

**DO DIBOSONS DREAM OF SEMILEPTONIC SHEEP?
SEARCHING FOR HEAVY WH RESONANCES AND OPTIMIZING
TRACK RECONSTRUCTION WITH THE ATLAS DETECTOR**

A Dissertation Presented

by

M. A. VESSELLA

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2024

Physics



© Copyright by M. A. Vessella 2024

All Rights Reserved

**DO DIBOSONS DREAM OF SEMILEPTONIC SHEEP?
SEARCHING FOR HEAVY WH RESONANCES AND OPTIMIZING
TRACK RECONSTRUCTION WITH THE ATLAS DETECTOR**

A Dissertation Presented

by

M. A. VESSELLA

Approved as to style and content by:

Carlo Dallapiccola, Chair

Stephane Willocq, Member

Scott Hertel, Member

Grant Wilson, Member

Anthony Dinsmore, Department Chair
Physics

DEDICATION

For every black sheep, electric or otherwise.

Though both are bound in the spiral dance, I would rather be a cyborg than a goddess.

- Donna Haraway, *A Cyborg Manifesto*

*And indeed there will be time
For the yellow smoke that slides along the street,
Rubbing its back upon the window-panes;
There will be time, there will be time
To prepare a face to meet the faces that you meet;
There will be time to murder and create,
And time for all the works and days of hands
That lift and drop a question on your plate;
Time for you and time for me,
And time yet for a hundred indecisions,
And for a hundred visions and revisions,
Before the taking of a toast and tea.*

- T.S Eliot, *The Love Song of J. Alfred Prufrock*

ACKNOWLEDGMENTS

Some may say it takes a village to raise a child, but it’s a lesser-known aphorism that it also takes a village to raise a 28-year-old PhD candidate. To that end, I would most like to thank Nora Pettersson, to whom I owe most of my professional successes in one way or another. The most important lessons you taught me about surviving in this unusual world of ours had very little to do with the physics, though you taught me so much of that, too. I would never have realized how much I would miss the daily ~~gossiping~~ intellectual conversations over MatterMost, or even that I might miss compiling...colorful...debug messages into the ResonanceFinder. It’s hopeless to try to fill your shoes in ATLAS – but I hope, at least, I can keep the spirit alive. 🐶 💜 🐮 I promise I’ll come visit Finland someday.

I also owe my advisor, Carlo Dallapiccola, for a near-decade of mentorship, advice, and good conversation. I can hardly believe it’s been almost ten years. You’ve been there every step of the way, watching me go from a confused undergraduate learning how to make my very first ROOT histogram to (soon-to-be) Dr. Vessella. You’ve shown me the importance of putting in the work where it counts but never losing sight of what’s good and important (friends and family, limoncello, etc. 😊) in this one short life we get to live, a lesson I will carry with me for the rest of my career. I’m grateful for the many opportunities you’ve given me and the trust you’ve placed in me to mostly run my own little operation these past couple of years. I’m proud to have been your student.

Of course, I also owe enormous gratitude to our incredible Vh analysis team, without whom the result in this thesis would never have seen the light of day. In particular, I’d like to thank Dominik Duda for his tireless coordination and seemingly endless availability, Tong Qiu for his monumental efforts in the 0- and 2-lepton channels, the toy limits, and camaraderie against the unending demands of the “peanut gallery”, Andreas Hoenle for his breathtaking scripting skills and name-pronunciation solidarity, Stefan Maschek for tolerating endless iterations of re-deriving the theory uncertainty templates, Tong Li for the tedious task of producing the multijet templates, and Wade Fisher for soothing our anxieties about the statistics and fit model with great regularity.

On a professional level, other individual folks I owe a great deal of thanks to include:

- Max Goblirsch, for spearheading the tracking performance optimization efforts and giving years of support, guidance, and a sense of humor in dealing with my eldritch software development. A mutual love of the "fun stuff" has brought us together in a few unrelated endeavors since, which has given me many opportunities to learn from the best.
- Jackson Burzynski, for helping me in many ways over the years. From answering my inane first-year newbie questions in the UMass office to extending your TVPD convenership through your other important commitments so that I could write this thesis before taking up the mantle myself. I owe you more than a few beers! (One for each week of extra mandate? 😊)
- Gabriel Facini, Valentina Cairo, Noemi Calace, and Christian Greife for vigorously supporting my professional development through their Tracking CP convenerships and beyond.
- Stephane Willocq, also for always unwaveringly supporting my professional development during my time at UMass and not getting *too* annoyed at my incessant travel requests. (I'm sorry I had to go to Australia!)
- Per Johansson, for being a great IDGlobal DQ co-coordinator and teaching me the ropes during the chaotic beginning of Run 3.

I'd also like to give big thanks to my ~~fellow-squirrels~~ tracking performance collaborators Stephen Swatman, Prajita Bhattarai, and Johannes Jungeburth, as well as the many other folks who contributed in some way to the optimization campaign. I also thank the rest of the ATLAS Tracking and DQ community members (you know who you are, and I'm sorry I can't name you all!). I also thank my dHNL collaborators for the fun and exciting analysis after VH – Sagar Addepali, Christian Appelt, Audrey Kvam, Dominique Trischuk, Marzieh Bahmani, Guglielmo Frattari, Margaret Lutz, and Matthias Danniger. This thesis would also have not been possible without the rest of the ATLAS collaboration, who have all contributed in some vital way to every result, and I also thank CERN for the successful operation of the LHC.

Lastly, I can't forget everyone I've met at UMass who has shaped my life in a significant way over the past six years. I thank Verena Martinez Outschoorn, Rafael Coelho de Lopez de Sa, and Ben Brau each for their part in advising us students and leading the UMass group into a prosperous era, as well as Ed Moyse for his bottomless ATLAS computing expertise. I also thank all the ATLAS grad students, postdocs, and staff past and present who aren't acknowledged in some other way in this monologue – Dale Abbott, Yuan-Tang Chou, Samyukta Krishnamurthy, Guy Rosin, Nathan

Bernard, Dan Humphreys, Jay Sandesara, Will Leight, Guillermo Loustau De Linares, Matt Maroun, Andrew Toler, Shani Perera, and anybody else I may be forgetting.

On a personal level, I'd like to thank all my friends and family who have stuck with me on this incredible journey. I'd like to thank Matt Harris for a listening ear and friendly banter in my final years at UMass and Garrett Bell-Paolucci for years of fun times since the good old days. I thank Joyce and Alek Tenerowicz for taking up space in their house while I finished writing this work – hope you will enjoy the peace and solitude soon! I particularly thank my co-conspirator, Cooper Wagner, for the years of steadfast friendship. Oceans, then continents, separated us, but we've always had each other's backs through the best and the worst of it. You know by now that underneath this prickly New Englander exterior of mine is someone irritatingly sentimental – so I'll just say I'm grateful for the many great times we've had over the years and, hopefully, many more to come.

I thank my parents, Jesse and Amy Vessella, for their lifelong encouragement to follow my dreams and their unwavering support through what probably has seemed like an endless journey. When your little girl dreamed of the stars and said she wanted to be a scientist when she grew up, I'm sure you didn't imagine constant travel to far-flung places so far away from our tiny hometown and a life filled with so much uncertainty – but I am so incredibly blessed to have had the opportunity. I love you both and I hope I've made you proud.

Lastly, I thank my fiance, Tristan Tenerowicz, for years of unconditional love, dedication, and support—through both good times and trying ones. It has been such a privilege to learn and grow together with you. I would not be here today without you by my side, and words simply cannot do this justice. I love you more than anything, bug, and I can't wait for this new chapter in our lives. Thanks for sticking with me. 

ABSTRACT

DO DIBOSONS DREAM OF SEMILEPTONIC SHEEP? SEARCHING FOR HEAVY WH RESONANCES AND OPTIMIZING TRACK RECONSTRUCTION WITH THE ATLAS DETECTOR

SEPTEMBER 2024

M. A. VESSELLA

B.Sc., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Carlo Dallapiccola

A search for new heavy vector resonances decaying into a W boson and a Standard-Model Higgs boson (h), targeting a semileptonic final state where the W boson decays into a lepton-neutrino pair and the Higgs boson decays into a b-quark pair, is performed using 139 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ proton-proton collision data collected by the ATLAS detector at the Large Hadron Collider (LHC) during LHC Run 2 (2015-2018). The search probes a wide range of potential heavy resonance masses, from 400 GeV to 5 TeV, by examining the invariant and transverse mass distributions of Wh candidates for a localized excess. Many beyond Standard Model (BSM) theories generically predict such diboson vector resonances, such as the Heavy Vector Triplet (HVT) benchmark model through which the results are interpreted. No significant excess over Standard Model predictions is observed, and 95% confidence level upper limits are placed on the production cross-section times branching ratio. These limits are also converted into constraints on the parameter space of the HVT model. Furthermore, reconstructing the trajectories of charged particles through the ATLAS Inner Detector is a critical component of both this search and the broader ATLAS physics program, but represents an enormous computational challenge due to its inherent combinatorial complexity. This work also presents successful efforts to improve and optimize this process, in both computational resources and physics performance, for LHC Run 3 (2022-2025). Under design accelerator conditions, the improved ATLAS track reconstruction is twice as fast as the legacy version, with no significant reduction in overall efficiency and a more than two-fold reduction in “fake” tracks.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	vi
ABSTRACT	ix
LIST OF TABLES	xiv
LIST OF FIGURES	xvi
 CHAPTER	
INTRODUCTION	1
1. THE STANDARD MODEL AND BEYOND	5
1.1 Standard Model Lagrangian: Symmetries and Groups	6
1.2 Quantum Electrodynamics	9
1.3 Quantum Chromodynamics	10
1.4 Electroweak Theory	12
1.5 Higgs Mechanism	16
1.5.1 Weak Gauge Boson (W/Z) Masses	17
1.5.2 Fermion masses	18
1.5.3 The Higgs Boson	19
1.6 Problems with the Standard Model	20
1.6.1 The Hierarchy Problem	21
1.7 Heavy Vector Triplet (HVT) Model	23
1.7.1 Lagrangian and Model Parameters	23
1.7.2 Benchmark Models	25
1.8 Collision Phenomenology and Monte Carlo Simulations	25
1.8.1 Cross Sections and Decay Rates	26
1.8.2 Hard Scatter and Parton Distribution Functions	28
1.8.3 Parton Showering and Hadronization	29
1.8.4 Underlying Event and Simulation	31
2. THE ATLAS EXPERIMENT	32
2.1 The Large Hadron Collider	32
2.1.1 Injection Chain and Accelerator Complex	34

2.1.2	LHC System	35
2.1.3	Luminosity and Pileup.....	37
2.1.4	LHC Run Schedule and Operations.....	39
2.2	The ATLAS Detector	41
2.2.1	The ATLAS Coordinate System	42
2.2.2	Magnet System	44
2.2.3	The Inner Detector	45
2.2.3.1	Pixel Detector.....	46
2.2.3.2	Semiconductor Tracker (SCT)	47
2.2.3.3	Transition Radiation Tracker (TRT)	49
2.2.4	The Calorimeters	50
2.2.4.1	EM Calorimeter	51
2.2.4.2	Hadronic Calorimeter	53
2.2.5	The Muon Spectrometer	55
2.2.5.1	Monitored Drift Tubes (MDTs)	56
2.2.5.2	Cathode Strip Chambers (CSCs)	58
2.2.5.3	Resistive Plate Chambers (RPCs).....	59
2.2.5.4	Thin Gap Chambers (TGCs)	59
2.2.6	Trigger and Data Acquisition	60
2.2.6.1	Level 1 (L1) Trigger.....	61
2.2.6.2	High-Level Trigger (HLT)	61
2.2.6.3	Trigger Menu and Selections	62
2.3	Event Reconstruction	63
2.3.1	Track and Vertex Reconstruction	65
2.3.1.1	Track Parameterization and Propagation	66
2.3.1.2	Clustering and Space Point Formation	67
2.3.1.3	Track Seeding	69
2.3.1.4	Track Finding	69
2.3.1.5	Ambiguity Resolution	71
2.3.1.6	TRT Extension	72
2.3.1.7	Backtracking	73
2.3.1.8	Large Radius Tracking	75
2.3.1.9	Primary Vertex Reconstruction	77
2.3.2	Topological Calorimeter Clusters	78
2.3.3	Muons	80
2.3.3.1	Muon Reconstruction	81
2.3.3.2	Muon Identification and Calibration.....	84
2.3.4	Jets	85
2.3.4.1	Jet Types and Calibrations	87
2.3.4.2	b-Tagging.....	93
2.3.4.3	Hadronic Taus.....	95

2.3.5	Electrons	96
2.3.5.1	Electron Reconstruction	96
2.3.5.2	Electron Identification and Calibration	97
2.3.6	Missing Transverse Momentum	99
3.	OPTIMIZATION OF THE ATLAS TRACK RECONSTRUCTION	102
3.1	ATLAS Software and Computing Infrastructure	104
3.1.1	Computing Resources	104
3.1.1.1	Tier-0 and HLT	104
3.1.1.2	The Worldwide LHC Computing Grid (WLCG)	105
3.1.2	Software Resources	106
3.1.2.1	Data Processing Workflows	107
3.1.2.2	Athena	111
3.1.2.3	Event Data Model	113
3.1.2.4	Builds, Releases, and Monitoring	114
3.2	Data, MC Simulations, and Benchmarking Methodology	117
3.2.1	Data and MC	117
3.2.2	Benchmarking Methodology	119
3.3	Software Optimization	120
3.3.1	Track Seeding and Finding Optimization	121
3.3.2	TRT Extension and Backtracking Optimization	123
3.3.3	Primary Vertexing Optimization with AMVF and ACTS	124
3.3.4	LRT and Other Optimization	127
3.4	Performance Results	128
3.5	Conclusions and Outlook	135
4.	SEARCH FOR Wh RESONANCES	138
4.1	Data and Monte Carlo Simulations	140
4.1.1	Data and Trigger	140
4.1.2	Signal and Background MC	141
4.2	Event Selection and Analysis Strategy	145
4.2.1	Object Definition and Preselection	145
4.2.2	Overlap Removal and Event Cleaning	148
4.2.3	Higgs Candidate Reconstruction and Selection	149
4.2.4	Event Selection Requirements	151
4.2.4.1	1-Lepton Channel	151
4.2.4.2	0-Lepton Channel	153
4.3	Background Estimation	156
4.4	Systematic Uncertainties	158

4.4.1	Experimental Uncertainties	158
4.4.2	Modeling Uncertainties	162
4.4.2.1	Theory Uncertainties	162
4.4.2.2	Additional Uncertainties	167
4.4.3	Smoothing, Symmetrization, and Pruning	168
4.4.3.1	Smoothing	168
4.4.3.2	Averaging and symmetrization	171
4.4.3.3	Pruning	171
4.5	Statistical Analysis	172
4.5.1	Likelihood Function	172
4.5.2	Statistical Tests and Limit Setting	174
4.5.3	Nuisance Parameter Pulls, Constraints, and Rankings	177
4.6	Results	179
4.7	Conclusion and Outlook	194

APPENDICES

A. MOMENT MORPHING PROCEDURE	197
---	------------

B. BACKGROUND MODELING STUDIES	206
---	------------

BIBLIOGRAPHY	222
---------------------------	------------

LIST OF TABLES

Table	Page
4.1 Summary of the missing transverse momentum and single electron triggers that have been used for the 2015-2018 datataking periods. The number following xe or e corresponds to the minimum E_T^{miss} or E_T requirement in GeV. Text following lh is the isolation requirement, while mht or $pufit$ refer to the calorimeter clustering method used to calculate online E_T . The $pufit$ algorithm corrects for pileup effects before being used as input based on a μ dependent threshold. $nod0$ specifies that the transverse impact parameter is not used in the electron identification, while $etcut$ specifies that only E_T information is used in the trigger selection. $L1EM$ specify a Level 1 seed with the number specifying the p_T requirement. VH specifies several topologically dependent electron vetos. ^(*) available before subperiod D3 ^(**) available after subperiod D3	141
4.2 Summary of the Monte Carlo event generators used to produce the various background processes. The perturbative accuracy (in QCD and, if relevant, in EW corrections) of the total production cross section σ_{prod} is stated for each process. The order at which the corresponding matrix elements are calculated in the Monte Carlo simulation is not necessarily the same as for the cross section. Details on the final state are also provided, where $\ell = e, \mu$, and τ	145
4.3 A list of the signal and control regions included in the statistical analysis of the W' signal model hypothesis.....	151
4.4 Topological and kinematic selections for each channel and category as described in the text. ^(*) Applies in the case of only two central jets. ^(†) A higher threshold (80 GeV) is used for the single-electron channel. ^(††) Applied only for $m_{Wh} > 320$ GeV. ^(‡‡) Only the two leading VR track-jets matched by ghost-association to the large- R jet are considered when classifying events into b -tag categories. Events are further classified according to the number of b -tagged jets.	156
4.5 Relative MC generator modeling systematic uncertainties (at the pre-fit stage), connected to the normalization, cross-region extrapolation, and shape of the W/Z +jets background processes included in the fits described in the text. An ‘S’ indicates that a shape variation is included for the sources listed, ‘ $A \leftrightarrow B$ ’ indicates a transfer-factor uncertainty between the regions A and B , and ‘norm.’ is the sum of cross-section and acceptance variations. A value of ‘float’ indicates that the parameter is not constrained in the fit. A range of values means the value depends on the lepton channel. Parentheses indicate when the uncertainty applies only to a given fit or a given region.	169

4.6	Relative MC generator modeling systematic uncertainties (at the pre-fit stage), connected to the normalization, cross-region extrapolation, and shape of the signal, top, SM Vh , and SM diboson background processes included in the fits described in the text. An ‘S’ indicates that a shape variation is included for the sources listed, ‘ $A \leftrightarrow B$ ’ indicates a transfer-factor uncertainty between the regions A and B , and ‘norm.’ is the sum of cross section and acceptance variations. A value of ‘float’ indicates that the parameter is not constrained in the fit. A range of values means the value depends on the lepton channel. Parentheses indicate when the uncertainty applies only to a given fit or a given region.	170
4.7	Post-fit normalization scale factors (NF) and their uncertainties obtained from a combined background-only fit of all signal and control regions of the 0- and 1-lepton channels. Numbers are presented for the background components that are allowed to float in the likelihood fit.	181
A.1	Train and test group for the W' signal morphing quality study	202
B.1	Multijet fractions from template fit in different control regions in electron and muon channels for different datataking periods.	209

LIST OF FIGURES

Figure	Page
1.1 The full particle content of the Standard Model. Taken from [21]	7
1.2 The Higgs potential after spontaneous symmetry breaking. Taken from [27]	17
1.3 Sketch of QCD processes from a proton-proton collision. The red point in the center represents the hard interaction, representing the largest momentum transfer. The initial and final state radiation (ISR/FSR) of gluons and $q\bar{q}$ production are shown in blue and red, respectively. Partons then fragment into a parton shower (PS) and hadronize into mesons and baryons, shown in green. The underlying event (UE) from interactions between non-colliding partons, producing further activity, is shown in purple. Soft photon radiation during hadron decays is shown in yellow. Taken from Ref [30].	26
1.4 Parton distribution functions for a reference momentum transfer of $Q^2 = 10 \text{ GeV}^2$ (left) and $Q^2 = 10^4 \text{ GeV}^2$ (right) from the NNPDF3.1NNLO PDF set. Different flavors of constituent partons are shown in different colors. Taken from Ref [31].	29
2.1 View of the LHC and surrounding environment under the Swiss-French border and greater Geneva area. Taken from [37]	33
2.2 The complete CERN accelerator complex, including non-LHC experiments. Taken from [45].	35
2.3 Layout of the LHC with the eight experimental and service points labeled. Beam 1 is injected close to ALICE and circulates clockwise, while Beam 2 is injected close to LHCb and circulates counterclockwise. The two beams exchange positions between the interior and exterior rings at each experiment to ensure the same overall path length. Taken from [46]	36
2.4 Total integrated luminosity vs. time delivered to (green) and recorded by (yellow) ATLAS during Run 2 (a) and Run 3 (b). Total luminosity in the Run 2 GRL is shown in blue, while the GRLs are not yet finalized for Run 3 as datataking is in progress at the time of this writing. From [47] and [48].	38
2.5 Luminosity-weighted distributions of the mean number of interactions per bunch crossing for Run 2 (a) and Run 3 (b). Average values for each individual datataking year and summed totals are shown. From [47] and [48].	39
2.6 Planned run schedule for the LHC and HL-LHC, updated in 2022.	41
2.7 Schematic of ATLAS and its subdetectors with a cut-away view. Persons are shown for scale. Taken from [34].	42

2.8	Diagram of the ATLAS coordinate system with an example particle of momentum $\vec{p}$. Modified from [49].	43
2.9	Representational layout of the ATLAS magnet system. Taken from [50].	45
2.10	Cutaway render of the ATLAS Inner Detector with an example high- p_T particle trajectory. Taken from [51].	46
2.11	Schematic of the ATLAS Inner Detector in the r - z plane. The lower left shows an enlargement of the pixel detector, and the lower right shows the radial coverage of the three detector subsystems. Taken from [52].	47
2.12	Diagram of an ATLAS pixel module, showing the silicon pixel sensors alongside the readout and support electronics. Taken from [34].	48
2.13	Diagram of an SCT barrel module. The stereo angle offset between the two silicon microstrip layers is visible alongside the readout and support electronics. Taken from [34].	49
2.14	Generalized diagram of TRT drift tube operation, including transition radiation detection. Taken from [54].	50
2.15	Cutaway render of the ATLAS calorimeter system showing the various subdetectors. Taken from [34].	51
2.16	Technical layout of the ATLAS calorimeters with color-coded subsystems. The inner detector systems are also shown in black-and-white in the center, up to $\eta = 2.5$. Taken from [55].	52
2.17	Diagram of a LAr barrel module showing the accordion geometry of the absorber plates and readout panels. It also shows the granularity of cells in η and ϕ across each of the three layers and trigger towers. Taken from [34].	53
2.18	Diagram of a tile calorimeter module, showing the active and passive layers and the readout mechanism with wavelength-shifting fibers and photomultiplier tubes. “Source tubes” are used in the calibration system. Taken from [34].	54
2.19	Cutaway render of the ATLAS Muon Spectrometer (MS) showing the subsystems and toroid magnets. Taken from [34].	56
2.20	A cross-section schematic of a quadrant of the ATLAS Muon Spectrometer. RPCs, CSCs, and TGCs are labeled, while MDTs are shown with a 3-letter encoding based on whether they are a barrel or endcap (B or E) module, an inner, middle, outer, or “extra” module (I,M,O,E), with the last letter denoting that these are large-sector modules (L). Small-sector modules are not shown in the cross-section as they are shifted azimuthally by $2\pi/6$. Taken from [34].	57
2.21	Diagram of the mechanical structure of an MDT chamber. Each chamber consists of an aluminum space frame formed by three spacer bars linked by longitudinal beams, which support two multi-layers comprising three or four layers of drift tubes. Four optical alignment rays—two parallel and two diagonal—are used to monitor the internal geometry of the chamber. The acronyms RO and HV indicate the positions of the readout electronics and high voltage supplies, respectively. Taken from [34].	58

2.22	Diagram of a CSC, shown in both a view along the anode wires (left) and a view along the cathode strips (right). Taken from [34].	59
2.23	Cross-sectional view of the upper part of the barrel, highlighting the RPCs in color. In the middle chamber layer, RPC1 and RPC2 are positioned below and above their corresponding MDT partners, respectively. In the outer layer, RPC3 is located above the MDT in the large sectors and below the MDT in the small sectors. All measurements are given in millimeters. Taken from [34].	60
2.24	Flowchart of the ATLAS TDAQ system during Run 2, illustrating the components involved in triggering, along with detector read-out and data flow. The Fast Tracker was ultimately canceled and not used in the final Run 2 configuration. Taken from [60].	62
2.25	Trigger efficiency as a function of transverse energy of all combined single-electron triggers during the different years of Run 2. Efficiency is given with respect to the offline Tight electron identification working point and the FCTight isolation working point. The ratios of data to MC simulation efficiencies are also shown. Taken from [65].	63
2.26	Illustration of particle signatures as they traverse the ATLAS detector subsystems. Taken from [66].	64
2.27	Flowchart of the stages of ATLAS track reconstruction, showing the inside-out primary pass (top) and the outside-in secondary backtracking pass (bottom).	65
2.28	Diagram of the local perigee representation used to parameterize ATLAS ID tracks. Taken from [67].	67
2.29	Illustration of (a) single-particle pixel clusters on a pixel sensor and (b) a merged pixel cluster due to highly collimated charged particles. The different colors represent energy deposits from different charged particles, whose trajectories are shown as arrows. Taken from [51].	68
2.30	Illustration of the seeding (a) and road search (b) processes in track reconstruction using a cutaway of the pixel and SCT detectors. Triplets of space points that form seeds are outlined in orange, and the search roads formed for each seed are outlined in red.	69
2.31	Illustration of the combinatorial Kalman filter (CKF) algorithm in the ID track finding process. Seeds, in orange, are extended recursively in a branching, combinatorial fashion as the filter iteratively searches for measurements and updates the track state parameters.	70
2.32	Illustration of the ambiguity-solving algorithm (represented by the squirrel) used in ID-track reconstruction. The overlapping track candidates outlined in red have multiple shared hits.	71
2.33	Flowchart of tracks processed by the ambiguity solver. Taken from [51].	72
2.34	Illustration of the global χ^2 fitter used after the ambiguity resolution to finalize precision parameter measurements of ID tracks.	73

2.35	Illustration of the TRT extension step of track reconstruction, where the new TRT extension attached to the existing track (purple) is shown in orange.	74
2.36	Depiction of the different kinds of TRT hits. Precision hits have a residual within 2.5 times the drift radius uncertainty. A tube hit has a radius outside this threshold or no valid signal leading edge. An outlier is a hit on track where the fitted track passes 100 μm or more outside the straw wall. Holes are straws crossed by a fitted track but do not have a hit. Taken from [70].	75
2.37	Illustration of the backtracking reconstruction pass, with the two space point seeds and TRT segments outlined in orange, search roads in red, and a final fitted track in purple.	76
2.38	Example primary, LRT, and combined track reconstruction efficiencies for displaced charged particles produced by the decay of the LLP scalar a in a Higgs portal signal model. Efficiencies are shown as a function of (a) the generated transverse impact parameter $ d_0^{\text{truth}} $ and (b) the generated decay radius $R_{\text{prod}}^{\text{truth}}$. Taken from [71].	77
2.39	Flowchart of the IVF logic. Taken from [74].	79
2.40	Stages of TopoCluster formation in an FCAL module for a simulated dijet event. In (a), seed cells with signal significance $\zeta_{\text{cell}}^{EM} > 4$ are shown, then in (b) cells with $\zeta_{\text{cell}}^{EM} > 2$ controlling the TopoCluster growth, and then in (c) all clustered cells with the final TopoClusters outlined in black. All clusters shown in (c) that do not contain a seed cell from this module are seeded in adjacent modules, not shown. Cluster boundaries are displayed on a dimensionless grid using θ and ϕ coordinates, preserving the general calorimeter cell shapes and proportions in the projection. Taken from [75].	80
2.41	Simplified overview of the primary outside-in combined tracking chain and secondary inside-out chain used in ATLAS muon track reconstruction.	81
2.42	A sample parton-level event clustered with the anti- k_t algorithm with $R = 1$, illustrating the catchment areas of the resulting hard jets. Taken from [79].	87
2.43	Flowchart of the jet energy scale calibrations. Taken from [55]	88
2.44	A diagram illustrating the creation of seven different TCC objects. A simple track-cluster match is shown in (1), (2) shows a cluster without a matched track, (3) shows a track without a matching cluster, (4) and (5) each show tracks matching a single cluster but sharing that cluster's energy, and (6) and (7) show a complex scenario with multiple different track-cluster matches. Taken from [81]	90
2.45	Flowchart of large-R jet calibrations. TCC jets use TrackCaloClusters instead of LC topo-clusters (LCW) as depicted but follow the same procedure. Taken from [82]	91
2.46	Illustration of the jet trimming procedure. Taken from [83]	92
2.47	Illustration of fixed-radius jets (left) versus variable-radius jets (right) inside a matched large-radius jet. The secondary and tertiary vertices from b -quark decays, used for b -tagging, are shown. Taken from [87]	93

2.48	Output score of the MV2c10 b -tagging algorithm for b -jets (blue), c -jets (green), and light-flavor jets (red). Taken from [88]	95
2.49	Schematic illustration of the path of an electron through ATLAS. The red arrow indicates the path of the electron, including a secondary photon (dashed) produced by material interactions. Taken from [93].....	96
2.50	: The electron identification efficiency in $Z \rightarrow ee$ events in data as a function of E_T for the Loose, Medium, and Tight operating points. The error bars represent both statistical and systematic uncertainties. Taken from [94].	99
3.1	Projected evolution of ATLAS CPU and disk usage from 2020 until 2036, under the conservative (blue) and aggressive (red) R&D scenarios. The black lines indicate the impact of sustained year-on-year budget increases and improvements in new hardware, which together amount to a capacity increase of 10% (lower line) and 20% (upper line). The shaded bands indicate datataking periods. Taken from [99].	102
3.2	Fractions of the total CPU consumption of the full ATLAS Run 2 reconstruction for one data run at $\langle\mu\rangle = 50$	103
3.3	Illustration of the WLCG Tier-0, -1, and -2 sites. Additional Tier-2 sites have been added since this image was made in 2013, and not all Tier-2 sites are available to ATLAS. Taken from [101]	105
3.4	Evolution of the total ATLAS data volume since before Run 1. Taken from [100]	107
3.5	Flowchart of the ATLAS software workflow. Processing steps are shown in blue, while output formats are shown in red. Taken from [100].....	108
3.6	Flowchart illustrating the flow of ATLAS data from initial reconstruction during the calibration loop followed by the bulk processings at the Tier-0 and Tier-1 sites. Taken from [100].....	110
3.7	Flowchart of the structure of an Athena application. The solid lines on the left indicate ownership, while those on the right indicate the flow of data. The dotted line down the center indicates the control flow from Algorithm to Algorithm. The dashed lines indicate a non-owning reference. Taken from [100].	112
3.8	Flowchart of the static and dynamic stores used for xAOD objects. This figure shows a container of electrons stored in the event data store named El es , which is also used as the name of the ROOT TBranch when the objects are saved to ROOT. The associated auxiliary store contains the data members pt , eta , and so on. The dynamic auxiliary store on the right contains extra members attached (decorated) to the container. Solid lines with open arrows indicate inheritance, plain solid lines show references between the objects, and dotted lines map between names and objects in the event store. Taken from [100].	114
3.9	Snapshot of the SPOT monitoring webpage on May 29, 2024, highlighting the CPU and memory consumption of the primary reconstruction chain for both the <i>main</i> and <i>Release 24</i> Nightly release branches.	116
3.10	Example snapshot of the SPOTPlots monitoring page.	117

3.11	Example SPOTPlots distribution showing elapsed CPU vs time for reconstructing a set of 2017 data events, in both absolute time in ms (top) and a ratio with respect to the starting value pre-optimization campaign (bottom). Significant merge requests are labeled via the colored dashed lines with keys to the right of the figure.	118
3.12	Incremental improvement in execution speed for ID track reconstruction in a set of $\langle\mu\rangle = 50$ collision data events as improvements are added to the reconstruction software for Run 3. The blue shaded area indicates the time taken for the track reconstruction relative to the initial Run 3 software implementation. The purple area indicates the time added by the additional LRT. Due to changes required to ensure thread safety, the unoptimized initial performance of the Run 3 software, referred to as “Baseline”, was approximately 20% slower than the previous software used during Run 2.	121
3.13	Illustration of the implementation of confirmation space points in ID track seed formation. The initial seeds comprised of three space points are outlined in orange, whereas their associated confirmation space points are outlined in green. The confirmation space points are additional space points in an adjacent layer compatible with the original seed trajectory.	122
3.14	Flowchart of the AMVF logic. Taken from [74].	126
3.15	Distributions of the longitudinal separation between pairs of reconstructed primary vertices (a) and hard-scatter reconstruction efficiency as a function of local pileup density (b) for simulated $t\bar{t}$ events. Taken from [74].	127
3.16	Flowchart illustrating the Run 2 (top) and improved Run 3 (bottom) event reconstruction workflows for the main physics and the LLP streams. Taken from [71].	128
3.17	A comparison of the radial distributions of reconstructed secondary vertices in a simulated LLP sample using the legacy and updated track reconstruction configurations. The circular markers represent reconstructed vertices that are matched to truth-level LLP decay vertices (LLP). The dashed lines represent reconstructed vertices that are not matched to truth-level LLP decay vertices (non-LLP). Vertical lines indicate the radial position of the pixel and SCT detector layers. The lower panel shows the ratio of the number of vertices reconstructed in the updated reconstruction to those reconstructed in the legacy implementation for true LLP vertices (circular markers) and non-LLP vertices (dashed lines). The uncertainties shown are statistical only. Figure and caption taken from [71].	129
3.18	Relative computational performance improvement of key reconstruction steps obtained in the updated reconstruction software compared to that of the legacy implementation when reconstructing a set of LHC Run 2 data events, as a function of $\langle\mu\rangle$, for ID track reconstruction. The “Miscellaneous” category includes event data model handling, vertex reconstruction, reading of detector condition information, and cluster formation in the Pixel and SCT subdetectors.	130

3.19	Breakdown of the average CPU time required per event for ID track and vertex reconstruction when reconstructing a set of LHC Run 2 data events, as a function of $\langle\mu\rangle$, comparing the legacy (a) and updated (b) configurations. The bottom panel shows the fraction of the total track reconstruction time taken by each component. The "Miscellaneous" category includes event data model handling, vertex reconstruction, reading of detector condition information, and cluster formation in the Pixel and SCT subdetectors.	131
3.20	Processing time taken per event versus average pileup to reconstruct ID tracks and vertices in the same events for the updated (purple) and legacy (green) reconstruction software. The updated reconstruction is also evaluated when including the impact of the additional LRT step (blue). The bottom panel depicts the time taken relative to the legacy reconstruction time. The shaded area indicates data events from a 2018 machine development run that did not pass the full ATLAS data quality requirements.	132
3.21	Number of seed tracks processed divided by the final number of tracks created in the silicon-seeded tracking as a function of $\langle\mu\rangle$ (a) and average number of tracks reconstructed per event as a function of the value μ_{actual} , with a linear fit to the range $10 \leq \mu_{\text{actual}} \leq 30$ superimposed as a dashed line (b).....	133
3.22	Event size of the inner detector reconstruction output in the ATLAS event data format for the same set of reconstructed data events as a function of the average pileup, comparing the legacy and updated releases. The shaded area indicates data events from a 2018 machine development run that did not pass the full ATLAS data quality requirements.....	133
3.23	A comparison of primary vertex reconstruction processing time taken per event versus average pileup for the Athena version of the AMVF algorithm and the ACTS-provided implementation. The shaded area indicates data events from a 2018 machine development run that did not pass the full ATLAS data quality requirements.	134
3.24	Fractions of the total CPU consumption of the full ATLAS Run 3 reconstruction for one data run at $\langle\mu\rangle = 50$. In the updated reconstruction, several algorithms, such as jet reconstruction and hadronic tau decay candidate identification, are no longer included in the primary ATLAS reconstruction; instead, they are created downstream and, hence, no longer enter the sum.	135
3.25	Tracking efficiency as a function of p_T in simulated $t\bar{t}$ events with a flat $\langle\mu\rangle$ distribution between 0 and 80, in a comparison of the updated and legacy reconstructions. The error bars indicate statistical uncertainties.	136
3.26	Tracking efficiency for simulated single muons using the Run 2 detector geometry with $p_T = 10$ GeV as a function of η , in a comparison of the updated and legacy reconstructions. The error bars indicate statistical uncertainties.	137
4.1	Representative lowest-order Feynman diagram of the resonant production of a W boson and a Higgs boson via $q\bar{q}$ annihilation. The subsequent hadronic decay of the Higgs boson and leptonic decay of the W boson are also pictured.	138
4.2	Example tree-level Feynman diagrams for $t\bar{t}$ (a) and W/Z +jets (b) production, the two most dominant SM backgrounds for the Wh resonance search.....	143

4.3	Illustration of the resolved and merged $H \rightarrow b\bar{b}$ jet topologies.	150
4.4	Simplified flowchart of the one-lepton analysis selection into the eight one-lepton signal and control regions.	153
4.5	Efficiency of signal and main background components after successively applying analysis selections in the 1-lepton resolved (a) and merged (b) categories for selected W' masses of 700 GeV (resolved) and 3 TeV (merged). <i>AllCxAOD</i> refers to the initial data type used after the object definitions but before any analysis selections, and MET refers to E_T^{miss} . In the resolved, <i>mbbRestrict</i> corresponds to a requirement of $50 < m_{jj} < 200$ GeV, while <i>mbbCorr</i> refers to the mass window selection requirement in both categories. <i>VROR</i> refers to the overlap removal for overlapping variable-radius track jets.	154
4.6	Product of acceptance and efficiency for $W' \rightarrow Wh \rightarrow \ell^\pm \nu b\bar{b}/c\bar{c}$ as a function of the resonance mass for the 1-lepton (a) and 0-lepton (b) signal regions.	155
4.7	Pre-fit m_{Wh} distributions of the JET_CR_JET_JER_EffectiveNP1 (top) and 2 (bottom) uncertainties for 1 b -tag (left) and 2 b -tag (right) in the m_{jj} sideband control regions of the 1-lepton resolved regime. Normalized entries are shown in the primary panel and the relative variation is shown in the ratio panel, below. The x-axis is expressed as the fraction of the maximum tested m_{Wh} value, 5000 GeV.....	159
4.8	Pre-fit distributions of $m_{T,Wh}$ for $t\bar{t}$ (a) and W +heavy flavor jets (b) backgrounds in the 1-lepton, 1 b -tag merged signal region. The lower panel shows the relative variations with respect to the nominal induced by the resolution of the large- R jet after averaging and smoothing.	160
4.9	Comparisons of pre-fit m_{Wh} -distributions of $t\bar{t}$ inside the combined 1-lepton resolved signal regions as predicted by the nominal generation versus the alternative generations as described in the text, labeled in each ratio panel. Variations are normalized with respect to the nominal to study pure shape uncertainties and are shown after Gaussian kernel smoothing. For legibility, only 5 of the 100 NNPDF variations are shown.....	165
4.10	Comparisons of pre-fit m_{Wh} -distributions of $t\bar{t}$ inside the combined 1-lepton merged signal regions as predicted by the nominal generation versus the alternative generations as described in the text, labeled in each ratio panel. Variations are normalized with respect to the nominal to study pure shape uncertainties and are shown after Gaussian kernel smoothing. For legibility, only 5 of the 100 NNPDF variations are shown.....	166
4.11	Event distributions of m_{Wh} for the 1-lepton SR in the resolved (a, b) and merged (c, d) categories of the W' fit. The quantity on the vertical axis is the number of data events divided by the bin width in GeV. In each plot, the last bin contains the overflow. The background prediction is shown after a background-only maximum-likelihood fit to the data. The lower panels show the ratio of the observed to the estimated SM background. The background uncertainty band shows the post-fit statistical and systematic uncertainties added in quadrature. The signal for the benchmark HVT <i>Model A</i> with $m_{W'} = 1.4$ TeV is shown as a dashed line and normalized to a cross section times branching fraction of 0.1 pb.	180

4.12	Event distributions of m_{T,W_h} for the 0-lepton SR in the resolved (a, b) and merged (c, d) categories of the W' fit. The quantity on the vertical axis is the number of data events divided by the bin width in GeV. In each plot, the last bin contains the overflow. The background prediction is shown after a background-only maximum-likelihood fit to the data. The lower panels show the ratio of the observed to the estimated SM background. The background uncertainty band shows the post-fit statistical and systematic uncertainties added in quadrature. The signal for the benchmark HVT <i>Model A</i> with $m_{W'} = 1.4$ TeV is shown as a dashed line and normalized to a cross section times branching fraction of 0.1 pb.	182
4.13	Distributions of the m_{W_h} and m_{T,W_h} observables in the control regions of the resolved (a) and merged (b) event categories. The background prediction is shown after a background-only maximum-likelihood fit to the data. The lower panels show the ratio of the observed to the estimated SM background. The background uncertainty band shows the post-fit statistical and systematic uncertainties added in quadrature.	183
4.14	Comparisons of the observed data and the expected background event yields in the signal regions of the W' fit. The background predictions are obtained through a background-only simultaneous fit to the signal and control regions. The middle panel shows the ratios of the observed data to the post-fit background predictions. The bottom panel shows the ratios of the post-fit and pre-fit background predictions. The hatched bands represent uncertainties in the total background predictions, combining statistical and systematic contributions.	184
4.15	Nuisance parameter pulls after a conditional $\mu = 0$ fit to data in the signal and control regions of the W' combination. The blue, black, and red markers indicate the pull in the 0-lepton, combined, and 1-lepton fits, respectively. The first number shows the pull value, while the number after $\pm$ shows the ratio of pre-fit to post-fit constraint.	186
4.16	Nuisance parameter pulls after a conditional $\mu = 0$ fit to data in the signal and control regions of the W' combination. The blue, black, and red markers indicate the pull in the 0-lepton, combined, and 1-lepton fits, respectively. The first number shows the pull value, while the number after $\pm$ shows the ratio of pre-fit to post-fit constraint.	187
4.17	Ranking plots for the SR fit to observed data for $m_{W'} = 700$ GeV. Positive correlations between nuisance parameters and the parameter of interest are indicated by the blue bars, while negative ones are shown in green.	189
4.18	Ranking plots for the SR fit to observed data for $m_{W'} = 2000$ GeV. Positive correlations between nuisance parameters and the parameter of interest are indicated by the blue bars, while negative ones are shown in green.	190
4.19	Correlation matrix of the W' combination fit in the signal regions. Only NPs where the absolute value of at least one correlation is above 0.5 are shown for legibility.	191
4.20	Upper limits at the 95% CL on the product of the cross section for $pp \rightarrow W'$ and the branching fraction to Wh from the combination of the 0-lepton and 1-lepton channels. Limits from the 0- and 1-lepton channels individually are also shown for comparison.	192

4.21	Observed local p_0 values as a function of the tested W' mass point. The red curve indicates the standalone 0-lepton fit, the blue curve the standalone 1-lepton fit, and the black curve the 0- and 1-lepton combined fit.	192
4.22	Observed limits on the HVT model at 95% CL in the g_F versus g_H plane for W' masses of 2,3, and 4 TeV. The circles indicate the coupling values for Models A and B, and the grey region corresponds to the area of phase space where the decay width of the resonance is no longer negligible and the signal m_{Wh} shape is no longer expected to be dominated by the experimental resolution.	193
4.23	Upper limits at the 95% CL on the product of the cross section for $pp \rightarrow W'$ and the branching fraction to Wh from the combination of the 0-lepton and 1-lepton channels. Limits from the 36.1 fb^{-1} analysis are shown for comparison, both in their original form and scaled to the full Run-2 luminosity, for comparison.	194
4.24	Event display of an $X \rightarrow Wh \rightarrow \mu\nu b\bar{b}$ candidate in the merged 1 b-tag signal region. The W boson candidate is comprised of a muon and missing transverse energy. The muon is indicated by the red solid line and has a transverse momentum of $p_T = 632 \text{ GeV}$. The missing transverse energy has a value of 362 GeV and is indicated by the dashed white line. The Higgs-boson candidate is comprised of a large-R jet which contains one b-tagged VR track-jet. The large-R jet is indicated by the turquoise cone and has a transverse momentum of $p_T = 1.09 \text{ TeV}$. The yellow blocks correspond to energy depositions in the hadronic calorimeter, and the blue boxes correspond to muon chambers with hits from the muon. The invariant mass of the Wh candidate is $m_{Wh} = 3.7 \text{ TeV}$. The event was recorded in 2018 (Run 363710, Event 2531279786).	196
A.1	Mass residuals for $m_{WH} = 500 \text{ GeV}$ and $m_{WH} = 2400 \text{ GeV}$ in the resolved category (a,b) and for $m_{WH} = 1400 \text{ GeV}$ and $m_{WH} = 3000 \text{ GeV}$ for the merged category (c,d). Overlaid is the fitted Bukin distribution, and the fitted width of the distribution, σ , is labeled.	198
A.2	Mass resolutions per W' signal point for the merged and resolved categories in the 1 b-tag, 1L channel. The resolved category has a better mass resolution in the lower mass regime and loses resolution in the higher mass regime, and vice versa for the merged category.	199
A.3	Distributions of the local p -value as a function of the W' mass. The p -value distributions correspond to a fit to Asimov data injected with a hypothetical signal with a mass of (a) $m = 700 \text{ GeV}$, (b) 1400 GeV , and (c) 2200 GeV	200
A.4	Comparison of simulated (black) and morphed (blue) samples in the 1-lepton resolved channel for the HVT W' signal with masses of 0.7, 0.9, and 1.2 TeV. The selected region is indicated in the text of the individual figures.	203
A.5	Comparison of simulated (black) and morphed (blue) samples in the 1-lepton merged channel for the HVT W' signal with masses of 1.6, 2.4, and 3.5 TeV. The selected region is indicated in the text of the individual figures.	204
B.1	m_T^W distributions in multijet control region in the electron channel (a) and the muon channel (b) for the 1-lepton 1 b-tag resolved regions.	208
B.2	Postfit distribution of m_T^W for the 1-lepton, 1 b-tag resolved region.	209

B.3	Pre-fit distribution (a) of the $p_{T,jj}$ observable presented in the 0-tag control region with data-to-simulation ratio and shaded statistical uncertainty, as well as the piecewise correction function derived from this distribution (b).	210
B.4	Pre-fit distribution of the m_{Wh} (top row) and p_T^{miss} (bottom row) observables presented separately in the 1- (left column) and 2-tag (right column) CRs. The observables are presented after the $p_{T,jj}$ corrections have been applied.	212
B.5	Fits through the data-to-simulation ratios obtained from studying the p_T^{miss} observables within the CRs of the resolved 1- and 2-tag categories.	213
B.6	Pre-fit distribution of the large- R jet p_T observable presented in the 1-tag (left) and 2-tag (right) CRs (top row), as well as the piecewise correction functions derived from these distributions (bottom row).	215
B.7	The relative uncertainties for the p_T^W slope in the resolved regime for 1-tag (left) and 2-tag (right) in the CRs.	217
B.8	The relative difference for the large- R jet p_T slope in the resolved regime for 1-tag (left) and 2-tag (right) in the merged CRs.	217
B.9	The difference between the unrenormalized prefit distributions of stop DS/DR in the 1 b -tagged (left) and 2 b -tagged (right) resolved signal regions of the 1-lepton channel.	219
B.10	The difference between the unrenormalized prefit distributions of stop DS/DR in the 1 b -tagged (left) and 2 b -tagged (right) merged signal regions of the 1-lepton channel.	219
B.11	The prefit m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DR scheme) backgrounds compared to the m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DS scheme) backgrounds presented separately for the 1 b -tagged (left) and 2 b -tagged (right) resolved signal regions of the 1-lepton channel.	220
B.12	The prefit m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DR scheme) backgrounds compared to the m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DS scheme) backgrounds presented separately for the 1 b -tagged (left) and 2 b -tagged (right) merged signal regions of the 1-lepton channel.	220
B.13	Comparison between the nominal m_{Wh} distribution and the up and down variations of the <i>truthWZless than 2</i> (top row) and <i>truthWZmore than 3</i> (bottom row) uncertainties presented separately for the W +jets background within the 1- and 2-tag signal region.	221

INTRODUCTION

The Standard Model (SM) of particle physics is the most complete theory of nature to date, describing the fundamental interactions between all known elementary constituents of matter. The discovery of a new particle consistent with the Higgs boson by the ATLAS [1] and CMS [2] experiments in 2012 marked a monumental achievement for particle physics, confirming the final “missing” component predicted by the Standard Model. Subsequent precision measurements by the experiments so far indicate that this new particle is wholly consistent with the properties predicted for the Higgs boson as described by the Standard Model. However, it is known that the Standard Model is not a fully complete description of physical reality. Many unanswered questions remain, such as the nature of dark matter and dark energy, the matter-antimatter asymmetry of the universe, and the so-called “naturalness problem” where the Higgs mass is protected from large radiative corrections by an unexplained mechanism [3, 4, 5].

In the time since the Higgs discovery, a large-scale search program has emerged to probe whether the Higgs is embedded in a, so far unobserved, beyond the Standard Model (BSM) “extended Higgs” sector. Many theories addressing the naturalness problem assume new strong interactions at a high energy scale to dynamically break the electroweak symmetry. Such models, for example, Minimal Walking Technicolor [6, 7, 8], Little Higgs [9], or Composite Higgs [10, 11], generically predict new vector resonances with significant branching ratios for decays into a vector boson and a Higgs boson. These new heavy resonances have masses at the several-hundred GeV to TeV scale and thus would be accessible at the center-of-mass energies available in the LHC. However, as resonance searches are typically not sensitive to all free parameters of a specific underlying theory, and reinterpretability through different models is broadly desirable, simplified phenomenological models such as the Heavy Vector Triplet [12, 13] (HVT) model are used. This model predicts BSM heavy neutral vector bosons W' and Z' and adds an additional $SU(2)_L$ iso-triplet of resonances to the Standard Model with a restricted number of additional couplings to the new field.

This work describes a search for these BSM heavy resonances decaying into a vector boson and a Higgs boson in a “semileptonic” signature, where the Higgs boson decays into a pair of b -quarks and the vector boson decays into a pair of charged leptons and/or neutrinos, and discusses specifically the $W' \rightarrow Wh$ signature in the “1-lepton” (1L) channel with $\ell^\pm \nu b \bar{b}$ final states, of which the author

was the primary analyzer. This is situated as part of a larger comprehensive VH semileptonic search also including the “0-lepton” (0L) $\nu\bar{\nu}b\bar{b}$ and “2-lepton” (2L) $\ell^+\ell^-b\bar{b}$ final states with combined fits to both W' and Z' among other BSM scenarios. The search uses the 139 fb^{-1} pp collision dataset collected with the ATLAS detector during LHC Run-2 (between 2015 and 2018) at a center-of-mass energy of $\sqrt{s} = 13\text{ TeV}$. This combined effort led to the paper “*Search for heavy resonances decaying into a Z or W boson and a Higgs boson in final states with leptons and b -jets in 139 fb^{-1} of pp collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector.*” published in the Journal of High Energy Physics (JHEP) on June 5, 2023 [14].

Semileptonic final states are among the most sensitive probes of new resonances, as they have much larger branching fractions than fully leptonic final states while maintaining a cleaner detector signature than fully hadronic final states. Particularly, the $h \rightarrow b\bar{b}$ decay is targeted as it has the highest branching ratio of the Higgs decay modes, and the leptonic W decay provides efficient objects on which to trigger. The search is performed by examining the invariant (in the 1-lepton case) and transverse (in the 0-lepton case, used in the combined W' fit to extend sensitivity to hadronic τ lepton decays) mass distributions of the $\ell^\pm\nu b\bar{b}$ system for a localized excess, probing an extensive range of W' masses from 400 GeV to 5 TeV. The subsequent broad range of Higgs boson transverse momenta necessitates two methods to reconstruct the Higgs candidates from the hadronic jets produced by the b -quark hadronization process. The Higgs decay products are sufficiently separated at low transverse momenta to be resolved into two individual jets. In contrast, at high transverse momenta, they are reconstructed as merged into a single large-radius jet.

The results are interpreted in two HVT benchmark models, where the W' is produced via a Drell-Yan process. The first model, *Model A*, has comparable branching fractions to both fermions and gauge bosons, as in some models with an extended gauge symmetry [15]. The second, *Model B* [16], has fermionic couplings suppressed as in strong dynamical models such as minimal composite Higgs. Events are selected and categorized into signal regions (SR) and background-dominated control regions (CR) based on event-level quantities and the reconstructed final-state particles’ kinematic properties. The statistical interpretation is then performed with a binned-maximum likelihood fit on the invariant mass distributions of the reconstructed Wh system, m_{Wh} , using the combined signal and control regions, where the major backgrounds and signal processes are modeled using Monte Carlo (MC) simulations and compared to the observed data. This sets upper limits on the HVT W' production cross section times branching ratio, which are then converted into constraints on the coupling-strength scale factors to fermions and gauge bosons, g_f and g_h .

Reconstruction of the trajectories of charged particles through the innermost system of ATLAS, the Inner Detector (ID), into tracks is a critical foundational component of both this search and the broader ATLAS physics program. The majority of physics analysis objects, from leptons to jets, incorporate ID tracks. This is accomplished by an ordered series of algorithms to “connect the dots” between raw hits in the ID subdetectors and form high-quality tracks. These are then used to reconstruct the primary and secondary vertices of collision events and form the basis of other physics objects. This process represents a significant computational challenge as track reconstruction is fundamentally a combinatorial problem, with the number of possible track combinations scaling exponentially with the number of hits. As the ID is a complex, high-granularity system closest to the primary interaction point, a single inelastic pp collision leads to up to 200 individual measurements recorded in the silicon detectors. In an average ATLAS bunch crossing event with up to 70 simultaneous pp interactions, or “pile-up” ($\langle\mu\rangle$), this implies up to 14,000 silicon hits that must be processed and combined into clusters, combined into short track “seeds”, and then extended into full tracks with trajectory parameters. As such, ID tracking was, by far, the most computationally intensive component of the ATLAS reconstruction workflow in LHC Run 2, taking nearly 70% of the total CPU time per event. Additionally, the quality of track candidates becomes challenging to maintain under high pile-up. Collections of unrelated clusters can be combined erroneously into “fake” tracks that do not correspond to any real charged particle, and high cluster densities can lead to ambiguities in the assignment of clusters to tracks, potentially degrading the reconstructed track properties relative to their true values.

The peak instantaneous luminosity delivered to ATLAS during Run 2 was twice the original LHC design value for which the reconstruction software was initially optimized. The ongoing Run 3 aims to surpass this value and record up to 300 fb⁻¹ of collision data at $\sqrt{s} = 13.6$ TeV in the period 2022–2025.¹ Average pile-up during Run 2 ranged predominately from $\langle\mu\rangle = 20$ to $\langle\mu\rangle = 40$, while a large fraction of the Run 3 dataset is recorded at 50 or more pp interactions per bunch crossing, with peak values beyond 60 during 2022 data-taking. ATLAS computing resources are increasingly strained as larger volumes of data with more complex topologies must be processed and analyzed, potentially limiting the overall scope and capabilities of the physics program.

As track reconstruction is the most computationally intensive component of the ATLAS reconstruction chain, optimizations in this area have the largest potential impact on overall resource

¹Though, a series of unfortunate accidents and technical difficulties causing long losses of beamtime during the 2023 datataking period makes achieving this total value less likely as we get farther into Run 3.

consumption. Particularly, aborting the reconstruction of low-quality tracks as early as possible in the algorithm chain can yield cascading improvements as all succeeding algorithms have fewer downstream executions. Thus, a large-scale campaign to optimize the performance of the ATLAS track reconstruction algorithms with this overarching strategy was undertaken in advance of Run 3 by the ATLAS Tracking Combined Performance working group, for which the author developed the primary centralized monitoring framework and was a major contributor. This work led to the paper “*Software Performance of the ATLAS Track Reconstruction for LHC Run 3*” published in Computing and Software for Big Science (CSBS) on April 2, 2024 [17]. This paper also discusses the parallel optimization effort for the ATLAS muon reconstruction algorithms, due to their performance being fundamentally intertwined with the ID tracking optimization, but this remains largely outside this thesis’s scope.

This thesis is organized in several parts. Chapter 1 discusses the Standard Model and motivations for BSM heavy resonances. Chapter 2 discusses the experimental apparatus, giving an overview of the Large Hadron Collider, the ATLAS detector, and the reconstruction of physics objects used in the analysis. Chapter 3 then discusses the technical optimization of ATLAS track reconstruction algorithms for LHC Run-3 (2022-2025) in detail. Chapter 4 finally discusses the in-depth analysis strategy, statistical interpretation, and results of the analysis.

CHAPTER 1

THE STANDARD MODEL AND BEYOND

Throughout the mid-20th century, the discovery of numerous elementary particles that interact through the strong and electroweak nuclear forces paved the way for the development of the **Standard Model (SM)** of particle physics, a theoretical framework that elegantly describes all known particles and their interactions within a cohesive **quantum field theory (QFT)**. The collective effort of generations of physicists pushing collider experiments to progressively higher energy scales in search of an underlying theory of nature culminated in this landmark formulation.

The Standard Model describes the fundamental particle constituents of matter and three known forces that mediate interactions between them: the strong and weak nuclear forces and the electromagnetic force. The fourth and weakest fundamental force, gravity, remains absent from the Standard Model as the discovery of a quantum theory that successfully integrates it with the other three fundamental forces remains an unsolved problem in particle physics. Einstein’s theory of general relativity instead describes gravity as a consequence of the curvature of spacetime – while in contrast, Standard Model interactions are mediated by fundamental particles. These force-carrying particles are categorized as **bosons**, while the matter constituent particles are **fermions**. Each class of particles is characterized by their specific properties:

- **Fermions** have half-integer spin and obey Fermi-Dirac statistics, where no two identical fermions can occupy the same quantum state simultaneously as a consequence of the Pauli Exclusion Principle. Fermions are further categorized into **leptons** and **quarks**, each falling into three “**generations**” reflecting their various quantum numbers and masses. Charged leptons, with $q = 1$, include the electron (e^-), the muon (μ^-), and the tau (τ^-). Each charged lepton “**flavor**” has a corresponding neutral ($q = 0$) particle, the electron neutrino (ν_e), the muon neutrino (ν_μ), and the tau neutrino (ν_τ). Quarks also come in six flavors, where the “up-type” quarks carry an electric charge of $q = \frac{2}{3}$ and include the up (u), charm (c), and top (t) quarks, while the “down-type” quarks carry a charge of $q = -\frac{1}{3}$ and include the down (d), strange (s), and bottom (b) quarks. Each fermion also has a corresponding antilepton (e^+ , μ^+ , τ^+ , $\bar{\nu}_e$, $\bar{\nu}_\mu$, $\bar{\nu}_\tau$) or antiquark ($\bar{u}$, $\bar{c}$, $\bar{t}$, $\bar{d}$, $\bar{s}$, $\bar{b}$) with an inverted charge.

- **Bosons** have integer spin and obey Bose-Einstein statistics, where multiple identical bosons can occupy the same quantum state. The exchange of spin-1 vector bosons mediates the three fundamental forces described by the Standard Model. The electromagnetic interaction (Section 1.2) is mediated by the massless **photon** (γ), the strong nuclear force (Section 1.3) is mediated by the massless **gluon** (g), and the weak nuclear force (Section 1.4) is mediated by the massive $W^\pm$ **and** Z^0 **bosons**. The Standard Model also contains the spin-0 scalar **Higgs boson**, which is responsible for giving mass to other particles through the Higgs mechanism, discussed further in section 1.5.

The full particle content of the Standard Model is depicted in Figure 1.1. The interactions between these fundamental “building blocks” described by the Standard Model give rise to the universe’s vast complexity – the strong interaction binds quarks together to form hadrons, such as protons and neutrons, forming atomic nuclei. In turn, the electromagnetic force binds electrons to these nuclei, creating atoms, while the weak interaction mediates radioactive decay. In the following Chapter, we will discuss an abbreviated overview of the main concepts underpinning the Standard Model, as well as motivations for physics beyond the Standard Model alongside proton-proton collision phenomenology, to set the theoretical context for this work.¹

1.1 Standard Model Lagrangian: Symmetries and Groups

The Standard Model describes particles and their interactions in the language of QFT, where particles are described as excitations of a quantum field. Under quantum mechanics, the condition of a single particle is modeled by combinations of wave functions of discrete modes, and in QFT each mode is the excitation of an underlying quantum field. The kinematics of a field ϕ can be described through a **Lagrangian density** $\mathcal{L}$, whose equation of motion can be determined from the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = 0 \quad (1.1)$$

At its core, the Standard Model is a story of **symmetries**. A general renormalizable Lagrangian, $\mathcal{L}_{SM}$, can describe all fundamental particles and their interactions. The SM Lagrangian is invariant

¹It goes without saying that we cannot come anywhere close to fully developing the entire formalism of the Standard Model and pp collision phenomenology in a thesis introduction – for further reading, the author recommends Peskin and Schroeder’s *An Introduction to Quantum Field Theory* [18], Schwartz’s *Quantum Field Theory and the Standard Model* [19], and *Practical Collider Physics*[20], which this discussion has been largely drawn from.

Standard Model of Elementary Particles

three generations of matter (elementary fermions)						three generations of antimatter (elementary antifermions)						interactions / force carriers (elementary bosons)	
I		II		III		I		II		III			
mass charge spin	$\approx 2.2 \text{ MeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 1.28 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 173.1 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 2.2 \text{ MeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	$\approx 1.28 \text{ GeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	$\approx 173.1 \text{ GeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	$\approx 124.97 \text{ GeV}/c^2$ 0 0
QUARKS	u up	c charm	t top	$\bar{u}$ antiup	$\bar{c}$ anticharm	$\bar{t}$ antitop	g gluon					H higgs	
	d down	s strange	b bottom	$\bar{d}$ antidown	$\bar{s}$ antistrange	$\bar{b}$ antibottom	γ photon						
	$\approx 0.511 \text{ MeV}/c^2$ -1 $\frac{1}{2}$	$\approx 105.66 \text{ MeV}/c^2$ -1 $\frac{1}{2}$	$\approx 1.7768 \text{ GeV}/c^2$ -1 $\frac{1}{2}$	$\approx 0.511 \text{ MeV}/c^2$ 1 $\frac{1}{2}$	$\approx 105.66 \text{ MeV}/c^2$ 1 $\frac{1}{2}$	$\approx 1.7768 \text{ GeV}/c^2$ 1 $\frac{1}{2}$	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1
LEPTONS	e electron	μ muon	τ tau	e^+ positron	$\bar{\mu}$ antimuon	$\bar{\tau}$ antitau	Z^0 Z ⁰ boson					W ⁺ W ⁺ boson	
	$< 2.2 \text{ eV}/c^2$ 0 $\frac{1}{2}$	$< 0.17 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 18.2 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 2.2 \text{ eV}/c^2$ 0 $\frac{1}{2}$	$< 0.17 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 18.2 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$\approx 91.19 \text{ GeV}/c^2$ 1 1	$\approx 80.360 \text{ GeV}/c^2$ 1 1	$\approx 80.360 \text{ GeV}/c^2$ -1 1	$\approx 80.360 \text{ GeV}/c^2$ -1 1	$\approx 80.360 \text{ GeV}/c^2$ -1 1	$\approx 80.360 \text{ GeV}/c^2$ -1 1	$\approx 80.360 \text{ GeV}/c^2$ -1 1
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	$\bar{\nu}_e$ electron antineutrino	$\bar{\nu}_\mu$ muon antineutrino	$\bar{\nu}_\tau$ tau antineutrino	W ⁺ W ⁺ boson					W ⁻ W ⁻ boson	

Figure 1.1. The full particle content of the Standard Model. Taken from [21]

under the **spacetime symmetries** of the Poincare group ($\mathcal{P}$), meaning translations, Lorentz boosts, and rotations, as well as under internal local **gauge symmetries**.

Noether's theorem relates the global symmetries of a Lagrangian to conservation laws, where if the Lagrangian of a system is invariant under a continuous transformation

$$\phi(x) \rightarrow \phi'(x) = \phi(x) + \alpha \Delta \phi(x), \quad (1.2)$$

then the **Noether current**

$$j^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta \phi, \quad (1.3)$$

is a conserved ($\partial_\mu j^\mu = 0$) current, yielding a **conservation law** for each continuous symmetry of $\mathcal{L}$ and a corresponding quantity, the **Noether charge**, that is constant in time:

$$Q \equiv \int_{\text{all space}} j^0 d^3x. \quad (1.4)$$

This relationship can be used to derive all the familiar conservation laws and conserved quantities of the Standard Model, leading to the conservation of angular momentum, electric and color charges, weak isospin, etc.

The symmetries of the Standard Model are described by **group theory**, which determines how the various SM fields transform under different symmetry transformations. A group is defined as a set of elements $\{g_i\}$ equipped with a set of operations $(\times)$ that satisfies several axioms:

- **Closure:** The result of any group operation between two elements is also an element of the group, such that $g_i \times g_j = g_k$
- **Associativity:** The associative law holds for all elements in the group, such that $(g_i \times g_j) \times g_k = g_i \times (g_j \times g_k)$
- **Identity:** There exists an identity element such that $\mathbb{1} \times g_i = g_i$
- **Inverse:** Each element in the group has an inverse such that $g_i \times g_i^{-1} = \mathbb{1}$

The Standard Model symmetries are described by **Lie groups**, where a field is invariant under a transformation that defines the Lie algebra

$$U = e^{i\theta^a T^a} \mathbb{1}, \quad (1.5)$$

where θ^a are parameters and T^a are the **generators** of the Lie group which satisfy a commutation relation

$$[T^a, T^b] = i f^{abc} T^c, \quad (1.6)$$

where f^{abc} are the structure constants of the group. A group is **Abelian** if the $f^{abc} = 0$ and **non-Abelian** otherwise. Lie groups are labeled reflecting their algebraic properties and structure, where, for example, a group is **special (S)** if $|U| = 1$ and/or **unitary (U)** if they preserve complex inner products such that $U^\dagger U = \mathbb{1}$. **Orthogonal (O)** groups in n dimensions deal with rotations and reflections while preserving Euclidean or Lorentzian metrics.

The Standard Model is invariant under the symmetries of spacetime defined in special relativity, namely spacetime **translations**, **rotations**, and **Lorentz boosts**. The set of transformations that respect this symmetry is the **Poincare group**, which is itself a semi-direct product of spacetime translations and the Lorentz transformations. Through Noether's theorem, the translational symmetries lead to energy and momentum conservation, while the Lorentz symmetries give rise to conservation of angular momentum. The **Lorentz group**, $SO(3,1)$, can be defined by its **representations** that connect spin with relativistic transformations, where each particle in the SM is classified using pairs of half-integer values that characterize how they transform. The vector $(W^\pm/Z)$

bosons are defined by the $(\frac{1}{2}, \frac{1}{2})$ representation, the scalar Higgs boson by the $(0, 0)$ representation, and the fermions as a combination of two left- and right-handed Weyl spinors with the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation.

The Standard Model is a **non-Abelian gauge theory**, whose Lagrangian is invariant under the direct products of the individual gauge groups of quantum chromodynamics (QCD) and electroweak (EW) theory

$$SM = \overbrace{SU(3)_C}^{QCD} \times \overbrace{SU(2)_L \times U(1)_Y}^{EW}. \quad (1.7)$$

The following sections will develop this SM gauge formalism further as we discuss the components of the Standard Model Lagrangian.

1.2 Quantum Electrodynamics

Quantum electrodynamics (QED) describes the electromagnetic interaction between electrically charged fermions mediated by photons. It is an Abelian gauge theory with $U(1)$ symmetry and is thus invariant under the $U(1)$ global gauge transformation

$$\phi \rightarrow \phi' = e^{iq\alpha(\mathbf{x})}\phi, \quad (1.8)$$

where q is a parameter of the transformation and $\alpha(\mathbf{x})$ is a real coordinate-dependent phase. The Dirac Lagrangian describes the free propagation of the spin 1/2 fermion field ψ with mass m

$$\mathcal{L} = \bar{\psi}i\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \quad (1.9)$$

where γ^μ are the Dirac gamma matrices. Under the $U(1)$ transformation (Eq. 1.8), this becomes

$$\mathcal{L} \rightarrow \mathcal{L}' - q\bar{\psi}\gamma^\mu\partial_\mu\alpha(\mathbf{x})\psi, \quad (1.10)$$

As this Lagrangian now includes a coordinate-dependent term in the derivative, it is not explicitly invariant as specified. To make this Lagrangian invariant under the local gauge transformation, we can introduce a new **gauge field** A_μ that transforms like

$$A_\mu \rightarrow A_\mu + \partial_\mu\alpha. \quad (1.11)$$

The new gauge field A_μ couples to fermionic fields $\psi(x, t)$ with a strength q , and we can replace the standard derivative ∂_μ with the **gauge covariant derivative**

$$D_\mu = \partial_\mu + iqA_\mu, \quad (1.12)$$

and q gives the coupling strength of QED and, through Noether's theorem, the conserved electric charge e . The Dirac Lagrangian can then be written as

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi, \quad (1.13)$$

which is invariant under the $U(1)$ gauge transformation. The gauge field itself, A_μ , can be interpreted as the photon field and constructs the electromagnetic (EM) field tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (1.14)$$

We can then add a new term using the EM field tensor, the kinetic term of the photon $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, to obtain the full $U(1)$ invariant QED Lagrangian

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi. \quad (1.15)$$

All electromagnetic interactions can be derived via applying the Lorenz gauge condition $\partial_\mu A^\mu = 0$ and solving for the equations of motion via Eq. 1.1. The relative strengths of QED interactions are determined via the dimensionless coupling constant, the fine structure constant

$$\alpha_0 = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137}. \quad (1.16)$$

Divergent contributions to observables, such as the self-energy of the electron or the vacuum polarization, can be removed through **renormalization**, which involves redefining the parameters of the theory at different energy scales, and probabilities of QED interactions and decays can be calculated as a perturbative series in α_0 .

1.3 Quantum Chromodynamics

Quantum chromodynamics (QCD) describes the strong nuclear interactions between quarks and gluons. In addition to their electrical charge, quarks and gluons also carry units of color charge (r, g, b) or anticolor charge $(\bar{r}, \bar{g}, \bar{b})$. QCD is a non-Abelian gauge theory, unlike QED, and instead is invariant under an $SU(3)_C$ gauge symmetry, where C indicates the three color indices added to the $SU(3)$ Lie algebra. The $SU(3)$ Lie algebra has $3^2 - 1 = 8$ dimensions, and thus is constructed

via eight corresponding 3×3 generators T^a with $a = 1, \dots, 8$, collectively known as the **Gell-Mann matrices**. We can again write the local gauge transformation

$$\psi(x, t) \rightarrow e^{i\alpha^a(x)T^a} \psi(x, t). \quad (1.17)$$

We then must similarly introduce eight new gluon gauge fields A_μ^a that transform as

$$A_\mu^a \rightarrow A_\mu^a - \frac{1}{g_s} \partial_\mu \alpha_a - f_{abc} \alpha_b A_\mu^c, \quad (1.18)$$

and then we can define the covariant derivative

$$D_\mu = \partial_\mu - ig_s T^a A_\mu^a, \quad (1.19)$$

where g_s is the QCD coupling constant, often expressed as $\alpha_s = \frac{g_s^2}{4\pi}$. The Dirac Lagrangian can then also be written for a quark similarly as in QED, but with the addition of the color charges

$$\mathcal{L} = \sum_c \bar{\psi}_c (i\gamma^\mu D_\mu - m) \psi_c, \quad (1.20)$$

where

$$\psi_c = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}. \quad (1.21)$$

We can also construct a gauge-invariant kinetic term, including a self-interaction term between the gluon fields due to the non-Abelian nature of $SU(3)_C$, by defining the QCD field strength tensor

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g_s f_{abc} A_\mu^b A_\nu^c, \quad (1.22)$$

and thus, we can finally write down the full QCD Lagrangian

$$\mathcal{L}_{\text{QCD}} = \sum_c \bar{\psi}_c (i\gamma^\mu D_\mu - m) \psi_c - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (1.23)$$

QCD gives rise to some complex emergent behaviors that are particularly relevant in the collider physics landscape. Like QED and QFTs in general, perturbative predictions for QCD observables like cross sections and decay rates (see Section 1.8.1) are expressed in terms of the renormalized coupling constant, which is a function of the **renormalization scale** parameter μ_R . The μ_R parameter itself is not a physical quantity, but it can be taken to be the *scale* of the momentum transfer in some

process, Q , such that the coupling $\alpha_s(Q)$ is the effective interaction strength of that process. The coupling constant changing continuously as the interaction scale changes is one of the fundamental features of QCD, known as the **running of the coupling constant**. For a one-loop interaction, the renormalization group equation² [22]

$$\mu_R^2 \frac{d\alpha_s}{d\mu_R^2} = \beta(\alpha_s) = -\beta_0\alpha_s(\mu_R)^2 - \beta_1\alpha_s(\mu_R)^3 - \beta_2\alpha_s(\mu_R)^4 + \dots, \quad (1.24)$$

ultimately yields the energy-dependent strong coupling

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \beta_0\alpha_s(\mu^2) \ln\left(\frac{Q^2}{\mu^2}\right)}, \quad (1.25)$$

where β_0 is the one-loop beta function coefficient, related to the number of quark flavors n_f below the energy scale $b_0 = (33 - 2n_f)/12\pi$. This can be written more simply by defining the **QCD scale parameter** Λ , below which a_s begins to diverge, such that

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \ln\left(\frac{Q^2}{\Lambda^2}\right)}. \quad (1.26)$$

The value of Λ can be measured experimentally but varies depending on the number of quark flavors included in the computation, from approximately 100-300 MeV [22]. The negative sign in the first term of Eq. 1.24 implies that the strength of the coupling constant decreases as the energy scale increases, resulting in **asymptotic freedom** for values well above Λ . This results in bound quarks inside hadrons behaving as nearly free particles, resulting in frequent gluon-induced QCD processes at hadron colliders such as the LHC. Conversely, at low energy scales, the coupling is so large that quarks cannot be isolated. It is instead energetically more favorable to create a new quark-antiquark pair and thus a new bound state, a property known as **color confinement**. This results in the characteristic **jet** final state from hadronic interactions in colliders, resulting from the **hadronization** of individual quarks producing collimated sprays of secondary particles.

1.4 Electroweak Theory

The formalism begins to become more complex for the weak nuclear force, which controls nuclear decay processes such as radioactive β -decay. Enrico Fermi first developed an initial theory of weak interactions using a four-fermion interaction vertex between the neutron, electron, neutrino, and

²In the minimal subtraction ($\overline{\text{MS}}$) scheme.

proton. The formulation was further developed over time to include other experimental observations like the discovery of the muon and successfully described β -decay at low energies, but was not renormalizable and gave unremovable divergences at high energy scales. As experimental evidence already suggested spin-1 massive gauge bosons mediated the weak force, it was challenging to develop a local gauge invariant theory that was both well-behaved at high energy scales and included such particles.

In the 1960s, Glashow, Weinberg, and Salam [23, 24, 25] developed a model that unifies the weak force and electromagnetism into a single **electroweak model**, or GWS model, a non-Abelian gauge theory that obeys a $SU(2)_L \times U(1)_Y$ local symmetry. Like the electric charge q conserved in QED by the original $U(1)$ symmetry, the **weak hypercharge**, a function of the electric charge and weak isospin (T_3), is conserved by the $U(1)_Y$ symmetry. In the electroweak model, fermions are grouped into $SU(2)_L$ doublets

$$\text{Leptons} = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}, \quad (1.27)$$

$$\text{Quarks} = \begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} s \\ c \end{pmatrix}, \begin{pmatrix} b \\ t \end{pmatrix}. \quad (1.28)$$

the L in $SU(2)_L$ indicates that only left-handed fermions transform as doublets, while right-handed fermions transform as singlets. Thus, the weak interaction is a **chiral** theory. Any Dirac fermion field can be projected into its left- or right-handed components through the chiral operator, defined as $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, via

$$\psi_R = \frac{1 + \gamma^5}{2}\psi, \psi_L = \frac{1 - \gamma^5}{2}\psi. \quad (1.29)$$

The doublets in Eq. 1.31 are further decomposed into their Weyl spinor components, such that for the first generation of leptons

$$\chi_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \psi_R = e_R, \quad (1.30)$$

and for quarks

$$\chi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \psi_R = u_R, d_R, \quad (1.31)$$

and correspondingly for the second and third generations. The fermion fields transform under the local $SU(2)_L$ and $U(1)_Y$ symmetries as

$$\chi_L \rightarrow e^{i\beta(x)Y_L + i\alpha^a(x)T^a} \chi_L, \quad (1.32)$$

$$\psi_R \rightarrow e^{i\beta(x)Y_R} \psi_R, \quad (1.33)$$

where $\beta(x)$ and $\alpha(x)$ are the real phase transformations for $U(1)_Y$ and $SU(2)_L$. The generator of $U(1)_Y$, $Y_{L,R}$, is the weak hypercharge operator, and T^a are the $SU(2)$ generators as seen previously.

As with QED and QCD, to preserve local gauge invariance, we introduce three new gauge bosons W_μ^a for $SU(2)_L$ and a single new gauge boson B_μ for $U(1)_Y$, for which the covariant derivatives are

$$D_\mu \chi_L = (\partial_\mu + igW_\mu^a T^a + ig' Y_L B_\mu) \chi_L, \quad (1.34)$$

and

$$D_\mu \psi_R = (\partial_\mu + ig' Y_R B_\mu) \psi_R, \quad (1.35)$$

yielding field strength tensors

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (1.36)$$

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + gf_{abc} W_\mu^b W_\nu^c, \quad (1.37)$$

where the non-Abelian term in Eq. 1.37 again yields self-interactions between the weak gauge bosons with coupling strength g . We could write down an electroweak Lagrangian for both the fermion and gauge boson components here, but the electroweak gauge fields are generally reparameterized more physically like

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}}, \quad (1.38)$$

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \quad (1.39)$$

where charged-current and neutral-current interactions arise via the W^{pm} and Z bosons and the photon field A_μ . The **weak mixing angle** θ_W , which quantifies the total amount of *mixing* between the gauge bosons, can be expressed in terms of the $SU(2)_L$ coupling g' and the $U(1)_Y$ coupling g as

$$\sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}. \quad (1.40)$$

QED can be re-obtained from the photon fields A_μ through relationships between the weak hypercharge, the electric charge and QED coupling constant, and the projection of the weak isospin T_3

$$Y = Q - T_3, e = g \sin \theta_W = g' \cos \theta_W. \quad (1.41)$$

We can then consider the GWS Lagrangian [26]

$$\mathcal{L} = \sum_j^3 i \bar{\psi}_j \gamma^\mu D_\mu \psi_j, \quad (1.42)$$

where j runs over the three fermion generations and ψ includes both the left- and right-handed fields. There also exists the standard kinetic term between the field tensors

$$\mathcal{L}_{\text{kinetic}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu}, \quad (1.43)$$

which also generates both cubic and quartic self-interaction terms between the gauge bosons.³ Importantly, Eq. 1.42 leads to charged- and neutral-current interaction terms between the fermions and the gauge bosons. Plugging the parameterized definitions into the covariant derivative ultimately yields the charged-current interaction (shown here for a single generation)

$$\mathcal{L}_{\text{CC}} = -\frac{g}{2\sqrt{2}} \{ W_\mu^+ [\bar{u} \gamma^\mu (1 - \gamma^5) d + \bar{\nu} \gamma^\mu (1 - \gamma^5) e] + \text{h.c.} \}. \quad (1.44)$$

An interesting feature of the charged-current interaction is that only left-handed fermions and right-handed antifermions couple to the $W^\pm$ bosons, breaking both charge and parity symmetries. The neutral current interaction similarly can be derived

$$\mathcal{L}_{\text{NC}} = - \sum_j \bar{\psi}_j \gamma^\mu \{ A_\mu [g T_3 \sin \theta_w + g' Y_j \cos \theta_w] + Z_\mu [g T_3 \cos \theta_w - g' Y_j \sin \theta_w] \} \psi_j, \quad (1.45)$$

which can be decomposed into the QED Lagrangian and the neutral weak interaction through the relations in Eq 1.41 where

³Not derived here, for brevity!

$$\mathcal{L}_{\text{NC}} = \mathcal{L}_{\text{QED}} + \mathcal{L}_{\text{NC}}^Z, \quad (1.46)$$

where

$$\mathcal{L}_{\text{QED}} = -eA_\mu \sum_j \bar{\psi}_j \gamma^\mu Q_j \psi_j, \quad (1.47)$$

and

$$\mathcal{L}_{\text{NC}}^Z = -\frac{e}{2 \sin \theta_W \cos \theta_W} \left[\sum_j 2 \bar{\psi}_j \gamma^\mu (T_3 - \sin^2 \theta_W Q_j) \psi_j \right] Z_\mu. \quad (1.48)$$

The gauge bosons, as defined in the original theory, are massless. However, the observed short range of the weak force can only be explained by massive gauge bosons for the electroweak interactions, implying the electroweak symmetry is broken by some mechanism at a high energy scale. This **electroweak symmetry breaking (EWSB)** is ultimately understood through the **Higgs mechanism**, detailed in the next section.

1.5 Higgs Mechanism

The Higgs mechanism allows the gauge bosons and fermions to acquire mass while still preserving the $SU(2)_L \times U(1)_Y$ symmetry by introducing a complex scalar field ϕ in the spinor representation of the $SU(2)_L$ doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \quad (1.49)$$

with a new term in the SM Lagrangian

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2} (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) = \frac{1}{2} (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (1.50)$$

which is also invariant under $SU(2)_L \times U(1)_Y$. The first term describes the kinematics of the new field while the **Higgs potential**, $V(\phi)$, describes the quartic self-interactions of the scalar field and leads to **spontaneous symmetry breaking**. When $\mu > 0$, the minimum energy configuration of the fields is the trivial $\phi = 0$ case, describing a massive scalar particle with mass and coupling λ . However, more interestingly, when $\mu < 0$, the scalar potential develops a nonzero **vacuum expectation value (VEV)**. Due to the symmetry of $V(\phi)$, an infinite number of degenerate states with potential v exist with minimum energy

$$\phi^\dagger \phi = \frac{v^2}{2} = \frac{\mu^2}{2\lambda}, \quad (1.51)$$

as depicted in Figure 1.2.

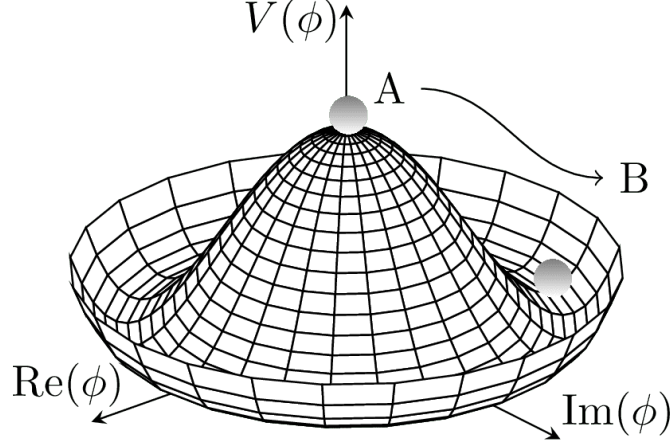


Figure 1.2. The Higgs potential after spontaneous symmetry breaking. Taken from [27]

Expanding the Higgs field around the minimum potential with a small $h(x), \xi(x)$ like

$$\phi(x) = (v + h(x))e^{i\xi(x)}, \quad (1.52)$$

and then substituting this into the Lagrangian (Eq. 1.50) generates three massless scalar **Goldstone bosons** $\xi(x)$ such that

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \xi_1 + i\xi_2 \\ v + h(x) + i\xi_3 \end{pmatrix}. \quad (1.53)$$

A unitary gauge transformation can eliminate the remaining degrees of freedom, where the Goldstone bosons are ultimately absorbed into the longitudinal degrees of freedom of the weak gauge bosons.

The real massive scalar Higgs field (h) then arises from the vacuum expectation

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (1.54)$$

1.5.1 Weak Gauge Boson (W/Z) Masses

Armed now with our vacuum expectation value, the kinetic component of Eq. 1.50 in the unitary gauge [26] becomes

$$(D_\mu \phi)^\dagger (D^\mu \phi) \rightarrow \frac{1}{2} \partial_\mu h \partial^\mu h + (v + h)^2 \left\{ \frac{g^2}{4} W_\mu^- W^{+\mu} + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right\}. \quad (1.55)$$

Note that the vacuum expectation value of the new scalar generates quadratic terms for the $W^\pm$ and Z , implying the gauge bosons have acquired masses

$$m_W = \frac{vg}{2}, m_Z = \frac{v}{2} \sqrt{g'^2 + g^2}, m_A = 0, \quad (1.56)$$

which implies that adding the new scalar Lagrangian to the electroweak Lagrangian gives masses to the weak gauge bosons, spontaneously breaking the $SU(2)_L \times U(1)_Y$ symmetry, but the photon preserves $U(1)$ symmetry and thus remains massless. Before EWSB, the massless $W^\pm$ and Z bosons possess 6 degrees of freedom due to the two possible polarizations of the massless spin-1 field and four related to the complex scalar field ϕ . After EWSB, the three Goldstone bosons are “eaten” by the $W^\pm$ and Z bosons, which become massive and therefore acquire a longitudinal polarization as well, leaving 9 degrees of freedom for the weak gauge bosons and one remaining degree of freedom from the Higgs boson, summing to the same total number of degrees of freedom.

1.5.2 Fermion masses

Fermion masses can be generated similarly. In the Dirac Lagrangian, the fermion mass terms, when expanded in Weyl spinor components, are

$$\mathcal{L}_{\text{F, mass}} = m(\bar{\chi}_L \psi_R + \bar{\psi}_R \chi_L). \quad (1.57)$$

Since this term contains couplings between the left-handed doublets and the right-handed singlets, it violates gauge invariance under $SU(2)_L$. However, this can be circumvented after adding the additional Higgs scalar field. The **Yukawa interaction** Lagrangian gives masses to the fermions through the fermion-scalar coupling. Taking a single term as an example, for leptons

$$\mathcal{L}_{\text{Yukawa}}^e = -\lambda_e \begin{pmatrix} \nu_e & e \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \text{h.c.}, \quad (1.58)$$

with new dimensionless coupling constants for each lepton flavor $\lambda_{\ell=e,\mu,\tau}$. The addition of the complex scalar results in singlets, which then are gauge invariant under the transformation, and after EWSB, this results in

$$\mathcal{L}_{\text{Yukawa}}^e = -\frac{\lambda_e v}{\sqrt{2}} \bar{e}e + \text{h.c.}, \quad (1.59)$$

where we have thus generated a mass term for the electron, which is also proportional to the Higgs vacuum expectation

$$m_e = \frac{\lambda_e v}{\sqrt{2}}. \quad (1.60)$$

Note that as there are no right-handed neutrinos in the Standard Model, a consequence of this formulation is that they cannot obtain their masses through the Yukawa interaction, yielding one of the unsolved problems in particle physics in light of the experimental observation of neutrino oscillations. The quark masses are generated analogously,

$$\mathcal{L}_{\text{Yukawa}}^{u,d} = -\frac{\lambda_u v}{\sqrt{2}} \bar{u}u - \frac{\lambda_d v}{\sqrt{2}} \bar{d}d + \text{h.c.}, \quad (1.61)$$

$$m_u = \frac{\lambda_u v}{\sqrt{2}}, m_d = \frac{\lambda_d v}{\sqrt{2}}. \quad (1.62)$$

However, the way this is written only allows quark pairs within the same generation to interact in their weak eigenstates. To allow for quark mixing, the Yukawa interaction must be written in the mass eigenstates, which is achieved by a unitary transformation of the quark fields. The **Cabibbo-Kobayashi-Maskawa (CKM) matrix** is then defined as the unitary matrix that diagonalizes the mass matrix, relating the weak eigenstates (d', s', b') and the mass eigenstates (d, s, b) via

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (1.63)$$

where the off-diagonal terms allow for transitions between the different quark generations, allowing for weak flavor-changing currents with the $W^\pm$ bosons.

1.5.3 The Higgs Boson

Lastly, we return again to the new Lagrangian for the scalar Higgs field in Eq. 1.50. After EWSB, the Higgs Lagrangian can be written

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2} \partial^\mu h \partial_\mu h - \mu^2 h^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4, \quad (1.64)$$

where the h field itself is a scalar particle, the **Higgs boson**, with mass

$$m_h = \sqrt{\frac{\lambda}{2}}v, \tag{1.65}$$

and cubic and quartic self-interaction terms. A new scalar particle with properties consistent with those of the Higgs boson with mass $m_h = 125$ GeV was observed simultaneously by the ATLAS and CMS experiments [1, 2] in 2012, after the completion of LHC Run 1. Since its observation, it has become the primary focus of a broad program to understand its properties and interactions, as these could potentially unlock an understanding of some unsolved problems in the Standard Model, discussed briefly in the next section.

1.6 Problems with the Standard Model

The Standard Model is the most complete and successful theory of nature to date. Its countless predictions, such as production cross sections and branching ratios for all SM particles, have been observed to high precision. The Higgs boson has been observed in all of its predicted dominant production modes, and no significant deviation from prediction has been found. This indicates that the interaction strengths between the Higgs and the rest of the SM are so far consistent with the model in every tested regime.⁴ Despite this, several significant unresolved issues with the SM suggest that it is not a fully complete theory of nature, leading to searches for **Beyond the Standard Model (BSM)** physics to address these issues and further our understanding of the fundamental structure of nature. Some of the significant limitations of the SM include:

- **Gravity:** Famously, the Standard Model does not describe a theory of gravity, which is instead described by General Relativity. The two theories are incompatible, but thankfully, the effect of gravity is negligibly weak on the scale of a collider experiment and does not have any tangible impact on experimental outcomes. However, a successful quantum theory of gravity that integrates the two is needed to describe the behavior of physics at the Planck scale.
- **Matter-Antimatter Asymmetry:** While an equal amount of matter and antimatter was necessarily created at the beginning of the universe, our current universe is now dominated by matter *far* beyond the amount predicted via the CP violation that is allowed in the SM via the CKM matrix. Additional sources of CP violation are required to make this observation consistent with Standard Model predictions.

⁴Ok, the controversial CDF W mass measurement [28] doesn't count.

- **Dark Matter:** Astrophysical observations such as galactic rotation curves and gravitational lensing suggest that the matter density of the universe is far greater than what is supported by the Standard Model, suggesting that up to 85% of the matter density of the universe is made of yet-unknown “dark matter” that does not interact meaningfully with known particles. No SM particle is a dark matter candidate due to cosmological theory constraints, such that there may be a hidden unknown “dark sector” extending the SM.
- **Neutrino Masses:** In the SM, neutrinos are massless, purely left-handed particles by construction. However, experimental observation of neutrino flavor oscillations implies they must have an exceedingly small but nonzero mass. Neutrino masses can be accommodated with extensions to the SM, but whether neutrinos are Dirac or Majorana particles remains an open question.

The motivation of the work presented in this thesis concerns an additional problem in the SM, **the hierarchy problem** (or *naturalness* problem), relating to unmitigated divergences in the Higgs self-coupling.

1.6.1 The Hierarchy Problem

Particle cross-sections are calculated by summing the matrix elements of all possible contributing Feynman diagrams resulting from the SM formulation. The calculation of leading order diagrams yields a finite amplitude, but those with *loop corrections*, i.e., self-energy contributions, yield unphysical infinite amplitudes. Without delving into the full formalism for brevity, renormalization, as discussed briefly earlier, removes unphysical infinities by redefining the observable parameters in terms of the renormalization scale μ_R . Particle masses are redefined as infinite “bare” parameters that absorb the higher-order infinities. The bare Higgs mass can be written as

$$m_0 = (1 + \delta)m_h, \tag{1.66}$$

where δ is an infinite **counterterm** and m_h is the observed normalized Higgs mass at a certain energy scale.⁵ The infinite integrals up to maximally high energy scales can be “cutoff” at a chosen energy scale parameter, Λ , assuming that results at lower, well-understood energy scales are not dependent on physics at extremely high energy scales above the cutoff.

⁵Physical observables are independent of the renormalization scale, μ_R , if they are calculated to all orders. As it’s not possible to do an infinite calculation, μ_R is generally chosen to minimize the contributions from higher-order diagrams. This choice of renormalization scale introduces a source of systematic uncertainty in Monte Carlo simulations of physics processes, discussed in a later section.

These self-energy contributions pose a particular problem for the Higgs mass, however, as the loop divergence is *quadratic* for scalar particles rather than logarithmically divergent for fermions. Taking an example of the one-loop self-energy contribution from a generic fermion with coupling λ_f , then the two-point correction is

$$\begin{aligned}\Pi_{hh}(0) &= (-1) \int \frac{d^4k}{(2\pi)^4} \text{Tr} \left(\frac{-i\lambda_f}{\sqrt{2}} \frac{i}{\not{k} - m_f} \frac{-i\lambda_f}{\sqrt{2}} \frac{i}{\not{k} - m_f} \right), \\ &= -2\lambda_f^2 \int \frac{d^4k}{(2\pi)^4} \frac{k^2 + m_f^2}{(k^2 - m_f^2)^2}, \\ &= -2\lambda_f^2 \int \frac{d^4k}{(2\pi)^4} \left(\frac{1}{k^2 - m_f^2} + \frac{2m_f^2}{(k^2 - m_f^2)^2} \right).\end{aligned}\tag{1.67}$$

The first term in the result is quadratically divergent, but the Higgs boson has been experimentally observed to have a mass around $m_h = 125$ GeV. Taking the cutoff of known physics to be at the Planck scale, around 10^{19} GeV where quantum gravity effects are believed to become dominant and SM predictions break down, one one-loop correction to m_h would have to be 30 orders of magnitude larger than m_h itself. For fermions, the corrections needed are on the same order of magnitude as their masses such that $\delta m \approx m$ as they are “protected” from large divergences by their chiral symmetry that produces cancellations. This implies one of two things – either there is new physics at some energy scale much below the Planck scale that contributes additional terms to stabilize the divergence, or there is a massive amount of **fine-tuning** such that the bare Higgs mass is so enormous it near-perfectly cancels the divergence at $\Lambda \approx 10^{19}$ GeV. This is not necessarily a mathematical problem but an aesthetic one – is the universe truly arranged such that one fundamental parameter is 10^{30} times larger than all of the other fundamental parameters? This is considered **unnatural** and conceptually suggests there should be new physics at some energy scale above 125 GeV.

To avoid fine-tuning, the scale of new physics should be within a few orders of magnitude of the measured Higgs mass before it fails to produce a natural result- up to approximately 10 TeV and thus within the maximum reach of the LHC. The most popular BSM theories that addressed this problem were **supersymmetric models (SUSY)**, which created a sector of bosonic “superpartners” to the SM fermions that each canceled out the divergent loop contributions and protected the Higgs mass from radiative corrections through the new boson-to-fermion symmetry being related to fermion masses protected by their chiral symmetry. Experimental results from ATLAS and CMS have broadly ruled out the most preferred formulations of SUSY, so a suite of other models has been proposed that solve fine-tuning at similar energies.

1.7 Heavy Vector Triplet (HVT) Model

As discussed in the previous section, many BSM theories attempt to solve the hierarchy problem by introducing new physics between the electroweak scale and the Planck scale, many of which generically predict the existence of heavy spin-1 resonances with small widths that couple electroweakly to SM particles, including the SM vector bosons and the Higgs boson. These new strong interactions would dynamically break the electroweak symmetry and stabilize the Higgs boson mass against large radiative corrections. These frameworks also make, in theory, testable predictions for physics processes that would be accessible at the center of mass energies studied at the LHC. However, not only would every specific model need its own dedicated analysis and individual interpretation of results, but each model depends on such a large number of free parameters that a direct fit to data would be practically impossible. On the other hand, resonance searches (“bump hunts”) are the simplest way to search for new particles and are ultimately not sensitive to all underlying free parameters of a given theory, instead being generally sensitive to only those parameters that control the mass and coupling strengths of the new resonance.

In light of this, a simplified phenomenological model can be developed that only includes these properties as free parameters. The **Heavy Vector Triplet (HVT)** model [12] introduces such a Lagrangian which predicts new heavy vector bosons, W' and Z' , that have a significant branching fraction to $V' \rightarrow WH$. Experimental resonance searches, such as the one presented in this thesis, can give results interpreted through this model and presented in terms of the phenomenological parameter space, which can then be easily translated into any specific “fundamental” model that can compute the HVT Lagrangian parameters analytically.⁶

1.7.1 Lagrangian and Model Parameters

In addition to the Standard Model fields and interactions, the HVT model considers a real vector V_μ^a , where $a = 1, 2, 3$ in the adjoint representation of $SU(2)_L$ with vanishing hypercharge. This describes one charged and one neutral heavy spin-one resonance with the charge eigenstate fields defined as:

$$V_\mu^\pm = \frac{V_\mu^1 \mp iV_\mu^2}{\sqrt{2}}, \quad V_\mu^0 = V_\mu^3. \quad (1.68)$$

A simple phenomenological Lagrangian for the new vector is defined as

⁶Care must be taken, however, that the simplified model is not a complete theory and should not be used outside its realm of validity. It only well-describes on-shell resonance production and decay, and as such, searches should only be sensitive to on-shell effects and insensitive to off-shell ones.

$$\begin{aligned}\mathcal{L}_V = & -\frac{1}{4}D_{[\mu}V_{\nu]}^a D^{[\mu}V^{\nu]}_a + \frac{m_V^2}{2}V_\mu^a V^{\mu a} + ig_V c_H V_\mu^a H^\dagger \tau^a \overleftrightarrow{D}^\mu H + \frac{g^2}{g_V} c_F V_\mu^a J_F^{\mu a}, \\ & + \frac{g_V}{2} c_{VVV} \epsilon_{abc} V_\mu^a V_\nu^b D^{[\mu}V^{\nu]}_c + g_V^2 c_{VVHH} V_\mu^a V^{\mu a} H^\dagger H - \frac{g^2}{2} c_{VW} \epsilon_{abc} W^{\mu\nu a} V_\nu^b V_\mu^c,\end{aligned}\quad (1.69)$$

where the first two terms above contain the V_μ^a kinetic and mass terms, as well as trilinear and quadrilinear interactions with the vector bosons from the covariant derivatives

$$D_{[\mu}V_{\nu]}^a = D_\mu V_\nu^a - D_\nu V_\mu^a, \quad D_\mu V_\nu^a = \partial_\mu V_\nu^a + g\epsilon^{abc}W_\mu^b V_\nu^c, \quad (1.70)$$

where g is the SM $SU(2)$ gauge coupling. As the new V_μ^a fields mix with the gauge boson fields W_μ^a after EWSB, they are not mass eigenstates and m_V is not the physical resonance mass. Diagonalizing the mass matrices of the new vector bosons gives their physical masses and the relationship between their masses $(m_{W'}, m_{Z'})$ and the SM vector boson masses (m_W, m_Z) ,

$$m_W^2 m_{W'}^2 = \cos\theta_W^2 m_Z^2 m_{Z'}^2. \quad (1.71)$$

As lower masses for these new bosons have already been excluded [29], the measured value of the weak mixing angle is only restored when the masses are (nearly) degenerate. The following two terms contain the direct interactions of V_μ^a with the Higgs current

$$iH^\dagger \tau^a \overleftrightarrow{D}^\mu H = iH^\dagger \tau^a D_\mu H - iD^\mu H^\dagger \tau^a H, \quad (1.72)$$

and the left-handed fermionic currents

$$J_a^\mu = \sum_f \bar{f}_L \gamma^\mu \tau^a f_L, \quad (1.73)$$

where $\tau^a = \sigma^a/2$ are the generators of the $SU(2)$ Lie algebra. This Lagrangian introduces several coupling parameters, where it is simpler in terms of experimental limit setting to treat the combinations that enter the vertices as fundamental parameters, e.g. the **bosonic coupling** $g_H = g_V c_H$ and the **fermionic coupling** $g_F = g^2/g_V c_F$ which assumes a universal value for the different fermion types in the context of the Wh resonance search. The g_H term introduces vertices that involve the physical Higgs field and the unphysical Goldstone bosons resulting from EWSB, which represent the longitudinal degrees of freedom of the $Z, W^\pm$ gauge fields. Thus, it describes the couplings of the new vector field to the Higgs boson as well as the SM gauge bosons, determining the decay rates of the new heavy vector bosons into Zh and Wh . The g_F parameter controls the resonance production via quark-antiquark annihilation as well as direct fermionic decays. The remaining terms of the

Lagrangian contain the quadratic and cubic couplings, which do not contribute directly to the V' decays or single production processes that are relevant for LHC phenomenology. Therefore, the only free parameters relevant for a search interpreted through the HVT framework are g_H, g_F , and m_V . In the search presented in this work, a wide range of possible values are tested for m_V ($m_{W'}$), and the search is interpreted through two benchmark models.

1.7.2 Benchmark Models

Two simplified benchmark models are proposed in Ref [12], **Model A** and **Model B**. These models fix c_H and c_F to specific values and leave the resonance mass and the coupling strength g_V as free parameters. The HVT simulations used in the Wh resonance search additionally fix g_V at specific benchmark points for each model, leaving the resonance mass as the final discriminant.

Model A is based on the light gauge model from Ref [15], which gives a weakly coupled extension to the SM gauge group and introduces a new $SU(2)$ gauge symmetry where g_V is typically small. The benchmark point sets $g_V = 1$, and the fermionic and bosonic couplings are comparable such that $g_F \approx g_H$. *Model B*, on the other hand, is based on the composite Higgs model from Ref [16], which gives a strongly coupled extension to the SM and has larger values for g_V . The benchmark point for this model is set to $g_V = 3$, and the fermionic couplings are suppressed such that $g_H \approx -g_V$ and $g_F \approx g^2/g_V$. This implies that *Model A* has comparable branching fractions to fermions and diboson final states, while *Model B* is dominated by diboson decays.

1.8 Collision Phenomenology and Monte Carlo Simulations

Within ATLAS and collider experiments more broadly, data collection focuses on **scattering** interactions between colliding particles and their constituents, resulting in a wide variety of final states that facilitate both precision measurements of SM processes and searches for potential new BSM particles. The physics of pp collisions that determine the final states from the countless fundamental interactions in a given event is enormously complex, ultimately governed by the predictions of the Standard Model and the underlying QFT. As protons are bound states of quarks and gluons, QCD effects dominate pp interactions and thus must be well-modeled to accurately simulate both SM background and BSM signal using **Monte Carlo (MC)** simulations, which are critical to all aspects of ATLAS analyses, from optimizing analysis selections to performing statistical analysis on observed data. A sketch of the QCD interactions from a pp collision simulated in an MC generator is shown in Figure 1.3. It is not possible to compute all effects from first principles directly. Instead, event modeling is done in several stages using specialized algorithms that work to estimate higher-

order perturbative and non-perturbative effects. This section will provide a generalized overview of pp collision phenomenology and MC event generation.

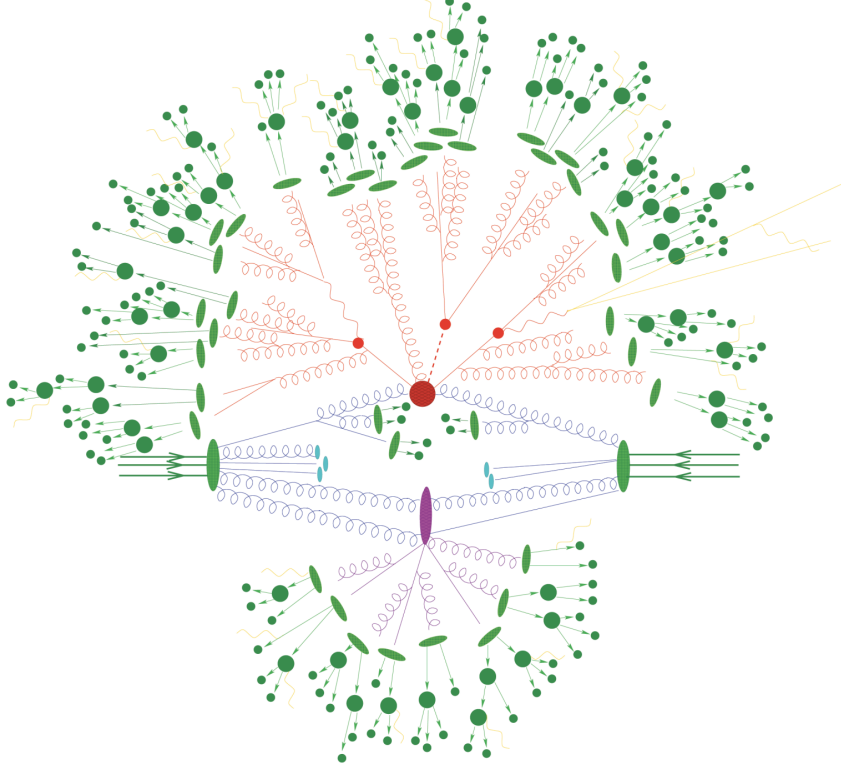


Figure 1.3. Sketch of QCD processes from a proton-proton collision. The red point in the center represents the hard interaction, representing the largest momentum transfer. The initial and final state radiation (ISR/FSR) of gluons and $q\bar{q}$ production are shown in blue and red, respectively. Partons then fragment into a parton shower (PS) and hadronize into mesons and baryons, shown in green. The underlying event (UE) from interactions between non-colliding partons, producing further activity, is shown in purple. Soft photon radiation during hadron decays is shown in yellow. Taken from Ref [30].

1.8.1 Cross Sections and Decay Rates

The total rate of observed events for a given scattering process depends on the beam properties (described in more detail in Section 2.1) as well as interactions between individual particles. This **event rate** is related to the probability for a given scattering process to occur by the scattering **cross section**, σ . This is classically associated with the physical cross-sectional area of interaction for two particle beams:

$$\sigma = \frac{N_i A}{N_a N_b}, \quad (1.74)$$

where N_i is the number of total scattering interactions, A is the cross-sectional area of the beam, and $N_{a,b}$ is the number of particles per beam in the given unit time.

Cross sections are measured in **barns**, a specialized unit of area where $1\text{b} = 10^{-24}\text{cm}^2$.⁷ These principles hold analogously in QFT, where the scattering cross-section can be calculated with the *matrix elements* (amplitudes) for the desired process $|\mathcal{M}|$ calculated from the perturbative expansions and the available *phase space* using **Fermi's Golden Rule**. When two particles collide, producing up to n new particles, e.g

$$1 + 2 \rightarrow 3 + 4 + \cdots + n, \quad (1.75)$$

the differential cross section is then given by

$$d\sigma = \frac{|\mathcal{M}|^2}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}} \frac{S}{(2\pi)^3} \left(\prod_{i=3}^n \frac{d^3 \vec{p}_i}{(2\pi)^3 2E_i} \right) (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - \cdots - p_n), \quad (1.76)$$

where $p_i = (E_i, \vec{p}_i)$ is the four-momentum of a particle i with mass m_i , S is a statistical factor $(1/j!)$ for each group of j identical particles in the final state, and the delta function enforces conservation of energy and momentum when integrated for the total cross section.

The decay rate, Γ , which is the probability of a particle decaying in a given unit time, can also be calculated similarly. If a particle decays into n new particles

$$1 \rightarrow 2 + 3 + 4 + \cdots + n, \quad (1.77)$$

the differential decay rate is again given by the Golden Rule via

$$d\Gamma = |\mathcal{M}|^2 \frac{S}{2m_1} \left(\prod_{i=2}^n \frac{d^3 \vec{p}_i}{(2\pi)^3 2E_i} \right) (2\pi)^4 \delta^4(p_1 - p_2 - p_3 - \cdots - p_n), \quad (1.78)$$

The **branching ratio** of a process is then just the decay rate of the individual mode divided by the total decay rate

$$\text{BR}(P_1 \rightarrow X) = \frac{\Gamma_X}{\Gamma_{\text{Total}}}. \quad (1.79)$$

⁷This is literally named after “the broad side of a barn”, owing to the relatively large size of the base unit. Manhattan project physicists, lamenting a lack of a proper unit for the cross section, couldn't settle on a choice of “great physicist” to bestow an eponym, and so settled on a witty quip instead. In the LHC era, most processes of interest have very small cross sections, and thus are more commonly referenced in terms of femtobarns (fb) and picobarns (pb).

1.8.2 Hard Scatter and Parton Distribution Functions

To model a proton-proton collision event, the first stage is to simulate the primary **hard-scatter** interaction. Protons are composite particles made of bound states of quarks and gluons, collectively referred to as **partons**. At the high energies found in a pp collision, asymptotic freedom allows the quarks and gluons within each proton to essentially act as freely interacting particles. Thus, each parton individually carries a fraction of the total momentum x_i of the incoming proton, distributed among all constituent partons. The process with the most significant momentum transfer Q is the hard-scatter interaction, producing the physics of interest for a given collision. The cross section for a hard scattering process of two initial-state protons with momenta p_1 and p_2 producing a final state X is given by

$$d\sigma_{p_1 p_2 \rightarrow X} = \sum_i \sum_j \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1, \mu_F) f_j(x_2, \mu_F) d\sigma_{ij \rightarrow X}(x_1 p_1, x_2 p_2, \mu_F, \mu_R), \quad (1.80)$$

where i and j are the individual partons and $d\sigma_{ij \rightarrow X}$ is the specific partonic cross section for a given initial parton state i, j . The functions $f(x, \mu_F)$ are experimentally-derived **parton distribution functions (PDFs)**, which give a probability density of a parton inside a proton carrying a longitudinal momentum fraction x evaluated at a specific energy scale μ_F which is generally chosen to be around the total momentum transfer in the hard scattering process such that $\mu_F = |Q|$. The partonic cross sections can be calculated perturbatively as an expansion in terms of the strong coupling constant α_s

Due to initial-state radiation (ISR), infrared divergences occur during the cross section calculation, where μ_F is defined as a cutoff value. If emitted particle momenta are larger than this value, they are included in the perturbative calculation directly. Below the threshold, the radiative effects are included in the PDF. Ultraviolet divergences also occur due to loop diagrams, as discussed in previous sections, controlled by the renormalization scale parameter μ_R . Physics itself is independent of the choice of μ_F and μ_R , but only when considering an infinite number of orders in perturbation theory. Thus, the finite-order cross section predictions used in collider physics depend on the two energy scales, and therefore the values are set around the scale of the particular process being evaluated to minimize overall scale dependence effectively. Any residual scale dependence is taken as a systematic uncertainty on the predicted cross sections by varying these parameters and assessing how the simulations evolve.

As perturbative QCD calculations fail at low energy scales, PDFs are determined by dedicated deep inelastic scattering experiments. The PDF of a particular parton depends on the value of Q , as shown in the NNPDF3.1NNLO distributions in Figure 1.4. These differences arise from QCD effects relating to gluon emission and $g \rightarrow q\bar{q}$ production. The “valence” quarks (u, d) generally contain a large portion of the momentum fraction, as expected, while the “sea” quarks and gluons are more likely to contain a low portion of the momentum fraction. The fraction of gluons and sea quarks increases at small x values with a larger total momentum transfer.

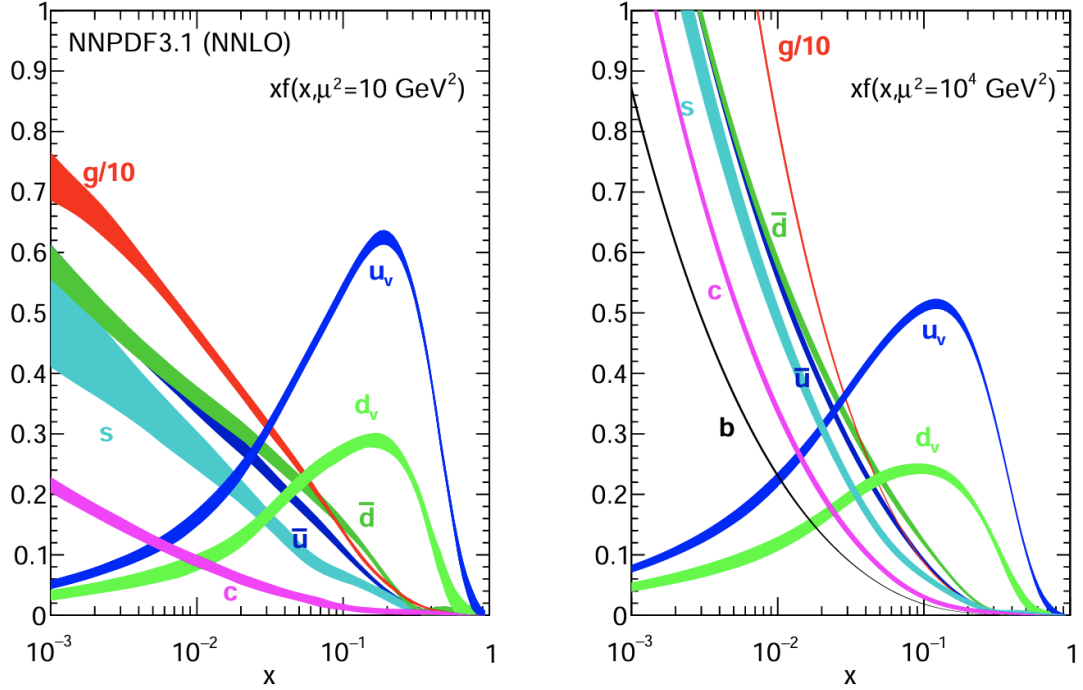


Figure 1.4. Parton distribution functions for a reference momentum transfer of $Q^2 = 10 \text{ GeV}^2$ (left) and $Q^2 = 10^4 \text{ GeV}^2$ (right) from the NNPDF3.1NNLO PDF set. Different flavors of constituent partons are shown in different colors. Taken from Ref [31].

1.8.3 Parton Showering and Hadronization

Similar to bremsstrahlung radiation from electrically charged particles, partons with color charge emit gluon radiation when accelerated. As gluons themselves have color charge, they too emit radiation, and the process repeats, forming a **parton shower (PS)**. Parton showers result in both initial-state radiation from the incoming parton and final-state radiation from the outgoing partons produced in the collision. While this process results in some complex topologies, **parton shower algorithms** can approximate this, mainly when the emitted gluon is nearly collinear to the original

parton. This splitting process is a higher-order effect and is largely independent of the rest of the scattering event, such that the cross section of the parton emission diagram can be factorized as a function of the leading-order cross section times a **splitting function** that parameterizes the fraction of initial parton energy as a function of the emission opening and azimuthal angles. The parton showering algorithm repeats successively until the average energy of the products falls below the hadronization scale.

The perturbative calculation of the matrix elements described in the previous sections performs well at simulating lower-order effects and can simulate several “hard” gluon emissions that are well-separated, while the parton showering simulates the higher-order splitting effects in the collinear limit. These simulations can overlap, however, and duplicate elements must be accounted for to avoid double counting. This **matching** of the tree-level **matrix elements to the parton shower** is performed using a variety of techniques and introduces a source of systematic uncertainty into searches and measurements performed with Monte Carlo simulations.

Once the parton showering is complete, the final step in simulating the hard-scatter interaction is the formation of colorless bound hadron states from the quarks and gluons produced in the parton shower. Once the energy falls below the QCD scale parameter, they can no longer freely propagate, and the subsequent **hadronization** must be simulated. This low-energy regime where the strong coupling constant is large means that perturbative calculations are no longer valid and, instead, must be simulated using non-perturbative models. It is not possible to compute hadronization analytically at this time; thus, event generation broadly relies on two phenomenological models, the string and cluster models. The string model assumes that the potential energy between a quark-antiquark pair increases linearly as the distance between them grows, where a color-field “string” that connects two quarks under “tension” breaks at the point where it is energetically favorable to create a new quark-antiquark pair. This process is iterative, where the new quark pairs create new “strings” that can “stretch” and break again to produce a new quark pair, ending when the available energy is expended, and no new pairs can be formed. Cluster hadronization instead clusters the shower partons into colorless groups, as in the transition energy between confinement and the parton shower, particles with color charge begin to cluster along these groups due to the strong interaction. These phenomenological models involve many different tunable parameters that control the underlying dynamics to match better the observed data – specific sets of these parameters are known as **tunes**.

1.8.4 Underlying Event and Simulation

Lastly, any remaining interactions in the event that do not come from the primary hard-scatter interaction must be simulated. Soft interactions between partons not participating in the scattering process form the **underlying event (UE)**, including final state particles from the ISR and FSR, residual beam remnants, etc. Additionally, multiple proton-proton collisions occur during each bunch crossing, known as **pileup** (See Section 2.1.3). These interactions are modeled using simulated soft inelastic collisions, known as **minimum bias** events, which are *overlaid* on top of the primary hard-scattering interaction based on a standardized pileup profile with a specific amount of simultaneous interactions based on the average pileup in a given datataking period.

Once the event generation is complete, the entire set of simulated “truth” particles are processed with an advanced GEANT4 [32] simulation of the full ATLAS detector in the **simulation** step. Each component of the detector response and material interactions as each simulated particle traverses the ATLAS volume is modeled, and then in the **digitization** step, the corresponding signals from the simulated detector electronics are read out. Once this is done, the complete ATLAS reconstruction chain (See Section 2.3) can be run on these signals precisely as is done for the actual collected data, and the Monte Carlo production is complete.

CHAPTER 2

THE ATLAS EXPERIMENT

Operated by the **European Center for Nuclear Research (CERN)** in Geneva, Switzerland, the **Large Hadron Collider (LHC)** is the world’s largest particle accelerator.¹ Searches for TeV-scale new physics such as the one presented in this thesis are only accessible using extremely high-energy particle-particle collisions, made possible by the LHC’s unprecedented large scale and sophisticated design.

As of this writing, nine experiments collect data from the hadron-hadron collisions produced at the LHC, each designed to explore a different range of SM and BSM phenomena. The **ATLAS (A Toroidal LHC ApparatuS) Experiment**² is one of two general-purpose detectors at the LHC that are designed to be sensitive to a wide range of physics processes, from precision measurements of Standard Model properties to searches for BSM physics at the “energy frontier”. More than 6000 physicists, engineers, students, technicians, and staff contribute to the design, data analysis, and daily operations of the experiment, with more than 3000 signing authors among them.

This chapter presents an overview of the LHC and the ATLAS detector, discussing the key details of the experimental design and data reconstruction underpinning the work featured in this thesis. More detailed descriptions of the LHC and the ATLAS detector can be found in Refs [33] and [34], respectively.

2.1 The Large Hadron Collider

The LHC is a multi-purpose particle accelerator situated approximately 100 m under the Swiss-French border near Geneva, Switzerland, designed to accelerate and collide hadron pairs from two counter-directed particle beams. The circular accelerator complex spans a 27 km circumference, integrated into the 3.8 m wide concrete tunnels originally designed for the Large Electron-Positron (LEP) collider, the predecessor of the LHC [35, 36]. It extends from the CERN laboratory site in Meyrin, Switzerland, to the quiet agricultural village of Cessy, France, with the four major LHC

¹It is also the world’s largest *machine*, in general!

²A particularly egregious example of questionable acronyms coined by physicists.

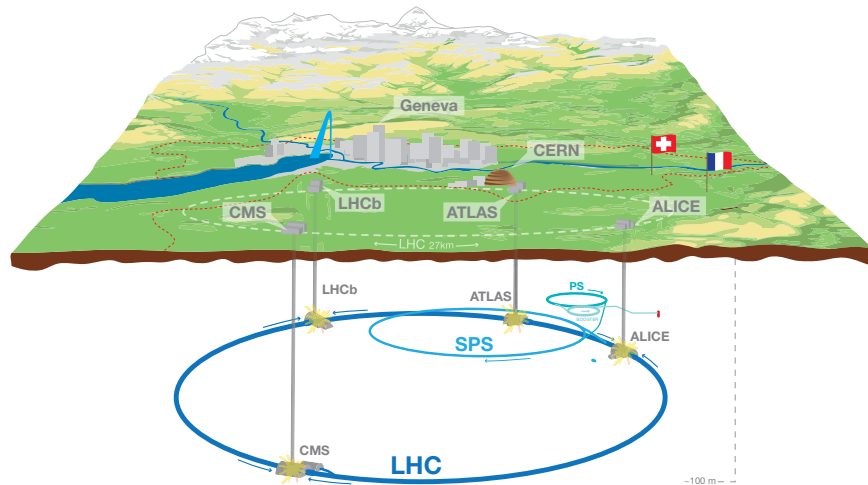


Figure 2.1. View of the LHC and surrounding environment under the Swiss-French border and greater Geneva area. Taken from [37]

experiments situated at distant points along the ring, as shown in Figure 2.1. Unlike LEP, a particle-antiparticle collider, the LHC is a particle-particle collider primarily designed for proton-proton collisions but additionally equipped to accelerate heavy ions. This allows for proton-proton (pp), proton-lead (p -Pb), and lead-lead (Pb-Pb) collisions in varying campaigns throughout the LHC run schedule, at a design maximum center-of-mass energy of $\sqrt{s} = 14$ TeV for proton-proton collisions and 5.5 TeV for heavy ion collisions.^{3 4}

The four major LHC experiments, ATLAS, CMS [38], ALICE [39], and LHCb [40], are uniquely aimed to be sensitive to varying physics processes. As stated previously, ATLAS and CMS are general-purpose detectors designed to capture an extensive range of final states for both precision Standard Model measurements and potential BSM searches. Both experiments have relatively similar designs and overall sensitivity, meant to ensure rigor and reproducibility in results by enabling

³Due to technical complications relating to maintaining superconductivity in the dipole magnets, LHC operations have not been able to reach the maximum intended design energy and instead proton-proton runs have been operated at 7-8 TeV in Run 1, 13 TeV in Run 2, and 13.6 TeV in Run 3. This analysis presented in this work uses data from the Run 2 datataking period.

⁴Unless otherwise specified, this work assumes pp collisions from this point. Heavy ion physics is outside the scope of the author's expertise.

cross-verification of observations by two independent communities. ALICE studies quark-gluon plasma using heavy ion collisions, and LHCb uses a forward detector to study heavy-flavor physics and CP violation through both searches and precision measurements. Five other smaller experiments are also operated at the LHC, including FASER, a new detector inserted 480 m downstream from ATLAS [41]. This experiment is aimed at long-lived particle searches and neutrino direct detection, using extremely forward light decay products that escape through the ATLAS beampipe.

2.1.1 Injection Chain and Accelerator Complex

Particles are first accelerated from rest to successively higher energies by a sequence of smaller accelerators, collectively known as the **injection chain** [42], before they enter the main LHC ring. In pp collisions, protons are produced by ionizing atoms from bottles of stored hydrogen gas, which are then fed through an initial linear accelerator (**LINAC 2** for Runs 1 and 2, replaced by **LINAC 4** in 2020 for Run 3 and beyond⁵) to bring their energy up to 50 (Runs 1-2) or 160 MeV (Run 3). They are subsequently brought to 1.4 GeV (Runs 1-2) or 2 GeV (Run 3) by the **Proton Synchrotron Booster** (PSB, BOOSTER) and then injected into the **Proton Synchrotron** (PS) to reach 25 GeV (Runs 1-2) or 26 GeV (Run 3).⁶ The PSB is made of four superimposed synchrotron rings with a 25 m radius, where the four “beamlets” are merged into a single beam before being injected into the 100 m radius PS. The increased injection energy supplied by the PSB allows more protons to be injected into the PS, with a higher luminosity at the end of the accelerator chain and thus sufficient for the needs of the LHC.⁷

The PS accelerates the protons through its radio-frequency (RF) cavities using 277 conventional electromagnets, including 100 dipole magnets to bend the beam into a circular path, and then passes them to the **Super Proton Synchrotron** (SPS), which accelerates them to 450 GeV. The SPS is the last accelerator in the chain before the LHC and the second-largest machine in the CERN

⁵LINAC 2 accelerated protons for multiple CERN experiments from its construction in 1978, but was finally retired during Long Shutdown 2 to accommodate higher-intensity beams in preparation for the HL-LHC era. This subsequently changed the energies for each stage in the injection chain. LINAC 4 also technically accelerates hydrogen anions, H^- , rather than bare protons, H^+ , as in LINAC 2. The extra electron is subsequently removed by a stripping foil at the PSB injection point.

⁶In heavy ion runs, the PS instead receives particles delivered from the Low Energy Ion Ring (LEIR).

⁷The PS is CERN’s oldest accelerator still in operation, originally designed as CERN’s flagship synchrotron in 1959. Technological advances in the coming decades would render it obsolete for state-of-the-art experiments, but it gained a new lease on life in supplying particles to many newer accelerators from the 70s to the present.

CERN's Accelerator Complex

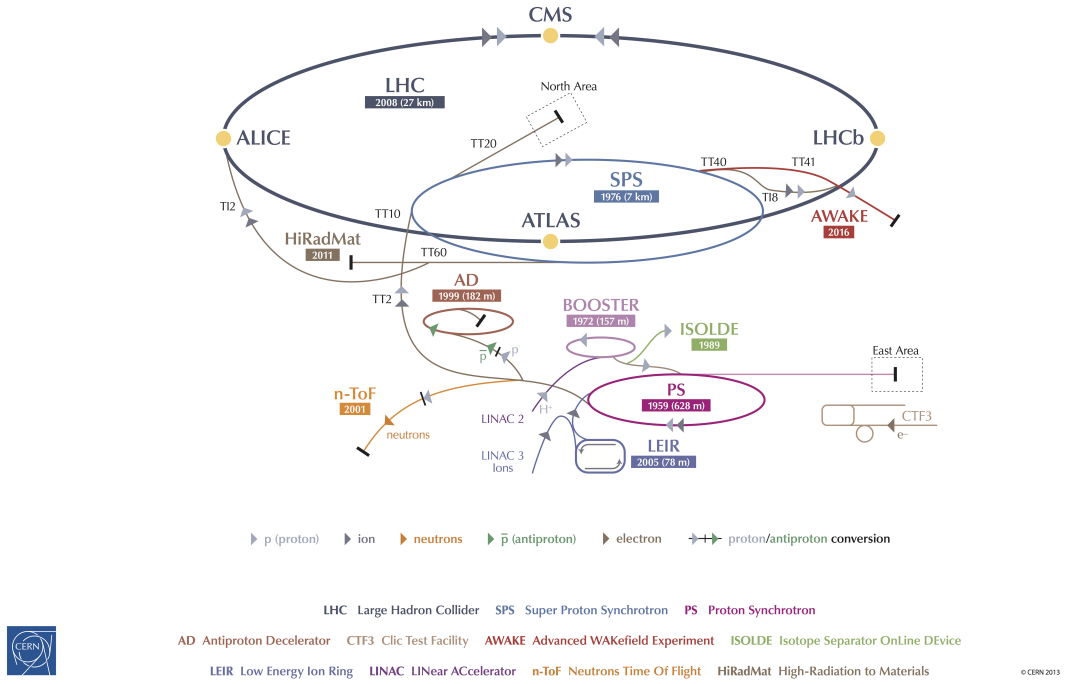


Figure 2.2. The complete CERN accelerator complex, including non-LHC experiments. Taken from [45].

accelerator complex with a 1.1 km radius and more than 1300 conventional electromagnets.⁸ A schematic of the various stages of the complex is shown in Figure 2.2.

2.1.2 LHC System

After reaching the final injection energy of 450 GeV, protons are serially injected into the LHC ring in opposing directions through two 2 km-long tunnels. Eight “points” along the beamline allow access from the surface and contain experimental and service facilities, shown in figure 2.3. The four main experiments, ATLAS, ALICE, CMS, and LHCb, are housed at Points 1, 2, 5, and 8, respectively. At these **interaction points (IPs)**, the two opposing beams are brought together to produce collisions to be recorded at the respective experiments. The remaining points are service caverns, which contain various infrastructure and support facilities. Points 3 and 7 contain the

⁸Originally commissioned in 1976, the SPS operated as a proton-antiproton collider from 1981 to 1991, resulting in several landmark discoveries including the discovery of the W and Z bosons in the UA1 [43] and UA2 [44] experiments. Like the PS, its illustrious career as a standalone collider is over, but it remains a crucial component of the LHC injection chain and supplies beams for several fixed-target experiments.

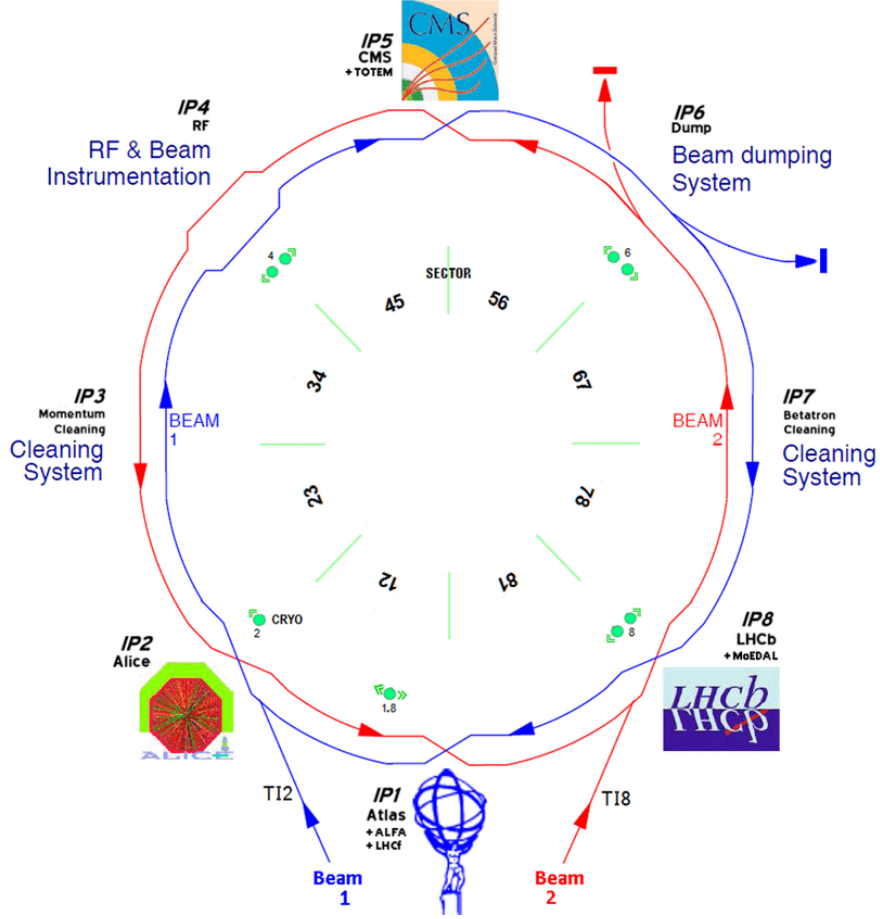


Figure 2.3. Layout of the LHC with the eight experimental and service points labeled. Beam 1 is injected close to ALICE and circulates clockwise, while Beam 2 is injected close to LHCb and circulates counterclockwise. The two beams exchange positions between the interior and exterior rings at each experiment to ensure the same overall path length. Taken from [46]

beam-cleaning systems, where the beams are collimated and particles that stray too far from the main beam paths are absorbed, protecting delicate accelerator components from potential damage. Point 4 contains the beam instrumentation and RF cavities that accelerate the protons from the injection energy to a final energy of 6.5 (Run 2) or 6.8 (Run 3) TeV over around 20 minutes. Finally, Point 6 houses the beam dumping system, where kicker magnets can be quickly inserted to safely deflect the beam out of the LHC and into an external 8.5 m long cylindrical graphite absorber. Beams are routinely dumped during normal runtime to prepare for a new fill of protons, but this also protects the equipment in case of a malfunction or other emergency.

The beams are kept confined in the LHC ring through a series of 1232 superconducting niobium-titanium dipole magnets. These 15 m long magnets, along with both beam pipes, are housed inside

a cryostat cooled to 1.9 K using superfluid helium. Each magnet produces an 8.33 T magnetic field to bend the beam into a circular orbit. Additionally, 392 3 m long quadrupole magnets focus and collimate the beam in the transverse plane alongside thousands of smaller corrector magnets. These keep the beam collimated to within $15\mu\text{m}$ prior to the interaction points and maintain the beam radius through varying energies during the ramp-up stage.

The protons comprising each beam are grouped into “bunches” of 10^{11} particles by the RF cavities, which produce an oscillating electromagnetic field at 400 MHz. RF cavities shape the beam by keeping protons synchronized with the RF frequency and at a stable energy – a proton that arrives with the target energy will not be accelerated, while those with different energies arriving earlier or later will be accelerated or decelerated by the RF voltage at the corresponding point in the oscillation. These envelopes are **RF buckets**, which are separated with a filled bunch spacing of 25 ns for an approximately 40MHz collision rate.⁹

2.1.3 Luminosity and Pileup

The number (N) of pp collisions in a detector at a given unit time t is a function of the **instantaneous luminosity** and the interaction cross-section:

$$N = \sigma_{pp} \int \mathcal{L}(t) dt. \quad (2.1)$$

The instantaneous luminosity $\mathcal{L}(t)$ delivered to the experiments is contingent on several accelerator and beam conditions and can be approximately written as:

$$\mathcal{L} = f \frac{k_b n_1 n_2}{4\pi\sigma_x\sigma_y}, \quad (2.2)$$

where n_i are the number of particles per bunch, k_b is the number of bunches per beam, f is the frequency of revolution, and σ_x and σ_y are widths of the bunches in the transverse plane, assuming a Gaussian distribution. Thus, the total **integrated luminosity** produced by the collider over a period T , which is largely proportional to the amount of data collected by each detector, is simply the integral:

⁹Intentionally, not all RF buckets are filled with protons, due to both the fact that the collision rate would be too high for the detector technologies and that a series of unfilled buckets ensures an “abort gap” to account for the small but nonzero amount of time to activate the kicker magnets in the event of a beam dump. Collisions also deplete the RF buckets throughout a fill.

$$L = \int_0^T \mathcal{L} dt. \quad (2.3)$$

The LHC was originally designed to deliver an instantaneous luminosity of $\mathcal{L} = 10^{34} \text{cm}^{-2} \text{s}^{-1}$, but the peak luminosities delivered to ATLAS during Run 2 and Run 3 have exceeded this value by a large margin.¹⁰ The peak luminosity recorded during Run 2 was $2.1 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$, increasing further in the ongoing Run 3 to $2.6 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$ at the time of this writing. Shown in figure 2.4a, the total integrated luminosity delivered by the LHC over Run 2 amounted to 156fb^{-1} , with 147fb^{-1} total recorded by ATLAS and 139fb^{-1} in the “**Good Run List**” (GRL) with a fully functional detector and normal data-taking conditions. The total integrated luminosity delivered so far in Run 3 during the 2022 and 2023 data-taking periods corresponds to 70fb^{-1} , with 66fb^{-1} recorded by ATLAS, as seen in figure 2.4b.

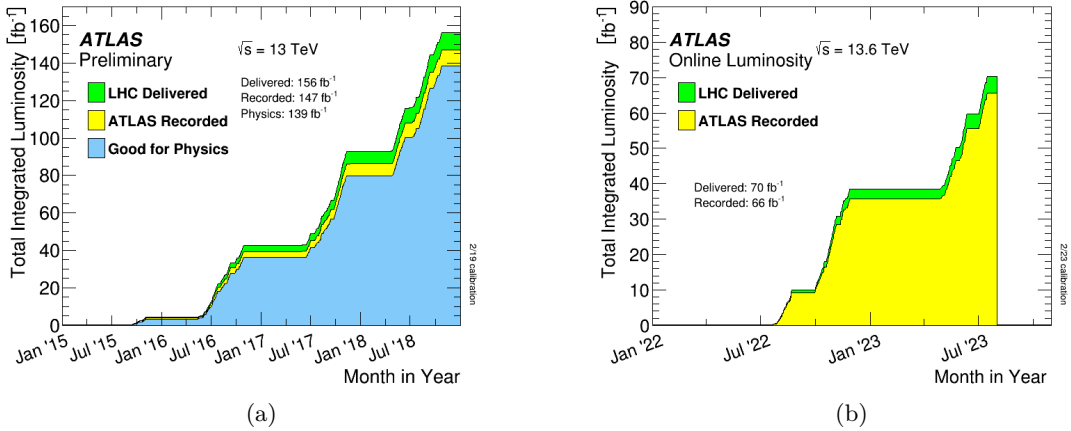


Figure 2.4. Total integrated luminosity vs. time delivered to (green) and recorded by (yellow) ATLAS during Run 2 (a) and Run 3 (b). Total luminosity in the Run 2 GRL is shown in blue, while the GRLs are not yet finalized for Run 3 as data-taking is in progress at the time of this writing. From [47] and [48].

The approximately 10^{11} total protons per bunch crossing previously discussed is necessary to achieve these high target luminosities, following naturally from Equation 2.2. However, this does not come without repercussions, as multiple pp interactions co-occur during each filled bunch crossing. Of these interactions, only the primary, inelastic **hard-scatter** collision produces physics of

¹⁰Several special dedicated detectors, ALFA and LUCID, are located forward of ATLAS to measure the delivered luminosity with high precision. ALFA is located 240 m downstream from the ATLAS IP, allowing it to be inserted extremely close to the beam-pipe and measure elastic pp scattering down to minimal angles to determine absolute luminosity. In contrast, LUCID is inserted 17m from the IP and measures “online” bunch-by-bunch luminosity and number of interactions per bunch crossing.

interest. The ancillary elastic, soft interactions are **pileup** and serve as a source of contamination from unwanted signals in a given event. Maintaining efficiency and purity in a high-pileup environment while conserving computational resources represents a significant challenge for physics object reconstruction, discussed further in Section 2.3. Pileup is determined using the average number of interactions per bunch crossing, $\langle\mu\rangle$, over a given **luminosity block (lb)**, a small time interval in data-taking where instantaneous luminosity is roughly constant.¹¹ The average value of $\langle\mu\rangle$ during Run 2 was $\langle\mu\rangle = 33.7$, while the increased luminosity during Run 3 has resulted in a higher average pileup of $\langle\mu\rangle = 46.5$ during 2022 and 2023 data-taking. The luminosity-weighted distributions of $\langle\mu\rangle$ for both Runs 2 and 3 are shown in figure 2.5.

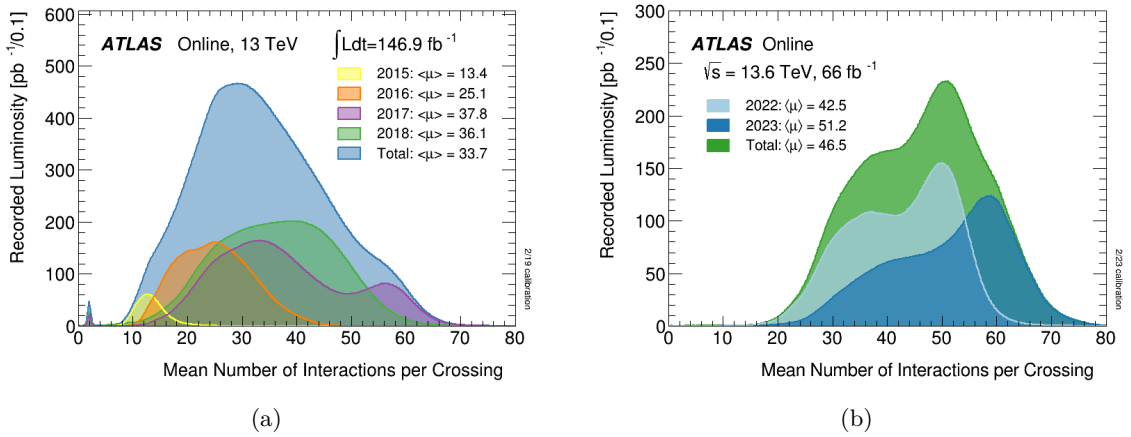


Figure 2.5. Luminosity-weighted distributions of the mean number of interactions per bunch crossing for Run 2 (a) and Run 3 (b). Average values for each individual datataking year and summed totals are shown. From [47] and [48].

2.1.4 LHC Run Schedule and Operations

As mentioned throughout this work, long-term LHC datataking periods are separated into overarching **Runs**. These Runs are separated by **Long Shutdown (LS)** periods where upgrades and maintenance are performed on the LHC and the individual detectors.¹² The LHC first began opera-

¹¹Technically, the true pileup value is $\langle\mu\rangle - 1$, as the total number of average interactions per bunch crossing also includes the hard-scatter, but pileup is often referred to as just $\langle\mu\rangle$ for simplicity.

¹²During the winter holiday period each datataking year, there is also a short Year End Technical Stop (YETS) where smaller-scale maintenance is performed.

tion in late 2008 and achieved its first collisions in March 2010¹³, reaching a center-of-mass energy of $\sqrt{s} = 7$ TeV. **Run 1** would continue up to 2012 and the center-of-mass energy was increased midway to $\sqrt{s} = 8$ TeV. This period recorded the data that would lead to the Higgs discovery by ATLAS and CMS and was further analyzed during the LS1 period, where the magnet system was upgraded to support close to the design beam energies. **Run 2** subsequently began in 2015 and continued through 2018 at $\sqrt{s} = 13$ TeV. The WH resonance search described in this thesis analyzes the full ATLAS Run 2 dataset and was performed through the whole of LS2 and into the beginning of **Run 3**, which began in mid-2022. During LS2, installation of the first phase of upgrades to prepare for the **high-luminosity LHC (HL-LHC)** began, which is scheduled to begin datataking in 2029 and is planned to record up to 4000fb^{-1} of data at pileup values up to $\langle\mu\rangle = 200$ onwards through 2040. Smaller upgrades were also performed during LS2 in advance of Run 3, on a shorter timescale, where the increased instantaneous luminosity and larger center-of-mass energy of $\sqrt{s} = 13.6$ TeV also required optimization of ATLAS reconstruction software to handle the increased volume of data and subsequent strain on computing resources. The ATLAS track reconstruction optimization described in this work was also performed during this period and evaluated on both Run 2 data reprocessed with the optimized configuration, used to develop and tune the optimizations, and Run 3 data, used to demonstrate the “production” performance for publication. Figure 2.6 shows the currently scheduled LHC and HL-LHC run plans.

Each datataking period is separated into many **fills**, where after luminosity drops below a threshold during collisions, the existing beams are dumped, and the injection chain is “filled” with a fresh batch of protons. This is done when the downtime needed to load a new fill is smaller than the loss of luminosity from the reduced collision rate due to depleted RF buckets. In ATLAS, portions of fills are assigned a **run number** and further split into the aforementioned luminosity blocks. As ATLAS is a highly complex machine with thousands of interdependent components, detector malfunctions or periods of sub-optimal performance occasionally occur. Subsequently, a small fraction of data recorded by ATLAS is not fit for physics analysis. The ATLAS **Data Quality** team evaluates each run for each subdetector and assigns **defects** at the per-lb level based on observed anomalies or areas of degraded performance. So-called **tolerable defects** are minor problems, such as an inefficiency in a single Pixel module, where data is still fit for physics analysis. Conversely, **intolerable defects**

¹³First collisions were delayed so long from startup because shortly after beginning operations, a faulty electrical connection led to a magnet quench. The warm-up of the sector caused six tons of superfluid helium to become gas, expand, and burst the 10-ton magnets away from their mountings, causing massive damage. Subsequent repairs and safety checks would take another 14 months to complete.

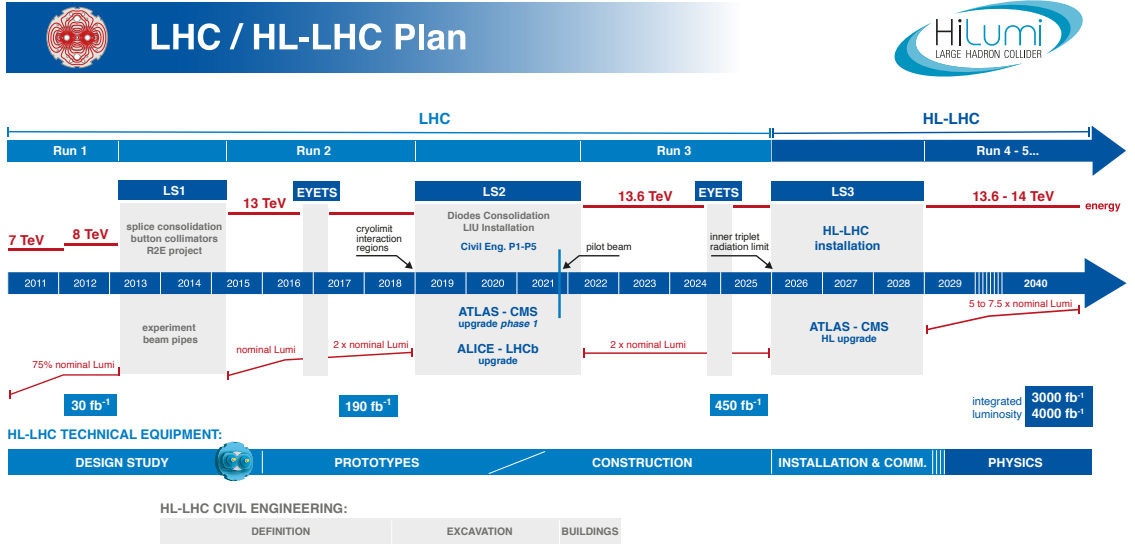


Figure 2.6. Planned run schedule for the LHC and HL-LHC, updated in 2022.

are more significant problems that significantly reduce the data quality. All runs and lbs signed off with only tolerable defects are used to create the GRL, which is the list of data considered suitable for physics analysis and used by all analysis teams.

2.2 The ATLAS Detector

Proposed in 1994 and constructed from 2003 to 2008, the ATLAS experiment is the largest general-purpose particle detector ever built, spanning 44m in length and 25m in diameter. It covers nearly the full solid angle around the central interaction point and consists of three concentric **subdetectors** oriented coaxially around the beampipe. The cylindrical detector volume is divided into three main regions: the central **barrel** region and two disk-shaped **endcaps**. This arrangement, shown in figure 2.7, is forward-backward symmetric along the beamline and azimuthally symmetric around it; each subdetector in this geometry fills a specific role in detecting particle signatures corresponding to its detector technologies and radial distance from the IP. Going outwards from the beamline, the **Inner Detector (ID)** measures the trajectories of charged particles and the position of primary and secondary vertices, the **Electromagnetic Calorimeter (ECAL)** and **Hadronic Calorimeter (HCAL)** measure the energy of electromagnetic and hadronic physics objects, and the **Muon Spectrometer (MS)** identifies muons and additionally measures their momentum. This

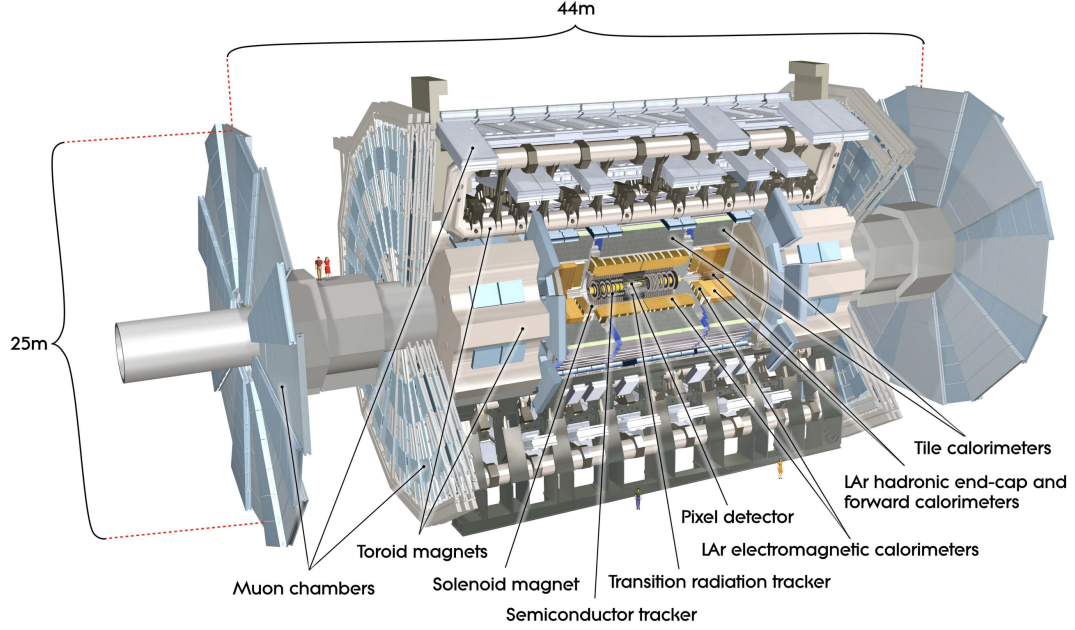


Figure 2.7. Schematic of ATLAS and its subdetectors with a cut-away view. Persons are shown for scale. Taken from [34].

section will present a detailed overview of each subdetector alongside other key ATLAS systems and concepts.

2.2.1 The ATLAS Coordinate System

ATLAS uses a right-handed coordinate system whose origin is the nominal interaction point, close to the center of the detector. In Cartesian coordinates, the positive x , y , and z axes point horizontally towards the center of the LHC ring, upwards towards the surface, and out along the beamline, respectively. Polar coordinates are defined in the standard transformation:

$$r = \sqrt{x^2 + y^2}, \quad (2.4)$$

$$\phi = \arctan\left(\frac{y}{x}\right), \quad (2.5)$$

$$\theta = \arctan\left(\frac{z}{\sqrt{x^2 + y^2}}\right). \quad (2.6)$$

The azimuthal angle ϕ is measured around the beamline in the xy plane, where $\phi = 0$ points along the x axis. Meanwhile, the spherical polar angle θ is measured relative to the beamline where $\theta = 0$

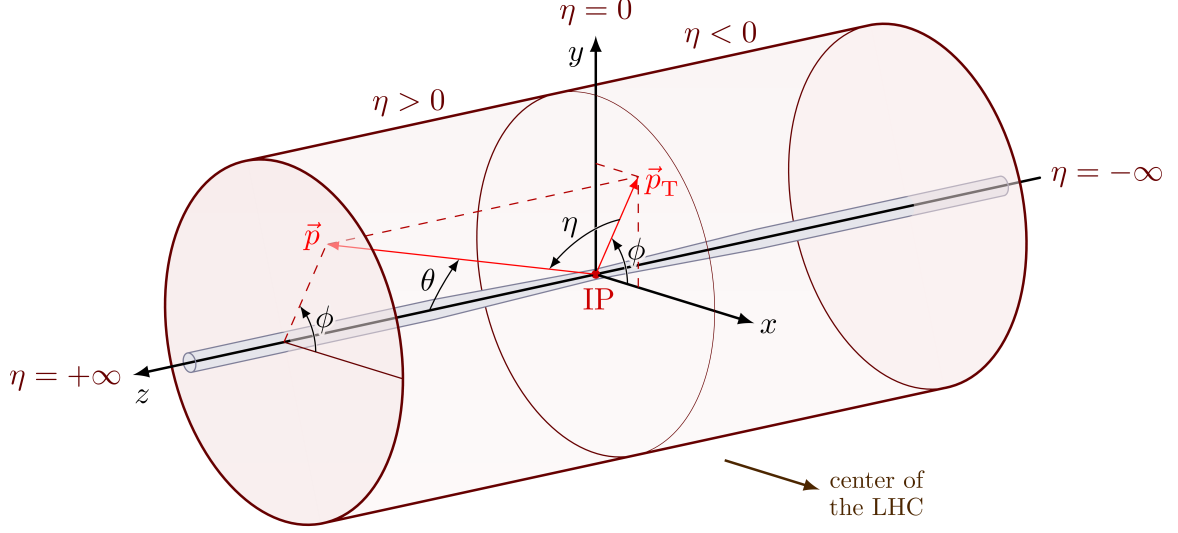


Figure 2.8. Diagram of the ATLAS coordinate system with an example particle of momentum $\vec{p}$. Modified from [49].

points along the z axis. However, the true polar angle is a less useful quantity in collider physics than the **pseudorapidity**, which is Lorentz invariant under boosts for massless particles. The **rapidity** y of a particle is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (2.7)$$

where E is the energy of the particle and p_z is its momentum in the z direction. In the ultrarelativistic limit, the rapidity approaches the pseudorapidity η , which can be expressed purely in terms of the polar angle:

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (2.8)$$

Pseudorapidity of $\eta = 0$ corresponds to a particle perpendicular to the beamline in the transverse plane, while η approaches ∞ parallel to the beamline. Higher values of pseudorapidity correspond to the “**forward**” region of the detector.¹⁴ The angular separation between objects can thus be expressed as a Lorentz invariant geometric quantity:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (2.9)$$

¹⁴The Inner Detector only has η coverage up to $\eta \approx 2.5$, limiting full reconstruction of physics objects forward of this value, but the calorimeters have extended coverage up to $\eta = 4.9$. The Inner Tracker (ITk) upgrade, which will replace the Inner Detector in the HL-LHC era, will have coverage up to $\eta = 4$ owing to improvements in silicon detector design and radiation hardness.

Quantities such as momenta and energies are typically measured with respect to the transverse plane, as the initial **transverse momentum** of the colliding system along the beamline is known to sum to zero, and thus, any transverse quantity is from a physics process produced in the collision rather than from the initial colliding parton. Transverse momentum p_T and transverse energy E_t are defined as:

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad (2.10)$$

$$E_T = E \sin \theta. \quad (2.11)$$

Taking this together, particles are usually described in terms of (p_T, η, ϕ) , as shown in figure 2.8. Combined with either the mass or energy, these quantities construct the standard four-vector notation used in ATLAS.

2.2.2 Magnet System

A powerful magnet system is a central component of any general-purpose particle detector, as the curvature of the trajectories of charged particles in an intense magnetic field allows their momenta to be measured with high precision. The ATLAS system is composed of two types of magnets: a central **solenoid** and three **toroidal** magnets through the outer barrel and endcaps. Constructed from superconducting niobium-titanium, the magnets are cooled to 4.7K by liquid helium to maintain superconductivity. The 5.3 m long, 2.4 m wide, 4.5cm thick solenoid surrounds the Inner Detector and supplies it with a near-constant 2T axial magnetic field; it is also designed to be as thin and lightweight as possible to minimize the total number of undesirable material interactions. In contrast, the three toroid magnets, each consisting of eight air-core coils arranged radially symmetrically, supply the magnetic field to the Muon Spectrometer. The barrel toroid is 25.3m long and 20.1m in diameter, surrounding the MS in the barrel region and providing an approximately 0.5 T field. The two endcap toroids are 5 meters long and have a diameter of 10.7 meters. They provide approximately 1 T and are inserted half-periodically to the barrel coils, as seen in figure 2.9, to minimize the drop in field strength in the transition region.¹⁵

¹⁵In contrast to the solenoid, the magnetic field strength of the much larger toroid magnets is highly inhomogeneous. Dedicated sensors around the MS measure the magnetic field strength to create a detailed “map” to be used in particle reconstruction.

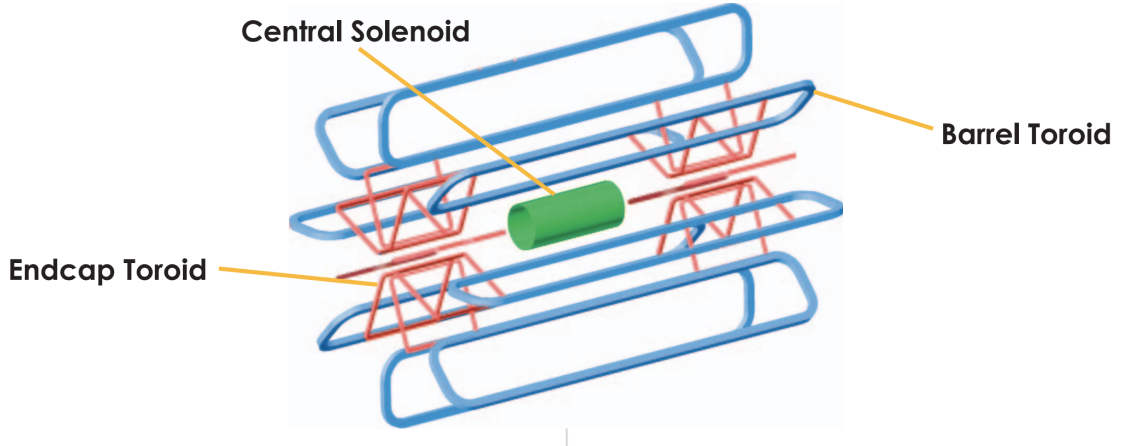


Figure 2.9. Representational layout of the ATLAS magnet system. Taken from [50].

2.2.3 The Inner Detector

Incident particles first encounter the Inner Detector, the innermost ATLAS subsystem. Located just 33 mm from the beamline and extending out to $r = 1.1$ m, the ID is a high-granularity system designed to capture the overlapping trajectories of hundreds of particles. It extends across $|z| < 3.5$, covering up to $|\eta| < 2.5$, and consists of three separate sensor technologies, each designed to maximize performance at their respective r range. As a charged particle moves through the ID, it curves in an azimuthal trajectory under the solenoid’s magnetic field, allowing measurement of its charge and momentum. The energy depositions, or “**hits**”, in the detector elements are used to construct trajectories, known as “**tracks**”, which serve as the foundational element of the majority of ATLAS physics objects used in analyses, including the determination of the hard-scatter primary vertex. Their reconstruction is described in detail in Section 2.3.1. The three ID subsystems include the silicon **pixel detector**¹⁶, the silicon-microstrip **semiconductor tracker (SCT)**, and the gas straw-tube **transition radiation tracker (TRT)**. Mirroring the general ATLAS geometry, each subsystem is additionally divided into barrel and endcap regions with concentric sensor layers. Figure 2.10 shows a render of the overall subsystem arrangement and a schematic view is shown in figure 2.11.

¹⁶Including the **insertable b-layer (IBL)**

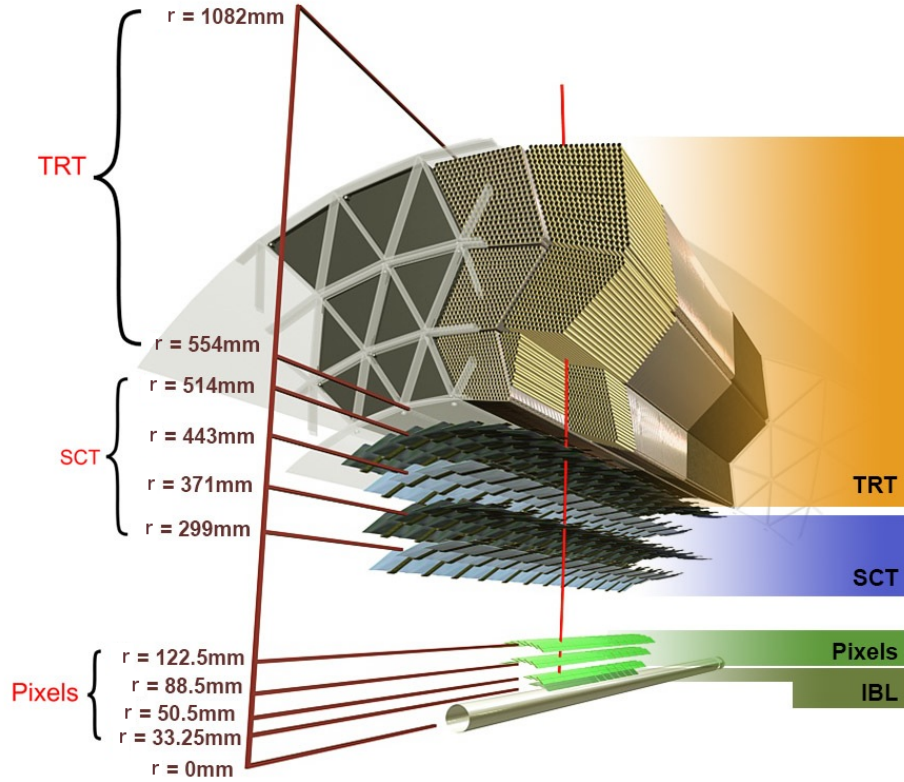


Figure 2.10. Cutaway render of the ATLAS Inner Detector with an example high- p_T particle trajectory. Taken from [51].

2.2.3.1 Pixel Detector

The innermost portion of the ID, the pixel detector, is comprised of four¹⁷ concentric layers of identical modules, shown in figure 2.12, each composed of silicon semiconductor sensors and readout electronics. Semiconductor sensors produce a signal when a charged particle passes through their sensing material and excites electrons to the conduction band, creating “holes” in the valence band. When a high bias voltage is applied, the charge pairs are pulled to the corresponding electrodes, and the readout electronics measure the resulting current pulse, forming a hit. True to its name, the pixel detector uses rectangular “**pixel**” sensing elements that together give a three-dimensional position measurement.

The first layer of the pixel detector, closest to the interaction point, is the **Insertable B-Layer (IBL)**, which was installed between Run 1 and Run 2 to improve impact parameter and vertex resolution at higher pileup as well as help protect the original first pixel layer, the **B-layer**, from

¹⁷Technically three in the endcap, as the IBL is only present in the barrel region.

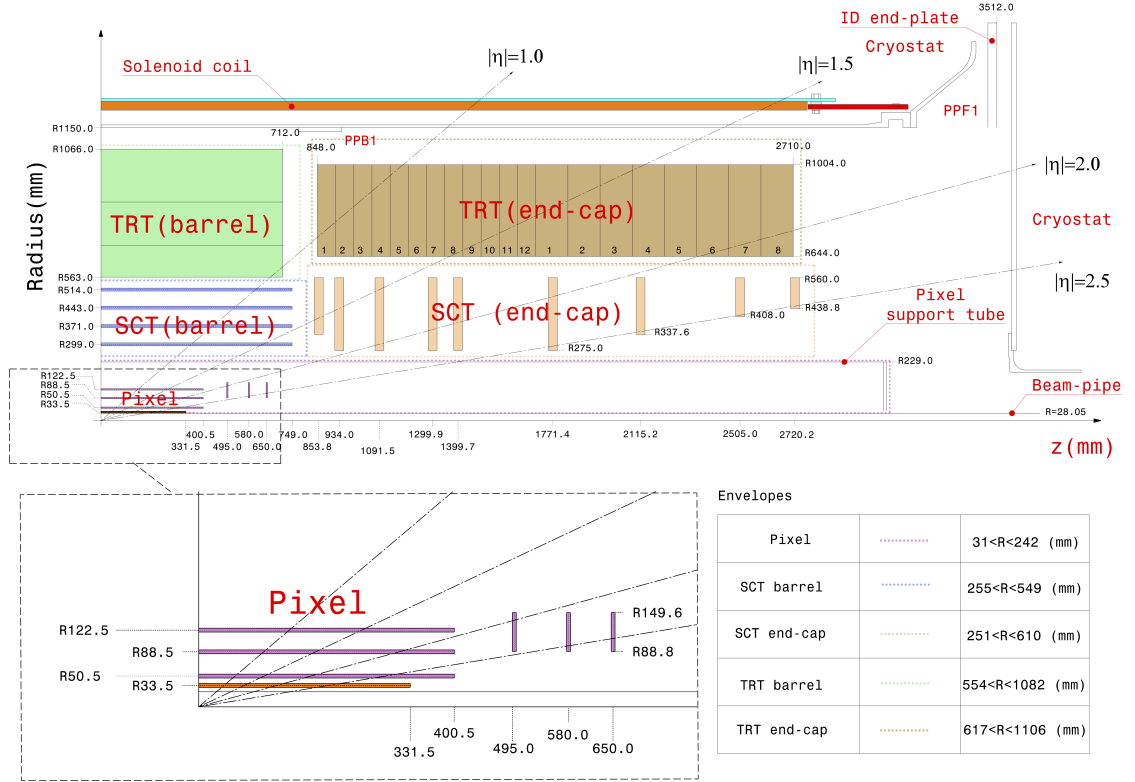


Figure 2.11. Schematic of the ATLAS Inner Detector in the r - z plane. The lower left shows an enlargement of the pixel detector, and the lower right shows the radial coverage of the three detector subsystems. Taken from [52].

radiation damage. The IBL is exceptionally granular, consisting of 12 million $50 \times 250 \mu\text{m}^2$ pixels that give a spatial resolution of $8 \mu\text{m}$ in r - ϕ and $40 \mu\text{m}$ in z .

The remaining three original pixel barrel layers surrounding the IBL are located 50.5, 88.5, and 122.5 mm away from the beam pipe. Each of these layers is roughly 4 meters long and together comprises a total of 1744 modules, each containing 46,080 $50 \times 400 \mu\text{m}^2$ pixel elements, giving a spatial resolution of $10 \mu\text{m}$ in r - ϕ and $115 \mu\text{m}$ in z .¹⁸ The barrel layers contain 1456 modules and 67 million pixels, while the remaining 288 modules, with a total of 13.2 million pixels, are located in the endcaps.

2.2.3.2 Semiconductor Tracker (SCT)

Surrounding the pixel detector is the SCT, which also uses silicon semiconductor sensors to detect charged particles using the same physical mechanism. Unlike the pixel detector, however,

¹⁸The larger pixel size and coarser granularity compared to the IBL is sufficient for distances more than 5cm from the beamline.

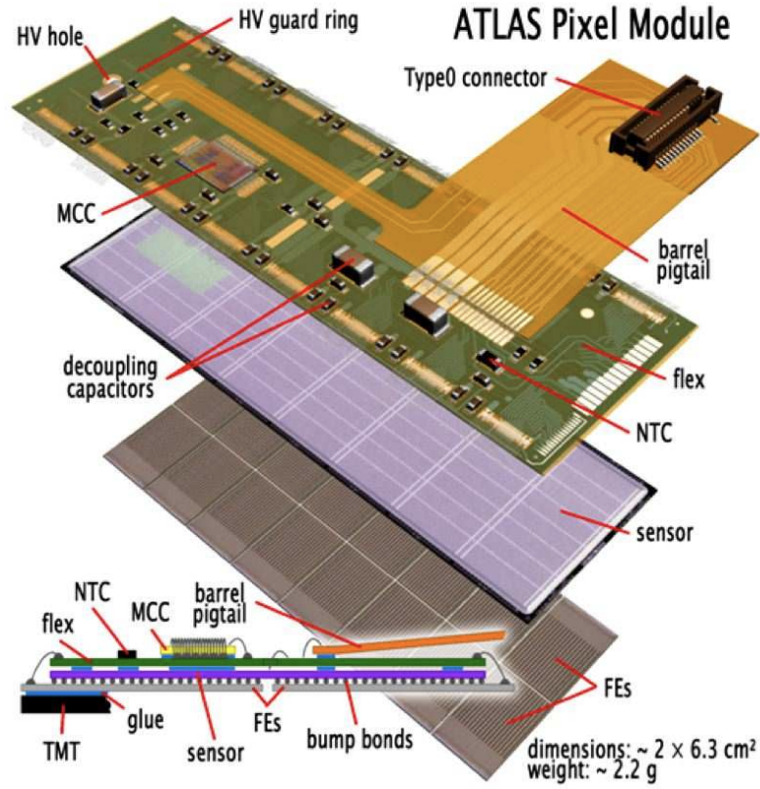


Figure 2.12. Diagram of an ATLAS pixel module, showing the silicon pixel sensors alongside the readout and support electronics. Taken from [34].

these sensors are arranged in silicon microstrips rather than individual pixels. Microstrip detectors are made up of numerous identical silicon strips arranged along a single axis of the detector structure; for the SCT, a total of 4088 modules cover four concentric cylindrical layers in the barrel and nine planar discs in the endcaps. In the barrel, these are located 299, 371, 443, and 514 mm from the beampipe. Each module contains two 12.8 cm strips that are glued together in pairs with a 40 mrad stereo offset, shown in figure 2.13, allowing for a two-dimensional position measurement when combined. The modules are rectangular with an $80 \mu\text{m}$ pitch in the barrel, where they are parallel to the beam axis, and “fan-shaped” with a variable pitch of $57\text{--}94 \mu\text{m}$ in the endcaps, where they are oriented radially. Each module has 1536 readout channels for a total of 6.2 million channels that give a $17 \mu\text{m}$ resolution in $r\text{--}\phi$ and $580 \mu\text{m}$ in z .

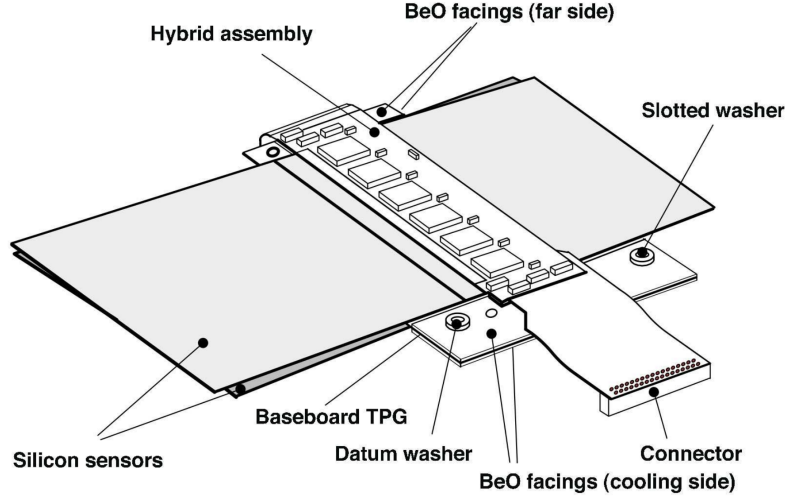


Figure 2.13. Diagram of an SCT barrel module. The stereo angle offset between the two silicon microstrip layers is visible alongside the readout and support electronics. Taken from [34].

2.2.3.3 Transition Radiation Tracker (TRT)

The outermost layer of the ID is the TRT, which uses gas drift-tube detector technology to improve charged particle identification, unlike the other two ID subsystems. It is composed of nearly 300,000 4mm wide Kapton straw tubes each filled with a 70% xenon (Xe), 27% carbon dioxide (CO₂), and 3% oxygen (O₂) gas mixture¹⁹, surrounding a gold-plated tungsten wire 31 microns in diameter. When a charged particle passes through a tube, it ionizes the gas, causing the newly free electrons to drift towards the 1.53 kV wire, as shown in figure 2.14. The resulting current is amplified and converted into a hit by the readout electronics.

Additionally, lying between the straws is a layer of radiator material with a large dielectric constant. When a particle traverses the boundary between the two media, it emits transition radiation. The Xe atoms reabsorb these x-ray photons, and the subsequent ionization produces a much stronger signal than that from a regular traversing particle, creating “**high threshold**” (HT) hits. A light particle, such as an electron, emits more photons at the same momentum than a heavier particle, such as a pion, which gives discrimination power in particle identification during reconstruction as electrons are much more likely than pions to have a high fraction of HT hits. The TRT also aids in

¹⁹The TRT has suffered some age-related issues related to frequent leakages of the (very expensive) Xe gas mixture, of up to 150L/day, in a handful of specific channels. These channels are now flushed with much cheaper argon gas as an alternative. Consequently, they, unfortunately, lose their transition radiation detection capabilities and associated electron-pion discrimination power as the Ar mixture does not reabsorb the emitted X-rays in the same fashion as the active Xe mixture [53].

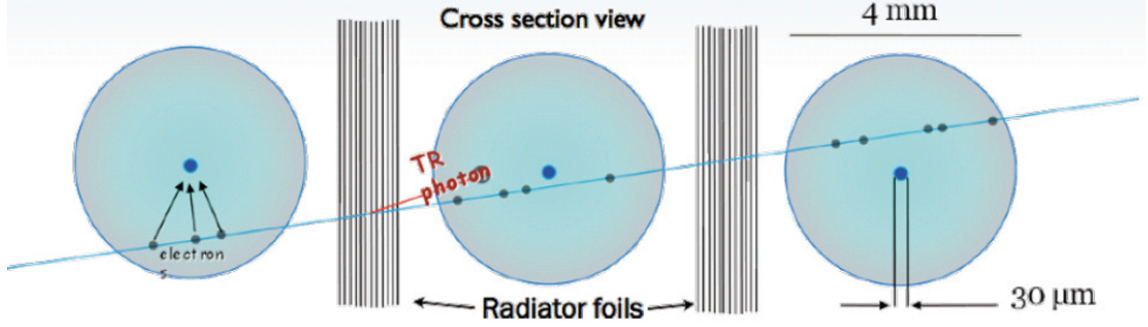


Figure 2.14. Generalized diagram of TRT drift tube operation, including transition radiation detection. Taken from [54].

identifying secondary electron tracks produced by photon conversions in the ID detector material, detailed in Section 2.3.1.7.

The TRT barrel region contains 52,544 1.46 m long tubes oriented parallel to the beamline, extending from 554 mm to 1.08m out from the IP. In contrast, the endcaps together contain 122,880 39cm long tubes oriented perpendicular to the beamline, extending out to $|z| < 2.7$ m. The TRT gives relatively poor resolution per measurement, providing only r - ϕ (barrel) and z - ϕ (endcap) measurements with a $120 \mu\text{m}$ resolution, but effectively covers a very large fiducial volume, which is sufficient to measure the curvature of tracks at a larger distance from the IP due to a large total number of TRT hits per track.

2.2.4 The Calorimeters

The ATLAS calorimeter system, shown in figure 2.15, is designed to completely absorb and measure the energy of all Standard Model particles except muons and neutrinos. It surrounds the ID, covering up to $|\eta| < 4.9$, and comprises two subsystems: the **Electromagnetic** and **Hadronic Calorimeters**. The electromagnetic calorimeter measures the energy of electrons and photons, while the hadronic calorimeter measures the energy of charged and neutral hadrons. They both use two different detector technologies, **liquid argon (LAr)** and **Tile** calorimeters, strategically deployed to maximize efficiency; within the $|\eta| < 2.5$ range of the ID, both technologies enable precision measurements of electrons and photons as well as central jets, while forward of this range they are used for forward jet reconstruction and determining missing transverse energy. Both technologies are **sampling calorimeters** that differ in their choice of passive and active materials. In this type of calorimeter, layers of dense passive, or absorbing, material are alternated with active material; as particles pass through the absorbing material, they initiate electromagnetic or hadronic showers

through multiple scattering. The active layers then measure the energy produced in the shower through ionization, and this process is repeated until all of the energy from the incident particle has been fully absorbed.²⁰ Calorimeters are characterized by the radiation length X_0 (ECAL) or interaction length λ (HCAL) of their active material, determining the total volumes needed before the energy is completely absorbed. Figure 2.16 shows a technical diagram of the calorimetry system and its constituents, detailed further in the following section.

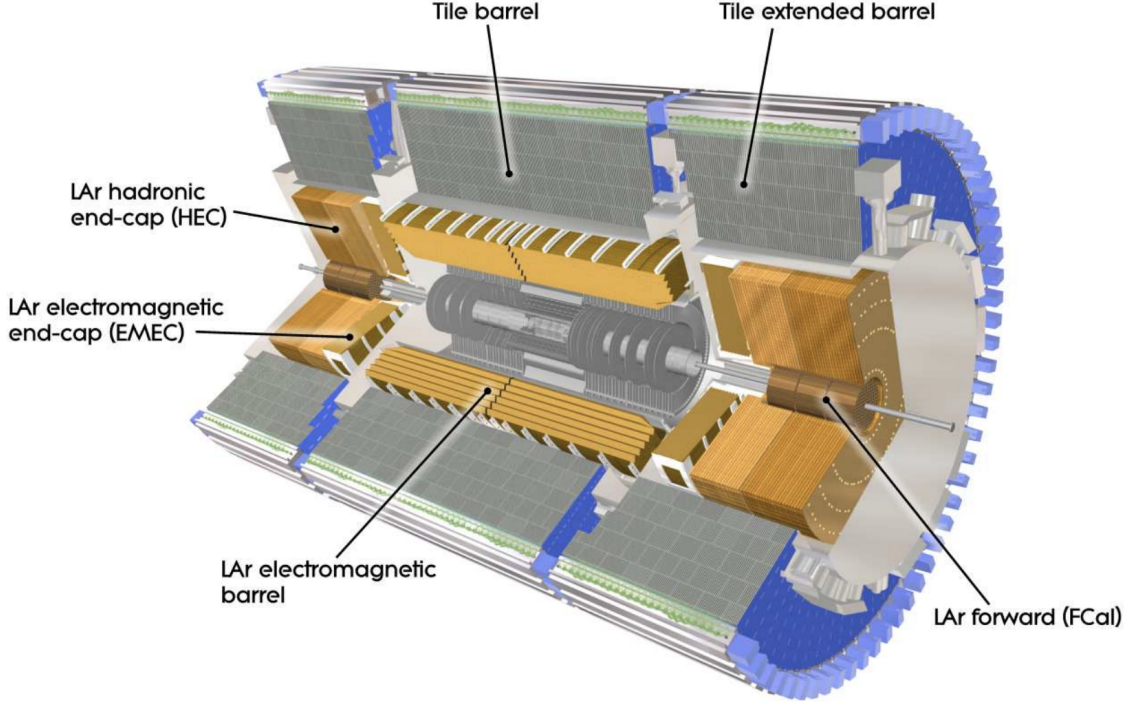


Figure 2.15. Cutaway render of the ATLAS calorimeter system showing the various subdetectors. Taken from [34].

2.2.4.1 EM Calorimeter

The EM Calorimeter is the interior layer of the calorimetry system, enveloping the ID and solenoid. It exclusively uses LAr calorimetry, where liquid argon is the active material and lead is the passive absorbing material. The barrel region, between $|\eta| < 1.475$ and $1.25 < r < 2.05$ m, consists

²⁰As energy is lost to the absorbing material, the calorimeter can only “sample” the particle energies. The total incident energy must be inferred by combining information from the presampler layer and statistically expected material losses, reducing resolution compared to a heterogeneous calorimeter. Hadronic calorimeters only exist as sampling calorimeters.

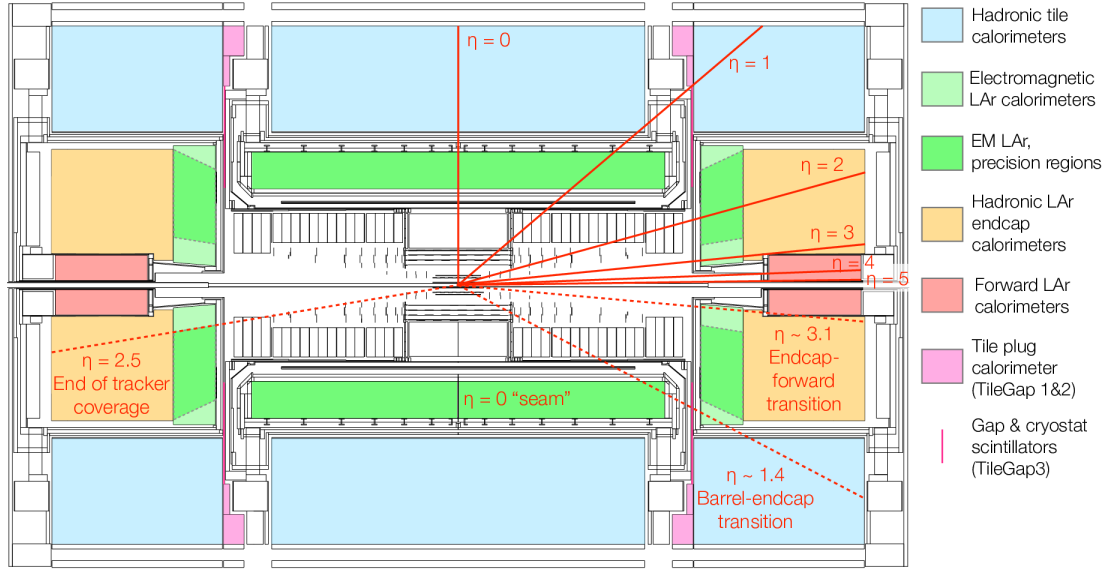


Figure 2.16. Technical layout of the ATLAS calorimeters with color-coded subsystems. The inner detector systems are also shown in black-and-white in the center, up to $\eta = 2.5$. Taken from [55].

of two half-barrel regions with 16 modules each, separated by a 4mm wide seam at $z=0$. Combined, the barrel has approximately 110,000 total readout channels. The endcaps are constructed from two coaxial wheels, each divided into eight modules, and cover between $1.375 < |\eta| < 3.2$. In the central “precision” region, between $|\eta| < 1.8$, a presampler layer corrects for the energy lost by electrons and photons before reaching the calorimeter modules. The presampler has an angular resolution of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$.

Each LAr module comprises three distinct layers with their active and passive volumes arranged in an “accordion” geometry, as shown in figure 2.17. This maximizes the total active volume without azimuthal gaps. The first layer closest to the IP has the finest granularity to provide precision measurements of photons and to resolve diphoton decays, with a granularity of $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$. It extends for a total of 4.3 radiation lengths X_0 before the second layer, which contains at least 16 X_0 depending on the location in η . Most of the incident energy is absorbed in this layer, with a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The remaining third layer, with a thickness of 2 X_0 , provides information to distinguish EM and hadronic showers and estimates any amount of “punch-through” energy that may escape the ECAL. The majority of cells in this layer have a resolution of $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$, but several special coarse cells in this layer are reserved for use in the

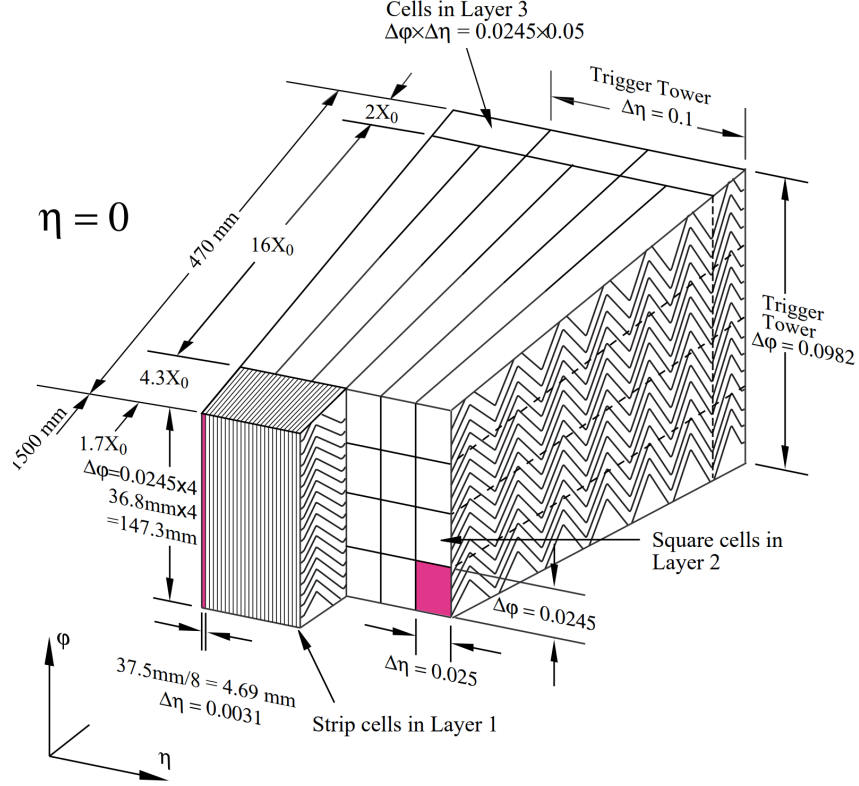


Figure 2.17. Diagram of a LAr barrel module showing the accordion geometry of the absorber plates and readout panels. It also shows the granularity of cells in η and ϕ across each of the three layers and trigger towers. Taken from [34].

trigger system (detailed further in Section 2.2.6) in so-called “**trigger towers**”, with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$.

2.2.4.2 Hadronic Calorimeter

The Hadronic Calorimeter forms the exterior layer of the calorimetry system. Unlike the ECAL, which only interacts with charged particles and photons²¹, the HCAL induces hadronic showers through incident particles interacting with the absorber nuclei via the strong interaction. It uses Tile calorimetry before $|\eta| < 1.7$ and LAr calorimetry in the forward regions up to $|\eta| < 4.9$.

The **Tile calorimeter**, used in the central regions, uses plastic scintillator tiles as the active material and steel tiles as the absorbing material in a volume ratio of 4.7 : 1. The hadronic showers, produced by incident particles interacting with the nuclei in the absorber, then hit the scintillating tiles and produce high-energy photons. These photons interact with wavelength-shifting fibers, which

²¹Charged hadrons, such as pions, do interact with the EM calorimeter, but are not wholly absorbed by it.

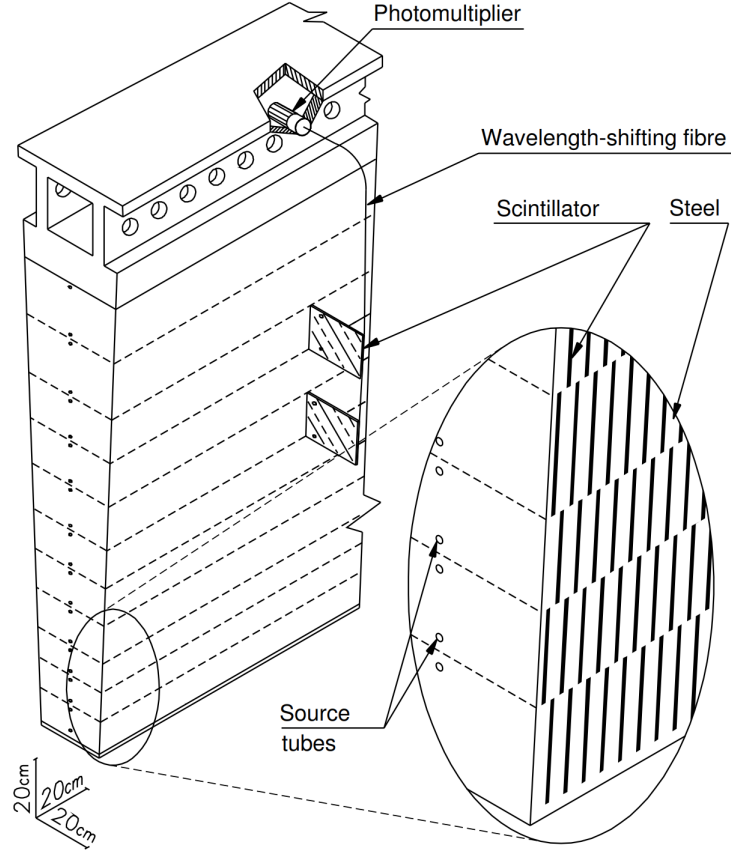


Figure 2.18. Diagram of a tile calorimeter module, showing the active and passive layers and the readout mechanism with wavelength-shifting fibers and photomultiplier tubes. “Source tubes” are used in the calibration system. Taken from [34].

re-emit photons of a lower wavelength that are subsequently read out as a signal using photomultiplier tubes (PMTs), shown in figure 2.18.

It contains a 5.8 m long central barrel region between $|\eta| < 1.0$ and two 2.6 m long extended barrel regions prior to the endcaps, between $0.8 < |\eta| < 1.7$. Each region contains 64 modules, each separated into three layers with analogous functions to those in the ECAL, corresponding to approximately 10 total radiation lengths λ . Additionally, the Tile calorimeter, broadly, has a coarser resolution than the LAr-based ECAL, as hadronic showers are topologically wider than EM showers. The first two layers have a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and the coarser third layer has a granularity of $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$.

The **Hadronic Endcap Calorimeter (HEC)** extends the $|\eta|$ coverage of the Tile calorimeter into the forward regions, using LAr calorimetry with copper absorbers (rather than lead, as in the ECAL). It covers between $1.5 < |\eta| < 3.2$ and contains two disks per endcap with 32 modules each.

In the extreme forward region, between $3.2 < |\eta| < 4.9$, the **Foward Calorimeter (FCal)** takes over, which also uses LAr calorimetry. Extending out to this region leaves only 1.3 degrees uncovered by the calorimeter system, allowing for a nearly complete measurement of missing transverse momentum. However, the increased radiation exposure in the region very close to the beamline imposes additional constraints to handle the high interaction rates. The FCal is composed of three layers, the first of which is an EM layer that uses a copper absorber similar to the HEC. The second and third layers, the hadronic layers, instead use tungsten as it is more radiation-hard than copper or steel. Its granularity ranges from $\Delta\eta \times \Delta\phi = 0.15 \times 0.15$ to 0.3×0.3 , varying with increasing $|\eta|$.

2.2.5 The Muon Spectrometer

Muons are highly penetrating due to their high mass, which causes Bremsstrahlung radiation to be highly suppressed compared to an electron or other light charged particle. As such, they leave minimally ionizing deposits in the calorimeters and escape their fiducial volumes nearly unimpeded, giving their characteristically “clean” detector signature. The **Muon Spectrometer**, shown in figure 2.19, surrounds the calorimeters and is explicitly designed to identify muons and give an additional measure of their momenta by using the toroid magnets to bend their trajectories under an orthogonal magnetic field. The MS is composed of four types of gaseous subdetectors: **Monitored Drift Tubes (MDTs)** [56], **Cathode Strip Chambers (CSCs)** [57], **Resistive Plate Chambers (RPCs)** [58], and **Thin Gap Chambers (TGCs)** [59].

The MDTs are the primary precision muon tracking system, while the CSCs provide additional precision tracking in the high- η region. In contrast, the RPCs and TGCs give a fast, coarse measurement and compose the muon trigger system, where the precision timing allows for correct assignment to individual bunch crossings. They also provide an additional measurement to determine the azimuthal coordinate of a muon segment, which is orthogonal to the bending plane and the arrangement of the precision MDTs. For Run-3, the inner endcap layers were replaced with the **New Small Wheel (NSW)**, which uses new MicroMegas (MM) gas detector and small-strip Thin Gap Chamber (sTGCs) technologies to provide *both* precision tracking and triggering capabilities in the forward regions, in preparation for the HL-LHC.²² A detailed schematic of the MS is shown in Figure 2.20.

²²We won’t go in-depth on the NSW here, as the physics analysis presented in this work uses only Run 2 data and the tracking performance work explicitly excludes the NSW from comparisons involving muons, as to provide a like-to-like comparison for computational metrics. Additionally, the NSW muon reconstruction was not yet finalized at the time of evaluation.

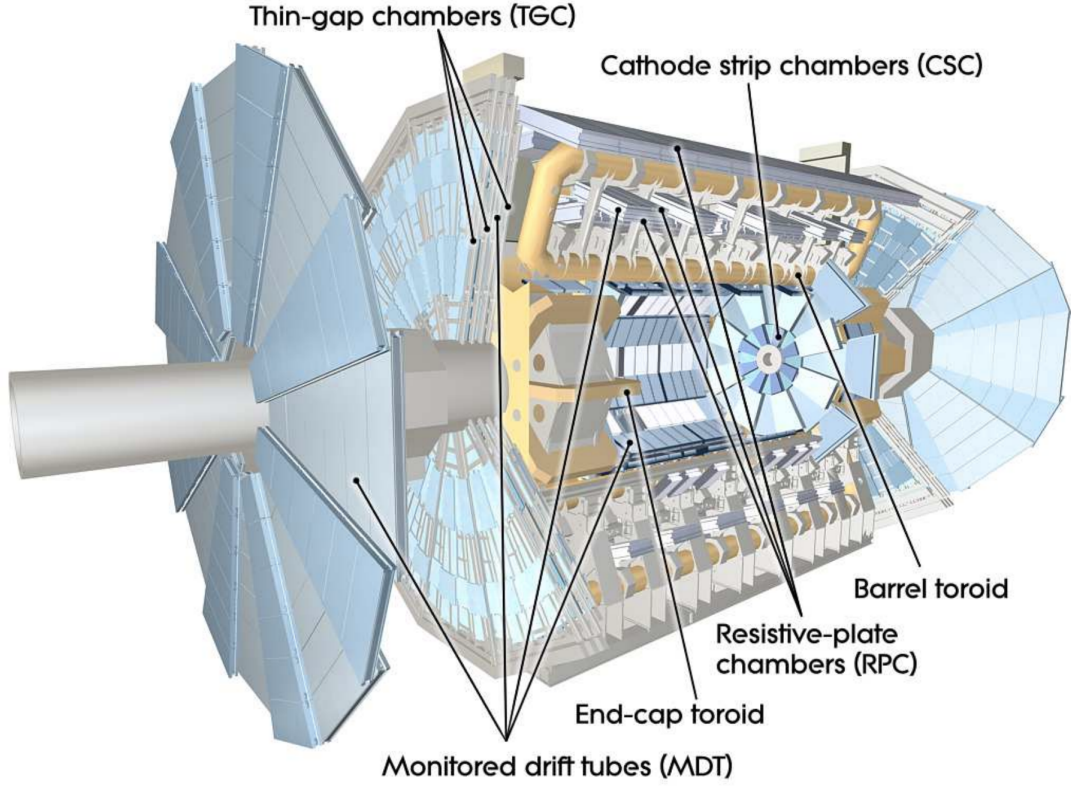


Figure 2.19. Cutaway render of the ATLAS Muon Spectrometer (MS) showing the subsystems and toroid magnets. Taken from [34].

2.2.5.1 Monitored Drift Tubes (MDTs)

MDTs give the bulk of precision muon tracking measurements in ATLAS, covering a range of up to $|\eta| < 2.7$, except in the innermost layer of the endcaps, where they cover only up to $|\eta| < 2.0$. A total of 1150 MDT chambers with approximately 354,000 total drift tube channels are stationed across the MS. They are arranged in three concentric, cylindrical layers in the barrel at $r = 5\text{ m}$, 7.5 m , and 10 m from the beam axis, while the endcap MDT modules are arranged in four layers, including one “extra” layer located outside the outer toroid magnet in r , at $|z| = 7.4\text{ m}$, 10.8 m , 14 m , and 21.5 m from the nominal interaction point.²³ Chambers in the barrel region are positioned around and between the eight coils of the barrel toroid, while those in the endcap regions are situated both in front of and behind the two endcap toroid magnets. The toroids’ ϕ symmetry is mirrored in the symmetric design of the muon chamber system, which is divided into eight octants. Each of

²³The additional layer provides a third measurement for transition-region muons that lie before the $|\eta|$ range of the outer wheel but are beyond the barrel.

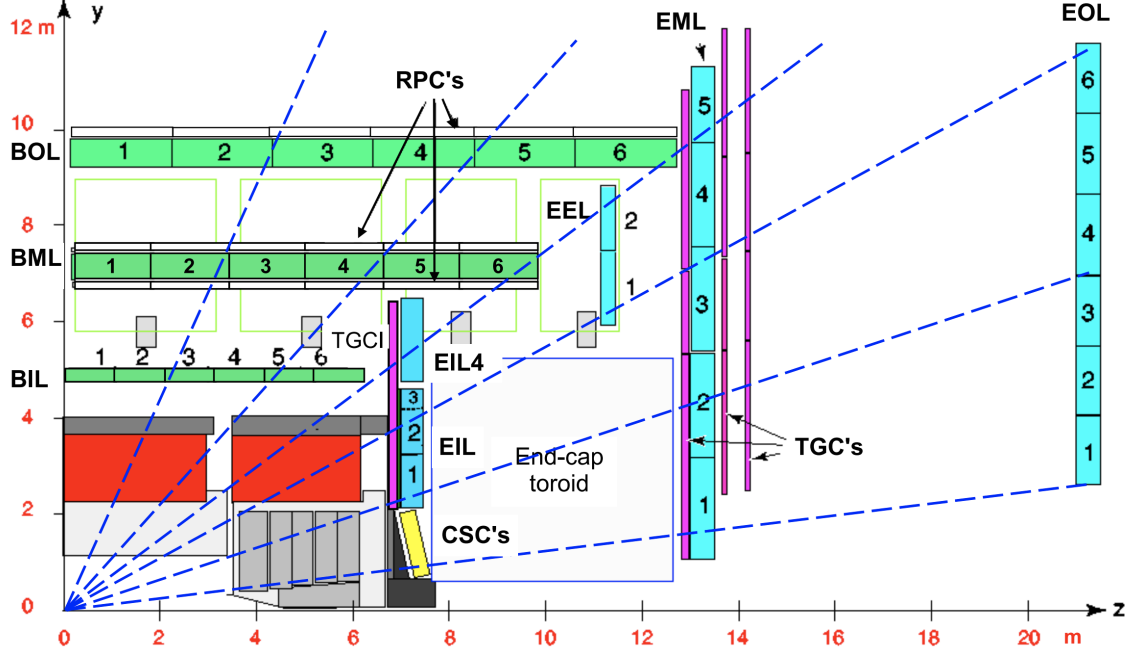


Figure 2.20. A cross-section schematic of a quadrant of the ATLAS Muon Spectrometer. RPCs, CSCs, and TGCs are labeled, while MDTs are shown with a 3-letter encoding based on whether they are a barrel or endcap (B or E) module, an inner, middle, outer, or “extra” module (I,M,O,E), with the last letter denoting that these are large-sector modules (L). Small-sector modules are not shown in the cross-section as they are shifted azimuthally by $2\pi/6$. Taken from [34].

these octants is further split azimuthally into two sectors of differing widths—a larger and a smaller sector—creating overlapping areas in ϕ for a total of 16 sectors. This overlap at the chamber edges reduces gaps in the detector’s coverage and facilitates the alignment of adjacent sectors by utilizing the muon tracks recorded in both a large and a small chamber.

Individual drift tubes operate under similar principles as those described previously for the TRT, with some design differences in the choice of materials. The 3 cm wide aluminum tubes are filled with a 97% argon and 3% CO_2 gas mixture at a pressure of 3 bar, surrounding a gold-plated tungsten-rhenium anode wire kept at a bias voltage of 3kV. They achieve an average resolution of $80 \mu\text{m}$ per tube or approximately $35 \mu\text{m}$ per chamber. Each chamber consists of two “multi-layers” composed of three (in the barrel) or four (in the endcaps) drift tubes separated by a spacer, as shown in Figure 2.21. The multiple layers allow for a precise measurement of a muon’s momentum by determining the sagitta of the muon track as it passes through the chamber.

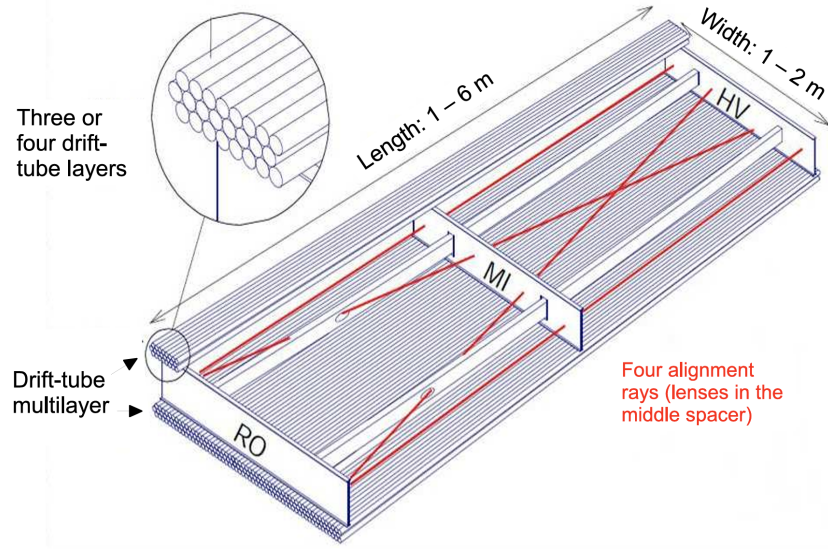


Figure 2.21. Diagram of the mechanical structure of an MDT chamber. Each chamber consists of an aluminum space frame formed by three spacer bars linked by longitudinal beams, which support two multi-layers comprising three or four layers of drift tubes. Four optical alignment rays—two parallel and two diagonal—are used to monitor the internal geometry of the chamber. The acronyms RO and HV indicate the positions of the readout electronics and high voltage supplies, respectively. Taken from [34].

2.2.5.2 Cathode Strip Chambers (CSCs)

Complementing the MDTs are the CSCs, providing additional coverage in the forward region between $2.0 < |\eta| < 2.7$. The maximum safe operational limit for MDTs is at counting rates of about 150 Hz/cm^2 , but this is surpassed in the endcap's innermost layer. In this specific region, the MDTs are substituted with CSCs, which offer superior spatial and temporal resolution while sustaining hit rates of up to 1 kHz/cm^2 . Like the MDTs, the CSCs are divided into large and small chambers in ϕ . The CSC system comprises two disks per endcap, each with eight chambers, for a total of 32 chambers with 31,000 readout channels. Each chamber houses four CSC planes, allowing four independent measurements along each track in both η and ϕ .

The CSCs are multiwire proportional chambers with radially oriented wires. These operate under a similar principle as a drift tube, with a 30-micron wide anode wire kept at 1.9 kV inside an 80% argon / 20% CO_2 gas mixture. The cathodes are segmented into strips, with one set perpendicular to the wires for precise measurements and the other parallel for transverse measurements, as shown in Figure 2.22. Track positions are determined by interpolating the charges induced from the drifting electrons on adjacent cathode strips, providing a resolution of $60 \mu\text{m}$ per plane in the bending direction and 5 mm in the non-bending direction.

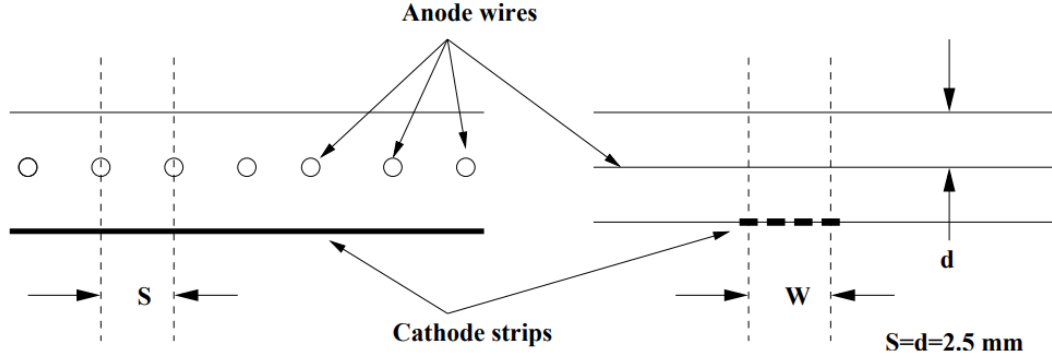


Figure 2.22. Diagram of a CSC, shown in both a view along the anode wires (left) and a view along the cathode strips (right). Taken from [34].

2.2.5.3 Resistive Plate Chambers (RPCs)

In the barrel region between $|\eta| < 1.05$, the muon trigger system uses three concentric cylindrical layers, or trigger stations, containing 606 RPCs with a combined total of 373,000 readout channels. RPCs are gaseous detectors that operate without wires using a parallel-plate configuration. It comprises two resistive plates made of phenolic-melaminic plastic laminate, maintained parallel at a 2mm gap by insulating spacers. The area between the plates is filled with a gas mixture of 94.7% tetrafluoroethane ($C_2H_2F_4$), 5% isobutane (Iso- C_4H_{10}), and 0.3% sulfur hexafluoride (SF_6), which is chosen for its relatively low operating voltage, non-flammability, and cost-effectiveness. A 9.8kV electric field across the plates enables the formation of avalanches along ionizing muon tracks heading toward the anode. Signals are then captured through capacitive coupling to metallic readout strips attached to the outer surfaces of the resistive plates, giving a fast signal response width of 5ns. The RPCs are attached to the outside of the MDT chambers in a “sandwich” style configuration, shown in Figure 2.23.²⁴

2.2.5.4 Thin Gap Chambers (TGCs)

Finally, in the forward regions between $1.05 < |\eta| < 2.4$, the trigger system uses TGCs, which serve the dual role of both triggering and determining the azimuthal measurement to complement the MDT measurements. The middle layer of MDTs in the endcap is paired with seven layers of TGCs, while the inner layer features only two TGC layers, comprising a total of 3588 TGC chambers with 318,000 readout channels.

²⁴A few RPCs are not attached to MDTs in select “gap” regions provide complete trigger coverage.

drive and tape reel²⁵ on earth within just a few months of data-taking. Storing this amount of data is both highly impractical and, more importantly, extremely inefficient, as the majority of events are soft interactions that do not produce “interesting” physics. To select interesting events and reduce the overall data collection rate, ATLAS employs a two-level **trigger** system that preselects notable events for storage and further processing.

2.2.6.1 Level 1 (L1) Trigger

The Level 1 (L1) trigger is a hardware-based system that uses very simple algorithms programmed onto custom FPGAs that process reduced-granularity information from the calorimeters (**L1Calo**) and the MS (**L1Muon**) to make a fast triggering decision within $2\ \mu\text{s}$, reducing the overall event rate to 100 kHz.

The L1Calo identifies objects of interest using signals from the calorimeters, which are first digitized and calibrated by the preprocessor before being sent in parallel to the Cluster Processor (CP) and the Jet/Energy-sum Processor (JEP). The CP system identifies electron, photon, and τ -lepton candidates above a programmable threshold, and the JEP system identifies jet candidates and produces global sums of total and missing transverse energy. The L1Muon system estimates the transverse momentum of muon candidates from signals from the RPCs and TGCs, combining signals to count the number of muon candidates passing a predefined threshold. Additionally, the **L1Topo** system triggers on geometric or kinematic combinations among the trigger objects received from the L1Calo or L1Muon systems. The final trigger decision is determined by the Central Trigger Processor (CTP) after receiving input from the L1 systems.

2.2.6.2 High-Level Trigger (HLT)

When the L1 trigger accepts an event, it is transferred from the event buffers in the detector electronics to a temporary storage area in the **High-Level Trigger (HLT)** computing farm, which is equipped with 40,000 processing units (PUs) that makes filtering decisions within a few hundred milliseconds. The HLT is software-based, using the regions of interest (RoI) in η and ϕ provided by the L1 trigger to perform a localized reconstruction of physics objects using the full detector granularity. Dedicated fast trigger algorithms run initially in order to provide fast rejections, then “online” reconstruction algorithms that are similar to the “offline” reconstruction (with some simpli-

²⁵This may strike the reader as surprising or old-fashioned, but magnetic tape is indeed used for long-term archival data storage in ATLAS, for e.g., raw data files that are not expected to be directly accessed on a regular basis. Tape can store large amounts of data in an extremely energy-efficient and cost-effective way, at the expense of slow recall times where the physical tape must be “staged” and read out to disk.

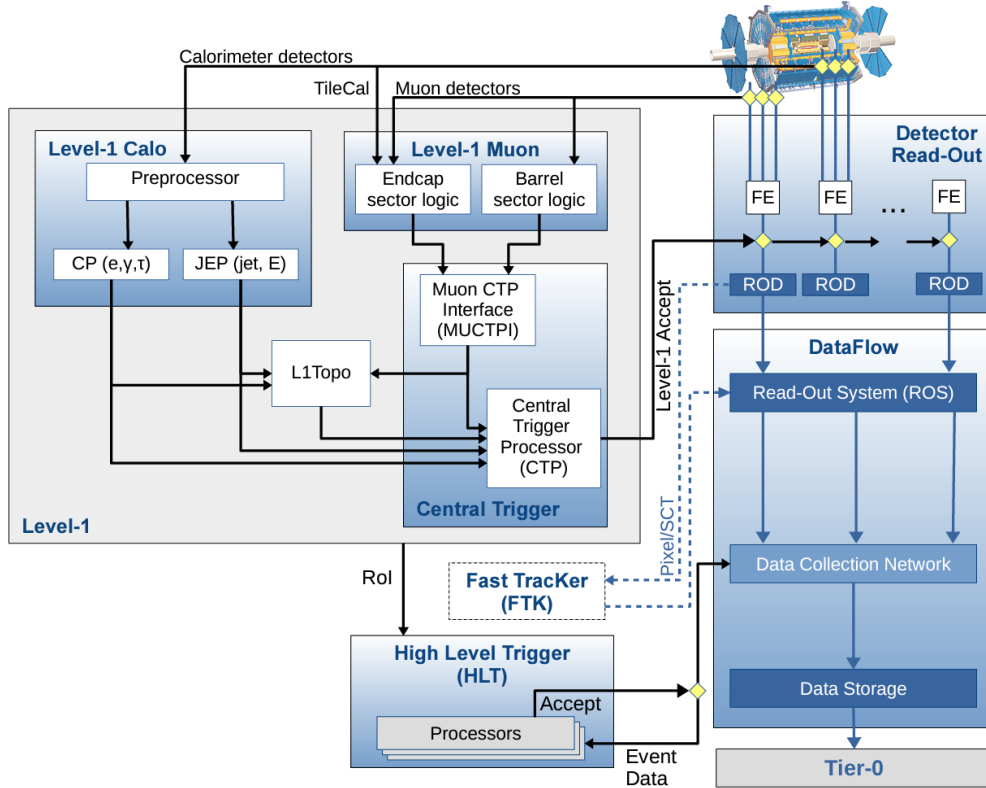


Figure 2.24. Flowchart of the ATLAS TDAQ system during Run 2, illustrating the components involved in triggering, along with detector read-out and data flow. The Fast TracKer was ultimately canceled and not used in the final Run 2 configuration. Taken from [60].

fications for efficiency) are run to fully reconstruct the objects and make the final selection, reducing the total recorded event rate down to 100 kHz. Once the HLT accepts an event, the data is moved to permanent storage for offline analysis and is forwarded to the **Tier-0** computing facility. A flowchart of the atlas TDAQ system is shown in Figure 2.24.

2.2.6.3 Trigger Menu and Selections

The thresholds that are applied to various objects by the L1 trigger and the HLT are defined in the **trigger menus** for each datataking year. ATLAS analyzers work together with the trigger group to determine specific event selections that cover a broad range of analyses while maintaining a sufficiently reduced overall event rate. Events are written based on these triggers into several “streams”, with a “main” stream for data used in primary physics analyses and several special streams used for calibrations and performance measurements. The complete trigger menus used throughout Run 2 datataking can be found in Refs [61],[62],[63], and [64].

The resonance search presented in this work uses a combination of triggers that select for single-electrons and “missing” transverse momentum (introduced in section 2.3.6), while the tracking software performance optimization results are shown using data from all available triggers in the main physics stream. Due to the differences between coarse online reconstruction and full offline reconstruction, triggers have a “turn-on” regime before they reach their full efficiency, as shown in Figure 2.25. Trigger selections are also applied to simulations to properly emulate the data behavior – however, as the online reconstruction algorithms are not additionally run on them, trigger scale factors are derived to account for the discrepancies between data and simulation.

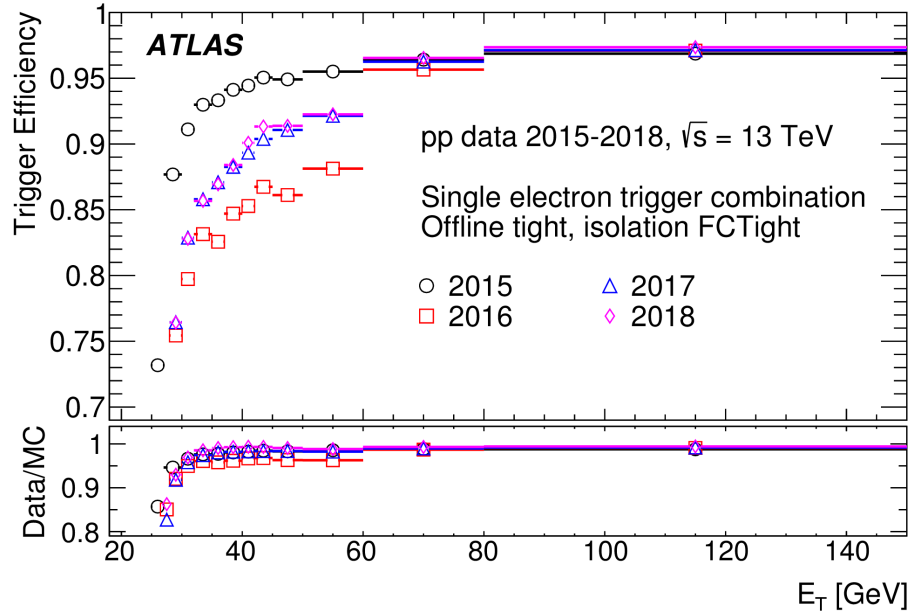


Figure 2.25. Trigger efficiency as a function of transverse energy of all combined single-electron triggers during the different years of Run 2. Efficiency is given with respect to the offline Tight electron identification working point and the FCTight isolation working point. The ratios of data to MC simulation efficiencies are also shown. Taken from [65].

2.3 Event Reconstruction

Raw data from the combined ATLAS subsystems must be transformed into **physics objects**, distinguishing the individual particle types from their specific signatures. This process is referred to as **reconstruction** and is performed by a series of algorithms designed to identify particles and measure their properties, creating collections that can be used in physics analysis. A broad overview of the variety of particle signatures is illustrated in Figure 2.26, showing the combinations of

subsystem signals that contribute to each object type. Some particles, such as muons²⁶, leave “clean” signatures that can be reconstructed directly, while hadronic signatures can only be reconstructed as “jets” from calorimeter deposits. This section describes the many different physics objects used throughout this thesis, beginning first with the fundamental primary objects. Reconstruction begins by identifying charged particle trajectories through the Inner Detector and clusters of energy deposits in the EM and hadronic calorimeters, which are then used to construct the higher-level objects such as muons, electrons, jets, and missing transverse momentum, detailed in the following.

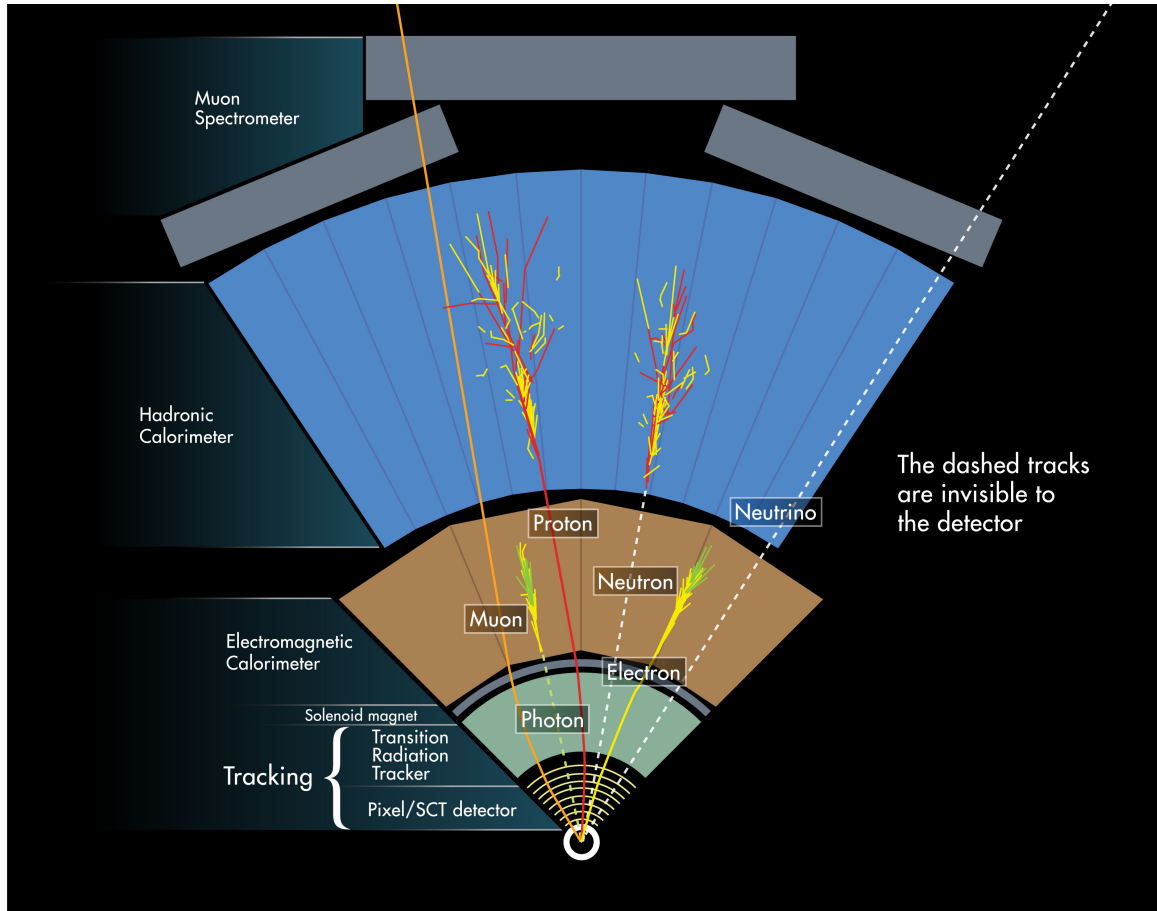


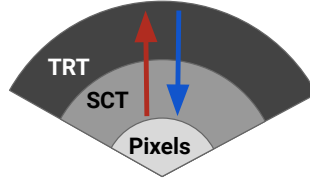
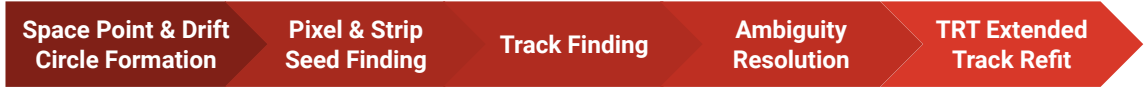
Figure 2.26. Illustration of particle signatures as they traverse the ATLAS detector subsystems. Taken from [66].

²⁶Note that in ATLAS “electrons” and “muons” are umbrella terms that refer to positrons and antimuons as well, for simplicity.

2.3.1 Track and Vertex Reconstruction

The reconstruction of charged particle trajectories, or “tracks”, in the Inner Detector is the first stage in reconstructing an ATLAS collision event. Tracks are the primary foundational element of the overwhelming majority of ATLAS physics objects used in analyses, reconstructed from combinations of hits in the inner detector. Long gone are the days of hand-tracing trajectories from bubble chamber photographs, and instead, tracks are formed by “connecting the dots” between the hits left in the ID subsystems. As a single inelastic pp interaction leaves up to 200 individual measurements in the silicon detectors, up to 15,000 hits need to be ultimately combined into final track candidates with as much accuracy as possible, implying that track reconstruction is fundamentally a combinatorial problem that scales exponentially with increasing pileup. Efficiently turning these many thousands of hits into high-quality tracks is accomplished through an ordered execution of algorithms as shown in Figure 2.27, first with a primary **inside-out** pass and then a secondary **outside-in** recovery pass.

ATLAS Primary Tracking



ATLAS Back-Tracking

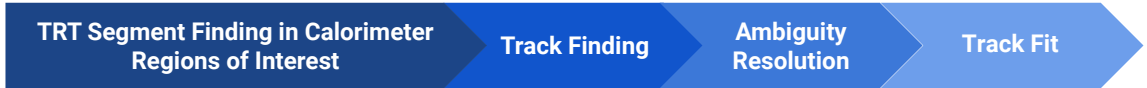


Figure 2.27. Flowchart of the stages of ATLAS track reconstruction, showing the inside-out primary pass (top) and the outside-in secondary backtracking pass (bottom).

This section will provide a general overview of the algorithms involved in track reconstruction in order. When discussing specific implementation details, this section will discuss the **legacy** implementation used throughout Run 2. Chapter 3 details a large-scale campaign of changes to optimize this track reconstruction and create an **updated** version for Run 3 that is better equipped for a high pileup environment.²⁷ Associated with this, some language in this section describing the

²⁷This section will also have a higher level of detail than some of the other stages of event reconstruction, to set the context for this work better.

stages of track reconstruction has been published in *Software Performance of the ATLAS Track Reconstruction for LHC Run 3* [17].

2.3.1.1 Track Parameterization and Propagation

Charged particle tracks are described by five parameters and a reference point (the beamspot location) using a **perigee representation**, as shown in Figure 2.28. The five parameters,

$$\text{Trk} = (d_0, z_0, \phi, \theta, \frac{q}{p}), \quad (2.12)$$

include both the transverse and longitudinal impact parameters, d_0 and z_0 , which are the distances of the single point of closest approach to the reference point in the transverse and longitudinal plane. Tracks are also defined by the azimuthal angle ϕ and the polar angle θ of the track momentum at the reference point, as well as the ratio q/p of the charge of the track divided by its momentum. The reconstruction process determines these parameters for each track candidate by identifying the sets of hits originating from a common charged particle. Covariance matrices relating to determining these parameters in the track fitting process and their associated uncertainties and correlations are also defined and associated with the track object, alongside other track-level properties, which will be described in the following. Tracks are reconstructed for $p_T > 500 \text{ MeV}^{28}$ and $|\eta| < 2.5$, where the minimum p_T threshold is driven both by the track curvature in the magnetic field and the associated number of combinatorial fakes at very low p_T values.

The charged particle trajectory is often extrapolated from a given state in the various reconstruction steps. In a homogeneous magnetic field without material interaction, a charged particle track follows an ideal helical trajectory, bending in the transverse plane and following a straight trajectory in the z direction. The ATLAS solenoid magnetic field is mainly homogeneous, producing a near-constant 2T field strength, but it has some inhomogeneities at the boundaries. Solutions for the equations of motion under the magnetic field are performed using standard numerical integration techniques with a Runge-Kutta propagator.

Material interactions are also essential to track propagation through the various stages. Charged particles interact with portions of the detector material, both the active sensing elements and the passive support structures. ATLAS uses a detailed detector **geometry model** that includes the

²⁸This threshold is a bit fuzzy and varies for different algorithms in the legacy track reconstruction – e.g., in the seeding and ambiguity solving steps, the threshold is 400 MeV, but only tracks above 500 MeV are kept at the end after the global track fit. The intention was that the track parameters could fluctuate after the fit and ambiguity resolution and not prematurely do away with candidates that might ultimately succeed. This doesn't end up recovering many good quality tracks, though, and is tightened throughout in the updated reconstruction to save CPU cycles.

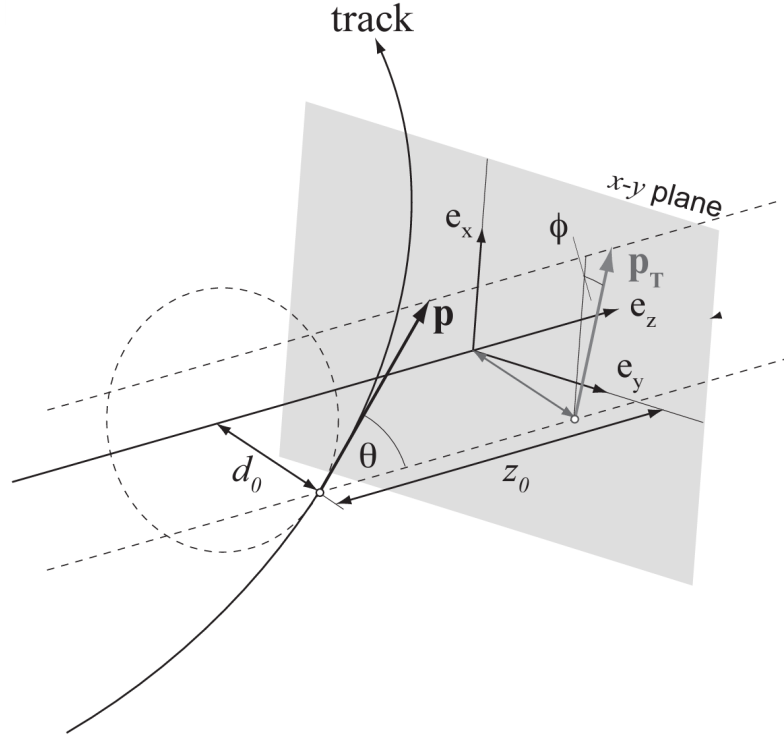


Figure 2.28. Diagram of the local perigee representation used to parameterize ATLAS ID tracks. Taken from [67].

material composition and thickness of each detector component, alongside a continually updated **alignment** model that describes the relative positions of the detector components at a given time with a high degree of precision using an extensive series of optical sensors arranged throughout the detector. When a particle reaches a detector surface²⁹ in a propagation, summarized modifications to the trajectory are applied. The effects of **multiple scattering**, which can change the trajectory of a track slightly, are modeled by increasing the uncertainty. For bremsstrahlung effects, which are very non-negligible for lighter charged particles like electrons, specialized techniques are used, which will be described in a subsequent section.

2.3.1.2 Clustering and Space Point Formation

Track reconstruction begins by combining raw pixel and SCT measurements into **clusters**, which are formed from connected groups of pixels and strips in each sensor. A connected component

²⁹It's a bit more complicated than this, track reconstruction doesn't use the "full" geometry for most purposes and instead uses a simplified geometry with a map of the summarized material properties projected onto thin surfaces for simplicity, as loading the whole geometry is very resource-intensive.

analysis (CCA)³⁰ is used to group elements sharing a common edge or corner, with a charge above a certain threshold, into clusters [51] as shown in Figure 2.29a. The particle’s interaction point with a pixel sensor is determined by taking a linear interpolation of the charges from the constituent elements. In highly dense environments, such as the cores of high-energy jets, charged particles can be collimated enough such that charge deposits from multiple particles are combined by the CCA algorithm into a single cluster, as seen in Figure 2.29b. A neural network (NN) [68] uses the original CCA clusters as input and identifies potential merged clusters while attempting to split them into sub-clusters by assigning a probability for one, two, or more particles to have contributed to its energy deposits.

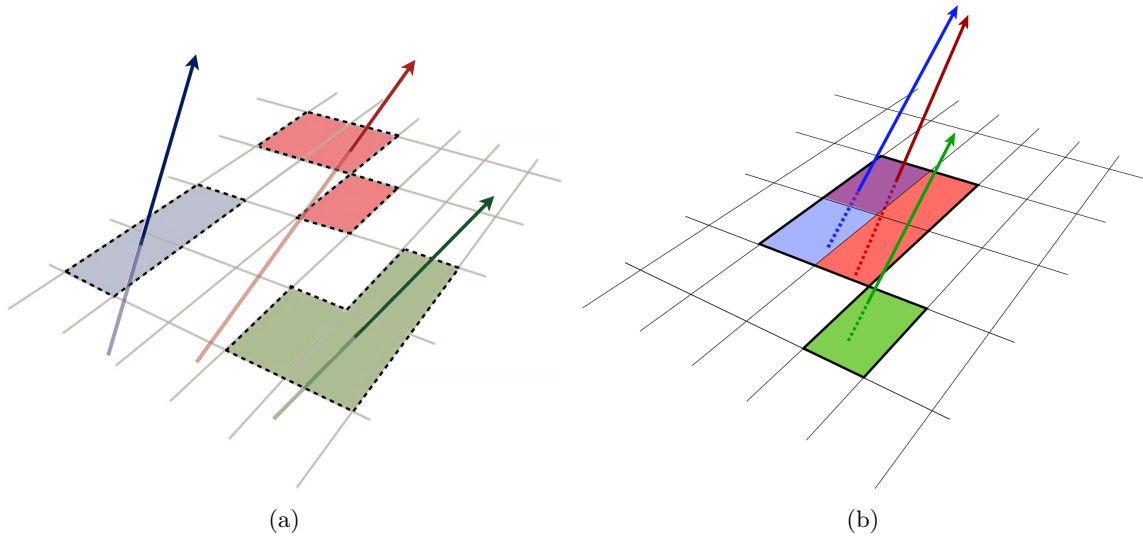


Figure 2.29. Illustration of (a) single-particle pixel clusters on a pixel sensor and (b) a merged pixel cluster due to highly collimated charged particles. The different colors represent energy deposits from different charged particles, whose trajectories are shown as arrows. Taken from [51].

Three-dimensional **space-points** are then formed from these clusters, which represent the point where a charged particle traversed the ID material for a given sensor. Individual pixel clusters, which provide 3D measurements on their own, are used to create one space point. For the SCT, 1D SCT clusters are converted into 3D space points by combining clusters on either side of a strip layer, where the stereo angle offset is used to construct a 3D measurement with position uncertainties determined by the detector geometry and sensor pitch.

³⁰CCA is an image processing algorithm based on graph theory – this is actually the same type of algorithm that determines the connected pixels when you use the magic wand tool in tools like Adobe Photoshop or MS Paint! Here, charge thresholds take the place of pixel color values.

2.3.1.3 Track Seeding

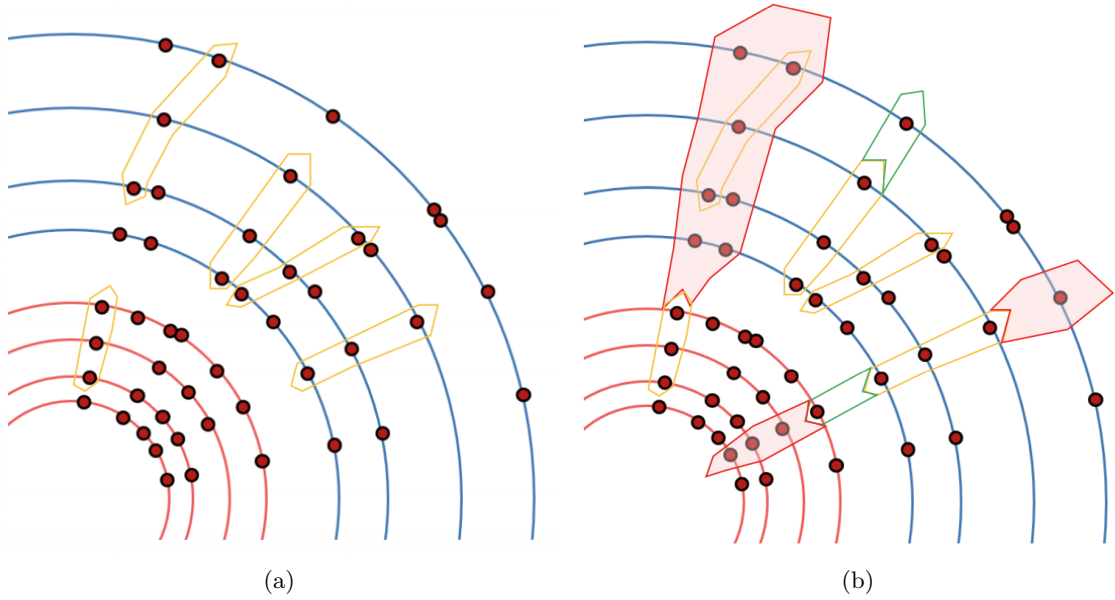


Figure 2.30. Illustration of the seeding (a) and road search (b) processes in track reconstruction using a cutaway of the pixel and SCT detectors. Triplets of space points that form seeds are outlined in orange, and the search roads formed for each seed are outlined in red.

In the following algorithm, triplets of space points are combined into track **seeds**. This approach maximizes the total number of combinations while allowing for a coarse momentum estimate to begin filtering the seeds that are unlikely to result in a good-quality track candidate. All possible combinatorial triplets with increasing radii are tested, and their momentum and impact parameters are estimated by assuming a perfect helical extrapolated trajectory through the seed measurements. Seeds are required to pass coarse quality thresholds to ensure that they are loosely compatible with having originated from the IP, such as requiring $d_0 < 10$ mm and $z_0 < 250$ mm, to be passed into further downstream processing. An illustration of the seeding is shown in Figure 2.30a.

2.3.1.4 Track Finding

After the collection of viable track seeds has been created, 3D **search roads** are built through the remaining detector based on the estimated seed trajectory. These search roads, which are sets of detector modules that are expected to contain clusters compatible with the seed, reduce total combinatorics by only considering clusters on the specific subset of modules in the path of the seed, shown in Figure 2.30b.

Track candidates are then built via a **combinatorial Kalman Filter (CKF)** [69] that progressively incorporates additional space points along the search roads. The CKF iteratively calculates a track estimate in a two-phase process, a **prediction** phase and an **update** phase. The prediction phase estimates the current track state variables and then predicts the locations of the following space points along the track trajectory, both forward and backward. If a compatible hit is found within the prediction, it is added to the track candidate. The update phase then recomputes the track state parameters after incorporating the new measurement and its uncertainty. This iteratively increases the accuracy of the track parameter estimate until the final detector layers are encountered and the track finding ends. At each detector layer, the search can “branch” combinatorially when searching for compatible measurements, creating a tree-like structure of possible track candidates from each seed when multiple sets of compatible measurements are available. An illustration of this process is shown in Figure 2.31.

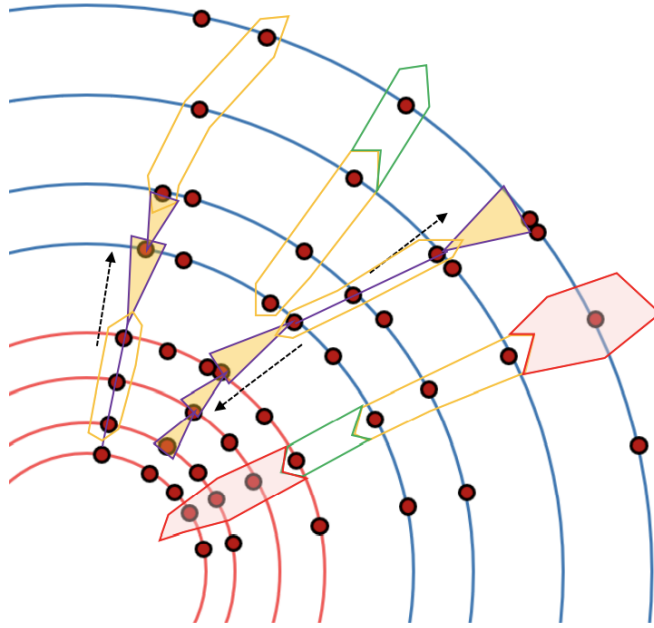


Figure 2.31. Illustration of the combinatorial Kalman filter (CKF) algorithm in the ID track finding process. Seeds, in orange, are extended recursively in a branching, combinatorial fashion as the filter iteratively searches for measurements and updates the track state parameters.

Once the collection of track candidates has been created, some basic quality requirements are again enforced to reduce the number of low-quality tracks entering downstream processing, such as requiring a minimum of at least seven silicon clusters per track.

2.3.1.5 Ambiguity Resolution

Due to the track-finding process’s combinatorial nature, a vast number of total track candidates are created. This has excellent efficiency for reconstructing real charged particles but also produces many random combinatorial fakes from collections of unrelated space points. Multiple track candidates can arise from a single seed, necessarily sharing a set of measurements, and some tracks originating from different seeds can overlap. These features necessitate a dedicated ambiguity-resolution step to resolve overlaps between track candidates and reject fake tracks.

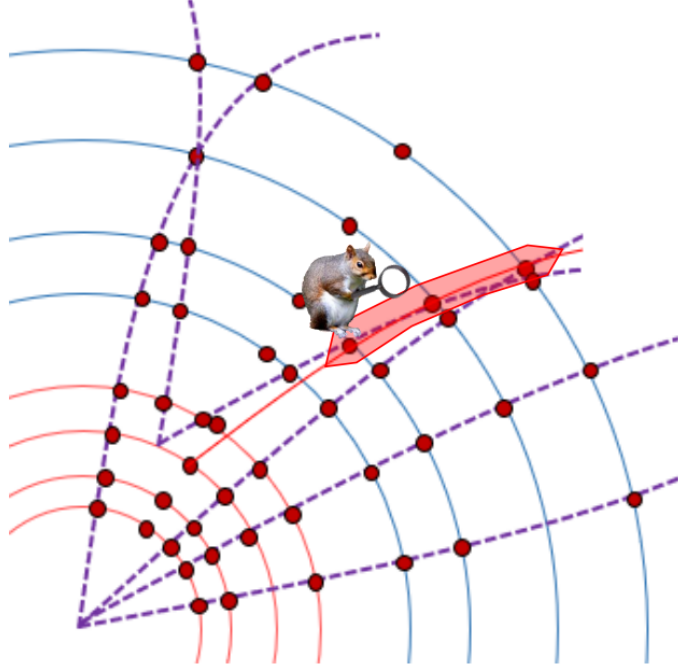


Figure 2.32. Illustration of the ambiguity-solving algorithm (represented by the squirrel) used in ID-track reconstruction. The overlapping track candidates outlined in red have multiple shared hits.

The ambiguity-solver, depicted in Figure 2.32, scores track candidates based on a range of quality criteria, such as momentum and the number of hits, shared modules, and **holes**.³¹ The number of hits on track increases the total track score, with pixel hits given a higher weight than SCT hits due to their increased precision. The number of holes on track reduces the overall score, and tracks that have more than 1 pixel hole or 2 total combined pixel and SCT holes are discarded. Tracks with

³¹Holes occur when the charged particle trajectory intersects with a functioning detector module but does not record a hit. The track finding searches all layers radially and does not terminate when an individual layer does not have a measurement, instead recording the “expected” number of hits as an additional track quantity versus the actual hit count. Modules that are known to be “dead” are counted as hits on track, to prevent localized inefficiencies in the detector.

a high score are more likely to have originated from a real well-measured charged particle, while tracks with a low score are more likely to be combinatorial **fakes**.

Tracks are allowed to have several **shared hits**, up to 1 pixel and 2 SCT, in order to maintain high performance in dense topologies. Tracks with a larger number of shared hits have their scores compared, with the higher quality track accepted and the lower quality one rejected. The aforementioned NN determines the positions and uncertainties of the crossing for each track with a shared hit. A flowchart of the ambiguity solver process is shown in Figure 2.33. Tracks that fail the ambiguity solver can also be reprocessed, allowing for an additional “kink” of up to 30% energy loss at a measurement surface to allow for bremsstrahlung radiation in electron reconstruction.

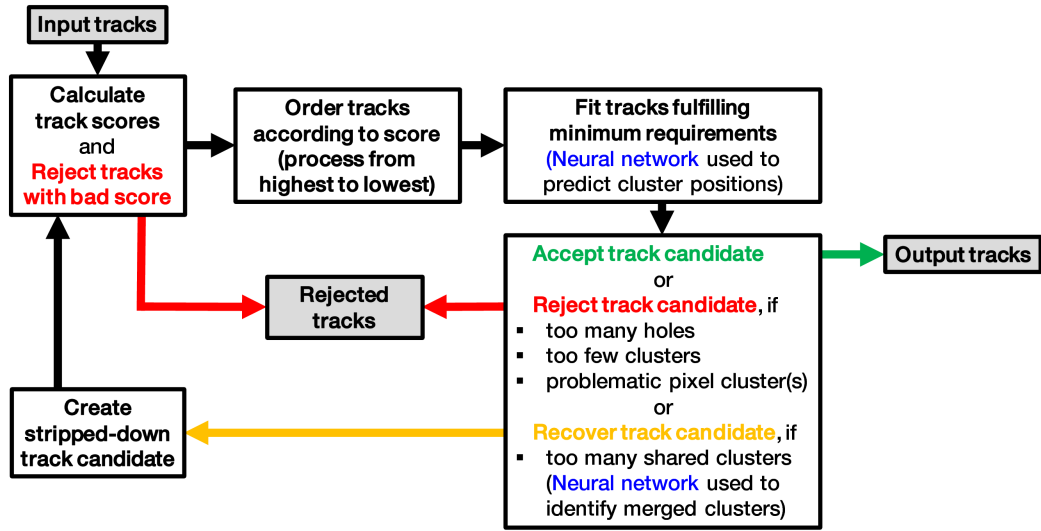


Figure 2.33. Flowchart of tracks processed by the ambiguity solver. Taken from [51].

The refined and purified track candidates collected after the ambiguity resolution are refit using a global χ^2 method, depicted in Figure 2.34, to obtain the final, high-precision track parameter estimates, accounting for errors on the measurements and their expected uncertainties due to multiple scattering and other effects. This precision fit runs several iterations until it either aborts or achieves the precision required. Tracks that abort the fit due to too large χ^2 values are discarded.

2.3.1.6 TRT Extension

Finally, tracks that pass the global refit are attempted to be extended into the TRT, depicted in Figure 2.35 by collecting drift circles close to the extrapolated trajectory of the track and performing an iterative refit of the entire track newly including the TRT drift circles, keeping the new TRT-extended track if the overall score improves. A track extension may be rejected if the track fit χ^2

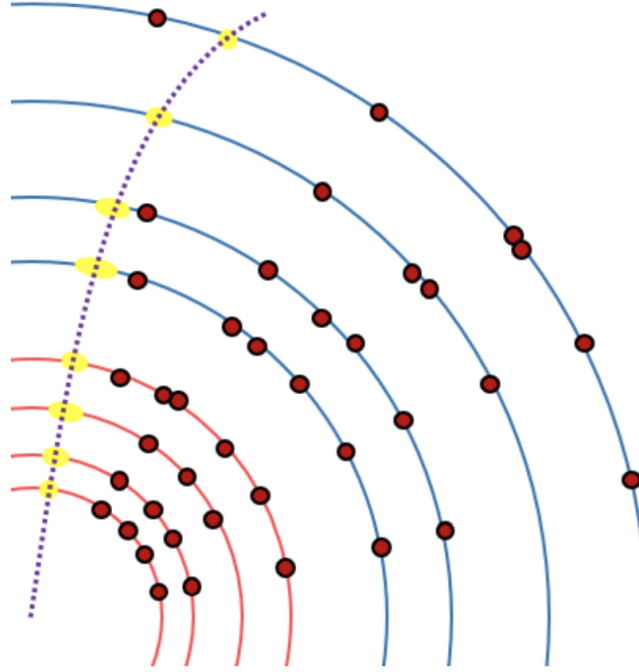


Figure 2.34. Illustration of the global χ^2 fitter used after the ambiguity resolution to finalize precision parameter measurements of ID tracks.

decreases, if there are too many **TRT outliers** (where the track passes outside the TRT straw tube), if more than 50% of the TRT hits are **tube hits** that have no leading edge, or the track does not pass through the drift circle within the straw tube, as depicted in Figure 2.36.

The additional measurements benefit extended tracks by improving the momentum resolution and allowing for particle identification using the transition-radiation detection capabilities of the TRT. To prevent spurious or low-quality extensions that can degrade the momentum resolution, the fraction of **precision hits** is considered. Precision hits are hits such that the trajectory of the incoming track has to coincide with the measured drift circle radius of the TRT hit within 2.5 times the drift radius uncertainty, and the precision hit fraction is updated up to the third iteration of the TRT extension refit. For tracks with at least 15 overall TRT hits, at least 30% of the TRT hits along a track have to be classified as precision hits for the TRT extension to be retained. The rate of successful TRT extensions depends some on the detector occupancy but is generally above 90% for muon tracks that traverse the whole ID.

2.3.1.7 Backtracking

The primary track reconstruction pass is optimized for particles produced at the primary interaction point. To increase acceptance for particles produced at a greater distance from the beamline,

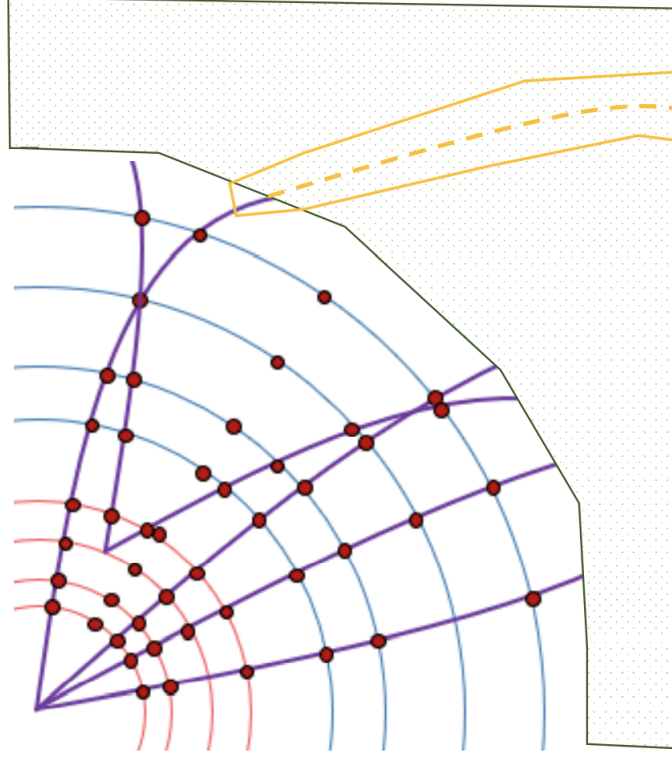


Figure 2.35. Illustration of the TRT extension step of track reconstruction, where the new TRT extension attached to the existing track (purple) is shown in orange.

such as electrons originating from photon conversions in the detector material, a secondary back-tracking pass is performed using leftover detector hits not already assigned to tracks. The back-tracking reconstruction is only attempted in regions of interest (ROI) seeded by deposits in the EM calorimeter and starts by building sets of mutually compatible hits in the TRT, known as **TRT segments**, that are compatible with the ROIs. In the presence of such a segment, short silicon track seeds consisting of two unused space points are constructed and then extended through the silicon detector volume in the same CKF, ambiguity resolution, and refit procedure as the primary pass. This process is illustrated in Figure 2.37. TRT segments that are not ultimately incorporated into an outside-in track are referred to as *TRT-Standalone* tracks. These are typically very low quality with high fake rates but are useful for some niche applications. Similarly, pixel “tracklets” are also created using a dedicated algorithm from leftover hits, which can be used to reconstruct muon trajectories at extremely high η for use in some particular applications.

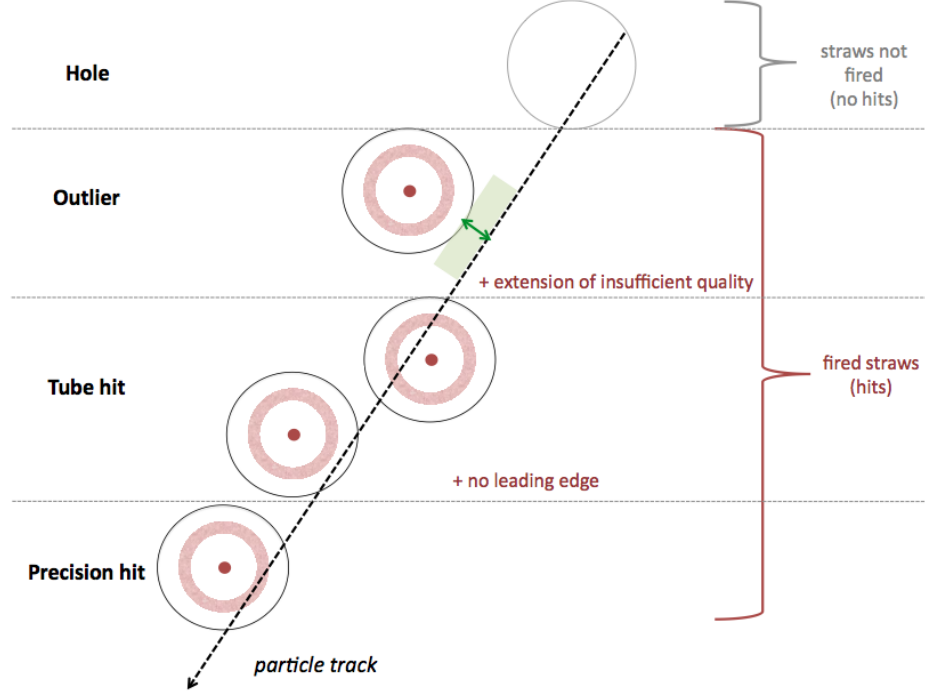


Figure 2.36. Depiction of the different kinds of TRT hits. Precision hits have a residual within 2.5 times the drift radius uncertainty. A tube hit has a radius outside this threshold or no valid signal leading edge. An outlier is a hit on track where the fitted track passes $100\ \mu\text{m}$ or more outside the straw wall. Holes are straws crossed by a fitted track but do not have a hit. Taken from [70].

2.3.1.8 Large Radius Tracking

Though not the topic of this thesis³², searches for **long-lived particles (LLPs)** with lifetimes greater than a few millimeters are exploding in popularity in both ATLAS and the broader HEP community. As increasingly stronger limits are set on many popular models in traditional prompt searches and no answers to the unsolved questions of the SM in sight, some of the community is turning to long-lived particles as a new frontier in the search for physics beyond the Standard Model. There's no well-motivated reason why new physics must necessarily be prompt, and long-lived particles can be a natural consequence of many models. However, the ATLAS track reconstruction, and the ATLAS detector as a whole, were explicitly designed to capture *prompt* decay signatures. The track reconstruction places strict requirements on the transverse and longitudinal impact parameters, as described above, to suppress fake tracks and maintain efficient use of available computational resources. To supplement the existing track reconstruction and gain sensitivity to highly displaced

³²Though the author has since moved into LLP searches after the resonance search presented in this work! It's a tremendously exciting area with fun, unique challenges for a tracking and vertexing enthusiast.

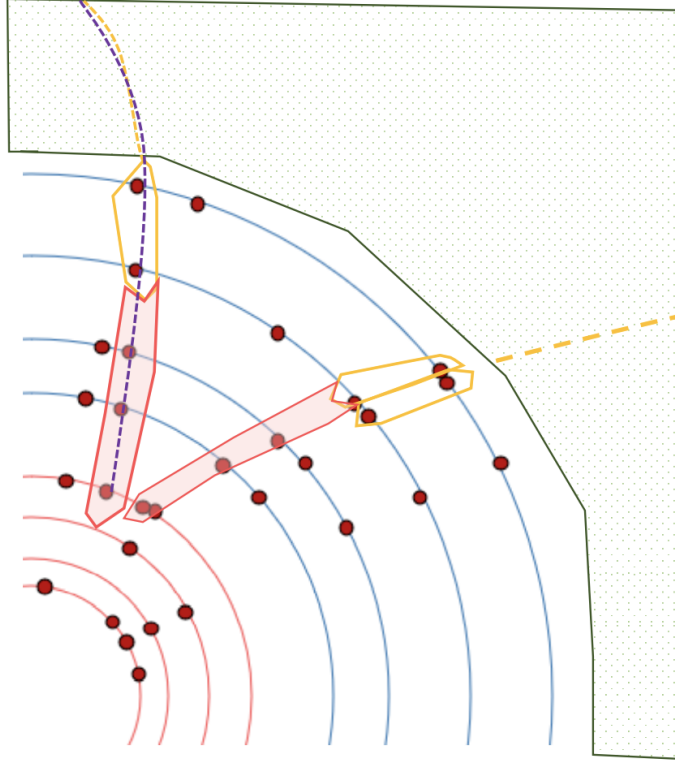


Figure 2.37. Illustration of the backtracking reconstruction pass, with the two space point seeds and TRT segments outlined in orange, search roads in red, and a final fitted track in purple.

tracks, such as those that would originate from an LLP decay within the ID volume, a tertiary track reconstruction pass known as **Large Radius Tracking (LRT)** is run on remaining hits not used in the primary or secondary passes. This pass is similar mechanically to the primary inside-out track reconstruction but seeded from only SCT clusters (as highly displaced tracks may be produced outside the pixel detector) and configured with a much broader search space in the transverse and longitudinal impact parameters, allowing up to $d_0 < 300$ mm and $z_0 < 500$ mm. This dramatically boosts efficiency for LLP searches, as seen in Figure 2.38, by enabling reconstruction of displaced tracks at the price of a dramatically slowed execution speed. In the legacy software used throughout Run 2, this additional pass would only be run in a special workflow on a pre-selected filtered subset of events comprising $\mathcal{O}(1\%)$ of the total dataset to remain computationally feasible. This preselection relied on quantities from the trigger and offline reconstruction without LRT tracks was driven by specific targeted physics analysis signatures.³³ The large-scale optimization campaigns for

³³For example, a search targeting a long-lived heavy neutral lepton produced in association with a high p_T prompt lepton would have this “DRAW” filter select for the hard lepton, where LRT would then reconstruct the displaced tracks from the HNL decay vertex.

both the primary and LRT passes [17, 71] would allow for LRT to be run by default in the primary reconstruction, over all events, for Run 3 and the subsequent reprocessed Run 2 data to be used in current analyses. This will be discussed further in the following chapter.

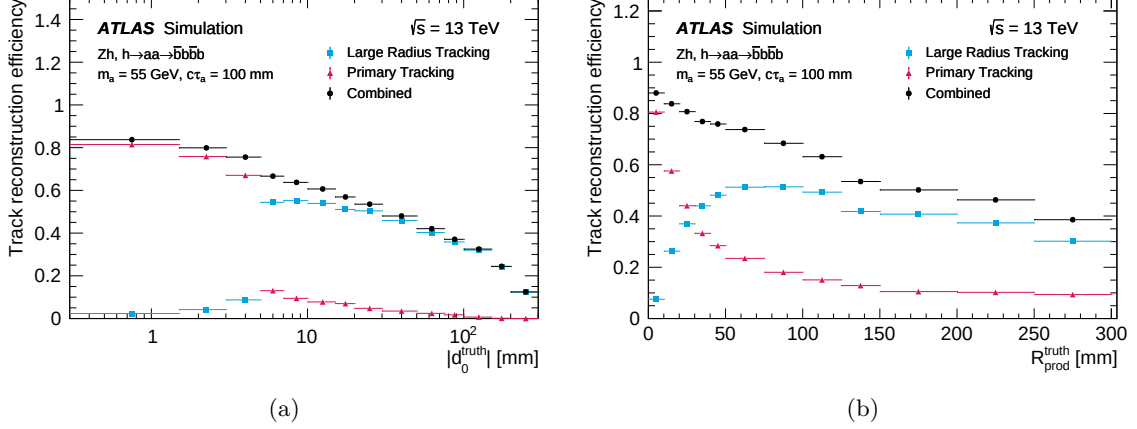


Figure 2.38. Example primary, LRT, and combined track reconstruction efficiencies for displaced charged particles produced by the decay of the LLP scalar a in a Higgs portal signal model. Efficiencies are shown as a function of (a) the generated transverse impact parameter $|d_0^{\text{truth}}|$ and (b) the generated decay radius $R_{\text{prod}}^{\text{truth}}$. Taken from [71].

2.3.1.9 Primary Vertex Reconstruction

After the primary and secondary track candidate searches have been finalized, the locations of the underlying pp interactions, or vertices, are identified by a dedicated vertex reconstruction procedure [72] using an **Iterative Vertex Finder (IVF)** algorithm throughout Run 2. This algorithm has a two-stage vertex finding and fitting procedure that iteratively assigns tracks to single vertex candidates.

First, the initial seed position for the vertex fit is assigned to the beamspot position in the transverse plane. The x and y coordinates of the starting point are taken from the center of the beamspot, and the z -coordinate is calculated as the mode of the z -coordinates of the tracks at their points of closest approach to the center of the beamspot, using the Half-Sample Mode algorithm [73]. Then, the algorithm begins to determine the best vertex location using the seed position as an initial parameter through an iterative annealing procedure. Each input track is assigned a weight that reflects its compatibility with the vertex estimate, and an iterative χ^2 minimization is performed. The vertex position is recalculated using the weighted tracks, and then the process is

repeated, where the track weights are again recalculated with respect to the updated vertex position. Individual track weights are calculated as

$$\omega(\chi^2) = \frac{1}{1 + \exp\left\{\frac{\chi^2 - \chi_{\text{cutoff}}^2}{2T}\right\}}, \quad (2.13)$$

where χ_{cutoff}^2 specifies the threshold where the weight of an individual track is equal to 0.5 and the annealing temperature T controls the smoothness of the weighting procedure with each iteration, starting high and lowering over time to prevent convergence in local minima. Tracks with large weights have a substantial impact on the vertex location, while those with low weights correspondingly have a low impact.

Once the final iteration is complete, the final weight of each track used to determine the vertex position is considered. Those that are incompatible with the vertex by more than seven standard deviations are removed from the candidate and returned to the pool of tracks that will be used to locate the next vertex. Tracks that are compatible with the vertex are removed from further consideration. This procedure is repeated until either no unassociated tracks remain or no additional vertex can be found with the remaining tracks in the pool. A diagram of the IVF logic is shown in Figure 2.39. The vertex with the highest Σp_T^2 of all associated tracks is chosen as the hard scatter vertex, which will be generally referred to as simply the primary vertex in the rest of the work.³⁴

2.3.2 Topological Calorimeter Clusters

Topological calorimeter clusters, or “**TopoClusters**”, are the other fundamental objects, alongside ID tracks, used as “building blocks” to construct the wide range of physics objects ultimately studied in ATLAS analyses. Three-dimensional collections of energy deposits are created from topologically connected EM and hadronic calorimeter cells using a nearest-neighbor algorithm [75]. The primary observable used to seed cluster formation is the cell signal significance, ζ_{cell}^{EM} , which is the ratio of the cell energy, E_{cell}^{EM} , to the expected background noise, $\sigma_{\text{noise, cell}}^{EM}$, in a cell

$$\zeta_{\text{cell}}^{EM} = \frac{E_{\text{cell}}^{EM}}{\sigma_{\text{noise, cell}}^{EM}}. \quad (2.14)$$

TopoClusters are formed by a growing-volume algorithm, starting from a highly significant seed calorimeter cell with $\zeta_{\text{cell}}^{EM} > 4$. Each seed cell forms a *protocluster*, where a “seed-and-collect”

³⁴This simple metric selects the hard scatter vertex with excellent efficiency for nearly every type of physics process. The one very notable exception is processes that result in final states with very high energy photons, such as $H \rightarrow \gamma\gamma$, as a significant fraction of the energy creates no tracks. Analyses targeting these final states use a unique NN-based approach. Some work is being done to try to create a new single ML-based approach that can identify the HS for all processes!

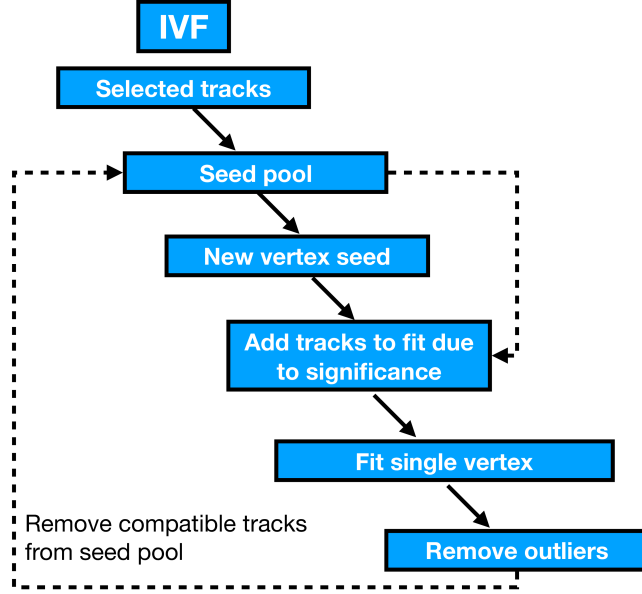


Figure 2.39. Flowchart of the IVF logic. Taken from [74].

procedure then iteratively adds neighboring cells in 3D space to the protocluster if they satisfy $\zeta_{cell}^{EM} > 2$. If a neighbor is itself also a seed cell, or if a neighboring cell is attached to two different protoclusters, the two protoclusters are merged. This is repeated on further neighbors until the last set of neighboring cells with $\zeta_{cell}^{EM} > 0$ is collected, and the cluster formation stops. This creates a “core” of cells with highly significant signals surrounded by a region of cells with less significant signals; using the correlations between energies in adjacent cells in this way allows for the periphery of showers to be captured while maintaining an appropriate level of noise suppression. The cluster formation stages can be seen in Figure 2.40. To support more detailed substructure analysis and avoid biases in jet and E_T^{miss} reconstruction, if a cluster has two or more local maxima with $E_{cell}^{EM} > 500$ MeV in the EM calorimeter, it is split between the two signal peaks in all three spatial dimensions using a dedicated algorithm that assigns weights to shared cells.

TopoClusters are assigned a four-vector spanned by their position in η, ϕ and energy, with corrections applied to account for the position of the primary vertex. The energies are assigned to clusters based on the scale of the EM calorimeter and then calibrated to account for signal losses due to inactive material and differences between responses in the EM and hadronic calorimeters. In addition to the four-vector, *cluster moments* are defined, which characterize the geometric shape of the cluster, shower direction of flight, center-of-gravity, and other relevant quantities.

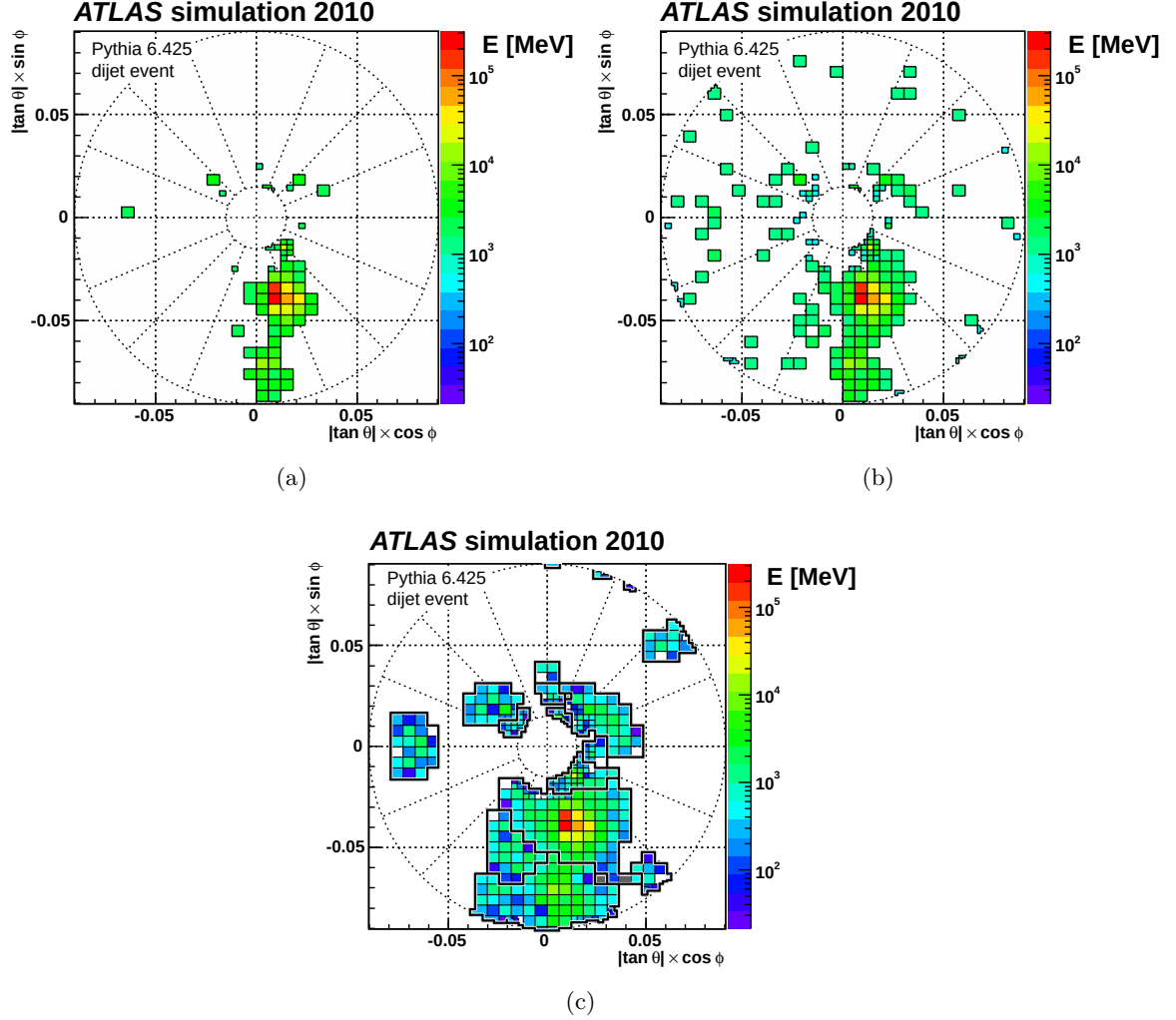


Figure 2.40. Stages of TopoCluster formation in an FCAL module for a simulated dijet event. In (a), seed cells with signal significance $\zeta_{cell}^{EM} > 4$ are shown, then in (b) cells with $\zeta_{cell}^{EM} > 2$ controlling the TopoCluster growth, and then in (c) all clustered cells with the final TopoClusters outlined in black. All clusters shown in (c) that do not contain a seed cell from this module are seeded in adjacent modules, not shown. Cluster boundaries are displayed on a dimensionless grid using θ and ϕ coordinates, preserving the general calorimeter cell shapes and proportions in the projection. Taken from [75].

2.3.3 Muons

Muons are some of the most important final-state objects reconstructed in ATLAS, characterized by distinct, “clean” trajectories that traverse the entire detector volume and (usually) leave complementary signatures in both the ID and the MS. Muons are reconstructed primarily for $p_T > 3$ GeV and $|\eta| < 2.7$, with the momentum reach limited by energy losses traversing the calorimeters and the $|\eta|$ range of the ID fiducial volume. Muon reconstruction coarsely resembles ID track reconstruction

but reversed: muon segments are first built in the MS and combined with ID tracks in a primary “outside-in” pass, with a secondary “inside-out” recovery, pass seeded from ID tracks [76, 77]. An illustrated flowchart of the muon reconstruction chain is shown in Figure 2.41, and the stages will be discussed in this section. Like the ID tracking, the muon reconstruction chain was intensely optimized for Run 3. This section will discuss the “legacy” configuration used throughout Run 2.³⁵

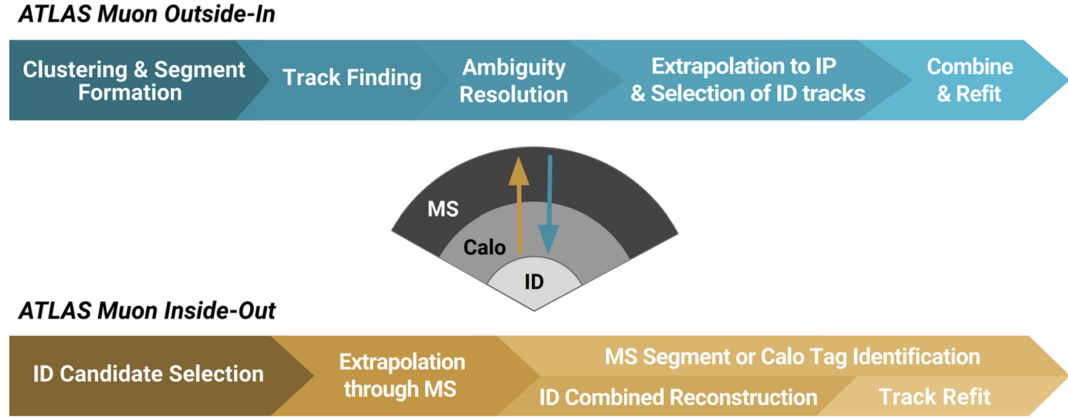


Figure 2.41. Simplified overview of the primary outside-in combined tracking chain and secondary inside-out chain used in ATLAS muon track reconstruction.

2.3.3.1 Muon Reconstruction

The muon reconstruction chain begins with a preprocessing step, where signals from adjacent CSC channels are combined into clusters and combined with the MDT drift circles, TCG, and RPC clusters to create a collection of “hits”. ID tracks that will be considered in the creation of muon candidates are also pre-selected by requiring a minimum number of silicon detector hits. Their trajectories are then extrapolated outward through the detector to obtain the expected location where they enter the MS fiducial volume and associated momentum at the intersection.

The formation of straight-line **muon segments** in the MS is initiated in 0.4×0.4 cells in $\eta - \phi$ around activity in the trigger chambers. A Hough transform [78] is used in each MDT chamber and associated nearby trigger chambers to search for hits aligned on a common trajectory along the MS bending plane, which are then combined into local muon segments by fitting them to a straight-line trajectory. The non-bending coordinate is determined using the trigger chambers if available, while

³⁵The optimizations do not change any of the basic principles of muon reconstruction and largely relate to changes in thresholds to abort the reconstruction of low-quality muon tracks “early”, alongside specific algorithmic implementation changes and other software-level improvements.

segments using only MDT information are assumed to be at the center of the given MDT chamber with a significant associated uncertainty on ϕ .

Primary Pass: Outside-In

The initial outside-in reconstruction pass is responsible for reconstructing the large bulk of muons. First, segments from different layers of the MS are combined to form MS tracks in a segment-seeded combinatorial search. Seed segments from the middle chamber layer with the most available trigger chambers are propagated between the adjacent outer and inner MS layers using an estimated parabolic trajectory. Segments are chosen based on hit count and fit quality, matched by their relative positions and angles. Generally, at least two matching segments are needed for a track, except in the barrel-endcap transition, where one high-quality segment suffices. A global χ^2 fit is performed using the complete track propagation through the inhomogeneous toroid field, material interactions, and the Ms alignment model. Incompatible measurements are removed from each MS track candidate, and compatible measurements that have not been yet added are assigned to the candidate, which is then iteratively refitted. Initially, a segment may contribute to multiple track candidates. Later, an ambiguity resolution algorithm selects the best assignment, using information from the fit quality and hit multiplicities, or allows the segment to be shared between two tracks. To ensure efficiency for low mass dimuon systems, tracks with segments in three different spectrometer layers are retained if they match in two layers but have no hits in the outermost layer.

MS tracks with $|\eta| < 2.5$ are combined with ID tracks to enhance the track parameter resolution, as the ID is highly granular. This process begins by gathering ID track candidates around the MS track within an angular window of 0.5×1.0 in η and ϕ . Candidates are initially ranked through a χ^2 comparison of their parameters, followed by attempting a trajectory fit to the complete set of ID and MS track measurements, employing global minimization to produce a unified track candidate. The fit accommodates energy loss within the calorimeter and alignment uncertainties, including relative misalignment between ID and MS. MS measurements may be added to or removed from the candidate based on overall fit quality. Subsequently, the best combined fit between the MS track and the ID track is retained. These **Combined (CB)** muons from the outside-in process constitute over 95% of those reconstructed in ATLAS.

Recovery Pass: Inside-Out

To complement the outside-in approach, an inside-out recovery chain is utilized for reconstructing muons with low transverse momenta or those entering regions of the detector where the Muon Spectrometer (MS) is not fully instrumented. This chain consists of three algorithms: **segment-tagged (ST)**, **calorimeter-tagged (CT)**, and **inside-out combined (CB)** muon reconstruction.

All these algorithms begin with ID track candidates and aim to either classify them as muons or utilize them as starting points for combined muon candidates. In preparation for this process, ID track trajectories are extended through the entire MS, recording their anticipated crossing points and momenta within muon chambers.

Segment-tagged muon reconstruction is tailored to recover muons with extremely low transverse momenta that do not traverse beyond the initial MS layer. Here, ID tracks are flagged as muons if compatible muon segments are identified. Initially, the trajectories of ID tracks are projected onto a simplified representation of the MS, modeled as a series of hollow cylinders for the barrel layers and disks for the endcap layers. Compatible segments based on location and momenta are sought within these representations, and if found, the track is labeled as a muon. In this approach, MS data serves solely for particle identification without engaging in a combined fit.

Calorimeter-tagged muons exploit the characteristic of muons as minimum ionizing particles (MIPs) to regain sensitivity in the “gap” region, $|\eta| < 0.1$, which exists in the MS for calorimeter and ID cabling. Clusters in the calorimeter along the projected ID track trajectory are examined for MIP signatures. If these signatures are found, the ID track is identified as a muon, again independent of MS data. These have a relatively low overall purity.

The inside-out combined reconstruction chain aims to recover combined muons with low transverse momenta that evade detection by the outside-in reconstruction. Beginning with ID track candidates, predicted crossing points and trajectory directions within muon chambers are utilized to establish search roads within the MS. This is then used in a dedicated muon segment-finding pass. Information from the ID-based search roads is used to refine MDT chamber measurements, allowing for the anticipated location of charge deposition along the wire direction during drift time measurement to be accounted for. Identified segments are integrated with the track, and a global refit akin to the outside-in approach is executed to obtain combined muon candidates.

When a well-defined MS track is not reconstructed in one of these ways, it can be added to a collection of **Extrapolated (ME)** muons. The muon trajectory is extrapolated backward solely based on the information from the MS track, with a loose constraint for ensuring compatibility with having originated from the IP. Muon track parameters are established at the interaction point, considering the estimated energy loss within the calorimeters. Typically, the muon must traverse at least two layers of MS chambers to provide a track measurement, although three layers are required in the forward region. ME muons primarily serve to expand the muon reconstruction coverage into the region $2.5 < |\eta| < 2.7$, which lies beyond the ID coverage. For MS tracks that still fail all of these requirements, they are added to a special collection of **StandAlone muons (SA)**.

2.3.3.2 Muon Identification and Calibration

Reconstructed muons are assigned a quality metric using four identification working points (WPs): *Loose*, *Medium*, *Tight*, and *High- p_T* . In-flight decays of light charged hadrons (such as pions and kaons) contribute background to the identification of prompt muons, which can be suppressed with varying losses of efficiency with a few key variables, where a set of selections are defined centrally for all ATLAS analyses [76]. Particularly, pions and kaon decays often result in a “kinked” topology, which can result in poor fit quality and less-compatible momenta between the ID and MS. For *Combined* muons, the variables used are:

- The q/p **significance**, which is defined as the absolute value of the difference between the charge-to-momentum ratios of muons measured in the ID and MS, divided by the corresponding uncertainties summed in quadrature;

$$\frac{|(q/p)_{ID} - (q/p)_{MS}|}{\sqrt{\sigma(p_{T,ID})^2 + \sigma(p_{T,MS})^2}}. \quad (2.15)$$

- The ρ' , defined as the absolute value of the difference between the two p_T measurements of the separate ID and MS tracks divided by the p_T of the combined track;

$$\frac{|p_{T,ID} - p_{T,MS}|}{p_{T,Comb}}. \quad (2.16)$$

- The normalized χ^2 of the combined track fit.

Ensuring a reliable momentum measurement requires some specific additional quality criteria in both the ID and MS. In the ID, quality thresholds require a minimum of one pixel hit, at least five SCT hits, fewer than three pixel or SCT holes, and the inclusion of at least 10% of originally assigned TRT hits in the final fit, with the last condition applying solely for $|\eta|$ between 0.1 and 1.9, within the whole TRT acceptance region. The identification criteria, which are inclusive such that tighter requirements are included in the looser requirements (except for the special *High- p_T* WP), are as follows:

- **Medium** muons are the default ATLAS muon selection³⁶, minimizing systematic uncertainties linked to muon reconstruction and calibration. It exclusively uses *Combined* and *Extrapolated*

³⁶Though actually, not explicitly used in the Wh resonance search presented in this thesis, which instead selects *Tight* and *High- p_T* muons and uses *Loose* muons as a veto in the zero-lepton channel. However, these WPs are based on the medium criteria as a starting point.

muons, where Combined muons must have ≥ 3 hits in at least two MDT layers, except for $|\eta| < 0.1$, where those with at least one MDT layer but no more than one MDT hole layer are permitted. ME tracks, required to possess at least three MDT or CSC layers, are utilized solely in the $2.5 < |\eta| < 2.7$ region to extend coverage outside the ID geometry. A loose constraint is also placed on q/p significance < 7 , which mitigates hadron misidentification.

- **Loose** muon criteria aim to maximize reconstruction efficiency while ensuring reasonably robust muon track quality. All muon types are considered, with those satisfying the *Medium* requirements included. Calo- and segment-tagged muons are then additionally allowed but are confined to the $|\eta| < 0.1$ gap region.
- **Tight** muons prioritize purity over efficiency, selecting only *Combined* muons meeting Medium criteria and having hits in at least two MS stations. A normalized $\chi^2 < 8$ is required, as well as a two-dimensional cut on the ρ' and q/p significance as a function of muon pT. This enhances background rejection for momenta below 20 GeV, where misidentification probability is elevated.
- The **High- p_T** selection maximizes momentum resolution for muons with $p_T > 100$ GeV and is specifically tailored for searches involving high-mass Z' and W' resonances such as the one presented in this thesis. It selects *Combined* muons passing Medium criteria and having at least three hits in three MS stations, vetoing specific MS regions with suboptimal alignment. Though it reduces overall reconstruction efficiency by about 20%, requiring three MS stations improves p_T resolution for muons above 1.5 TeV by approximately 30%.

Following reconstruction, corrections are also made to muon candidates to ensure consistency between data and simulation concerning muon momentum scale and resolution. These corrections are derived from comparisons between muon p_T in data versus simulation in specific detector regions in $\eta - \phi$ that are homogeneous in expected performance. These corrections are associated with systematic uncertainties. Additionally, like electrons, muons also are assigned isolation variables via the same p_T^{cone} and E_T^{cone} treatments as described previously.

2.3.4 Jets

Due to the inherent nature of hadronization, it is impossible to reconstruct single hadrons directly. Instead, hadronic final states are characterized by **jets**, collimated sprays of particles formed in the hadronization process that leave cone-shaped energy depositions in the calorimeters. Jets are formed

from various **jet clustering algorithms** that create localized, topologically connected groups from input objects, such as TopoClusters, ID tracks, or generator-level truth particles.

The most commonly used jet clustering algorithm in ATLAS is the **anti- k_t** algorithm [79], which is a sequential recombination algorithm that iteratively combines input objects into jets. This is done based on two distance metrics, the pairwise distance between two input objects d_{ij} and the distance between an object i and the beam axis d_{iB} :

$$d_{ij} = \min \left(\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,j}^2} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad (2.17)$$

$$d_{iB} = \frac{1}{k_{T,i}^2}, \quad (2.18)$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and $k_{T,i}$, y_i , and ϕ_i are the transverse momentum, rapidity, and azimuthal position of an input object i . R is the radius parameter controlling the maximum width of the jet and varies from implementation to implementation depending on the targeted topology. For each input object i , the algorithm computes these quantities for each other input object. If the smallest quantity is d_{ij} , then the two objects are merged to form a single compound object, and the separate “original” objects are removed from consideration. This process is repeated until the minimum quantity is d_{iB} , where the combined i object is removed and taken to be a complete jet. This results in an ensemble of recombined jet objects, as seen in Figure 2.42. The anti- k_t algorithm is preferred over other schemas, such as the “standard” inclusive k_t algorithm or Cambridge/Aachen algorithms, as it preferentially clusters “hard,” high p_T inputs and absorbs all soft clusters in a given R value around the hard cluster. This limits the effect of noisy soft radiation on the final shape of jets and creates predominantly conical jets with clear boundaries. As many targeted final state topologies contain a few well-separated hard particles with high p_T accompanied by many soft particles, this approach strikes the appropriate balance.

A wide range of algorithms, the vast majority based on modifications of the anti- k_t algorithm with varying input types, are used to reconstruct jets in ATLAS. Each has its strengths and weaknesses and is optimized to target widely different signatures and topologies – for brevity, the discussion in this section will be limited to those used in the Wh resonance search discussed in this thesis.³⁷

³⁷Jet reconstruction is a vibrant, diverse field that is difficult to do justice to in a few pages. The discussion here leaves out more modern Particle Flow (PFlow) jet algorithms, which have now become ubiquitous in ATLAS, or any discussion on Unified Flow Objects (UFOs), which unfortunately postdate the work presented but will ultimately be studied in a future iteration of the Wh analysis.

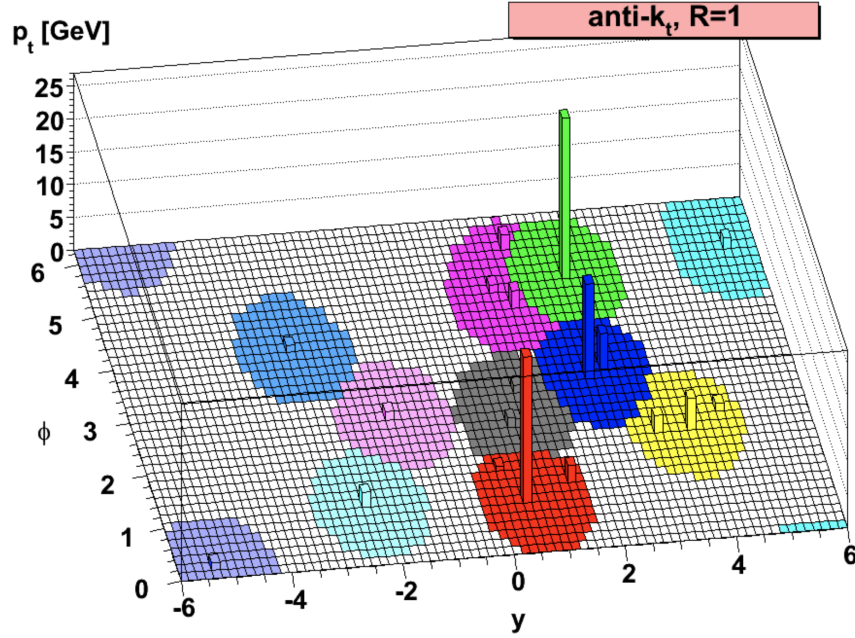


Figure 2.42. A sample parton-level event clustered with the anti- k_t algorithm with $R = 1$, illustrating the catchment areas of the resulting hard jets. Taken from [79]

2.3.4.1 Jet Types and Calibrations

Three types of reconstructed jets are employed in the Wh resonance search. **Calorimeter jets** are formed from TopoClusters, **track jets** from ID tracks, and a specialized combination of both, **TrackCaloClusters**, form **TCC jets**. Each type also has its associated set of calibrations, which follow a similar overarching strategy. However, they must be performed for each individually as they are subject to different sets of energy losses and biases. Each jet type is chosen for a specific purpose in the analysis related to the identification of the $H \rightarrow b\bar{b}$ final state in different regimes, detailed in the following.

Small-R ($R = 0.4$) Calorimeter jets

Small-R calorimeter jets are the “base” jet type used throughout ATLAS, constructed from the anti- k_T algorithm described above using calorimeter TopoClusters as the input object and a radius parameter of $R=0.4$. These jets can be used to broadly reconstruct hadronic decays from a wide range of signatures, including non-boosted (resolved) $H \rightarrow b\bar{b}$ decays. A sequence of corrections [55] are applied to small-R jets to account for pileup dependence and calibrate the **jet energy scale**

(JES) and **resolution (JER)**, correcting the four-momentum of the reconstructed jets to match the jets at the generator level. The steps used in the calibration are shown in Fig 2.43.

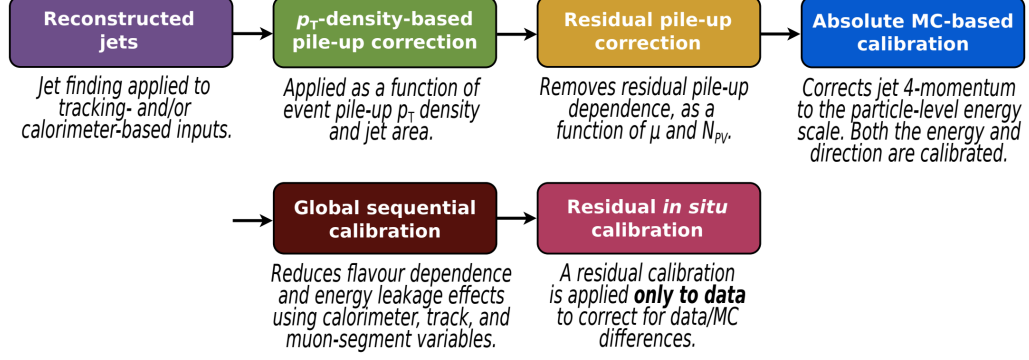


Figure 2.43. Flowchart of the jet energy scale calibrations. Taken from [55]

Pileup activity significantly affects the reconstruction of jets, as it contributes to the energy deposits in the calorimeter cells used to create the TopoClusters. These pileup contributions are subtracted from the reconstructed jet p_T through two corrections. The initial calculation determines the median p_T density, $\rho = \langle p_T/A \rangle$, of jets in the y - ϕ plane in the central region between $|\eta| < 2$. Here, A represents the jet's vulnerability to pileup, calculated based on the relative number of ghost-associated [80] particles linked to a jet following clustering. The ratio of the ρ -subtracted jet p_T to the uncorrected jet p_T can then be applied as a scale factor to the jet four-momentum. The central region is chosen to create a more meaningful measurement of the average p_T density, as occupancy drops rapidly in the forward regions due to their coarser calorimeter segmentation. However, this implies that some pileup dependence remains from the forward regions, and thus, a residual correction is derived to remove the remaining component. The p_T response depends linearly on the multiplicity of the primary vertices in an event, N_{PV} , and the average number of interactions per bunch crossing μ , which are sensitive to in-time and out-of-time pileup, respectively. The total **pileup-corrected** jet p_T is then given by

$$p_T^{corr} = p_T^{reco} - \rho \times A - \alpha \times (N_{PV} - 1) - \beta \times \mu, \quad (2.19)$$

where α and β are linear coefficients determined through linear fits to the truth jet p_T in coarse bins of $|\eta|$.

Next, the **absolute MC-based correction** is performed, which adjusts the reconstructed jet four-momenta to the particle-level energy scale, addressing the non-compensating calorimeter response, energy losses in passive materials, and biases in jet η reconstruction. These biases typically arise from transitions between different calorimeter technologies and abrupt changes in calorimeter granularity. The average energy response, $\mathcal{R}$, is derived as the mean of a Gaussian fit to the E^{reco}/E^{truth} distribution. Then, the jet calibration factor is taken as the inverse of the average energy response. The transitions between different detector technologies with different energy responses can also cause a bias in the η measurement, causing one side of the jet to have a higher energy artificially, thus biasing the four-momentum. A second correction is derived as the difference between η^{reco} and η^{true} parameterized as a function of E^{true} , which is then also inverted to give a correction in E^{reco} .

The final steps of the calibration procedure correct for remaining variations in responses from jet-to-jet due to factors such as the flavor and energy distributions of their constituent particles as well as differences in composition and shower shape between quark- and gluon-initiated jets. The **Global Sequential Calibration** (GSC) employs a series of multiplicative corrections to mitigate these variations and enhance the jet resolution without altering the average jet energy response. The GSC applies independent corrections derived from five global jet observables, such as the longitudinal structure of the calorimeter energy deposits as well as neighboring ID track and muon track segments that can be associated with the jet. Following the GSC, *in situ* calibrations are applied to data only to account for the final remaining differences in the jet response between data and simulation caused by imperfect simulations of both detector materials and relevant physics processes. Correction scale factors are derived from a statistical combination of three measurements to ensure a smooth calibration applicable across a range of momenta. The first, the η intercalibration, corrects the energy scale of forward ($0.8 \leq |\eta_{det}| < 4.5$) jets to match those of central ($|\eta_{det}| < 0.8$) jets by balancing the p_T in dijet events. The other two derive calibrations from balancing jet p_T against multiple well-calibrated reference objects. The calibration workflow introduces many sources of systematic uncertainty associated with the now-calibrated jets, which are discussed further in Section 4.4.1.

Large-R ($R = 1.0$) TCC jets

TrackCaloClusters [81] are designed to improve resolution in resolvable energy structures within a jet, known as **jet substructure**, by combining the superior angular resolution and granularity of the Inner Detector with the exceptional energy resolution of the calorimeters. This combined approach

is well-justified in the very high-energy regimes where jets are extremely collimated; the ID has far superior angular resolution than the calorimeter but poor energy resolution and vice versa.³⁸

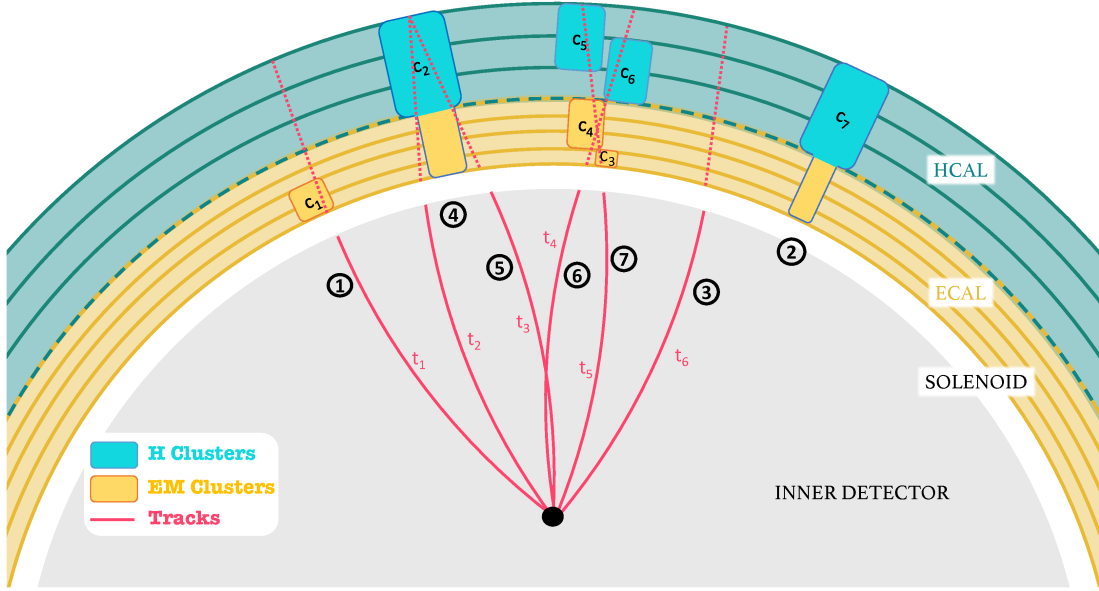


Figure 2.44. A diagram illustrating the creation of seven different TCC objects. A simple track-cluster match is shown in (1), (2) shows a cluster without a matched track, (3) shows a track without a matching cluster, (4) and (5) each show tracks matching a single cluster but sharing that cluster’s energy, and (6) and (7) show a complex scenario with multiple different track-cluster matches. Taken from [81]

A matching procedure builds four-vectors from a combination of track and TopoCluster information, sequentially attempting to match every track of sufficient quality to every TopoCluster. First, the uncertainty on the track extrapolation to the calorimeter is compared to the width of the TopoCluster—if the uncertainty is larger, that track-cluster combination is discarded. Otherwise, the track-cluster pair is considered matched when their angular separation satisfies

$$\Delta R < \sqrt{\sigma_{cluster}^2 + \sigma_{track}^2}, \quad (2.20)$$

where $\sigma_{cluster}$ is the TopoCluster width and σ_{track} the track extrapolation uncertainty. At high p_T , track extrapolation uncertainties are quite small, such that this criterion reduces to approximately $\Delta R \leq |\sigma_{cluster}|$ in the high p_T limit. This procedure uses tracks originating from any primary vertex rather than only tracks matched to the hard scatter vertex. Unlike the ID, the calorimeters do not

³⁸Much of the improvement over just increased statistics from the 36 fb⁻¹ *Wh* analysis can be attributed to the addition of TCC jets and VR track jets in place of “standard” EMTopo jets and fixed-radius track jets, particularly in the high-mass regime targeted by the merged category.

have sufficient resolving power to distinguish between pileup vertices and the hard scatter (thus necessitating the complex calibrations to remove pileup contributions, detailed above for small-R jets). Tracks from pileup vertices matched to TopoClusters are not used to build independent TCC objects but rather allow for better determination of energy overlaps in the calorimeter and calculating the total energy assigned to final TCCs. Tracks that are compatible with the primary vertex and matched to TopoClusters are referred to as “*combined*” TCCs, while TopoClusters not matched to any track are “*neutral*” TCCs, and tracks from the primary vertex not matched to clusters are defined as “*charged*” TCCs. All varieties are used in the jet clustering, and various examples can be seen in Figure 2.44.

Four-vectors for the TCC objects are constructed from the combined information through a variety of methods, ranging from a simple combination of the TopoCluster energy and track trajectory parameters in a one-to-one match to a more complicated calculation involving energy-sharing scale factors when multiple tracks are assigned to numerous overlapping TopoClusters. Further details on the formation of TCCs can be found in Ref [81].

Large-R TCC jets are clustered from TCC objects using the anti- k_T algorithm with an enlarged radius parameter of $R = 1.0$. Large-R jets allow for the reconstruction of highly boosted decays where the decay products are emitted with too small an angular separation to be sufficiently resolved into two distinct structures, such as $H \rightarrow b\bar{b}$ decays resulting from a hypothetical high-mass resonance.

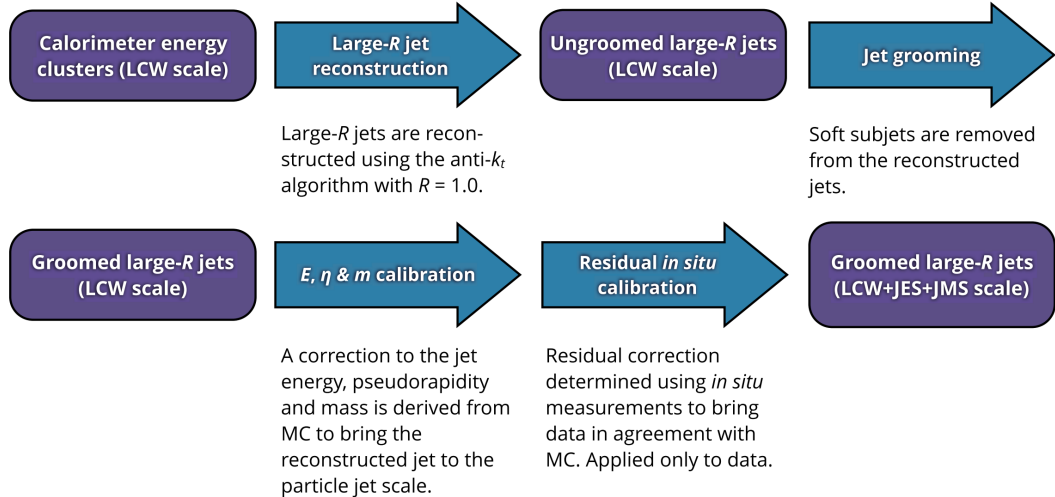


Figure 2.45. Flowchart of large-R jet calibrations. TCC jets use TrackCaloClusters instead of LC topo-clusters (LCW) as depicted but follow the same procedure. Taken from [82]

Large-R TCC jets follow a very similar calibration procedure as described for small-R calorimeter jets above, shown in Figure 2.45. However, they are particularly susceptible to pileup effects due to their large area, which gives a wider “surface” for nuisance contributions to accumulate, and thus require an additional **grooming** procedure where the jet constituents are reclustered using the anti- k_T algorithm with a smaller radius parameter $R_{sub} = 0.2$. Any subjects that are successfully created with $p_T^{sub}/p_T^{jet} < f_{cut}$ are removed, where f_{cut} is a defined threshold, usually 5%. This “trimming” procedure is shown in Figure 2.46.

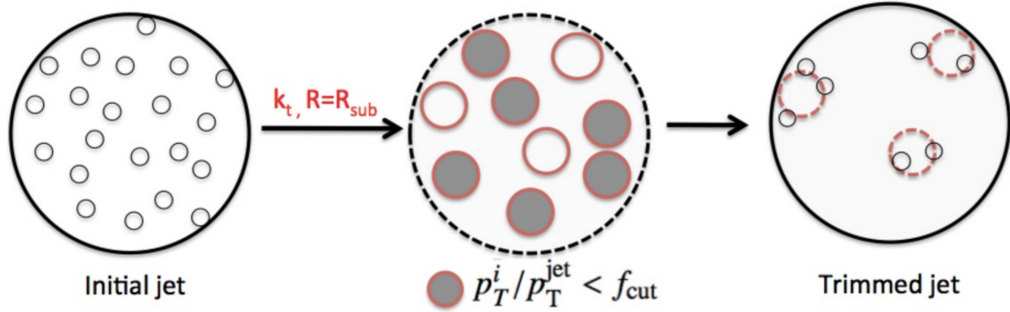


Figure 2.46. Illustration of the jet trimming procedure. Taken from [83]

Variable-Radius (VR) Track Jets

Unlike the other two jet types outlined above, VR track jets are clustered with the anti- k_T algorithm from ID tracks only, without using calorimeter information [84]. They have a distinguishing feature that the radius parameter R in Equation 2.17 is modified to scale with the transverse momentum of the forming proto-jet:

$$R \rightarrow R_{eff}(p_T) = \frac{\rho}{p_T}, \quad (2.21)$$

where ρ is a defined parameter that controls how fast the jet shrinks with increasing p_T , in ATLAS $\rho = 30$ GeV. [85] VR jets are also given a minimum and maximum radius parameter, $R_{max} = 0.4$ and $R_{min} = 0.02$, to prevent them from shrinking below the detector resolution or becoming overly large. VR track jets are used to identify the flavor content of large-R jets, to identify boosted $H \rightarrow b\bar{b}$ decays, with a better resolution than fixed-radius jets as depicted in Figure 2.47. The b -tagging algorithms, discussed in the next section, use these associated track jets rather than the large-R jet directly. VR track jets are matched to large-R TCC jets through ghost-association [80, 86], where “ghost” copies of the VR jets are created that have the same direction parameters but infinitesimal

p_T . The large-R TCC jets are then reclustered, adding these ghost jets to the possible constituents. If the large-R jet “picks up” one of these ghost jets in the re-formation, they are considered ghost-associated. This procedure provides a robust matching criterion between the two types of jets.

As they do not use calorimeter information, VR track jets are not subject to the extensive series of JES calibrations outlined above, though they do use a *Loose Primary* track selection to ensure jets are not clustered from very low-quality tracks.

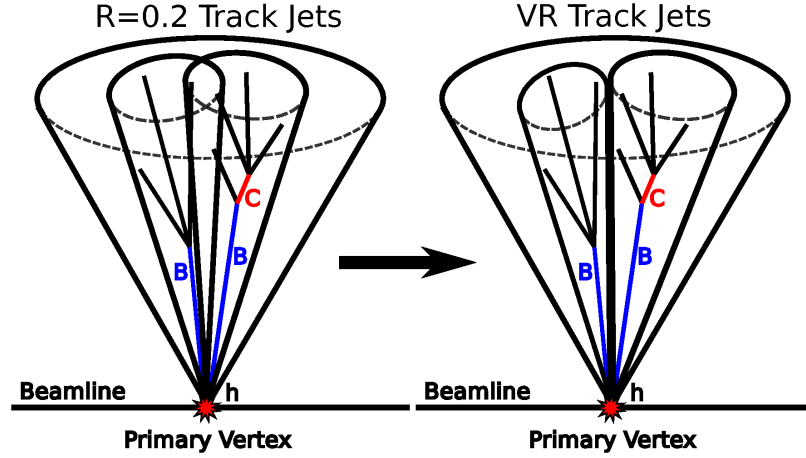


Figure 2.47. Illustration of fixed-radius jets (left) versus variable-radius jets (right) inside a matched large-radius jet. The secondary and tertiary vertices from b -quark decays, used for b -tagging, are shown. Taken from [87]

2.3.4.2 b -Tagging

Identifying jets that are likely to have originated from a b -hadron is essential to the reconstruction of $H \rightarrow b\bar{b}$ decays. This decay mode is particularly interesting for searches and measurements as not only does it have the largest branching ratio of the Higgs decay modes, but the long-lived nature of b -quarks gives rise to unique kinematic properties that can be exploited on the analysis level to “tag” b -jets and dramatically improve sensitivity. The suppressed b -decay results in decay lengths up to a few millimeters, which can be properly reconstructed as a secondary vertex within the high-granularity ID alongside other key distinguishing kinematic variables. Many algorithms for b -tagging are used throughout ATLAS, the majority built with machine learning techniques trained on these unique kinematics. The resonance search presented in this thesis uses the **MV2 algorithm**, which uses a multivariate classifier built from a boosted decision tree (BDT) trained on $t\bar{t}$ MC simulations [88, 89]. The resulting output gives a score from -1 to 1, with 1 being maximally signal-like and -1

minimally signal-like. The working point chosen in the Wh resonance search corresponds to a 70% b -jet selection efficiency.

The output of three complementary sub-algorithms are combined to develop the MV2 discriminant, each using ID tracks throughout. For calorimeter jets, tracks are associated based on their angular separation varying with jet p_T . Tracks are associated with only one jet, where if multiple jets satisfy the matching criteria for a track, the one with the smallest ΔR is chosen.

The first set of algorithms uses the signed impact parameter significance of the matched tracks, where the sign is defined as positive if the point of the track's closest approach to the primary vertex is in front of the primary vertex with respect to the jet direction (and vice versa, for a negative sign) [90]. Probability density functions are constructed from the distributions of the significance for different qualities of tracks and used to define ratios of the b - and light-flavor jet hypotheses, which are then combined in a log-likelihood ratio discriminant (LLR). The IP3D algorithm uses both the transverse and longitudinal impact parameters, while the IP2D algorithm uses only the transverse impact parameter, and outputs from both pass to the final discriminant.

The second algorithm, SV1, [90] reconstructs an inclusive secondary vertex inside the jet. Initially, it reconstructs two-track vertices using pairs of candidate tracks passing a loose quality selection. If tracks create a secondary vertex that appears to originate from the decay of a long-lived particle (such as K_S or Λ), from photon conversions, or from interactions with the detector material, they are discarded. Following this initial selection, a single vertex is constructed from the remaining tracks.

The third algorithm, JetFitter [91], is designed for multi-vertex reconstruction of decay chains and utilizes the topological characteristics of weak b - and c -hadron decays within the jet to reconstruct the entire decay sequence from the primary vertex (PV) through the b -hadron to the c -hadron. It employs a Kalman filter to determine a common trajectory that approximates the flight path of the b -hadron and locates the primary, bottom, and charm vertices along it. This method enables the resolution of b - and c -hadron vertices, even in cases where each vertex is associated with only a single track, as long as the resolution permits.

The input variables collected from three primary algorithms are merged using a boosted decision tree (BDT) algorithm to differentiate b -jets from lighter jets and c -jets. The MV2c10 algorithm is defined as the output of this BDT where the training is performed by classifying b -jets as the signal

and a mix of 93% light-flavor jets and 7% c -jets as the background.³⁹ The output distribution of MV2c10 for b , c , and light-flavor jets is shown in Figure 2.48. Correction scale factors and their associated uncertainties are derived to account for differences in selection efficiency between data and simulation in the performance of the algorithm, which is trained on MC only.

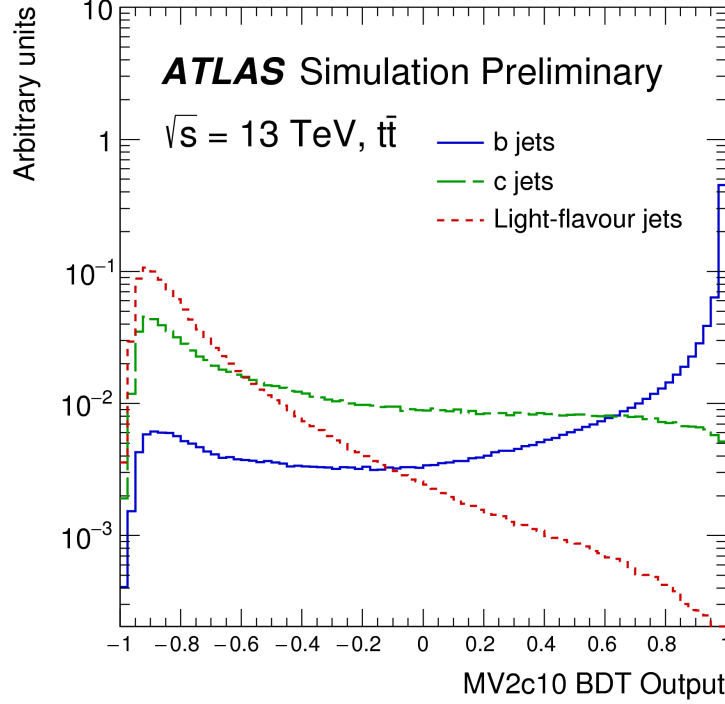


Figure 2.48. Output score of the MV2c10 b -tagging algorithm for b -jets (blue), c -jets (green), and light-flavor jets (red). Taken from [88]

2.3.4.3 Hadronic Taus

While leptonically decaying τ -leptons are reconstructed with the resulting electron or muon (with some loss of momentum from the neutrino), reconstruction of hadronically decaying τ -leptons follows a similar method to b -tagging and broadly uses multivariate classifiers developed with machine learning techniques exploiting the slightly long lifetime and produced secondary vertices [92]. Hadronic τ decays are reconstructed using small-R calorimeter jets and are uniquely characterized by having an odd number of associated tracks creating “1-prong” and “3-prong” signatures. Three identification

³⁹The “ cXX ” designation directly referred to the background charm fraction in the first iterations of the algorithm, but further studies showed that keeping the optimal charm vs light jet rejection power for the different available variations didn’t settle on a nice round number. It was decided to keep the old nomenclature anyway.

working points, *Tight*, *Medium*, and *Loose*, are defined corresponding to signal efficiencies of 75%, 85%, and 95%, respectively. The Wh resonance search only uses reconstructed taus to veto events with multiple identified τ decays, using the broad *Loose* working point. A further in-depth discussion about hadronic τ identification can be found in Ref [92].

2.3.5 Electrons

When electrons traverse ATLAS, they generally leave both a track in the Inner Detector and a shower in the EM calorimeter, which is reconstructed as the aforementioned TopoClusters. These two signatures are combined to produce **electron candidates** [93, 94], whose reconstruction is detailed below.

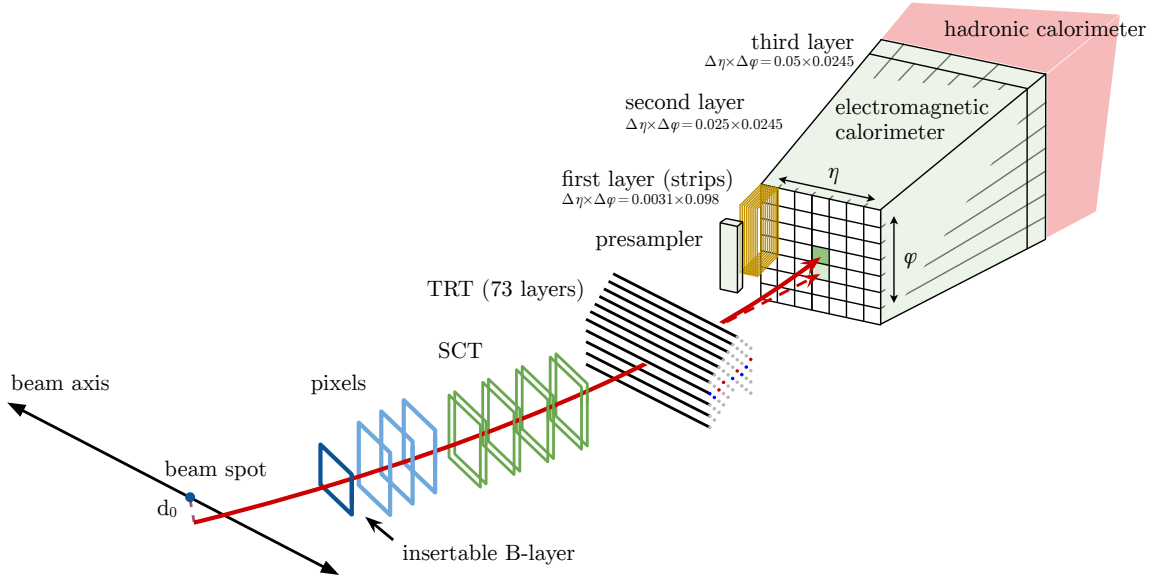


Figure 2.49. Schematic illustration of the path of an electron through ATLAS. The red arrow indicates the path of the electron, including a secondary photon (dashed) produced by material interactions. Taken from [93].

2.3.5.1 Electron Reconstruction

Reconstruction of electron candidates begins by using the reconstructed TopoClusters to search for potential ID tracks within a loosely matching $\eta - \phi$ region. Measurements from tracks that satisfy this loose matching criterion are subsequently refitted using a **Gaussian Sum Filter (GSF)** [95], which is better tailored than the traditional CKF used to build the original track to more accurately accommodate the energy loss of charged particles in the material layers. The GSF method builds

upon a generalized version of the Kalman filter and accommodates the nonlinear energy-loss effects associated with bremsstrahlung.

After this track candidate refit, TopoClusters are used to form **superclusters**. In the first stage, available TopoClusters in each event are ordered by E_T , using the EM scale, and each one in descending order is individually evaluated for use as a *seed cluster*. For a cluster to be designated as a seed, it must have a minimum E_T of 1 GeV and be matched to a track that has at least four pixel or SCT hits.⁴⁰

Once a cluster meets the seed criteria, the algorithm searches for *satellite clusters*. A cluster is considered a satellite if it falls within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycenter, which typically indicates secondary EM showers from the same electron. A cluster can also be a satellite if it falls within a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ around the seed cluster barycenter and shares the best-matched track with the seed cluster. The seed clusters and their associated satellite clusters together form superclusters.

The final step involves assigning calorimeter cells to each supercluster. Only cells from the presampler and the first three LAr calorimeter layers are considered⁴¹. To reduce sensitivity to pileup noise, the width of each TopoCluster is limited to a maximum of 0.075 or 0.125 in the η direction in the barrel or endcap regions, respectively. As interactions with detector material typically cause EM showers to spread in the ϕ direction due to the magnetic field orientation, no restriction is applied. The final superclusters are then matched to GSF tracks again using the same procedure as the initial matching with the original TopoClusters.

2.3.5.2 Electron Identification and Calibration

As the majority of electrons of interest for physics analysis are so-called “prompt electrons” produced in the primary hard scatter interaction, rather than background electrons from photon conversions in detector material, misidentified jet signatures, and secondary electrons produced in heavy-flavor hadron decays, several **electron identification WPs** are defined [93]. These WPs use a **likelihood-based (LH)** discriminant constructed from the product of n probability density functions for a set of observables $\mathbf{x}$, each entry a specific quantity with useful discrimination power, for both signal (L_S) and background (L_B):

⁴⁰A cluster also cannot serve as a seed if it has already been assigned as a satellite cluster to another seed cluster.

⁴¹Except in the transition region of $1.4 < |\eta| < 1.6$, where energy from the scintillator between the calorimeter cryostats is also included

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^n P_{S(B),i}(x_i). \quad (2.22)$$

For each electron candidate, a discriminant d_L is calculated as:

$$d_L = \frac{L_S}{L_S + L_B}. \quad (2.23)$$

The electron LH identification is based on this discriminant, where d_L peaks sharply at unity for signal and zero for background, making it difficult to set operating points. Therefore, an inverse sigmoid function is applied to transform the discriminant distribution:

$$d'_L = -\tau^{-1} \ln(d_L^{-1} - 1), \quad (2.24)$$

where the τ parameter is fixed at 15. The signal pdfs are obtained through a tag-and-probe method using recorded data from $Z \rightarrow e^+e^-$ events (for $E_T > 15$ GeV) and $J/\psi \rightarrow e^+e^-$ events (for $E_T < 15$ GeV). Background PDFs are derived from data events containing at least one electron, primarily consisting of dijet events. Input observables include the ratio of track momentum to calorimeter energy deposit, the angular distance between the cluster and the extrapolated track, the number of hits in each ID subsystem, TRT transition radiation probabilities, and dozens of other relevant quantities.

Based on this likelihood (LH) discriminant, the three WPs, *Loose*, *Medium*, and *Tight*, are defined. Each offers greater background reduction at the cost of reduced signal efficiency. The *Loose LH* WP achieves an average signal efficiency of 93%, while the *Medium LH* WP achieves 88% and *Tight LH* 80% at $E_T = 40$ GeV. The efficiency of all three WPs increases with higher E_T . In the energy range of $20\text{GeV} < E_T < 50$ GeV, the background rejection of the *Tight LH* WP compared to the *Loose LH* WP improves by a factor of 3.5.

Electron candidates, like the other physics objects, are subject to a series of calibrations and corrections. Corrections are applied to account for upstream energy losses before the calorimeters and differences between data and simulation. Additionally, similar corrections as to the jet calibrations described previously are applied to correct the energy scale of the calorimeter deposits as well as the series of *in situ* calibrations that correct for pileup effects, η intercalibration, inhomogeneities in calorimeter response, etc. Systematic uncertainties are defined and associated with each of these calibrations.

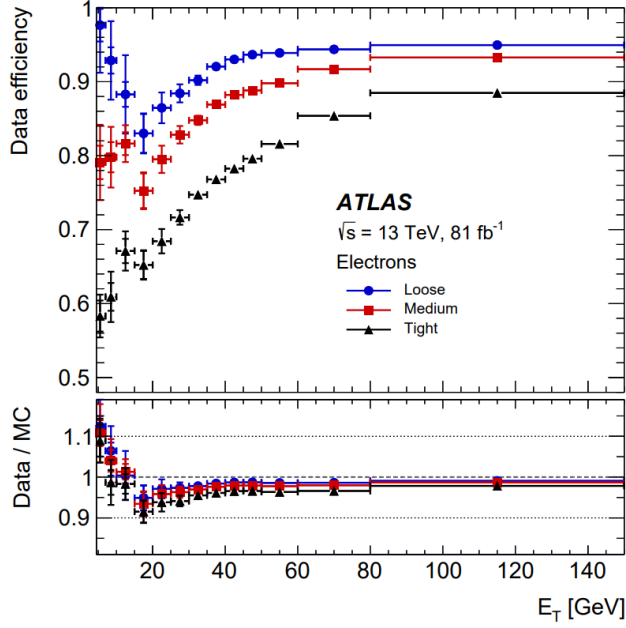


Figure 2.50. : The electron identification efficiency in $Z \rightarrow ee$ events in data as a function of E_T for the Loose, Medium, and Tight operating points. The error bars represent both statistical and systematic uncertainties. Taken from [94].

2.3.6 Missing Transverse Momentum

As neutrinos do not interact on the scale of the ATLAS detector volume, they escape entirely undetected by any subsystem. However, momentum conservation in the transverse plane implies that the sum of all transverse momenta from all real particles in a collision final state should sum to approximately zero. Thus, any observed imbalance after reconstructing all other categories of final state objects, known as **missing transverse momentum**, can be used as a proxy for neutrinos or other (BSM) invisible particles [96, 97]. This $E_T^{\vec{\text{miss}}}$ is defined for the x and y components from the corresponding components of the transverse momentum vectors ($\vec{p}_T$) like

$$E_{x(y)}^{\text{miss}} = - \sum_{i \in \text{hard objects}} p_{x(y),i} - \sum_{j \in \text{soft signals}} p_{x(y),j}, \quad (2.25)$$

where then the magnitude is simply

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}. \quad (2.26)$$

The hard term encompasses all reconstructed, calibrated final state objects detailed above, including electrons, photons⁴², τ -leptons, muons, and jets. As all of these objects are reconstructed separately, an ambiguity resolution algorithm is run to perform an overlap removal, as well as some minimum selection quality criteria on each object type to be included in the calculation. The soft term includes all tracks matched to the hard scatter vertex, passing some loose quality requirements, that are not associated with another hard object.

As E_T^{miss} is dependent on the accurate and complete reconstruction of every other type of object, fake contributions to E_T^{miss} can arise from interacting particles that are outside the acceptance of the detector or are inaccurately reconstructed. An object-based E_T^{miss} significance variable [98] is constructed that can help distinguish between whether it is more likely that the E_T^{miss} is due to a real invisible particle or missed reconstruction effects, taking into account the expected detector resolutions and reconstruction efficiencies. This is computed using an event-by-event likelihood calculation, testing the hypothesis that the total transverse momentum carried by invisible particles (p_{inv}^T) is zero against the alternative hypothesis that p_{inv}^T is different from zero. Given the likelihood function $\mathcal{L}(E_{\text{miss}}^T | p_{\text{inv}}^T)$ for the 2D parameter p_{inv}^T at a given measured E_{miss}^T , which depends on the types, kinematics, and multiplicities of the objects measured in each event, the log-likelihood ratio and significance $\mathcal{S}$ can be defined as:

$$\mathcal{S}^2 = 2 \ln \left(\frac{\max_{p_{\text{inv}}^T \neq 0} \mathcal{L}(E_{\text{miss}}^T | p_{\text{inv}}^T)}{\max_{p_{\text{inv}}^T = 0} \mathcal{L}(E_{\text{miss}}^T | p_{\text{inv}}^T)} \right), \quad (2.27)$$

which is calculated by maximizing the likelihood parameters under the constraints $p_{\text{inv}}^T \neq 0$ and $p_{\text{inv}}^T = 0$. The event-by-event likelihood function is then calculated under the following assumptions:

- The measurement of each reconstructed object i is independent from each other object in the event;
- For each object, the probability distribution of measuring $p_{T,i}$ given the true transverse momentum value associated with the hard physics process $\pi_{T,i}$ follows a Gaussian probability distribution with covariance matrix V_i ;
- Due to conservation of momentum in the x-y plane, $\sum_i \pi_{T,i} = -p_{\text{inv}}^T$.

Under these assumptions, the likelihood function takes the form of a 2D Gaussian:

⁴²We didn't discuss photons in detail because they are not used in either the Wh resonance search or the tracking software optimization. They follow a nearly identical process to electron reconstruction but without requiring a track (or requiring exactly two from pair production). Electron and photon reconstruction are intertwined and are managed by a single working group in ATLAS.

$$\mathcal{L}(E_{\text{miss}}^T | p_{\text{inv}}^T) \propto \exp \left[-\frac{1}{2} (E_{\text{miss}}^T - p_{\text{inv}}^T)^T \left(\sum_i V_i^{-1} \right) (E_{\text{miss}}^T - p_{\text{inv}}^T) \right]. \quad (2.28)$$

The log-likelihood ratio becomes a χ^2 variable with two degrees of freedom:

$$\mathcal{S}^2 = 2 \ln \left(\frac{\mathcal{L}(E_{\text{miss}}^T | E_{\text{miss}}^T)}{\mathcal{L}(E_{\text{miss}}^T | 0)} \right) = (E_{\text{miss}}^T)^T \left(\sum_i V_i^{-1} \right) E_{\text{miss}}^T. \quad (2.29)$$

Here, the subscript i denotes each reconstructed object contributing to the E_{miss}^T calculation and V_i is the corresponding covariance matrix. Small values of $\mathcal{S}$ indicate that E_{miss}^T is likely fake or consistent with reconstruction and resolution effects, while large values suggest it is likely genuine E_{miss}^T .⁴³

⁴³For a more in-depth discussion about likelihood ratios and related statistical concepts, please see Chapter 4.

CHAPTER 3

OPTIMIZATION OF THE ATLAS TRACK RECONSTRUCTION

As we look towards the future of the ATLAS physics campaign, particularly in view of the HL-LHC, many concerns have been raised regarding the overall resource consumption of the ATLAS computing model, regarding both budgetary and environmental concerns as well as physics performance. Projections show [99] that without significant budgetary increases and aggressive research and development, both the CPU and disk resources required will outstrip those available by the beginning of Run 4 datataking (2029), as shown in Figure 3.1. The average pileup is expected to reach up to $\langle\mu\rangle = 200$, recording data at a rate 7-10 times greater than current datataking. An open question remains: how to develop and innovate new software techniques to reconstruct these dense, highly complex topologies all while collecting more data than ever before - in other words, how can ATLAS do *more* with *less*?

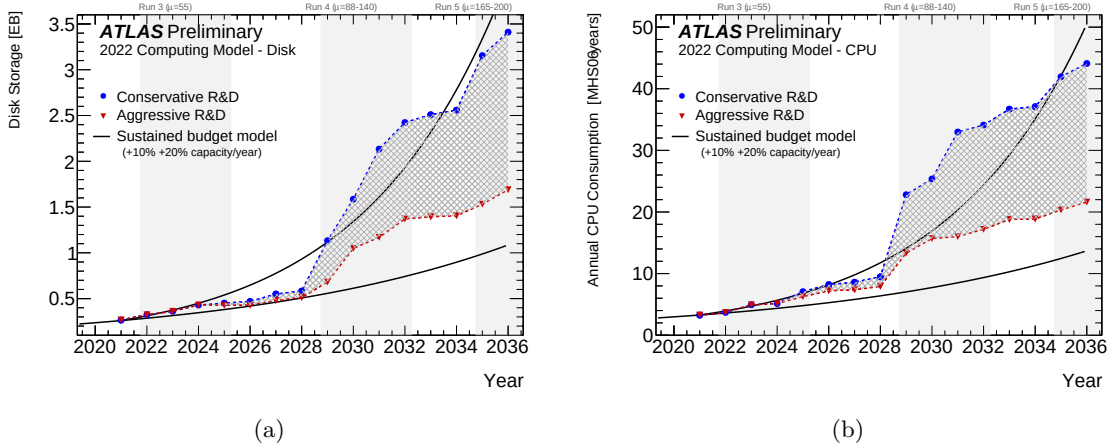


Figure 3.1. Projected evolution of ATLAS CPU and disk usage from 2020 until 2036, under the conservative (blue) and aggressive (red) R&D scenarios. The black lines indicate the impact of sustained year-on-year budget increases and improvements in new hardware, which together amount to a capacity increase of 10% (lower line) and 20% (upper line). The shaded bands indicate datataking periods. Taken from [99].

On a smaller scale than the HL-LHC, Run 3 had a much higher design pileup than previous datataking years, with peak values around $\langle\mu\rangle = 60$ compared to $\langle\mu\rangle = 30$ in Run 2 (see Section

2.1.3). As such, the Run 3 version of the reconstruction software needed to be much faster and lighter on storage than the legacy Run 2 version, necessitating an effort to target the most significant consumers.

Track reconstruction, and the ATLAS reconstruction chain in general, has been extensively documented in the previous Chapter. ID track reconstruction is a central element of the ATLAS physics program, serving as the foundation of the vast majority of physics objects used throughout ATLAS analyses, from precision measurements to searches for physics beyond the Standard Model. In an average ATLAS event with $\langle\mu\rangle = 70$, up to 15,000 silicon hits need to be processed into high-quality track candidates. Track reconstruction is thus fundamentally a combinatorial problem and requires exponentially more computational resources as pileup grows linearly. Additionally, the high density of hits leads to incorrect cluster-to-track association, potentially degrading the performance of the reconstructed track parameters and properties relative to their true values, and collections of unrelated clusters can be combined into combinatorial “fake” tracks, which also grow exponentially with pileup.

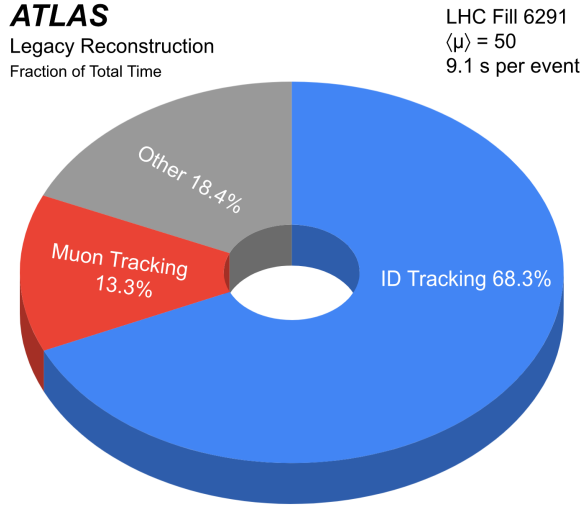


Figure 3.2. Fractions of the total CPU consumption of the full ATLAS Run 2 reconstruction for one data run at $\langle\mu\rangle = 50$.

Due to these factors, ID track and vertex reconstruction is the most computationally intensive portion of the entire ATLAS reconstruction chain, accounting for nearly 70% of the total CPU time per event during Run 2 as seen in Figure 3.2. Optimizations in ID tracking, therefore, scale almost directly to an optimization of the total ATLAS event reconstruction and yield the most significant possible improvements. This Chapter will present a large-scale effort to optimize the ID track

reconstruction in advance of Run 3, improving resource consumption while maintaining high-quality physics performance. It will first discuss a highly abbreviated overview of some critical components of the ATLAS software and computing infrastructure. Then, it will describe the optimization effort for Run 3 with a suite of benchmarks and performance comparisons. For a more comprehensive overview of ATLAS software and computing in Run 3, please see Ref [100]. The results and optimization methodology portions of this Chapter have been previously published in *Software Performance of the ATLAS Track Reconstruction for LHC Run 3* [17], supplemented with additional context and discussion.

3.1 ATLAS Software and Computing Infrastructure

An experiment as large and sophisticated as ATLAS requires an enormous system of software and computing infrastructure to support its operation, reflecting the overall complexity of its physics program. This section will discuss a brief overview of both the distributed computing resources and the software suites used for data reconstruction, detector operation, and physics analysis.

3.1.1 Computing Resources

ATLAS computing resources are distributed worldwide, and different data centers, or **sites**, are employed for a specific subset of use cases. Combined, they total nearly one million computing cores, 350 PB of disk space, and almost 450 PB of archival tape storage, all used to process and store the total volume of collected and simulated data produced by ATLAS.

3.1.1.1 Tier-0 and HLT

The central **Tier-0 (T0)** computing site, located at CERN, is where prompt data processing directly after collection is done, as well as where nightly software releases are built and tested. This prompt processing consists of the first “express” data reconstruction pass as well as most of the object calibration workflows, which rely on the initial results of the express processing. It contains around 40,000 cores in an HTCondor-managed batch system and can temporarily store intermediate datasets with 2PB of disk space, with around 100 GB of persistent disk space to store the T0 software and other files needed for operation [100]. The express reconstruction software shares commonalities with the offline reconstruction software used throughout the rest of the experiment, but the overall T0 operational software is designed to be stable and self-contained to minimize service interruptions. The T0 software is not the focus of the tracking optimization work, though the online

track reconstruction run for the trigger at the nearby HLT computing farm also ultimately inherits some of the optimizations.

3.1.1.2 The Worldwide LHC Computing Grid (WLCG)

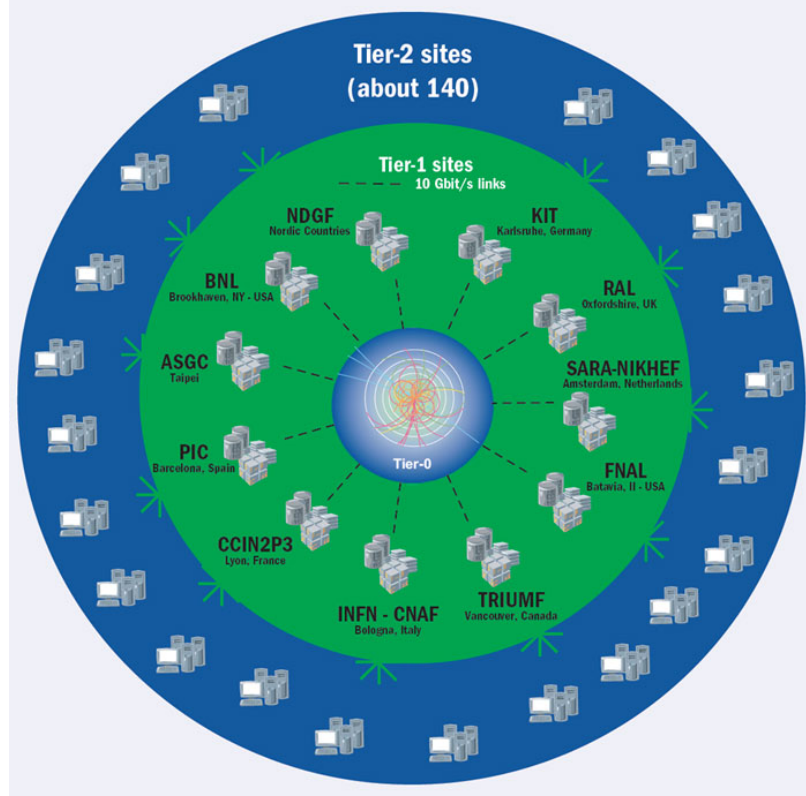


Figure 3.3. Illustration of the WLCG Tier-0, -1, and -2 sites. Additional Tier-2 sites have been added since this image was made in 2013, and not all Tier-2 sites are available to ATLAS. Taken from [101]

Once passed through the initial T0 processing, ATLAS data are further processed and stored over a worldwide network of sites known as **The Worldwide LHC Computing Grid (WLCG)**, colloquially called the **Grid** by the collaboration. These sites are organized into a hierarchical **tier** structure, as depicted in Figure 3.3, each with varying capacities, roles, and responsibilities. The 11 ATLAS **Tier-1** sites are linked to CERN via dedicated national research and education networks, typically using optical private networks with bandwidths ranging from 10 to 100 Gbit/s [102]. Tier-1 sites are tasked with providing both disk and tape storage, ensuring the long-term preservation of detector data. Approximately 70 ATLAS **Tier-2** sites, primarily located at national universities and laboratories, offer disk storage for data processing and user analysis. In addition, several other computing centers contribute to ATLAS’s storage and computing capacity, including

“opportunistic“ resources and sites leveraged for end-user data analysis, such as cloud computing providers and general-purpose high-performance computing (HPC) centers. As illustrated in Figure 3.4, ATLAS had over 800 PB of data in 2023, evenly divided between tape and disk storage.

The **Rucio** [103] distributed data management system is utilized for organizing, managing, transferring, accessing, and tracking all of this data, handling an average of 1.5 million file transfers per day for a total of 8PB of moved data volume on average [100]. ATLAS datasets follow a specific nomenclature that details their properties in a condensed form. This includes a dataset ID number identifying the run or simulated process, the stage in the processing workflow (detailed in the next section), and a “tag” corresponding to the software release and configuration it was processed with, among other critical information. Some specific operations done on a daily basis with Rucio include registering new files and datasets, searching for datasets, downloading specific files for processing jobs, modifying metadata, and scheduling obsolete files for deletion.¹

The Grid resources, while sizable, are ultimately finite. Since the volume of total data collected, processed, and stored by ATLAS only ever increases, maintaining the efficient performance of the reconstruction software is critically important for the continued successful operation of the ATLAS computing model. Minimizing unnecessary usage of computing resources frees these resources for the rest of the activities of the collaboration, speeding up analysis workflows² and allowing for the development of new techniques that may otherwise be computationally prohibitive.

3.1.2 Software Resources

An extensive software suite [104] is used throughout the ATLAS experiment, including the reconstruction and analysis of real and simulated data, detector operations, and the trigger and data acquisition systems. The terms *online* and *offline* are used to differentiate between software and services used during real-time datataking for the trigger and data acquisition systems and those used for subsequent data processing and analysis, respectively. The work presented in this Chapter concerns the offline reconstruction software, contained within the ATLAS **Athena** project (Section 3.1.2.2) [105].

¹This does not mean deleting RAW data, which is kept indefinitely in archival tape storage. Rather, outdated processings no longer in use are eventually obsoleted to free disk space, as datasets are *immutable* for reproducibility reasons and cannot be modified once produced. As all RAW data, Athena releases, and tags are archived, any dataset is always (in theory) completely reproducible.

²Less time spent on reconstruction means, for example, analysis teams do not have to spend valuable time waiting weeks or months for their samples!

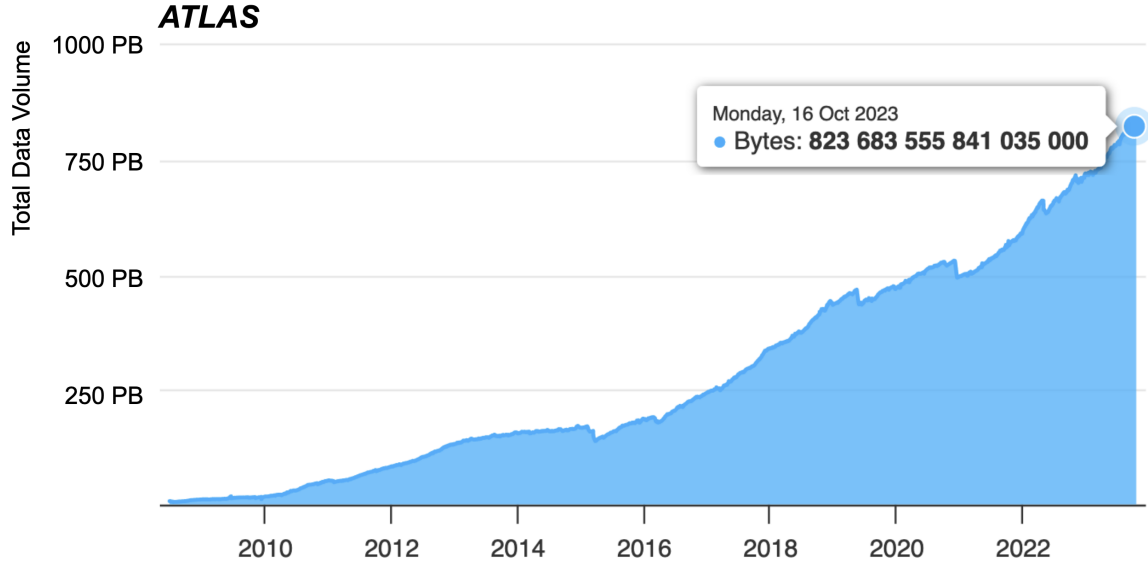


Figure 3.4. Evolution of the total ATLAS data volume since before Run 1. Taken from [100]

3.1.2.1 Data Processing Workflows

The software workflows for both MC simulation production and collision data processing are shown schematically in Figure 3.5. Both sets of inputs are passed through a series of processing and reconstruction steps, ultimately resulting in the final data formats used in physics analysis.

MC simulation production begins with the event generation, simulation, and digitization steps outlined previously in Section 1.8. The event generator creates HepMC [106] output, which is then written to an **EVNT** file, which contains the collection of generated particles alongside metadata storing the details of the generator job configuration. These resulting particles are then passed through the detector simulation, where a **HITS** file is created that includes records of energy deposits from each particle alongside timing information for each subsystem, **truth** information about the simulated particle during the event, and additional metadata related to the simulation job’s configuration. Next, these simulated energy deposits are processed through a digitization step, where the detector electronics are modeled, resulting in **Raw Data Object (RDO)** output, which is similar to the actual raw data collected by ATLAS. During digitization, pileup is also modeled by adding background events overlaying additional simulated minimum bias events. The RDO files also carry the truth records from the simulation and include additional records correlating generator-level particles with detector signals, as well as further configuration metadata. These RDO files are also then processed through a trigger simulation stage, where decisions emulating both L1 and HLT triggers, as well as simulated trigger object collections, are appended to the output. The RDO Trigger

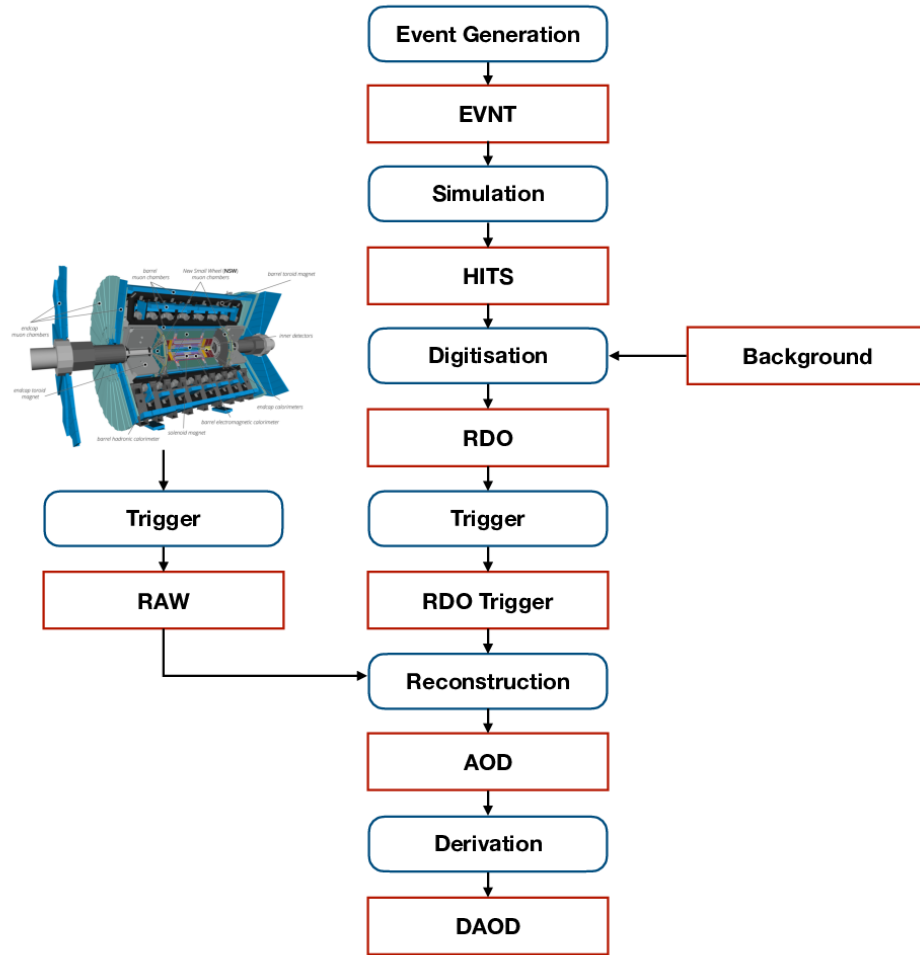


Figure 3.5. Flowchart of the ATLAS software workflow. Processing steps are shown in blue, while output formats are shown in red. Taken from [100].

output is now functionally the same as the detector data and can be passed to the reconstruction chain.

After passing the trigger decision, raw detector data is promptly reconstructed at the Tier-0 facility before being distributed to the Grid for further analysis, *derivation* production, and reprocessings. Data are organized into **streams** defined by a specific set of trigger selections that contain all passing events recorded to disk. The major streams³ include

- **Physics** streams, which contain the primary data that will be used in physics analysis;

³Many other specialty streams exist, not detailed here for brevity.

- The **Express** stream, which contains a slimmed, representative subset of the physics streams;
- **Calibration** streams, which are designed to provide specific additional statistics to be used in certain calibrations.

Before any processing, all data is stored in **RAW** files, which use a custom *bytestream* format. The flow of data from collection to the reconstruction steps in the various streams is shown in Figure 3.6. Before the primary bulk processings of the main physics streams, a **calibration loop** begins where the express and calibration streams are initially processed at T0. This calibration loop uses the express processing to update calibration constants and detector conditions iteratively. These include but are not limited to updating the precise Inner Detector alignment model, running a fit to the results of the initial primary vertex finding to determine the precise ATLAS beamspot position (such that it can be later used in the bulk processing to better determine the primary vertex positions), and locating corrupted, disabled, and noisy modules for a given run such that they can be accounted for in the bulk processing. Further details on the calibration loop can be found in Ref. [107]. Once finished, the improved conditions information is passed to the **conditions database** for use in the subsequent processings. In addition to the initial bulk processing after the express processing at T0, physics data is repeatedly reprocessed in varying campaigns in the years after its initial collection. This is motivated by improved knowledge of detector conditions and geometry, such as updates to dead material maps and minor alignment corrections, and performance updates to the reconstruction algorithms, such as the ones described in this Chapter.

Once both the RAW and Trigger RDO files are available, reconstruction begins, turning real and simulated detector signals into physics objects with the processes described extensively in Chapter 2. This output is first written to an intermediate **Event Summary Data (ESD)** file⁴ and then written out to an **Analysis Object Data (AOD)** file, the primary data type used throughout ATLAS. These files contain a summary of the reconstructed event with all its constituent physics objects written into the xAOD object format, which is readable by both the ATLAS Athena software and ROOT [108], making it portable and adaptable for end-user analysis.

Physics analysis is generally performed on an additional **derivation** reconstruction step (**DAOD**) produced from the primary AODs. The main function of this step is to reduce the total data output size that needs to be stored and manipulated by end-users by removing unused information, which can be done in a combination of three operations:

⁴These files are enormous and are generally only used for specific reconstruction studies where low-level cluster etc. information is needed, not physics analysis.

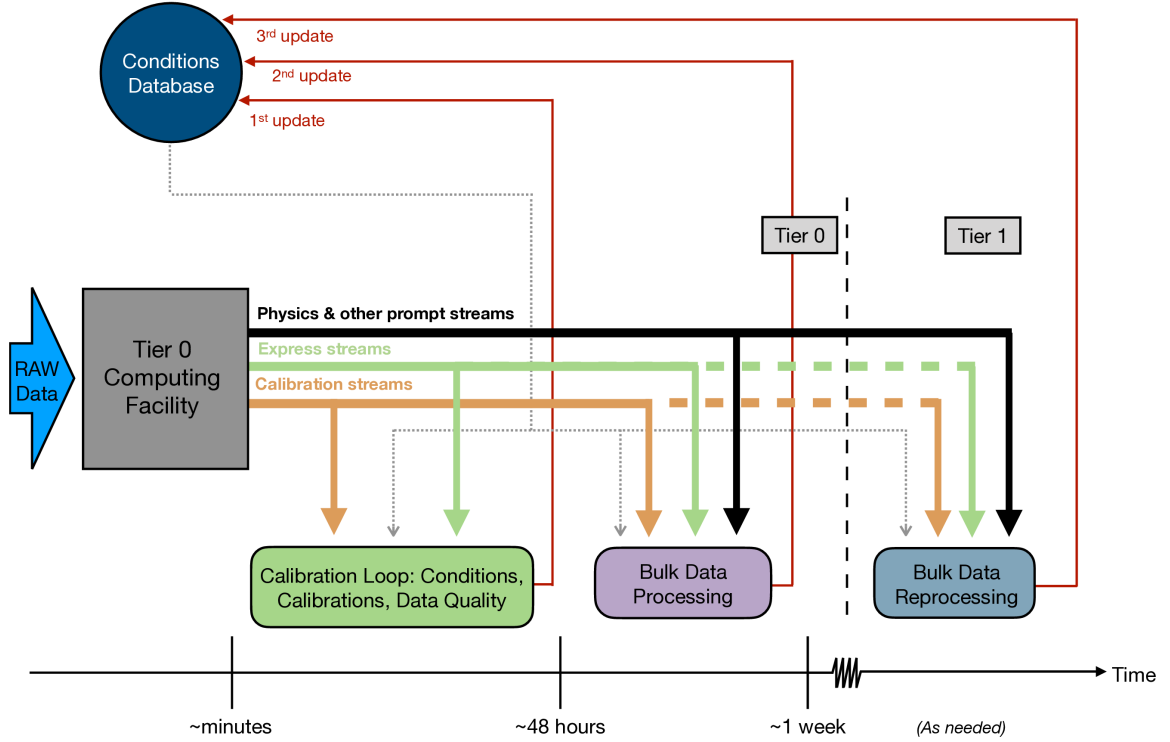


Figure 3.6. Flowchart illustrating the flow of ATLAS data from initial reconstruction during the calibration loop followed by the bulk processings at the Tier-0 and Tier-1 sites. Taken from [100].

- **Skimming:** This involves removing entire events based on specific criteria related to the characteristics of the event.
- **Thinning:** This process removes individual objects within an event based on criteria related to the characteristics of those objects.
- **Slimming:** This entails removing certain variables within a given object type uniformly across all objects of that type and across all events. Unlike skimming and thinning, slimming does not depend on any event or object properties, as the same variables are removed for every event and object.

The DAOD production also includes the reconstruction of secondary physics objects, such as secondary vertices and other specialized items that use the primary objects as input – in particular, in Run 3 jet reconstruction and b -tagging runs in this step rather than in the primary reconstruction step. Output files are also merged to distribute an optimal number of events to each file.

The *resource* performance studies shown in this Chapter begin with RAW data as the software performance of the reconstruction itself on data is being evaluated, where specifically the evaluated elapsed time is for the “RAWtoALL” chain, which processes RAW data first to an intermediate ESD before creating the output AOD. The results demonstrating the impacts on *physics* performance are evaluated at the AOD level after the reconstruction is complete on both data and simulation.

3.1.2.2 Athena

The ATLAS offline software framework, Athena [105], contains the entire codebase used for offline reconstruction and analysis tasks. It is an open-source software framework, available on GitLab [109], under the APACHE 2.0 license [110] with the copyright held by CERN on behalf of the ATLAS collaboration. Every ATLAS collaboration member can contribute to the software via merge requests to the active branches of the GitLab repository, which are reviewed by a team of shifters and experts to ensure that only high-quality code is merged into production. Additionally, merge requests are required to pass a suite of Continuous Integration (CI) tests representing most of the major tasks performed by the ATLAS software to ensure no unintended changes to the physics output are introduced. The codebase is organized into around 2000 packages, each containing code with common functionality or goals [100]. It additionally depends on approximately 200 external general high-energy physics, data science, and Linux software packages. Athena’s codebase includes over 50,000 unique files, predominantly written in C++, configured using PYTHON, and built with CMAKE [111].

Athena is based on the older Gaudi project [112], which was developed in collaboration with LHCb and other experiments. An Athena application comprises dynamically loadable *components* implementing **Algorithms**, **Services**, and **Tools** as shown in Figure 3.7. These have the following properties:

- **Algorithms:** These components process data residing in a shared **event store**. They read objects from the store (identified by type and a string key) and write new objects back to the store. Ideally, Algorithms only communicate with the event store and do not contain data themselves or interact with other Algorithms.
- **Services:** These are objects utilized by multiple other components. Examples include the event store itself, error logging, random number generation, etc.
- **Tools:** These serve as helpers for other components and may be owned by Algorithms, Services, or other Tools.

All three component types can declare **Properties** to be initialized during job configuration. Athena originally only ran in a serial mode, where one event is processed at a time, and the order of Algorithm execution was fixed in the initial PYTHON job configuration. For Run 3, the software was migrated to be thread-safe, where Algorithms can run in parallel in separate threads. The different **AthenaMT** Algorithms can process the same event data in parallel, and the same Algorithm can process different events in parallel, depending on the job configuration. Each Algorithm declares the data objects it reads from and writes to the event store, enabling the construction of a dependency graph during application initialization. In addition to 'data flow' information, 'control flow' rules can explicitly state the sequence in which specific Algorithms must run, allowing for early termination of event processing if certain conditions are not met.

Data dependencies for Algorithms and Tools are declared through their data access mechanisms. To read or write an object of type T from the event store, an Algorithm declares a Property of type `SG::ReadHandleKey<T>` or `SG::WriteHandleKey<T>`. During execution, these key objects can be combined with an `EventContext` object to form smart pointers for accessing data objects. These declarations help build the data flow graph used by the scheduler. Tools declare dependencies similarly, and these dependencies propagate to the owning Algorithm.

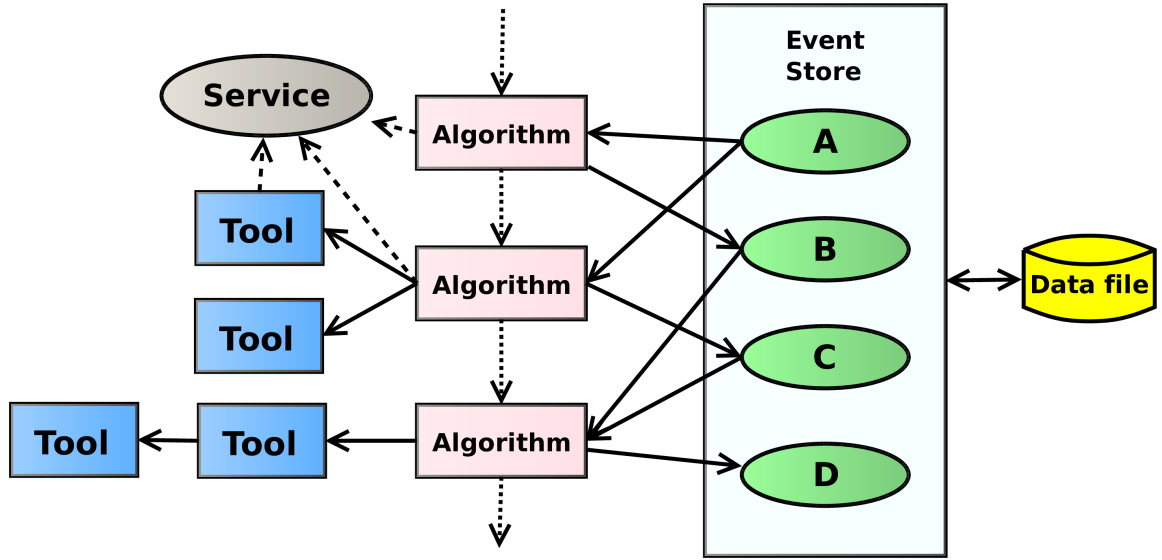


Figure 3.7. Flowchart of the structure of an Athena application. The solid lines on the left indicate ownership, while those on the right indicate the flow of data. The dotted line down the center indicates the control flow from Algorithm to Algorithm. The dashed lines indicate a non-owning reference. Taken from [100].

Returning to the discussion of the track reconstruction software, the various stages of the process each are run by a specific Algorithm that takes in the products of the previous stage and produces

its own output. As one particular example, the track seeding and finding stage is performed by the Algorithm `SiSPSeededTrackFinder`. This Algorithm takes in both the `SpacePointContainers` with keys `SCT_SpacePoints` and `Pixel_SpacePoints` from the event store, then initializes and owns several Tool instances used during the Algorithm such as `SiSpacePointsSeedMaker` and `SiTrackMaker`. After the Algorithm is complete, it produces a `TrackCollection` (aliased to an `xAOD::TrackParticleContainer`, discussed in the next section) with the key `SiSPSeededTracks`, which is then released to the event store for the ambiguity-solving Algorithm to process, which produces its own output and so on until the final track collections are created.

3.1.2.3 Event Data Model

Processed data must be represented in a way that maintains common access interfaces and internal consistency throughout the collaboration. The ATLAS **Event Data Model (EDM)** is a set of C++ classes that define the structure of the data objects used in the event store. The EDM is designed to be flexible and extensible, allowing for the addition of new data objects and the modification of existing ones. The object stores, such as the event store and conditions store, are implemented via instances of the Service `StoreGateSvc`. This service provides a type-safe mapping of string-based identifiers to arbitrary C++ objects. Event data containers are typically represented by the `DataVector<T>` type, similar to `std::vector<T*>` but with several additional functionalities, including optional ownership of constituent elements and auxiliary data store of an arbitrary type which can be attached to the elements of the vector.

For objects intended for analysis, nearly all data are stored as **auxiliary** data rather than class members, referred to as **xAOD objects**. For example, `xAOD::Electron` objects are stored in `DataVector<xAOD::Electron>`, aliased as `xAOD::ElectronContainer`, with static data in an `xAOD::ElectronAuxContainer` implementing the `IAuxStore` interface. An `AuxContainer` class can also forward requests for variables it does not contain to another `IAuxStore` instance. The EDM supports the dynamic addition of variables, called **decorations**, allowing xAOD objects to have a set of core variables with the flexibility to add more as needed, as depicted in Figure 3.8. The scheduler is aware of decorations read and written by Algorithms, factoring them into scheduling decisions. Auxiliary store objects are saved separately, e.g., the auxiliary store for `Electrons` is named `ElectronsAux`. The `xAOD::Object` class is a base class that defines common properties and methods for all data objects in the EDM, containing methods for accessing and modifying the properties of the data object as well as for performing common operations on the data object.

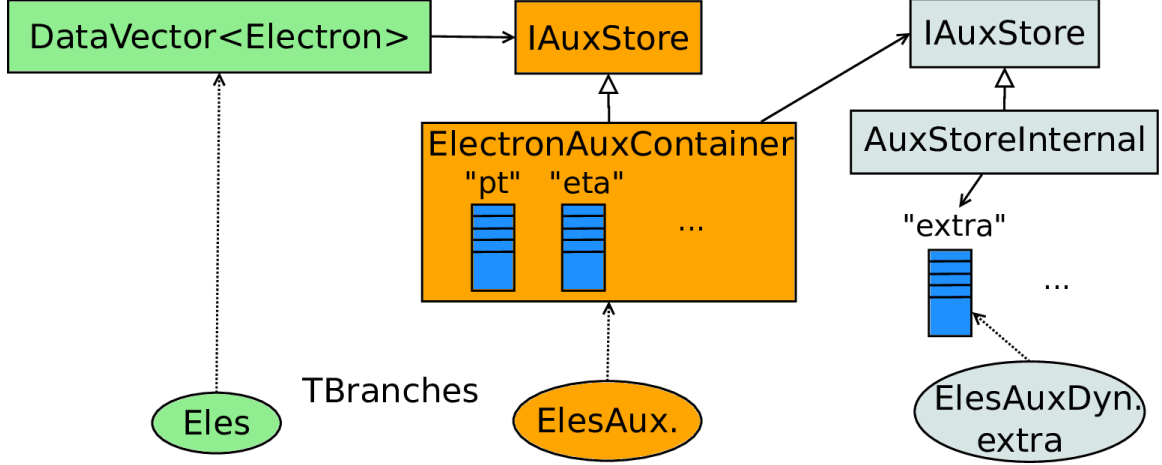


Figure 3.8. Flowchart of the static and dynamic stores used for xAOD objects. This figure shows a container of electrons stored in the event data store named **Eles**, which is also used as the name of the ROOT TBranch when the objects are saved to ROOT. The associated auxiliary store contains the data members **pt**, **eta**, and so on. The dynamic auxiliary store on the right contains extra members attached (decorated) to the container. Solid lines with open arrows indicate inheritance, plain solid lines show references between the objects, and dotted lines map between names and objects in the event store. Taken from [100].

Again, returning to track reconstruction, the default container for the combined final standard tracks in ATLAS is `InDetTrackParticles`, a `TrackCollection` which is aliased to a `xAOD::TrackParticleContainer`. The `xAOD::TrackParticles` in each container have associated `xAOD::SummaryType` values in their `AuxContainer`, which contains an extensive collection of member variables from the track reconstruction process itself such as `numberOfContribPixelLayers` and `numberOfSCTHoles`. This highly detailed tracking information, and their analogs in other types of physics objects, is not generally used directly by analysis end-consumers and is instead “baked in” to the centrally provided recommendation and selection tools developed by the respective Combined Performance working groups. This auxiliary summary information written to the individual track objects via the reconstruction algorithms was continually monitored during the optimization campaign, detailed in the next section.

3.1.2.4 Builds, Releases, and Monitoring

Athena is compiled and built into numbered **stable releases** on a regular schedule, representing the centrally validated state of the codebase at a particular point in time. The **Nightly** System, separately, performs daily overnight builds of all ATLAS software projects from the Athena GitLab repository on multiple platforms. This system supports a parallel development environment with numerous simultaneous branches, evaluating how changes from accepted merge requests integrate.

The Nightlies are compiled on a 1400-core build farm with a dynamic monitoring system that hosts information about build results and installation status, held in the central ATLAS Nightlies database.

The system builds releases from the GitLab **main** (**master**) branch and from separate dedicated branches for online operations, legacy branches for older versions of the reconstruction, and production branches for simulation and analysis. Changes from these branches are “swept” on a routine basis to the **main** branch. Special development Nightlies also exist that use new versions of ROOT, GCC, and other externals to prepare for upcoming changes. All production releases until fall 2023 were built on the CENTOS7 Linux operating system, and those after were migrated to Alma Linux 9, with some legacy releases still built on CENTOS7.

Upon successful completion of nightly build compilation and tests, CPACK [111] generates RPMs, which are uploaded to EOS [113] disk for staging. These are then fetched and installed using AYUM, an ATLAS-specific YUM wrapper, in the **atlas-nightlies** repository on the CernVM FileSystem (CVMFS) [114], making them globally accessible to developers. CVMFS distributes all ATLAS development and production software. Rigorous testing of nightly releases is conducted using the ATLAS Release Tester (ART) Grid-based framework [115]. Nightly releases successfully meeting development goals are transformed into the stable releases, assigned unique numeric identifiers, and maintained indefinitely. The numbering pattern A.B.C(D) is used, where A indicates the major release version, B indicates the release purpose, C is the minor release version, and D is an optional patch version. Patches are applied as needed to fix bugs or read new file versions while maintaining compatibility with the release used during data taking. In the following Chapter, “Legacy” refers to the last applicable *Release 21* version of Athena, used for the Run 2 reconstruction from around 2018 up until 2022, while “Updated” refers to the applicable *Release 22* version, used to reconstruct the first Run 3 datataking, at the time of evaluation.⁵

In addition to the ART tests, the ATLAS Software Performance Optimization Team (SPOT) monitors the performance of the various Nightly builds, storing and displaying the results on a central monitoring page. The **SPOT tests** monitor the *resource* performance of the Nightlies on a single consistent machine, rather than the Grid as in ART, with a menu of tests that are representative of the different types of reconstruction and simulation tasks that are performed throughout the collaboration. Monitored variables include total CPU consumption per Algorithm, total CPU

⁵As of May 2024, the latest major release version is *Release 24*, as the versioning schedule has been modified a bit to correspond better with the year. (It’s actually somewhat of a coincidence that “Release 22” ended up roughly corresponding with 2022!) The jump from *Release 21* to *22* was an enormous paradigm shift with the introduction of multithreading, the migration to Component Accumulator job configurations, large-scale changes in reconstruction, etc. The jump from *Release 22* to Releases *23* and *24* was much more routine.

utilization percentage, memory usage per thread, and rate of change of memory usage over time, among other quantities shown for a two-week period. A snapshot of the SPOT monitoring page is shown in Figure 3.9.

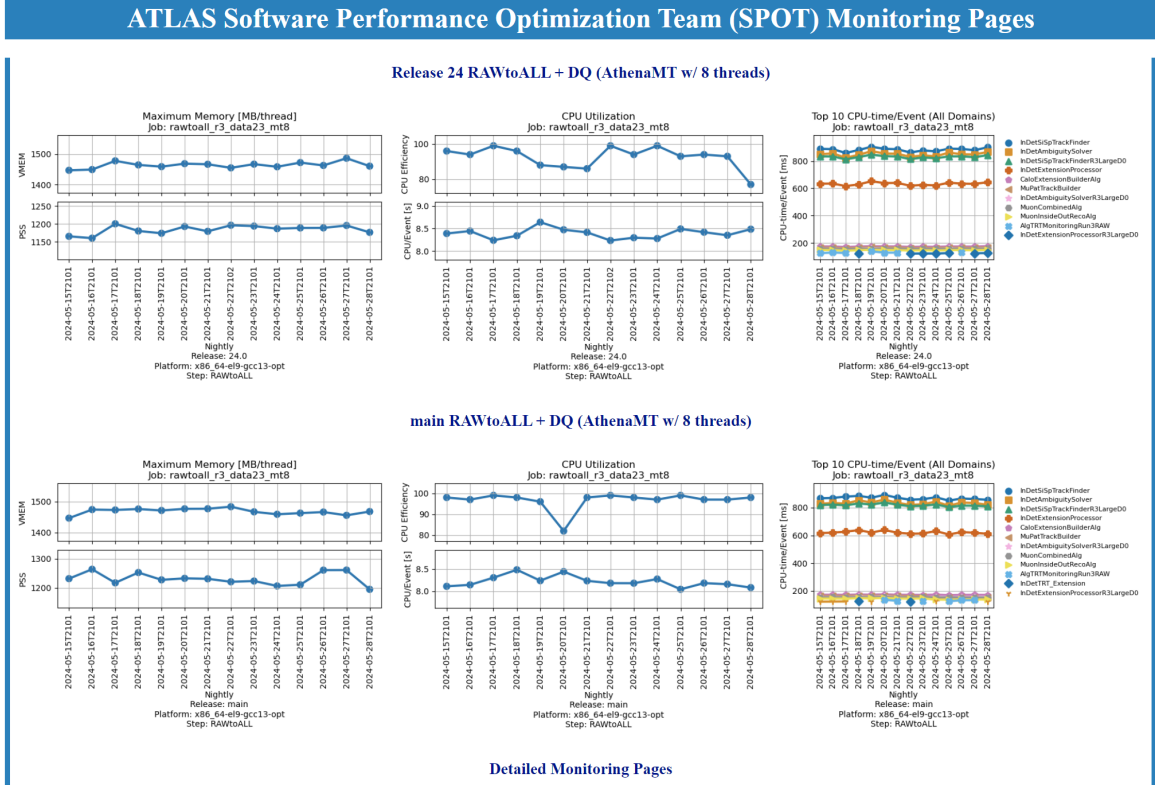


Figure 3.9. Snapshot of the SPOT monitoring webpage on May 29, 2024, highlighting the CPU and memory consumption of the primary reconstruction chain for both the *main* and *Release 24* Nightly release branches.

The algorithms consuming the most significant fraction of CPU are shown at the top right of the monitoring page. Several ID tracking algorithms appear, including *InDetSiSpTrackFinder*, *InDetAmbiguitySolver*, *InDetSiSpTrackFinderLargeD0*, and *InDetExtensionProcessor*, reflecting their status as the most resource-intensive portion of the ATLAS software chain. Prior to the optimization campaign, the existing SPOT infrastructure was only equipped to monitor the machine-level resource variables described above. However, the software package used to profile the performance, IDPERFMON, creates extensive log files as a byproduct, saved alongside the test results, where the track *SummaryType* values can be extracted. The **SPOTPlots** software package was developed in the early stages of the tracking optimization campaign, automatically parsing the combined outputs of these nightly SPOT logs into a portable data format. It presented the total CPU, event size, and relevant tracking summary statistics over time while indicating individual

merge requests in a user-friendly searchable web interface, allowing for centralized coordination of the optimization effort and complete preservation of performance results across multiple years of development within the Tracking Combined Performance working group, leading to the profiled results described in this Chapter. An example snapshot of the associated webpage is shown in Figure 3.10, and an example of the raw monitoring distributions is shown in Figure 3.11.



Figure 3.10. Example snapshot of the SPOTPlots monitoring page.

3.2 Data, MC Simulations, and Benchmarking Methodology

The performance of the ATLAS track reconstruction was evaluated both in real-time during the optimization campaign by the aforementioned SPOT profiling and in a more comprehensive study after the optimized version was complete and went into production. Results from both sets of information are presented in this work, and the samples and methodology are detailed below.

3.2.1 Data and MC

The impact of the software improvements discussed in the following sections on the reconstruction performance is evaluated using both pp collision data and MC simulations of top-quark pair production ($t\bar{t}$). To assess the data reconstruction performance for a given value of pileup, the raw data of a set of 500-1000 consecutive collision events from a common luminosity block (LB) are

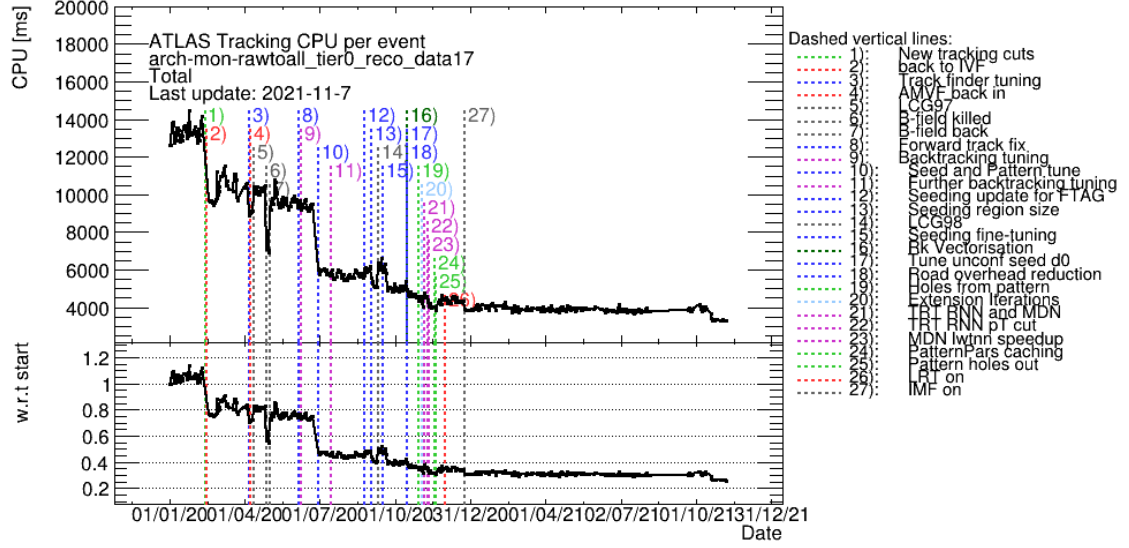


Figure 3.11. Example SPOTPlots distribution showing elapsed CPU vs time for reconstructing a set of 2017 data events, in both absolute time in ms (top) and a ratio with respect to the starting value pre-optimization campaign (bottom). Significant merge requests are labeled via the colored dashed lines with keys to the right of the figure.

reconstructed, and the processing time taken for the reconstruction, as well as the size of the output written to disk, are recorded. This procedure is repeated for sets of events at different pileup values, such that a full range of representative values is probed. The samples are taken from a single LHC fill (fill 6291) recorded towards the end of 2017 datataking and covering a range of average pileup values between $\langle\mu\rangle = 15.5$ to 60. Additionally, a set of samples from an LHC fill recorded during 2022 as part of Run 3 (fill 8112) with pileup values between 22.4 and 54.8 allows this information to be compared with the reconstruction of Run 2 data using the same software suite. Keeping within a single LHC fill ensures consistent datataking conditions across the tested pileup values, which can cause slight fluctuations in the reconstruction performance from run to run. Finally, a set of 300 events from LHC fill 5824, recorded in 2017 at $\langle\mu\rangle = 50$, is used in the nightly SPOT monitoring described in the previous section. The run times measured during this monitoring are used to evaluate the incremental speed-up of track reconstructions as sequential changes are made over time. All events from these two fills are part of the GRL and satisfy ATLAS data quality requirements. An additional LHC fill with up to $\langle\mu\rangle$ recorded in late 2018 (fill 7358) as part of a machine-development campaign is used in addition to extend the study towards even larger pileup values. Unlike the 2017 and 2022 data, this run is not part of the ATLAS physics dataset due to the nonstandard datataking conditions. However, since the ID was fully operational, it's possible to use

these events to extrapolate the scaling behavior of track reconstruction performance under extreme pileup conditions far beyond that of regular datataking conditions.

Monte-Carlo simulations of $t\bar{t}$ production are used to evaluate the fraction of successfully reconstructed charged particles. These are modeled using the POWHEG BOX v2 [116, 117, 118, 119] generator at next-to-leading order (NLO) QCD precision with the NNPDF3.0_{NLO} PDF set [120] and the h_{damp} parameter⁶ set to $1.5m_{\text{top}}$ [121]. The events are interfaced to PYTHIA 8.230 [122] to model the parton shower, hadronization, and underlying event using the NNPDF2.3_{LO} PDF set [123] and the A14 set of tuned parameters [124]. The decays of bottom and charm hadrons were performed by EVTGEN 1.6.0 [125]. Pileup effects were modeled by overlaying the simulated hard-scatter event with inelastic pp events generated with PYTHIA 8.816 [126] using the NNPDF2.3_{LO} PDF set [127] and the A3 tune [128]. To study the reconstruction performance at varying levels of pileup, the number of events overlaid for each simulated $t\bar{t}$ event was sampled from a flat distribution between 0 and 80.

3.2.2 Benchmarking Methodology

The benchmark studies described in this chapter were run as the only active user on a dedicated machine equipped with an AMD EPYC™7302 16-core processor, running the CERN CENTOS7 operating system. The processor was operated in “performance” mode, with simultaneous multi-threading (SMT) and frequency boosting disabled. An HS06 score⁷ of 22.6 per core was obtained for the processor in this configuration [129]. The machine was kept at a stable 50% capacity by running an appropriate number of reconstruction tasks simultaneously. To exclude the impact of multi-threading from the comparison, all tests were run in single-threaded mode.

The executable binary files tested are identical to those most commonly used for ATLAS’ regular production data reconstruction. As a result, different versions of the software differ not only in terms of their own programming but also in terms of the method of compilation as well as external libraries and their associated compilations. All software versions were compiled with the default compilation settings as they were defined in Athena at the time of their release. The exact details may differ between compiler versions since the optimization flags are partially set according to compiler presets.

⁶The h_{damp} parameter is a resummation damping factor and one of the parameters that control the matching of the POWHEG matrix elements to the parton shower, thus effectively regulating the high- p_T radiation against which the $t\bar{t}$ system recoils.

⁷HS06 is a standardized CPU benchmark used in high-energy physics, developed by the HEPiX Benchmarking Working Group. It is based on industry-standard benchmarks with modifications to match the overall average percentage of floating point operations in large experiment codebases.

However, the code is compiled for a generic x86-64 architecture in all cases. The legacy Run 2 reconstruction binaries were compiled using version 6.2.0 of the GNU Compiler Collection, while the updated Run 3 version of the software utilizes version 11.2.0.

The reconstruction times shown represent the particular set of test events on which they were evaluated. Due to statistical fluctuations in the number of charged particles produced per collision event and the particle flux through the various regions of the detector, the run time is expected to vary for evaluations of different sets of events from the same LB. Based on trials performed by sampling different, equal-sized sets of events, this variation is found to be on the order of 15%. However, when evaluating the run time differences between the different software releases, the variation is less than 10% in both cases, as the changes in absolute run time due to different particle multiplicities largely cancel out in this metric. For a given set of events, the stability of the measured run time across repeated reconstruction runs on the same machine was evaluated to be better than 5%.

3.3 Software Optimization

Many changes to the tracking software were introduced to optimize the computational performance and the size of the generated output for the conditions expected during Run 3 data-taking. In addition to general algorithmic improvements, the guiding principle is to abort the track reconstruction as early as possible for candidates not expected to result in high-quality tracks. This minimizes the number of executions of the downstream algorithms, thus saving time and resources. The experience obtained by operating the reconstruction software throughout Runs 1 and 2 also allows the algorithms to be operated with less conservative tolerances than were initially implemented, gaining execution speed without sacrificing reconstruction acceptance.

A summary of the nightly SPOT profiling results over time as optimizations were incrementally added is shown in Figure 3.12.⁸ The starting point for the optimizations, referred to as “Baseline” in the figure, is a version of the legacy reconstruction with an initial set of modifications required to ensure thread-safety, and as a consequence is slower than the original legacy implementation used throughout Run 2 by around 20%. The following sections will discuss the specific optimizations in broad order of the stages of track reconstruction.⁹

⁸It is unfortunately not possible to temporally uncouple the sets of improvements from each other for this figure as changes are often interdependent on those implemented previously, or sets of changes relating to a single stage of the track reconstruction are implemented throughout several different merge requests. Creating separate instances of the original legacy reconstruction with an isolated single change would be prohibitively difficult and not necessarily representative. Presenting the improvements over time in this fashion is fully representative of the ATLAS track reconstruction as a “living” codebase with dozens of distinct contributors.

⁹Grouping them like this makes for a more coherent narrative than the real-time order, which is very disorganized!

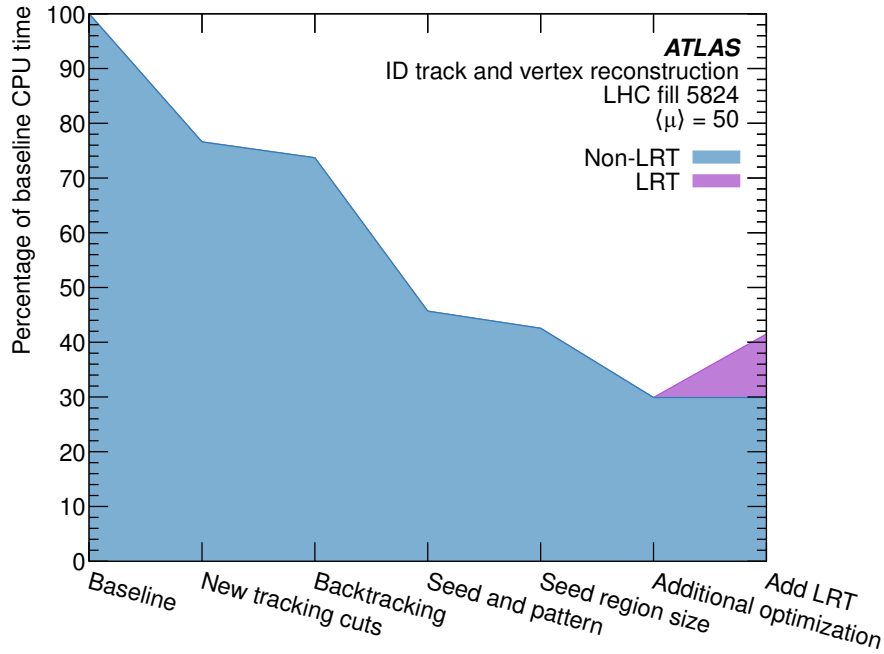


Figure 3.12. Incremental improvement in execution speed for ID track reconstruction in a set of $\langle\mu\rangle = 50$ collision data events as improvements are added to the reconstruction software for Run 3. The blue shaded area indicates the time taken for the track reconstruction relative to the initial Run 3 software implementation. The purple area indicates the time added by the additional LRT. Due to changes required to ensure thread safety, the unoptimized initial performance of the Run 3 software, referred to as “Baseline”, was approximately 20% slower than the previous software used during Run 2.

3.3.1 Track Seeding and Finding Optimization

The seeding formation time of the inside-out track reconstruction was optimized to prevent seeds unlikely to result in tracks from being passed into downstream processing. The optimizations include narrowing the acceptance threshold for the estimated impact parameters of SCT seeds from 2cm to 5mm and reducing the tolerance in the search road used to extend the seeds from 20 to 12 mm. Additionally, pixel seeds sharing space points with another seed with a lower estimated transverse impact parameter are prevented from passing into further processing. The reduced redundancy from this change was mitigated by exploiting the availability of four pixel layers since the installation of the IBL at the beginning of Run 2 by introducing **confirmation space-points**. As seeds are formed from only three space points, confirmation space points are a fourth space point from a different detector layer, which, if used to replace the outermost space point on a given seed, results in a new seed with a curvature compatible with the original, indicating it is likely part of the same charged-particle trajectory. This new feature is illustrated in Figure 3.13. All confirmed seeds are accepted

for further processing in the updated seed formation procedure, even if they share space points with other seeds. Finally, the window in the transverse impact parameter for accepting seeds dynamically adjusts depending on the deviation of the seed from a straight line in the $r - z$ plane, where no deflection due to the magnetic field is expected at the length scale of a seed. The criterion becomes stricter for seeds with a substantial deviation from a straight line, as these are more likely to be combinatorial artifacts. The introduction of these improvements corresponds to “Seed and pattern” in figure 3.12.

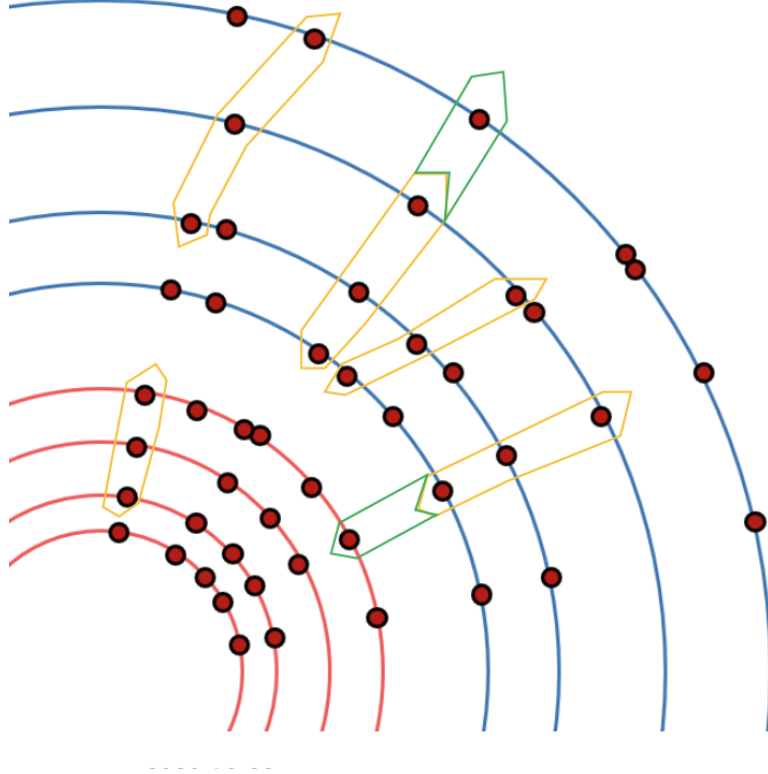


Figure 3.13. Illustration of the implementation of confirmation space points in ID track seed formation. The initial seeds comprised of three space points are outlined in orange, whereas their associated confirmation space points are outlined in green. The confirmation space points are additional space points in an adjacent layer compatible with the original seed trajectory.

The seeding strategy was also further optimized to speed up the seed formation step itself, as it contributes a significant fraction of the overall track reconstruction run time due to the process’s combinatorial complexity. The size of the angular regions within which seeds are formed was reduced to correspond to the maximum track curvature in the magnetic field expected to occur at the lowest track transverse momentum to be reconstructed (500 MeV in the primary pass) instead of the overly

conservative wider regions used previously. Additionally, the knowledge of the detector geometry is applied more rigorously, removing space points from consideration as the central point of a seed if they are placed on the outer- or innermost layers of a subdetector or in η regions where the subsequent layers provide no acceptance. This improves execution speed without significantly changing the number of tracks being reconstructed. This set of changes is labeled as “Seed region size” in Figure 3.12. Together, the improvements in the seeding phase reduce the overall ID track reconstruction time by half while only having a minor impact on the number of correctly reconstructed half and strongly suppressing falsely reconstructed tracks, which will be shown in the following section.

The number of incorrect tracks propagating beyond the initial seeding stage was further reduced by re-tuning the selection criteria applied to the track candidates. Instead of seven silicon (pixel or SCT) clusters per track, at least eight are now required in the updated tracking, and the allowed transverse impact parameter range of silicon-seeded tracks is restricted to $|d_0| < 5$ mm instead of < 10 mm. This reduces the acceptance of track reconstruction in terms of both displacement and production radius by a small fraction but significantly lowers the rate of low-quality tracks being reconstructed and written to storage. The reduction in initial track candidates achieved by this optimization also reduces the total number of executions of the subsequent ambiguity resolution procedure and TRT extension phases per event, reducing the total track reconstruction time per event by approximately 25%. This tuning of the track selection criteria is referred to as “New tracking cuts” in Figure 3.12.

3.3.2 TRT Extension and Backtracking Optimization

The TRT extension was sped up significantly by aborting the iterative track fit procedure early for candidates with insufficient precision hits in the TRT. While the required precision hit fraction of 30% has remained unchanged, the compatibility threshold for precision hits was reduced from 2.5 times the drift tube radius to 1.75, lowering the fraction of spurious TRT extensions and benefiting momentum resolutions by reducing poor-quality TRT extensions. In addition, the iterative fit is now aborted after the third iteration if the final precision hit fraction obtained at this stage does not satisfy the requirement, rather than continuing the fit iterations up to convergence before deciding whether to retain the extension as in the legacy reconstruction. These changes do not impact the reconstruction efficiency or the rate of fake tracks but speed up the TRT extension step by nearly 30%. These improvements to the TRT extension are included under the “Additional optimization” entry in Figure 3.12.

The TRT-seeded backtracking step previously generated a large number of falsely reconstructed tracks. This was reduced by limiting the backtracking to regions of interest seeded by energy deposits in the EM calorimeter with $E_T > 6$ GeV. The recovery of late-appearing tracks from electron conversions, which is the primary purpose of this reconstruction step, is degraded at a negligible level since these topologies generally coincide with significant calorimetric deposits. However, the number of fake tracks is reduced and the execution speed of the backtracking phase is improved by a factor of 20. As backtracking comprises a smaller fraction of the total track reconstruction CPU time, this improves the overall timing on the order of 5% at $\langle\mu\rangle = 50$. These improvements are listed as “Backtracking” in Figure 3.12. More considerable gains are achieved at higher pileup values as the detector occupancy grows.

3.3.3 Primary Vertexing Optimization with AMVF and ACTS

The iterative vertex finding (IVF) algorithm used for primary vertex identification described in the previous chapter was replaced with an **adaptive multi-vertex fitter** algorithm (AMVF) [74], where vertex candidates are allowed to “compete” for tracks in order to reduce the chance of nearby pp interactions being reconstructed as a single merged vertex. Tracks that are not yet assigned to a vertex candidate form an initial “seed pool” of unassigned tracks. The initial vertex locations are estimated with high accuracy using a Gaussian resolution model for the track impact parameter, using the density of tracks along the beam axis to determine the most likely location of a pp interaction vertex from tracks in the seed pool. The density function for a track is given by a normalized one-dimensional Gaussian distribution of width $\sigma(z_0)$, centered at z_0 , with an additional z -independent multiplicative factor that reduces the contribution based on its measurement errors and d_0 significance:

$$\lim_{\sigma(d_0, z_0) \rightarrow 0} \rho(z) = \left(\frac{1}{\sqrt{2\pi}\sigma(d_0)} e^{-\frac{1}{2}d_0^2/\sigma^2(d_0)} \right) \left(\frac{1}{\sqrt{2\pi}\sigma(z_0)} e^{-\frac{1}{2}(z-z_0)^2/\sigma^2(z_0)} \right). \quad (3.1)$$

The total density $W(z)$ as a function of z is the sum of all nearby track densities. The global maxima of this density are the vertex seed positions. Taking the first and second derivatives along z around each track location determines if a given location is near a local maximum, which can then be iteratively stepped through to determine the locations of all maxima. This also provides an estimate of the seed width, which is supplied to the vertex finder to constrain the track to vertex fit to the approximate position of the seed, to avoid unwanted migration when tracks from other nearby vertices might be considered compatible:

$$\sigma(z) = \sqrt{-\frac{W(z_{\max})}{W''(z_{\max})}}. \quad (3.2)$$

After the seed locations are found, the set of nearby tracks to fit is chosen. All tracks passing a selection based on a sliding z constraint around the vertex seed width and an impact parameter significance-based cut

$$s = \sqrt{(d_0/\sigma(d_0))^2 + (z_0/\sigma(z_0))^2}. \quad (3.3)$$

are eligible for assignment to the new vertex candidates, and unlike the IVF, tracks can be assigned to multiple vertices with varying weights. The helical parameters of the assigned tracks are used to fit the position of the vertex candidate using a weighted adaptive Kalman filter, itself subject to transverse and longitudinal constraints from the beamspot and seed positions. Again, unlike the IVF strategy, each time a new vertex candidate is fit, all the other candidates associated with it via any other track are also simultaneously refit. Track weights ω to each vertex i are annealed in six steps based on χ^2 compatibility. After each decrease of the annealing temperature T , the weights are recalculated, and all associated connecting vertices are refit. The total weight for an individual track across all vertices it is assigned to is normalized to one, computed from its compatibility with being unassigned:

$$\omega_i(\chi_i^2, T) = \frac{e^{-\frac{1}{2}\chi_i^2/T}}{\sum_j e^{-\frac{1}{2}\chi_j^2/T} + e^{-\frac{1}{2}\chi_0^2/T}}. \quad (3.4)$$

Once a vertex candidate satisfies several conditions, it is accepted or rejected. A vertex candidate must have at least two compatible tracks ($\chi^2 > 10^{-4}$) from the unassigned seeding track pool, such that a new vertex cannot be formed from only tracks that are already compatible with a previously accepted vertex. The new vertex position must also not be within 3σ of another vertex based on the calculated fit error of both vertices, and the weighted average of track weights in the fit ($\Sigma\omega^2/\Sigma\omega$) must be greater than $2/3$. If the candidate is accepted, all of the compatible tracks are removed from the initial seeding pool, and if it is not accepted, the single most compatible track is removed from the seeding pool. This ensures that the pool of remaining seed tracks is always reduced in each iteration of the vertex-finding, ensuring it does not continue infinitely. The vertex-finding terminates when either less than two tracks remain in the seeding pool, the seed finder can no longer return a seed, or the maximum allowed number of iterations is exceeded. A simplified flowchart of this new vertex-finding algorithm is shown in Figure 3.14.

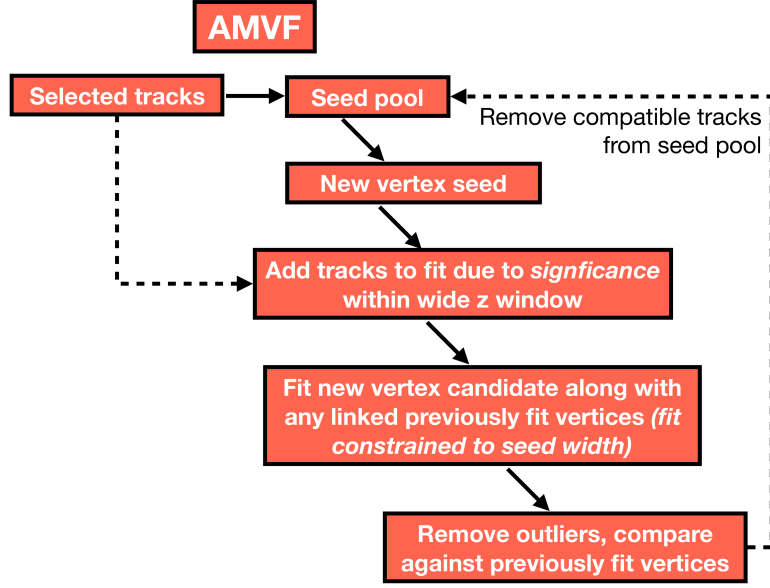


Figure 3.14. Flowchart of the AMVF logic. Taken from [74].

The AMVF performs much better in a high pileup environment due to its reduced tendency to merge nearby vertices. Due to the track assignment’s adaptive weighted nature, it can better resolve vertices with smaller longitudinal separations, as shown in Figure 3.15a. This significantly reduces the impact of pileup on the efficiency of reconstructing the hard scatter vertex. As shown in Figure 3.15b, the selection efficiency, for instance, in $t\bar{t}$ events, dramatically improves with the new strategy as a function of local pileup density.

However, this algorithm is computationally more intensive than the IVF due to its longer convergence time. This was mitigated partially by implementing this algorithm within the *A Common Tracking Software* (ACTS) framework [130]. ACTS is a modular track reconstruction framework based on the ATLAS track reconstruction workflow that is rewritten from the ground up to be thread-safe, optimized for vectorization, easily parallelizable, and agnostic to specific detector technologies and design such that it can be adapted for any experiment.¹⁰ ACTS will entirely comprise the backend of ATLAS track reconstruction during the HL-LHC era, and this implementation represents the first production use of this framework in an LHC experiment. The improved vertex finding is included under “Additional optimization” in Figure 3.12.

¹⁰It already has early adopters from smaller experiments that can exploit this newly available breadth of ATLAS tracking expertise, such as FASER, PANDA, and sPHENIX. ACTS is the future of tracking, and we will not be able to do it appropriate justice here!

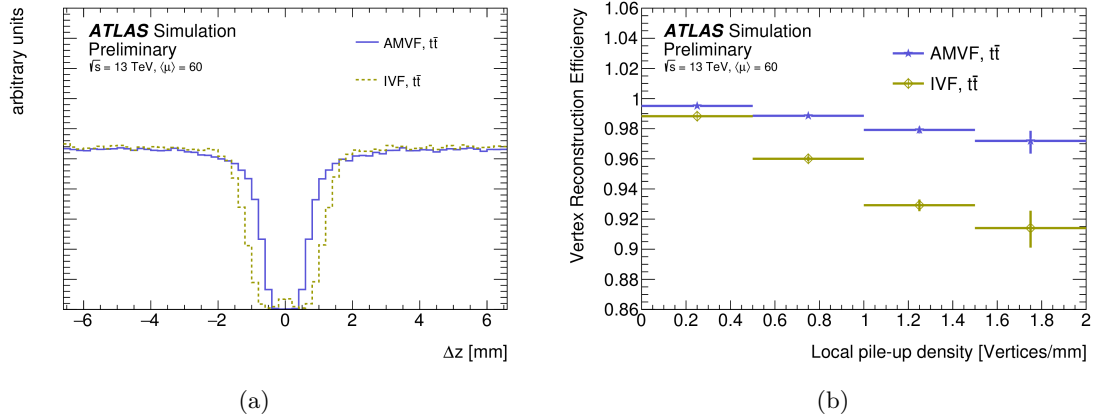


Figure 3.15. Distributions of the longitudinal separation between pairs of reconstructed primary vertices (a) and hard-scatter reconstruction efficiency as a function of local pileup density (b) for simulated $t\bar{t}$ events. Taken from [74].

3.3.4 LRT and Other Optimization

Further execution speed was gained by carefully optimizing the software implementation of each reconstruction step individually. Notable examples include a reorganization of the search for holes on tracks performed as part of the precision fit, exploitation of the navigation between detector surfaces already being conducted by the track fit procedure, optimization of the space-point formation by minimizing the number of required matrix multiplications, and the rewriting of parts of the Runge-Kutta propagator implementation used to extrapolate trajectories through the inhomogeneous magnetic field of the detector to exploit vectorized instructions where possible. These changes are also part of the entry “Additional optimization” in Figure 3.12.

The improvements to the track seeding and finding steps also directly benefited the LRT reconstruction. Together with some dedicated improvements to the LRT strategy [71], this made it feasible to integrate fully the LRT pass in the updated ATLAS track reconstruction that is performed for all events. A flowchart comparing the Run 2 and Run 3 workflows is shown in Figure 3.16. As it reconstructs long-lived particle tracks from SCT seeds only with a massively loosened requirement on impact parameters, LRT is highly susceptible to combinatorial fake tracks. The optimizations described in this chapter, along with the additional LRT-specific tunings, dramatically reduce the reconstruction of combinatorial fakes and contamination from particles produced in secondary material interactions while *increasing* the overall acceptance for charged particles from simulated LLP decays. This can be seen in Figure 3.17. Removing the need for a custom workflow involving a pre-selection step represents a significant extension of the ATLAS physics potential for long-lived

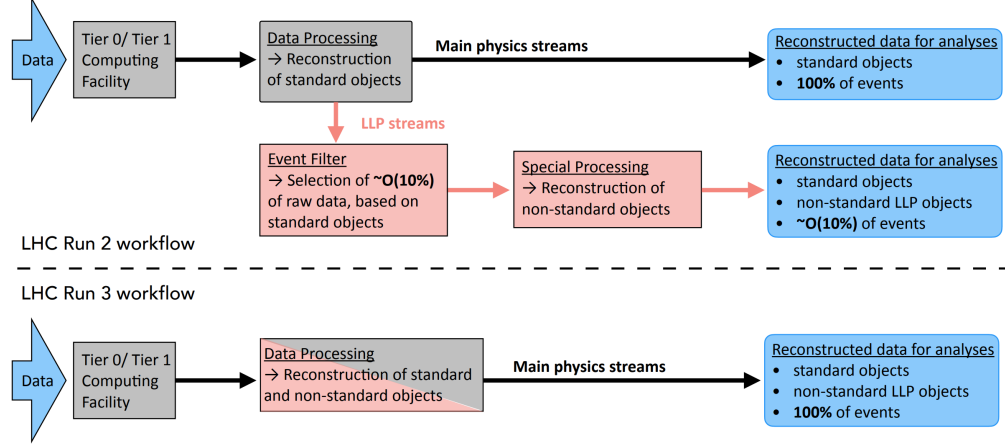


Figure 3.16. Flowchart illustrating the Run 2 (top) and improved Run 3 (bottom) event reconstruction workflows for the main physics and the LLP streams. Taken from [71].

particle (LLP) searches. In the following studies of reconstruction performance, the impact of the LRT step will be pointed out separately to facilitate a like-to-like comparison, as it was not run by default in the legacy reconstruction and adds back a non-negligible amount of CPU time in the final optimized configuration. The impact of adding LRT into the default reconstruction chain is shown by a dedicated entry denoted “Add LRT” in Figure 3.12.

3.4 Performance Results

The updated reconstruction’s performance is compared to the original, single-threaded software version used to reconstruct the Run 2 dataset at the time of data collection, with an emphasis on the evolution of the results as a function of pileup. Figure 3.18 shows the relative speedup obtained for the main components of track reconstruction when compared to the original legacy software as a function of the amount of total pileup. The run time of the tracking pattern recognition step has been reduced by up to a factor of four, and the other steps of ID tracking are faster by factors of up to two. This includes drift circle creation, TRT segment making, TRT-seeded backtracking, and TRT extension finding and fitting.

A comparison of the average total track reconstruction time per event between the updated and legacy reconstruction software is shown as a function of pileup in Figure 3.19. The time is decomposed into the key reconstruction steps, and the bottom panels indicate the fraction of the total time contributed by each step. For the legacy reconstruction, the track-finding step was by far the largest CPU consumer for track and total ATLAS reconstruction, scaling superlinearly with $\langle\mu\rangle$.

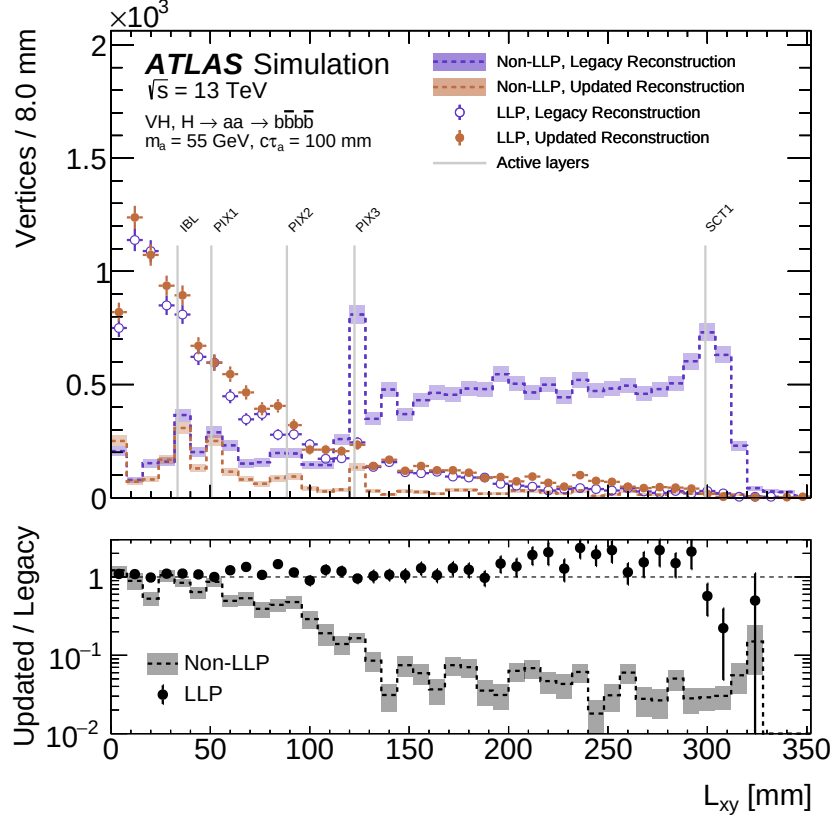


Figure 3.17. A comparison of the radial distributions of reconstructed secondary vertices in a simulated LLP sample using the legacy and updated track reconstruction configurations. The circular markers represent reconstructed vertices that are matched to truth-level LLP decay vertices (LLP). The dashed lines represent reconstructed vertices that are not matched to truth-level LLP decay vertices (non-LLP). Vertical lines indicate the radial position of the pixel and SCT detector layers. The lower panel shows the ratio of the number of vertices reconstructed in the updated reconstruction to those reconstructed in the legacy implementation for true LLP vertices (circular markers) and non-LLP vertices (dashed lines). The uncertainties shown are statistical only. Figure and caption taken from [71].

This behavior has been corrected in the updated reconstruction, where after the optimizations, the pileup dependency of reconstruction time is much closer to linear in the range relevant for Run 3 datataking, showing that the updated tracking software is now well suited for datataking at higher $\langle\mu\rangle$. A reduction of the total tracking and vertexing CPU time per event by a factor of three is visible for $\langle\mu\rangle = 60$, and the CPU time taken is distributed more evenly between the individual components.

Figure 3.20 shows the total track reconstruction time as a function of $\langle\mu\rangle$, including the additional ultra-high $\langle\mu\rangle$ machine-development run. The updated reconstruction is more than twice as fast as the previous version, even after including the new LRT step. Nearly linear scaling of the CPU

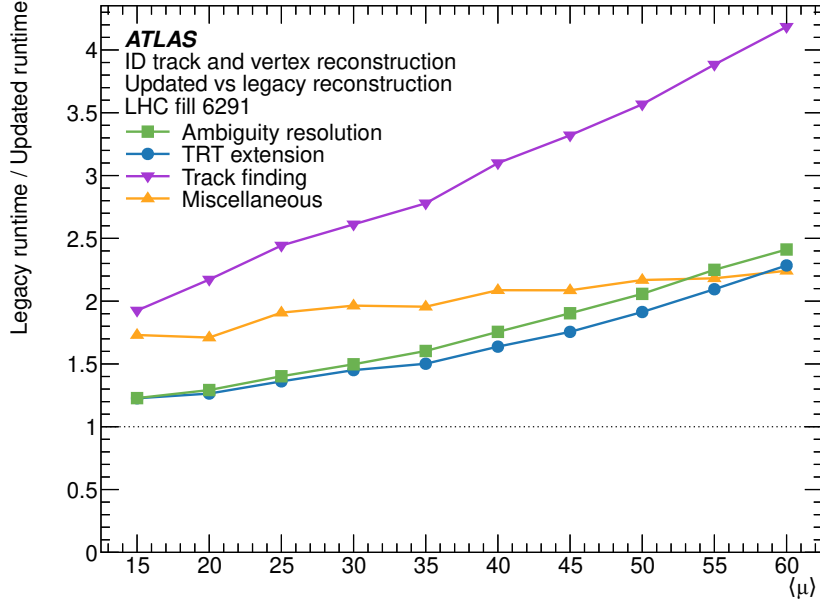


Figure 3.18. Relative computational performance improvement of key reconstruction steps obtained in the updated reconstruction software compared to that of the legacy implementation when reconstructing a set of LHC Run 2 data events, as a function of $\langle\mu\rangle$, for ID track reconstruction. The “Miscellaneous” category includes event data model handling, vertex reconstruction, reading of detector condition information, and cluster formation in the Pixel and SCT subdetectors.

consumption with $\langle\mu\rangle$ is now observed compared to the exponential behavior seen for the legacy reconstruction, and even when reconstructing data taken at $\langle\mu\rangle = 90$, the updated reconstruction achieves a processing speed comparable to the legacy reconstruction at $\langle\mu\rangle = 60$ when applied to a similar workload. The benefits from the improvements become more evident with increasing pileup. The CPU time taken to reconstruct a set of Run 3 collision events is also studied, with the LRT step included. The timing of the Run 3 reconstruction is consistent with the expectation obtained by evaluating the updated reconstruction on Run 2 data, demonstrating the optimizations developed on reprocessed Run 2 data were successfully fully representative of authentic Run 3 performance. The slightly faster reconstruction compared to the Run 2 reference is due to temporary failures of pixel B-layer detector modules during 2022 data-taking, which resulted in a reduced number of space points to process.

The improved efficiency of the track reconstruction is further illustrated in Figure 3.21. Figure 3.21a shows the number of processed track seeds divided by the final number of tracks created in the silicon-seeded track-finding process as a function of $\langle\mu\rangle$. The ratio of seeds to output tracks is reduced by a factor of almost two in the updated reconstruction. To reconstruct the same number

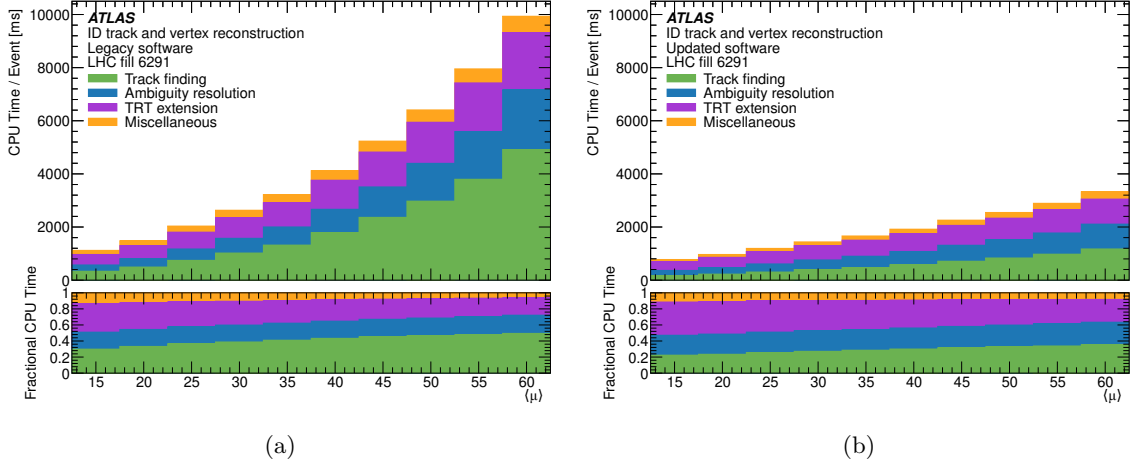


Figure 3.19. Breakdown of the average CPU time required per event for ID track and vertex reconstruction when reconstructing a set of LHC Run 2 data events, as a function of $\langle\mu\rangle$, comparing the legacy (a) and updated (b) configurations. The bottom panel shows the fraction of the total track reconstruction time taken by each component. The “Miscellaneous” category includes event data model handling, vertex reconstruction, reading of detector condition information, and cluster formation in the Pixel and SCT subdetectors.

of tracks, the costly algorithms responsible for extending a seed through the ID into a candidate are thus called only half the number of times on average, allowing ATLAS to achieve the considerable speedup of the track-finding stage shown in Figure 3.18. Figure 3.21 depicts the average number of tracks reconstructed per event as a function of the *instantaneous* pileup, μ_{actual} .¹¹ A linear fit to $10 \leq \mu_{\text{actual}} \leq 30$ is superimposed as a dashed line on the full range of μ_{actual} . The number of genuine track candidates per event is proportional to the number of charged particles produced in the pp collisions and thus scales linearly with μ_{actual} . The number of fake tracks scales superlinearly with μ_{actual} , reflecting the growth of combinatorics. A clear non-linear component amounting to up to 30% of the total number of tracks is visible for the legacy reconstruction, whereas the updated reconstruction shows nearly ideal linear behavior. The formation of combinatorial fake tracks has become very rare in the updated tracking, further improving computational efficiency, benefiting the reconstruction of track-based physical observables, and reducing the disk storage requirements per event.

¹¹The distinction here is a bit subtle – LUCID allows ATLAS to have a near-exact event-by-event determination of pileup values, here μ_{actual} , which is used in this figure computing the average number of tracks on a per-event level. But the Figures showing machine-level metrics are taking, e.g., average timings over *several hundred events* selected from a single lb of an average pileup $\langle\mu\rangle$, where the individual events may vary slightly. It would not be very meaningful to try to compute an average representative CPU timing for a specific μ_{actual} as this would just be dominated by statistical fluctuations in the charged particle multiplicity and conditions differences, etc.

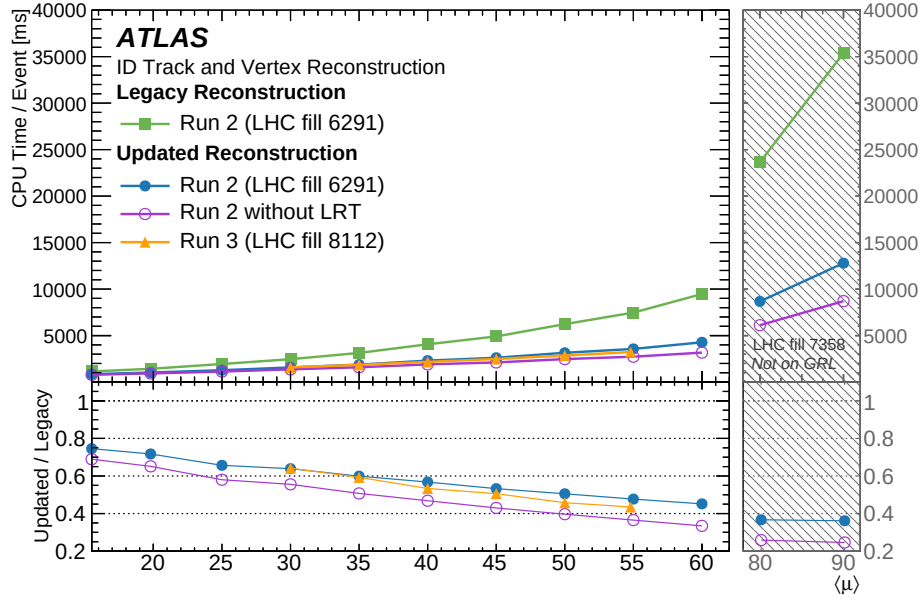


Figure 3.20. Processing time taken per event versus average pileup to reconstruct ID tracks and vertices in the same events for the updated (purple) and legacy (green) reconstruction software. The updated reconstruction is also evaluated when including the impact of the additional LRT step (blue). The bottom panel depicts the time taken relative to the legacy reconstruction time. The shaded area indicates data events from a 2018 machine development run that did not pass the full ATLAS data quality requirements.

The improvement in data storage requirements resulting from the updated ID track reconstruction is also visible in Figure 3.22, which shows the average disk space required to store ID tracks in kilobytes per event for the standard ATLAS event data format. The observed 20-50% reduction compared to the legacy reconstruction is a direct consequence of reducing the rate of misreconstructed ID tracks by a combination of the algorithmic improvements, stricter silicon hit count, and impact parameter requirements discussed above. Even after including the additional tracks from the LRT, a reduction of up to 40% is achieved at $\langle\mu\rangle = 60$. Additionally, the growth of the disk usage with increasing pileup is observed to be smaller in the updated software release, leading to more considerable improvements at even higher values of $\langle\mu\rangle$.

The new deployment of the ACTS software framework in ATLAS for primary vertex reconstruction significantly reduces the primary vertexing CPU time. Figure 3.23 shows a comparison of the primary vertex reconstruction time required per event between the ACTS-provided AMVF algorithm and the previously deployed Athena implementation. An average reduction of CPU time of more than a factor of 1.5 is seen, while the ACTS implementation yields physics results that

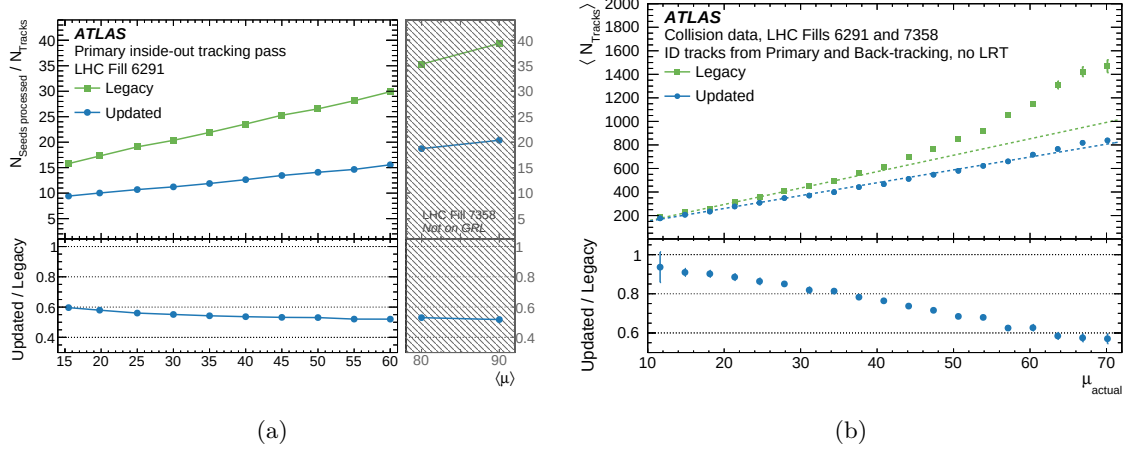


Figure 3.21. Number of seed tracks processed divided by the final number of tracks created in the silicon-seeded tracking as a function of $\langle \mu \rangle$ (a) and average number of tracks reconstructed per event as a function of the value μ_{actual} , with a linear fit to the range $10 \leq \mu_{\text{actual}} \leq 30$ superimposed as a dashed line (b).

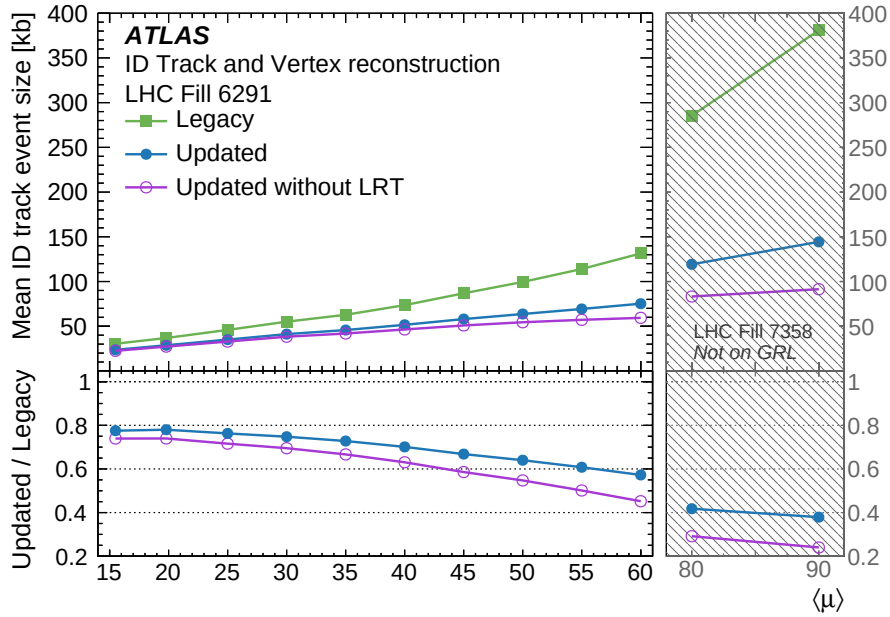


Figure 3.22. Event size of the inner detector reconstruction output in the ATLAS event data format for the same set of reconstructed data events as a function of the average pileup, comparing the legacy and updated releases. The shaded area indicates data events from a 2018 machine development run that did not pass the full ATLAS data quality requirements.

are identical to those from the original Athena version. The now fully integrated ACTS vertexing

software makes additional functionality with further potential for CPU improvements available for future developments.

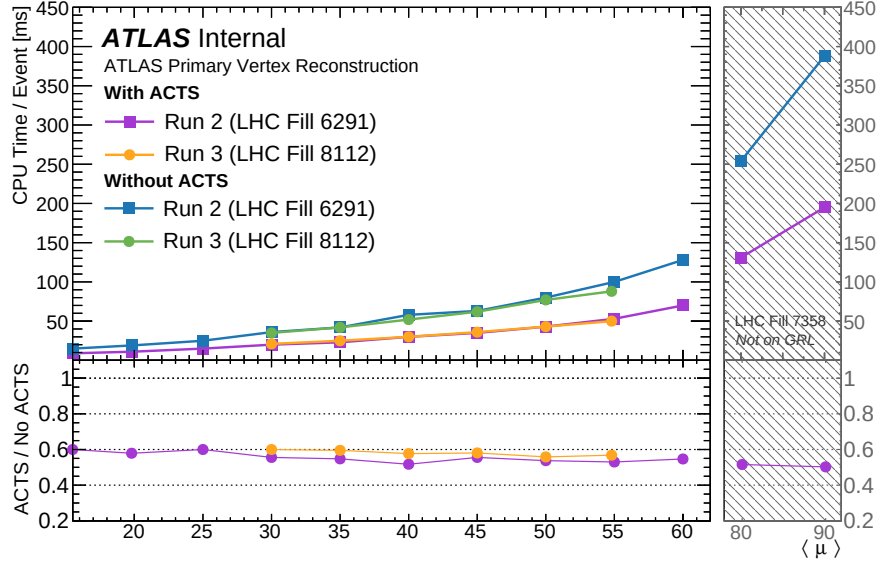


Figure 3.23. A comparison of primary vertex reconstruction processing time taken per event versus average pileup for the Athena version of the AMVF algorithm and the ACTS-provided implementation. The shaded area indicates data events from a 2018 machine development run that did not pass the full ATLAS data quality requirements.

In the updated reconstruction, the nearly 70% fraction of total reconstruction CPU time shown in Figure 3.2 is approximately retained when including the additional LRT contribution, shown in Figure 3.24. This is because several other algorithms, such as jet reconstruction and hadronic tau decay candidate identification, which contributed around 3% of the total reconstruction time, have been moved from the primary reconstruction pass into the downstream processing in the updated event reconstruction. Additionally, a large number of the most time-consuming non-tracking algorithms, such as streaming the reconstructed event data to disk, scale in run time with the number of reconstructed track candidates and thus benefit themselves proportionally from the reduced rate of fake tracks. Accordingly, the total reconstruction time per event has decreased from 9.1s to 4.8s on the machine used for producing these benchmarks. This indicates that track reconstruction remains the most promising part of event reconstruction to achieve further improvements in overall throughput.

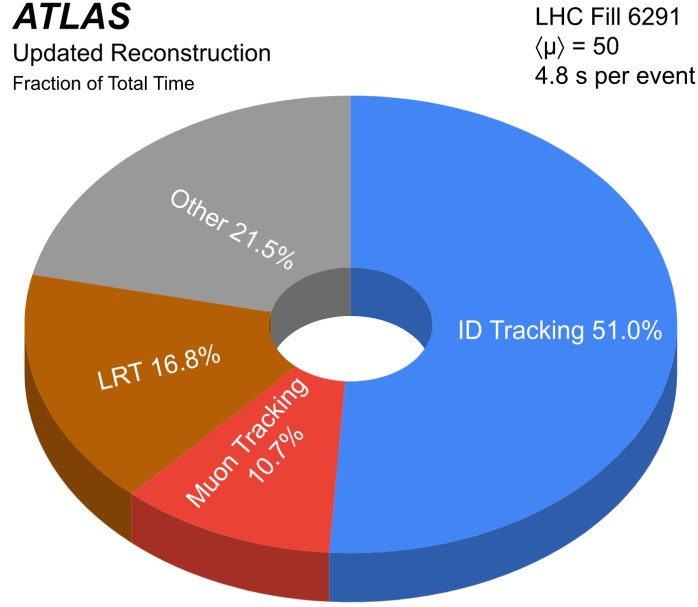


Figure 3.24. Fractions of the total CPU consumption of the full ATLAS Run 3 reconstruction for one data run at $\langle\mu\rangle = 50$. In the updated reconstruction, several algorithms, such as jet reconstruction and hadronic tau decay candidate identification, are no longer included in the primary ATLAS reconstruction; instead, they are created downstream and, hence, no longer enter the sum.

Finally, the software improvements have not negatively impacted the track reconstruction physics performance compared to the legacy reconstruction. The tracking efficiency, defined as the fraction of charged particles originating from the primary pp interaction successfully reconstructed, is shown in Figure 3.25 as a function of the charged particle transverse momentum. Simulated $t\bar{t}$ events are studied, leading to a track sample largely dominated by pions. The efficiency loss is smaller than 4% at the lowest reconstructed p_T and smaller than 1% at larger values. The slight reduction compared to the legacy implementation is a result of stricter requirements on the number of silicon clusters and impact parameters for a track to be retained and does not affect the reconstruction of muons as minimum ionizing particles. For simulated single muons with $p_T = 10$ GeV, which are generally ultimately reconstructed as the *Combined* muons used in the majority of ATLAS analyses involving muons, the updated track reconstruction achieves an efficiency within five permille of the legacy reconstruction across the full instrumented pseudorapidity range.

3.5 Conclusions and Outlook

Run 3 challenged the ATLAS event reconstruction with a significant increase in the amount of pileup activity. The inner detector track reconstruction software has been reoptimized to meet this

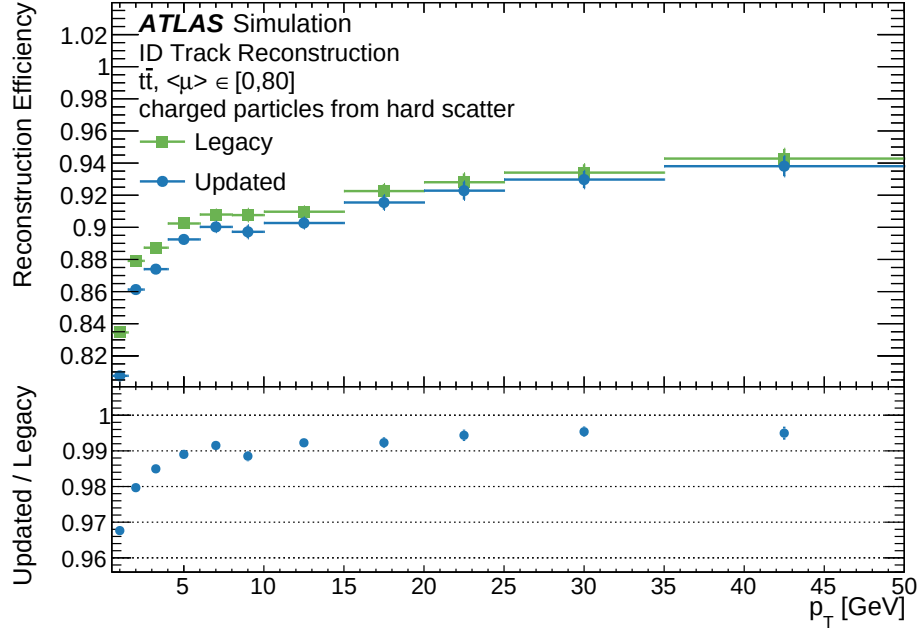


Figure 3.25. Tracking efficiency as a function of p_T in simulated $t\bar{t}$ events with a flat $\langle\mu\rangle$ distribution between 0 and 80, in a comparison of the updated and legacy reconstructions. The error bars indicate statistical uncertainties.

challenge, ensuring that prompt event reconstruction remains feasible within available computing resources. The improvements exploit the increased experience in operating the experiment by employing stricter tolerances where possible. Non-promising track candidates are rejected as early as possible, minimizing the number of calls to downstream algorithms within the reconstruction chain, thus yielding cascading improvements. An execution speed improvement between a factor of 2 and 4, depending on pileup, is achieved, and the number of fake tracks has been reduced to less than half the previous amount, and up to 40% less storage space is required for the tracks. At the same time, the reconstruction efficiency is nearly wholly retained. The performance and reconstruction speed remain stable at pileup values well beyond the design values for Run 3, giving confidence in the robustness of ATLAS event reconstruction to evolving operating conditions. The improvements also allow the execution of the additional LRT reconstruction pass to recover tracks not pointing towards the beamline in standard reconstruction, greatly enriching the ATLAS LLP search program.

Track reconstruction is a vibrant and diverse field that is fundamental to the operation of general-purpose collider experiments like ATLAS and CMS. In the HL-LHC era, through Run 4 and beyond, the current ATLAS ID will be replaced with the all-silicon Inner Tracker (ITk), and the existing Athena implementation of the track reconstruction software will be migrated to the full ACTS im-

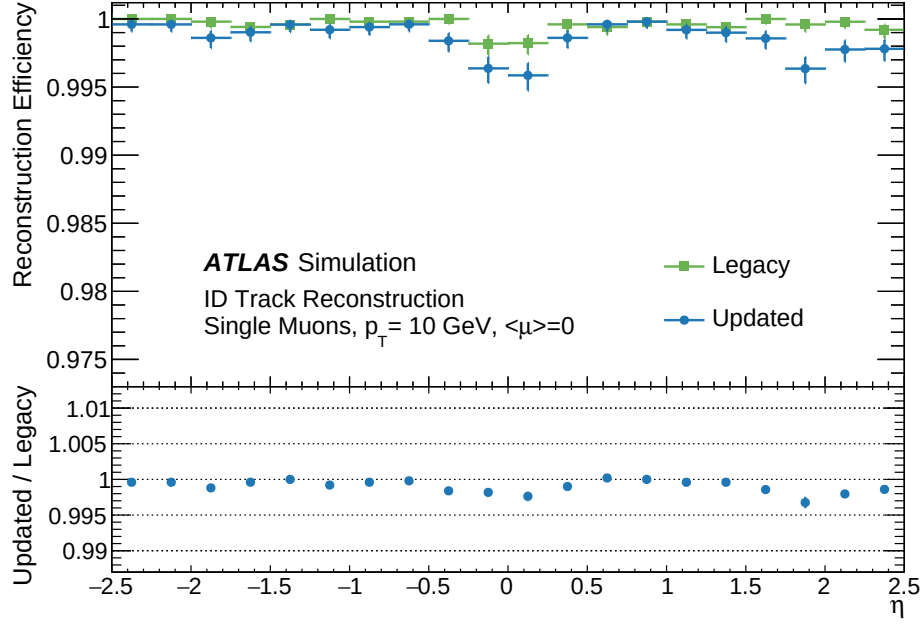


Figure 3.26. Tracking efficiency for simulated single muons using the Run 2 detector geometry with $p_T = 10$ GeV as a function of η , in a comparison of the updated and legacy reconstructions. The error bars indicate statistical uncertainties.

plementation, as has been done for the primary vertexing already. This “upgrade tracking” has been in development for many years and is well-prepared for the integration of the all-silicon detector technologies following the same software principles, with lessons and features taken from the optimizations for higher pileup values described herein. ACTS itself is the future of track reconstruction more broadly, and its flexible and modular approach opens new paths for innovations both inside and outside of ATLAS. Particularly, novel machine learning techniques such as graph neural networks (GNNs) are ever-increasingly explored for a vast range of tracking-related applications, ranging from track finding itself to hard-scatter vertex identification. The ATLAS tracking community is well-prepared to keep innovating as the field continues to evolve and looks forward to the challenges and opportunities the HL-LHC will bring.

CHAPTER 4

SEARCH FOR Wh RESONANCES

We have now established all of the required context to describe the search for Wh resonances with the ATLAS detector. As discussed in the introduction, this work presents a search for a new heavy vector resonance decaying into a W boson and a Higgs boson, interpreted through the Heavy Vector Triplet phenomenological model. It targets the semileptonic final state, where the Higgs boson decays hadronically into a b -quark pair and the W boson decays leptonically into a lepton-neutrino pair, and ultimately sets 95% CL upper limits on the production cross section of the HVT W' boson. A representative lowest-order Feynman diagram of the $q\bar{q} \rightarrow W'$ production mode is shown in Figure 4.1.

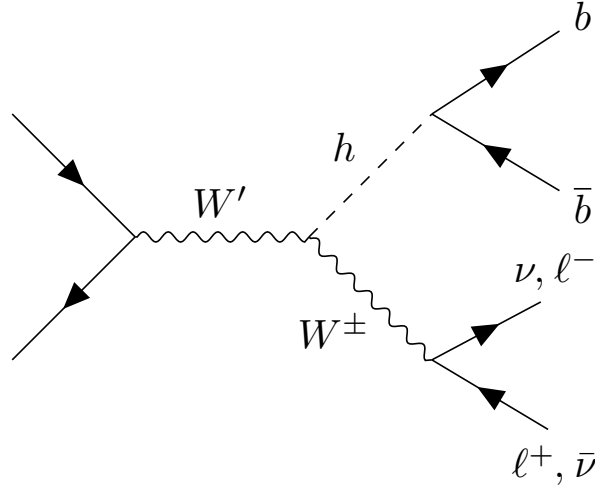


Figure 4.1. Representative lowest-order Feynman diagram of the resonant production of a W boson and a Higgs boson via $q\bar{q}$ annihilation. The subsequent hadronic decay of the Higgs boson and leptonic decay of the W boson are also pictured.

The search is performed in two overarching analysis channels based on the reconstructed lepton multiplicity, with a primary **1-lepton** channel selecting an electron or muon produced in association with missing transverse momentum and a secondary **0-lepton** channel selecting large missing transverse momentum with signal-like topological requirements, to increase acceptance to hadronic τ lepton decays and events where the lepton is not well-reconstructed. This analysis searches for

a localized excess in the reconstructed invariant (1-lepton) or transverse (0-lepton) mass distributions of the Wh system. It probes a wide hypothetical W' mass range from 400 GeV to 5 TeV and subsequently employs two methods to reconstruct the Higgs boson candidates. At low transverse momenta, the decay products of the Higgs boson are reconstructed as individual jets in a **resolved** topology, while at high transverse momenta, they are reconstructed as a single **merged** large-radius jet. These resolved and merged topologies are further divided into 1- and 2 b -tag categories based on the output of a multivariate classifier.

The search selects events in signal regions (SR) and background-dominated “sideband” control regions (CR) based on requirements placed on the kinematic properties of the final-state particles and event-level quantities. The final statistical interpretation is performed using a simultaneous binned maximum likelihood fit to data in all signal and control regions, where the major backgrounds are modeled using Monte Carlo simulations whose normalizations are determined by the fit. The results are interpreted in terms of the two HVT benchmarks *Model A* and *Model B*, detailed in Section 1.7.2.

Prior to the initial publication of this search, searches for the same final state had been performed at $\sqrt{s} = 13$ TeV with an integrated luminosity of 36.1 fb⁻¹ by ATLAS [29] as well as by CMS [131] with 137 fb⁻¹ [131].¹ Results in other final states have also been published: ATLAS performed a search in the fully hadronic final state [132] with the full Run 2 139fb-1 dataset, and CMS has published results in the fully hadronic final state [133] and the $\tau^+\tau^-\ell^+\ell^-$ final state [134] using a partial Run-2 35.9fb⁻¹ dataset. Apart from the significantly larger dataset used in this search, several improvements have been implemented since the previous publication, including improved b -tagging, lepton isolation, and jet reconstruction, as well as a new selection based on the missing transverse momentum significant in the 0-lepton channel. The refined b -tagging techniques lead to a significantly higher signal acceptance for W' masses with multiple TeV, while the other changes allow for a much-improved background rejection.

This Chapter will present this search, first providing a description of the data and simulated event samples in Section 4.1 and an in-depth description of the event selection and analysis strategy in Section 4.2. The background estimation techniques and systematic uncertainties are then discussed in detail in Sections 4.3 and 4.4. Section 4.5 gives a pedagogical discussion of the statistical analysis techniques used to derive the results, which are finally presented in Section 4.6.

¹For the curious reader, CMS has 2fb⁻¹ less integrated total Run 2 integrated luminosity than ATLAS due to some differences in detector downtime periods and respective data quality assessments.

This search has been previously published in “*Search for heavy resonances decaying into a Z or W boson and a Higgs boson in final states with leptons and b -jets in 139fb^{-1} of pp collisions at $\sqrt{s} = 13\text{TeV}$ with the ATLAS detector*”[14]. The version presented in this thesis adds a great deal of additional context and discussion to what has been presented in the published work.

4.1 Data and Monte Carlo Simulations

4.1.1 Data and Trigger

The data used in this analysis corresponds to the full ATLAS Run 2 dataset, collected between 2015 and 2018 from pp collisions at a center of mass energy of $\sqrt{s} = 13$ TeV, amounting to a total integrated luminosity of 139fb^{-1} . Events are required to be in the GRL for each datataking period, ensuring events are recorded with a fully functional detector and suitable for physics analysis.

As discussed in Section 2.2.6, the extremely high event rate in the detector is reduced by a two-level trigger system, where the specific triggers are selectively chosen in this search to maximize efficiency for each lepton multiplicity channel. Notably, the high efficiencies of leptonic triggers are a major motivating factor for targeting the Wh semileptonic decay mode, even though it does not have the largest branching fraction. For the 0-lepton channel, E_T^{miss} HLT triggers are used with thresholds between 70-110 GeV, increasing with each datataking period to adapt to a higher trigger rate due to the increased luminosity. These provide a trigger efficiency of around 80% for an offline E_T^{miss} value of 150 GeV and approach 100% above 200 GeV.

For the 1-lepton channel, a combination of single-electron and E_T^{miss} triggers are used with varying p_T and E_T thresholds and isolation requirements. The single-electron triggers, used in the e channel, require transverse energy (E_T) thresholds between 24-300 GeV, scaling with datataking period and increasing with looser isolation requirements, which are used to recover efficiency for electrons with $p_T > 60$ GeV. The reconstructed electron ultimately used in the analysis selection must be matched to the trigger object and satisfy a minimum p_T requirement of 27 GeV. For μ channel events, the E_T^{miss} triggers are used as they have a significantly higher efficiency for the relatively large momentum of the W boson required by the analysis, $p_{T,W} > 150$ GeV, compared to the available single-muon triggers.² The combined trigger efficiency for selected 1-lepton signal events is above 90% and approaches unity at high $p_{T,W}$. The specific trigger selections used in the analysis, separated by year, are detailed in Table 4.1.

²This choice is well-motivated from an analysis standpoint; as muons are not included in the online E_T^{miss} calculation, E_T^{miss} and $p_{T,W}$ are functionally equivalent at the trigger level, thus essentially triggering on $p_{T,W}$ in the muon channel.

E_T^{miss} Trigger	Electron trigger
2015	
HLT_xe70	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose
2016	
HLT_xe90_mht_L1XE50 ^(*) HLT_xe110_mht_L1XE50 ^(**)	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0 HLT_e300_etcut
2017	
HLT_xe110_pufit_L1XE55	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0 HLT_e300_etcut
2018	
HLT_xe110_pufit_xe70_L1XE50	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0 HLT_e300_etcut

Table 4.1. Summary of the missing transverse momentum and single electron triggers that have been used for the 2015-2018 datataking periods. The number following *xe* or *e* corresponds to the minimum E_T^{miss} or E_T requirement in GeV. Text following *lh* is the isolation requirement, while *mht* or *pufit* refer to the calorimeter clustering method used to calculate online E_T . The *pufit* algorithm corrects for pileup effects before being used as input based on a μ dependent threshold. *nod0* specifies that the transverse impact parameter is not used in the electron identification, while *etcut* specifies that only E_T information is used in the trigger selection. *L1EM* specify a Level 1 seed with the number specifying the p_T requirement. *VH* specifies several topologically dependent electron vetos. ^(*) available before subperiod D3 ^(**) available after subperiod D3

Data is compared to predictions from MC simulations, which are also required to “fire” the trigger. Small scale factors are applied in the analysis at lowest E_T^{miss} and p_T^ℓ to correct for discrepancies between data and simulation in the trigger “turn-on” regime.

4.1.2 Signal and Background MC

Simulated events of **HVT W' boson production** via $q\bar{q}$ annihilation were modeled at leading-order (LO) accuracy by MADGRAPH5 2.3.3 [135], interfaced to PYTHIA 8.186 [126] in the A14 [124] tune for simulation of the parton showers and subsequent hadronization. The NNPDF2.3LO PDF set [123] was used in all stages, and events were generated for a range of hypothetical W' masses from 300 GeV to 5 TeV using the benchmark HVT *Model A*. As the expected kinematic

differences between *Model A* and *Model B* are negligible in the analysis regime, both giving mass peaks in the Wh system with a width dominated by the experimental resolution, it is unnecessary to simulate Model B events separately. Additionally, as it is computationally infeasible to generate a full grid of resonant signal points with mass resolution spacing, an interpolation was performed using a moment morphing technique [136] [137], detailed in Appendix A, to increase the discovery potential of signals between simulated mass points. Even though the search does not directly target $h \rightarrow c\bar{c}$ decays, it is sensitive to such events. As such, Higgs boson decays into both $b\bar{b}$ and $c\bar{c}$ were generated, assuming branching fractions $\text{BR}(h \rightarrow b\bar{b}/c\bar{c}) = 0.569/0.0287$ ³ and a Higgs boson mass of $m_h = 125$ GeV. The W bosons are required to decay leptonically via $W \rightarrow \ell^\pm \nu$, where $\ell = e, \mu$, and τ .

Simulations are also generated for all non-negligible SM background processes, used to develop the analysis selection, derive systematic uncertainties, and fit observed m_{VH} distributions against SM predictions. The production of W and Z bosons in association with jets recoiling in the transverse plane (collectively V +**jets**, shown in Figure 4.2b) was modeled by SHERPA 2.2.1 [138] with the NNPDF3.0NLO PDF set [120]. Diagrams with up to two additional parton emissions were simulated at next-to-leading-order (NLO) precision, and those with three or four additional parton emissions at LO precision, using the COMIX [139] and OPENLOOPS [140, 141, 142] libraries. The matrix elements are merged with the SHERPA parton shower [143] using the MEPS@NLO formalism [144, 145, 146, 147] with the set of tuned parameters developed by the SHERPA authors. The inclusive production cross section is normalized with a global k -factor to match the next-to-next-to-leading-order (NNLO) prediction [148]. The total number of heavy-flavor (HF) quarks in the final samples has a strong impact on the overall shape of kinematic variables, including the final discriminants m_{VH} and $m_{T,VH}$. V +jets samples are thus further classified by the multiplicity of HF vs light (l) jets into $V+l$, $V+hl$ (bl, cl) and $V+hf$ (bb, bc, cc).⁴

The production of **top-quark pair** ($t\bar{t}$) events, the most significant background for the W' search, was modeled using the POWHEG BOX v2 generator [116, 117, 118, 119]. The matrix elements were calculated at NLO precision in QCD, assuming a top-quark mass of 172.5 GeV, using the NNPDF3.0NLO PDF set and the five-flavor scheme. Parton-showering, hadronization, and the

³An astute reader will notice that these are not the current 2024 state-of-the-art Higgs branching ratios but were the up to date value at the time of signal production. As we quote an upper limit on the *total* cross section and the branching ratios used in the generation are a simple multiplicative factor that is ultimately “divided out” in the final calculation, this does not meaningfully impact the final result.

⁴This is determined by checking if the final small- R jets can be *ghost-associated* [80] to a generator-level b – or c –quark.

underlying event were modeled with the PYTHIA 8.230 generator [122] using the NNPDF2.3LO PDF set and the A14 tune. The cross section is calculated at NNLO precision, including the resummation of the next-to-next-to-leading logarithmic (NNLL) soft gluon terms with TOP++ 2.0 [149, 150, 151, 152, 153, 154], and the h_{damp} parameter is set to $1.5 m_{top}$ [121]. An example tree-level Feynman diagram is shown in Figure 4.2a. Top quark pair production in association with a Higgs, W , or Z boson, which contribute a relatively small amount of background in the search but are modeled for completeness, were also generated at NLO accuracy. The $t\bar{t} + h$ samples were generated using MADGRAPH5_AMC@NLO 2.3.2 and interfaced to PYTHIA 8.186, while the $t\bar{t} + V$ samples were modeled using MADGRAPH5_AMC@NLO 2.3.3 and interfaced to PYTHIA 8.230. Both use the NNPDF3.0_{NLO} PDF set during generation and NNPDF2.3_{LO} with the A14 tune during the PYTHIA interfacing.

Similarly to $t\bar{t}$, production of top quarks in association with W bosons (Wt) and the s - and t -channel production of **single top quarks** was modeled at NLO precision in QCD using the POWHEG BOX v2 generator, using the NNPDF3.0_{NLO} PDF set and the five-flavor scheme, and subsequently interfaced to PYTHIA 8.230 using the NNPDF2.3_{LO} PDF set and A14 tune. The diagram removal (DR) scheme [155] was used to remove overlap and interference with $t\bar{t}$ production.

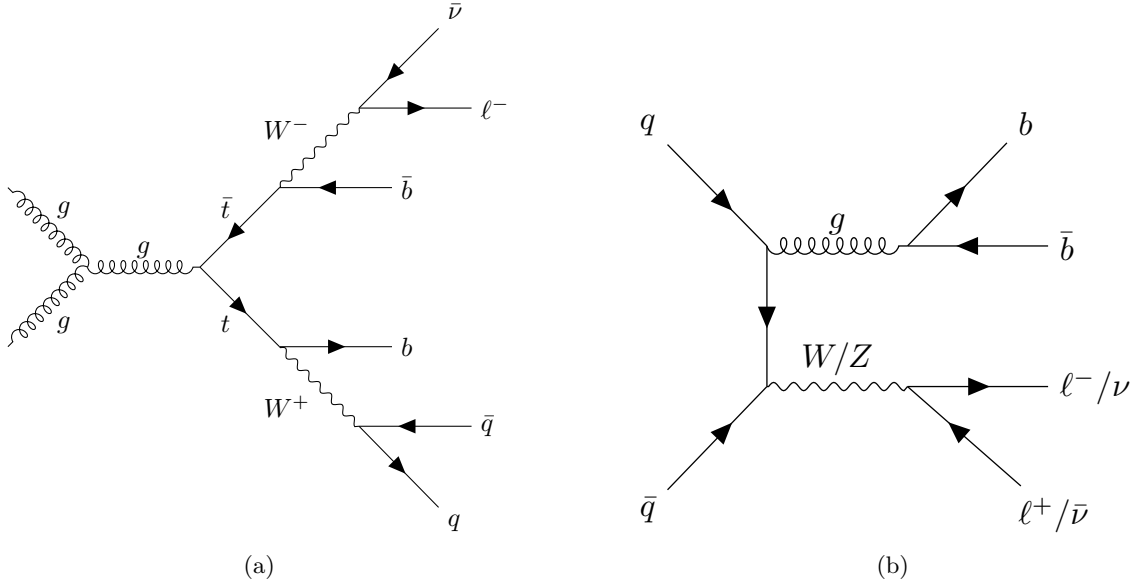


Figure 4.2. Example tree-level Feynman diagrams for $t\bar{t}$ (a) and W/Z +jets (b) production, the two most dominant SM backgrounds for the Wh resonance search.

Standard Model diboson events (WW, WZ, ZZ , collectively VV) with semileptonic decays were simulated with SHERPA 2.2.1, while those with leptonic decays were simulated with SHERPA

2.2.2. The NNPDF3.0NNLO PDF set was used in both productions, and off-shell effects and Higgs boson contributions were included where appropriate. Diagrams with up to one additional parton emission were simulated at NLO precision in QCD, while diagrams with two or three parton emissions were simulated at LO precision [156]. The matrix elements were merged and matched with the parton shower using the MEPS@NLO prescription. Loop-induced diboson processes that are initiated via the gg production mode were simulated at LO for diagrams with up to one additional parton emission using OPENLOOPS in SHERPA 2.2.2, also using the NNPDF3.0NNLO PDF set.

Finally, Standard Model production of a Higgs boson in association with a vector boson (**SM** Vh) was simulated using POWHEG BOX v2, interfaced with PYTHIA 8.212 for parton showering and non-perturbative effects, using the PDF4LHC15 PDF set [157] and the AZNLO tune [158]. The POWHEG prediction is accurate to NLO for Vh plus one jet production. The loop-induced $gg \rightarrow Zh$ process was generated separately at LO, and its production cross section was calculated at NLO precision, including the resummation of next-to-leading logarithmic (NLL) soft gluon terms. The Higgs boson mass is set to $m_h = 125$ GeV during the generation of all Vh events. Standard Model Vh and VV both give minimal contributions to the overall background composition due to their very low cross sections, but they do provide a nonzero contribution in some analysis categories.

A summary of all MC generators and programs used to model the parton shower, hadronization, and the underlying event is shown in Table 4.2. All simulated event samples include pileup by overlaying simulated minimum-bias events on each generated signal or background event, generated with single-, double-, and non-diffractive pp processes of PYTHIA 8.186 using the NNPDF2.3LO PDF set [127] and A3 tune [128]. All MadGraph and Powheg Box samples use the EvtGen 1.6.0 program [125] for the bottom and charm hadron decays. Generated particle-level samples are all processed through the full GEANT4-based ATLAS detector simulation [159, 32] and then reconstructed with the same algorithms used for data. They are then finally reweighted to compensate for discrepancies between data and simulation in dedicated performance measurements, accounting for the efficiencies for reconstruction, isolation, and triggering of leptons, tagging efficiency for heavy-flavor jets, and the energy/momentum scale and resolution of leptons and jets [55, 160, 161].

Alternate sets of samples, such as Wt production using the diagram subtraction (DS) scheme rather than the DR scheme, are also produced to estimate and derive systematic uncertainties from generator-level effects. To reduce overall resource consumption, as “standard” full simulation is computationally expensive, they are simulated with the ATLAS Fast Simulation [162] rather than the full GEANT4 simulation.

Table 4.2. Summary of the Monte Carlo event generators used to produce the various background processes. The perturbative accuracy (in QCD and, if relevant, in EW corrections) of the total production cross section σ_{prod} is stated for each process. The order at which the corresponding matrix elements are calculated in the Monte Carlo simulation is not necessarily the same as for the cross section. Details on the final state are also provided, where $\ell = e, \mu$, and τ .

Process	Generator	Perturbative accuracy of σ_{prod}
$q\bar{q} \rightarrow W' \rightarrow Wh \rightarrow \ell\nu + b\bar{b}/c\bar{c}$	MADGRAPH5 2.3.3 + PYTHIA 8.186	LO
$W \rightarrow \ell\nu, Z \rightarrow \ell\ell, Z \rightarrow \nu\nu$	SHERPA 2.2.1	NNLO
$t\bar{t}$	POWHEG BOX v2 + PYTHIA 8.230	NNLO+NNLL
single top quarks (s - and t -channels)	POWHEG BOX v2 + PYTHIA 8.230	NLO
single top quarks (Wt -channel)	POWHEG BOX v2 + PYTHIA 8.230	approx. NNLO
$t\bar{t} + h$	MADGRAPH5_AMC@NLO 2.3.2 + PYTHIA 8.186	NLO (QCD) and NLO (EW)
$t\bar{t} + V$	MADGRAPH5_AMC@NLO 2.3.3 + PYTHIA 8.230	NLO
$qg/q\bar{q} \rightarrow VV \rightarrow \ell\ell/\ell\nu/\nu\nu + q\bar{q}$	SHERPA 2.2.1	NLO
$gg \rightarrow VV \rightarrow \ell\ell/\ell\nu/\nu\nu + q\bar{q}$	SHERPA 2.2.2	NLO
$qg/q\bar{q} \rightarrow \ell\ell\nu\nu$	SHERPA 2.2.2	NLO
$qq \rightarrow Wh \rightarrow \ell\nu + b\bar{b}$	POWHEG BOX v2 + PYTHIA 8.212	NNLO (QCD) and NLO (EW)
$qq \rightarrow Zh \rightarrow \ell\ell/\nu\nu + b\bar{b}$	POWHEG BOX v2 + PYTHIA 8.212	NNLO (QCD) and NLO (EW)
$gg \rightarrow Zh \rightarrow \ell\ell/\nu\nu + b\bar{b}$	POWHEG BOX v2 + PYTHIA 8.212	NLO+NNL

QCD “multijet” events, which are events containing only quarks and gluons in the final state of the hard interaction, are not possible to directly simulate with a sufficient degree of accuracy. As such events do not contain any true E_T^{miss} from neutrinos nor any real high- p_T charged leptons, their contributions are heavily suppressed by the analysis selection. Still, they give a nonzero contribution in some phase space regions of the 1L channel. These are estimated using a data-driven template method, detailed further in Section 4.3.

4.2 Event Selection and Analysis Strategy

To maximize sensitivity to the $W' \rightarrow Wh \rightarrow \ell\nu b\bar{b}$ signature and discriminate signal from background, many event-level selections on the physics objects, whose reconstruction using the central ATLAS algorithms is extensively detailed in Chapter 2, are applied. Appropriate working points (WP) are selected for each object based on their respective performance at the analysis level, and events are categorized into the various signal and control regions based on set criteria described in this section.

4.2.1 Object Definition and Preselection

The **primary vertex** in the event is chosen with the standard Run 2 configuration as described in Chapter 2, using the Iterative Vertex Finder and selecting the vertex with the highest $\sum p_T^2$ of associated tracks.

Electron and muon candidates are both used in the primary selection for the 1-lepton channel and to actively veto events for the 0-lepton channel.⁵ **Electron candidates** are required to have a minimum p_T of 7 GeV and lie within the ID fiducial region of $|\eta| < 2.47$, additionally excluding the “gap” region between the barrel and endcaps ($1.37 < |\eta| < 1.52$) of the EM calorimeter.⁶ To ensure that electron candidates originate from the PV, they are required to satisfy a maximum threshold on the transverse impact parameter significance, $|d_0|/\sigma(d_0) < 5.0$, and on the longitudinal impact parameter, $|z_0 \sin \theta| < 0.5$ mm. For the 1-lepton e channel selection, the *Tight* identification WP is used to select high-quality electrons, while any electron passing these criteria and the *Loose* WP is used in the veto in the 0-lepton channel. Electron candidates are also further required to satisfy isolation requirements to ensure the suppression of backgrounds from hadrons faking electron signatures or from heavy-flavor decays inside jets. An electron is isolated from additional detector activity if the energy sum within a cone of $\Delta R = 0.2$ in the calorimeter system around the electron cluster axis is smaller than 0.06 times the electron E_T . Additionally, the sum of the transverse momenta of all ID tracks within a variable-radius cone up to $\Delta R = 0.2$, shrinking with increasing p_T , around the electron track is required to be smaller than 0.06 times the electron E_T – these working points are referred to as *topoetcone20* and *ptvarcone20* respectively.

Muon candidates are also required to have a minimum p_T of 7 GeV and lie within the MS fiducial volume of $|\eta| < 2.5$, and demonstrate compatibility with the PV by fulfilling impact parameter requirements: $|d_0|/\sigma(d_0) < 3.0$ and $|z_0 \sin \theta| < 0.5$ mm. Like electrons, selected muons in the 1-lepton channel must also pass the *Tight* identification WP, while the 0-lepton veto requires only a *Loose* WP. For the 1-lepton selection, muon candidates with $p_T > 300$ GeV must additionally satisfy the tighter *HighPT* identification requirement to improve the p_T resolution in the highest-energy regime [?]. They are also required to pass isolation criteria, where the sum of the transverse momenta of all ID tracks within a variable-radius cone up to $\Delta R = 0.3$ are required to be smaller than 0.06 times the muon p_T .

⁵This may seem counterintuitive in a search for the $\ell\nu b\bar{b}$ final state, but recall that this W' search is situated in a larger context, outside of the scope of this thesis, involving the $\nu\nu b\bar{b}$ final state as well where an active veto is very well motivated. The 0-lepton and 1-lepton channels have dramatically different background compositions, and they must remain orthogonal to create two well-understood regimes that can be combined in a single fit. While a veto on loosely identified and low p_T leptons loses some efficiency in recovering “missed” reconstruction of leptons from the 1-lepton signature, another primary function of the inclusion of the 0-lepton channel in the W' fit is to recover sensitivity from hadronic τ lepton decays through $W' \rightarrow Wh \rightarrow \tau_{had}\nu b\bar{b}$. To this end, while the “default” 0-lepton selection used in the other fit models of the broader analysis also actively vetoes events with any reconstructed τ leptons, the version used in the W' fit is loosened to allow for one reconstructed τ .

⁶This is due to a high fake rate and overall poor performance from electron candidates from the gap region and is a relatively standard procedure in ATLAS for this type of search.

To select the Higgs boson candidates, three jet definitions are employed in the analysis, as detailed in Chapter 2, all using the anti- k_T algorithm implemented in the FASTJET [163] package. **Small-R jets**, used to reconstruct **low-momentum Higgs candidates** in the **resolved** category, are reconstructed from noise-suppressed calorimeter TopoClusters with radius parameter $R = 0.4$ and required to have $|\eta| < 4.5$. Forward small-R jets ($2.5 < |\eta| < 4.5$) are required to have $p_T > 30$ GeV while central jets ($|\eta| < 2.5$) are required to have $p_T > 20$ GeV. Additionally, to suppress spurious jets resulting from pileup interactions, central small-R jets with $p_T < 60$ GeV are also required to pass the jet vertex tagger [164] selection. Large-R TCC jets, with $R = 1$, are used to reconstruct **high-momentum Higgs candidates** in the **merged category**, where the b -quarks are emitted with small angular separations. They are built from combined TrackCaloClusters [81] and are required to have $p_T > 250$ GeV and $|\eta| < 2.0$. Trimming [165] is also applied to large-R jets to remove the energy of contaminating clusters that originate from pileup, the underlying event, or initial-state radiation, as the large radius increases sensitivity to noise. The momenta of both small-R and large-R jets are corrected for energy losses in passive material in the calorimeters and differences in response scale between the EM and hadronic calorimeters [55]. Lastly, **VR track jets**, which are clustered from ID tracks using a variable radius that shrinks with increasing p_T , are used to tag b -jets in the merged category by requiring them to be spatially matched to the selected large-R jets through ghost association [86].⁷ They must contain at least two ID tracks that are compatible with the PV and have $p_T > 10$ GeV and $|\eta| < 4.5$.

Small-R jets and VR track-jets that contain b -hadrons are identified using the MV2C10 **b -tagging algorithm** [161, 166, 167, 168] with a working point that corresponds to a b -jet selection efficiency of 70% as measured in simulated $t\bar{t}$ events.

While not actively selected for in any analysis category, τ **lepton candidates** are used in the 0-lepton channel as a veto to help suppress backgrounds with multiple real hadronic τ decays. They are reconstructed using small-R calorimeter jets and are required to have either one or three associated tracks to capture the most frequent hadronic τ decay modes. They are also required to have $p_T > 20$ GeV and lie between $|\eta| < 2.5$, again excluding the calorimeter gap region ($1.37 < |\eta| < 1.52$), and are identified using a multivariate tagging algorithm with the *Medium* working point which is 40-55% efficient for three- and one-track τ candidates, respectively [92].

⁷Ghost-associated track jets are used in addition to the large-R jets because the traditional b -tagging algorithms are not effective on large-radius jets, owing to the way jet substructure variables are used in the algorithms. New machine-learning-based “double- b taggers” have been developed but largely postdate this search.

Though they cannot be observed directly, the presence of **highly energetic neutrinos** in the final state can be inferred from the **missing transverse momentum**, E_T^{miss} . The momentum imbalance in the transverse plane is calculated as the negative vector sum of the transverse momenta of calibrated leptons and jets, as well as a track-based “soft term” built from all tracks compatible with the PV but not associated with any object [96, 97]. Additionally, a track-based missing transverse momentum estimator p_T^{miss} is independently built as the negative vectorial sum of all tracks associated to the PV. An object-based E_T^{miss} **significance** $\mathcal{S}$ is also exploited to reduce backgrounds from mismeasurement of lepton and jet energies by taking the resolutions of objects entering the E_T^{miss} calculation and quantifying how likely it is there is significant E_T^{miss} from undetectable particles such as neutrinos [98].

4.2.2 Overlap Removal and Event Cleaning

Physics objects used in the analysis are reconstructed and identified independently, which can lead to ambiguous identification when these objects are spatially close to each other. An **overlap removal** procedure based on the angular separation is therefore applied to resolve these ambiguities and prevent double-counting of these objects. Additionally, to maintain rigor and quality in the analysis selections, some events are removed from consideration entirely if these ambiguities cannot be sufficiently resolved or if certain features would meaningfully impact the analysis quality.

Small-R jets are discarded if they are within a cone of $\Delta R = 0.2$ around an electron or τ lepton or if they have fewer than three associated tracks and are within $\Delta R = 0.2$ around a muon candidate, which preserves efficiency for leptonic b -hadron decays. A τ -lepton is removed if it is within $\Delta R = 0.2$ around an electron or muon unless, in the muon case, the τ lepton has p_T above 50 GeV and the reconstructed muon is not a *Combined* muon. Electrons and muons are discarded if they are within $\Delta R = \min(0.4, 0.04 + 10\text{GeV}/p_t^\ell)$ of the central axis of any jet surviving the other overlap requirements, where p_t^ℓ is the transverse momentum of the electron or muon. Electrons are also removed if they share the same ID track as a muon.

Events are removed from consideration entirely if they contain two or more overlapping VR track jets. Two VR track jets are considered overlapping if the angular separation between their central axes $\Delta R(j_i, j_j)$ is smaller than their jet radii R_i and R_j :

$$\Delta R(j_i, j_j) < \min R_i, R_j, \quad (4.1)$$

where i runs over all jets used for b -tagging and j considers all VR track jets with a $p_T > 5$ GeV and at least two associated tracks. This removal aims to prevent the selection of jets with an ambiguous

association of tracks from heavy flavor decays, as the standard b -tagging calibration scale factors are not applicable in these topologies.⁸ Events are also finally removed if they contain any jet that is likely to stem from noise bursts in the calorimeter system.⁹

4.2.3 Higgs Candidate Reconstruction and Selection

As discussed previously, to identify the $h \rightarrow b\bar{b}$ candidate decays, two different sets of reconstructed objects are used to target both the low-mass **resolved** topology, where the dijet system is sufficiently separated, and the high-mass, **merged** topology where the dijet system is highly boosted. In the resolved event topologies, the Higgs candidate consists of the two highest p_T (“leading” and “subleading”) b -tagged small-R jets, whose four-momentum vector is the sum of the individual jet four-momenta. If only one b -tagged small-R jet is present in the event, the highest p_T jet from all non b -tagged jets is additionally used. These respectively define the **2-tag** and **1-tag** analysis categories. The leading jet in the pair must have at least $p_T > 45$ GeV. In the merged event topologies, the leading large-R TCC jet is used if it has at least two ghost-associated VR track jets where at least one is b -tagged. An illustration of the two different topologies is shown in Figure 4.3.

Two dedicated correction procedures are applied to improve the mass resolution of the reconstructed Higgs candidates. “Muon-in-jet” energy corrections [169] are applied to the jets to account for semileptonic b -hadron decays, where if any muons are found within a cone of p_T dependent size around the jet axis, the four-momentum of the closest muon in ΔR with $p_T > 5\text{GeV}$ is added to the jet four-momentum after subtracting the muon’s expected energy depositions in the calorimeter. Additional flavor-specific energy corrections, determined from simulated SM $Vh(h \rightarrow b\bar{b})$ events and referred to as “ p_T reconstruction corrections”, are applied to account for biases in b -jet responses. These are derived by calculating the ratio of the p_T of the generator-level b -jets to the p_T of the reconstructed b -tagged jets after the muon-in-jet correction is applied.

To create the analysis **signal regions (SRs)** and **control regions (CRs)**, requirements are placed on the invariant mass of the dijet system m_{jj} or the large-R jet mass m_J . In the resolved regime, the dijet mass m_{jj} is determined via the sum of the four vectors of the two selected small-R calorimeter jets, while in the merged regime, the large-R TCC jet mass m_J is calculated using a

⁸This overlap removal does, unfortunately, lead to a non-negligible loss of signal acceptance, on the order of a few percent varying by channel and hypothetical HVT mass. Improvements in jet reconstruction and b -tagging could recover some of this efficiency for a future iteration of the search!

⁹A special veto is applied to help mitigate an oversaturation problem in the forward EM calorimeters during specific high pileup runs during the 2015 and 2016 datataking periods. This is affectionately referred to as the “*bad Batman veto*” due to the Batman-logo-like shape of the resulting pathological η distributions.

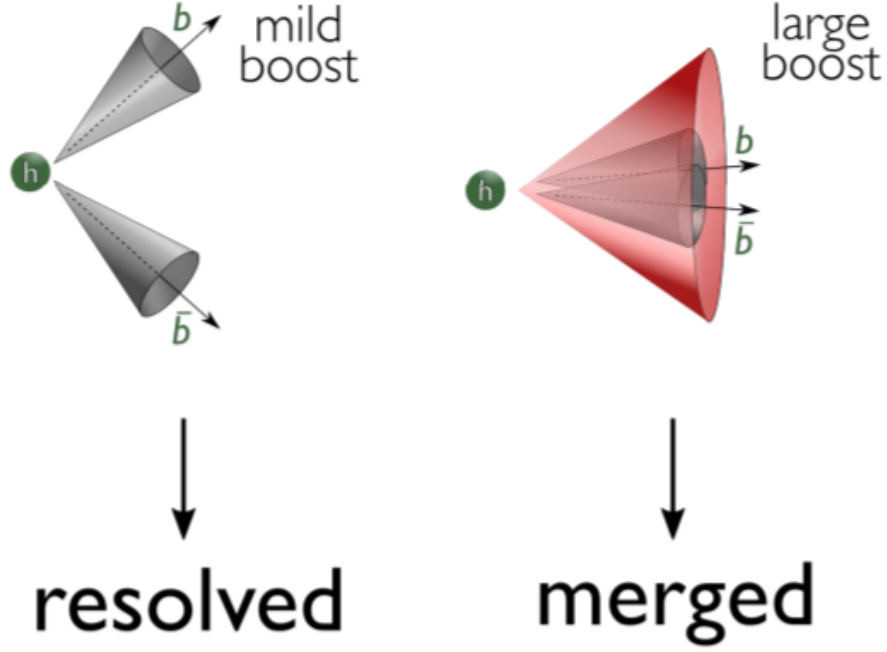


Figure 4.3. Illustration of the resolved and merged $H \rightarrow b\bar{b}$ jet topologies.

combination of track and calorimeter information depending on the number and type of contributing TCCs (see Section 2.3.4.1) as outlined in Ref [81]. The signal regions are defined in a central window around the SM Higgs boson mass of $m_h = 125$ GeV, requiring m_{jj} within 110-140 GeV in the resolved regime and m_J within 75-145 GeV in the merged regime. The control regions are defined in the mass sidebands, requiring m_{jj} or m_J between 50-200 GeV while not falling into the central Higgs mass window, which results in very similar kinematic distributions for the major backgrounds while remaining nearly signal-free. These are used in the final fit to constrain the normalization of $t\bar{t}$ and W/Z +jets backgrounds as well as improve the modeling of W/Z +jets kinematics.¹⁰ In the case that a given event passes both resolved and merged selections, priority is given to the resolved category as the resolved reconstruction method has a superior resolution.¹¹

¹⁰These are also used extensively to develop the analysis strategy where examining actual data is required, as ATLAS analyses are **blinded** until an advanced stage to avoid biasing results.

¹¹This is uncommon due to the overlap removal and the relatively orthogonal selection definitions but does occur for some events in the “transition” region between the two regimes, around $m_{Wh} = 1.2$ TeV.

4.2.4 Event Selection Requirements

The search for Wh resonances is carried out in two distinct lepton channels based on the number of selected charged lepton candidates. The primary channel, the **1-lepton** channel, selects one charged lepton and E_T^{miss} to reconstruct the leptonic W decay and constitutes the substantial majority of the search sensitivity. The **0-lepton** channel, selecting exactly zero charged leptons and large E_T^{miss} , adds a small amount of additional sensitivity by recovering hadronic τ -lepton decays and “missed” reconstruction of the primary lepton from the W decay. As previously discussed, events are further classified into either **resolved** or **merged** regimes depending on the Higgs candidate topology and **1-** and **2-tag** categories based on the number of associated b -tagged jets. All signal and control region categories used in the final statistical analysis are summarized in Table 4.3.

Channel	Region	Signal regions	Control regions
Resolved			
0-lepton	b -tags	1, 2 b -tags	1, 2 b -tags
	Mass window	$110 < m_{jj} \text{ (GeV)} < 140$	$50 < m_{jj} \text{ (GeV)} < 110 \parallel 140 < m_{jj} \text{ (GeV)} < 200$
1-lepton	b -tags	1, 2 b -tags	1, 2 b -tags
	Mass window	$110 < m_{jj} \text{ (GeV)} < 140$	$50 < m_{jj} \text{ (GeV)} < 110 \parallel 140 < m_{jj} \text{ (GeV)} < 200$
Merged			
0-lepton	b -tags	1, 2 b -tags	1, 2 b -tags
	Mass window	$75 < m_J \text{ (GeV)} < 145$	$50 < m_J \text{ (GeV)} < 75 \parallel 145 < m_J \text{ (GeV)} < 200$
1-lepton	b -tags	1, 2 b -tags	1, 2 b -tags
	Mass window	$75 < m_J \text{ (GeV)} < 145$	$50 < m_J \text{ (GeV)} < 75 \parallel 145 < m_J \text{ (GeV)} < 200$

Table 4.3. A list of the signal and control regions included in the statistical analysis of the W' signal model hypothesis.

4.2.4.1 1-Lepton Channel

Selected events must have exactly one electron or muon with $p_T > 27$ GeV. To select for the neutrino and suppress the QCD multijet background, events in the resolved category are required to have $E_T^{miss} > 80$ or 40 GeV in events with an electron or muon, respectively. In the merged category, this requirement is increased to $E_T^{miss} > 100$ GeV for both the e and μ channels. In the

resolved category only, events with more than three b -tagged small-R jets are also vetoed to suppress $t\bar{t}$ background.

The primary analysis observable, the reconstructed Wh resonance mass m_{Wh} , is calculated from the sum of the four-momenta of the W boson and Higgs boson candidates. The W boson is itself reconstructed from the sum of the four-momenta of the charged lepton and the neutrino, whose transverse component is measured as E_T^{miss} . The momentum of the neutrino in the z direction, $p_{z,\nu}$, is not possible to measure directly but is obtained by imposing a W boson mass constraint on the lepton-neutrino system, where $p_{z,\nu}$ is taken to be either the real component of the complex solutions or the real solution with the smallest magnitude from the resulting quadratic equation.

A m_{Wh} -dependent requirement is placed on $p_{T,W}$, where it must be larger than both 150 GeV and $[710 - (3.3 \cdot 10^5)/m_{Wh}]$ GeV in the resolved category and more than 150 GeV and $[394 \cdot \ln m_{Wh} - 2350]$ GeV in the merged category, with m_{Wh} in GeV. This exploits the fact that, broadly, signal decays emerge from very high mass W' mediators, produced nearly at rest, resulting in “back-to-back” decay topologies with very high p_T . Meanwhile, background processes give rise to lower p_T decays from a wide range of lower-mass initial states. This results in a stronger correlation for m_{Wh} and $p_{T,W}$ in signal than in the various backgrounds, which can be used in a two-dimensional selection to boost discrimination power.

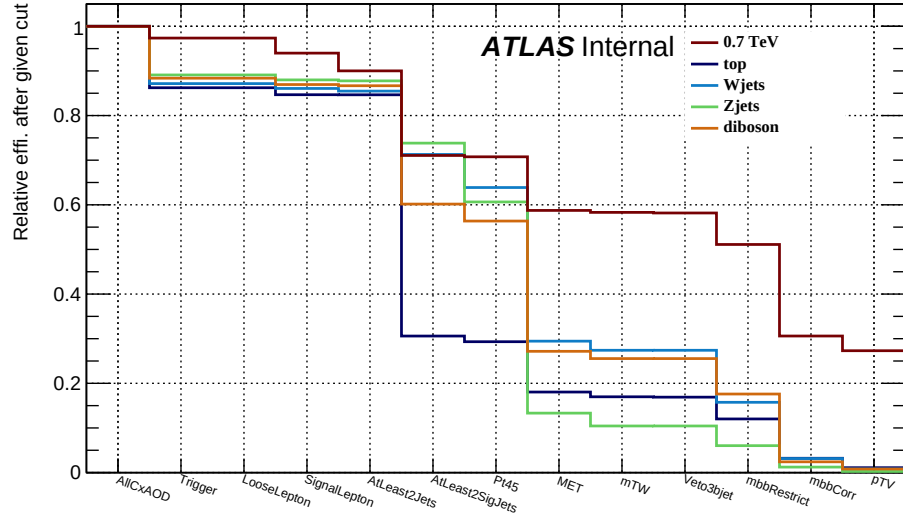
The transverse mass of the W candidate,

$$m_{T,W} = \sqrt{(p_{T,\ell} + E_T^{miss})^2 - (\vec{p}_{T,\ell} + \vec{E}_T^{miss})^2}, \quad (4.2)$$

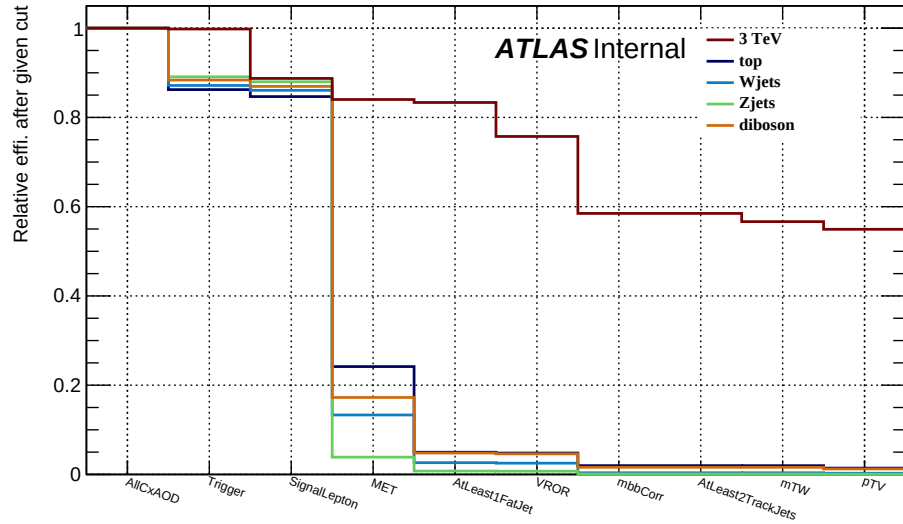
is required to be below 300 GeV to remove events where the E_T^{miss} or lepton momentum is not well reconstructed. Finally, the angular separation between the charged lepton and the Higgs boson candidate must satisfy $\Delta R(\ell, h) > 2.0$ to select signal-like topologies.¹² A simplified flowchart of the 1-lepton analysis selection into the signal and control regions is shown in Figure 4.4. The analysis “cutflows”, showing relative proportions of signal and background in the signal regions after successively applying the selections, is shown in Figure 4.5. The selections reduce the total background by a large fraction while leaving a much higher proportion of signal.

The overall kinematic acceptance times reconstruction efficiency as a function of the resonance mass is shown for the 1-lepton channel in Figure 4.6a. Acceptance is defined as the fraction of simulated signal events where the final state particles pass all of the object definition requirements,

¹²This requirement has the additional, secondary purpose of maintaining orthogonality with other ATLAS searches that use an HVT interpretation, such that they can be more easily incorporated into a more extensive combination. Such a combination, including this result, is now published! (discussed further at the conclusion of this Chapter)



(a)



(b)

Figure 4.5. Efficiency of signal and main background components after successively applying analysis selections in the 1-lepton resolved (a) and merged (b) categories for selected W' masses of 700 GeV (resolved) and 3 TeV (merged). *AllCxAOD* refers to the initial data type used after the object definitions but before any analysis selections, and MET refers to E_T^{miss} . In the resolved, *mbbRestrict* corresponds to a requirement of $50 < m_{jj} < 200$ GeV, while *mbbCorr* refers to the mass window selection requirement in both categories. *VROR* refers to the overlap removal for overlapping variable-radius track jets.

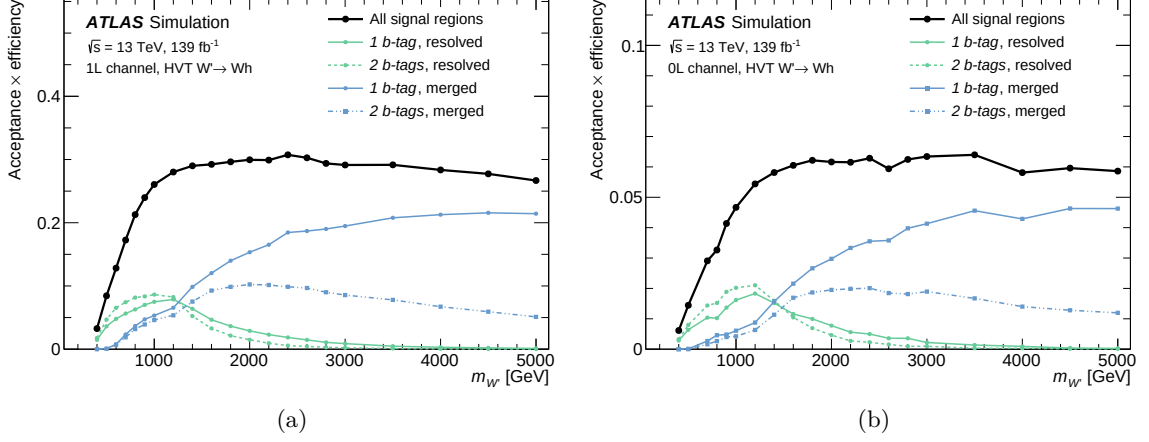


Figure 4.6. Product of acceptance and efficiency for $W' \rightarrow Wh \rightarrow \ell^\pm \nu b \bar{b} / c \bar{c}$ as a function of the resonance mass for the 1-lepton (a) and 0-lepton (b) signal regions.

where $E_{T,h}$ and $\vec{p}_{T,h}$ are the energy and momentum vector of the Higgs candidate in the transverse plane.

To preferentially select for E_T^{miss} that is likely to come from an “invisible” particle, like that from a neutrino, a requirement is placed on E_T^{miss} significance $\mathcal{S}$ from 9.0-13.6, increasing linearly as a function of the resonance mass beginning at $m_{T,Wh} = 240 \text{ GeV}$ and reaching the maximum at 700 GeV.

Several topological selections on the azimuthal separations are employed between the analysis objects to select for signal-like, “back-to-back” final states in the absence of a well-reconstructed lepton. The separation $\Delta\phi$ between $E_T^{\vec{\text{miss}}}$ and $p_T^{\vec{\text{miss}}}$ is required to satisfy $\Delta\phi(E_T^{\vec{\text{miss}}}, p_T^{\vec{\text{miss}}}) < \pi/2$, the separation between $E_T^{\vec{\text{miss}}}$ and the Higgs candidate is required to satisfy $\Delta\phi(E_T^{\vec{\text{miss}}}, h) > 2\pi/3$, and the minimum separation between $E_T^{\vec{\text{miss}}}$ and any reconstructed small- R jets in the event must satisfy $\min[\Delta\phi(E_T^{\vec{\text{miss}}}, \text{small-}R \text{ jets})] > \pi/9$ for events with up to three jets and $\pi/6$ for those with more than three jets. For the resolved category, the separation between the two Higgs candidate jets is required to be $\Delta\phi < 7\pi/9$. The scalar sum S_T of all jet p_T values must satisfy $\sum |p_T| > 120 \text{ GeV}$ if the event contains two central jets or more than 150 GeV if the event includes more than two.

The acceptance times efficiency for the 0-lepton channel is shown in Figure 4.6b, showing similar behavior as the 1-lepton channel with the resolved category dominating at low W' masses and vice versa for the merged, with the total efficiency plateauing at around 6%. A summary of the complete selection criteria for the two channels is shown in Table 4.4.

Variable	Resolved	Merged
Common selection		
Number of jets	≥ 2 small- R jets (0-lep.) 2 or 3 small- R jets (1-lep.)	≥ 1 large- R jet ≥ 1 VR track-jets (matched to leading large- R jet) ^{††}
Leading jet p_T [GeV]	> 45	> 250
m_h [GeV]	110–140	75–145
1-lepton selection		
Leading lepton p_T [GeV]	> 27	> 27
E_T^{miss} [GeV]	> 40 (80 [†])	> 100
$p_{T,W}$ [GeV]	$> \max[150, 710 - (3.3 \cdot 10^5 \text{ GeV})/m_{Wh}]$	$> \max[150, 394 \cdot \log(m_{Wh}/(1 \text{ GeV})) - 2350]$
$m_{T,W}$ [GeV]	< 300	
$\Delta R(\ell, h)$	> 2.0	
0-lepton selection		
E_T^{miss} [GeV]	> 150	> 200
S_T [GeV]	> 150 (120*)	–
$\Delta\phi_{jj}$	$< 7\pi/9$	–
p_T^{miss} [GeV]	> 60	
$\Delta\phi(\vec{E}_T^{\text{miss}}, \vec{p}_T^{\text{miss}})$	$< \pi/2$	
$\Delta\phi(\vec{E}_T^{\text{miss}}, h)$	$> 2\pi/3$	
$\min[\Delta\phi(\vec{E}_T^{\text{miss}}, \text{small-}R \text{ jet})]$	$> \pi/9$ (2 or 3 jets), $> \pi/6$ (≥ 4 jets)	
$N_{\tau_{\text{had}}}$	≤ 1	
E_T^{miss} significance $\mathcal{S}$	$\begin{cases} > 9 & \text{if } m_{Wh} < 240 \text{ GeV}, \\ > 6.6 + 0.01 \cdot m_{Wh} & \text{if } 240 \text{ GeV} \leq m_{Wh} < 700 \text{ GeV}, \\ > 13.6 & \text{if } m_{Wh} > 700 \text{ GeV}, \end{cases}$	

Table 4.4. Topological and kinematic selections for each channel and category as described in the text. (*) Applies in the case of only two central jets. (†) A higher threshold (80 GeV) is used for the single-electron channel. (††) Applied only for $m_{Wh} > 320$ GeV. (†††) Only the two leading VR track-jets matched by ghost-association to the large- R jet are considered when classifying events into b -tag categories. Events are further classified according to the number of b -tagged jets.

4.3 Background Estimation

The background composition in each signal region depends strongly on the charged lepton and b -jet multiplicities, and the compositions are thus studied on a per-category basis. The dominant backgrounds for the 1-lepton channel are $t\bar{t}$, single-top Wt production, and W +jets, with the relative contribution from top quark backgrounds decreasing slightly with increasing m_{Wh} . For the 0-lepton channel, $t\bar{t}$ and Z +jets predominate with a moderate contribution from W +jets. Contribution from SM VV and Vh production and $t\bar{t} + V/h$ are small but non-negligible in both channels.

Contributions from QCD multijet events were studied and found to be negligible in the 0-lepton channel and the merged category of the 1-lepton channel, with a small contribution found in the resolved 1 b -tag category of the 1-lepton channel only. As the Monte Carlo simulations of QCD events have poor modeling and limited statistics in the analysis phase-space region, this contribution is estimated using a data-driven template method, where a dedicated control region is created

by inverting the lepton isolation requirements described in Section 4.2.1. Leptons from signal processes are expected to be well-isolated from hadronic activity since the “back-to-back” topology generally has the W boson very well separated from the Higgs decay. On the other hand, leptons from multijet events originate inside jets and, thus, by definition, are very unlikely to be isolated; therefore, inverting the isolation requirements creates a regime highly enriched in multijet contributions. The shape of the multijet m_{WH} template is obtained by subtracting the remaining Monte Carlo background contributions in the control region. The normalization of the template is derived from a two-component fit of the multijet and non-multijet contributions to data in the m_{jj} sideband control region. Further details on the template method are available in Appendix B and Ref [170].

The remaining background distributions are estimated from the MC simulations, with normalizations estimated from the data. In the 0-lepton channel, to correct for observed differences between data and simulation in several kinematic distributions, the simulated W/Z +jets samples are reweighted as a function of the dijet system’s transverse momentum, $p_{T,jj}$, using correction scale factors derived from fits to data. Further details on this reweighting procedure are also available in Appendix B.

Background processes enter the statistical analysis, which is discussed further in Section 4.5, grouped into different components based on their expected contributions from category to category and their expected event yield. The $t\bar{t}$, single-top quark, $t\bar{t} + V$, and $t\bar{t} + h$ processes are combined into a single fit component, referred to as the combined “top” background. As discussed previously, the W/Z +jets events are split into three different jet flavor components to help improve the final fit modeling as the contributions of heavy-flavor jets vary between the 1 and 2 b -tag categories. If in the resolved category, the two selected small- R jets have two ghost-associated generator-level b - or c -quark “truth” jets, they are classified as W/Z +hf, while one or zero are categorized as W/Z +hl and W/Z +lf, respectively. The merged category follows a similar strategy using the associated VR track-jets. In the statistical analysis, the global normalization of the heavy-flavor and “heavy-light” W/Z +jets backgrounds are determined via a fit to data depending on the lepton channel, while the normalization of the light-flavor components are taken directly from simulation for both channels as their expected contributions are much smaller.

The global normalizations of the dominant background contributions are determined separately for the 0- and 1-lepton channels via a combined fit to data in both signal and control regions. The m_{jj}/m_J sideband regions are used to provide additional information to constrain the top-quark and W/Z +jets background normalizations.

4.4 Systematic Uncertainties

It is, unfortunately, impossible to perform perfect measurements and simulations¹³, particularly with an experimental apparatus as intricate and complex as ATLAS. As such, the distributions of m_{Wh} and $m_{T,Wh}$ are affected by a vast number of systematic uncertainties that enter the final fit as **nuisance parameters**, which will be discussed further in Section 4.5. A given source of uncertainty can modify either the **shape** or **normalization** of the final discriminant distributions, defined through the creation of one standard deviation **variations** with respect to the **nominal** distributions which then enter the statistical analysis framework.

Systematic uncertainties can be roughly classified into two groups: **experimental** and **modeling** uncertainties. Experimental uncertainties arise from particle reconstruction, identification, and calibration inefficiencies, as well as measurement uncertainties in particle energies and momenta. Modeling uncertainties arise from theory predictions, such as renormalization and factorization scales and QCD calculation order, as well as generator-level and simulation-level effects, such as the choice of MC generator, parton shower model, PDF set, hadronization, etc. Uncertainties are broadly correlated in varying proportions between signal and background processes among differing channels, kinematic regions, and observable distributions. Experimental uncertainties, as they arise from the detector response, are correlated across all regions and between physics processes. In contrast, modeling uncertainties are generally more specific to a given process or analysis category.

4.4.1 Experimental Uncertainties

The relevant experimental uncertainties for the analysis are listed below, in approximate order of overall size and impact. The effects of experimental uncertainties on the final fit parameterization are, on average, substantially smaller than the modeling uncertainties, discussed in the next section.¹⁴

Small-R jets:

A large number of uncertainties are defined for small-R jets from *in situ* performance analyses associated with the calibration of the small-R jet energy scale (JES), mass scale (JMS), and energy resolution (JER) [55, 171, 172, 173]. Determining the JES and JMS is a complex procedure that introduces multiple sources of uncertainty in each stage. The calibrations are performed with a large number of free parameters that are decomposed into a smaller number of effective parameters

¹³If only!

¹⁴They are very often small enough to be *pruned* by the fit!

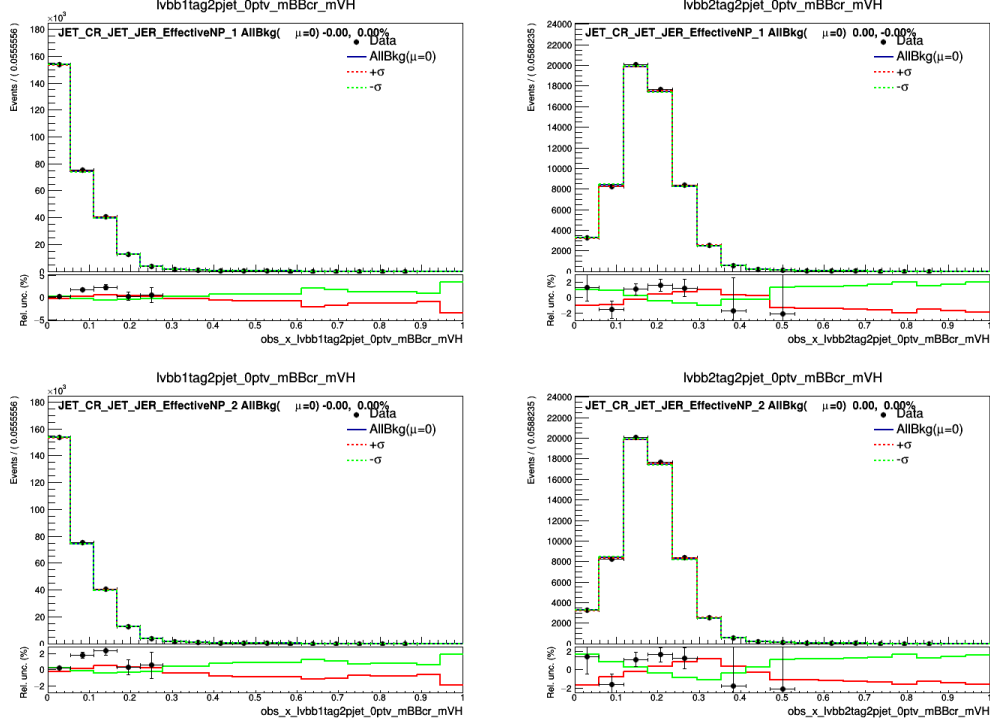


Figure 4.7. Pre-fit m_{Wh} distributions of the JET_CR_JET_JER_EffectiveNP1 (top) and 2 (bottom) uncertainties for 1 b -tag (left) and 2 b -tag (right) in the m_{jj} sideband control regions of the 1-lepton resolved regime. Normalized entries are shown in the primary panel and the relative variation is shown in the ratio panel, below. The x-axis is expressed as the fraction of the maximum tested m_{Wh} value, 5000 GeV.

grouped based on their origin, resulting in p_T and η dependent uncertainties from detector effects, choice of fragmentation and hadronization model, overall statistical uncertainty, and a mixed term. They also encompass uncertainties from the differences between data and simulation in different kinematic phase space regions, dijet η inter-calibration to account for η -dependent mismodeling effects, “punch-through” energy estimation, high- p_T jet behavior, b - and c -jet response and composition, and pileup dependency. The width of the JER is determined through Gaussian smearing of jet energies in MC simulation, where the fit covariance is decomposed into the seven leading effective terms plus one additional residual uncertainty. This effect is propagated to the m_{Wh} distributions through an additional smearing corresponding to each effective parameter, resulting in eight separate JER uncertainties. Additionally, one uncertainty is defined to account for the efficiency of the jet vertex tagger. Example pre-fit uncertainty variations for two of the small-R JER effective terms for the 1-lepton resolved regime are shown in Figure 4.7.

Each of these uncertainties is treated as an individual parameter in the final statistical analysis. Together, in quadrature, they result in an absolute maximum of approximately 3% normalization difference in the resolved signal regions and 2% in the merged signal regions.

Large-R jets:

Several systematic uncertainties are defined for the calibration of the large-R jet energy scale and determination of the large-R jet energy resolution and reconstructed mass resolution (JMR). The large-R JES uncertainties are estimated by comparing the ratio of calorimeter-based to track-based measurements in observation and prediction, where the covariance matrix describing the uncertainties in the calibration metric is decomposed into five orthogonal components [174, 175]. These include a baseline term corresponding to the track-to-calorimeter ratio, residual ratio uncertainties, track measurement uncertainties, the difference between SHERPA and PYTHIA samples in the hadronization, and statistical uncertainties. The large-R JER uncertainty is determined through a similar smearing method as described for the small-R jets. The large-R JMR uncertainty is also produced through a similar process where the jet mass, rather than the energy, is smeared around the nominal resolution value by artificially modifying the four-momentum with Gaussian distributed random values representing a 20% worse mass resolution. Example JMR variations for the 1-lepton, 1 b -tag merged signal region for $t\bar{t}$ and W +heavy flavor jets backgrounds are shown in Figure 4.8.

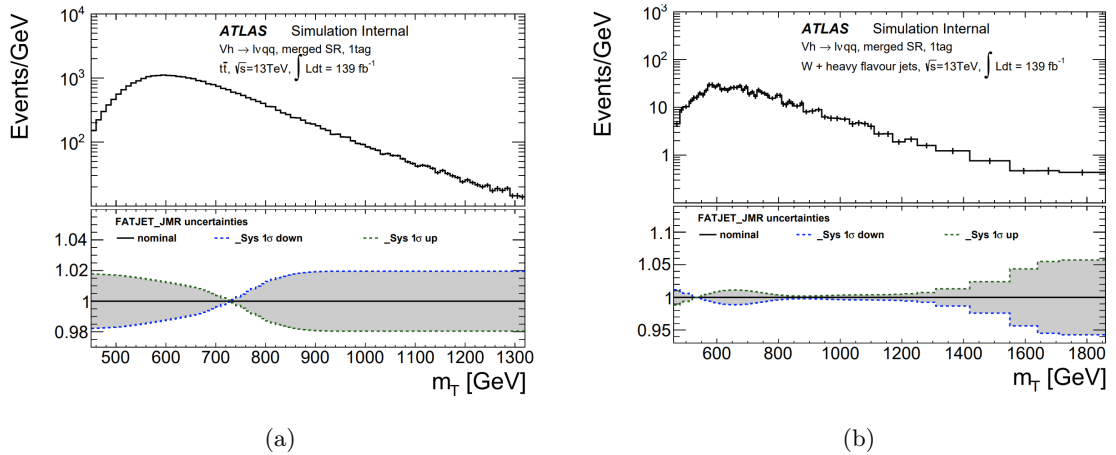


Figure 4.8. Pre-fit distributions of $m_{T,Wh}$ for $t\bar{t}$ (a) and W +heavy flavor jets (b) backgrounds in the 1-lepton, 1 b -tag merged signal region. The lower panel shows the relative variations with respect to the nominal induced by the resolution of the large-R jet after averaging and smoothing.

***b*-tagging:**

Uncertainties on the *b*-tagging efficiency can affect the m_{Wh} distributions, particularly between the different *b*-tag multiplicity categories. Differences between the efficiency measured in data versus simulation result in separate correction factors for *b*-jets, *c*-jets, and light-flavor jets [161, 167, 168], resulting in three, four, and five uncertainties, respectively, when reduced into uncorrelated components. An additional uncertainty is also introduced from the extrapolation of the measured uncertainties to the high- p_T regime outside the calibration regime. This is determined from simulation by considering variations of the quantities that affect *b*-tagging performance, such as impact parameter resolution, rate of random ID track crossings, and track multiplicities per jet. These uncertainties are defined for both the small-R calorimeter jets used in the resolved category and the VR track jets used in the merged category, resulting in two separate “sets”.

Electrons:

Reconstructed electrons carry separate uncertainties on the reconstruction, identification, and isolation efficiencies, derived from differences between data and simulation in “tag-and-probe” measurements of SM $Z \rightarrow \ell\ell$ data events [93]. They additionally have uncertainties derived for their energy scale and resolution and uncertainties on the single-electron trigger scale factors used to correct for the differences between data and simulation in the “turn-on” regime. The overall effect of these uncertainties on the final fit is minimal, amounting to a sub-1% maximum normalization difference.

Muons:

Muons carry analogous uncertainties to electrons derived through similar measurements, with some additional elements [76]. They also have additional uncertainties relating to the track-to-vertex association efficiency, as well as separate reconstruction efficiency uncertainties for the low- p_T and high- p_T regimes and separate energy resolution uncertainties for both the ID and MS. Several dedicated additional uncertainties are also introduced for the muon energy resolution relating to the *sagitta bias*, a spatial-resolution effect pertaining to the accuracy of the detector alignment model. Very small discrepancies in the alignment can cause a gradient of systematic bias in muon trajectory parameters over the detector, which can cause a non-negligible impact in a muon momentum measurement for high- p_T muon tracks as they have a nearly straight-line trajectory that is more sensitive to measurement errors. As for electrons, the overall effect of these uncertainties is also

minimal, amounting to a sub-1% maximum normalization difference.

Missing Transverse Momentum:

As E_T^{miss} is calculated using the reconstructed objects, the systematic uncertainties related to each object propagate to the final E_T^{miss} . Additionally, modeling of the underlying event and the momentum scale, resolution, and reconstruction efficiency of the unassociated tracks used to compute the soft-term introduces additional uncertainties, as well as another uncertainty on the energy scale of tracks in jets [96, 97]. The E_T^{miss} trigger efficiency also contributes another final uncertainty [169].

τ leptons:

Uncertainties are defined for the τ lepton reconstruction efficiency and calibration [92] but have a negligible impact on the final result as τ leptons only enter the analysis as a counting veto.

Luminosity and Pileup:

In simulation, events are generated assuming a static pileup profile for each datataking period, while real average pileup in data varies from run to run within a period. This is accounted for with a *pileup reweighting* procedure, which carries an associated uncertainty. Finally, a global luminosity uncertainty of 1.7% is applied to the normalization of all simulated events [176].

4.4.2 Modeling Uncertainties

Modeling uncertainties in this analysis come in two overarching varieties: one type comes from theory considerations and generator-level effects derived from alternate samples, and the other comes from analysis-level modeling considerations necessitating the creation of bespoke uncertainties. These uncertainties can significantly impact the final fit parameterization, as detailed in the next section.

4.4.2.1 Theory Uncertainties

Theoretical uncertainties are derived for all simulated signal and background processes, leading to variations in the shape and normalization of the templates that enter the statistical analysis. These uncertainties encompass a wide range of effects introduced by various choices made at the generation level, including the matrix element generator and the associated matching to the parton shower, parton shower model, hadronization, underlying event, PDF sets, and QCD renormalization

and factorization scales. These uncertainties affect the distribution of kinematic variables, including the final discriminants, and are evaluated on the expected signal and background yields in the corresponding phase space regions.

For the major backgrounds, the theory uncertainties translate into pre-fit probability distributions (priors) for nuisance parameters. Direct normalization uncertainties, such as the ones defined for more minor backgrounds, cover effects on both the cross section and acceptance.

Systematic uncertainties corresponding to the matching of the matrix element and the parton shower, the parton shower modeling, hadronization, or underlying event, are estimated by comparing the prediction of the nominal MC generator process to those of an alternative generator for each separate process. In general, the perturbative precision and the PDF sets used in the alternative configurations match those of the nominal generators unless otherwise stated, while uncertainties resulting from the choice of PDF set are separately evaluated. QCD scale uncertainties are evaluated by individually shifting the renormalization and factorization scales μ_R and μ_F up and down up to a factor of 2 in varying combinations, such that the final QCD scale uncertainty is obtained as the most prominent variation with respect to the central value.

Further specifics on individual uncertainties for each simulated process are listed below.

HVT W' signal:

Uncertainties for the W' signal acceptance were derived by replacing the nominal PDF set by the MSTW2008LO [177] and CT10 [178] PDF sets, by using the eigenvariations of the PYTHIA 8.186 A14 tune [124], and by replacing PYTHIA 8.186 with HERWIG 7.0.4 [179]. QCD scale variations were also considered. The effects of these variations are ultimately minimal for the W' signal, particularly in the 1-lepton channel, amounting to a normalization uncertainty of 2-7% depending on channel and region.

$t\bar{t}$:

As the dominant background of the search with high statistics in the signal regions, the normalization of $t\bar{t}$ is allowed to “float” in the fit, meaning that the normalization is left unconstrained and is freely determined by the fit without a penalty (discussed further in the next section). Thus, theory uncertainties on the $t\bar{t}$ background are relevant as shape uncertainties only. To assess uncertainties in the matching of the matrix elements to the parton shower, the nominal sample is compared with an

alternate sample where POWHEG BOX is replaced by MADGRAPH5_AMCNLO. The parton shower modeling uncertainty is assessed by replacing PYTHIA 8.230 with HERWIG 7.04 using the H7UE tune. The impact of a decreased or increased amount of initial and final state radiation (ISR/FSR) produced in association with the $t\bar{t}$ system is determined by recalculating the event weights using adjusted settings for the renormalization and factorization scales in combination with changing the eigenvariations in the PYTHIA A14 tune.

Uncertainties in the PDF were determined through several comparisons, two from comparing the yields from the nominal set in a particular phase space region with those from the MMHT2014NNLO and CT14NNLO PDF sets [180]. The third is determined by taking the standard deviation of the 100 NNPDF3.0NLO replicas with varying α_s [120]. The largest of these three variations is then taken to be the PDF uncertainty. Example distributions of the $t\bar{t}$ theory uncertainties for the 1-lepton signal regions are shown in Figures 4.9 and 4.10 for the resolved and merged categories, respectively. $t\bar{t} + h$ and $t\bar{t} + V$ production have extremely small contributions to the combined top background and are assigned a simple flat normalization uncertainty of 50%.

Single-top quark:

As single-top production is a relatively small background with a lower impact on the final fit result, it is combined with $t\bar{t}$ in the final fit and assigned an overall normalization variation uncertainty of 19%. In the nominal samples, the interference between single top Wt and $t\bar{t}$ production is removed through the diagram removal (DR) scheme by setting the amplitudes of the overlapping diagrams to zero before the simulation. An alternate method to account for this involves subtracting the cross section instead through the diagram subtraction (DS) scheme. The effect is negligible in the 0-lepton channel, but as Wt production comprises nearly 10% of the total top background in the 1-lepton channel, a special uncertainty is derived to account for this difference, which becomes non-negligible in the high-mass tails of the m_{Wh} distributions. Further details on the derivation of this uncertainty are available in Appendix B.

W/Z+jets:

To account for variations of both the matrix element and the parton shower, the nominal samples were compared with samples with LO-accurate matrix elements that contain up to four final-state partons, generated with MADGRAPH5_AMCNLO 2.2.2 using the NNPDF3.0NLO PDF set, interfaced to PYTHIA 8.186 using the CKKW-L merging procedure [181, 182]. These also use the PYTHIA A14 tune with the NNPDF2.3LO PDF set, with decays of bottom and charm hadrons performed

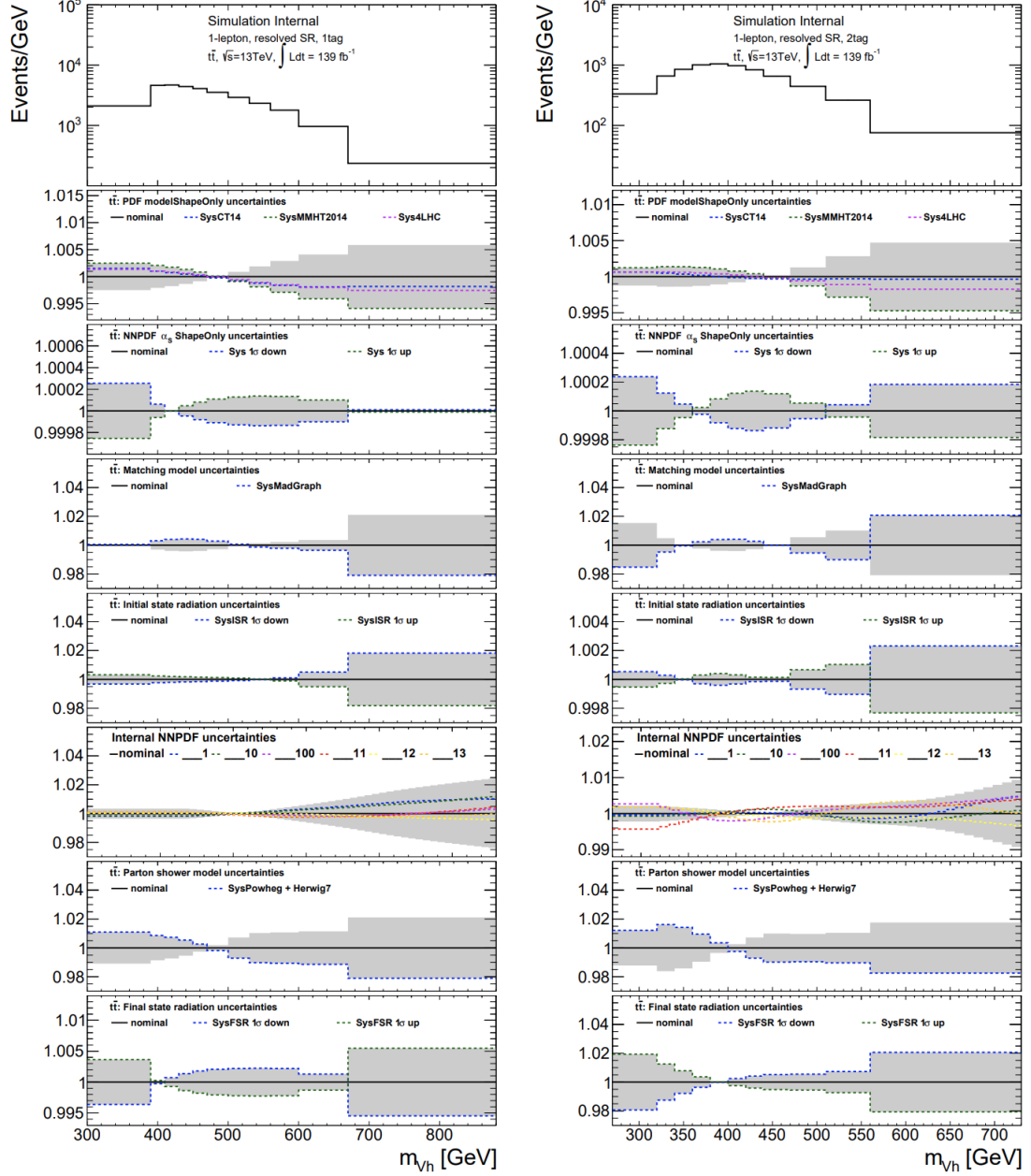


Figure 4.9. Comparisons of pre-fit m_{Wh} -distributions of $t\bar{t}$ inside the combined 1-lepton **resolved** signal regions as predicted by the nominal generation versus the alternative generations as described in the text, labeled in each ratio panel. Variations are normalized with respect to the nominal to study pure shape uncertainties and are shown after Gaussian kernel smoothing. For legibility, only 5 of the 100 NNPDF variations are shown.

by EVTGEN 1.2.0. Uncertainties in the PDF were determined using the same prescription as for $t\bar{t}$, as well as QCD scale uncertainties.

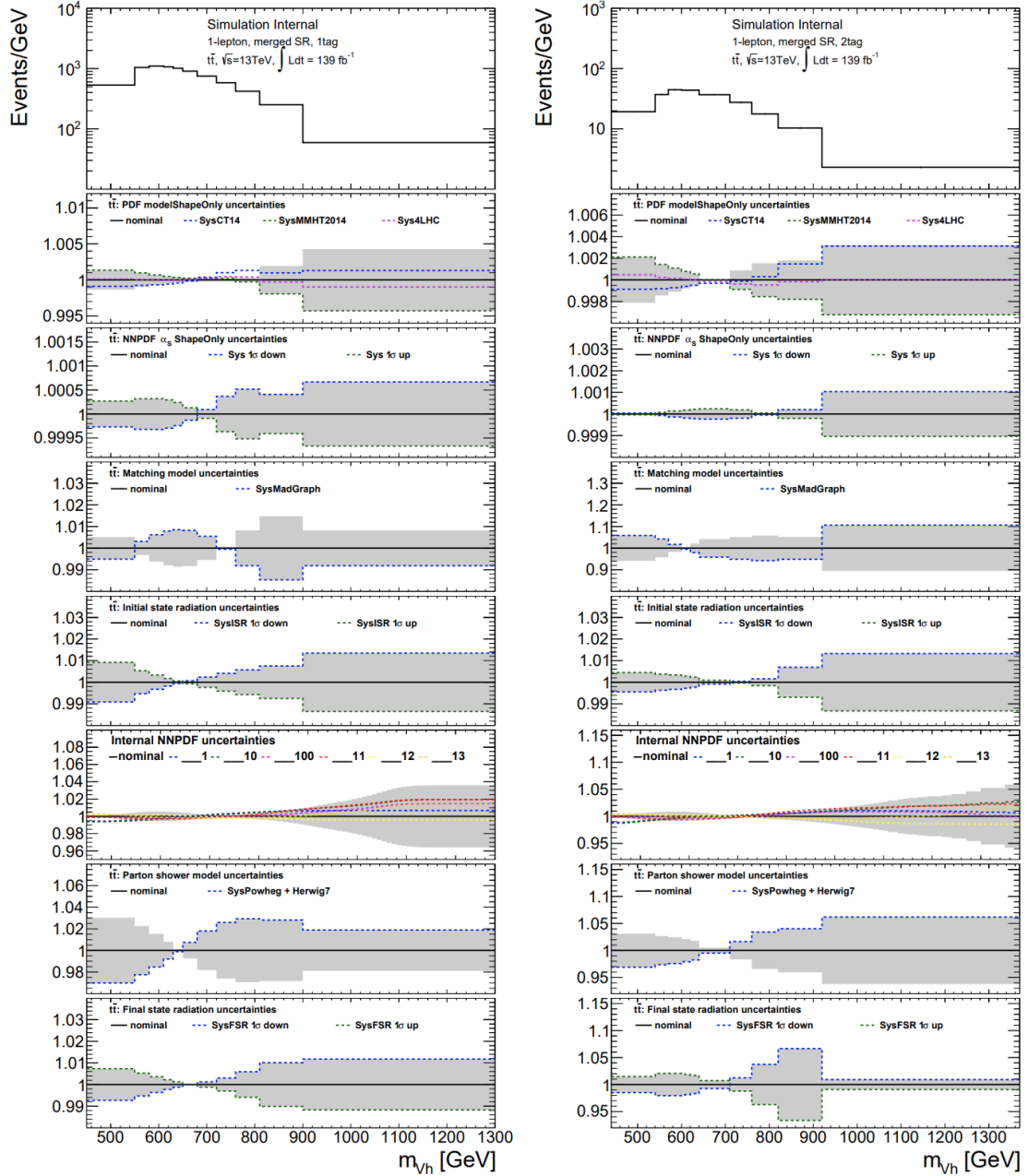


Figure 4.10. Comparisons of pre-fit m_{Wh} -distributions of $t\bar{t}$ inside the combined 1-lepton **merged** signal regions as predicted by the nominal generation versus the alternative generations as described in the text, labeled in each ratio panel. Variations are normalized with respect to the nominal to study pure shape uncertainties and are shown after Gaussian kernel smoothing. For legibility, only 5 of the 100 NNPDF variations are shown.

In the 1-lepton channel, due to the active veto of additional (more than 3) reconstructed small-R jets in the selection criteria, additional *JetBinMigration* uncertainties are derived for the W +jets

backgrounds to account for the modeling of extra jets from QCD contributions etc [183]. These are evaluated by varying the contributions from W +jets events with either less than two jets or more than three jets at the generator level by $\pm 50\%$. In the 0-lepton channel, additional uncertainties are derived for the $p_{T,jj}$ reweighting of the W/Z +jets backgrounds (discussed previously in Section 4.3) from the difference between the original and reweighted distributions. Further details on both of these uncertainties can be found in Appendix B.

The uncertainties are split among the three flavor composition categories for both W and Z +jets. Normalization of the W +jets distributions is allowed to float in the 1-lepton fit, while the normalization of Z +jets distributions is allowed to float in the 0-lepton fit, as these respective channels have sufficient statistics to determine the normalizations properly.¹⁵

SM Diboson (VV, Vh):

Flat 50% normalization uncertainties are assigned for the various SM diboson backgrounds, as they contribute a very small proportion of the overall background composition.

QCD Multijet:

Several conservative uncertainties are assigned to the data-driven multijet background estimate, where the normalization uncertainty is set to 50%, and a shape uncertainty is derived from the differences when varying the E_T^{miss} selection requirement by $\pm 25\%$, yielding a shape uncertainty of up to 40% for the m_{Wh} template. However, as the multijet background is only non-negligible in the 1 b -tag resolved regions of the 1-lepton channel and has a minimal contribution there regardless, its uncertainties ultimately have a negligible impact on the final result.

4.4.2.2 Additional Uncertainties

Several additional modeling uncertainties, not described as theory uncertainties above, are defined for multiple processes below. All modeling systematic uncertainties at the pre-fit stage are detailed in Tables 4.5 and 4.6.

Non-Closure Uncertainties:

Residual mismodelings, which are non-closures in comparisons between data and simulation, are observed in the p_T^{miss} distribution in the 0-lepton channel, in the resolved $p_{T,W}$ distribution in the 1-lepton channel, and the $p_{T,J}$ of the leading large- R jet in both channels. These are not explicitly reweighted, and instead, data-to-simulation ratios are used to define systematic uncertainties in the

¹⁵Both sets of normalizations are global free parameters in the combined W' fit, with the final post-fit normalization scale factors being determined by the respective channel, as will be discussed in Section 4.5.

relevant shapes and applied to all backgrounds as uncorrelated nuisance parameters. The additional degrees of freedom allow the fit to “pull” on these nuisance parameters to correct the mismodelings in the final parameterization, with an associated penalty in the likelihood calculation. Additional details on the derivation of these non-closure uncertainties are also available in Appendix B.

Transfer Factor Uncertainties:

Extrapolation effects between the signal and control regions, resolved and merged categories, and different lepton channels are also taken into account, where additional transfer-factor uncertainties are defined to constrain the normalization ratios of the two regions. These provide additional degrees of freedom to the maximum likelihood fit and help describe the extrapolation of normalization scale factors from one region to another. The transfer factor for a particular process p is defined as:

$$R_{B \rightarrow A}^p = \frac{n_A^p}{n_B^p}, \quad (4.4)$$

where n_A^p and n_B^p are the pre-fit event yields of a process p in regions A and B , respectively. The corresponding uncertainty on a given transfer factor is calculated via:

$$\sigma_{R_{TF}} = \sqrt{\sum_i^M \left(\frac{R_{B \rightarrow A}^{\text{var},i} - R_{B \rightarrow A}^{\text{nominal}}}{R_{B \rightarrow A}^{\text{nominal}}} \right)^2}, \quad (4.5)$$

where $R_{B \rightarrow A}^{\text{nominal}}$ is the transfer factor corresponding to the nominal simulation model and the $R_{B \rightarrow A}^{\text{var},i}$ are the transfer factors corresponding to one of the M alternative simulation models considered for the calculation of theory uncertainties of the process p .

4.4.3 Smoothing, Symmetrization, and Pruning

Several modifications are made to the various derived uncertainty templates, as the statistical model assumes a priori that systematic uncertainties are both reasonably smooth over the ranges of m_{Wh} and $m_{T,Wh}$ and are centered symmetrically around the nominal value. Failing to account for uncertainty templates with limited statistics and asymmetric distributions can cause fit instability and artificial constraints.

4.4.3.1 Smoothing

Limited statistics in the uncertainty templates can cause unwanted behavior, such as artificial pulls and constraints. Individual bins with overestimation of uncertainties due to statistical fluctuations will cause the fit to spuriously constrain the associated nuisance parameters, and shape variations may be exaggerated by statistical fluctuations such that they are not an accurate representation of the “true” variation. In the final fit model, the discriminant distributions are rebinned

Table 4.5. Relative MC generator modeling systematic uncertainties (at the pre-fit stage), connected to the normalization, cross-region extrapolation, and shape of the W/Z +jets background processes included in the fits described in the text. An ‘S’ indicates that a shape variation is included for the sources listed, ‘ $A \leftrightarrow B$ ’ indicates a transfer-factor uncertainty between the regions A and B , and ‘norm.’ is the sum of cross-section and acceptance variations. A value of ‘float’ indicates that the parameter is not constrained in the fit. A range of values means the value depends on the lepton channel. Parentheses indicate when the uncertainty applies only to a given fit or a given region.

Process	Quantity/source	Value
Z+hf	0-lep. norm.	float
	1-lep. norm.	50%
	0-lep. resolved $\leftrightarrow$ merged	10%–19%
	1-lep. resolved $\leftrightarrow$ merged	9%–18%
	0-lep. SR $\leftrightarrow$ CR	5%–12%
	1-lep. SR $\leftrightarrow$ CR	6%–20%
	0-lep. $\leftrightarrow$ 1-lep.	7%–27%
	generator, PDF, scale	S
	$p_{T,jj}$ reweighting	S
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S
	0-lep. norm.	float
	1-lep. norm.	50%
Z+hl	0-lep. resolved $\leftrightarrow$ merged	15%–28%
	1-lep. resolved $\leftrightarrow$ merged	12%–13%
	0-lep. SR $\leftrightarrow$ CR	3%–20%
	1-lep. SR $\leftrightarrow$ CR	5%–7%
	0-lep. $\leftrightarrow$ 1-lep.	7%–27%
	generator, PDF, scale	S
	$p_{T,jj}$ reweighting	S
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S
	norm.	19%
	0-lep. resolved $\leftrightarrow$ merged	8%–50%
	1-lep. resolved $\leftrightarrow$ merged	10%–29%
	0-lep. SR $\leftrightarrow$ CR	5%–20%
Z+lf	1-lep. SR $\leftrightarrow$ CR	29%–99%
	0-lep. $\leftrightarrow$ 1-lep.	8%–39%
	generator, PDF, scale	S
	$p_{T,jj}$ reweighting	S
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S
	0-lep. norm.	20%
	1-lep. norm.	float
	0-lep. resolved $\leftrightarrow$ merged	14%–18%
	1-lep. resolved $\leftrightarrow$ merged	20%–21%
	0-lep. SR $\leftrightarrow$ CR	4%–20%
	1-lep. SR $\leftrightarrow$ CR	2%–4%
	generator, PDF, scale	S
	$p_{T,jj}$ reweighting (0-lep.)	S
W+hf	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S
	JetBinMigration (1-lep.)	S
	0-lep. norm.	30%
	1-lep. norm.	float
	0-lep. resolved $\leftrightarrow$ merged	13%–28%
	1-lep. resolved $\leftrightarrow$ merged	15%–22%
	0-lep. SR $\leftrightarrow$ CR	5%–28%
	1-lep. SR $\leftrightarrow$ CR	3%
	0-lep. $\leftrightarrow$ 1-lep.	4%
	generator, PDF, scale	S
	$p_{T,jj}$ reweighting (0-lep.)	S
	p_T^{miss} non-closure (0-lep.)	S
W+hl	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S
	JetBinMigration (1-lep.)	S
	0-lep. norm.	30%
	1-lep. norm.	float
	0-lep. resolved $\leftrightarrow$ merged	2%–43%
	1-lep. resolved $\leftrightarrow$ merged	12%–13%
	0-lep. SR $\leftrightarrow$ CR	2%–20%
	1-lep. SR $\leftrightarrow$ CR	1%–2%
	0-lep. $\leftrightarrow$ 1-lep.	4%
	generator, PDF, scale	S
	$p_{T,jj}$ reweighting (0-lep.)	S
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
W+lf	Large- R jet p_T non-closure	S
	JetBinMigration (1-lep.)	S
	0-lep. norm.	20%
	1-lep. norm.	float
	0-lep. resolved $\leftrightarrow$ merged	14%–18%
	1-lep. resolved $\leftrightarrow$ merged	20%–21%
	0-lep. SR $\leftrightarrow$ CR	4%–20%
	1-lep. SR $\leftrightarrow$ CR	2%–4%
	generator, PDF, scale	S
	$p_{T,jj}$ reweighting (0-lep.)	S
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S
	JetBinMigration (1-lep.)	S

to maximize search sensitivity in a given phase space region while minimizing statistical fluctuations; thus, the uncertainty templates are derived in a fine binning that can then be subsequently

Table 4.6. Relative MC generator modeling systematic uncertainties (at the pre-fit stage), connected to the normalization, cross-region extrapolation, and shape of the signal, top, SM Vh , and SM diboson background processes included in the fits described in the text. An ‘S’ indicates that a shape variation is included for the sources listed, ‘ $A \leftrightarrow B$ ’ indicates a transfer-factor uncertainty between the regions A and B , and ‘norm.’ is the sum of cross section and acceptance variations. A value of ‘float’ indicates that the parameter is not constrained in the fit. A range of values means the value depends on the lepton channel. Parentheses indicate when the uncertainty applies only to a given fit or a given region.

Process	Quantity/source	Value
Signal	PS, ISR/FSR, PDF	2%–7%, S
SM Vh	0/1-lep norm.	32%
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S
Diboson	0/1-lep norm.	50%
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S

Process	Quantity/source	Value
Top quark	0/1-lep. norm.	float
	single-top-quark contribution	19%
	$t\bar{t} + V$ contribution	50%
	$t\bar{t} + h$ contribution	50%
	0-lep. resolved $\leftrightarrow$ merged	9%–20%
	1-lep. resolved $\leftrightarrow$ merged	18%–20%
	0-lep. SR $\leftrightarrow$ CR	2%–12%
	1-lep. SR $\leftrightarrow$ CR	2%–3%
	PS, ISR/FSR, ME, PDF	S
	DS vs DR scheme (Wt)	S
	p_T^{miss} non-closure (0-lep.)	S
	$p_{T,W}$ non-closure (1-lep.)	S
	Large- R jet p_T non-closure	S

rebinned in the same fashion. Uncertainty templates are then *smoothed* to maintain fit stability while retaining an accurate shape model of a given variation.

For all experimental uncertainties, a **basic smoothing** algorithm is applied that first takes bins with a relative integral statistical error of greater than 5% and combines them into a single bin, then identifies all local extrema of the ratio between the nominal and the statistical variation. The smallest of these local fluctuations are iteratively merged with the neighboring bin until only four local extrema remain. The final step rebins each distribution until a statistical uncertainty smaller than 5% per bin is obtained.

For the theory uncertainties, during the derivation of the variation templates, the distributions are again rebinned such that the statistical uncertainty per bin is less than 5%. Then a **Gaussian kernel smoothing** [184] is applied to the relative variations, where the Gaussian kernels are defined via

$$G(x, x_0) = \exp \left\{ - \left(\frac{x - x_0}{2\sigma} \right)^2 \right\}, \quad (4.6)$$

with a bandwidth parameter σ set to 35 GeV to provide an appropriate level of smoothing while maintaining a good description of the shape variations. Each bin of the original, non-rebinned

relative uncertainty is then filled by the Gaussian weighted sum of adjacent bins using these kernels as

$$C_{\text{smooth}}(x) = \frac{1}{\sum_{x_0} G(x, x_0)} \sum_{x_0} G(x, x_0) \cdot C(x_0), \quad (4.7)$$

where $C_{\text{smooth}}(x)$ is the smoothed relative uncertainty distribution of m_{Wh} and $m_{T,Wh}$, x , from the values of the original bin centers x_0 .

4.4.3.2 Averaging and symmetrization

The statistical model assumes that the $+\sigma$ and $-\sigma$ variations of an uncertainty are centered with respect to the nominal distribution.¹⁶ Some uncertainties are only derived as a one-sided variation, such as the jet energy scale and resolution experimental uncertainties, as well as the matrix element and parton shower matching theory uncertainties. Others create uncertainties where both the up and down variations are on the same side of the nominal in certain phase space regions. These are **averaged** between the two values per bin, and then both types are **symmetrized** in each bin. For those derived with both $+\sigma$ and $-\sigma$ variations where both variations lie on the same side of the nominal, this is done via

$$c_{\pm\sigma}^{\text{sym}} = c_{\text{nominal}} \pm \frac{c_{+\sigma} - c_{-\sigma}}{c_{+\sigma} + c_{-\sigma}}, \quad (4.8)$$

where c represents the per-bin content in the specified distribution. Absolute magnitudes for the individual distributions are reduced by design to reflect the newly constructed negative variation. Those that are directly derived as a single-sided uncertainty have their variation with respect to the nominal simply mirrored to create the negative variation.

4.4.3.3 Pruning

Many uncertainties, particularly experimental uncertainties, yield variations that are extremely small in magnitude and have a negligible impact on the final result. To simplify the final fit parameterization, these uncertainties are removed (“**pruned**”) by the statistical framework if they satisfy one of two criteria:

- Normalization uncertainties are removed if they lead to a variation smaller than 1% with respect to the nominal.

¹⁶On a technical level, failing to symmetrize uncertainties that have both up and down variations on the same side of the nominal can cause extremely undefined behavior from the minimizer – the fit often takes millions of iterations to ultimately converge on a nonsensical value, or fails to converge at all!

- Shape uncertainties are removed if fewer than two bins have a deviation over 1% with respect to the nominal after the overall normalization has been removed.

This procedure is performed after the previously described smoothing, averaging, and symmetrization.

4.5 Statistical Analysis

To obtain the final results of this search for Wh resonances, a statistical analysis must be performed that tests the likelihood of the combined **signal+background hypothesis** with the **background-only hypothesis** for each hypothetical W' mass point. This is conceptually simple but rather complex in implementation – searches for heavy resonances, like this one, are canonical “bump-hunts”. This is historically the simplest way to discover new particles by searching the reconstructed invariant and transverse mass distributions of the Wh system for a localized excess, which would be evident as a “bump” against the expected background. In this search, this is performed using a **binned maximum likelihood fit** to data performed simultaneously in all signal and control regions, implemented with the ROOFIT [185] software package. All information about the vast array of systematic uncertainties detailed in the previous section is ultimately encoded in the final parameterizations of the post-fit m_{Wh} and $m_{T,Wh}$ distributions. This section will give an abbreviated overview of the statistical model and limit-setting procedure.¹⁷

4.5.1 Likelihood Function

Ultimately, the **parameter of interest** in this search is the production cross section of the HVT W' resonance, $\sigma(pp \rightarrow W' \rightarrow Wh)$, which is linearly related to the **signal strength** parameter, μ . After constructing the histogram $\mathbf{n}$ of reconstructed m_{Wh} for a given signal sample, the expectation value [186] for the number of events in each bin n_i can be written as

$$n_{exp} = \mu s_i + b_i, \quad (4.9)$$

where s_i and b_i are the mean number of signal and background events in each bin. Thus, a signal strength of $\mu = 0$ corresponds to the background-only hypothesis and $\mu = 1$ the nominal signal hypothesis. The mean number of entries per bin can also be written as

¹⁷The discussion presented largely follows Ref [186]. It also will use m_{Wh} as a shorthand for both $m_{T,Wh}$ and m_{Wh} for simplicity.

$$s_i = s_{\text{tot}} \int_{\text{bin } i} f_s(m_{Wh}; \boldsymbol{\theta}_s) d(m_{Wh}), \quad (4.10)$$

$$b_i = b_{\text{tot}} \int_{\text{bin } i} f_b(m_{Wh}; \boldsymbol{\theta}_b) d(m_{Wh}), \quad (4.11)$$

where $f_s(m_{Wh}; \boldsymbol{\theta}_s)$ and $f_b(m_{Wh}; \boldsymbol{\theta}_b)$ are the probability density functions (PDFs) of m_{Wh} and s_{tot} while b_{tot} are the total numbers of signal and background events. Here, $\boldsymbol{\theta}$ represents the total ensemble of **nuisance parameters (NPs)** that characterize the shape of the PDFs and affect the signal and background yields. The PDFs, which give the probability to observe a certain number of events given the prediction, follow a Poisson distribution. For an idealized single measurement, this can be written per-bin as

$$\text{Pois}(n_i | n_{\text{exp}}) = \text{Pois}(n_i | \mu s_i + b_i) = \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}. \quad (4.12)$$

Thus, the likelihood function would be the product of all bins

$$\mathcal{L}(\mu) = \prod_i \text{Pois}(n_i | \mu s_i + b_i), \quad (4.13)$$

However, each NP has its own PDF that modifies the overall likelihood. Most uncertainties discussed in the previous section are determined through a subsidiary measurement that gives a “prior” **constraint**. These nuisance parameters, the ensemble of which will be referred to as $\vec{\alpha}$, have their likelihoods constrained by a function $f_k(a_k | \alpha_k)$ of the subsidiary measurements a . For shape uncertainties, each subsidiary measurement defines $\pm 1\sigma$ up and down variations, such that the explicit parameterization of each shape uncertainty modifying the background PDF varying with the NP value α_k becomes

$$h_k(x) = \begin{cases} h_k^0(x) + \alpha_k (h_k^+(x) - h_k^0(x)), & \alpha_k \geq 0, \\ h_k^0(x) - \alpha_k (h_k^-(x) - h_k^0(x)), & \alpha_k < 0 \end{cases}, \quad (4.14)$$

where $h_k^0(x)$ is the nominal distribution and $h_k^\pm(x)$ are the $\pm 1\sigma$ variations of the corresponding uncertainty. The majority of the parameters in the search are **log-normal constrained**, where f_k assumes the logarithm of the NP is Gaussian-distributed around the nominal value with the width equal to its prior. This approach is chosen over a “standard” Gaussian constraint so that expected event yields remain positive for all values of a corresponding NP, even for bins with low event counts. This construction implies that f_k is maximized at the nominal prediction, and a value

$\alpha_k \neq a_k$ results in a “penalty” in the likelihood calculation. When performing the fit, all α_k are allowed to float within their prior constraints.

The NPs corresponding to the main background normalizations are set to “float”, meaning that they are left entirely unconstrained by the likelihood function such that $f_k = 1$, as they have sufficient statistics to be freely determined by the fit without penalty. In contrast, the normalization of more minor backgrounds cannot be sufficiently well-determined from data alone, so their normalizations are taken from simulation and assigned a conservative prior uncertainty that is also log-normal constrained.

Finally, an additional Poisson term must be included to account for the statistical uncertainty, with a per-bin expectation value γ_i , where $\vec{\gamma}$ is an additional set of NPs. Thus, we arrive at our final (perhaps inelegant but functional) likelihood function from the combination of these three factors

$$\mathcal{L}(\mu, \vec{\alpha}, \vec{\gamma}) = \prod_c \prod_b \text{Pois} \left(n_{cb}^{\text{obs}} | n_{cb}^{\text{sig}}(\mu, \vec{\alpha}) + n_{cb}^{\text{bkg}}(\vec{\alpha}) \right) \text{Pois}(n_{cb}^{\text{bkg}}(\vec{\alpha}) | \gamma_{cb}) \prod_k f_k(a_k | \alpha_k), \quad (4.15)$$

where the likelihood is the triple product over all bins b of all regions c and nuisance parameters k . The final fit to data maximizes this likelihood function by varying both the signal strength and NPs, where two types of fit are ultimately performed. An unconditional fit allows the signal strength to vary freely to achieve the absolute maximum possible likelihood, which is used as the signal-plus-background hypothesis. Defining the total ensemble of nuisance parameters as $\vec{\theta} = (\vec{\alpha}, \vec{\gamma})$, the maximized likelihood function can be written as $\mathcal{L}(\hat{\mu}, \hat{\theta})$ and the parameters that globally maximize the likelihood, $\hat{\mu}$ and $\hat{\theta}$, are the **maximum likelihood estimators (MLE)**. In contrast, a conditional fit fixes the signal strength to a specific value. The NPs that maximize the likelihood in the background-only conditional fit, with $\mu = 0$, are the “**profiled**” parameters $\hat{\hat{\theta}}$. Determining the MLEs and obtaining the post-fit parameterization is a large-scale numerical minimization problem performed with the MINUIT minimizer in ROOFIT [185].

4.5.2 Statistical Tests and Limit Setting

To test a hypothesized value of μ , we use the **profile likelihood ratio**

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})}. \quad (4.16)$$

This search assumes that the signal strength should always be non-negative, so a slightly modified definition is used:

$$\lambda(\mu) = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\theta})}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, & \hat{\mu} \geq 0, \\ \frac{\mathcal{L}(\mu, \hat{\theta})}{\mathcal{L}(0, \hat{\theta})}, & \hat{\mu} < 0. \end{cases} \quad (4.17)$$

From the definition of $\lambda(\mu)$, maximum agreement between the data and the hypothetical μ value approaches unity and ranges between 0 and 1. The one-sided negative log-likelihood likelihood test statistic, q_μ , is the most widely used in analyses of this type and can be defined as

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu), & \hat{\mu} \leq 0, \\ 0 & \hat{\mu} > 0. \end{cases} \quad (4.18)$$

We can compute the p -value

$$p_\mu = \int_{\mu, \text{obs}}^{\infty} f(q_\mu | \mu) dq_\mu, \quad (4.19)$$

where $q_{\mu, \text{obs}}$ is the value of the test statistic q_μ observed from the data and $f(q_\mu | \mu)$ is the conditional probability density function of q_μ under the assumption of the signal strength μ . Determining $f(q_\mu | \mu)$ exactly is nontrivial – one can perform “**toy**” **pseudoexperiments** where the parameters are randomly sampled many times from the total model PDF to obtain a distribution of the test statistic, scanning over a full range of values of μ . However, this approach is computationally costly as the fit must be performed many thousands of times for each signal hypothesis. Thus, an **asymptotic approximation** [187] is used, which is valid in the high-statistics limit, which is appropriate for this analysis.¹⁸ The asymptotic approximation assumes that $f(q_\mu | \mu)$ can be approximated with a Gaussian distribution.

To make a statement about a *discovery*, we define the special test statistic q_0 where the p -value will quantify the level of disagreement between the data and the background-only hypothesis.

$$q_0 = \begin{cases} -2 \ln \lambda(0), & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases} \quad (4.20)$$

with a corresponding p -value

$$p_0 = \int_{0, \text{obs}}^{\infty} f(q_0 | \mu) dq_0. \quad (4.21)$$

¹⁸The asymptotic approximation does begin to diverge some in the very highest mass limit, around $m_{W_h} = 4$ TeV and above, where the statistics do begin to become limited. Toy limits were tested for a select number of signal points in this regime and were found to diverge a maximum of 15% at 5TeV from the asymptotic approximation, which was considered acceptable. It would not be practically possible to use toy limits over the full W' mass grid, as just running 8 points locked up the entire shared high-performance computing cluster at Brookhaven National Lab for one week, prompting a deluge of angry emails to the author’s inbox. All following results are shown using the asymptotic limits.

The cumulative distribution can then be approximated by

$$F(q_0|0) = \Phi(\sqrt{q_0}). \quad (4.22)$$

Thus, the p -value is simply

$$p_0 = 1 - F(q_0|0), \quad (4.23)$$

where Φ is the cumulative Gaussian. The discovery significance Z_0 using this approximation is then

$$Z_0 = \Phi^{-1}(1 - p_0) = \sqrt{q_0}. \quad (4.24)$$

The standard for discovery in particle physics is 5σ , corresponding to a p -value of $3 \cdot 10^{-7}$. This search, unfortunately, cannot exclude the background-only hypothesis and does not claim a discovery¹⁹, so this statistic is interesting only for estimating the discovery potential of the analysis model, should a signal be observed. (See Appendix A for more details on these signal injection tests) The more relevant test aims to set a **confidence interval** with a given **confidence level (CL)**. A value of μ can be excluded at a certain CL α if $p_\mu < \alpha$ – the standard is to set $\alpha = 0.05$ such that confidence intervals are reported at $(1 - \alpha)=95\%$. The largest value of μ , and thus the W' production cross section that the analysis cannot reject, is the upper limit at 95% CL. This is the *frequentist* confidence interval, CL_{s+b} . However, in the statistical regimes where modern searches are performed, statistical fluctuations can lead to a result where the number of events fluctuates much below the expected background value. When searching for small signals, the PDFs of the signal-plus-background hypothesis and the background-only hypothesis may be very similar. Thus, the p_μ can become very small regardless of whether the search actually has any real sensitivity. To avoid setting over-aggressive limits in regimes with limited sensitivity, the **modified-frequentist “CL_s” method** [188] is used instead, where we instead define:

$$\text{CL}_s = \frac{p_\mu}{1 - p_0}, \quad (4.25)$$

where the value of μ is then determined by solving for $\text{CL}_s=5\%$. Limits quoted using this method are, by design, more conservative but less susceptible to statistical fluctuation. To estimate the **expected limit**, which is the median exclusion that can be expected given the background-only hypothesis is true, the **“Asimov” dataset** [187] is used, which creates a pseudodataset where

¹⁹Spoilers! If it did, the reader would likely have already known of this work before reading this thesis.

the data is replaced with a PDF is equal to the sum of the background MC simulations and all variations are fixed to their nominal values.²⁰ The median expected upper limits in the asymptotic approximation can be calculated with the MLE of the Asimov dataset, $\hat{\mu}=0$ (by design), after solving for μ by

$$\mu = \hat{\mu} + \sigma\Phi^{-1}(1 - \alpha) = \sigma\Phi^{-1}(1 - \alpha), \quad (4.26)$$

where the uncertainty width can also be determined from the Asimov dataset

$$\sigma_A^2 = \frac{\mu}{q_{\mu,A}}, \quad (4.27)$$

which conveniently encodes the covariance matrix of the estimators for all nuisance parameters that determine the width of the Gaussian in the asymptotic approximation.²¹ It is also important to calculate the overall error bands of the expected limit, as real data will always contain statistical fluctuations and not, in general, be exactly equal to the median. These correspond to the variations of $\pm 1\sigma$ or $\pm 2\sigma$ around $\mu = 0$ such that

$$\mu_{1\sigma} = \sigma(\Phi^{-1}(1 - \alpha) \pm 1), \quad (4.28)$$

$$\mu_{2\sigma} = \sigma(\Phi^{-1}(1 - \alpha) \pm 2). \quad (4.29)$$

As the generated signal templates are generated according to their predicted cross section from the HVT *Model A*, a signal strength of $\mu = 1$ corresponds directly to this value. Thus, the 95% excluded signal strength parameter can be directly mapped to an exclusion on the W' cross section.

4.5.3 Nuisance Parameter Pulls, Constraints, and Rankings

It is essential to assess the quality of the post-fit results to ensure scientific rigor and validity. After all, the statistical analysis will always attempt to maximize the likelihood and spit out a value for the excluded signal strength regardless of how well the background and prior uncertainties are modeled. This is assessed comprehensively using several metrics.

²⁰Named after the short story *Franchise* by Isaac Asimov, about a society where elections are held by selecting the single most representative voter to replace the rest of the electorate.

²¹It may seem strange that σ itself depends on the tested signal strength parameter μ – in practice, it can also be determined numerically from the Asimov fit covariance matrix.

As described above, each individual constrained nuisance parameter also acquires a post-fit value for its MLE, $\hat{\theta}$, used in the final parameterization of the post-fit m_{Wh} distributions. The **pull** P of a given NP is defined by

$$P = \frac{\hat{\theta} - \theta_0}{\sigma_\theta}, \quad (4.30)$$

where σ_0 is the pre-fit value of θ and σ_θ is its pre-fit uncertainty. Each “healthy” NP is generally expected to have pulls roughly within ± 1 in the conditional background-only ($\mu = 0$) fit to data and ideally close to zero. Larger pull values can be indicative of more severe mismodelings, as the likelihood receives a penalty from the constraint when allowing parameters to acquire a very non-nominal value and will only do so if it “corrects” a large discrepancy in some region.²² This can sometimes be intentional; for example, the $p_{T,W}$ mismodeling uncertainties defined in the 1-lepton channel are pulled by design to correct for the W boson p_T mismodeling without directly reweighting the distributions, accepting the associated penalty in the signal strength determination.

The MINUIT minimizer provides a corresponding Hessian matrix in addition to the MLE outputs, which is a square $k \times k$ matrix equal to the second-order partial derivatives of the parameters θ_k . In the large-sample limit, the inverse of this Hessian matrix corresponds to the covariance matrix of the MLEs (including both μ and all NPs) such that

$$\text{Cov}(\theta_i, \theta_j) = \left(-\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} \right)^{-1}. \quad (4.31)$$

The post-fit uncertainties can be extracted from the diagonal elements of the covariance matrix. If the post-fit uncertainty is significantly reduced compared to the pre-fit uncertainty, it is considered “**overconstrained**”. Overconstrained NPs imply that the fit concludes it has a better determination of the value of a given uncertainty than the subsidiary measurement, which can have concerning implications. This is not necessarily an issue if a derived uncertainty is intentionally over-conservative, but otherwise, it should be investigated to ensure the subsidiary measurements are correctly performed and accurately represent the actual variation. For this reason, the smoothing procedure described in the previous section helps to prevent spurious overconstraints from statistical fluctuations in the variation templates.

The off-diagonals of the covariance matrix give the **correlations** between nuisance parameters in the fit. The majority of NPs in the analysis should not be strongly correlated with other NPs, being

²²A skeptical reader may wonder about the validity of assuming a priori that the background-only hypothesis will be the true distribution and any deviations necessarily represent mismodeling to be judiciously corrected rather than some indication of new physics. This is a matter of philosophical debate – still, one can argue it can reasonably be assumed that new physics will not happen to be shaped in an invariant mass distribution exactly like, e.g., a JES uncertainty.

largely statistically independent effects. These correlations are examined through a correlation matrix where all NPs with a single correlation above a certain threshold are shown. While correlations may not necessarily be pathological (To choose a random example, one might reasonably expect that the single top normalization NP may be correlated to some degree with the single top DS/DR scheme uncertainty), unexpected large values are noted for further investigation to determine if a real correlated effect between two parameters is not being appropriately well-described.

Finally, the last fit quality metrics are the **NP rankings**. The strength of the impact of a given individual parameter on the determined value of μ can be determined by fixing the NP to both the pre-fit and post-fit values, adding and subtracting their respective pre-fit or post-fit uncertainties. Then, the fit procedure is redone, and the impact $\Delta\mu$ is calculated as the difference between the original fitted μ and the new value for both the pre-fit case and the post-fit case. This is done sequentially for each NP, such that a ranking is produced detailing the NPs with the largest impact in order, which helps determine which uncertainty is primarily driving the fit model.

The fit quality metrics for this search will be shown and discussed alongside the final post-fit results in the following section.

4.6 Results

Having saved the best for last, the combined binned maximum likelihood fit was performed simultaneously in all signal and control regions of the 0-lepton and 1-lepton channels, with the signal cross section and the global normalizations of the top-quark, $W + hf$, $W + hl$, $Z + hf$, and $Z + hl$ backgrounds as free parameters. The final binning in each category and region was chosen to maximize search sensitivity while minimizing statistical fluctuations, such that a bin has a width of at least two times the expected mass resolution at a given W' mass and is not allowed to have a statistical uncertainty greater than 70%. The obtained post-fit normalization scale factors are described in Table 4.7, where the observed deviations from unity are consistent with other ATLAS analyses on the same dataset that target a similar phase space [169, 170]. The m_{Wh} and $m_{T,Wh}$ distributions in the signal regions after a fit of the background templates to data are shown in Figures 4.11 and 4.12 for the 1-lepton and 0-lepton channels, respectively. The control regions are shown in Figure 4.13. The post-fit distributions show excellent agreement between data and simulation in all regions, and a summary of the total event yields in each category is shown in Figure 4.14.

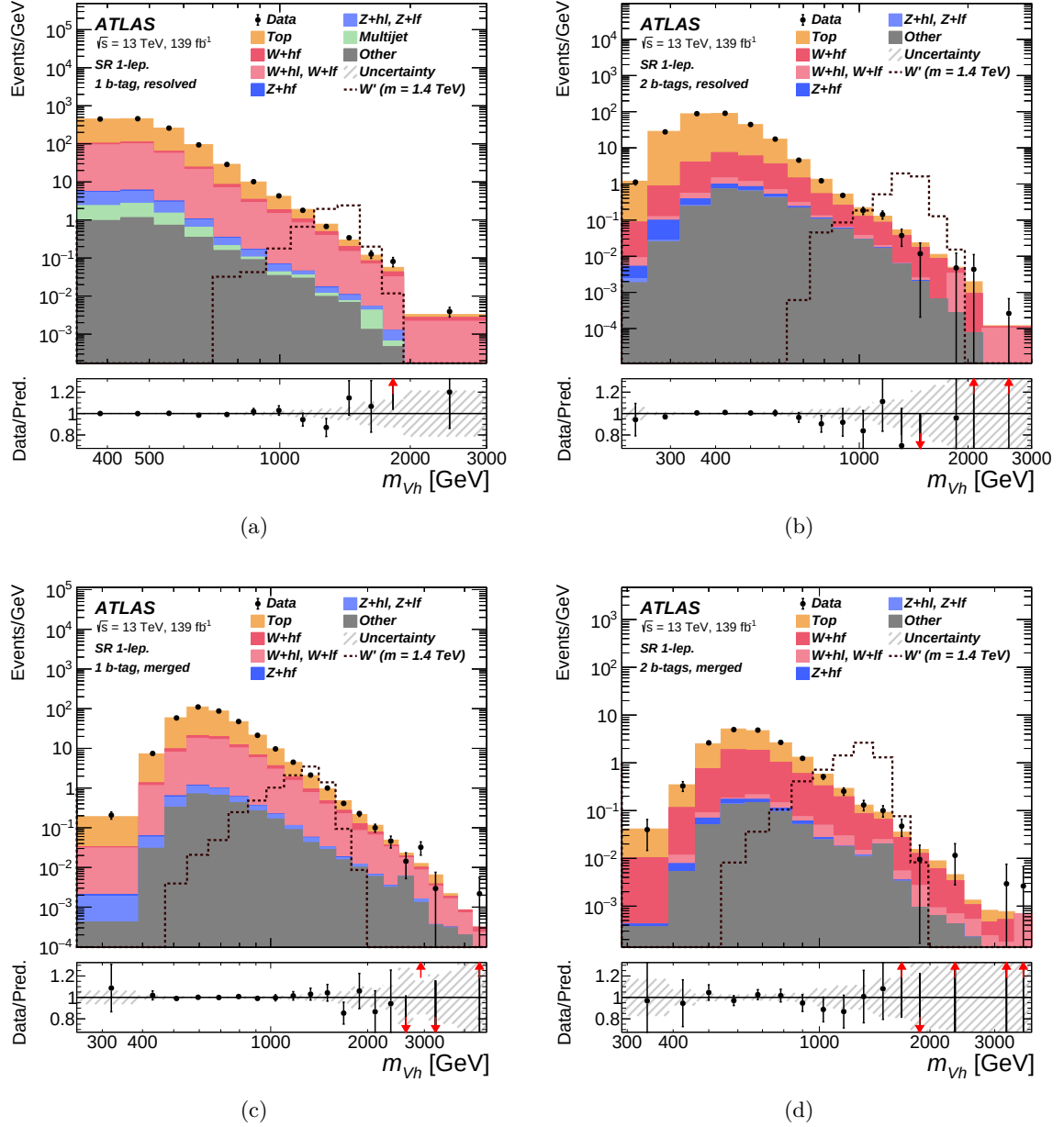


Figure 4.11. Event distributions of $m_{W'h}$ for the 1-lepton SR in the resolved (a, b) and merged (c, d) categories of the W' fit. The quantity on the vertical axis is the number of data events divided by the bin width in GeV. In each plot, the last bin contains the overflow. The background prediction is shown after a background-only maximum-likelihood fit to the data. The lower panels show the ratio of the observed to the estimated SM background. The background uncertainty band shows the post-fit statistical and systematic uncertainties added in quadrature. The signal for the benchmark HVT *Model A* with $m_{W'} = 1.4$ TeV is shown as a dashed line and normalized to a cross section times branching fraction of 0.1 pb.

Table 4.7. Post-fit normalization scale factors (NF) and their uncertainties obtained from a combined background-only fit of all signal and control regions of the 0- and 1-lepton channels. Numbers are presented for the background components that are allowed to float in the likelihood fit.

Background	0-lepton	1-lepton
Combined W' fit of 0- and 1-lepton channels		
top	0.92 ± 0.04	0.93 ± 0.04
$Z+hf$	1.30 ± 0.16	–
$Z+hl$	1.30 ± 0.16	–
$W+hf$	–	1.35 ± 0.18
$W+hl$	–	1.34 ± 0.12

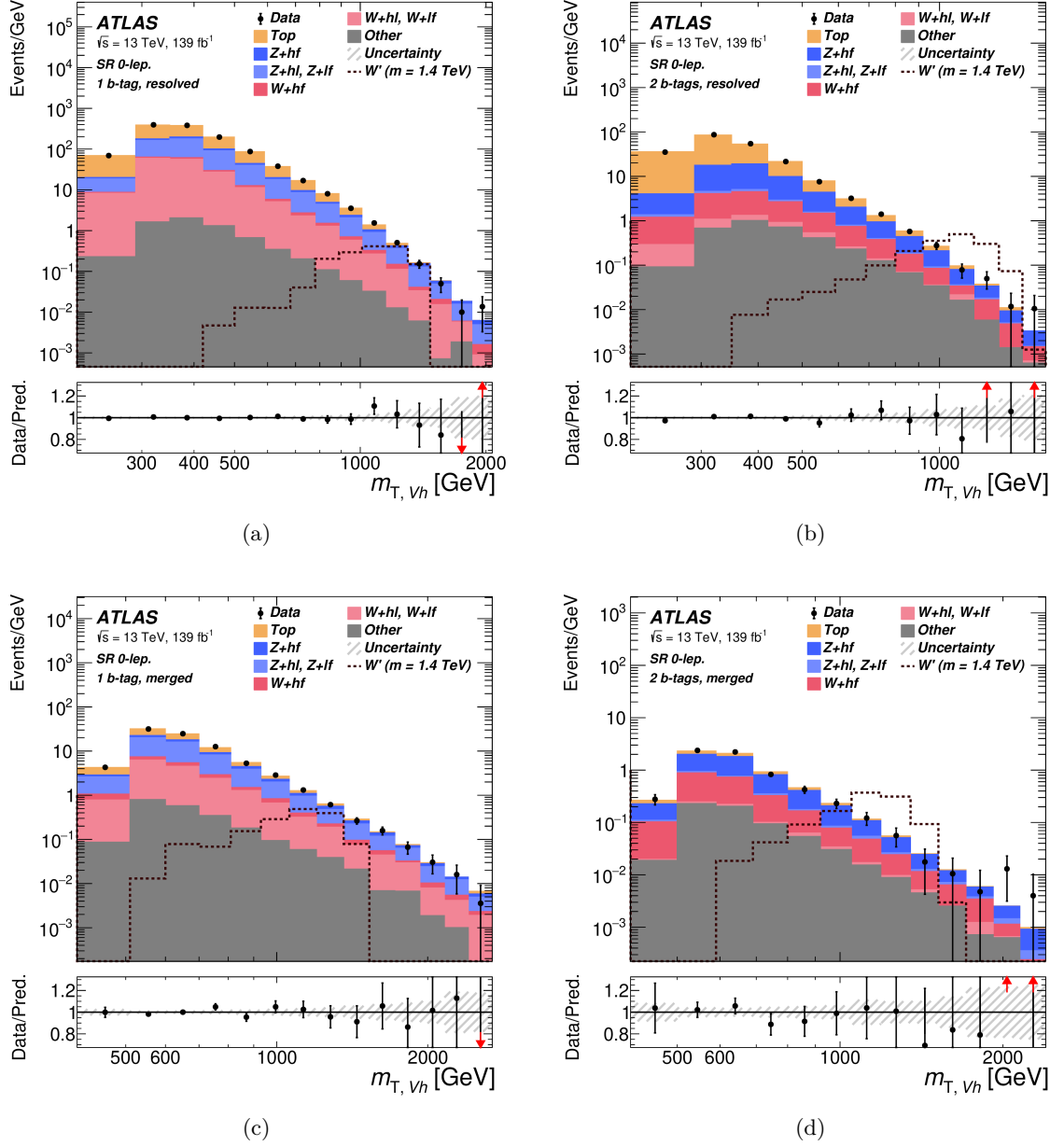


Figure 4.12. Event distributions of $m_{T,Wh}$ for the 0-lepton SR in the resolved (a, b) and merged (c, d) categories of the W' fit. The quantity on the vertical axis is the number of data events divided by the bin width in GeV. In each plot, the last bin contains the overflow. The background prediction is shown after a background-only maximum-likelihood fit to the data. The lower panels show the ratio of the observed to the estimated SM background. The background uncertainty band shows the post-fit statistical and systematic uncertainties added in quadrature. The signal for the benchmark HVT *Model A* with $m_{W'} = 1.4$ TeV is shown as a dashed line and normalized to a cross section times branching fraction of 0.1 pb.

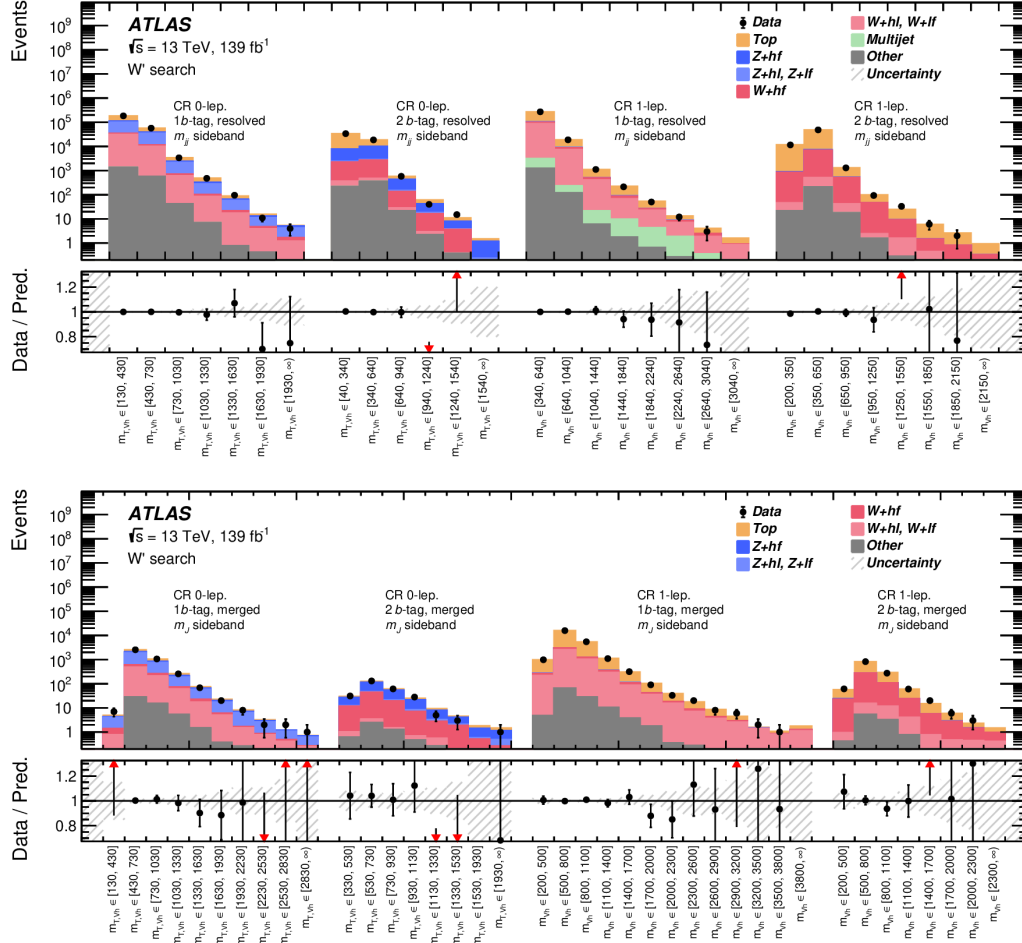


Figure 4.13. Distributions of the m_{Wh} and $m_{T,Wh}$ observables in the control regions of the resolved (a) and merged (b) event categories. The background prediction is shown after a background-only maximum-likelihood fit to the data. The lower panels show the ratio of the observed to the estimated SM background. The background uncertainty band shows the post-fit statistical and systematic uncertainties added in quadrature.

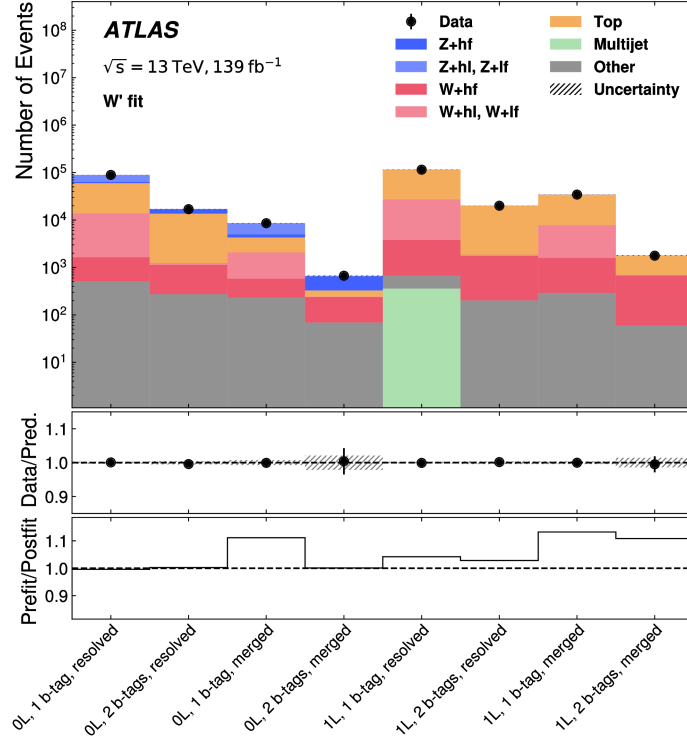


Figure 4.14. Comparisons of the observed data and the expected background event yields in the signal regions of the W' fit. The background predictions are obtained through a background-only simultaneous fit to the signal and control regions. The middle panel shows the ratios of the observed data to the post-fit background predictions. The bottom panel shows the ratios of the post-fit and pre-fit background predictions. The hatched bands represent uncertainties in the total background predictions, combining statistical and systematic contributions.

The fit quality was assessed by evaluating nuisance parameter pulls, rankings, and correlations shown in Figures 4.15 and 4.16, 4.17 and 4.18, and 4.19, respectively. The pull comparisons show that the nuisance parameters behave healthily in the fit, with the vast majority of values between -1 and 1. The few NPs with pulls beyond these values, `MPTRweight2tagZ1`, `Whf_resolvedPTV2tag`, `top_resolvedpTV1tag`, and `top_resolvedpTV2tag`, represent the bespoke shape correction uncertainties where a larger pull value shows the fit is successfully using these degrees of freedom to improve the m_{Wh} distributions, as intended. These NPs are also slightly constrained, indicating that the original estimate of the mismodeling uncertainties was conservative. The additional constraints seen, such as on `norm_top_L1` and the other global normalization parameters, indicate that the fit has enough information to strongly constrain the overall background normalizations, also by design. No unexpected large correlations between NPs are seen, with the few larger values relating to relationships in W/Z +jets flavor component normalizations and between the shape correction

uncertainties for different major background components in a single region. The NP ranking order shows that, as can be expected, the shape correction uncertainties have some of the largest impacts on the overall signal strength. Beyond this, the dominant uncertainties at low W' masses are due to the modeling of the parton shower in $t\bar{t}$ background events, and those at high W' masses relate to the matrix element and parton shower matching for the various W +jets contributions as well as scale uncertainties on large- R jets.²³ The search is statistically limited, rather than being limited by systematic uncertainties, beyond 600 GeV.

²³The $t\bar{t}$ parton shower and matrix element modeling uncertainty, `ttbarMatchModShapeOnly`, impacts the fits strongly for a specific reason that is not necessarily related to some fundamental issue with the parton shower matching. There is a slight pre-fit overestimation of background versus data in the lowest mass bins of the 1-lepton 1-tag resolved categories, and `ttbarMatchModShapeOnly` is one of the few available uncertainties whose shape variation is inverted in this lowest-mass regime; thus the fit chooses this to correct the discrepancy. It was decided to exclude the two lowest mass bins of this region in the final fit, decreasing the overall sensitivity slightly, as they were not appropriately well-modeled.

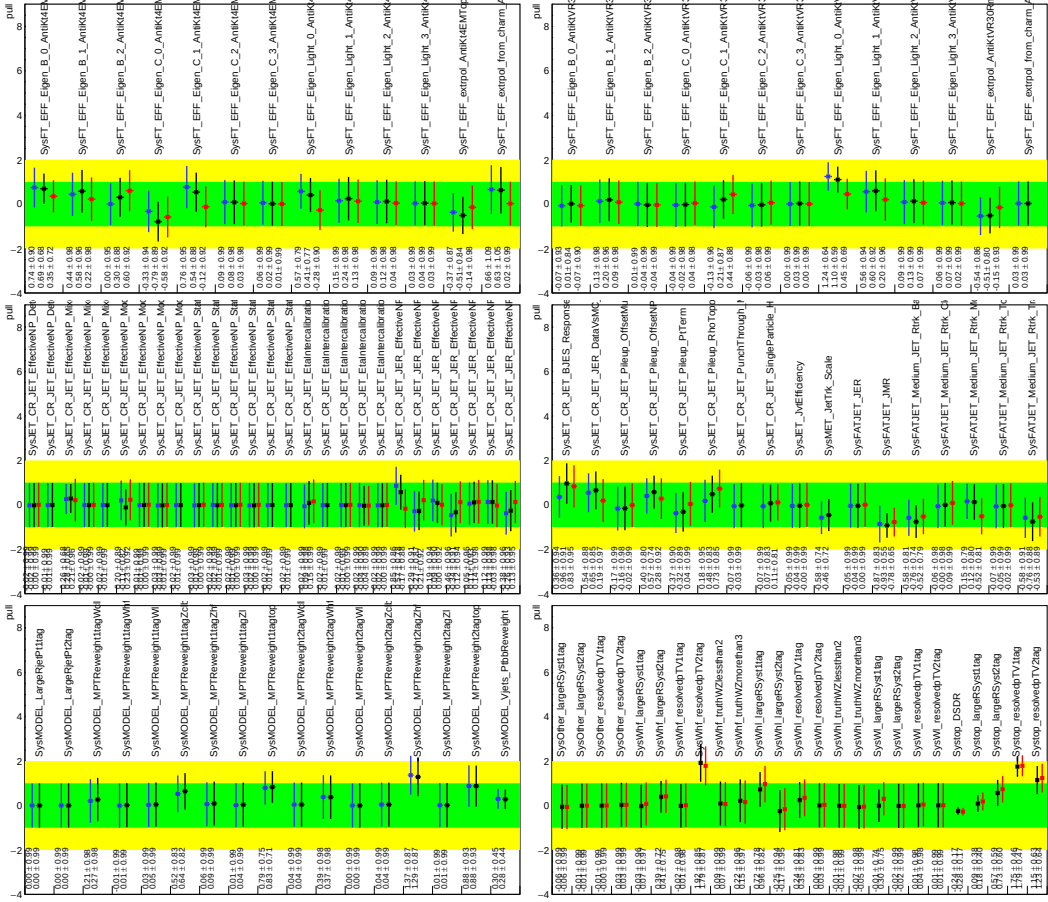


Figure 4.15. Nuisance parameter pulls after a conditional $\mu = 0$ fit to data in the signal and control regions of the W' combination. The blue, black, and red markers indicate the pull in the 0-lepton, combined, and 1-lepton fits, respectively. The first number shows the pull value, while the number after $\pm$ shows the ratio of pre-fit to post-fit constraint.

The 95% CL upper limits on the production cross section for $W' \rightarrow Wh$ as a function of the resonance mass are shown in Figure 4.20. The limits range from 1.3 pb for $m_{W'} = 400$ GeV to 0.4 fb for $m_{W'} = 5$ TeV, and exclude W' masses below 2.95 TeV for HVT benchmark *Model A* with coupling constant $g_V = 1$ [12]. For *Model B* with coupling constant $g_V = 3$, W' masses below 3.3 TeV are excluded. The limits from non-combined fits to the individual channels are also shown, where it can be seen that the combined limits are dominated by the 1-lepton channel over the full range of tested mass hypotheses. The observed limits agree with the Standard Model expectations across all hypotheses, with mild excesses observed at $m_{W'} = 400$ GeV (about 2.0 standard deviations, p_0 -value of 0.02) and 3 TeV (about 1.7 standard deviations, p_0 -value of 0.05). Local p_0 values for the background-only fit are shown for all resonance masses in Figure 4.21. The upper limits are also used to set 95% CL exclusion contours in the HVT parameter plane $\{g_F, g_H\}$ ²⁴ [189] Exclusion contours are shown in Figure 4.22 for W' masses of 2,3, and 4 TeV.

²⁴The coupling constants g_H and g_F are related to those in Ref [12] as follows: the Higgs boson coupling is $g_H = g_V c_H$ and the universal fermion coupling is $g_F = g^2 c_F / g_V$, where g is the SM $SU(2)_L$ gauge coupling.

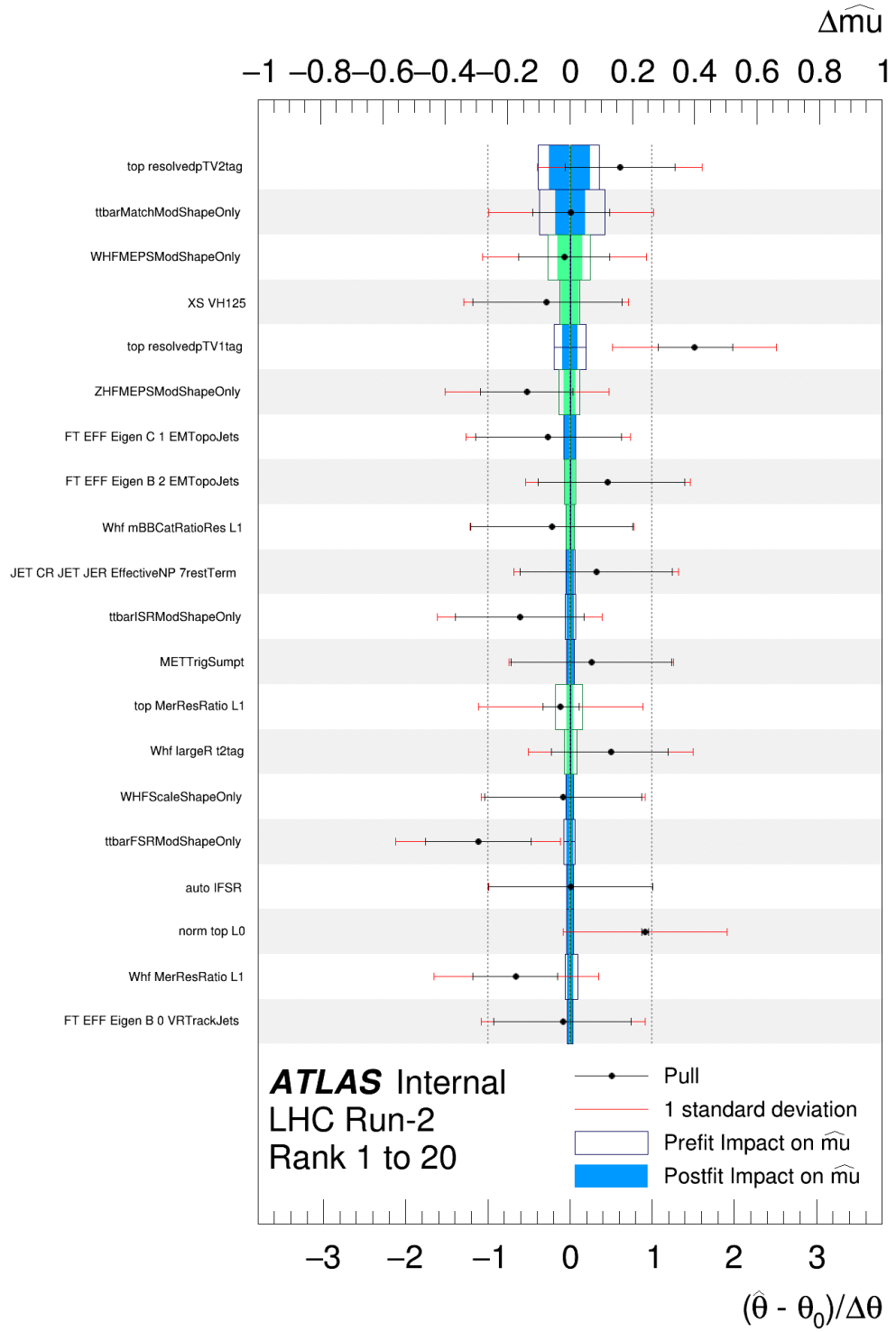


Figure 4.17. Ranking plots for the SR fit to observed data for $m_{W'} = 700$ GeV. Positive correlations between nuisance parameters and the parameter of interest are indicated by the blue bars, while negative ones are shown in green.

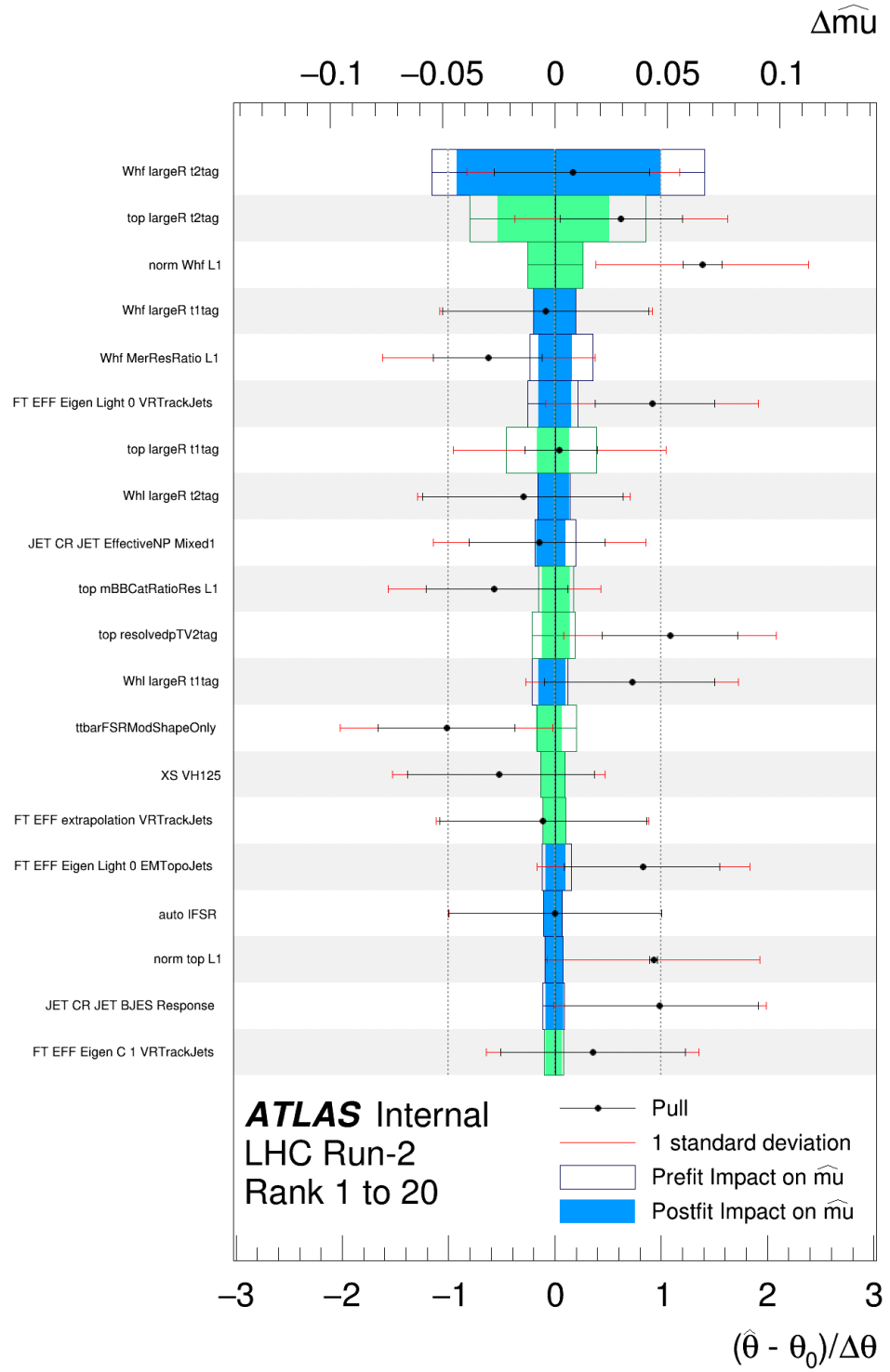
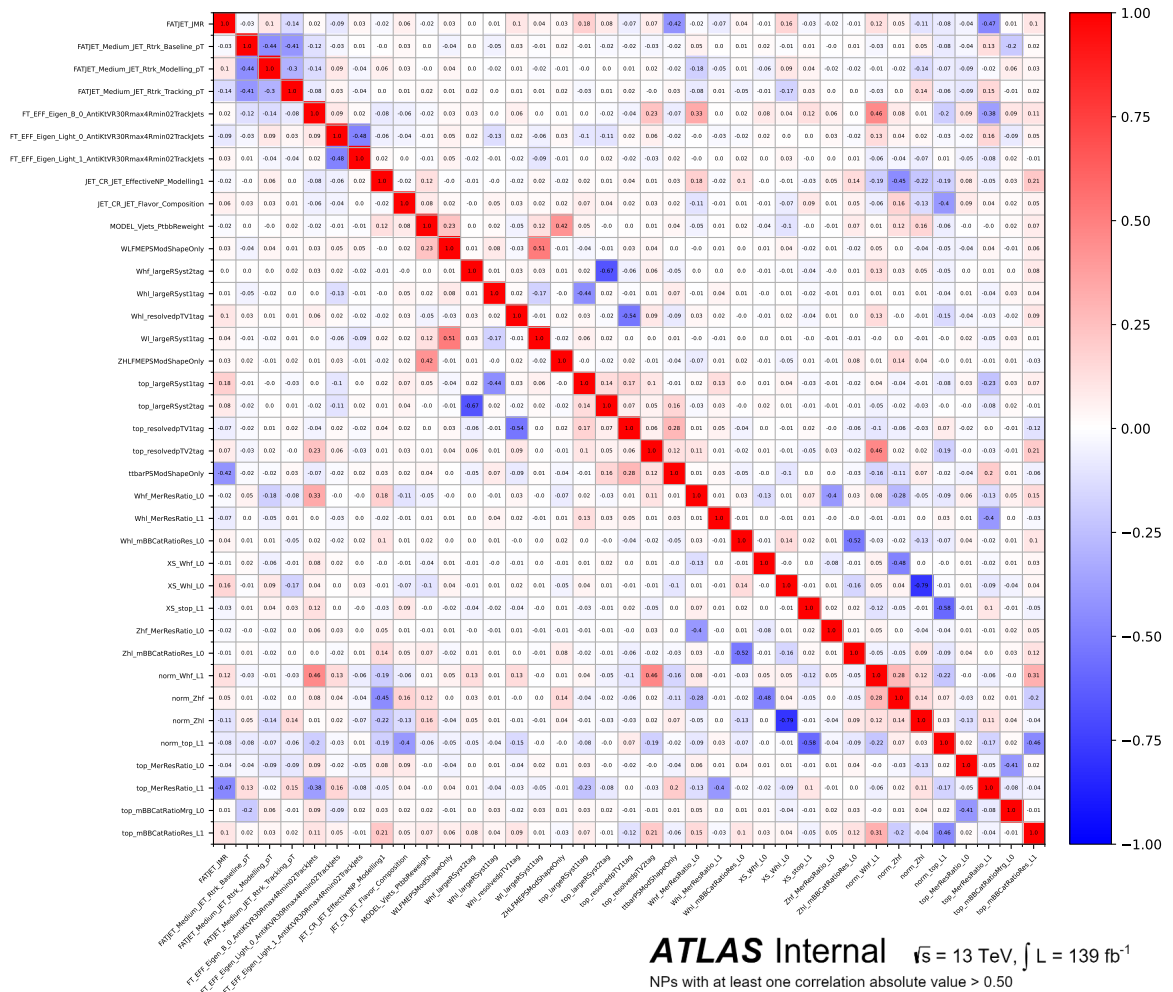


Figure 4.18. Ranking plots for the SR fit to observed data for $m_{W'} = 2000$ GeV. Positive correlations between nuisance parameters and the parameter of interest are indicated by the blue bars, while negative ones are shown in green.



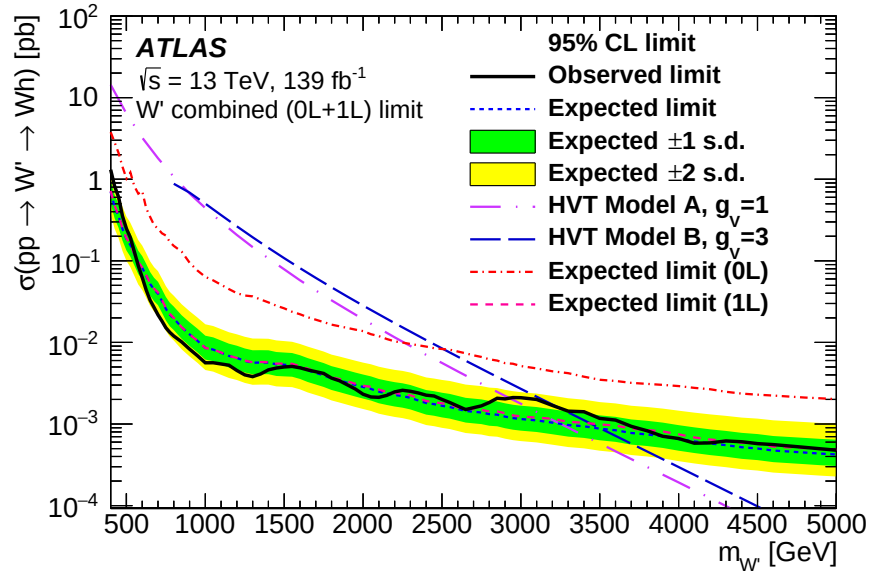


Figure 4.20. Upper limits at the 95% CL on the product of the cross section for $pp \rightarrow W'$ and the branching fraction to Wh from the combination of the 0-lepton and 1-lepton channels. Limits from the 0- and 1-lepton channels individually are also shown for comparison.

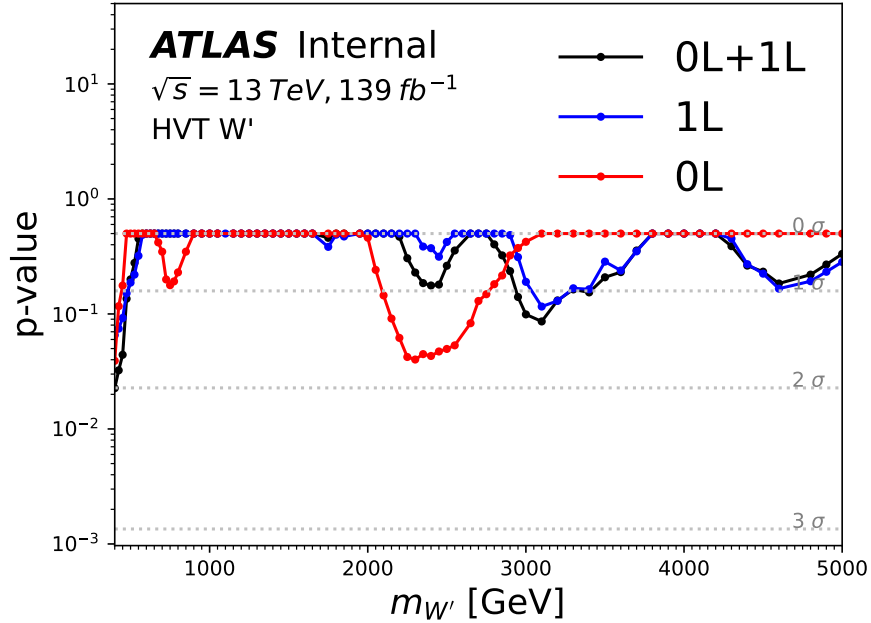


Figure 4.21. Observed local p_0 values as a function of the tested W' mass point. The red curve indicates the standalone 0-lepton fit, the blue curve the standalone 1-lepton fit, and the black curve the 0- and 1-lepton combined fit.

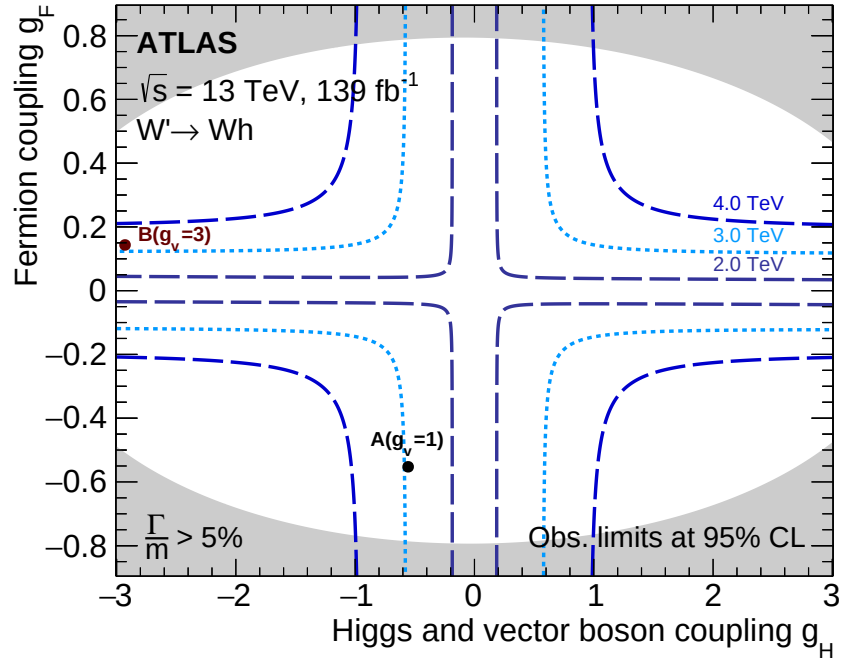


Figure 4.22. Observed limits on the HVT model at 95% CL in the g_F versus g_H plane for W' masses of 2, 3, and 4 TeV. The circles indicate the coupling values for Models A and B, and the grey region corresponds to the area of phase space where the decay width of the resonance is no longer negligible and the signal m_{Wh} shape is no longer expected to be dominated by the experimental resolution.

4.7 Conclusion and Outlook

We have presented a search for HVT W' bosons in the $\ell\nu b\bar{b}$ final state using 139 fb⁻¹ of 13 TeV pp collision data collected with the ATLAS detector. No significant excess of events was observed above SM predictions, and upper limits were set on the production cross sections of the W' resonance for hypothesized masses of 400 GeV to 5 TeV.

This result improves significantly on the previous 36.1 fb⁻¹ ATLAS result for the Wh semileptonic final state [29], ranging from around a 50% improvement at a W' mass of 400 GeV to about 500% at 5 TeV. Much of the improvement across the board is due to the larger dataset; however, the significant improvements at high resonance masses can be attributed to the inclusion of TCC jets over “standard” Calo-only large-R jets as well as refined b -tagging techniques using VR track jets. Figure 4.23 shows the upper limits overlaid with those from the partial Run 2 result for comparison. It also shows the partial Run 2 limits scaled to the total Run 2 luminosity, highlighting the big methodological improvements at higher masses.

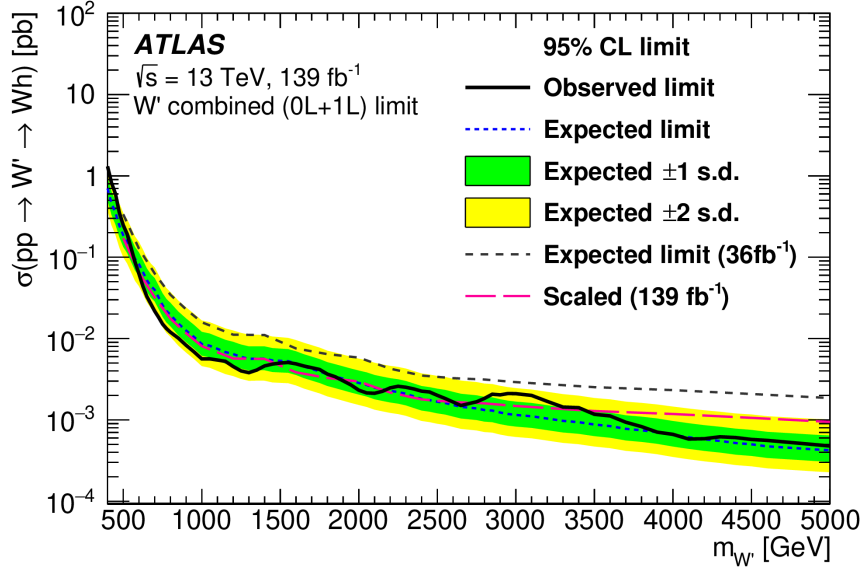


Figure 4.23. Upper limits at the 95% CL on the product of the cross section for $pp \rightarrow W'$ and the branching fraction to Wh from the combination of the 0-lepton and 1-lepton channels. Limits from the 36.1 fb⁻¹ analysis are shown for comparison, both in their original form and scaled to the full Run-2 luminosity, for comparison.

At the time of first publication, these limits were the strongest to date for an HVT W' released by ATLAS. More recently, in April 2024, another ATLAS result was released that includes this result in a broader combination that sets constraints on the broader HVT model using eleven ATLAS HVT

searches, including the fully hadronic VV and Vh decay modes [190]. This combination takes the W' and Z' masses to be degenerate, which is well-motivated by experimental constraints on the weak mixing angle (see Section 1.7), and excludes V' masses below 5.8 TeV for HVT *Model A* and 4.4 TeV for *Model B*.

The CMS experiment published a comparable result in 2022, which sets slightly stronger limits relative to our search in the 1 TeV to 4 TeV regime that it covers [131]. This slightly narrower range might contribute some to the difference, but the CMS search is also unique in that it performs a two-dimensional fit in the $(m_{V,\text{lep}}, m_{\text{jet}})$ plane, combining both the VV and Vh HVT decay modes, which may be a more powerful search technique than the separate one-dimensional searches performed in ATLAS. Motivated by this result, a similar strategy is now being investigated in ATLAS in a “legacy” full Run-2 analysis, following the search presented in this work.²⁵ As the Wh semileptonic search is still statistically limited, rather than limited by systematic uncertainties, in the intermediate- and high-mass regimes, it will likely also benefit from an improved analysis, including Run 3 data in the near future – and, of course, Run 4 and beyond in the more distant future.

²⁵The author is excited to see the developments, but has personally moved on from resonance searches and is not directly involved aside from occasional code archaeology requests.

This result was the culmination of six years of nonstop efforts from a very small team of dedicated collaborators scattered around the world. Countless days were whittled away into the quiet hours of the evening, most illuminated only by the faint glow of a terminal window. The years were punctuated with both triumphs and heartbreaks in spades which were a wellspring of professional, as well as personal, growth. It was a long and challenging journey, saying our goodbyes and well-wishes to many wonderful and irreplaceable colleagues (and friends) along the way. After all of it, we released a result to be proud of – and there's beauty to be found within if you know where to look.

Thanks for everything.

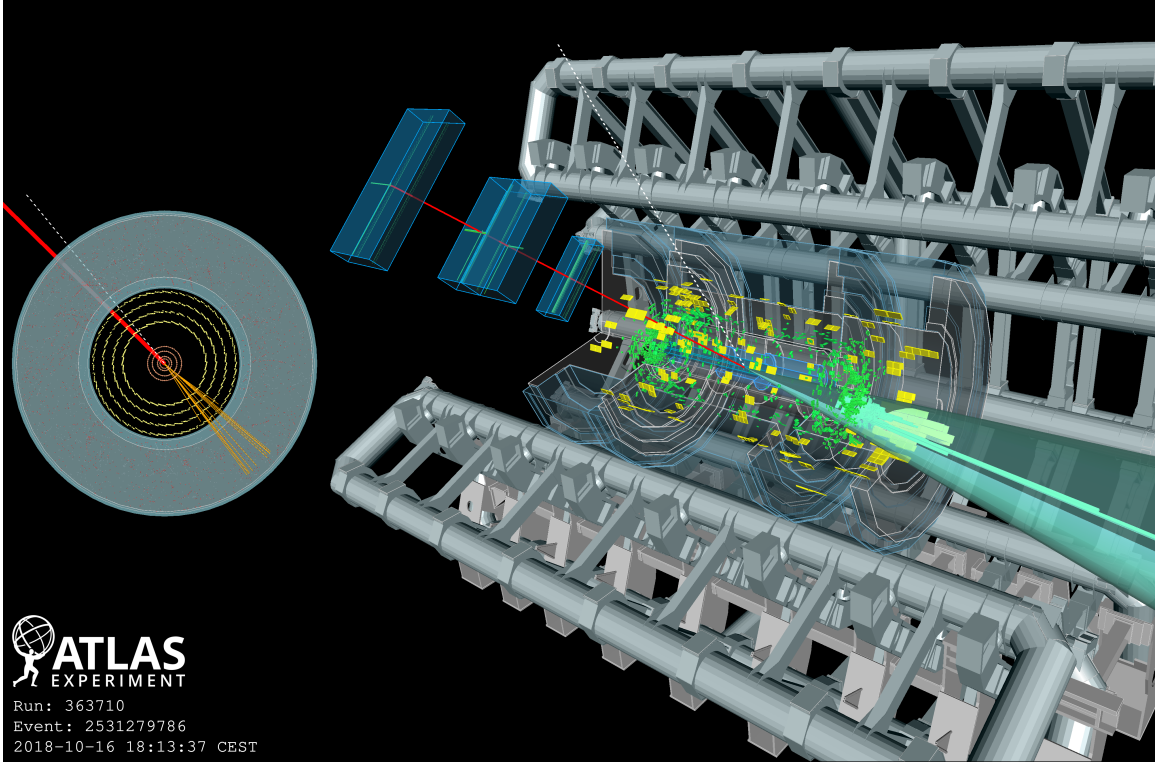


Figure 4.24. Event display of an $X \rightarrow Wh \rightarrow \mu\nu b\bar{b}$ candidate in the merged 1 b-tag signal region. The W boson candidate is comprised of a muon and missing transverse energy. The muon is indicated by the red solid line and has a transverse momentum of $p_T = 632$ GeV. The missing transverse energy has a value of 362 GeV and is indicated by the dashed white line. The Higgs-boson candidate is comprised of a large-R jet which contains one b-tagged VR track-jet. The large-R jet is indicated by the turquoise cone and has a transverse momentum of $p_T = 1.09$ TeV. The yellow blocks correspond to energy depositions in the hadronic calorimeter, and the blue boxes correspond to muon chambers with hits from the muon. The invariant mass of the Wh candidate is $m_{Wh} = 3.7$ TeV. The event was recorded in 2018 (Run 363710, Event 2531279786).

APPENDIX A

MOMENT MORPHING PROCEDURE

An analysis with excellent signal mass resolution, though a sign of great performance and overall success in developing the analysis strategy, poses some unique practical challenges when targeting an expansive grid of resonant signal points. Notably, as the standard for discovery in high-energy physics is a 5σ significance, rather than 3σ as in many other fields, any decrease in discovery potential for a real signal with a mass between the simulated points could have severe implications on the ability to make a claim, were a “real” excess to be seen. The simplest solution to this problem would be to simulate additional mass points until the spacing no longer impacts the discovery potential – however, as computational resources for the broader ATLAS collaboration are stretched to their limit, it is not practically feasible to directly simulate enough W' signal to reach mass-resolution spacing. This appendix details an investigation into, and ultimately an innovative solution to, this phenomenon, focusing on the 1L channel.

A.1 Mass Resolutions

The **mass resolution** of a reconstructed signal is determined by fitting the distributions of the mass residuals given by

$$\frac{m_{recon} - m_{truth}}{m_{truth}}, \tag{A.1}$$

to a Bukin function [191], where m_{recon} is the reconstructed signal mass in an event and m_{truth} is the “truth” generated signal mass for the same event. The Bukin function is a convolution of Gaussian and exponential distributions designed to fit asymmetric peaks, where the fitting parameter σ is equal to the peak width at half-maximum of the distribution and taken to be the mass resolution for a given signal point. Residual distributions for several W' mass points for 1L, 1 b -tag merged and resolved categories are shown in Figure A.1, overlaid with the Bukin fit.

A summary of the mass resolutions obtained from these residuals can be seen in Figure A.2. As expected, the resolved category has a better mass resolution in the lower mass regime and loses power in the higher mass regime, and vice versa for the merged category. The resolutions hold at around

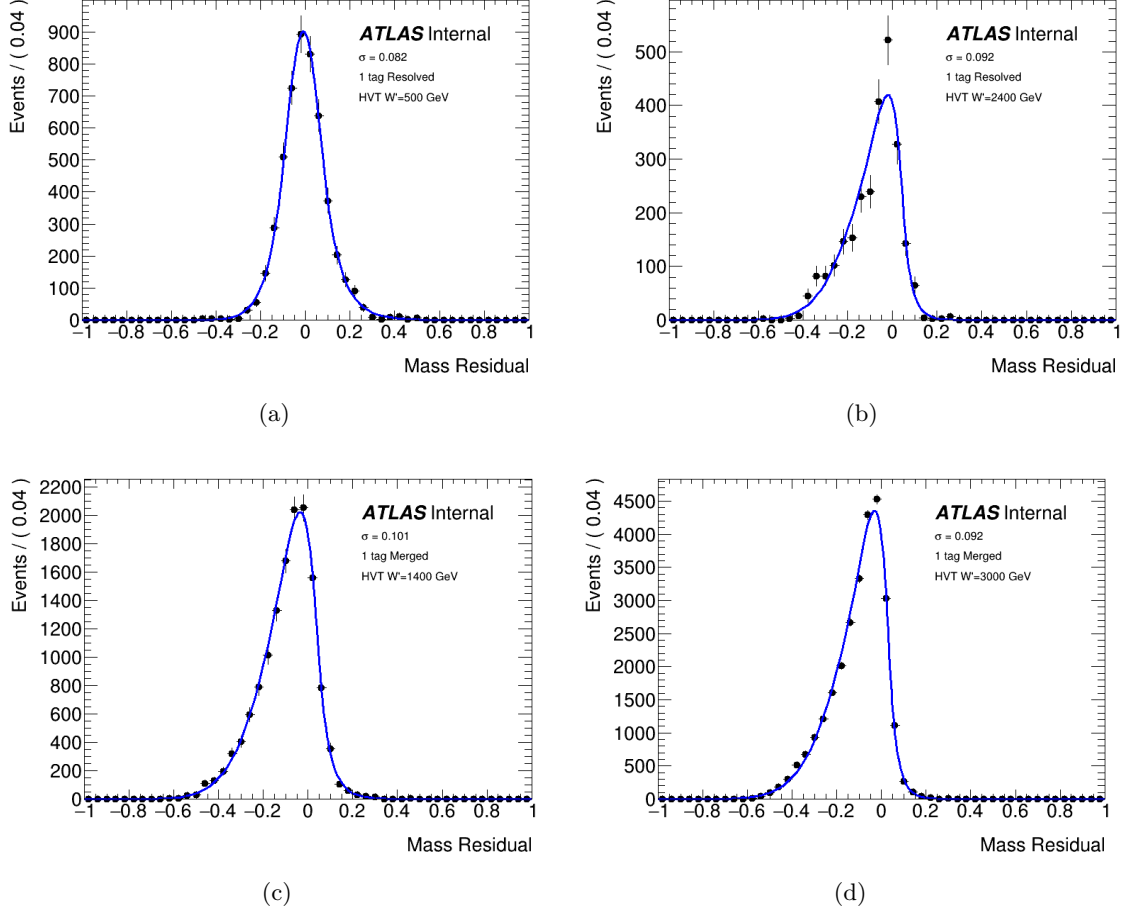


Figure A.1. Mass residuals for $m_{WH} = 500$ GeV and $m_{WH} = 2400$ GeV in the resolved category (a,b) and for $m_{WH} = 1400$ GeV and $m_{WH} = 3000$ GeV for the merged category (c,d). Overlaid is the fitted Bukin distribution, and the fitted width of the distribution, σ , is labeled.

10% of the resonant signal mass, from around 7% at the smallest and 12% at the largest. To account for the anticipated mass resolutions a priori, the simulated grid points were produced in increments of 100 GeV from $m_{W'} = 400$ -1000 GeV, in increments of 200 GeV from $m_{W'} = 1000$ -3000 GeV, and in increments of 500 GeV from $m_{W'} = 3000$ -5000 GeV. A potential issue can immediately be spotted. If there were to be a 3200 GeV W' , this is at the edge of the mass resolution of either neighboring simulated point. Thus, the corresponding hypothesis tests will have much-reduced sensitivity to the “intermediate” signal. This is confirmed through signal injection tests, which are detailed in the following section.

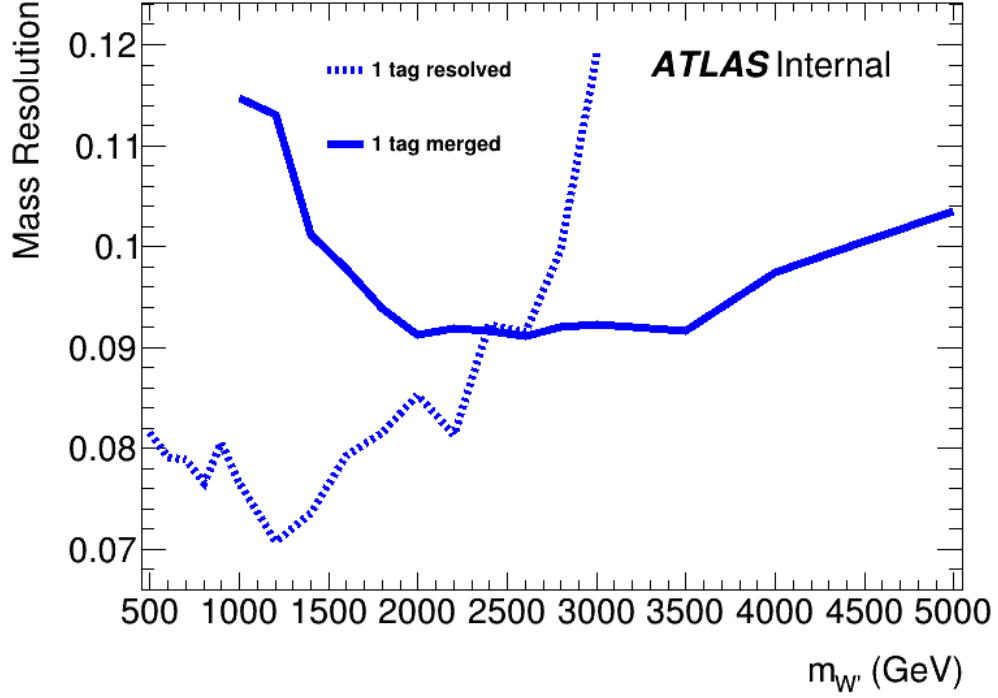


Figure A.2. Mass resolutions per W' signal point for the merged and resolved categories in the 1 b -tag, 1L channel. The resolved category has a better mass resolution in the lower mass regime and loses resolution in the higher mass regime, and vice versa for the merged category.

A.2 Signal Injection Tests

To assess the discovery potential of the analysis and validate the overall fitting procedure, **signal injection tests** were performed. This is accomplished through building several “pseudodatasets” from the sum of MC backgrounds and different signal mass hypotheses, scaled to three times the expected upper limit on the production cross section from the background-only fit. (in this case, $m_W = 700, 1400, \text{ and } 2200$ GeV). These are then run through the same binned maximum likelihood fitting procedure using the RESONANCEFINDER as used in the standard analysis, but with the pseudodataset instead of the real dataset. The local p -values are then calculated for the fit to each signal mass hypothesis as they are in the standard analysis, which demonstrates the ability of the search to claim a discovery were a localized excess to exist in the genuine dataset. As shown in Figure A.3, these tests show that while strong discovery potential exists if the signal is close to the simulated mass point, it rapidly drops to below a 5σ excess for each neighboring mass point. This effect is particularly severe in the low mass regime – injection of a 700 GeV signal only results in a

2σ local p -value when testing the 600 GeV hypothesis. This strongly suggested the need to enhance the signal grid without directly producing new simulations.

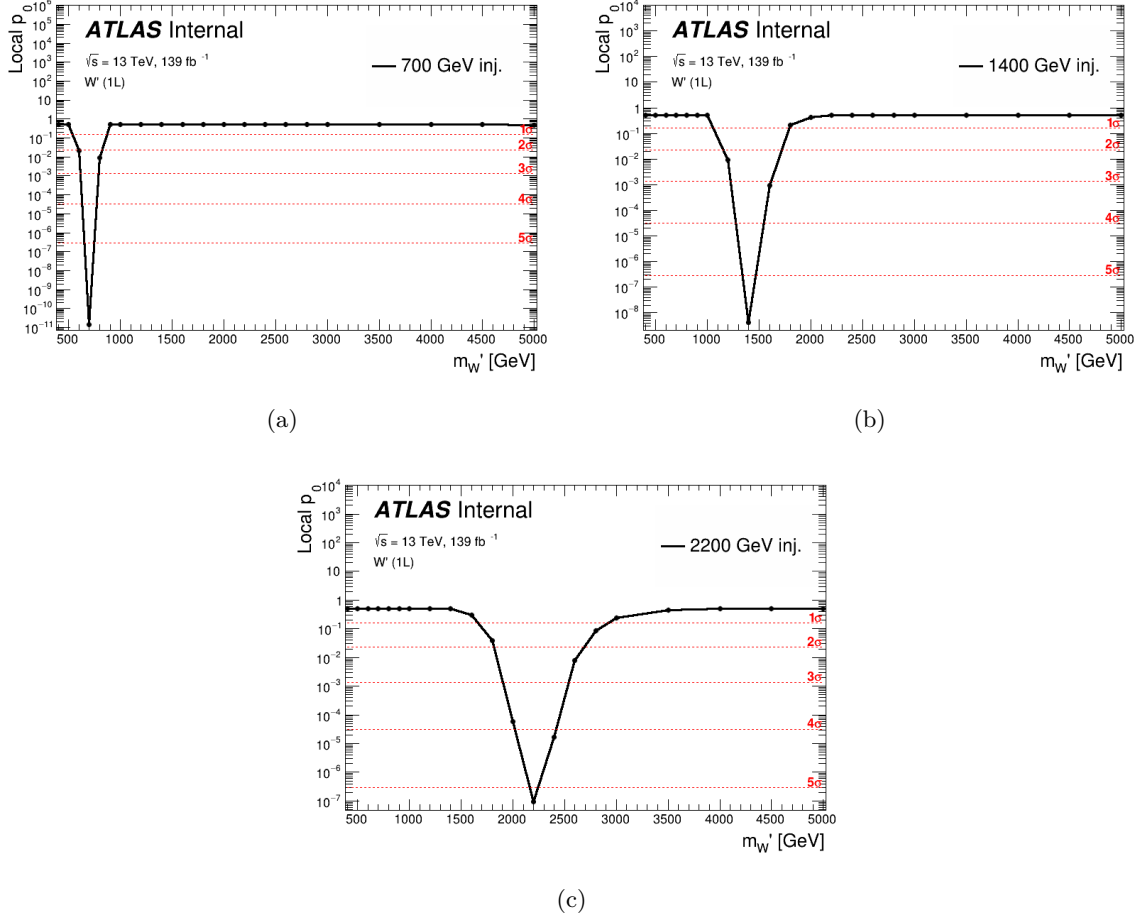


Figure A.3. Distributions of the local p -value as a function of the W' mass. The p -value distributions correspond to a fit to Asimov data injected with a hypothetical signal with a mass of (a) $m = 700$ GeV, (b) 1400 GeV, and (c) 2200 GeV.

A.3 Moment Morphing

It has been established that injection studies demonstrate that the expected local p_0 of an excess depends strongly on how close the nearest simulated signal mass point is, and “missing” the excess by only one mass resolution can decrease a $> 5\sigma$ excess to a 3σ one, severely limiting the discovery potential. Yet, it has also been established that computationally, additional signal generation is prohibitively expensive and, more importantly, extremely inefficient. The solution to this problem exists in the form of a **nonlinear moment morphing technique** based off of

Ref [192] and implemented through the ROOMOMENTMORPH class of ROOFIT.¹ This technique constructs a morphing probability density function (PDF) that can transform an observable $\vec{x}$ based on a single morphing parameter m – here, the fit observable m_{VH} and the hypothesized resonance mass, respectively.

A.3.1 Morphing Technique Summary

The core concept behind the moment morphing technique is that there exists a *morphing PDF*, $f(\vec{x}|m)$, that can transform the observables $\vec{x}$ based on a single morphing parameter m . While the method can transform several observables $\vec{x} = (x_1, x_2, x_3, \dots)$ at once, this search is particularly well suited to the technique as it can be reduced to the minimal version, as a canonical “bump-hunt” only requires a direct fit on a single kinematic observable – again, here the invariant mass distribution of the Wh system, m_{Wh} , with the hypothesized resonance mass $m_{W'}$ as the morphing parameter. It is not possible to obtain an exact form for $f(\vec{x}|m)$ directly,² instead f has been sampled at regular values of m with each $f(\vec{x}|m_i)$ representing a known input template shape for a given m_i . The aim is to thus construct a parametric approximation of $f(\vec{x}|m)$ that is smooth and continuous. Given n known values of m , this is accomplished by expanding in a Taylor series up to order $(n - 1)$ around a reference point m_0 . One can write the vector equation

$$f(\vec{x}|m_i) \approx \sum_{j=0}^{n-1} (m_i - m_0)^j f'_j(\vec{x}|m_0) = \sum_{j=0}^{n-1} M_{ij} f'_j(\vec{x}|m_0). \quad (\text{A.2})$$

Here, M_{ij} is the $n \times n$ transformation matrix $M_{ij} = (m_i - m_0)^j$. Inverting the equation allows the prediction of the template shape at any given model parameter m' :

$$f_{pred}(\vec{x}|m') = \sum_{j=0}^{n-1} c_j(m') f'_j(\vec{x}|m_0). \quad (\text{A.3})$$

In the analysis, the $f(\vec{x}|m_i)$ are given by the simulated signal samples for the different mass points, and c_i incorporate the inverse of the transformation matrix

$$c_i(m') = \sum_{j=0}^{n-1} (m' - m_0)^j (M^{-1})_{ji}. \quad (\text{A.4})$$

Therefore, predictions for any value of m' can be made through Eq. A.3, normalized to a linear interpolation of the total event count in the neighboring simulated points.

¹Particularly notable as, though it has now become more commonplace in other ATLAS resonance searches, this analysis was one of the first in ATLAS to develop and implement this technique in a published result of this type.

²It goes without saying that particle physics would undoubtedly be very different if we could write down analytically exact functions for hypothetical data distributions without observing them!

A.3.2 Validation and Quality Studies

Now that the morphing procedure has been established, it can be used to create new m_{VH} distributions for any value of W' mass for each analysis category, thus allowing for hypothesis tests at mass resolution spacing. To validate that this procedure creates sufficiently accurate distributions within systematic and statistical uncertainties, a test is performed by creating morphed versions of simulated mass points and comparing the resulting distributions with their real counterparts. The original samples are divided into train and test groups as summarized in Table A.1.

Table A.1. Train and test group for the W' signal morphing quality study

Group	W' Signal Mass Points [GeV]	Total
Train	400, 600, 800, 1000, 1400, 1800, 2200, 2600, 3000, 4000, 5000	11
Test	500, 700, 900, 1200, 1600, 2000, 2400, 2800, 3500, 4500	10

Selected validation distributions are shown in Figures A.4 and A.5 for the resolved and merged categories, demonstrating that the morphed samples approximate the simulated distributions accurately. It is important to note that these quality checks are inherently very conservative by design. To validate a morphed mass point against the same simulated points, the morphing must use information skipping the available simulated points being validated against, resulting in information gaps that are twice as wide as the “production” morphing that uses all available signal samples to generate the PDFs with much higher accuracy. Performance rapidly degrades, however, when the morphed mass point is beyond the range of simulated mass points – that is, the technique does not perform well at *extrapolations* and is only well-suited for *interpolations*. Performance is also less optimal in regimes with low event yields in the training m_{VH} distributions, such as the resolved categories for higher W' masses and merged categories for lower W' masses. This is due to the interpolating nature of the morphing technique, as the shape of the morphed template is intrinsically smoothed and thus over-pronounces the bin-by-bin variations of the simulated sample, which is addressed through a conservative systematic uncertainty.

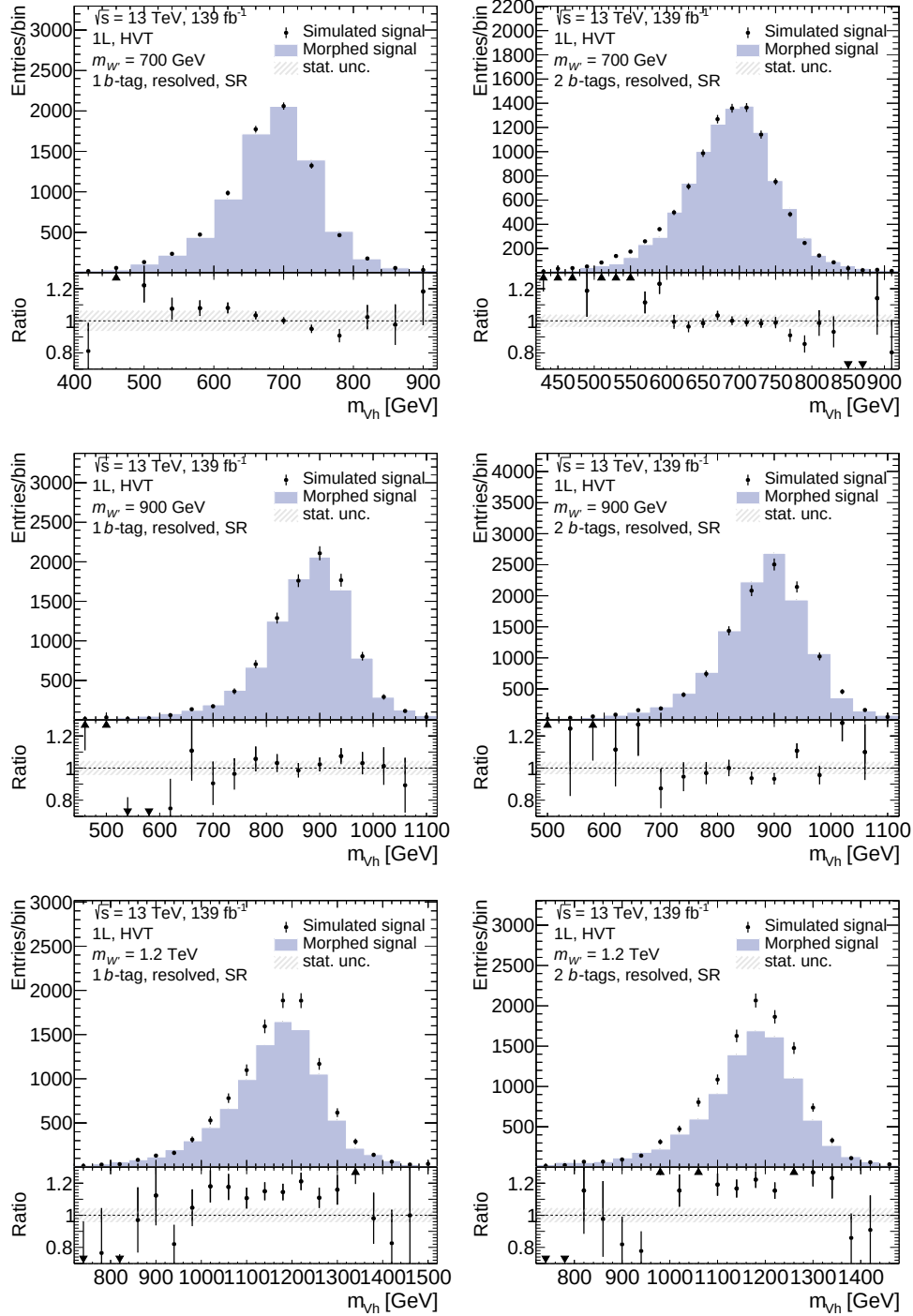


Figure A.4. Comparison of simulated (black) and morphed (blue) samples in the 1-lepton resolved channel for the HVT W' signal with masses of 0.7, 0.9, and 1.2 TeV. The selected region is indicated in the text of the individual figures.

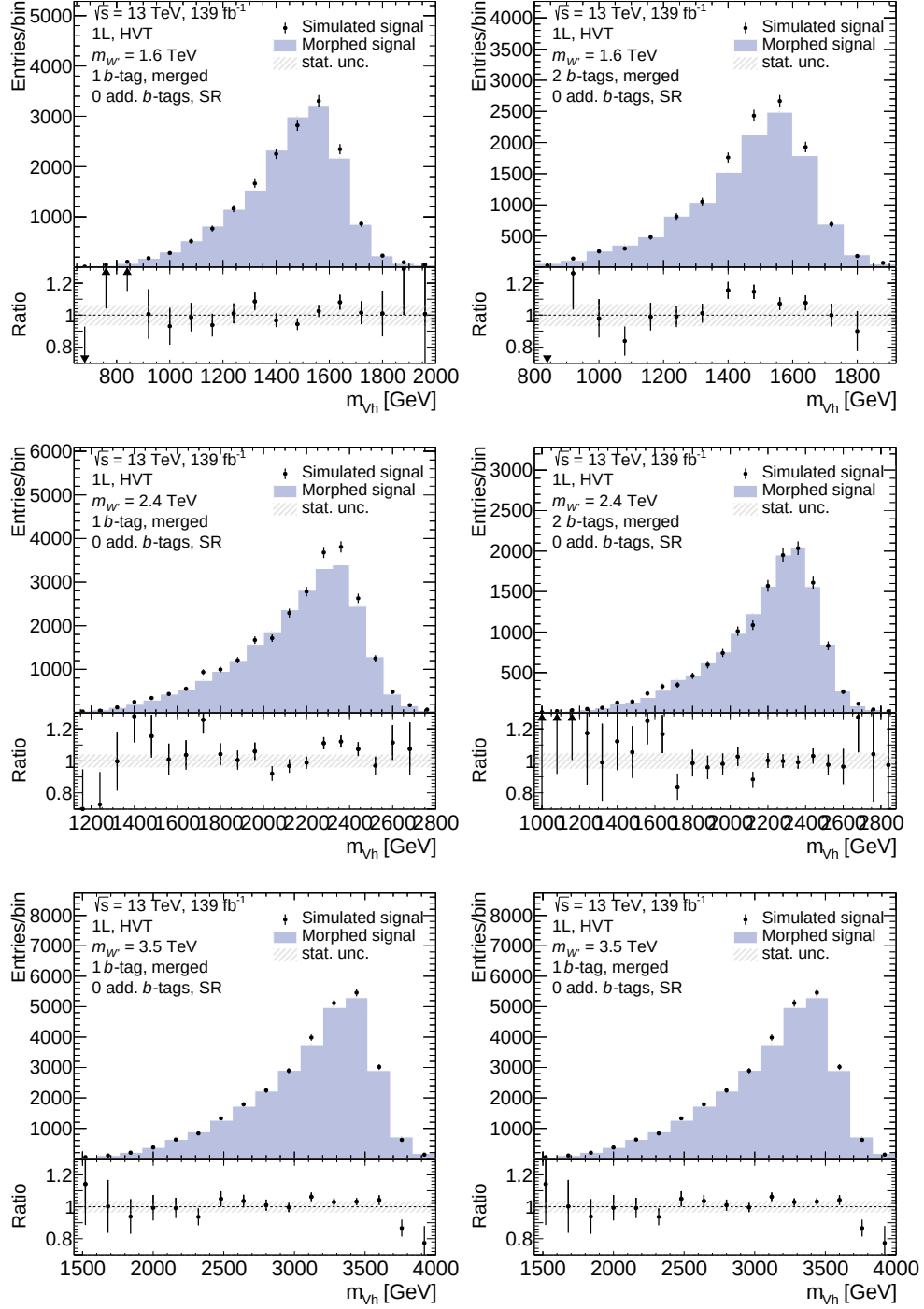


Figure A.5. Comparison of simulated (black) and morphed (blue) samples in the 1-lepton merged channel for the HVT W' signal with masses of 1.6, 2.4, and 3.5 TeV. The selected region is indicated in the text of the individual figures.

A.3.3 Uncertainties and Performance

All systematic uncertainties assigned to the simulated samples are also propagated to the morphed samples. This is accomplished by repeating the morphing procedure individually for each alternative distribution for each uncertainty variation in place of the nominal distribution. Morphing all mass points for the nominal distribution and each uncertainty variation per analysis category is computationally very inexpensive, taking on the order of seconds for each. Still, due to the sheer total volume of systematic variations in the analysis, the entire procedure takes around 24 hours of CPU time. As the morphing of each systematic variation is entirely independent, the procedure is well-suited to multithreading.

Additionally, the statistical uncertainty of the simulated samples is propagated to the morphed samples using a conservative approach. This is calculated by taking the two closest simulated mass points above and below the morphed sample of mass M and defining the uncertainty using their integrals I and statistical uncertainties σ :

$$\frac{\sigma_M}{I_M} = \frac{1}{\sqrt{2}} \left(\frac{\sigma_{m < M}}{I_{m < M}} + \frac{\sigma_{m > M}}{I_{m > M}} \right), \quad (\text{A.5})$$

which is the arithmetic mean, increased by an additional factor of $\sqrt{2}$. The statistical uncertainty is propagated to every bin of the morphed histogram by multiplying it with the bin content.

The expected upper limits on the signal production cross section were compared for simulated and morphed samples, and morphed samples were found to result in upper limits close to, but consistently slightly (10%) higher (more conservative) than the corresponding limits calculated from simulated samples. This can be attributed to the additional factor of $\sqrt{2}$ assigned to the statistical uncertainty of the morphed sample, which is an intentional design choice to yield more conservative results in kinematic regions not covered by the simulated samples. Morphed samples are successfully produced in 25 GeV increments from $m_{W'} = 400$ GeV to 800 GeV, in 50 GeV increments from 800 GeV to 3000 GeV, and 100 GeV increments from 3000 GeV to 5000 GeV and used in the final analysis and limit setting.

APPENDIX B

BACKGROUND MODELING STUDIES

While it would be ideal to have perfect simulations applicable in every possible region of phase space, we unfortunately do not yet live in a world where this is reality. In the meantime, this Appendix details some of the various modeling techniques, corrections, and uncertainties derived throughout the Wh resonance search.

B.1 QCD Multijet Estimation

QCD “multijet” production is the most dominant process of proton-proton collisions, owing to the asymptotic freedom at high energies in QCD. Contributions from these events in the search regions are intentionally heavily suppressed by the analysis selections, which place strict requirements on the reconstruction, identification, and isolation criteria of the physics objects, particularly the “hard” prompt lepton from the W boson decay in the 1-lepton channel. Still, due to the sheer volume of this type of event, it is possible for some contamination to leak through, which is true only non-negligibly for the 1-lepton channel in this search.

A small fraction of multijet events have two real b -quarks from gluon splitting or have their light jets misidentified as b -jets due to the small inefficiencies of the b -tagging discriminant. Prompt and well-isolated leptons, though, are generally not expected in multijet processes but can mimic the electron signature in several ways. At high momentum, non-prompt leptons from weak decays inside the parton shower of the jet dominate the multijet contribution, while at low momentum, the large number of light jets gives rise to a slight possibility of an individual jet mimicking an electron signature. Photons from pion decays can also fake an electron signature, though this is less of a contributor. The muon signature suffers a similar issue as to the electron case at high momentum, but the multijet contribution for muons is much smaller and easier to account for due to their unique signature and cleaner identification.

Multijet events are also characterized by a small amount of E_T^{miss} compared to other background processes, as they generally deposit most of their energy in the calorimeter and do not produce “hard” neutrinos. Thus, the 0-lepton channel does not suffer any significant multijet contamination.

The selection for E_T^{miss} varies for the electron and muon channels in the 1-lepton channel for this reason, requiring 80 GeV for electrons and only 40 GeV for muons as they require less suppression power.

B.1.1 Template Method

MC samples for QCD multijet processes suffer from limited statistics and poor modeling in this analysis’s targeted phase space. This is broadly true for most ATLAS searches targeting these high-mass regimes, so a template method is used as the standard prescription. This method exploits the lepton’s isolation criteria to create a control region (**MJCR**) from data that is naturally enriched in multijet events. Leptons from the signal process are expected to be well-isolated from hadronic activity since events will contain a W boson in one hemisphere and a Higgs boson in the other due to the “back-to-back” nature of the high mass $W' \rightarrow Wh$ decay. Additional contributions from ISR, etc., surrounding the lepton are expected to be small. On the other hand, leptons in multijet events are located inside jets and thus are not well-isolated from additional hadronic activity.

The electron isolation criteria are defined by $topoetcone20 < 0.06$, which requires the energy sum within a cone of $\Delta R = 0.2$ in the calorimeter system around the electron cluster axis to be smaller than 0.06 times the electron E_T , and $ptvarcone20 < 0.06$, where the sum of the transverse momenta of all ID tracks within a variable-radius cone up to $\Delta R = 0.2$, shrinking with increasing p_T , around the electron track is required to be smaller than 0.06 times the electron E_T , respectively. For the MJCR, the calorimeter $topoetcone20$ is inverted for electrons. Additionally, as the overall statistics are low, the E_T^{miss} requirement is removed.

For muons, only the tracking-based $ptvarcone30$ component (same as above but with $\Delta R = 0.3$) is used as they do not use the calorimeter information. The MJCR for the muon case is selected by $0.06 < ptvarcone30 < 0.15$. Both the electron and muon cases are required to pass the *Tight* identification criteria in the MJCR, as in the SRs.

The multijet estimation from the non-isolated region is based on two assumptions. First, it is expected that the shape of the distributions of the observables between isolated and non-isolated regions will agree well; thus, templates of the multijet contribution can be derived in the non-isolated region and applied in the signal region. Second, background from the non-multijet processes, such as top and W +jet samples, are assumed to be well-modeled in the non-isolated region; thus, any differences between data and MC predictions in this region can be attributed to the multijet contribution. The templates are then built by subtracting the MC prediction of the non-multijet contributions from the data. Separate categories are constructed to account for potential differences

in the shape of the multijet template, for resolved vs. merged, electron vs muon, and number of b -tagged jets.

Inclusive categories are built for the number of jets in the resolved category. The templates from the MJCR are derived and applied separately to the signal regions and sideband control regions. Since the overall normalization of the templates is given by the ratio of events with isolated and non-isolated leptons, a fit to data using the `TFractionFitter` package in ROOT is performed in the $m_{T,W}$ distribution as it yields the largest shape difference between multijet and non-multijet backgrounds, taking into account only the statistical uncertainties. The normalization of the multijet templates from this fit is used as input to the final statistical analysis. Ultimately, the only non-negligible multijet contribution is in the 1 b -tag resolved regions. The pre-fit MJCR for these regions is shown in Figure B.1.

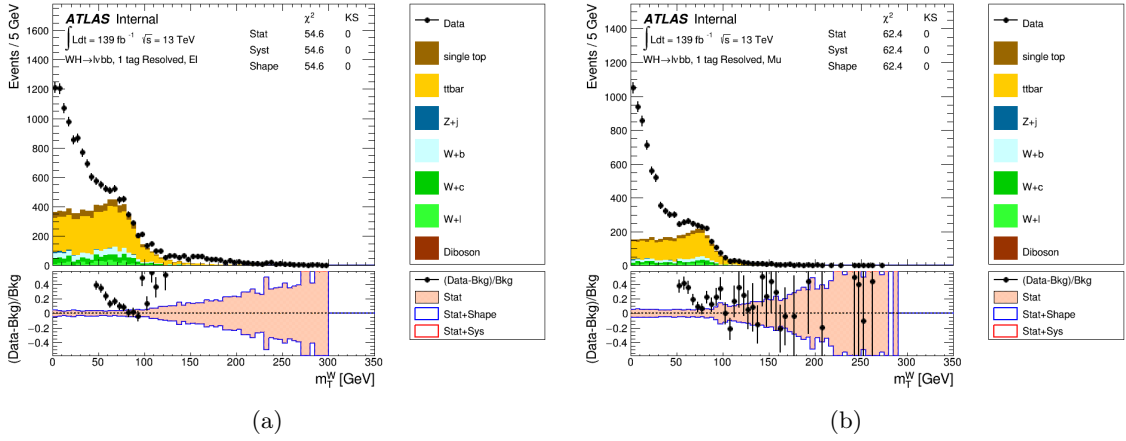


Figure B.1. m_T^W distributions in multijet control region in the electron channel (a) and the muon channel (b) for the 1-lepton 1 b -tag resolved regions.

B.1.2 Results

An example postfit distribution for the MJ template fit in $m_{T,W}$ in the muon channel is shown in Figure B.2.

To check the stability of the template fit, the multijet fractions were derived individually in different datataking periods, shown in Table B.1. Note that these studies were performed at an earlier stage of the analysis, where the sideband control regions were initially split into separate low and high-mass sideband categories. These were later merged throughout for fit stability reasons, but the conclusions remain valid.

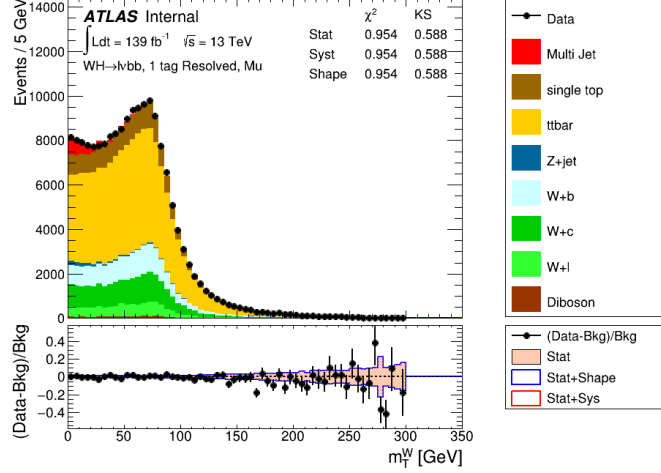


Figure B.2. Postfit distribution of m_T^W for the 1-lepton, 1 b -tag resolved region.

Periods	Low m_{jj} CR, Electron	High m_{jj} CR, Electron	Low m_{jj} CR, Muon	High m_{jj} CR, Muon
Data1516	0 ± 0.0055	0.0092 ± 0.0137	0.0230 ± 0.0060	0.0297 ± 0.0073
Data17	0 ± 0.0077	0.0083 ± 0.0117	0.0199 ± 0.0066	0.0264 ± 0.0073
Data18	0 ± 0.0062	0.0140 ± 0.0079	0.0121 ± 0.0059	0.0262 ± 0.0060
Full Run2	0 ± 0.0051	0.0114 ± 0.0052	0.0172 ± 0.0035	0.0277 ± 0.0039

Table B.1. Multijet fractions from template fit in different control regions in electron and muon channels for different data-taking periods.

Shape uncertainties for the multijet templates were derived by shifting the E_T^{miss} selections by 25% and varying the normalization of the $t\bar{t}$ background by 10%. These are ultimately negligible due to the minimal contribution of the multijet background and are pruned by the final fit.

B.2 0-Lepton Mismodeling Corrections

Mismodeling effects were observed in some key pre-fit observables in both the 0- and 1-lepton channels. Significant deviations from unity in data vs. MC comparisons in the control regions that increase in magnitude with the value of the observable, leading to a “slope” in the corresponding ratio, indicate that a process is not appropriately well-modeled. This can be examined on a constituent level for observables corresponding to the individual objects entering the m_{Wh} and $m_{T,Wh}$ distributions, and corrections and/or uncertainties can be derived from these observed differences in the control regions. This section describes the corrections for the 0-lepton channel.

B.2.1 $p_{T,jj}$ Reweighting

Significant slopes were observed in the m_{jj} observables in the W and Z +jets backgrounds in the 0-lepton resolved channel, stronger for Z +jets than W +jets.¹ To enrich the Z +jets background composition and derive a correction, a special “0 b -tag” category is used as W and Z +jets backgrounds are dominant and equally large in this regime. The distribution of the $p_{T,jj}$ observable within this 0-tag region is shown in Figure B.3.

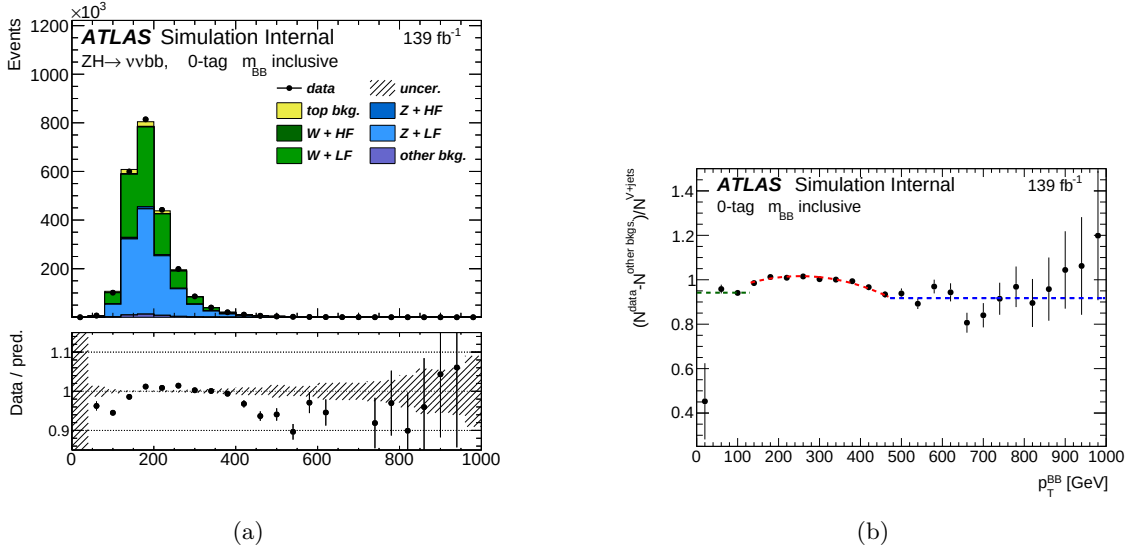


Figure B.3. Pre-fit distribution (a) of the $p_{T,jj}$ observable presented in the 0-tag control region with data-to-simulation ratio and shaded statistical uncertainty, as well as the piecewise correction function derived from this distribution (b).

A clear modulation of the data-to-simulation ratio is observed. A piecewise function is fit through various parts of the $p_{T,jj}$ spectrum to obtain correction scale factors for three separate $p_{T,jj}$ intervals:

$$w(p_{T,jj}) = \begin{cases} 0.946 & \text{for } 0 \text{ GeV} < p_{T,jj} \leq 130 \text{ GeV}, \\ 0.888 + 0.001 * p_{T,jj} - 2.003 \cdot 10^{-6} (p_{T,jj})^2 & \text{for } 130 \text{ GeV} < p_{T,jj} \leq 470 \text{ GeV}, \\ 0.917 & \text{for } p_{T,jj} \geq 470 \text{ GeV}, \end{cases} \quad (\text{B.1})$$

¹In the interest of full disclosure, this effect was primarily observed in the 2-lepton channel of the broader analysis, where Z +jets backgrounds are much more dominant and this represented a much larger problem. These corrections were derived for the 0-lepton channel largely to harmonize the procedure for the combined Z' and 2HDM fits, which are outside the scope of this thesis – this is why the Figures in this section are labeled $Zh \rightarrow \nu\nu b\bar{b}$. But they are discussed here for completeness as they do enter the W' fit as well.

These correction factors are then applied to all components of the $W + \text{jets}$ and $Z + \text{jets}$ backgrounds within the 1- and 2-tag CR and SRs. The prefit distribution of the m_{Wh} and p_T^{miss} observables are presented in Figure B.4 separately for the 1- and 2-tag CRs. Both observables are presented after the $p_{T,jj}$ corrections have been applied. While the $p_{T,jj}$ reweighting sufficiently corrects the shape of the m_{Wh} distribution, residual shape differences between the data and the MC predictions are observed for the p_T^{miss} observables. Dedicated fits to their data-to-simulation ratios are performed in order to derive an additional uncertainty dedicated to covering the modeling of this observable. However, as the origin of the slope seen in the data-to-simulation ratios of the p_T^{miss} observable can not be clearly determined from these distributions, these additional uncertainties are assigned to all background contributions equally. The p_T^{miss} mismodeling uncertainties are described via

$$w(p_T^{\text{miss}}) = \begin{cases} 1.200 - 0.003 \cdot p_T^{\text{miss}} + 5.801 \cdot 10^{-6} (p_T^{\text{miss}})^2 & \text{for } 50 \text{ GeV} < p_T^{\text{miss}} \leq 180 \text{ GeV}, \\ 0.926 & \text{for } p_T^{\text{miss}} \geq 220 \text{ GeV}, \end{cases} \quad (\text{B.2})$$

$$w(p_T^{\text{miss}}) = \begin{cases} 1.116 - 0.001 \cdot p_T^{\text{miss}} & \text{for } 50 \text{ GeV} < p_T^{\text{miss}} \leq 220 \text{ GeV}, \\ 0.901 & \text{for } p_T^{\text{miss}} \geq 220 \text{ GeV}, \end{cases} \quad (\text{B.3})$$

for the 1- and 2 b -tag categories, respectively. These functions are shown in Figure B.5.

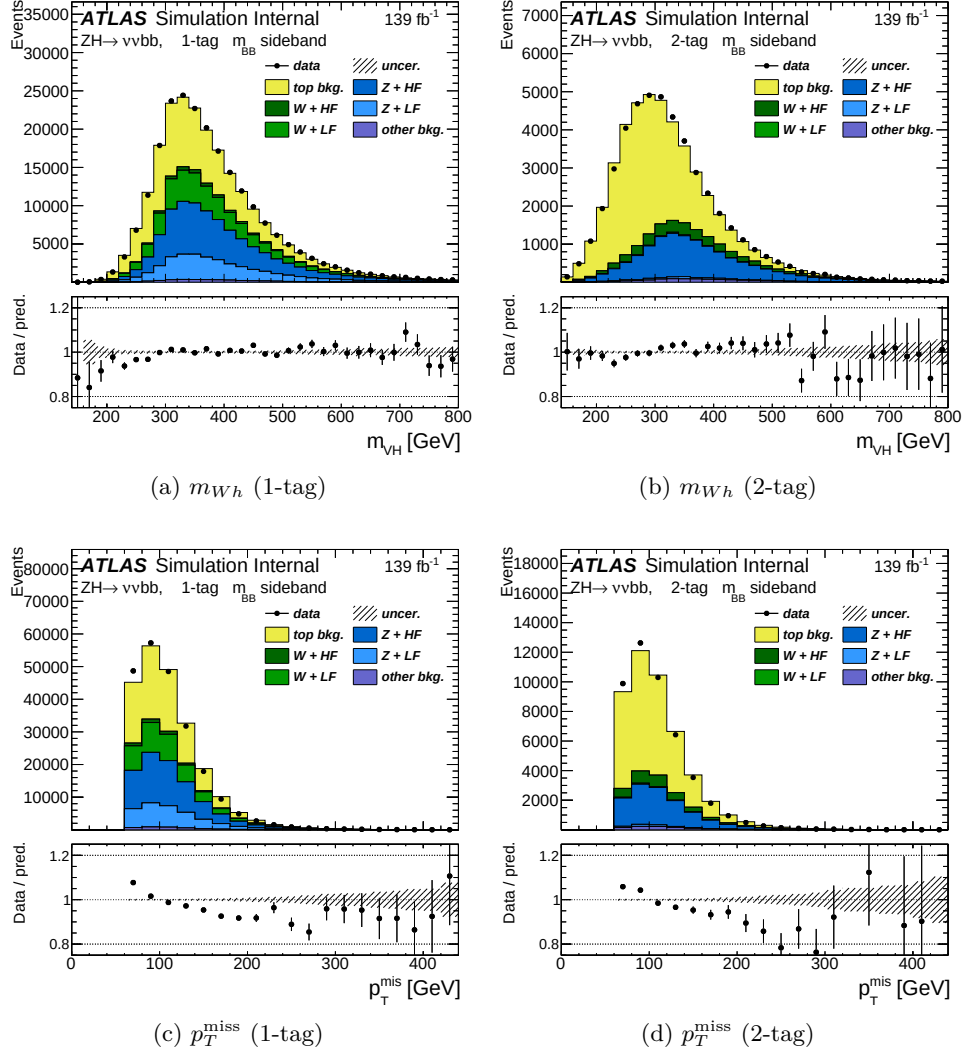
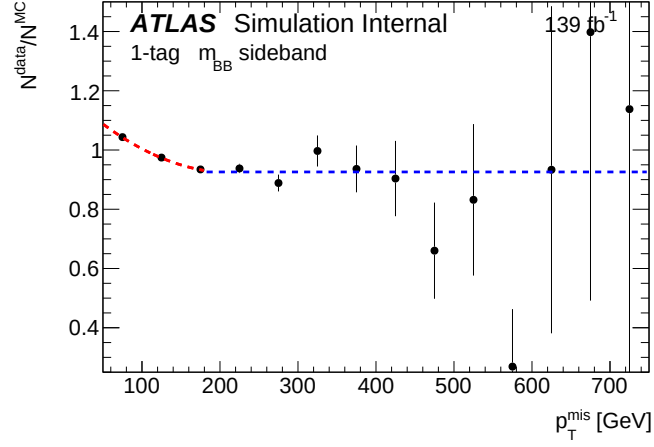
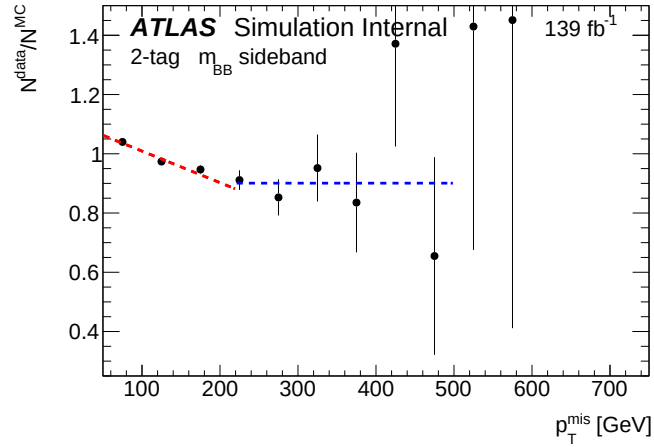


Figure B.4. Pre-fit distribution of the m_{Wh} (top row) and p_T^{miss} (bottom row) observables presented separately in the 1- (left column) and 2-tag (right column) CRs. The observables are presented after the $p_{T,jj}$ corrections have been applied.



(a) p_T^{miss} (1-tag)



(b) p_T^{miss} (2-tag)

Figure B.5. Fits through the data-to-simulation ratios obtained from studying the p_T^{miss} observables within the CRs of the resolved 1- and 2-tag categories.

B.2.2 Large-R jet p_T Uncertainties

In addition, significant data-to-simulation slopes were observed in the large-R jet p_T distributions in both the 1- and 2 b -tag merged categories in all backgrounds. The corresponding distributions are shown in Figure B.6. Uncertainties were derived from the data-to-simulation ratios observed in the control regions via

$$w(p_{T,J}) = \begin{cases} 1.3651 - 0.0010 \cdot +3.4051 \cdot 10^{-8} (p_{T,J})^2 & \text{for } 250 \text{ GeV} < p_{T,J} \leq 850 \text{ GeV}, \\ 0.5340 & \text{for } p_{T,J} \geq 850 \text{ GeV}, \end{cases} \quad (\text{B.4})$$

and

$$w(p_{T,J}) = \begin{cases} 1.0622 + 0.0002 \cdot p_{T,J} - 1.0254 \cdot 10^{-6} (p_{T,J})^2 & \text{for } 250 \text{ GeV} < p_{T,J} \leq 700 \text{ GeV}, \\ 0.7499 & \text{for } p_{T,J} \geq 700 \text{ GeV}, \end{cases} \quad (\text{B.5})$$

for the 1- and 2-tag event categories, respectively. These additional uncertainties are separately applied to all relevant background components.

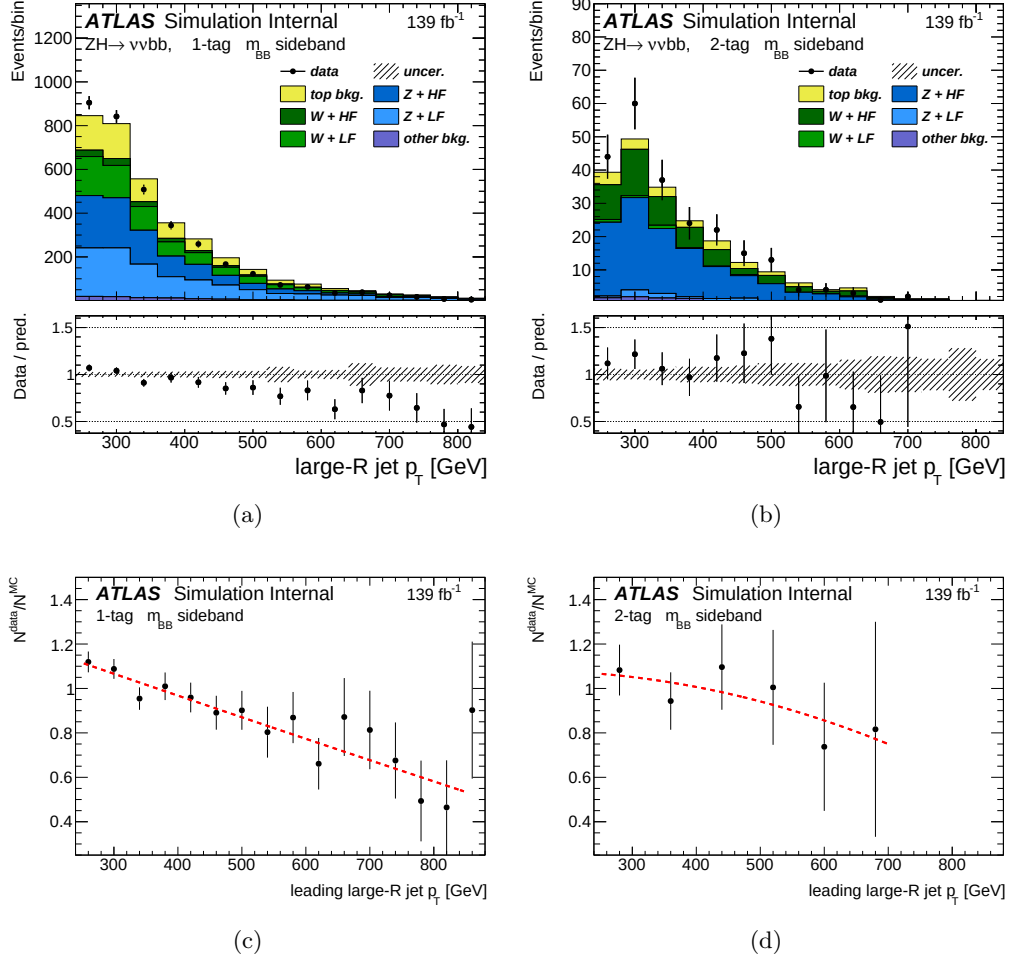


Figure B.6. Pre-fit distribution of the large- R jet p_T observable presented in the 1-tag (left) and 2-tag (right) CRs (top row), as well as the piecewise correction functions derived from these distributions (bottom row).

B.3 1-Lepton Modeling Uncertainties

Some shape mismodelings were also observed in the 1-lepton channel in both the merged and resolved categories. Unlike the 0-lepton channel, no direct reweightings are performed. Instead, systematic uncertainties are derived for all effects and included in the final fit as NPs.²

B.3.1 Resolved p_T^W

Significant slopes were observed in the resolved regime in the p_T^W distributions in the CR sidebands for both 1- and 2-tags. The piecewise uncertainties derived from the data-to-simulation ratios for p_T^W are shown in B.7.

The fitted function for 1-tag is described by:

$$w(p_T^W) = \begin{cases} 1.034 - 2.666 \cdot 10^{-5} \cdot p_T^W - 6.817 \cdot 10^{-7} \cdot (p_T^W)^2 & \text{for } p_T^W < 600 \text{ GeV} \\ 0.91 & \text{otherwise.} \end{cases} \quad (\text{B.6})$$

and for 2-tag by:

$$w(p_T^W) = \begin{cases} 1.062 - 3.971 \cdot 10^{-5} \cdot p_T^W - 1.387 \cdot 10^{-6} \cdot (p_T^W)^2 & \text{for } p_T^W < 400 \text{ GeV} \\ 0.89 & \text{otherwise.} \end{cases} \quad (\text{B.7})$$

For the fit, the p_T^W related uncertainties are split by the major backgrounds and treated as individual nuisance parameters for $W+1$, $W+hl$, $W+hf$, top, and the rest. No reweighting of the p_T^W distribution is done.

B.3.2 Large-R Jet p_T

In the merged regime, the p_T^W is better modeled, and no significant slopes are observed. However, similarly to the 0-lepton channel, the large-R jet p_T distributions show significant slopes in the control regions. A piecewise function is again derived for use as an uncertainty, described for 1-tag by:

²Reweightings were attempted in earlier iterations, but it was found that the CR-only fit would “undo” the correction by pulling the uncertainties accompanying the reweighting in the opposite direction. This indicated the mismodelings were not severe enough to warrant a complete reweighting treatment.

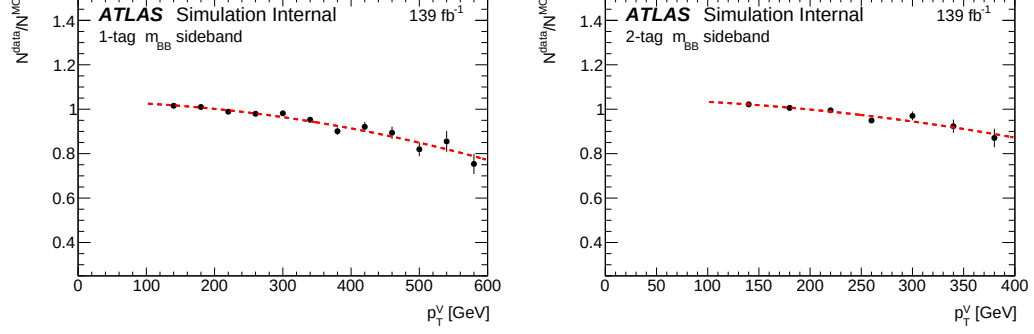


Figure B.7. The relative uncertainties for the p_T^W slope in the resolved regime for 1-tag (left) and 2-tag (right) in the CRs.

$$w(p_{T,J}) = \begin{cases} 1.235 - 0.001 \cdot p_{T,J} + 8.291 \cdot 10^{-8} \cdot (p_{T,J})^2 & \text{for } p_{T,J} < 1300 \\ 0.481 & \text{otherwise.} \end{cases} \quad (\text{B.8})$$

and for the 2-tag by:

$$w(p_{T,J}) = \begin{cases} 1.596 - 0.002 \cdot p_{T,J} + 1.014 \cdot 10^{-6} \cdot (p_{T,J})^2 & \text{for } p_{T,J} < 850 \text{ GeV} \\ 0.569 & \text{otherwise.} \end{cases} \quad (\text{B.9})$$

For the fit, the uncertainties are again split by the major backgrounds and treated as individual nuisance parameters for $W+1$, $W+hl$, $W+hf$, top, and the rest.

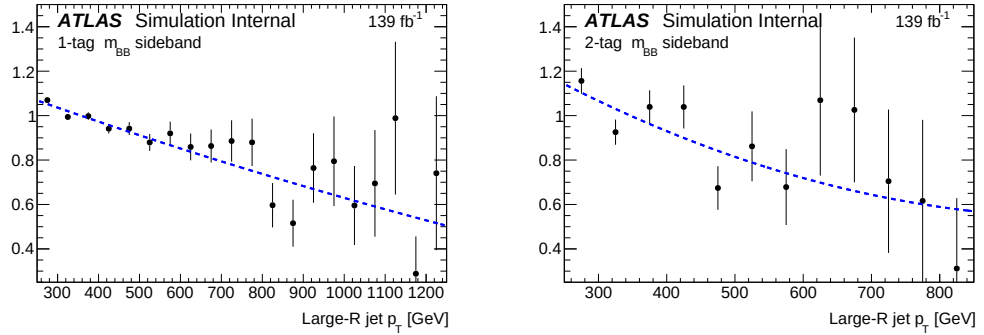


Figure B.8. The relative difference for the large-R jet p_T slope in the resolved regime for 1-tag (left) and 2-tag (right) in the merged CRs.

B.4 Single-top Wt DS/DR Uncertainties

The standard procedure for accounting for interference between single top Wt channel production and $t\bar{t}$ production is removed by setting the amplitudes of the overlapping diagrams to zero before

the simulation, known as the *diagram removal* (DR) scheme. This method is simple to implement but can lead to gauge-dependent results because the interference terms are not handled carefully. An alternative to this method is to compute the amplitudes of the overlapping diagrams and then explicitly subtract the cross section after the fact, known as the *diagram subtraction* (DS) scheme. This is harder to simulate and characterize and, as such, is not used by default in ATLAS simulation but is a more methodologically “correct” approach. As single top production is not a significant background for the 0-lepton channel, the difference between these two approaches has a negligible impact. However, Wt production is a more substantial background for the 1-lepton channel, comprising nearly 10% of the total top background. The differences in the Wt m_{Wh} distributions become non-negligible in the high mass tails, as shown in the resolved and merged regimes in Figure B.9 and Figure B.10 respectively.

An uncertainty is derived from these shape differences and applied in the fit to the total sum of the top background, where weights are applied to each event through a second-order polynomial fit to the differences in the total top background in the four signal regions. As all top backgrounds are combined in the fit, this uncertainty is derived from the total shape difference in all combined top backgrounds rather than the shape difference in single top Wt only, where $t\bar{t}$ comprises the vast majority of the total top background (as well as minimal additional contributions from single top s - and t - channel production.) These combined $t\bar{t}$ and single top Wt distributions for resolved and merged regimes are shown in Figure B.11 and Figure B.12 respectively. The fit functions are described for the resolved 1- and 2-tag categories (respectively) via:

$$w(m_{Wh}) = \left\{ \begin{array}{ll} 1.0 & \text{for } m_{Wh} < 600\text{GeV}, \\ 1 + ((-1.04 \cdot 10^{-7} * m_{Wh}^2) + (6.66 \cdot 10^{-5} * m_{Wh}) - 0.0063) & \text{for } 600\text{GeV} \leq m_{Wh} \leq 1400\text{GeV}, \\ 0.8831, & \text{for } 1400\text{GeV} < m_{Wh}. \end{array} \right\} \quad (\text{B.10})$$

$$w(m_{Wh}) = \left\{ \begin{array}{ll} 1.0 & \text{for } m_{Wh} < 500\text{GeV}, \\ 1 + ((-2.77 \cdot 10^{-7} * m_{Wh}^2) + (1.12 \cdot 10^{-4} * m_{Wh}) - 0.00447) & \text{for } 500\text{GeV} \leq m_{Wh} \leq 1000\text{GeV}, \\ 0.8384, & \text{for } 1000\text{GeV} < m_{Wh}. \end{array} \right\} \quad (\text{B.11})$$

and the merged 1- and 2- tag via:

$$w(m_{Wh}) = \left\{ \begin{array}{ll} 1.018 & \text{for } m_{Wh} < 550\text{GeV}, \\ 1 + ((-7.09 \cdot 10^{-8} * m_{Wh}^2) + (3.22 \cdot 10^{-5} * m_{Wh}) - 0.0057) & \text{for } 550\text{GeV} \leq m_{Wh} \leq 2000\text{GeV}, \\ 0.7089, & \text{for } 2000\text{GeV} < m_{Wh}. \end{array} \right\} \quad (\text{B.12})$$

$$w(m_{Wh}) = \left\{ \begin{array}{ll} 1.06 & \text{for } m_{Wh} < 550\text{GeV}, \\ 1 + ((1.10 \cdot 10^{-7} * m_{Wh}^2) - (5.81 \cdot 10^{-4} * m_{Wh}) + 0.34644) & \text{for } 550\text{GeV} \leq m_{Wh} \leq 1600\text{GeV}, \\ 0.698, & \text{for } 1600\text{GeV} < m_{Wh}. \end{array} \right\} \quad (\text{B.13})$$

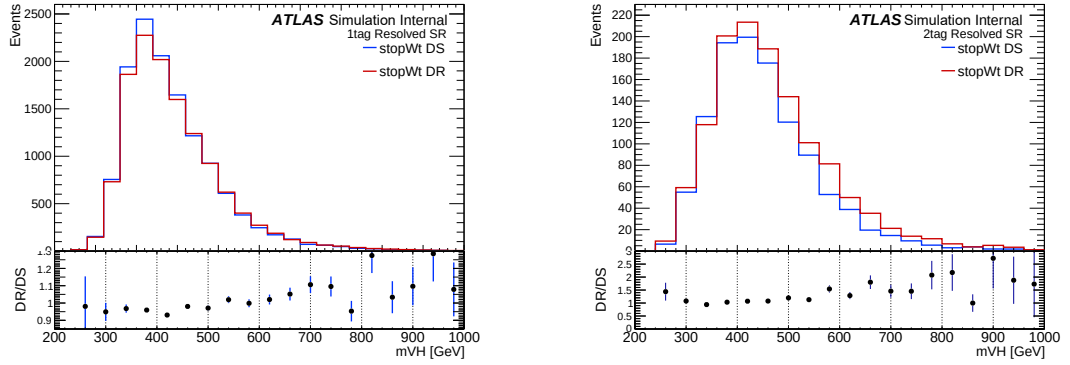


Figure B.9. The difference between the unrenormalized prefit distributions of stop DS/DR in the 1 b -tagged (left) and 2 b -tagged (right) resolved signal regions of the 1-lepton channel.

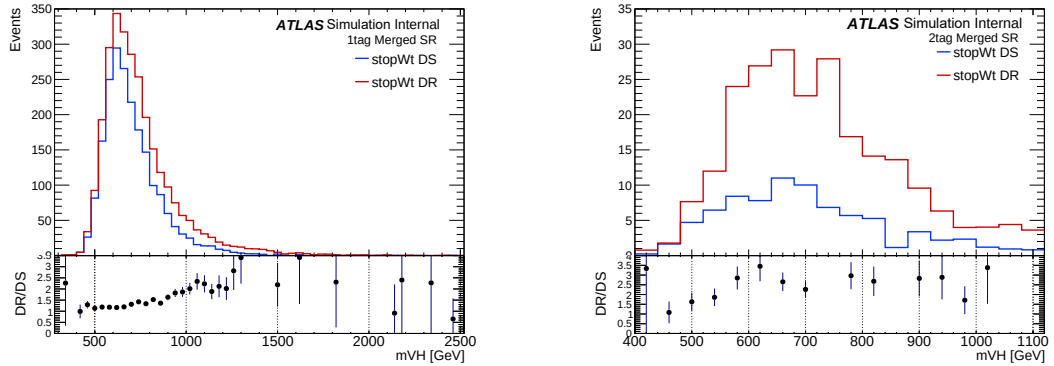


Figure B.10. The difference between the unrenormalized prefit distributions of stop DS/DR in the 1 b -tagged (left) and 2 b -tagged (right) merged signal regions of the 1-lepton channel.

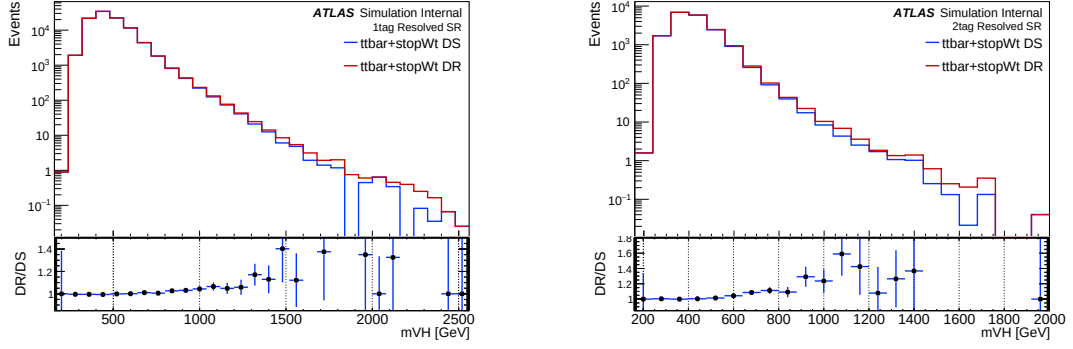


Figure B.11. The prefit m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DR scheme) backgrounds compared to the m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DS scheme) backgrounds presented separately for the 1 b -tagged (left) and 2 b -tagged (right) resolved signal regions of the 1-lepton channel.

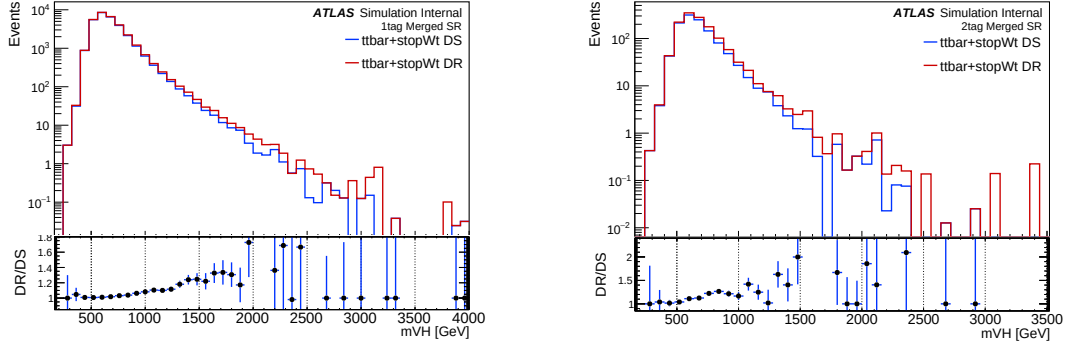


Figure B.12. The prefit m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DR scheme) backgrounds compared to the m_{Wh} distribution for the combined $t\bar{t} + Wt$ (DS scheme) backgrounds presented separately for the 1 b -tagged (left) and 2 b -tagged (right) merged signal regions of the 1-lepton channel.

B.5 1-Lepton JetBinMigration Uncertainty

In contrast to the 0-lepton channel, which selects $N_{\text{Jets}} \geq 2$, the 1-lepton channel uses an *exclusive* jet bin selection requirement ($2 \leq N_{\text{Jets}} \leq 3$) to further suppress $t\bar{t}$ background. This implies the impact of jet bin migration effects due to additional QCD radiation has to be considered for the most dominant backgrounds. This is addressed with a simplified but conservative approach with respect to other “standard” techniques used in similar searches, which explicitly perform jet veto measurements or vary resummation parameters in the simulation to create a migration matrix [183] etc. Here, the true generator jet multiplicity is determined per event by applying the same jet selection requirements on the generator-level “truth” jets as the reconstructed jets. Two separate

uncertainties are then introduced for the W +jets background: The weight of events with less than two jets at the generator level get scaled up and down by 50% (*truthWZlessthan2*) and those with more than three jets at the generator level are also scaled up and down by 50% (*truthWZmorethan3*). The corresponding shape uncertainties are shown in Figure B.13.

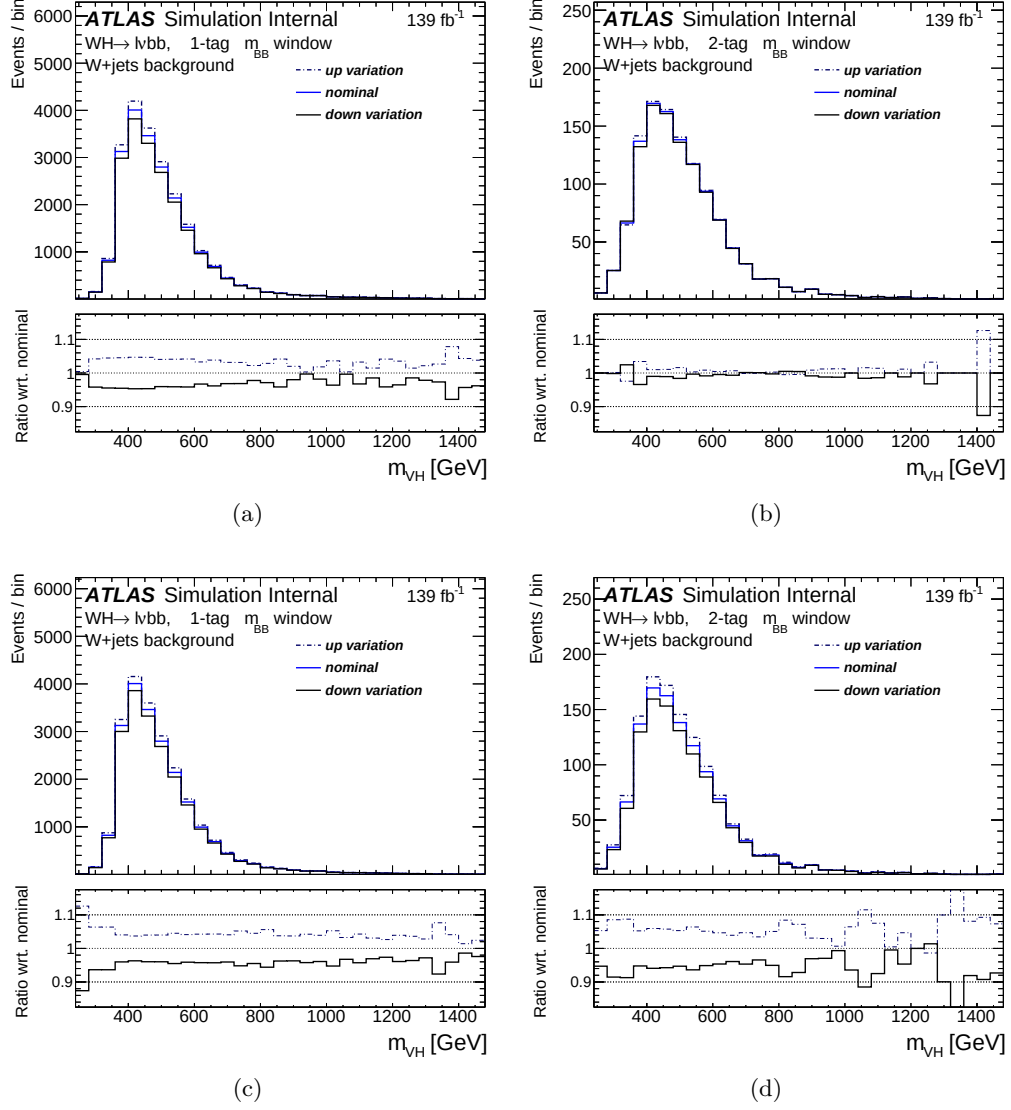


Figure B.13. Comparison between the nominal m_{W_h} distribution and the up and down variations of the *truthWZlessthan2* (top row) and *truthWZmorethan3* (bottom row) uncertainties presented separately for the W +jets background within the 1- and 2-tag signal region.

...

BIBLIOGRAPHY

- [1] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716(1):1–29, September 2012. Available online at <https://doi.org/10.1016/j.physletb.2012.08.020>.
- [2] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716(30):30–61, September 2012. Available online at <https://doi.org/10.1016/j.physletb.2012.08.021>.
- [3] C.H.Llewellyn Smith and G.G. Ross. The real gauge hierarchy problem. *Physics Letters B*, 105(1):38–40, 1981. Available online at [https://doi.org/10.1016/0370-2693\(81\)90035-6](https://doi.org/10.1016/0370-2693(81)90035-6).
- [4] Martinus J.G Veltman. The Infrared-Ultraviolet Connection. *Origin of Symmetries*, pages 526–546, 1991.
- [5] Steven Weinberg. Gauge hierarchies. *Phys. Lett. B*, 82(3–4):387–391, April 1979.
- [6] Francesco Sannino and Kimmo Tuominen. Orientifold theory dynamics and symmetry breaking. *Phys. Rev. D*, 71:051901, Mar 2005. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.71.051901>.
- [7] Roshan Foadi, Mads T. Frandsen, Thomas A. Rytto, and Francesco Sannino. Minimal walking technicolor: Setup for collider physics. *Phys. Rev. D*, 76:055005, Sep 2007. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.76.055005>.
- [8] Alexander Belyaev, Roshan Foadi, Mads T. Frandsen, Matti Järvinen, Francesco Sannino, and Alexander Pukhov. Technicolor walks at the LHC. *Phys. Rev. D*, 79:035006, Feb 2009. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.79.035006>.
- [9] Martin Schmaltz and David Tucker-Smith. LITTLE HIGGS THEORIES. *Annual Review of Nuclear and Particle Science*, 55(1):229–270, 2005. Available online at <https://doi.org/10.1146/annurev.nucl.55.090704.151502>.
- [10] Michael J. Dugan, Howard Georgi, and David B. Kaplan. Anatomy of a composite Higgs model. *Nuclear Physics B*, 254:299–326, 1985. Available online at <https://www.sciencedirect.com/science/article/pii/0550321385902214>.
- [11] Kaustubh Agashe, Roberto Contino, and Alex Pomarol. The minimal composite Higgs model. *Nuclear Physics B*, 719(1):165–187, 2005. Available online at <https://www.sciencedirect.com/science/article/pii/S0550321305003445>.
- [12] Duccio Pappadopulo, Andrea Thamm, Riccardo Torre, and Andrea Wulzer. Heavy vector triplets: bridging theory and data. *Journal of High Energy Physics*, 2014(9):60, Sep 2014. Available online at [https://doi.org/10.1007/JHEP09\(2014\)060](https://doi.org/10.1007/JHEP09(2014)060).
- [13] J. de Blas, J. M. Lizana, and M. Pérez-Victoria. Combining searches of Z' and W' bosons. *Journal of High Energy Physics*, 2013(1):166, Jan 2013. Available online at [https://doi.org/10.1007/JHEP01\(2013\)166](https://doi.org/10.1007/JHEP01(2013)166).
- [14] The ATLAS Collaboration. Search for heavy resonances decaying into a Z or W boson and a Higgs boson in final states with leptons and b-jets in 139 fb⁻¹ of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2023(6):16, Jun 2023. Available online at [https://doi.org/10.1007/JHEP06\(2023\)016](https://doi.org/10.1007/JHEP06(2023)016).

- [15] V. Barger, W. Y. Keung, and Ernest Ma. Gauge model with light W and Z bosons. *Phys. Rev. D*, 22:727–737, Aug 1980. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.22.727>.
- [16] Roberto Contino, Duccio Pappadopulo, David Marzocca, and Riccardo Rattazzi. On the effect of resonances in composite Higgs phenomenology. *Journal of High Energy Physics*, 2011(10):81, Oct 2011. Available online at [https://doi.org/10.1007/JHEP10\(2011\)081](https://doi.org/10.1007/JHEP10(2011)081).
- [17] The ATLAS Collaboration. Software performance of the atlas track reconstruction for lh run 3. *Computing and Software for Big Science*, 8(1):9, Apr 2024. Available online at <https://doi.org/10.1007/s41781-023-00111-y>.
- [18] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995. Reading, USA: Addison-Wesley (1995) 842 p.
- [19] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 3 2014.
- [20] Andy Buckley, Christopher White, and Martin White. *Practical Collider Physics*. IOP, 12 2021.
- [21] Wikimedia Commons. Standard Model of Elementary Particles, 2019. File: Standard_Model_of_Elementary_Particles.svg. Available online at https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.
- [22] The Particle Data Group. Review of Particle Physics. *Phys. Rev. D*, 98:030001, Aug 2018. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [23] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961. Available online at <https://www.sciencedirect.com/science/article/pii/0029558261904692>.
- [24] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. Available online at <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [25] Abdus Salam. *Weak and electromagnetic interactions*, pages 244–254. Available online at https://www.worldscientific.com/doi/abs/10.1142/9789812795915_0034.
- [26] Antonio Pich. The Standard Model of Electroweak Interactions. The Standard model of electroweak interactions. pages 1–49, 2007.
- [27] Janosh Riebesell. Higgs Potential. Available: <https://tikz.net/higgs-potential/>.
- [28] The CDF Collaboration. High-precision measurement of the W boson mass with the CDF II detector. *Science*, 376(6589):170–176, 2022. Available online at <https://www.science.org/doi/abs/10.1126/science.abk1781>.
- [29] The ATLAS Collaboration. Search for heavy resonances decaying into a w or z boson and a higgs boson in final states with leptons and b -jets in 36 fb $^{-1}$ of $\sqrt{s} = 13\text{tev}$ pp collisions with the atlas detector. *Journal of High Energy Physics*, 2018(3):174, Mar 2018. Available online at [https://doi.org/10.1007/JHEP03\(2018\)174](https://doi.org/10.1007/JHEP03(2018)174).
- [30] John Campbell, Joey Huston, and Frank Krauss. *The Black Book of Quantum Chromodynamics: A Primer for the LHC Era*. Oxford University Press, 12 2017. Available online at <https://doi.org/10.1093/oso/9780199652747.001.0001>.
- [31] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Emma Slade, and Maria Ubiali. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77(10):663, Oct 2017. Available online at <https://doi.org/10.1140/epjc/s10052-017-5199-5>.

- [32] The Geant4 Collaboration. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. Available online at <https://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [33] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, August 2008. Available online at <https://dx.doi.org/10.1088/1748-0221/3/08/S08001>.
- [34] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, August 2008. Available online at <https://dx.doi.org/10.1088/1748-0221/3/08/S08003>.
- [35] *LEP design report Vol. 1: The LEP Injector Chain*. Report. CERN, Geneva, 1983. By the LEP Injector Study Group. Available online at <https://cds.cern.ch/record/98881>.
- [36] *LEP design report Vol. 2: The LEP Main Ring*. Report. CERN, Geneva, 1984. Copies shelved as reports in LEP, PS and SPS libraries.
- [37] CERN Service graphique. Overall view of the LHC. Vue d’ensemble du LHC. 2014. General Photo. Available online at <https://cds.cern.ch/record/1708849>.
- [38] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, August 2008. Available online at <https://dx.doi.org/10.1088/1748-0221/3/08/S08004>.
- [39] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, August 2008. Available online at <https://dx.doi.org/10.1088/1748-0221/3/08/S08002>.
- [40] The LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, August 2008. Available online at <https://dx.doi.org/10.1088/1748-0221/3/08/S08005>.
- [41] The FASER Collaboration. Technical Proposal for FASER: ForwArd Search ExpeRiment at the LHC. Technical report, CERN, Geneva, 2018.
- [42] Michael Benedikt, Paul Collier, V Mertens, John Poole, and Karlheinz Schindl. *LHC Design Report Vol 3: The LHC Injector Chain*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.
- [43] The UA1 Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Physics Letters B*, 122(1):103–116, 1983. Available online at <https://www.sciencedirect.com/science/article/pii/0370269383911772>.
- [44] The UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $p\bar{p}$ collider. *Physics Letters B*, 122(5):476–485, 1983. Available online at <https://www.sciencedirect.com/science/article/pii/0370269383916052>.
- [45] Julie Haffner. The CERN accelerator complex. Complexe des accélérateurs du CERN. 2013. General Photo. Available online at <http://cds.cern.ch/record/1621894>.
- [46] J. Wenninger. Machine Protection and Operation for LHC. *CERN Yellow Reports*, page Vol 2 (2016): Proceedings of the 2014 Joint International Accelerator School: Beam Loss and Accelerator Protection, 2016. Available online at <https://e-publishing.cern.ch/index.php/CYR/article/view/242>.
- [47] ATLAS Collaboration. Luminosity Public Results Run 2. Available: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.

- [48] ATLAS Collaboration. Luminosity Public Results Run 3. Available: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3>.
- [49] Neutelings, Izaak. CMS Coordinate System. Available: https://tikz.net/axis3d_cms/.
- [50] ATLAS Outreach. ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC. 2010.
- [51] The ATLAS Collaboration. Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2. *The European Physical Journal C*, 77(10):673, October 2017. Available online at <https://doi.org/10.1140/epjc/s10052-017-5225-7>.
- [52] The ATLAS Collaboration. Study of the material of the ATLAS inner detector for Run 2 of the LHC. *Journal of Instrumentation*, 12(12):P12009, December 2017. Available online at <https://dx.doi.org/10.1088/1748-0221/12/12/P12009>.
- [53] Bartosz Mindur. ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 845:257–261, 2017. Proceedings of the Vienna Conference on Instrumentation 2016.
- [54] Ahmet Bingül. The ATLAS TRT and its Performance at LHC. *Journal of Physics: Conference Series*, 347(1):012025, February 2012. Available online at <https://dx.doi.org/10.1088/1742-6596/347/1/012025>.
- [55] The ATLAS Collaboration. Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *The European Physical Journal C*, 81(8):689, August 2021. Available online at <https://doi.org/10.1140/epjc/s10052-021-09402-3>.
- [56] F. Bauer, U. Bratzler, H. Dietl, H. Kroha, Th. Lagouri, A. Manz, A. Ostapchuk, R. Richter, S. Schael, S. Chouridou, M. Deile, O. Kortner, A. Staude, R. Ströhmer, and T. Trefzger. Construction and test of MDT chambers for the ATLAS muon spectrometer. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 461(1):17–20, 2001. 8th Pisa Meeting on Advanced Detectors.
- [57] T. Argyropoulos, K. A. Assamagan, B. H. Benedict, V. Chernyatin, E. Cheu, J. Deng, A. Gordeev, I. Gough Eschrich, V. Grachev, K. Johns, S. Junnarkar, A. Kandasamy, V. Kaushik, A. Khodinov, A. Lankford, X. Lei, R. Murillo Garcia, K. Nikolopoulos, P. O’Connor, C. L. Parnell-Lampen, W. Park, V. Polychronakos, R. Porter, M. Purohit, M. Schernau, B. K. Toggerson, and D. Tompkins. Cathode strip chambers in ATLAS : Installation, commissioning and in situ performance. In *2008 IEEE Nuclear Science Symposium Conference Record*, pages 2819–2824, 2008.
- [58] Davide Boscherini. The ATLAS RPC System for the LHC Run-3. Technical report, CERN, Geneva, 2023.
- [59] Koichi Nagai. Thin gap chambers in ATLAS. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 384(1):219–221, 1996. BEAUTY ’96.
- [60] The ATLAS collaboration. Operation of the ATLAS trigger system in Run 2. *Journal of Instrumentation*, 15(10):P10004, Oct 2020. Available online at <https://dx.doi.org/10.1088/1748-0221/15/10/P10004>.
- [61] The ATLAS Collaboration. Performance of the atlas trigger system in 2015. *The European Physical Journal C*, 77(5):317, May 2017. Available online at <https://doi.org/10.1140/epjc/s10052-017-4852-3>.

- [62] The ATLAS Collaboration. Trigger Menu in 2016. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2017-001>.
- [63] The ATLAS Collaboration. Trigger Menu in 2017. Technical report, CERN, Geneva, 2018. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2018-002>.
- [64] The ATLAS Collaboration. Trigger menu in 2018. Technical report, CERN, Geneva, 2019. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-DAQ-PUB-2019-001>.
- [65] The ATLAS Collaboration. Performance of electron and photon triggers in atlas during lh run 2. *The European Physical Journal C*, 80(1):47, Jan 2020. Available online at <https://doi.org/10.1140/epjc/s10052-019-7500-2>.
- [66] Joao Pequeno. Event Cross Section in a computer generated image of the ATLAS detector. 2008.
- [67] T G Cornelissen, N Van Eldik, M Elsing, W Liebig, E Moyse, N Piacquadio, K Prokofiev, A Salzburger, and A Wildauer. Updates of the ATLAS Tracking Event Data Model (Release 13). Technical report, CERN, Geneva, 2007. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-SOFT-PUB-2007-003>.
- [68] The ATLAS collaboration. A neural network clustering algorithm for the ATLAS silicon pixel detector. *Journal of Instrumentation*, 9(09):P09009, Sep 2014. Available online at <https://dx.doi.org/10.1088/1748-0221/9/09/P09009>.
- [69] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444–450, 1987. Available online at <https://www.sciencedirect.com/science/article/pii/0168900287908874>.
- [70] The ATLAS Collaboration. Performance of the reconstruction of large impact parameter tracks in the ATLAS inner detector. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-014>.
- [71] The ATLAS Collaboration. Performance of the reconstruction of large impact parameter tracks in the inner detector of ATLAS. *The European Physical Journal C*, 83(11):1081, Nov 2023. Available online at <https://doi.org/10.1140/epjc/s10052-023-12024-6>.
- [72] The ATLAS Collaboration. Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC. *The European Physical Journal C*, 77(5):332, May 2017. Available online at <https://doi.org/10.1140/epjc/s10052-017-4887-5>.
- [73] Tim Robertson and J. D. Cryer. An Iterative Procedure for Estimating the Mode. *Journal of the American Statistical Association*, 69(348):1012–1016, 1974.
- [74] The ATLAS Collaboration. Development of ATLAS Primary Vertex Reconstruction for LHC Run 3. Technical report, CERN, Geneva, 2019. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2019-015>.
- [75] The ATLAS Collaboration. Topological cell clustering in the atlas calorimeters and its performance in lh run 1. *The European Physical Journal C*, 77(7):490, Jul 2017. Available online at <https://doi.org/10.1140/epjc/s10052-017-5004-5>.

- [76] The ATLAS Collaboration. Muon reconstruction performance of the atlas detector in proton–proton collision data at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 76(5):292, May 2016. Available online at <https://doi.org/10.1140/epjc/s10052-016-4120-y>.
- [77] The ATLAS Collaboration. Muon reconstruction and identification efficiency in atlas using the full run 2 pp collision data set at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 81(7):578, Jul 2021. Available online at <https://doi.org/10.1140/epjc/s10052-021-09233-2>.
- [78] Richard O. Duda and Peter E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15, Jan 1972. Available online at <https://doi.org/10.1145/361237.361242>.
- [79] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-kt jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, April 2008. Available online at <https://dx.doi.org/10.1088/1126-6708/2008/04/063>.
- [80] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The catchment area of jets. *Journal of High Energy Physics*, 2008(04):005, April 2008. Available online at <https://dx.doi.org/10.1088/1126-6708/2008/04/005>.
- [81] The ATLAS Collaboration. Improving jet substructure performance in ATLAS using Track-CaloClusters. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-015>.
- [82] The ATLAS Collaboration. In situ calibration of large-radius jet energy and mass in 13 tev proton–proton collisions with the atlas detector. *The European Physical Journal C*, 79(2):135, Feb 2019. Available online at <https://doi.org/10.1140/epjc/s10052-019-6632-8>.
- [83] The ATLAS Collaboration. Performance of jet substructure techniques for large- r jets in proton-proton collisions at $\sqrt{s} = 7$ tev using the atlas detector. *Journal of High Energy Physics*, 2013(9):76, Sep 2013. Available online at [https://doi.org/10.1007/JHEP09\(2013\)076](https://doi.org/10.1007/JHEP09(2013)076).
- [84] The ATLAS Collaboration. Boosted Object Tagging with Variable- R Jets in the ATLAS Detector. Technical report, CERN, Geneva, 2016. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-013>.
- [85] The ATLAS Collaboration. Variable Radius, Exclusive- k_T , and Center-of-Mass Subjet Reconstruction for Higgs($\rightarrow b\bar{b}$) Tagging in ATLAS. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-010>.
- [86] The ATLAS Detector. Identification of boosted Higgs bosons decaying into b-quark pairs with the ATLAS detector at 13 TeV. *The European Physical Journal C*, 79(10):836, Oct 2019. Available online at <https://doi.org/10.1140/epjc/s10052-019-7335-x>.
- [87] The ATLAS Collaboration. Variable Radius, Exclusive- k_T , and Center-of-Mass Subjet Reconstruction for Higgs($\rightarrow b\bar{b}$) Tagging in ATLAS. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-010>.
- [88] The ATLAS Collaboration. Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run. Technical report, CERN, Geneva, 2016. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-012>.

- [89] The ATLAS Collaboration. Expected performance of the ATLAS b -tagging algorithms in Run-2. Technical report, CERN, Geneva, 2015. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-022>.
- [90] The ATLAS Collaboration. Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data. Technical report, CERN, Geneva, 2011. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2011-102>.
- [91] The ATLAS Collaboration. Expected performance of the atlas experiment - detector, trigger and physics, 2009.
- [92] The ATLAS Collaboration. Reconstruction of hadronic decay products of tau leptons with the atlas experiment. *The European Physical Journal C*, 76(5):295, May 2016. Available online at <https://doi.org/10.1140/epjc/s10052-016-4110-0>.
- [93] The ATLAS Collaboration. Electron reconstruction and identification in the atlas experiment using the 2015 and 2016 lhc proton–proton collision data at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 79(8):639, Aug 2019. Available online at <https://doi.org/10.1140/epjc/s10052-019-7140-6>.
- [94] The ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006, Dec 2019. Available online at <https://dx.doi.org/10.1088/1748-0221/14/12/P12006>.
- [95] The ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. Technical report, CERN, Geneva, 2012. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-047>.
- [96] The ATLAS Collaboration. Performance of missing transverse momentum reconstruction with the atlas detector using proton–proton collisions at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 78(11):903, Nov 2018. Available online at <https://doi.org/10.1140/epjc/s10052-018-6288-9>.
- [97] The ATLAS Collaboration. E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions. Technical report, CERN, Geneva, 2018. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-023>.
- [98] The ATLAS Collaboration. Object-based missing transverse momentum significance in the ATLAS detector. Technical report, CERN, Geneva, 2018. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-038>.
- [99] The ATLAS Collaboration. ATLAS Software and Computing HL-LHC Roadmap. Technical report, CERN, Geneva, 2022.
- [100] The ATLAS Collaboration. Software and computing for Run 3 of the ATLAS experiment at the LHC. Technical report, CERN, Geneva, 2024. 175 pages in total, author list starting page 158, 42 figures, 5 tables, submitted to EPJC. All figures including auxiliary figures are available at <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/SOFT-2022-02>.
- [101] CERN Courier Volume 53, Number 2, March 2013. 2013. Available online at <https://cds.cern.ch/record/1734919>.

- [102] E Martelli and S Stancu. LHCOPN and LHCONE: Status and Future Evolution. *Journal of Physics: Conference Series*, 664(5):052025, Dec 2015. Available online at <https://dx.doi.org/10.1088/1742-6596/664/5/052025>.
- [103] Martin Barisits, Thomas Beermann, Frank Berghaus, Brian Bockelman, Joaquin Bogado, David Cameron, Dimitrios Christidis, Diego Ciangottini, Gancho Dimitrov, Markus Elsing, Vincent Garonne, Alessandro di Girolamo, Luc Goossens, Wen Guan, Jaroslav Guenther, Tomas Javurek, Dietmar Kuhn, Mario Lassnig, Fernando Lopez, Nicolo Magini, Angelos Molfetas, Armin Nairz, Farid Ould-Saada, Stefan Prenner, Cedric Serfon, Graeme Stewart, Eric Vaandering, Petya Vasileva, Ralph Vigne, and Tobias Wegner. Rucio: Scientific Data Management. *Computing and Software for Big Science*, 3(1):11, Aug 2019. Available online at <https://doi.org/10.1007/s41781-019-0026-3>.
- [104] The ATLAS Collaboration. The ATLAS Collaboration Software and Firmware. Technical report, CERN, Geneva, 2021. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-SOFT-PUB-2021-001>.
- [105] The ATLAS Collaboration. Athena, May 2021. Zenodo, Version 21.0.127. Available online at <https://doi.org/10.5281/zenodo.4772550>.
- [106] Matt Dobbs and Jørgen Beck Hansen. The HepMC C++ Monte Carlo event record for High Energy Physics. *Computer Physics Communications*, 134(1):41–46, 2001. Available online at <https://www.sciencedirect.com/science/article/pii/S0010465500001892>.
- [107] The ATLAS Collaboration. ATLAS data quality operations and performance for 2015–2018 data-taking. *Journal of Instrumentation*, 15(04):P04003, April 2020. Available online at <https://dx.doi.org/10.1088/1748-0221/15/04/P04003>.
- [108] Rene Brun and Fons Rademakers. ROOT — An object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1):81–86, 1997. New Computing Techniques in Physics Research V.
- [109] The ATLAS Collaboration. Athena GitLab Repository, 2024. Available online at <https://gitlab.cern.ch/atlas/athena>.
- [110] Apache Software Foundation. Apache License, Version 2.0, 2004. Available online at <https://www.apache.org/licenses/LICENSE-2.0>.
- [111] Cmake. Available online at <https://cmake.org/download/>.
- [112] G. Barrand, I. Belyaev, P. Binko, M. Cattaneo, R. Chytrcek, G. Corti, M. Frank, G. Gracia, J. Harvey, E. van Herwijnen, P. Maley, P. Mato, S. Probst, and F. Ranjard. GAUDI — A software architecture and framework for building HEP data processing applications. *Computer Physics Communications*, 140(1):45–55, 2001. CHEP2000.
- [113] EOS Open Storage. Available online at <https://eos-web.web.cern.ch/eos-web/>.
- [114] CernVM-FS. Available online at <https://cvmfs.readthedocs.io/en/stable/>.
- [115] Cuhadar Donszelmann, Tulay, Lampl, Walter, and Stewart, Graeme A. ART ATLAS Release Tester using the Grid. *EPJ Web Conf.*, 245:05015, 2020.
- [116] Stefano Frixione, Giovanni Ridolfi, and Paolo Nason. A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *Journal of High Energy Physics*, 2007(09):126, Sep 2007. Available online at <https://dx.doi.org/10.1088/1126-6708/2007/09/126>.
- [117] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *Journal of High Energy Physics*, 2004(11):040, Dec 2004. Available online at <https://dx.doi.org/10.1088/1126-6708/2004/11/040>.

- [118] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007(11):070, Nov 2007. Available online at <https://dx.doi.org/10.1088/1126-6708/2007/11/070>.
- [119] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6):43, Jun 2010. Available online at [https://doi.org/10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043).
- [120] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, Maria Ubiali, and The NNPDF collaboration. Parton distributions for the lhc run ii. *Journal of High Energy Physics*, 2015(4):40, Apr 2015. Available online at [https://doi.org/10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- [121] The ATLAS Collaboration. Studies on top-quark Monte Carlo modelling for Top2016. Technical report, CERN, Geneva, 2016. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-020>.
- [122] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, 2015. Available online at <https://www.sciencedirect.com/science/article/pii/S0010465515000442>.
- [123] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, José I. Latorre, Juan Rojo, and Maria Ubiali. Impact of heavy quark masses on parton distributions and LHC phenomenology. *Nuclear Physics B*, 849(2):296–363, 2011. Available online at <https://www.sciencedirect.com/science/article/pii/S0550321311001842>.
- [124] The ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. Technical report, CERN, Geneva, 2014. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-021>.
- [125] David J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1):152–155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.
- [126] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, 2008. Available online at <https://www.sciencedirect.com/science/article/pii/S0010465508000441>.
- [127] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions with LHC data. *Nuclear Physics B*, 867(2):244–289, 2013. Available online at <https://www.sciencedirect.com/science/article/pii/S0550321312005500>.
- [128] The ATLAS Collaboration. The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model. Technical report, CERN, Geneva, 2016. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-017>.
- [129] The HEPiX Benchmarking Working Group. Results for SL7 x86_64 (GCC 4.8) Benchmark environment. Technical report, CERN, Geneva. <https://w3.hepik.org/benchmarking/sl7-x8664-gcc48.html>.

- [130] Xiaocong Ai, Corentin Allaire, Noemi Calace, Angéla Czirkos, Markus Elsing, Irina Ene, Ralf Farkas, Louis-Guillaume Gagnon, Rocky Garg, Paul Gessinger, Hadrien Grasland, Heather M. Gray, Christian Gumpert, Julia Hrdinka, Benjamin Huth, Moritz Kiehn, Fabian Klimpel, Bernadette Kolbinger, Attila Krasznahorkay, Robert Langenberg, Charles Leggett, Georgiana Mania, Edward Moyse, Joana Niermann, Joseph D. Osborn, David Rousseau, Andreas Salzburger, Bastian Schlag, Lauren Tompkins, Tomohiro Yamazaki, Beomki Yeo, and Jin Zhang. A Common Tracking Software Project. *Computing and Software for Big Science*, 6(1):8, Apr 2022. Available online at <https://doi.org/10.1007/s41781-021-00078-8>.
- [131] The CMS Collaboration. Search for heavy resonances decaying to WW , WZ , or WH boson pairs in a final state consisting of a lepton and a large-radius jet in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 105:032008, Feb 2022. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.105.032008>.
- [132] The ATLAS Collaboration. Search for resonances decaying into a weak vector boson and a Higgs boson in the fully hadronic final state produced in proton – proton collisions at $\sqrt{s} = 13$ TeV with the atlas detector. *Phys. Rev. D*, 102:112008, Dec 2020. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.102.112008>.
- [133] The CMS Collaboration. Search for heavy resonances that decay into a vector boson and a Higgs boson in hadronic final states at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 77(9):636, Sep 2017. Available online at <https://doi.org/10.1140/epjc/s10052-017-5192-z>.
- [134] The CMS Collaboration. Search for a heavy pseudoscalar Higgs boson decaying into a 125 GeV Higgs boson and a Z boson in final states with two tau and two light leptons at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2020(3):65, Mar 2020. Available online at [https://doi.org/10.1007/JHEP03\(2020\)065](https://doi.org/10.1007/JHEP03(2020)065).
- [135] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7):79, Jul 2014. Available online at [https://doi.org/10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- [136] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke. Interpolation between multi-dimensional histograms using a new non-linear moment morphing method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 771:39–48, 2015. Available online at <https://www.sciencedirect.com/science/article/pii/S0168900214011814>.
- [137] Lorenzo Moneta, Kevin Belasco, Kyle Cranmer, Sven Kreiss, Alfio Lazzaro, Danilo Piparo, Gregory Schott, Wouter Verkerke, and Matthias Wolf. The RooStats Project, 2011.
- [138] Enrico Bothmann, Gurpreet Singh Chahal, Stefan Höche, Johannes Krause, Frank Krauss, Silvan Kuttimalai, Sebastian Liebschner, Davide Napoletano, Marek Schönherr, Holger Schulz, Steffen Schumann, and Frank Siegert. Event generation with Sherpa 2.2. *SciPost Phys.*, 7:034, 2019. Available online at <https://scipost.org/10.21468/SciPostPhys.7.3.034>.
- [139] Tanju Gleisberg and Stefan Höche. Comix, a new matrix element generator. *Journal of High Energy Physics*, 2008(12):039, Dec 2008. Available online at <https://dx.doi.org/10.1088/1126-6708/2008/12/039>.
- [140] Federico Buccioni, Jean-Nicolas Lang, Jonas M. Lindert, Philipp Maierhöfer, Stefano Pozzorini, Hantian Zhang, and Max F. Zoller. OpenLoops 2. *The European Physical Journal C*, 79(10):866, Oct 2019. Available online at <https://doi.org/10.1140/epjc/s10052-019-7306-2>.
- [141] F. Cascioli, P. Maierhöfer, and S. Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, Mar 2012. Available online at <https://link.aps.org/doi/10.1103/PhysRevLett.108.111601>.

- [142] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. Collier: A fortran-based complex one-loop library in extended regularizations. *Computer Physics Communications*, 212:220–238, 2017. Available online at <https://www.sciencedirect.com/science/article/pii/S0010465516303320>.
- [143] S. Schumann and F. Krauss. A parton shower algorithm based on Catani-Seymour dipole factorisation. *Journal of High Energy Physics*, 2008(03):038, mar 2008. Available online at <https://dx.doi.org/10.1088/1126-6708/2008/03/038>.
- [144] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. A critical appraisal of NLO+PS matching methods. *Journal of High Energy Physics*, 2012(9):49, Sep 2012. Available online at [https://doi.org/10.1007/JHEP09\(2012\)049](https://doi.org/10.1007/JHEP09(2012)049).
- [145] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. QCD matrix elements + parton showers. The NLO case. *Journal of High Energy Physics*, 2013(4):27, Apr 2013. Available online at [https://doi.org/10.1007/JHEP04\(2013\)027](https://doi.org/10.1007/JHEP04(2013)027).
- [146] Stefano Catani, Frank Krauss, Bryan R. Webber, and Ralf Kuhn. QCD Matrix Elements + Parton Showers. *Journal of High Energy Physics*, 2001(11):063, Jan 2002. Available online at <https://dx.doi.org/10.1088/1126-6708/2001/11/063>.
- [147] Stefan Höche, Frank Krauss, Steffen Schumann, and Frank Siegert. QCD matrix elements and truncated showers. *Journal of High Energy Physics*, 2009(05):053, may 2009. Available online at <https://dx.doi.org/10.1088/1126-6708/2009/05/053>.
- [148] Charalampos Anastasiou, Lance Dixon, Kirill Melnikov, and Frank Petriello. High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order. *Phys. Rev. D*, 69:094008, May 2004. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.69.094008>.
- [149] M. Beneke, P. Falgari, S. Klein, and C. Schwinn. Hadronic top-quark pair production with NNLL threshold resummation. *Nuclear Physics B*, 855(3):695–741, 2012. Available online at <https://www.sciencedirect.com/science/article/pii/S0550321311005803>.
- [150] Matteo Cacciari, Michał Czakon, Michelangelo Mangano, Alexander Mitov, and Paolo Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Physics Letters B*, 710(4):612–622, 2012. Available online at <https://www.sciencedirect.com/science/article/pii/S0370269312002766>.
- [151] Peter Bärnreuther, Michał Czakon, and Alexander Mitov. Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t}+X$. *Phys. Rev. Lett.*, 109:132001, Sep 2012. Available online at <https://link.aps.org/doi/10.1103/PhysRevLett.109.132001>.
- [152] Michał Czakon and Alexander Mitov. NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction. *Journal of High Energy Physics*, 2013(1):80, Jan 2013. Available online at [https://doi.org/10.1007/JHEP01\(2013\)080](https://doi.org/10.1007/JHEP01(2013)080).
- [153] Michał Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$. *Phys. Rev. Lett.*, 110:252004, Jun 2013. Available online at <https://link.aps.org/doi/10.1103/PhysRevLett.110.252004>.
- [154] Michał Czakon and Alexander Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Computer Physics Communications*, 185(11):2930–2938, 2014. Available online at <https://www.sciencedirect.com/science/article/pii/S0010465514002264>.
- [155] Stefano Frixione, Eric Laenen, Patrick Motylinski, Chris White, and Bryan R. Webber. Single-top hadroproduction in association with a W boson. *Journal of High Energy Physics*, 2008(07):029, July 2008. Available online at <https://dx.doi.org/10.1088/1126-6708/2008/07/029>.

- [156] The ATLAS Collaboration. Multi-Boson Simulation for 13 TeV ATLAS Analyses. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-005>.
- [157] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, Albert De Roeck, Joël Feltesse, Stefano Forte, Jun Gao, Sasha Glazov, Joey Huston, Zahari Kassabov, Ronan McNulty, Andreas Morsch, Pavel Nadolsky, Voica Radescu, Juan Rojo, and Robert Thorne. PDF4LHC recommendations for LHC Run II. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, Jan 2016. Available online at <https://dx.doi.org/10.1088/0954-3899/43/2/023001>.
- [158] The ATLAS Collaboration. Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2014(9):145, Sep 2014. Available online at [https://doi.org/10.1007/JHEP09\(2014\)145](https://doi.org/10.1007/JHEP09(2014)145).
- [159] The ATLAS Collaboration. The ATLAS Simulation Infrastructure. *The European Physical Journal C*, 70(3):823–874, Dec 2010. Available online at <https://doi.org/10.1140/epjc/s10052-010-1429-9>.
- [160] The ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 81(7):578, Jul 2021. Available online at <https://doi.org/10.1140/epjc/s10052-021-09233-2>.
- [161] Atlas b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 79(11):970, Nov 2019. Available online at <https://doi.org/10.1140/epjc/s10052-019-7450-8>.
- [162] W Lukas. Fast Simulation for ATLAS: Atlfast-II and ISF. Technical report, CERN, Geneva, 2012.
- [163] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *The European Physical Journal C*, 72(3):1896, Mar 2012. Available online at <https://doi.org/10.1140/epjc/s10052-012-1896-2>.
- [164] The ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *The European Physical Journal C*, 76(11):581, Oct 2016. Available online at <https://doi.org/10.1140/epjc/s10052-016-4395-z>.
- [165] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet trimming. *Journal of High Energy Physics*, 2010(2):84, Feb 2010. Available online at [https://doi.org/10.1007/JHEP02\(2010\)084](https://doi.org/10.1007/JHEP02(2010)084).
- [166] The ATLAS Collaboration. Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run. Technical report, CERN, Geneva, 2017. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-013>.
- [167] The ATLAS Collaboration. Measurement of b -tagging Efficiency of c -jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector. Technical report, CERN, Geneva, 2018. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-001>.
- [168] The ATLAS Collaboration. Calibration of light-flavour b -jet mistagging rates using ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2018. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-006>.
- [169] The ATLAS Collaboration. Measurements of wh and zh production in the $h \rightarrow b\bar{b}$ decay channel in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector. *The European Physical Journal C*, 81(2):178, Feb 2021. Available online at <https://doi.org/10.1140/epjc/s10052-020-08677-2>.

- [170] The ATLAS Collaboration. Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector. *Physics Letters B*, 786:59–86, 2018. Available online at <https://www.sciencedirect.com/science/article/pii/S0370269318307056>.
- [171] The ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the atlas detector. *Phys. Rev. D*, 96:072002, Oct 2017. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.96.072002>.
- [172] The ATLAS Collaboration. Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector. *The European Physical Journal C*, 73(3):2306, Mar 2013. Available online at <https://doi.org/10.1140/epjc/s10052-013-2306-0>.
- [173] The ATLAS Collaboration. Jet mass reconstruction with the ATLAS Detector in early Run 2 data. Technical report, CERN, Geneva, 2016. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2016-035>.
- [174] The ATLAS Collaboration. Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS. Technical report, CERN, Geneva, 2015. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-033>.
- [175] The ATLAS Collaboration. Jet mass reconstruction with the ATLAS Detector in early Run 2 data. Technical report, CERN, Geneva, 2016. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2016-035>.
- [176] The ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical report, CERN, Geneva, 2019. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-021>.
- [177] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the lhc. *The European Physical Journal C*, 63(2):189–285, Sep 2009. Available online at <https://doi.org/10.1140/epjc/s10052-009-1072-5>.
- [178] Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, Zhao Li, Pavel Nadolsky, Jon Pumplin, Daniel Stump, and C.-P. Yuan. CT10 next-to-next-to-leading order global analysis of QCD. *Phys. Rev. D*, 89:033009, Feb 2014. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.89.033009>.
- [179] Johannes Bellm, Stefan Gieseke, David Grellscheid, Simon Plätzer, Michael Rauch, Christian Reuschle, Peter Richardson, Peter Schichtel, Michael H. Seymour, Andrzej Siódmok, Alexandra Wilcock, Nadine Fischer, Marco A. Harrendorf, Graeme Nail, Andreas Papaefstathiou, and Daniel Rauch. Herwig 7.0/herwig++ 3.0 release note. *The European Physical Journal C*, 76(4):196, Apr 2016. Available online at <https://doi.org/10.1140/epjc/s10052-016-4018-8>.
- [180] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, Albert De Roeck, Joël Feltesse, Stefano Forte, Jun Gao, Sasha Glazov, Joey Huston, Zahari Kassabov, Ronan McNulty, Andreas Morsch, Pavel Nadolsky, Voica Radescu, Juan Rojo, and Robert Thorne. PDF4LHC recommendations for LHC Run II. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, Jan 2016. Available online at <https://dx.doi.org/10.1088/0954-3899/43/2/023001>.
- [181] Leif Lönnblad. Correcting the Colour-Dipole Cascade Model with Fixed Order Matrix Elements. *Journal of High Energy Physics*, 2002(05):046, Jun 2002. Available online at <https://dx.doi.org/10.1088/1126-6708/2002/05/046>.
- [182] Leif Lönnblad and Stefan Prestel. Matching tree-level matrix elements with interleaved showers. *Journal of High Energy Physics*, 2012(3):19, Mar 2012. Available online at [https://doi.org/10.1007/JHEP03\(2012\)019](https://doi.org/10.1007/JHEP03(2012)019).

- [183] Iain W. Stewart and Frank J. Tackmann. Theory uncertainties for Higgs mass and other searches using jet bins. *Phys. Rev. D*, 85:034011, Feb 2012. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.85.034011>.
- [184] Kyle Cranmer. Kernel estimation in high-energy physics. *Computer Physics Communications*, 136(3):198–207, 2001. Available online at <https://www.sciencedirect.com/science/article/pii/S0010465500002435>.
- [185] Wouter Verkerke and David Kirkby. The roofit toolkit for data modeling, 2003.
- [186] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1554, Feb 2011. Available online at <https://doi.org/10.1140/epjc/s10052-011-1554-0>.
- [187] Abraham Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54:426–482, 1943. Available online at <https://api.semanticscholar.org/CorpusID:54174575>.
- [188] A L Read. Presentation of search results: the CL_s technique. *J. Phys. G*, 28(10):2693–704, 2002.
- [189] The ATLAS Collaboration. Combination of searches for heavy resonances decaying into bosonic and leptonic final states using 36 fb^{−1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the atlas detector. *Phys. Rev. D*, 98:052008, Sep 2018. Available online at <https://link.aps.org/doi/10.1103/PhysRevD.98.052008>.
- [190] The ATLAS Collaboration. Combination of searches for heavy spin-1 resonances using 139 fb^{−1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2024(4):118, Apr 2024. Available online at [https://doi.org/10.1007/JHEP04\(2024\)118](https://doi.org/10.1007/JHEP04(2024)118).
- [191] A. D. Bukin. Fitting function for asymmetric peaks, 2007. Available online at <https://doi.org/10.48550/arXiv.0711.4449>.
- [192] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke. Interpolation between multi-dimensional histograms using a new non-linear moment morphing method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 771:39–48, 2015. Available online at <https://www.sciencedirect.com/science/article/pii/S0168900214011814>.