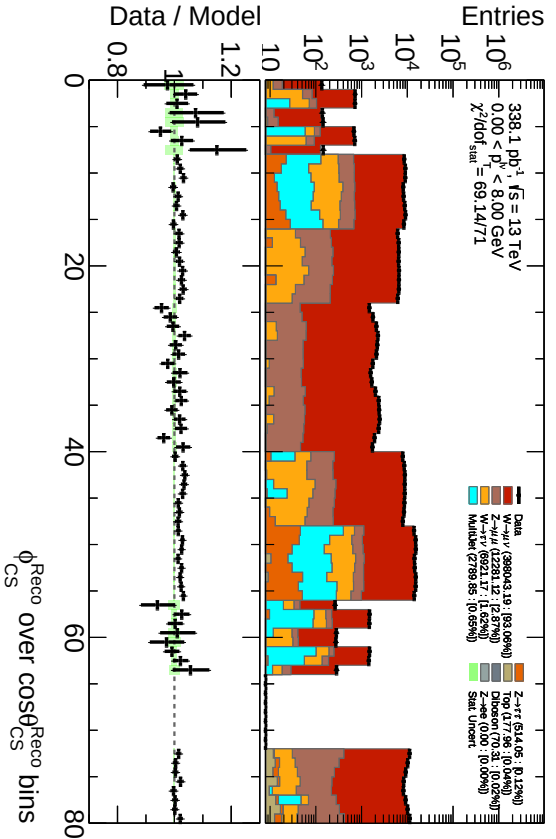
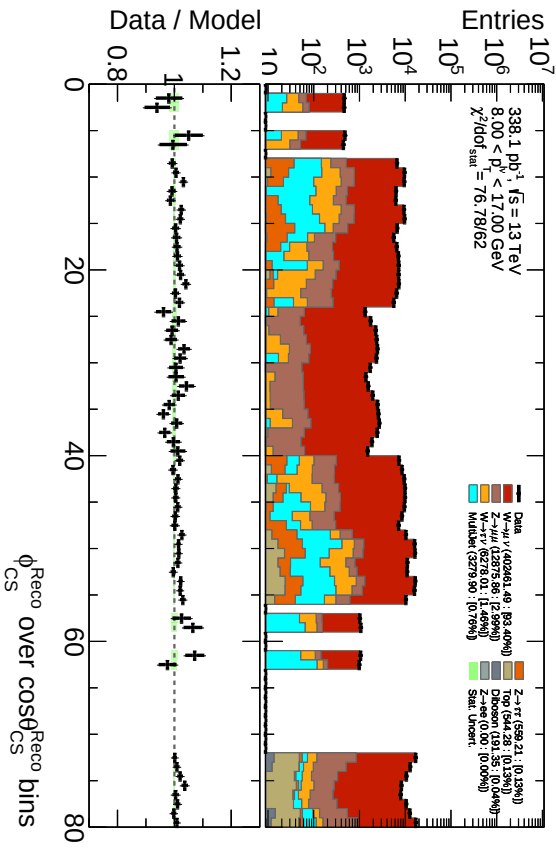


Figure G.33: The kinematic distributions for a) m_T , b) $p_T^{l\nu}$ in a fine binning, c) $p_T^{l\nu}$ in the binning used for the measurement, and d) $|y^{l\nu}|$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.



(a) $0 < p_T^{\ell\nu} < 8 \text{ GeV}$



(b) $8 < p_T^{\ell\nu} < 17 \text{ GeV}$

Figure G.34: The two-dimensional ($\cos\theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{\ell\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $0 < p_T^{\ell\nu} < 8 \text{ GeV}$ and b) $8 < p_T^{\ell\nu} < 17 \text{ GeV}$. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

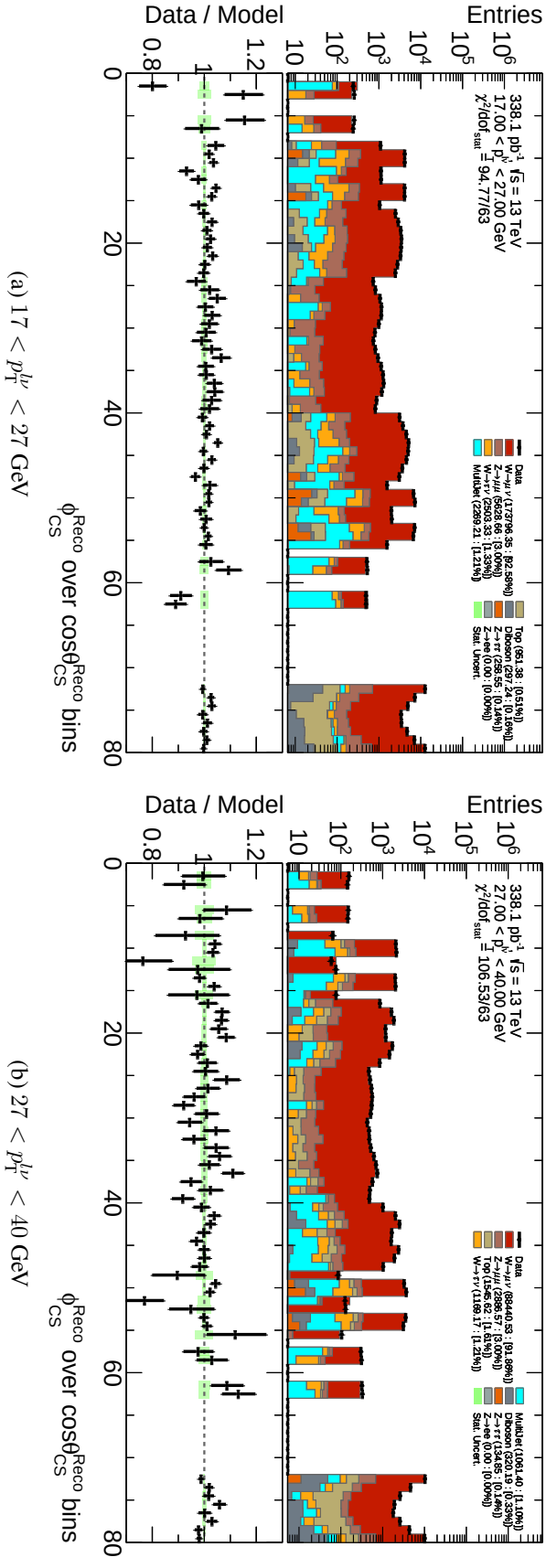


Figure G.35: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{\ell\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $17 < p_T^{\ell\nu} < 27 \text{ GeV}$ and b) $27 < p_T^{\ell\nu} < 40 \text{ GeV}$. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

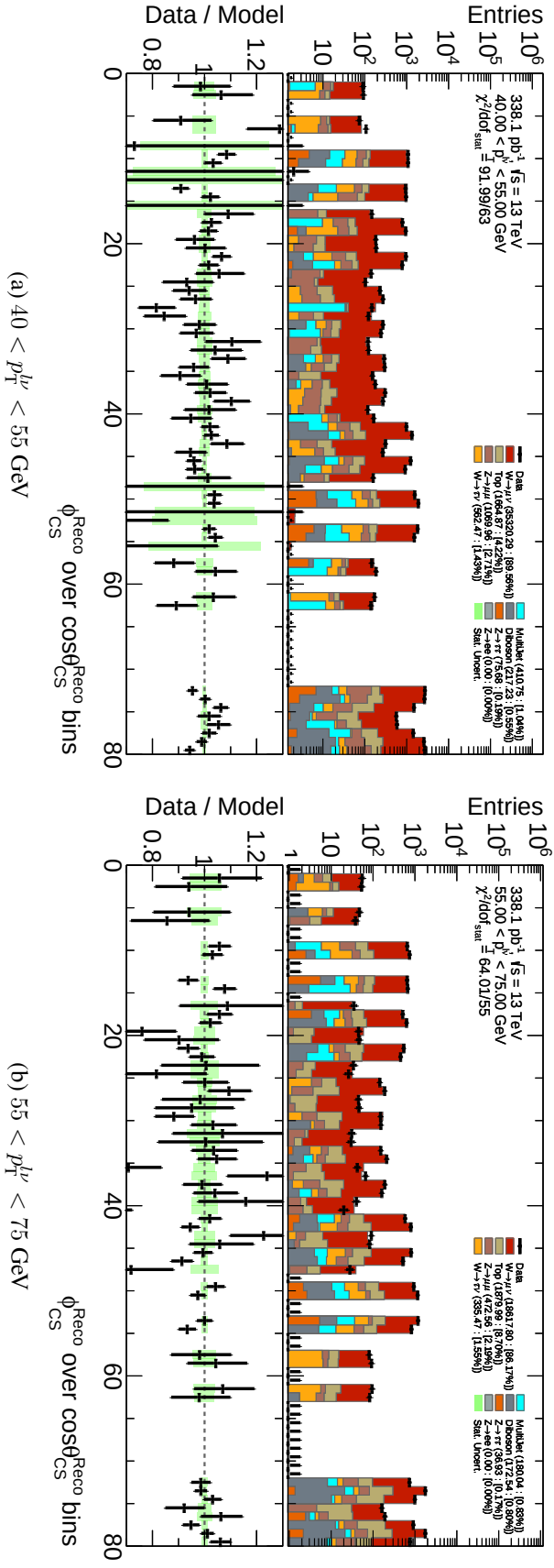


Figure G.36: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{\ell\nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $40 < p_T^{\ell\nu} < 55 \text{ GeV}$ and b) $55 < p_T^{\ell\nu} < 75 \text{ GeV}$. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

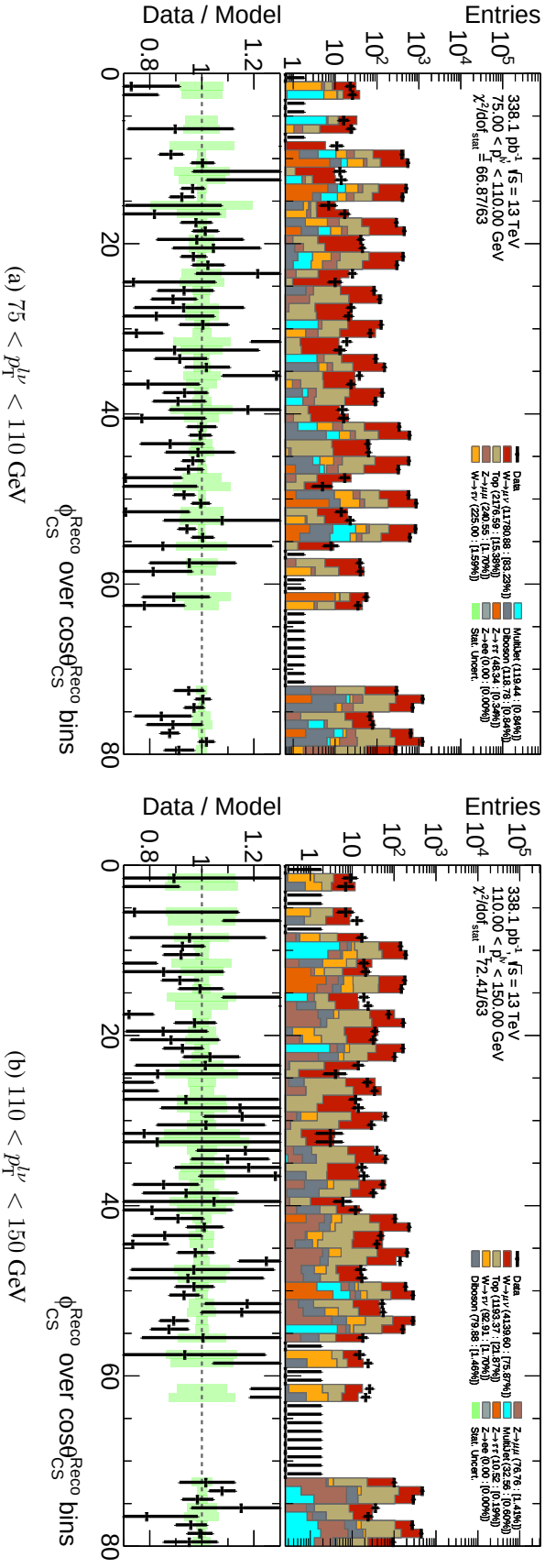


Figure G.37: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{\mu \nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $75 < p_T^{\mu \nu} < 110$ GeV and b) $110 < p_T^{\mu \nu} < 150$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.

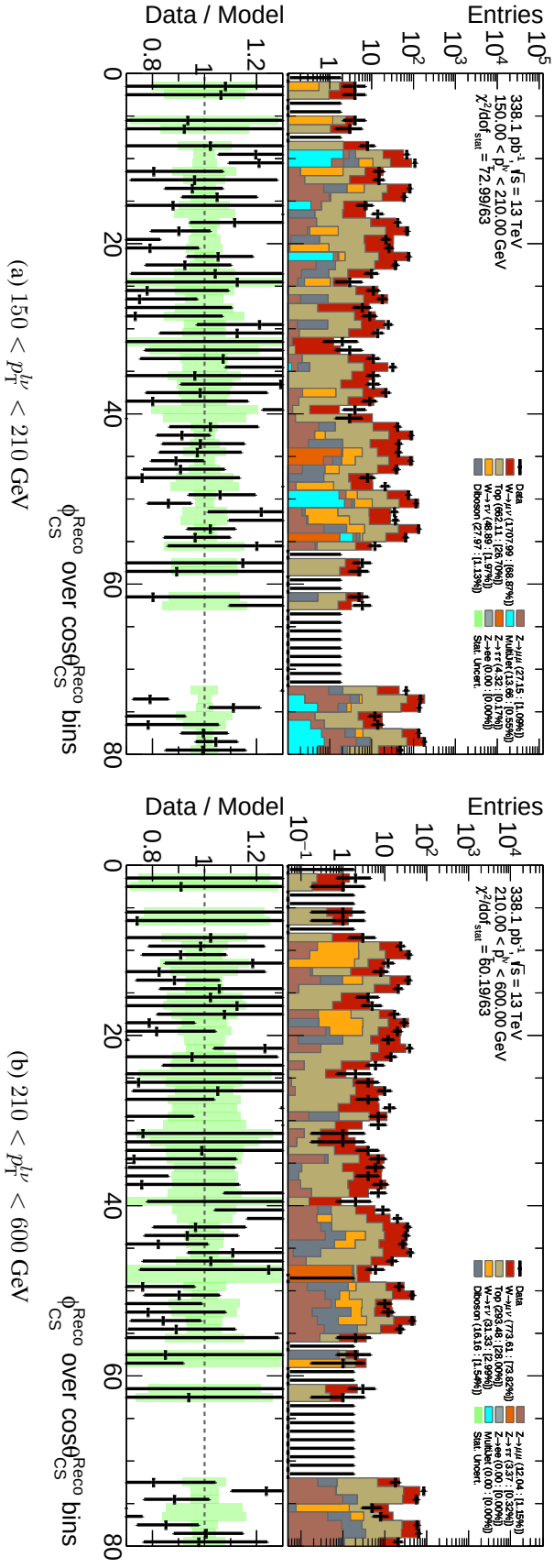
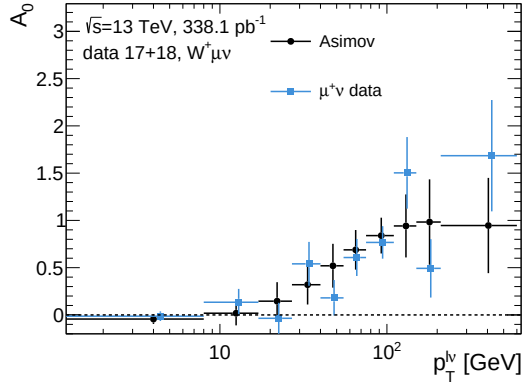
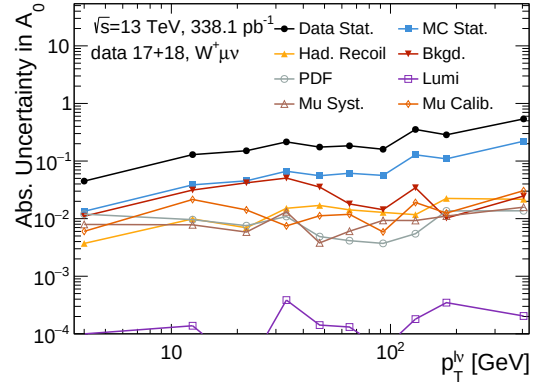


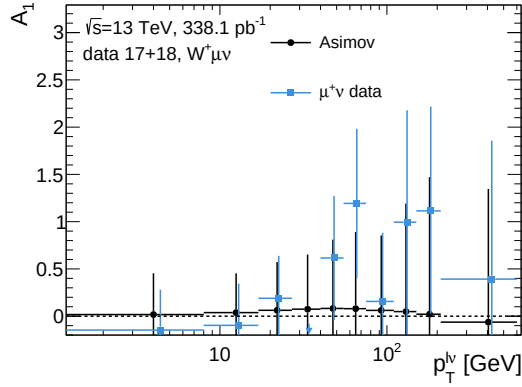
Figure G.38: The two-dimensional ($\cos \theta_{\text{CS}}^{\text{Reco}}$, $\phi_{\text{CS}}^{\text{Reco}}$) distribution unwrapped into a one-dimensional distribution as a function of $p_T^{\mu \nu}$ for $W^+ \rightarrow \mu^+ \nu$ events that pass selection criteria in Table 8.3 shown on a logarithmic scale for two bins: a) $150 < p_T^{\mu \nu} < 210$ GeV and b) $210 < p_T^{\mu \nu} < 600$ GeV. The black points represent the measured data and its statistical uncertainty. The MC signal and background pre-fit distributions as well as data-driven MJ estimate are normalized to the data luminosity and are denoted by the filled histograms. The bottom plots show the ratio of the data over the model to show the pre-fit agreement, where the uncertainty on the points is the statistical uncertainty of the data and the error band is the systematic uncertainty of the data-driven MJ.



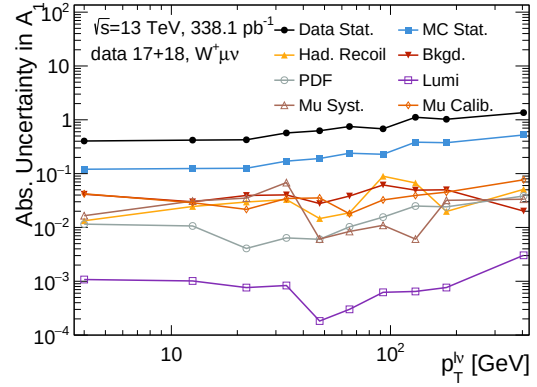
(a) A_0



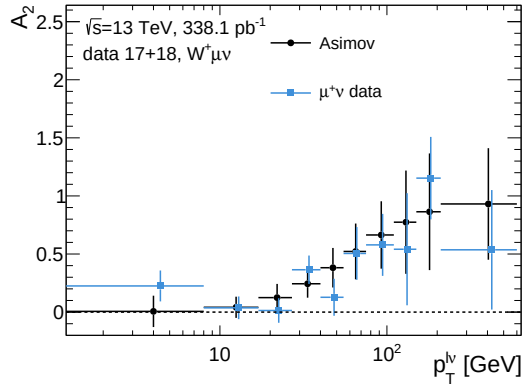
(b) Uncertainties for A_0



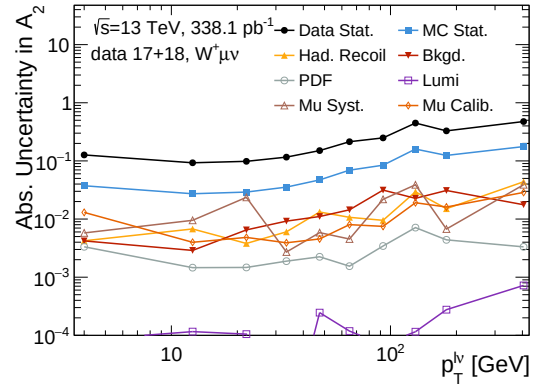
(c) A_1



(d) Uncertainties for A_1

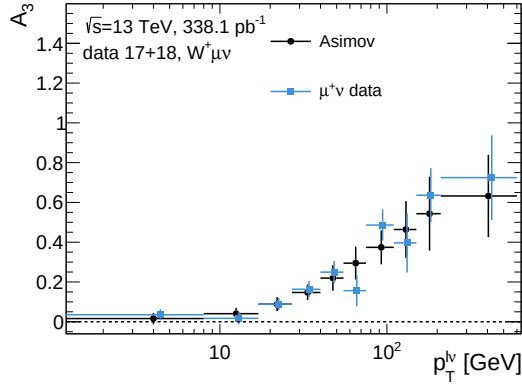


(e) A_2

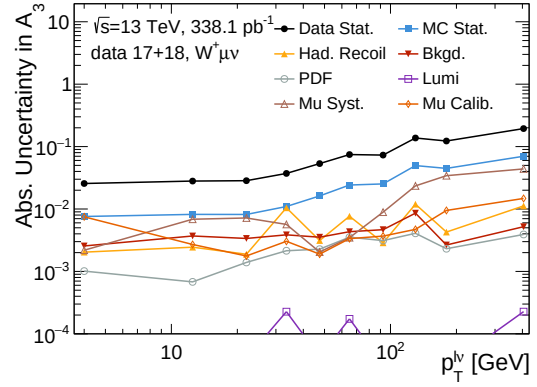


(f) Uncertainties for A_2

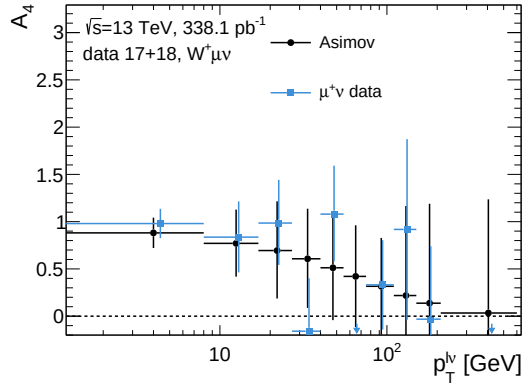
Figure G.39: The measured coefficients – A_0 , A_1 , A_2 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow \mu^+ \nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



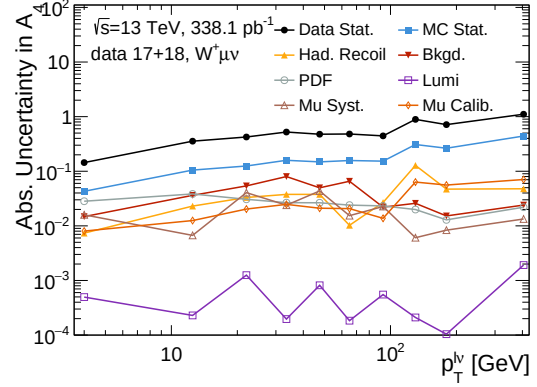
(a) A_3



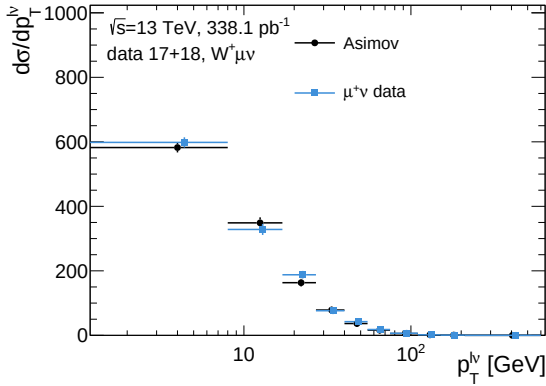
(b) Uncertainties for A_3



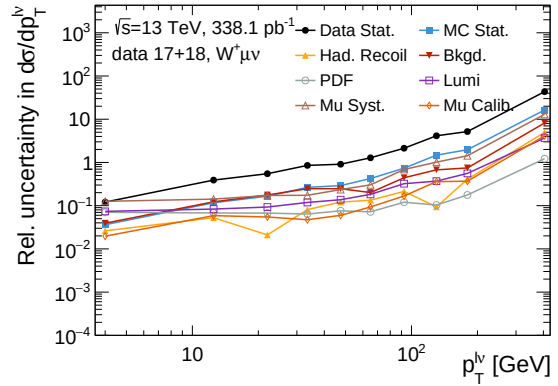
(c) A_4



(d) Uncertainties for A_4

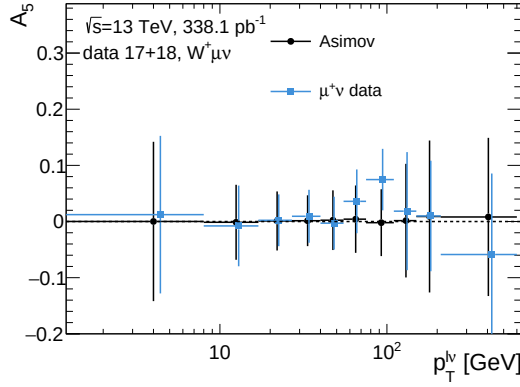


(e) Differential cross-section

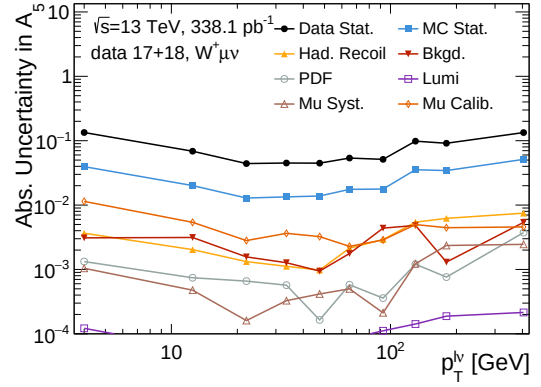


(f) Uncertainties for the differential cross-section

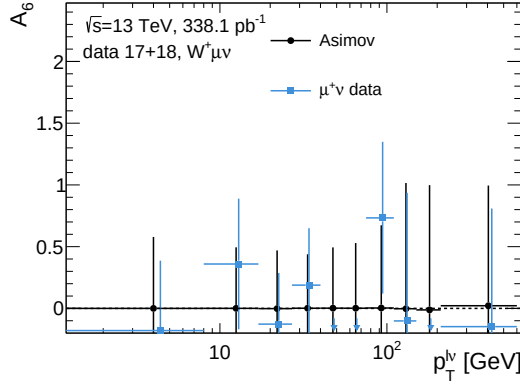
Figure G.40: The measured coefficients – A_3 and A_4 –, the differential cross-section, and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow \mu^+ \nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



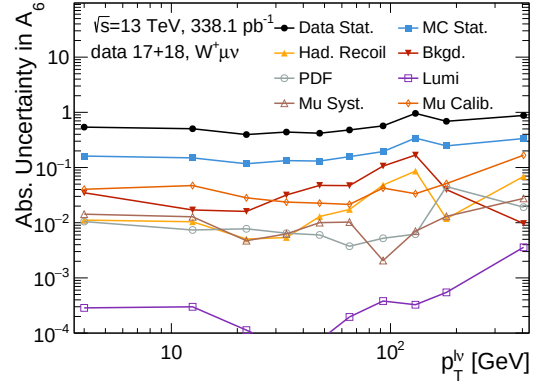
(a) A_5



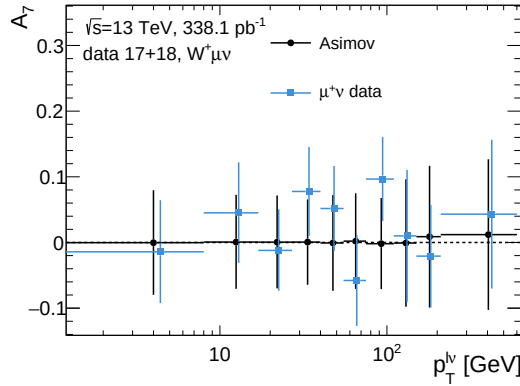
(b) Uncertainties for A_5



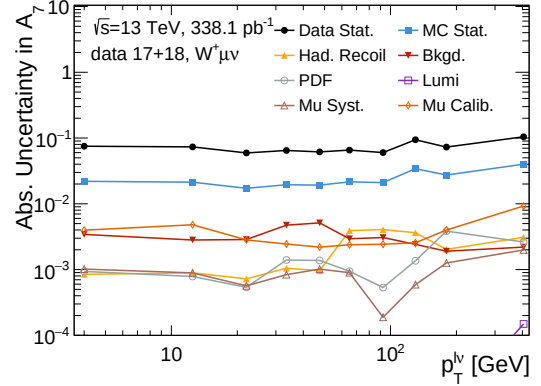
(c) A_6



(d) Uncertainties for A_6

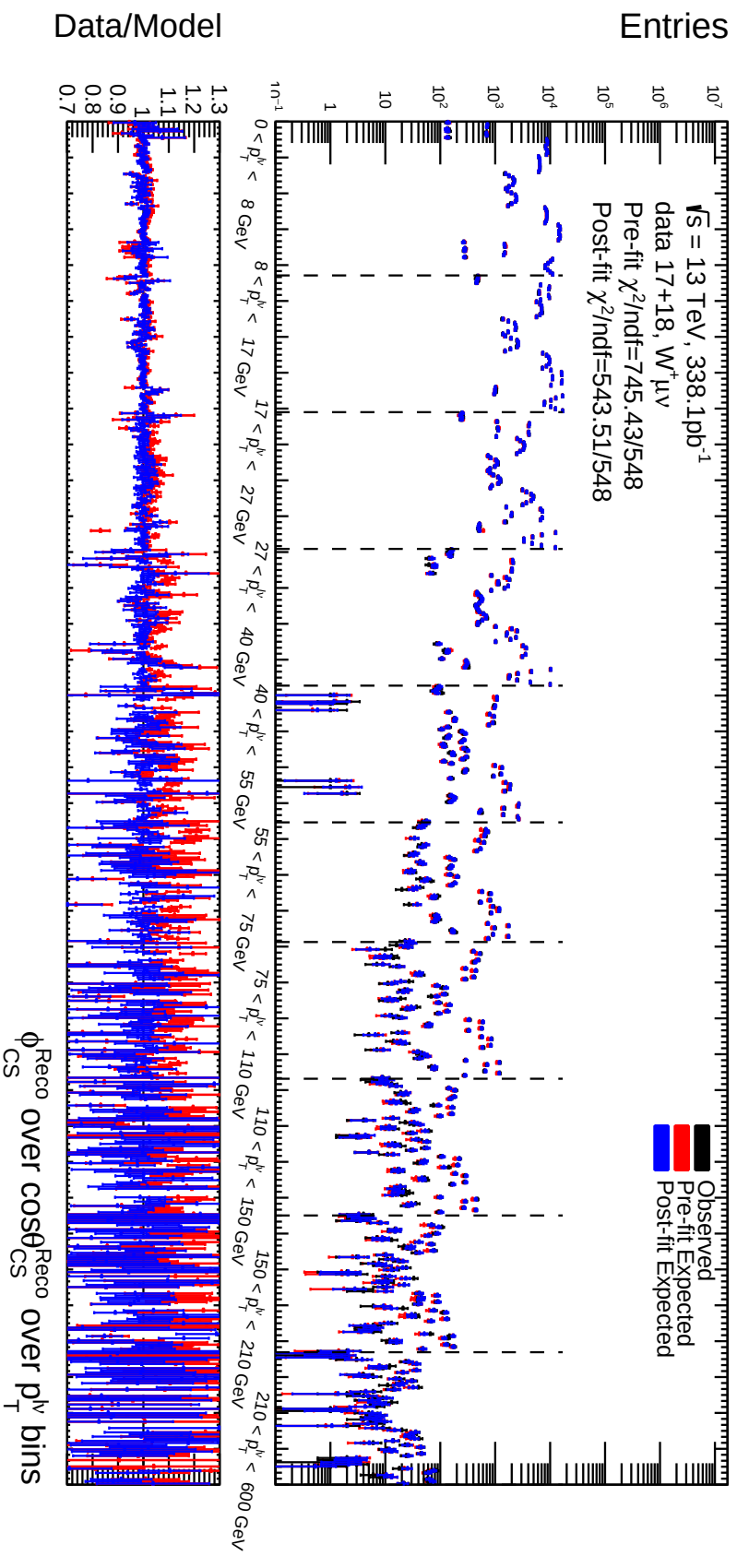


(e) A_7



(f) Uncertainties for A_7

Figure G.41: The measured coefficients – A_5 , A_6 , A_7 – and their uncertainties as a function of $p_T^{l\nu}$ for the $W^+ \rightarrow \mu^+ \nu$ channel using data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.



G.5 Additional W^- Combined Channels Figures

The only figures of the combined $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels not included in Chapters 9 and 11 were for coefficients A_1 , A_5 , A_6 , and A_7 . This section will present the measured values and uncertainties for these coefficients, which are measured but not statistically different from zero. This section will present the plot of the NPs pulls and constraints that occurred during the combined channel fit.

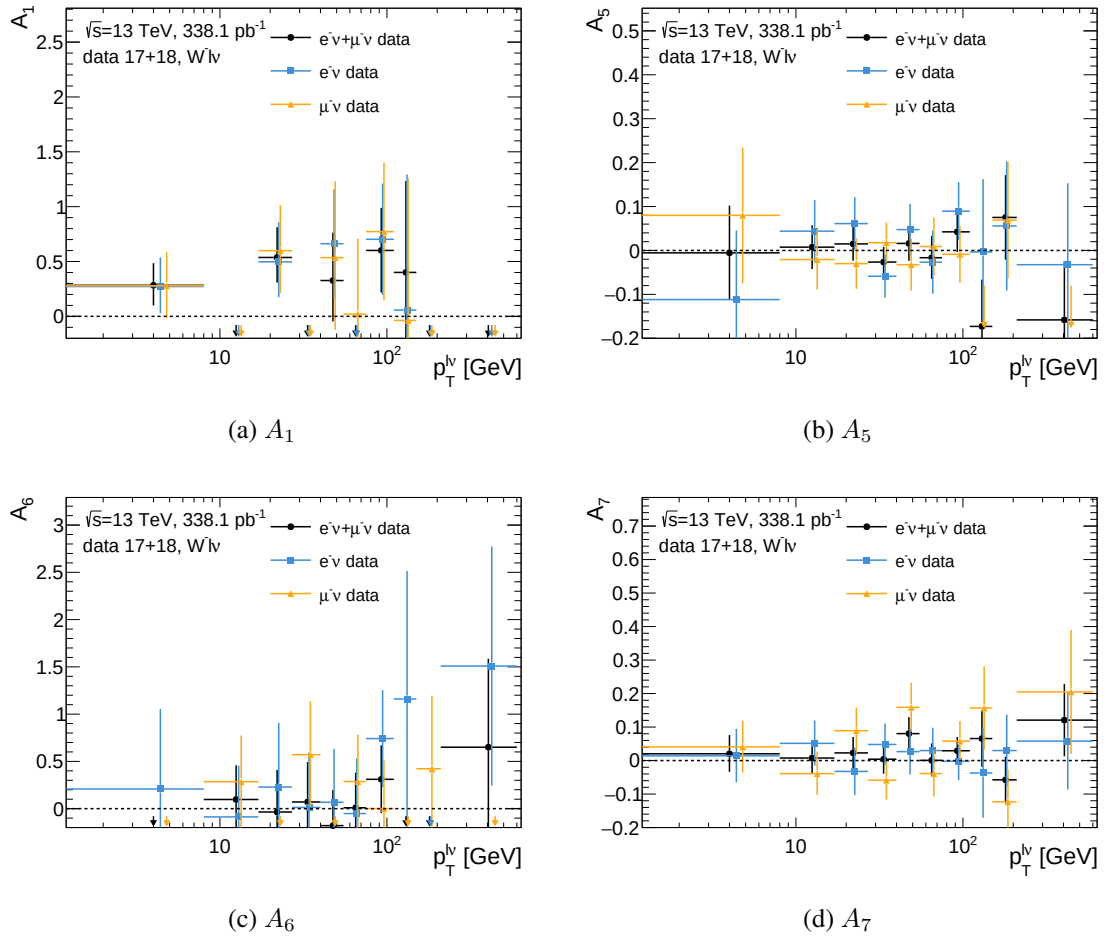
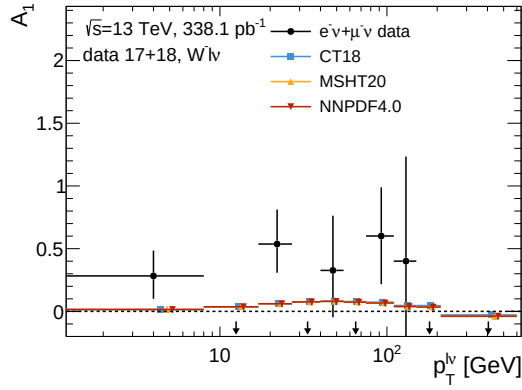
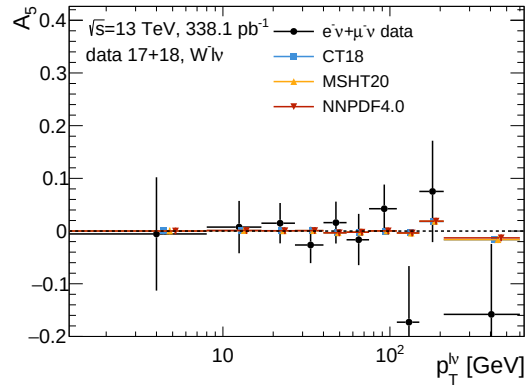


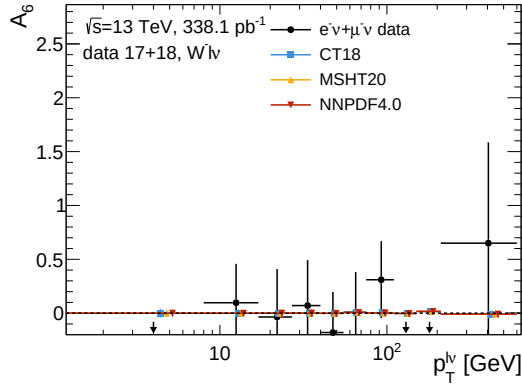
Figure G.44: The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The combined channel measurements can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



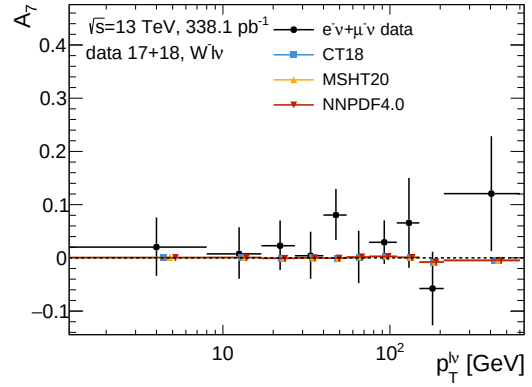
(a) A_1



(b) A_5

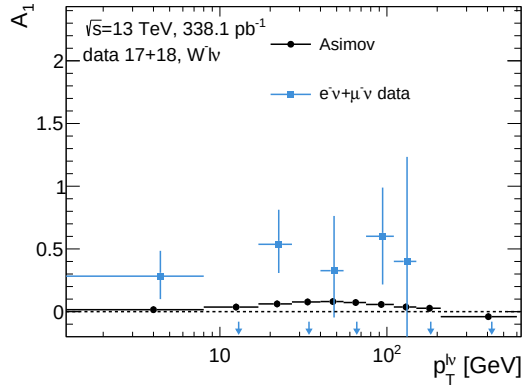


(c) A_6

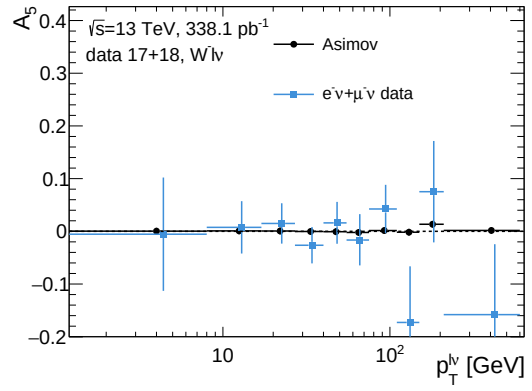


(d) A_7

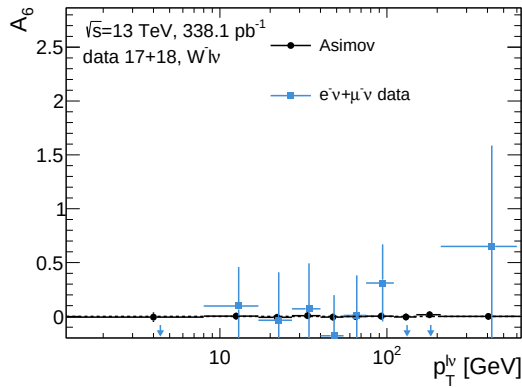
Figure G.45: The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The combined channel measurements can be seen in black and are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



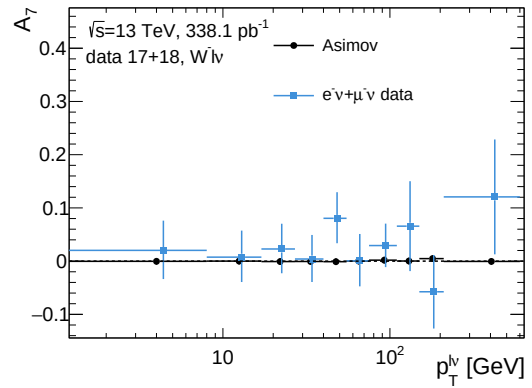
(a) A_1



(b) A_5

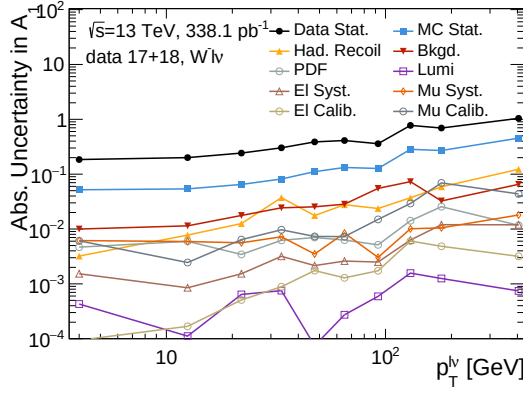


(c) A_6

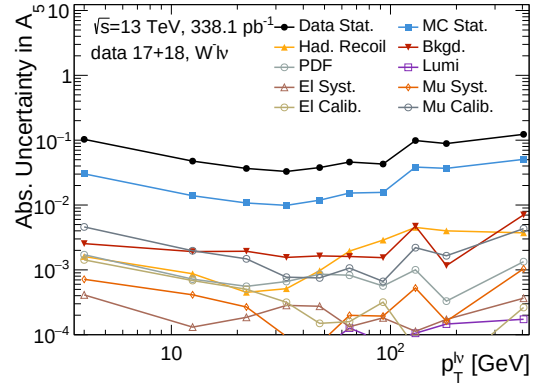


(d) A_7

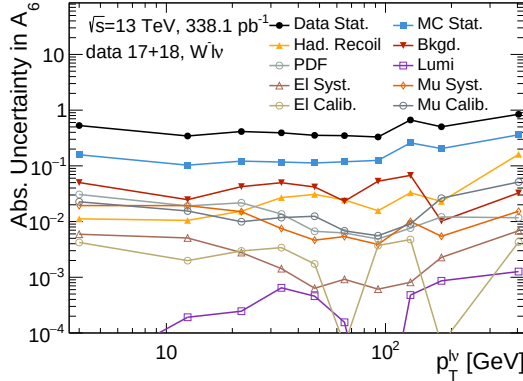
Figure G.46: The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combination of the $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



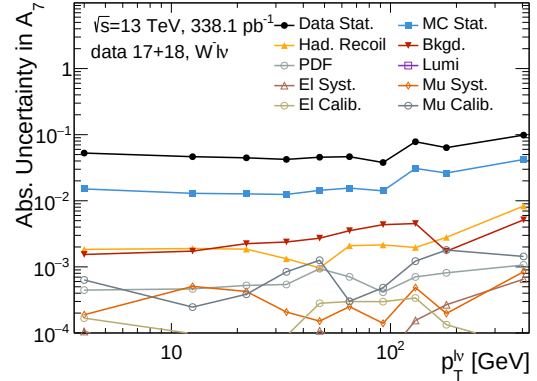
(a) Uncertainties for A_1



(b) Uncertainties for A_5



(c) Uncertainties for A_6



(d) Uncertainties for A_7

Figure G.47: A breakdown of all sources of measured uncertainty for the measured angular coefficients, A_1 , A_5 , A_6 , and A_7 , as a function of $p_T^{l\nu}$ for Figure G.46 for the combination of $W^- \rightarrow e^- \bar{\nu}$ and $W^- \rightarrow \mu^- \bar{\nu}$ channels. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.

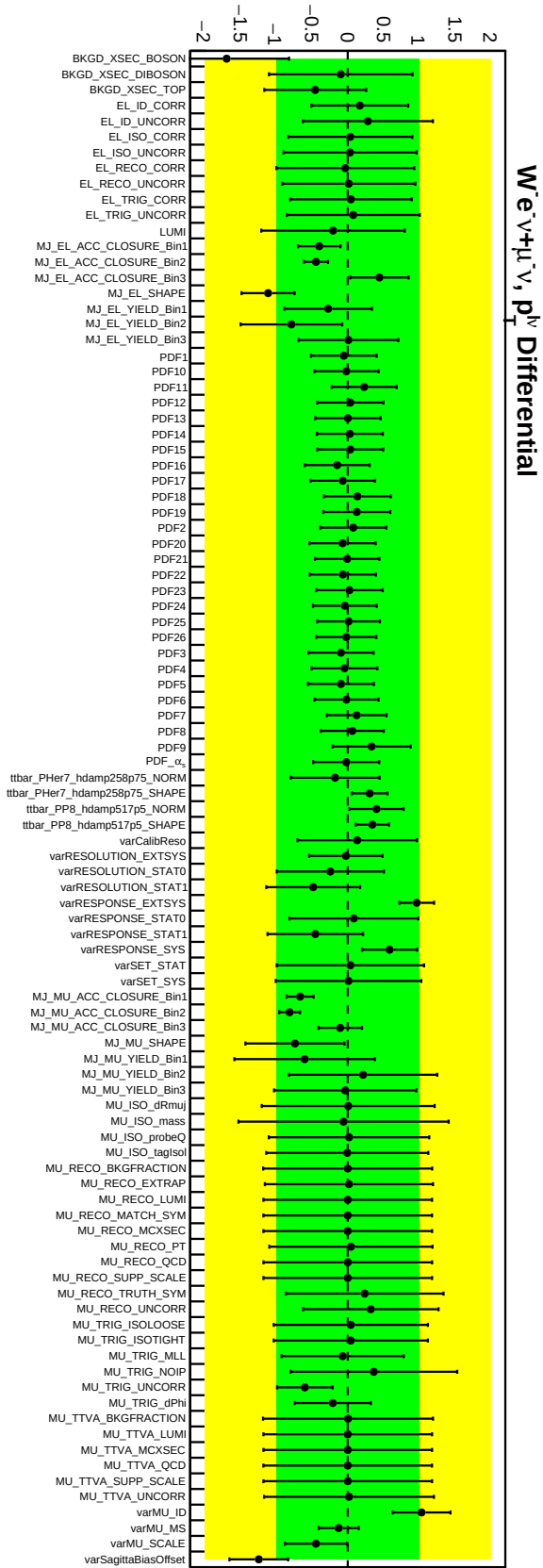
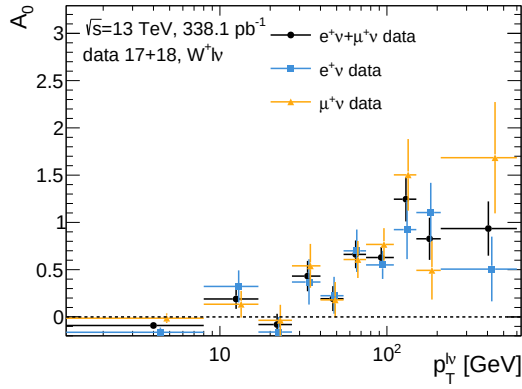


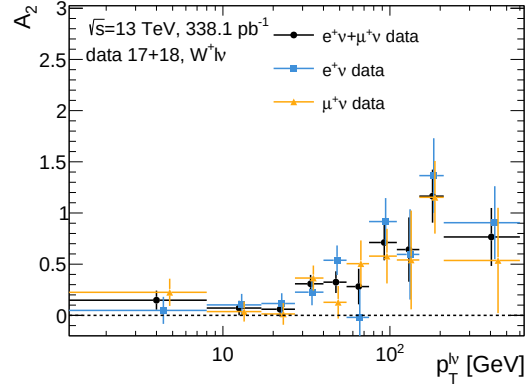
Figure G.48: The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of $p_T^{\ell\nu}$ for the combination of $W^- \rightarrow e^-\bar{\nu}$ and $W^- \rightarrow \mu^-\bar{\nu}$ channels. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty.

G.6 Additional W^+ Combined Channels Figures

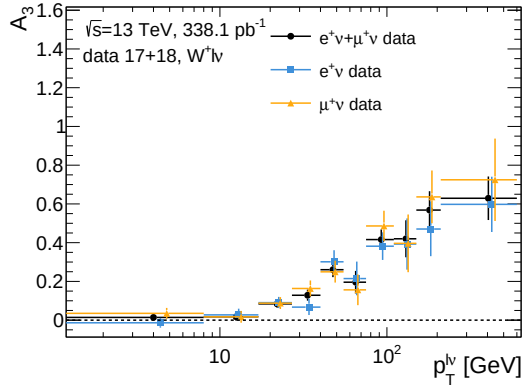
The only figures previously shown for the combined $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels were the measured values and uncertainties of the four major coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section in Chapter 11. Only a comparison to Asimov data for these measured coefficients and differential cross-section were shown. This section will show the comparisons of these measured values to the individual channels as well the other predictions of Asimov data reweighted used different PDF sets. The measured values and uncertainties for the coefficients – A_1 , A_5 , A_6 , and A_7 –, which are measured but not statistically different from zero, will also be shown. This section will present the plot of the NPs pulls and constraints that occurred during the combined channel fit.



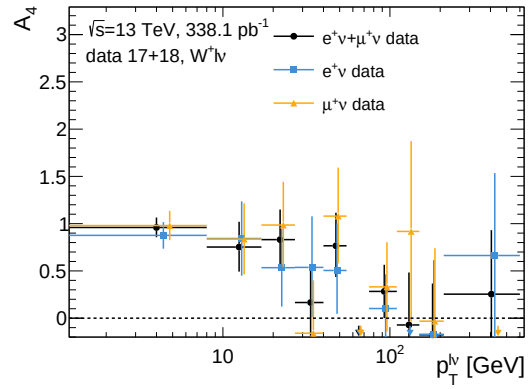
(a) A_0



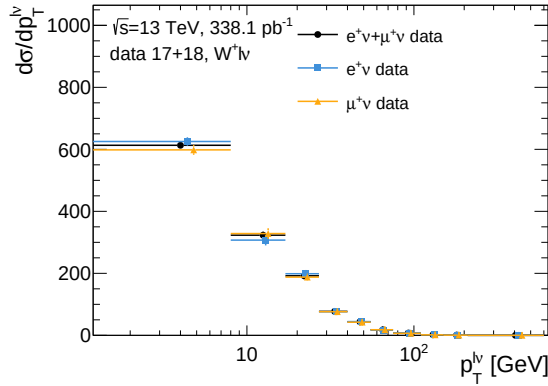
(b) A_2



(c) A_3

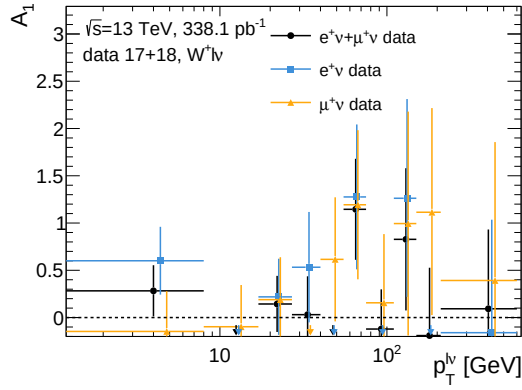


(d) A_4

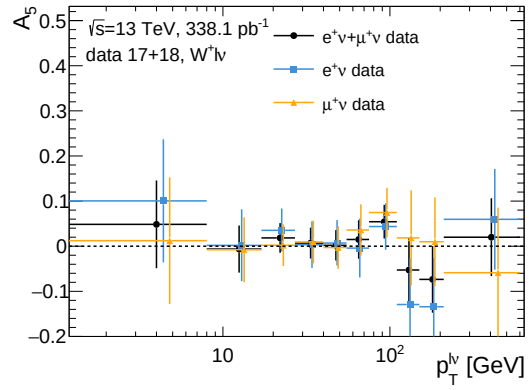


(e) Differential cross-section

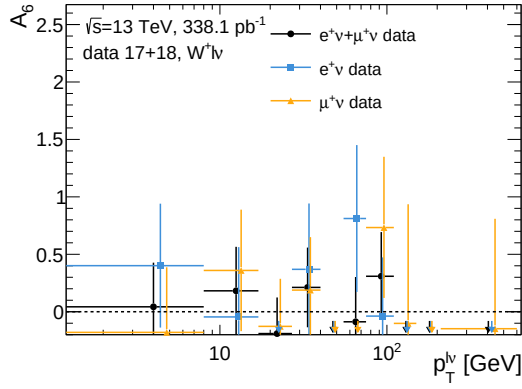
Figure G.49: The measured values of the four major coefficients – A_0 , A_2 , A_3 , A_4 – and the differential cross-section as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The combined channel measurement can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



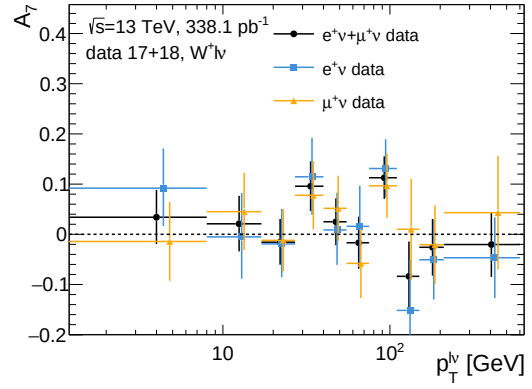
(a) A_1



(b) A_5



(c) A_6



(d) A_7

Figure G.50: The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The combined channel measurement can be seen in black, while the individual channels for comparison are shown in yellow and blue for $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$, respectively. The blue and yellow points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.

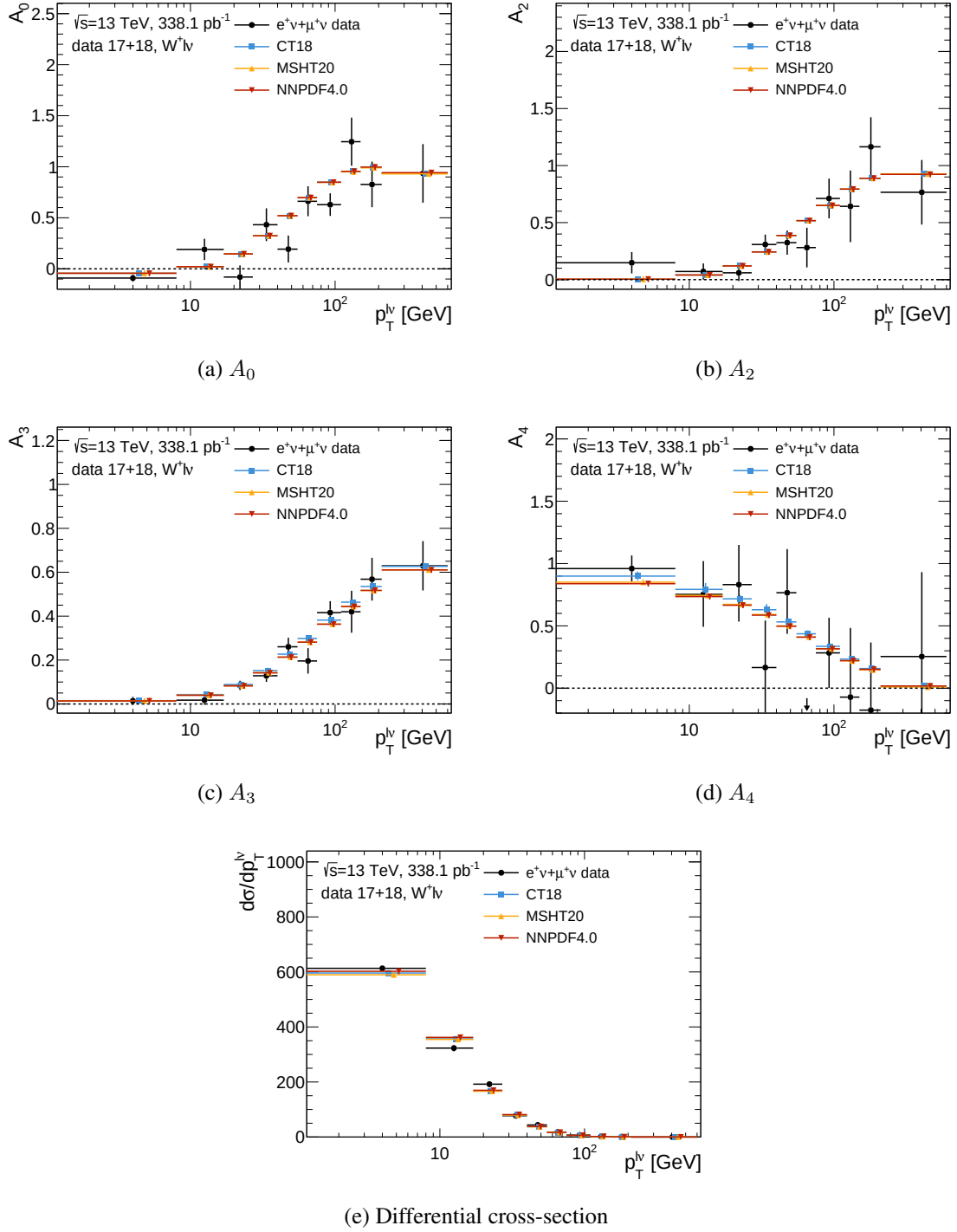
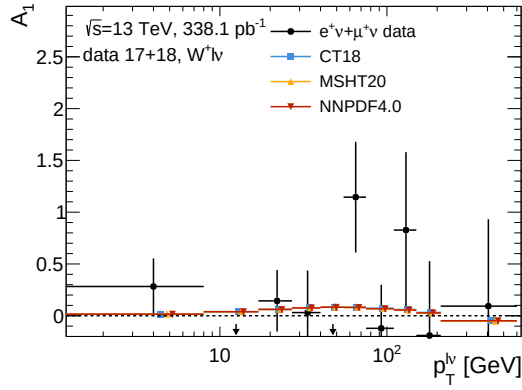
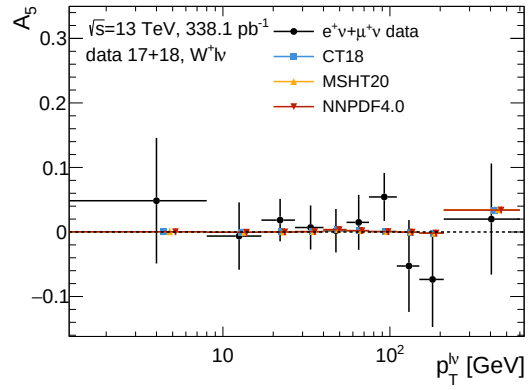


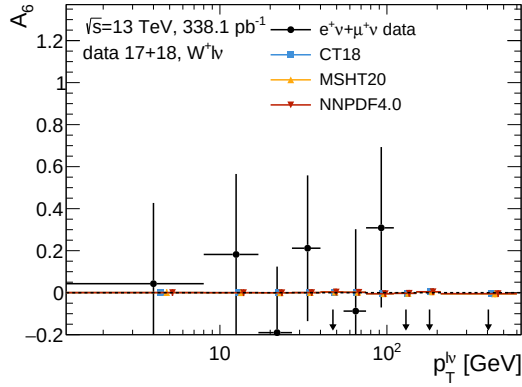
Figure G.51: The measured values of coefficients A_0 , A_2 , A_3 , A_4 , and the differential cross-section as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The combined channel measurements can be seen in black and are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



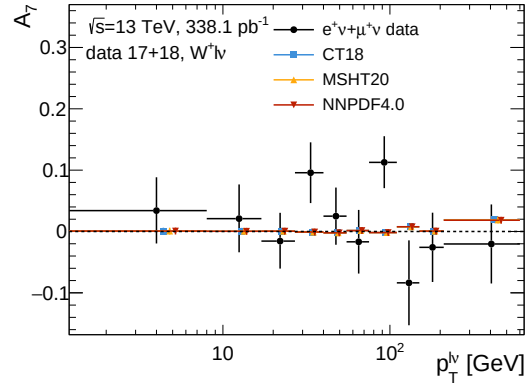
(a) A_1



(b) A_5

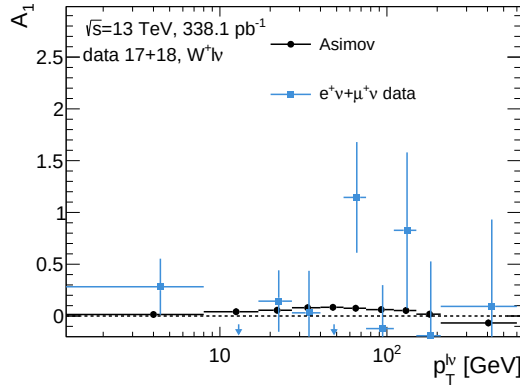


(c) A_6

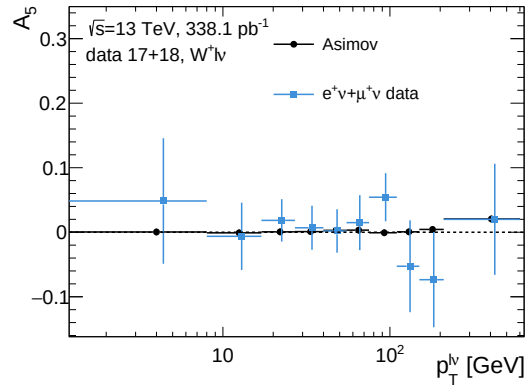


(d) A_7

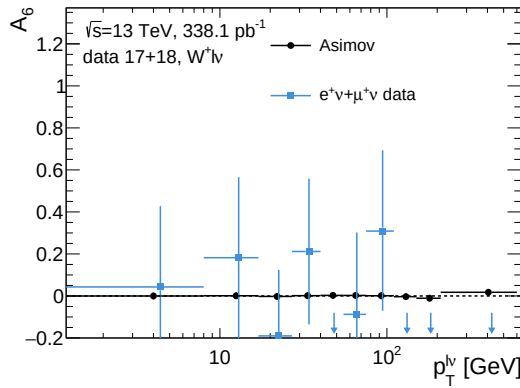
Figure G.52: The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combined fit of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The combined channel measurements can be seen in black and are compared to different predictions of Asimov data reweighted using different PDF sets: CT18 [114] in blue, MSHT20 [115] in yellow, and NNPDF4.0 [116] in red. The coloured points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



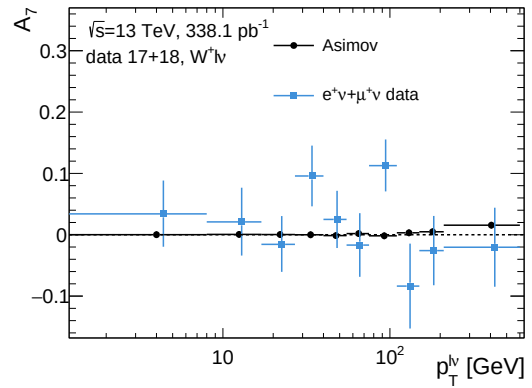
(a) A_1



(b) A_5

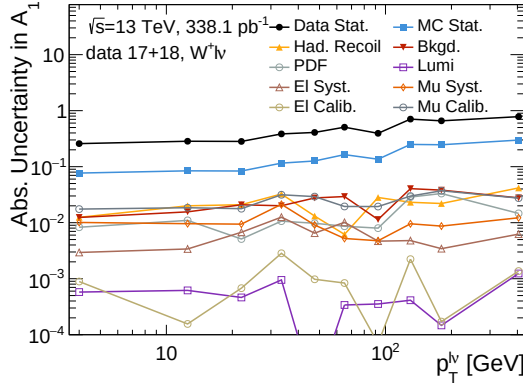


(c) A_6

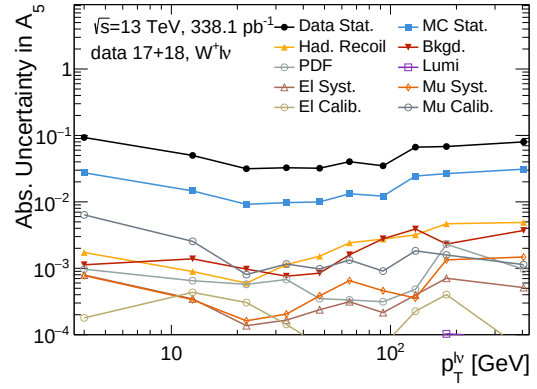


(d) A_7

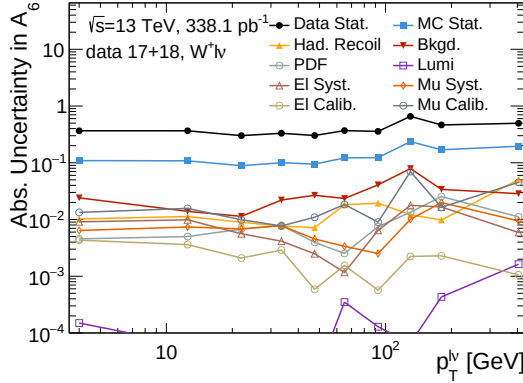
Figure G.53: The measured values of coefficients A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for the combination of the $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The measured values with data can be seen in blue squares and are compared to the expected coefficients from Asimov data in black. The blue points are slightly shifted horizontally to avoid overlap with the black points and their uncertainties.



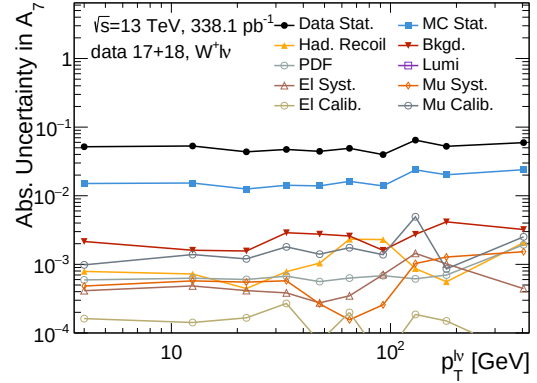
(a) Uncertainties for A_1



(b) Uncertainties for A_5



(c) Uncertainties for A_6



(d) Uncertainties for A_7

Figure G.54: A breakdown of all sources of measured uncertainty for the measured angular coefficients, A_1 , A_5 , A_6 , and A_7 as a function of $p_T^{l\nu}$ for Figure G.53 for the combination of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The data statistical uncertainty is outlined in Section 10.1, denoted by the black points, and all sources of systematic uncertainty are outlined in Section 10.2.

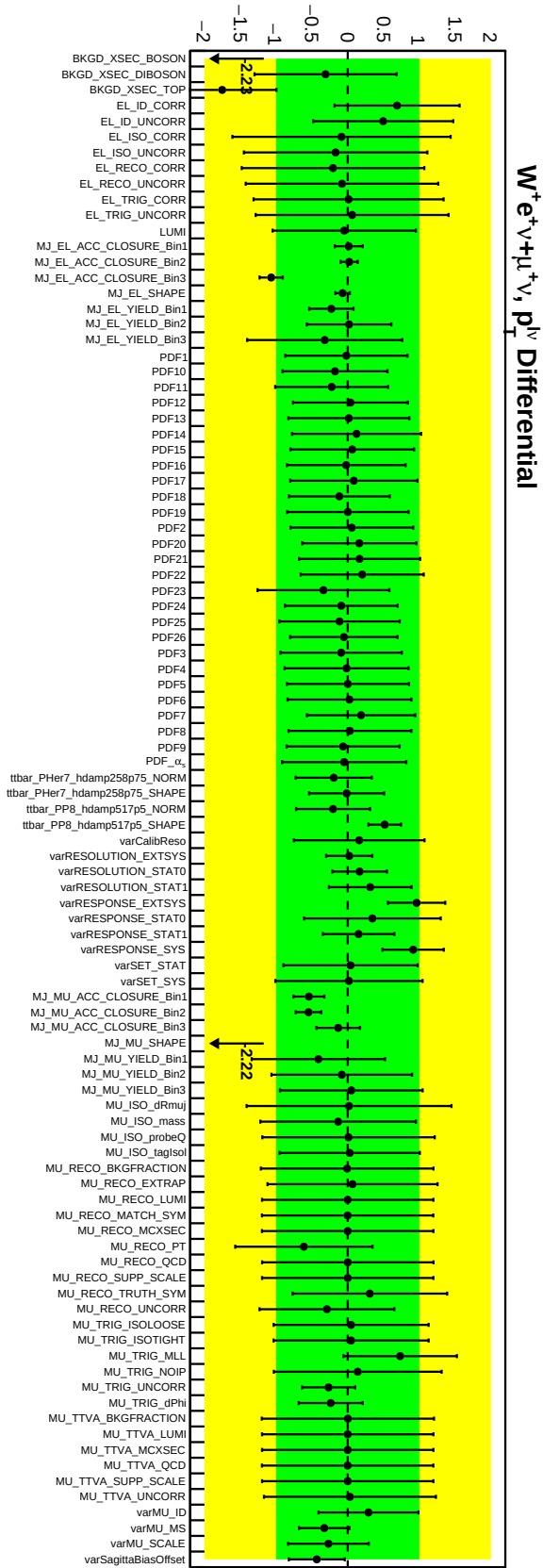


Figure G.55: The post-fit pulls of the systematic NPs listed in alphanumeric order from the measurement of the Drell–Yan angular coefficients and differential cross-section as a function of p_T^{ν} for the combination of $W^+ \rightarrow e^+\nu$ and $W^+ \rightarrow \mu^+\nu$ channels. The pre-fit estimation of the systematic NPs is described with the black dashed line of mean zero and unit standard deviation described by the green band. The post-fit pulls are denoted by the black points and any constraint on the post-fit uncertainty is described by the error bars. The systematics are appropriately modeled as most points are within one standard deviation of the pre-fit uncertainty.