

1

[ATLAS Semivisible Jets]

2

[Elena Laura Busch]

3

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

4

5

6

7

8

COLUMBIA UNIVERSITY

9

2024

10

© 2024

11

[Elena Laura Busch]

12

All Rights Reserved

13

Abstract

14

[ATLAS Semivisible Jets]

15

[Elena Laura Busch]

16

Abstract of dissertation (place-holder).

Table of Contents

18	Acknowledgments	xi
19	Dedication	xii
20	Introduction or Preface	1
21	I Theory	2
22	Chapter 1: The Standard Model	3
23	1.1 Phenomenology: Particles and Forces	3
24	1.1.1 Particles	3
25	1.1.2 Forces	4
26	1.2 QCD and Jets	7
27	1.3 Symmetries	8
28	1.3.1 Spontaneous Symmetry Breaking and The Higgs Mechanism	9
29	1.4 Experimental Validation of the Standard Model	10
30	1.5 Limitations of the Standard Model	11
31	Chapter 2: Physics Beyond the Standard Model	13
32	2.1 Hidden Valley Models	13
33	2.2 Dark QCD	14

34	2.3 Semi-visible Jets	15
35	II Experiment	17
36	Chapter 3: The Large Hadron Collider	18
37	3.1 Accelerator Physics	19
38	3.1.1 The Journey of a Proton	19
39	3.1.2 Magnets	20
40	3.2 Luminosity	21
41	3.3 LHC Timeline	24
42	Chapter 4: The ATLAS Detector	26
43	4.1 Coordinate System and Geometry	26
44	4.2 Inner Detector	28
45	4.2.1 Pixel Detector	28
46	4.2.2 Semiconductor Tracker	29
47	4.2.3 Transition Radiation Tracker	29
48	4.3 Calorimeters	30
49	4.3.1 Liquid Argon Calorimeter	31
50	4.3.2 Tile Calorimeter	34
51	4.4 Muon Spectrometer	35
52	4.5 Magnet System	37
53	4.6 Forward Detectors	38
54	4.7 Trigger and Data Acquisition	39

55	Chapter 5: Particle Reconstruction and Identification	41
56	5.1 Inner Detector Tracks	41
57	5.2 Photons and Electrons	42
58	5.3 Muons	44
59	5.4 Jets	46
60	5.4.1 Calorimeter Clusters	47
61	5.4.2 Particle Flow Algorithm	48
62	5.4.3 Jet Clustering	49
63	5.4.4 Ghost Track Association	52
64	5.5 Missing Transverse Energy	53
65	III Search	55
66	Chapter 6: Monte Carlo and Data	56
67	6.1 Data	56
68	6.2 Simulation	57
69	6.2.1 Simulated Backgrounds	57
70	6.2.2 Signal Simulation	58
71	Chapter 7: Machine Learning Tools	61
72	7.1 Introduction	61
73	7.1.1 Particle Flow Network (Supervised)	62
74	7.1.2 ANTELOPE (Semi-supervised)	72
75	Chapter 8: Analysis Strategy	80

76	8.1 Event Selection	80
77	Conclusion or Epilogue	81
78	References	85
79	Appendix A: Trigger Studies	90
80	Appendix B: Machine Learning Approaches	94
81	B.1 Unsupervised: AE vs. ANTELOPE	94
82	B.2 PFN Optimality Checks	95
83	B.3 Supervised: BDT vs. PFN	98
84	B.4 Supervised: Variable Correlations	98
85	B.5 Single Jet vs Jet System ML Approach	100
86	B.6 PFN Training Composition	101

List of Figures

88	1.1	Diagram of the 17 particles comprising the Standard Model	4
89	1.2	Fundamental particle interactions of the three fundamental forces described by the	
90		Standard Model [2].	6
91	1.3	An example Feynmann diagram of jet production	7
92	1.4	An illustration of the “hat shaped” potential of the Higgs field, resulting in a non-	
93		zero vacuum expectation value.	9
94	2.1	Illustration of the hidden valley potential.	14
95	2.2	The massive mediator particle Z' of the s-channel realization of a HV model . . .	14
96	3.1	The LHC accelerator complex at CERN [27]	20
97	3.2	The octants of the LHC and location of various beam activities [26]. Stars indicate	
98		the locations of beam collisions, and the associated detectors recording the	
99		outcome of those collisions.	21
100	3.3	(Left) Total integrated luminosity over the course of Run 2. (Right) Average num-	
101		ber of pp interactions per bunch crossing in Run 2. Each curve is weighted by the	
102		integrated luminosity for the year.	23
103	3.4	Timeline of LHC and HL-LHC activities [29]. Integrated luminosity estimates are	
104		approximate, and not reflective of the exact amount delivered to each experiment. .	25
105	4.1	ATLAS coordinate system and geometry	28
106	4.2	A 3D visualization of the structure of the ID in the barrel region [33]	29
107	4.3	ATLAS calorimetery system [34]	30

108	4.4	Diagram of a segment of the EMB, demonstrating the accordion plate arrangement [35]	32
109			
110	4.5	A LAr pulse as produced in the detector (triangle) and after shaping (curve) [35]	33
111			
112	4.6	Readout gap structure in HEC [35]	33
113			
114	4.7	TileCal wedge module [38]	35
115			
116	4.8	Cross section view of the muon spectrometer system [39]	36
117			
118	4.9	Layout of the barrel and endcap toroid magnets [32]	38
119			
120	5.1	Graphic illustrating the various objects and high level features identified by ATLAS object reconstruction, and their interaction with different systems of the ATLAS detector [42]	42
121			
122	5.2	Track reconstruction seeding, finding and fitting illustration [43]	43
123			
124	5.3	Three types of EM object candidates [45].	44
125			
126	5.4	Four types of muon track candidates [47].	46
127			
128	5.5	The fragmentation and hadronization processes undergone by a quark produced in a proton-proton collision [49].	47
129			
130	5.6	A flow chart illustrating the particle flow algorithm progression [54].	50
131			
132	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 [50].	51
133			
134	5.8	A comparison of MC simulation and data for $Z \rightarrow \mu\mu$ events where real $E_T^{\text{miss}} = 0$ [61]. The resolution of the missing energy in the transverse ($x - y$) plane is observed to increase with increasing total $\sum E_T$	54
135			
136	6.1	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 is shown on the right, illustrating the enhanced statistics for the high p_T range.	58
137			

134	6.2	Background processes relevant to the SVJ signal. The agreement between the black line (data) and grey line (all MC processes combined) illustrates that this collection of background processes is sufficient to model the expected E_T^{miss} in the selected data events.	59
138	7.1	The Energy/Particle Flow Network concept, from Ref. [68].	63
139	7.2	An annotated diagram of the PFN architecture. y and ϕ represent geometric in- formation for the input particles, z represents energy information, and PID encom- passes any other particle ID information in the input.	63
142	7.3	A illustration of the expected dijet behavior of semi-visible jets, where one jet is closely aligned with E_T^{miss}	64
144	7.4	Illustration of track coordinates d_0 and z_0	65
145	7.5	Distributions of the track multiplicity in the leading and subleading jets, comparing signal and background PFN training samples.	65
147	7.6	A diagram demonstrating how the two jet system is rotated in (ϕ, η)	66
148	7.7	The 6 PFN track variables in background MC and signal MC. There are some differences between signal and background, but the track kinematics are largely similar.	66
151	7.8	The 6 PFN track variables in data and background MC, after the scaling and ro- tation procedure is applied. There is excellent modeling of the data by the MC within the track variables. The slight discrepancy in the phi distribution is due to the modeling of dead TileCal cells by the QCD MC, which will be discussed in Chapter 8. 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.	67
158	7.9	PFN score for background MC, data, and signal, comparing a PFN training on QCD-only vs all-background MC samples. The average AUC for the QCD-only training (left) is 0.93, while the average AUC for the mixed background training (right) is 0.84. The sensitivity estimate across the grid is better for the QCD-only training - from the distribution we can conclude that this is because the sensitivity to MET enhanced signals is greatly reduced.	69
164	7.10	PFN architecture loss during training as a function of epoch (left) and the evaluated loss over the signal and background (right).	69

166	7.11 ROC the PFN score for combined signal (true positive) and QCD background (false positive).	70
167		
168	7.12 AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD-only training sample.	71
169		
170	7.13 PFN score for two signals and the total background MC (top), and between data and MC (bottom). The difference between data and MC efficiency is minimal (< 5%).	71
171		
172		
173	7.14 A visual representation of the 64 PFN latent space variables which create the input of the VAE component of ANTELOPE. The left shows a 2D histogram of the PFN latent space index (0-63) versus the value assumed by that index. The right shows 1D histograms of two particular PFN latent space variables.	73
174		
175		
176		
177	7.15 An annotated diagram of the ANTELOPE architecture.	74
178		
179	7.16 ANTELOPE architecture loss during training as a function of epoch.	75
180		
181		
182	7.17 ANTELOPE score distribution comparing data and the total background MC (left), with good agreement observed between data and simulated background, and comparing all background MC to signals (right), revealing good discrimination power.	76
183		
184	7.18 AUC from the ANTELOPE score for each signal in the SVJ grid.	77
185		
186		
187		
188	7.19 Comparing data and the alternate signal models for the PFN score (top left), ANTELOPE score (top right), E_T^{miss} (bottom left), and m_T (bottom right). The emerging jet signal is an example of the gain of the model-independent ANTELOPE approach, where it has a bimodal shape in PFN score but is clearly tagged as anomalous by ANTELOPE.	78
189		
190		
191		
192		
193	7.20 Comparing data and the alternate signal models in terms of sensitivity (S/\sqrt{B}) for the PFN and ANTELOPE tools, applying the selection that is used in the analysis. The ANTELOPE network is found to provide significant added sensitivity to alternate signals such as the gluino \rightarrow R-hadron and emerging jets, which have higher E_T^{miss} than the SVJs.	79
194		
195		
196		
197	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.	90
198		
199	A.2 The factor of improvement in S/\sqrt{B} for each trigger method compared to the untriggered case.	91
200		

199	A.3 The ratio of S/\sqrt{B} of jet trigger over E_T^{miss} trigger selection.	91
200	A.4 Analysis variables where high R_{inv} signals a clearly distinct from background and	
201	low R_{inv} variables. On the contrary, leading jet p_T is one of the only variables	
202	where low R_{inv} signals are distinct from background.	93
203	A.5 OR of jet and E_T^{miss} triggers.	93
204	B.1	94
205	B.2	95
206	B.3 AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD-	
207	only training sample (top) and the total MC background (bottom). Note the three	
208	missing points will be added shortly - they were delayed due to a DAOD production	
209	mistake.	96
210	B.4 Comparison of PFN AUC (top), SIC (middle), and sensitivity in the m_T mass win-	
211	dow (bottom) for a single PFN model (left) vs. two models, trained on $R_{inv} < 0.5$	
212	and > 0.5 separately.	97
213	B.5 Preferred cuts on the PFN score for each point in the grid, comparing the effect of	
214	adding the NCB preselection.	98
215	B.6 Scans done to check for optimality of PFN training parameters.	99
216	B.7	99
217	B.8 ϕ orientation variables in the CR and VR	100
218	B.9 Correlation between ϕ orientation variables and PFN score	101
219	B.10 ϕ Performance comparison between single jet and jet system ML approach	102
220	B.11 ϕ Comparison in the AUC score across the grid for the mixed background strategy	
221	vs the QCD only strategy. The bottom table highlights that the QCD only strategy	
222	gives superior sensitivity across the signal grid.	103

List of Tables

224	6.1 Fixed parameters in the Pythia8 HV model	60
225	6.2 Values for m_{dark}	60
226	6.3 Mass points and cross sections of the SVJ search signal grid	60

227

Acknowledgements

228 Insert your acknowledgements text here. This page is optional, you may delete it if not
229 needed.

230

Dedication

231

Dedicated to my friends and family

232

Introduction or Preface

233 Insert your preface text here if applicable. This page is optional, you may delete it if not
234 needed. If you delete this page make sure to move page counter comment in thesis.tex to correct
235 location.

236

Part I

237

Theory

Chapter 1: The Standard Model

240 The Standard Model of particle physics is a universally accepted framework which explains
 241 the interactions of fundamental particles. All known fundamental particles, outlined in Figure
 242 1.1, are represented in the Standard Model. The model describes three of the four known forces:
 243 the electromagnetic force, the weak force, and the strong force. Gravity, the fourth fundamental
 244 force, is not addressed by the Standard Model. The Standard Model was primarily developed over
 245 the course of the 1960s and 1970s, by combining the work of many physicists into one coherent
 246 model. The Standard Model has been established as a well-tested theory by decades of experimen-
 247 tal physics research.

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

252 **1.1 Phenomenology: Particles and Forces**

253 **1.1.1 Particles**

254 A classic representation of the particles comprising the Standard Model is shown in Figure
 255 1.1. The two primary particles classes are bosons (gauge bosons and the scalar Higgs boson) and
 256 fermions (leptons and quarks). The bosons are carriers of fundamental forces, while the fermions
 257 are the building blocks of matter. Fermions are sorted into three *generations*, and each fermion is
 258 identified by a unique *flavor*.

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

Standard Model of Elementary Particles

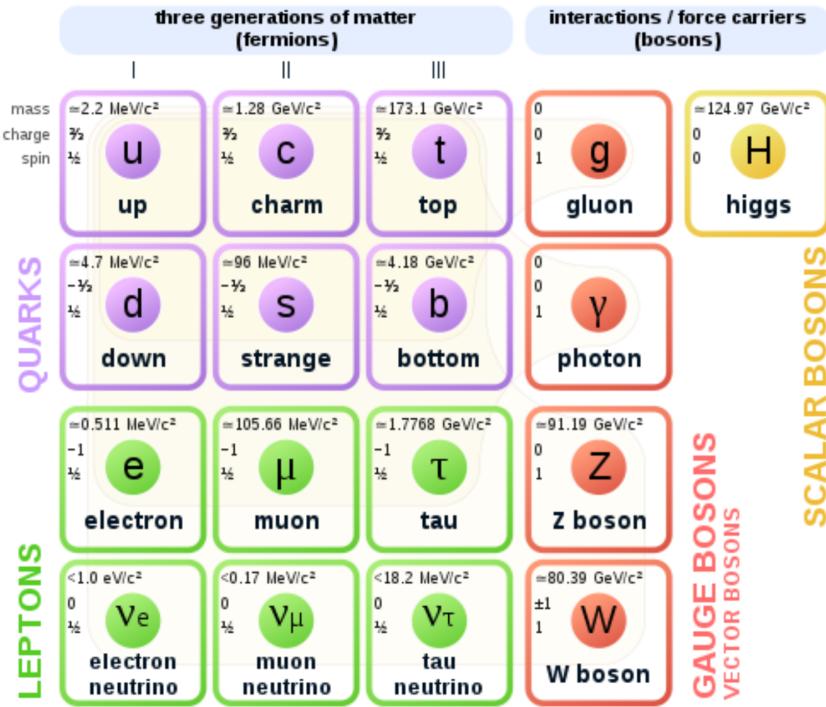


Figure 1.1: Diagram of the 17 particles comprising the Standard Model

261 tion, with the exception of photons and gluons which are known to be massless. Charge refers to
 262 the electromagnetic charge in the case of leptons and W bosons, and to color charge in the case
 263 of quarks and gluons. Spin is an intrinsic form of angular momentum carried by fundamental
 264 particles; all fermions have half integer spin, while bosons have integer spin.

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

270 1.1.2 Forces

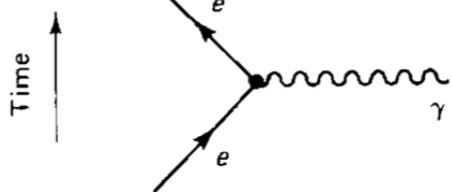
271 The three fundamental forces explained by the Standard Model are the electromagnetic force,
 272 the strong force, and the weak force. The photon is the carrier of the electromagnetic force, which

273 dictates the nature of interactions between electrically charged particles, and is widely covered by
274 introductory physics courses. The electromagnetic force has an infinite interaction range, a result
275 of the massless and non-self interaction nature of the photon. The electromagnetic interaction is
276 described by the theory of quantum electrodynamics (QED).

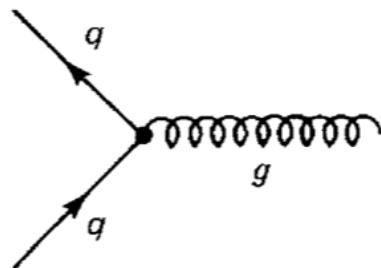
277 The weak force gives rise to atomic radiation and decay. It allows for the processes of beta
278 decay, which enables conversion between neutrons and protons within the nucleus of an atom. In
279 the process of beta decay, a proton decays into a neutron, a positron, and a neutrino; or, a neutron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

286 The strong force confines quarks into hadron particles, such as protons and neutrons. The
287 strong force also allows for the creation of atomic nuclei by binding protons and neutrons together,
288 and is generally referred to as the “nuclear force” in this context. The gluon is the mediator of
289 the strong force, which is a short-range force which acts at subatomic distances on the order of
290 10^{-15} m. At this range, the strong force is about 100x as strong as the electromagnetic force,
291 which allows for the creation of positively charged nuclei [2]. The strong force is described by the
292 theory of quantum chromodynamics (QCD). In the same way that QED dictates the interaction of
293 electrically charges particles, QCD dictates the interactions of *color-charged* particles. Due to the
294 particular importance of QCD in this thesis, this topic will be explored in detail in section 1.2.

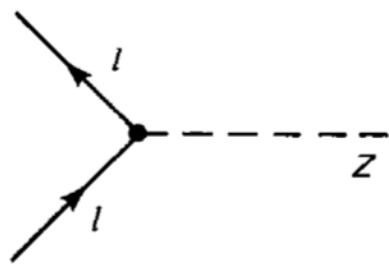
295 The fundamental Feynmann diagram for each of the three forces discussed here is depicted
296 in Figure 1.2. The fourth fundamental force, gravity, is not currently explained by any known
297 mechanism within the Standard Model.



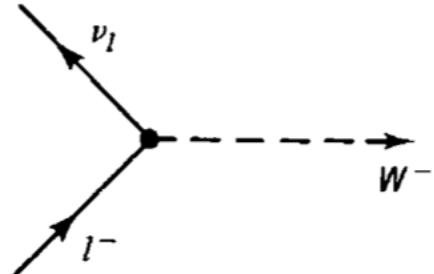
(a) The electromagnetic force



(b) The strong force



(c) The neutral weak force



(d) The charged weak force

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

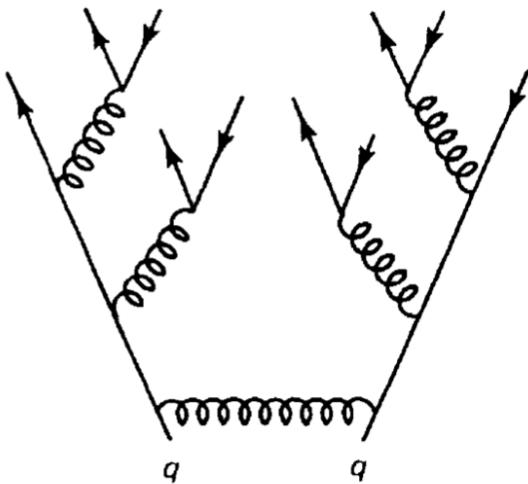


Figure 1.3: An example Feynmann diagram of jet production

298 1.2 QCD and Jets

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

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

311 **1.3 Symmetries**

312 The Standard Model is a renormalizable quantum field theory that obeys the local symmetry

313 G_{SM} :

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

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

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

329 The SM Lagrangian is invariant under CPT symmetry, or charge, parity, and time reversal.
330 Charge conjugation (C) transform a particle into it's corresponding antiparticle by reversing the
331 charge and other quantum numbers. Parity conjugation (P) reverses spatial coordinates, which
332 transforms left-handed particles into right-handed particles and vice-versa. Time reversal (T) is
333 the theoretical process of reversing time. The L subscript in the $SU(2)_L$ group indicates that this
334 symmetry only applies to left-handed fermions. As a result, the $W^{1,2,3}$ gauge bosons of $SU(2)_L$

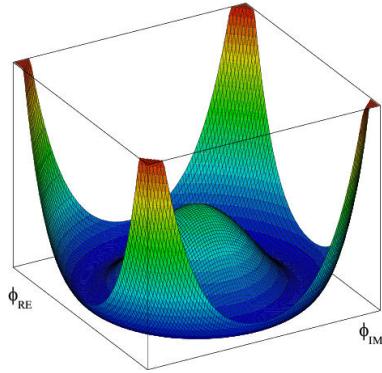


Figure 1.4: An illustration of the “hat shaped” potential of the Higgs field, resulting in a non-zero vacuum expectation value.

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]. 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 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 γ is created from the components of the

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

357 **1.4 Experimental Validation of the Standard Model**

358 The theoretical framework of the Standard Model coalesced into a unified theory in the mid-
359 20th century. A cascade of discoveries providing empirical evidence for the model followed
360 closely. In the 1960s, three quarks (up, down and strange) and four leptons (electron, muon,
361 and their associated neutrinos) were the known particulate building blocks of matter and the Stan-
362 dard Model. The discovery of the charm quark in 1974, through the observation of the J/ψ meson
363 [6][7], confirmed the existence of a fourth quark flavor. The discovery of the τ in 1975 [8] provided
364 the first evidence of a 3rd generation of matter. This was quickly followed by the observation of
365 the Υ meson in 1977 [9], which provided evidence for the existence of a fifth quark, the b quark
366 (bottom, or beauty). The existence of a 3rd generation of fermion also explained the observation
367 of CP violation in the weak force, as it allowed for the addition of a complex phase in the CKM
368 matrix (a unitary matrix which describes flavor mixing in the weak interaction). The top quark
369 (t) and tau neutrino (ν_τ) were predicted at this point as the final building blocks of three complete
370 generations of fermions, and they were discovered by experimental observation around the turn of
371 the 21st century [10] [11] [12].

372 The W and Z bosons were predicted by the Standard Model, but to observe them required the
373 construction of a particle accelerator powerful enough to produce them. They were finally observed
374 at CERN in 1983 by the UA1 and UA2 experiments [13] [14] at the newly constructed Super Proton
375 Synchrotron (SPS). Their masses were observed to be compatible with the masses predicted by the
376 Standard Model nearly a decade earlier. The final missing piece then was confirming the existence
377 of the Higgs, which again required the construction of a newer and more powerful collider. CERN
378 achieved this with the construction of the Large Hadron Collider (LHC), and in 2012 the ATLAS

379 and CMS experiments announced the discovery of the Higgs particle [15] [16].

380 **1.5 Limitations of the Standard Model**

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

- 384 • Gravity - the Standard Model does not account for the fourth fundamental force of gravity.
- 385 • Dark Matter - there is no viable candidate to explain the existence of dark matter, a non-
386 interacting form of matter which must exist to account for gravitational observations which
387 cannot be explained by general relativity, such as the motion of galaxies, gravitational lens-
388 ing, and the structure of the universe [17].
- 389 • Matter-Antimatter asymmetry - the level of CP violation in the Standard Model isn't suf-
390 ficient to explain the large discrepancy between the amount of matter and the amount of
391 antimatter in the universe today, and the origins of this imbalance are not understood.
- 392 • Neutrino masses - the Standard Model assumes that neutrinos are massless and provides
393 no mechanism for them to acquire mass. However, observations of neutrino oscillations
394 indicates they posses some small non-zero mass [18].

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

- 400 • Strong CP - while CP symmetry is violated in the weak force, observations indicate that it
401 is preserved by the strong force [19]. The Standard Model predicts that CP violation in the

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

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

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

Chapter 2: Physics Beyond the Standard Model

417 In light of the various phenomenon unexplained by the Standard Model, physicists have pro-
 418 posed various extensions to the Standard Model, collectively termed *Beyond the Standard Model*
 419 (BSM) theories. A particular focus of the physic programs at the Large Hadron Collider (LHC) are
 420 BSM models which suggest dark matter candidate particles. If these particles couple to Standard
 421 Model, they could be produced and observed at the LHC.

422 **2.1 Hidden Valley Models**

423 Hidden Valley (HV) models are a category of BSM models that allow for dark matter (DM)
 424 production at the LHC. They extend the Standard Model with an additional non-Abelian gauge
 425 group [20]. This introduces the possibility of a complex dark sector, which mirrors the complexities
 426 of Standard Model QCD, and introduces the possibility of dark quarks and gluons. The term
 427 “hidden valley” refers to the idea that the DM is hidden from the SM by a high-energy barrier, as
 428 illustrated in Figure 2.1. The dark sector is assumed to communicate with the Standard Model via
 429 a “portal”, or “messenger particle”, that can interact with both Standard Model and HV forces. For
 430 the s-channel scenario, the portal is considered to be a new massive mediator particle Z' .

431 The portal particle allows for the production of dark sector particles at hadron colliders. If
 432 dark quarks are produced via the decay $Z' \rightarrow q_D q_D$ they can hadronize and form dark jets. The
 433 properties of the dark jets are determined by the dynamics of the dark sector, which are explored in
 434 the subsequent section. Depending on the details of the model, the jets formed by the dark hadrons
 435 can be categorized as fully dark, semi-visible, leptonic, emerging, or other [20].

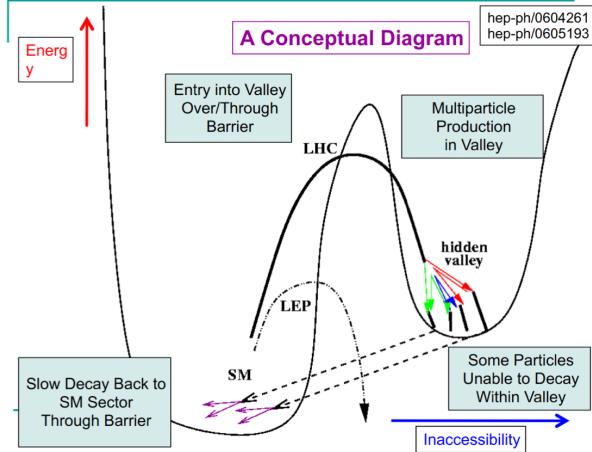


Figure 2.1: Illustration of the hidden valley potential.

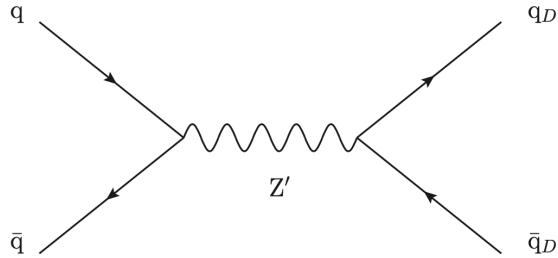


Figure 2.2: The massive mediator particle Z' of the s-channel realization of a HV model

436 2.2 Dark QCD

437 The theoretical underpinning of the semi-visible jet phenomenology is a dark sector with a
 438 gauge group $SU(N_d)$ leading to confinement at a scale Λ_d . For illustration, let's consider the
 439 case of an $SU(2)_d$ gauge theory, which gives rise to two dark fermionic generations $\chi_a = \chi_1, \chi_2$.
 440 Following the work of Timothy Cohen, et.al. we can write the fundamental dark Lagrangian as:

$$\mathcal{L}_{dark} \supset -\frac{1}{2} \text{Tr } G_{\mu\nu}^d G^{d\mu\nu} - \bar{\chi}_a (i \not{D} - M_{d,a}) \chi_a \quad (2.1)$$

441 The first term allows for the dark gluons to self-interact, while the second term enables the dark
 442 quarks to hadronize and acquire mass. The dark quarks are assumed to have a common mass M_d .
 443 The coupling strength of the strongly interacting dark quarks is termed α_d . At the confinement

444 scale Λ_d , the dark quarks can form bound states. At the scale $M_d \approx \Lambda_d$ a QCD-like show occurs.

445 The properties of the hadrons formed by the dark quarks are of particular importance to the
446 observed dark QCD dynamics. Dark-isospin number $U(1)_{1-2}$ and dark-baryon number $U(1)_{1+2}$
447 are accidental symmetries of the theory which determine the stability of the hadrons. In the case
448 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
449 two baryons ($\bar{\chi}_1\bar{\chi}_2$, $\bar{\chi}_1\bar{\chi}_2$). The mesons $\chi_1\bar{\chi}_2$ and $\bar{\chi}_1\chi_2$ are charged under dark-isospin and will be
450 stable if this symmetry is unbroken. The baryons would also be stable as they are charged under
451 the dark-baryon number. These four stable hadrons become dark matter candidates of the theory.
452 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.
453 The unstable mesons can decay into stable dark mesons, or into an off-shell Z' . The off-shell Z'
454 will then decay into two DM quarks or two SM quarks, and its products will continue to shower
455 until the final state particles are stable.

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

461 2.3 Semi-visible Jets

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

468 While there are a myriad of HV and dark QCD models, a handful of model parameters are most
469 important in determining the observable of these showers within a particle detector. The coupling

470 strength α_d is one of the most important, as it controls the fraction of dark hadrons emitted in the
471 shower and their average p_T . The mass of the dark quarks directly impacts the jet mass. If the
472 masses of the dark quark flavors are comparable, the ratio of stable to unstable dark hadrons will
473 be approximately 1:1. However, if there is a mass splitting, stable or unstable dark hadrons may
474 be favored, which impacts the amount of missing energy observed.

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

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

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

484

Part II

485

Experiment

Chapter 3: The Large Hadron Collider

488 The Large Hadron Collider (LHC) is a 26.7 km circular high-energy particle accelerator, span-
 489 ning the Swiss-French border near the city of Geneva, Switzerland [21]. The LHC occupies the
 490 tunnel constructed in 1989 for the Large Electron-Positron (LEP) Collider, and reaches a maxi-
 491 mum depth of 170m below the surface. The LHC is operated by the European Organization for
 492 Nuclear Research (CERN), the largest international scientific collaboration in the world.

493 The LHC accelerates protons and heavy ions, and collides them at four interaction points
 494 around the ring, with a design center-of-mass energy per collision of $\sqrt{s} = 14$ TeV. Each interaction
 495 point is home to one of four detector experiments, which study the products of the collisions. The
 496 largest of these experiments is the ATLAS detector, a general purpose detector designed to study
 497 the Standard Model and search for new physics that could be produced in LHC collisions [22].
 498 The CMS detector is another general purpose detector, designed and operated independently of the
 499 ATLAS detector, but intended to probe the same range of physics [23]. The ALICE experiment is
 500 a dedicated heavy ion experiment, and the LHC-b experiment is a dedicated *b*-physics experiment
 501 [24] [25].

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

506 **3.1 Accelerator Physics**

507 **3.1.1 The Journey of a Proton**

508 From 2010 - 2018, the protons which fed the LHC started as hydrogen gas. The electrons were
509 removed from the hydrogen atoms through the use of strong electric fields. The linear accelerator
510 LINAC2 then accelerated the protons to an energy of 50 MeV. Between 2018 and 2020, LINAC2
511 was replaced with LINAC4, which instead accelerates H^- ions, hydrogen atoms with two electrons.
512 LINAC4 is capable of accelerating the H^- ions to 160 MeV. Before injection to the next part of
513 the acceleration chain, both electrons are stripped from the H^- ions, leaving just protons. From
514 here the protons enter the Proton Synchrotron booster, where they are accelerated up to 1.4 GeV of
515 energy. Subsequently they are sorted into bunches separated in time by 25 ns, where each bunch
516 contains approximately 10^{11} protons. Next the bunches pass through the Proton Synchrotron (PS)
517 and the Super Proton Synchrotron (SPS), where they reach energies of 25 GeV and 450 GeV
518 respectively. Finally they are injected into the LHC as two beams traveling in opposite direction.
519 The original design allowed each beam to be accelerated up to 7 TeV of energy. Due to limitations
520 in the performance of the superconducting LHC magnets, the highest energy actually achieved by
521 the LHC beams during Run 2 was 6.5 TeV, giving a collision center-of-mass energy of $\sqrt{s} = 13$
522 TeV [26]. Figure 3.1 shows the full LHC accelerator complex.

523 Acceleration in the LHC is performed by eight radio frequency (RF) cavities located around the
524 ring. Each RF cavity produces a 2 MV electric field oscillating at 40 MHz. The 40MHz oscillation
525 produces a point of stable equilibrium every 2.5 ns. These points of equilibrium are synchronized
526 with the occurrence of the proton bunches produced in the PS – a proton bunch occupies one out
527 of every ten points of stable equilibrium, such that the bunches maintain a 25 ns spacing [26].

528

The CERN accelerator complex Complexe des accélérateurs du CERN

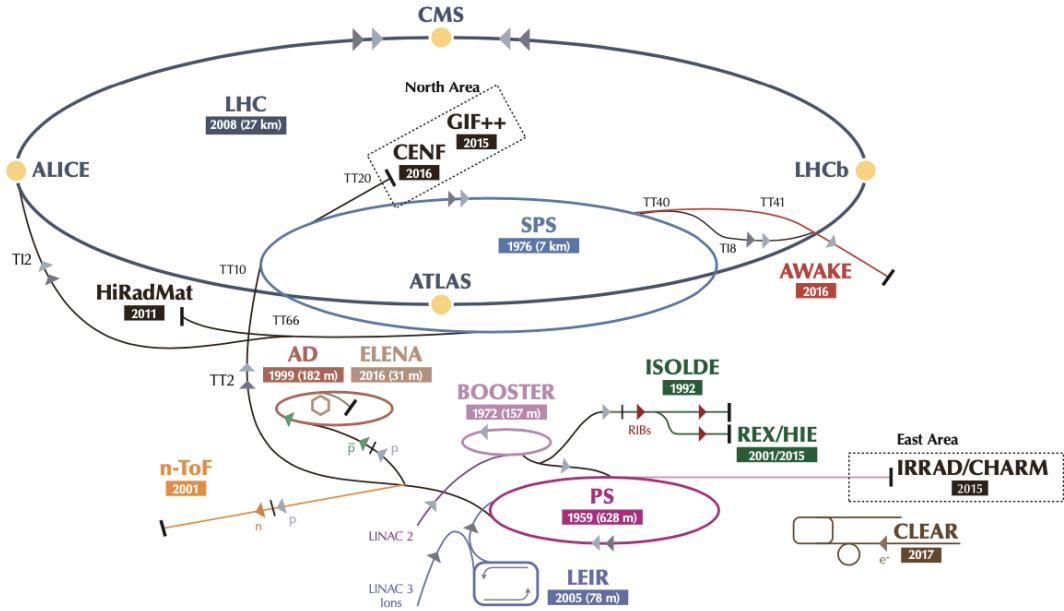


Figure 3.1: The LHC accelerator complex at CERN [27]

529 3.1.2 Magnets

530 In addition to the acceleration cavities, the LHC houses 9593 superconducting magnets which
 531 direct and focus the proton beam on its 27 kilometer journey. The magnets are comprised of super-
 532 conducting Niobium-Titanium coils cooled to 1.9K by superfluid helium. As the beams approach
 533 one of the four collision points around the ring, multipole magnets focus and squeeze the beam for
 534 optimal collisions [26].

535 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
 536 to follow the roughly circular 27 km path. The main dipoles generate powerful 8.3 tesla magnetic
 537 fields to achieve this bend. Each dipole magnet is 15 meters long and weighs 35 tonnes. The
 538 dipoles work in conjunction with quadrupole magnets, which keep the particles in a focused beam,
 539 and smaller sextupole, octupole and decapole magnets which tune the magnetic field at the ends of
 540 the dipole magnets [28].

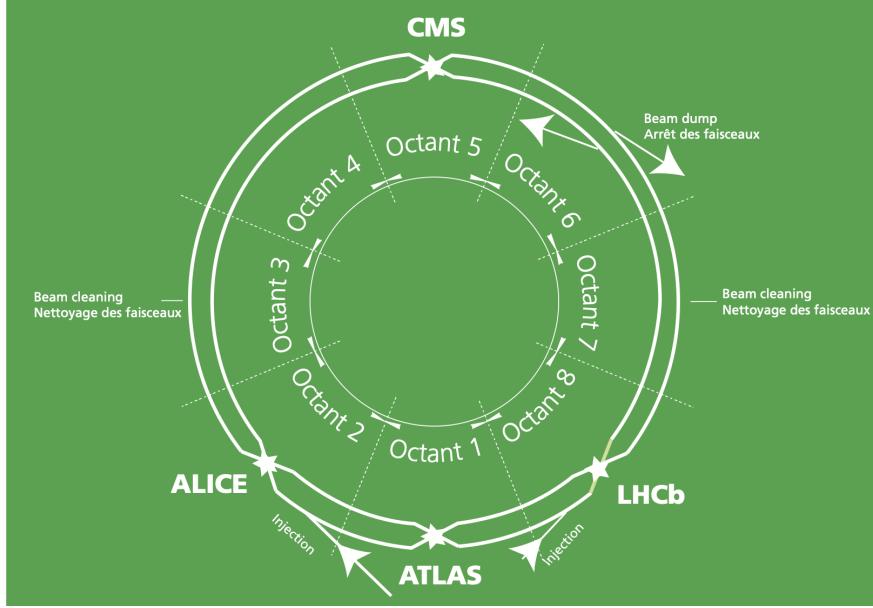


Figure 3.2: The octants of the LHC and location of various beam activities [26]. Stars indicate the locations of beam collisions, and the associated detectors recording the outcome of those collisions.

542 The straight insertion sections have different purposes depending on their location around the
 543 ring: beam collisions, beam injection, beam dumping, or beam cleaning. At the four collision
 544 points, insertion magnets squeeze the beam to ensure a highly focused collision. This is accom-
 545 plished with a triplet of quadrupole magnets, which tighten the beam from 0.2 millimeters to just
 546 16 micrometers in diameter. Insertion magnets also clean the beam, which prevents stray particles
 547 from hitting sensitive components throughout the LHC. When the LHC is ready to dispose of a
 548 beam of particles, beam dump magnets deflect the path of the beam into a straight line towards
 549 a block of concrete and graphite that stops the beam. A dilution magnet then reduces the beam
 550 intensity by a factor of 100,000 before the final stop [28]. Figure 3.2 shows the locations various
 551 beam activities.

552 3.2 Luminosity

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

555 formula for which is given in equation 3.1.

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

556 Here f is the revolution frequency, N_1 and N_2 are the number of particle per bunch for each
557 beam, and σ_x , σ_y are the horizontal and vertical beam widths.

558 The instantaneous luminosity gives the number of the collisions that could be produced at the
559 interaction point per unit of cross-sectional area per unit of time, generally expressed in $\text{cm}^{-2}\text{s}^{-1}$.
560 The integrated luminosity is obtained by integrating the instantaneous luminosity over a given
561 block of time, and measures the total number of collisions which have occurred during that op-
562 eration period. The total integrated luminosity is directly correlated with the size of the datasets
563 collected by the LHC experiments. Total integrated luminosity for Run 2 is illustrated in Figure
564 3.3.

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

572 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
573 peak instantaneous luminosity was $0.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Over the course of Run 1 the LHC collected
574 a total integrated luminosity of 5.46 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$, and 22.8 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$. Following the
575 first long shutdown and upgrade phase of operations, the LHC achieved a center of mass energy
576 $\sqrt{s} = 13 \text{ TeV}$ at the beginning of Run 2 in 2015. The LHC was also able to deliver 2.0×10^{34}
577 $\text{cm}^{-2}\text{s}^{-1}$ peak instantaneous luminosity, double the design value. During LHC Run 2, from 2015-
578 2018, the LHC delivered 156 fb^{-1} of integrated luminosity for proton-proton collisions. Run 3 of

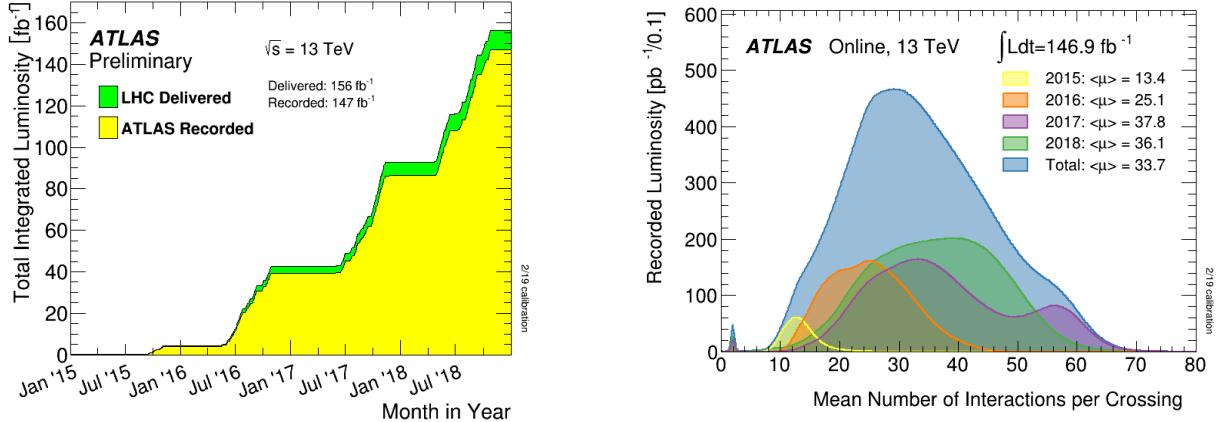


Figure 3.3: (Left) Total integrated luminosity over the course of Run 2. (Right) Average number of pp interactions per bunch crossing in Run 2. Each curve is weighted by the integrated luminosity for the year.

579 the LHC began in 2022, and is expected to deliver 250 fb^{-1} of integrated luminosity to the ATLAS
 580 and CMS experiments by 2026 [29].

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

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

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

591 **3.3 LHC Timeline**

592 The first proton-proton collisions at the LHC were achieved in 2010 with a center-of-mass
593 energy of $\sqrt{s} = 7$ TeV. Run 1 of the LHC took place between 2010 and early 2013, during which
594 time the center-of-mass collision energy increased from 7 TeV to 8 TeV. Figure 3.4 shows an
595 overview of LHC activities beginning in 2011, in the midst of Run 1. The data collected during
596 Run 1 led to the discovery of the Higgs Boston in 2012 [30].

597 Between 2013 and 2015 the LHC underwent the first Long Shutdown (LS1) during which
598 time maintenance and renovation was performed on the accelerator chain, including the repair and
599 consolidation of the high-current splices which connect the super-conducting LHC magnets. Run
600 2 of the LHC took place from 2015 to 2018 and achieved a center-of-mass energy of $\sqrt{s} = 13$ TeV.
601 Analysis of data collected in Run 2 is still on going, and is the subject of study in this thesis.

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

607 Run 3 is scheduled to continue through 2026, at which point the LHC machine and detectors
608 will undergo upgrades for the *high luminosity* LHC (HL-LHC). The HL-LHC will increase the
609 instantaneous machine luminosity by a factor of 5 - 7.5 with respect to the nominal LHC design.
610 The bottom panel of Figure 3.4 shows an overview of the preparation work for the HL-LHC that
611 has been going on concurrently with Run 1, 2, and 3 of the LHC [31].

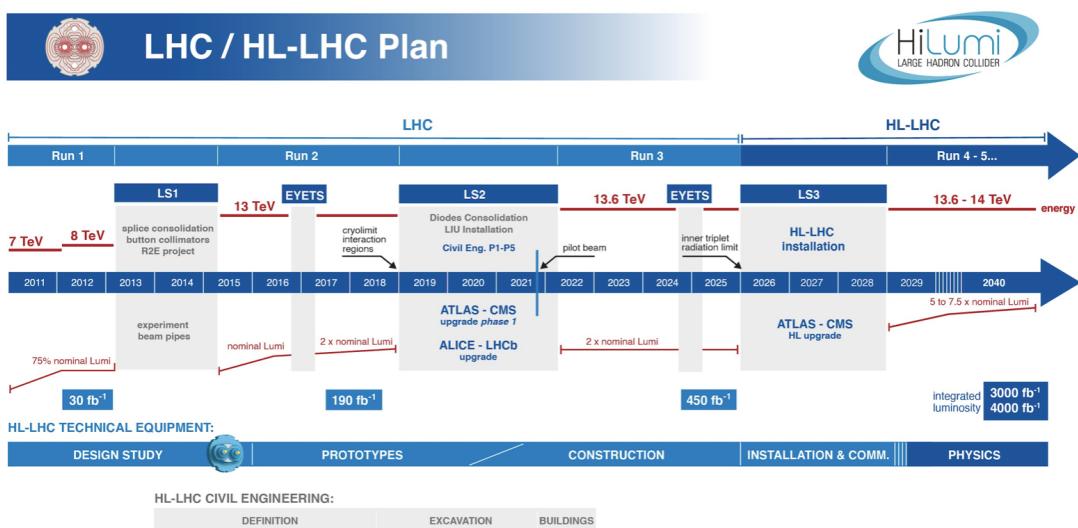


Figure 3.4: Timeline of LHC and HL-LHC activities [29]. Integrated luminosity estimates are approximate, and not reflective of the exact amount delivered to each experiment.

612

613

Chapter 4: The ATLAS Detector

614 The ATLAS detector (**A Toroidal LHC ApparatuS**) is one of two general purpose physics
615 detectors designed to study the products of proton-proton collisions at the LHC. The detector is
616 composed of a variety of specialized subsystems, designed to fully capture a wide array of physics
617 processes. The apparatus is 25m high, 44m in length, and weighs over 7000 tons [32]. The LHC
618 beam pipes direct proton beams to an interaction point at the center of ATLAS, and the cylindrical
619 detector design captures a complete 360° view of the *event*, tracking all particles that result from
620 the collision.

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

631 **4.1 Coordinate System and Geometry**

632 The ATLAS detector employs a right hand cylindrical coordinate system. The z axis is aligned
633 with the beam line, and the x-y plane sits perpendicular to the beam line. The coordinate system
634 origin is centered on the detector, such that the origin corresponds with the interaction point of the

635 two colliding beams. The detector geometry is usually characterized by polar coordinates, where
636 the azimuthal angle ϕ spans the x-y plane. The polar angle θ represents the angle away from the
637 beam line, or z axis. $\theta = 0$ aligns with the positive z -axis, and $\phi = 0$ aligns with the positive x-axis.

638 The polar coordinate θ is generally replaced by the Lorentz invariant quantity *rapidity* or y :

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

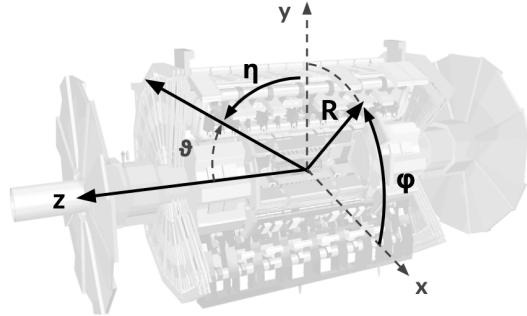
639 This substitution is advantageous because objects in the detector are traveling at highly rela-
640 tivistic speeds. The relativistic speed also means that the masses of the particles are generally small
641 compared to their total energy. In the limit of zero mass, the rapidity y reduces to the pseudorapid-
642 ity η , which can be calculated directly from the polar angle θ :

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

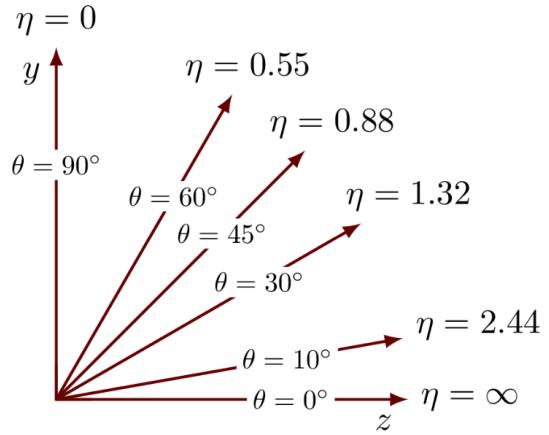
643 The distance between physics objects in the detector is generally expressed in terms of the solid
644 angle between them ΔR :

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

645 Figure 4.1a depicts the orientation of the coordinate system with respect to the ATLAS detector,
646 while Figure 4.1b illustrates the relationship between θ , η , and the beamline axis z . Direct or “head
647 on” proton-proton collisions are more likely to result in objects whose momentum is directed
648 along transverse plane (low $|\eta|$); glancing proton-proton collisions are more likely to result in
649 objects whose momentum is directed along the z -axis (high $|\eta|$). Due to the difference in the
650 nature of these collisions, as well as the cylindrical design of the ATLAS detector, the detector
651 is divided into regions of low and high η . Each subsystem has a “central” or “barrel” region
652 covering low $|\eta|$, while the “forward” or “end-cap” regions cover the area up to $|\eta| = 4.9$. Each of
653 the three main ATLAS subsystems will be discussed in the following sections.



(a) The ATLAS geometry



(b) Relationship between η and θ

Figure 4.1: ATLAS coordinate system and geometry

654 4.2 Inner Detector

655 The Inner Detector (ID) is the ATLAS subsystem closest to the interaction point. The primary
 656 purpose of the ID is to determine the charge, momentum, and trajectory of charged particles pass-
 657 ing through the detector. With this information the ID is also able to precisely determine interaction
 658 vertices.

659 The ID is composed of three sub-detectors; the Pixel Detector, the Semiconductor Tracker
 660 (SCT) and the Transition Radiation Tracker (TRT). Figure 4.2 shows the location of these three
 661 subsystems with respect to each other and the interaction point.

662 4.2.1 Pixel Detector

663 The pixel detector is the first detector encountered by particles produced in LHC collisions.
 664 The original pixel detector consists of 3 barrel layers of silicon pixels, positioned at 5 cm, 9 cm
 665 and 12 cm from the beamline. There are also 3 disks on each end-cap positioned 50 - 65 cm from
 666 the interaction point, providing full coverage for $|\eta| < 2.2$. The layers are comprised of silicon
 667 pixels each measuring $50 \times 400 \mu\text{m}^2$, with 140 million pixels in total. The pixels are organized
 668 into modules, which each contain a set of radiation hard readout electronics chips. In 2014, the
 669 Insertable B-layer (IBL) was installed, creating a new innermost layer of the pixel detector sitting

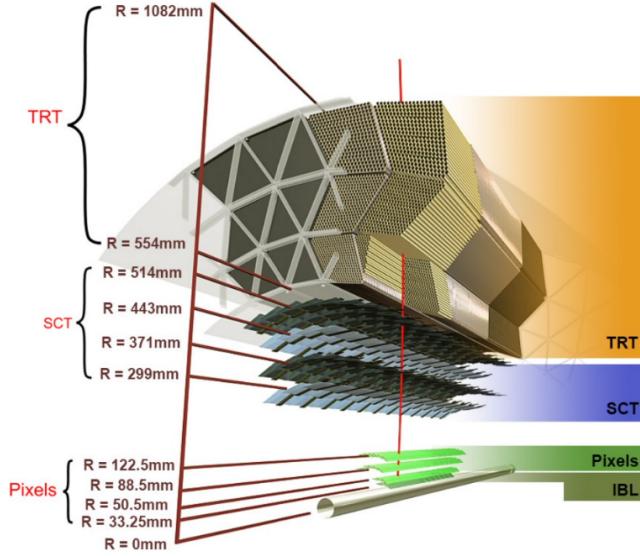


Figure 4.2: A 3D visualization of the structure of the ID in the barrel region [33]

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 with short-lived particles such as the b-quark. The improved vertex identification also helped compensate for increasing pile-up in Run 2 [32].

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

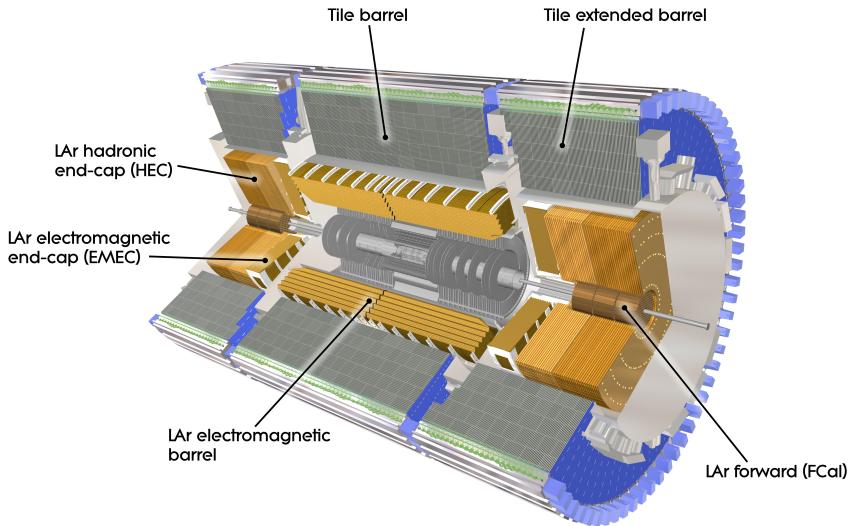


Figure 4.3: ATLAS calorimetry system [34]

684 50000 barrel region straws and 640000 end-cap straws, comprising 420000 electronic readout
 685 channels. Each channel provides a drift time measurement with a spatial resolution of $170\ \mu\text{m}$ per
 686 straw. As charged particles pass through the many layers of the detector, transition radiation is
 687 emitted. The use of two different drift time thresholds allows the detector to distinguish between
 688 tracking hits and transition radiation hits.

689 4.3 Calorimeters

690 The ATLAS calorimeter system is responsible for measuring the energy of electromagnetically
 691 interacting and hadronically interacting particles passing through the detector. The calorimeters are
 692 located just outside the central solenoid magnet, which encloses the inner detectors. The calorime-
 693 ters also stop most known particles, which the exception of muons and neutrinos, preventing them
 694 from traveling to the outermost layers of the detector. The ATLAS calorimetry system is composed
 695 of two subsystems - the Liquid Argon (LAr) calorimeter for electromagnetic calorimetry and the
 696 Tile calorimeter for hadronic calorimetry. The full calorimetry system is shown in Figure 4.3.

697 4.3.1 Liquid Argon Calorimeter

698 The LAr calorimeter is a sampling calorimeter designed to trigger on and measure the ener-
699 gies of electromagnetic (EM) particles, as well as hadronic particles in the high η regions. It is
700 divided in several regions, as shown in Figure 4.3. For the region $|\eta| < 1.4$, the electromagnetic
701 barrel (EMB) is responsible for EM calorimetry, and provides high resolution energy, timing,
702 and position measurements for electrons and photons passing through the detector. The elec-
703 tromagnetic endcap (EMEC) provides additional EM calorimetry up to $|\eta| < 3.2$. In the re-
704 gion $1.4 < |\eta| < 3.2$, the hadronic endcap (HEC) provides hadronic calorimetry. For hadronic
705 calorimetry in the region $|\eta| < 1.4$, corresponding to a detector radii > 2.2 m, the less expensive
706 tile calorimeter (discussed in the next section) is used instead. A forward calorimeter (FCAL)
707 extends the hadronic calorimetry coverage up to $3.1 < |\eta| < 4.8$ [35].

708 The LAr calorimeter is composed of liquid argon sandwiched between layers of absorber mate-
709 rial and electrodes. Liquid argon is advantageous as a calorimeter active medium due to its natural
710 abundance and low cost, chemical stability, radiation tolerance, and linear response over a large
711 energy range [36]. The calorimeter is cooled to 87k by three cryostats: one barrel cryostat encom-
712 passing the EMB, and two endcap cryostats. The barrel cryostat also encloses the solenoid which
713 produces the 2T magnetic field for the inner detector. Front-end electronics are housed outside the
714 cryostats and are used to process, temporarily store, and transfer the calorimeter signals.

715 **Electromagnetic Calorimeter**

716 For the electromagnetic calorimeters, the layers of electrodes and absorber materials are ar-
717 ranged in an an accordion shape, as illustrated in Figure 4.4. The accordion shape ensures that
718 each half barrel is continuous in the azimuthal angle, which is a key feature for ensuring consistent
719 high resolution measurements. Liquid argon permeates the space between the lead absorber plates,
720 and a multilayer copper-polymide readout board runs through the center of the liquid argon filled
721 gap.

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

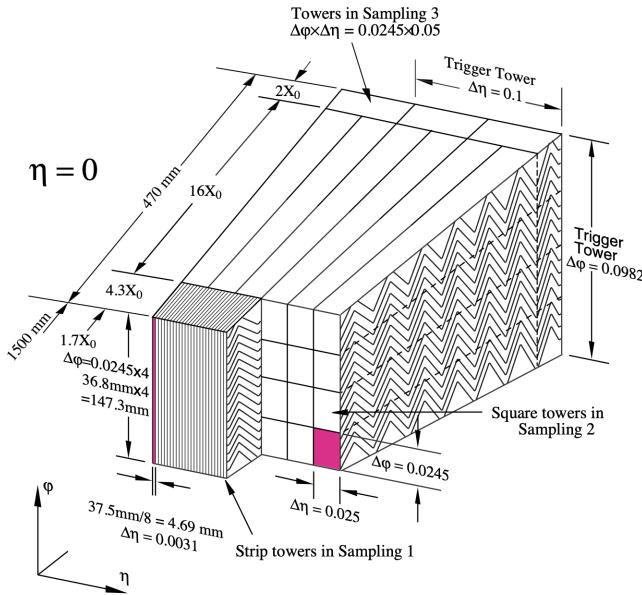


Figure 4.4: Diagram of a segment of the EMB, demonstrating the accordion plate arrangement [35]

723 released when a charged particle ionizes the liquid argon. In the barrel region, the electrons are
 724 driven towards the center electrodes by a 2000 V potential with a drift time of less than 450 ns [37].
 725 In the end-caps the voltage varies as a function of the radius in order to maintain a flat response
 726 [35]. The amount of current produced by the ionized electrons is proportional to the energy of
 727 the particle creating the signal. Figure 4.5 shows the shape of the signal produced in the LAr
 728 calorimeter, before and after it undergoes shaping during the readout process. The shaping of the
 729 pulse enforces a positive peak and a negative tail, which ensures that subsequence pulses can be
 730 separated with the precision required for the 25 ns LHC bunch spacing [35].

731 **Hadronic End-cap Calorimeter**

732 The HEC sits radially beyond the EMEC. The copper absorber plates in the HEC are oriented
 733 perpendicular to the beamline, with LAr as the active medium. Each end-cap is divided into two
 734 independent wheels; the inner wheel uses 25 mm copper plates, while the outer wheel uses 50 mm
 735 plates as a cost saving measure. In each wheel, the 8.5 mm plate gap is crossed by three parallel

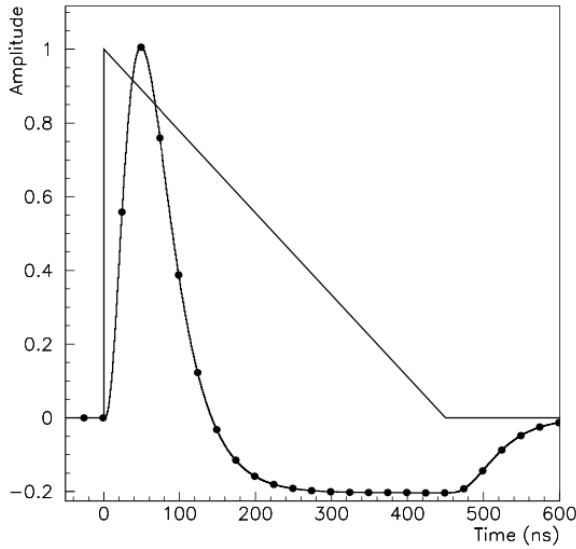


Figure 4.5: A LAr pulse as produced in the detector (triangle) and after shaping (curve) [35]

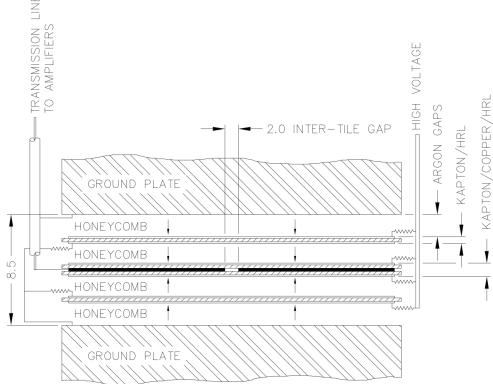


Figure 4.6: Readout gap structure in HEC [35]

736 electrodes, creating an effective drift distance of 1.8 mm. This gap is illustrated in Figure 4.6.
 737 Each wheel is divided into 32 wedge-shaped modules, each containing their own set of readout
 738 electronics.

739 **Forward Calorimeter**

740 The forward range is covered by the FCal, which provides both EM and hadronic calorimetry.
 741 It is composed of three active cylindrical modules; one EM module with copper absorber plates,
 742 and two hadronic modules with tungsten absorber plates. The plates are oriented perpendicular to

743 the beamline, and LAr is used as the active material throughout. The electrodes of the FCal consist
744 of tubes that run parallel to the beam line, arranged in a honeycomb pattern. The resulting LAr
745 gaps are as small as $250\ \mu\text{m}$, which enables the FCal to handle the large influx of particles in the
746 forward region [35].

747 4.3.2 Tile Calorimeter

748 The Tile Calorimeter (TileCal) provides hadronic calorimetry in the region $\eta < 1.7$, and sur-
749 rounds the LAr calorimeter. It is responsible for measurements of jet energy and jet substructure,
750 and also plays an important role in electron isolation and triggering (including muons) [38]. Tile-
751 Cal is composed of 3 sections, as shown in Figure 4.3; a barrel calorimeter sits directly outside the
752 LAr EMB and provides coverage up to $\eta < 1.0$. Two extended barrel sections sit outside the LAr
753 endcaps and cover the region $0.8 < \eta < 1.7$.

754 TileCal is a sampling calorimeter composed of steel and plastic scintillator plates as illustrated
755 in Figure 4.7. A total of 460,000 scintillators are read out by wavelength-shifting fibers. The
756 fibers are gathered to define cells and in turn read out by photomultiplier tubes, which amplify
757 the signal and convert it to an electrical signature. Each cell has an approximate granularity of
758 $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. Each barrel is divided azimuthally into 64 independent modules, an example
759 of which is show in Figure 4.7. The modules are each serviced by front-end electronic housed in a
760 water-cooled drawer on the exterior of the module.

761 The detection principle of the TileCal is the production of light from hadronic particles inter-
762 acting with the scintillating tiles. When a hadronic particle hits the steel plate, a shower of particles
763 are produced. The interaction of the shower with the plastic scintillator produces photons, the num-
764 ber and intensity of which are proportional to the original particle's energy.

765

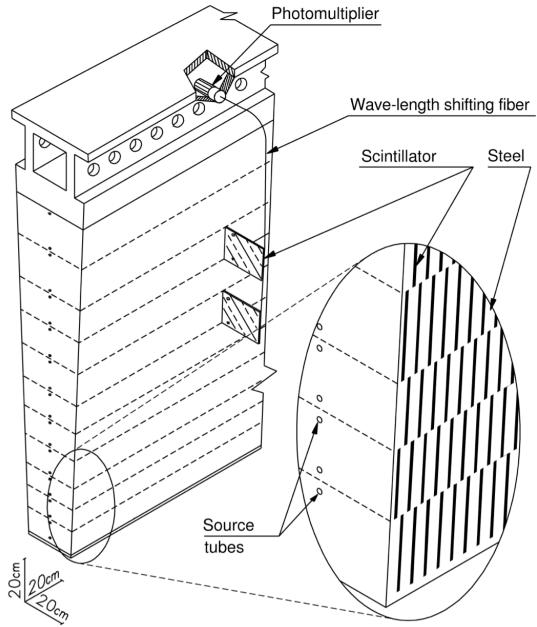


Figure 4.7: TileCal wedge module [38]

766 4.4 Muon Spectrometer

767 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 768 Spectrometer (MS) provides additional tracking information to improve the identification and measurement of muons. The MS comprises the outermost layers of the detector, and is interspersed 769 with toroid magnets (discussed in Section 4.5), which provide a magnetic field of approximately 770 0.5 T. The magnetic field bends the trajectory of the muons as they pass through the detector, and 771 the degree of the bend is directly correlated with the muon momentum. The path of the muon is 772 primarily measured by hits in three layers of Monitored Drift Tube (MDT) precision chambers, 773 which cover the range $|\eta| < 2.7$. The barrel layout of the MS is show in Figure 4.8. 774

776 Muon triggering is provided by three layers of Resistive Plate Chambers (RPC) in the barrel 777 ($|\eta| < 1.05$), and 3 - 4 layers of Thin Gap Chambers (TGC) in the end-caps ($1.05 < |\eta| < 2.4$). 778 RPCs and TGCs also provide muon track measurements in the non-bending coordinate (ϕ). RPCs 779 are constructed from two parallel resistive plates separated by a 2mm gap filled with a sensitive

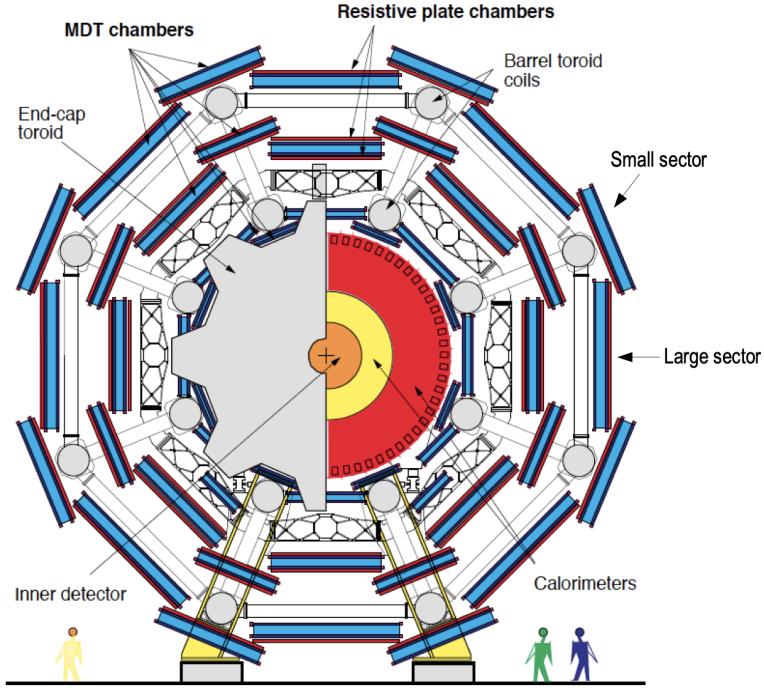


Figure 4.8: Cross section view of the muon spectrometer system [39]

780 gas mixture. This provides a total of six independent measurements for each muon track, with a
 781 spatial resolution of ~ 1 cm and a time resolution of ~ 1 ns. Time measurements from the RPCs
 782 are primarily associated to hits in the MDT precision chambers to determine the bunch crossing.
 783 The time measurement is also used to reject cosmic muons, and to search for delayed signals.
 784 TCGs provide triggering in the end-cap regions, and consist of parallel $30\ \mu\text{m}$ wires suspended
 785 in a sensitive gas mixture. TCGs provide high radiation tolerance and a fast response time, both
 786 features that are necessary for handling the high flux of muons in the forward region [39].

787 Precision measurements of muon momentum and position are primarily achieved by MDTs.
 788 The MDTs are constructed from 30 mm diameter tubes, permeated by a gas mixture of 93% Ar and
 789 7% CO₂. The average single-tube spatial resolution is $80\ \mu\text{m}$. Each chamber consists of six drift
 790 tube layers, which together provide a muon track segment resolution of $35\ \mu\text{m}$. The momentum
 791 of the muons can be calculated from the bend in the muon trajectory as they pass through the
 792 0.5T magnetic field provided by the toroids. For a $p_T = 1$ TeV track, the average p_T resolution is
 793 11%. In the inner most end-cap wheels, Cathode Strip Chambers (CSC) are used instead of MDTs,

794 covering the region $2.0 < |\eta| < 2.7$. CSCs are multi-wire proportional chambers, with a cathode
795 strip readout. The CSCs have a spatial resolution in the range of $50 \mu\text{m}$, and a maximum drift time
796 of about 30 ns, which makes them superior for handling the high flux of particles in the forward
797 region [40].

798 **4.5 Magnet System**

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

804 The inner solenoid magnet is composed of over 9 km of niobium-titanium superconductor
805 wires, which are imbedded into strengthen pure aluminum strips. The solenoid is just 4.5 cm
806 thick, which minimizes interactions between the magnet material and particles passing through the
807 detector. It is housed in the LAr cryostat, as described in section 4.3.1, which further reduces the
808 amount of non-detector material required to support the solenoid. The return yoke of the magnet
809 is provided by the iron absorber of the TileCal [41].

810 The central ATLAS toroid magnet, providing the magnetic field for the barrel region of the MS,
811 is the largest toroidal magnet ever constructed at 25 m in length. The toroid is composed of eight
812 individual coils, each housed in their own cryostat. The toroidal magnetic field is advantageous
813 as the direction of the field is almost perpendicular to the path of the charged particles. 56 km of
814 aluminum stabilized niobium-titanium-copper superconductor wire compose the magnet. In each
815 end-cap, eight smaller superconducting coils extend the toroidal magnetic field to particles leaving
816 the detector in the forward direction [41]. Figure 4.9 shows the layout of the toroid magnets.

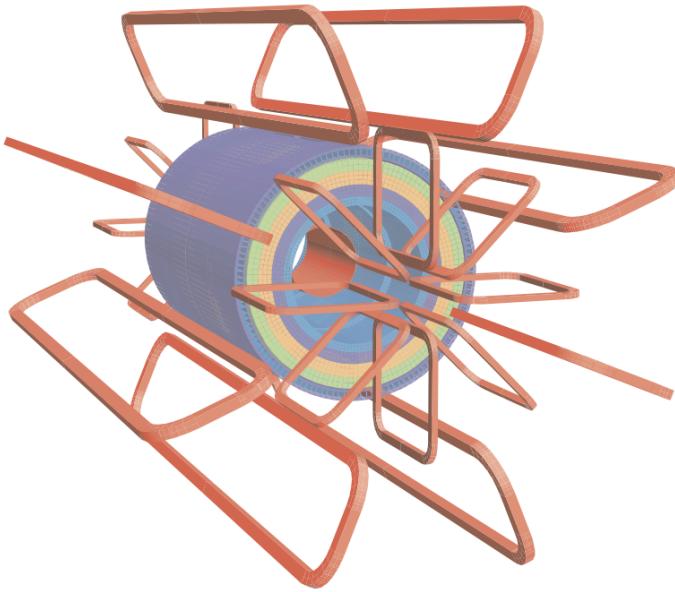


Figure 4.9: Layout of the barrel and endcap toroid magnets [32]

817 4.6 Forward Detectors

818 In addition to the inner detector, calorimeters, and muon spectrometer, three smaller detectors
819 provide coverage in the very forward region. The innermost forward detector, at 17 m from the
820 interaction point, is the **L**Uminosity measurement using **C**erenkov **I**ntegrating **D**etector (LUCID).
821 LUCID's primary purpose is to measure the relative online-luminosity for the ATLAS detector,
822 from inelastic $p - p$ scattering. The detector is composed of 20 aluminum Cerenkov tubes which
823 surround the beam pipe and face towards the interaction point.

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

827 The third forward detector is the Absolute Luminosity For ATLAS (ALFA) system, located 240
828 m from the interaction point in both directions. ALFA determines luminosity by measuring elastic
829 scattering at small angles, from which luminosity can be calculated via the optical theorem. The
830 detector is built from scintillating fiber trackers. These are connected to the accelerator vacuum

831 via Roman pots, which allow the detector to come as close as 1mm to the beam without disrupting
832 the machine vacuum. The LUCID and ALFA detectors are crucial to determining the real-time
833 conditions of the beams and the total luminosity delivered to the ATLAS detector [32].

834 **4.7 Trigger and Data Acquisition**

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

841 The L1 trigger is implemented in the hardware of the ATLAS calorimeter and muon systems.
842 The primary modality of the L1 trigger is to identify muons, electrons, photons, jets, and τ -leptons
843 with high transverse momentum. Particles with high transverse momentum are more likely to
844 originate from direct, high energy collisions, which are most likely to produce interesting physics
845 processes. The L1 trigger also identifies events with large missing transverse energy, which could
846 be indicative of new physics. The L1 muon trigger (L1Muon) relies on RPC and TGC trigger
847 chambers in the barrel and end-cap regions of the muon spectrometer. The L1 Calorimeter Trigger
848 (L1Calo) uses reduced granularity information collected by all the calorimeter subsystems. Results
849 from the L1Muon and L1Calo triggers are combined by the Central Trigger Processor (CTP),
850 which implements a trigger ‘menu’, listing various combinations of trigger requirements. The
851 maximum L1 acceptance rate is 75 kHz, and the L1 trigger decision must reach the front-end
852 electronics within $2.5 \mu\text{s}$ of its associated bunch-crossing [32].

853 The L1 trigger defines a Region-of-Interest (RoI) for each passing event. The ROI is repre-
854 sented by the η - ϕ detector region where interesting features were identified by the L1 selection
855 process. Information about the type of feature identified and the threshold which was exceeded to
856 trigger the L1 response is also recorded. The ROI data is sent to the L2 trigger, which uses all of

857 the available information within the ROI at full granularity and precision. The L2 trigger reduces
858 the event rate from 75 kHz to 3.5 kHz, with an average processing time of 40 ms. The final stage of
859 the HLT is the event filter, which reduces the event rate to 200 Hz. The event filter uses an offline
860 analysis process to select fully rebuilt events which will be saved for further analysis.

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

873 The DCS allows centralized control of all detector subsystems simultaneously. It continually
874 monitors operational conditions, reports any abnormal behavior to the operator, and can perform
875 both automatic and manual interventions. The DCS reports on real time detector conditions such
876 as high or low voltage detector electronics, gas and cooling systems, magnetic field conditions,
877 humidity and temperature. This information is continually monitored by experts in the ATLAS
878 control room, so that action can be taken immediately to correct any issues that arise. The DCS also
879 handles communication between detector systems, and other systems such as the LHC accelerator,
880 the ATLAS magnets, and CERN technical services [32].

881

882

Chapter 5: Particle Reconstruction and Identification

883 With a design luminosity of $1.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, and a peak Run-2 instantaneous luminosity of
884 $2.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, reconstructing and identifying the products of LHC $p p$ collisions is one of the
885 most complex tasks for each LHC experiment. The accurate reconstruction and identification of
886 physics objects lays the ground work for all subsequent physics analyses, so it is also one of the
887 most fundamentally important tasks performed by an experiment.

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

898 **5.1 Inner Detector Tracks**

899 As the inner most layer of the detector, the ID measures charged particles close to the interac-
900 tion point. The various hits of these charged particles throughout the ID are used to reconstruct
901 *tracks* which give the trajectories of charged particles [43]. Track reconstruction begins by clus-
902 tering hits in the Pixel and SCT detectors, and combining clusters from different radial layers of
903 these detectors. The multi-layer clusters form track *seeds*, which provide initial estimates of mea-

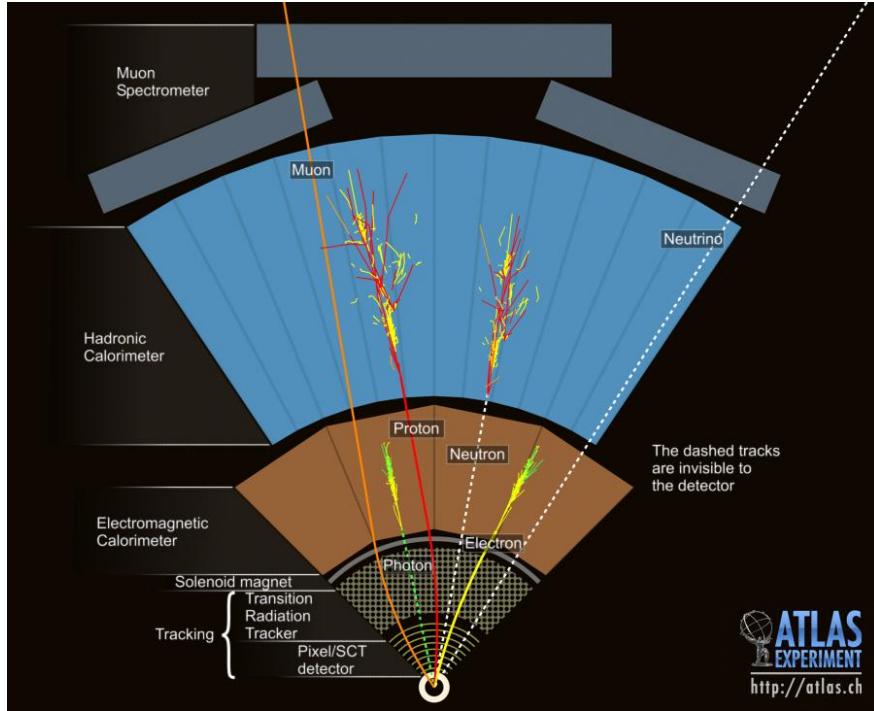


Figure 5.1: Graphic illustrating the various objects and high level features identified by ATLAS object reconstruction, and their interaction with different systems of the ATLAS detector [42]

904 surements belonging to an individual track. The requirement of three points allows for a rough
 905 estimate of the track p_T to be made by calculating the curvature of the track and accounting of the
 906 magnetic field in the ID.

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

911 **5.2 Photons and Electrons**

912 Photons and electrons shower in the LAr calorimeter, and are identified by the energy deposits
 913 they leave there. Energy deposits in a collection of nearby cells are termed *clusters*, which become
 914 the starting point for electron and photon reconstruction [44]. The clustering algorithm begins
 915 when the energy deposit in a certain cell exceeds the noise threshold with a significance of 4σ .

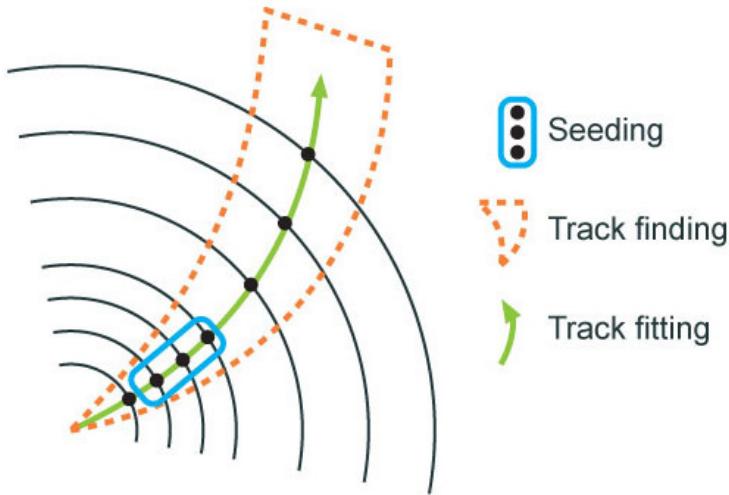


Figure 5.2: Track reconstruction seeding, finding and fitting illustration [43]

916 The algorithm then collects neighboring cells which have an energy deposit exceeding the noise
 917 threshold with a significance of 2σ , creating a topo-cluster. Next, these topo-clusters are matched
 918 to ID tracks, created as described in Section 5.1. The location of the topo-cluster defines a region
 919 of interest (ROI) in the ID, where additional modified track reconstruction algorithms are run in the
 920 case that no associated tracks are found. Any ID tracks associated to the topo-cluster are retrofitted to
 921 allow for additional energy loss due to bremsstrahlung. A converted photon track reconstruction
 922 algorithm is run to check for tracks coming from secondary vertices consistent with converted
 923 photons. The secondary vertices are constructed from two oppositely charged tracks consistent
 924 with a massless particle, or from one track without any hits in the innermost layer of the ID.

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

929 *Superclusters* are built separately for photons and electrons, based on the combined topo-cluster
 930 and ID track information. First, the EM topo-clusters are tested to see if they meet the minimum
 931 requirements to become electron or photon seed clusters. For electrons, the cluster must have a
 932 minimum E_T of 1 GeV, and must be matched to a track with at least 4 hits in the silicon tracking

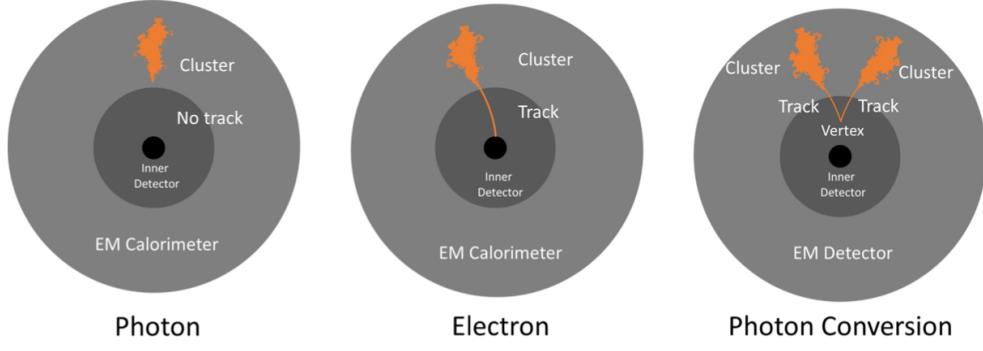


Figure 5.3: Three types of EM object candidates [45].

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

Electron and photon objects are identified from the superclusters after energy calibration and position corrections are applied. 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). 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 [46]. A χ^2 fit is

951 performed on each track candidate. Based on the χ^2 criteria, hits are removed or added such that
952 the track contains as many hits as possible while satisfying the fit criteria.

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

- 956 • Combined muon: a muon track identified through independent track reconstruction in the
957 ID and MS, where the combined track is formed using a global refit that uses hit information
958 from both detectors. Most muons are constructed through an outside-in procedure, in which
959 a muon track candidate is identified in the MS and then an associated track is found in the ID.
960 A complementary inside-out procedure is also implemented and identifies additional muons.
- 961 • Segment-tagged muon: an ID track is identified as a muon if when extrapolated out to the
962 MS (following the inside-out global fit procedure) it is matched to at least one local MS
963 segment.
- 964 • Calorimeter-tagged muon: an ID track is identified as a muon if it is matched to a calorimeter
965 energy deposit that is compatible with a minimum-ionizing particle. This muon identifica-
966 tion has the lowest purity, but it used in regions where the MS has only partial coverage due
967 to cabling and service access routes.
- 968 • Extrapolated muons: the muon is reconstruction only from the MS track and a requirement
969 on compatibility with the primary interaction point. The muon track is required to cross at
970 least two layers of the MS, and three layers in the forward region. These muons are mainly
971 used to extend muon acceptance into the region $2.5 < |\eta| < 2.7$ where ID track information
972 is not available.

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

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

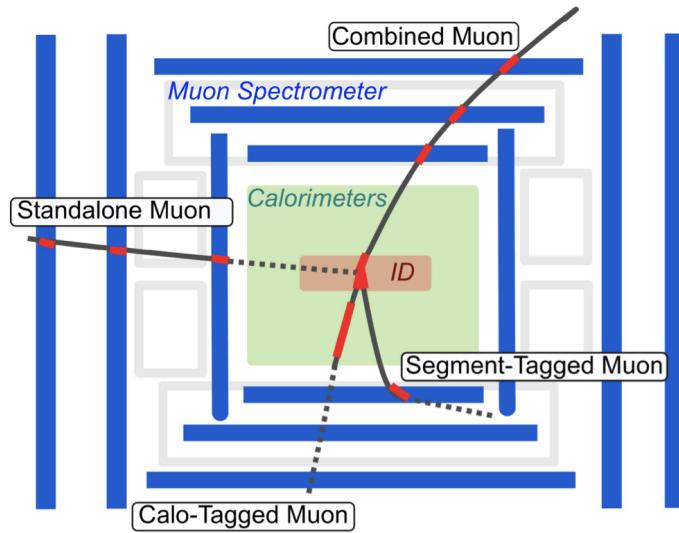


Figure 5.4: Four types of muon track candidates [47].

982 5.4 Jets

983 The protons accelerated in the LHC are composed of quarks and gluons, and thus their colli-
 984 sions often result in the release of energetic quarks and gluons, collectively termed *partons*. The
 985 energetic partons can radiate additional gluons, and these gluons can pair produce quarks in a pro-
 986 cess called *fragmentation*. Fragmentation continues until the energy drops sufficiently that color
 987 conservation plays a dominant role. At that point, additional quarks and gluons are produced from
 988 vacuum to create neutral color states for the fragmented collection of partons. This process is
 989 known as *hadronization* [48]. The hadronized partons compose a collimated stream of particles,
 990 known as a *jet*, which is then observed in the detector. The full process that produces jets is known
 991 as a *parton shower*, and is illustrated in Figure 5.5.

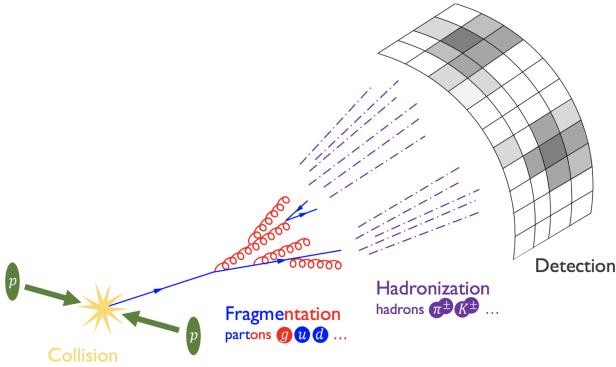


Figure 5.5: The fragmentation and hadronization processes undergone by a quark produced in a proton-proton collision [49].

992 Jets are identified by the energy deposits they leave in the calorimeter, which are then matched
 993 to the tracks they leave in the ID. Jet reconstruction generally begins in the calorimeters, with the
 994 identification of *topo-clusters*. Then jet reconstruction algorithms combine calorimeter informa-
 995 tion with tracking information. The anti- k_t algorithm [50] as provided by the FastJet library [51] is
 996 generally used by the ATLAS experiment, with varying reconstruction radius settings. There are a
 997 variety of jet collections depending on the exact usage of calorimeter and tracking information in
 998 the reconstruction. Some common collections include particle flow jets (PFlow), track calo-cluster
 999 jets (TCC), EM topo-cluster jets (EMTopo), and unified flow object jets (UFO). Only particle flow
 1000 jets will be discussed in greater detail due to their importance in this analysis. The following sec-
 1001 tions discuss jet identification in the calorimeters, particle flow jet construction using the anti- k_t
 1002 algorithm, jet clustering and jet substructure characteristics.

1003 5.4.1 Calorimeter Clusters

1004 Jets are first identified by the energy deposits they leave in the calorimeters. The reconstruc-
 1005 tion of jets in ATLAS begins with the construction of *topo-clusters*, which are topologically-grouped
 1006 noise-suppressed clusters of calorimeter cells [52]. The topo-cluster seed is a cell with an energy
 1007 that is at least 4× the average background noise level for the cell. Any cells adjacent to the seed
 1008 cell in three dimensions are added to the cluster if they have an energy deposit that is at least 2x

1009 the average expected noise. This process is repeated, growing the cluster, until no adjacent cells
1010 exceeding the energy deposit threshold remain. As a final step, all adjacent cells are added to the
1011 topo-cluster, irrespective of their energy.

1012 The construction process for topo-clusters allows for the possibility that several independent
1013 signatures are grouped into one topo-cluster. To correct for this, the topo-cluster is scanned for
1014 local maximum, defined by any cell with energy > 500 MeV, and no neighboring cells with greater
1015 energy. If more than one local maximum is identified, the topo-cluster is split among the corre-
1016 sponding energy peaks [53]. In the event that one cell neighbors two or more local maxima, the
1017 cell is assigned to the two highest-energy clusters that is neighbors. This means each cell is shared
1018 at most once, between at most two post-splitting topo-clusters.

1019 Two measurements for the total energy of the topo-cluster are considered. The raw, or electro-
1020 magnetic (EM), scale simply considers the sum of energy from all cells in the topo-cluster. The
1021 local cell weighting (LCW) scale first classifies clusters as electromagnetic or hadronic, and then
1022 applies appropriate corrections for hadronic interactions in the jet energy calculation [52]. The
1023 corrections are derived from Monte Carlo simulations, and account for the weaker response of
1024 ATLAS calorimeters to hadronic interactions (ATLAS calorimeters are *non-compensating*), and
1025 hadronic energy losses due to interactions with dead material [53].

1026 5.4.2 Particle Flow Algorithm

1027 The calorimeters provide excellent jet energy resolution for high energy jets. However, the
1028 granularity of the hadronic calorimeter is restricted to 0.1×0.1 in $\eta \times \phi$. Combining the infor-
1029 mation from the calorimeter with tracking information provides superior angular resolution and
1030 energy resolution. The particle flow (PFlow) algorithm is one of a handful of algorithms which
1031 can perform this task.

1032 An overview of the process is given in Figure 5.6. Tracks from the ID which are selected for the
1033 PFlow algorithm are required to have at least 9 hits in the silicon detector, and missing pixel hits in
1034 places where a hit would be expected. Additionally, the tracks have $p_T > 0.5$ GeV, and $|\eta| < 2.5$.

1035 The algorithm then attempts to match these tracks to EM scale calorimeter topo-clusters. This
1036 matching is performed using the distance metric

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

1037 where σ_η and σ_ϕ represent the angular widths of the topo-clusters, and $\Delta\eta$ and $\Delta\phi$ represent the
1038 distance between the track, extrapolated to the second layer of the EM calorimeter, and barycenter
1039 of the topo-cluster [54]. The topo-cluster closest to the track as measured by $\Delta R'$ is considered
1040 matched to the track. If no topo-cluster is found within the cone size of $\Delta R' = 1.64$, it is assumed
1041 that particle which left the track did not form a topo-cluster in the calorimeter.

1042 The PFlow algorithm predicts the expected single topo-cluster energy for a given track, based
1043 on the track momentum and topo-cluster position. This value is then compared to the observed
1044 energy of the topo-cluster, and the probability that the particle energy was deposited in more than
1045 one topo-cluster is evaluated. If necessary, the algorithm adds more topo-clusters to the track/topo-
1046 cluster system, in order to account of the full shower energy of the track particle.

1047 To reduce the impact of double counting the energy of a given particle by including both its
1048 tracker and calorimeter energy measurements, the calorimeter energy measurements associated to a
1049 given track are subtracted from the total calorimeter measurement. If the expected energy deposited
1050 by the particle exceeds the topo-cluster energy, the full topo-cluster is removed. If the expected
1051 energy is less than the EM scale energy of all the considered topo-clusters, topo-cluster cells are
1052 removed one by one, until the full expected energy deposit of the particle has been removed from
1053 the calorimeter information. The resulting set of tracks and topo-clusters represent the event with
1054 no double-counting of energy between subdetectors [54]. This information is passed to the jet-
1055 finding algorithm.

1056 5.4.3 Jet Clustering

1057 When a parton decays in the detector, its energy deposits often result in multiple calorimeter
1058 clusters. For physics purposes, it is useful to combine clusters that likely resulted from an individ-

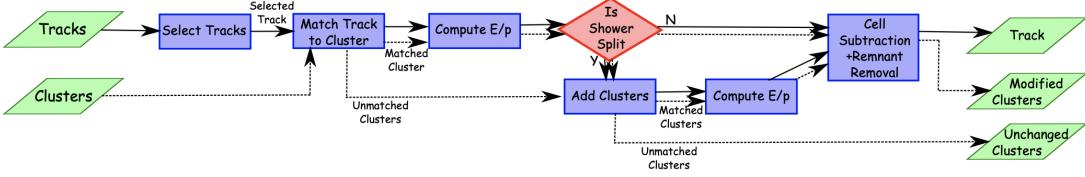


Figure 5.6: A flow chart illustrating the particle flow algorithm progression [54].

1059 ual parton decay, in order to reconstruct the parton. The process of grouping topo-clusters which
 1060 were produced by the same parton decay is *jet clustering*.

1061 The anti- k_t algorithm is the most commonly used algorithm for jet clustering in ATLAS. The
 1062 anti- k_t algorithm is based on sequential recombination algorithms [55]. A sequential recombina-
 1063 tion considers the distance d_{ij} between objects i and j (particles or pseudojets), and the distance
 1064 d_{iB} between an object i and the beam line B . If d_{ij} between two objects is the smallest distance
 1065 among those considered, i and j are combined into a pseudojet. The process continues until the
 1066 smallest distance is d_{iB} at which point the object i is determined to be a jet and removed from the
 1067 objects in consideration. The procedure is repeated with the remaining objects until there are none
 1068 remaining [50].

1069 The anti- k_t algorithm adopts this procedure, but modifies the distance measurements d_{ij} and
 1070 d_{iB} to consider the transverse momentum k_t :

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

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

1071 The addition of the term p allows adjustments to algorithm which govern the relative power of
 1072 the momentum versus the geometrical scale $\Delta_{i,j}$, which is defined as $\Delta_{i,j} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$
 1073 where y_i and ϕ_i are respectively the rapidity and azimuth of particle i [50]. The radius parameter
 1074 R is chosen and determines the geometric cone size [55].

1075 In the case $p = 1$ the inclusive k_t algorithm [55] is recovered, which is a standard sequential

combination jet clustering algorithm. In the case $p = 0$, the Cambridge/Aachen sequential combination algorithm [56] 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 [57], which checks for sets of stable cones compatible with the observed radiation.

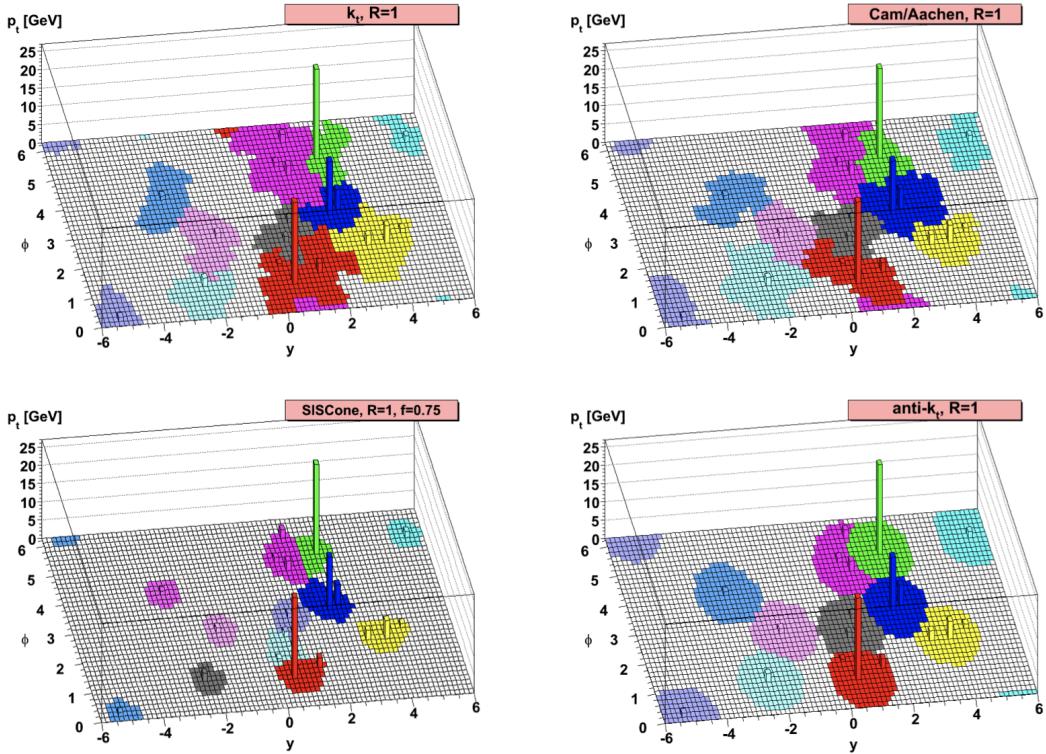


Figure 5.7: A comparison of jet clustering with four different jet algorithms. The anti- k_t algorithm is observed to create the most conical jets, where the shape of the jet is immune to the presence of soft radiation [50].

Any useful jet clustering algorithm must satisfy the requirements of infrared safety and collinear

1087 (IRC) safety. Infrared safety implies that the resulting set of jets is unaltered by the presence of
1088 additional soft particles in the list of seed clusters. As explained above, the anti- k_t algorithm is
1089 natural infrared safe. Collinear safety requires that the final set of jets is not impacted by collinear
1090 splitting of one of the jets. If the hardest particle p_1 is split into a collinear pair (p_{1a}, p_{1b}) (as
1091 is common in the fragmentation process for a hard parton), the jet clustering algorithm must
1092 still recognize (p_{1a}, p_{1b}) as the hardest jet in the collision. If another softer particle p_2 with
1093 $p_{t,1a}, p_{t,1b} < p_{t,2} < p_{t,1}$ is instead considered the hardest particle in the event, a different final
1094 set of jets would be returned. Collinear safety is a requirement of perturbative QCD, to ensure
1095 non-divergent higher-order calculations [58]. The anti- k_t algorithm's tendency to cluster hard par-
1096 ticles first ensures its collinear safety. By satisfying the IRC safety requirement, anti- k_t jets can be
1097 calculated using perturbative QCD, which improves comparisons with theory.

1098 5.4.4 Ghost Track Association

1099 Once a collection of jets has been created, the jet objects can be studied at both the event-level
1100 and the jet-level. In the event-level picture, the momentum, energy, and geometric orientation
1101 of the jets within an event are considered. This yields important information about decay of any
1102 resonant heavy objects, the total energy in the event, and the distribution of energy amongst the jets.
1103 At the jet-level picture, the particles composing the jet are considered. The momentum, energy,
1104 and geometric orientation of the particle tracks provides a low-level picture of the jet, which can
1105 help determine if the properties of the jet are consistent with standard QCD, or if new physics
1106 processes might be represented within the patterns found in the constituent particles. Jet-level
1107 analysis is also widely used in flavor tagging.

1108 For anti- k_t jets with a radius parameter $R = 0.4$, one way of studying the jet-level picture
1109 is through considering the ghost-associated tracks. Track association is process of determining
1110 which tracks should be considered associated with a given jet. In the ghost association algorithm,
1111 the anti- k_t clustering algorithm is used for the collection of tracks and calorimeter clusters [59].
1112 However, the tracks are considered to have infinitesimal momentum (*ghosts*), so their addition to

1113 a jet object does not alter the four-momentum of the jet. This ensures the final jet collection is not
1114 altered by the presence of the ghost tracks in the reclustering, but information about the associated
1115 tracks for each reconstructed jet becomes available [60].

1116 Ghost tracks are of particular importance to this analysis, as a means of providing a low-level
1117 picture of the shape of $R = 0.4$ jets, and discriminating Standard Model QCD-like jets from dark
1118 QCD-like jets.

1119 5.5 Missing Transverse Energy

1120 A simple principle leveraged in ATLAS physics analyses is checking for conservation of mo-
1121 mentum among the products of any pp collisions. The initial state transverse momentum of any
1122 pp collision is always zero, so the transverse momentum of all final state particles should likewise
1123 be zero. The missing transverse energy, E_T^{miss} , is determined by the magnitude of the negative
1124 momentum vector sum of all final state objects resulting from the pp collision.

1125 Specifically, the objects considered in the E_T^{miss} calculation are photons, electrons, muons, jets,
1126 and soft terms. The first four items comprise the hard components of the E_T^{miss} calculation, and
1127 have been discussed previously in this chapter. The final item represents a collection of *soft terms*,
1128 comprising any detector signals not associated to hard detector objects. These can be based on
1129 unassociated tracks, or unassociated soft calorimeter clusters. Both are generally not used in the
1130 same calculation to avoid double counting of soft terms. In this analysis the calorimeter cluster
1131 soft terms are considered in the E_T^{miss} calculation.

1132 E_T^{miss} can arise due to non-interacting Standard Model objects such as neutrinos, fake sources
1133 such as mis-reconstructed objects and dead detector regions, or in some theories, non-interacting
1134 BSM objects such as a dark matter candidate particles. To understand the amount of E_T^{miss} at-
1135 tributable to detector noise and mis-reconstruction, E_T^{miss} is studied in $Z \rightarrow \mu\mu$ where little real
1136 E_T^{miss} is expected [61]. As Figure 5.8 illustrates, the resolution of E_T^{miss} generally decreases as
1137 E_T^{miss} increases, due to detector resolution effects. As E_T^{miss} is an important quantity for most dark
1138 QCD analyses, limitations in the accuracy of the E_T^{miss} calculation must be considered.

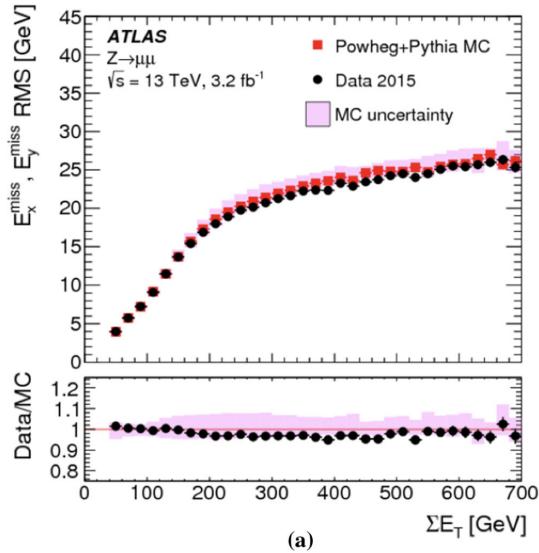


Figure 5.8: A comparison of MC simulation and data for $Z \rightarrow \mu\mu$ events where real $E_T^{\text{miss}} = 0$ [61]. The resolution of the missing energy in the transverse ($x - y$) plane is observed to increase with increasing total $\sum E_T$.

1139

Part III

1140

Search

1141

1142

Chapter 6: Monte Carlo and Data

1143 The search for semi-visible jets via s-channel production presented in the following chapters is
1144 performed with 139 fb^{-1} of proton-proton collision data collected by the ATLAS detector during
1145 Run 2 (2015 - 2018). The full Run-2 dataset is used for the final interpretation. Monte Carlo
1146 (MC) simulations of background processes and the semi-visible jet signal process are used in the
1147 development of the analysis strategy, and in the final interpretation to set limits on the observed
1148 cross-section of the signal model. This chapter will provide details about the full Run-2 dataset,
1149 and the background MC simulations, and the signal MC simulations used in this search.

1150 6.1 Data

1151 The 139 fb^{-1} of proton-proton collision data is selected according to the *good runs list* (GRL),
1152 which identify runs with good detector conditions that are suitable for physics evaluation. Events
1153 are further selected to pass a single-jet trigger selection, where events are required to have a jet at
1154 trigger-level with a p_T that exceeds a certain value. The lowest p_T unprescaled single jet trigger
1155 threshold for each period is as follows:

- 1156 • 2015: $p_T \geq 360 \text{ GeV}$
- 1157 • 2016 & 2017: $p_T \geq 380 \text{ GeV}$
- 1158 • 2017: $p_T \geq 380 \text{ GeV}$
- 1159 • 2017 & 2018: $p_T \geq 420 \text{ GeV}$

1160 A post-trigger selection of jet $p_T > 450 \text{ GeV}$ ensures all these triggers are fully within their
1161 efficiency plateaus. The jet collection used is anti- k_t EM particle flow jets with a radius parameter
1162 of $R = 0.4$, also referred to as small-R jets.

1163 Due to the variance in visible and invisible momenta due to the R_{inv} parameter of the signal
1164 model, many signals also have significant E_T^{miss} . The use of a E_T^{miss} trigger to select events was
1165 considered, and the single jet approach described here was found to preserve more signal events
1166 across the grid, particularly in the high resonance mass and low R_{inv} region of phase space. These
1167 studies are documented in Appendix A.

1168 The data are subject to a blinding strategy throughout the analysis design so as to mitigate
1169 analyzer-induced bias. Blinded and unblinded region definitions are described further in Sec-
1170 tion 8.1.

1171 6.2 Simulation

1172 Simulated events are generated with a variety of Monte Carlo (MC) generator processes that
1173 run in stages. The pp hard scatter physics process is simulated, and the final state particles are
1174 subsequently showered and decayed. This full description of the event is then propagated through
1175 a detailed detector simulation based on GEANT4 [62]. The MC simulation is weighted to match
1176 the distribution of the average number of interactions per bunch crossing μ observed in collision
1177 data.

1178 All simulated samples included in this analysis were produced with three different campaigns:
1179 `mc20a` corresponds to 2015-2016 data-taking conditions, `mc20d` to 2017, and `mc20e` to 2018.
1180 These three campaigns are weighted to the integrated luminosities of their respective data-taking
1181 periods and combined to produce simulation for the entire Run 2 dataset. Simulated events are
1182 reconstructed with the same algorithms run on collision data.

1183 6.2.1 Simulated Backgrounds

1184 Though the final background estimation is data-driven, background MC is studied for analysis
1185 optimization and machine learning tool development.

1186 Dijet QCD is the dominant background process. QCD is simulated with PYTHIA8 [63], and
1187 generated in approximate slices of p_T , to ensure high statistics across the momentum spectrum.

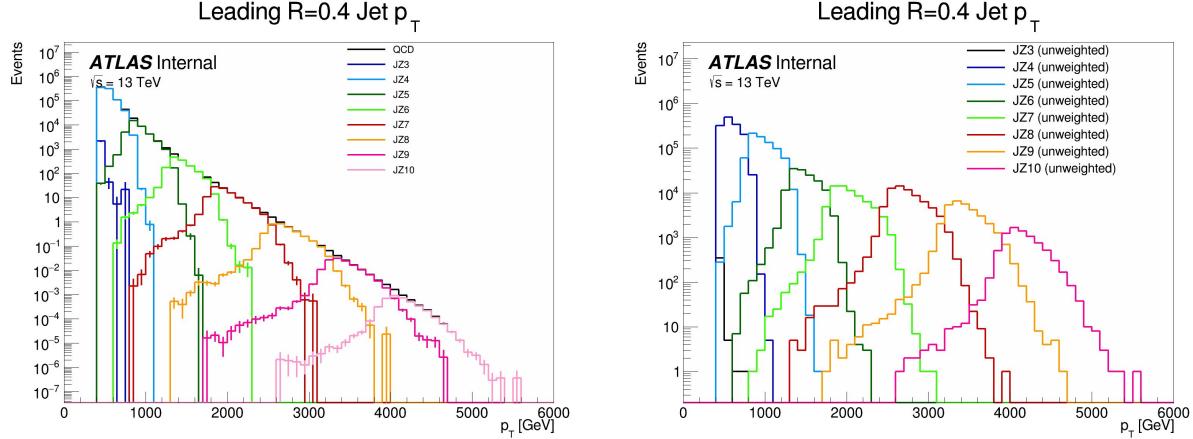


Figure 6.1: The transverse momentum slices of the QCD MC simulation, overlayed to show how they come together to create a smooth distribution (left) once weighted properly. The original unweighted distribution is shown on the right, illustrating the enhanced statistics for the high p_T range.

1188 The slices are then reweighted using MC generated event weights to create a physical distribution.

1189 Figure 6.1 illustrates the 8 momentum slices used in this analysis.

1190 Due to presence of E_T^{miss} in the SVJ signals, additional MC background processes are required
 1191 to create a full picture of the relevant background. The $Z \rightarrow vv$ process contributes to the back-
 1192 ground due to its high missing energy. Leptonic W/Z decays and W/Z+jets are also included as
 1193 they can contribute both additional missing energy and significant hadronic activity. Single top
 1194 and $t\bar{t}$ processes are also considered for their contribution to hadronic activity,. The background
 1195 composition in the selected region (discussed in Section 8.1) is 76% QCD, 12% W/Z+jets, 8%
 1196 top and $t\bar{t}$ processes, and 4% $Z \rightarrow vv$. Figure 6.2 illustrates the background composition for the
 1197 analysis.

1198 6.2.2 Signal Simulation

1199 The Hidden Valley signal model implementation is based on Ref [64]. The s-channel semi-
 1200 visible jet model is governed by a number of parameters. The mass of the mediator $m_{Z'}$ can be set,
 1201 together with the couplings of the Z' to the visible and dark quarks g_q and g_{q_D} . The dark sector
 1202 shower is governed by the number of dark colors N_{c_D} , the number of dark flavors N_{f_D} , and the

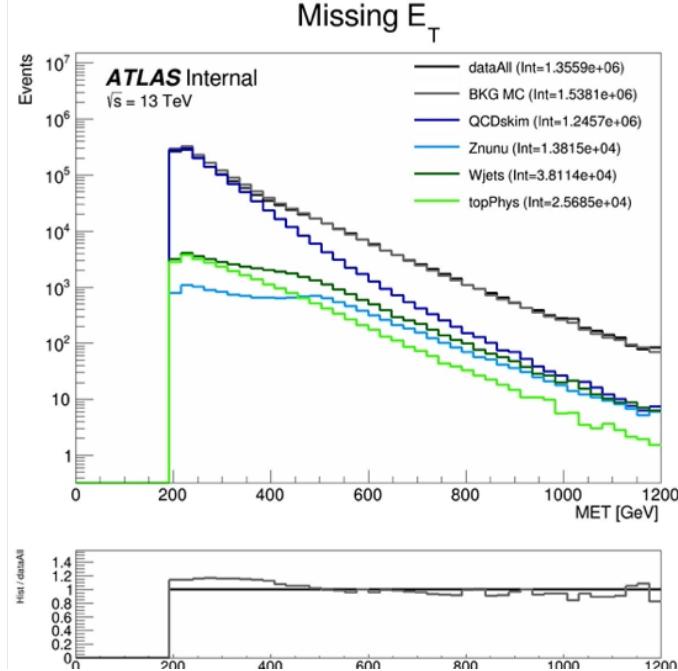


Figure 6.2: Background processes relevant to the SVJ signal. The agreement between the black line (data) and grey line (all MC processes combined) illustrates that this collection of background processes is sufficient to model the expected E_T^{miss} in the selected data events.

1203 dark sector confinement scale Λ_D . There is also the characteristic scale of the dark hadrons m_D ,
 1204 which determines the mass of the dark hadrons, which can be pseudoscalars m_{π_D} or vectors m_{ρ_D} .
 1205 Finally, the average fraction of invisible particles in the final state jet is dictated by R_{inv} .

1206 The chosen parameters for this model were carefully selected in collaboration with theorists to
 1207 be compatible with the new benchmarks established in the 2021 Snowmass process [20]. These pa-
 1208 rameters reflect extensive communication with the Snowmass, CMS, and theory teams. The signal
 1209 generation allows for up to two initial state radiation jets, and uses an MLM merging scheme [63]
 1210 to match jets to the original partons.

1211 The choices of fixed parameters for the Pythia8 HV model are summarized in Table 6.1. A
 1212 detailed discussion of these parameters and their implications on the dark shower topology can be
 1213 found in Ref. [20]. The mass choices for the dark quark and the dark hadrons are also summarized
 1214 in Table 6.2.

1215 Note that the number of dark flavors differs from the Snowmass recommendation of $N_{f_D} = 4$.
 1216 This change is minimal in impact because R_{inv} is set explicitly (rather than allowing it to arise

Parameter	Value
HiddenValley:Ngauge	3
HiddenValley:alphaOrder	1
HiddenValley:Lambda	10.0
HiddenValley:nFlav	2
HiddenValley:spinFv	0
HiddenValley:FSR	on
HiddenValley:fragment	on
HiddenValley:pTminFSR	1.1
HiddenValley:probVector	0.58

Table 6.1: Fixed parameters in the Pythia8 HV model

Parameter	Value [GeV]
m_{π_D}	17
m_{ρ_D}	31.77
m_{q_D}	10

Table 6.2: Values for m_{dark}

naturally from the HV theory), and allows us to remain more comparable with the CMS semi-visible jets analysis and the ATLAS t-channel analysis.

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.

Samples are generated using MADGRAPH5 [65] version 2.9.9 interfaced to PYTHIA8.244P3 [63] for shower and hadronization with NNPDF23LO PDF [66] and the ATLAS A14 [67] to tune the underlying event data.

$m_{Z'}$ (GeV)	Cross section (fb)
2000	2.52e+2
2500	7.42e+1
3000	2.45e+1
3500	8.83e+0
4000	3.49e+0
5000	7.57e-1

Table 6.3: Mass points and cross sections of the SVJ search signal grid

1226

1227

Chapter 7: Machine Learning Tools

1228 **7.1 Introduction**

1229 The search for semi-visible jets presents an opportunity to use novel machine learning (ML)
1230 tools to uncover patterns in the behavior of dark QCD. The subtlety of the shower differences be-
1231 tween dark and SM QCD motivates a complex model that can accept high-dimensional low-level
1232 inputs to best understand key differences between signal and background correlations. Addition-
1233 ally, the large number of theory parameters which can be chosen arbitrarily and affect the shape of
1234 the dark QCD shower motivate exploring a data-driven machine learning approach, which could
1235 be sensitive to a wider variety of dark QCD behavior.

1236 To this end, two machine learning approaches are developed for this search, which are used
1237 in tandem. The first is a supervised ML method where the ML algorithm is built to maximize
1238 exclusion sensitivity to the specific generated SVJ signal models used in this analysis. Here, su-
1239 pervised refers to the use of full and correct labels for all events considered during model training,
1240 which necessitates training over simulated data. The second is a semi-supervised method, where
1241 training of the model is data-driven and labels are only partially provided during training. The
1242 semi-supervised ML algorithm broadens the discovery sensitivity of the search, and reduces the
1243 dependence on the exact theory parameters chosen for signal model simulation.

1244 The two different ML algorithms used in this approach will be explained in the following
1245 sections, along with their application in the SVJ analysis strategy.

1246 7.1.1 Particle Flow Network (Supervised)

1247 **Architecture Fundamentals**

1248 A Particle Flow Network (PFN) [68] architecture is selected for two reasons: *permutation in-*
1249 *variant input modeling* to best describe the events consisting of an unordered set of particles, and a
1250 *low-level input modeling* using tracks to take advantage of the available high-dimensional informa-
1251 tion to best exploit available correlations within the event. Permutation invariant input modeling
1252 is an architecture priority as ordered input modeling has been observed to bias the performance of
1253 low-level modeling tools as in [69]. Low-level input modeling is an architecture priority to capture
1254 the intricacies of dark QCD showers which may not express themselves in higher level variables,
1255 as explored in [64]. A comparison to a high-level *boosted decision tree* (BDT) is available in
1256 Appendix B.3.

1257 The PFN is used to model input events as an unordered set of tracks. Given the inherently
1258 unordered and variable-length nature of particles in an event, this choice of modeling as a *set*
1259 can enable the model to better learn the salient features of the dataset that enable a signal-to-
1260 background classification. Constructing the PFN involves the creation of new basis variables \oplus
1261 for each particle in the event. Permutation invariance is enforced by summing over the \oplus basis for
1262 every particle in the event to create a new permutation invariant latent space basis O . Finally the
1263 classifier F is a function of the sum over this new basis. The creation of the latent space basis O
1264 from M particles \vec{p} with d features each can be expressed as:

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

1265 where $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^l$ is a per particle mapping, with l being the dimension of the new basis O .
1266 Figure 7.1 gives a graphical representation of the use of summation in the PFN over per-particle
1267 information to create a permutation-invariant event representation.

1268 Figure 7.2 provides an annotated diagram of the PFN architecture as used in this analysis.

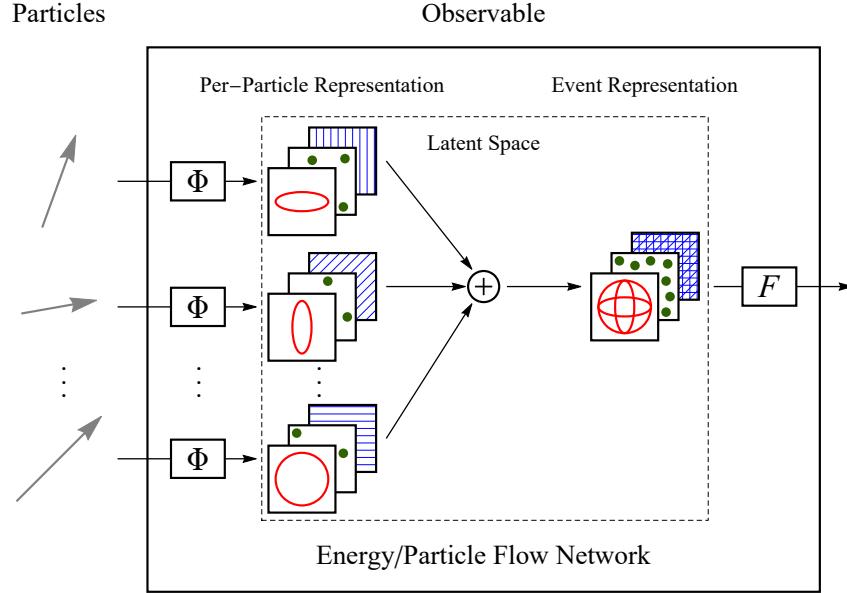


Figure 7.1: The Energy/Particle Flow Network concept, from Ref. [68].

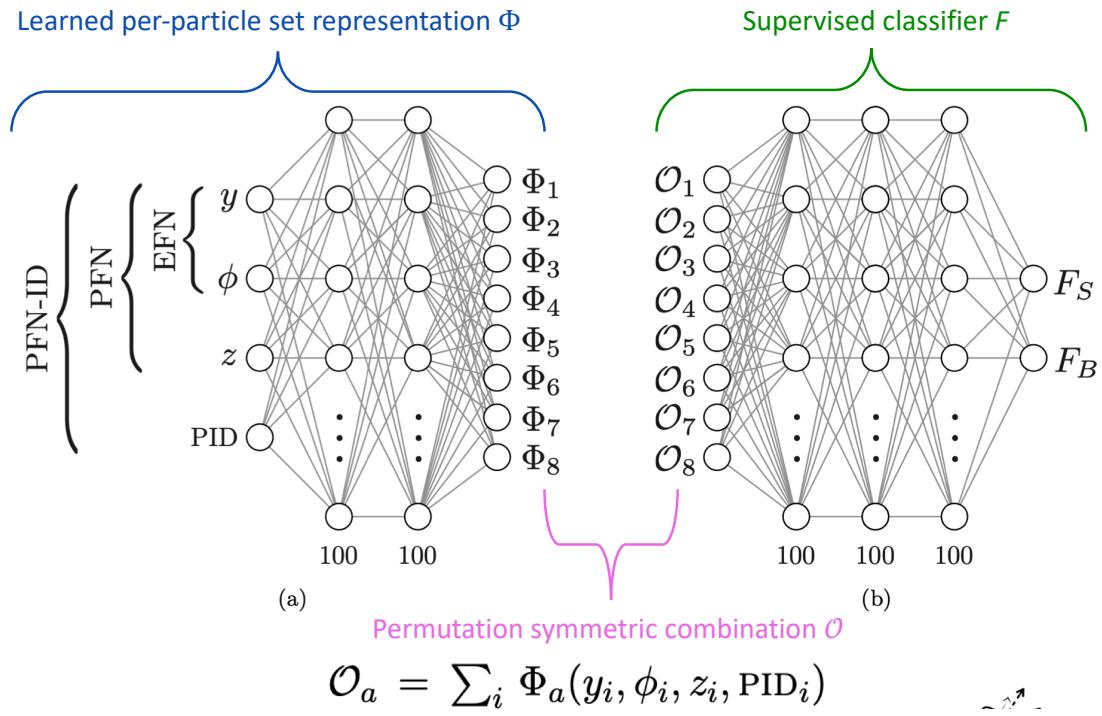


Figure 7.2: An annotated diagram of the PFN architecture. y and ϕ represent geometric information for the input particles, z represents energy information, and PID encompasses any other particle ID information in the input.

1269 **Input Modeling, Scaling, and Rotation**

1270 In this implementation, the particle input information comes from all tracks associated to the
1271 leading and subleading jets. The track association method is Ghost association, as discussed in
1272 Section 5.4.4. A single jet tagger strategy was also considered, but utilizing tracks from both
1273 leading jets creates a complete low-level picture of the event, which both focuses on the objects
1274 most likely to be associated to the decay of the dark quark (as will be justified in Chapter 8)
1275 and the relationship between those objects. If we consider the dijet topology of semi-visible jets
1276 as illustrated in Figure 7.3, the advantage of modeling both leading jets simultaneously becomes
1277 clear. In the semi-visible jet model presented in [64], E_T^{miss} in the event is expected to arise due to
1278 an imbalance in the number of visible tracks of the two jets associated to the dark quark decay.

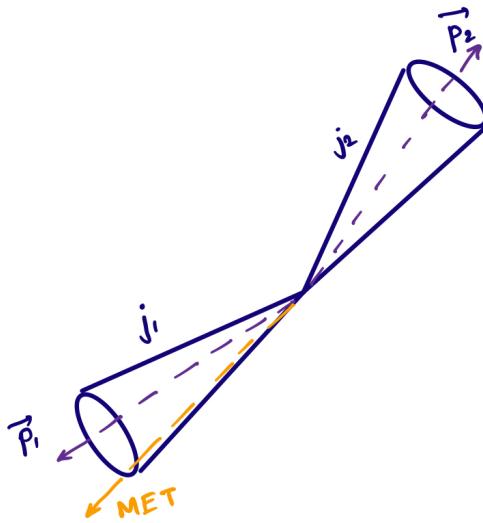


Figure 7.3: A illustration of the expected dijet behavior of semi-visible jets, where one jet is closely aligned with E_T^{miss} .

1279 Each track is described using six variables: the four-vector of the track (p_T , η , ϕ , E), and the
1280 track displacement parameters d_0 and z_0 , where d_0 measures displacement in the radial direction
1281 from the beamline and z_0 measures displacement along the beamline from the primary interac-
1282 tion point. Figure 7.4 illustrates these coordinates. Up to 80 tracks per jet are allowed, which is
1283 a threshold chosen to generally include all the tracks in the jet, which leads to maximal perfor-

1284 mance. Figure 7.5 shows the track multiplicity in the leading and subleading jet for the signal and
 1285 background samples used in training.

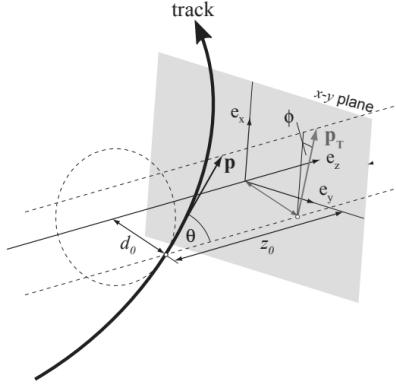


Figure 7.4: Illustration of track coordinates d_0 and z_0 .

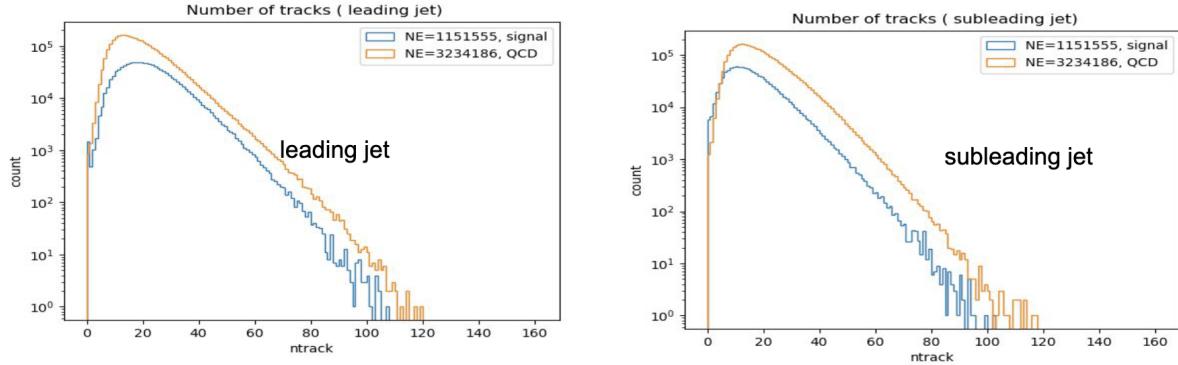


Figure 7.5: Distributions of the track multiplicity in the leading and subleading jets, comparing signal and background PFN training samples.

1286 These tracks (up to 160 total) are the input to the PFN. Referencing Equation 7.1, this corre-
 1287 sponds to $M = 160$ and $d = 6$. The two leading jets and their associated tracks are rotated so
 1288 that the center of the system is aligned with $(\eta, \phi) = (0, 0)$. Each track is normalized to its relative
 1289 fraction of the total dijet system energy and transverse momentum- this enforces agnosticism to the
 1290 total energy and transverse momentum of the event. The rotation and scaling are motivated by the
 1291 procedures described in [68] to improve the optimality of the PFN learning. Figure 7.6 illustrates
 1292 the rotation process.

1293 Finally, each of the 6 track variables is scaled so that its range is $[0,1]$. This is a common

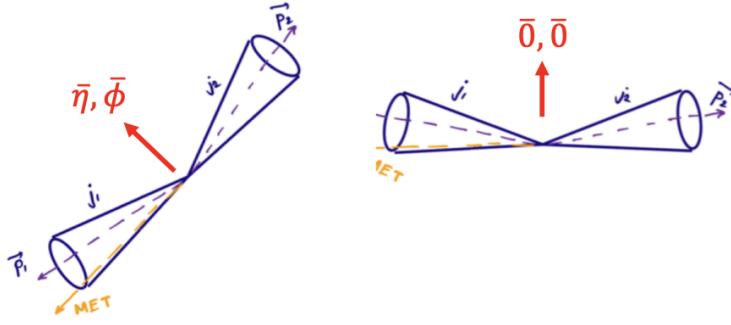


Figure 7.6: A diagram demonstrating how the two jet system is rotated in (ϕ, η) .

1294 preprocessing step that ensures the input data is bounded over a similar range, so that arbitrarily
 1295 large values don't develop an outsized impact on the model. Figure 7.7 show each of 6 track
 1296 variables before and after scaling and rotation have been applied, demonstrating the impact of
 1297 these procedures, as well as the track level similarities differences between the background SM
 1298 QCD processes and the signal SVJ processes. Figure 7.8 illustrates that the data is well modeled
 1299 by the MC at track level.

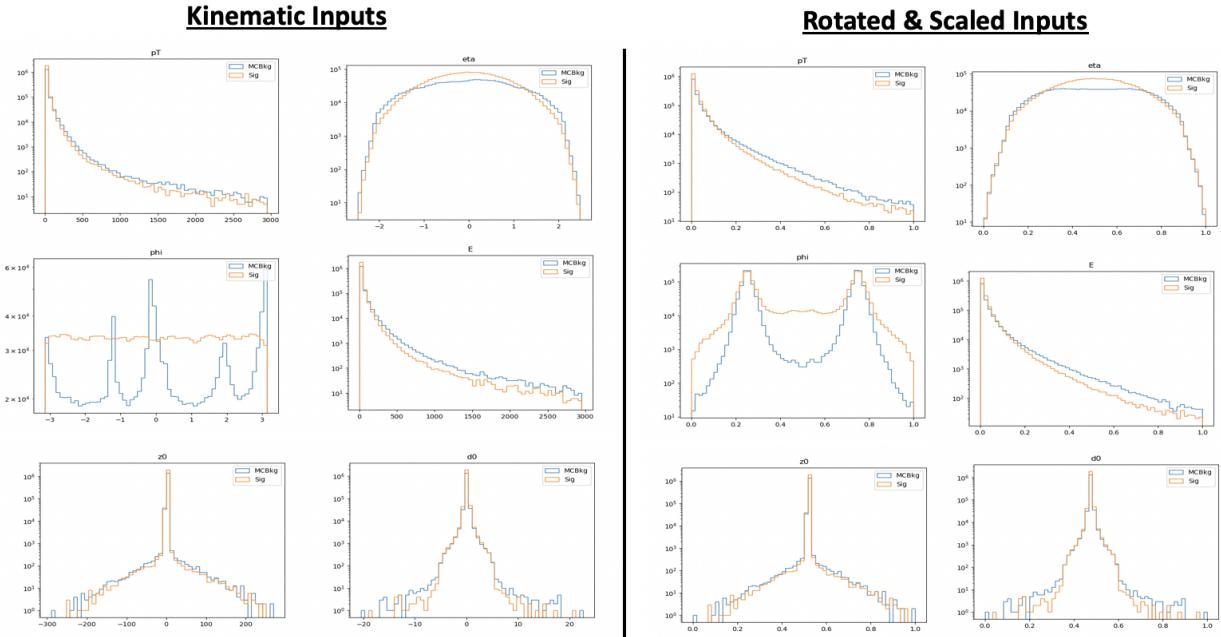


Figure 7.7: The 6 PFN track variables in background MC and signal MC. There are some differences between signal and background, but the track kinematics are largely similar.

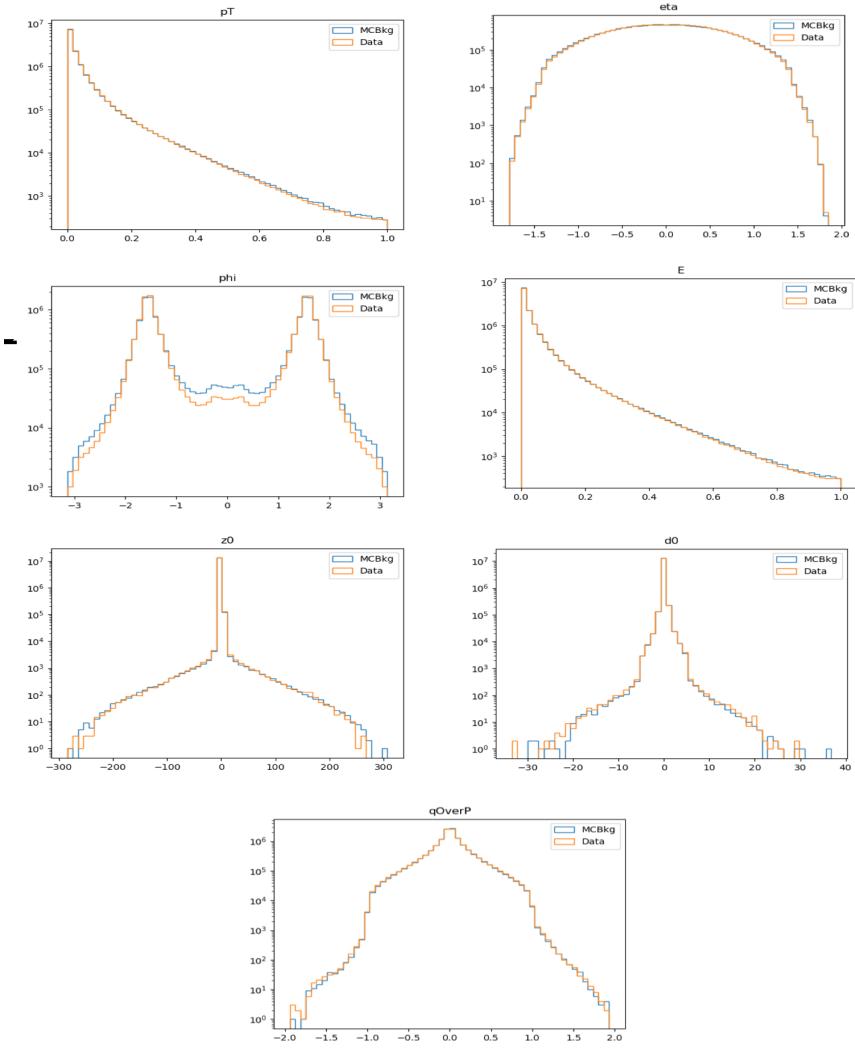


Figure 7.8: The 6 PFN track variables in data and background MC, after the scaling and rotation procedure is applied. There is excellent modeling of the data by the MC within the track variables. The slight discrepancy in the phi distribution is due to the modeling of dead TileCal cells by the QCD MC, which will be discussed in Chapter 8. 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.

1300 **Training**

1301 As seen in Figure 7.2, two separate architectures are defined and combined to do the super-
1302 vised training. The PFN uses a masking layer to suppress any zero-padded inputs, making the
1303 architecture length agnostic. The masking layer ignores any all-zero inputs in the summation layer.
1304 Additionally, The summation layer in the PFN enforces permutation invariance, so the network is
1305 unordered. The Φ network has 3 dense layers of dimensionality 75 with RELU activation, with
1306 27.5k trainable parameters and an output Φ latent space dimension of 64.

1307 The classifier F network similarly has 3 dense layers with 75 nodes with RELU activation, and
1308 a final softmax layer to determine the event-level classification with a categorical cross-entropy
1309 loss. The Adam optimizer is used with an initial learning rate of 0.001.

1310 The PFN is trained in a fully supervised way using SVJ signal MC and QCD MC events. Al-
1311 though several SM processes are expected to contaminate the SR (see Chapter 8), QCD is the dom-
1312 inant background. Training against a QCD-only sample is determined to produced better results
1313 than training on a more complete background - when training with a background which repre-
1314 sents samples that are more enriched in E_T^{miss} , the ability of the PFN to identify high E_T^{miss} signals
1315 is reduced. When training with a QCD-only background, there is greater contamination from
1316 E_T^{miss} enhanced backgrounds in the final SR - however the increased signal acceptance means that
1317 overall sensitivity is still higher with a QCD-only training. This can be seen in the comparison of
1318 output classifier distributions in Figure 7.9.

1319 500k events from both background and signal are used in training, where the signal is a com-
1320 bined file of all simulated signal points and the full QCD background which is sampled according
1321 to it's MC weights to produce the proper p_T input shape. A study was done to check the optimality
1322 of the inclusive signal model PFN as compared to one trained on high and low R_{inv} points sepa-
1323 rately, to better capture the differences in high and low E_T^{miss} across signals and backgrounds, but
1324 a small effect is found and the decision is taken to keep the inclusive model (Appendix B.2).

1325 The network is trained for 100 epochs. A train/test/validation split of 78%, 20%, and 2% is used
1326 for the final PFN training. Figure 7.10 shows the loss during training, which is stable and flattens

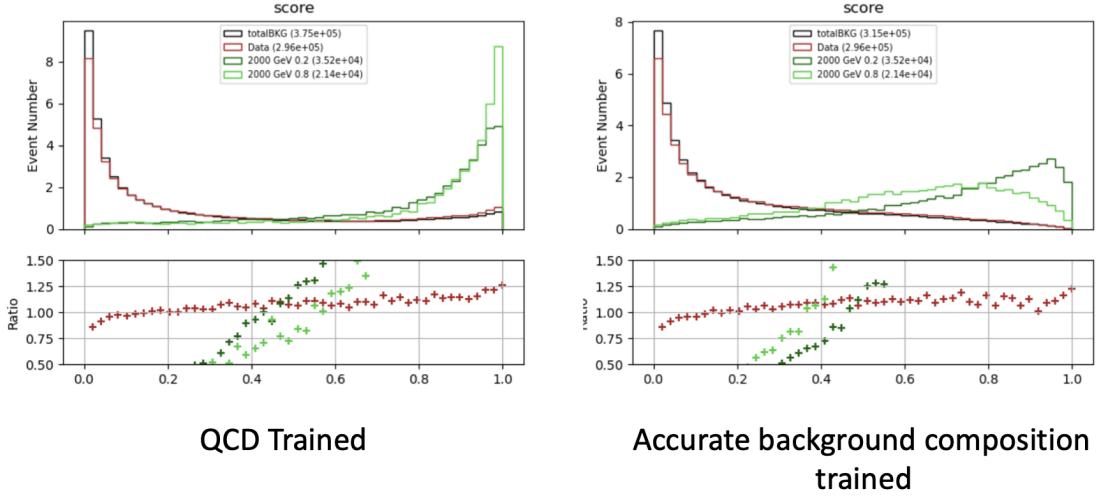


Figure 7.9: PFN score for background MC, data, and signal, comparing a PFN training on QCD-only vs all-background MC samples. The average AUC for the QCD-only training (left) is 0.93, while the average AUC for the mixed background training (right) is 0.84. The sensitivity estimate across the grid is better for the QCD-only training - from the distribution we can conclude that this is because the sensitivity to MET enhanced signals is greatly reduced.

1327 by the end of training, and the final evaluated losses that provide signal-background discrimination
 1328 over the test set.

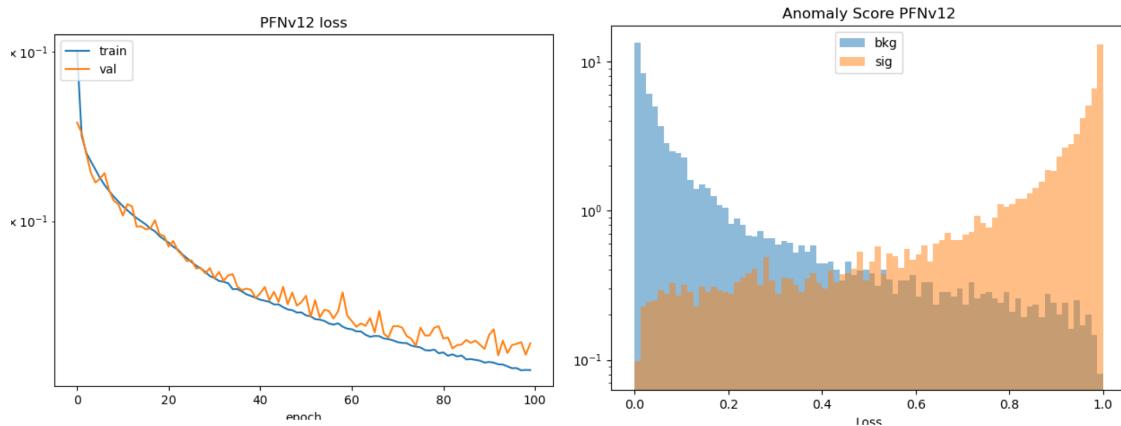


Figure 7.10: PFN architecture loss during training as a function of epoch (left) and the evaluated loss over the signal and background (right).

1329 Optimization studies were performed on the PFN, varying the number of training epochs, num-
 1330 ber of training events, batch size, learning rate, number of neurons, and dimension of the Φ space.
 1331 A summary of these studies is presented in Appendix B.2. The model presented here represents an
 1332 optimal choice across these parameters.

1333 **Performance**

1334 The performance of the PFN can be assessed via the area-under-curve (AUC) of the receiver
1335 operating characteristic (ROC) associated to evaluating the PFN on the test set of signal and back-
1336 ground events. Figure 7.11 shows the ROC curve of the PFN when classifying the QCD back-
1337 ground from the combined signal, with an AUC of 0.93. Figure 7.12 shows the AUC of the PFN
1338 across the SVJ signal grid, demonstrating strong discrimination capability even in the varying cor-
1339 ners of phase space. Figure 7.13 shows the output score distribution in two signals, data, and the
1340 total background MC. A selection of PFN score > 0.6 for all SR events is chosen to maximize
signal sensitivity across the grid.

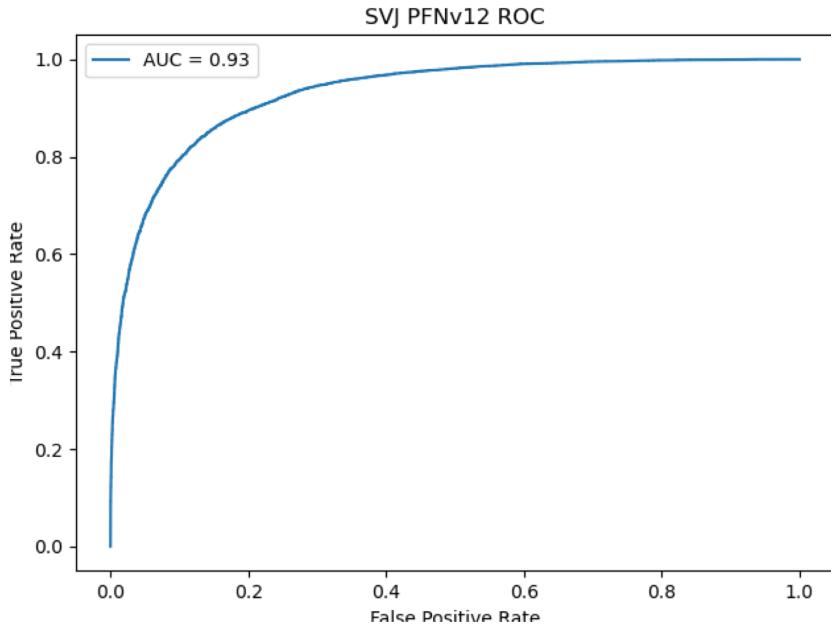


Figure 7.11: ROC the PFN score for combined signal (true positive) and QCD background (false positive).

1341

1342 Another supervised approach was studied using a BDT as the primary selection tool, trained
1343 over high-level variables describing each event. Studies comparing the PFN and BDT approaches
1344 are provided in Appendix B.3. Ultimately the low-level high-dimensional approach offered by the
1345 PFN was selected for its increased performance and lessened kinematic dependence.

1346 Appendix B shows more studies on the ML methods and comparisons of varying approaches.

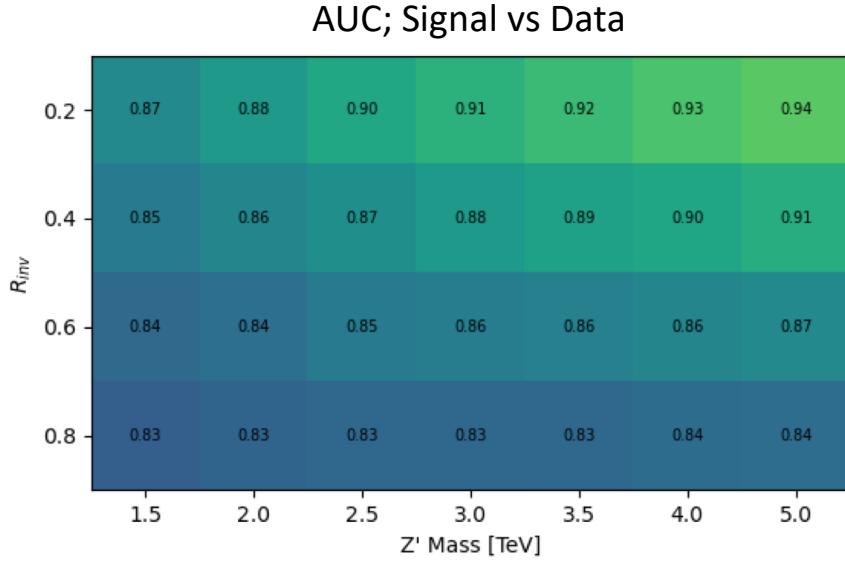


Figure 7.12: AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD-only training sample.

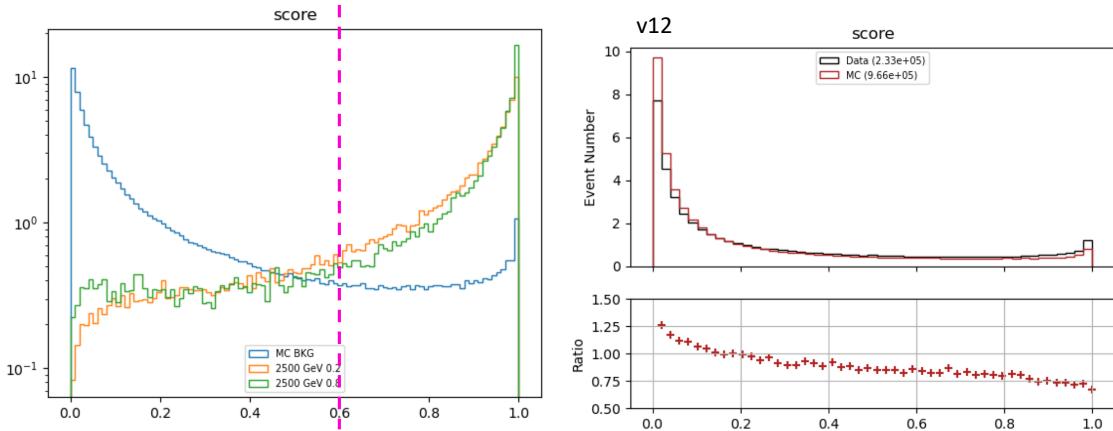


Figure 7.13: PFN score for two signals and the total background MC (top), and between data and MC (bottom). The difference between data and MC efficiency is minimal (< 5%).

1347 7.1.2 ANTELOPE (Semi-supervised)

1348 The semi-supervised analysis approach broadens the discovery sensitivity of the search through
1349 the use of semi-supervised ML, where training of the model is data-driven and labels are only
1350 partially provided during training. While broad sensitivity is a general key goal of LHC searches,
1351 it is particularly motivated in the case of dark QCD models, which can lead to widely varying
1352 topologies depending on the values of model parameters. In the case of SVJs, the R_{inv} fraction in
1353 the jet can dramatically vary the E_T^{miss} , shower shape, and other key features, making it difficult to
1354 find a single standard analysis variable that can distinguish all signal topologies from QCD.

1355 **Architecture Fundamentals**

1356 The model-independent search region of this analysis is implemented with a novel ML ap-
1357 proach that builds on the ANTELOPE architecture to construct a tool that is capable of performing
1358 low-level anomaly detection with permutation-invariant inputs. This tool, referred to as **ANomaly**
1359 **deTEction on particLe flOw latent sPacE (ANTELOPE)**, is a custom solution designed for this
1360 analysis.

1361 ANTELOPE uses the supervised signal vs. background training of the PFN network described
1362 in the previous section to generate a permutation invariant latent space that is representative of the
1363 original input variables, encodes the input events into these latent space variables O , and trains a
1364 *variational autoencoder* (VAE) over the events modeled as PFN latent space variables. A VAE
1365 is a common architecture used for anomaly detection and data-driven ML training. It has been
1366 used in previous ATLAS searched to model jet level information, such as the search presented
1367 in [70] using the recurrent architecture described in [69]. One of the limitations of a recurrent
1368 architecture is the need to order the low level inputs, which affects the performance of the tool. Jet
1369 constituent information is intrinsically unordered, and therefore a permutation invariant approach
1370 removes this element of arbitrary decision making from the modeling process. A visual example
1371 of the ANTELOPE inputs is given in Figure 7.14.

1372 The input to the model is the same 6 track variables for the leading 160 tracks of the leading

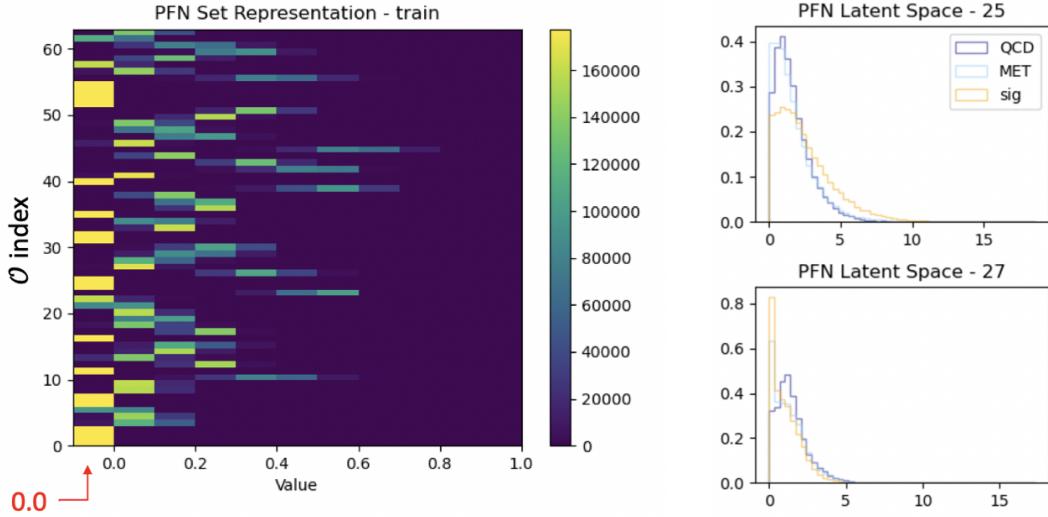


Figure 7.14: A visual representation of the 64 PFN latent space variables which create the input of the VAE component of ANTELOPE. The left shows a 2D histogram of the PFN latent space index (0-63) versus the value assumed by that index. The right shows 1D histograms of two particular PFN latent space variables.

1373 two jets, as presented in Section 7.1.1. The track information is encoded to the PFN Φ latent
 1374 space using the pre-trained Φ network (trained according to the steps outline in Section 7.1.1. The
 1375 resulting Φ basis is summed to created the fixed length symmetric representation O . The VAE is
 1376 then trained in an unsupervised way using inputs encoded to O from data events only. The VAE is
 1377 given no knowledge of the signal model during training. It is able to perform anomaly detection
 1378 through an encoder stage which does a lossy compression on the input to a lower-dimensional
 1379 latent space, and a decoder stage that samples from that latent space and generates an output of the
 1380 original dimensionality. By using the reconstruction error as a loss, this process enables the VAE
 1381 to develop a knowledge of the underlying data structure, thereby isolating new out-of-distribution
 1382 events by their high reconstruction error. This strategy is semi-supervised because the tool has
 1383 some knowledge of correct labels (eg. through the PFN latent space embedding) but is followed
 1384 by a data-driven unsupervised stage.

1385 Figure 7.15 provides a diagram of the ANTELOPE architecture.

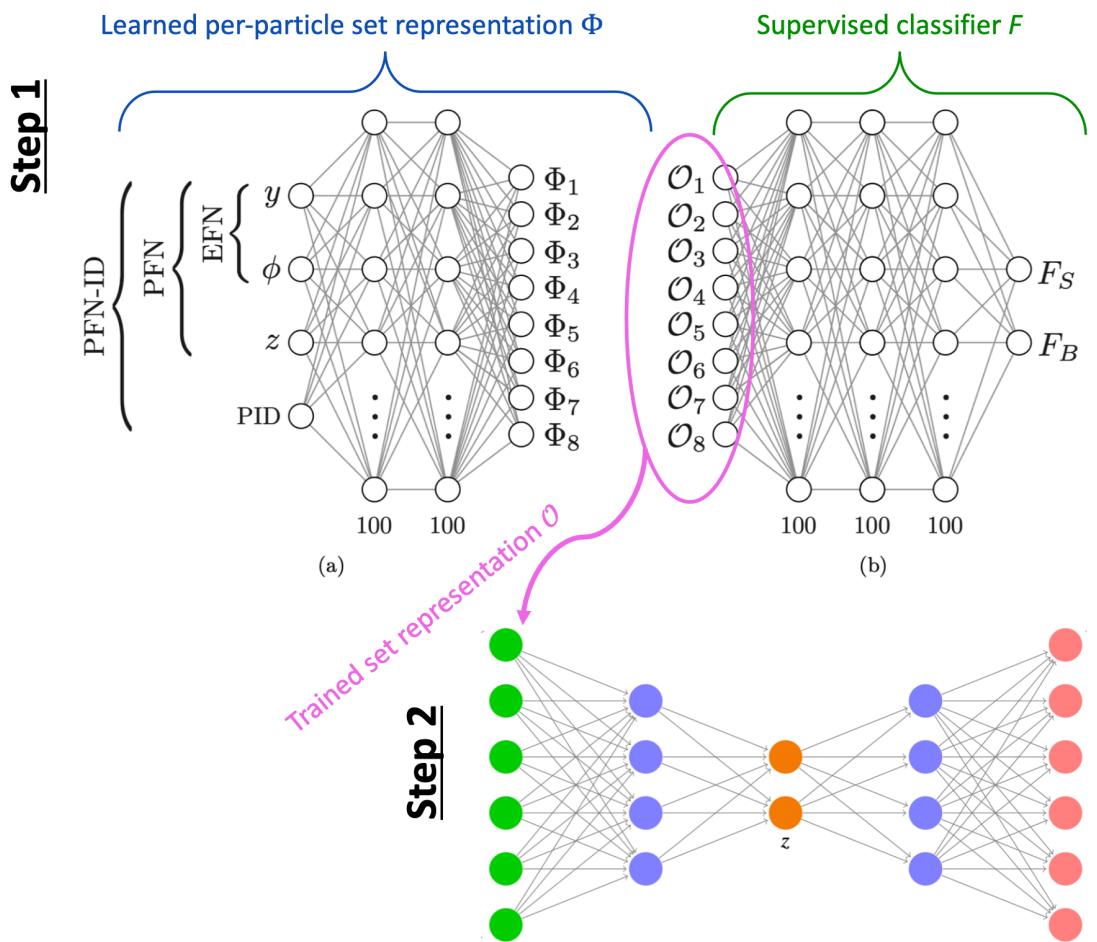


Figure 7.15: An annotated diagram of the ANTELOPE architecture.

1386 **Training**

1387 The VAE stage of the ANTELOPE network is trained directly over a subset of data events
1388 at preselection (6.7 million available, 500,000 used, with a 80% / 20% training/test split). The
1389 input dimensionality of the VAE has to match the encoded Φ dimension of the PFN, in this case
1390 64. The encoder has an encoding layer that brings the dimensionality to 32, and a final layer that
1391 compresses to the latent space dimension of 12. The network is trained for 50 epochs, with a
1392 learning rate of 0.00001. The loss \mathcal{L} is the sum of two terms, the mean-squared error (MSE) of
1393 input-output reconstruction, and the Kullback-Leibler divergence (KLD).

$$\mathcal{L} = \sum_i L_i = \sum_i |\Phi_i^2 - \Phi'_{\ell i}|^2 + \lambda D_{\text{KL}} \quad (7.2)$$

1394 As the PFN inputs are sufficiently normalized to remove any spurious information from train-
1395 ing, no additional normalization is applied to the PFN encoded inputs. The final ANTELOPE score
1396 used in the analysis is produced by applying a log + sigmoid transformation function to the total
1397 evaluated loss \mathcal{L} .

Figure 7.16 shows the loss during training.

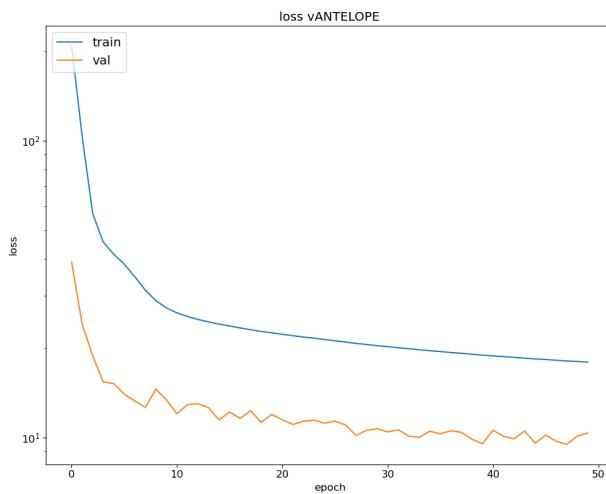


Figure 7.16: ANTELOPE architecture loss during training as a function of epoch.

1399 **Performance**

1400 As with the PFN, the ANTELOPE performance is assessed via the area-under-curve (AUC) of
1401 the receiver operating characteristic (ROC) associated to evaluating the ANTELOPE on the test
1402 set of signal and background events. Figure 7.17 shows the output score distribution in data and
1403 total background MC, showing a very flat ratio and motivating the use of MC for studies of the
1404 ANTELOPE score.

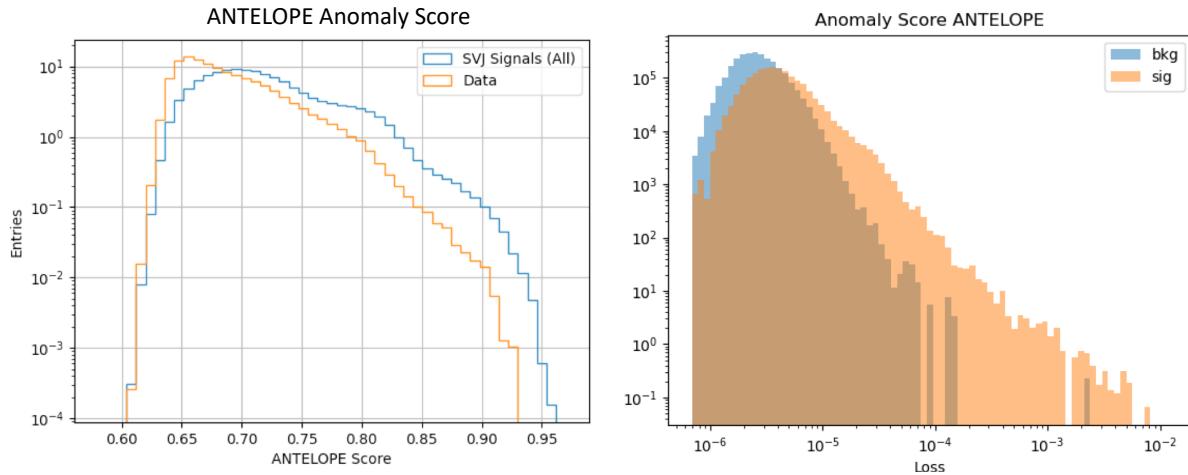


Figure 7.17: ANTELOPE score distribution comparing data and the total background MC (left), with good agreement observed between data and simulated background, and comparing all background MC to signals (right), revealing good discrimination power.

1405 Figure 7.18 shows the AUC of the ANTELOPE across the SVJ signal grid, demonstrating
1406 strong discrimination capability even in the varying corners of phase space. Compared to the
1407 supervised PFN method, the ANTELOPE is not as performant (as expected due to the absence of
1408 signal model in training). However, a selection on events with high ANTELOPE score nonetheless
1409 provides a 10-40% increase in signal significance by removing background and isolating the long
1410 tail of anomalous events.

1411 **Model Independence** The unsupervised component of training the ANTELOPE network is ex-
1412 pected to give it a more generalized sensitivity to new physics with E_T^{miss} and jet activity, beyond
1413 the scope of the SVJ grid. To assess this, alternative signal models are evaluated with the trained

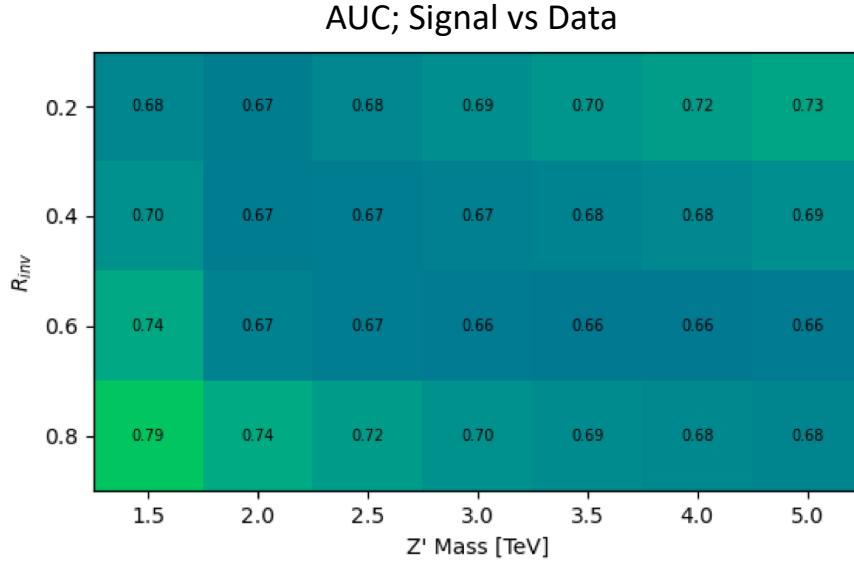


Figure 7.18: AUC from the ANTELOPE score for each signal in the SVJ grid.

1414 ANTELOPE network, as optimized for the SVJ grid, and their sensitivity in the analysis selection
1415 is evaluated.

1416 The following alternate signal models were considered:

1417 • $Z' \rightarrow t\bar{t}$

1418 • $W' \rightarrow WZ$

1419 • Gluino pair production \rightarrow R-hadron + LSP (E_T^{miss}) with gluino masses 2000/3000 GeV, LSP
1420 mass 100 GeV, and lifetime 0.03 ns (LSP = *lightest supersymmetric particle*)

1421 • Emerging jets s-channel with mass 1000 GeV and lifetime 1ns

1422 Figure 7.19 shows the distribution of these signals in several key analysis variables, namely
1423 E_T^{miss} , the PFN score, and the ANTELOPE score. This comparison reveals that ANTELOPE is
1424 sensitive to E_T^{miss} in the event; it classifies signals with no real E_T^{miss} , like the all-hadronic Z'
1425 and W' decays (given our imposed lepton veto) as data-like, but the distributions for signals with
1426 E_T^{miss} such as SVJs, R-hadrons, and emerging jets have distributions with higher anomaly score
1427 tails.

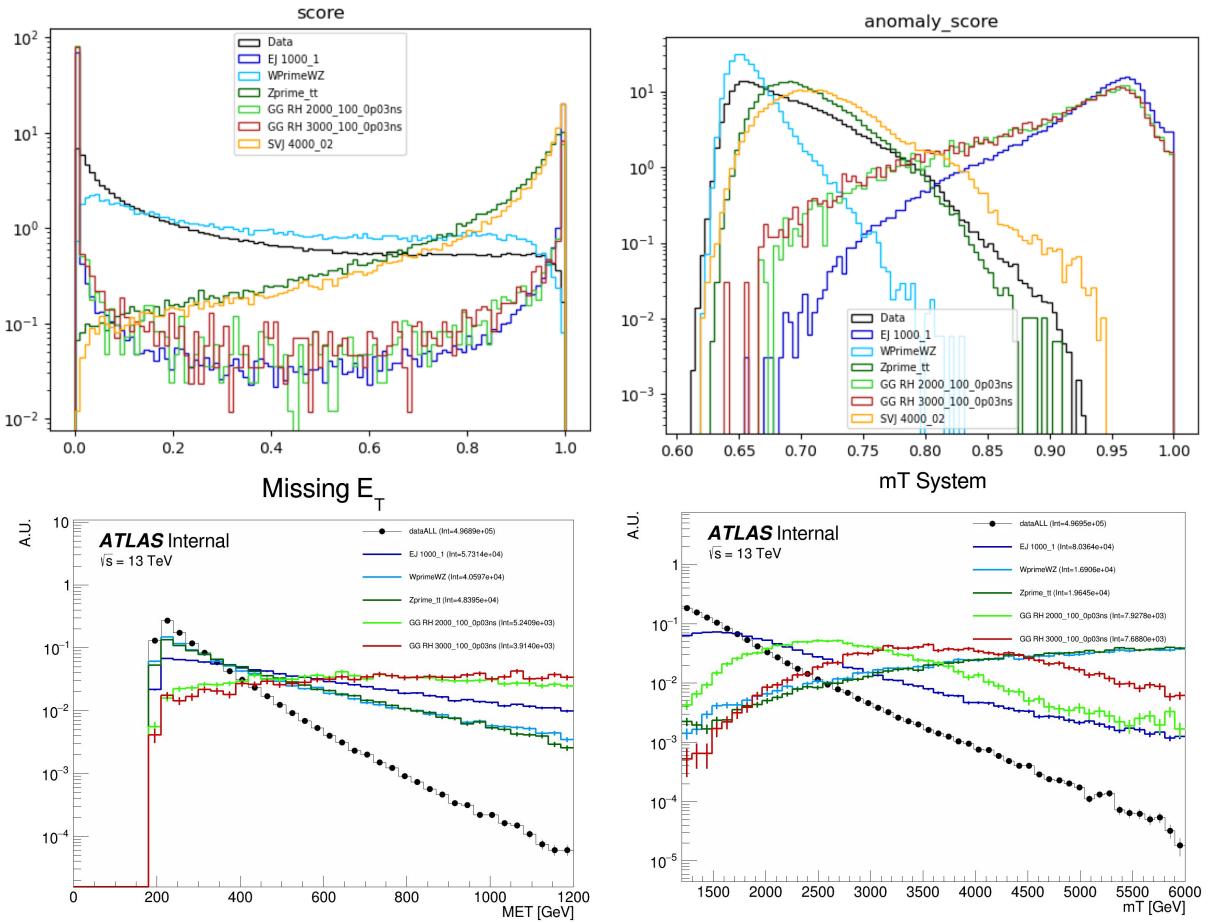


Figure 7.19: Comparing data and the alternate signal models for the PFN score (top left), ANTELOPE score (top right), E_T^{miss} (bottom left), and m_T (bottom right). The emerging jet signal is an example of the gain of the model-independent ANTELOPE approach, where it has a bimodal shape in PFN score but is clearly tagged as anomalous by ANTELOPE.

1428 Figure 7.20 shows a comparison of the sensitivity of the PFN and ANTELOPE regions across
 1429 a variety of signals, including the combined SVJ signal used to train the PFN. The benefit of the
 1430 unsupervised stage of ANTELOPE in enhancing model independence is clearly seen through the
 1431 boost in performance for other signal models, namely the gluino and emerging jet signals, which
 1432 have more E_T^{miss} than the W' and Z' signals (all-hadronic) that were also tested. As commented
 1433 above, the PFN outperforms ANTELOPE as expected, because it was designed explicitly for the
 1434 task of classifying SVJs from background, demonstrating the power of supervised learning for the
 1435 model-specific approach.

$\text{sig eff} / \sqrt{\text{bkg eff}}$ for respective score cut (0.6 PFN, 0.7 ANTELOPE)

	EJ 1000	WprimeWZ	Zprime tt	GG RH 2000	GG RH 3000	SVJ
PFN	0.57	0.65	1.92	0.30	0.32	1.97
ANTELOPE	1.73	0.07	0.84	1.72	1.72	1.13

Figure 7.20: Comparing data and the alternate signal models in terms of sensitivity (S/\sqrt{B}) for the PFN and ANTELOPE tools, applying the selection that is used in the analysis. The ANTELOPE network is found to provide significant added sensitivity to alternate signals such as the gluino→R-hadron and emerging jets, which have higher E_T^{miss} than the SVJs.

1436 Studies on the ANTELOPE architecture and comparisons to other methods can be found in
 1437 Appendix B.1.

1438

1439

Chapter 8: Analysis Strategy

1440 **8.1 Event Selection**

1441

Conclusion or Epilogue

1442 Use this page for your epilogue or conclusion if applicable; please use only one of the titles
1443 for this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1444 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1445 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1446 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1447 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1448 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1449 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1450 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1451 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1452 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1453 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1454 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1455 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1456 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1457 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1458 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1459 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1460 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1461 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete

1462 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1463 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1464 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1465 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1466 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1467 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1468 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1469 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1470 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1471 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1472 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1473 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1474 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1475 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1476 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1477 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1478 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1479 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1480 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1481 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1482 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1483 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1484 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1485 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1486 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1487 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1488 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,

1489 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1490 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1491 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1492 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1493 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1494 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1495 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1496 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1497 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1498 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1499 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1500 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1501 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1502 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1503 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1504 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1505 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1506 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1507 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1508 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1509 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1510 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1511 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1512 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1513 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1514 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1515 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.

1516 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1517 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1518 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1519 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1520 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1521 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1522 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1523 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1524 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1525 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1526 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1527 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1528 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1529 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1530 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1531 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1532 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for
1533 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if
1534 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this
1535 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.
1536 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please
1537 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your
1538 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,
1539 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one
1540 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or
1541 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete
1542 it.

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 2π Decay of the K_2^0 Meson”. In: *Phys. Rev. Lett.* 13 (1964), pp. 138–140.
- [6] J. E. Augustin et al. “Discovery of a Narrow Resonance in e^+e^- Annihilation”. In: *Phys. Rev. Lett.* 33 (1974), pp. 1406–1408.
- [7] J. J. Aubert et al. “Experimental Observation of a Heavy Particle J ”. In: *Phys. Rev. Lett.* 33 (1974), pp. 1404–1406.
- [8] Martin L. Perl et al. “Evidence for Anomalous Lepton Production in e+ - e- Annihilation”. In: *Phys. Rev. Lett.* 35 (1975), pp. 1489–1492.
- [9] 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.
- [10] 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.
- [11] S. Abachi et al. “Observation of the top quark”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. arXiv: hep-ex/9503003.
- [12] K. Kodama et al. “Observation of tau neutrino interactions”. In: *Phys. Lett. B* 504 (2001), pp. 218–224. arXiv: hep-ex/0012035.
- [13] G. Arnison et al. “Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider”. In: *Phys. Lett. B* 126 (1983), pp. 398–410.

- 1570 [14] P. Bagnaia et al. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ Collider”. In: *Phys. Lett. B* 129
1571 (1983), pp. 130–140.
- 1572 [15] Serguei Chatrchyan et al. “Observation of a New Boson at a Mass of 125 GeV with the
1573 CMS Experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 30–61. arXiv: 1207.7235
1574 [hep-ex].
- 1575 [16] Georges Aad et al. “Observation of a new particle in the search for the Standard Model
1576 Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29.
1577 arXiv: 1207.7214 [hep-ex].
- 1578 [17] K. G. Begeman, A. H. Broeils, and R. H. Sanders. “Extended rotation curves of spiral galaxies:
1579 Dark haloes and modified dynamics”. In: *Mon. Not. Roy. Astron. Soc.* 249 (1991), p. 523.
- 1580 [18] Y. Ashie et al. “Evidence for an oscillatory signature in atmospheric neutrino oscillation”.
1581 In: *Phys. Rev. Lett.* 93 (2004), p. 101801. arXiv: hep-ex/0404034.
- 1582 [19] C. Abel et al. “Measurement of the Permanent Electric Dipole Moment of the Neutron”. In:
1583 *Phys. Rev. Lett.* 124.8 (2020), p. 081803. arXiv: 2001.11966 [hep-ex].
- 1584 [20] Guillaume Albouy et al. “Theory, phenomenology, and experimental avenues for dark showers:
1585 a Snowmass 2021 report”. In: *The European Physical Journal C* 82.12 (Dec. 2022).
- 1586 [21] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08
1587 (2008), S08001.
- 1588 [22] “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008). Also
1589 published by CERN Geneva in 2010, S08003.
- 1590 [23] “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008),
1591 S08004.
- 1592 [24] “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008),
1593 S08002.
- 1594 [25] “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005.
- 1595 [26] Ana Lopes and Melissa Loyse Perrey. *FAQ-LHC The guide*. 2022.
- 1596 [27] Esma Mobs. “The CERN accelerator complex in 2019. Complexe des accélérateurs du
1597 CERN en 2019”. In: (2019). General Photo.
- 1598 [28] *Pulling together: Super Conducting electromagnets*. <https://home.cern/science/engineering/pulling-together-superconducting-electromagnets>.
1599 Accessed: 2024-01-05.

- 1601 [29] *The High-Luminosity LHC*. <https://voisins.web.cern.ch/en/high-luminosity-lhc-hl-lhc>. Accessed: 2024-01-05.
- 1602
- 1603 [30] Aad G., et al. (ATLAS Collaboration and CMS Collaboration). “Combined Measurement of
1604 the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS
1605 Experiments”. In: *Phys. Rev. Lett.* 114 (19 2015), p. 191803.
- 1606 [31] O. Aberle et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design re-*
1607 *port*. CERN Yellow Reports: Monographs. Geneva: CERN, 2020.
- 1608 [32] The ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”.
1609 In: *Journal of Instrumentation* 3.08 (2008), S08003.
- 1610 [33] G Aad, B Abbott, and ATLAS Collaboration. “Performance of the reconstruction of large
1611 impact parameter tracks in the inner detector of ATLAS”. In: *Eur. Phys. J. C Part. Fields*
1612 83.11 (Nov. 2023).
- 1613 [34] Joao Pequenao. *Computer Generated image of the ATLAS calorimeter*. 2008.
- 1614 [35] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. AT-
1615 LAS. Geneva: CERN, 1996.
- 1616 [36] H A Gordon. “Liquid argon calorimetry for the SSC”. In: () .
- 1617 [37] Henric Wilkens and (on behalf of the ATLAS LArg Collaboration). “The ATLAS Liquid
1618 Argon calorimeter: An overview”. In: *Journal of Physics: Conference Series* 160.1 (2009),
1619 p. 012043.
- 1620 [38] *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*. Tech.
1621 rep. Geneva: CERN, 2017.
- 1622 [39] “Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer”. In:
1623 () .
- 1624 [40] L Pontecorvo. “The ATLAS Muon Spectrometer”. In: (2004). revised version number 1
1625 submitted on 2003-07-27 16:31:16.
- 1626 [41] *ATLAS magnet system: Technical Design Report, 1*. Technical design report. ATLAS. Geneva:
1627 CERN, 1997.
- 1628 [42] Joao Pequenao. “Event Cross Section in a computer generated image of the ATLAS detec-
1629 tor.” 2008.
- 1630 [43] ATLAS Collaboration. “ATLAS Experiment Implements Heterogeneous Particle Recon-
1631 struction with Intel oneAPI Tools”. General Photo. 2023.

- 1632 [44] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS
 1633 detector using the 2015–2017 LHC proton-proton collision data”. In: *Journal of Instrumentation*
 1634 14.12 (2019), P12006.
- 1635 [45] Chiara Deponte. “Studies on the properties of non-prompt photons at the ATLAS experi-
 1636 ment”. Presented 16 Aug 2022. Technische Universitaet Dortmund (DE), 2022.
- [46] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in pro-
 ton–proton collision data at
 \sqrt{s}
 1637 s = 13 TeV”. In: *The European Physical Journal C* 76.5 (2016).
- 1638 [47] Sebastien Rettie. *Muon identification and performance in the ATLAS experiment*. Tech. rep.
 1639 Geneva: CERN, 2018.
- 1640 [48] B. R. Webber. *Fragmentation and Hadronization*. 1999. arXiv: hep-ph/9912292 [hep-ph].
- 1641 [49] Eric M. Metodiev. *The Fractal Lives of Jets* | Eric M. Metodiev — ericmetodiev.com. <https://www.ericmetodiev.com/post/jetformation/>. 2019, note = [Accessed 18-
 1642 05-2024],
 1643
- 1644 [50] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti-k_Tjet clustering algorithm”.
 1645 In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), 063–063.
- 1646 [51] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet user manual: (for version
 1647 3.0.2)”. In: *The European Physical Journal C* 72.3 (Mar. 2012).
- 1648 [52] Steven Schramm. *ATLAS Jet Reconstruction, Calibration, and Tagging of Lorentz-boosted
 1649 Objects*. Tech. rep. Geneva: CERN, 2017.
- 1650 [53] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its per-
 1651 formance in LHC Run 1”. In: *The European Physical Journal C* 77.7 (July 2017).
- 1652 [54] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the
 1653 ATLAS Detector”. In: *The European Physical Journal C* 77.7 (July 2017).
- 1654 [55] Stephen D. Ellis and Davison E. Soper. “Successive combination jet algorithm for hadron
 1655 collisions”. In: *Physical Review D* 48.7 (Oct. 1993), 3160–3166.
- 1656 [56] M. Wobisch and T. Wengler. *Hadronization Corrections to Jet Cross Sections in Deep-
 1657 Inelastic Scattering*. 1999. arXiv: hep-ph/9907280 [hep-ph].
- 1658 [57] Gavin P Salam and Grégory Soyez. “A practical seedless infrared-safe cone jet algorithm”.
 1659 In: *Journal of High Energy Physics* 2007.05 (May 2007), 086–086.

- 1660 [58] Gavin P. Salam. “Towards jetography”. In: *The European Physical Journal C* 67.3–4 (May
1661 2010), 637–686.
- 1662 [59] *A Monte Carlo study of track association to jets for b-tagging*. Tech. rep. Geneva: CERN,
1663 2021.
- 1664 [60] *Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector*. Tech. rep.
1665 All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/P>
1666 PHYS-PUB-2014-013. Geneva: CERN, 2014.
- 1667 [61] ATLAS Collaboration. “Performance of missing transverse momentum reconstruction with
1668 the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C*
1669 78.11 (2018), p. 903. arXiv: 1802.08168.
- 1670 [62] GEANT4 Collaboration, S. Agostinelli, et al. “GEANT4 – a simulation toolkit”. In: *Nucl.
1671 Instrum. Meth. A* 506 (2003), p. 250.
- 1672 [63] Christian Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*.
1673 2022. arXiv: 2203.11601 [hep-ph].
- 1674 [64] Timothy Cohen et al. “LHC searches for dark sector showers”. In: *Journal of High Energy
1675 Physics* 2017.11 (Nov. 2017).
- 1676 [65] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differ-
1677 ential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014),
1678 p. 079. arXiv: 1405.0301 [hep-ph].
- 1679 [66] Jon Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *J. Phys. G* 43
1680 (2016), p. 023001. arXiv: 1510.03865 [hep-ph].
- 1681 [67] Peter Skands, Stefano Carrazza, and Juan Rojo. “Tuning PYTHIA 8.1: the Monash 2013
1682 Tune”. In: *Eur. Phys. J. C* 74.8 (2014), p. 3024. arXiv: 1404.5630 [hep-ph].
- 1683 [68] Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler. “Energy flow networks: deep sets
1684 for particle jets”. In: *Journal of High Energy Physics* 2019.1 (2019).
- 1685 [69] A. Kahn et al. “Anomalous jet identification via sequence modeling”. In: *Journal of Instru-
1686 mentation* 16.08 (Aug. 2021), P08012.
- 1687 [70] Georges Aad et al. “Anomaly detection search for new resonances decaying into a Higgs
1688 boson and a generic new particle X in hadronic final states using $\sqrt{s} = 13$ TeV pp collisions
1689 with the ATLAS detector”. In: *Phys. Rev. D* 108 (2023), p. 052009. arXiv: 2306.03637
1690 [hep-ex].

Appendix A: Trigger Studies

1693 Both the lowest unprescaled E_T^{miss} and single small-R jet triggers were considered for this
 1694 analysis. The E_T^{miss} trigger is observed to have higher efficiency for low mass, high R_{inv} points,
 1695 while the single small-R trigger favors high mass, low R_{inv} points. Figure A.1 shows the yields
 1696 and signal efficiencies across the grid for both these strategies.

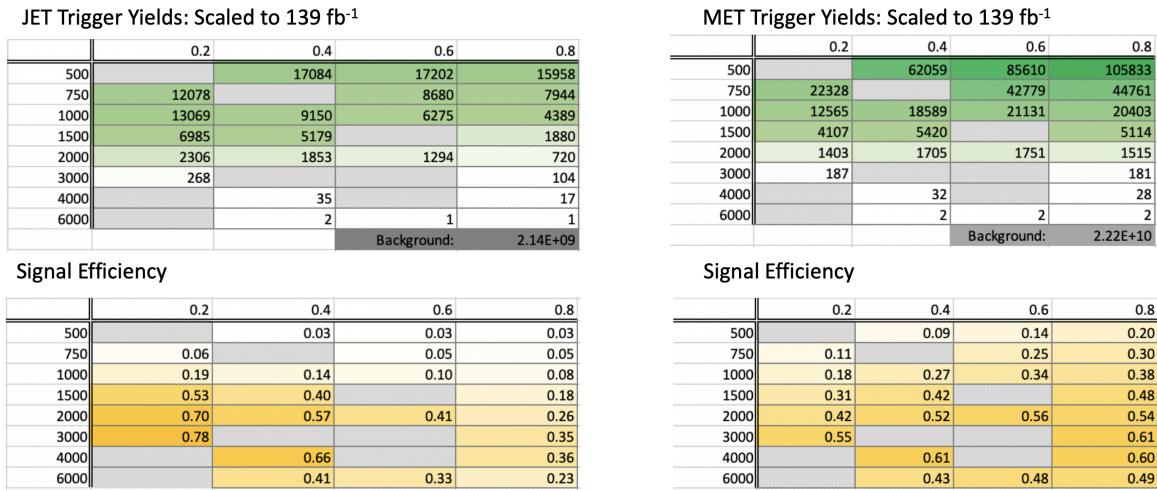


Figure A.1: Trigger yield and efficiency for both the MET trigger and small-R jet trigger approach. Each entry represent a signal point, labelled by the Z' mass and the R_{inv} fraction.

1697 The cross section is higher for the lower Z' mass signal points. As a result, our sensitivity
 1698 to these points and ability to set limits on them is already naturally enhanced. Figure A.2 shows
 1699 the factor of improvement in the inclusive S/\sqrt{B} using each trigger strategy. A cross-check was
 1700 also done calculating S/\sqrt{B} in windows around the mT mass. The results of this cross-check
 1701 confirmed the results shown in Figure A.2.

1702 Figure A.3 shows the ratio of S/\sqrt{B} across the signal grid for the jet trigger compared to the
 1703 E_T^{miss} trigger.

1704 This plot makes clear that the E_T^{miss} trigger favors the low mass, high R_{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^{miss} trigger selection.

1705 the single jet trigger favors the high mass, low R_{inv} signal points. A number of considerations led
1706 us to selecting the single jet trigger. First, Figure A.2 illustrates that the sensitivity enhancement is
1707 greater for signals favored by the jet trigger than signals favored by the E_T^{miss} trigger. This indicates
1708 we have more to gain from focusing on the region of our signal grid where the jet trigger is most
1709 efficient. We combined this with the observation that the jet trigger is beneficial for high mass
1710 points where the cross-section is lower and we need to maximize sensitivity to hope to set limits
1711 on these points.

1712 Second, there are a number of analysis variables at our disposal which are more discriminant for
1713 high R_{inv} signals than for low R_{inv} signals. A collection of these variables are shown in Figure A.4.
1714 Given that we had multiple avenues available to enhance sensitivity to low R_{inv} signals through
1715 analysis strategy cuts, we opted for a trigger strategy that benefitted the low R_{inv} signal points.
1716 Additionally, we know E_T^{miss} is highly correlated with these discriminant analysis variables, and a
1717 E_T^{miss} trigger would reduce the effectiveness of these variables. Not implementing the E_T^{miss} trigger
1718 allows us to explore other ways of leveraging E_T^{miss} in the analysis using a cut that is more finely
1719 tuned to the specifics of the signal model than the E_T^{miss} trigger is.

1720 Third, we know that fitting the high R_{inv} mass points would be challenging given their very
1721 broad shapes in the key analysis variable m_T . Given the analysis decision to do a search for
1722 resonant features in m_T , we chose to design a search that emphasized maximizing sensitivity
1723 to the signal points that we were mostly to appear as a resonant feature in m_T . The shape of
1724 m_T and our reduced ability to set limits on high R_{inv} points is demonstrated in the body of this
1725 note. Another search strategy may be more optimal for setting limits on high R_{inv} semi-visible jet
1726 signals, but it is currently outside the scope of this analysis.

1727 A jet or E_T^{miss} trigger was also considered, but abandoned due to concerns about sculpting
1728 features in the smoothly falling E_T^{miss} or leading jet p_T distributions, as shown in Figure A.5.

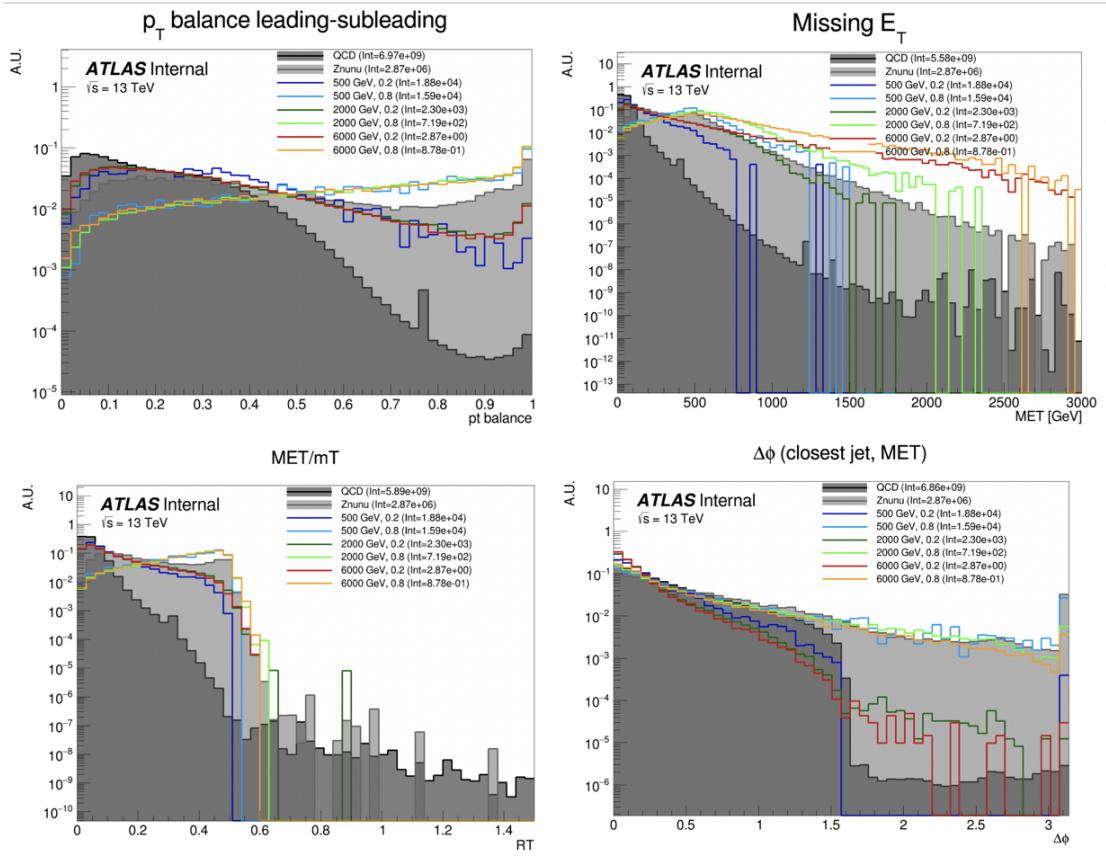


Figure A.4: Analysis variables where high R_{inv} signals a clearly distinct from background and low R_{inv} variables. On the contrary, leading jet p_T is one of the only variables where low R_{inv} signals are distinct from background.

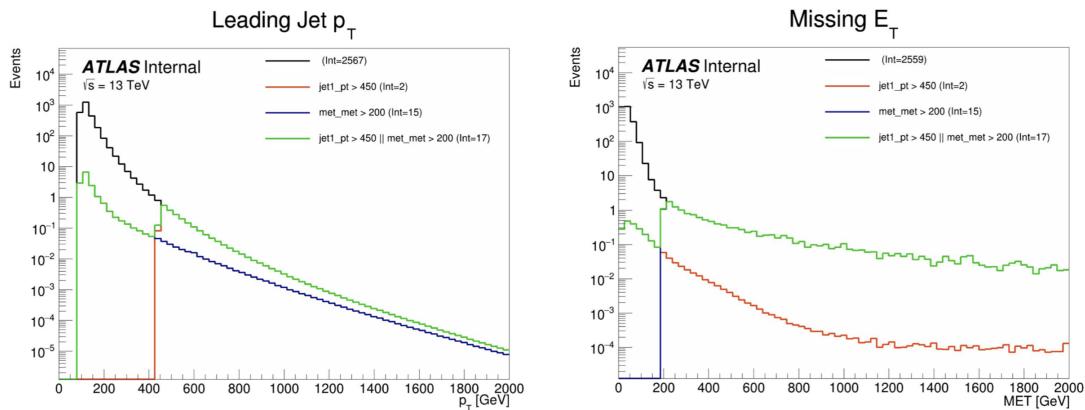


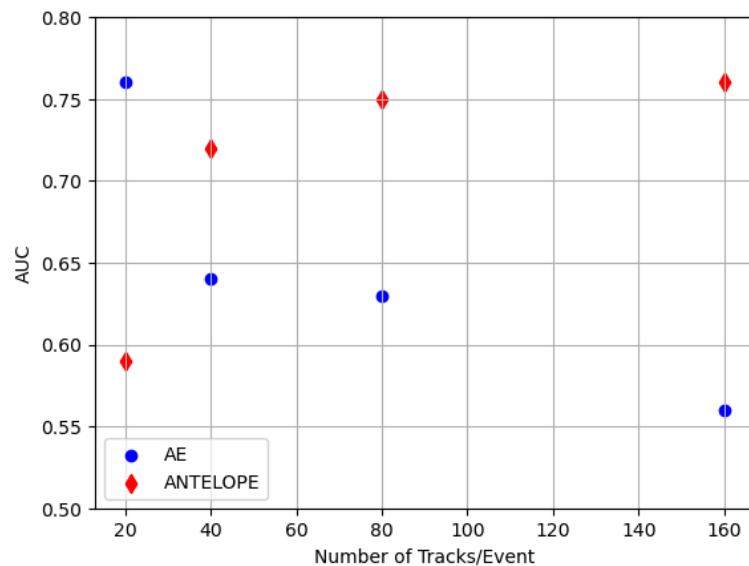
Figure A.5: OR of jet and E_T^{miss} triggers.

Appendix B: Machine Learning Approaches

1731 B.1 Unsupervised: AE vs. ANTELOPE

1732 To understand the benefits of the semi-supervised ANTELOPE approach, we study the AN-
 1733 TELOPE in comparison to a traditional anomaly detection architecture like an autoencoder. The
 1734 autoencoder cannot accommodate variable length or permutation invariant inputs.

1735 Figure B.1 shows the AUC determined by each of these two tools as a function of number
 1736 of tracks. The trend is that the AE suffers when more information is given, due to the presence
 1737 of 0-padding. In contrast, the ANTELOPE architecture performs better with more information,
 motivating the use of high dimensional input modelin with this method.

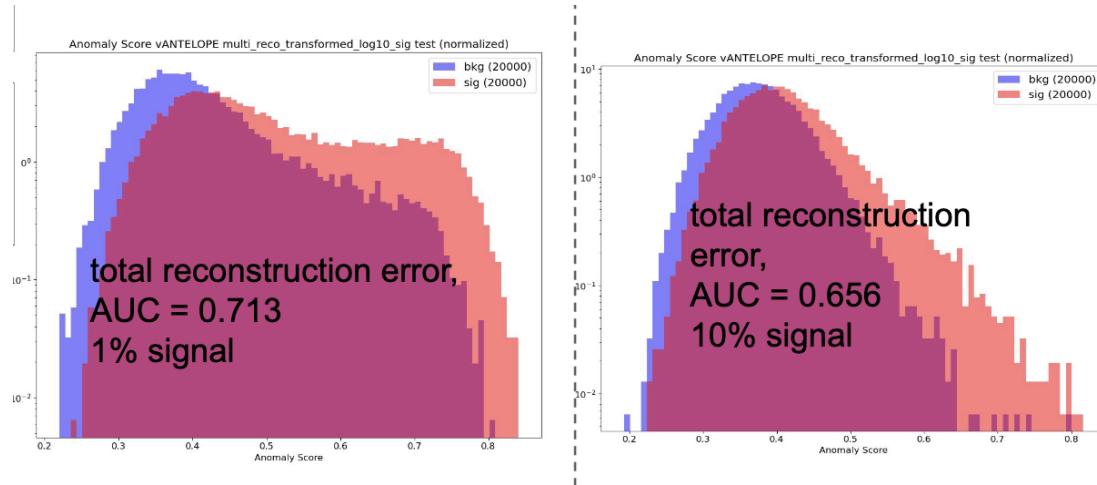


- Confirms that AE is negatively sensitive to zero padding while ANTELOPE sees a performance improvement as it sees more “information”

Figure B.1

1739 **Signal Contamination**

1740 To understand the effect of signal contamination in training on the ANTELOPE score, we inject
1741 a percent of signal events into the data used to train the ANTELOPE autoencoder stage and look
1742 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%.



1743 Figure B.2

1744

B.2 PFN Optimality Checks

1745 The PFN is trained using QCD as the background. A study was done to compare the perfor-
1746 mance of the tool in the analysis context if it trains against QCD or a representative MC background
1747 considering the small fractions of other processes ($V+jets$, top) that would contribute at preselec-
1748 tion. Figure B.3 shows the AUC across the grid for both training approaches, revealing better
1749 performance if the tool focuses on learning QCD differences.

1750 Further studies were done to ensure the relatively optimality of the single PFN model, trained
1751 over combined signals, across the grid. As the grid spans signals with a large range of E_T^{miss} , their
1752 varying input features and background composition may be conducive to separate PFN models
1753 trained on high and low E_T^{miss} signal points to better capture the signal-background differences.
1754 Figure B.4 shows a comparison of the signal-inclusive PFN model performance and the perfor-

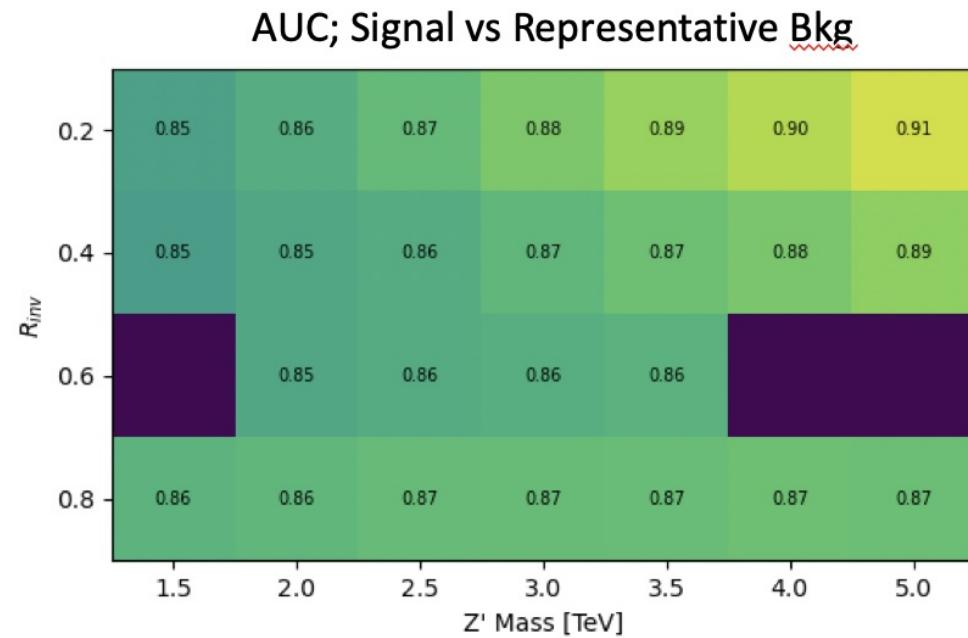
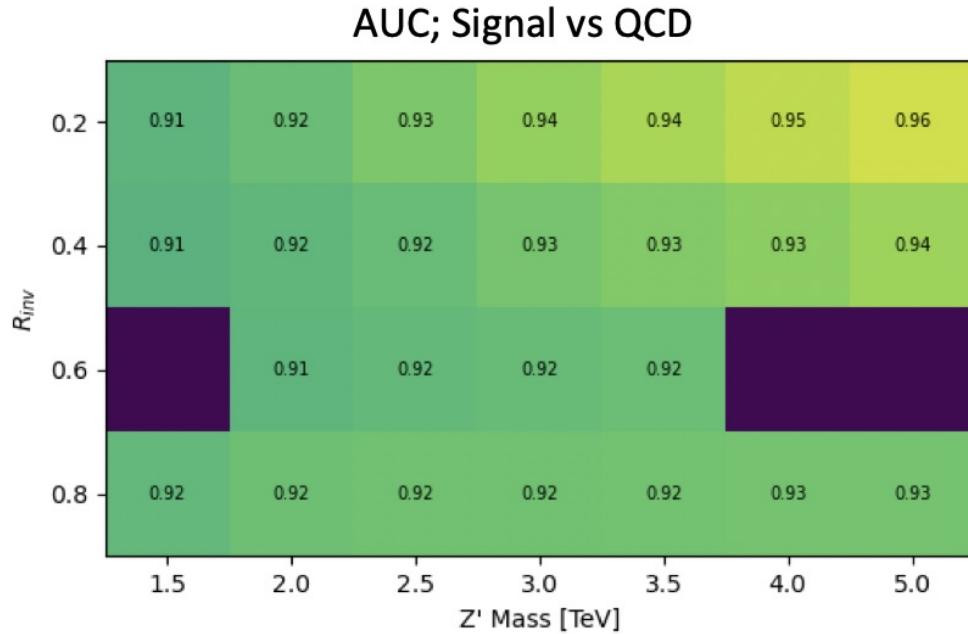


Figure B.3: AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD-only training sample (top) and the total MC background (bottom). Note the three missing points will be added shortly - they were delayed due to a DAOD production mistake.

1755 mance of models separated into high and low R_{inv} signals in training. The most notable impact is
 1756 found for the low R_{inv} and high mass points, indicating that the signal-inclusive PFN is learning
 1757 morning about the distinction between high- E_T^{miss} signals and backgrounds. However, these high
 1758 mass points are also the most challenging to find due to their very wide resonance on top of m_T ,
 1759 and in the final projected sensitivity in the m_T window the differences are $< 10\%$ across the grid.
 1760 To maintain a harmonized strategy with the ANTELOPE region we keep the inclusive PFN model
 1761 as the final version.

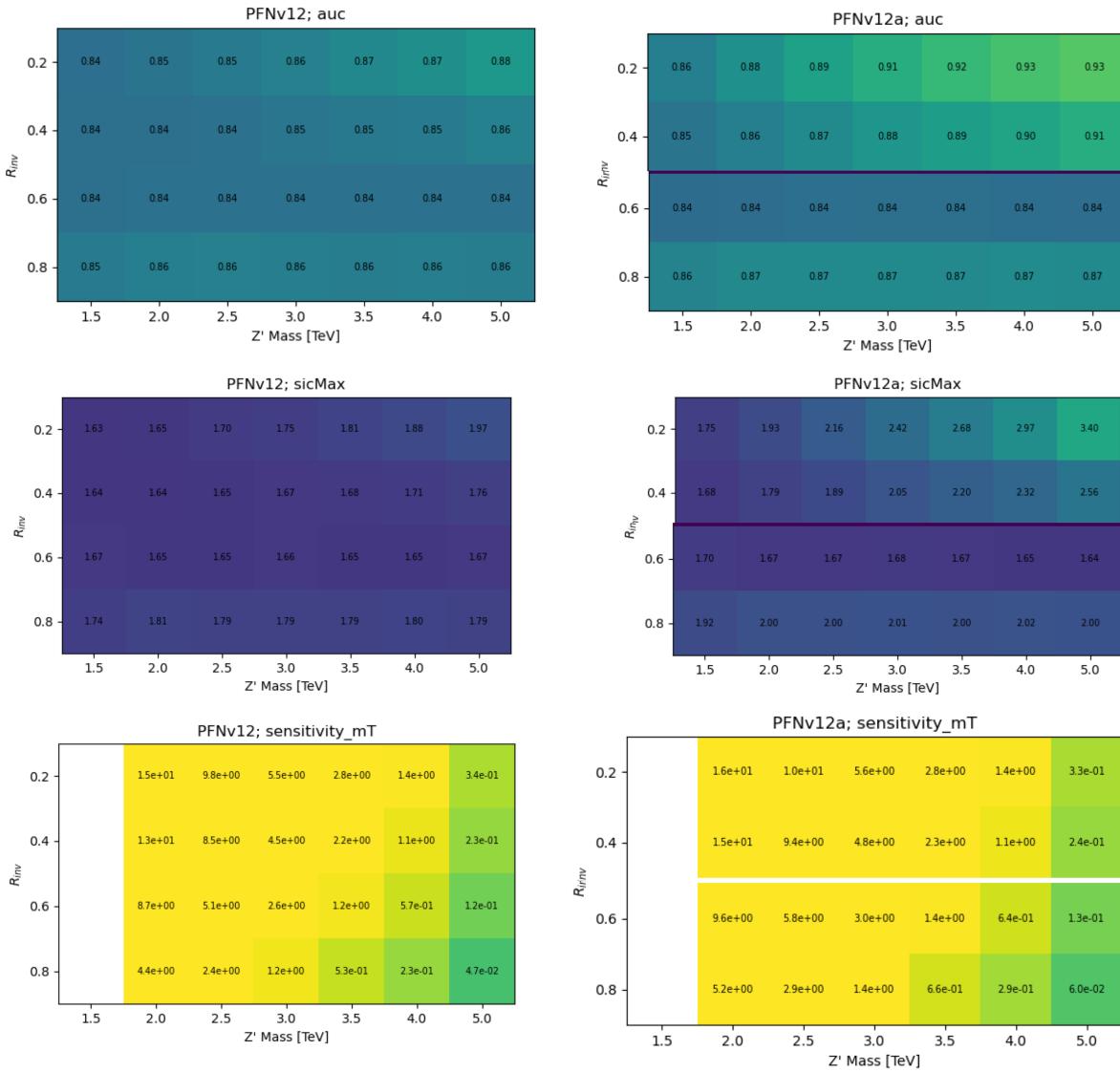


Figure B.4: Comparison of PFN AUC (top), SIC (middle), and sensitivity in the m_T mass window (bottom) for a single PFN model (left) vs. two models, trained on $R_{inv} < 0.5$ and > 0.5 separately.

1762 Figure B.5 shows the optimal cut on the PFN score for each point in the signal grid, motivating
 the loose inclusive choice used to define the SR.

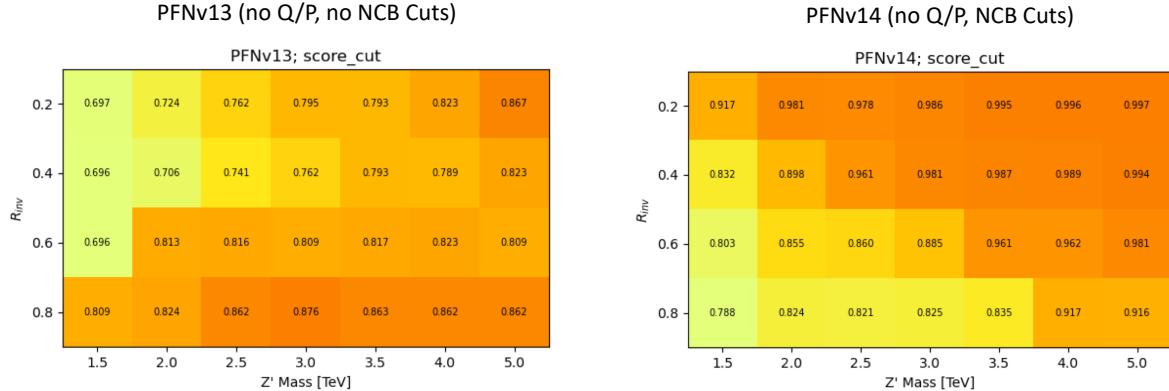


Figure B.5: Preferred cuts on the PFN score for each point in the grid, comparing the effect of adding the NCB preselection.

1763
 1764 Grid cans for optimality were also performed on the number of training epochs, number of
 1765 training events, batch size, learning rate, number of neurons, and dimension of the Φ space. The
 1766 results of these scans are summarized in the tables in Figure B.6. The selected or default parameters
 1767 were found to be optimal, or close enough to optimal to justify not increasing the training time or
 1768 complexity of the network for negligible increases in performance.

1769 B.3 Supervised: BDT vs. PFN

1770 Studies of the BDT compared to the PFN performance, where training over events modeled
 1771 with jet-related HLVs (high-level track variables, etas, angles, etc.) are compared to events mod-
 1772 eled by the tracks of the 2 leading jets. Figure B.7 shows the performance of the BDT with and
 1773 without explicit use of energy scale variables.

1774 B.4 Supervised: Variable Correlations

1775 The correlation between the PFN score and analysis variables known to be unique to semi-
 1776 visible jets are studied, for the sake of better understanding the topology of the tool. $\Delta\phi(j, E_T^{\text{miss}})$

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

	default 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

	default	batchsize_pfn 250	batchsize_pfn 1000
AUC	.905	.903	.902

Figure B.6: Scans done to check for optimality of PFN training parameters.

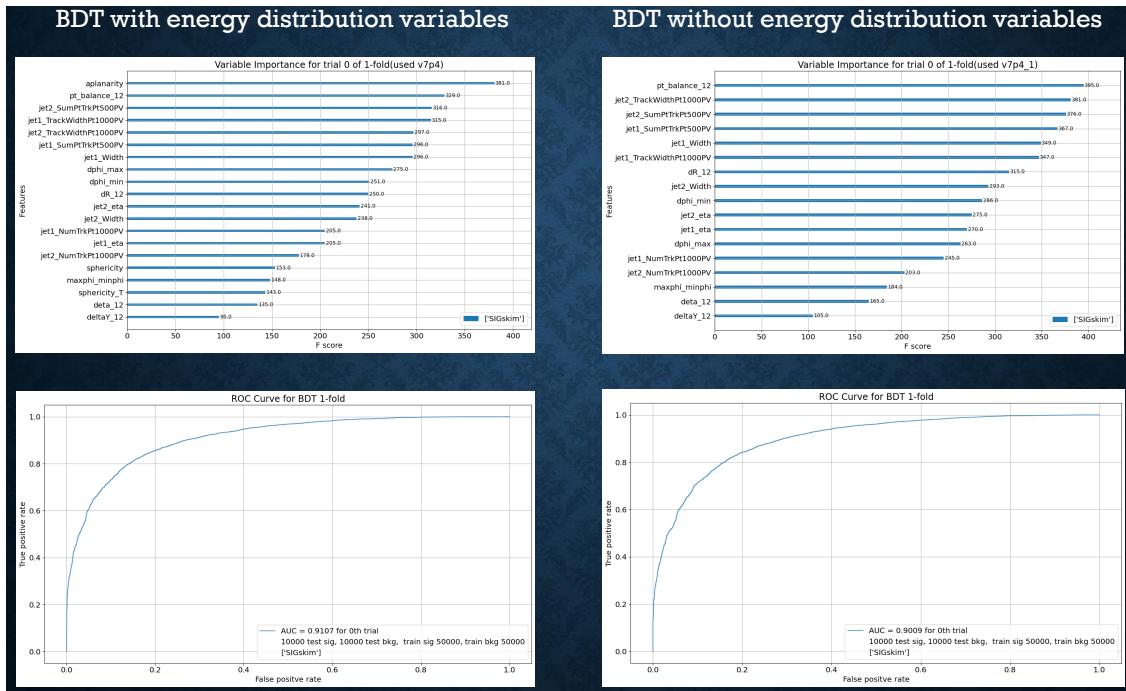


Figure B.7

1777 and $\Delta\phi(j_1, j_2)$ in the CR and VR are highlighted in Figure B.8 (eg. in different bins of the PFN
 1778 score), and 2D plots of the score with respect to the orientation variables are in Figure B.9.

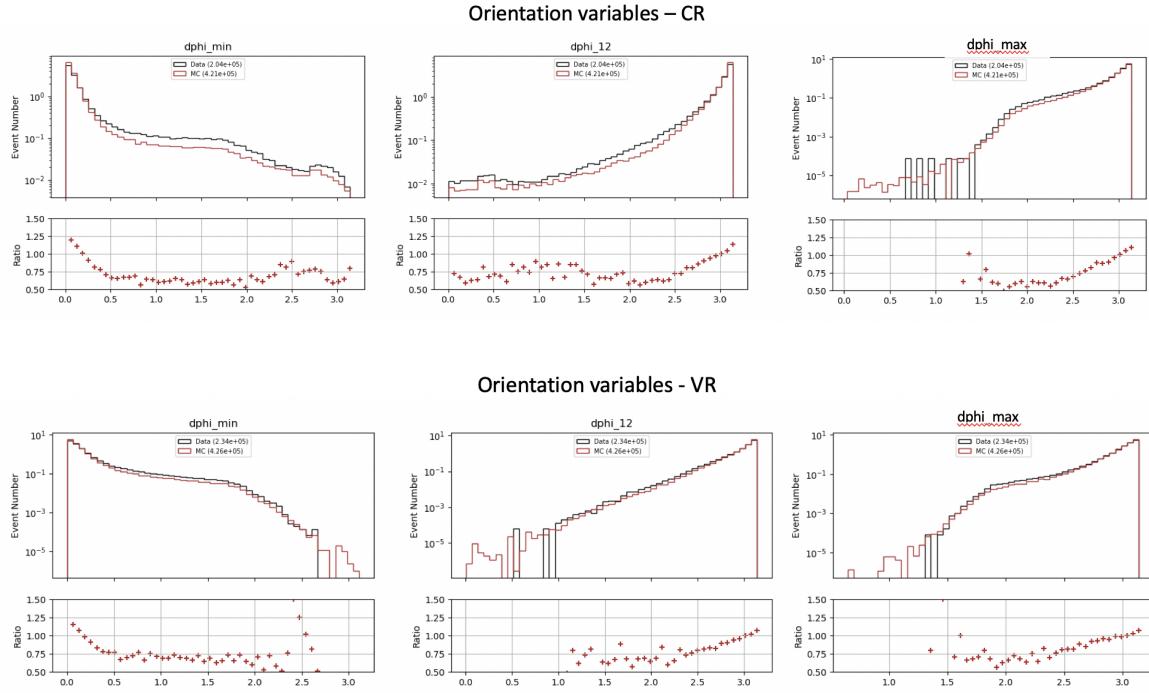


Figure B.8: ϕ orientation variables in the CR and VR

1779 B.5 Single Jet vs Jet System ML Approach

1780 The analysis considered both a single jet and jet system ML approach. A jet system approach,
 1781 where the leading two jets and their orientation with respect to each other was selected for a variety
 1782 of reasons. The jet system approach captures the MET information which is crucial to identifying
 1783 SVJs. In the topology where the dark quarks come from a heavy Z' decay and are back to back,
 1784 the measurable MET will have to be aligned with one or the other.

1785 Additionally, the performance of both a supervised PFN approach and an unsupervised AE ap-
 1786 proach was studied in the case of a single jet tagger. While the PFN approach was still performant
 1787 on a single jet case, the unsupervised approach was significantly improved by using both jets. This
 1788 is shown in Figure B.10.

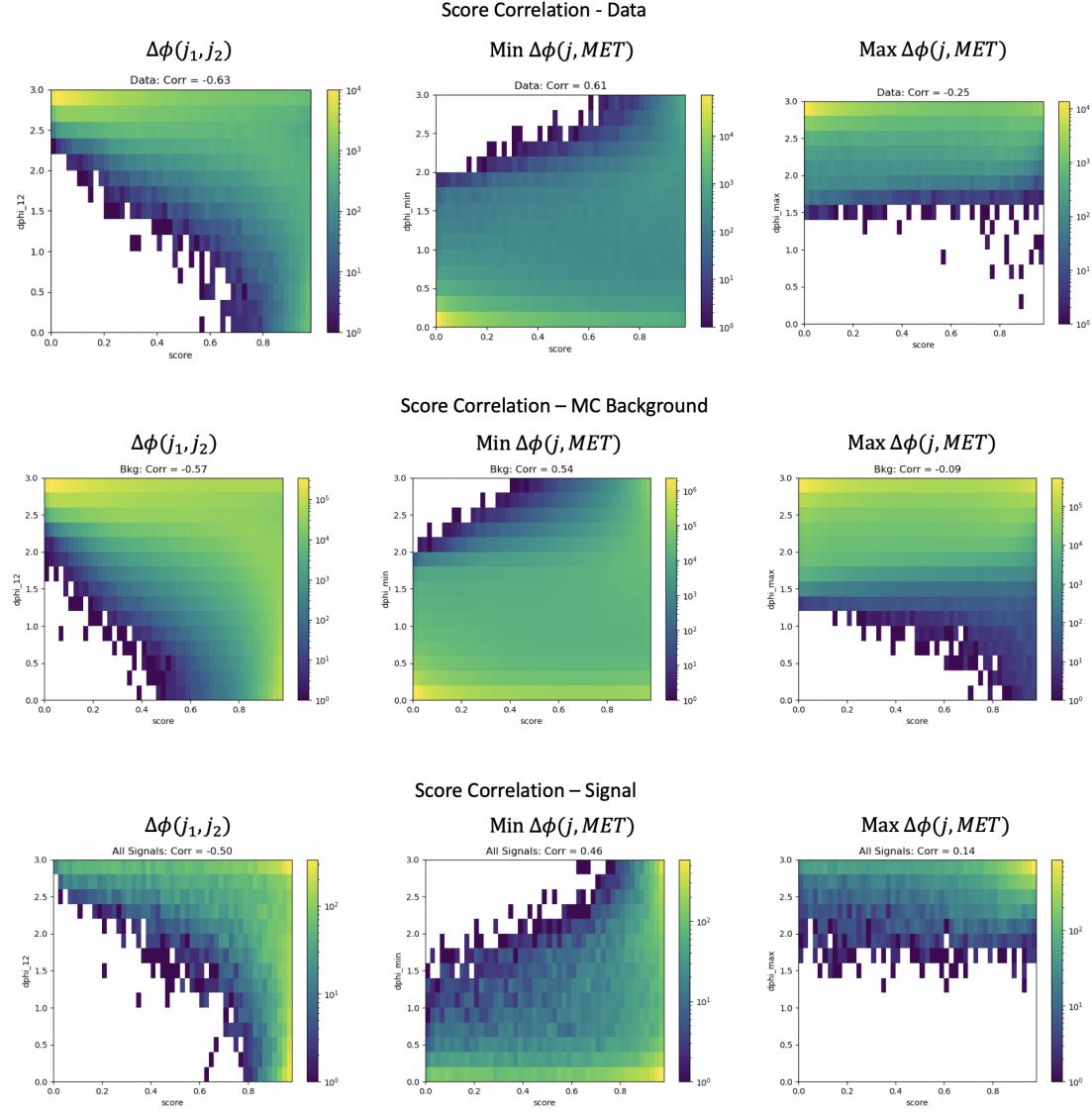


Figure B.9: Correlation between ϕ orientation variables and PFN score

1789 B.6 PFN Training Composition

1790 The overall sensitivity and stability across the signal grid is observed to benefit by training
 1791 the ML tool to reject only the dominant background, QCD. This is evidenced by the PFN re-
 1792 sponse plots shown in Section 7.1.1 and the following AUC and sensitivity comparison plots in
 1793 Figure B.11.

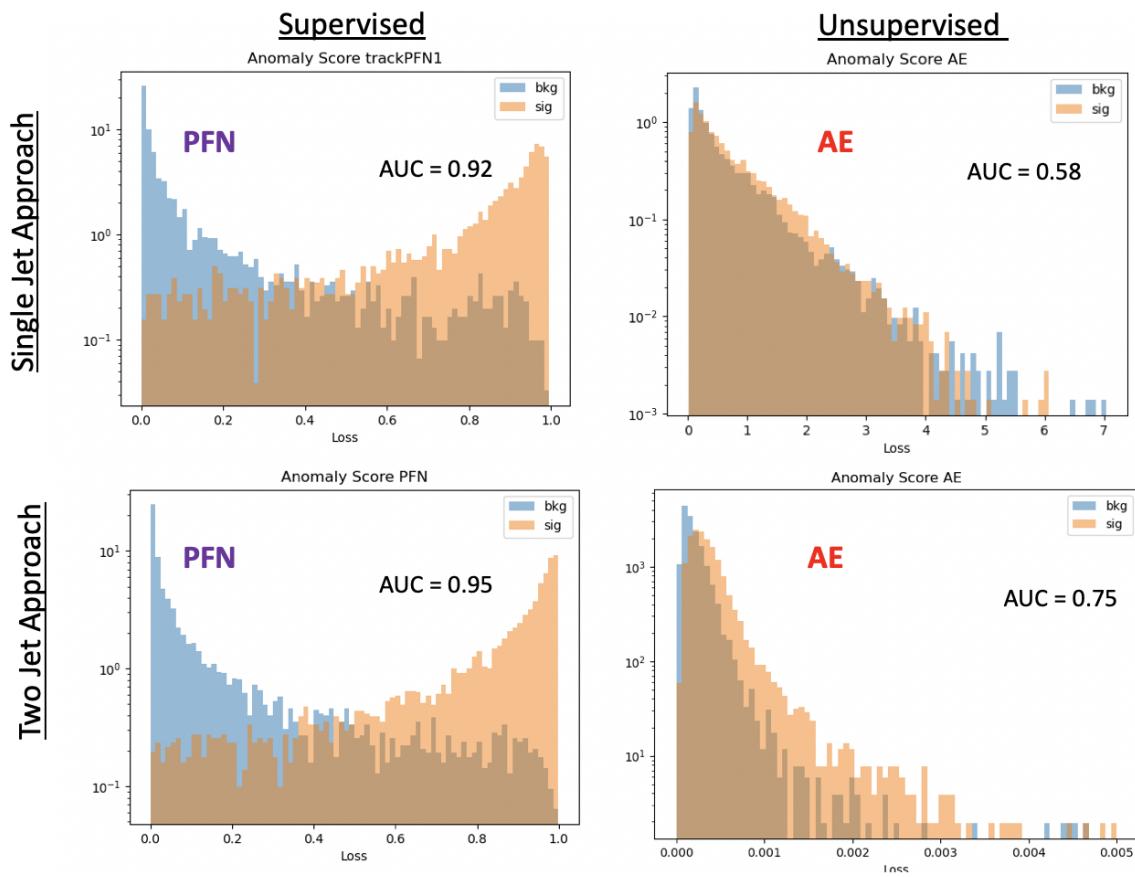


Figure B.10: ϕ Performance comparison between single jet and jet system ML approach

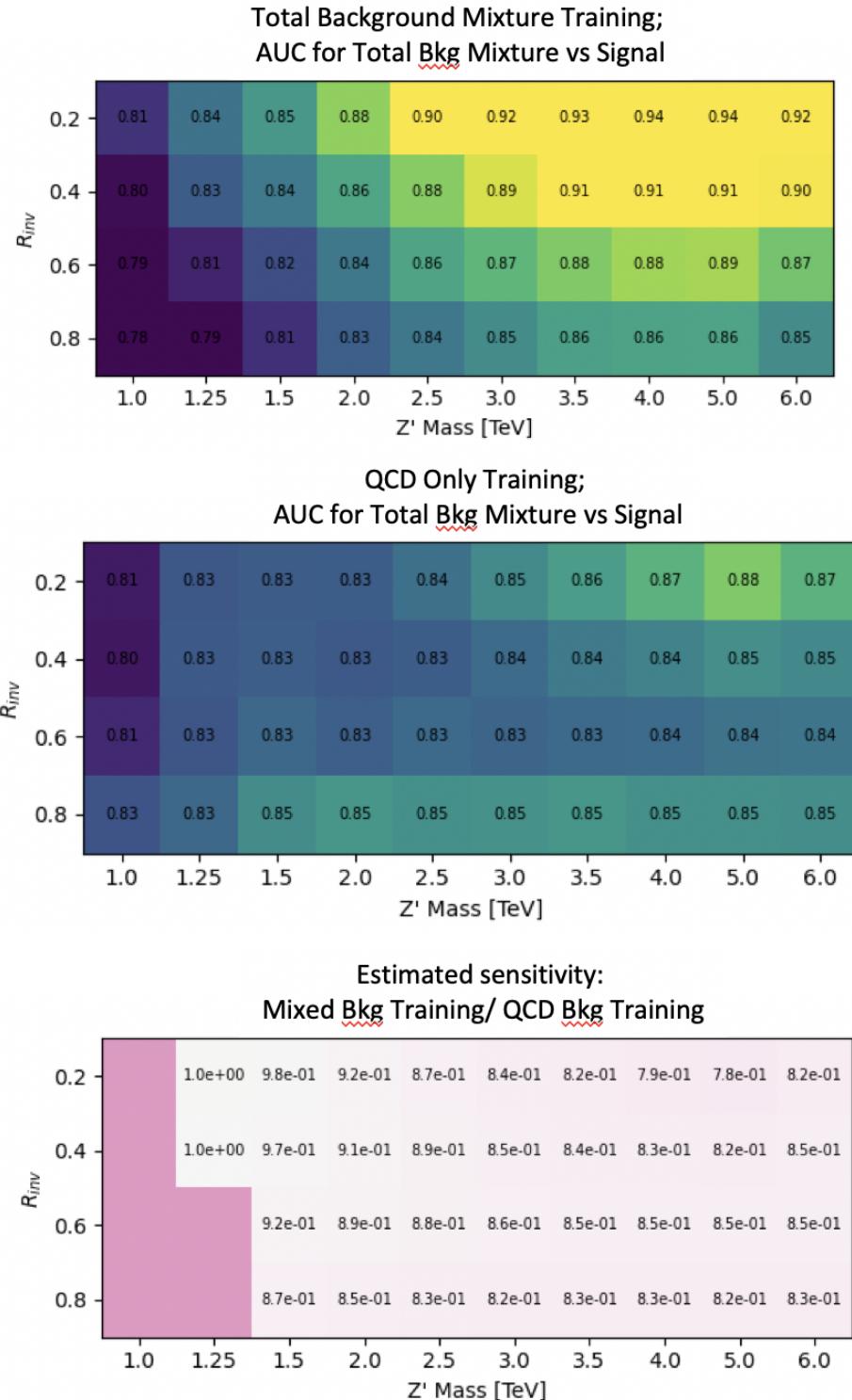


Figure B.11: ϕ Comparison in the AUC score across the grid for the mixed background strategy vs the QCD only strategy. The bottom table highlights that the QCD only strategy gives superior sensitivity across the signal grid.