

THE UNIVERSITY OF CHICAGO

A GLOBAL VIEW OF JETS WITH THE ATLAS DETECTOR: FROM HARDWARE
TRIGGERS TO PRECISION MEASUREMENTS AND BEYOND

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
EMILY ANN SMITH

CHICAGO, ILLINOIS
DECEMBER 2023



Copyright © 2023 by Emily Ann Smith

All Rights Reserved

TABLE OF CONTENTS

LIST OF FIGURES	viii
LIST OF TABLES	xix
ACKNOWLEDGMENTS	xxi
ABSTRACT	xxiii
1 INTRODUCTION	1
2 THE STANDARD MODEL	3
2.1 Introduction to the Standard Model	4
2.2 The Electroweak Sector: Quantum Electrodynamics	5
2.3 Electroweak Symmetry Breaking and The Higgs Mechanism	8
2.4 Quantum Chromodynamics: Life in Color	11
2.5 Beyond the Standard Model	16
2.5.1 Supersymmetry	19
3 THE LARGE HADRON COLLIDER	21
3.1 Motivation	22
3.2 Machine Overview	23
3.3 Luminosity and Pile-Up	25
3.4 LHC Upgrades and Schedule	27
4 THE ATLAS EXPERIMENT	31
4.1 ATLAS Geometry	33
4.2 Inner Detector	35

4.3	Calorimeters and Calorimetry	38
4.3.1	Electromagnetic Calorimetry	39
4.3.2	Hadronic Calorimetry	41
4.3.3	Liquid Argon Calorimeter	42
4.3.4	Tile Calorimeter	44
4.3.5	Calorimeter Energy Resolution	44
4.3.6	Calorimeter Upgrades for Run 3	47
4.4	Muon Detectors	49
4.5	Trigger and Data Acquisition	50
4.5.1	Level-1 Calorimeter Trigger	53
4.5.2	TDAQ Run 3 Upgrades	56
5	GLOBAL FEATURE EXTRACTOR	61
5.1	Overview	61
5.2	Hardware Design	62
5.2.1	Virtex 9 Ultrascale+ FPGA	66
5.2.2	Zynq+ System on Chip	68
5.3	Firmware Infrastructure	69
5.3.1	Fiber Control & Monitoring	71
5.4	Firmware Algorithms	75
5.4.1	Tower Builder	75
5.4.2	Jet Finder: Large Radius Jets Algorithm	76
5.4.3	Missing Transverse Energy Finder	82
5.5	gFEX Trigger Efficiencies	83
5.6	Custom Operating System	91
5.6.1	OpenEmbedded	92

5.6.2	The Yocto Project	92
5.6.3	Bitbake	94
5.6.4	Xilinx Yocto-Manifests	94
5.6.5	gFEX Custom OS Implementation	95
5.7	Control and Monitoring	99
5.7.1	I2C Monitoring	99
5.7.2	IPMC and ATCA Shelf Monitoring	101
5.7.3	FPGA-Based Monitoring and Control	103
5.7.4	Ironman IPBus Communication	104
5.8	Detector Control Systems	108
5.9	Integration and Commissioning	119
5.9.1	Surface Test Facility	121
5.9.2	gFEX Installation	122
5.9.3	Fiber and Channel Mapping	123
5.10	gFEX in Run 3	125
6	JETS: HADRONIC PHYSICS WITH THE ATLAS EXPERIMENT	128
6.1	Monte Carlo Modeling	129
6.2	Jet Algorithms and Reconstruction	130
7	SEARCH FOR CHARGINO AND NEUTRALINO PRODUCTION IN FINAL STATES WITH A HIGGS BOSON AND MISSING TRANSVERSE MOMENTUM AT $\sqrt{s} =$ 13 TeV WITH THE ATLAS DETECTOR	133
7.1	Introduction	133
7.2	Analysis Methods	135
7.2.1	MC Samples	137
7.2.2	Data Samples	139

7.2.3	Object Definition and Reconstruction	142
7.2.4	Discriminating Variables	146
7.3	Preselection	148
7.4	Signal Region Selection	152
7.4.1	Resolved Signal Regions	154
7.4.2	Boosted Signal Regions	161
7.5	Resolved and Boosted Combined Sensitivity	167
7.6	Boosted Channel Background Estimation	169
7.7	Boosted Channel Systematic Uncertainties	171
7.8	Boosted Channel Results	172
8	PRECISION MEASUREMENT OF LUND AND PRIMARY LUND MULTIPLICI- TIES IN JETS	176
8.1	Motivation	177
8.2	Observable Definition	178
8.3	Data	182
8.4	Simulation	183
8.5	Jets, Tracks, and Event Selection	190
8.5.1	Jets	190
8.5.2	Tracks	194
8.5.3	Event Selection	194
8.5.4	Binning and Normalization	195
8.6	Unfolding and Related Quantities	197
8.6.1	Iterative Bayesian Unfolding	197
8.6.2	Response Matrices	200
8.6.3	Fake and Efficiency Factors	210

8.6.4	Technical Closure	215
8.6.5	Data Driven Non-Closure test for IBU	220
8.7	Systematic Uncertainties	229
8.7.1	Statistical uncertainties	229
8.7.2	Pile-up reweighting uncertainty	234
8.7.3	Disabled TILE Calorimeter modules	239
8.7.4	Jet energy scale and resolution (JES/JER) uncertainties	245
8.7.5	Tracking-related uncertainties	254
8.7.6	Uncertainty due to the choice of Monte Carlo models	260
8.8	Results	273
8.8.1	Multiplicity Distributions	273
8.8.2	Average Multiplicities	284
9	CONCLUSION	292
	APPENDICES	294
A	GFEX MONITORED HARDWARE SENSORS	295
B	GFEX CLOCK CHIP RELEVANT REGISTERS	302
C	OPTIMIZATIONS OF ITERATIONS IN UNFOLDING	306
D	MULTIPLICITIES REBINNING STUDY	331
E	MULTIPLICITIES REBINNING AVERAGES STUDY	342
	BIBLIOGRAPHY	350

LIST OF FIGURES

2.1	Summary of ATLAS measurements compared with calculated theory cross sections for many SM processes [197].	3
2.2	A summary of Standard Model particles and their properties. Image from Wikipedia.	4
2.3	Running of α_s as a function of the energy scale [197].	13
2.4	Measurements of the strong coupling constant α_s at different experiments (a) and at hadron colliders (b) [197].	14
2.5	Summary of the ordinary matter vs dark matter in the universe, as observed by Planck 2015 CMB measurement. Image by David Miller [167].	17
3.1	A diagram of the LHC showing the four experiments ATLAS, CMS, LHCb, and ALICE. Also shown is the Super Proton Synchrotron which is the penultimate stage of the injector chain before the beam enters the LHC [161].	21
3.2	A diagram of the LHC injection chain [158].	24
3.3	Mean number of interactions per LHC Run [94].	27
3.4	LHC luminosity for Runs 1 - 3 [94].	29
3.5	LHC long term schedule, available online at the lhc-commisioning website. . . .	30
4.1	A labeled diagram of the ATLAS detector [91].	31
4.2	A transverse view of the ATLAS detector [168].	32
4.3	ATLAS coordinate systems, modified from ATLAS detector images in Ref [91]. .	34
4.4	The ATLAS inner detector [91].	36
4.5	The ATLAS inner detector components, viewed radially [86].	37
4.6	The ATLAS calorimeters, labeled [91].	39
4.7	Fractional energy loss per radiation length in lead as a function of electron or positron energy [197].	40
4.8	Hadronic shower with the characteristic electromagnetic shower component [144].	42
4.9	A diagram of a LAr module which shows the accordion geometry and the different layers radially, and in ϕ and η . The granularity is also shown [91].	43

4.10	A diagram of a Tile module which shows shape and granularity of the modules [91].	45
4.11	Run 2 to Run 3 Liquid Argon calorimeter cell upgrades [4].	48
4.12	The ATLAS muon detector, labeled [91].	50
4.13	Overview of the Run 2 Trigger and Data Acquisition System for the ATLAS detector. Shows the Level-1 Trigger, the HLT, and the data readout path. [89] .	51
4.14	Breakdown of the overall trigger rate from the hardware and software systems to the offline readout. Image Credit: Maurizio Pierini.	52
4.15	Run 2 L1Calo Schematic showing the flow of data from the calorimeters through to the L1CTP.	53
4.16	Visualization of L1Calo Trigger Towers, corresponding to summed electromagnetic and hadronic calorimeter cells.	54
4.17	Run 2 $\langle\mu\rangle$ in green versus the early Run 3 $\langle\mu\rangle$ from 2022.	57
4.18	L1Calo Schematic showing the flow of data from the calorimeters through to the L1CTP for Run 3.	58
5.1	Higgs boson hadronic decay with boost	61
5.2	ATCA shelf setup with IPMC and Shelf Manager shown [156].	62
5.3	Full gFEX schematic [65].	64
5.4	gFEX faceplate schematic [65].	65
5.5	gFEX pFPGA link connections [65].	66
5.6	Diagram of Virtex FPGA components [198].	67
5.7	Comparison of Ultrascale+ FPGA resources [198].	68
5.8	Zynq+ System on Chip architecture [198].	69
5.9	gFEX full clock tree [65].	71
5.10	gFEX input fiber alignment. Diagram Credit: Greg Myers.	73
5.11	gFEX Firmware Block Diagram	76
5.12	gFEX input data formats.	77

5.13 gFEX gTower η and ϕ index, and the overall tower indexing scheme. Figure Credit: Greg Myers.	79
5.14 gFEX gBlock Event Display. Image Credit: David Strom and David Miller. . . .	81
5.15 gFEX gJet Event Display. Image Credit: David Strom and David Miller.	81
5.16 η dependent ρ for gFEX pileup subtraction. Plot credit: Aparajita Dattagupta.	82
5.17 Early gFEX Trigger Efficiency Study [95].	84
5.18 gFEX Efficiency Study Samples	86
5.19 Kinematic relationships between gJets, gBlocks, and Offline jets.	87
5.20 gFEX efficiency compared to Run 2 trigger efficiency with no pileup subtraction applied.	88
5.21 gFEX rates with pileup subtraction applied, using different ρ calculation methods.	89
5.22 gFEX efficiencies with different methods of calculating the ρ used for pileup subtraction, including median, mean and the mean under a certain threshold. . .	90
5.23 The Yocto Project and OpenEmbedded relationship [171].	91
5.24 The Yocto Project workflow [171]	93
5.25 Workflow for setting up and building the gFEX OS	99
5.26 Monitoring overview, including several different graphical interfaces, and low level on-board hardware monitoring. Image Credit: Cecilia Tosiri.	100
5.27 Diagram of the GPIO which interfaces between the PS (processing system) and PL (programmable logic, FPGA) on the Zynq to provide temperature values to the IPMC.	102
5.28 Diagram of the IPBus protocol implemented in different sized systems [143]. . .	105
5.29 Overview of ironman functionality [186].	106
5.30 Overview of ironman server [186].	106
5.31 Overview of ironman Hardware Interface [186].	107
5.32 Overview of the DCS system for L1Calo. Image from Paul Thompson.	109
5.33 gFEX OPC-UA server class diagram.	110

5.34	High level DCS panels for TDAQ and L1Calo that include links and graphics representing gFEX.	112
5.35	High level DCS panels for the gFEX shelf, gfc0.	113
5.36	gFEX main DCS panel with ATCA datapoints, and gFEX temperature sensor panel.	115
5.37	gFEX power modules DCS panel, and OS utilization panel.	116
5.38	gFEX MiniPOD DCS panel and example of archived data.	117
5.39	gFEX expert panels for ATCA and SoC specific applications.	118
5.40	gFEX early data taken during 2018 in Run 2. Specifically this data was taken during the period 15 Oct at 22h to 16 Oct at 11h UTC. This corresponds to run 363664 with the tag data18_13TeV. Figure Credit: Greg Myers.	120
5.41	TDAQ partition, often referred to as the "GUI", which controls and monitors any hardware included in the run. Also used in the ATLAS control room to run and monitor the whole detector.	122
5.42	Overview of gFEX mapping. Image Credit: Cecilia Tosiri.	124
5.43	gFEX small-radius jet efficiencies compared to legacy jet efficiencies in Run 456016. Plot Credit: Kristin Dona.	125
5.44	gFEX small-radius jet trigger rate versus legacy jet trigger rate in Run 456016. Plot Credit: Kristin Dona.	126
5.45	gFEX small-radius jet trigger rate compared to both jFEX small-radius jet trigger rate, and legacy jet trigger rate in Run 456016. Plot Credit: Kristin Dona.	126
7.1	Diagrams for the targeted electroweakino production and final states [90].	134
7.2	Graphical representation of the mass points included in this search for a variety of $\tilde{\chi}_{\text{heavy}}$ and $\tilde{\chi}_{\text{light}}$ masses. Often referred to as a signal grid.	136
7.3	The efficiency of wino-NLSP/bino-LSP signal samples with the dead tile module veto.	140
7.4	The efficiency of wino-NLSP/bino-LSP signal samples with the non-collision background veto.	141
7.5	The efficiency of wino-NLSP/bino-LSP signal samples with the dead tile jet event veto and the non-collision background veto.	141

7.6	The boson-tagging efficiency for jets arising from $W/Z/h$ bosons decaying into $q\bar{q}$ or $b\bar{b}$ (signal jets) and the rejection factor (inverse of the efficiency) for jets that have other origins (background jets) are shown.	145
7.7	Distributions of key observables for the resolved channel with the preselection applied as described in Table 7.7. Figure continued on next page.	149
7.8	Distributions of key observables for the boosted channel with the preselection applied as described in Table 7.7. Plot Credit: Boosted Channel Team.	151
7.9	Distributions of key observables for the Resolved 2B2Q with the SRHad-Low signal region selection.	156
7.10	Distributions of key observables for the Resolved 2B2Q with the SRHad-High signal region selection.	157
7.11	Original SRHad-LM and SRHad-HM signal region estimated sensitivities.	158
7.12	Final SRHad-Diag, SRHad-LM, and SRHad-HM signal region estimated sensitivities.	160
7.13	Signal region definitions as a function of the masses of the two leading large- R jets. The mass window cuts of the W_{qq}/Z_{qq} -tagging shown in the plot only indicate the typical values, while variable cut values along p_T are applied in the analysis. The inclusive SRs, defined by the logical union of a few mutually overlapping SRs, are indicated by the gray dashed lines. [90]	162
7.14	Kinematic plots for the SR4Q-VV signal region. The overlaid signal is $(\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm) \rightarrow WW + (\tilde{\chi}_1^\pm, \tilde{\chi}_2^0) \rightarrow WZ$ with $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_2^0}) = (800, 100)$ GeV. Plot credit: boosted channel team.	163
7.15	Expected sensitivity for the SR4Q-VV inclusive signal region with the $(\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm) \rightarrow WW + (\tilde{\chi}_1^\pm, \tilde{\chi}_2^0) \rightarrow WZ$ signal grid. Red dotted contours correspond to exclusion ($Z_N = 1.64$) and the solid contour correspond to 3σ sensitivity. Plot Credit: Boosted Channel Team.	164
7.16	Kinematic variables for the 2B2Q boosted Wh signal-region. Compared to signal shown in gray: $(\tilde{\chi}_1^\pm, \tilde{\chi}_2^0) \rightarrow Wh$ with $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_2^0}) = (800, 200)$ GeV.	165
7.17	The expected sensitivity for Wh signal search. Signal and background yields are estimated from simulation. Points inside the solid red curve have at least 3σ sensitivity, and points inside the dashed red curve have at least $Z_N = 1.64$ (exclusion). Previously observed limits are shown in purple.	166
7.18	Estimated sensitivity of the boosted+resolved signal regions.	168

7.19	Schematics illustrating the relation between the signal regions, validation regions, and control regions [90].	170
7.20	Comparison between the observed data and the post-fit SM background prediction in the validation regions. "Top" includes $t\bar{t}$, $t+X$ and $t\bar{t}+X$ production. The total systematic uncertainty of the background prediction is shown by the hatched band. The bottom panel shows the statistical significance of the discrepancy between the observed number of events and the SM expectation [90]. . . .	170
7.21	Summary of the observed data and predicted SM background in all SRs. The total systematic uncertainty in the background prediction is shown by the hatched area. Distributions of a few representative signals are overlaid, with $(m(\tilde{\chi}_1^\pm), m(\tilde{\chi}_1^0)) = (900, 100)$ GeV. The bottom panel shows the statistical significance of the discrepancy between the observed number of events and the SM expectation, following the prescription described in Ref. [97]. [90].	173
7.22	Exclusion limits for the (\tilde{W}, \tilde{B}) and (\tilde{H}, \tilde{B}) models shown as a function of the wino/higgsino chargino mass $m(\tilde{\chi}_1^\pm)$ and the bino LSP mass $m(\tilde{\chi}_1^0)$. Expected (dashed) and observed (solid red) 95% CL exclusion limits are shown for the (\tilde{W}, \tilde{B}) and (\tilde{H}, \tilde{B}) models with a representative branching ratio $\mathcal{B}(\tilde{\chi}_2^0 \rightarrow Z\tilde{\chi}_1^0) = 50\%$. [90]	174
7.23	The boosted channel exclusion limit results compared to several other Run 1 and Run 2 searches for electroweakinos.	175
8.1	ATLAS Lund Jet plane result, with areas of perurbative and hadronization effects labeled. The direction of increasing k_t is also labeled as this illustrates the choice of multiplicity observables.	179
8.2	Depiction of Lund subjet multiplicities, reproduced and modified from Ref. [154]: a schematic Lund jet plane is provided for a jet with one primary and two secondary emissions. The Lund multiplicity of this jet is $N_{\text{Lund}} = 4$, while the Primary Lund multiplicity is $N_{\text{Lund}}^{\text{Primary}} = 2$. The lower edge of the Lund jet plane is truncated, indicating the region of emission k_t that is removed when imposing a k_t requirement.	181
8.3	Validation that p_T spectrum is smooth in PYTHIA 8 and POWHEG +PYTHIA 8 samples.	186
8.4	Validation that p_T spectrum is smooth in SHERPA samples.	187
8.5	Validation that p_T spectrum is smooth in SHERPA 2.2.11 and DIRE samples. . .	188
8.6	Validation that p_T spectrum is smooth in HERWIG 7 samples.	189

8.7	Particle-level jet kinematics.	191
8.8	Detector-level jet kinematics.	193
8.9	Example plot of one multiplicity observable and its associated uncertainties. . .	197
8.10	The rebinned PYTHIA unfolding matrices for Lund multiplicity with various k_t cuts. Figure continued on next page.	202
8.11	The rebinned PYTHIA unfolding matrix diagonal fraction for Lund multiplicity with various k_t cuts. Figure continued on next page.	204
8.12	The rebinned PYTHIA unfolding matrices for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	206
8.13	The rebinned PYTHIA unfolding matrix diagonal fraction for primary Lund multiplicity. Figure continued on next page.	208
8.14	The fake and efficiency factors for Lund multiplicity with various k_t cuts. Figure continued on next page.	211
8.15	The fake and efficiency factors for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	213
8.16	Technical closure cross-check results for the Lund multiplicity. The ratio of the unfolded PYTHIA results with the particle-level PYTHIA distribution is 1, indicating good technical closure. Figure continued on next page.	216
8.17	Technical closure cross-check results for the primary Lund multiplicity. The ratio of the unfolded PYTHIA results with the particle-level PYTHIA distribution is 1, indicating good technical closure. Figure continued on next page.	218
8.18	Data-driven non-closure diagnostic for Lund multiplicity, demonstrating the performance of the reweighting applied to the PYTHIA MC. Figure continued on next page.	221
8.19	Data-driven non-closure diagnostic for primary Lund multiplicity, demonstrating the performance of the reweighting applied to the PYTHIA MC. Figure continued on next page.	223
8.20	Data-driven non-closure uncertainty for Lund multiplicity. Figure continued on next page.	225
8.21	Data-driven non-closure uncertainty for primary Lund multiplicity. Figure continued on next page.	227
8.22	Statistical uncertainties from data & MC for Lund multiplicity. Figure continued on next page.	230

8.23	Statistical uncertainties from data & MC, for primary Lund multiplicity. Figure continued on next page.	232
8.24	Pile-up reweighting uncertainties for Lund multiplicity. Figure continued on next page.	235
8.25	Pile-up reweighting uncertainties for primary Lund multiplicity. Figure continued on next page.	237
8.26	Disabled tile-module uncertainty, for Lund multiplicity. Figure continued on next page.	241
8.27	Disabled tile-module uncertainty for primary Lund multiplicity. Figure continued on next page.	243
8.28	Fractional jet energy scale systematic uncertainty components for anti- k_t $R = 0.4$ jets as a function of jet p_T at $\eta = 0$, reconstructed from particle-flow objects. The total uncertainty, determined as the quadrature sum of all components, is shown as a filled region topped by a solid black line. Figure from public Jet/EtMiss plots: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2022-005/	248
8.29	(a) Relative jet energy resolution and (b) absolute uncertainty in the relative resolution as a function of p_T for PFlow jets in the central region of the detector, measured using the dijet balance method. The resolution in data is shown in black points with error bars indicating statistical uncertainties; the resolution in detector-level simulated events is shown by the blue curve with total systematic uncertainty given by the blue band. The systematic uncertainty is dominated by terms propagated from the JES uncertainty, while additional terms arise from the analysis selection, pile-up rejection (JVT), physics modelling (comparison with alternative generator), and non-closure effects. The bump in uncertainty around 800 GeV comes from the single-particle uncertainty in the earlier in situ jet-energy scale calibration used when making the in situ JER measurement [23].	249
8.30	Uncertainties related to the jet energy scale and resolution (JES/JER) for the Lund multiplicity with various k_t cuts. Figure continued on next page.	250
8.31	Uncertainties related to the jet energy scale and resolution (JES/JER) for the primary Lund multiplicity with various k_t cuts. Figure continued on next page.	252
8.32	Tracking uncertainties for Lund multiplicity for several k_t requirements. Figure continued on next page.	256
8.33	Tracking uncertainties for primary Lund multiplicity for several k_t requirements. Figure continued on next page.	258

8.34	Uncertainty due to the use of an alternative MC model (SHERPA 2.2.11 with re-tuned cluster hadronisation) on Lund multiplicity distributions. Figure continued on next page.	261
8.35	Uncertainty due to the use of an alternative MC model (SHERPA 2.2.11 with re-tuned cluster hadronisation) on primary Lund multiplicity distributions. Figure continued on next page.	263
8.36	Uncertainty due to the use of an alternative MC model (HERWIG 7.1 with angle-ordered parton shower) on Lund multiplicity distributions. Figure continued on next page.	265
8.37	Uncertainty due to the use of an alternative MC model (HERWIG 7.1 with angle-ordered parton shower) on primary Lund multiplicity distributions. Figure continued on next page.	267
8.38	Total uncertainty due to the use of alternative MC models SHERPA 2.2.11 and HERWIG 7.1 on Lund multiplicity distributions. Figure continued on next page. .	269
8.39	Total uncertainty due to the use of alternative MC models SHERPA 2.2.11 and HERWIG 7.1 on primary Lund multiplicity distributions. Figure continued on next page.	271
8.40	The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.	276
8.41	The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.	280
8.42	The (a) $\langle N_{\text{Lund}} \rangle$ and (b) $\langle N_{\text{Lund}}^{\text{Primary}} \rangle$ distributions as a function of the emission k_t requirement, for an inclusive p_T selection ($p_T > 300$ GeV).	285
8.43	The average Lund multiplicity ($\langle N_{\text{Lund}} \rangle$) as a function of the k_t requirement is shown in bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area. Figure continued on next page.	286
8.44	The average primary Lund multiplicity ($\langle N_{\text{Lund}}^{\text{Primary}} \rangle$) as a function of the k_t requirement is shown bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area. Figure continued on next page.	289
C.1	The data-driven non-closure uncertainty for Lund multiplicity with various k_t cuts. Figure continued on next page.	307

C.2	The data-driven non-closure uncertainty for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	309
C.3	The statistical uncertainty from data for Lund multiplicity with various k_t cuts. Figure continued on next page.	311
C.4	The statistical uncertainty from data for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	313
C.5	The statistical uncertainty from MC for Lund multiplicity with various k_t cuts. Figure continued on next page.	315
C.6	The statistical uncertainty from MC for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	317
C.7	The Monte Carlo modelling uncertainty for Lund multiplicities with various k_t cuts. Figure continued on next page.	319
C.8	The Monte Carlo modelling uncertainty for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	321
C.9	The distributions with 2 unfolding iterations for Lund multiplicity with various k_t cuts. Figure continued on next page.	323
C.10	The distributions with 2 unfolding iterations for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	325
C.11	The distributions with 4 unfolding iterations for Lund multiplicity with various k_t cuts. Figure continued on next page.	327
C.12	The distributions with 4 unfolding iterations for primary Lund multiplicity with various k_t cuts. Figure continued on next page.	329
D.1	A selection of multiplicity observables in different p_T and η bins, showing a comparison of data rebinned both before (pre) and after (post) the unfolding procedure.	332
D.2	The Lund multiplicity $k_t > 0.5 \text{ GeV}$ in several p_T and η bins, showing a comparison of data rebinned both before (pre) and after (post) the unfolding procedure.	333
D.3	The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and bottom figure is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page.	334

D.4	The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and on the bottom is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page. . .	338
E.1	(Top) The response matrices and (bottom) purities for $k_t > 1$ GeV, (left) before rebinning, and (right) after rebinning.	343
E.2	The (left) fine-binned and (right) coarse-binned multiplicity distributions for (top) $300 < p_T < 500$ GeV, and (bottom) $1250 < p_T < 4500$ GeV	344
E.3	The (left) fine-binned and (right) coarse-binned multiplicity distributions for (top) $300 < p_T < 500$ GeV, and (bottom) $1250 < p_T < 4500$ GeV where the multiplicity distribution has been fit using the fine-binned distribution.	347
E.4	The (left) fine-binned and (right) coarse-binned multiplicity distributions for (top) $300 < p_T < 500$ GeV, and (bottom) $1250 < p_T < 4500$ GeV where the multiplicity distribution has been fit using the coarse-binned distribution.	348

LIST OF TABLES

5.1	Fiber Control and Status Registers	72
5.2	gFEX to FELIX elinks and their several associated naming schemes. Table Credit: Greg Myers.	75
5.3	Jet Finder Registers	77
5.4	Jet Finder MAIN_CSR Bit Definitions	78
5.5	Relevant Yocto versions compared to Xilinx Vivado versions.	95
5.6	gFEX IPMC temperatue sensor threshold values.	103
5.7	gFEX FIFO registers and relevant bits.	104
5.8	gFEX Board Location	123
7.1	Location of dead tile modules over the course of Run 2. Partially dead modules are not listed here.	139
7.2	Summary of MET triggers used in the analysis.	142
7.3	Summary of single-lepton triggers used in the analysis.	142
7.4	Nomenclature of object multiplicity and the kinematic variables used for common preselection.	146
7.5	Kinematic variables used in the resolved category selection.	147
7.6	Object multiplicity and kinematic variables used in the boosted category selection.	147
7.7	Definition of the common preselection	148
7.8	Starting selection requirements for the signal regions of the resolved 2B2Q channel.	155
7.9	Final selection requirements for the signal regions of the resolved 2B2Q channel.	159
7.10	Definition of each SR in the 4Q and 2B2Q categories, where $n(W_{qq})$, $n(Z_{qq})$, $n(V_{qq})$, $n(Z_{bb})$, and $n(h_{bb})$ are respectively the number of large- R jets passing the W_{qq} -, Z_{qq} -, V_{qq} -, Z_{bb} -, and h_{bb} -tagging of the two leading large- R jets. SR- 4Q-WZ requires $n(W_{qq}), n(Z_{qq}) \geq 1$ instead of $n(W_{qq}) = n(Z_{qq}) = 1$ because the selections in W_{qq} - and Z_{qq} -tagging are not exclusive.	161
7.11	Definition of extra cuts as applied in the boosted 4Q SRs.	162
7.12	Definition of extra cuts as applied in the boosted 2B2Q SRs.	163

8.1	Summary of the single-jet triggers used during each year of Run 2 data-taking. The indicated trigger is used to select events in the corresponding year for use in this analysis.	183
8.2	Summary table of MC generators used in this analysis and their various models.	190
8.3	Summary of disabled tile modules during Run 2 data-taking and in Run 2 simulation. Adapted from https://gitlab.cern.ch/atlas/athena/-/blob/master/Reconstruction/Jet/JetAnalysisTools/JetTileCorrection/data/Tile_maskedDB_Run2.conf	239
A.1	Boolean Health Flags describing the state of the gFEX system	295
A.2	FPGA Monitoring	296
A.3	OS-Utilization Sensors	296
A.4	ADM7417 Temperature Sensors	296
A.5	DCDC Converter Sensors	297
A.6	pFPGA Power Modules	297
A.7	Zynq and Global Power Modules	298
A.8	Zynq & Spare RX and TX MiniPODs	298
A.9	pFPGA A MiniPODs	299
A.10	pFPGA B MiniPODs	300
A.11	pFPGA C MiniPODs	301
B.1	SI5345 Sticky Error Bits	304
B.2	SI5345 Error Bits	304
B.3	Other useful SI5345 registers.	305
E.1	Comparison of the raw averages between the fine-binned and coarse-binned distributions.	345
E.2	Comparison of the raw averages between the fine-binned distribution and the averages from the fit to the coarse-binned distribution.	349

ACKNOWLEDGMENTS

A PhD can be a long and winding path and there have been many people along the way who have offered their support and guidance throughout the process. First and foremost, to my advisor David Miller, thank you for the opportunity to work with you on the wild and exciting ride that is gFEX, and many thanks for your support and guidance over the last six (six!?) years through gFEX and many other projects.

Many thanks also to the UChicago ATLAS group: Mel Shochet, Mark Oreglia, Young-Kee Kim, Cecilia Tosciri, Ben Rosser, Lacey Rainbolt, Jan Offerman, and Kristin Dona. This group has been a wonderful environment in which to broaden my knowledge about physics and also my ability to explain my work. Your comments and questions over the years have been invaluable to my growth as a physicist.

To Matt LeBlanc and Jennifer Roloff, thank you for a remarkable introduction into the intriguing world of jet substructure. Working with you both has been such a fantastic experience and I look forward to the possibility of continued collaboration.

To everyone working on gFEX, many many thanks for all the years of collaboration. Thank you to Simone Sottocornola and Kristin Dona for swooping in to take over gFEX work when I had to concentrate on other things (among them the document before you). Thank you also to the project leaders and managers David Miller, Sabine Lammers, and Michael Begel. Many thanks to Giordon Stark both for your expert help and introduction to the gFEX world, and for your friendship. Finally, and especially to Greg Myers, my collaborator in the insanity that was hardware work at CERN during COVID, many thanks for your vast knowledge and skill, but also for your demeanor, positive outlook, and your friendship over the last several years.

To my UChicago cohort thank you for your support and the collaborative approach to the

world of physics, especially during the first year of graduate school. Thanks to the brunch crew for many enjoyable Sunday mornings. To Claire Baum and Nina Coyle, thank you for everything. Your friendship and support has been unmatched over the years, even after my departure to Geneva.

I have been lucky enough to have many colleagues and friends at CERN since my move to Geneva in 2020 and to all of you: thank you! Especially to my office mate and my dear friend Kristin Dona, your company, your support (and your impeccable taste in decorating) has made the CERN office a pleasure to work in. To my unofficial office mate Meg Duell, I have so appreciated your support and friendship over the last several years. To Mason, thank you for everything.

To my extended family, thank you for the many years of love and support. And of course to my parents, thank you for your encouragement and love not only for the last six years but always. Your belief in me has never wavered and you have never had any doubts. I'm so lucky to have parents like you.

It is the end, but it is also the beginning again.

ABSTRACT

This thesis presents substantial work done on the ATLAS calorimeter hardware trigger upgrades and the results of two analyses. These all show the importance of understanding the hadronic final state in ATLAS at a global level. This is relevant from an experimental perspective which contributes to our ability to take data, and also helps to enable a thorough understanding of the Standard Model and searches of physics beyond the Standard Model.

My work on the ATLAS hardware trigger encompasses significant contributions to the commissioning of the new module the global Feature Extractor (gFEX). This includes the full development of the gFEX Detector Control Systems, contributions to the custom operating system used in gFEX, and additional significant contributions to custom software written to interface with the gFEX firmware. The gFEX is now installed in ATLAS and triggering on real physics data in the ongoing Run 3.

The two analyses documented in this thesis were performed using 139 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ data taken with the ATLAS detector during Run 2. A search for supersymmetric partners of the Standard Model electroweak gauge bosons is presented first, and shows the significant gains in sensitivity possible by utilizing the high branching fractions of the fully hadronic final state. The precision measurement of Lund multiplicities in dijet events can help to constrain Monte Carlo models, and highlight important higher order effects missing from our current Monte Carlo models.

CHAPTER 1

INTRODUCTION

The Standard Model of particle physics at first glance appears to fully describe the structure and dynamics of the fundamental particle universe, with the discovery of the Higgs boson in 2012 the final piece of the puzzle. However, there remain many questions that the Standard Model cannot answer, and several more related to phenomena that it does not describe. In addition to these questions, there are many aspects of the Standard Model itself such as Quantum Chromodynamics (QCD), where our ability to precisely characterize dynamics and to perform accurate and precise empirical tests is still limited, and we must instead rely heavily on models and simulation to describe important features.

Particle colliders have historically played an integral role in the experimental discovery of fundamental particles, culminating in the Large Hadron Collider (LHC) and the Higgs boson discovery. The LHC has successfully taken two full runs of data and is in the midst of a third. The LHC is a hadron collider and as such quarks and gluons are extremely common. Due to QCD confinement these quarks and gluons become collimated sprays of particles, or jets, in our detectors. Again, as the LHC is a hadron collider jets are basically inescapable in every aspect of LHC physics. The detector must be prepared to measure jets such that we can reconstruct them, and the trigger which filters events must be able to reconstruct on the order of nanoseconds interesting physics objects like jets and recognize which of these objects are worth saving.

Jets themselves are complicated to reconstruct and difficult to model. They are strongly coupled, non-perturbative systems which can be intensely interesting in their own right, but which also have a global impact on LHC physics due to the hadronic nature of the collider, and can be powerful tools in other types of physics analyses. In the presented search for

Supersymmetric electroweakinos, the highest exclusionary mass limits were placed by the fully hadronic analysis. Here, jets were invaluable in the search for physics beyond the Standard Model. Precision measurements of jets can also produce very interesting results which can help deepen our understanding of QCD, but the application of these results can also extend our ability to model and understand jets in other contexts, such as searches for beyond the Standard Model physics.

This thesis presents a global look at the relevance of jets from their impact on hardware trigger design, to the fully hadronic search for Supersymmetric electroweakinos and the precision measurement of jet substructure. To start, an overview of the Standard Model is given in Chapter 2. Chapter 3 and Chapter 4 provide an experimental overview of the LHC and the ATLAS detector. A detailed look at the commissioning process of the global Feature Extractor is given in Chapter 5, and jets themselves and details of their simulation and reconstruction are discussed in Chapter 6. Chapter 7 summarizes the fully hadronic search for Supersymmetric partners of the Standard Model gauge bosons, and Chapter 8 describes a measurement of subjet multiplicities and the impacts this could have on jet modeling. Concluding remarks on the impacts of the work presented and its potential in the future are provided in Chapter 9.

THE STANDARD MODEL

[illegible]

1. gravity would become non-negligible at energies around the Plank scale, or roughly 10^{19} GeV.

2.1 Introduction to the Standard Model

Fundamental particles as described by the Standard Model are point-like objects with certain characteristics (mass, spin, charge, etc) which distinguish them. The formalism for describing these particles and their interactions is Quantum Field Theory (QFT), where particles are described by quantum fields which respect certain symmetries. The types of particles in the Standard Model can be grouped into *fermions* which are constituents of matter and interact via the three forces listed above and *bosons* which mediate these forces. The photon (γ) is the electromagnetic mediator, the W^\pm and Z bosons are the charged and neutral mediators of the weak force, and the gluon (g) is the strong force mediator. The Higgs Boson is a unique particle which is the mediator of the bosonic field which gives all other particles mass. These particles and their characteristics are summarized in Figure 2.2.

mass →	$\approx 2.3 \text{ MeV}/c^2$	$\approx 1.275 \text{ GeV}/c^2$	$\approx 173.07 \text{ GeV}/c^2$	0	$\approx 126 \text{ GeV}/c^2$
charge →	2/3	2/3	2/3	0	0
spin →	1/2	1/2	1/2	1	0
	u up	c charm	t top	g gluon	H Higgs boson
QUARKS	$\approx 4.8 \text{ MeV}/c^2$	$\approx 95 \text{ MeV}/c^2$	$\approx 4.18 \text{ GeV}/c^2$	0	
	-1/3	-1/3	-1/3	0	
	1/2	1/2	1/2	1	
	d down	s strange	b bottom	γ photon	
	$0.511 \text{ MeV}/c^2$	$105.7 \text{ MeV}/c^2$	$1.777 \text{ GeV}/c^2$	$91.2 \text{ GeV}/c^2$	
	-1	-1	-1	0	
	1/2	1/2	1/2	1	
	e electron	μ muon	τ tau	Z Z boson	
LEPTONS	$< 2.2 \text{ eV}/c^2$	$< 0.17 \text{ MeV}/c^2$	$< 15.5 \text{ MeV}/c^2$	$80.4 \text{ GeV}/c^2$	
	0	0	0	± 1	
	1/2	1/2	1/2	1	
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	W W boson	
					GAUGE BOSONS

Figure 2.2: A summary of Standard Model particles and their properties. Image from Wikipedia.

Mathematically the Standard Model is a gauge theory which is described by $SU(3)_c \times$

$SU(2)_L \times U(1)_Y$ gauge symmetry and contains 26 free parameters which must be determined through experimental measurements. These symmetries may be split into different sectors for simplification. The electroweak sector is associated with the $SU(2)_L \times U(1)_Y$ symmetry and is described by the Glashow-Weinberg-Salam Electroweak Theory. Quantum Chromodynamics (QCD) describes the dynamics of the strong force and is associated with the $SU(3)_c$ symmetry group. A Lagrangian density which describes the dynamics of these theories can be developed based on the symmetry group, and combined into the overall Lagrangian for the Standard Model as follows:

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{Higgs} + \mathcal{L}_{QCD} \quad (2.1)$$

2.2 The Electroweak Sector: Quantum Electrodynamics

The electroweak Lagrangian can be broken into the following components for additional simplicity:

$$\mathcal{L}_{EW} = \mathcal{L}_{fermions} + \mathcal{L}_{gauge} \quad (2.2)$$

The gauge symmetries $SU(2)_L \times U(1)_Y$ can also be separated into weak isospin ($SU(2)_L$) and weak hypercharge ($U(1)_Y$) respectively.

The simplest starting point to examine the structure of a fermion under some symmetry is a fermion ψ which transforms under $U(1)$ symmetry with the fundamental representation:

$$\psi(x) \rightarrow e^{i\alpha(x)}\psi(x) \quad (2.3)$$

To preserve invariance under the $U(1)$ transformation, one must define a gauge field $A_\mu(x)$:

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{g} \partial_\mu \alpha(x) \quad (2.4)$$

A covariant derivative on ψ can then be defined as:

$$D_\mu \psi = (\partial_\mu + ig A_\mu(x)) \psi \quad (2.5)$$

The field strength tensor for this gauge field can be defined as:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.6)$$

For non-abelian groups (which all the SM groups are), the generators of the group must be introduced into the covariant derivative in place of $\alpha(x)$ such that they satisfy $[t^a, t^b] = t^a t^b - t^b t^a = if^{abc} t^c$. This results in the following covariant derivative and field strength tensor:

$$D_\mu \psi = (\partial_\mu + igt^a A_\mu^a(x)) \psi \quad (2.7)$$

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - gf^{abc} A_\mu^b A_\nu^c \quad (2.8)$$

Now, let us shift from the general $U(1)$ symmetry discussed above to the more specific $SU(2)_L \times U(1)_Y$ of the Electroweak sector. For the $SU(2)_L$ symmetry there are $2 \times 2 - 1 = 3$ gauge fields written as $W_\mu^{a=1,2,3}$. For the $U(1)_Y$ symmetry there is one gauge field B_μ . The field tensors of these respective fields are:

$$B_\mu = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.9)$$

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - gW_\mu \times W_\nu \quad (2.10)$$

where g is the $SU(2)_L$ coupling constant. This results in the Lagrangian:

$$\mathcal{L}_{gauge} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu} \cdot W^{\mu\nu} \quad (2.11)$$

To formulate the fermionic component of the electroweak sector, one can start with the Dirac equation, which describes relativistic spin-1/2 particles:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad (2.12)$$

where γ^μ are the Dirac gamma matrices. A conjugate spinor can also be defined as $\bar{\psi} = \psi^\dagger \gamma^0$.

To express these dynamics as a relativistic (ie Lorentz invariant) Lagrangian, one can write:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (2.13)$$

Using the transformation $i\partial_\mu \rightarrow i\partial_\mu - gA_\mu$ and defining a covariant derivative as in Equation 2.5 one recovers a Lagrangian which is defined as follows:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \quad (2.14)$$

All that is needed is invariance under $SU(2)_L \times U(1)_Y$ transformations which can be done in a similar way to the example explored above with $U(1)$ symmetry, with the relevant gauge fields for the full electroweak sector, $B_{\mu\nu}$ and $W_{\mu\nu}$.

The full fermionic Lagrangian for Quantum Electrodynamics requires combining these dif-

ferent areas explored above, and results in:

$$\mathcal{L}_{fermion} = \bar{\psi} \gamma^\mu D_\mu \psi \quad (2.15)$$

$$D_\mu = \partial_\mu - ig' \frac{Y}{2} B_\mu - i \frac{g}{2} (\tau \cdot W_\mu) \quad (2.16)$$

where τ are the Pauli matrices, g' is the $U(1)_Y$ coupling constant, and g is again the $SU(2)_L$ coupling constant.

The full electroweak Lagrangian is then:

$$\mathcal{L}_{fermion} = \bar{\psi} \gamma^\mu D_\mu \psi - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu} \cdot W^{\mu\nu} \quad (2.17)$$

This is clearly lacking in some areas, particularly a fermion mass term. Experimentally, fermions have mass so there must be a mechanism which provides it, namely the Higgs mechanism.

2.3 Electroweak Symmetry Breaking and The Higgs Mechanism

The gauge structure of the SM is complex enough that one cannot simply write down mass terms for fermions. Since left- and right-handed fermions are $SU(2)$ doublets and singlets respectively, they will transform differently under the $SU(2)_L$ symmetry of the Standard Model. Experimentally, it has been observed that fermions and bosons have mass, so we are left with a puzzle: the fermions must have mass, but how can these terms be written? The solution to this puzzle is a Standard Model which obeys the underlying $SU(2)$ gauge symmetry and a specific choice of vacuum which causes this symmetry to break. This is referred to as *spontaneous symmetry breaking* and is the Higgs mechanism: the mechanism

by which all fermions and bosons obtain mass.

First the Higgs field can be introduced as a double of complex scalar fields with hypercharge $Y = \frac{1}{2}$:

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} \quad (2.18)$$

The kinetic and potential terms of the Higgs Lagrangian are then given by:

$$\mathcal{L}_{higgs,bosons} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(x) \quad (2.19)$$

where the Higgs potential is defined as:

$$V(x) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (2.20)$$

and the covariant derivative is:

$$D_\mu = \partial_\mu - i\frac{g}{2}\tau^a W_\mu^a - i\frac{g'}{2}B_\mu \quad (2.21)$$

The Higgs potential is determined by μ and λ . If $\lambda > 0$ then the potential will posses a stable ground state. If $\mu^2 < 0$, the Higgs potential will have a minimum at the *vacuum expectation value (vev)* of:

$$v = \sqrt{\frac{-\mu^2}{\lambda}} \quad (2.22)$$

Putting this all together, and with a choice of gauge that simplifies the result, gives:

$$\phi(x) = \begin{pmatrix} 0 \\ \frac{v+H}{\sqrt{2}} \end{pmatrix} \quad (2.23)$$

Adding this into the Higgs kinetic term in the Lagrangian gives:

$$(D^\mu \phi)^\dagger (D_\mu \phi) = |(\partial_\mu - i\frac{g}{2}\tau^a W_\mu^a - i\frac{g'}{2}B_\mu)\phi|^2 \quad (2.24)$$

After expanding the above equation with the ϕ doublet, it can be seen how the familiar gauge bosons appear. The W^\pm boson, Z boson, and photon (A boson) are defined as linear combinations of the W_μ and B_μ gauge fields:

$$W_\mu^\pm = W_\mu^1 \mp iW_\mu^2 \quad (2.25)$$

$$Z_\mu = \frac{1}{\sqrt{(g^2 + (g')^2)}}(gW_\mu^2 - g'B_\mu) \quad (2.26)$$

$$A_\mu = \frac{1}{\sqrt{(g^2 + (g')^2)}}(g'W_\mu^2 - gB_\mu) \quad (2.27)$$

If Equation 2.24 is expanded, it can be clearly seen that the Higgs term H , is only coupled to the W^\pm and Z bosons, and not the A_μ photon term. Thus the photon remains massless and the W and Z bosons acquire the following masses:

$$m_W = \frac{v}{2}g \quad (2.28)$$

$$m_Z = \frac{v}{2}\sqrt{g^2 + (g')^2} \quad (2.29)$$

Fermion masses are also generated via the Higgs through the introduction of Yukawa interactions to the Lagrangian, which take the form:

$$\mathcal{L}_{higgs,fermions} = -\lambda_f(\bar{f}_L \phi f_R + \bar{f}_R \phi^\dagger f_L) \quad (2.30)$$

where λ_f is the Yukawa coupling of the particular fermion, and f_L , f_R , and ϕ are the fermion left-handed doublet, right-handed singlet, and Higgs scalar doublet. After inserting the Higgs doublet, this becomes:

$$\mathcal{L}_{higgs,fermions} = -\frac{\lambda_f v}{\sqrt{2}}(f\bar{f} + 2Hf\bar{f}) \quad (2.31)$$

Thus the fermionic masses can be defined as :

$$m_f = -\frac{\lambda_f v}{\sqrt{2}} \quad (2.32)$$

2.4 Quantum Chromodynamics: Life in Color

The final $SU(3)_C$ gauge symmetry corresponds to the strong force, or Quantum Chromodynamics (QCD). It is arranged similarly to QED in the sense that fermions interact through dynamics mediated by force carriers. For QCD the interacting fermions are quarks, and the force carrier is the gluon. Based on the specific type of gauge symmetry, it is expected to have $3 \times 3 - 1 = 8$ gauge bosons. Just as the QED symmetries are associated with isospin and hypercharge, QCD has a corresponding charge with three possible values, referred to as color. The possible color charges are r, g, b and the anti-color charges are $\bar{r}, \bar{g}, \bar{b}$. Quarks have color, anti-quarks have anti-color, and gluons carry both a color and anti-color charge. The possible types of gluons are then: $r\bar{g}, r\bar{b}, g\bar{r}, g\bar{b}, b\bar{r}, b\bar{g}, r\bar{r}, g\bar{g}, b\bar{b}$. However the last three result in a color-neutral particle and are functionally the same. The color singlet state is a mix of these three options: $\frac{1}{\sqrt{3}}(r\bar{r} + g\bar{g} + b\bar{b})$ and does not represent a physical gluon observed in nature. The remaining two gauge bosons of QCD are the non-colorless linear combinations of these three states: $\frac{1}{\sqrt{2}}(r\bar{r} - g\bar{g})$ and $\frac{1}{\sqrt{6}}(r\bar{r} + g\bar{g} - 2b\bar{b})$.

The tensor for these eight gluons associated with the $SU(3)_C$ gauge symmetry of QCD is:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c \quad (2.33)$$

where the index a runs over the 8 QCD gluons, and g_s and f_{abc} are the $SU(3)_C$ coupling and structure constants.

The QCD Lagrangian then reads:

$$\mathcal{L}_{QCD} = -\frac{1}{4} G_{\mu\nu}^a G^{a\mu\nu} + \bar{\psi} i \gamma^\mu D_\mu \psi \quad (2.34)$$

where the covariant derivative is:

$$D_\mu = \partial_\mu + i g_s \frac{\lambda_a}{2} G_\mu^a \quad (2.35)$$

The behavior of the QCD coupling constant $g_s(\mu)$ or equivalently (and perhaps more familiarly) $\alpha_s(\mu)$ is integral to the dynamics of QCD. This coupling constant is dependant upon the energy scale, μ , and runs with this energy:

$$\alpha_s(\mu) = \frac{4\pi}{\beta_0 \ln(\mu^2/\Lambda_{QCD})} \quad (2.36)$$

where β_0 and Λ_{QCD} are experimentally determined constants. The running of the strong coupling constant has been measured experimentally and the results can be seen in Figure 2.3. The inverse logarithm dependence on the scale as shown above causes α_s to be to be large at low energies, and $\ll 1$ at high energies. This means that QCD itself is strongly coupled at low energies and weakly coupled at high energies. α_s has been measured by many experiments and is often done so using the mass of the Z boson as a reference scale, as seen in Figure 2.4.

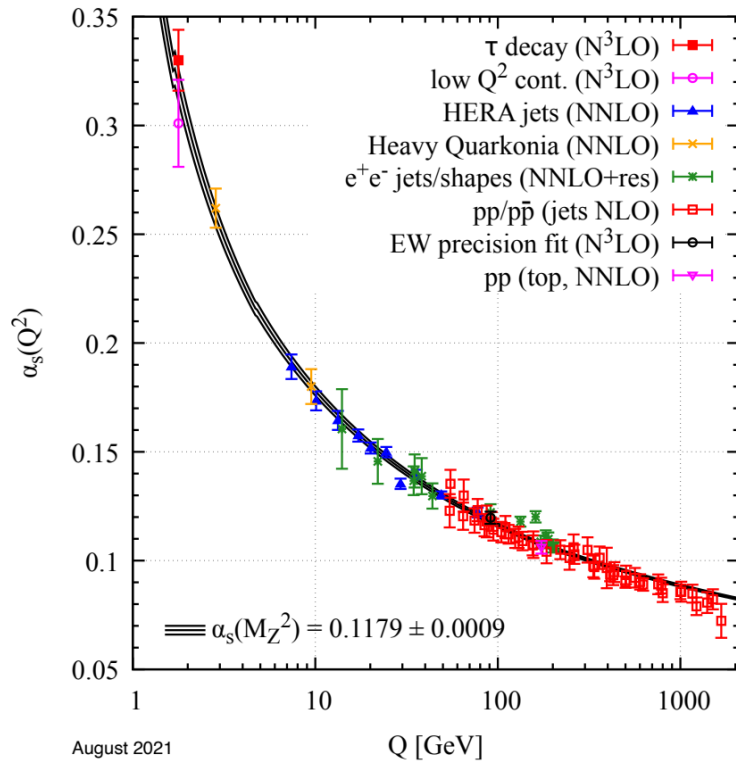
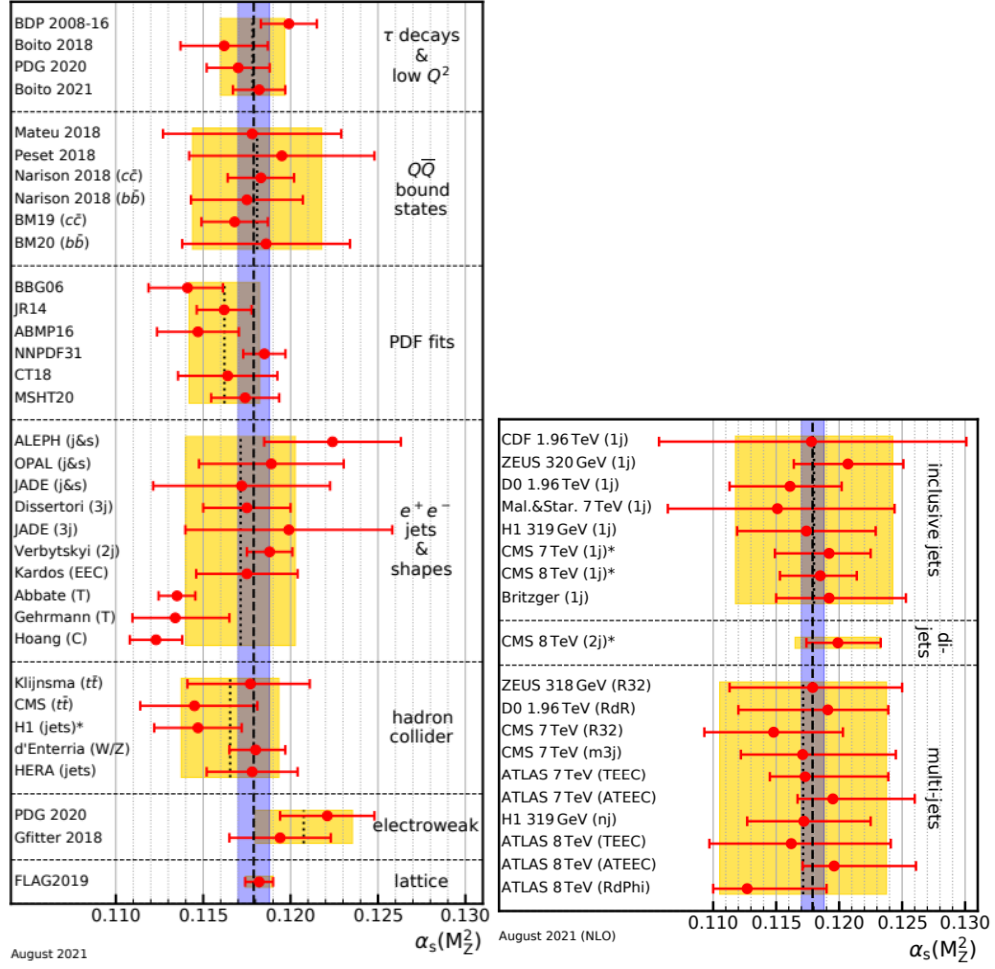


Figure 2.3: Running of α_s as a function of the energy scale [197].



(a) $\alpha_s(M_z^2)$ measured in several sub-fields. (b) $\alpha_s(M_z^2)$ measured at hadron colliders.

Figure 2.4: Measurements of the strong coupling constant α_s at different experiments (a) and at hadron colliders (b) [197].

A consequence of the structure of the strong coupling constant α_s is what is known as *asymptotic freedom*: at large energy scale (and equivalently small distance scales) the strong force becomes negligible and can be well described by perturbation theory. However at small energy scales (and large distance scales) the strong force approaches infinity rendering it extremely nonperturbative. *Color confinement* is a direct result of this behavior, and affects quarks and gluons. As the distance between two particles carrying a color charge increases, so does the strength of the QCD force between them. Eventually this binding energy hits or passes the threshold for quark-antiquark pair production and new quark and gluon pairs are created. These new particles couple to each other and the original quarks and gluons to produce color-singlet hadrons in what is known as *hadronization*. In this scenario, each of the resulting hadrons retains some percentage of the original particle's momentum. These hadrons form the **jets** of particles that are found in detectors like ATLAS at the LHC, and are the experimentally observable result of QCD interactions between quarks and gluons in collided protons at the LHC.

Another relevant consequence of the confinement is the structure of the proton, which is particularly relevant for proton collider experiments. Protons are often considered to be composed of three quarks, but in reality these "valence" quarks account for only about 1% of the proton mass. More accurately, protons consist of a cloud of constantly changing, constantly interacting quarks, antiquarks (sometimes referred to as *sea* quarks), and gluons. The probability of a hard-scatter process with constituent quarks and gluons must be described by a set of *parton distribution functions* (PDFs) which describe the probability of interacting with a specific particle inside a proton, characterized by the momentum fraction of the particle, at the energy scale of the proton (Q^2).

2.5 Beyond the Standard Model

The Standard Model is a complete theory of particle physics, but as hinted at the beginning of this chapter, there are obvious pieces missing from the Standard Model's description of the universe. This theory is incredibly precise, but there are things it neglects or things that don't quite fit which hint at the possibility of new physics beyond the Standard Model.

One of the biggest missing pieces is a description of quantum gravity. As discussed earlier this is most relevant at the plank scale which remains above the current experimental reach, but nevertheless gravity is clearly missing from the Standard Model. All attempts thus far to include quantum gravity have been unsuccessful, though the unification of QFT and gravity remains a goal in the theory community.

In addition to the lack of gravity, the experimental evidence of the existence of Dark Matter has provided a concrete experimental result which points to another piece missing from the Standard Model. Based on Astrophysical observational results of galaxy rotation speed, there must be matter in addition to ordinary visible matter that is affecting these galaxies through gravitational forces [167, 177, 173, 7, 127]. The resulting conclusion is that only 5% of the matter in the universe is ordinary matter, and a large component of the remaining percentage is Dark Matter, as shown in Figure 2.5.

A Dark Matter candidate particle must satisfy the experimental observations thus far. It must be stable on long timescales and be massive such that it can contribute to the matter density of the universe [50]. Additionally it must be weakly interacting via the strong, and electromagnetic forces. Neutrinos and black holes have both been proposed as Dark Matter candidates, but neither have the correct collection of qualities. Thus it is widely agreed that new physics is needed to explain Dark Matter, with a number of models having been proposed over the last several years. An example of one possible model which provides a

Energy Content of the Universe

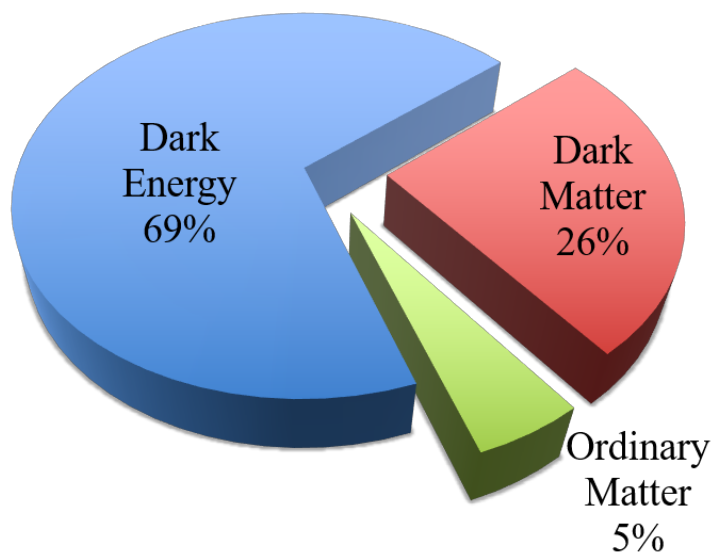


Figure 2.5: Summary of the ordinary matter vs dark matter in the universe, as observed by Planck 2015 CMB measurement. Image by David Miller [167].

Dark Matter candidate is discussed in Subsection 2.5.1.

The matter/anti-matter asymmetry of the universe is the last remaining external factor relating to the limitations of the Standard Model which will be discussed here. The universe we see around us is almost entirely made up of matter, with very little anti-matter to be found. There are existing theoretical explanations of how this imbalance could occur such as electroweak baryogenesis [159], leptogenesis [102], and axiogenesis [84]. None of these quite fit within the Standard Model, and would require new physics to explain the matter/anti-matter asymmetry observed.

Within the Standard Model itself there remain a number of areas of tension between the theory itself and experimental results, and even just questions about why certain things in the Standard Model exist in the precise arrangement that is seen.

Neutrinos are a good example of internal tension to the Standard Model. There is no SM

mechanism for neutrino masses, however in the theory describing neutrino oscillations ² it can be seen that these oscillations vanish for massless neutrinos. Neutrino oscillation has been observed several times with many different experiments, and shows experimentally that neutrinos must have masses, though they are expected to be very small. This discrepancy is yet another missing link in the Standard Model, though there exist theoretical solutions to this issue which have not yet been experimentally verified. The size of the neutrino mass and the difference between the neutrino mass and that of the other Standard Model particles is an additional puzzle.

The other oft-cited issue with the Standard Model is *the hierarchy problem* (or perhaps problems). This is sometimes used to refer to the issue of very small neutrino masses relative to other Standard Model particles. It is also used when describing the difference in masses between the quark generations, particularly the large mass held by the top quark relative to other quark masses. However, most often this phrase refers to the difference in electroweak scale vs gravitational scale ($\Lambda_{ewk}/\Lambda_{Planck} \simeq \mathcal{O}(10^{17})!$).

The masses of the fundamental scalar particles are sensitive to the relevant scale (Λ_{ewk}) here, and this is often discussed in terms of the Higgs boson mass. The mass of the Higgs boson receives higher-order corrections in perturbation theory which depend on Λ^2 , thus making the Higgs mass sensitive to the scale Λ , and without any cancellation or protections, roughly equivalent to the electroweak scale, for example. The possibilities for seeing the low Higgs mass of $m_h = 125$ GeV that have been experimentally observed are twofold: the Higgs mass could just be fine-tuned to the result seen through happenstance, or there could be an intermediate scale contributing to the mass. The possibility of particles at an intermediate scale is an interesting one, as it means any new bosons would contribute higher-order terms of the opposite sign which would cancel nicely the Standard Model terms in the calculation

2. Neutrino theory is not discussed in detail in this thesis, for a nice overview, please see Ref [149], which is technically a book chapter, but stands alone nicely as well.

of the Higgs mass. If the Yukawa couplings for new fermions remained the same as those of the SM fermions, the same cancellation effect would occur.

This argument is often referred to as *naturalness* and grasps at the hope and belief that the world around us is not just a consequence of the happenstance of mathematics, but has some underlying structural organization, a description of which can be codified into theory. Whether this remains a futile hope³ is yet to be seen.

2.5.1 Supersymmetry

One class of theories which extend the Standard Model and have been quite popular over the years is Supersymmetry (SUSY) [124, 191, 195, 194, 115, 175]. In the simplest models, new particles analogous to the existing Standard Model particles are introduced. These "superpartners" have the same quantum numbers as their SM compatriots, but have spins which differ by 1/2, meaning that fermions have bosonic superpartners, and bosons have fermionic superpartners. One of the nicest effects of SUSY is that the hierarchy and naturalness issues previously discussed cease to exist in the SUSY world. Contributions from the new scalar particles add terms in the Higgs mass which exactly cancel those of the original SM particles, to result in the small Higgs mass seen experimentally.

Of course, with the assumption that the superpartners have the same mass as the SM particles, one can raise the question of why these partners haven't been discovered experimentally. Logically, it can be concluded that they must have higher masses: but does this fit within the SUSY theoretical models? It certainly can, if Supersymmetry is taken to be a "softly" broken symmetry. That is to say, it is broken in a specific way which avoids the addition of

3. Hope is the thing with feathers
That perches in the soul
And sings the tune without the words
And never stops - at all [103]

divergent terms⁴. However, if the superpartner masses become too high, the terms relating to them may become larger than the original terms we wished to cancel; the sweet spot lies around the TeV scale, which makes it possible that we simply haven't yet reached high enough energies at the LHC to find these particles experimentally.

Certain SUSY models conserve a quantity called R-parity [112] which is +1 for SM particles and -1 for SUSY particles. This requires that SUSY particles must be produced in pairs and also ensures that the Lightest Supersymmetric Particle (LSP) is stable. In this scenario the LSP provides a good Dark Matter candidate⁵ [123, 108].

The Minimal Supersymmetric Standard Model (MSSM) [114, 113] is the smallest extension that can be made to the Standard Model to accommodate SUSY. In this model the bosonic superpartners have *-ino* added to the end of their name such as the bino (\tilde{B}), wino (\tilde{W}) and Higgsino (\tilde{H}). These are the superpartners of the $U(1)_Y$, $SU(2)_L$, and Higgs fields, and are collectively referred to as electroweakinos. The fermionic content of the SM receives spin-0 partners with the prefix *s-* added to their names⁶. Additionally, two complex Higgs doublets are needed in the MSSM to cancel chiral anomalies. The unconstrained MSSM has 105 degrees of freedom, but this can be constrained if certain assumptions are imposed to prevent CP-violating and flavor-changing interactions. This reduces the available degrees of freedom down to 22 and make the MSSM a nice model to constrain experimental searches for Supersymmetry particles, like those that occur at the Large Hadron Collider.

4. one may again think back to the concept of *naturalness* here and wonder yet again if this qualifies as natural or not

5. Specifically in the form of a weakly interacting massive particle (WIMP).

6. We all sound a little mad when we start discussing the sleptons like sneutrinos, stop and sup squarks.

CHAPTER 3

THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) [110, 72, 73, 68] is, as of 2023, the largest particle collider; the largest machine in the world. It sits at the European Center for Nuclear Research (CERN) about 100 m under the border of France and Switzerland (surely a tumultuous construction project) and is 27 km in circumference. The LHC accelerates hadrons, mostly protons, up to near the speed of light and collides them at a rate of roughly 40 million collisions per second. These collisions occur at 4 specified interaction points corresponding to the detectors ATLAS, CMS, LHCb, and ALICE which are arranged around the LHC ring as can be seen in Figure 3.1.

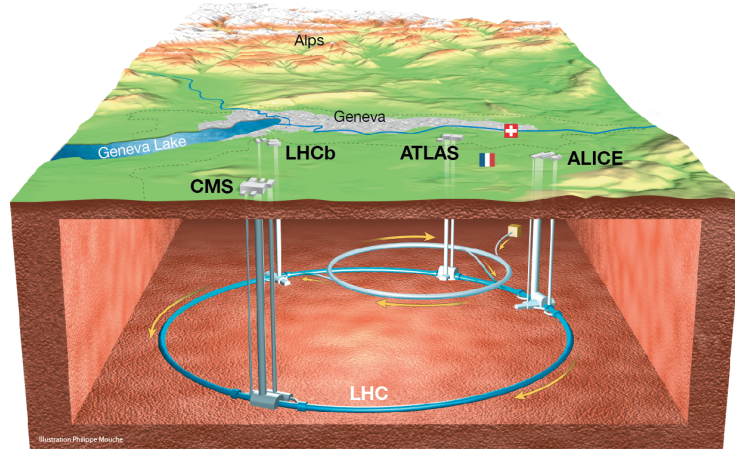


Figure 3.1: A diagram of the LHC showing the four experiments ATLAS, CMS, LHCb, and ALICE. Also shown is the Super Proton Synchrotron which is the penultimate stage of the injector chain before the beam enters the LHC [161].

3.1 Motivation

Despite the precision of the Standard Model there remain many questions that are as yet unanswered, several of which were discussed in Section 2.5. Many of these questions can be investigated from several directions experimentally, but high energy particle colliders provide an environment that cannot be replicated elsewhere. Particularly with hadron colliders the wealth of new fundamental particles that have been discovered speaks for itself; the Tevatron at Fermilab and now the LHC at CERN have discovered such particles as the top quark and the Higgs boson. In the early days, physicists didn't know what they were looking for, only that there were missing pieces and that particle colliders were providing a wealth of information and functioning at an energy scale which produced the very particles we didn't even know we were looking for. The LHC was originally devised partially to find the Higgs boson, and this was accomplished in 2012 [36, 82]. We stand now again on the precipice of knowing that there is something to find, but not quite sure where to look, though there are many hints. I am confident that colliders are one of the many places that must be explored to answer our remaining questions about particle physics.

It is striking that the largest machine in the world is required to study the smallest known particles in the universe. This is ultimately due to the tenet of $E = mc^2$ which constrains the particles produced according to the input energy of the collision. The machinery to manufacture and precisely direct this type of energy is no small feat of engineering, and took several years to originally construct, not to mention the many years dedicated to upgrading the accelerator and its associated detectors.

3.2 Machine Overview

The acceleration and collision of protons to the speed and precision indicated above requires an entire complex of accelerators and specialized magnets and hardware components to control the particle beam. These specialized hardware components include radio-frequency cavities which perform the actual acceleration on individual protons, but also the full cooling infrastructure which is needed to keep the magnets operating at a superconducting level and the power infrastructure which keeps the entire system running. The full injection chain starts with the Linac2 (Linac4 after 2020) which accelerates protons to 50 MeV. The protons are then injected into the Proton Synchrotron Booster (PSB), accelerated to 1.4 GeV and sent to the Proton Synchrotron (PS) where they are pushed to 25 GeV. The protons are then transmitted to the Super Proton Synchrotron (SPS) and accelerated to 450 GeV before finally arriving at the LHC where each proton beam can be accelerated up to 6.5 TeV. Post-2020 the acceleration steps in the injection chain are Linac4 (160 MeV) \rightarrow PSB (2 GeV) \rightarrow PS (26 GeV) \rightarrow SPS (450 GeV) [110]. A diagram of the full injection chain can be seen in Figure 3.2.

The LHC itself is responsible for a large proportion of the beam acceleration (SPS: 0.45 GeV \rightarrow LHC: 6.5 TeV). This is achieved through the use of 16 superconducting radio-frequency (RF) cavities. These metallic cavities contain a precisely oscillating electric field which produces a force upon particles in the cavity, and this force changes direction at a set frequency [80]. Particles with the correct energy and timing are not affected, but particles with slightly different energies will be accelerated or decelerated naturally as needed.

To keep the proton beam on the circular LHC path, dipole magnets are used to bend the trajectory of the beam. These dipoles have an extremely powerful magnetic field of 8.3 Tesla and are superconducting electromagnets using 11,080 Amperes of current to generate a tight enough turn [81]. There are also higher order magnets which correct for edge effects of the

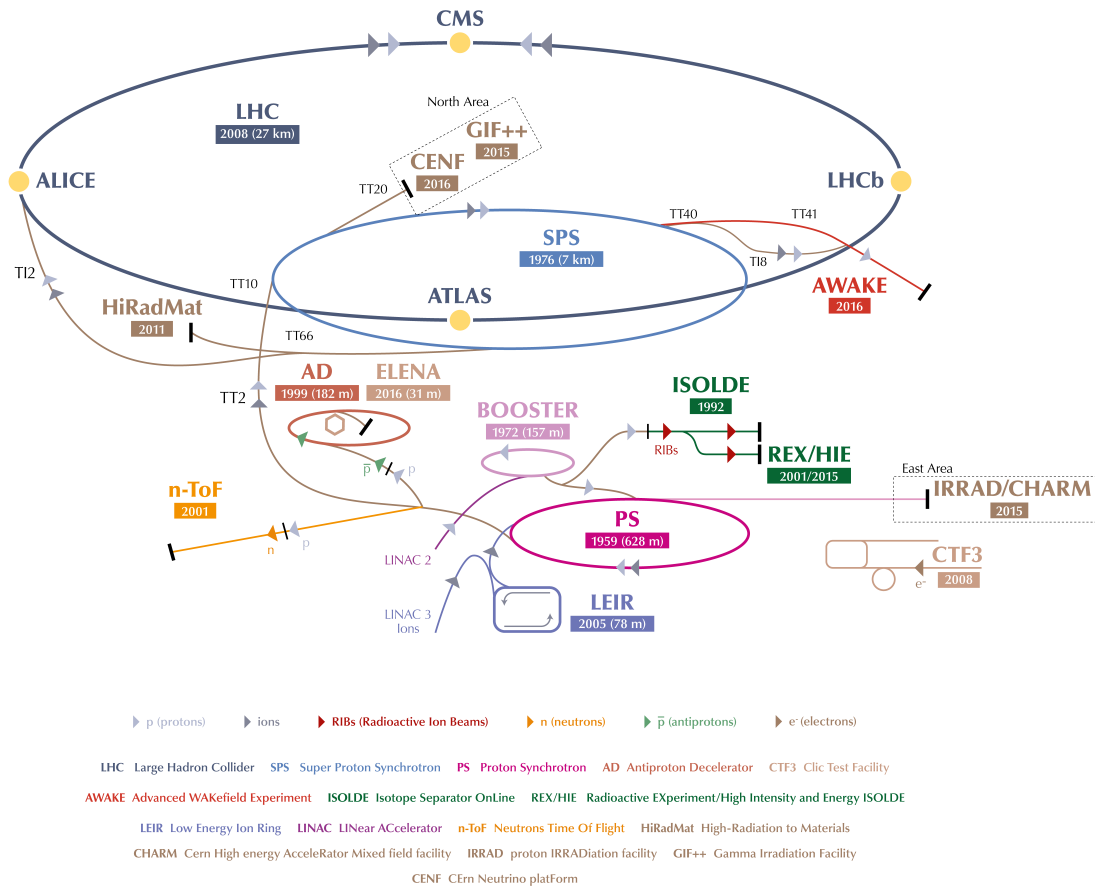


Figure 3.2: A diagram of the LHC injection chain [158].

dipole magnetic fields.

The LHC proton beam is actually composed of multiple bunches of protons. This structure increases the likelihood of a collision. Quadrupole magnets are used to achieve the bunch spacing, with four poles surrounding the beam pipe that can squeeze the beam either vertically or horizontally as needed. These bunches of protons are separated by 25 nanoseconds, and the collisions are often referred to as bunch crossings. The proton beam at the LHC has a specific bunch structure, meaning that some bunches are filled, while some are deliberately left empty. This is done for operational reasons, and is something the LHC experimental systems must take into account.

3.3 Luminosity and Pile-Up

The LHC and its associated experiments measure the amount of data recorded using a concept called luminosity. Luminosity is proportional to the number of collisions in a certain time frame. Often the total luminosity will be referenced, but there is also the concept of instantaneous luminosity (\mathcal{L}_i) which refers to the rate of collisions. Thus the instantaneous luminosity at the LHC is proportional to the bunch crossing rate of the LHC proton beams. The total or integrated luminosity can then be defined as the integral of the instantaneous luminosity over time, usually over a specific run period, for example Run 2 which took place from 2015 - 2018:

$$\mathcal{L} = \int \mathcal{L}_i dt = \int_{2015}^{2018} \mathcal{L}_i dt = 139 \text{ fb}^{-1} \quad (3.1)$$

The luminosity itself depends only on the machine parameters, and can be defined as:

$$\mathcal{L}_i = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \mathcal{E}_n \beta^*} F \quad (3.2)$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, \mathcal{E}_n the normalized transverse beam emittance, β^* the beta function at the collision point, and F the geometric luminosity reduction factor due to the crossing angle at the interaction point [110]. The smaller the β^* , the narrower the beam which increases the number of collisions and thus the luminosity. During a Run at the LHC, the peak luminosity occurs near the beginning, with the luminosity decreasing as protons are used up in collisions. The higher the luminosity of a given dataset, the more data there is to be used for physics analysis.

A final note on the units of instantaneous luminosity: since luminosity is events / time / area, it is usually measured in $\text{cm}^{-1} \text{s}^{-1}$. Due to the scale of this type of experiment, it is inconvenient to actually use centimeters, so often the unit of inverse barns is used, where a single barn is equivalent to 10^{-24} cm^2 . In reality, inverse femtobarns (fb^{-1}) are the actual unit associated with most luminosities, which corresponds to 10^{-15} barns.

As the luminosity increases, it becomes more likely that multiple proton-proton interactions will be observed in a given bunch crossing. These multiple interactions are referred to as "pile-up" and encompasses both multiple interactions from the current bunch crossing, and interactions originated in previous bunch crossings. Pile-up is often characterized by the mean number of interactions $\langle \mu \rangle$ which is defined as:

$$\langle \mu \rangle = \frac{\langle \mathcal{L}_i \rangle \sigma_{inel}}{f_{rev}} \quad (3.3)$$

and thus increases as luminosity increases [26]. It is also affected by certain accelerator parameters like β^* and the narrowness of the beam. The peak μ (and peak luminosity)

is seen early in a run, but as protons interact and fewer protons remain in the circulating bunches, the μ and luminosity can decrease, though again certain beam properties can be tweaked to get the maximum luminosity possible. Figure 3.3 shows the mean number of interactions over the course of Runs 1-3. As seen here, $\langle \mu \rangle$ has increased from 18.5 in Run 1 to 44.6 in Run 3, and that's with a large portion of the Run 3 data taken during commissioning runs in 2022.

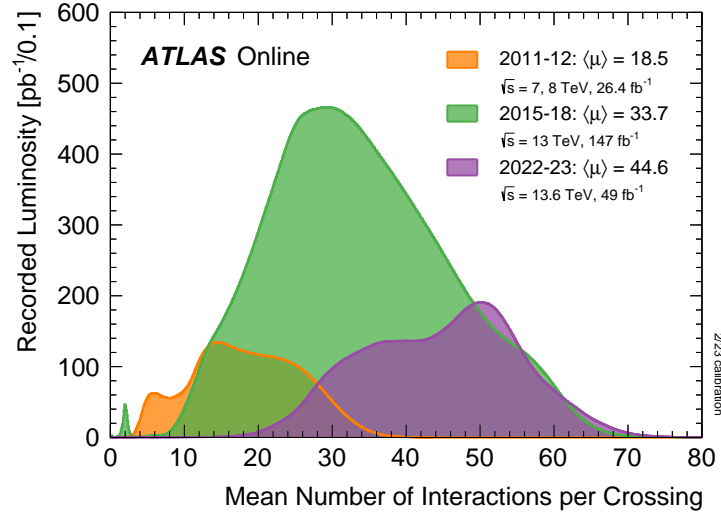


Figure 3.3: Mean number of interactions per LHC Run [94].

3.4 LHC Upgrades and Schedule

The LHC has undertaken two multi-year runs in the past and is currently in the middle of a third run in 2023. In between these runs are periods where the LHC is not running, which are referred to as "Long Shutdowns" (LS). There are existing plans for future runs and future shutdowns which are described below.

Run 1: 2009 - 2013. Center of mass energy of $\sqrt{s} = 7$ TeV which was increased to $\sqrt{s} = 8$ TeV in 2012. As shown in Figure 3.4a, 4.57 fb^{-1} of data at 7 TeV and 20.3 fb^{-1}

of data at 8 TeV were collected by the ATLAS detector and usable for physics over the course of Run 1.

Long Shutdown 1: 2013 - 2015. Upgrades to the accelerator and magnet systems. Design energy was planned to be $\sqrt{s} = 14$ TeV but delays resulted in a decrease to $\sqrt{s} = 13$ TeV. ATLAS added the pixel "insertable b-layer" to the inner detector, and replaced the Minimum Bias Trigger Scintillators (MBTS), and LUCID. ATLAS also upgraded a number of other components including power supplies for the calorimeter, muon system components, ALFA, and the High Level Trigger. Additional new hardware, the L1Topological Trigger and the L1Muon Trigger, were added to the Level-1 trigger system.

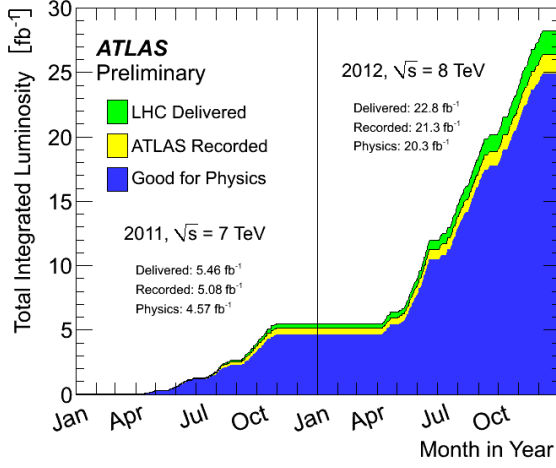
Run 2: 2015 - 2018. Center of mass energy of $\sqrt{s} = 13$ TeV. Figure 3.4b shows that 139 fb^{-1} of data was taken by ATLAS that is usable for physics analysis in Run 2. Data from this run is used for the analysis in this thesis, presented in Chapter 8.

Long Shutdown 2: 2018 - 2022. Extensive ATLAS TDAQ upgrades occurred. Two new muon detectors were installed in the forward region, the New Small Wheels (NSW). The cooling for the Tile calorimeters was upgraded. The Liquid Argon trigger path electronics were extensively upgraded to provide digital signal to the Level-1 Calorimeter trigger system, collectively referred to as the "LAr digital trigger". The entirety of the Level-1 trigger system was upgraded to include a set of three new Feature Extractors to identify physics objects, with the new system running alongside the old at the beginning of Run 3, as discussed in Subsection 4.5.2.

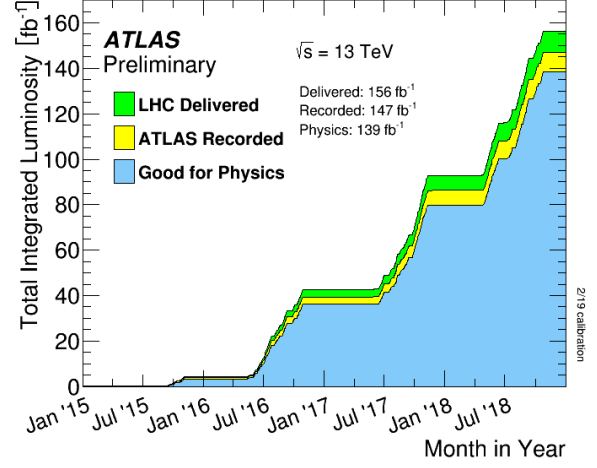
Run 3: 2022 - 2025. Center of mass energy of $\sqrt{s} = 13.6$ TeV. In Figure 3.4c it can be seen that 66 fb^{-1} of data was taken by ATLAS so far in Run 3.

LS3, Run 4, LS4, Run 5, LS5, Run 6: 2026 - 2041. Run 4 marks the beginning of the High Luminosity LHC, and as such LS3 will be spent preparing for the several

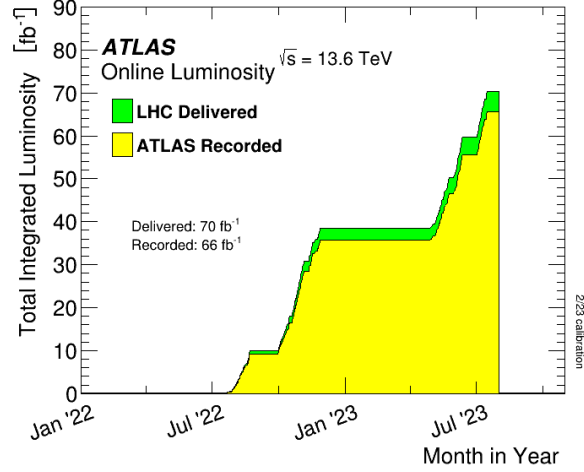
upgrades necessary. LS3 is planned to last 3 full years from 2026 - early 2029 to allow for this, though the schedule of the LHC is subject to change. The full LHC future schedule can be seen in Figure 3.5.



(a) Run 1 luminosity summary [94].



(b) Run 2 luminosity summary [94].



(c) Run 3 luminosity summary[94].

Figure 3.4: LHC luminosity for Runs 1 - 3 [94].

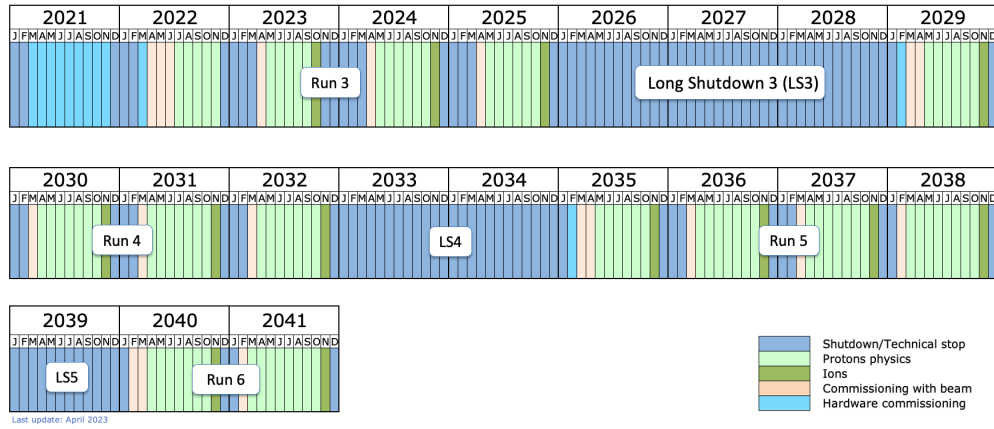


Figure 3.5: LHC long term schedule, available online at the [lhc-commissioning](https://lhc-commissioning.web.cern.ch/) website.

CHAPTER 4

THE ATLAS EXPERIMENT

The ATLAS detector [91, 87] is one of four detectors on the LHC. It is located at Interaction Point 1 (IP1) on the LHC, often colloquially referred to as Point 1. Point 1 is located directly across the street from the CERN Meyrin campus. The ATLAS detector is about 40 m long and 25 m high and is cylindrical in shape with the beam pipe running through the center of the detector. With over 100 million data channels and 3000 km of cabling, ATLAS is one of the largest particle detectors in the world and is an extremely complex machine. As a general purpose detector, ATLAS was designed to characterize all components of particle collision events and to do so it incorporates a full spectrum of sophisticated subdetectors that work together to record all necessary physics data for a variety of physics analysis topics.

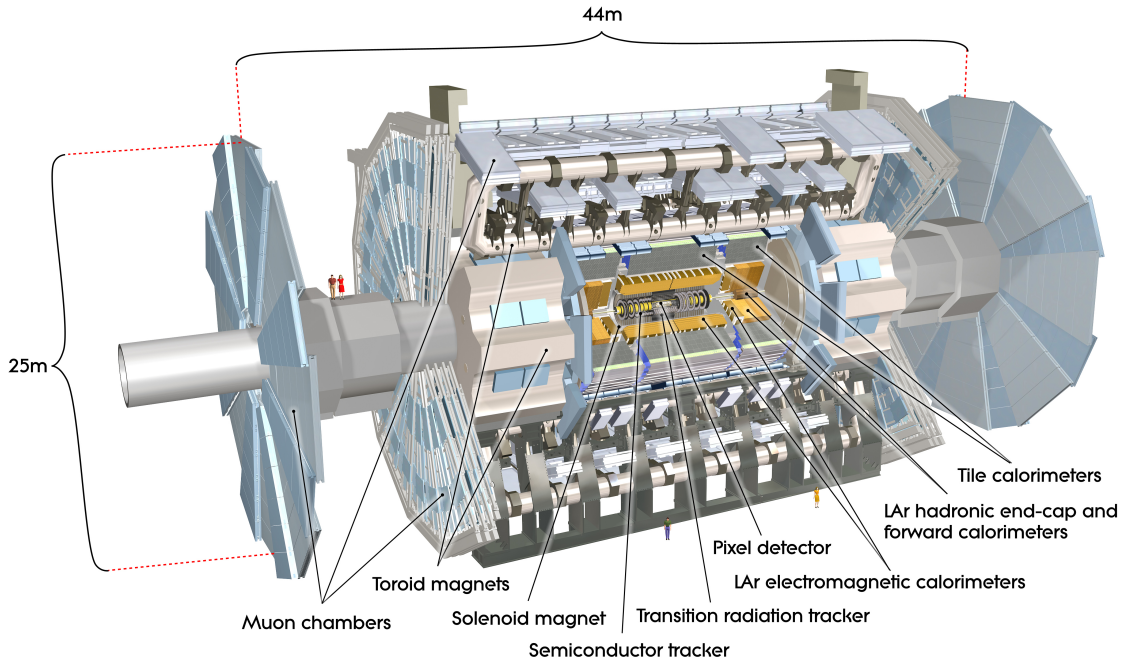


Figure 4.1: A labeled diagram of the ATLAS detector [91].

Figure 4.1 displays these subdetectors which include an inner tracking detector, electromagnetic and hadronic calorimeters, and a muon spectrometer. The ATLAS inner tracker measures the trajectories of charged particles, and the calorimeters measure the energy of electrons, photons, and hadrons. The muon spectrometer measures the trajectories and energy of muons. Neutrinos are not detected by ATLAS but their presence can be inferred using the momentum balance of the other existing particles in the transverse plane (ie. perpendicular to the beam pipe). This is illustrated by Figure 4.2.

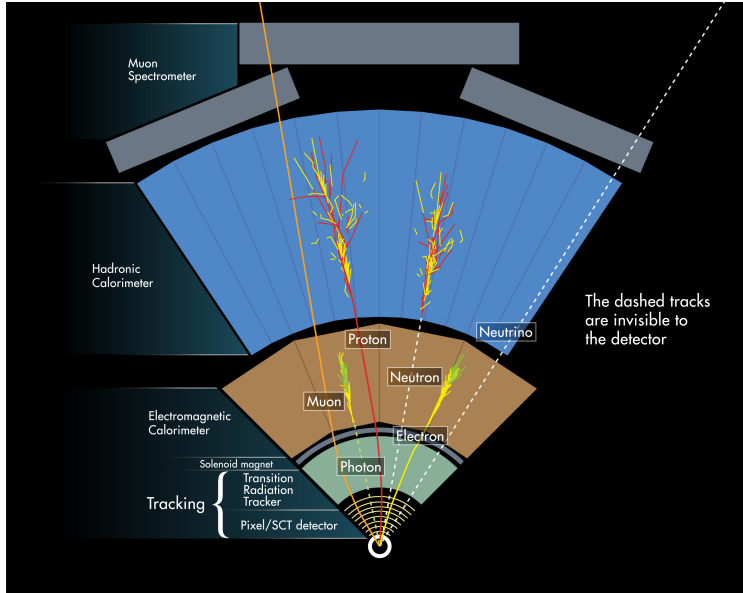


Figure 4.2: A transverse view of the ATLAS detector [168].

The ATLAS detector, like the LHC itself, underwent a number of upgrades for Run 3. In general, the detector will be described as it existed for Run 2 since the data used in this thesis is from Run 2. The hardware work described in this thesis, however, is part of the ATLAS Run 3 upgrade, and as such the upgrades of relevant subsystems may be discussed in specifically marked sections.

4.1 ATLAS Geometry

The ATLAS detector can be characterized by a few different coordinate systems. Cartesian coordinates and cylindrical coordinates are used in different circumstances. In practice, the commonly used coordinates are some combination of several systems.

The Cartesian axes are shown in Figure 4.3a and are defined as follows:

- the positive x-axis points from the interaction point towards the center of the LHC
- the positive y-axis points vertically up from the interaction point
- the z-axis runs along the center of the beam pipe and the positive z points towards LHCb, which is also towards downtown Geneva

The x-y plane is perpendicular to the beam pipe (and the z-axis) and is referred to as the transverse plane. The version of cylindrical coordinates used in the detector are shown in Figure 4.3b and are defined as:

- the same z-axis as in Cartesian coordinates, running along the beam pipe with positive z pointing towards downtown Geneva
- the azimuthal angle ϕ running around the z axis in the transverse plane
- the angle θ from the z axis
- the radial distance r from the z-axis

Particles in the ATLAS detector are often described using their p_T , or momentum in the transverse (x-y) plane. The full particle momentum can be described as $\vec{p} = (p_T, p_z)$, however p_z is not measured by the ATLAS detector. An additional quantity which is measured

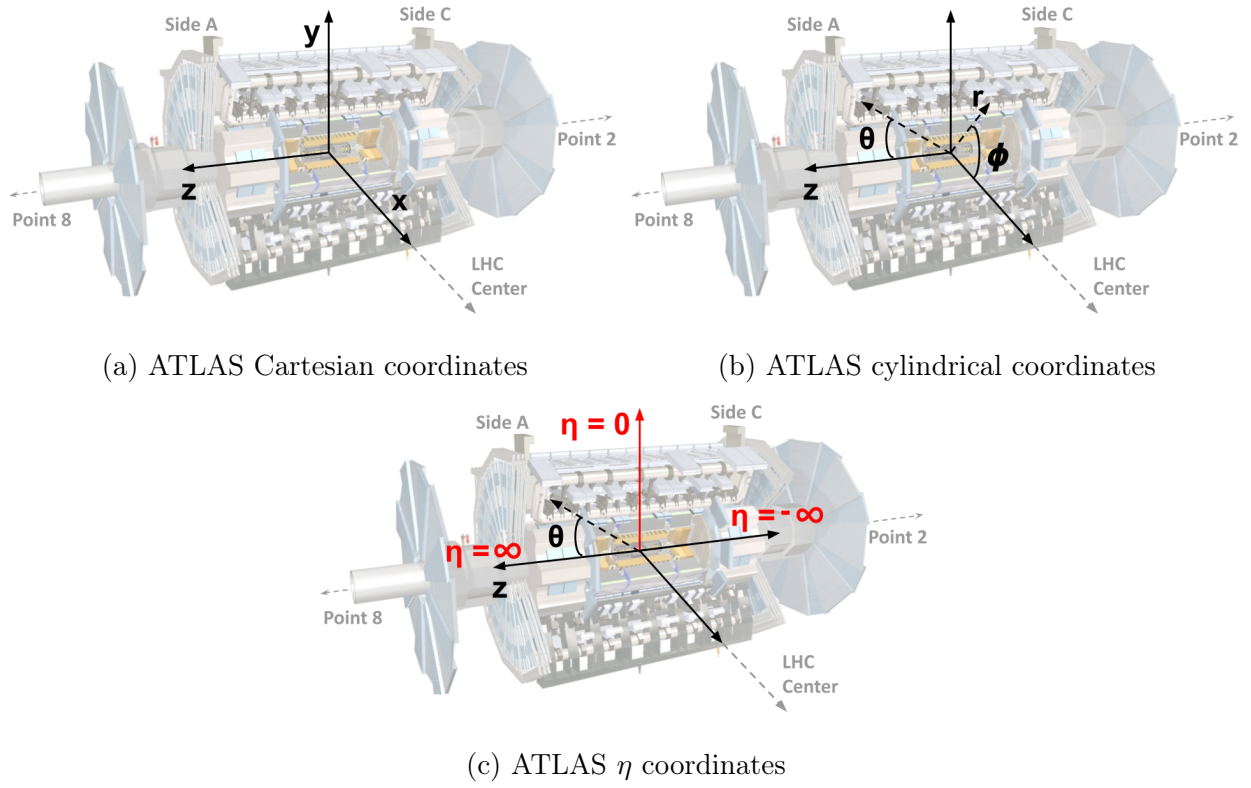


Figure 4.3: ATLAS coordinate systems, modified from ATLAS detector images in Ref [91].

in the transverse plane is missing transverse momentum (the magnitude of which is often referred to as Missing Transverse Energy, E_T^{miss} , or MET). MET is calculated from the momentum conservation of all existing particles in the event and can indicate a missing particle that was not detected, like a neutrino.

An additional coordinate that is often used is η , or pseudorapidity. This is defined as:

$$\eta = -\tan\left(\frac{\theta}{2}\right) \quad (4.1)$$

As shown in Figure 4.3c, particles perpendicular to the z-axis have $\eta = 0$, while those parallel to the z-axis have $\eta \rightarrow \pm\infty$. In the limit of a particle traveling close to the speed of light (or equivalently when the mass of a particle is negligible) η is invariant under Lorentz transformations, and converges to the definition of full rapidity y , defined as:

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right) \quad (4.2)$$

The most commonly used coordinates are z , ϕ , and η . It is particularly common to talk about the position of a physics object in $\eta - \phi$ space. The difference in this space is denoted $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, and is often used to denote difference in particle location, as the $\eta - \phi$ plane is perpendicular to the transverse plane.

4.2 Inner Detector

The purpose of the inner detector is to reconstruct the path of charged particles, often referred to as *tracking*. This is primarily done through the ionization of electrons which occurs when a charged particle passes through a tracking detector. With multiple layers of

trackers, the particle path can be tracked as it moves radially away from the interaction point. The inner detector itself is surrounded by a solenoid magnet which produces a 2 T magnetic field, which produces a force on charged particles in the detector via $\vec{F} = q\vec{v} \times \vec{B}$. This force causes them to curve in the transverse (x-y) plane, and provides additional information about the transverse momentum and charge of the particle.

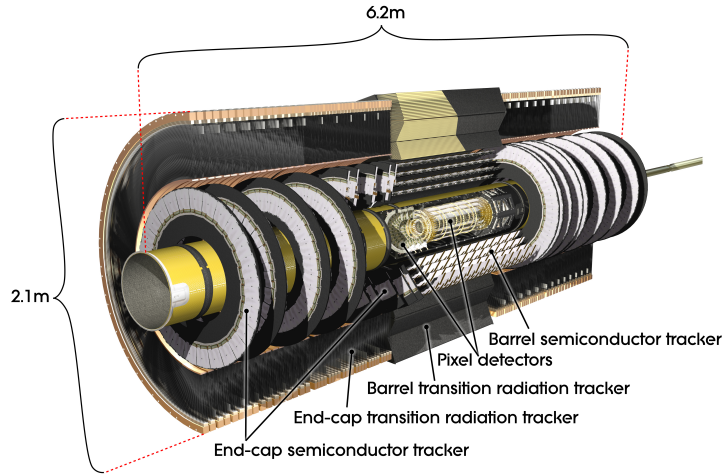


Figure 4.4: The ATLAS inner detector [91].

The ATLAS inner detector [46] is the first component of ATLAS after the beam pipe and is a cylindrical detector as shown in Figure 4.4. The detector is designed to have very good momentum resolution for charged particles with $p_T > 500$ MeV within $|\eta| < 2.5$, and has three main components: the pixel detector, the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT).

The pixel detector [193] is made of 80 million silicon pixels and provides extremely high-granularity tracking. As the closest detector to the beam pipe, it generates extremely important positional information. These pixels cover a 1.7 m^2 area and cover $|\eta| < 2.5$. The pixel detector has four barrel layers, including the insertable B-layer which was added prior to Run 2, which can all provide a measurement for a charged particle track. These can be seen in green in Figure 4.5. The pixel detector is crucial for the reconstruction of primary

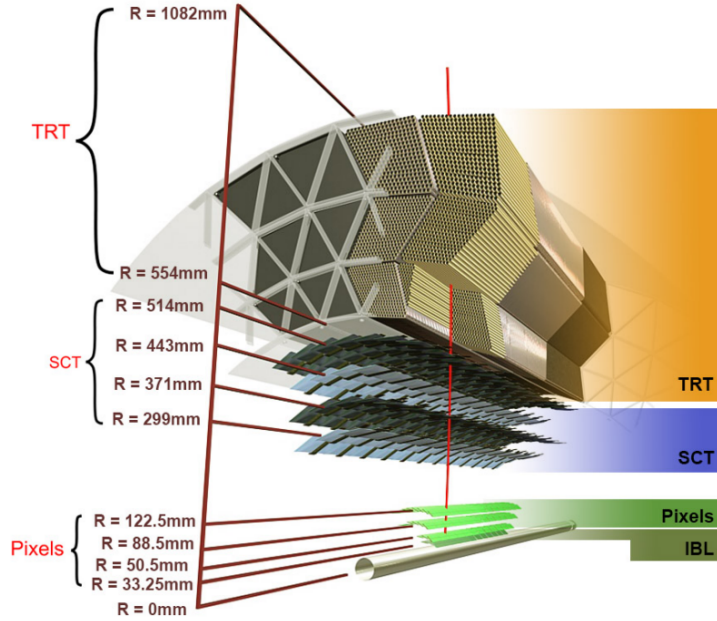


Figure 4.5: The ATLAS inner detector components, viewed radially [86].

and secondary vertices and is thus very important also for things like b-tagging, which are important components of many ATLAS analyses.

The next layer of the inner detector is the SCT [85], which can be seen in blue in Figure 4.5. The SCT is a silicon microstrip tracker with over 4,000 modules, each with over six million strips covering a total area of 62 m^2 also for $|\eta| < 2.5$. There are a total of four layers in the barrel SCT, which are designed to provide 4 measurements per charged track. The SCT endcaps have a total of 9 "wheels" on each side of the cylindrical barrel modules and are designed to provide 9 measurements per charged track. The SCT helps to measure charged particle momentum, as well as vertex identification along with the pixel detector.

The final component of the inner detector is the Transition Radiation Tracker [92, 93] shown in gold in Figure 4.5. The TRT is composed of over 350,000 drift tubes covering a volume of 12 m^3 for $|\eta| < 2.0$. These drift tubes are filled with Xenon gas and have a wire at the center which collects electrons displaced by incoming particles. The TRT includes materials with

widely varying indices of refraction (xenon, carbon dioxide, and oxygen), which leads to the production of transition radiation, or radiation which is produced when a charged particle passes through an inhomogeneous media. The energy of this radiation depends on the mass of the particle, allowing for particle identification to be performed. This is particularly useful for distinguishing electrons from charged hadrons.

4.3 Calorimeters and Calorimetry

The ATLAS calorimeters are radially located after the inner detector and have the job of measuring the energy of particles produced in collisions, which is done by inducing a particle shower and subsequently absorbing a large fraction of the total energy of the original particle. The ATLAS calorimeter system is separated into two components, the mainly electromagnetic Liquid Argon (LAr) calorimeter [47] and the hadronic Tile calorimeter [49]. Both of these systems are implemented as sampling calorimeters, meaning that there are layers of energy absorber material interspersed with layers of active material which provide a detectable signal. This combination allows for a detailed measurement of the particle shower and the corresponding particle energy. However, with a sampling calorimeter only a fraction of the actual particle energy is measured and a calibration is required to recover the relationship between the absorbed energy and the initial particle energy. This calibration is generated by studying the calorimeter response and plays a role in the energy scale and resolution of measured objects, particularly jets. Jets are characterized by parton showers and hadronization as discussed in Section 2.4, but there are additionally many experimental challenges in measuring the energy of jet which are heavily related to the calorimeters and their energy calibrations. The jet energy scale and resolution are important source of systematic uncertainties in ATLAS measurements and a large amount of work goes into taking into account these experimental detector constraints properly so that the measured data can

be used in physics analyses. Further discussions on the jet energy scale and resolution and their impact on ATLAS measurements can be found in Subsection 8.7.4.

The LAr calorimeter uses lead and liquid argon, while the Tile calorimeter uses steel and plastic scintillators. A full overview of the ATLAS calorimeter system can be seen in Figure 4.6.

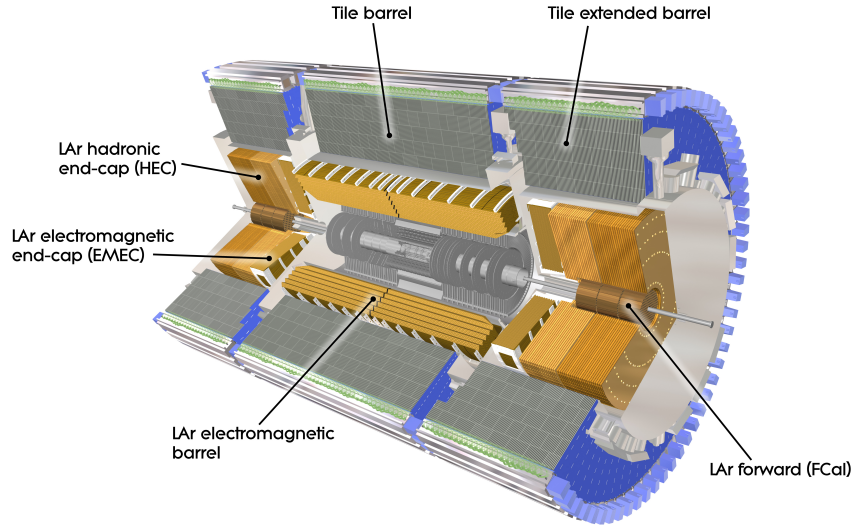


Figure 4.6: The ATLAS calorimeters, labeled [91].

4.3.1 *Electromagnetic Calorimetry*

Electromagnetic calorimeters (EMCal) are designed to measure the energy of particles interacting via the electromagnetic force. At all but very low energies (less than 10 MeV) bremsstrahlung or "braking radiation" is the primary process through which this occurs, as shown in Figure 4.7. Bremsstrahlung is electromagnetic radiation produced by a charged particle that is decelerating, usually through interaction or deflection by another charged particle. This is often an electron which is deflected by the electric field of an atomic nucleus, particularly a heavy nucleus. Electromagnetic radiation is ultimately a photon, and

this photon may go on to produce an electron-positron pair, which starts the bremsstrahlung cycle all over again. The radiation length X_0 is equal to the mean distance over which a high-energy electron loses all but $\frac{1}{e}$ of its energy by bremsstrahlung, and this distance is often used when describing electromagnetic showers.

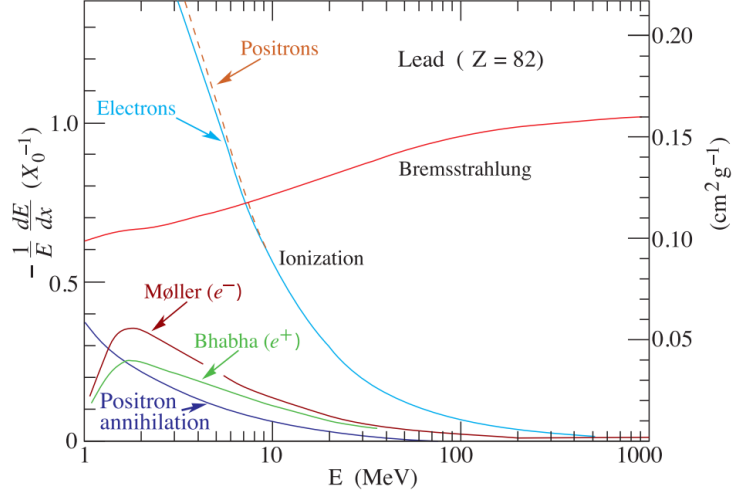


Figure 4.7: Fractional energy loss per radiation length in lead as a function of electron or positron energy [197].

At lower energies the process of ionization is much more likely to occur, with electrons forming ions with the atoms of the absorber material. The energy at which the rates of ionization and bremsstrahlung are equal is sometimes referred to as the critical energy, which for lead is around 10 MeV as seen in Figure 4.7. The depth of the shower, or how long before the electron switches from bremsstrahlung to ionization, can be approximately defined as shown in Equation 4.3 [197].

$$x = X_0 \frac{\ln(\frac{E}{E_c})}{\ln(2)} \quad (4.3)$$

which for a 100 GeV electron in lead results in a shower depth of $x \approx 13X_0$, or 13 radiation lengths. Radiation lengths are often reported in g cm^{-2} , but the result in centimeters can

be obtained by dividing by the density. For lead $X_0(Pb) = 6.37 \text{ g cm}^{-2} = \frac{6.37 \text{ g cm}^{-2}}{11.4 \text{ g cm}^{-3}} = 0.56 \text{ cm}$. So, this 100 GeV electron shower could be captured in $13X_0(Pb) \simeq 7.3 \text{ cm}$ of lead. A shower being "captured" here means that the repeating process of bremsstrahlung has given way to ionization, which guarantees that the remaining energy is below the 10 MeV threshold mentioned above. So for a 100 GeV electron, 99.99% of the original energy has been captured by the calorimeter. One could perhaps round up to around 10 cm which is a relatively small region of space compared to the overall ATLAS detector diameter of 25 m. The distance needed does however change as a function of the energy of the incoming particle. For photons, pair production and the resulting bremsstrahlung are also the dominant process for energies above 10 MeV, while for those below 1 MeV the photoelectric effect dominates. Electromagnetic showers are not deterministic, but due to the small number of processes available, the EM showers tend to look quite similar for similar energies.

4.3.2 Hadronic Calorimetry

Hadronic calorimeters (HCal) are designed to measure the energy of particles interacting via the strong nuclear force. There are a significantly higher number of processes involved than in EM calorimeters, and these processes generally have more widely varying length scales. Additionally hadrons have higher masses than electrons which results in the suppression of bremsstrahlung as it is inversely proportion to the square of the mass of the particle. These hadrons can experience energy loss through ionizing interactions with electrons. However, the full hadronic shower only begins after the interaction of hadrons with atomic nuclei. This can lead to the production of both charged pions (π^\pm) and neutral pions (π^0). π^0 has a very short lifetime and immediately decays to two photons, which subsequently produce an electromagnetic shower as described previously. The π^\pm travel further and can experience additional strong interactions with other nuclei, producing additional particles with different

lifetimes and properties, thus resulting in a complicated shower with both an electromagnetic and hadronic component, as shown in Figure 4.8.

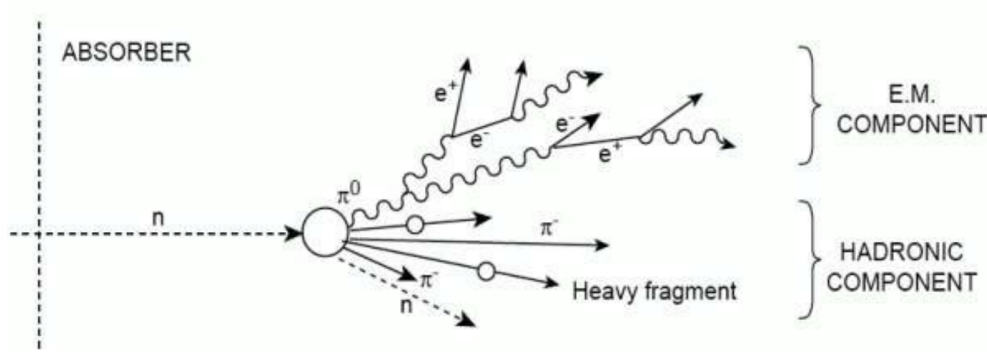


Figure 4.8: Hadronic shower with the characteristic electromagnetic shower component [144].

It is clear from this that hadronic showers are both much more complex and much more variable than electromagnetic showers. The interactions are mainly via the strong force, which results in smaller cross-sections and lower likelihoods of interaction. This means that the length scales involved in hadronic showers are larger, requiring more distance to be covered by detectors. The length scale or interaction length λ_I again depends on the material, for iron (steel being an iron alloy) $\lambda_I(Fe) = 16.77cm$ [197].

4.3.3 Liquid Argon Calorimeter

The LAr electromagnetic barrel (EMB) and electromagnetic endcap (EMEC) are sampling calorimeters made of lead and liquid argon which have an "accordion geometry" as shown in Figure 4.9. The EMB and EMEC cover $|\eta| < 3.2$ and perform precision electromagnetic shower measurements. The unique geometry of the module provides uniform coverage over ϕ without gaps, and has three radial layers. The first sampling layer is composed of strips with a very fine granularity in η , while the second sampling layer, which collects the majority of the shower energy, has a fine granularity in both η and ϕ . The final radial layer is coarser

in granularity and collects the tail end of the shower energy. The LAr EMB is composed of two half cylinders (or "barrels") and covers $|\eta| < 1.475$, while the EMEC is made of two "wheels" (one on either end of the EMB) and covers $1.375 < |\eta| < 3.2$. There is additionally a thin LAr pre-sampler covering $|\eta| < 1.8$ which provides the ability to correct for energy losses incurred prior to the EM calorimeter.

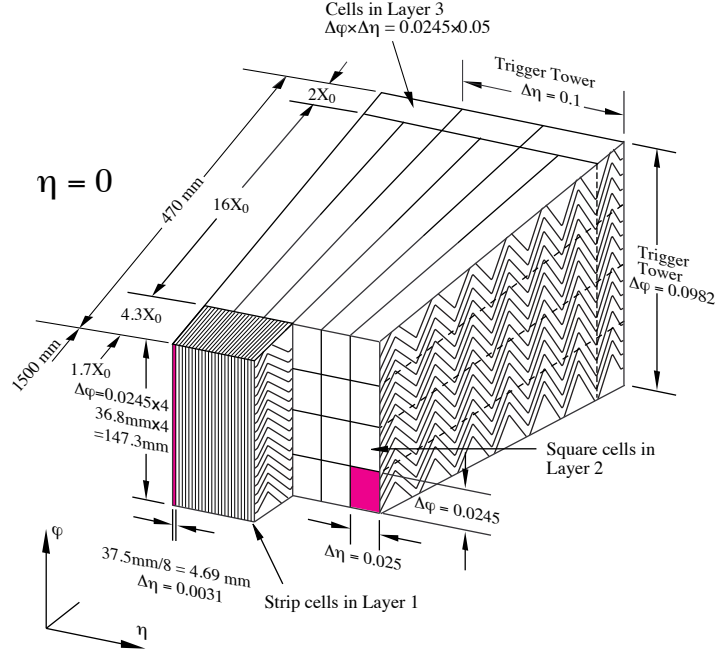


Figure 4.9: A diagram of a LAr module which shows the accordion geometry and the different layers radially, and in ϕ and η . The granularity is also shown [91].

The LAr calorimeter also has a hadronic component in the endcap and forward regions. The LAr hadronic endcap (HEC) is a sampling calorimeter composed of copper and liquid argon and covers $1.5 < |\eta| < 3.2$. It consists of two wheels, each formed from 32 individual wedge shaped modules, which are located after the LAr EMEC modules (further along the z-axis). The final LAr modules are the forward calorimeters (FCal) which provides EMEC and HCal coverage using 3 total modules with coarse granularity in the region $3.1 < |\eta| < 4.9$. One of these modules provides EMEC coverage and is composed of copper and liquid argon, while

the other two provide HCal coverage and are composed of tungsten and liquid argon.

4.3.4 *Tile Calorimeter*

The Tile calorimeter is a hadronic sampling calorimeter made of steel and scintillator covering the region $|\eta| < 1.7$. It is located after the EMB and EMEC calorimeters and has a central barrel covering $|\eta| < 1.0$ and two extended barrels which cover $0.8 < |\eta| < 1.7$. Each barrel consists of 64 wedge shaped modules, as shown in Figure 4.10. The steel tiles are utilized as absorber layers and are about 4-5 mm thick. The active tiles are about 3 mm and are constructed from doped polystyrene scintillators attached to photomultiplier tubes with optical fibers. Overall the Tile calorimeter provides three layers of measurement with varying granularity. The inner two layers are $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ while the third is $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$. Tile is about 8-12 interactions lengths (λ as discussed above for hadronic calorimetry) depending on the pseudorapidity of the incoming particle.

The scintillating tiles are under constant bombardment and thus experience radiation damage, which decreases the light produced by the tiles. This requires that ongoing scans are done to calibrate the system. These scans utilize a laser system to calibrate the PMTs directly by producing pulses of light similar to signals from the scintillator itself.

4.3.5 *Calorimeter Energy Resolution*

The energy resolution of the calorimeters has a profound effect on the way that the measured data can be used in physics analyses. Generally, energy resolution is better as the energy increases. This is due to the fact that the resolution is better with a higher number of distinct particles, and you get more interactions and more subsequent splittings with higher energy incident particles. This is particularly relevant for sampling calorimeters where the measured

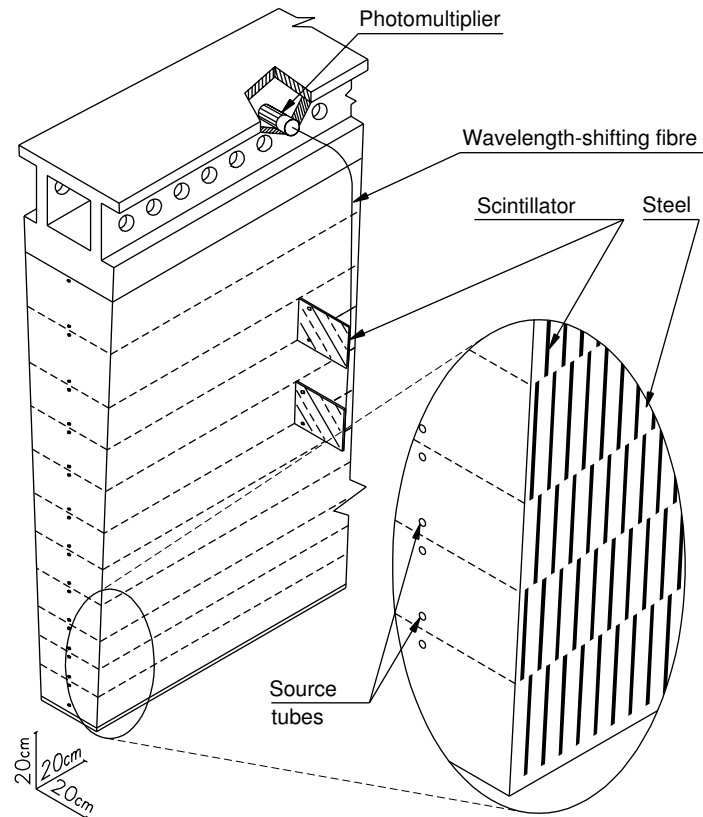


Figure 4.10: A diagram of a Tile module which shows shape and granularity of the modules [91].

energy depends strongly on the number of particles which pass through the sampling medium. Overall the resolution of a calorimeter can be roughly described by the following:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (4.4)$$

where the first term is often referred to as the stochastic term and depends on fluctuations of the shower as discussed above [111]. The second term is the noise term and mainly comes from electronic noise in the readout chain, which is strongly dependent on the electronics used. The last term is the constant term and includes effects which are not dependent on the energy of the particle, like constant instrumentation effects, or effects from detector geometries.

Returning to the noise term for a moment, it is clear that the calorimeter energy response depends somewhat on the level of noise present. Scintillator-based calorimeters (like the Tile calorimeter in ATLAS) can achieve very low levels of noise when the first step of the electronic chain is a photosensitive device, since these devices can provide high-gain multiplication with almost no noise. When the signal is collected in the form of charge (like the Liquid Argon calorimeters) noise tends to be high because the first element in the readout chain is a preamplifier. There is a fundamental limit to the possible noise mitigation when measuring charge which is referred to as Johnson Noise, or noise which is generated by the thermal agitation of the charge carriers:

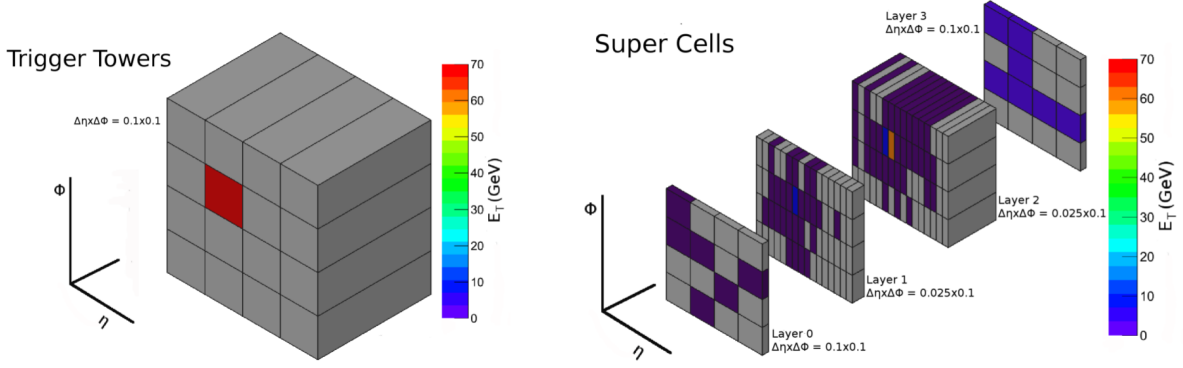
$$Q = \sqrt{4kTR\Delta F} \quad (4.5)$$

where Q is the equivalent noise charge, k is the Boltzman constant, T is the temperature, R the equivalent noise resistance of the preamplifier, and ΔF the bandwidth [111]. In a sampling calorimeter it is often possible to increase the sampling fraction and thus increase the signal-to-noise ratio, however there are limitations in bandwidth from read-out electronics which constrain this mitigation strategy.

The energy resolution of the calorimeters is quite complex as a whole, and these effects must be carefully considered when using calorimeter data in physics analyses. There is an entire group in ATLAS that is devoted to making sure we are treating calorimeter data properly. There will be additional discussions of the Jet Energy Scale (JES) and Jet Energy Resolution (JER) in Subsection 8.7.4 and how these calorimeter effects carryover to physics analyses, particularly their uncertainties.

4.3.6 Calorimeter Upgrades for Run 3

Upgrades to the calorimeter systems are mainly motivated by the desire to provide higher granularity information to the Level-1 calorimeter trigger system (L1Calo). The movement of the digitization process earlier in the trigger path also reduces the expected effects of noise and radiation damage, particularly for the HL-LHC. L1Calo is also being upgraded extensively for Run 3 which is discussed in detail in Subsection 4.5.2. The main upgrades to LAr come in the form of updated read out architecture and updates to the L1Calo trigger interface [4]. The Run 2 LAr-L1Calo interface was based on the concept of a Trigger Tower which sums the energy deposition in the area $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ over several stages of analog electronics located on-detector. The new granularity scheme implemented for Run 3 is based on supercells which provide layer-by-layer information with finer segmentation from the LAr EMB and EMEC, the details of which can be seen in Figure 4.11b.



(a) Trigger Towers used in the LAr-L1Calo interface during Run 2 [4].

(b) Supercells used in the LAr-L1Calo interface for Run 3 [4].

Figure 4.11: Run 2 to Run 3 Liquid Argon calorimeter cell upgrades [4].

To provide high granularity and high precision digital data to L1Calo, new LAr Trigger Digitizer Boards (LTDBs) are installed along with the LAr front-end electronics. These new LTDBs digitize finer granularity information up to $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ in the first two layers of the EMCal, and also recreate the legacy trigger towers to send through the legacy LAr-L1Calo interface [109]. The LAr Digital Processing System (LDPS) takes in the digitized signals and converts them to calibrated energies (where calibration here refers to the process discussed at the beginning of Section 4.3) which are sent to L1Calo. The LDPS consists of the Liquid Argon Trigger Mezzanine card (LATOME) and the Liquid Argon ATCA carrier blade (LArC). The LATOME firmware is where the bulk of the summing and energy conversion occur, and these LATOMEs are the direct interface to the Feature Extractors of L1Calo, which are described in detail in Subsection 4.5.2 [109].

The Tile calorimeter electronics are planned to be fully upgraded for Run 4 with the HL-LHC. For the new digital L1Calo system, the TileCal signal did need to be digitized, but this was incorporated into the Tile Rear Extension Module (TREM), which is described in Subsection 4.5.2.

4.4 Muon Detectors

Muons interact electromagnetically, but are extremely heavy compared to the electron ($m_\mu = 106$ MeV, $M_e = 0.510$ MeV). As shown in Figure 4.7 Bremsstrahlung is the primary interaction in electromagnetic showers, but the energy radiated due to Bremsstrahlung is inversely dependent on the mass of the particle so muons lose less energy in this way and they are not stopped by the electromagnetic calorimeter. A dedicated muon detector is required to fully characterize muons from particle collisions in ATLAS. The muon spectrometer [48] is designed for the detection and measurement of muon quantities and is the outermost layer of the ATLAS detector and by far the largest, spatially. It has four components: two types of triggering chambers, and two sets of high-precision tracking chambers.

The trigger chambers are composed of Resistive Plate Chambers (RPC) [3] and Thin Gap Chambers (TGC) [150] which detect the presence of a muon, and perform a coordinate measurement of said muon. The tracking chambers include the Monitored Drift Tubes (MDT) [57] and the Cathode Strip Chambers (CSC) [9] which provide precision measurements of muon tracks and of muon p_T . These tracking chambers function similarly to those of the inner detector and use both barrel and endcap superconducting toroidal magnets to provide a magnetic field of 2-6 T across the majority of the detector. These magnets provide the deflection of muon tracks that is then measured by the MDT and CSC detectors, using three layers of barrel and endcap chambers.

All together, the muon spectrometer covers an η range of $|\eta| < 2.4$ for the trigger chambers and $|\eta| < 2.7$ for the tracking chambers. The full muon spectrometer can be seen in Figure 4.12.

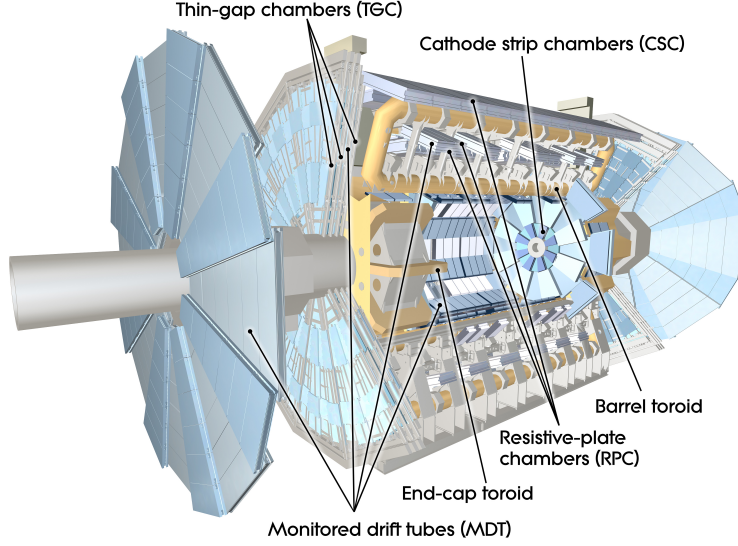


Figure 4.12: The ATLAS muon detector, labeled [91].

4.5 Trigger and Data Acquisition

The Trigger and Data Acquisition system (TDAQ) is a vital component of the ATLAS detector which is responsible for selecting the subset of the available collision data that will be saved for offline physics analysis. The LHC collides proton bunches at a rate of 40 MHz and has about 120 billion protons per bunch. Assuming only a subset of these protons interact at any given "bunch crossing", this results in around 1 billion collisions per second! With each bunch crossing in atlas about 1MB of data is read out so this results in about a petabyte of data per second, which would not only be difficult to save but would also require a huge amount of storage and a very high number of CPU hours to process and reconstruct. Thus the job of the trigger system is very important as it must decide what data is considered interesting enough to save for offline reconstruction and use in physics analyses.

The ATLAS trigger system is designed to select events by identifying the signatures of electrons, photons, tau leptons, muons, and jets. During Run 1 the system was designed to function at a center of mass energy of $\sqrt{s} = 7$ TeV and instantaneous luminosities around

$8 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. This luminosity resulted from (among other things) the 40 MHz LHC bunch crossing rate and 50 ns bunch spacing which produced an average of 25 interactions per bunch crossing. Over the course of Run 1 the trigger could save events at a rate of 200 Hz or up to 600 Hz for short periods of time [88]. During the second Run of the LHC the instantaneous luminosity did not increase significantly but the center of mass energy increased from $\sqrt{s} = 7 \text{ TeV}$ to $\sqrt{s} = 13 \text{ TeV}$. This resulted in a $2.5\times$ increase in trigger rates when using the same trigger criteria used in Run 1, necessitating updates to the trigger hardware and configuration in Run 2 [89]. The data used in this thesis was taken during Run 2 so I will describe the system as it functioned at this time. The upgrades for Run 3 will be discussed in Subsection 4.5.2 and Chapter 5.

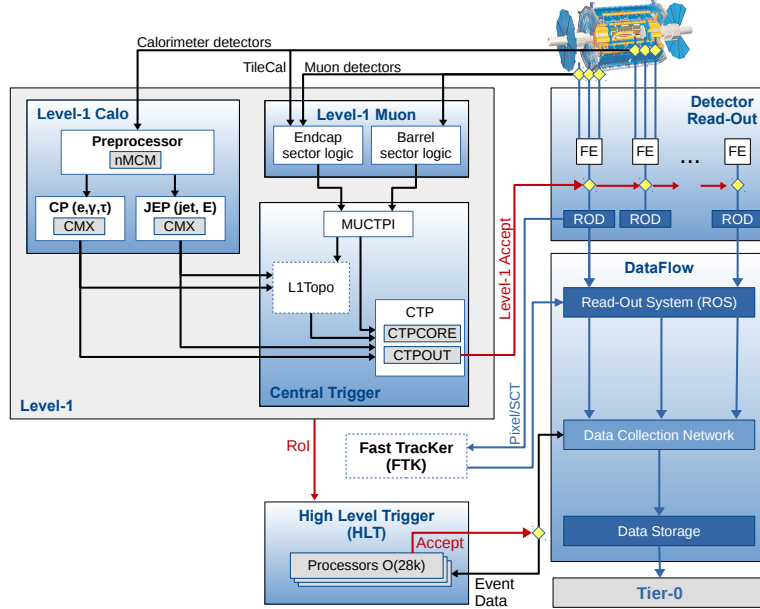


Figure 4.13: Overview of the Run 2 Trigger and Data Acquisition System for the ATLAS detector. Shows the Level-1 Trigger, the HLT, and the data readout path. [89]

The TDAQ system is composed of the Level-1 hardware trigger (L1), and the software based High Level Trigger (HLT) as can be seen in Figure 4.13. The Level-1 trigger is composed of the Level-1 Calorimeter trigger, the Level-1 Muon trigger (L1Muon), and the Central Trigger Processor (L1CTP). The L1Calo and L1Muon systems process input data from their respective subdetectors and send the resulting Regions of Interest (ROIs) or Trigger Objects (TOBs) the L1Topological Trigger and the Muon-to-Central Trigger Processor Interface(MuCTPI), which forward this information to L1CTP. The L1CTP makes a trigger decision based on these inputs and generates a Level-1 Accept (L1A) signal with a maximum rate of 100 kHz. The L1A is sent to all detector and trigger components and prompts these components to read out the data corresponding to the event which caused the L1A, via the readout system. This data is then buffered and processed by the HLT which performs further analysis on the TOBs received from the L1 trigger. As shown in Figure 4.14 the L1 trigger system has a maximum acceptance rate of 100 kHz. The HLT further filters this data and has a maximum acceptance rate of 1 kHz for Run 2, meaning the ATLAS detector is able to save only 1 event out of every 40,000 produced by LHC collisions in Run 2. The trigger system thus must be extremely precisely optimized to select interesting and rare events for offline physics analysis, and as such is a crucial component of the ATLAS detector.

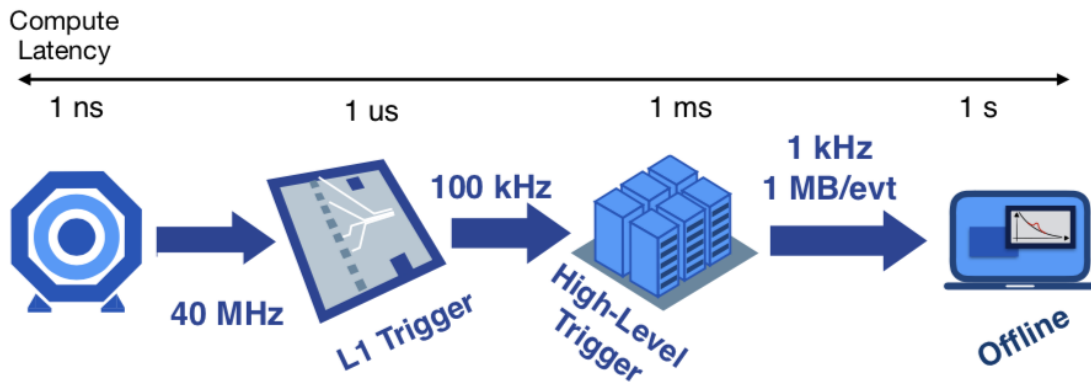


Figure 4.14: Breakdown of the overall trigger rate from the hardware and software systems to the offline readout. Image Credit: Maurizio Pierini.

4.5.1 Level-1 Calorimeter Trigger

The Level-1 Calorimeter trigger has the very difficult job of identifying interesting objects such as electrons, photons, taus, jets, and E_T^{miss} , and furthermore doing all of this with the very limited latency of $2.2 \mu s$.

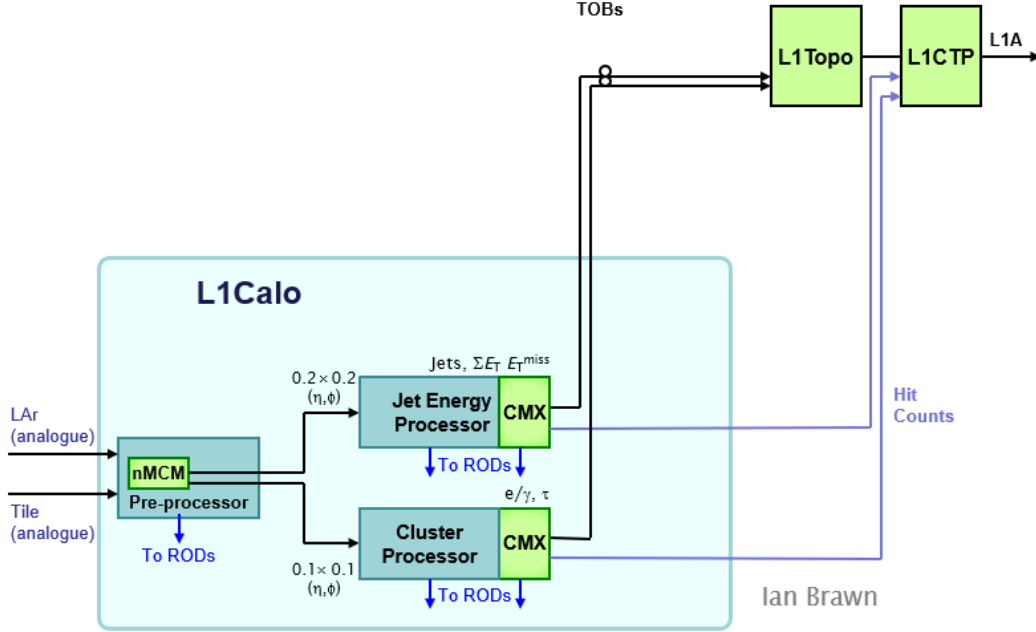


Figure 4.15: Run 2 L1Calo Schematic showing the flow of data from the calorimeters through to the L1CTP.

L1Calo receives input from both the Liquid Argon and TILE calorimeters of ATLAS, as described in Section 4.3. Due to the extremely low latency requirement, and a strong preference for flexibility, this system is implemented in hardware, primarily programmable hardware chips called FPGAs. FPGAs allow for a fixed latency implementation of complex algorithms which is vital in a system like L1Calo that must identify physics objects in $2.2 \mu s$. They also provide the flexibility for algorithm modifications and upgrades as required by the system.

The Level-1 Calorimeter system is composed of three types of submodules: the Pre-Processor

Module (PPM), Jet Energy Processor (JEP) and Cluster Processor (CP) as seen in Figure 4.15.

The PPM digitizes and calibrates the received signal from the calorimeters through the use of a pedestal subtraction scheme. The new Multi Chip Modules (nMCM) on the PPM process and sum this data into Trigger Towers (TT) which are 0.1×0.1 in $\Delta\eta \times \Delta\phi$, as visualized in Figure 4.16. The Trigger Tower E_T and associated bunch crossing (BCID) is sent to the JEP and CP modules for the identification of physics objects.

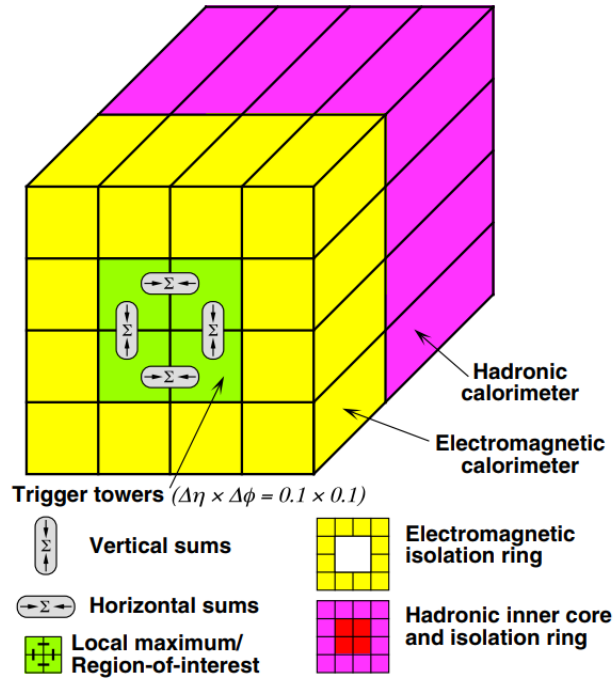


Figure 4.16: Visualization of L1Calo Trigger Towers, corresponding to summed electromagnetic and hadronic calorimeter cells.

The Cluster Processor uses the input TTs (~ 7200 to cover the entire calorimeter) and a 4×4 sliding window algorithm to produce six basic building blocks which are used when identifying TOBs [107]. These six building blocks are defined as follows:

1. Four overlapping 2×1 **EM clusters** used to measure the E_T of electromagnetic showers

(the green towers in Figure 4.16).

2. A **hadronic core** created by summing the four central hadronic towers (shown in red in Figure 4.16), used for isolation tests in the hadronic calorimeter.
3. Four **hadronic clusters** made by summing each individual hadronic tower in (2) with the EM clusters in (1), used to measure the E_T of hadronic showers.
4. An **EM isolation ring** utilizing the towers surrounding the EM clusters in (1), shown in yellow in Figure 4.16.
5. A **hadronic isolation ring** made of the towers behind the EM isolation ring.
6. A 2×2 tower **cluster ROI or TOB**, created by summing the centered EM clusters (1) with the hadronic core (2) and used to identify trigger objects.

The CP uses these calculated values to search for electromagnetic and hadronic trigger candidates by requiring that the central clusters pass an E_T threshold while the isolation ring E_T remains under the specified threshold. For the tau/hadron and electron/photon algorithms respectively the shower is required to extend or not extend into the hadronic towers. If these criteria are met, then a trigger candidate is found.

The JEP sums over 2×2 TTs to create jet elements (0.2×0.2 in $\Delta\eta \times \Delta\phi$). These elements are summed to create potential jet clusters (0.4×0.4 in $\Delta\eta \times \Delta\phi$) and jets windows ranging from 0.4×0.4 , 0.6×0.6 , 0.8×0.8 in $\Delta\eta \times \Delta\phi$. To select a jet candidate, the jet cluster E_T must be a local maximum and the jet window E_T must be greater than the given threshold.

The Common Merger Extended Module (CMX) receives the TOBs from the CP and JEP modules and sends the number of TOBs and the corresponding energy and position of these TOBs to the L1Topological (L1Topo) trigger. The Level-1 Topological Trigger can apply kinematic and angular requirements to the trigger objects to form topological combinations

of different types of objects. L1Topo improves background event rejection and acceptance of signal events with the increased center of mass energy present during Run 2. Based on the output of L1Topo, the L1CTP generates a Level-1 Accept signal which causes all data to be read out and sent to the High Level Trigger.

4.5.2 TDAQ Run 3 Upgrades

The LHC and ATLAS are undergoing a number of upgrades in preparation for Run 3. The center-of-mass energy has increased to $\sqrt{s} = 13.6$ TeV, and the instantaneous luminosity is expected to reach values of $10 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$. In addition, the LHC will run with roughly double the number of proton bunches, resulting in increased pileup with an expected $\langle\mu\rangle$ of around 60, double that seen in Run 2. This increase can already be seen in the preliminary data from Run 3 shown in Figure 4.17, despite the fact that much of the LHC beam in 2022 was at lower energies and with fewer bunches as part of the restarting and commissioning process for the LHC.

As mentioned in Section 4.3, the ATLAS Liquid Argon Calorimeter has undergone major upgrades in preparation for Run 3. The TDAQ system will need to adapt to these upgrades, while also dealing with an environment significantly beyond that of the original design specifications of the system. Namely the TDAQ system will have to cope with higher event rates and increased pileup, as demonstrated in Figure 4.17.

The L1Calo trigger system as described in Subsection 4.5.1 is primarily composed of custom hardware and firmware, the latter of which was updated in preparation for Run 2. In the face of the increased luminosity and pileup of Run 3, the custom hardware will also need to be upgraded to take full advantage of the detector upgrades and the improvements in FPGA technology over the last 25 years. Additionally, the algorithms implemented on this

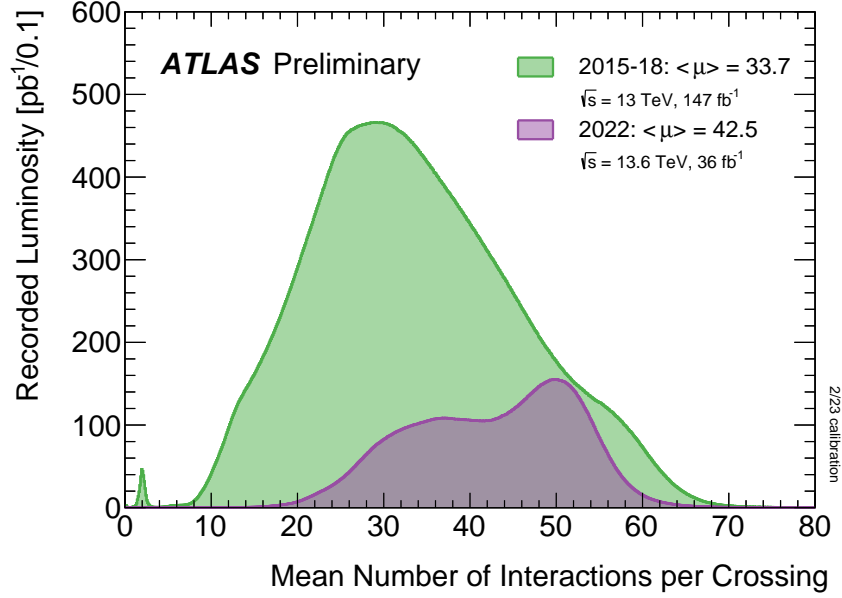


Figure 4.17: Run 2 $\langle\mu\rangle$ in green versus the early Run 3 $\langle\mu\rangle$ from 2022.

hardware can be updated and redesigned to operate as efficiently as possible with the new calorimeter upgrades. Specifically, the electromagnetic calorimeter information from the LAr calorimeter will need to be processed at a higher granularity to preserve trigger efficiency on electroweak physics objects such as electrons, photons, and taus. Improved processing and this higher granularity will allow increased efficiency and better identification of jets in Run 3. The use of a wider search window for jets will also allow larger jets to be identified and the overall topology of events to factor into the L1Calo trigger decision.

The CP and JEP modules as described in Subsection 4.5.1 are being replaced by a totally new suite of Feature Extractor (FEX) modules, with new custom hardware and algorithms. Data from the upgraded LAr calorimeters is now transmitted optically to L1Calo, which processes this data and transmits the output optically to L1Topo. Similarly to the Run 2 system, L1Topo receives this information and transmits it to the L1CTP which produces the L1A signal to indicate to all the subdetectors that data should be read out for the corresponding bunch crossing.

The new FEX modules have been designed to perform identification of physics objects similar to the CP and JEP systems, with increased resolution and efficiency at higher pileup, energies, and luminosities. The new L1Calo trigger system will be composed of the electromagnetic Feature Extractor (eFEX), the jet Feature Extractor (jFEX), and the global Feature Extractor (gFEX), as well as additional upgraded support modules for data transfer and readout. The new modules will be installed alongside the Run 2 L1Calo system for the duration of the integration and commissioning process so that the system can be fully commissioned and calibrated before becoming the sole ATLAS trigger. The full Run 3 system is displayed in Figure 4.18 below.

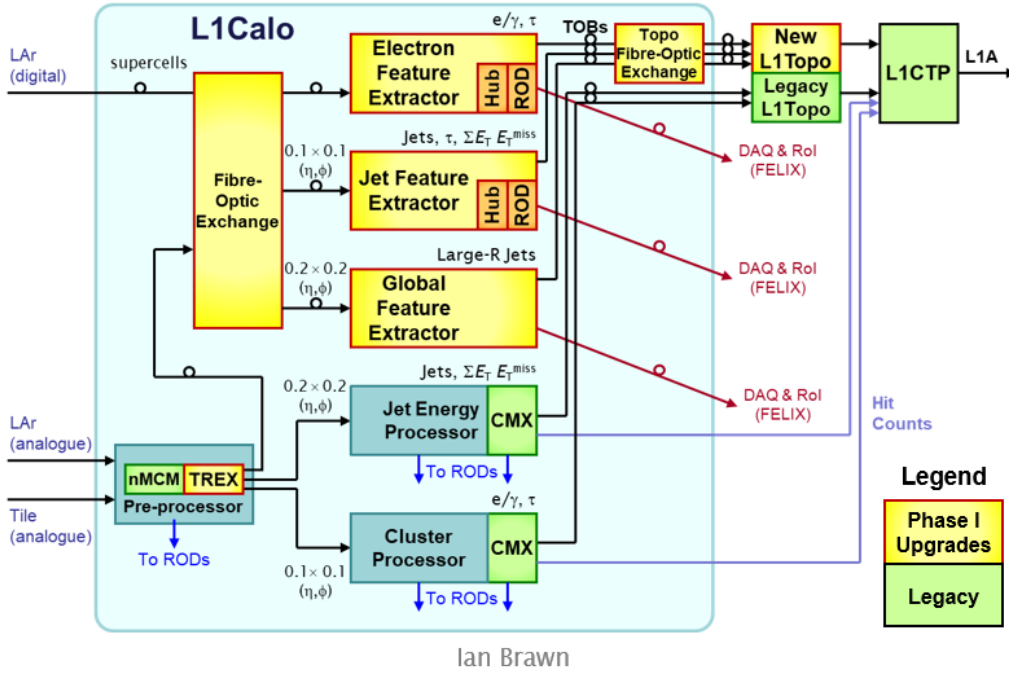


Figure 4.18: L1Calo Schematic showing the flow of data from the calorimeters through to the L1CTP for Run 3.

The eFEX system replaces the CP from the Run 2 system and is composed of 24 modules that utilize the full calorimeter granularity. The upgraded calorimeter granularity ranges from 0.1×0.1 to 0.025×0.1 in $\Delta\eta \times \Delta\phi$ depending on the supercell layer of the calorimeter

that the data is from. eFEX is designed to identify isolated energy clusters that are indicative of an electron, photon, or tau particle.

The jFEX system was designed to replace the JEP system from the Run 2 L1Calo trigger and is composed of 6 modules which currently in Run 3 identify small radius jets, taus, E_T^{miss} , and ΣE_T . The jFEX inputs are calorimeter towers of size 0.1×0.1 in $\Delta\eta \times \Delta\phi$. jFEX uses a sliding window of size 0.9×0.9 to identify local maxima in energy as jets and taus.

The gFEX system does not directly replace any L1Calo module from Run 2, but is instead a completely new system designed to look at an event globally and to identify large radius jets, small radius jets, E_T^{miss} , and ΣE_T . To accomplish this, the gFEX system reads in the data from all ATLAS calorimeters on one single board, with a decreased granularity of 0.2×0.2 in $\Delta\eta \times \Delta\phi$.

In addition to the FEX modules, there are a number of new modules to facilitate the transfer of optical data to L1Calo and through the entire readout chain of the system. These include a new Tile Rear Extension Module (TREM), a Fiber Optic Exchange box (FOX), a Front End Link Exchange (FELIX), and a software read out device (SWROD). The TREM digitizes the Tile calorimeter data, and this data along with digital data from the LAr calorimeter is routed through the FOX from the calorimeters to the correct FEX module. The FELIX receives the optical output from the FEX systems (either directly or via an additional module in some cases), and the SWROD receives data from the FELIX and forwards this data to both the HLT, and ultimately to the CERN Tier 0 data storage facility.

The L1Calo trigger upgrade for Run 3 is one of the biggest ATLAS upgrades performed at this time and is extremely important for the continuation of efficient data taking with the ATLAS detector. At this point in time the upgraded L1Calo system is being commissioned and is in some cases being used as the sole ATLAS trigger. The new FEX systems all

provided triggers for ATLAS during the 2023 proton and heavy ion LHC runs.

CHAPTER 5

GLOBAL FEATURE EXTRACTOR

5.1 Overview

In my graduate work I was heavily involved in almost every aspect of the global Feature Extractor (gFEX) system, outside of the hardware design and construction which were completed before I joined the project in 2018. It was a pleasure and a learning experience to bring a totally new system into existence and taking data with the ATLAS detector.

With the increased pile-up in Run 3 efficient triggering on interesting physics events will continue to be a difficult goal to achieve. The goal of the gFEX system is to target global phenomena by triggering more efficiently on large radius jets that contain interesting substructure, as well as other global quantities like E_T^{miss} . Interesting jet substructure is often seen in boosted objects, for example a boosted Higgs boson or top quark that decay hadronically to additional jets, as shown in Figure 5.1.

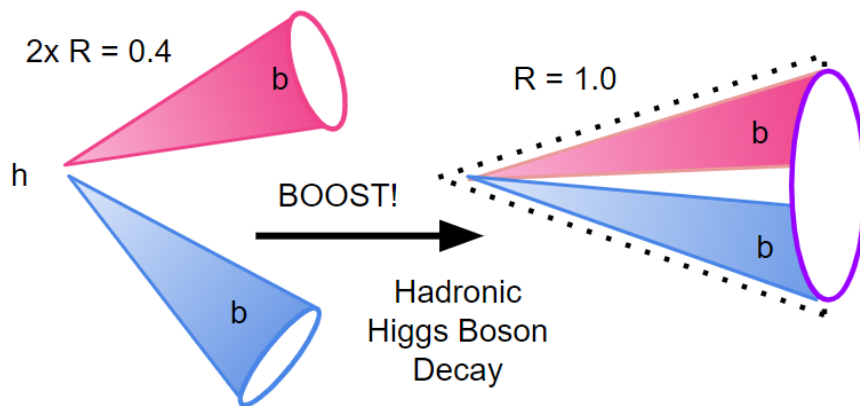


Figure 5.1: Higgs boson hadronic decay with boost

5.2 Hardware Design

The gFEX system is implemented on a single Advanced Telecommunications Computing Architecture (ATCA) board. There are three full production boards total, one that is meant to be installed and in use, and two backups which were designed to be functionally identical. These three production boards are identified as the v4B, v4C, and v4D boards. There is also one partial production board, the v4A board.

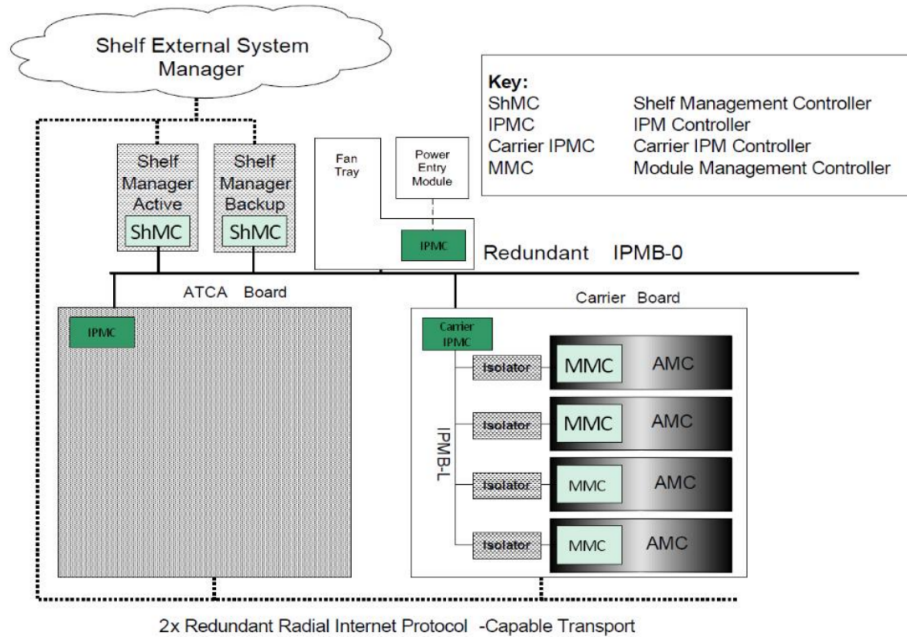


Figure 5.2: ATCA shelf setup with IPMC and Shelf Manager shown [156].

All of these ATCA boards can be inserted into a double slot in an ATCA shelf. The ATCA shelf consists of a shelf manager, backplane, cooling fans, and power entry modules (PEMs). The shelf manager allows for the management of individual modules, fans, and PEMs, while also monitoring the status of these components. The backplane allows for direct connections and signal routing, and can also provide power to individual modules. Each module also has an Intelligent Platform Management Controller (IPMC) that communicates with the shelf manager and helps to monitor the status of the board, as shown in Figure 5.2.

The gFEX board has a number of components but the largest and most important are the Field Programmable Gate Arrays (FPGAs) and the Zynq+ Multi-Processor System on Chip (SoC). These chips are where the majority of the data processing and algorithms are run. In addition to these, there are many peripherals that facilitate the data flow through the board as well as components that regulate the voltage and current and that monitor these values in addition to the temperature. The full gFEX schematic which shows all of these components can be seen in Figure 5.3.

The gFEX system needs a number of external optical fibers for both receiving and sending data. This is implemented through a number of connections on the faceplate of the board which are organized into 8 MTP 48 connectors, one MTP 24 connector and one spare MTP 12 connector. The MTP connectors are a trademarked version of the generic multi-fiber push on (MPO) connectors. Each of these connectors are labeled on the faceplate of the gFEX board with either TX or RX, and the FPGA that the connector is connected to, as shown in Figure 5.4.

Each of the connectors has a small fiber bundle which routes 12 fibers from the interior side of the connector on the faceplate to individual miniPODs. A miniPOD is a high performance fiber optics module that can be used for short-range parallel multi-lane data communication [157]. The miniPODs used on the gFEX board are twelve channel pluggable modules that route signals between the faceplate optical fiber connectors and the appropriate FPGA. There are a total of 35 miniPODs on the board, 26 of which receive external signals and 11 of which transmit signals off of the board. These miniPODs are connected to the MGTs (Multi Gigabit Transceivers) on the FPGAs which then allow for the firmware to decode and send signals to and from the miniPODs. A visual summary of the gFEX FPGA link connections can be seen in Figure 5.5.

The gFEX board has one main power supply module in the form of a DCDC converter.

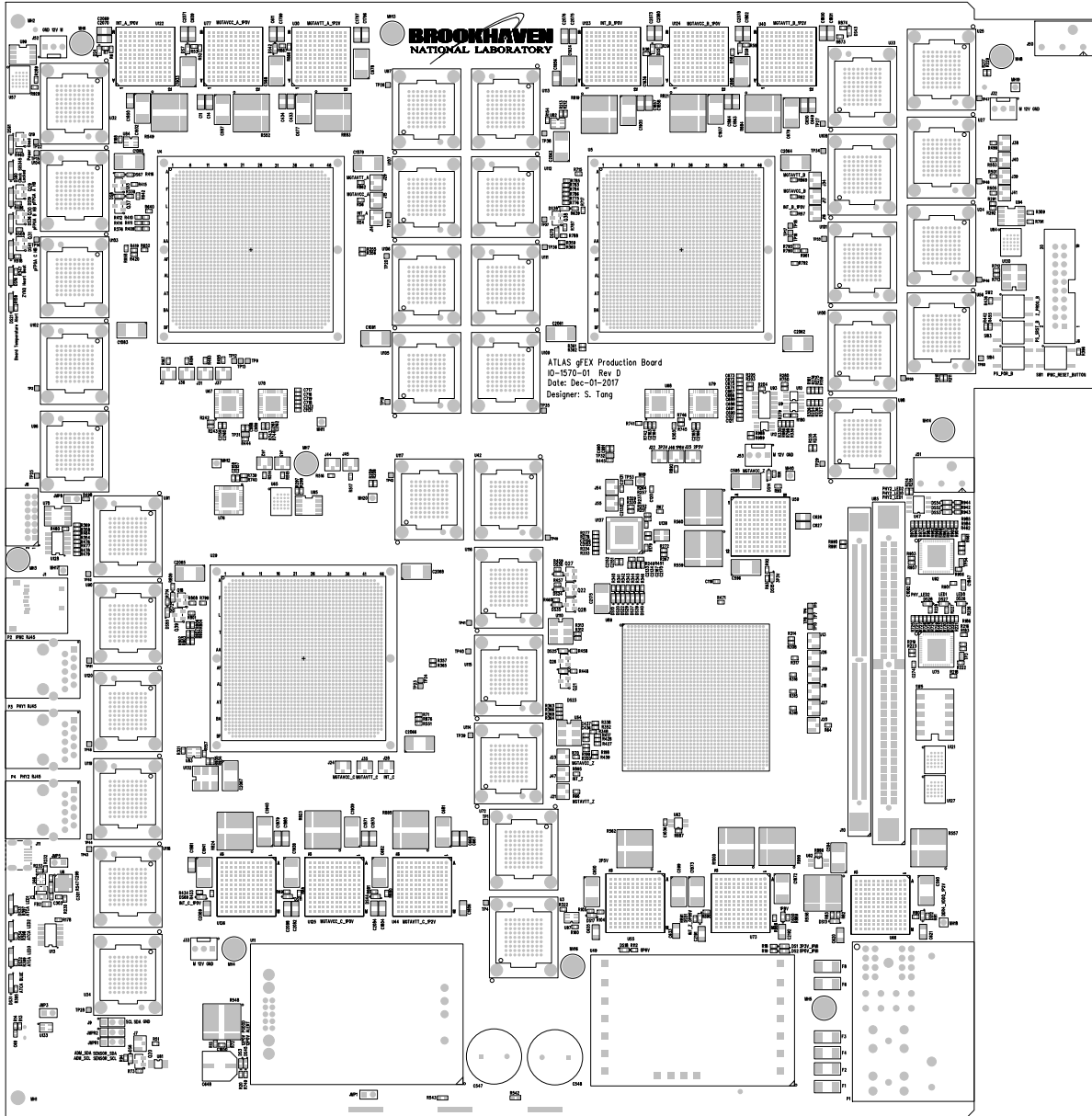


Figure 5.3: Full gFEX schematic [65].

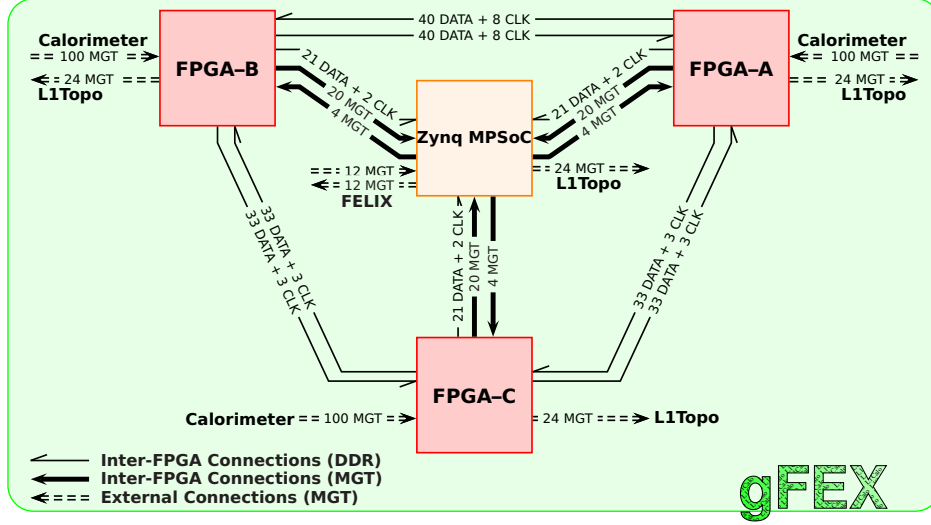


Figure 5.5: gFEX pFPGA link connections [65].

Power is then distributed by thirteen power modules on the board to various other components, including the FPGAs and SoC. Each of the main chips (FPGAs and SoC) has a VCCINT (common collector voltage), MGTAVCC (MGT supply voltage), and MGTAVTT (MGT termination voltage). The MGT voltage types power the MGT rails on the FPGAs. There are a few additional global power modules which provide power to the Zynq+ SoC and to the rest of the board. The power modules also allow for constant monitoring of voltage and current levels to make sure these remain within specified parameters.

5.2.1 Virtex 9 Ultrascale+ FPGA

The gFEX board has three Ultrascale+ VU9P FPGAs which are used for high speed data processing and algorithm implementation [190]. The FPGAs are labeled A, B, and C, and are often referred to processor FPGAs or pFPGAs. These chips are some of the best available commercially and allow for the performance in extreme conditions that is required for the full functionality of the hardware trigger system.

The Xilinx Ultrascale architecture focuses on providing high performance FPGAs and SoCs while lowering total power consumption. The Ultrascale+ are the latest and greatest in FPGA technology and provide the highest transceiver bandwidth, highest digital signal processor (DSP) count, and largest memory available, as shown in Figure 5.6 and Figure 5.7.

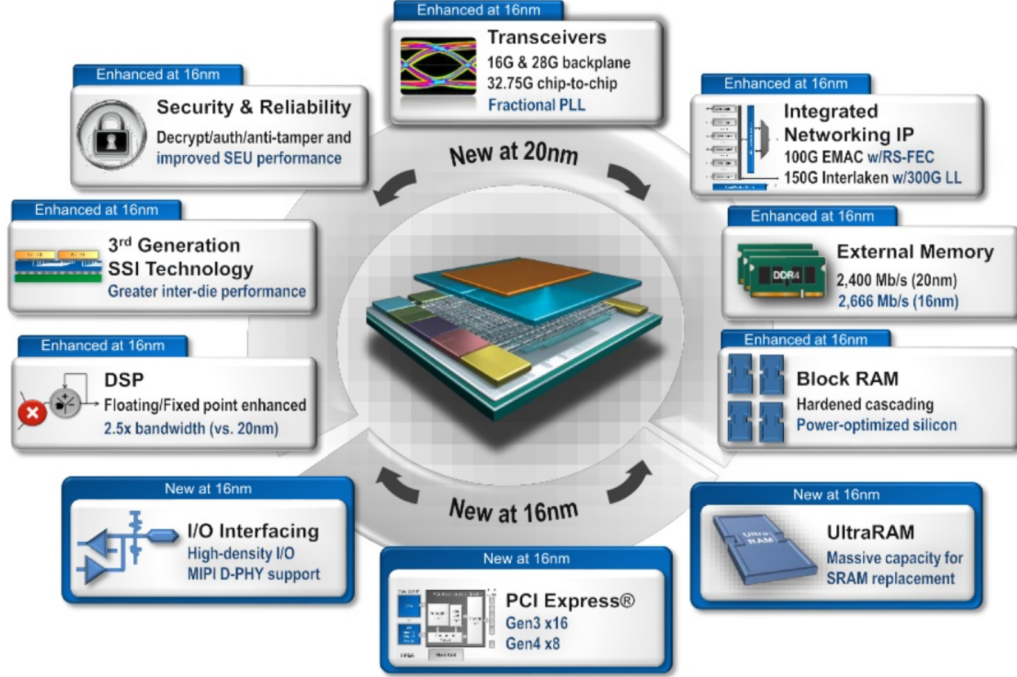


Figure 5.6: Diagram of Virtex FPGA components [198].

These FPGAs are exceedingly important for the functionality of the hardware trigger system, and for gFEX specifically. Their ability to perform fixed latency calculations is central to the implementation of a triggering system where the full system latency is limited. Of additional importance are the many high speed serial transceivers included on each FPGA. These are organized into groups of four known as Quads. These multi gigabit transceivers (MGTs) are the sole path for data into and out of the FPGAs on the gFEX board and are thus integral to the operation of the system, which runs on input data from the calorimeters, and must send output data to subsequent modules in the L1Calo trigger chain.

Foundation											
Device Name	VU3P	VU5P	VU7P	VU9P	VU11P	VU13P	VU19P	58G PAM4			
System Logic Cells (K)	862	1,314	1,724	2,586	2,835	3,780	8,938	2,252	2,835	3,780	
CLB Flip-Flops (K)	788	1,201	1,576	2,364	2,592	3,456	8,172	2,059	2,592	3,456	
CLB LUTs (K)	394	601	788	1,182	1,296	1,728	4,086	1,030	1,296	1,728	
Max. Dist. RAM (Mb)	12.0	18.3	24.1	36.1	36.2	48.3	58.4	14.2	36.2	48.3	
Total Block RAM (Mb)	25.3	36.0	50.6	75.9	70.9	94.5	75.9	74.3	70.9	94.5	
UltraRAM (Mb)	90.0	132.2	180.0	270.0	270.0	360.0	90.0	99.0	270.0	360.0	
DSP Slices	2,280	3,474	4,560	6,840	9,216	12,288	3,840	1,320	9,216	12,288	
Peak INT8 DSP (TOP/s)	7.1	10.8	14.2	21.3	28.7	38.3	10.4	4.1	28.7	38.3	
PCIe® Gen3 x16	2	4	4	6	3	4	0	0	1	1	
PCIe Gen3 x16/Gen4 x8 / CCIX ⁽¹⁾	—	—	—	—	—	—	8	4	—	—	
150G Interlaken	3	4	6	9	6	8	0	0	8	8	
100G Ethernet w/ KR4 RS-FEC	3	4	6	9	9	12	0	2	15	15	
Max. Single-Ended HP I/Os	520	832	832	832	624	832	1,976	572	676	676	
Max. Single-Ended HD I/Os	0	0	0	0	0	0	96	72	0	0	
GTY 32.75Gb/s Transceivers	40	80	80	120	96	128	80	34	32	32	
GTM 58Gb/s PAM4 Transceivers	—	—	—	—	—	—	—	4	48	48	
100G / 50G KP4 FEC	—	—	—	—	—	—	—	2 / 4	24 / 48	24 / 48	
Extended ⁽²⁾	-1 -2 -2L -3	-1 -2 -2L -3	-1 -2 -2L -3	-1 -2 -2L -3	-1 -2 -2L -3	-1 -2 -2L -3	-1 -2	-1 -2 -2L -3	-1 -2 -2L -3	-1 -2 -2L -3	
Industrial	-1 -2	-1 -2	-1 -2	-1 -2	-1 -2	-1 -2	—	-1 -2	-1 -2	-1 -2	
Footprint ^(3,4,5)	Dim. (mm)		HP I/O, GTY				HP I/O, HD I/O, GTY		HP I/O, HD I/O, GTY, GTM		
A1365 ⁽⁶⁾	35x35								364, 0, 34 ⁽⁷⁾ , 4		
C1517	40x40	520, 40									
J1760	42.5x42.5								572, 72, 34, 4		
F1924 ⁽⁸⁾	45x45					624, 64					
A2104	47.5x47.5		832, 52	832, 52	832, 52		832, 52				
	52.5x52.5 ⁽⁷⁾										
B2104	47.5x47.5		702, 76	702, 76	702, 76	572, 76					
	52.5x52.5 ⁽⁷⁾						702, 76				
C2104	47.5x47.5		416, 80	416, 80	416, 104	416, 96					
	52.5x52.5 ⁽⁷⁾						416, 104				
D2104	47.5x47.5				676, 76	572, 76					
	52.5x52.5 ⁽⁷⁾						676, 76		676, 16, 30	676, 16, 30	
H2104	47.5x47.5										
A2577	52.5x52.5			448, 120	448, 96	448, 128			448, 32, 48	448, 32, 48	
A3824	65x65						1976, 96, 48				
B3824	65x65						1664, 96, 80				

Virtex® UltraScale+™ FPGAs

1. This block operates in compatibility mode for 16.0GT/s (Gen4) operation. See [PG213](#).
2. -2LE (T) = 0°C to 110°C. See Ordering Information in DS890.
3. For full part number details, see DS890, *UltraScale Architecture and Product Overview*.
4. All packages are 1.0mm ball pitch, with the exception of A1365, which is 0.92mm.
5. Consult [UG583](#), *UltraScale Architecture PCB Design User Guide* for specific migration details.
6. The GTY transceiver line rate in the F1924 footprint is package limited to 16.3Gb/s. Refer to data sheet for details.
7. These 52.5x52.5mm packages have the same PCB ball footprint as the 47.5x47.5mm packages and are footprint compatible.
8. GTYs in quads 224-230 and 232 are limited to 16Gb/s.



Figure 5.7: Comparison of UltraScale+ FPGA resources [198].

5.2.2 Zynq+ System on Chip

The Zynq UltraScale+ Multi-Processor System on Chip (SoC) is a chip that combines a quad-core Arm Cortex-A53 64-bit processor with the Arm Cortex-R5F real-time processor, the Mali-400 MP2 GPU, and the UltraScale+ FPGA architecture to create a multi-processor system on chip that has all the functionality of both a FPGA system and processing system as shown in Figure 5.8. The chip also has a number of peripheral devices with dedicated functions including a multi-protocol dynamic memory controller, a direct memory access controller, a NAND controller, an SD/eMMC controller, and a quad SPI controller. The Processing System (PS) also includes 4 channels of transmit (TX) and receive (RX) transceivers with a maximum data rate of 6.0 Gb/s. The PS also includes two USB 2.0 controllers, an I2C controller, a UART, and a CAN2.0B controller. There are also 2 triple speed Ethernet MACs which can be used to access the PS remotely.

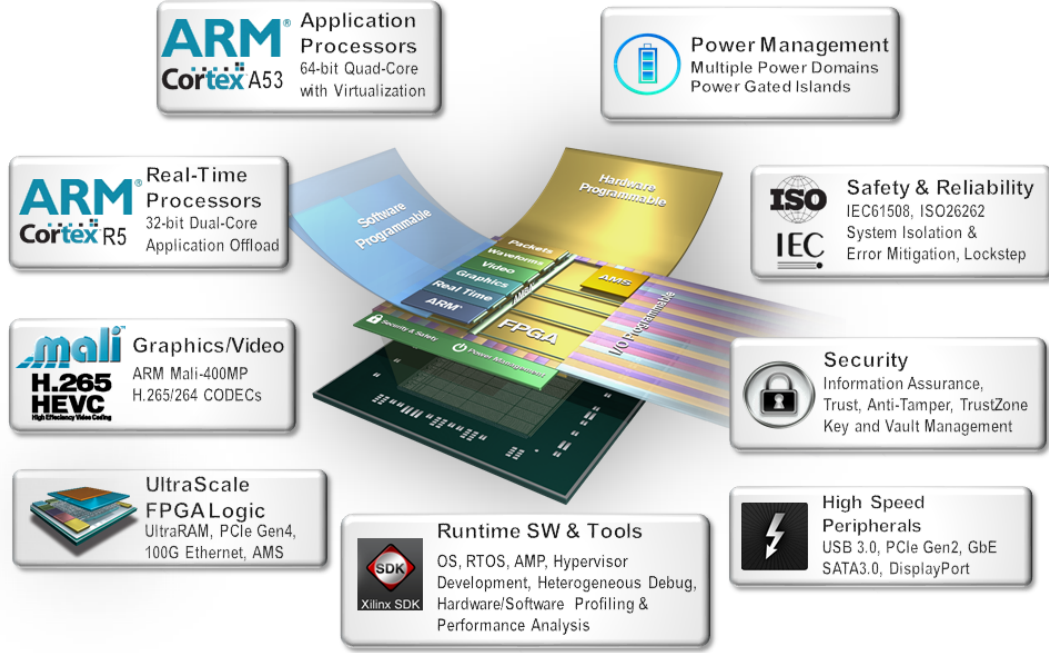


Figure 5.8: Zynq+ System on Chip architecture [198].

The combination of this high speed, cutting edge technology and the innovative integration of hardware results in a chip that allows for the fixed-latency processing power of a FPGA combined with the flexibility and ease of implementation that a processing system offers, which is perfect for the environment of the hardware trigger modules.

One of the major benefits of the processing system is the ability to implement a custom operating system that helps add flexibility and opportunity for control and monitoring of the SoC itself and the wider hardware environment of the full board.

5.3 Firmware Infrastructure

The global Feature Extractor was designed to identify global physics objects like large radius jets, small radius jets, E_T^{miss} , and $\sum E_T$. There are two main algorithms which produce these objects: Jet Finder, and MET finder. There is also the gTower Builder which builds the

inputs to these algorithms. The details of Jet Finder and MET Finder will be discussed in more detail in Section 5.4.

Besides these algorithms which compute physics quantities, there are a number of different firmware pieces that manage things like clock signal handling, link management, input signals and output signals, and the decoding and response to communication from the central ATLAS control mechanism.

The gFEX derives its clock signal from the FELIX. The FELIX data header is used to derive a clock signal which is routed through the SI5345 chip for jitter cleaning and then used to derive the various clock frequencies needed for the gFEX firmware. A visual representation of the gFEX clock tree can be seen in Figure 5.9. The FELIX signal is also used to derive a number of other values which are used in various ways, described below:

- BCR (bunch crossing reset) - this signal indicates that the bunch crossing counter should be reset from 3563 to 0.
- ECR (event counter reset) - this signal indicates that the counter of the number of L1As should be reset by clearing the 24 bit L1A counter such that it returns to zero.
- L1A (Level-1 Accept) - the Level-1 Accept signal is generated by the L1CTP and indicates that every module should read out the data from the associated BCID.

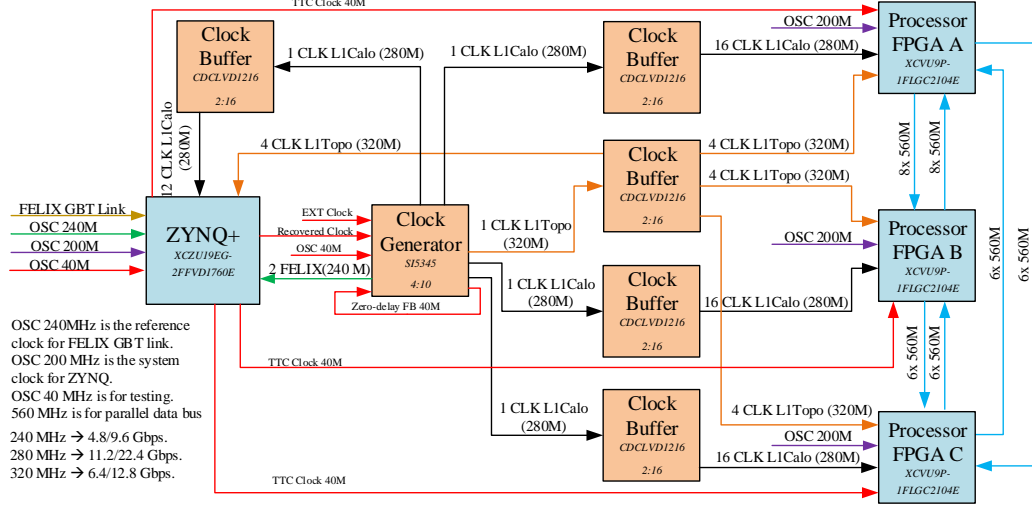


Figure 5.9: gFEX full clock tree [65].

5.3.1 Fiber Control & Monitoring

The input fibers have a number of registers dedicated to monitoring and control. It is important that all fibers are locked onto the input signal mainly through the use of the synchronization character or *comma character*¹ which must be detected in the correct place for the link to be lock onto the input signal. All fibers must be checked for a good status and good timing, and any unused fibers must be masked to avoid nonphysical noise contaminating the algorithm calculations. Additionally all input links must be aligned to the same clock tick and bunch crossing to ensure that the physics algorithms are processing the correct data from the same collision event. Table 5.1 lists the registers exposed to the user that are relevant for fiber control and monitoring. In addition to these registers, there are a number of other registers which extract relevant information from the input data and alignment frames and can be used to debug various issues [189].

1. In 8b/10b encoding, 8-bit words are mapped to 10-bit symbols. The K28.5 character is referred to as a comma character and is used for synchronization and alignment of the 8b/10b codes in a bit-stream. In hex this symbol is BC, in decimal 188.

Register	Number	Bits/fiber	Type	Description
MGT_ALGN	4	1	read-only	Fiber lock status
CLK_ALGN	25	8	read-only	Fine alignment counter
MGT_CRCERR	25	8	read-only	CRC error counter
MGT_LINKERR	25	8	read-only	Link error counter
MGT_MASK	4	1	read/write	MGT mask status
CLK_DLY	25	8	read/write	Fine alignment delay (in clock ticks)
BCID_DLY	25	8	read/write	Coarse alignment delay (in clock ticks, 7 per BC)
FIB_DSBL	1	-	read/write	One bit to enable fiber input
ERR_CTRL_CLR	4	1	read/write	Toggle to clear error counters
BCID_INIT	1	-	read/write	BCID initialization of masked fibers

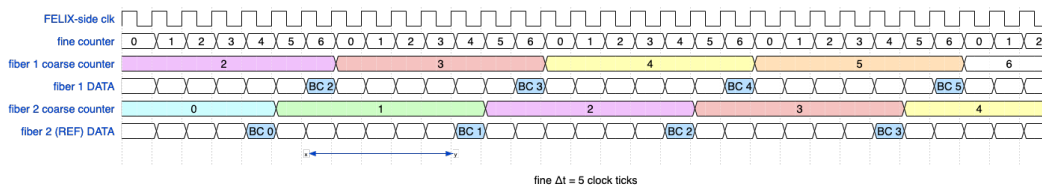
Table 5.1: Fiber Control and Status Registers

The MGT_ALGN register displays the lock status for all input fibers. Any fibers that are not locked need to be masked by writing a 1 to each bit in the MGT_MASK register that corresponds to the fiber that should be masked. This includes the spare fibers which have no inputs, and show a bad status in the appropriate bits of the MGT status. After bad fibers are masked, the input fibers must be aligned properly.

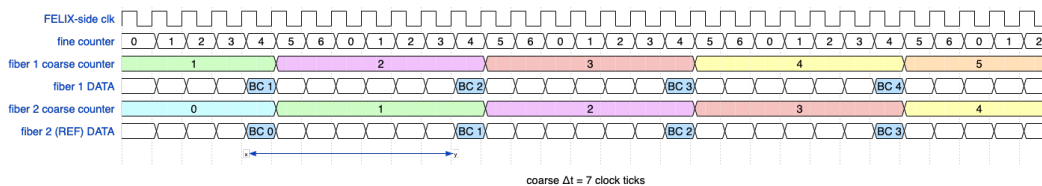
Link errors including CRC and disparity errors (8b10b out-of-table errors) are counted in MGT_CRCERR and MGT_LINKERR. These counters can be reset by toggling the corresponding bit in the ERR_CTRL_CLR.

The gFEX input signals consist of 7 32-bit data words per bunch crossing on each fiber. The gFEX firmware needs all of these input fibers to be aligned to the same BCID to function properly. However, the disparate signal paths from the various parts of the calorimeter introduce timing differences, and the the MGTs also have an uncertainty associated with the the latency that could affect the fiber timing. An example of a possible input timing can be seen in Figure 5.10a. The gFEX firmware counts the number of 280 MHz clock ticks

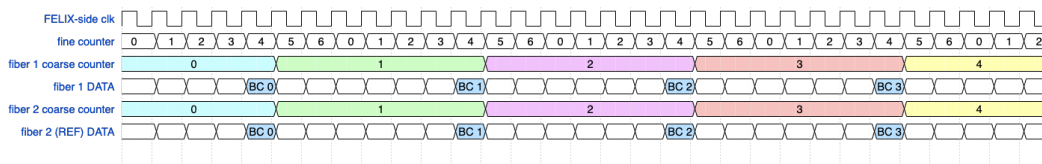
until the end of the data frame (indicated by the K28.5 comma character) and stores this number in `CLOCK_ALGN`. A BCID is calculated from the data frame and stored in `BCID_ALGN`. These values can be compared across all input fibers, and the minimum value found, which corresponds to the slowest fiber. All other fibers can be delayed to match the slowest fiber by writing the difference (calculated in the number of clock phases) to the `CLOCK_DLY` register. Figure 5.10b shows the signals after this fine alignment (also known as clock alignment) has been performed. The same alignment procedure is performed using the BCID alignment values, and writing the differences (in terms of clock ticks, 7 280 MHz clock ticks to one BC) to the `BCID_DLY` registers. After this coarse alignment procedure (also known as BCID alignment), the alignment is considered to be complete, and all signals should be aligned to the bunch crossing boundary and to the same BCID as shown in Figure 5.10c. At this point the fiber input to the algorithms can be enabled by changing the state of bit 0 of `FIB_DSBL` (the value of the bit does not matter, only the change in either direction).



(a) Input fiber clock sequence before any alignment.



(b) Input fiber clock sequence after the fine alignment to the clock boundary.



(c) Input fiber clock sequence after the coarse alignment to the correct BCID.

Figure 5.10: gFEX input fiber alignment. Diagram Credit: Greg Myers.

After the input fibers have been aligned and data processing has begun in the algorithms, the fibers from the pFPGAs to the Zynq FPGA (zFPGA) need to also be aligned. There are 2 jet and 2 E_T^{miss} links from each pFPGA. The process is similar to that of the input fiber alignment and uses several additional registers to hold relevant values. The E_T^{miss} output here must all be aligned together and the two jet links from each pFPGA must be aligned with each other. After this alignment is performed, the data is stored in a circular buffer (one per pFPGA \rightarrow zFPGA link) addressed by BCID. When an L1A is received, the relevant event is read from this circular buffer and sent to FELIX. The buffer depth is the size of a full orbit of data to prevent the wrong event from being read out. This allows the gFEX to read out up to 255 bunch crossings at once if the L1As are spaced out enough to allow it, which helps tremendously with the commissioning process.

The gFEX data is read out using 12 MGTs connected to FELIX. These links are often referred to as "elinks" on the FELIX side, and each gFEX elink carries a specific type of data corresponding to one of the gFEX trigger objects, as shown in Table 5.2 The FELIX forwards this data on to the Software Rod which combines all the different trigger objects into a full event. This event is transmitted to the high-level trigger, where it is processed by the HLT software running on generic linux PCs.

zFPGA Name	Data Type	Mnemonic	Link ID	Elink ID	Link Num.
ttc_tx_data[1]	Global TOB A	GBLA	0x6B1A	0x1C0	7
ttc_tx_data[2]	Jet TOB C	TOBC	0x20BC	0x180	6
ttc_tx_data[3]	Jet TOB B	TOBB	0x20BB	0x140	5
ttc_tx_data[4]	Jet TOB A	TOBA	0x20BA	0x100	4
daq_tx_data[1]	Input Data B	TWRB	0xFEDB	0x280	10
daq_tx_data[2]	Input Data A	TWRA	0xFEDA	0x2C0	11
daq_tx_data[3]	Global TOB C	GBLC	0x6B1C	0x240	9
daq_tx_data[4]	Global TOB B	GBLB	0x6B1B	0x200	8
daq_tx_data[5]	Spare	SPR3	N/A	0x000	0
daq_tx_data[6]	Spare	SPR2	N/A	0x040	1
daq_tx_data[7]	Input Data C	TWRC	0xFEDC	0x080	2
daq_tx_data[8]	Spare	SPR1	N/A	0x0C0	3

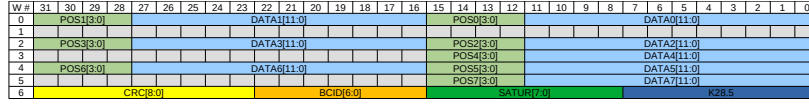
Table 5.2: gFEX to FELIX elinks and their several associated naming schemes. Table Credit: Greg Myers.

5.4 Firmware Algorithms

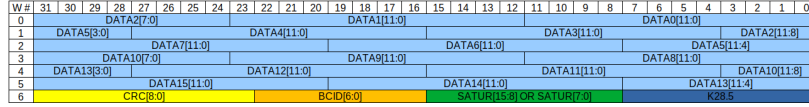
There are two main algorithms implemented in the gFEX firmware. The E_T^{miss} algorithm Jets without Jets [69], and the large-radius jet algorithm Jet Finder. These are both preceded by the gTower Builder, which is less algorithmic in nature, but is included in the descriptions below as it is an important firmware component which contributes to overall latency and data processing ability. The full firmware block diagram can be seen in Figure 5.11.

5.4.1 Tower Builder

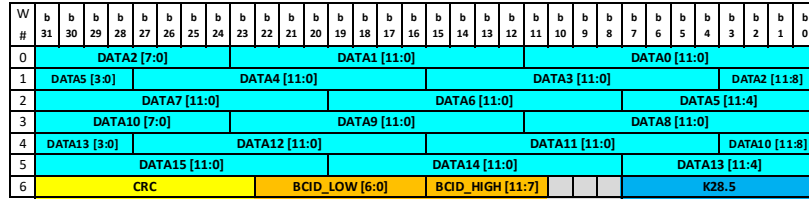
The tower builder firmware is external to the algorithm firmware and functions mostly independently of the control of the user. This firmware is implemented on the pFPGAs, each of which process a subset of the input calorimetry data from a specific η range, as shown in Figure 5.15. The input data format from the Liquid Argon and TILE calorimeters respectively is shown in Figure 5.12. The tower builder decodes the input data with this format, and utilizes the necessary bits denoting energy and saturation to construct the gTowers, which



(a) LATOME EMB/EMEC data format.



(b) LATOME EMEC/HEC data format.



(c) TREX data format

Figure 5.12: gFEX input data formats.

Register	Type	Description
DATE_CODEM	read-only	Most significant date code bits
DATE_CODEL	read-only	Least significant date code bits
HASH_CODE	read-only	Gitlab commit hash
FW_VER_CODE	read-only	Firmware version code
VIVADO_VERSION	read-only	Vivado version of build
MAIN_CSR	read/write	Configure pileup correction, inter-fpga coms, multilinear encoding, run mode selection
RUN_LEN	read/write	Number of BCs in playback run
SEED_THRSH	read/write	Threshold required for gBlock seed to build a gJet
JET_THRSH	read/write	Threshold for gJet TOB
GBLK_THRSH	read/write	Threshold for gBlock TOB
RST_BITS	read/write	[0] Run enable bit which controls FSM according to MAIN_CSR settings

Table 5.3: Jet Finder Registers

MAIN_CSR Bits	Name	Functionality
[2]	PUC_CLR	zero pileup correction to jet engines (1), normal operation (0)
[3]	IFC_MODE	disable inter-fpga (1), normal operation (0)
[4]	EN_FIB	enables fibers (1), redundant, may not be in use
[5]	LCONV_EN	enable multi-linear decoding (1), disable multi-linear decoding (0)
[6]	TOP_CAP_EN	can be toggled 0 \rightarrow 1 during some playback runs
[7]	BCID_TRIGGER	can be toggled 0 \leftrightarrow 1 to capture output TOB spy memory
[11 : 8]	RUN_MODE_SEL	0000: RUN_WAIT_LAR , wait for lar_dat_eof signal 0001: RUN_NO_LAR , normal run without lar_dat_eof signal 0010: RUN_PLAYBACK , playback run with generated η values 0011 : wait for lar_frm_bcr, then normal run 0100 : BCID Triggered Normal Run, sync to gtb_eta clock, then do TOB spy memory capture of set BCID 0101 : playback, synchronize to gtb_eta, wait for eta_ind = 6, then run until specified limit set by RUN_LEN register is reached

Table 5.4: Jet Finder MAIN_CSR Bit Definitions

		eta bound:																																													
		4.9	4.5	3.9		3.5	3.3	3.1	2.9	2.7	2.5		2.2	2.0	1.8	1.6	1.4	1.3	0.0	0.8	0.6	0.4	0.2	0.2	0.6	0.4	0.8	1.0	1.2	1.4	1.8	1.8	2.0	2.2		2.5	2.7	2.9	3.1	3.3	3.5		3.8		4.5		4.9
		eta index:																																													
		0	0	1		1	2	3	4	5	0	1	2	3	4	5	6	7	8	9	10	11	0	1	2	3	4	5	6	7	8	9	10	11	6	7	8	9	10	11		11					
ghi bound	ghi index																																														
0.00	0																																														
0.20	1																																														
0.39	2																																														
0.59	3																																														
0.79	4																																														
0.98	5																																														
1.18	6																																														
1.37	7																																														
1.57	8																																														
1.77	9																																														
1.96	10																																														
2.18	11																																														
2.36	12																																														
2.55	13																																														
2.75	14																																														
2.95	15																																														
3.14	16																																														
3.34	17																																														
3.53	18																																														
3.73	19																																														
4.12	20																																														
4.32	21																																														
4.52	22																																														
4.71	23																																														
4.91	24																																														
5.11	25																																														
5.30	26																																														
5.50	27																																														
5.69	28																																														
5.89	29																																														
6.09	30																																														
6.29	31																																														
6.49	32																																														
6.69	33																																														
6.89	34																																														
7.09	35																																														
7.29	36																																														
7.49	37																																														
7.69	38																																														
7.89	39																																														
8.09	40																																														
8.29	41																																														
8.49	42																																														
8.69	43																																														
8.89	44																																														
9.09	45																																														
9.29	46																																														
9.49	47																																														
9.69	48																																														
9.89	49																																														
10.09	50																																														
10.29	51																																														
10.49	52																																														
10.69	53																																														
10.89	54																																														
11.09	55																																														
11.29	56																																														
11.49	57																																														
11.69	58																																														
11.89	59																																														
12.09	60																																														
12.29	61																																														
12.49	62																																														
12.69	63																																														
12.89	64																																														
13.09	65																																														
13.29	66																																														
13.49	67																																														
13.69	68																																														
13.89	69																																														
14.09	70																																														
14.29	71																																														
14.49	72																																														
14.69	73																																														
14.89	74																																														
15.09	75																																														
15.29	76																																														
15.49	77																																														
15.69	78																																														
15.89	79																																														
16.09	80																																														
16.29	81																																														
16.49	82																																														
16.69	83																																														
16.89	84																																														
17.09	85																																														
17.29	86																																														
17.49	87																																														
17.69	88																																														
17.89	89																																														
18.09	90																																														
18.29	91																																														
18.49	92																																														
18.69	93																																														
18.89	94																																														
19.09	95																																														
19.29	96																																														
19.49	97																																														
19.69	98																																														
19.89	99																																														
20.09	100																																														
20.29	101																																														
20.49	102																																														
20.69	103																																														
20.89	104																																														
21.09	105																																														
21.29	106																																														
21.49	107																																														
21.69	108																																														
21.89	109																																														
22.09	110																																														
22.29	111																																														
22.49	112																																														
22.69	113																																														
22.89	114																																														
23.09	115																																														
23.29	116																																														
23.49	117																																														
23.69	118																																														

79

Each pFPGA is divided up into 64 "jet engines" which each handle a single ϕ bin, and 6 of the 12 possible η bins in that ϕ bin. The Jet Finder algorithm calculates all possible gBlocks first, which are 3×3 sums of gTowers². The four highest p_T gBlocks are output as small radius jet TOBs from the algorithm, an example of which is shown in Figure 5.14. Next a gJet is calculated as a sum of 69 gTowers for each gBlock passing the seed threshold, which can be seen in Figure 5.15. The seed threshold can be set in the SEED_THRSH register, and the default value is zero, such that a gJet is built for every existing gBlock. During this summing process, information on the borders of the pFPGA regions is shared between pFPGAs using inter-FPGA communication over serial lines; this is the main functionality of the pFPGA C jet engines as jets are not built for the forward region. The one highest p_T gJet and two highest p_T gBlocks from each jet engine (2 and 4 per pFPGA respectively) are the outputs of the algorithm. These output TOBs are sent directly over the high-speed pFPGA MGTs to L1Topo, and internally to the Zynq FPGA to be read out using FELIX. Jet Finder also provides the already computed gBlocks to the Jets without Jets algorithm.

The Jet Finder algorithm has the capability to apply pileup correction to the gTower inputs in the form of ρ -subtraction. ρ is calculated on an event-by-event basis by taking the average E_T of the gTowers in rings of consistent η , shown in Equation 5.1.

$$\rho = \frac{\sum_{i \in gTowers} E_T^i}{\sum_{i \in gTowers} Area} \quad (5.1)$$

This ρ is applied such that $E_T^{i,\rho} = E_T^i - A_i \rho$, and $E_T^{i,\rho}$ is saved as the new gTower energy and used as needed in the Jet and E_T^{miss} finding algorithms. A few simulations of η dependent ρ calculations can be seen in Figure 5.16.

2. Except on the boundaries of the pFPGAs, where the gBlocks are computed as 3×2 gTowers.

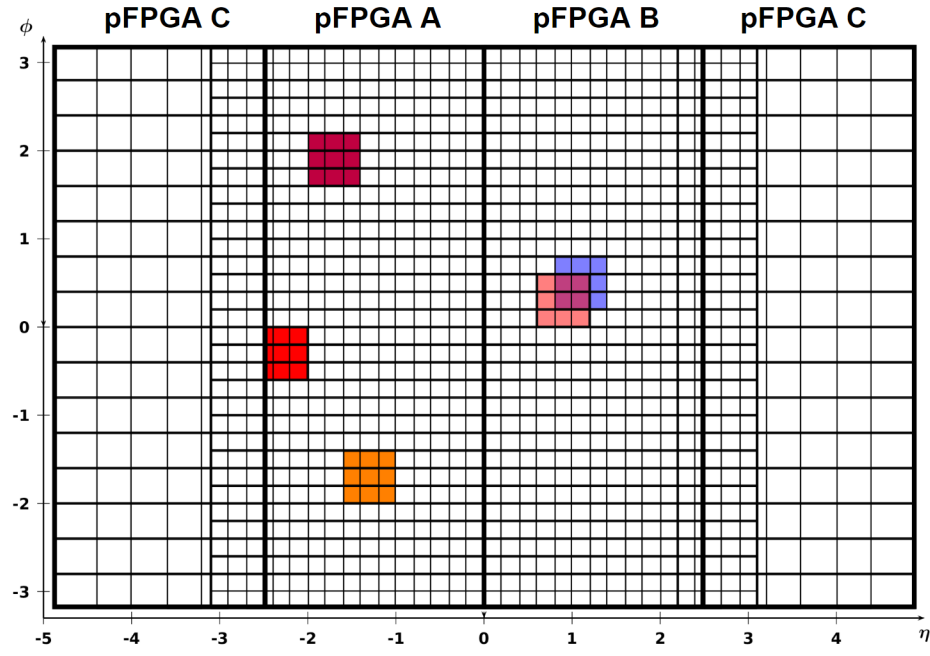


Figure 5.14: gFEX gBlock Event Display. Image Credit: David Strom and David Miller.

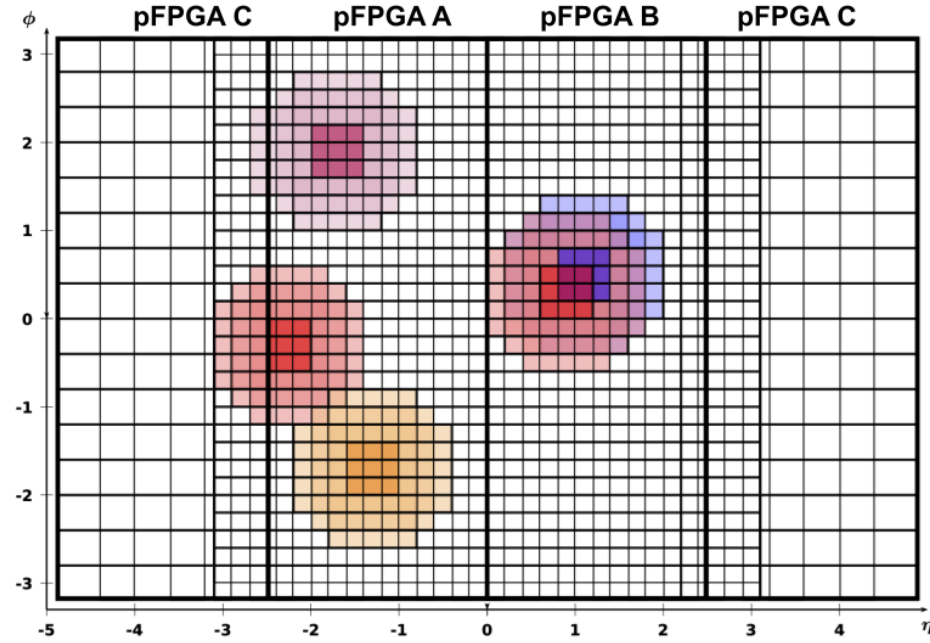


Figure 5.15: gFEX gJet Event Display. Image Credit: David Strom and David Miller.

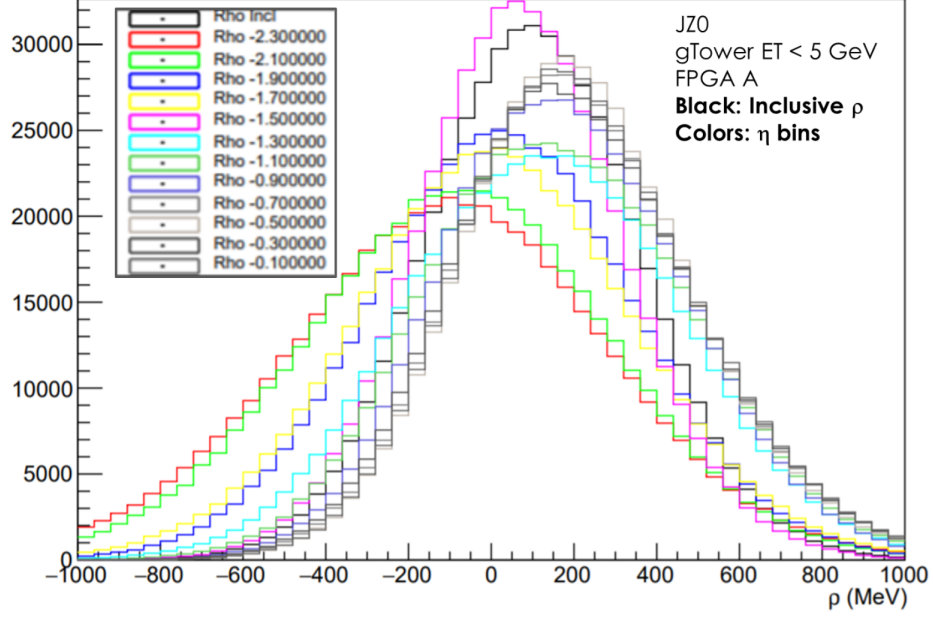


Figure 5.16: η dependent ρ for gFEX pileup subtraction. Plot credit: Aparajita Dattagupta.

5.4.3 Missing Transverse Energy Finder

The MET Finder algorithm on gFEX is actually composed of three separate algorithms. The gFEX output data protocol can only fit the output of two algorithms total. The baseline algorithm for MET finding on gFEX is implemented as Jets without Jets (JwoJ) [69]. JwoJ considers the E_T of each gTower as well as the corresponding gBlock E_T and follows the following procedure to compute the E_T^{miss} [188]:

1. Process every gTower and the corresponding gBlock centered on a given gTower.
2. if $E_T^{gBlock} > E_T^{\text{threshold}}$ then add E_T^{gTower} to the hard term MHT.
3. if $E_T^{gBlock} < E_T^{\text{threshold}}$ then add E_T^{gTower} to the soft term MST.
4. MET x and y components are constructed from MHT and MST using a look-up table (LUT) which stores the cosine and sine values at each azimuthal angle ϕ

5. Compute a weighted sum of the soft and hard terms: $E_{Tx,y} = a \cdot \text{MHT}_{x,y} + b \cdot \text{MST}_{x,y}$ where weights a and b are different for each pFPGA but the same for the x and y components. Send x and y components to the zFPGA.
6. In the zFPGA, sum all pFPGA terms for x and y components and compute the full missing transverse energy: $E_{\text{T}}^{\text{miss}^2} = E_{Tx}^2 + E_{Ty}^2$
7. Send the final $E_{\text{T}}^{\text{miss}}$ values to L1Topo and to FELIX for readout.

The factorization of the $E_{\text{T}}^{\text{miss}}$ into MHT and MST separates out pileup and noise from the hard term, and the weights a and b allow for calibration and tuning to better match offline quantities.

The two additional MET algorithms are noise-cut MET and ρ +RMS MET. The noise-cut MET algorithm calculates the noise using the RMS of the gTower E_{T} in minimum bias events. Modified JwoJ firmware blocks are used to calculate the vector sum of all gTower x and y components of E_{T} with $\text{noise} < 4\sigma$. A final MET value is then computed in the same way as for JwoJ [188]. The noise-cut MET algorithm is implemented in the firmware, and is included in the data frame that is sent out from gFEX.

The ρ +RMS MET has not been implemented in the firmware yet. For this algorithm the pileup corrected gTowers and gBlocks are used, and a 3σ noise cut is applied and $E_{\text{T}}^{\text{miss}}$ calculated as a vector sum of gTower E_{T} , as described for the noise-cut MET algorithm [188].

5.5 gFEX Trigger Efficiencies

Several studies of the simulated gFEX trigger efficiencies utilizing the planned gFEX trigger algorithms (described above) have been done over the years. The original study by Giordon

Stark (Figure 5.17) showed that the simulated gFEX trigger was significantly more efficient than the Run 2 Level-1 jet trigger for jets with more than one subjet. This is particularly important for many boosted physics objects such as a boosted Higgs ($h \rightarrow b\bar{b}$), or boosted top quark ($t \rightarrow Wb$).

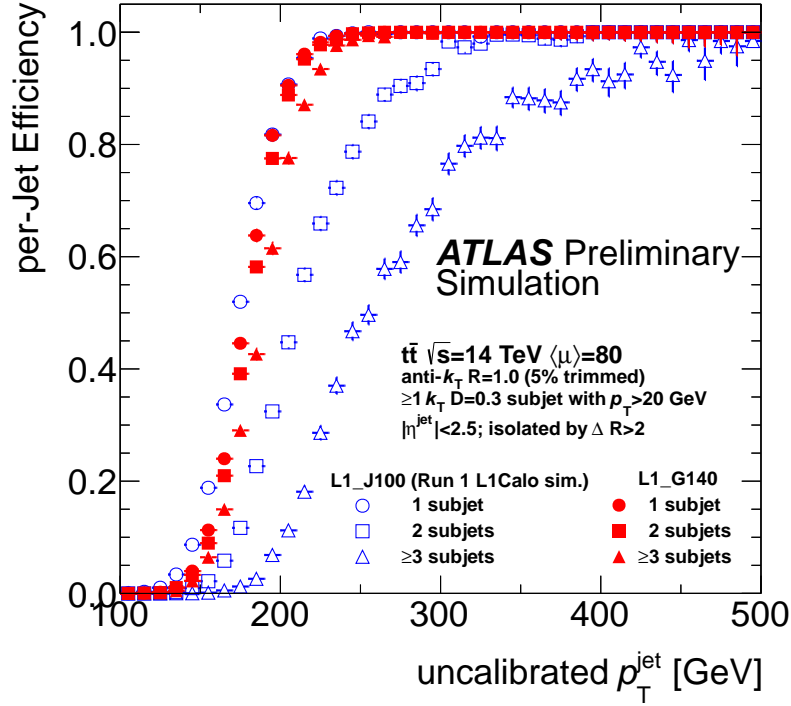


Figure 5.17: Early gFEX Trigger Efficiency Study [95].

I performed a very similar study with updated samples and a finalized calorimetry timing scheme which also focused on the gFEX large radius jets trigger efficiency, as well as the rates with various gFEX jet thresholds. I also looked at the efficiency as a function of the subjet multiplicity, as well as the general relationship of the p_T and energy response of the offline jets and the gJets and gBlocks. Generally these studies were performed using `OfflineAntiKT10` jets, which have a radius of $R=1.0$.

The gFEX trigger efficiency for these studies is defined as:

$$\epsilon = \frac{N_{Trigger}^{Offline}}{N_0^{Offline}} \quad (5.2)$$

where $N_{Trigger}^{Offline}$ is taken to be the distribution of offline jets with selections applied to the corresponding gJet, binned in the jet p_T , and $N_0^{Offline}$ is the full p_T distribution of the offline jets.

The trigger rate is defined as shown in Equation 5.3.

$$rate = L_{instantaneous} \times \sigma \times \frac{N_{Trigger}}{N_0} \quad (5.3)$$

where the instantaneous luminosity is taken to be $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, σ is the production cross section, $N_{Trigger}$ is the number of events passing gJet selections, and N_0 is the total number of events produced with cross section σ . The rate is important because there is a finite Level-1 bandwidth that gFEX must share with the other L1Calo submodules. The maximum Level-1 rate that the HLT can handle is 100 kHz in Run 3. The rate is also important for comparisons of efficiency between two systems; if the rate is drastically different it is hard to directly compare the efficiencies because these systems have such different resource usage. The realities of actually running the trigger system are import to consider while performing this type of efficiency study.

The signal samples used for this study were $Z^* \rightarrow Zh$, $Z \rightarrow \nu\nu$, $h \rightarrow b\bar{b}$, as shown in Figure 5.18. This allows for jet studies and MET studies to be performed using the same signal samples. For the rate calculations, p_T sliced dijet samples were used as minimum bias samples to represent the background rate that would result from a given trigger selection scheme.

The sample name is:

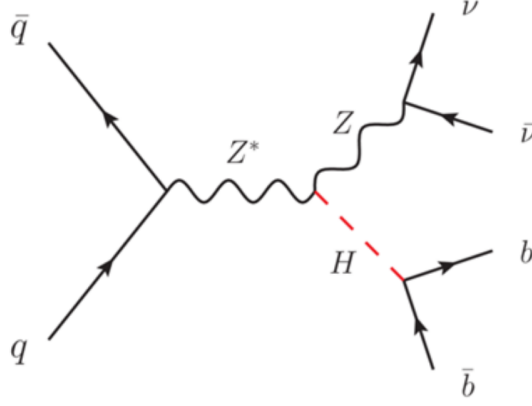


Figure 5.18: gFEX Efficiency Study Samples

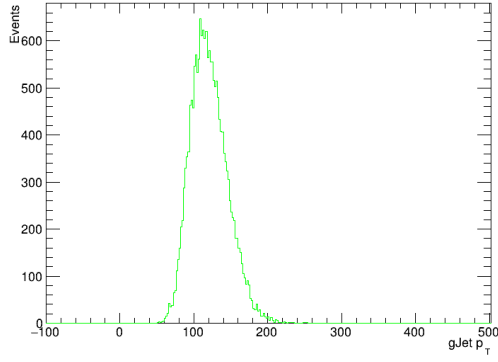
user.cylin.341101.Pythia8EvtGen_A14NNPDF23L0_ZvvH125_bb.e5902_s3142_s3143_r9700.mu60.v02_OUTPUT

Figure 5.19 shows some general plots of the relationship between the `OfflineAntiKT10Jets` and the `gJets`, including distributions of the `gJet` and `Offline jet p_T` , and the energy response of these `gJets` as a function of the jet p_T , where the energy response is defined as in Equation 5.4.

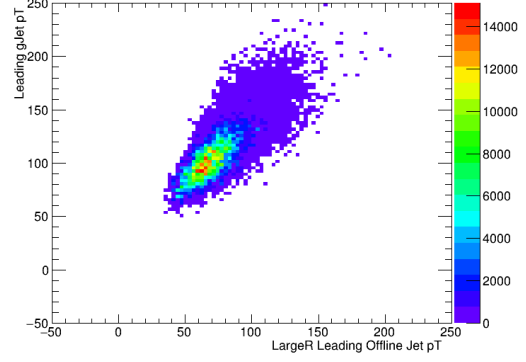
$$\text{energy response} = \frac{p_T^{gJet}}{p_T^{Offline}} \quad (5.4)$$

Equation 5.4 shows that `gJets` tend to slightly overestimate the p_T of the offline jets, though slightly less so for higher p_T jets. The resolution of this response is not bad considering that the `gFEX` input data resolution is quite large compared to most systems, since all data from the entire calorimeter is processed.

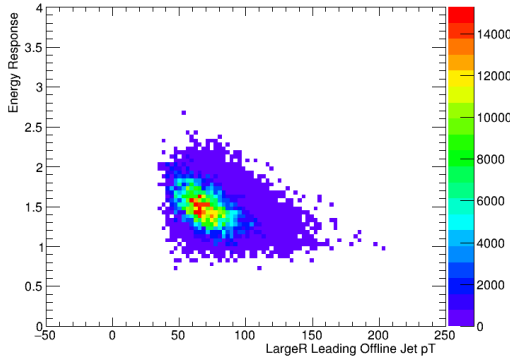
One of the biggest results of this study is a comparison of the `gFEX` trigger efficiency to the Run 2 trigger efficiency, shown in Figure 5.20. This efficiency is calculated with specific subjet requirements placed on the offline jet. For all choices of the subjet requirement, the `gFEX` trigger is more efficient, significantly more so as the number of subjets increases. The



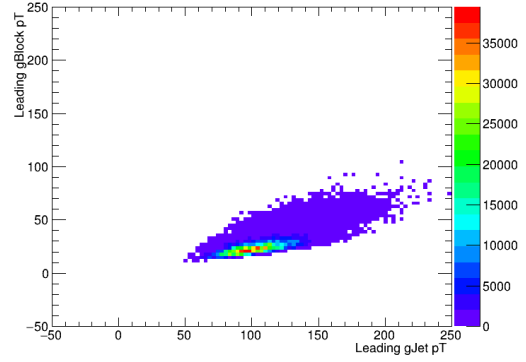
(a) gJet p_T



(b) gJet p_T versus offline jet p_T



(c) gJet energy response vs offline jet p_T



(d) gJet p_T versus gBlock p_T

Figure 5.19: Kinematic relationships between gJets, gBlocks, and Offline jets.

choice of gJet threshold was chosen such that the rates are similar between the gFEX trigger and the Run 2 trigger object, L1_J100. As the gFEX gJet threshold is lowered, of course the efficiency improves for the gFEX triggers, and the rate tends to increase.

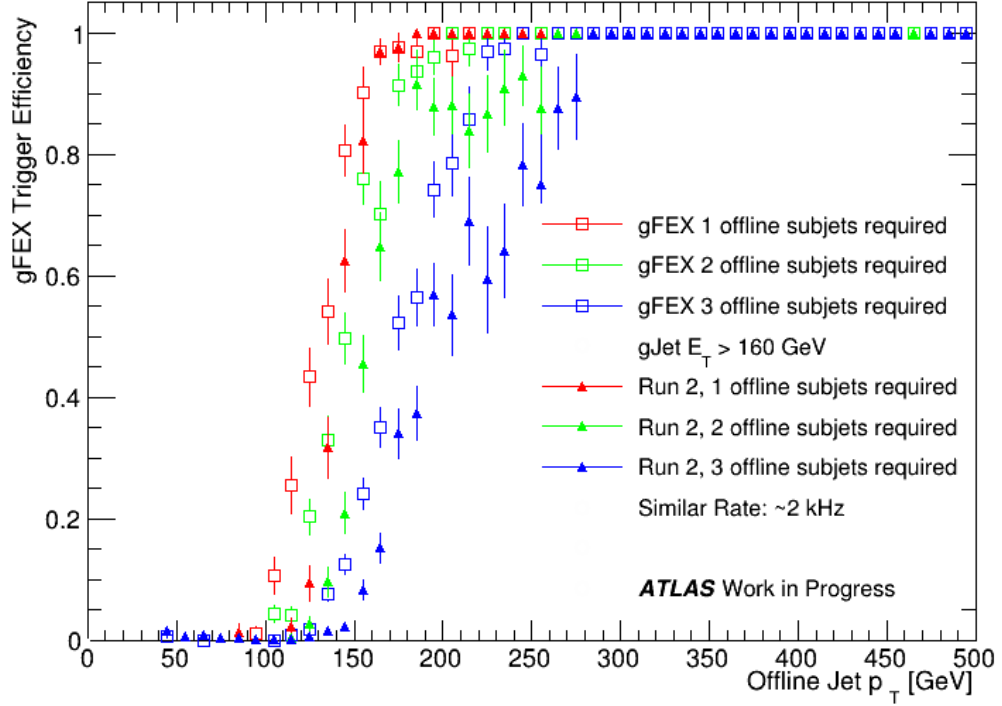


Figure 5.20: gFEX efficiency compared to Run 2 trigger efficiency with no pileup subtraction applied.

Additionally, the gFEX system has the ability to perform pileup subtraction on the inputs to the algorithm, the gTowers. This pileup subtraction is applied by calculating the pileup density ρ , and applying the subtraction to each gJet. This has a large effect on the rate as shown in Figure 5.21, and also has an effect on the efficiency as well, as shown in Figure 5.22. Ideally the pileup density ρ would be calculated by taking the median of all the gTowers, but the median is a very difficult calculation to make in the firmware. A running mean is much easier to calculate, and it is clear from Figure 5.21 that the mean taken from gTowers with an $E_T < 3$ GeV is very similar to the median. It is also clear that some sort of pileup

suppression is very necessary to reduce the rate down to a manageable level.

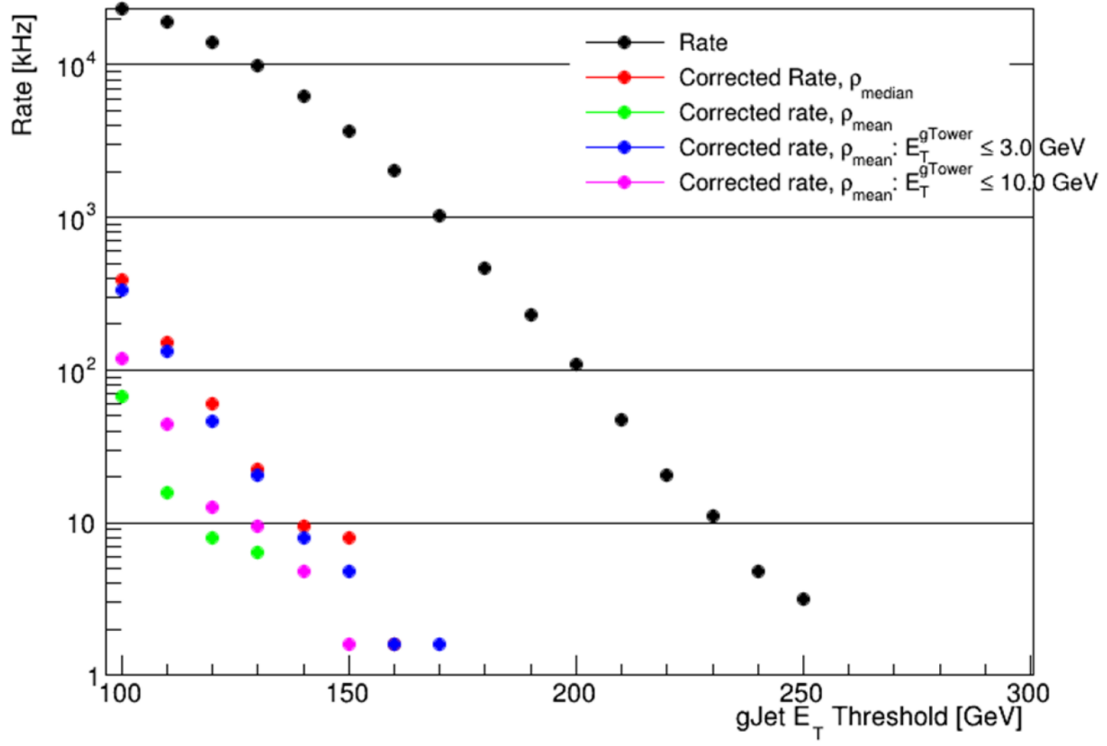
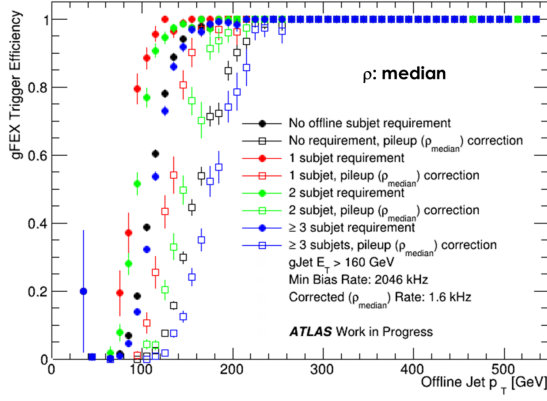
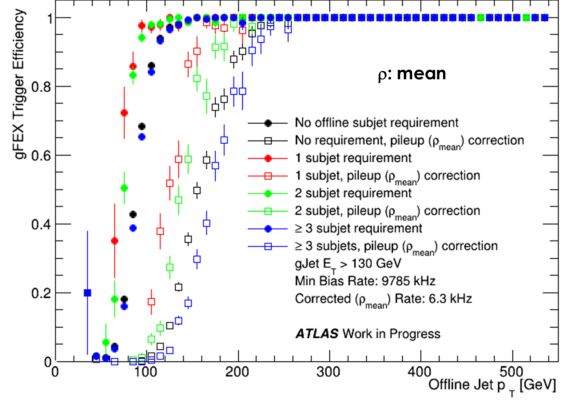


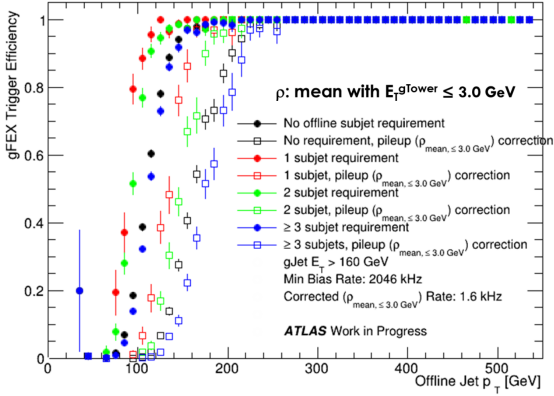
Figure 5.21: gFEX rates with pileup subtraction applied, using different ρ calculation methods.



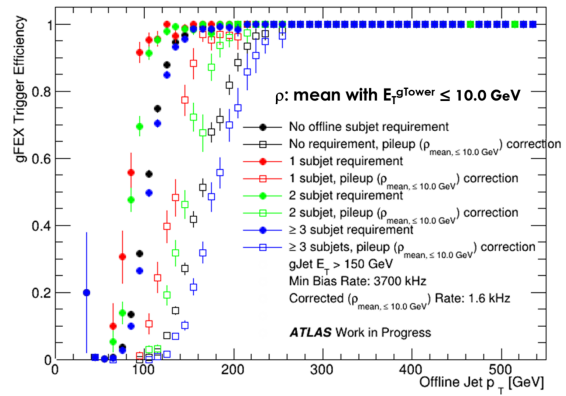
(a) gFEX efficiency with pileup subtraction using median pileup density



(b) gFEX efficiency with pileup subtraction using mean pileup density



(c) gFEX efficiency with pileup subtraction using $\rho, E_T^{gTower} < 3 \text{ GeV}$



(d) gFEX efficiency with pileup subtraction using $\rho, E_T^{gTower} < 10 \text{ GeV}$

Figure 5.22: gFEX efficiencies with different methods of calculating the ρ used for pileup subtraction, including median, mean and the mean under a certain threshold.

5.6 Custom Operating System

Part of what allow the gFEX board to function and to provide increased efficiency for global objects is the custom hardware designed and implemented for this purpose. The presence of a processor on the gFEX board, specifically on the Zynq+ SoC allows for the implementation of a Linux Operating System (OS), albeit a fairly stripped down version for the limited resources and embedded environment that the SoC provides. There are multiple ways to build an embedded OS but one of the most comprehensive, most configurable, and most flexible options is the combined application of the OpenEmbedded framework [118] and the Yocto Project [171]. The entire idea of these combined frameworks, as visualized in Figure 5.23, is that the same packages can be build for different supported hardware and distributions without having to dramatically change the implementation of the OS build. Additionally, this modular design allows for the building of embedded Operating Systems for custom hardware at an extremely high level, without requiring assembly or hardware languages and without extremely detailed knowledge of the hardware that is most likely only available to the manufacturer.

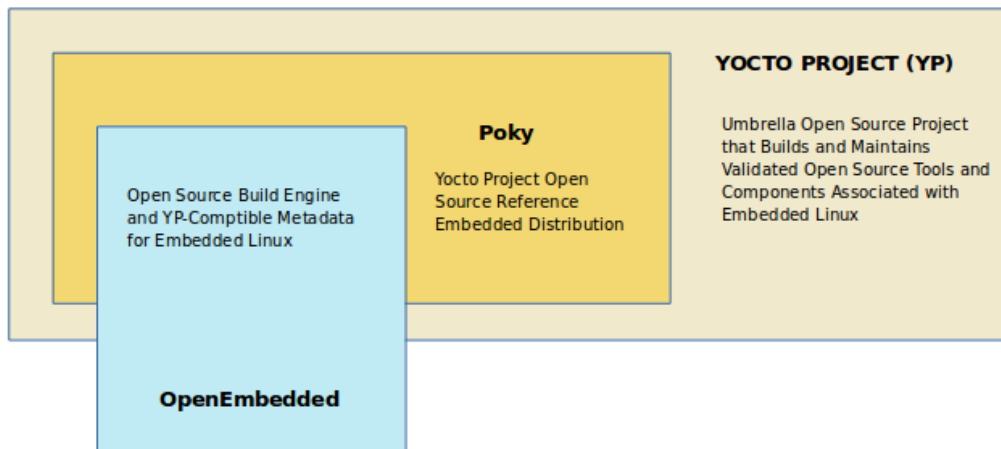


Figure 5.23: The Yocto Project and OpenEmbedded relationship [171].

5.6.1 *OpenEmbedded*

The OpenEmbedded framework is a build automation framework and cross compilation environment for embedded Linux distributions. Cross compilation refers to the act of compiling software on a host machine for use on a target machine with a different architecture than the host machine. OpenEmbedded helps to facilitate this process by factorizing hardware dependence from the implementation of specific packages. The framework does this through the use of "recipes". An OpenEmbedded recipe specifies how to build a given package, or packages, as well as source code locations, dependencies, and instructions for installation and removal of said package. These recipes are grouped into "layers" which help to organize recipes by functionality and level of specificity. Architecture-specific, application-specific and distribution-dependent recipes are grouped into different layers that are supported by the appropriate groups. The layers that contain the basics required to run an embedded OS are meta-openembedded and openembedded-core, which are implemented primarily as platform-independent and distribution-independent recipes.

5.6.2 *The Yocto Project*

The Yocto Project is an open source collaboration that helps users create custom embedded linux systems for any hardware. It provides flexible tools to make this possible, and a space for developers to interact and share software, configurations, and best practices. The Yocto Project maintains and validates three key component for embedded linux development: a set of integrated tools, a reference embedded linux distribution called Poky, and the OpenEmbedded build system, which is co-maintained with OpenEmbedded. Yocto uses the same Layer model as OpenEmbedded, with different layers containing different types of recipes with instructions on building and installing specific packages. The Yocto workflow is visualized in Figure 5.24. To build an embedded linux system, the developer starts by specifying

the architecture, software, patches, and configuration. The build system fetches and downloads the necessary source code, extracts them, applies any requested patches and configures and compiles the software. The binary package format is then applied, and package feed created which is used to make the final root file system. Then, the file system image is created. Throughout this process Quality Assurance tests and sanity checks are performed. Most developers use a Linux host machine but there are other options including Docker, and a web interface called Toaster.

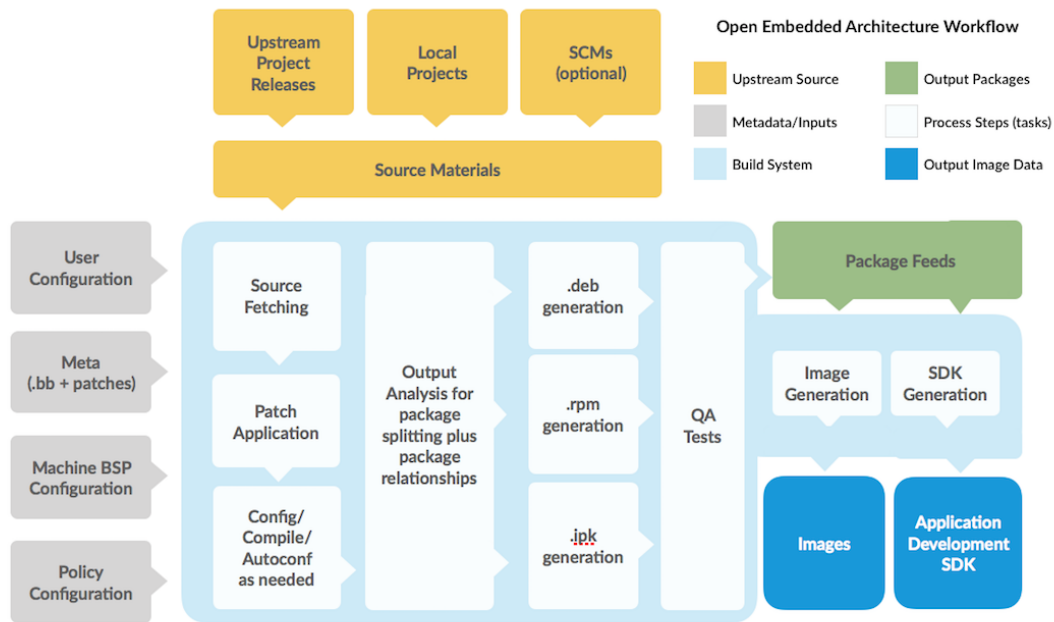


Figure 5.24: The Yocto Project workflow [171]

The Yocto Project has major releases twice a year in roughly April and October. Functionally, each layer is its own git repository and these releases are implemented as different branches in these repositories. This allows for a level of confidence that different layers with the same branch name (and thus release) are likely to work well together. Table 5.5 lists the most recent and most relevant Yocto versions, as well as some additional information which will be explained in the next section.

5.6.3 *Bitbake*

Bitbake is a make-like tool that builds or "bakes" a Yocto / OpenEmbedded based system (though it can be used for more general purposes as well). This tool is co-maintained by the Yocto Project and OpenEmbedded [117]. Bitbake is fundamentally a task execution engine that allows shell and python tasks to be run efficiently while working within an environment of complex inter-task dependencies. Bitbake executes tasks according to specific metadata provided in the form of recipe files (.bb), recipe append files (.bbappend), configuration files (.conf), include files (.incl), and class files (.bbclass). A fetcher library is included in Bitbake to obtain source code from a variety of source types, including git repos, tarballs, local files, or other websites. These recipes are the same recipes that are found in Yocto and OpenEmbedded layers, and Bitbake is the main build tool for building embedded Linux systems in this context.

5.6.4 *Xilinx Yocto-Manifests*

Xilinx is the manufacture of the Zynq SoC on the gFEX board as well as the other FPGAs. This is particularly relevant from the OS perspective because Xilinx as a manufacturer has a number of Yocto / OpenEmbedded layers that they maintain [166]. These layers are meta-xilinx, meta-xilinx-tools, and meta-petalinux [166]. The majority of the board support packages (BSP) that help to define the hardware are in meta-xilinx/meta-xilinx-bsp, including the definitions of the Zynq that are used to define the gFEX hardware. These layers can be integrated with a regular Yocto/OE build, but in addition to this, Xilinx offers a more integrated build system that automates the hardware descriptions used by the OS build procedure. To do this, Xilinx has made a fork of many relevant layers that are made to be compatible with the meta-xilinx-tools layer, which integrates the proprietary Xilinx software Vivado. Vivado is used to build firmware as well as other hardware specific files

for their FPGAs, SoCs, and other modules. These layers have their own versioning system which is tied to the Vivado version, and also linked to the Yocto release version as shown in Table 5.5. The biggest benefit to this setup is that the XSA file (previously called HDF for hardware description file) from the gFEX firmware build can be input into the OS build and the device tree (at least as it pertains to the firmware) can be automatically created and added into the OS build. This is hugely simplified from the manual custom device tree files as previously used in the gFEX system and makes adapting to firmware changes exceedingly simpler than before, which is incredibly important for the system as a whole. Xilinx has extensive documentation that describes how this system can be used to build embedded Linux Operating Systems [145].

Xilinx Branch	Yocto Code Name	Yocto Version	Release Date	Support Level
N/A	Scarthgap	4.4	April 2024	LTS (April 2028)
N/A	Nanbield	4.3	October 2023	7 months (April 2024)
N/A	Mickledore	4.2	May 2023	7 months (November 2023)
rel-v2023.1	Langdale	4.1	October 2022	EOL
N/A	Kirkstone	4.0	May 2022	LTS (April 2026)
rel-v2021.1 rel-v2021.2	Gatesgarth	3.2	October 2020	EOL
N/A	Dunfell	3.1	April 2020	LTS (April 2024)
rel-v2020.1 rel-v2020.2	Zeus	3.0	October 2019	EOL
rel-v2018.3 rel-v2018.2 rel-v2018.1	Rocko	2.4	October 2017	EOL

Table 5.5: Relevant Yocto versions compared to Xilinx Vivado versions.

5.6.5 *gFEX Custom OS Implementation*

The gFEX custom OS is built using the Xilinx versions of Yocto/OpenEmbedded. This provides the benefit of the open source Yocto/OE framework as well as a more integrated

OS build that produces all the files necessary, even hardware specific files thanks to the Xilinx integration.

For use with the gFEX board, a new layer called meta-l1calo was created by Giordon Stark before I joined the project. This layer holds the custom description of the gFEX hardware and device tree, as well as custom software needed to integrate and run the board with the rest of the ATLAS trigger system. Right now this layer is compatible with the Yocto release Zeus and the Xilinx release rel-v2020.2, though this will probably be updated in the future, and should be clear in the meta-l1calo/conf/layer.conf file, as well as the meta-l1calo/scripts/setup.sh file.

This layer has a number of components that can be described as follows:

- **meta-l1calo/conf/machines:** This is where the custom gFEX "machines" are defined. From a Yocto perspective, an OS image can be build for any machine that is defined. These .conf files are based on Xilinx's configuration files for the Zynq SoC, since this is the chip on the gFEX board with the processor that runs the OS. This also allows for certain versions of recipe configuration or include files to be used for certain machines. At this point, it is primarily used to distinguish between the gFEX at P1 (gfex-production-p1) and the gFEX at the Surface Test Facility (gfex-production-stf). In the past it was also used to distinguish between different versions of gFEX prototypes.
- **meta-l1calo/docs:** This is where the detailed documentation lives! Hopefully it's still up to date.
- **meta-l1calo/recipes-bsp:** All of the hardware related so-called "board support packages" live here. This includes things like the custom gFEX device tree and the meta-l1calo appended version of the Xilinx external-hdf recipe. The custom gFEX device

tree is limited to hardware components on the board that aren't related to the firmware on the Zynq FPGA. At the moment this means the two ethernet devices on the board, and the SD card slot among others. The external-hdf .bbappend file only points to the correct XSA file. At the moment this is a "golden image" which contains all the infrastructure needed for monitoring, but none of the production firmware that would normally live on the Zynq SoC.

- **meta-l1calo/recipes-core:** Here lives the gFEX image definition (which decides what is included in the OS build), as well as all of the infrastructure to run scripts and programs on initialization of the OS. There are many things that need to be run and they are arranged at the moment to all run at different times, in a specified order to prevent any conflicts between these processes.
- **meta-l1calo/recipes-gfex:** This is a folder for specifically gFEX software. Mostly custom software loaded from git repositories that was written specifically for gFEX functionality.
- **meta-l1calo/recipes-kernel:** This implements slight modifications to the Linux kernel configuration. Mostly modifications of various variables in the kernel build that are necessary for different applications or functionalities.
- **meta-l1calo/recipes-python:** It will come as no surprise that this folder holds all the python modules that are required for gFEX. The basics are installed using the regular python3 recipe from OpenEmbedded that is included in the image definition file, but the recipes in this directory are additional modules that didn't already have a recipe in another one of the dependent layers.
- **meta-l1calo/recipes-support:** The only recipe in this folder is Chrony, the NTP manager to keep the time on the board correct. It accesses the NTP server available

at P1 and the Surface Test Facility to make sure the board time is always correctly set³.

- **meta-l1calo/scripts**: This has only one file, the setup.sh script which pulls all the necessary files, downloads repo, checks out the correct branches of all the Xilinx Yocto repositories, and prepares to build the OS.
- **meta-l1calo/wic**: The file here defines the gFEX wic image which can be used for flashing the OS to an SD card. This includes the Image and rootfs, so can be quite large if the rootfs is large.

I will not go in to technical detail on how to build the custom gFEX Operating System as this information should be fully covered in the meta-l1calo documentation, but in general it is the same as the Yocto workflow described above, and shown in Figure 5.25. On a host system, the existing layers that are needed are cloned. These can be from Yocto, OpenEmbedded, or Xilinx. A new layer is created, or expanded with custom recipes or modifications to existing recipes. Once all the configurations are done, the build is "baked" with bitbake, and then installed onto the target system.

The custom operating system on gFEX allows for significant functionality to be implemented in the OS rather than in firmware which reduces firmware complexity. The majority of the offloaded functionality has to do with control and monitoring, which will be discussed in detail in the following section.

3. The time to which the OS is set is controlled by the relevant NTP server in either P1 and the STF, and currently sets the board to UTC time.

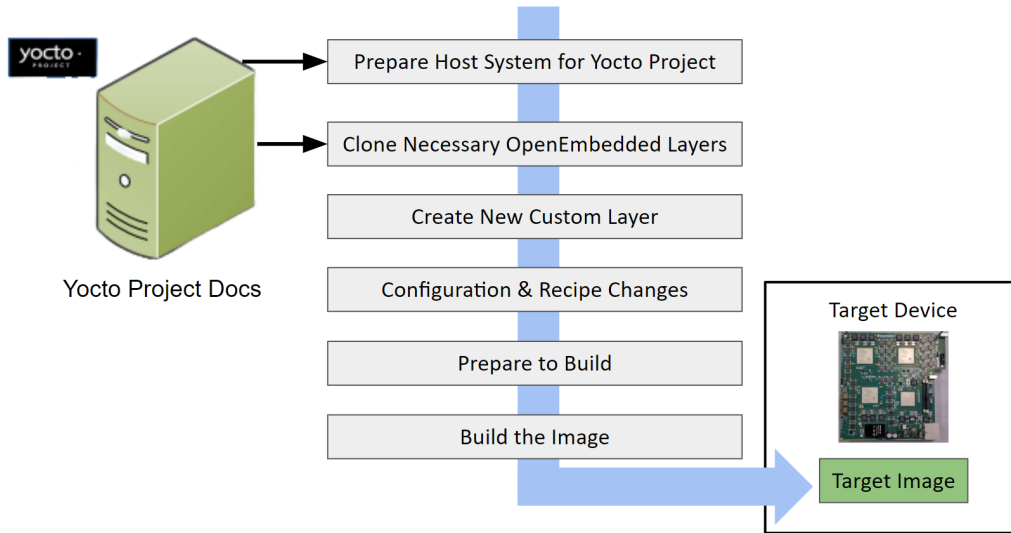


Figure 5.25: Workflow for setting up and building the gFEX OS

5.7 Control and Monitoring

The control and monitoring of the full gFEX board is a large and somewhat complex undertaking. There are many different kinds of monitoring as shown in Figure 5.26, but this section will focus on monitoring low-level hardware sensors, as well as monitoring of register values in firmware, and how to configure these values. This section will also discuss the IPMC and shelf monitoring setup, as well as the IPBus based system that allows the board itself to be reachable from the ATLAS Control Room.

5.7.1 I²C Monitoring

The I²C monitoring refers to monitoring by the OS of a number of low level hardware sensors that are connected to the Zynq SoC processor over an I²C(I2C, inter-integrated circuit) bus. The I2C bus is a synchronous multi-target/multi-controller serial communications bus. In this case it is used to connect a number of low-level voltage, current, and temperature sensors

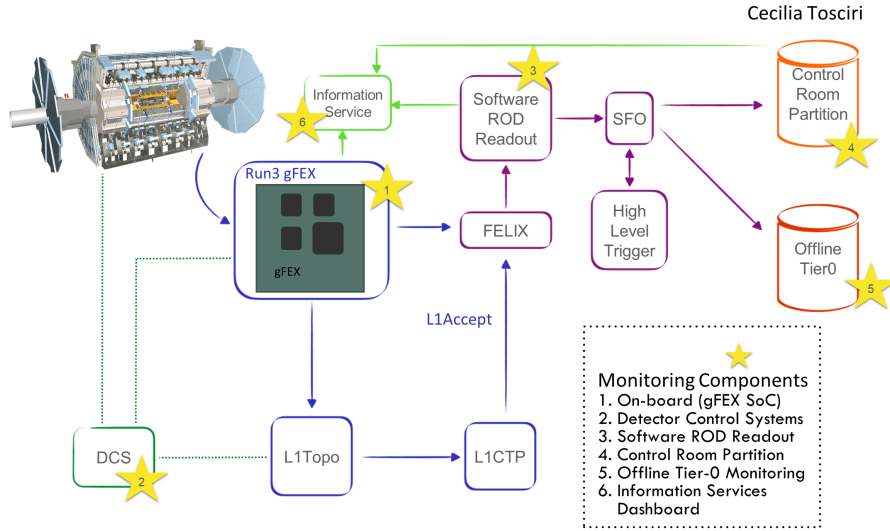


Figure 5.26: Monitoring overview, including several different graphical interfaces, and low level on-board hardware monitoring. Image Credit: Cecilia Tosciri.

to the processor. For a full list of the I2C monitored sensors, please see Appendix A. There is a second I2C bus that is connected to the SI5345 clock chip that is on the gFEX board which allows access to the registers on this clock chip for both monitoring purposes and configuration purposes. See Appendix B for more information on the SI5345 registers.

The I2C bus and associated sensors can be accessed using the Python module periphery and its associated I2C method which allows for an I2C bus object to be defined based on the /dev address of the I2C bus. The registers of the sensors on that I2C bus can be read and written using this same mechanism. This allows for the voltage, current, and temperature to be read from the power modules, temperatures sensors, and minipods. The Loss of Signal (LOS) can also be read from the minipods. Methods for reading and writing the various types of sensors were developed jointly between myself and a BNL engineer Shaochun Tang, and were integrated into the OS. They were also integrated into a monitoring script that is run on initialization of the board which constantly monitors these sensors and contains a backup fail-safe to disconnect the onboard DCDC converter from power if the temperature

risers above the configured threshold. This would effectively power down the board and thus reduce the temperature to protect the hardware, particularly the FPGAs.

5.7.2 IPMC and ATCA Shelf Monitoring

The IPMC is a small removable mezzanine which can be inserted into an existing slot on the gFEX board. The IPMC is what interacts with the Shelf Manager and allows for some level of monitoring via the shelf. Depending on the firmware of the IPMC, a number of sensors can be monitored and given thresholds: Non Critical High/Low, Critical High/Low, Non Recoverable High/Low. The gFEX IPMC monitors a sensor on the IPMC, as well as the maximum temperature of all minipods, power modules, and temperature sensors, as well as the Zynq temperature. These temperatures are gathered in the OS using the I2C monitoring scripts and the maximum is found, then these maximums are written to a general purpose GPIO in the Zynq FPGA. Then, the IPMC firmware gathers these values from this GPIO in the Zynq firmware. If the values have not been updated in a certain time frame, the Zynq firmware reverts to a default value of 0 degrees Celsius. This prevents a situation where the process on the Zynq freezes or dies and this failure does not get propagated to the IPMC values. A block diagram of this setup is shown in Figure 5.27.

The LAPP IPMC was originally used on the board, but this was switched to the CERN IPMC due to a known issues with the LAPP IPMC which can cause a board to be unresponsive to Shelf Manager commands in specific circumstances, and requires the board to be re-seated before control is regained. Of course, with a board like the gFEX that will be installed in the ATLAS electronics cavern, it is no small feat to travel underground to reseat the board. Thus it was felt that the switch to the CERN IPMC was worth the effort. The major pin mappings are compatible between the LAPP and CERN IPMCs. A small subset of minor pins on the CERN IPMC were modified for gFEX specific uses:

Sensor	Upper Non-Critical	Upper Critical	Upper Non-Recoverable
IPMC Internal	45°C	55°C	70°C
Zynq Temperature	55°C	70°C	85°C
Temperature Sensors	55°C	65°C	70°C
Power Module Sensor	90°C	95°C	100°C
MiniPOD Temperature	65°C	75°C	80°C

Table 5.6: gFEX IPMC temperatue sensor threshold values.

5.7.3 *FPGA-Based Monitoring and Control*

In addition to the low level monitoring of hardware sensors there is extensive monitoring of registers in the Firmware. These registers also allow for configuration and control of the algorithms. Full reading and writing capabilities are needed for these firmware registers to have full control over the gFEX firmware algorithms and the setup procedure. This section will describe the custom protocol used on the gFEX for several years until ongoing updates implemented in 2023 and 2024.

Registers on the Zynq can be read from the Zynq OS using `/dev/mem`. `/dev/mem` is a device file that holds an image of the main memory of the system. Physical memory addresses can be referenced, and non existent addresses return an error. Generally access is only allowed to memory mapped regions though this depends on the configuration of the kernel.

Registers on the pFPGAs can be read from the Zynq OS utilizing the custom FIFO implemented in the Zynq and pFPGA firmware. The memory mapped interface on the Zynq ARM processor is connected to the FIFO, with specific registers on the Zynq that control the RX (Zynq \rightarrow pFPGA) and TX (pFPGA \rightarrow Zynq) transactions [138]. There are three 32-bit registers for both RX and TX: DATA, CTRL, and STAT. These registers hold the FIFO data, FIFO control bits, and FIFO status bits respectively. The control and status words have specific bits representing the ready signal (RDY), start of frame (SOF), end of frame (EOF), and read enable (RDEN) as listed in Table 5.7.

FIFO Register	Relevant Bits
IPB_TX_CTRL	[0] RDY, [1] EOF, [2] SOF
IPB_TX_STAT	[0] RDY
IPB_TX_DATA	[32:0] Data
IPB_RX_CTRL	[0] RDEN
IPB_RX_STAT	[0] RDY, [1] EOF, [2] SOF
IPB_RX_DATA	[32:0] Data

Table 5.7: gFEX FIFO registers and relevant bits.

The FIFO protocol requires that specific control bits be toggled in a particular order to send the contents of the DATA register to the pFPGA, or to read the result of a transaction from the pFPGA from the DATA register. I wrote and maintained a driver to perform these operations, along with Greg Myers, which I integrated into the Operating System and which has been used for several years, including while the LHC was running and full Physics data was taken from 2023-2024 [183].

5.7.4 *Ironman IPBus Communication*

The gFEX board along with the rest of the L1Calo trigger system is installed in the ATLAS electronics cavern which is 100 m underground. The system must be configured from the ATLAS Control Room on the surface, so this control and configuration of the system must take place over an internet based connection of some kind, usually as an Ethernet connection. Ethernet uses the Internet Protocol (IP) which is a set of rules that determines how packets can be routed and addressed to the correct destination. The IPBus protocol is a simple IP-based protocol used to communicate with hardware devices. Most often the target system implements firmware to receive IPBus packets, though for gFEX this is implemented on the software side in the Operating System. Due to the most common implementation being in firmware, most often User Datagram Protocol (UDP) is used, since it is more easily implemented in firmware. This requires the use of a control hub module to intermediate

IPBus packets, as shown in Figure 5.28. gFEX is unique in its implementation of the IPBus protocol in software. This allows the firmware to be simplified and adds additional flexibility in the implementation and debugging of the setup. Additionally the more reliable Transmission Control Protocol (TCP) can be used easily, while also supporting UDP.

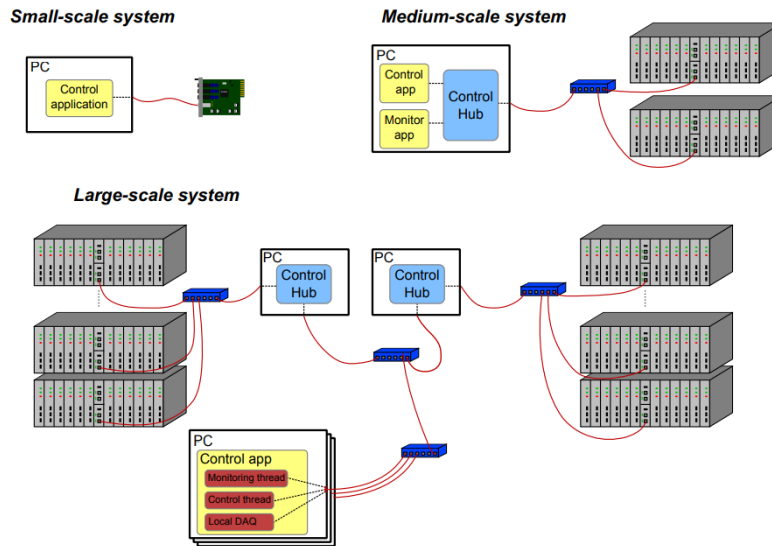


Figure 5.28: Diagram of the IPBus protocol implemented in different sized systems [143].

IPBus packet handling was implemented in Python by a previous UChicago student working on gFEX, Giordon Stark, in the ironman Python module [185]. This software manages the logic of connecting incoming IPBus packets with outgoing IPBus responses. Ironman provides several standard protocols for read and write packets, as well as a hardware manager that allows for the definition of custom hardware maps specifying the layout of the hardware and custom communication protocols to interface with this hardware [186]. The Ironman software also implements a single-threaded reactor model which listens for incoming packets and responds appropriately. An overview of the Ironman functionality is shown in Figure 5.29 [186].

The Ironman server listens for incoming packets and upon receipt of a packet initiates a chain

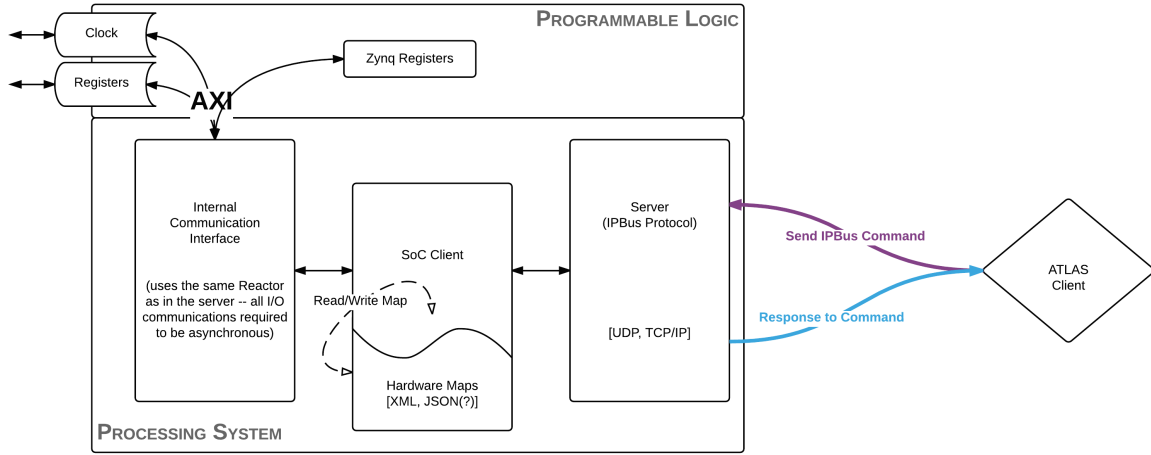


Figure 5.29: Overview of ironman functionality [186].

of deferred callbacks and proceeds to unpack the data and check the validity of the data and packet headers. If the packet is a request for hardware information the packet is passed to the custom hardware communication protocols which are defined by the implementation on the target system as seen in Figure 5.30. The server additionally maintains a log of incoming and outgoing packets and can produce this information when prompted [186].

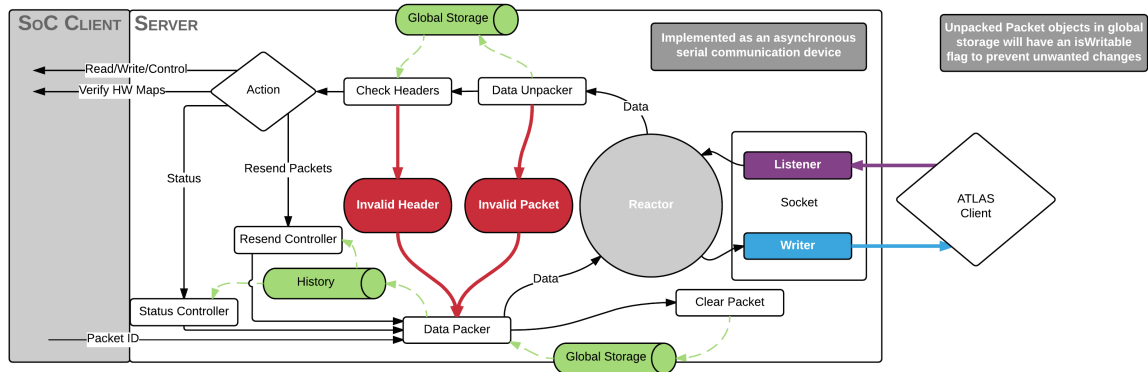


Figure 5.30: Overview of ironman server [186].

The Ironman Hardware Manager offers a way to customize Ironman implementations to any type of custom hardware. The Hardware Manager allows for custom hardware maps to be added to the Manager, several different custom maps can be used if needed (such as

Register Maps). It checks to ensure that the maps are compatible, with no conflicts, as well as being parsable and valid [186]. An example of how the Hardware Manager works is shown in Figure 5.31. The Ironman client performs most of the work of routing packets where they need to go internally. The client utilizes the Hardware Interface to validate the packet request and routes it to the correct custom software implementation used to access the hardware.

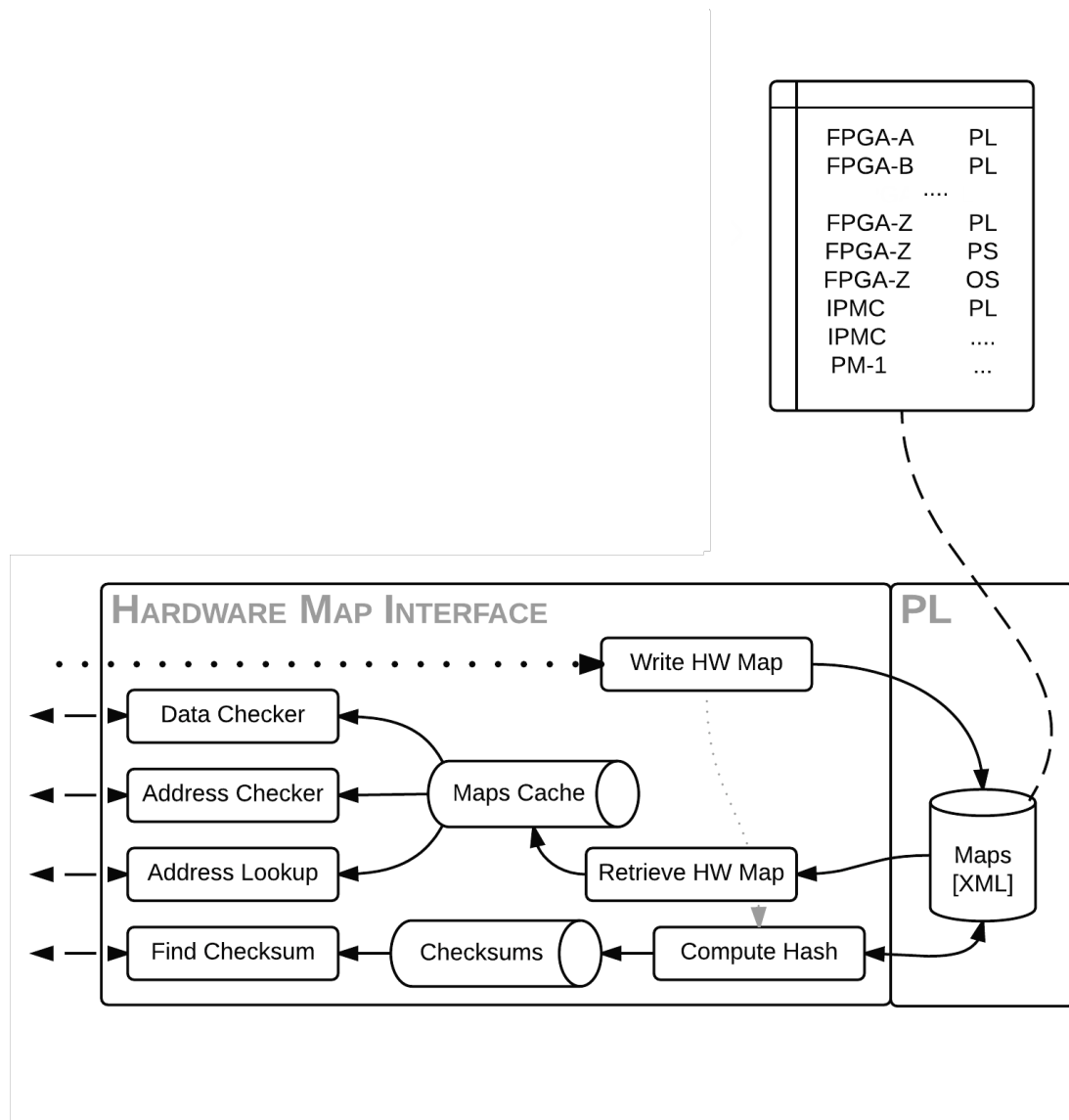


Figure 5.31: Overview of ironman Hardware Interface [186].

Ironman as a whole is a very modular and nonspecific piece of software which implements a full working system to receive and respond to IPBus packets. It is composed of a server, client, hardware interface, and custom software to access the hardware. This general implementation allows for a high level of flexibility, but does make a specific implementation a bit more confusing. A full working example can be found in the gFEX software [183].

The implementation of ironman on gFEX utilizes the Hardware Manager to add a separate register map for each FPGA (Zynq, A, B, C). Each of these then have their own class that specifies how reading and writing is implemented for each piece of hardware, which uses the already existing drivers written for gFEX which are used to access FPGA registers. The gFEX implementation of Ironman also includes methods for building response packets, printing packets for debugging, and management of the deferred chain of callbacks. It also creates a server that listens to several different possible protocols at different ports, including UDP, TCP, and an L1Calo specific version of TCP. Ideally there would be additional error handling added to the gFEX implementation to return packets with a proper error code to the external client. At the moment, an error is printed to a log file, and the client receives a timeout error. This implementation on gFEX has been used for several years, including with the ATLAS control room during official physics data runs.

5.8 Detector Control Systems

The Detector Control Systems (DCS) is used in the ATLAS Control Room and remotely to monitor the status of all hardware in the ATLAS detector. DCS includes a graphical interface and a client and server system that guarantees real-time updated monitoring to ensure safety of the hardware. It also includes an alarm system to notify users of ongoing alerts, and a fail-safe that can deactivate modules if necessary. The DCS graphical interface is often referred to as the "FSM" (which stands for Finite State Machine), and the different interfaces that

can be accessed graphically are referred to as panels. The underlying architecture is based on the SIMATIC WinCC Open Architecture software (WinCCOA) and on the Open Platform Communication Unified Architecture (OPCUA) client and server.

The OPCUA server obtains the low level hardware values from the sensors being monitored. Depending on the module, this server run on the module, or on a nearby PC and uses the IPBus protocol over Ethernet to gather sensor data. The gFEX module runs an OPCUA server on the SoC OS, and runs an additional OPCUA server on a nearby PC that gathers data from the Shelf Manager. The OPCUA client runs on a specific machine and interfaces with the servers over Ethernet to gather datapoint values. This client is connected to the WinCCOA project which is tied to the displayed graphical objects on the FSM panels. An overview of how this system works can be seen in Figure 5.32.

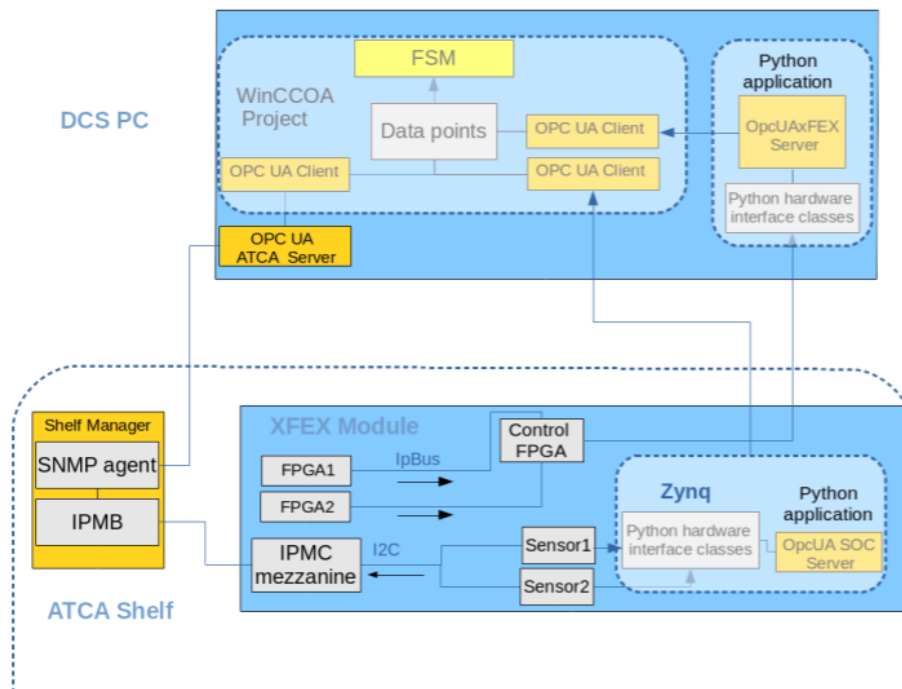


Figure 5.32: Overview of the DCS system for L1Calo. Image from Paul Thompson.

The gFEX DCS project is based on existing libraries created by the ATLAS central DCS group, such as fwATCA. These centralized libraries make it possible to standardize ATCA-

based OPCUA servers that utilize IPMC data points delivered by the shelf manager. These are implemented for the entirety of the L1Calo system, including gFEX. gFEX also utilizes a standalone OPCUA server that runs on the OS of the SoC.

The standalone gFEX OPCUA server was originally implemented using the CERN Quasar framework, which is a quick OPCUA generation framework that is C++ based [160]. The Quasar setup allows the user to define custom classes in the Design.xml file and create specific data points belonging to these classes in the config.xml file. The gFEX class structure, Figure 5.33, includes classes for the different types of physical sensors (FPGA, MiniPODs, Power Modules, Temperature Sensors) and for the different types of data being monitored (temperature, voltage, current, loss of signal, and health flags).

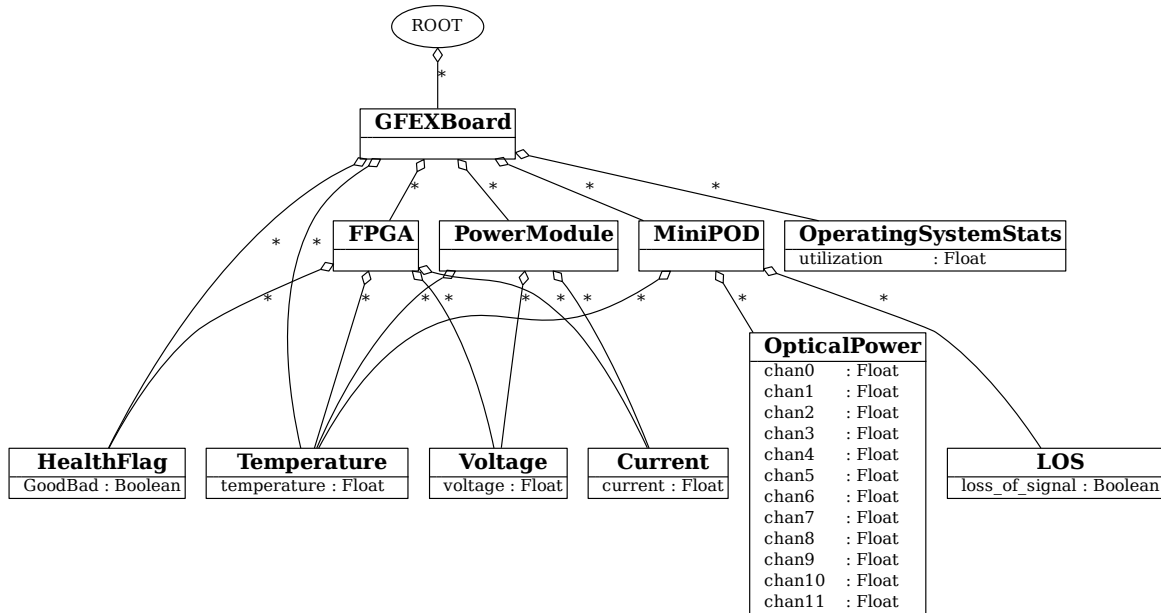


Figure 5.33: gFEX OPC-UA server class diagram.

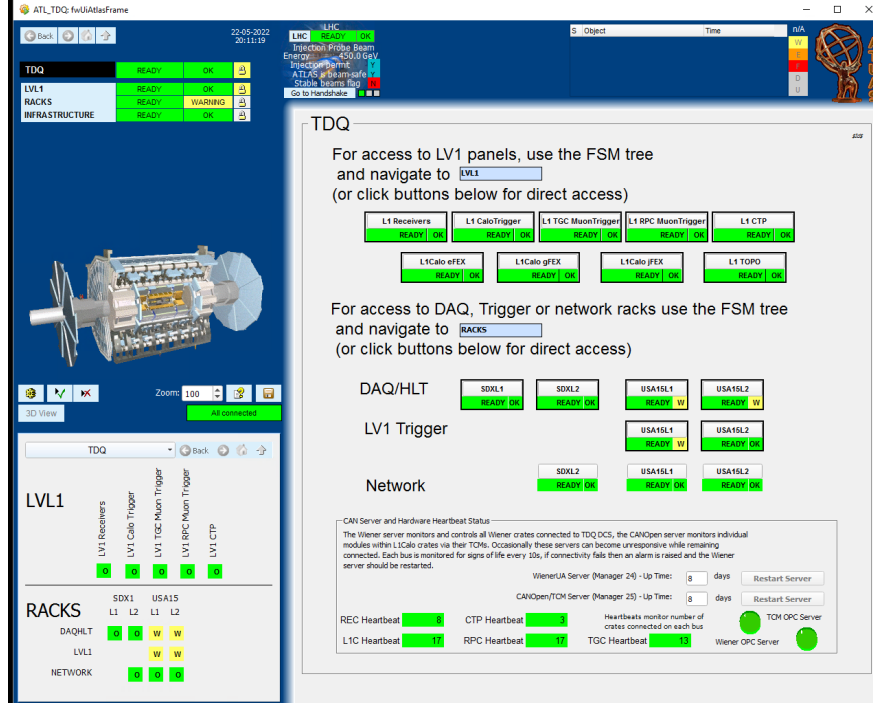
The Quasar group at CERN developed a python-based OPCUA server with the MilkyWay project [125]. A python based gFEX server was created based on MilkyWay that is analogous to the C++ based server. The python server is much easier to integrate into the OS as it

doesn't need to be compiled. This python server is deployed on the Zynq SoC, and uses the same I2C monitoring code that the IPMC and on-board monitoring to supply low-level sensor data to the OPCUA server datapoints. The ATCA based OPCUA server is built using central DCS libraries and is based on the sensors available in a specified IPMC and slot number for a specific board.

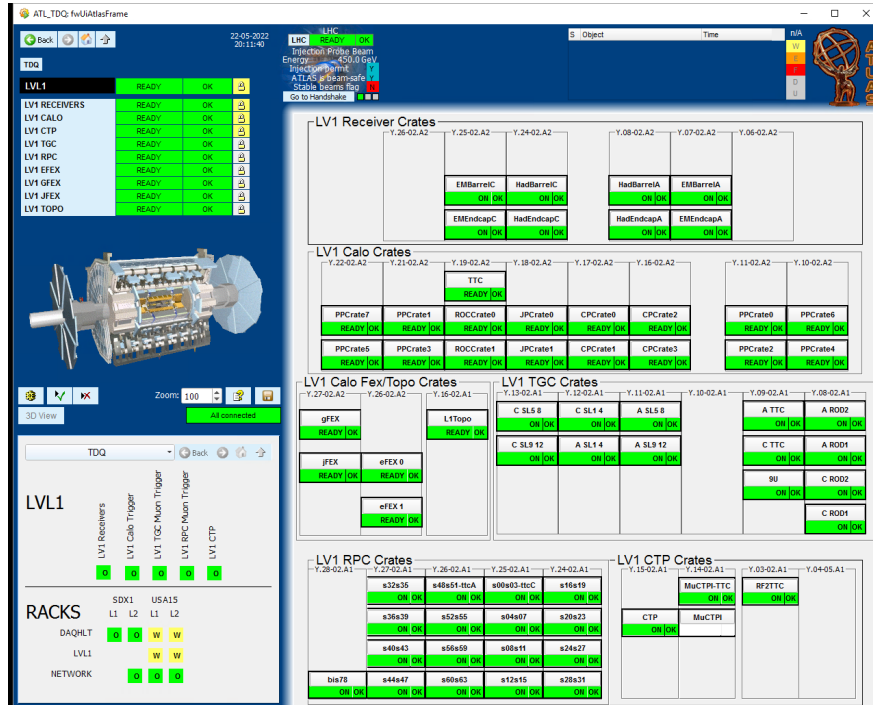
The gFEX build script builds the ATCA core server, the fwATCA client, the SoC server based client, and the FSM. There are several custom WinCCOA scripts that I modified and wrote to facilitate the build of the gFEX clients and FSMs. There was also extensive panel design that went into the gFEX DCS system. The original gFEX panel design was very graphical in nature but did not take advantage of the central DCS developed tools. The currently implemented panels take full advantage of central DCS tools and display most data in a table format. The high level DCS panels for TDAQ and L1Calo are shown in Figure 5.34. Figure 5.35 shows the gFEX shelf panels which monitor additional datapoints like the Shelf Manager and Fan Tray temperatures and voltages.

The main gFEX panel, Figure 5.36a, shows the values from the ATCA server which are supplied by the gFEX IPMC, as well as a schematic of the gFEX board. This panel also has buttons that lead to the secondary gFEX panels displaying different types of data from the SoC server. Figure 5.36b and Figure 5.37 show temperature, power module, and OS related datapoints. The miniPOD panel, Figure 5.38a, displays graphical LOS objects in the upper right corners, and has a table containing temperatures, voltages and optical power. The alerts for optical power have been disabled, so these datapoints are gray in the table, but the values are archived for long term data on the optical power of the miniPODs. Figure 5.38b shows the plot of archived data that is available for every datapoint in the panel, when the datapoint object is right clicked.

In addition to just graphical display panels, the gFEX DCS project also includes two expert

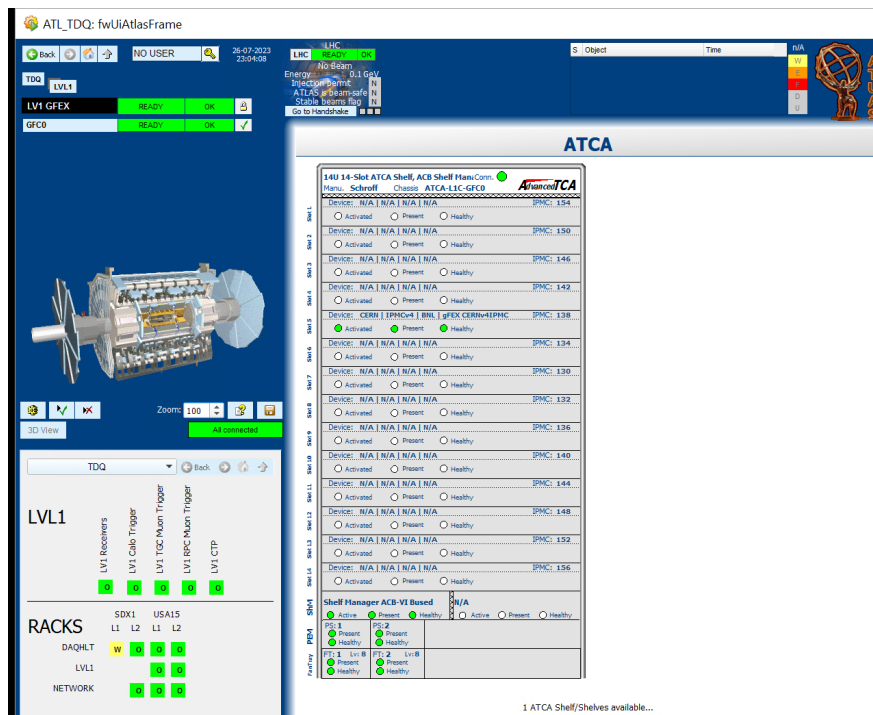


(a) TDAQ DCS panel

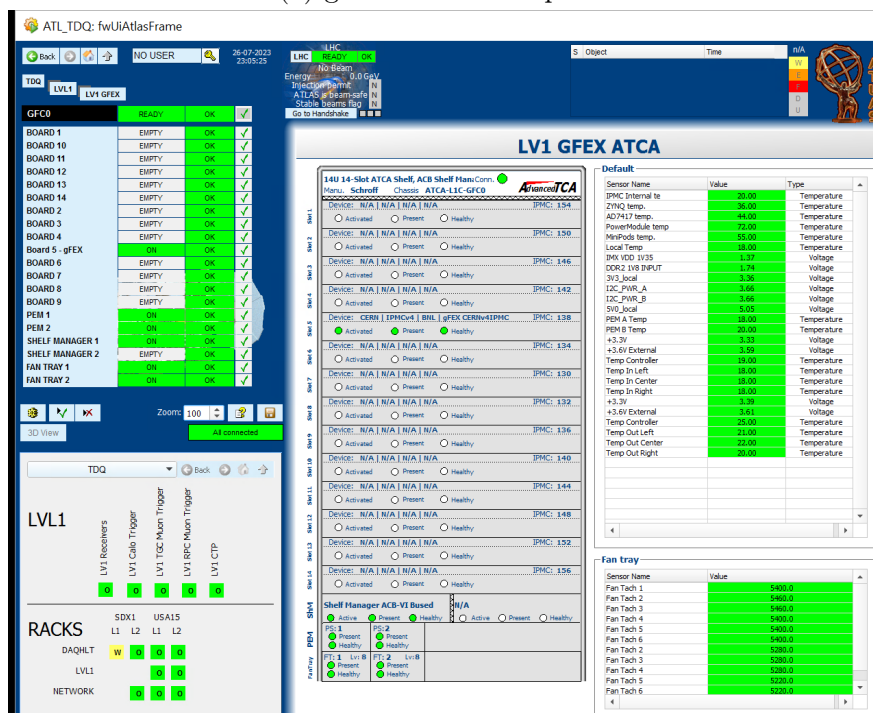


(b) L1Calo DCS panel

Figure 5.34: High level DCS panels for TDAQ and L1Calo that include links and graphics representing gFEX.



(a) gFEX shelf DCS panel

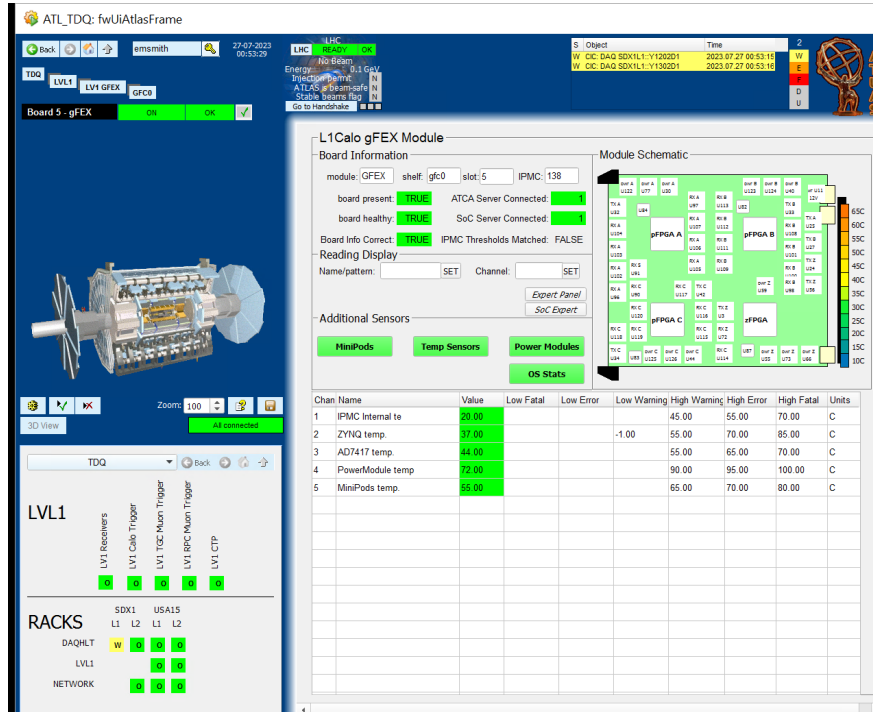


(b) gFEX shelf detailed DCS panel

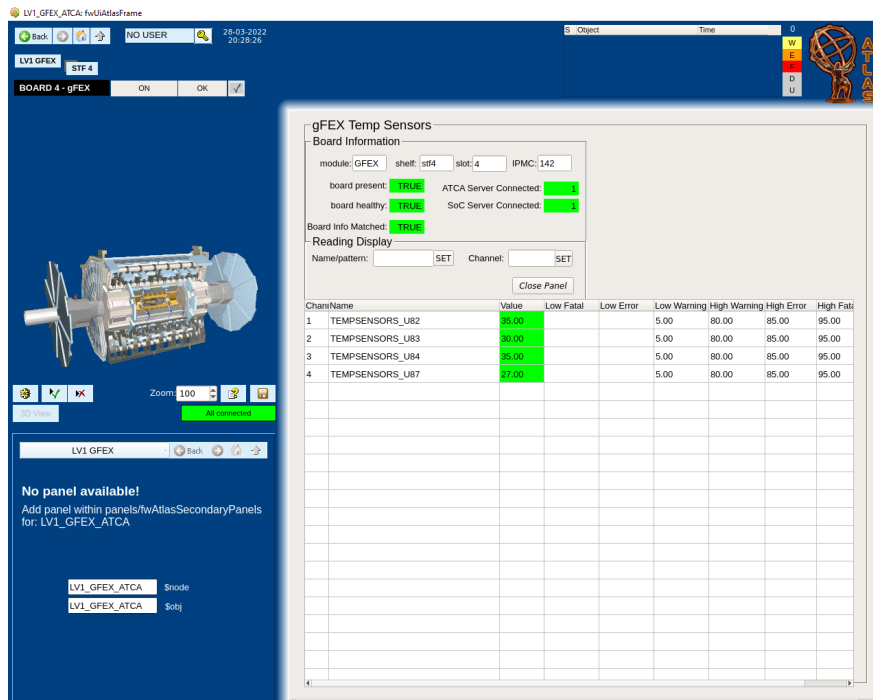
Figure 5.35: High level DCS panels for the gFEX shelf, gfc0.

panels that allow an expert user to modify the displayed data and alerts. When an expert is logged in using the key icon in the upper left corner of the FSM, the two expert panel buttons are available on the main gFEX panel as seen in Figure 5.36a. The ATCA expert panel for gFEX offers interactive buttons which check the shelf manager information and compare it to the ATCA settings in the DCS project. If they differ, there is a button to update the DCS information. For the gFEX project there is one anomaly because there are two inserted alerts around the value 0 which gives the alert "No Hardware Update" from the IPMC. This means that there are usually two more alerts than there are thresholds, but the mismatch is set to zero. This could be improved in the future to be less confusing. An example of this setup is shown in Figure 5.39a. The SoC expert panel provides an interactive list of the alerts that are enabled, and allows for certain groups of alerts to be disabled if preferred by the expert user, as shown in Figure 5.39b. The miniPOD optical power alerts are generally kept disabled, but this interface provides a way to enable them should this ever be wanted in the future.

All of the ATLAS TDAQ DCS implementations are kept in the same repository [187]. Each subsystem has an individual folder to hold their specific implementation of the WinCCOA client and FSM. There are individual folders holding WinCCOA scripts as well as regular bash scripts to facilitate building the various components of the DCS project.

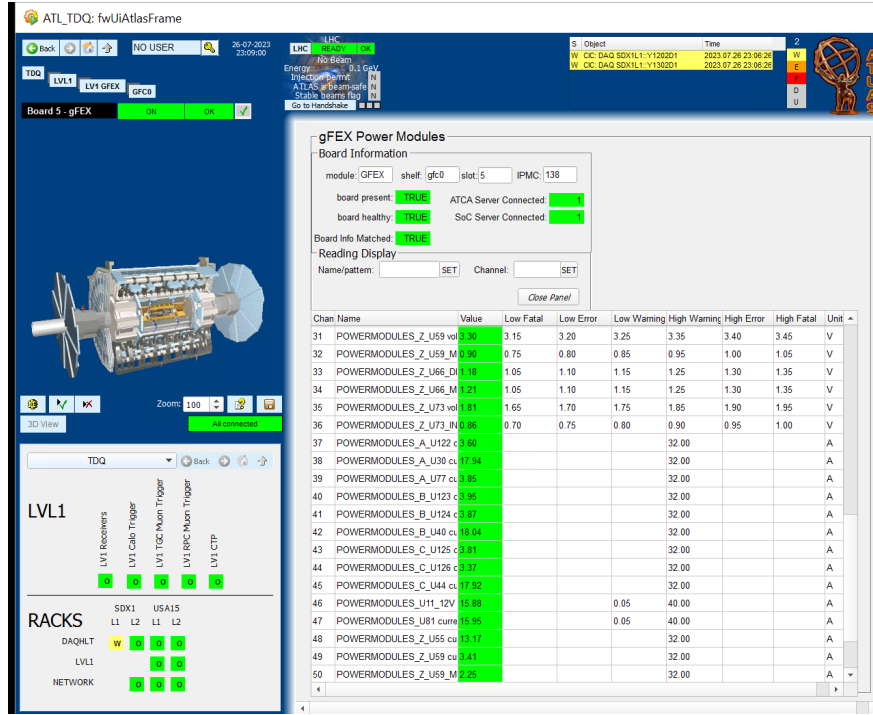


(a) gFEX main panel, with expert buttons.

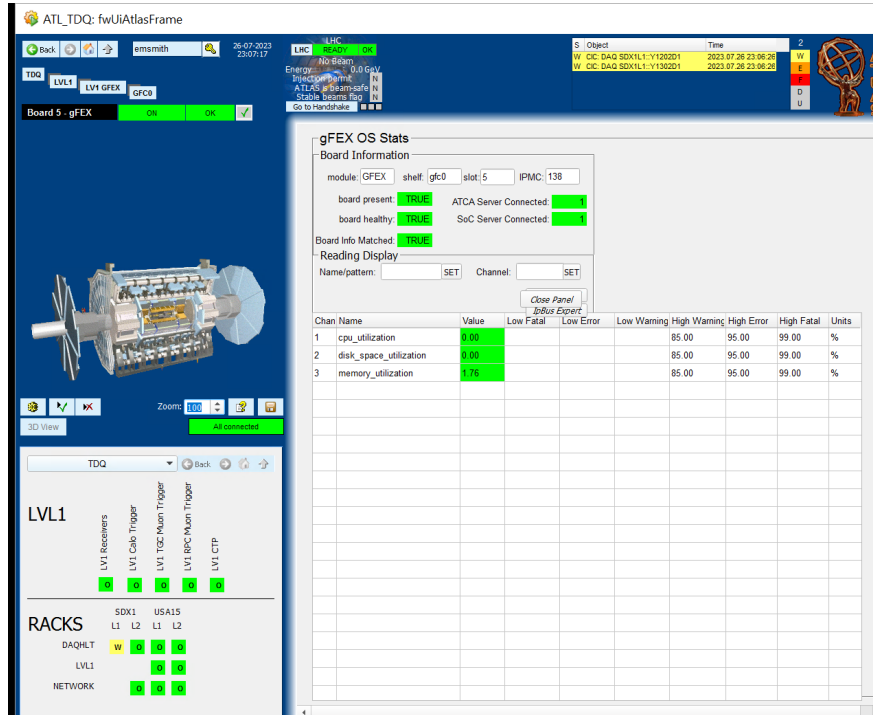


(b) gFEX temperature sensor panel.

Figure 5.36: gFEX main DCS panel with ATCA datapoints, and gFEX temperature sensor panel.

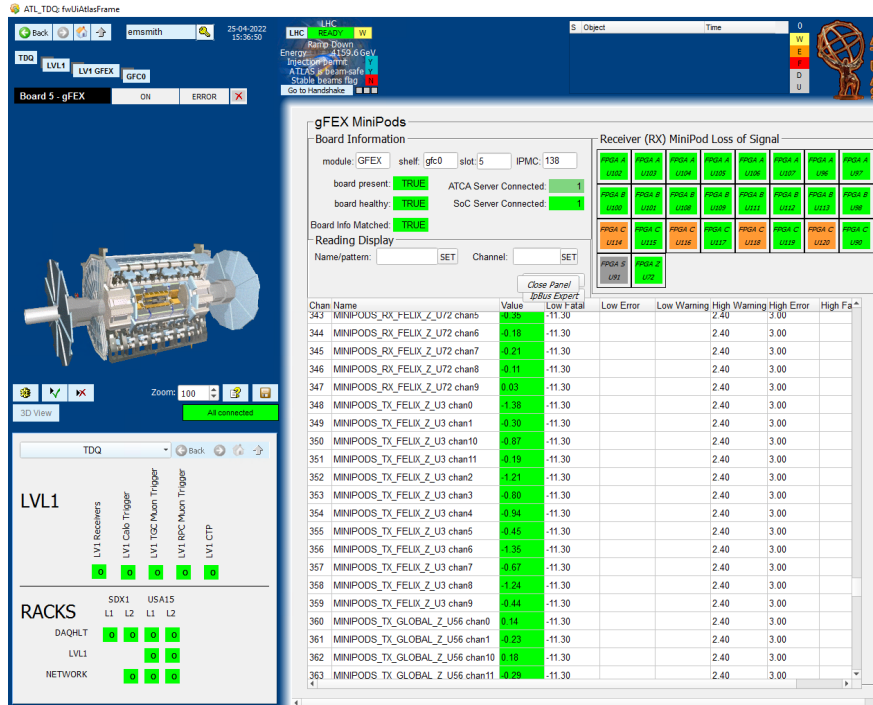


(a) gFEX power modules panel

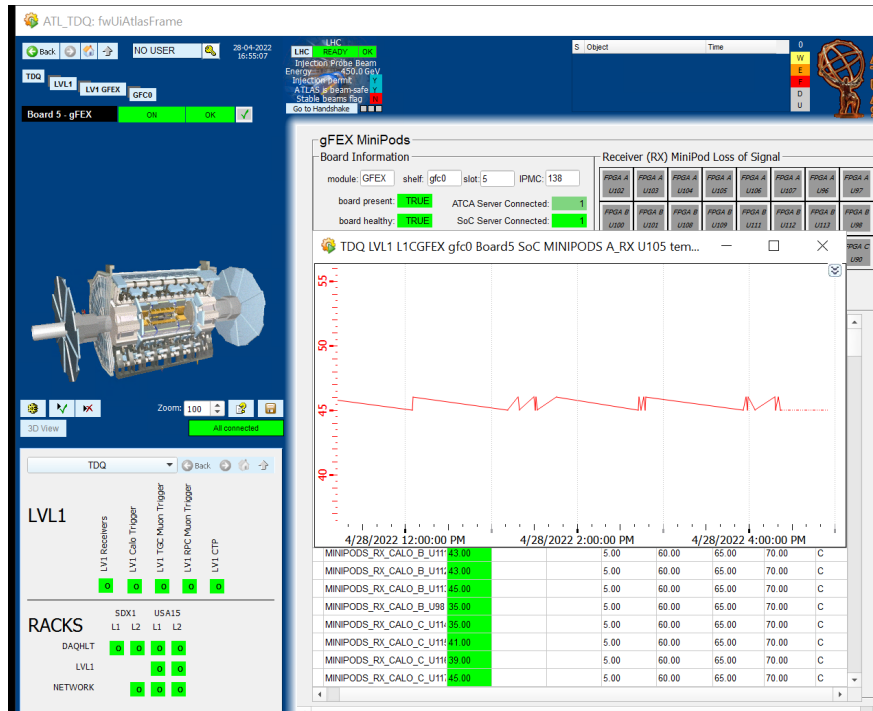


(b) gFEX OS utilization panel

Figure 5.37: gFEX power modules DCS panel, and OS utilization panel.

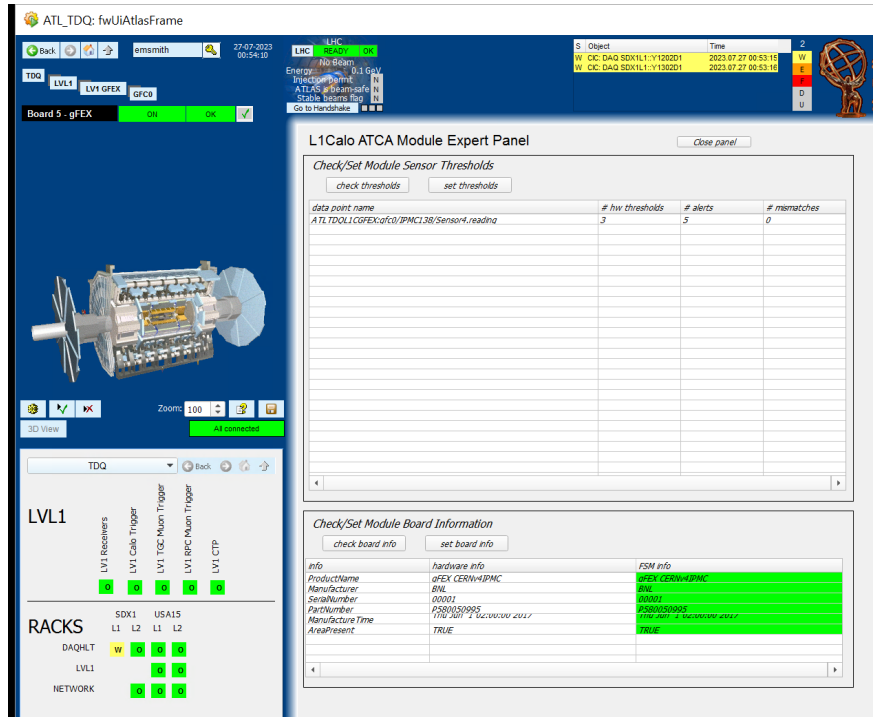


(a) gFEX MiniPOD panel

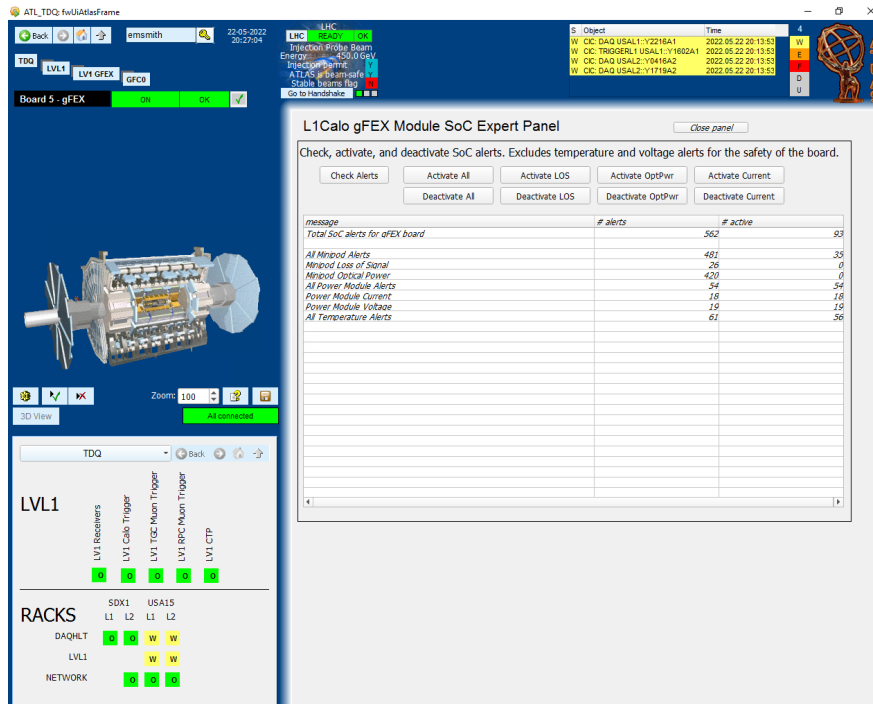


(b) Plot of archived data

Figure 5.38: gFEX MiniPOD DCS panel and example of archived data.



(a) gFEX ATCA expert panel



(b) gFEX SoC expert panel

Figure 5.39: gFEX expert panels for ATCA and SoC specific applications.

5.9 Integration and Commissioning

The Integration and Commissioning process refers to the process of bringing the new hardware, firmware, and software into a final running state, and also making sure that all the different modules can run well together and as a part of the larger ATLAS trigger system and the ATLAS detector as a whole. The final production gFEX boards were brought to CERN as early as 2018 and actually took some data with the Liquid Argon demonstrator board which is shown in Figure 5.40. This used very preliminary and simplified firmware, but was demonstrative of the gFEX hardware functionality and ability to run with ATLAS data. More targeted module tests began in 2019 as well as integration tests between different L1Calo modules.

The gFEX system is directly connected to a number of other modules. The gFEX inputs come from the LATOME and TREX boards routed through the Fiber Optic Exchange (FOX) box, which provide calorimeter data from the Liquid Argon and TILE calorimeters respectively. The gFEX output on the real-time path are sent to the L1Topological trigger module, and the outputs on the read-out path are sent directly to FELIX and then to the SWROD, and are then used as inputs to the ATLAS High Level Trigger. Extensive testing was needed with all of these boards to ensure that data protocols were implemented correctly and that interface between modules functioned properly.

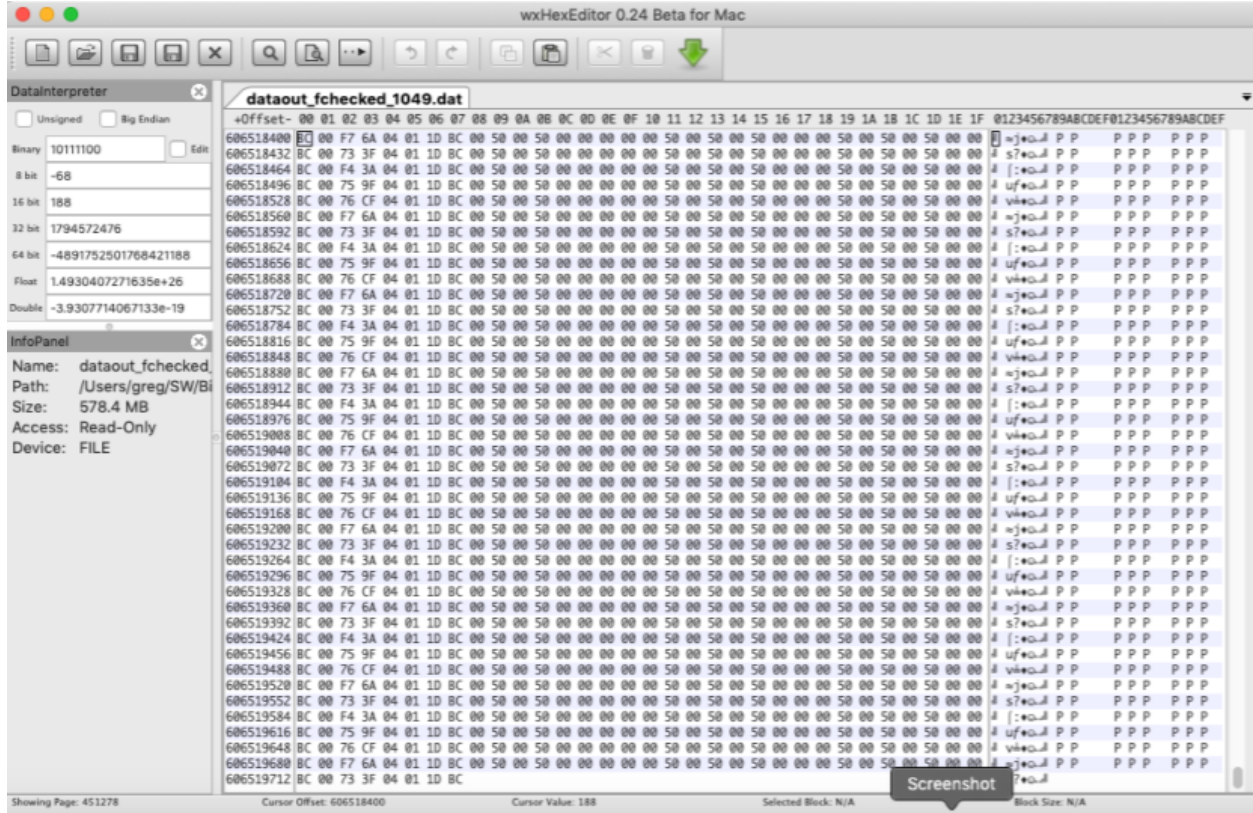


Figure 5.40: gFEX early data taken during 2018 in Run 2. Specifically this data was taken during the period 15 Oct at 22h to 16 Oct at 11h UTC. This corresponds to run 363664 with the tag data18_13TeV. Figure Credit: Greg Myers.

5.9.1 *Surface Test Facility*

The Surface Test Facility (STF) is a lab in the basement of building 4 on the CERN Meyrin site. It is "surface-level" compared to the detector and electronics cavern which are 100 m below the surface. This lab contains three racks with several shelves containing at least one module of each type that exists in the L1Calo system. One of the gFEX production boards was installed in the STF in June of 2019. A FELIX was also installed and some time was devoted to getting the gFEX-FELIX interface working. The FELIX readout was used to validate the gFEX firmware for the majority of the firmware testing. There were also dedicated tests with LATOME, and L1Topo, as well as L1Calo "Slice Tests" which had the goal of getting a full "slice" of the system working, from TREX/LATOME \rightarrow FEX modules \rightarrow L1Topo.

The gFEX \leftrightarrow FELIX interface is particularly important because the gFEX board derives its clock signal from the optical links connected to the FELIX board. For this reason the FELIX must be functional to utilize any of the other firmware operations. The FELIX derives its clock source from the overall L1Calo/ATLAS clock source. This overall clock source is generated by the ATLAS Local Trigger Interface (ALTI) and is controlled by the TDAQ partition. This partition is a GUI which implements software for the control of the entire ATLAS detector, and is shown in Figure 5.41. The TDAQ partition requires a working IPBus connection to the module, as well as software written to perform specific setup at a given time. Initial gFEX firmware development and testing was performed using local operations run on the OS for several months until the full IPBus interface and partition software was completely functional.

Firmware development at the individual engineer level was ongoing throughout this time, but full scale firmware testing and development started in 2019 when the gFEX was installed in the STF. This included significant testing of the Jets Without Jets algorithm, as well as the

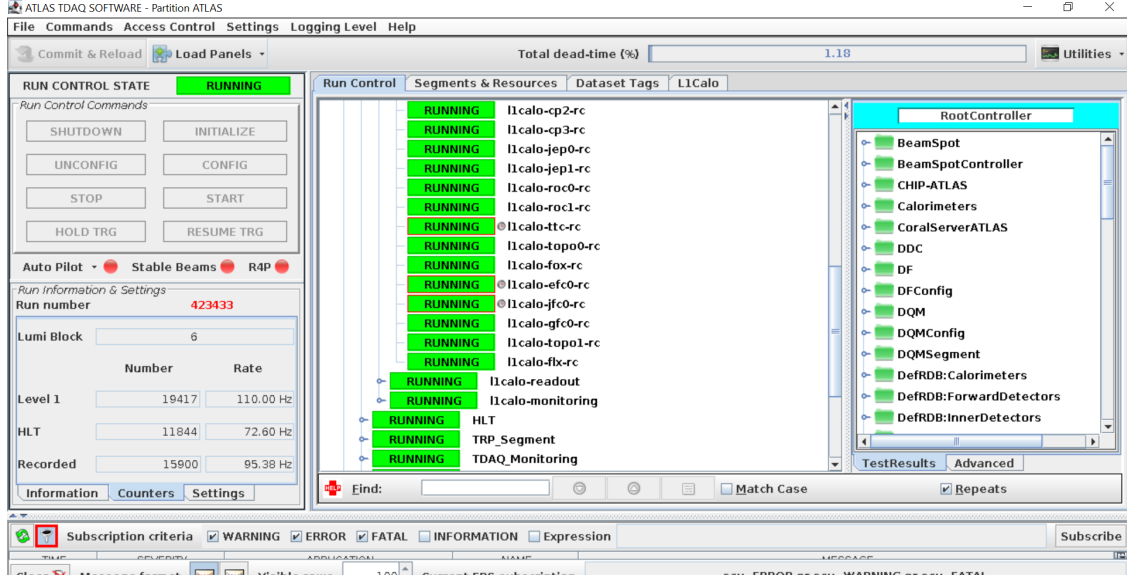


Figure 5.41: TDAQ partition, often referred to as the "GUI", which controls and monitors any hardware included in the run. Also used in the ATLAS control room to run and monitor the whole detector.

MGT interfaces, and link alignment and setup procedures. Eventually this included testing of the Large Radius Jets algorithm and general pFPGA firmware. This required the OS → pFPGA driver to be written and tested, which was done in conjunction with the ongoing firmware testing.

I was particularly involved in the firmware testing on the Zynq FPGA since it is more directly related to the OS, but I was also very involved in the Large Radius Jets algorithm testing and the validation and setup of the associated playback memories.

5.9.2 *gFEX Installation*

The new L1Calo hardware has been installed near the ATLAS detector in the ATLAS electronics cavern 100 m underground at P1, the first interaction point on the LHC. The gFEX board has its own shelf, gfc0, and shares its rack with a second shelf containing jFEX boards.

The gFEX production boards have changed location a number of times, for various reasons as detailed in Table 5.8. The v4C board was originally installed at P1, but was removed when a high current reading of 62.5 A was seen on a power module, when the maximum should be 32 A. This issue was traced back to a resistor that was not assembled well on the monitoring sensor, and was repaired by Schaochun Tang at Brookhaven National Lab (BNL). After its repair the v4C was returned to the STF, and is now the final installed module at P1. The v4B board was removed from P1 due to an issue with the reset continually failing, which was not seen in the replacement board. The v4D board saw some issues with the SD card slot and with the pFPGAs clock signal, which was solved by again installing the v4C board at P1.

gFEX Board	Swap	Reason	Date
v4D	BNL \rightarrow STF	Original STF Installation	\sim June 2019
v4C	BNL \rightarrow P1	Original P1 Installation	7-28-2021 elog
v4C	P1 \rightarrow BNL	High current reading	4-25-2022 elog
v4B	BNL \rightarrow P1	Replace v4C at P1	4-25-2022 elog
v4B	P1 \rightarrow STF	Reset failing	11-22-2022 elog
v4D	STF \rightarrow P1	Replace v4B at P1	11-22-2022 elog
v4B	STF \rightarrow BNL	Investigate reset failing	\sim November 2022
v4C	BNL \rightarrow STF	Replace v4B in STF	\sim November 2022
v4D	P1 \rightarrow STF	SD card error, lost pFPGA clock	2-27-2023 elog
v4C	STF \rightarrow P1	Replace v4D at P1	2-27-2023 elog
v4B	BNL \rightarrow CERN	No replacement.	8-01-2023

Table 5.8: gFEX Board Location

5.9.3 Fiber and Channel Mapping

The fiber and channel mapping of the gFEX system underwent an extensive validation process to ensure that all calorimeter data was routed properly. The calorimeter data comes from Liquid Argon supercells, which are created on the LATOME boards, and TILE towers which are summed in the pre-processors and digitized on the TREX modules. The fibers

are connected from the LATOMEs and TREXs to the Fiber Optic eXchange box (FOX) where the signals are rerouted and regrouped into the signals that are sent to the FEX systems. On gFEX, each fiber comes from a different LATOME board, and there are 96 total input fibers to each pFPGA. These signals are routed from the gFEX faceplate, through miniPODs to the MGTs and their associated FPGAs, as shown in Figure 5.42. There is a specified alignment frame that is sent from the LATOME and TREX that holds designated mapping information like module number, fiber number, and FEX number, which allows the mapping to be fully verified. The gFEX firmware decodes these signals, and fills registers with the relevant information. To fully validate the mapping, the LATOME \rightarrow FOX \rightarrow gFEX Hardware mapping must be understood, but the firmware mapping of the signals from each fiber must also be fully understood since this impacts the decoded mapping information.

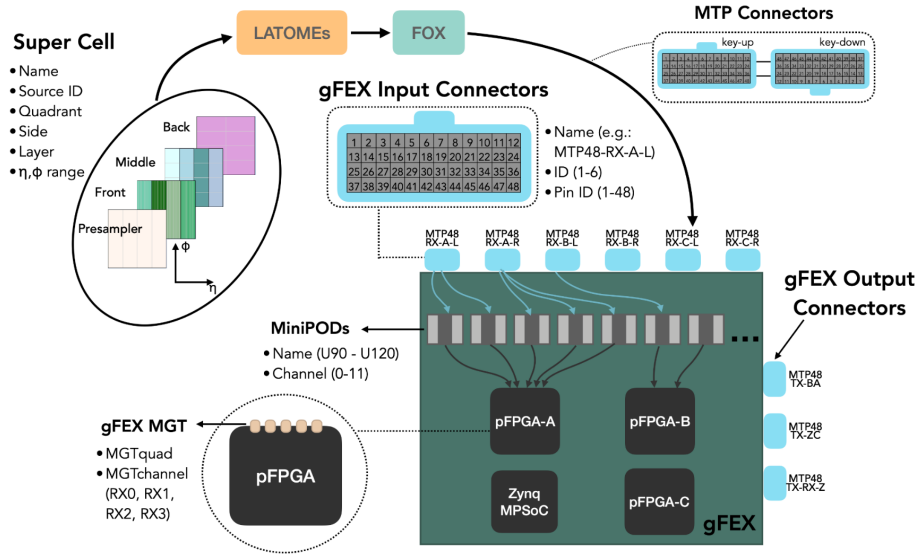


Figure 5.42: Overview of gFEX mapping. Image Credit: Cecilia Tosciri.

The gFEX input mapping from the Liquid Argon and TILE calorimeters has been fully verified. The output mapping to L1Topo was designed less strictly but has now been verified to match the expected mapping on the L1Topo side. This mapping is verified each time the gFEX board has been replaced at P1, and continues to show 100% agreement with the

expected mapping.

5.10 gFEX in Run 3

The global Feature Extractor has been installed and running alongside the legacy system since the beginning of Run 3 in July of 2022. A large portion of the 2022 runs were at lower energies for commissioning purposes of the LHC and experiments. As we moved into 2023 and full physics runs, several of the new FEX triggers were enabled including the gFEX small radius jet triggers and large radius jet triggers.

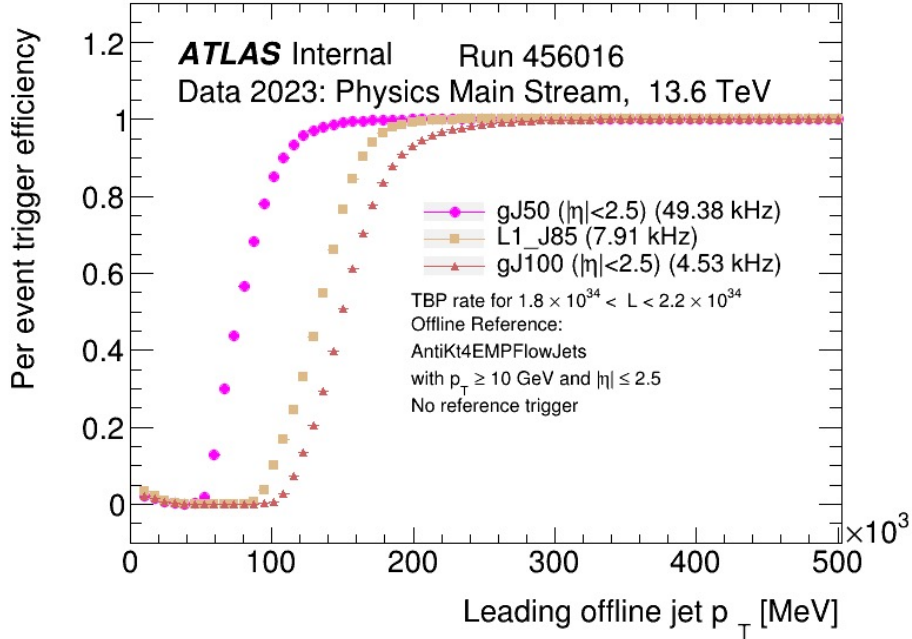


Figure 5.43: gFEX small-radius jet efficiencies compared to legacy jet efficiencies in Run 456016. Plot Credit: Kristin Dona.

The gFEX small radius jet triggers have L1_gJ as a prefix, the gFEX large radius jet triggers use the prefix L1_gLJ and the legacy jet triggers have the L1_J prefix. The gFEX triggers have been functioning nicely and have efficiencies better than the comparable legacy triggers,

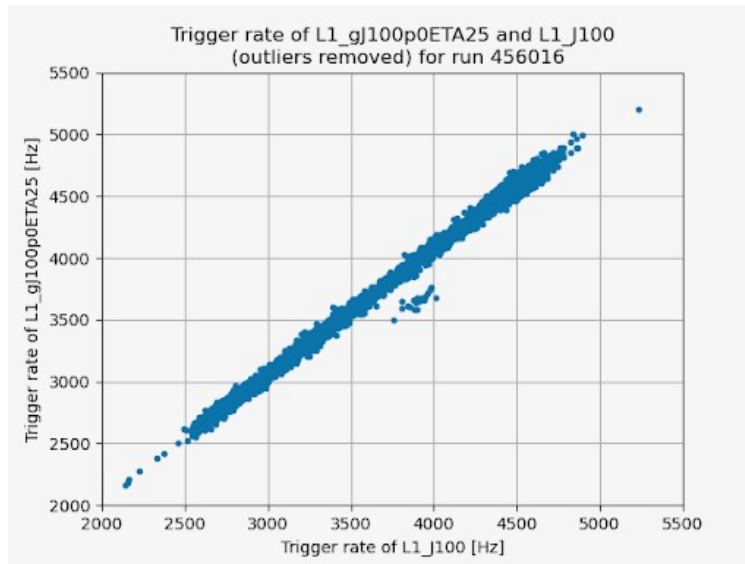


Figure 5.44: gFEX small-radius jet trigger rate versus legacy jet trigger rate in Run 456016. Plot Credit: Kristin Dona.

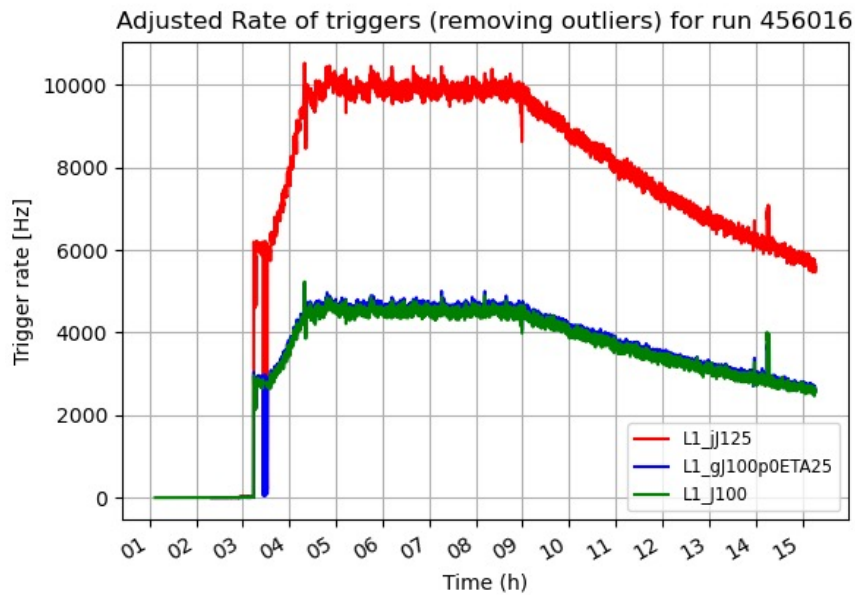


Figure 5.45: gFEX small-radius jet trigger rate compared to both jFEX small-radius jet trigger rate, and legacy jet trigger rate in Run 456016. Plot Credit: Kristin Dona.

as shown in Figure 5.43, with similar rates, as shown in Figure 5.44. A direct comparison of identical triggers between gFEX and the legacy system can be difficult for a variety of reasons. The thresholds are not identical, and the objects are fundamentally different, but comparison of the most analogous triggers can still be made. An additional comparison of rates to the legacy jet triggers and to jFEX triggers can be seen in Figure 5.45.

gFEX is regularly included in physics runs with the ATLAS detector. Due to electrical issues at the LHC caused by thunderstorms and falling trees, 2023 will have very few proton beams in the second half of the year, but gFEX will be included for heavy ion running, and in proton beam physics runs next year. Commissioning continues for the remaining gFEX large-radius jet and E_T^{miss} triggers that have not yet been enabled in ATLAS. In 2024 L1Calo plans to enable all FEX triggers and slowly disable the remaining legacy triggers. I look forward to seeing the progress made by the rest of the gFEX team towards this goal, and to seeing gFEX triggers used in physics analyses in the future!

CHAPTER 6

JETS: HADRONIC PHYSICS WITH THE ATLAS EXPERIMENT

The LHC is fundamentally a proton collider, meaning that any conversation around LHC physics must interact with or at the very least acknowledge the importance of QCD and the hadronic final state. This often involves a discussion of quarks and gluons, but fundamentally these two particles are never directly observable. Due to the nature of QCD, quarks and gluons fragment and hadronize almost immediately, resulting in a collimated spray of energetic particles that when grouped together is called a **jet** [176].

Protons are often considered to be composed of three quarks, but in reality these "valence" quarks account for only about 1% of the proton mass. More accurately, protons consist of a cloud of constantly changing, constantly interacting quarks, antiquarks, and gluons. Quarks and gluons generally have divergent branching probabilities in perturbative QCD, and so must be described by a set of *parton distribution functions* (PDFs) which give the probability density of finding constituent quarks or gluons with a certain momentum fraction at a given energy scale (Q^2). These PDFs and a specified physics process characterize the hard scattering events that occur when two protons interact, for example during an LHC collision event. The hard scatter is only the first of several corresponding processes that occur in a proton-proton collisions. The partons (quarks and gluons) undergo many collinear (parallel) splittings as part of the process of fragmentation. These high energy quarks and gluons may radiate particles, particularly gluons, which subsequently radiate additional gluons or decay to quarks. There are often some number of soft emissions in this process. These effects can be highly non-perturbative and essentially occur randomly, making predictions very difficult. This process is referred to overall as the parton shower

and encompasses both perturbative and non-perturbative effects, many of which must be modeled as opposed to analytically calculated. The quarks and gluons resulting from the parton shower are then hadronized, a process during which the quarks and gluon combine to form color-neutral hadrons. This process is inherently non perturbative, and again must be modeled. This shower of hadrons is what is experimentally measured in a detector, and this is what is referred to as a **jet**.

6.1 Monte Carlo Modeling

As described above and in Section 2.4, many of the interactions which occur in the fragmentation and hadronization process are governed by the nonperturbative regime of QCD. These processes, however, are vital to the experimental physics research which involves colliding hadrons at high energy.

Modern High Energy Physics relies heavily on simulation for comparisons and expected values of these QCD processes. Simulations in this area are usually performed using random sampling Monte Carlo (MC) techniques to produce predictions of expected values. Several MC generators exist and different ones may be used in different scenarios depending on the underlying physics being probed, or on the resource usage available at a given time.

Parton showering is the first stage of jet formation, as described above, and is generally modelled using an ordered shower. This ordered shower helps to avoid divergences which appear in soft gluon radiation and issues which are related to double counting emissions. Showers are usually either angular ordered, with the widest emissions occurring first, or p_T ordered, with the hardest emissions occurring first.

Additional models are needed for the nonperturbative hadronization process. Two of the most common are the Lund string model and the cluster model of hadronization. In the

Lund string model [52], a "string" (or flux tube) exists between the quarks in space. As the quarks move further apart, the energy in the string increases, with some probability at every point to "break" into a vertex of a quark/anti-quark pairs, and this process continues. The cluster model [53] uses the idea of *preconfinement* to form color singlet clusters consisting of the quarks and anti-quarks (after all gluons have decayed) that resulted from the parton shower, which then decay into the final state hadrons. The clusters are fragmented into several final state hadrons using a number of assumptions in different scenarios.

Pythia [179] and Herwig [96] are two ubiquitous MC generators which are often used. Pythia uses a p_T ordered parton shower and the Lund string model of hadronization, while Herwig uses an angular ordered parton shower and the cluster model of hadronization. Other relevant generators include Sherpa [122], MadGraph [8], and MC@NLO [119] which use a mix of these and other models for parton showering and hadronization.

6.2 Jet Algorithms and Reconstruction

A **jet** is generally considered to be a collimated shower of particles in a detector, originating from the fragmentation and hadronization of a quark or gluon. The jet itself is not uniquely defined but depends upon several things, both the detector used to record it, as well as the algorithm used to reconstruct it.

The requirements for jet algorithms have been the consideration of many years and the work of many individuals [134], but generally these algorithms require the following:

- Ease of implementation, experimentally and theoretically
- Defined at any order of perturbation theory
- Yield finite cross sections at any order of perturbation theory

- Be relatively insensitive to hadronization

Ideally the definition of a jet is thus infrared and collinear (IRC) safe, meaning that the jet doesn't change with the addition of soft radiation, and collinear splittings.

Most modern jet definitions are sequential recombination algorithms which use a distance metric and a recursive algorithm to step through the particles present and combine them together. The three main modern algorithms use a distance metric defined as:

$$d_{ij} = \min(p_{Ti}^{2P}, p_{Tj}^{2P}) \frac{\Delta R_{ij}^2}{R} \quad (6.1)$$

where R is a configurable parameter which determines the size of the jet, and P is a configurable parameter which ultimately corresponds to different algorithms, and will be discussed more below. Common choices of jet size in ATLAS are $R = 0.4$ and $R = 1.0$ jets.

The algorithm using this metric is performed as follows:

1. Find the distance d_{ij} between all pairs of particles such that $i \neq j$
2. Find the pair of particles i, j which minimizes this distance.
3. If the minimum d_{ij} is less than a threshold placed on d , then combine i and j into one pseudo-jet
4. If no pairs of particles meet this criteria, the algorithm is terminated.

The k_t algorithm [174] uses $P = 1$ which corresponds to soft particles being clustered first. The anti- k_t algorithm [77] uses $P = -1$ which corresponds to hard radiation being clustered first. The Cambridge/Aachen (C/A) algorithm [199, 196] use $P = 0$ which corresponds to no dependence on p_T , and results in only an angular dependence of the clustering algorithm.

The k_t algorithm can be nice for calibration but the anti- k_t and C/A algorithms are much nicer in terms of theoretical calculations. Anti- k_t is the default clustering algorithm currently used in ATLAS, though C/A is often used for specific types of studies.

CHAPTER 7

SEARCH FOR CHARGINO AND NEUTRALINO PRODUCTION IN FINAL STATES WITH A HIGGS BOSON AND MISSING TRANSVERSE MOMENTUM AT $\sqrt{s} = 13$ TeV WITH THE ATLAS DETECTOR

This analysis is one that I worked on a great deal in the earlier days of my graduate degree. It is a search for the electroweakino superpartners of the W, Z, and Higgs bosons. This search was performed in both a boosted channel and a resolved channel, which will be explained below, and I worked primarily on the resolved channel. Between channels, we shared samples and a general analysis strategy, but in the end the resolved channel did not get published, as it did not significantly increase sensitivity to the search in the targeted phase space. This will be discussed in more detail in Section 7.5, but this chapter will primarily document my work on the resolved channel. I will also present the boosted channel results so naturally those sections are heavily inspired by the published results from the boosted channel which can be found in Ref [90].

7.1 Introduction

Supersymmetry as discussed in Subsection 2.5.1 is a theoretical extension of the Standard Model which introduces supersymmetric partners to the SM particles. In the MSSM the bino, wino, and higgsino are the main electroweakinos and these form the chargino ($\tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm$) and neutralino ($\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$) mass eigenstates through mixing¹.

1. The subscripts here indicate increasing mass

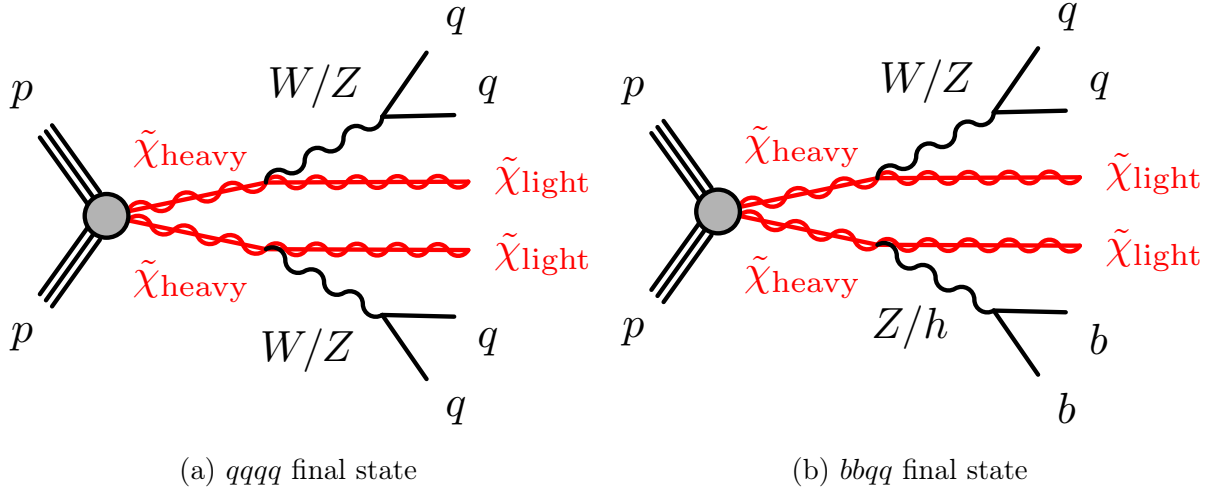


Figure 7.1: Diagrams for the targeted electroweakino production and final states [90].

SUSY can be motivated in many ways, one such way is that it provides solutions to many phenomenological issues in the Standard Model. Weak scale SUSY is particularly interesting in contexts where R-parity conservation is assumed [112]. Light electroweakinos then provide a dark matter candidate with the lightest neutralino ($\tilde{\chi}_1^0$) as the LSP.

This search targets the pair production of electroweakinos ($\tilde{\chi}_{\text{heavy}}$) where each of them decays into a lighter one ($\tilde{\chi}_{\text{light}}$) and an on-shell SM W , Z or Higgs boson (h). The $\tilde{\chi}_{\text{heavy}}$ can be either wino- or higgsino-like, and $\tilde{\chi}_{\text{light}}$ can be a bino-, wino-, higgsino-like chargino/neutralino, gravitino or axino depending on the model². Cases where both $\tilde{\chi}_{\text{heavy}}$ and $\tilde{\chi}_{\text{light}}$ are wino- or higgsino-like are not considered as they lead to very small $\Delta m(\tilde{\chi}_{\text{heavy}}, \tilde{\chi}_{\text{light}})$. The $\tilde{\chi}_{\text{light}}$ as the LSP leads to missing transverse momentum ($\mathbf{p}_T^{\text{miss}}$, with magnitude E_T^{miss}) in the decay signature, as can be seen in Figure 7.1.

Fully hadronic final states are considered with $W \rightarrow q\bar{q}$, $Z \rightarrow q\bar{q}$ or $Z \rightarrow b\bar{b}$ and $h \rightarrow b\bar{b}$, such that overall two final states are considered, that of $qqqq$ (4Q) and that of $qqbb$ (2B2Q), as shown in Figure 7.1.

2. The gravitino and axino interpretations will not be discussed in this thesis but can be found in Ref [90]

The fully hadronic final state is particularly advantageous in searching for signals with substantial mass splitting between the produced gauginos and the LSP. Since the SM background can be highly suppressed by requiring tight kinematic selections instead of requiring leptons to be present in the final state, the large hadronic branching ratio of $W/Z/h$ can provide enhanced acceptance for certain signals. The number of produced signals are typically $10 \sim 40$ times higher than the multi-lepton final state mode.

The heavy electroweakinos in the diagram in Figure 7.1 can be:

- $\tilde{\chi}_1^\pm \tilde{\chi}_1^\pm \rightarrow WW \tilde{\chi}_1^0 \tilde{\chi}_1^0$
- $\tilde{\chi}_1^\pm \tilde{\chi}_2^0 \rightarrow WZ \tilde{\chi}_1^0 \tilde{\chi}_1^0$
- $\tilde{\chi}_1^\pm \tilde{\chi}_2^0 \rightarrow Wh \tilde{\chi}_1^0 \tilde{\chi}_1^0$

7.2 Analysis Methods

This analysis is separated out into two channels, a boosted and a resolved channel. In the resolved channel the $qqqq$ and $qqbb$ final states all require 4 individual jets, with 2 b-tagged jets for the $qqbb$ final state, whereas the boosted channel requires two large-radius jets which are tagged as W , Z , or Higgs bosons using jet-substructure-based boson tagging techniques.

Both the resolved and boosted channels use data collected by triggers targeting large missing transverse momentum. Monte Carlo (MC) simulations are used to estimate the SM background and the signals used in the analysis. A common preselection is defined which makes certain cuts based on kinematic variables. Each of the channels have independent signal regions (SRs) for both the 4Q and 2Q2B final states. These signal regions are also defined using kinematic variables and are designed to limit contributions from background events and to contain a high percentage of signal events.

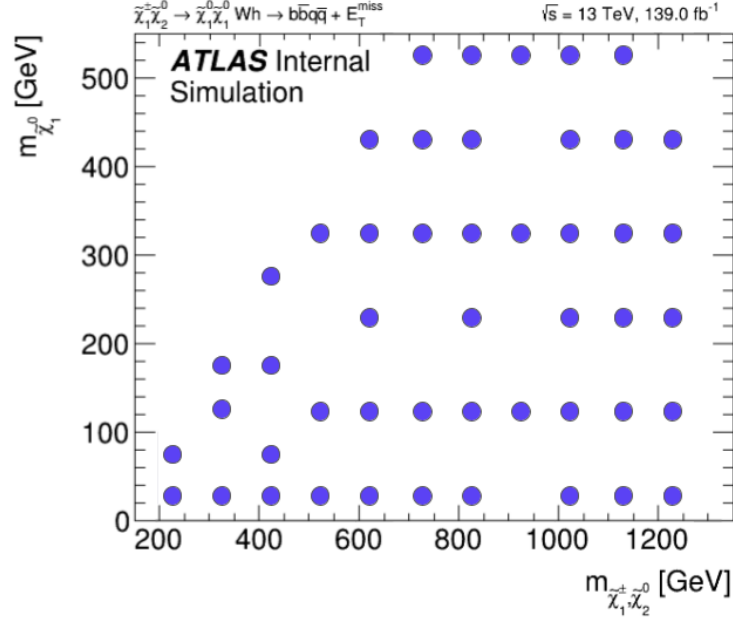


Figure 7.2: Graphical representation of the mass points included in this search for a variety of $\tilde{\chi}_{\text{heavy}}$ and $\tilde{\chi}_{\text{light}}$ masses. Often referred to as a signal grid.

MC samples of the supersymmetric signal processes are generated at a variety of mass points for different $\tilde{\chi}_{\text{heavy}}$ and $\tilde{\chi}_{\text{light}}$ masses. When put together these values create a signal "grid" where one axis (usually the x-axis) corresponds to the $\tilde{\chi}_{\text{heavy}}$ ($\tilde{\chi}_1^\pm, \tilde{\chi}_2^0$) mass and one axis (usually the y-axis) corresponds to the $\tilde{\chi}_{\text{light}}$ ($\tilde{\chi}_1^0$) mass. An example of this signal grid can be seen in Figure 7.2, with a purple dot on each mass point included. This type of signal grid plot is often used to display efficiency or sensitivity values for all mass points at once, and will often be shown with a numerical value at each of the points in the grid.

The background estimation strategy uses control regions (CRs) aimed at maximizing the contributions from the main SM backgrounds. Which backgrounds are relevant depends on the analysis channel. The backgrounds are estimated using MC in the CRs which are then corrected using a data-driven method which is validated in the validation regions (VRs). These control regions are defined by either the mass side-band region of the reconstructed bosons, or the equivalent 1 lepton/2 lepton regions.

The resolved channel focus is mainly on the $qqbb$ final state for initial studies. The dominant backgrounds in this region are $t\bar{t}$, single-top and Z +jets. Each of these is given a dedicated control region to verify the background contributions, and which can be used to generate a data-driven estimation if required. Sub-dominant backgrounds are W +jets, diboson production and $t\bar{t}+X$, where X is a vector boson.

The boosted channel's dominant background in all signal regions is Z +jets, where $Z \rightarrow \nu\nu$. The sub-dominant backgrounds are diboson production and W +jets production in the $qqqq$ final state and $t\bar{t}+X$ in the $qqbb$ final state. The side-band control and validation regions are defined in the mass side-band of either or both large-radius jets.

All samples are processed using a common central derivation managed by the SUSY working group. This derivation is fully defined in the SUSY derivation framework, but requires that the events pass the following skimming criterion:

- $H_T > 150$ GeV where H_T is the scalar sum of the p_T of all jets in the event
- at least one electron, muon, or photon with $p_T > 100$ GeV
- at least two electrons or two muons with $p_T > 20$ GeV
- at least two photons with $p_T > 50$ GeV

7.2.1 MC Samples

Centrally produced MC samples are used for modeling all signals and SM backgrounds. The processes considered as backgrounds are listed below, with the corresponding generators.

- $t\bar{t}$ - pair production of top quarks. POWHEG + PYTHIA.
- Wt - single top-quark production in conjunction with a W boson. POWHEG + PYTHIA.

- $t\bar{t}+X$ - pair production of top quarks with an associated vector bosons, MADGRAPH + PYTHIA.
- W +jets - W boson production with associated jets, SHERPA 2.2.1.
- Z +jets - Z boson production with associated jets, SHERPA 2.2.1.
- diboson production - processes with two bosons, mainly WW, WZ, ZZ , SHERPA 2.2.2.
- dijet production - QCD dijet production, PYTHIA 8, estimated using data.

The SUSY signal samples are simulated with leading-order (LO) matrix elements with up to two extra partons using MADGRAPH 2.6.1 + PYTHIA 8.230. The Parton Distribution Function set used for the generation of the signal samples is NNPDF2.3LO [165]. The matrix element to parton shower matching was performed with the CKKW-L prescription [148] with a matching scale set to $1/4$ of the mass of the produced electroweakinos. All signal cross sections are calculated to approximate next-to-next-to-leading order in the strong coupling constant, adding the resummation of soft gluon emission at next-to-next-to-leading-logarithmic accuracy [63, 61, 64, 62]. The nominal cross section and the uncertainty are derived using the PDF4LHC15_mc PDF set, following the recommendations of Ref. [75].

All MC samples were generated centrally by ATLAS and were propagated through the ATLAS detector simulation [43] based on GEANT4 [4] [2]. Multiple proton-proton collisions in the same and neighboring bunch crossings (pileup) were modeled by overlaying the hard-scatter events with minimum-bias events simulated by PYTHIA 8.186 [44]. The simulated events are processed with the same trigger and reconstruction algorithms as the data.

7.2.2 Data Samples

The dataset used in this analysis is from proton–proton (pp) collisions provided by the LHC with $\sqrt{s} = 13$ TeV between 2015 - 2018 and was collected by the ATLAS detector, which is discussed in detail in Chapter 4. The total integrated luminosity after applying specific criteria ensuring the data can be used for physics analysis is 140 fb^{-1} . The uncertainty in the combined 2015–2018 integrated luminosity is 0.83% [26], obtained using the LUCID-2 detector [51] for the primary luminosity measurements. Due to the high instantaneous luminosity and the large total inelastic pp cross section, there are, on average, 33.7 simultaneous (‘pileup’) collisions in each bunch crossing. Data which are included in this dataset but identified as contaminated with errors from the liquid argon calorimeter, tile calorimeter or semiconductor tracker (SCT) are rejected using event-level flags during analysis [13].

Events with jets oriented to dead tile modules in Table 7.1 are vetoed to avoid pathological mis-measurement of $E_{\text{T}}^{\text{miss}}$. This is common when using jet or $E_{\text{T}}^{\text{miss}}$ objects and prevents issues that may arise when physics objects are pointed at these dead modules. The efficiencies of wino-NLSP/bino-LSP signal samples are about 95% after this veto is applied and can be seen in Figure 7.3.

Year	RunNumber	η region	ϕ region
2015	266904 - 284484	$0. < \eta < 0.9$	$0.8 < \phi < 1.0$
		$-1.6 < \eta < -0.9$	$1.9 < \phi < 2.1$
2016	302053 - 311481 306988 - 311481	$0. < \eta < 0.9$	$-1.33 < \phi < -1.13$
		$-0.9 < \eta < 0.$	$0.34 < \phi < 0.54$
2017	325713 - 340453	$-0.9 < \eta < 0.$	$-0.25 < \phi < -0.05$
		$0.8 < \eta < 1.7$	$0.14 < \phi < 0.34$
2018	350310 - 352514 355261 - 364292	$0. < \eta < 0.9$	$2.7 < \phi < -3.0$
		$0. < \eta < 0.9$	$\phi < -3.0 \text{ or } 3.0 < \phi$

Table 7.1: Location of dead tile modules over the course of Run 2. Partially dead modules are not listed here.

Additional cleaning cuts are implemented to reject non-collision backgrounds, such as beam-

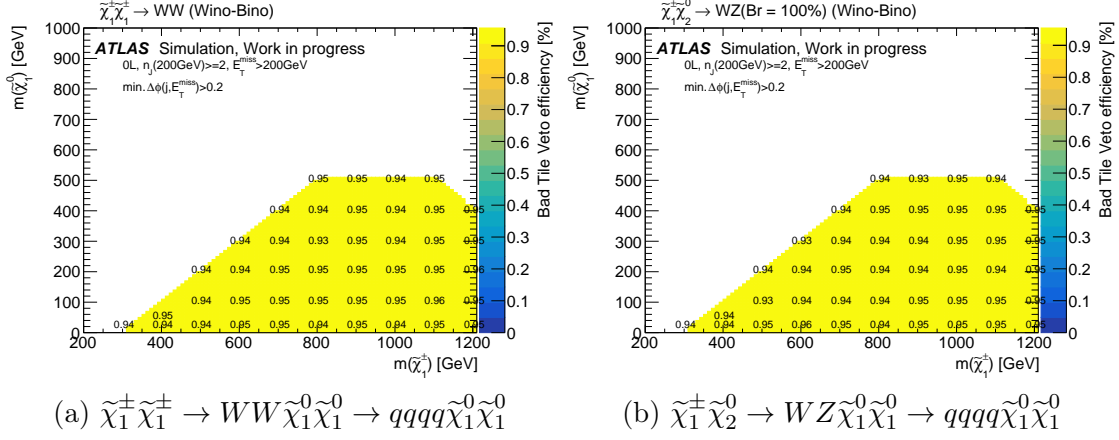


Figure 7.3: The efficiency of wino-NLSP/bino-LSP signal samples with the dead tile module veto.

induced background, cosmic particles and detector noise. These backgrounds cause abnormal data excesses in the tails of some distributions, and can have very large energies. This is a non-negligible effect for a fully hadronic final state. Track-based MET is known to be useful for repressing this type of background since the non-collision jet background is usually found via calorimeter interactions.

The following cleanings cuts are used to veto the non-collision background:

- $E_{T,\text{track}}^{\text{miss}} > 75 \text{ GeV}$
- $\min \Delta R(E_T^{\text{miss}}, E_{T,\text{track}}^{\text{miss}}) < 2.0$

Signal efficiencies of simulated samples with non-collision background veto are shown in Fig 7.4. The efficiencies are at least 90% in low Δm region ($\Delta m \sim 300 \text{ GeV}$), and are 95% elsewhere. This low Δm region is targeted by the multi-lepton electroweakino search, so efficiency here is not as big of a concern.

The signal efficiencies of wino-NLSP/bino-LSP signals with the dead tile jet event veto and the non-collision background veto are about 90%, shown in Fig 7.5.

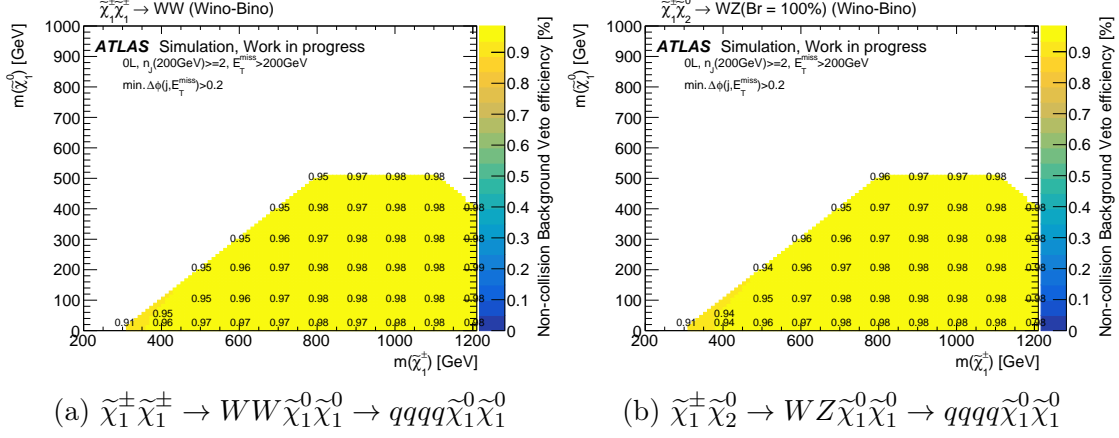


Figure 7.4: The efficiency of wino-NLSP/bino-LSP signal samples with the non-collision background veto.

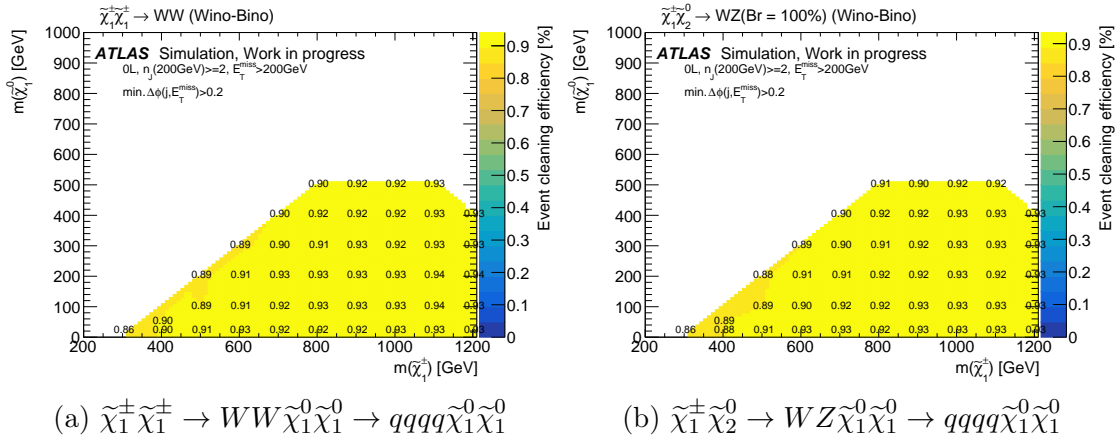


Figure 7.5: The efficiency of wino-NLSP/bino-LSP signal samples with the dead tile jet event veto and the non-collision background veto.

The triggers used to collect this data are mainly based on the E_T^{miss} in an event but some lepton triggers are used for events which fill the control and validation regions.

Events with $E_T^{\text{miss}} > 200$ GeV (offline) are triggered by the E_T^{miss} triggers listed in Table 7.2, where the lowest unprescaled triggers are chosen from the period. The efficiency is $> 98\%$ for $E_T^{\text{miss}} > 200$ GeV. The lowest-unprescaled single-lepton triggers are used to record 1- and 2-lepton events and are listed in Table 7.3.

Chain name	Unprescaled period
MET	
xe70	2015
xe90_mht_L1XE50	2016 A-D1
xe100_mht_L1XE50	2016 D1-F1
xe110_mht_L1XE50	2016 F2-
xe110_pufit_L1XE55	2017
xe110_pufit_xe70_L1XE50	2018 B-C5
xe110_pufit_xe65_L1XE50	2018 C5-

Table 7.2: Summary of MET triggers used in the analysis.

Chain name	Unprescaled period
1e	
e24_lhmedium_L1EM20VH e60_lhmedium e120_lhloose	2015
e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e140_lhloose_nod0	2016-18
1μ	
mu20_loose_L1MU15 mu50	2015
mu26_ivarmedium mu50	2016-18

Table 7.3: Summary of single-lepton triggers used in the analysis.

7.2.3 Object Definition and Reconstruction

The main objects used in this analysis are small-radius jets, and large-radius jets, and E_T^{miss} . These objects will be discussed in detail below. To a lesser extent the analysis also uses reconstructed leptons (electrons and muons) and photons for kinematic selections, validation of the background estimation and the E_T^{miss} calculation. Electron candidates are reconstructed from energy clusters which are consistent with electromagnetic showers in the ECAL and are matched to tracks in the inner detector [18]. Muon candidates are reconstructed by matching tracks in the muon spectrometer to tracks in the inner detector [35]. Photon candidates are reconstructed either as electromagnetic clusters with no matching tracks or as e^+e^- pairs

from photon conversion in the inner detector [18].

Small-radius jets (j) are reconstructed from particle flow objects [25] using the anti- k_t algorithm [77] with radius parameter $R = 0.4$, implemented in the FastJet package [76], which corresponds to jet collection `AntiKt4PFlow`. Small-radius jets must satisfy $p_T > 20$ GeV and $|\eta| < 2.8$. Pileup mitigation is achieved by using jet vertex taggers which associate a jet with a given reconstructed primary vertex. Jets within $|\eta| < 2.5$ originating from b -hadrons are identified with the MV2c10 b -tagging algorithm [12]. This algorithm is a multivariate discriminator that utilizes track impact parameters, the presence of secondary vertices, and the trajectories of b - and c -hadrons inside the jet. It is implemented with the 85% efficiency working point, which means that b -jets from simulated $t\bar{t}$ events are identified with 85% efficiency, with rejection factors of 3 against c -jets and 33 against jets originating from other light-flavor quarks or gluons.

Large-radius jets (J) are clustered using the locally calibrated topo-clusters [45] and the anti- k_t algorithm with radius parameter $R = 1.0$, corresponding to jet collection `AntiKt10LCTopoTrimmedPtFrac5SmallR20Jets`. A trimming algorithm [20] is applied to mitigate the effects of pileup and soft radiation. The constituents of the jet are reclustered using the k_t algorithm with a smaller distance parameter, $R_{\text{subjet}} = 0.2$, to produce a collection of subjets. These subjets are discarded if they carry less than a specific fraction, $f_{\text{cut}} < 5\%$, of the original jet p_T .

The jet mass, m_J , is calculated according to the combined mass prescription [24]. At very high- p_T it is possible for two particles in the decay chain to be reconstructed as one topological cluster due to the angular resolution of the calorimeter cells. The missing angular information makes the jet mass resolution worse. However, in this case charged particles can be reconstructed separately as tracks from the inner detector. Using the track information can help to improve the mass resolution [24]. A combination is used: at high- p_T this track-

assisted mass calculation is implemented, while at low- p_T the purely calorimeter calculated mass gives the best resolution, and is thus utilized.

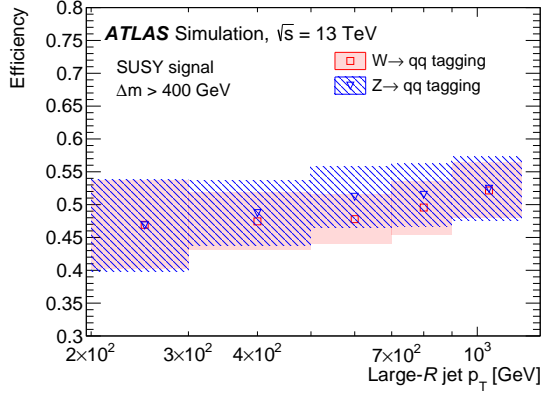
This analysis considers large- R jets with $p_T > 200$ GeV, $|\eta| < 2.0$, $m_j > 40$ GeV.

Track-jets are used to identify subjets in the large-radius jets that contain b-hadrons, specifically trackjets with $p_T > 20$ GeV, $|\eta| < 2.5$. The MV2c10 b-tagging algorithm is again used to do this b-tagging, with the same 85% efficiency working point used for the small-radius jets. A b-tagged trackjet is contained in a large-radius jet if the two jet axes have an angular separation of $\Delta R < 1.0$.

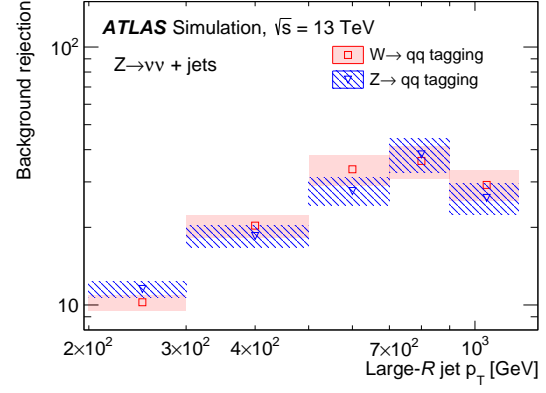
Two types of boosted boson tagging are employed for the preselected large-radius jets. W_{qq}/Z_{qq} -tagging targeting $W \rightarrow qq$ and $Z \rightarrow qq$, and Z_{bb}/h_{bb} -tagging targeting $Z \rightarrow bb$ and $h \rightarrow bb$. The W_{qq}/Z_{qq} -tagging utilizes cuts on m_J , the energy correlation function ratio D_2 [142, 141], and the track multiplicity n_{trk} [41, 15], and requires that fewer than two b-tagged subjets were found. The selections are optimized separately for targeting the W and Z bosons, and the performance of this tagging is summarized in Figure 7.6a and Figure 7.6b. V_{qq} -tagging is defined by a selection satisfying either the W_{qq} or Z_{qq} -tagging. For a sample of preselected large-radius jets the tagging efficiency is about 50% for the signal jets origination from electroweakino decays, while the background rejection is usually a factor of 10 at $p_T = 200$ GeV and a factor of 40 at $p_T = 1000$ GeV per jet in Z+jets events.

The Z_{bb}/h_{bb} -tagging is applied to large-radius jets containing exactly two b-jets by applying a jet mass window cut that selects the peak consistent with the Z or h boson. The Z-tagging and h-tagging require the jet mass to satisfy $70 \text{ GeV} < m_J < 100 \text{ GeV}$ and $100 \text{ GeV} < m_J < 135 \text{ GeV}$ respectively. The performance of the Z_{bb}/h_{bb} -tagging is summarized in Figure 7.6c and Figure 7.6d.

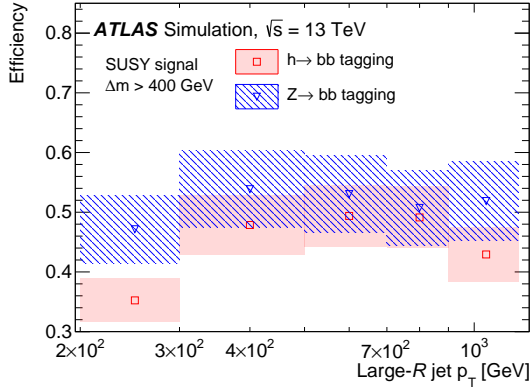
Missing transverse momentum or p_T^{miss} with magnitude E_T^{miss} is constructed as the



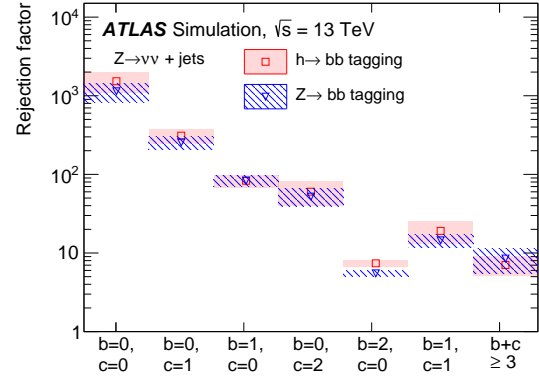
(a) W_{qq}/Z_{qq} -tagging efficiency



(b) W_{qq}/Z_{qq} -tagging rejection



(c) Z_{bb}/h_{bb} -tagging efficiency



(d) Z_{bb}/h_{bb} -tagging rejection

Figure 7.6: The boson-tagging efficiency for jets arising from $W/Z/h$ bosons decaying into $q\bar{q}$ or $b\bar{b}$ (signal jets) and the rejection factor (inverse of the efficiency) for jets that have other origins (background jets) are shown.

negative vectorial sum of the transverse momenta of all objects in the event calibrated to their respective energy scales, and an additional "soft term" constructed from tracks originating from the primary vertex but not associated with any specific objects [38].

To prevent the reconstruction of a single particle as multiple objects, an overlap removal procedure is applied to leptons, photons and jets. Any electron sharing a track with a muon or other electrons is removed. Photons around the remaining electrons and muons are removed if the photon-lepton separation is $\Delta R < 0.4$. Small-radius jets are removed if they are $\Delta R < 0.2$ from a remaining electron or photon or are $\Delta R < 0.4$ from a muon, and the jet has fewer than three tracks with $p_T > 500$ MeV. Leptons are removed if they are close enough with a low enough p_T to small-radius jets, and large-radius jets are removed if they are separated by $\Delta R < 1.0$ from any remaining electrons.

7.2.4 Discriminating Variables

Specific kinematic variables are used to discriminate signal from background. These variables and their nomenclatures are described below. Table 7.4 defines the variables used for the common preselection. The variables used to define signal/control/validation regions for the boosted and resolved channels respectively are shown in Table 7.5 and Table 7.6.

Variable	Definition / Comments
$n_{\text{baseline lepton}}$	Number of baseline leptons.
$n_{\text{Small-}R \text{ jets}}$	Number of small-radius jets.
$n_{b\text{-jet}}$	Number of b-tagged small-radius jets.
$n_{\text{Large-}R \text{ jets}}$	Number of large-radius jets.
E_T^{miss}	Magnitude of the missing E_T vector.
$\min \Delta\phi(j, E_T^{\text{miss}})$	Minimum azimuthal angle between the E_T^{miss} vector and jets. Useful to reject processes with no real E_T^{miss} (e.g. QCD multi-jets) or with high jet multiplicity. Also denoted as $\Delta\phi_{\min}^{4j}$

Table 7.4: Nomenclature of object multiplicity and the kinematic variables used for common preselection.

Variable	Definition / Comments
$m_{b\bar{b}}$	Invariant mass of the two small-radius b -jets. Used in resolved 2B2Q SRs.
m_{jj}	Invariant mass of the leading two small-radius non- b jets.
m_{eff}	Effective mass defined by a scalar sum of p_{T} of small-radius jets and $E_{\text{T}}^{\text{miss}}$: $m_{\text{eff}} := \sum_i p_{\text{T}}(j_i) + E_{\text{T}}^{\text{miss}}$
m_{CT}	Contraverse mass used for $t\bar{t}$ rejection in resolved 2B2Q SRs, defined by: $m_{\text{CT}} := 2 p_{\text{T}}(b_1) p_{\text{T}}(b_2) (1 + \cos \Delta R_{bb})$. The distribution has an end-point at $m_t^2 - m_W^2/m_t$ for $t\bar{t}$.
$m_{\text{T}}^{b,\text{min}}$	Variable used for $t\bar{t}$ rejection in resolved 2B2Q SRs, defined by: $m_{\text{T}}^{b,\text{min}} := \min_i m_{\text{T}}(b_i, E_{\text{T}}^{\text{miss}})$ where b_i is i -th small- R b -jet. The $t\bar{t}$ distribution has an end-point corresponding to the top-mass constraint.

Table 7.5: Kinematic variables used in the resolved category selection.

Variable	Definition / Comments
$n_{b\text{-jet}}^{\text{trk}}$ (inside J)	Number of $R = 0.2$ track b -jets ghost-associated with any of the large- R jets in the event.
$n_{b\text{-jet}}^{\text{unmatched}}$	Number of $R = 0.2$ track b -jets NOT ghost-associated with any of the large- R jets in the event.
$m(J_{1,2})$	Combined track associated + calo mass of large-radius jet.
$m_{\text{eff}}(J)$	Effective mass defined by large-radius jets: $m_{\text{eff}} := \sum_i p_{\text{T}}(J_i) + E_{\text{T}}^{\text{miss}}$.
$m_{\text{T}2}$	The two reconstructed bosons are assigned to the visible legs. The jet masses are also fed to the visible leg masses. Missing particle test mass set to 100 GeV ($m_{\chi} = 100$ GeV) and off-set is subtracted off : $m_{\text{T}2} := m_{T2}(\text{vis1} = J_1, \text{vis2} = J_2, E_{\text{T}}^{\text{miss}}; m_{\text{mis1}} = m_{\text{mis2}} = m_{\chi}) - m_{\chi}$, Powerful BG vs signal separation particularly in the 2B2Q category and after the boson tagging selection where BGs tend to have lower values. The end-point also corresponds to the signal ΔM in case of the signal distributions. [146, 56]

Table 7.6: Object multiplicity and kinematic variables used in the boosted category selection.

7.3 Preselection

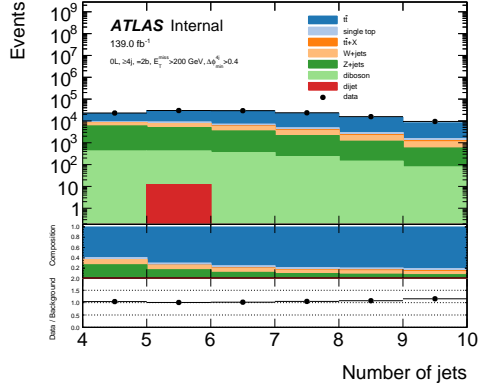
A common preselection is used between the boosted and resolved channels. With this preselection applied, different variables that could be used in each channel to define signal regions can be studied. MC and data can be compared at the preselection level and different kinematic variables can be checked for mis-modeling, or other issues. The cuts defining the common preselection are documented in Table 7.7 below.

Relevant observables for the resolved channel with the preselection applied are shown in Figure 7.7. The data and MC results match quite well in the preselection region.

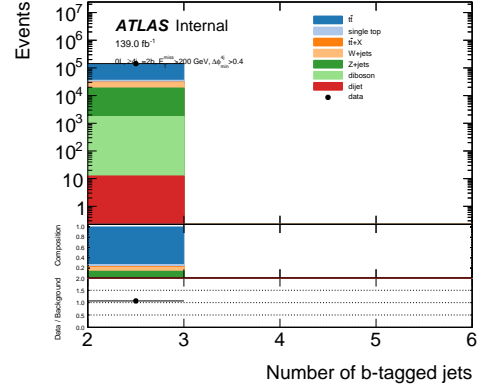
Relevant observables for the boosted channel with the preselection applied are shown in Figure 7.8.

Variable	Requirement
E_T^{miss} Trigger	✓
Event cleaning	✓
$n_{\text{baseline lepton}}$	0
$n_{\text{Small-}R \text{ jets}}$	= 4
$n_{b\text{-jet}}$	= 2
E_T^{miss}	> 200 GeV
$\Delta\phi_{\min}^{4j}$	>= 0.2
$n_{\text{Large-}R \text{ jets}}$	≥ 2 (Boosted only)
$n_{b\text{-jet}}^{\text{unmatched}}$	= 0 (Boosted only)

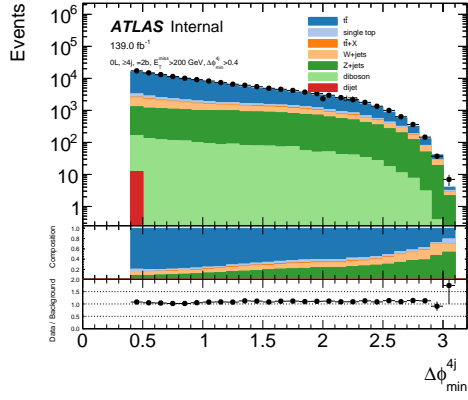
Table 7.7: Definition of the common preselection



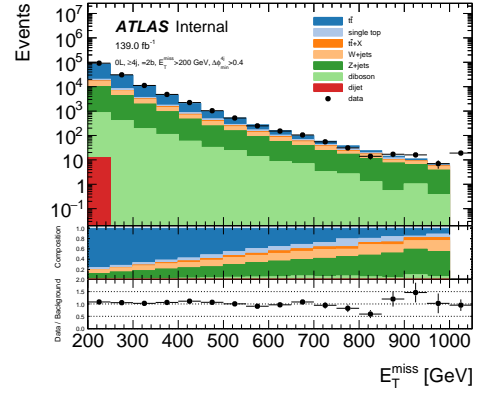
(a) Number of jets



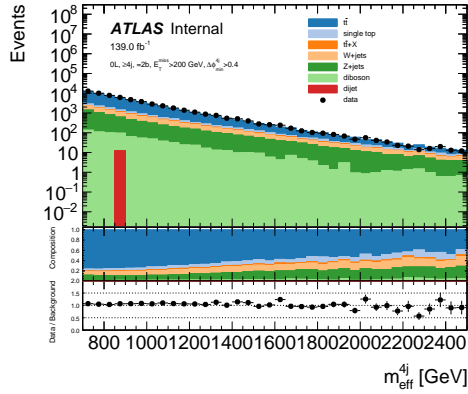
(b) Number of b -tagged jets



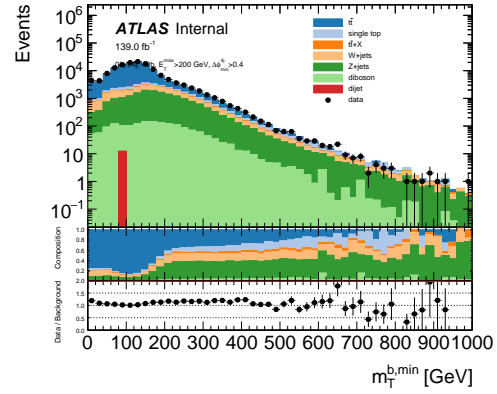
(c) $\min \Delta\phi(E_T^{\text{miss}}, j)$



(d) E_T^{miss}

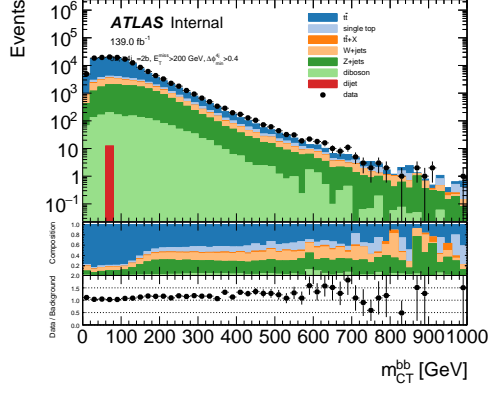


(e) m_{eff}

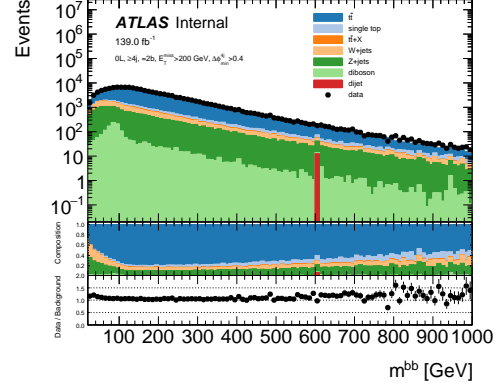


(f) $m_T^{b,\text{min}}$

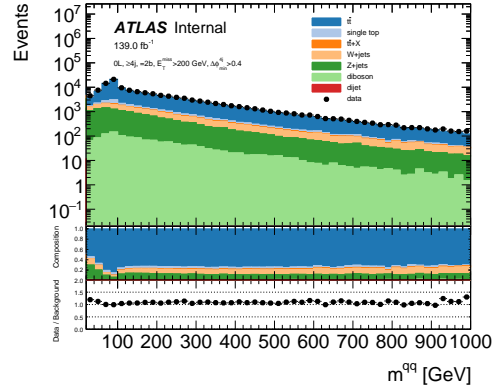
Figure 7.7: Distributions of key observables for the resolved channel with the preselection applied as described in Table 7.7. Figure continued on next page.



(g) m_{CT}^{bb}

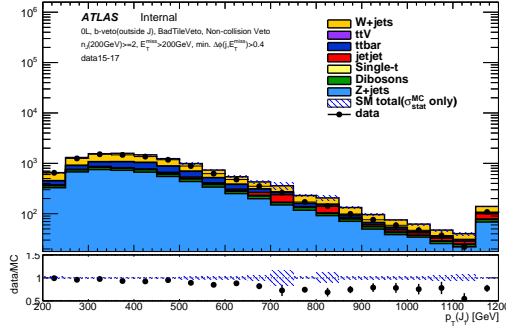


(h) m_{bb}

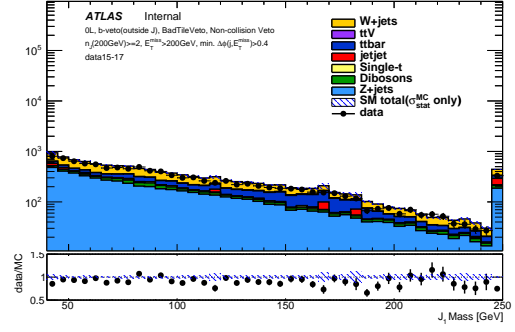


(i) m_{jj}

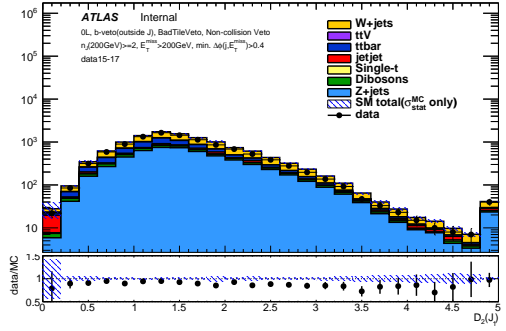
Figure 7.7: (continued) Distributions of key observables for the resolved channel with the preselection applied as described in Table 7.7.



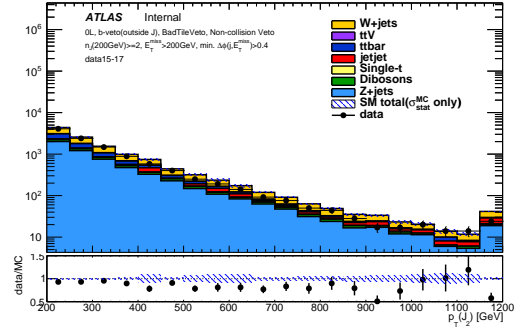
(a) leading large- R jet p_T



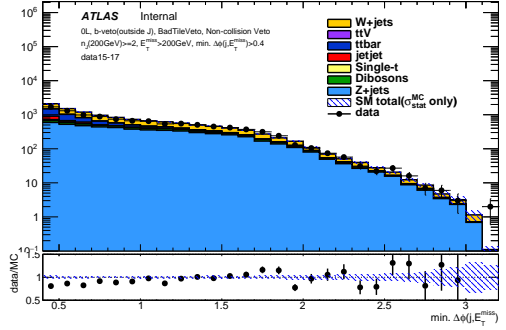
(b) leading large- R jet mass



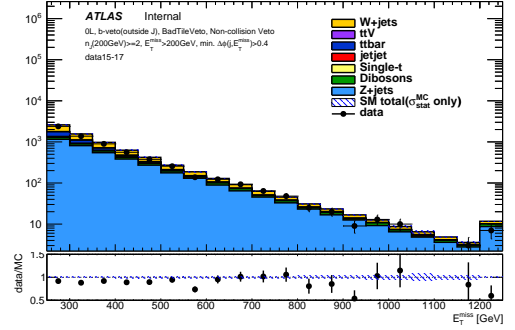
(c) leading large- R jet D_2



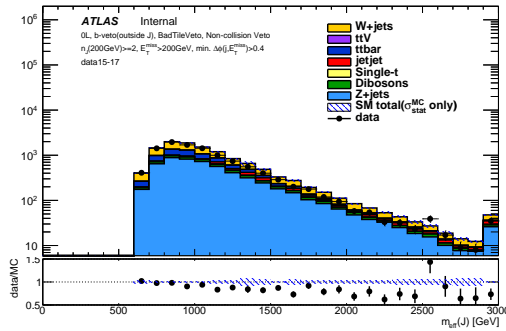
(d) sub-leading large- R jet p_T



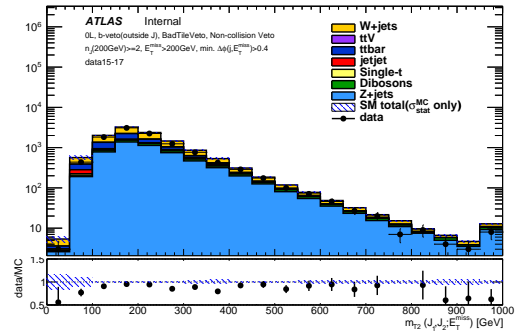
(e) $\min \Delta\phi(j, E_T^{\text{miss}})$



(f) E_T^{miss}



(g) $m_{\text{eff}}(J)$



(h) m_{T2}

Figure 7.8: Distributions of key observables for the boosted channel with the preselection applied as described in Table 7.7. Plot Credit: Boosted Channel Team.

7.4 Signal Region Selection

The resolved signal regions target a moderate $\Delta m(\tilde{\chi}_{\text{heavy}}, \tilde{\chi}_{\text{light}})$ whereas the boosted region targets a large $\Delta m(\tilde{\chi}_{\text{heavy}}, \tilde{\chi}_{\text{light}}) > 400$ GeV. There are two main categories of signal regions based on the two final states. The $qqqq$ or 4Q final state, and the $qqbb$ or 2B2Q final state. Inclusive signal regions are also defined which suppress model dependency and accommodate signatures from several models where the electroweakinos can decay into various boson modes. These can be further specified based on the exact bosons which are producing the final state jets.

Before diving in to the specifics of the signal-regions for this analysis, I first want to note that when defining and optimizing signal regions, the *significance* is a very important component. Significance is often used to understand an experimental result, and gives an idea of how incompatible new results are with the understood and existing body of data. In ATLAS analyses the significance \mathcal{Z} is thought of as the equivalent fluctuation of a standard Gaussian variable [98], which is similar to the Z-score: the number of standard deviations (σ) away from the mean. This is slightly different than many fields, which consider the p-value³ when deciding if a result is compatible with the current understanding. For reference, results with $Z = 3$ are considered to show evidence of new physics, but $Z = 5$ is required for a "discovery"⁴. Z can be defined in terms of the p-value as below:

$$\mathcal{Z} = \Phi^{-1}(1 - p) \quad (7.1)$$

where Φ is the cumulative distribution of the standard Gaussian with a mean of zero and a variance of one, and Φ^{-1} is the quantile of the standard Gaussian, which is the inverse of

3. $p = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \Phi(Z)$. The p-value represents the probability of obtaining a result equal to or more extreme than what was observed [98].

4. $Z = 3 \leftrightarrow p = 1.35 \times 10^{-3}$, $Z = 5 \leftrightarrow p = 2.87 \times 10^{-7}$

the cumulative.

In the limit $s \ll b$ with sufficiently large b , the significance becomes:

$$\mathcal{Z} = \frac{s}{\sqrt{b}} \quad (7.2)$$

where s is the number of signal events and b is the number of background events. However, in searches like the one described here, the s and b values come from finite data or Monte Carlo distributions and thus have an associated uncertainty. Additionally, in searches like this one the signal region is designed to have a very low background contribution so the $s \ll b$ condition is not necessarily met. A better approximation for the significance utilizes the beta function, which has a close relationship to the gamma function:

$$\mathcal{B}(z_1, z_2) = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1 + z_2)} \quad (7.3)$$

More precisely, it is the regularized incomplete beta function (defined in terms of the incomplete beta function and the complete beta function) which is used:

$$\mathcal{B}_{RI}(x; z_1, z_2) = \frac{B_I(x; z_1, z_2)}{B(z_1, z_2)} \quad (7.4)$$

The significance is defined using the regularized incomplete beta function as:

$$\mathcal{Z} = \Phi^{-1}\left(1 - \mathcal{B}_{RI}\left(\frac{1}{1 + \tau}; s + b, 1 + b\tau\right)\right) \quad (7.5)$$

where τ is defined as:

$$\tau = \frac{b}{\sigma_b^2} \quad (7.6)$$

In this definition, s is the number of signal events, b is the number of background events and σ_b^2 is the uncertainty on b .

In this analysis a specific implementation of \mathcal{Z} is used: `BinomialExpZ` implemented with ROOT [74]. For this section estimated sensitivities will be discussed, and these are generally computed using a flat 30% uncertainty as σ_b in the definition above.

7.4.1 Resolved Signal Regions

The resolved channel focused mainly on the 2B2Q final state during the signal-region definition and optimization procedure. This channel builds off the experience from the early Run 1 analysis approach and seeks to provide a complementary view of the signal events with respect to the boosted reconstruction approach.

In the 2B2Q region, four small-radius jets are required to be present, two of which are required to be b-tagged. Additional kinematic cuts are made to limit background contribution in the signal region. Different variables are particularly good at limiting the contribution of specific background processes. For example, m_{CT} and $m_T^{b,\min}$ are both ideal for removing large portions of the $t\bar{t}$ background. The number of leptons in the event and $m_{b\bar{b}}$ are particularly suited for reducing the contributions from single-top quark production and Z +jets production.

The previous Run 1 analysis regions were used as the starting point for the optimization procedure, as described in Table 7.8. These regions from the previous Run 1 analysis are used, but the plots shown below utilize the updated signal grid, updated samples and production, as well as updated code and uncertainties of the current analysis. Plots of the relevant kinematic variables with the signal region selections applied can be seen in Figure 7.9 and Figure 7.10 for the two signal regions. The estimated significance for these signal regions is shown in Figure 7.11.

	SRHad-High	SRHad-Low
N_{lep}	$= 0$	$= 0$
$N_{jet} (p_T > 30 \text{ GeV})$	$\in [4, 5]$	$\in [4, 5]$
N_{b-jet}	$= 2$	$= 2$
$\min \Delta\phi(E_T^{\text{miss}}, j)$	> 0.4	> 0.4
$E_T^{\text{miss}} [\text{GeV}]$	> 250	> 200
$m_{\text{eff}} [\text{GeV}]$	> 900	> 700
$m_{b\bar{b}} [\text{GeV}]$	$\in [105, 135]$	$\in [105, 135]$
$m_{q\bar{q}} [\text{GeV}]$	$\in [75, 90]$	$\in [75, 90]$
$m_{CT} [\text{GeV}]$	> 140	< 190
$m_T^{b,\min} [\text{GeV}]$	> 160	< 180

Table 7.8: Starting selection requirements for the signal regions of the resolved 2B2Q channel.

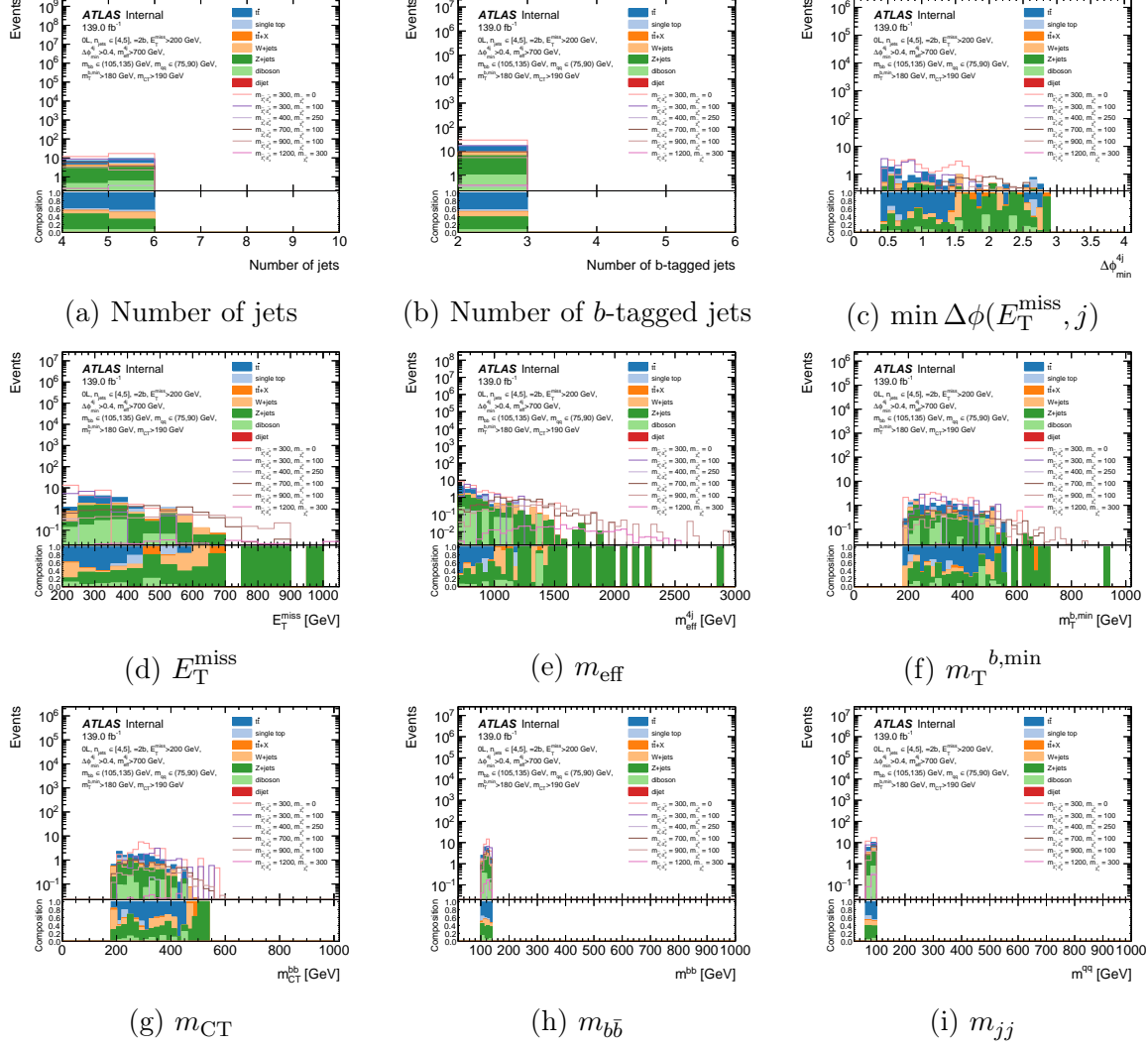


Figure 7.9: Distributions of key observables for the Resolved 2B2Q with the SRHad-Low signal region selection.

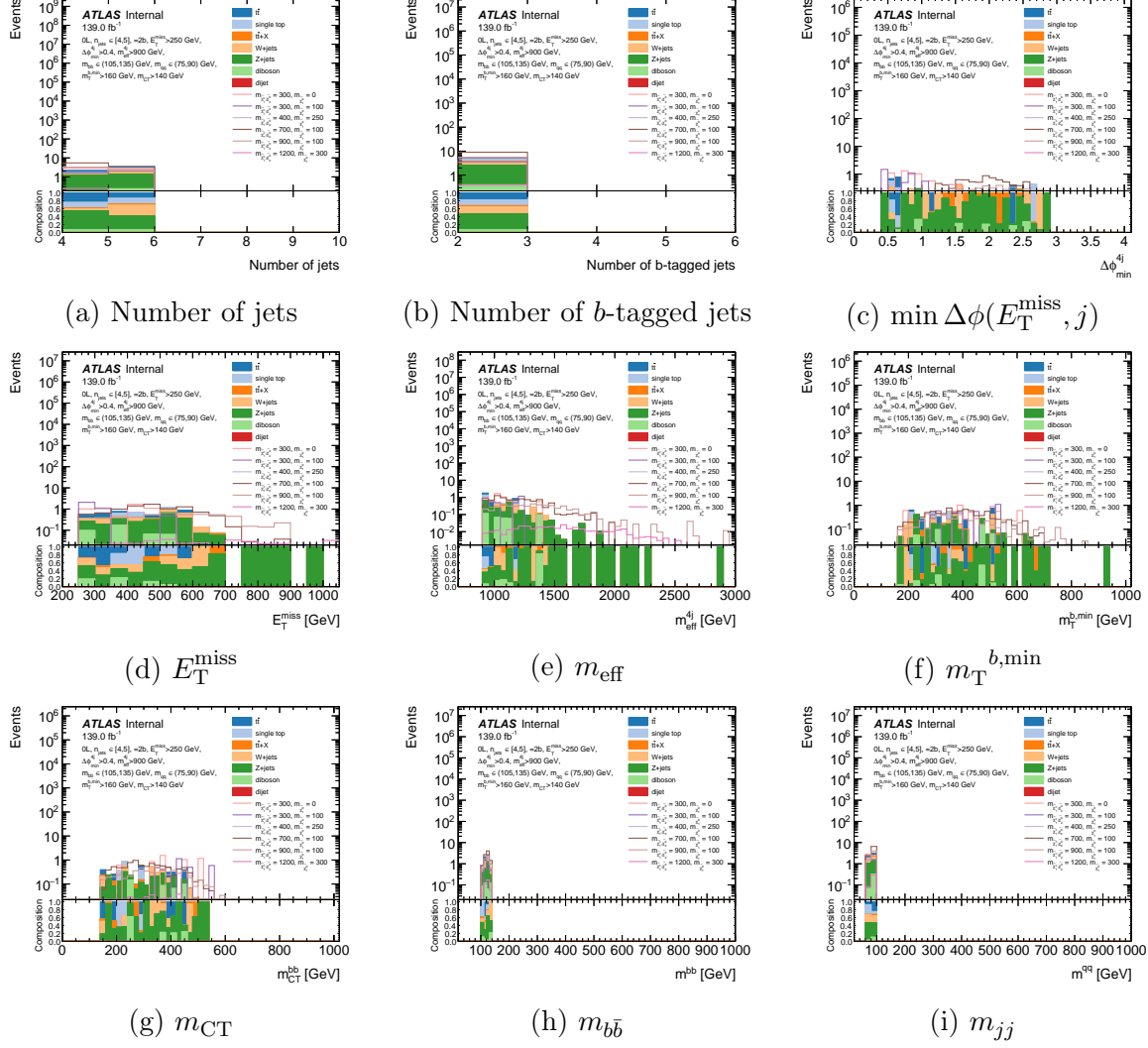
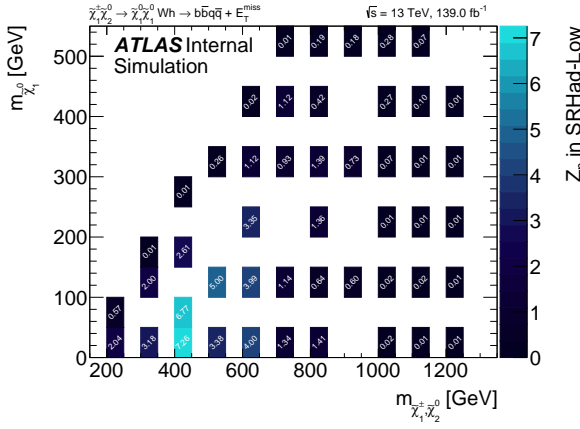
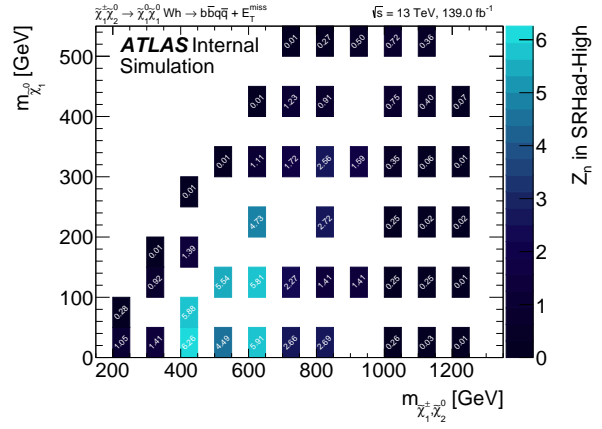


Figure 7.10: Distributions of key observables for the Resolved 2B2Q with the SRHad-High signal region selection.



(a) SRHad-LM signal region significance



(b) SRHad-HM signal region significance

Figure 7.11: Original SRHad-LM and SRHad-HM signal region estimated sensitivities.

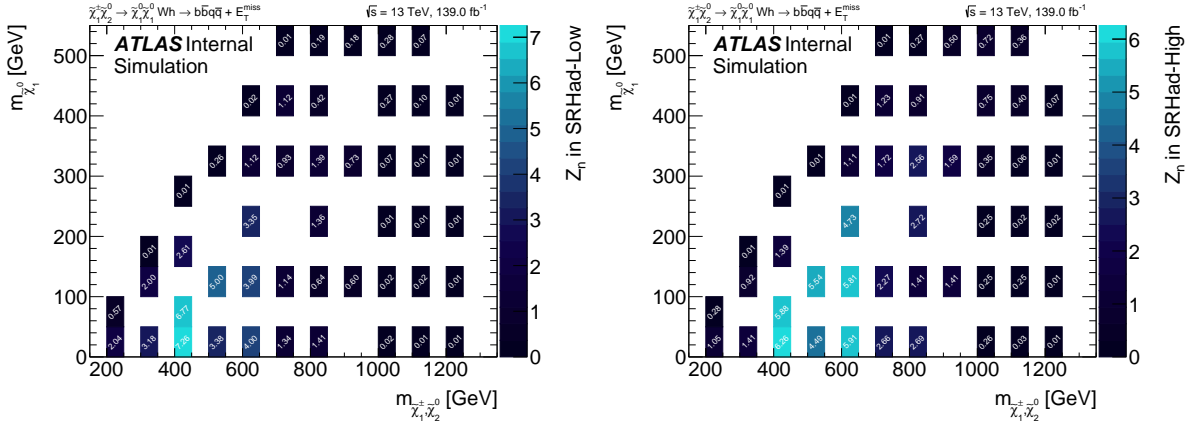
The optimization procedure was run several times, initially with two regions targeting high and low masses of the $\tilde{\chi}_1^\pm$, and $\tilde{\chi}_2^0$ electroweakinos. Over the course of this optimization the focus of the signal regions was expanded to target the diagonal area of the signal grid, as this is where the boosted channel is the least sensitive and where sensitivity is missing in other leptonic SUSY electroweakino search channels. This is a difficult region, due to the much smaller $\Delta m(\tilde{\chi}_{\text{heavy}}, \tilde{\chi}_{\text{light}})$ in many cases.

The resulting signal regions after the optimization procedure are shown in Table 7.9. The sensitivity for these optimized regions is shown in Figure 7.12.

	SRHad-Diag	SRHad-High	SRHad-Low
N_{lep}	= 0	= 0	= 0
$N_{jet} (p_T > 30 \text{ GeV})$	= 4	$\in [4, 5]$	$\in [4, 5]$
$N_{b\text{-jet}}$	= 2	= 2	= 2
$\min \Delta\phi(E_T^{\text{miss}}, j)$	> 0.4	> 0.4	> 0.4
E_T^{miss} [GeV]	> 200	> 250	> 200
m_{eff} [GeV]	> 800	> 900	> 700
$m_{b\bar{b}}$ [GeV]	$\in [105, 135]$	$\in [105, 135]$	$\in [105, 135]$
$m_{q\bar{q}}$ [GeV]	$\in [75, 90]$	$\in [75, 90]$	$\in [75, 90]$
m_{CT} [GeV]	> 260	> 140	< 190
$m_T^{b,\min}$ [GeV]	> 220	> 160	< 180

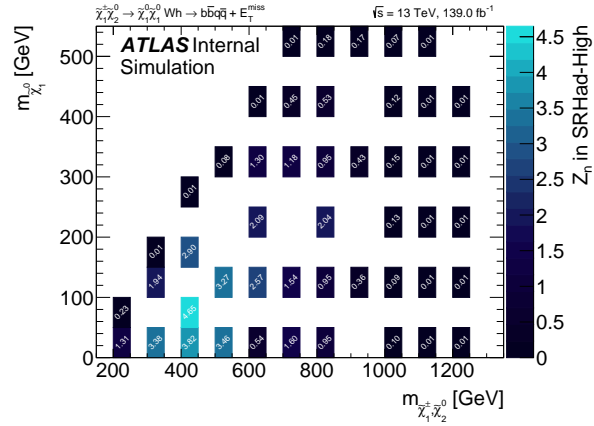
Table 7.9: Final selection requirements for the signal regions of the resolved 2B2Q channel.

The optimization procedure for the resolved signal regions was extensive, but ultimately it is very difficult to progress into the diagonal region of the signal grid.



(a) SRHad-LM signal region significance

(b) SRHad-HM signal region significance



(c) SRHad-Diag signal region significance

Figure 7.12: Final SRHad-Diag, SRHad-LM, and SRHad-HM signal region estimated sensitivities.

7.4.2 Boosted Signal Regions

The Boosted channel has ten signal regions in total, four of which come from the 4Q final state and six from the 2B2Q final state. Boson tagging is required for the two leading large-radius jets. They must pass W/Z-tagging in the 4Q signal regions, while the 2B2Q signal regions require that J_{bb} be tagged as a Z/Higgs boson and J_{qq} must be tagged as a W/Z-boson (ie satisfy the V_{qq} -tagging requirements). The selections for all ten signal regions are listed in Table 7.10, and illustrated schematically in Figure 7.15.

	$n(W_{qq})$	$n(Z_{qq})$	$n(V_{qq})$	$n(Z_{bb})$	$n(h_{bb})$
4Q-WW	= 2	-	= 2	= 0	= 0
4Q-WZ	≥ 1	≥ 1	= 2	= 0	= 0
4Q-ZZ	-	= 2	= 2	= 0	= 0
4Q-VV	-	-	= 2	= 0	= 0
2B2Q-WZ	= 1	-	= 1	= 1	= 0
2B2Q-ZZ	-	= 1	= 1	= 1	= 0
2B2Q-Wh	= 1	-	= 1	= 0	= 1
2B2Q-Zh	-	= 1	= 1	= 0	= 1
2B2Q-VZ	-	-	= 1	= 1	= 0
2B2Q-Vh	-	-	= 1	= 0	= 1

Table 7.10: Definition of each SR in the 4Q and 2B2Q categories, where $n(W_{qq})$, $n(Z_{qq})$, $n(V_{qq})$, $n(Z_{bb})$, and $n(h_{bb})$ are respectively the number of large- R jets passing the W_{qq^-} , Z_{qq^-} , V_{qq^-} , Z_{bb^-} , and h_{bb^-} -tagging of the two leading large- R jets. SR-4Q-WZ requires $n(W_{qq}), n(Z_{qq}) \geq 1$ instead of $n(W_{qq}) = n(Z_{qq}) = 1$ because the selections in W_{qq^-} and Z_{qq^-} -tagging are not exclusive.

Additional cuts are applied to further suppress the main SM backgrounds and are listed in Table 7.11 and Table 7.12. These cuts were selected after a full optimization procedure which considered several options for each observable. The boson-tagging requirement on the two large-radius jets plays the most important role in rejecting background from the signal region.

Using the 4Q-VV signal region as an example, plots of the variables and their corresponding cuts used for the inclusive 4Q-VV signal region can be seen in Figure 7.14. The expected

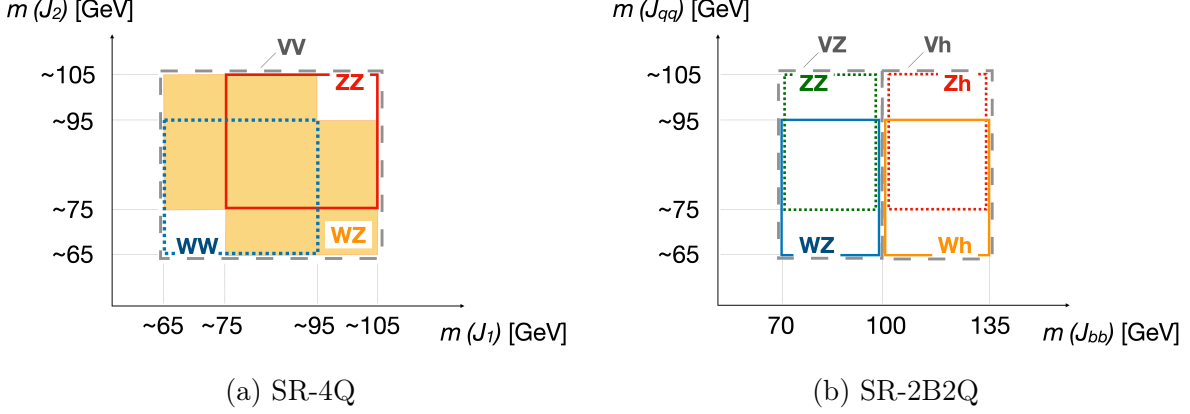


Figure 7.13: Signal region definitions as a function of the masses of the two leading large- R jets. The mass window cuts of the W_{qq}/Z_{qq} -tagging shown in the plot only indicate the typical values, while variable cut values along p_T are applied in the analysis. The inclusive SRs, defined by the logical union of a few mutually overlapping SRs, are indicated by the gray dashed lines. [90]

sensitivity of this same signal region can be seen using the signal grid in Figure 7.14.

	$\Delta\phi_{min}^{4j}$	m_{eff}	m_{T2}	E_T^{miss}	n_b^{trk}
4Q-WW	> 1.1	> 1100	—	> 200	< 2
4Q-WZ	> 1.1	> 1100	—	> 200	< 2
4Q-ZZ	> 0.9	> 800	> 250	> 300	< 2
4Q-VV	> 1.1	> 1100	—	> 200	< 2

Table 7.11: Definition of extra cuts as applied in the boosted 4Q SRs.

Similarly to the SR4Q signal regions the SR2B2Q regions are based on large-radius jet reconstruction, with kinematic cuts which are used for additional background suppression. As described in Table 7.10, two large-radius jets are required and one is required to have two matched b-jets (the decay from the $Z/h \rightarrow b\bar{b}$), and one can have at most one matched b-jets (the decay from the $W/Z \rightarrow qq$). The additional cuts used to suppress SM backgrounds were found using an optimization procedure and are shown for the SR2B2Q regions in Table 7.12.

Using the boosted SR2B2Q-Wh as an example, plots of the variables used for the kinematic cuts can be seen in Figure 7.16. The expected sensitivity of this same signal region can be

	$\Delta\phi_{min}^{4j}$	m_{eff}	m_{T2}	$n_J(n_b^{\text{trk}} \leq 1)$	$n_J(n_b^{\text{trk}} = 2)$	m_J
2B2Q-WZ	> 1.0	$= 1$	> 250	$= 1$	$= 1$	$\in [70, 100]$
2B2Q-ZZ	> 1.0	$= 1$	> 250	$= 1$	$= 1$	$\in [70, 100]$
2B2Q-Wh	> 1.0	$= 1$	> 250	$= 1$	$= 1$	$\in [100, 135]$
2B2Q-Zh	> 1.0	$= 1$	> 250	$= 1$	$= 1$	$\in [100, 135]$
2B2Q-VZ	> 1.0	$= 1$	> 250	$= 1$	$= 1$	$\in [70, 135]$
2B2Q-Vh	> 1.0	$= 1$	> 250	$= 1$	$= 1$	$\in [70, 135]$

Table 7.12: Definition of extra cuts as applied in the boosted 2B2Q SRs.

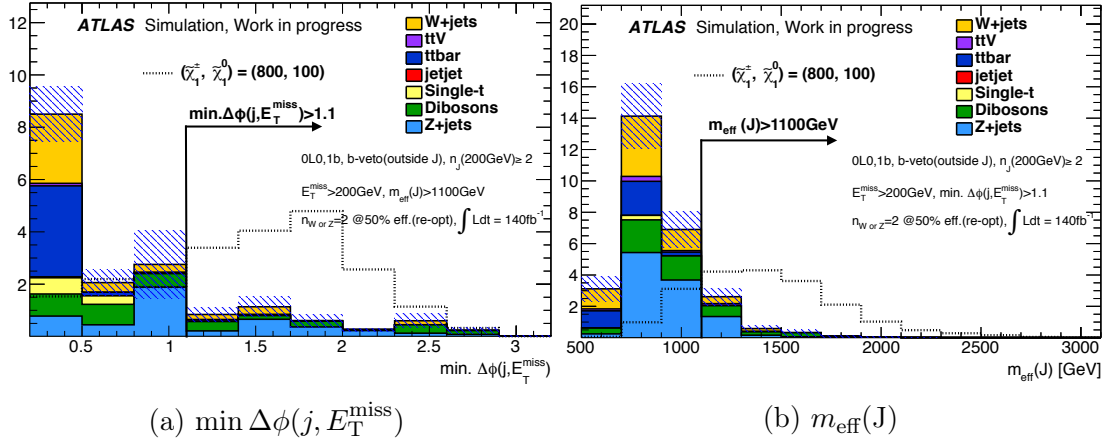


Figure 7.14: Kinematic plots for the SR4Q-VV signal region. The overlaid signal is $(\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm) \rightarrow WW + (\tilde{\chi}_1^\pm, \tilde{\chi}_2^0) \rightarrow WZ$ with $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_2^0}) = (800, 100)$ GeV. Plot credit: boosted channel team.

seen display using the signal grid in Figure 7.17.

The acceptance times efficiency for signal events ranges from 1% to 4% depending on $\Delta m(\tilde{\chi}_{\text{heavy}}, \tilde{\chi}_{\text{light}})$ and the SR. For example, it is about 1.5% in the SR-2B2Q-Vh region for the C1N2-Wh signals with $\Delta m(\tilde{\chi}_{\text{heavy}}, \tilde{\chi}_{\text{light}}) = 600$ GeV, and 2%–3% with $\Delta m(\tilde{\chi}_{\text{heavy}}, \tilde{\chi}_{\text{light}}) = 1$ TeV.

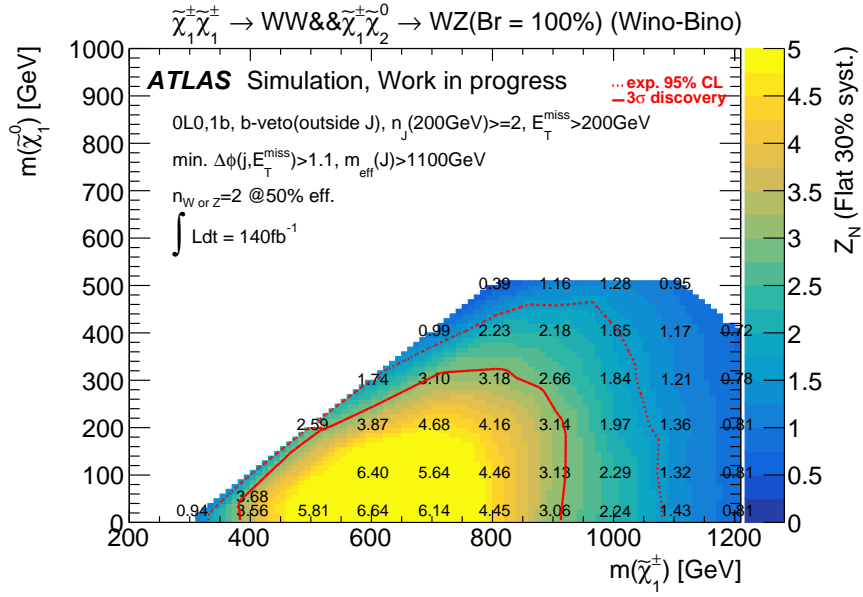
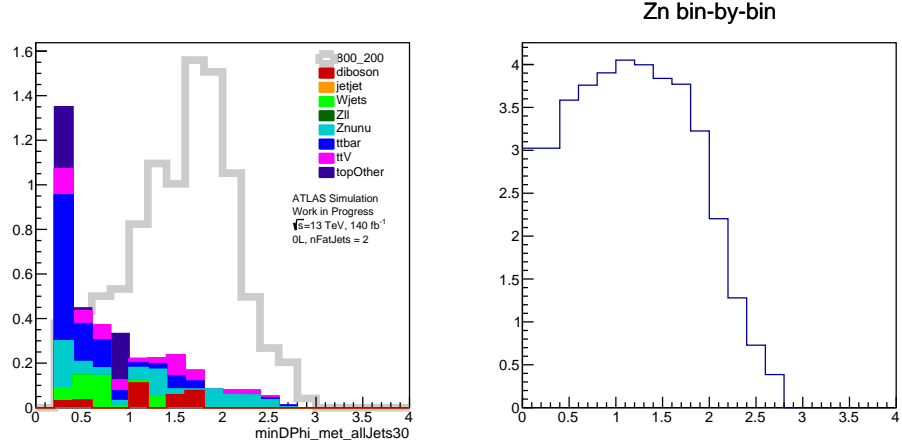
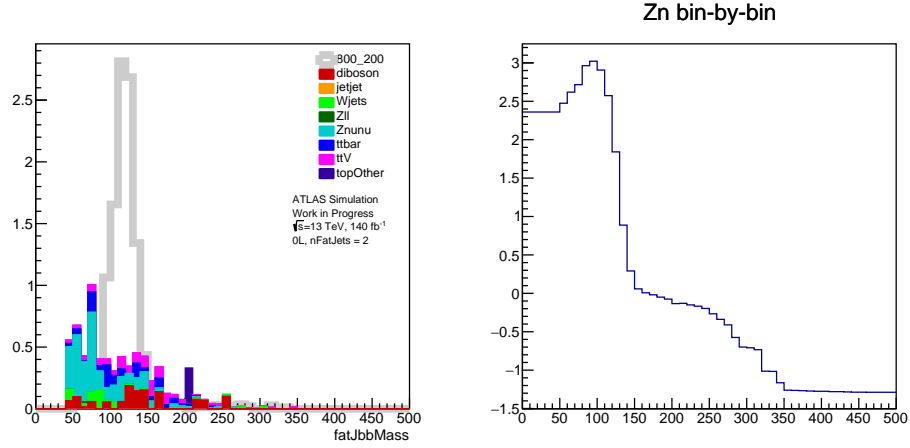


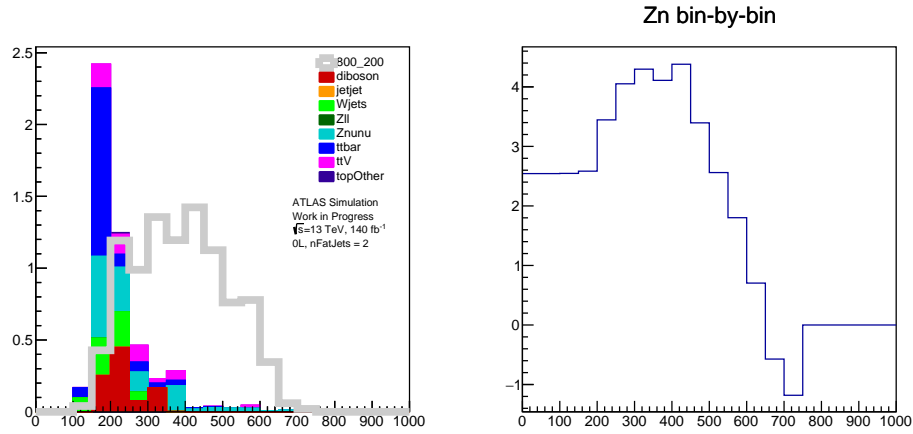
Figure 7.15: Expected sensitivity for the SR4Q-VV inclusive signal region with the $(\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm) \rightarrow WW + (\tilde{\chi}_1^\pm, \tilde{\chi}_2^0) \rightarrow WZ$ signal grid. Red dotted contours correspond to exclusion ($Z_N = 1.64$) and the solid contour correspond to 3σ sensitivity. Plot Credit: Boosted Channel Team.



(a) $\min \Delta\phi(j, E_T^{\text{miss}})$



(b) $2b\text{-Large-}R$ Jet Mass



(c) m_{T2}

Figure 7.16: Kinematic variables for the 2B2Q boosted Wh signal-region. Compared to signal shown in gray: $(\tilde{\chi}_1^\pm, \tilde{\chi}_2^0) \rightarrow Wh$ with $(m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_2^0}) = (800, 200)$ GeV.

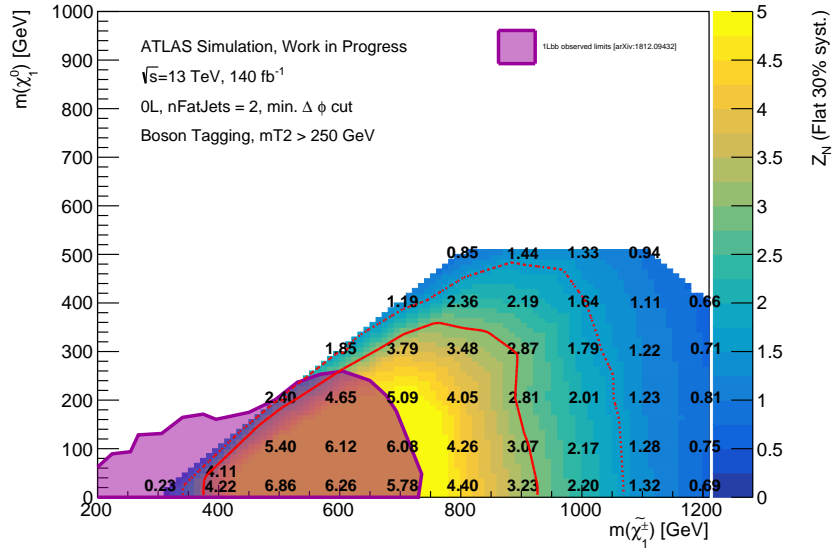


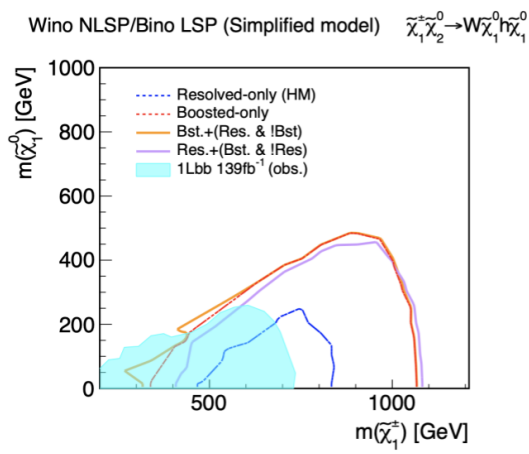
Figure 7.17: The expected sensitivity for Wh signal search. Signal and background yields are estimated from simulation. Points inside the solid red curve have at least 3σ sensitivity, and points inside the dashed red curve have at least $Z_N = 1.64$ (exclusion). Previously observed limits are shown in purple.

7.5 Resolved and Boosted Combined Sensitivity

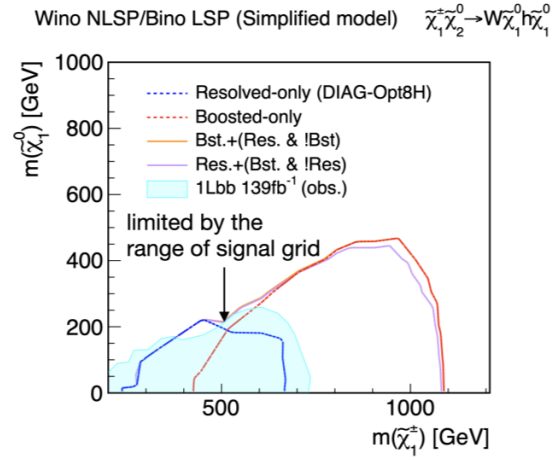
The boosted channel has increased sensitivity due to the boosted nature of the bosons and boosted boson tagging used to define the signal regions. Between this sensitivity of the boosted channel and the sensitivity of the existing one-lepton electroweakino search, the sensitivity added by the resolved channel signal regions is minimal, as shown in Figure 7.18a and Figure 7.18b. This can additionally be seen in the ratio of the combined sensitivity to the boosted channel sensitivity, shown in Figure 7.18c.

At this point in the analysis, the resolved channel was unable to add notable sensitivity after a significant optimization effort. It is possible that with additional time, person power, and new analysis techniques, such as the application of machine learning algorithms, the resolved channel sensitivity may have continued to improve. At the time of this analysis in 2020, the resolved channel only had one professor, one post-doctoral researcher, and one graduate student contributing significantly, all of whom had additional time constraints. It was ultimately decided that the boosted analysis would proceed with publication and that the resolved analysis would not.

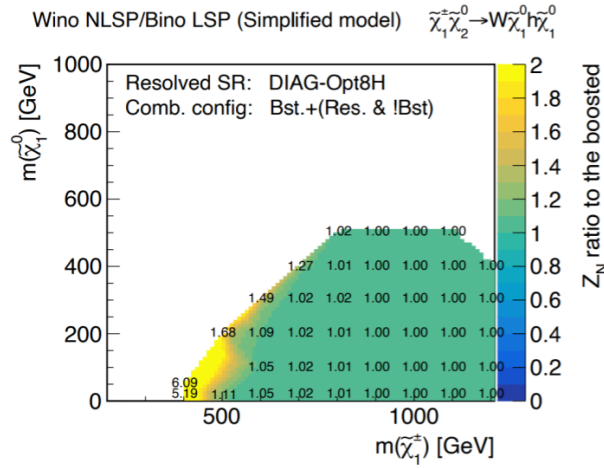
In the remaining sections of this chapter, I will summarize the remaining strategies and results of the boosted channel, though for a full overview I suggest reading the published work directly [90].



(a) Combined sensitivity with resolved SRHad-HM.



(b) Combined sensitivity with resolved SRHad-Diag.



(c) Ratio of combined boosted and resolved SRHad-Diag region sensitivity to boosted region sensitivity alone.

Figure 7.18: Estimated sensitivity of the boosted+resolved signal regions.

7.6 Boosted Channel Background Estimation

The main SM backgrounds in the boosted signal regions are $Z(\rightarrow \nu\nu) + \text{jets}$ ($\sim 50\%$), followed by $W(\rightarrow \ell\nu) + \text{jets}$ (15%–20%), and diboson production (10%–20%). The rest of the background consists of triboson events in SR-4Q (5%–10%), and $t\bar{t}$, single-top and $t\bar{t}+X$ events in SR-2B2Q (10%–20%). The irreducible backgrounds in this search are due to SM events including at least two hadronic $W/Z/h$ decays and large $E_{\text{T}}^{\text{miss}}$ from high- p_{T} neutrinos. These consist of triboson and $t\bar{t}+X$, and are estimated using MC simulation. Due to the high $E_{\text{T}}^{\text{miss}}$ requirement, events from fully-hadronic diboson production and $t\bar{t}$ are negligible.

The reducible backgrounds are characterized by the presence of at least one large-radius jet which comes from a process other than the targeted W/Z decay. A partial data-driven method is used to estimate these backgrounds through the use of a control region defined in the phase space adjacent to the signal regions. This control region contains mostly events from a specific background process. The MC sample for a given process targeted in a specific CR is normalized to the control region data, and the SR expectation is obtained by using the normalized MC sample. The normalized MC sample is tested against the data in a variety of validation regions (VR). The control and validation regions used are graphically represented in Figure 7.19. The validation regions in particular use data events with exactly one lepton or one photon, as in these events the main backgrounds have similar jet kinematics

A background-only fit is performed for each set of control regions. The results of these fits (including systematic uncertainties) are used to generate the background expectations in validation regions, and these are shown graphically in Figure 7.20. The data agree reasonably well with the estimated backgrounds throughout the validation regions.

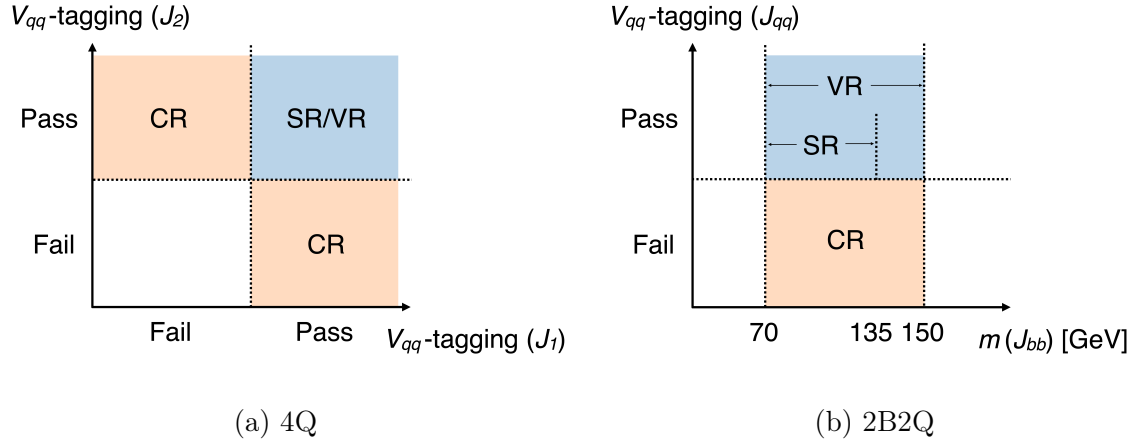


Figure 7.19: Schematics illustrating the relation between the signal regions, validation regions, and control regions [90].

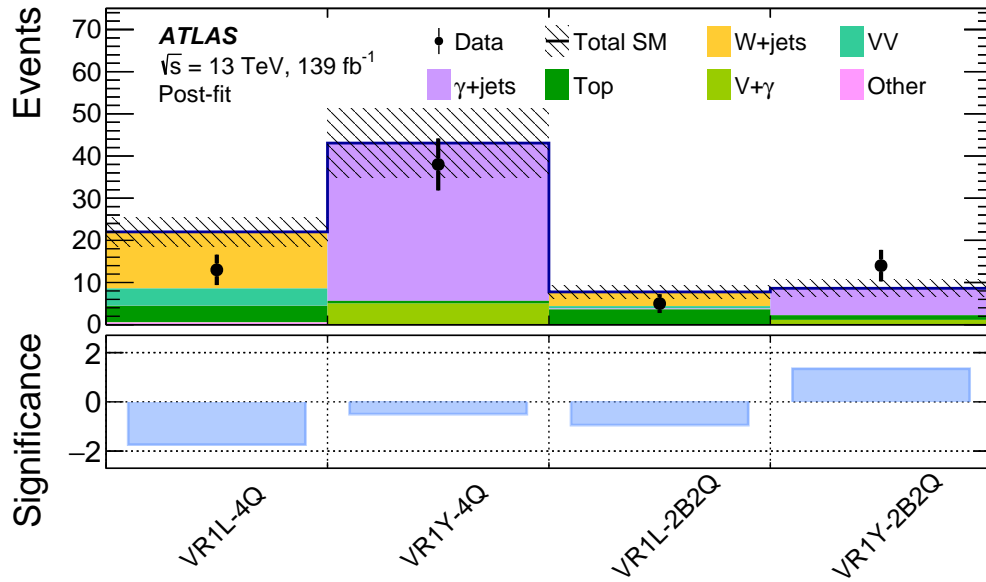


Figure 7.20: Comparison between the observed data and the post-fit SM background prediction in the validation regions. "Top" includes $t\bar{t}$, $t+X$ and $t\bar{t}+X$ production. The total systematic uncertainty of the background prediction is shown by the hatched band. The bottom panel shows the statistical significance of the discrepancy between the observed number of events and the SM expectation [90].

7.7 Boosted Channel Systematic Uncertainties

Relevant uncertainties include statistical uncertainties in the MC samples, experimental systematic uncertainties associated with the detector and the reconstruction process, and systematic uncertainties related to the theory in the modeling of the MC samples.

The largest systematic uncertainty is the statistical uncertainty from the MC generation, due mostly to the limited size of the Z+jets background sample. This is not the largest overall source of uncertainty in the analysis; this honor goes to the statistical uncertainty in the SRs due to the low number of data events present.

The experimental uncertainties include the uncertainty in the W_{qq}/Z_{qq} -tagging, originating ultimately from the MC modeling uncertainty present in the jet substructure variables [15]. There is also an uncertainty present in the Z_{bb}/h_{bb} -tagging, due to the b-tagging efficiency factors and the uncertainty on the shape of the mass distributions of the large radius jets [22]. Uncertainties from the energy/momentum determination of physics objects are also taken into account; these often come from simulation-based or in-situ calibrations and affect the calculation of E_T^{miss} in this analysis. An uncertainty in the integrated luminosity used to normalize the MC samples is also considered, as well as a pileup modeling uncertainty which is assigned to account for differences between predicted and measured inelastic cross-sections.

Additional experimental uncertainties from triggering, object identification and reconstruction, isolation requirements of electrons [18, 37], muons, and photons, and the jet vertex tagger were investigated and found to be negligible.

Theoretical uncertainties are estimated by either varying specific parameters of MC generators, or by utilizing a second MC generator as a comparison. In this analysis both methods are used.

Parameters which are varied include the QCD renormalization and factorization scales, which are varied up and down by a factor of 2 [33]. The CKKW merging scale is also shifted up and down from 20 GeV to 15 GeV and 30 GeV. The POWHEG+PYTHIA 8 $t\bar{t}$ sample is compared to the MADGRAPH5_AMC@NLO sample to estimate the hard-scatter modeling uncertainty and to the POWHEG+HERWIG sample [66] to show the uncertainty due to the choice of parton shower and hadronization models. Variations in $t\bar{t}$ initial and final state radiation modeling, renormalization and factorization scales are also taken into account. An additional uncertainty is taken into account for the potential mis-modeling of the relative background composition. Uncertainties on cross-sections are also included for some background processes and for the signal yields. The signal yields also have an uncertainty due to the shape which are mainly driven by the PDF uncertainty [75].

7.8 Boosted Channel Results

The final background estimates were calculated by performing a simultaneous profile log-likelihood fit [99] in all CRs and SRs relevant to a specific interpretation. The HISTFITTER framework [54] was used for this process. Systematic uncertainties are treated as Gaussian-distributed nuisance parameters and statistical MC uncertainties are treated as Poisson-distributed nuisance parameters.

A *background-only fit*, *discovery fit*, and *exclusion fit* were all performed. The *background only fit* considers only the control regions and assumes no contribution from signal. The *discovery fit* performs the hypothesis test for a generic beyond the Standard Model signal, and sets upper limits on the event numbers and cross sections of the signal. This type of fit uses one signal region and its associated control region. The *exclusion fit* sets the exclusion limit for a specific signal model. The signal regions and corresponding control regions are fit simultaneously.

For each discovery or exclusion fit, the compatibility of the data with different hypotheses (background only, or signal+background) are quantified by calculating a p-value with the profile likelihood ratio used as a test statistic [99]. Here, a 95% confidence level exclusion is defined by $CL < 0.05$ [172].

The observed data yields and expected SM backgrounds in each signal region and control region are shown in Figure 7.21. No significant excess is found in any signal-region.

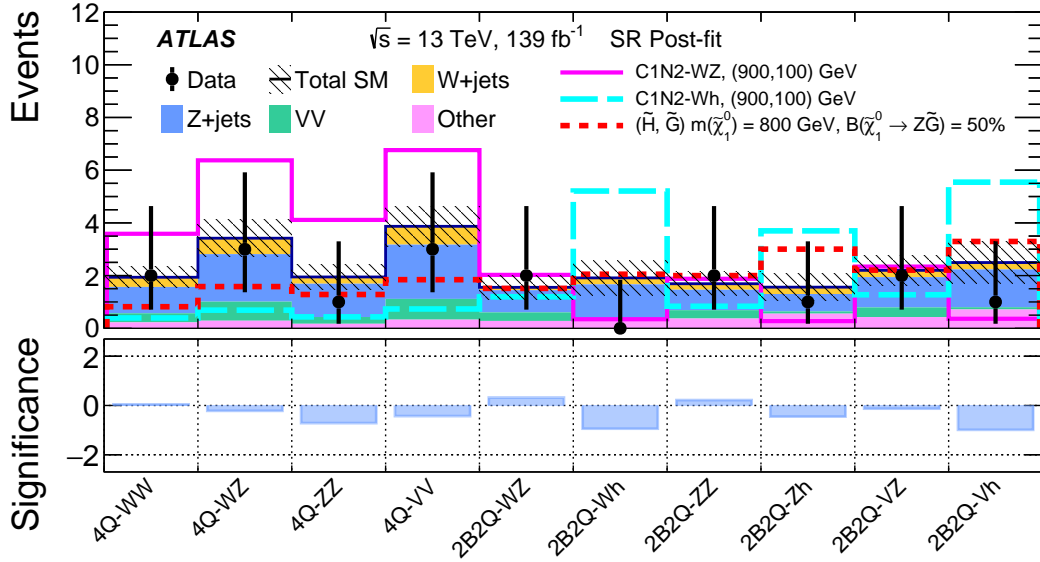


Figure 7.21: Summary of the observed data and predicted SM background in all SRs. The total systematic uncertainty in the background prediction is shown by the hatched area. Distributions of a few representative signals are overlaid, with $(m(\tilde{\chi}_1^\pm), m(\tilde{\chi}_1^0)) = (900, 100)$ GeV. The bottom panel shows the statistical significance of the discrepancy between the observed number of events and the SM expectation, following the prescription described in Ref. [97]. [90].

Exclusion fits were run for specific models. The (\tilde{W}, \tilde{B}) -SIM and (\tilde{H}, \tilde{B}) model exclusion limits are shown in Figure 7.22.

Additional models and interpretations are considered in Ref [90], but the (\tilde{W}, \tilde{B}) model result is the focus of this thesis.

In Figure 7.23 the exclusion limits on the $\tilde{\chi}_1^\pm$, $\tilde{\chi}_2^0$, and $\tilde{\chi}_1^0$ masses in the (\tilde{W}, \tilde{B}) model are

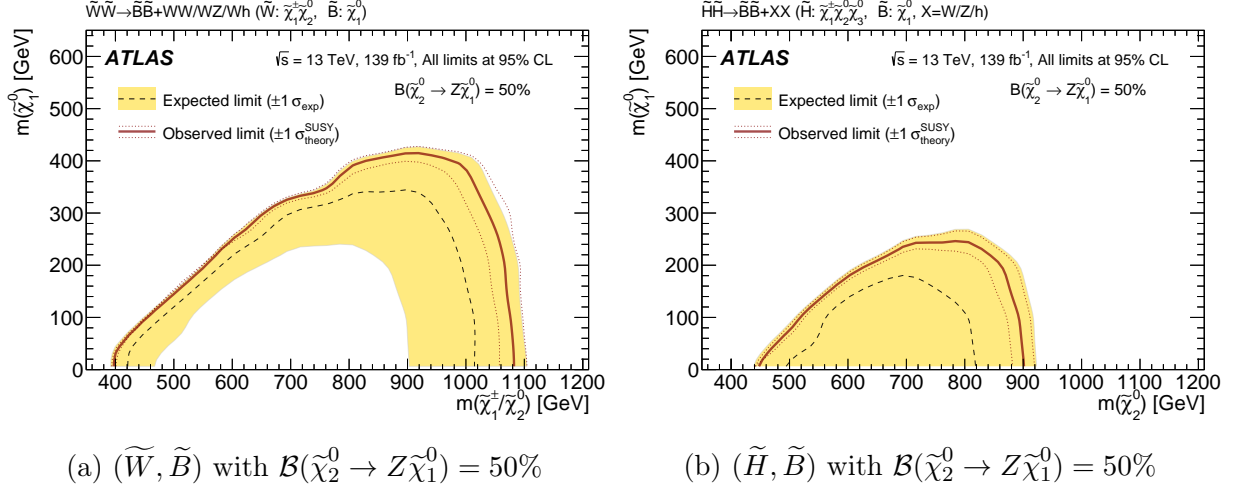


Figure 7.22: Exclusion limits for the (\tilde{W}, \tilde{B}) and (\tilde{H}, \tilde{B}) models shown as a function of the wino/higgsino chargino mass $m(\tilde{\chi}_1^\pm)$ and the bino LSP mass $m(\tilde{\chi}_1^0)$. Expected (dashed) and observed (solid red) 95% CL exclusion limits are shown for the (\tilde{W}, \tilde{B}) and (\tilde{H}, \tilde{B}) models with a representative branching ratio $\mathcal{B}(\tilde{\chi}_2^0 \rightarrow Z\tilde{\chi}_1^0) = 50\%$. [90]

compared to the same results from previous Run 1 analysis and additional channels in the Run 2 analysis, most of which are looking at final states requiring at least one lepton. The boosted channel result increases the limit placed on the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ masses by ~ 400 GeV compared to the fully hadronic Run 1 analysis, and by ~ 350 GeV compared to the closest limit, placed by the 1-lepton channel of the analysis.

This is a huge accomplishment, and really showcases the power of hadronic final states. The Run 1 fully hadronic electroweakino search was the first electroweakino search performed without leptons in the final state, and it had an equally large impact, expanding the limits by almost 40% compared to the next highest limit existing at the time. Hadronic final states are extremely powerful tools in physics searches, even in the electroweak SUSY sector where one might not immediately think of jets as the most relevant of objects. This analysis shows that jets are powerful tools in almost any analysis at the LHC, and the modeling uncertainties and SM hadronic backgrounds also emphasize the need for continued studies of jets themselves to increase our understanding of these objects, and our ability to model

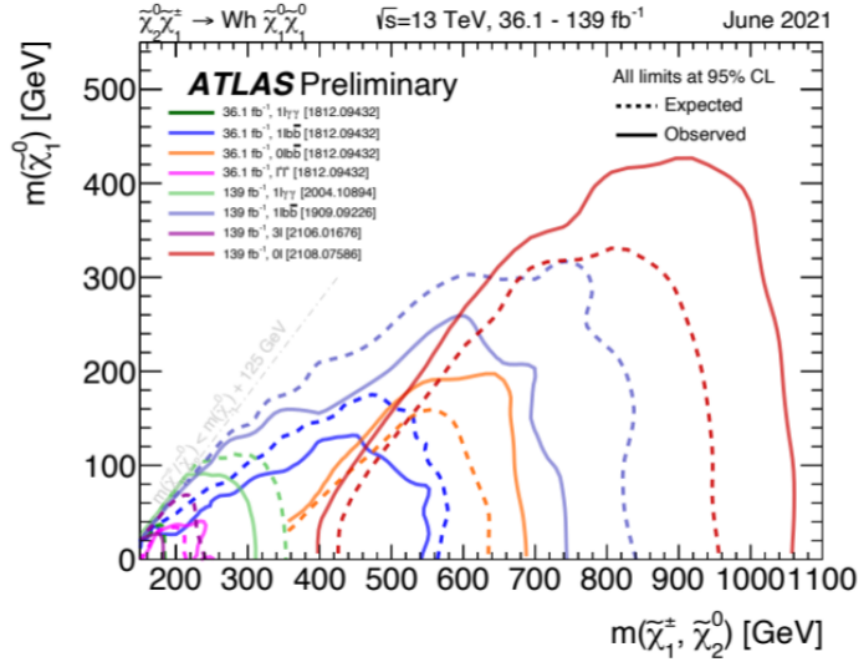


Figure 7.23: The boosted channel exclusion limit results compared to several other Run 1 and Run 2 searches for electroweakinos.

them well, as this modeling has profound impacts in all ATLAS and LHC analyses.

CHAPTER 8

PRECISION MEASUREMENT OF LUND AND PRIMARY LUND MULTIPLICITIES IN JETS

This chapter presents a precision measurement on two specific jet substructure observables: the primary and Lund subjet multiplicities. These two types of multiplicities count the emissions within a jet according to different parameters. This is the kind of measurement which allows for precise tests of QCD, the theory of the strong interaction.

This analysis was performed with the full Run 2 dataset, collected with the ATLAS detector from 2015 - 2018 at the LHC, which contains 149 fb^{-1} of data. As of the 8 November 2023 the analysis has received group-level approval from the Standard Model group, is in the process of internal collaboration review with ATLAS. Full publication is planned for Winter 2023.

My work on this analysis is organized into several sections below. In Section 8.1 the underlying motivations of this analysis are discussed and Section 8.2 defines the measured observables in detail. In Section 8.3 and Section 8.4 the datasets and Monte Carlo samples used are described. Additional information on the event selections and binnings of the measurement are given in Section 8.5. Section 8.6 describes the unfolding procedure used to account for detector effects on the measured distributions, and Section 8.7 discusses the relevant uncertainties on the measurement. Finally, Section 8.8 describes the results of the measurement and the comparisons to new predictions.

8.1 Motivation

At hadron colliders like the LHC, high energy quarks and gluons fragment and hadronize producing collimated *jets* of particles as discussed in detail in Chapter 6. Jets are complex objects both experimentally and theoretically. The energy scales of their constituent particles can span many orders of magnitude, and the complexities of QCD theory often require additional techniques for calculating substructure observables such as resummation. This makes jets a very valuable tool for precise tests of QCD, the theory of the strong interaction, in the collinear limit¹ [153].

Hadronic radiation within collider events is modelled using Parton Shower Monte Carlos (PSMCs) which are widely used in both theoretical and experimental particle physics. PSMCs are extremely important tools, and their accuracy directly affects the precision of data analysis at the LHC in a number of ways. The most direct effect is the uncertainties on the models themselves which are often limited by *ad hoc* comparisons between discrepant algorithms [21].

There has been recent exciting progress towards improved understanding of higher-order QCD effects in the collinear limit, which is partially motivated by the ongoing development of PSMC algorithms with higher logarithmic accuracy than existing options [71, 101, 116, 162, 163, 70, 59, 60, 130, 58]. One of these higher order effects is the incorporation of so called 'double-soft' splittings in QCD where pairs of soft particles are emitted with commensurate energy and angle to their emitter. Such terms have been occasionally incorporated at lower order into several existing PSMC algorithms [137, 129, 135, 136, 147, 133, 131, 105, 78, 121], but are now a focus for PSMCs with higher logarithmic accuracy. The primary and Lund subjet multiplicities are jet substructure observables which are directly sensitive

1. "in the collinear limit" = inside of a jet

to these contributions. Resummed analytical predictions of these multiplicities have been newly performed in Refs. [154, 155] at next-to-next-to-double-logarithmic (NNDL) accuracy. These predictions will be subsequently used to assess the complete implementation of these higher-order contributions within one new PSMC being developed by the PANGLOBAL collaboration² [101].

As theoretical understanding rapidly advances, it is important to promptly verify these new predictions with collider data that exposes such higher-order effects as those discussed above. The Lund multiplicities are closely related to the Cambridge multiplicity, which has an alternative definition that is suitable for e^+e^- annihilation, and was measured at LEP by the OPAL Collaboration [169]. In this analysis, the first differential cross-section measurement of the Lund subjet multiplicities at a hadron collider is performed, and the first comparisons are performed between these measured data to predictions for these observables from both the the new ALARIC parton shower, produced by the authors of the SHERPA MC. These results may also be compared in the future (though not in this thesis) to additional new analytical predictions from the PANGLOBAL collaboration [155].

8.2 Observable Definition

The measured multiplicity observables are described here, but a more thorough overview of the phenomenological interest of these observables can be found in Refs. [154, 101, 155], which provide a helpful overview, as well as the full NNDL analytical calculation of the Lund multiplicity.

The subjet multiplicities discussed here are part of a class of Lund jet plane (LJP) related observables [104] which have become popular in the last several year due to their ability to

2. PANGLOBAL is a collaboration of QCD theorists with the goal of improving PSMCs [website].

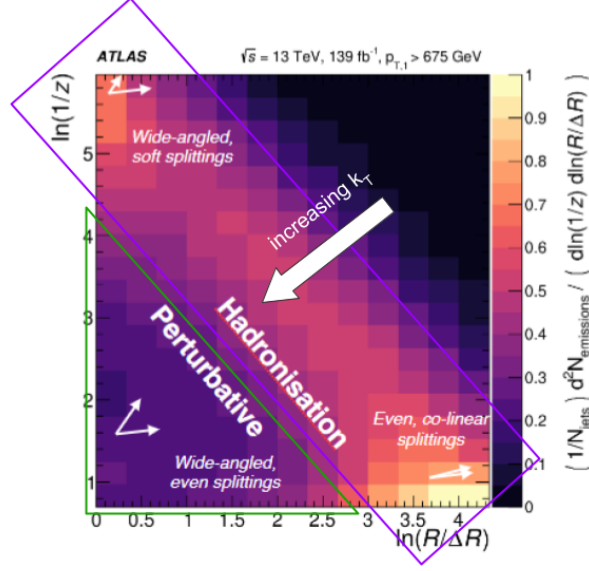


Figure 8.1: ATLAS Lund Jet plane result, with areas of perturbative and hadronization effects labeled. The direction of increasing k_t is also labeled as this illustrates the choice of multiplicity observables.

examine the underlying structure of a jet. The Lund jet plane probes the entire emission history of a parton and parameterizes these emissions in terms of their relative energies³ and angles (ΔR). This produces a picture like the one shown in Figure 8.1, which is the measurement of the Lund Jet Plane performed by ATLAS. As can be seen in Figure 8.1, specific regions of the LJP are sensitive to perturbative effects and effects due to hadronization (which are non-perturbative).

The primary and Lund multiplicities are computed by counting the number of emissions above a specified relative transverse momentum requirement in a jet's angle-ordered clustering history [154], building on the Lund jet plane representation of jet substructure (JSS) [104]. Such a clustering history is obtained by reclustering an anti- k_t jet's constituents with the Cambridge-Aachen (C/A) algorithm [199, 196]. The relative transverse momentum of an

3. often k_t , defined in Equation 8.1, or the relative momentum fraction, z , is used for this

emission, k_t , is defined to be

$$k_t = p_T^{\text{emission}} \Delta R(\text{emission, core}) \quad (8.1)$$

where the ‘emission’ and ‘core’ of the clustering sequence are respectively defined to be the softer and harder branches at a given node of the angular-ordered tree.

The angle-ordered picture is obtained experimentally by reclustering the inner-detector tracks inside of anti- k_t jets with the Cambridge/Aachen algorithm [199, 196], following established methodology from earlier measurements of jet substructure observables by ATLAS [30, 27] and CMS [83] during Run 2⁴

Emissions may be counted either solely in the primary Lund jet plane or in the full, fractal Lund jet plane, denoted $N_{\text{Lund}}^{\text{Primary}}$ and N_{Lund} , respectively. These observables are illustrated schematically in terms of the Lund jet plane in Figure 8.2.

By varying the k_t cut that is applied during this counting procedure, these observables can be made more- or less-sensitive to perturbative and non-perturbative radiation within a jet, by preferentially removing the region of the Lund jet plane which collects non-perturbative emissions produced via hadronisation. This is nicely illustrated by the dashed horizontal line at the bottom of Figure 8.2 representing the cut on k_t that is applied when calculating the multiplicities.

In order to resolve the most collinear splittings within jets, this measurement reconstructs the Lund jet plane using charged tracks within jets, following the approach from Ref. [29]. As this approach omits neutral information within the jet, the scale of the reconstructed emission k_t must be restored from the charged-only scale to the all-particles scale in order

4. In particular, this reclustering is performed on all tracks associated with the jet. This association is performed with ΔR based matching dependent on the size of the jet ($R=0.4$ in this analysis).

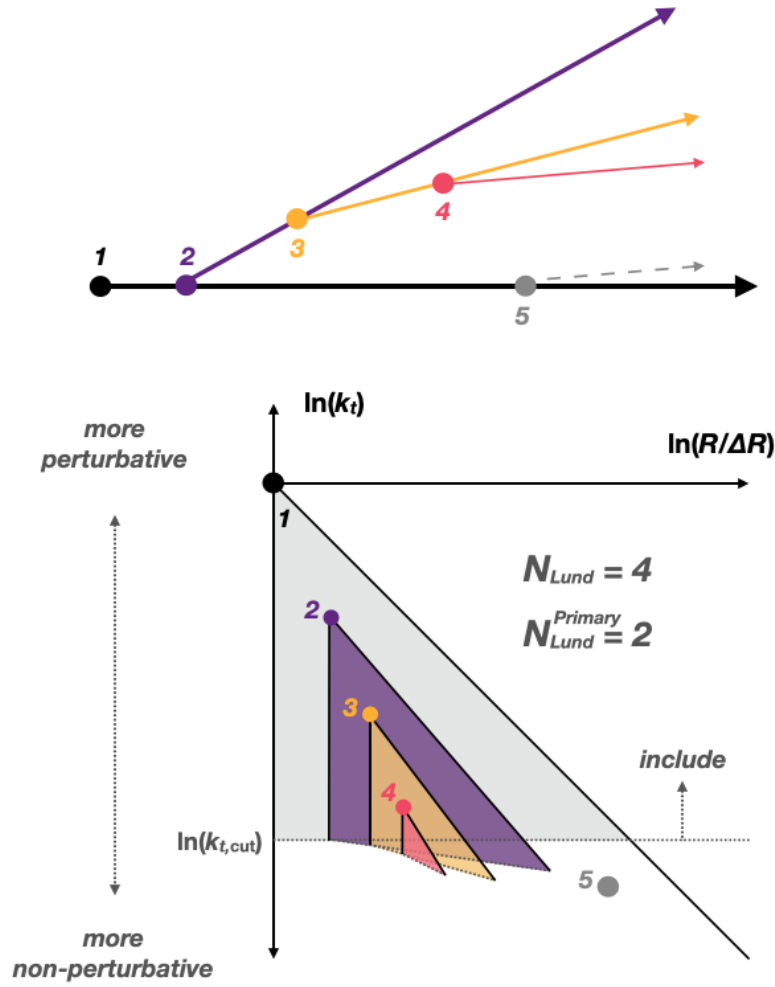


Figure 8.2: Depiction of Lund subjet multiplicities, reproduced and modified from Ref. [154]: a schematic Lund jet plane is provided for a jet with one primary and two secondary emissions. The Lund multiplicity of this jet is $N_{\text{Lund}} = 4$, while the Primary Lund multiplicity is $N_{\text{Lund}}^{\text{Primary}} = 2$. The lower edge of the Lund jet plane is truncated, indicating the region of emission k_t that is removed when imposing a k_t requirement.

for comparisons between the measured distributions and theoretical predictions to be made. The k_t of each emission is therefore rescaled by the ratio of the transverse momentum of the reconstructed jet (all-particle) to the transverse momentum of the matched track jet (charged-particle), as shown in Equation 8.2:

$$k_t = \frac{p_T^{\text{all}}}{p_T^{\text{charged}}} k_t^{\text{charged}}. \quad (8.2)$$

This restores the average scale to the all-particles level. This change is particularly helpful in the case of a comparison to an analytic prediction, which may be done in the future for this measurement.

The N_{Lund} and $N_{\text{Lund}}^{\text{Primary}}$ distributions are measured with 16 different choices of the required minimum k_t of each emission: $k_t > 0.5$ GeV, 1 GeV, 2 GeV, 5 GeV, 10 GeV, 20 GeV, 50 GeV, 100 GeV, resulting in a total of 16 different triple-differential measurements. These 16 observables are measured differentially in 5 bins of jet p_T beginning at 300 GeV, and differentially in the relative η of the jet. The average multiplicity is extracted from these differential distributions and provided as a function of the k_t requirement, which can then be compared to predictions from PSMC models or to calculated predictions.

8.3 Data

The dataset used in this analysis is from pp collisions provided by the LHC with $\sqrt{s} = 13$ TeV between 2015 - 2018 and was collected by the ATLAS detector, which is discussed in detail in Chapter 4. The total integrated luminosity after applying specific criteria ensuring the data can be used for physics analysis is 140 fb^{-1} . The uncertainty in the combined 2015–2018 integrated luminosity is 0.83% [26], obtained using the LUCID-2 detector [51] for the

primary luminosity measurements. Due to the high instantaneous luminosity and the large total inelastic pp cross section, there are, on average, 33.7 simultaneous (‘pileup’) collisions in each bunch crossing. Data which are included in this dataset but identified as contaminated with errors from the liquid argon calorimeter, tile calorimeter or semiconductor tracker (SCT) are rejected using event-level flags during analysis [13].

Data are collected using a set of single-jet triggers with varying thresholds. The unscaled triggers used to collect data at high- p_T vary depending on the data-taking period during Run 2. A specific single-jet trigger is used to fill each p_T bin used in this analysis; the appropriate trigger must have fired in order for an event in data to enter the measurement. These triggers are summarised in in Table 8.1.

p_T bin	Trigger	Year	Prescaled?
240 – 300 GeV	HLT_j175	All	Yes
300 – 500 GeV	HLT_j260	All	Yes
> 500 GeV	HLT_j360	2015	No
> 500 GeV	HLT_j380	2016	No
> 500 GeV	HLT_j420	2017	No
> 500 GeV	HLT_j420	2018	No

Table 8.1: Summary of the single-jet triggers used during each year of Run 2 data-taking. The indicated trigger is used to select events in the corresponding year for use in this analysis.

8.4 Simulation

Samples of Monte Carlo (MC) simulated dijet and multijet events are used in this analysis. As the jet cross-section is vastly larger than that of electroweak processes, the consideration of only these signal samples is sufficient to describe the data – all other processes are neglected. In order to fully populate a wide range of jet p_T , these samples are generated in slices of the particle-level $R = 0.4$ jet p_T [151, 152]. These samples also include the full ATLAS detector simulation and reconstruction for the reco level information.

PYTHIA 8.230 [180, 179] is used as the nominal MC generator for this analysis. Samples of $2 \rightarrow 2$ dijet events are simulated using the A14 tune [14], Lund string hadronisation model and NNPDF 2.3 [165] parton distribution function (PDF) set. The PYTHIA parton shower algorithm uses a dipole-style p_T -ordered evolution, and its renormalisation and factorisation scales are set to the geometric mean of the squared transverse masses of the outgoing particles. EvtGen [140] is used to model decays of heavy flavor hadrons.

Two sets of SHERPA 2.2.5 [122] dijet events are used with the default AHADIC⁵ cluster hadronization model [192] or using the SHERPA interface to the Lund string hadronisation model as implemented in PYTHIA 6.4, and its decay tables. These samples include leading order matrix element calculations for $2 \rightarrow 2$ processes, and use the CSS SHERPA parton shower algorithm [178]. The CT14NNLO PDF [106] set is used for matrix element calculations and CT10 is used for multi-parton interactions [139].

Two additional SHERPA samples are generated with SHERPA 2.2.11. One is produced using with the same settings as the sample with cluster-based hadronisation described above, except that the hadronisation model’s parameters have been re-tuned to improve the level of agreement with LEP data. This improvement particularly improves the description of the baryon production rate inside of jets, and was found to make the description of the ATLAS jet energy response more consistent between SHERPA and PYTHIA jets.⁶ The other is produced using the new DIRE parton shower [132] with dipole resummation, and the AHADIC cluster hadronisation model, and has not been re-tuned.

Two sets of HERWIG 7.1.3 [55, 66, 67] multijet events are generated with the MMHT2014NLO PDF set [128], cluster hadronisation model and either the default angle-ordered PS or alternative dipole PS [192]. These samples model $2 \rightarrow 2$ matrix elements with NLO accuracy

5. A HADronisation model In C++.

6. Further details can be found in Ref. [16].

and $2 \rightarrow 3$ matrix elements with LO accuracy. Both parton shower models are matched to the matrix element calculation using the MC@NLO matching scheme. The p_T of the leading jet is taken as the renormalisation scale.

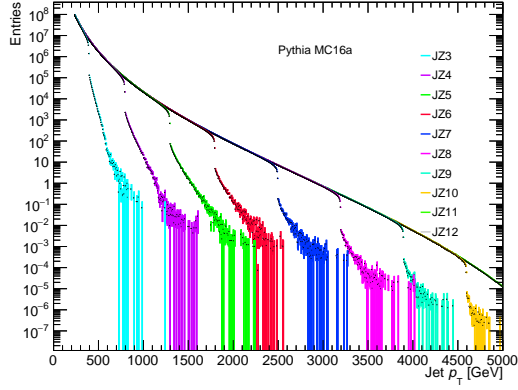
A sample of dijet events with NLO matrix element accuracy is produced with POWHEG V2 [164, 120, 6] and matched to PYTHIA 8 with the dijet process implemented in POWHEG BOX V2 [5]. The renormalisation and factorisation scales in this sample are set to the p_T of the underlying Born-level configuration. The default Lund string hadronisation model is used with the NNPDF30NLO PDF set. The parton shower and multi-parton interactions are modelled using PYTHIA 8.230, configured as for the nominal MC sample described above.

All simulation is reconstructed using the full ATLAS detector simulation. Additional superimposed simulated minimum-bias interactions are generated using PYTHIA 8 with the A3 set of tuned parameters [44] and NNPDF2.3 LO PDF set [165] to represent multiple pp interactions during the same or nearby bunch crossings (pileup). The distribution of the average number of pile-up interactions in simulation is re-weighted during data analysis to match that observed in the Run 2 data. This procedure follows the standard ATLAS pileup reweighting guidelines⁷, and in practice reweights the *average* μ distribution in mc16a samples, and the *actual* μ distribution in mc16d and mc16e samples.

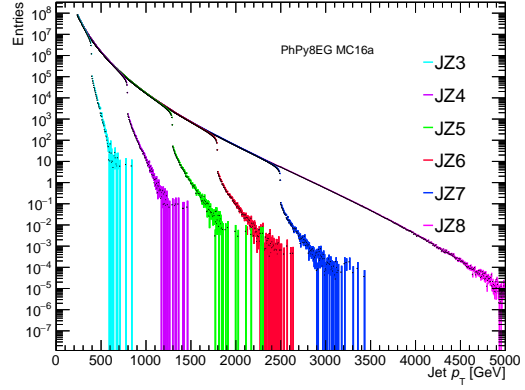
The MC samples used in this analysis are summarized in Table 8.2, and additional details on these samples may be found in Ref. [34].

Validation that the weights related to multijet sample slicing and that the slicing procedure is correctly implemented in this analysis [151, 152] and produces a smoothly falling underlying jet p_T spectrum can be seen in Figure 8.3 - Figure 8.6

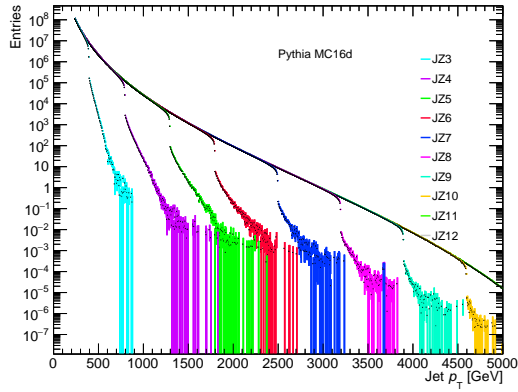
7. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ExtendedPileupReweighting>



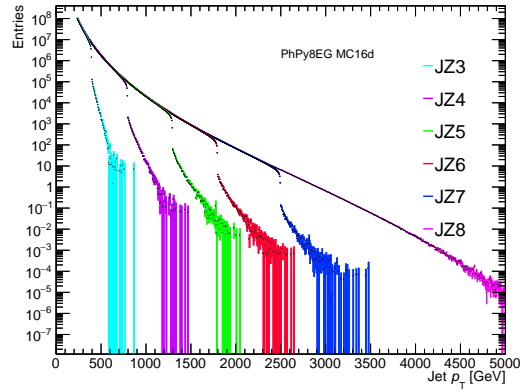
(a) MC16a PYTHIA



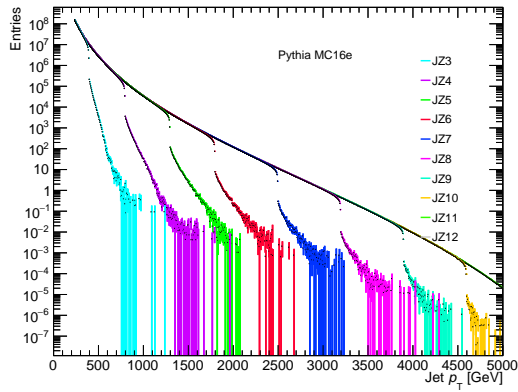
(b) MC16a POWHEG +PYTHIA



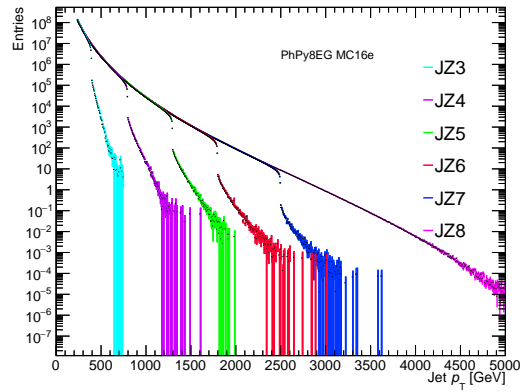
(c) MC16d PYTHIA



(d) MC16d POWHEG +PYTHIA

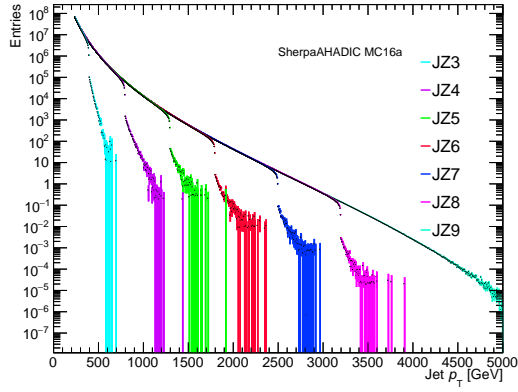


(e) MC16e PYTHIA

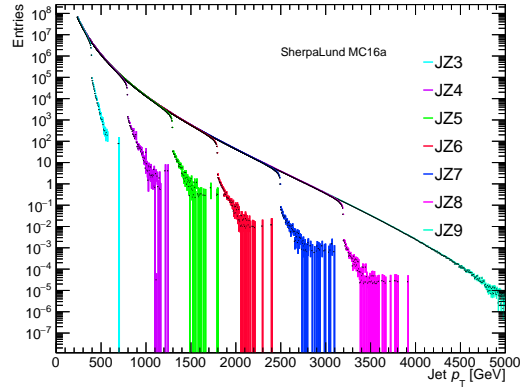


(f) MC16e POWHEG +PYTHIA

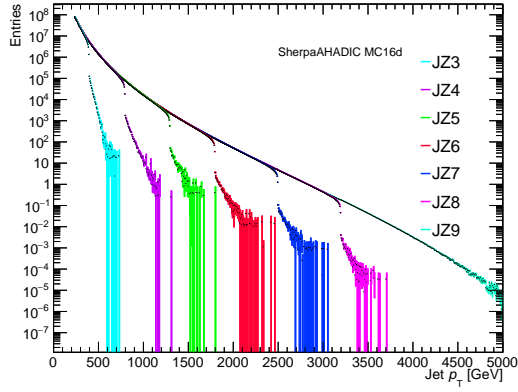
Figure 8.3: Validation that p_T spectrum is smooth in PYTHIA 8 and POWHEG +PYTHIA 8 samples.



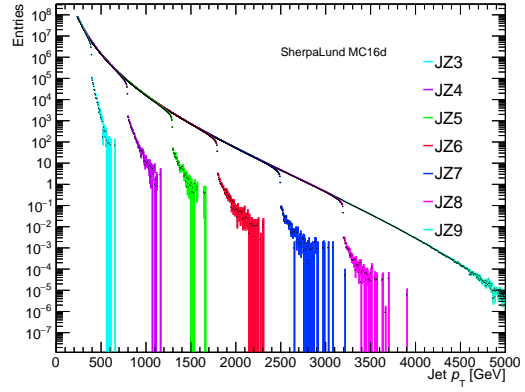
(a) MC16a AHADIC



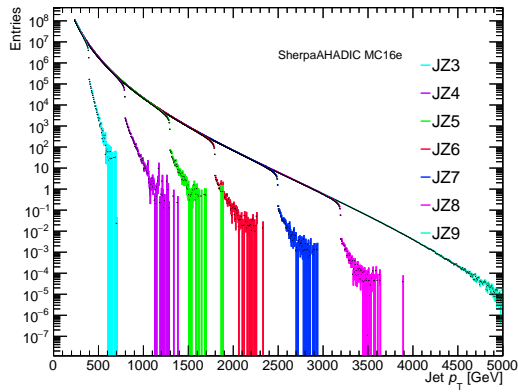
(b) MC16a Lund



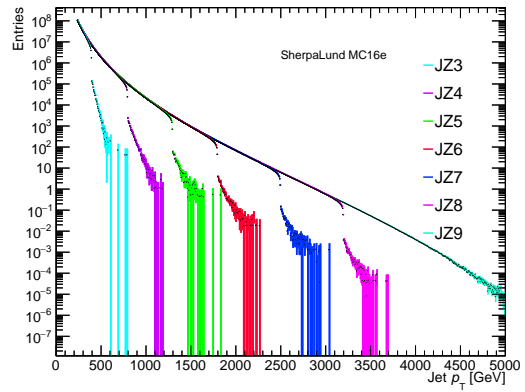
(c) MC16d AHADIC



(d) MC16d Lund

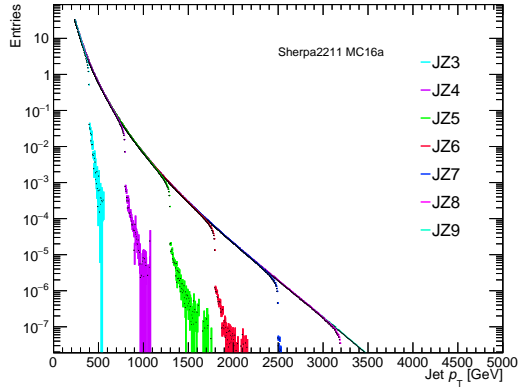


(e) MC16e AHADIC

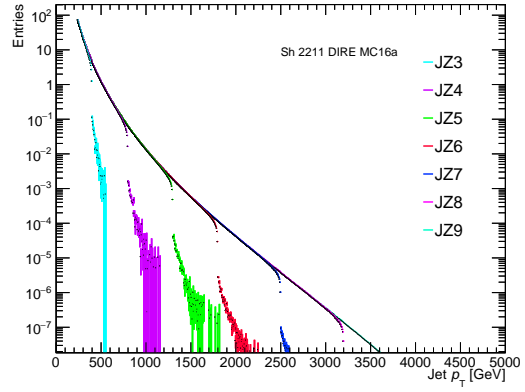


(f) MC16e Lund

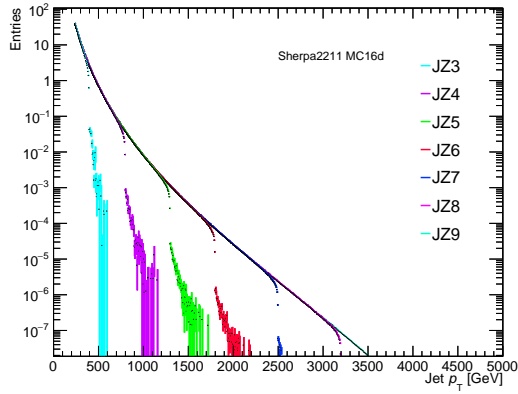
Figure 8.4: Validation that p_T spectrum is smooth in SHERPA samples.



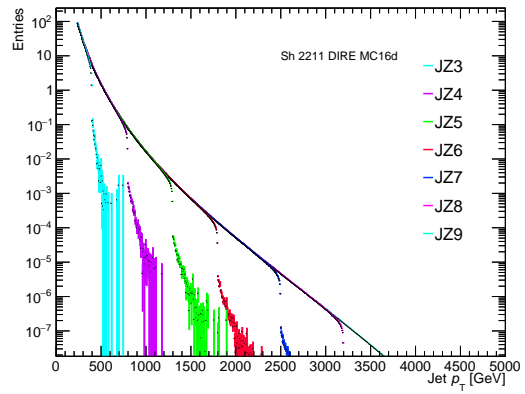
(a) MC16a 2.2.11



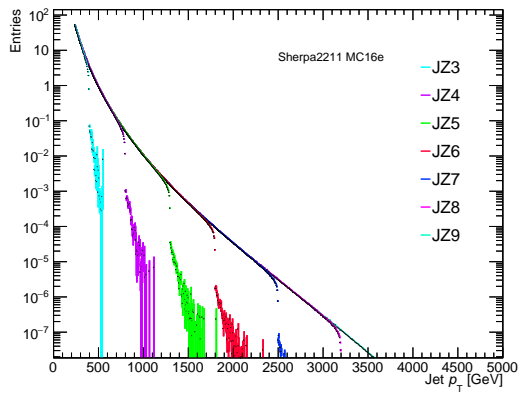
(b) MC16a DIRE



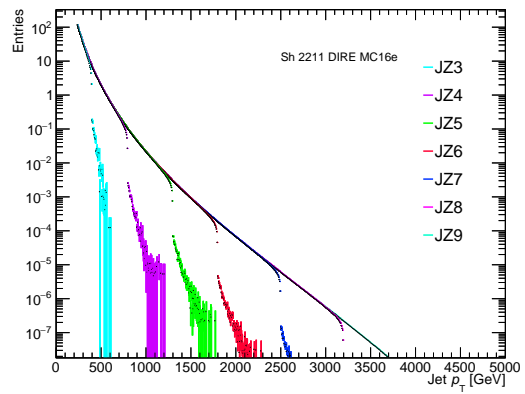
(c) MC16d 2.2.11



(d) MC16d DIRE

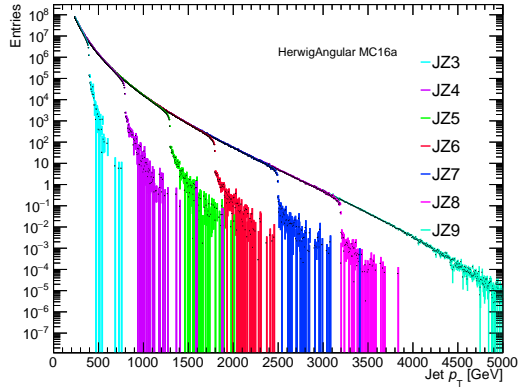


(e) MC16e 2.2.11

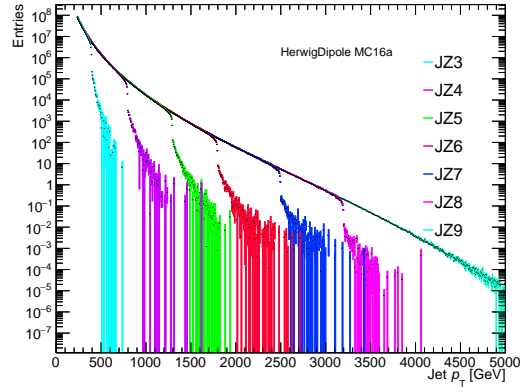


(f) MC16e DIRE

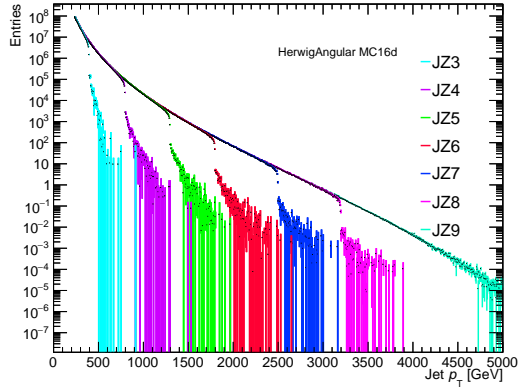
Figure 8.5: Validation that p_T spectrum is smooth in SHERPA 2.2.11 and DIRE samples.



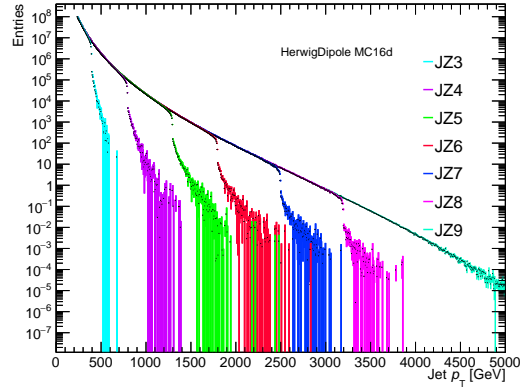
(a) MC16a Ang. Ord.



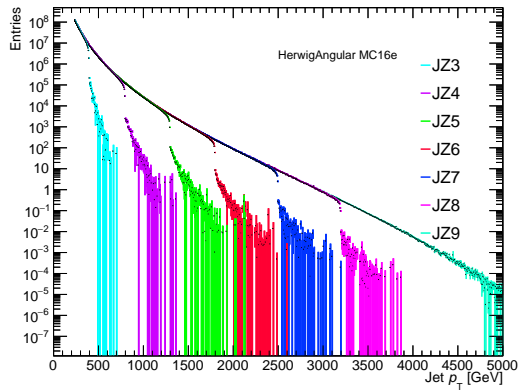
(b) MC16a Dipole



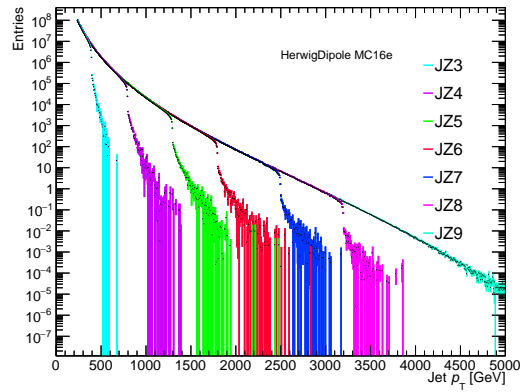
(c) MC16d Ang. Ord.



(d) MC16d Dipole



(e) MC16e Ang. Ord.



(f) MC16e Dipole

Figure 8.6: Validation that p_T spectrum is smooth in HERWIG 7 samples.

Generator	PS Model	Hadronization Model	PDFs
PYTHIA 8.230	dipole style, p_T ordered	Lund string	NNPDF 2.3
POWHEG V2 + PYTHIA 8.230	dipole style, p_T ordered	Lund string	NNPDF30NLO
SHERPA 2.2.5	CSS	Lund string	CT14NNLO
SHERPA 2.2.5	CSS	Cluster (AHADIC)	CT14NNLO
SHERPA 2.2.11	CSS	Cluster (AHADIC)	CT14NNLO
SHERPA 2.2.11	DIRE	Cluster (AHADIC)	CT14NNLO
HERWIG 7.1.3	angular-ordered	Cluster	MMHT2014NLO
HERWIG 7.1.3	dipole	Cluster	MMHT2014NLO

Table 8.2: Summary table of MC generators used in this analysis and their various models.

8.5 Jets, Tracks, and Event Selection

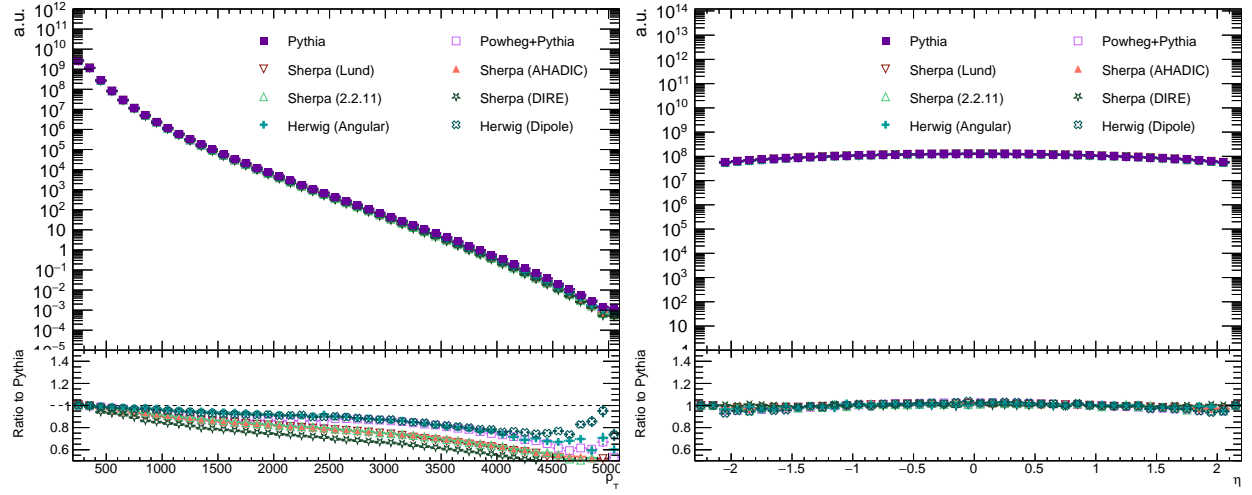
8.5.1 Jets

All jets in this analysis are reconstructed using the anti- k_t algorithm as implemented in FASTJET [76], using a jet radius parameter of $R = 0.4$.

Particle-level jets, or ‘truth jets’, are reconstructed in simulated events at generator level. Particles that are expected to leave only negligible energy depositions in the calorimeter, *i.e.* muons and neutrinos, are excluded. Particle-level jets are required to have a $p_T > 120$ GeV and a rapidity y satisfying $|y| < 2.1$ to be selected for this analysis.

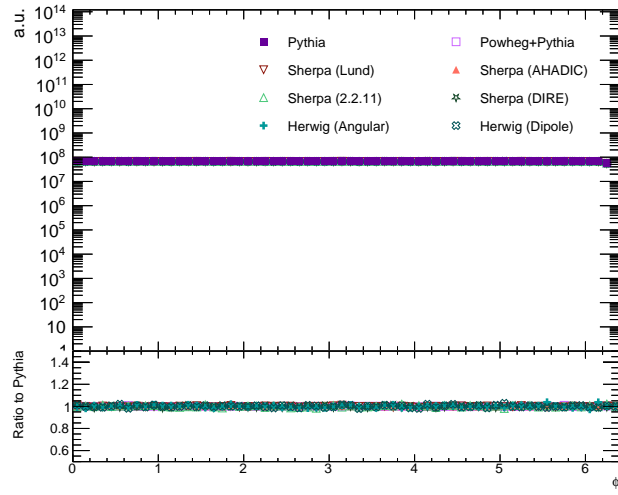
Standard jet kinematic information is shown for the particle-level jets selected in this analysis in Figure 8.7.

Detector-level jets, or ‘reco jets’, are reconstructed from particle-flow objects [25] that combine measurements from the ATLAS inner detector and calorimeter systems in order to improve the jet energy resolution (JER) and improve jet reconstruction efficiency, especially at low jet p_T . See Subsection 8.7.4 for additional details.



(a) jet p_T

(b) jet η



(c) jet ϕ

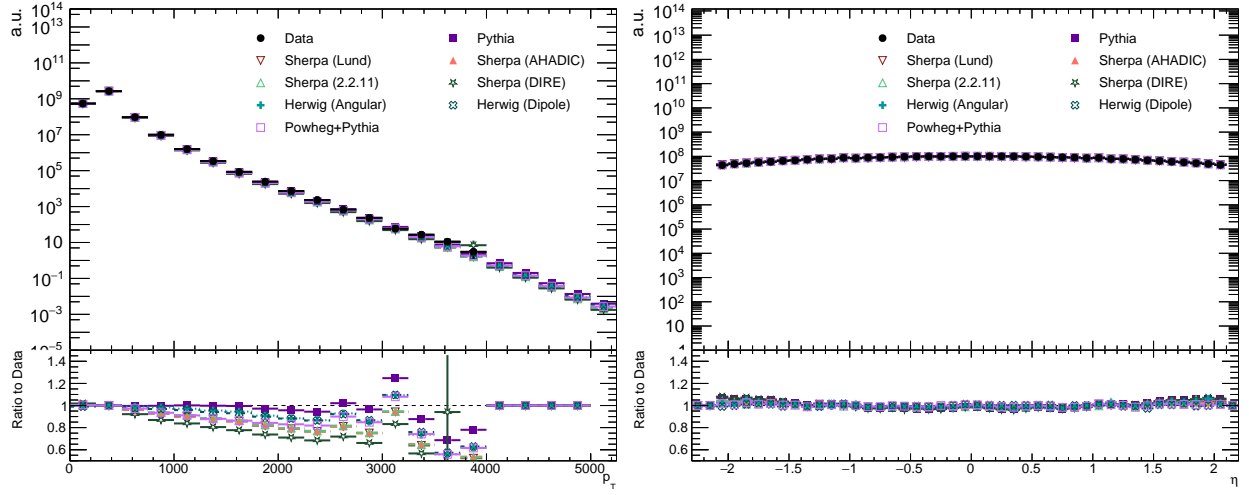
Figure 8.7: Particle-level jet kinematics.

Detector-level jets are required to have a $p_T > 120$ GeV and a rapidity y satisfying $|y| < 2.1$ to enter this analysis. In order to reject events which are contaminated by non-collision backgrounds (calorimeter noise bursts, readout errors and other pathologies) a jet cleaning criteria is applied, following the methodology described in Ref. [42]. Events are rejected if any leading or subleading jet fails the ATLAS jet cleaning criteria. The ATLAS jet cleaning working points are documented in Ref. [42], and include subsequent updates⁸. Following these cleaning cuts, the impact of noise is negligible in the analysis. The likelihood that a particle-flow jet originates from a pileup interaction following the kinematic selections is sufficiently low such that no additional pileup jet rejection is applied [39, 19]⁹.

Standard jet kinematic information is shown for the detector-level jets selected in this analysis in Figure 8.8.

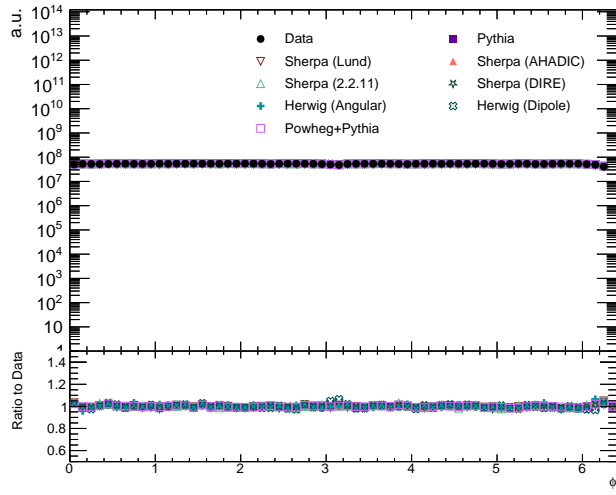
8. This analysis uses the 'tightBad' working point, documented in the jet cleaning twiki.

9. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PileupJetRecommendations>



(a) jet p_T

(b) jet η



(c) jet ϕ

Figure 8.8: Detector-level jet kinematics.

8.5.2 Tracks

Tracks at detector-level are reconstructed from charged-particle hits within the silicon- and straw-tube based inner tracking detectors, discussed in Chapter 4. The primary vertex (PV) of an event is defined as the reconstructed vertex with the highest sum of associated track p_T^2 . Tracks are required to pass a quality check based on p_T , η and the number of hits in specific inner detector modules [17]. They must also be associated with the PV; specifically tracks must be within 0.5 mm of the PV in both the y and z directions¹⁰. Tracks are required to have a p_T of at least 500 MeV to be included in this measurement. The track η and ϕ coordinates come from the five-parameter track fit to $(\eta, \phi, q/p, d_0, z_0)$, and thus correspond to the track coordinates at the origin.

Only charged particle-level information is used when calculating multiplicities. In particular, charged hadrons with $p_T > 500$ MeV are used to calculate the N_{Lund} and $N_{\text{Lund}}^{\text{Primary}}$ values for the truth-jets described in Section 8.5.1. The particle-level multiplicities are computed using the same charged-hadron association and C/A reclustering procedure as for detector-level jets.

8.5.3 Event Selection

All detector-level events are required to have at least one vertex reconstructed from two or more inner-detector tracks with $p_T > 500$ MeV, and to pass the data quality requirements described in Section 8.3. Events in data must be selected by the single-jet trigger that corresponds to the p_T of the leading jet in the event (Section 8.3). To ensure that dijet topologies are selected, a kinematic requirement is placed on the p_T of the leading jet pair

10. This corresponds to the 'tight' PV-association criteria as described by the `TrackVertexAssociationTool`.

in the event:

$$\frac{p_{\text{T}}^{\text{leading}}}{p_{\text{T}}^{\text{subleading}}} < 1.5.$$

After the event selection, over 175 million jets enter the analysis. The efficiency of the complete event selection in data, including the trigger requirements, data quality requirements and dijet balance selection is $\geq 93\%$.

8.5.4 Binning and Normalization

In this analysis, both N_{Lund} and $N_{\text{Lund}}^{\text{Primary}}$ were each calculated separately with k_t cuts [> 0.5 , > 1 , > 2 , > 5 , > 10 , > 20 , > 50 , > 100] GeV, resulting in a total of 16 observables. They are measured differentially in bins of the jet p_{T} , relative rapidity-ordering of the jets, and the multiplicity itself.

The jet p_{T} bins are (exclusively) [240, 300, 500, 750, 1000, 1250, ∞] GeV. The lowest p_{T} has been included to catch migrations into and out of the fiducial region during the unfolding procedure, and will not be published. The performance of the measurement in this bin has not been emphasised during the analysis optimisation.

The jets are separated into two relative rapidity bins according to which is the more-central and more-forward jet in the dijet system. This is done to account for differences in the quark and gluon composition of the jets in these regions, and also to enable demixed interpretations in terms of the underlying quark and gluon-labeled jets (see *e.g.* Ref. [27], Section 9.5), though this type of interpretation will not be published with this analysis.

The Lund multiplicity binning ranges from 0 to 60, with a different variable binning scheme chosen for each k_t cut and in each p_{T} bin. No jets in the Run 2 data have a multiplicity above

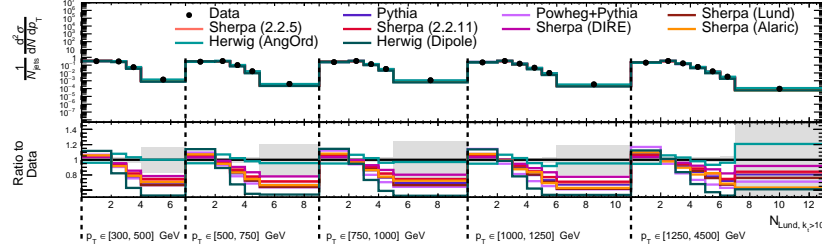
60 for any of the p_T bins or k_t cuts studied in this measurement. Nevertheless, the upper end of the binning is inclusive, as some jets in simulated samples possess higher multiplicities than those observed in data.

For the unfolding procedure, multiplicity bins of width one are used. The unfolded multiplicity distributions are rebinned post-unfolding to ensure that the bins used to present the differential cross-section results have tolerable detector-to-truth-level correspondence. The unfolding fake and efficiency correction factors, and the diagonal-fraction (‘purity’) of the rebinned response matrix are used to determine the coarser binning in this procedure, which targets efficiency and fake factors that are as high as possible, with a diagonal fraction *after* rebinning of roughly $\sim 40\%$. The data-driven non-closure uncertainty (Section 8.6.5) is also considered in this rebinning procedure, and is evaluated with the fine binning used for the unfolding procedure.

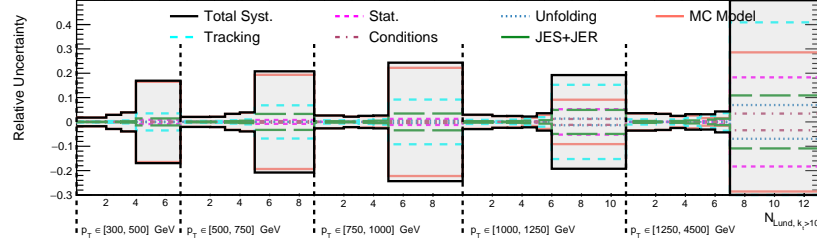
The unfolded distributions are normalised to the number of jets entering the measurement. Distributions are re-normalised during each systematic variation.

To calculate the average multiplicity as a function of the k_t cut (Subsection 8.8.2), the unfolded results using the binning scheme with unit-width bins is used. This approach avoids creating bin-bias uncertainties that would occur during the unfolding if the shape of the data and the Monte Carlo distributions differed within non-unit-width bins, at the expense of potentially introducing larger model-dependence during the unfolding procedure (Subsection 8.7.6).

This analysis has 16 observables, each measured triple-differentially. There is a lot of information to present, and thus the plots have been structured in a very specific way. An example of one set of plots for one observable is shown in Figure 8.9. The plots are presented inclusive in eta, with five jet p_T bins, marked with vertical dotted lines and labeled below the x-axis. The numerical axis labels inside of each of these p_T bins, along the x-axis, refer



(a) Full distribution of one observable.



(b) Full distribution of the relative uncertainty of one observable.

Figure 8.9: Example plot of one multiplicity observable and its associated uncertainties.

to the value of the multiplicity. The relative uncertainties are displayed in a separate plot, with identical binning structure, with the total uncertainty and the breakdown of the relative uncertainty shown.

8.6 Unfolding and Related Quantities

8.6.1 Iterative Bayesian Unfolding

In experimental physics, reconstructed objects are the main focus of most data analyses, since this is what is experimentally measured. However, observables calculated from reconstructed objects can often differ from truth level or theoretically calculated observables due several different effects. These include to efficiencies and fakes related to reconstruction, the scale and resolution of the object or observable, and other detector effects. The goal of unfolding is to correct for these effects and produce an observable that can be compared directly with

a calculated value, or across different experiments.

In this analysis an Iterative Bayesian Unfolding (IBU) technique is used to remove detector effects.

The process of unfolding utilizes Monte Carlo simulations at both the truth and reconstructed (reco) levels. This requires the full detector simulation to obtain the reco distributions.

The problem in its simplest form can be written as the following

$$\mathbf{R} \cdot \mathbf{t} = \mathbf{m} \quad (8.3)$$

where \mathbf{t} is the distribution of interest, the observable without any detector effects, and \mathbf{m} is the distribution measured. \mathbf{R} is clearly then a matrix which describes the probability of reconstructing a value into a bin i of the reco distribution, with a given true value in bin j , and is often referred to as the Response matrix. To write this another way:

$$R_{ij} = P(M_i|T_j) \quad (8.4)$$

Generally the matrix \mathbf{R} cannot be easily inverted without biasing the result due to the amplification of fluctuations. Instead, the process of Iterative Bayesian Unfolding is used to generate this matrix [100].

Essentially the goal is to produce a matrix by which one can multiply the measured distribution to receive back the "true" distribution. In other words, the goal is to find the matrix θ_{ij} which provides the probability of finding a truth value in bin i , given a measured value in bin j :

$$t_i = \sum_j P(T_i|M_j)m_j = \sum_j \theta_{ij} m_j \quad (8.5)$$

So instead of inverting the response matrix, one can instead use Bayes' theorem to define the matrix θ_{ij} in terms of generalized probabilities:

$$\theta_{ij} = P(T_i|M_j) = \frac{P(M_j|T_i) \cdot P(T_i)}{\sum_i P(M_j|T_i) \cdot P(T_i)} = \frac{R_{ji} \cdot P(T_i)}{\sum_i R_{ji} \cdot P(T_i)} , \quad (8.6)$$

where R_{ij} are the elements of the response matrix from Equation 8.3, and $P(T_i)$ is the probability of finding an event in bin i of the true distribution, and is often referred to as the prior.

The truth distributions of Monte Carlo samples can be used to estimate the starting prior $P(T_i)$. The process is performed iteratively to reduce dependence on this prior, such that:

$$\theta_{ij}^n = \frac{R_{ji} \cdot P(T_{i,n-1})}{\sum_i R_{ji} \cdot P(T_{i,n-1})} \quad (8.7)$$

where $t_{i,0}$ is the truth distribution produced initially by the Monte Carlo simulation. Additional iterations can decrease the dependence on the prior, but they can also create statistical fluctuations that can increase as the iterative process continues. Thus the number of iterations is really a balance between the prior dependence and this statistical effect, which are related to the finite nature of the data and Monte Carlo distributions.

Other quantities which are important in this process are the efficiency and fake factors. The efficiency is the probability of a given truth value being reconstructed in the reco distribution, and is defined as:

$$\epsilon_j = \sum_i R_{ij} \quad (8.8)$$

The fake factor can be defined similarly as the probability that a reconstructed value corre-

sponds to any truth value in the truth distribution:

$$f_i = \sum_j R_{ij} \quad (8.9)$$

The iterative Bayesian unfolding procedure used in this analysis is implemented using the package `RooUnfold2` [1]. ROOT version `6.14.04-x86_64-slc6-gcc62-opt` is used to set-up and run the unfolding code.

When dealing with actual the actual distributions in this analysis and not idealized mathematical probabilities, the response matrix becomes a 2D histogram of truth-level and reconstructed-level jets. Equally, the efficiency becomes the fraction of truth jets which are reconstructed and the fake rate becomes the fraction of reconstructed jets which exist at truth level.

8.6.2 Response Matrices

For this analysis, the number of iterations in the unfolding procedure was optimized minimizing the combined data-driven non closure, statistical, and modeling uncertainties. A full overview of this process can be found in Appendix C. Four unfolding iterations provides the best performance according to these metrics for all of the k_t cut values studied.

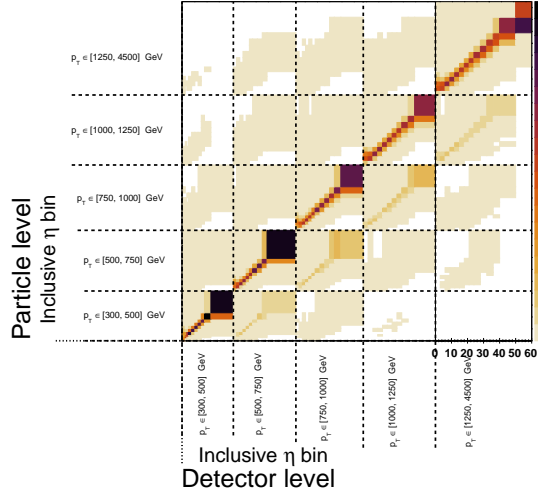
The unfolding in this analysis is performed with fine-binned (unit-width) histograms in multiplicity, which are then rebinned *after unfolding* in order to ensure adequate detector-to-truth-level correspondence in the differential measurement bins. The rebinned response matrices and diagonal fractions are shown in Figure 8.10 and Figure 8.11 for the Lund multiplicity, and in Figure 8.12 and Figure 8.13 for the primary Lund multiplicity.

This approach to unfolding has been discussed in Ref. [184], and is used in order to avoid

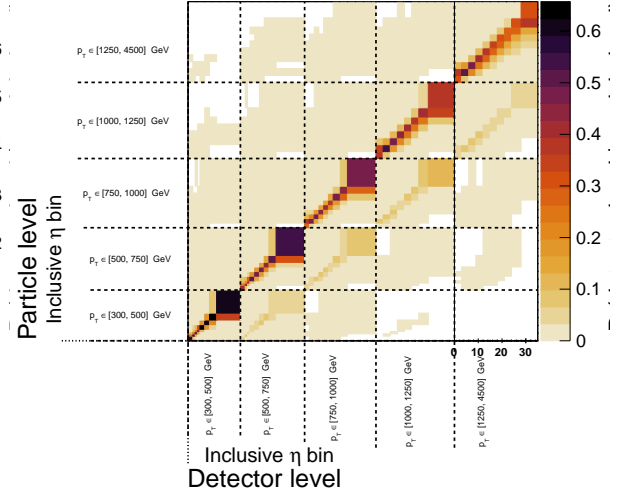
the potential bias inherent to the wide bins that are necessary in the tails of the multiplicity distribution. Note that the rebinning procedure improves the diagonal fraction substantially, providing regularization, while also ensuring that the measured distributions are not sensitive to sub-detector-resolution effects. Several cross-checks, documented in Appendix D, were performed to demonstrate that the unfolded distributions where the rebinning is done before and after the unfolding are compatible.

The average multiplicity is also determined from the fine-binned distributions, removing the need for a binning correction on the computation of the average. Averaging is a form of regularization, meaning that the same arguments that apply to rebinning the distribution after the unfolding also apply for the average. In addition, several checks were performed to ensure that this determination of the average is not extracting information that could not be extracted using the coarse-binned distribution, and that this determination does not neglect modeling uncertainties that would have arisen using the binning correction. These studies are shown in Appendix E.

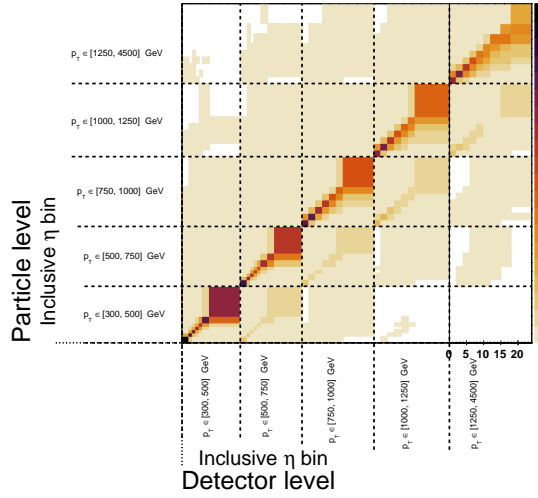
In all cases, the response matrices are normalised such that the integral of each detector-level bin is equal to unity, so that migrations between detector- and truth-level can be quantified in terms of the diagonal fraction of the matrix. A target of 40% of the entries in the rebinned unfolding matrices should remain on the main diagonal in each case, although this diagonal fraction can be slightly lower in certain bins in the tails of distributions. When the purity is low, it is important to verify that the unfolding closes technically (true by construction: see Subsection 8.6.4), and that the data-driven non-closure uncertainty is not large (Subsection 8.6.5). In some cases where there is low purity after rebinning, the cause is extremely low statistics in the tails of the measured distributions. In these cases, the large statistical uncertainty also covers potential unfolding biases (Subsection 8.7.1).



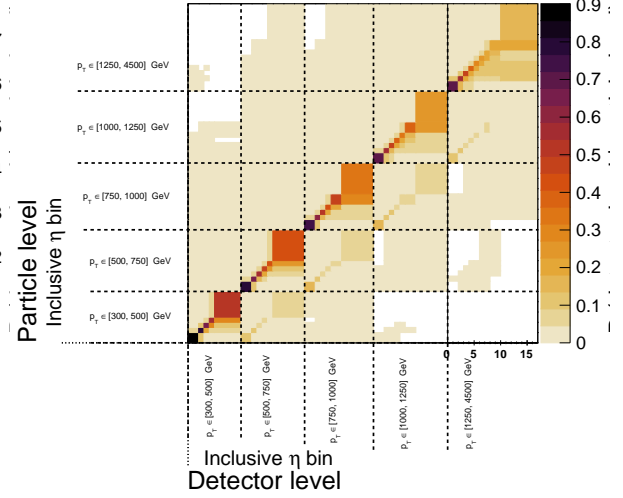
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

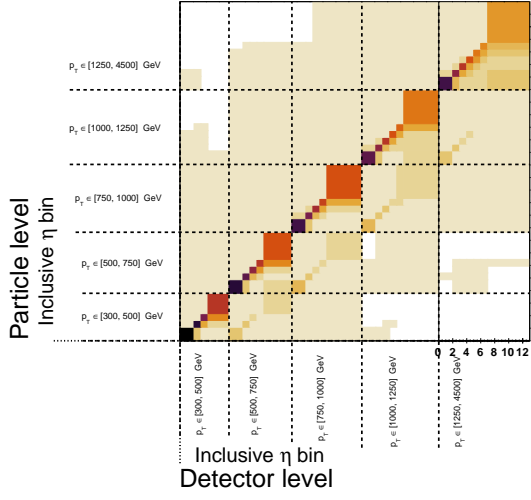


(c) $k_t > 2$ GeV, N_{Lund}

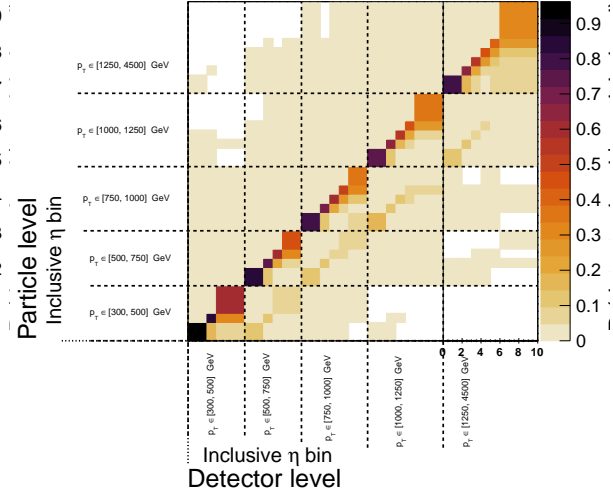


(d) $k_t > 5$ GeV, N_{Lund}

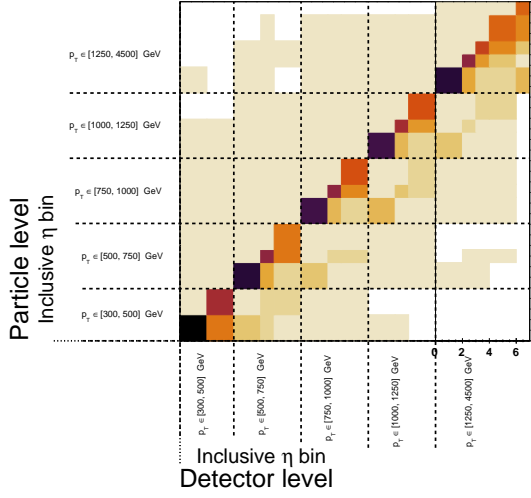
Figure 8.10: The rebinned PYTHIA unfolding matrices for Lund multiplicity with various k_t cuts. Figure continued on next page.



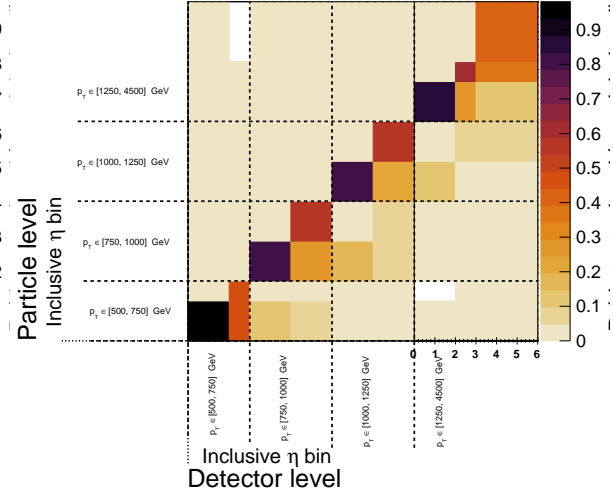
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

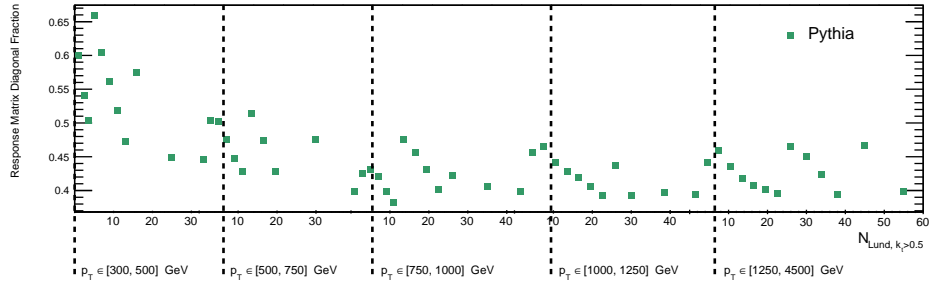


(g) $k_t > 50$ GeV, N_{Lund}

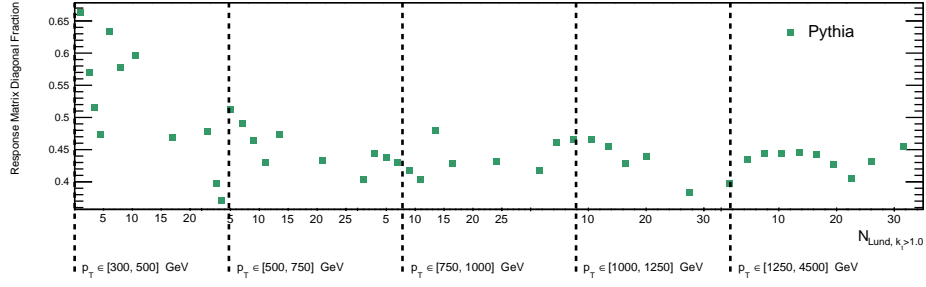


(h) $k_t > 100$ GeV, N_{Lund}

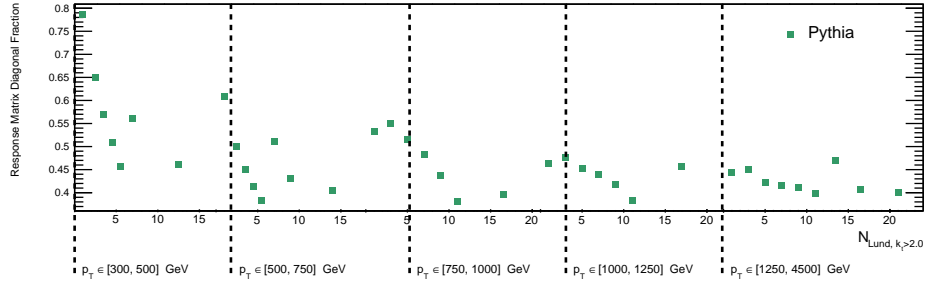
Figure 8.10: (continued) The rebinned PYTHIA unfolding matrices for Lund multiplicity with various k_t cuts.



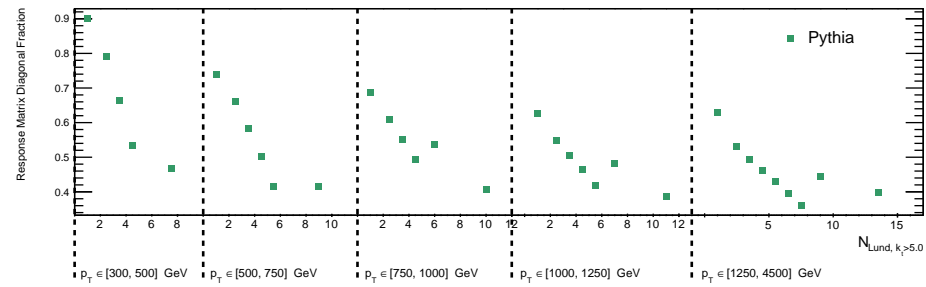
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

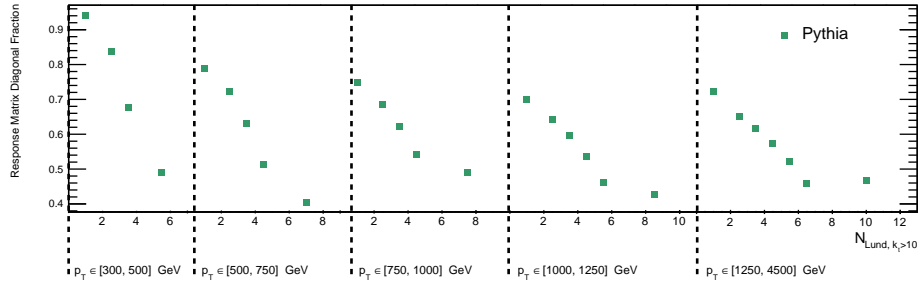


(c) $k_t > 2$ GeV, N_{Lund}

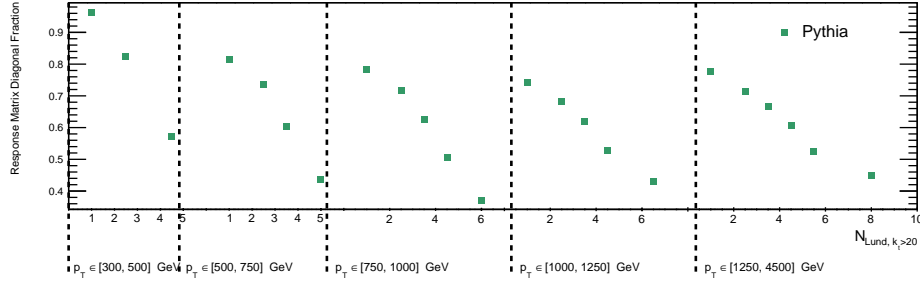


(d) $k_t > 5$ GeV, N_{Lund}

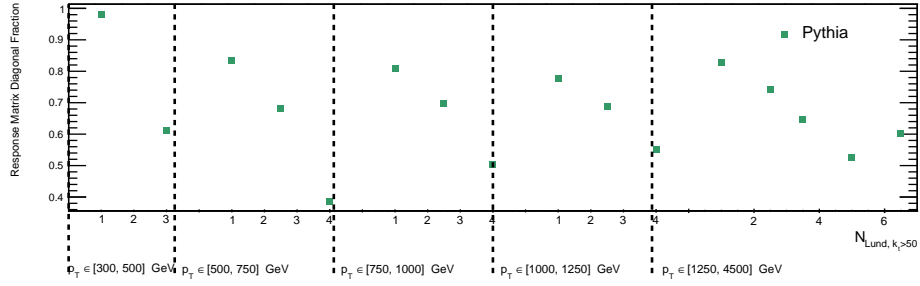
Figure 8.11: The rebinned PYTHIA unfolding matrix diagonal fraction for Lund multiplicity with various k_t cuts. Figure continued on next page.



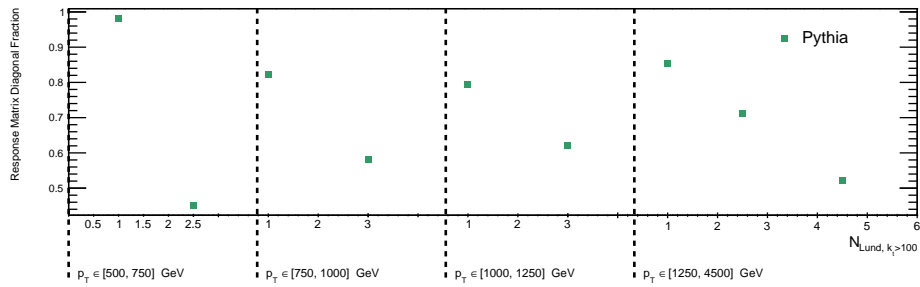
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

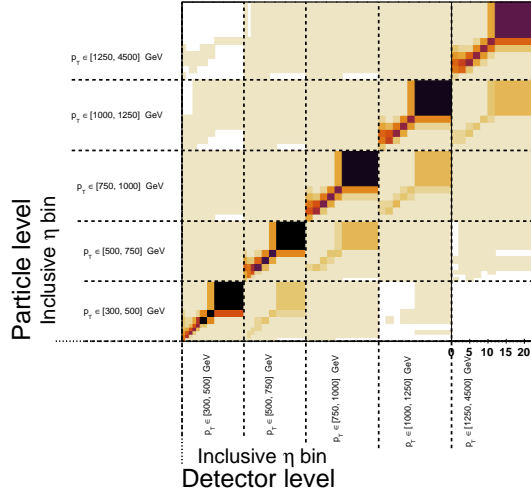


(g) $k_t > 50$ GeV, N_{Lund}

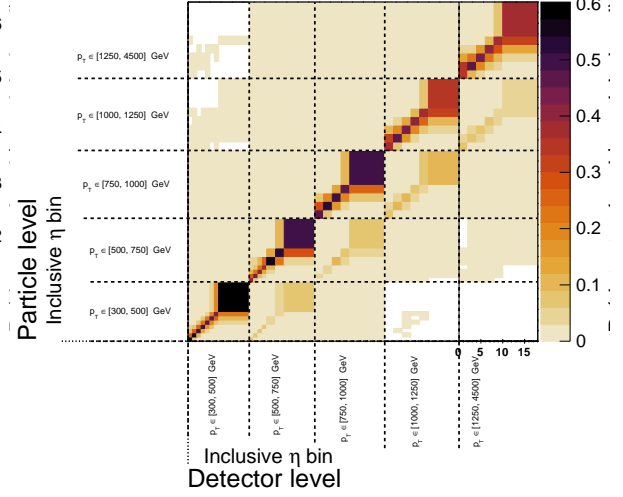


(h) $k_t > 100$ GeV, N_{Lund}

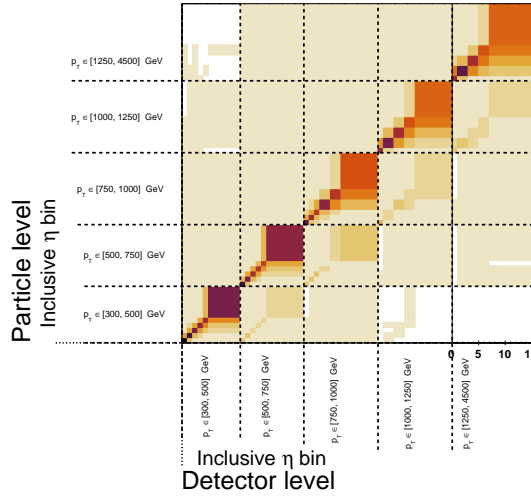
Figure 8.11: (continued) The rebinned PYTHIA unfolding matrix diagonal fraction for Lund multiplicity with various k_t cuts.



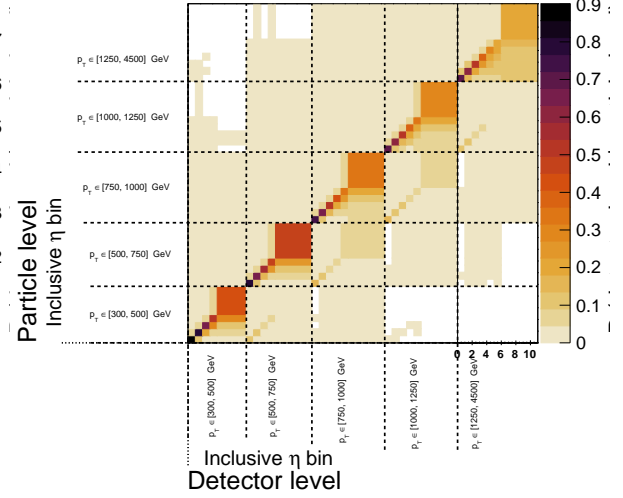
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

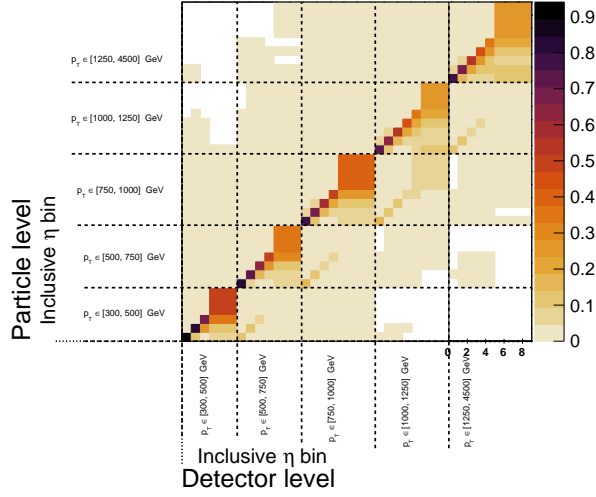


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

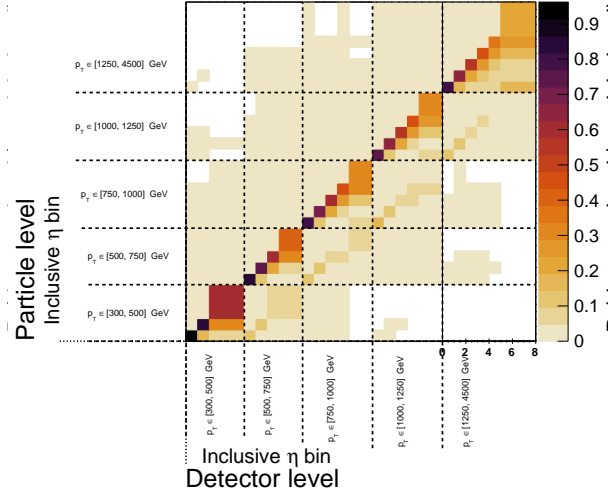


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

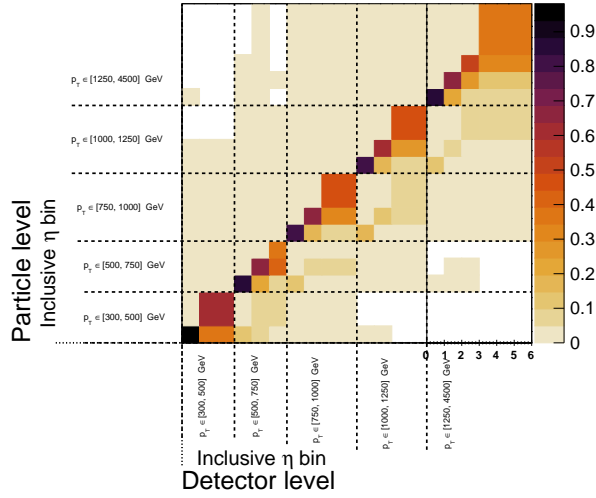
Figure 8.12: The rebinned PYTHIA unfolding matrices for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



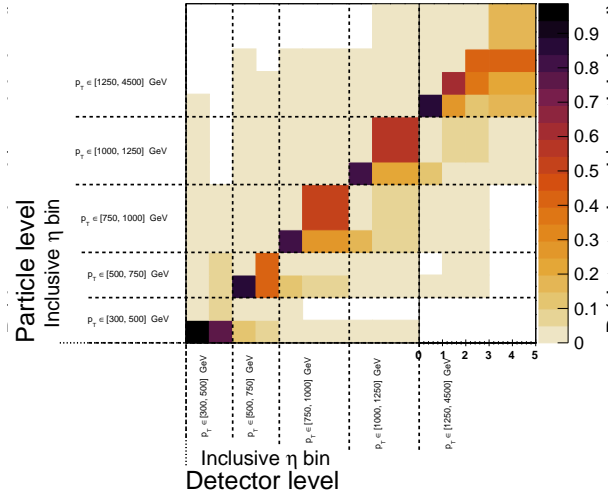
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

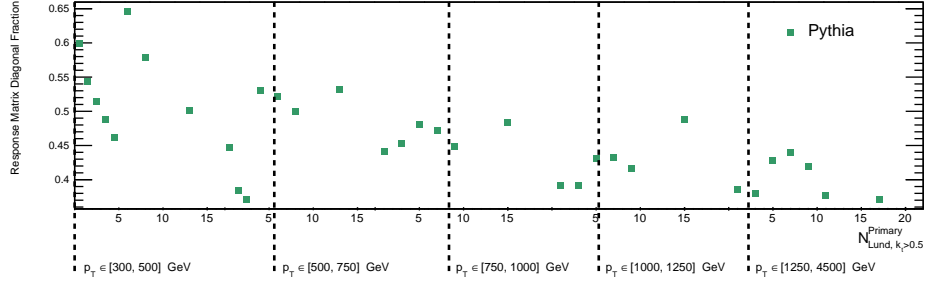


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

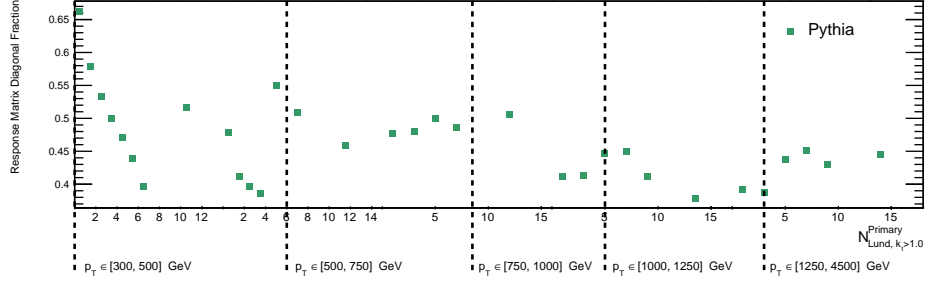


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

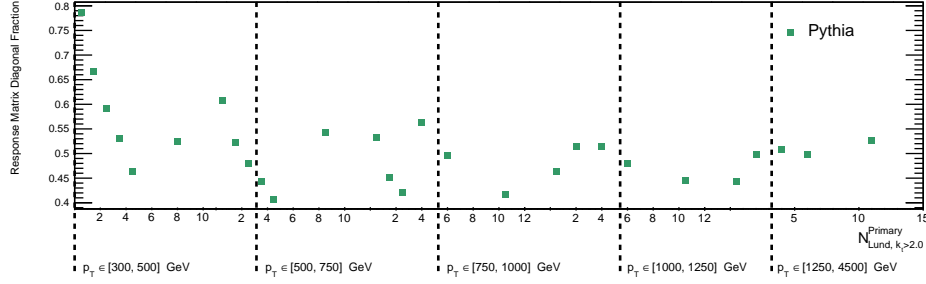
Figure 8.12: (continued) The rebinned PYTHIA unfolding matrices for primary Lund multiplicity with various k_t cuts.



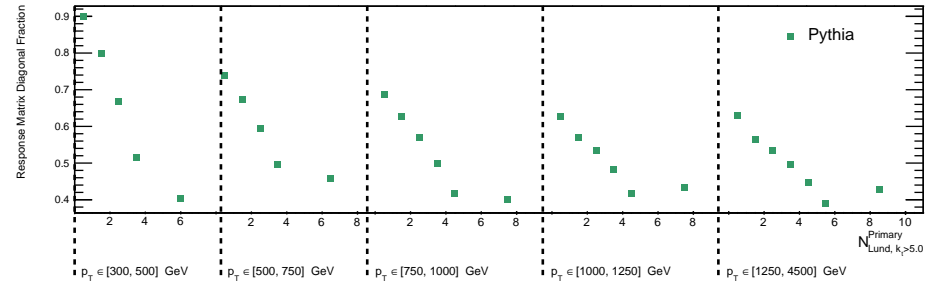
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

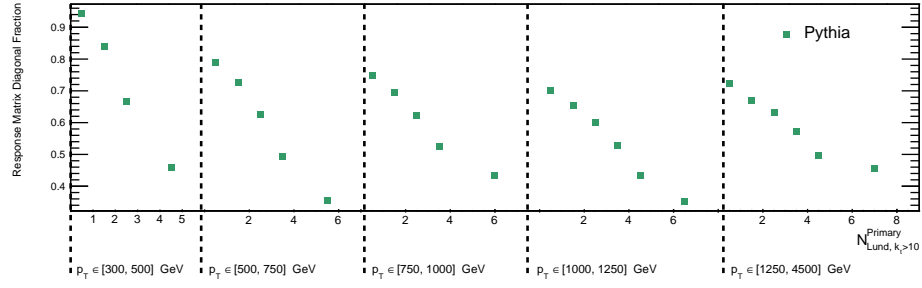


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

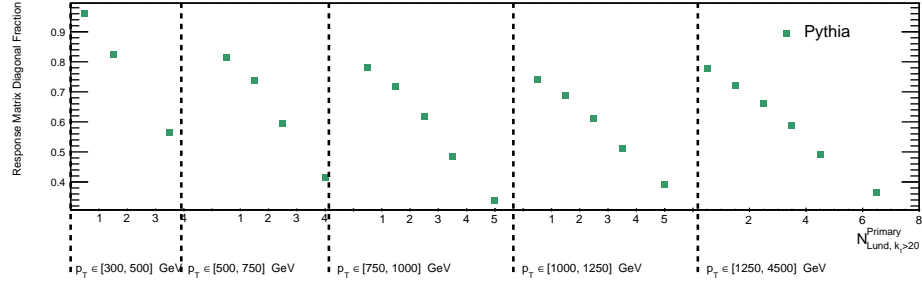


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

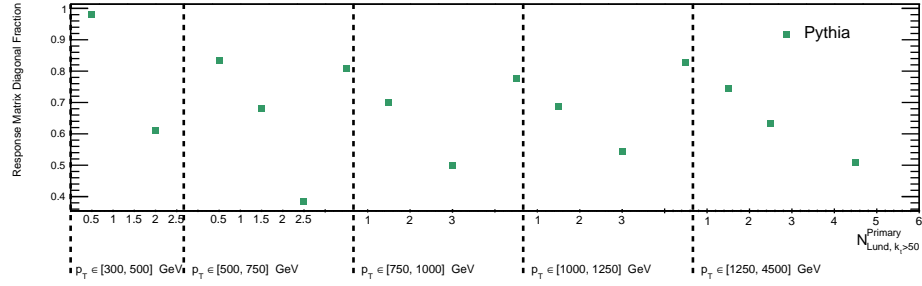
Figure 8.13: The rebinned PYTHIA unfolding matrix diagonal fraction for primary Lund multiplicity. Figure continued on next page.



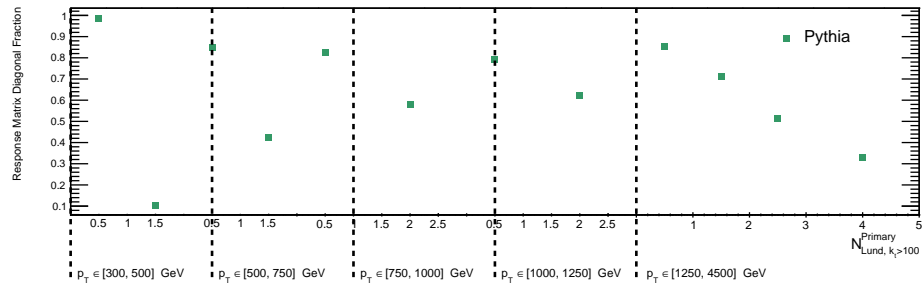
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



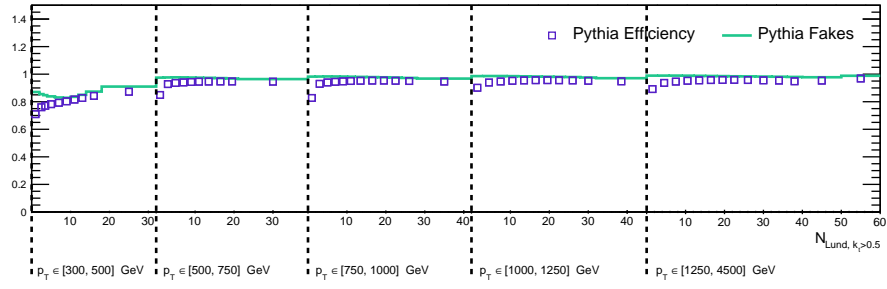
(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.13: (continued) The rebinned PYTHIA unfolding matrix diagonal fraction for primary Lund multiplicity.

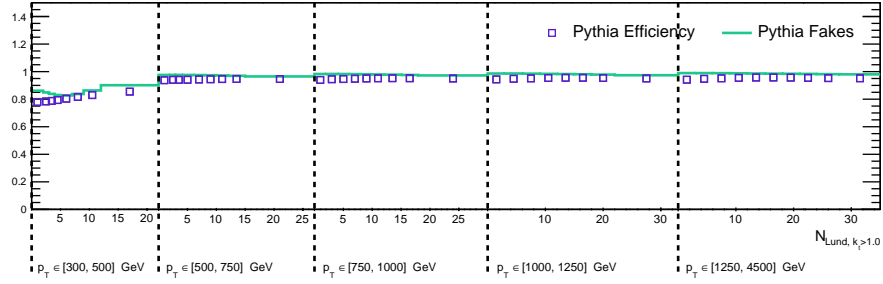
8.6.3 Fake and Efficiency Factors

The fake and efficiency factors for N_{Lund} and $N_{\text{Lund}}^{\text{Primary}}$ for all observables are shown in Figure 8.14 and Figure 8.15. They are typically above 90%, with the exception of the lowest jet transverse momentum bin where they are more often between 65% to 80% due to migrations out of the analysis fiducial region.

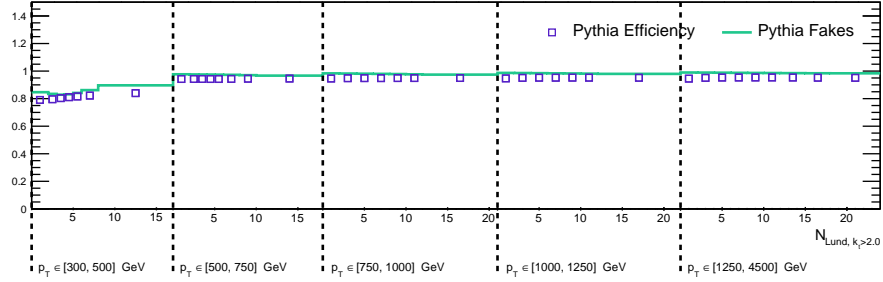
This is indicative that the binning used in the measurement is not too fine, and that the acceptance at particle- and detector-level is similar.



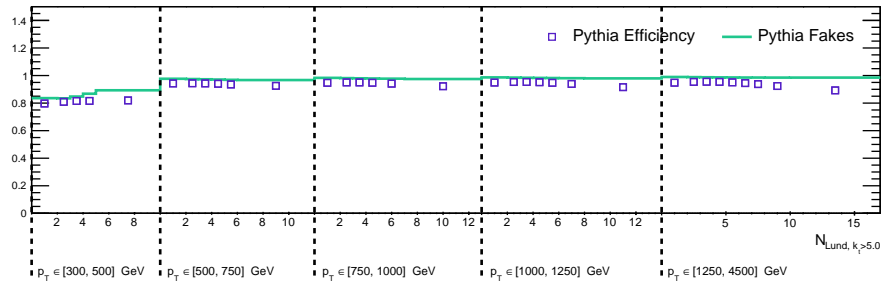
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

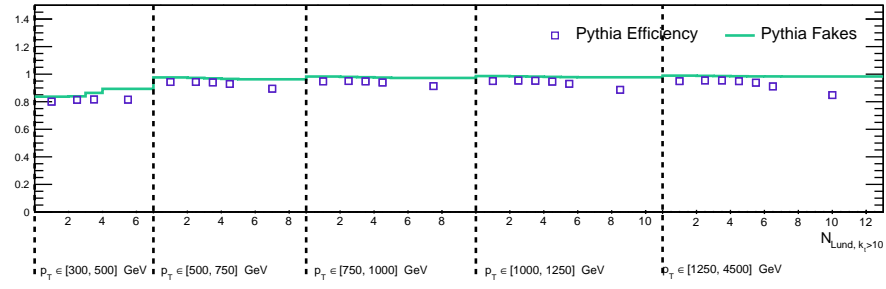


(c) $k_t > 2$ GeV, N_{Lund}

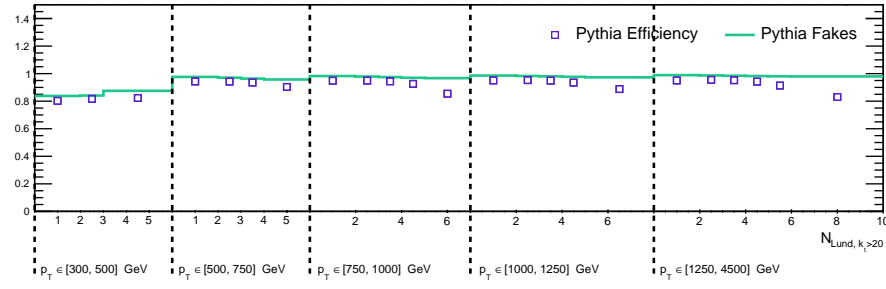


(d) $k_t > 5$ GeV, N_{Lund}

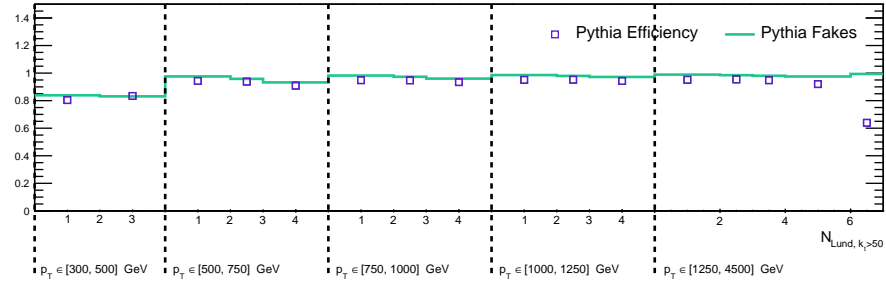
Figure 8.14: The fake and efficiency factors for Lund multiplicity with various k_t cuts. Figure continued on next page.



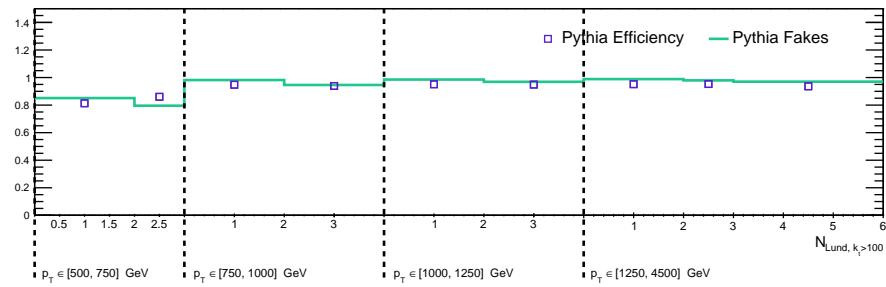
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

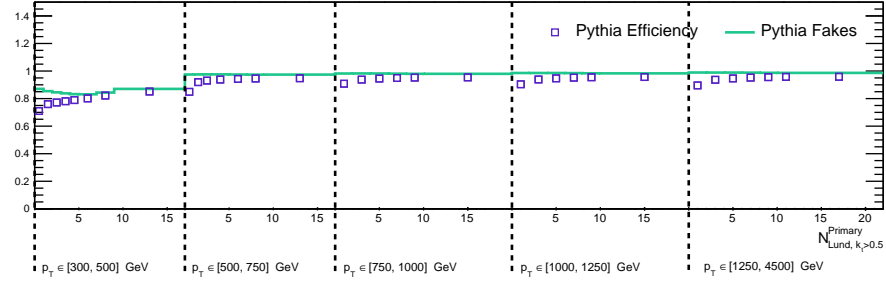


(g) $k_t > 50$ GeV, N_{Lund}

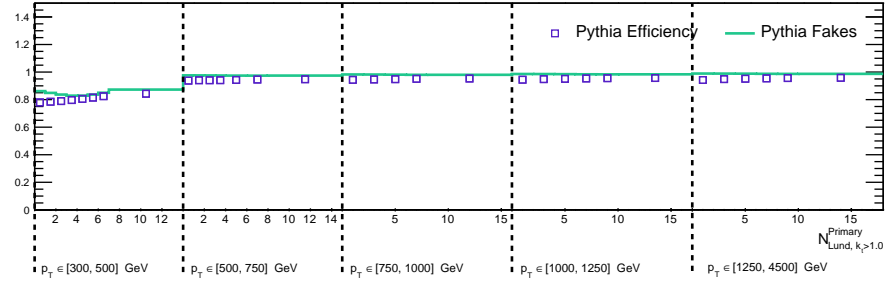


(h) $k_t > 100$ GeV, N_{Lund}

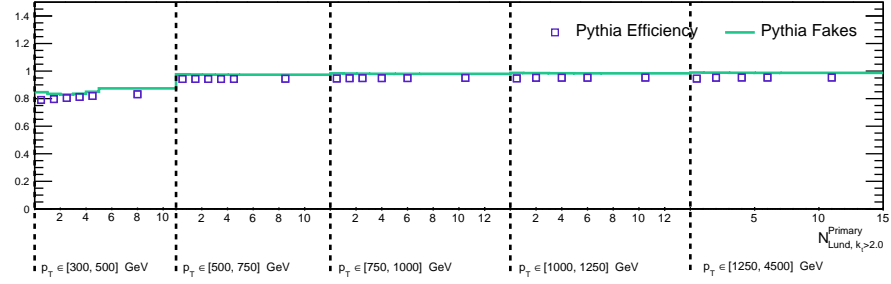
Figure 8.14: (continued) The fake and efficiency factors for Lund multiplicity with various k_t cuts.



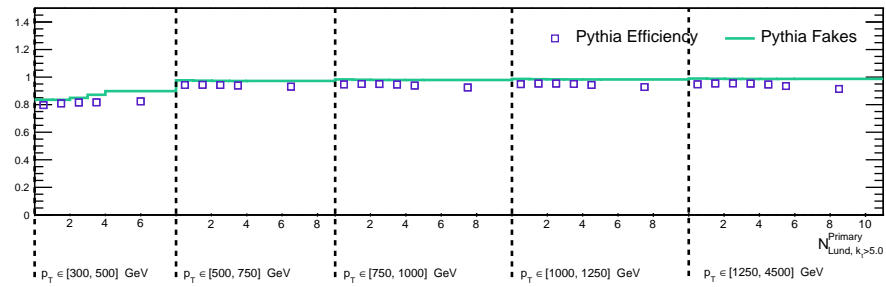
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

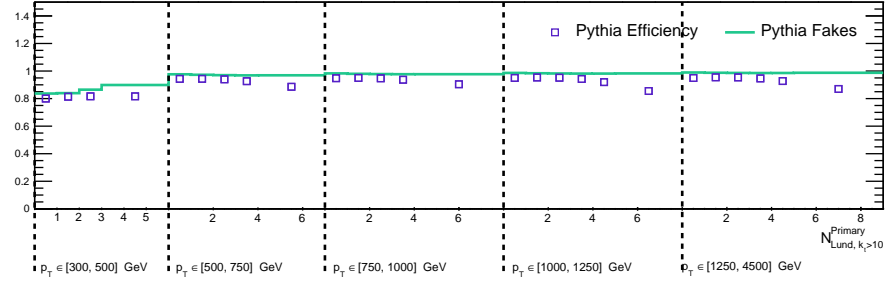


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

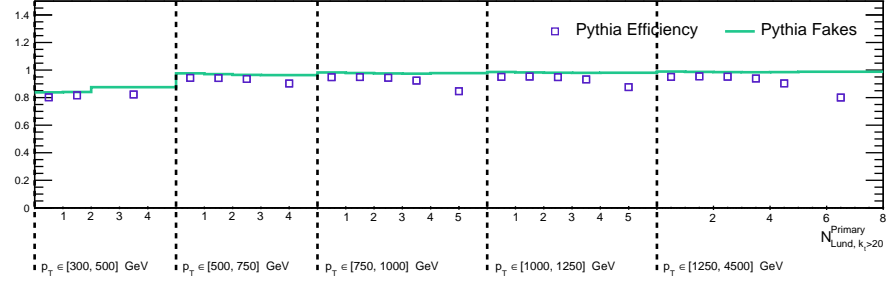


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

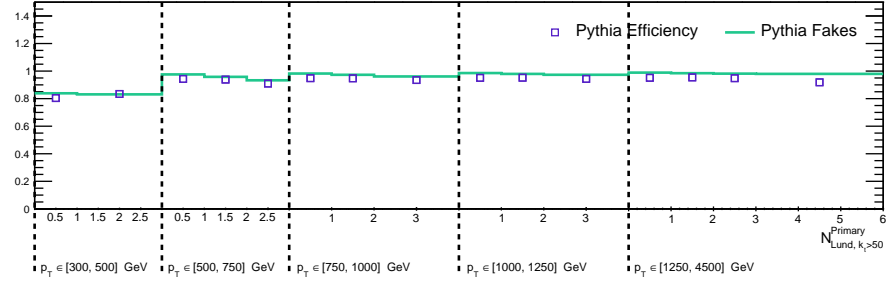
Figure 8.15: The fake and efficiency factors for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



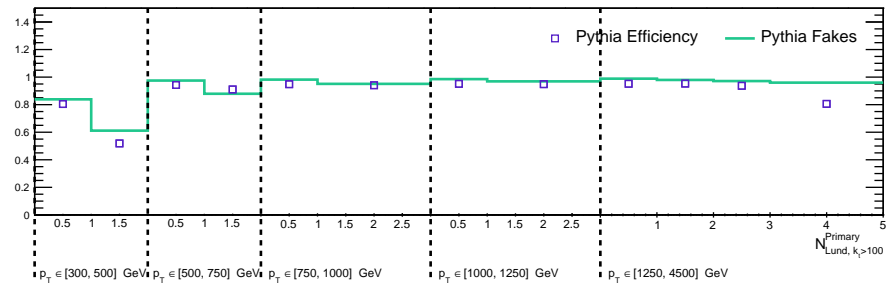
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



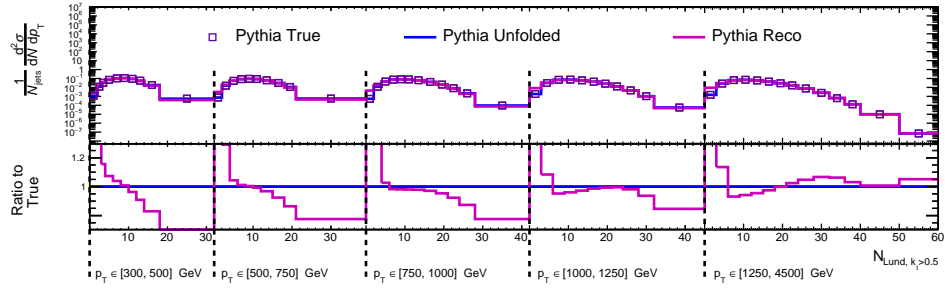
(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.15: (continued) The fake and efficiency factors for primary Lund multiplicity with various k_t cuts.

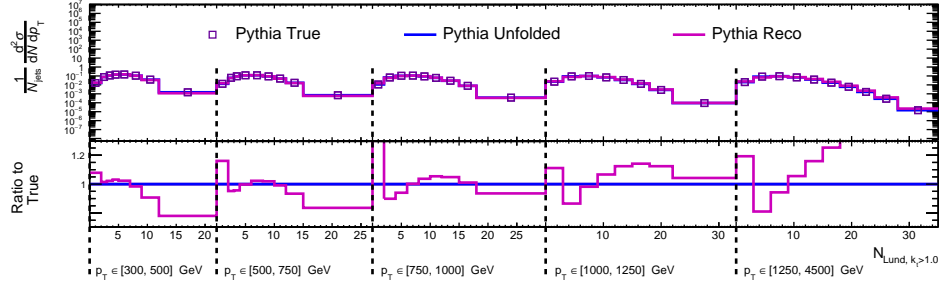
8.6.4 *Technical Closure*

In order to confirm that the unfolding methodology applied in this analysis is unbiased and functional, the *technical closure* of the procedure is verified by unfolding the nominal reconstructed simulation (PYTHIA v8) with its own response matrix. Perfect closure is expected from this test.

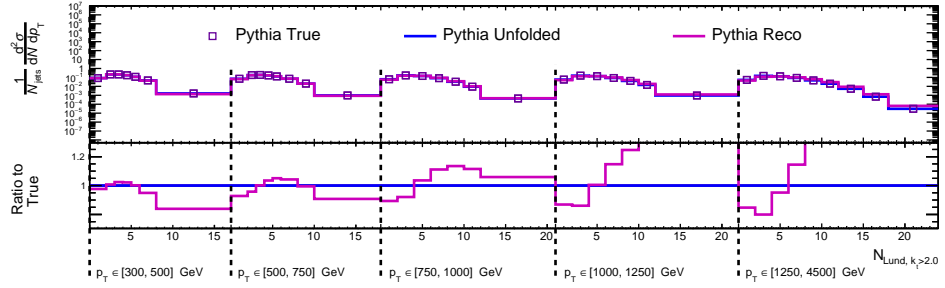
The results of this diagnostic cross-check are shown in Figure 8.16 and Figure 8.17 for the complete set of measured observables. The truth-level and unfolded distributions are identical in all cases (the solid blue line is at 1 in the ratio panels, for each bin), indicating technical closure. These plots also indicate the approximate size of the corrections made by the unfolding procedure (compare the unfolded and the reco distributions), which tend to show fair agreement except in the tails of multiplicity distributions, where differences of 20% or larger can manifest.



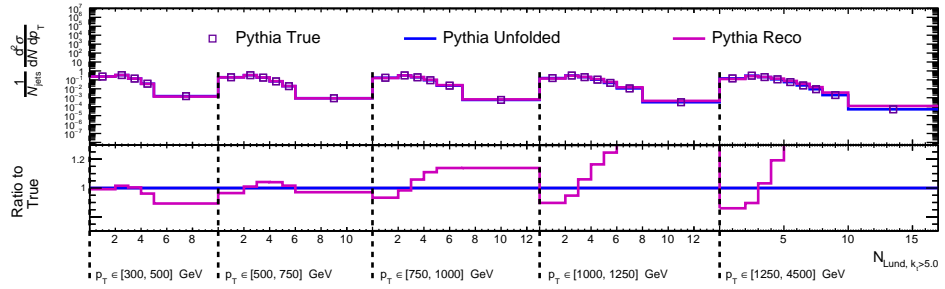
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

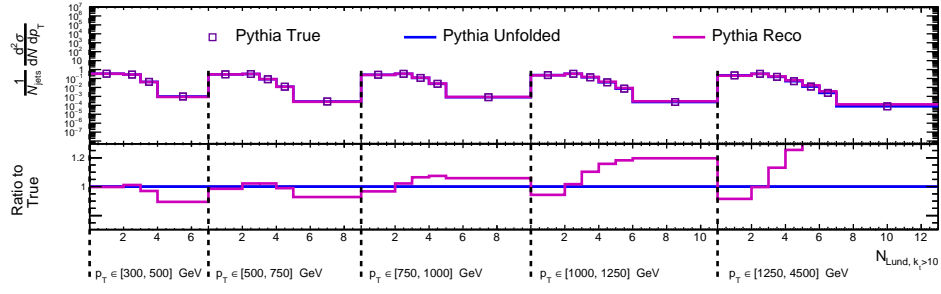


(c) $k_t > 2$ GeV, N_{Lund}

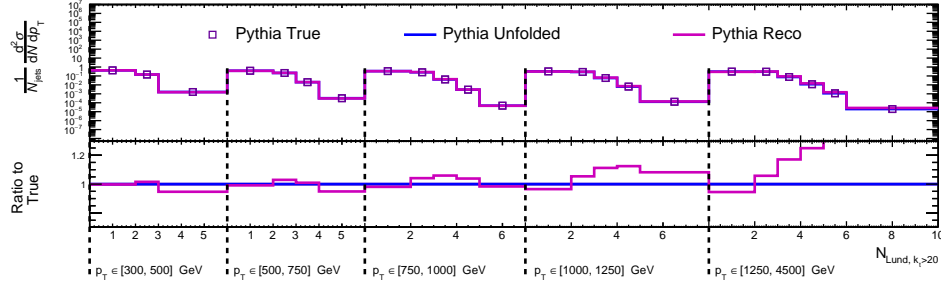


(d) $k_t > 5$ GeV, N_{Lund}

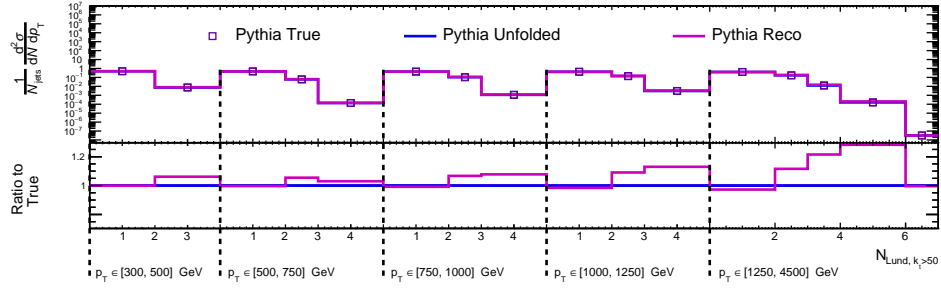
Figure 8.16: Technical closure cross-check results for the Lund multiplicity. The ratio of the unfolded PYTHIA results with the particle-level PYTHIA distribution is 1, indicating good technical closure. Figure continued on next page.



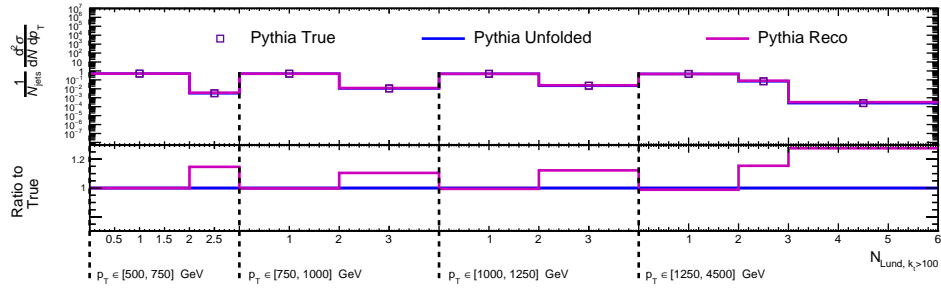
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

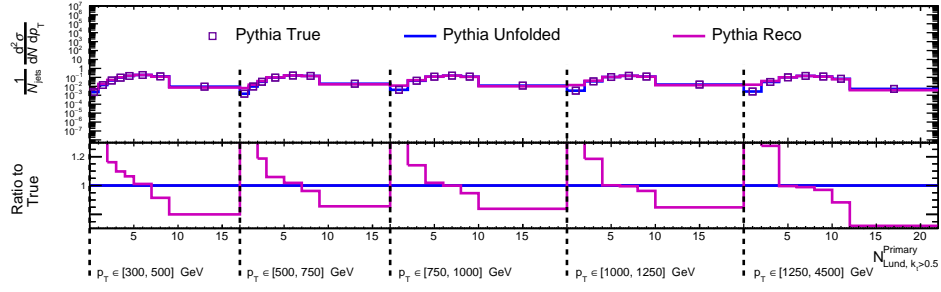


(g) $k_t > 50$ GeV, N_{Lund}

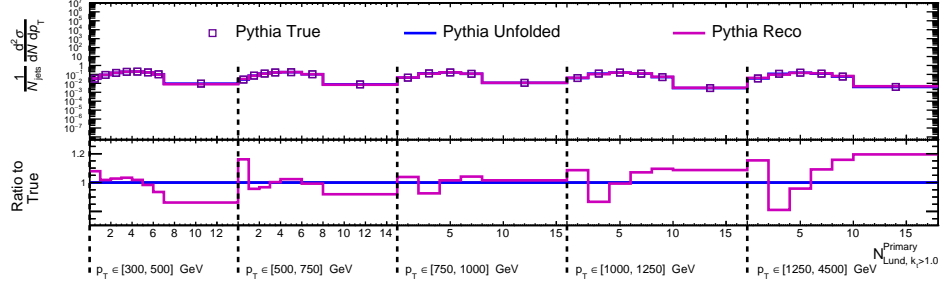


(h) $k_t > 100$ GeV, N_{Lund}

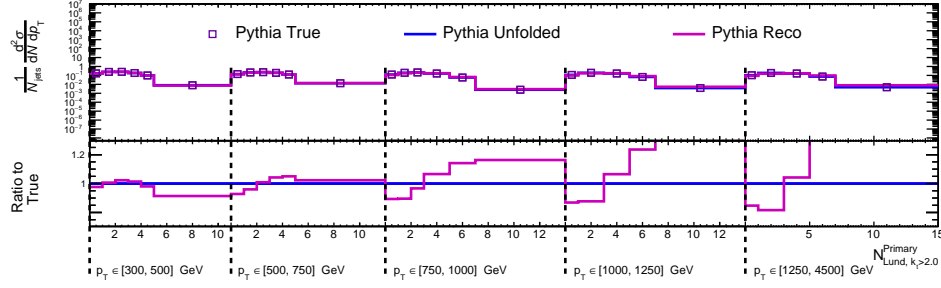
Figure 8.16: (continued) Technical closure cross-check results for the Lund multiplicity. The ratio of the unfolded PYTHIA results with the particle-level PYTHIA distribution is 1, indicating good technical closure.



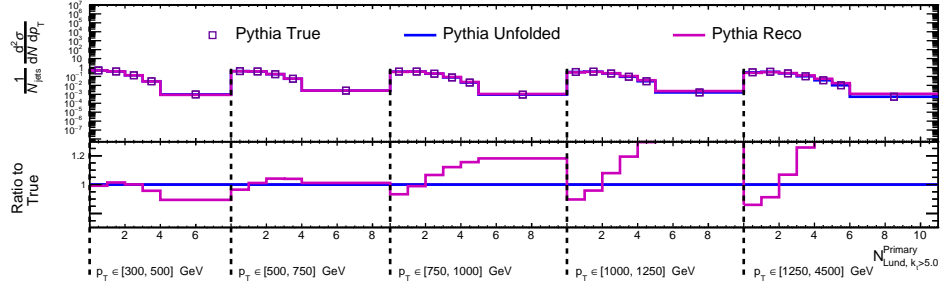
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

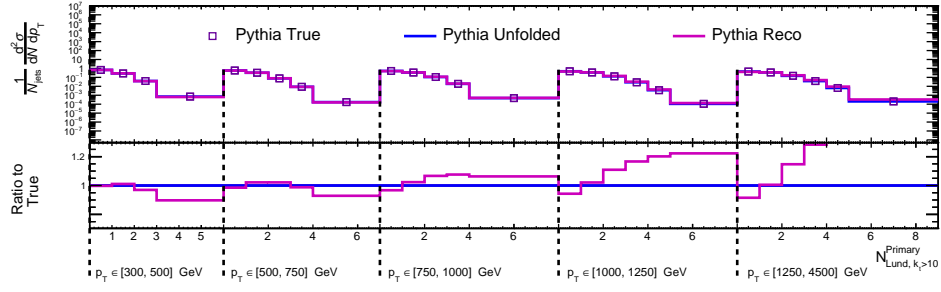


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

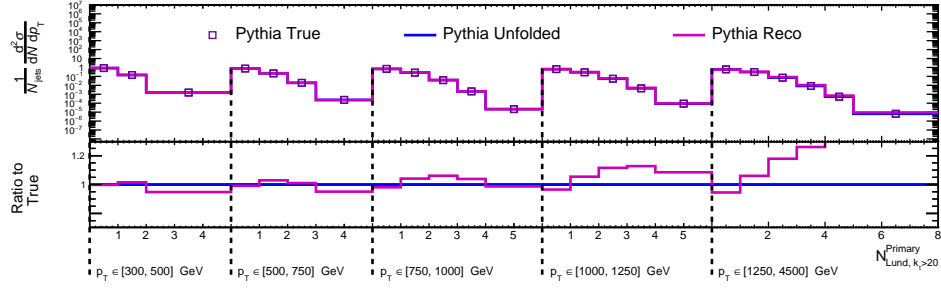


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

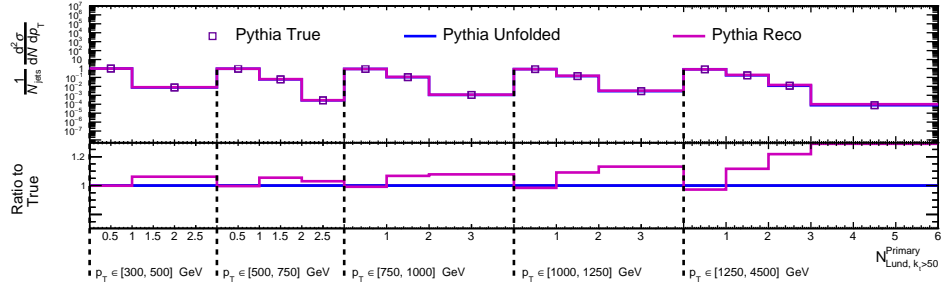
Figure 8.17: Technical closure cross-check results for the primary Lund multiplicity. The ratio of the unfolded PYTHIA results with the particle-level PYTHIA distribution is 1, indicating good technical closure. Figure continued on next page.



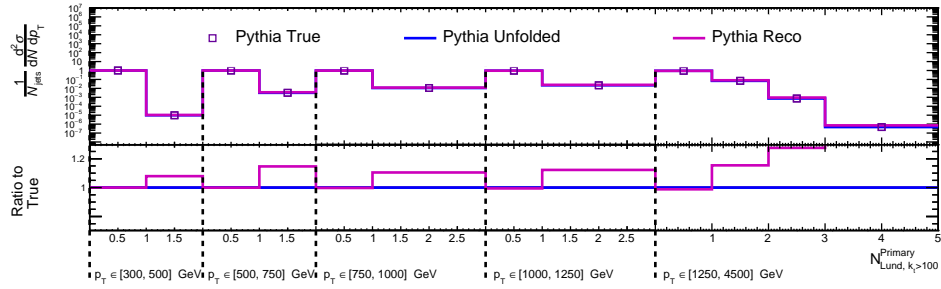
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

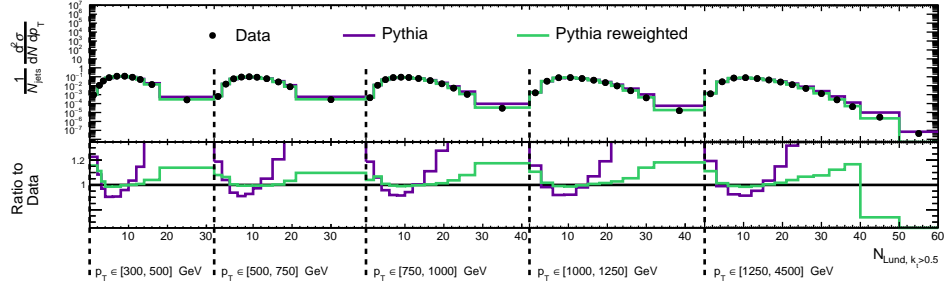
Figure 8.17: (continued) Technical closure cross-check results for the primary Lund multiplicity. The ratio of the unfolded PYTHIA results with the particle-level PYTHIA distribution is 1, indicating good technical closure.

8.6.5 Data Driven Non-Closure test for IBU

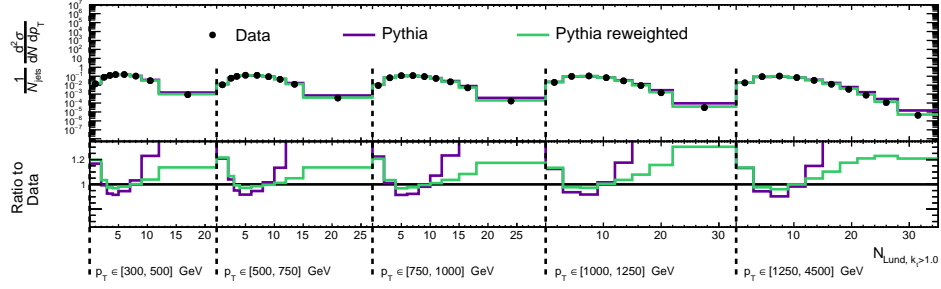
The standard method [170, 126] for evaluating the systematic uncertainty from the unfolding procedure is applied in this analysis. This method involves reweighting the truth-level MC distribution and recovering a reconstructed MC distribution by folding the truth distribution using the response matrix. The reweighting is chosen such that the reconstructed MC more closely matches the reconstructed data distribution. The data and the new reconstructed MC distribution are then unfolded with the original response matrix. The difference between the two is taken as an uncertainty.

Comparisons between the unfolded data, PYTHIA, and reweighted, unfolded PYTHIA distributions are provided in Figure 8.18 and Figure 8.19 for all measured observables. In general, these diagnostic figures show that the reweighting during the data driven non-closure (DDNC) procedure brings the unfolded reweighted PYTHIA distribution closer to the unfolded data, indicating that the process is working as intended.

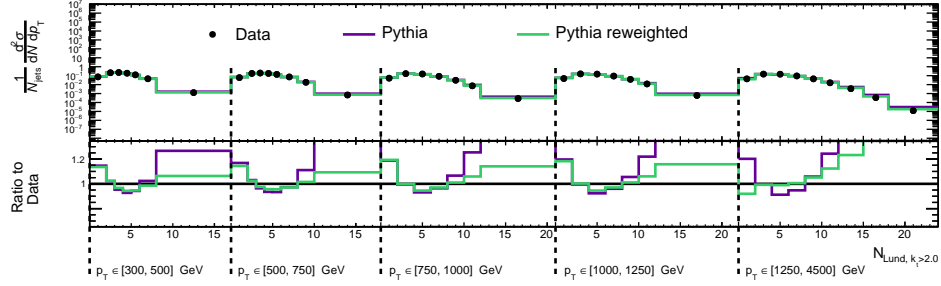
The uncertainty arising from this procedure is generally small ($\lesssim 5\%$) throughout the measurement, though it can increase substantially in the tails of distributions ($> 15\%$). The size of this uncertainty is shown in Figure 8.20 and Figure 8.21 for the complete set of measured distributions.



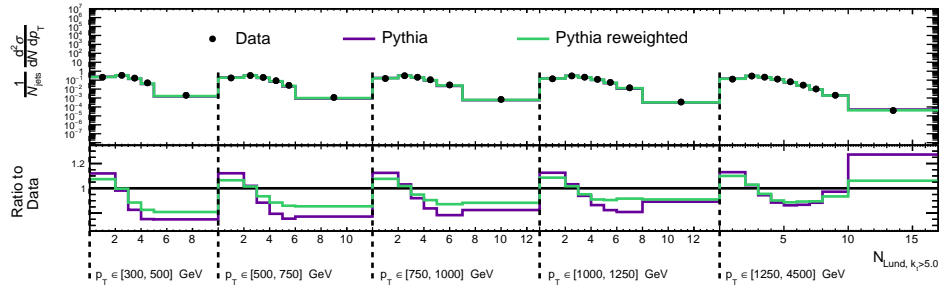
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

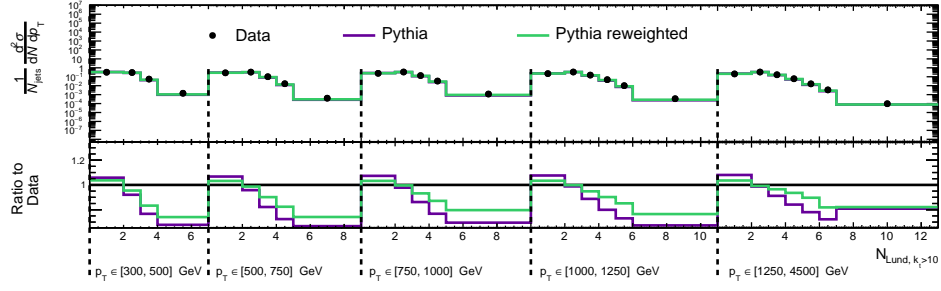


(c) $k_t > 2$ GeV, N_{Lund}

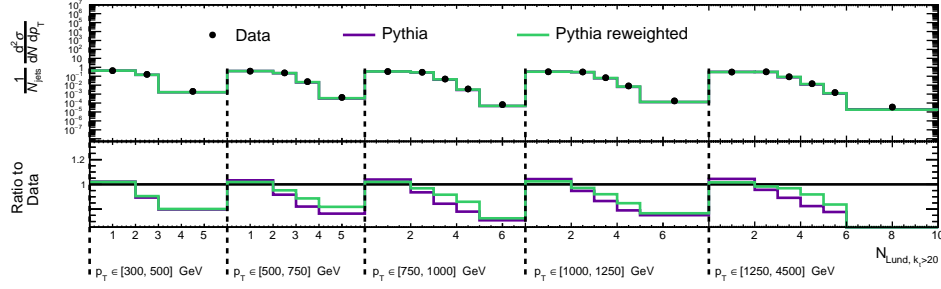


(d) $k_t > 5$ GeV, N_{Lund}

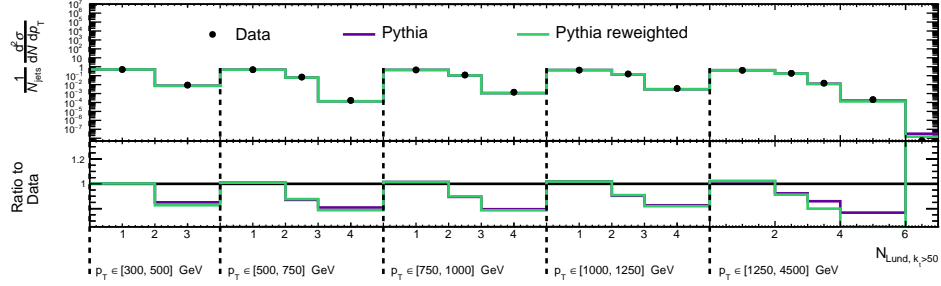
Figure 8.18: Data-driven non-closure diagnostic for Lund multiplicity, demonstrating the performance of the reweighting applied to the PYTHIA MC. Figure continued on next page.



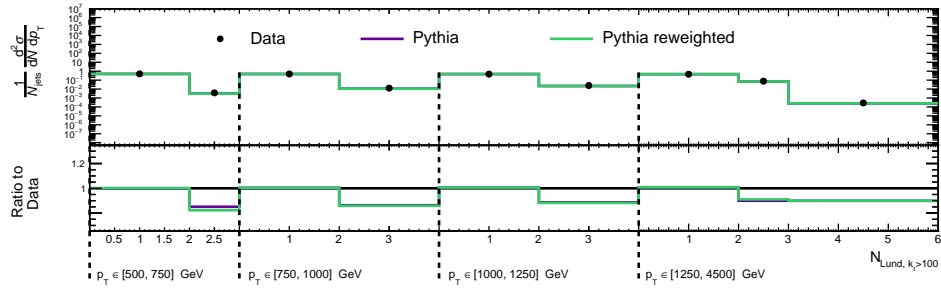
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

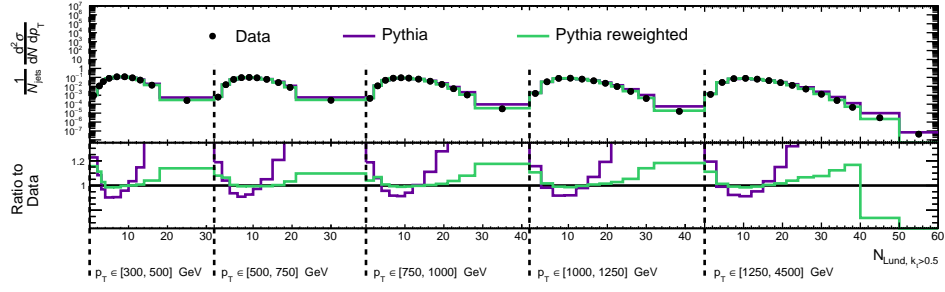


(g) $k_t > 50$ GeV, N_{Lund}

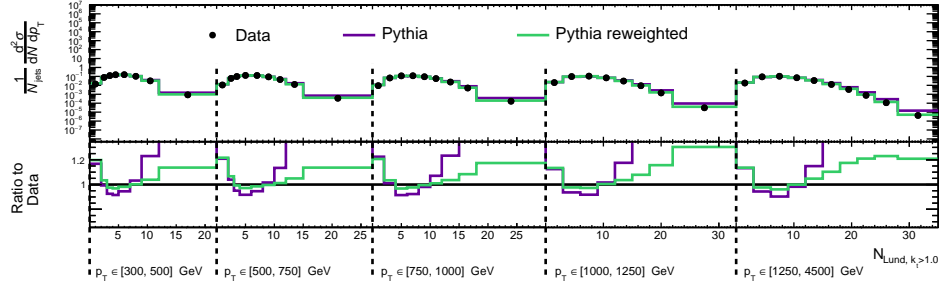


(h) $k_t > 100$ GeV, N_{Lund}

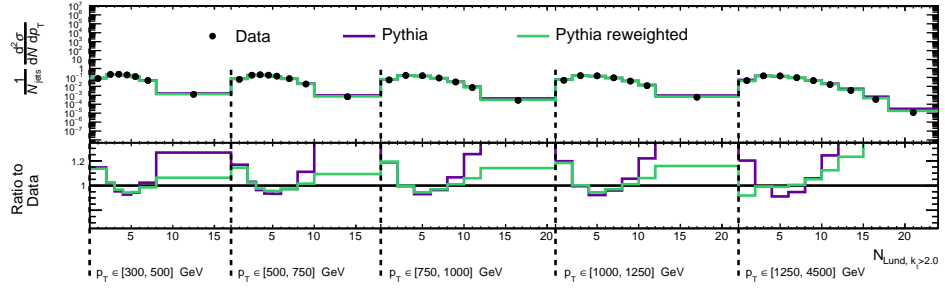
Figure 8.18: (continued) Data-driven non-closure diagnostic for Lund multiplicity, demonstrating the performance of the reweighting applied to the PYTHIA MC.



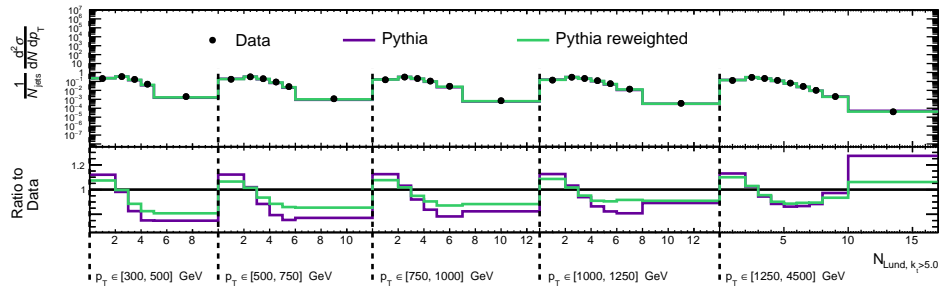
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

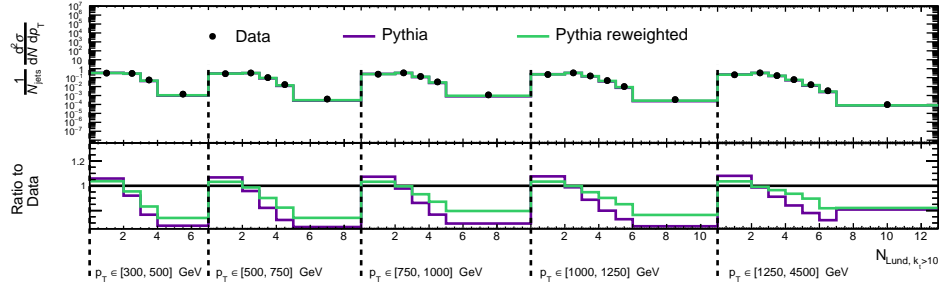


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

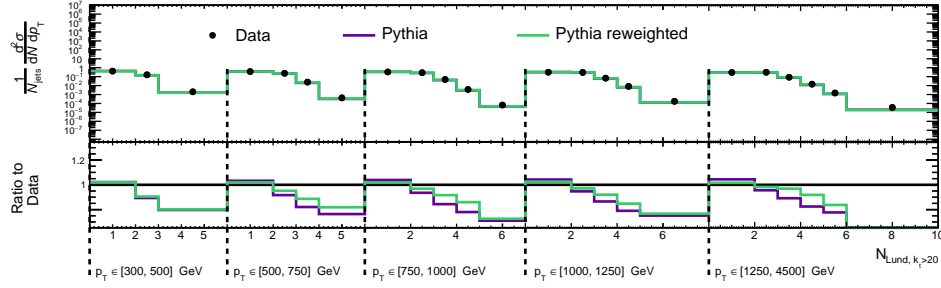


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

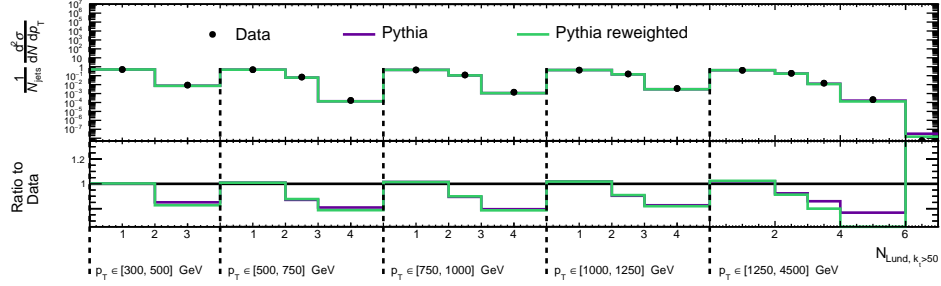
Figure 8.19: Data-driven non-closure diagnostic for primary Lund multiplicity, demonstrating the performance of the reweighting applied to the PYTHIA MC. Figure continued on next page.



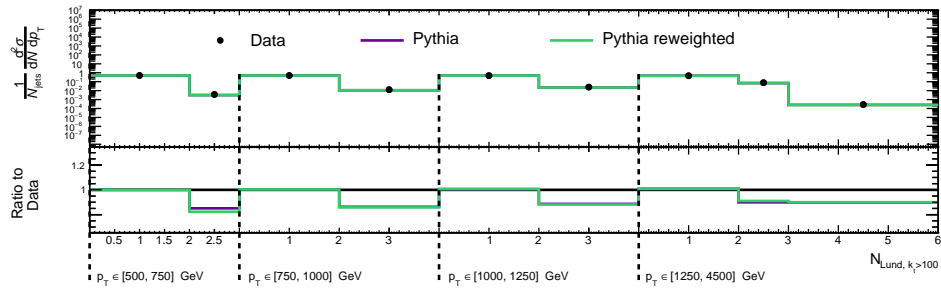
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

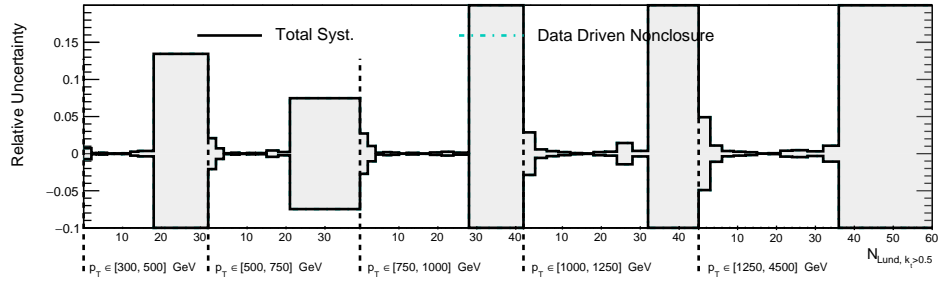


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

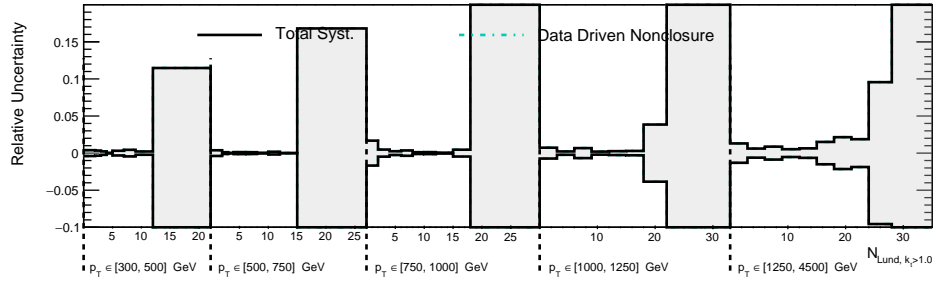


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

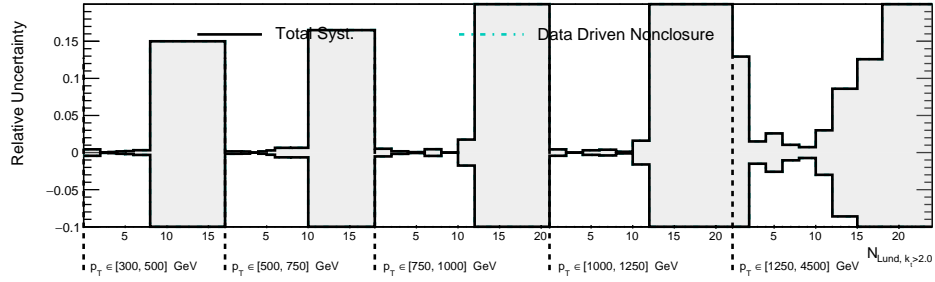
Figure 8.19: (continued) Data-driven non-closure diagnostic for primary Lund multiplicity, demonstrating the performance of the reweighting applied to the PYTHIA MC.



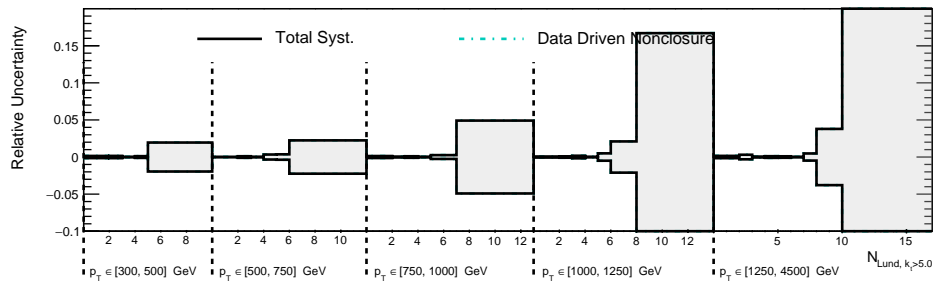
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

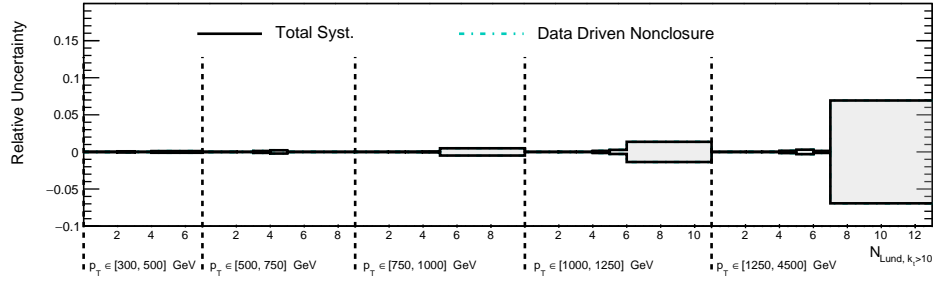


(c) $k_t > 2$ GeV, N_{Lund}

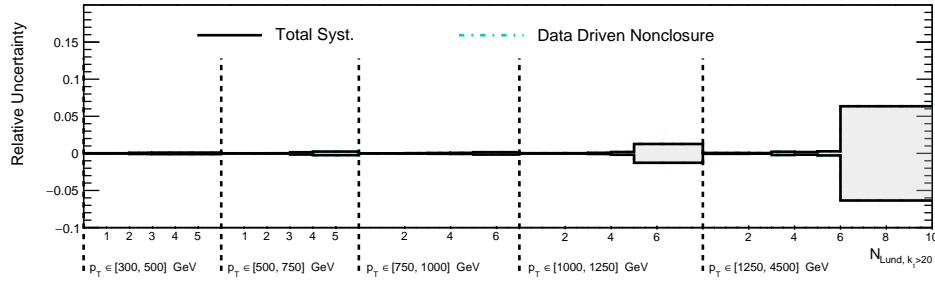


(d) $k_t > 5$ GeV, N_{Lund}

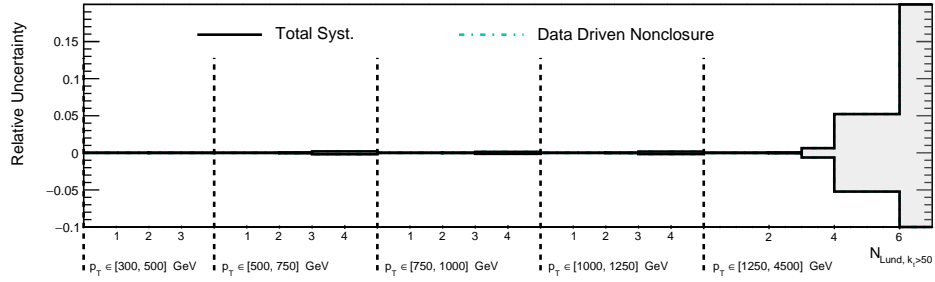
Figure 8.20: Data-driven non-closure uncertainty for Lund multiplicity. Figure continued on next page.



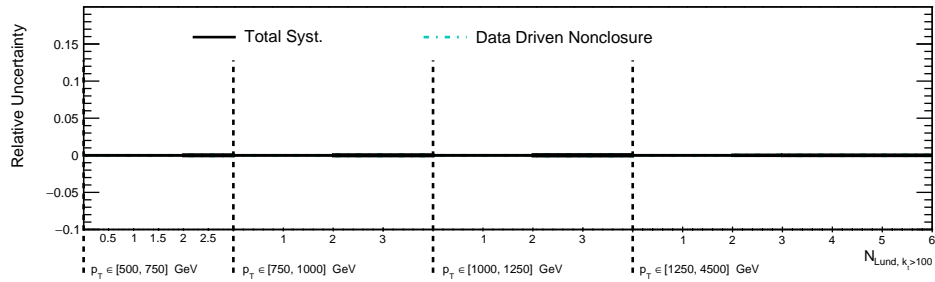
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

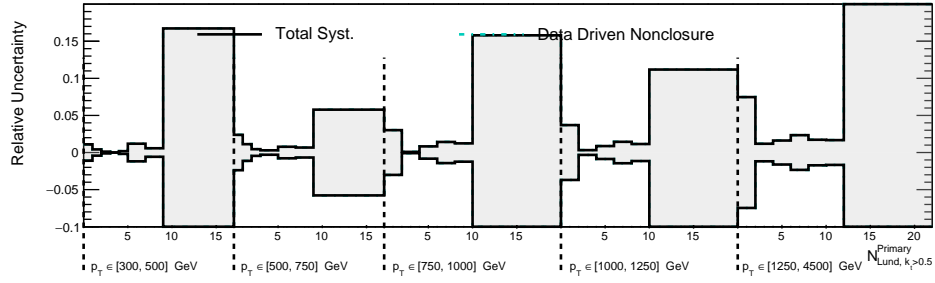


(g) $k_t > 50$ GeV, N_{Lund}

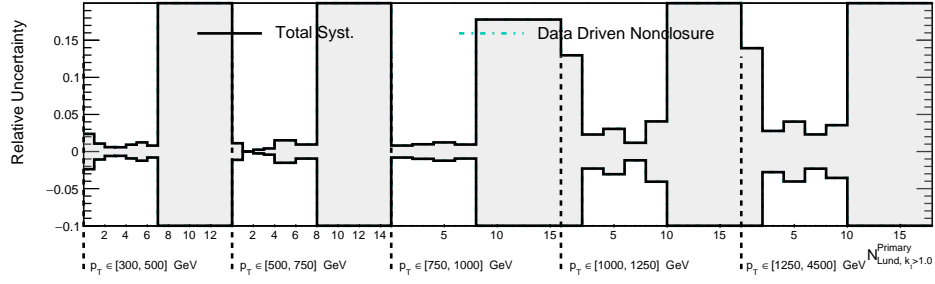


(h) $k_t > 100$ GeV, N_{Lund}

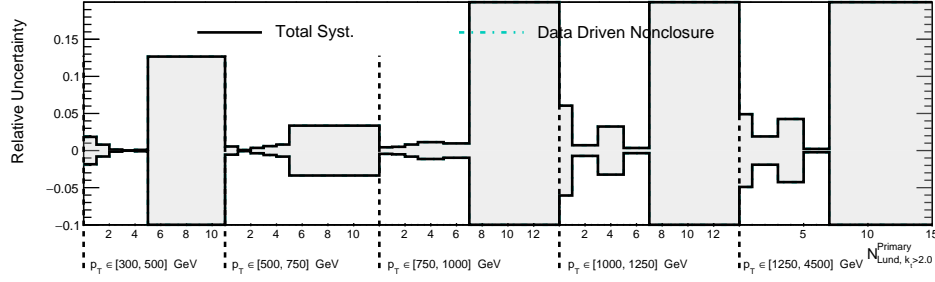
Figure 8.20: (continued) Data-driven non-closure uncertainty, for Lund multiplicity.



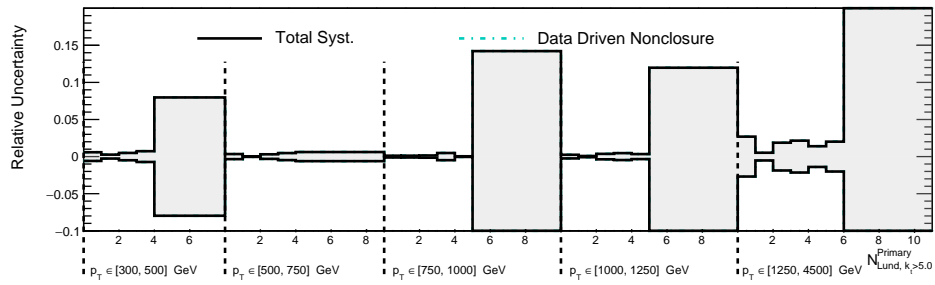
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

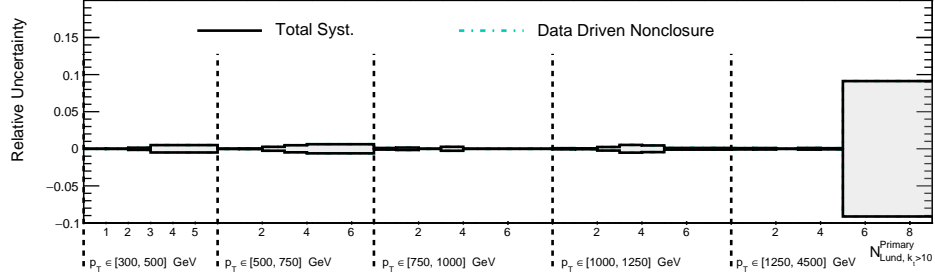


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

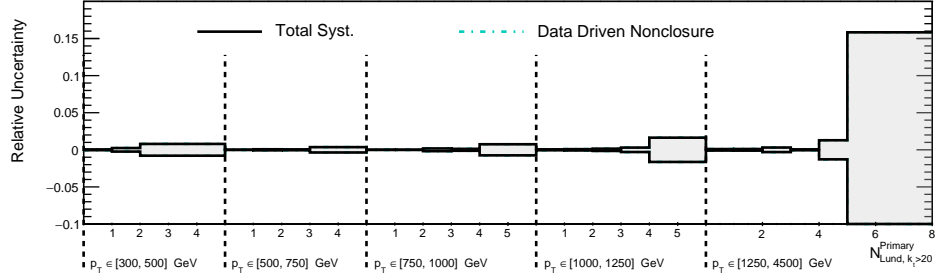


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

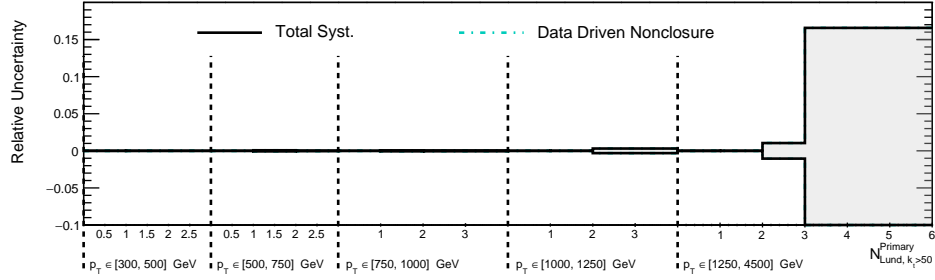
Figure 8.21: Data-driven non-closure uncertainty for primary Lund multiplicity. Figure continued on next page.



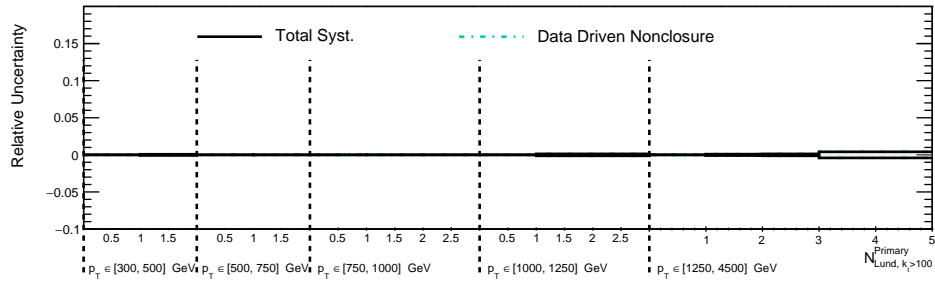
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.21: (continued) Data-driven non-closure uncertainty for primary Lund multiplicity.

8.7 Systematic Uncertainties

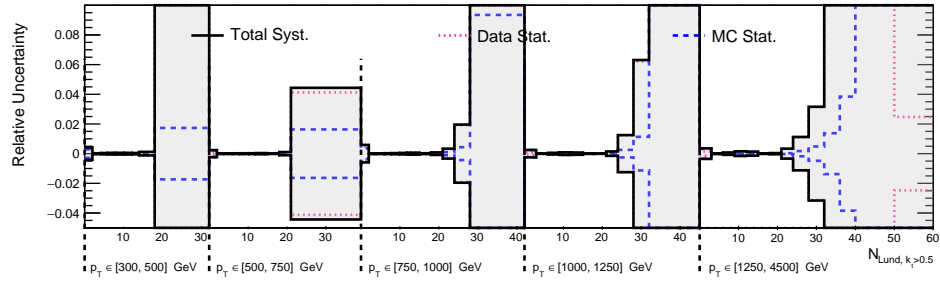
Unless otherwise indicated, systematic uncertainties are evaluated by replacing the nominal PYTHIA response matrix during the unfolding procedure with a varied version.

8.7.1 *Statistical uncertainties*

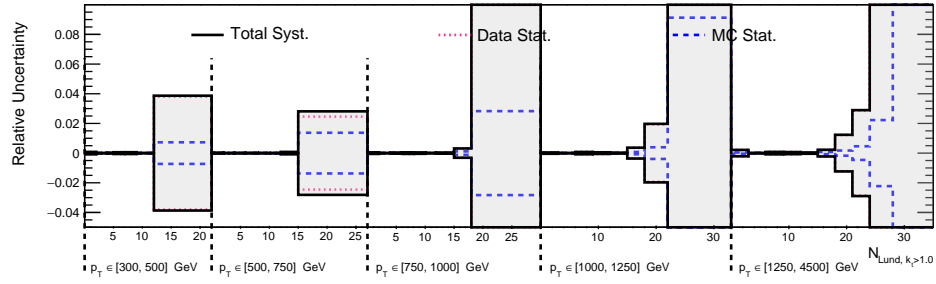
Statistical uncertainties related to the finite Monte Carlo and Data sample statistics used during the unfolding procedure for this measurement are estimated with pseudoexperiments using the bootstrapping method¹¹. Fifty pseudoexperiments are created using the nominal PYTHIA v8 sample used in the unfolding procedure. For the Monte Carlo, the response matrix, efficiency and acceptances are all allowed to vary during this process. For the data statistical uncertainty, the input data distributions to the unfolding procedure are varied. The 68% inter-quantile range (IQR) of the output distributions generated as a result of these variations is taken as the corresponding statistical uncertainty.

The MC and Data statistical uncertainties are provided in Figure 8.22 and Figure 8.23 for the complete set of measured distributions. These uncertainties are usually subdominant, but can increase in the tails of multiplicity distributions. The data statistical uncertainty is *almost always* the dominant source of statistical uncertainty in the measurement.

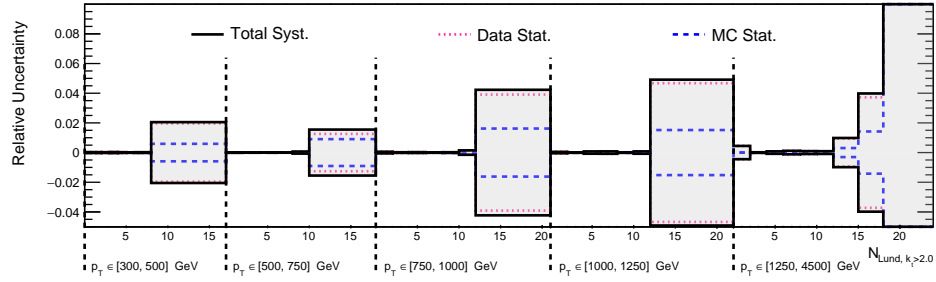
11. The implementation of this uncertainty follows the ATLAS SM Group recommendations: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/StandardModelUnfoldingNew>



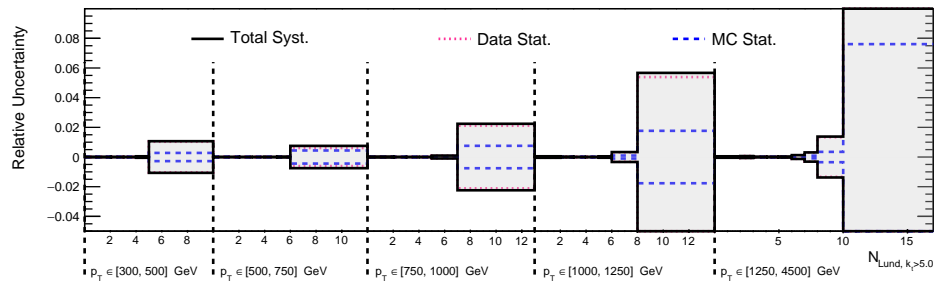
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

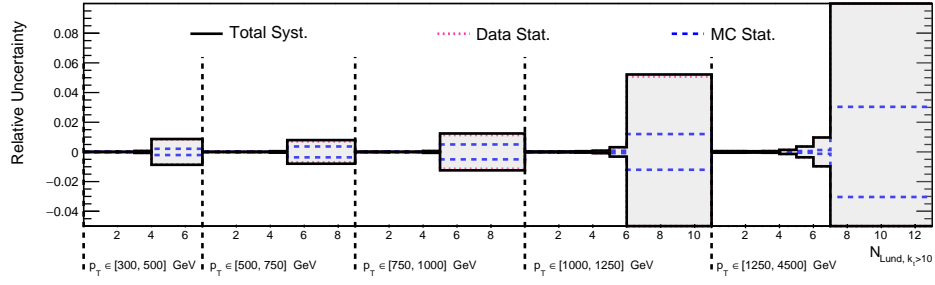


(c) $k_t > 2$ GeV, N_{Lund}

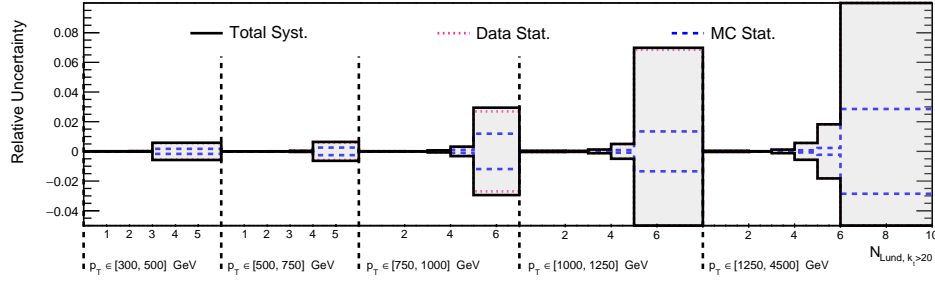


(d) $k_t > 5$ GeV, N_{Lund}

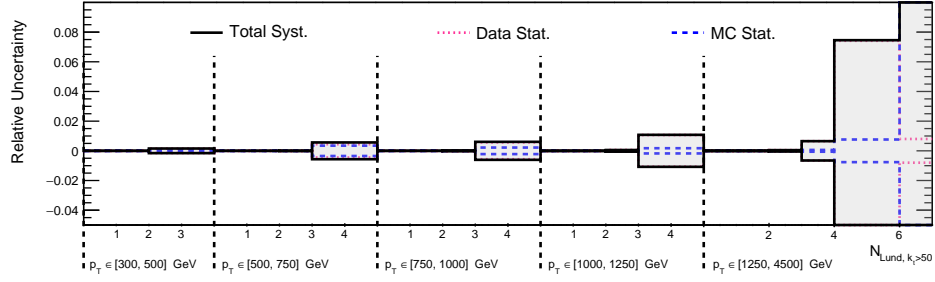
Figure 8.22: Statistical uncertainties from data & MC for Lund multiplicity. Figure continued on next page.



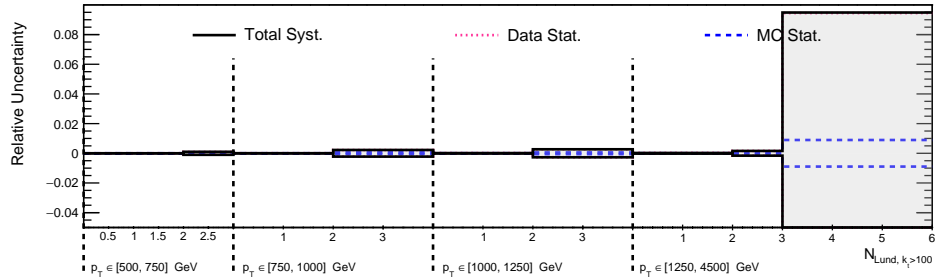
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

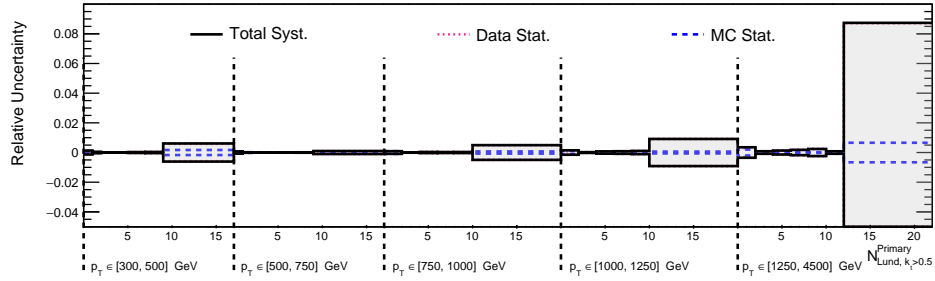


(g) $k_t > 50$ GeV, N_{Lund}

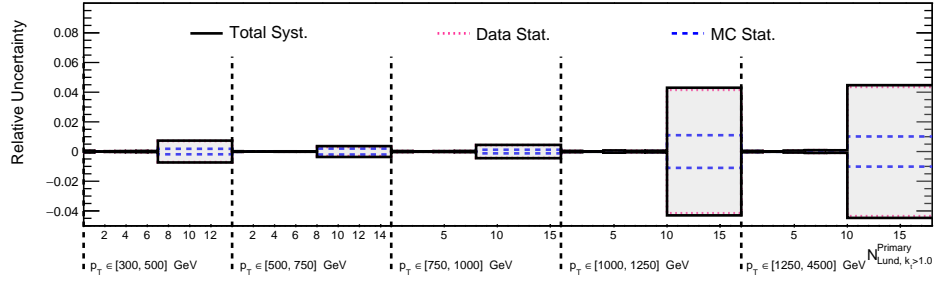


(h) $k_t > 100$ GeV, N_{Lund}

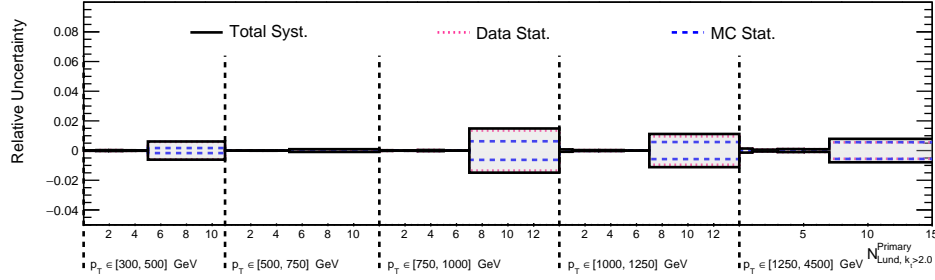
Figure 8.22: (continued) Statistical uncertainties from data & MC for Lund multiplicity.



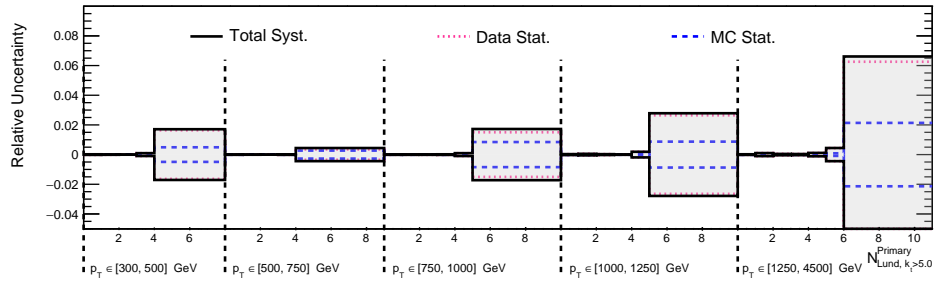
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

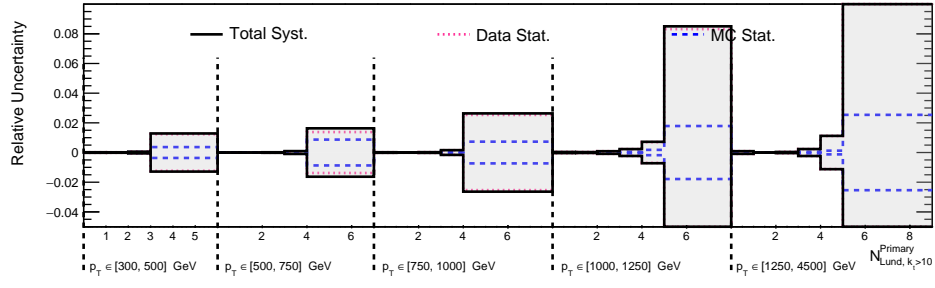


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

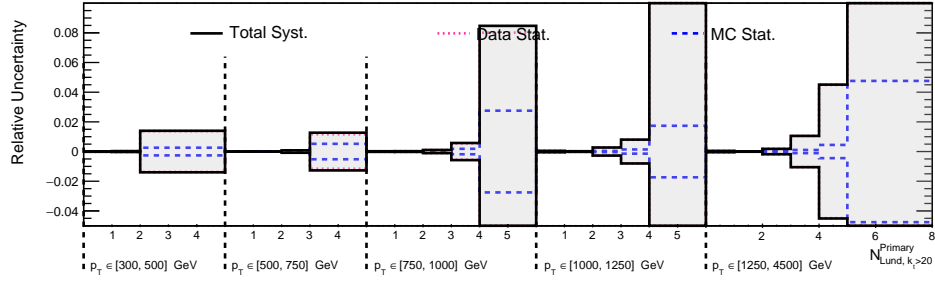


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

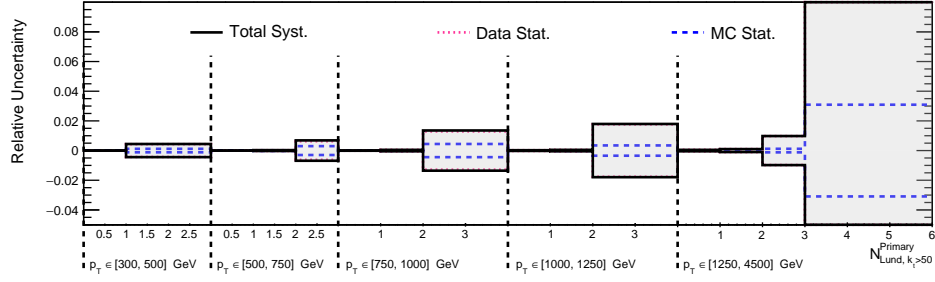
Figure 8.23: Statistical uncertainties from data & MC, for primary Lund multiplicity. Figure continued on next page.



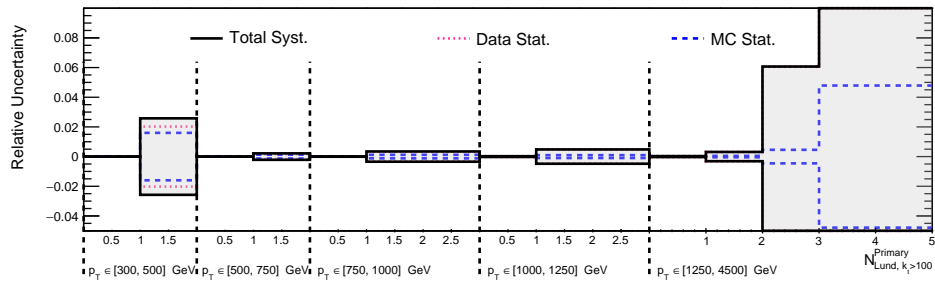
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



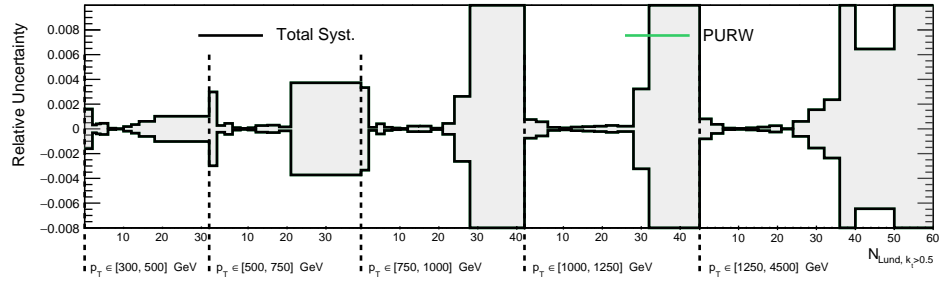
(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.23: (continued) Statistical uncertainties from data & MC, for primary Lund multiplicity.

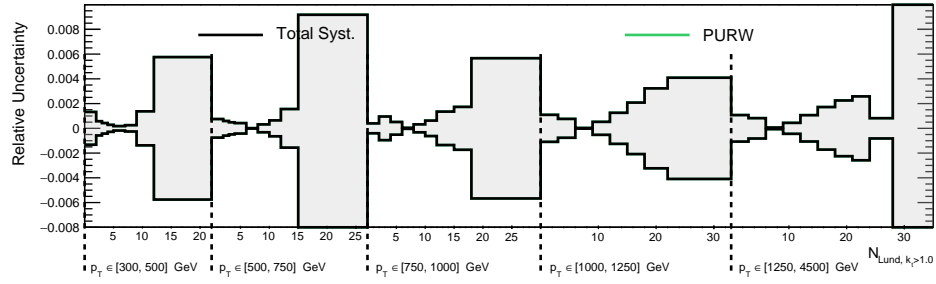
8.7.2 *Pile-up reweighting uncertainty*

The scale factor used in the pileup reweighting procedure (Section 8.4) is varied from 1.03 to 1.00 and 1.06, following ATLAS recommendations¹². This uncertainty is almost always negligible throughout the analysis, but it is nevertheless included in the final result. It is shown in Figure 8.24 and Figure 8.25 for the complete set of measured distributions.

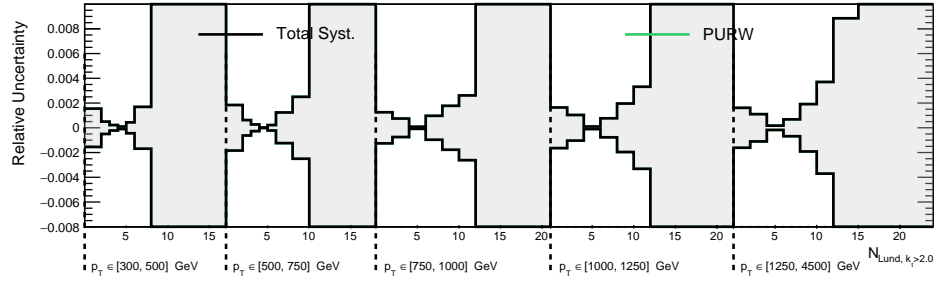
12. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ExtendedPileupReweighting>



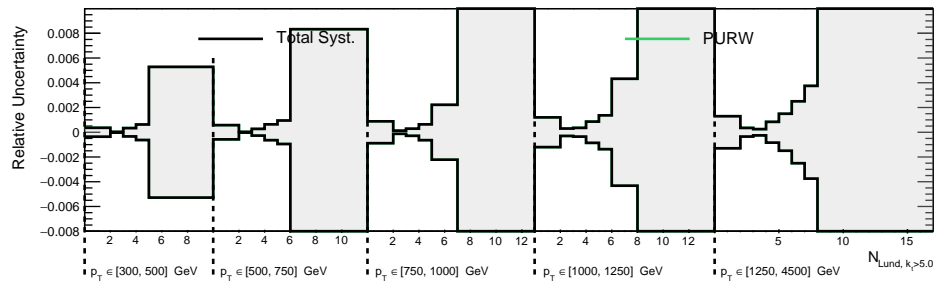
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

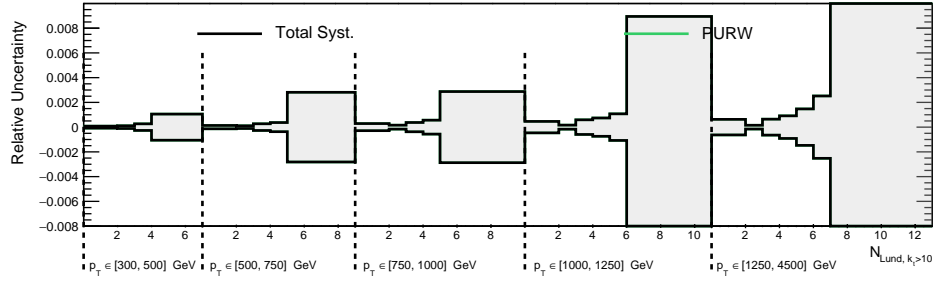


(c) $k_t > 2$ GeV, N_{Lund}

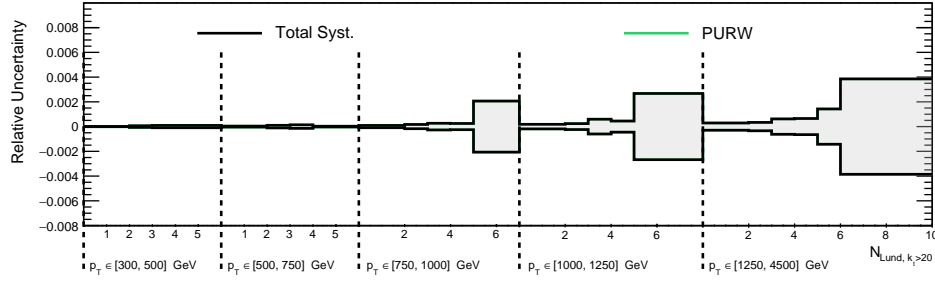


(d) $k_t > 5$ GeV, N_{Lund}

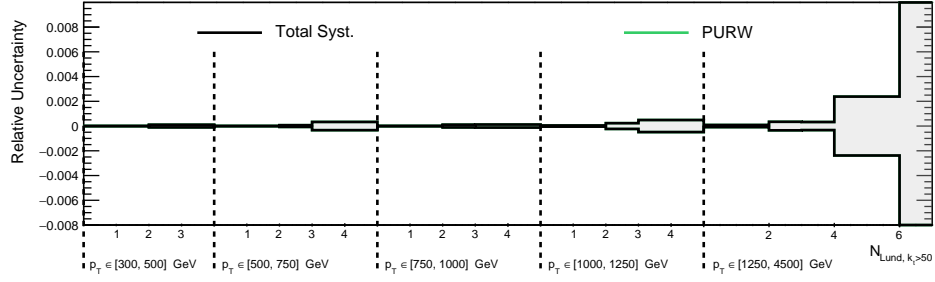
Figure 8.24: Pile-up reweighting uncertainties for Lund multiplicity. Figure continued on next page.



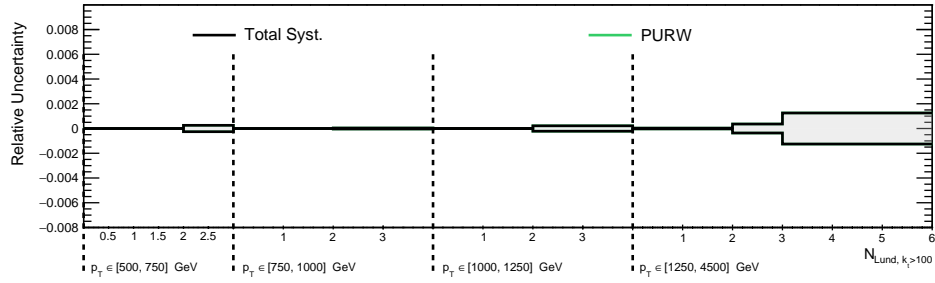
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

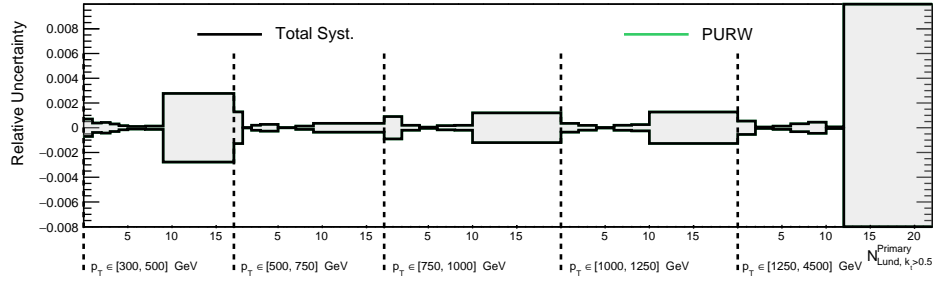


(g) $k_t > 50$ GeV, N_{Lund}

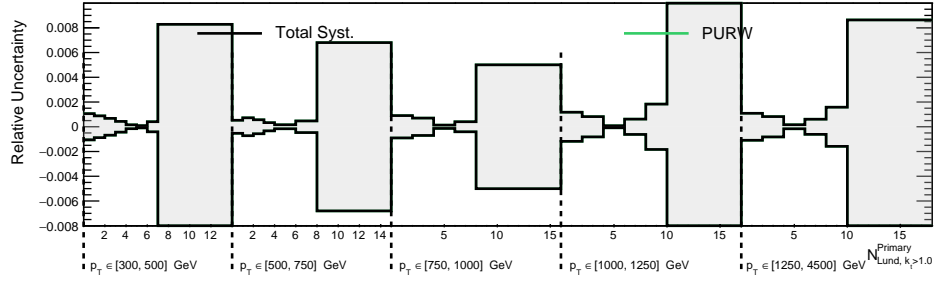


(h) $k_t > 100$ GeV, N_{Lund}

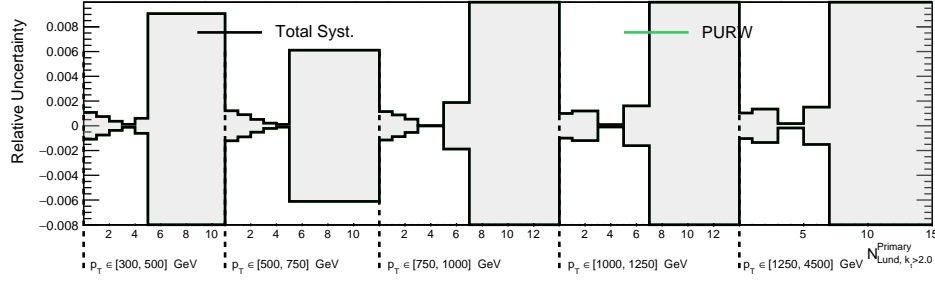
Figure 8.24: (continued) Pile-up reweighting uncertainties for Lund multiplicity.



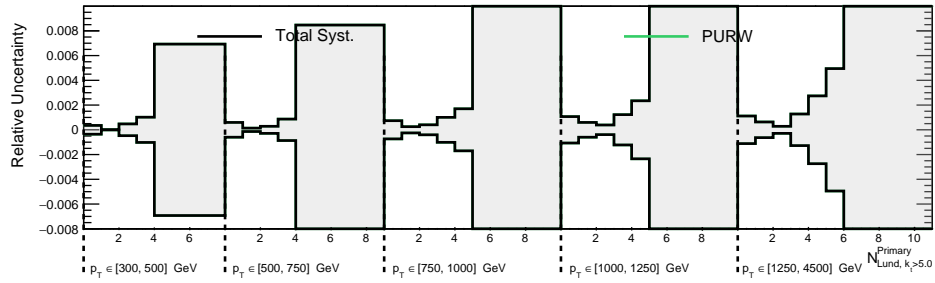
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

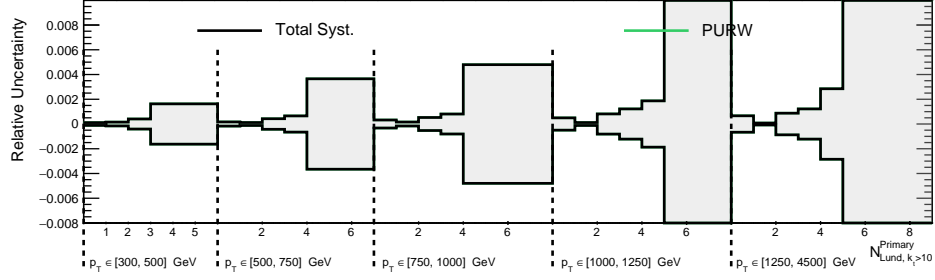


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

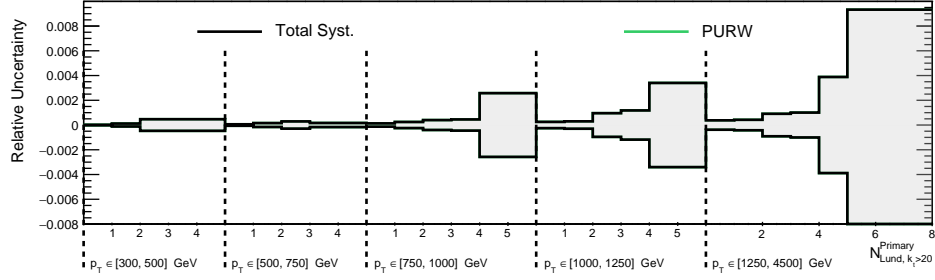


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

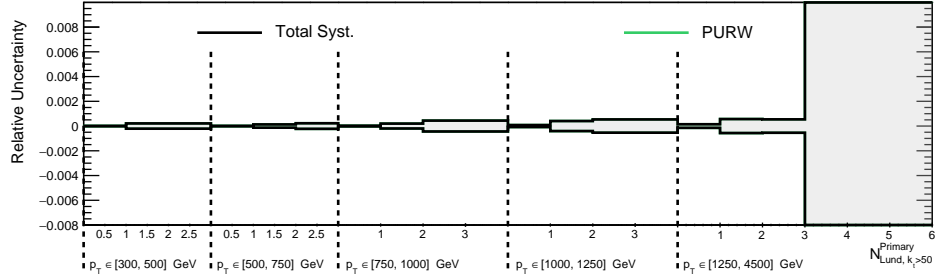
Figure 8.25: Pile-up reweighting uncertainties for primary Lund multiplicity. Figure continued on next page.



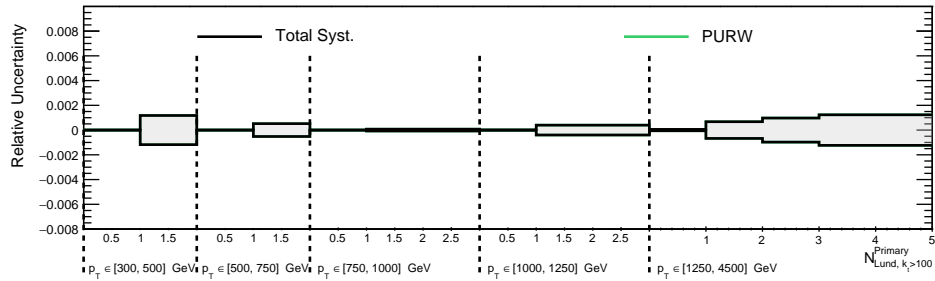
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.25: (continued) Pile-up reweighting uncertainties for primary Lund multiplicity.

8.7.3 Disabled TILE Calorimeter modules

During certain data-taking periods, specific modules of the tile calorimeter were disabled due to technical problems. Some such modules are included in the MC simulation which corresponds to a given data-taking period, while other modules exist only during specific runs in data. The tile modules which are disabled in Run 2 data and simulation are summarised in Table 8.3.

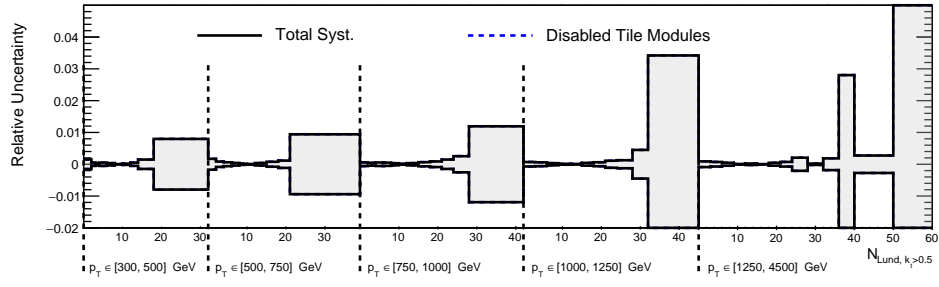
Table 8.3: Summary of disabled tile modules during Run 2 data-taking and in Run 2 simulation. Adapted from https://gitlab.cern.ch/atlas/athena/-/blob/master/Reconstruction/Jet/JetAnalysisTools/JetTileCorrection/data/Tile_maskedDB_Run2.conf.

Partition	Module	Start (Run)	End (Run)	Name
2015				
0	9	276262	284484	LBA10
3	20	276262	284484	EBC21
1	37	283155	283155	LBC38
2016				
1	4	306988	311481	LBC5
0	51	307856	311481	LBA52
2017				
2	2	325713	340453	EBA03
1	62	325713	340453	LBC63
2018				
0	28	351310	353395	LBA29
0	29	350361	353395	LBA30
0	31	353395	364292	LBA32
MC16a				
0	51	284500	284501	LBA52
1	4	284500	284501	LBC5
MC16d				
2	2	300000	300001	EBA3
1	62	300000	300001	LBC63
MC16e				
0	31	310000	310001	LBA32

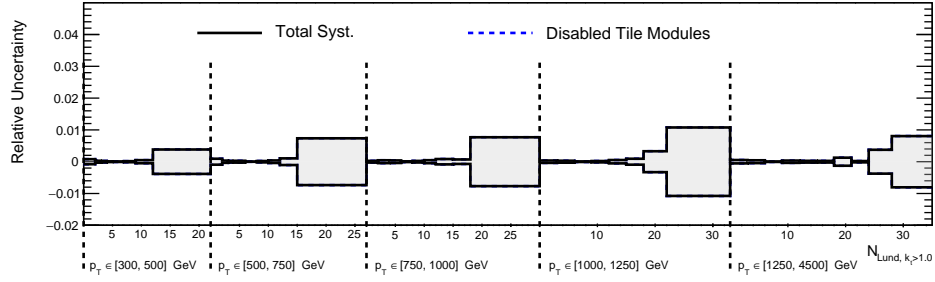
Following ATLAS recommendations, no correction is applied to the p_T of jets which may have deposited energy in disabled tile modules. The impact of disabled tile modules on the

event shape distributions studied in this analysis is evaluated by repeating the measurement and vetoing events with jets directed at disabled modules in both data and PYTHIA. This means that the data with the veto is unfolded using the PYTHIA response matrix constructed with the veto. The difference between this result and the nominal result will be taken as an uncertainty on the final measurement.

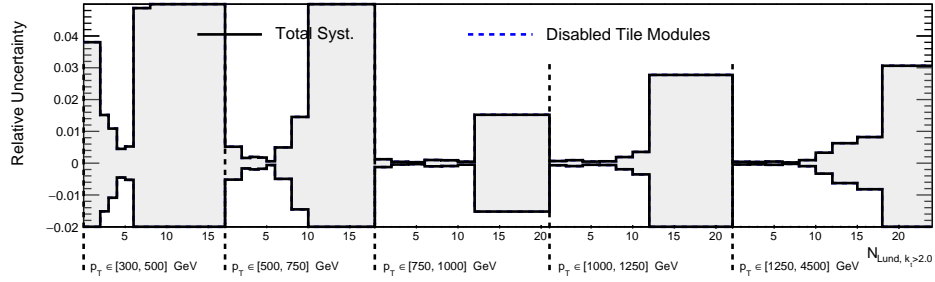
The size of the resultant uncertainty is shown in Figure 8.26 and Figure 8.27 for the complete set of measured distributions. In the bulk of the distributions measured in this analysis, this uncertainty is negligible.



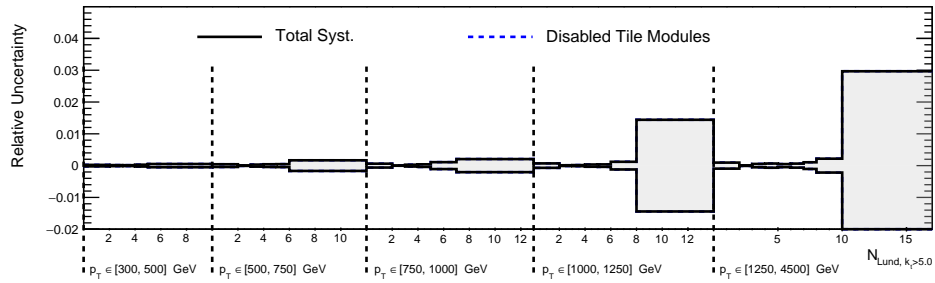
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

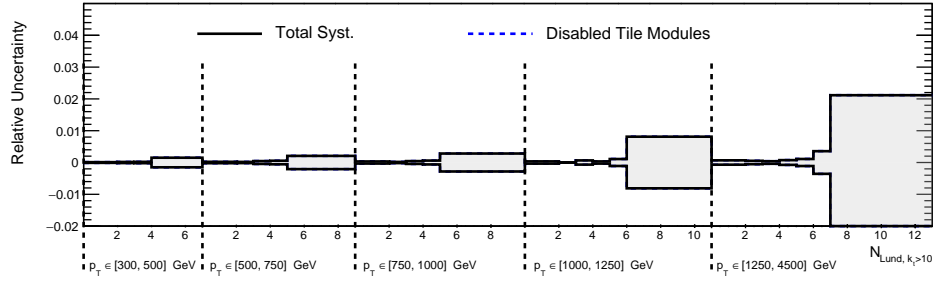


(c) $k_t > 2$ GeV, N_{Lund}

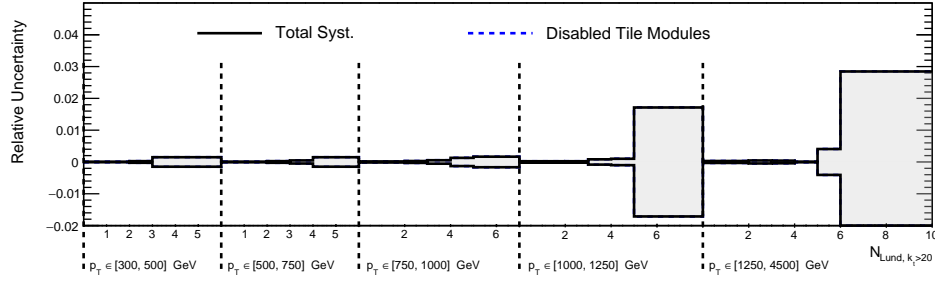


(d) $k_t > 5$ GeV, N_{Lund}

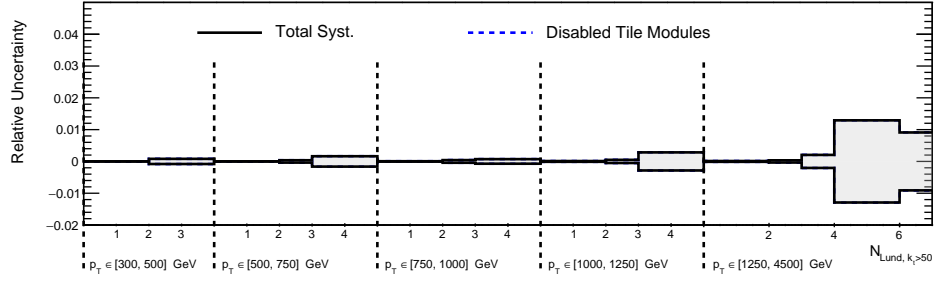
Figure 8.26: Disabled tile-module uncertainty, for Lund multiplicity. Figure continued on next page.



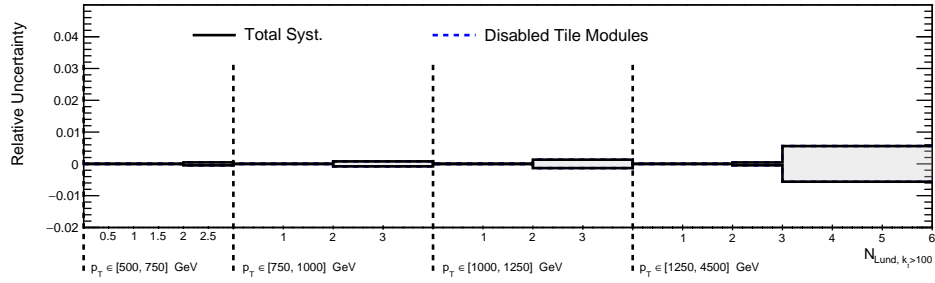
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

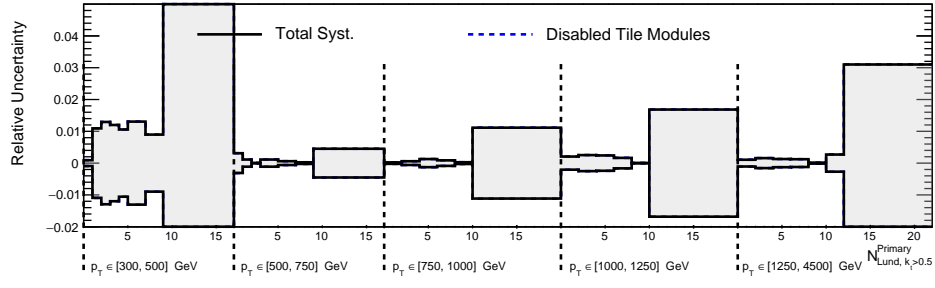


(g) $k_t > 50$ GeV, N_{Lund}

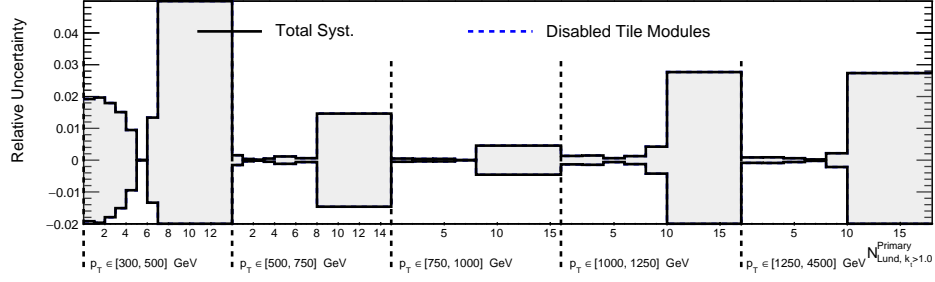


(h) $k_t > 100$ GeV, N_{Lund}

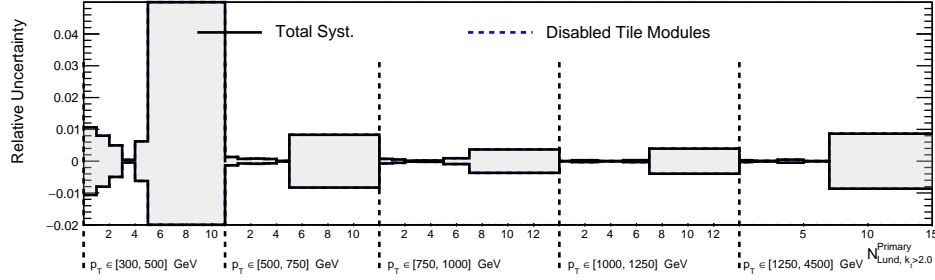
Figure 8.26: (continued) Disabled tile-module uncertainty, for Lund multiplicity.



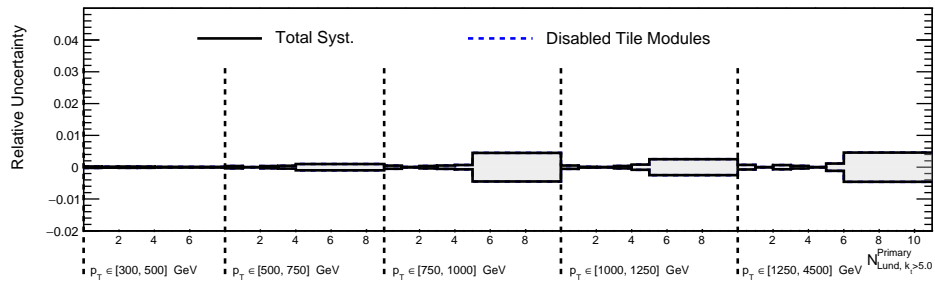
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

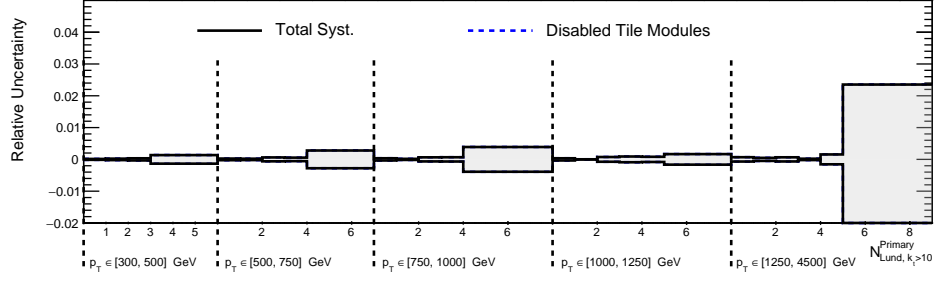


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

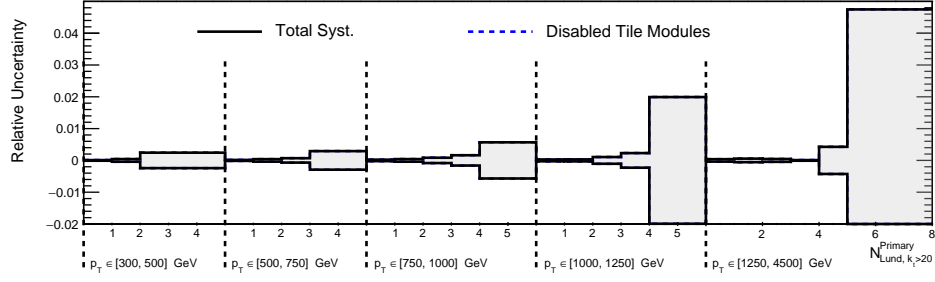


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

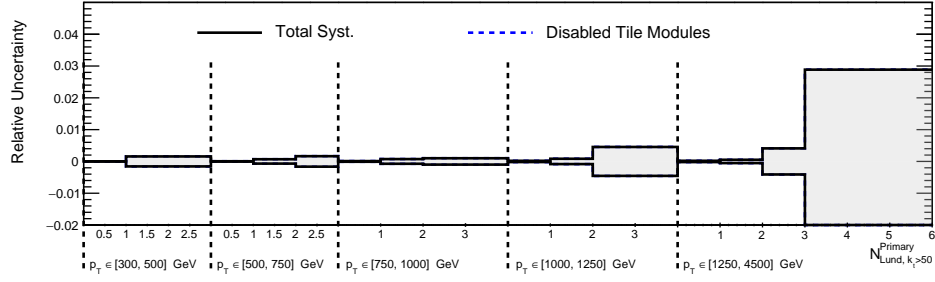
Figure 8.27: Disabled tile-module uncertainty for primary Lund multiplicity. Figure continued on next page.



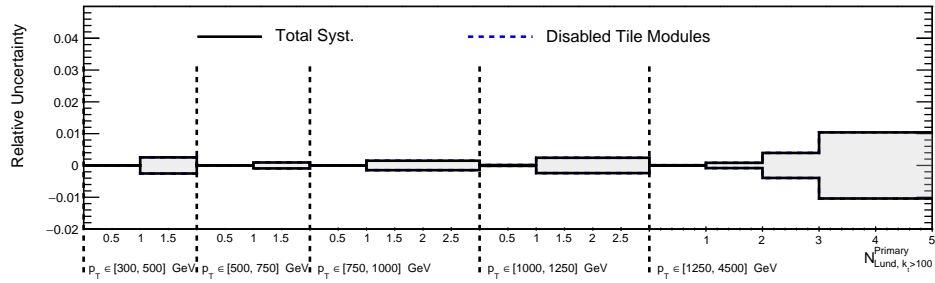
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.27: (continued) Disabled tile-module uncertainty for primary Lund multiplicity.

8.7.4 Jet energy scale and resolution (*JES/JER*) uncertainties

Jets in ATLAS are complex objects which are formulated using complex algorithms like the ATLAS particle flow reconstruction algorithm. The calibration procedure is an important part of any reconstruction process and specific recommendations for this calibration and its associated uncertainty are implemented by the ATLAS Jet/ $E_{\text{T}}^{\text{miss}}$ group. The JES and JER are calibrated using a series of *in situ* measurements and simulation-based techniques. I will summarize these below, but a thorough description can be found in Ref. [23].

The main sources of JES uncertainties are:

dijet eta-intercalibration : calibration of the scale of forward jets with respect to central jets, depends on p_{T} and η and is mostly relevant for forward jets

Z+jet balance : calibration of the scale of jets with respect to a well understood object like a Z boson decaying leptonically, $Z \rightarrow ee, Z \rightarrow \mu\mu$. Depends on p_{T} and is mostly relevant for low p_{T} jets below 100 GeV.

γ +jet balance : calibration of the scale of jets with respect to a well understood object like a photon (γ). Depends on p_{T} and is mostly relevant for medium p_{T} jets with a p_{T} between 100 GeV and 1 TeV.

multi-jet balance : calibration of the scale of jets with respect to a well understood object such as a recoil system of lower p_{T} , previously calibrated jets. Depends on p_{T} and is mostly relevant for very energetic jets with a p_{T} from 1 TeV to 2 TeV.

propagation of single particle and test beam uncertainties : Depends on p_{T} , mostly relevant for jets with a p_{T} above 2 TeV.

pile-up uncertainties : Depends on η , p_{T} , the number of primary vertices, and potentially

other variable depending on the exact uncertainty component. Mostly relevant for very low p_T jets below 50 GeV.

flavor-related uncertainties : highly dependent on tagging requirements and multiplicity-flavor relationships, such as if b-jets are being identified, if flavor composition is specified, if the flavor composition depends on jet multiplicity, and if the jet multiplicity flavor composition is specified by process. Depends on p_T and η and is potentially relevant for low to medium p_T jets with up to several hundred GeV of p_T .

punch-through uncertainties : depends on p_T , η , and the number of muon segments behind the jet. Relevant for analyses which are sensitive to muon spectrometer activity close to jets, but otherwise only relevant in the high p_T regime above 1 TeV.

MC non-closure uncertainty : depends on p_T and η . This is zero if you are using ATLAS full simulation, and non zero with fast simulation. This analysis uses only full simulation, so these uncertainties are not used.

The JER also has a number of sources of uncertainty which are summarized below:

data vs MC differences : accounts for the differences between the jet energy resolution of the nominal MC and the data. If the MC has a better resolution than the data, the MC is smeared to match the data resolution. If the data has a better resolution than the MC, it is not smeared since this would degrade the data. In this case, there is a non-zero uncertainty since the nominal data resolution is superior to that of the MC. This uncertainty is different depending on the type of ATLAS simulation used (fast or full simulation), and depends on the jet p_T and η .

noise uncertainties : the noise component of the uncertainty is evaluated from random cones in zero bias data. This depends on p_T and η , and is mostly relevant for low p_T jets below 100 GeV.

dijet p_T balance asymmetry : depends on p_T and η and is relevant for all jets starting at a few tens of GeV.

These systematic uncertainties on the $R = 0.4$ JES and JER are evaluated following the ‘consolidated recommendations’¹³ developed in late 2018 by the Jet/ E_T^{miss} group¹⁴ (summarised in Figure 8.28 and Figure 8.29). These uncertainties were later updated in two ways: the first was an improvement to the flavour response uncertainty by using updated MC comparisons to derive it in a factorised way. The second was a reduction of the single-particle deconvolution uncertainty using new in situ measurements from $W \rightarrow \tau\nu$ decays [28]. The **Category Reduction JES** and **Full JER** schemes for $R = 0.4$ EMPFlow jets are used. The resultant JES and JER nuisance parameters are treated as uncorrelated variations, following the standard ATLAS approach. This analysis uses the updated jet flavour response and single-particle uncertainty treatments described above, which can also be found on the Jet/ E_T^{miss} twiki¹⁵.

The JES/JER uncertainties are the second largest uncertainty in many bins of this measurement. Their impact is detailed in Figure 8.30 and Figure 8.31 for the complete set of measured distributions. The distinct sources of JES and JER uncertainties are broken down. The largest single source of uncertainty is related to the effective JES and JER nuisance parameters.

13. JES/JER Recommendations documentation

14. JES/JER twiki

15. Reduced flavor uncertainties in JES documentation

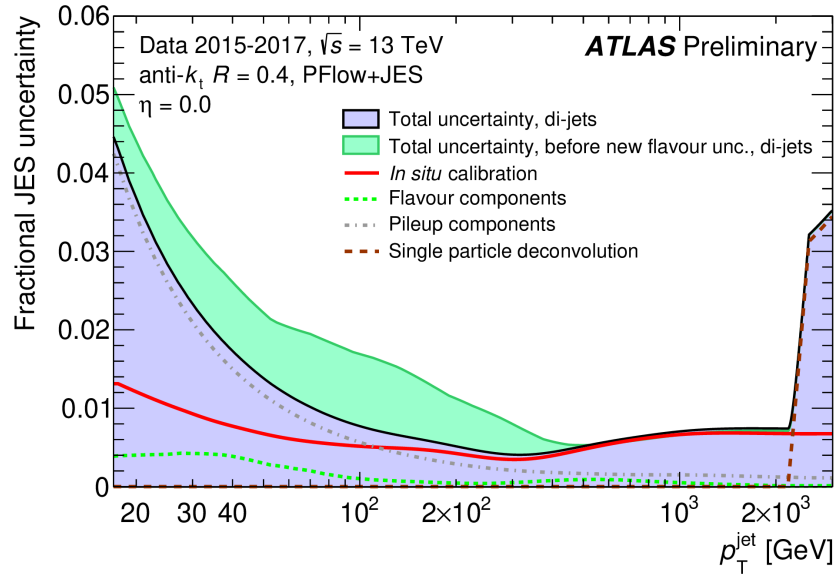


Figure 8.28: Fractional jet energy scale systematic uncertainty components for anti- k_t $R = 0.4$ jets as a function of jet p_T at $\eta = 0$, reconstructed from particle-flow objects. The total uncertainty, determined as the quadrature sum of all components, is shown as a filled region topped by a solid black line. Figure from public Jet/EtMiss plots: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2022-005/>

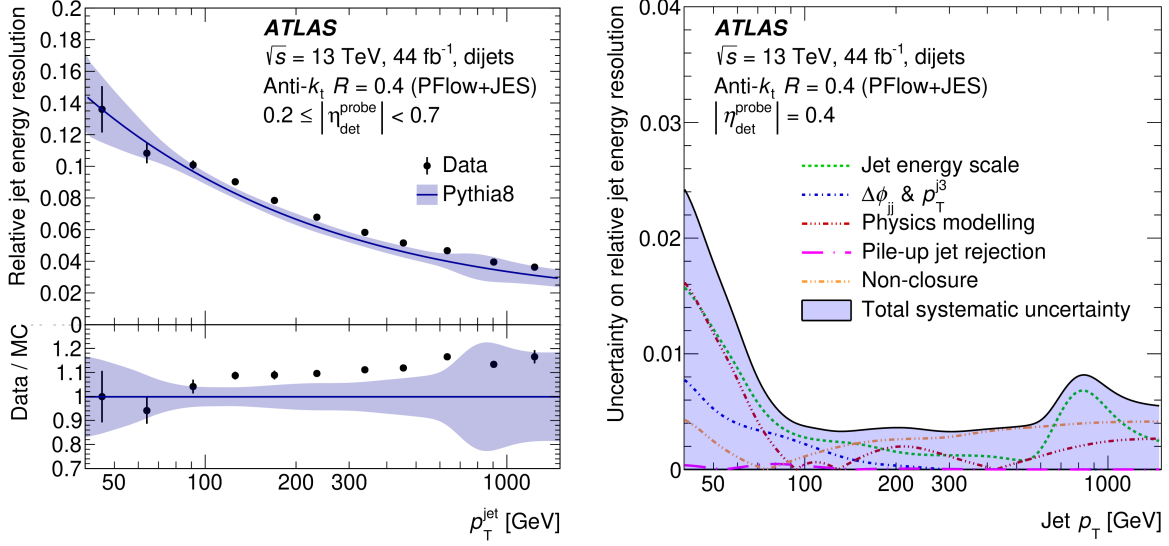
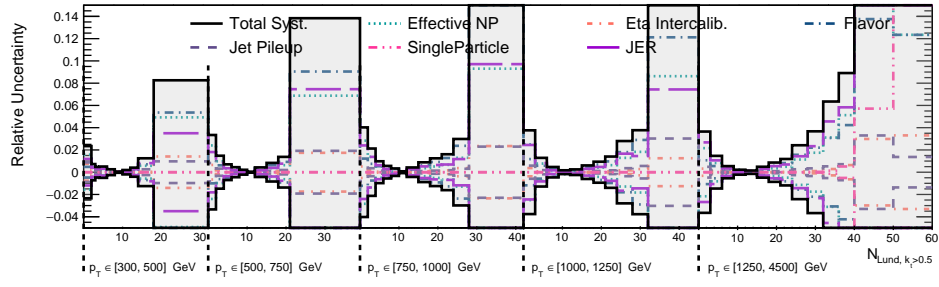
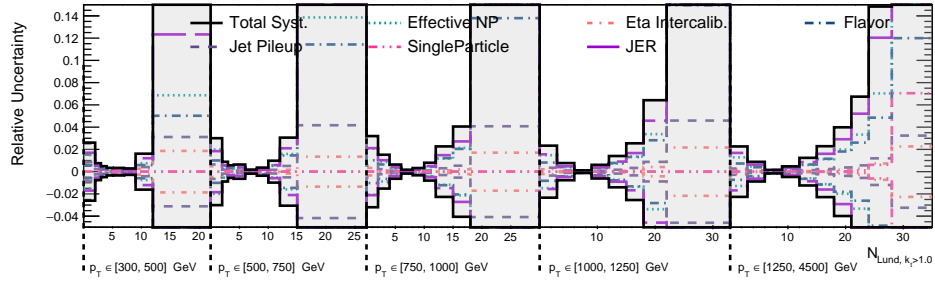


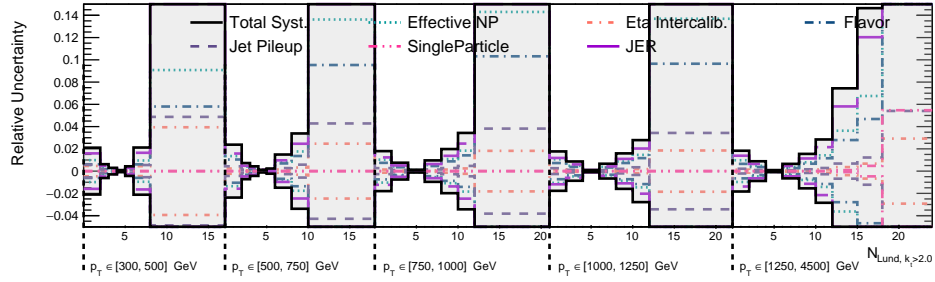
Figure 8.29: (a) Relative jet energy resolution and (b) absolute uncertainty in the relative resolution as a function of p_T for PFlow jets in the central region of the detector, measured using the dijet balance method. The resolution in data is shown in black points with error bars indicating statistical uncertainties; the resolution in detector-level simulated events is shown by the blue curve with total systematic uncertainty given by the blue band. The systematic uncertainty is dominated by terms propagated from the JES uncertainty, while additional terms arise from the analysis selection, pile-up rejection (JVT), physics modelling (comparison with alternative generator), and non-closure effects. The bump in uncertainty around 800 GeV comes from the single-particle uncertainty in the earlier in situ jet-energy scale calibration used when making the in situ JER measurement [23].



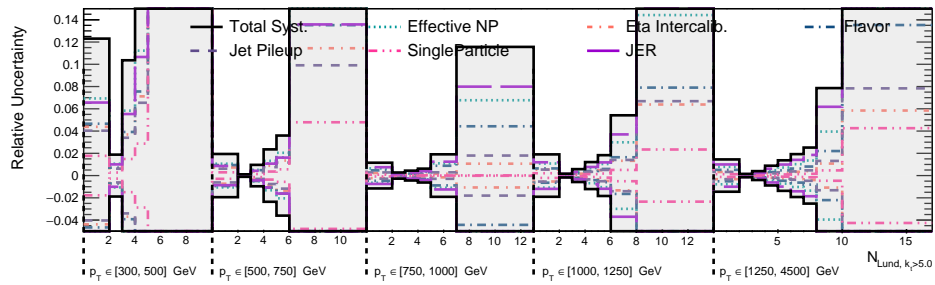
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

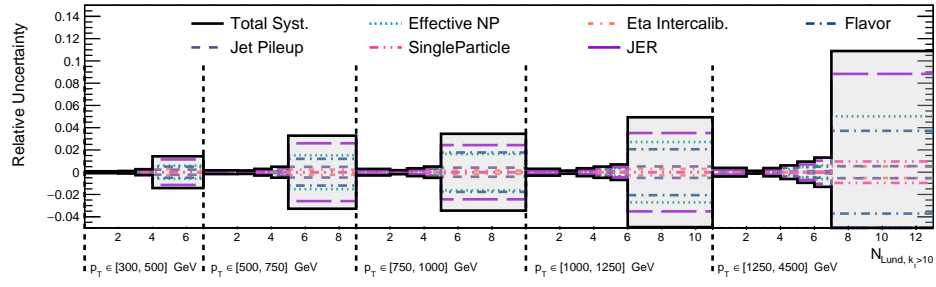


(c) $k_t > 2$ GeV, N_{Lund}

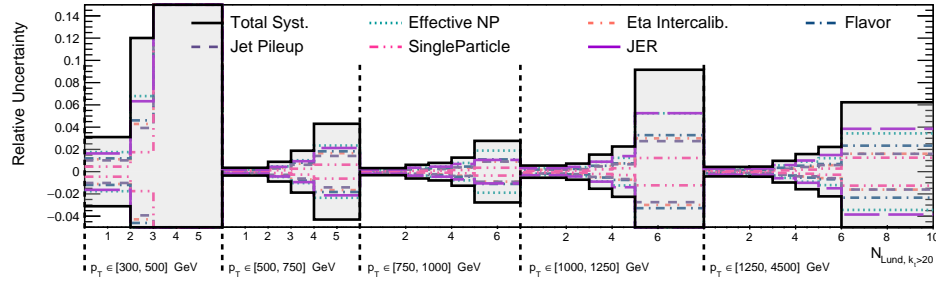


(d) $k_t > 5$ GeV, N_{Lund}

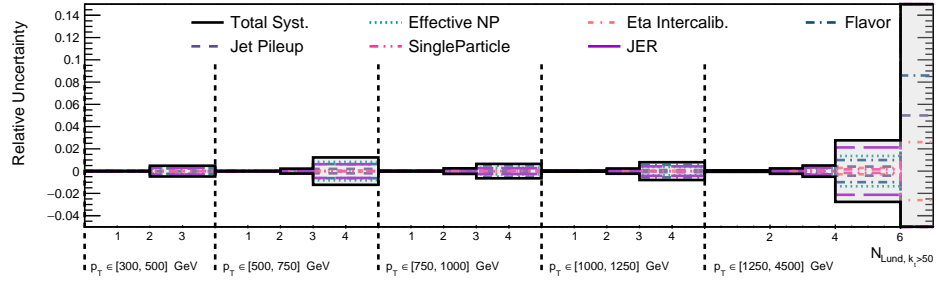
Figure 8.30: Uncertainties related to the jet energy scale and resolution (JES/JER) for the Lund multiplicity with various k_t cuts. Figure continued on next page.



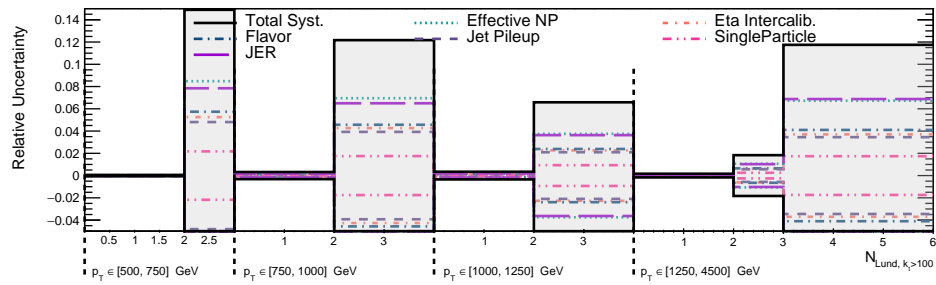
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

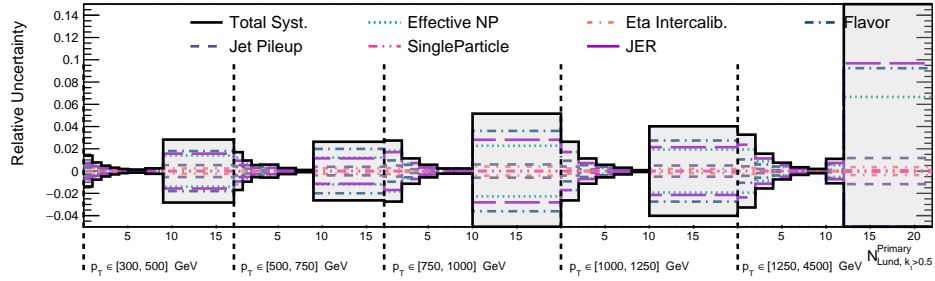


(g) $k_t > 50$ GeV, N_{Lund}

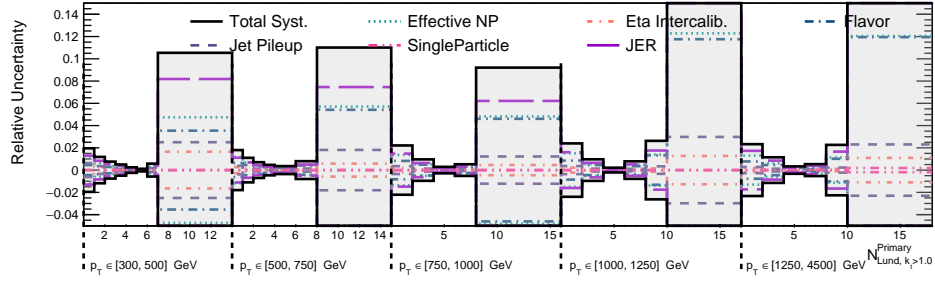


(h) $k_t > 100$ GeV, N_{Lund}

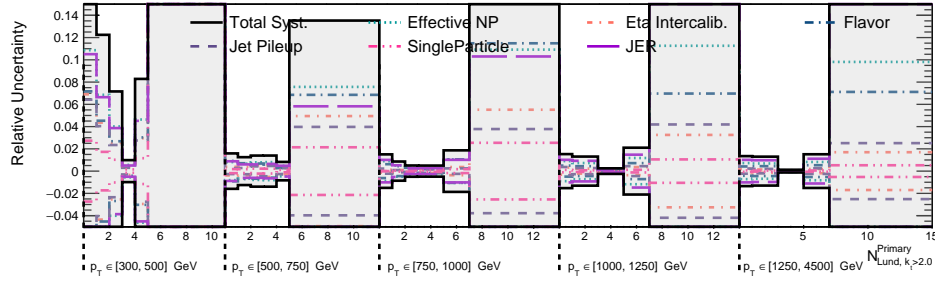
Figure 8.30: (continued) Uncertainties related to the jet energy scale and resolution (JES/JER) for the Lund multiplicity with various k_t cuts.



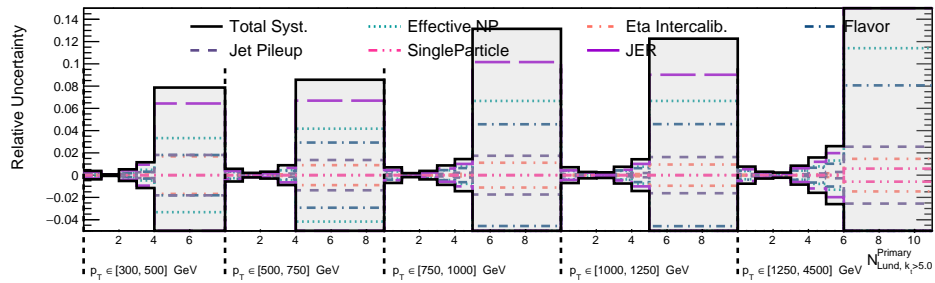
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

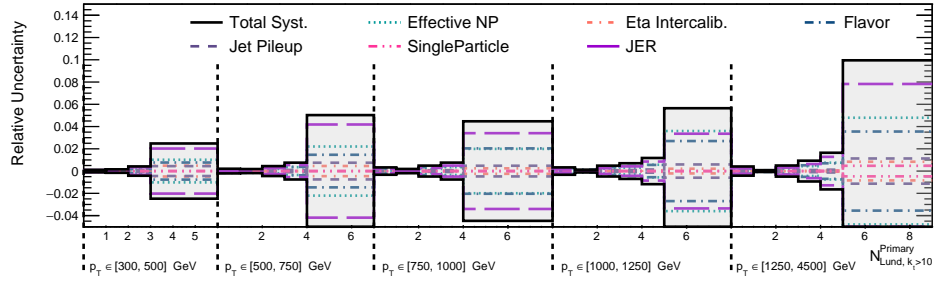


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

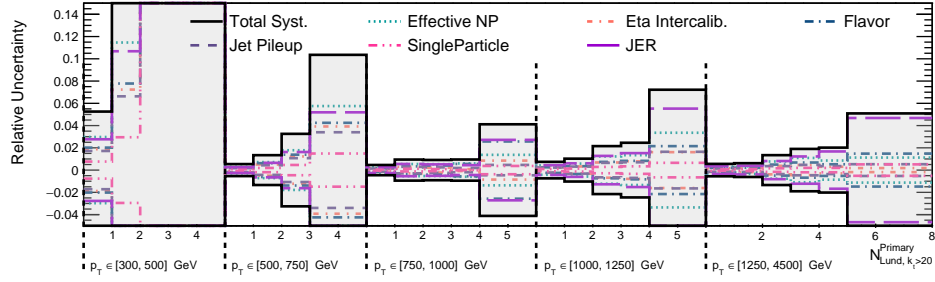


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

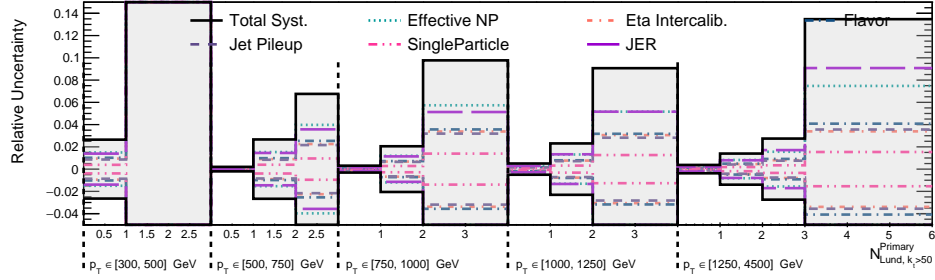
Figure 8.31: Uncertainties related to the jet energy scale and resolution (JES/JER) for the primary Lund multiplicity with various k_t cuts. Figure continued on next page.



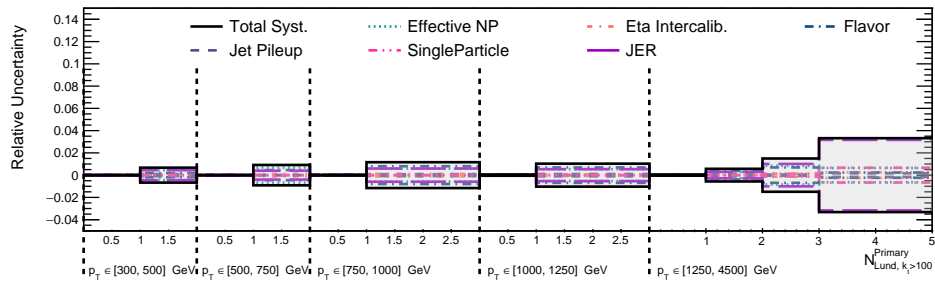
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.31: (continued) Uncertainties related to the jet energy scale and resolution (JES/JER) for the primary Lund multiplicity with various k_t cuts.

8.7.5 *Tracking-related uncertainties*

Systematic uncertainties relevant to the measurement and reconstruction of tracks from the inner detector are considered in this analysis. Tracks are the objects utilized as emissions within the C/A clustering history, and as such these uncertainties are important to ensure that track-based biases do not significantly impact the analysis.

The tracking uncertainties are evaluated by computing their impact on the complete container of tracks, and repeating the track-to-jet association and then calculation of the subjet multiplicities. This is performed following the ATLAS Tracking group’s recommendations for Full Run 2 analyses¹⁶.

The track-related sources of uncertainty are:

inclusive tracking efficiencies : due mainly to uncertainties related to material interactions, and the physics model used in simulation. Their effect is studied by comparing the efficiencies utilizing samples with different material configurations, and changes in models. The overall systematic uncertainties range from 0.5% for $|\eta| < 0.1$, and 1.8% $2.3 < |\eta| < 2.5$.

efficiency inside jets : measured by the ATLAS Tracking in Dense Environments (CTIDE) group. In high- p_T jets the local track density can be very high, which can negatively affect track energy resolution. Additional studies on pixel clustering and dedicated tuning are required to mitigate these effects [40, 31, 32].

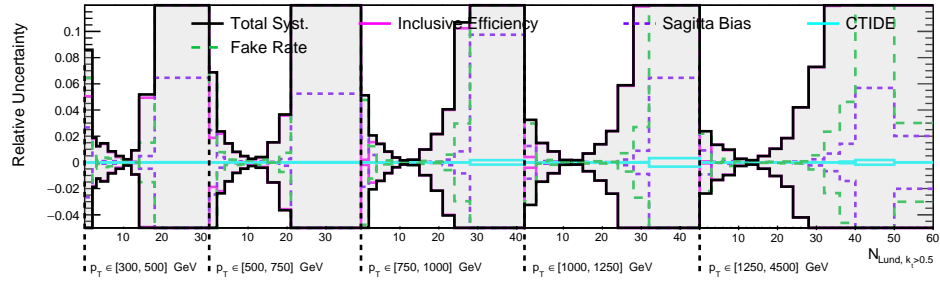
reconstruction fake rate : calculated by defining a fake-enriched control region based on MC information, and performing a fit to extract an estimate of the real and fake track fractions. This can be compared between data and MC [32].

16. Tracking group recommendation documentation

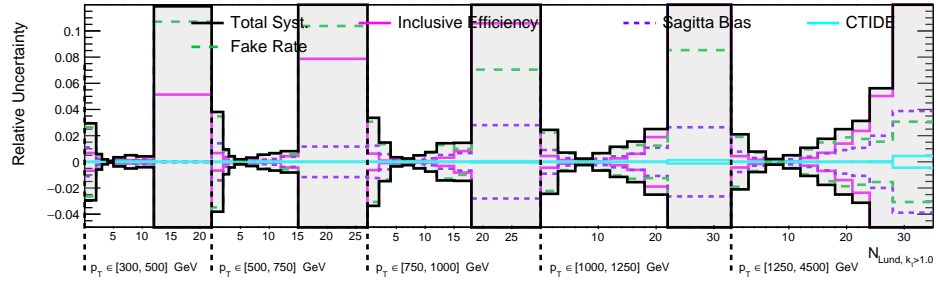
sagitta bias : slight mis-alignment of the ATLAS inner detector can cause sagitta biases, effects that impact oppositely charged particles in different ways. These effects can also differ in different environmental conditions (like temperature, which varies across runs). A small correction is applied to account for this bias, derived from the measurement of isolated electrons [11]. This is a small source of uncertainty, and is implemented as a smearing of the track p_T .

The uncertainties due to inclusive tracking efficiency, efficiency inside jets, and fake rate of track reconstruction are all implemented as a parameterised tracking inefficiency, where the tracks inside of jets are dropped depending on their parameters and environment using the tool created by the ATLAS Tracking group to propagate these uncertainties. The correction from the sagitta bias is implemented as a smearing of the track p_T .

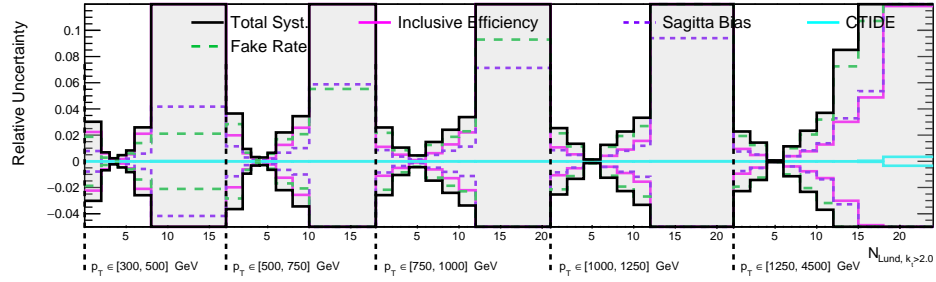
The impact of each of these tracking-related variations is shown in Figure 8.32 for the complete set of measured distributions. The overall impact of these tracking-related uncertainties is typically less than 2%, although they are occasionally larger in specific bins of the analysis. In general, the uncertainties from the inclusive efficiency are dominant, with subdominant contributions from the fake rate, and negligible uncertainties from CTIDE and the sagitta bias.



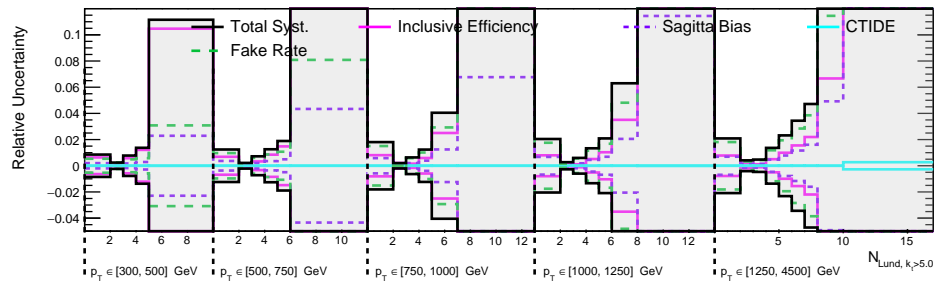
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

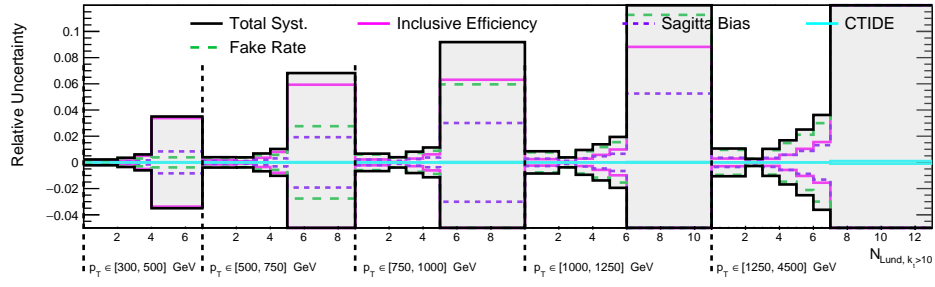


(c) $k_t > 2$ GeV, N_{Lund}

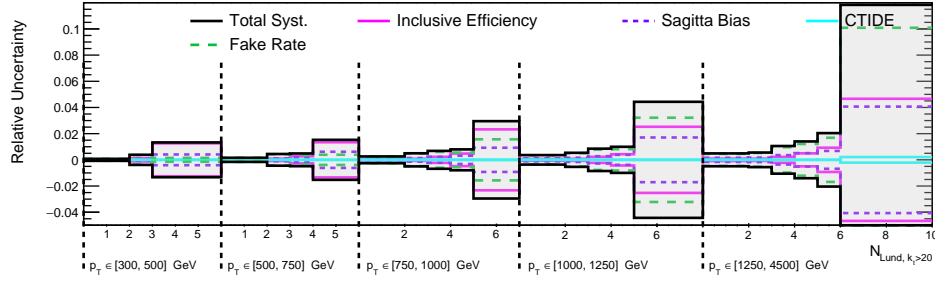


(d) $k_t > 5$ GeV, N_{Lund}

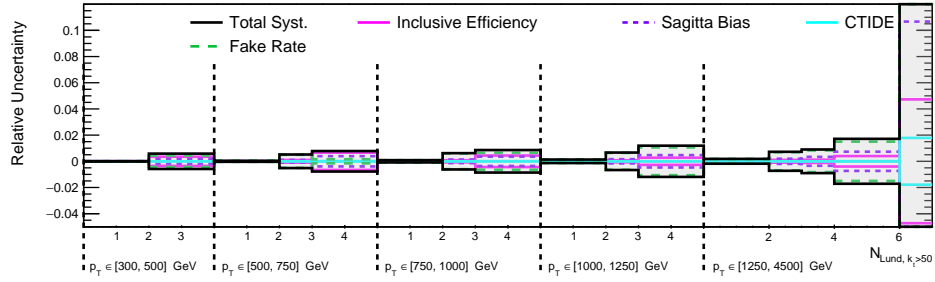
Figure 8.32: Tracking uncertainties for Lund multiplicity for several k_t requirements. Figure continued on next page.



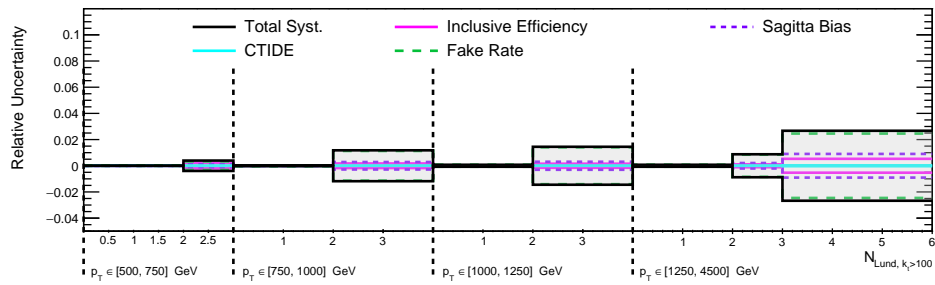
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

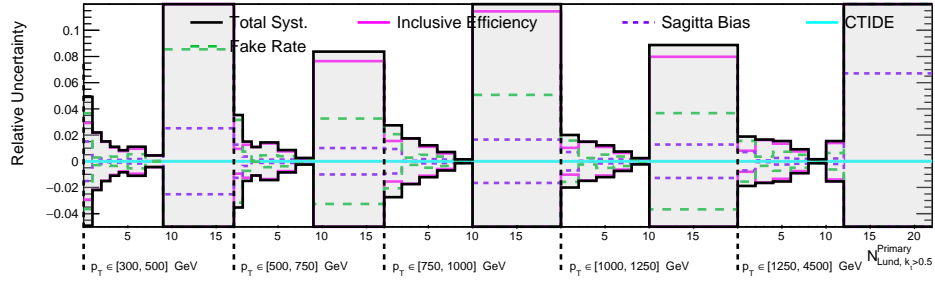


(g) $k_t > 50$ GeV, N_{Lund}

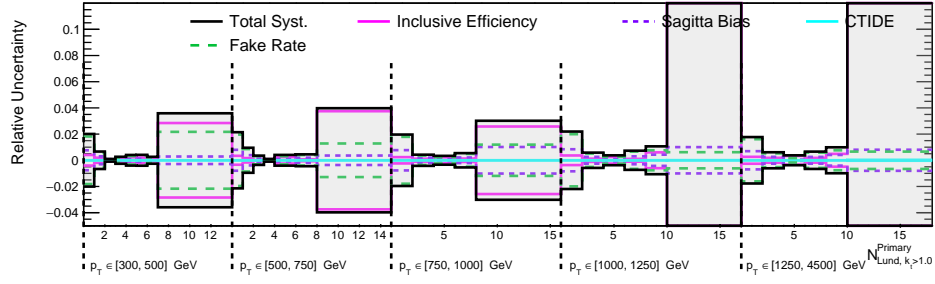


(h) $k_t > 100$ GeV, N_{Lund}

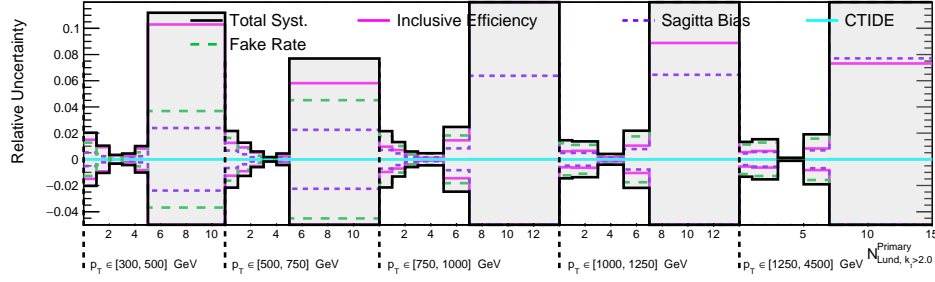
Figure 8.32: (continued) Tracking uncertainties for Lund multiplicity for several k_t requirements



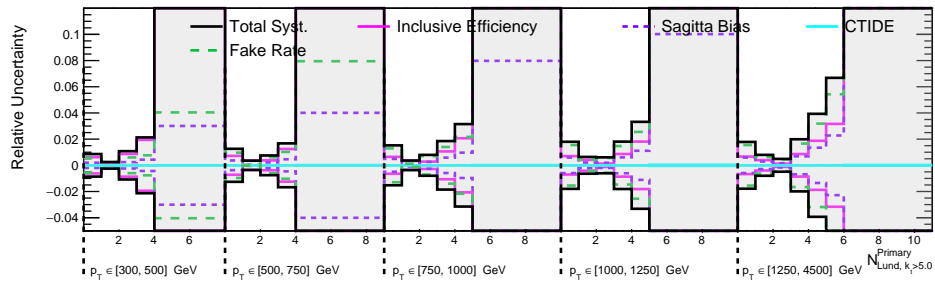
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

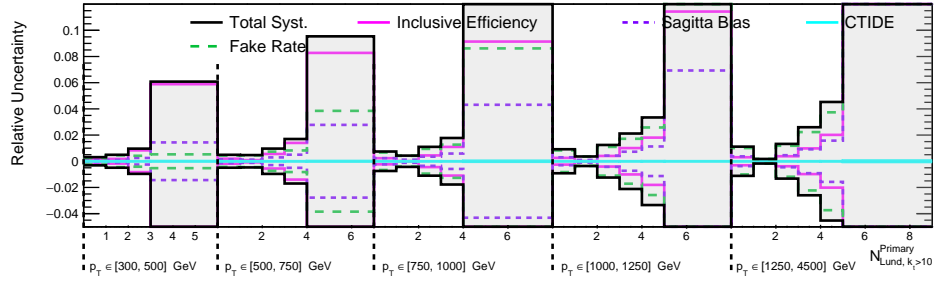


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

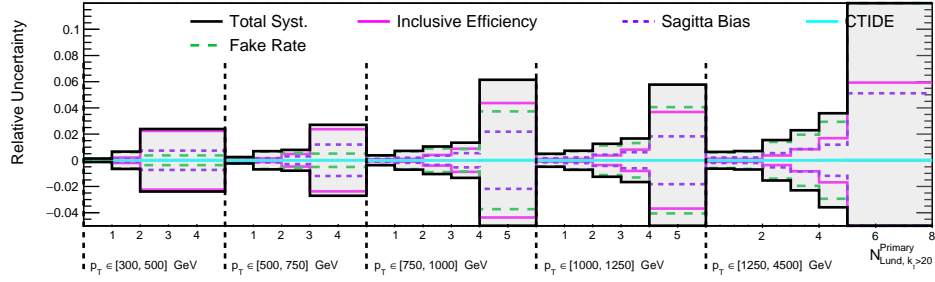


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

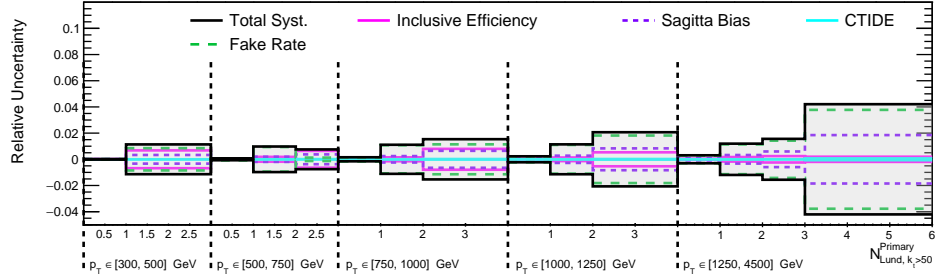
Figure 8.33: Tracking uncertainties for primary Lund multiplicity for several k_t requirements. Figure continued on next page.



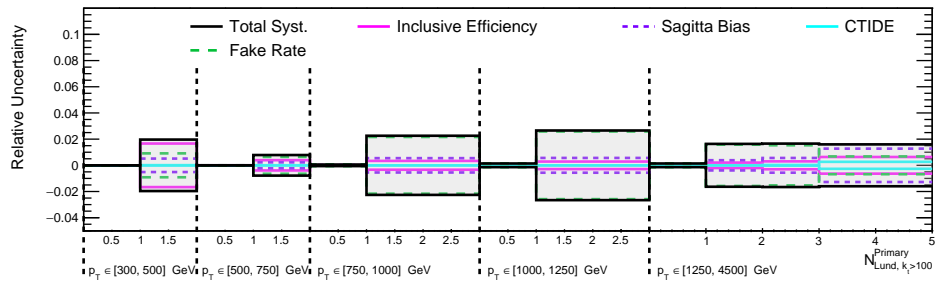
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.33: (continued) Tracking uncertainties for primary Lund multiplicity for several k_t requirements.

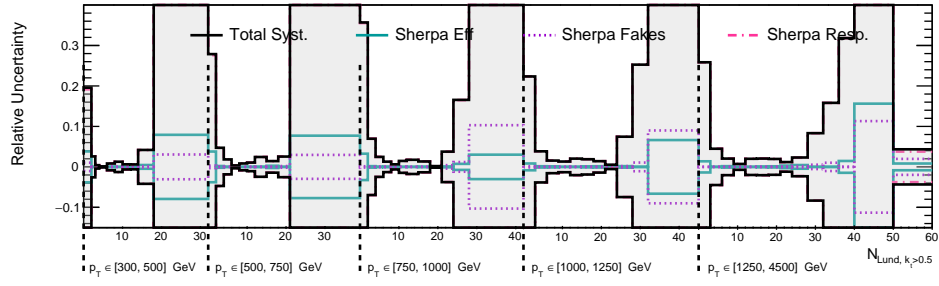
8.7.6 *Uncertainty due to the choice of Monte Carlo models*

In order to unfold a measurement, one relies on some selected nominal Monte Carlo simulation to construct the response matrix applied to data. No particular MC models data perfectly, and so different results may be obtained if a different MC model is used to define the unfolding procedure. In order to account for uncertainty related to the choice of the nominal MC model, the unfolding procedure is repeated with the nominal PYTHIA prior but a response matrix, fake and efficiency factors which are constructed using an alternative (or multiple alternative) MC samples. A conservative uncertainty is defined by adding these three sources of uncertainty in quadrature for each model used.

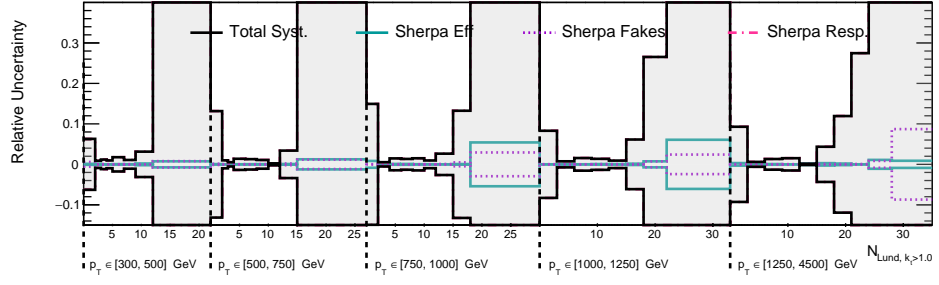
Specifically the fake and efficiency uncertainties are derived from the the differences in the fake and efficiency factors between the nominal unfolding setup and the setup using the alternative model response matrix. The uncertainty due to the response matrix is derived from the result when the unfolding is performed with the alternative response matrix but the nominal PYTHIA fake and efficiency correction factors. Uncertainties on prior variations are considered via the data-driven non-closure uncertainty (Section 8.6.5), and so they are not considered in the modelling error calculation to avoid double-counting.

The alternative samples used to define this uncertainty are the tuned SHERPA 2.2.11 sample with the default cluster hadronization model and the HERWIG 7.1.3 sample with the angle-ordered PS. For each of these samples the procedure described above is performed, and an envelope of the results from each alternative MC setup is taken as the final uncertainty.

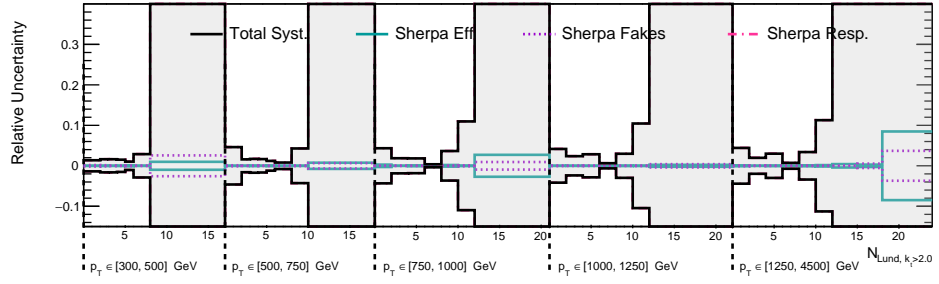
Breakdowns of these components for both SHERPA and HERWIG are shown in Figure 8.34 - Figure 8.37 for the complete set of measured distributions. The final modelling uncertainty takes an envelope of the uncertainties derived using either SHERPA or HERWIG, and is shown in Figure 8.38 and Figure 8.39.



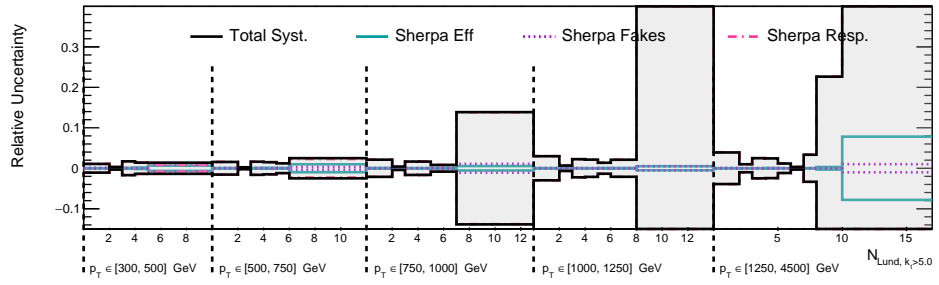
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

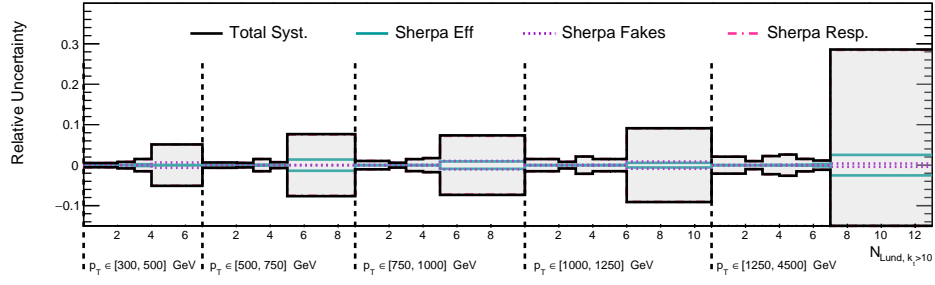


(c) $k_t > 2$ GeV, N_{Lund}

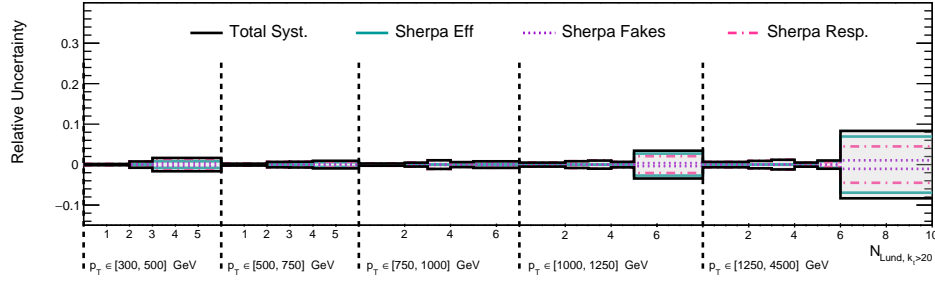


(d) $k_t > 5$ GeV, N_{Lund}

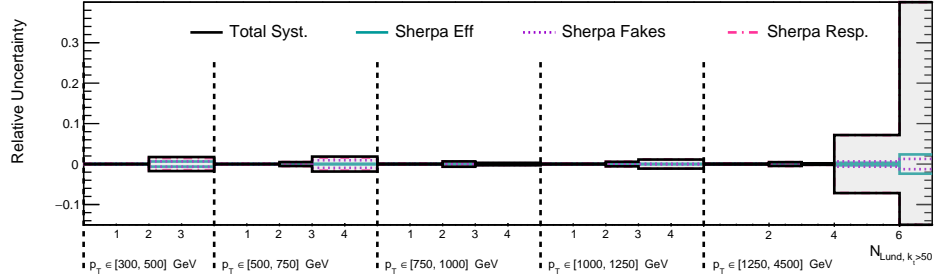
Figure 8.34: Uncertainty due to the use of an alternative MC model (SHERPA 2.2.11 with re-tuned cluster hadronisation) on Lund multiplicity distributions. Figure continued on next page.



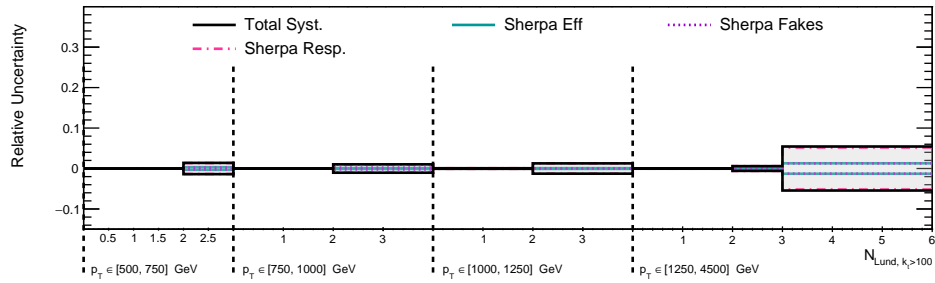
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

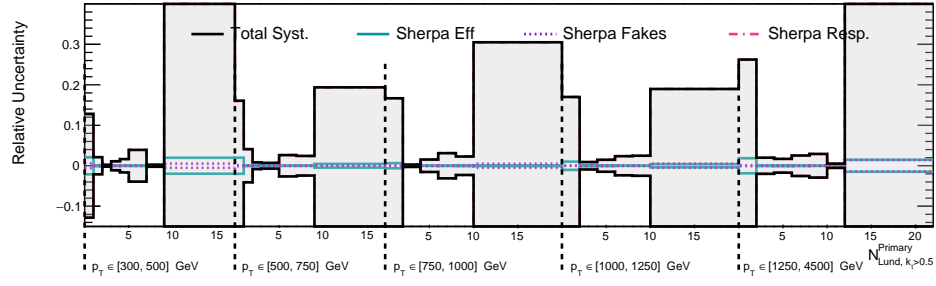


(g) $k_t > 50$ GeV, N_{Lund}

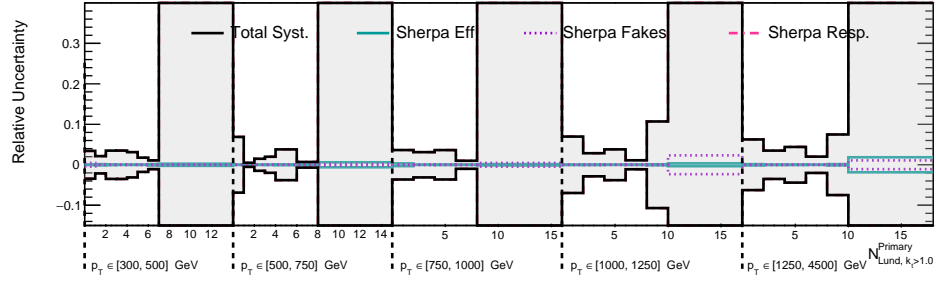


(h) $k_t > 100$ GeV, N_{Lund}

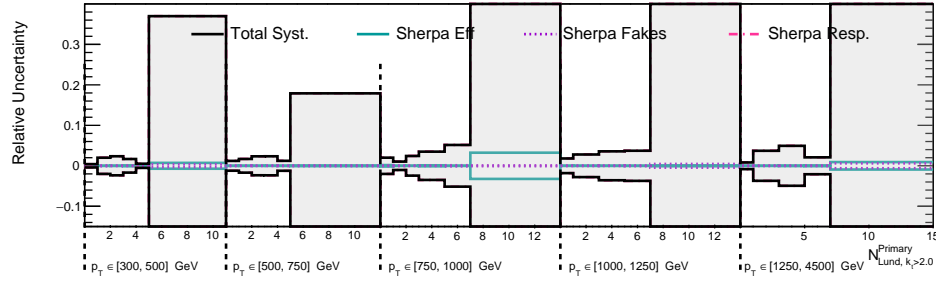
Figure 8.34: (continued) Uncertainty due to the use of an alternative MC model (SHERPA 2.2.11 with re-tuned cluster hadronisation) on Lund multiplicity distributions.



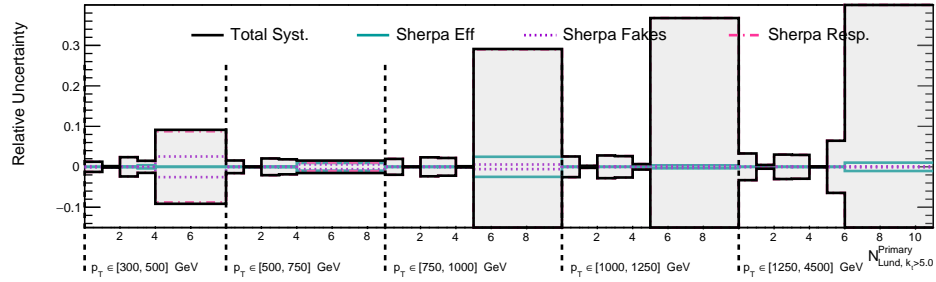
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

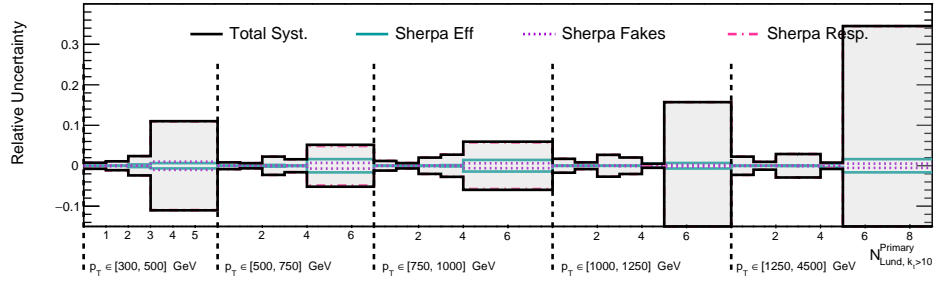


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

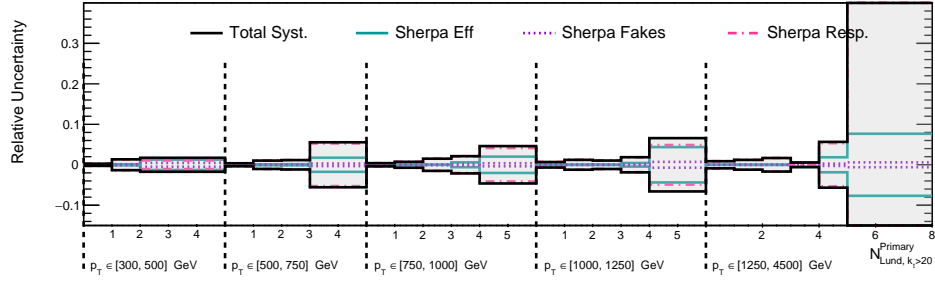


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

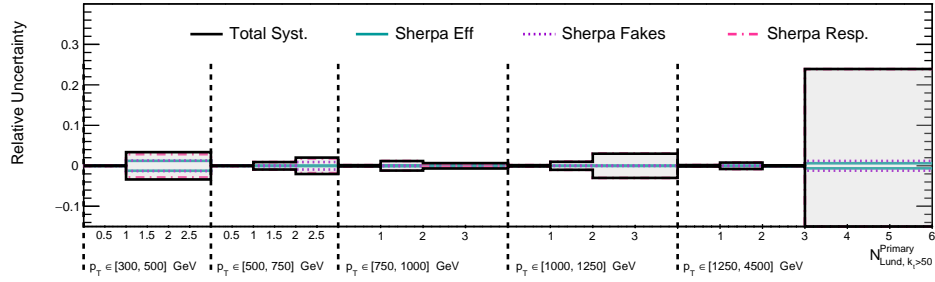
Figure 8.35: Uncertainty due to the use of an alternative MC model (SHERPA 2.2.11 with re-tuned cluster hadronisation) on primary Lund multiplicity distributions. Figure continued on next page.



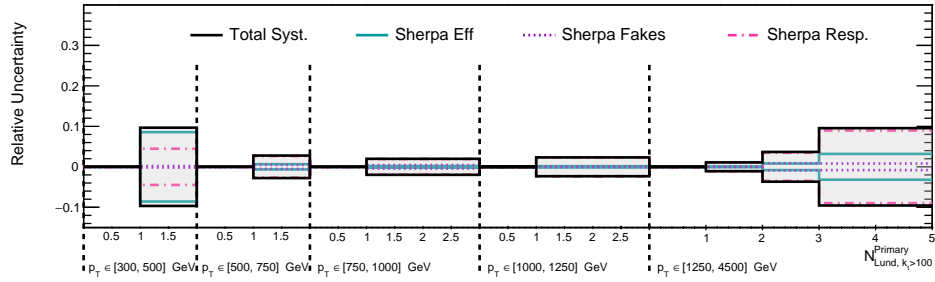
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

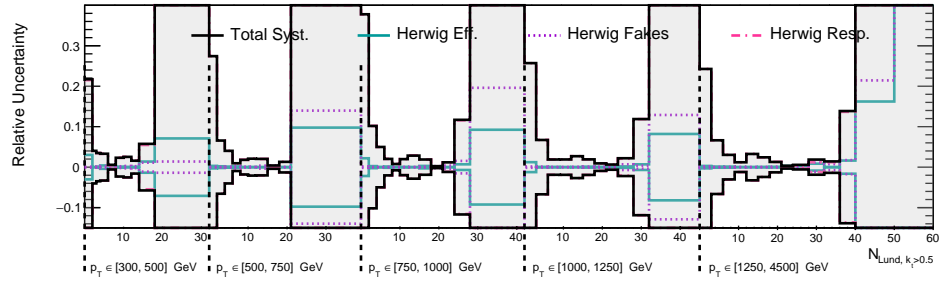


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

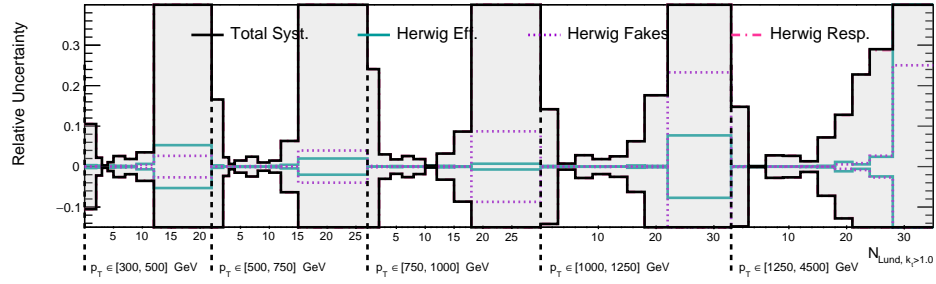


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

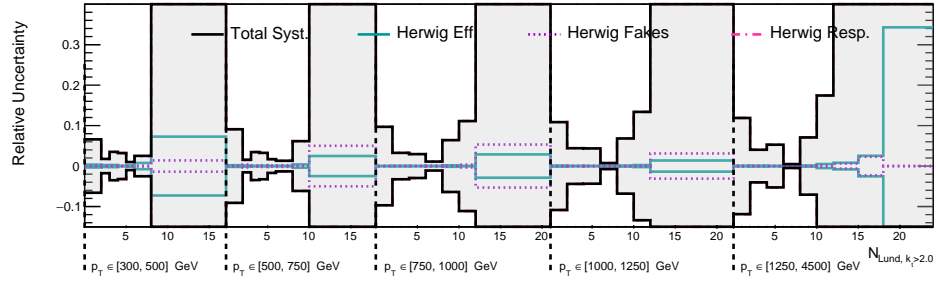
Figure 8.35: (continued) Uncertainty due to the use of an alternative MC model (SHERPA 2.2.11 with re-tuned cluster hadronisation) on primary Lund multiplicity distributions.



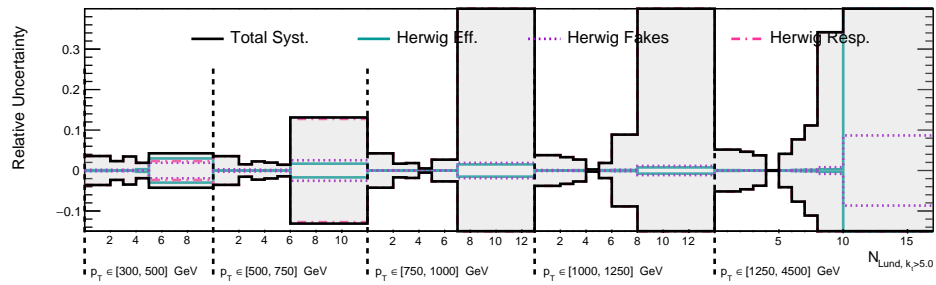
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

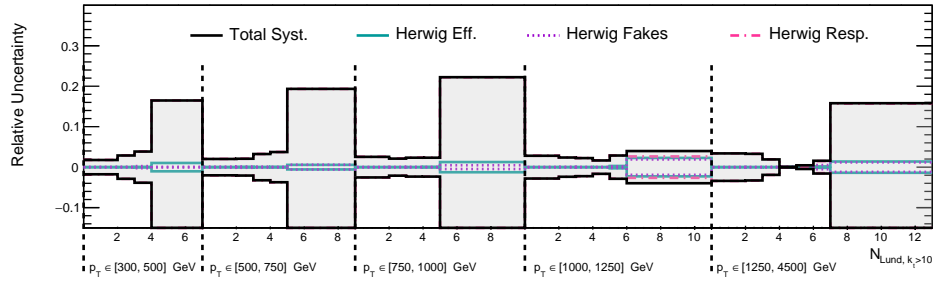


(c) $k_t > 2$ GeV, N_{Lund}

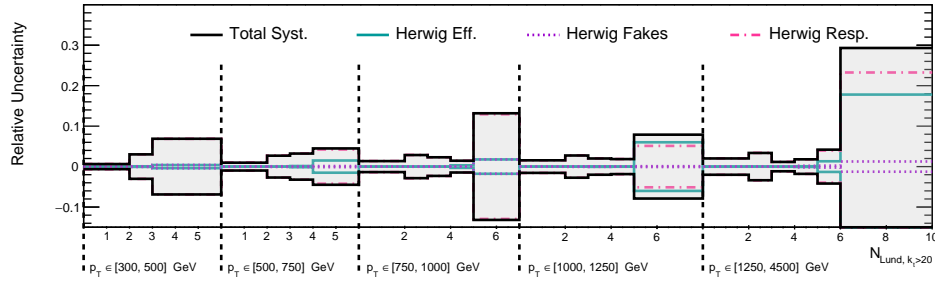


(d) $k_t > 5$ GeV, N_{Lund}

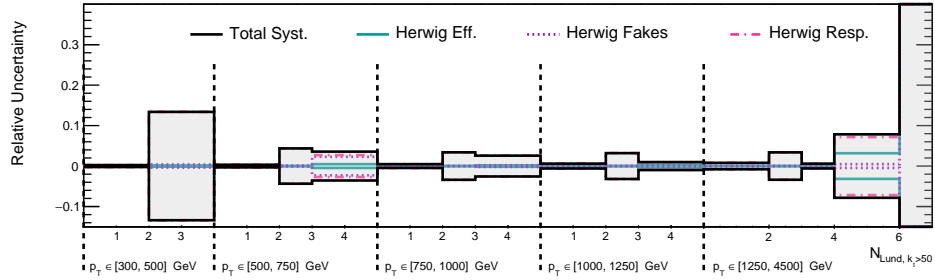
Figure 8.36: Uncertainty due to the use of an alternative MC model (HERWIG 7.1 with angle-ordered parton shower) on Lund multiplicity distributions. Figure continued on next page.



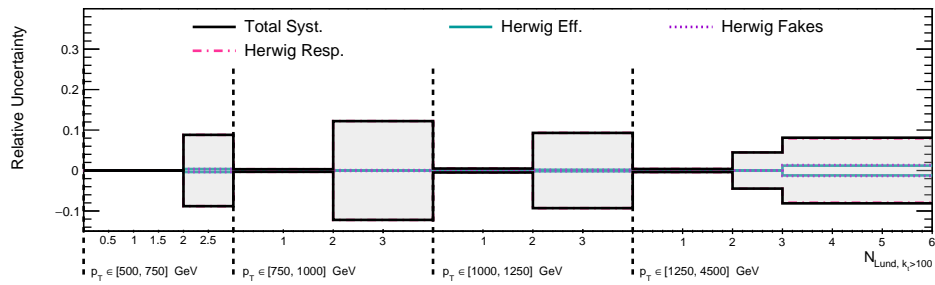
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

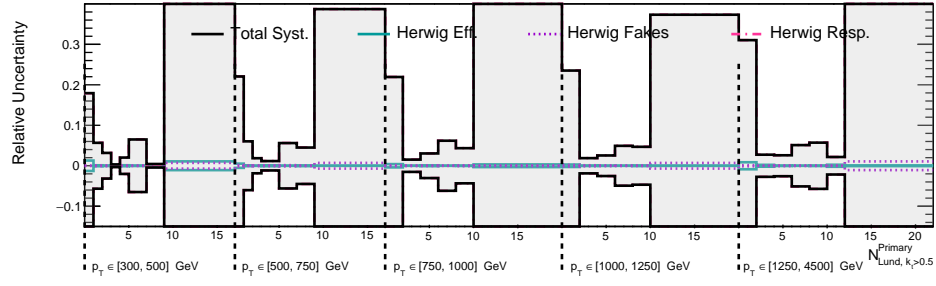


(g) $k_t > 50$ GeV, N_{Lund}

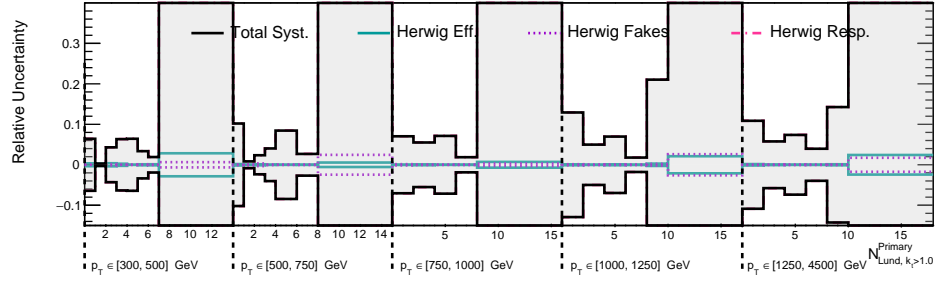


(h) $k_t > 100$ GeV, N_{Lund}

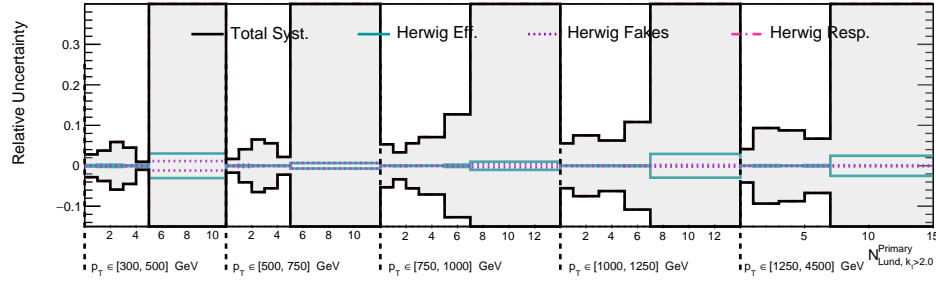
Figure 8.36: (continued) Uncertainty due to the use of an alternative MC model (HERWIG 7.1 with angle-ordered parton shower) on Lund multiplicity distributions.



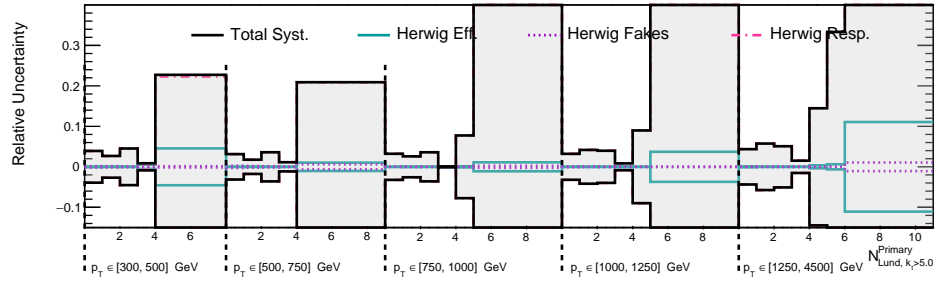
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

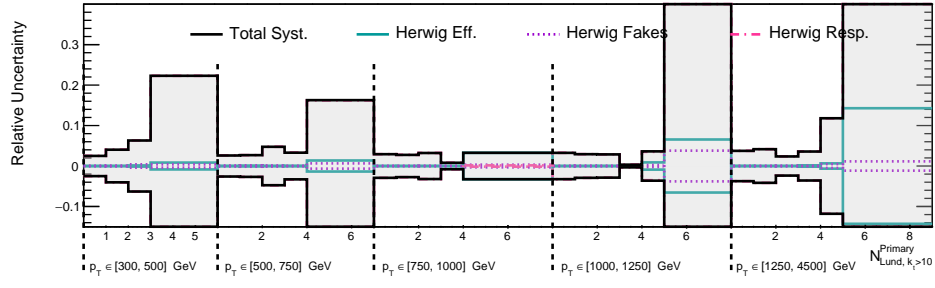


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

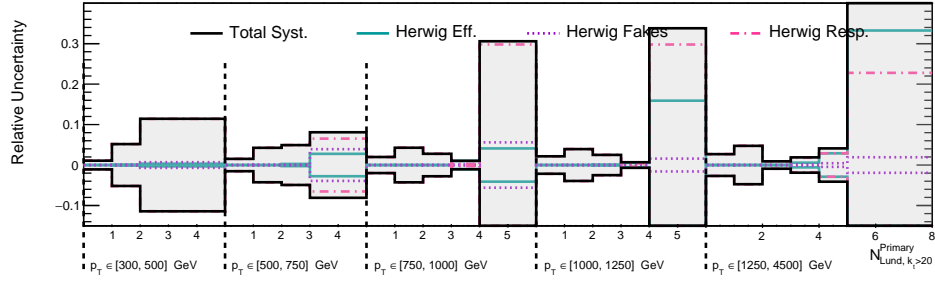


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

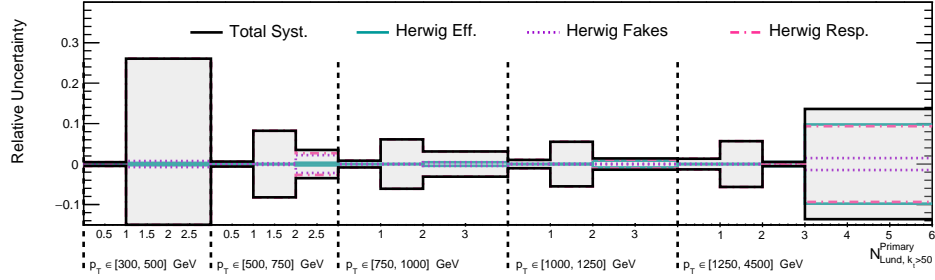
Figure 8.37: Uncertainty due to the use of an alternative MC model (HERWIG 7.1 with angle-ordered parton shower) on primary Lund multiplicity distributions. Figure continued on next page.



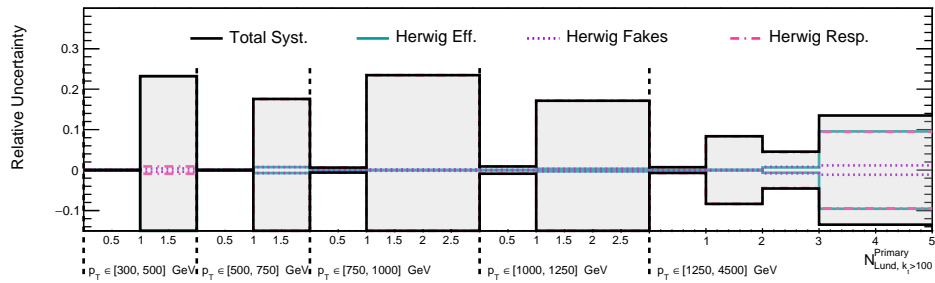
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

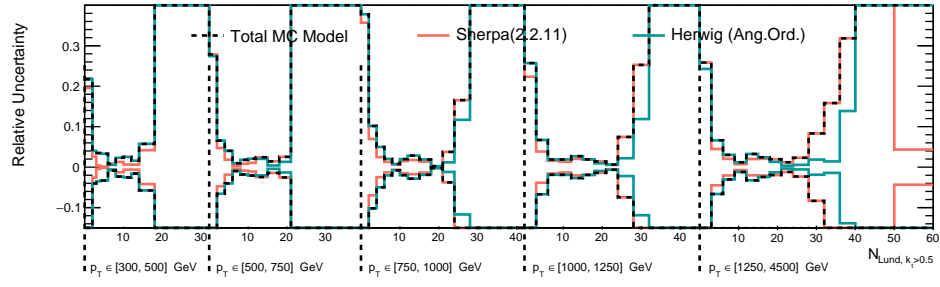


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

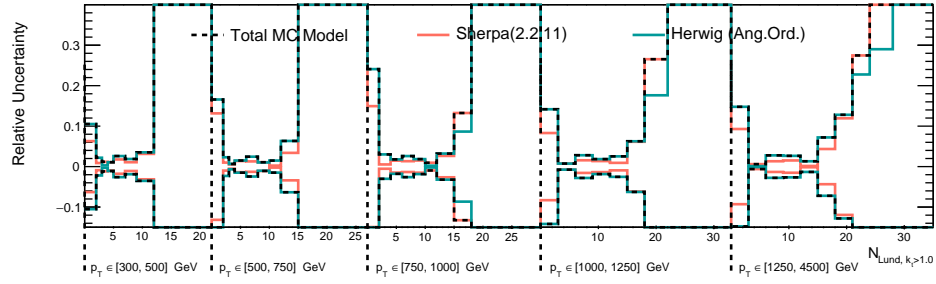


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

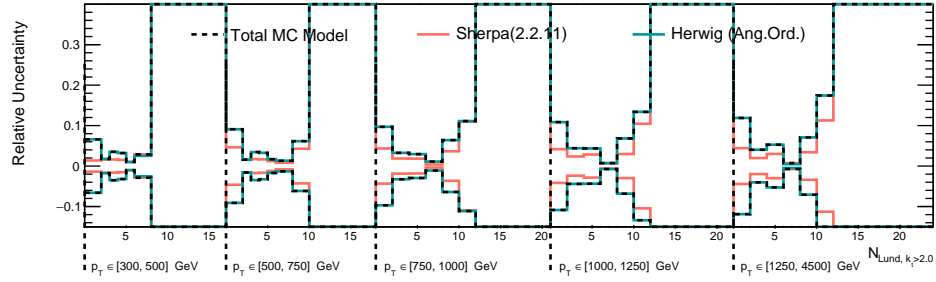
Figure 8.37: (continued) Uncertainty due to the use of an alternative MC model (HERWIG 7.1 with angle-ordered parton shower) on primary Lund multiplicity distributions.



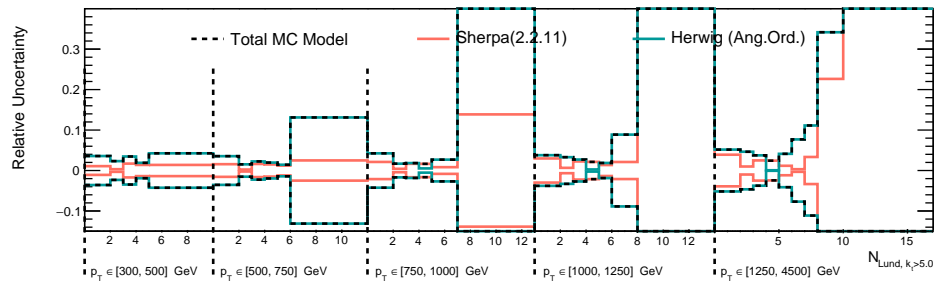
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

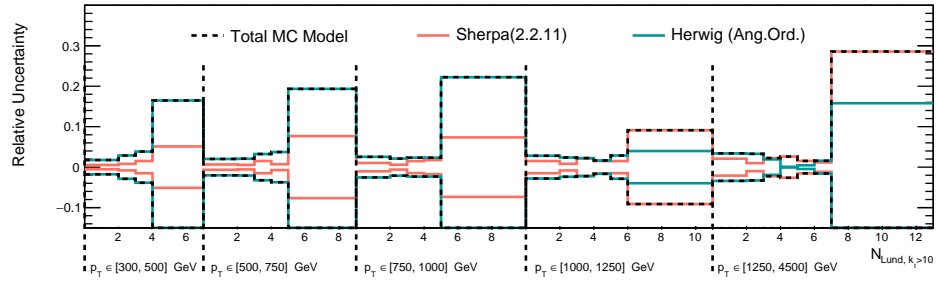


(c) $k_t > 2$ GeV, N_{Lund}

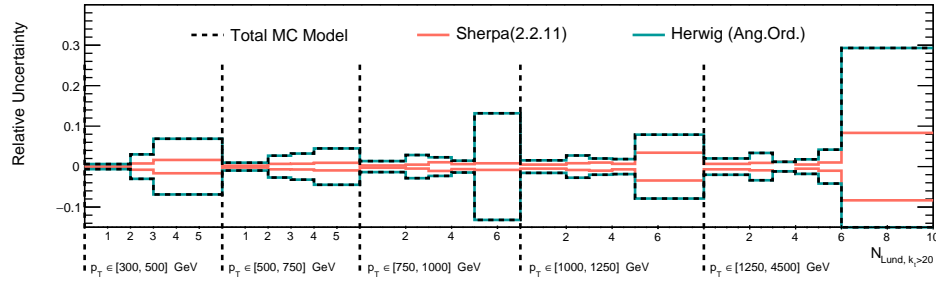


(d) $k_t > 5$ GeV, N_{Lund}

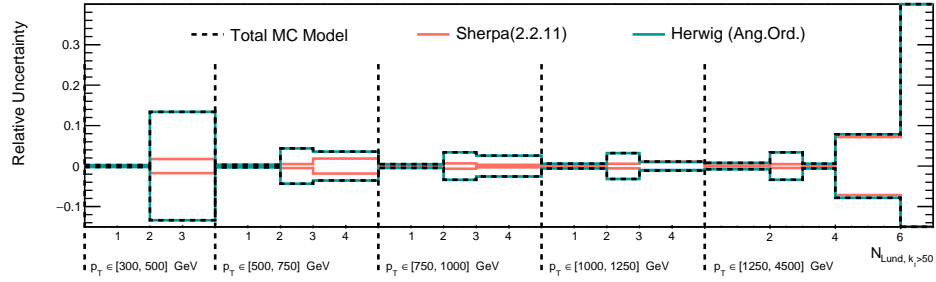
Figure 8.38: Total uncertainty due to the use of alternative MC models SHERPA 2.2.11 and HERWIG 7.1 on Lund multiplicity distributions. Figure continued on next page.



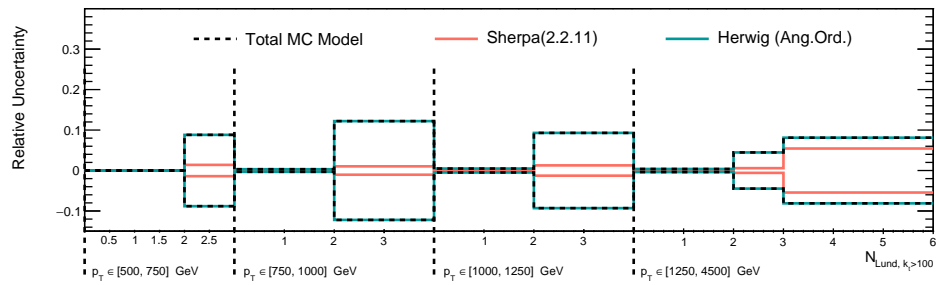
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

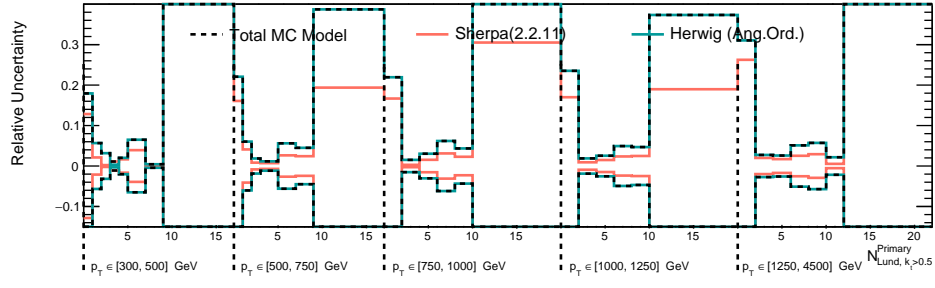


(g) $k_t > 50$ GeV, N_{Lund}

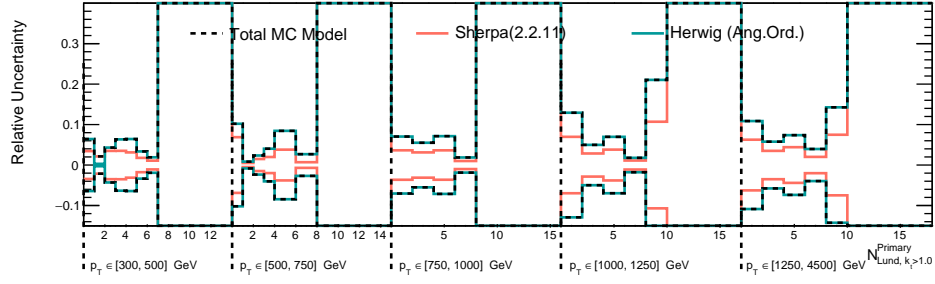


(h) $k_t > 100$ GeV, N_{Lund}

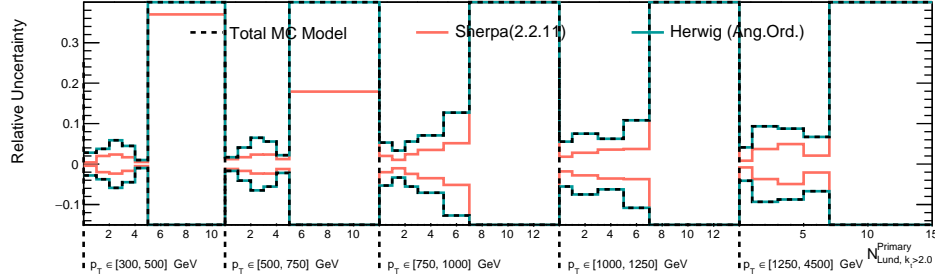
Figure 8.38: (continued) Total uncertainty due to the use of alternative MC models SHERPA 2.2.11 and HERWIG 7.1 on Lund multiplicity distributions.



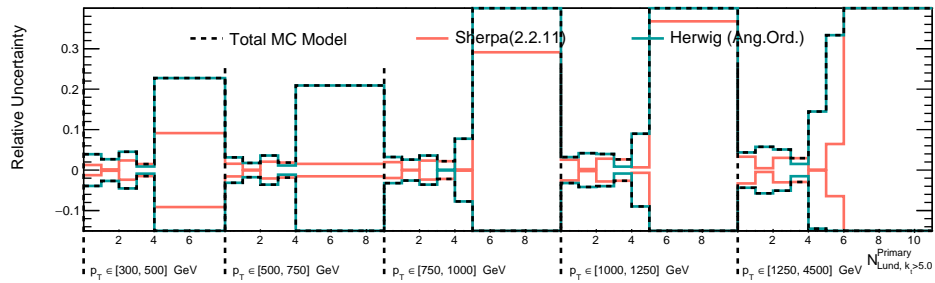
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

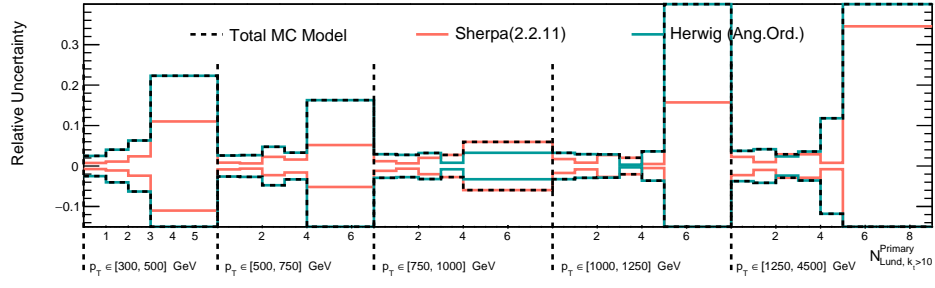


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

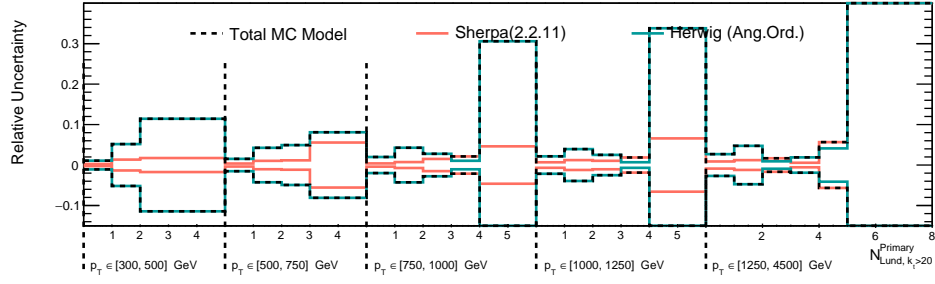


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

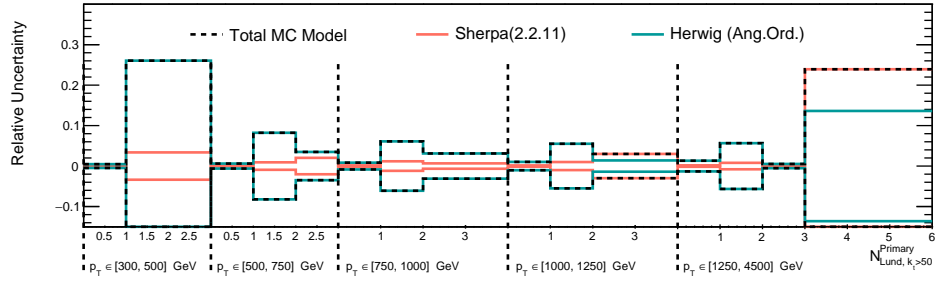
Figure 8.39: Total uncertainty due to the use of alternative MC models SHERPA 2.2.11 and HERWIG 7.1 on primary Lund multiplicity distributions. Figure continued on next page.



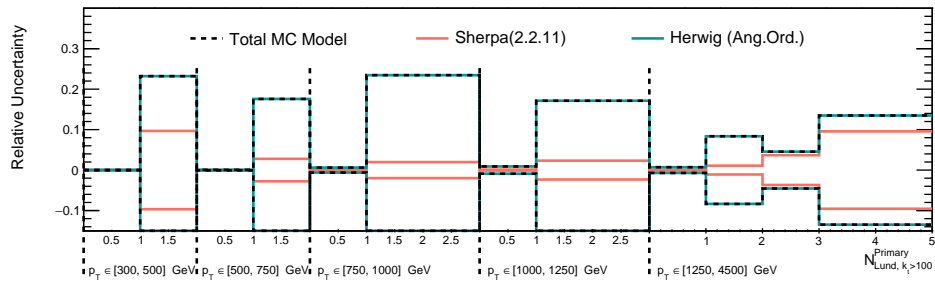
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.39: (continued) Total uncertainty due to the use of alternative MC models SHERPA 2.2.11 and HERWIG 7.1 on primary Lund multiplicity distributions.

8.8 Results

8.8.1 Multiplicity Distributions

A full selection of the measured distributions are presented in this section for both Lund (Figure 8.40) and primary Lund (Figure 8.41) multiplicities for various emission k_t requirements. These results are presented differentially in bins of jet p_T and the unfolded data are compared to predictions from several Monte Carlo models. For the majority of both the measured N_{Lund} and $N_{\text{Lund}}^{\text{Primary}}$ distributions, the largest contributor to the uncertainty is related to the choice of Monte Carlo model used for the unfolding, though a subset of higher transverse momentum bins see a larger impact from the JES/JER uncertainty.

Broadly speaking, the nominal PYTHIA MC sample provides an accurate description of the unfolded data within the core of multiplicity distributions given the precision of the differential cross-section measurement, except in the case of $N_{\text{Lund}}^{\text{Primary}}$ with small k_t requirements where agreement in the core of the distribution is poor. The agreement is found to degrade for both large and small multiplicity values for all k_t requirements that are studied, for both N_{Lund} and $N_{\text{Lund}}^{\text{Primary}}$. The overall best agreement between PYTHIA and the unfolded N_{Lund} data is found for an emission k_t requirement of 5 GeV. For smaller k_t requirements, the PYTHIA sample predicts significantly more high-multiplicity events for both N_{Lund} and $N_{\text{Lund}}^{\text{Primary}}$ than are observed in the data. When the k_t requirement is larger than 5 GeV, a deficit of high-multiplicity jets is instead observed. The differences observed between the PYTHIA and POWHEG+PYTHIA setups appear to generally be small, and are largest for the N_{Lund} distributions. This observation is expected, as it indicates that the additional matrix-element accuracy does not contribute strongly to the description of $R = 0.4$ jet substructure. Differences between the two samples may be attributable to the A14 set of tuned parameters also being used for the POWHEG+PYTHIA MC, despite being designed for use

with the PYTHIA sample. Agreement at large and small multiplicity values is still poor for the POWHEG+PYTHIA sample, although it is found to be slightly better than the standalone PYTHIA sample (particularly for large k_t requirements, *e.g.* Figure 8.40f).

Four different setups of the SHERPA generator are compared to the unfolded data. Agreement of these configurations with the measured data depends strongly on the phase-space region that is being studied. For example, Figure 8.40c illustrates that the SHERPA 2.2.11 sample with retuned cluster-based hadronisation agrees best with the measured data in the lowest p_T bin, but that agreement between the data and this setup degrades as the jet p_T increases. In the highest jet p_T bins, the older hadronisation tune in the SHERPA 2.2.5 sample performs significantly better. One possible explanation for this is that LHC data was included in the initial tune, while the re-tuning campaign was performed only with LEP fragmentation measurements. This observation emphasizes the need for substructure measurements from the LHC to be included when improvements to Monte Carlo models are being made for LHC applications. The SHERPA 2.2.11 AHADIC hadronisation sample is found to be very consistent with the SHERPA 2.2.5 sample with Lund string hadronisation throughout the measurements, confirming earlier simulation-based studies that found the ATLAS jet energy scale of the re-tuned sample to be more consistent with that used in PYTHIA MC [16]. The SHERPA 2.2.5 AHADIC sample is found to provide the best overall description of the unfolded N_{Lund} data for configurations with increased amounts of soft radiation, when the k_t requirement is low. For $N_{\text{Lund}}^{\text{Primary}}$; however, all of the SHERPA samples perform worse: the DIRE setup is found to have better agreement the others at low k_t , but other generators typically describe these distributions better.

The HERWIG 7.1.3 sample with the default HERWIG angle-ordered parton shower provides the best description of the N_{Lund} distributions for $k_t > 10$ GeV (*e.g.* Figure 8.40b). For these measurements with large k_t requirements, it is consistent with the measured data and

does not exhibit the same decreasing trend found in all other models that were considered. However, it does behave similarly for $N_{\text{Lund}}^{\text{Primary}}$ at high p_{T} until higher emission k_t requirements, around 50 GeV (*e.g.* Figures 8.41f and 8.41g). In lower p_{T} bins, the agreement between this sample and the unfolded data is improved for slightly lower k_t requirements (around 10 GeV). The HERWIG 7.1.3 dipole parton shower sample generally performs poorly when compared to other models, and often exhibits the largest disagreement of the models considered. This disagreement is most pronounced for the N_{Lund} cross-sections, and for $N_{\text{Lund}}^{\text{Primary}}$ at high emission k_t (> 10 GeV).

These results also present the first (external) comparison of the SHERPA ALARIC [130, 10] parton shower to collider data. The aim of this parton shower is to correctly simulate the two-loop soft radiation pattern, which requires that the kinematics mapping be modified to provide resolved parton evolution [105, 79]. This technique allows the NLL precision of the evolution to be analytically proven, and results in splitting functions which acquire an azimuthal angle dependence. This sample is hadronized using the Lund string model via an interface to PYTHIA. The SHERPA ALARIC PS is found to perform consistently with the other SHERPA models studied for large k_t requirements ($k_t \geq 5$ GeV), where non-perturbative contributions are suppressed. At low k_t this sample looks very similar to PYTHIA, which is expected since PYTHIA hadronization is used.

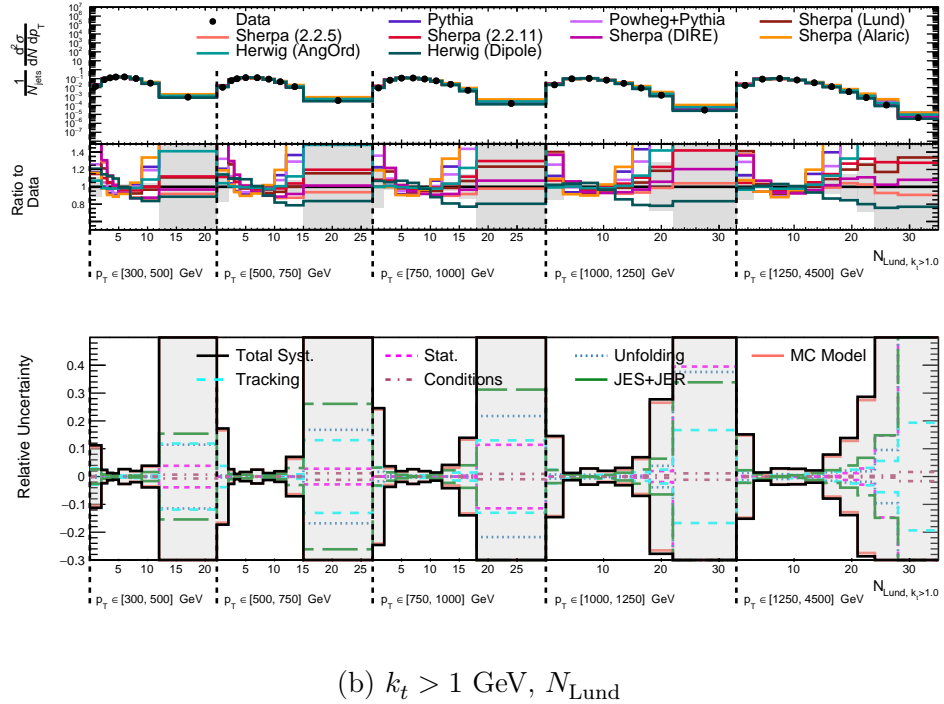
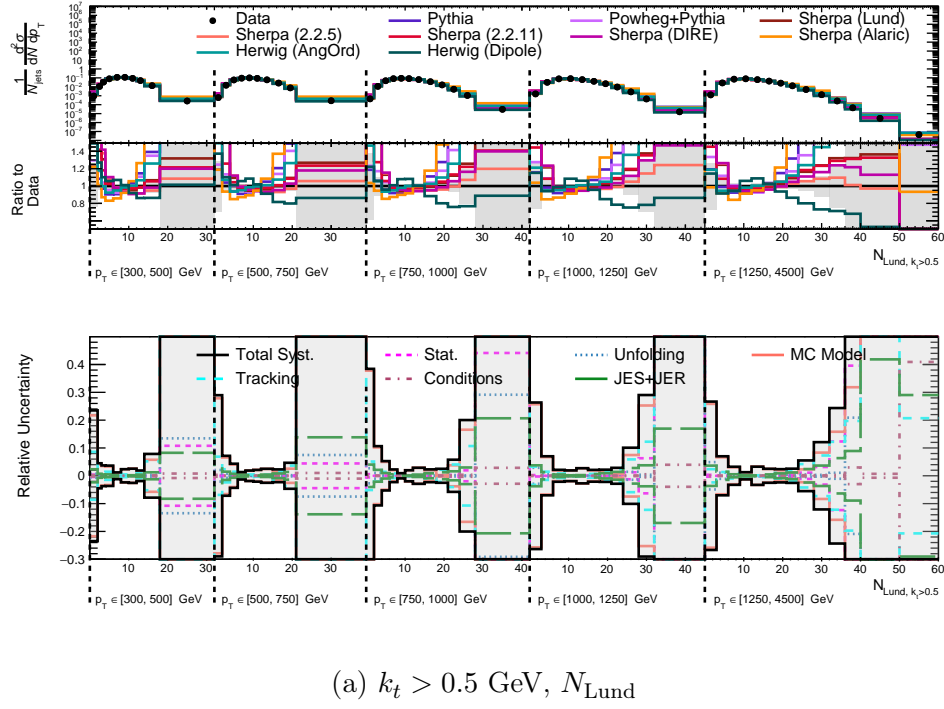
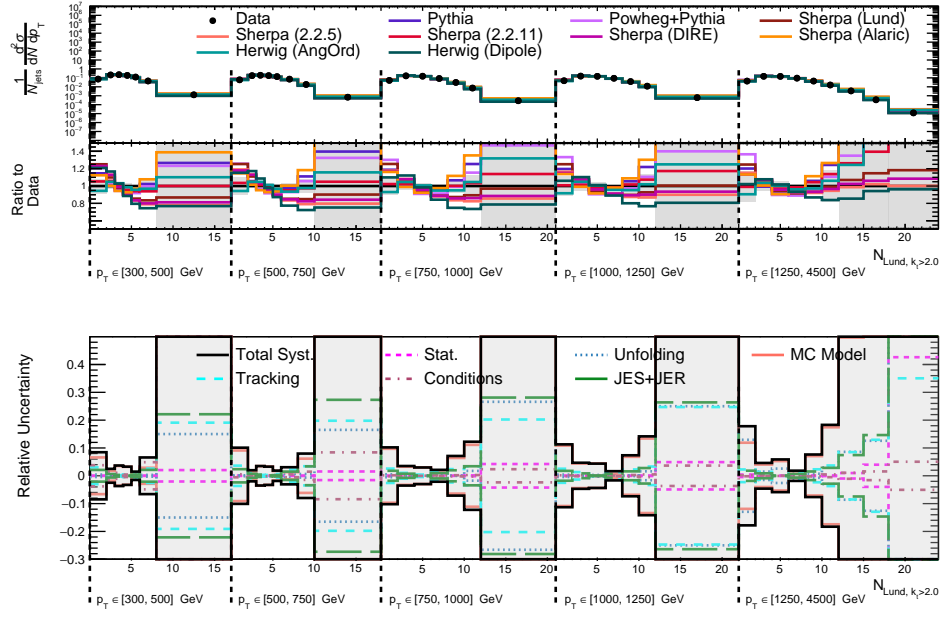
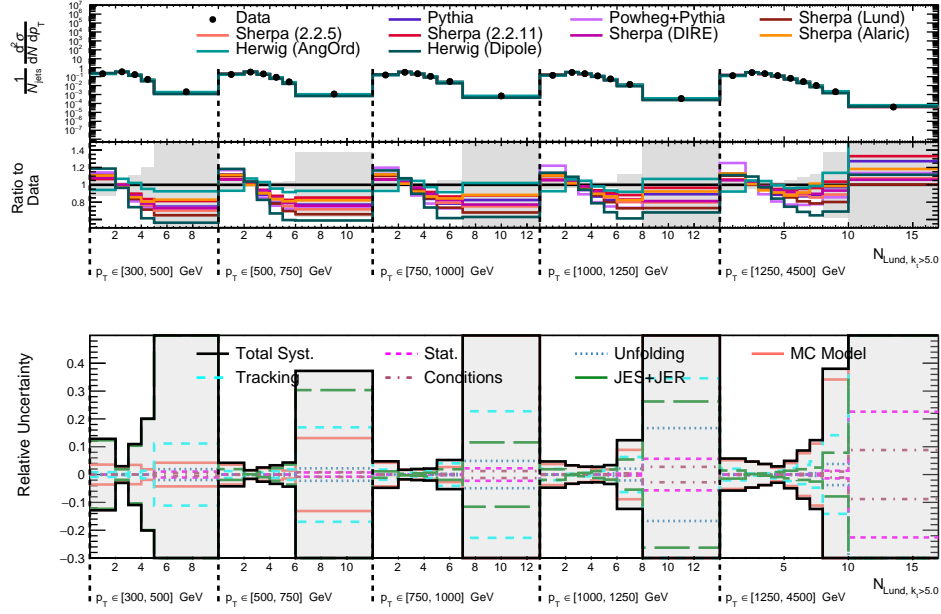


Figure 8.40: The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.



(c) $k_t > 2$ GeV, N_{Lund}



(d) $k_t > 5$ GeV, N_{Lund}

Figure 8.40: (continued) The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.

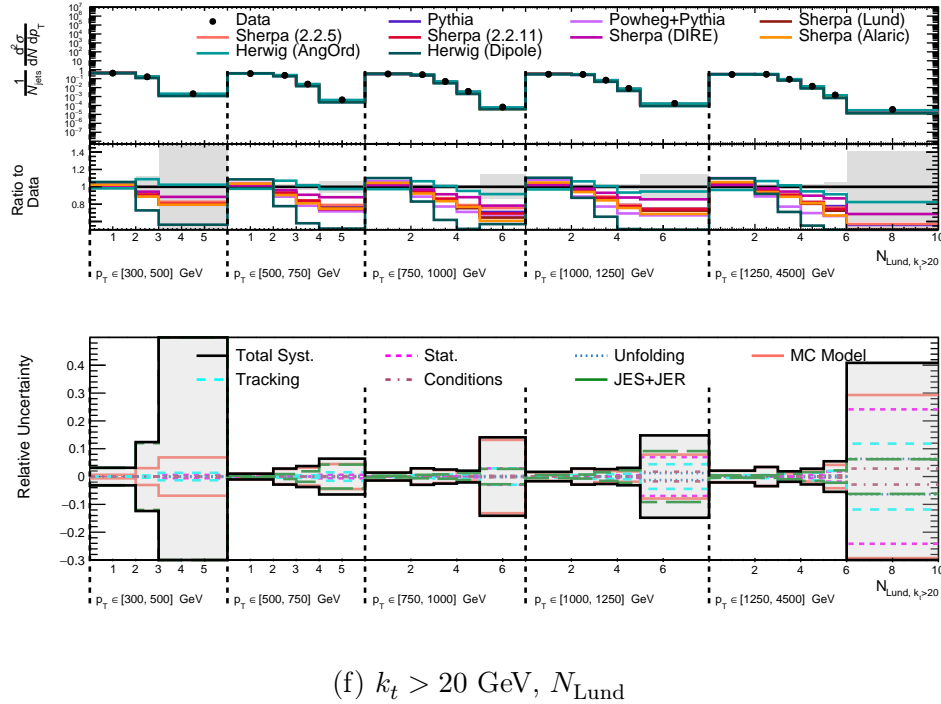
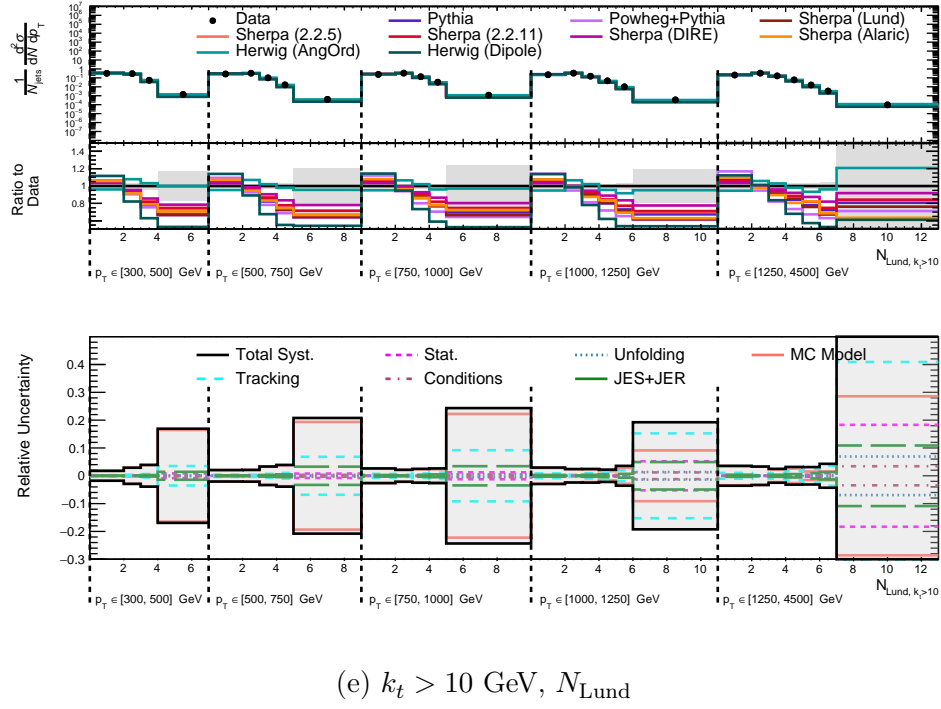
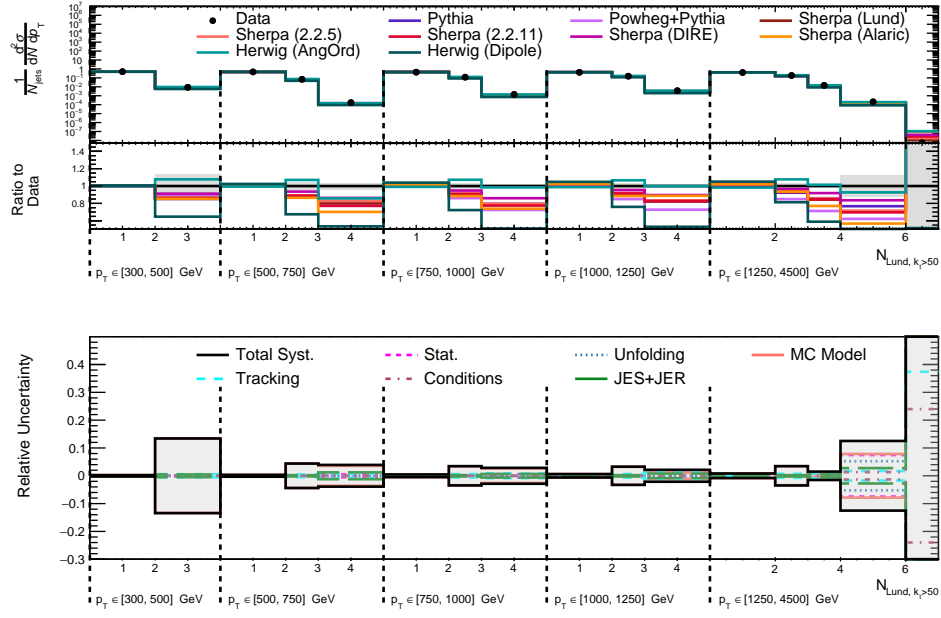
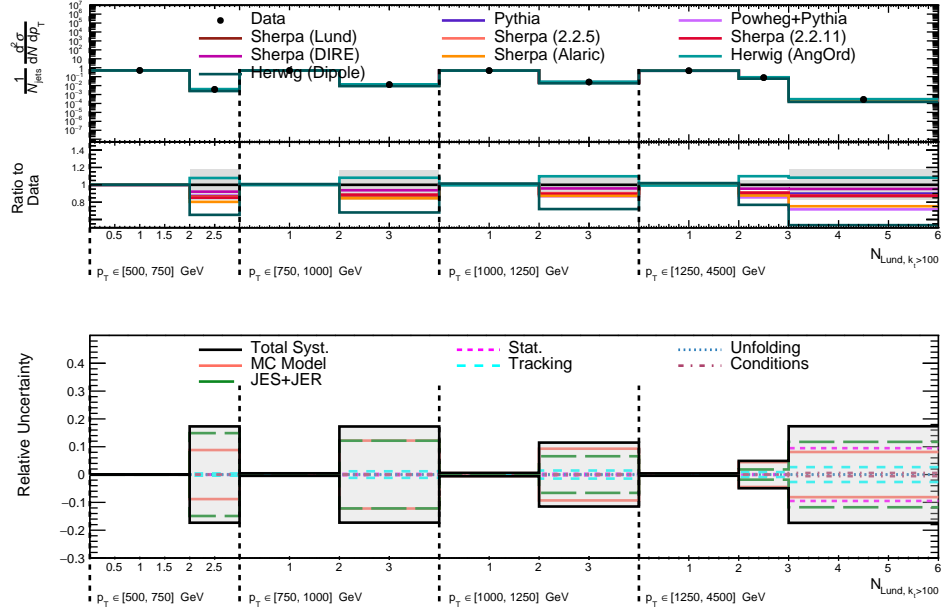


Figure 8.40: (continued) The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.



(g) $k_t > 50$ GeV, N_{Lund}



(h) $k_t > 100$ GeV, N_{Lund}

Figure 8.40: (continued) The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement.

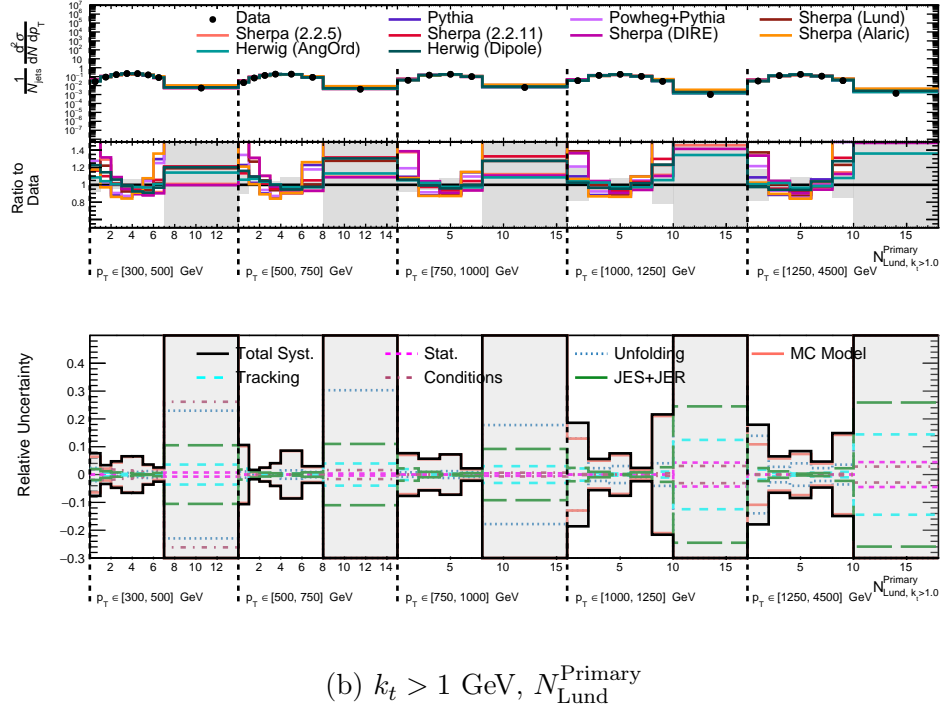
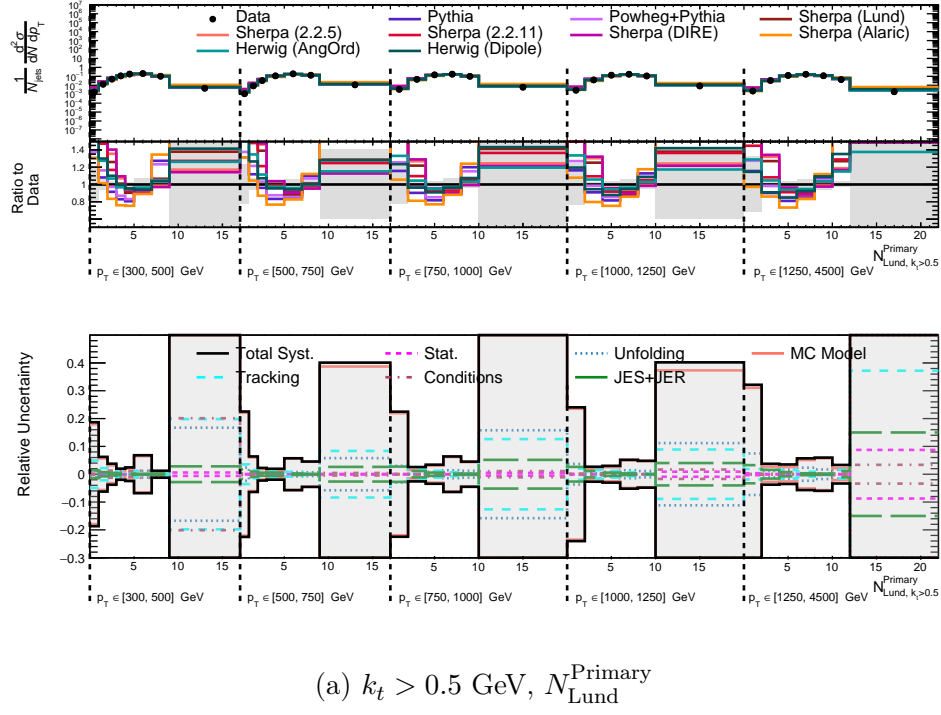


Figure 8.41: The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.

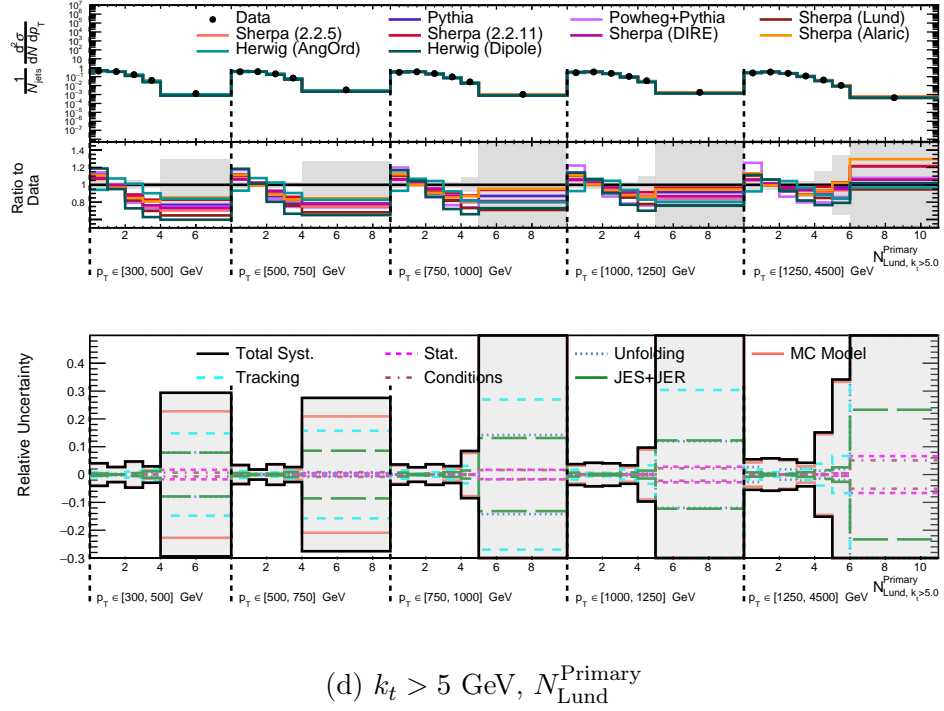
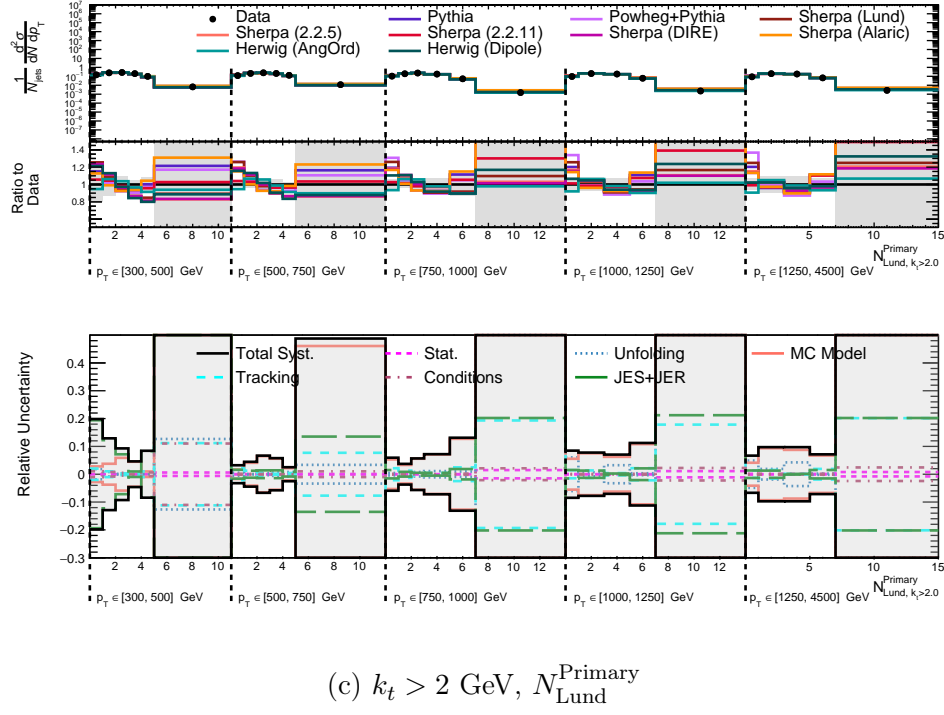


Figure 8.41: (continued) The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.

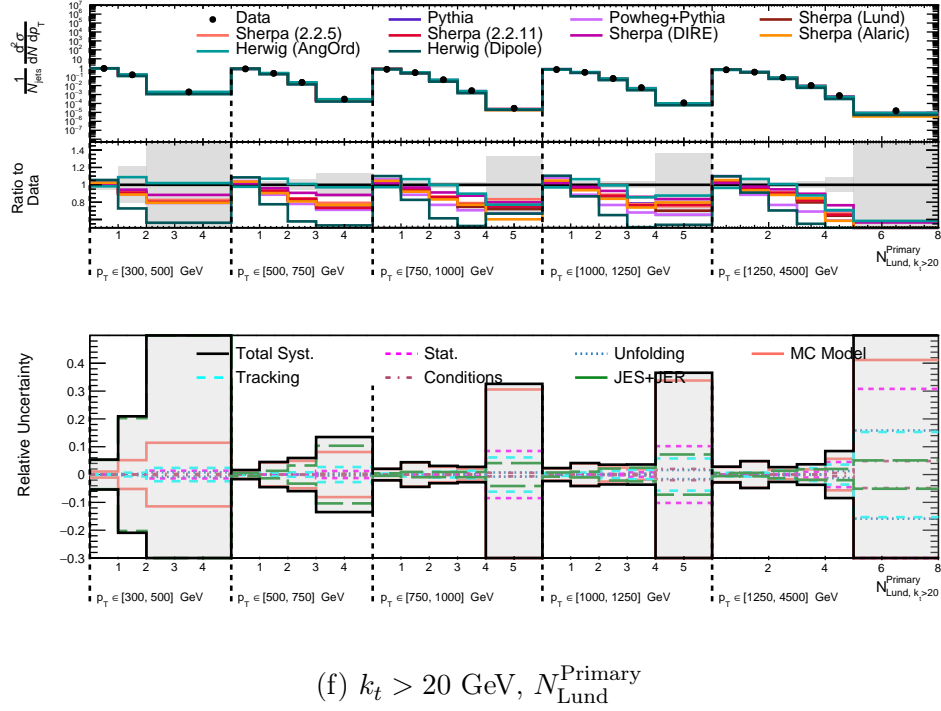
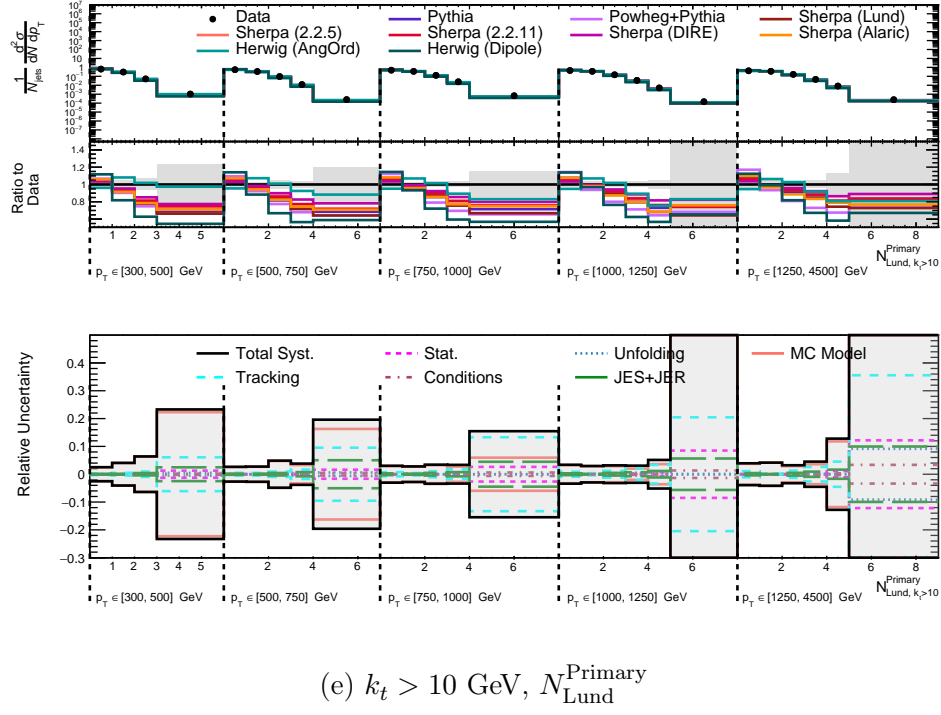


Figure 8.41: (continued) The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement. Figure continued on next page.

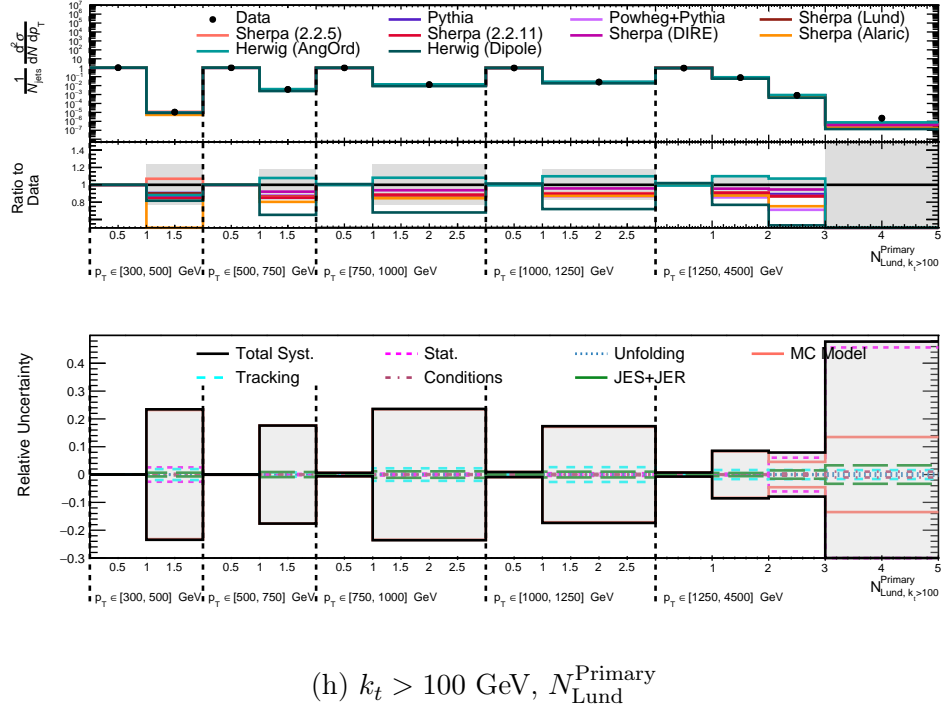
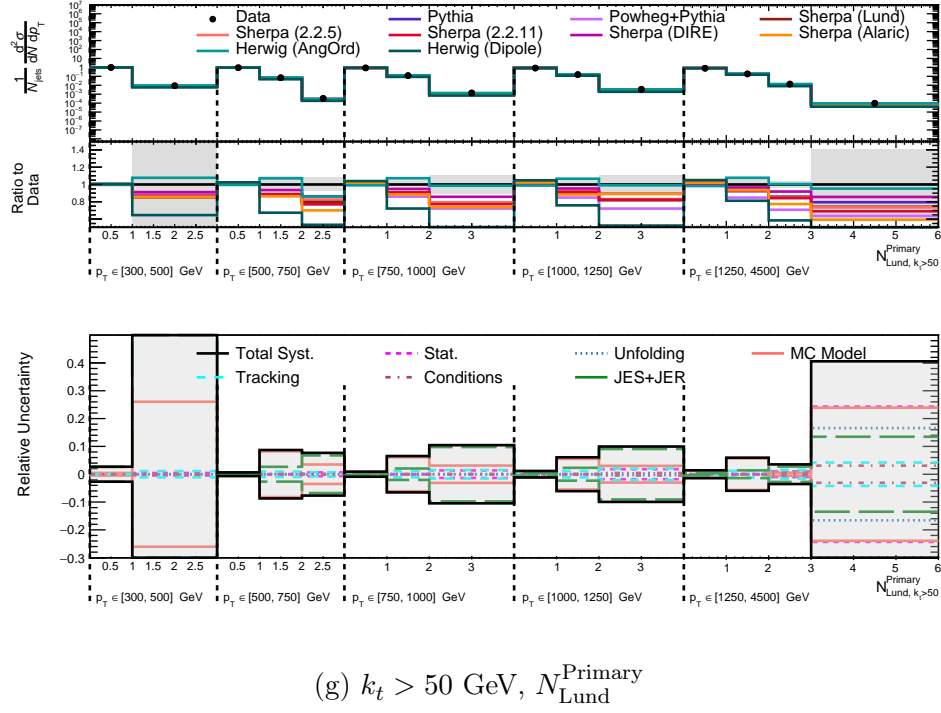


Figure 8.41: (continued) The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators, and the corresponding total uncertainty on the measurement.

8.8.2 Average Multiplicities

In this section, the distributions of the average values of N_{Lund} ($\langle N_{\text{Lund}} \rangle$, Figure 8.43) and $N_{\text{Lund}}^{\text{Primary}}$ ($\langle N_{\text{Lund}}^{\text{Primary}} \rangle$, Figure 8.44) as a function of the emission k_t requirement are also presented for each of the inclusive bins in the jet p_T and the relative jet rapidity-ordering. These results are also presented for an inclusive jet p_T selection above 300 GeV in Figure 8.42. These distributions are smoothly falling, from a maximum value of roughly 10 or more down to values of approximately unity as the emission k_t requirement is increased.

Both the $\langle N_{\text{Lund}} \rangle$ and $\langle N_{\text{Lund}}^{\text{Primary}} \rangle$ distributions show similar qualitative trends for the various MC models when compared to the unfolded data, which also do not depend strongly on the jet p_T bin. Most of the models tend to underpredict the average value in data for intermediate emission k_t requirements, then predict comparatively more emissions as the k_t requirement is increased and decreased to the highest and lowest values.

The HERWIG 7.1 prediction with the angle-ordered PS is found to be in the best overall agreement with the unfolded $\langle N_{\text{Lund}} \rangle$ data, accurately predicting the average values for nearly all emission k_t requirements across the wide range of jet p_T probed by the measurement. When compared to the unfolded $\langle N_{\text{Lund}}^{\text{Primary}} \rangle$ data, the agreement is degraded in the region of intermediate emission k_t and low jet p_T , where this MC over-predicts the measured cross-section while all others under-predict it. Many models tend to agree better for the highest and lowest k_t requirements than in the intermediate k_t region. Large differences between this model and the alternative HERWIG 7.1 sample using the dipole parton shower are observed, particularly in the region of perturbative emissions ($k_t \geq 5 - 50$ GeV). This model also disagrees significantly with other MC setups in this region. When more non-perturbative emissions are able to contribute to $\langle N_{\text{Lund}}^{\text{Primary}} \rangle$, the angle-ordered and dipole PS models come into agreement ($k_t < 5$ GeV), although for the $\langle N_{\text{Lund}} \rangle$ distribution a small difference remains between them in this region. Similarly, in the low- k_t region, large differences are seen

across the different SHERPA models, which primarily differ in their hadronization, and these differences are reduced as the k_t increases.

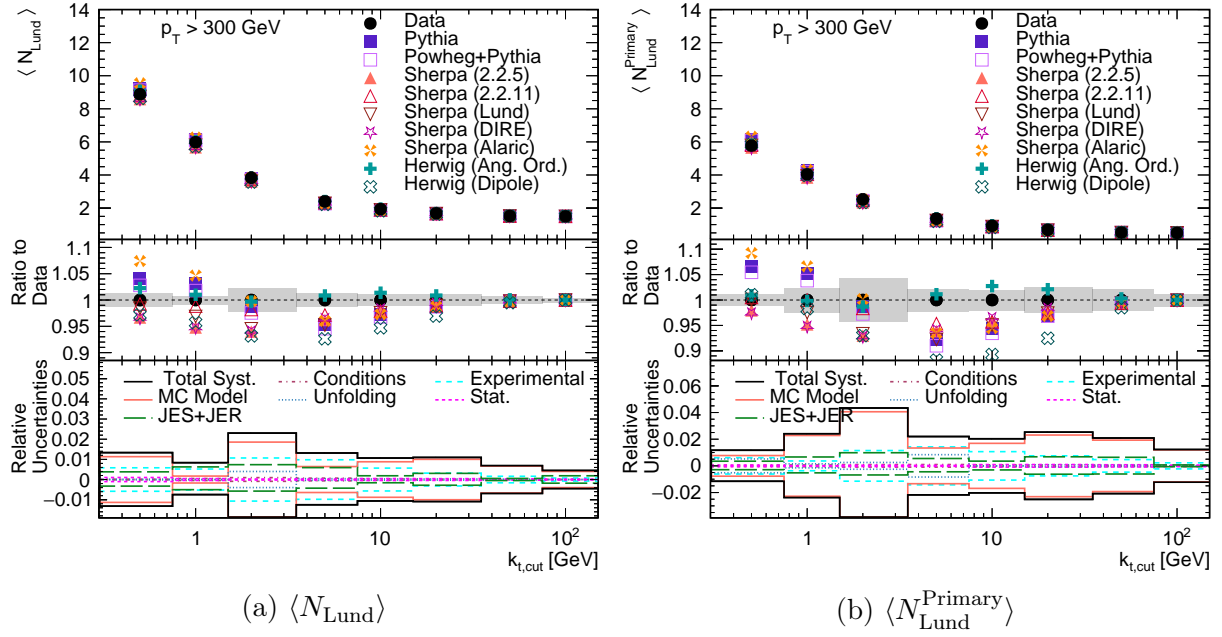
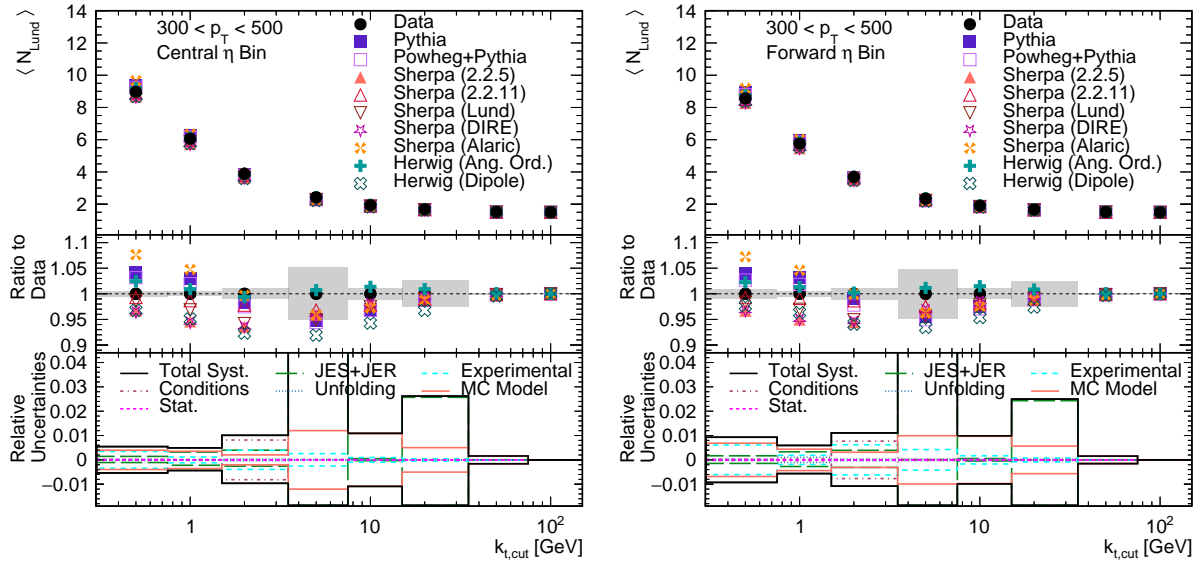
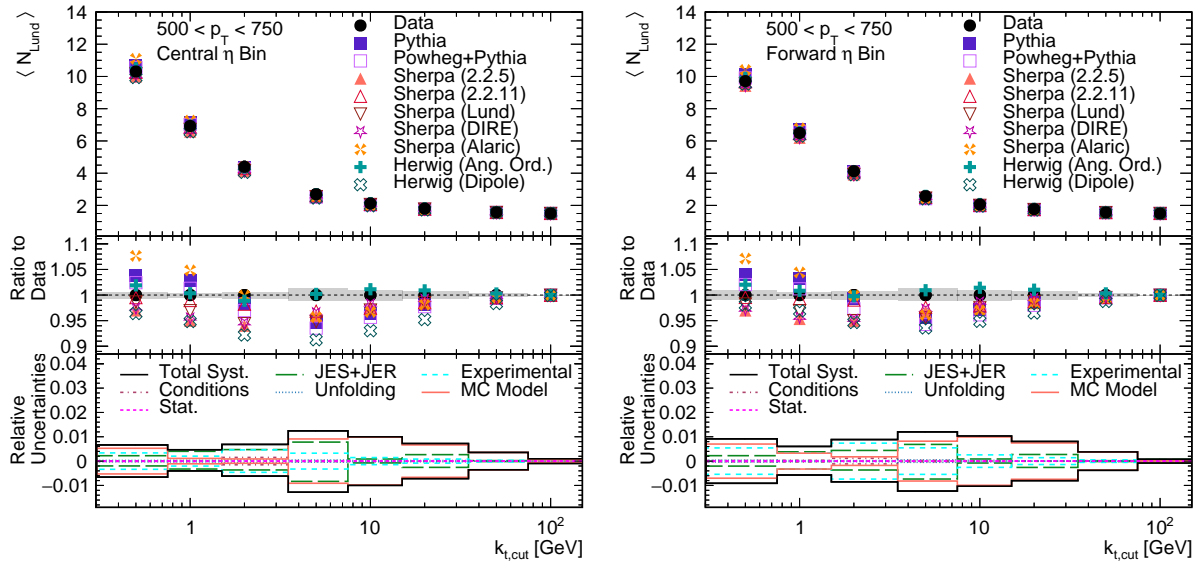


Figure 8.42: The (a) $\langle N_{\text{Lund}} \rangle$ and (b) $\langle N_{\text{Lund}}^{\text{Primary}} \rangle$ distributions as a function of the emission k_t requirement, for an inclusive p_T selection ($p_T > 300 \text{ GeV}$).

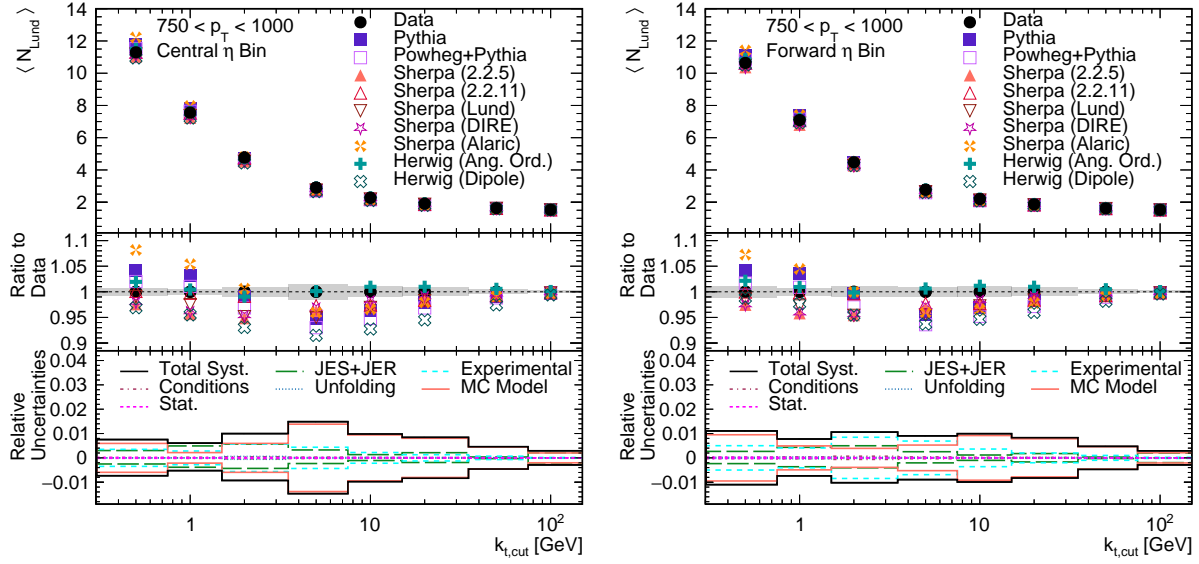


(a) $300 \text{ GeV} < p_T < 500 \text{ GeV}$, N_{Lund}

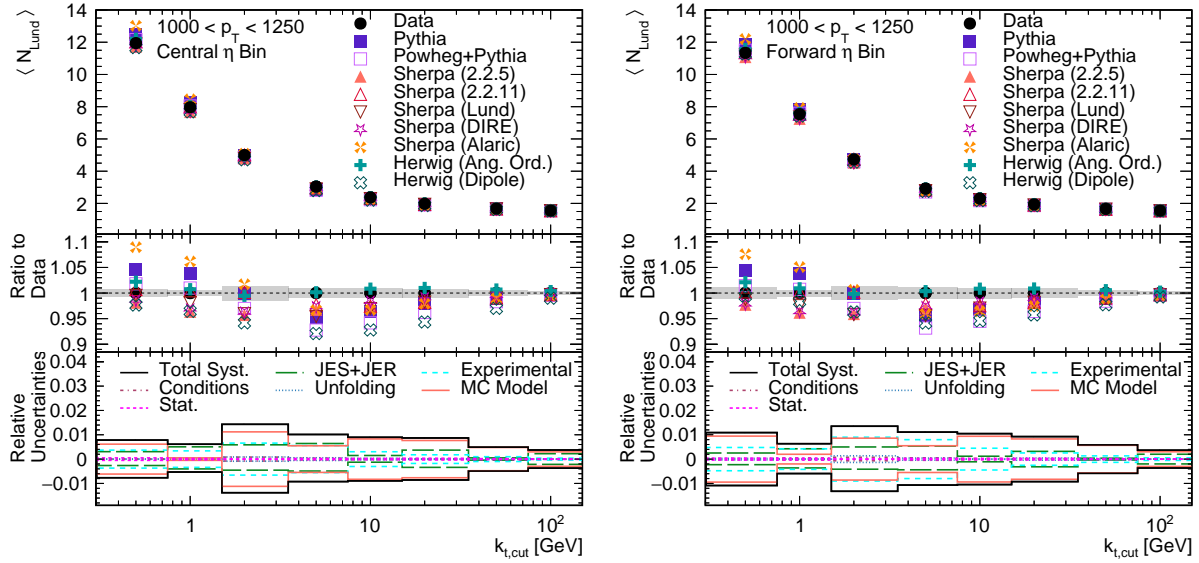


(b) $500 \text{ GeV} < p_T < 750 \text{ GeV}$, N_{Lund}

Figure 8.43: The average Lund multiplicity ($\langle N_{\text{Lund}} \rangle$) as a function of the k_t requirement is shown in bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area. Figure continued on next page.

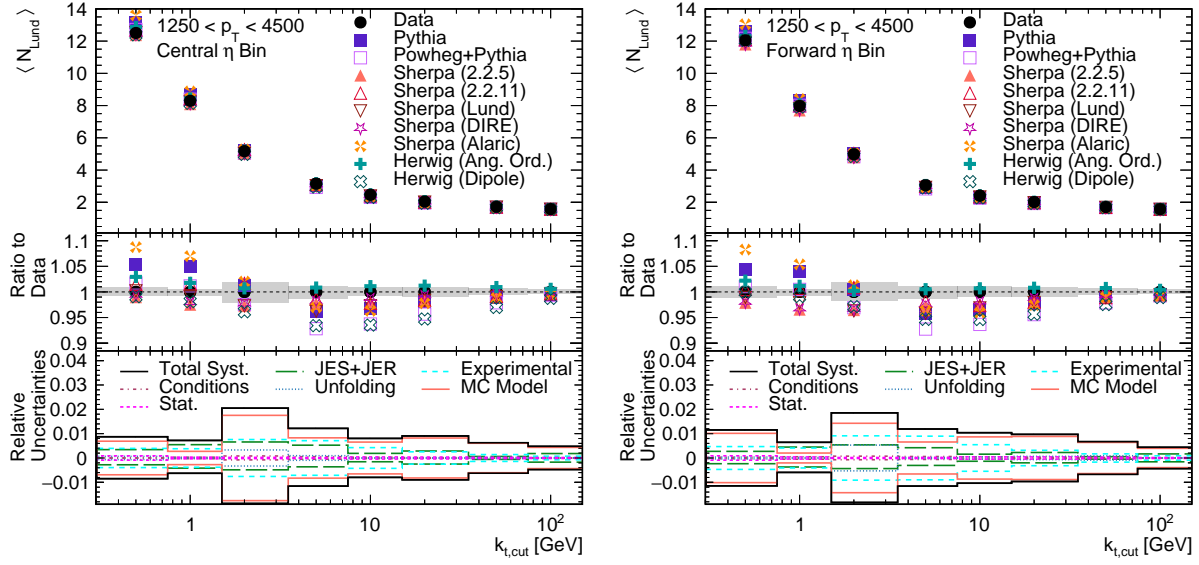


(c) $750 \text{ GeV} < p_T < 1000 \text{ GeV}$, N_{Lund}



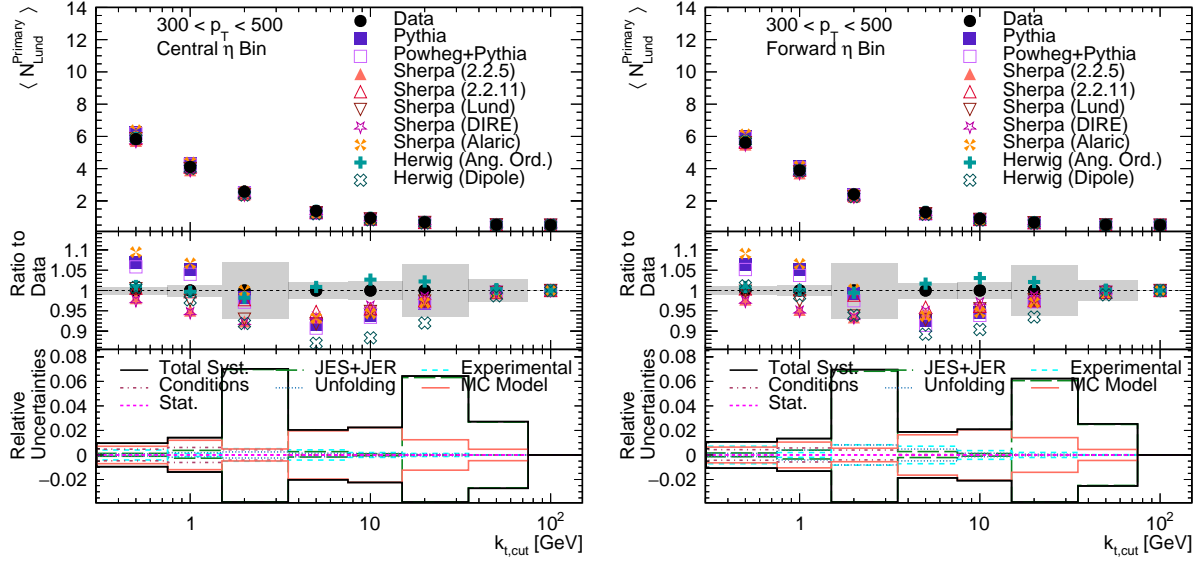
(d) $1000 \text{ GeV} < p_T < 1250 \text{ GeV}$, N_{Lund}

Figure 8.43: (continued) The average Lund multiplicity ($\langle N_{\text{Lund}} \rangle$) as a function of the k_t requirement is shown in bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area. Figure continued on next page.

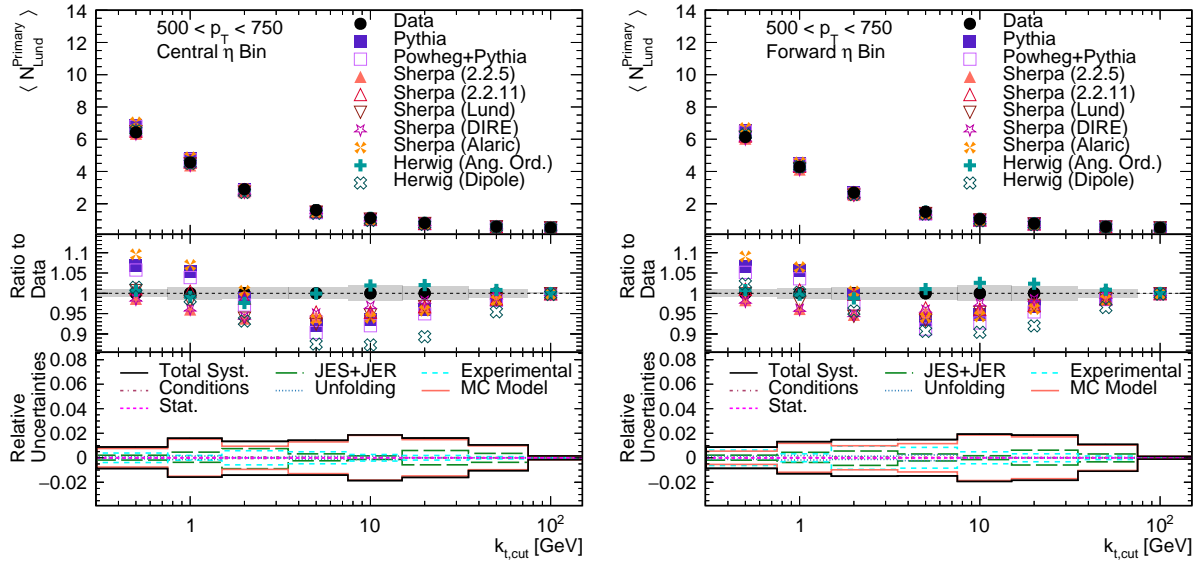


(e) $1250 \text{ GeV} < p_T < 4500 \text{ GeV}$, N_{Lund}

Figure 8.43: (continued) The average Lund multiplicity ($\langle N_{\text{Lund}} \rangle$) as a function of the k_t requirement is shown in bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area.

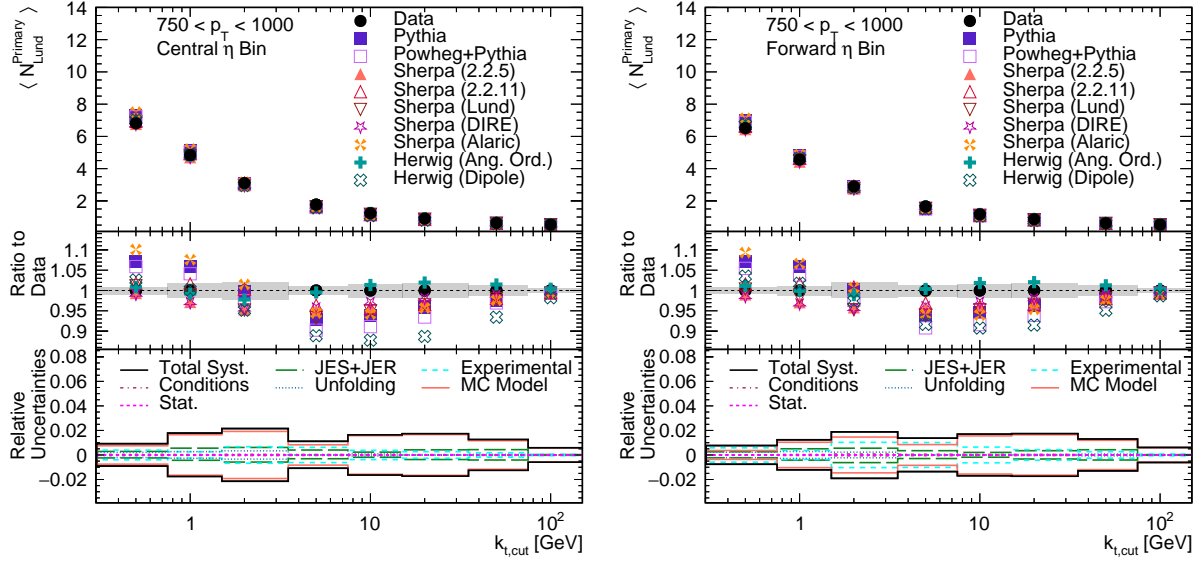


(a) $300 \text{ GeV} < p_T < 500 \text{ GeV}$, $N_{\text{Lund}}^{\text{Primary}}$

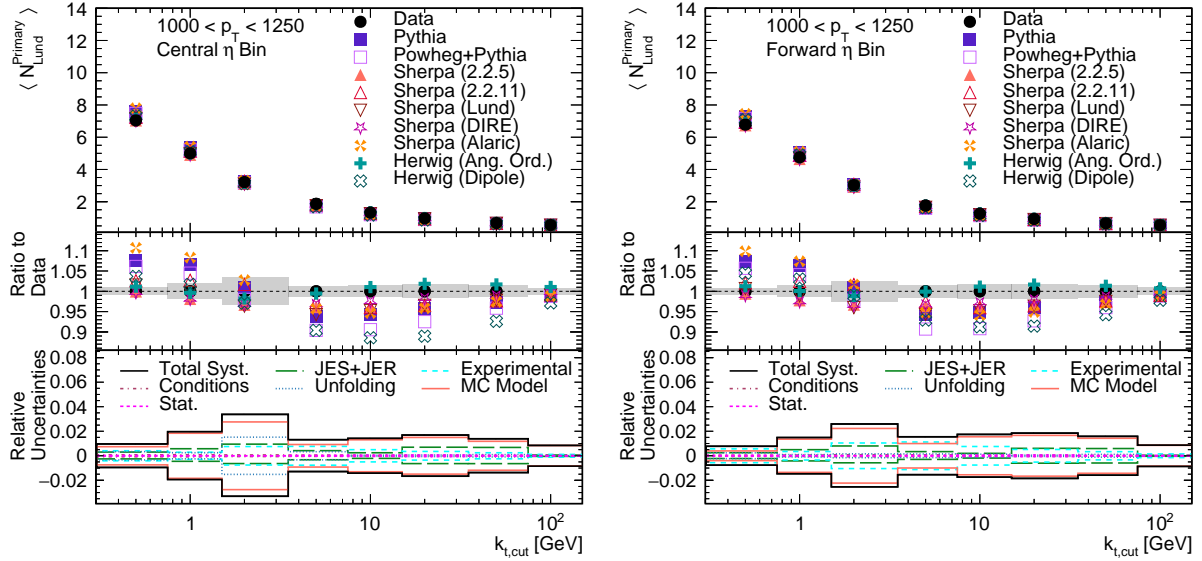


(b) $500 \text{ GeV} < p_T < 750 \text{ GeV}$, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.44: The average primary Lund multiplicity ($\langle N_{\text{Lund}}^{\text{Primary}} \rangle$) as a function of the k_t requirement is shown bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area. Figure continued on next page.

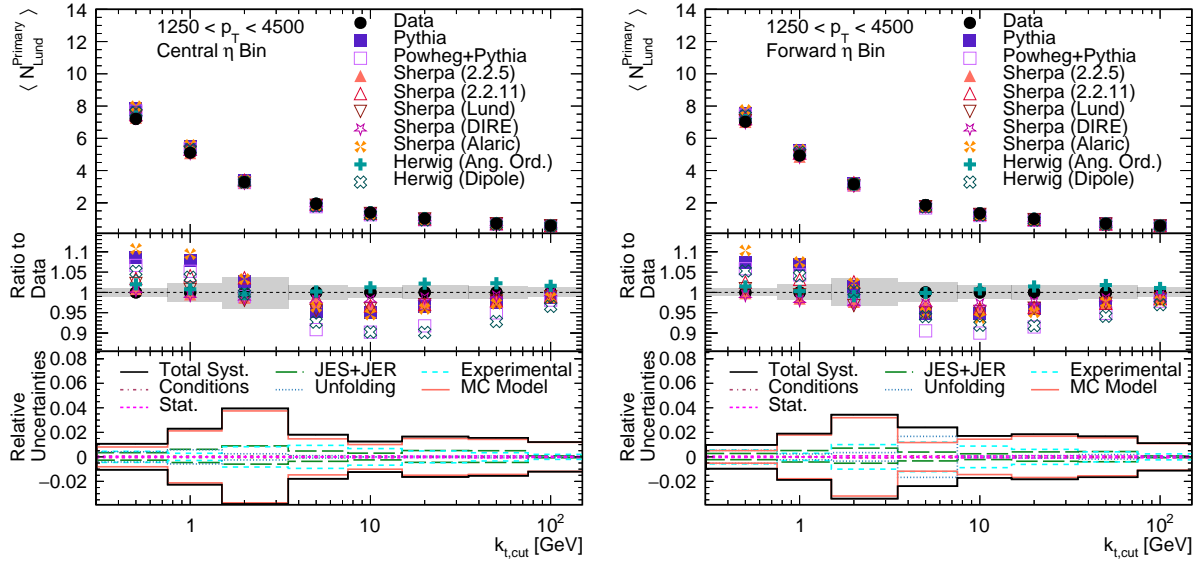


(c) $750 \text{ GeV} < p_T < 1000 \text{ GeV}$, $N_{\text{Lund}}^{\text{Primary}}$



(d) $1000 \text{ GeV} < p_T < 1250 \text{ GeV}$, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.44: (continued) The average primary Lund multiplicity ($\langle N_{\text{Lund}}^{\text{Primary}} \rangle$) as a function of the k_t requirement is shown bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area. Figure continued on next page.



(e) $1250 \text{ GeV} < p_T < 4500 \text{ GeV}$, $N_{\text{Lund}}^{\text{Primary}}$

Figure 8.44: (continued) The average primary Lund multiplicity ($\langle N_{\text{Lund}}^{\text{Primary}} \rangle$) as a function of the k_t requirement is shown bins of p_T . The unfolded data values are compared to predictions from several Monte Carlo generators. The corresponding total uncertainty on the measurement is shown as a shaded area.

CHAPTER 9

CONCLUSION

This thesis represents a global look at jets and their impact in ATLAS. From triggering to data analyses, jets are complicated objects which play an integral role in LHC physics.

The global Feature Extractor has been installed and participating in Run 3 since its beginning. It will remain installed throughout Run 3 and will continue to operate during Run 4 at the HL-LHC as part of the overall hardware calorimeter trigger upgrade. The Run 4 upgraded trigger system will include the existing FEX modules as well as additional modules which use System on Chips in their design. As gFEX is the first module to fully rely on the included SoC and its custom operating system, the functionality and design of gFEX will be particularly relevant as these systems continue the development of Run 4 prototypes. The work done on the gFEX custom operating system and monitoring processes as well as the lessons learned about firmware development and integration with the processing system form a baseline for the Run 4 systems, and help to inform their design and implementation. The calorimeter trigger will continue to be an important component of the ATLAS detector and boosted jets will continue to be important physics objects for many analyses, particularly those with hadronic final states.

Searches for Supersymmetry at the LHC have been performed for several years, and continue to be done. The fully hadronic search for supersymmetric partners to the electroweak gauge bosons presented in this thesis increased the exclusion limits all the way to a \sim TeV. This is an impressive increase in sensitivity and demonstrates the impact that hadronic final states can have. Above several TeV, SUSY begins to cause problems that it used to solve. Many remain convinced that Supersymmetry is the answer to several questions about the SM, and while I admit that it this could be the case, I am interested to see what other regimes have

to offer. I am convinced that the hadronic final state will continue to be as relevant and as powerful in other areas as it was in this analysis.

Jets, while being powerful tools to investigate different areas of physics, are themselves complicated and interesting objects. They tend to be difficult to reconstruct and complex to model. They are strongly coupled, non-perturbative systems which have a global impact on LHC physics due to the hadronic nature of the collider. The precision measurement of Lund multiplicities presented in this thesis provides a comparison between LHC data and the latest available predictions of this observable. This comparison could be used to constrain future Monte Carlo generators which hope to be more precise, and to decrease the modeling uncertainties which affect almost all LHC analyses.

The third run of the LHC will shape the near future of particle physics, but beyond this there are many exciting possibilities on the horizon. The field of particle physics is full of interesting problems to solve, hints of new physics, and exciting new experimental plans. I look forward to the new developments that will come, and I am excited to study them myself from new perspectives.

Appendices

APPENDIX A

GFEX MONITORED HARDWARE SENSORS

The gFEX DCS monitoring has a number of datapoints which are constantly monitored and which have alarm thresholds placed on them to protect the safety of the hardware. Specific datapoints require firmware implementations to monitor (the majority of which correspond to the SysMon implementation from Xilinx) and this has not been integrated yet into the gFEX firmware. As such, the sensors shown in Table A.1 and Table A.2 exist in the gFEX DCS project but are not displayed and do not have alarms set. This can be updated whenever the necessary changes are made to the gFEX firmware.

The remaining hardware sensors are both incorporated into the DCS project and actually measured and monitored with alarms. Most of the datapoints correspond to hardware sensors which are measured over the I2C bus on the gFEX board, as explained in Subsection 5.7.1. However, a subset of the DCS datapoints are instead calculated in the OS, specifically the OS Utilization datapoints which are listed in Table A.3. The remaining monitored sensors are listed in Table A.4 - Table A.11.

Health Flag Datapoints					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
led_all_pfpga_status	N/A	HealthFlag	×	✓	×
led_clock_generator	N/A	HealthFlag	×	✓	×
led_board_temp_alert	N/A	HealthFlag	×	✓	×
led_power_good	N/A	HealthFlag	×	✓	×
ipmc_communication	N/A	HealthFlag	×	✓	×
i2c_zynq_status	N/A	HealthFlag	×	✓	×
run_mode_on	N/A	HealthFlag	×	✓	×

Table A.1: Boolean Health Flags describing the state of the gFEX system

FPGA Monitoring					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
FPGA_Z	U89	temperature	×	✓	×
		voltage	×	✓	×
		current	×	✓	×
		LED heartbeat	×	✓	×
FPGA_A	U4	temperature	×	✓	×
		voltage	×	✓	×
		current	×	✓	×
		LED heartbeat	×	✓	×
FPGA_B	U5	temperature	×	✓	×
		voltage	×	✓	×
		current	×	✓	×
		LED heartbeat	×	✓	×
FPGA_C	U29	temperature	×	✓	×
		voltage	×	✓	×
		current	×	✓	×
		LED heartbeat	×	✓	×

Table A.2: FPGA Monitoring

Operating System Utilization					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
Disk Space Utilization	N/A	OS	×	✓	✓
Memory (RAM) Utilization	N/A	OS	×	✓	✓
CPU Utilization	N/A	OS	×	✓	✓

Table A.3: OS-Utilization Sensors

ADM7417 Temperature Sensors					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
Temperature Sensor A	U84	temperature	✓	✓	✓
Temperature Sensor B	U82	temperature	✓	✓	✓
Temperature Sensor C	U83	temperature	✓	✓	✓
Temperature Sensor Z	U87	temperature	✓	✓	✓

Table A.4: ADM7417 Temperature Sensors

DCDC Converters					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
DCDC Converter Positive	U11	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
DCDC Converter Negative	U11	voltage	✓	✓	✓
DCDC Monitoring Chip	U81	voltage	✓	✓	✓
		current	✓	✓	✓

Table A.5: DCDC Converter Sensors

pFPGA Power Modules					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
pFPGA A VCCINT	U122	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA A MGTAVCC	U77	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA A MGTAVTT	U30	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA B VCCINT	U123	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA B MGTAVCC	U124	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA B MGTAVTT	U40	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA C VCCINT	U126	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA C MGTAVCC	U125	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓
pFPGA C MGTAVTT	U44	temperature	✓	✓	✓
		voltage	✓	✓	✓
		current	✓	✓	✓

Table A.6: pFPGA Power Modules

Zynq and Global Power Modules					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
Zynq VCCINT	U73	temperature voltage current	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
Zynq MGTAVCC	U59	temperature voltage current	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
Zynq MGTAVTT	U66	temperature voltage current	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
Global 3.3V	U59	temperature voltage current	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
Global 2.5V	U55	temperature voltage current	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
Global 1.8V	U73	temperature voltage current	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
DDR4 1.2V	U66	temperature voltage current	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓

Table A.7: Zynq and Global Power Modules

Zynq & Spare MiniPODs					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
TX_L1TOPO_Z	U24	temperature optical power	✓ ✓	✓ ✓	✓ ✓
TX_GLOBAL_Z	U56	temperature optical power	✓ ✓	✓ ✓	✓ ✓
TX_FELIX_Z	U3	temperature optical power	✓ ✓	✓ ✓	✓ ✓
RX_FELIX_Z	U72	temperature optical power LOS	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
RX_CALO_S	U91	temperature optical power LOS	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓

Table A.8: Zynq & Spare RX and TX MiniPODs

pFPGA A MiniPODs					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
TX_L1TOPO_A	U32	temperature	✓	✓	✓
		optical power	✓	✓	✓
TX_L1TOPO_A	U25	temperature	✓	✓	✓
		optical power	✓	✓	✓
RX_CALO_A	U96	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_A	U102	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_A	U103	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_A	U104	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_A	U105	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_A	U106	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_A	U107	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_A	U97	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓

Table A.9: pFPGA A MiniPODs

pFPGA B MiniPODs					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
TX_L1TOPO_B	U33	temperature	✓	✓	✓
		optical power	✓	✓	✓
TX_L1TOPO_B	U27	temperature	✓	✓	✓
		optical power	✓	✓	✓
RX_CALO_B	U98	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_B	U100	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_B	U101	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_B	U108	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_B	U109	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_B	U111	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_B	U112	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_B	U113	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓

Table A.10: pFPGA B MiniPODs

pFPGA C MiniPODs					
Sensor Name	Schematic Label	Type	I2C	DCS	DCS Alarms
TX_L1TOPO_C	U34	temperature	✓	✓	✓
		optical power	✓	✓	✓
TX_L1TOPO_C	U42	temperature	✓	✓	✓
		optical power	✓	✓	✓
RX_CALO_C	U114	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_C	U115	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_C	U116	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_C	U117	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_C	U118	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_C	U119	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_C	U120	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓
RX_CALO_C	U90	temperature	✓	✓	✓
		optical power	✓	✓	✓
		LOS	✓	✓	✓

Table A.11: pFPGA C MiniPODs

APPENDIX B

GFEX CLOCK CHIP RELEVANT REGISTERS

The SI5345 clock chip on the gFEX board is a jitter attenuating clock multiplier which uses digital signal processing phase locked loop (DSPLL) and MultiSynth technologies to generate any frequency of clock for applications that require an extremely precise clock signal with very low jitter. The chip offers free-run, synchronous, and holdover modes of operation, and can generate any combination of output frequencies from any input frequency [182]. The SI5345 has 4 inputs, 5 MultiSynths, and 10 possible outputs. On the gFEX board input 0 corresponds to the TTC clock signal derived from FELIS, input 1 is an oscillator signal, input 2 is the external signal to RF connectors, and input 3 is a feedback signal. The TTC signal is the desired input clock, and the feedback clock is necessary operationally.

In the gFEX firmware the clock interrupt signal (INTRb) depends on the status of the clock chip. The b in this signal name indicates that a high signal (bit = 1) is good and a low signal (bit = 0) is bad. It was discovered that sticky error bits in the clock chip caused this interrupt signal to go low, and we realized that these sticky error bits needed to be cleared. Mostly these errors indicated Out of Frequency (OOF), Loss of Signal (LOS), and occasionally Loss of Lock (LOL) errors, and were often occurring for IN1 and IN2. The chip used on the gFEX board is Revision B, which is an important fact to note, because Revision B of the SI5345 clock chip has an erratum which states that "No individual LOS or OOF sticky status bit may be cleared when any other LOS or OOF alarm is asserted" [181]. This meant that the unneeded inputs IN1 and IN2 needed to be masked because any errors from these inputs made it impossible to clear any of the sticky error bits in the clock chip. This was done manually to start, but was eventually made part of the clock chip configuration, which is done in the Operating System, as part of the boot procedure.

The ClockBuilder Pro software (CBPro), which is provided by the company which produces the chip, can be used to generate a project¹ which produces a configuration file² with a list of register commands which program the chip in the desired way. The gFEX configuration file masks IN1 and IN2, and additionally configures the input frequency to 240 MHz and the output frequency to 280 MHz.

After this configuration is run, the sticky error bits listed in Table B.1 must be cleared. They must also be cleared after the gFEX board is reset which occurs at the beginning of every ATLAS run, during any clock switches that occur throughout the run, and during the L1Calo TTC-restart procedure which is manually triggered in the ATLAS control room. The clock chip registers are accessible over I2C, so the same Python scripts as those used for monitoring are used to access the clock chip. The non-sticky error bits listed in Table B.2 can be checked for reference and debugging purposes, and the additional registers listed in Table B.3 are useful for validation purposes. For example, the DESIGN_ID registers verify that the correct project configuration file has been loaded, as it sets the DESIGN_ID to the unique value of "GF240280".

1. The gFEX project can be found in the gfex-management-scripts repository.

2. The gFEX configuration file is also in the gfex-management-scripts repository, as well as the meta-l1calo repository for the OS.

Register Name	Address [bits]	Functionality
OOF_FLG	0x0012 [7:4]	If the clock input is OOF (out of frequency) for a given input
LOS_FLG	0x0012 [3:0]	If the clock input is in LOS (loss of signal) for a given input
HOLD_FLG	0x0013 [5]	If the DSPLL is in holdover or free run
LOL_FLG	0x0013 [1]	If the DSPLL is not locked
CAL_PLL_FLG	0x0014 [5]	If the internal calibration was busy
SYSINCAL_FLG	0x0011 [0]	Asserted when in calibration
LOSXAXB_FLG	0x0011 [1]	LOS status monitor for XAXB.
XAXB_ERR_FLG	0x0011 [3]	If there is a problem locking to XAXB input signal
SMBUS_TIMEOUT_FLG	0x0011 [5]	If SMBUS timeout error

Table B.1: SI5345 Sticky Error Bits

Register Name	Address [Bits]	Functionality
OOF	0x000D [7:4]	If the clock input is OOF (out of frequency) for a given input (non-sticky)
LOS	0x000D [3:0]	If the clock input is in LOS (loss of signal) for a given input (non-sticky)
LOL	0x000E [1]	If the DSPLL is not locked (non-sticky)
SYSINCAL	0x000C [0]	Asserted when in calibration
LOSXAXB	0x000C [1]	LOS status monitor for the STAL or REFCLK at the XA/XB pins
XAXB_ERR	0x000C [3]	If problem locking to the XAXB input
SMBUS_TIMEOUT	0x000C [5]	If there is an SMBus timeout error

Table B.2: SI5345 Error Bits

Register Name	Address [Bits]	Functionality
PN_BASE (0x45)	0x0002 [7:0]	Four-digit "base" part number
PN_BASE (0x53)	0x0002 [15:8]	Four-digit "base" part number
DEVICE_REV	0x0005 [7:0]	One ASCII char, 0=A, 1=B, etc.
DESIGN_ID0 (G)	0x026B [7:0]	ASCII char, defined in CBPro by user
DESIGN_ID1 (F)	0x026C [15:8]	ASCII char, defined in CBPro by user
DESIGN_ID2 (2)	0x026D [23:16]	ASCII char, defined in CBPro by user
DESIGN_ID3 (4)	0x026E [31:24]	ASCII char, defined in CBPro by user
DESIGN_ID4 (0)	0x026F [39:32]	ASCII char, defined in CBPro by user
DESIGN_ID5 (2)	0x0270 [47:40]	ASCII char, defined in CBPro by user
DESIGN_ID6 (8)	0x0271 [55:48]	ASCII char, defined in CBPro by user
DESIGN_ID7 (0)	0x0272 [63:56]	ASCII char, defined in CBPro by user

Table B.3: Other useful SI5345 registers.

APPENDIX C

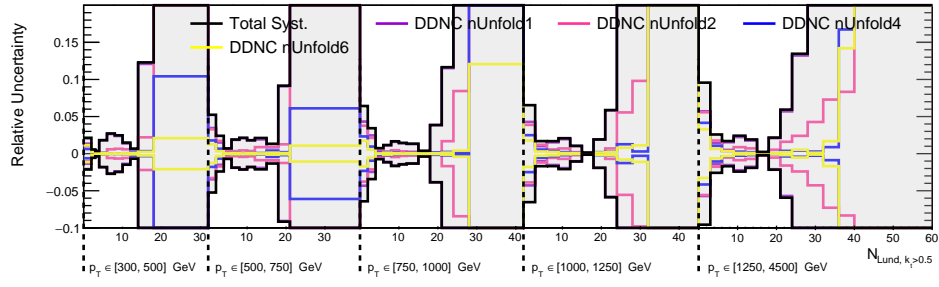
OPTIMIZATIONS OF ITERATIONS IN UNFOLDING

This appendix describes the optimization of the number of iterations used for the unfolding procedure as defined in Section 8.6. This section provides additional detail on the unfolding procedure.

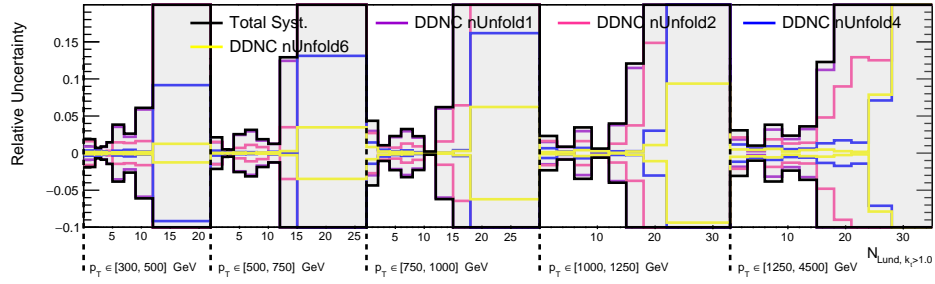
The number of iterations is chosen such that the data-driven non-closure (DDNC), statistical, and modelling uncertainties are minimized throughout the unfolding procedure. The DDNC uncertainty is shown in Figure C.1 and Figure C.2. The unfolding statistical uncertainties are in Figure C.3 and Figure C.6, and the modelling uncertainty, which is the dominant uncertainty in the measurement, in Figure C.7-Figure C.8.

Generally the data driven non closure uncertainty can be minimized by increasing the number of unfolding iterations, while the opposite is true of the statistical uncertainties. In this instance, the difference between the different number iterations is generally small. 1, 2, 4, and 6 iterations were studied. The total uncertainty for the examined sources listed above for unfolding iterations of 2 and 4 can be found respectively in Figure C.9 and Figure C.12.

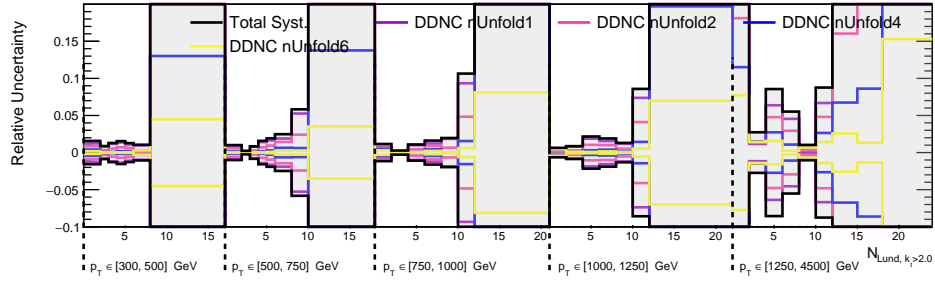
The uncertainties are very similar between 2 and 4 iterations, such that a subset of bins has a lower uncertainty with 2 and a different subset with 4. Since these options are so similar the number of iterations which minimizes the data-driven-non-closure was chosen. Based on these results, 4 iterations of unfolding were used to perform the analysis.



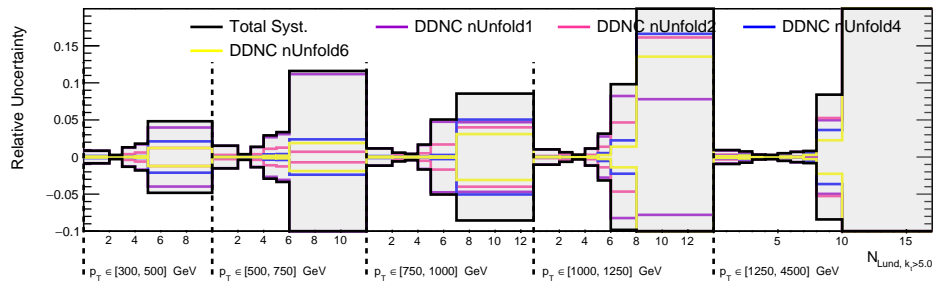
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

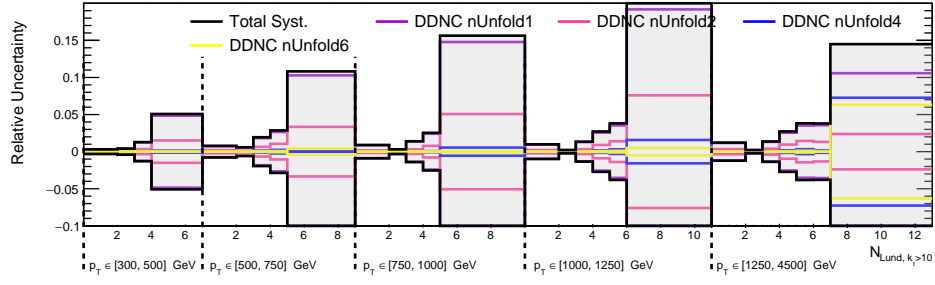


(c) $k_t > 2$ GeV, N_{Lund}

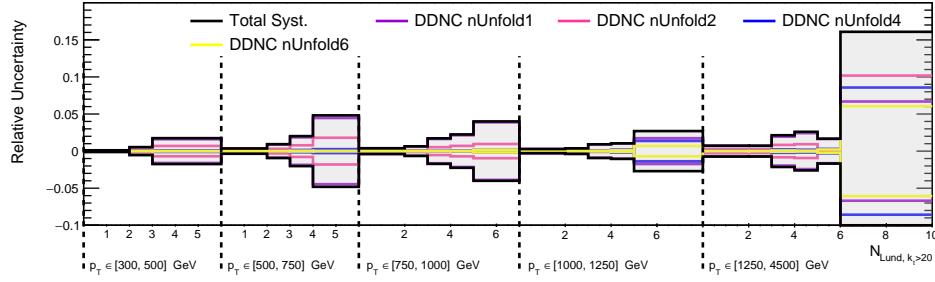


(d) $k_t > 5$ GeV, N_{Lund}

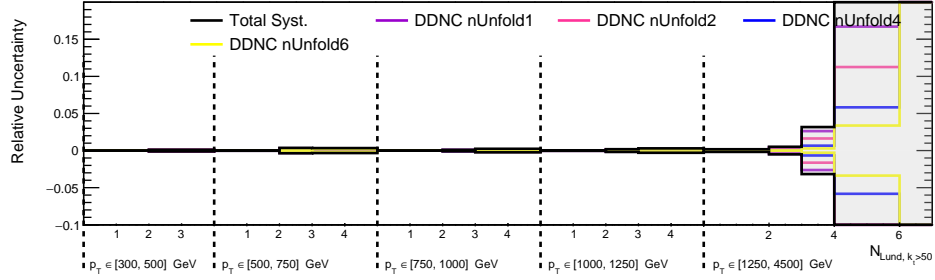
Figure C.1: The data-driven non-closure uncertainty for Lund multiplicity with various k_t cuts. Figure continued on next page.



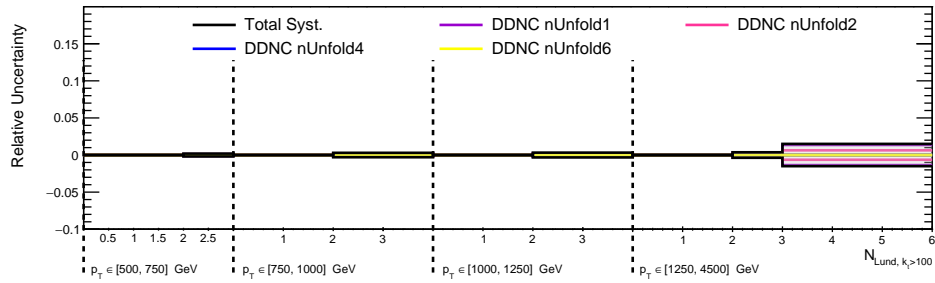
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

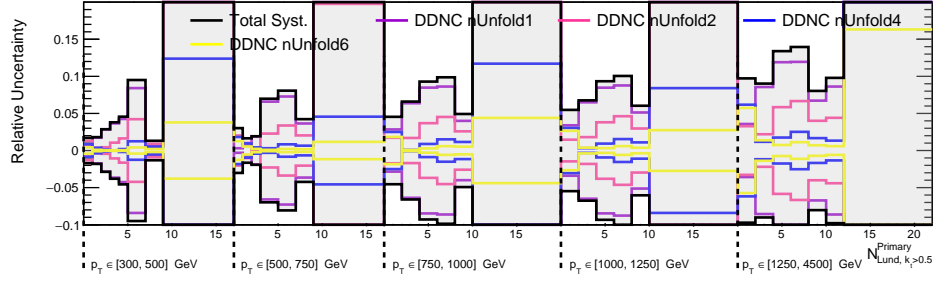


(g) $k_t > 50$ GeV, N_{Lund}

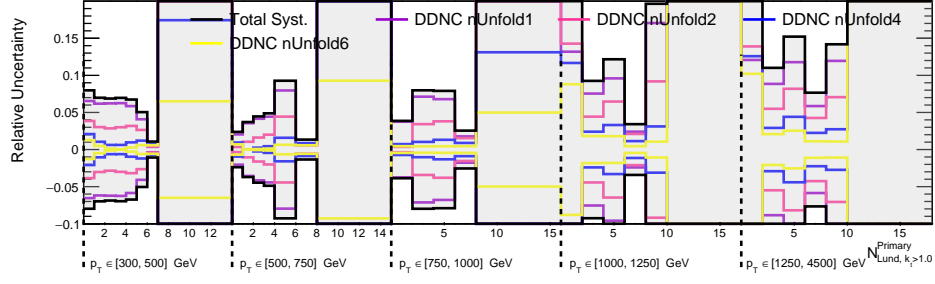


(h) $k_t > 100$ GeV, N_{Lund}

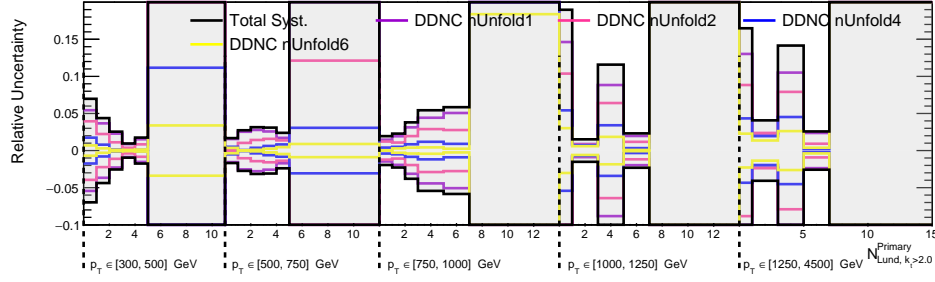
Figure C.1: (continued) The data-driven non-closure uncertainty for Lund multiplicity with various k_t cuts.



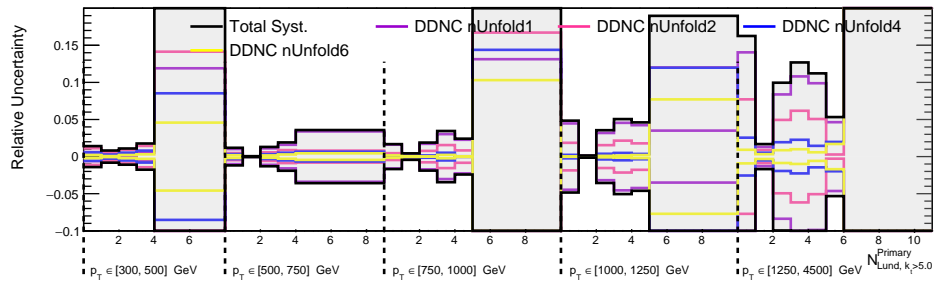
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

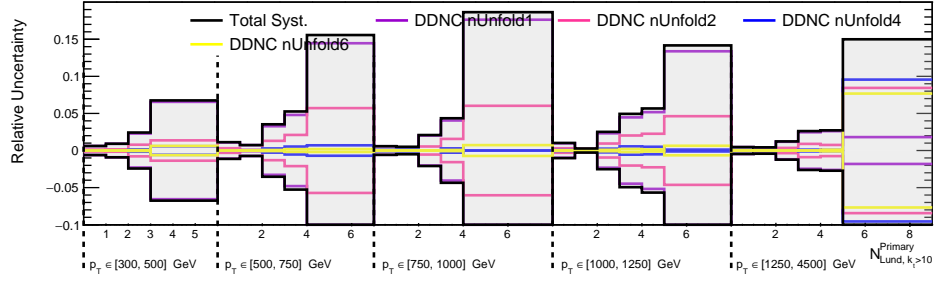


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

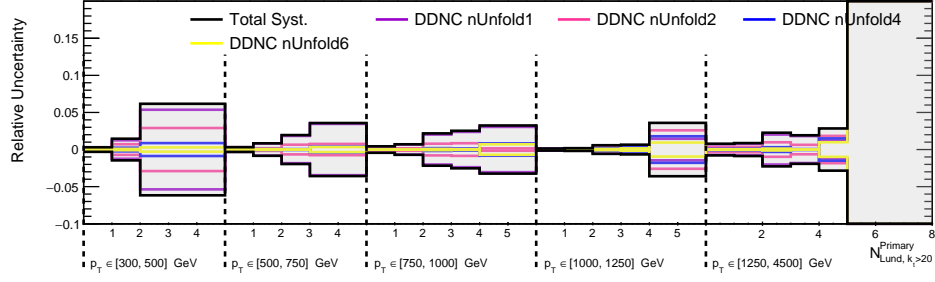


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

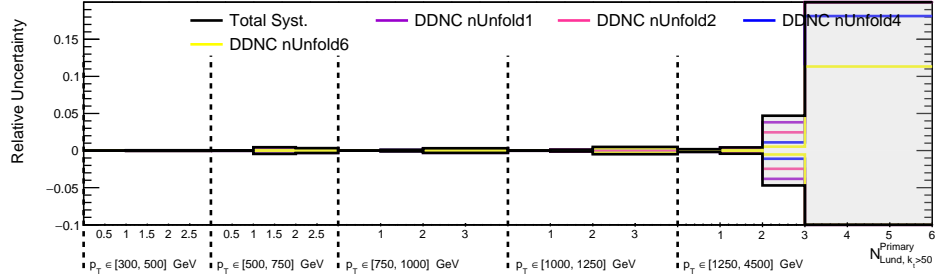
Figure C.2: The data-driven non-closure uncertainty for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



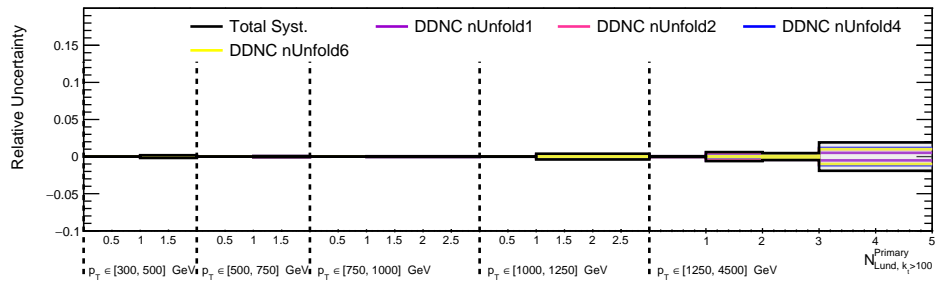
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

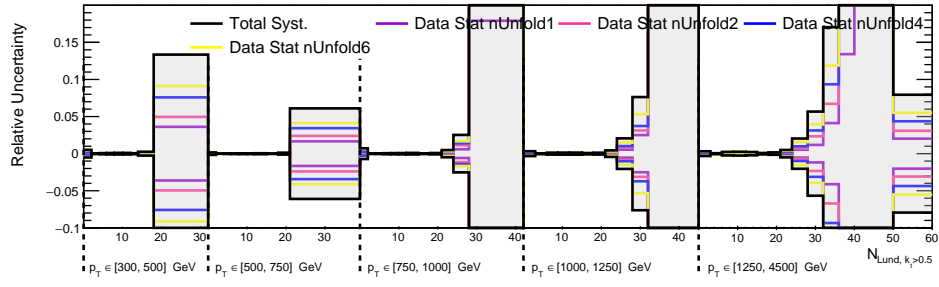


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

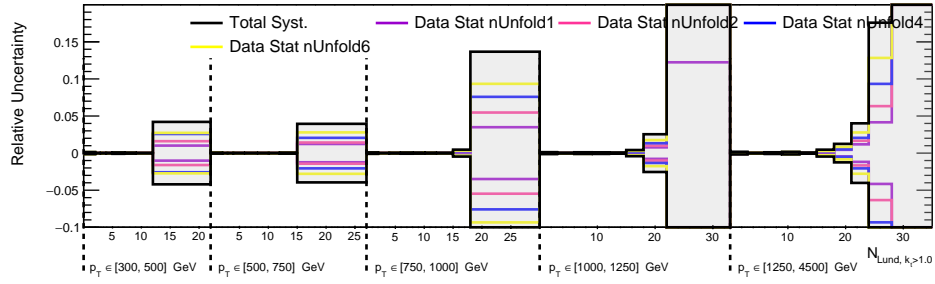


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

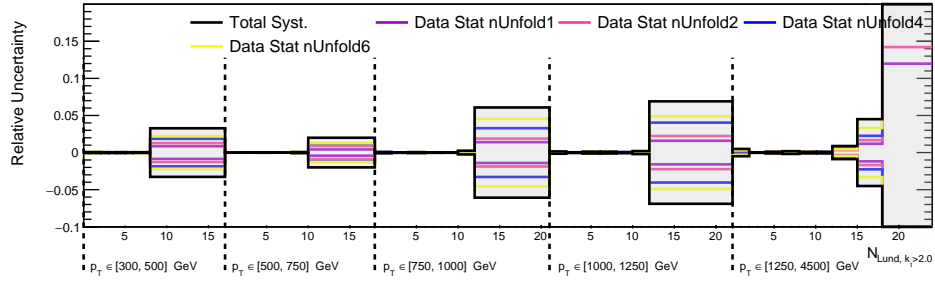
Figure C.2: (continued) The data-driven non-closure uncertainty for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



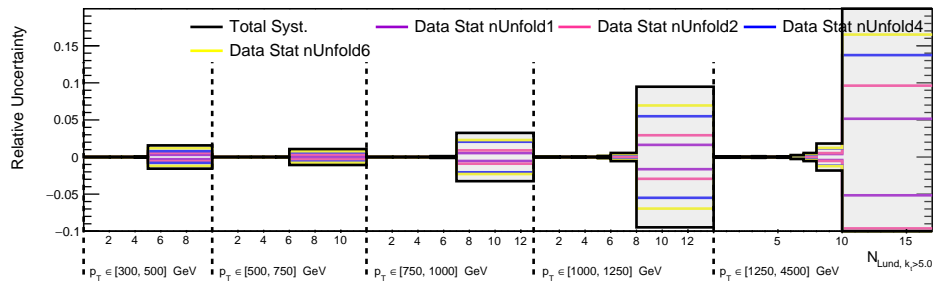
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

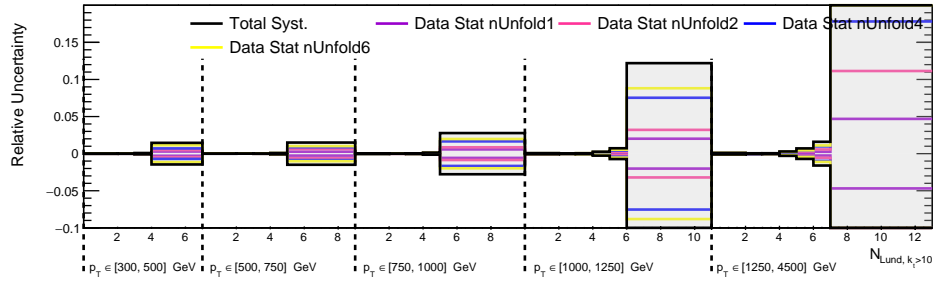


(c) $k_t > 2$ GeV, N_{Lund}

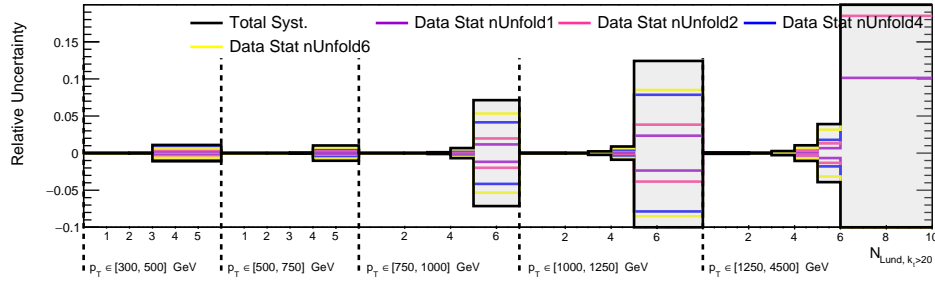


(d) $k_t > 5$ GeV, N_{Lund}

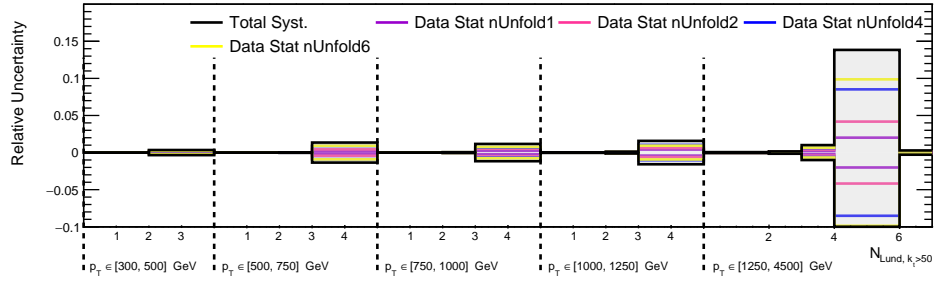
Figure C.3: The statistical uncertainty from data for Lund multiplicity with various k_t cuts. Figure continued on next page.



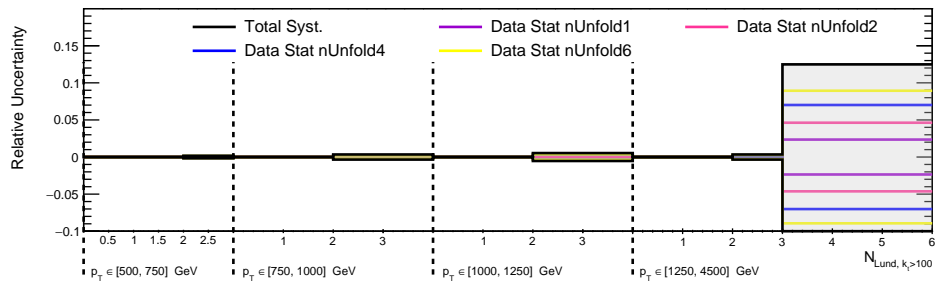
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

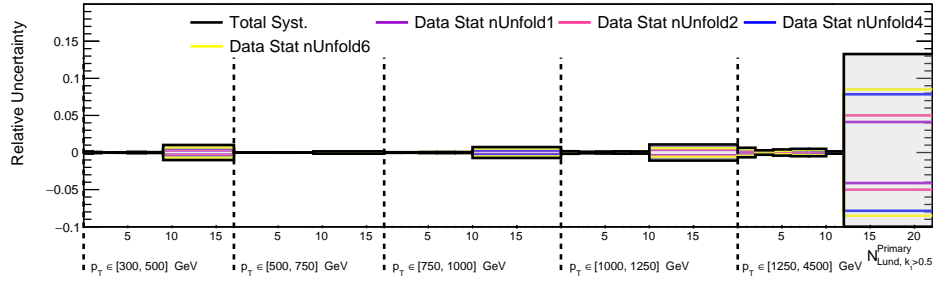


(g) $k_t > 50$ GeV, N_{Lund}

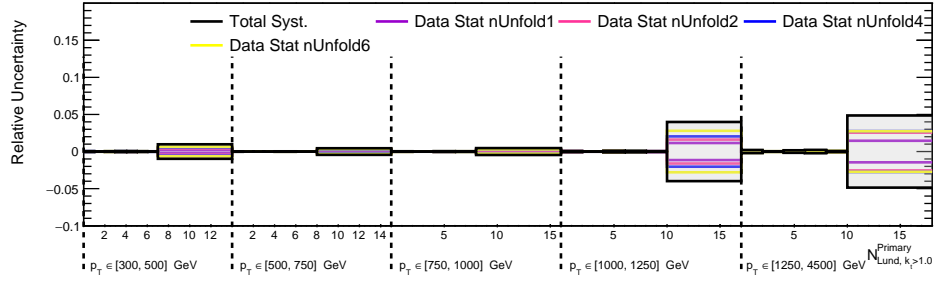


(h) $k_t > 100$ GeV, N_{Lund}

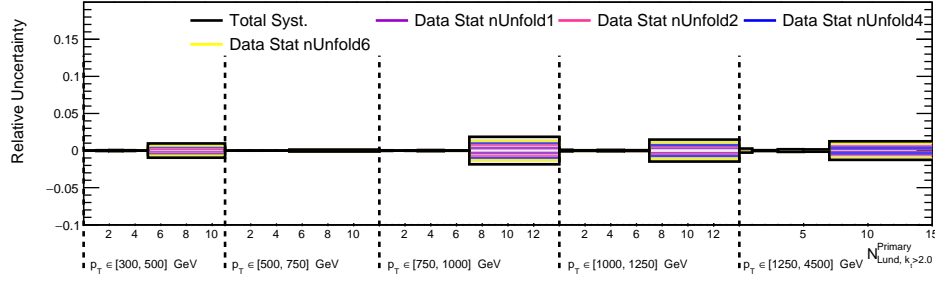
Figure C.3: (continued) The statistical uncertainty from data for Lund multiplicity with various k_t cuts.



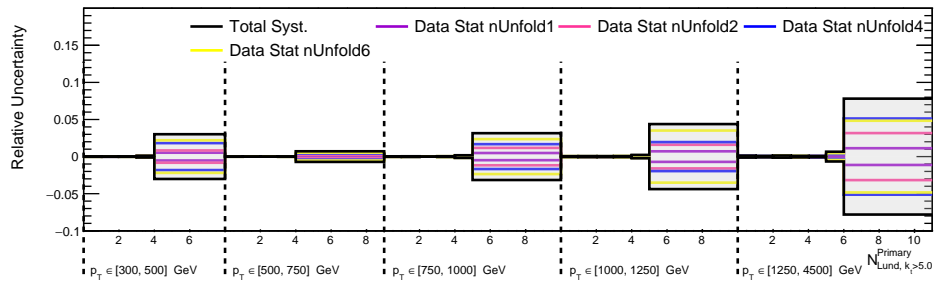
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 0.5$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 1.0$

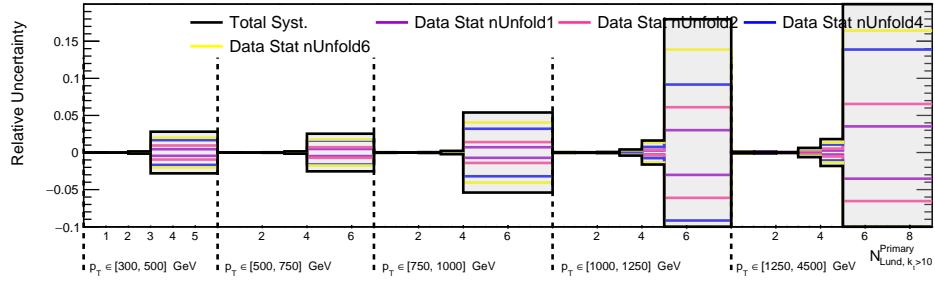


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 2.0$

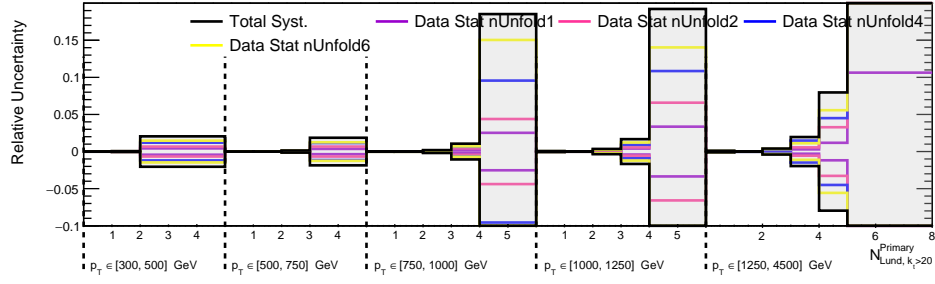


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 5.0$

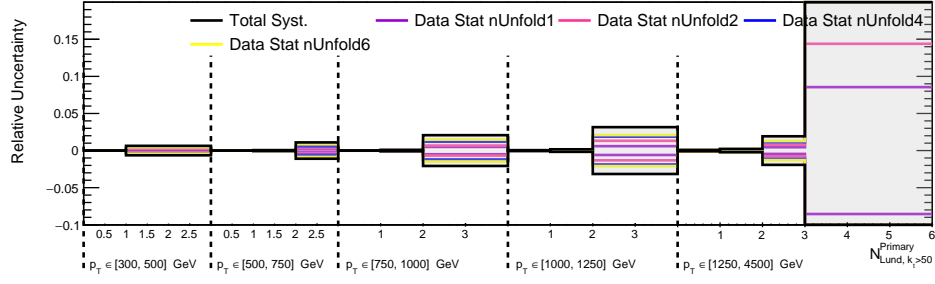
Figure C.4: The statistical uncertainty from data for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



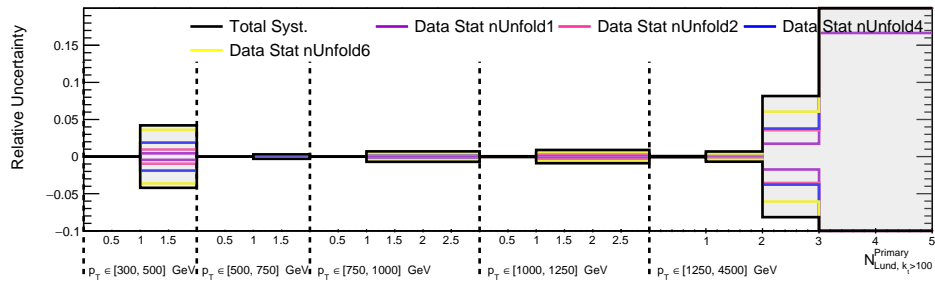
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

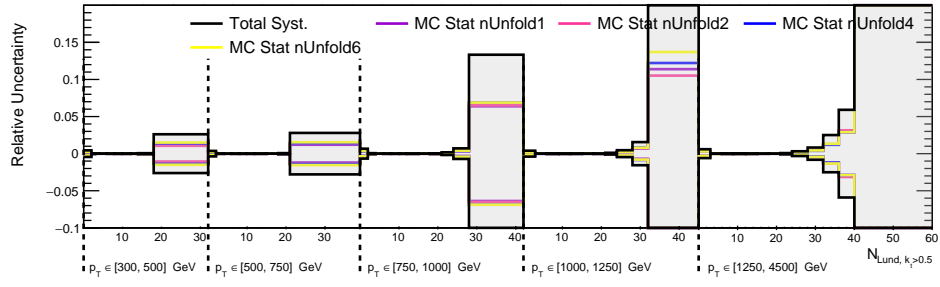


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

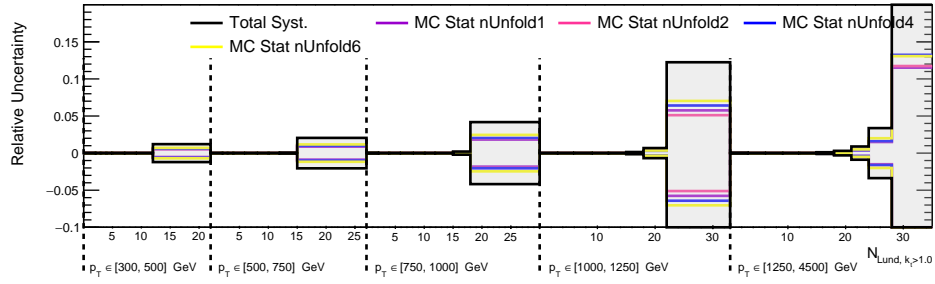


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

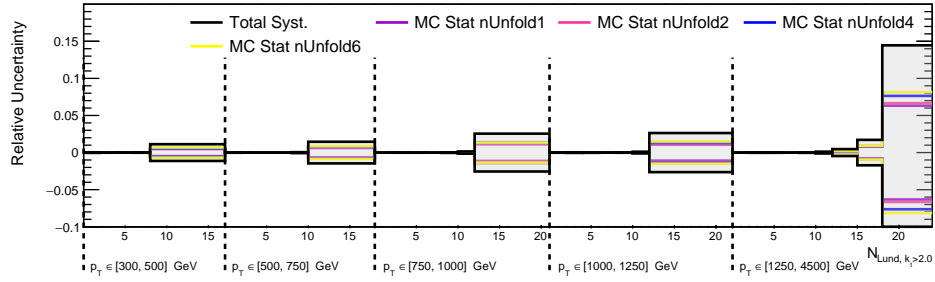
Figure C.4: (continued) The statistical uncertainty from data for primary Lund multiplicity with various k_t cuts.



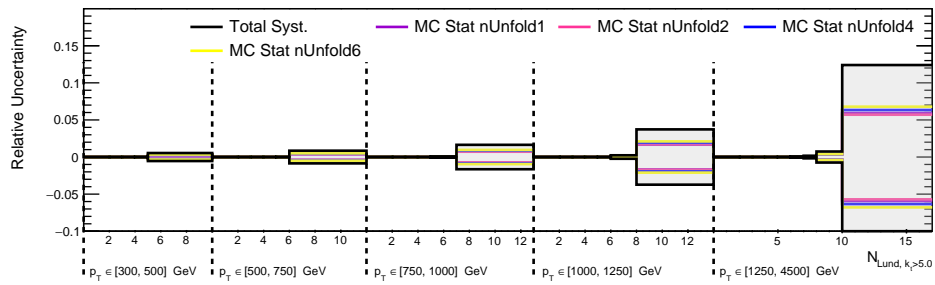
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

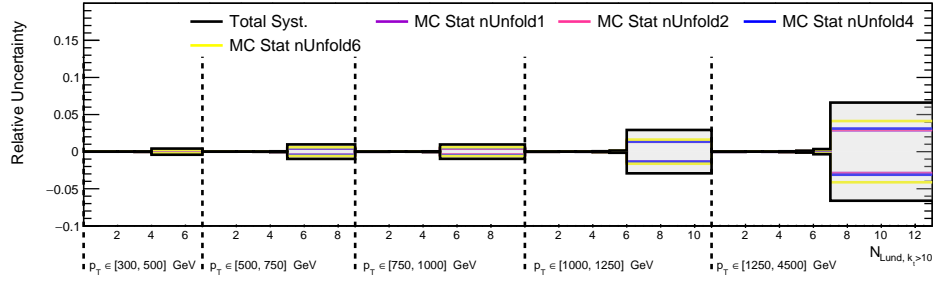


(c) $k_t > 2$ GeV, N_{Lund}

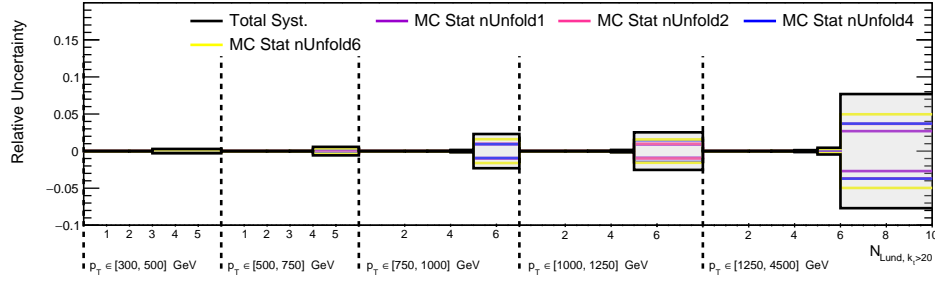


(d) $k_t > 5$ GeV, N_{Lund}

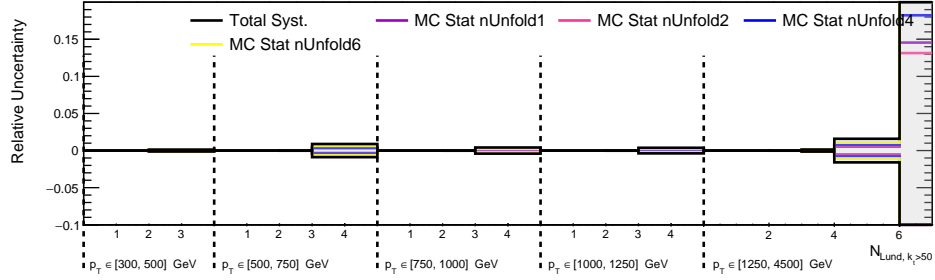
Figure C.5: The statistical uncertainty from MC for Lund multiplicity with various k_t cuts. Figure continued on next page.



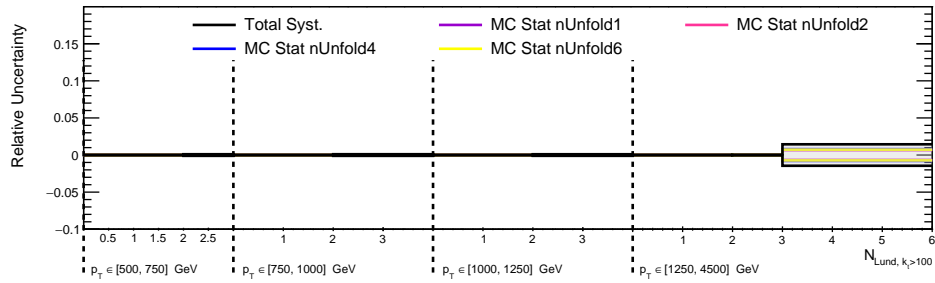
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

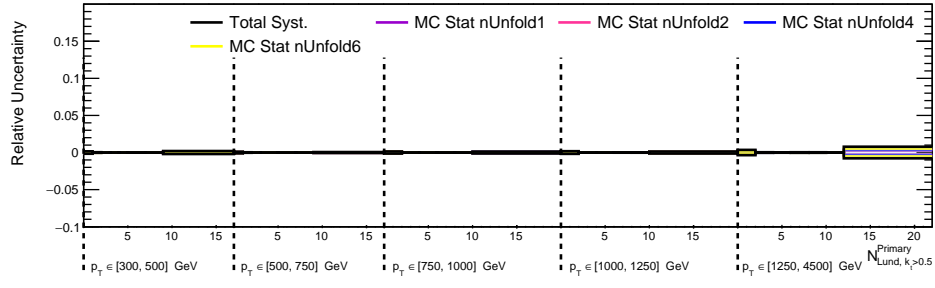


(g) $k_t > 50$ GeV, N_{Lund}

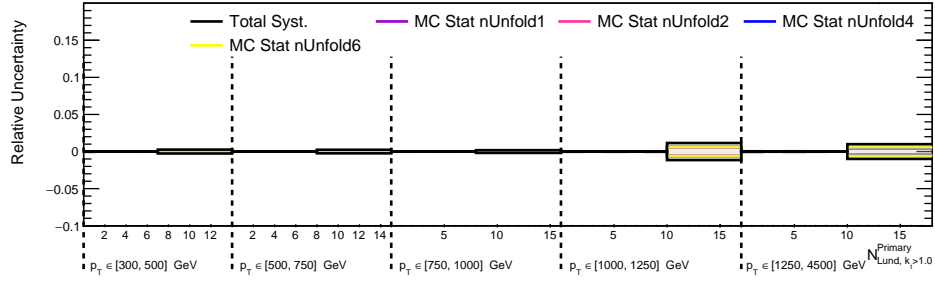


(h) $k_t > 100$ GeV, N_{Lund}

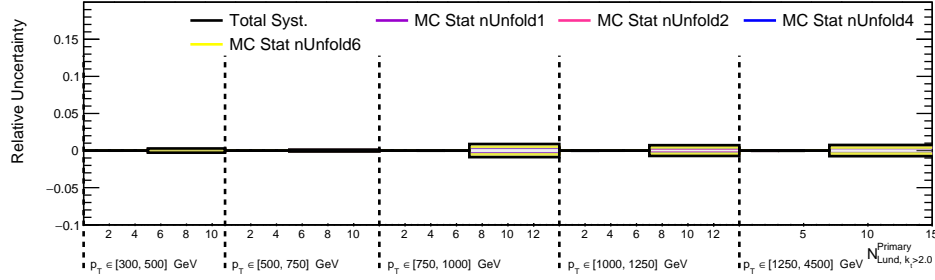
Figure C.5: (continued) The statistical uncertainty from MC for Lund multiplicity with various k_t cuts.



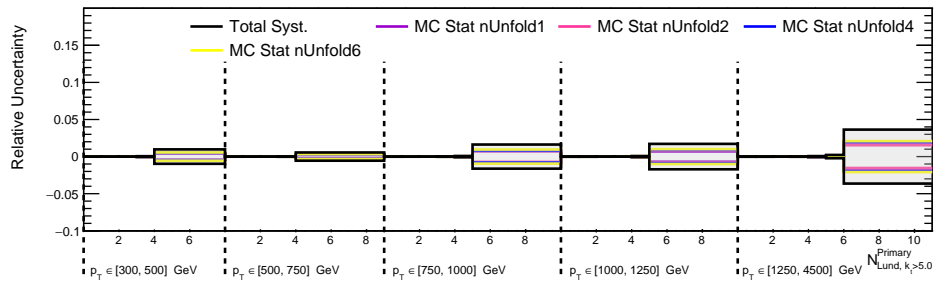
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

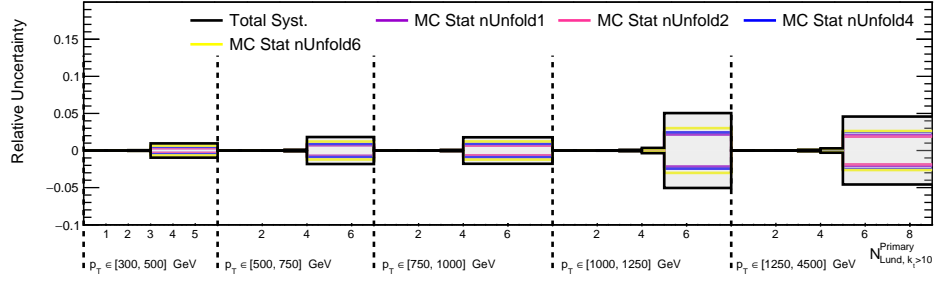


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

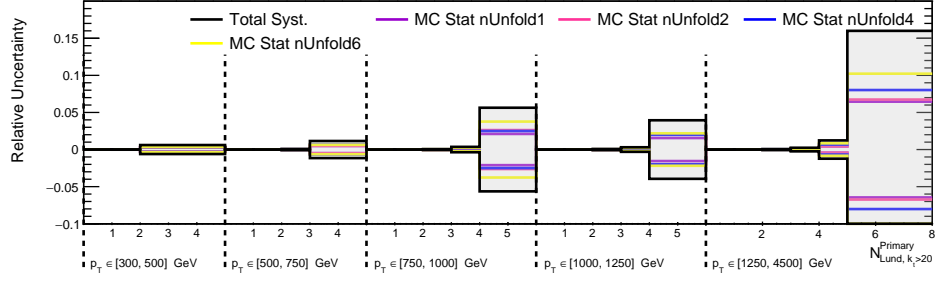


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

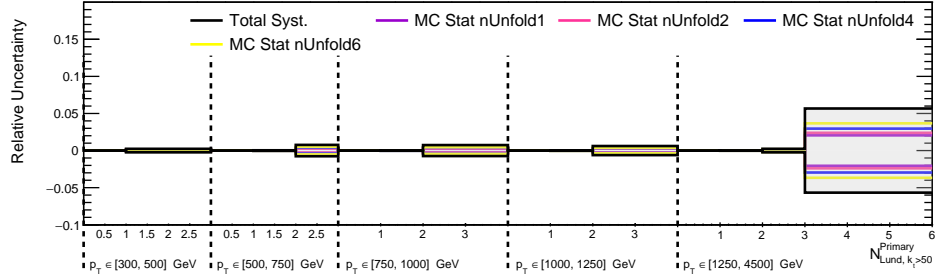
Figure C.6: The statistical uncertainty from MC for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



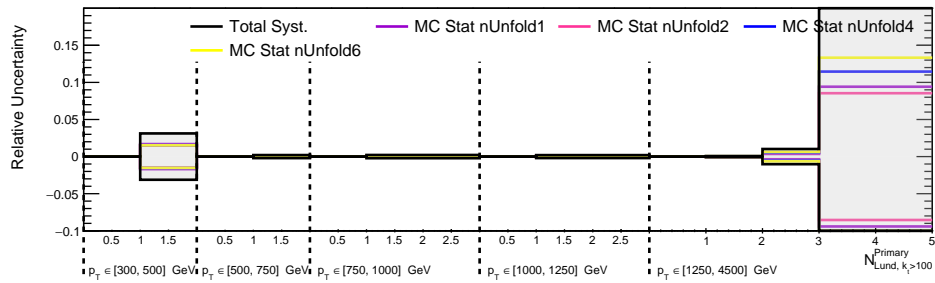
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 10$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 20$

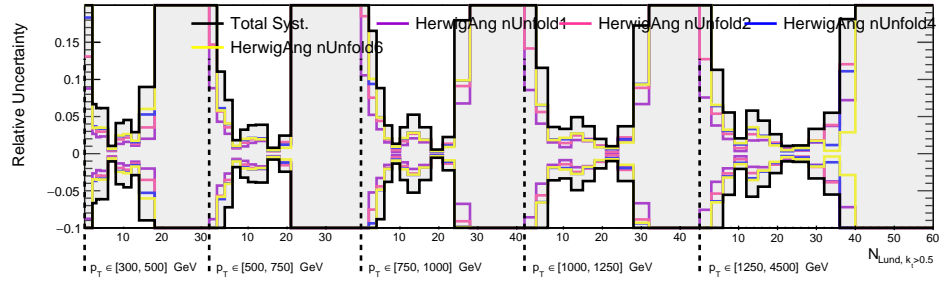


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 50$

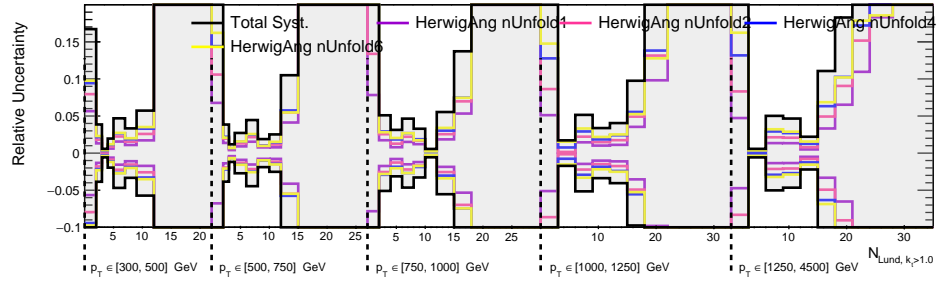


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 100$

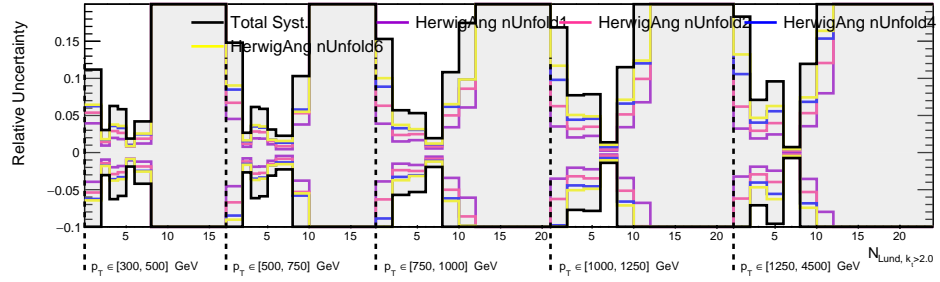
Figure C.6: (continued) The statistical uncertainty from MC for primary Lund multiplicity with various k_t cuts.



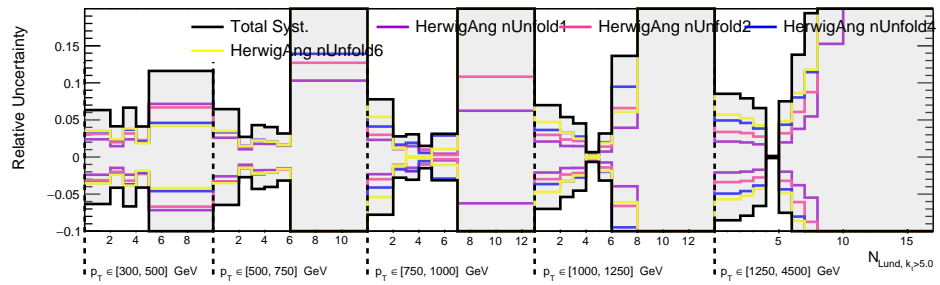
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

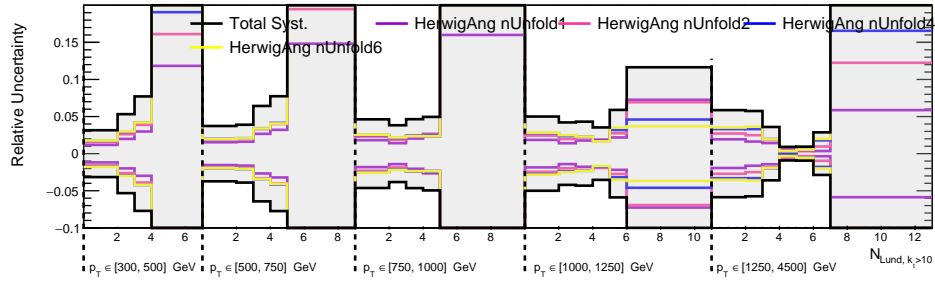


(c) $k_t > 2$ GeV, N_{Lund}

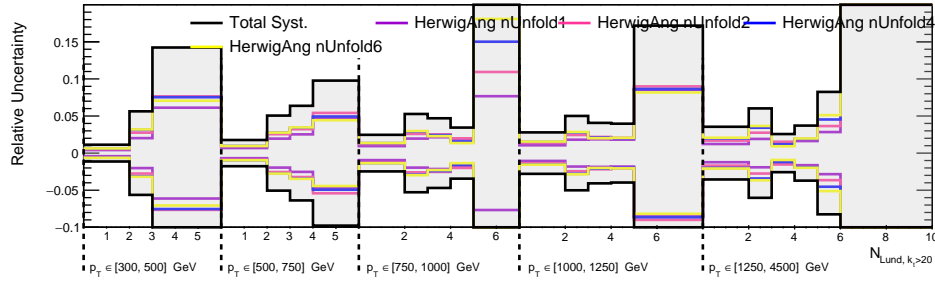


(d) $k_t > 5$ GeV, N_{Lund}

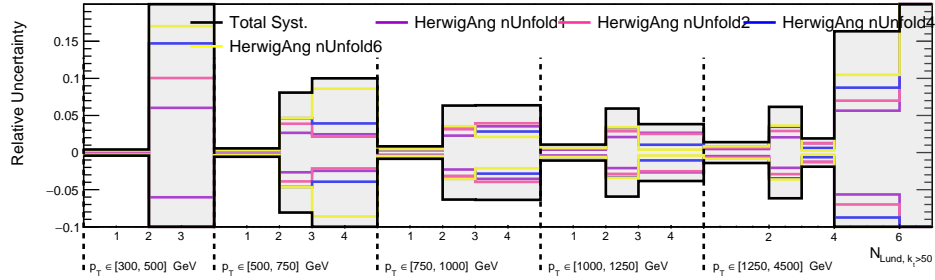
Figure C.7: The Monte Carlo modelling uncertainty for Lund multiplicities with various k_t cuts. Figure continued on next page.



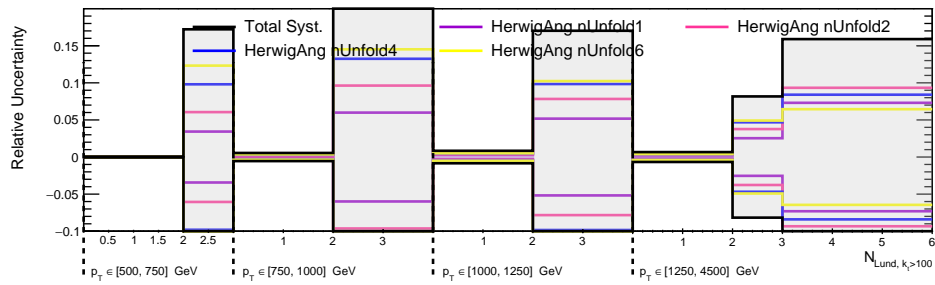
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

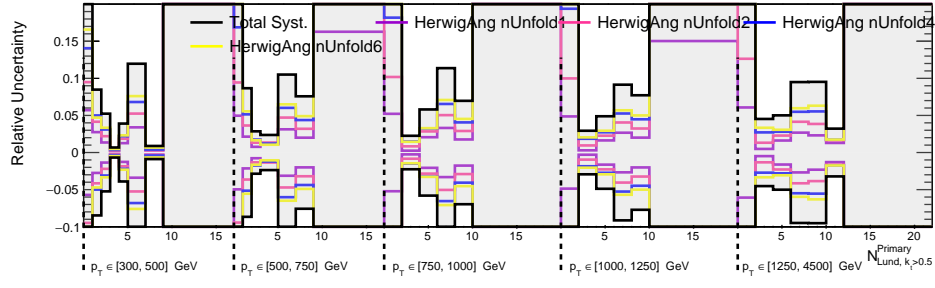


(g) $k_t > 50$ GeV, N_{Lund}

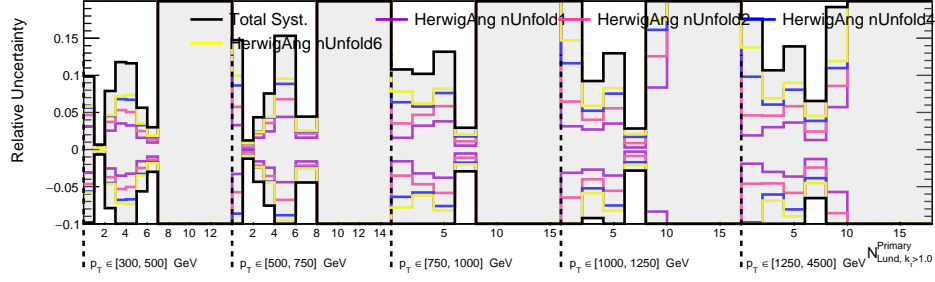


(h) $k_t > 100$ GeV, N_{Lund}

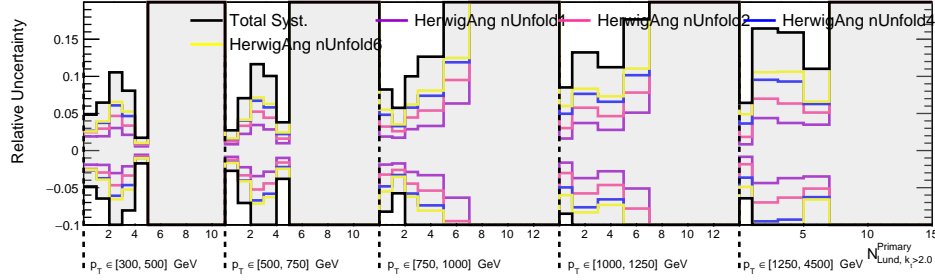
Figure C.7: (continued) The Monte Carlo modelling uncertainty for Lund multiplicities with various k_t cuts.



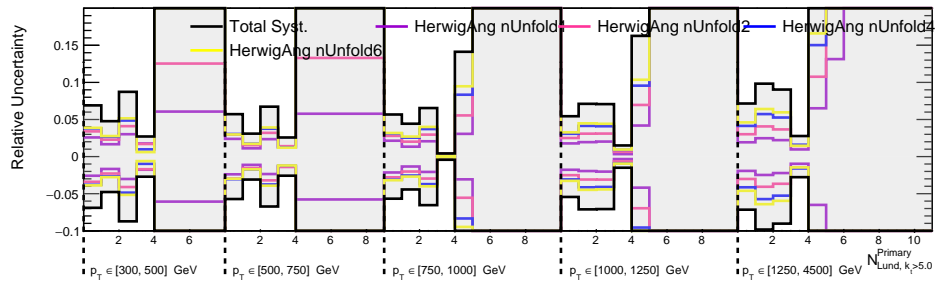
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

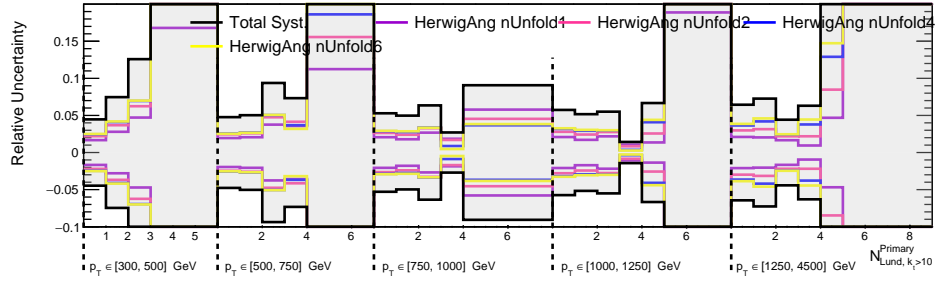


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

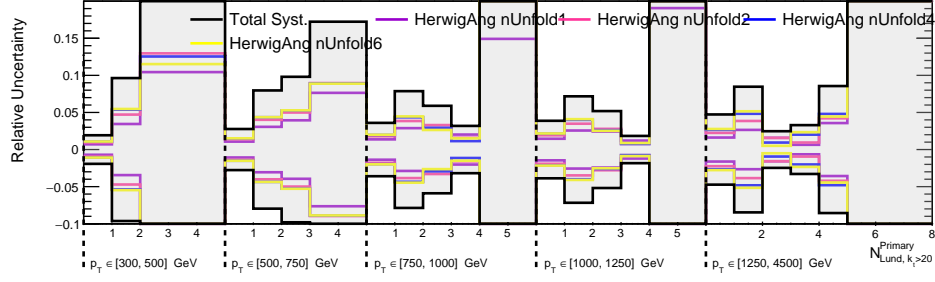


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

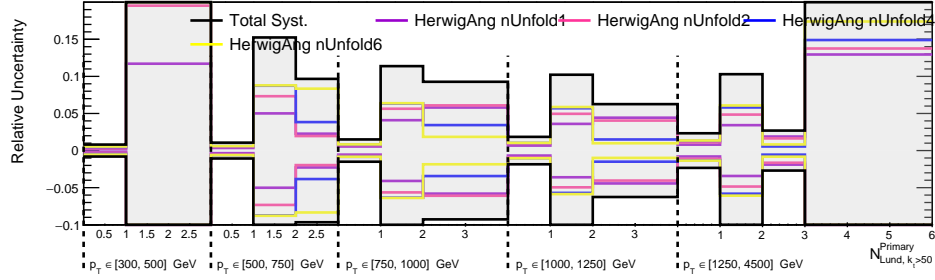
Figure C.8: The Monte Carlo modelling uncertainty for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



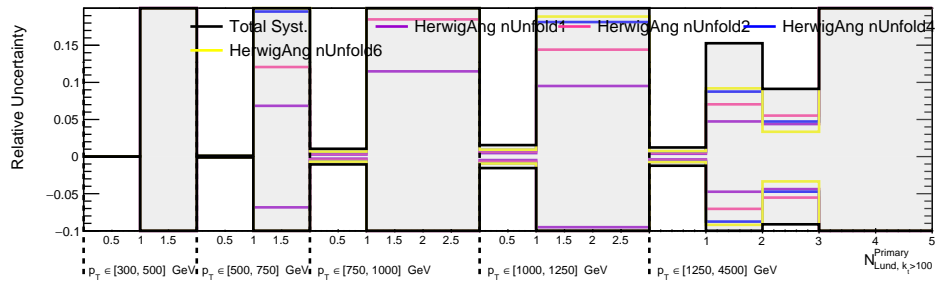
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

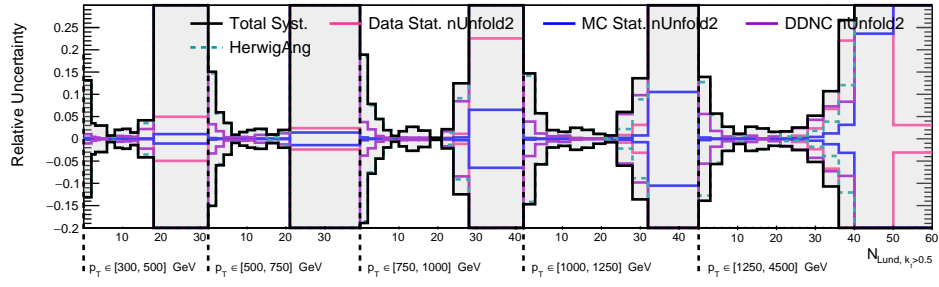


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

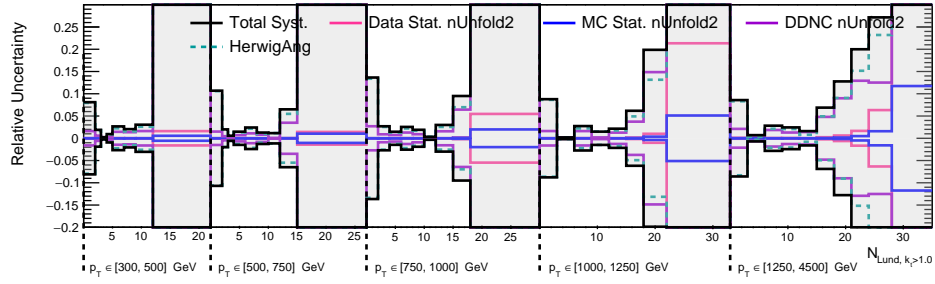


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

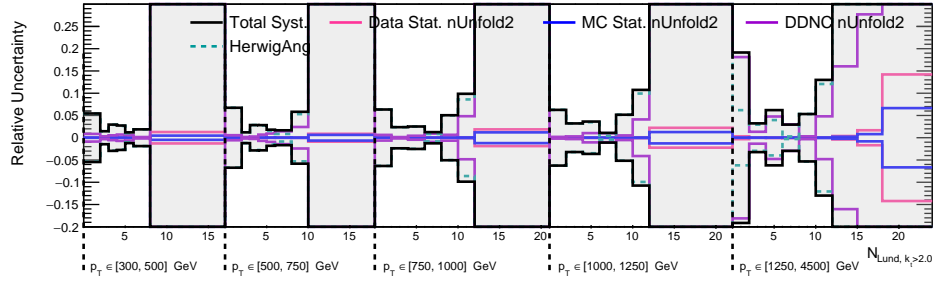
Figure C.8: (continued) The Monte Carlo modelling uncertainty for primary Lund multiplicity with various k_t cuts.



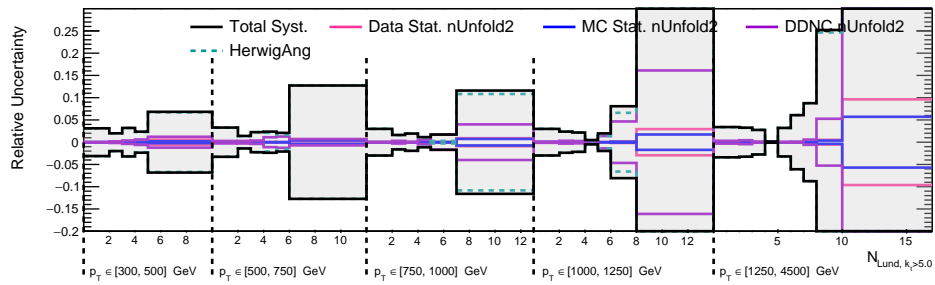
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

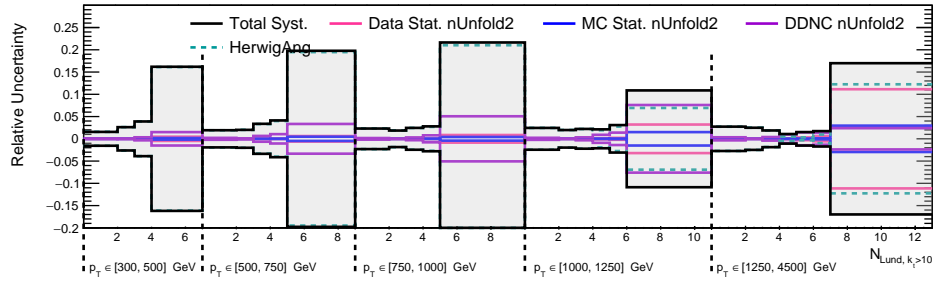


(c) $k_t > 2$ GeV, N_{Lund}

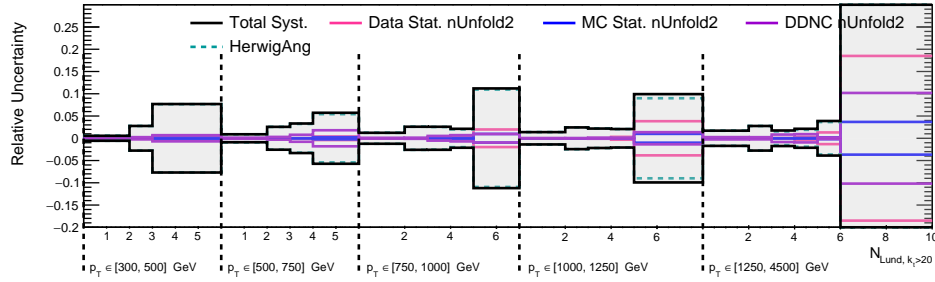


(d) $k_t > 5$ GeV, N_{Lund}

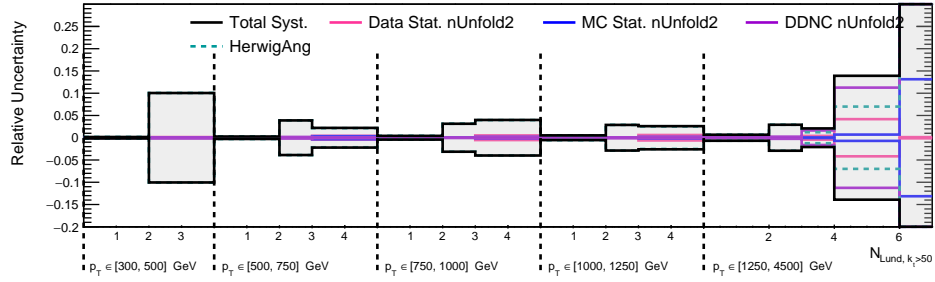
Figure C.9: The distributions with 2 unfolding iterations for Lund multiplicity with various k_t cuts. Figure continued on next page.



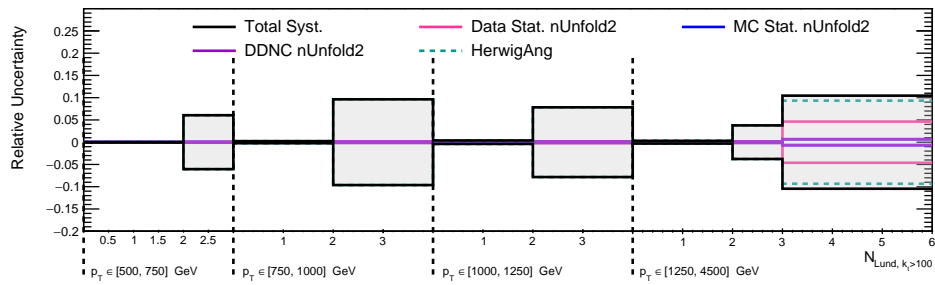
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

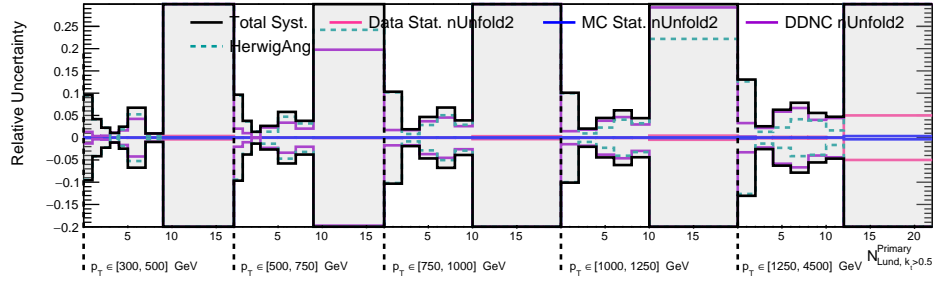


(g) $k_t > 50$ GeV, N_{Lund}

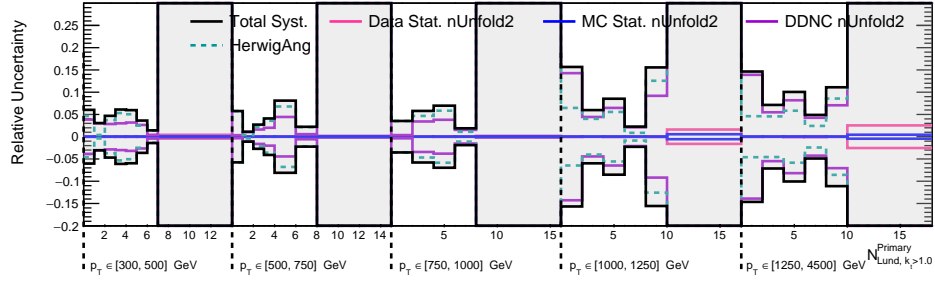


(h) $k_t > 100$ GeV, N_{Lund}

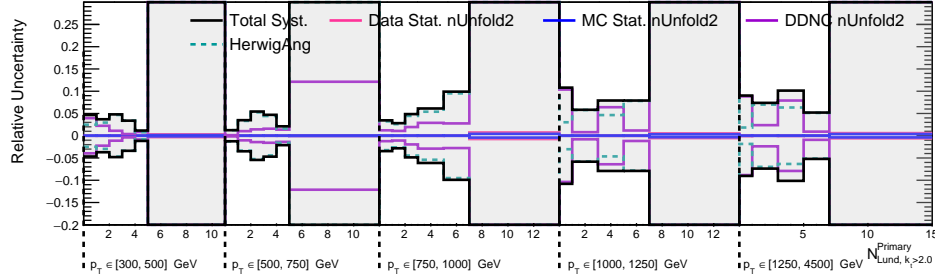
Figure C.9: (continued) The distributions with 2 unfolding iterations for Lund multiplicity with various k_t cuts.



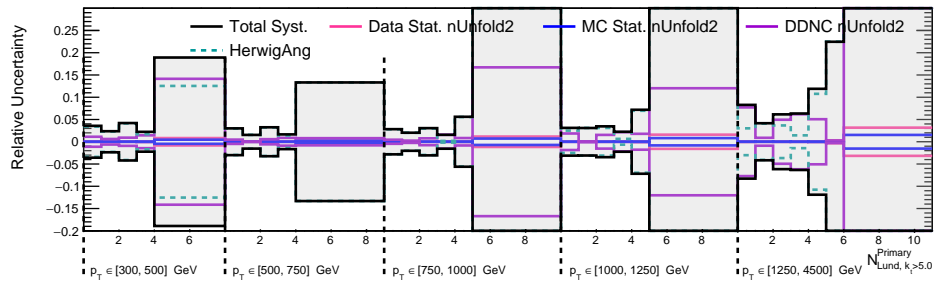
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 0.5$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 1.0$

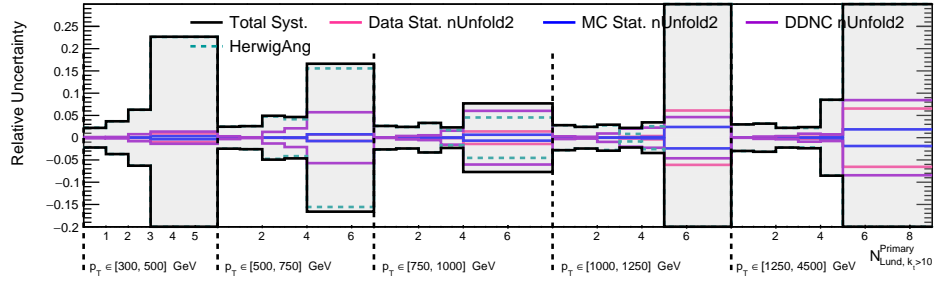


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 2.0$

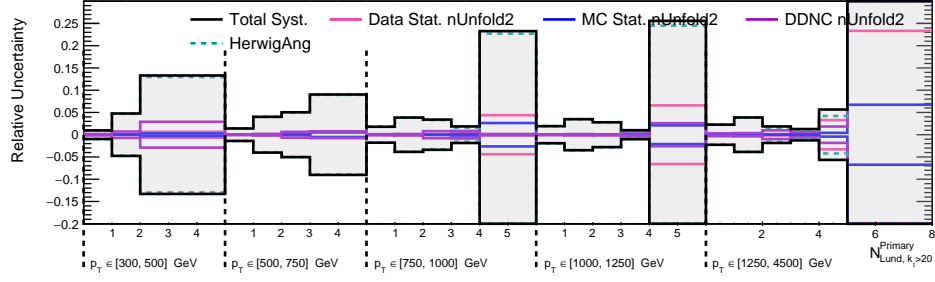


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 5.0$

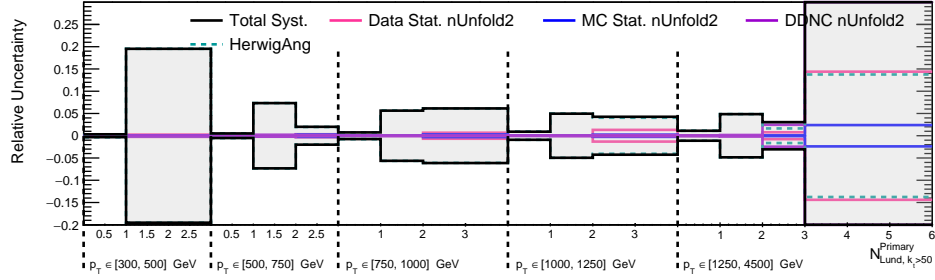
Figure C.10: The distributions with 2 unfolding iterations for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



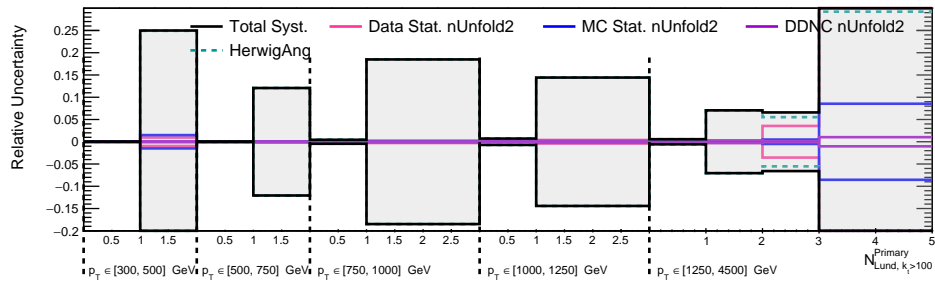
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 10$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 20$

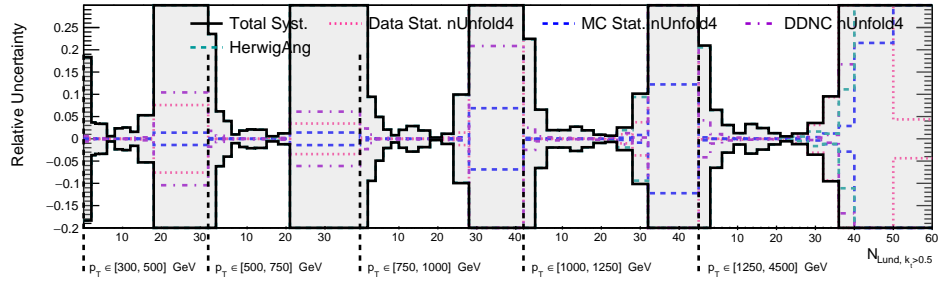


(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 50$

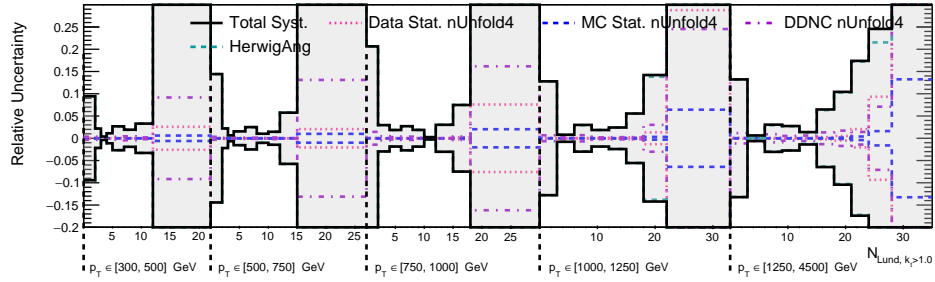


(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 100$

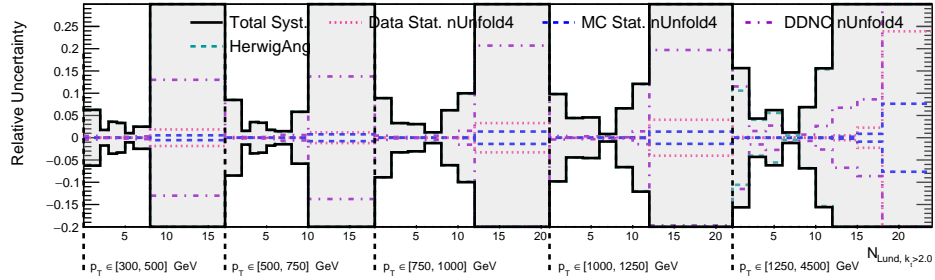
Figure C.10: (continued) The distributions with 2 unfolding iterations for primary Lund multiplicity with various k_t cuts.



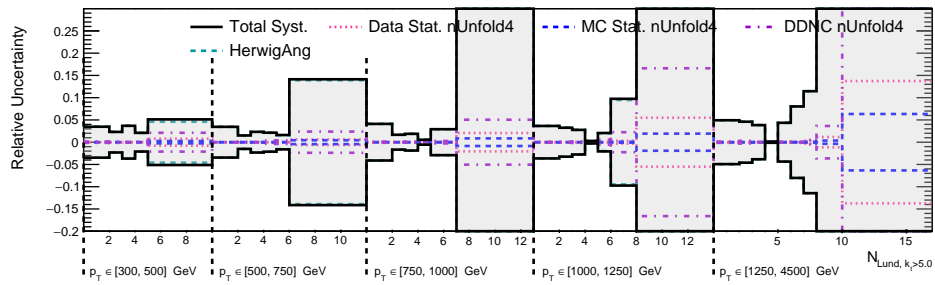
(a) $k_t > 0.5$ GeV, N_{Lund}



(b) $k_t > 1$ GeV, N_{Lund}

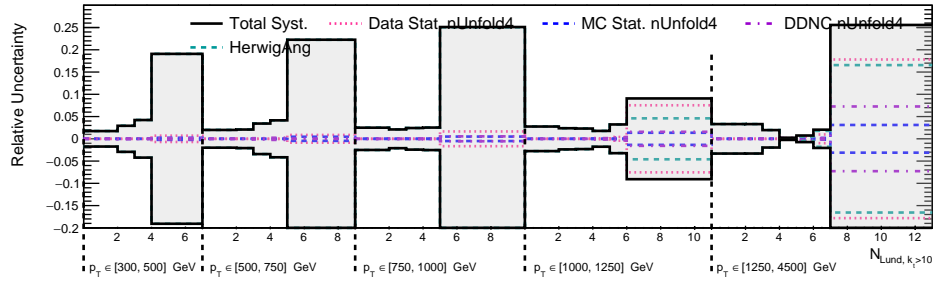


(c) $k_t > 2$ GeV, N_{Lund}

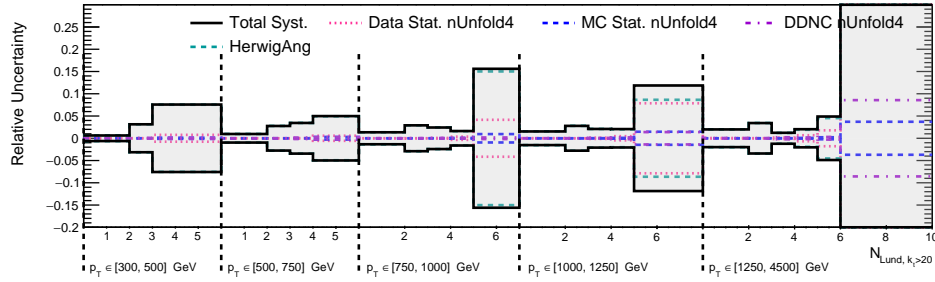


(d) $k_t > 5$ GeV, N_{Lund}

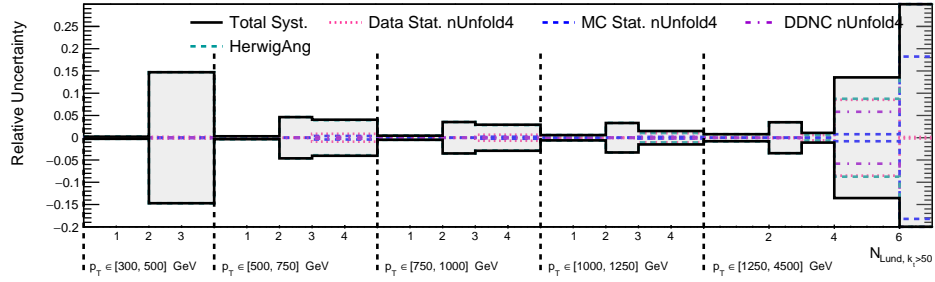
Figure C.11: The distributions with 4 unfolding iterations for Lund multiplicity with various k_t cuts. Figure continued on next page.



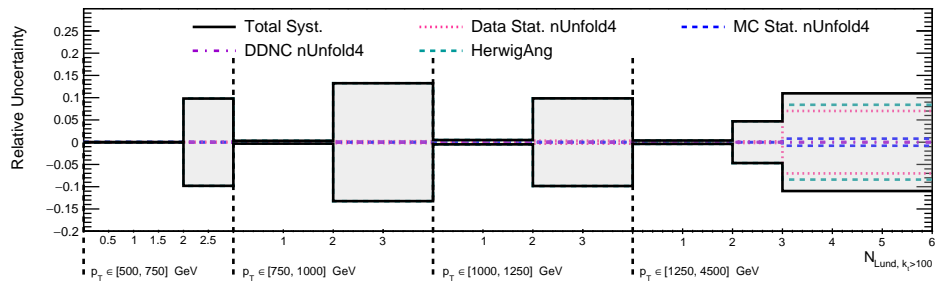
(e) $k_t > 10$ GeV, N_{Lund}



(f) $k_t > 20$ GeV, N_{Lund}

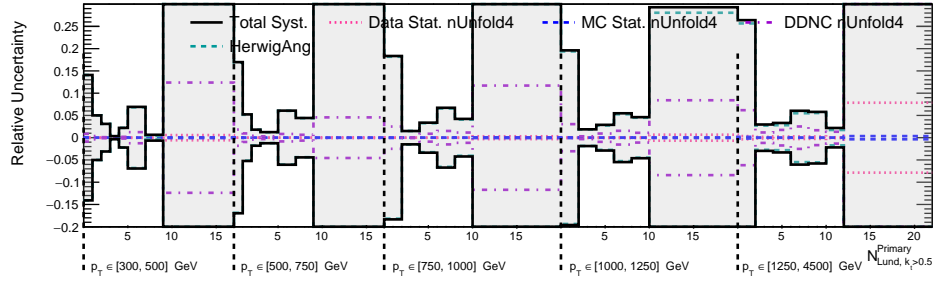


(g) $k_t > 50$ GeV, N_{Lund}

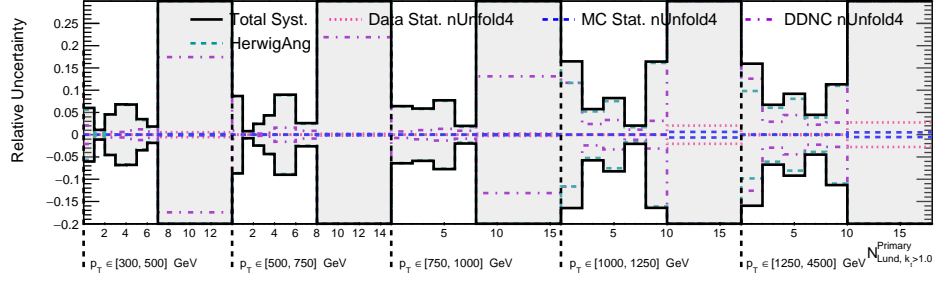


(h) $k_t > 100$ GeV, N_{Lund}

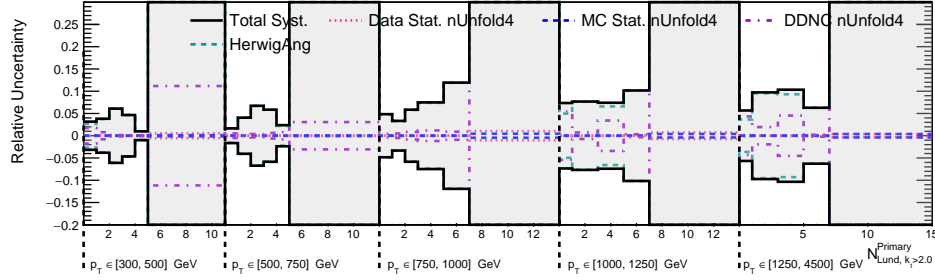
Figure C.11: (continued) The distributions with 4 unfolding iterations for Lund multiplicity with various k_t cuts.



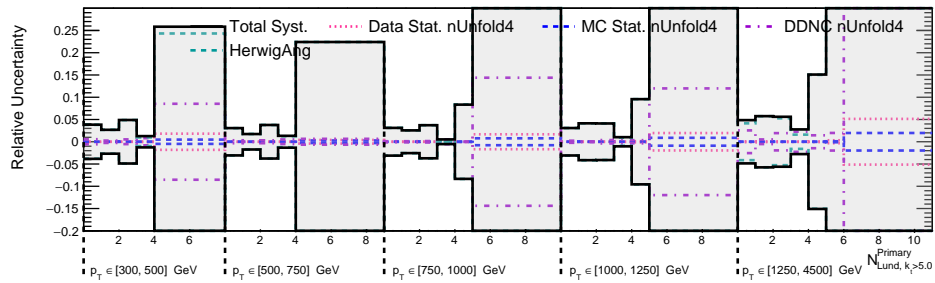
(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 0.5$



(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 1.0$

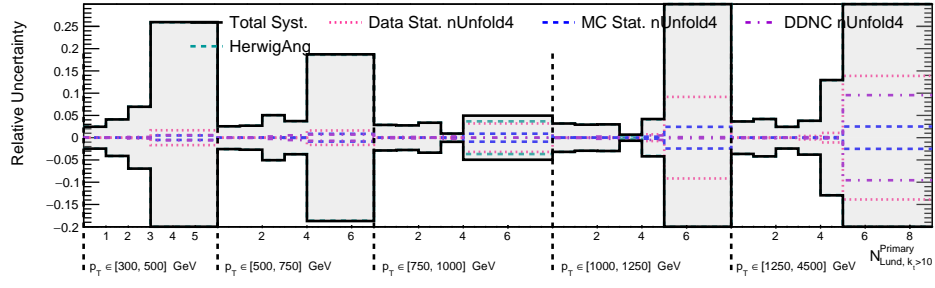


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 2.0$

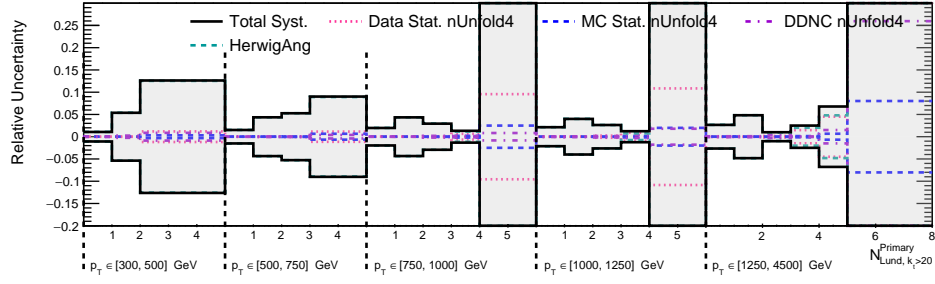


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}} > 5.0$

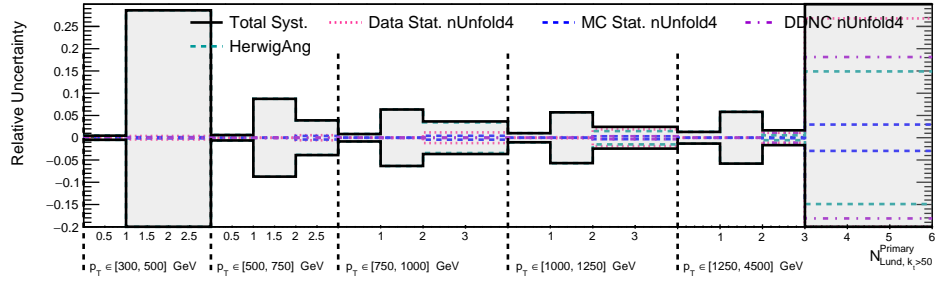
Figure C.12: The distributions with 4 unfolding iterations for primary Lund multiplicity with various k_t cuts. Figure continued on next page.



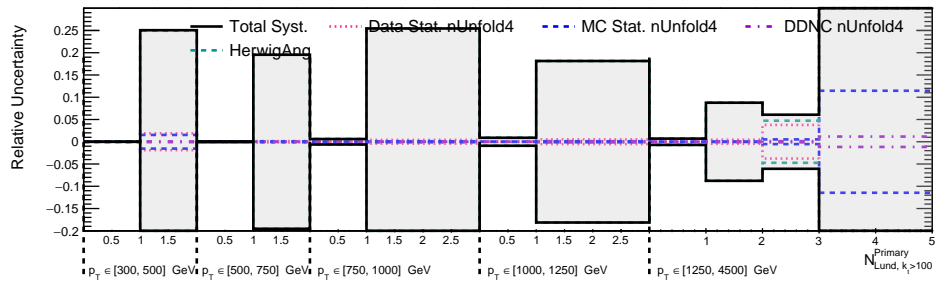
(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$

Figure C.12: (continued) The distributions with 4 unfolding iterations for primary Lund multiplicity with various k_t cuts.

APPENDIX D

MULTIPLICITIES REBINNING STUDY

This appendix details the study done to verify that the rebinning procedure used in the analysis agrees with rebinning procedure normally used for these types of measurements. The observables in this analysis are well suited to a fine-binned unfolding procedure, as multiplicities and counting-variables naturally lend themselves to the use of unit-width binning. This measurement uses bins of width 1 to perform the unfolding, and subsequently rebins post-unfolding to ensure good reco-to-truth-level correspondence and small fake and efficiency corrections in the bins used to report the fiducial cross-section measurement. This study compares the results of the unfolding procedure when performed with fine bins and then rebinned, versus the results of the unfolding when done originally with the coarse bins used in the rebinning scheme.

Using this configuration all of the observables except for the Lund multiplicity with the lowest k_t cut of 0.5 GeV show good agreement between the two types of rebinning procedures. Typically this means perfect correspondence, and in some cases this means that the agreement occurs within the relevant uncertainties (Data-Driven Non-Closure, Data and MC statistical uncertainties, and the uncertainty related to the choice of MC model). A selection of different p_T and η bins for these observables are shown in Figure D.1.

A more thorough selection of bins from the Lund multiplicity with $k_t > 0.5$ GeV is shown in Figure D.2. For only this observable, there are measured bins where the difference between the data rebinned pre-unfolding versus rebinned post-unfolding are not covered by the uncertainties mentioned above. A comparison of the uncertainties in the two cases can be seen in Figure D.3 and Figure D.4, though these plots are not quite graphically finalized.

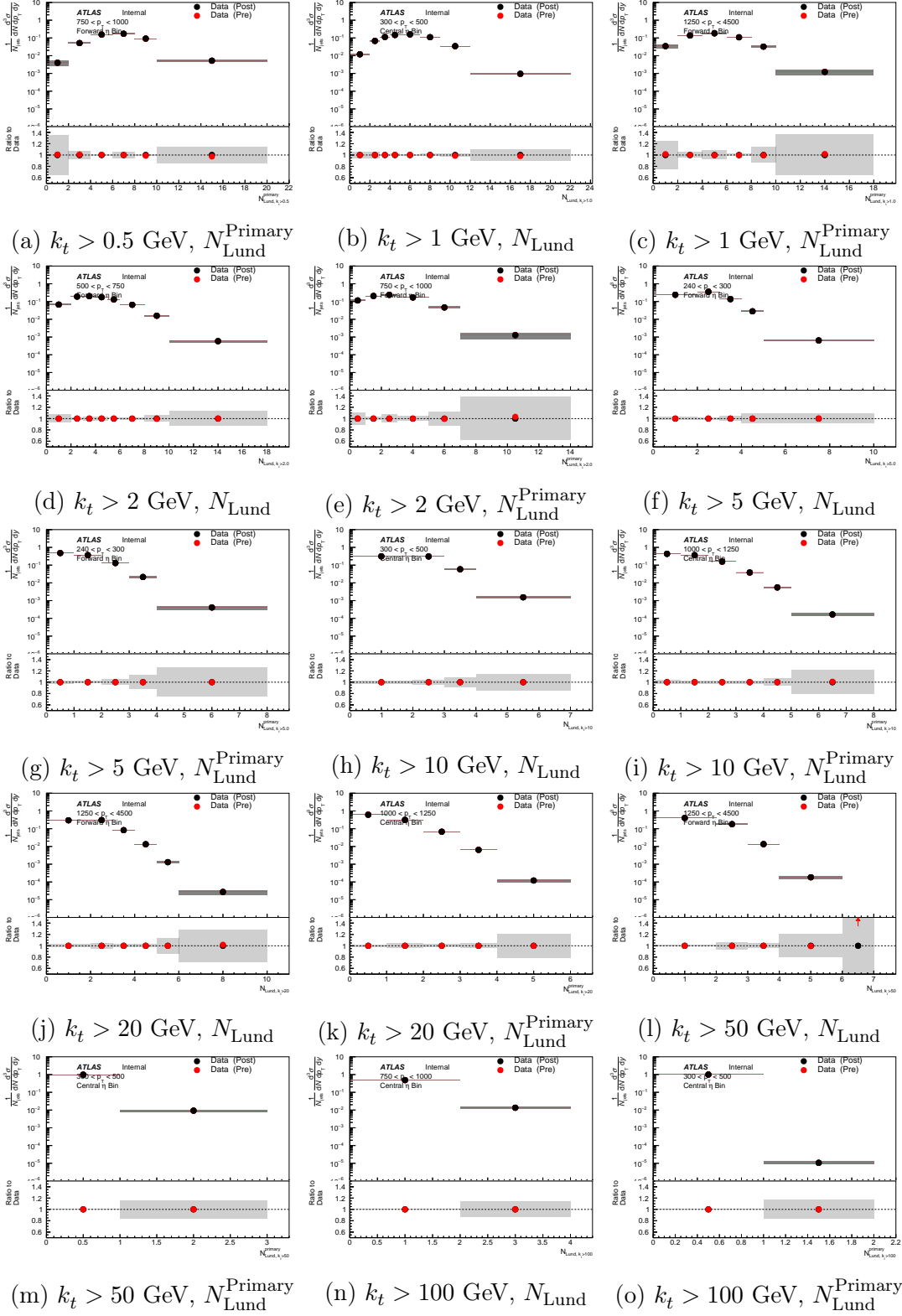
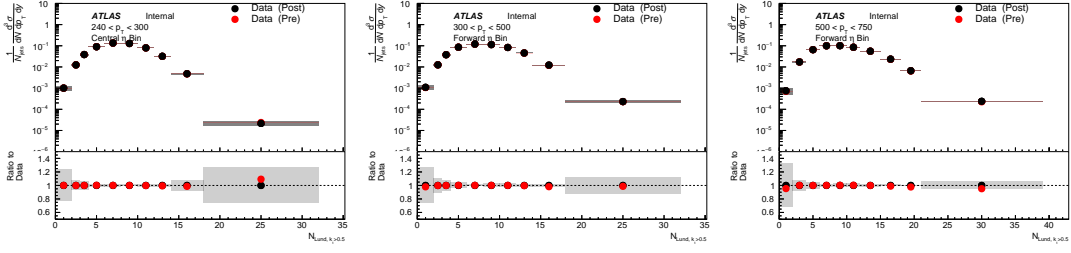
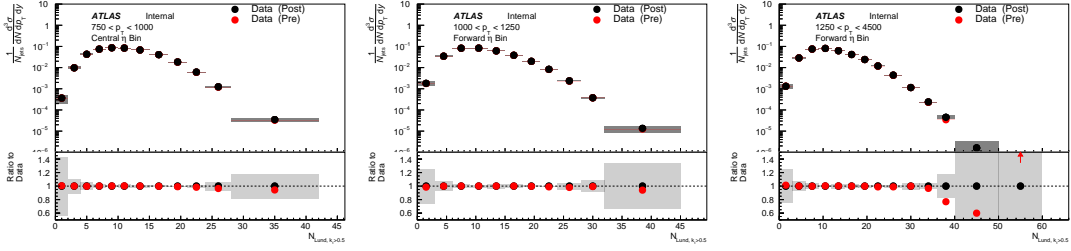


Figure D.1: A selection of multiplicity observables in different p_T and η bins, showing a comparison of data rebinned both before (pre) and after (post) the unfolding procedure.

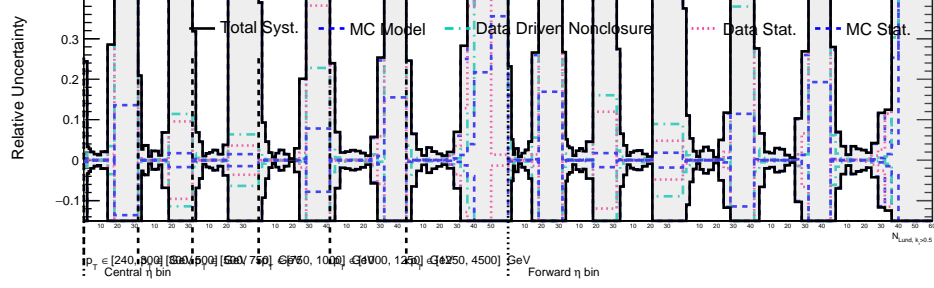
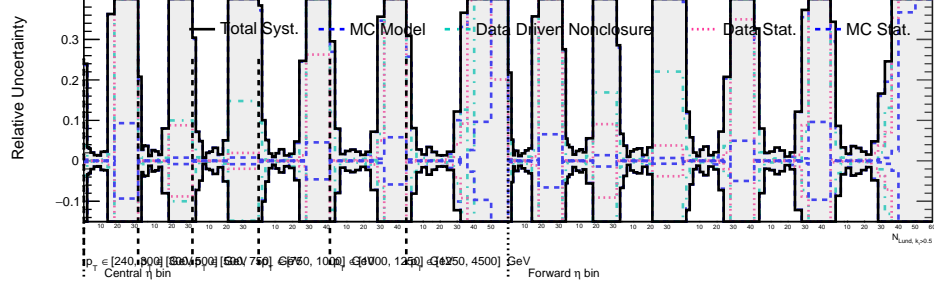


(a) Central η bin, 240 GeV $< p_T < 300$ GeV (b) Forward η bin, 300 GeV $< p_T < 500$ GeV (c) Forward η bin, 500 GeV $< p_T < 750$ GeV

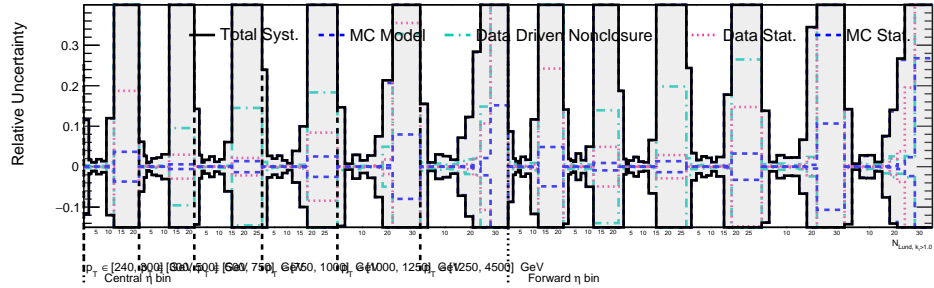
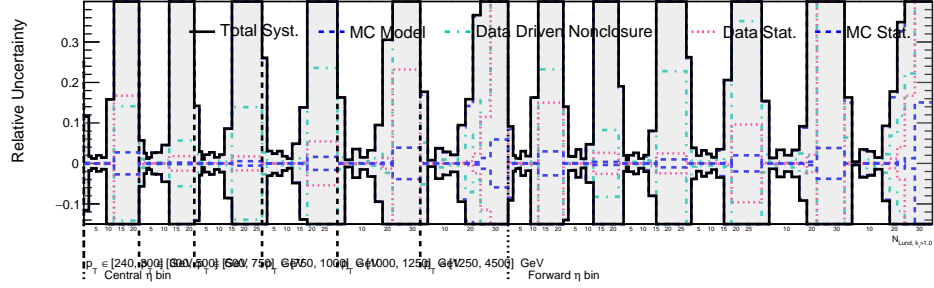


(d) Central η bin, 750 GeV $< p_T < 1000$ GeV (e) Forward η bin, 1000 GeV $< p_T < 1250$ GeV (f) Forward η bin, 1250 GeV $< p_T < 4500$ GeV

Figure D.2: The Lund multiplicity $k_t > 0.5$ GeV in several p_T and η bins, showing a comparison of data rebinned both before (pre) and after (post) the unfolding procedure.

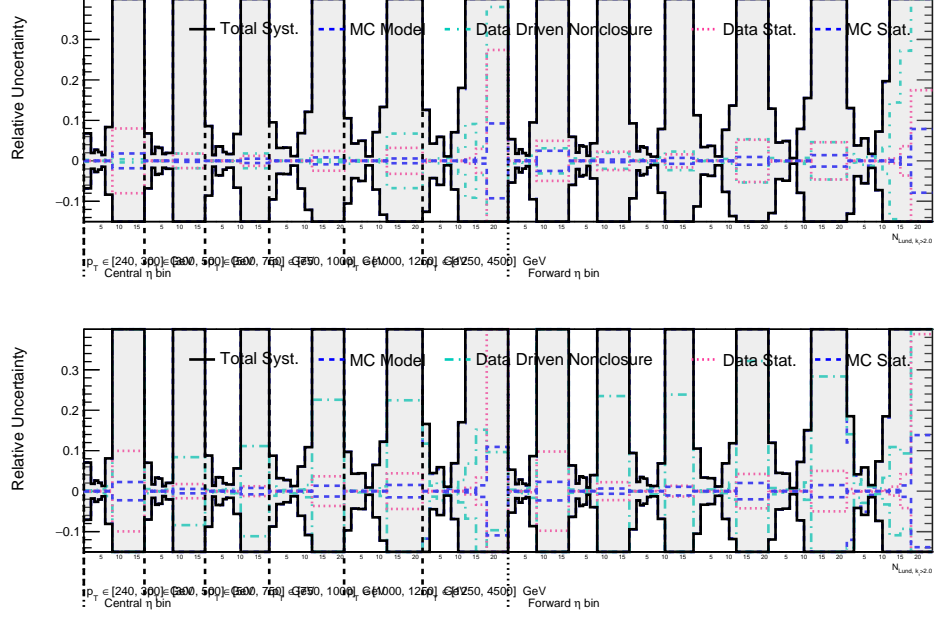


(a) $k_t > 0.5$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

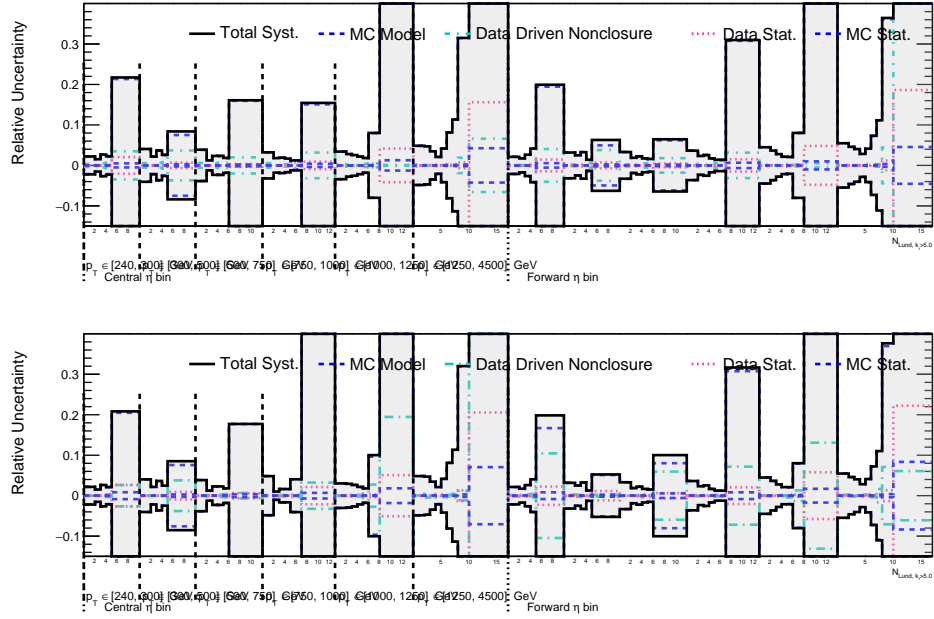


(b) $k_t > 1$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

Figure D.3: The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and bottom figure is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page.

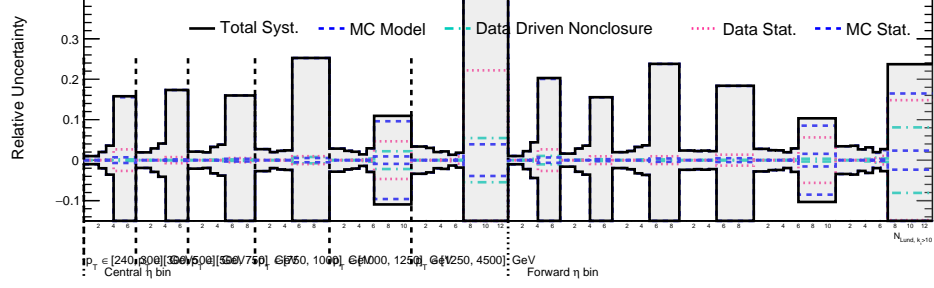
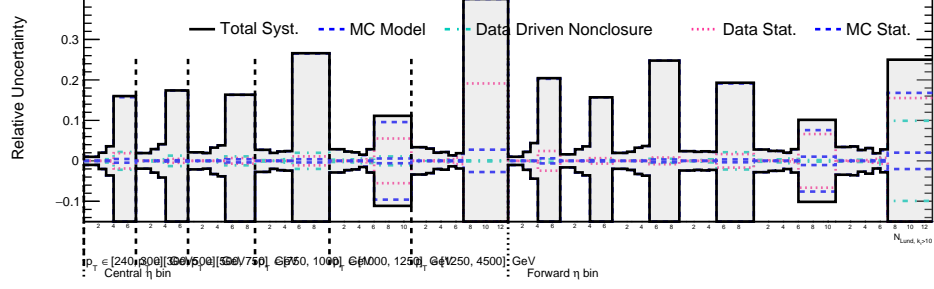


(c) $k_t > 2$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

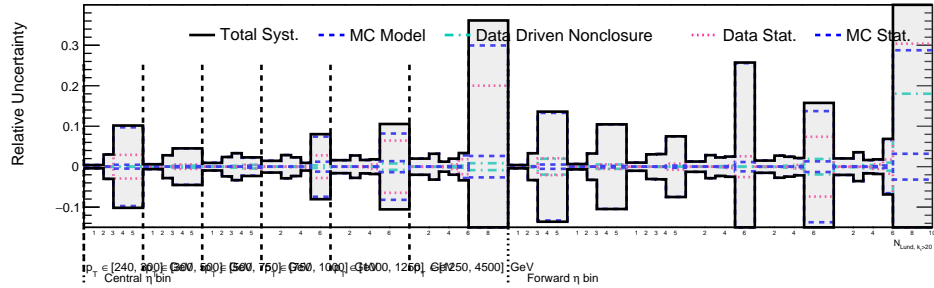
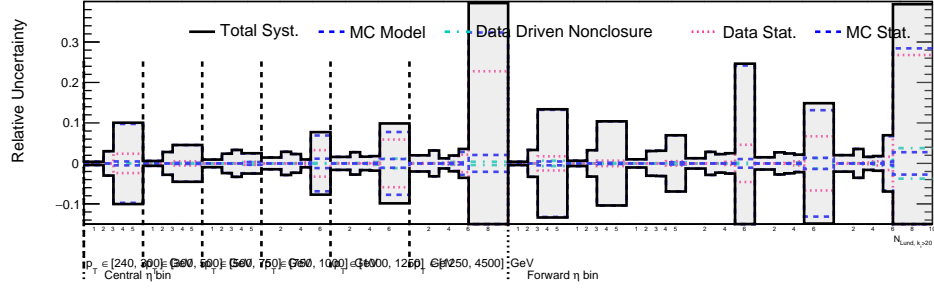


(d) $k_t > 5$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

Figure D.3: (continued) The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and bottom figure is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page.

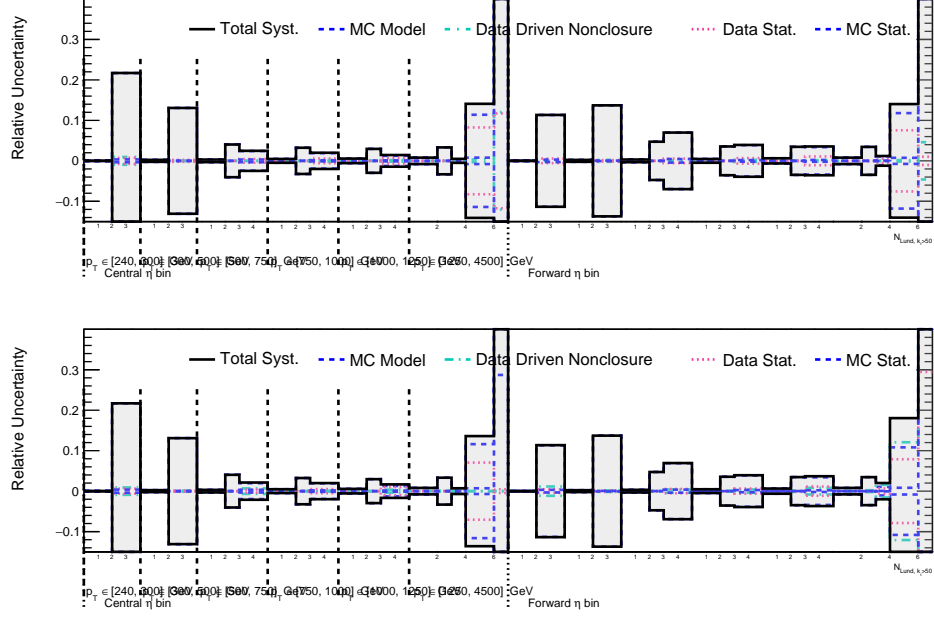


(e) $k_t > 10$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

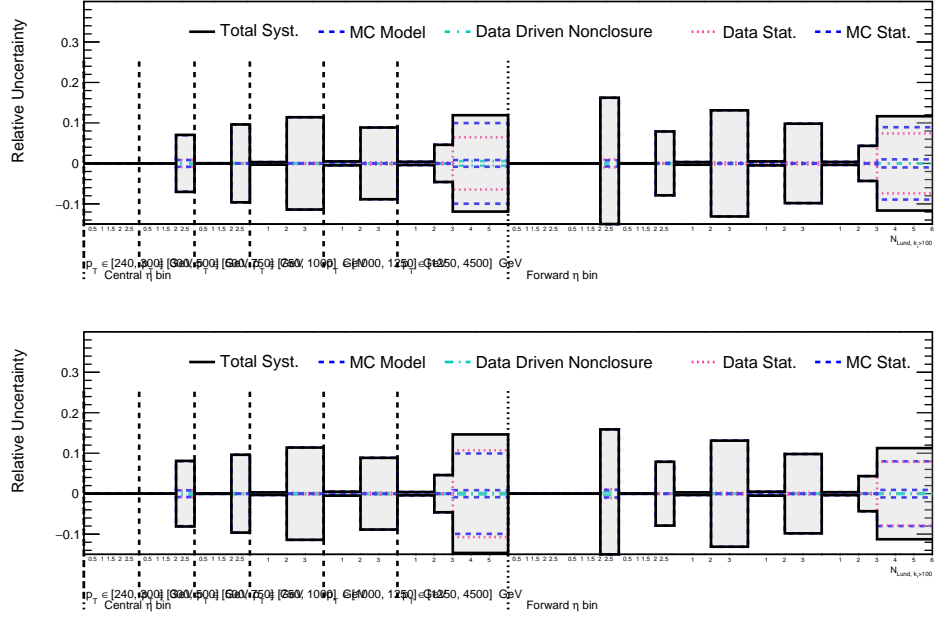


(f) $k_t > 20$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

Figure D.3: (continued) The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and bottom figure is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page.

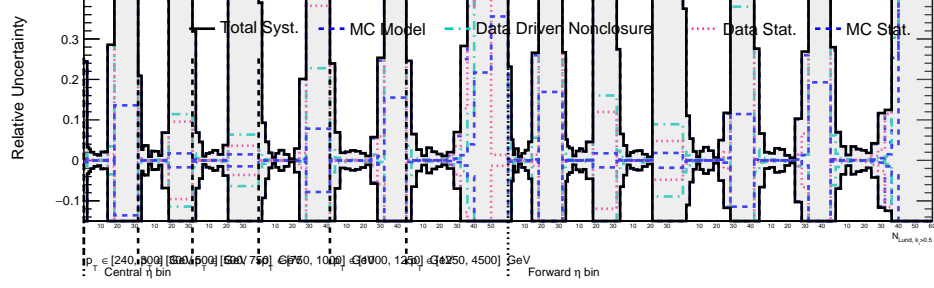
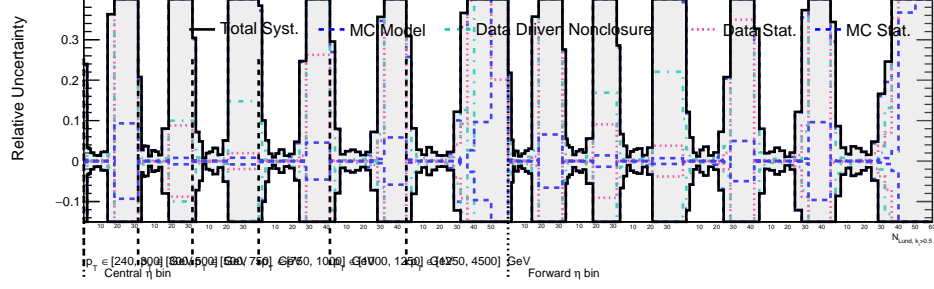


(g) $k_t > 50$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

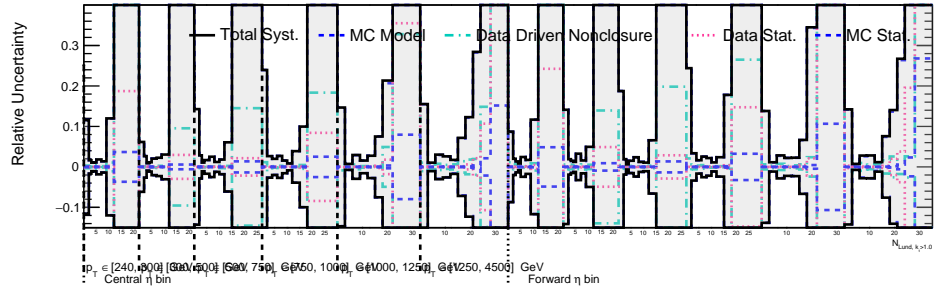
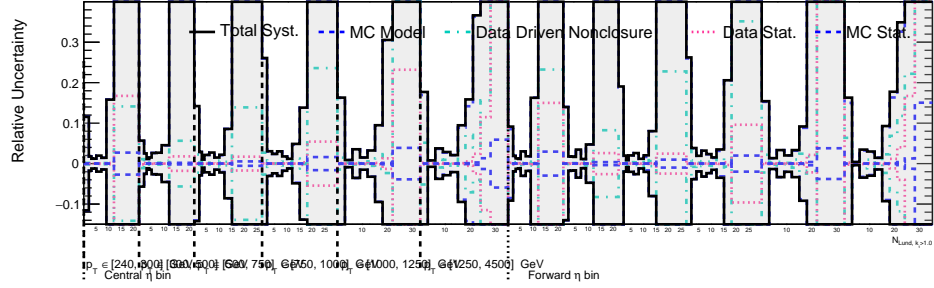


(h) $k_t > 100$ GeV, N_{Lund} , pre-unfolding (top), post-unfolding (bottom)

Figure D.3: (continued) The Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and bottom figure is the uncertainty breakdown when the rebinning procedure is performed post-unfolding.

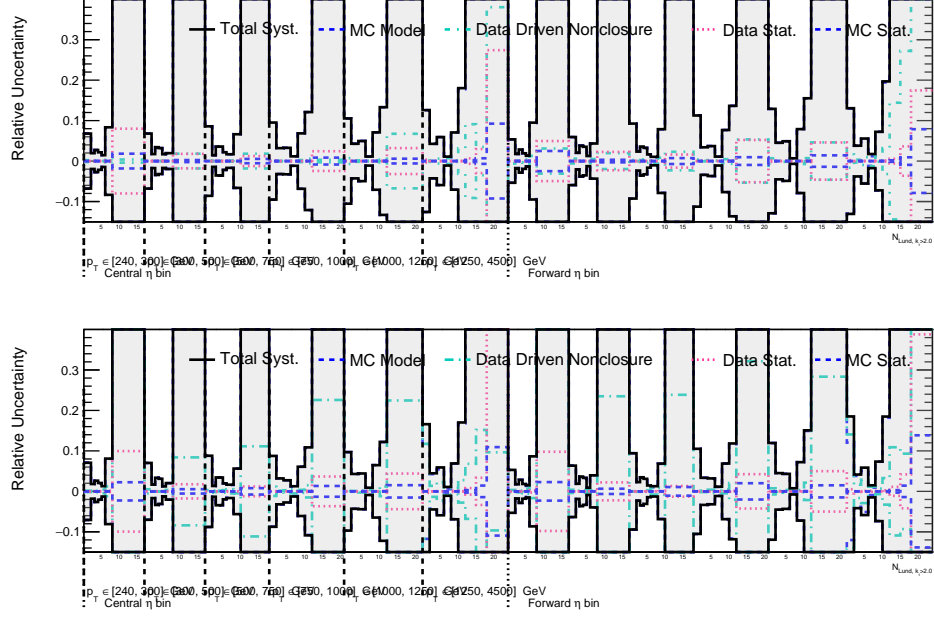


(a) $k_t > 0.5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)

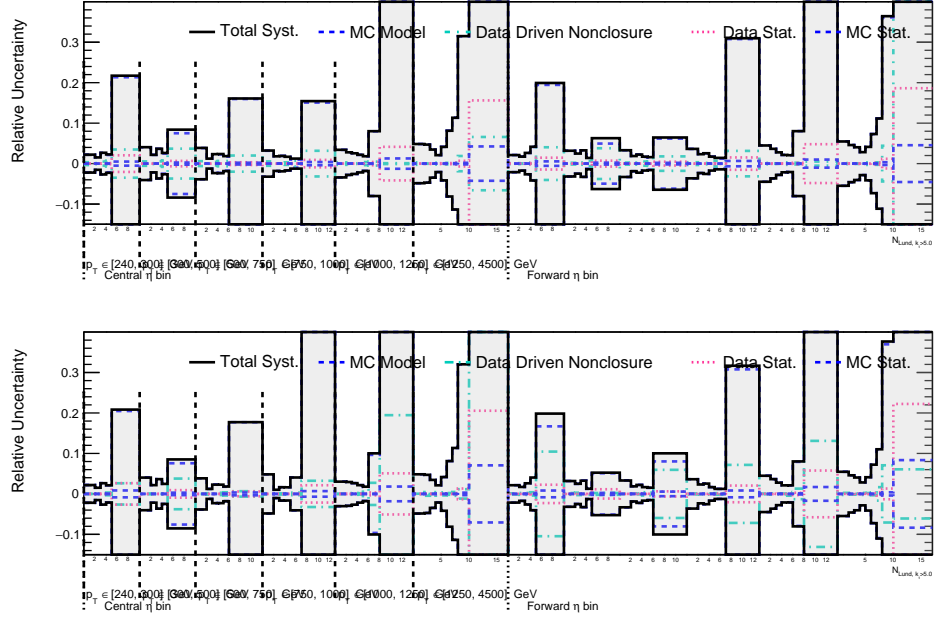


(b) $k_t > 1$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)

Figure D.4: The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and on the bottom is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page.

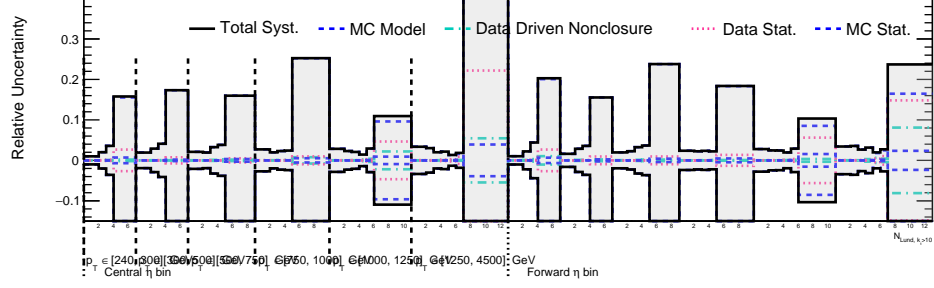
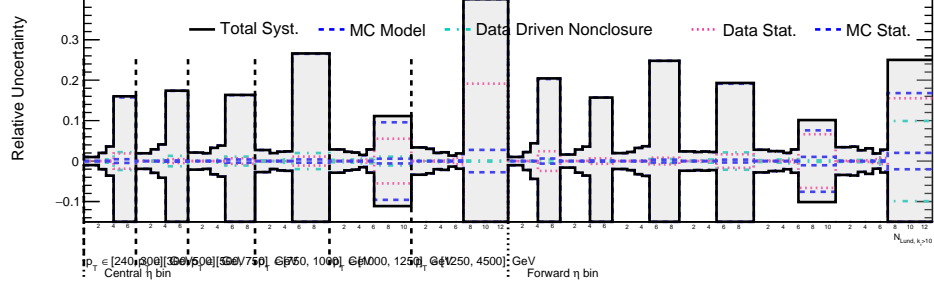


(c) $k_t > 2$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)

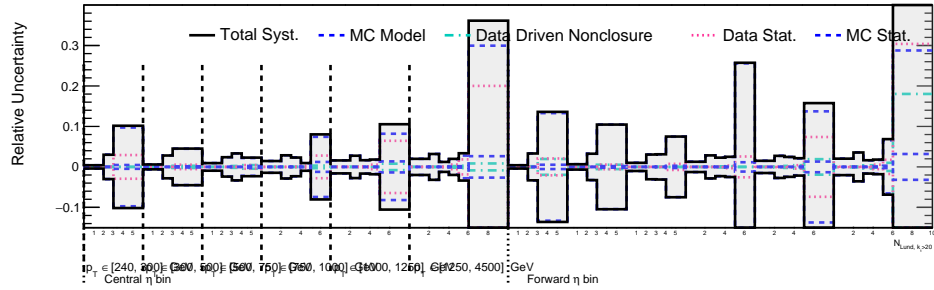
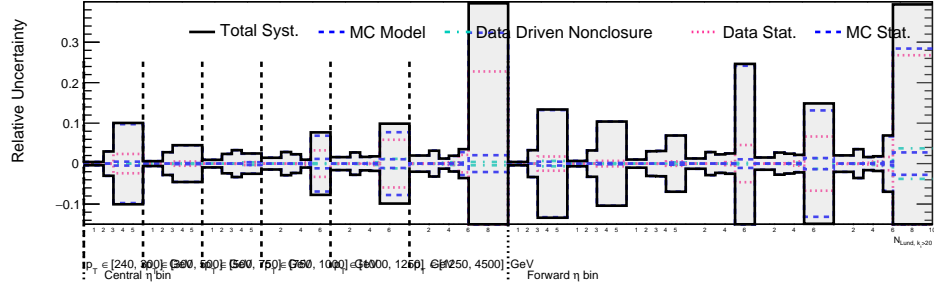


(d) $k_t > 5$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)

Figure D.4: (continued) The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and on the bottom is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page.

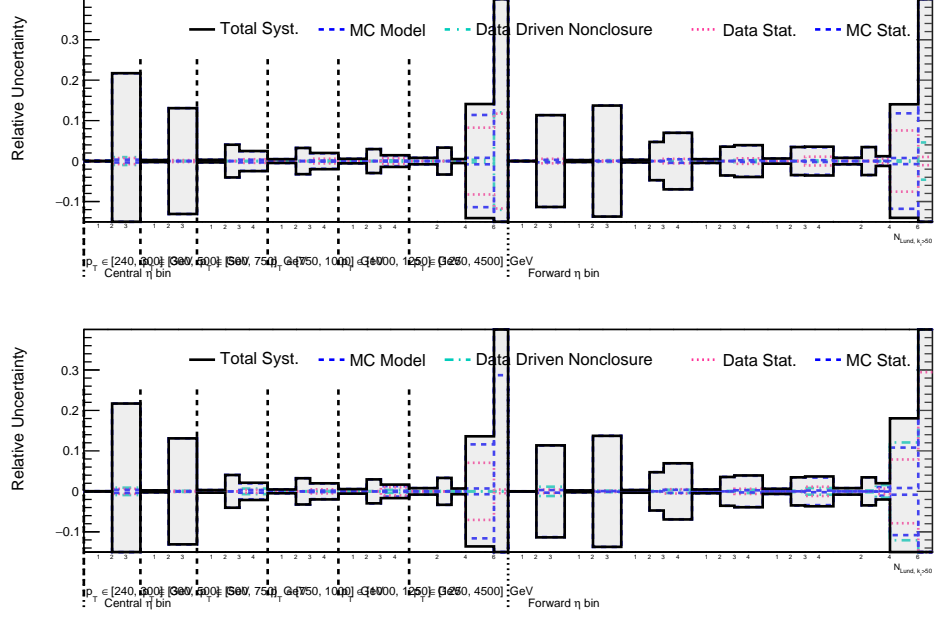


(e) $k_t > 10$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)

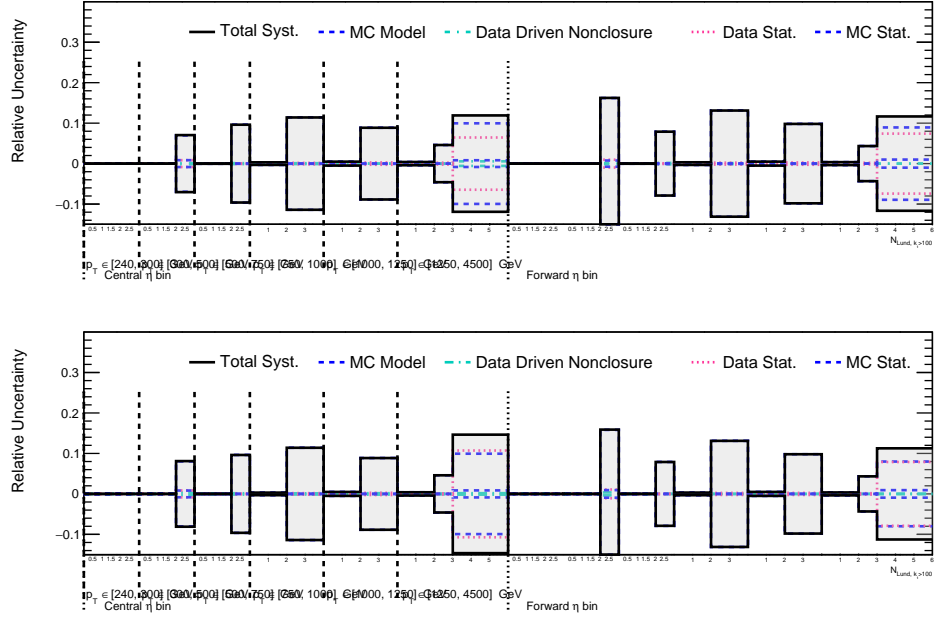


(f) $k_t > 20$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)

Figure D.4: (continued) The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and on the bottom is the uncertainty breakdown when the rebinning procedure is performed post-unfolding. Figure continued on next page.



(g) $k_t > 50$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)



(h) $k_t > 100$ GeV, $N_{\text{Lund}}^{\text{Primary}}$, pre-unfolding (top), post-unfolding (bottom)

Figure D.4: (continued) The primary Lund multiplicity distributions with various k_t cuts compared to predictions to several Monte Carlo generators. In each set of two figures the top figure is the uncertainty breakdown when the rebinning procedure is performed pre-unfolding, and on the bottom is the uncertainty breakdown when the rebinning procedure is performed post-unfolding.

APPENDIX E

MULTIPLICITIES REBINNING AVERAGES STUDY

The averages in this analysis are determined before the rebinning procedure is performed, requiring further studies to determine if this strategy is valid for the determination of the mean. The averaging of a distribution does provides a regularizing effect, which reduces the sensitivity to sub-resolution effects. In addition, the use of unit binning removes the need for a binning correction on the average, based on the inherit bias of bin centers that do not correspond to the average value of entries within that bin.

In order to test the validity of performing the average on the non-rebinned distribution, several tests were performed. These studies aim to test two potential effects: whether the non-rebinned results are sensitive to sub-detector-resolution information (i.e. if we are extracting information that would not be present in the coarsely-binned results), and whether this strategy results in an underestimation of the modeling uncertainties by ignoring any resolution-dependent effects.

To answer these questions, one single observale is used with a more optimized binning for a single k_t cut (1 GeV) for the primary multiplicity. A low value of the k_t cut is used in order to test the impact of the rebinning across a wide range of multiplicities, particularly since the model differences decrease at high k_t values. The binning is optimized such that the purity is between 30-40% for all bins. The one exception to this is that the first bin always ranges from 0-1, since there are very few entries in this bin, and so merging it with the 1-2 bin would create an inherit bias, since the bin center would be at 1, instead of slightly below 2.¹ The response matrices and purities pre- and post-rebinning are shown in Figure E.1, which shows that the purities are significantly higher after the rebinning.

1. If this strategy were used in the analysis, this effect could be treated more correctly, and so the overall conclusions would not be strongly affected.

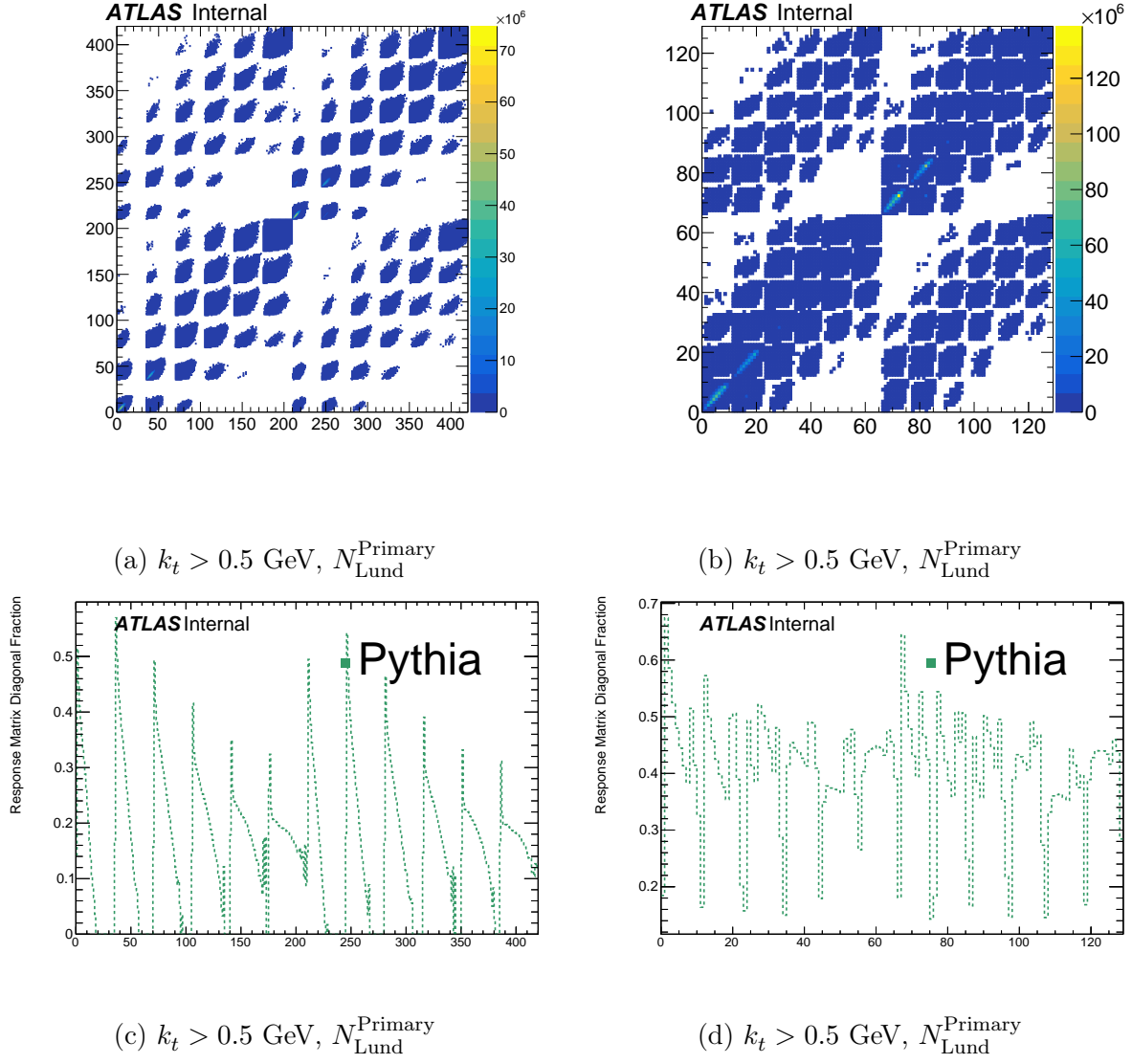


Figure E.1: (Top) The response matrices and (bottom) purities for $k_t > 1$ GeV, (left) before rebinning, and (right) after rebinning.

Figure E.2 shows a few examples of the detector-level distributions for PYTHIA, SHERPA2.2.11, HERWIG with the angular-ordered parton shower, and data, before and after the rebinning. These studies are done at detector-level, in order to include a comparison to data, and the conclusions are not expected to change at truth or unfolded level. The binning correction is defined as the ratio of the fine-binned average to the coarse-binned average, and the results of this are shown in table E.1. In general, the binning correction is between 2-10%, decreasing at higher p_T as the multiplicity distribution falls less steeply at higher values, reducing the impact of the last bin, which typically is very wide.

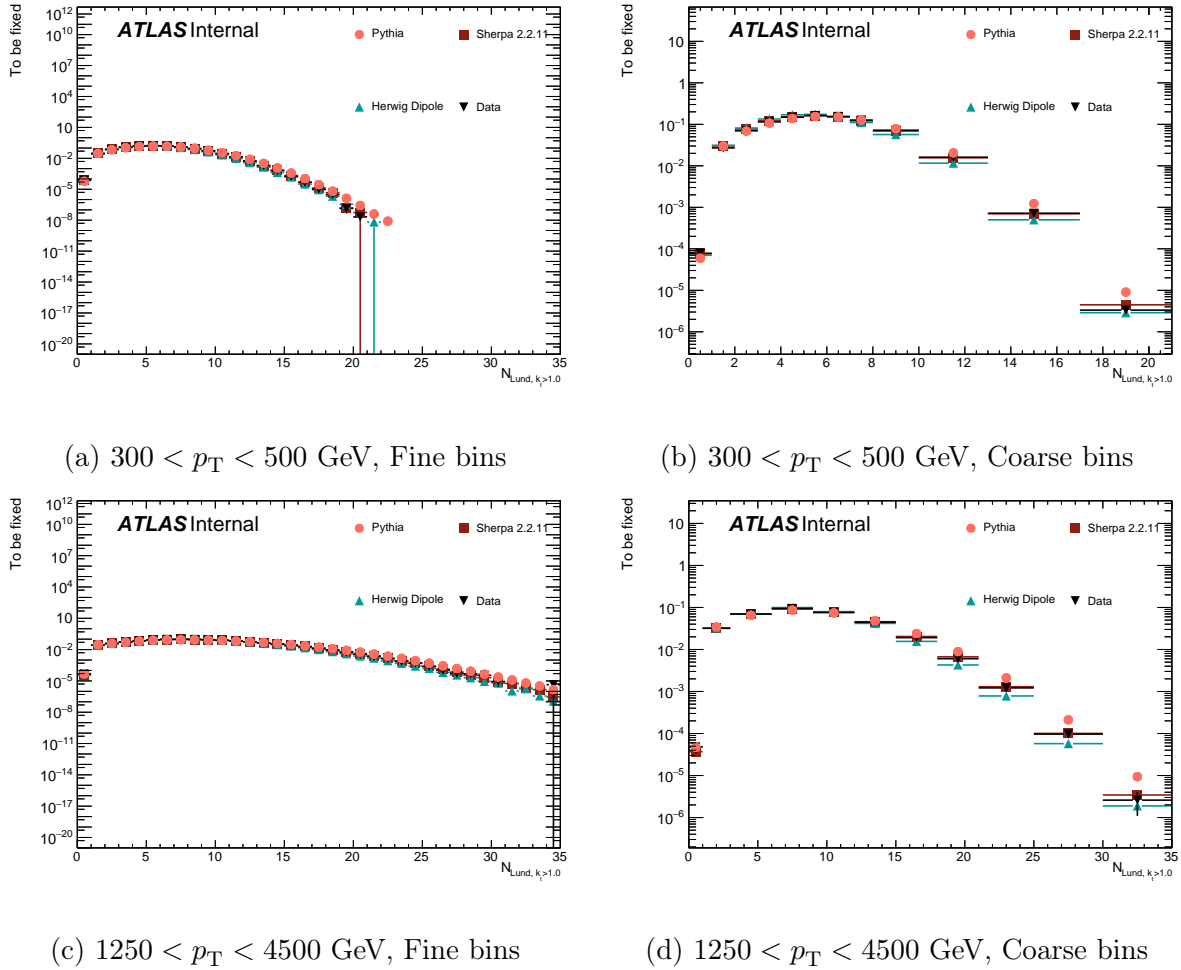


Figure E.2: The (left) fine-binned and (right) coarse-binned multiplicity distributions for (top) $300 < p_T < 500$ GeV, and (bottom) $1250 < p_T < 4500$ GeV

p_T bin	MC	Fine-bin average	Coarse-bin average	Binning correction
300 – 500 GeV	Pythia	6.1435	5.6555	1.08629
	Sherpa	5.89467	5.48769	1.07416
	Herwig	5.652	5.31953	1.0625
	Data	5.99212	5.57775	1.07429
500 – 750 GeV	Pythia	7.11274	6.47571	1.09837
	Sherpa	6.8601	6.26349	1.09525
	Herwig	6.54029	5.98384	1.09299
	Data	6.93468	6.35501	1.09122
750 – 1000 GeV	Pythia	7.91185	7.42699	1.06528
	Sherpa	7.64496	7.25178	1.05422
	Herwig	7.3081	6.98687	1.04598
	Data	7.68916	7.30002	1.05331
1000 – 1250 GeV	Pythia	8.53571	7.89786	1.08076
	Sherpa	8.27549	7.74492	1.06851
	Herwig	7.92318	7.48946	1.05791
	Data	8.26036	7.74154	1.06702
1250 – 40000 GeV	Pythia	9.16619	8.91656	1.028
	Sherpa	8.87924	8.66645	1.02455
	Herwig	8.58055	8.38711	1.02306
	Data	8.8026	8.59608	1.02402

Table E.1: Comparison of the raw averages between the fine-binned and coarse-binned distributions.

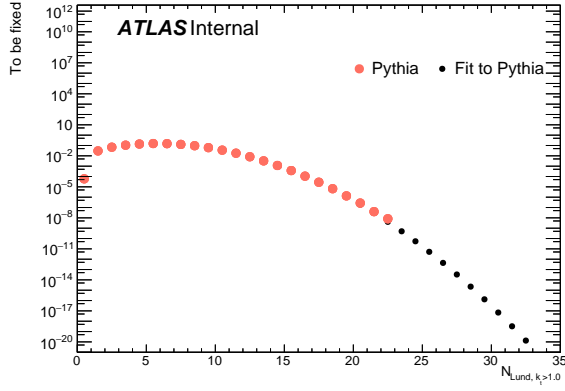
The raw average demonstrates that the modeling dependence is not significantly underestimated by our modeling uncertainties, even though the non-rebinned distributions are used. However, a more realistic analysis would have smaller modeling uncertainties when using a binned distribution.

The multiplicity distributions have additional information that can be used to determine the average more precisely than simply using the bin center. In particular, fitting these distributions can provide a more precise estimate of the average, by providing a more accurate bin center. In an analysis, the fits could be used either directly, as is done here, or to

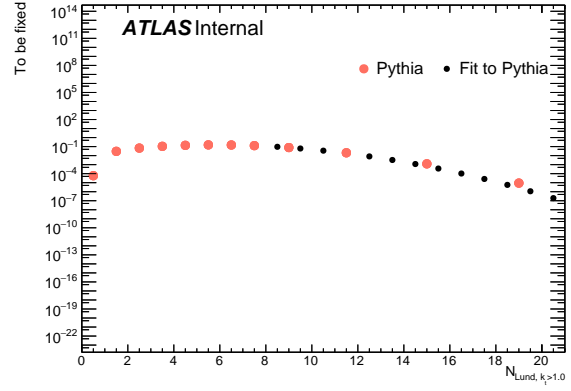
determine the bin center, which is probably more realistic, but the differences between these two approaches is small if the fits describe the distributions.

For this cross-check, the left side of the plot was found to be well-modeled by a Poisson distribution, and the right side by a Gaussian distribution. To illustrate that this models the fine-binned distribution well, the fine-binned distributions are fit, which is shown in Figure E.3. Then, these fits were performed on the coarse-binned distributions. These fits are shown in Figure E.4, where they are compared to the coarse-binned distributions that were fit, as well as the fine-binned distributions. As shown in the figure, both distributions are modeled reasonably well, with some deviations in the tails when compared to the fine-binned distribution.

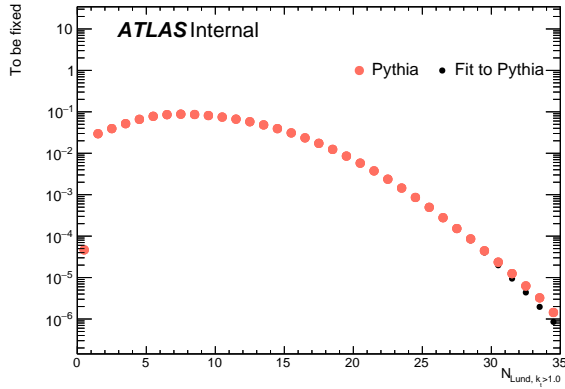
The results of the binning corrections determined from the fits to the coarse-binned distributions are shown in Table E.2. The binning corrections determined using the fits is significantly smaller than the raw average binning correction, and it is typically less than 1%, with less than 2% variation across Monte Carlo predictions. We note that further improvements may be possible with a more detailed study, and so this is an upper bound on the impact of the binning. These studies demonstrate that we are not extracting information beyond what is available by the detector resolution, and that the modeling differences that would have arisen from a coarse binning are significantly smaller than the existing modeling uncertainties.



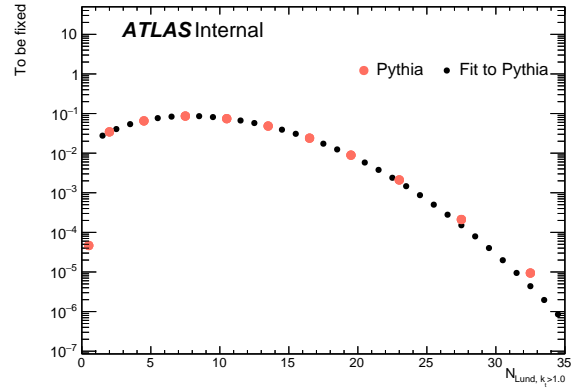
(a) $300 < p_T < 500$ GeV, Fine bins



(b) $300 < p_T < 500$ GeV, Coarse bins

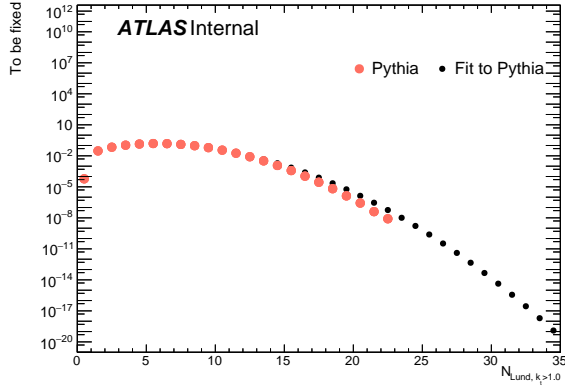


(c) $1250 < p_T < 4500$ GeV, Fine bins

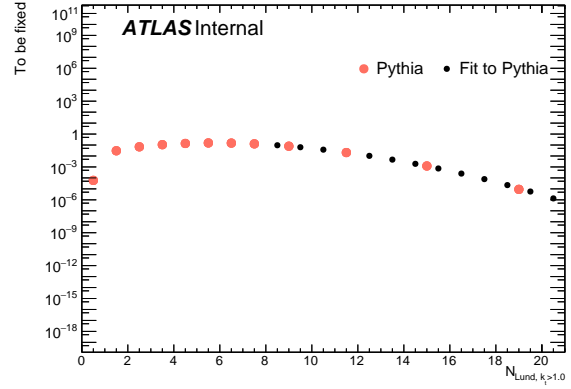


(d) $1250 < p_T < 4500$ GeV, Coarse bins

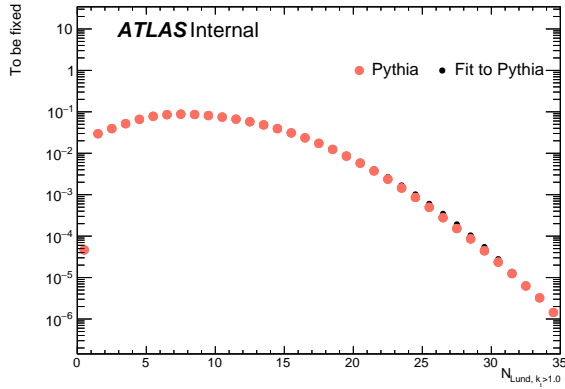
Figure E.3: The (left) fine-binned and (right) coarse-binned multiplicity distributions for (top) $300 < p_T < 500$ GeV, and (bottom) $1250 < p_T < 4500$ GeV where the multiplicity distribution has been fit using the fine-binned distribution.



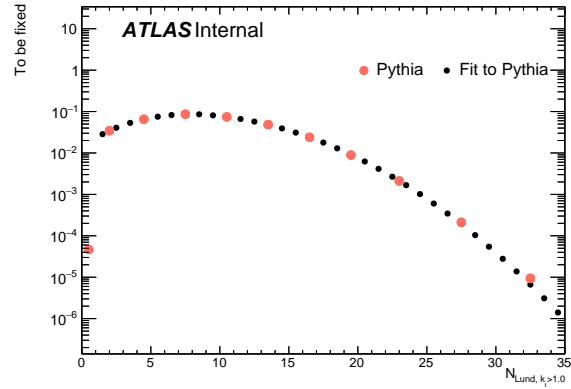
(a) $300 < p_T < 500$ GeV, Fine bins



(b) $300 < p_T < 500$ GeV, Coarse bins



(c) $1250 < p_T < 4500$ GeV, Fine bins



(d) $1250 < p_T < 4500$ GeV, Coarse bins

Figure E.4: The (left) fine-binned and (right) coarse-binned multiplicity distributions for (top) $300 < p_T < 500$ GeV, and (bottom) $1250 < p_T < 4500$ GeV where the multiplicity distribution has been fit using the coarse-binned distribution.

p_T bin	MC	Fine-bin average	Coarse-bin fit average	Binning correction
300 – 500 GeV	Pythia	6.1435	6.20242	1.00959
	Sherpa	5.89467	5.95288	1.00987
	Herwig	5.652	5.71149	1.01053
	Data	5.99212	6.05061	1.00976
500 – 750 GeV	Pythia	7.11274	7.19013	1.01088
	Sherpa	6.8601	6.9329	1.01061
	Herwig	6.54029	6.59978	1.0091
	Data	6.93468	7.00546	1.01021
750 – 1000 GeV	Pythia	7.91185	7.99141	1.01006
	Sherpa	7.64496	7.71907	1.00969
	Herwig	7.3081	7.37013	1.00849
	Data	7.68916	7.76031	1.00925
1000 – 1250 GeV	Pythia	8.53571	8.60505	1.00812
	Sherpa	8.27549	8.34053	1.00786
	Herwig	7.92318	7.97773	1.00688
	Data	8.26036	8.32371	1.00767
1250 – 40000 GeV	Pythia	9.16619	9.21868	1.00573
	Sherpa	8.87924	8.93075	1.0058
	Herwig	8.58055	8.62996	1.00576
	Data	8.8026	8.85359	1.00579

Table E.2: Comparison of the raw averages between the fine-binned distribution and the averages from the fit to the coarse-binned distribution.

BIBLIOGRAPHY

- [1] Tim Adye. “Unfolding algorithms and tests using RooUnfold”. In: *Proceedings, 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding (PHYSTAT 2011)* (CERN, Geneva, Switzerland, Jan. 17–20, 2011), pp. 313–318. DOI: 10.5170/CERN-2011-006.313. arXiv: 1105.1160 [physics.data-an] (cit. on p. 200).
- [2] S. Agostinelli et al. “GEANT4 – a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 138).
- [3] G. Aielli et al. “The RPC first level muon trigger in the barrel of the ATLAS experiment”. In: *Nuclear Physics B - Proceedings Supplements* 158 (2006). Proceedings of the 8th International Workshop on Resistive Plate Chambers and Related Detectors, pp. 11–15. ISSN: 0920-5632. DOI: <https://doi.org/10.1016/j.nuclphysbps.2006.07.031>. URL: <http://www.sciencedirect.com/science/article/pii/S0920563206004178> (cit. on p. 49).
- [4] M (CERN) Aleksa et al. “ATLAS Liquid Argon Calorimeter Phase-I Upgrade Technical Design Report”. In: CERN-LHCC-2013-017. ATLAS-TDR-022 (Sept. 2013). Final version presented to December 2013 LHCC. URL: <https://cds.cern.ch/record/1602230> (cit. on pp. 47, 48).
- [5] Simone Alioli 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] (cit. on p. 185).
- [6] Simone Alioli et al. “Jet pair production in POWHEG”. In: *JHEP* 04 (2011), p. 081. DOI: 10.1007/JHEP04(2011)081. arXiv: 1012.3380 [hep-ph] (cit. on p. 185).
- [7] Steven W. Allen, August E. Evrard, and Adam B. Mantz. “Cosmological Parameters from Observations of Galaxy Clusters”. In: *Annual Review of Astronomy and Astrophysics* 49.1 (Sept. 2011), pp. 409–470. DOI: 10.1146/annurev-astro-081710-102514. URL: <https://doi.org/10.1146%5C%2Fannurev-astro-081710-102514> (cit. on p. 16).
- [8] Johan Alwall et al. “MadGraph 5 : Going Beyond”. In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph] (cit. on p. 130).
- [9] T. Argyropoulos et al. “Cathode strip chambers in ATLAS : Installation, commissioning and in situ performance”. In: *2008 IEEE Nuclear Science Symposium Conference Record*. Oct. 2008, pp. 2819–2824. DOI: 10.1109/NSSMIC.2008.4774958 (cit. on p. 49).
- [10] Benoit Assi and Stefan Höche. “A new approach to QCD evolution in processes with massive partons”. In: (July 2023). arXiv: 2307.00728 [hep-ph] (cit. on p. 275).

- [11] ATLAS Collaboration. “Alignment of the ATLAS Inner Detector in Run-2”. In: *Eur. Phys. J. C* 80 (2020), p. 1194. DOI: 10.1140/epjc/s10052-020-08700-6. arXiv: 2007.07624 [hep-ex] (cit. on p. 255).
- [12] ATLAS Collaboration. “ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex] (cit. on p. 143).
- [13] ATLAS Collaboration. “ATLAS data quality operations and performance for 2015–2018 data-taking”. In: *JINST* 15 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det] (cit. on pp. 139, 183).
- [14] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: <https://cds.cern.ch/record/1966419> (cit. on p. 184).
- [15] ATLAS Collaboration. *Boosted hadronic vector boson and top quark tagging with ATLAS using Run 2 data*. ATL-PHYS-PUB-2020-017. 2020. URL: <https://cds.cern.ch/record/2724149> (cit. on pp. 144, 171).
- [16] ATLAS Collaboration. *Dependence of the Jet Energy Scale on the Particle Content of Hadronic Jets in the ATLAS Detector Simulation*. ATL-PHYS-PUB-2022-021. 2022. URL: <https://cds.cern.ch/record/2808016> (cit. on pp. 184, 274).
- [17] ATLAS Collaboration. *Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-051. 2015. URL: <https://cds.cern.ch/record/2110140> (cit. on p. 194).
- [18] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data”. In: *JINST* 14 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex] (cit. on pp. 142, 143, 171).
- [19] ATLAS Collaboration. “Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector”. In: *Eur. Phys. J. C* 77 (2017), p. 580. DOI: 10.1140/epjc/s10052-017-5081-5. arXiv: 1705.02211 [hep-ex] (cit. on p. 192). Erratum: in: *Eur. Phys. J. C* 77 (2017), p. 712. DOI: 10.1140/epjc/s10052-017-5245-3.
- [20] ATLAS Collaboration. *Identification of Boosted, Hadronically-Decaying W and Z Bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS*. ATL-PHYS-PUB-2015-033. 2015. URL: <https://cds.cern.ch/record/2041461> (cit. on p. 143).
- [21] ATLAS Collaboration. *Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2*. ATL-PHYS-PUB-2018-009. 2018. URL: <https://cds.cern.ch/record/2630327> (cit. on p. 177).

- [22] ATLAS Collaboration. “In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector”. In: *Eur. Phys. J. C* 79 (2019), p. 135. DOI: 10.1140/epjc/s10052-019-6632-8. arXiv: 1807.09477 [hep-ex] (cit. on p. 171).
- [23] ATLAS Collaboration. “Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 81 (2020), p. 689. DOI: 10.1140/epjc/s10052-021-09402-3. arXiv: 2007.02645 [hep-ex] (cit. on pp. 245, 249).
- [24] ATLAS Collaboration. *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*. ATLAS-CONF-2016-035. 2016. URL: <https://cds.cern.ch/record/2200211> (cit. on p. 143).
- [25] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex] (cit. on pp. 143, 190).
- [26] ATLAS Collaboration. “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”. In: (2022). arXiv: 2212.09379 [hep-ex] (cit. on pp. 26, 139, 182).
- [27] ATLAS Collaboration. “Measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV”. In: *Phys. Rev. D* 101 (2020), p. 052007. DOI: 10.1103/PhysRevD.101.052007. arXiv: 1912.09837 [hep-ex] (cit. on pp. 180, 195).
- [28] ATLAS Collaboration. “Measurement of the energy response of the ATLAS calorimeter to charged pions from $W^\pm \rightarrow \tau^\pm(\rightarrow \pi^\pm \nu_\tau)\nu_\tau$ events in Run 2 data”. In: *Eur. Phys. J. C* 82 (2021), p. 223. DOI: 10.1140/epjc/s10052-022-10117-2. arXiv: 2108.09043 [hep-ex] (cit. on p. 247).
- [29] ATLAS Collaboration. “Measurement of the Lund Jet Plane Using Charged Particles in 13 TeV Proton–Proton Collisions with the ATLAS Detector”. In: *Phys. Rev. Lett.* 124 (2020), p. 222002. DOI: 10.1103/PhysRevLett.124.222002. arXiv: 2004.03540 [hep-ex] (cit. on p. 180).
- [30] ATLAS Collaboration. “Measurement of the Soft-Drop Jet Mass in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Rev. Lett.* 121 (2018), p. 092001. DOI: 10.1103/PhysRevLett.121.092001. arXiv: 1711.08341 [hep-ex] (cit. on p. 180).
- [31] ATLAS Collaboration. *Measurement of track reconstruction inefficiencies in the core of jets via pixel dE/dx with the ATLAS experiment using $\sqrt{s} = 13$ TeV pp collision data*. ATL-PHYS-PUB-2016-007. 2016. URL: <https://cds.cern.ch/record/2140460> (cit. on p. 254).

- [32] ATLAS Collaboration. *Modelling of Track Reconstruction Inside Jets with the 2016 ATLAS $\sqrt{s} = 13$ TeV pp Dataset*. ATL-PHYS-PUB-2017-016. 2017. URL: <https://cds.cern.ch/record/2275639> (cit. on p. 254).
- [33] ATLAS Collaboration. *Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2*. ATL-PHYS-PUB-2016-003. 2016. URL: <https://cds.cern.ch/record/2120133> (cit. on p. 172).
- [34] ATLAS Collaboration. *Multijet simulation for 13 TeV ATLAS Analyses*. ATL-PHYS-PUB-2019-017. 2019. URL: <https://cds.cern.ch/record/2672252> (cit. on p. 185).
- [35] ATLAS Collaboration. “Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 81 (2020), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex] (cit. on p. 142).
- [36] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex] (cit. on p. 22).
- [37] ATLAS Collaboration. “Performance of electron and photon triggers in ATLAS during LHC Run 2”. In: *Eur. Phys. J. C* 80 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex] (cit. on p. 171).
- [38] ATLAS Collaboration. “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 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex] (cit. on p. 146).
- [39] ATLAS Collaboration. “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 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex] (cit. on p. 192).
- [40] ATLAS Collaboration. “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2”. In: *Eur. Phys. J. C* 77 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex] (cit. on p. 254).
- [41] ATLAS Collaboration. “Performance of top-quark and W -boson tagging with ATLAS in Run 2 of the LHC”. In: *Eur. Phys. J. C* 79 (2019), p. 375. DOI: 10.1140/epjc/s10052-019-6847-8. arXiv: 1808.07858 [hep-ex] (cit. on p. 144).
- [42] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: <https://cds.cern.ch/record/2037702> (cit. on p. 192).

- [43] ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *Eur. Phys. J. C* 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det] (cit. on p. 138).
- [44] ATLAS Collaboration. *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*. ATL-PHYS-PUB-2016-017. 2016. URL: <https://cds.cern.ch/record/2206965> (cit. on pp. 138, 185).
- [45] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *Eur. Phys. J. C* 77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex] (cit. on p. 143).
- [46] “ATLAS inner detector: Technical Design Report, 1”. In: Technical Design Report ATLAS (1997). URL: <http://cds.cern.ch/record/331063> (cit. on p. 36).
- [47] “ATLAS liquid-argon calorimeter: Technical Design Report”. In: Technical Design Report ATLAS (1996). URL: <https://cds.cern.ch/record/331061> (cit. on p. 38).
- [48] “ATLAS muon spectrometer: Technical Design Report”. In: Technical Design Report ATLAS (1997). URL: <https://cds.cern.ch/record/331068> (cit. on p. 49).
- [49] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331062> (cit. on p. 38).
- [50] Benjamin Audren et al. “Strongest model-independent bound on the lifetime of Dark Matter”. In: *Journal of Cosmology and Astroparticle Physics* 2014.12 (Dec. 2014), p. 028. DOI: 10.1088/1475-7516/2014/12/028. URL: <https://dx.doi.org/10.1088/1475-7516/2014/12/028> (cit. on p. 16).
- [51] G. Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017 (cit. on pp. 139, 182).
- [52] G. Ingelman B. Andersson G. Gustafson and T. Sjostrand. “Parton Fragmentation and String Dynamics”. In: *Physics Reports* 97 (1983), pp. 31–145. DOI: doi:10.1016/0370-1573(83)90080-7 (cit. on p. 130).
- [53] B. R. Webber. “A QCD Model for Jet Fragmentation Including Soft Gluon Interference”. In: *Nuclear Physics B* 238 (1984), pp. 492–528. DOI: doi:10.1016/0550-3213(84)90333-X (cit. on p. 130).
- [54] M. Baak et al. “HistFitter software framework for statistical data analysis”. In: *Eur. Phys. J. C* 75 (2015), p. 153. DOI: 10.1140/epjc/s10052-015-3327-7. arXiv: 1410.1280 [hep-ex] (cit. on p. 172).

- [55] M. Bähr et al. “Herwig++ physics and manual”. In: *Eur. Phys. J. C* 58 (2008), p. 639. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph] (cit. on p. 184).
- [56] Alan Barr, Christopher Lester, and P. Stephens. “A variable for measuring masses at hadron colliders when missing energy is expected; 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 (cit. on p. 147).
- [57] F. Bauer et al. “Construction and test of MDT chambers for the ATLAS muon spectrometer”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 461.1 (2001). 8th Pisa Meeting on Advanced Detectors, pp. 17–20. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(00\)01156-6](https://doi.org/10.1016/S0168-9002(00)01156-6). URL: <http://www.sciencedirect.com/science/article/pii/S0168900200011566> (cit. on p. 49).
- [58] Melissa van Beekveld and Silvia Ferrario Ravasio. “Next-to-leading-logarithmic PanScales showers for Deep Inelastic Scattering and Vector Boson Fusion”. In: (May 2023). arXiv: 2305.08645 [hep-ph] (cit. on p. 177).
- [59] Melissa van Beekveld et al. “PanScales parton showers for hadron collisions: formulation and fixed-order studies”. In: *JHEP* 11 (2022), p. 019. DOI: 10.1007/JHEP11(2022)019. arXiv: 2205.02237 [hep-ph] (cit. on p. 177).
- [60] Melissa van Beekveld et al. “PanScales showers for hadron collisions: all-order validation”. In: *JHEP* 11 (2022), p. 020. DOI: 10.1007/JHEP11(2022)020. arXiv: 2207.09467 [hep-ph] (cit. on p. 177).
- [61] W. Beenakker et al. “Stop production at hadron colliders”. In: *Nucl. Phys. B* 515 (1998), pp. 3–14. DOI: 10.1016/S0550-3213(98)00014-5. arXiv: hep-ph/9710451 (cit. on p. 138).
- [62] Wim Beenakker et al. “NNLL resummation for stop pair-production at the LHC”. In: *JHEP* 05 (2016), p. 153. DOI: 10.1007/JHEP05(2016)153. arXiv: 1601.02954 [hep-ph] (cit. on p. 138).
- [63] Wim Beenakker et al. “NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation”. In: *JHEP* 12 (2016), p. 133. DOI: 10.1007/JHEP12(2016)133. arXiv: 1607.07741 [hep-ph] (cit. on p. 138).
- [64] Wim Beenakker et al. “Supersymmetric top and bottom squark production at hadron colliders”. In: *JHEP* 08 (2010), p. 098. DOI: 10.1007/JHEP08(2010)098. arXiv: 1006.4771 [hep-ph] (cit. on p. 138).

- [65] Michael Begel et al. “Global Feature Extractor of the Level-1 Calorimeter Trigger: ATLAS TDAQ Phase-I Upgrade gFEX Final Design Report”. In: (2016). URL: <http://cds.cern.ch/record/2233958> (cit. on pp. 64–66, 71).
- [66] Johannes Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph] (cit. on pp. 172, 184).
- [67] Johannes Bellm et al. “Herwig 7.1 Release Note”. In: (2017). arXiv: 1705.06919 [hep-ph] (cit. on p. 184).
- [68] Michael Benedikt et al. *LHC Design Report Volume 3: The LHC Injector Chain*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-3. URL: <http://cds.cern.ch/record/823808> (cit. on p. 21).
- [69] Daniele Bertolini, Tucker Chan, and Jesse Thaler. “Jet observables without jet algorithms”. In: *Journal of High Energy Physics* 2014.4 (Apr. 2014). DOI: 10.1007/jhep04(2014)013. URL: [https://doi.org/10.1007/JHEP04\(2014\)013](https://doi.org/10.1007/JHEP04(2014)013) (cit. on pp. 75, 82).
- [70] Gavin Bewick et al. “Initial state radiation in the Herwig 7 angular-ordered parton shower”. In: *JHEP* 01 (2022), p. 026. DOI: 10.1007/JHEP01(2022)026. arXiv: 2107.04051 [hep-ph] (cit. on p. 177).
- [71] Gavin Bewick et al. “Logarithmic accuracy of angular-ordered parton showers”. In: *JHEP* 04 (2020), p. 019. DOI: 10.1007/JHEP04(2020)019. arXiv: 1904.11866 [hep-ph] (cit. on p. 177).
- [72] Oliver Sim Bruning et al. *LHC Design Report Volume 1: The Main Ring*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: <https://cds.cern.ch/record/782076> (cit. on p. 21).
- [73] Oliver Sim Bruning et al. *LHC Design Report Volume 2: The LHC Infrastructure and General Services*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-2. URL: <https://cds.cern.ch/record/815187> (cit. on p. 21).
- [74] Rene Brun and Fons Rademakers. “ROOT — An object oriented data analysis framework”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 81–86. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X). URL: <https://www.sciencedirect.com/science/article/pii/S016890029700048X> (cit. on p. 154).

- [75] Jon Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph] (cit. on pp. 138, 172).
- [76] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet user manual”. In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph] (cit. on pp. 143, 190).
- [77] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “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] (cit. on pp. 131, 143).
- [78] John M. Campbell et al. “Towards NNLO+PS matching with sector showers”. In: *Phys. Lett. B* 836 (2023), p. 137614. DOI: 10.1016/j.physletb.2022.137614. arXiv: 2108.07133 [hep-ph] (cit. on p. 177).
- [79] Stefano Catani and Michael H. Seymour. “A General Algorithm for Calculating Jet Cross Sections in NLO QCD”. In: 485 (Feb. 1997). DOI: 10.1016/S0550-3213(96)00589-5. URL: [https://doi.org/10.1016/S0550-3213\(96\)00589-5](https://doi.org/10.1016/S0550-3213(96)00589-5) (cit. on p. 275).
- [80] CERN. *Accelerating: Radiofrequency cavities*. Accessed: 2 February 2023. URL: <http://home.cern/science/engineering/accelerating-radiofrequency-cavities> (cit. on p. 23).
- [81] CERN. *Pulling Together: Superconducting Electromagnets*. Accessed: 20 July 2023. URL: <https://home.cern/science/engineering/pulling-together-superconducting-electromagnets> (cit. on p. 23).
- [82] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex] (cit. on p. 22).
- [83] CMS Collaboration. “Study of quark and gluon jet substructure in Z +jet and dijet events from pp collisions”. In: *JHEP* 01 (2021), p. 188. DOI: 10.1007/JHEP01(2021)188. arXiv: 2109.03340 [hep-ex] (cit. on p. 180).
- [84] Raymond T. Co and Keisuke Harigaya. “Axiogenesis”. In: *Physical Review Letters* 124.11 (Mar. 2020). DOI: 10.1103/physrevlett.124.111602. URL: <https://doi.org/10.1103/PhysRevLett.124.111602> (cit. on p. 17).
- [85] ATLAS Collaboration. “The silicon microstrip sensors of the ATLAS semiconductor tracker”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 578.1 (2007), pp. 98–118. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2007.04.157>. URL:

- <http://www.sciencedirect.com/science/article/pii/S0168900207007644> (cit. on p. 37).
- [86] The ATLAS Collaboration. “Alignment of the ATLAS Inner Detector in Run-2”. In: *Eur. Phys. J. C* 80 (12 2020), p. 1194. DOI: 10.1140/epjc/s10052-020-08700-6. arXiv: 2007.07624 [hep-ph] (cit. on p. 37).
 - [87] The ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 2*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: <http://cds.cern.ch/record/391177> (cit. on p. 31).
 - [88] The ATLAS Collaboration. “Performance of the ATLAS Trigger System in 2010”. In: *The European Physical Journal C - Particles and Fields* 72.1849 (2012). URL: <https://doi.org/10.1140/epjc/s10052-011-1849-1> (cit. on p. 51).
 - [89] The ATLAS Collaboration. “Performance of the ATLAS trigger system in 2015”. In: *The European Physical Journal C - Particles and Fields* 77.317 (2017). URL: <https://doi.org/10.1140/epjc/s10052-017-4852-3> (cit. on p. 51).
 - [90] The ATLAS Collaboration. “Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physical Review D* 104.11 (Dec. 2021). DOI: 10.1103/physrevd.104.112010. URL: <https://doi.org/10.1103/PhysRevD.104.112010> (cit. on pp. 133, 134, 162, 167, 170, 173, 174).
 - [91] The ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003. URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08003> (cit. on pp. 31, 34, 36, 39, 43, 45, 50).
 - [92] The ATLAS TRT collaboration. “The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance”. In: *Journal of Instrumentation* 3.02 (2008), P02013. URL: <http://stacks.iop.org/1748-0221/3/i=02/a=P02013> (cit. on p. 37).
 - [93] The ATLAS TRT collaboration. “The ATLAS TRT Barrel Detector”. In: *Journal of Instrumentation* 3.02 (2008), P02014. URL: <http://stacks.iop.org/1748-0221/3/i=02/a=P02014> (cit. on p. 37).
 - [94] The ATLAS Collaborations. *ATLAS Experiment - Public Results*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults> (cit. on pp. 27, 29).
 - [95] The ATLAS Collaborations. *Level-1 Calorimeter Trigger Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/L1CaloTriggerPublicResults#ATLAS_Level_1_calorimeter_trigge (cit. on p. 84).

- [96] G. Corcella et al. “HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)”. In: *JHEP* 01 (2001), p. 010. DOI: 10.1088/1126-6708/2001/01/010. arXiv: hep-ph/0011363 (cit. on p. 130).
- [97] Robert D. Cousins, James T. Linnemann, and Jordan Tucker. “Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process”. In: *Nucl. Instrum. Meth. A* 595.2 (2008), p. 480. DOI: 10.1016/j.nima.2008.07.086. arXiv: physics/0702156 [physics.data-an] (cit. on p. 173).
- [98] Glen Cowan and Eilam Gross. “Discovery significance with statistical uncertainty in the background estimate”. In: *ATLAS Statistics Forum* (May 2008). URL: <https://www.pp.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf> (cit. on p. 152).
- [99] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011). [Erratum: *Eur. Phys. J. C* **73** (2013) 2501], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an] (cit. on pp. 172, 173).
- [100] G. D’Agostini. “A multidimensional unfolding method based on Bayes’ theorem”. In: *Nucl. Instrum. Meth. A* 362 (1995), pp. 487–498. DOI: 10.1016/0168-9002(95)00274-X (cit. on p. 198).
- [101] Mrinal Dasgupta et al. “Parton showers beyond leading logarithmic accuracy”. In: *Phys. Rev. Lett.* 125.5 (2020), p. 052002. DOI: 10.1103/PhysRevLett.125.052002. arXiv: 2002.11114 [hep-ph] (cit. on pp. 177, 178).
- [102] Sacha Davidson, Enrico Nardi, and Yosef Nir. “Leptogenesis”. In: *Physics Reports* 466.4-5 (Sept. 2008), pp. 105–177. DOI: 10.1016/j.physrep.2008.06.002. URL: <https://doi.org/10.1016%5C%2Fj.physrep.2008.06.002> (cit. on p. 17).
- [103] Emily Dickinson. *Complete Poems of Emily Dickinson*. Ed. by Thomas H. Johnson. The Belknap Press of Harvard University Press, 1951 (cit. on p. 19).
- [104] Frédéric A. Dreyer, Gavin P. Salam, and Grégory Soyez. “The Lund Jet Plane”. In: *JHEP* 12 (2018), p. 064. DOI: 10.1007/JHEP12(2018)064. arXiv: 1807.04758 [hep-ph] (cit. on pp. 178, 179).
- [105] Falko Dulat, Stefan Höche, and Stefan Prestel. “Leading-Color Fully Differential Two-Loop Soft Corrections to QCD Dipole Showers”. In: *Phys. Rev. D* 98.7 (2018), p. 074013. DOI: 10.1103/PhysRevD.98.074013. arXiv: 1805.03757 [hep-ph] (cit. on pp. 177, 275).
- [106] Sayipjamal Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Phys. Rev. D* 93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph] (cit. on p. 184).

- [107] E F Eisenhandler. “ATLAS Level-1 Calorimeter Trigger Algorithms”. In: (2004). URL: <http://cds.cern.ch/record/792528> (cit. on p. 54).
- [108] John Ellis et al. “Supersymmetric relics from the big bang”. In: *Nucl. Phys. B* 238 (1984), p. 453. DOI: 10.1016/0550-3213(84)90461-9 (cit. on p. 20).
- [109] Yuji Enari and on behalf of the ATLAS Collaboration. “The Phase-1 Trigger Readout Electronics Upgrade of the ATLAS Liquid Argon Calorimeter”. In: *Journal of Physics: Conference Series* 1162.1 (Jan. 2019), p. 012041. DOI: 10.1088/1742-6596/1162/1/012041. URL: <https://dx.doi.org/10.1088/1742-6596/1162/1/012041> (cit. on p. 48).
- [110] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *JINST* 3.08 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08001> (cit. on pp. 21, 23, 26).
- [111] Christian Wolfgang Fabjan and F Gianotti. “Calorimetry for Particle Physics”. In: *Rev. Mod. Phys.* 75 (2003), pp. 1243–1286. DOI: 10.1103/RevModPhys.75.1243. URL: <https://cds.cern.ch/record/692252> (cit. on p. 46).
- [112] Glennys R. Farrar and Pierre Fayet. “Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry”. In: *Phys. Lett. B* 76 (1978), p. 575. DOI: 10.1016/0370-2693(78)90858-4 (cit. on pp. 20, 134).
- [113] Pierre Fayet. “Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions”. In: *Phys. Lett. B* 69 (1977), p. 489. DOI: 10.1016/0370-2693(77)90852-8 (cit. on p. 20).
- [114] Pierre Fayet. “Supersymmetry and weak, electromagnetic and strong interactions”. In: *Phys. Lett. B* 64 (1976), p. 159. DOI: 10.1016/0370-2693(76)90319-1 (cit. on p. 20).
- [115] S. Ferrara and B. Zumino. “Supergauge invariant Yang-Mills theories”. In: *Nucl. Phys. B* 79 (1974), p. 413. DOI: 10.1016/0550-3213(74)90559-8 (cit. on p. 19).
- [116] Jeffrey R. Forshaw, Jack Holguin, and Simon Plätzer. “Building a consistent parton shower”. In: *JHEP* 09 (2020), p. 014. DOI: 10.1007/JHEP09(2020)014. arXiv: 2003.06400 [hep-ph] (cit. on p. 177).
- [117] Linux Foundation and The Yocto Project. *Bitbake Documentation*. Accessed: 23 July 2023. URL: <https://docs.yoctoproject.org/bitbake.html> (cit. on p. 94).
- [118] The OpenEmbedded Framework. *OpenEmbedded Wiki*. Accessed: 15 July 2023. URL: http://www.openembedded.org/wiki/Main_Page (cit. on p. 91).
- [119] S. Frixione and B. R. Webber. “The MC@NLO Event Generator”. In: (2002). arXiv: hep-ph/0207182 [hep-ph] (cit. on p. 130).

- [120] Stefano Frixione, Paolo Nason, and Carlo Oleari. “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] (cit. on p. 185).
- [121] Leif Gellersen, Stefan Höche, and Stefan Prestel. “Disentangling soft and collinear effects in QCD parton showers”. In: *Phys. Rev. D* 105.11 (2022), p. 114012. DOI: 10.1103/PhysRevD.105.114012. arXiv: 2110.05964 [hep-ph] (cit. on p. 177).
- [122] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph] (cit. on pp. 130, 184).
- [123] H. Goldberg. “Constraint on the Photino Mass from Cosmology”. In: *Phys. Rev. Lett.* 50 (1983), p. 1419. DOI: 10.1103/PhysRevLett.50.1419 (cit. on p. 20). Erratum: in: *Phys. Rev. Lett.* 103 (2009), p. 099905. DOI: 10.1103/PhysRevLett.103.099905.
- [124] Y.A. Golfand and E.P. Likhtman. “Extension of the Algebra of Poincare Group Generators and Violation of P Invariance”. In: *JETP Lett.* 13 (1971). [*Pisma Zh. Eksp. Teor. Fiz.* **13** (1971) 452], p. 323 (cit. on p. 19).
- [125] ATLAS Quasar Group. Quasar MilkyWay GitLab. URL: <https://github.com/quasar-team/MilkyWay> (cit. on p. 110).
- [126] ATLAS Standard Model Working Group. <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/StandardModelUnfolding> (cit. on p. 220).
- [127] A. J. S. Hamilton and Max Tegmark. “The real-space power spectrum of the PSCz survey from 0.01 to 300 h Mpc⁻¹”. In: *Monthly Notices of the Royal Astronomical Society* 330.3 (Mar. 2002), pp. 506–530. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2002.05033.x. eprint: <https://academic.oup.com/mnras/article-pdf/330/3/506/18409902/330-3-506.pdf>. URL: <https://doi.org/10.1046/j.1365-8711.2002.05033.x> (cit. on p. 16).
- [128] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph] (cit. on p. 184).
- [129] L. Hartgring, E. Laenen, and P. Skands. “Antenna Showers with One-Loop Matrix Elements”. In: *JHEP* 10 (2013), p. 127. DOI: 10.1007/JHEP10(2013)127. arXiv: 1303.4974 [hep-ph] (cit. on p. 177).
- [130] Florian Herren et al. “A new approach to color-coherent parton evolution”. In: (Aug. 2022). arXiv: 2208.06057 [hep-ph] (cit. on pp. 177, 275).
- [131] Stefan Höche, Frank Krauss, and Stefan Prestel. “Implementing NLO DGLAP evolution in Parton Showers”. In: *JHEP* 10 (2017), p. 093. DOI: 10.1007/JHEP10(2017)093. arXiv: 1705.00982 [hep-ph] (cit. on p. 177).

- [132] Stefan Höche and Stefan Prestel. “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] (cit. on p. 184).
- [133] Stefan Höche and Stefan Prestel. “Triple collinear emissions in parton showers”. In: *Phys. Rev. D* 96.7 (2017), p. 074017. DOI: 10.1103/PhysRevD.96.074017. arXiv: 1705.00742 [hep-ph] (cit. on p. 177).
- [134] *et al.* J. E. Huth. “Toward a standardization of jet definitions”. In: *1990 DPF Summer Study on High-energy Physics (Snowmass 1990) : Research directions for the decade (1990)*, pp. 134–136 (cit. on p. 130).
- [135] S. Jadach et al. “NLO corrections in the initial-state parton shower Monte Carlo”. In: *Acta Phys. Polon. B* 44.11 (2013), pp. 2179–2187. DOI: 10.5506/PhysPolB.44.2179. arXiv: 1310.6090 [hep-ph] (cit. on p. 177).
- [136] S. Jadach et al. “On the dependence of QCD splitting functions on the choice of the evolution variable”. In: *JHEP* 08 (2016), p. 092. DOI: 10.1007/JHEP08(2016)092. arXiv: 1606.01238 [hep-ph] (cit. on p. 177).
- [137] S. Jadach et al. “Two real parton contributions to non-singlet kernels for exclusive QCD DGLAP evolution”. In: *JHEP* 08 (2011), p. 012. DOI: 10.1007/JHEP08(2011)012. arXiv: 1102.5083 [hep-ph] (cit. on p. 177).
- [138] Brandon Kunkler. *gFEX Production Readout*. Indiana University. June 2016. URL: https://gitlab.cern.ch/atlas-l1calo/gfex/firmware/-/blob/devel/doc/gFEX_Production_Readout.docx (cit. on p. 103).
- [139] H.-L. Lai et al. “New parton distributions for collider physics”. In: *Phys. Rev. D* 82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph] (cit. on p. 184).
- [140] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152. DOI: 10.1016/S0168-9002(01)00089-4 (cit. on p. 184).
- [141] Andrew J. Larkoski, Ian Moutl, and Duff Neill. “Power counting to better jet observables”. In: *JHEP* 12 (2014), p. 009. DOI: 10.1007/JHEP12(2014)009. arXiv: 1409.6298 [hep-ph] (cit. on p. 144).
- [142] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. “Energy correlation functions for jet substructure”. In: *JHEP* 06 (2013), p. 108. DOI: 10.1007/JHEP06(2013)108. arXiv: 1305.0007 [hep-ph] (cit. on p. 144).
- [143] C. Ghabrous Larrea et al. “IPbus: a flexible Ethernet-based control system for xTCA hardware”. In: *Journal of Instrumentation* 10.02 (Feb. 2015), p. C02019. DOI: 10.1088/1748-0221/10/02/C02019. URL: <https://dx.doi.org/10.1088/1748-0221/10/02/C02019> (cit. on p. 105).

- [144] Sehwook Lee. “On the limits of the hadronic energy resolution of calorimeters”. In: *Journal of Physics: Conference Series* 1162 (2018). DOI: 10.1088/1742-6596/1162/1/012043. URL: <https://iopscience.iop.org/article/10.1088/1742-6596/1162/1/012043> (cit. on p. 42).
- [145] Emil Lenchak. *Install and Build with Xilinx Yocto*. Accessed: 25 July 2023. URL: <https://xilinx-wiki.atlassian.net/wiki/spaces/A/pages/18841862/Install+and+Build+with+Xilinx+Yocto> (cit. on p. 95).
- [146] C. G. Lester and D. J. Summers. “Measuring masses of semi-invisibly 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 (cit. on p. 147).
- [147] Hai Tao Li and Peter Skands. “A framework for second-order parton showers”. In: *Phys. Lett. B* 771 (2017), pp. 59–66. DOI: 10.1016/j.physletb.2017.05.011. arXiv: 1611.00013 [hep-ph] (cit. on p. 177).
- [148] Leif Lönnblad. “Correcting the Colour-Dipole Cascade Model with Fixed Order Matrix Elements”. In: *JHEP* 05 (2002), p. 046. DOI: 10.1088/1126-6708/2002/05/046. arXiv: hep-ph/0112284 (cit. on p. 138).
- [149] Pedro A. N. Machado. *Neutrino Properties and Interactions*. 2022. arXiv: 2206.13449 [hep-ph] (cit. on p. 18).
- [150] S. Majewski et al. “A THIN MULTIWIRED CHAMBER OPERATING IN THE HIGH MULTIPLICATION MODE”. In: *Nucl. Instrum. Meth.* 217 (1983), pp. 265–271. DOI: 10.1016/0167-5087(83)90146-1 (cit. on p. 49).
- [151] Z Marshall. *Re-defining the Standard QCD Di-Jet Samples: Event Weights Make Me Grumpy*. Tech. rep. ATL-COM-PHYS-2011-992. Geneva: CERN, 2011. URL: <https://cds.cern.ch/record/1370089> (cit. on pp. 183, 185).
- [152] Zach Marshall. *Re-re-defining the Standard QCD Di-Jet Samples: Beginning to Like Event Weights*. Tech. rep. ATL-COM-PHYS-2015-417. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2016630> (cit. on pp. 183, 185).
- [153] Simone Marzani, Gregory Soyez, and Michael Spannowsky. *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*. Vol. 958. Springer, 2019. DOI: 10.1007/978-3-030-15709-8. arXiv: 1901.10342 [hep-ph] (cit. on p. 177).
- [154] Rok Medves, Alba Soto-Ontoso, and Gregory Soyez. “Lund and Cambridge multiplicities for precision physics”. In: *JHEP* 10 (2022), p. 156. DOI: 10.1007/JHEP10(2022)156. arXiv: 2205.02861 [hep-ph] (cit. on pp. 178, 179, 181).
- [155] Rok Medves, Alba Soto-Ontoso, and Gregory Soyez. “Lund multiplicity in QCD jets”. In: (Dec. 2022). arXiv: 2212.05076 [hep-ph] (cit. on p. 178).

- [156] Julian Mendez. *CERN-IPMC: Technical overview*. 2017. URL: https://espace.cern.ch/ph-dep-ESE-BE-ATCAEvaluationProject/PP_IPMC/Public%20documents/CERN-IPMC%20-%20Technical%20overview.pdf (cit. on p. 62).
- [157] *MiniPOD AFBR-814VxyZ, AFBR-82VxyZ 14 Gbps/Channel Twelve Channel, Parallel Fiber Optics Modules*. Avago Technologies Confidential. Restricted under NDA. Avago Technologies. Feb. 2013 (cit. on p. 63).
- [158] Esma Mobs. “The CERN accelerator complex. Complexe des accélérateurs du CERN”. In: (2016). General Photo. URL: <https://cds.cern.ch/record/2197559> (cit. on p. 24).
- [159] David E Morrissey and Michael J Ramsey-Musolf. “Electroweak baryogenesis”. In: *New Journal of Physics* 14.12 (Dec. 2012), p. 125003. DOI: 10.1088/1367-2630/14/12/125003. URL: <https://dx.doi.org/10.1088/1367-2630/14/12/125003> (cit. on p. 17).
- [160] Paris Moschovakos. *Quasar Documentation*. Accessed: 26 July 2023. URL: <https://quasar.docs.cern.ch/> (cit. on p. 110).
- [161] Philippe Mouche. “Overall view of the LHC. Vue d’ensemble du LHC”. In: (2014). General Photo. URL: <https://cds.cern.ch/record/1708847> (cit. on p. 21).
- [162] Zoltán Nagy and Davison E. Soper. “Summations by parton showers of large logarithms in electron-positron annihilation”. In: (Nov. 2020). arXiv: 2011.04777 [hep-ph] (cit. on p. 177).
- [163] Zoltán Nagy and Davison E. Soper. “Summations of large logarithms by parton showers”. In: *Phys. Rev. D* 104.5 (2021), p. 054049. DOI: 10.1103/PhysRevD.104.054049. arXiv: 2011.04773 [hep-ph] (cit. on p. 177).
- [164] Paolo Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 (cit. on p. 185).
- [165] NNPDF Collaboration, Richard D. Ball, et al. “Parton distributions with LHC data”. In: *Nucl. Phys. B* 867 (2013), p. 244. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph] (cit. on pp. 138, 184, 185).
- [166] Terry O’Neal. *Yocto*. Accessed: 23 July 2023. URL: <https://xilinx-wiki.atlassian.net/wiki/spaces/A/pages/18841883/Yocto> (cit. on p. 94).
- [167] and P. A. R. Ade et al. “Planck 2015 results”. In: *Astronomy and Astrophysics* 594 (Sept. 2016), A15. DOI: 10.1051/0004-6361/201525941. URL: <https://doi.org/10.1051/0004-6361/201525941> (cit. on pp. 16, 17).

- [168] Joao Pequenaio and Paul Schaffner. “How ATLAS detects particles: diagram of particle paths in the detector”. 2013. URL: <https://cds.cern.ch/record/1505342> (cit. on p. 32).
- [169] P. Pfeifenschneider et al. “QCD analyses and determinations of $\alpha(s)$ in e^+e^- annihilation at energies between 35-GeV and 189-GeV”. In: *Eur. Phys. J. C* 17 (2000), pp. 19–51. DOI: 10.1007/s100520000432. arXiv: hep-ex/0001055 (cit. on p. 178).
- [170] “Practical considerations for unfolding”. In: *ATL-COM-PHYS-2014-277* (2014). URL: <https://cds.cern.ch/record/1694351> (cit. on p. 220).
- [171] The Yocto Project. *The Yocto Project*. Accessed: 23 July 2023. URL: <https://www.yoctoproject.org/> (cit. on pp. 91, 93).
- [172] Alexander L. Read. “Presentation of search results: the CL_S technique”. In: *J. Phys. G* 28 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313 (cit. on p. 173).
- [173] V. C. Rubin, Jr. Ford W. K., and N. Thonnard. “Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 ($R=4\text{kpc}$) to UGC 2885 ($R=122\text{kpc}$).” In: *The Astrophysical Journal* 238 (June 1980). Provided by the SAO/NASA Astrophysics Data System, pp. 471–487. DOI: 10.1086/158003. URL: <https://ui.adsabs.harvard.edu/abs/1980ApJ...238..471R> (cit. on p. 16).
- [174] M. H. Seymour S. Catani Y. L. Dokshitzer and B. R. Webber. “Longitudinally invariant K_t clustering algorithms for hadron hadron collisions”. In: *Nuclear Physics B* 406 (1993), pp. 187–224. DOI: doi:10.1016/0550-3213(93)90166-M (cit. on p. 131).
- [175] Abdus Salam and J. Strathdee. “Super-symmetry and non-Abelian gauges”. In: *Phys. Lett. B* 51 (1974), p. 353. DOI: 10.1016/0370-2693(74)90226-3 (cit. on p. 19).
- [176] Gavin P. Salam. “Towards Jetography”. In: *Eur. Phys. J. C* 67 (2010), pp. 637–686. DOI: 10.1140/epjc/s10052-010-1314-6. arXiv: 0906.1833 [hep-ph] (cit. on p. 128).
- [177] Paolo Salucci. “The distribution of dark matter in galaxies”. In: *The Astronomy and Astrophysics Review* 27 (1 2019). DOI: 10.1007/s00159-018-0113-1. URL: <https://doi.org/10.1007/s00159-018-0113-1> (cit. on p. 16).
- [178] Steffen Schumann and Frank Krauss. “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] (cit. on p. 184).
- [179] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. “PYTHIA 6.4 physics and manual”. In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 (cit. on pp. 130, 184).

- [180] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph] (cit. on p. 184).
- [181] Skyworks. *AN1006: Differences Between Si534x/8x Revision B and Revision D Silicon*. URL: <https://www.skyworksinc.com/-/media/Skyworks/SL/documents/public/application-notes/AN1006-Si534x-8x-RevB-RevD-Differences.pdf> (cit. on p. 302).
- [182] Skyworks. *Si5345, Si5344, Si5342 Rev. D Family Reference Manual*. URL: <https://www.skyworksinc.com/-/media/Skyworks/SL/documents/public/reference-manuals/Si5345-44-42-D-RM.pdf> (cit. on p. 302).
- [183] Emily Smith and Greg Myers. GitLab Repository. URL: <https://gitlab.cern.ch/atlas-l1calo/gfex/gfex-register-access/-/tree/master> (cit. on pp. 104, 108).
- [184] Michael Stanley, Pratik Patil, and Mikael Kuusela. “Uncertainty quantification for wide-bin unfolding: one-at-a-time strict bounds and prior-optimized confidence intervals”. In: *JINST* 17.10 (2022), P10013. DOI: 10.1088/1748-0221/17/10/P10013. arXiv: 2111.01091 [stat.AP] (cit. on p. 200).
- [185] Giordon Stark. Python Module. URL: <https://pypi.org/project/ironman/> (cit. on p. 105).
- [186] Giordon Stark. “THE SEARCH FOR SUPERSYMMETRY IN HADRONIC FINAL STATES USING BOOSTED OBJECT RECONSTRUCTION”. PhD thesis. University of Chicago, 2018. URL: <https://kratsg.github.io/thesis/thesis.pdf> (cit. on pp. 105–107).
- [187] ATLAS TDAQ. ATLAS TDQ DCS GitLab. URL: https://gitlab.cern.ch/atlas-dcs-subsystems/atlas_dcs_tdq (cit. on p. 114).
- [188] gFEX Team. “ATLAS Phase I Global Feature Extraction (gFEX) Trigger Algorithms”. In: (2020) (cit. on pp. 82, 83).
- [189] gFEX Team. *gFEX Firmware Register Map*. URL: <https://docs.google.com/spreadsheets/d/1w2UuTHTV-Ks9uX0QDzmEI0Zia4EWIYva74c7AlX7IS8/edit?usp=sharing> (cit. on p. 71).
- [190] *Ultrascale Architecture and Product Data Sheet: Overview*. Xilinx. June 2023. URL: <https://docs.xilinx.com/v/u/en-US/ds890-ultrascale-overview> (cit. on p. 66).
- [191] D.V. Volkov and V.P. Akulov. “Is the neutrino a goldstone particle?” In: *Phys. Lett. B* 46 (1973), p. 109. DOI: 10.1016/0370-2693(73)90490-5 (cit. on p. 19).

- [192] B. R. Webber. “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 (cit. on p. 184).
- [193] Norbert Wermes and G Hallewel. “ATLAS pixel detector: Technical Design Report”. In: Technical Design Report ATLAS (1998). URL: <http://cds.cern.ch/record/381263> (cit. on p. 36).
- [194] J. Wess and B. Zumino. “Supergauge invariant extension of quantum electrodynamics”. In: *Nucl. Phys. B* 78 (1974), p. 1. DOI: 10.1016/0550-3213(74)90112-6 (cit. on p. 19).
- [195] J. Wess and B. Zumino. “Supergauge transformations in four dimensions”. In: *Nucl. Phys. B* 70 (1974), p. 39. DOI: 10.1016/0550-3213(74)90355-1 (cit. on p. 19).
- [196] M. Wobisch and T. Wengler. “Hadronization corrections to jet cross-sections in deep inelastic scattering”. In: *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*. 1998, pp. 270–279. arXiv: [hep-ph/9907280](https://arxiv.org/abs/hep-ph/9907280) (cit. on pp. 131, 179, 180).
- [197] R. L. Workman et al. “Review of Particle Physics”. In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097 (cit. on pp. 3, 13, 14, 40, 42).
- [198] Xilinx. URL: <https://www.xilinx.com/products/silicon-devices.html> (cit. on pp. 67–69).
- [199] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber. “Better jet clustering algorithms”. In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: [hep-ph/9707323](https://arxiv.org/abs/hep-ph/9707323) (cit. on pp. 131, 179, 180).