Differential cross-section measurements of $WbWb$ production in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS detector

by:

Bohan CHEN
Bachelor of Science, MCGILL UNIVERSITY, 2017
Master of Science, MCGILL UNIVERSITY, 2019

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Physics and Astronomy

University of Victoria

Differential cross-section measurements of $WbWb$ production in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS detector

by:

Bohan CHEN
Bachelor of Science, MCGILL UNIVERSITY, 2017
Master of Science, MCGILL UNIVERSITY, 2019

SUPERVISORY COMMITTEE

Dr. Robert KOWALEWSKI, Supervisor
Department of Physics and Astronomy

Dr. Robert MCPHERSON, Departmental Member
Department of Physics and Astronomy

Dr. Henning STRUCHTRUP, Outside Member
Department of Mechanical Engineering

ABSTRACT

The top quark provides a unique opportunity to test the Standard Model of particle physics. Its heavy mass causes it to decay before it can hadronize, allowing for the study of the properties of a bare quark. This thesis presents differential cross-section measurements stemming from various top quark processes. This measurement is conducted in the presently uncovered semileptonic channel, where one W boson decays leptonically and the other decays hadronically. The analysis is motivated by the need to improve the modeling of the interference effects between singly-resonant and doubly-resonant top quark decays. This effect has been the source of significant modeling uncertainties in other top-sector analyses. The data and simulation used to perform this analysis correspond to the full ATLAS Run 2 dataset, collected 2015-2018. This data corresponds to an integrated luminosity of 140 fb^{-1} and a center of mass energy of $\sqrt{s} = 13$ TeV. The measured cross-sections are compared to predictions using different combinations of MC generators. Together with a complementary analysis in the dileptonic channel, this thesis provides a means to constrain modeling uncertainties towards the development of an all-inclusive $bb4l$ generator. This generator could model all top quark processes while accurately accounting for all interference effects.

Table of Contents

Supervisory Committee . . . ii

Abstract . . . iii

Table of Contents . . . iv

List of Tables . . . ix

List of Figures . . . xi

List of Acronyms . . . xiv

Declaration . . . xvi

Acknowledgements . . . xix

Dedication . . . xx

1 Introduction . . . 1

2 Motivation . . . 5

3 Theory . . . 11

3.1 The Standard Model . . . 11

3.1.1 Fermions and Bosons . . . 11

3.1.2 Mathematical Definition of the Standard Model . . . 13

3.1.3 Weak Interactions . . . 13

3.1.4 Quantum Electro-Dynamics . . . 15

3.1.5 Quantum Chromo-Dynamics . . . 16

3.2 Top Quark Properties . . . 18

3.3 $pp \to WbWb$ Processes in the SM . . . 22

3.4 Methods for Computing Top Quark Predictions . . . 24

3.5 $t\bar{t}$ and tW Quantum Interference . . . 25

4 LHC and ATLAS . . . 31

4.1 Large Hadron Collider . . . 31
4.1.1 Pre-Accelerator Chain . . . 32
4.1.2 Radio-Frequency System . . . 33
4.1.3 Magnets System . . . 34
4.1.4 Luminosity . . . 34
4.2 ATLAS Detector . . . 36
4.2.1 Coordinate System . . . 39
4.2.2 Inner Detector . . . 40
4.2.2.1 Pixel Detector . . . 41
4.2.2.2 Semi-Conductor Tracker . . . 41
4.2.2.3 Transition Radiation Tracker . . . 42
4.2.3 Calorimeters . . . 42
4.2.3.1 Electromagentic Calorimeter . . . 44
4.2.3.2 Hadronic Calorimeter . . . 45
4.2.4 Muon Spectrometer . . . 46
4.2.4.1 Monitored Drift Tubes . . . 47
4.2.4.2 Cathode Strip Chambers . . . 49
4.2.4.3 Resistive Plate Chambers . . . 50
4.2.4.4 Thin Gap Chambers . . . 52
4.2.5 Trigger and Data Acquisition . . . 53
4.2.5.1 L1 Trigger . . . 54
4.2.5.2 High-Level Trigger . . . 55

5 Data and Monte-Carlo Samples . . . 57
5.1 Monte-Carlo Event Generation . . . 57
5.1.1 Cross-Section Calculations . . . 58
5.1.2 Hard Subprocess . . . 59
5.1.3 Parton Showering . . . 59
5.1.3.1 General-Purpose vs. ME/NLO+PS Event Generators . . . 61
5.1.4 Hadronization . . . 61
5.1.5 Multiple Particle Interactions . . . 62
5.1.6 Hadron Decay . . . 63
5.1.7 Detector Simulation . . . 63
5.1.8 Detector, Parton, Particle-Level . . . 63
5.1.9 Monte-Carlo Reweighting and Scaling . . . 64
5.1.9.1 Statistical Uncertainties . . . 66
5.2 Data Samples . . . 66
5.3 Monte-Carlo Samples . . . 66
5.4 Fixed vs. Dynamical Scaling . . . 68

6 Object Reconstruction and Definition . . . 71
6.1 Baseline Objects Reconstruction . . . 71
6.1.1 Primary Vertex . . . 72
6.1.1.1 Primary Vertex Requirements . . . 73
6.1.2 Electrons . . . 73
6.1.2.1 Electron Reconstruction . . . 73
6.1.2.2 Electron Calibration . . . 74
6.1.2.3 Electron Identification . . . 75
6.1.2.4 Electron Isolation . . . 76
6.1.2.5 Electron Requirements . . . 78
6.1.3 Muons . . . 79
6.1.3.1 Muon Reconstruction . . . 79
6.1.3.2 Muon Calibration . . . 80
6.1.3.3 Muon Identification . . . 81
6.1.3.4 Muon Isolation . . . 82
6.1.3.5 Muon Requirements . . . 83
6.1.4 Jets . . . 84
6.1.4.1 Jet Tagging . . . 85
6.1.4.2 Jet Requirements . . . 89
6.1.5 Missing Transverse Energy . . . 90
6.1.6 Overlap Removal . . . 93
6.2 Leptonic and Hadronic W Reconstruction . . . 93
6.2.1 Advantages of a Semileptonic Final State . . . 93
6.2.2 Leptonic-W Reconstruction . . . 94
6.2.3 Hadronic-W Reconstruction . . . 96
6.3 Minimax Algorithm . . . 101
6.4 Fake Lepton Estimation . . . 101
6.4.1 Fake-Factor Method . . . 102
6.4.2 Fakes Enriched Region Definition . . . 103
6.4.3 Electron Fake Factors . . . 104
6.4.4 Muon Fake Factors . . . 105
6.4.5 Fake Factors Uncertainties . . . 105
6.4.5.1 Electron Fake Factors Uncertainties . . . 106
6.4.5.2 Muon Fake Factors Uncertainties . . . 108

7 Region Definitions . . . 111
7.1 Baseline Selection . . . 111
7.2 Measurement Regions . . . 115
7.3 Selected Observables of Interest . . . 117
7.4 Detector-Level Control Plots . . . 119

8 Systematic Uncertainties . . . 123
8.1 Experimental Systematics . . . 123
8.2 Theory Systematics . . . 127
8.2.1 Signal modeling . . . 128
8.2.2 Background modeling . . . 129
8.3 Systematics Summary Plots . . . 130

9 Unfolding . . . 135
9.1 Unfolding Description . . . 135
9.1.1 Regularization . . . 136
9.2 Unfolding Ingredients . . . 139
9.3 Propagation of Uncertainties . . . 144
9.3.1 Propagation of Statistical Uncertainties . . . 144
9.3.2 Propagation of Systematic Uncertainties . . . 145
9.4 Covariance and Correlation Matrices . . . 146
9.5 Unfolded Cross-Sections . . . 150
9.5.1 Absolute Differential Cross-Sections . . . 150
9.5.2 Normalized Differential Cross-Sections . . . 154
9.6 Discussion of Results . . . 157
9.6.1 Discussion of MiNNLO Predictions . . . 160

10 Conclusion . . . 161

Bibliography . . . 163

A Neutrino Momentum Derivation . . . 175

B Additional Cutflow Tables . . . 177
B.1 Baseline Selection . . . 177
B.2 Measurement Regions . . . 179

C Alternative Signal Regions . . . 183
C.1 Detector-Level Control Plots . . . 183
C.2 Systematics Summary Plots . . . 186
C.3 Signal Modeling Systematics Plots . . . 189
C.4 Migration Matrices . . . 192
C.5 Acceptance/Efficiency Correction Factors . . . 195
C.6 Covariance Matrices . . . 198
C.7 Correlation Matrices . . . 201
C.8 Absolute Differential Cross Sections . . . 204
C.9 Normalized Differential Cross Sections . . . 210
C.10 Chi-Squared Tables . . . 216

D Unfolding Closure Tests 219

E Parton-Level Measurement Region 221

F Leading/Subleading *b*-jet Asymmetry 229

List of Tables

4.1 Selected LHC Beam Parameters for pp Collisions During Run 2 . . . 35
4.2 General Performance Goals of the ATLAS Detector . . . 38

5.1 Run 2 Integrated Luminosity and GRL Summary . . . 66
5.2 Summary of Monte-Carlo Samples used in the Analysis . . . 67

6.1 Baseline Electron Requirements . . . 78
6.2 Baseline Muon Requirements . . . 84
6.3 Baseline Jet Requirements . . . 90

7.1 Baseline Selection Weighted Event Yields . . . 114
7.2 Measurement Regions Weighted Event Yields . . . 116
7.3 Measurement Regions Final Event Yields . . . 117

8.1 Electron Uncertainties Nuisance Parameters . . . 124
8.2 Muon Uncertainties Nuisance Parameters . . . 125
8.3 Jet Uncertainties Nuisance Parameters . . . 126
8.4 b-tagging Uncertainties Nuisance Parameters . . . 126
8.5 MET Uncertainties Nuisance Parameters . . . 127
8.6 μ_r and μ_f Scales Uncertainties Nuisance Parameters . . . 128
8.7 W+jets Modeling Uncertainties Nuisance Parameters . . . 130
8.8 Other V/VV Modeling Uncertainties Nuisance Parameters . . . 130

9.1 χ^2 and p-values for the $W_{\text{had.}}$ (particle) Region . . . 158

B.1 Baseline Selection Absolute Event Yields . . . 178
B.2 Baseline Selection Marginal Event Yields . . . 178
B.3 Measurement Regions Absolute Event Yields . . . 180
B.4 Measurement Regions Marginal Event Yields . . . 181

C.1 χ^2 and p-values for the $W_{\text{had.}}$ (loose) Region . . . 217
C.2 χ^2 and p-values for the $W_{\text{lep.}}$ (loose) Region . . . 217

E.1 χ^2 and p-values for the $W_{had.}$ (parton) Region . . . 227

F.1 χ^2 and p-values for the b-jet p_{T} Asymmetry and Fraction 234

List of Figures

2.1 SM Production Cross-Sections in the ATLAS Detector . . . 6
2.2 SM Production Cross-Sections in the ATLAS Detector Compared to Theory . . . 7
2.3 Predicted DS-to-DR Ratio as a Function of Dilepton Transverse Momentum . . . 8
2.4 Predicted *bb4l* Differential Cross-Section Distributions at LO and NLO . . . 10

3.1 Standard Model of Elementary Particles . . . 12
3.2 Fundamental Charged Current Interaction Vertices . . . 14
3.3 Fundamental Neutral Current Interaction Vertex . . . 15
3.4 Fundamental QED Interaction Vertex . . . 16
3.5 Fundamental QCD Interaction Vertices . . . 18
3.6 Measured $t\bar{t}$ Inclusive Cross-Sections at ATLAS and CMS . . . 19
3.7 Measured Single-Top Inclusive Cross-Sections at ATLAS and CMS . . . 20
3.8 Top Mass Measurements from Direct top quark Decay Processes at ATLAS and CMS . . . 21
3.9 Representative $t\bar{t}$ Production Modes at the LHC . . . 22
3.10 Representative tW Production Modes at the LHC . . . 23
3.11 Representative Non-Resonant Production Modes at the LHC . . . 23
3.12 Minimax Distribution Plotting the DR/DS Interference Effect . . . 29
3.13 Unfolded Minimax Distribution Plotting the DR/DS Interference Effect . . . 30

4.1 CERN Accelerator Complex and LHC Injection Systems . . . 33
4.2 Selected LHC/ATLAS Luminosity Summary Plots . . . 36
4.3 Labeled Cutaway of the ATLAS Detector . . . 38
4.4 Schematic of the Common ATLAS Coordinate System . . . 39
4.5 Labeled Cutaway of the ATLAS Inner Detector . . . 40
4.6 Labeled Cutaway of the ATLAS Calorimeter System . . . 43
4.7 Barrel Module of the ATLAS Electromagnetic Calorimeter . . . 44
4.8 Module of the Tile Calorimeter in the ATLAS Hadronic Calorimeter . . . 46
4.9 Labeled Cutaway of the ATLAS Muon System . . . 47
4.10 Module of the MDT in the ATLAS Muon System . . . 48
4.11 Illustration of a Fitted Muon Track Passing Through Several Drift Tubes . . . 49
4.12 Cross-Sectional View of the Multiwire Proportional Chamber of the CSC . . . 50
4.13 Module of the RPC in the ATLAS Muon System . . . 51
4.14 Layout of the RPC and TGC Modules in the Muon Trigger System . . . 52

4.15 Block Diagram of the ATLAS Trigger and Data Acquisition System 54

5.1 Total $t\bar{t}$ Cross-Sections Computed with Different Dynamic Scales and PDF Sets 69

6.1 Mass, Lifetime and Stability of Various Standard Model Particles 72
6.2 Illustrated Path of an Electron Through the ATLAS Detector 74
6.3 Electron Identification Efficiencies for Various Working Points 76
6.4 Electron Isolation Efficiencies for Various Working Points 77
6.5 Muon Identification Efficiencies for Various Working Points 82
6.6 Distribution of the Isolation Variable Divided by Muon Transverse Momentum 83
6.7 Identifying Characteristics of a Jet Initiated by a b-quark 86
6.8 DL1r b-tagging Probabilities and Discriminant in $t\bar{t}$ Events 87
6.9 Background Rejection as a Function of the b-tagging Efficiency 88
6.10 DL1r b-tagging and c-jet/light-jet Rejection as a Function of Jet p_{T} 89
6.11 Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ Observables Evaluated Using a $Z \to \mu\mu$ Sample 92
6.12 Leptonic-W Reconstruction Efficiency as a Function of p_{T} and η 96
6.13 Hadronic-W Reconstruction Efficiency as a Function of p_{T} and η 98
6.14 Hadronic-W Reconstruction Efficiency as a Function of p_{T} and η with an Additional W Mass Requirement 100
6.15 Electron Fake-Factors Plotted as a Function of p_{T} in Different Bins of η 104
6.16 Muon Fake-Factors Plotted as a Function of p_{T} in Different Bins of η 105
6.17 Electron Fake-Factors with Systematic Uncertainties 107
6.18 Muon Fake-Factors with Systematic Uncertainties 109

7.1 Event Weighted/Normalized Baseline Selection Cutflow 112
7.2 Detector-Level Observables in the $W_{\mathrm{had.}}$ (particle) Region 120

8.1 Systematic Uncertainties Summary for the $W_{\mathrm{had.}}$ (particle) Region 131
8.2 Summary of Signal Modeling Uncertainties for the $W_{\mathrm{had.}}$ (particle) Region 133

9.1 Representative L-Curve Plot 139
9.2 Migration Matrices for the $W_{\mathrm{had.}}$ (particle) Region 142
9.3 Acceptance/Efficiency Factors for the $W_{\mathrm{had.}}$ (particle) Region 143
9.4 Covariance Matrices for the $W_{\mathrm{had.}}$ (particle) Region 148
9.5 Correlation Matrices for the $W_{\mathrm{had.}}$ (particle) Region 149
9.6 Absolute Differential Cross-Sections (Nominal MC) for the $W_{\mathrm{had.}}$ (particle) Region . . . 152
9.7 Absolute Differential Cross-Sections (Alternative MC) for the $W_{\mathrm{had.}}$ (particle) Region . . 153
9.8 Normalized Differential Cross-Sections (Nominal MC) for the $W_{\mathrm{had.}}$ (particle) Region . 155
9.9 Normalized Differential Cross-Sections (Alternative MC) for the $W_{\mathrm{had.}}$ (particle) Region 156

C.1 Detector-Level Observables in the $W_{\mathrm{had.}}$ (loose) Region 184
C.2 Detector-Level Observables in the $W_{\mathrm{lep.}}$ (loose) Region 185

C.3 Systematic Uncertainties Summary for the $W_{\text{had.}}$ (loose) Region . . . 187
C.4 Systematic Uncertainties Summary for the $W_{\text{lep.}}$ (loose) Region . . . 188
C.5 Summary of Signal Modeling Uncertainties for the $W_{\text{had.}}$ (loose) Region . . . 190
C.6 Summary of Signal Modeling Uncertainties for the $W_{\text{lep.}}$ (loose) Region . . . 191
C.7 Migration Matrices for the $W_{\text{had.}}$ (loose) Region . . . 193
C.8 Migration Matrices for the $W_{\text{lep.}}$ (loose) Region . . . 194
C.9 Acceptance/Efficiency Factors for the $W_{\text{had.}}$ (loose) Region . . . 196
C.10 Acceptance/Efficiency Factors for the $W_{\text{lep.}}$ (loose) Region . . . 197
C.11 Covariance Matrices for the $W_{\text{had.}}$ (loose) Region . . . 199
C.12 Covariance Matrices for the $W_{\text{lep.}}$ (loose) Region . . . 200
C.13 Correlation Matrices for the $W_{\text{had.}}$ (loose) Region . . . 202
C.14 Correlation Matrices for the $W_{\text{lep.}}$ (loose) Region . . . 203
C.15 Absolute Differential Cross-Sections (Nominal MC) for the $W_{\text{had.}}$ (loose) Region . . . 205
C.16 Absolute Differential Cross-Sections (Nominal MC) for the $W_{\text{lep.}}$ (loose) Region . . . 206
C.17 Absolute Differential Cross-Sections (Alternative MC) for the $W_{\text{had.}}$ (loose) Region . . . 208
C.18 Absolute Differential Cross-Sections (Alternative MC) for the $W_{\text{lep.}}$ (loose) Region . . . 209
C.19 Normalized Differential Cross-Sections (Nominal MC) for the $W_{\text{had.}}$ (loose) Region . . . 211
C.20 Normalized Differential Cross-Sections (Nominal MC) for the $W_{\text{lep.}}$ (loose) Region . . . 212
C.21 Normalized Differential Cross-Sections (Alternative MC) for the $W_{\text{had.}}$ (loose) Region . . . 214
C.22 Normalized Differential Cross-Sections (Alternative MC) for the $W_{\text{lep.}}$ (loose) Region . . . 215

D.1 Unfolding Closure Test for the $W_{\text{had.}}$ (particle) Region . . . 220

E.1 Migration Matrices Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions . . . 222
E.2 Acceptance/Efficiency Factors Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions . . . 224
E.3 Differential Cross-Sections Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions . . . 225
E.4 Differential Cross-Sections Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions . . . 226

F.1 b-jet p_{T} Asymmetry and Fraction Detector-Level Distributions . . . 230
F.2 b-jet p_{T} Asymmetry and Fraction Systematics Uncertainties Summary . . . 231
F.3 b-jet p_{T} Asymmetry and Fraction Absolute Differential Cross-Sections (Nominal MC) . . . 232
F.4 b-jet p_{T} Asymmetry and Fraction Absolute Differential Cross-Sections (Alternative MC) 233

List of Acronyms

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC ApparatuS
BSM	Beyond the Standard Model
CDF	Collider Detector at Fermilab
CERN	Conseil Européen pour la Recherche Nucléaire
CKKW	Catani-Krauss-Kuhn-Webber
CKM	Cabibbo-Kobayashi-Maskawa
CMS	Compact Muon Solenoid
CP	Charge-Parity
CSC	Cathode Strip Chamber
DL1	Deep-Learning 1
DR	Diagram-Removal
DS	Diagram-Subtraction
DØ	D-Zero
EM	Electro-Magnetic
FSR	Final State Radiation
GPMC	General-Purpose Monte-Carlo
GRL	Good Runs List
GSF	Gaussian-Sum Filter
HCAL	Hadronic Calorimeter
HLT	High-Level Trigger
ID	Inner Detector
ISR	Initial State Radiation
JVT	Jet Vertex Tagger
L1	Level-1 Trigger
LAr	Liquid Argon
LEP	Large Electron-Positron
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LINAC2	Linear Accelerator 2
LINAC4	Linear Accelerator 4
LO	Leading-Order
MC	Monte-Carlo
MC@NLO	Monte-Carlo at Next-to-Leading-Order
MDT	Monitored Drift Tube
ME	Matrix Element
MET	Missing Transverse Energy

MS	Muon Spectrometer
NDF	Number of Degrees of Freedom
NLO	Next-to-Leading-Order
NNLL	Next-to-Next-Leading-Logarithmic
NNLO	Next-to-Next-Leading-Order
NP	Nuisance Parameter
NSW	New Small Wheel
NWA	Narrow Width Approximation
PDF	Parton Distribution Function
PDG	Particle Data Group
PFO	Particle Flow Object
PS	Parton Shower
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster
PV	Primary Vertex
QCD	Quantum Chromo-Dynamics
QED	Quantum Electro-Dynamics
QFT	Quantum Field Theory
RF	Radio Frequency
RNN	Recurrent Neural-Network
RPC	Resistive Plate Chamber
SF	Scale-Factor
SM	Standard Model
SCT	Semi-Conductor Tracker
SPS	Super Proton Synchrotron
sTGC	Small-Strip Thin Gap Chamber
SV	Secondary Vertex
TDAQ	Trigger and Data Acquisition
TGC	Thin Gap Chamber
TRT	Transition Radiation Tracker

DECLARATION

The work presented in this thesis uses data and simulations corresponding to the Run 2 data-taking period provided by the ATLAS collaboration. The collaboration organizes its physics program into different subgroups. This analysis falls under the Top Cross-Section subgroup, which produces differential and total cross-section measurements of various top quark processes. This subgroup is part of the larger Top Physics working group, which studies top quark properties and production at the LHC. The lepton+jets $WWbb$ analysis team comprises approximately ten active analyzers, supervisors, and faculty members. The analysis team is responsible for the development of the analysis strategy, the implementation of the analysis, and the interpretation of the results. The author of this thesis is a member of the lepton+jets $WWbb$ analysis team and has contributed to the analysis and collaboration efforts in the following ways:

Analysis Contributions

- Developed an initial analysis framework used to process the entire Run 2 dataset. This framework implemented all of the tools and cuts used to perform the preliminary characterization studies of the data. After a detailed comparison of independently developed frameworks, a common framework was adopted to facilitate collaboration on the analysis.
- Optimized the event selection criteria by studying the impact of different cuts on the signal and background yields. The final event selection presented in this document follows from these studies and is the result of a collaborative effort by the author and another analysis team member.
- Implemented the leptonic-W reconstruction algorithm by combining a lepton and missing energy using a W-mass kinematic constraint.
- Developed and implemented different hadronic-W reconstruction algorithms. The unique feature of this topology allows for the reconstruction of the hadronic-W using different combinations of jets. The author studied the effects of selecting the hadronic-W based on different figures of merit.
- Verified the performance of the various hadronic-W reconstruction algorithms by comparing the kinematics of the reconstructed W boson to the true W boson in simulation. The final analysis strategy uses the best-performing algorithm to define various signal regions and to produce the final cross-section distributions.
- Introduced a modified minimax variable using the leading and subleading b-jets, the hadronic-W and the leading lepton as a variable that might be sensitive to the interference effects between singly- and doubly-resonant top quark processes.
- Assigned the b-jets to their corresponding top quark parents using the minimax variable.

- Reconstructed the hadronically decaying top quark using the hadronic-W and the selected b-jet output from the minimax variable. A proxy for the leptonically decaying top quark is reconstructed using the lepton and the other b-jet.
- In collaboration with other analysis team members, the author compiled the set of interesting observables that are characteristic of the $WWbb$ process.
- Produced detector-level control plots of these observables of interest and also studied the impact of the various systematic uncertainties on the final cross-section measurement.
- Optimized the binning of the final cross-section measurement by studying the statistical uncertainty and purity in each bin.
- Developed the strategy for unfolding the detector-level distributions to the particle-level through studies of the unfolding software, TUnfold. Specifically, the author's work on developing a scheme for filling the migration matrices, defining the acceptance and efficiency corrections, and selection of the regularization parameter using the L-curve method is implemented into the common analysis framework and used by the analysis team to perform the unfolding.
- Using the common analysis framework and plotting tools, the author produced the detector-level plots, the unfolding diagnostic plots, and the final cross-section measurement plots.
- Contributed to the collective internal note that documents the analysis strategy by writing different sections, reviewing the note, and providing feedback to other analyzers.
- Participated in regular weekly meetings with the analysis team to present the progress of the analysis.
- Presented the progress and results of the analysis to the Top Cross-Section subgroup.

Collaboration Contributions

- Completed an ATLAS qualification task investigating novel approaches to jet clustering using Particle-Flow Objects (PFOs) towards calculating Missing Transverse Energy (MET). The author presented the results of this project to the general trigger community.
- Performed Jet/Calo/MET trigger expert shifts during the start of the Run 3 data-taking period. The author was responsible for monitoring the performance of these three trigger signatures during daily data-taking runs and logging any defects that would arise.
- Active participant in the MET trigger group. The author contributed to the development of the MET trigger monitoring tools and the MET trigger performance studies. The author was also previously responsible for validating the MET signature for periodic data reprocessings and signing off on these reprocessings.
- The author currently acts as the MET trigger software expert. The author is responsible for maintaining the MET trigger software and investigating any bugs in the software.

- Studied the potential of using machine-learning in discriminating between true and fake MET using calorimeter cells information. The author developed the model architecture and code used to train and test the model. The project is currently in an advanced testing phase where it will be implemented into the ATLAS trigger framework and tested on real data.

ACKNOWLEDGEMENTS

This thesis would not be possible without the support of the many people that I have worked with throughout my academic career. Firstly, I would like to thank my supervisor, **Prof. Robert Kowalewski**, for his guidance and support throughout my time in the physics and astronomy department. From you, I have learned to be a better physicist by honing my critical thinking, independent research, and problem-solving skills. I always knew that your support was just a short email away. I am grateful for the opportunity to have worked and to have learned from you. I would also like to thank my committee members, **Prof. Robert McPherson**, **Prof. Henning Struchtrup**, and **Prof. Matthew Jones**, for their time and effort in reviewing my work and providing valuable feedback.

Every step along the way, **Dr. Kenji Hamano** never hesitated to help me with any questions, no matter how small. You have supported and guided me from the very beginning with my qualification task, during the development of my analysis, and throughout the writing of this thesis. Your expertise with the ATLAS Collaboration, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger group, and conducting a physics analysis have been invaluable.

My work on the analysis would not have progressed as smoothly without the support of the other *WWbb* analysis team members. I want to specifically thank **Dr. Daniel Britzger** at the Max Planck Institute for Physics for allowing me the privilege of working on the MPPMU cluster. Together with **Johannes Hessler**, we have worked to develop the analysis and to produce the results that are presented in this thesis. Thank you for your hospitality whenever I visited Munich and the institution. I would also like to acknowledge the support offered by **Dr. Tadej Novak**. Your expertise in the development of the software tools and the analysis framework has been instrumental in the success of the analysis. I am grateful for the time you have spent to help me understand the intricacies of the analysis.

I want to extend a special thanks to **Justin Chiu** and **Danika MacDonell** who from the very beginning helped me with the ATLAS infrastructure, the software tools, and for providing me with the necessary resources to conduct my own analysis. I knew that I could always count on you both to help me with any questions I had.

Finally, to my friends, we shared many laughs and definitely made the second floor of the Elliott building a lot noisier (sorry to Bob next door). **Alexandre Beaubien**, **Maheyer Shroff**, **Alberto Nava**, **Caleb Miller**, **Michael Gennari**, and **Sahar Gholipourverki**, I will always cherish your friendship and company. I thank you for the many discussions we have had, the many meals we have shared, and the many memories we have created. I would also like to thank my friends from McGill University, **John McGowan**, **John Flores**, **Tony Kwan**, and **Rapahel Hulsken** for their support along the way.

Dedication

In loving memory of my Mom,
who taught me the virtue of hard-work.

"Abandon the idea that you are ever going to finish.
Lose track of the 400 pages and write just one page for each day, it helps.
Then when it gets finished you are always surprised."

— John Steinbeck

CHAPTER 1

Introduction

The current prevailing theory in the field of particle physics is the Standard Model (SM), which represents the primary classification of all known elementary particles and three out of four fundamental forces. Physicists compiled the present-day Standard Model in stages, adding more elementary particles through experimental observations. The model has proven to be successful in predicting the existence and basic properties of several well-known particles, including the W, Z, and the H (Higgs) bosons. Experiments later verified these properties with a high degree of precision.

The top quark (t) was one of the last particles added to the SM, partly due to its high mass. Experimentalists faced several challenges when trying to make a top quark measurement: (1) producing enough energy in a collision event to produce a top quark; (2) reconstructing a top quark from its decay products in amongst the other background processes; and (3) generating a sufficiently high number of top quark events to perform a meaningful measurement. The discovery of the top quark was officially announced in 1995 at the Tevatron particle accelerator. The two major experiments located at the Tevatron, Collider Detector at Fermilab (CDF) and D-Zero (DØ), independently reported a top mass of 176 ± 8 (stat.) $\pm$ 10 (syst.) GeV [1] and 199^{+19}_{-21} (stat.) $\pm$ 22 (syst.) GeV [2], respectively. For context, the accepted value for the top mass as reported by the Particle Data Group (PDG) is 172.69 ± 0.30 GeV [3]. The precision of this measurement underscores the rapid development of accelerator and particle detectors. Simultaneously, simulations of SM processes have also improved, allowing for more accurate predictions of the final state particles produced in a collision event.

A Toroidal LHC ApparatuS (ATLAS) [4] is the culmination of this rapid development. ATLAS is one of four major experiments at the Large Hadron Collider (LHC) and serves as a general-purpose particle detector. The ATLAS physics program is multifaceted and spans SM measurements, exotic physics searches, and dark matter searches. Physicists using ATLAS data study the properties of the final state particles, such as momentum, energy, mass, etc. Researchers then compare the distributions of these observables to the distributions predicted by theory. Any deviations between data and theory are particularly interesting to the broader physics community. The model could be better constrained to match the data in these regions. These constraints result in models that better represent the data and reduce uncertainties for future analyses.

The results in this thesis use ATLAS data collected between 2015 and 2018 during the Run 2 data-

taking period at a center-of-mass energy of $\sqrt{s} = 13$ TeV. This analysis studies the various SM processes that result in a semileptonic final state of $WWbb$. The semileptonic final state consists of one of the W-bosons decaying leptonically into a lepton and its corresponding neutrino while the second W-boson decays hadronically into a pair of quarks.

The physics observables of interest will naturally correspond to combinations of the $WWbb$ final state. The goal of the analysis is to calculate differential cross-sections of these observables. Differential cross-sections are widely used in particle physics, as they can be sensitive to potential new physics Beyond the Standard Model (BSM). Similar analyses, studying the same $WWbb$ final state, have only focussed on a single process. This analysis adopts a more comprehensive approach by considering all eligible processes that result in a $WWbb$ final state. The result is a holistic picture of the $pp \to WWbb$ process. Theorists and phenomenologists can then use these measurements to better constrain future models. The dissertation is organized as follows.

Chapter 2 motivates the study of the $WWbb$ final state in the context of top quark physics. The chapter also provides a brief overview of the interference effect between the $t\bar{t}$ and tW processes.

Chapter 3 provides a theoretical background on the SM and the SM processes that result in a $WWbb$ final state. The chapter delves deeper into the interference effect between the $t\bar{t}$ and tW processes.

Chapter 4 provides an overview of the LHC and the various subsystems making up the ATLAS detector.

Chapter 5 discusses the data and simulation samples used in this analysis. In doing so, this chapter provides a general description of the MC process. Several relevant generator-specific tunes and parameters are also discussed in this chapter.

Chapter 6 discusses the reconstruction of the physics objects used in this analysis. In addition to the basic object definitions, this chapter discusses the reconstruction of the leptonic and hadronic W-bosons. This chapter also introduces the minimax variable as a method to assign b-jets. Finally, the fake-factor method, used to estimate the background from fake leptons, is also discussed.

Chapter 7 describes the cuts used to define the different measurement regions. The chapter also lists the primary observables of interest and presents detector-level control plots of these observables in the nominal measurement region.

Chapter 8 describes the experimental and theoretical uncertainties considered in this analysis. The chapter lists the Nuisance Parameters that are varied to estimate the systematic uncertainties.

Chapter 9 describes the unfolding procedure used to extract the differential cross-sections. Unfolding is conducted using the TUnfold package and the L-curve scan method to select the regularization strength. The final absolute and normalized differential cross-sections are presented in this chapter. These cross-sections are compared to the predictions provided by different generators.

Chapter 10 concludes the dissertation by summarizing the main observations and results and how they can be used in the development of better models.

Appendices A-F present additional material that supplements the main text. These appendices include the derivation of the neutrino momentum, additional cutflow tables, collection of plots for the alternative signal regions, a closure test for the unfolding procedure, comparisons of the parton vs particle-level hadronic-W definitions, and additional observables related to b-jet p_{T} asymmetry and fraction.

CHAPTER 2

Motivation

The physics effort at the LHC roughly falls under two categories: (1) precision measurements of the SM, and (2) dark-matter/BSM searches. In recent years, dark-matter (and other BSM) processes have garnered much of the attention. Precision measurements are an equally important endeavor for several reasons. As mentioned, theorists use these measurements to constrain their models to better agree with data. These measurements can also serve as inspiration for the development of future Monte-Carlo (MC) generators. These generators are essential for many dark-matter and BSM searches to estimate SM backgrounds.

The properties of the top quark (t) continue to be a key focus area. In particular, top quark pair production ($t\bar{t}$) is extensively studied at the LHC. Differential cross-sections of various top quark observables probe the fundamental parameters of the SM, such as the top quark mass. These properties are sensitive to potential new physics wherever significant discrepancies between data and simulation arise. Naturally, it's particularly important to focus on the regions where the models require further tuning.

Top quark pair production is one of the most abundant processes observed in proton-proton collisions due to its relatively large cross-section at high energies. Figure 2.1 compares the cross-section for various SM processes across the ATLAS Run 1 and Run 2 data-taking periods.

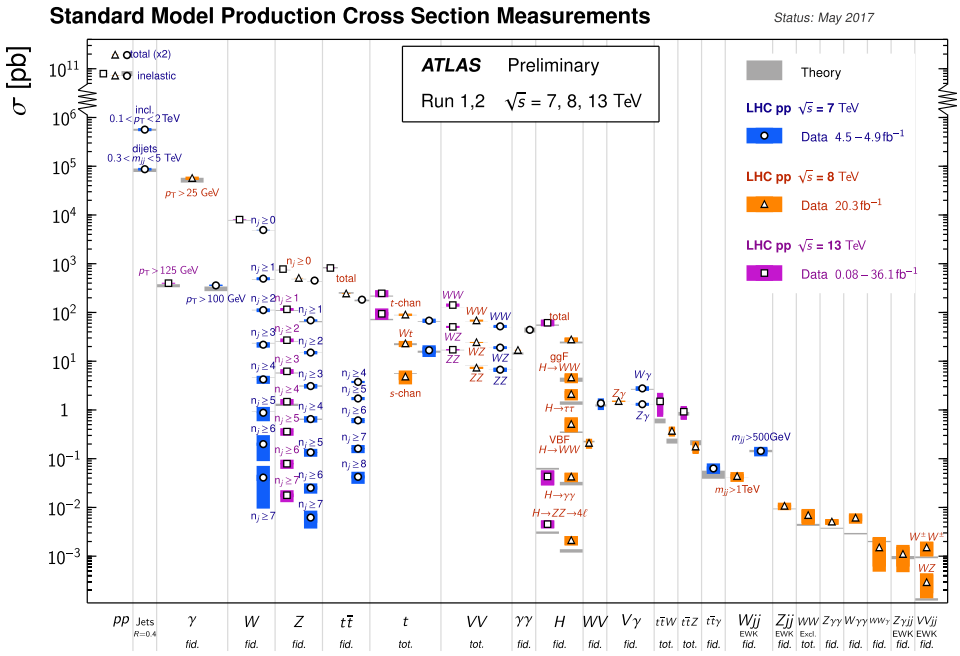


Figure 2.1: SM Production Cross-Sections in the ATLAS Detector. Cross-sections correspond to measurements taken during the Run 1 and Run 2 ATLAS data-taking periods. The center-of-mass energy during these run periods were $\sqrt{s} = 7,\ 8$ (Run 1), and 13 (Run 2) TeV. The gray-shaded boxes indicate the cross-sections as predicted by theory. The vertical length of the shaded boxes represents the total uncertainty on the cross-section measurements. Plot is taken from Ref. [5].

At the high energies of the LHC, the pp collisions in the ATLAS detector produce top quark pairs in abundance. This large cross-section is an advantage for analyses studying top quark properties. This property is especially important for some BSM analyses, where the $t\bar{t}$ process constitutes a significant proportion of the background. ATLAS continues to make precision measurements of various top quark properties for these reasons. Figure 2.2 shows the level of agreement between data and theory for various SM processes across the ATLAS Run 1 and Run 2 data-taking periods.

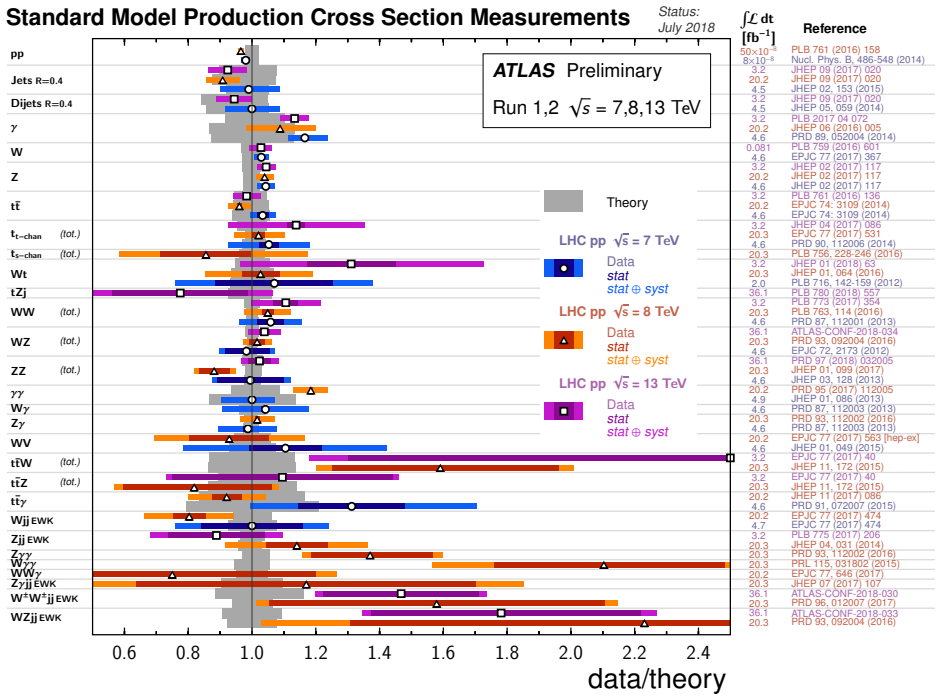


Figure 2.2: SM Production Cross-Sections in the ATLAS Detector Compared to Theory. Cross-sections correspond to measurements taken during the Run 1 and Run 2 ATLAS data-taking periods. The center-of-mass energy during these run periods were $\sqrt{s} = 7$, 8 (Run 1), and 13 (Run 2) TeV. Measurements are compared to the SM theory predictions, where the gray-shaded boxes represent the uncertainty in the theoretical predictions. The darker-colored error bars represent the purely statistical uncertainty, while the lighter-colored error bars account for all uncertainties, including systematics and luminosity. Plot is taken from Ref. [5].

Top quark pair production is one of the most precisely measured multi-particle processes in ATLAS. This precision is partly due to its high production cross-section, which significantly reduces statistical uncertainties. This process is also relatively well-modeled, which is why the $t\bar{t}$ cross-section is said to be a precision test of the SM. The slight deviation in the data-to-theory ratio from unity likely indicates that models could be tuned to represent the data better.

Top quark pair production is also the dominant process in this analysis. The number of resonant top quarks in the process defines other channels contributing to the $WWbb$ final state. The next most dominant signal process belongs in the class of singly-resonant top quark diagrams, with the dominant process being tW. Other signal contributions include diagrams with zero-resonant top quarks and higher-ordered processes with off-shell top quarks. Previous measurements targeting the same $WWbb$ final state considered either the $t\bar{t}$ [6–9] or tW [10, 11] processes as signal and treats the other processes as background.

At higher orders in perturbative Quantum Chromo-Dynamics (QCD), an interference effect exists between the doubly-resonant $t\bar{t}$ and singly-resonant tW processes because they share an identical final state. Evaluating the different approaches to modeling the single-top samples is an additional challenge for the analyses referenced above. For example, analyses performing a measurement using

the $t\bar{t}$ process would need a single-top sample decoupled from its $t\bar{t}$ counterpart. The two strategies for generating tW samples are the Diagram-Removal (DR) and the Diagram-Subtraction (DS) methods [12]. Figure 2.3 demonstrates the differences identified by the original authors of the DR/DS tW calculations as a function of the transverse dilepton momentum.

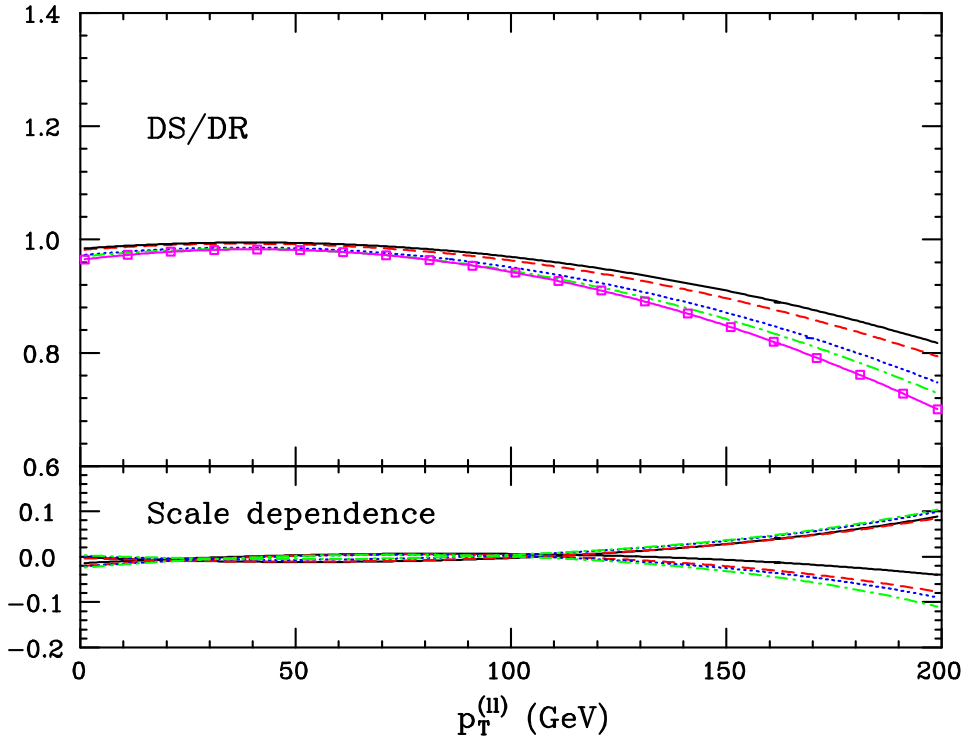
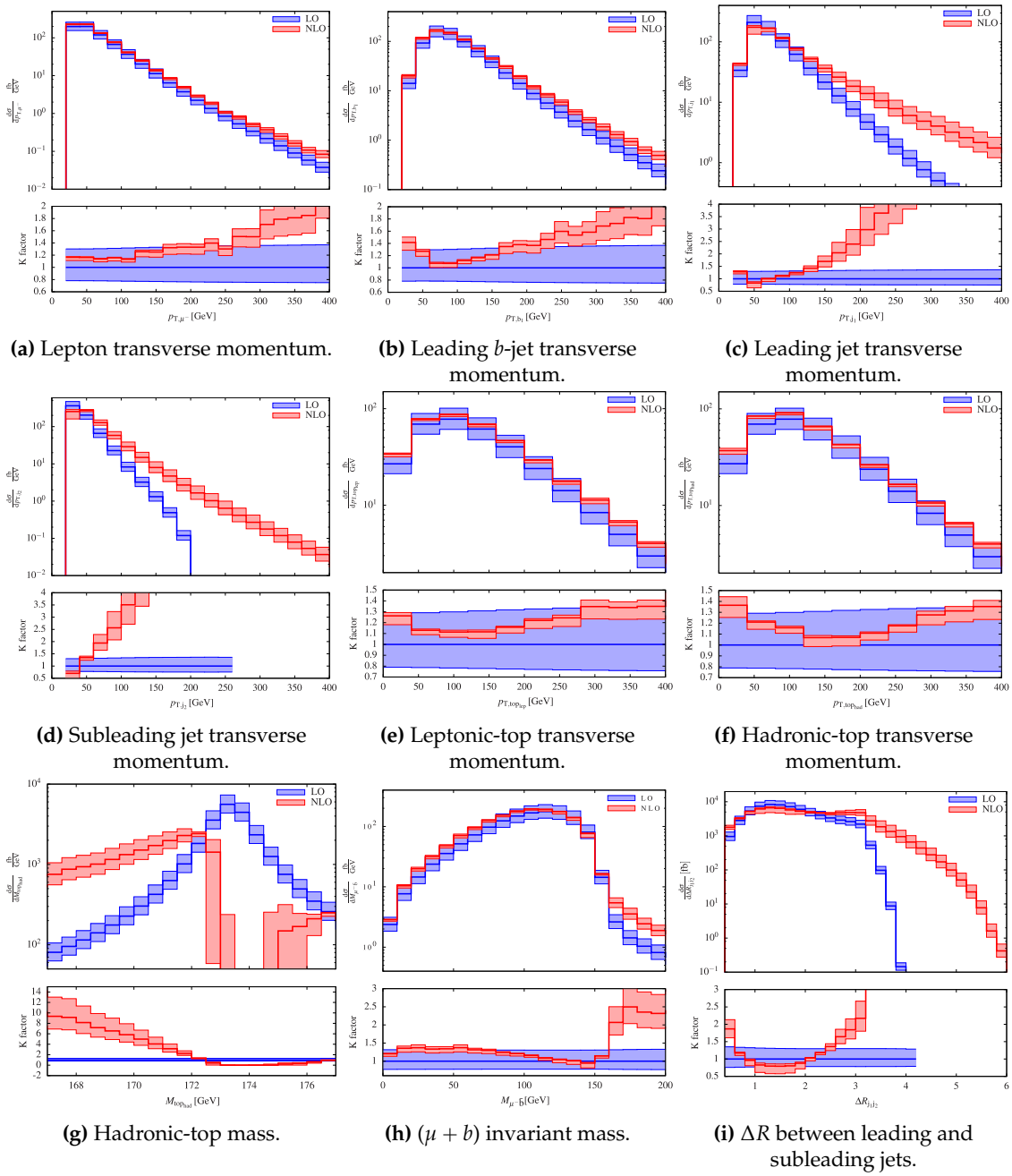


Figure 2.3: Predicted DS-to-DR Ratio as a Function of Dilepton Transverse Momentum. The upper pane plots the predicted DS/DR ratio as a function of the dilepton transverse momentum ($p_T^{(\ell\ell)}$). The different colored lines correspond to different transverse momentum veto cuts. The bottom pane shows the scale dependence obtained by varying a multiplicative factor on the top mass (m_t). These plots demonstrate the potential differences that could arise for the DR and DS approaches as $p_T^{(\ell\ell)}$ increases. Analyses targeting a single-top process as the signal must address this complication. In these analyses, it is necessary to evaluate the differences between the two methods. Plot is taken from Ref. [12].

Evaluating the DR and DS approaches is necessary to estimate the uncertainties associated with this quantum interference. The modeling of this interference can be a significant component of the total uncertainty, especially for BSM physics [13–17]. In the search for top squarks [13], the modeling uncertainty from the interference between $t\bar{t}$ and tW is estimated using inclusive $WWbb$ events. The theoretical uncertainties quoted in that paper were 19-26% for $t\bar{t}$ events and 38-57% for tW events. This uncertainty is a significant component of the total uncertainty, which would only increase after accounting for other background modeling uncertainties. A direct measurement of $WWbb$ would be sensitive to this interference effect. By treating both $t\bar{t}$ and tW contributions as signal, the results could be used to improve future MC generators. A comprehensive generator that accurately describes both $t\bar{t}$ and tW production, including quantum interference effects at higher-orders, would be of great use

to the larger top physics community. In that regard, there exists a $pp \to \ell^+ \nu_l \ell^- \overline{\nu}_\ell b\bar{b}$ ($bb4l$) generator [18] which by itself is a self-contained $t\bar{t}$ and tW generator. Likewise, predictions of the semileptonic version of the $bb4l$ also exists [19], and efforts are ongoing to implement this into the existing dileptonic $bb4l$ framework. Discrepancies between data and $bb4l$ predictions would be of most interest to the original authors of the $bb4l$ generator.

The authors of Ref. [19] have presented predictions at NLO for various $bb4l$ observables. After including NLO corrections, several observables show significant deviations from the corresponding LO predictions. These include transverse momentum, invariant mass, angular, and rapidity distributions. As shown in Figure 2.4, these discrepancies arise in the high-energy regions and the tails of the selected distributions. These high-energy regions can be accurately probed with the increased dataset and sufficiently high energies of the Run 2 data-taking period. This analysis measures the distributions in data and compares them to existing MC predictions.



(a) Lepton transverse momentum.

(b) Leading b-jet transverse momentum.

(c) Leading jet transverse momentum.

(d) Subleading jet transverse momentum.

(e) Leptonic-top transverse momentum.

(f) Hadronic-top transverse momentum.

(g) Hadronic-top mass.

(h) $(\mu + b)$ invariant mass.

(i) ΔR between leading and subleading jets.

Figure 2.4: Predicted *bb4l* Differential Cross-Section Distributions at LO and NLO. Selected differential cross-section predictions are shown at LO and NLO corresponding to a center-of-mass energy of $\sqrt{s} = 13$ TeV. The upper panel plots the LO and NLO predictions, while the lower panel plots the ratio of the NLO and LO predictions. Significant LO-NLO discrepancies are observed in the tails of the selected distributions. The following observables are plotted: **(a)** lepton transverse momentum, **(b)** leading b-jet transverse momentum, **(c)** leading jet transverse momentum, **(d)** subleading jet transverse momentum, **(e)** leptonic-top transverse momentum, **(f)** hadronic-top transverse momentum, **(g)** hadronic-top mass, **(h)** $(\mu + b)$ invariant mass, and **(i)** angular separation (ΔR) between leading and subleading jets. Plots are taken from Ref. [19].

CHAPTER 3

Theory

This chapter lays the theoretical foundation for the measurements conducted in this analysis, beginning with a description of the SM. The discussion will cover the properties of the top quarks in the SM, including various production and decay modes. Next, we will describe the various processes contributing to the $WWbb$ final state. Since the final measurements will be compared to various SM predictions, an overview of the methods used to derive these predictions will also be provided. Many of these calculations have referenced the interference effects between $t\bar{t}$ and tW. Here, this chapter extends the discussion of this interference effect that was introduced previously and elaborates on the theoretical source of this effect. Thus, there will also be a description of the DR and DS methods for decoupling this interference.

3.1 The Standard Model

The Standard Model (SM) of particle physics is the theoretical framework that serves to classify all known elementary particles and three out of four fundamental forces. Figure 3.1 shows the SM organized into two groups: fermions and bosons. The fermions are the basic building blocks of matter and have a half-integer spin. The bosons, representing the forces, have integer spin values and are exchanged during particle interactions. Gluons (g) mediate the strong nuclear force, the $W^{\pm}/Z$ mediates the weak nuclear force, and the photon (γ) mediates the Electro-Magnetic (EM) force. Finally, the Higgs (H) boson couples to the other SM particles through which the SM particles acquire their mass.

3.1.1 Fermions and Bosons

The SM further categorizes fermions into quarks and leptons, organized into three groups of two, known as generations. Quarks have a fractional electrical charge and are relatively massive particles in the SM. The top quark, a focus of this analysis is the most massive of all elementary particles, with an approximate mass of 173 GeV. In addition to electric charge, quarks and gluons also contain a color charge, which can only come in three types: red, green, and blue. During particle interactions, quarks and gluons actively exchange color charge. Since gluons also carry a color charge, they can

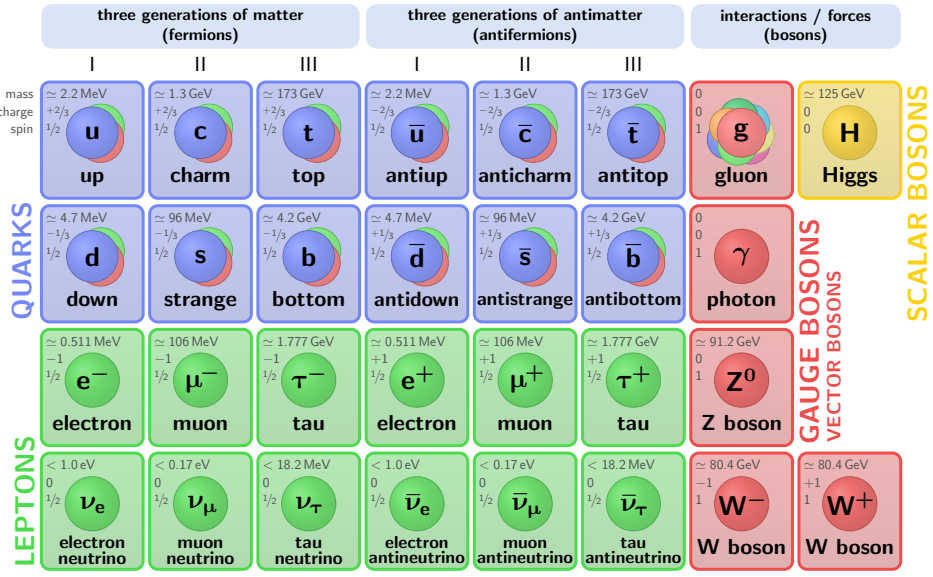


Figure 3.1: Standard Model of Elementary Particles. The Standard Model (SM) of particle physics classifies all known elementary particles into fermions and bosons (and their respective antiparticles). Fermions are comprised of the quarks and leptons and constitutes the vast majority of the visible matter in the universe. Bosons represent the three major forces included in the theory of the SM. The Higgs boson arises from the Higgs mechanism which is a component of Electroweak theory. Particles interact with the Higgs field to acquire their masses. The plot lists, for each particle, the charge, spin and estimated mass. Plot is taken from Ref. [20].

self-interact. Therefore, pairs of gluons can initiate interactions and pair-produce a quark-antiquark pair. In the context of this analysis, this gluon-fusion process is the primary production mode for top quark pairs.

A color charge also results in an effect known as color-confinement. This effect ensures that quarks and gluons cannot exist as free particles. Instead, quarks must "hadronize" with other quarks and gluons to form (color-neutral) composite particles. Hadrons composed of two quarks are known as "mesons" and three quarks are known as "baryons" and exist in a color-neutral state. Protons and neutrons fall under the category of baryons, consisting of a bound state of (u, u, d) and (u, d, d), respectively. As a result, quarks created by the high energy proton collisions at the LHC quickly hadronize and form structures known as jets, which are subsequently captured and reconstructed by the ATLAS detector.

Three generations of leptons comprise the other subgroup of fermions. Leptons have an integer electrical charge and are relatively light compared to quarks. They interact via the electromagnetic force through the exchange of a photon or via the weak nuclear force through the exchange of a $W^{\pm}/Z$. Unlike quarks, leptons do not carry color charges and can exist as free particles in nature. Particle interactions that result in a purely-leptonic final state are said to be "clean", avoiding much of the messiness associated with jets.

3.1.2 *Mathematical Definition of the Standard Model*

The SM is formulated in the language of Quantum Field Theory (QFT), which treats the fundamental particles as fields in all points of spacetime rather than discrete point-like particles. Mathematically, the SM is summed up by Eq. 3.1 as a gauge-invariant quantum field theory [21].

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &+ i\bar{\psi}\not{D}\psi + \text{h.c.} \\ &+ \bar{\psi}_i y_{ij}\psi_j\phi + \text{h.c.} \\ &+ |D_\mu\phi|^2 - V(\phi)\end{aligned} \tag{3.1}$$

In broad terms, each line of the above equation translates to the current understanding of the elementary particles and their interactions with the fundamental forces. The first line describes the action of forces (electromagnetism, weak and strong nuclear forces) neatly captured in the electromagnetic tensor, $F_{\mu\nu}$. The second line describes the fermions (using the Dirac spinor, ψ representation) under the actions of the fundamental forces. The third line corresponds to the "Yukawa Interaction" [22] and refers to the interaction of Dirac particles with a scalar field (ϕ). In the SM, this scalar field is the scalar Higgs field, and the couplings of fermions to this field result in the quarks and leptons acquiring their mass. The final line of that equation enables self-interactions of the Higgs field ($V(\phi)$), whereby an effect known as "spontaneous symmetry breaking" results in the vector bosons acquiring their mass.

The lagrangian is both gauge-invariant and Lorentz-invariant. A gauge, as it relates to mathematics and physics, refers to choosing a mathematical representation of a theory. Often, mathematicians/physicists choose a specific gauge for the convenience of simplifying the mathematics and making it easier to understand and apply. Gauge invariance is the idea that the underlying physics of a theory should not depend on a specific choice of the mathematical representation of that theory. The SM is both gauge-invariant and Lorentz-invariant, so it remains consistent regardless of its mathematical representation or choice of reference frame. In other words, two observers in different reference frames measuring the same physical properties should get the same answer regardless of their choice of gauges.

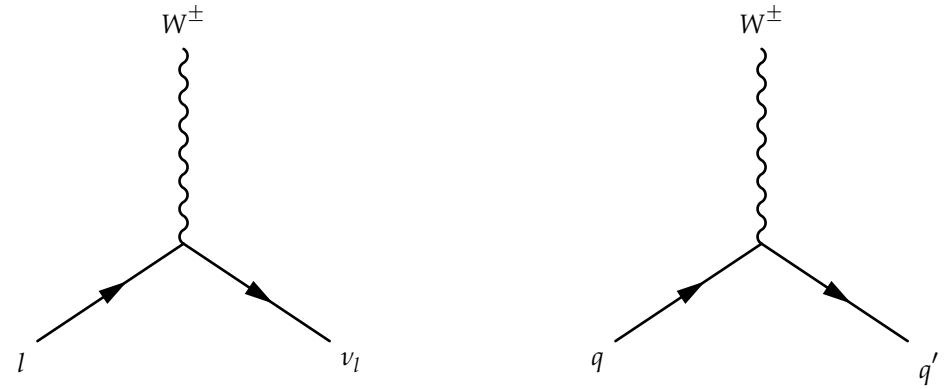
The symmetries of gauge invariance within the SM are organized using group theory. A group is a mathematical structure where all elements satisfy several properties: closure, associativity, identity, and inverse. The lagrangian of the SM is expressed in a gauge identified by the $SU(3) \otimes SU(2) \otimes U(1)$ groups [23]. The symmetries prescribed by these three groups roughly correspond to the three fundamental interactions described by the SM.

3.1.3 *Weak Interactions*

As the name implies, the weak nuclear force is the weakest of the three fundamental forces included in the SM. The weak interaction is responsible for radioactive beta decay, which converts a neutron

into a proton ($n \rightarrow p + e^- + \nu_e$). A neutrino was included in the theory of beta decay by Wolfgang Pauli to conserve momentum and energy. The SM incorporates the weak interaction as a $SU(2) \otimes SU(1)$ symmetry group component.

The massive vector bosons, W^+, W^-, and Z, mediate the weak interaction. The two fundamental weak interaction processes include the charged and neutral current interactions. The charged current interaction is so named because the process involves the exchange of electrical charge through the exchange of the $W^\pm$ bosons. Radioactive beta decay proceeds through a charged current interaction as one of the d quarks in the neutron converts to a u quark by emitting a W^- boson. Another charged current interaction involves a charged lepton converting to a neutrino of the same lepton flavor by emitting a W. Figure 3.2 illustrates the fundamental vertices of the charged current interaction.



(a) Leptonic charged current interaction. **(b)** Flavor changing charged current interaction.

Figure 3.2: Fundamental Charged Current Interaction Vertices. The charged current weak interaction proceeds through one of these two fundamental vertices. The emittance or absorption of a $W^\pm$ is responsible for converting **(a)** leptons to neutrinos, and **(b)** converting quarks to another flavor.

The semileptonic final state of this analysis contains both of these fundamental vertices. The two W-bosons undergo weak decay through the charged current interaction. In this process, one W decays into a charged lepton and its corresponding neutrino, while the other W decays hadronically into a quark and antiquark.

The other type of weak interaction is the neutral current interaction mediated by the exchange of a Z-boson. Unlike the $W^\pm$, interactions of the Z do not involve any transfer of electric charge, hence the name neutral current. In the neutral current interaction, a fermion such as a neutrino can emit or absorb a Z-boson and remain in the same flavor state while only experiencing changes in energy. Neutral current processes are important in neutrino scattering experiments where an incoming neutrino scatters off the electrons in a medium [24]. Figure 3.3 shows the fundamental vertex for the neutral current process.

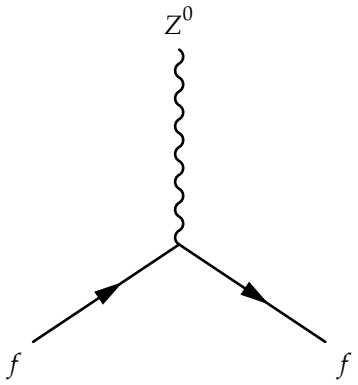


Figure 3.3: Fundamental Neutral Current Interaction Vertex. The neutral current weak interaction proceeds via the exchange of a Z-boson without the exchange of electrical charge.

The neutral current process does not contribute to the signal in this analysis. However, the production of Z-boson alongside several quarks (which quickly hadronize into jets) is one of the background processes considered (Z + jets).

3.1.4 Quantum Electro-Dynamics

Quantum Electro-Dynamics (QED) is the quantum field theory describing interactions between light and matter. Electrically charged particles interact electromagnetically through the exchange of photons. Unlike their $W^{\pm}$ and Z counterparts, photons are massless as a direct consequence of the QED Lagrangian lacking any mass terms coupled to the vector fields. A m^2-like term would negate the gauge-invariance of the existing Lagrangian. The Lagrangian density describing QED interactions is shown in Eq. 3.2 and is represented in the SM as a $U(1)$ symmetry.

$$\mathcal{L}_{QED} = \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \overline{\psi}(i\not{D} - m)\psi \tag{3.2}$$

The above equation is written in terms of the aforementioned electromagnetic tensor, $F_{\mu\nu}$, a Dirac spinor, and its corresponding adjoint, ψ ($\overline{\psi}$). The Dirac spinors are connected by the gamma matrices, γ^{μ}, which Dirac introduced in his relativistic wave equation to describe the interactions of fermions. Finally, D_{μ} is known as the gauge covariant derivative, which when expanded, yields $D_{\mu} \equiv \partial_{\mu} - ieA_{\mu}$. This gauge-invariant derivative introduces A_{μ} as the potential produced by the electron itself. When solved, the equations of motion of the QED Lagrangian yields the following two equations.

$$\Box A^{\mu} = ej^{\mu}, \; \Box \equiv \partial_{\mu}\partial^{\mu} \tag{3.3a}$$

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = e\gamma^{\mu}A_{\mu}\psi \tag{3.3b}$$

Eq. 3.3a represents the QED version of the classical Maxwell's Equations, a set of equations that form the basis of classical electromagnetism. Eq. 3.3b is the Dirac equation including a non-zero potential

term. The Dirac spinor (representing fermions) couples to the gauge electromagnetic potential in this formulation. This coupling of the Dirac fermions to a massless field allows the charged fermions to interact via the exchange of a photon. Figure 3.4 shows the resulting fundamental QED vertex.

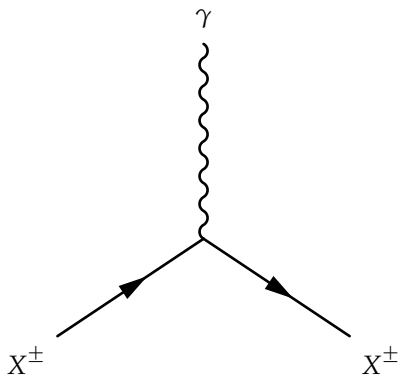


Figure 3.4: Fundamental QED Interaction Vertex. The QED interaction is the quantum field theory version of classical electromagnetism. Two charged particles (shown by $X^{\pm}$) interact by exchanging a photon.

Physicists unified the electromagnetic and weak interactions into a comprehensive description known as the Electroweak theory [25, 26]. The SM represents Electroweak theory as the $SU(2) \otimes U(1)$ symmetry group. The combination of the two gauge groups yields a set of three (massless) W-bosons ($W_{1,2,3}$) and a fourth B-vector boson that is also massless. The addition of these bosons maintains the symmetries of the electroweak theory. The symmetries of the electroweak theory can be spontaneously broken, leading to the generation of the masses of the vector bosons via the Higgs mechanism [27]. It begins with the Higgs potential as a function of the Higgs field (ϕ), as shown in Eq. 3.4.

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{3.4}$$

Interactions of the initially massless electroweak W and B bosons with this field result in the vector boson fields obtaining the masses of the SM $W^{\pm}$, Z, and γ particles.

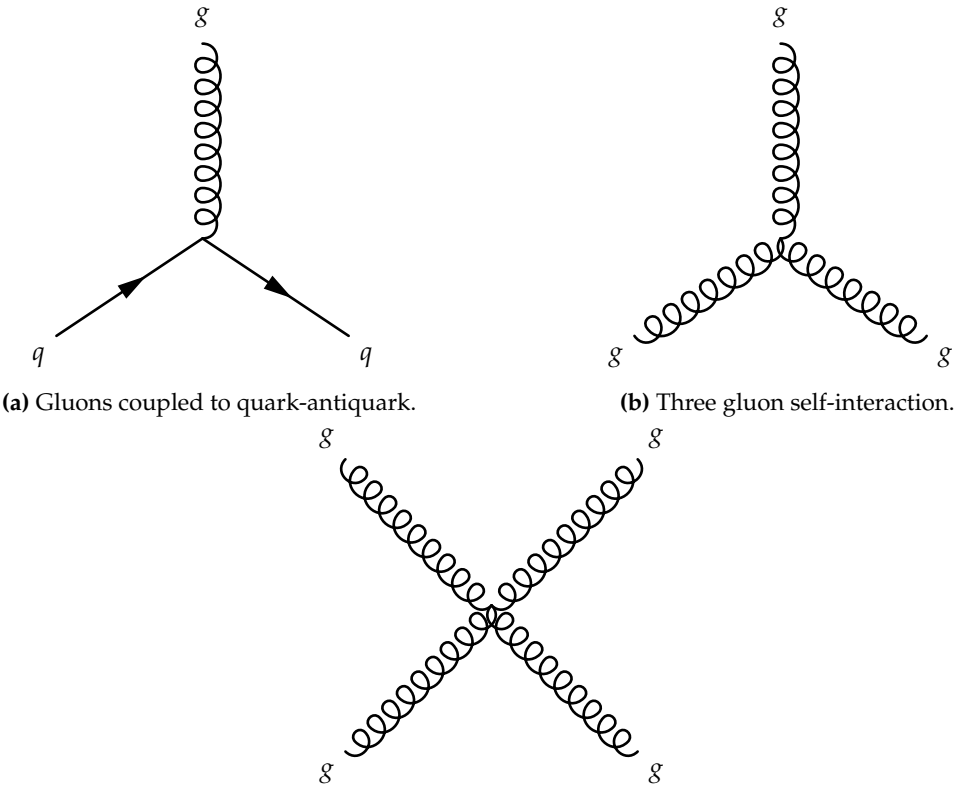
3.1.5 Quantum Chromo-Dynamics

Quantum Chromo-Dynamics (QCD) is the quantum field theory describing the strong interactions between quarks and gluons. The strong force is responsible for maintaining the structural integrity of the nucleus in atoms. Thus, the strong force has the shortest range of all of the fundamental forces at approximately the size of the nucleus of an atom. Gluons are the massless force carriers of the strong interaction and are exchanged during quark-quark (and gluon-gluon) interactions. The quarks also carry an additional quantum property known as color, a property of quarks and gluons that can take on values of red, blue, green, and their anticolor counterparts. Color-confinement is a feature that prevents quarks from existing as free particles. The QCD lagrangian, shown in Eq. 3.5, is represented in the SM by the $SU(3)$ symmetry group.

$$\mathcal{L}_{QCD} = \overline{\psi}_i(i\gamma^\mu(D_\mu)_{ij} - m\delta_{ij})\psi_j - \frac{1}{4}G^\alpha_{\mu\nu}G^{\mu\nu}_\alpha \tag{3.5}$$

The QCD Lagrangian shares some similarities with the previous QED Lagrangian. The quark field spinor ψ_i ($\overline{\psi}_i$) replaces the Dirac spinor, which describes fermions in QED. Note, the subscripted i takes on the values of 1, 2, and 3 roughly corresponding to the three possible colors. The same gamma matrices, γ^μ also appear here. The covariant derivative, $(D_\mu)_{ij} = \partial_\mu\delta_{ij} - ig(T_\alpha)_{ij}A^\alpha_\mu$, has a similar form to the QED covariant derivative. The differences begin with the g-constant which measures the coupling strength between quarks and gluons. QCD also introduces the $(T_\alpha)_{ij}$ matrices, which are 3×3 matrices that act as the generators of $SU(3)$. In group theory, a generator denotes the smallest set of group elements that generate all other elements through linear combinations of these generators. The set of eight $(T_\alpha)_{ij}$ matrices are responsible for producing the eight gluon variations exchanged during strong interactions. In the QCD covariant derivative, the gluon potential, A^α_μ characterizes these eight gluon variations. Finally, the gluon field strength tensor $G^\alpha_{\mu\nu}$ is the QCD analog of the electromagnetic tensor.

Gluons must contain a color and an anticolor (or some superposition of color and anticolor states). However, one expects that there should be $3 \times 3 = 9$ total gluons instead of the eight described by theory. It is possible to formulate a list of nine unique gluon colorless states using a linear combination of three colors and three anticolors. However, the so-called color singlet state (defined by a linear combination of $\sim (r\overline{r} + b\overline{b} + g\overline{g})$ is omitted since such a state does not permit interactions with other color states. During the strong interaction, interacting particles exchange color charges so that the initial and final states remain color-neutral. This condition leads to the color confinement property inherent to QCD. Figure 3.5 shows the fundamental vertices for QCD.



(a) Gluons coupled to quark-antiquark.

(b) Three gluon self-interaction.

(c) Four gluon self-interaction.

Figure 3.5: Fundamental QCD Interaction Vertices. The QCD interaction is mediated by gluons and the exchange of color charges. The emittance or absorption of a g converts **(a)** quarks into their antiquark counterparts. Gluons carry color, allowing for **(b)** three-gluon and **(c)** four-gluon self-interactions.

The gluon-gluon fundamental vertex dominates top quark pair production at higher energies, while the quark-antiquark annihilation process contributes more at lower energies.

3.2 Top Quark Properties

The top quark has a mass of approximately 173 GeV, making it the heaviest of all elementary particles. This large mass leads to several interesting properties unique to the top, the measurements of which serve as tests of the SM. ATLAS and CMS have independently measured the single and double top quark production cross-sections at the LHC and observed good agreement with predicted cross-sections. Figure 3.6 shows the inclusive cross-section from $t\bar{t}$ production, while Figure 3.7 shows the inclusive cross-section from single-top production. A consistent physics effort has been put into studies of the top at both ATLAS and CMS throughout the last decade.

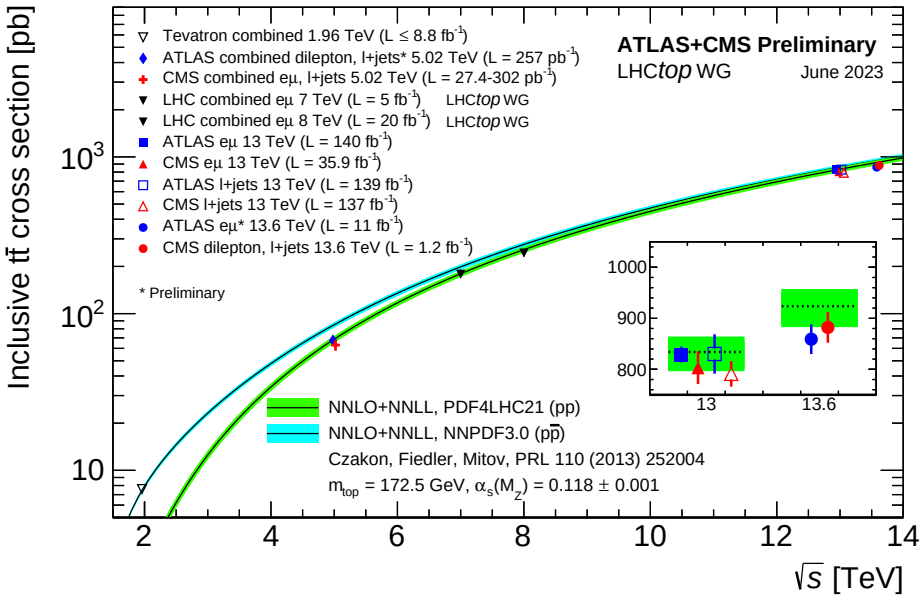


Figure 3.6: Measured $t\bar{t}$ Inclusive Cross-Sections at ATLAS and CMS. The inclusive $t\bar{t}$ cross-sections have been measured at ATLAS and CMS in different decay channels and center-of-mass energies. Measurements are compared to NNLO predictions. The error bars on the data point represent the total uncertainty on the cross-section for each measurement. The colored bands represent the uncertainties associated with various MC parameters. Plot is taken from Ref. [3].

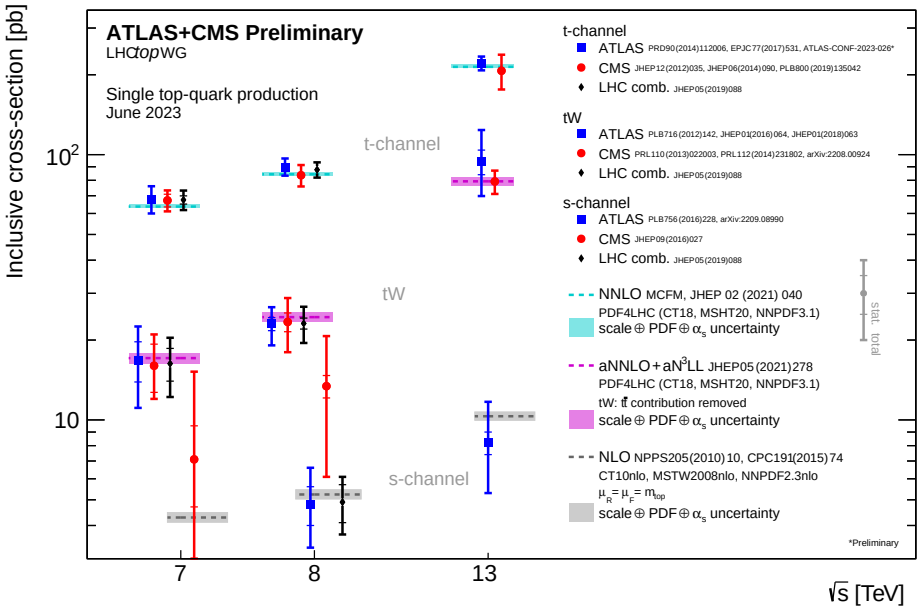


Figure 3.7: Measured Single-Top Inclusive Cross-Sections at ATLAS and CMS. The inclusive single-top cross-sections have been measured at ATLAS and CMS in various decay channels and center-of-mass energies. Measurements are compared to NLO and NNLO predictions. For each measurement, the inner error bars on the data point represent the statistical uncertainty. The outer error bars represent the total uncertainty after summing all remaining uncertainties. The colored bands represent the uncertainties associated with various MC parameters. Plot is taken from Ref. [3].

The top mass is an important parameter in the formulation of the SM. ATLAS and CMS have both conducted multiple measurements of this quantity. Figure 3.8 shows a summary of top mass measurements made by ATLAS and CMS in various decay channels. While a direct measurement of the top mass is not conducted in this analysis, we will present differential cross-sections of various top quantities (including the mass). The semileptonic final state of the $WWbb$ process allows for the full reconstruction of the top quark. A focus on a semi-boosted final state probes kinematic regions where the theoretical calculations based on the SM still contain large uncertainties. This is important for analyses with large backgrounds from top processes, where modelling uncertainties can be a significant source of systematic uncertainty.

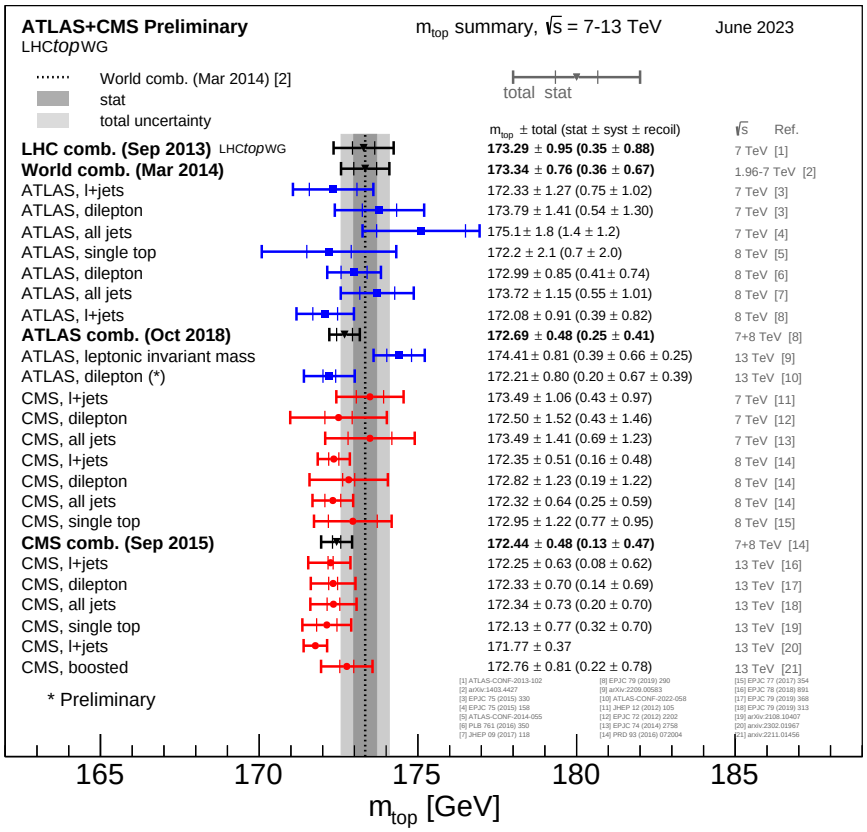


Figure 3.8: Top Mass Measurements from Direct top quark Decay Processes at ATLAS and CMS. The top quark mass has been measured at ATLAS and CMS in various top quark decay channels and center-of-mass energies. Individual measurements are plotted relative to the current combined average of $m_t = 173.34 \pm 0.76$ GeV. For each measurement, the inner error bars on the data point represent the statistical uncertainty. The outer error bars represent the total uncertainty after summing all remaining uncertainties. Plot is taken from Ref. [28].

A unique property of the top is its short lifetime, due to its large mass. The lifetime of the top is shorter than the time for hadronization, meaning the top quark does not form any quarkonia ($q\bar{q}$) states, unlike the other five flavors of quarks. This characteristic is unique to the top quark, allowing analysts to study a quark in its natural state and not subject to the property of color confinement. Upon decaying, the top quark will decay almost 100% of the time to a W-boson and a b-quark via the weak force. Decays to other light quarks are also possible but are highly suppressed. This level of suppression is quantified by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, whose elements quantify the probability of flavor-changing weak interactions. The nine elements of the CKM matrix are another important measurable property of the top. The current accepted CKM matrix values are shown below [3].

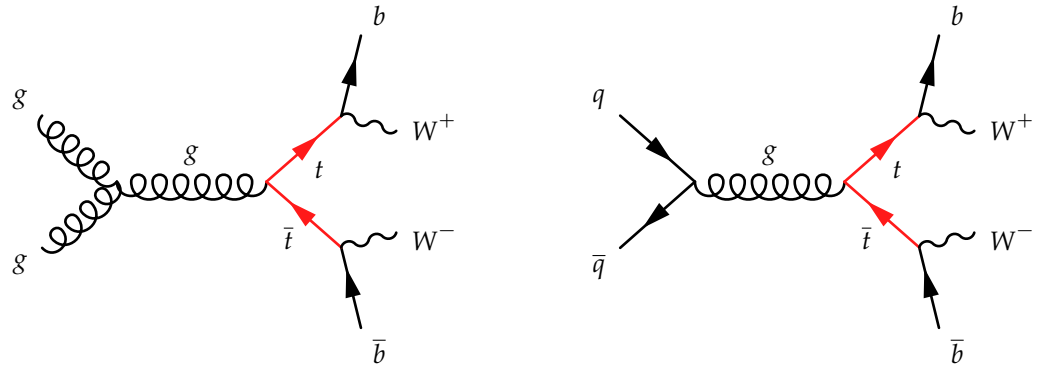
$$|V_{\text{CKM}}| = \begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} \tag{3.6a}$$

$$|V_{\text{CKM}}| = \begin{bmatrix} 0.97435 \pm 0.00016 & 0.22500 \pm 0.00067 & 0.00369 \pm 0.00011 \\ 0.22486 \pm 0.00067 & 0.97349 \pm 0.00016 & 0.04182^{+0.00085}_{-0.00074} \\ 0.00857^{+0.00020}_{-0.00018} & 0.04110^{+0.00083}_{-0.00072} & 0.999118^{+0.000031}_{-0.000036} \end{bmatrix} \tag{3.6b}$$

Here, it is evident that the $|V_{tb}| \gg |V_{ts}|, |V_{td}|$ explaining why the top quark will almost always convert to a b-quark alongside the release of a W.

3.3 $pp \rightarrow WbWb$ Processes in the SM

The final state of interest contains two W-bosons and two b-quarks. Several processes could contribute to such a final state, categorized by the number of top quark resonances. Top quarks are mainly produced in $t\bar{t}$ pairs; at the LHC 90% of the $t\bar{t}$ processes occur via gluon-gluon fusion ($gg \rightarrow t\bar{t}$) [3]. Another production mechanism, although less favored, is the production of $t\bar{t}$ through quark-antiquark annihilation ($q\bar{q} \rightarrow t\bar{t}$). These processes proceed through the strong interaction and form the dominant process of this analysis. A representative Leading-Order (LO) diagram of these processes is shown in Figure 3.9.

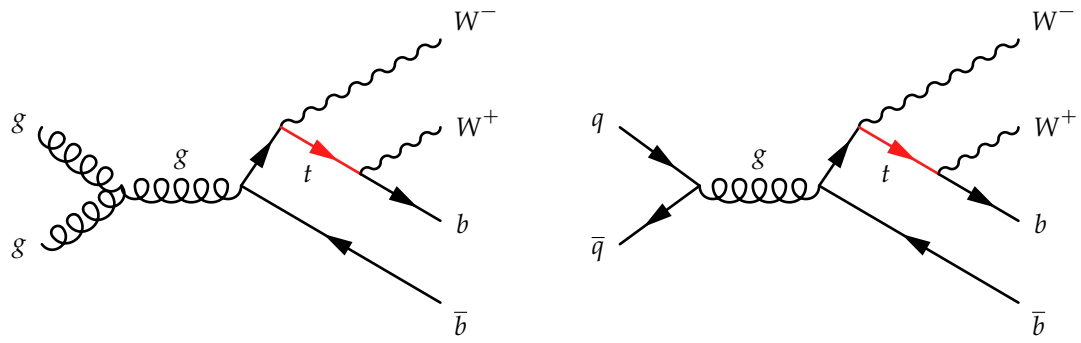


(a) $t\bar{t}$ production via gluon-gluon fusion.

(b) $t\bar{t}$ production via quark-antiquark annihilation.

Figure 3.9: Representative $t\bar{t}$ Production Modes at the LHC. $t\bar{t}$ production at the LHC is dominated by **(a)** gluon-gluon fusion, while **(b)** quark-antiquark annihilation is another mechanism by which $t\bar{t}$ is produced but at a much lower rate.

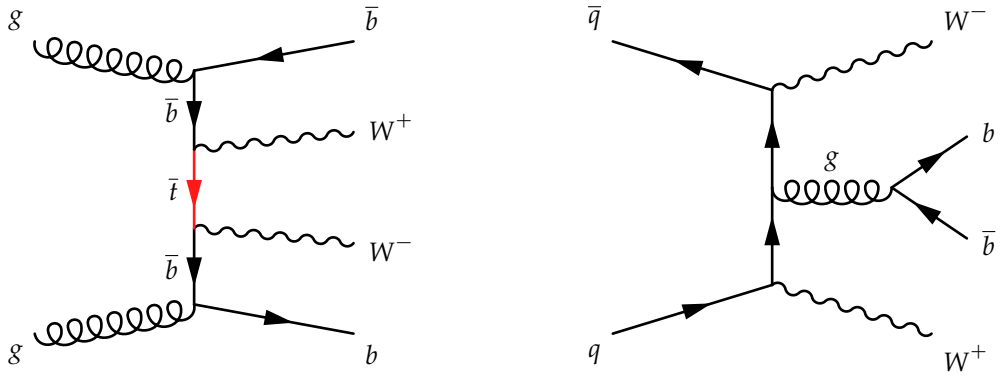
The above diagrams illustrate doubly-resonant top quark pair production. This process produces both top quarks on-shell with a mass consistent with the "true" mass of the top. Similarly, singly-resonant processes produce a single on-shell top quark. These processes are known as tW processes and proceed via the electroweak interaction. Figure 3.10 presents sample diagrams for the singly-resonant top quark process.



(a) tW production via gluon-gluon fusion. **(b)** tW production via quark-antiquark annihilation.

Figure 3.10: Representative tW Production Modes at the LHC. Singly-resonant processes can also contribute to the same $WWbb$ final state, using the same **(a)** gluon-gluon fusion process, or **(b)** quark-antiquark annihilation. These processes produce a single on-shell intermediate top quark, but occur less frequently than their doubly-resonant counterparts.

Other rarer processes can create the $WWbb$ final state containing zero on-shell tops. These processes fall under non-resonant processes and may involve a virtual top quark (off-shell). Some rarer SM processes can produce a $WWbb$ final state without involving a top quark. These processes can include more exotic mechanisms such as the example processes shown in Figure 3.11.



(a) $WWbb$ production involving an off-shell antitop. **(b)** $WWbb$ production involving gluon-splitting.

Figure 3.11: Representative Non-Resonant Production Modes at the LHC. Non-resonant processes include zero on-shell top quarks, implying virtual particles or no top quarks whatsoever. Two examples are shown where the $WWbb$ final state is achieved through the production of **(a)** an intermediate off-shell antitop, or via **(b)** gluon-splitting. Nonetheless, these processes must be considered to produce an inclusive $WWbb$ cross-section measurement.

These diagrams are shown at tree-level, without any high-order (self-interactions) loops included. Theorists have computed higher-ordered predictions and included these in the present set of MC event generators. A fundamental goal of this analysis is to compare the measurements of the observables of interest with the predictions of $WWbb$ at these higher-orders. Therefore, it is necessary to briefly

describe the process by which various top properties are computed.

3.4 Methods for Computing Top Quark Predictions

Predictions for $t\bar{t}$ production have been computed at the Next-to-Leading Order (NLO) [29–31]. These references proceed using the Narrow Width Approximation (NWA). The NWA enforces the assumption of an on-shell top and that the mass of the top is much greater than the decay width, $\Gamma_t/m_t \to 0$. This limit greatly reduces required computations by removing diagrams which contribute at higher orders of Γ_t/m_t. This is an important feature for future MC generators looking to expand to higher-orders. For context, to compute the NLO QCD corrections to $pp \to WWbb$ requires the inclusion of 300 $q\bar{q} \to WWbb$ one-loop diagrams and 800 $gg \to WWbb$ one-loop diagrams [32]. This number of diagrams grows exponentially with the order of the calculations, requiring the input of thousands of two-loop diagrams.

The NWA approximation begins by first writing down the amplitude for the processes of interest. To compute the cross-section, the amplitude is squared to obtain the probability for the process. The squared amplitude is then integrated over the phase space to get the final fiducial cross-section. To demonstrate the usefulness of the NWA, we start with the amplitude of the tree-level process of top quark pair production in the fully-leptonic final state ($ij \to t\bar{t} \to (\bar{b}\ell^-\bar{\nu})(\bar{b}\ell^+\nu)$) shown in Eq. 3.7 [33].

$$\begin{aligned} A^{\text{tree}} = &\left(\tilde{A}(t \to b\ell^+\nu) \frac{i(\not{p}_t + m_t)}{p_t^2 - m_t^2 + i m_t \Gamma_t} \right) \\ &\times \tilde{A}(ij \to t\bar{t}) \\ &\times \left(\frac{i(-\not{p}_{\bar{t}} + m_t)}{p_{\bar{t}}^2 - m_t^2 + i m_t \Gamma_t} \tilde{A}(\bar{t} \to \bar{b}\ell^-\bar{\nu}) \right) \end{aligned} \tag{3.7}$$

The amplitude does not make any assumptions about having the top quark on-shell. This requirement is enforced after squaring the amplitude and applying the NWA. As written, the amplitude is expressed in terms of several sub-amplitudes, $\tilde{A}$. These correspond to the production of the intermediate $t\bar{t}$ process and subsequent decay to the dilepton final state. The remaining terms in the amplitude represent propagators[1]. After squaring the amplitude and applying the NWA, the propagators have the form shown in Eq. 3.8.

$$\left. \frac{1}{(p_t^2 - m_t^2)^2 + m_t^2 \Gamma_t^2} \right|_{\Gamma_t/m_t \to 0} = \frac{2\pi}{2 m_t \Gamma_t} \delta(p_t^2 - m_t^2) \tag{3.8}$$

The propagator is defined in terms of the top's transverse momentum, mass, and decay width (p_{T}, m_t and Γ_t). Adding the Dirac delta function after applying the NWA enforces the on-shell top mass. The

[1] In QFT, the propagator is a mathematical expression describing a particle's propagation from one state in spacetime to another. Feynman diagrams represent these as internal lines connecting pairs of vertices. As a result, propagators often enter into cross-section calculations. The denominators of propagators often carry a $i\epsilon$-like term. This term is introduced to maintain causality and ensure that any virtual particles represented by the propagator remain in the light-cone (i.e., virtual particles are still subject to the speed of light limit).

amplitude can then be simplified into a form that contains a term with zero-orders of Γ_t/m_t and higher orders of Γ_t/m_t, which do not contribute to the approximation, as shown in Eq. 3.9.

$$A^{\text{tree}} = \overline{U}(p_t)\tilde{A}(ij \to t\bar{t})V(p_{\bar{t}}) + \mathcal{O}(\frac{\Gamma_t}{m_t}) \tag{3.9}$$

Here, $\overline{U}(p_t)$, $V(p_{\bar{t}})$ correspond to the simplified on-shell Dirac spinors for top and anti-top, respectively. The advantage of such spinors is that the calculation can proceed without further prescriptions to differentiate between the top and the anti-top. Furthermore, any terms carrying Γ_t/m_t factor will also reduce to zero, eliminating non-resonant diagrams and significantly reducing the computation required. After defining the amplitude, the calculation proceeds as any other computation, and any number of physical properties, such as the cross-section, mass, etc., can be calculated by taking the square of the amplitude.

Despite its usefulness in simplifying the problem, the NWA is not without its downfalls. Firstly, the application of the NWA comes with an associated accuracy parameterized by the ratio (Γ_t/m_t) and is approximately $\mathcal{O}(\Gamma_t/m_t) \approx 1\%$ [34]. This error is important for analyses targeting phase spaces sensitive to off-shell top mass effects. Thus, the NWA becomes less applicable in regions where the limit of $\Gamma_t \ll m_t$ does not apply. Secondly, the NWA cannot be applied to calculations focusing on an inclusive final state, requiring contributions from singly-resonant and non-resonant top processes. For example, the previously mentioned *bb4l* generator that features a calculation leading to a $pp \to b\bar{b}\ell^+\ell'^-\nu_\ell\bar{\nu}_\ell$ will not be able to apply the simplifying powers of the NWA as this would eliminate a large subset of diagrams that could contribute to this final state.

3.5 $t\bar{t}$ and tW Quantum Interference

Up to this point, the methods for computing predictions have only mentioned $t\bar{t}$. There is an added challenge in performing computations involving single-top production. A quantum-interference effect exists at higher-orders in QCD that results from the same $WWbb$ final state between $t\bar{t}$ and tW processes [12]. In the computation of tW production, the calculation must consider all possible processes resulting in the following final state.

$$t + W + \sum_i X_i \tag{3.10}$$

The $\{X_i\}$ set corresponds to a set of particles that can be produced alongside the base tW final state. At LO in the SM, the production of tW is represented as follows.

$$bg \to tW \tag{3.11}$$

At NLO, the set becomes $\{X_i\} \equiv \bar{b}$. Processes such as gluon-gluon fusion, which produces a top quark pair can also contribute. In this process, the produced antitop decays to a $W\bar{b}$ ($gg \to tW\bar{b}$). At leading-order, the two processes are well-separated as the cross-section for the doubly-resonant $t\bar{t}$ process is much larger than the singly-resonant tW process, $\sigma_{t\bar{t}} > \sigma_{tW}$. When one includes NLO corrections to

tW, the addition of doubly-resonant processes begins to have a sizeable contribution. It turns out that the NLO corrections can be larger than the LO terms in the region where the $W\bar{b}$ mass approaches the top quark resonance.

$$M^2_{W\bar{b}} \equiv (p_W + p^2_{\bar{b}}) \simeq m^2_t \tag{3.12}$$

At the resonance, a divergence can occur in the propagator. The divergence is avoided by including a regulatory term defined by the decay width of the top, Γ_t. Including a non-zero top width comes with an inherent challenge since this factor must be included in all electroweak correction orders to avoid this divergence [12]. This same approach has already been illustrated in top quark pair production propagators shown in Eq. 3.7.

There have been a few approaches to deal with this interference effect during the calculations of single-top processes. The first approach is to conclude that tW is not an independent process but a subprocess (alongside $t\bar{t}$) that leads to a final state of WWb or $WWbb$. This way, the interference effects are naturally avoided, as there is no need to decouple the two resonant top processes. The two are treated as contributions to the same final state. One of the challenges associated with this approach is that there are non-negligible corrections at NLO to the $t\bar{t}$ process, which would degrade the accuracy of the prediction if not included [35]. This approach also lacks the flexibility of offering an exclusive single-top generator, especially for analyses strictly focusing on tW as their signal. The more flexible approaches are the DR and DS methods [12].

Diagram Removal (DR) The DR approach removes all doubly-resonant diagrams contributing to the same final state as singly-resonant tW at NLO accuracy.

Diagram Subtraction (DS) The DS approach introduces a subtractive term that cancels out the contributions of local NLO doubly-resonant $t\bar{t}$.

To illustrate these two approaches for isolating tW, we begin by writing down the amplitude ($\mathcal{A}_{\alpha\beta}$) for all partonic processes of $tW\bar{b}$.

$$\alpha(p_1) + \beta(p_2) \to t(k_1) + W(k_2) + \delta(k) \tag{3.13}$$

The interacting particles with their corresponding momenta are labeled as $\alpha(p_1)$ and $\beta(p_2)$, while the $\bar{b}$ is represented by the $\delta(k)$ term. The resulting amplitude will then contribute to both tW and $t\bar{t}$.

$$\mathcal{A}_{\alpha\beta} = \mathcal{A}^{(tW)}_{\alpha\beta} + \mathcal{A}^{(t\bar{t})}_{\alpha\beta} \tag{3.14}$$

To calculate the probability, the next step is to square the amplitude.

$$\begin{aligned} |\mathcal{A}_{\alpha\beta}|^2 &= |\mathcal{A}^{(tW)}_{\alpha\beta}|^2 + 2\mathcal{R}\left\{\mathcal{A}^{(tW)}_{\alpha\eta}\mathcal{A}^{(t\bar{t})*}_{\alpha\beta}\right\} + |\mathcal{A}^{(t\bar{t})}_{\alpha\beta}|^2 \\ &\equiv \mathcal{S}_{\alpha\beta} + \mathcal{I}_{\alpha\beta} + \mathcal{D}_{\alpha\beta} \end{aligned} \tag{3.15}$$

In the above, $S_{\alpha\beta}$ represents the singly-resonant contributions, $\mathcal{I}_{\alpha\beta}$ quantifies the interference between the two processes, and $\mathcal{D}_{\alpha\beta}$ represents the doubly-resonant contributions. We define the NLO differential cross-section by factoring in this probability.

$$d\hat{\sigma}_{\alpha\beta} = \frac{1}{2s}\left(\hat{\mathcal{S}}_{\alpha\beta} + \mathcal{I}_{\alpha\beta} + \mathcal{D}_{\alpha\beta}\right) d\phi_3 \tag{3.16}$$

Here, the hat denotes that infrared singularities have already been subtracted. These are singularities caused by intermediate soft gluons, where "soft" refers to gluons carrying away small amounts of energy or momentum in an interaction. Soft gluons force some internal propagators on-shell, resulting in singularities that either need to be renormalized or, in this calculation, subtracted away [36]. The $d\phi_3$ term represents the three-body final state phase space.

We obtain the complete hadro-production differential cross-section from the above cross-section by summing over all possible α, β processes.

$$\begin{aligned} d\sigma &= d\sigma^{(2)} + \sum_{\alpha,\beta}\int dx_1 dx_2 \mathcal{L}_{\alpha\beta} d\hat{\sigma}_{\alpha\beta} \\ &= d\sigma^{(2)} + \sum_{\alpha,\beta}\int \frac{dx_1 dx_2}{2x_1 x_2 s}\mathcal{L}_{\alpha\beta}\left(\hat{\mathcal{S}}_{\alpha\beta} + \mathcal{I}_{\alpha\beta} + \mathcal{D}_{\alpha\beta}\right) d\phi_3 \end{aligned} \tag{3.17}$$

The quantity $d\sigma^{(2)}$ represents all other contributions to the process defined by the process defined in Eq. 3.13. s is the squared center of mass energy and $\mathcal{L}_{\alpha\beta}$ is the parton-level luminosity. The x_i represents the momentum fractions of the interacting particles, integrated over all possible momentum states of the scattering particles.

The next step is to match the NLO computations to the parton showers according to the specific parton showering packages such as Monte-Carlo at NLO (MC@NLO) [37]. In the process, MC counterterms[2] are subtracted from the above equation and absorbed into the $\hat{\mathcal{S}}_{\alpha\beta}$ term. This action is permissible because this term already has particular prescriptions for dealing with such singularities. The resulting differential cross-section is as follows.

$$d\sigma^{(DR)} = d\sigma^{(2)} + \sum_{\alpha,\beta}\int \frac{dx_1 dx_2}{2x_1 x_2 S}\mathcal{L}_{\alpha\beta}\hat{\mathcal{S}}_{\alpha\beta} d\phi_3 \tag{3.18}$$

Now, the cross-section does not involve any contributions from $\mathcal{I}_{\alpha\beta}$ or $\mathcal{D}_{\alpha\beta}$. Doubly-resonant contributions have been removed from the cross-section. The diagram subtraction method starts from Eq. 3.17 and appends an additional term to subtract away the doubly-resonant term.

[2]MC evolution can result spurious NLO terms leading to different types of divergences (e.g., collinear divergences). Counterterms are constructed and appended to the cross-section calculations to cancel out such divergences [38].

$$d\sigma^{(\mathrm{DS})} = d\sigma - d\sigma^{\mathrm{subt}} \tag{3.19a}$$

$$d\sigma^{\mathrm{subt}}_{\alpha\beta} = \frac{1}{2s}\tilde{\mathcal{D}}_{\alpha\beta}d\phi_3 \tag{3.19b}$$

$$d\sigma^{(\mathrm{DS})} = d\sigma^{(2)} + \sum_{\alpha,\beta}\int \frac{dx_1 dx_2}{2x_1x_2S}\mathcal{L}_{\alpha\beta}\left(\hat{\mathcal{S}}_{\alpha\beta} + \mathcal{I}_{\alpha\beta} + \mathcal{D}_{\alpha\beta} - \tilde{\mathcal{D}}_{\alpha\beta}\right)d\phi_3 \tag{3.19c}$$

Now, we can evaluate the difference between the DR and DS approaches.

$$d\sigma^{(\mathrm{DS})} - d\sigma^{(\mathrm{DR})} = d\sigma^{(2)} + \sum_{\alpha,\beta}\int \frac{dx_1 dx_2}{2x_1x_2S}\mathcal{L}_{\alpha\beta}\left(\mathcal{I}_{\alpha\beta} + \mathcal{D}_{\alpha\beta} - \tilde{\mathcal{D}}_{\alpha\beta}\right)d\phi_3 \tag{3.20}$$

As expected, the difference is characterized by the interference term, $\mathcal{I}_{\alpha\beta}$, and the difference between the subtraction term and the doubly-resonant contribution. The difficulty is in defining such a subtraction term that cancels out all doubly-resonant contributions. The first condition is to ensure the difference, $\mathcal{D}_{\alpha\beta} - \tilde{\mathcal{D}}_{\alpha\beta}$ is as close to zero as possible, and the DR and DS differ only in the size of the interference. It is impossible to force a difference of exactly zero because of the requirements of local gauge invariance. However, one can guess the ingredients required to form a subtraction term. It should contain such a term to account for doubly-resonant production, and the subsequent decay of the $\bar{t} \to W\bar{b}$.

$$\tilde{D}_{\alpha\beta} = |\mathcal{A}^{(0)}(\alpha\beta \to t\bar{t})|^2 \times BW(M_{W\bar{b}}) \times |\mathcal{A}^{(0)}(\bar{t} \to W\bar{b})|^2 \tag{3.21}$$

where $\mathcal{A}^{(0)}(\alpha\beta \to t\bar{t})$ corresponds to the amplitude of the LO top pair production, and $\mathcal{A}^{(0)}(\bar{t} \to W\bar{b})$ is the decay amplitude of the antitop to $W\bar{b}$. The $BW(M_{W\bar{b}})$ term is the Breit-Wigner probability density function that includes the familiar Γ term to prevent divergences in mass resonances.

A measurement of the size of this interference has been previously conducted [39]. The referenced analysis combines $t\bar{t}$ + $tW\bar{b}$ events, resulting in a final state with two leptons and two b-jets. The observable of interest is a minimax quantity formed using the two leptons and the two b-jets.

$$m_{b\ell}^{\mathrm{minimax}} \equiv \min\{\max(m_{b_1\ell_1}, m_{b_2\ell_2}), \max(m_{b_1\ell_2}, m_{b_2\ell_1})\} \tag{3.22}$$

This quantity was inspired by previous analyses measuring the masses of supersymmetric particles [40, 41]. The minimax observable is particularly useful when there is a combinatorial ambiguity in the pairings of particles. In the dilepton context, the minimax variable sets a kinematic minimum for the b-ℓ pairing. It has the useful property of being bounded by the top quark's mass for $t\bar{t}$ events but can be much larger for tW events in cases where the wrong pairing is chosen. As a result, the minimax variable is sensitive to the interference effects between $t\bar{t}$ and tW, particularly in the region $m_{b\ell}^{\mathrm{minimax}} > m_{\mathrm{top}}$.

The distribution of this variable shown in Figure 3.12 already shows the differences between DR and DS. The difference is especially relevant in the tails of the minimax distribution. In the semilep-

tonic channel, we cannot construct an identical minimax variable using two leptons and two b-jets as input since there is only one lepton in each event. Therefore, this analysis constructs a modified minimax variable by replacing the second lepton with the hadronic-W. The construction and distribution of this variable will be described in Section 6.3.

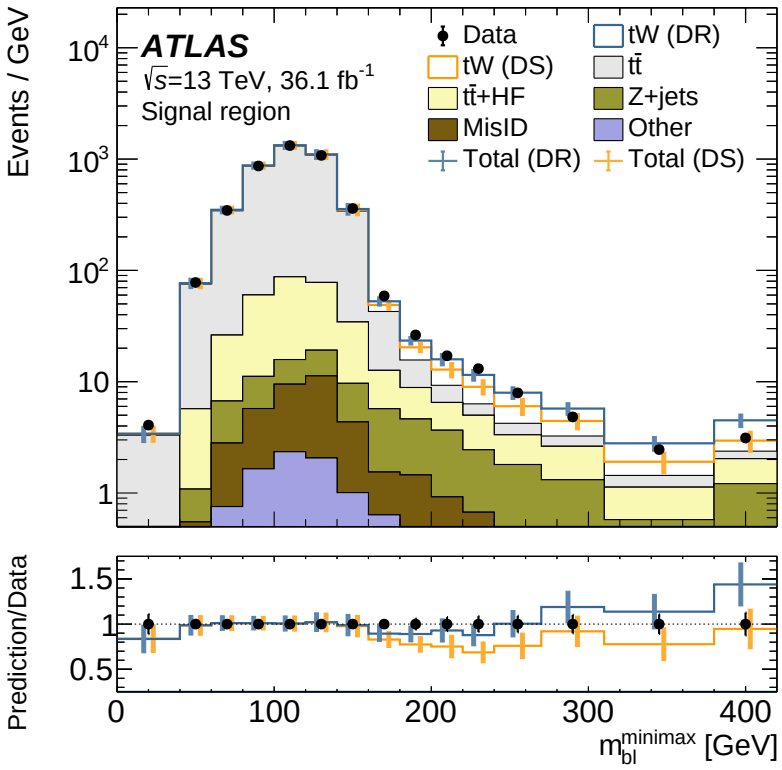


Figure 3.12: Minimax Distribution Plotting the DR/DS Interference Effect. The detector-level $m_{b\ell}^{\text{minimax}}$ distribution is plotted for both signal and background samples. The total MC prediction is evaluated using both the DR and DS approaches for modeling the single-top tW samples. The main panel plots the data compared against the MC simulated processes. The bottom panel plots the ratio between the total predicted distribution and data. The blue curve is the ratio when using the DR approach, while the orange curve shows the ratio when using the DS approach. The differences between DR and DS can be clearly observed at large minimax values. Plot is taken from Ref. [39].

To obtain the underlying truth distribution, the detector effects must be mathematically abstracted away in a process known as unfolding. The unfolded distribution of the $m_{b\ell}^{\text{minimax}}$ variable is shown in Figure 3.13. The measured distributions are compared to different combinations of MC generators. Each generator uniquely handles the parton showering and interference calculations. The previously mentioned $bb4l$ generators have not yet become available for comparison against real data.

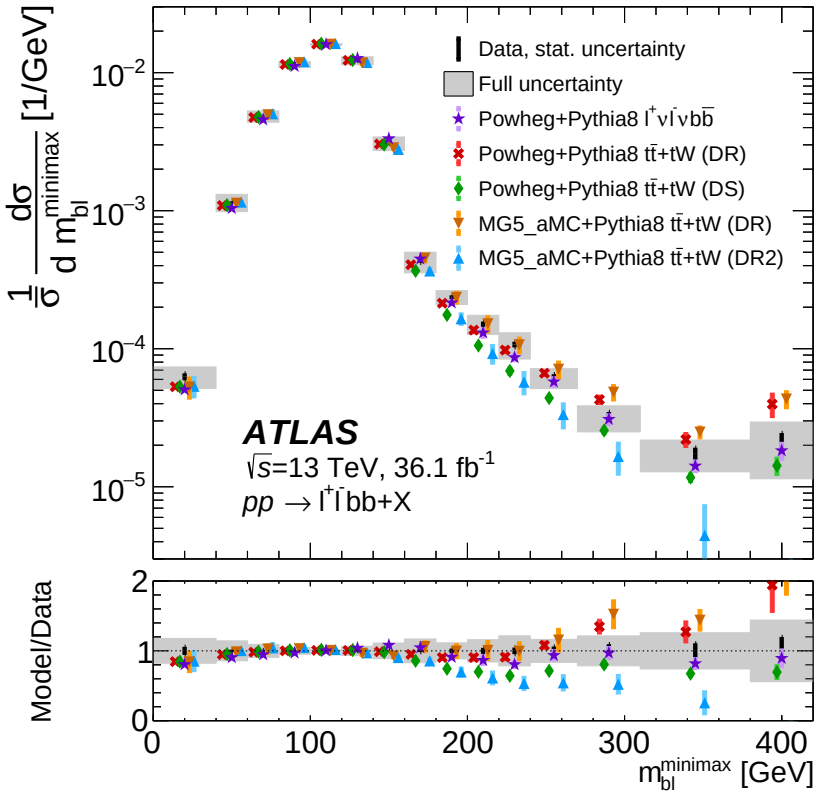


Figure 3.13: Unfolded Minimax Distribution Plotting the DR/DS Interference Effect. The detector-level distribution of the $m_{b\ell}^{\text{minimax}}$ quantity is unfolded to reveal the underlying cross-sections. The main panel plots the unfolded and normalized differential cross-section measurement represented by the black points. The various colored points represent the different combinations of MC generators for the DR and DS interference. DR2 is an additional method for computing this interference, which includes the doubly-resonant contributions directly in the interference term. DR2 is much less used than the DR and DS schemes and is not explored further in this analysis. The ratio plot measures the ratio between the different MC generators and the measured distribution. Plot is taken from Ref. [39].

CHAPTER 4

LHC and ATLAS

The LHC is the world's most powerful particle accelerator situated at the Conseil Européen pour la Recherche Nucléaire (CERN) complex which straddles the Swiss-French border near Geneva. Two counter-rotating beams of protons accelerate within a 27 km ring, reaching a final speed at the moment of collision approaching the speed of light limit. Collisions occur at four interaction points in the ring surrounded by the four major experiments. The previously mentioned ATLAS and CMS detectors are the two general-purpose experiments designed to serve a multi-faceted physics program. A Large Ion Collider Experiment (ALICE) studies the products of heavy-ion ($Pb - Pb$) collisions and is uniquely suited for studying the conditions of the very early universe. Large Hadron Collider beauty (LHCb) is the fourth major experiment at the LHC and has a concentrated focus on the physics of the b-quark. This analysis uses data collected by the ATLAS detector. This chapter will briefly describe the LHC and the ATLAS collider.

4.1 Large Hadron Collider

The LHC [42] results from an immense input of effort by a multinational collaboration of scientists to try to fill the remaining gaps in the SM. Construction of the LHC began in earnest after the decommissioning of the Large Electron-Positron (LEP) project in 2001 [43]. The LHC repurposed the LEP tunnels to house the ring. The project completed construction and began firing its first beams of protons in 2008. Within two years of completing construction, the main physics program began fully when the collision energy increased to 7 TeV. Between 2008 and 2012, there would only be a modest increase in energy to 8 TeV. During this time, the ATLAS and CMS collaborations reported the discovery of the Higgs boson [44]. The next big jump in energy occurred in 2015 when the energy would reach a new energy frontier of 13 TeV. Recently, the energy would increase again in 2022 to its present-day collision energy of 13.6 TeV.

4.1.1 *Pre-Accelerator Chain*

The LHC ring is the final step in a series of linear and circular pre-accelerators to increase the energy in steps up to the final 13.6 TeV energy. Significant upgrades have been made to the CERN accelerator complex since the release of the LHC design report [45, 46]. Before 2020, the protons that would eventually end up in the LHC ring began life as atomic hydrogen, consisting of a single proton and a single electron. A bare proton is created by ionizing the hydrogen gas to strip away the electrons. Recently, the LHC replaced the atomic hydrogen with negatively-charged ionic hydrogen (H^-), which is atomic hydrogen but with an additional electron. H^- ions are produced at the Linear Accelerator 4 (Linac4) facility which replaced the previously-used Linear Accelerator 2 (Linac2) facility.

Many accelerator complexes (including CERN) have begun operating using H^- ions because of the advantages it offers over atomic hydrogen [47]. To start with, H^- and H^+ beams are naturally electromagnetically attractive. Injecting H^- ions from Linac4 into the already circulating proton beam in the PSB creates a single tightly-packed proton beam with high intensity. The H^- ions still carry extra electrons, but these are easily stripped away by passing the beam through a stripping foil with near 100% efficiency. Previously, a physical kicker magnet stripped away the electrons with much lower efficiency. These reasons make using H^- ions ideal for filling the LHC storage rings.

Linac4 creates the H^- ions and accelerates them to 160 MeV before entering the next link in the pre-accelerator chain. The 160 MeV ions will have their two extra electrons stripped before entering the Proton Synchrotron Booster (PSB) [48]. The now bare protons are then accelerated to 2 GeV before entering the Proton Synchrotron (PS), increasing the energy of the proton beam to 26 GeV. The Super Proton Synchrotron (SPS) accepts the protons and accelerates them to 450 GeV before they are separated into the two beam-pipes that make up the LHC. Figure 4.1 shows the layout of the CERN pre-accelerator chain.

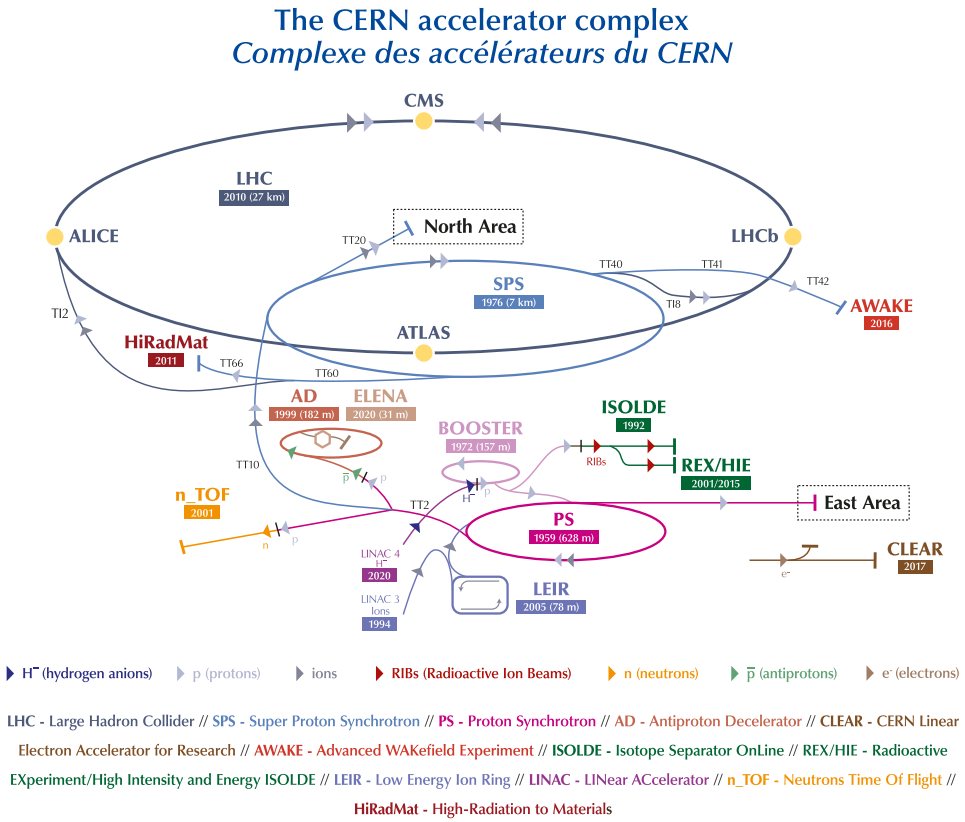


Figure 4.1: CERN Accelerator Complex and LHC Injection Systems. The CERN accelerator complex comprises several pre-accelerators and the main LHC ring. The chain begins in the Linac4 facility, which sources the H^- ions. The H^- ions have their excess electrons stripped, leaving a bare proton. The proton experiences incremental jumps in energy as they progress through the pre-accelerator rings. Protons enter the LHC ring at 450 GeV and reach their final energy of 6.8 TeV in approximately 20 minutes. Plot is taken from Ref. [49].

4.1.2 *Radio-Frequency System*

Upon reaching the LHC ring, the protons circulate the 27 km ring approximately 10000 revolutions every second. Protons gain energy every time they traverse a Radio-Frequency (RF) cavity. An RF cavity, in its most basic form, is two conducting plates that contain an oscillating electric-field. The period of oscillation is set to perfectly time the direction of the electric field with the incoming proton to maximize the efficiency of energy transfer to the proton. The oscillation of the electric fields has a secondary effect of concentrating the protons into bunches. Protons with more energy experience less acceleration, while those with less energy experience more acceleration. As a result, the protons do not form a continuous beam but are organized into bunches. At the LHC, each RF system has eight such cavities, and each cavity carries $V = 2$ MV of accelerating voltage. Therefore, a proton traversing the complete chain of eight cavities gains a $E = qV = 16$ MeV of energy per revolution. The maximum 6.5 TeV of energy is reached after the proton has made approximately ten million revolutions.

4.1.3 Magnets System

A combination of dipole and quadrupole magnets bend and focus the protons into a circular orbit in the LHC rings. The magnet system at the LHC uses superconducting *NbTi* superconducting magnets, which operate below a temperature of 2 K and generate a magnetic field above 8 T. Superconducting materials can carry large currents with near-zero resistance, reducing energy losses while simultaneously generating a large magnetic field. A total of 1232 dipole magnets surround the ring to bend the beam of protons into a circular trajectory. These magnets correct minor deviations in the circular trajectory during the protons' millions of revolutions. Quadrupole magnets, of which there are 474, conduct fine beam focussing. They compress the beam of protons in one direction while simultaneously expanding the beam in the other direction. Multiple quadrupole magnets with rotated polarities are strategically placed along the ring, ensuring equal compression of the proton beam from all directions. The result is a high density of protons, significantly increasing the chance of collisions.

4.1.4 Luminosity

Each of the two 6.5 TeV beams during the Run 2 data-taking period contains 2556 bunches organized by the RF system. The bunches are approximately 8 cm in length and contain 11×10^{10} protons in each bunch. At regular 25 ns intervals, two counter-rotating bunches of protons collide at the four interaction points. The collision probability is rare even with billions of protons in each beam. The number of events of interest ($N_{\text{exp.}}$) over some time is related to the cross-section of the process of interest and a quantity known as the luminosity, $\mathcal{L}$.

$$N_{\text{exp.}} = \sigma_{\text{exp.}} \cdot \int \mathcal{L}(t)dt \tag{4.1}$$

Specifically, $\mathcal{L}$ is the instantaneous luminosity of the beam and measures the number of particles per area per time. Instantaneous luminosity carries units of $\text{cm}^{-2}\ \text{s}^{-1}$ and is often quoted as the figure of merit for particle accelerators. Increasing the instantaneous luminosity is a goal for many particle physics experiments so that more data can be collected in a shorter amount of time. With more data, the occurrences of rare events with low production cross-sections increase. $\mathcal{L}$ can only be adjusted by tweaking the parameters of the experiment and beam as shown in Eq. 4.2.

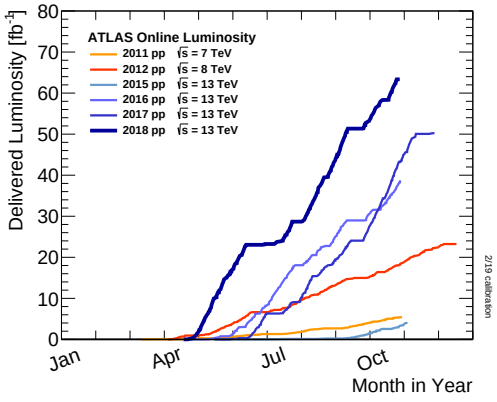
$$\mathcal{L} = f_{\text{coll}} \frac{N_1 N_2}{4\pi\sigma_x^* \sigma_y^*} \mathcal{F} \tag{4.2}$$

In this equation, f_{coll} is the frequency of colliding bunches, each containing N_1 and N_2 number of particles. $\sigma_{\text{x,y}}^*$ measures the transverse beam sizes in the horizontal and vertical directions at the interaction point. The proton beams at the LHC have a $\sigma_{\text{x,y}}^* \approx 220\ \mu\text{m}$ [50]. Finally, the $\mathcal{F}$ factor of order 1 considers certain inefficiencies of the beam dynamics, such as the crossing angle and finite bunch length. Table 4.1 lists the relevant factors for the LHC between the years $2015 - 2018$ corresponding to the Run 2 data-taking period. During this period, the peak instantaneous luminosity is on the order of $10^{34}\ \text{cm}^{-2}\ \text{s}^{-1}$.

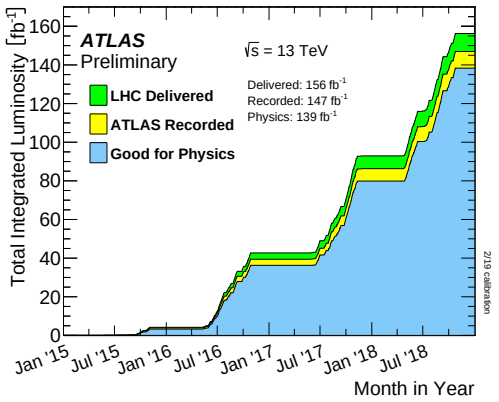
Table 4.1: Selected LHC Beam Parameters for *pp* Collisions During Run 2. During Run 2 ($2015 - 2018$), there were several adjustments to the LHC beam parameters. These beam parameters are used to compute the final peak luminosity values. A particular parameter of interest is the β^* parameter, which quantifies the degree of focusing at the interaction point. A unique situation occurred in 2017 where the "8b4e" filling scheme was used, where eight bunches separated by 25 ns were followed by a four bunch-slot gap. Table is taken from Ref. [51].

Parameter	2015	2016	2017	2018
Maximum number of colliding bunch pairs (n_b)	2232	2208	2544/1909	2544
Bunch spacing [ns]	25	25	25/8*b*4*e*	25
Typical bunch population [10^{11} protons]	1.1	1.1	1.1/1.2	1.1
β^* [m]	0.8	0.4	0.3	$0.3 - 0.25$
Peak luminosity $\mathcal{L}_{\text{peak}}$ [10^{33} cm^{-2} s^{-1}]	5	13	16	19
Peak number of inelastic interactions/crossing ($\langle\mu\rangle$)	~ 16	~ 41	$\sim 45/60$	~ 55
Luminosity-weighted mean inelastic interactions/crossing	13	25	38	36
Total delivered integrated luminosity [fb^{-1}]	4.0	39.0	50.6	63.8

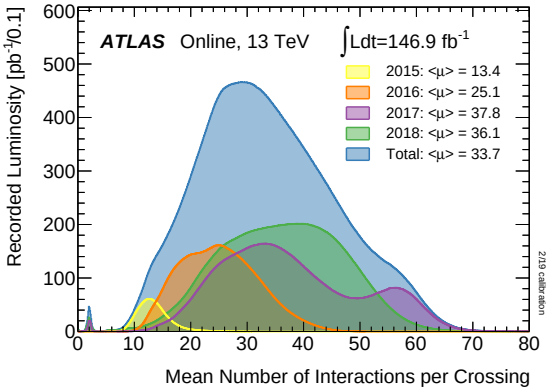
When the instantaneous luminosity is integrated over a specific period of time, the resulting quantity is now the integrated luminosity. The result measures the number of particle collisions and is directly related to the amount of data collected by the experiment over that period. The integrated luminosity has units of inverse area and is often reported in units of inverse femtobarn (1 fb^{-1} = 10^{-43} m^2). For example, the total cross-section for Higgs boson production is approximately 6000 fb at 13 TeV. For every 1 fb^{-1} of data collected, one can expect to produce 6000 Higgs boson events. This example illustrates the importance of collecting and aggregating data from many years to increase the number of potential events in the entire dataset. The total integrated luminosity of the entire Run 2 dataset is approximately 140 fb^{-1}. Figure 4.2 presents various luminosity summary plots delivered by the LHC and collected by ATLAS.



(a) Delivered luminosity versus time.



(b) Total integrated luminosity and data-quality versus time.



(c) Total integrated luminosity versus mean number of interactions.

Figure 4.2: Selected LHC/ATLAS Luminosity Summary Plots. The plots summarizes various luminosity figures reported by the LHC and ATLAS during 2011 – 2018. **(a)** plots the total integrated luminosity which was delivered by the LHC during both Run 1 (2011 – 2012) and Run 2 (2015 – 2018). During Run 2, the LHC delivered a total of 156 fb^{-1}. As shown in **(b)**, ATLAS collected a total of 147 fb^{-1}. A requirement of good object quality reduces the data set to the final 140 fb^{-1} total. Lastly, **(c)** plots the recorded luminosity by ATLAS as a function of the mean number of interactions per crossing, $\langle\mu\rangle$. This value represents the average number of proton-proton interactions when the bunches collide. Plots are taken from Ref. [51].

4.2 ATLAS Detector

The ATLAS detector is one of two general-purpose particle detectors placed around the LHC ring [4]. The experiment was specifically designed to capture data that can be used across a broad physics scope. The physics program includes precision measurements of the SM, as in the context of this analysis, to more exotic physics and probing for theorized processes that have yet to be directly measured. The basic design principle has mostly stayed the same since the project's conception. ATLAS is a near-hermetic structure divided into several subsystems, each with its own purpose and functionality. The

main components of the detector are concentrically placed surrounding the primary interaction point.

Inner Detector (ID) The Inner Detector is the subsystem closest to the interaction point. The ID is primarily responsible for reconstructing the trajectories and momenta of particles produced during the collision. It is characterized by its high granularity, which is crucial for achieving the high precision needed to resolve many multi-track events.

Calorimeters The ATLAS calorimeter system is two-tiered, comprised of the electromagnetic and hadronic calorimeters. The calorimeters measure the energy of a passing particle and provide a second measurement of the particle's position complementary to the ID. The energy of an incoming particle is determined by measuring the energy of charged secondary particles created by the interaction of the primary particle and the calorimeter material.

Muon Spectrometers (MS) The Muon Spectrometers are the outermost subsystem of the detector and are used to track the passage of muons. The trajectory of muons is inferred by collecting the ionization of gas molecules in gas-filled drift tubes.

Trigger and Data Acquisition (TDAQ) The ATLAS Trigger and Data Acquisition system selects interesting events for storage. The trigger system is two-tiered. A primary Level-1 (L1) trigger is implemented at the hardware level and performs a rapid preliminary selection. The secondary High-Level Trigger (HLT) is implemented at the software level and applies more sophisticated algorithms to select events.

The goal is to capture all information associated with the diverse array of SM particles produced during a pp collision. These include the quarks and gluons (observed as jets), leptons, $W^{\pm}$, Z, H-bosons, photons, and other hadrons composed of quarks. Throughout its 15-year existence, the detector has undergone several shutdowns. During these shutdowns, upgrades were made to several subsystems to prepare for increases in both luminosity and energy delivered by the LHC. Figure 4.3 shows an illustrated drawing of the ATLAS detector at the onset of Run 3.

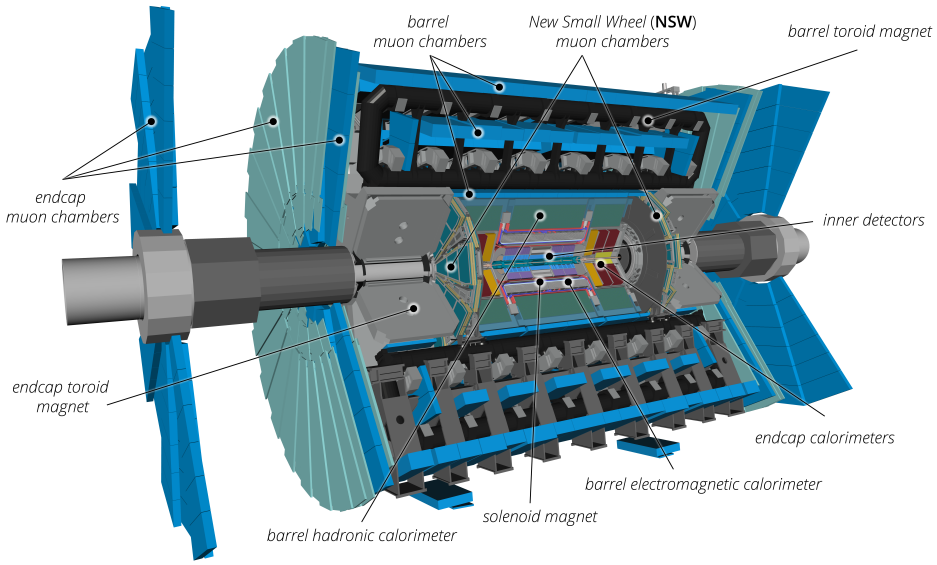


Figure 4.3: Labeled Cutaway of the ATLAS Detector. A cutaway of the ATLAS detector at the start of Run 3 is shown with each of the main subsystems labeled. Each subsystem is specialized for a unique task during a proton-proton collision. The Inner Detector reconstructs hits into particle tracks with high granularity. A combination of electromagnetic and hadronic calorimeters are used to measure energy deposition. Weakly interacting muons passing through the bulk of the active detector are registered by the specialized muon chambers. Specifically highlighted is the addition of the New Small Wheel (NSW). These upgrades to the muon system can accommodate the higher rates of Run 3 and improve muon triggering and tracking. The entire detector is immersed in a magnetic field created by a combination of solenoid and toroid magnets to bend the particle trajectories. The result is a complete collection of the physics observables from the collision. Plot is taken from Ref. [52].

The subsystems work cooperatively to reconstruct the signals the detector collects. Table 4.2 lists the main performance goals of the four main detector subsystems.

Table 4.2: General Performance Goals of the ATLAS Detector. The expected resolution for each of the four main detector subsystems is listed along with their pseudorapidity (η) coverages. Table is taken from Ref. [4].

		η Coverage	
Detector Component	Required Resolution	Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\%\ p_T \oplus 1\%$	± 2.5	
EM Calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic Calorimetry (jets)			
barrel and endcap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

The result is a set of observables corresponding to the potentially thousands of physics objects produced during a single collision event. Additional data-cleaning steps are needed to calibrate the data according to the current detector conditions, and tagging procedures are needed to characterize

the origin and quality of the events collected. At this point, it is up to the analysts to use this dataset to complete their analysis. The remainder of this section describes each of the main subsystems.

4.2.1 *Coordinate System*

This section defines the standard coordinate system used throughout ATLAS, and illustrated in Figure 4.4. Naturally, the shape of the ATLAS detector lends itself to using a cylindrical-based coordinate system. The interaction point defines the origin of the coordinate system, with the z-axis running through the interaction point and along the beam line. The positive x-axis points from the interaction point to the center of the LHC ring. The positive y-axis points upwards towards the surface.

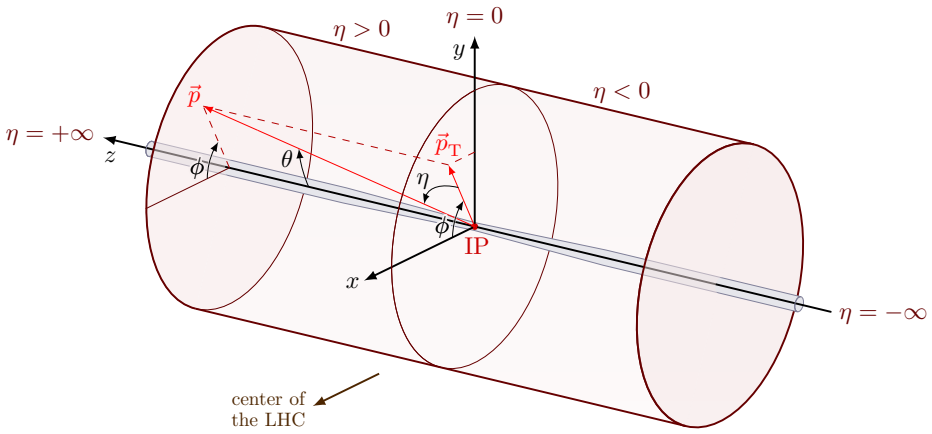


Figure 4.4: Schematic of the Common ATLAS Coordinate System. The ATLAS detector uses a right-handed coordinate system, which places the origin at the primary interaction point. The z-axis passes through the origin and lies along the beamline. The positive x-axis points from the interaction point to the center of the LHC ring. The positive y-axis points from the origin to the surface. Particle positions and momenta are defined using the standard cylindrical coordinate angles, θ and ϕ. The polar angle, θ, is used to define the much more convenient pseudorapidity quantity to measure the angle of a particle relative to the beam axis. Plot is taken from Ref. [20].

The polar angle, θ, and azimuthal angle, ϕ, are defined as usual in the standard cylindrical coordinate system. The transverse $x - y$ plane runs perpendicular to the beam line. Transverse momentum (p_{T}) is a particularly important quantity because a large p_{T} indicates that a hard-scatter interaction occurred between the partons of the two colliding protons, which carried no momentum in the transverse plane before the collision. The polar angle is also used to define another useful quantity known as pseudorapidity.

$$\eta = -\ln(\tan(\theta/2)) \tag{4.3}$$

Pseudorapidity, η, is a useful quantity in hadron colliders to define the angle of a particle with respect to the beam axis. Particles with a larger $|\eta|$ are closer to the beamline and are likely to be detected near the endcap regions of the detector. The range of $|\eta|$ defines two exclusive regions in the detector, $|\eta| < 2.5$ defines the central region, while the $|\eta| > 2.5$ defines the forward regions of the detector.

Pseudorapidity is a much more useful quantity than simply θ because it can be shown that differences in η are Lorentz-invariant[1]. Therefore, when making measurements as a function of $\Delta\eta$, a Lorentz boost can be applied to the plot without changing the underlying physics. For example, in situations with a large imbalance of momentum that produces particles at one end of the detector, performing a translation in η back to the center-of-mass frame is trivial.

4.2.2 *Inner Detector*

The ATLAS Inner Detector (ID) [53, 54] refers to the set of subdetectors closest to the beam axis and is responsible for tracking charged particles by measuring the particle's charge and momentum. Figure 4.5 shows a labeled schematic of the ID and its components. A central solenoid surrounding the ID generates a 2 T magnetic field, immersing the entire ID. When a charged particle enters the magnetic field, the path of the particle (track) is deflected, creating an arced trajectory. The direction of curvature determines the particle's charge and the radius of curvature determines the particle's momentum.

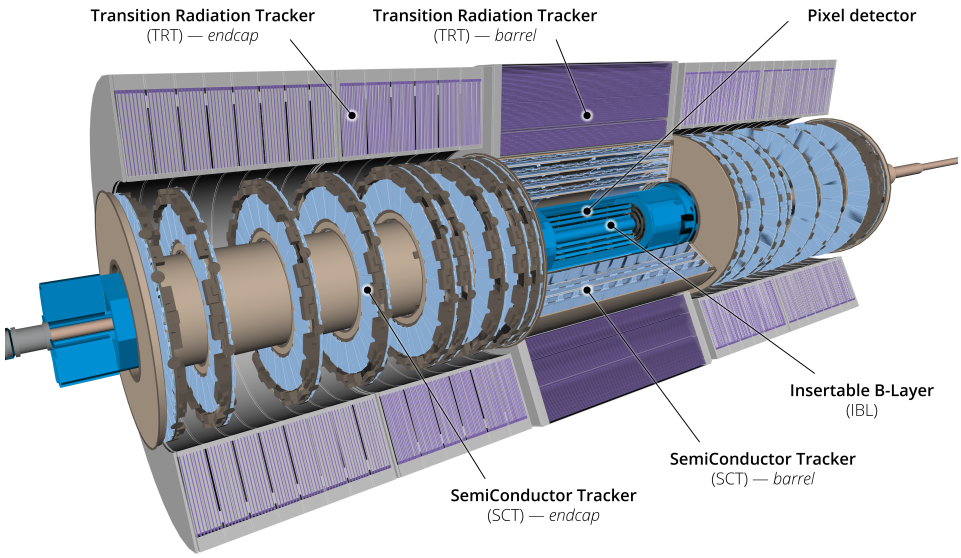


Figure 4.5: Labeled Cutaway of the ATLAS Inner Detector. A cutaway of the ATLAS Inner Detector is shown with its main components labeled. The main components of the ID include the Pixel Detectors, Semi-Conductor Trackers (SCT), and the Transition Radiation Trackers (TRT). One of the most recent upgrades is the addition of the Insertable B-Layer. This layer of silicon detectors is the closest layer to the interaction point and was added to improve the tracking of low-momentum particles. Plot is taken from Ref. [52].

There are several requirements the ID must possess by being so close to the primary interaction point. Protons collide in bunches every 25 ns within the ATLAS detector; the electronics of the ID

[1] Pseudorapidity, η is an approximation of the much more general rapidity quantity, $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) = \tanh^{-1}\left(\frac{p_z}{E}\right)$. Unlike pseudorapidity, whose differences are only Lorentz-invariant, rapidity is by itself Lorentz-invariant. However, it is a much more difficult quantity to work within an experiment setting where the total energy (E) and complete 3-momentum vector must be known. Often, the longitudinal (p_z) can not be precisely measured, especially for missing energy. Luckily, one can make the approximation that for near-massless particles (particles with momentum much greater than their rest mass, $m \ll p$), the pseudorapidity is equal to rapidity.

must have a fast enough readout speed to cope with this 25 ns collision window. In addition, these electronics must be sufficiently radiation-hard while constantly facing the incoming flux of particles. It is also necessary that within the ID, the amount of dead material (i.e., non-active detector material used for structural purposes) should be reduced to near zero. Any dead material near the interaction point can produce unwanted secondary particles and photons not associated with the initial hard-scatter event.

To achieve the precise momentum reconstructions used for physics analyses, the ID is segmented into three components: the pixel detectors, the Semi-Conductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The three subdetectors work cooperatively to cover the entire central detector ($|\eta| < 2.5$) region and complete 2π coverage in ϕ with high-precision tracking capabilities. A brief description of each of these systems is provided below.

4.2.2.1 Pixel Detector

The innermost set of subdetectors is the pixel detectors that cover a radial distance from 45.5 mm to 242 mm, offering the highest-granularity tracking capabilities [55]. Before the start of the Run 2 data-taking period, there were three active pixel detector layers in the barrel and two identical endcap layers. A fourth barrel layer, the insertable b-layer [56], was installed as the closest layer to the beam pipe during the long shutdown before Run 2. The layer aims to improve track resolution by introducing an additional layer of high-precision detectors. In particular, this improves the resolution of low p_{T} particles, which are more sensitive to multiple scattering effects. Jet tagging is one area that will greatly benefit from this improvement.

The basic building blocks of the pixel detector are modules consisting of silicon pixels. A single silicon pixel has a size of $50 \times 400\ \mu\mathrm{m}^2$ (in the insertable b-layer, the size of a pixel sensor is $50 \times 250\ \mu\mathrm{m}^2$). A single module comprises 46080 individual pixel readout channels. The result is 80×10^6 readout channels alone in the pixel detector. The operating principle of a silicon sensor relies on creating a depletion region by applying a reverse-biased voltage to a PN (positive-negative) junction. When charged particles traverse the depletion region, they create electron-hole pairs (negative and positive charge carriers) in the material. These pairs drift to opposite sides of the sensor under the influence of the electric field, thereby generating a current. The amount of charge collected and the pixel's position correspond to the amount of energy lost by the particle and the particle's position, respectively. Therefore, the spatial resolution of the pixel detector is mainly determined by the dimensions of the individual pixel.

4.2.2.2 Semi-Conductor Tracker

The middle layer of the ID after the pixel detectors is the Semi-Conductor Tracker (SCT), which covers a radial region from 255 mm to 549 mm [57]. Like the pixel detectors, the SCT uses silicon as the active detector material in the form of 4088 silicon-strip detectors. These are arranged in four concentric layers in the barrel and two layers at the endcaps. The silicon sensors are further organized into modules. The construction of each silicon-strip module is very simple; two planes of silicon sensors

are glued together with a baseboard material in the middle. The baseboard material provides thermal and mechanical strength of the module and gives each module a length of 128 mm (there are different dimensions for modules on the barrel and the endcaps). Using strips instead of rectangular pixels, the SCT has less resolution (in the longitudinal direction) than the inner pixel detector and contains a total of 6.3×10^6 readout channels. The overall spatial resolution of each sensor is 17 μm in the $r - \phi$ direction and 580 μm in the z direction.

4.2.2.3 Transition Radiation Tracker

The outermost layer of the ID is the Transition Radiation Tracker (TRT) and covers the radial distance from 554 mm to 1082 mm [58]. The TRT consists of approximately 370000 straw tubes, each with a diameter of 4 mm, which run parallel to the beamline. A single straw tube operates as a cylindrical drift chamber made of Kapton and filled with a mixture of Xenon and Argon gas. Running down the central axis of each tube is a 31 μm diameter Tungsten wire kept at a potential difference relative to the straw tube walls. This potential difference creates an electric field so that the primary electrons accelerate towards the central wire. The accelerating electrons produce secondary ionization in the gas, which generates a measurable current signal. The TRT also allows for electron identification using polypropylene fibers (barrel) and foils (endcap) dispersed between the straw tubes to measure the charge produced by the transition radiation. When particles move between two different materials, transition radiation can be produced. The radiated photons interact with the gas and free more electrons resulting in the generation of a large signal. Particles with a higher relativistic speed (electrons) will produce more transition radiation than slower-moving particles (pions), which provides an easy solution for separating electrons from the sea of other charged particles [59].

4.2.3 Calorimeters

Calorimeters are devices that measure the energy of a passing particle by collecting the energy of the secondary showers initiated by the incident particle. The ATLAS calorimeter serves two purposes: the first is to measure the energy of the incoming particle, and the second is to aid in the measurement of the particle's position alongside the ID. There are two main classes of calorimeters: sampling and homogenous. Sampling calorimeters use alternating layers of absorbing material (such as lead), which produces an avalanche of particle showers, and active detector material which collects the showering energy and measures the total energy. On the other hand, a homogenous calorimeter combines these two functions into a single complete material. The ATLAS calorimeter system, shown in Figure 4.6, uses the sampling calorimeter construction for both the electromagnetic and hadronic calorimeters.

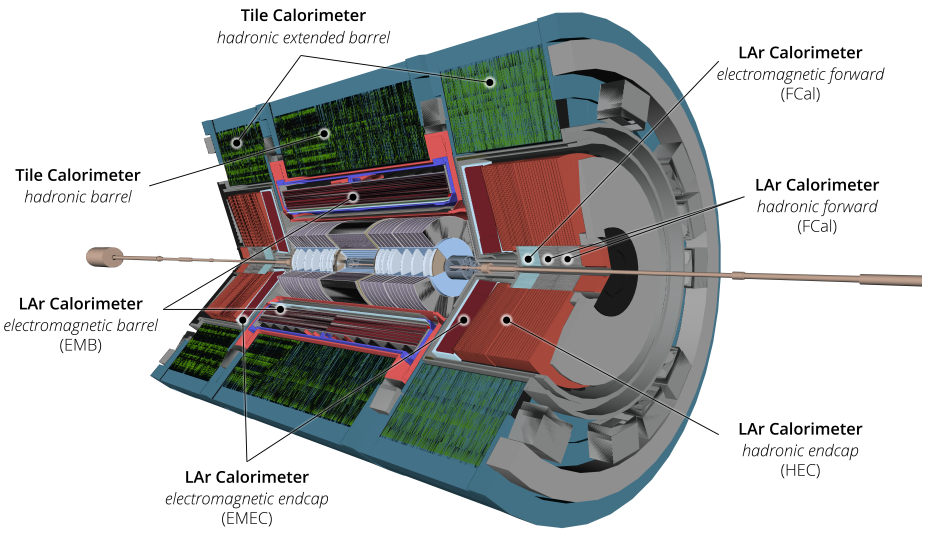


Figure 4.6: Labeled Cutaway of the ATLAS Calorimeter System. A cutaway of the ATLAS calorimeter system is shown with its main components labeled. Two main calorimeters used to identify different types of particles: an electromagnetic calorimeter for electrons and photons and a hadronic calorimeter for measuring jets' energy. Liquid Argon is used throughout the electromagnetic calorimeter as the active medium, while the hadronic calorimeter uses scintillators connected to photomultiplier tubes for measurement. Plot is taken from Ref. [52].

The energy resolution of a calorimeter is naturally a function of the energy of the incident particle. Interactions of particles with higher energy with the absorbing material produce more secondary showers, generating a larger signal response. There are other contributions to resolution which depend on the geometry and construction of the calorimeter system. These contributions are known as the stochastic, noise, and constant terms. The total energy resolution is a quadratic sum of all of these contributions as shown in Eq. 4.4 [60].

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{4.4}$$

In the above equation, the first term on the right-hand side is the stochastic term, which accounts for fluctuations in the signal-generating process (i.e., the number of secondary photo-electrons produced). The second contribution to the energy resolution is the noise term, which accounts for the inherent electronic noise in the readout chain. The constant term, which has no dependence on energy, characterizes the level of the calorimeter response non-uniformity. This is mainly due to imperfections in the detector geometry and construction, damage from long-term radiation exposure, or unaccounted-for secondary ionization from dead material. The resolution of the ATLAS calorimeters was measured using test beams and yielded the following values: $a = 10\% \sqrt{\text{GeV}}$, $b \approx 200$ MeV, and $c = 0.2\%$ [61].

4.2.3.1 *Electromagentic Calorimeter*

The EM calorimeter measures the energy of charged particles such as electrons and photons and covers a total spatial region up to $|\eta| < 3.2$ [62]. The EM calorimeter has two components: the barrel, which covers up to $|\eta| < 1.475$, and the two endcaps, which extend the coverage to $|\eta| < 3.2$. The EM calorimeter uses lead as its passive absorber and Liquid Argon (LAr) as the active detection media. Liquid Argon was ultimately chosen because of several properties: (1) Argon can be purified to a high level of homogeneity, ensuring consistent performance in different parts of the calorimeter, (2) a linear response over a wide range of energies, and (3) high stability in a radiation-intense environment. When a charged particle impacts a layer of lead and creates a particle shower, the secondary particles generate ionization in the liquid Argon layers that drift under high voltage to anode wires that collect the charge. Since the amount of charge released by the ionization event is proportional to the energy, the primary particle's energy can be deduced.

The barrel portion of the EM calorimeter is segmented into cells with different granularities depending on the layer in which the cell is located. Figure 4.7 is a labeled drawing of a module found in the EM calorimeter.

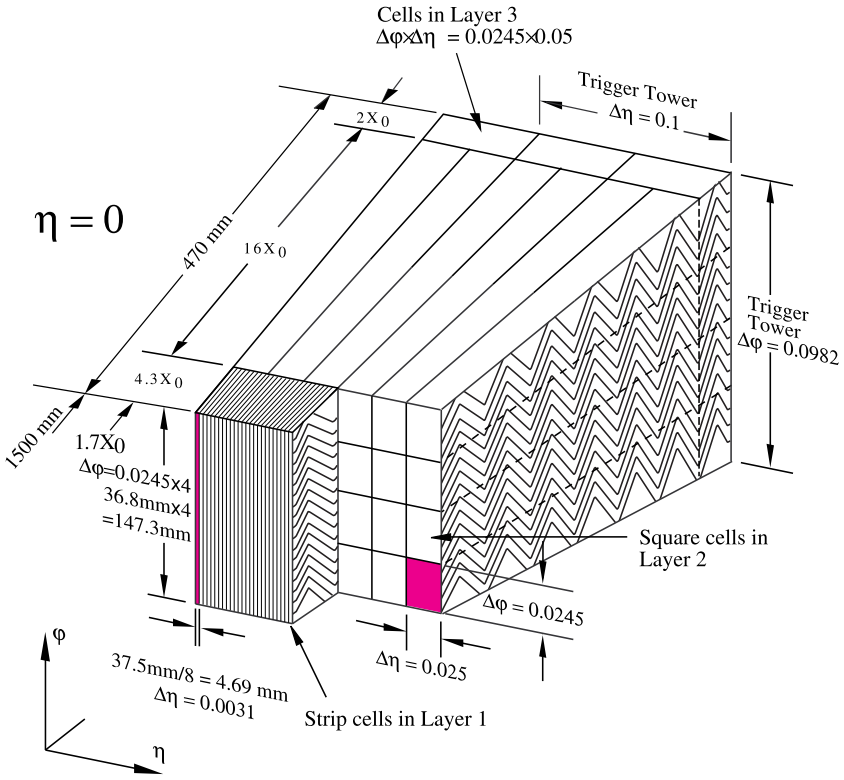


Figure 4.7: Barrel Module of the ATLAS Electromagnetic Calorimeter. The basic building block of the ATLAS EM calorimeter is a barrel calorimeter module. Non-uniform-sized cells are found in the three layers of the module, with cells having the highest granularity being located in the single Presampler layer. The considerable depth of the second layer ensures that most of the energy is deposited in the second layer. Plot is taken from Ref. [4].

Cells in the initial "Presampler" layer have the highest granularity in $\eta - \phi$ to account for any energy loss that may occur before reaching the bulk of the remaining calorimeter cells. The majority of the deposited energy occurs in the second layer and thus is the layer with the largest depth[2]. Finally, any remaining energy deposited by highly energetic particles ends up in the third layer of the calorimeter with the lowest $\eta - \phi$ granularity. The high degree of segmentation in the EM calorimeter aids in tracking the positions of the particles in combination with the ID.

4.2.3.2 *Hadronic Calorimeter*

Quarks and gluons produced during pp collisions quickly hadronize and appear as jets in the detector. The Hadronic Calorimeter (HCAL) surrounds the EM calorimeter and is responsible for measuring the energy of these strongly interacting particles and consists of three sections: the tile calorimeter, the hadronic endcap calorimeters, and the forward calorimeters. Hadronic interactions with matter in the HCAL produce secondary hadronic showers proportional to the energy of the incident particle. Unlike the EM showers, the secondary showers produced by hadrons involve the strong interaction and produce a more extensive variety of particles. The result is a broader shower, which is characteristic of a hadronic shower.

Before reaching the HCAL, the primary hadrons first pass through the EM calorimeter without depositing large amounts of energy[3][4]. This deposition occurs because of the intrinsically large interaction lengths[5] of these particles. The tile calorimeter, whose module is shown in figure 4.8, covers a region up to $|\eta| < 1.7$ and is also constructed as a sampling calorimeter. The absorbing material is steel, and plastic scintillators are the active detector material. The two hadronic endcap calorimeters use liquid Argon as the detector medium interleaved with copper places acting as the absorber and covers a range from $1.5 < |\eta| < 3.2$. Finally, the two forward calorimeters cover the remaining $3.1 < |\eta| < 4.9$ range. Three modules make up the forward calorimeters. One module made of copper and placed closest to the interaction point is optimized for EM interactions. The other two modules are made of tungsten and are optimized for hadronic interactions. Notice that the hadronic endcap calorimeters overlap with the tile and forward sections of the HCAL. The overlap in these regions is necessary to compensate for a drop in material density.

[2] Calorimeter depth can also be stated in terms of a quantity known as radiation length, X_0 given in areal-density units of $\mathrm{g \cdot cm^{-2}}$. Radiation length is defined as the amount of material traversed by a particle before losing all but $1/e$ of its original energy. This $1/e$ factor comes from the fact that the energy loss of a particle through matter follows exponential decay. The radiation length of pure Argon is 19.5489 $\mathrm{g \cdot cm^{-2}}$ [63].

[3] For charged hadrons such as the charged pions ($\pi^{\pm}$), a large fraction of their energy can also be deposited in the EM calorimeters. Neutral hadrons such as neutrons or hadrons interacting primarily through the strong force will deposit most of their energy in the hadronic calorimeter by passing the EM calorimeter without significant energy deposition.

[4] Distinguishing energy deposits by different particles in the EM calorimeter is crucial for accurate object reconstructions (e.g., pion vs. electron). The large mass difference between a pion ($\sim$ 130 MeV) and an electron ($\sim$ 0.5 MeV) manifests itself in the track information where an electron will have a smaller radius of curvature than a pion. Other differences in the electromagnetic shower profile can also be used to differentiate between pions and electrons.

[5] For hadronic interactions, the distance traveled by a particle is defined using the interaction length, λ. The interaction length quantifies the average distance a particle travels in a material before undergoing a significant interaction (e.g., nuclear scattering). This length depends on the type of interaction, the properties of the primary particle, and the characteristics of the medium.

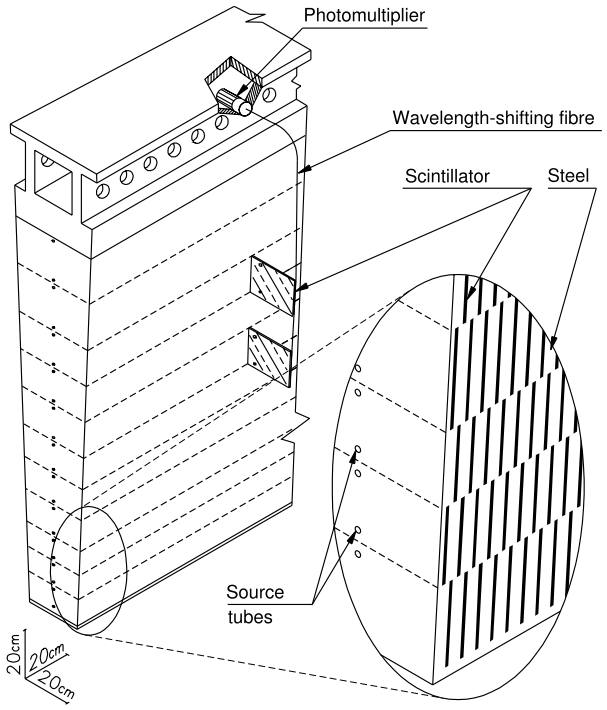


Figure 4.8: Module of the Tile Calorimeter in the ATLAS Hadronic Calorimeter. An individual module of the tile calorimeter in the ATLAS HCAL is shown with the major parts labeled. The tile calorimeter is a sampling calorimeter that uses steel as the absorbing material and scintillators coupled to photomultiplier tubes for active detection. Plot is taken from Ref. [4].

4.2.4 Muon Spectrometer

Muons are the electron's much heavier cousin and penetrate the two layers of calorimeters without depositing any significant amounts of energy. The Muon Spectrometer (MS) [64] is the outermost subsystem of the ATLAS detector and is specialized for measuring muon momenta over a total $|\eta| < 2.7$ range. Like the ID, the MS measures momenta by bending the particles and measuring the radius of the curvature. This bending is produced by a set of toroid magnets, where a single large barrel toroid magnet generates a magnetic field in the $|\eta| < 1.4$ region. Smaller endcap magnets generate the magnetic field in the $1.6 < |\eta| < 2.7$ range. In the transition region, $1.4 < |\eta| < 1.6$, there is a reduction in the bending power where the two magnets overlap. This configuration of magnets generates a field that is primarily orthogonal to the trajectory of the muons while reducing the effects of multiple scattering from the muon traversing different materials. After deflection, a set of four subdetectors measures the muons: Monitored Drift Tubes (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC). These subsystems are labeled in Figure 4.9 and described briefly below.

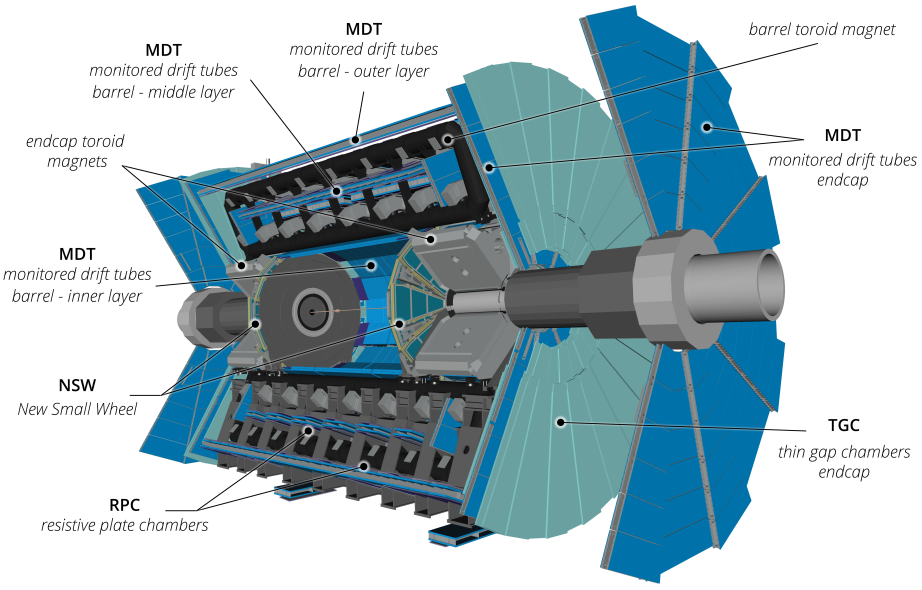


Figure 4.9: Labeled Cutaway of the ATLAS Muon System. A cutaway of the ATLAS muon system is shown with its main components labeled. There are three main components of the muon system. Monitored Drift Tubes (MDT) are used for precision tracking purposes. Thin Gap Chambers (TGC) and Resistive Plate Chambers (RPC) are used for triggering. Finally, the New Small Wheel (NSW) is the most recent major upgrade to the muon system placed at each of the endcaps closest to the interaction point. The NSW is primarily used for triggering and associating muons with the primary vertex. New detector technologies enable them to handle the increased rates in Run 3. Plot is taken from Ref. [52].

4.2.4.1 Monitored Drift Tubes

MDTs comprise most of the detectors in the $|\eta| < 2.7$ range. These function as cylindrical drift chambers, allowing for precision measurements of the trajectories of the deflected muons. A drift tube in the MDT is a cylindrical structure of various lengths ranging from 1 m to 6 m, with an outer diameter of 30 mm [65]. A gas mixture of Ar/CO_2 (93%/7%) fills each tube and is the source of the ionized particles produced by a passing muon. A potential difference is placed between the anode wire at the center of each tube and the outer walls. This potential difference causes the ionization to accelerate toward the wire (consistently building up more ionization as it travels). The basic building block of a MDT module, as shown in Figure 4.10, uses multiple layers of stacked drift tubes.

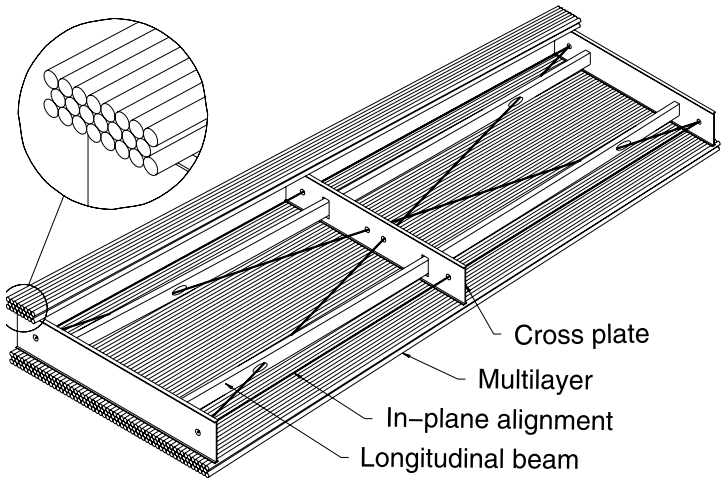


Figure 4.10: Module of the MDT in the ATLAS Muon System. A MDT module uses drift tubes and the ionization of the gas-filled tubes to measure the trajectory of a passing muon. A single layer of drift tubes only provides directionality in one axis, but multiple layers can produce a three-dimensional path. Plot is taken from Ref. [65].

Figure 4.11 illustrates a muon traversing three layers of drift tubes, inducing an ionization reaction in tubes crossed by the muon. The response time of a signal generated by a passing muon gives a measurement of the radial distance, $r(t)_{\text{tube}}$ in each of the activated tubes. Combining all three activated tubes gives a precise measurement of the trajectory in the plane of the tubes. The spatial resolution of the MDT is estimated by examining the residuals between the measured drift radius and the minimum distance between the fitted muon track and the wire, $d(\text{wire}, \text{track})$ [66].

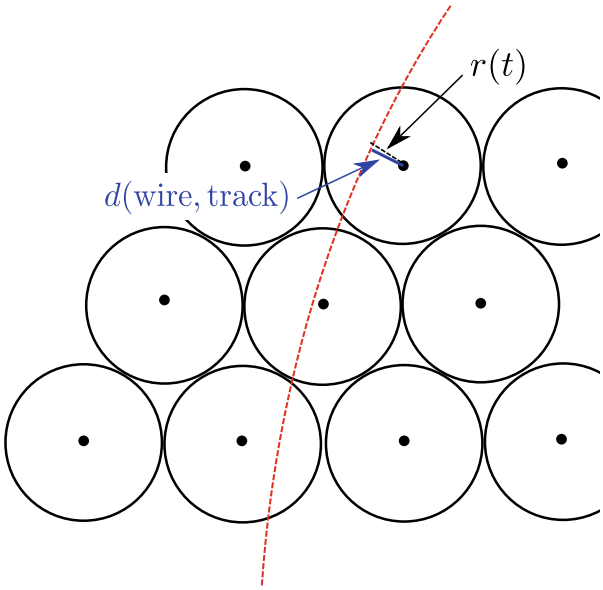


Figure 4.11: Illustration of a Fitted Muon Track Passing Through Several Drift Tubes. A muon passing through a tube filled with gas will generate ionization that will travel to an anode collection wire. Measuring the drift distance, d(wire, track) enables the reconstruction of the muon's trajectory through the drift tubes. This distance determines the spatial resolution of the drift chambers. Plot is taken from Ref. [66].

This muon tracking algorithm is identical across the 1088 chambers and 339 000 channels, which comprise the MDT system. The result is a a single-tube resolution of 80 μm and a chamber position resolution of 40 μm [67]. Extensive work in precisely aligning these large chambers provides the high level of precision required by the detector.

4.2.4.2 Cathode Strip Chambers

At the larger $2.0 < |\eta| < 2.7$ range, the counting rate of particles increases up to 1000 Hz/cm^2, which exceeds the optimal rate for the MDT system (150 Hz/cm^2) [68]. In this region, the CSC system replaces the MDT, where the former is more resistant to this increased rate while offering improved tracking resolution. The CSC is a multiwire proportional chamber, with cathodes organized into strips. The detection method of the multiwire chamber in the CSC is similar to the drift tube construction of the MDT. Detection occurs when a passing particle passes through a gaseous medium, producing an ionization shower. An anode wire then collects the products of this shower. However, in the case of the CSC, multiple wires are strung in the radial direction sandwiched between two sets of conducting plates. A similar gas mixture of Ar/CO_2 (80%/20%) fills the volume between the two cathode plates. ATLAS creates further segmentation by dividing the cathode plates into strips, with strips on one of the cathode plates running parallel to the direction of the wires and the strips on the other cathode plate running in the perpendicular direction. A cross-sectional view of the multiwire proportional chamber is shown in Figure 4.12.

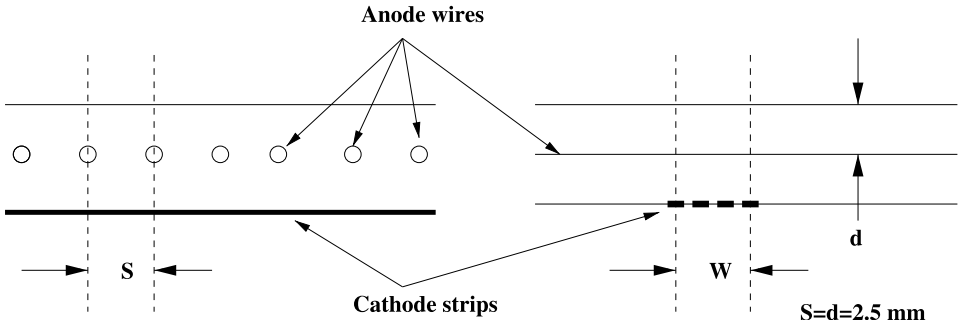


Figure 4.12: Cross-Sectional View of the Multiwire Proportional Chamber of the CSC. The CSC is a multiwire proportional chamber comprised of wires and strips running perpendicularly to each other. (**Left**) shows a cross-sectional view of the chamber with the wires running into/out of the plane of the page. (**Right**) shows the rotated view with the wires now running left/right in the plane of the page. This view shows the strips running into/out of the plane of the page. The wire pitch s is equal to the spacing between the anode and cathode, $s = 2.54$ mm. The spacing of the readout strips is set at $w = 5.08$ mm. This spacing generates a signal produced from a positive charge to be spread over the $3 - 5$ strips. Plot is taken from Ref. [4].

The positive charges created through ionization of the gaseous medium move towards the negative cathodes. The stripes running perpendicularly to the direction of the wires are more segmented than the strips running parallel to the wires. Therefore, the resolution in the perpendicular strip configuration is 60 μm, an improvement over the 80 μm in the MDT. For the strips running parallel to the wires, the segmentation is much coarser, resulting in a spatial resolution of 5 mm. The entire CSC system contains 32 such chambers, each containing four planes of CSC modules and a total of 31000 readout channels. Combining multiple layers enables the assignment of a precise $\eta - \phi$ coordinate to each track in a multi-track event.

4.2.4.3 Resistive Plate Chambers

The RPC system is the first of two triggering systems in the ATLAS muon system. RPC modules have a simple design and good spatial and time resolution. The RPC has a similar construction to the multiwire proportional chambers of the CSC but without any wires present. Two parallel plate electrodes create an electric field that collects the charge produced during an ionization event. A gas mixture of $C_2H_2F_4/C_4H_10/SF_6$ (94.7%/5%/0.3) fills the space between the electrodes and is the source of the ionized charges. Two planes of strips, oriented in perpendicular directions, measure both the transverse (ϕ) and the longitudinal (η) direction. Figure 4.13 shows a module in the RPC system.

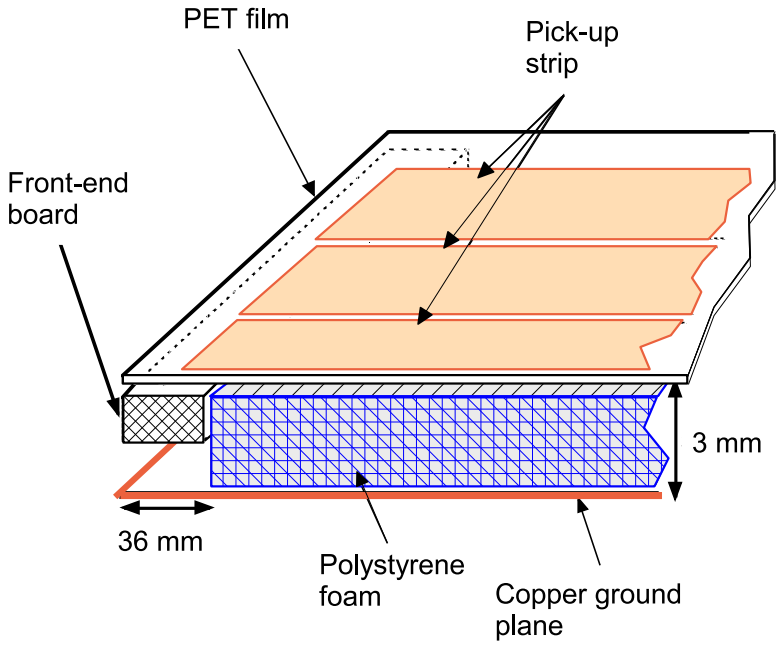


Figure 4.13: Module of the RPC in the ATLAS Muon System. A module of the RPC system contains a plane of readout strips. A second readout strip plane is oriented perpendicularly to the first plane to create a coincidence mechanism that can be used for triggering purposes in the barrel region of the muon system. Plot is taken from Ref. [4].

There are several requirements the trigger system must fulfill: rapid response time (small separation between plates to avoid long drift times), bunch-crossing identification, event discrimination based on muon transverse momentum, and acting as a complementing orthogonal measurement to the downstream precision trackers described above. In the barrel, $-1.05 < \eta < 1.3$, multiple modules of RPC modules are layers to satisfy these requirements while providing a complete 2π coverage. The placement of the RPC modules relative to the other triggering detectors in the muon system is shown in Figure 4.14.

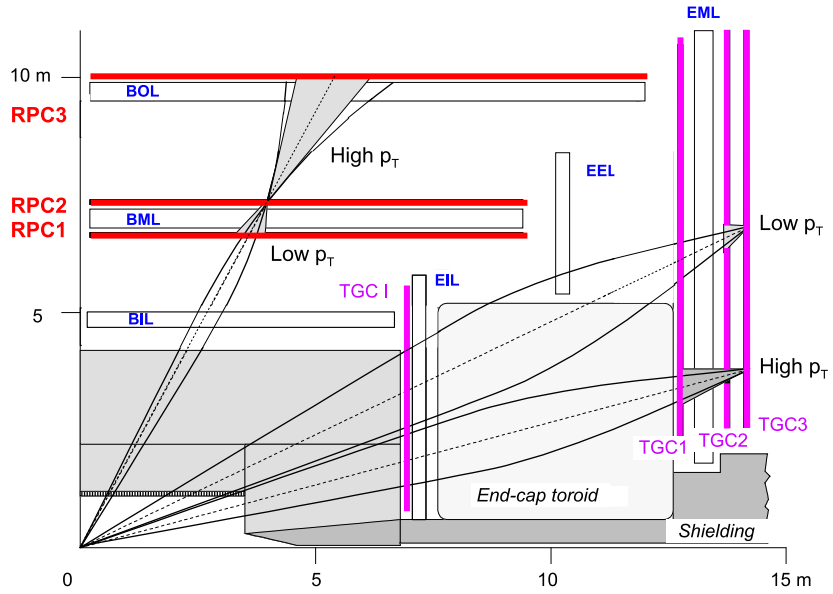


Figure 4.14: Layout of the RPC and TGC Modules in the Muon Trigger System. The muon trigger system is comprised of a combination of RPC and TGC modules. Multiple layers of each module type creates a coincidence system for a trigger that can operate with high efficiency. Plot is taken from Ref. [4].

One of the main tasks of a trigger is to reject fake tracks originating from noise or multiple scattering events unrelated to the initial hard scatter event. To create a veto system, two RPC plates sandwiching a MDT create a coincidence scheme in which a track must pass through multiple subdetectors, producing multiple $\eta - \phi$ measurements. Tracks originating from noise may only pass through a single RPC-MDT layer and would quickly be rejected by such a trigger system. Placing the RPC-MDT pairs in this coincidence configuration provides a means to separate low and high-p_{T} muons. A low-p_{T} ($6 - 9$ GeV) muon trigger is created by using a 3-out-of-4 coincidence in the RPC 1 and RPC 2 stations. A high-p_{T} ($9 - 35$ GeV) muon trigger is created by selecting a 1-out-of-2 coincidence in the outer RPC 3 stations [69].

4.2.4.4 Thin Gap Chambers

The second component of the muon trigger system is the TGC system. These are organized in groups to provide a means of vetoing tracks that do not originate from the primary collision event. In the endcap[6] ($1.05 < |\eta| < 1.92$), there are seven layers of TGCs, while the inner region ($1.92 < |\eta| < 2.4$) contains two layers of TGC modules. These modules are organized into units and mounted radially to form a wheel. The ATLAS detector mounts two such wheels, with the big EM wheel positioned farther away from the interaction point while the small wheel is placed much closer to the interaction point. Each module of the TGC system is constructed as a multiwire proportional chamber similar to the chambers installed in the CSC. The ionization source is a mixture of CO_2 and n-C_5H_{12} (n-pentane) gas.

[6] In the Muon Spectrometer, there are two distinguishable regions, the barrel ($|\eta| < 1$) and the end-cap ($|eta| > 1$).

A plane of anode wires collects the negatively charged ions, while readout strips lining the cathode boards collect the positive ions.

During the long shutdown after the completion of Run 2, the instantaneous luminosity of the LHC nearly doubled. One of the major challenges of coping with this luminosity was the increase in the muon trigger rate, which quickly surpassed the readout capabilities of the electronics used in Run 2. The current muon system was also very susceptible to fakes produced by unwanted interaction in the dead material between the small and big wheels. To address both of these problems, the New Small Wheel (NSW) [70] was installed during the long shutdown before the start of Run 3, replacing the muon system's current small wheel. The NSW uses two main detector technologies: small-strip Thin Gap Chambers (sTGC), and Micromegas. The sTGCs are similar to the multiwire proportional chambers previously installed but with a much smaller strip pitch, hence the name "small-strip". For Micromegas detectors, the wire plane characteristic of a multiwire chamber is replaced by a fine mesh which offers greater resolving two-dimensional power than a wire plane. The addition of the Micromegas technology greatly improves the precision of the NSW. The result is a muon trigger that can cope with the higher rates of Run 3 while reducing fakes without being detrimental to physics.

4.2.5 Trigger and Data Acquisition

Several reasons necessitate a dedicated Trigger and Data Acquisition (TDAQ) [71] system. Firstly, the data-readout rate limits the overall efficiency of the trigger. The readout of digitized data beyond 100 kHz will lead to large amounts of deadtime. At this point, storing the data becomes very costly. Secondly, during collisions, most events contain only soft QCD objects such as quarks and gluons. These events are characterized by minimal transverse momentum and are unlikely to contain larger mass SM particles. Events with these soft QCD signatures are generally not very useful for further analysis as they are unlikely to contain any interesting new physics. Another reason to discard these events is to reduce the rate of data entering the acquisition system, which will become quickly overwhelmed by the influx of data. A quick calculation reveals that for the nominal 40 MHz bunch crossing rate and a single event requiring approximately 25 MB of space, there would be approximately 1 petabyte of data streaming in every second. The goal of the TDAQ system is to reduce this rate to a reasonable rate that the offline data storage center can handle. The data flow from the detector to the offline data storage is shown in Figure 4.15.

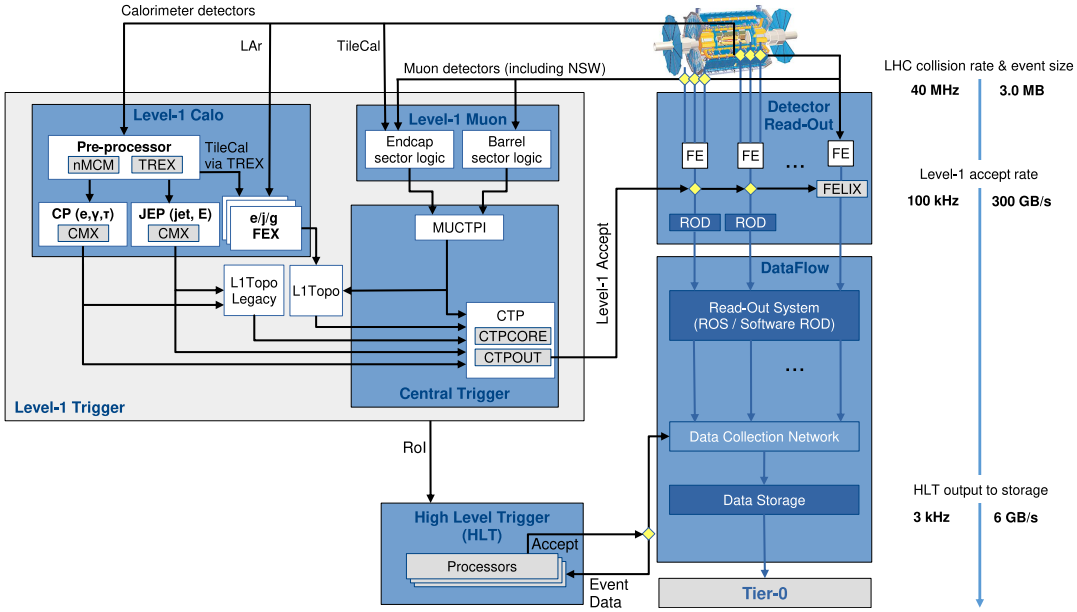


Figure 4.15: Block Diagram of the ATLAS Trigger and Data Acquisition System. The decision on whether to keep an event is made by a two-tiered trigger system. The L1 trigger system accepts input from the calorimeters and the muon detectors. The system rapidly reconstructs physics objects and checks their properties against the current trigger menu configurations. Successfully passed events and their regions of interest move onto the HLT. The HLT trigger applies more sophisticated selection algorithms to the physics objects. Events passing this final trigger level are saved in long-term storage for offline analysis. The result of such a system is a reduction of the data rate from 40 MHz down to 3 kHz. Plot is taken from Ref. [52].

A two-tier system, comprising a hardware-based Level-1 (L1) trigger and a software-based High-Level Trigger (HLT) system, makes the decision on which events are deemed interesting or not. A trigger chain evaluates the event and decides whether to keep or to discard it. A trigger chain refers to the set of selection criteria that, if passed, will save the event for further study. The set of all trigger chains operating during a run period is recorded in the trigger menu. An event in the ATLAS detector is retained and directed to offline storage if it passes any of the trigger chains listed in the trigger menu. A separate trigger menu exists for the L1 and HLT systems.

4.2.5.1 L1 Trigger

The first level, L1, receives the subset of trigger data flux from the various subsystems at the nominal 40 MHz rate. Decision-making relies on quickly reconstructing hits and clusters into various physics objects. These include high-p_{T} muons from the muon system, electrons/photons registered in the EM calorimeters, jets, and taus decaying into hadrons registered in the hadronic calorimeters. Another critical component of the trigger system is a precise calculation of the Missing Transverse Energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) from particles that leave no traces in the detector (e.g., neutrinos). L1 combines these objects and evaluates them against the list of trigger chains in the L1 trigger menu. After applying the L1 triggers, the rate reduces from the input 40 MHz rate down to a maximum of 100 kHz rate, which is then passed

onto the HLT.

4.2.5.2 *High-Level Trigger*

Successful events selected by the L1 trigger pass onto the HLT. The HLT accepts these events and applies complex software-based selection algorithms. Recall that along with the event data, regions of interest in the detector are also accepted by the HLT. By only applying these algorithms within the regions of interest, the HLT greatly reduces the computation requirements (computation time, bandwidth, and data readout) by not exerting any effort on regions with no interesting events. Within these regions of interest, the HLT has access to the full granularity of the detector to help make its decisions. The result is that the initial rate of 100 kHz received from the L1 trigger reduces to a final rate of $\mathcal{O}(100\text{ Hz})$. Events passing the HLT trigger move onto long-term storage for offline analysis.

CHAPTER 5

Data and Monte-Carlo Samples

Monte-Carlo (MC) simulations play an important role in many aspects of experimental high-energy physics. One of its main uses is to simulate the expected event yields of various physics processes according to SM theory. The decay products of the various pp collisions register as electronic signals in the ATLAS detector. These signal are reconstructed into physics objects such as electrons, muons, jets, and missing energy. When working with real-life detector events, it is impossible to determine which of the dozens of physics processes could have resulted in the same detector response. We differentiate between two types of processes: signal and background. Signal processes originate from the process of interest (in this case, $WWbb$ production), while background processes are all other processes that can mimic the signal process by inducing a similar response in the detector.

By simulating background and signal events, analysts better understand the proportions of each in their kinematic distributions. The kinematic distributions from MC are compared against real ATLAS data. Regions of phase-space displaying significant discrepancies between real ATLAS data and the predicted distributions are strong indicators of new physics. Theorists are particularly interested in these phase-spaces as potential targets requiring further tuning of their models. This chapter describes the MC process to generate the underlying signal and background events. A description of the various MC samples used in this analysis will be provided. The data samples correspond to the Run 2 data-taking period.

5.1 Monte-Carlo Event Generation

MC event generation is the preferred method many high-energy particle physics experiments employ to predict the expected yields of their kinematic distributions. In broad terms, MC experiments refer to the class of computational algorithms that uses random sampling to produce many numerical results to solve a problem that would otherwise require a complex calculation. The MC method is prevalent among particle physicists because the basic algorithms are built on a foundation of probability. Given that particle physics processes are probabilistic by nature (e.g., amplitudes, cross-sections, branching ratios, etc.), the MC method is well-suited for simulating these processes. The sequence of steps in the MC event generation chain is introduced below, with a more detailed description of each step to

follow.

Hard Subprocess The Hard Subprocess describes the initial hard-scatter interaction between constituents of the colliding particles. In the pp context, constituent quarks and gluons of colliding protons interact to produce other SM particles.

Parton Showering Partons carrying color produced during the Hard Subprocess, such as quarks and gluons, can radiate additional quarks and gluons. These additional particles can lead to a cascade of successive emissions known as a parton shower. Each successively emitted parton carries less and less energy.

Hadronization Partons undergo Hadronization upon reaching a critical energy during the parton showering step. At this point, the strong interaction coupling increases enough to bind individual partons into colorless hadrons.

Multiple Particle Interactions Besides the primary hard-scatter interaction, additional partons in the hadron can also interact. The additional partons can interact in secondary collisions concurrent with the primary interaction. These multiple parton interactions produce the underlying event.

Hadron Decay Following Hadronization, the hadrons produced are mainly unstable. This final step concerns the decay of these hadrons into their final state daughter particles.

The final daughter particles from this sequence pass through a simulated ATLAS detector. Generally, a particle's mass is inversely proportional to its decay time. Therefore, heavy particles such as top-quarks and W/Z-bosons will decay before reaching the detector. Stable particles, such as electrons, muons, and jets, produced in the initial interaction or from an intermediate decay, will reach the detector and leave a signal. The virtual signals from this simulated detector represent the best theoretical prediction of a physics event, which could also be observed in the real ATLAS detector. There are various general-purpose and specialized event generators on the market, characterized by different implementations of event generation. The three main programs used for simulating the hard-scatter process are SHERPA, PYTHIA, and HERWIG.

5.1.1 Cross-Section Calculations

Before any simulations, a calculation of the cross-section of the process of interest is performed. For example, one can calculate the top-quark pair production process via quark-antiquark annihilation ($q\bar{q} \rightarrow t\bar{t}$) following the Feynman rules from the QCD Lagrangian. From here, the scattering amplitudes, matrix elements, and cross-section for this event can all be computed. $2 \rightarrow 2$ (2 input partons and 2 output partons) processes, such as $q\bar{q} \rightarrow t\bar{t}$, are relatively simple to compute. When one includes more partons in the final state (e.g., $2 \rightarrow n$, $n > 2$), the cross-section calculation complexity increases. Further difficulties arise when one includes higher-order corrections to these processes, which may include intermediate self-interactions in the form of loops. As a result, the number of contributing diagrams increases exponentially as one moves to higher and higher orders, which is an analytical and

computational hurdle. The benefit of including these higher-order corrections is a far more accurate prediction. As shown previously in Figure 2.4, the addition of NLO corrections can drastically change the distribution of certain physical observables.

5.1.2 *Hard Subprocess*

The probability of an event occurring is directly proportional to the cross-section of the underlying process. Physicists compute these cross-sections perturbatively up to a particular order in perturbation theory. The cross-sections need to be derived for all physics processes of interest. The next step is to enter the cross-sections into the event generator to encode the production probabilities. Often, the cross-section depends on measurable quantities (i.e., mass, energy, momentum, decay angle, etc.). Therefore, in practice, a certain phase-space is defined using these quantities when making differential cross-section measurements. Integration of the quantities over the complete phase space gives the total cross-section value [72].

The hard subprocess defines the first step in event generation. In this step, interactions between partons of the simulated colliding hadrons form the basis of the hard process of interest. The momenta of the interacting partons are selected by drawing from known Parton Distribution Functions (PDF). These functions are probability densities that define the probability of finding the parton carrying a certain fraction of the total momentum of the hadron. Different collaborations have previously tabulated these PDFs. A commonly used PDF set comes from the NNPDF collaboration, which uses a neural-network approach. These PDFs are constrained using measurements from many different physics processes. The set of physics processes include deep-inelastic scattering, Drell-Yan processes, inclusive jet production, heavy quark production, and electroweak boson production, among others [73]. The interaction of any color-carrying partons can initiate the next step in the simulation, parton showering.

5.1.3 *Parton Showering*

When simulating events at the same energy scales as the LHC, each of the incoming partons carries a large momentum. The resulting hard subprocess event involves large momentum transfers to the outgoing partons, meaning the outgoing partons are substantially boosted relative to the lab frame. Color-carrying partons undergoing this rapid acceleration will emit QCD radiation as gluons. This radiation is the QCD analog of the corresponding QED Bremsstrahlung process, where an accelerating electron will radiate a photon. Unlike photons, a gluon carries color and can initiate a cascade of further radiation, creating a parton shower. Each radiated generation of partons carries less and less momentum than the previous generation of partons. Once the partons reach a critical energy, color confinement effects dominate, and parton showering stops.

If we consider the near-collinear splitting of a parton of type i into $j + k$, the differential cross-section of the $n + 1$ parton system ($d\sigma_{n+1}$) can be computed perturbatively at leading order [74].

$$d\sigma_{n+1} \approx d\sigma_n \frac{\alpha_s}{2\pi} \frac{d\theta^2}{\theta} dz d\phi \mathcal{P}_{ji}(z,\phi) \tag{5.1}$$

where $d\sigma_n$ is the n-parton differential cross-section, α_s is the strong coupling constant , θ is the polar angle, and ϕ is the azimuthal angle. Finally, $\mathcal{P}_{ji}$ is the splitting function that measures the probability of parton i with momentum fraction z to evolve into a parton of type j. Eq. 5.1 can be applied to each colored parton from the hard subprocess to iteratively generate a complete parton shower, where the values of z, θ, and ϕ are generated using MC for each step.

An important feature relevant to the progression of the parton shower is the evolution variable. This variable characterizes the progression of the parton shower over momentum scales. The choice of the evolution variable depends on the specific MC generator used, with the simplest variable being the virtuality (squared mass of the virtual partons in the shower, q^2). The evolution variable determines the momentum cutoff to terminate the parton showering. This point is the infrared cutoff, which is defined as the point where the evolution variable falls below the hadronization scale, $q^2 < Q_0^2 \sim 1$ GeV^2.

The evolution variable enters into the Sudakov form factor. This factor accounts for quantum loop (self-interaction) effects at the same order in perturbation theory. As shown below, it describes the probability of not undergoing splitting during evolution from the scale q_1^2 to q_2^2.

$$\Delta_i(q_1^2, q_2^2) = \exp\left\{ -\int_{q_2^2}^{q_1^2} \frac{dq^2}{q^2} \frac{\alpha_s}{2\pi} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} dz \int_0^{2\pi} d\phi \mathcal{P}_{ji}(z,\phi) \right\} \tag{5.2}$$

The bounds of integration in the Sudakov form factor describe the range in which the splitting is resolvable. Contributions with too little energy or at too small an angle are unresolvable and are not included in the parton shower. As written, the Sudakov form factor describes a Markov process where the characteristics of the next iteration depend only on the state of the current iteration. Starting at an initial state, Q^2, the next evolution variable q_1^2 is computed by solving the following equation.

$$\Delta_i(Q^2, q_1^2) = \mathcal{R}_1 \tag{5.3}$$

In the above, $\mathcal{R}_1$ is a random number uniformly generated in the interval $[0,1]$. This random number aims to introduce the inherent randomness characteristic of a parton shower. If the computed q_1^2 value is less than the infrared cutoff, Q_0^2, the splitting between partons is unresolvable and the showering terminates. Otherwise, the splitting proceeds as $i \to j + k$ and the Sudakov form factors are calculated identically for partons j and k, $\Delta_{j,k}(q_1^2, q_2^2) = \mathcal{R}_2$. This process continues until all splittings fall below the critical cutoff and are no longer resolvable. The values of z and ϕ at each splitting are selected using MC using the pre-determined probability distributions $\mathcal{P}_{ji}(z,\phi)$, assuming z falls inside the allowed ranges set in the Sudakov form factors.

The calculation of q^2 and z for each parton allows a near-complete reconstruction of the kinematics of the parton shower. In practice, modifications to the $\mathcal{P}_{ji}$ functions do not allow an analytical solution to the Sudakov integrals and different techniques must be used to overcome this hurdle. For example,

recent versions of PYTHIA and SHERPA use a dipole evolution model to simulate the parton showers. A dipole in this context is a pair of matching color-anticolor partons, and the radiation occurs as the dipole radiates away a parton. This model simplifies shower implementations and agrees well with real data [75].

5.1.3.1 *General-Purpose vs. ME/NLO+PS Event Generators*

At this point, it is necessary to point out that MC generators generally fall under two categories: General-Purpose Monte-Carlo (GPMC) generators, and Matrix-Element (ME)/Next-to-leading-Order (NLO) combined with parton shower (PS) matching generators. GPMCs, as their name implies, are among the most common and widely-used type of MC generators available in high-energy physics. GPMCs offer a complete simulation of the physics processes by implementing the complete MC event generation, including the hard-scattering process, parton showering, hadronization, and the underlying event.

On the other hand, the ME/NLO + PS generators begin with an extremely precise calculation of the matrix-elements of the various hard subprocesses. This results in a precise model for well-separated jet products but is less accurate when the jets are collinear and soft. These soft jets fall under the regime where parton showering models in GPMCs perform well. As discussed above, GPMCs simulate the parton showering using a perturbative approach. Therefore, to get the best of both worlds, a ME/NLO generator is first used to compute the matrix-elements of the hard subprocess. The products are then interfaced (matched) to a GPMC to perform the parton showering step. For example, in this analysis, the POWHEG BOX v2 generator is used to compute the matrix-elements of the various physics processes, but the parton showering and hadronization are handled by PYTHIA 8. A complete table of the MC samples used in this analysis is provided later in this chapter.

5.1.4 *Hadronization*

Upon reaching the infrared limit, the parton shower terminates as the perturbation theory that governs the evolution of the parton shower breaks down. The simulation enters a non-perturbative regime in which confinement effects dominate, and the final generation of partons forms the observed final state hadrons. Perturbation theory is not applicable in this regime, and developing an exact theory to describe hadronization has yet to be successful. There are two choices for modeling the hadronization. The Lund string model [76] connects quark-antiquark pairs by a string representing the color field between the quark and antiquark. The quark-antiquark pair can move apart, "stretching" the string. The string eventually breaks and creates a new quark-antiquark pair at the breakpoint. String formation continues until it is energetically-unfavorable to create another string and the existing quarks and antiquarks then form the color-neutral hadron. The second model is the Cluster model [77], which roughly assumes that a local cluster of quarks and antiquarks can be grouped into a colorless cluster. This cluster, known as a proto-hadron, then decays into the final state observed hadrons.

5.1.5 *Multiple Particle Interactions*

For events containing a hard subprocess event, additional hadronic activity are characterized by partons not participating in the initial hard-scatter interaction. These interactions do not generate the high transverse momentum objects characteristic of the hard-scatter. Softer interactions still contribute to the total momentum transfer of the event and are collectively described as the underlying event. One of the most common interactions relevant to hadron colliders is t-channel gluon-gluon elastic scattering, $gg \rightarrow gg$ (and similar variations). Models of the underlying event build upon this basic process. In the soft QCD regime, this process is modeled perturbatively. The t-channel propagator for this process is forced (almost) on-shell, resulting in the differential cross-sections to increase sharply at low energies [72].

$$d\sigma_{2\rightarrow 2} \propto \frac{dt}{t^2} \sim \frac{dp_{\mathrm{T}}^2}{p_{\mathrm{T}}^4} \tag{5.4}$$

This cross-section represents a single parton-parton interaction in a hadron-hadron scattering event. Each additional parton-parton interaction in the same hadron-hadron scattering event will contribute twice to the parton-parton cross-section $\sigma_{2\rightarrow 2}$ but only once to the hadron-hadron cross-section, σ_{tot}. If all of the parton-parton interactions are independent, then the $2 \rightarrow 2$ cross-section is expected to behave as follows.

$$\sigma_{2\rightarrow 2}(p_{\mathrm{T,min}}) = \langle n \rangle (p_{\mathrm{T,min}}) \sigma_{\mathrm{tot}} \tag{5.5}$$

The cross-section for parton-parton scattering scales according to the average number of parton-parton interactions, $\langle n \rangle (p_{\mathrm{T,min}})$. This number, in the case of hadron-hadron scattering, could be greater than one. Additionally, the parton-parton interaction must have a minimum transverse momentum, $p_{\mathrm{T}} > p_{\mathrm{T,min}}$ so that the parton-parton cross-section remains finite. The probability for n parton-parton scatterings follows a Poisson distribution.

$$\mathcal{P}_n = \langle n \rangle^n \left[\frac{e^{-\langle n \rangle}}{n!} \right] \tag{5.6}$$

As written, unitarity is not violated in the limit $p_{\mathrm{T,min}} \rightarrow 0$. Instead of the parton-parton cross-section diverging, the expected number of interactions in the collision diverges. There are several techniques to handle this divergence. Naturally, the total momentum of the interactions cannot exceed the momentum of the parent hadron. Therefore, a cutoff in the number of interactions suppresses events with high-n. PYTHIA orders the soft QCD interactions by p_{T}. Then, the parton distributions for each successive interaction are explicitly constructed so that the sum of the momentum fractions in each parton is never greater than unity. In HERWIG, the expected number of interactions according to Poisson statistics, $\langle n \rangle$, is used as an initial guess. Interactions continue to be generated until the energy-momentum conservation limit is reached.

The color screening effect is needed as a secondary factor to regulate the divergence. Screening

refers to a phenomenon where the interactions between partons are weakened or screened at low energy scales. At these scales, the result is that individual partons become indistinguishable. To account for color screening, the following factor is introduced in various MC generators.

$$\frac{\alpha_s^2(p_{\mathrm{T},0}^2 + p_{\mathrm{T}}^2)}{\alpha_s^2(p_{\mathrm{T}}^2)} \frac{p_{\mathrm{T}}^4}{(p_{\mathrm{T},0}^2 + p_{\mathrm{T}}^2)^2} \tag{5.7}$$

In the above, α_s represents the strong coupling constants, and $p_{\mathrm{T},0}$ is a free parameter that must be tuned to data. Theorists tune the strong coupling constants in their models to constrain the divergence.

5.1.6 *Hadron Decay*

The final stage of event generation involves the decay of any unstable final-stage hadrons created during the hadronization step. The PDG contains much of the information needed to accurately describe parton decays, decay widths, branching fractions, cross-sections, etc. However, it is not as trivial as simply keying in all the different parameters of the PDG. One must make several distinct choices regarding what to include in the simulation. Analysts must first define the set of hadrons to be generated. In doing so, one must be careful to avoid any biases toward certain parton flavors. As a result, all variations of spin and parity must be included for some flavor multiplets. When deciding which hadrons to include, another challenge arises from the difficulties of modeling properties of theoretical flavor states that have yet to be experimentally verified. Once the set of hadrons has been defined, it is then necessary to define the set of acceptable decays and how exactly to simulate these.

5.1.7 *Detector Simulation*

The decay products from the hadron decays pass through a simulated detector geometry implemented in GEANT4 [78]. The particles pass through the simulated detector, registering their energy deposits on simulated detector components. The data acquisition chain digitizes the signals and applies the reconstruction algorithms. The same reconstruction algorithms applied to data are also applied to MC to produce the final dataset. The result is a dataset that contains an identical set of information that would be output by the real-life ATLAS detector. From here, an analyst compares the distributions of the simulated physics observables against the same distributions obtained from real data and notes any significant discrepancies.

5.1.8 *Detector, Parton, Particle-Level*

These terms correspond to the types of objects one can expect to obtain from a complete MC dataset. A MC production allows analysts access to the initial list of generated particles and their daughters (also known as the truth record), which would not otherwise be possible with real detector data. The following is a description of each of the different terms.

Detector-Level This level refers to the set of physics objects registered by the detector after passing through the reconstruction chain. The data at detector-level provides the most accurate pre-

diction of the response that an identical response would produce in the real ATLAS detector. Detector-level objects are available in both MC and real data.

Particle-Level This level refers to the record of final state quasi-stable particles before entering the detector. This record includes many of the same objects found at detector-level such as jets (following hadronization of quarks and gluons), leptons, and photons. It is also possible to reconstruct parent objects at particle level where the same algorithms at detector-level are applied at the particle-level using particle-level objects. For example, the reconstruction of a particle-level W uses the same detector-level algorithms but uses particle-level objects as input.

Parton-Level This level refers to the theoretical description of SM processes for the first generation of daughters originating from the pp-collision. These mainly include many of the unstable elementary particles before decay. For example, in this analysis, parton-level objects may include top-quarks, bottom-quarks, and W-bosons produced immediately following the primary interaction.

There are several advantages to keeping a truth record during MC production: (1) allows opportunities to refine analysis algorithms while being able to compare to a "known solution", (2) maintains a record of different particle properties, which can then be used for later identification (e.g., parent/daughter particles), (3) facilitates the calculation of difficult to reconstruct objects such as missing energy and neutrinos which relies on complex algorithms in real data and (4) provides the opportunity to evaluate the detector response to that type of object through the use of a migration matrix.

5.1.9 Monte-Carlo Reweighting and Scaling

When comparing distributions between data and MC, the data events are always filled using a weight of 1. For MC events, the weight can be a continuous range of numbers and can even be negative. Several reweighting factors account for generator, detector, and reconstruction effects, which often differ between data and simulation[1].

Generator Weight ($w_{\mathbf{gen}}$) This weight varies between MC generators and is applied to correct for various generator effects. Corrections may include removing double-counted events, accounting for QCD corrections at higher-orders in the calculation, and adjusting the simulated events' overall normalization to match measured cross-sections

Pileup Reweighting Weight ($w_{\mathbf{prw}}$) This weight is applied to account for the effects of pileup interactions in the simulation. The weight ensures the distribution of pileup events in the simulation accurately reflects the same conditions in the real ATLAS detector.

Lepton Scale-Factor Weight (w_{ℓ}) This SF weight accounts for various efficiencies associated with the reconstruction of leptons in the simulation.

[1] During event generation, certain simulation characteristics can differ significantly from the observed or expected behavior. To adjust the simulated events to match experimental observations better, a Scale-Factor (SF) is derived. The SF is computed as the ratio of the efficiency in data (ϵ_{data}) to the efficiency in simulation (ϵ_{MC}), $SF = \epsilon_{\text{data}}/\epsilon_{\text{MC}}$.

Jet Vertex tagger Weight (w_{jvt}) This weight accounts for inefficiencies associated with the Jet Vertex Tagger (JVT)[2] discriminant.

***b*-tagging Scale-Factor Weight (w_{btag})** The b-tagging algorithm labels certain jets as originating from b-quarks or other lighter-mass quarks. This SF weight is applied to account for differences in b-tagging efficiencies between data and simulation.

The total event weight is the product of these individual factors.

$$w_{\text{total}} = w_{\text{gen}} \times w_{\text{prw}} \times w_{\ell} \times w_{\text{jvt}} \times w_{\text{btag}} \tag{5.8}$$

In addition to the event-by-event weights, the contents of each bin must be scaled to the same integrated luminosity of the data sample ($\mathcal{L}_{\text{data}}$). This second multiplicative factor (M) is again a combination of different factors.

$$M = \frac{\mathcal{L}_{\text{data}}}{\mathcal{L}_{\text{MC}}} \tag{5.9a}$$

$$\mathcal{L}_{\text{MC}} = \frac{\sum_i W_i}{\sigma \cdot \epsilon \cdot k} \tag{5.9b}$$

The individual factors in the above equation are described as follows.

Total Sum of Weights ($\sum_i W_i$) The total sum of weights is a normalization factor that counts the total (weighted) number of generated events. To compare against other MC generators, the total number of weighted events is often not a useful quantity and is therefore divided away.

Cross-Section (σ) This variable refers to the generator-specific cross-section value for the process of interest. These values vary from generator to generator because of differences in theoretical models, algorithms, input parameters, and handling of non-perturbative effects such as multiple-parton interactions.

Generator Filter Efficiency (ϵ) This efficiency factor is associated with filters applied during event generation. If the generator produced events with no specific biases, this factor will be 1. However, filters will sometimes be applied to remove events that do not pass certain criteria or are otherwise uninteresting for further analysis.

***k*-Factor (k)** This factor accounts for any higher-order processes omitted in the theoretical calculations. During event generation, a certain accuracy in perturbative QCD is assumed for theoretical purposes. However, for processes where higher-order QCD terms are more impactful, the k-factor is used to correct from NLO to Next-to-Next Leading Order (NNLO). Like the generator weight

[2] The Jet Vertex Tagger is designed to improve the identification and rejection of jets originating from pileup. It does this by considering the fraction of tracks associated with the jet coming from the primary vertex. Then, it combines this with additional information, such as the scalar p_{T} sum of the tracks associated with the jet. These variables are entered into a likelihood algorithm to construct the JVT discriminant. See Ref. [79] for more details.

described above, the k-factor may or may not be unity depending on the specific MC generator used.

5.1.9.1 *Statistical Uncertainties*

The statistical uncertainty on each histogram bin depends on whether the histogram is filled with weighted or unweighted events. Assuming that each event is independent, the height of a bin is distributed according to Poisson statistics. Given bin i with height h_i, the uncertainty on the height of bin i (σ_i) will be the standard deviation of a Poisson distribution with an expected value of h_i, $\sigma_i = \sqrt{h_i}$. The Poisson uncertainty applies for unweighted events, as is the case when filling histograms with real data. For weighted events, the uncertainty on the height of bin i instead follows Binomial statistics. The uncertainty on the height of bin i is computed by taking the square root of the sum of the weights squared, $\sigma_i = \sqrt{\sum_j w_j^2}$ where the inner sum runs over all events j that fall into bin i.

5.2 Data Samples

This analysis uses the full Run 2 dataset, which collected 140 fb^{-1}. During normal LHC operations, certain issues can arise which can negatively affect the performance of the ATLAS detector. The Good Runs List (GRL) records luminosity blocks that are deemed acceptable for analyses for each active data-taking year. Table 5.1 lists the integrated luminosity and GRLs for each data-taking year.

Table 5.1: Run 2 Integrated Luminosity and GRL Summary. The integrated luminosity collected and GRL are shown for each year of the Run 2 data-taking period. The uncertainties correspond to the total uncertainty on the measured integrated luminosity figures.

Year	Int. Lumi. (fb^{-1})	GRL
2015	3.24 ± 0.04	`data15_13TeV/20170619/physics_25ns_21.0.19.xml`
2016	33.40 ± 0.30	`data16_13TeV/20170605/physics_25ns_21.0.19.xml`
2017	44.63 ± 0.50	`data17_13TeV/20180619/physics_25ns_Triggerno17e33prim.xml`
2018	58.79 ± 0.64	`data18_13TeV/20181111/physics_25ns_Triggerno17e33prim.xml`
Run 2	$\mathbf{140.07 \pm 1.17}$	

5.3 Monte-Carlo Samples

This section describes the dominant MC processes contributing to the $WWbb$ final state. There are three categories of sample types: (1) the nominal signal samples constitute the primary events of interest, (2) the alternative signal samples used to evaluate modeling systematic uncertainties and comparing with data, and (3) the nominal background samples whose events can potentially mimic the characteristics of signal events but are not the main events of interest. Table 5.2 lists the characteristics for each MC sample.

Table 5.2: Summary of Monte-Carlo Samples used in the Analysis. The various Monte-Carlo samples used to simulate the dominant physics processes are listed. The MC samples are organized into nominal, alternative, and background samples. The nominal signal and background samples are used to perform the main analysis such as filling control distributions and migration matrices used to perform the unfolding. The alternative samples are included in the final unfolded differential cross-sections as comparisons to data. The first column lists the dominant physics processes considered in the analysis. The second column corresponds to the labels used to refer to the different samples in the plots in the remainder of the document. The remaining columns list the generators and a selection of various defining parameters used in the simulation of each sample. These columns are briefly introduced after the table.

Nominal Signal Samples						
Physics Process	Legend Entry(s)	Generator	PDF	Parton Shower & Hadronization	Tune	σ Norm. [pb]
$t\bar{t}$	PH+PYTHIA 8	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	832^{+41}_{-51}
tW (DR, DS)	PH+PYTHIA 8 (DR, DS)	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	71.7 ± 3.8
Alternative Signal Samples						
Physics Process	Legend Entry(s)	Generator	PDF	Parton Shower & Hadronization	Tune	σ Norm. [pb]
$t\bar{t}$	aMCNLO+PYTHIA 8	MADGRAPH 5 _aMC@NLO 2.3.3	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	832^{+41}_{-51}
	PH+HERWIG 7.1	POWHEG BOX v2	NNPDF3.0NLO	HERWIG 7.1.6, MMHT2014LO	H7.1-Default	832^{+41}_{-51}
	MiNNLO	POWHEG BOX v2	NNPDF3.0NNLO	PYTHIA 8.230, NNPDF2.3LO	A14	832^{+41}_{-51}
tW (DR, DS)	aMCNLO+PYTHIA 8 (DR, DS)	MADGRAPH 5 _aMC@NLO 2.3.3	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	71.7 ± 3.8
	PH+HERWIG 7.1 (DR, DS)	POWHEG BOX v2	NNPDF3.0NLO	HERWIG 7.1.6, MMHT2014LO	H7.1-Default	71.7 ± 3.8
Background Samples						
Physics Process	Legend Entry(s)	Generator	PDF	Parton Shower & Hadronization	Tune	σ Norm. [pb]
W+jets	W+jets	SHERPA 2.2.11	NNPDF3.0NNLO	SHERPA	SHERPA	20100 ± 1000
Z+jets	Other V/VV	SHERPA 2.2.11	NNPDF3.0NNLO	SHERPA	SHERPA	1906 ± 95
Diboson (VV)	Other V/VV	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA	SHERPA	211 ± 8
s-channel single-top (no tW)	Other t	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	$6.35^{+0.23}_{-0.20}$
s-channel single-anti-top (no tW)	Other t	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	$3.97^{+0.19}_{-0.17}$
t-channel single-top (no tW)	Other t	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	$134.2^{+2.6}_{-1.7}$
t-channel single-anti-top (no tW)	Other t	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	$80.0^{+1.8}_{-1.4}$
$t\bar{t}V$	Other t	MADGRAPH 5 _aMC@NLO 2.3.3	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	1.5 ± 0.1
$t\bar{t}H$	Other t	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230, NNPDF2.3LO	A14	0.51 ± 0.05

The following is a description of the columns in Table 5.2.

Physics Process The particular physics process the sample is modeled after.

Legend Entry(s) The label(s) that will be used to refer to each MC sample later in the document. The same physics process may have different legend entries corresponding to different generators.

Generator The program used to model the various processes and compute the corresponding matrix elements. In this analysis the GPMC generators include SHERPA [80], PYTHIA [81], and HERWIG [82, 83]. The ME/NLO generators include POWHEG BOX v2 [84–87] and MADGRAPH 5 _aMC@NLO 2.3.3 [88]. The MiNNLO$_{ps}$ sample [89] combines for the first time QCD corrections at NNLO for the $t\bar{t}$ process and parton shower models implemented in PYTHIA. This approach has been validated to agree well with ATLAS and CMS data for various $t\bar{t}$ decay modes [90]. The resulting corrections are in the order of $\mathcal{O}(10\%)$ relative to the NLO predictions.

PDF The PDF set used as input into the generator. The accuracy of the PDF set is also identified in its name. As discussed previously, various collaborations make the PDF sets available. The PDF sets used in this analysis include NNPDF3.0NLO and NNPDF3.0NNLO [91].

Parton Shower & Hadronization The program which models and simulates the parton showering and hadronization processes. If available, the PDF set used for this step of the simulation is also listed. The following generators are used in this analysis to simulate the parton showering and hadronization: PYTHIA 8.230, HERWIG, and the default SHERPA parton showering model [92]. As input to the parton showering, the following sets of PDFs are used: NNPDF2.3LO [93] and MMHT2014LO [94].

Tune The tune corresponds to the set of parameters to control the various steps of the MC generation process. A specific tune is chosen for parameters that cannot be analytically calculated (generally non-perturbative) to ensure better agreement between simulation and data. The set of generator tunes includes A14 [95], H7.1-Default [96] and the default SHERPA tune.

σ Norm. [pb] The generation of MC events needs to be normalized to the cross-sections expected at a certain center-of-mass energy. The values in this column correspond to cross-sections expected at the Run 2 center-of-mass energy of $\sqrt{s} = 13$ TeV. These values are extracted from Refs. [97–101].

5.4 Fixed vs. Dynamical Scaling

A commonly tuned parameter associated with the hard-scatter simulation is the POWHEG h_{damp} parameter. The definition for this parameter is $h_{\text{damp}} = h \cdot m_{\text{top}}$ with the h values ranging from $0.5 - 3.0$. h_{damp} is associated with the computation of the NLO QCD matrix elements and matching with the subsequent parton showers [102]. This analysis uses a nominal h_{damp} parameter, $h_{\text{damp}} = 1.5m_t$.

The h_{damp} parameter is often associated with the tuneable renormalization and factorization scale parameters, μ_{r} and μ_{f}, respectively. These scales are introduced into perturbative QCD calculations to

handle certain types of divergences that arise when calculating the cross-section matrix elements. In practice, the μ_r and μ_f scales are used to estimate the uncertainties associated with the modeling and generation of Initial State Radiation (ISR) and Final State Radiation (FSR). To evaluate the effects of renormalization and factorization, μ_r and μ_f are varied by 0.5 or 2.0 times the default value (μ_0) [102].

The choice of the default value of the μ_r and μ_f scales can affect the final (top-quark) cross-section calculations [103]. The μ_r and μ_f scales can be fixed or dynamically scaled. Nominally, the default value for μ_r and μ_f is fixed to the mass of the top quark ($m_t = 172.5$ GeV). However, one can also dynamically set the value of μ_r and μ_f depending on the kinematics of the process being simulated. Eqs. 5.10a–5.10d lists several functional forms for the dynamical scale.

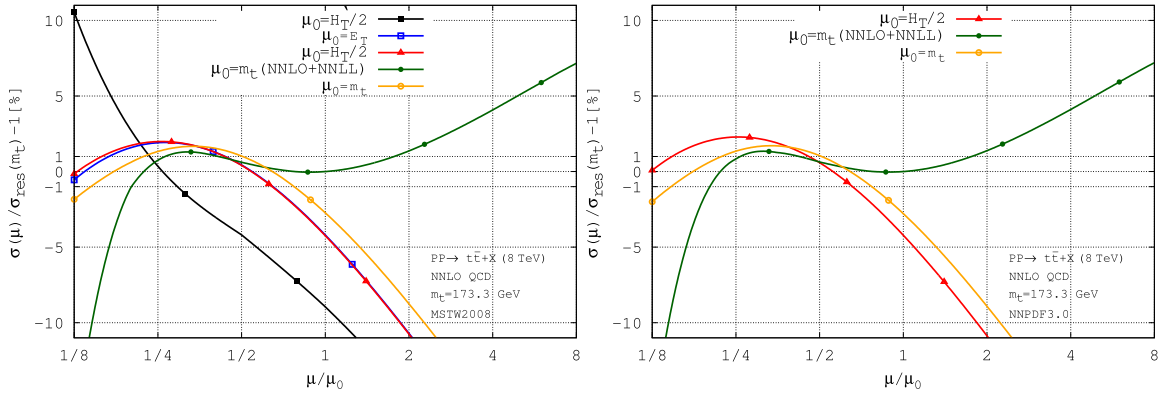
$$\mu_0 \sim m_{\mathrm{T}} = \sqrt{m_t^2 + p_{\mathrm{T}}^2} \tag{5.10a}$$

$$\mu_0 \sim H_{\mathrm{T}} = \sqrt{m_t^2 + p_{\mathrm{T},t}^2} + \sqrt{m_t^2 + p_{\mathrm{T},\bar{t}}^2} \tag{5.10b}$$

$$\mu_0 \sim H'_{\mathrm{T}} = \sqrt{m_t^2 + p_{\mathrm{T},t}^2} + \sqrt{m_t^2 + p_{\mathrm{T},\bar{t}}^2} + \sum_i p_{\mathrm{T},i} \tag{5.10c}$$

$$\mu_0 \sim E_{\mathrm{T}} = \sqrt{\sqrt{m_t^2 + p_{\mathrm{T},t}^2}\sqrt{m_t^2 + p_{\mathrm{T},\bar{t}}^2}} \tag{5.10d}$$

In the above, p_{T} is the momentum of the top or the anti-top while the sum in Eq. 5.10c runs over all massless partons in the final state. Figure 5.1 plots the total top-quark pair inclusive cross-sections evaluated with the different dynamic scales and with two different PDF sets, MSTW2008NNLO [104] and NNPDF3.0NNLO [91].



(a) MSTW2008NNLO. **(b)** NNPDF3.0NNLO.

Figure 5.1: Total $t\bar{t}$ Cross-Sections Computed with Different Dynamic Scales and PDF Sets. The total cross-section for $t\bar{t}$ production is computed by varying the default values of different dynamic scales and two separate PDF sets, **(a)** MSTW2008NNLO and **(b)** NNPDF3.0NNLO. Each curve has been normalized to the NNLO + NNLL cross-section computed using the respective PDF set and the fixed $\mu_0 = m_t$. The plot is used to evaluate the convergence speed for the different choices of dynamic scales. Plot is taken from Ref. [103].

The authors of Ref. [103] emphasize that the best choice of the dynamic scale is the one that pro-

duces the fastest perturbative convergence to the total cross-section computed with the fixed scale (at NNLO+NNLL). In this context, convergence is defined as the difference between the total cross-section computed with the dynamic scale and the total cross-section computed with the reference scale. Using this metric to evaluate the different dynamic scales yields the best choice of scale as $\mu_0 = m_{\mathrm{T}}$ and $\mu_0 = H_{\mathrm{T}}/2$, which occurs at approximately $\mu = \mu_0/2$. In this analysis, the dynamic scale shown in Eq. 5.10a is used to produce the single-top DR and DS samples. In later plots, both the fixed and dynamic scales are shown, referred to in the legend as *tW (DR/DS, fix.)* and *tW (DR/DS)* respectively.

CHAPTER 6

Object Reconstruction and Definition

The main goal of the ATLAS detector is to collect all information associated with the proton-proton collision event. The collision products appear as electronic signals in each detector subsystem. The detector's job is to translate these electronic signals into their corresponding physics objects. These baseline objects and their kinematics form the foundation for this and many other analyses. This chapter will describe the methods through which ATLAS reconstructs physics objects on an event-by-event basis. There will also be a description of the reconstruction of the leptonic and hadronic W-bosons, which are a particular focus of this analysis. Finally, we describe the fake-factor method as the preferred method for estimating the background due to fake leptons.

6.1 Baseline Objects Reconstruction

The ATLAS reconstruction chain is a step-by-step process that transforms the raw electronic signals from the millions of read-out channels into meaningful information about the particles produced in the collision event. The reconstruction software identifies clusters of hits that are likely associated with the passage of a particle. During track reconstruction, multiple clusters are aligned to reconstruct the trajectory of a particle as it moves through the detector. Next, vertex reconstruction extrapolates particle trajectories backward to a common point in space. This process defines the primary vertices originating from the hard-scatter event and secondary vertices from intermediate particle decays. Particle identification combines all this information to label particles (electrons, muons, hadrons, etc.). In its most basic form, particle identification uses the fundamental characteristics to deduce the most likely particle candidate. Even more sophisticated algorithms are applied to differentiate between closely related particles. Figure 6.1 plots the stability of various SM particles. These variables are used extensively in particle identification.

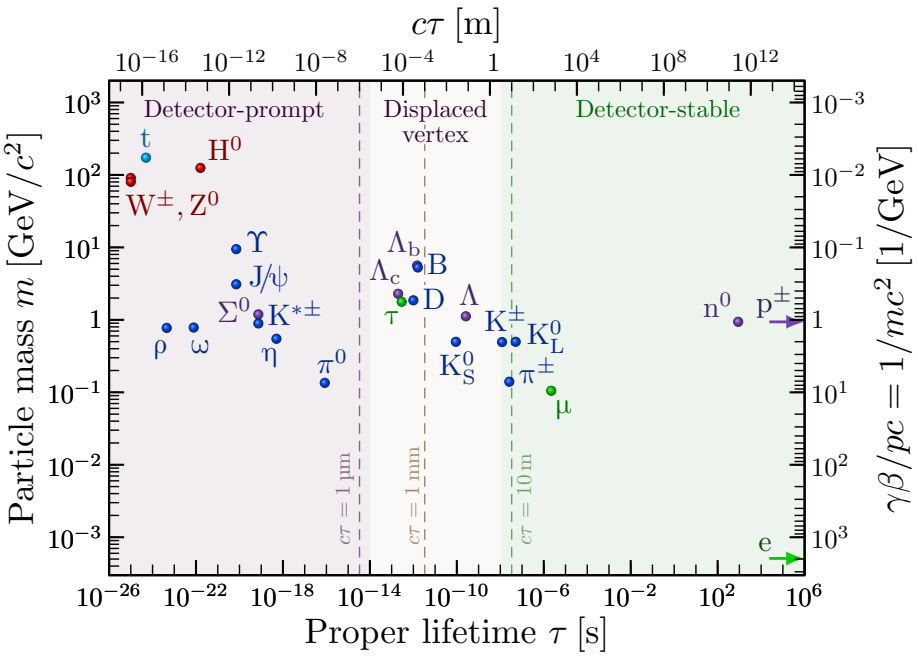


Figure 6.1: Mass, Lifetime and Stability of Various Standard Model Particles. Particle identification algorithms rely on accurate measurements of fundamental properties to accurately label particles. These properties are plotted for various Standard Model particles as a function of the mass (m), proper lifetime (τ), the average distance traveled ($c\tau$), and a dimensionless relativistic quantity ($\gamma\beta/pc$). The particles are further categorized by their stability and "promptness" of their decay in the detector. This property is especially useful for identifying decaying B-mesons during jet-tagging. Plot is taken from Ref. [20].

This section describes the identification and reconstruction of the physics objects relevant to this analysis. For each baseline object, the specific kinematic and quality requirements are also listed[1].

6.1.1 *Primary Vertex*

In the context of the collision, a vertex refers to the point from which a set of tracks originate. There can be many vertices in the same event. However, the primary vertex is defined as the vertex carrying the largest sum of squared transverse momenta of the tracks associated with it, $\sum p_{\mathrm{T}}^2$. A primary vertex (with sufficient total transverse momentum) in the event is often enforced to minimize contributions from pileup and other non-collision events. In ATLAS, the procedure for vertex reconstruction is an iterative process beginning by identifying a set of viable particle trajectories [105]. The first iteration fits the tracks to the best vertex position, after which any incompatible tracks (tracks not associated with the specific vertex) are removed, and a new best vertex position is calculated. The next iteration proceeds identically with this new best vertex position, removing incompatible tracks. This process

[1]The requirements described for each baseline object are based on Run 2 recommendations from the various ATLAS Combined Performance and Trigger groups listed in https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasPhysicsRun2Planning2019 and https://twiki.cern.ch/twiki/bin/view/Atlas/TriggerRecommendationsForAnalysisGroupsFullRun2.

continues until no unassociated tracks remain in the event or no additional vertices can be found. The primary vertex is the vertex with the largest sum of transverse momentum.

6.1.1.1 *Primary Vertex Requirements*

We require one primary vertex with at least two associated tracks each with transverse momentum greater than 500 MeV.

6.1.2 *Electrons*

ATLAS reconstructs an electron candidate by combining clusters of cells in the EM calorimeter with tracks from the Inner Detector [106]. Electron candidates are then calibrated to account for specific detector effects applied at the time of reconstruction. To accommodate many varieties of analyses, the electrons are then grouped into various working points identified by stringency of the selection criteria, also known as the identification and isolation criteria.

6.1.2.1 *Electron Reconstruction*

The first step in electron reconstruction is the topological cell clustering algorithm. This algorithm generates a set of topo-clusters: a group of calorimeter cells with sufficiently high energy. It starts with a seed cell and iteratively adds neighboring cells if they exceed a significance threshold ς_{cell}[2], defined as the ratio of the energy of the cell (E_{cell}) to the noise value in the cell $\sigma_{\text{noise,cell}}$. The usefulness of topo-clusters extends beyond electron reconstruction. Topo-clusters are also used in the reconstruction of other objects which interact with the calorimeters, such as photons and missing energy ($E_{\text{T}}^{\text{miss}}$).

$$\varsigma_{\text{cell}} = \left| \frac{E_{\text{cell}}}{\sigma_{\text{noise,cell}}} \right| \tag{6.1}$$

At the same time the topo-clusters are being formed, the track reconstruction algorithm runs to build tracks using hits from the Inner Detector [107]. Charged particle interactions in the pixel and SCT detectors register as clusters from which three-dimensional measurements known as space-points can be created. Multiple space-points are then combined to form a track seed which can be fit with a known model (pion hypothesis) for energy loss within detector materials. The track can fail the fit to the pion hypothesis, in which case the solution is to fall back on a model which gives more room for energy loss due to bremsstrahlung. To better account for energy losses in the material, a Gaussian-Sum Filter (GSF) [108] is applied. The GSF uses a weighted sum of Gaussian functions to describe the material-induced losses. Successive losses in energy cause the charged particle's radius of curvature to decrease. The GSF method improves track parameters relevant to the bending plane. The result is a GSF-track candidate from a passing charged particle.

[2]The absolute value of the cell significance is used because it is possible for cell signals to be negative. These negative cell signals are the result of fluctuations arising from pileup and electronic noise. However, cells with negative signals can still enter into a topo-cluster. These cells improve noise suppression by balancing out random positive (upward) noise fluctuations.

After identifying a set of topo-clusters and tracks, the matching between clusters and tracks can proceed. A cluster-track combination is considered matched if the following spatial conditions are satisfied: $|\eta_{\text{cluster}} - \eta_{\text{track}}| < 0.05$ and $-0.10 < -q \times (\phi_{\text{cluster}} - \phi_{\text{track}}) < 0.05$, where q is the charge of the particle[3]. These requirements may be satisfied by multiple tracks. The system selects the "best" track based on detector conditions, such as the number of hits in the various Inner Detector layers. This last step completes the electron reconstruction process. Figure 6.2 illustrates the path of a hypothetical electron passing through the detector.

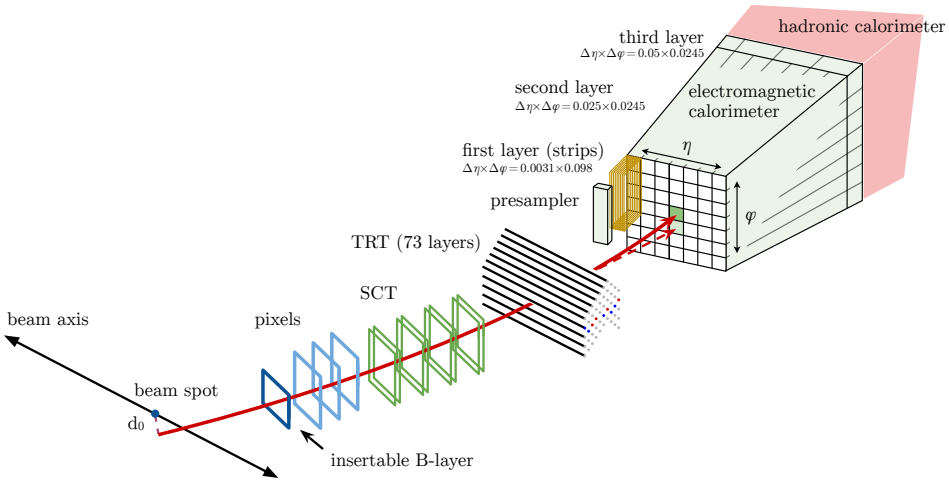


Figure 6.2: Illustrated Path of an Electron Through the ATLAS Detector. ATLAS reconstructs electrons by combining tracks in the ID and clusters in the EM calorimeter. The electron candidate first passes through the Inner Detector before depositing its energy in the high-granular calorimeter. The solid red line indicates the path of an electron candidate. The dashed red line indicates the path of a possible photon that could be produced from interactions of the electron with the detector materials. Plot is taken from Ref. [109].

6.1.2.2 Electron Calibration

The identified electron candidates are then calibrated to account for specific detector and reconstruction effects. Correction factors are derived using simulation and then applied to data [110, 111]. The following effects must be corrected for to properly calibrate an electron (or photon).

Calorimeter Layer Differences The calorimeter is segmented differently between the various layers, resulting in different energy responses in each layer. Therefore, different calibration factors must be applied to each calorimeter layer. The calibration factors are derived using muons coming from $Z \to \mu\mu$ decays. Muons are used because they pass through the passive material in front of the EM calorimeter. The muon's path through the calorimeter is extrapolated by using the muon's track measured in the ID. The muon energy response in each layer can be modeled by the convolution of a Landau distribution and a noise distribution. A fit is applied to this distribution

[3]The second condition is asymmetric to mitigate the effects of energy loss due to bremsstrahlung, and the potential mismatches in energy between the cluster and the track.

to extract the mean of the deposited energy ($\langle E \rangle$). This process can be repeated in data and in MC to derive a relative calibration factor between the two (largest) calorimeter layers, $\alpha_{1/2} = \left(\langle E_1 \rangle^{\text{data}} / \langle E_1 \rangle^{\text{MC}}\right) / \left(\langle E_2 \rangle^{\text{data}} / \langle E_2 \rangle^{\text{MC}}\right)$.

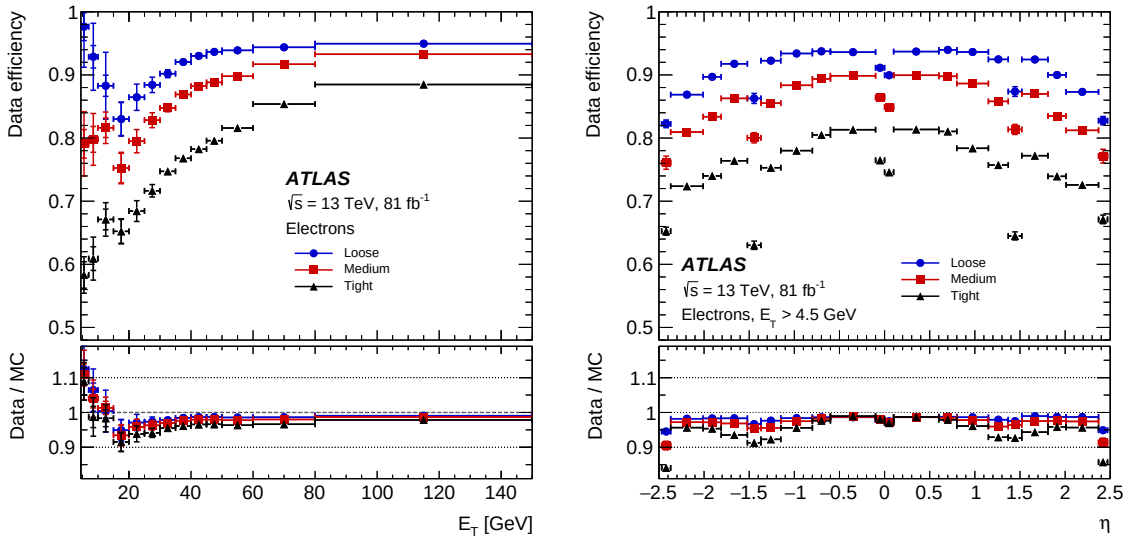
Electron/Photon Energy Estimates The electron/photon energy is computed from energy deposits in the EM calorimeter. This energy must be corrected to account for energy deposits/losses in unexpected areas of the detector. This can include correcting for energy losses in the passive material before the calorimeter, energy deposits in cells neighboring the central cluster, and energy deposited beyond the LAr calorimeter. A multivariate approach is used to derive a calibration factor to account for these effects. A boosted decision tree is trained using the relative energy deposits in each calorimeter layer and the location of the shower barycenters. This trained model is applied to simulation and data to measure the relative energy estimates.

Detector Non-Uniformity Certain localized non-uniformities can appear in specific calorimeter areas. These include non-optimal high voltage, azimuthal non-uniformities, or biases in the calorimeter's electronics calibration. The first effect arises in certain small areas of the calorimeter where a non-nominal high voltage value is set to account for short-circuits in the LAr gaps. The correction factors for this effect are typically 1% to 7%. The second effect is mechanical in nature due to gravity. Gaps between absorbers at the top of the detector are larger than gaps found at the bottom. The correction factor for this effect is $\lesssim 2\%$. The final effect results from a non-linear electronic response to the electron's energy during signal digitization. The electron's energy can differ when running the electronics at different readout gains. A correction factor comparing the high and medium gains is computed to account for this. The correction factor for this effect is 0.1% in the barrel and 0.4% in the endcaps.

Overall Energy Scale After correcting for the effects described above, there may still be a difference in the energy scales between data and MC. To correct this, a sample of $Z \to ee$ events is used to measure the overall calorimeter response. A distribution of the invariant mass of the Z is first constructed in both MC and data. Then, estimates of the correction factors are obtained by minimizing the χ^2 between data and MC. The derived correction factors are typically between $0.001\% - 0.01\%$.

6.1.2.3 Electron Identification

To maximize flexibility, ATLAS reconstructs electrons with various identification efficiencies. The identification efficiency is defined as the number of actual electrons correctly identified out of the total number of generated electrons. These identification efficiencies correspond to working points: *Loose*, *Medium*, and *Tight*. Electrons satisfying each tighter working point is a subset of the electrons in the previous working points. Figure 6.3 plots the different identification working points as a function of the electron candidate's E_{T} and η.



(a) Electron identification efficiency as a function of E_{T} for different working points.

(b) Electron identification efficiency as a function of η for different working points.

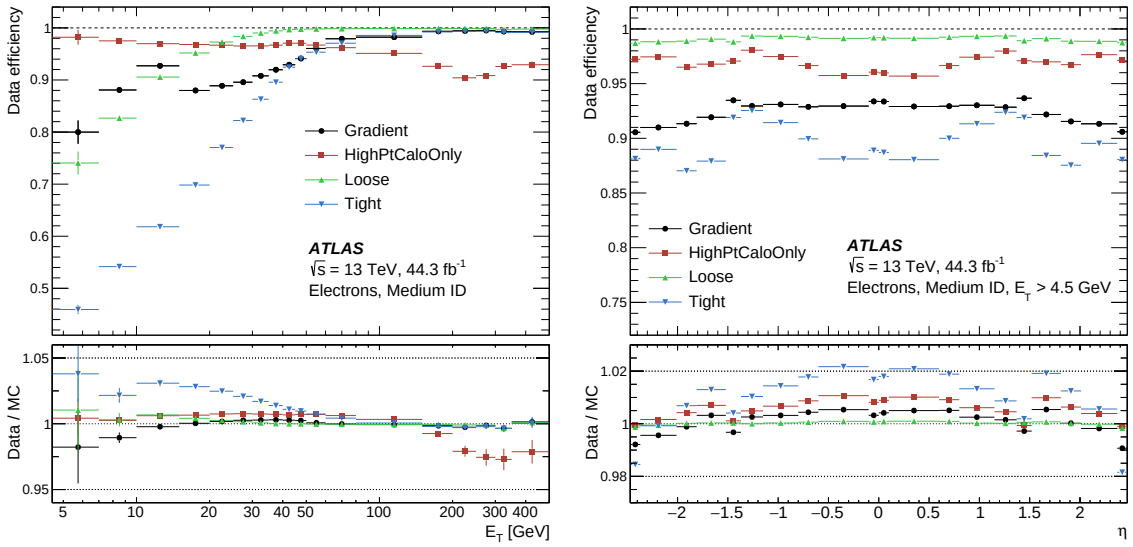
Figure 6.3: Electron Identification Efficiencies for Various Working Points. The electron identification efficiency is measured as a function of the electron's **(a)** transverse energy and **(b)** η for the three main electron identification working points. The efficiency for a Tight selection is nominally between 60%-70% in the $E_{\mathrm{T}} < 20$ GeV range. Since the Tight working point is often used for signal electrons, a cut on the electron's transverse energy is placed to avoid this region of lower efficiency. Plots are taken from Ref. [112].

The identification working points are computed using a likelihood approach. This method determines a discriminating variable based on the quality of the track reconstruction and the distribution of the energy deposition in the various layers of the calorimeter (see Table 1 in Ref. [112]). As the name suggests, the Loose working point has the loosest set of criteria and thus the highest identification efficiency (when evaluated on $Z \to ee$ events). This higher efficiency comes at the cost of lower background rejection, leading to a higher probability that an object that was not an electron (e.g., jet) is incorrectly identified as an electron. In practice, a signal electron is nominally selected using the Tight working point to ensure the lowest probability of misidentification.

6.1.2.4 Electron Isolation

Electron isolation characterizes the amount of separation between an electron candidate and all of the other decay products in an event. A well-separated electron is also known as a prompt electron, identified as originating directly from the main signal process. In this analysis specifically, a prompt electron originates from the decay of a W. On the other hand, non-prompt electrons can originate from different sources: (1) the product of a heavy-flavored quark decays (e.g., a b-quark decays to a c-quark via the exchange of a W which subsequently decays into a lepton and neutrino), (2) pair production from a photon converting to a pair of e^+e^-, and (3) from other particles being misidentified as electrons. Like electron identification, ATLAS provides several isolation working points for flexibility. Figure 6.4

plots the different isolation working points as a function of the electron candidate's E_T and η.



(a) Electron isolation efficiency as a function of E_T for different working points.

(b) Electron isolation efficiency as a function of η for different working points.

Figure 6.4: Electron Isolation Efficiencies for Various Working Points. The electron isolation efficiency is measured as a function of the electron's **(a)** transverse energy and **(b)** η for the three main electron identification working points. The different isolation working points labeled in the figure represent different requirements on the calorimeter and track isolation variables. Plots are taken from Ref. [112].

The figures of merit for electron isolation are the calorimeter and track isolation variables, E_T^{coneXX} and p_T^{coneXX} variables respectively [113]. The former represents the amount of energy deposited in the calorimeter cells found within a cone of size $\Delta R < XX/100$[4].

$$E_T^{\text{coneXX}} = E_{T,\text{raw}}^{\text{isolXX}} - E_{T,\text{core}} - E_{T,\text{leakage}}(E_T, \eta, \Delta R) - E_{T,\text{pileup}}(\eta, \Delta R) \tag{6.2}$$

In the above equation, $E_{T,\text{raw}}^{\text{isolXX}}$ is the raw isolated calorimeter energy from the sum of the transverse energy of topo-clusters whose barycenters fall within a $\Delta R < XX/100$ cone of the electron cluster barycenter. $E_{T,\text{core}}$ represents the energy of the clusters directly associated with the electron candidate. This term may not include all of the energy of the electron candidate, so a correction term, $E_{T,\text{leakage}}$, is required. Lastly, a corrective term accounting for pileup and the underlying event is also subtracted away. The track isolation variable is much simpler to compute as p_T^{coneXX} is determined by simply adding up the transverse momentum of tracks not associated with the electron that falls within the $\Delta R < XX/100$ cone. These two variables' criteria can be defined to determine the various isolation working points.

[4] $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is a distance metric used to determine the proximity of two objects in $\eta - \phi$ space.

6.1.2.5 *Electron Requirements*

Table 6.1 lists the baseline requirements for the signal and background electrons.

Table 6.1: Baseline Electron Requirements. The baseline requirements for the signal and background electrons are listed in the table. The background electrons are identified as failing the identification and isolation requirements, while the two sets of electrons share all other requirements.

e Requirement	Signal	Background
Trigger		
2015	`HLT_e24_lhmedium_L1EM20VH`	
	`HLT_e60_lhmedium`	
	`HLT_e120_lhloose`	
2016/2017/2018	`HLT_e26_lhtight_nod0_ivarloose`	
	`HLT_e60_lhmedium_nod0`	
	`HLT_e140_lhloose_nod0`	
Trigger matched	Yes	
Identification	`LHTight`	`LHLoose`
Isolation	`PLImprovedTight`	fail `PLImprovedTight`
p_{T}	> 10 GeV	
$\lvert\eta\rvert$	< 2.47	
Crack veto	veto $1.37 < \lvert\eta\rvert < 1.52$	
$\lvert d_0\rvert/\sigma_{d_0}$	< 5.0	
$\lvert z_0 \sin(\theta) \rvert$	< 0.5 mm	
Bad cluster veto	Yes	

The following is a description of each of the electron requirements listed in the table above.

Trigger The set of single electron triggers[5] is listed for each data-taking year, where a logical OR is used for the set of triggers. Each of the listed triggers in the above table is unprescaled[6].

Trigger Matched This condition requires the online objects to be matched to their offline counterparts.

Identification The identification requirements on the quality of the electron reconstruction.

Isolation The isolation requirements on the quality of the electron reconstruction[7].

[5] The name of the trigger nominally identifies the requirements of the trigger itself. These indicate the level at which the trigger is applied (HLT versus L1), specific kinematic requirements, identification/isolation requirements, and other requirements based on other subsystems. For example, the `HLT_e60_lhmedium_nod0` indicates that this trigger is applied at the HLT level requiring a single electron with $p_{\mathrm{T}} > 60$ GeV passing the medium likelihood identification with no additional constraints on the transverse impact parameter, d_0.

[6] A prescale factor is a number, N, that defines the fraction of events that passed the trigger are saved. A prescale factor of $N > 1$ means that only 1 out of N events are saved for later analysis, while $N = 1$ refers to an unprescaled trigger in which all events are saved. Implementing a prescale factor comes about because of limited computing resources and maximizing efficient data rates.

[7] The `PLImprovedTight` algorithm is the current recommended isolation working point (https://twiki.cern.ch/twiki/bin/view/Main/PLImprovedWPs). This algorithm is an improvement over previous algorithms at rejecting non-prompt electron and muon candidates from the decay of heavy hadrons.

p_T Transverse momentum requirements.

$|\eta|$ Absolute pseudorapidity requirements.

Crack Veto This condition removes leptons that fall into the crack region of the EM calorimeter in the $1.37 < |\eta| < 1.52$ range. This region contains services and infrastructure for the Inner Detector. Therefore, this region contains very few active detector materials.

$|d_0|/\sigma_{d_0}$ This is a requirement on the significance of the transverse impact parameter[8].

$z_0 \sin\theta$ This is a requirement on the longitudinal impact parameter[9].

Bad Cluster Veto This condition removes leptons that were not reconstructed using good-quality clusters in the calorimeter due to poor pulse shapes in the individual calorimeter cells.

6.1.3 Muons

Muon reconstruction also uses tracking information from hits in the Inner Detector. However, because muons are minimally ionizing particles, they pass through the calorimeters without significant energy deposition. Instead, the reconstruction algorithm combines hits from the outer muon spectrometers with tracks to form a muon candidate. Muon candidates are calibrated to account for detector effects before being assigned to the proper identification and isolation working points.

6.1.3.1 Muon Reconstruction

The tracks used during muon reconstruction use the same tracking algorithms used during electron reconstruction. In the case of muons, the energy loss due to bremsstrahlung is much less than that of electrons. At the same time, the system combines hits in the subdetectors of the Muon Spectrometer to construct various segments [114]. For signals registered in the MDTs, a straight-line fit is applied to the hits in each of the layers of a MDT module. Hits in the RPC or TGC modules measure the coordinate orthogonal to the bending plane. Finally, a combinatorial search in η and ϕ is used to build a segment from hits in the CSC. The result is a set of segments compiled from each subdetector which enters into a fitting algorithm to piece together a precise muon trajectory through the Muon Spectrometer.

The segment-fitting algorithm uses a segment-seeded combinatorial search [115]. The search starts with a seed segment selected from the middle layers of the detector where most of the trigger hits are registered. Segments from the inner and outer layers of the detector are progressively added to the seed, provided that they meet specific quality requirements such as the number of hits relative to their position and angle. A minimum of two matched segments are required to form a track candidate. It is possible the same segments are used in constructing multiple track candidates, in which case a

[8] The transverse impact parameter, d_0 is a measurement of the distance of the closest approach of a particle's trajectory to the point of collision in the transverse plane and its associated uncertainty, σ_{d_0}

[9] The longitudinal impact parameter, z_0 measures the distance between the primary vertex and the point on the particle's trajectory to be used to evaluate d_0. $z_0 \sin\theta$ is the vertical component of z_0 (where θ is the usual polar angle measured from the positive z-axis) and measures the degree of curvature due to deviations from a purely longitudinal trajectory.

downstream overlap removal algorithm selects the best set of segments for the track. Track candidates then pass through a final global χ^2 fit to determine if any of the reconstructed tracks pass the χ^2 selection criteria. During this iterative process, hits that negatively impact the χ^2 are removed and the fit is repeated. The result is a set of track candidates in the Muon Spectrometer ready to be matched to their counterparts in the Inner Detector.

The final muon candidate is created by combining the tracks from the Inner Detector and Muon Spectrometer and energy deposits in the calorimeters. Depending on which subdetectors are used, there are four types of muons.

Combined (CB) Muon Track candidates from the ID are matched to a track candidate in the MS. A global refit combines hits from the ID and MS, adding and removing hits to improve the overall fit quality. Successfully matched tracks are considered to be muons.

Segment-Tagged (ST) Muon A track candidate from the ID is considered a muon if it is associated with at least one local track segment in the MDT or CSC.

Calorimeter-Tagged (CT) Muon A track candidate from the ID is considered a muon if it can be associated with a cluster in the calorimeter that matches the shower profile of a minimum-ionizing particle. This approach has the lowest purity out of the four reconstruction algorithms. This approach recovers muons in regions where the MS contains no active detector materials.

Extrapolated (ME) Muon A track candidate in the MS is considered a muon if it extends back to the primary interaction point. The track is nominally required to pass through at least two layers of the MS in the barrel region, while three layers are required in the forward region. In this approach, tracks in the ID are not required to be associated with the track candidate in the MS.

6.1.3.2 Muon Calibration

Like the calibration procedures discussed for electrons, corrective factors must be applied to the muon candidates to correct the energy and momentum scales [116]. The first effect is the charge-dependent bias in the muon momentum measurement due to detector misalignment. This asymmetry causes positively and negatively charged muons to behave slightly differently in the bending plane of the MS. The charge-dependent bias is parameterized as follows.

$$\frac{q}{\hat{p}} = \frac{q}{p} + q \cdot \delta_s \tag{6.3}$$

where $q = \pm 1$ is the charge of the muon, p is the corrected momentum, $\hat{p}$ is the uncorrected momentum, and δ_s is the strength of the bias. Evaluating the bias term is an iterative process that begins with a large sample of $Z \to \mu\mu$ events. It can be shown that the reconstructed dimuon mass can be expressed in terms of δ_s.

$$m^2_{\mu\mu} = \hat{m}^2_{\mu\mu}(1 + \delta_s(\eta, \phi)p^+_{\mathrm{T}} - \delta_s(\eta, \phi)p^-_{\mathrm{T}}) \tag{6.4}$$

The momentums of the positive and negative muons are approximately equal, $p_{\mathrm{T}}^{+} \sim p_{\mathrm{T}}^{-}$. Therefore, the charge-dependent bias effect does not affect the average dimuon invariant mass. Instead, it manifests in the width of the dimuon invariant mass. The bias term, $\delta_s(\eta, \phi)$, is evaluated by minimizing the width of the invariant mass distribution. This value is then used in the calculation of the corrected transverse momentum.

$$p_{\mathrm{T}} = \frac{\hat{p}_{\mathrm{T}}}{1 - q\delta_s(\eta, \phi)\hat{p}_{\mathrm{T}}} \tag{6.5}$$

The corrected transverse momentum is then used to recalculate the invariant mass distribution, thus starting the next iteration. This process continues until a certain threshold for δ_s is reached. The result is that the bias is reduced from 0.4 TeV^{-1} to 2×10^{-4} TeV^{-1} for muons with $p_{\mathrm{T}} < 450$ GeV.

The second effect is the correction to the scale and resolution of the muon momentum measurement. This effect is studied using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events and is applied to simulation. These samples are used to construct the dimuon invariant mass distribution in both simulation and data. The simulated invariant mass distribution is used as a template that can then be fit onto the same distribution measured in data. An iterative minimization process is then used to determine the best parameters of the template. During each iteration, the level of agreement between simulation and data is evaluated using two approaches: (1) a simple comparison of the means and standard deviations and (2) calculating a binned χ^2 value. The best parameters result in the best agreement between the simulated distribution and the data. This calibration process improves the precision of the momentum scale by a factor of 2. The resolution of the muon p_{T} is measured to be 1.8% (2.3%) at small values of $|\eta|$ and 3.0% (3.4%) at large values of $|\eta|$ for J/ψ (Z) decays. These figures agree (within uncertainty) with the muon resolutions measured in simulation.

6.1.3.3 Muon Identification

Muon identification working points are similarly defined according to a set of requirements passed by the muon candidate. The criteria characterizing each working point is the momentum of the muon candidate relative to its uncertainty (q/p compatibility). This compatibility factor defines a significance-like discriminant. There are further requirements regarding the quality of the muon reconstruction in the Muon Spectrometer. For example, these include the minimum number of layers that need to be traversed by a passing muon and the goodness of the final χ^2 fit. There are three nominal muon identification working points labeled as *Loose*, *Medium*, and *Tight* [114]. These three working points accommodate the majority of physics analyses.

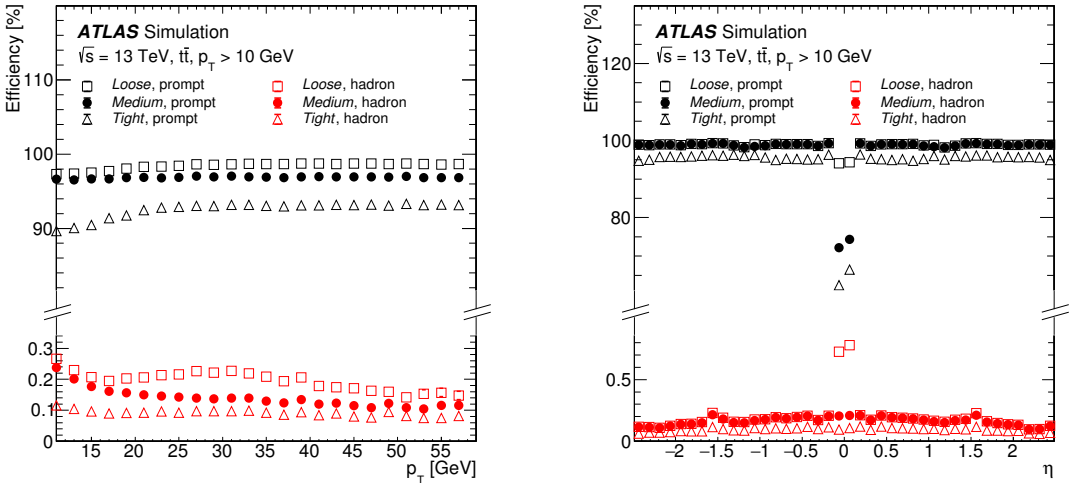
Loose The Loose working point is especially suited for processes with multiple muons in the final state. An example would be the $H \to \mu\mu\mu\mu$ process which contains four muons in the final state. The high muon multiplicity requires a selection with the highest possible efficiency to maximize statistics. The disadvantages of this selection are a lower purity and a higher misidentification rate.

Medium The muons selected by the Medium working point are a subset of those in the Loose working

point. This selection is the nominal selection used by most physics analyses and is suitable for a wide variety of use cases.

Tight The Tight working point uses the harshest requirements to yield the highest purity of actual muons. This selection is suited for analyses where statistics is not a limiting factor but may have larger backgrounds from non-prompt muons. This selection ensures the lowest number of non-prompt muons in the final selection.

There are two additional working points labeled as *High-*p_{T} and *Low-*p_{T}, which target selections requiring muons with $p_{\mathrm{T}} > 100$ GeV and muons with the lowest possible p_{T}. Specialized algorithms are used in these two extremes to select muons where the detector resolution may be limited, leading to poorer track quality. Figure 6.5 plots the different identification working points as a function of the muons candidate's E_{T} and η.



(a) Muon identification efficiency as a function of p_{T} for different working points.

(b) Muon identification efficiency as a function of η for different working points.

Figure 6.5: Muon Identification Efficiencies for Various Working Points. The muon identification efficiency is measured as a function of the muon's **(a)** transverse momentum and **(b)** η in $t\bar{t}$ events for the three main muon identification working points. The efficiencies of both prompt and non-prompt muons from the decay of hadrons are measured. Plots are taken from Ref. [114].

6.1.3.4 *Muon Isolation*

Similar to electrons, muon isolation is measured using calorimeter or track-based isolation variables, $E_{\mathrm{T}}^{\mathrm{topoconeXX}}$ and $p_{\mathrm{T}}^{\mathrm{coneXX}}$ respectively. The former measures the amount of transverse energy in the calorimeter cells in a $\Delta R < XX/100$ cone around the center of the muon energy cluster after subtracting away the energy from the muon itself and correcting for pileup effects. The latter is a sum of the tracks around the muon track candidate after subtracting away the transverse momentum of the muon itself.

These two isolation variables define the various isolation working points, the details of which are defined in Table 2 of Ref. [114]. Figure 6.6 plots the distribution of the different combinations of the isolation requirements.

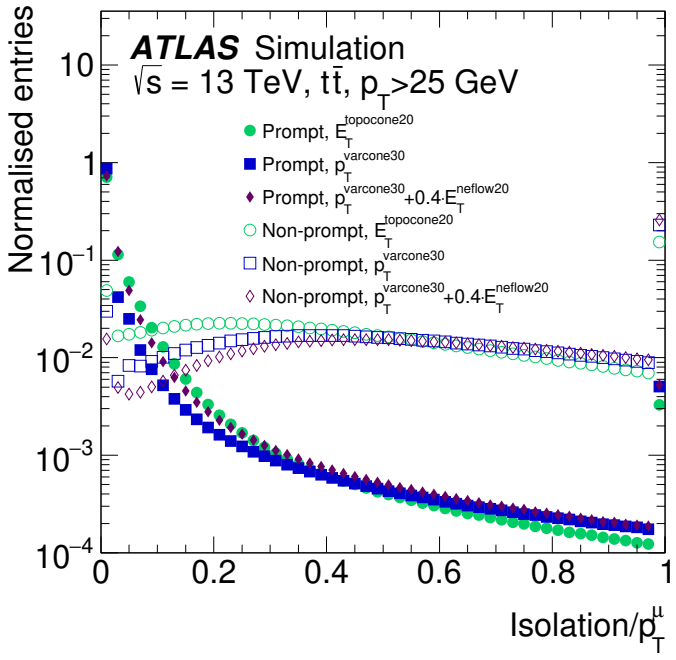


Figure 6.6: Distribution of the Isolation Variable Divided by Muon Transverse Momentum. The distribution of the isolation variables is shown for prompt and non-prompt muons studied using $t\bar{t}$ events. The distributions are normalized to unit area to illustrate the differences in the shape of the distributions between prompt versus non-prompt muons. Plot is taken from Ref. [114].

6.1.3.5 Muon Requirements

Table 6.2 lists the baseline requirements for the signal and background muons.

Table 6.2: Baseline Muon Requirements. The baseline requirements for the signal and background muons are listed in the table. The background muons are identified as failing the isolation requirements, while the two sets of muons share all other requirements.

μ Requirement	Signal	Background
Trigger		
2015	`HLT_mu20_iloose_L1MU15` OR `HLT_mu50`	
2016/2017/2018	`HLT_mu26_ivarmedium` OR `HLT_mu50`	
Trigger matched	Yes	
Identification	`Medium`	
Isolation	`PLImprovedTight`	fail `PLImprovedTight`
p_{T}	> 10 GeV	
$\lvert\eta\rvert$	< 2.5	
$\lvert d_0\rvert/\sigma_{d_0}$	< 3.0	
$z_0 \sin(\theta)$	< 0.5 mm	
Bad muon veto	Yes	

The "bad muon veto" requirement is unique to muons and removes muons with poor reconstruction qualities identified in either the Inner Detector or Muon Spectrometer. The other requirements in the table are common to both electrons and muons; the reader should refer to the descriptions of these requirements in the previous section.

6.1.4 Jets

Jets are collimated sprays of particles produced when a high-energy quark or gluon undergoes fragmentation and hadronization due to the property of color confinement. Jets are identified in ATLAS as the result of a specific jet-clustering algorithm that combines sufficiently close-together objects into a single jet object. There are numerous varieties of jet-clustering algorithms available, which differ on which objects are used as input and the strategies used to combine them. For this analysis, the selected jet-clustering algorithm is the anti-k_t [117] algorithm using Particle Flow Objects (PFO)[10] [118].

The anti-k_t algorithm is a recursive algorithm that iteratively combines pairs of objects until a specific condition is reached and a jet is returned. The algorithm starts by defining two distance quantities:

[10] PFOs are composite objects created by combining tracking and calorimeter signals and come in two types: charged and neutral. The former is created from calorimeter topo-clusters successfully associated with a charged track, while the latter corresponds to topo-clusters not associated with any charged track. PFOs are useful tools that have resulted in improved performance over using tracking and calorimetry individually. For low-energy charged particles, the momentum resolution of the tracker is much better than the energy resolution of the calorimeter. Meanwhile, the larger energy deposition in the calorimeter yields a superior energy resolution for high-energy particles. Therefore, combining the capabilities of the tracking and calorimeter subsystems is a natural next step. As a result of the many performance advantages, the use of PFOs extends beyond jet clustering and is incorporated in the calculation of other objects, such as $E_{\mathrm{T}}^{\mathrm{miss}}$.

$$d_{ij} = \min\left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2}\right)\frac{\Delta_{ij}^2}{R^2} \tag{6.6a}$$

$$d_{iB} = \frac{1}{k_{t,i}^2} \tag{6.6b}$$

In the above, $\Delta_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ where $k_{t,i}$, y_i, and ϕ_i are the transverse momentum, rapidity, and azimuthal angle of the i-th particle. The final size of the jet is set by the radius parameter, R, and denotes the maximum distance that two objects can be clustered into a single jet. The two distance quantities represent the distance between the pairwise (i, j) entities and the distance from the i-th entity to the beam, respectively. The d_{iB} is analogous to the original definition in the standard k_t algorithm [119] for an impact-parameter-like value for the i-th particle. The procedure starts by computing both distance measures and then compares d_{ij} and d_{iB}. If the smaller of the two values is a d_{ij}, then the two entities are combined into one object, and the algorithm continues its pairwise calculation. If the smaller of the two distance measures is a d_{iB}, then the i-th object is removed and declared a jet. This process continues until the complete list of entities is exhausted. Jet clustering prioritizes objects with higher p_T because of the $k_{t,i}$ in the denominator. The jets in this analysis are constructed using the anti-k_t algorithm with a jet radius of $R = 0.4$ and PFOs as input to generate the set of `AntiKt4EMPFlowJets`.

6.1.4.1 *Jet Tagging*

Jet tagging refers to labeling jets as originating from a specific flavor of quark. Jet tagging algorithms aim to separate jets originating from the decay of b-hadrons (b-jets) from the population of the other light quark decays (light-jets). The decay of the b-hadrons is relatively long compared to the other quark flavors, leaving a distinct displaced vertex in the detector. Accurate b-tagging is necessary for many analyses (including this one), which use the number of b-jets as a cut in the definition of their signal regions. Therefore, there have been significant developments to develop an accurate b-tagging algorithm [120]. Low-level algorithms tagged the displaced vertices and large impact parameters characteristic of decaying b-hadrons while currently recommended b-tagging algorithms use deep-learning architectures to make their predictions. Figure 6.7 illustrates various identifying characteristics of a b-jet.

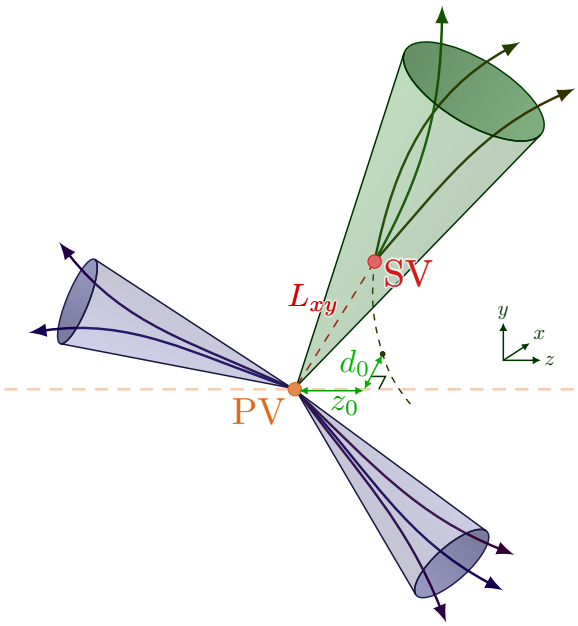
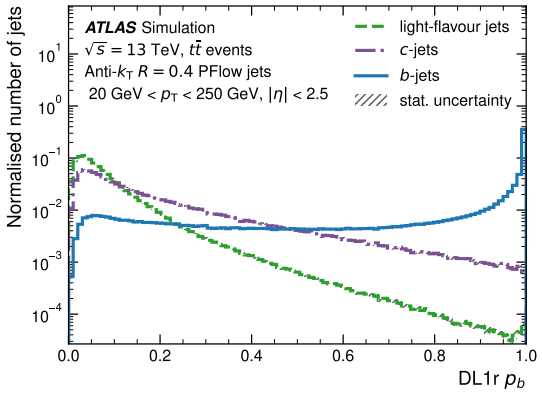


Figure 6.7: Identifying Characteristics of a Jet Initiated by a *b*-quark. The signature of jets produced by *b*-quarks is a displaced secondary vertex, SV. The decay length parameter, L_{xy}, determines the distance from the secondary vertex to the original primary vertex, PV. Other distance parameters of interest are the transverse and longitudinal impact parameters, d_0 and z_0, respectively. These parameters are entered into various *b*-tagging algorithms for jet classification. Plot is taken from Ref. [20].

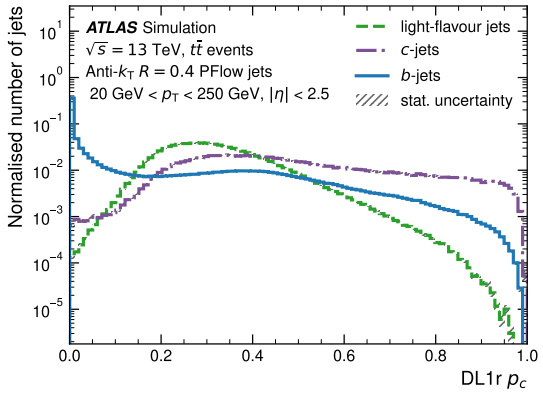
The *b*-tagging algorithm used to select *b*-jets in this analysis is the Deep-Learning 1 (DL1) algorithm which combines the outputs of the previously defined low-level algorithms into a fully-connected neural-network architecture [121]. In addition to the low-level algorithms' outputs, the jets' p_{T} and $|\eta|$ measurements are also included as input. The model is trained on a 70%/30% combined sample of $t\bar{t}$ and $Z' \rightarrow q\bar{q}$. The model outputs a set of probabilities corresponding to the probabilities for the jet to be a *b*-jet, *c*-jet, or light-jet. Using these probabilities, the DL1 discriminant, D_{DL1} is as follows.

$$D_{\mathrm{DL1}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\mathrm{light}}}\right) \tag{6.7}$$

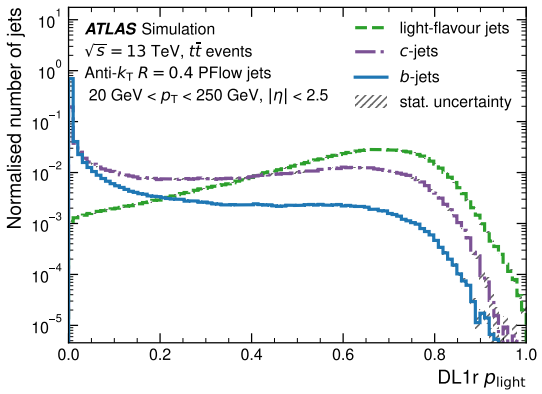
In the above, $p_{b,c,\mathrm{light}}$ corresponds to the *b*-jet, *c*-jet, and light-jet probabilities determined by the model. f_c is the effective *c*-jet fraction in the background hypothesis. This value is chosen depending on the specific needs of the physics analyses using the tagger. Figure 6.8 plots the distributions of the $p_{b,c,\mathrm{light}}$ and D_{DL1} for *b*-jets, *c*-jets, and light-jets.



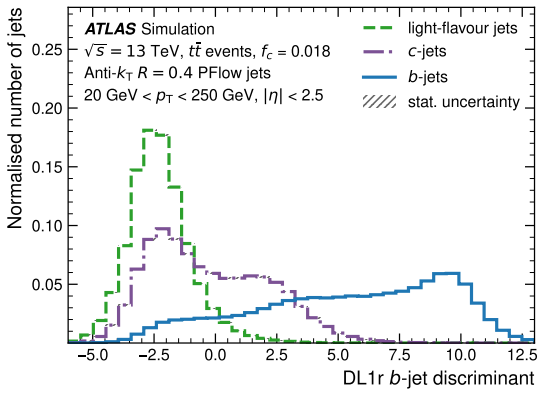
(a) Distribution of b-jet probabilities.



(b) Distribution of c-jet probabilities.



(c) Distribution of light-jet probabilities.



(d) Distribution of the DL1 b-jet discriminant.

Figure 6.8: DL1r b-tagging Probabilities and Discriminant in $t\bar{t}$ Events. The DL1r b-tagging algorithm is a multidimensional classifier whose outputs correspond to the probabilities a jet is a b-jet, c-jet, or light-jet. The distributions of the **(a)** b-jet probability, **(b)** c-jet probability, and **(c)** light-jet probability is shown. These probabilities are used in the calculation of the final b-tagging discriminant; the distribution is shown in **(d)**. These distributions are studied using simulated $t\bar{t}$ events and evaluated for each of the three groups of jets. Plots are taken from Ref. [121].

The performance of the DL1 tagger is often set at a fixed b-jet tagging efficiency, known as fixed-cut operating points. The efficiencies correspond to the true-positive rate of the tagger. For example, a 77% operating point corresponds to 77% of true b-jets would be correctly identified as such. Naively, setting the efficiency of a b-tagging algorithm to 100% (or near 100%) might seem desirable at first glance as it would mean that every true b-jet is correctly tagged. However, this greatly reduces the discriminatory power since non-b-jets with characteristics similar to true b-jets would also be tagged as a b-jet, leading to exponentially more misidentified jets. Therefore, setting an optimal b-tagging operating point is a fine balancing act to maximize high-b-jet efficiency and low c-jet/light-jet misidentification. There are four main fixed operating points, set at 60%, 70%, 77%, and 85%. Figure 6.9 plots the background rejection as a function of the b-jet tagging efficiency.

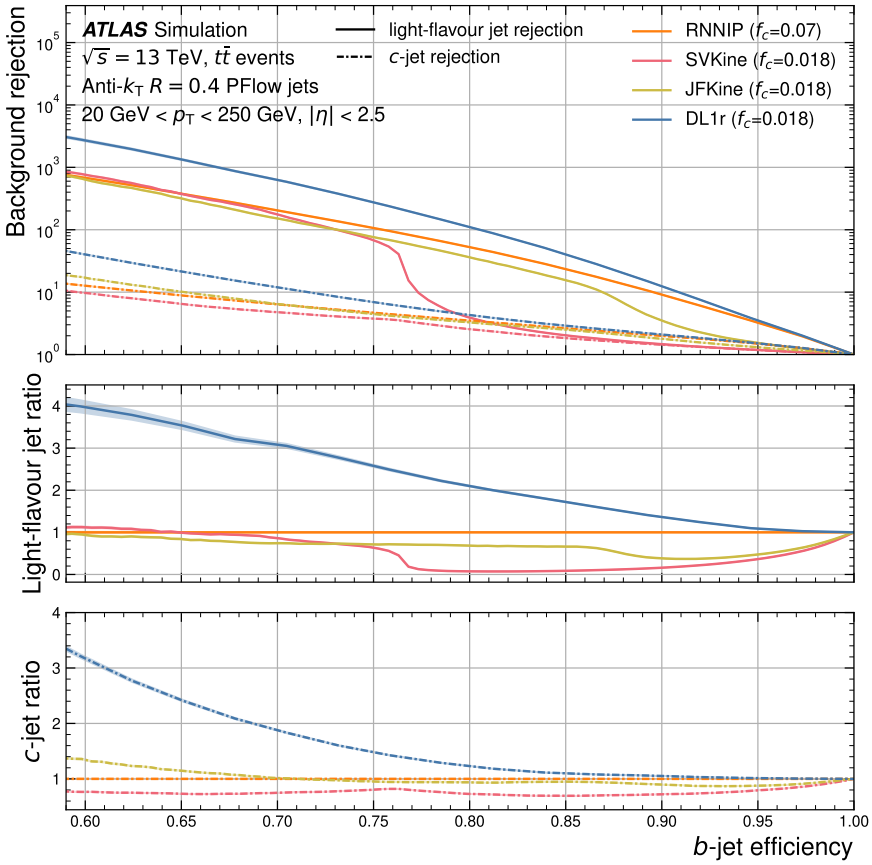
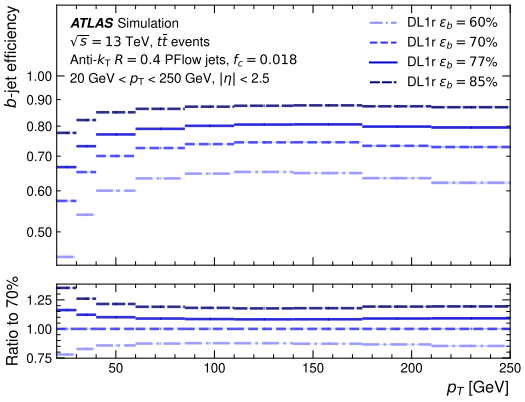
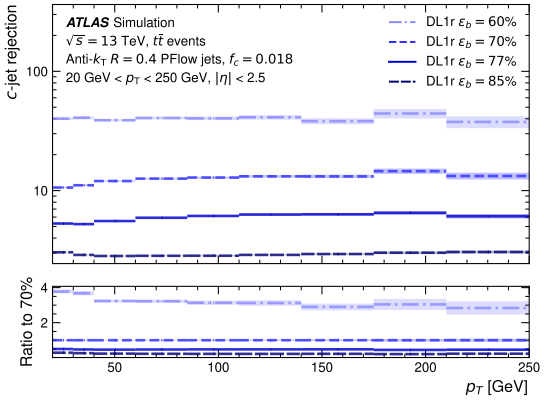


Figure 6.9: Background Rejection as a Function of the *b*-tagging Efficiency. The top panel shows the *c*-jet and light-jet rejection as a function of the *b*-tagging efficiency for various low-level and high-level algorithms. The bottom two panels show the ratio of the strength *c*-jet and light-jet rejection to the reference RNNIP algorithm, which uses a Recurrent Neural Network (RNN) to make jet tagging predictions. Across the ϵ_b range, the DL1r (the suffix "r" indicates the output of RNNIP is used as an input into DL1) outperforms all other lower-level taggers. Plot is taken from Ref. [121].

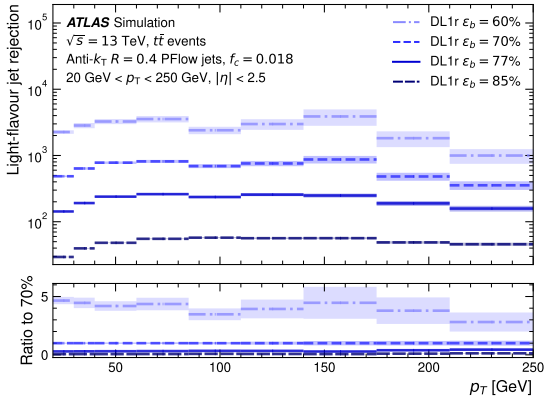
The efficiency of the *b*-tagging and background rejection power also varies with the jet's p_{T}. Specifically, the *b*-tagging efficiency is lower in the $[0, 50]$ GeV range and can fall below 50% for low (approximately 20 GeV) jets in the case of the 60% operating point. This behavior is shown in Figure 6.10 which plots the dependence of the *b*-tagging efficiency and *c*-jet/light-jet rejection on the jet's p_{T} for various operating points.



(a) The b-tagging efficiency as a function of the jet p_{T}.



(b) The c-jet rejection as a function of the jet p_{T}.



(c) The light-jet rejection as a function of the jet p_{T}.

Figure 6.10: DL1r b-tagging and c-jet/light-jet Rejection as a Function of Jet p_{T}. The performance of the DL1 b-jet tagger is dependent on the p_{T} of the jet. The **(a)** b-jet tagging efficiency, **(b)** c-jet rejection, and **(c)** light-jet rejection is shown as a function of the jet p_{T} for the four main ϵ_b operating points. The lower panel in each figure plots the ratio relative to the 70% operating point. Plots are taken from Ref. [121].

6.1.4.2 *Jet Requirements*

Table 6.3 lists the requirements for the signal and baseline jets.

Table 6.3: Baseline Jet Requirements. The signal and baseline requirements for jets are listed in the table. Baseline jets are defined as bare reconstructed jet objects representing the loosest jet selections, while the signal jets have stricter requirements and are specifically used in the analysis.

Jet Requirement	Signal	Baseline
Jet type	`AntiKt4EMPFlowJets`	
JVT working point	`Tight`	
p_{T}	> 25 GeV	> 20 GeV
$\lvert\eta\rvert$	< 2.5	< 4.5
b-tagging	DL1r (77%)	-

The following is a description of each of the jet requirements listed in the table above.

Jet Type The specific jet clustering algorithm used to reconstruct the jets. The `AntiKt4EMPFlowJets` reconstructs jets using EM-scale PFOs as input with a radius parameter of $R = 0.4$.

JVT Working Point The selected operating point of the JVT algorithm. The algorithm is applied to jets with p_{T} in the $[20, 60]$ GeV range and $|\eta| < 2.4$.

***b*-tagging** The selected b-tagging algorithm and the corresponding operating point. Any jet passing the tagger is classified as a b-jet, while those that fail are classified as a light-jet.

6.1.5 Missing Transverse Energy

The missing transverse energy represents the imbalance of energy/momentum of the final state particles in the transverse plane. In pp collisions, conservation of momentum dictates that the total momentum in any given direction should be conserved throughout the collision. Before the collision, there is zero momentum in the transverse plane. Therefore, conservation of momentum necessitates that after the collision, the sum of the transverse momenta should also equal zero. Weakly-interacting particles, like neutrinos, may escape the detector without leaving a signal. Since these particles still carry momentum, their absence can result in a non-zero value of the transverse momentum. The $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is then constructed as a two-component vector and a direction in the transverse plane defined by its azimuthal angle.

$$E_x^{\mathrm{miss}} = -\sum_i p_{x_i} \tag{6.8a}$$

$$E_y^{\mathrm{miss}} = -\sum_i p_{y_i} \tag{6.8b}$$

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = \{E_x^{\mathrm{miss}}, E_y^{\mathrm{miss}}\} \tag{6.8c}$$

$$E_{\mathrm{T}}^{\mathrm{miss}} = \left|\vec{E}_{\mathrm{T}}^{\mathrm{miss}}\right| = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \tag{6.8d}$$

$$\phi^{\mathrm{miss}} = \arctan\left(E_y^{\mathrm{miss}}/E_x^{\mathrm{miss}}\right) \tag{6.8e}$$

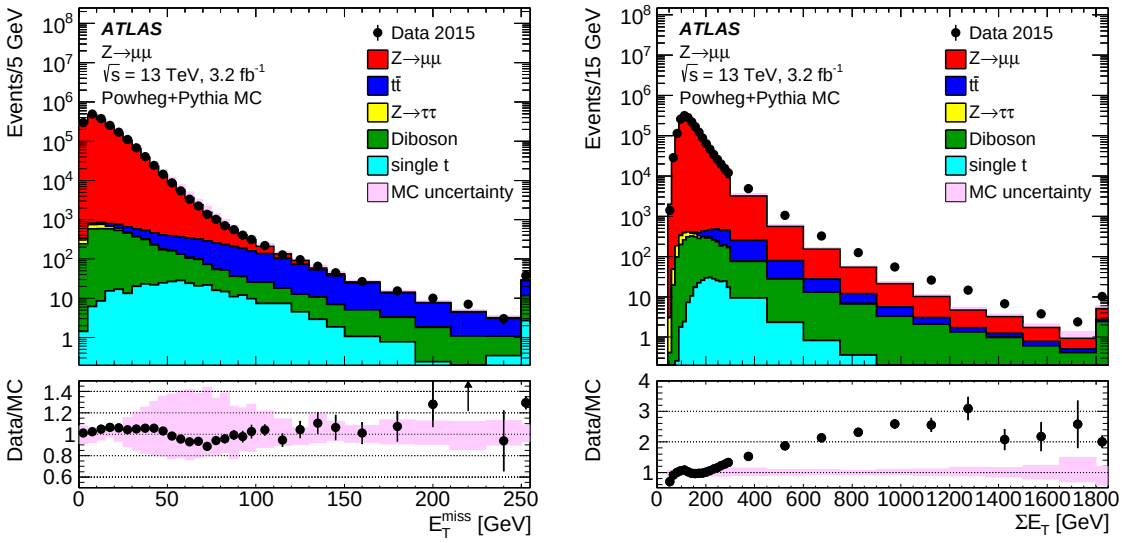
ATLAS computes the $E_{\mathrm{T}}^{\mathrm{miss}}$ observables with this approach using the visible objects in the event. The contributions to the $E_{\mathrm{T}}^{\mathrm{miss}}$ fall under two categories: (1) the hard-event signals and (2) the soft-event signals. The hard-event objects include the complete set of reconstructed and calibrated particles and jets. The reconstructed particles include electrons, photons, τ-leptons, and muons. The soft-event signals consider reconstructed tracks that are not associated with the hard-scatter vertex but are not associated with any of the hard objects. We calculate the missing transverse momentum by vectorially summing the transverse momentum of all visible particles and assigning this vector an overall negative sign [122].

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\left[\sum \vec{p}_{\mathrm{T}}^{\,e} + \sum \vec{p}_{\mathrm{T}}^{\,\gamma} + \sum \vec{p}_{\mathrm{T}}^{\,\tau} + \sum \vec{p}_{\mathrm{T}}^{\,\mathrm{jets}} + \sum \vec{p}_{\mathrm{T}}^{\,\mu} + \sum \vec{p}_{\mathrm{T}}^{\,\mathrm{soft}}\right] \tag{6.9}$$

Another relevant quantity is the scalar sum of all transverse momenta, which is defined as the scalar sum of all hard and soft objects. This quantity is particularly useful because it provides an overall scale of the "hardness" of the event.

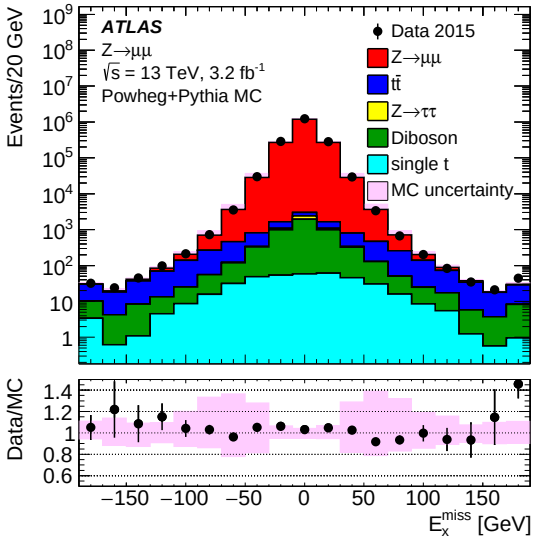
$$\sum E_{\mathrm{T}} = \sum_{i\in\{\mathrm{hard}\}} p_{\mathrm{T,i}} + \sum_{j\in\{\mathrm{soft}\}} p_{\mathrm{T,j}} \tag{6.10}$$

The $E_{\mathrm{T}}^{\mathrm{miss}}$ observables are evaluated using $Z \to \mu\mu$ events. $Z \to \mu\mu$ is the preferred channel for evaluating the fidelity of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ since the $Z \to \mu\mu$ final state can be selected with a high signal-to-background ratio. In these events, the muons are considered to be visible, and any neutrinos are produced through very rare heavy-flavor decays with an accompanying hadronic recoil. There should be no real sources of $E_{\mathrm{T}}^{\mathrm{miss}}$ in this channel, making it ideal for evaluating the quality and resolution of the reconstructed observables. Figure 6.11 plots the quality of the MC modeling of the various $E_{\mathrm{T}}^{\mathrm{miss}}$ observables.

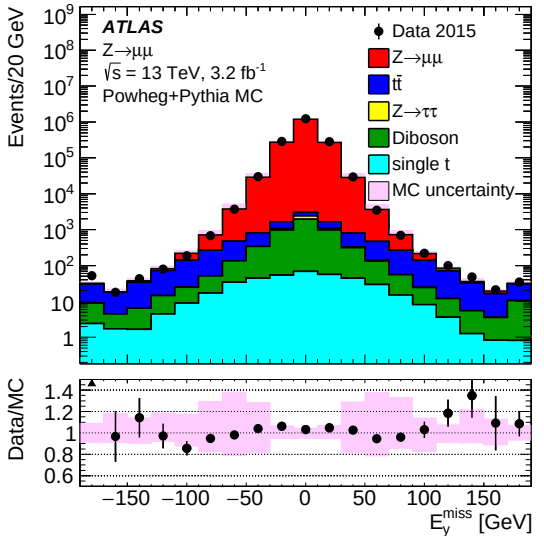


(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

(b) Scalar sum E_{T} distribution.



(c) $E_{\mathrm{X}}^{\mathrm{MISS}}$ distribution.



(d) $E_{\mathrm{Y}}^{\mathrm{MISS}}$ distribution.

Figure 6.11: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ Observables Evaluated Using a $Z \to \mu\mu$ Sample. A $Z \to \mu\mu$ sample is ideally used to evaluate the quality and resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction since it is expected to contain no real $E_{\mathrm{T}}^{\mathrm{miss}}$ events. The measured distributions can be easily compared to the expected $E_{\mathrm{T}}^{\mathrm{miss}} = 0$, with the widths of the distributions quantifying the resolution. The following observables are measured in both data and MC: **(a)** $E_{\mathrm{T}}^{\mathrm{miss}}$, **(b)** scalar sum E_{T}, **(c)** $E_{\mathrm{X}}^{\mathrm{miss}}$, and **(d)** $E_{\mathrm{Y}}^{\mathrm{miss}}$. The relevant backgrounds are also included. The pink shaded areas in the top panels of each plot shows the total MC uncertainty. The bottom panel of each plot show the data to MC ratio with the relative uncertainties shaded in pink. Plots are taken from Ref. [122].

The mismatch between data and MC increases towards the tails of the various distributions. The discrepancy in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution can be attributed to the mismodeling of the $t\bar{t}$ MC, the dominant

background in the simulation. The discrepancy in the scalar sum E_{T} distribution likely comes from mismodeling of the composition of hard objects in the final state. It should be noted here that pileup can also contribute to the disagreement. Multiple parton interactions in pileup events can add significant amounts of energy, these interactions can often be difficult to model. The $E_{\mathrm{X,Y}}^{\mathrm{miss}}$ distributions show good data/MC agreement, with differences appearing after $\pm$100 GeV.

6.1.6 *Overlap Removal*

Overlap removal refers to procedures for resolving double-counting of energy from different reconstructions of the same object. For example, an electron may be reconstructed as a jet resulting in twice the expected amount of energy deposited in the calorimeter. The following overlap removal procedures are relevant for this analysis[11].

Muons-Electrons Any calorimeter-tagged muon sharing a track with an electron is removed. Any electron subsequently found to share an Inner Detector track with a muon is removed.

Electrons-Jets Any jet found within a ΔR of 0.2 of an electron is removed. Any electron subsequently found with ΔR of 0.4 of a jet is removed.

Muons-Jets Any jet with less than three tracks associated with it and also found within ΔR of 0.2 of a muon is removed. Jets with $p_{\mathrm{T}}^{\mu}/p_{\mathrm{T}}^{\mathrm{jet}} > 0.5$ and $p_{\mathrm{T}}^{\mu}/\sum p_{\mathrm{T}}^{\mathrm{jet\ tracks}} > 0.7$ are removed. Any muons subsequently found within ΔR of 0.4 of a jet are removed.

6.2 Leptonic and Hadronic *W* Reconstruction

The topology of the semileptonic final state provides the unique opportunity to directly reconstruct and measure the W-boson kinematics. This analysis performs a dedicated measurement of the observables associated with the fully reconstructed $WWbb$ final state. A complementary analysis focuses on the fully leptonic channel of the $WWbb$ final state. This section briefly compares the semileptonic final state of this analysis and the fully leptonic channel. Next, the methods to reconstruct the leptonic and hadronic W-bosons are given. The W-boson reconstruction efficiency is evaluated using truth and particle-level events.

6.2.1 *Advantages of a Semileptonic Final State*

The semileptonic final state is defined by the decay of $WWbb$ to two b-jets, two light-jets, a lepton, and a neutrino. Conversely, the fully leptonic channel final state replaces the two light-jets with a second lepton and neutrino pair. Several advantages are gained by considering a semileptonic topology over its fully leptonic counterpart.

[11] The listed overlap removal procedures are part of recommendations offered by the Isolation and Fakes Forum subgroup, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/OverlapRemoval.

Second Neutrino Event Tagging In the fully-leptonic channel, the second neutrino contributing to the total $E_{\mathrm{T}}^{\mathrm{miss}}$ implies the $WWbb$ event kinematics cannot be fully reconstructed. There is no way to decouple the overall $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector in two separate momentum vectors representing the two neutrino tracks. In the semileptonic channel, there is only a single source of $E_{\mathrm{T}}^{\mathrm{miss}}$ representing the neutrino. Therefore, the $WWbb$ event kinematics can be fully reconstructed.

Fewer Misreconstructed Leptons The second lepton increases the chances of W mis-reconstructions when one attempts to correctly associate the two leptons with the two decaying W-bosons. The presence of jets in the final state of the semileptonic channel presents a natural method to assign jets and leptons to their respective parent W-bosons, thereby reducing the possibility of combinatorial mismatches.

Increased Statistics The branching ratios of the $W \to \ell\nu$ and $W \to q\bar{q}$ processes are approximately 32% and 68% resulting in increased statistics in the semileptonic channel.

6.2.2 Leptonic-W Reconstruction

The leptonic-W is reconstructed by combining the momenta of the lepton and the neutrino. The first step is to calculate the longitudinal component of the neutrino's momentum. This quantity is not well measured in the ATLAS detector and must be computed by other means following some assumptions. For a process with a single neutrino, it can be assumed that it is the only contributor to the $E_{\mathrm{T}}^{\mathrm{miss}}$. It can be shown that the longitudinal component of the neutrino's momentum can be expressed in the form of a quadratic equation.

$$Ap_{\nu,\mathrm{z}}^2 + Bp_{\nu,\mathrm{z}} + C = 0 \tag{6.11}$$

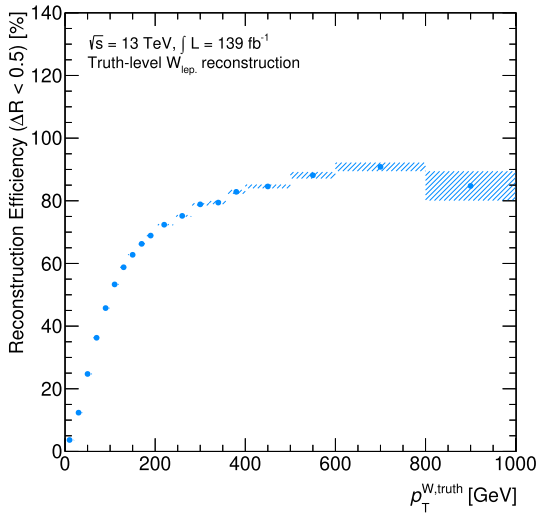
where the constants, A, B, and C, are functions of the energy of the lepton, the transverse momenta of the lepton and neutrino ($p_{\ell,\mathrm{x}}$, $p_{\ell,\mathrm{y}}$, $p_{\nu,\mathrm{x}}$, $p_{\nu,\mathrm{y}}$), and the constrained mass of the W (m_W) (a complete derivation of Eq. 6.11 with the complete definitions of A, B, C is shown in Appendix A). The quadratic equation can be solved to yield two solutions for $p_{\nu,\mathrm{z}}$. This two-fold ambiguity is expected and is a common feature in analyses performing similar reconstructions. Different approaches exist to select one of the two solutions for the neutrino momentum. In this analysis, which solution is chosen depends on whether there are two real solutions or zero real solutions based on the sign of the discriminant, $D = B^2 - 4AC$.

$\boldsymbol{D > 0}$ The discriminant is positive, which results in two real solutions, $p_{\nu,\mathrm{z}}^{1,2}$. The selected solution is the smallest $p_{\nu,\mathrm{z}}^{i}$ in absolute value.

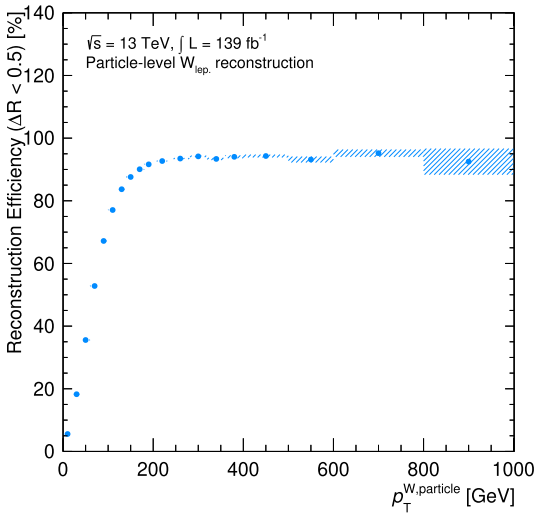
$\boldsymbol{D < 0}$ The discriminant is negative, which results in two complex solutions. The selected solution is the real part of the two roots, $p_{\nu,\mathrm{z}} = B/2A$.

This algorithm results in high reconstruction efficiencies compared to the same objects at the truth and particle level. An object at the detector-level is considered to be successfully matched to its par-

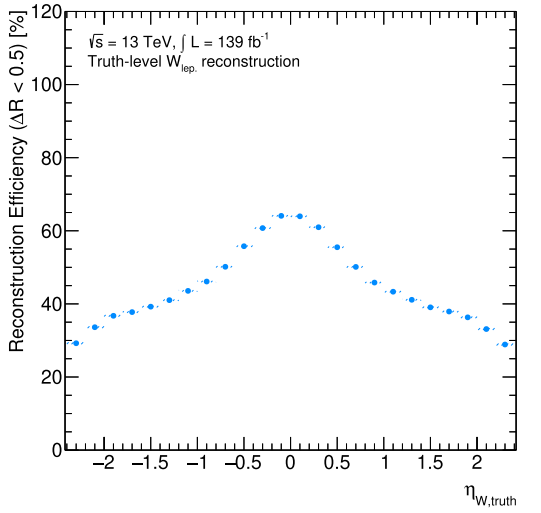
ticle/truth counterparts if the two objects satisfy the requirement $\Delta R < 0.5$. The reconstruction efficiency is then defined as the ratio of successful matches to the total number of events in the sample. Figure 6.12 plots the reconstruction efficiency as a function of the truth/particle-level W p_{T} and η. This measurement was conducted using $t\bar{t}$ events with the only condition being a minimum leading lepton p_{T} of 30 GeV in each event.



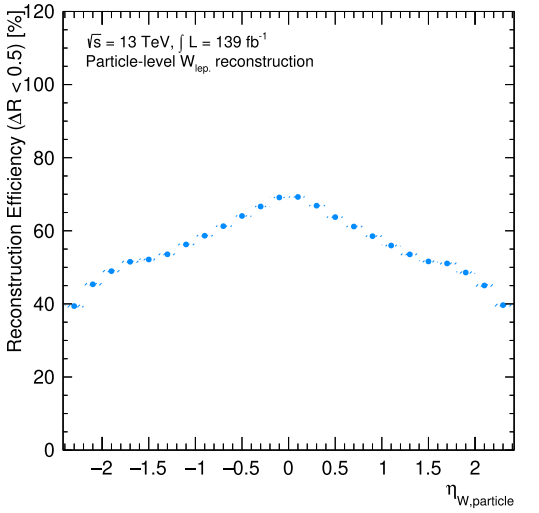
(a) Truth-level leptonic-W reconstruction efficiency as a function of p_{T}.



(b) Particle-level leptonic-W reconstruction efficiency as a function of p_{T}.



(c) Truth-level leptonic-W reconstruction efficiency as a function of η.



(d) Particle-level leptonic-W reconstruction efficiency as a function of η.

Figure 6.12: Leptonic-W Reconstruction Efficiency as a Function of p_{T} and η. The reconstruction efficiency of the leptonic-W is measured using $t\bar{t}$ events as a function of the **(a)** truth and **(b)** particle-level W-boson p_{T}. The reconstruction efficiency is also plotted as a function of the **(c)** truth and **(d)** particle-level W-boson η. The shaded area show the statistical uncertainty on the reconstruction efficiency.

6.2.3 Hadronic-W Reconstruction

Compared to the leptonic-W, the reconstruction of the hadronic-W is more difficult as it involves combining the set of light-jets in different ways. The starting point for all algorithms is a set of possible

W candidates formed from all 1-, 2-, and 3-jet combinations of light-jets passing the requirements for signal jets while failing the b-tagging criteria. The following algorithms have been investigated for reconstructing the hadronic-W. The overall goal is to select the algorithm that displays consistently high reconstruction efficiency across a broad range of the W kinematics. The reconstruction efficiency will be later used to guide cuts in certain signal region definitions.

2j, Highest p_{T} The 2-jet combination with the highest p_{T} is selected as the hadronic-W.

2j, Closest Mass The 2-jet combination with a mass closest to the true W mass is selected as the hadronic-W.

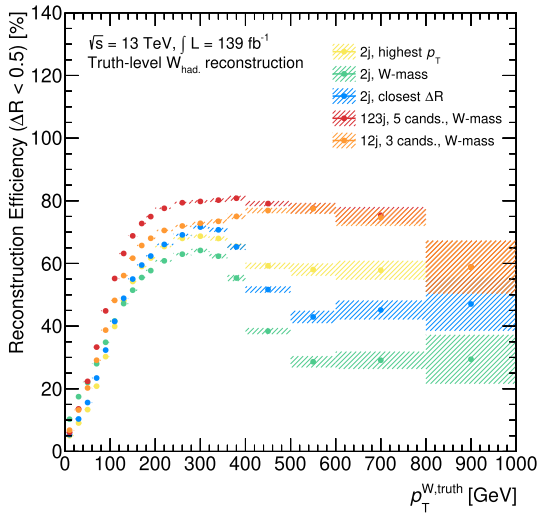
2j, Closest ΔR The 2-jet combination with the two constituent jets closest in ΔR is selected as the hadronic-W.

123j, 5 Cands., Closest Mass A list of five candidates comprises the two highest-p_{T} 1-jet combinations, the two highest-p_{T} 2-jet combinations, and the highest-p_{T} 3-jet combination. The hadronic-W is selected from this list of five candidates as the candidate possessing a mass closest to the true mass of the W.

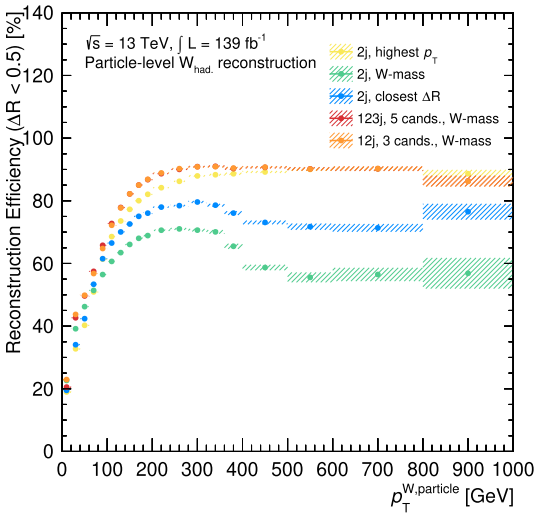
12j, 3 Cands., Closest Mass A list of three candidates comprises the two highest-p_{T} jets and the sum of these two jets. The hadronic-W is selected from this list as the candidate possessing a mass closest to the true mass of the W.

Nominally, the hadronic-W decays into two quarks, so the most natural approach would be to reconstruct the W from a combination of two light-jets. However, some algorithms consider 1-jet and 3-jet combinations to cover a broader range of physics phenomena. The 1-jet candidates pertain to events where the hadronic-W is produced with especially high p_{T}. When the boosted W decays, the two daughter quarks are also produced with relatively high momentum, resulting in a small opening angle between the two quarks. The two jets become highly collimated so that it is no longer possible to resolve two distinct jets. In this situation, the jet-clustering algorithm clusters the two quarks into a single jet. The 3-jet candidates pertain to events where one of the daughter quarks radiates a gluon, resulting in a third light-jet in the event.

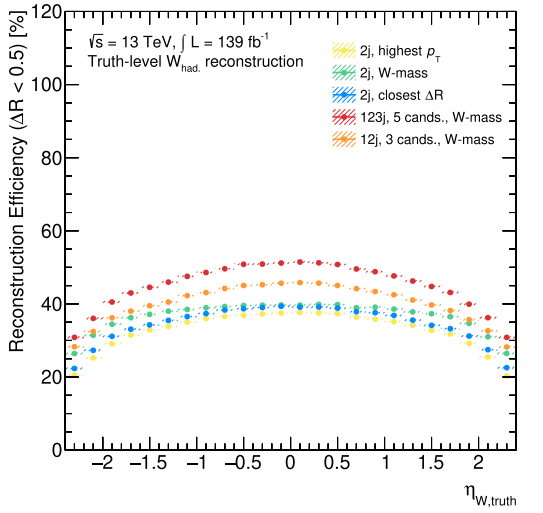
Like the leptonic-W, the reconstruction efficiency is evaluated on truth and particle-level events. The same $\Delta R < 0.5$ requirement is used to select for successfully matched reconstructions. Note that a successfully matched detector-to-particle level W indicates the constituent jets are successfully matched and does not guarantee the W was reconstructed correctly. Figure 6.13 plots the reconstruction efficiency for each previously mentioned algorithm as a function of the truth and particle-level hadronic-W p_{T} and η.



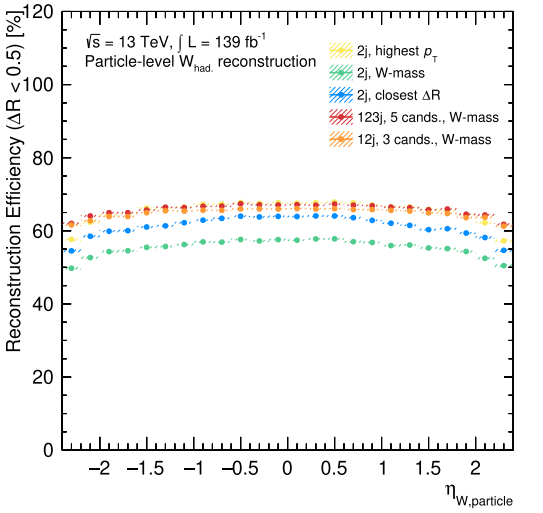
(a) Truth-level hadronic-W reconstruction efficiency as a function of p_{T}.



(b) Particle-level hadronic-W reconstruction efficiency as a function of p_{T}.



(c) Truth-level hadronic-W reconstruction efficiency as a function of η.



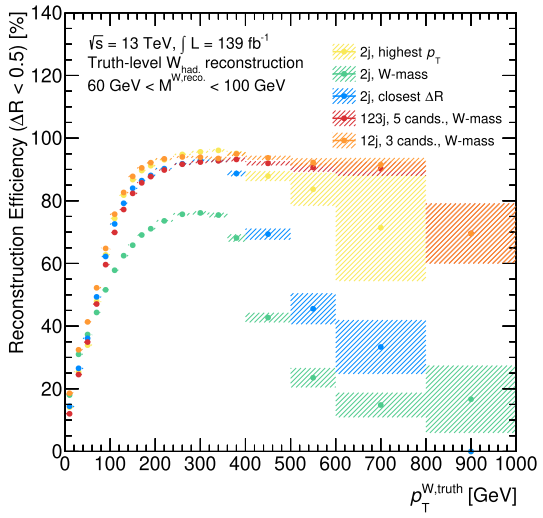
(d) Particle-level hadronic-W reconstruction efficiency as a function of η.

Figure 6.13: Hadronic-W Reconstruction Efficiency as a Function of p_{T} and η. The reconstruction efficiency of the hadronic-W is measured using $t\bar{t}$ events as a function of the **(a)** truth and **(b)** particle-level W-boson's p_{T}. The reconstruction efficiency is also plotted as a function of the **(c)** truth and **(d)** particle-level W-boson η. The shaded areas show the statistical uncertainty on the reconstruction efficiency.

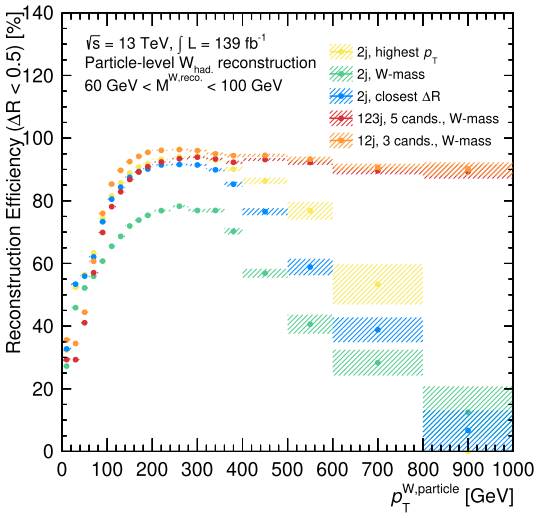
In the above plots, the reconstruction efficiency is lower in the low-p_{T} regions before plateauing at $p_{\mathrm{T}} > 100$ GeV. In these low-p_{T} regions, the W is produced with low momentum, leading to correspondingly low momenta of the daughter particles. The low-p_{T} values of the decay products may

cause them to fail the baseline requirements, making an accurate reconstruction impossible. In high-p_{T} regions ($p_{\mathrm{T}} > 300$ GeV), there is a noticeable drop observed in the algorithms considering only 2-jet candidates. As described above, in these regions, the two jets become collimated and are reconstructed as a single jet object, causing the 2-jet selections to become less valid. From these plots, either the *"123j, 5 cands., closest mass"* or the *"12j, 3 cands., closest mass"* algorithms should be selected as the best-performing algorithms.

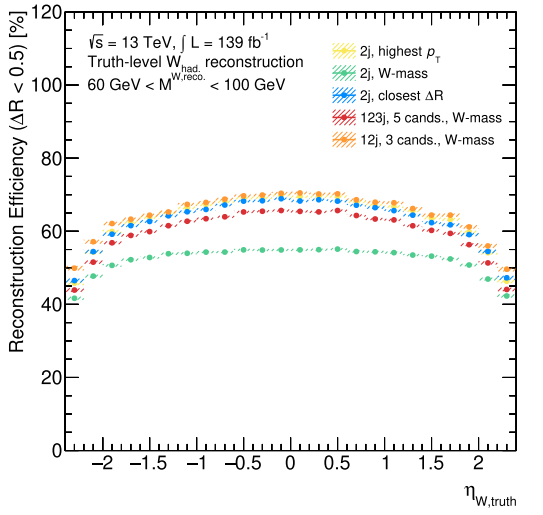
The reconstruction efficiency is measured again after requiring the mass of the reconstructed W fall within a 20 GeV window of the true W mass (i.e., 60 GeV $< M^{W,\mathrm{reco}} <$ 100 GeV). Figure 6.14 plots the reconstruction efficiency in truth and particle-level events after enforcing this mass-window requirement. Note that this efficiency is measured after applying the same 20 GeV mass window requirement to the particle-level events.



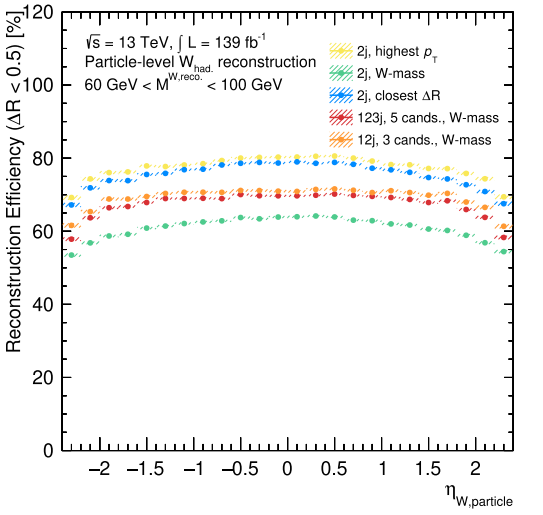
(a) Truth-level hadronic-W reconstruction efficiency as a function of p_{T} after requiring the mass of the reconstructed W be within $\pm$20 GeV of the true W mass.



(b) Particle-level hadronic-W reconstruction efficiency as a function of p_{T} after requiring the mass of the reconstructed W be within $\pm$20 GeV of the true W mass.



(c) Truth-level hadronic-W reconstruction efficiency as a function of η after requiring the mass of the reconstructed W be within $\pm$20 GeV of the true W mass.



(d) Particle-level hadronic-W reconstruction efficiency as a function of η after requiring the mass of the reconstructed W be within $\pm$20 GeV of the true W mass.

Figure 6.14: Hadronic-W Reconstruction Efficiency as a Function of p_{T} and η with an Additional W Mass Requirement. The reconstruction efficiency of the hadronic-W is measured using $t\bar{t}$ events as a function of the **(a)** truth and **(b)** particle-level W-boson's p_{T}. The reconstruction efficiency is also plotted as a function of the **(c)** truth and **(d)** particle-level W-boson η. A requirement of 60 GeV $< M_W^{\mathrm{reco}} <$ 100 GeV is also added. The shaded areas show the statistical uncertainty on the reconstruction efficiency.

After requiring 60 GeV $< M^{W,\text{reco}} <$ 100 GeV, all algorithms see an increase in the reconstruction efficiency. The *"123j, 5 cands., closest mass"* and *"12j, 3 cands., closest mass"* algorithms have both shown to be consistently well-performing over a large p_{T} range. However, the latter is chosen because it is simpler to understand from a physics standpoint and simpler to apply from an implementation standpoint. For the remainder of the analysis, the hadronic-W is therefore reconstructed using the *"12j, 3 cands., closest mass"* algorithm.

6.3 Minimax Algorithm

The $m_{b\ell}^{\text{minimax}}$ variable was previously shown to be sensitive to the interference between the singly- and doubly-resonant top-quark processes. The canonical minimax variable is constructed by combining two b-jets and two leptons, as was shown in Eq. 3.22. Given that there is only a single lepton in the event selection of this analysis, an identical minimax variable cannot be constructed. However, a modified minimax variable can be constructed by considering combinations of the single lepton (ℓ), the leading (b_1) and subleading b-jets (b_2), and the hadronic-W (W). This modified minimax variable ($m_{\text{mod.}}^{\text{minimax}}$) is defined as follows.

$$m_{\text{mod.}}^{\text{minimax}} \equiv \min\{\max(m_{b_1\ell}, m_{b_2W}), \max(m_{b_1W}, m_{b_2\ell})\} \tag{6.12}$$

The variable is also used to assign the two leading b-jets to the hadronically or leptonically decaying top quarks. Depending on which of the four combinations is used to form the minimax quantity, the selected b-jet would be assigned with the W or the lepton. For example, if the minimax observable is formed as $m_{b_1\ell}$, then the leading b-jet is assigned to the leptonically decaying top-quark. At the same time, the subleading b-jet is assigned to the hadronically decaying top-quark. Several observables are formed using this selection process, which will later be identified by the "selected" in the observable name.

6.4 Fake Lepton Estimation

During the reconstruction process, it is possible for an object that was not a lepton to begin with (i.e., jet) to be incorrectly reconstructed as a lepton. These misidentified leptons are therefore referred to as "fake leptons" and their contributions constitute another source of background in this analysis. Fake leptons are usually not produced in isolation and will likely be removed by tighter lepton identification and isolation requirements. Nevertheless, a non-negligible number of leptons can pass the tighter selection and negatively affect the signal purity. This section begins by describing the fake-factor method used to estimate the contributions from fake leptons. This method is applied to electrons and muons with the fake-factors parameterized as a function of the electron and muon kinematics.

6.4.1 *Fake-Factor Method*

Several different methods exist for estimating the background coming from fake leptons. This analysis uses the fake-factor method because of its simple and intuitive application. The fake-factor method begins by defining a control region corresponding to a set of selections orthogonal to those used for the primary analysis signal regions. The events in this region are enriched in the background of interest; in this case, the events contain a large proportion of fake leptons. An extrapolation factor is then derived using events in this region and then applied to the signal regions to determine the number of events containing a fake lepton. Even though the control region and signal regions are orthogonal in phase space, the fake-factor can be reasonably applied because the production of fake leptons is assumed to be identical in both regions.

In this analysis, the baseline selection requires a single lepton to pass the tight identification and isolation requirements. The number of leptons that pass the tight (N^t) and loose (N^l) selections can be expressed in terms of the number of real (N_r) and fake (N_f) leptons. In matrix form, the following system of equations connecting the two sets of quantities can be written as follows [123].

$$\begin{pmatrix} N^t \\ N^l \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_r \\ N_f \end{pmatrix} \tag{6.13}$$

The matrix above defines the connection between tight/loose and real/fake leptons using the real (r) and fake (f) efficiencies. The real efficiency is the probability that a real lepton passes the tight selection. In contrast, the fake efficiency is defined as the probability that a fake lepton passes the tight selection. The numbers N^t and N^l can be easily measured in data, so if the real/fake efficiencies can be determined, then an estimate of the number of real/fake leptons can be calculated. To determine the number of fake leptons passing the tight selection (N_f^t), we start by multiplying both sides of Eq. 6.13 by the row vector $(1, -\frac{1}{1-f})$.

$$N^t - \frac{f}{1-f}N^l = rN_r + fN_f - \frac{f}{1-f}(1-r)N_r - \frac{f}{1-f}(1-f)N_f \tag{6.14a}$$

$$rN_r + N_f^t - \frac{f}{1-f}N^l = rN_r - \frac{f}{1-f}N_r^l \tag{6.14b}$$

$$N_f^t = \frac{f}{1-f}(N^l - N_r^l) \tag{6.14c}$$

To go from Eq. 6.14a to Eq. 6.14b, the following equations are inserted: $N^t = N_r^t + N_f^t$, $N_r^t = rN_r$, and $N_r^l = (1-r)N_r$. This results in Eq. 6.14c which defines the number of fake leptons passing the tight selection as a function of the fake factor, $F = \frac{f}{1-f}$, the total number of loose leptons (N^l) and the number of real leptons passing the loose selection (N_r^l). The fake factor can be rewritten in terms of the number of fake leptons passing/failing the tight selection by using $N_f^t = fN_f$.

$$F = \frac{f}{1-f} = \frac{N_f^t}{N_f - N_f^t} = \frac{N_f^t}{N_f^l} \tag{6.15}$$

The numerator counts the number of fake leptons passing the tight selection, while the denominator counts the number of fake leptons failing the tight selection. In practice, the fake-factor is parameterized as a function of the properties of the lepton. Then, it is applied on an event-by-event basis as an event weight, making it pragmatically easier to fill histograms. The sum of the weights from the fake leptons then constitutes the total background due to the fake leptons [124].

$$N_f^t = \sum_{\text{data},i=1}^{N_l} F_i - \sum_{\text{MC},j=1}^{N_{\text{MC}}^l} w_{\text{MC,j}} F_j \tag{6.16}$$

In the above, the first sum runs over events in data containing a loose lepton while the second sum runs over MC events containing a loose lepton ($w_{\text{MC,j}}$ is the weight assigned to the MC event j). Since both sums only consider events with a loose lepton, events with tight leptons do not contribute to the total fake leptons background.

6.4.2 *Fakes Enriched Region Definition*

A fakes-enriched region with a large proportion of fake leptons is defined to measure the fake factors. It has been shown previously that the quantity $E_{\text{T}}^{\text{miss}} + m_{\text{T}}^{W}$ is particularly sensitive to QCD events which are the primary source of fake leptons [125]. The quantity m_{T} is known as the transverse mass and can be defined in terms of a particle's rest mass and transverse momentum.

$$m_{\text{T}}^2 = m^2 + p_x^2 + p_y^2 = E^2 - p_z^2 \tag{6.17}$$

While the rest mass is invariant under a Lorentz boost in all directions, the transverse mass is only invariant under boosts along the z-direction. This fact makes the transverse mass particularly useful for highly boosted events with particles traveling close to the beamline where the longitudinal component of the particle's momentum cannot be accurately measured. For a single heavy particle of mass M decaying to two daughter particles, the transverse mass of the parent particle is as follows.

$$M_{\text{T}}^2 \equiv \left(E_{\text{T}}^1 - E_{\text{T}}^2\right) - \left(\vec{p}_{\text{T}}^1 - \vec{p}_{\text{T}}^2\right) \tag{6.18}$$

where E_{T}^i, $\vec{p}_{\text{T}}^i$ correspond to the total energy and transverse momentum of the two daughter particles. If the decay occurs semi-invisibly (where one of the two daughter particles goes undetected), then the transverse momentum of one of the two daughters is equal to the missing transverse momentum in the event, $\vec{p}_{\text{T}}^1 = \vec{E}_{\text{T}}^{\text{miss}}$. This analysis measures the transverse mass using the kinematics of the leading lepton and the $E_{\text{T}}^{\text{miss}}$ in the event.

Nominally, a requirement of $E_{\text{T}}^{\text{miss}} + m_{\text{T}}^{W} > 60$ GeV is used to reject background from QCD. A control region with enhanced QCD contributions (and higher probabilities of fake leptons) is defined

by reversing the cut. In addition, the event should contain a minimum of three jets, of which at least one should be a b-jet. The fake-factor is calculated using the procedures defined in the previous section by measuring the relative fractions of the number of leptons passing tight and loose selections in this region. The fake factors are parameterized using the electron/muon p_{T} and $|\eta|$.

6.4.3 *Electron Fake Factors*

Several different sources could result in a fake electron. A fake electron could result from the mis-reconstruction of a highly-collimated jet, which occurs in highly energetic events resulting in boosted decay products leaving a narrow shower profile in the EM calorimeter. Other sources can produce so-called non-prompt electrons, which are actually electrons but are produced in secondary reactions apart from the prompt decay of interest. In this analysis, prompt electrons are expected to come only from the leptonic decay of a W-boson. A possible source of non-prompt electrons is photon conversion, which produces electron-positron pairs. Another source of non-prompt electrons comes from the decay of heavy-flavour decays in which hadrons containing $b-$ or c-quarks decay semileptonically to produce an electron. While non-prompt electrons are correctly reconstructed as electrons, these are still categorized as fake electrons in this study. Figure 6.15 plots the electron fake-factors parameterized by p_{T} in four bins of $|\eta|$.

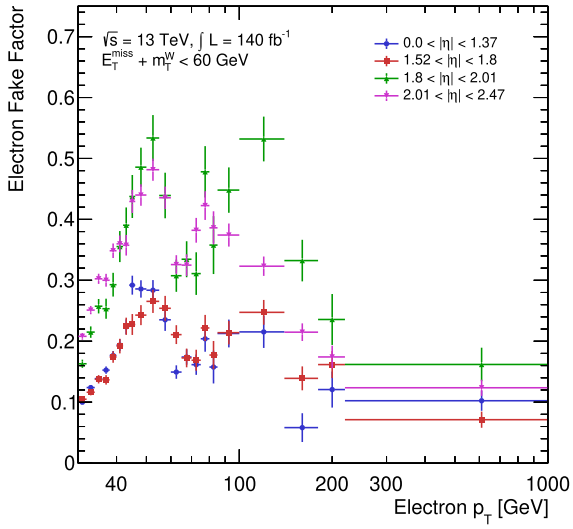


Figure 6.15: Electron Fake-Factors Plotted as a Function of p_{T} in Different Bins of η. Electron fake-factors are measured as a function of electron p_{T} in four different bins of $|\eta|$. Measurements are made in a specific fakes-enriched region created by reversing the anti-QCD cut. Any prompt electrons in this region are estimated using MC and subtracted away before calculating the fake-factors. The error bars on the data points correspond only to statistical uncertainties.

6.4.4 Muon Fake Factors

The decay of mesons in-flight within jets and the semileptonic decay of heavy-flavored hadrons dominate the production of fake muons. The fake-factors for muons are measured using the same fakes-enriched region used for electrons. The tag-and-probe method targets dijet events, possibly containing a fake muon. The highest-p_{T} light-jet in the event is identified as the "tag", while the reconstructed muon is the "probe". It is required the two objects in the event are back-to-back (i.e., $\Delta\phi(\mu, \mathrm{jet}) > 2.7$) and have the following minimum p_{T} values, $p_{\mathrm{T}}^{\mu} > 25$ GeV and $p_{\mathrm{T}}^{\mathrm{jet}} > 35$ GeV. Figure 6.16 plots the muon fake-factors parameterized by p_{T} in three bins of $|\eta|$.

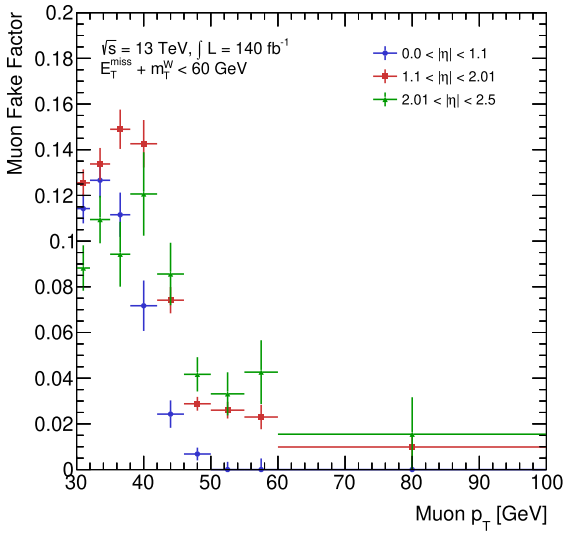


Figure 6.16: Muon Fake-Factors Plotted as a Function of p_{T} in Different Bins of η. Muon fake-factors are measured as a function of muon p_{T} in three bins of $|\eta|$. Measurements are made in a specific fakes-enriched region created by reversing the anti-QCD cut. A tag-and-probe approach selects dijet events where the highest-p_{T} light-jet acts as the tag while the reconstructed muon acts as the probe. The error bars on the data points correspond only to statistical uncertainties.

6.4.5 Fake Factors Uncertainties

Systematic uncertainties are estimated by varying aspects of the MC or cuts and then recalculating the fake factors. The following systematic variations are used to estimate the uncertainties associated with the electron and muon fake factors.

MC Scale The overall MC normalization is evaluated by varying the MC normalization by $\pm 10\%$.

2+ *b*-jets The fake factors are recalculated after tightening the b-jet requirement to 2+ b-jets.

$E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}^{W} < 50, 70$ GeV The fake factors are recalculated after adjusting the anti-QCD threshold to 50 and 70 GeV.

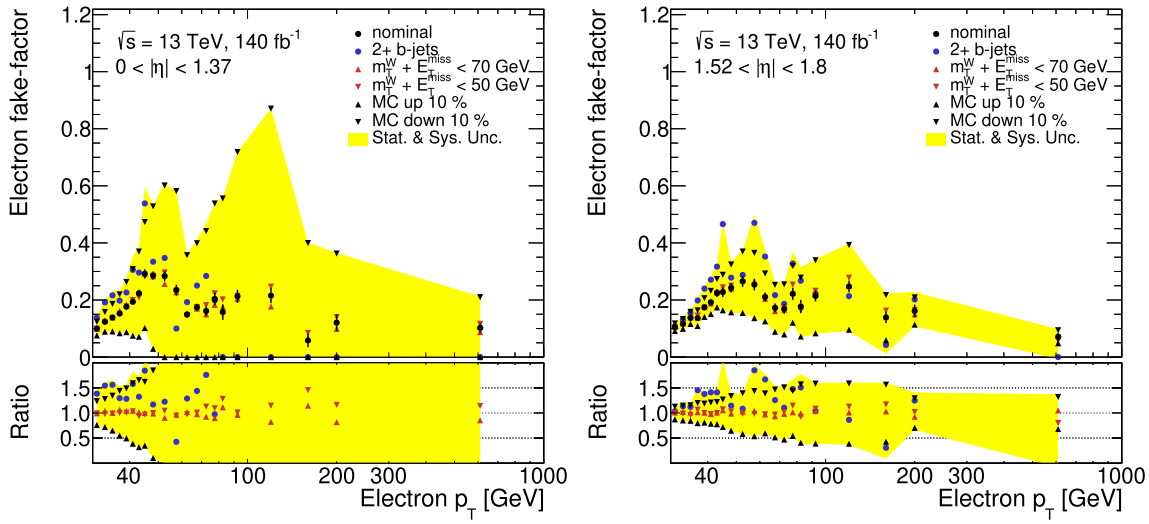
No Light-Jet Req. The fake-factors are recalculated after removing the minimum light-jet p_{T} requirement.

$\boldsymbol{p_{\mathrm{T}}(\mathbf{jet}) > 30, 40\ \mathbf{GeV}}$ The fake-factors are recalculated after adjusting the minimum light-jet p_{T} to 30 and 40 GeV.

$\boldsymbol{\Delta\phi(j,\mu) > 2.6, 2.8}$ The fake-factors are recalculated after adjusting the minimum $\Delta\phi$ thresholds to 2.6 and 2.8.

6.4.5.1 Electron Fake Factors Uncertainties

Figure 6.17 plots the recalculated electron fake factors after applying the systematic shifts defined above. The total uncertainty is evaluated by summing up the statistical and the systematic variations in quadrature.



(a) Electron fake-factor uncertainties in the first $|\eta|$ bin.

(b) Electron fake-factor uncertainties in the second $|\eta|$ bin.

(c) Electron fake-factor uncertainties in the third $|\eta|$ bin.

(d) Electron fake-factor uncertainties in the fourth $|\eta|$ bin.

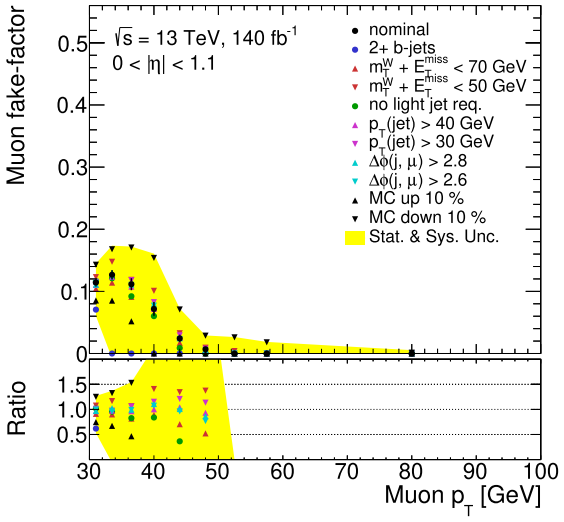
Figure 6.17: Electron Fake-Factors with Systematic Uncertainties. The electron fake-factors are recalculated after different systematics variations in the **(a)** $0 < |\eta| < 1.37$, **(b)** $1.52 < |\eta| < 1.8$, **(c)** $1.8 < |\eta| < 2.01$, and **(d)** $2.01 < |\eta| < 2.47$. The yellow band in each plot represents the total uncertainty on the fake factor calculated by combining the variations and the statistical uncertainty in quadrature.

The largest systematic uncertainty originates from varying the MC normalization. This effect stems from modeling of subtracted leptons, differences in the composition of fake electrons in the fakes-enriched region and the signal region, and the overall normalization of the MC samples in the fakes-enriched region. On the other hand, varying the anti-QCD threshold seems to have minimal effect.

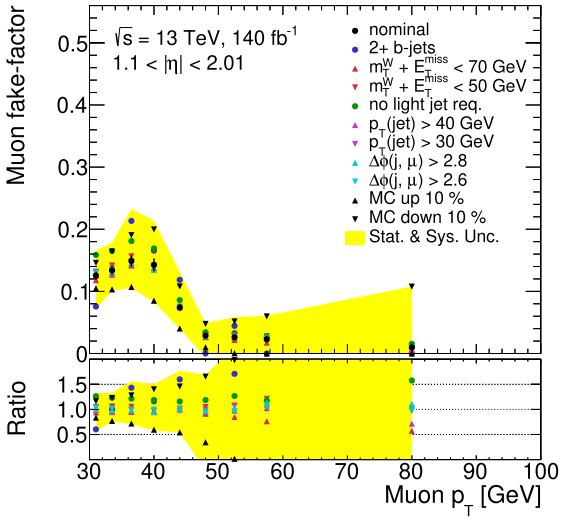
Therefore, the scale of the anti-QCD cut is set appropriately at 50-70 GeV. However, the total uncertainty can surpass 100% in several p_{T} bins.

6.4.5.2 Muon Fake Factors Uncertainties

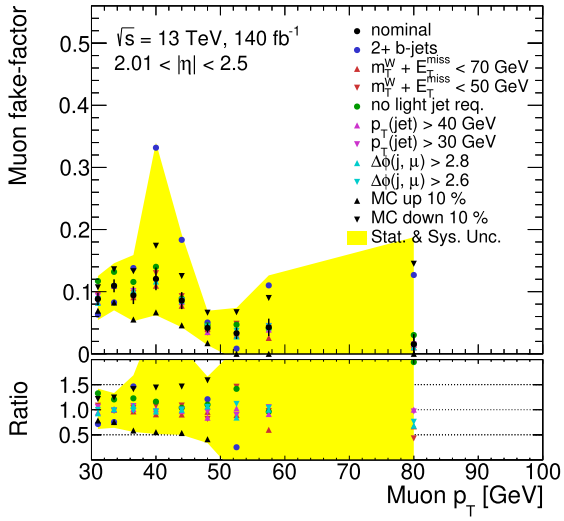
The uncertainties for the muon fake-factors are estimated using the same approach. In addition to the same systematic variations defined for the electron fake-factors, several other variations targeting the tag-and-probe method are included. Figure 6.18 plots the muon fake factors recalculated after applying these systematic shifts.



(a) Muon fake-factor uncertainties in the first $|\eta|$ bin.



(b) Muon fake-factor uncertainties in the second $|\eta|$ bin.



(c) Muon fake-factor uncertainties in the third $|\eta|$ bin.

Figure 6.18: Muon Fake-Factors with Systematic Uncertainties. The muon fake-factors are recalculated after different systematics variations in the **(a)** $0 < |\eta| < 1.1$, **(b)** $1.1 < |\eta| < 2.01$, and **(c)** $2.01 < |\eta| < 2.5$. The yellow band in each plot represents the total uncertainty on the fake factor calculated by combining the variations and the statistical uncertainty in quadrature.

Like the electron, the largest impact on the muon fake-factors comes from varying the MC normalization. The kinematic variations associated with the tag-and-probe method do not significantly impact on the overall muon fake factors. Again, the resulting uncertainties can be large, surpassing 100% in several muon p_{T} bins.

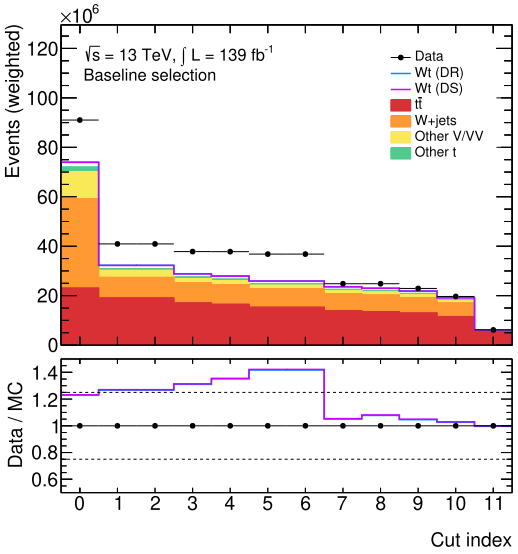
CHAPTER 7

Region Definitions

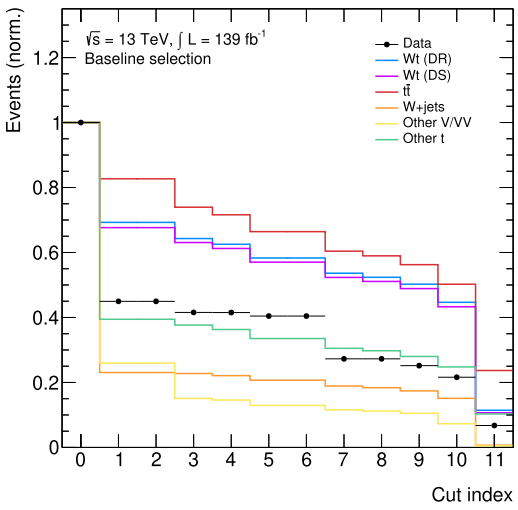
After reconstructing all necessary physics objects, the next step is applying additional requirements (cuts) using these objects to define specific measurement regions. The regions of interest where the final measurements will be performed are known as signal regions. In designing such regions, the selection criteria are chosen to maximize the number of signal events while removing as much background as possible. This analysis selects signal regions designed to maximize top processes, resulting in a $WWbb$ final state while reducing background contamination from $W+$jets, $Z+$jets, and diboson processes. This chapter describes the baseline section common to all signal regions explored in this analysis. The measurement regions are then described, and several kinematic variables of interest are shown as control plots.

7.1 Baseline Selection

The baseline selection refers to the initial requirements applied to all events. These remove the bulk of uninteresting events and filter out background events. An event must pass these requirements before more detailed signal selection can be applied. Figure 7.1 plots the cutflow histograms counting the number of MC and data events that pass each baseline requirement.



(a) Event weighted baseline selection cutflow histogram.



(b) Normalized baseline selection cutflow histogram.

Figure 7.1: Event Weighted/Normalized Baseline Selection Cutflow. The baseline selection refers to the initial set of requirements. The baseline selection cutflow counts the number of events that pass each of the cuts in the baseline selection. The main panel in **(a)** plots the number of (weighted) events passing each baseline cut for MC and data. The stacked MC is computed separately for both the DR and DS single-top samples. The bottom panel plots the ratio between the data and the total MC prediction. Again, this ratio is computed separately for DR and DS. The cutflow in **(b)** plots the normalized versions of the previous cutflow histograms.

Each bin in the above plots corresponds to the number of events that pass the cut referenced by the bin number. The following describes each of the baseline requirements, relevant for both the single-electron and single-muon channels.

(Bin 0) All events The initial number of data/MC events.

(Bin 1) $N_{jets} \geqslant 3$ Events must contain a minimum of three jets.

(Bin 2) Pass electron crack veto The region of $|\eta|$ between 1.37 and 1.52 is known as the "crack" region. This region defines the transition between the barrel and endcap calorimeters. This region contains cables and other services meant for the inner detector. Particles interacting with the passive material in this region are prone to producing non-prompt backgrounds through unwanted secondary interactions. These secondary interactions are not well-modeled, so vetoing electrons that fall into this region is common. Events are vetoed if an electron fails this requirement.

(Bin 3) One loose lepton: Events must contain at least one lepton (electron or muon) passing the loose identification and isolation requirements.

(Bin 4) Pass single lepton triggers Events must pass one of the single lepton triggers defined for the specific run period. Depending on the run year, these triggers impose a minimum lepton p_T of 24 or 26 GeV.

(Bin 5) Lepton is trigger matched The online leptons identified by the ATLAS trigger must match a reconstructed offline electron or muon candidate. The matching procedure involves comparing the kinematics and reconstruction qualities of the online and offline leptons. Trigger matching ensures that the particles the trigger identifies correspond to real physical leptons. All leptons in an event are required to be trigger-matched.

(Bin 6) Pass bad muon veto The reconstruction chain tags particular muons as "bad" due to poor reconstruction qualities. These muons have significantly worse momentum resolution than "good" muons. A very small number of muons fail to meet this requirement. However, accepting these bad muons can have a noticeably negative impact on certain analyses. Therefore, events with a bad muon are removed.

(Bin 7) One tight lepton Events must contain at least one lepton (electron or muon) passing the tight identification and isolation requirements.

(Bin 8) Exactly one lepton Events must contain exactly one lepton to be consistent with the semileptonic final state.

(Bin 9) Lepton $p_{\mathrm{T}} > 30$ GeV The single lepton in the event must have $p_{\mathrm{T}} \geqslant 30$ GeV.

(Bin 10) $m_{\mathrm{T}}^{W} + E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV This is the previously introduced anti-QCD cut aimed at reducing contributions from the main background process, W+jets.

(Bin 11) $\mathrm{N}_{b\text{-jets}} \geqslant 2$ Events must contain at least 2 b-jets. This selection is an inclusive b-jet selection where a range of b-jets multiplicities are accepted. On the other hand, an exclusive b-jet selection would require exactly 2 b-jets. A b-jet requirement is often imposed in many analyses to suppress background events that do not have the required number of b-jets. We also prefer the inclusive b-jet selection over the exclusive selection to allow for the possibility of additional b-jets produced from radiation. The inclusive selection increases the statistics of the final event yields.

Table 7.1 lists the number of (weighted) events that pass each of the baseline requirements. The table lists the event yields for each of the MC and data sources. Additional cutflow tables showing the absolute and marginal percentages are presented in Appendix B.

Table 7.1: Baseline Selection Weighted Event Yields. The weighted event yields for the MC and data samples after each baseline requirement are listed. Each entry has been scaled down by 10^3 for readability. The total summed MC is computed separately for the single-top DR and DS samples.

Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
All events	36140.1	10960.5	2001.5	1875.3	1771.9	23042.8	74020.1	73916.7	91034.0
$N_{jets} \geqslant 3$	8331.2	2843.9	789.5	1299.0	1198.7	19051.0	32314.6	32214.3	40929.4
Pass electron crack veto	8331.0	2843.8	789.5	1299.0	1198.7	19050.5	32313.8	32213.5	40929.4
One loose lepton	8223.5	1653.7	753.7	1205.7	1117.4	17038.0	28874.6	28786.3	37826.2
Pass single lepton triggers	7984.5	1598.2	726.5	1172.7	1085.1	16498.1	27980.0	27892.3	37795.8
Lepton is trigger matched	7483.1	1417.7	670.8	1093.5	1010.5	15304.9	25970.0	25887.1	36812.6
Pass bad muon veto	7482.9	1417.7	670.8	1093.5	1010.5	15304.6	25969.5	25886.6	36811.3
One tight lepton	6830.9	1266.8	610.6	1005.6	927.2	13921.4	23635.3	23556.9	24824.4
Exactly one lepton	6644.5	1226.8	595.5	982.1	905.3	13587.8	23036.7	22960.0	24824.4
Lepton $p_T > 30$ GeV	6291.3	1152.4	560.5	942.2	866.2	12961.7	21908.1	21832.0	22916.2
$m_T^W + E_T^{miss} > 60$ GeV	5462.7	798.9	496.0	837.6	767.0	11574.9	19170.0	19099.5	19670.1
$N_{b\text{-jets}} \geqslant 2$	243.4	53.2	206.3	214.6	189.2	5455.9	6173.4	6148.0	6148.2

7.2 Measurement Regions

Following the baseline cuts, the remaining sets of cuts define the individual measurement regions. This analysis is performed in three main distinct regions, characterized by cuts on the kinematics of the hadronic-W. These cuts are applied on both detector-level and particle-level so that the unfolded distributions represent the fiducial differential cross-sections. Multiple measurement regions are defined to observe the effects of varying the kinematic requirements of the hadronic-W on the unfolded $WWbb$ observables. A specific measurement region is also defined to measure differential cross-sections upon switching to a parton-level definition for specific observables. The following is a brief description of each of the three measurement regions.

$W_{\text{had.}}$ (particle) This is the nominal measurement region and employs the strictest cuts on the kinematics of the hadronic-W. The region is characterized by a semi-boosted hadronic-W, thereby maximizing the matching efficiency between detector-level and particle-level objects. There is an additional requirement that the mass of the reconstructed hadronic-W is found within a 20 GeV window of the true W mass to improve reconstruction efficiencies further (see Chapter 6.2.3). A particle-level definition of objects is used during the unfolding process.

$W_{\text{had.}}$ (loose) This region relaxes the kinematic cuts on the hadronic-W defined in "$W_{\text{had.}}$ (particle)". This region guarantees the largest acceptance and captures events in the low p_{T} regions of the hadronic-W. A particle-level definition of objects is used during the unfolding process.

$W_{\text{lep.}}$ (loose) This region applies the same cuts as "$W_{\text{had.}}$ (loose)" but adds additional requirements for the $E_{\text{T}}^{\text{miss}}$ and transverse momentum of the leptonic-W. These additional requirements further restrict the phase-space, targeting a more holistic picture of the $WWbb$ topology. A particle-level definition of objects is used during the unfolding process.

A fourth measurement region referred to as $W_{\text{had.}}$ (parton) has also been investigated. This region applies the same requirements on detector-level as "$W_{\text{had.}}$ (particle)" but instead uses a parton-level definition for certain observables during the unfolding process. The differences between parton and particle-level can then be examined for selected $WWbb$ observables. The observables measured in this "parton-level" region are shown in Appendix E.

Table 7.2 lists the number of (weighted) events that pass the various requirements in each of the three main measurement regions. The table lists the event yields for each of the MC and data sources. Additional cutflow tables showing the absolute and marginal percentages are presented in Appendix B.

Table 7.2: Measurement Regions Weighted Event Yields. The weighted event yields for the MC and data samples after each selection in the three measurement regions are listed. Each entry has been scaled down by 10^3 for readability. The total summed MC is computed separately for the single-top DR and DS samples.

				$W_{\text{had.}}$ (particle)					
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
Baseline events	243.4	53.2	206.3	214.6	189.2	5455.9	6173.4	6148.0	6148.2
$p_{\text{T}}^{W_{\text{had.}}} > 60$ GeV	96.4	19.2	88.3	106.4	86.6	2821.3	3131.5	3111.8	3038.1
$60 < M^{W_{\text{had.}}} < 100$ GeV	9.4	2.2	9.3	35.0	28.4	848.2	904.2	897.6	849.2
				$W_{\text{had.}}$ (loose)					
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
Baseline events	243.4	53.2	206.3	214.6	189.2	5455.9	6173.4	6148.0	6148.2
$p_{\text{T}}^{W_{\text{had.}}} > 30$ GeV	208.3	45.1	179.7	194.3	169.5	4973.2	5600.5	5575.8	5544.1
$M^{W_{\text{had.}}} > 40$ GeV	46.7	11.2	37.1	83.0	70.8	2193.7	2371.7	2359.5	2304.2
				$W_{\text{lep.}}$ (loose)					
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
$W_{\text{had.}}$ (loose)	46.7	11.2	37.1	83.0	70.8	2193.7	2371.7	2359.5	2304.2
$E_{\text{T}}^{\text{miss}} > 30$ GeV	40.0	8.9	32.9	74.6	63.2	1973.7	2130.2	2118.8	2060.6
$p_{\text{T}}^{W_{\text{lep.}}} > 30$ GeV	35.7	8.1	30.2	70.0	58.5	1819.2	1963.2	1951.8	1891.5

Table 7.3 lists the final (weighted) event yields for the three main measurement regions. This table also includes the contributions from misidentified leptons. As expected, regions with looser selections have higher event yields but also admit more background events coming from W+jets, Z+jets, and diboson processes. The tightest selection, "$W_{\text{had.}}$ (particle)", has sufficient statistics to perform the final measurements. Therefore, "$W_{\text{had.}}$ (particle)" is designated as the nominal measurement region. The other two regions are labeled as alternative regions.

Table 7.3: Measurement Regions Final Event Yields. The final (weighted) event yields are listed for the MC and data samples in the four measurement regions. Each entry has been scaled down by 10^3 for readability. The first row lists the contributions from misidentified leptons. The total summed MC is computed separately for the single-top DR and DS samples. The errors on each entry represent the statistical uncertainties only.

Sample ($\times 10^3$)	$W_{\text{had.}}$ (particle)	$W_{\text{had.}}$ (loose)	$W_{\text{lep.}}$ (loose)
Mis. ID leptons	14.7 ± 0.1	43.5 ± 0.1	35.9 ± 0.1
W+jets	9.4 ± 0.1	45.0 ± 0.2	35.7 ± 0.2
Other V/VV	2.2 ± 0.0	10.7 ± 0.1	8.1 ± 0.1
Other t	9.2 ± 0.1	35.5 ± 0.2	30.2 ± 0.2
tW (DR)	35.0 ± 0.2	81.0 ± 0.3	70.0 ± 0.3
tW (DS)	28.3 ± 0.2	69.0 ± 0.3	58.5 ± 0.2
$t\bar{t}$	846.0 ± 0.9	2137.8 ± 1.5	1819.2 ± 1.3
$\sum$ MC (DR)	916.4 ± 1.0	2353.5 ± 1.5	1999.1 ± 1.4
$\sum$ MC (DS)	909.8 ± 0.9	2341.5 ± 1.5	1987.6 ± 1.4
Data	846.6 ± 0.9	2241.1 ± 1.5	1891.5 ± 1.4

7.3 Selected Observables of Interest

The following variables are selected to characterize the $WWbb$ final state and are used to validate the agreement between simulation and data. These quantities will be later unfolded to extract the final differential cross-section distributions.

Lepton transverse momentum The transverse momentum of the single lepton (electron or muon). In plots, this observable is labeled as "Lepton p_{T}".

Leading b-jet transverse momentum The transverse momentum of the highest-p_{T} b-jet. In plots, this observable is labeled as "Leading b-jet p_{T}".

Subleading b-jet transverse momentum The transverse momentum of the second highest-p_{T} b-jet. In plots, this observable is labeled as "Subleading b-jet p_{T}".

Leading light-jet transverse momentum The transverse momentum of the leading (largest transverse momentum) light-jet. In plots, this observable is labeled as "Leading light-jet p_{T}".

Hadronic-W transverse momentum The transverse momentum of the reconstructed hadronic-W. In plots, this observable is labeled as "Hadronic-W p_{T}".

Selected ($b + \ell$) invariant mass The invariant mass of the leading lepton and the selected b-jet combination. This combination is sensitive to variations in the top quark mass and is often used as a proxy for the top quark. We avoid the two-fold combinatorial problems associated with reconstructing the leptonic-W. The b-jet is selected using the minimax algorithm described in Section 6.3. In plots, this observable is labeled as "$m_{\text{bl}}^{\text{selected}}$".

Selected ($b + W_{\text{had.}}$) invariant mass The invariant mass of the hadronic-W and the selected b-jet combination. The combination represents the hadronically-decaying top quark. The b-jet is selected using the minimax algorithm described in Section 6.3. In plots, this observable is labeled as "$m_{\text{bWhad}}^{\text{selected}}$".

Selected ($W_{\text{had.}} + b_1 + b_2 + \ell$) invariant mass The invariant mass of the hadronic-W, the leading b-jet, the subleading b-jet, and the lepton combination. This observable is sensitive to the full $WWbb$ final state and represents the invariant mass of the two top quarks. In plots, this observable is labeled as "$m_{\text{Whadbbl}}^{\text{selected}}$".

Modified minimax mass The modified minimax variable adapted from the fully-leptonic channel replaces the second lepton with the hadronic-W. This quantity is sensitive to interference effects between DR/DS. In most cases, one expects this minimax variable to be identical to the selected ($b + W_{\text{had.}}$) observable. The mass of the ($b + \ell$) pairing will almost always be less than the ($b + W_{\text{had.}}$) mass due to the missing neutrino. In plots, this observable is labeled as "$\text{minimax}_{\text{Whadbbl}}^{\text{selected}}$".

The choice of these observables is motivated by their sensitivity to mismodeling of the $WWbb$ final state. As discussed in Chapter 2, several of these observables have been predicted to NLO accuracy. Therefore, the agreement between the simulation and data for these observables is of particular interest. This analysis is also uniquely positioned to reconstruct the hadronic-W and study the distributions of combinations involving the hadronic-W. Whenever referenced, the term "kinematic observables of interest" refers to this variable list.

In addition to the main kinematic observables, we also choose to unfold several spatial quantities. This analysis also studies the following spatial variables.

Hadronic-W rapidity The rapidity of the reconstructed hadronic-W. In plots, this observable is labeled as "Hadronic-W y".

Selected $\Delta R(b, \ell)$ The ΔR separation between the selected b-jet and the leading lepton. In plots, this observable is labeled as "ΔR between selected bjet and ℓ" and "Selected $\Delta R(b, \ell)$".

Selected $\Delta R(b, W_{\text{had.}})$ The ΔR separation between the selected b-jet and the hadronic-W. In plots, this observable is labeled as "ΔR between selected b-jet and $W_{\text{had.}}$" and "Selected $\Delta R(b, W_{\text{had.}})$".

Recall the rapidity is defined as $y = \frac{1}{2} \ln \left(\frac{E+p_z}{E-p_z} \right)$, where E is the energy and p_z is the momentum in the z-direction. Differences in rapidity are Lorentz-invariant with respect to boosts along teh beam

axis. The rapidity is often used to describe the spatial distribution of particles. A rapidity distribution is sensitive to the boost of the system; a rapidity close to 0 indicates a particle with large transverse momentum, while a rapidity tending to $+\infty$ or $-\infty$ indicates a particle moving in the forward or backward direction close to the beamline. This observable is studied for the hadronic-W to assess whether there is a potential bias in predictions of the location of the hadronic-W.

The ΔR separation is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, where $\Delta\eta$ is the difference in pseudorapidity and $\Delta\phi$ is the difference in azimuthal angle. The ΔR separation measures the angular separation between two objects. The ΔR separation is studied for the selected b-jet and lepton, and the selected b-jet and hadronic-W. These observables are sensitive to the kinematics of the top-quark decay. A small ΔR separation indicates the objects are close in phase-space, while a large ΔR separation indicates the objects are far apart. Whenever referenced, the term "spatial observables of interest" refers to this variable list.

7.4 Detector-Level Control Plots

Figure 7.2 plots the detector-level distributions for the kinematic and spatial observables of interest. These variables are only shown for the nominal $W_{\text{had.}}$ (particle) region. The control plots for the alternative signal regions are shown in Appendix C.1.

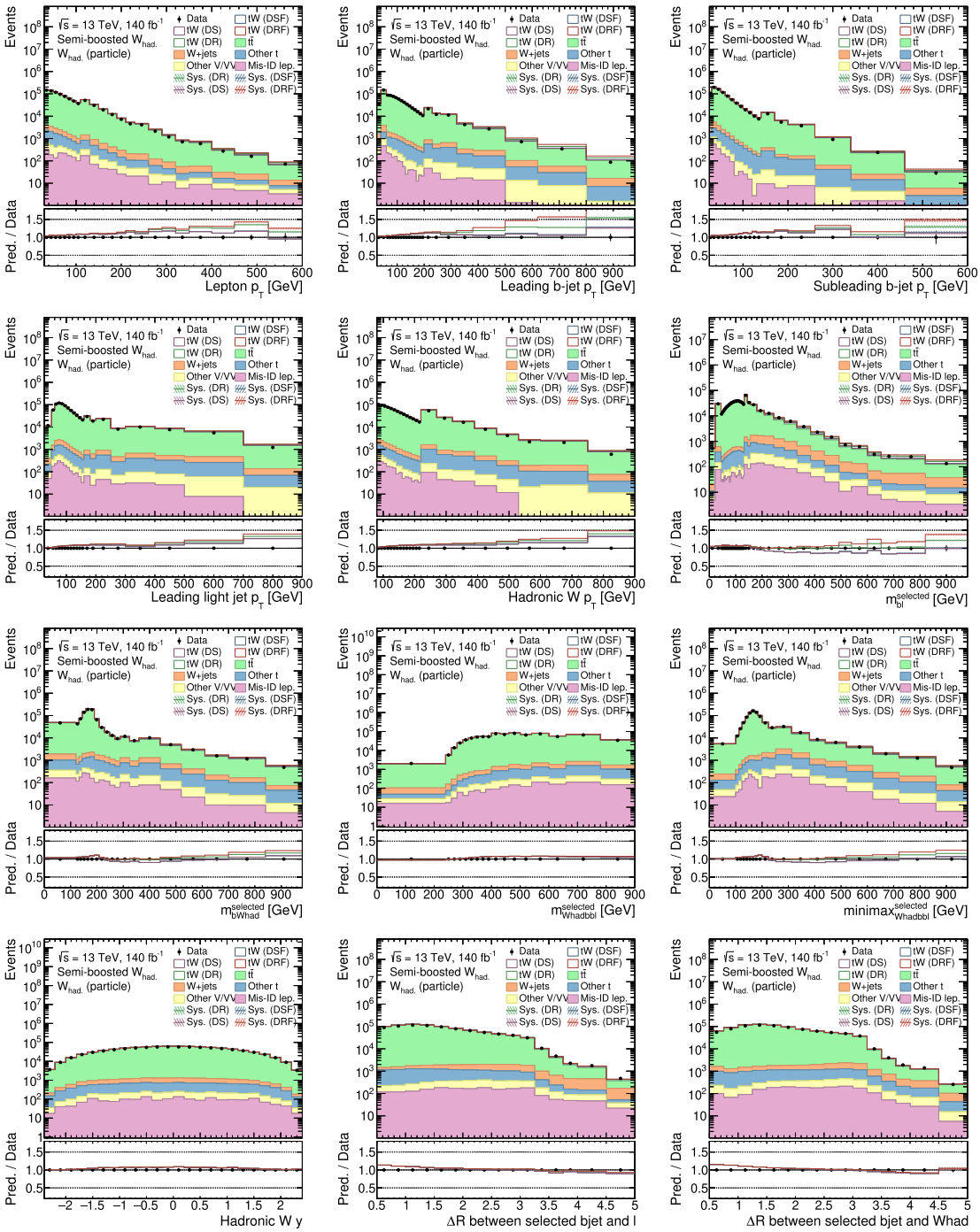


Figure 7.2: Detector-Level Observables in the $W_{\text{had.}}$ (particle) Region. The detector-level control distributions of the observables of interest plot the full Run 2 data and the expected MC contributions. The shaded bands represent the statistical uncertainties in the total MC prediction. The bottom panel in each plot shows the ratio between the data and the MC prediction, computed separately for the DR and DS single-top samples. Refer to Section 7.3 for a description of the observables of interest.

Several observations can be drawn from the detector-level control plots. The agreement between the simulation and data is generally good for most of the observables. The agreement only starts to degrade in the tails of the distributions, where mismodeling effects start to play a larger role. In these regions, the simulation tends to overestimate the data. The $m_{b\ell}$ distribution is of particular interest as it shows the DS prediction underestimating the data while the DR model overestimates the data. It is expected that the DR and DS techniques yield different results in terms of agreement with the observations. Other NLO predictions studying the same interference effect have also reported significant differences between the two techniques [126]. For this measurement, the DS model for single-top production models the data better than the DR model. When comparing the fixed-scale and dynamically-scaled single-top samples, the dynamically-scaled sample generally models the data better. Given the discussion in Section 5.4, the better performance when using dynamic scales is expected.

CHAPTER 8

Systematic Uncertainties

In addition to statistical uncertainties, many systematic uncertainties can affect the final measurement. However, unlike statistical uncertainties, one cannot decrease systematic uncertainties by taking more measurements. It is necessary to understand the response of the measurement to these systematic variations. We divide these uncertainties into two broad categories: experimental and theoretical. Experimental uncertainties are associated with the limitations of the apparatus and any experimental techniques. The beam conditions of the LHC can also contribute to experimental uncertainty. Theoretical uncertainties are associated with the modeling of the physics processes. This chapter will describe the sources of systematic uncertainties.

8.1 Experimental Systematics

Experimental systematics quantify the limited precision and accuracy of the detector subsystems in the simulated model of the ATLAS detector. Each physics object (e.g., leptons, jets, b-tagging, and missing transverse energy) carries its own systematic uncertainties. These are associated with the techniques used to reconstruct and calibrate the objects. A source of uncertainty is captured in the form of a Nuisance Parameter (NP). To measure the effect of these uncertainties, the nominal value of the observable is varied by $\pm 1\sigma$ of the corresponding NP. The measurement is then repeated using these varied values, and the difference between the nominal and the varied measurement is taken as the systematic uncertainty. For a binned quantity, the relative systematic uncertainty on bin i contributed by a single NP (ϵ_i^{NP}) is computed as shown in Eq. 8.1.

$$\epsilon_i^{\mathrm{NP}} = |h_i^{\mathrm{NP}} - h_i^{\mathrm{nom.}}|/h_i^{\mathrm{nom.}} \tag{8.1}$$

In the above, h_i^{NP} and $h_i^{\mathrm{nom.}}$ corresponds to the bin heights of bin i for the varied and nominal measurements, respectively. The total systematic uncertainty ($\epsilon_i^{\mathrm{tot.}}$) is then computed by summing the contributions of the NPs in a single category in quadrature. This procedure assumes that the systematic uncertainties are largely uncorrelated between the different sources. Correlations in uncertainties that impact the shape or normalization of the distributions, such as luminosity and PDF uncertainties,

are expected to be small and are disregarded due to the magnitude of these uncertainties.

$$\epsilon_i^{\text{tot.}} = \sqrt{\sum_j (\epsilon_i^{\text{NP}_j})^2} \tag{8.2}$$

In the above, the inner sum j runs over all of the NPs in a specific category of experimental uncertainties. For this analysis, the following categories of systematics uncertainties are considered alongside a list of all NPs under each category.

Electrons The energy resolution and scale of simulated electrons are adjusted to account for differences between data and simulation. A correction factor is applied to each electron four-momentum to account for these differences. Efficiencies associated with electron identification, isolation, and reconstruction are also considered. Corrections to efficiencies are applied as scale factors to the simulated events, which affects the overall event weight normalization. Lastly, there are also uncertainties related to the efficiencies of the electron triggers. Table 8.1 lists the NPs associated with electrons.

Table 8.1: Electron Uncertainties Nuisance Parameters. The NPs associated with electron uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
`EG_RESOLUTION_ALL`	Energy resolution uncertainty
`EG_SCALE_ALL`	Energy scale uncertainty
`EG_SCALE_AF2`	
`EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR`	ID efficiency uncertainty
`EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR`	Isolation efficiency uncertainty
`EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR`	Reconstruction efficiency uncertainty
`EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR`	Trigger efficiency uncertainty
`EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR`	Trigger scale-factor uncertainty

Muons Similar to electrons, the simulated muons' energy resolution, and scale are adjusted to account for differences between data and simulation. A correction factor is applied to each muon four-momentum to account for these differences. Efficiencies associated with muon isolation, reconstruction, track-to-vertex association, and bad muon veto are also considered. Corrections to efficiencies are applied as scale factors to the simulated events, which affects the overall event weight normalization. Lastly, there are uncertainties related to the efficiencies of the muon triggers. Table 8.2 lists the NPs associated with muons.

Table 8.2: Muon Uncertainties Nuisance Parameters. The NPs associated with muon uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
MUON_SAGITTA_DATASTAT	Energy scale uncertainty
MUON_SAGITTA_RESBIAS	
MUON_EFF_Iso_STAT	Isolation efficiency uncertainty
MUON_EFF_Iso_SYS	
MUON_EFF_Reco_STAT	Reconstruction efficiency uncertainty
MUON_EFF_Reco_SYS	
MUON_EFF_TTVA_STAT	Track-to-vertex association efficiency uncertainty
MUON_EFF_TTVA_SYS	
MUON_EFF_BADMUON_SYS	Bad muon veto efficiency uncertainty
MUON_EFF_TrigStatUncertainty	Trigger efficiency uncertainty
MUON_EFF_TrigSystUncertainty	

Jets The jets' energy scale and resolution uncertainties are evaluated by varying the jet energies in situ. The uncertainty in the jet energy scale is captured in 28 NPs, which aims to assess various effects that could affect jet energy calibration. Such effects include the pile-up conditions, differences in response due to jet flavor, and differences in response due to the jet p_{T} and η. The uncertainty in the jet energy resolution is evaluated using 12 independent NPs. A thirteenth jet energy resolution NP is used to correct for differences between simulation and data. Finally, since the JVT algorithm is used to identify hard-scatter jets, a NP is included to account for the uncertainty in the JVT efficiency. The NPs associated with jets are listed in Table 8.3.

Table 8.3: Jet Uncertainties Nuisance Parameters. The NPs associated with jet uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
JET_EffectiveNP_Detector[1-2]	Energy scale uncertainty
JET_EffectiveNP_Mixed[1-3]	
JET_EffectiveNP_modeling[1-4]	
JET_EffectiveNP_Statistical[1-6]	
JET_EtaIntercalibration_modeling	
JET_EtaIntercalibration_NonClosure_2018data	
JET_EtaIntercalibration_NonClosure_highE	
JET_EtaIntercalibration_NonClosure_negEta	
JET_EtaIntercalibration_NonClosure_posEta	
JET_EtaIntercalibration_TotalStat	
JET_Flavor_Composition_prop	
JET_Flavor_Response_prop	
JET_Pileup_OffsetMu	
JET_Pileup_OffsetNPV	
JET_Pileup_PtTerm	
JET_Pileup_RhoTopology	
JET_PunchThrough_MC16	
JET_JER_EffectiveNP_[1-12]	Energy resolution uncertainty
JET_JER_DataVsMC_MC16	
JET_JvtEfficiency	JVT efficiency uncertainty

***b*-tagging** The DL1r algorithm is used in this analysis to tag b-jets. Uncertainties associated with b-tagging are evaluated separately for b-jets, c-jets, and light-jets with independent sets of NPs for each type of jet. There are 9 such NPs for b-jets, 4 for c-jets, and 4 for light-jets. Finally, there are two additional NPs to account for efficiencies associated with uncertainties on the extrapolation on high-p_{T} jets and tau-jets (τ-jets). All of these uncertainties enter as scale-factors applied to the simulated events. The NPs associated with b-tagging are listed in Table 8.4.

Table 8.4: *b*-tagging Uncertainties Nuisance Parameters. The NPs associated with b-tagging uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
FT_EFF_Eigen_B_[0-8]	b-jets efficiency component uncertainty
FT_EFF_Eigen_C_[0-3]	c-jets efficiency component uncertainty
FT_EFF_Eigen_Light_[0-3]	Light-jets efficiency component uncertainty
FT_EFF_extrapolation	Extrapolation on high-p_{T} jets uncertainty
FT_EFF_extrapolation_from_charm	Extrapolation on τ-jets uncertainty

Missing Transverse Energy The uncertainty in the scale and resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is estimated by varying the soft-track component in calculating $E_{\mathrm{T}}^{\mathrm{miss}}$. Variations of these parameters can affect the overall value of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and its direction. The NPs associated with the $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties are listed in Table 8.5.

Table 8.5: MET Uncertainties Nuisance Parameters. The NPs associated with $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
`MET_SoftTrk_Scale`	Momentum scale uncertainty
`MET_SoftTrk_ResoPerp`	Momentum resolution uncertainty
`MET_SoftTrk_ResoPara`	

Pileup For a simulated event, the pileup must also be simulated to match the same conditions in the real ATLAS detector. In practice, the simulated distribution of the average number of interactions per bunch crossing does not match what is observed in the data. To improve the agreement of the distributions between data and simulation, a procedure known as "pileup reweighting" is applied. This procedure corrects the distribution of the average number of interactions per bunch crossing in the simulated events by applying certain scale factors. The uncertainty associated with this procedure is evaluated by varying the scale factors and is captured in a single NP, `PRW_DATASF`.

Luminosity The total integrated luminosity normalizes the simulated events to the full Run 2 luminosity. Therefore, a precise measurement of the total integrated luminosity and its corresponding uncertainty is required. Table 5.1 presented the integrated luminosity for each year during the Run 2 data-taking period. The total integrated luminosity for Run 2 was 140.07 ± 1.17 fb^{-1}, corresponding to a relative uncertainty of approximately 0.83%. This constant uncertainty is propagated throughout the remaining analysis.

8.2 Theory Systematics

Theory systematics quantify the uncertainties associated with the modeling of the physics processes. These uncertainties are typically associated with the choice of the generator used to simulate the events and the specific parameters used in the simulation. The theory systematics are evaluated by varying the parameters during event generation and comparing the results to the nominal simulation. For this analysis, we categorize theory systematics into two categories: signal modeling systematics and background modeling systematics. The former relates to uncertainties associated with the modeling of the signal processes ($t\bar{t}$ and tW) while the latter relates to uncertainties associated with the modeling of the background processes.

8.2.1 *Signal modeling*

The following categories associated with signal modeling systematics are considered alongside a list of all NPs under each category.

Cross-Section The uncertainty in the predicted cross-sections of the $t\bar{t}$ and tW processes (listed in Table 5.2) is evaluated by varying the normalization factor of the MC samples. We adjust the normalization according to the uncertainties in the predicted cross-sections. This effect is captured in a single NP, `THEORY_CROSS_SECTION_signal`.

Parton Showering The uncertainty in the parton showering process is evaluated by adjusting the $p_{\mathrm{T,hard}}$ parameter in the POWHEG generator [86]. This parameter is associated with the probability of suppressing further emissions during the showering process. In general, higher values of $p_{\mathrm{T,hard}}$ are associated with suppressing further emissions. To evaluate the effect of this parameter, the $p_{\mathrm{T,hard}}$ parameter is increased from nominally 0 to 1. This uncertainty is captured in the NP, `THEORY_PTHARD_signal`.

Hadronization Model The uncertainty in the hadronization process is evaluated by comparing the results of the nominal simulation to the results using an alternative hadronization model. The alternative hadronization model is evaluated using the POWHEG + HERWIG 7.1 generator and is captured in the NP, `THEORY_SHOWERING_HERWIG7_signal`.

μ_r and μ_f Scales The renormalization and factorization scales are nonphysical quantities introduced during cross-section calculations to absorb the divergences in the perturbative expansion. This analysis considers both the fixed and dynamically-scaled μ_r and μ_f. During event generation, the μ_r and μ_f are associated with the production of initial/final state radiation (ISR and FSR). To evaluate the uncertainties associated with μ_r and μ_f, the nominal values are multiplied by a factor of 2 and 0.5 to obtain the upper and lower variations. The NPs associated with the μ_r and μ_f scales are listed in Table 8.6.

Table 8.6: μ_r and μ_f Scales Uncertainties Nuisance Parameters. The NPs associated with μ_r and μ_f scales uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
`THEORY_SCALE_FACTORISATION_signal`	μ_f uncertainty
`THEORY_SCALE_RENORMALISATION_signal`	μ_r uncertainty
`THEORY_ISR_signal`	ISR uncertainty
`THEORY_FSR_signal`	FSR uncertainty

h_{damp} The h_{damp} parameter is used to regulate the $t\bar{t}$ production in the generator by controlling the matching of the $t\bar{t}$ matrix element to the parton shower. The uncertainty in the h_{damp} parameter is evaluated by multiplying the nominal value by 2. This variation corresponds to the NP, `THEORY_HDAMP_signal`.

Top Recoil The top-quark decays into a W-boson and a b-quark, $t \rightarrow W + b$. Gluons can be radiated during this decay. After the first gluon radiation, all secondary gluon emissions are modeled as recoiling against the b-quark (recoiler). This procedure is implemented in the nominal POWHEG+PYTHIA 8 generator [127]. This procedure tends to produce radiation that is too collinear with the b-quark, resulting in jets that are too narrow. To improve the modeling of the radiated jets, a modified approach is to use the top quark as the recoiler. This model changes the shape of the radiated b-jets and results in a narrower top-mass distribution [128]. This analysis uses an alternative sample implementing the top recoil procedure. The differences due to top recoil is captured in the NP, `THEORY_TOPRECOIL_signal`.

PDF The choice of the specific PDF set may affect the final unfolded cross-section distributions. To evaluate the uncertainty associated with the PDF set, this analysis adopts the procedures and PDF sets provided by the PDF4LHC15 collaboration [129]. The PDF4LHC15 combined PDF set contains replicas of many of the most commonly used PDF sets. In practice, one could calculate the predicted cross-section for each replica and compare the results when using the nominal PDF set. To reduce computational time, it is possible to compute the PDF uncertainties by instead reweighting the nominal PDF set, such that an alternative event weight is obtained. The uncertainty is then evaluated by comparing the nominally and alternatively weighted events. This uncertainty is captured in the NP, `THEORY_PDF4LHC_VARIATION_signal`.

DR/DS This analysis evaluates the differences between the DR and DS approaches to remove the overlap between the $t\bar{t}$ and tW samples. The uncertainty is measured by taking the envelope (highest/lowest values) from among the different DR and DS samples as the total uncertainty. This uncertainty is captured in the NP, `THEORY_DR_DS_signal`.

***W*-Mass** The accepted mass of the W-boson is used in the reconstruction of the leptonic and hadronic W-bosons. Using a hard-coded mass may result in a bias towards reconstructions that are biased towards the accepted mass. To evaluate the uncertainty associated with the W-mass, the true mass of the W-boson is varied by $\pm 2 \times \sigma_W$ where $\sigma_W = 2.085$ GeV corresponds to the measured width of the W-mass distribution [3]. The associated NP is `WMASS_VAR_signal`.

8.2.2 Background modeling

The following categories associated with background modeling systematics are considered alongside a list of all NPs under each category.

***W*+ jets** Similar to the signal modeling systematics, the uncertainty in the W+ jets background is evaluated by varying the normalization factor of the MC generation parameters. These variations include modeling uncertainties on the cross-section, PDFs, scales and electroweak corrections. Table 8.7 lists the NPs associated with the W+ jets background modeling.

Table 8.7: W+jets Modeling Uncertainties Nuisance Parameters. The NPs associated with W+jets modeling uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
`THEORY_CROSS_SECTION_Wjets`	Cross-section uncertainty
`THEORY_SCALE_COMBINED_Wjets`	Scales uncertainties
`THEORY_PDF4LHC_VARIATION_Wjets`	Choice of PDF set uncertainty
`THEORY_EWK_Wjets`	Electroweak correction uncertainty

Other V/VV The scales, PDF, and electroweak correction uncertainties are also assessed for the multiboson samples. Table 8.8 lists the NPs associated with the multiboson background modeling.

Table 8.8: Other V/VV Modeling Uncertainties Nuisance Parameters. The NPs associated with multiboson modeling uncertainties are listed alongside a brief description.

Nuisance Parameter	Description
`THEORY_SCALE_COMBINED_multiboson_noW`	Scales uncertainties
`THEORY_PDF4LHC_VARIATION_multiboson_noW`	Choice of PDF set uncertainty
`THEORY_EWK_multiboson_noW`	Electroweak correction uncertainty

Other t A constant 25% uncertainty is applied to the normalization of the non-resonant top samples. This uncertainty is captured in the "cross-section" NP, `THEORY_CROSS_SECTION_other_top_noWt`.

Fake Leptons The uncertainty in the misidentified leptons was conducted by varying aspects of the MC or cuts. The uncertainties were assessed independently for electrons and muons as shown in Section 6.4.5. The corresponding fake leptons NPs are `FAKES_Electron` and `FAKES_Muon`.

8.3 Systematics Summary Plots

For the observables introduced in the previous section, the bin-by-bin systematics are evaluated for each of the main categories listed above. The "total uncertainty" is computed by summing all other contributions in quadrature, and this is shown as a gray band in each plot. The "expected statistical" uncertainty is shown as a dark gray band and is the statistical uncertainty from data. Figure 8.1 plots the systematic uncertainties for the kinematic and spatial observables of interest in the $W_{\text{had.}}$ (particle) region. The corresponding systematics for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions are presented in Appendix C.2.

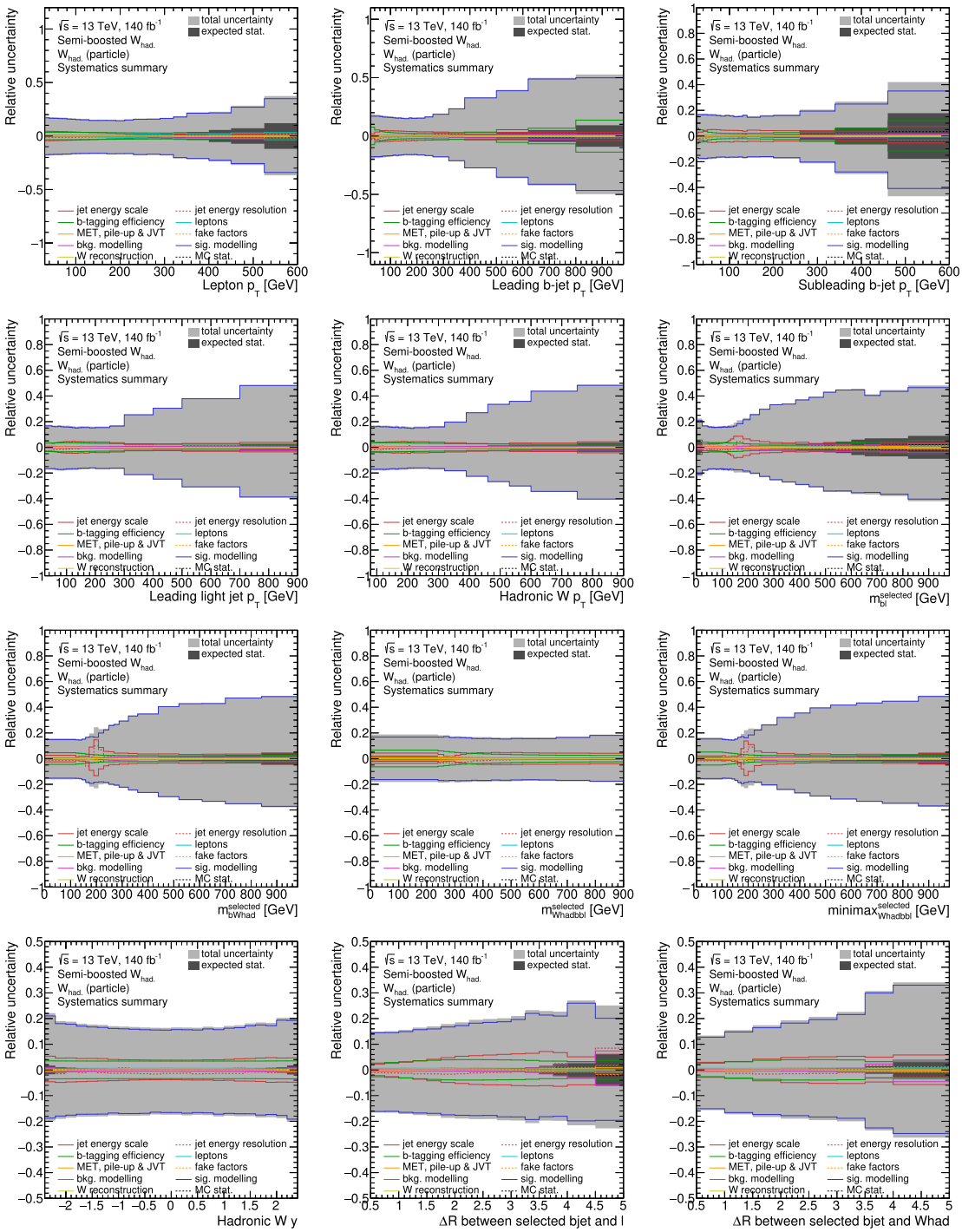
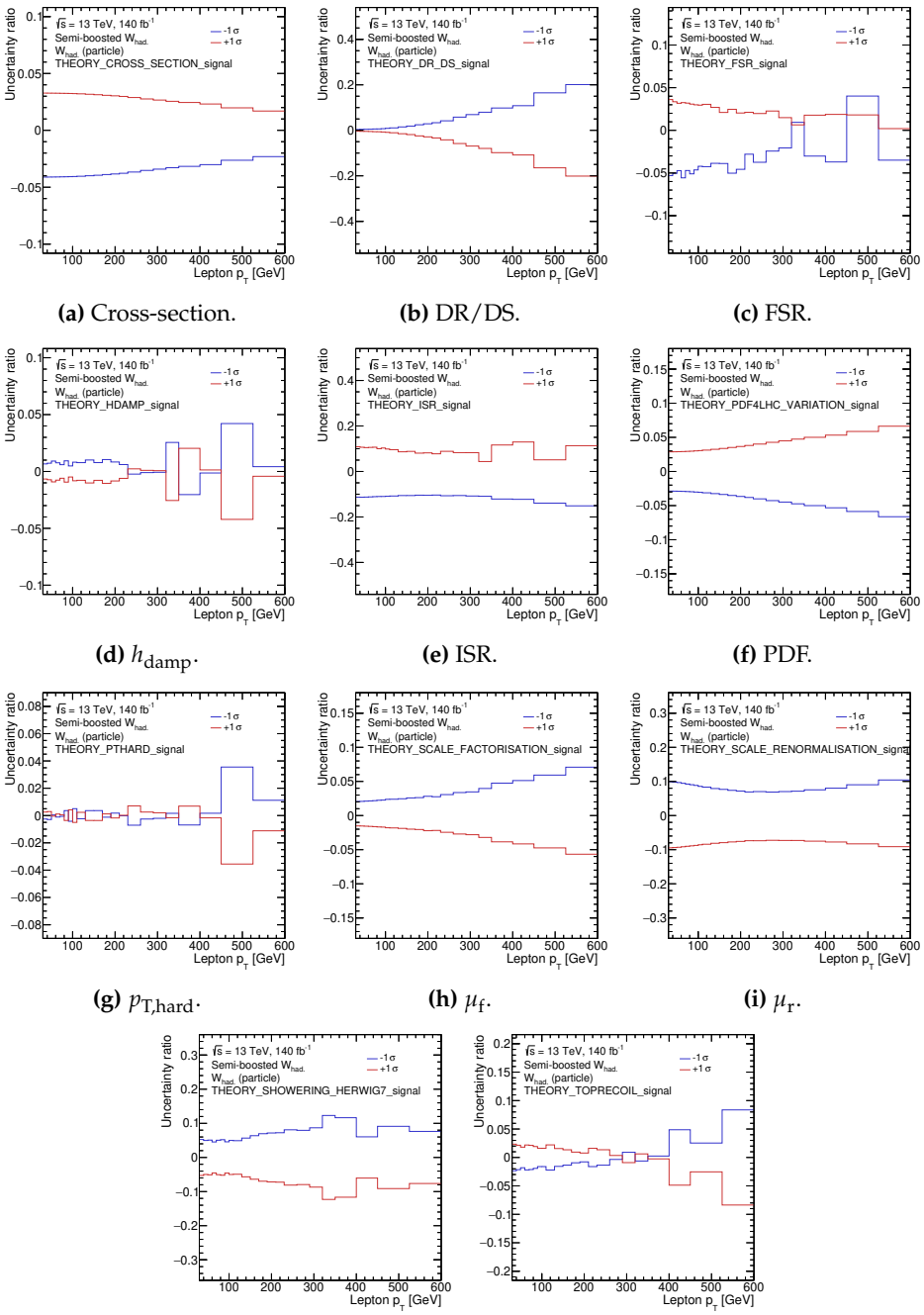


Figure 8.1: Systematic Uncertainties Summary for the $W_{\text{had.}}$ (particle) Region. Systematic uncertainties include contributions from both detector-level and modeling uncertainties. The total systematic uncertainty is computed by adding the statistical and systematic uncertainties in quadrature. Refer to Section 7.3 for a description of the observables of interest.

The above plots correspond to the systematic uncertainties prior to unfolding. In each of the presented observables, signal model is the dominant source of uncertainty. The other sources of uncertainty generally do not exceed 10%, while the signal modeling uncertainties increase past 50% in the tails of these distributions. Figure 8.2 plots the contributions of the NPs associated with the signal modeling uncertainty. For each NP, the ratio of the distributions of the up/down variations to the nominal distribution is calculated. For illustration purposes, only the lepton transverse momentum plot in the $W_{\text{had.}}$ (particle) region is shown. The other regions have similar contributions and are shown in Appendix C.3.



(a) Cross-section. **(b)** DR/DS. **(c)** FSR.

(d) h_{damp}. **(e)** ISR. **(f)** PDF.

(g) $p_{\text{T,hard}}$. **(h)** μ_{f}. **(i)** μ_{r}.

(j) Showering. **(k)** Top recoil.

Figure 8.2: Summary of Signal Modeling Uncertainties for the $W_{\text{had.}}$ (particle) Region. Signal modeling uncertainties make up the largest contribution to the total systematic uncertainty. The contributions of the NPs associated with the signal modeling uncertainties are computed for each bin as a ratio of the up/down variations to the nominal distribution. The following NPs associated with the signal modeling uncertainty are plotted: **(a)** cross-section, **(b)** DR/DS, **(c)** FSR, **(d)** h_{damp}, **(e)** ISR, **(f)** PDF, **(g)** $p_{\text{T,hard}}$, **(h)** μ_{f}, **(i)** μ_{r}, **(j)** parton-showering, and **(k)** Top recoil.

The most significant source of uncertainty in the signal modeling category comes from DR and DS uncertainties, which can be as large as 30% of the total uncertainty. Recall that the DR/DS uncertainty is calculated by taking the envelope of the various DR and DS samples. Therefore, this large uncertainty is likely contributed by the fixed-scale DR/DS samples, which were shown to disagree with predictions more so than their dynamically-scaled counterparts. Therefore, the final cross-sections can be particularly sensitive to the modeling of the interference effects between the $t\bar{t}$ and tW samples. A dedicated MC generator including these interference effects (e.g., *bb4l*) could reduce this uncertainty.

In the case of the detector uncertainties, the dominant uncertainty depends on the observables. For the lepton transverse momentum, at higher p_{T} (> 300 GeV) ranges, uncertainties associated with the fake factors become more dominant. In these high-p_{T} regions, the statistical uncertainty also increases to greater than 20%. For the b-jet and light-jet transverse momentum observables, the dominant uncertainty naturally comes from the b-tagging efficiency. For composite observables, the jet energy scale and resolution uncertainties begin to play a more significant role. This is particularly true for the selected $(b + \ell)$, $(b + W_{\mathrm{had.}})$, and the modified minimax mass distributions. For these distributions, there is a clear increase in the jet energy scale uncertainties around the 200 GeV. It is assumed this arises because of differences in the uncertainties of the jet energy scale and resolution in different jet p_{T} ranges [130].

When comparing the uncertainties across the different regions (see Appendix C.2), the scales of the total uncertainties are relatively similar. For observables such as the lepton and b-jet transverse momenta with large statistical uncertainties in the tails, the total uncertainty of the stricter $W_{\mathrm{had.}}$ (particle) region is slightly larger than the other two looser regions. However, the effects of the W-mass variation are more pronounced in the looser regions. The $W_{\mathrm{had.}}$ (loose) and $W_{\mathrm{lep.}}$ (loose) regions have a simple $p_{\mathrm{T}}^{W} > 40$ GeV cut while the $W_{\mathrm{had.}}$ (particle) region uses a $60 < p_{\mathrm{T}}^{W} < 100$ GeV window to improve the overall reconstructed-to-particle level matching. A lower cut allows potentially misreconstructed W-bosons to enter the selection. For example, there may be events carrying a single light-jet. By default, this single jet will be reconstructed as the hadronic-W despite not having a mass close to the true W-mass.

CHAPTER 9

Unfolding

The detector-level distributions are subject to various detector effects such as acceptance, efficiency, resolution, misreconstruction, and misidentification. Unfolding refers to the mathematical procedures used to correct these effects to obtain the true particle-level distributions from the observed detector-level distributions. Unfolding is a necessary step to compute the various differential cross-section distributions that will be used to compare with various theoretical predictions. Different packages exist to perform unfolding, but in this analysis, the TUnfold package [131] is used. The unfolding procedure seems straightforward in principle, requiring only a few ingredients. However, in practice, challenges can arise in constructing the migration matrix, selecting the regularization parameter, and propagating the systematic uncertainties. This chapter begins by describing the unfolding procedure, the required ingredients, and the inherent challenges of unfolding. Next, we will discuss the unfolding process as it pertains to this analysis. Finally, we will present the unfolded differential cross-sections and the associated uncertainties compared to various theoretical predictions.

9.1 Unfolding Description

In high-energy physics, measurements are generally performed as counting experiments. Observations are grouped into bins of a certain observable, such as the transverse momentum of a particle. However, the detector is subject to various inherent inefficiencies and resolutions that can distort the true underlying particle-level distributions. Unfolding aims to correct for these detector effects and obtain a detector-independent distribution. The unfolding problem can be formulated as a linear inverse problem, which may be written as follows.

$$y_i = \sum_{j=1}^{m} A_{ij} x_j + b_i,\ 1 \leqslant i \leqslant n \tag{9.1}$$

In the above, y_i is the observed number of counts in bin i (with n total bins), A_{ij} is the migration matrix which measures the probability that an event in bin j (with m total bins) at the particle-level is reconstructed in bin i at the detector level, x_j is the true number of events in bin j, and b_i is the background contribution in bin i. The goal of unfolding is to determine the true number of events x_j given

the observed counts y_i. The migration matrix A_{ij} is a key ingredient in the unfolding procedure and is determined using MC simulations. It encodes the detector response and relates the true particle-level distributions to the observed detector-level distributions[1]. The migration matrix's diagonal elements represent the purity, i.e., the probability that an event in a given particle-level bin is correctly reconstructed in the corresponding detector-level bin. The off-diagonal elements represent the migration of events between different bins. A perfect detector would have a diagonal migration matrix.

As written, unfolding seems to be a trivial matrix inversion problem requiring only a few input vectors and a single matrix. However, unfolding is inherently an ill-posed problem [132]. A direct solution via simple matrix inversion would correctly yield the desired true distribution.

$$\hat{\vec{x}} = A^{-1}\hat{\vec{y}} \tag{9.2}$$

Here, $\hat{\vec{x}}$ is an estimator of the underlying true distribution, while $\hat{\vec{y}}$ is the observed distribution with added statistical fluctuations (e) such that $\hat{\vec{y}} = \vec{y} + \vec{e}$. We can prove that the solution from this approach is statistically sound and does not introduce any unwanted biases. In other words, the expectation value of the estimator of $\vec{x}$ should yield the true underlying distribution, $\vec{x}$.

$$\begin{aligned} E[\hat{\vec{x}}] &= A^{-1}E[\hat{\vec{y}}] \\ &= A^{-1}E[\vec{y} + \vec{e}] \\ &= A^{-1}E[\vec{y}] + A^{-1}E[\vec{e}] \\ &= A^{-1}E[\vec{y}] \\ &= A^{-1}AE[\vec{x}] \\ &= \vec{x} \end{aligned} \tag{9.3}$$

In going from the third line to the fourth line, we have used the fact that the expectation value of the statistical fluctuations is zero, $E[\vec{e}] = 0$. This method does not yield satisfactory results since it can produce unstable outputs with large statistical fluctuations. Relatively small changes in the input vector or migration matrices can result in large changes in the unfolded distributions. Moreover, matrix inversion can often result in negative terms in the inverted matrix, potentially leading to a non-physical negative event count in one or more bins. Lastly, a priori, it cannot be assumed that the inverse of the migration matrix should exist. Often, migration matrices tend to be large and sparse and, therefore, can be singular.

9.1.1 Regularization

Regularization techniques are added to address the issues associated with simple matrix inversion. Regularization involves adding constraints to the unfolding problem to stabilize the solution and remove unwanted large fluctuations. In doing so, an unavoidable bias will be introduced into the final

[1] The least squares minimization employed by TUnfold requires a non-zero number of degrees of freedom, $n - m$. Therefore, it is required that the data y_i is measured in finer bins than the final particle-level distributions, x_j.

unfolded result. The size of the bias depends on the strength of the regularization. In TUnfold, the regularization is implemented as a penalty term (τ), which controls the strength of the regularization. If τ is chosen to be too large, the bias skews the unfolded distribution away from the true particle-level spectrum. If τ is chosen to be too small, the unfolded distribution will be unstable and exhibit large fluctuations. TUnfold formulates this optimization problem as a minimization of the following Lagrangian.

$$\begin{aligned} \mathcal{L}(x,\lambda) &= \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \\ \mathcal{L}_1 &= (\vec{y} - A\vec{x})^{\mathrm{T}} V_{yy}^{-1} (\vec{y} - A\vec{x}) \\ \mathcal{L}_2 &= \tau^2 (\vec{x} - f_b \vec{x}_0)^{\mathrm{T}} (L^{\mathrm{T}} L)(\vec{x} - f_b \vec{x}_0) \\ \mathcal{L}_3 &= \lambda (Y - \vec{e}^{\mathrm{T}} \vec{x}) \end{aligned} \tag{9.4}$$

The first Lagrangian, $\mathcal{L}_1$ represents the standard least-squares minimization between the observed distribution and the predicted distribution connected by the covariance matrix of the observations, V_{yy}. In many cases, the covariance matrix is diagonal and holds the squares of the total uncertainties.

The second Lagrangian, $\mathcal{L}_2$ encodes the regularization whose strength is determined by the parameter τ. The bias vector $f_b \vec{x}_0$ consists of a normalization factor, f_b and a reference vector, $\vec{x}_0$. Depending on the choice of f_b, the regularization can reduce deviations of $\vec{x}$ from 0 or $\vec{x}_0$. L is a matrix that encodes the regularization conditions. There are three basic types of regularization conditions implemented in TUnfold.

Size TUnfold is set to regularize the amplitude of the unfolded distribution. L is initialized to be the identity matrix with size $n_{\vec{x}}$.

First Derivative TUnfold is set to regularize the unfolded distribution's first derivative (slope). L is initialized with $L_{i,i} = -1$ and $L_{i,i+1} = 1$ and has $n_{\vec{x}} - 1$ rows.

Second Derivative TUnfold is set to regularize the unfolded distribution's second derivative (concavity). L is initialized with $L_{i,i} = 1$, $L_{i,i+1} = -2$, and $L_{i,i+2} = 1$ and has $n_{\vec{x}} - 2$ rows.

The simplest case corresponds to the regularization of the size of the unfolded distribution with $f_b = 0$ and L being the identity matrix. In this case, $\mathcal{L}_2$ simplifies to $\tau^2 ||x||^2$, which penalizes deviations of the unfolded distribution from zero. This analysis includes the regularization of the size during the unfolding process. This analysis does not uses the regularization of the first and second derivatives.

The third Lagrangian represents an optional area constraint. This area constraint adds two additional parameters to the unfolding problem. The first parameter, $Y \equiv \sum_i y_i$, counts the number of observed events. The second parameter, $e_j \equiv \sum_i A_{ij}$, represents an efficiency vector that counts the number of observations in each column of the migration matrix. When $\mathcal{L}_3$ is activated, the normalization of $\vec{x}$ is corrected to match the total number of events, Y using the efficiencies, $\vec{e}$. This constraint is activated in this analysis. Minimizing the complete Lagrangian yields the following solution for $\vec{x}$ (see Ref. [131] for details).

$$\vec{x} = \begin{cases} \vec{x}|_{\lambda=0}, & \text{without area constraint} \\ \vec{x}|_{\lambda=0} + (\lambda/2)E\vec{e}, & \text{with area constraint} \end{cases} \tag{9.5}$$

The following variables are defined in terms of the components of Eq. 9.4.

$$\vec{x}|_{\lambda=0} = E\left[A^{\mathrm{T}}V_{yy}^{-1}\vec{y} + \tau^2(L^{\mathrm{T}}L)f_b\vec{x}_0\right] \tag{9.6a}$$

$$E = \left(A^{\mathrm{T}}V_{yy}^{-1}A + \tau^2(L^{\mathrm{T}}L)\right)^{-1} \tag{9.6b}$$

$$\frac{\lambda}{2} = \frac{Y - \vec{e}^{\mathrm{T}}\vec{x}|_{\lambda=0}}{e^{\mathrm{T}}E\vec{e}} \tag{9.6c}$$

The final step to determine the unfolded distribution is to select the optimal value of τ. It is selected by running the unfolding iteratively with different values of τ and selecting the value that minimizes the discrepancy between the observed and predicted distributions. There are different methods to select τ with different figures of merit to optimize[2]. In the present implementation, TUnfold uses the L-curve scan method to optimize τ [133]. The L-curve scan method defines two variables, L_x^{curve} and L_y^{curve}.

$$L_x^{\text{curve}} = \log \mathcal{L}_1 \tag{9.7a}$$

$$L_y^{\text{curve}} = \log \frac{\mathcal{L}_2}{\tau^2} \tag{9.7b}$$

L_x^{curve} measures the agreement of $\vec{x}$ with the observations, while L_y^{curve} measures the agreement of $\vec{x}$ with the regularization conditions. A parametric plot of L_y^{curve} against L_x^{curve} will yield an L-shaped curve. The optimal value of τ corresponds to the point on the curve that maximizes the curvature. At this point, the regularization is strong enough to stabilize the solution but not so strong as to introduce large biases. The unfolded distribution is consistent with what would be output by a simple least-squares minimization while also being compatible with the bias. Figure 9.1 shows an example of a typical L-curve plot with different values of the regularization parameter (labeled here as λ).

[2] In addition to the L-curve scan presented in the text, another popular method to choose τ is to minimize the average global correlation coefficient. The global correlation coefficient in bin i (ρ_i) is defined as $\rho_i = \sqrt{1 - \left((V_{yy})_{ii}(V_{yy}^{-1})_{ii}\right)^{-1}}$. The average global correlation coefficient over N bins is $\rho_{\text{avg.}} = \frac{1}{N}\sum_{i=1}^{N}\rho_i$. Similar to the L-curve scan, the unfolding is performed iteratively for different values of τ. However, the best value of τ is chosen as the iteration that gives the smallest $\rho_{\text{avg.}}$. This method minimizes the bin-to-bin fluctuations post-unfolding.

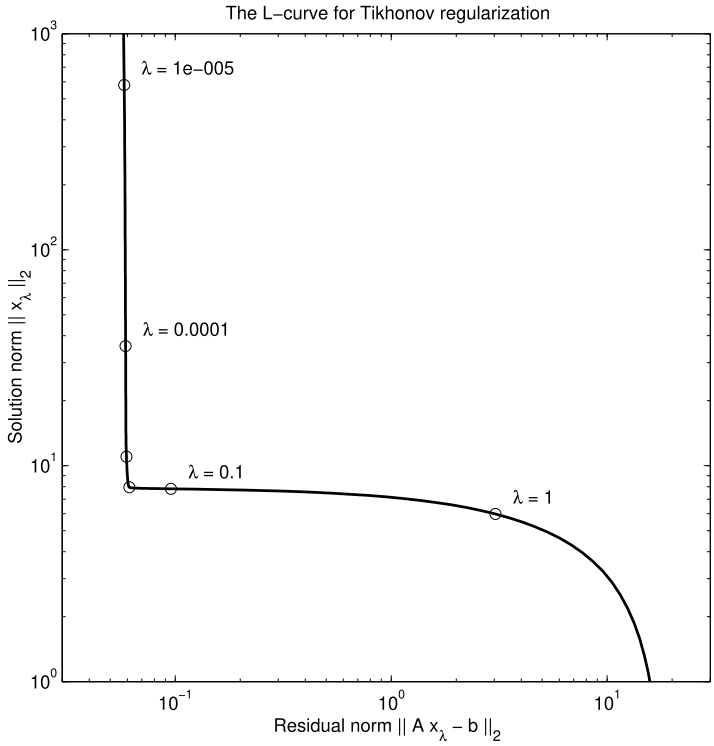


Figure 9.1: Representative L-Curve Plot. A representative L-curve plot shown with various regularization parameter (λ) values is presented. The optimal value of the regularization parameter is the point corresponding to maximum curvature. This point corresponds to a kink in the L-curve. Generally, this number is a fractional number that is less than 1. In the text, the λ-parameter is equivalently labeled as τ. Plot is taken from Ref. [133].

In practice, the user defines a constant number of iterations to scan between a range of possible τ values. At each iteration, the unfolding is performed with a different τ value, and the corresponding L_x^{curve} and L_y^{curve} values are recorded. The curvature of the L-curve is then calculated at each point, and the τ value that maximizes the curvature is selected as the optimal value. The optimal value is then used during the final iteration of the unfolding to obtain the unfolded distribution. This final step completes the unfolding process. The unfolding process is correctly verified with a closure test, with the results shown in Appendix D.

9.2 Unfolding Ingredients

For each observable of interest, the following ingredients are required to perform unfolding.

Full Detector-Level Distribution, *R* The distribution comprises all reconstructed events filled with MC simulated events that pass the detector-level selection.

Full Particle-Level Distribution, *T* The distribution comprises all particle-level events filled with MC simulated events that pass the particle-level selection.

Data Distribution, D The distribution of observed events is filled with real data events that pass the detector-level selection.

Background Distribution, B The distribution sums up all contributions from simulated background events that pass the detector-level selection filled using MC simulations.

Response Matrix, M_r The response matrix is constructed using MC simulation and is filled with all events that simultaneously pass the detector-level and particle-level selections.

Migration Matrix, M The migration matrix is the normalized version of the response matrix such that each row in the migration matrix sums to unity. It encodes the probability that an event in a given particle-level bin is reconstructed in a detector-level bin. The migration matrix's diagonal elements quantifies the purity, i.e., the probability that an event in a given particle-level bin is correctly reconstructed in the corresponding detector-level bin. The off-diagonal elements represent the migration of events between different bins. There are approximately twice the number of bins at detector-level compared to truth-level so as to align with the recommendations of the TUnfold package.

The unfolding process can be summarized in the following equation which calculates the unfolded distribution ($D_i^{\text{unf.}}$) in given bin i.

$$D_i^{\text{unf.}} = \frac{1}{\epsilon_i} \sum_j M_{i,j}^{-1} A_j (D_j - B_j) \tag{9.8}$$

In addition to the previously defined ingredients, two additional ingredients are required for the unfolding process. These are the acceptance (A) and efficiency (ϵ) correction factors. The acceptance correction factor accounts for events that pass the detector-level selection but fail the particle-level selection. The acceptance correction in bin j (A_j) can be calculated using the response matrix (M_r) and the full detector-level distribution (R) as follows.

$$A_j = \frac{\sum_i M_{r,i,j}}{R_j} \tag{9.9}$$

In other words, for the j-th bin, the acceptance correction factor is the sum of the response matrix elements in the j-th column divided by the number of events in the j-th bin of the full detector-level distribution. On the other hand, the efficiency correction factor accounts for events that pass the particle-level selection but are not reconstructed at the detector-level. The efficiency correction in bin i (ϵ_i) can also be calculated using the response matrix (M_r) and the full particle-level distribution (T) as follows.

$$\epsilon_i = \frac{\sum_j M_{r,i,j}}{T_i} \tag{9.10}$$

The efficiency correction factor in the i-th bin is the sum of the response matrix elements in the i-th row divided by the number of events in the i-th bin of the full particle-level distribution. The acceptance and efficiency correction factors correct the normalization of the unfolded distribution due to events

passing/failing the detector-level and particle-level selections. Figures 9.2 and 9.3 plots the migration matrices and acceptance/efficiency correction factors for the $W_{\text{had.}}$ (particle) region. The corresponding migration matrices for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions are found in Appendix C.4, while the acceptance and efficiency factors are found in Appendix C.5.

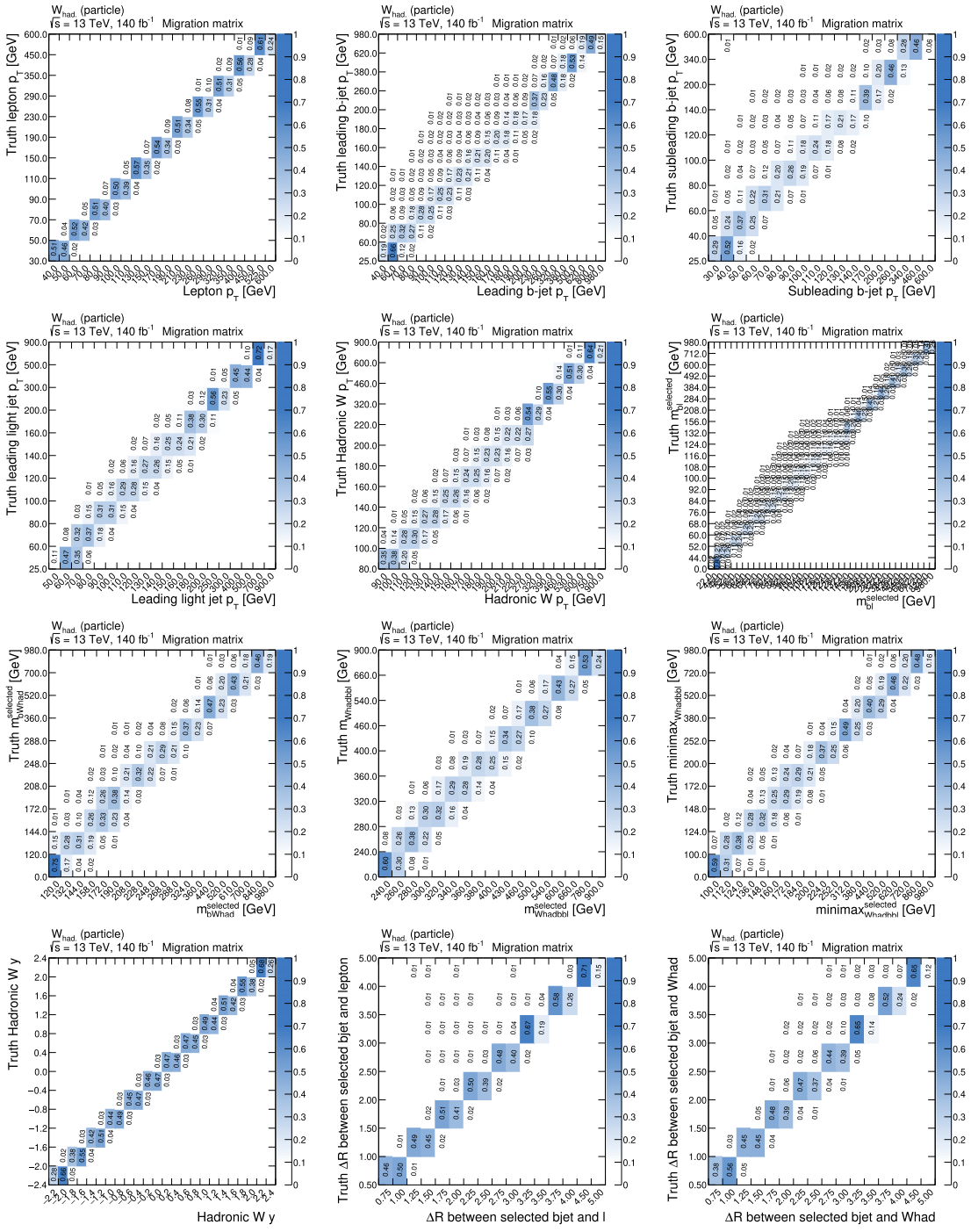


Figure 9.2: Migration Matrices for the $W_{\text{had.}}$ (particle) Region. Each bin in the migration matrix represents the fraction of events in the reconstructed bin that originate from the true bin. The diagonal elements represent the fraction of correctly reconstructed events, while the off-diagonal elements represent the fraction of misreconstructed events. Migration matrices are plotted for the observables of interest. Refer to Section 7.3 for a description of the observables of interest.

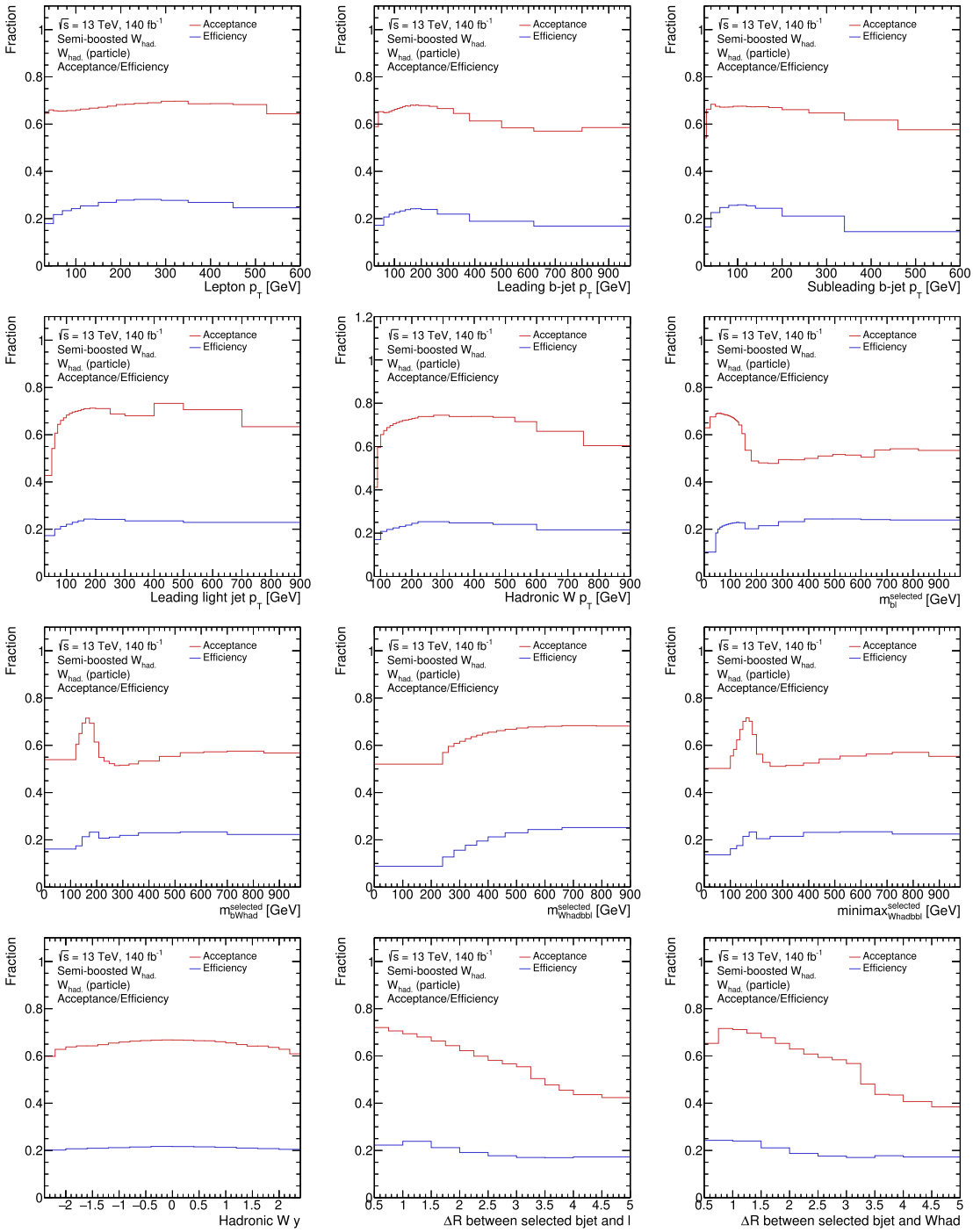


Figure 9.3: Acceptance/Efficiency Factors for the $W_{\text{had.}}$ (particle) Region. The acceptance and efficiency factors correct the normalization of the unfolded distribution. For each bin, the acceptance correction factor accounts for events that pass the detector-level selection but fail the particle-level selection. The efficiency correction factor accounts for events that pass the particle-level selection but are not reconstructed at the detector-level. Refer to Section 7.3 for a description of the observables of interest.

The following are several reasons why the efficiency correction factors are much smaller than unity.

Finite detector acceptance/efficiency Physics objects at the particle-level may fall out of the detector acceptance and, therefore, may not be reconstructed. For example, in the case of leptons and jets, this effect may cause the entire event to be rejected if they do not pass the required p_{T} and η cuts.

Finite detector resolution The kinematics of various physics objects may be smeared due to the finite resolution of the detector. This smearing effect causes the reconstructed objects to have different kinematics than the true particle-level objects. This effect may cause specific objects to fail certain selection cuts at detector-level. For example, in the case of low-momentum jets and low missing transverse energy.

Misreconstruction The result of certain physics objects failing to be reconstructed may have a trickle-down effect on later reconstruction algorithms. For example, jets falling out of the detector acceptance may cause the hadronic-W to be misreconstructed and cause large migrations between bins. This results in an off-diagonal migration matrix with a lower efficiency.

Reconstruction inefficiencies The various reconstruction algorithms carry inefficiencies that cause physics objects not to be reconstructed. For example, the b-tagging algorithm has several working points with efficiencies lower than 100%. Therefore, particle-level b-jets may not be tagged at the detector-level. As a result, this leads to events being rejected at the detector-level if they fail the minimum two b-jet requirement.

Failing quality criteria The physics objects may fail certain quality requirements. For example, the identification and isolation requirements require that leptons to be well-separated from other physics objects. At detector-level, if the lepton is reconstructed close to a jet, the lepton may fail the isolation requirement and be rejected.

9.3 Propagation of Uncertainties

Uncertainties in the unfolded distributions are propagated from the uncertainties in the input distributions. The uncertainties in the unfolded distributions are divided into two categories: statistical and systematic. The statistical uncertainties arise from the finite number of events in the data and MC samples and are evaluated using pseudo-data. The systematic uncertainties are evaluated by systematically shifting the ingredients of the unfolding and recomputing the unfolded distribution. The total uncertainty is then the sum in quadrature of the statistical and systematic uncertainties. Both of these methods are described in more detail below.

9.3.1 Propagation of Statistical Uncertainties

To evaluate the effects of statistical fluctuations propagated through the unfolding process, the unfolding is performed multiple times using pseudo-data. In particle physics, this procedure is often termed

"throwing toys" at a distribution and is a common method to evaluate the statistical uncertainties[3]. Given a distribution, a toy distribution is generated by shifting the number of events in each bin according to a Poisson distribution. Given the number of events in bin i (N_i), the number of events in the toy distribution (N_i^{toy}) is generated as follows.

$$N_i^{\text{toy}} = \text{Poisson}(\lambda = N_i) \tag{9.11}$$

In this analysis, the number of toys is set to 1000. The unfolding is then performed for each toy distribution, and the unfolded distribution is recorded. For each bin, there are 1000 newly generated bin heights. Although these 1000 values were generated from a Poisson distribution, the large λ values mean that the distribution of the unfolded bin heights will be approximately Gaussian. To evaluate the statistical uncertainty in each bin, we evaluate the quantiles of the unfolded bin heights. The $\pm 1\sigma$ statistical uncertainty corresponds to the 15.87^{th} and 84.13^{th} percentiles of the unfolded bin heights.

$$\Pr(Z \leqslant 1) = 0.8413 \tag{9.12a}$$

$$\Pr(Z \leqslant -1) = 0.1587 \tag{9.12b}$$

It will be shown in the unfolded cross-sections that the statistical uncertainties in the MC are on the order of 5% or less. The statistical uncertainties in the data are larger and can be as large as 10% in the tails of the distributions.

9.3.2 Propagation of Systematic Uncertainties

The systematic uncertainties are evaluated by systematically shifting the ingredients of the unfolding and recomputing the unfolded distribution. The systematic uncertainties are evaluated for each source of systematic uncertainty. The discrepancy between the unfolded distributions with and without the systematic shifts is then taken as the systematic uncertainty. Firstly, the relative shift (in bin i) corresponding to a single source of systematic uncertainty ($\Delta^{\text{rel.}}_{\text{sys.},i}$) is calculated as follows.

$$\Delta^{\text{rel.}}_{\text{sys.},i} = \theta(x_{\text{sys.}}) \cdot |x_{\text{sys.}}| \cdot \sigma^{\text{up}}_{\text{sys.,rel.},i} + \theta(-x_{\text{sys.}}) \cdot |x_{\text{sys.}}| \cdot \sigma^{\text{down}}_{\text{sys.,rel.},i} \tag{9.13}$$

In the above, $\theta(x)$ is the Heaviside step function, equal to 1 for $x \geqslant 0$ and 0 for $x < 0$. $x_{\text{sys.}}$ is a random number drawn from a standard Gaussian distribution. $\sigma^{\text{up}}_{\text{sys.,rel.},i}$ and $\sigma^{\text{down}}_{\text{sys.,rel.},i}$ are the relative differences between the (up/down) shifted distributions relative to the nominal distribution.

[3] As an aside, the use of toy distributions goes beyond estimating statistical fluctuations. Another common use of toy distributions arises when determining p-values to test the significance of a potential discovery. For example, in searches, one will generate many toy datasets (with similar signal and background yields) from a Poisson distribution according to a known PDF set. The parameter of interest can be determined for each toy dataset; these "toy parameters" can be used to populate a probability distribution. Once unblinded and the observed value of the parameter is measured, the p-value of the observation can be calculated using this distribution. The result is deemed statistically significant if the p-value exceeds the standard 5σ requirement.

$$\sigma^{\text{up/down}}_{\text{sys.,rel.,}i} = \frac{h_i^{\text{up/down}} - h_i^{\text{nom.}}}{h_i^{\text{nom.}}} \tag{9.14}$$

As before, the $h_i^{\text{up/down}}$ represent the height of bin i of the up/down varied distributions, while $h_i^{\text{nom.}}$ represents the height of bin i of the nominal distribution. This procedure generates a large set of relative shifts for each source of systematic uncertainty that can be used to recompute the full-detector-level distributions (R'), the migration matrices (M') and the background distributions (B'). The full particle-level distributions ($T' = T$) and the data distribution ($D' = D$) remain unchanged. The unfolded distribution is then recomputed using the shifted ingredients.

$$D_i'^{\text{unf.}} = \frac{1}{\epsilon_i'} \sum_j M_{i,j}'^{-1} A_j' (D_j - B_j') \tag{9.15}$$

The corresponding shifted quantities in the unfolding equation are defined as follows.

$$R_j' = R_j \left[1 + \sum_{\text{sys.}} \Delta^{\text{rel.}}_{\text{sys.},j} \right] \tag{9.16a}$$

$$M_{i,j}' = M_{i,j} \left[1 + \sum_{\text{sys.}} \Delta^{\text{rel.}}_{\text{sys.},i,j} \right] \tag{9.16b}$$

$$B_j' = B_j \left[1 + \sum_{\text{sys.}} \Delta^{\text{rel.}}_{\text{sys.},j} \right] \tag{9.16c}$$

In the above, each sum runs over all sources of systematic uncertainties. The acceptance and efficiency corrections are then recomputed using the shifted quantities and entered into the modified unfolding equation, Eq. 9.15.

$$A_j' = \frac{\sum_i M_{r,i,j}'}{R_j'} \tag{9.17a}$$

$$\epsilon_i' = \frac{\sum_j M_{r,i,j}'}{T_i} \tag{9.17b}$$

We obtain a shifted unfolded distribution corresponding to the various NPs for each observable. Each shifted unfolded distributions can then be compared to the nominal unfolded distribution to determine the systematic shift.

9.4 Covariance and Correlation Matrices

The total uncertainty from statistical and systematic sources can be summed together to form a covariance matrix. The elements of the covariance matrix are computed as follows.

$$\text{Cov}_{i,j} = \sigma_i \times \sigma_j \tag{9.18}$$

where σ_i and σ_j represent the larger of the two absolute shifts in bin i and bin j respectively. The absolute shift is the difference between the up/down shifted distributions and the nominal distribution.

$$\delta_i^{\text{up/down}} = h_i^{\text{up/down}} - h_i^{\text{nom.}} \tag{9.19}$$

The magnitude of σ_i is taken to be the larger of the two absolute shifts.

$$\sigma_i = \begin{cases} |\delta_i^{\text{up}}|, & \text{if } |\delta_i^{\text{up}}| \geqslant |\delta_i^{\text{down}}| \\ |\delta_i^{\text{down}}|, & \text{otherwise} \end{cases} \tag{9.20}$$

The sign of σ_i is determined by the sign of the product of δ_i^{up} and δ_i^{down} and their relative magnitudes. The following conditions determine the sign of σ_i.

$$\text{sign}(\sigma_i) = \begin{cases} \text{sign}(\delta_i^{\text{up}}), & \text{if } \delta_i^{\text{up}} \times \delta_i^{\text{down}} < 0 \\ \text{sign}(\delta_i^{\text{up}}), & \text{if } \delta_i^{\text{up}} \times \delta_i^{\text{down}} \geqslant 0 \text{ and } |\delta_i^{\text{up}}| \geqslant |\delta_i^{\text{down}}| \\ -\text{sign}(\delta_i^{\text{down}}), & \text{if } \delta_i^{\text{up}} \times \delta_i^{\text{down}} \geqslant 0 \text{ and } |\delta_i^{\text{up}}| \leqslant |\delta_i^{\text{down}}| \end{cases} \tag{9.21}$$

The elements falling on the main diagonal of the covariance matrix then represent the (squared) total uncertainty in bin i of the unfolded distribution. Once the covariance matrices have been calculated, the correlation matrix can be calculated from the elements of the covariance matrix as follows.

$$\text{Corr}_{i,j} = \frac{\text{Cov}_{i,j}}{\sqrt{\text{Cov}_{i,i} \times \text{Cov}_{j,j}}} \tag{9.22}$$

Unlike the covariance matrix, which measures the total uncertainty in each bin, the correlation matrix measures the correlation between bins. It can be used to assess the degree to which the entries in one bin correlate with those in another. Figures 9.4 and 9.5 plot the covariance and correlation matrices for the observables of interest in the $W_{\text{had.}}$ (particle) region. The corresponding covariance and correlation matrices for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions can be found in Appendix C.6 and Appendix C.7.

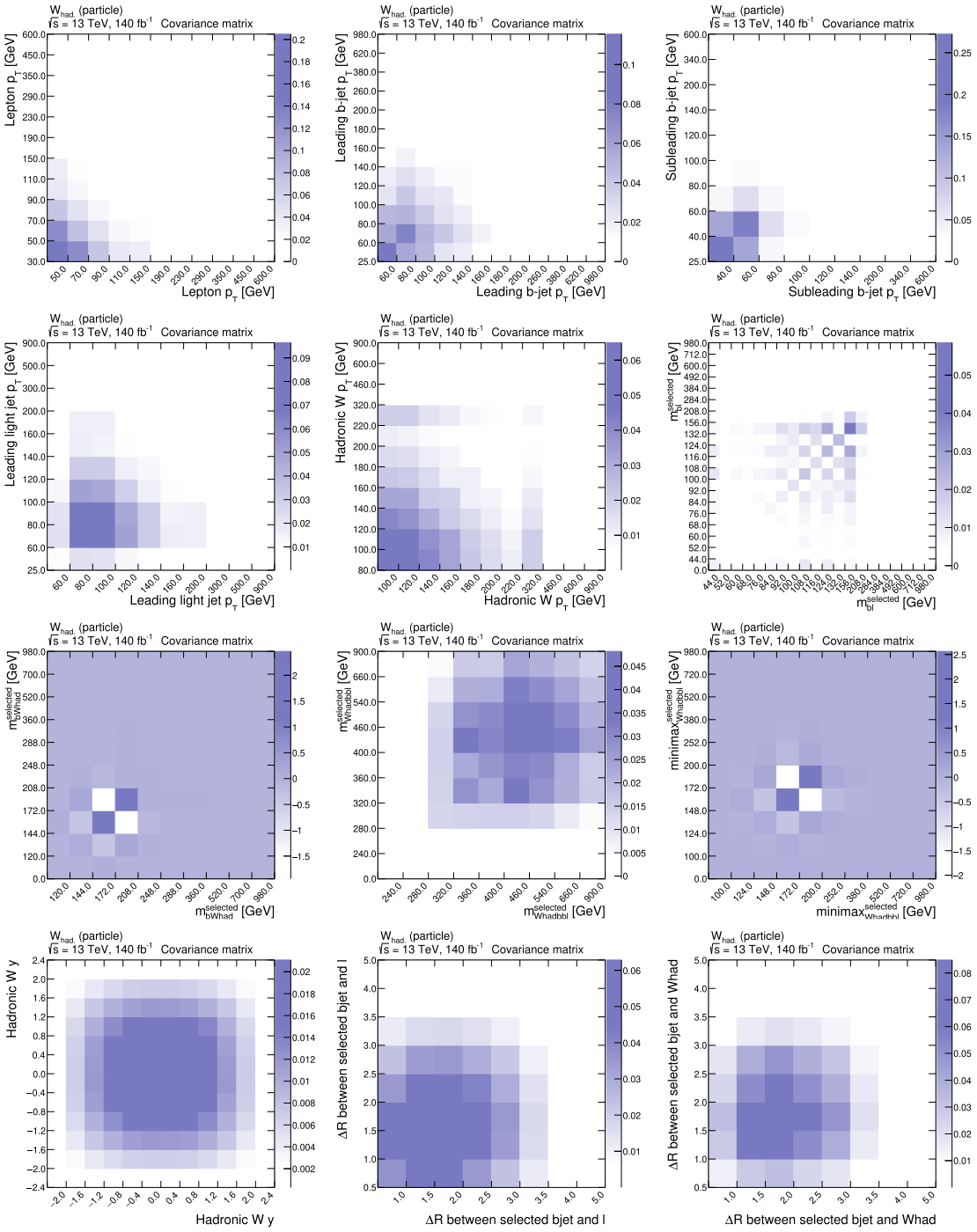


Figure 9.4: Covariance Matrices for the $W_{\text{had.}}$ (particle) Region. The covariance matrix measures the covariance between bins of the unfolded distribution. The elements on the covariance matrix's main diagonal indicate the bins' total squared uncertainties in the unfolded distribution. Refer to Section 7.3 for a description of the observables of interest.

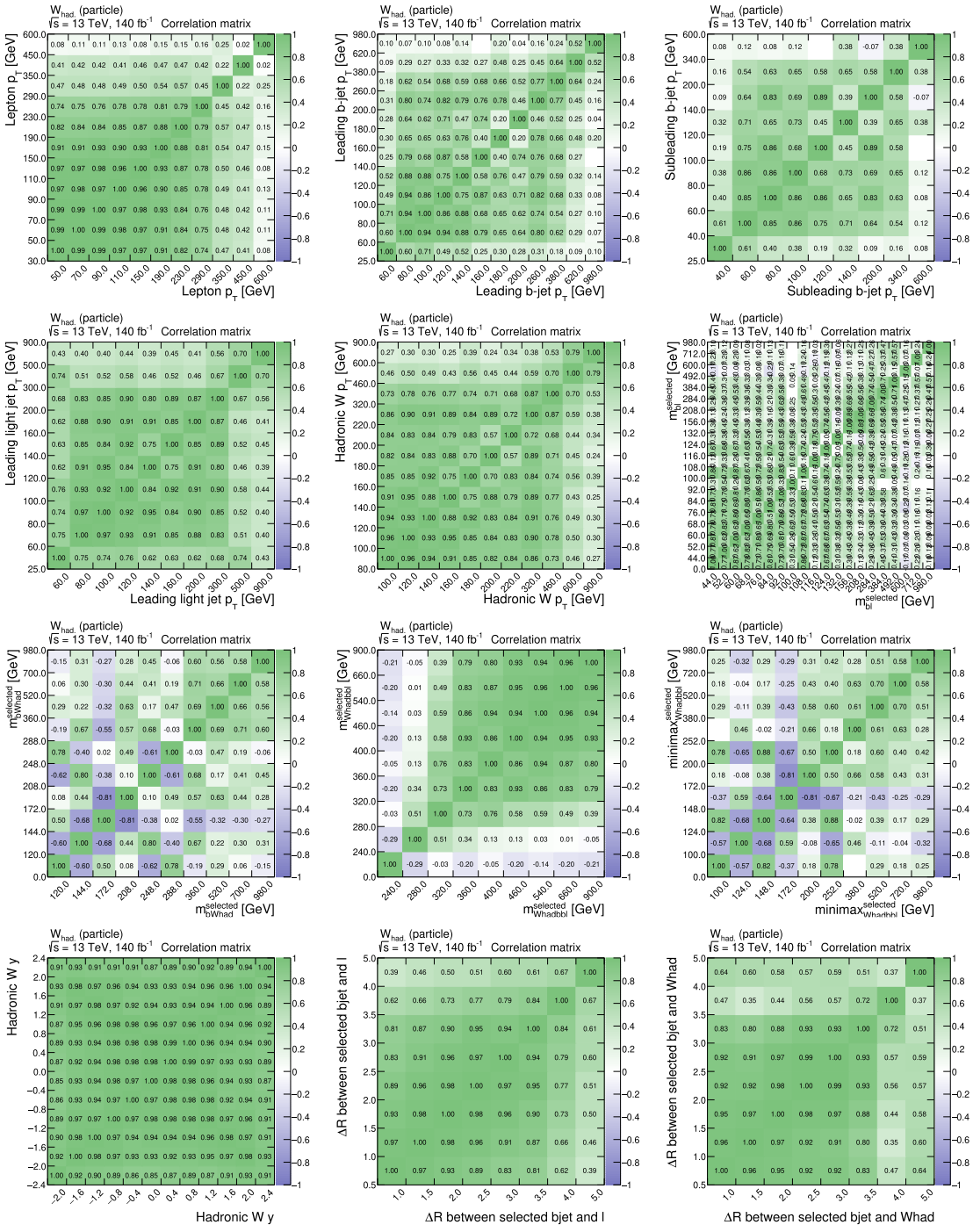


Figure 9.5: Correlation Matrices for the $W_{\text{had.}}$ (particle) Region. The correlation measures the total correlation between bins of the unfolded distribution. A large positive number equates to positive correlations, while a large negative number equates to negative correlations. Large correlations (positive or negative) in distant bins suggest some common factor or influence affecting both bins. Refer to Section 7.3 for a description of the observables of interest.

There are generally two forms the correlation matrices take on for the observables of interest: matrices with uniformly large positive correlations and matrices with alternating positive and negative correlations. Note that these patterns for correlation matrices are not uncommon and are often seen in unfolding analyses. For example, the corresponding $pp \to WbWb$ dilepton analysis also exhibits these patterns in the correlation matrices for different variables [134]. The large positive correlations indicate a global shift in the unfolded distribution, likely due to some systematic uncertainty that has been propagated through the unfolding. In the second case, the alternating positive and negative correlations indicate unfolding with little to no regularization. Recall that while unfolding can be correctly performed using a simple matrix inversion, the result is often unstable and can lead to large fluctuations in the uncertainties of the unfolded distribution. Regularization is a technique used to stabilize the unfolding and reduce the fluctuations. The appearance of these alternating correlations suggests that the unfolding is still under-regularized. Further tuning of the regularization parameter may be necessary to stabilize the unfolding and reduce the fluctuations in the correlations.

9.5 Unfolded Cross-Sections

The unfolded cross-sections are presented in this section. After conducting the unfolding procedure, the absolute differential cross-section in bin i can be calculated after normalizing to the integrated luminosity and the bin width. The result is the absolute differential cross-section in the fiducial phase space defined by the particle-level selections. The absolute differential cross-section in bin i for an observable X ($\frac{d\sigma_{\text{fid.}}}{dX_i}$) is given by the following equation.

$$\frac{d\sigma_{\text{fid.}}}{dX_i} = \frac{1}{L \times \Delta X_i} \times N_i^{\text{unf.}} \tag{9.23}$$

In the above, $N_i^{\text{unf.}}$ is the unfolded number of events in bin i, L is the integrated luminosity (corresponding to the full Run 2 luminosity, ~ 140 fb^{-1}), and ΔX_i is the bin width of the observable X in bin i. A normalized differential cross-section can also be calculated by dividing the absolute differential cross-section by the integral of the differential cross-section distribution. These distributions are shown for the observables of interest in the following plots. For each observable, the unfolded data is compared to the nominal and alternative MC samples, previously listed in Table 5.2. The distinction is made clear in the titles of the captions and the legends of each plot.

9.5.1 Absolute Differential Cross-Sections

The absolute differential cross-sections for the $W_{\text{had.}}$ (particle) region are calculated using Eq. 9.23. In each plot, the unfolded data is represented by black points and the error bars represent the statistical uncertainty. The total uncertainty propagated through the unfolding is shown as a gray band. The various MC predictions are shown as colored points, with corresponding error bars representing the respective modeling uncertainties. The ratio of the unfolded data to the MC predictions is shown in the lower panel of each plot. Figure 9.6 plots the absolute differential cross-sections using the nominal MC generators, while Figure 9.7 plots the absolute differential cross-sections using the alternative

MC generators. The corresponding absolute differential cross-sections for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions can be found in Appendix C.8.

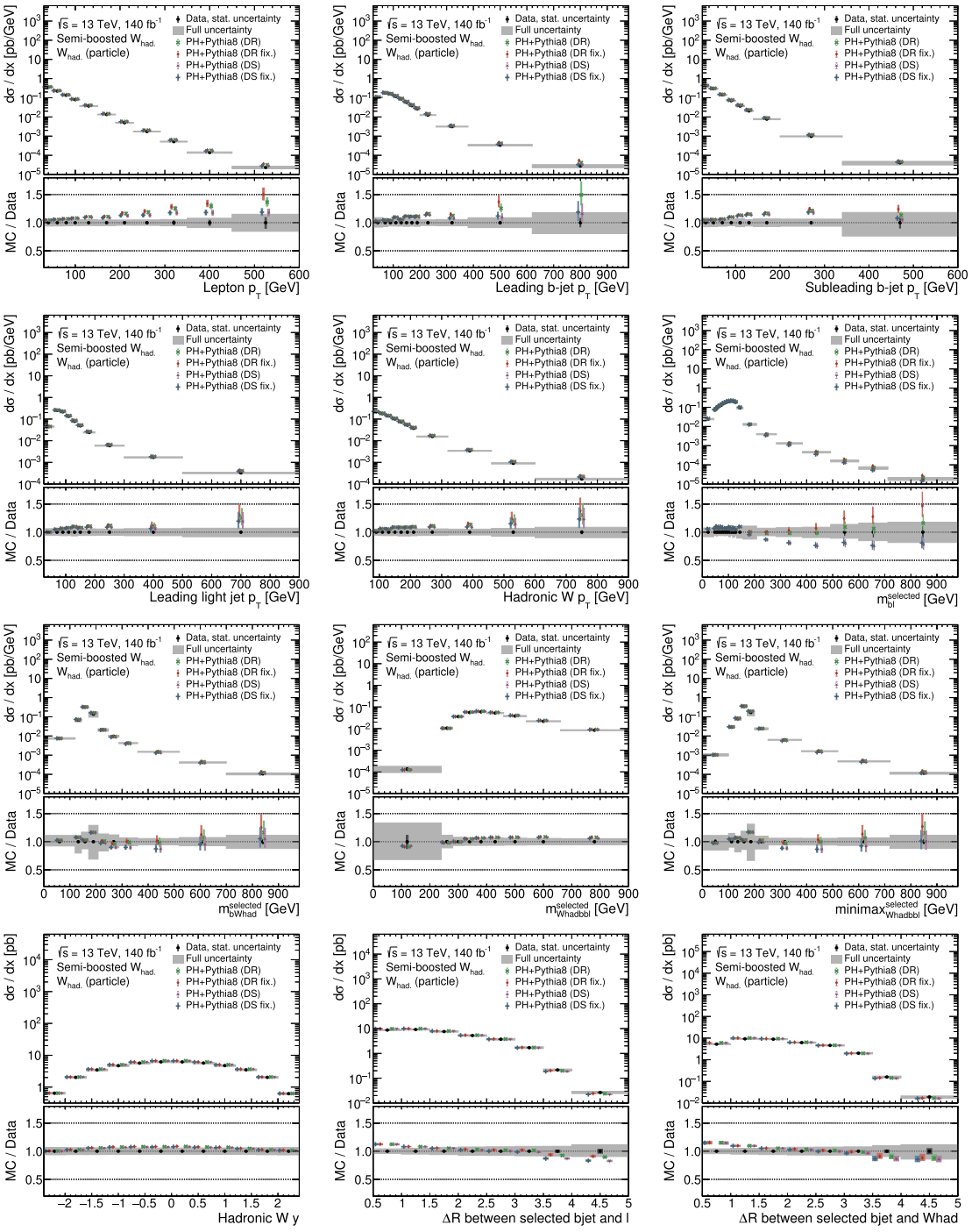


Figure 9.6: Absolute Differential Cross-Sections (Nominal MC) for the $W_{\text{had.}}$ (particle) Region. The absolute differential cross-section is calculated using each bin's unfolded number of events. The measured cross-sections are then compared individually to the set of nominal MC generators. Refer to Section 7.3 for a description of the observables of interest.

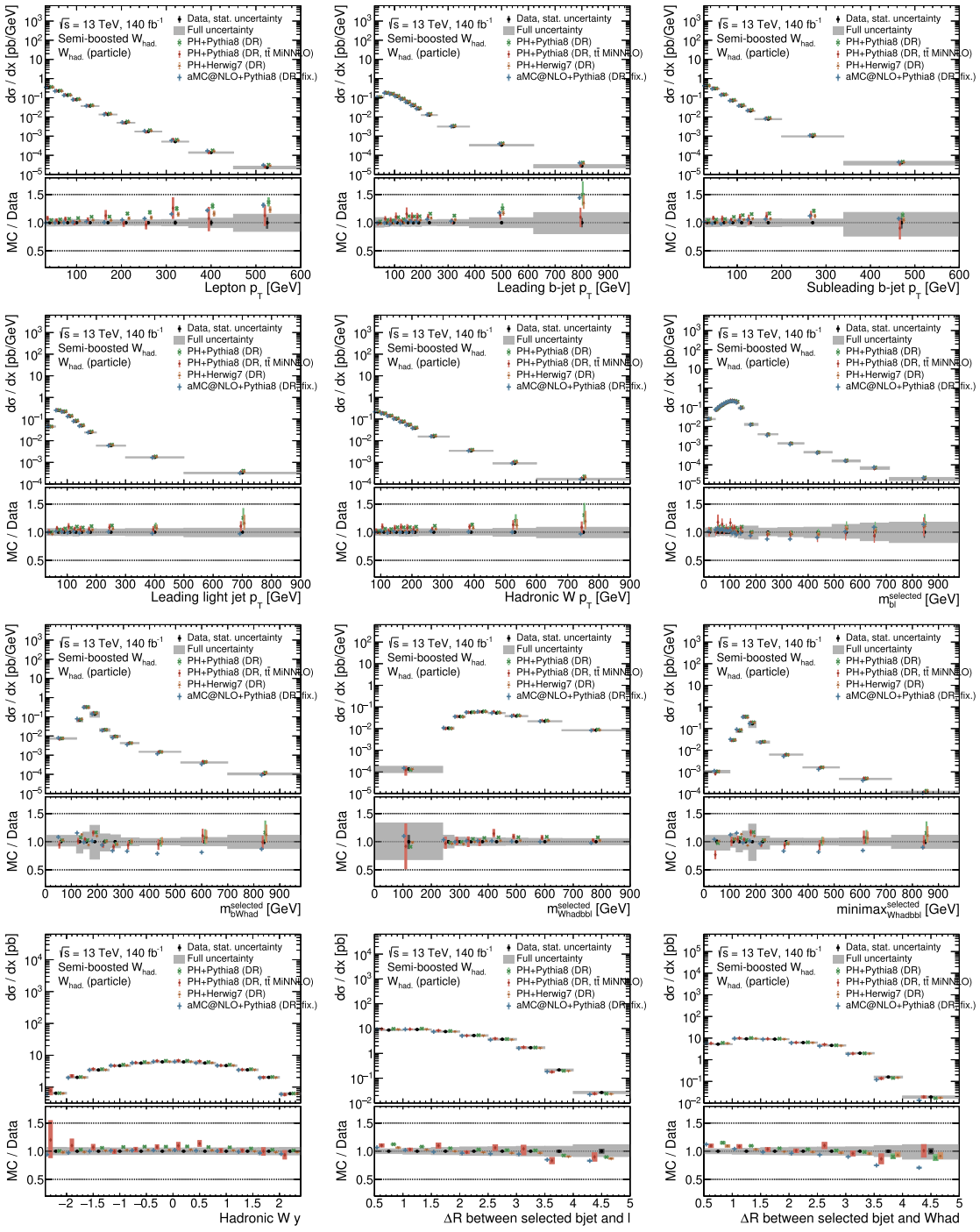


Figure 9.7: Absolute Differential Cross-Sections (Alternative MC) for the $W_{\text{had.}}$ (particle) Region. The absolute differential cross-section is calculated using each bin's unfolded number of events. The measured cross-sections are then compared individually to the set of alternative MC generators. Refer to Section 7.3 for a description of the observables of interest.

9.5.2 *Normalized Differential Cross-Sections*

In addition to the absolute differential cross-sections, one can also normalize each distribution to the integral of the differential cross-section to compute the normalized differential cross-section ($1/\sigma d\sigma/dx$). The normalized differential cross-section is a useful quantity as it allows for a direct comparison of the shape of the distributions. It scales out the overall normalization of the cross-section and is, therefore, equivalent to a probability distribution of the observable. We can easily compare the same observables across different experiments and processes. The plots are labeled in the same manner as was described for the absolute differential cross-sections. Figure 9.8 plots the normalized differential cross-sections using the nominal MC generators, while Figure 9.9 plots the normalized differential cross-sections using the alternative MC generators. The corresponding normalized differential cross-sections for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions can be found in Appendix C.9.

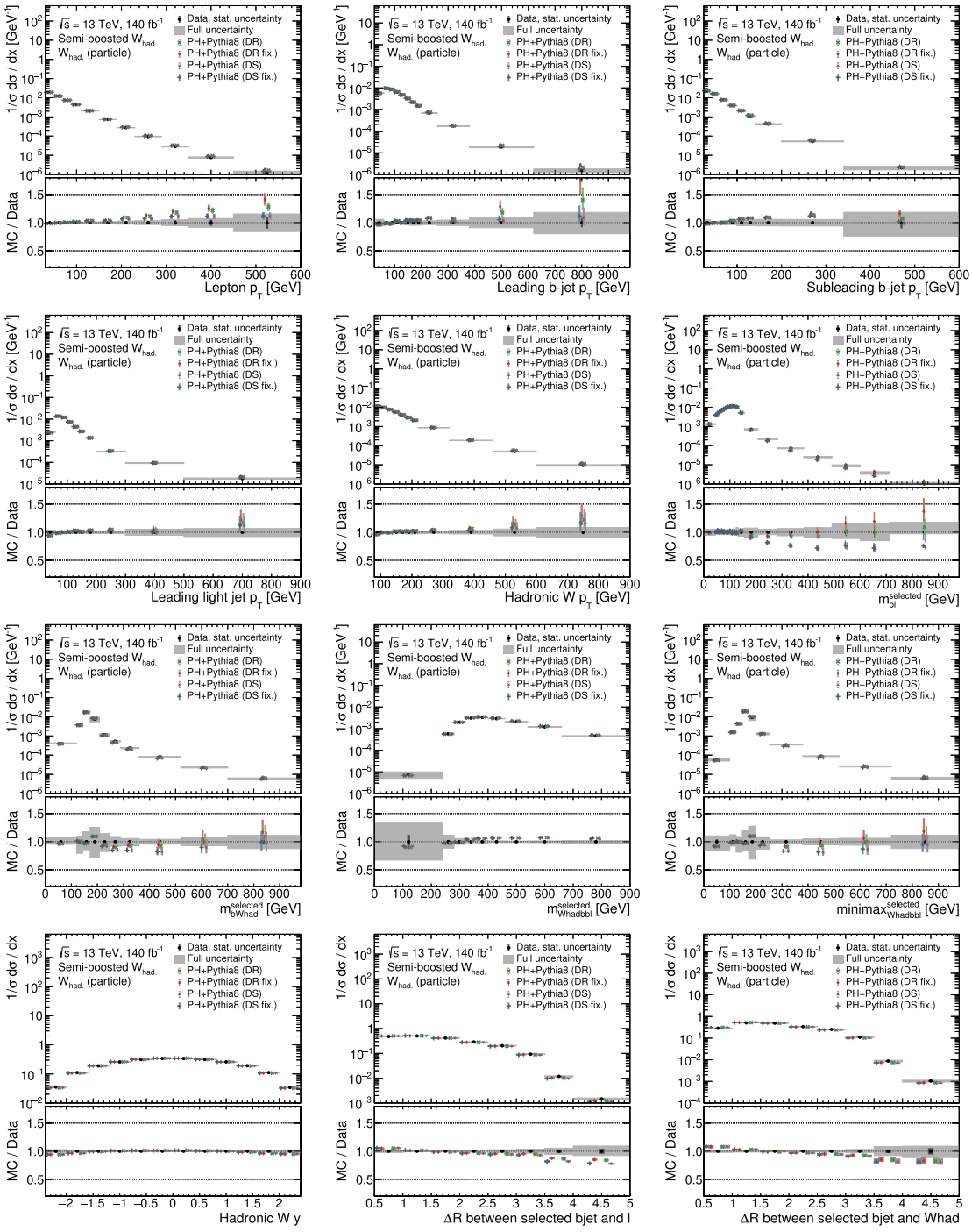


Figure 9.8: Normalized Differential Cross-Sections (Nominal MC) for the $W_{\text{had.}}$ (particle) Region. The normalized differential cross-section is calculated by normalizing the absolute differential cross-section to unit area. The measured cross-sections are then compared individually to the set of nominal MC generators. Refer to Section 7.3 for a description of the observables of interest.

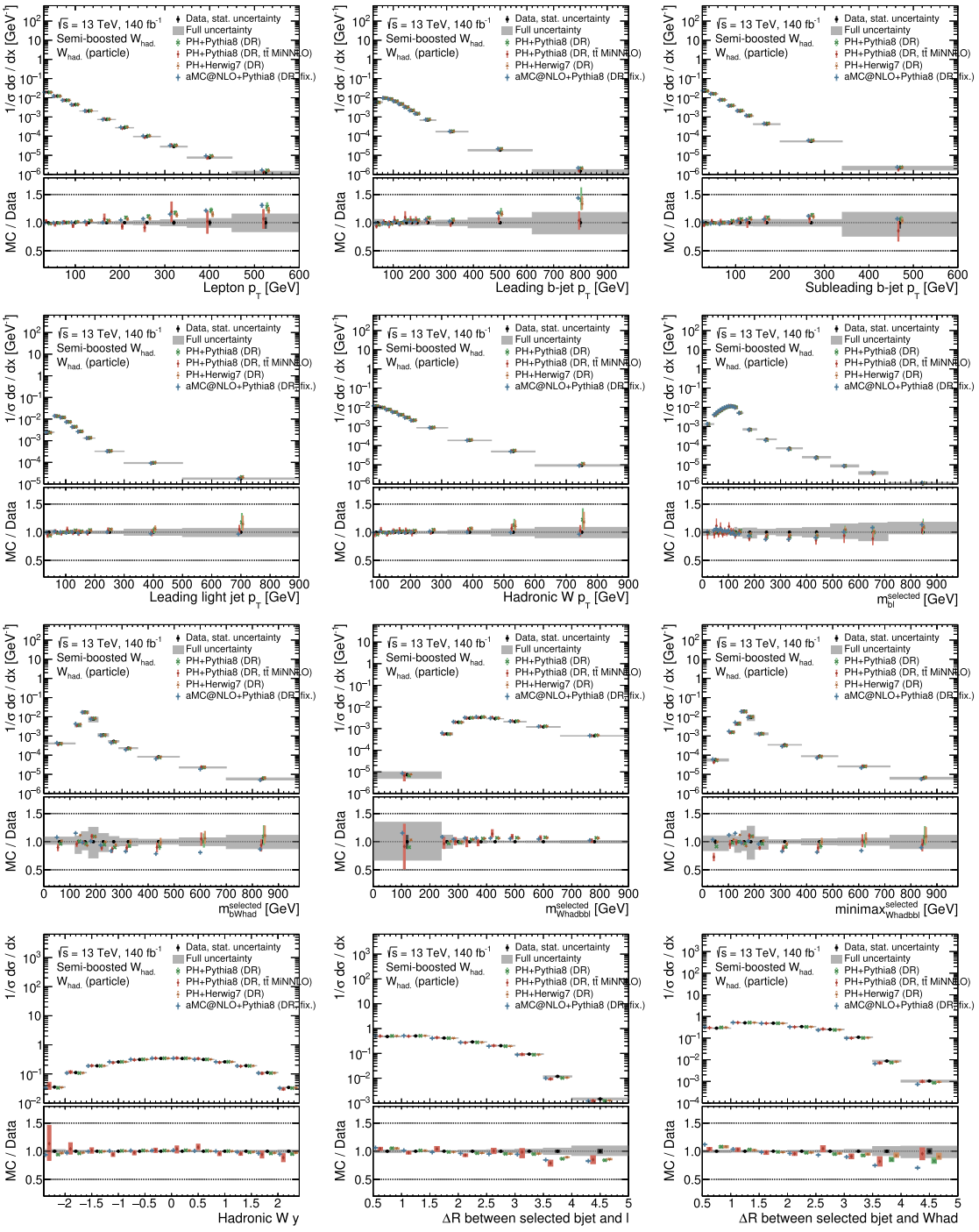


Figure 9.9: Normalized Differential Cross-Sections (Alternative MC) for the $W_{\text{had.}}$ (particle) Region. The normalized differential cross-section is calculated by normalizing the absolute differential cross-section to unit area. The measured cross-sections are then compared individually to the set of alternative MC generators. Refer to Section 7.3 for a description of the observables of interest.

9.6 Discussion of Results

The agreement of the measured cross-sections with the MC predictions is quantified using the χ^2 test. The χ^2 value is calculated using the covariance matrix and the differences between the measured cross-sections and the predictions.

$$\chi^2 = (x-\mu)^T \mathrm{Cov}^{-1} (x-\mu) \tag{9.24}$$

In the above, x corresponds to the measured cross-sections, while μ corresponds to the predicted cross-sections. The corresponding χ^2 value can then be normalized by the Number of Degrees of Freedom (NDF) to obtain the χ^2/NDF value. For a binned quantity, the number of degrees of freedom corresponds to the number of bins. The χ^2/NDF value can be considered a measure of the goodness-of-fit of the measured cross-sections to the MC predictions. One can then calculate the p-value of the χ^2/NDF value. The p-value is the probability of observing a χ^2/NDF value as extreme as the one observed by chance alone, given that the null hypothesis (predictions are correct) is true. The p-value is calculated as the complement of the cumulative distribution function of the χ^2 distribution. The χ^2/NDF and p-values for the $W_{\mathrm{had.}}$ (particle) region are shown in Table 9.1. The corresponding χ^2/NDF and p-values for the $W_{\mathrm{had.}}$ (loose) and $W_{\mathrm{lep.}}$ (loose) regions can be found in Appendix C.10.

Table 9.1: χ^2 and p-values for the $W_{\text{had.}}$ (particle) Region. The χ^2 and p-values are calculated by comparing the measured cross-sections to the MC predictions. The χ^2 values are then normalized to the number of degrees of freedom. Using the χ^2 value and NDF, the p-value is calculated. The p-value is the probability of observing a χ^2/NDF value as extreme as the one observed by chance alone. The χ^2/NDF and p-values are computed for the observables of interest in the $W_{\text{had.}}$ (particle) region.

	Generator													
	PH+PYTHIA 8 (DR)		PH+PYTHIA 8 (DR, fix.)		PH+PYTHIA 8 (DS)		PH+PYTHIA 8 (DS, fix.)		PH+PYTHIA 8 (DR, $t\bar{t}$ MINNLO)		PH+HERWIG 7.1 (DR)		aMCNLO+PYTHIA 8 (DR, fix.)	
Observable	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
Lepton p_{T}	45.44 / 11	<0.001	55.26 /11	<0.001	25.36 / 11	0.008	26.12 /11	0.006	78.94 / 11	<0.001	21.89 / 11	0.025	20.77 / 11	0.036
Leading b-jet p_{T}	30.59 / 12	0.002	41.35 / 12	<0.001	24.28 / 12	0.019	24.57 / 12	0.017	29.67 / 12	0.003	14.39 / 12	0.277	14.73 / 12	0.257
Subleading b-jet p_{T}	32.76 / 9	<0.001	36.88 / 9	<0.001	27.33 / 9	0.001	28.61 / 9	<0.001	8.78 / 9	0.457	11.30 / 9	0.256	16.84 / 9	0.051
Leading light-jet p_{T}	15.22 / 10	0.124	20.25 / 10	0.027	9.21 / 10	0.512	9.59 / 10	0.477	23.54 / 10	0.009	8.99 / 10	0.533	2.74 / 10	0.987
Hadronic-W p_{T}	17.87 / 11	0.085	22.70 / 11	0.019	12.03 / 11	0.362	12.35 / 11	0.338	30.53 / 11	0.001	10.66 / 11	0.473	3.40 / 11	0.984
Selected $(b + \ell)$ mass	21.04 / 20	0.395	33.20 / 20	0.032	41.95 / 20	0.003	41.13 / 20	0.004	99.21 / 20	<0.001	11.46 / 20	0.934	33.77 /20	0.028
Selected $(b + W_{\text{had.}})$ mass	24.22 / 10	0.007	26.74 / 10	0.003	50.13 / 10	<0.001	49.81 / 10	<0.001	78.01 / 10	<0.001	12.83 / 10	0.233	59.37 / 10	<0.001
Selected $(W_{\text{had.}} + b_1 + b_2 + \ell)$ mass	4.08 / 9	0.906	4.01 / 9	0.911	5.08 / 9	0.827	4.79 / 9	0.852	153.5 / 9	<0.001	1.60 / 9	0.996	6.98 / 9	0.639
Modified minimax mass	18.96 / 10	0.041	25.95 / 10	0.004	45.30 / 10	<0.001	44.57 / 10	<0.001	45.02 / 10	<0.001	20.17 / 10	0.028	65.90 / 10	<0.001
Hadronic-W y	14.95 / 12	0.244	15.26 / 12	0.228	14.24 / 12	0.285	14.17 / 12	0.290	418.3 / 12	<0.001	7.85 / 12	0.796	9.83 / 12	0.630
$\Delta R(b, \ell)$	41.25 / 8	<0.001	42.04 / 8	<0.001	42.88 / 8	<0.001	43.86 / 8	<0.001	268.5 / 8	<0.001	22.55 / 8	0.004	33.41 / 8	<0.001
$\Delta R(b, W_{\text{had.}})$	103.0 / 8	<0.001	105.1 / 8	<0.001	107.3 / 8	<0.001	110.2 / 8	<0.001	159.7 / 8	<0.001	67.66 / 8	<0.001	148.8 / 8	<0.001

When one examines the various differential cross-section distributions in Figures 9.6–9.9, there are several distributions with predictions that describe the data poorly. For example, the lepton (top row, top left) and leading b-jet (top row, middle) transverse momentum predictions generally overestimate the hardness of these objects. The discrepancy is most pronounced in the high-p_{T} region. There is also an upward-trending slope to the predictions across all generators. In these distributions, the DS predictions generally describe the data better than the DR predictions. In particular, the DS predictions models the leading b-jet p_{T} well.

Inspecting the subleading b-jet p_{T} distribution (top row, right), the same pattern of an upward-trending slope is observed. This pattern continues until the last bin when the predictions match data well. The leading light-jet p_{T} (second row, left) and hadronic-W p_{T} (second row, middle) distributions show a similar trend, with the predictions overestimating the data in the high-p_{T} region with the DS predictions generally outperforming their DR counterparts. However, predictions provided by the MADGRAPH5_AMC@NLO+PYTHIA generator are a particularly good match to the data throughout the hadronic-W p_{T} spectrum.

The selected $(b + \ell)$ invariant mass (second row, right) shows good agreement with data near the top-mass resonance. However, there is a sharp transition immediately after the top-mass resonance where the DS generators begin to underestimate the data. Curiously, the DR predictions do not exhibit this behavior and better model the spectrum of this mass variable. As expected, this quantity shows high sensitivity to single-top interference effects. This quantity is a modified version of a $m_{b\ell}^{\text{min.}} \equiv \min\left[M(b_i + \ell)\right]$ variable which is also sensitive to interference effects. At high values of $m_{b\ell}^{\text{min.}}$, doubly-resonant contributions are suppressed.

The selected $(b + W_{\text{had.}})$ invariant mass (third row, left) also sees good agreement near the top-mass resonance and underestimation in the region immediately after. However, at larger values of $m_{\text{bWhad}}^{\text{selected}}$ the predictions fall within the uncertainties. This same pattern is observed in the modified minimax mass distribution (third row, right).

The selected $(W_{\text{had.}} + b_1 + b_2 + \ell)$ invariant mass (third row, middle) shows good agreement with the data for all generators. There is a noticeable transition near the $\mathrm{M}(W_{\text{had.}} + b_1 + b_2 + \ell) \approx 300$ GeV region after which the predictions begin to overestimate the data slightly. However, this quantity is the best modeled observable as shown by the χ^2 and p-values in Table 9.1. This quantity also has relatively low uncertainties, even in the high mass tails.

Regarding the spatial observables of interest, the hadronic-W rapidity (bottom row, left) presents a slight overestimation in the regions of $|\eta| < 1.5$. The $\Delta R(b, \ell)$ (bottom row, middle) and $\Delta R(b, W_{\text{had.}})$ (bottom row, right) distributions display a transition near the $\Delta R \approx 1.5$ region where the predictions transition from overestimating the data to underestimating the data. The ΔR distributions are sensitive to the "hardness" of top quark production. A top quark carrying larger momentum results in the decay products becoming more collimated resulting in a smaller ΔR. On the other hand, a top quark carrying smaller momentum results in decay products that have a larger separation and, thus, a larger ΔR. Therefore, from these distributions, one can infer that the predictions overestimate the hardness of the top quark production for boosted top quarks and underestimate the hardness for softer top quarks.

9.6.1 *Discussion of MiNNLO Predictions*

The MiNNLO predictions are of particular interest as they represent recent efforts to match NNLO QCD calculations with parton showers [89]. This analysis uses the MiNNLO algorithm to simulate $t\bar{t}$ events with the POWHEG+PYTHIA generator. For several observables, one can visually observe that the MiNNLO predictions describe the data well. For example, the lepton p_{T}, leading b-jet p_{T}, and selected $(b + W_{\mathrm{had.}})$ mass distributions are well-modeled by the MiNNLO predictions. However, this is often accompanied by a larger modeling uncertainty in the MiNNLO predictions. Larger uncertainties are particularly evident in the high-p_{T} tails of the lepton p_{T} distribution.

This larger uncertainty is partly due to the MiNNLO algorithm adopting a more conservative uncertainty estimate compared to other fixed-order approaches (refer to Ref. [89] for more details). Firstly, the MiNNLO algorithm induces larger scale uncertainties due to probing much lower scales in both the PDFs and α_s than in fixed-order calculations. Secondly, the MiNNLO algorithm includes additional scale-dependent terms not present in fixed-order calculations. Specifically, these terms are included in the Sudakov form factor to probe additional sources of higher-order corrections. These factors contribute to the larger theoretical uncertainties observed in the MiNNLO predictions.

CHAPTER 10

Conclusion

The top quark is the heaviest known elementary particle. It is the only quark that decays before hadronization. This property allows for the study of the top quark in isolation. At the LHC, top quarks are commonly produced in pairs and are among the most abundant processes. As such, it offers a reliable testing ground for the parameters of the Standard Model. Any deviations from the Standard Model predictions are strong indicators of new physics. The top quark decays predominantly into a W-boson and a bottom quark, so top quark pairs can produce the $WWbb$ final state. Each of the W-bosons can then decay hadronically or leptonically.

Differential cross-sections considering the $WWbb$ final state have been measured previously. However, those measurements considered just one top quark production process (e.g., $t\bar{t}$ or tW) leading to the $WWbb$ final state as signal while treating all other processes (e.g., single-top tW) as background. The sensitivity of any measured top quark properties is reduced due to lower statistics and larger backgrounds. An interference effect also exists between doubly-resonant $t\bar{t}$ processes and higher-ordered single-top tW processes. This interference effect is the source of a large systematic uncertainty in several SM and BSM measurements.

For the first time, measurements of differential cross-sections of the $WWbb$ inclusive production process are presented. All top quark processes are considered, including doubly-resonant, singly-resonant, and non-resonant contributions. We can perform measurements in the high-p_{T} tails of distributions previously inaccessible due to statistics limitations. These regions are susceptible to the interference effect between $t\bar{t}$ and tW processes and will be useful for BSM analyses in the future. The measured differential cross-sections are compared to several MC generators implementing different methods for simulating the interference effect. The measurements in this thesis and their consistency with present models can be used to constrain the parameters associated with modeling this interference effect.

The results presented in this thesis are based on the full Run 2 (representing 140 fb^{-1}) dataset collected by the ATLAS detector at the LHC. This analysis presents the first measurements in the semileptonic final state of the $WWbb$ process. The final state includes a charged lepton (electron or muon), a neutrino (in the form of $E_{\mathrm{T}}^{\mathrm{miss}}$), two b-jets, and two light-jets. These physics objects are used to reconstruct the $WWbb$ final and define several observables that are sensitive to the interference

effect. This analysis is also unique in that the hadronic-W can be fully reconstructed efficiently. This allows for the $WWbb$ final state to be fully reconstructed.

The baseline selection places cuts on the momentum of the charged lepton, the number of jets, and the number of tagged b-jets. From this, several measurement regions are constructed based on cuts on the hadronic-W kinematics. The nominal measurement region requires a hadronic-W to have significant transverse momentum and an invariant mass consistent with the W-boson mass. Alternative signal regions are also defined to study the effects of loosening these cuts. In each of these regions, several observables of interest are studied. These include kinematics of the base physics objects (e.g., lepton, jets, $E_{\mathrm{T}}^{\mathrm{miss}}$), the hadronic-$W$ kinematics, and the kinematics of the top quark system.

The detector-level observables are unfolded to particle-level by using the TUnfold package. TUnfold uses a regularized unfolding method to correct for detector effects encoded in a migration matrix. The migration matrix is constructed from MC samples and measures the probability of a particle-level observable being reconstructed at the detector level. A separate measurement region was defined to study the effects of using a parton-level definition to construct the migration matrix. The unfolded distributions are then used to compute absolute and normalized differential cross-sections binned in the observables of interest. The cross-sections are compared to several MC generators at NLO accuracy: POWHEG+PYTHIA, POWHEG+HERWIG, and MADGRAPH5_AMC@NLO+PYTHIA.

Several distributions of observables of interest are presented in this measurement. Distributions such as the leading lepton and leading b-jet p_{T} show significant disagreement between the data and the MC generators. There is a clear upward trend where there is an overestimation of the data in the high-p_{T} regions. Other distributions such as the $M(W_{\mathrm{had.}} + b_1 + b_2 + \ell)$ are better described by the MC generators. Also of interest are the noticeable differences between the DR and DS methods. Generally, the DS method is more consistent with the data than the DR method. However, there are other observables where the DR method is more consistent with the data. The $m_{b\ell}^{\mathrm{selected}}$ observable using the dynamically-scaled DR method is the most consistent with the data.

The results of this measurement can be used to tune the parameters of MC generators to better align with the observations. Presently, the interference effects between $t\bar{t}$ and tW processes are poorly modeled and present a large source of uncertainty for many physics analyses. The results of this measurement can be used to constrain these uncertainties towards the development of a combined $bb4l$ model. Rather than treating the $t\bar{t}$ and tW processes as separate entities, the $bb4l$ model treats them as a single process. This model can then be applied throughout the LHC physics program wherever the $t\bar{t}$ and tW processes are relevant.

Bibliography

[1] Abe F. et al. "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab". In: Phys. Rev. Lett. 74 (14 Apr. 1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[2] Abachi S. et al. "Observation of the Top Quark". In: Phys. Rev. Lett. 74 (14 Apr. 1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2632.

[3] Workman R. L. et al. "Review of Particle Physics". In: PTEP 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[4] Aad G. et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: JINST 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[5] Kretzschmar Jan. "Standard Model Physics at the LHC". In: From My Vast Repertoire ...: Guido Altarelli's Legacy. Ed. by Levy Aharon, Forte Stefano, and Ridolfi Giovanni. 2019, pp. 153–171. DOI: 10.1142/9789813238053_0009. arXiv: 1803.10800 [hep-ex].

[6] Aaboud M. et al. "Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: JHEP 11 (2017), p. 191. DOI: 10.1007/JHEP11(2017)191. arXiv: 1708.00727 [hep-ex].

[7] Aad Georges et al. "Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: Eur. Phys. J. C 79.12 (2019). [Erratum: Eur.Phys.J.C 80, 1092 (2020)], p. 1028. DOI: 10.1140/epjc/s10052-019-7525-6. arXiv: 1908.07305 [hep-ex].

[8] Sirunyan A. M. et al. "Measurement of normalized differential $\mathrm{t\bar{t}}$ cross sections in the dilepton channel from pp collisions at $\sqrt{s} = 13$ TeV". In: JHEP 04 (2018), p. 060. DOI: 10.1007/JHEP04(2018)060. arXiv: 1708.07638 [hep-ex].

[9] Sirunyan Albert M et al. "Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton+jets events from pp collisions at $\sqrt{s} = 13$ TeV". In: Phys. Rev. D 97.11 (2018), p. 112003. DOI: 10.1103/PhysRevD.97.112003. arXiv: 1803.08856 [hep-ex].

[10] Aad Georges et al. "Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: Phys. Lett. B 717 (2012), pp. 330–350. DOI: 10.1016/j.physletb.2012.09.031. arXiv: 1205.3130 [hep-ex].

[11] Sirunyan Albert M et al. "Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV". In: Phys. Lett. B 772 (2017), pp. 752–776. DOI: 10.1016/j.physletb.2017.07.047. arXiv: 1610.00678 [hep-ex].

[12] Frixione Stefano et al. "Single-top hadroproduction in association with a W boson". In: JHEP 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph].

[13] Aaboud Morad et al. "Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: Phys. Rev. D 94.5 (2016), p. 052009. DOI: 10.1103/PhysRevD.94.052009. arXiv: 1606.03903 [hep-ex].

[14] Aaboud M. et al. "Search for direct top squark pair production in final states with two leptons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: Eur. Phys. J. C 77.12 (2017), p. 898. DOI: 10.1140/epjc/s10052-017-5445-x. arXiv: 1708.03247 [hep-ex].

[15] Aaboud Morad et al. "Search for bottom squark pair production in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: Eur. Phys. J. C 76.10 (2016), p. 547. DOI: 10.1140/epjc/s10052-016-4382-4. arXiv: 1606.08772 [hep-ex].

[16] Aaboud Morad et al. "Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s}$=13 TeV with the ATLAS detector". In: JHEP 12 (2017), p. 085. DOI: 10.1007/JHEP12(2017)085. arXiv: 1709.04183 [hep-ex].

[17] Aad Georges et al. "Search for gluinos in events with an isolated lepton, jets and missing transverse momentum at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: Eur. Phys. J. C 76.10 (2016), p. 565. DOI: 10.1140/epjc/s10052-016-4397-x. arXiv: 1605.04285 [hep-ex].

[18] Ježo Tomáš et al. "An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects". In: Eur. Phys. J. C 76.12 (2016), p. 691. DOI: 10.1140/epjc/s10052-016-4538-2. arXiv: 1607.04538 [hep-ph].

[19] Denner Ansgar and Pellen Mathieu. "Off-shell production of top-antitop pairs in the lepton+jets channel at NLO QCD". In: JHEP 02 (2018), p. 013. DOI: 10.1007/JHEP02(2018)013. arXiv: 1711.10359 [hep-ph].

[20] Neutelings Isaac. Various Physics Graphics. Graphics with TikZ in LaTeX. 2019–2023. URL: https://tikz.net/.

[21] Peskin Michael E. and Schroeder Daniel V. An Introduction to quantum field theory. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.

[22] Yukawa Hideki. "On the Interaction of Elementary Particles I". In: Proc. Phys. Math. Soc. Jap. 17 (1935), pp. 48–57. DOI: 10.1143/PTPS.1.1.

[23] Griffiths David. Introduction to elementary particles. New York, NY: Wiley, 2008. ISBN: 978-3-527-40601-2.

[24] Ismail Ahmed, Mammen Abraham Roshan, and Kling Felix. "Neutral current neutrino interactions at FASERν". In: Phys. Rev. D 103.5 (2021), p. 056014. DOI: 10.1103/PhysRevD.103.056014. arXiv: 2012.10500 [hep-ph].

[25] Weinberg Steven. "A Model of Leptons". In: Phys. Rev. Lett. 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[26] Salam Abdus. "Weak and Electromagnetic Interactions". In: Conf. Proc. C 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.

[27] Horejsi Jiri. Fundamentals of electroweak theory. 2002. arXiv: 2210.04526 [hep-ph].

[28] ATLAS Collaboration. "Top quark mass and properties summary plots June 2023". In: Cern Document Server (June 2023). URL: https://cds.cern.ch/record/2862523.

[29] Bernreuther W. et al. "Top quark pair production and decay at hadron colliders". In: Nucl. Phys. B 690 (2004), pp. 81–137. DOI: 10.1016/j.nuclphysb.2004.04.019. arXiv: hep-ph/0403035.

[30] Melnikov Kirill and Schulze Markus. "NLO QCD corrections to top quark pair production in association with one hard jet at hadron colliders". In: Nucl. Phys. B 840 (2010), pp. 129–159. DOI: 10.1016/j.nuclphysb.2010.07.003. arXiv: 1004.3284 [hep-ph].

[31] Bernreuther Werner and Si Zong-Guo. "Distributions and correlations for top quark pair production and decay at the Tevatron and LHC." In: Nucl. Phys. B 837 (2010), pp. 90–121. DOI: 10.1016/j.nuclphysb.2010.05.001. arXiv: 1003.3926 [hep-ph].

[32] Denner A. et al. "NLO QCD corrections to WWbb production at hadron colliders". In: Phys. Rev. Lett. 106 (2011), p. 052001. DOI: 10.1103/PhysRevLett.106.052001. arXiv: 1012.3975 [hep-ph].

[33] Melnikov Kirill and Schulze Markus. "NLO QCD corrections to top quark pair production and decay at hadron colliders". In: JHEP 08 (2009), p. 049. DOI: 10.1088/1126-6708/2009/08/049. arXiv: 0907.3090 [hep-ph].

[34] Uhlemann C. F. and Kauer N. "Narrow-width approximation accuracy". In: Nucl. Phys. B 814 (2009), pp. 195–211. DOI: 10.1016/j.nuclphysb.2009.01.022. arXiv: 0807.4112 [hep-ph].

[35] White Chris D. et al. "Isolating Wt production at the LHC". In: JHEP 11 (2009), p. 074. DOI: 10.1088/1126-6708/2009/11/074. arXiv: 0908.0631 [hep-ph].

[36] Gardi Einan and Magnea Lorenzo. "Infrared singularities in QCD amplitudes". In: Nuovo Cim. C 32N5-6 (2009). Ed. by Greco Mario, pp. 137–157. DOI: 10.1393/ncc/i2010-10528-x. arXiv: 0908.3273 [hep-ph].

[37] Frixione Stefano et al. "The MCaNLO 4.0 Event Generator". In: (Oct. 2010). arXiv: 1010.0819 [hep-ph].

[38] Frixione Stefano and Webber Bryan R. "Matching NLO QCD computations and parton shower simulations". In: JHEP 06 (2002), p. 029. DOI: 10.1088/1126-6708/2002/06/029. arXiv: hep-ph/0204244.

[39] Aaboud Morad et al. "Probing the quantum interference between singly and doubly resonant top-quark production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: Phys. Rev. Lett. 121.15 (2018), p. 152002. DOI: 10.1103/PhysRevLett.121.152002. arXiv: 1806.04667 [hep-ex].

[40] Lester C. G. and Summers D. J. "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders". In: Phys. Lett. B 463 (1999), pp. 99–103. DOI: 10.1016/S0370-2693(99)00945-4. arXiv: hep-ph/9906349.

[41] Barr Alan, Lester Christopher, and Stephens P. "m(T2): The Truth behind the glamour". In: J. Phys. G 29 (2003), pp. 2343–2363. DOI: 10.1088/0954-3899/29/10/304. arXiv: hep-ph/0304226.

[42] "LHC Machine". In: JINST 3 (2008). Ed. by Evans Lyndon and Bryant Philip, S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[43] Llewellyn Smith Chris. "Genesis of the Large Hadron Collider". In: Phil. Trans. Roy. Soc. Lond. A 373.2032 (2014), p. 20140037. DOI: 10.1098/rsta.2014.0037.

[44] Aad Georges et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: Phys. Lett. B 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[45] "LHC Design Report Vol.1: The LHC Main Ring". In: (June 2004). Ed. by Bruning Oliver S. et al. DOI: 10.5170/CERN-2004-003-V-1.

[46] "LHC Design Report. 2. The LHC infrastructure and general services". In: (Nov. 2004). Ed. by Buning O. et al. DOI: 10.5170/CERN-2004-003-V-2.

[47] Welton Robert et al. "Negative hydrogen ion sources for particle accelerators: Sustainability issues and recent improvements in long-term operations". In: J. Phys. Conf. Ser. 2244.1 (2022), p. 012045. DOI: 10.1088/1742-6596/2244/1/012045.

[48] CERN. "CERN's Accelerator Complex". In: CERN Science Home Page (2022). URL: https://home.cern/science/accelerators/accelerator-complex.

[49] Lopienska Ewa. "The CERN accelerator complex". In: CERN Document Server (2022). URL: https://cds.cern.ch/record/2800984.

[50] Alexopoulos A. et al. "Noninvasive LHC transverse beam size measurement using inelastic beam-gas interactions". In: Phys. Rev. Accel. Beams 22.4 (2019), p. 042801. DOI: 10.1103/PhysRevAccelBeams.22.042801.

[51] Aad G. et al. "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. DOI: 10.1140/epjc/s10052-023-11747-w. arXiv: 2212.09379 [hep-ex].

[52] Aad Georges et al. "The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3". In: (May 2023). arXiv: 2305.16623 [physics.ins-det].

[53] ATLAS Collaboration. "ATLAS inner detector: Technical design report. Vol. 1". In: Cern Document Server (Apr. 1997). URL: https://cds.cern.ch/record/331064.

[54] ATLAS Collaboration. "ATLAS inner detector: Technical design report. Vol. 2". In: Cern Document Server (Apr. 1997). URL: https://cds.cern.ch/record/331064.

[55] Aad G. et al. "ATLAS pixel detector electronics and sensors". In: JINST 3 (2008), P07007. DOI: 10.1088/1748-0221/3/07/P07007.

[56] Abbott B. et al. "Production and Integration of the ATLAS Insertable B-Layer". In: JINST 13.05 (2018), T05008. DOI: 10.1088/1748-0221/13/05/T05008. arXiv: 1803.00844 [physics.ins-det].

[57] Aad Georges et al. "Operation and performance of the ATLAS semiconductor tracker". In: JINST 9 (2014), P08009. DOI: 10.1088/1748-0221/9/08/P08009. arXiv: 1404.7473 [hep-ex].

[58] Mindur Bartosz. "ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider". In: Nucl. Instrum. Meth. A 845 (2017). Ed. by Badurek G. et al., pp. 257–261. DOI: 10.1016/j.nima.2016.04.026.

[59] Aad G. et al. "The ATLAS Inner Detector commissioning and calibration". In: Eur. Phys. J. C 70 (2010), pp. 787–821. DOI: 10.1140/epjc/s10052-010-1366-7. arXiv: 1004.5293 [physics.ins-det].

[60] Fabjan C. W. and Gianotti F. "Calorimetry for particle physics". In: Rev. Mod. Phys. 75 (2003), pp. 1243–1286. DOI: 10.1103/RevModPhys.75.1243.

[61] Aleksa Martin and Diemoz Marcella. "Discussion on the electromagnetic calorimeters of ATLAS and CMS". In: Nucl. Instrum. Meth. A 732 (2013), pp. 442–450. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2013.05.195. URL: https://www.sciencedirect.com/science/article/pii/S0168900213008553.

[62] Koletsou Iro. "The ATLAS liquid argon calorimeter at the LHC". In: Nucl. Instrum. Meth. A 628 (2011). Ed. by Bergauer Thomas et al., pp. 351–354. DOI: 10.1016/j.nima.2010.06.351.

[63] Tsai Yung-Su. "Pair Production and Bremsstrahlung of Charged Leptons". In: Rev. Mod. Phys. 46 (1974). [Erratum: Rev.Mod.Phys. 49, 421–423 (1977)], p. 815. DOI: 10.1103/RevModPhys.46.815.

[64] ATLAS Collaboration. "ATLAS muon spectrometer: Technical design report". In: Cern Document Server (June 1997). URL: https://cds.cern.ch/record/331068.

[65] Dubbert J. et al. "Integration, commissioning and installation of monitored drift tube chambers for the ATLAS barrel muon spectrometer". In: Nucl. Instrum. Meth. A 572 (2007). Ed. by Cervelli Franco et al., pp. 53–56. DOI: 10.1016/j.nima.2006.10.230.

[66] Köhler Nicolas Maximilian. "Searches for the Supersymmetric Partner of the Top Quark, Dark Matter and Dark Energy at the ATLAS Experiment". PhD thesis. Munich, Tech. U., 2018. DOI: 10.1007/978-3-030-25988-4.

[67] Bauer F. et al. "Construction and Test of MDT Chambers for the ATLAS Muon Spectrometer". In: Nucl. Instrum. Meth. A 461 (2001). Ed. by Batignani G. et al., pp. 17–20. DOI: 10.1016/S0168-9002(00)01156-6. arXiv: 1604.02000 [physics.ins-det].

[68] Nikolopoulos Konstantinos et al. "Cathode strip chambers in ATLAS : Installation, commissioning and in situ performance". In: 2008 IEEE Nuclear Science Symposium. 2008, pp. 2819–2824. DOI: 10.1109/NSSMIC.2008.4774958.

[69] Cattani Giordano. "The resistive plate chambers of the ATLAS experiment: Performance studies". In: J. Phys. Conf. Ser. 280 (2011). Ed. by Cannuccia E. et al., p. 012001. DOI: 10.1088/1742-6596/280/1/012001.

[70] D'Amico Valerio. "The New Small Wheel upgrade project of the ATLAS Experiment". In: PoS EPS-HEP2021 (2022), p. 756. DOI: 10.22323/1.398.0756.

[71] Scannicchio D. A. "ATLAS trigger and data acquisition: Capabilities and commissioning". In: Nucl. Instrum. Meth. A 617 (2010). Ed. by Chiarelli Giorgio et al., pp. 306–309. DOI: 10.1016/j.nima.2009.06.114.

[72] Buckley Andy et al. "General-purpose event generators for LHC physics". In: Phys. Rept. 504 (2011), pp. 145–233. DOI: 10.1016/j.physrep.2011.03.005. arXiv: 1101.2599 [hep-ph].

[73] Ball Richard D. et al. "The path to proton structure at 1% accuracy". In: Eur. Phys. J. C 82.5 (2022), p. 428. DOI: 10.1140/epjc/s10052-022-10328-7. arXiv: 2109.02653 [hep-ph].

[74] Dobbs M. A. et al. "Les Houches guidebook to Monte Carlo generators for hadron collider physics". In: 3rd Les Houches Workshop on Physics at TeV Colliders. Mar. 2004, pp. 411–459. arXiv: hep-ph/0403045.

[75] Höche Stefan and Prestel Stefan. "The midpoint between dipole and parton showers". In: Eur. Phys. J. C 75.9 (2015), p. 461. DOI: 10.1140/epjc/s10052-015-3684-2. arXiv: 1506.05057 [hep-ph].

[76] Andersson Bo et al. "Parton Fragmentation and String Dynamics". In: Phys. Rept. 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[77] Webber B. R. "A QCD Model for Jet Fragmentation Including Soft Gluon Interference". In: Nucl. Phys. B 238 (1984), pp. 492–528. DOI: 10.1016/0550-3213(84)90333-X.

[78] Agostinelli S. et al. "GEANT4–a simulation toolkit". In: Nucl. Instrum. Meth. A 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[79] Aad Georges et al. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: Eur. Phys. J. C 76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[80] Bothmann Enrico et al. "Event Generation with Sherpa 2.2". In: SciPost Phys. 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[81] Sjöstrand Torbjörn et al. "An introduction to PYTHIA 8.2". In: Comput. Phys. Commun. 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[82] Bellm Johannes et al. "Herwig 7.0/Herwig++ 3.0 release note". In: Eur. Phys. J. C 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[83] Bahr M. et al. "Herwig++ Physics and Manual". In: Eur. Phys. J. C 58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[84] Frixione Stefano, Nason Paolo, and Ridolfi Giovanni. "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: JHEP 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].

[85] Nason Paolo. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: JHEP 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[86] Frixione Stefano, Nason Paolo, and Oleari Carlo. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: JHEP 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[87] Alioli Simone et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: JHEP 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[88] Alwall J. et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: JHEP 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[89] Monni Pier Francesco et al. "MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers". In: JHEP 05 (2020). [Erratum: JHEP 02, 031 (2022)], p. 143. DOI: 10.1007/JHEP05(2020)143. arXiv: 1908.06987 [hep-ph].

[90] Mazzitelli Javier et al. "Top-pair production at the LHC with MINNLO$_{PS}$". In: JHEP 04 (2022), p. 079. DOI: 10.1007/JHEP04(2022)079. arXiv: 2112.12135 [hep-ph].

[91] Ball Richard D. et al. "Parton distributions for the LHC Run II". In: JHEP 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[92] Schumann Steffen and Krauss Frank. "A Parton shower algorithm based on Catani-Seymour dipole factorisation". In: JHEP 03 (2008), p. 038. DOI: 10.1088/1126-6708/2008/03/038. arXiv: 0709.1027 [hep-ph].

[93] Ball Richard D. et al. "Parton distributions with LHC data". In: Nucl. Phys. B 867 (2013), pp. 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[94] Harland-Lang L. A. et al. "Parton distributions in the LHC era: MMHT 2014 PDFs". In: Eur. Phys. J. C 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph].

[95] Buckley Andrew. "ATLAS Pythia 8 tunes to 7 TeV data". In: 6th International Workshop on Multiple Partonic Interactions at the LHC. Dec. 2014, p. 29.

[96] Seymour Michael H. and Siodmok Andrzej. "Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data". In: JHEP 10 (2013), p. 113. DOI: 10.1007/JHEP10(2013)113. arXiv: 1307.5015 [hep-ph].

[97] Glasman Claudia et al. "Measurements of top-quark pair differential cross sections in the ℓ+jets channel in pp collisions at $\sqrt{s}$ = 13 TeV using 139 fb^{-1} of ATLAS data". In: Cern Document Server (Aug. 2021). URL: https://cds.cern.ch/record/2765035.

[98] Campbell John, Neumann Tobias, and Sullivan Zack. "Single-top-quark production in the t-channel at NNLO". In: JHEP 02 (2021), p. 040. DOI: 10.1007/JHEP02(2021)040. arXiv: 2012.01574 [hep-ph].

[99] Ball Richard D. et al. "The PDF4LHC21 combination of global PDF fits for the LHC Run III". In: J. Phys. G 49.8 (2022), p. 080501. DOI: 10.1088/1361-6471/ac7216. arXiv: 2203.05506 [hep-ph].

[100] Kidonakis Nikolaos and Yamanaka Nodoka. "Higher-order corrections for tW production at high-energy hadron colliders". In: JHEP 05 (2021), p. 278. DOI: 10.1007/JHEP05(2021)278. arXiv: 2102.11300 [hep-ph].

[101] Kidonakis Nikolaos. "NNLL resummation for s-channel single top quark production". In: Phys. Rev. D 81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph].

[102] ATLAS Collaboration. "Studies on top-quark Monte Carlo modeling for Top2016". In: Cern Document Server (Oct. 2016). URL: https://cds.cern.ch/record/2216168.

[103] Czakon Michal, Heymes David, and Mitov Alexander. "Dynamical scales for multi-TeV top-pair production at the LHC". In: JHEP 04 (2017), p. 071. DOI: 10.1007/JHEP04(2017)071. arXiv: 1606.03350 [hep-ph].

[104] Martin A. D. et al. "Parton distributions for the LHC". In: Eur. Phys. J. C 63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[105] Meloni Federico. "Primary vertex reconstruction with the ATLAS detector". In: JINST 11.12 (2016). Ed. by Gemme Claudia and Rossi Leonardo, p. C12060. DOI: 10.1088/1748-0221/11/12/C12060.

[106] ATLAS Collaboration. "Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach". In: Cern Document Server (Dec. 2017). URL: https://cds.cern.ch/record/2298955.

[107] Aaboud M. et al. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2". In: Eur. Phys. J. C 77.10 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].

[108] ATLAS Collaboration. "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung". In: Cern Document Server (May 2012). URL: https://cds.cern.ch/record/1449796.

[109] Aaboud Morad et al. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s}$ = 13 TeV". In: Eur. Phys. J. C 79.8 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [physics.ins-det].

[110] Aad Georges et al. "Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data". In: JINST 19.02 (2024), P02009. DOI: 10.1088/1748-0221/19/02/P02009. arXiv: 2309.05471 [hep-ex].

[111] Aaboud Morad et al. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: JINST 14.03 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex].

[112] Aad Georges et al. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: JINST 14.12 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[113] Aad Georges et al. "Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment". In: (Aug. 2023). arXiv: 2308.13362 [hep-ex].

[114] Aad Georges et al. "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. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex].

[115] Aad Georges et al. "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. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[116] Aad Georges et al. "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. DOI: 10.1140/epjc/s10052-023-11584-x. arXiv: 2212.07338 [hep-ex].

[117] Cacciari Matteo, Salam Gavin P., and Soyez Gregory. "The anti-k_t jet clustering algorithm". In: JHEP 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[118] Aaboud Morad et al. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: Eur. Phys. J. C 77.7 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[119] Catani S. et al. "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions". In: Nucl. Phys. B 406 (1993), pp. 187–224. DOI: 10.1016/0550-3213(93)90166-M.

[120] ATLAS Collaboration. "Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run". In: Cern Document Server (Aug. 2016). URL: https://cds.cern.ch/record/2160731.

[121] Aad Georges et al. "ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset". In: Eur. Phys. J. C 83.7 (2023), p. 681. DOI: 10.1140/epjc/s10052-023-11699-1. arXiv: 2211.16345 [physics.data-an].

[122] Aaboud Morad et al. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: Eur. Phys. J. C 78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[123] Lehmann Konstantin and Stelzer Bernd. "The Fake Factor Method and its relation to the Matrix Method". In: Nucl. Instrum. Meth. A 1054 (2023), p. 168376. DOI: 10.1016/j.nima.2023.168376.

[124] Aad G. et al. "Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC". In: JINST 18.11 (2023), T11004. DOI: 10.1088/1748-0221/18/11/T11004. arXiv: 2211.16178 [hep-ex].

[125] Arnold Hannah et al. "Measurement of the b-jet identification efficiency for high transverse momentum jets in $t\bar{t}$ events in the lepton + jets channel with the ATLAS detector using Run 2 data". In: Cern Document Server (May 2020). URL: https://cds.cern.ch/record/2718513.

[126] Demartin Federico et al. "tWH associated production at the LHC". In: Eur. Phys. J. C 77.1 (2017), p. 34. DOI: 10.1140/epjc/s10052-017-4601-7. arXiv: 1607.05862 [hep-ph].

[127] ATLAS Collaboration. "Measurement of the top-quark mass in $t\bar{t}$ $\rightarrow$ dilepton events with the ATLAS experiment using the template method in 13 TeV pp collision data". In: Cern Document Server (Sept. 2022). URL: https://cds.cern.ch/record/2826701.

[128] Brooks Helen and Skands Peter. "Coherent showers in decays of colored resonances". In: Phys. Rev. D 100.7 (2019), p. 076006. DOI: 10.1103/PhysRevD.100.076006. arXiv: 1907.08980 [hep-ph].

[129] Butterworth Jon et al. "PDF4LHC recommendations for LHC Run II". In: J. Phys. G 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[130] Aad Georges et al. "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. DOI: 10.1140/epjc/s10052-021-09402-3. arXiv: 2007.02645 [hep-ex].

[131] Schmitt Stefan. "TUnfold: an algorithm for correcting migration effects in high energy physics". In: JINST 7 (2012), T10003. DOI: 10.1088/1748-0221/7/10/T10003. arXiv: 1205.6201 [physics.data-an].

[132] Blobel Volker. "An Unfolding method for high-energy physics experiments". In: Conference on Advanced Statistical Techniques in Particle Physics. Aug. 2002, pp. 258–267. arXiv: hep-ex/0208022.

[133] Hansen Per Christian. "The L-curve and its use in the numerical treatment of inverse problems". English. In: Computational Inverse Problems in Electrocardiology. WIT Press, 2000.

[134] ATLAS Collaboration. "Measurements of differential cross-sections of $WbWb$ production in the dilepton channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: Cern Document Server (Mar. 2023). URL: https://cds.cern.ch/record/2853756.

[135] Godbole Rohini et al. "Boosted Top quark polarization". In: Phys. Rev. D 100.5 (2019), p. 056010. DOI: 10.1103/PhysRevD.100.056010. arXiv: 1902.08096 [hep-ph].

[136] Shelton Jessie. "Polarized tops from new physics: signals and observables". In: Phys. Rev. D 79 (2009), p. 014032. DOI: 10.1103/PhysRevD.79.014032. arXiv: 0811.0569 [hep-ph].

[137] Perelstein Maxim and Weiler Andreas. "Polarized Tops from Stop Decays at the LHC". In: JHEP 03 (2009), p. 141. DOI: 10.1088/1126-6708/2009/03/141. arXiv: 0811.1024 [hep-ph].

[138] Tweedie Brock. "Better Hadronic Top Quark Polarimetry". In: Phys. Rev. D 90.9 (2014), p. 094010. DOI: 10.1103/PhysRevD.90.094010. arXiv: 1401.3021 [hep-ph].

[139] Kitadono Yoshio and Li Hsiang-nan. "Jet substructures of boosted polarized hadronic top quarks". In: Phys. Rev. D 93.5 (2016), p. 054043. DOI: 10.1103/PhysRevD.93.054043. arXiv: 1511.08675 [hep-ph].

[140] Krohn David, Shelton Jessie, and Wang Lian-Tao. "Measuring the Polarization of Boosted Hadronic Tops". In: JHEP 07 (2010), p. 041. DOI: 10.1007/JHEP07(2010)041. arXiv: 0909.3855 [hep-ph].

[141] Kitadono Yoshio and Li Hsiang-nan. "Jet substructures of boosted polarized top quarks". In: Phys. Rev. D 89.11 (2014), p. 114002. DOI: 10.1103/PhysRevD.89.114002. arXiv: 1403.5512 [hep-ph].

[142] Bhattacherjee Biplob, Mandal Sourav K., and Nojiri Mihoko. "Top Polarization and Stop Mixing from Boosted Jet Substructure". In: JHEP 03 (2013), p. 105. DOI: 10.1007/JHEP03(2013)105. arXiv: 1211.7261 [hep-ph].

APPENDIX A

Neutrino Momentum Derivation

In the decay of the W to a lepton and a neutrino, the longitudinal component of the neutrino ($p_{\nu,z}$) can be calculated by enforcing momentum conservation. The four-momentum of the leptonically-decaying W-boson (p_W) can be written in terms of the daughter lepton (p_ℓ) and neutrino (p_ν) momenta.

$$p_W = p_\ell + p_\nu \tag{A.1}$$

The four-momentum vector of a particle is the relativistic analog of the classical momentum vector, which describes the energy and momentum of a particle in spacetime. Mathematically, the four-momentum vector is expressed as $p_\mu = (E, \vec{p})$ with the index, $\mu = 1, 2, 3, 4$. The four-momentum of the lepton and neutrino can be expanded as follows.

$$p_\ell = (E_\ell, p_{\ell,\mathrm{x}}, p_{\ell,\mathrm{y}}, p_{\ell,\mathrm{z}}) \tag{A.2a}$$
$$p_\nu = (E_\nu, p_{\nu,\mathrm{x}}, p_{\nu,\mathrm{y}}, p_{\nu,\mathrm{z}}) \tag{A.2b}$$

The magnitude of the four-momentum gives the invariant mass of the particle.

$$|p| = p^\mu p_\mu = \sqrt{E^2 - |\vec{p}|^2} = m \tag{A.3}$$

Squaring both sides of Eq. A.1 and applying the energy-momentum relation ($E^2 = m^2 + |\vec{p}|^2$) yields the following equations.

$$p_W^2 = (p_\ell + p_\nu)^2 \tag{A.4a}$$
$$m_W^2 = (p_\ell + p_\nu)^2 \tag{A.4b}$$

The W is assumed to decay on-shell with the mass constrained to its true mass, $m_W^2 = 80.385\,\mathrm{GeV}$. The right-hand side of Eq. A.4b can then be expanded to yield the following.

$$m_W^2 = p_\ell^2 + p_\nu^2 + 2p_\ell p_\nu \tag{A.5}$$

In the ultra-relativistic limit ($p_{\ell,\nu} \gg m_{\ell,\nu}$), the invariant masses of the lepton and neutrino can be considered to be infinitesimally small compared to its momentum ($m_{\ell,\nu} \to 0$).

$$m_W^2 = 2p_\ell p_\nu \tag{A.6}$$

The four-momenta components of the lepton and neutrino can then be inserted into the right-hand side of Eq. A.6.

$$m_W^2 = 2\left(E_\ell E_\nu - p_{\ell,x}p_{\nu,x} - p_{\ell,y}p_{\nu,y} - p_{\ell,z}p_{\nu,z}\right) \tag{A.7a}$$

$$m_W^2 = 2\left(E_\ell\sqrt{m_\nu^2 + |\vec{p}_\nu|^2} - p_{\ell,x}p_{\nu,x} - p_{\ell,y}p_{\nu,y} - p_{\ell,z}p_{\nu,z}\right) \tag{A.7b}$$

$$m_W^2 = 2\left(E_\ell|\vec{p}_\nu| - p_{\ell,x}p_{\nu,x} - p_{\ell,y}p_{\nu,y} - p_{\ell,z}p_{\nu,z}\right) \tag{A.7c}$$

$$m_W^2 = 2\left(E_\ell\sqrt{p_{\nu,x}^2 + p_{\nu,y}^2 + p_{\nu,z}^2} - p_{\ell,x}p_{\nu,x} - p_{\ell,y}p_{\nu,y} - p_{\ell,z}p_{\nu,z}\right) \tag{A.7d}$$

In Eq. A.7b, the energy-momentum relation is inserted for the total energy of the neutrino ($E_\nu = \sqrt{m_\nu^2 + |\vec{p}_\nu|^2}$). The ultra-relativistic limit is again applied in Eq. A.7c ($p_\nu \gg m_\nu$). This results in Eq. A.7d where the entire equation is only a function of the energy of the lepton, three-momentum components of the neutrino and the constrained mass of the W. The above can be expanded to yield a quadratic equation for $p_{\nu,z}$.

$$\begin{aligned}
&\underbrace{(p_{\ell,z}^2 - E_\ell^2)}_{\text{A}} \times p_{\nu,z}^2 + \\
&\underbrace{(m_W^2 + 2p_{\ell,z}\vec{p}_{\ell,T}\cdot\vec{p}_{\nu,T})}_{\text{B}} \times p_{\nu,z} + \\
&\underbrace{\frac{m_W^4}{4} + (\vec{p}_{\ell,T}\cdot\vec{p}_{\nu,T})^2 + m_W^2\vec{p}_{\ell,T}\cdot\vec{p}_{\nu,T} - E_\ell^2|\vec{p}_{\nu,T}|^2}_{\text{C}} = 0
\end{aligned} \tag{A.8}$$

The constants, A, B, C, can then be inserted into the standard quadratic equation to solve for the two solutions for $p_{\nu,z}$. After one of the two solutions is selected, the full neutrino momentum vector can be reconstructed.

APPENDIX B

Additional Cutflow Tables

This appendix lists additional cutflow tables for the various measurement regions. Specifically, the absolute and marginal event yield cutflow tables are shown. The absolute event yield is defined as the number of events that pass the i-th cut relative to the total number of events. The marginal event yield is defined as the number of events that pass the i-th cut relative to the $(i-1)$th cut.

B.1 Baseline Selection

Tables B.1 and B.2 lists the absolute and marginal event yields, respectively, for each of the baseline requirements.

Table B.1: Baseline Selection Absolute Event Yields. The baseline selection represents the cuts that all events must pass before moving on to the region-specific cuts. The absolute event yield is computed for data and each MC source, while the total summed MC is computed separately for the single-top DR and DS samples.

Requirement	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
All events	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$\geqslant 3$ jets	0.23	0.26	0.39	0.70	0.68	0.83	0.44	0.44	0.45
Pass electron crack veto	0.23	0.26	0.39	0.70	0.68	0.83	0.44	0.44	0.45
One loose lepton	0.23	0.15	0.38	0.65	0.63	0.74	0.39	0.39	0.42
Pass lepton triggers	0.22	0.15	0.36	0.63	0.61	0.72	0.38	0.38	0.42
Trigger match loose lepton	0.21	0.13	0.34	0.59	0.57	0.66	0.35	0.35	0.40
Pass bad muon veto	0.21	0.13	0.34	0.59	0.57	0.66	0.35	0.35	0.40
One tight lepton	0.19	0.12	0.31	0.54	0.52	0.60	0.32	0.32	0.27
Exactly one lepton	0.18	0.11	0.30	0.53	0.51	0.59	0.31	0.31	0.27
Lepton $p_{\mathrm{t}} > 30$ GeV	0.17	0.11	0.28	0.50	0.49	0.56	0.30	0.30	0.25
$E_{\mathrm{t}}^{\mathrm{miss}} + m_{\mathrm{t}}^{W} > 60$ GeV	0.15	0.07	0.25	0.45	0.43	0.50	0.26	0.26	0.22
$\geqslant 2$ b-jets	0.01	0.00	0.10	0.12	0.11	0.24	0.08	0.08	0.07

Table B.2: Baseline Selection Marginal Event Yields. The baseline selection represents the cuts that all events must pass before moving on to the region-specific cuts. The marginal event yield is computed for data and each MC source, while the total summed MC is computed separately for the single-top DR and DS samples.

Requirement	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
All events	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$\geqslant 3$ jets	0.23	0.26	0.39	0.70	0.68	0.83	0.44	0.44	0.45
Pass electron crack veto	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
One loose lepton	0.99	0.58	0.95	0.93	0.93	0.89	0.89	0.89	0.92
Pass lepton triggers	0.97	0.97	0.96	0.97	0.97	0.97	0.97	0.97	1.00
Trigger match loose lepton	0.94	0.89	0.92	0.93	0.93	0.93	0.93	0.93	0.97
Pass bad muon veto	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
One tight lepton	0.91	0.89	0.91	0.92	0.92	0.91	0.91	0.91	0.67
Exactly one lepton	0.97	0.97	0.98	0.98	0.98	0.98	0.97	0.97	1.00
Lepton $p_{\mathrm{t}} > 30$ GeV	0.95	0.94	0.94	0.96	0.96	0.95	0.95	0.95	0.92
$E_{\mathrm{t}}^{\mathrm{miss}} + m_{\mathrm{t}}^{W} > 60$ GeV	0.87	0.69	0.88	0.89	0.89	0.89	0.88	0.87	0.86
$\geqslant 2$ b-jets	0.04	0.07	0.42	0.26	0.25	0.47	0.32	0.32	0.31

B.2 Measurement Regions

Tables B.3 and B.4 lists the absolute and marginal event yields, respectively, for each of the four measurement regions defined previously. The absolute event yields are computed relative to the total number of events. The first row of marginal event yields in each measurement region is computed relative to the number of events passing the final baseline requirement.

Table B.3: Measurement Regions Absolute Event Yields. A unique set of cuts applied to events passing the baseline selection defines a measurement region. The absolute event yields are listed for each requirement in the three measurement regions. The total summed MC is computed separately for the single-top DR and DS samples.

$W_{\text{had.}}$ (particle)									
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
$p_{\text{T}}^{W_{\text{had.}}} > 60$ GeV	0.00	0.00	0.04	0.06	0.05	0.12	0.04	0.04	0.03
$60 < M^{W_{\text{had.}}} < 100$ GeV	0.00	0.00	0.00	0.02	0.02	0.04	0.01	0.01	0.01
$W_{\text{had.}}$ (loose)									
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
$p_{\text{T}}^{W_{\text{had.}}} > 30$ GeV	0.01	0.00	0.09	0.10	0.10	0.22	0.08	0.08	0.06
$M^{W_{\text{had.}}} > 40$ GeV	0.00	0.00	0.02	0.04	0.04	0.10	0.03	0.03	0.03
$W_{\text{lep.}}$ (loose)									
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
$W_{\text{had.}}$ (loose)	0.00	0.00	0.02	0.04	0.04	0.10	0.03	0.03	0.03
$E_{\text{T}}^{\text{miss}} > 30$ GeV	0.00	0.00	0.02	0.04	0.04	0.09	0.03	0.03	0.02
$p_{\text{T}}^{W_{\text{lep.}}} > 30$ GeV	0.00	0.00	0.02	0.04	0.03	0.08	0.03	0.03	0.02

Table B.4: Measurement Regions Marginal Event Yields. A unique set of cuts applied to events passing the baseline selection defines a measurement region. The absolute event yields are listed for each requirement in the three measurement regions. The total summed MC is computed separately for the single-top DR and DS samples.

			$W_{\text{had.}}$ (particle)						
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
$p_{\text{T}}^{W_{\text{had.}}} > 60$ GeV	0.40	0.36	0.43	0.50	0.46	0.52	0.51	0.51	0.49
$60 < M^{W_{\text{had.}}} < 100$ GeV	0.10	0.12	0.10	0.33	0.33	0.30	0.29	0.29	0.28
			$W_{\text{had.}}$ (loose)						
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
$p_{\text{T}}^{W_{\text{had.}}} > 30$ GeV	0.86	0.85	0.87	0.91	0.90	0.91	0.91	0.91	0.90
$M^{W_{\text{had.}}} > 40$ GeV	0.22	0.25	0.21	0.43	0.42	0.44	0.42	0.42	0.42
			$W_{\text{lep.}}$ (loose)						
Requirement ($\times 10^3$)	W+jets	Other V/VV	Other t	tW (DR)	tW (DS)	$t\bar{t}$	$\sum$ MC (DR)	$\sum$ MC (DS)	Data
$p_{\text{T}}^{W_{\text{had.}}} > 30$ GeV	0.86	0.85	0.87	0.91	0.90	0.91	0.91	0.91	0.90
$W_{\text{had.}}$ (loose)	0.22	0.25	0.21	0.43	0.42	0.44	0.42	0.42	0.42
$E_{\text{T}}^{\text{miss}} > 30$ GeV	0.89	0.83	0.93	0.92	0.92	0.92	0.92	0.92	0.92
$p_{\text{T}}^{W_{\text{lep.}}} > 30$ GeV	0.89	0.91	0.92	0.94	0.93	0.92	0.92	0.92	0.92

APPENDIX C

Alternative Signal Regions

This appendix holds additional plots for the alternative signal regions, $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose). The plots include the detector-level control plots, migration matrices, acceptance/efficiency correction factors, covariance/correlation matrices, and the final differential cross-section plots. The each section, the following kinematic and spatial observables are shown: lepton transverse momentum, leading b-jet transverse momentum, subleading b-jet transverse momentum, leading light-jet transverse momentum, hadronic-W transverse momentum, selected $(b + \ell)$ invariant mass, selected $(b + W_{\text{had.}})$ invariant mass, selected $(W_{\text{had.}} + b_1 + b_2 + \ell)$ invariant mass, modified minimax mass, hadronic-W y, selected $\Delta R(b, \ell)$, and selected $\Delta R(b, W_{\text{had.}})$. These variables are first introduced and described in Section 7.3.

C.1 Detector-Level Control Plots

Figures C.1 and C.2 plot the detector-level observables of interest for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

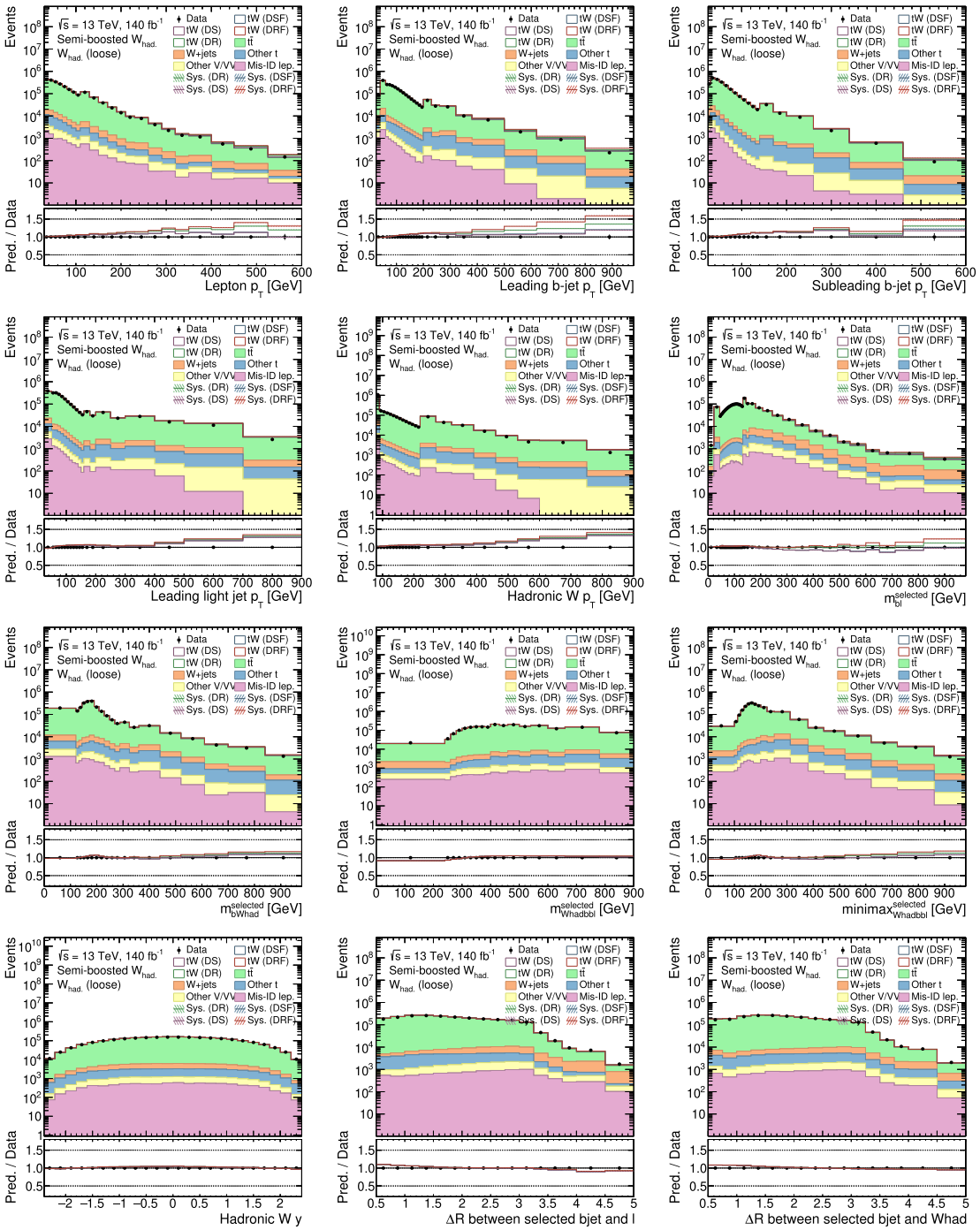


Figure C.1: Detector-Level Observables in the $W_{\text{had.}}$ (loose) Region. The detector-level control distributions of the observables of interest plot the full Run 2 data and the expected MC contributions. The shaded bands represent the statistical uncertainties in the total MC prediction. The bottom panel in each plot shows the ratio between the data and the MC prediction, computed separately for the DR and DS single-top samples. Refer to Section 7.3 for a description of the observables of interest.

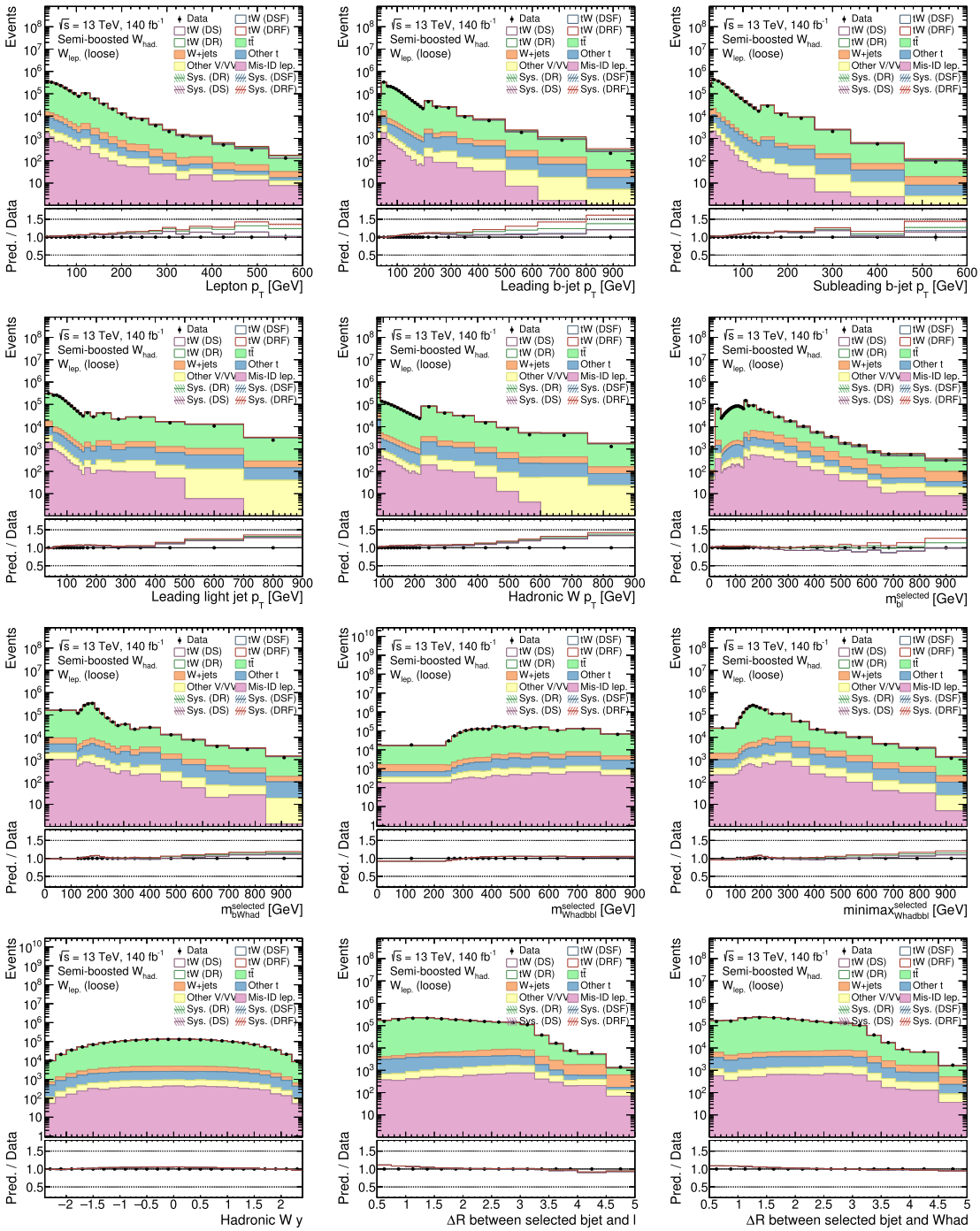


Figure C.2: Detector-Level Observables in the $W_{\text{lep.}}$ (loose) Region. The detector-level control distributions of the observables of interest plot the full Run 2 data and the expected MC contributions. The shaded bands represent the statistical uncertainties in the total MC prediction. The bottom panel in each plot shows the ratio between the data and the MC prediction, computed separately for the DR and DS single-top samples. Refer to Section 7.3 for a description of the observables of interest.

C.2 Systematics Summary Plots

Figures C.3 and C.4 plot the summary of the systematics for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

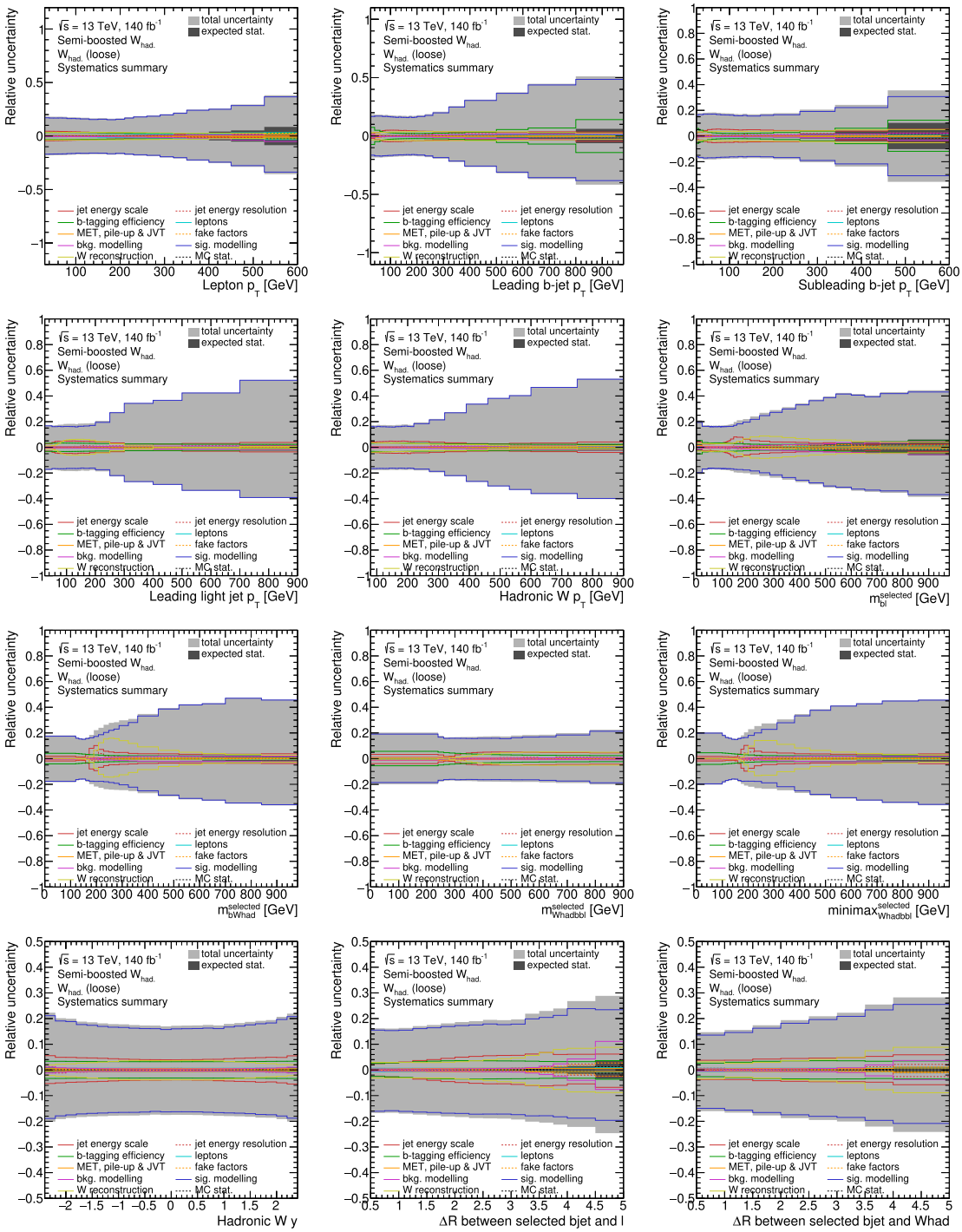


Figure C.3: Systematic Uncertainties Summary for the $W_{\text{had.}}$ (loose) Region. Systematic uncertainties include contributions from both detector-level and modeling uncertainties. The total systematic uncertainty is computed by adding the statistical and systematic uncertainties in quadrature. Refer to Section 7.3 for a description of the observables of interest.

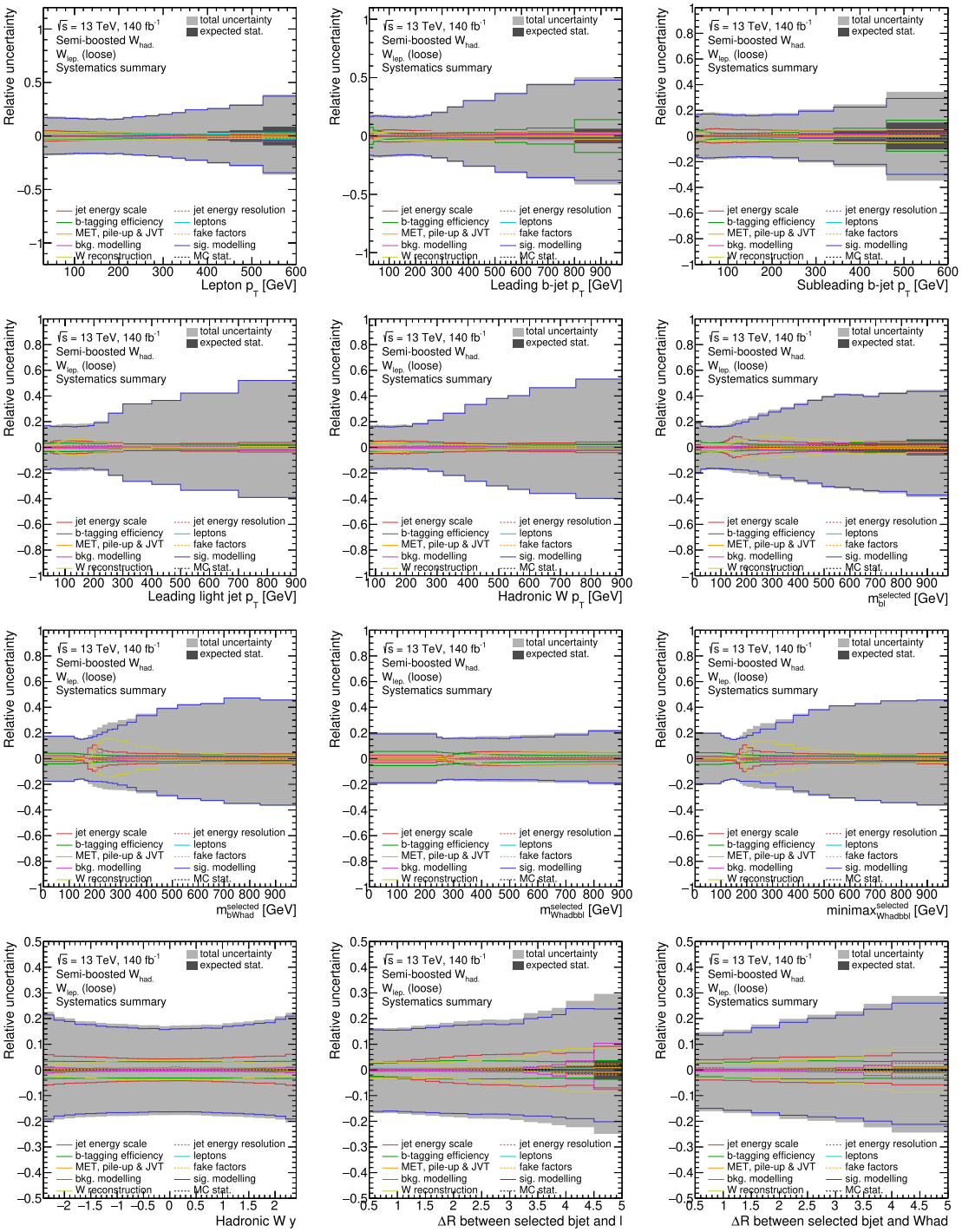
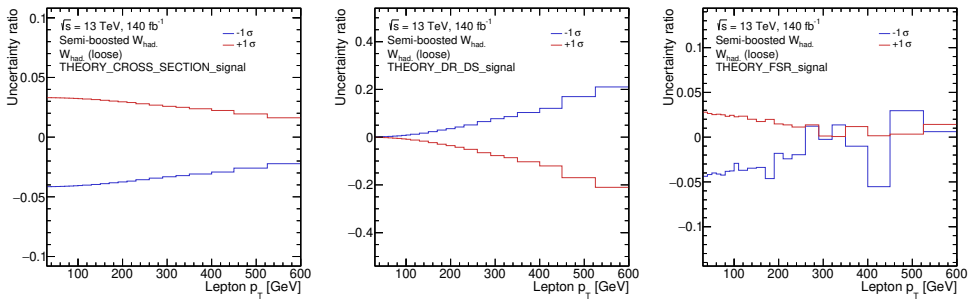


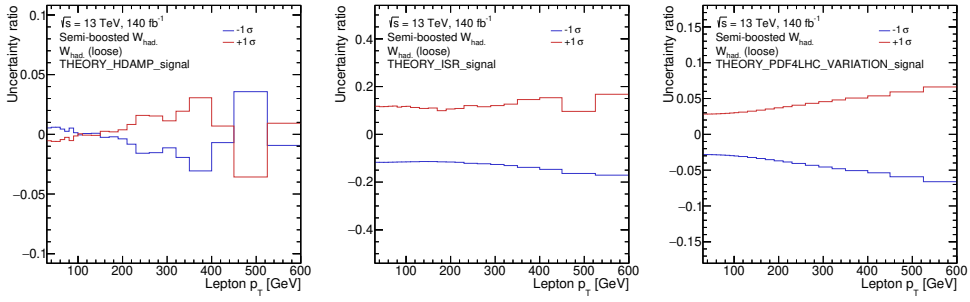
Figure C.4: Systematic Uncertainties Summary for the $W_{\text{lep.}}$ (loose) Region. Systematic uncertainties include contributions from both detector-level and modeling uncertainties. The total systematic uncertainty is computed by adding the statistical and systematic uncertainties in quadrature. Refer to Section 7.3 for a description of the observables of interest.

C.3 Signal Modeling Systematics Plots

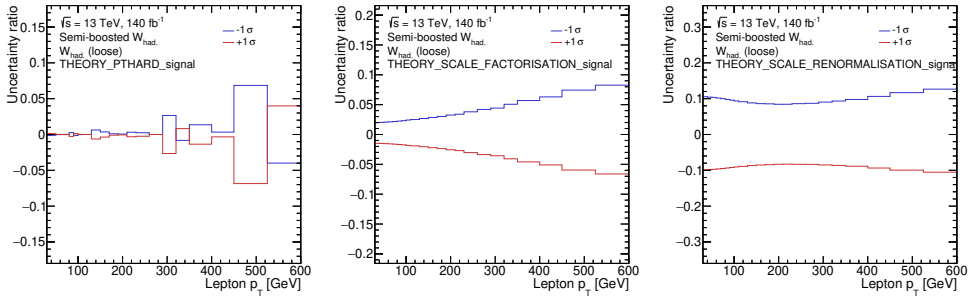
Figures C.5 and C.6 plot the NPs associated with signal modeling systematics for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively. In these plots, the following NPs are plotted: cross-section, DR/DS, FSR, h_{damp}, ISR, PDF variations, $p_{\text{T,hard}}$, μ_{f}, μ_{r}, parton-showering, and top recoil.



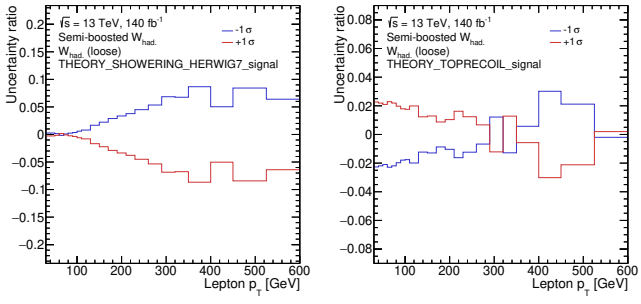
(a) Cross-section. **(b)** DR/DS. **(c)** FSR.



(d) h_{damp}. **(e)** ISR. **(f)** PDF.

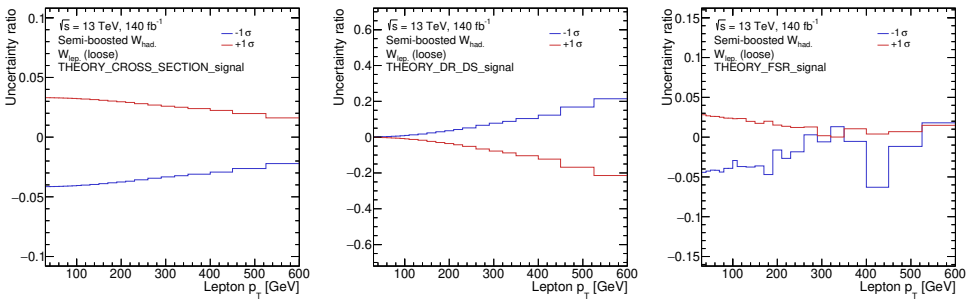


(g) $p_{\text{T,hard}}$. **(h)** μ_{f}. **(i)** μ_{r}.



(j) Showering. **(k)** Top recoil.

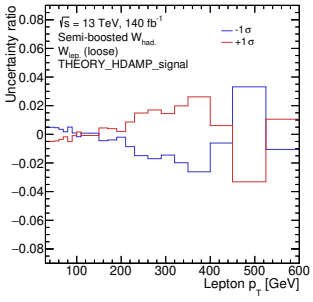
Figure C.5: Summary of Signal Modeling Uncertainties for the $W_{\text{had.}}$ (loose) Region. Signal modeling uncertainties make up the largest contribution to the total systematic uncertainty. The contributions of the NPs associated with the signal modeling uncertainties are computed for each bin as a ratio of the up/down variations to the nominal distribution.



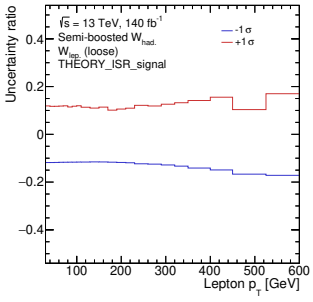
(a) Cross-section.

(b) DR/DS.

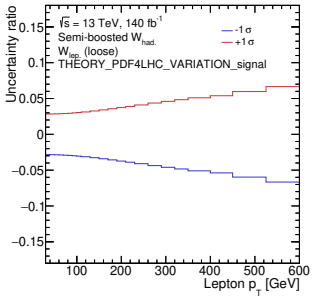
(c) FSR.



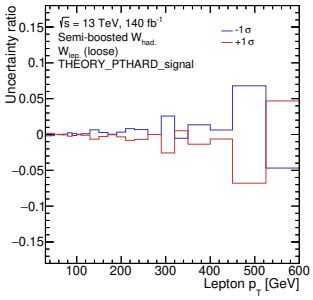
(d) h_{damp}.



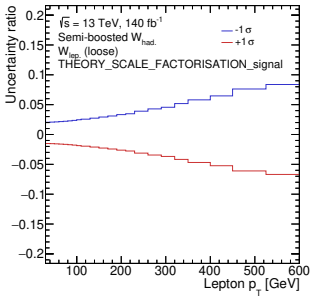
(e) ISR.



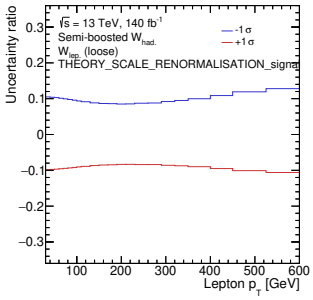
(f) PDF.



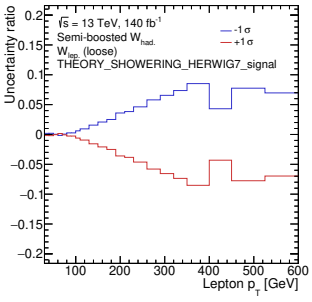
(g) $p_{\text{T,hard}}$.



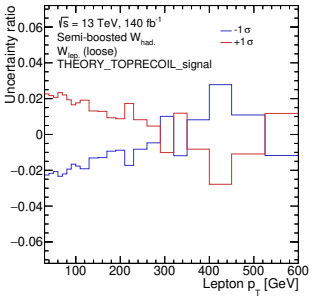
(h) μ_{f}.



(i) μ_{r}.



(j) Showering.



(k) Top recoil.

Figure C.6: Summary of Signal Modeling Uncertainties for the $W_{\text{lep.}}$ (loose) Region. Signal modeling uncertainties make up the largest contribution to the total systematic uncertainty. The contributions of the NPs associated with the signal modeling uncertainties are computed for each bin as a ratio of the up/down variations to the nominal distribution.

C.4 Migration Matrices

Figures C.7 and C.8 plot the migration matrices for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

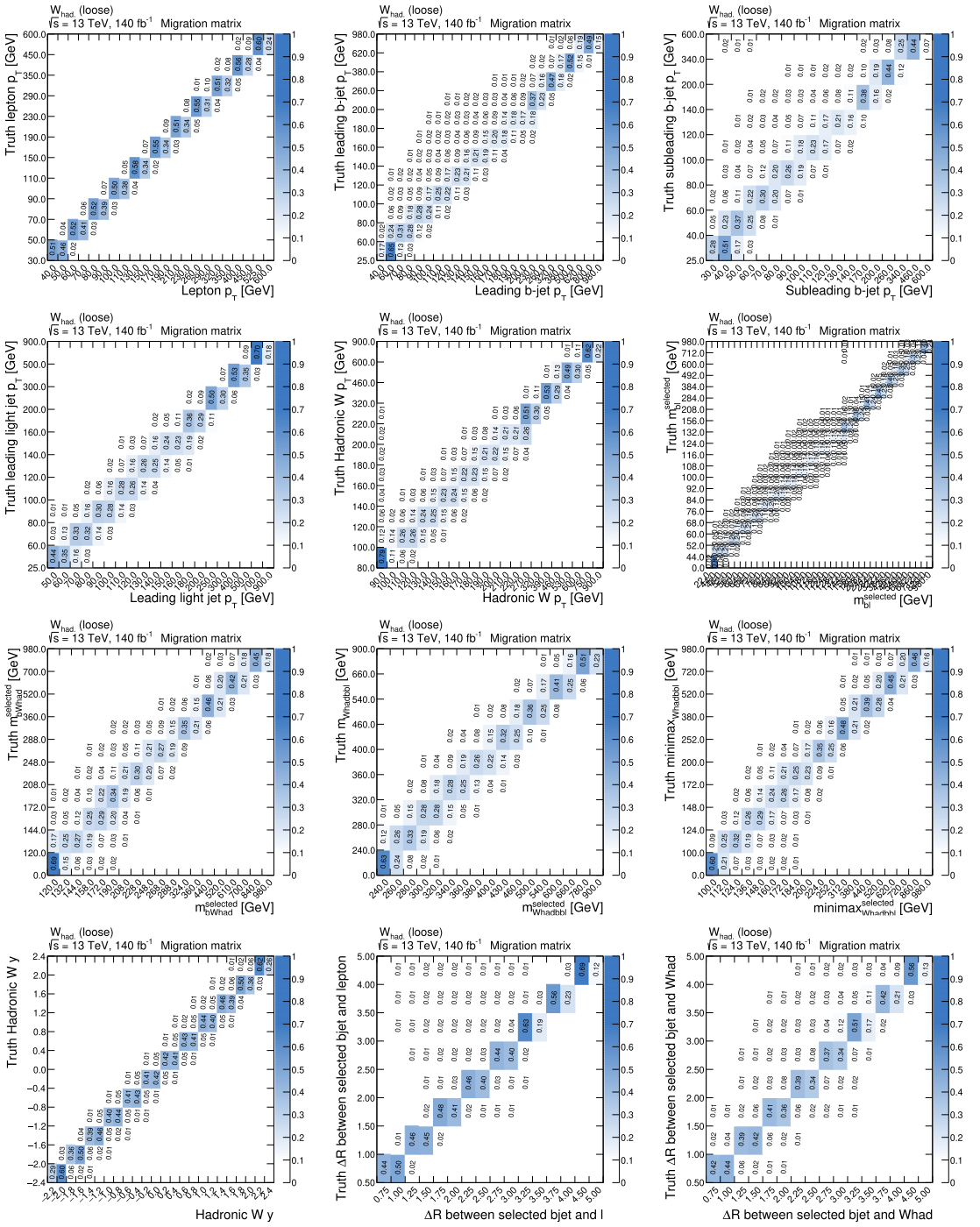


Figure C.7: Migration Matrices for the $W_{\text{had.}}$ (loose) Region. Each bin in the migration matrix represents the fraction of events in the reconstructed bin that originate from the true bin. The diagonal elements represent the fraction of correctly reconstructed events, while the off-diagonal elements represent the fraction of misreconstructed events. Migration matrices are plotted for the observables of interest. Refer to Section 7.3 for a description of the observables of interest.

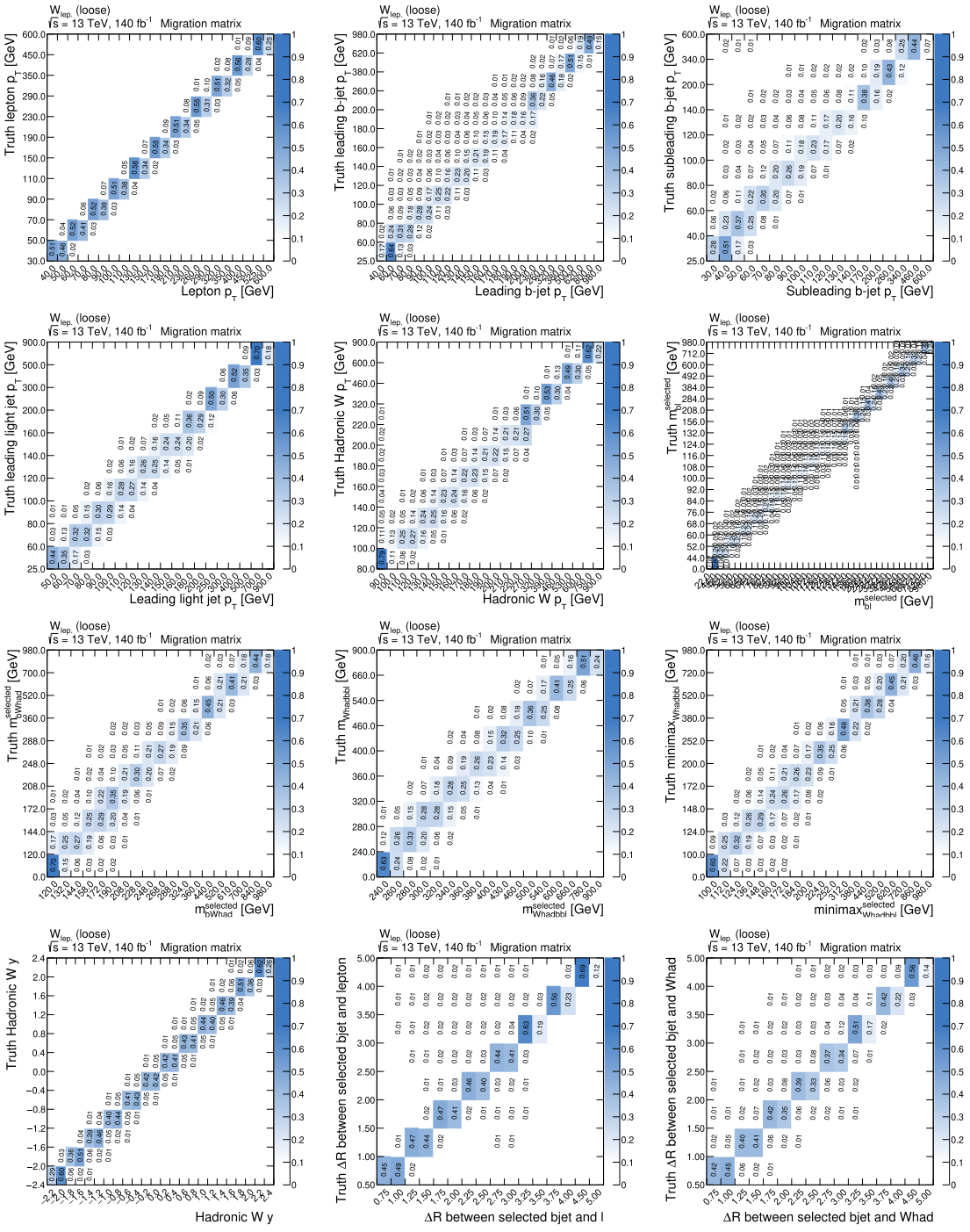


Figure C.8: Migration Matrices for the $W_{\text{lep.}}$ (loose) Region. Each bin in the migration matrix represents the fraction of events in the reconstructed bin that originate from the true bin. The diagonal elements represent the fraction of correctly reconstructed events, while the off-diagonal elements represent the fraction of misreconstructed events. Migration matrices are plotted for the observables of interest. Refer to Section 7.3 for a description of the observables of interest.

C.5 Acceptance/Efficiency Correction Factors

Figures C.9 and C.10 plot the acceptance/efficiency correction factors for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

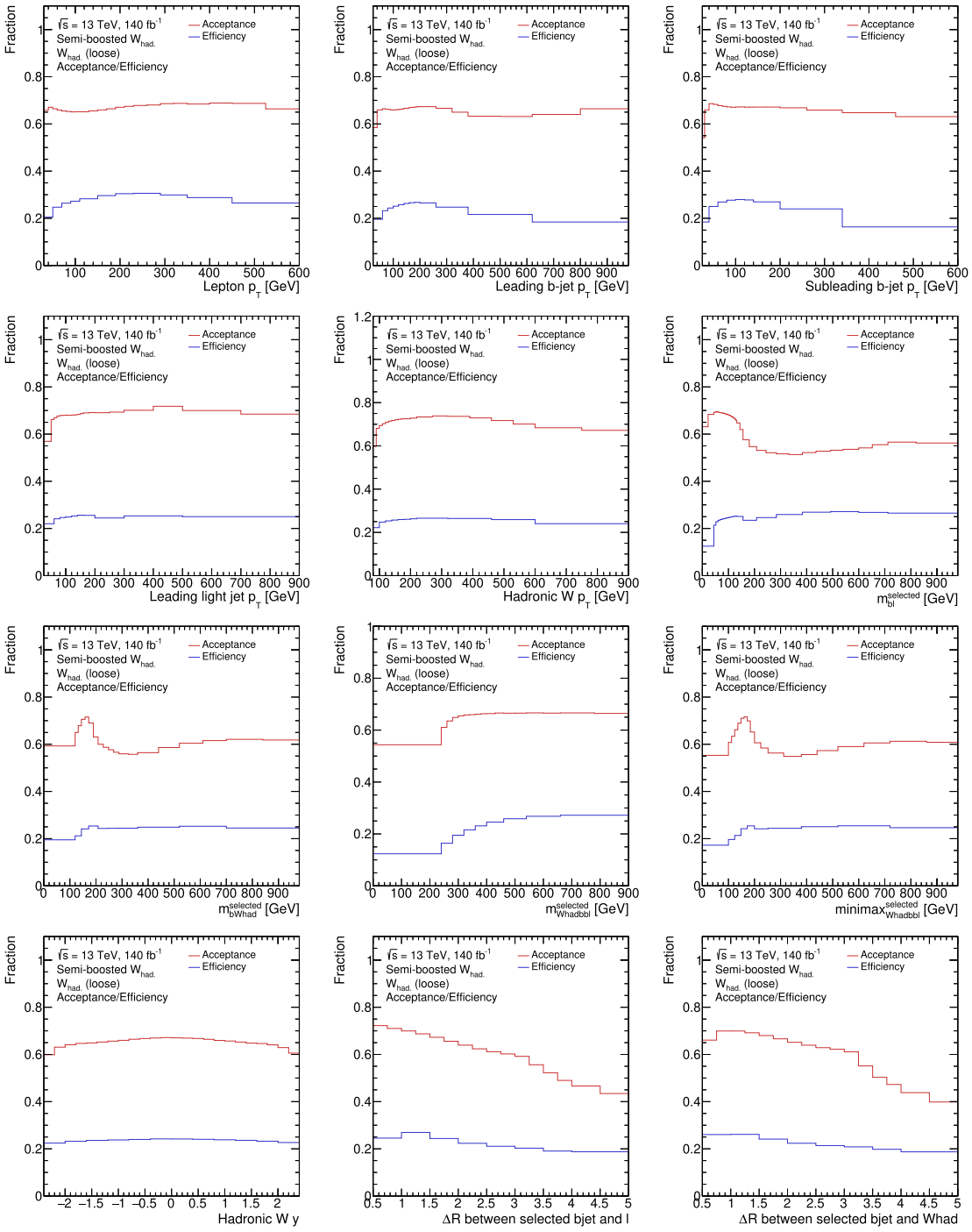


Figure C.9: Acceptance/Efficiency Factors for the $W_{\text{had.}}$ (loose) Region. The acceptance and efficiency factors correct the normalization of the unfolded distribution. For each bin, the acceptance correction factor accounts for events that pass the detector-level selection but fail the particle-level selection. The efficiency correction factor accounts for events that pass the particle-level selection but are not reconstructed at the detector-level. Refer to Section 7.3 for a description of the observables of interest.

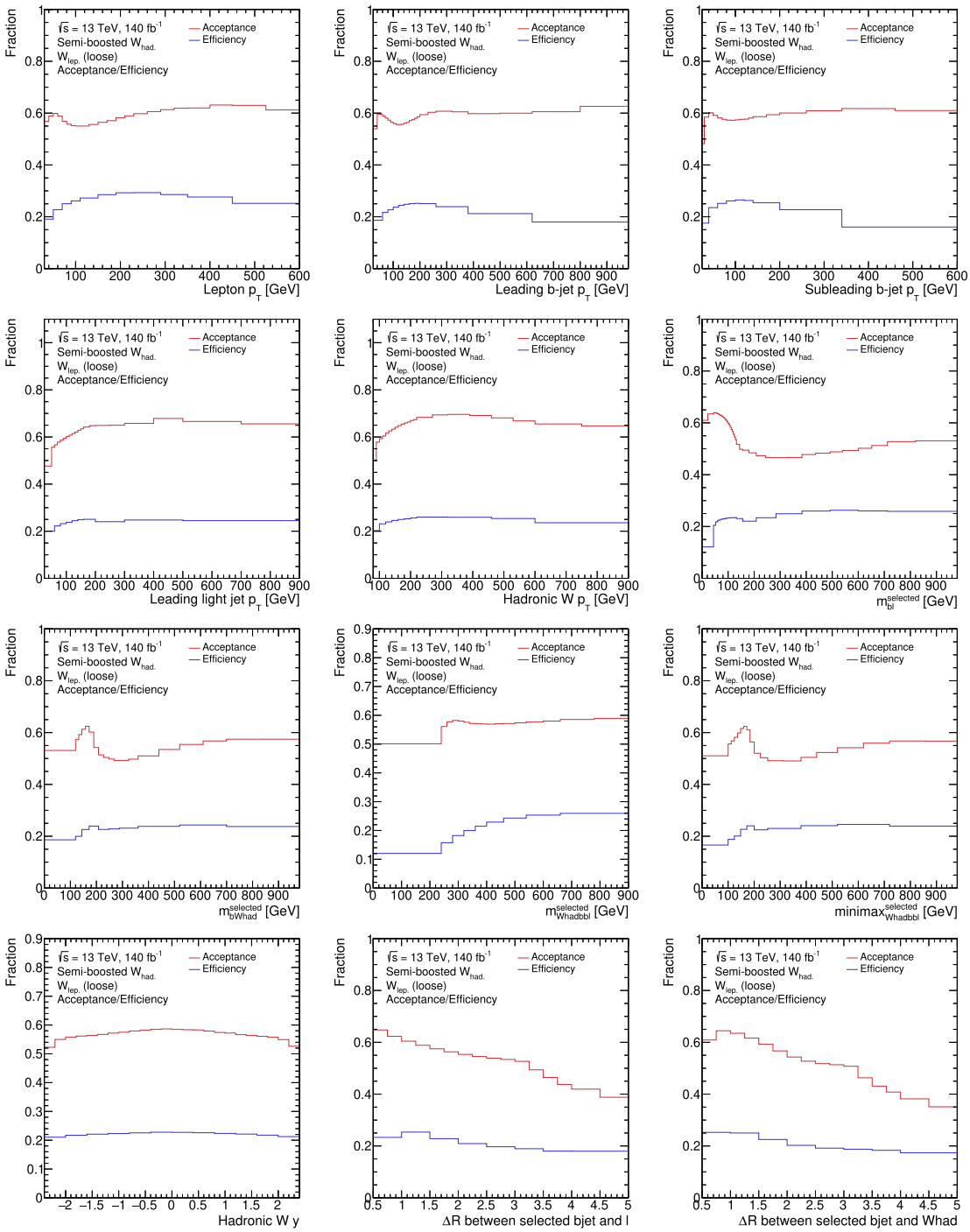


Figure C.10: Acceptance/Efficiency Factors for the $W_{\text{lep.}}$ (loose) Region. The acceptance and efficiency factors correct the normalization of the unfolded distribution. For each bin, the acceptance correction factor accounts for events that pass the detector-level selection but fail the particle-level selection. The efficiency correction factor accounts for events that pass the particle-level selection but are not reconstructed at the detector-level. Refer to Section 7.3 for a description of the observables of interest.

C.6 Covariance Matrices

Figures C.11 and C.12 plot the covariance matrices for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

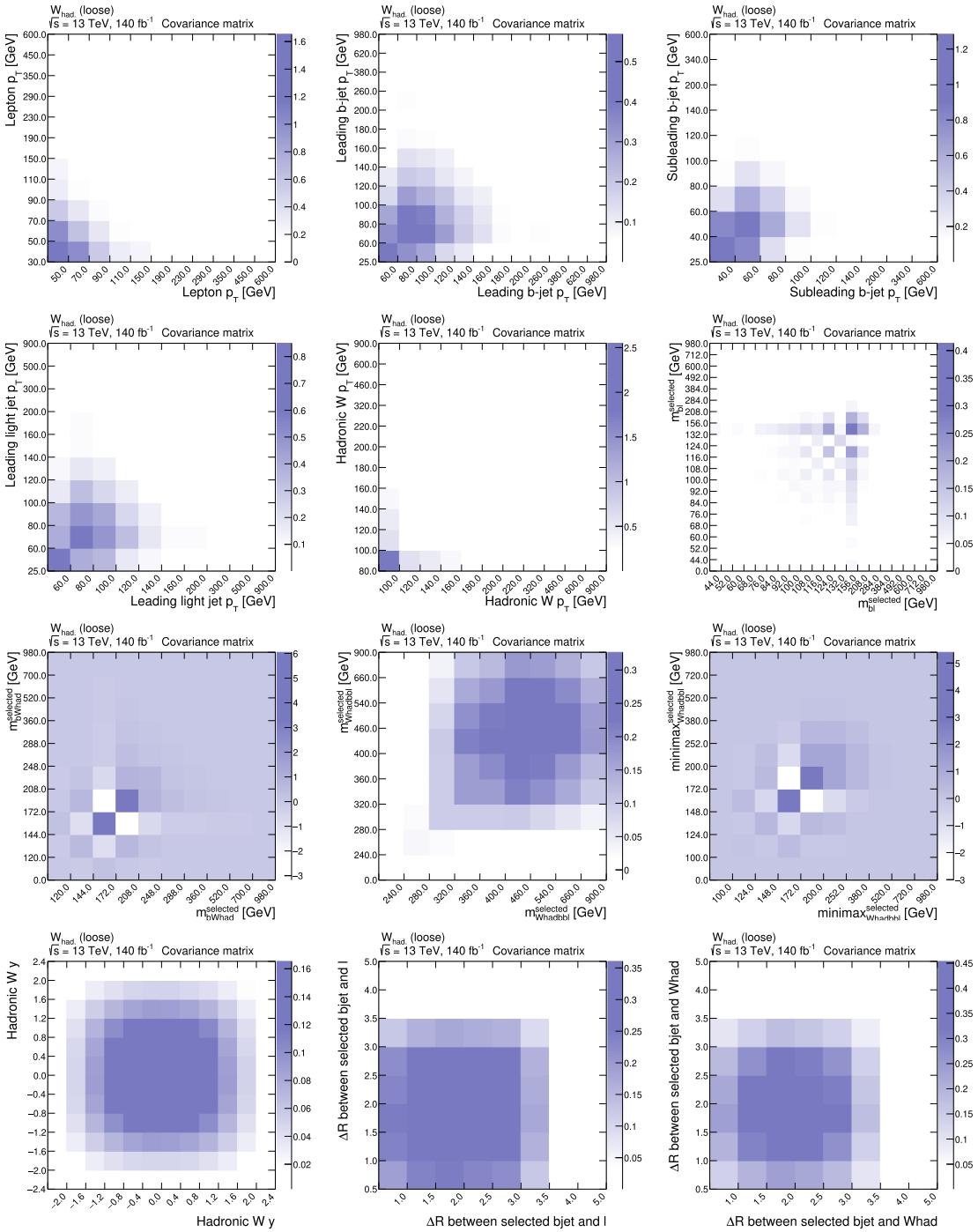


Figure C.11: Covariance Matrices for the $W_{\text{had.}}$ (loose) Region. The covariance matrix measures the covariance between bins of the unfolded distribution. The elements on the covariance matrix's main diagonal indicate the bins' total squared uncertainties in the unfolded distribution. Refer to Section 7.3 for a description of the observables of interest.

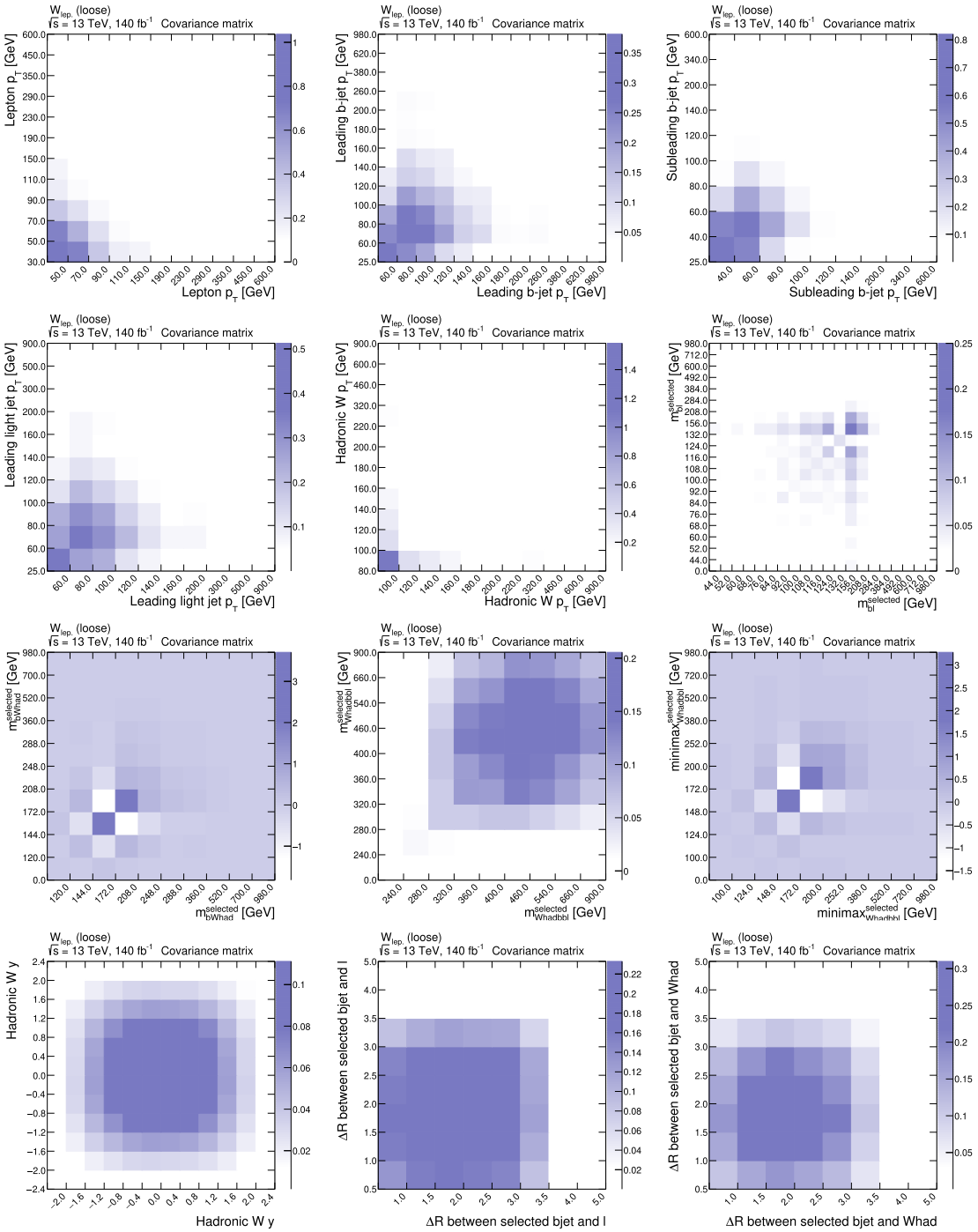


Figure C.12: Covariance Matrices for the $W_{\text{lep.}}$ (loose) Region. The covariance matrix measures the covariance between bins of the unfolded distribution. The elements on the covariance matrix's main diagonal indicate the bins' total squared uncertainties in the unfolded distribution. Refer to Section 7.3 for a description of the observables of interest.

C.7 Correlation Matrices

Figures C.13 and C.14 plot the covariance matrices for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

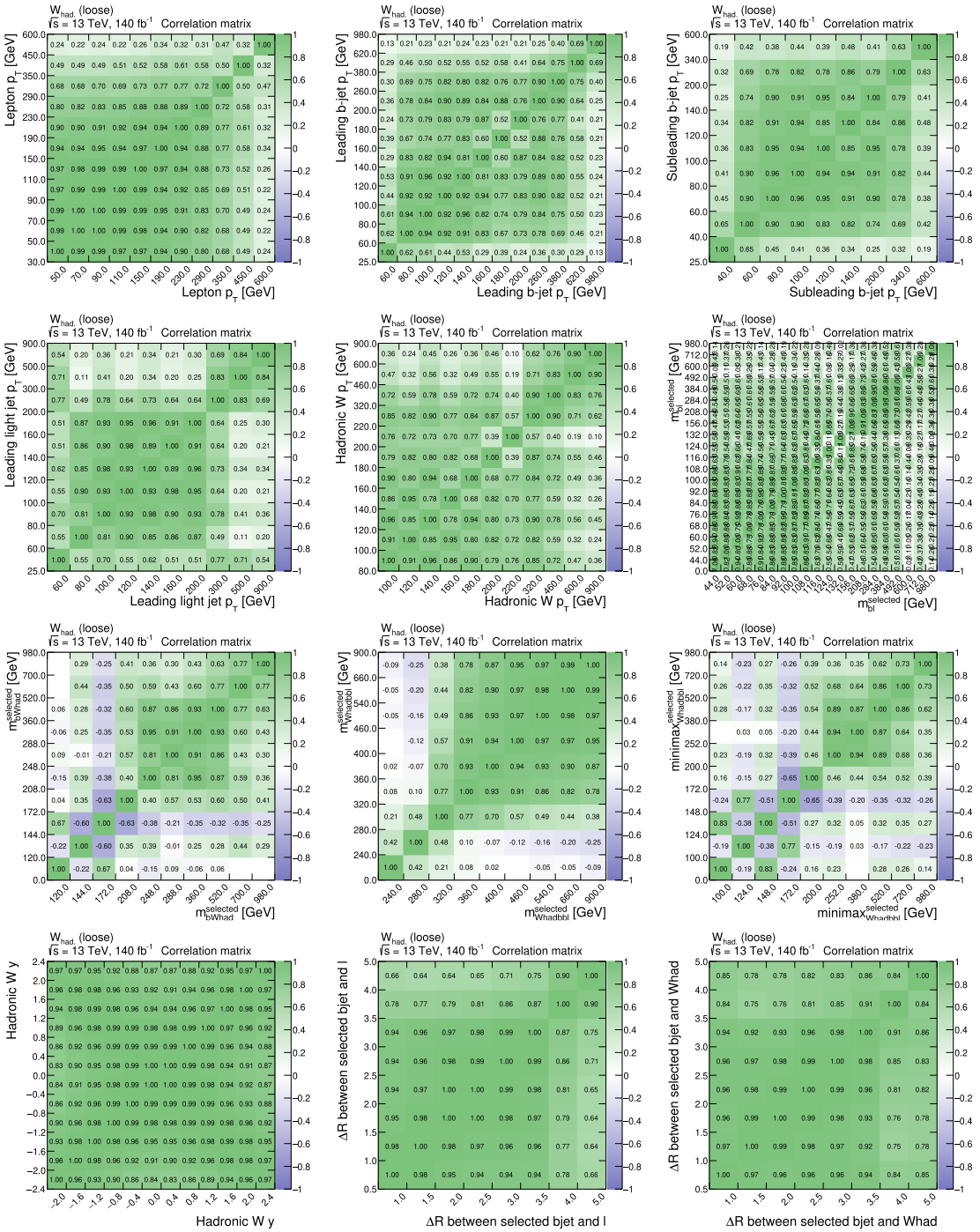


Figure C.13: Correlation Matrices for the $W_{\text{had.}}$ (loose) Region. The correlation measures the total correlation between bins of the unfolded distribution. A large positive number equates to positive correlations, while a large negative number equates to negative correlations. Large correlations (positive or negative) in distant bins suggest some common factor or influence affecting both bins. Refer to Section 7.3 for a description of the observables of interest.

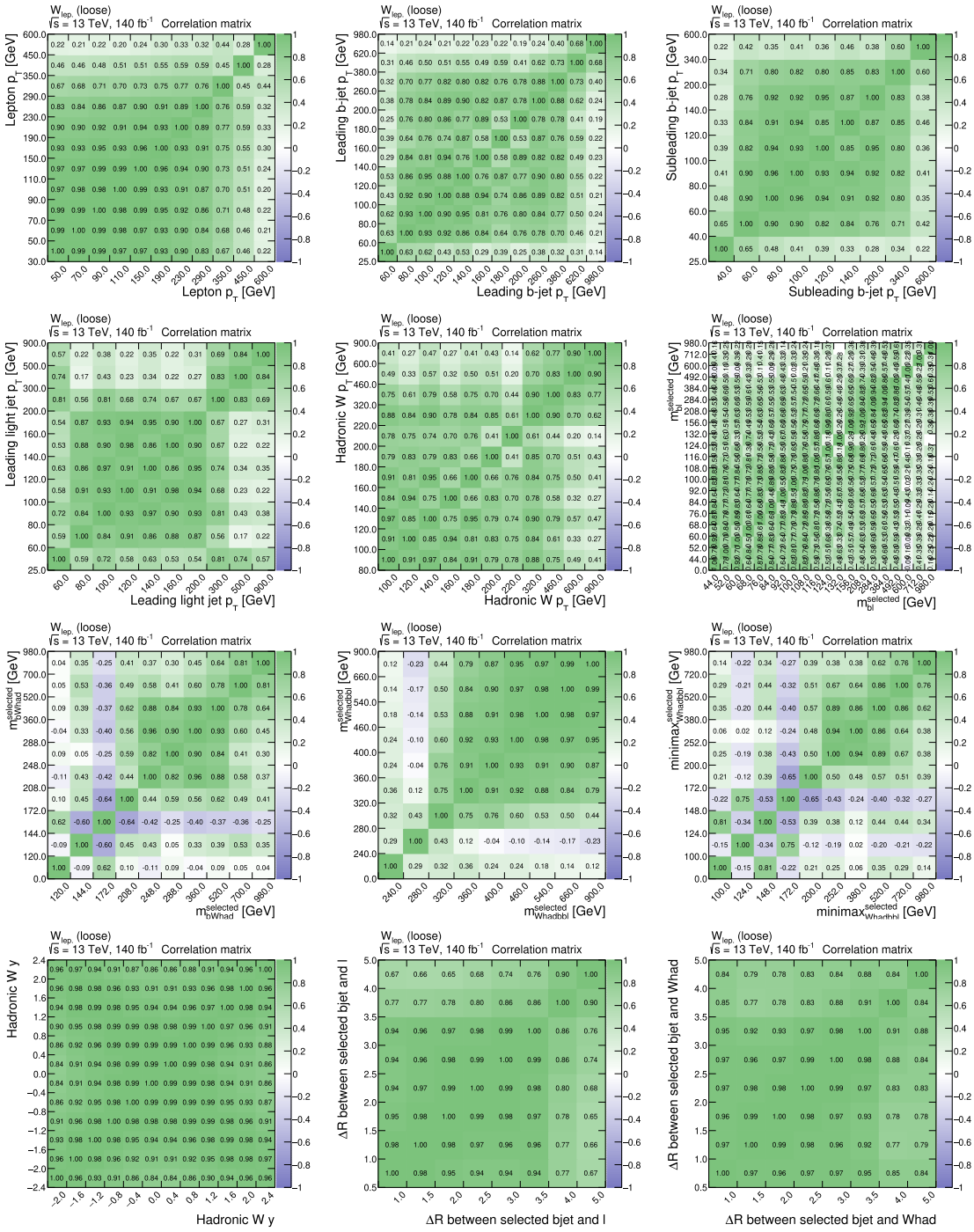


Figure C.14: Correlation Matrices for the $W_{\text{lep.}}$ (loose) Region. The correlation measures the total correlation between bins of the unfolded distribution. A large positive number equates to positive correlations, while a large negative number equates to negative correlations. Large correlations (positive or negative) in distant bins suggest some common factor or influence affecting both bins. Refer to Section 7.3 for a description of the observables of interest.

C.8 Absolute Differential Cross Sections

Figures C.15 and C.16 plot the absolute differential cross-sections (using the nominal MC samples) for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

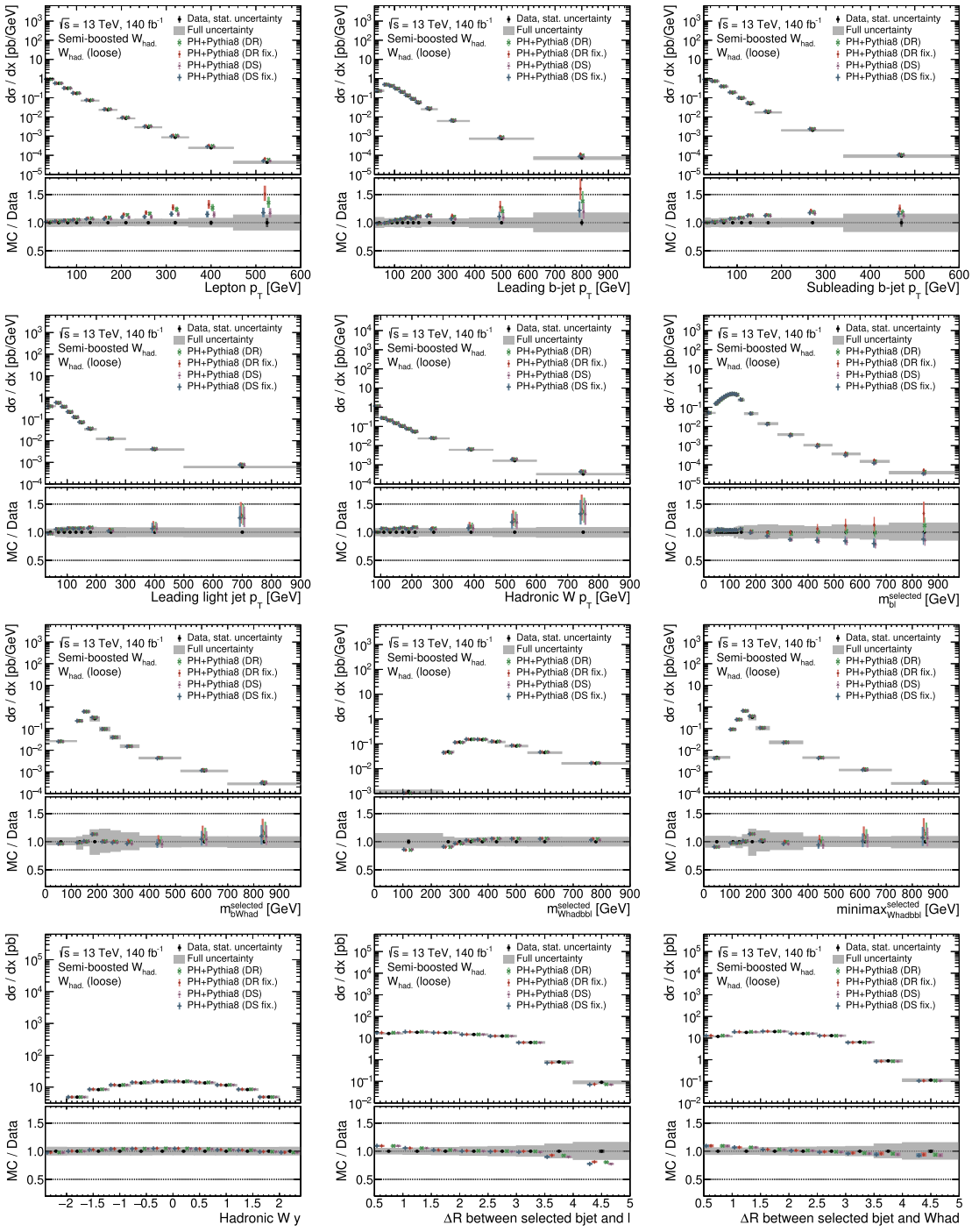


Figure C.15: Absolute Differential Cross-Sections (Nominal MC) for the $W_{\text{had.}}$ (loose) Region. The absolute differential cross-section is calculated using each bin's unfolded number of events. The measured cross-sections are then compared individually to the set of nominal MC generators. Refer to Section 7.3 for a description of the observables of interest.

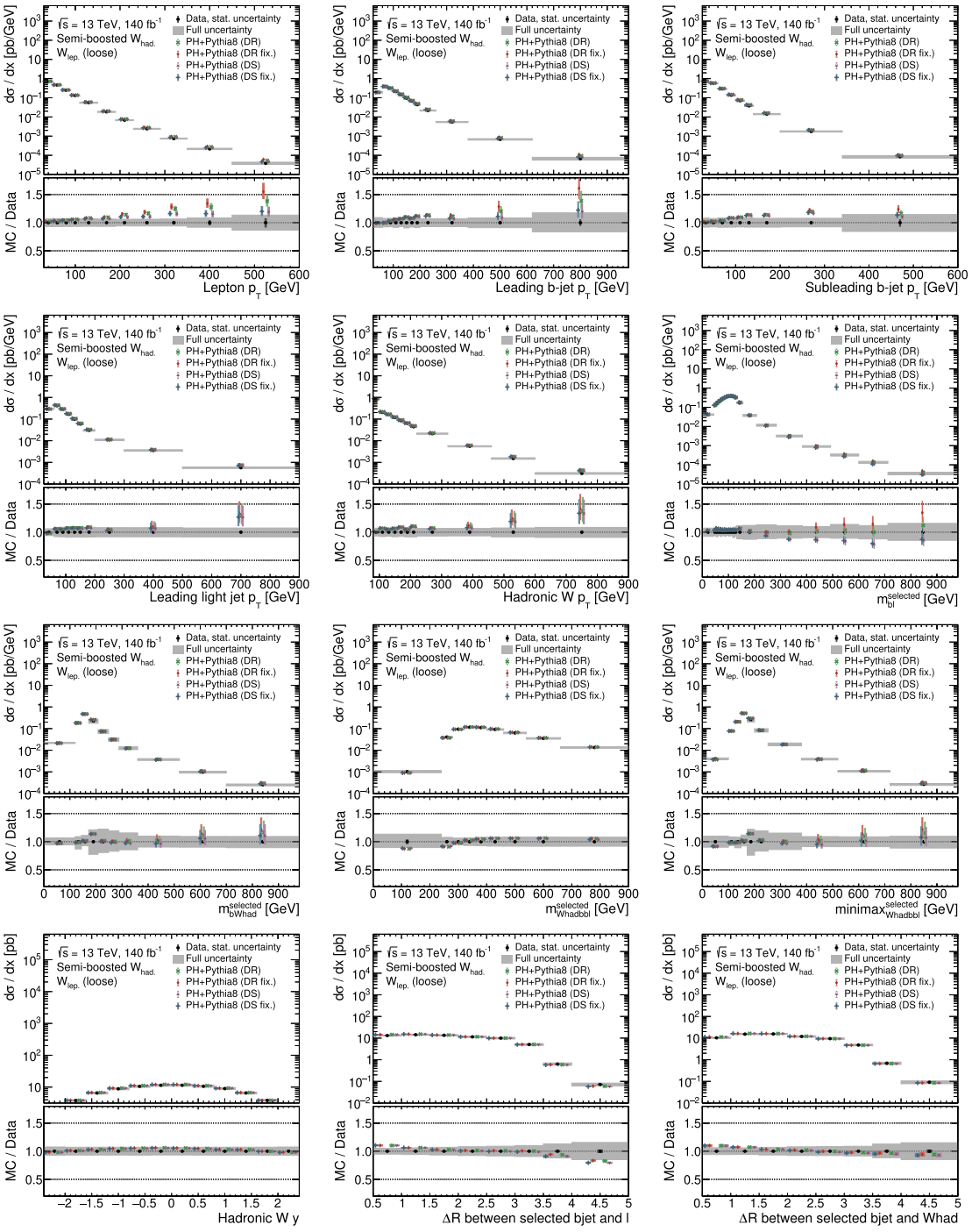


Figure C.16: Absolute Differential Cross-Sections (Nominal MC) for the $W_{\text{lep.}}$ (loose) Region. The absolute differential cross-section is calculated using each bin's unfolded number of events. The measured cross-sections are then compared individually to the set of nominal MC generators. Refer to Section 7.3 for a description of the observables of interest.

Figures C.17 and C.18 plot the absolute differential cross-sections (using the alternative MC samples) for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

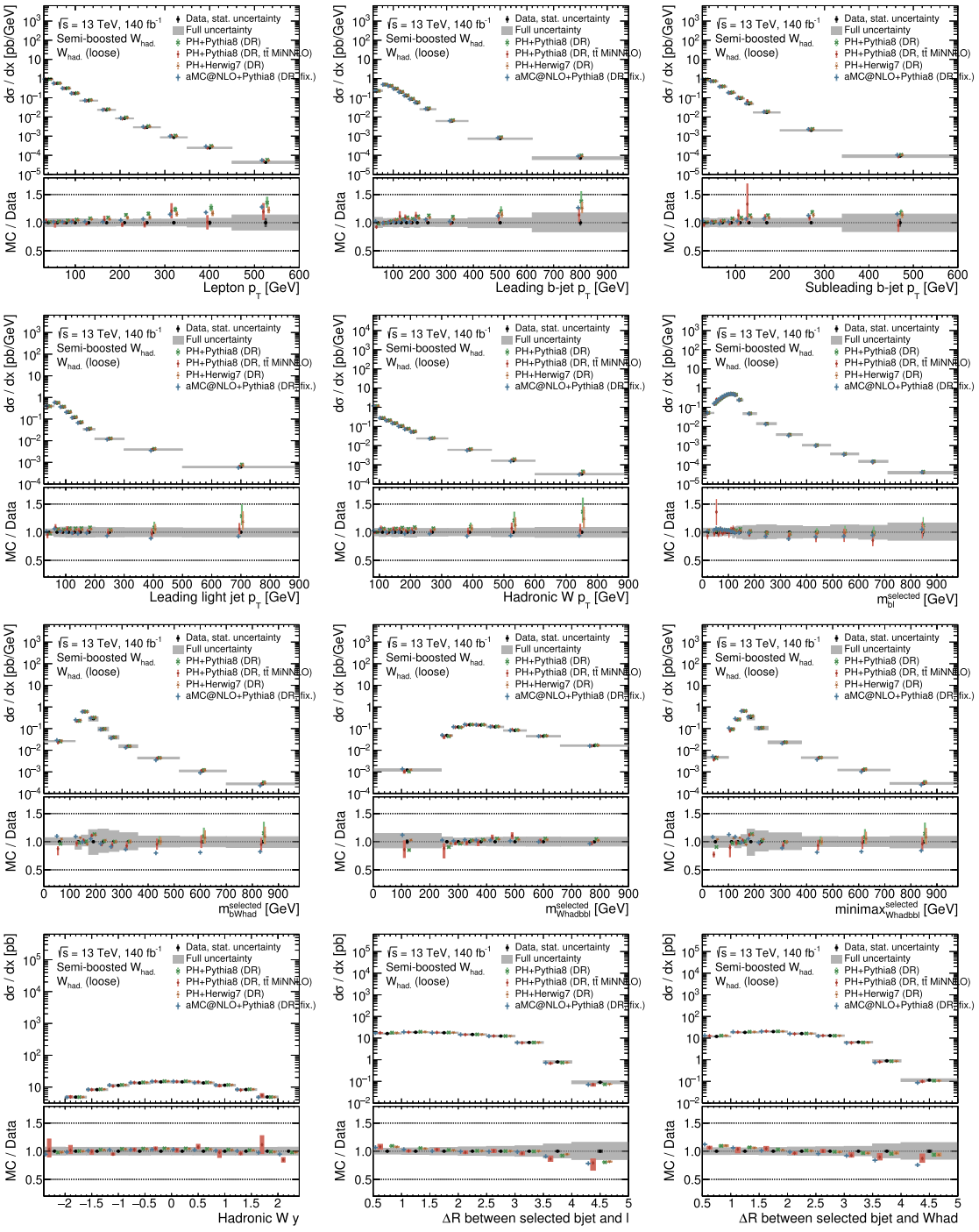


Figure C.17: Absolute Differential Cross-Sections (Alternative MC) for the $W_{\text{had.}}$ (loose) Region. The absolute differential cross-section is calculated using each bin's unfolded number of events. The measured cross-sections are then compared individually to the set of alternative MC generators. Refer to Section 7.3 for a description of the observables of interest.

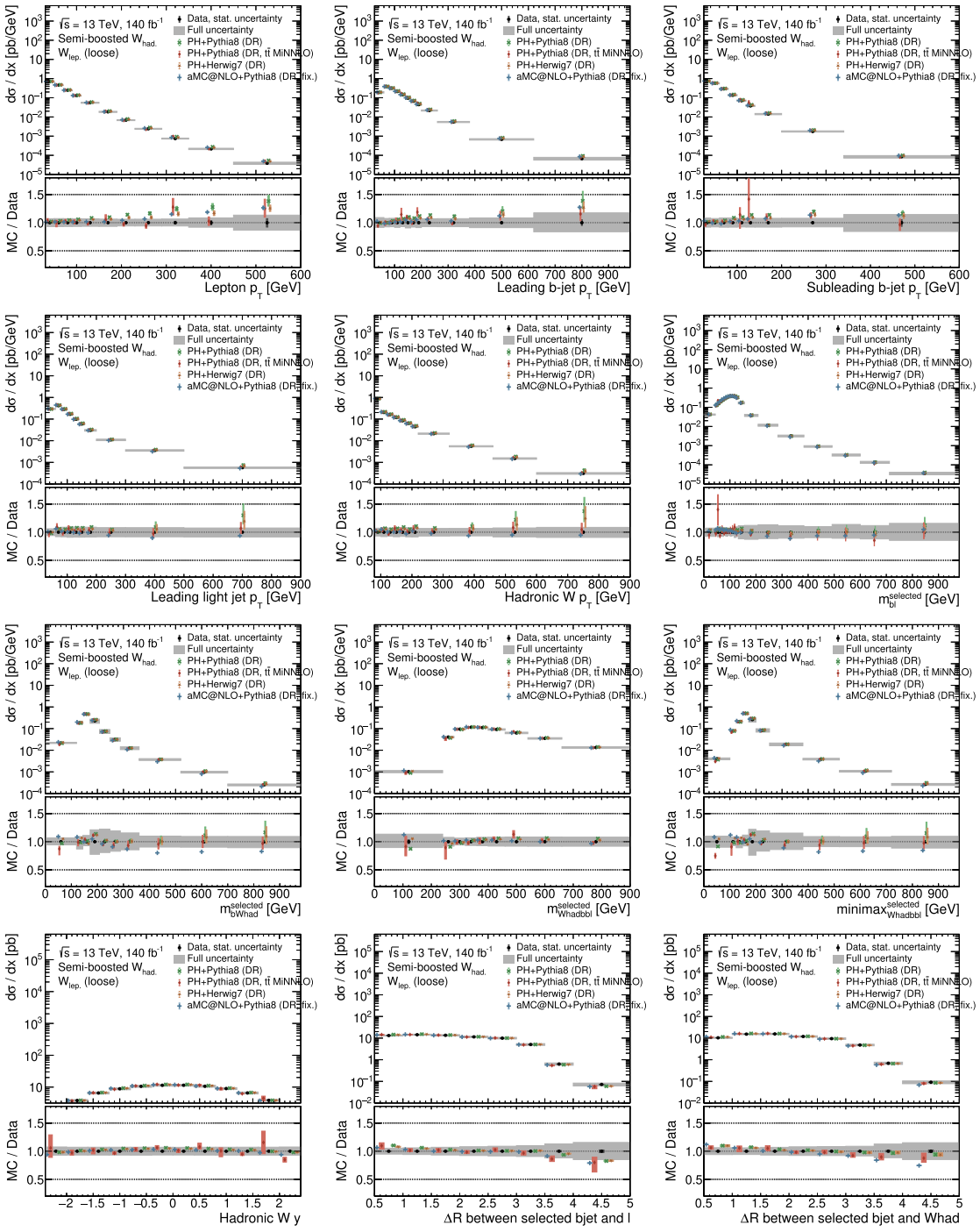


Figure C.18: Absolute Differential Cross-Sections (Alternative MC) for the $W_{\text{lep.}}$ (loose) Region. The absolute differential cross-section is calculated using each bin's unfolded number of events. The measured cross-sections are then compared individually to the set of alternative MC generators. Refer to Section 7.3 for a description of the observables of interest.

C.9 Normalized Differential Cross Sections

Figures C.19 and C.20 plot the normalized differential cross-sections (using the nominal MC samples) for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

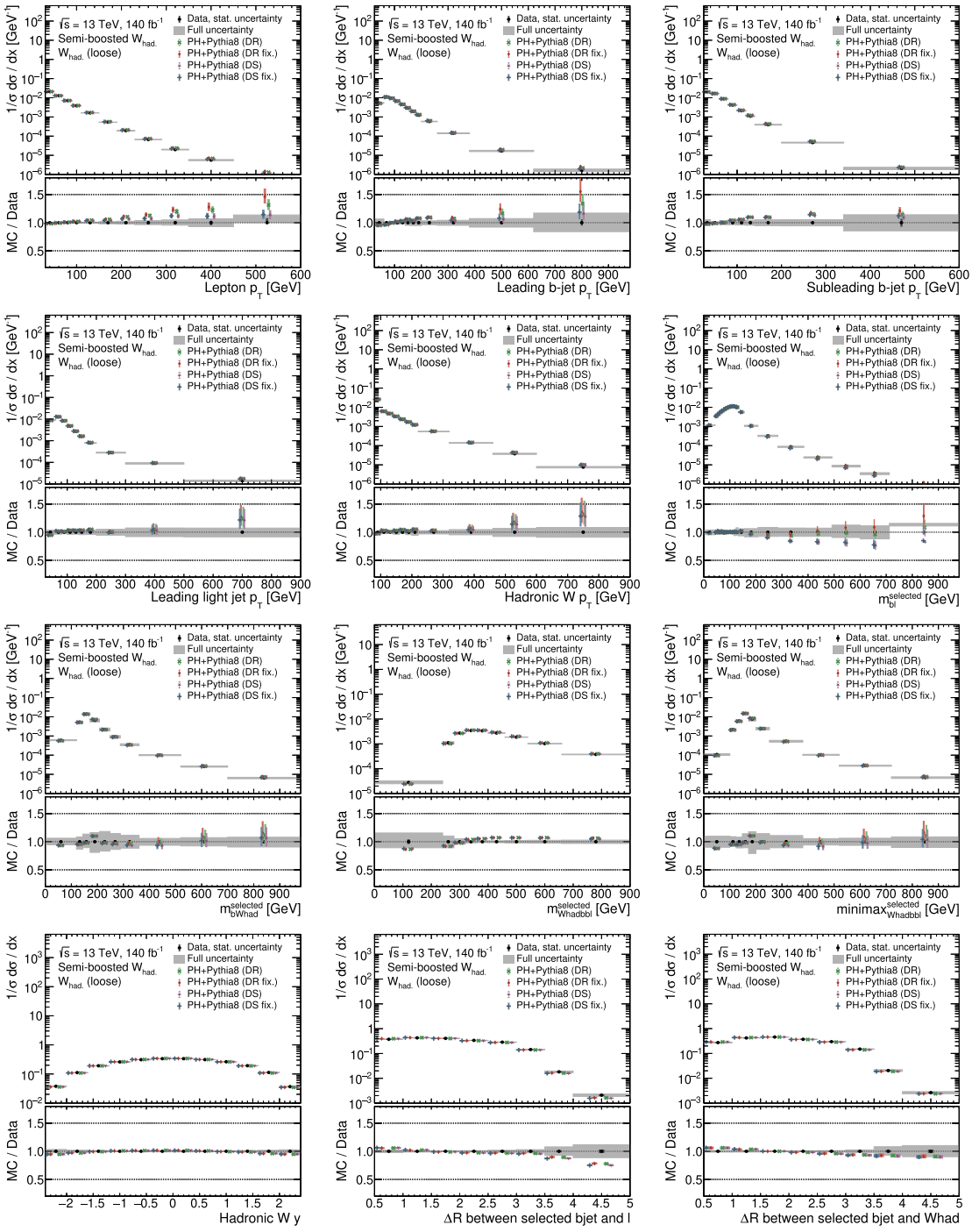


Figure C.19: Normalized Differential Cross-Sections (Nominal MC) for the $W_{\text{had.}}$ (loose) Region. The normalized differential cross-section is calculated by normalizing the absolute differential cross-section to unit area. The measured cross-sections are then compared individually to the set of nominal MC generators. Refer to Section 7.3 for a description of the observables of interest.

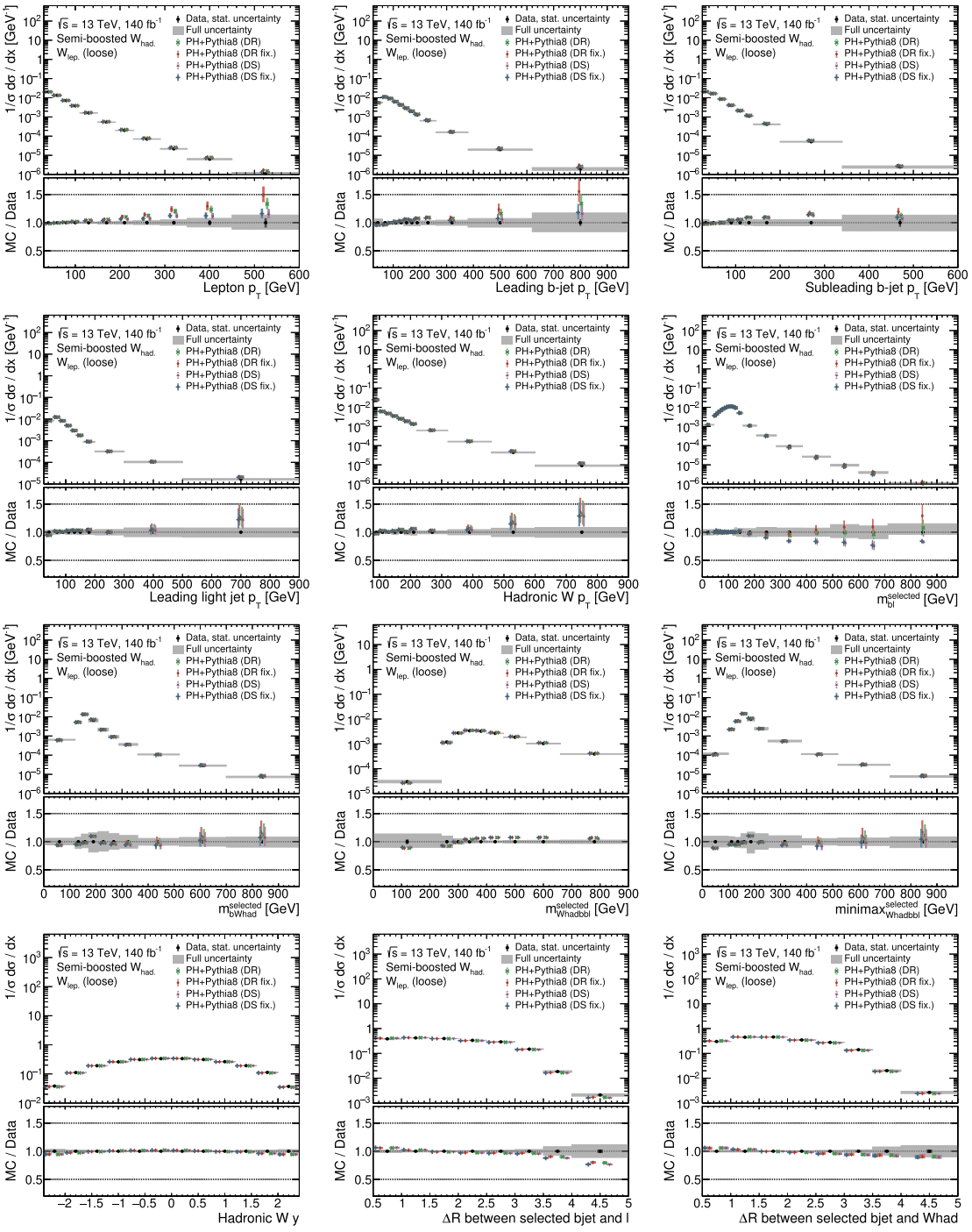


Figure C.20: Normalized Differential Cross-Sections (Nominal MC) for the $W_{\text{lep.}}$ (loose) Region. The normalized differential cross-section is calculated by normalizing the absolute differential cross-section to unit area. The measured cross-sections are then compared individually to the set of nominal MC generators. Refer to Section 7.3 for a description of the observables of interest.

Figures C.21 and C.22 plot the normalized differential cross-sections (using the alternative MC samples) for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

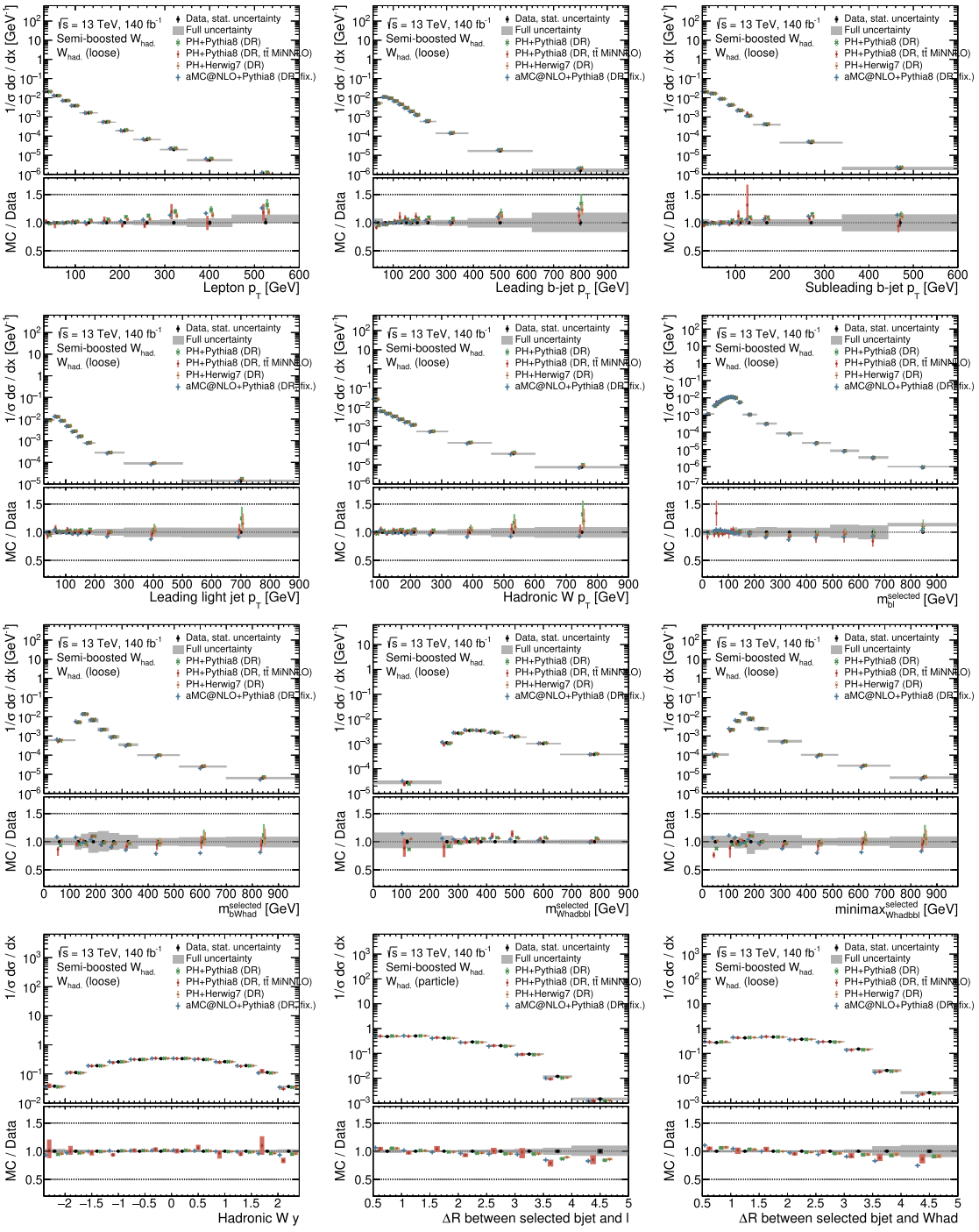


Figure C.21: Normalized Differential Cross-Sections (Alternative MC) for the $W_{\text{had.}}$ (loose) Region. The normalized differential cross-section is calculated by normalizing the absolute differential cross-section to unit area. The measured cross-sections are then compared individually to the set of alternative MC generators. Refer to Section 7.3 for a description of the observables of interest.

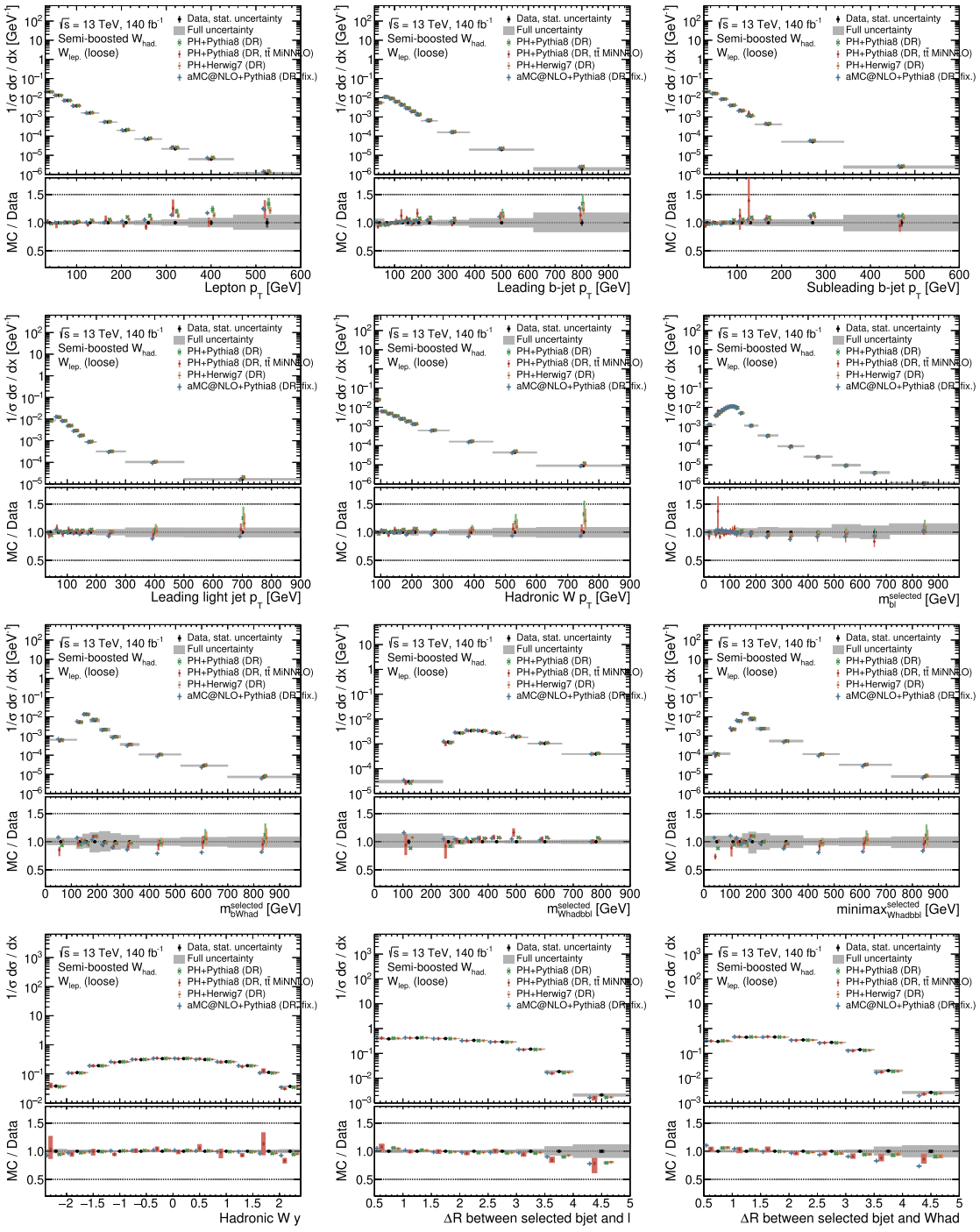


Figure C.22: Normalized Differential Cross-Sections (Alternative MC) for the $W_{lep.}$ (loose) Region. The normalized differential cross-section is calculated by normalizing the absolute differential cross-section to unit area. The measured cross-sections are then compared individually to the set of alternative MC generators. Refer to Section 7.3 for a description of the observables of interest.

C.10 Chi-Squared Tables

Tables C.1 and C.2 list the χ^2 and p-values for the $W_{\text{had.}}$ (loose) and $W_{\text{lep.}}$ (loose) regions, respectively.

Table C.1: χ^2 and p-values for the $W_{\text{had.}}$ (loose) Region. The χ^2 and p-values are calculated by comparing the measured cross-sections to the MC predictions. The *chi*^{2} values are then normalized to the number of degrees of freedom. Using the χ^2 value and NDF, the p-value is calculated. The p-value is the probability of observing a χ^2/NDF value as extreme as the one observed by chance alone. The χ^2/NDF and p-values are computed for the observables of interest in the $W_{\text{had.}}$ (loose) region.

	Generator													
	PH+PYTHIA 8 (DR)		PH+PYTHIA 8 (DR, fix.)		PH+PYTHIA 8 (DS)		PH+PYTHIA 8 (DS, fix.)		PH+PYTHIA 8 (DR, $t\bar{t}$ MINNLO)		PH+HERWIG 7.1 (DR)		aMCNLO+PYTHIA 8 (DR, fix.)	
Observable	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
Lepton p_{T}	31.83 / 11	<0.001	41.34 / 11	<0.001	15.36 / 11	0.166	16.05 / 11	0.139	76.58 / 11	<0.001	15.23 / 11	0.172	23.01 / 9	0.006
Leading b-jet p_{T}	21.30 / 12	0.046	28.86 / 12	0.004	18.21 / 12	0.109	18.28 / 12	0.107	75.28 / 12	<0.001	12.09 / 12	0.438	11.45 / 12	0.491
Subleading b-jet p_{T}	23.01 / 9	0.006	24.80 / 9	0.003	22.18 / 9	0.008	23.14 / 9	0.006	389.9 / 9	<0.001	11.09 / 9	0.270	20.13 / 9	0.017
Leading light-jet p_{T}	35.65 / 10	<0.001	42.26 / 10	<0.001	33.11 / 10	<0.001	33.96 / 10	<0.001	37.58 / 10	<0.001	13.12 / 10	0.217	21.82 / 10	0.016
Hadronic-W p_{T}	53.37 / 11	<0.001	62.23 / 11	<0.001	48.70 / 11	<0.001	49.64 / 11	<0.001	18.01 / 11	0.081	18.49 / 11	0.071	9.02 / 11	0.620
Selected $(b+\ell)$ mass	9.82 / 20	0.971	16.61 / 20	0.678	18.41 / 20	0.561	17.72 / 20	0.606	673.6 / 20	<0.001	7.91 / 20	0.992	33.50 / 20	0.030
Selected $(b+W_{\text{had.}})$ mass	18.21 / 10	0.052	18.89 / 10	0.042	26.53 / 10	0.003	26.31 / 10	0.003	63.96 / 10	<0.001	10.43 / 10	0.403	59.23 / 10	<0.001
Selected $(W_{\text{had.}}+b_1+b_2+\ell)$ mass	5.81 / 9	0.758	5.07 / 9	0.827	9.22 / 9	0.416	8.69 / 9	0.466	192.4 / 9	<0.001	4.33 / 9	0.888	15.35 / 9	0.082
Modified minimax mass	12.44 / 10	0.256	13.77 / 10	0.184	21.56 / 10	0.018	21.09 / 10	0.020	59.13 / 10	<0.001	12.84 / 10	0.233	40.68 / 10	<0.001
Hadronic-W y	14.55 / 12	0.267	14.50 / 12	0.270	14.47 / 12	0.272	14.47 / 12	0.272	920.9 / 12	<0.001	14.46 / 12	0.272	16.38 / 12	0.174
$\Delta R(b,\ell)$	38.01 / 8	<0.001	38.46 / 8	<0.001	40.07 / 8	<0.001	41.13 / 8	<0.001	232.8 / 8	<0.001	25.39 / 8	0.001	30.01 / 8	<0.001
$\Delta R(b,W_{\text{had.}})$	52.85 / 8	<0.001	53.56 / 8	<0.001	55.56 / 8	<0.001	56.81 / 8	<0.001	107.3 / 8	<0.001	58.48 / 8	<0.001	118.7 / 8	<0.001

Table C.2: χ^2 and p-values for the $W_{\text{lep.}}$ (loose) Region. The χ^2 and p-values are calculated by comparing the measured cross-sections to the MC predictions. The *chi*^{2} values are then normalized to the number of degrees of freedom. Using the χ^2 value and NDF, the p-value is calculated. The p-value is the probability of observing a χ^2/NDF value as extreme as the one observed by chance alone. The χ^2/NDF and p-values are computed for the observables of interest in the $W_{\text{lep.}}$ (loose) region.

	Generator													
	PH+PYTHIA 8 (DR)		PH+PYTHIA 8 (DR, fix.)		PH+PYTHIA 8 (DS)		PH+PYTHIA 8 (DS, fix.)		PH+PYTHIA 8 (DR, $t\bar{t}$ MINNLO)		PH+HERWIG 7.1 (DR)		aMCNLO+PYTHIA 8 (DR, fix.)	
Observable	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
Lepton p_{T}	32.76 / 11	<0.001	44.34 / 11	<0.001	13.76 / 11	0.25	14.43 / 11	0.21	101.50 / 11	<0.001	15.50 / 11	0.16	13.16 / 11	0.28
Leading b-jet p_{T}	19.08 / 12	0.09	26.99 / 12	0.01	15.27 / 12	0.23	15.35 / 12	0.22	104.70 / 12	<0.001	11.01 / 12	0.53	10.17 / 12	0.60
Subleading b-jet p_{T}	22.25 / 9	0.01	24.22 / 9	0.00	20.62 / 9	0.01	21.53 / 9	0.01	369.70 / 9	<0.001	18.80 / 9	0.03	10.54 / 9	0.31
Leading light-jet p_{T}	35.36 / 10	<0.001	41.56 / 10	<0.001	32.81 / 10	<0.001	33.71 / 10	<0.001	21.77 / 10	0.02	19.79 / 10	0.03	12.94 / 10	0.23
Hadronic-W p_{T}	47.91 / 11	<0.001	55.84 / 11	<0.001	42.73 / 11	<0.001	43.64 / 11	<0.001	27.47 / 11	0.00	8.46 / 11	0.67	16.80 / 11	0.11
Selected $(b+\ell)$ mass	9.50 / 20	0.98	16.51 / 20	0.68	17.54 / 20	0.62	16.75 / 20	0.67	573.10 / 20	<0.001	24.93 / 20	0.20	7.69 / 20	0.99
Selected $(b+W_{\text{had.}})$ mass	18.54 / 10	0.05	19.07 / 10	0.04	25.84 / 10	0.00	25.59 / 10	0.00	76.78 / 10	<0.001	58.81 / 10	<0.001	10.70 / 10	0.38
Selected $(W_{\text{had.}}+b_1+b_2+\ell)$ mass	4.86 / 9	0.85	4.30 / 9	0.89	7.64 / 9	0.57	7.17 / 9	0.62	149.90 / 9	<0.001	14.82 / 9	0.10	4.23 / 9	0.90
Modified minimax mass	12.80 / 10	0.24	14.78 / 10	0.14	21.99 / 10	0.02	21.54 / 10	0.02	75.05 / 10	<0.001	44.37 / 10	<0.001	15.85 / 10	0.10
Hadronic-W y	16.52 / 12	0.17	16.46 / 12	0.17	16.38 / 12	0.18	16.37 / 12	0.18	649.30 / 12	<0.001	19.31 / 12	0.08	17.00 / 12	0.15
$\Delta R(b,\ell)$	41.89 / 8	<0.001	42.28 / 8	<0.001	44.51 / 8	<0.001	45.71 / 8	<0.001	211.60 / 8	<0.001	27.95 / 8	<0.001	25.49 / 8	0.00
$\Delta R(b,W_{\text{had.}})$	55.13 / 8	<0.001	55.94 / 8	<0.001	57.83 / 8	<0.001	59.25 / 8	<0.001	45.80 / 8	<0.001	128.00 / 8	<0.001	57.96 / 8	<0.001

APPENDIX D

Unfolding Closure Tests

The unfolding procedure is validated by comparing the unfolded distributions to the truth-level distributions. We perform the unfolding using a MC sample and comparing the unfolded distribution to the corresponding truth-level distribution using the same MC sample. Figure D.1 plots the closure tests for the observables of interest in the $W_{\text{had.}}$ (particle) region. The closure test is meant to test both the unfolding implementation and check that the unfolded distributions are consistent with the same truth-level distributions used to construct the migration matrix. All observables show perfect agreement between the unfolded and truth-level distributions, which are said to be successfully closed.

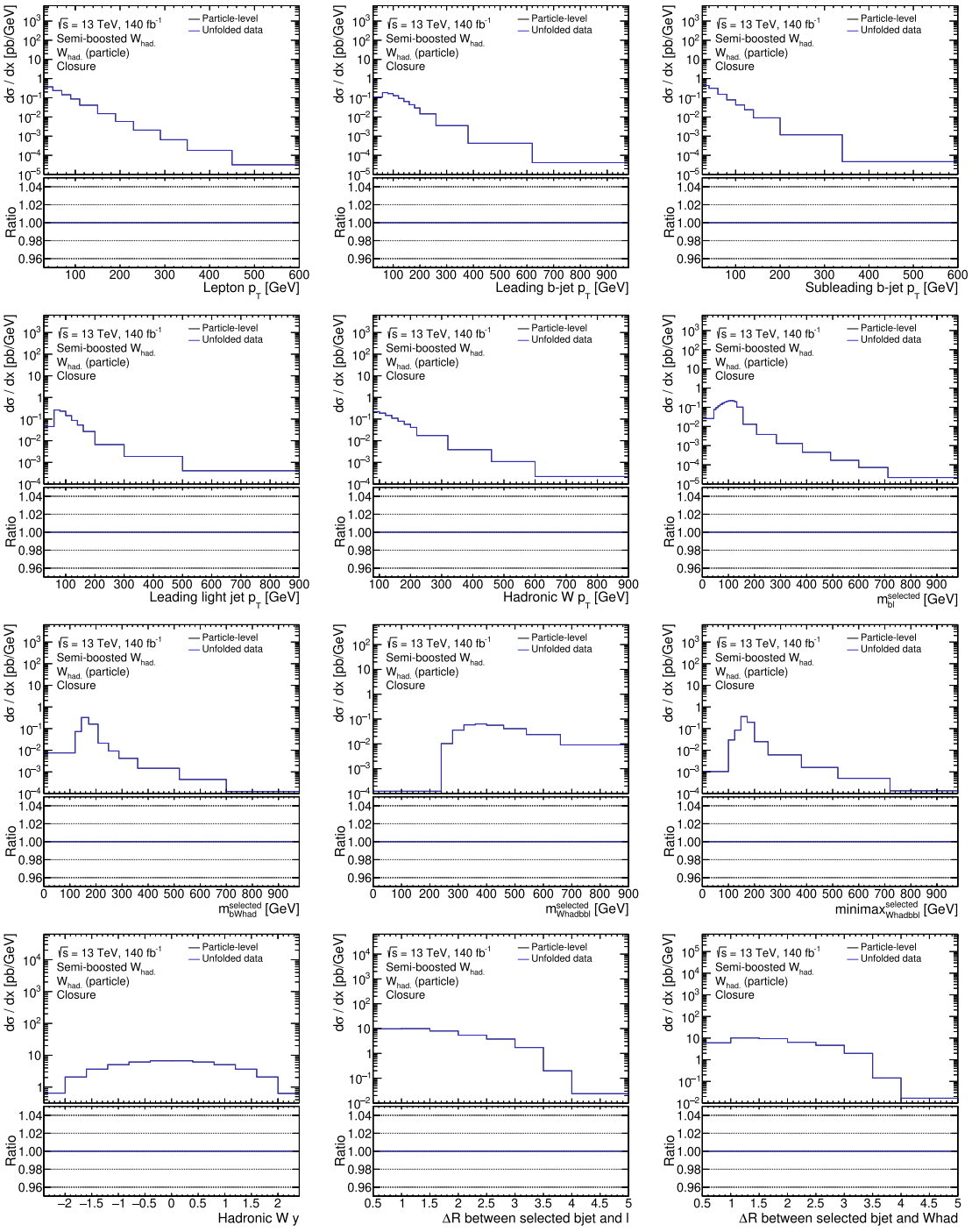


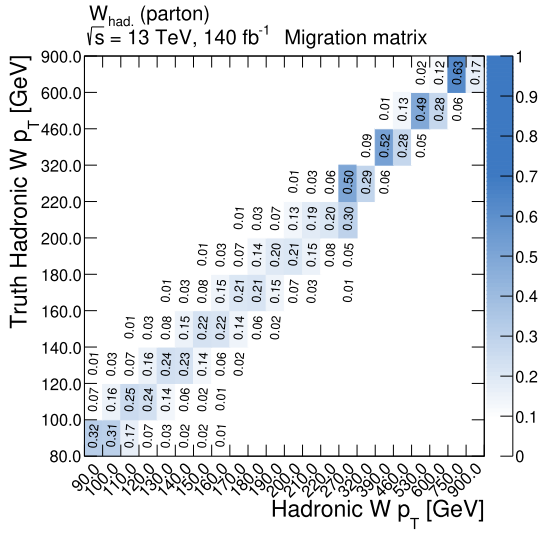
Figure D.1: Unfolding Closure Test for the $W_{\text{had.}}$ (particle) Region. The closure test validates the correctness of the unfolding procedure. The test is conducted by unfolding MC events and comparing the unfolded distribution to the truth-level distribution using the same MC sample. The closure test is passed if the two distributions agree within statistical uncertainties. Refer to Section 7.3 for a description of the observables of interest.

APPENDIX E

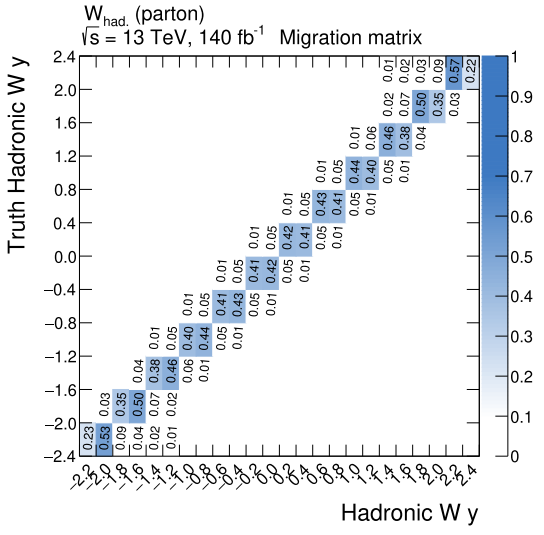
Parton-Level Measurement Region

An additional measurement region is defined using parton-level definition of the hadronic-W, labeled as $W_{\text{had.}}$ (parton). Recall that the parton-level refers to objects contained in the MC truth record at the time of event generation. In this analysis, we are particularly interested in the simulation of the hadronic-W and its decay. Therefore, one of the motivations for defining a parton-level measurement region is to compare the unfolded results when using the parton-level definition of the hadronic-W to the unfolded results when using the nominal particle-level definition of the hadronic-W.

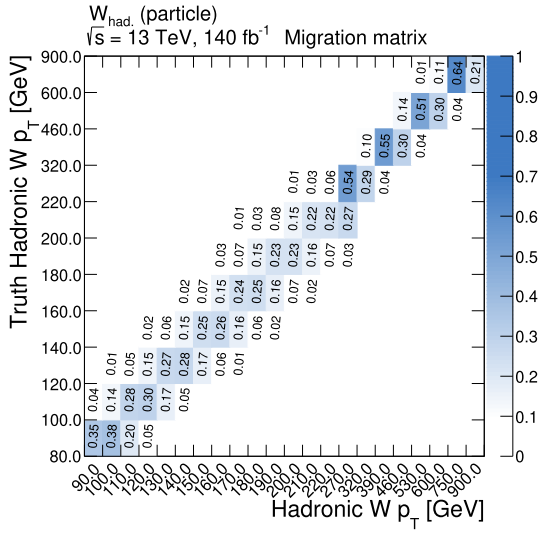
The parton-level measurement region applies the same detector-level selection but replaces all instances of the particle-level hadronic-W with the parton-level hadronic-W. Any cuts on the kinematics of the hadronic-W (i.e., p_{T}, η, mass) are applied using the parton-level hadronic-W. Similarly, when constructing the migration matrices, acceptance, and efficiency correction factors, the parton-level hadronic-W is used. Figure E.1 compares the migration matrices for the particle-level and parton-level measurement regions for the hadronic-W transverse momentum and rapidity.



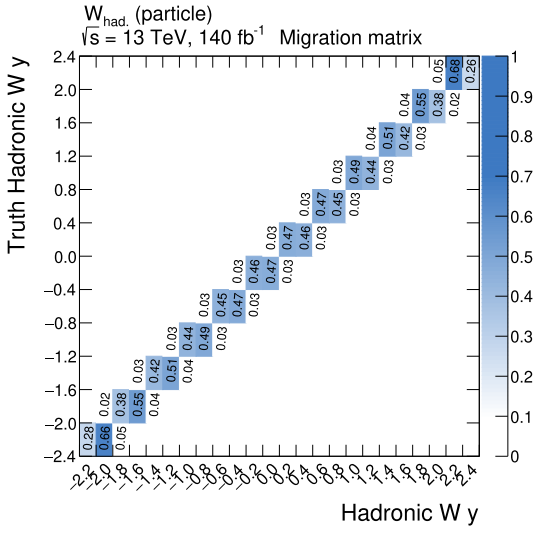
(a) Hadronic-W (parton) p_{T} migration matrix.



(b) Hadronic-W (parton) y migration matrix.



(c) Hadronic-W (particle) p_{T} migration matrix.

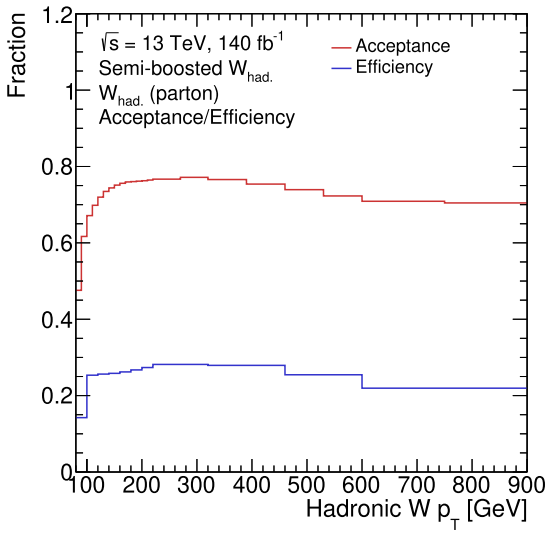


(d) Hadronic-W (particle) y migration matrix.

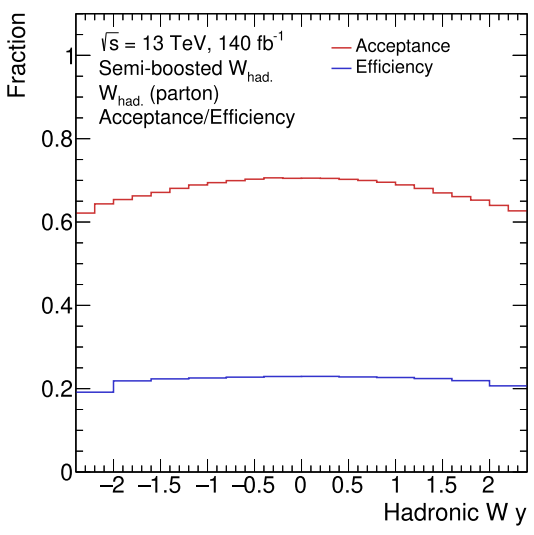
Figure E.1: Migration Matrices Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions. At truth-level, the migration matrices for the hadronic-W transverse momentum can be constructed using the parton-level definition of the hadronic-W. The migrations matrices for the parton-level hadronic-W **(a)** transverse momentum and **(b)** rapidity are compared to the particle-level hadronic-W **(c)** transverse momentum and **(d)** rapidity.

When comparing the migration matrices, the immediate observation is that the particle-level migration matrix is more diagonal than the parton-level migration matrix. The lower p_{T} bins (< 180 GeV) of the parton-level matrix show significantly more migrations between bins. However, in the higher p_{T} bins (> 270 GeV), the migration matrices are nearly identical. Large migrations at lower p_{T} indicate that the parton-level definition of the hadronic-W is less well-defined at lower p_{T} values. This

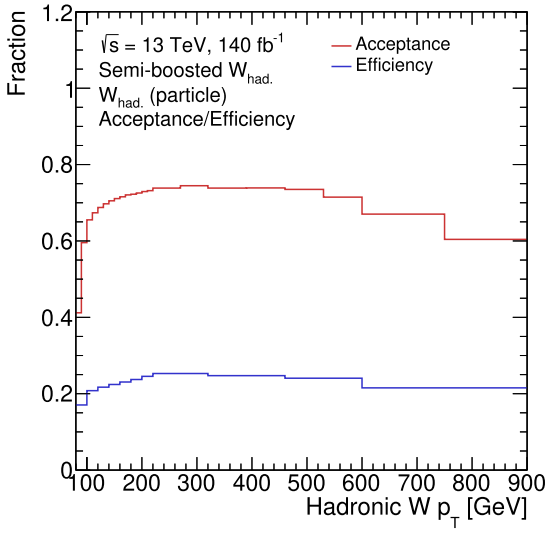
lower purity is expected given that the studies of the reconstruction efficiency of the hadronic-W (Section 6.2.3) showed that the reconstruction efficiency of the hadronic-W is lower at lower p_{T} values. In addition, these studies showed that the reconstruction efficiency of the hadronic-W is generally higher for the particle-level definition of the hadronic-W than the parton-level definition of the hadronic-W This is reflected in the more diagonal migration matrix. Figure E.2 compares the acceptance and efficiency correction factors for the hadronic-W using the parton and particle-level definitions.



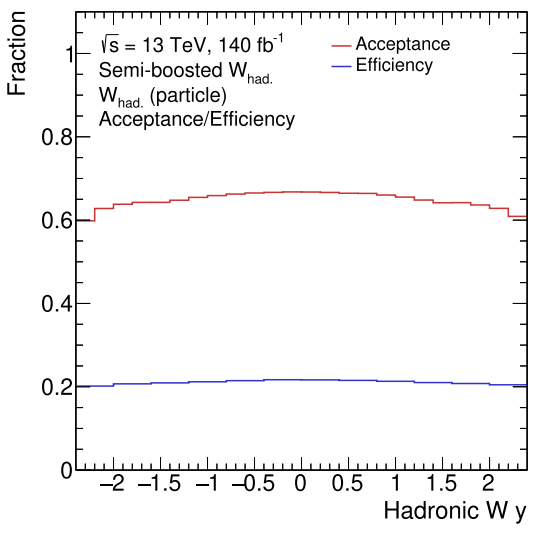
(a) Hadronic-W (parton) p_{T} acceptance/efficiency.



(b) Hadronic-W (parton) y acceptance/efficiency.



(c) Hadronic-W (particle) p_{T} acceptance/efficiency.



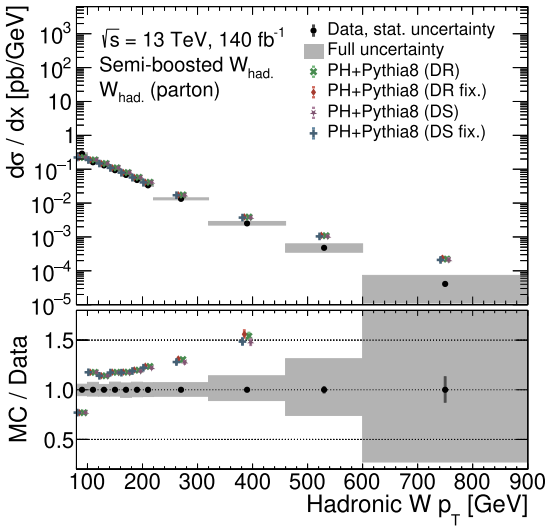
(d) Hadronic-W (particle) y acceptance/efficiency.

Figure E.2: Acceptance/Efficiency Factors Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions. The acceptance and efficiency correction factors can also be evaluated for the parton and particle-level definitions of the hadronic-W. The acceptance and efficiency correction factors for the parton-level hadronic-W **(a)** transverse momentum and **(b)** rapidity are compared to the particle-level hadronic-W **(c)** transverse momentum and **(d)** rapidity.

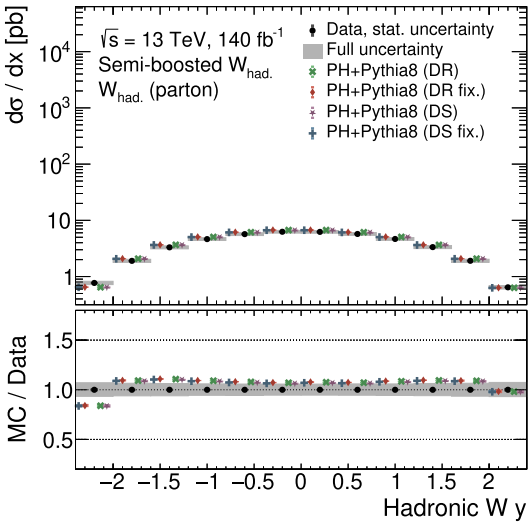
The parton-level and particle-level acceptance and efficiency correction factors are similar. The acceptance and efficiency correction factors are generally flat as a function of p_{T}, except the lowest p_{T} bin, a common feature between both definitions. The acceptance/efficiency factors are also slightly lower in the particle-level definition.

These ingredients are supplied to the unfolding procedure to obtain cross-sections comparing the

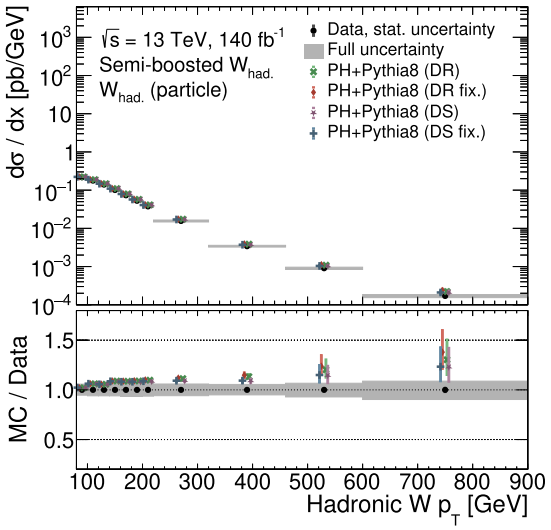
parton and particle-level differential cross-section distributions. Figures E.3 and E.4 plots the absolute differential cross-sections for the nominal and alternative MC samples, respectively.



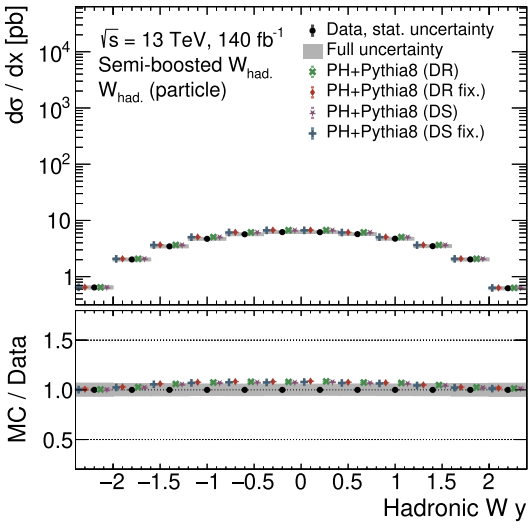
(a) Hadronic-W (parton) p_{T} cross-sections.



(b) Hadronic-W (parton) y cross-sections.

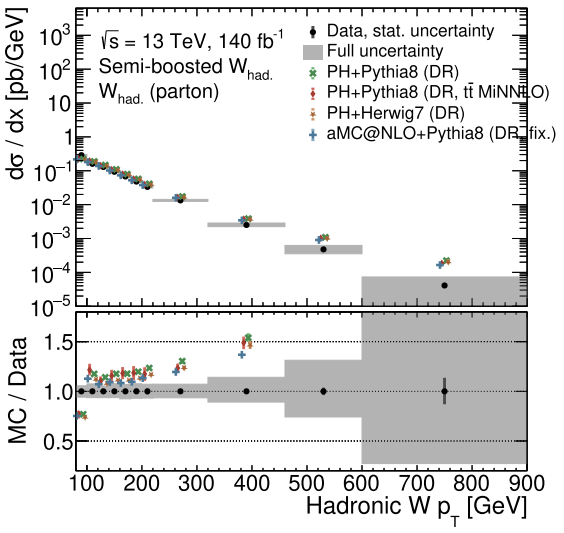


(c) Hadronic-W (particle) p_{T} cross-sections.

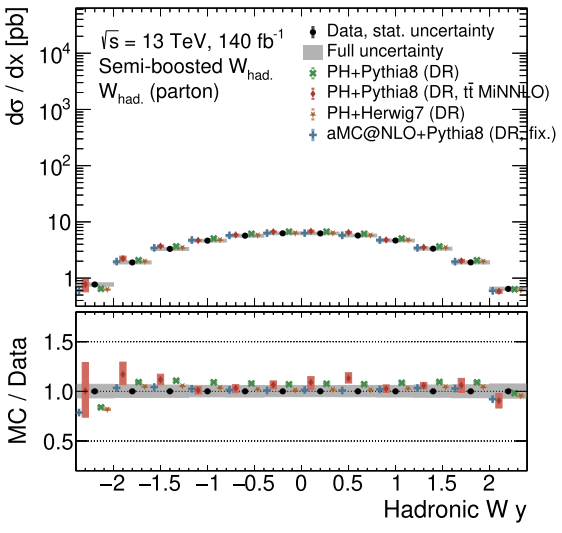


(d) Hadronic-W (particle) y cross-sections.

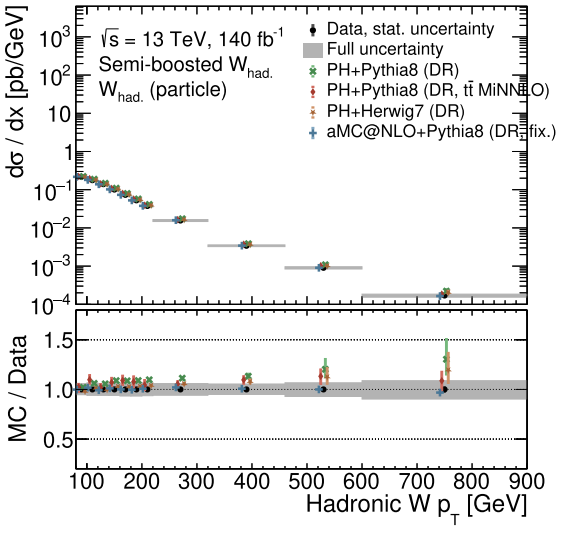
Figure E.3: Differential Cross-Sections Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions. The differential cross-sections are shown comparing the parton-level and particle-level definitions of the hadronic-W. The parton-level differential cross-sections include the hadronic-W **(a)** transverse momentum and **(b)** rapidity. The particle-level differential cross-sections include the hadronic-W **(c)** transverse momentum and **(d)** rapidity.



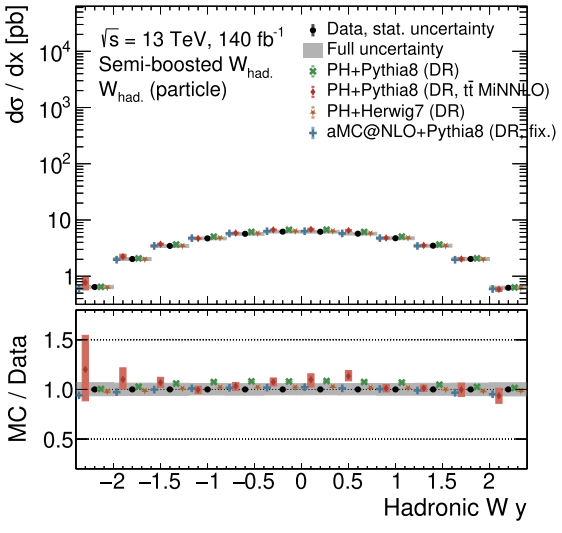
(a) Hadronic-W (parton) p_{T} cross-sections.



(b) Hadronic-W (parton) y cross-sections.



(c) Hadronic-W (particle) p_{T} cross-sections.



(d) Hadronic-W (particle) y cross-sections.

Figure E.4: Differential Cross-Sections Comparing the Hadronic-W (parton) and Hadronic-W (particle) Definitions. The differential cross-sections are shown comparing the parton-level and particle-level definitions of the hadronic-W. The parton-level differential cross-sections include the hadronic-W **(a)** transverse momentum and **(b)** rapidity. The particle-level differential cross-sections include the hadronic-W **(c)** transverse momentum and **(d)** rapidity.

When using the parton-level definition, the predicted cross-sections are generally more poorly modeled than the particle-level definition. There may be challenges in certain regions of phase space where non-perturbative effects become important. The predictions systematically overestimate the data, which is especially evident in Figure E.3(a). The uncertainties are also much larger in the parton-level distributions. The total uncertainties can become as large as 100% in the highest p_{T} bins. In

contrast, the particle-level definition have the uncertainties that are generally less than 10%. Table E.1 shows the χ^2 and p-values for the parton-level and particle-level definitions of the hadronic-W.

Table E.1: χ^2 and p-values for the $W_{had.}$ (parton) Region. The χ^2 and p-values are calculated by comparing the measured cross-sections to the MC predictions. The χ^2 values are then normalized to the number of degrees of freedom. Using the χ^2 value and NDF, the p-value is calculated. The p-value is the probability of observing a χ^2/NDF value as extreme as the one observed by chance alone. The χ^2/NDF and p-values are computed for the observables of interest in the $W_{\text{had.}}$ (parton) region. Included for comparison are the same χ^2 and p-values for the $W_{\text{had.}}$ (particle) region that was presented earlier in the document.

	Hadronic-W p_{T}			
	$W_{\text{had.}}$ (parton)		$W_{\text{had.}}$ (particle)	
Generator	χ^2/NDF	p-value	χ^2/NDF	p-value
PH+PYTHIA 8 (DR)	206.3 / 11	<0.001	17.87 / 11	0.085
PH+PYTHIA 8 (DR, fix.)	213.6 / 11	<0.001	22.70 / 11	0.019
PH+PYTHIA 8 (DS)	211.3 / 11	<0.001	12.03 / 11	0.362
PH+PYTHIA 8 (DS, fix.)	210.6 / 11	<0.001	12.35 / 11	0.338
PH+PYTHIA 8 (DR, $t\bar{t}$ MiNNLO)	199.3 / 11	<0.001	30.53 / 11	0.001
PH+HERWIG 7.1 (DR)	189.9 / 11	<0.001	10.66 / 11	0.473
aMCNLO+PYTHIA 8 (DR, fix.)	178.8 / 11	<0.001	3.40 / 11	0.984
	Hadronic-W y			
	$W_{\text{had.}}$ (parton)		$W_{\text{had.}}$ (particle)	
Generator	χ^2/NDF	p-value	χ^2/NDF	p-value
PH+PYTHIA 8 (DR)	98.97 / 12	<0.001	14.95 / 12	0.244
PH+PYTHIA 8 (DR, fix.)	99.39 / 12	<0.001	15.26 / 12	0.228
PH+PYTHIA 8 (DS)	97.70 / 12	<0.001	14.24 / 12	0.285
PH+PYTHIA 8 (DS, fix.)	97.81 / 12	<0.001	14.17 / 12	0.290
PH+PYTHIA 8 (DR, $t\bar{t}$ MiNNLO)	439.3 / 12	<0.001	418.3 / 12	<0.001
PH+HERWIG 7.1 (DR)	80.48 / 12	<0.001	7.85 / 12	0.796
aMCNLO+PYTHIA 8 (DR, fix.)	91.39 / 12	<0.001	9.83 / 12	0.630

The χ^2 values for the parton-level definition of the hadronic-W are significantly larger than the particle-level definition. The p-values are also much smaller, indicating that the parton-level definition of the hadronic-W is not as well-modeled as the particle-level definition. The particle-level MADGRAPH5_AMC@NLO generator does a particularly good job modeling the p_{T} distribution, with a χ^2/NDF value of 4.23/11 and a p-value of 0.963. Similarly, the PH+HERWIG 7.1 generator performs the best at modeling the rapidity distribution, with a χ^2/NDF value of 9.140/12 and a p-value of 0.691.

APPENDIX F

Leading/Subleading b-jet Asymmetry

We have also studied additional variables of particular interest in the context of top quark polarization. In this context, polarization refers to the preferred orientation of the top quark's spin along a specific direction. Unlike the other quarks, the top quark is a unique laboratory to study polarization effects. The top quark decays before hadronization; this ensures that information regarding its helicity is passed onto its decay products. Therefore, to study top quark polarization, one can study the various distributions of the decay products. Polarization carries important implications for Charge-Parity (CP) symmetry within the SM.

Recall that the two dominant production modes for top quarks at the LHC correspond to $t\bar{t}$ pairs and singly-resonant tW. In the case of $t\bar{t}$ production, the top quarks are predominantly unpolarized since the production of top quarks pairs (via gluon-gluon fusion or quark-antiquark annihilation) is mediated by the strong force and is parity conserving. Generally, this means the top quarks select no preferential polarization state. On the other hand, singly-resonant top production proceeds via the electroweak process, and the weak force is known to violate parity. The production of the top quark is purely left-handed as a result [135].

There have been previous studies that have measured the polarization of top quarks both in SM and BSM contexts [136–142]. It is common in these studies to measure the differences in polarization (as a means of studying CP violation) through various asymmetry and fractional observables. For example, rapidity and energy asymmetry can be measured between the top and anti-top quarks can be used to measure charge asymmetry. In this analysis, we have selected the leading and subleading b-jet p_{T} asymmetry and fraction as a means to study top quark polarization. The b-jet asymmetry measures the difference in the p_{T} of the leading and subleading b-jets in the event. The b-jet fraction measures the fraction of the total jet p_{T} carried by the leading and subleading b-jets.

The leading and subleading b-jet p_{T} asymmetry ($A_{b1,2}$) is defined as the difference in the p_{T} of the leading ($p_{\mathrm{T}}(b_1)$) and subleading ($p_{\mathrm{T}}(b_2)$) b-jets divided by their sum.

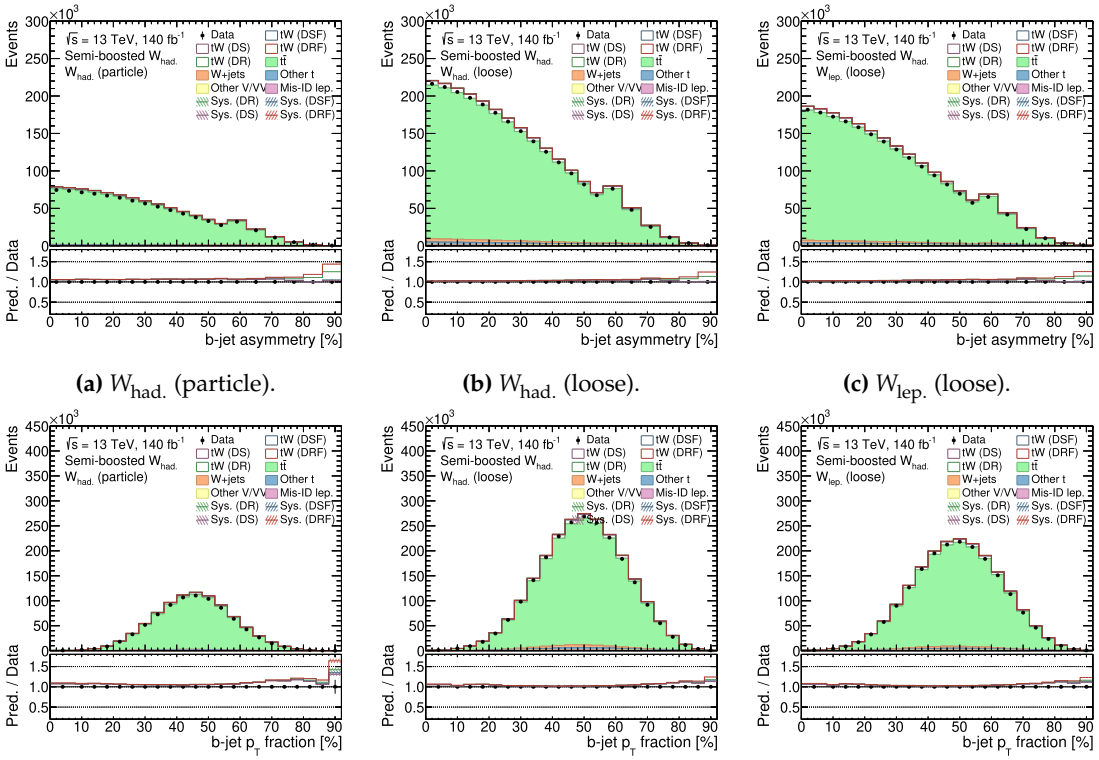
$$A_{b1,2} = [p_{\mathrm{T}}(b_1) - p_{\mathrm{T}}(b_2)] / [p_{\mathrm{T}}(b_1) + p_{\mathrm{T}}(b_2)] \tag{F.1}$$

The leading and subleading b-jet fraction ($F_{b1,2}$) is defined as the sum of the p_{T} of the leading and

subleading b-jets divided by the sum of the p_{T} of all the jets in the event ($H_{\mathrm{T}} = \sum_i p_{\mathrm{T}}(j_i)$).

$$F_{b1,2} = \left[p_{\mathrm{T}}(b_1) + p_{\mathrm{T}}(b_2)\right] / H_{\mathrm{T}} \tag{F.2}$$

Figure F.1 plots the detector-level distributions of the b-jet p_{T} asymmetry and fraction for the $W_{\mathrm{had.}}$ (particle), $W_{\mathrm{had.}}$ (loose), and $W_{\mathrm{lep.}}$ (loose) regions. The following figures are plotted using a linear scale to visualize the shapes of the distributions better.



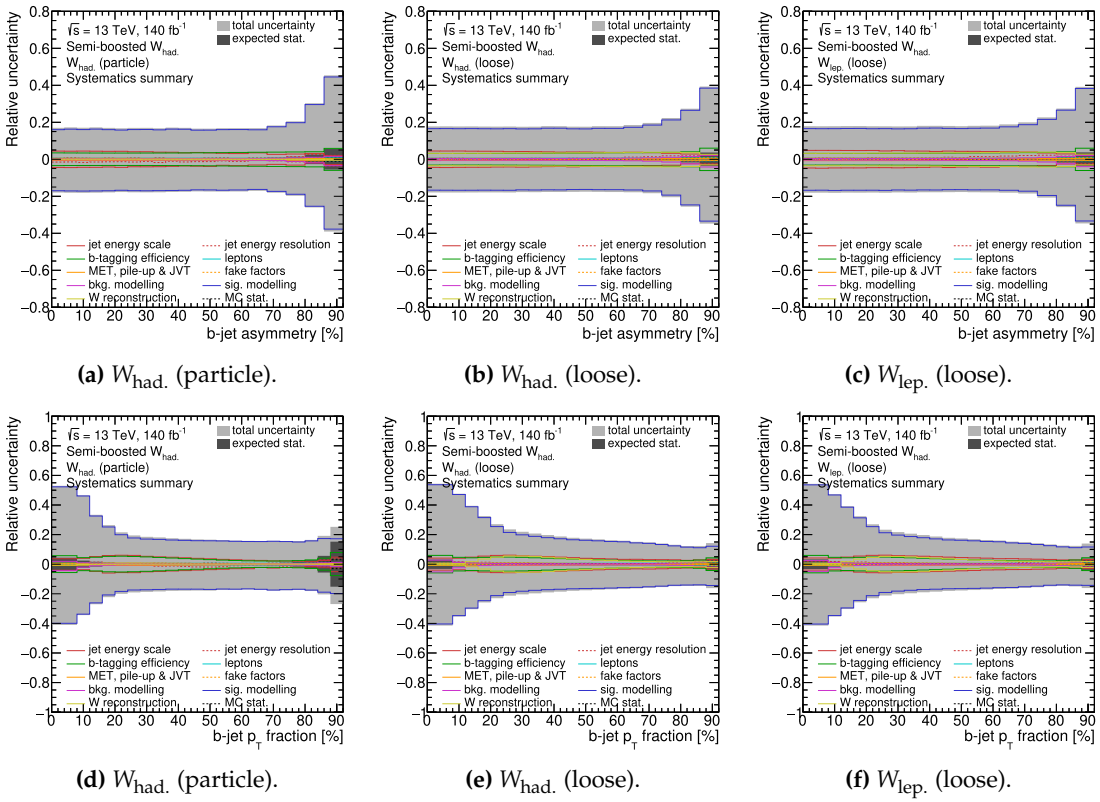
(a) $W_{\mathrm{had.}}$ (particle). **(b)** $W_{\mathrm{had.}}$ (loose). **(c)** $W_{\mathrm{lep.}}$ (loose).

(d) $W_{\mathrm{had.}}$ (particle). **(e)** $W_{\mathrm{had.}}$ (loose). **(f)** $W_{\mathrm{lep.}}$ (loose).

Figure F.1: b-jet p_{T} Asymmetry and Fraction Detector-Level Distributions. The b-jet p_{T} asymmetry measures the difference in the p_{T} of the leading and subleading b-jets in the event. The b-jet fraction measures the fraction of the total jet p_{T} carried by the leading and subleading b-jets. The detector-level distributions are shown for the $W_{\mathrm{had.}}$ (particle) (left), $W_{\mathrm{had.}}$ (loose) (middle), and $W_{\mathrm{lep.}}$ (loose) (right) regions. The uncertainties shown are statistical only.

The b-jet p_{T} asymmetry in all three regions has a maximum at zero, meaning that the leading and subleading b-jets in most events have similar p_{T}. The distributions then decrease almost linearly at higher values of the b-jet asymmetry. There are very few events with very disparate p_{T} values. The b-jet fraction distributions in all three regions have a maximum centered at approximately 0.5. The sum of the p_{T} of the leading and subleading b-jets is approximately half of the total jet p_{T} in the event with a distribution that resembles a Gaussian distribution. The MC agrees with data within uncertainties but is slightly overestimating the data in the $W_{\mathrm{had.}}$ (particle) region. The largest disagreements with MC

come only in the tails of each distribution. Figure F.2 plots the breakdown of the various systematic uncertainties.

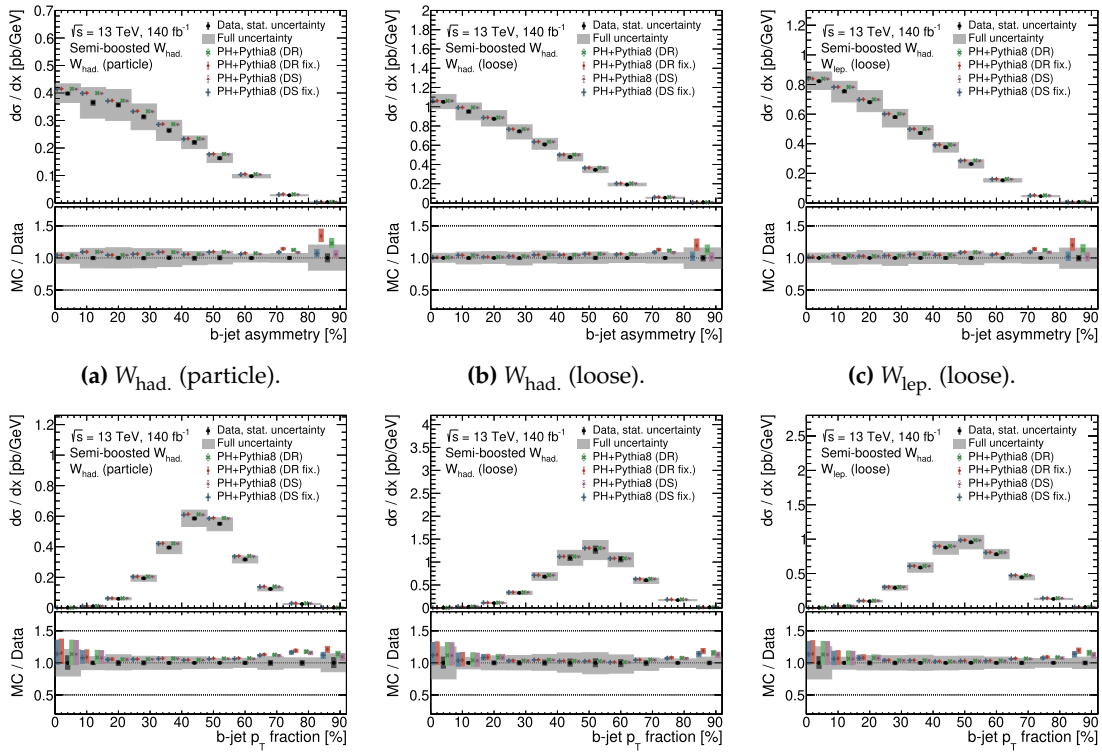


(a) $W_{\text{had.}}$ (particle). **(b)** $W_{\text{had.}}$ (loose). **(c)** $W_{\text{lep.}}$ (loose).

(d) $W_{\text{had.}}$ (particle). **(e)** $W_{\text{had.}}$ (loose). **(f)** $W_{\text{lep.}}$ (loose).

Figure F.2: b-jet p_{T} Asymmetry and Fraction Systematics Uncertainties Summary. The various systematics uncertainties are summarized for the b-jet p_{T} asymmetry and fraction. The systematics are plotted for the $W_{\text{had.}}$ (particle) (left), $W_{\text{had.}}$ (loose) (middle), and $W_{\text{lep.}}$ (loose) (right) regions.

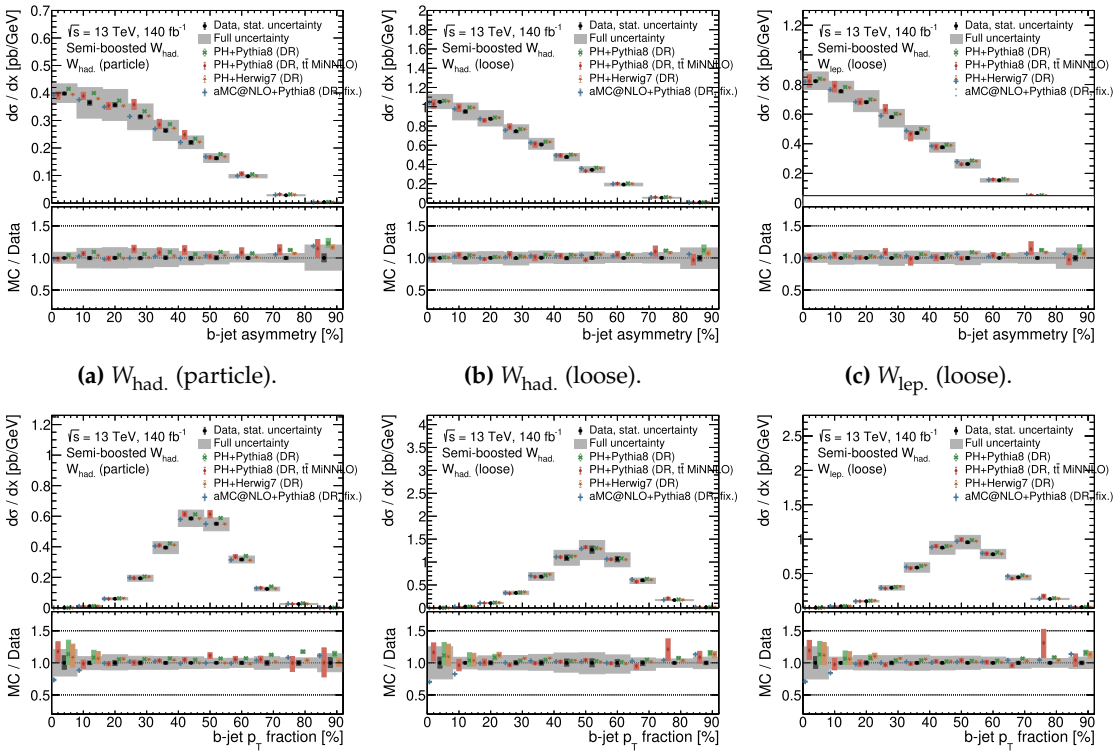
As before, signal modeling contributes the largest source of uncertainty in the b-jet p_{T} asymmetry and fraction. The uncertainty in the b-jet p_{T} asymmetry is relatively constant at 0.15 until the asymmetry reaches approximately 0.7. On the other hand, the uncertainty is largest for events with a small b-jet p_{T} fraction and exponentially decreases towards larger b-jet p_{T} fraction values. These observables can also be unfolded following the procedures described in Chapter 9. Figures F.3 and F.4 plot the absolute cross-sections of the unfolded b-jet p_{T} asymmetry and fraction using the nominal and alternative sets of MC, respectively.



(a) $W_{\text{had.}}$ (particle). **(b)** $W_{\text{had.}}$ (loose). **(c)** $W_{\text{lep.}}$ (loose).

(d) $W_{\text{had.}}$ (particle). **(e)** $W_{\text{had.}}$ (loose). **(f)** $W_{\text{lep.}}$ (loose).

Figure F.3: *b*-jet p_{T} Asymmetry and Fraction Absolute Differential Cross-Sections (Nominal MC). The absolute cross-sections are plotted for the $W_{\text{had.}}$ (particle) (left), $W_{\text{had.}}$ (loose) (middle), and $W_{\text{lep.}}$ (loose) (right) regions using the nominal set of MC generators.



(a) $W_{\text{had.}}$ (particle). **(b)** $W_{\text{had.}}$ (loose). **(c)** $W_{\text{lep.}}$ (loose).

(d) $W_{\text{had.}}$ (particle). **(e)** $W_{\text{had.}}$ (loose). **(f)** $W_{\text{lep.}}$ (loose).

Figure F.4: b-jet p_{T} Asymmetry and Fraction Absolute Differential Cross-Sections (Alternative MC). The absolute cross-sections are plotted for the $W_{\text{had.}}$ (particle) (left), $W_{\text{had.}}$ (loose) (middle), and $W_{\text{lep.}}$ (loose) (right) regions using the alternative set of MC generators.

Table F.1 lists the χ^2 and p-values for the b-jet p_{T} asymmetry and fraction. The χ^2 values are calculated using the unfolded data and each of the MC predictions using the same procedure described in Chapter 9.

Table F.1: χ^2 and p-values for the b-jet p_T Asymmetry and Fraction. The χ^2 and p-values are shown for the b-jet p_T asymmetry and fraction in the $W_{had.}$ (particle), $W_{had.}$ (loose), and $W_{lep.}$ (loose) regions.

	b-jet p_T Asymmetry					
	$W_{had.}$ (particle)		$W_{had.}$ (loose)		$W_{lep.}$ (loose)	
Generator	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
PH+PYTHIA 8 (DR)	4.51 / 10	0.92	3.04 / 10	0.98	3.62 / 10	0.96
PH+PYTHIA 8 (DR, fix.)	6.54 / 10	0.77	3.82 / 10	0.96	4.61 / 10	0.92
PH+PYTHIA 8 (DS)	2.78 / 10	0.99	2.75 / 10	0.99	2.99 / 10	0.98
PH+PYTHIA 8 (DS, fix.)	2.93 / 10	0.98	2.74 / 10	0.99	3.01 / 10	0.98
PH+PYTHIA 8 (DR, $t\bar{t}$ MiNNLO)	43.00 / 10	<0.001	20.38 / 10	0.03	13.91 / 10	0.18
aMCNLO+PYTHIA 8 (DR, fix.)	1.98 / 10	1.00	2.36 / 10	0.99	2.88 / 10	0.98
PH+HERWIG 7.1 (DR)	2.30 / 10	0.99	2.89 / 10	0.98	3.14 / 10	0.98
	b-jet p_T Fraction					
	$W_{had.}$ (particle)		$W_{had.}$ (loose)		$W_{lep.}$ (loose)	
Generator	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
PH+PYTHIA 8 (DR)	68.50 / 11	<0.001	66.78 / 11	<0.001	33.31 / 11	<0.001
PH+PYTHIA 8 (DR, fix.)	70.70 / 11	<0.001	67.69 / 11	<0.001	33.72 / 11	<0.001
PH+PYTHIA 8 (DS)	60.21 / 11	<0.001	59.38 / 11	<0.001	29.60 / 11	0.001
PH+Pythia[8] (DS, fix.)	60.55 / 11	<0.001	59.57 / 11	<0.001	29.64 / 11	0.001
PH+PYTHIA 8 (DR, $t\bar{t}$ MiNNLO)	196.50 / 11	<0.001	94.47 / 11	<0.001	101.00 / 11	<0.001
aMCNLO+PYTHIA 8 (DR, fix.)	87.14 / 11	<0.001	262.80 / 11	<0.001	191.00 / 11	<0.001
PH+HERWIG 7.1 (DR)	28.32 / 11	0.001	32.20 / 11	<0.001	9.04 / 11	0.62

The b-jet p_T asymmetry is a particularly well-modeled quantity. The unfolded cross-sections show that the predictions are well-matched to the data for all values of the b-jet asymmetry. As the detector-level plots show, the various MC generators slightly overestimate the asymmetry. Regarding the b-jet p_T fraction, this quantity shows MC-data agreement near the peak of 0.5, but the MC predictions are not as well-matched to the data in the tails of the distributions. This is particularly true for the aMCNLO+PYTHIA 8 (DR, fix.) generator at low b-jet fractions where agreement with observations is generally poor.