

Semi-Supervised Learning for Semi-VISIBLE Jets: A Search for Dark Matter Jets at the LHC with  
the ATLAS Detector

Elena Laura Busch

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy  
under the Executive Committee  
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

© 2024

Elena Laura Busch

All Rights Reserved

## Abstract

Semi-Supervised Learning for Semi-Visible Jets: A Search for Dark Matter Jets at the LHC with  
the ATLAS Detector

Elena Laura Busch

A search is presented for hadronic signatures of a strongly-coupled hidden dark sector, accessed via resonant production of a  $Z'$  mediator. The analysis uses  $139 \text{ fb}^{-1}$  of proton-proton collision data collected by the ATLAS experiment during Run 2 of the LHC. The  $Z'$  mediator decays to two dark quarks, which each hadronize and decay to showers containing both dark and Standard Model particles; these showers are termed “semi-visible” jets. The final state expects missing energy aligned with one of the jets, a topology that is ignored by most dark matter searches. A supervised machine learning method is used to select these dark showers and reject the dominant background of mis-measured multijet events. A complementary semi-supervised anomaly detection approach introduces broad sensitivity to a variety of strongly coupled dark matter models. A resonance search is performed by fitting the transverse mass spectrum with a polynomial background estimation function. Results are presented as limits on the effective cross section of the  $Z'$ , parameterized by the fraction of invisible particles in the decay and the  $Z'$  mass. No structure in the transverse mass spectrum compatible with the signal hypothesis is observed.  $Z'$  mediator masses from ranging from 2.0 TeV to 3.5 TeV are excluded at the 95% confidence level.

## Table of Contents

Acknowledgments . . . . . xvii

### I Theory 1

Chapter 1: The Standard Model . . . . . 2

1.1 Phenomenology: Particles and Forces . . . . .	2
1.1.1 Particles . . . . .	2
1.1.2 Forces . . . . .	4
1.2 QCD and Jets . . . . .	6
1.3 Symmetries . . . . .	7
1.3.1 Spontaneous Symmetry Breaking and The Higgs Mechanism . . . . .	8
1.4 Experimental Validation of the Standard Model . . . . .	9
1.5 Limitations of the Standard Model . . . . .	10

Chapter 2: Physics Beyond the Standard Model . . . . . 12

2.1 Hidden Valley Models . . . . .	12
2.2 Dark QCD . . . . .	13
2.3 Semi-visible Jets . . . . .	14

### II Experiment 16

Chapter 3: The Large Hadron Collider . . . . . 17

3.1 Accelerator Physics . . . . .	18
3.1.1 The Journey of a Proton . . . . .	18
3.1.2 Magnets . . . . .	19
3.2 Luminosity . . . . .	20
3.3 LHC Timeline . . . . .	23
 Chapter 4: The ATLAS Detector . . . . .	25
4.1 Coordinate System and Geometry . . . . .	26
4.2 Inner Detector . . . . .	27
4.2.1 Pixel Detector . . . . .	28
4.2.2 Semiconductor Tracker . . . . .	30
4.2.3 Transition Radiation Tracker . . . . .	30
4.3 Calorimeters . . . . .	30
4.3.1 Liquid Argon Calorimeter . . . . .	31
4.3.2 Tile Calorimeter . . . . .	34
4.4 Muon Spectrometer . . . . .	35
4.5 Magnet System . . . . .	37
4.6 Forward Detectors . . . . .	38
4.7 Trigger and Data Acquisition . . . . .	39
 Chapter 5: Particle Reconstruction and Identification . . . . .	43
5.1 Inner Detector Tracks . . . . .	43
5.2 Photons and Electrons . . . . .	45
5.3 Muons . . . . .	47

5.4	Jets . . . . .	49
5.4.1	Calorimeter Clusters . . . . .	50
5.4.2	Particle Flow Algorithm . . . . .	51
5.4.3	Jet Clustering . . . . .	52
5.4.4	Ghost Track Association . . . . .	55
5.5	Missing Transverse Energy . . . . .	56
<b>III</b>	<b>Search</b>	<b>58</b>
Chapter 6:	Monte Carlo and Data . . . . .	59
6.1	Data . . . . .	59
6.2	Simulation . . . . .	61
6.2.1	Simulated Backgrounds . . . . .	61
6.2.2	Signal Simulation . . . . .	62
Chapter 7:	Machine Learning Tools . . . . .	65
7.1	Introduction . . . . .	65
7.1.1	Machine Learning Fundamentals . . . . .	66
7.2	Particle Flow Network (Supervised) . . . . .	69
7.2.1	Architecture Fundamentals . . . . .	69
7.2.2	Input Modeling, Scaling, and Rotation . . . . .	70
7.2.3	Training . . . . .	77
7.2.4	Performance . . . . .	78
7.3	ANTELOPE (Semi-supervised) . . . . .	82
7.3.1	Architecture Fundamentals . . . . .	82

7.3.2	Training . . . . .	85
7.3.3	Performance . . . . .	86
Chapter 8:	Analysis Strategy . . . . .	89
8.1	Transverse Mass . . . . .	89
8.2	Preselection . . . . .	91
8.3	SVJ Fit and Discovery Analysis Strategies . . . . .	96
8.4	Analysis Regions . . . . .	97
8.4.1	Control and Validation Regions . . . . .	97
8.4.2	Signal Region . . . . .	100
8.5	Background Estimation . . . . .	100
8.6	Fit Strategy and Validation . . . . .	102
8.6.1	SVJ Fit Strategy . . . . .	102
8.6.2	Discovery Strategy . . . . .	112
Chapter 9:	Results . . . . .	118
9.1	SVJ Fit Result . . . . .	118
9.1.1	Systematics . . . . .	119
9.1.2	Interpretation . . . . .	124
9.2	Discovery Result . . . . .	126
Conclusion . . . . .		127
References . . . . .		128

Appendix A: Trigger Studies . . . . .	135
Appendix B: Machine Learning Approaches . . . . .	139
B.1 Unsupervised: AE vs. ANTELOPE . . . . .	139
B.1.1 Signal Contamination . . . . .	140
B.2 PFN Optimality Checks . . . . .	140
B.3 Single Jet vs Jet System ML Approach . . . . .	141
B.4 PFN Training Composition . . . . .	143
Appendix C: $E_T^{miss}$ and $E_T^{miss}\phi$ Shapes . . . . .	145
C.1 Tight Cleaning . . . . .	145
C.2 NCB Preselection . . . . .	147
C.3 TileCal Correction . . . . .	147
Appendix D: Truth Studies . . . . .	150
D.1 Jet dR Matching . . . . .	150
Appendix E: MC Fitting . . . . .	152
Appendix F: BumpHunter . . . . .	153
F.1 Signal Mass Resolution $m_T$ Binning . . . . .	153

## List of Figures

1.1	Diagram of the 17 particles comprising the Standard Model . . . . .	3
1.2	Fundamental particle interactions of the three fundamental forces described by the Standard Model [2]. . . . .	5
1.3	An example Feynmann diagram of jet production . . . . .	6
1.4	An illustration of the “hat shaped” potential of the Higgs field, resulting in a non-zero vacuum expectation value. . . . .	8
2.1	Illustration of the hidden valley potential. . . . .	13
2.2	The massive mediator particle $Z'$ of the s-channel realization of a HV model . . . .	13
3.1	The LHC accelerator complex at CERN [29] . . . . .	19
3.2	The octants of the LHC and location of various beam activities [28]. Stars indicate the locations of beam collisions, and the associated detectors recording the outcome of those collisions. . . . .	20
3.3	(Left) Total integrated luminosity over the course of Run 2. (Right) Average number of $p p$ interactions per bunch crossing in Run 2. Each curve is weighted by the integrated luminosity for the year. . . . .	22
3.4	Timeline of LHC and HL-LHC activities [31]. Integrated luminosity estimates are approximate, and not reflective of the exact amount delivered to each experiment. .	24
4.1	The ATLAS detector [34]. . . . .	25
4.2	ATLAS coordinate system and geometry . . . . .	28
4.3	A 3D visualization of the structure of the ID in the barrel region [35] . . . . .	29

4.4	A 3D visualization of the structure of the ID end-caps [34]. The paths of two charged particles ( $\eta = 1.4$ and $2.2$ ) are illustrated. . . . .	29
4.5	ATLAS calorimetry system [36] . . . . .	31
4.6	Diagram of a segment of the EMB, demonstrating the accordion plate arrangement and the cell segmentation pattern [37]. Three layers are illustrated, and the variation in the cell granularity between the front layer (Sampling Layer 1) and the back layer (Sampling Layer 3) is shown. . . . .	32
4.7	A LAr pulse as produced in the detector (triangle) and after shaping (curve) [37] . .	33
4.8	Readout gap structure in HEC [37] . . . . .	34
4.9	TileCal wedge module [40] . . . . .	35
4.10	Cross section view of the muon spectrometer system [41] . . . . .	36
4.11	Layout of the barrel and end-cap toroid magnets [34] . . . . .	38
4.12	Block diagram of the L1 trigger process [34]. The overall L1 trigger decision is made by the CTP. . . . .	40
4.13	Block diagram of the trigger and data acquisition processes, including L1 and HLT processes [45]. . . . .	41
5.1	This graphic illustrates a slice of ATLAS detector barrel [46]. A photon, electron, and two jets (associated to the proton and neutron) are shown, illustrating each object's interaction with various ATLAS subsystems. The path of the charged particles is curved on account of the strong magnetic field produced by the solenoid and toroid magnets. A neutrino, generally representing the reconstruction of missing transverse energy $E_T^{\text{miss}}$ , is also illustrated. . . . .	44
5.2	Track reconstruction seeding, finding and fitting illustration [47] . . . . .	45
5.3	Three types of EM object candidates [49]. . . . .	46
5.4	The four types of muon track candidates [51]. The red bands indicate where detector measurements for each muon track candidate are made. The dashed lines indicate the path of the muon through detectors where no measurement is made. Standalone muon is another term for an extrapolated muon. . . . .	49
5.5	The fragmentation and hadronization processes undergone by a quark produced in a proton-proton collision [53]. . . . .	50

5.6	A flow chart illustrating the particle flow algorithm progression [56]. The solid lines indicate the progression of tracks through the algorithm, while the dotted lines indicate the progression of clusters. The process begins with track selection and continues until the energy associated with the tracks has been removed from the calorimeter. At the end charged particle tracks, unmodified topo-clusters, and the remnants of topo-clusters which have had part of their energy removed remain.	53
5.7	A comparison of jet clustering with four different jet algorithms. The anti- $k_t$ algorithm is observed to create the most conical jets, where the shape of the jet is immune to the presence of soft radiation [57].	54
5.8	A comparison of MC simulation and data for $Z \rightarrow \mu\mu$ events where real $E_T^{\text{miss}} = 0$ [65]. The resolution of the missing energy in the transverse ( $x - y$ ) plane is observed to increase with increasing total $\sum E_T$ .	57
6.1	Integrated luminosity for the ATLAS experiment as a function of time during Run 2 [66]	60
6.2	The transverse momentum slices of the QCD MC simulation, overlayed to show how they come together to create a smooth distribution (left) once weighted properly. The original unweighted distribution (right) illustrates the enhanced statistics for the high $p_T$ range.	62
6.3	Background processes relevant to the SVJ signal.	63
7.1	A diagram of a deep neural network architecture.	66
7.2	An example score distribution for a binary classifier. A higher score indicates a greater probability of the event being signal-like. Most signal events (orange) receive a high score while most background events (blue) receive a low score, indicating good classification.	67
7.3	Several example ROC curves. The AUC is also illustrated [76].	68
7.4	A diagram of auto-encoder architecture. The loss is computed as a difference (often the <i>mean squared error</i> or MSE) between the input $x$ and the output $y$ [78].	69
7.5	The Energy/Particle Flow Network concept, from Ref. [79]. The physics input information is represented as arrows on the left, for an arbitrary number of particles. The $\Phi$ transformation converts these arrows to 3 graphs, indicating the $\Phi$ basis dimension $l$ is 3 in this example. The graphs are then summed for all particles to create $O$ , or the event representation.	71

7.6 An annotated diagram of the PFN architecture [79]. $y$ and $\phi$ represent geometric trajectory information for the input particles, $z$ represents energy information, and PID encompasses any other particle ID information in the input. PID is presented in the diagram as a 1-dimensional input, but could represent multiple input dimensions. . . . .	72
7.7 An illustration of the expected dijet behavior of semi-visible jets, where one jet is closely aligned with $E_T^{\text{miss}}(\text{MET})$ . In the figure two jet cones $j_1$ and $j_2$ are illustrated, along with their associated momentum vectors $\vec{p}_1$ and $\vec{p}_2$ . . . . .	72
7.8 Illustration of track coordinates $d_0$ and $z_0$ . . . . .	73
7.9 A diagram demonstrating how the two jet system is rotated in $(\eta, \phi)$ . The jet cones and associated jet tracks are illustrated. The dashed tracks represent dark hadrons while the solid tracks represent SM hadrons. The system average $(\bar{\eta}, \bar{\phi})$ is shown in red and an example track with coordinates $(\eta_i, \phi_i)$ is shown in purple. . . . .	74
7.10 The 6 PFN track variables in background MC (blue) and data (orange), after the scaling and rotation procedure is applied. There is excellent modeling of the data by the MC within the track variables. The slight discrepancy in the $\phi$ distribution due to the inaccuracies of modeling dead TileCal cells in the QCD MC is considered. The level of discrepancy is determined to be within tolerance given that the final result will be data driven and the QCD model is used in the PFN training only. . . . .	75
7.11 The 6 PFN track variables in background MC (blue) and signal MC (orange) before scaling and rotation. The track kinematics are largely similar, and the variation in the $\phi$ distribution is explained in the text. . . . .	76
7.12 The 6 PFN track variables in background MC (blue) and signal MC (orange) after scaling and rotation. The $\phi$ distribution is modified by the rotation procedure, as explained in the text. . . . .	76
7.13 PFN score for full-background MC (black), data (red), and 2 representative signal points (green). The left plot is from a QCD-only training, while the right plot is from a full-background training. The histograms have been normalized to visualize the shapes better - the actual number of plotted events is shown in the legend. In the left plot we observe that both signal points are strongly classified as signal-like. In the right plot we observe less background contamination in the high score region, but worse signal classification. Both PFN trainings were tested for their effect on the analysis sensitivity and the QCD-only training was found to be favorable. . . . .	78
7.14 PFN architecture loss during training as a function of epoch (left) and the evaluated score for signal and background training samples (right). The loss vs. epoch plot shows that the network is not overtrained. The score plot shows a good separation between signal and background. . . . .	79

7.15 ROC for the PFN, using SVJ signal events (true positive) and data (false positive).	80
7.16 AUC for the PFN, shown for each signal in the SVJ grid.	80
7.17 Illustration of the PFN score selection, showing the separation between data (black) and 4 signal points (blue and green). The legend information takes the form “ $m_Z$ , $R_{inv}$ ” for the signal. The PFN score selection value is shown by the pink line. Only events with a score $> 0.6$ will be accepted for use in the analysis. We see that most background (data) is rejected, while most signal is accepted.	81
7.18 An annotated diagram of the ANTELOPE architecture. Step 1 illustrates the PFN which is fully trained before its use in the ANTELOPE network. Step 2 illustrates the variational auto-encoder. The Gaussian sampling of the latent space is shown, illustrating how the VAE differs from the AE shown in Figure 7.4.	83
7.19 A visual representation of the 64 PFN $O$ which create the input for the VAE component of ANTELOPE. The plot is 2D histogram of the PFN $O$ index (0-63) versus the value assumed by that index. Many entries have a $O$ value of exactly 0.0. To visually separate these from entries with a small but non-zero $O$ value, any entries with value = 0.0 are moved to value = -0.01 (leftmost column) for the purpose of the plot only.	84
7.20 ANTELOPE architecture loss during training as a function of epoch.	85
7.21 Anomaly score distribution (left), comparing all data (orange) and all SVJ signals (blue). The signals have a small but consistently higher score than the data, indicating that they are tagged as more anomalous by ANTELOPE. A ROC curve for an example signal point is also shown (right).	86
7.22 AUC from the ANTELOPE score for each signal in the SVJ grid.	87
7.23 Comparing data and the alternate signal models for the PFN score (left) and ANTELOPE score (right). The emerging jet signal (dark blue) and gluino R-hadron signals (red, light green) are an example of the advantage of the model-independent ANTELOPE approach. These signals have a bimodal shape in PFN score but are clearly tagged with a high anomaly score by the ANTELOPE.	88

8.1	The resonant shape of the SVJ signals (color) in $m_T$ , in contrast to the smoothly falling $m_T$ background (grey). The top row illustrates unit normal shapes, so that the shape of the signals is more easily seen. The bottom row illustrates the signal and background scaled to their expected yield at preselection, illustrating the relative expected statistics. The $R_{inv} = 0.8$ signals (right) boast a wider shape, making them more difficult to detect, while the $R_{inv} = 0.2$ signals (left) produce a more narrow resonance in $m_T$ . The signal models are identified in the legend as “ $m_Z'$ , $R_{inv}$ ”.	90
8.2	Energy and momentum analysis variables at preselection, for data (black), background MC (grey), and representative signal models (color). The signal models are identified in the legend as “ $m_Z'$ , $R_{inv}$ ”.	94
8.3	Orientation analysis variables at preselection, for data (black), background MC (grey), and representative signal models (color). The signal models are identified in the legend as “ $m_Z'$ , $R_{inv}$ ”.	95
8.4	The two fitting strategies. The SVJ Fit (left) illustrates how SVJ signal shapes will be considered in the fit to search for SVJ specific signal shapes, where “s+b fit” indicates a fit that considers the shape of the signal. The Discovery Fit (right) illustrates how the data is compared to a background-only hypothesis to search for any kind of $m_T$ bump, where “b fit” indicates a background-only fit with no signal hypothesis.	97
8.5	Recall the construction of anti- $k_t$ jets as described in Section 5.4.3 and illustrated in Figure 5.7. On the right, we zoom in on two jets, illustrating the narrow cluster pattern in the green jet, and the wide cluster pattern in the yellow jet.	98
8.6	Distributions of the subleading jet width $\text{width}_{j2}$ (left) and leading jet width $\text{width}_{j1}$ (right) in data, background MC and signals at preselection. All SVJ signals are seen to be wider than the background in $\text{width}_{j2}$ . The same is not true for $\text{width}_{j1}$ , where some signals are observed to closely match the background.	98
8.7	$m_T$ in simulation across the CR, VR, and SR for both PFN (left) and ANTELOPE (right) selections. While there is variation in the slope of the distribution, no sculpting of bumps is observed.	99
8.8	Distribution of data events amongst the CR, VR, and SR regions, along with the fractional population of each region. The SVJ Fit region is shown left with the PFN score on the x-axis, and Discovery region is shown right, with the ANTELOPE score on the x-axis.	100

8.9 Flow of analysis selections and fitting strategy. From preselection, events with $\text{Jet2Width} < 0.05$ are set aside for the CR. Events with $\text{Jet2Width} \geq 0.05$ are split according the ML score. Events with low ML score create the VR, while events with high ML score create the SR. Events with high PFN score are fitted to determine if they are compatible with the SVJ signal shape. Events with high ANTELOPE score are fitted for a background estimation, and a search for any general data bump is performed. . . . .	101
8.10 Background-only $m_T$ fits using data in the full statistics CR and VR regions. The fit is observed to converge with $p\text{-value} > 0.05$ . The distribution of residuals is reasonably flat. The number of events in the data histogram, $p\text{-value}$ and $\chi^2$ value (x) are reported in the legend. . . . .	104
8.11 $m_T$ distribution in the data CR, before (left) and after (right) smoothing. . . . .	105
8.12 Background-only $m_T$ fits using pseudo-data from the CR template. All three fits are seen to successfully converge, with varying $p$ -values. The distribution of residuals is reasonably flat for all three fits. . . . .	106
8.13 $p$ -value histograms from 100 fits to Asimov data in the CR. The even distribution of $p$ -values between 0 and 1 indicates that the behavior of the fit is healthy. 98 $p$ -values are shown, as two are excluded due to fits that did not converge on the first try. These fits later converged after the initial parameters were adjusted. . . . .	106
8.14 Example S+B fit on a background $m_T$ spectrum with injected signal from the point ( $4000 \text{ GeV}$ , $R_{inv} = 0.2$ ). The shape of the injected signal can be seen in Figure 8.1. The ability of the s+b fit to capture the shape of the signal and accurately measure the amount of injected signal is observed. . . . .	107
8.15 Measured signal at a variety of injected values (1x, 2x, and 5x $\sqrt{b}$ ), for all signal points in the grid, $R_{inv} = 0.2$ (top left), 0.4 (top right), 0.6 (bottom left), and 0.8 (bottom right). The x-axis values are slightly shifted from their true value so that all points can be viewed simultaneously. The error bars indicate the standard deviation of the number of fitted events across the 50 Asimov experiments. While the errors are large for some points, the strong linear relationship of the means, illustrated by the dashed lines, is the key feature. . . . .	108
8.16 95% C.L. upper limits on the $Z' \rightarrow q_D q_{\bar{D}}$ process cross section, derived from the $m_T$ spectrum in the CR. The red line indicates the theoretical cross section, while the blue line indicates the observed 95% C.L. upper limit on the cross section given the data spectrum. The black line indicates the expected limit given the background shape provided by the fit. The green and yellow bands indicate the uncertainty bands. All signal models across $Z'$ mass and four different $R_{inv}$ fractions are shown. . . . .	110

8.17	95% C.L. observed limit for signal models across $Z'$ mass, with varying amounts of signal injected. The increasing observed limit indicates the desired behavior.	111
8.18	Post-fit function and residuals for the ANTELOPE CR and VR.	112
8.19	BumpHunter fits on the ANTELOPE $m_T$ spectra for both the CR (left) and VR (right). In a signal-depleted region, good agreement with the background estimation is observed.	114
8.20	BumpHunter p-values extracted for 100 Asimov toys for both the ANTELOPE CR (left) and VR (right). Fits where the polynomial fit initially failed to converge are excluded from the plot. These fits were later recovered through modification of the initial parameters.	115
8.21	Example injected gaussian signal, with mean = 3000 GeV and width of 20%.	115
8.22	Response of the BumpHunter framework to gaussian signal injection is shown. The local significance (top plot) and bump location in GeV (bottom plot) are shown. With the exception of the 5.0 TeV 20% width signal (expressed as 0.2), the BH identifies bumps with a significance $> 2.0\sigma$ in approximately the correct location.	116
8.23	Example BH response to gaussian signal injection at 4000 GeV with width of 10%.	117
9.1	$m_T$ in the unblinded SVJ Fit SR with a background-only fit (p-value = 0.265).	118
9.2	Spurious signal metric as a function of resonance mass. The requirement $S_{\text{spur}}/\sigma_{\text{fit}} < 0.5$ is satisfied for all signal points. 100 pseudo-data experiments are used for the measurement.	120
9.3	$m_T$ of the 3500 GeV $Z'$ , $R_{\text{inv}} = 0.2$ signal point, shown with an example JES noise uncertainty variation. The nominal shape, $1\sigma$ up, and $1\sigma$ down variations are shown. The variation is seen to have minimal impact on the signal shape. Signal only (no background) is shown.	122
9.4	$m_T$ of the 3500 GeV $Z'$ , $R_{\text{inv}} = 0.2$ signal point, shown with the sum in quadrature of all JES and JER variations. The nominal shape before systematic variations, the maximal $1\sigma$ “up”, and maximal $1\sigma$ “down” variations are shown. Even summed in quadrature, the effect of the JES and JER variations on the shape of the signal is seen to be small.	123
9.5	Signal distribution of $m_T$ , varying the ISR (Var1), FSR (Var2) and MPI (Var3a-c) configurations.	123

9.6	Expected and observed 95% C.L. limits in the unblinded SR, as a function of $Z'$ mass for $R_{inv} = 0.2$ (top left), 0.4 (top right), 0.6 (bottom left), 0.8 (bottom right). The red line indicates the theoretical cross section, while the blue line indicates the observed 95% C.L. upper limit on the cross section given the data spectrum. The black line indicates the expected limit given the background shape provided by the fit. The green and yellow bands indicate the uncertainty bands. . . . .	125
9.7	$m_T$ in the unblinded ANTELOPE SR with a background-only fit ( $p\text{-value} = 0.74$ ), left. BumpHunter test selecting the most significant data excess with a $p\text{-value}$ of 0.8098, right. . . . .	126
A.1	Trigger yield and efficiency for both the MET trigger and small-R jet trigger approach. Each entry represent a signal point, labelled by the $Z'$ mass and the $R_{inv}$ fraction. . . . .	135
A.2	The factor of improvement in $S/\sqrt{B}$ for each trigger method compared to the un-triggered case. . . . .	136
A.3	The ratio of $S/\sqrt{B}$ of jet trigger over $E_T^{\text{miss}}$ trigger selection. . . . .	136
A.4	Analysis variables where high $R_{inv}$ signals a clearly distinct from background and low $R_{inv}$ variables. On the contrary, leading jet $p_T$ is one of the only variables where low $R_{inv}$ signals are distinct from background. . . . .	138
A.5	OR of jet and $E_T^{\text{miss}}$ triggers. . . . .	138
B.1	. . . . .	139
B.2	. . . . .	140
B.3	Comparison of PFN AUC (top) and sensitivity in the $m_T$ mass window (bottom) for a single PFN model (left) vs. two models trained on $R_{inv} < 0.5$ and $> 0.5$ separately (right). . . . .	141
B.4	Scans done to check for optimality of PFN training parameters. . . . .	142
B.5	$\phi$ Performance comparison between single jet and jet system ML approach . . . . .	143
B.6	$\phi$ Comparison in the AUC score across the grid for the mixed background strategy vs the QCD only strategy. The bottom table highlights that the QCD only strategy gives superior sensitivity across the signal grid. . . . .	144

C.1	$E_T^{\text{miss}}$ in data before and after Tight event cleaning is applied.	145
C.2	$E_T^{\text{miss}}$ vs jet1 $p_T$ in data before and after Tight event cleaning is applied.	146
C.3	$\eta$ vs $\phi$ for leading and subleading jets, before and after the application of tight cleaning.	146
C.4	Added NCB preselection and impact on $E_T^{\text{miss}}$ shape.	147
C.5	NCB preselection impact on data and signal yields.	148
C.6	Impact of tight cleaning and non-collision background preselection.	149
C.7	$E_T^{\text{miss}}\phi$ in data, before (left) and after (right) application of the TileCal correction tool.	149
D.1	Index of jets truth matched (by requirement of $\Delta R < 0.4$ ) with dark quark.	150
D.2	Percent of jets with $\Delta R(j, E_T^{\text{miss}}) < 0.4$ comparing two jet identification strategies. Leading and subleading jets are seen to be the better metric for identifying jets associated with the dark quark decay.	151
E.1	Background-only $m_T$ fits using representative MC in the CR (left), VR (middle), and SR (right).	152
F.1	Example determinations of the 60% mass window means for several signal points.	154
F.2	Signal mass resolution for $m_T$ binning for the signal grid in ( $R_{inv}$ , mass) space.	154
F.3	$m_T$ bins based on the signal mass resolution and the minimum 100 GeV width requirement, for each $R_{inv}$ signal category.	155

## List of Tables

4.1	General performance goals of the ATLAS detector [34]. . . . .	26
6.1	Fixed parameters in the Pythia8 HV model . . . . .	64
6.2	Values for $m_{dark}$ . . . . .	64
6.3	Mass points and cross sections of the SVJ search signal grid. The cross section is determined by the $Z'$ mass and its SM coupling $g_q$ ; it is not impacted by the $R_{inv}$ fraction. . . . .	64
8.1	Preselection cuts for data (left) and signal (right). . . . .	92
8.2	Post-fit parameters for the PFN CR and VR. $p1$ can also be considered $N_{bkg}$ or the normalization factor. . . . .	104
8.3	Post-fit parameters for the ANTELOPE CR and VR. . . . .	113
9.1	Post-fit parameters for the PFN SR. $p1$ can also be considered $N_{bkg}$ or the normalization factor. . . . .	119
9.2	Summary of Experimental Uncertainties and their impact on the yield of MC signal events. . . . .	121

## Acknowledgements

I would first and foremost like to thank my advisor, Michael Tuts, for his support and guidance throughout the past few years. Mike, you have provided me with unwavering counsel through two international moves, the selection of this analysis topic, and the chaotic journey towards completing this thesis. I am very grateful for the creative direction I have been allowed to take this work in, making it truly my own.

I would also like to thank John Parsons and Gustaaf Brooijmans for their mentorship during my time at Columbia. I have learned so much from both of you, through my work at Nevis Laboratories, through my coursework, and through the variety of Columbia ATLAS projects I have been fortunate to work on during my time at Columbia.

These acknowledgements would be tragically incomplete without the mention of Julia Gonski, my colleague, former post-doc, and personal role model. Julia, I have learned so much from you about how to navigate scientific problem-solving, collaboration, and life in general. From the bottom of my heart, thank you for your heroic efforts and unwavering dedication to this project and to my success, even through your own career transition.

I am very grateful to the many ATLAS students I have grown with and learned from during my time at Columbia. Andrew, Daniel, Gabriel, Kiryeong, Eleanor, Alan and Kiley, thank you so much for always answering my Skype messages, offering encouragement and wisdom, and to Gabe and Kiryeong in particular for your hard work on this analysis.

To my friends and family who have helped me celebrate my victories and supported me in my times of need over the past 6 years, I could never have done this without you. Thank you

Mom and Dad for believing in me every time I doubted myself, and encouraging me to pursue my passions. And to my fiancé Tayo - words cannot express how grateful I am for your steadfast support, whether you were 10 feet or 1,000 miles away. Thank you for always being there for me.

Lastly, I would like to thank Huckleberry Roasters and Tonantzin Casa de Café, my two wonderful local coffee shops. Without your delicious caffeine, fast wifi, and distraction free work environment, I would never have managed to complete this thesis.

# **Part I**

## **Theory**

# Chapter 1: The Standard Model

The Standard Model of particle physics is a universally accepted framework which explains the interactions of fundamental particles. All known fundamental particles, outlined in Figure 1.1, are represented in the Standard Model. The model describes three of the four known forces: the electromagnetic force, the weak force, and the strong force. Gravity, the fourth fundamental force, is not addressed by the Standard Model. The Standard Model was primarily developed over the course of the 1960s and 1970s, by combining the work of many physicists into one coherent model. The Standard Model has been established as a well-tested theory by decades of experimental physics research.

This chapter will seek to introduce the phenomenology and mathematical foundations of the Standard Model, and present the supporting experimental evidence. Phenomenon which are unexplained by the Standard Model such as gravity will be considered at the end of the chapter, leading to an exploration of theories beyond the Standard Model in the subsequent chapter.

## 1.1 Phenomenology: Particles and Forces

### 1.1.1 Particles

A classic representation of the particles comprising the Standard Model is shown in Figure 1.1. The two primary particles classes are bosons (gauge bosons and the scalar Higgs boson) and fermions (leptons and quarks). The bosons are carriers of fundamental forces, while the fermions are the building blocks of matter. Fermions are sorted into three *generations*, and each fermion is identified by a unique *flavor*.

Each entry in the table in Figure 1.1 is accompanied by 3 characteristic numbers: mass, charge, and spin. The mass of each particle is determined to limited precision by experimental observation,

## Standard Model of Elementary Particles

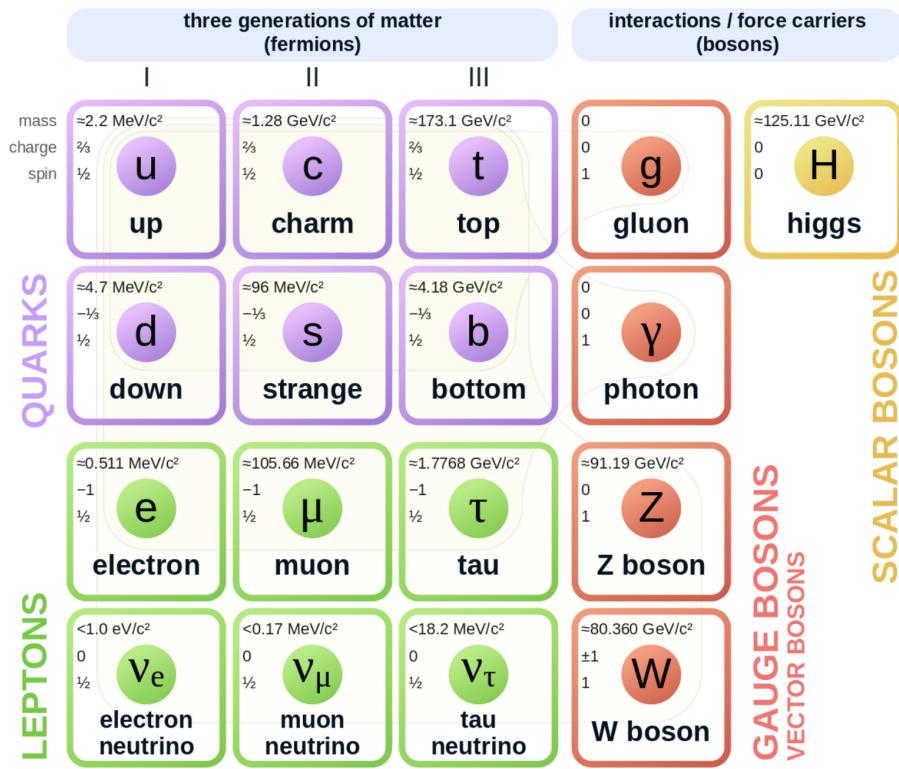


Figure 1.1: Diagram of the 17 particles comprising the Standard Model

with the exception of photons and gluons which are known to be massless. Charge refers to the electromagnetic charge, which is integer for leptons and fractional for quarks. Spin is an intrinsic form of angular momentum carried by fundamental particles; all fermions have half integer spin, while bosons have integer spin.

Each particle is also known to have an *antiparticle*. Each antiparticle has the same mass but the opposite charge of their Standard Model counter part; for example, the antiparticle of the electron is the positron, which has all the same properties but a positive charge. The photon, Z boson, and Higgs are each their own antiparticle. The nature of antineutrinos is an open question driving neutrino physics research, as it is not currently known whether neutrinos are their own antiparticle.

### 1.1.2 Forces

The three fundamental forces explained by the Standard Model are the electromagnetic force, the strong force, and the weak force. The photon is the carrier of the electromagnetic force, which dictates the nature of interactions between electrically charged particles, and is widely covered by introductory physics courses. The electromagnetic force has an infinite interaction range, a result of the massless and non-self interaction nature of the photon. The electromagnetic interaction is described by the theory of quantum electrodynamics (QED).

The weak force gives rise to atomic radiation and decay. It allows for the processes of beta decay, which enables conversion between neutrons and protons within the nucleus of an atom. In the process of beta decay, a proton decays into a neutron, a positron, and a neutrino; or, a neutron decays into a proton, an electron and an antineutrino. The weak interaction allows for quark flavor mixing, the which enables beta decay. The  $W^+$ ,  $W^-$ , and  $Z^0$  are the force carriers of the weak force. The effective range of the weak force is limited to subatomic distances, as a result of the massive nature of the mediator bosons. The unified theory of the electroweak interaction posits that at high enough energies the electromagnetic interaction and the weak force merge into the same force. This threshold is termed the unification energy and calculated to be about 246 GeV [1].

The strong force confines quarks into hadron particles, such as protons and neutrons. The

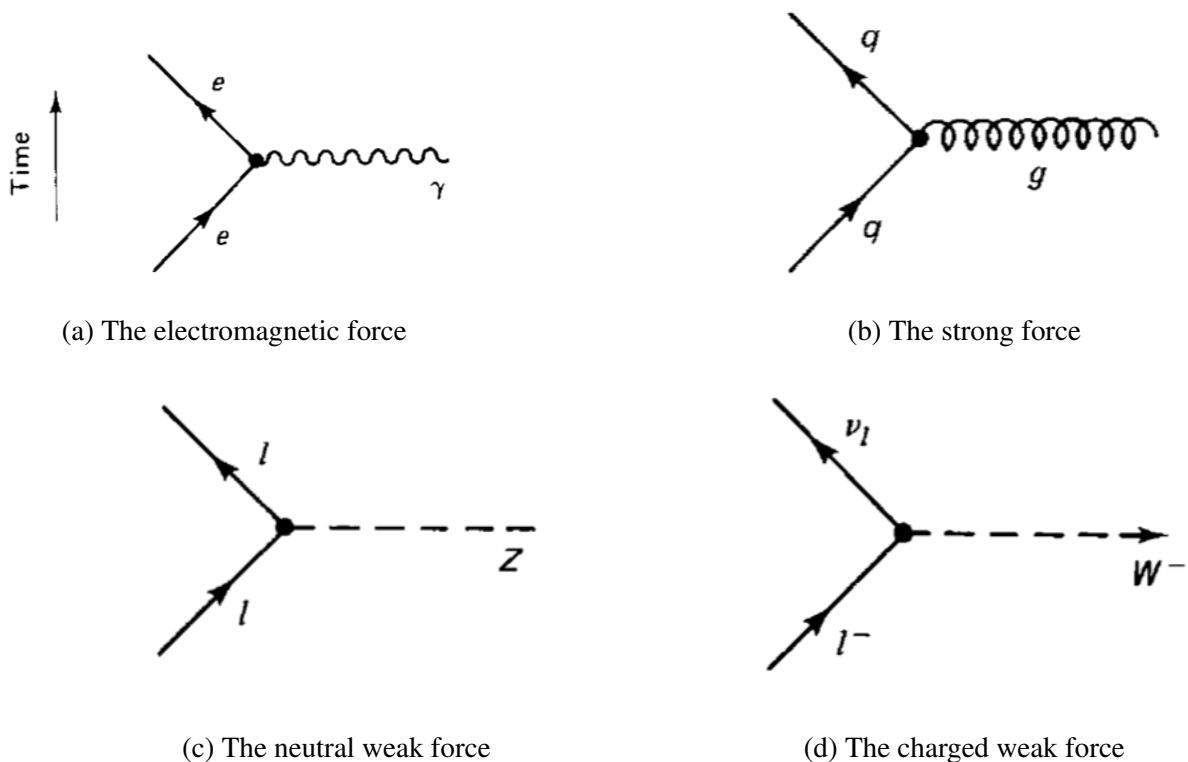


Figure 1.2: Fundamental particle interactions of the three fundamental forces described by the Standard Model [2].

strong force also allows for the creation of atomic nuclei by binding protons and neutrons together, and is generally referred to as the “nuclear force” in this context. The gluon is the mediator of the strong force, which is a short-range force which acts at subatomic distances on the order of  $10^{-15}$  m. At this range, the strong force is about 100x as strong as the electromagnetic force, which allows for the creation of positively charged nuclei [2]. The strong force is described by the theory of quantum chromodynamics (QCD). In the same way that QED dictates the interaction of electrically charges particles, QCD dictates the interactions of *color-charged* particles. Due to the particular importance of QCD in this thesis, this topic will be explored in detail in section 1.2.

The fundamental Feynmann diagram for each of the three forces discussed here is depicted in Figure 1.2. The fourth fundamental force, gravity, is not currently explained by any known mechanism within the Standard Model.

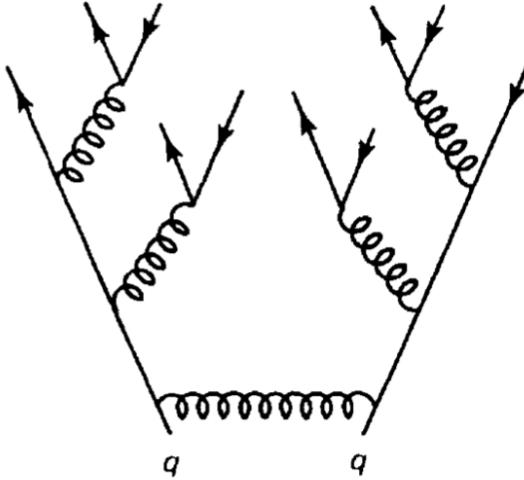


Figure 1.3: An example Feynmann diagram of jet production

## 1.2 QCD and Jets

While there is only one type of electric charge, there are three types of color charge; red, green, and blue. In the process  $q \rightarrow q + g$ , the color of the quark can change. In order to conserve color charge, gluons are bicolored, and always carry some positive color charge and some negative color charge.

Color charged particles can only exist in bound states which result in a neutral total color charge, a principle known as confinement. This requires that quarks and gluons exist in group states known as hadrons; either mesons in the case of two quarks or baryons in the case of three quarks. When a quark is separated from a hadron, confinement dictates that other colored objects are produced around the quark to obey confinement. An example of this process is shown in Figure 1.3. This ensemble of objects, generally a mixture of quarks and gluons, is termed a *jet*. Jets are among the most common phenomenon observed by detectors at hadron colliders, and their complex structure makes them a key focus of many physics analyses.

### 1.3 Symmetries

The Standard Model is a renormalizable quantum field theory that obeys the local symmetry  $G_{SM}$ :

$$G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y. \quad (1.1)$$

The  $SU(3)_C$  symmetry component represents the non-Abelian gauge group of QCD. There are 8 generators for the  $SU_C(3)$  group which correspond to 8 types of gluon, each representing a different superposition of color charge [3]. The  $SU(2)_L \times U(1)_Y$  symmetry group represents the electroweak sector of the Standard Model, which can be spontaneously broken into the electromagnetic and weak sectors. There are 4 generators for this group, which correspond to four massless gauge bosons  $W^1$ ,  $W^2$ ,  $W^3$ , and  $B$ . From these massless gauge bosons are formed the massive mediators of the weak force, the  $W^-$ ,  $W^+$  and  $Z^0$  bosons, and the massless electromagnetic force carrier, the photon  $\gamma$ . Spontaneous symmetry breaking and the process by which gauge bosons acquire mass will be addressed in section 1.3.1.

Noether's theorem [4] stipulates that any continuous symmetry is associated with a conserved quantity. In the Standard Model, this means that the  $SU(3)_C$  symmetry gives rise to conservation of color charge. The  $SU(2)_L \times U(1)_Y$  symmetry gives rise to conservation of electromagnetic charge. Conservation of spin results from the Poincaré symmetry described by the theory of special relativity, which combined with Noether's theorem gives us the conservation of energy, momentum, and angular momentum.

The SM Lagrangian is invariant under  $CPT$  symmetry, or charge, parity, and time reversal. Charge conjugation ( $C$ ) transform a particle into its corresponding antiparticle by reversing the charge and other quantum numbers. Parity conjugation ( $P$ ) reverses spatial coordinates, which transforms left-handed particles into right-handed particles and vice-versa. Time reversal ( $T$ ) is the theoretical process of reversing time. The  $L$  subscript in the  $SU(2)_L$  group indicates that this symmetry only applies to left-handed fermions. As a result, the  $W^{1,2,3}$  gauge bosons of  $SU(2)_L$

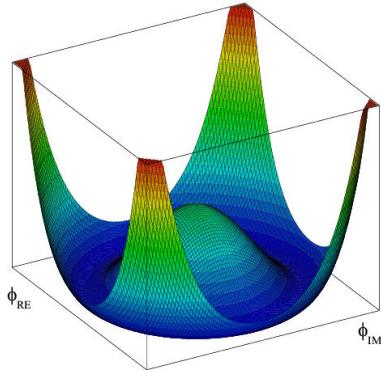


Figure 1.4: An illustration of the “hat shaped” potential of the Higgs field, resulting in a non-zero vacuum expectation value.

only interact with left handed particles, a process which maximally violates P-symmetry in the weak force. A small amount of the CP symmetry violation is also observed in the Standard Model, through the decays of strange flavored mesons [5] and  $b$ -mesons [6]. The CPT theorem posits that the violation of CP symmetry implies that T-symmetry must also be violated, so that CPT is a preserved symmetry.

### 1.3.1 Spontaneous Symmetry Breaking and The Higgs Mechanism

Spontaneous symmetry breaking is the process by which a Lagrangian obeys a symmetry at high energies, but exhibits asymmetric behavior at lower energies. The electroweak symmetry group is spontaneously broken as  $SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}$ . The quantity conserved by the  $SU(2)_L$  symmetry is weak isospin  $T_{1,2,3}$ , while the quantity conserved by  $U(1)_Y$  symmetry is weak hypercharge  $Y$ . Below very high energies, the presence of the Higgs field causes the electroweak symmetry to break. The Higgs field is a scalar field which forms a complex doublet of the  $SU(2)$  symmetry group, with four degrees of freedom. The shape of the Higgs field potential, shown in Figure 1.4, results in a ground state with a non-zero vacuum expectation value; thus the Higgs field takes a non-zero value throughout all space, which breaks the symmetry of the weak isospin  $SU(2)$  group.

The interaction with the Higgs field mixes the four massless gauge bosons  $W^{1,2,3}$  and  $B$ . Three Higgs field degrees of freedom mix with the massless gauge bosons, resulting in three massive

gauge bosons  $W^-$ ,  $W^+$  and  $Z^0$ . The massless photon  $\gamma$  is created from the components of the massless gauge bosons which do not interact with the Higgs field. The scalar Higgs boson arises from the one unmixed degree of freedom the Higgs field. Spontaneous symmetry breaking also violates the conservation of weak isospin and weak hypercharge, leaving only electromagnetic charge ( $Q = T_3 + \frac{1}{2}Y$ ) as a conserved quantity associated with the  $U(1)_{EM}$  symmetry.

## 1.4 Experimental Validation of the Standard Model

The theoretical framework of the Standard Model coalesced into a unified theory in the mid-20th century. A cascade of discoveries providing empirical evidence for the model followed closely. In the 1960s, three quarks (up, down and strange) and four leptons (electron, muon, and their associated neutrinos) were the known particulate building blocks of matter and the Standard Model. The discovery of the charm quark in 1974, through the observation of the  $J/\psi$  meson [7][8], confirmed the existence of a fourth quark flavor. The discovery of the  $\tau$  in 1975 [9] provided the first evidence of a 3rd generation of matter. This was quickly followed by the observation of the  $\Upsilon$  meson in 1977 [10], which provided evidence for the existence of a fifth quark, the  $b$  quark (bottom, or beauty). The existence of a 3rd generation of fermion also explained the observation of CP violation in the weak force, as it allowed for the addition of a complex phase in the CKM matrix (a unitary matrix which describes flavor mixing in the weak interaction). The top quark ( $t$ ) and tau neutrino ( $\nu_\tau$ ) were predicted at this point as the final building blocks of three complete generations of fermions, and they were discovered by experimental observation around the turn of the 21st century [11] [12] [13].

The W and Z bosons were predicted by the Standard Model, but to observe them required the construction of a particle accelerator powerful enough to produce them. They were finally observed at CERN in 1983 by the UA1 and UA2 experiments [14] [15] at the newly constructed Super Proton Synchrotron (SPS). Their masses were observed to be compatible with the masses predicted by the Standard Model nearly a decade earlier. The final missing piece then was confirming the existence of the Higgs, which again required the construction of a newer and more powerful collider. CERN

achieved this with the construction of the Large Hadron Collider (LHC), and in 2012 the ATLAS and CMS experiments announced the discovery of the Higgs particle [16] [17].

## 1.5 Limitations of the Standard Model

While the Standard Model has enjoyed decades of experimental results which confirm its predictions, there are several glaring shortcomings. The observed phenomenon for which the Standard Model provides no explanation are summarized below.

- Gravity - the Standard Model does not account for the fourth fundamental force of gravity.
- Dark Matter - there is no viable candidate to explain the existence of dark matter, a non-interacting form of matter which must exist to account for gravitational observations which cannot be explained by general relativity, such as the motion of galaxies, gravitational lensing, and the structure of the universe [18].
- Matter-Antimatter asymmetry - the level of CP violation in the Standard Model isn't sufficient to explain the large discrepancy between the amount of matter and the amount of antimatter in the universe today, and the origins of this imbalance are not understood.
- Neutrino masses - the Standard Model assumes that neutrinos are massless and provides no mechanism for them to acquire mass. However, observations of neutrino oscillations indicates they posses some small non-zero mass [19].

In addition to these unexplained natural phenomenon, there are several questions about the *naturalness* of the Standard Model. The principle of naturalness states that dimensionless ratios between physical constants should be of order 1, and that nature should not be arbitrarily fine-tuned. While this is largely an aesthetic argument, it points to many aspects of the Standard Model for which there exists no natural explanation.

- Strong CP - while CP symmetry is violated in the weak force, observations indicate that it is preserved by the strong force [20]. The Standard Model predicts that CP violation in the

strong force is possible. There is no principle which motivates this incongruity between the weak force and strong force.

- Hierarchy Problem - The wide range of masses for elementary particles and the wide range of scales at which the four fundamental forces operate is not motivated by the SM. Specifically, it is not understood why the Higgs mass is observed to be well below the Plank scale  $\lambda$ , which is the energy level at which the effects of quantum gravity become significant. QFT indicates that the Higgs mass is determined by contributions from all energy scales including  $\lambda$ , meaning that its observed mass is inexplicably small.

The limitations of the Standard Model provide a road map for theoretical and experimental particle physicists, who seek to develop new theories which account for these observations, and then to find evidence which might support these *Beyond the Standard Model* (BSM) theories. The next chapter will introduce the BSM theories which motivate the physics search presented in this thesis.

## Chapter 2: Physics Beyond the Standard Model

In light of the various phenomenon unexplained by the Standard Model, physicists have proposed various extensions to the Standard Model, collectively termed *Beyond the Standard Model* (BSM) theories. A particular focus of the physics programs at the Large Hadron Collider (LHC) are BSM models which suggest dark matter candidate particles. If these particles couple to Standard Model, they could be produced and observed at the LHC. This chapter will explore Hidden Valley models, and in particular the potential for Hidden Valley models to produce *semi-visible jets*. This will set the theoretical foundations for the experimental search presented in the later chapters of this thesis. The mechanisms of dark QCD that arise from these models and allow for the production of semi-visible jets will also be discussed.

### 2.1 Hidden Valley Models

Hidden Valley (HV) models are a category of BSM models that allow for dark matter (DM) production at the LHC. They extend the Standard Model with an additional non-Abelian gauge group [21]. This introduces the possibility of a complex dark sector, which mirrors the complexities of Standard Model QCD, and introduces the possibility of dark quarks and gluons. The term “hidden valley” refers to the idea that the DM is hidden from the SM by a high-energy barrier, as illustrated in Figure 2.1. The dark sector is assumed to communicate with the Standard Model via a “portal”, or “messenger particle”, that can interact with both Standard Model and HV forces. For the s-channel scenario, the portal is considered to be a new massive mediator particle  $Z'$ .

The portal particle allows for the production of dark sector particles at hadron colliders. If dark quarks are produced via the decay  $Z' \rightarrow q_D \bar{q}_D$  they can hadronize and form dark jets. The properties of the dark jets are determined by the dynamics of the dark sector, which are explored in

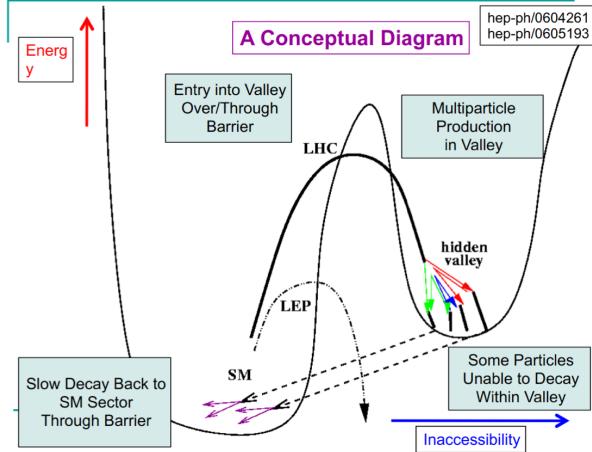


Figure 2.1: Illustration of the hidden valley potential.

the subsequent section. Depending on the details of the model, the jets formed by the dark hadrons can be categorized as fully dark, semi-visible, leptonic, emerging, or other [21].

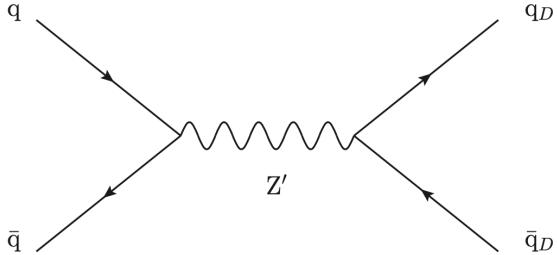


Figure 2.2: The massive mediator particle  $Z'$  of the s-channel realization of a HV model

## 2.2 Dark QCD

The theoretical underpinning of the semi-visible jet phenomenology is a dark sector with a gauge group  $SU(N)_d$  leading to confinement at a scale  $\Lambda_d$ . For illustration, let's consider the case of an  $SU(2)_d$  gauge theory, which gives rise to two dark fermionic generations  $\chi_a = \chi_1, \chi_2$ . Following the work of Ref. [22] we can write the fundamental dark Lagrangian as:

$$\mathcal{L}_{dark} \supset -\frac{1}{2} \text{Tr} G_{\mu\nu}^d G^{d\mu\nu} - \bar{\chi}_a (i\cancel{D} - M_{d,a}) \chi_a \quad (2.1)$$

The first term allows for the dark gluons to self-interact, while the second term enables the dark quarks to hadronize and acquire mass. The dark quarks are assumed to have a common mass  $M_d$ . The coupling strength of the strongly interacting dark quarks is termed  $\alpha_d$ . At the confinement scale  $\Lambda_d$ , the dark quarks can form bound states. At the scale  $M_d \approx \Lambda_d$  a QCD-like shower occurs.

The properties of the hadrons formed by the dark quarks are of particular importance to the observed dark QCD dynamics. Dark-isospin number  $U(1)_{1-2}$  and dark-baryon number  $U(1)_{1+2}$  are accidental symmetries of the theory which determine the stability of the hadrons. In the case of two dark flavors, six dark hadrons can be formed: four mesons ( $\chi_1\bar{\chi}_1$ ,  $\chi_2\bar{\chi}_2$ ,  $\chi_1\bar{\chi}_2$ ,  $\bar{\chi}_1\chi_2$ ) and two baryons ( $\bar{\chi}_1\bar{\chi}_2$ ,  $\bar{\chi}_1\chi_2$ ). The mesons  $\chi_1\bar{\chi}_2$  and  $\bar{\chi}_1\chi_2$  are charged under dark-isospin and will be stable if this symmetry is unbroken. The baryons would also be stable as they are charged under the dark-baryon number. These four stable hadrons become dark matter candidates of the theory. The  $\chi_1\bar{\chi}_1$  and  $\chi_2\bar{\chi}_2$  mesons are not charged under either symmetry and are thus expected to decay. The unstable mesons can decay into stable dark mesons, or into an off-shell  $Z'$ . The off-shell  $Z'$  will then decay into two DM quarks or two SM quarks, and its products will continue to shower until the final state particles are stable.

The number of stable and unstable dark states varies substantially depending on the details of the model. The model discussed above can be generalized from  $SU(2)_d$  to  $SU(N)_d$ , with any number of colors  $N_c$  or flavors  $N_f$ . This affects the ratio of possible stable to unstable mesons, which can directly impact the amount of missing energy. The fraction of missing energy is a variable in many dark QCD models, and is especially important in the case of semi-visible jets.

### 2.3 Semi-visible Jets

A “semi-visible jet” occurs when the heavy  $Z'$  messenger particle decays into dark quarks, which then hadronize in a QCD-like shower. If some of the dark hadrons are stable while others decay to SM quarks via the off-shell  $Z'$ , a collimated mixture of visible and dark matter is formed – this is termed a semi-visible jet. If the  $Z'$  messenger particle is produced at rest, the two jets will be back-to-back in the transverse plane. If there is an imbalance in the amount of invisible particles

between the two jets, one of the jets will be observed to be aligned with missing transverse energy.

While there are a myriad of HV and dark QCD models, a handful of model parameters are most important in determining the observable of these showers within a particle detector. The coupling strength  $\alpha_d$  is one of the most important, as it controls the fraction of dark hadrons emitted in the shower and their average  $p_T$ . The mass of the dark quarks directly impacts the jet mass. If the masses of the dark quark flavors are comparable, the ratio of stable to unstable dark hadrons will be approximately 1:1. However, if there is a mass splitting, stable or unstable dark hadrons may be favored, which impacts the amount of missing energy observed.

The ratio of stable to unstable dark hadrons in the shower is a critical variable for capturing the behavior of dark showers. This value is termed  $R_{inv}$ :

$$R_{inv} = \frac{\# \text{ of stable hadrons}}{\# \text{ of hadrons}} \quad (2.2)$$

Events containing jets aligned with missing transverse momentum are generally considered to be misreconstructed by other DM searches, and therefore discarded. This class of final states is therefore largely uncovered by existing DM searches. The nature of the dark hadron shower is determined by the following parameters: the  $Z'$  mass  $m_{Z'}$ , the  $Z'$  couplings to visible and dark quarks  $g_q$  and  $g_{q_D}$ , the number of dark colors and flavors, the characteristic scale of the dark sector confinement  $\Lambda_D$ , the mass scale of the dark hadrons  $m_D$ , and the average fraction of stable hadrons in the decay  $R_{inv}$ . The coupling to SM quarks determines the  $Z'$  production cross section.

# **Part II**

# **Experiment**

## Chapter 3: The Large Hadron Collider

The Large Hadron Collider (LHC) is a 26.7 km circular high-energy particle accelerator, spanning the Swiss-French border near the city of Geneva, Switzerland [23]. The LHC occupies the tunnel constructed in 1989 for the Large Electron-Positron (LEP) Collider, and reaches a maximum depth of 170m below the surface. The LHC is operated by the European Organization for Nuclear Research (CERN), the largest international scientific collaboration in the world.

The LHC accelerates protons and heavy ions, and collides them at four interaction points around the ring, with a design center-of-mass energy per collision of  $\sqrt{s} = 14$  TeV. Each interaction point is home to one of four detector experiments, which study the products of the collisions. The largest of these experiments is the ATLAS detector, a general purpose detector designed to study the Standard Model and search for new physics that could be produced in LHC collisions [24]. The CMS detector is another general purpose detector, designed and operated independently of the ATLAS detector, but intended to probe the same range of physics [25]. The ALICE experiment is a dedicated heavy ion experiment, and the LHC-b experiment is a dedicated *b*-physics experiment [26] [27].

This chapter will cover the multi-component accelerator complex powering the LHC, the state-of-the-art magnets which steer the particle beams, measurements of the intensity and number of collisions produced by the LHC, and finally an overview of LHC activities in the past, present, and future.

### 3.1 Accelerator Physics

#### 3.1.1 The Journey of a Proton

From 2010 - 2018, the protons which fed the LHC started as hydrogen gas. The electrons were removed from the hydrogen atoms through the use of strong electric fields. The linear accelerator LINAC2 then accelerated the protons to an energy of 50 MeV. Between 2018 and 2020, LINAC2 was replaced with LINAC4, which instead accelerates  $H^-$  ions, hydrogen atoms with two electrons. LINAC4 is capable of accelerating the  $H^-$  ions to 160 MeV. Before injection to the next part of the acceleration chain, both electrons are stripped from the  $H^-$  ions, leaving just protons. From here the protons enter the Proton Synchrotron booster, where they are accelerated up to 1.4 GeV of energy. Subsequently they are sorted into bunches separated in time by 25 ns, where each bunch contains approximately  $10^{11}$  protons. Next the bunches pass through the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), where they reach energies of 25 GeV and 450 GeV respectively. Finally they are injected into the LHC as two beams traveling in opposite direction. The original design allowed each beam to be accelerated up to 7 TeV of energy. Due to limitations in the performance of the superconducting LHC magnets, the highest energy actually achieved by the LHC beams during Run 2 was 6.5 TeV, giving a collision center-of-mass energy of  $\sqrt{s} = 13$  TeV [28]. Figure 3.1 shows the full LHC accelerator complex.

Acceleration in the LHC is performed by eight radio frequency (RF) cavities located around the ring. Each RF cavity produces a 2 MV electric field oscillating at 40 MHz. The 40MHz oscillation produces a point of stable equilibrium every 2.5 ns. These points of equilibrium are synchronized with the occurrence of the proton bunches produced in the PS – a proton bunch occupies one out of every ten points of stable equilibrium, such that the bunches maintain a 25 ns spacing [28].

## The CERN accelerator complex Complexe des accélérateurs du CERN

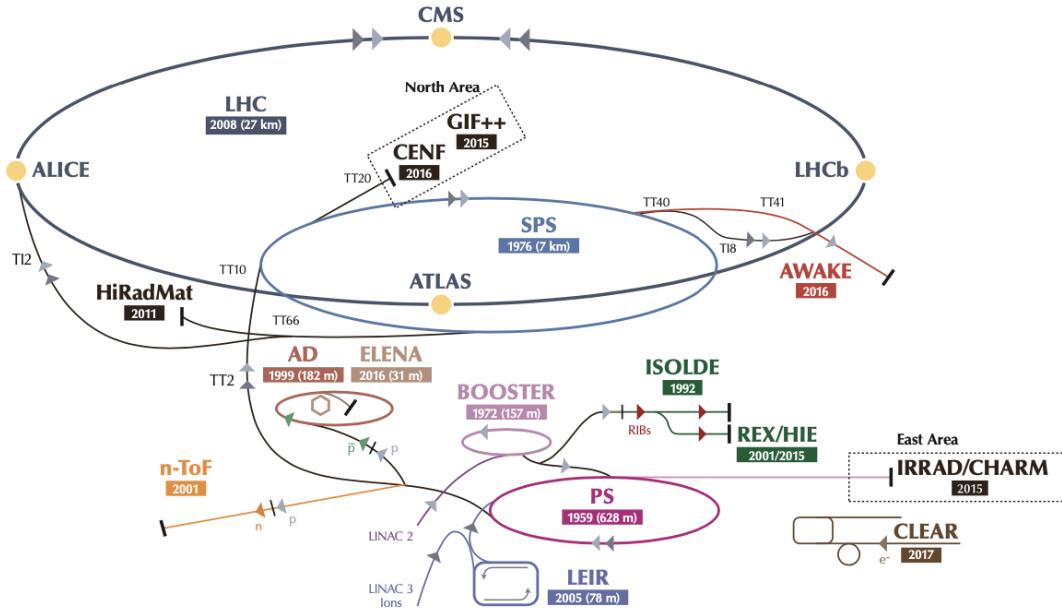


Figure 3.1: The LHC accelerator complex at CERN [29]

### 3.1.2 Magnets

In addition to the acceleration cavities, the LHC houses 9593 superconducting magnets which direct and focus the proton beam on its 27 kilometer journey. The magnets are comprised of superconducting Niobium-Titanium coils cooled to 1.9K by superfluid helium. As the beams approach one of the four collision points around the ring, multipole magnets focus and squeeze the beam for optimal collisions [28].

The LHC is divided into sections, where each section contains an arc and a straight insertion. The arcs are composed of 1232 large dipole magnets which bend the beam to follow the roughly circular 27 km path. The main dipoles generate powerful 8.3 tesla magnetic fields to achieve this bend. Each dipole magnet is 15 meters long and weighs 35 tonnes. The dipoles work in conjunction with quadrupole magnets, which keep the particles in a focused beam, and smaller sextupole, octupole and decapole magnets which tune the magnetic field at the ends of the dipole magnets [30].

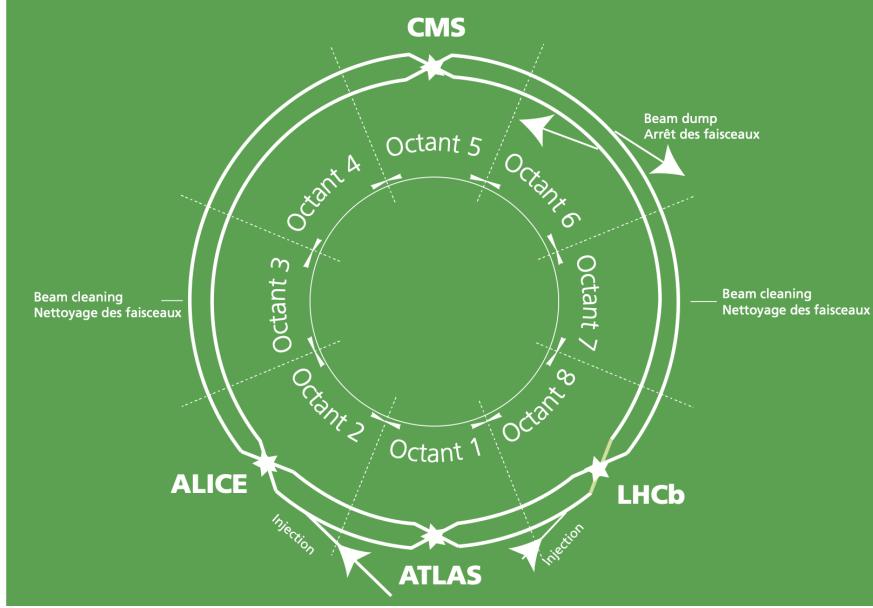


Figure 3.2: The octants of the LHC and location of various beam activities [28]. Stars indicate the locations of beam collisions, and the associated detectors recording the outcome of those collisions.

The straight insertion sections have different purposes depending on their location around the ring: beam collisions, beam injection, beam dumping, or beam cleaning. At the four collision points, insertion magnets squeeze the beam to ensure a highly focused collision. This is accomplished with a triplet of quadrupole magnets, which tighten the beam from 0.2 millimeters to just 16 micrometers in diameter. Insertion magnets also clean the beam, which prevents stray particles from hitting sensitive components throughout the LHC. When the LHC is ready to dispose of a beam of particles, beam dump magnets deflect the path of the beam into a straight line towards a block of concrete and graphite that stops the beam. A dilution magnet then reduces the beam intensity by a factor of 100,000 before the final stop [30]. Figure 3.2 shows the locations various beam activities.

### 3.2 Luminosity

Collisions at the LHC occur when the two beams of proton bunches cross at one of the four interaction points. The intensity of collisions is described by the instantaneous luminosity, the

formula for which is given in equation 3.1.

$$L = \frac{fN_1N_2}{4\pi\sigma_x\sigma_y} \quad (3.1)$$

Here  $f$  is the revolution frequency,  $N_1$  and  $N_2$  are the number of particle per bunch for each beam, and  $\sigma_x$ ,  $\sigma_y$  are the horizontal and vertical beam widths.

The instantaneous luminosity gives the number of the collisions that could be produced at the interaction point per unit of cross-sectional area per unit of time, generally expressed in  $\text{cm}^{-2}\text{s}^{-1}$ . The integrated luminosity is obtained by integrating the instantaneous luminosity over a given block of time, and measures the total number of collisions which have occurred during that operation period. The total integrated luminosity is directly correlated with the size of the datasets collected by the LHC experiments. Total integrated luminosity for Run 2 is illustrated in Figure 3.3.

High levels of instantaneous luminosity result in multiple  $pp$  collisions per bunch crossing, which leads to an effect known as *pileup*. Pileup poses a challenge for detector physics, as reconstructing the products of multiple simultaneous events is far more challenging than reconstructing a single event with no pileup. Pileup conditions vary from year-to-year and run-to-run of LHC operation, and the impact of these conditions are taken into account when analyzing the data, as will be discussed further in Chapter 5. Measurement of pileup conditions during Run 2 are illustrated in Figure 3.3.

The design peak luminosity of the LHC is  $1.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . During Run 1 of the LHC the peak instantaneous luminosity was  $0.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . Over the course of Run 1 the LHC collected a total integrated luminosity of  $5.46 \text{ fb}^{-1}$  at  $\sqrt{s} = 7 \text{ TeV}$ , and  $22.8 \text{ fb}^{-1}$  at  $\sqrt{s} = 8 \text{ TeV}$ . Following the first long shutdown and upgrade phase of operations, the LHC achieved a center of mass energy  $\sqrt{s} = 13 \text{ TeV}$  at the beginning of Run 2 in 2015. The LHC was also able to deliver  $2.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  peak instantaneous luminosity, double the design value. During LHC Run 2, from 2015-2018, the LHC delivered  $156 \text{ fb}^{-1}$  of integrated luminosity for proton-proton collisions. Run 3 of

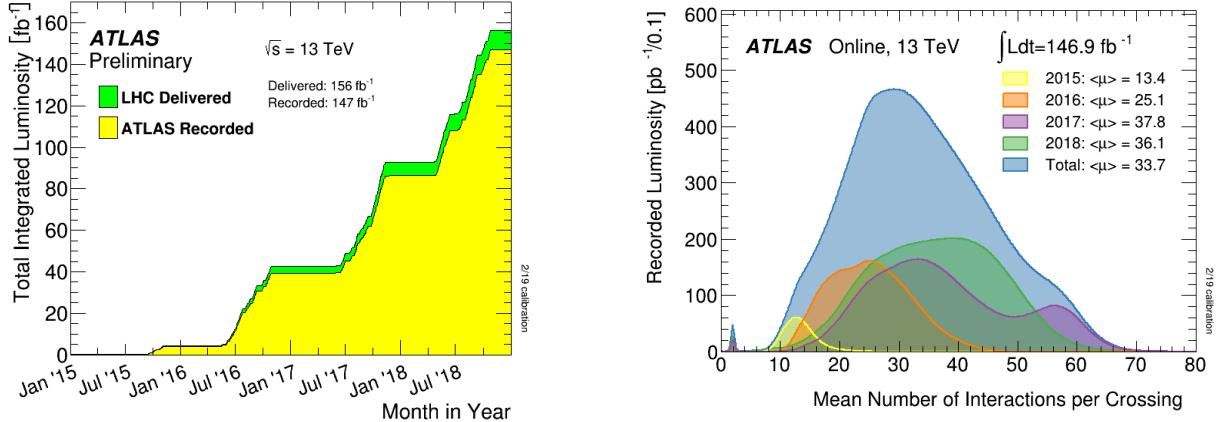


Figure 3.3: (Left) Total integrated luminosity over the course of Run 2. (Right) Average number of  $pp$  interactions per bunch crossing in Run 2. Each curve is weighted by the integrated luminosity for the year.

the LHC began in 2022, and is expected to deliver  $250 \text{ fb}^{-1}$  of integrated luminosity to the ATLAS and CMS experiments by 2026 [31].

The goal of LHC physic analyses is to find and study rare events produced by interesting physics processes. The cross section  $\sigma$  of a given process indicates the probability of that process occurring given the beam conditions of the LHC. Multiplying the cross section by the integrated luminosity of a dataset gives the expected number of events for that process within the dataset.

$$N_{\text{events}} = \int \sigma L(t) dt = \mathcal{L} \times \sigma \quad (3.2)$$

The cross section for most processes of interest, especially BSM processes, is several orders of magnitude below the total cross section for the LHC. Therefore maximizing the number of events produced in collisions is crucial to increase the likelihood of producing events from processes of interest. For this reason, maximizing instantaneous luminosity is a key factor in accelerator design and operation, while mitigating the resulting pileup effects is a key component in detector design and operation.

### 3.3 LHC Timeline

The first proton-proton collisions at the LHC were achieved in 2010 with a center-of-mass energy of  $\sqrt{s} = 7$  TeV. Run 1 of the LHC took place between 2010 and early 2013, during which time the center-of-mass collision energy increased from 7 TeV to 8 TeV. Figure 3.4 shows an overview of LHC activities beginning in 2011, in the midst of Run 1. The data collected during Run 1 led to the discovery of the Higgs boson in 2012 [32].

Between 2013 and 2015 the LHC underwent the first Long Shutdown (LS1) during which time maintenance and renovation was performed on the accelerator chain, including the repair and consolidation of the high-current splices which connect the super-conducting LHC magnets. Run 2 of the LHC took place from 2015 to 2018 and achieved a center-of-mass energy of  $\sqrt{s} = 13$  TeV. Analysis of data collected in Run 2 is still on going, and is the subject of study in this thesis.

Between 2018 and 2022 the LHC underwent the second Long Shutdown (LS2), allowing for further detector and accelerator maintenance and upgrades. Key improvements to the LHC included the improvement of the insulation for over 1200 diode magnets, and the upgrade from LINAC2 to LINAC4 mentioned in Section 3.1.1. Run 3 of the LHC began in 2022 and achieved a center-of-mass energy of  $\sqrt{s} = 13.6$  TeV.

Run 3 is scheduled to continue through 2026, at which point the LHC machine and detectors will undergo upgrades for the *high luminosity* LHC (HL-LHC). The HL-LHC will increase the instantaneous machine luminosity by a factor of 5 - 7.5 with respect to the nominal LHC design. The bottom panel of Figure 3.4 shows an overview of the preparation work for the HL-LHC that has been going on concurrently with Run 1, 2, and 3 of the LHC [33].



Figure 3.4: Timeline of LHC and HL-LHC activities [31]. Integrated luminosity estimates are approximate, and not reflective of the exact amount delivered to each experiment.

## Chapter 4: The ATLAS Detector

The ATLAS detector (**A** **T**oroidal **L**H**C** **A**pparatu**S**) is one of two general purpose physics detectors designed to study the products of proton-proton collisions at the LHC. The detector is composed of a variety of specialized subsystems, designed to fully capture a wide array of physics processes. A diagram of the detector is shown in Figure 4.1. The apparatus is 25m high, 44m in length, and weighs over 7000 tons [34]. The LHC beam pipes direct proton beams to an interaction point at the center of ATLAS, and the cylindrical detector design captures a complete  $360^\circ$  view of the *event*<sup>1</sup>, tracking all particles that result from the collision.

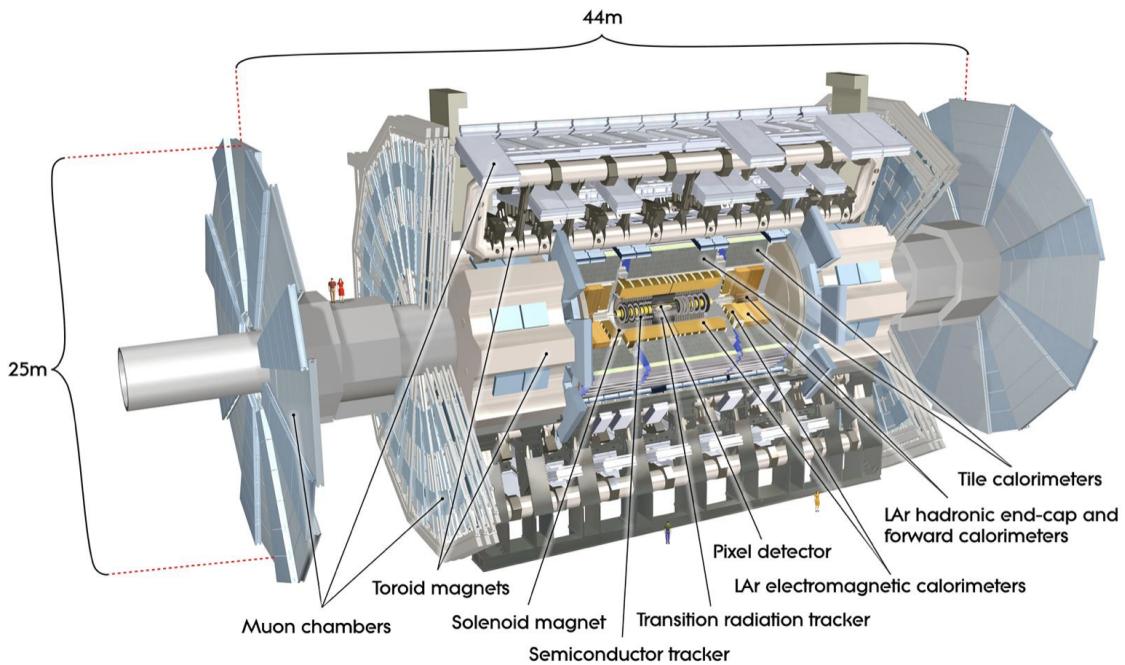


Figure 4.1: The ATLAS detector [34].

<sup>1</sup>An ATLAS event is a snapshot of all the measurements associated to a given  $p p$  collision

The main components of the ATLAS detector are the Inner Detector (ID) which provides high precision tracking of charged particles leaving the collision vertex, the calorimeter system which measures the energy of electromagnetic and hadronic objects, and the Muon Spectrometer (MS) which gives detailed information about muons that reach the outer radii of the detector. A summary of the resolution and angular coverage of each of these systems is presented in Table 4.1. Two magnet systems, a 2 T solenoid magnet surrounding the ID, and a 0.5-1.0 T toroid magnet system situated throughout the MS, produce magnetic fields which bend the trajectory of charged particles traversing the detector. In addition to the main detector components, dedicated forward detectors monitor beam conditions and instantaneous luminosity, and an online trigger system reduces the data rate to a manageable level for storage. Each of these components will be discussed in further detail in this chapter.

<b>Detector component</b>	<b>Required resolution</b>	<b><math>\eta</math> coverage</b>	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	$\pm 2.5$	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\pm 3.2$	$\pm 2.5$
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\pm 3.2$	$\pm 3.2$
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 <  \eta  < 4.9$	$3.1 <  \eta  < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\% \text{ at } p_T = 1 \text{ TeV}$	$\pm 2.7$	$\pm 2.4$

Table 4.1: General performance goals of the ATLAS detector [34].

## 4.1 Coordinate System and Geometry

The ATLAS detector employs a right hand cylindrical coordinate system. The  $z$  axis is aligned with the beam line, and the  $x - y$  plane sits perpendicular to the beam line. The coordinate system origin is centered on the detector, such that the origin corresponds with the interaction point of the two colliding beams. The detector geometry is usually characterized by polar coordinates, where the azimuthal angle  $\phi$  spans the  $x - y$  plane. The polar angle  $\theta$  represents the angle away from the beam line, or  $z$  axis.  $\theta = 0$  aligns with the positive  $z$  axis, and  $\phi = 0$  aligns with the positive  $x$  axis.

The polar coordinate  $\theta$  is generally replaced by the Lorentz invariant quantity *rapidity* or  $y$ :

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right). \quad (4.1)$$

This substitution is advantageous because objects in the detector are traveling at highly relativistic speeds. The relativistic speed also means that the masses of the particles are generally small compared to their total energy. In the limit of zero mass, the rapidity  $y$  reduces to the pseudorapidity  $\eta$ , which can be calculated directly from the polar angle  $\theta$ :

$$\eta = -\ln\left(\frac{\theta}{2}\right). \quad (4.2)$$

The distance between physics objects in the detector is generally expressed in terms of the solid angle between them  $\Delta R$ :

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \quad (4.3)$$

Figure 4.2a depicts the orientation of the coordinate system with respect to the ATLAS detector, while Figure 4.2b illustrates the relationship between  $\theta$ ,  $\eta$ , and the beamline axis  $z$ . Direct or “head on” proton-proton collisions are more likely to result in objects whose momentum is directed along transverse plane (low  $|\eta|$ ); glancing proton-proton collisions are more likely to result in objects whose momentum is directed along the  $z$ -axis (high  $|\eta|$ ). Due to the difference in the nature of these collisions, as well as the cylindrical design of the ATLAS detector, the detector is divided into regions of low and high  $|\eta|$ . Each subsystem has a “central” or “barrel” region covering low  $|\eta|$ , while the “forward” or “end-cap” regions cover the area up to  $|\eta| = 4.9$ . Each of the three main ATLAS subsystems will be discussed in the following sections.

## 4.2 Inner Detector

The Inner Detector (ID) is the ATLAS subsystem closest to the interaction point. The primary purpose of the ID is to determine the charge, momentum, and trajectory of charged particles pass-

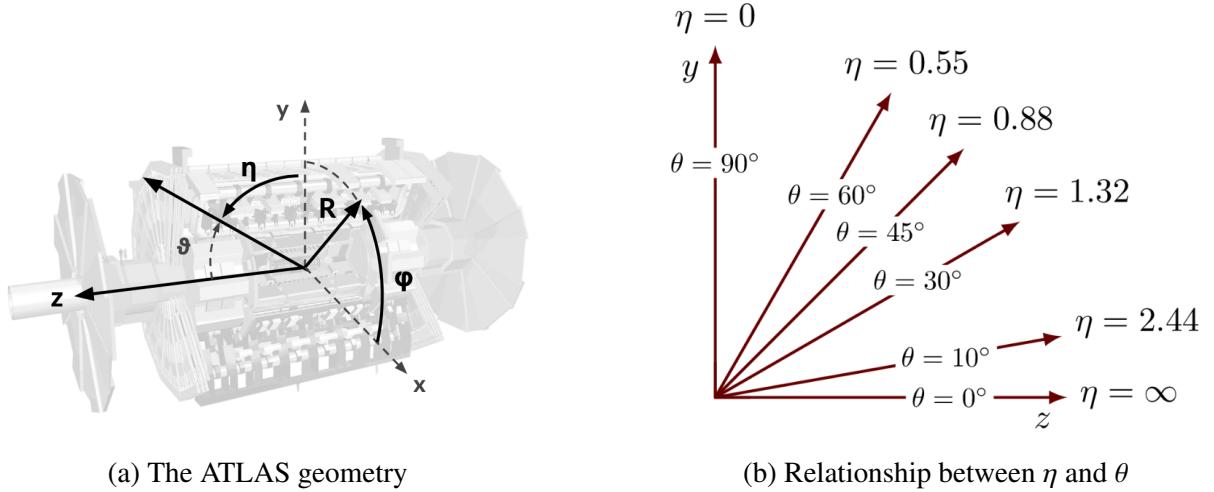


Figure 4.2: ATLAS coordinate system and geometry

ing through the detector. With this information the ID is also able to precisely determine interaction vertices.

The ID is composed of three sub-detectors; the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT) [34]. Figure 4.3 shows the location of these three subsystems with respect to each other and the interaction point.

#### 4.2.1 Pixel Detector

The pixel detector is the first detector encountered by particles produced in LHC collisions. The original pixel detector consists of 3 barrel layers of silicon pixels, positioned at 5 cm, 9 cm and 12 cm from the beamline. There are also 3 disks on each end-cap positioned 50 - 65 cm from the interaction point, providing full coverage for  $|\eta| < 2.5$ . Figure 4.4 illustrates the end-cap geometry for the ID. The layers are comprised of silicon pixels each measuring  $50 \times 400 \mu\text{m}^2$ , with 140 million pixels in total. The pixels are organized into modules, which each contain a set of radiation hard readout electronics chips. In 2014, the Insertable B-layer (IBL) was installed, creating a new innermost layer of the pixel detector sitting just 3.3 cm from the beamline. The pixels of the IBL measure  $50 \mu\text{m}$  by  $250 \mu\text{m}$ , and cover a pseudo-rapidity range up to  $|\eta| < 3$ . The IBL upgrade enhances the pixel detector's ability to reconstruct secondary vertices associated

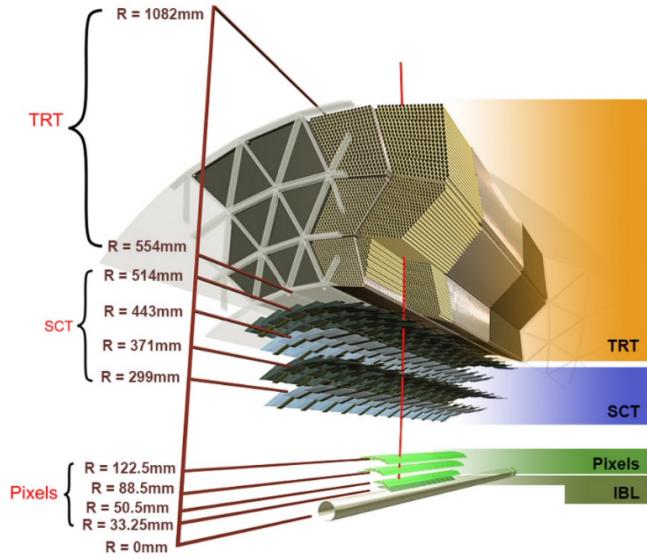


Figure 4.3: A 3D visualization of the structure of the ID in the barrel region [35]

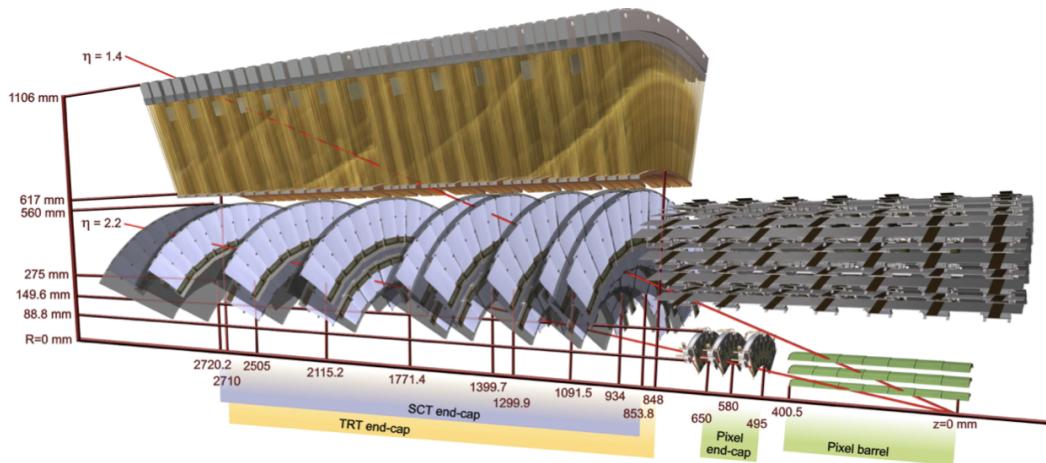


Figure 4.4: A 3D visualization of the structure of the ID end-caps [34]. The paths of two charged particles ( $\eta = 1.4$  and  $2.2$ ) are illustrated.

with short-lived particles such as the b-quark. The improved vertex identification also helped compensate for increasing pile-up in Run 2.

#### 4.2.2 Semiconductor Tracker

The SCT provides at least 4 additional measurements of each charged particle. It employs the same silicon technology as the Pixel Detector, but utilizes larger silicon strips which measure 80  $\mu\text{m}$  by 12.4 cm. The SCT is composed of 4 barrel layers, located between 30 cm and 52 cm from the beamline, and 9 end-cap layers on each side. The SCT can distinguish tracks that are separated by at least 200  $\mu\text{m}$ .

#### 4.2.3 Transition Radiation Tracker

The TRT provides an additional 36 hits per particle track. The detector relies on gas filled straw tubes, a technology which is intrinsically radiation hard. The straws which are each 4 mm in diameter and up to 150 cm in length and filled with xenon gas. The detector is composed of about 50,000 barrel region straws and 640,000 end-cap straws, comprising 420,000 electronic readout channels. Each channel provides a drift time measurement with a spatial resolution of 170  $\mu\text{m}$  per straw. As charged particles pass through the many layers of the detector, transition radiation is emitted. The use of two different drift time thresholds allows the detector to distinguish between tracking hits and transition radiation hits.

### 4.3 Calorimeters

The ATLAS calorimeter system is responsible for measuring the energy of electromagnetically interacting and hadronically interacting particles passing through the detector. The calorimeters are located just outside the central solenoid magnet, which encloses the inner detectors. The calorimeters also stop most known particles, with the exception of muons and neutrinos, preventing them from traveling to the outermost layers of the detector. The ATLAS calorimetry system is composed of two subsystems - the Liquid Argon (LAr) calorimeter for electromagnetic calorimetry and the

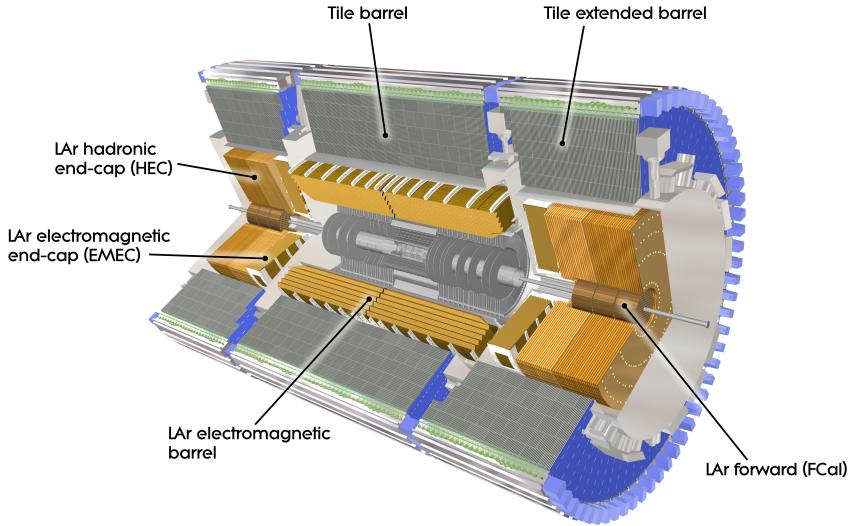


Figure 4.5: ATLAS calorimetry system [36]

Tile calorimeter for hadronic calorimetry. The full calorimetry system is shown in Figure 4.5.

#### 4.3.1 Liquid Argon Calorimeter

The LAr calorimeter is a sampling calorimeter designed to trigger on and measure the energies of electromagnetic (EM) particles, as well as hadronic particles in the high  $|\eta|$  regions [37]. It is divided in several regions, as shown in Figure 4.5. For the region  $|\eta| < 1.4$ , the electromagnetic barrel (EMB) is responsible for EM calorimetry, and provides high resolution energy, timing, and position measurements for electrons and photons passing through the detector. The electromagnetic end-cap (EMEC) provides additional EM calorimetry up to  $|\eta| < 3.2$ . In the region  $1.4 < |\eta| < 3.2$ , the hadronic end-cap (HEC) provides hadronic calorimetry. For hadronic calorimetry in the region  $|\eta| < 1.4$ , corresponding to a detector radii  $> 2.2$  m, the less expensive tile calorimeter (discussed in the next section) is used instead. A forward calorimeter (FCAL) extends the hadronic calorimetry coverage up to  $3.1 < |\eta| < 4.9$ .

The LAr calorimeter is composed of liquid argon sandwiched between layers of absorber material and electrodes. Liquid argon is advantageous as a calorimeter active medium due to its natural abundance and low cost, chemical stability, radiation tolerance, and linear response over a large

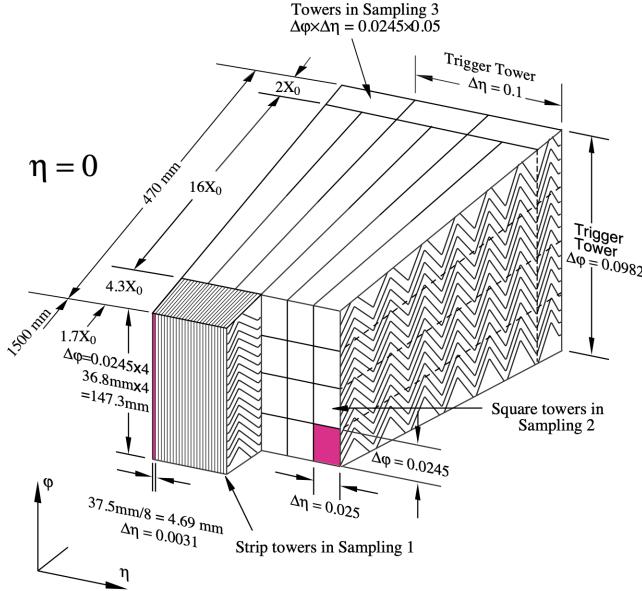


Figure 4.6: Diagram of a segment of the EMB, demonstrating the accordion plate arrangement and the cell segmentation pattern [37]. Three layers are illustrated, and the variation in the cell granularity between the front layer (Sampling Layer 1) and the back layer (Sampling Layer 3) is shown.

energy range [38]. The calorimeter is cooled to 87k by three cryostats: one barrel cryostat encompassing the EMB, and two end-cap cryostats. The barrel cryostat also encloses the solenoid which produces the 2T magnetic field for the inner detector. Front-end electronics are housed outside the cryostats and are used to process, temporarily store, and transfer the calorimeter signals.

## Electromagnetic Calorimeter

For the electromagnetic calorimeters, the layers of electrodes and absorber materials are arranged in an accordion shape, as illustrated in Figure 4.6. The detector is divided into  $\eta - \phi$  regions or *cells*, which are each read out independently. The The accordion shape ensures that each half barrel is continuous in the azimuthal angle, which is a key feature for ensuring consistent high resolution measurements. Liquid argon permeates the space between the lead absorber plates, and a multilayer copper-polymide readout board runs through the center of the liquid argon filled gap.

The detection principle for the LAr calorimeter is the current created by electrons which are

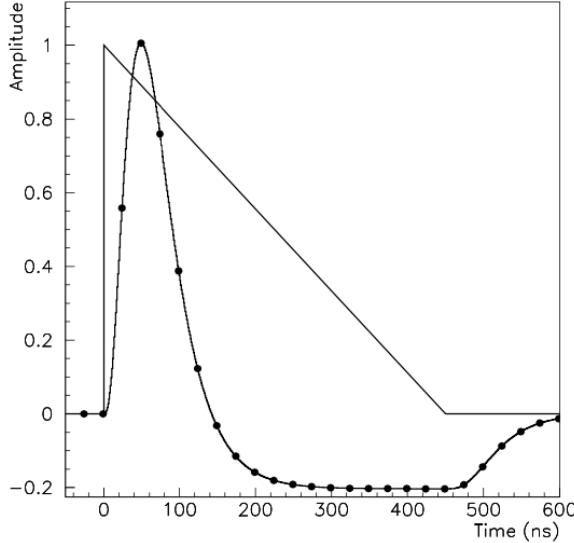


Figure 4.7: A LAr pulse as produced in the detector (triangle) and after shaping (curve) [37]

released when a charged particle ionizes the liquid argon. In the barrel region, the electrons are driven towards the center electrodes by a 2,000 V potential with a drift time of less than 450 ns [39]. In the end-caps the voltage varies as a function of the radius in order to maintain a flat response [37]. The amount of current produced by the ionized electrons is proportional to the energy of the particle creating the signal. Figure 4.7 shows the shape of the signal produced in the LAr calorimeter, before and after it undergoes shaping during the readout process. The shaping of the pulse enforces a positive peak and a negative tail, which ensures that subsequent pulses can be separated with the precision required for the 25 ns LHC bunch spacing.

### Hadronic End-cap Calorimeter

The HEC sits radially beyond the EMEC. The copper absorber plates in the HEC are oriented perpendicular to the beamline, with LAr as the active medium. Each end-cap is divided into two independent wheels; the inner wheel uses 25 mm copper plates, while the outer wheel uses 50 mm plates as a cost saving measure. In each wheel, the 8.5 mm plate gap is crossed by three parallel electrodes, creating an effective drift distance of 1.8 mm. This gap is illustrated in Figure 4.8. Each wheel is divided into 32 wedge-shaped modules, each containing their own set of readout

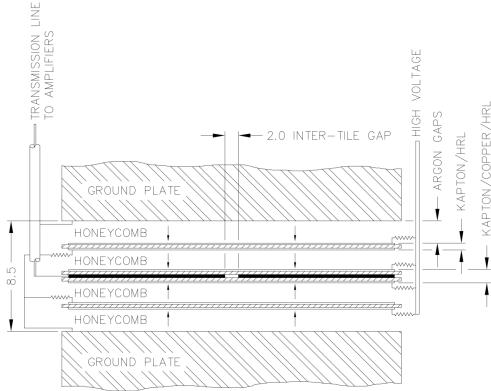


Figure 4.8: Readout gap structure in HEC [37]

electronics.

## Forward Calorimeter

The forward range ( $3.1 < |\eta| < 4.9$ ) is covered by the FCal, which provides both EM and hadronic calorimetry. It is composed of three active cylindrical modules; one EM module with copper absorber plates, and two hadronic modules with tungsten absorber plates [37]. The plates are oriented perpendicular to the beamline, and LAr is used as the active material throughout. The electrodes of the FCal consist of tubes that run parallel to the beam line, arranged in a honeycomb pattern. The resulting LAr gaps are as small as  $250 \mu\text{m}$ , which enables the FCal to handle the high particle flux in the forward region.

### 4.3.2 Tile Calorimeter

The Tile Calorimeter (TileCal) provides hadronic calorimetry in the region  $|\eta| < 1.7$ , and surrounds the LAr calorimeter. It is responsible for measurements of jet energy and jet substructure, and also plays an important role in electron isolation and triggering (including muons) [40]. TileCal is composed of 3 sections, as shown in Figure 4.5; a barrel calorimeter sits directly outside the LAr EMB and provides coverage up to  $|\eta| < 1.0$ . Two extended barrel sections sit outside the LAr end-caps and cover the region  $0.8 < |\eta| < 1.7$ .

TileCal is a sampling calorimeter composed of steel and plastic scintillator plates as illustrated

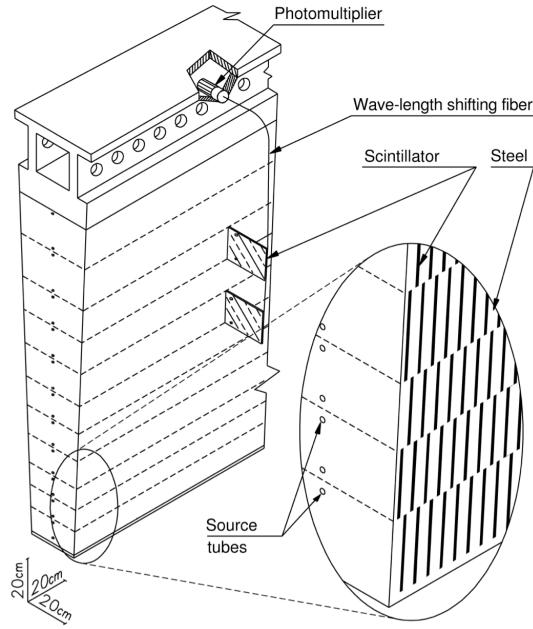


Figure 4.9: TileCal wedge module [40]

in Figure 4.9. A total of 460,000 scintillators are read out by wavelength-shifting fibers. The fibers are gathered to define cells and in turn read out by photomultiplier tubes, which amplify the scintillation light and convert it to an electrical signal. Each cell has an approximate granularity of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ . Each barrel is divided azimuthally into 64 independent modules, an example of which is show in Figure 4.9. The modules are each serviced by front-end electronic housed in a water-cooled drawer on the exterior of the module.

The detection principle of the TileCal is the production of light from hadronic particles interacting with the scintillating tiles. When a hadronic particle hits the steel plate, a shower of particles are produced. The interaction of the shower with the plastic scintillator produces photons, the number and intensity of which are proportional to the original particle's energy.

#### 4.4 Muon Spectrometer

Unlike electrons, photons, and hadrons, muons interact minimally with the ATLAS calorimeters, and can pass through large amounts of detector material without stopping. The ATLAS Muon

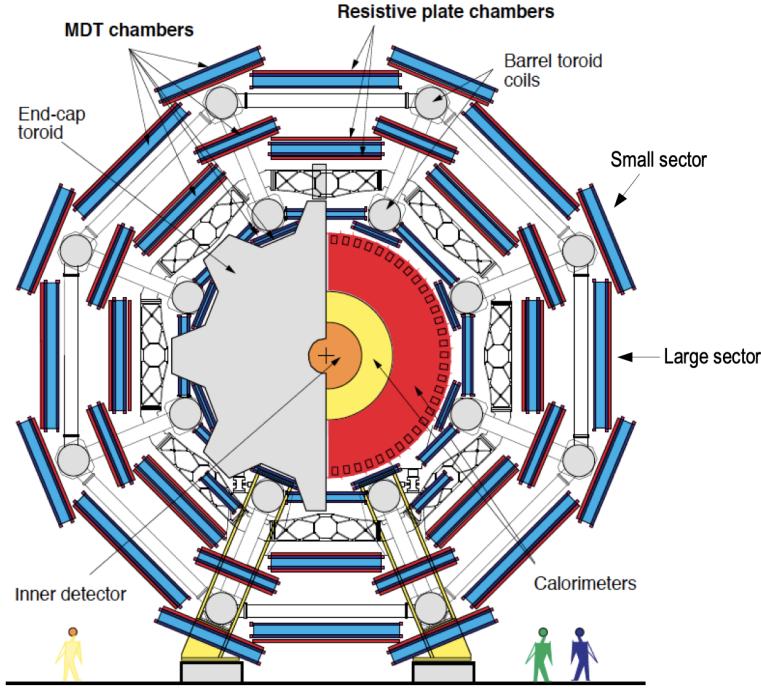


Figure 4.10: Cross section view of the muon spectrometer system [41]

Spectrometer (MS) provides additional tracking information to improve the identification and measurement of muons [41]. The MS comprises the outermost layers of the detector, and is interspersed with toroid magnets (discussed in Section 4.5), which provide a magnetic field of approximately 0.5 T. The magnetic field bends the trajectory of the muons as they pass through the detector, and the degree of the bend is directly correlated with the muon momentum. The path of the muon is primarily measured by hits in three layers of Monitored Drift Tube (MDT) precision chambers, which cover the range  $|\eta| < 2.7$ . The barrel layout of the MS is shown in Figure 4.10.

Muon triggering is provided by three layers of Resistive Plate Chambers (RPC) in the barrel ( $|\eta| < 1.05$ ), and 3 - 4 layers of Thin Gap Chambers (TGC) in the end-caps ( $1.05 < |\eta| < 2.4$ ). RPCs and TGCs also provide muon track measurements in the non-bending coordinate ( $\phi$ ). RPCs are constructed from two parallel resistive plates separated by a 2mm gap filled with a sensitive gas mixture. This provides a total of six independent measurements for each muon track, with a spatial resolution of  $\sim 1$  cm and a time resolution of  $\sim 1$  ns. Time measurements from the RPCs are primarily associated to hits in the MDT precision chambers to determine the bunch crossing.

The time measurement is also used to reject cosmic muons, and to search for delayed signals. TCGs provide triggering in the end-cap regions, and consist of parallel 30  $\mu\text{m}$  wires suspended in a sensitive gas mixture. TCGs provide high radiation tolerance and a fast response time, both features that are necessary for handling the high flux of muons in the forward region.

Precision measurements of muon momentum and position are primarily achieved by MDTs. The MDTs are constructed from 30 mm diameter tubes, permeated by a gas mixture of 93% Ar and 7% CO<sub>2</sub> [42]. The average single-tube spatial resolution is 80  $\mu\text{m}$ . Each chamber consists of six drift tube layers, which together provide a muon track segment resolution of 35  $\mu\text{m}$ . The momentum of the muons can be calculated from the bend in the muon trajectory as they pass through the 0.5T magnetic field provided by the toroids. For a  $p_T = 1$  TeV track, the average  $p_T$  resolution is 11%. In the inner most end-cap wheels, Cathode Strip Chambers (CSC) are used instead of MDTs, covering the region  $2.0 < |\eta| < 2.7$ . CSCs are multi-wire proportional chambers, with a cathode strip readout. The CSCs have a spatial resolution in the range of 50  $\mu\text{m}$ , and a maximum drift time of about 30 ns, which makes them superior for handling the high flux of particles in the forward region.

## 4.5 Magnet System

The ATLAS magnet system consists of four sets of superconducting magnets: a barrel solenoid, a barrel toroid, and two end-cap toroids [43]. The solenoid magnet produces a 2T magnetic field responsible for bending the trajectories of charged particles as they pass through the inner detector. The three toroid magnets provide a field of 0.5 - 1 T and curve the path of muons passing through the muon spectrometer.

The inner solenoid magnet is composed of over 9 km of niobium-titanium superconductor wires, which are embedded into strengthened pure aluminum strips. The solenoid is just 4.5 cm thick, which minimizes interactions between the magnet material and particles passing through the detector. It is housed in the LAr cryostat, as described in section 4.3.1, which further reduces the amount of non-detector material required to support the solenoid. The return yoke of the magnet

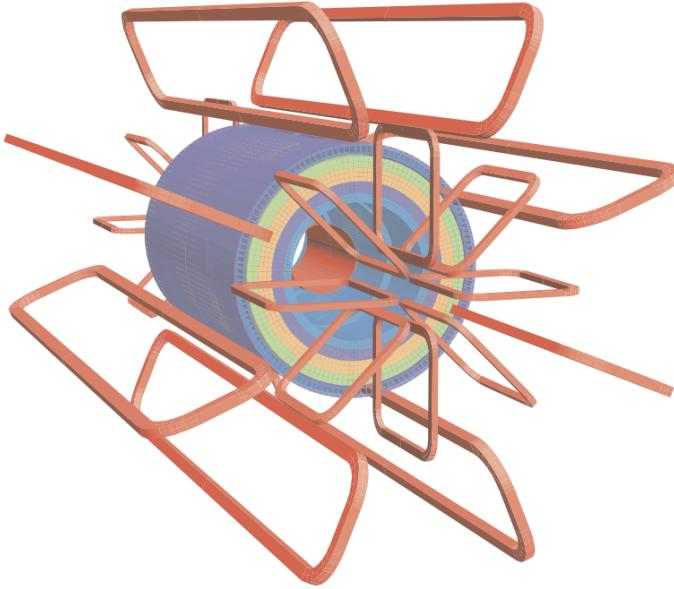


Figure 4.11: Layout of the barrel and end-cap toroid magnets [34]

is provided by the iron absorber of the TileCal.

The central ATLAS toroid magnet, providing the magnetic field for the barrel region of the MS, is the largest toroidal magnet ever constructed at 25.3 m in length [44]. The toroid is composed of eight individual coils, each housed in their own cryostat. The toroidal magnetic field is advantageous as the direction of the field is almost perpendicular to the path of the charged particles. 56 km of aluminum stabilized niobium-titanium-copper superconductor wire compose the magnet. In each end-cap, eight smaller superconducting coils extend the toroidal magnetic field to particles leaving the detector in the forward direction [43]. Figure 4.11 shows the layout of the toroid magnets.

## 4.6 Forward Detectors

In addition to the inner detector, calorimeters, and muon spectrometer, three smaller detectors provide coverage in the very forward region. The innermost forward detector, at 17 m from the interaction point, is the **L**uminosity measurement using **C**erenkov **I**ntegrating **D**etector (LUCID). LUCID's primary purpose is to measure the relative online-luminosity for the ATLAS detector,

from inelastic  $p - p$  scattering. The detector is composed of 20 aluminum Cerenkov tubes which surround the beam pipe and face towards the interaction point.

The second forward detector is the Zero-Degree Calorimeter (ZDC), located 140 m from the interaction point in both directions, at the point where the LHC beam-pipe divides into two separate pipes. The ZDC's primary purpose is to detect forward neutrons from heavy ion collisions.

The third forward detector is the Absolute Luminosity For ATLAS (ALFA) system, located 240 m from the interaction point in both directions. ALFA determines luminosity by measuring elastic scattering at small angles, from which luminosity can be calculated via the optical theorem. The detector is built from scintillating fiber trackers. These are connected to the accelerator vacuum via Roman pots, which allow the detector to come as close as 1mm to the beam without disrupting the machine vacuum. The LUCID and ALFA detectors are crucial to determining the real-time conditions of the beams and the total luminosity delivered to the ATLAS detector [34].

## 4.7 Trigger and Data Acquisition

The trigger and Data Acquisition systems (TDAQ) are responsible for selecting the most interesting events to save for further downstream processing. Because of the high luminosities delivered to the ATLAS detector, not all events recorded can be saved; the 40 MHz bunch crossing rate must be reduced by 5 orders of magnitude to an event storage rate of  $\sim 1$  kHz. The trigger system is composed of three distinct levels: Level 1 (L1), Level 2 (L2) and the event filter [34]. Collectively the L2 trigger and the event filter form the High Level Trigger (HLT).

The L1 trigger is implemented in the hardware of the ATLAS calorimeter and muon systems. A block diagram of the L1 trigger is available in Figure 4.12. The primary modality of the L1 trigger is to identify muons, electrons, photons, jets, and  $\tau$ -leptons with high transverse momentum. Particles with high transverse momentum are more likely to originate from direct, high energy collisions, which are most likely to produce interesting physics processes. The L1 trigger also identifies events with large missing transverse energy, which could be indicative of new physics. The L1 muon trigger (L1Muon) relies on RPC and TGC trigger chambers in the barrel and end-

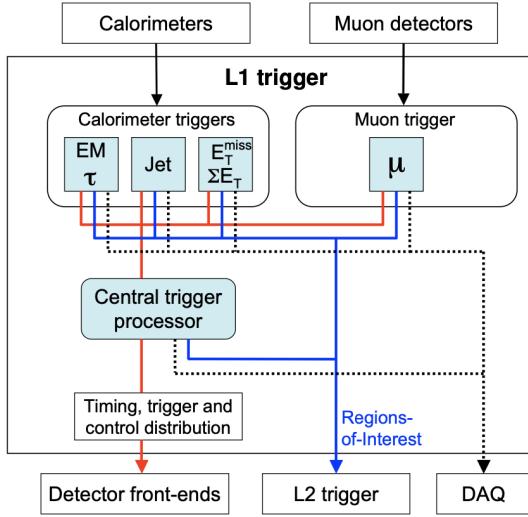


Figure 4.12: Block diagram of the L1 trigger process [34]. The overall L1 trigger decision is made by the CTP.

cap regions of the muon spectrometer. The L1 Calorimeter Trigger (L1Calo) uses reduced granularity information collected by all the calorimeter subsystems. Results from the L1Muon and L1Calo triggers are combined by the Central Trigger Processor (CTP), which implements a trigger ‘menu’, listing various combinations of trigger requirements. The maximum L1 acceptance rate is 100 kHz, and the L1 trigger decision must reach the front-end electronics within  $2.5 \mu\text{s}$  of its associated bunch-crossing.

The L1 trigger defines a Region-of-Interest (RoI) for each passing event. The ROI is represented by the  $\eta$ - $\phi$  detector region where interesting features were identified by the L1 selection process. Information about the type of feature identified and the threshold which was exceeded to trigger the L1 response is also recorded. The ROI data is sent to the L2 trigger, which uses all of the available information within the ROI at full granularity and precision. The L2 trigger reduces the event rate from 100 kHz to 3.5 kHz, with an average processing time of 40 ms. The final stage of the HLT is the event filter, which reduces the event rate to 200 Hz (<1 kHz). The event filter uses an offline analysis process to select fully rebuilt events which will be saved for further analysis. A diagram of the full ATLAS TDAQ system is shown in Figure 4.13.

All levels of the ATLAS trigger system depend on specialized electronics. Each detector front-

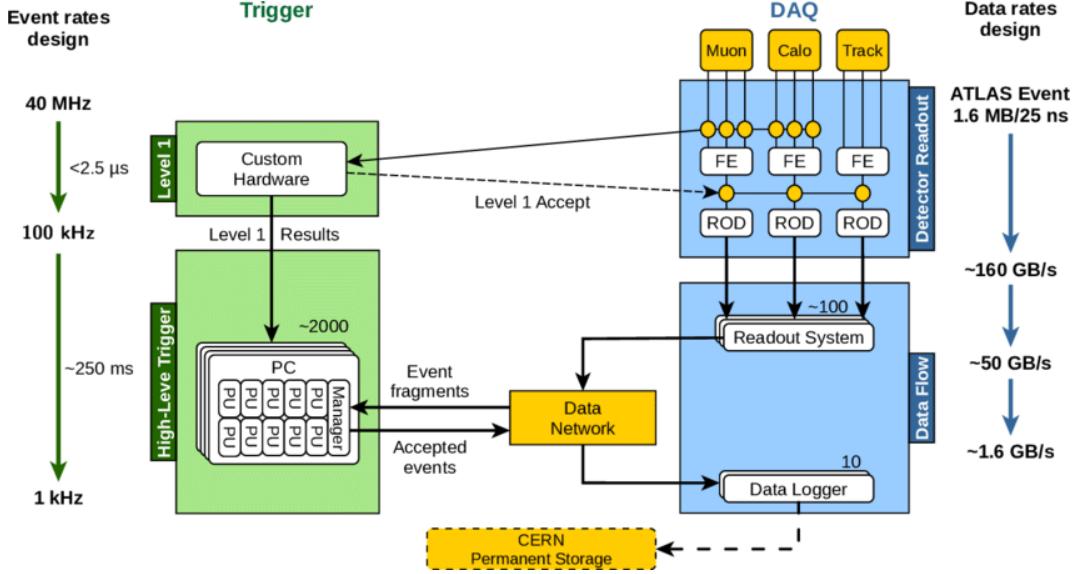


Figure 4.13: Block diagram of the trigger and data acquisition processes, including L1 and HLT processes [45].

end system has a specialized Readout Driver (ROD) which collects information from several front-end data streams at once. The ROD is composed of front-end analogue processing, a L1 buffer which retains the information long enough for the L1 trigger decision, and dedicated links which send the front-end L1 triggered data to Data Acquisition System (DAQ). Any digital signals are formatted as raw data before being transferred to the DAQ. The first stage of the DAQ temporarily stores the L1 data in local buffers. The ROI data is then requested by the L2 trigger, after which selected events are transferred to an event building system, before events passing the event filter are sent to the CERN computer center for permanent storage. The DAQ system not only allows for the readout of detector data, but is also responsible for the monitoring and configuration of the hardware and software components which make up the data readout system via the Detector Control System (DCS).

The DCS allows centralized control of all detector subsystems simultaneously. It continually monitors operational conditions, reports any abnormal behavior to the operator, and can perform both automatic and manual interventions. The DCS reports on real time detector conditions such as high or low voltage detector electronics, gas and cooling systems, magnetic field conditions, humidity and temperature. This information is continually monitored by experts in the ATLAS

control room, so that action can be taken immediately to correct any issues that arise. The DCS also handles communication between detector systems, and other systems such as the LHC accelerator, the ATLAS magnets, and CERN technical services [34].

## Chapter 5: Particle Reconstruction and Identification

With a design luminosity of  $1.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , and a peak Run 2 instantaneous luminosity of  $2.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , reconstructing and identifying the products of LHC  $p p$  collisions is one of the most complex tasks for the ATLAS detector. The accurate reconstruction and identification of *physics objects*<sup>1</sup> lays the ground work for all subsequent physics analyses, so it is also one of the most fundamentally important tasks performed by an experiment.

Reconstruction is the process of combining raw and uncalibrated hits across various subsystems into unique objects. Two particular subsystems, the Inner Detector (ID) tracker and the calorimeters play particularly important roles and will be discussed in detail. Analysis of the properties of the reconstructed objects identifies them as photon, electrons, muons, or jets. While photons, electrons, and muons are fundamental particles, jets represent a collimated shower of many hadronic particles, whose definition is more flexible. Jet reconstruction, clustering and track association are all of particular import to jet identification, and to the later content of this thesis. Finally, reconstruction also identifies *missing transverse energy*  $E_T^{\text{miss}}$  in events, which is a crucial variable for BSM physics searches. Figure 5.1 shows how the physics objects listed here interact with various systems in the ATLAS detector.

### 5.1 Inner Detector Tracks

As the inner most layer of the detector, the ID measures charged particles close to the interaction point. The various hits of these charged particles throughout the ID are used to reconstruct *tracks* which give the trajectories of charged particles [47]. Track reconstruction begins by clustering hits in the Pixel and SCT detectors, and combining clusters from different radial layers of

---

<sup>1</sup>A physics object is a particle or group of particles which has appeared in the ATLAS detector and whose properties have been estimated via the measurements left in various detector subsystems.

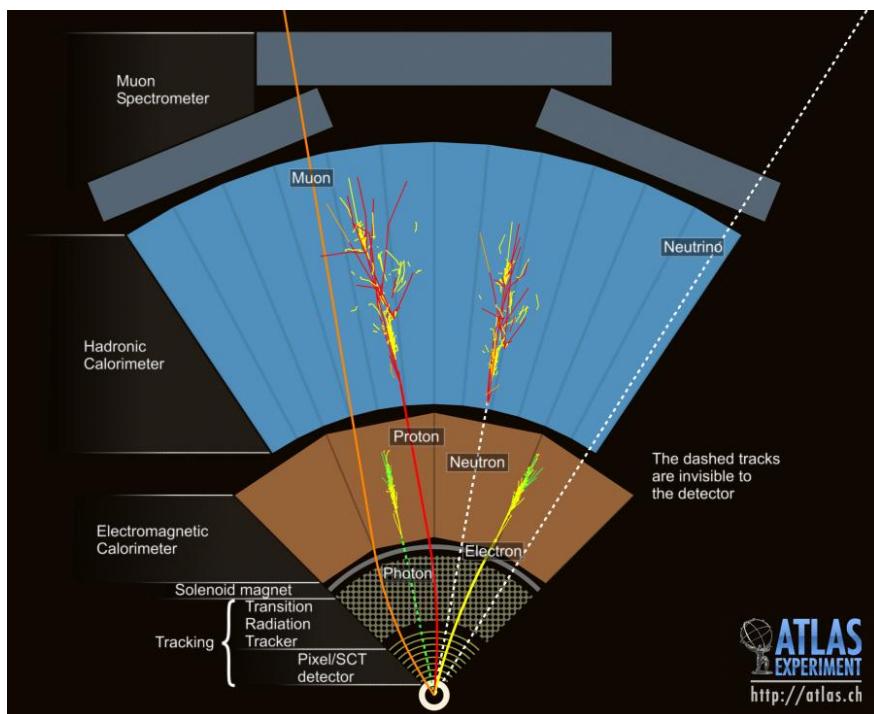


Figure 5.1: This graphic illustrates a slice of ATLAS detector barrel [46]. A photon, electron, and two jets (associated to the proton and neutron) are shown, illustrating each object's interaction with various ATLAS subsystems. The path of the charged particles is curved on account of the strong magnetic field produced by the solenoid and toroid magnets. A neutrino, generally representing the reconstruction of missing transverse energy  $E_T^{\text{miss}}$ , is also illustrated.

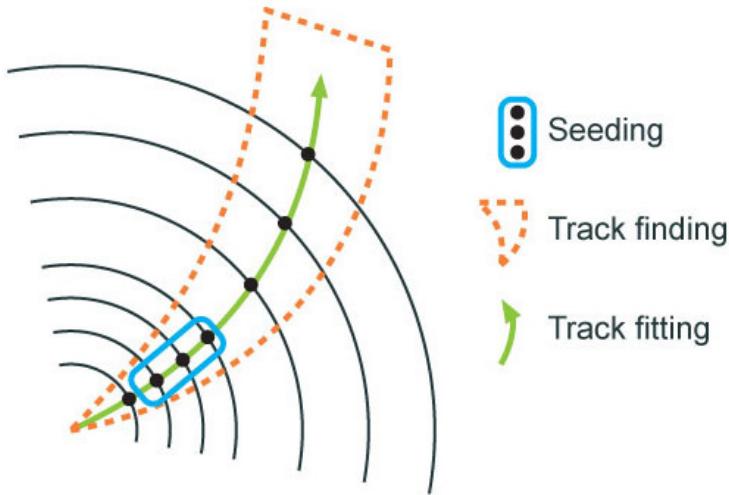


Figure 5.2: Track reconstruction seeding, finding and fitting illustration [47]

these detectors. The multi-layer clusters form track *seeds*, which provide initial estimates of measurements belonging to an individual track. The requirement of three points allows for a rough estimate of the track  $p_T$  to be made by calculating the curvature of the track and accounting of the magnetic field in the ID.

Track seeds are subject to a variety of quality requirements, such as having a minimum estimated  $p_T$  and passing interaction region compatibility criterion. If these requirements are satisfied, the track seeds are passed to the track finding and fitting algorithms. The interplay of these three track reconstruction steps is illustrated in Figure 5.2.

## 5.2 Photons and Electrons

Photons and electrons shower in the LAr calorimeter, and are identified by the energy deposits they leave there. Energy deposits in a collection of nearby cells are termed *proto-clusters*, which become the starting point for electron and photon reconstruction [48]. The clustering algorithm begins when the energy deposit in a certain cell exceeds the noise threshold with a significance of  $4\sigma$ . The algorithm then collects neighboring cells which have an energy deposit exceeding the noise threshold with a significance of  $2\sigma$ , creating a *topo-cluster*<sup>2</sup>. Next, these topo-clusters are

---

<sup>2</sup>A topo-cluser is a topological grouping of neighboring calorimeter cells based on their energy deposits

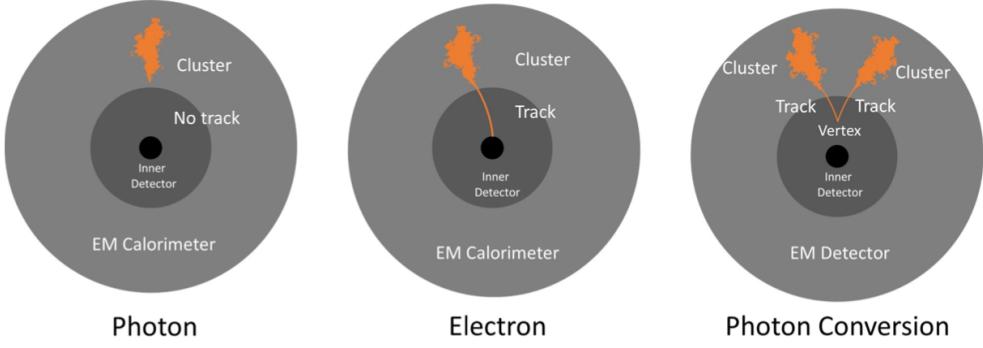


Figure 5.3: Three types of EM object candidates [49].

matched to ID tracks, created as described in Section 5.1. The location of the topo-cluster defines a region of interest (ROI) in the ID, where additional modified track reconstruction algorithms are run in the case that no associated tracks are found. Any ID tracks associated to the topo-cluster are retrofitted to allow for additional energy loss due to bremsstrahlung. A converted photon track reconstruction algorithm is run to check for tracks coming from secondary vertices consistent with converted photons. The secondary vertices are constructed from two oppositely charged tracks consistent with a massless particle, or from one track without any hits in the innermost layer of the ID.

For electron identification, the EM cluster is required to match ID tracks that originate from the primary vertex at the interaction point. For photon identification, the EM cluster can either be matched to tracks coming from a secondary vertex (converted photon), or matched to no tracks (unconverted photon). Figure 5.3 illustrates these three cases for EM object identification.

*Superclusters* are built separately for photons and electrons, based on the combined topo-cluster and ID track information. First, the EM topo-clusters are tested to see if they meet the minimum requirements to become electron or photon seed clusters. For electrons, the cluster must have a minimum  $E_T$  of 1 GeV, and must be matched to a track with at least 4 hits in the silicon tracking detectors. For photons, the cluster must have an  $E_T$  greater than 1.5 GeV. If the seed cluster requirements are met, the algorithm searches for satellite clusters, which can arise from bremsstrahlung radiation. If the satellite clusters pass the positional, energy and tracking requirements to be asso-

ciated with the proto-cluster, they are combined into a supercluster.

Electron and photon objects are identified from the superclusters after the energy calibration is applied, which accounts for the energy resolution of each subdetector measurement. Because photon and electron superclusters are built independently, some clusters can produce both a photon and an electron. In this case an ambiguity resolution procedure is applied to determine if the supercluster can be easily identified as only a photon (no tracks present) or only an electron (good tracks pointing to the primary vertex). In some cases, the identity of the cluster is still ambiguous, in which case both a photon and electron object are created for analysis and flagged as ambiguous. Energy, shower shape, and other analysis variables are calculated from the supercluster and saved with the electron or photon object.

### 5.3 Muons

Muons are identified through the tracks and energy deposits they leave in the ID, calorimeters, and Muon Spectrometer (MS). Muons experience minimum ionizing loss, meaning they do not deposit much of their energy in the calorimeters (recall Figure 5.1), and therefore reach the outer regions of the detector where the MS is housed. Muon identification begins in the Muon Drift Tube chambers by performing a straight line fit between the hits found in each layer, creating *segments*. Segments in the middle layers are then used as seeds for the track building algorithm, which searches for compatible combinations of segments based on their relative positions and angles [50]. A  $\chi^2$  fit is performed on each track candidate. Based on the  $\chi^2$  criteria, hits are removed or added such that the track contains as many hits as possible while satisfying the fit criteria.

The MS track candidates are combined with track information from the ID and calorimeters according to various algorithms based on the information available from each subdetector. Four different types of muons arise from the various reconstruction algorithms:

- Combined muon: a muon track identified through independent track reconstruction in the ID and MS, where the combined track is formed using a global refit that uses hit information

from both detectors. Most muons are constructed through an outside-in procedure, in which a muon track candidate is identified in the MS and then an associated track is found in the ID. A complementary inside-out procedure is also implemented and identifies additional muons.

- Segment-tagged muon: an ID track is identified as a muon if when extrapolated out to the MS (following the inside-out global fit procedure) it is matched to at least one local MS segment.
- Calorimeter-tagged muon: an ID track is identified as a muon if it is matched to a calorimeter energy deposit that is compatible with a minimum-ionizing particle. This muon identification has the lowest purity, but it used in regions where the MS has only partial coverage due to cabling and service access routes.
- Extrapolated muons: the muon is reconstruction only from the MS track and a requirement on compatibility with the primary interaction point. The muon track is required to cross at least two layers of the MS, and three layers in the forward region. These muons are mainly used to extend muon acceptance into the region  $2.5 < |\eta| < 2.7$  where ID track information is not available.

Figure 5.4 illustrates the four types of muon reconstruction. Overlap between reconstructed muons using ID tracks is resolved by giving preference to combined muons, then segment tagged muons, and finally calorimeter tagged muons. Overlap with extrapolated muons is resolved by giving preference to the muon with a better fit quality and higher number of tracks.

All muon track candidates are required to pass a series of quality selections to be identified in the final muon collection. The primary qualities considered are the  $\chi^2$  goodness of fit for the global track, the difference in  $p_T$  measurement between the ID and MS tracks, and the ratio between the charge and momentum of the tracks. The quality requirements help reject hadrons, primarily from kaon and pion decays. Muons candidates consistent with cosmic rays are also rejected.

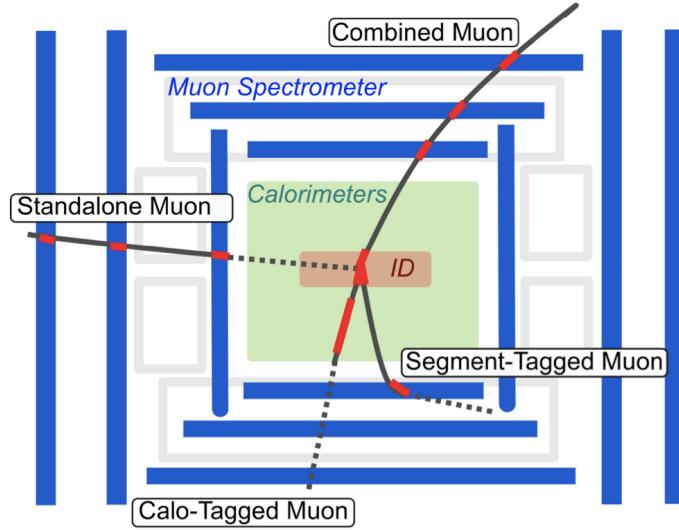


Figure 5.4: The four types of muon track candidates [51]. The red bands indicate where detector measurements for each muon track candidate are made. The dashed lines indicate the path of the muon through detectors where no measurement is made. Standalone muon is another term for an extrapolated muon.

## 5.4 Jets

The protons accelerated in the LHC are composed of quarks and gluons, and thus their collisions often result in the release of energetic quarks and gluons, collectively termed *partons*. The energetic partons can radiate additional gluons, and these gluons can pair produce quarks in a process called *fragmentation*. Fragmentation continues until the energy drops sufficiently that color conservation plays a dominant role. At that point, additional quarks and gluons are produced from vacuum to create neutral color states for the fragmented collection of partons. This process is known as *hadronization* [52]. The hadronized partons compose a collimated stream of particles, known as a *jet*, which is then observed in the detector. The full process that produces jets is known as a *parton shower*, and is illustrated in Figure 5.5.

Jets are identified by the energy deposits they leave in the calorimeter, which are then matched to the tracks they leave in the ID. Jet reconstruction generally begins in the calorimeters, with the identification of *topo-clusters*. Then jet reconstruction algorithms combine calorimeter information with tracking information. There are a variety of jet collections depending on the exact

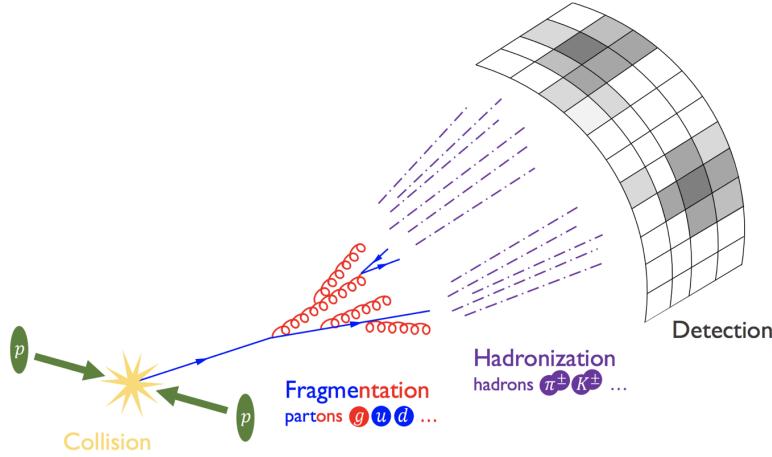


Figure 5.5: The fragmentation and hadronization processes undergone by a quark produced in a proton-proton collision [53].

usage of calorimeter and tracking information in the reconstruction. Some common collections include particle flow jets (PFlow), track calo-cluster jets (TCC), EM topo-cluster jets (EMTopo), and unified flow object jets (UFO). Only particle flow jets will be discussed in greater detail due to their importance in this analysis. The following sections discuss jet identification in the calorimeters, particle flow jet construction using the *anti- $k_t$  algorithm*, jet clustering and jet substructure characteristics.

#### 5.4.1 Calorimeter Clusters

Jets are first identified by the energy deposits they leave in the calorimeters. As for photons and electrons, the reconstruction of jets in ATLAS begins with the construction of topo-clusters, which are topologically-grouped noise-suppressed clusters of calorimeter cells [54]. The topo-cluster seed is a cell with an energy that exceeds the noise threshold for the cell with a significance of at least  $4\sigma$ . Any cells adjacent to the seed cell in three dimensions are added to the cluster if they have an energy deposit of at least  $2\sigma$ . This process is repeated, growing the cluster, until no adjacent cells exceeding the energy deposit threshold remain. As a final step, all adjacent cells are added to the topo-cluster, irrespective of their energy.

The construction process for topo-clusters allows for the possibility that several independent

signatures are grouped into one topo-cluster. To correct for this, the topo-cluster is scanned for local maxima, defined by any cell with energy  $> 500$  MeV, and no neighboring cells with greater energy. If more than one local maximum is identified, the topo-cluster is split among the corresponding energy peaks [55]. In the event that one cell neighbors two or more local maxima, the cell is assigned to the two highest-energy clusters that it neighbors. This means each cell is shared at most once, between at most two post-splitting topo-clusters.

Two measurements for the total energy of the topo-cluster are considered. The raw, or electromagnetic (EM), scale simply considers the sum of energy from all cells in the topo-cluster. The local cell weighting (LCW) scale first classifies clusters as electromagnetic or hadronic, and then applies appropriate corrections for hadronic interactions in the jet energy calculation [54]. The corrections are derived from Monte Carlo simulations, and account for the weaker response of ATLAS calorimeters to hadronic interactions (ATLAS calorimeters are *non-compensating*<sup>3</sup>), and hadronic energy losses due to interactions with dead material [55].

#### 5.4.2 Particle Flow Algorithm

The calorimeters provide excellent jet energy resolution for high energy jets. However, the granularity of the hadronic calorimeter is restricted to  $0.1 \times 0.1$  in  $\eta \times \phi$ . Combining the information from the calorimeter with tracking information provides superior angular resolution and energy resolution. The particle flow (PFlow) algorithm is one of a handful of algorithms which can perform this task.

An overview of the process is given in Figure 5.6. Tracks from the ID which are selected for the PFlow algorithm are required to have at least 9 hits in the silicon detector, and missing pixel hits in places where a hit would be expected. Additionally, the tracks have  $p_T > 0.5$  GeV, and  $|\eta| < 2.5$ . The algorithm then attempts to match these tracks to EM scale calorimeter topo-clusters. This matching is performed using the distance metric

---

<sup>3</sup>The response of ATLAS calorimeters is different for EM showers and hadronic showers, since the calorimeter response to hadronic showers is energy dependent

$$\Delta R' = \sqrt{\left(\frac{\Delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\Delta\eta}{\sigma_\eta}\right)^2} \quad (5.1)$$

where  $\sigma_\eta$  and  $\sigma_\phi$  represent the angular widths of the topo-clusters, and  $\Delta\eta$  and  $\Delta\phi$  represent the distance between the track (extrapolated to the second layer of the EM calorimeter) and the barycenter of the topo-cluster [56]. The topo-cluster closest to the track as measured by  $\Delta R'$  is considered matched to the track. If no topo-cluster is found within the cone size of  $\Delta R' = 1.64$ , it is assumed that particle which left the track did not form a topo-cluster in the calorimeter.

The PFlow algorithm predicts the expected single topo-cluster energy for a given track, based on the track momentum and topo-cluster position. This value is then compared to the observed energy of the topo-cluster, and the probability that the particle energy was deposited in more than one topo-cluster is evaluated. If necessary, the algorithm adds more topo-clusters to the track/topo-cluster system, in order to account of the full shower energy of the track particle.

To reduce the impact of double counting the energy of a given particle by including both its tracker and calorimeter energy measurements, the calorimeter energy measurements associated to a given track are subtracted from the total calorimeter measurement. If the expected energy deposited by the particle exceeds the topo-cluster energy, the full topo-cluster is removed. If the expected energy is less than the EM scale energy of all the considered topo-clusters, topo-cluster cells are removed one by one, until the full expected energy deposit of the particle has been removed from the calorimeter information. The resulting set of tracks and topo-clusters represent the event with no double-counting of energy between subdetectors [56]. This information is passed to the jet-finding algorithm.

#### 5.4.3 Jet Clustering

When a parton decays in the detector, its energy deposits often result in multiple calorimeter clusters. For physics purposes, it is useful to combine clusters that likely resulted from an individual parton decay, in order to reconstruct the parton. The process of grouping topo-clusters which were produced by the same parton decay is *jet clustering*.

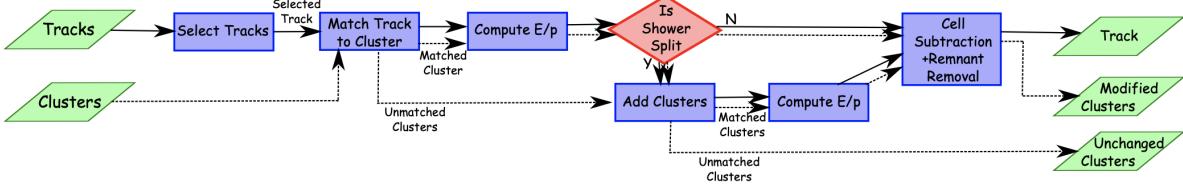


Figure 5.6: A flow chart illustrating the particle flow algorithm progression [56]. The solid lines indicate the progression of tracks through the algorithm, while the dotted lines indicate the progression of clusters. The process begins with track selection and continues until the energy associated with the tracks has been removed from the calorimeter. At the end charged particle tracks, unmodified topo-clusters, and the remnants of topo-clusters which have had part of their energy removed remain.

The anti- $k_t$  algorithm [57] as provided by the FastJet library [58] is most commonly used for jet clustering in the ATLAS experiment, with varying reconstruction radius settings. The anti- $k_t$  algorithm is based on sequential recombination algorithms [59]. A sequential recombination considers the distance  $d_{ij}$  between objects  $i$  and  $j$  (particles or pseudojets), and the distance  $d_{iB}$  between an object  $i$  and the beam line  $B$ . If  $d_{ij}$  between two objects is the smallest distance among those considered,  $i$  and  $j$  are combined into a pseudojet. The process continues until the smallest distance is  $d_{iB}$  at which point the object  $i$  is determined to be a jet and removed from the objects in consideration. The procedure is repeated with the remaining objects until there are none remaining [57].

The anti- $k_t$  algorithm adopts this procedure, but modifies the distance measurements  $d_{ij}$  and  $d_{iB}$  to consider the transverse momentum  $k_t$ :

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad (5.2a)$$

$$d_{iB} = k_{ti}^{2p}. \quad (5.2b)$$

The addition of the term  $p$  allows adjustments to algorithm which govern the relative power of the momentum versus the geometrical scale  $\Delta_{i,j}$ , which is defined as  $\Delta_{i,j} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  where  $y_i$  and  $\phi_i$  are respectively the rapidity and azimuth of particle  $i$  [57]. The radius parameter

$R$  is chosen and determines the geometric cone size [59].

In the case  $p = 1$  the inclusive  $k_t$  algorithm [59] is recovered, which is a standard sequential combination jet clustering algorithm. In the case  $p = 0$ , the Cambridge/Aachen sequential combination algorithm [60] is recovered. The case  $p = -1$  gives rise to the anti- $k_t$  algorithm. The impact of this choice means that the distance  $d_{ij}$  between many soft particles is larger than between soft and hard particles. Therefore, soft particles tend to cluster with hard ones before they cluster with other soft particles. They key feature of this behavior is that soft particles do not modify the shape of the jets. This leads to the creation of circular conical jets, a desirable feature which sequential combination algorithms and cone algorithms struggle to achieve. Figure 5.7 compares anti- $k_t$  jet formation with the inclusive  $k_t$  and Cambridge/Aachen algorithms mentioned here, as well as the SIScone algorithm [61], which checks for sets of stable cones compatible with the observed radiation.

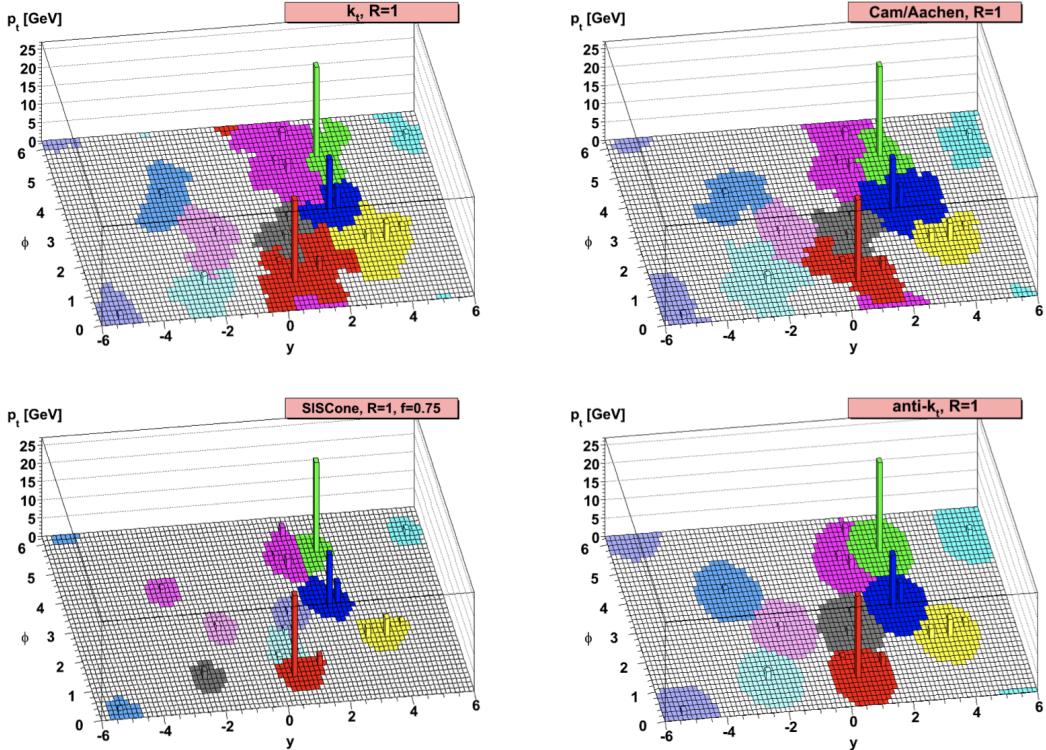


Figure 5.7: A comparison of jet clustering with four different jet algorithms. The anti- $k_t$  algorithm is observed to create the most conical jets, where the shape of the jet is immune to the presence of soft radiation [57].

Any useful jet clustering algorithm must satisfy the requirements of infrared safety and collinear (IRC) safety. Infrared safety implies that the resulting set of jets is unaltered by the presence of additional soft particles in the list of seed clusters. As explained above, the anti- $k_t$  algorithm is naturally infrared safe. Collinear safety requires that the final set of jets is not impacted by collinear splitting of one of the jets. If the hardest particle  $p_1$  is split into a collinear pair ( $p_{1a}, p_{1b}$ ) (as is common in the fragmentation process for a hard parton), the jet clustering algorithm must still recognize ( $p_{1a}, p_{1b}$ ) as the hardest jet in the collision. If another softer particle  $p_2$  with  $p_{t,1a}, p_{t,1b} < p_{t,2} < p_{t,1}$  is instead considered the hardest particle in the event, a different final set of jets would be returned. Collinear safety is a requirement of perturbative QCD, to ensure non-divergent higher-order calculations [62]. The anti- $k_t$  algorithm's tendency to cluster hard particles first ensures its collinear safety. By satisfying the IRC safety requirement, anti- $k_t$  jets can be calculated using perturbative QCD, which improves comparisons with theory.

#### 5.4.4 Ghost Track Association

Once a collection of jets has been created, the jet objects can be studied at both the event-level and the jet-level. In the event-level picture, the momentum, energy, and geometric orientation of the jets within an event are considered. This yields important information about decay of any resonant heavy objects, the total energy in the event, and the distribution of energy amongst the jets. In the jet-level picture, the particle constituents of the jet are considered. The momentum, energy, and geometric orientation of the associated particle tracks provides a low-level picture of the jet, which can help determine if the properties of the jet are consistent with standard QCD, or if new physics processes might be represented within the low-level patterns. Jet-level analysis is also widely used in flavor tagging.

For anti- $k_t$  jets with a radius parameter  $R = 0.4$ , one way of studying the jet-level picture is through considering the ghost-associated tracks. Track association is the process of determining which tracks should be considered associated with a given jet. In the ghost association algorithm, the anti- $k_t$  clustering algorithm is used for the collection of tracks and calorimeter clusters [63].

However, the tracks are considered to have infinitesimal momentum (*ghosts*), so their addition to a jet object does not alter the four-momentum of the jet. This ensures the final jet collection is not altered by the presence of the ghost tracks in the reclustering, but information about the associated tracks for each reconstructed jet becomes available [64].

Ghost tracks are of particular importance to this analysis, as a means of providing a low-level picture of the shape of  $R = 0.4$  jets, and discriminating Standard Model QCD-like jets from dark QCD-like jets.

## 5.5 Missing Transverse Energy

A simple principle leveraged in ATLAS physics analyses is checking for conservation of momentum among the products of any  $pp$  collisions. The initial state transverse momentum of any  $pp$  collision is always zero, so the transverse momentum of all final state particles should likewise be zero. The missing transverse energy,  $E_T^{\text{miss}}$ , is determined by the magnitude of the negative momentum vector sum of all final state objects resulting from the  $pp$  collision.

Specifically, the objects considered in the  $E_T^{\text{miss}}$  calculation are photons, electrons, muons, jets, and soft terms. The first four items comprise the hard components of the  $E_T^{\text{miss}}$  calculation, and have been discussed previously in this chapter. The final item represents a collection of *soft terms*, comprising any detector signals not associated to hard detector objects. These can be based on unassociated tracks, or unassociated soft calorimeter clusters. Both are generally not used in the same calculation to avoid double counting of soft terms. In this analysis the calorimeter cluster soft terms are considered in the  $E_T^{\text{miss}}$  calculation.

$E_T^{\text{miss}}$  can arise due to non-interacting Standard Model objects such as a neutrinos, fake sources such as mis-reconstructed objects and dead detector regions, or in some theories, non-interacting BSM objects such as a dark matter candidate particles. To understand the amount of  $E_T^{\text{miss}}$  attributable to detector noise and mis-reconstruction,  $E_T^{\text{miss}}$  is studied in  $Z \rightarrow \mu\mu$  where little real  $E_T^{\text{miss}}$  is expected [65]. As Figure 5.8 illustrates, the resolution of  $E_T^{\text{miss}}$  generally decreases as  $E_T^{\text{miss}}$  increases, due to detector resolution effects. As  $E_T^{\text{miss}}$  is an important quantity for most dark

QCD analyses, limitations in the accuracy of the  $E_T^{\text{miss}}$  calculation must be considered.

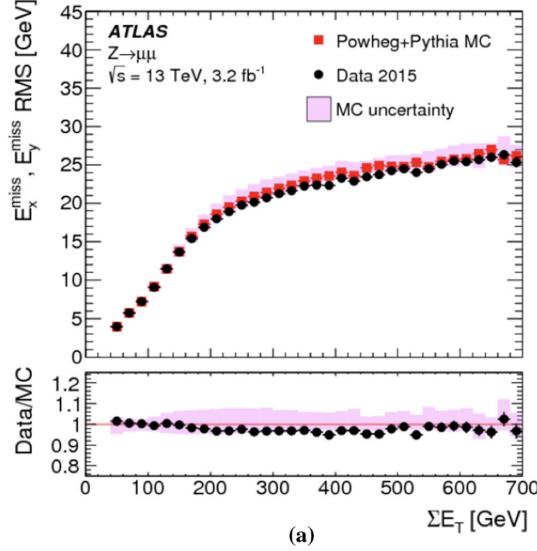


Figure 5.8: A comparison of MC simulation and data for  $Z \rightarrow \mu\mu$  events where real  $E_T^{\text{miss}} = 0$  [65]. The resolution of the missing energy in the transverse ( $x - y$ ) plane is observed to increase with increasing total  $\sum E_T$ .

## **Part III**

### **Search**

## Chapter 6: Monte Carlo and Data

The search for semi-visible jets via s-channel production presented in the following chapters is performed with an integrated luminosity of  $139 \text{ fb}^{-1}$  of proton-proton collision data collected by the ATLAS detector during Run 2 (2015 - 2018). The full Run 2 dataset is used for the final interpretation. Monte Carlo (MC) simulations of background processes and the semi-visible jet signal process are used in the development of the analysis strategy, and in the final interpretation to set limits on the observed cross section of the signal model. This chapter will provide details about the full Run 2 dataset, and the background MC simulations, and the signal MC simulations used in this search.

### 6.1 Data

The  $139 \text{ fb}^{-1}$  integrated luminosity of proton-proton collision data used for physics analyses are required to pass a set of data quality checks. In Run 2 94% of the  $pp$  collisions delivered by the LHC were successfully recorded by the ATLAS experiment, as illustrated in Figure 6.1. 95% of the data recorded by the ATLAS experiment was marked as “good for physics”, resulting in  $139 \text{ fb}^{-1}$  of integrated luminosity. Events are rejected if they are corrupted or incomplete, or if they were recorded during a subsystem malfunction.

Events for this analysis are further required to pass a single-jet trigger selection, where events are required to have at least one jet with a  $p_T$  that exceeds a certain value. The lowest  $p_T$  unprescaled<sup>1</sup> single jet trigger threshold for each period is as follows:

- 2015:  $p_T \geq 360 \text{ GeV}$

---

<sup>1</sup>An unprescaled trigger records every event that meets the trigger requirement. A prescaled trigger only records a fraction of events that meet the trigger requirement.

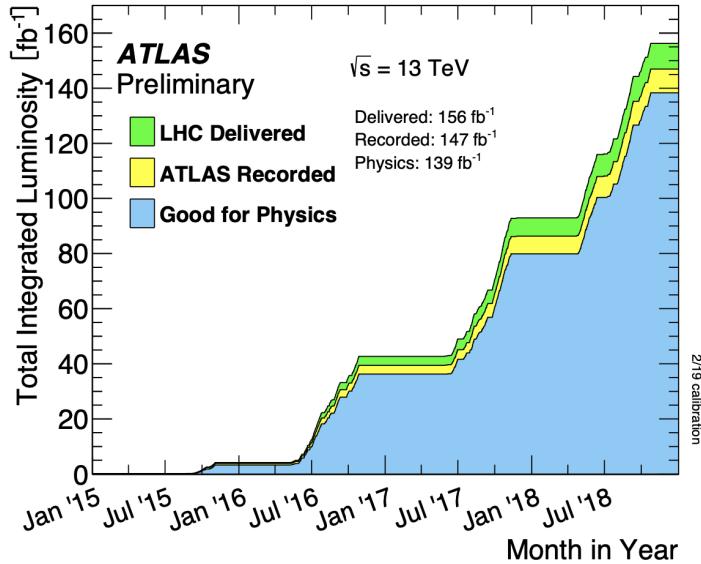


Figure 6.1: Integrated luminosity for the ATLAS experiment as a function of time during Run 2 [66]

- 2016 & 2017:  $p_T \geq 380$  GeV
- 2018:  $p_T \geq 420$  GeV

A post-trigger selection of *leading* jet  $p_T > 450$  GeV ensures these triggers are fully efficient, meaning that the jets are comfortably above the trigger threshold. The jet in the event with the highest  $p_T$  is termed the *leading jet* (or  $j_1$ ), while the jet with the second highest  $p_T$  is termed the *subleading jet* (or  $j_2$ ). The jet collection used is anti- $k_t$  EM particle flow jets with a radius parameter of  $R = 0.4$ , also referred to as small- $R$  jets.

Due to the variance in visible and invisible momenta due to the  $R_{inv}$  parameter of the signal model, many signals also have significant missing transverse energy ( $E_T^{\text{miss}}$ ). The use of a  $E_T^{\text{miss}}$  trigger to select events was considered, and the single jet approach described here was found to preserve more signal events across the grid, particularly in the high  $Z'$  mass and low  $R_{inv}$  region of phase space. These studies are documented in Appendix A.

The data are subject to a blinding strategy throughout the analysis design so as to mitigate analyzer-induced bias. Blinded and unblinded region definitions are described further in Section 8.2.

## 6.2 Simulation

Simulated events are generated with a variety of Monte Carlo (MC) generator processes that run in stages. The  $p p$  hard scatter physics process is simulated, and the final state particles are subsequently showered and decayed. This full description of the event is then propagated through a detailed detector simulation based on GEANT4 [67]. The MC simulation is weighted to match the distribution of the average number of interactions per bunch crossing  $\mu$  observed in collision data.

All simulated samples included in this analysis were produced with three different MC campaigns: `mc20a` corresponds to 2015-2016 data-taking conditions, `mc20d` to 2017, and `mc20e` to 2018. These three campaigns are weighted to the integrated luminosities of their respective data-taking periods and combined to produce simulation for the entire Run 2 dataset. Simulated events are reconstructed with the same algorithms run on collision data.

### 6.2.1 Simulated Backgrounds

Although the final background estimation is data-driven, MC simulated background is also studied for analysis optimization and machine learning tool development.

Dijet QCD is the dominant background process. QCD is simulated with PYTHIA8 [68], and generated in approximate slices of  $p_T$ , to ensure high statistics across the momentum spectrum. The slices are then reweighted using MC generated event weights to create a physical distribution. Figure 6.2 illustrates the 8 momentum slices used in this analysis.

Due to presence of  $E_T^{\text{miss}}$  in the semi-visible jet (SVJ) signals, additional MC background processes are required to create a full picture of the relevant background. The  $Z \rightarrow \nu\nu$  process contributes to the background due to its high missing energy. Leptonic W/Z decays and W/Z+jets are also included as they can contribute both additional missing energy and significant hadronic activity. Single top and  $t\bar{t}$  processes are also considered for their contribution to hadronic activity.

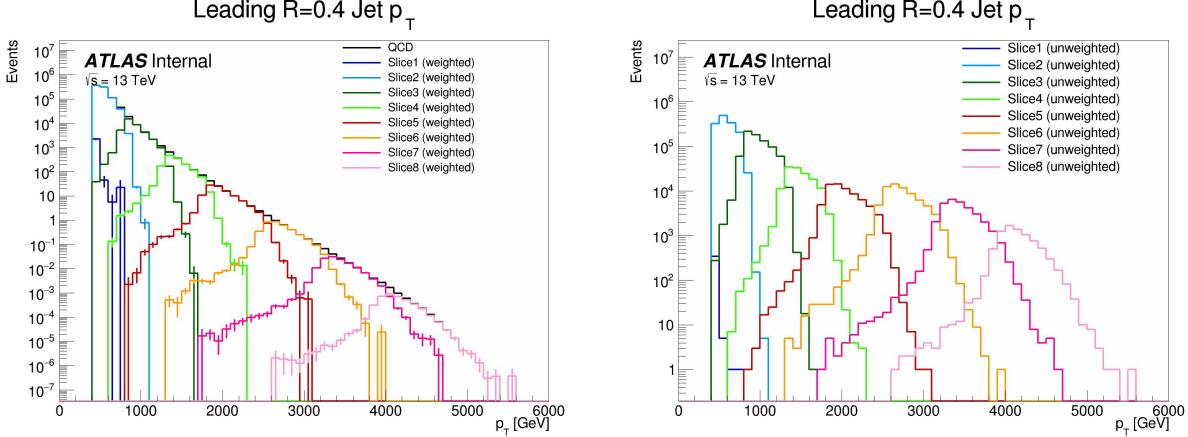


Figure 6.2: The transverse momentum slices of the QCD MC simulation, overlayed to show how they come together to create a smooth distribution (left) once weighted properly. The original unweighted distribution (right) illustrates the enhanced statistics for the high  $p_T$  range.

After the analysis *preselection*<sup>2</sup> is applied to isolate events most relevant to the SVJ topology, the background composition is 76% QCD, 12%  $W+jets$ , 8% top and  $t\bar{t}$  processes, and 4%  $Z \rightarrow \nu\nu$ . Figure 6.3 illustrates the background composition for the analysis. The lower panel in Figure 6.3 illustrates the ratio between data (black) and the combined MC processes (grey). While the agreement between data and MC is not perfect (ratio = 1.0 for all  $E_T^{\text{miss}}$  values), the difference is < 20% throughout the distribution. This is within tolerance for this analysis, since the final background estimation will be data driven, and background MC is only needed for approximate modeling. Analysis selections for high energy jets (discussed in Section 8.2) create some sculpting in the  $Z \rightarrow \nu\nu$  and  $W+jets$  distributions; however, the total  $E_T^{\text{miss}}$  distribution is smoothly falling so this is not an issue.

### 6.2.2 Signal Simulation

The Hidden Valley (HV) signal model implementation is based on Ref [22]. The s-channel semi-visible jet model, which was described in Chapter 2, is governed by a number of parameters. The mass of the mediator  $m_{Z'}$  can be set, together with the couplings of the  $Z'$  to the visible and dark quarks  $g_q$  and  $g_{q_D}$ . The dark sector shower is governed by the number of dark colors  $N_{c_D}$ ,

<sup>2</sup>A preselection is a set of cuts on physical observables used to isolate a collection of events which are most likely to contain the desired signal. The preselection for this analysis will be discussed in Section 8.2

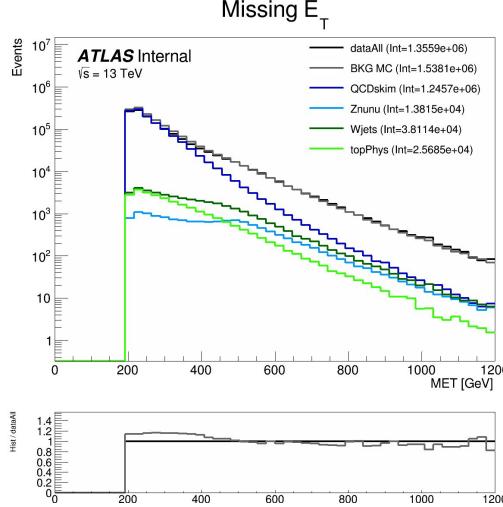


Figure 6.3: Background processes relevant to the SVJ signal.

the number of dark flavors  $N_{f_D}$ , and the dark sector confinement scale  $\Lambda_D$ . There is also the characteristic scale of the dark hadrons  $m_{dark}$ , determined by the mass of the dark quarks  $m_{q_D}$ . The characteristic scale determines the mass of the dark hadrons, which can be pseudoscalars  $m_{\pi_D}$  or vectors  $m_{\rho_D}$ . Finally, the average fraction of invisible particles in the final state jet is dictated by  $R_{inv}$ .

The chosen parameters for this model were carefully selected in collaboration with theorists to be compatible with the new benchmarks established in the 2021 Snowmass process [21]. The signal generation allows for up to two initial state radiation jets, and uses a jet-matching scheme described in Ref. [69] and implemented with PYTHIA MLM [68] to match jets to the original partons.

The choices of fixed parameters for the Pythia8 HV model are summarized in Table 6.1. A detailed discussion of these parameters and their implications on the dark shower topology can be found in Ref. [21]. The mass choices for the dark quark and the dark hadrons are also summarized in Table 6.2.

Note that the number of dark flavors differs from the Snowmass [21] recommendation of  $N_{f_D} = 4$ . This change is minimal in impact because  $R_{inv}$  is set explicitly (rather than allowing it to arise naturally from the HV theory), and allows this ATLAS analysis result to remain comparable

Parameter	Value
$N_{c_D}$	3.0
$\Lambda_D$	10.0 GeV
$N_{f_D}$	2.0
$g_q$	0.25
$g_{q_D}$	1.0

Table 6.1: Fixed parameters in the Pythia8 HV model

Parameter	Value [GeV]
$m_{\pi_D}$	17.0
$m_{\rho_D}$	31.77
$m_{q_D}$	10.0

Table 6.2: Values for  $m_{dark}$

with the CMS semi-visible jets s-channel analysis [70] and the ATLAS semi-visible jets t-channel analysis [71].

The mediator mass  $m_{Z'}$  and the fraction of invisible particles in the final state  $R_{inv}$  vary, and are used to define the search grid.  $m_{Z'}$  varies between 2.0 TeV and 5.0 TeV, while  $R_{inv}$  varies from 0.2 to 0.8.  $R_{inv}$  values of 0.2, 0.4, 0.6, and 0.8 are generated for each  $m_{Z'}$  mass point. Table 6.3 illustrates the signal grid and the associated cross section for each signal. There are a total of 24 signal points (6  $Z'$  masses  $\times$  4  $R_{inv}$  settings) considered in this analysis.

Samples are generated using MADGRAPH5 [72] version 2.9.9 interfaced to PYTHIA8.244P3 [68] for shower and hadronization with NNPDF23LO PDF [73] and the ATLAS A14 [74] to tune the underlying event data.

$m_{Z'}$ (GeV)	Cross section (fb)
2000	252
2500	74.2
3000	24.5
3500	8.83
4000	3.49
5000	0.757

Table 6.3: Mass points and cross sections of the SVJ search signal grid. The cross section is determined by the  $Z'$  mass and its SM coupling  $g_q$ ; it is not impacted by the  $R_{inv}$  fraction.

## Chapter 7: Machine Learning Tools

### 7.1 Introduction

The search for semi-visible jets presents an opportunity to use novel machine learning (ML) tools to uncover patterns in the behavior of dark QCD. The subtlety of the shower differences between dark QCD (signal) and SM QCD (background) motivates a complex model that can accept high-dimensional low-level information, such as particle track information, to understand key differences between signal and background patterns. Additionally, the large number of theory parameters which can be chosen arbitrarily and affect the shape of the dark QCD shower motivate exploring a data-driven machine learning approach, which could be sensitive to a wide variety of dark QCD behavior.

To this end, two machine learning approaches are developed for this search, which are used in tandem. The first is a *supervised* ML method where the ML algorithm is built to maximize sensitivity to the specific SVJ signal models described in Section 6.2.2. Here, supervised refers to the use of full and correct *labels*<sup>1</sup> for all events considered during model training. The second is a semi-supervised method, where training of the model is data-driven (no signal hypothesis used during training) and labels are only partially provided during training. The semi-supervised ML algorithm broadens the discovery sensitivity of the search, and reduces the dependence on the exact theory parameters chosen for signal model simulation.

The two different ML algorithms used in this approach will be explored in Section 7.2 and Section 7.3, along with their application in the SVJ analysis strategy. In the following Section 7.1.1, a brief overview of fundamental machine learning concepts is presented.

---

<sup>1</sup>In machine learning a label refers to the correct identification information for an input. In the case of the binary classifier algorithm discussed here, the label is either “signal” or “background”.

### 7.1.1 Machine Learning Fundamentals

The machine learning tools presented in this chapter depend on two basic *architectures*. An ML architecture refers to the specific neural network design used to create an ML *algorithm* (or *tool*).

The first basic architecture is a deep neural network (DNN) [75]. Figure 7.1 illustrates the concept. The hidden layers of a DNN allow the network to store information about the importance of each input feature and the importance of correlations amongst the input features. The elements of each layer are known as *nodes*. In a fully connected network like the one shown each node receives input from every node in the previous layer, represented by the arrows in the diagram. Each node input has an associated *weight* which is adjusted during training. The node combines the inputs and their associated weights according to an *activation function*. The output of the activation function becomes the value associated to the node, which is then used as input to the subsequent layer.

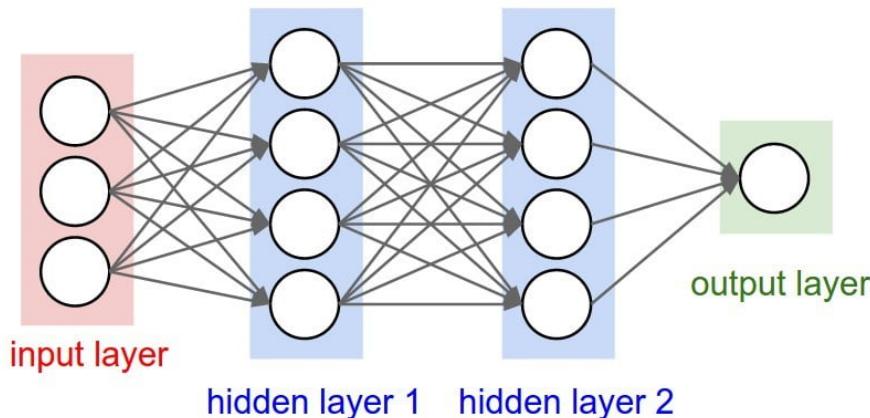


Figure 7.1: A diagram of a deep neural network architecture.

A *loss function* measures the performance of the model. The *loss* calculated by the loss function compares the output of the model to the correct response; a lower loss indicates better performance. In a *classifier* model the output layer is the probability that the input fits a certain category, for example “signal” (1.0) or “background” (0.0). This probability is called the *score*. The loss function calculates the accuracy of the scores. For example a signal input that receives a score of

0.9 would result in a small loss, while the same event given a score of 0.1 would result in a large loss. Figure 7.2 illustrates a typical classifier score response.

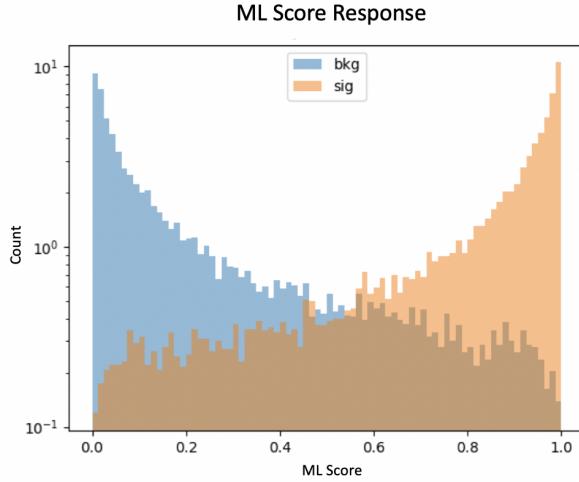


Figure 7.2: An example score distribution for a binary classifier. A higher score indicates a greater probability of the event being signal-like. Most signal events (orange) receive a high score while most background events (blue) receive a low score, indicating good classification.

The network improves by training over many *epochs*, which refers to the process of the ML algorithm evaluating all training events. After each epoch, the *optimizer* adjusts the weights to reduce the loss. The *learning rate* determines how big of an adjustment the network is allowed to make. During training, a set of events are set aside to use for *validation*. The purpose of the validation data is to prevent *overtraining*. If a network is sufficiently large and complex, the network could lose generality by perfectly learning (or “memorizing”) the correct response for every training event. This would minimize the training loss, but could result in the network failing to correctly classify events it hasn’t seen before. By evaluating the loss of the validation data the user can determine if the network is overtrained; the validation loss should not greatly exceed the training loss.

ML algorithms are often evaluated through a *receiver operating characteristic* (ROC). The ROC compares the true positive rate (correct classification) with the false positive rate (false classification). An example ROC curve is shown in Figure 7.3. If a classifier is performant, the true positive rate will be larger than the false positive rate for all possible false positive rates. If

the network has no classifying power, the true positive rate and false positive rate will be equal throughout. The *area under the curve* (AUC) is an important metric for evaluating the ROC. The AUC is the integral of the ROC curve. An AUC of 1.0 indicates perfect performance, while an AUC of 0.5 indicates that the network is no better than random guessing.

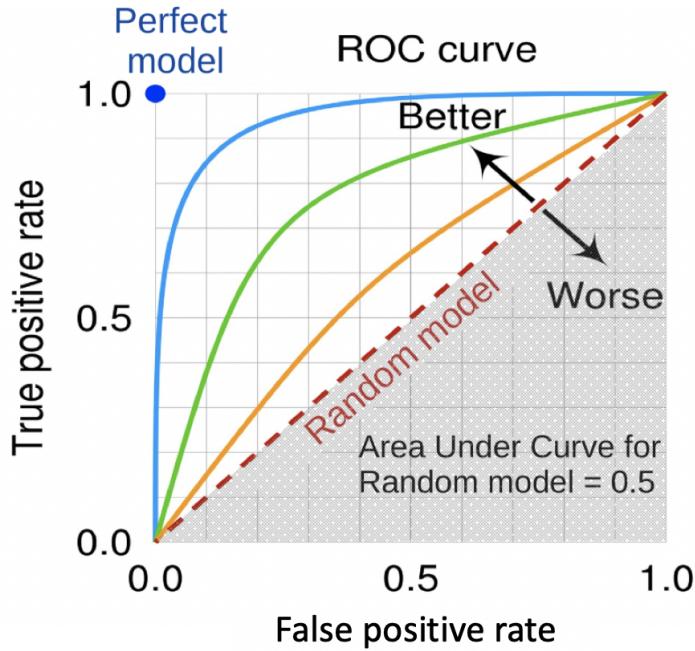


Figure 7.3: Several example ROC curves. The AUC is also illustrated [76].

The second architecture that is important to this thesis is the auto-encoder (AE) [77]. Unlike a DNN, which is a supervised network that depends on the use of correct labels to determine the loss, the AE calculates loss by comparing the input and output layers. Figure 7.4 illustrates the concept. The network is designed to extract the most salient features of the input via dimensionality reduction. This is achieved by compressing the input to a lower dimensional *latent space*, and then attempting to reconstruct the original input from that latent space. The loss is calculated by comparing the output of the network with the input. While the goal of a classifier is to correctly categorize the inputs, the goal of the AE is to correctly reconstruct the inputs. This allows the AE to be used for *anomaly detection*. The kinds of events that are seen most often during training will be reconstructed well by the algorithm, and therefore have the smallest loss. Events which

are anomalous or unusual in the training data will be more difficult for the AE to reconstruct, and therefore receive a larger loss. The loss can be used to create an *anomaly score*, which identifies unusual events with a higher anomaly score.

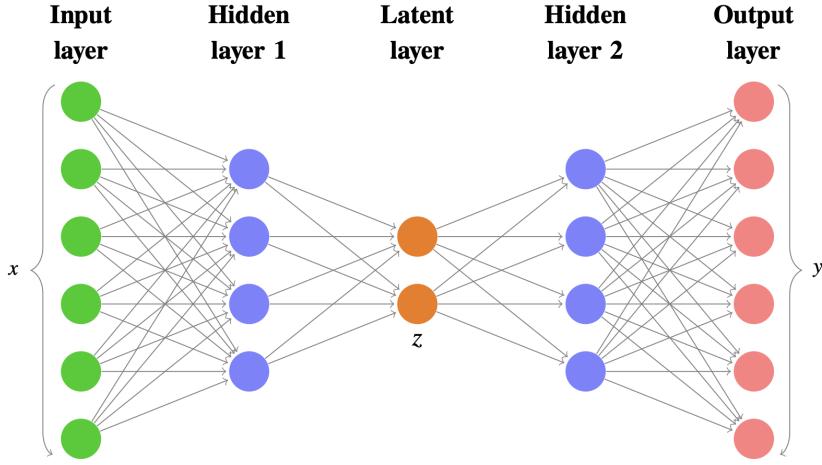


Figure 7.4: A diagram of auto-encoder architecture. The loss is computed as a difference (often the *mean squared error* or MSE) between the input  $x$  and the output  $y$  [78].

## 7.2 Particle Flow Network (Supervised)

The supervised machine learning approach maximizes discovery sensitivity for the SVJ signals considered in this thesis. The networks learn the features of the SVJ signals, allowing the network to be highly efficient in selecting events that resemble the SVJ signal.

### 7.2.1 Architecture Fundamentals

A Particle Flow Network (PFN) [79] architecture is selected for two reasons: *permutation invariant input modeling* to best describe the events consisting of an unordered set of particles, and a *low-level input modeling* to take advantage of the ability of neural networks to uncover patterns in high-dimensional data. *Low-level* refers to using detector level information such as individual particle tracks, rather than *high-level* information such as reconstructed jet objects. Low-level inputs are generally high-dimensional; for instance, an event may have only 2 jets (dim-2), but each jet consists of 70 particles (dim-140). Low-level input modeling is chosen to capture the

intricacies of dark QCD showers which may not express themselves in high level objects, as explored in Ref. [22]. Permutation invariant input modeling is chosen as the most accurate representation of a set of particles. In previous work such as Ref. [78], ordered input modeling has been observed to *bias* the performance of low-level modeling tools. In this case bias means that the performance of the tool was observed to change substantially depending on the input ordering; however, there is no physics motivation for choosing any particular order.

The input to the PFN is a collection of particles and their associated physics information, such as momentum and trajectory. Constructing the PFN involves the creation of new basis variables  $\Phi$  for each particle in the input event. This transformation is summarized as  $\vec{p}_i \rightarrow \vec{\Phi}_i$  where  $\vec{p}_i$  is the physics information for the  $i$ th particle in the event, and  $\vec{\Phi}_i$  is that same information encoded into the  $\Phi$  basis. Permutation invariance is enforced by summing over the  $\Phi$  basis for every particle in the event to create a new permutation invariant event representation  $O$ . The creation of  $O$  from  $M$  particles  $\vec{p}$  with  $d$  physics features each can be expressed as:

$$O(\{\vec{p}_1, \dots, \vec{p}_M\}) = \sum_{i=1}^M \Phi_i(\vec{p}_i) \quad (7.1)$$

where  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^l$  is a per particle mapping, with  $l$  being the dimension of the new basis  $O$ . Figure 7.5 gives a graphical representation of the use of summation in the PFN over per-particle information to create a permutation-invariant event representation.

The  $\Phi$  basis transformation is implemented via a deep neural network. The output of the neural network is summed as indicated in Equation 7.1 to create the new permutation invariant event representation  $O$ .  $O$  then becomes the input of a second deep neural network  $F$ .  $F$  is a classifier network which separates signal and background events. Figure 7.6 provides an annotated diagram of the PFN architecture as used in this analysis.

### 7.2.2 Input Modeling, Scaling, and Rotation

In this implementation, the particle input information comes from all tracks associated to the leading and subleading jets. The track association method is Ghost association, as discussed in

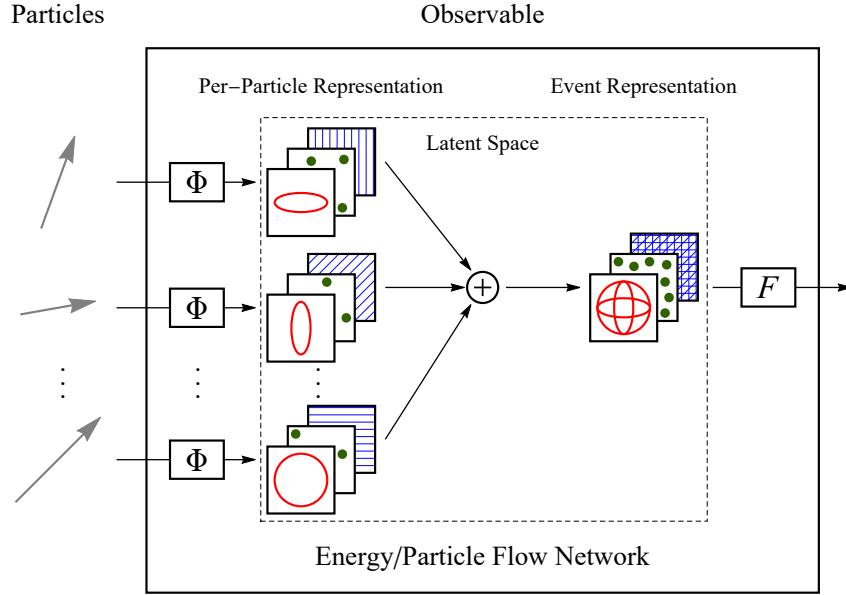


Figure 7.5: The Energy/Particle Flow Network concept, from Ref. [79]. The physics input information is represented as arrows on the left, for an arbitrary number of particles. The  $\Phi$  transformation converts these arrows to 3 graphs, indicating the  $\Phi$  basis dimension  $l$  is 3 in this example. The graphs are then summed for all particles to create  $O$ , or the event representation.

Section 5.4.4. A single jet tagger strategy was also considered, but utilizing tracks from both leading jets creates a more complete low-level picture of the event. The choice of the two leading jets is justified in Chapter 8. If we consider the dijet topology of semi-visible jets as illustrated in Figure 7.7, the advantage of modeling both leading jets simultaneously becomes clear. In the semi-visible jet model presented in Ref. [22],  $E_T^{\text{miss}}$  in the event is expected to arise due to an imbalance in the number of visible tracks of the two jets associated to the dark quark decay.

Each track is described using six variables: the four-vector of the track ( $p_T, \eta, \phi, E$ ), and the track displacement parameters  $d_0$  and  $z_0$ , where  $d_0$  measures displacement in the radial direction from the beamline and  $z_0$  measures displacement along the beamline from the primary interaction point. Figure 7.8 illustrates these coordinates. Up to 80 tracks per jet are allowed, which is a threshold chosen to generally include all the tracks in the jet, which leads to maximal performance.

These tracks (up to 160 total) are the input to the PFN. Referencing Equation 7.1, this corresponds to  $M = 160$  (number of particles) and  $d = 6$  (number of features per particle). The two leading jets and their associated tracks are rotated so that the vector sum of the jets, or system

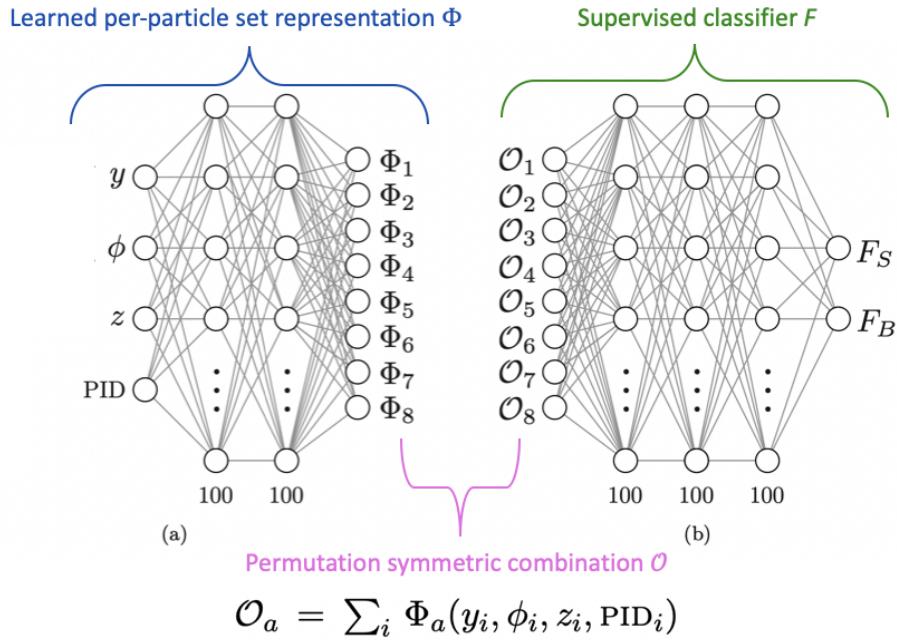


Figure 7.6: An annotated diagram of the PFN architecture [79].  $y$  and  $\phi$  represent geometric trajectory information for the input particles,  $z$  represents energy information, and PID encompasses any other particle ID information in the input. PID is presented in the diagram as a 1-dimensional input, but could represent multiple input dimensions.

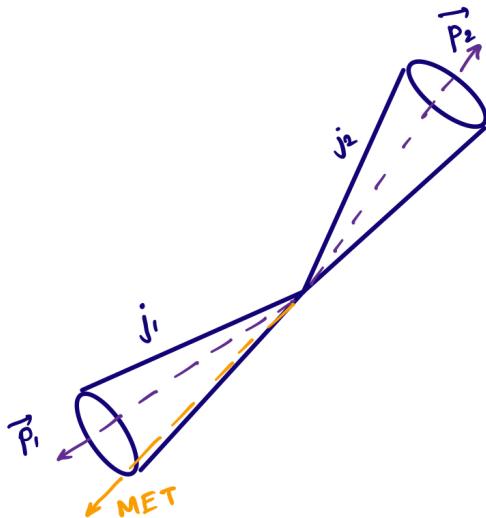


Figure 7.7: An illustration of the expected dijet behavior of semi-visible jets, where one jet is closely aligned with  $E_T^{\text{miss}}(\text{MET})$ . In the figure two jet cones  $j_1$  and  $j_2$  are illustrated, along with their associated momentum vectors  $\vec{p}_1$  and  $\vec{p}_2$ .

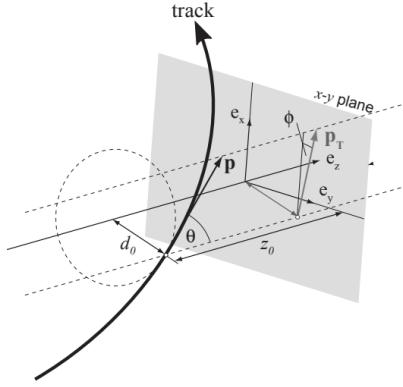


Figure 7.8: Illustration of track coordinates  $d_0$  and  $z_0$ .

average, is aligned with  $(\eta, \phi) = (0, 0)$ . The rotation can be summarized as

$$\eta'_i = \eta_i - \bar{\eta}, \quad (7.2a)$$

$$\phi'_i = \phi_i - \bar{\phi} \quad (7.2b)$$

where  $(\bar{\eta}, \bar{\phi})$  is the average angle of the dijet system,  $(\eta_i, \phi_i)$  are the original track coordinates, and  $(\eta'_i, \phi'_i)$  are the rotated track coordinates. Figure 7.9 illustrates the rotation process. The rotation ensures that the information used by the algorithm is the relative orientation of the jets (and associated tracks) to each other, not their absolute position in the detector. Each track is normalized to its relative fraction of the total dijet system energy and transverse momentum; this enforces agnosticism to the total energy and transverse momentum of the event. The rotation and scaling are motivated by the procedures described in Ref. [79] to improve the performance of the PFN.

Finally, each of the 6 track variables is scaled so that its range is [0,1]. This is a common preprocessing step that ensures the input data is bounded over a similar range, so that arbitrarily large values don't develop an outsized impact on the model. The track momentum and energy normalization mentioned above naturally enforces that these values are restrained between [0,1]. The  $\eta$  and  $\phi$  values are naturally bounded, so for these values the  $\eta$  tracking range<sup>2</sup> of [-2.5, 2.5] and the full  $\phi$  range  $[-\pi, \pi]$  are mapped to [0,1]. The displacement variables are restricted to [0,1]

---

<sup>2</sup>This range is dictated by the  $|\eta|$  coverage range of the Inner Detector, as shown in Table 4.1

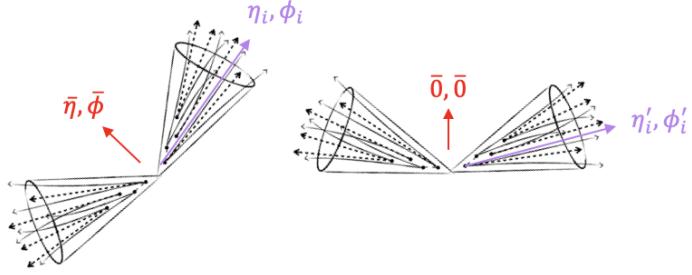


Figure 7.9: A diagram demonstrating how the two jet system is rotated in  $(\eta, \phi)$ . The jet cones and associated jet tracks are illustrated. The dashed tracks represent dark hadrons while the solid tracks represent SM hadrons. The system average  $(\bar{\eta}, \bar{\phi})$  is shown in red and an example track with coordinates  $(\eta_i, \phi_i)$  is shown in purple.

via the standard MINMAXSCALER [80] method which determines the minimum and maximum values observed in training, and maps those boundaries to 0 and 1 respectively.

Figure 7.10 illustrates that the data is well modeled by the MC at track level. Figure 7.11 shows the kinematics of each of 6 track variables for background and signal. Figure 7.12 shows each of the 6 track variables after scaling and rotation have been applied, demonstrating the impact of these procedures, as well as the track level similarities and differences between the background SM QCD processes and the signal SVJ processes.

The  $\phi$  distribution is of note for its jagged appearance in QCD MC. This arises due to dead tile calorimeter cells in certain  $\phi$  regions, the effects of which are seen in data and modeled in QCD MC but not modeled in SVJ signal MC. Appendix C.3 contains more information about how the issue was addressed in data. The distribution is not of concern for the PFN training because of the rotation process, which replaces the information about absolute detector  $\phi$  measurements with the relative  $\phi'$  measurement. This is illustrated in Figure 7.12, where it is observed that for both signal and background the tracks are clustered back to back, centered at  $-\pi/2$  and  $\pi/2$  (0.25 and 0.75 after scaling). The only remaining difference is that the signal tracks are more likely to be close to the system average  $\bar{\phi}$  than the background jet tracks. This is demonstrated by the excess of signal events in the center of the  $\phi'$  plot. This orientation difference is a real feature of the signal

model, confirmed in Figure 8.3 which illustrates that signal jets are more likely to have low  $\Delta\phi$  than background jets.

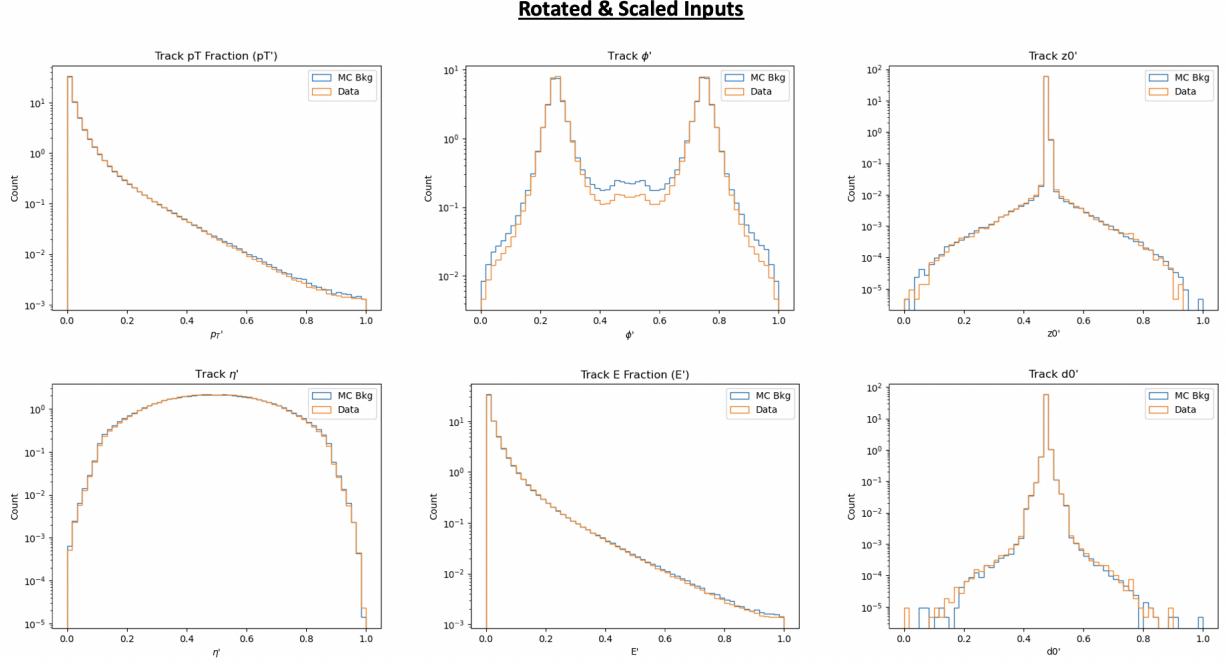


Figure 7.10: The 6 PFN track variables in background MC (blue) and data (orange), after the scaling and rotation procedure is applied. There is excellent modeling of the data by the MC within the track variables. The slight discrepancy in the  $\phi$  distribution due to the inaccuracies of modeling dead TileCal cells in the QCD MC is considered. The level of discrepancy is determined to be within tolerance given that the final result will be data driven and the QCD model is used in the PFN training only.

### Kinematic Inputs

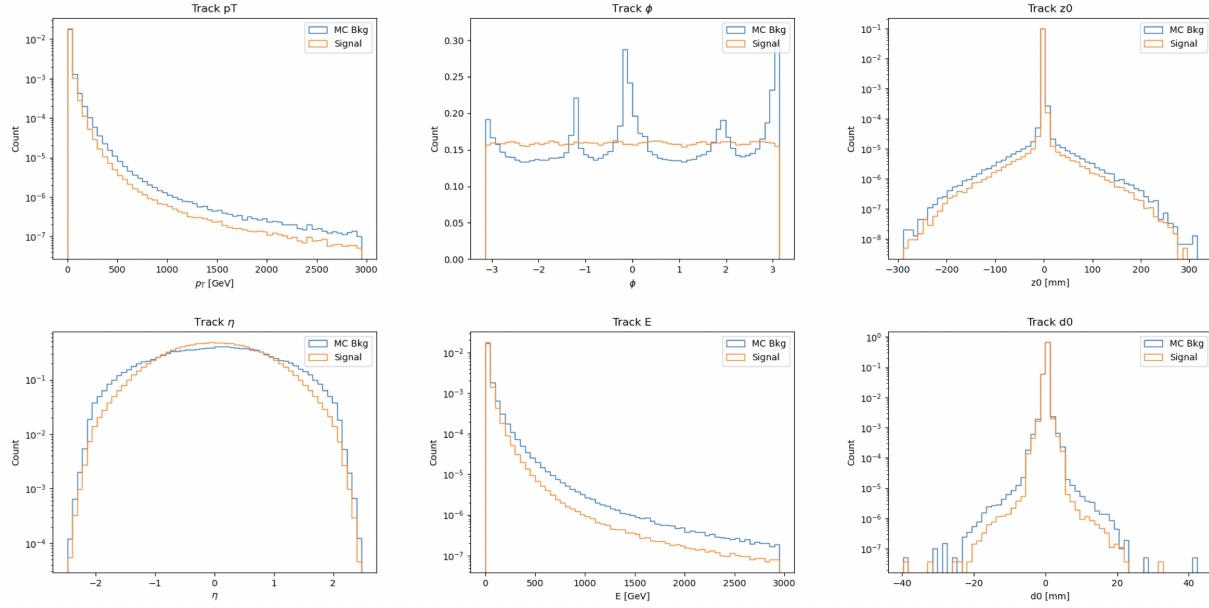


Figure 7.11: The 6 PFN track variables in background MC (blue) and signal MC (orange) before scaling and rotation. The track kinematics are largely similar, and the variation in the  $\phi$  distribution is explained in the text.

### Rotated & Scaled Inputs

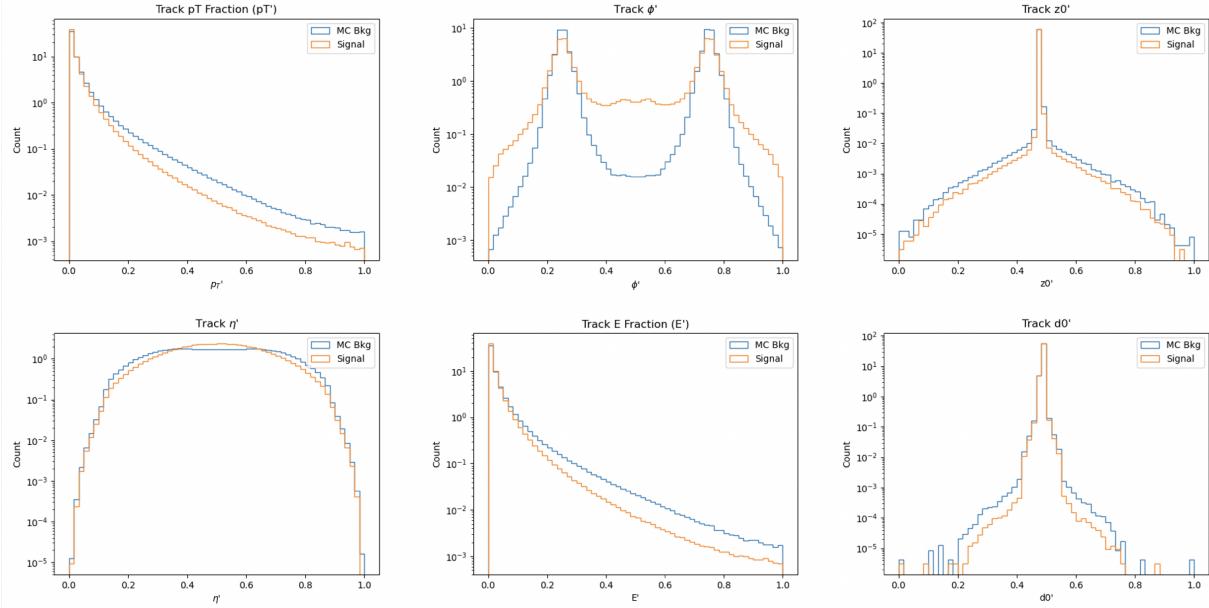


Figure 7.12: The 6 PFN track variables in background MC (blue) and signal MC (orange) after scaling and rotation. The  $\phi$  distribution is modified by the rotation procedure, as explained in the text.

### 7.2.3 Training

As seen in Figure 7.6, two networks are defined and combined for the PFN architecture. In our implementation the input layer has a dimension of 6, accounting for the 6 track variables described in the previous section. The first network, termed the  $\Phi$  network, creates the per-particle set representation as illustrated in Figure 7.5. The  $\Phi$  network has 2 hidden layers each of dimension 75, and an output later of dimension 64. These dimensions were chosen via an optimization procedure which balanced network complexity (achieved with more dimensions) against training time (achieved with fewer dimensions). The two hidden layers and  $\Phi$  output layer all use a RELU activation function [80], following the work of Ref. [79].

The input layer of the classifier  $F$  network is required to have the same dimension as the output layer of the  $\Phi$  network, and therefore takes dimension 64. This network contains 3 hidden layers with 75 nodes each, and again uses RELU activation [80]. The final layer is the binary classifier result with dimension 2, which uses a SOFTMAX activation [80] that is well suited for classification. The loss function for the complete PFN network is CATEGORICALCROSSENTROPY [80], which is a standard loss function for DNN classifiers. The standard Adam optimizer [81] [80] is used with a learning rate of 0.001. The learning rate was reduced from the nominal learning rate of 0.01 presented in Ref. [79] to prevent overtraining.

The PFN is a supervised algorithm, and is therefore trained on a labeled mixture of signal and background events. The signal input is an even mixture of all signal points considered in this analysis. Although the full simulated background for this analysis is composed of several SM processes as discussed in Section 6.2.1, QCD is the dominant background. Training with a QCD-only background sample is determined to produce better results than training using the full background mixture. Including MC backgrounds that are enriched in  $E_T^{\text{miss}}$  (recall Figure 6.3) reduces the ability of the PFN to classify SVJ signals. This is illustrated in the comparison of output classifier distributions in Figure 7.13. The signals used for training are the same in both cases. When training with a QCD-only background, high  $E_T^{\text{miss}}$  data and MC is more likely to be

classified as signal like; however the increased signal performance means that overall *sensitivity*<sup>3</sup> is higher with a QCD-only training. Additional studies on the optimal PFN training event mixture are available in Appendix B.2.

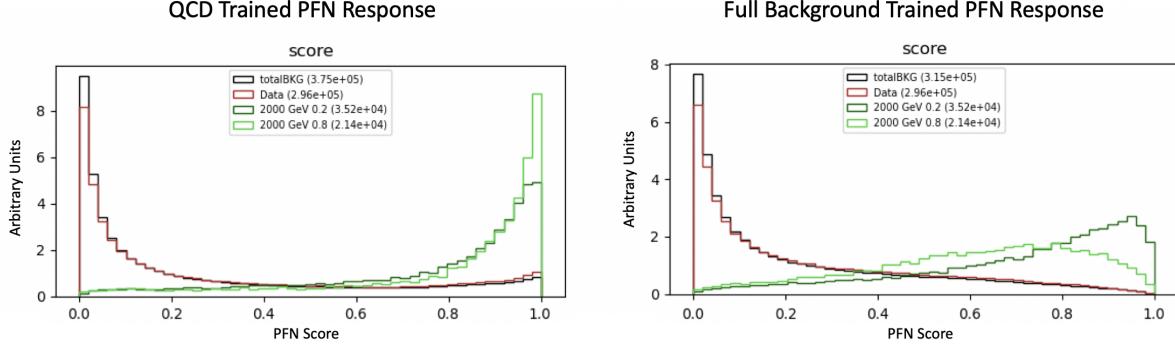


Figure 7.13: PFN score for full-background MC (black), data (red), and 2 representative signal points (green). The left plot is from a QCD-only training, while the right plot is from a full-background training. The histograms have been normalized to visualize the shapes better - the actual number of plotted events is shown in the legend. In the left plot we observe that both signal points are strongly classified as signal-like. In the right plot we observe less background contamination in the high score region, but worse signal classification. Both PFN trainings were tested for their effect on the analysis sensitivity and the QCD-only training was found to be favorable.

500k QCD MC background events and 500k SVJ signal events are used to train the network. The network is trained for 100 epochs. 20% of the training events are used for training validation. Figure 7.14 shows the loss during training, which is stable and shows no indication of overtraining, and the final score that provides signal-background discrimination.

Optimization studies were performed on the PFN, varying the number of training epochs, number of training events, learning rate, number of nodes, and dimension of the  $\Phi$  basis. A summary of these studies is presented in Appendix B.2. The model presented here represents an optimal choice across these parameters.

#### 7.2.4 Performance

The performance of the PFN is assessed via the AUC for each SVJ signal point. Although the PFN is trained against QCD MC only, the performance is evaluated using data as the background

---

<sup>3</sup>Sensitivity is a measure of the ability of an analysis to detect the signal, discussed further in Section 8.2

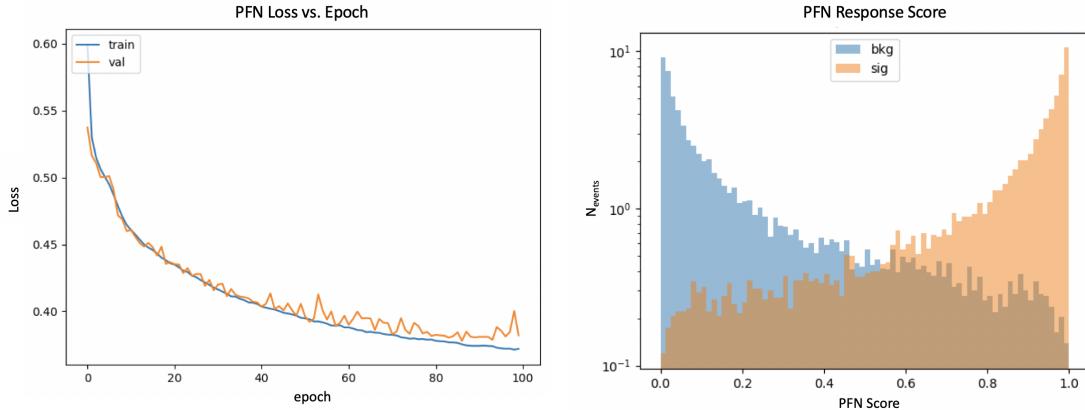


Figure 7.14: PFN architecture loss during training as a function of epoch (left) and the evaluated score for signal and background training samples (right). The loss vs. epoch plot shows that the network is not overtrained. The score plot shows a good separation between signal and background.

sample, since the ultimate task of the PFN is to separate SVJ signals from data, which is dominated by SM processes.

Figure 7.15 shows the ROC curve of one such signal point, illustrating a smooth response. Figure 7.16 shows the AUC of the PFN across the SVJ signal grid, demonstrating that  $AUC > 0.5$  is satisfied for all SVJ signals.

Figure 7.17 shows the output score distribution for data and four signals, illustrating the range of scores received by data events in comparison to signal events. As expected, most data events receive a background-like score (close to 0.0), indicating that the data is dominated by SM processes consistent with the background. Most signal events receive a signal-like score (close to 1.0). An optimization procedure determined that a selection of **PFN score > 0.6** can improve signal sensitivity across the grid. The optimization procedure considered the cut that would maximize sensitivity as measured by  $s/\sqrt{b}$ , where  $s$  the number of signal events accepted and  $b$  is the number of background events selected. This score selection is incorporated into the analysis design described in Chapter 8.

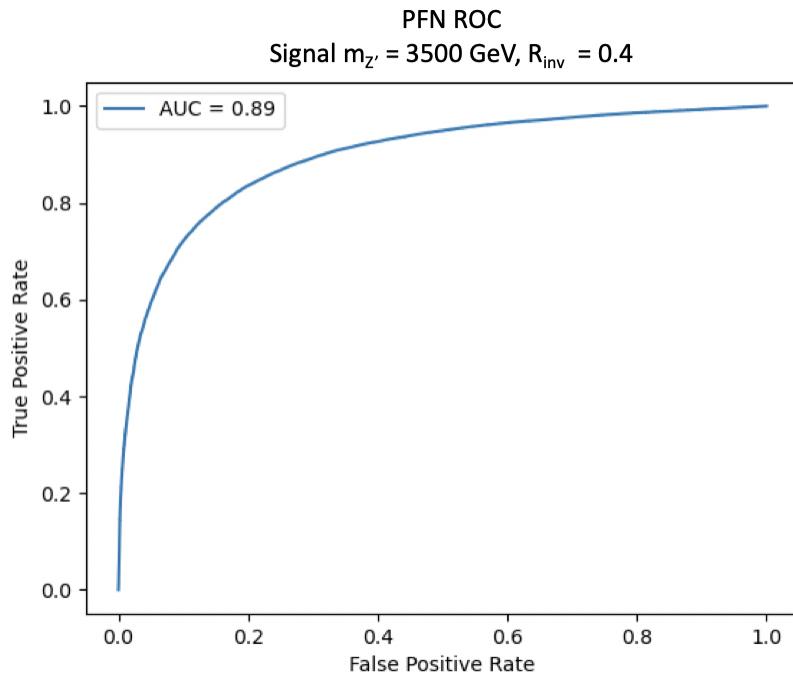


Figure 7.15: ROC for the PFN, using SVJ signal events (true positive) and data (false positive).

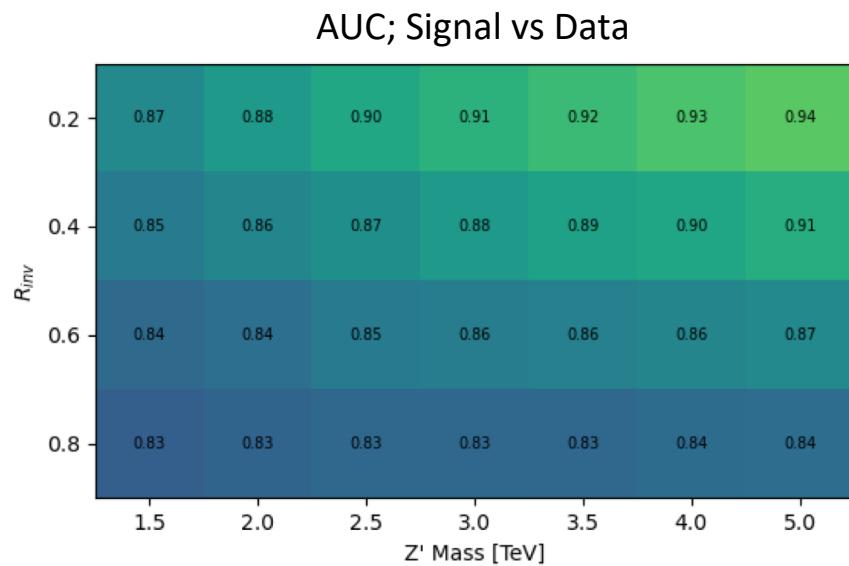


Figure 7.16: AUC for the PFN, shown for each signal in the SVJ grid.

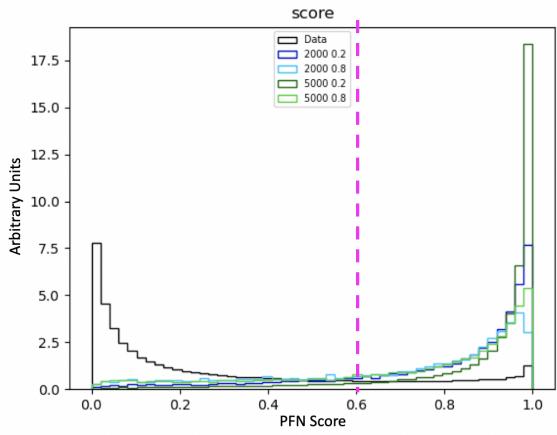


Figure 7.17: Illustration of the PFN score selection, showing the separation between data (black) and 4 signal points (blue and green). The legend information takes the form “ $m_{Z'} R_{inv}$ ” for the signal. The PFN score selection value is shown by the pink line. Only events with a score  $> 0.6$  will be accepted for use in the analysis. We see that most background (data) is rejected, while most signal is accepted.

## 7.3 ANTELOPE (Semi-supervised)

The semi-supervised machine learning approach broadens the discovery potential of the search through the use of data-driven training, where no signal model is provided. While broad sensitivity is a general goal of LHC searches, it is particularly motivated in the case of dark QCD models, which can lead to widely varying topologies depending on the values of model parameters.

### 7.3.1 Architecture Fundamentals

The model-independent search region of this analysis is implemented with a novel ML approach that builds on the PFN architecture to construct a tool that is capable of performing low-level anomaly detection with permutation-invariant inputs. This tool, referred to as **ANomaly deTEction on particLe flOw latent sPacE (ANTELOPE)**, is a custom solution designed for this analysis.

Figure 7.18 provides a diagram of the ANTELOPE architecture. ANTELOPE uses the trained PFN network described in the previous section to generate a permutation invariant event representation  $O$  from track level inputs. The  $O$  basis is used as the input for a *variational autoencoder* (VAE). A VAE is a common variation of a standard AE; the AE becomes *variational* if the latent space is constructed through Gaussian sampling rather than a vector of weights, as described further in Ref. [82]. VAEs have been used in previous ATLAS searches to model low-level particle information, such as the search presented in Ref. [83] which used the recurrent architecture described in Ref. [78]. One of the limitations of a recurrent architecture is the need to order the low-level inputs, which affects the performance of the tool. Jet track information is intrinsically unordered, and therefore a permutation invariant approach removes this element of arbitrary decision making from the modeling process.

The input to ANTELOPE architecture is the same 6 track variables for the leading two jets, as presented in Section 7.2.2. The track information is encoded to the PFN  $O$  event representation using the pre-trained  $\Phi$  neural network (trained according to the steps outline in Section 7.2.3). The

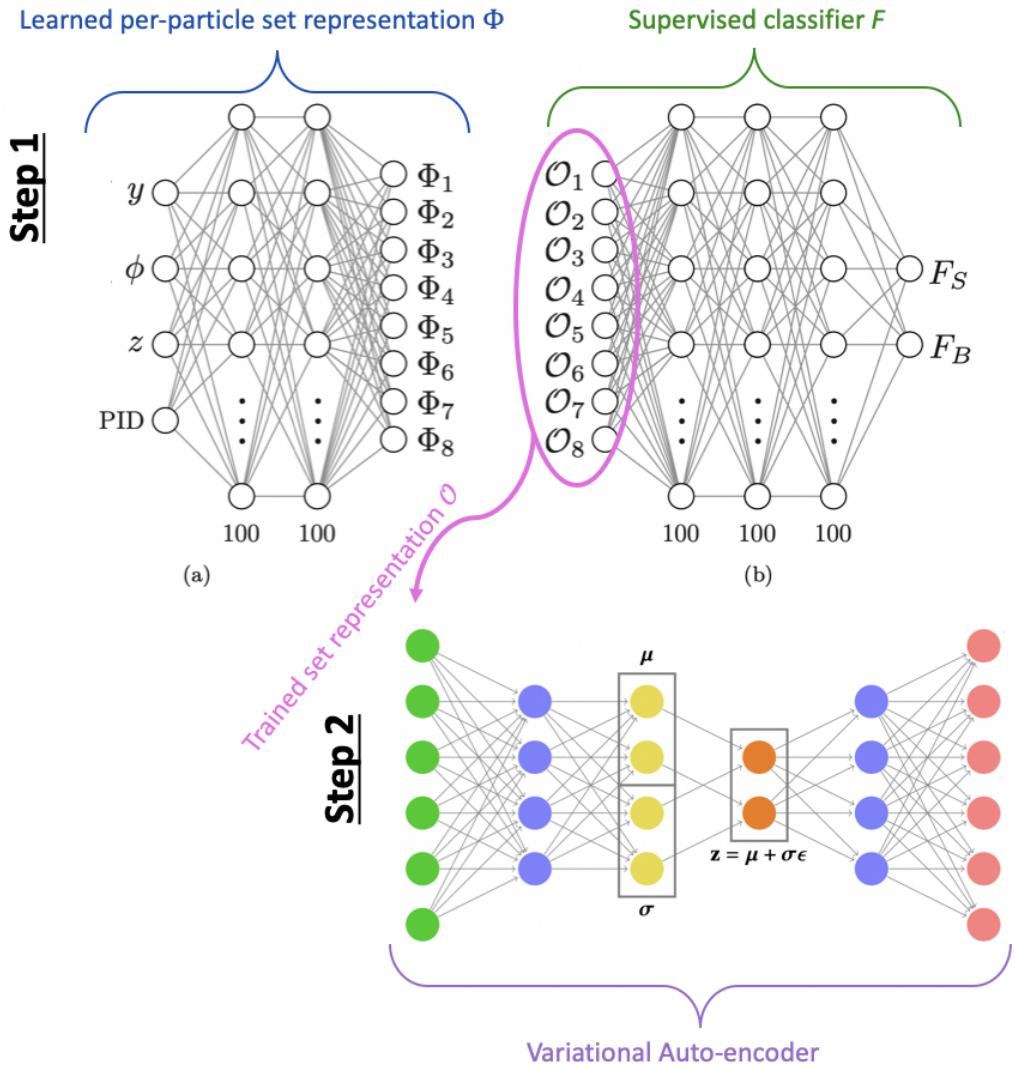


Figure 7.18: An annotated diagram of the ANTELOPE architecture. Step 1 illustrates the PFN which is fully trained before its use in the ANTELOPE network. Step 2 illustrates the variational auto-encoder. The Gaussian sampling of the latent space is shown, illustrating how the VAE differs from the AE shown in Figure 7.4.

VAE is then trained in an *unsupervised* way using inputs encoded to  $\mathcal{O}$  from data events only. Here *unsupervised* means that the VAE is given no knowledge of the signal model during training. There is implicit knowledge of the signal model in the  $\mathcal{O}$  encoding, so the full ANTELOPE network is considered semi-supervised, while the VAE component is unsupervised. A visual example of the  $\mathcal{O}$  input to the VAE portion of the ANTELOPE is given in Figure 7.19.

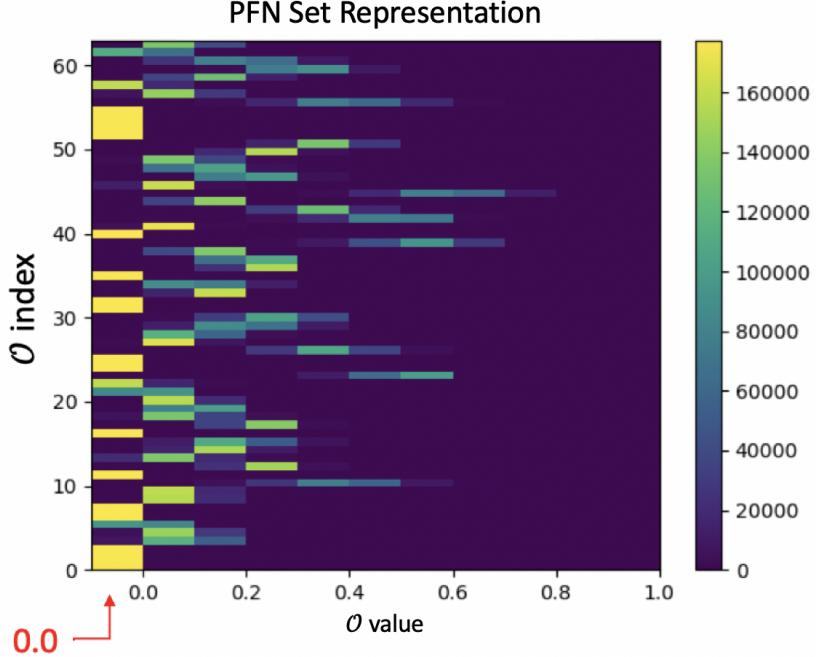


Figure 7.19: A visual representation of the 64 PFN  $\mathcal{O}$  which create the input for the VAE component of ANTELOPE. The plot is 2D histogram of the PFN  $\mathcal{O}$  index (0-63) versus the value assumed by that index. Many entries have a  $\mathcal{O}$  value of exactly 0.0. To visually separate these from entries with a small but non-zero  $\mathcal{O}$  value, any entries with value = 0.0 are moved to value = -0.01 (leftmost column) for the purpose of the plot only.

The VAE is trained to minimize the reconstruction error, or the difference between its input and output layer. This pushes it to uncover patterns in the data, which is predominantly composed of SM processes. Any rare events in the data which present patterns inconsistent with the majority of the data will receive a higher reconstruction error. This error is used to create the anomaly score.

### 7.3.2 Training

The VAE stage of the ANTELOPE network is trained over 500k data events. The input dimensionality of the VAE has to match the encoded  $\Phi$  dimension of the PFN, in this case 64. The encoding portion of the VAE has a hidden layer with 32 nodes, and a latent space dimension of 12. The decoding portion has another hidden layer of 32 nodes, and the output layer has a dimension of 64 to match the input layer. All layers use a RELU activation [80] except for the output layer which uses a SIGMOID activation [80], to restrict the output between 0 and 1. As in the PFN, the Adam optimizer [81] [80] is used.

The network is trained for 50 epochs, with a learning rate of 0.00001. The VAE was observed to need a very small learning rate to effectively minimize the loss. The loss  $\mathcal{L}$  is the sum of two terms, the mean-squared error (MSE) of input-output reconstruction, and the Kullback-Leibler divergence (KLD).

$$\mathcal{L} = \sum_i L_i = \sum_i |\Phi_i^2 - \Phi'_{i,i}|^2 + \lambda D_{\text{KL}} \quad (7.3)$$

Figure 7.20 shows the loss during training. The validation events are seen to have a lower loss than the training events, indicating there is no issue with overtraining.

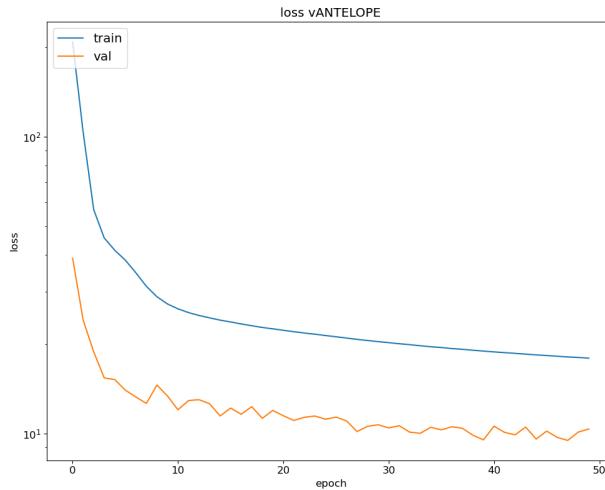


Figure 7.20: ANTELOPE architecture loss during training as a function of epoch.

### 7.3.3 Performance

As with the PFN, the ANTELOPE performance is assessed via the ROC and AUC. Figure 7.21 shows the anomaly score and an example ROC curve. The anomaly score is calculated from the loss as defined in 7.3. The score is produced by applying a sigmoid function to the loss to restrict its output between 0.0 and 1.0:

$$a = \frac{1}{1 + e^{-\mathcal{L}}} \quad (7.4)$$

where  $a$  is the anomaly score and  $\mathcal{L}$  is the VAE loss. Because the loss is always positive, the sigmoid transformation effectively restricts the anomaly score range between 0.5 and 1.0. The anomaly score is observed to range between 0.6 and 1.0, as the reconstruction loss is always non-zero. Following a similar sensitivity optimization as presented for the PFN score selection in Section 7.2.4, a selection of **anomaly score > 0.7** is chosen for use in the analysis.

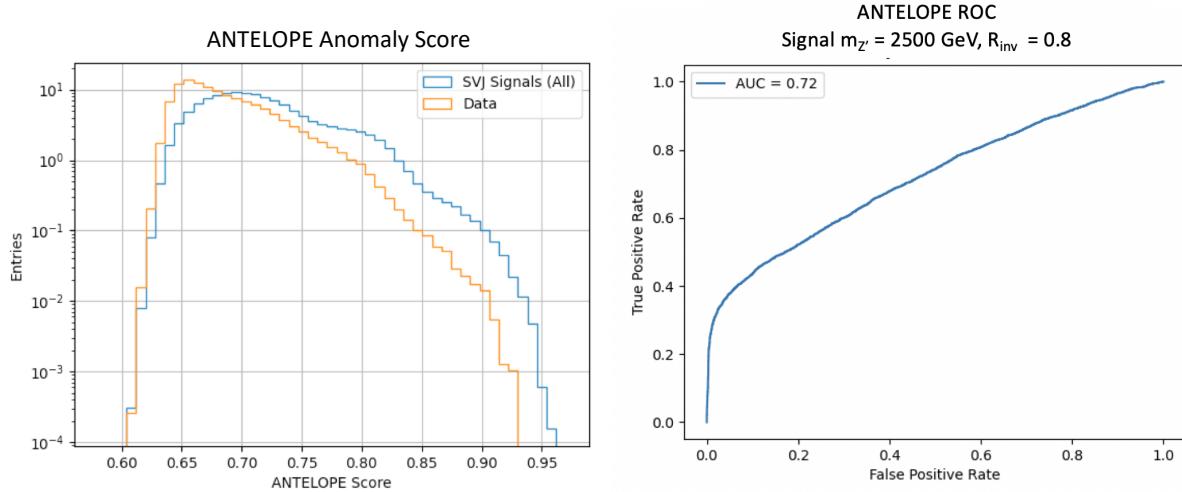


Figure 7.21: Anomaly score distribution (left), comparing all data (orange) and all SVJ signals (blue). The signals have a small but consistently higher score than the data, indicating that they are tagged as more anomalous by ANTELOPE. A ROC curve for an example signal point is also shown (right).

Figure 7.22 shows the AUC of the ANTELOPE across the SVJ signal grid, demonstrating discrimination capability across varying SVJ signal models. Compared to the supervised PFN

method, the ANTELOPE is not as performant (as expected due to the absence of signal model in training). However, the network is seen provide separation between signal and background for all signal points, as evidenced by  $AUC > 0.5$  across the signal grid.

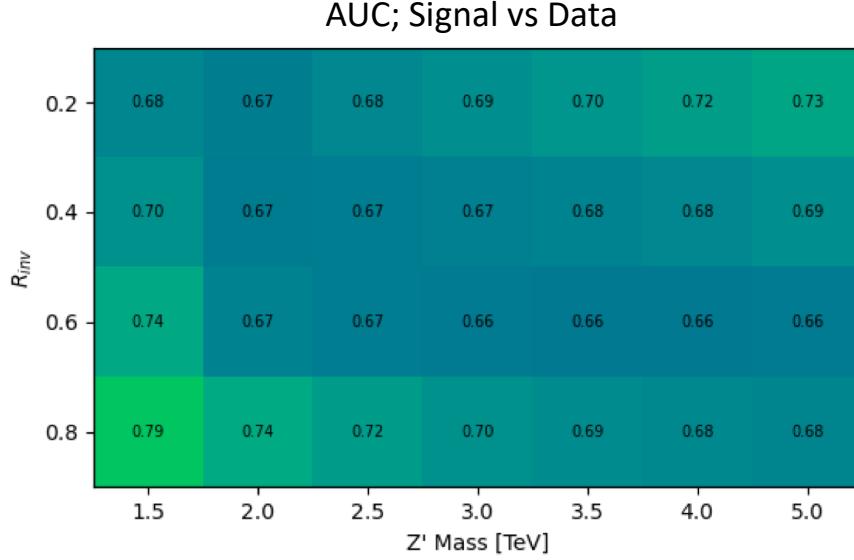


Figure 7.22: AUC from the ANTELOPE score for each signal in the SVJ grid.

**Model Independence** The unsupervised component of training the ANTELOPE network is expected to give it a more generalized sensitivity to new physics with  $E_T^{\text{miss}}$  and jet activity, beyond the scope of the SVJ grid. To assess this, alternative signal models are evaluated with the trained ANTELOPE network.

The following alternate signal models were considered:

- $Z' \rightarrow t\bar{t}$
- $W' \rightarrow WZ$
- Gluino pair production  $\rightarrow$  R-hadron + LSP ( $E_T^{\text{miss}}$ ) with gluino masses 2000/3000 GeV, LSP mass 100 GeV, and lifetime 0.03 ns (LSP = *lightest supersymmetric particle*)
- Emerging jets s-channel with mass 1000 GeV and lifetime 1ns

Figure 7.23 shows the distribution of these signals in the PFN score and the ANTELOPE anomaly score. The benefit of the ANTELOPE in enhancing model independence is clearly seen through the boost in performance for certain non-SVJ signal models. The gluino and emerging jet signals in particular are marked as highly anomalous by the ANTELOPE, but are marked as evenly background-like and signal-like by the PFN. This observation demonstrates that the use of the ANTELOPE network in this analysis has the potential to expand our sensitivity to include alternate signal models that could be marked as highly anomalous with the anomaly score.

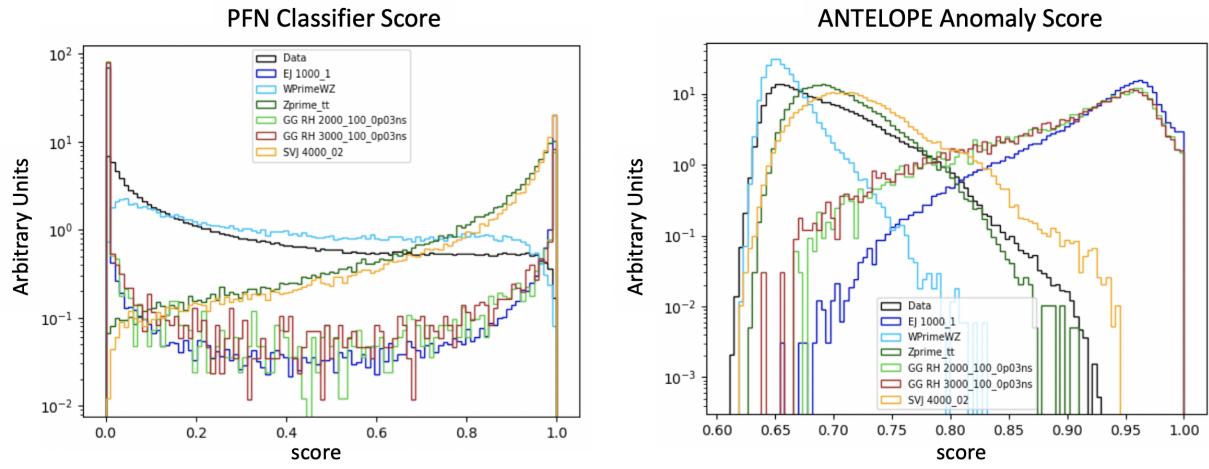


Figure 7.23: Comparing data and the alternate signal models for the PFN score (left) and ANTELOPE score (right). The emerging jet signal (dark blue) and gluino R-hadron signals (red, light green) are an example of the advantage of the model-independent ANTELOPE approach. These signals have a bimodal shape in PFN score but are clearly tagged with a high anomaly score by the ANTELOPE.

## Chapter 8: Analysis Strategy

This chapter will present the strategies used to isolate ATLAS data events most consistent with the SVJ model and to estimate the relevant background. The data and MC samples discussed in Chapter 6 are studied to create the analysis strategy, and the ML scores discussed in Chapter 7 are used to isolate the most signal like events. The background is estimated from the *transverse mass* ( $m_T$ ) spectrum, which captures the main components of the  $Z'$  decay. A *preselection* selects events consistent with the SVJ topology based on basic features of the jets and  $E_T^{\text{miss}}$ . Preselected events are then split into a *control region* (CR), *validation region* (VR), and *signal region* (SR). The CR and VR are used to validate the background estimation procedure. The SR is blinded during the development of the analysis strategy, and only unblinded to make the final measurements presented in Chapter 9. The final result is a polynomial fit of the transverse mass spectrum in the SR. The preselection, region definitions, and polynomial fit will be discussed in detail in the following sections.

### 8.1 Transverse Mass

The transverse mass  $m_T$  is chosen as the search variable due to the potential for the SVJ signal to create a resonant shape around the mass of the  $Z'$ .  $m_T$  is the total transverse mass of the two leading<sup>1</sup> jets and the  $E_T^{\text{miss}}$ , expressed in Equation 8.1 as:

$$m_T^2 = [E_{T,jj} + E_T^{\text{miss}}]^2 - [\vec{p}_{T,jj} + \vec{p}_T^{\text{miss}}]^2 \quad (8.1)$$

where  $E_{T,jj}$  is the transverse energy of the dijet system created by the two leading jets. We take  $E_{T,jj} = m_{jj}^2 + |\vec{p}_{T,jj}|^2$ , where  $m_{jj}^2$  is the invariant mass of the two leading jets, and  $\vec{p}_{T,jj}$  is

---

<sup>1</sup>Recall the definition of the leading jets from Chapter 6, meaning the two jets with highest  $p_T$  in the event.

the vector sum of the  $p_T$  of the two leading jets.  $m_T$  is selected as the search variable in place of simpler invariant mass  $m_{jj}$  because substantial energy from the  $Z'$  decay is captured in the  $E_T^{\text{miss}}$ .

Figure 8.1 illustrates the resonance in  $m_T$  of the SVJ signals. The smoothly falling background is shown in comparison to the resonant shape of the signals, which form a peak just below the associated  $Z'$  mass. The lower  $R_{inv}$  signals form a more distinctive resonance in  $m_T$ , while the high  $R_{inv}$  signals have a much wider  $m_T$  shape.

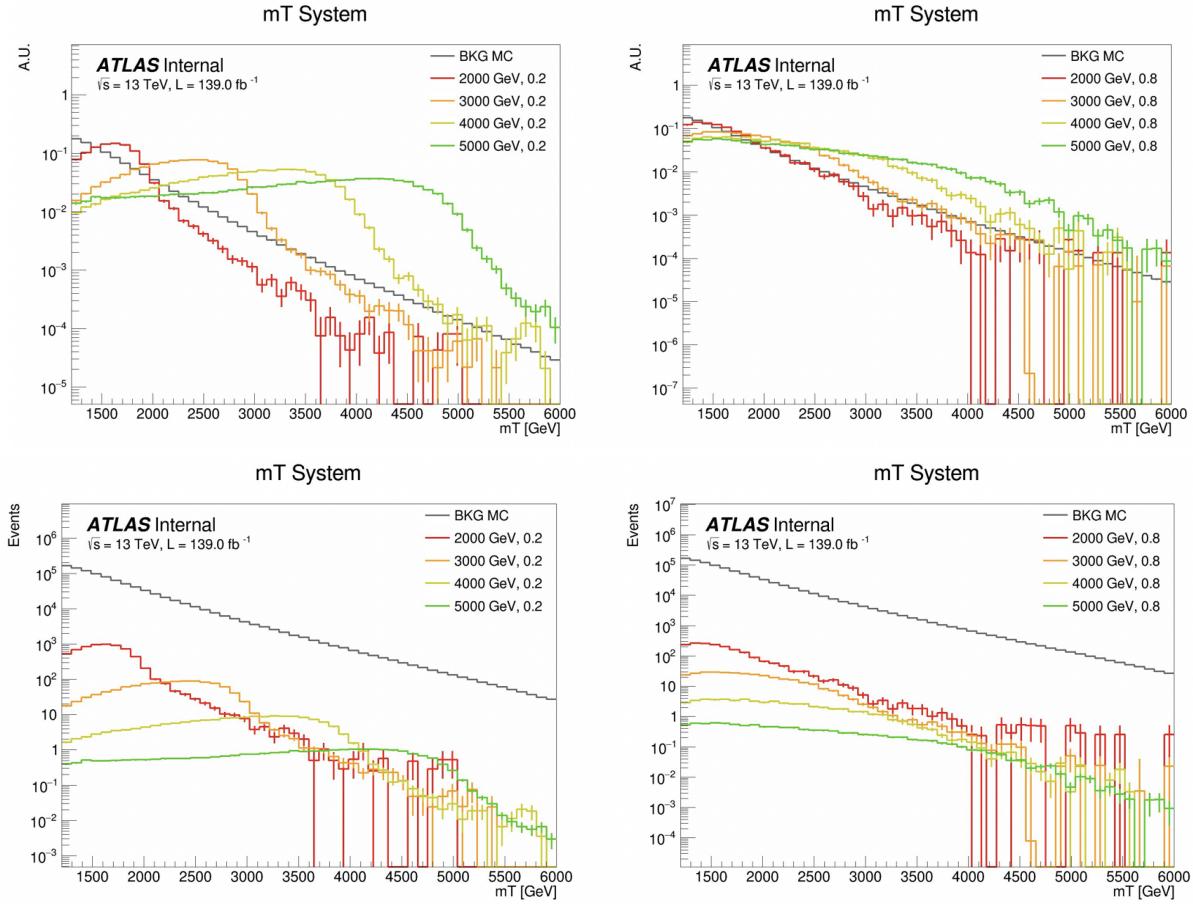


Figure 8.1: The resonant shape of the SVJ signals (color) in  $m_T$ , in contrast to the smoothly falling  $m_T$  background (grey). The top row illustrates unit normal shapes, so that the shape of the signals is more easily seen. The bottom row illustrates the signal and background scaled to their expected yield at preselection, illustrating the relative expected statistics. The  $R_{inv} = 0.8$  signals (right) boast a wider shape, making them more difficult to detect, while the  $R_{inv} = 0.2$  signals (left) produce a more narrow resonance in  $m_T$ . The signal models are identified in the legend as " $m_{Z'}$ ,  $R_{inv}$ ".

## 8.2 Preselection

The preselection isolates the phase space of events that most closely match the SVJ signal topology. Each cut was determined to reduce the background and enhance signal sensitivity. The list of preselection cuts and the motivation behind each cut are as follows. Here “jets” refer to anti- $k_t$  R=0.4 jets, as discussed in Chapter 5. Table 8.1 shows the impact of these cuts in sequence for data and signal.

- At least 2 jets; in order to reconstruct the resonance mass
- Leading jet ( $j_1$ )  $p_T > 450$  GeV; to ensure the trigger is fully efficient
- Subleading jet ( $j_2$ )  $p_T > 150$  GeV; to mitigate the presence of non-collision background (Appendix C.2)
- $|\eta_{j1,j2}| < 2.1$ ; to ensure jets are fully within the tracker
- $\Delta Y < 2.8$  (difference in rapidity between  $j_1$  and  $j_2$ ); to ensure central production associated with the hard scatter
- $E_T^{\text{miss}} > 200$  GeV; to restrict the phase space to events with possible dark particles
- $m_T > 1.2$  TeV, to ensure a smoothly falling  $m_T$  distribution for fitting (Section 8.5)
- At least 3 tracks for each of the two leading jets  $j_1$  and  $j_2$ ; to have adequate tracking information for the ML tools
- $\Delta\Phi(j_1, j_2) > 0.8$ ; to mitigate the presence of non-collision background (Appendix C.2).

With the exception of the cuts necessary to reduce the non-collision background, all cuts were verified to enhance signal sensitivity by improving  $s/\sqrt{b}$ , a standard estimate of discovery sensitivity, where  $s$  is the number of signal events and  $b$  is the number of background events. The cuts on  $\Delta Y$  and  $E_T^{\text{miss}}$  were optimized to enhance  $s/\sqrt{b}$ , and the other cuts were informed by the physics motivations provided above.

Data			Signals - All		
Cut	Statistics	Rel. Efficiency	Cut	Statistics	Rel. Efficiency
Initial	1.71E+10		Init	6.66E+05	
Trigger	3.45E+08	0.02	Trigger	2.83E+05	0.42
N. jets >= 2	2.84E+08	0.82	N. jets >= 2	2.80E+05	0.99
Leading Jet Pt > 450 GeV	1.49E+08	0.52	Leading Jet Pt > 450 GeV	2.21E+05	0.79
Lead Jet Eta < 2.1	1.42E+08	0.95	Lead Jet Eta < 2.1	2.19E+05	0.99
Subleading Jet Eta < 2.1	1.33E+08	0.94	Subleading Jet Eta < 2.1	2.12E+05	0.97
dY < 2.8	1.29E+08	0.96	dY < 2.8	2.11E+05	0.99
MET > 200 GeV	7.07E+05	0.01	MET > 200 GeV	1.13E+05	0.54
mT > 1200 GeV	5.38E+05	0.76	mT > 1200 GeV	7.94E+04	0.70
Jet2 pT > 150 GeV	5.03E+05	0.93	Jet2 pT > 150 GeV	6.14E+04	0.77
dPhi(j1,j2) > 0.8	4.97E+05	0.99	dPhi(j1,j2) > 0.8	5.83E+04	0.95

Table 8.1: Preselection cuts for data (left) and signal (right).

Vetos are applied to reject any events where an error for a subdetector is flagged. To reject non-collision backgrounds (NCB), such as calorimeter noise, beam halo interactions, or cosmic rays, the standard ATLAS event cleaning procedure is applied. As this analysis is very dependent on  $E_T^{\text{miss}}$  associated to jets, the TIGHT [84] event cleaning working point is applied. Tight cleaning requires jets to pass a stricter set of quality requirements compared to the LOOSE [84] cleaning option. Due to the alignment between jets and  $E_T^{\text{miss}}$  for SVJ events, it was found that two additional cuts (indicated above) are needed to remove NCB. The process for selecting these cuts is presented in Appendix C.2.

The leading and subleading jets in each event are considered the dark quark candidates from the  $Z' \rightarrow q_D \bar{q}_D$  decay. They are therefore the two jets of greatest interest in the event, and used in the computation of key analysis variables. This choice was determined through studies of the dark quark trajectory in simulation which determined that the leading and subleading jets are most often aligned with the dark quarks, and therefore most likely to capture the dark quark hadronization. This study can be found in Appendix D.

Figure 8.2 and Figure 8.3 show the distribution of signals, data and background MC in several key analysis variables after preselection is applied. The variables illustrated are:

- Transverse mass  $m_T$ : key analysis variable which reconstructs the  $Z'$  mass, as discussed in Section 8.1.

- Leading jet  $p_T$ : the trigger variable, and a component of  $m_T$ .
- Subleading jet  $p_T$ : dark quark candidate and component of  $m_T$ .
- Missing transverse energy  $E_T^{\text{miss}}$  (or MET): component of  $m_T$ , and an indication of the presence of dark hadrons. All signals are observed to have more  $E_T^{\text{miss}}$  than the background.
- $\Delta\phi(j1, j2)$ : difference in trajectory of the two leading jets, measured in the  $\phi$  plane (recall the ATLAS detector geometry of Figure 4.2a). Orientation of the jets is of importance to the ML model as discussed in Section 7.2.2.
- $\Delta Y(j1, j2)$ : difference in trajectory of the two leading jets, measured in the  $Y$  plane (recall Figure 4.2a and the definition of rapidity Equation 4.1). The signals are seen to have lower  $\Delta Y(j1, j2)$  than the background.
- $\Delta\phi(j1, E_T^{\text{miss}})$ : the angular separation between the leading jet and the  $E_T^{\text{miss}}$ . The leading jet is predominantly back-to-back with the  $E_T^{\text{miss}}$ .
- $\Delta\phi(j2, E_T^{\text{miss}})$ : the angular separation between the subleading jet and the  $E_T^{\text{miss}}$ . The subleading jet is predominantly aligned with the  $E_T^{\text{miss}}$ , which is a unique feature of this analysis as jets that are closely aligned with  $E_T^{\text{miss}}$  are often removed from other ATLAS analyses.

The data and background MC are both illustrated in Figure 8.2 and Figure 8.3. The agreement between them is generally observed to be good, particularly in the key analysis variable  $m_T$ . The agreement is not required to be perfect as the background MC is not used for the background estimation. The primary motivation for studying the background MC is to uncover and remove issues unique to data such as the non-collision background.

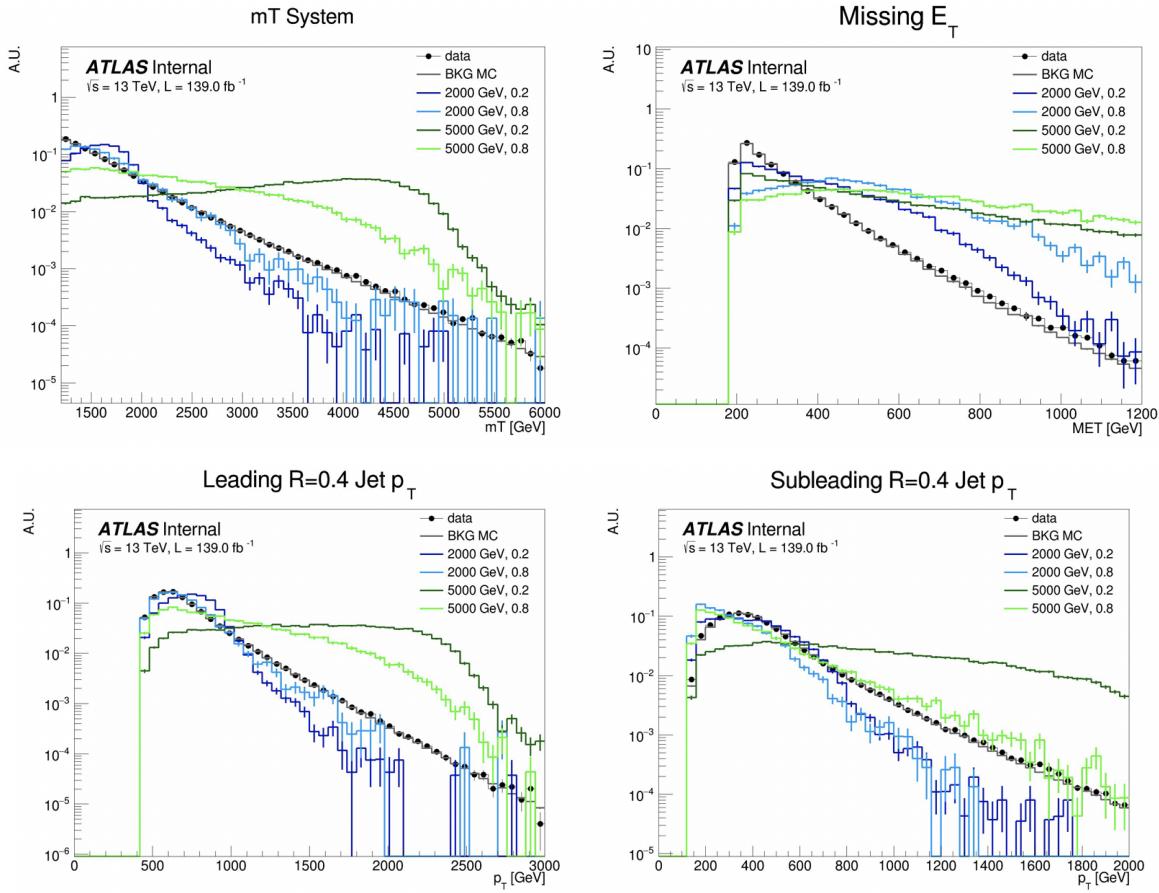


Figure 8.2: Energy and momentum analysis variables at preselection, for data (black), background MC (grey), and representative signal models (color). The signal models are identified in the legend as “ $m_{Z'}$ ,  $R_{inv}$ ”.

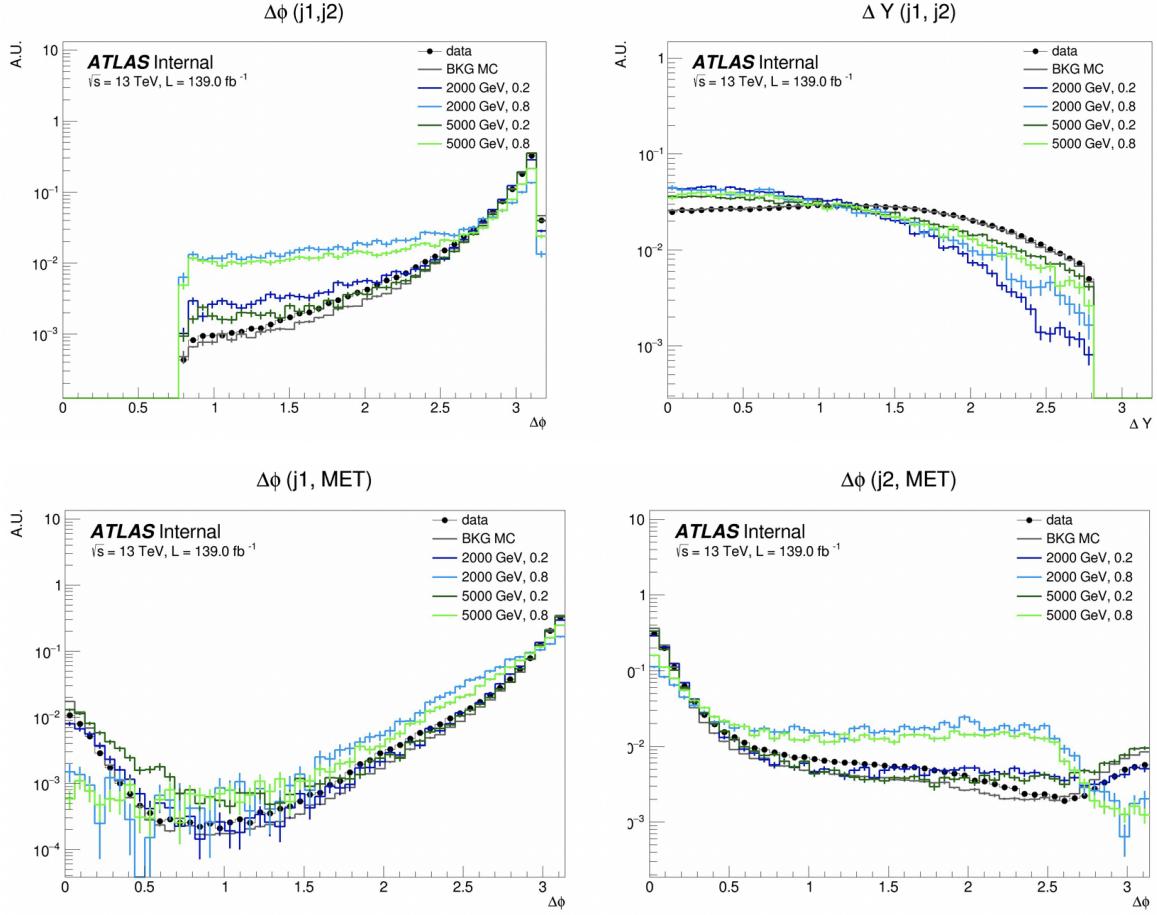


Figure 8.3: Orientation analysis variables at preselection, for data (black), background MC (grey), and representative signal models (color). The signal models are identified in the legend as “ $m_Z'$ ,  $R_{inv}$ ”.

### 8.3 SVJ Fit and Discovery Analysis Strategies

As was introduced in Chapter 7 this analysis is interested in achieving dual goals: to make the best possible measurement of the SVJ signal model generated for this analysis, and to broadly search for any signals consistent with dark QCD behavior and inconsistent with a Standard Model background hypothesis. To this end, two parallel analysis strategies are developed.

The SVJ Fit strategy uses the supervised PFN ML score in defining the signal region. Recall, the PFN is trained over simulated MC background and a combination of all MC SVJ signals. This gives this ML tool high sensitivity to the particular nuances of the SVJ shower predicted by the modeled theory. In addition to using the supervised ML tool, the SVJ Fit analysis strategy sets limits on the expected cross section of each signal point in the SVJ signal grid. To achieve this, the shape of the SVJ signals are considered in the final fit, as will be elaborated on Section 8.6.1. The combination of the supervised PFN ML score and the signal-shape sensitive fitting strategy allows for the greatest possible sensitivity to the modeled signal process, thus allowing the analysis the best chance at discovery of this model, or enabling the analysis to set the best possible limits on the observed cross section.

In contrast, the Discovery analysis strategy attempts to design a more general search, which could be sensitive to SVJs, but also to other possible hidden valley dark QCD models, such as fully dark jets or emerging jets [21]. The Discovery analysis strategy uses the semi-supervised ANTELOPE ML score in defining the signal region. Recall, the ANTELOPE is trained over ATLAS data only, with no explicit knowledge of the SVJ signal behavior. The Discovery fit strategy is also signal model agnostic, by employing a bump hunt [85] strategy, which searches a smoothly falling template for any bumps inconsistent with a background only hypothesis. Therefore any signal which could present a resonant signature in  $m_T$  could show up as an excess in this strategy.

The details of both strategies will be explored in the follow sections which detail the design of the signal regions and fit strategies. Figure 8.4 illustrates the difference in the fitting concept.

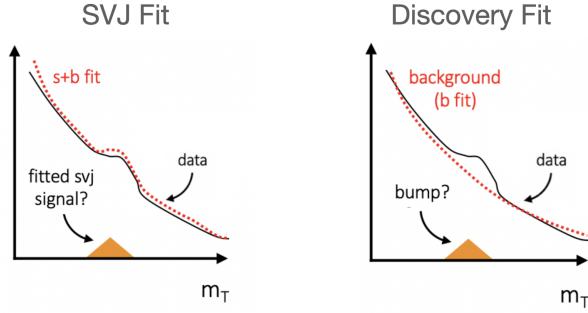


Figure 8.4: The two fitting strategies. The SVJ Fit (left) illustrates how SVJ signal shapes will be considered in the fit to search for SVJ specific signal shapes, where “s+b fit” indicates a fit that considers the shape of the signal. The Discovery Fit (right) illustrates how the data is compared to a background-only hypothesis to search for any kind of  $m_T$  bump, where “b fit” indicates a background-only fit with no signal hypothesis.

## 8.4 Analysis Regions

### 8.4.1 Control and Validation Regions

The final background estimation will come from a polynomial fit to the  $m_T$  distribution in the signal region. The control and validation regions are needed to develop and test this fit in data.

To define the CR selection, a variable is needed that isolates background from all signals across the  $(R_{inv}, m_Z)$  grid, which is challenging due to the varying nature of the signal models in quantities such as  $E_T^{\text{miss}}$  and  $p_T$ , as illustrated in Figure 8.2. The variable *jet width* is chosen, which is the calorimeter measurement of the spread of the clusters which are used to define the jet [86]. The concept is illustrated in Figure 8.5. Jets with only one very energetic cluster have a small width, while jets with many lower energy clusters have a large width.

Figure 8.6 shows jet width specifically for the subleading jet, in data, background MC and signal at preselection. The leading jet width, which was determined to be less useful for isolating signal from background, is also shown. The subleading jet is more likely to be aligned with  $E_T^{\text{miss}}$ , which is why the signal jet width is consistently wider in the subleading jet, but not the leading jet. A selection of  $\text{width}_{j2} < 0.05$  is chosen for the CR, with the VR and SR therefore having a selection of  $\text{width}_{j2} \geq 0.05$ .

While the CR was used to develop the polynomial strategy, and is the primary region used in

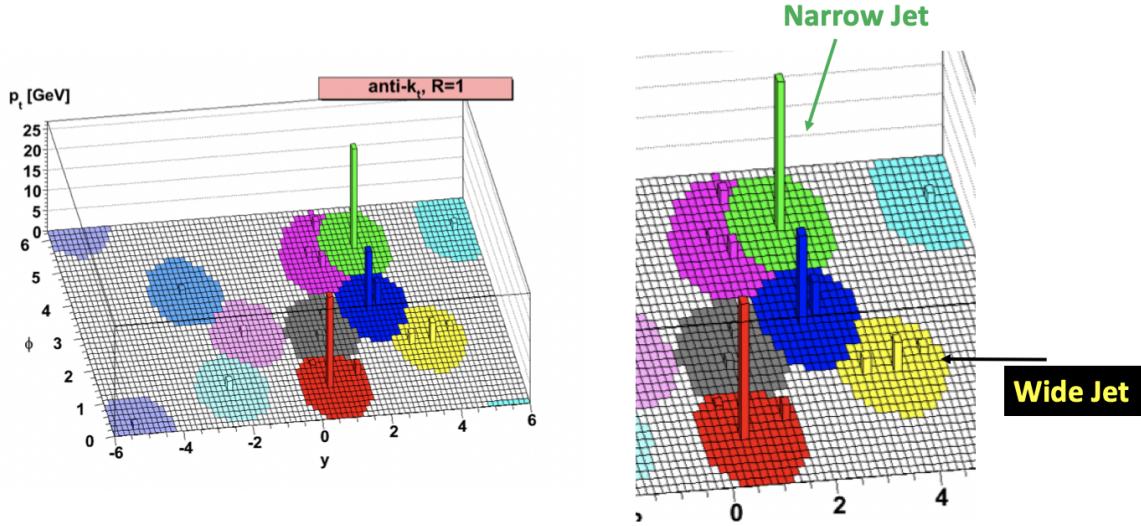


Figure 8.5: Recall the construction of anti- $k_t$  jets as described in Section 5.4.3 and illustrated in Figure 5.7. On the right, we zoom in on two jets, illustrating the narrow cluster pattern in the green jet, and the wide cluster pattern in the yellow jet.

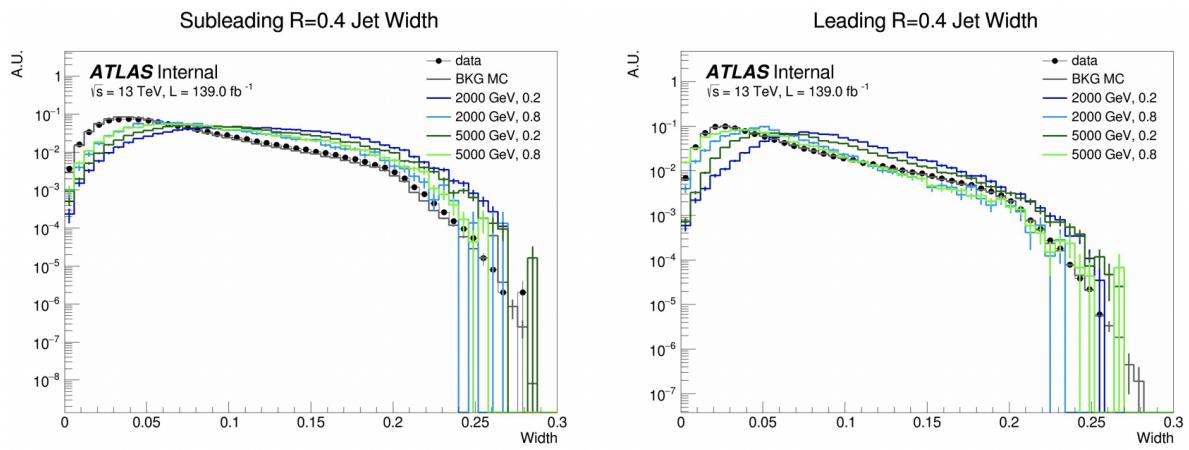


Figure 8.6: Distributions of the subleading jet width  $\text{width}_{j2}$  (left) and leading jet width  $\text{width}_{j1}$  (right) in data, background MC and signals at preselection. All SVJ signals are seen to be wider than the background in  $\text{width}_{j2}$ . The same is not true for  $\text{width}_{j1}$ , where some signals are observed to closely match the background.

many of the fit studies, a validation region is used as an additional check of the estimation strategy in data. The VR is defined using the region of events with low ML score by either the PFN or ANTELOPE networks. Here the analysis strategy splits into the two parallel strategies presented in Section 8.3: the SVJ fit strategy and the Discovery strategy. A selection of [PFN score  $\leq 0.6$  &  $\text{width}_{j2} \geq 0.05$ ] defines the SVJ Fit VR, while [ANTELOPE score  $\leq 0.7$  &  $\text{width}_{j2} \geq 0.05$ ] defines the discovery VR.

There are therefore three variables that are crucial to the analysis strategy:  $\text{width}_{j2}$ , ML score, and  $m_T$ . We check the expected shape of  $m_T$  across the CR, VR, and SR using background MC to ensure the shape is smoothly falling across all 3 regions. Figure 8.7 shows the distribution of  $m_T$  across the CR, VR, and SR, for both the PFN (supervised) and ANTELOPE (semi-supervised) strategies. No significant bumps or sculpting are observed. Some slope is observed in the ratio of the CR to the VR/SR shapes; however, the chosen background estimation strategy of polynomial fitting (to be discussed in Section 8.5) is expected to accommodate this slope. Further, testing the ability of the background polynomial to fit shapes with a variety of slopes increases our confidence in the ability to background polynomial to fit the blinded SR  $m_T$  distribution.

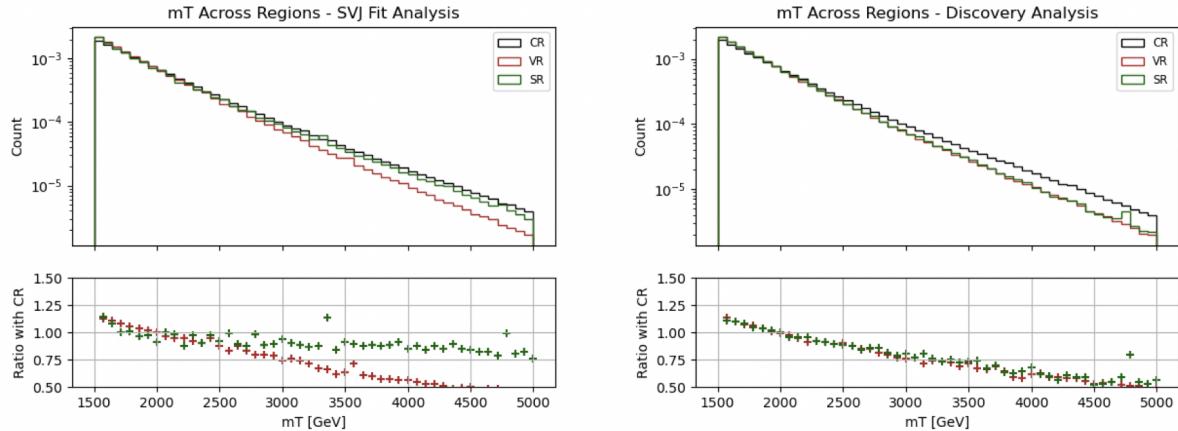


Figure 8.7:  $m_T$  in simulation across the CR, VR, and SR for both PFN (left) and ANTELOPE (right) selections. While there is variation in the slope of the distribution, no sculpting of bumps is observed.

### 8.4.2 Signal Region

A selection of PFN score  $> 0.6$  in the SVJ Fit region and ANTELOPE score  $> 0.7$  in the Discovery region is made to provide the primary signal-to-background enrichment, as motivated by Section 7.2. These values are determined to maximize  $s/\sqrt{b}$  in each region. The additional selection of  $\text{width}_{j2} \geq 0.05$  orthogonalizes the SR to the CR. Note that the PFN and ANTELOPE regions are not orthogonal; this is because the two analysis flows serve different purposes, their statistical treatments are different, and they will not be combined.

A summary of the SR, CR, and VR definitions can be seen in Figure 8.8, along with the relative data statistics in each region.

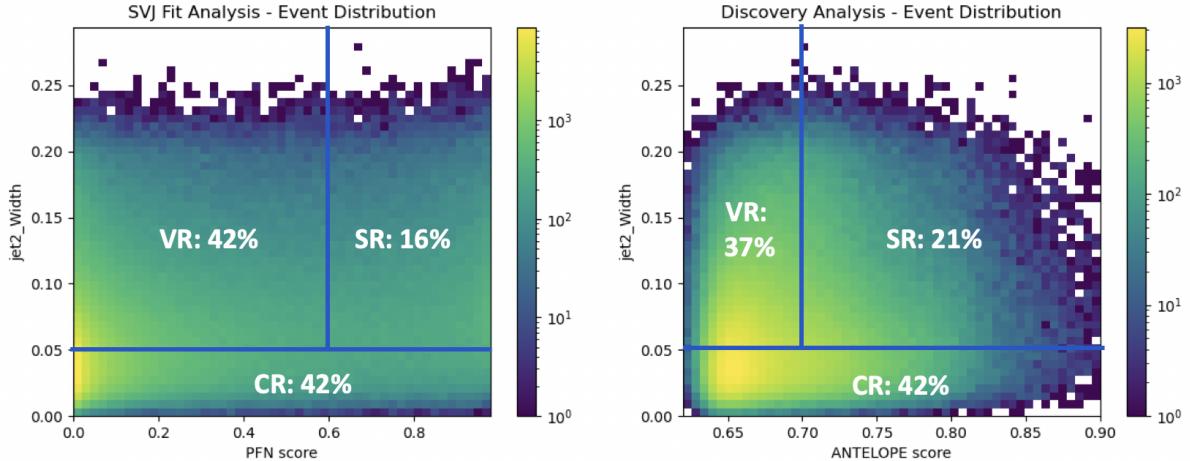


Figure 8.8: Distribution of data events amongst the CR, VR, and SR regions, along with the fractional population of each region. The SVJ Fit region is shown left with the PFN score on the x-axis, and Discovery region is shown right, with the ANTELOPE score on the x-axis.

A diagram demonstrating the complete analysis flow can be seen in Figure 8.9.

## 8.5 Background Estimation

The SM background in the SR is predominantly composed of QCD events, and due to the poor modeling of QCD at high energies by MC, it is estimated in a fully data-driven way. An empirical functional form is used for the background shape of  $m_T$ . The ability of this function to model

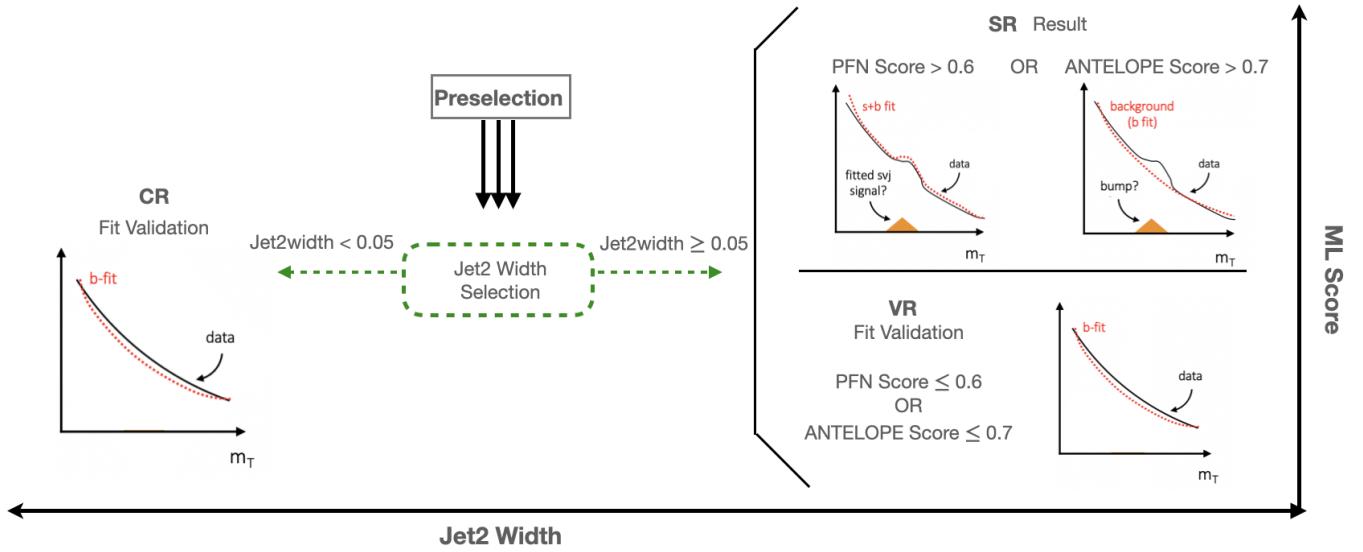


Figure 8.9: Flow of analysis selections and fitting strategy. From preselection, events with  $\text{Jet2Width} < 0.05$  are set aside for the CR. Events with  $\text{Jet2Width} \geq 0.05$  are split according the ML score. Events with low ML score create the VR, while events with high ML score create the SR. Events with high PFN score are fitted to determine if they are compatible with the SVJ signal shape. Events with high ANTELOPE score are fitted for a background estimation, and a search for any general data bump is performed.

the background behavior is tested both the CR and the VR for each analysis strategy. The shape parameters are left free in all the fits.

The fits are performed for  $1500 \text{ GeV} < m_T < 6000 \text{ GeV}$ . The polynomial chosen is a standard 5-parameter function used in several similar dijet search analyses such as [87] [88] [70] and shown in Equation 8.2:

$$f(x) = p_1(1 - x)^{p_2}x^{p_3+p_4\ln x+p_5\ln^2 x} \quad (8.2)$$

Here  $x = m_T/\sqrt{s}$  (transverse mass scaled to the  $pp$  collision center of mass energy) and  $p_i$  are free parameters. The fit function is required to be fully positive, and the  $m_T$  distribution is fit to 90 even-width bins. The resulting fit shape is used as the background estimation for both the SVJ Fit strategy and the Discovery strategy. Validation of the fit and its ability to both model the background and detect signal are shown in Section 8.6.

Higher order polynomials were also considered, but an F-test was performed according to the recommendations in Ref. [88], and the five parameter function was determined to be adequate and

optimal for capturing the shape of the background. The F-test uses the test statistic

$$F = \frac{\frac{\chi_{\text{nom}}^2 - \chi_{\text{alt}}^2}{n_{\text{alt}} - n_{\text{nom}}}}{\frac{\chi_{\text{alt}}^2}{n - n_{\text{alt}}}} \quad (8.3)$$

where  $\chi_{\text{nom}}^2$  and  $\chi_{\text{alt}}^2$  are the  $\chi^2$  values for the polynomial data fits using the nominal (in this case, 5 parameter) model and the alternate model (in this case, 6 or 7 parameter model).  $n_{\text{nom}}$  and  $n_{\text{alt}}$  are the number of free parameters in the nominal or alternate model, and  $n$  is the number of bins used for the fit. In the asymptotic limit  $F$  follows the Fisher-Snedecor distribution, so that  $F$  can be converted to a  $p$ -value  $p(F)$  to determine if the alternate model provides a better fit to the data, indicated by  $p(F) < 0.05$ . We find  $p(F) >> 0.05$  for a 6 and or 7 parameter polynomial function, indicating there is no improvement from increasing to 6 or 7 parameters. Comparing a nominal model with 4 parameters to the 5 parameter model, we find  $p(F) = 0.013$ , indicating that the 5 parameter model is preferable to the 4 parameter model.

## 8.6 Fit Strategy and Validation

The steps taken to validate the fitting approach for both the SVJ Fit strategy and the Discovery strategy will be outlined in the following sections. The signal region fits which comprise the final result will be presented in Chapter 9.

### 8.6.1 SVJ Fit Strategy

The ability of the five parameter fit function to capture the shape of the background is studied extensively, using data from the CR and VR. Signal injection tests are performed to determine the ability of the fit to recover and quantify any SVJ signal excess. Estimates of the expected sensitivity and the ability to set upper limits on the cross section of the signal process are also verified.

The fit results are primarily evaluated by their  $p$ -value, which dictates the probability of observing the given data spectrum given the fit hypothesis. A higher  $p$ -value is an indication of better agreement between the data and the fitted shape. The  $\chi^2/\text{d.o.f.}$  (or chi-square per degrees

of freedom, shown as just “ $\chi^2$ ” or “x”) is also presented. The  $\chi^2$  is checked to make sure it is not substantially larger or smaller than 1.0.  $\chi^2$  values close to 1.0 indicate that the fit is able to capture the data without overfitting. For each fit, the pattern of the residuals is also shown. The residual is calculated as the difference between the observed data in a bin and the fit estimation for the bin, divided by the statistical uncertainty, indicating the significance of the deviation from the fit estimation.

## Background Only Fits

The background fit polynomial is validated using the original data from the CR and VR, and pseudo-data generated from the CR.

The nature of the functional fitting method allows the fit to easily adapt to changes in slope of a smoothly falling distribution. Thus validation of the fit can be performed in data using the CR and the VR distributions to model the expected behavior in the SR, despite the expected differences in slope of the  $m_T$  distribution illustrated in Figure 8.7. Figure 8.10 shows the a successful fit performed on the full statistics CR and VR regions. The  $m_T$  spectrum is fit in 90 bins of width 50 GeV. For the purposes of this analysis, any converged fit with a  $p$ -value  $> 0.05$  is considered successful.

Table 8.2 shows the post-fit values of the fit parameters and their uncertainties for each fit. Fits of the MC background in the CR, VR, and SR were also performed and observed to be successful. These fit are available in Appendix E.

To further validate the stability of the fit against potential statistical fluctuations, *pseudo-data* (also known as *toy data*) are created from the CR data distribution. The pseudo-data is created following an *Asimov* prescription [89], using a template to generate a set of toys representing different possible statistical fluctuations. When studied as a group, the performance of the pseudo-data collection represents the range of possible behavior for an unknown distribution, such the the SR data, given its statistical uncertainties.

The template used to generate the pseudo-data is a *smoothed* and *scaled* version of the CR.

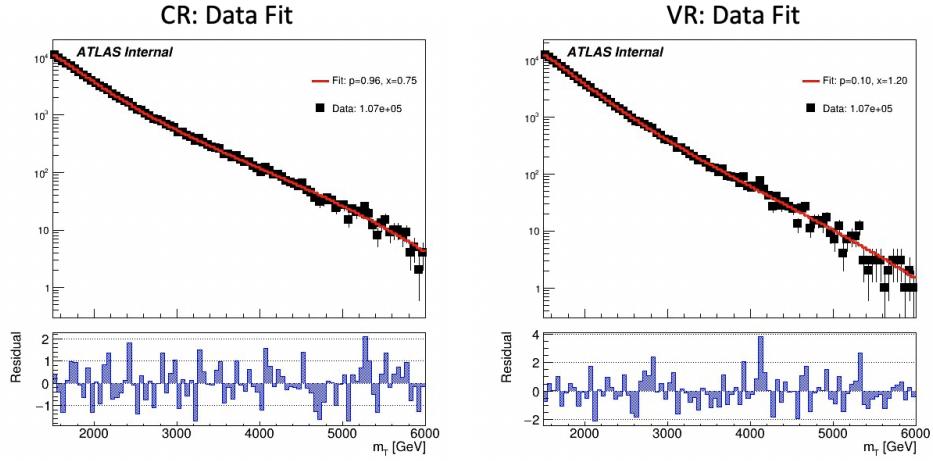


Figure 8.10: Background-only  $m_T$  fits using data in the full statistics CR and VR regions. The fit is observed to converge with  $p$ -value  $> 0.05$ . The distribution of residuals is reasonably flat. The number of events in the data histogram,  $p$ -value and  $\chi^2$  value (x) are reported in the legend.

Parameter	CR		VR	
	Value	Error	Value	Error
p1	1.0716e+05	4.64e+02	1.0686e+05	4.63e+02
p2	4.2479e+01	2.97e+00	3.4279e+01	3.71e+00
p3	5.3888e+01	3.81e+00	4.0798e+01	4.50e+00
p4	2.3804e+01	1.47e+00	1.9424e+01	1.71e+00
p5	3.6683e+00	2.18e-01	3.1583e+00	2.51e-01

Table 8.2: Post-fit parameters for the PFN CR and VR.  $p1$  can also be considered  $N_{bkg}$  or the normalization factor.

The smoothing applied follows the procedure for functional decomposition described in Ref. [90].

Figure 8.11 shows the impact of smoothing on the source data distribution in the CR.

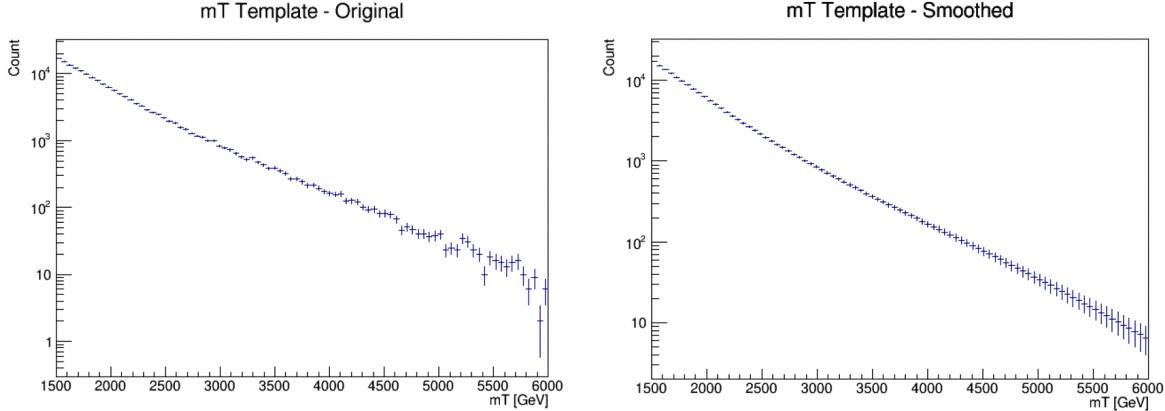


Figure 8.11:  $m_T$  distribution in the data CR, before (left) and after (right) smoothing.

The scaling adjusts the statistics of the smoothed template to the expected statistics of the SR. Recall Figure 8.8, which illustrates that the statistics (or number of events) in the CR and the VR are almost 3x the expected statistics of the SR. The polynomial fitting strategy is sensitive to the statistics of the fitted template, so its performance can vary substantially depending on the statistical power of the fitted distribution. To simulate this, the smoothed template is scaled to the expected statistics of the SR. Toys are then generated from the smoothed distribution, by varying each bin within its statistical uncertainty according to a Poisson distribution. Each toy has the same statistical power as the SR, within statistical uncertainty.

Figure 8.12 shows example fits to three such toy datasets. Figure 8.13 shows the resulting  $p$ -values after an ensemble of 100 Asimov pseudo-datasets are each individually fit. This test determines the likelihood of exceptionally good (high  $p$ -value) or poor (low  $p$ -value) fits due to random statistical fluctuations in the data. A flat distribution is observed, indicating that the data is compatible with the background fit function. By definition, if the null hypothesis is true (meaning the data is compatible with the background expectation) there is a 10% chance of a  $p$ -value less than 0.10, etc. This leads to a flat distribution of  $p$ -values across many tests.

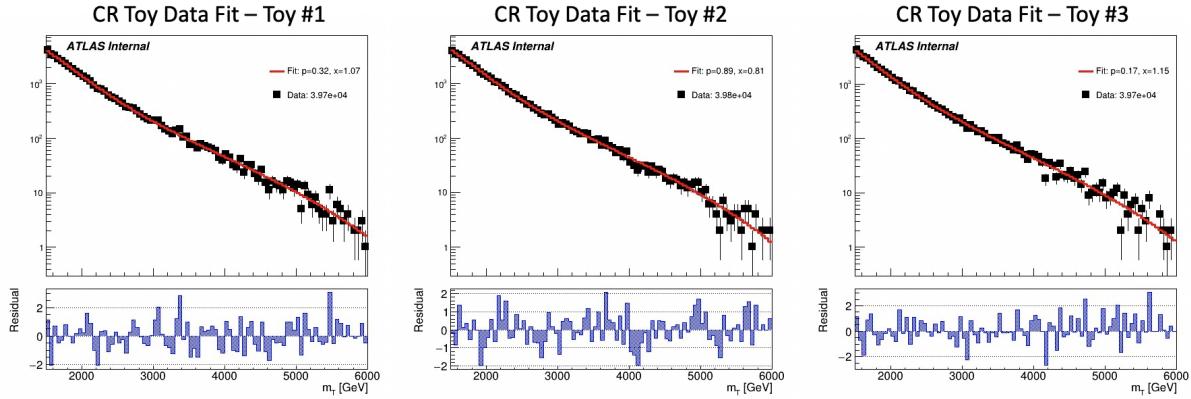


Figure 8.12: Background-only  $m_T$  fits using pseudo-data from the CR template. All three fits are seen to successfully converge, with varying  $p$ -values. The distribution of residuals is reasonably flat for all three fits.

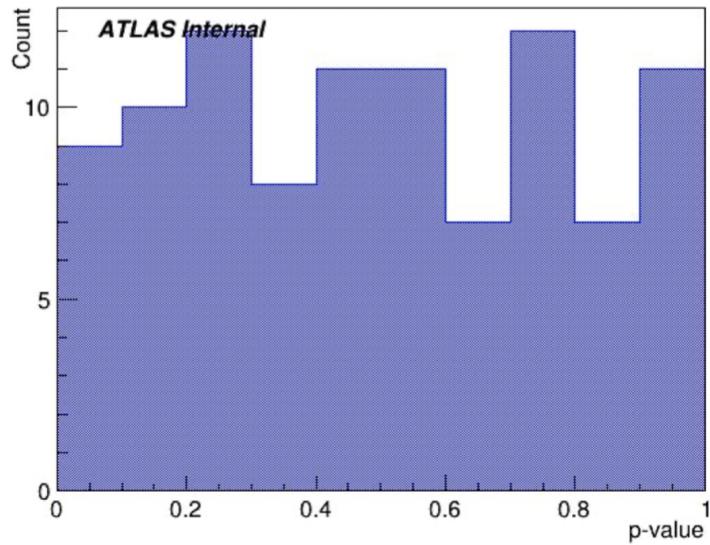


Figure 8.13:  $p$ -value histograms from 100 fits to Asimov data in the CR. The even distribution of  $p$ -values between 0 and 1 indicates that the behavior of the fit is healthy. 98  $p$ -values are shown, as two are excluded due to fits that did not converge on the first try. These fits later converged after the initial parameters were adjusted.

## Signal + Background Fits

Figure 8.14 shows an example of a signal into a pseudo-data CR  $m_T$  spectrum, and the ability of the fit framework to accurately fit the number of signal events. In a signal + background (or s+b) fit, the fit considers the shape of a given signal and estimates a magnitude associated to that signal.

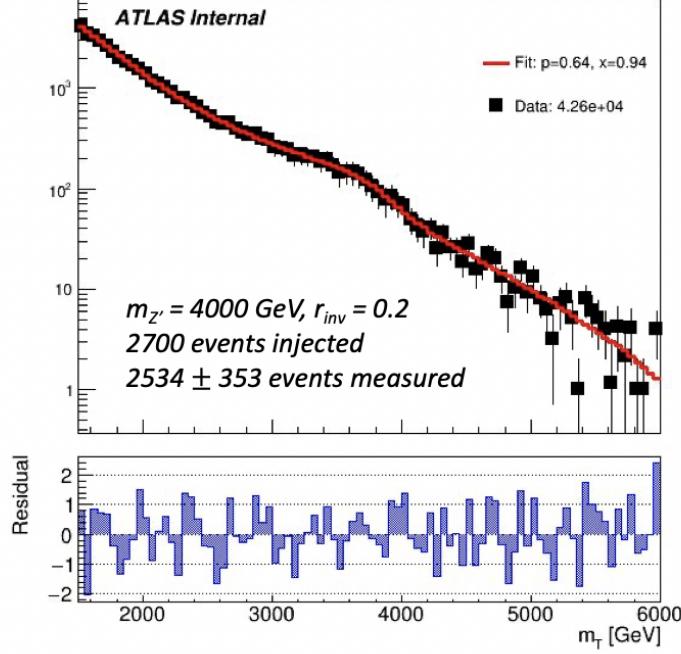


Figure 8.14: Example S+B fit on a background  $m_T$  spectrum with injected signal from the point ( $4000 \text{ GeV}$ ,  $R_{inv} = 0.2$ ). The shape of the injected signal can be seen in Figure 8.1. The ability of the s+b fit to capture the shape of the signal and accurately measure the amount of injected signal is observed.

Signal injection tests demonstrate the linear relationship between the amount of signal injected and the amount of signal measured by the fit. The signal injection tests are performed in Asimov datasets to counter the impact of statistical fluctuations in any given template. 50 Asimov trials are run for all signal points across  $Z'$  mass and  $R_{inv}$ .

Figure 8.15 provides the results of these tests. The uncertainty of the measurement varies according to the  $Z'$  mass, due to the larger relative background for lower mass points. However, a strong linear relationship between amount of signal injected and amount of signal measured is observed for all signal points, which is the key feature. The dashed lines illustrate the linear

relationship, showing that for all points, the amount of signal measured increases as more signal is injected. The variation in the y-intercept of the fitted dashed lines and the exact number of signal events measured for each injection level is not as significant as the overall linear behavior which is exhibited.

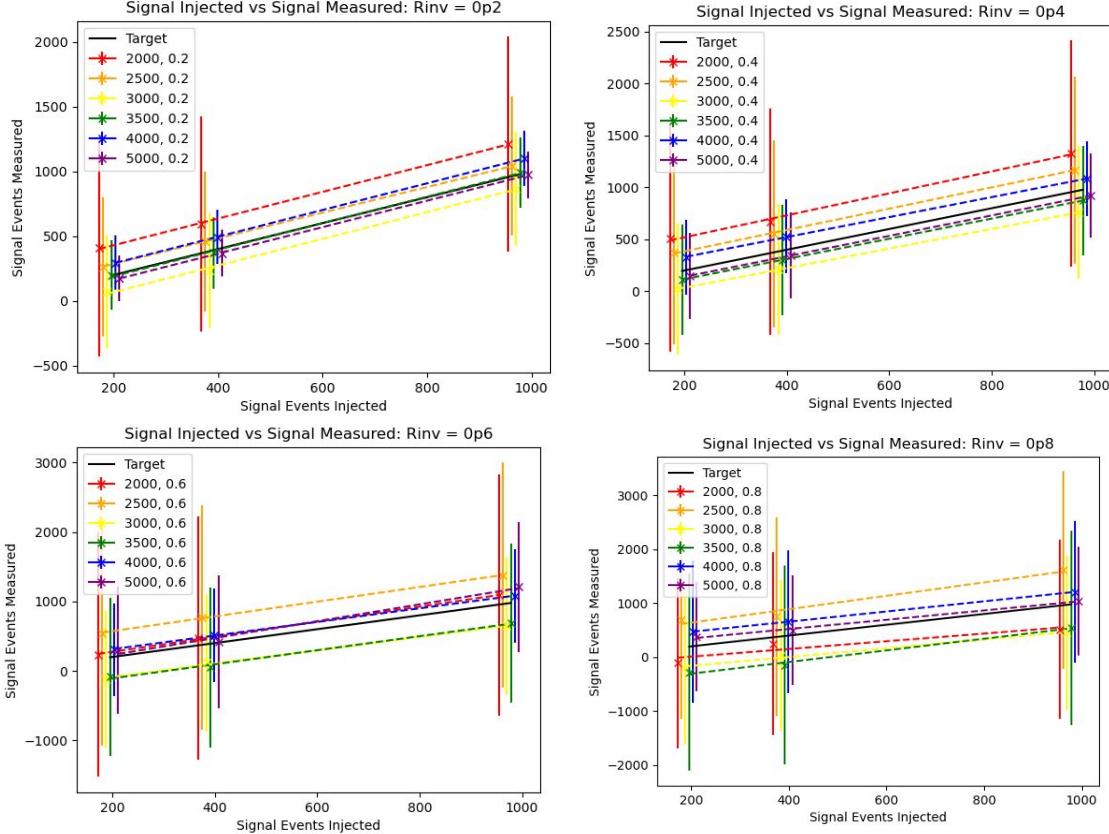


Figure 8.15: Measured signal at a variety of injected values ( $1x$ ,  $2x$ , and  $5x \sqrt{b}$ ), for all signal points in the grid,  $R_{inv} = 0.2$  (top left),  $0.4$  (top right),  $0.6$  (bottom left), and  $0.8$  (bottom right). The x-axis values are slightly shifted from their true value so that all points can be viewed simultaneously. The error bars indicate the standard deviation of the number of fitted events across the 50 Asimov experiments. While the errors are large for some points, the strong linear relationship of the means, illustrated by the dashed lines, is the key feature.

## Expected Sensitivity

Limits on the signal process are obtained by determining the cross section of the signal process  $Z' \rightarrow q_D q_{\bar{D}}$  that can be excluded at the 95% Confidence Level (CL). *Limits* refer to determining the maximum (or *limiting*) signal cross section compatible with the observed data spectrum, such that any theory resulting in a signal cross section above the limit is excluded at the 95% CL. The limit is determined from a maximum likelihood test statistic [91], which determines the likelihood of observing the given data spectrum using the background hypothesis and signal hypothesis. Compatibility of the signal model with the observed distribution is tested by generating pseudo-data based on the background estimation and including varying amounts of signal. Through analysis of these pseudo-data experiments, the maximum number of signals events that is compatible with the observed data distribution can be determined. The 95% confidence level is enforced by dictating that the number of signal events must be compatible with the observed data within  $2\sigma$  of uncertainty.

Figure 8.16 shows the expected limits obtained from an average of 50 Asimov data fits. The fits are signal + background fits performed on a background-only spectrum, which allows the fit to determine the level of signal compatible with the background-only hypothesis. The limits shown include a systematic uncertainty on the yield of the signal, arising from the *spurious signal*<sup>2</sup> which will be discussed in Section 9.1.1.

Considerable exclusion power is predicted for low  $R_{inv}$  signal points and lower mass points, indicated by any points where the theoretical cross section exceeds the observed limit on the cross section. Higher  $R_{inv}$  points present more difficulty due to the very broad shape of the signal in  $m_T$ . Recall Figure 8.1, which illustrates the expected shape of  $m_T$  for varying  $R_{inv}$  points. Higher  $Z'$  mass points are more difficult to exclude due to the lower theory cross sections, as given in Table 6.3.

The ability of the fit to identify a significant excess is tested by calculating the limits on signal injected toys.  $2\sigma$  and  $5\sigma$  of signal is injected for each signal point into 50 Asimov data toys.

---

<sup>2</sup>Spurious signal is the amount of signal measured by the fit in the absence of injected signal.

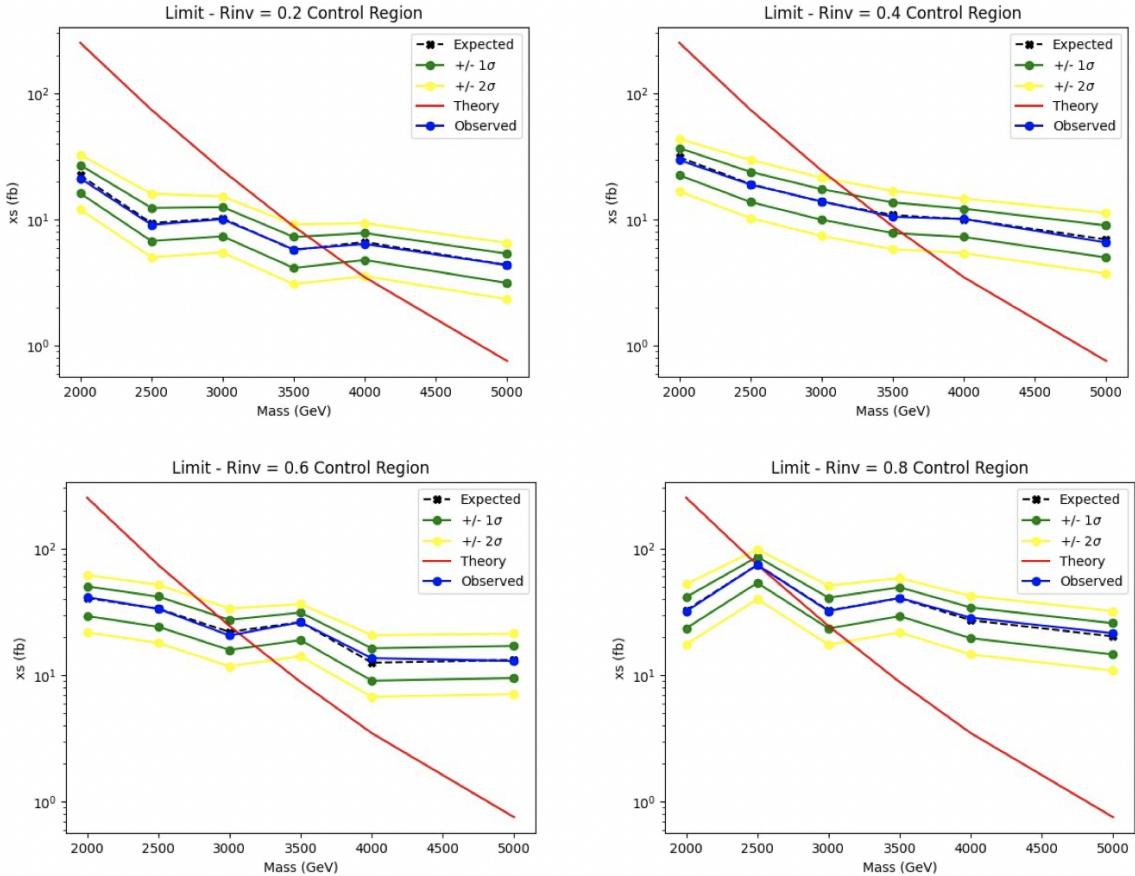


Figure 8.16: 95% C.L. upper limits on the  $Z' \rightarrow q_D \bar{q}_D$  process cross section, derived from the  $m_T$  spectrum in the CR. The red line indicates the theoretical cross section, while the blue line indicates the observed 95% C.L. upper limit on the cross section given the data spectrum. The black line indicates the expected limit given the background shape provided by the fit. The green and yellow bands indicate the uncertainty bands. All signal models across  $Z'$  mass and four different  $R_{inv}$  fractions are shown.

Figure 8.17 demonstrates the impact of this signal injection on the limit, using signals with  $R_{inv} = 0.2$ . The observed limit rises as more signal is injected, indicating the ability of the fit to identify a significant signal excess.

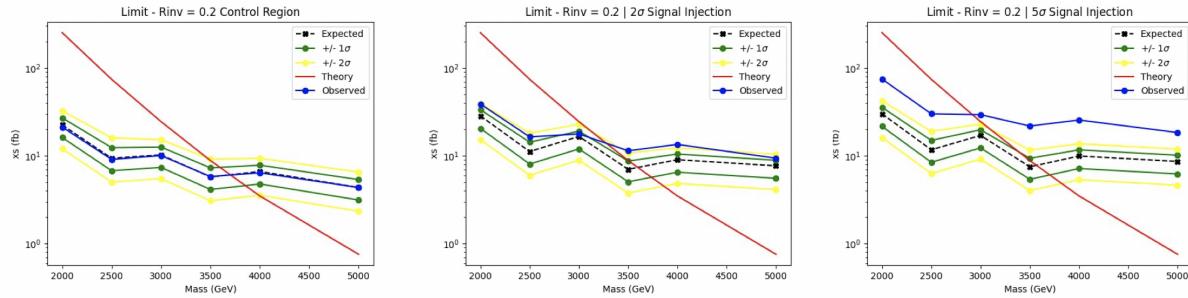


Figure 8.17: 95% C.L. observed limit for signal models across  $Z'$  mass, with varying amounts of signal injected. The increasing observed limit indicates the desired behavior.

### 8.6.2 Discovery Strategy

Model-independent fits for the discovery region are performed using PYBUMPHUNTER [85]. The strategy consists of comparing the data in a given  $m_T$  spectrum to a background estimation derived by performing the polynomial fit. This allows the background estimation method for the Discovery strategy to remain the same as for the SVJ Fit strategy, but introduces a bump hunt for model independent signal interpretation.

The polynomial fit is done to an  $m_T$  distribution with 180 bins (25 GeV wide), which is half the width of the bins used in the SVJ Fit strategy (50 GeV wide). The narrower bins allow for rebinning based on the *signal mass resolution* of the SVJ signals. The binning strategy is outlined in Appendix F.1.

Figure 8.18 shows the fit and residuals of the polynomial with the narrower binning in the CR and the Discovery VR data. Table 8.3 shows the post-fit values of the fit parameters and their uncertainties for the CR and VR. These results indicate good ability of the 5-parameter polynomial to model the ANTELOPE selected data.

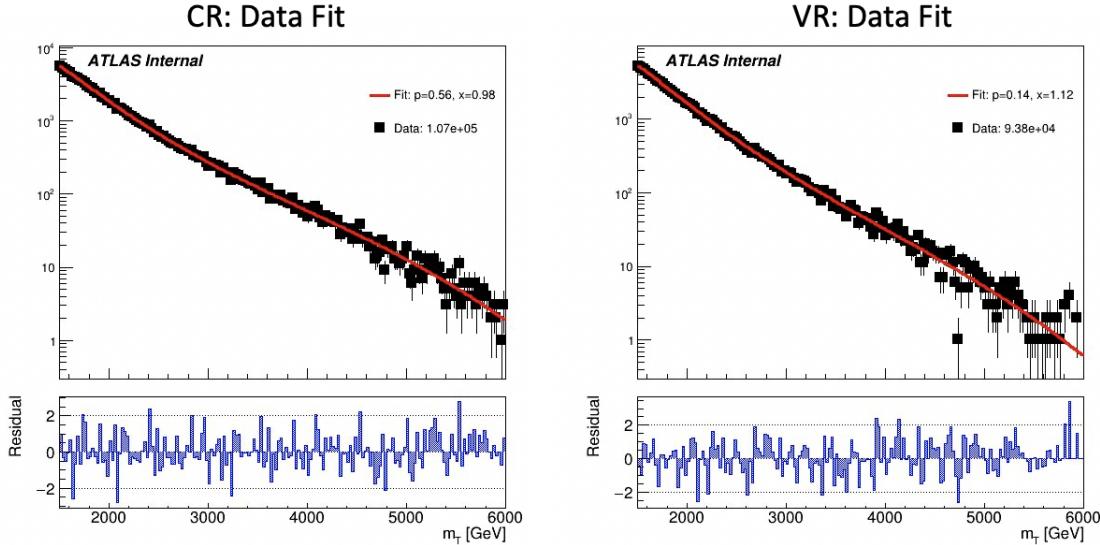


Figure 8.18: Post-fit function and residuals for the ANTELOPE CR and VR.

The studies previously shown in Section 8.6.1 validate the robustness of the background polynomial fit. The narrower bins are the only difference for polynomial fitting between the SVJ Fit

Parameter	CR		VR	
	Value	Error	Value	Error
p1	1.0709e+05	4.65e+02	9.3776e+04	4.34e+02
p2	4.5945e+01	3.03e+00	4.5787e+01	3.78e+00
p3	5.8566e+01	3.88e+00	5.6014e+01	4.58e+00
p4	2.5593e+01	1.50e+00	2.4914e+01	1.73e+00
p5	3.9196e+00	2.22e-01	3.8809e+00	2.53e-01

Table 8.3: Post-fit parameters for the ANTELOPE CR and VR.

and Discovery Fit strategies, and they are not observed to reduce the quality or consistency of the fit.

### BumpHunter Fits

The signal mass resolution binning strategy described in Appendix F.1 creates a monotonically increasing set of bins. While the SVJ signals help inform the binning, the binning is still broadly applicable to a variety of potential signal models. The mass resolution of any resonant signal generally widens as the mass of the mediator particle increases. A similar strategy and binning was used in the generic heavy resonance search presented in Ref. [83]. The resulting set of 15 bins to be used in the BumpHunter (BH) fits varies in width from 100 GeV at the  $m_T$  core to 925 GeV in the  $m_T$  tail.

Figure 8.19 shows the result of running BumpHunter over the rebinned CR and VR  $m_T$  spectra. The background estimation is given by the polynomial fit function. The BH  $p$ -values ( $>0.05$ ) indicate agreement with the background estimation. The BH  $p$ -value for the VR is somewhat low at 0.07, but within the realm of statistical fluctuations, and is thus not a cause for concern. The  $p$ -value of the polynomial fit which gives the background estimation is also reported. The location of the largest bump identified by BH is indicated on the plot.

Figure 8.20 shows BumpHunter  $p$ -values over 100 Asimov trials, where each toy is scaled to the statistics of the SR. The agreement is generally very good, as the  $p$ -values trend towards higher values. Since there is an upward trend in the  $p$ -values rather than a flat distribution as seen for

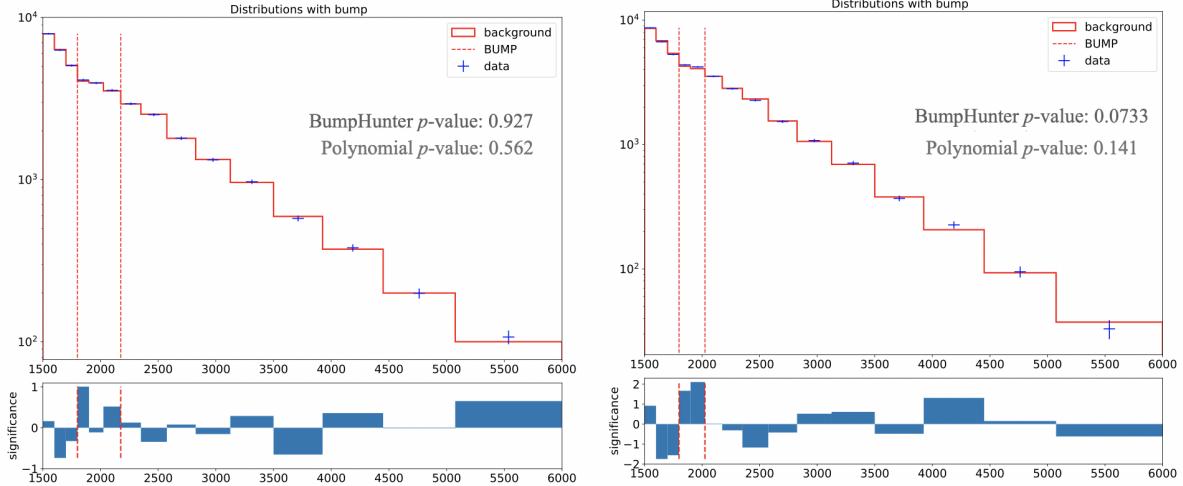


Figure 8.19: BumpHunter fits on the ANTELOPE  $m_T$  spectra for both the CR (left) and VR (right). In a signal-depleted region, good agreement with the background estimation is observed.

the polynomial fit, we can conclude that the background estimation captures the data slightly “too well”, reducing the frequency of low  $p$ -value BH fits. This indicates that some signal sensitivity (or ability to detect distributions with poor data/background agreement) is lost to the over-performant background estimation. This is accounted for in the signal injection studies presented in the next section, and ultimately was determined to not pose a significant issue for the analysis strategy. While the  $p$ -value distribution is not flat, we do see some fits with  $p$ -value  $< 0.10$ , indicating that low  $p$ -values are still possible from a background only distribution. No fits with a *spurious signal* are found. A spurious signal would be indicated by a fit with a BH  $p$ -value  $< 0.01$  or a maximum local significance of  $> 2\sigma$ .

### BumpHunter Signal Injection

To explore a model independent signal hypothesis, signal injection tests in the ANTELOPE region are done with generic Gaussian shapes. Two Gaussian models are built with a mean ranging from 2000 GeV to 5000 GeV and a standard deviation (or *width*) equal to 10 or 20% of the mean value. Figure 8.21 illustrates an injected Gaussian and its effect on the  $m_T$  distribution. The 20% gaussian represents the widest possible signals we might be sensitive to with a BH strategy, while the 10% injection represents a narrower signal peak. As the SVJ model and other dark QCD

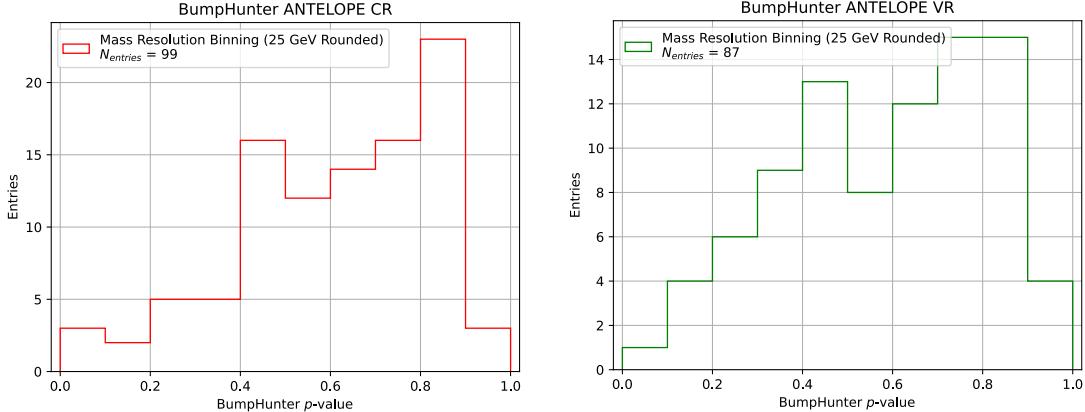


Figure 8.20: BumpHunter p-values extracted for 100 Asimov toys for both the ANTELOPE CR (left) and VR (right). Fits where the polynomial fit initially failed to converge are excluded from the plot. These fits were later recovered through modification of the initial parameters.

models generally predict wide signal shapes due to the presence of  $E_T^{\text{miss}}$  in the final state, narrower Gaussians were not explored.

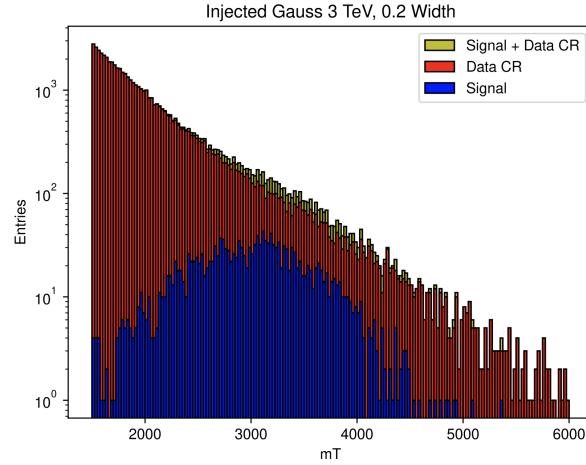


Figure 8.21: Example injected gaussian signal, with mean = 3000 GeV and width of 20%.

An estimated  $5\sigma$  of signal is injected for these tests. The estimate is derived from the amount of signal necessary to produce a  $5\sigma$  excess in the polynomial fit. It is an underestimate of the expected significance to the BumpHunter, as the flexibility of the background-only polynomial fit absorbs some of the signal. Therefore we do not expect to measure  $5\sigma$  significance with the BH approach, but rather hope to see that some level of signal (at least  $\geq 2\sigma$  local significance) is observed by the BumpHunter method.

Results are obtained by averaging over 100 toys for each injection. Figure 8.22 shows the resulting maximum local significance and the location of the determined bump, indicating a good response of the BumpHunter framework for detecting generic  $m_T$  resonances at the right location. Only the 5000 GeV 20% width point is not properly identified by the framework. While some sensitivity is lost due to the flexible nature of the fitting framework, the ability to identify a bump with substantial local significance in the correct location is observed. Figure 8.23 shows an example of the identified bump.

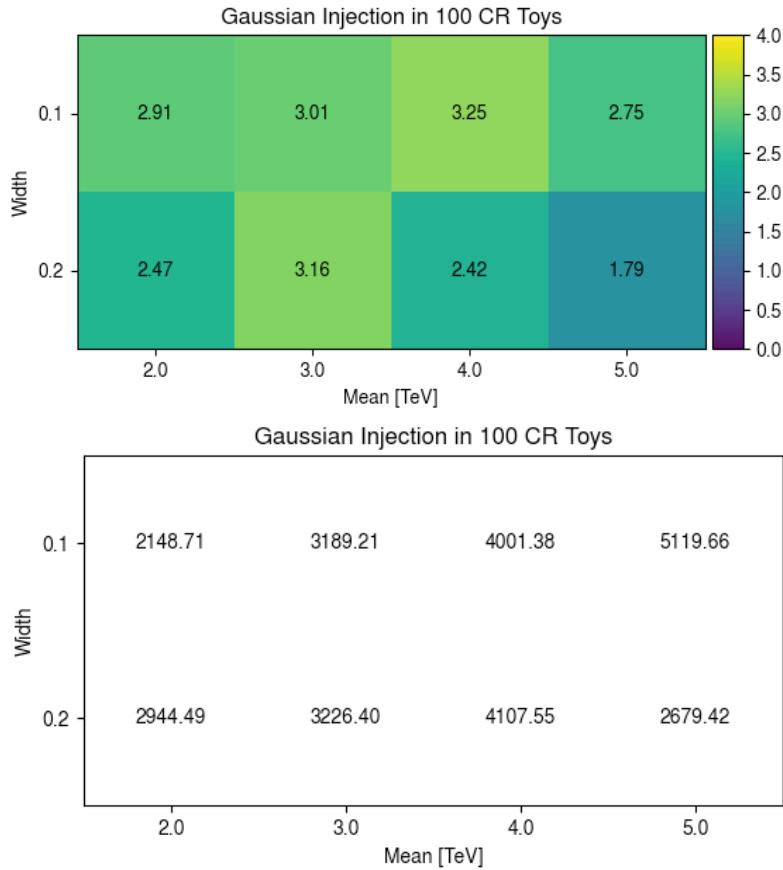


Figure 8.22: Response of the BumpHunter framework to gaussian signal injection is shown. The local significance (top plot) and bump location in GeV (bottom plot) are shown. With the exception of the 5.0 TeV 20% width signal (expressed as 0.2), the BH identifies bumps with a significance  $> 2.0\sigma$  in approximately the correct location.

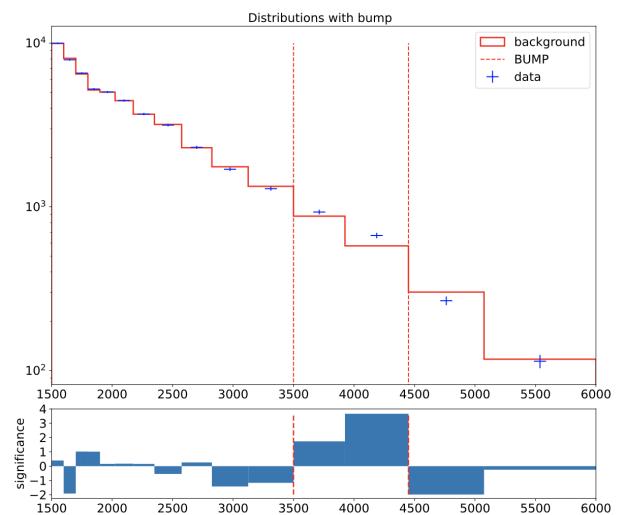


Figure 8.23: Example BH response to gaussian signal injection at 4000 GeV with width of 10%.

## Chapter 9: Results

The final results of this analysis are the polynomial fit to the  $m_T$  distribution in the SVJ Fit SR, and the BumpHunter evaluation of the  $m_T$  distribution in the Discovery SR. In the SVJ Fit region, systematic uncertainties are evaluated on the signal model, and limits on the observed  $Z'$  production cross section are set.

### 9.1 SVJ Fit Result

Figure 9.1 shows the unblinded  $m_T$  spectrum in the SVJ Fit SR with a background-only fit. The fit is successful and has a p-value of 0.265, indicating the data is compatible with the background hypothesis. Table 9.1 gives the values and uncertainties for the five parameters of the polynomial fit function given in Equation 8.2.

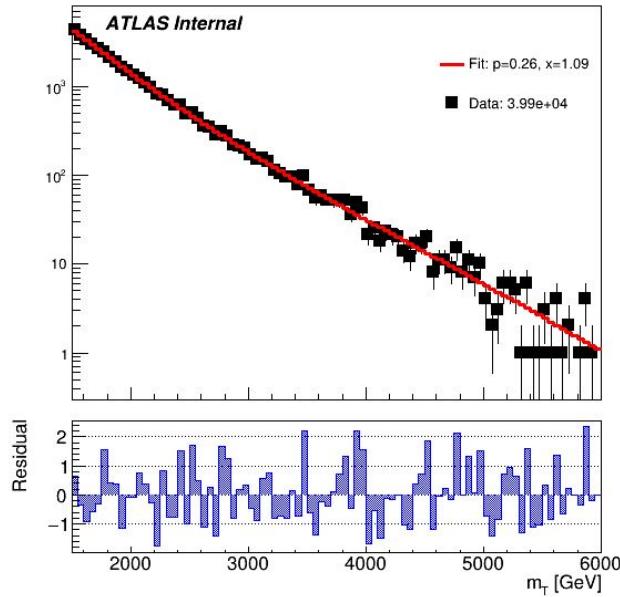


Figure 9.1:  $m_T$  in the unblinded SVJ Fit SR with a background-only fit (p-value = 0.265).

Parameter	SR	
	Value	Error
p1	3.9952e+04	2.83e+02
p2	1.8517e+01	9.08e-01
p3	1.4883e+01	4.43e-01
p4	8.0513e+00	1.21e-01
p5	1.3469e+00	3.62e-02

Table 9.1: Post-fit parameters for the PFN SR.  $p_1$  can also be considered  $N_{bkg}$  or the normalization factor.

### 9.1.1 Systematics

As is typically done in dijet resonance searches using a polynomial fit [92], the systematic uncertainties in this analysis are applied only to the signal and not to the background. This is because the background expectation is determined entirely from the data in the SR via the polynomial fit. Therefore the only uncertainty on the background is the statistical uncertainty, which is reflected in the uncertainty associated to each of the five freely floating parameters determined in the fit.

A variety of systematics on the signal shape and yield are considered. The most significant of these is the *spurious signal* systematic, which quantifies the level of signal observed in the absence of signal injection. Experimental uncertainties on the luminosity and jet reconstruction are studied. Finally, uncertainties on the MC simulation of the SVJ theory model are also considered.

#### Spurious Signal

The spurious signal uncertainty is assessed following the prescription in Ref. [88]. Asimov pseudo-datasets as described in Section 8.6.1 are used to estimate the spurious signal. The spurious signal uncertainty is included in the fit as a systematic uncertainty on the *yield* of each signal point.

The spurious signal  $S_{\text{spur}}$  is quantified for each signal as the mean number of signal events fitted across 100 signal-free pseudo-data experiments. To determine if the amount of spurious signal is tolerable, the threshold  $S_{\text{spur}}/\sigma_{\text{fit}} < 0.5$  is used as recommended in Ref. [88].  $\sigma_{\text{fit}}$  is the standard

deviation on the number of fitted signal events for each signal point across the 100 pseudo-data experiments, and represents the statistical uncertainty on the number of fitted signal events. The approximate total uncertainty on the fitted signal event yield is therefore  $\sigma_{\text{tot}} \approx \sqrt{\sigma_{\text{fit}}^2 + S_{\text{spur}}^2}$  with the addition of the spurious signal systematic. The requirement  $S_{\text{spur}}/\sigma_{\text{fit}} < 0.5$  enforces that the increase in the total measurement uncertainty  $\sigma_{\text{tot}}$  is tolerable at  $< 15\%$  ( $\sqrt{1+0.5^2} \approx 1.12$  for  $S_{\text{spur}} = 0.5\sigma_{\text{fit}}$ ).

Figure 9.2 shows the  $S_{\text{spur}}/\sigma_{\text{fit}}$  metric. The requirement for  $S_{\text{spur}}/\sigma_{\text{fit}} < 0.5$  is easily satisfied across the signal grid.

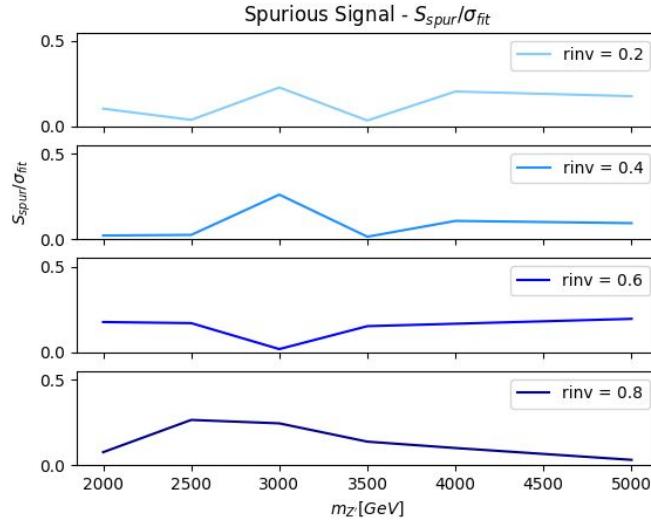


Figure 9.2: Spurious signal metric as a function of resonance mass. The requirement  $S_{\text{spur}}/\sigma_{\text{fit}} < 0.5$  is satisfied for all signal points. 100 pseudo-data experiments are used for the measurement.

The average size of  $S_{\text{spur}}$  is 209 events, and it ranges from 23 events for some  $R_{\text{inv}} = 0.2$  signals to 470 events for some  $R_{\text{inv}} = 0.8$  signals. For most points the spurious signal uncertainty is about 50% of the expected signal yield, though it ranges from 4.2% for  $m_{Z'} = 2000$  to over 100% for  $m_{Z'} = 5000$  GeV. The experimental and theory uncertainties presented in the following sections are generally negligible in the fit due to the size of the spurious signal uncertainty. They are included for completeness, and have some impact for the  $m_{Z'} = 2000$  GeV,  $R_{\text{inv}} = 0.2$  and  $m_{Z'} = 2500$  GeV,  $R_{\text{inv}} = 0.2$  signal points, where the spurious signal uncertainty is  $< 5\%$ .

Uncertainty	Effect on Yield [%]
Luminosity	0.83
JES	0.04 - 1.39
JER	0.01 - 0.64

Table 9.2: Summary of Experimental Uncertainties and their impact on the yield of MC signal events.

## Experimental Uncertainties

The main experimental uncertainties are on the recorded luminosity, *jet energy scale*, and *jet energy resolution*. The jet energy scale (JES) corrects for the non-compensating calorimeter response and jet energy losses in passive detector material [93]. The jet energy resolution (JER) applies a correction which accounts for the spread of possible reconstructed energies for a jet with a fixed true energy, as determined in simulation. Systematics uncertainties on the JES and JER processes must be considered for any analysis using reconstructed jets.

A flat yield uncertainty of 0.83% is applied for all signals, corresponding to the uncertainty reported on the luminosity measurement by the LUCID detector [94].

The JES and JER uncertainties are evaluated on each signal point for their impact on both the yield and shape of the  $m_T$  distribution. Table 9.2 summarizes the range impact on the yield for each uncertainty. The impact of these uncertainties on the signal yield is generally very small in comparison to the spurious signal systematic.

The impact of the JES and JER uncertainties on the shape of the  $m_T$  distribution is also considered. An example individual JES variation is shown in Figure 9.3, illustrating the minimal impact of this uncertainty on the shape of  $m_T$ . The “up” and “down” variations refer to varying the nominal JES or JER setting by  $+1\sigma$  or  $-1\sigma$ . In principle these uncertainties can shift the mean of the  $m_T$  distribution left or right, causing a substantial change in shape. However, the impact of these uncertainties on the shape of  $m_T$  in the SR is seen to be very small.

To make a conservative estimate of their impact on the shape, all shape uncertainty sources are summed in quadrature, bin-by-bin. This results in a maximum  $1\sigma$  “up” variation and a maximum  $1\sigma$  “down” variation. The the impact of these maximal shape variations on the  $Z'$  production cross

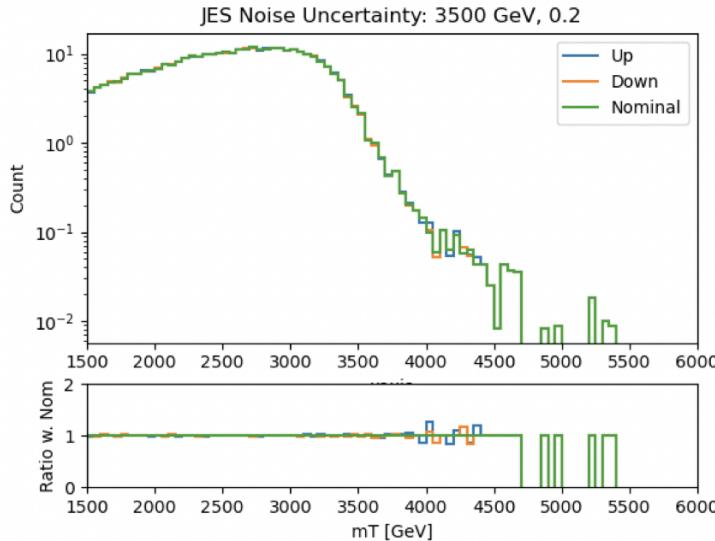


Figure 9.3:  $m_T$  of the 3500 GeV  $Z'$ ,  $R_{inv} = 0.2$  signal point, shown with an example JES noise uncertainty variation. The nominal shape,  $1\sigma$  up, and  $1\sigma$  down variations are shown. The variation is seen to have minimal impact on the signal shape. Signal only (no background) is shown.

section limit is evaluated, and uncertainty on this limit is propagated to the final limit bands. The impact is generally seen to be quite small, changing the limit variation by 0.61 fb at most. An example of the variations summed in quadrature is shown in Figure 9.4.

## Theory Uncertainty

Uncertainty on the parameters of the signal model are also considered. The primary theory uncertainty source is the tuning of the parton shower in PYTHIA8 [95]. Jet structure and extra jet production within the event depend on the modeling of initial state radiation (ISR), final state radiation (FSR) and behavior of multiple parton interactions (MPI) within an event. A variety of MC generation tuning parameters govern the behavior of ISR, FSR and MPI in the signal generation. Ref [96] describes how these parameters are condensed into 10 variations which capture the maximal range of impact for these tuning parameters.

The 10 variations (representing 5 up/down variation pairs) are evaluated for the SVJ signal shapes. Figure 9.5 provides a look at the effect of these variations on the SVJ  $m_T$  signal shape.

There is no substantial sculpting of the  $m_T$  shape from any of the 10 systematic variations.

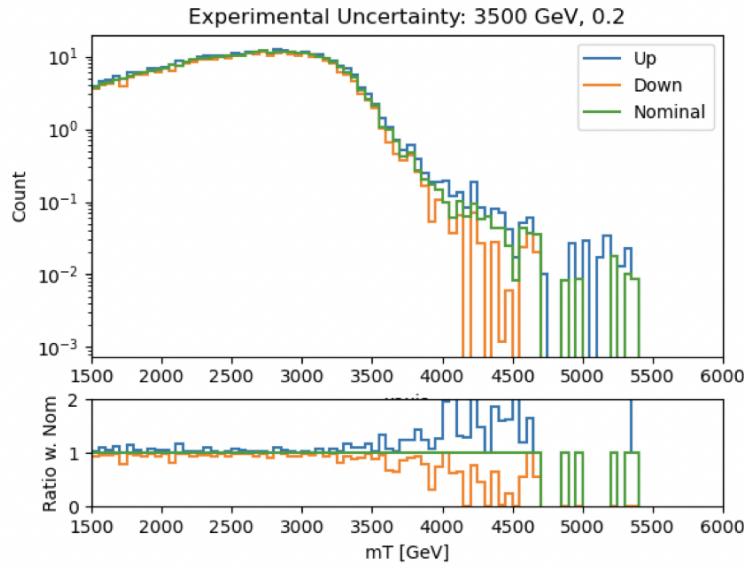


Figure 9.4:  $m_T$  of the 3500 GeV  $Z'$ ,  $R_{inv} = 0.2$  signal point, shown with the sum in quadrature of all JES and JER variations. The nominal shape before systematic variations, the maximal  $1\sigma$  “up”, and maximal  $1\sigma$  “down” variations are shown. Even summed in quadrature, the effect of the JES and JER variations on the shape of the signal is seen to be small.

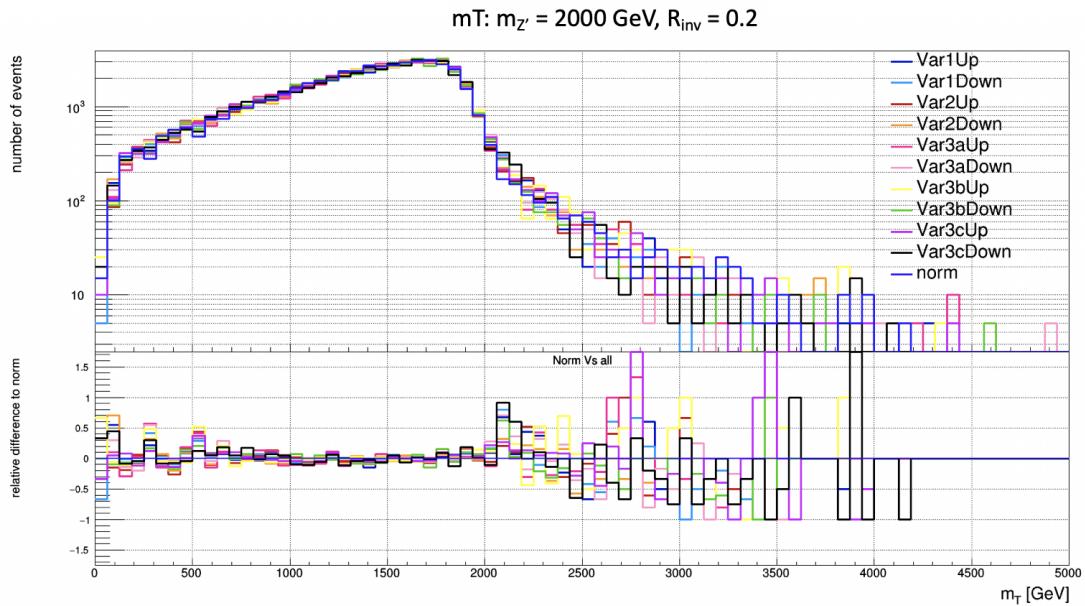


Figure 9.5: Signal distribution of  $m_T$ , varying the ISR (Var1), FSR (Var2) and MPI (Var3a-c) configurations.

Particularly through the core of the distribution, the net impact of the variations is flat. thus the systematic is considered for its impact on the signal yield. The variation in the signal yield is at most 1.2%. A conservative 2% yield uncertainty is applied to account for the uncertainty on the theory model. The spurious signal uncertainty is dominant for all but the lowest mass signal points.

### 9.1.2 Interpretation

Using a modified frequentist approach [97], *exclusion limits* at the 95% confident level (CL) are derived. The process was first described in Section 8.6.1 and is reviewed here. Exclusion limits refer to determining the maximum (or *limiting*) signal cross section compatible with the observed data spectrum, such that any theory resulting in a signal cross section above the limit is excluded with 95% confidence. The limit is determined from a maximum likelihood test statistic [91], which determines the likelihood of observing the given data spectrum using the background hypothesis, signal hypothesis, and uncertainty parameters. Compatibility of the signal model with the observed distribution is tested by generating pseudo-data based on the background estimation and including varying amounts of signal. Through analysis of these pseudo-data experiments, the maximum number of signals events that is compatible with the observed data distribution and background estimation can be determined. The 95% confidence level is enforced by dictating that the number of signal events must be compatible with the observed data within  $2\sigma$  of uncertainty.

The final limits on the  $Z'$  cross section after the implementation of the systematic uncertainties are shown in Figure 9.6. Exclusion of the theoretical model is observed for the 2000 GeV  $Z'$  mass point for all  $R_{inv}$  values. We are unable to exclude the highest mass points due to their low theoretical cross section, and relatively high spurious signal uncertainty. The most mass points are excluded for  $R_{inv} = 0.2$ , which excludes  $Z'$  masses up to 3500 GeV.

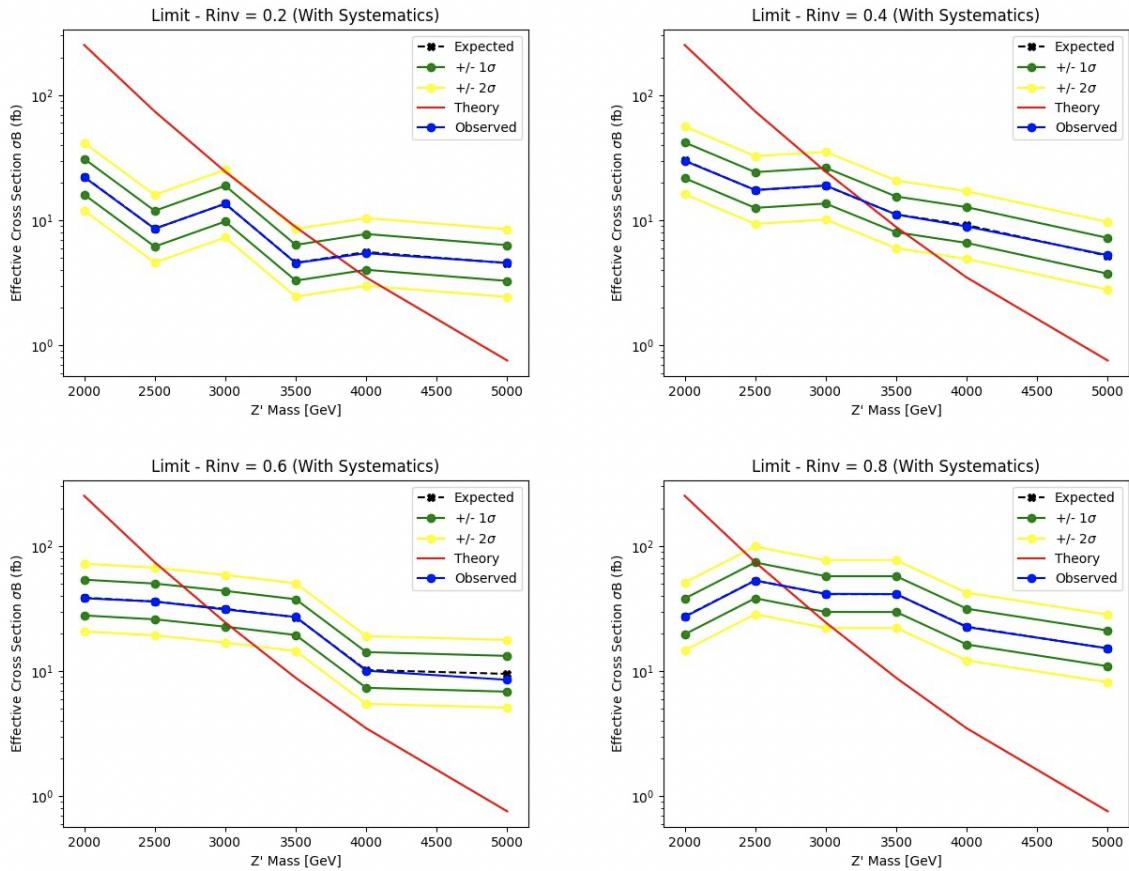


Figure 9.6: Expected and observed 95% C.L. limits in the unblinded SR, as a function of  $Z'$  mass for  $R_{inv} = 0.2$  (top left), 0.4 (top right), 0.6 (bottom left), 0.8 (bottom right). The red line indicates the theoretical cross section, while the blue line indicates the observed 95% C.L. upper limit on the cross section given the data spectrum. The black line indicates the expected limit given the background shape provided by the fit. The green and yellow bands indicate the uncertainty bands.

## 9.2 Discovery Result

Figure 9.7 shows the unblinded  $m_T$  spectrum in the Discovery SR with a background-only fit, and the resulting BumpHunter test. The polynomial fit is successful and has a background-only p-value of 0.74, indicating the data is compatible with the background hypothesis. The BumpHunter test gives a p-value of 0.8098, indicating no significant excess. The maximum local significance is  $0.877\sigma$ , located at 1700 GeV.

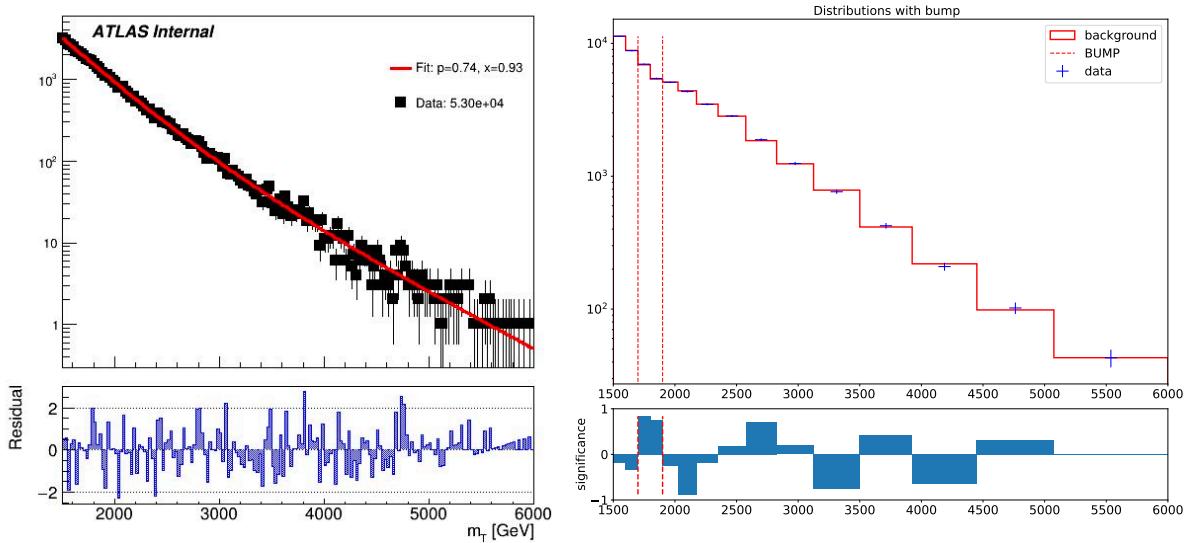


Figure 9.7:  $m_T$  in the unblinded ANTELOPE SR with a background-only fit (p-value = 0.74), left. BumpHunter test selecting the most significant data excess with a p-value of 0.8098, right.

Because there is no specific signal interpretation for the Discovery region and both the polynomial fit and BH analysis are entirely data driven, there are no systematics to consider in the interpretation of the BH result.

## Conclusion

This thesis presents a search for hadronic signatures of a strongly-coupled hidden dark sector, accessed via resonant production of a  $Z'$  mediator. A model in which the massive  $Z'$  mediator decays to two dark quarks which each shower partially back to the visible sector is presented. The resulting shower of dark and Standard Model particles creates a “semi-visible jet”.

The analysis is performed using  $139 \text{ fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13 \text{ TeV}$  of center-of-mass energy collected by the ATLAS experiment during Run 2 of the LHC. The analysis isolates events with visible and missing energy contributions which are consistent with the semi-visible jet topology. Two machine learning approaches are used to identify events consistent with the hadronic signatures of the hidden dark sector. The first is a supervised permutation invariant classifier approach which models jets according to their associated tracks. The second is a novel anomaly detection approach which couples a permutation invariant event modeling derived from the supervised approach with an unsupervised variational auto-encoder. The anomaly detection approach broadens the sensitivity of the analysis to a wider variety of possible dark jet signatures.

The background is estimated via a polynomial fit of the transverse mass spectrum. No significant signal excesses are observed with respect to the expected background processes. Upper limits at the 95% C.L. on the effective  $Z'$  cross section are set for signal models with  $m_{Z'}$  ranging from 2000 to 5000 GeV and  $R_{inv}$  ranging from 0.2 to 0.8. The anomaly detection approach reports no significant excess, with a maximum observed local significance of  $0.877\sigma$ .

The results presented here represent exploration of a new phase space in ATLAS Run 2 data. While no evidence of a strongly-coupled hidden dark sector is found, a novel architecture for permutation invariant low-level anomaly detection is presented. As ATLAS continues to collect data in Run 3 and beyond, further explorations of Hidden Valley models which go beyond the scope of this analysis are possible.

## References

- [1] Jens Erler and Paul Langacker. “Electroweak model and constraints on new physics”. In: (July 2004). arXiv: hep-ph/0407097.
- [2] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. New York, NY: Wiley, 2008.
- [3] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 2018), pp. 847–851.
- [4] E. Noether. “Invariante Variationsprobleme”. In: *Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse*, Seite 235-157 (1918). eprint: www.physics.ucla.edu/\~\cwp/articles/noether.trans/german/emmy235.html.
- [5] J. H. Christenson et al. “Evidence for the  $2\pi$  Decay of the  $K_2^0$  Meson”. In: *Phys. Rev. Lett.* 13 (1964), pp. 138–140.
- [6] Michael Gronau. “CP Violation in B Meson Decays”. In: *Nuclear Physics B - Proceedings Supplements* 142 (May 2005), 263–270.
- [7] J. E. Augustin et al. “Discovery of a Narrow Resonance in  $e^+e^-$  Annihilation”. In: *Phys. Rev. Lett.* 33 (1974), pp. 1406–1408.
- [8] J. J. Aubert et al. “Experimental Observation of a Heavy Particle  $J$ ”. In: *Phys. Rev. Lett.* 33 (1974), pp. 1404–1406.
- [9] Martin L. Perl et al. “Evidence for Anomalous Lepton Production in e+ - e- Annihilation”. In: *Phys. Rev. Lett.* 35 (1975), pp. 1489–1492.
- [10] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (1977), pp. 252–255.
- [11] F. Abe et al. “Observation of top quark production in  $\bar{p}p$  collisions”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. arXiv: hep-ex/9503002.
- [12] S. Abachi et al. “Observation of the top quark”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. arXiv: hep-ex/9503003.
- [13] K. Kodama et al. “Observation of tau neutrino interactions”. In: *Phys. Lett. B* 504 (2001), pp. 218–224. arXiv: hep-ex/0012035.

- [14] G. Arnison et al. “Experimental Observation of Lepton Pairs of Invariant Mass Around 95- $\text{GeV}/c^{**2}$  at the CERN SPS Collider”. In: *Phys. Lett. B* 126 (1983), pp. 398–410.
- [15] P. Bagnaia et al. “Evidence for  $Z^0 \rightarrow e^+e^-$  at the CERN  $\bar{p}p$  Collider”. In: *Phys. Lett. B* 129 (1983), pp. 130–140.
- [16] Serguei Chatrchyan et al. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 30–61. arXiv: 1207.7235 [hep-ex].
- [17] Georges Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29. arXiv: 1207.7214 [hep-ex].
- [18] K. G. Begeman, A. H. Broeils, and R. H. Sanders. “Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics”. In: *Mon. Not. Roy. Astron. Soc.* 249 (1991), p. 523.
- [19] Y. Ashie et al. “Evidence for an oscillatory signature in atmospheric neutrino oscillation”. In: *Phys. Rev. Lett.* 93 (2004), p. 101801. arXiv: hep-ex/0404034.
- [20] C. Abel et al. “Measurement of the Permanent Electric Dipole Moment of the Neutron”. In: *Phys. Rev. Lett.* 124.8 (2020), p. 081803. arXiv: 2001.11966 [hep-ex].
- [21] Guillaume Albouy et al. “Theory, phenomenology, and experimental avenues for dark showers: a Snowmass 2021 report”. In: *The European Physical Journal C* 82.12 (Dec. 2022).
- [22] Timothy Cohen et al. “LHC searches for dark sector showers”. In: *Journal of High Energy Physics* 2017.11 (Nov. 2017).
- [23] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008), S08001.
- [24] “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08003.
- [25] “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08004.
- [26] “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08002.
- [27] “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005.
- [28] Ana Lopes and Melissa Loyse Perrey. *FAQ-LHC The guide*. 2022.

- [29] Esma Mobs. “The CERN accelerator complex in 2019. Complexe des accélérateurs du CERN en 2019”. In: (2019). General Photo.
- [30] *Pulling together: Super Conducting electromagnets*. <https://home.cern/science/engineering/pulling-together-superconducting-electromagnets>. Accessed: 2024-01-05.
- [31] *The High-Luminosity LHC*. <https://voisins.web.cern.ch/en/high-luminosity-lhc-hl-lhc>. Accessed: 2024-01-05.
- [32] Aad G., et al. (ATLAS Collaboration and CMS Collaboration). “Combined Measurement of the Higgs Boson Mass in pp Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments”. In: *Phys. Rev. Lett.* 114 (19 2015), p. 191803.
- [33] O. Aberle et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2020.
- [34] The ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08003.
- [35] G. Aad, B. Abbott, and ATLAS Collaboration. “Performance of the reconstruction of large impact parameter tracks in the inner detector of ATLAS”. In: *Eur. Phys. J. C Part. Fields* 83.11 (Nov. 2023).
- [36] Joao Pequenao. *Computer Generated image of the ATLAS calorimeter*. 2008.
- [37] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996.
- [38] F. Lobkowicz. “Liquid argon calorimetry at the SSC”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 315.1 (1992), pp. 273–278.
- [39] Henric Wilkens and ATLAS LAr Collaboration. “The ATLAS Liquid Argon calorimeter: An overview”. In: *Journal of Physics: Conference Series* 160.1 (2009), p. 012043.
- [40] *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*. Tech. rep. Geneva: CERN, 2017.
- [41] *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*. Tech. rep. Geneva: CERN, 2017.
- [42] L. Pontecorvo. “The ATLAS muon spectrometer”. In: *Eur. Phys. J. C* 34 (2004). Ed. by Juergen Schuhkraft et al., S117–S128.

- [43] *ATLAS magnet system: Technical Design Report, 1.* Technical design report. ATLAS. Geneva: CERN, 1997.
- [44] *ATLAS Experiment Magnet System.* url=<https://atlas.cern/Discover/Detector/Magnet-System>, Accessed: 2024-03-05. 2024.
- [45] Tommaso Colombo. “Data-flow Performance Optimisation on Unreliable Networks: the ATLAS Data-Acquisition Case”. In: *Journal of Physics: Conference Series* 608 (May 2015), p. 012005.
- [46] Joao Pequenao. “Event Cross Section in a computer generated image of the ATLAS detector.” 2008.
- [47] ATLAS Collaboration. “ATLAS Experiment Implements Heterogeneous Particle Reconstruction with Intel oneAPI Tools”. 2023.
- [48] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.12 (2019), P12006.
- [49] Chiara Deponte. “Studies on the properties of non-prompt photons at the ATLAS experiment”. Presented 16 Aug 2022. Technische Universitaet Dortmund (DE), 2022.
- [50] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s}=13$  TeV”. In: *The European Physical Journal C* 76.5 (2016).
- [51] Sebastien Rettie. *Muon identification and performance in the ATLAS experiment*. Tech. rep. Geneva: CERN, 2018.
- [52] B. R. Webber. *Fragmentation and Hadronization*. 1999. arXiv: hep-ph/9912292 [hep-ph].
- [53] Eric M. Metodiev. *The Fractal Lives of Jets* | Eric M. Metodiev — ericmetodiev.com. <https://www.ericmetodiev.com/post/jetformation/>. 2019, note = Accessed 18-05-2024,
- [54] Steven Schramm. *ATLAS Jet Reconstruction, Calibration, and Tagging of Lorentz-boosted Objects*. Tech. rep. Geneva: CERN, 2017.
- [55] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *The European Physical Journal C* 77.7 (July 2017).
- [56] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *The European Physical Journal C* 77.7 (July 2017).

- [57] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti-kt jet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), 063–063.
- [58] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet user manual: (for version 3.0.2)”. In: *The European Physical Journal C* 72.3 (Mar. 2012).
- [59] Stephen D. Ellis and Davison E. Soper. “Successive combination jet algorithm for hadron collisions”. In: *Physical Review D* 48.7 (Oct. 1993), 3160–3166.
- [60] M. Wobisch and T. Wengler. *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering*. 1999. arXiv: hep-ph/9907280 [hep-ph].
- [61] Gavin P Salam and Gregory Soyez. “A practical seedless infrared-safe cone jet algorithm”. In: *Journal of High Energy Physics* 2007.05 (May 2007), 086–086.
- [62] Gavin P. Salam. “Towards jetography”. In: *The European Physical Journal C* 67.3–4 (May 2010), 637–686.
- [63] *A Monte Carlo study of track association to jets for b-tagging*. Tech. rep. Geneva: CERN, 2021.
- [64] *Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector*. Tech. rep. Geneva: CERN, 2014.
- [65] ATLAS Collaboration. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *Eur. Phys. J. C* 78.11 (2018), p. 903. arXiv: 1802.08168.
- [66] ATLAS Collaboration. *Public atlas luminosity results for run-2 of the LHC*. 2024.
- [67] GEANT4 Collaboration, S. Agostinelli, et al. “GEANT4 – a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250.
- [68] Christian Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. arXiv: 2203.11601 [hep-ph].
- [69] Michelangelo L. Mangano, Mauro Moretti, and Roberto Pittau. “Multijet matrix elements and shower evolution in hadronic collisions: W b bbar + b jets as a case study”. In: *Nuclear Physics B* 632.1–3 (June 2002), 343–362.
- [70] The CMS Collaboration. “Search for resonant production of strongly coupled dark matter in proton-proton collisions at 13 TeV”. In: *Journal of High Energy Physics* 2022.6 (June 2022).

- [71] The ATLAS Collaboration. “Search for non-resonant production of semi-visible jets using Run 2 data in ATLAS”. In: *Physics Letters B* 848 (Jan. 2024), p. 138324.
- [72] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. arXiv: 1405.0301 [hep-ph].
- [73] Jon Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *J. Phys. G* 43 (2016), p. 023001. arXiv: 1510.03865 [hep-ph].
- [74] Peter Skands, Stefano Carrazza, and Juan Rojo. “Tuning PYTHIA 8.1: the Monash 2013 Tune”. In: *Eur. Phys. J. C* 74.8 (2014), p. 3024. arXiv: 1404.5630 [hep-ph].
- [75] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *Nature* 521.7553 (2015), p. 436.
- [76] Juri Opitz. *Schroedinger’s Threshold: When the AUC doesn’t predict Accuracy*. 2024. arXiv: 2404.03344.
- [77] Dor Bank, Noam Koenigstein, and Raja Giryes. *Autoencoders*. 2021. arXiv: 2003.05991 [cs.LG].
- [78] A. Kahn et al. “Anomalous jet identification via sequence modeling”. In: *Journal of Instrumentation* 16.08 (Aug. 2021), P08012.
- [79] Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler. “Energy flow networks: deep sets for particle jets”. In: *Journal of High Energy Physics* 2019.1 (2019).
- [80] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.
- [81] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: 1412.6980 [cs.LG].
- [82] Diederik P. Kingma and Max Welling. *Auto-Encoding Variational Bayes*. 2022. arXiv: 1312.6114 [stat.ML].
- [83] Georges Aad et al. “Anomaly detection search for new resonances decaying into a Higgs boson and a generic new particle  $X$  in hadronic final states using  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS detector”. In: *Phys. Rev. D* 108 (2023), p. 052009. arXiv: 2306.03637 [hep-ex].
- [84] *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*. Tech. rep. Geneva: CERN, 2015.

- [85] Georgios Choudalakis. *On hypothesis testing, trials factor, hypertests and the BumpHunter*. 2011. arXiv: 1101.0390.
- [86] Peter Loch. “Jet measurements in ATLAS”. In: *J. Phys. Conf. Ser.* 323 (2011). Ed. by Giorgio Bellettini, p. 012002.
- [87] The ATLAS Collaboration. “Search for diboson resonances in hadronic final states in 139 fb<sup>-1</sup> of  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: *Journal of High Energy Physics* 2019.9 (Sept. 2019).
- [88] *Recommendations for the Modeling of Smooth Backgrounds*. Tech. rep. Geneva: CERN, 2020.
- [89] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (Feb. 2011).
- [90] Ryan Edgar et al. *Functional Decomposition: A new method for search and limit setting*. 2018. arXiv: 1805.04536 [physics.data-an].
- [91] *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. Geneva: CERN, 2011.
- [92] “Dijet Resonance Search with Weak Supervision Using  $\sqrt{s} = 13$   $pp$  Collisions in the ATLAS Detector”. In: *Physical Review Letters* 125.13 (Sept. 2020).
- [93] The ATLAS Collaboration. “Jet energy scale and resolution measured in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: *The European Physical Journal C* 81.8 (Aug. 2021).
- [94] The ATLAS Collaboration. “Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC”. In: *The European Physical Journal C* 83.10 (Oct. 2023).
- [95] S. Mrenna and P. Skands. “Automated parton-shower variations in pythia 8”. In: *Physical Review D* 94.7 (Oct. 2016).
- [96] *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-021>. Geneva: CERN, 2014.
- [97] A L Read. “Presentation of search results: the CLs technique”. In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), p. 2693.

## Appendix A: Trigger Studies

Both the lowest unprescaled  $E_T^{\text{miss}}$  and single small-R jet triggers were considered for this analysis. The  $E_T^{\text{miss}}$  trigger is observed to have higher efficiency for low mass, high  $R_{\text{inv}}$  points, while the single small-R trigger favors high mass, low  $R_{\text{inv}}$  points. Figure A.1 shows the yields and signal efficiencies across the grid for both these strategies.

JET Trigger Yields: Scaled to 139 fb <sup>-1</sup>					MET Trigger Yields: Scaled to 139 fb <sup>-1</sup>				
	0.2	0.4	0.6	0.8		0.2	0.4	0.6	0.8
500		17084	17202	15958		62059	85610	105833	
750	12078		8680	7944	22328		42779	44761	
1000	13069	9150	6275	4389	12565	18589	21131	20403	
1500	6985	5179		1880	4107	5420		5114	
2000	2306	1853	1294	720	1403	1705	1751	1515	
3000	268			104	187			181	
4000		35		17		32		28	
6000		2	1	1		2	2	2	
	Background:					Background:			
	2.14E+09					2.22E+10			
Signal Efficiency									
	0.2	0.4	0.6	0.8		0.2	0.4	0.6	0.8
500		0.03	0.03	0.03		0.09	0.14	0.20	
750	0.06		0.05	0.05	0.11		0.25	0.30	
1000	0.19	0.14	0.10	0.08	0.18	0.27	0.34	0.38	
1500	0.53	0.40		0.18	0.31	0.42		0.48	
2000	0.70	0.57	0.41	0.26	0.42	0.52	0.56	0.54	
3000	0.78			0.35	0.55			0.61	
4000		0.66		0.36		0.61		0.60	
6000		0.41	0.33	0.23		0.43	0.48	0.49	

Figure A.1: Trigger yield and efficiency for both the MET trigger and small-R jet trigger approach. Each entry represent a signal point, labelled by the  $Z'$  mass and the  $R_{\text{inv}}$  fraction.

The cross section is higher for the lower  $Z'$  mass signal points. As a result, our sensitivity to these points and ability to set limits on them is already naturally enhanced. Figure A.2 shows the factor of improvement in the inclusive  $S/\sqrt{B}$  using each trigger strategy. A cross-check was also done calculating  $S/\sqrt{B}$  in windows around the mT mass. The results of this cross-check confirmed the results shown in Figure A.2.

Figure A.3 shows the ratio of  $S/\sqrt{B}$  across the signal grid for the jet trigger compared to the  $E_T^{\text{miss}}$  trigger.

This plot makes clear that the  $E_T^{\text{miss}}$  trigger favors the low mass, high  $R_{\text{inv}}$  signal points, while

$(\text{JET Trigger } S/\sqrt{B}) / (\text{Untriggered } S/\sqrt{B})$				
	0.2	0.4	0.6	0.8
500		1.26	1.43	1.53
750	2.99		2.52	2.68
1000	9.24	6.74	5.03	4.07
1500	26.45	20.15		8.83
2000	34.66	28.24	20.57	12.85
3000	39.05			17.35
4000		32.94		17.96
6000		20.57	16.55	11.53

$(\text{MET Trigger } S/\sqrt{B}) / (\text{Untriggered } S/\sqrt{B})$				
	0.2	0.4	0.6	0.8
500		1.43	2.21	3.15
750	1.72		3.86	4.69
1000	2.76	4.26	5.26	5.89
1500	4.83	6.56		7.47
2000	6.55	8.08	8.65	8.40
3000	8.46			9.38
4000		9.51		9.32
6000		6.68	7.43	7.59

Figure A.2: The factor of improvement in  $S/\sqrt{B}$  for each trigger method compared to the untriggered case.

	0.2	0.4	0.6	0.8
500		0.89	0.65	0.49
750	1.74		0.65	0.57
1000	3.35	1.58	0.96	0.69
1500	5.47	3.07		1.18
2000	5.29	3.49	2.38	1.53
3000	4.62			1.85
4000		3.46		1.93
6000		3.08	2.23	1.52

Figure A.3: The ratio of  $S/\sqrt{B}$  of jet trigger over  $E_T^{\text{miss}}$  trigger selection.

the single jet trigger favors the high mass, low  $R_{inv}$  signal points. A number of considerations led us to selecting the single jet trigger. First, Figure A.2 illustrates that the sensitivity enhancement is greater for signals favored by the jet trigger than signals favored by the  $E_T^{\text{miss}}$  trigger. This indicates we have more to gain from focusing on the region of our signal grid where the jet trigger is most efficient. We combined this with the observation that the jet trigger is beneficial for high mass points where the cross section is lower and we need to maximize sensitivity to hope to set limits on these points.

Second, there are a number of analysis variables at our disposal which are more discriminant for high  $R_{inv}$  signals than for low  $R_{inv}$  signals. A collection of these variables are shown in Figure A.4. Given that we had multiple avenues available to enhance sensitivity to low  $R_{inv}$  signals through analysis strategy cuts, we opted for a trigger strategy that benefitted the low  $R_{inv}$  signal points. Additionally, we know  $E_T^{\text{miss}}$  is highly correlated with these discriminant analysis variables, and a  $E_T^{\text{miss}}$  trigger would reduce the effectiveness of these variables. Not implementing the  $E_T^{\text{miss}}$  trigger allows us to explore other ways of leveraging  $E_T^{\text{miss}}$  in the analysis using a cut that is more finely tuned to the specifics of the signal model than the  $E_T^{\text{miss}}$  trigger is.

Third, we know that fitting the high  $R_{inv}$  mass points would be challenging given their very broad shapes in the key analysis variable  $m_T$ . Given the analysis decision to do a search for resonant features in  $m_T$ , we chose to design a search that emphasized maximizing sensitivity to the signal points that we were mostly to appear as a resonant feature in  $m_T$ . The shape of  $m_T$  and our reduced ability to set limits on high  $R_{inv}$  points is demonstrated in the body of this note. Another search strategy may be more optimal for setting limits on high  $R_{inv}$  semi-visible jet signals, but it is currently outside the scope of this analysis.

A jet or  $E_T^{\text{miss}}$  trigger was also considered, but abandoned due to concerns about sculpting features in the smoothly falling  $E_T^{\text{miss}}$  or leading jet  $p_T$  distributions, as shown in Figure A.5.

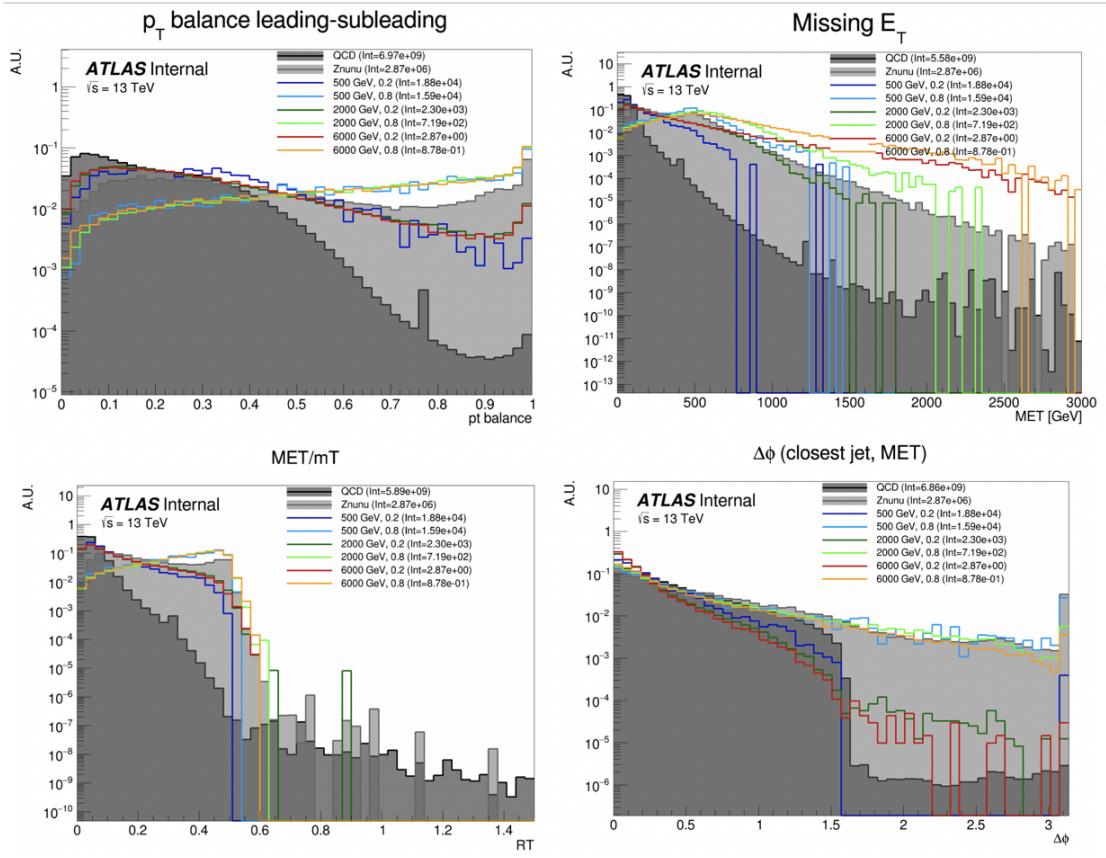


Figure A.4: Analysis variables where high  $R_{inv}$  signals a clearly distinct from background and low  $R_{inv}$  variables. On the contrary, leading jet  $p_T$  is one of the only variables where low  $R_{inv}$  signals are distinct from background.

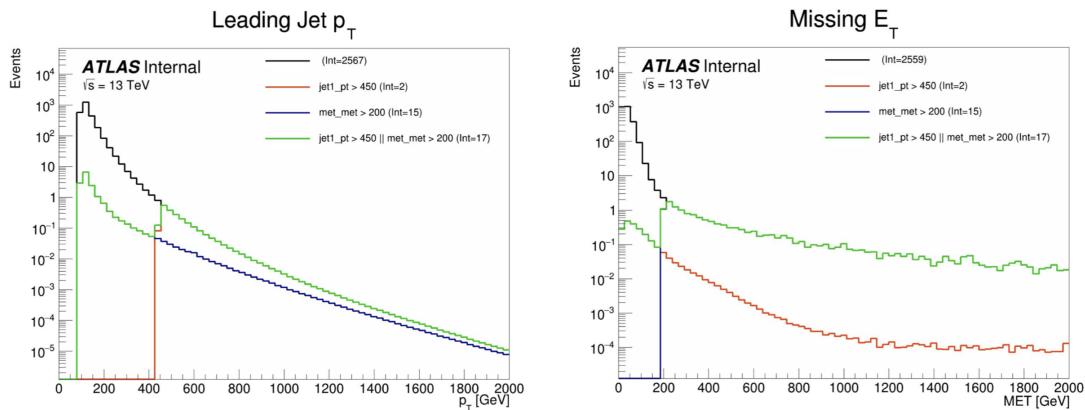


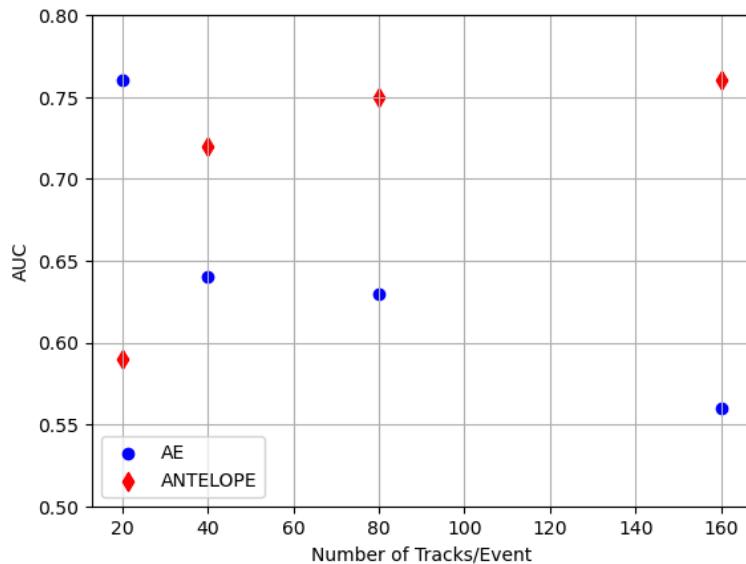
Figure A.5: OR of jet and  $E_T^{\text{miss}}$  triggers.

## Appendix B: Machine Learning Approaches

### B.1 Unsupervised: AE vs. ANTELOPE

To understand the benefits of the semi-supervised ANTELOPE approach, we study the ANTELOPE in comparison to a traditional anomaly detection architecture like an autoencoder. The autoencoder cannot accommodate variable length or permutation invariant inputs.

Figure B.1 shows the AUC determined by each of these two tools as a function of number of tracks. The trend is that the AE suffers when more information is given, due to the presence of 0-padding. In contrast, the ANTELOPE architecture performs better with more information, motivating the use of high dimensional input modelin with this method.



- Confirms that AE is negatively sensitive to zero padding while ANTELOPE sees a performance improvement as it sees more “information”

Figure B.1

### B.1.1 Signal Contamination

To understand the effect of signal contamination in training on the ANTELOPE score, we inject a percent of signal events into the data used to train the ANTELOPE autoencoder stage and look at the AUC on signals. Figure B.2 shows no variation in AUC with 1% contamination in training data, but a few % drop going up to 10%.



Figure B.2

## B.2 PFN Optimality Checks

Studies were done to ensure the relatively optimality of the single PFN model, trained over combined signals, across the grid. As the grid spans signals with a large range of  $E_T^{\text{miss}}$ , their varying input features and background composition may be conducive to separate PFN models trained on high and low  $E_T^{\text{miss}}$  signal points to better capture the signal-background differences. Figure B.3 shows a comparison of the signal-inclusive PFN model performance and the performance of models separated into high and low  $R_{\text{inv}}$  signals in training. The most notable impact is found for the low  $R_{\text{inv}}$  and high mass points, indicating that the signal-inclusive PFN is learning morning about the distinction between high- $E_T^{\text{miss}}$  signals and backgrounds. However, these high mass points are also the most challenging to find due to their very wide resonance on top of  $m_T$ , and in the final projected sensitivity in the  $m_T$  window the differences are < 10% across the grid. To maintain a

harmonized strategy with the ANTELOPE region we keep the inclusive PFN model as the final version.

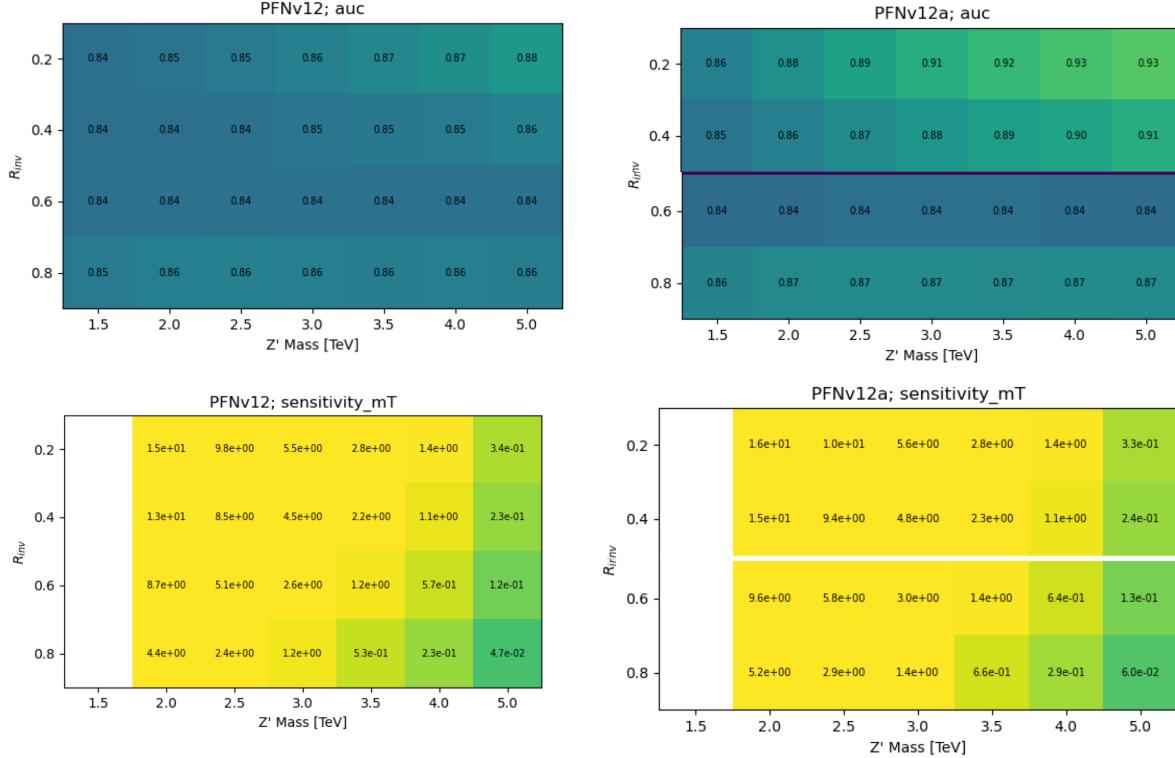


Figure B.3: Comparison of PFN AUC (top) and sensitivity in the  $m_T$  mass window (bottom) for a single PFN model (left) vs. two models trained on  $R_{inv} < 0.5$  and  $> 0.5$  separately (right).

Grid scans for optimality were also performed on the number of training epochs, number of training events, batch size, learning rate, number of neurons, and dimension of the  $\Phi$  space. The results of these scans are summarized in the tables in Figure B.4. The selected or default parameters were found to be optimal, or close enough to optimal to justify not increasing the training time or complexity of the network for negligible increases in performance.

### B.3 Single Jet vs Jet System ML Approach

The analysis considered both a single jet and jet system ML approach. A jet system approach, where the leading two jets and their orientation with respect to each other was selected for a variety of reasons. The jet system approach captures the MET information which is crucial to identifying

	<b>default</b> s_events=500096 b_events=501396	s_events:49135 b_events: 48220 *	s_events:246205 b_events:245136
AUC	.905	.874	.892

	<b>default</b>	n_neuro n 40	n_neuro n 150	phi_dim 32	phi_dim 128	learning _rate 0.0005	learning _rate 0.002	nepochs 50	nepochs 200*
AUC	.905	.898	.906	.902	.906	.898	.905	.893	.909

	<b>default</b>	batchsize_pfn 250	batchsize_pfn 1000
AUC	.905	.903	.902

Figure B.4: Scans done to check for optimality of PFN training parameters.

SVJs. In the topology where the dark quarks come from a heavy Z' decay and are back to back, the measurable MET will have to be aligned with one or the other.

Additionally, the performance of both a supervised PFN approach and an unsupervised AE approach was studied in the case of a single jet tagger. While the PFN approach was still performant on a single jet case, the unsupervised approach was significantly improved by using both jets. This is shown in Figure B.5.

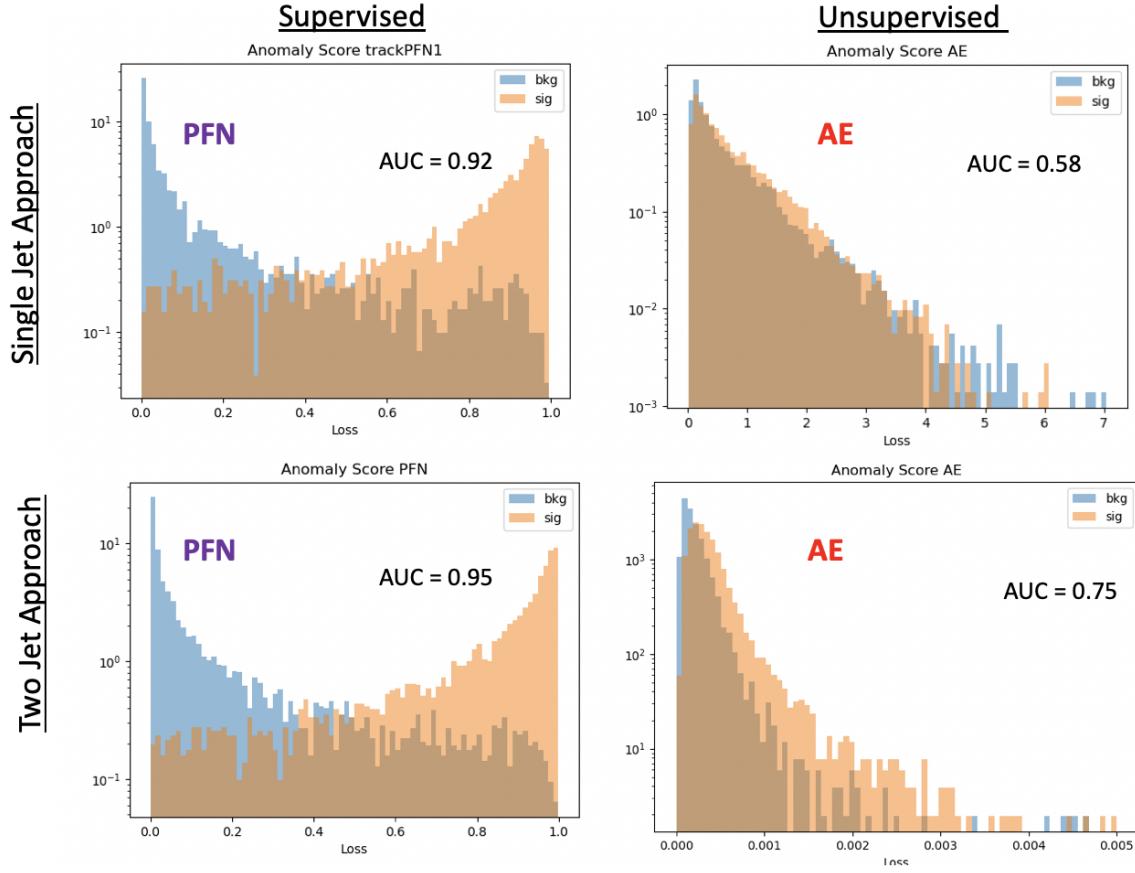


Figure B.5:  $\phi$  Performance comparison between single jet and jet system ML approach

#### B.4 PFN Training Composition

The overall sensitivity and stability across the signal grid is observed to benefit by training the ML tool to reject only the dominant background, QCD. This is evidenced by the PFN response plots shown in Section 7.2 and the following AUC and sensitivity comparison plots in Figure B.6.

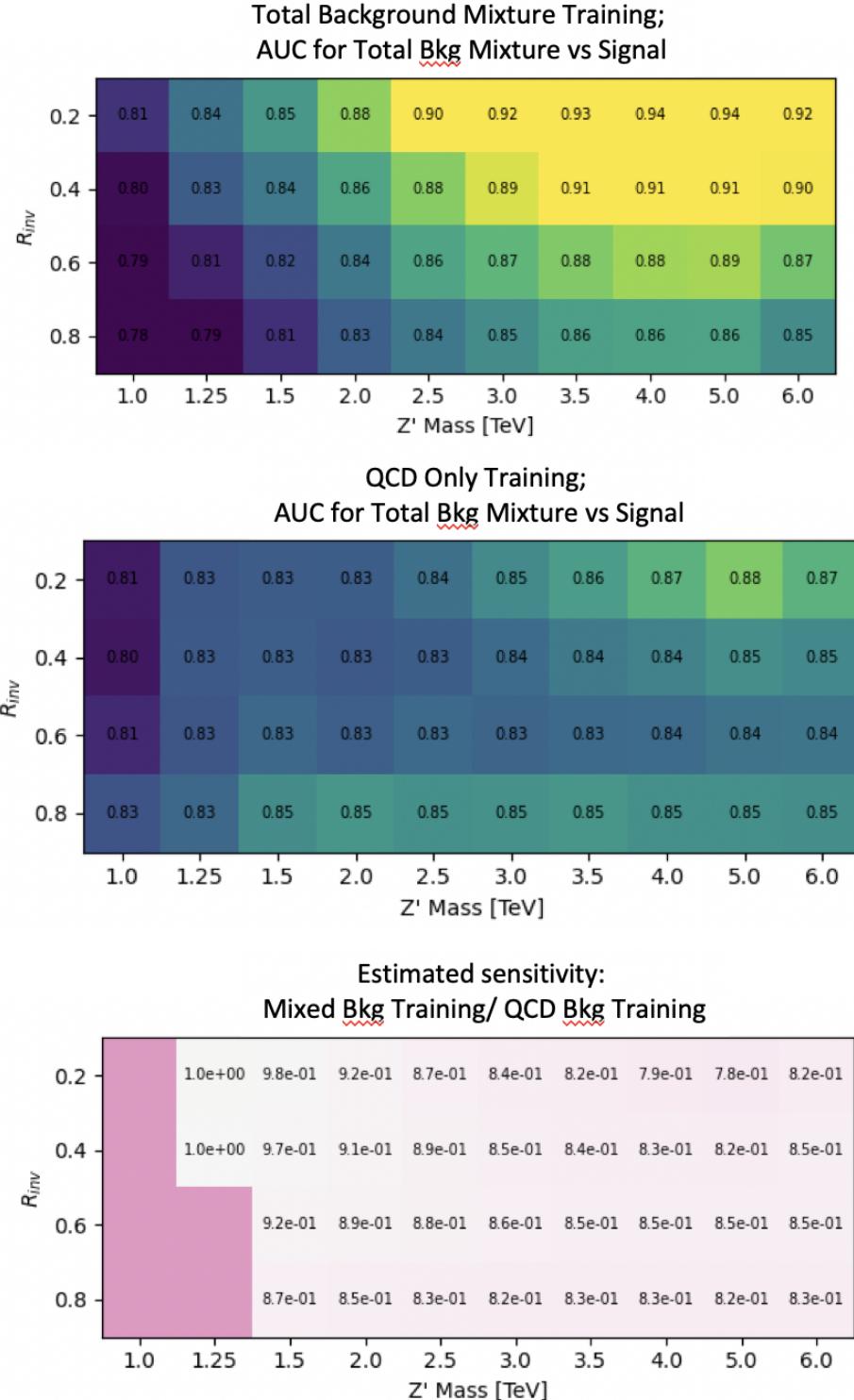


Figure B.6:  $\phi$  Comparison in the AUC score across the grid for the mixed background strategy vs the QCD only strategy. The bottom table highlights that the QCD only strategy gives superior sensitivity across the signal grid.

## Appendix C: $E_T^{\text{miss}}$ and $E_T^{\text{miss}}\phi$ Shapes

### C.1 Tight Cleaning

The Tight cleaning working point was found to be necessary due to the nature of our signal being  $E_T^{\text{miss}}$  and hadronic activity that are closely aligned, presenting a signature that is very affected by beam-induced background (BIB). As per the cleaning recommendations, any event with a jet that fails Tight criteria is rejected. Figure C.1 shows the effect of Tight cleaning on the shape of  $E_T^{\text{miss}}$  in data, fixing a feature present in Loose cleaning only.

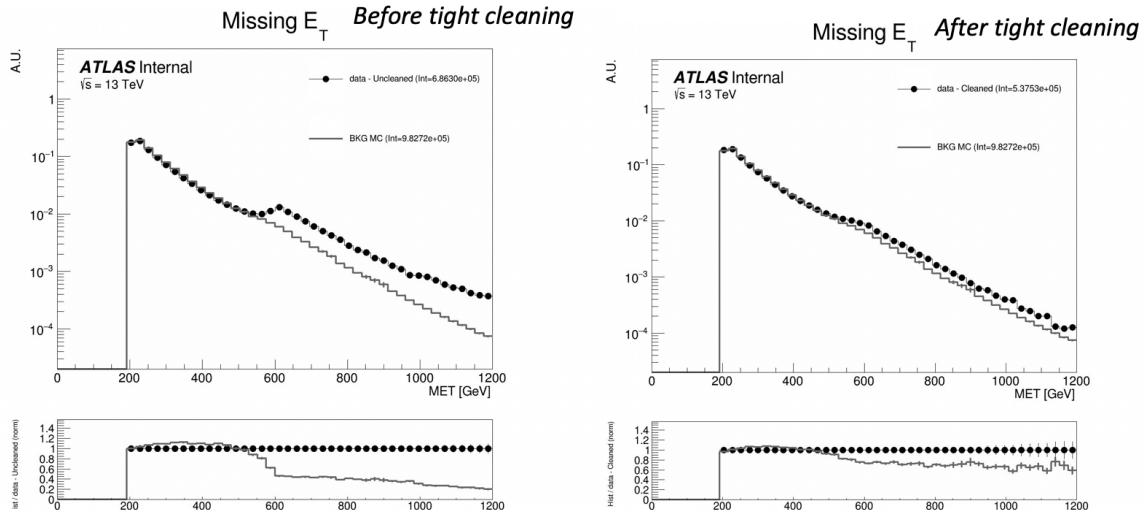


Figure C.1:  $E_T^{\text{miss}}$  in data before and after Tight event cleaning is applied.

Figure C.2 further illustrates the correlation between the excess events and the leading jet  $p_T$ , and illustrates the impact of the tight cleaning in reducing this feature.

Figure C.3 illustrates the 2D  $\eta$  vs  $\phi$  distribution of the leading and subleading jets before and after tight cleaning. No major spikes or hot areas are observed. One bright spot in the subleading jet map at  $\eta \approx 0$  and  $\phi \approx -1.0$  was studied and found to have no impact on the  $E_T^{\text{miss}}$  shape, indicating a likely missed spot in the Tile cleaning which does not affect  $E_T^{\text{miss}}$ .

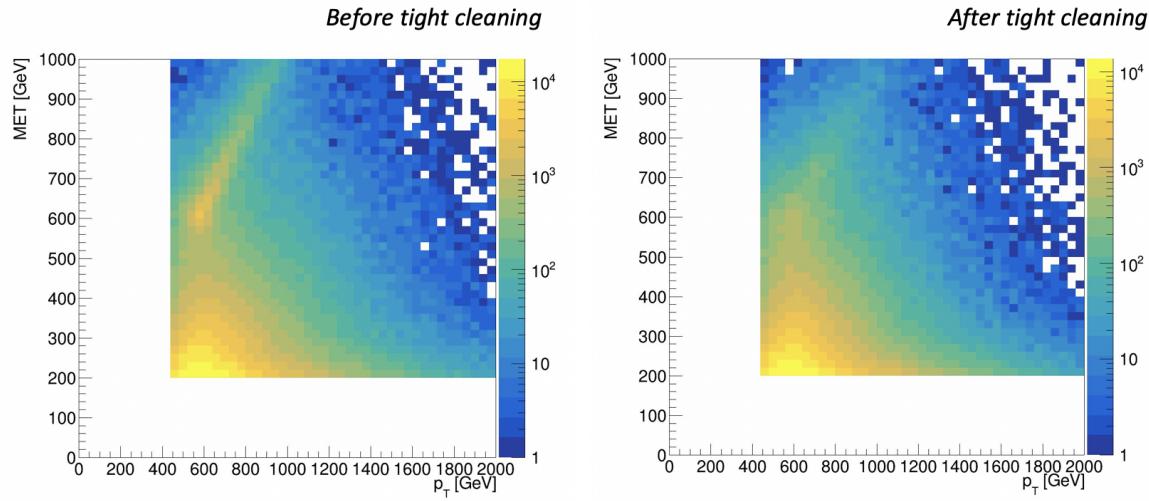


Figure C.2:  $E_T^{\text{miss}}$  vs jet1  $p_T$  in data before and after Tight event cleaning is applied.

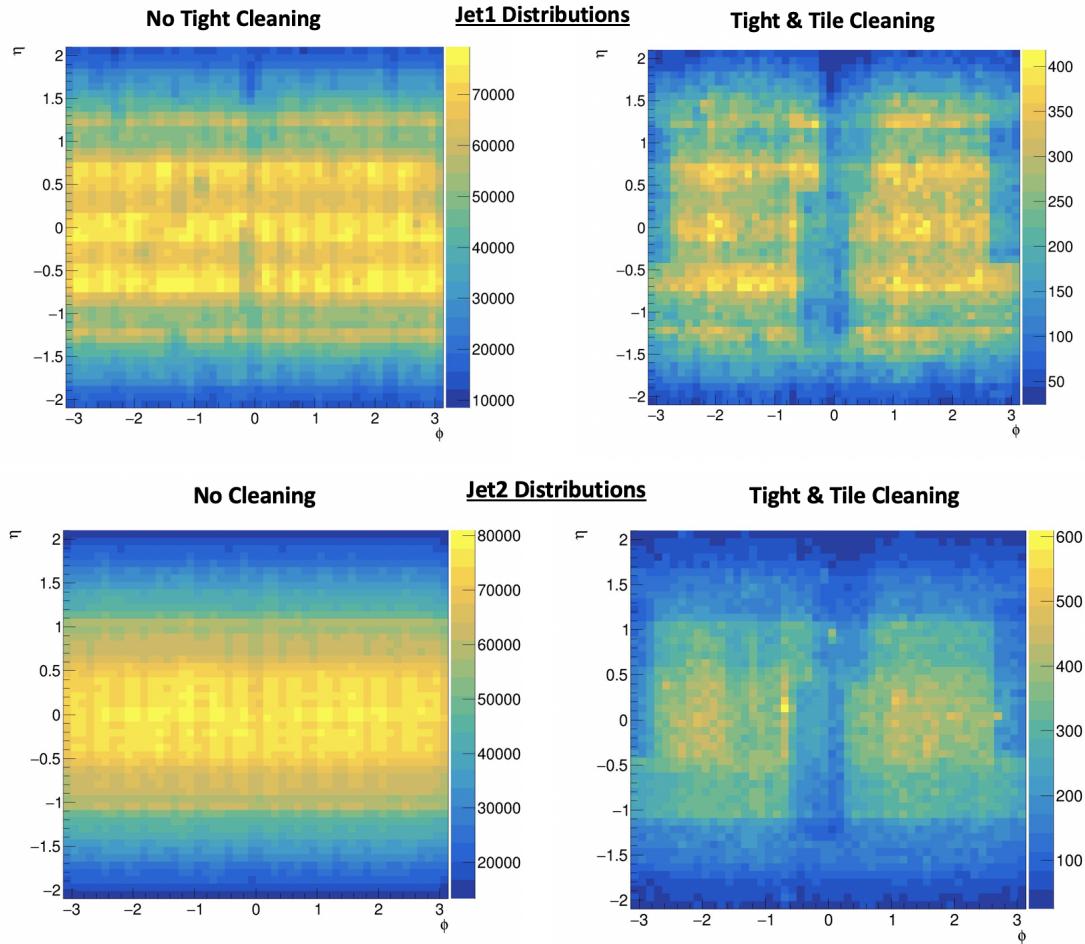


Figure C.3:  $\eta$  vs  $\phi$  for leading and subleading jets, before and after the application of tight cleaning.

## C.2 NCB Preselection

A final preselection was derived to entirely remove the presence of non-collision background, particularly noted through its impact on the  $E_T^{\text{miss}}$  shape. Cuts are added on the subleading jet  $p_T > 150 \text{ GeV}$  and  $\Delta\Phi(j1,j2) > 0.8$ . Figure C.4 shows the impact of these cuts to create a fully smoothly falling  $E_T^{\text{miss}}$  distribution. Figure C.5 shows the impact on the data yields in the CR and VR and several signals in the SR; the greatest inefficiencies are found for the signal points that are not sensitive in the analysis anyway. The reduction in background ultimately means that no impact is noticed on the limits.

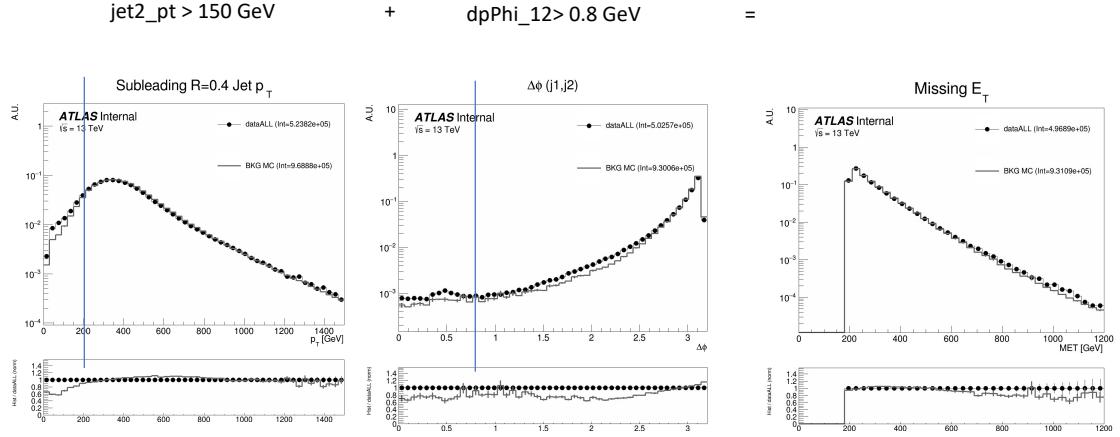


Figure C.4: Added NCB preselection and impact on  $E_T^{\text{miss}}$  shape.

Figure C.6 shows the impact of these cuts on the 2D  $E_T^{\text{miss}}$  vs jet1  $p_T$  distribution, where the feature is also observed to be smoothed.

This strategy was discussed and signed off by the Jet/ $E_T^{\text{miss}}$  CP group<sup>1</sup>.

## C.3 TileCal Correction

The  $E_T^{\text{miss}}\phi$  distribution was fixed through the implementation of an offline TileCal correction tool, which removes certain lumiblocks of certain runs based on poor functioning in TileCal modules. Figure C.7 shows the  $E_T^{\text{miss}}\phi$  distribution in data across runs, before and after the application of the tool, showing the ability of the tool to remove spikes due to instrumental problems.

<sup>1</sup><https://indico.cern.ch/event/1413217/>

Region	Before Extra Cleaning Cuts	After Extra Cleaning Cuts	Efficiency
Data - CR	108957	107435	0.99
Data - VR	116917	116008	0.99
2500, 0.2 - SR	1921	1709	0.89
2500, 0.4 - SR	1819	1424	0.78
2500, 0.6 - SR	1150	735	0.64
2500, 0.8 - SR	543	251	0.46
5000, 0.2 - SR	21.5	20.3	0.94
5000, 0.4 - SR	22.3	19.4	0.87
5000, 0.6 - SR	16.0	11.6	0.73
5000, 0.8 - SR	8.1	4.1	0.51

Figure C.5: NCB preselection impact on data and signal yields.

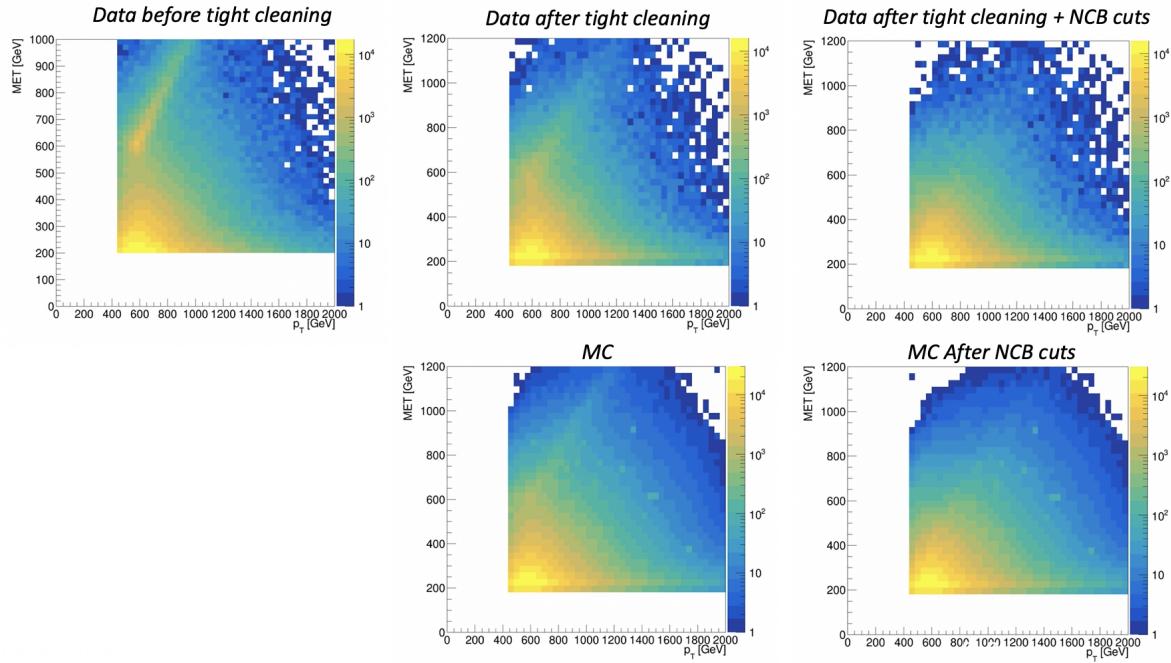


Figure C.6: Impact of tight cleaning and non-collision background preselection.

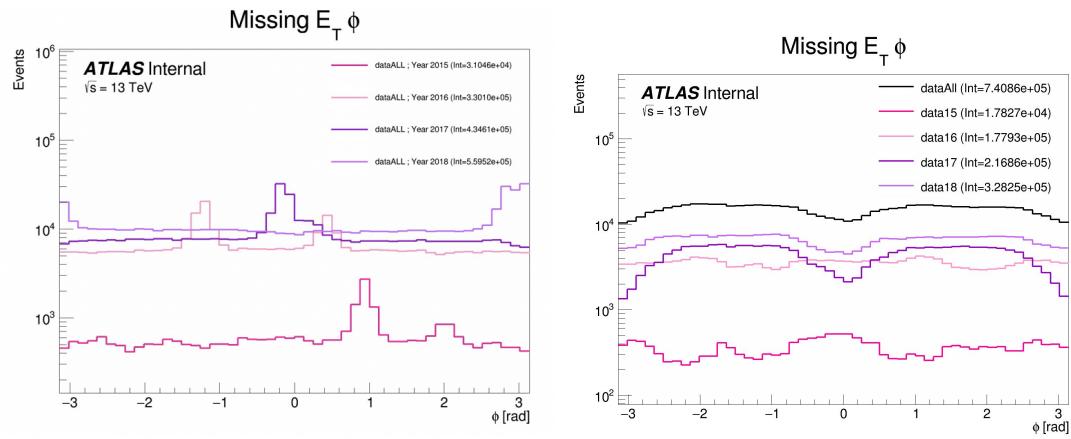


Figure C.7:  $E_T^{\text{miss}} \phi$  in data, before (left) and after (right) application of the TileCal correction tool.

## Appendix D: Truth Studies

### D.1 Jet dR Matching

Figure D.1 demonstrates that the leading and subleading jet are overwhelmingly the most likely jets to be matched to a dark quark.

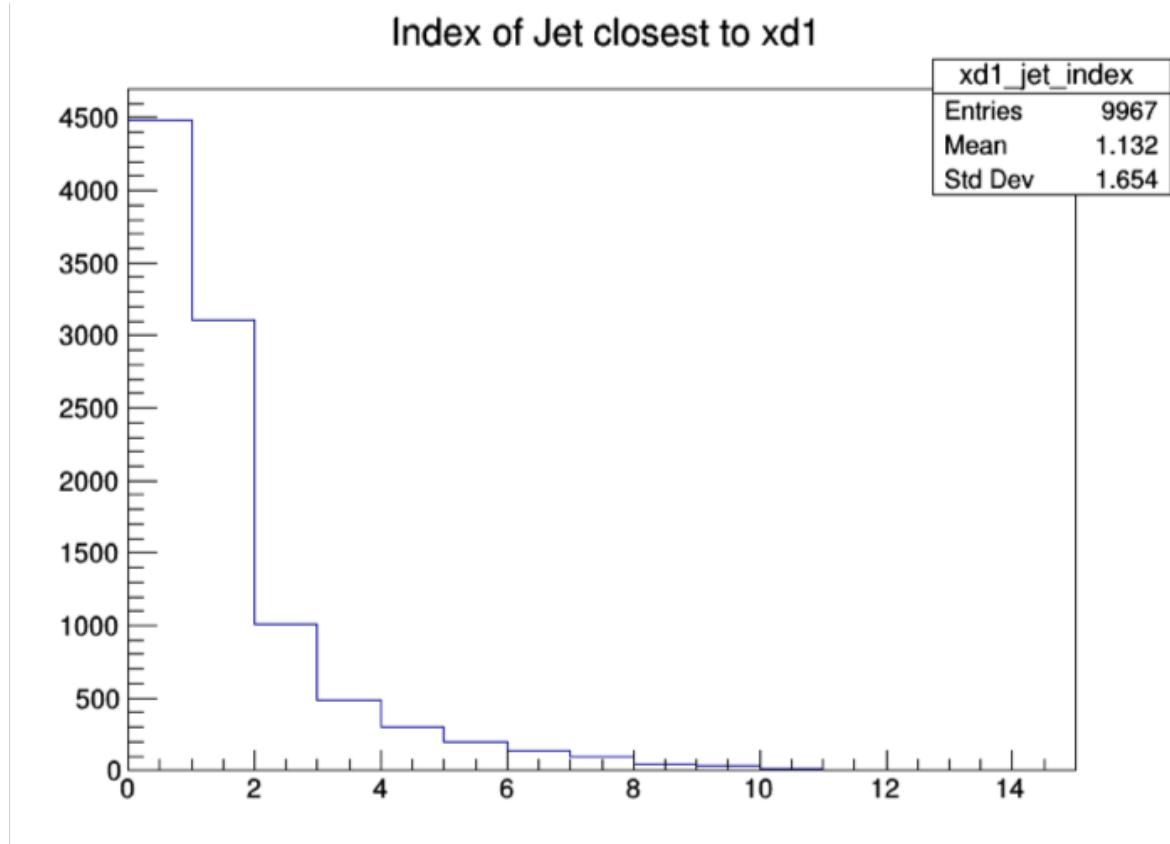


Figure D.1: Index of jets truth matched (by requirement of  $\Delta R < 0.4$ ) with dark quark.

Other matching strategies were explored, such as choosing the jet most aligned with  $E_T^{\text{miss}}(\min \Delta\phi(j, E_T^{\text{miss}}))$  and the jet most anti-aligned with  $E_T^{\text{miss}}(\max \Delta\phi(j, E_T^{\text{miss}}))$ . However, in most cases these two measurements correspond to the subleading and leading jets respectively. Additionally, as shown in Figure D.2 the leading/subleading strategy generally results in more matched jets than

the  $E_T^{\text{miss}}$  aligned and  $E_T^{\text{miss}}$  anti-aligned strategy. Therefore the leading and subleading jets were chosen for consideration in this analysis.

<b>Signal (<math>M_{Z'}</math>, <math>r_{\text{inv}}</math>)</b>	<b>% Leading Jets <math>dR</math> Matched</b>	<b>% MET anti- aligned Jets <math>dR</math> matched</b>	<b>% Subleading Jets <math>dR</math> Matched</b>	<b>% MET aligned jets <math>dR</math> matched</b>
750 GeV, 0.2	<b>0.432883</b>	0.3567	<b>0.3777</b>	0.3504
750 GeV, 0.8	<b>0.187819</b>	0.1589	<b>0.1826</b>	0.1803
3500 GeV, 0.2	<b>0.692931</b>	0.5408	<b>0.5097</b>	0.4144
3500 GeV, 0.8	<b>0.344057</b>	0.2927	<b>0.2634</b>	0.2585
6000 GeV, 0.2	<b>0.647237</b>	0.5047	<b>0.4975</b>	0.3919
6000 GeV, 0.8	<b>0.345542</b>	0.2882	<b>0.2517</b>	0.249

Figure D.2: Percent of jets with  $\Delta R(j, E_T^{\text{miss}}) < 0.4$  comparing two jet identification strategies. Leading and subleading jets are seen to be the better metric for identifying jets associated with the dark quark decay.

## Appendix E: MC Fitting

Figure E.1 shows the ability of this polynomial to fit the smoothly falling  $m_T$  background in simulation across all 3 analysis regions (CR, VR, SR). The  $m_T$  spectrum is fit in 90 bins of width 50 GeV. The high background-only  $p$ -value indicates a good fit. The  $p$ -values are exceptionally high for the MC fits because of the large number of statistics that are compressed to create the MC background spectrum (recall Figure 6.2) leading to a very smooth shape in  $m_T$ .

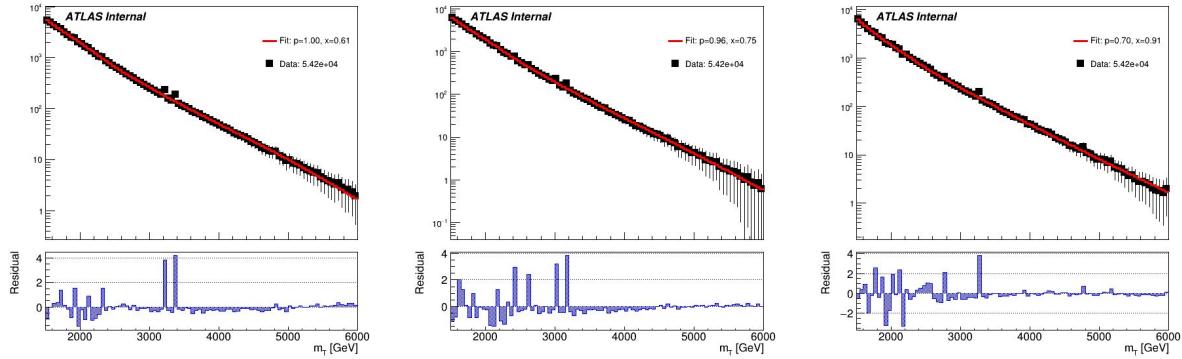


Figure E.1: Background-only  $m_T$  fits using representative MC in the CR (left), VR (middle), and SR (right).

A slight sinusoidal pattern in the residuals may be observed. This arises due to the “stitching” of the  $p_T$  slices for the QCD MC (as shown in Figure 6.2), which is picked up by the fit. For this reason, fitting to MC is only checked to verify that the differences in the slope of  $m_T$  between the three regions (as shown in Figure 8.7) do not pose a problem for the fitting strategy.

## Appendix F: BumpHunter

### F.1 Signal Mass Resolution $m_T$ Binning

In the discovery region, a binning for  $m_T$  is determined that is based on the expected signal width. This is done to improve the BumpHunter performance. The signal mass resolution for a given point is determined with a double-sided Crystal Ball fit to the mass. These fits are performed across  $Z'$  mass, and a linear fit to these values is performed to determine the optimal bin width across  $m_T$ .

The x-axis value used is a data-driven way to determine the appropriate value of  $m_T$  for a given signal, given that the considerable  $E_T^{\text{miss}}$  from the dark particles means that the truth  $Z'$  mass does not well approximate the peak  $m_T$  value. The  $E_T^{\text{miss}}$  in the final state is generally an underestimate of the amount of energy that could be attributed to dark hadrons, as any dark hadrons that are back-to-back in the transverse plane will cancel out and not register as  $E_T^{\text{miss}}$ . Therefore  $m_T$  is always an underestimate of the  $Z'$  mass, so the truth  $Z'$  mass can be used as an upper bound. An integral is then performed backwards from that value until 60% of the total signal yield is included. This window is referred to as the 60% mass window; the mean of this window then provides an approximate localization of the signal mass peak in  $m_T$ . Figure F.1 shows some examples of this algorithm on several signal points of varying  $R_{\text{inv}}$  and mass.

Figure F.2 shows the result of this linear fit for the four  $R_{\text{inv}}$  values considered in the signal grid. As expected, the resolution is considerably different for low and high  $R_{\text{inv}}$  points.

A single  $m_T$  binning for the final SR plotting and BumpHunting is determined by selecting a harmonized binning at low  $m_T$ , and moving to wider bins at high  $m_T$ . As for higher  $R_{\text{inv}}$  signal points the mass resolution linear fit gives negative results, we require each bin to have a width of at least 100 GeV. Figure F.3 shows the resulting bins for each  $R_{\text{inv}}$  category that comes from the

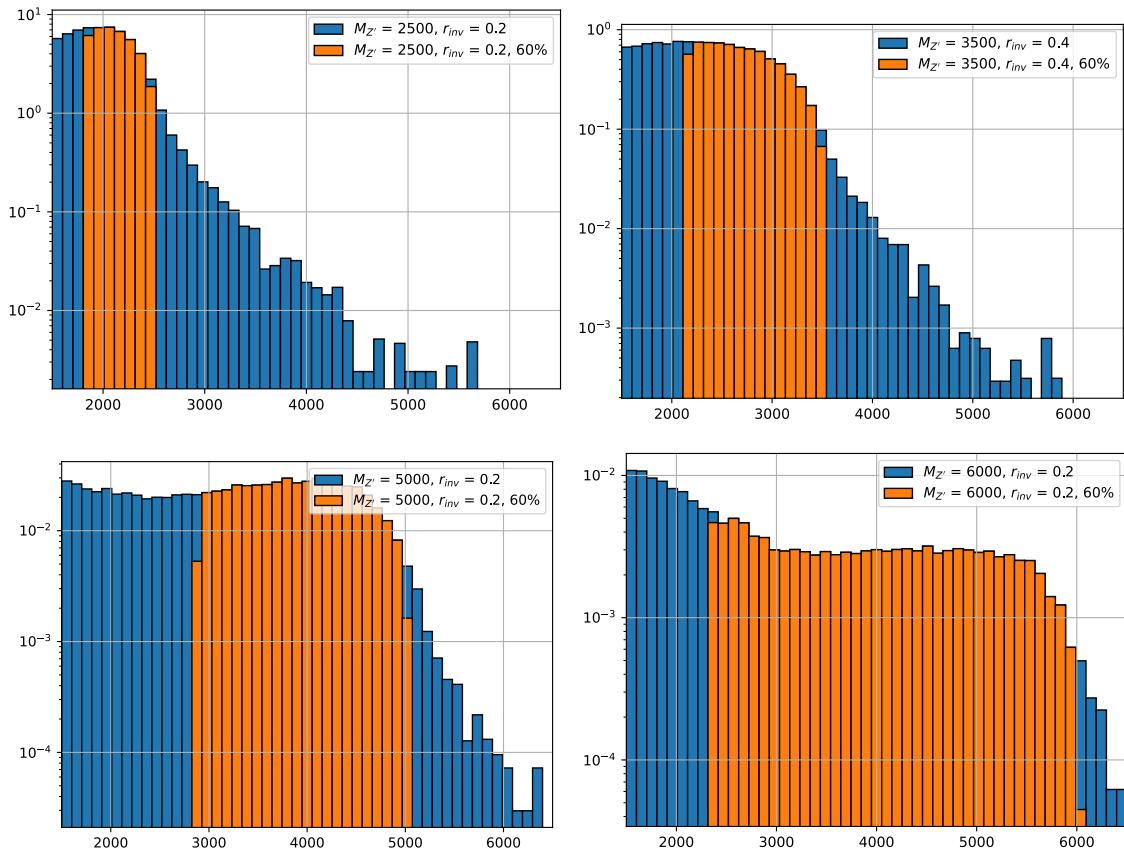


Figure F.1: Example determinations of the 60% mass window means for several signal points.

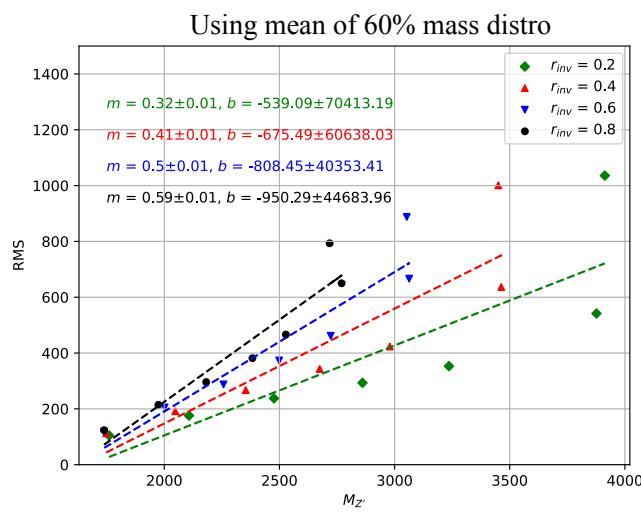


Figure F.2: Signal mass resolution for  $m_T$  binning for the signal grid in  $(R_{inv}, \text{mass})$  space.

mass resolution fits, with the addition of the minimum 100 GeV bin width requirement.

$$r_{inv} = 0.2 \rightarrow [1500, 1600, 1700, 1800, 1904, 2029, 2177, 2356, 2569, 2824, 3128, 3493, 3929, 4451, 5075, 5822, 6715]$$

$$r_{inv} = 0.4 \rightarrow [1500, 1600, 1700, 1800, 1910, 2053, 2238, 2477, 2787, 3187, 3704, 4374, 5240, 6361]$$

$$r_{inv} = 0.6 \rightarrow [1500, 1600, 1700, 1800, 1919, 2087, 2321, 2650, 3110, 3755, 4659, 5925, 7698]$$

$$r_{inv} = 0.8 \rightarrow [1500, 1600, 1700, 1800, 1927, 2117, 2400, 2823, 3455, 4400, 5811, 7918]$$

Figure F.3:  $m_T$  bins based on the signal mass resolution and the minimum 100 GeV width requirement, for each  $R_{inv}$  signal category.

In order to have a final  $m_T$  binning that is not highly model-dependent, we consolidate these four different bins into a single binning which is provided below:

**[1500, 1600, 1700, 1800, 1900, 2025, 2175, 2350, 2575, 2825, 3125, 3500, 3925, 4450, 5075, 6000]**