

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 . . . . .	xii
19	Dedication . . . . .	xii
20	Introduction or Preface . . . . .	1
21	<b>I Theory</b>	<b>2</b>
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	<b>II Experiment</b>	<b>17</b>
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	<b>III Search</b>	<b>55</b>
66	Chapter 6: Monte Carlo and Data . . . . .	56
67	Chapter 7: Machine Learning Tools . . . . .	57
68	7.1 Introduction . . . . .	57
69	7.1.1 Particle Flow Network (Supervised) . . . . .	58
70	7.1.2 ANTELOPE (Semi-supervised) . . . . .	68
71	Chapter 8: Analysis Strategy . . . . .	76
72	Conclusion or Epilogue . . . . .	77
73	References . . . . .	81
74	Appendix A: Machine Learning Approaches . . . . .	86

75	A.1 Unsupervised: AE vs. ANTELOPE . . . . .	86
76	A.2 PFN Optimality Checks . . . . .	87
77	A.3 Supervised: BDT vs. PFN . . . . .	90
78	A.4 Supervised: Variable Correlations . . . . .	90
79	A.5 Single Jet vs Jet System ML Approach . . . . .	92
80	A.6 PFN Training Composition . . . . .	93

## List of Figures

82	1.1	Diagram of the 17 particles comprising the Standard Model . . . . .	4
83	1.2	Fundamental particle interactions of the three fundamental forces described by the	
84		Standard Model [2]. . . . .	6
85	1.3	An example Feynmann diagram of jet production . . . . .	7
86	1.4	An illustration of the “hat shaped” potential of the Higgs field, resulting in a non-	
87		zero vacuum expectation value. . . . .	9
88	2.1	Illustration of the hidden valley potential . . . . .	14
89	2.2	The massive mediator particle $Z'$ of the s-channel realization of a HV model . . .	14
90	3.1	The LHC accelerator complex at CERN [27] . . . . .	20
91	3.2	The octants of the LHC and location of various beam activities [26]. Stars indicate	
92		the locations of beam collisions, and the associated detectors recording the	
93		outcome of those collisions. . . . .	21
94	3.3	(Left) Total integrated luminosity over the course of Run 2. (Right) Average num-	
95		ber of $pp$ interactions per bunch crossing in Run 2. Each curve is weighted by the	
96		integrated luminosity for the year. . . . .	23
97	3.4	Timeline of LHC and HL-LHC activities [29]. Integrated luminosity estimates are	
98		approximate, and not reflective of the exact amount delivered to each experiment. .	25
99	4.1	ATLAS coordinate system and geometry . . . . .	28
100	4.2	A 3D visualization of the structure of the ID in the barrel region [33] . . . . .	29
101	4.3	ATLAS calorimetery system [34] . . . . .	30

102	4.4	Diagram of a segment of the EMB, demonstrating the accordion plate arrangement [35] . . . . .	32
103			
104	4.5	A LAr pulse as produced in the detector (triangle) and after shaping (curve) [35] . . . . .	33
105			
106	4.6	Readout gap structure in HEC [35] . . . . .	33
107			
108	4.7	TileCal wedge module [38] . . . . .	35
109			
110	4.8	Cross section view of the muon spectrometer system [39] . . . . .	36
111			
112	4.9	Layout of the barrel and endcap toroid magnets [32] . . . . .	38
113			
114	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
115			
116	5.2	Track reconstruction seeding, finding and fitting illustration [43] . . . . .	43
117			
118	5.3	Three types of EM object candidates [45]. . . . .	44
119			
120	5.4	Four types of muon track candidates [47]. . . . .	46
121			
122	5.5	The fragmentation and hadronization processes undergone by a quark produced in a proton-proton collision [49]. . . . .	47
123			
124	5.6	A flow chart illustrating the particle flow algorithm progression [54]. . . . .	50
125			
126	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
127			
128	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
129			
130	7.1	The Energy/Particle Flow Network concept, from Ref. [62]. . . . .	59
131			
132	7.2	An annotated diagram of the PFN architecture. $y$ and $\phi$ represent geometric information for the input particles, $z$ represents energy information, and PID encompasses any other particle ID information in the input. . . . .	59
133			

128	7.3 A illustration of the expected dijet behavior of semi-visible jets, where one jet is 129 closely aligned with $E_T^{\text{miss}}$ . . . . .	60
130	7.4 Illustration of track coordinates $d_0$ and $z_0$ . . . . .	61
131	7.5 Distributions of the track multiplicity in the leading and subleading jets, comparing 132 signal and background PFN training samples. . . . .	61
133	7.6 A diagram demonstrating how the two jet system is rotated in $(\phi, \eta)$ . . . . .	62
134	7.7 The 6 PFN track variables in background MC and signal MC. There are some 135 differences between signal and background, but the track kinematics are largely 136 similar. . . . .	62
137	7.8 The 6 PFN track variables in data and background MC, after the scaling and ro- 138 tation procedure is applied. There is excellent modeling of the data by the MC 139 within the track variables. The slight discrepancy in the phi distribution is due to 140 the modeling of dead TileCal cells by the QCD MC, which will be discussed in 141 Chapter 8. The level of discrepancy is determined to be within tolerance given that 142 the final result will be data driven and the QCD model is used in the PFN training 143 only. . . . .	63
144	7.9 PFN score for background MC, data, and signal, comparing a PFN training on 145 QCD-only vs all-background MC samples. The average AUC for the QCD-only 146 training (left) is 0.93, while the average AUC for the mixed background training 147 (right) is 0.84. The sensitivity estimate across the grid is better for the QCD-only 148 training - from the distribution we can conclude that this is because the sensitivity 149 to MET enhanced signals is greatly reduced. . . . .	65
150	7.10 PFN architecture loss during training as a function of epoch (left) and the evaluated 151 loss over the signal and background (right). . . . .	65
152	7.11 ROC the PFN score for combined signal (true positive) and QCD background (false 153 positive). . . . .	66
154	7.12 AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD- 155 only training sample. . . . .	67
156	7.13 PFN score for two signals and the total background MC (top), and between data 157 and MC (bottom). The difference between data and MC efficiency is minimal (< 158 5%). . . . .	67

159	7.14 A visual representation of the 64 PFN latent space variables which create the input 160 of the VAE component of ANTELOPE. The left shows a 2D histogram of the PFN 161 latent space index (0-63) versus the value assumed by that index. The right shows 162 1D histograms of two particular PFN latent space variables. . . . .	69
163	7.15 An annotated diagram of the ANTELOPE architecture. . . . .	70
164	7.16 ANTELOPE architecture loss during training as a function of epoch. . . . .	71
165	7.17 ANTELOPE score distribution comparing data and the total background MC (left), 166 with good agreement observed between data and simulated background, and com- 167 paring all background MC to signals (right), revealing good discrimination power. 168 . . . . .	72
169	7.18 AUC from the ANTELOPE score for each signal in the SVJ grid. . . . .	73
170	7.19 Comparing data and the alternate signal models for the PFN score (top left), AN- 171 TELOPE score (top right), $E_T^{\text{miss}}$ (bottom left), and $m_T$ (bottom right). The emerg- 172 ing jet signal is an example of the gain of the model-independent ANTELOPE ap- 173 proach, where it has a bimodal shape in PFN score but is clearly tagged as anom- 174 alous by ANTELOPE. . . . .	74
175	7.20 Comparing data and the alternate signal models in terms of sensitivity ( $S/\sqrt{B}$ ) for 176 the PFN and ANTELOPE tools, applying the selection that is used in the analysis. 177 The ANTELOPE network is found to provide significant added sensitivity to alter- 178 nate signals such as the gluino $\rightarrow$ R-hadron and emerging jets, which have higher 179 $E_T^{\text{miss}}$ than the SVJs. . . . .	75
180	A.1 . . . . .	86
181	A.2 . . . . .	87
182	A.3 AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD- 183 only training sample (top) and the total MC background (bottom). Note the three 184 missing points will be added shortly - they were delayed due to a DAOD production 185 mistake. . . . .	88
186	A.4 Comparison of PFN AUC (top), SIC (middle), and sensitivity in the $m_T$ mass win- 187 dows (bottom) for a single PFN model (left) vs. two models, trained on $R_{inv} < 0.5$ 188 and $> 0.5$ separately. . . . .	89
189	A.5 Preferred cuts on the PFN score for each point in the grid, comparing the effect of 190 adding the NCB preselection. . . . .	90

191	A.6 Scans done to check for optimality of PFN training parameters. . . . .	91
192	A.7 . . . . .	91
193	A.8 $\phi$ orientation variables in the CR and VR . . . . .	92
194	A.9 Correlation between $\phi$ orientation variables and PFN score . . . . .	93
195	A.10 $\phi$ Performance comparison between single jet and jet system ML approach . . . . .	94
196	A.11 $\phi$ Comparison in the AUC score across the grid for the mixed background strategy vs the QCD only strategy. The bottom table highlights that the QCD only strategy gives superior sensitivity across the signal grid. . . . .	95
197		
198		

## List of Tables

200

## Acknowledgements

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

202        needed.

203

## **Dedication**

204

Dedicated to my friends and family

205

## **Introduction or Preface**

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

209

## **Part I**

210

## **Theory**

## Chapter 1: The Standard Model

213     The Standard Model of particle physics is a universally accepted framework which explains  
 214     the interactions of fundamental particles. All known fundamental particles, outlined in Figure  
 215     1.1, are represented in the Standard Model. The model describes three of the four known forces:  
 216     the electromagnetic force, the weak force, and the strong force. Gravity, the fourth fundamental  
 217     force, is not addressed by the Standard Model. The Standard Model was primarily developed over  
 218     the course of the 1960s and 1970s, by combining the work of many physicists into one coherent  
 219     model. The Standard Model has been established as a well-tested theory by decades of experimen-  
 220     tal physics research.

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

225     **1.1 Phenomenology: Particles and Forces**

226     1.1.1 Particles

227     A classic representation of the particles comprising the Standard Model is shown in Figure  
 228     1.1. The two primary particles classes are bosons (gauge bosons and the scalar Higgs boson) and  
 229     fermions (leptons and quarks). The bosons are carriers of fundamental forces, while the fermions  
 230     are the building blocks of matter. Fermions are sorted into three *generations*, and each fermion is  
 231     identified by a unique *flavor*.

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

# Standard Model of Elementary Particles

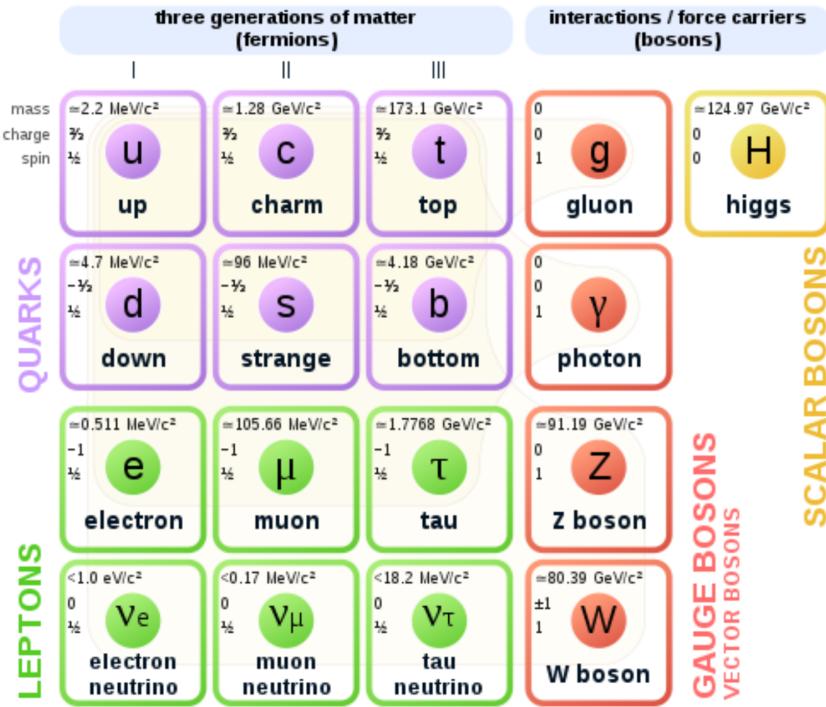


Figure 1.1: Diagram of the 17 particles comprising the Standard Model

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

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

## 243 1.1.2 Forces

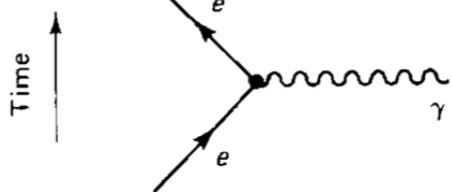
244     The three fundamental forces explained by the Standard Model are the electromagnetic force,  
 245 the strong force, and the weak force. The photon is the carrier of the electromagnetic force, which

<sup>246</sup> dictates the nature of interactions between electrically charged particles, and is widely covered by  
<sup>247</sup> introductory physics courses. The electromagnetic force has an infinite interaction range, a result  
<sup>248</sup> of the massless and non-self interaction nature of the photon. The electromagnetic interaction is  
<sup>249</sup> described by the theory of quantum electrodynamics (QED).

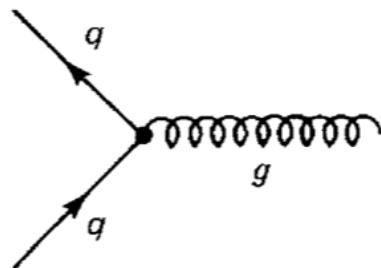
<sup>250</sup> The weak force gives rise to atomic radiation and decay. It allows for the processes of beta  
<sup>251</sup> decay, which enables conversion between neutrons and protons within the nucleus of an atom. In  
<sup>252</sup> the process of beta decay, a proton decays into a neutron, a positron, and a neutrino; or, a neutron  
<sup>253</sup> decays into a proton, an electron and an antineutrino. The weak interaction allows for quark flavor  
<sup>254</sup> mixing, the which enables beta decay. The  $W^+$ ,  $W^-$ , and  $Z^0$  are the force carriers of the weak force.  
<sup>255</sup> The effective range of the weak force is limited to subatomic distances, as a result of the massive  
<sup>256</sup> nature of the mediator bosons. The unified theory of the electroweak interaction posits that at high  
<sup>257</sup> enough energies the electromagnetic interaction and the weak force merge into the same force.  
<sup>258</sup> This threshold is termed the unification energy and calculated to be about 246 GeV [1].

<sup>259</sup> The strong force confines quarks into hadron particles, such as protons and neutrons. The  
<sup>260</sup> strong force also allows for the creation of atomic nuclei by binding protons and neutrons together,  
<sup>261</sup> and is generally referred to as the “nuclear force” in this context. The gluon is the mediator of  
<sup>262</sup> the strong force, which is a short-range force which acts at subatomic distances on the order of  
<sup>263</sup>  $10^{-15}$  m. At this range, the strong force is about 100x as strong as the electromagnetic force,  
<sup>264</sup> which allows for the creation of positively charged nuclei [2]. The strong force is described by the  
<sup>265</sup> theory of quantum chromodynamics (QCD). In the same way that QED dictates the interaction of  
<sup>266</sup> electrically charges particles, QCD dictates the interactions of *color-charged* particles. Due to the  
<sup>267</sup> particular importance of QCD in this thesis, this topic will be explored in detail in section 1.2.

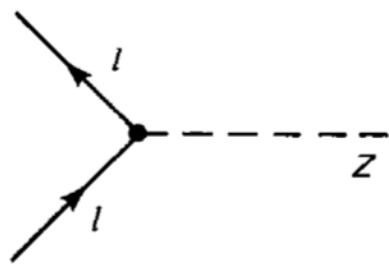
<sup>268</sup> The fundamental Feynmann diagram for each of the three forces discussed here is depicted  
<sup>269</sup> in Figure 1.2. The fourth fundamental force, gravity, is not currently explained by any known  
<sup>270</sup> mechanism within the Standard Model.



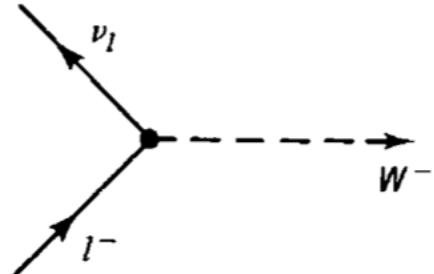
(a) The electromagnetic force



(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].

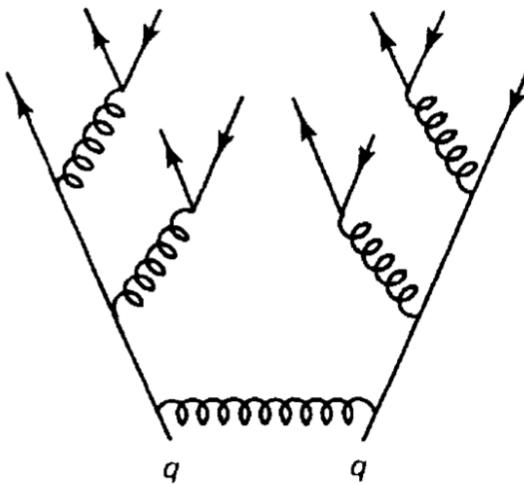


Figure 1.3: An example Feynmann diagram of jet production

<sup>271</sup> **1.2 QCD and Jets**

<sup>272</sup> While there is only one type of electric charge, there are three types of color charge; red, green,  
<sup>273</sup> and blue. In the process  $q \rightarrow q + g$ , the color of the quark can change. In order to conserve color  
<sup>274</sup> charge, gluons are bicolored, and always carry some positive color charge and some negative color  
<sup>275</sup> charge.

<sup>276</sup> Color charged particles can only exist in bound states which result in a neutral total color  
<sup>277</sup> charge, a principle known as confinement. This requires that quarks and gluons exist in group  
<sup>278</sup> states known as hadrons; either mesons in the case of two quarks or baryons in the case of three  
<sup>279</sup> quarks. When a quark is separated from a hadron, confinement dictates that other colored objects  
<sup>280</sup> are produced around the quark to obey confinement. An example of this process is shown in  
<sup>281</sup> Figure 1.3. This ensemble of objects, generally a mixture of quarks and gluons, is termed a *jet*.  
<sup>282</sup> Jets are among the most common phenomenon observed by detectors at hadron colliders, and their  
<sup>283</sup> complex structure makes them a key focus of many physics analyses.

284 **1.3 Symmetries**

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

286  $G_{SM}$ :

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

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

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

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

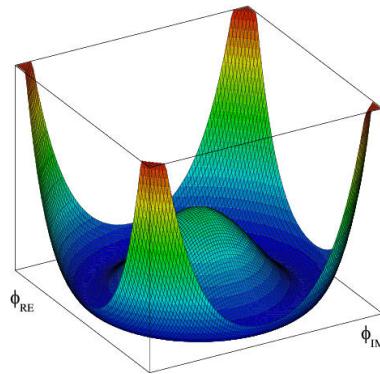


Figure 1.4: An illustration of the “hat shaped” potential of the Higgs field, resulting in a non-zero vacuum expectation value.

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  $\gamma$  is created from the components of the

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

330 **1.4 Experimental Validation of the Standard Model**

331 The theoretical framework of the Standard Model coalesced into a unified theory in the mid-  
332 20th century. A cascade of discoveries providing empirical evidence for the model followed  
333 closely. In the 1960s, three quarks (up, down and strange) and four leptons (electron, muon,  
334 and their associated neutrinos) were the known particulate building blocks of matter and the Stan-  
335 dard Model. The discovery of the charm quark in 1974, through the observation of the  $J/\psi$  meson  
336 [6][7], confirmed the existence of a fourth quark flavor. The discovery of the  $\tau$  in 1975 [8] provided  
337 the first evidence of a 3rd generation of matter. This was quickly followed by the observation of  
338 the  $\Upsilon$  meson in 1977 [9], which provided evidence for the existence of a fifth quark, the  $b$  quark  
339 (bottom, or beauty). The existence of a 3rd generation of fermion also explained the observation  
340 of CP violation in the weak force, as it allowed for the addition of a complex phase in the CKM  
341 matrix (a unitary matrix which describes flavor mixing in the weak interaction). The top quark  
342 ( $t$ ) and tau neutrino ( $\nu_\tau$ ) were predicted at this point as the final building blocks of three complete  
343 generations of fermions, and they were discovered by experimental observation around the turn of  
344 the 21st century [10] [11] [12].

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

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

353 **1.5 Limitations of the Standard Model**

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

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

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

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

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

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

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

## Chapter 2: Physics Beyond the Standard Model

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

395 **2.1 Hidden Valley Models**

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

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

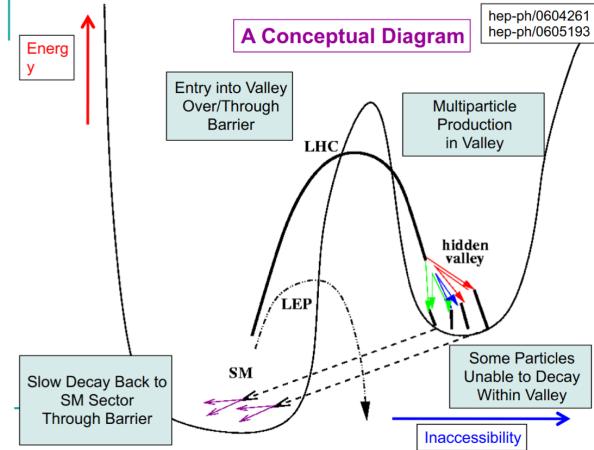


Figure 2.1: Illustration of the hidden valley potential

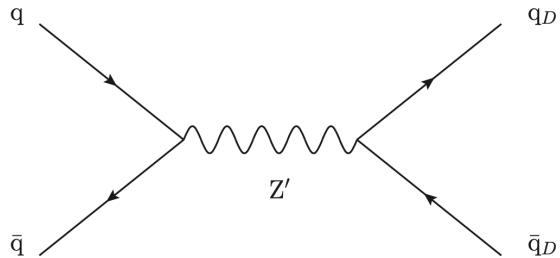


Figure 2.2: The massive mediator particle  $Z'$  of the s-channel realization of a HV model

## 409 2.2 Dark QCD

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

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

414 The first term allows for the dark gluons to self-interact, while the second term enables the dark  
 415 quarks to hadronize and acquire mass. The dark quarks are assumed to have a common mass  $M_d$ .  
 416 The coupling strength of the strongly interacting dark quarks is termed  $\alpha_d$ . At the confinement

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

418 The properties of the hadrons formed by the dark quarks are of particular importance to the  
419 observed dark QCD dynamics. Dark-isospin number  $U(1)_{1-2}$  and dark-baryon number  $U(1)_{1+2}$   
420 are accidental symmetries of the theory which determine the stability of the hadrons. In the case  
421 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  
422 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  
423 stable if this symmetry is unbroken. The baryons would also be stable as they are charged under  
424 the dark-baryon number. These four stable hadrons become dark matter candidates of the theory.  
425 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.  
426 The unstable mesons can decay into stable dark mesons, or into an off-shell  $Z'$ . The off-shell  $Z'$   
427 will then decay into two DM quarks or two SM quarks, and its products will continue to shower  
428 until the final state particles are stable.

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

### 434 2.3 Semi-visible Jets

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

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

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

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

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

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

457

## **Part II**

458

## **Experiment**

## Chapter 3: The Large Hadron Collider

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

466       The LHC accelerates protons and heavy ions, and collides them at four interaction points  
 467       around the ring, with a design center-of-mass energy per collision of  $\sqrt{s} = 14$  TeV. Each interaction  
 468       point is home to one of four detector experiments, which study the products of the collisions. The  
 469       largest of these experiments is the ATLAS detector, a general purpose detector designed to study  
 470       the Standard Model and search for new physics that could be produced in LHC collisions [22].  
 471       The CMS detector is another general purpose detector, designed and operated independently of the  
 472       ATLAS detector, but intended to probe the same range of physics [23]. The ALICE experiment is  
 473       a dedicated heavy ion experiment, and the LHC-b experiment is a dedicated *b*-physics experiment  
 474       [24] [25].

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

479 **3.1 Accelerator Physics**

480 **3.1.1 The Journey of a Proton**

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

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

501

## The CERN accelerator complex Complexe des accélérateurs du CERN

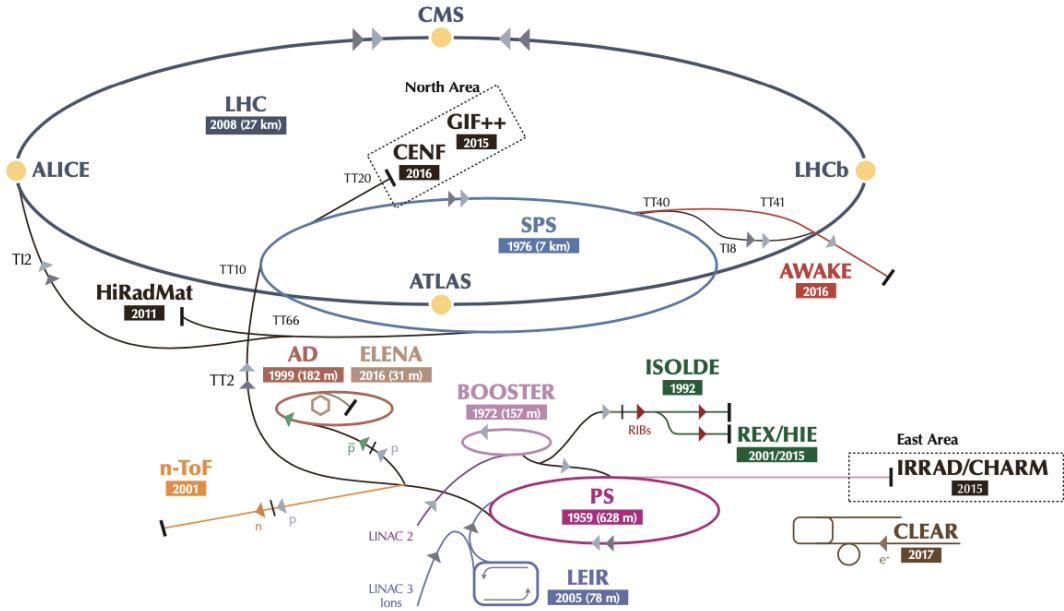


Figure 3.1: The LHC accelerator complex at CERN [27]

### 502 3.1.2 Magnets

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

508 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  
 509 to follow the roughly circular 27 km path. The main dipoles generate powerful 8.3 tesla magnetic  
 510 fields to achieve this bend. Each dipole magnet is 15 meters long and weighs 35 tonnes. The  
 511 dipoles work in conjunction with quadrupole magnets, which keep the particles in a focused beam,  
 512 and smaller sextupole, octupole and decapole magnets which tune the magnetic field at the ends of  
 513 the dipole magnets [28].

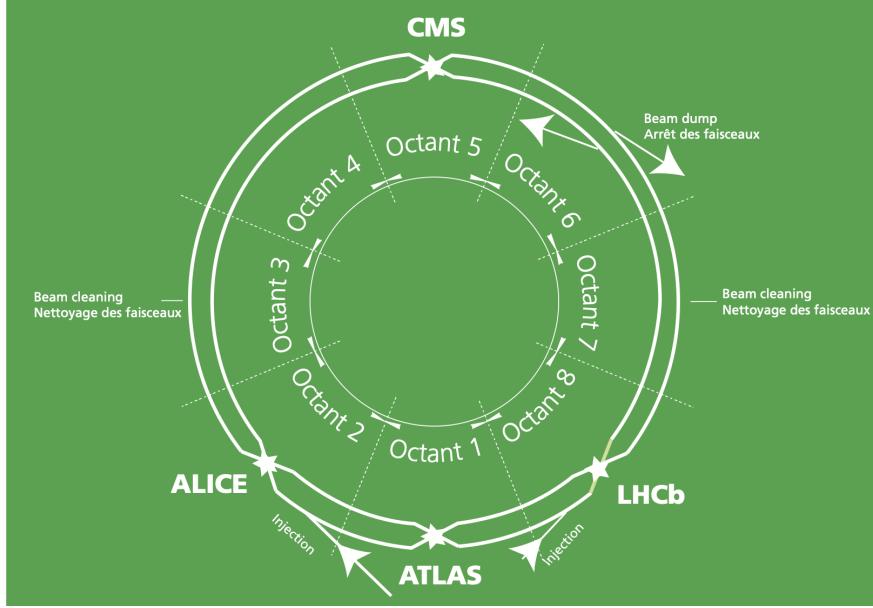


Figure 3.2: The octants of the LHC and location of various beam activities [26]. Stars indicate the locations of beam collisions, and the associated detectors recording the outcome of those collisions.

515     The straight insertion sections have different purposes depending on their location around the  
 516     ring: beam collisions, beam injection, beam dumping, or beam cleaning. At the four collision  
 517     points, insertion magnets squeeze the beam to ensure a highly focused collision. This is accom-  
 518     plished with a triplet of quadrupole magnets, which tighten the beam from 0.2 millimeters to just  
 519     16 micrometers in diameter. Insertion magnets also clean the beam, which prevents stray particles  
 520     from hitting sensitive components throughout the LHC. When the LHC is ready to dispose of a  
 521     beam of particles, beam dump magnets deflect the path of the beam into a straight line towards  
 522     a block of concrete and graphite that stops the beam. A dilution magnet then reduces the beam  
 523     intensity by a factor of 100,000 before the final stop [28]. Figure 3.2 shows the locations various  
 524     beam activities.

### 525     3.2 Luminosity

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

528 formula for which is given in equation 3.1.

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

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

531 The instantaneous luminosity gives the number of the collisions that could be produced at the  
532 interaction point per unit of cross-sectional area per unit of time, generally expressed in  $\text{cm}^{-2}\text{s}^{-1}$ .  
533 The integrated luminosity is obtained by integrating the instantaneous luminosity over a given  
534 block of time, and measures the total number of collisions which have occurred during that op-  
535 eration period. The total integrated luminosity is directly correlated with the size of the datasets  
536 collected by the LHC experiments. Total integrated luminosity for Run 2 is illustrated in Figure  
537 3.3.

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

545 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  
546 peak instantaneous luminosity was  $0.8 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . Over the course of Run 1 the LHC collected  
547 a total integrated luminosity of  $5.46 \text{ fb}^{-1}$  at  $\sqrt{s} = 7 \text{ TeV}$ , and  $22.8 \text{ fb}^{-1}$  at  $\sqrt{s} = 8 \text{ TeV}$ . Following the  
548 first long shutdown and upgrade phase of operations, the LHC achieved a center of mass energy  
549  $\sqrt{s} = 13 \text{ TeV}$  at the beginning of Run 2 in 2015. The LHC was also able to deliver  $2.0 \times 10^{34}$   
550  $\text{cm}^{-2}\text{s}^{-1}$  peak instantaneous luminosity, double the design value. During LHC Run 2, from 2015-  
551 2018, the LHC delivered  $156 \text{ fb}^{-1}$  of integrated luminosity for proton-proton collisions. Run 3 of

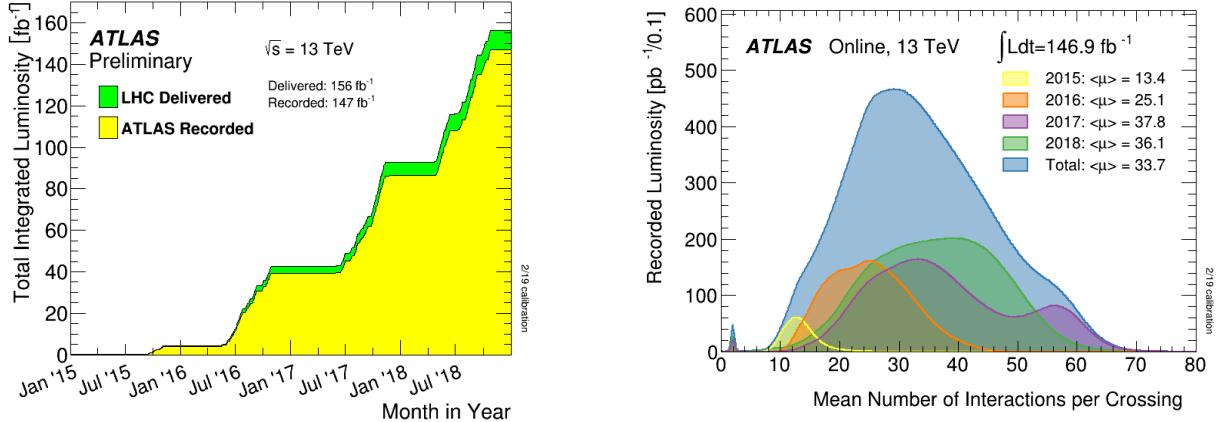


Figure 3.3: (Left) Total integrated luminosity over the course of Run 2. (Right) Average number of  $pp$  interactions per bunch crossing in Run 2. Each curve is weighted by the integrated luminosity for the year.

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

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

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

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

564    **3.3 LHC Timeline**

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

570    Between 2013 and 2015 the LHC underwent the first Long Shutdown (LS1) during which  
571    time maintenance and renovation was performed on the accelerator chain, including the repair and  
572    consolidation of the high-current splices which connect the super-conducting LHC magnets. Run  
573    2 of the LHC took place from 2015 to 2018 and achieved a center-of-mass energy of  $\sqrt{s} = 13$  TeV.  
574    Analysis of data collected in Run 2 is still on going, and is the subject of study in this thesis.

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

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

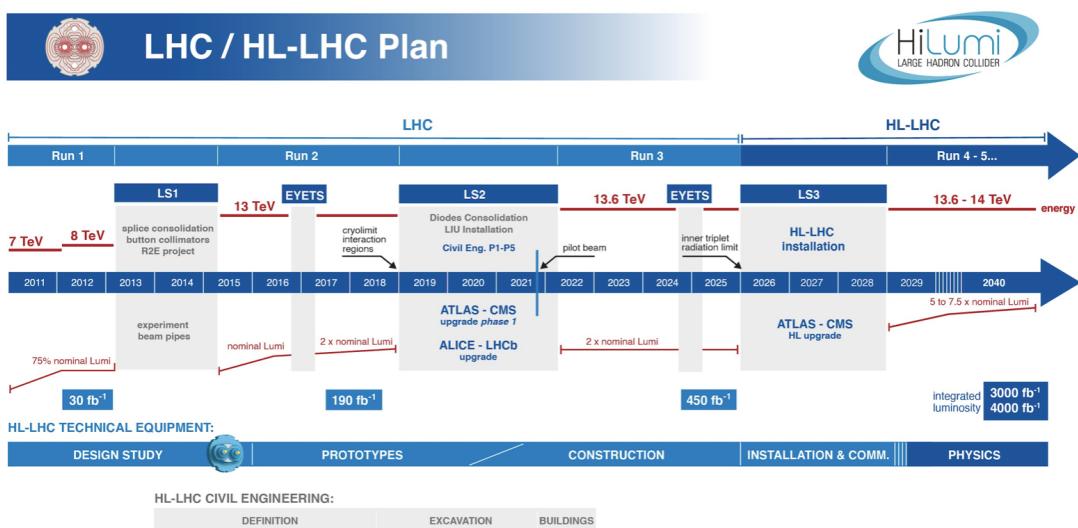


Figure 3.4: Timeline of LHC and HL-LHC activities [29]. Integrated luminosity estimates are approximate, and not reflective of the exact amount delivered to each experiment.

## Chapter 4: The ATLAS Detector

587     The ATLAS detector (**A** Toroidal **L**H**C** Apparatu**S**) is one of two general purpose physics  
 588 detectors designed to study the products of proton-proton collisions at the LHC. The detector is  
 589 composed of a variety of specialized subsystems, designed to fully capture a wide array of physics  
 590 processes. The apparatus is 25m high, 44m in length, and weighs over 7000 tons [32]. The LHC  
 591 beam pipes direct proton beams to an interaction point at the center of ATLAS, and the cylindrical  
 592 detector design captures a complete 360° view of the *event*, tracking all particles that result from  
 593 the collision.

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

### 604     4.1 Coordinate System and Geometry

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

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

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

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

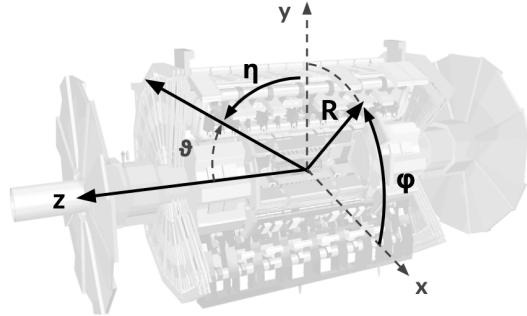
612 This substitution is advantageous because objects in the detector are traveling at highly rela-  
613 tivistic speeds. The relativistic speed also means that the masses of the particles are generally small  
614 compared to their total energy. In the limit of zero mass, the rapidity  $y$  reduces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

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

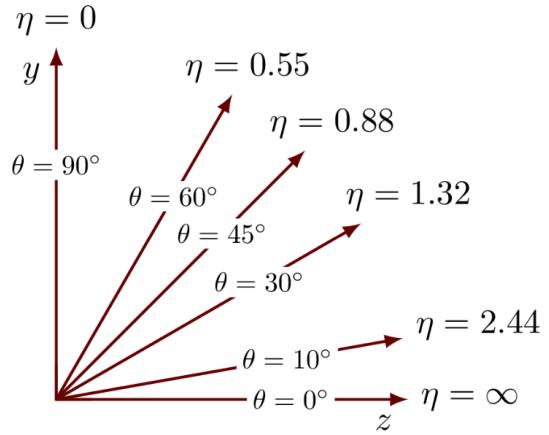
616 The distance between physics objects in the detector is generally expressed in terms of the solid  
617 angle between them  $\Delta R$ :

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

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



(a) The ATLAS geometry



(b) Relationship between  $\eta$  and  $\theta$

Figure 4.1: ATLAS coordinate system and geometry

## 627 4.2 Inner Detector

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

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

### 635 4.2.1 Pixel Detector

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

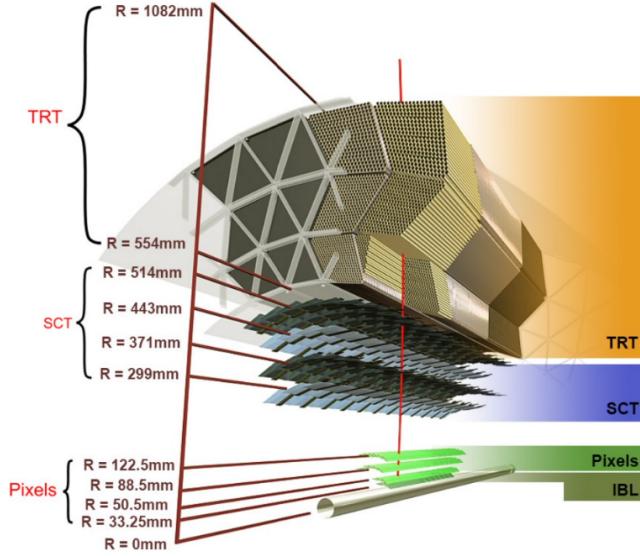


Figure 4.2: A 3D visualization of the structure of the ID in the barrel region [33]

643 just 3.3 cm from the beamline. The pixels of the IBL measure  $50 \mu\text{m}$  by  $250 \mu\text{m}$ , and cover  
 644 a pseudo-rapidity range up to  $|\eta| < 3$ . The IBL upgrade enhances the pixel detector's ability  
 645 to reconstruct secondary vertices associated with short-lived particles such as the b-quark. The  
 646 improved vertex identification also helped compensate for increasing pile-up in Run 2 [32].

647 4.2.2 Semiconductor Tracker

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

653 4.2.3 Transition Radiation Tracker

654 The TRT provides an additional 36 hits per particle track. The detector relies on gas filled  
 655 straw tubes, a technology which is intrinsically radiation hard. The straws which are each 4 mm in  
 656 diameter and up to 150 cm in length and filled with xenon gas. The detector is composed of about

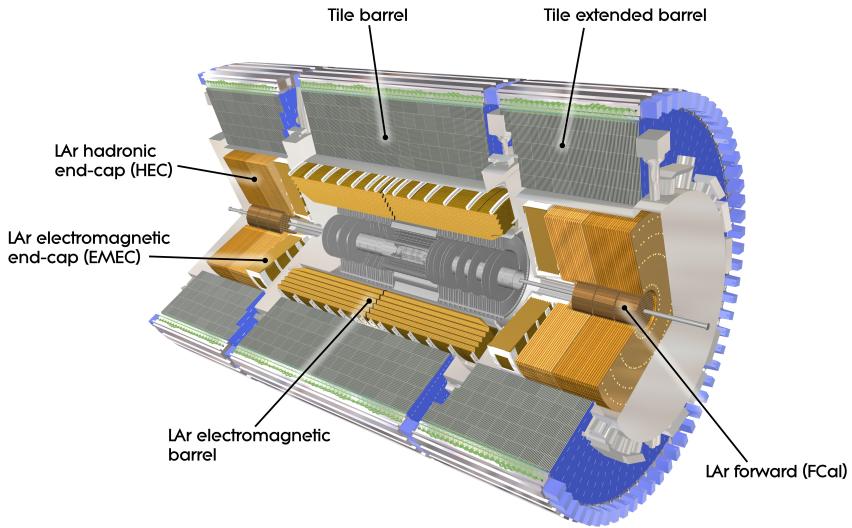


Figure 4.3: ATLAS calorimetry system [34]

657 50000 barrel region straws and 640000 end-cap straws, comprising 420000 electronic readout  
 658 channels. Each channel provides a drift time measurement with a spatial resolution of  $170\ \mu\text{m}$  per  
 659 straw. As charged particles pass through the many layers of the detector, transition radiation is  
 660 emitted. The use of two different drift time thresholds allows the detector to distinguish between  
 661 tracking hits and transition radiation hits.

### 662 4.3 Calorimeters

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

670 4.3.1 Liquid Argon Calorimeter

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

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

688 **Electromagnetic Calorimeter**

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

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

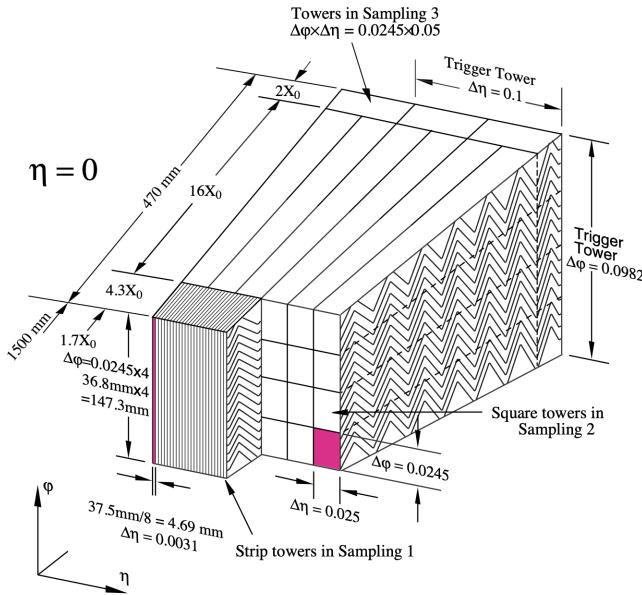


Figure 4.4: Diagram of a segment of the EMB, demonstrating the accordion plate arrangement [35]

696 released when a charged particle ionizes the liquid argon. In the barrel region, the electrons are  
 697 driven towards the center electrodes by a 2000 V potential with a drift time of less than 450 ns [37].  
 698 In the end-caps the voltage varies as a function of the radius in order to maintain a flat response  
 699 [35]. The amount of current produced by the ionized electrons is proportional to the energy of  
 700 the particle creating the signal. Figure 4.5 shows the shape of the signal produced in the LAr  
 701 calorimeter, before and after it undergoes shaping during the readout process. The shaping of the  
 702 pulse enforces a positive peak and a negative tail, which ensures that subsequence pulses can be  
 703 separated with the precision required for the 25 ns LHC bunch spacing [35].

#### 704 **Hadronic End-cap Calorimeter**

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

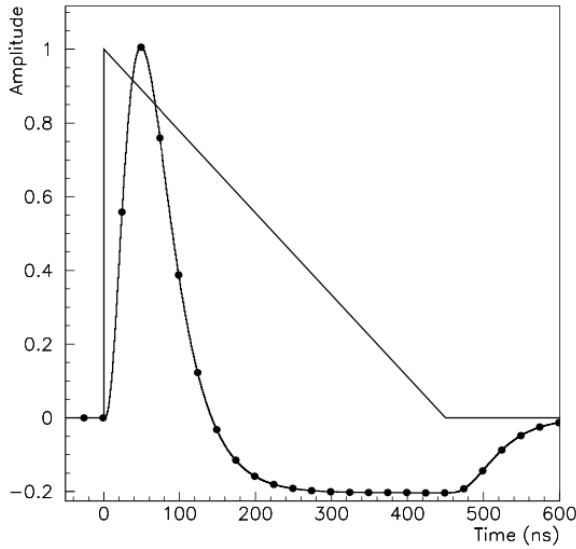


Figure 4.5: A LAr pulse as produced in the detector (triangle) and after shaping (curve) [35]

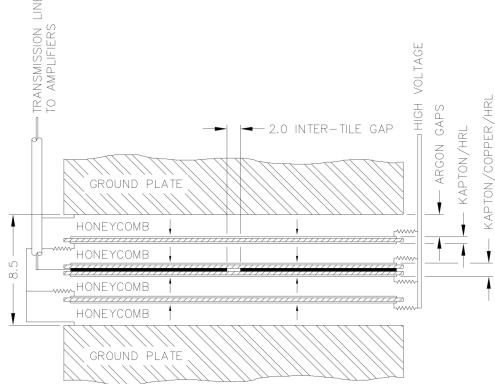


Figure 4.6: Readout gap structure in HEC [35]

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

## 712 **Forward Calorimeter**

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

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

720 4.3.2 Tile Calorimeter

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

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

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

738

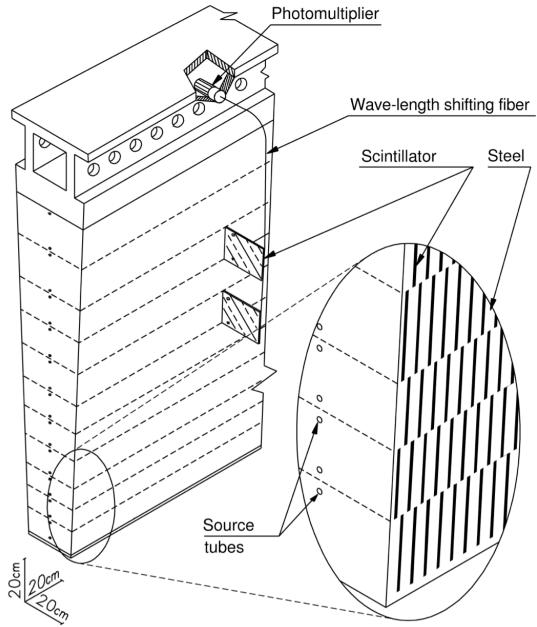


Figure 4.7: TileCal wedge module [38]

#### 739 4.4 Muon Spectrometer

740 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 741 Spectrometer (MS) provides additional tracking information to improve the identification and measurement 742 of muons. The MS comprises the outermost layers of the detector, and is interspersed 743 with toroid magnets (discussed in Section 4.5), which provide a magnetic field of approximately 744 0.5 T. The magnetic field bends the trajectory of the muons as they pass through the detector, and 745 the degree of the bend is directly correlated with the muon momentum. The path of the muon is 746 primarily measured by hits in three layers of Monitored Drift Tube (MDT) precision chambers, 747 which cover the range  $|\eta| < 2.7$ . The barrel layout of the MS is show in Figure 4.8. 748

749 Muon triggering is provided by three layers of Resistive Plate Chambers (RPC) in the barrel 750 ( $|\eta| < 1.05$ ), and 3 - 4 layers of Thin Gap Chambers (TGC) in the end-caps ( $1.05 < |\eta| < 2.4$ ). 751 RPCs and TGCs also provide muon track measurements in the non-bending coordinate ( $\phi$ ). RPCs 752 are constructed from two parallel resistive plates separated by a 2mm gap filled with a sensitive

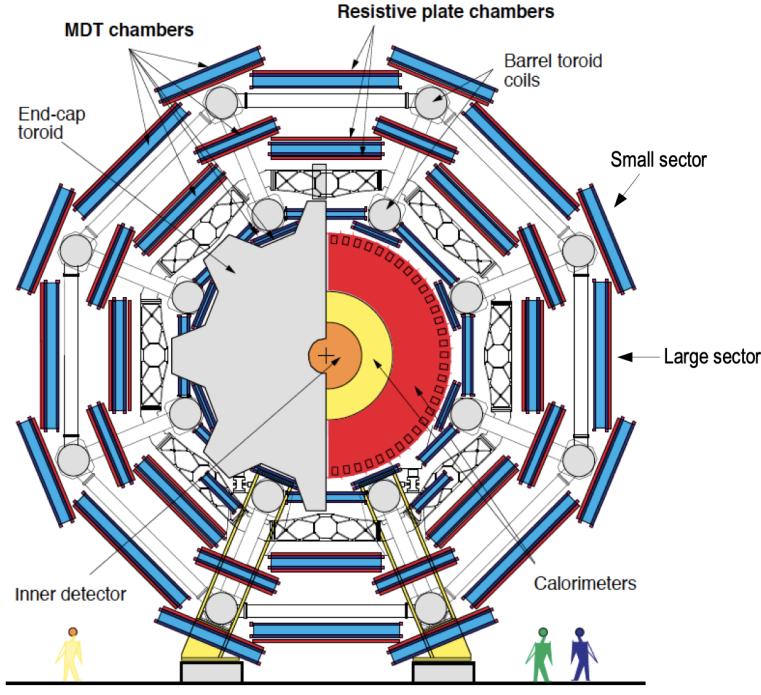


Figure 4.8: Cross section view of the muon spectrometer system [39]

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

760 Precision measurements of muon momentum and position are primarily achieved by MDTs.  
 761 The MDTs are constructed from 30 mm diameter tubes, permeated by a gas mixture of 93% Ar and  
 762 7% CO<sub>2</sub>. The average single-tube spatial resolution is  $80\ \mu\text{m}$ . Each chamber consists of six drift  
 763 tube layers, which together provide a muon track segment resolution of  $35\ \mu\text{m}$ . The momentum  
 764 of the muons can be calculated from the bend in the muon trajectory as they pass through the  
 765 0.5T magnetic field provided by the toroids. For a  $p_T = 1$  TeV track, the average  $p_T$  resolution is  
 766 11%. In the inner most end-cap wheels, Cathode Strip Chambers (CSC) are used instead of MDTs,

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

771 **4.5 Magnet System**

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

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

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

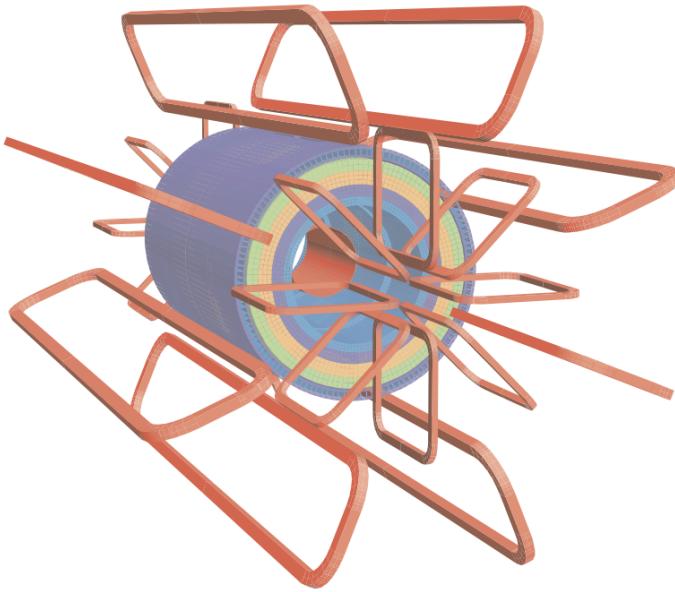


Figure 4.9: Layout of the barrel and endcap toroid magnets [32]

## 790 4.6 Forward Detectors

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

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

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

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

807 **4.7 Trigger and Data Acquisition**

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

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

826 The L1 trigger defines a Region-of-Interest (RoI) for each passing event. The ROI is repre-  
827 sented by the  $\eta$ - $\phi$  detector region where interesting features were identified by the L1 selection  
828 process. Information about the type of feature identified and the threshold which was exceeded to  
829 trigger the L1 response is also recorded. The ROI data is sent to the L2 trigger, which uses all of

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

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

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

## Chapter 5: Particle Reconstruction and Identification

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

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

### 5.1 Inner Detector Tracks

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

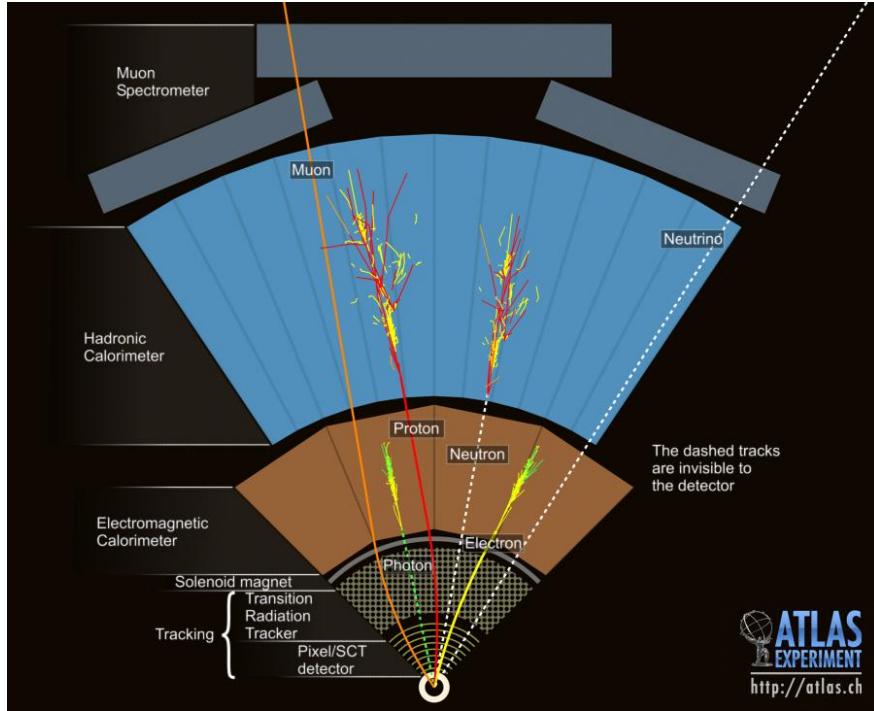


Figure 5.1: Graphic illustrating the various objects and high level features identified by ATLAS object reconstruction, and their interaction with different systems of the ATLAS detector [42]

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

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

## 884   **5.2 Photons and Electrons**

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

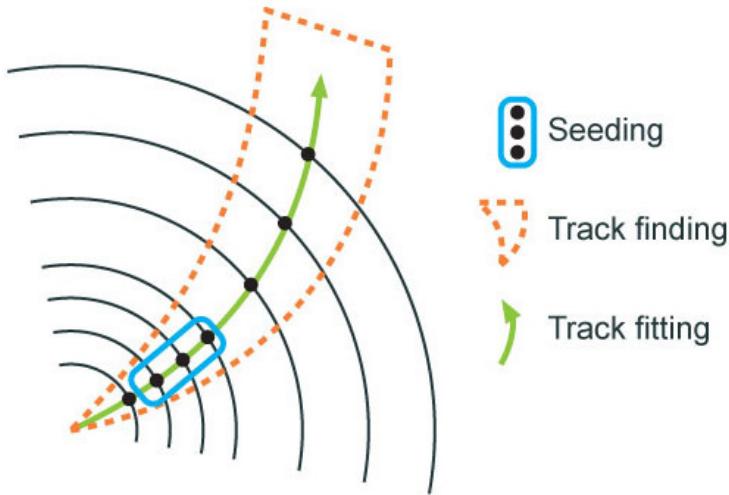


Figure 5.2: Track reconstruction seeding, finding and fitting illustration [43]

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

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

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

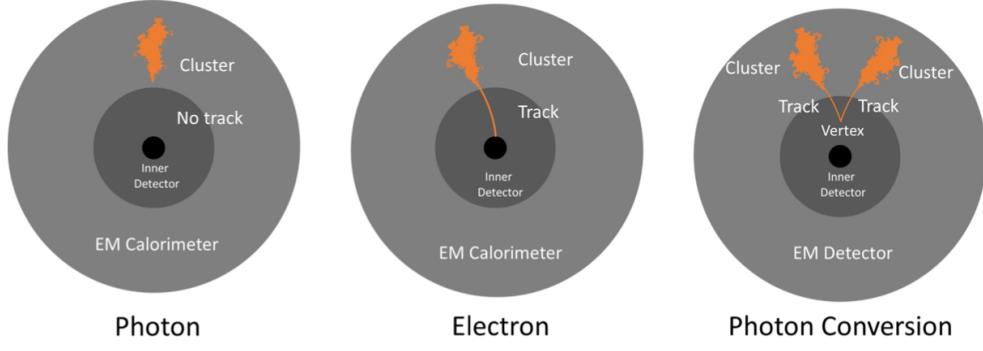


Figure 5.3: Three types of EM object candidates [45].

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  $\chi^2$  fit is

924 performed on each track candidate. Based on the  $\chi^2$  criteria, hits are removed or added such that  
925 the track contains as many hits as possible while satisfying the fit criteria.

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

- 929 • Combined muon: a muon track identified through independent track reconstruction in the  
930 ID and MS, where the combined track is formed using a global refit that uses hit information  
931 from both detectors. Most muons are constructed through an outside-in procedure, in which  
932 a muon track candidate is identified in the MS and then an associated track is found in the ID.  
933 A complementary inside-out procedure is also implemented and identifies additional muons.
- 934 • Segment-tagged muon: an ID track is identified as a muon if when extrapolated out to the  
935 MS (following the inside-out global fit procedure) it is matched to at least one local MS  
936 segment.
- 937 • Calorimeter-tagged muon: an ID track is identified as a muon if it is matched to a calorimeter  
938 energy deposit that is compatible with a minimum-ionizing particle. This muon identifica-  
939 tion has the lowest purity, but it used in regions where the MS has only partial coverage due  
940 to cabling and service access routes.
- 941 • Extrapolated muons: the muon is reconstruction only from the MS track and a requirement  
942 on compatibility with the primary interaction point. The muon track is required to cross at  
943 least two layers of the MS, and three layers in the forward region. These muons are mainly  
944 used to extend muon acceptance into the region  $2.5 < |\eta| < 2.7$  where ID track information  
945 is not available.

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

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

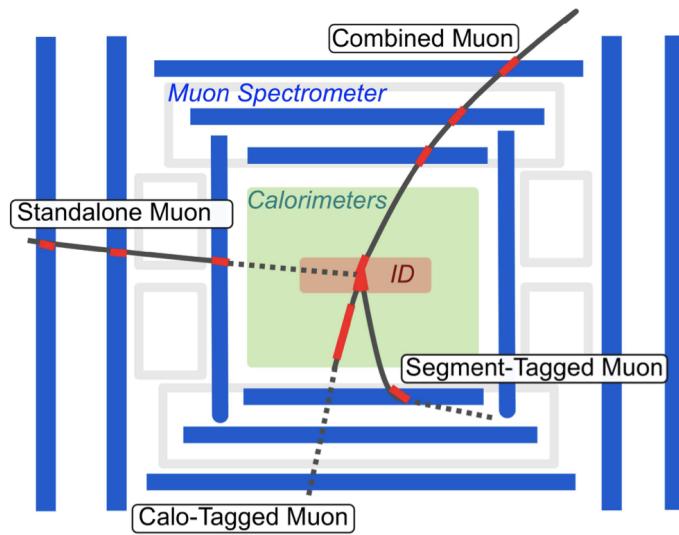


Figure 5.4: Four types of muon track candidates [47].

## 955 5.4 Jets

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

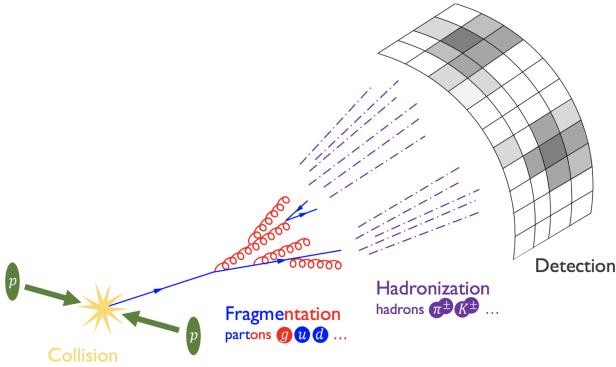


Figure 5.5: The fragmentation and hadronization processes undergone by a quark produced in a proton-proton collision [49].

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

#### 976        5.4.1 Calorimeter Clusters

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

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

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

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

#### 999 5.4.2 Particle Flow Algorithm

1000 The calorimeters provide excellent jet energy resolution for high energy jets. However, the  
1001 granularity of the hadronic calorimeter is restricted to  $0.1 \times 0.1$  in  $\eta \times \phi$ . Combining the infor-  
1002 mation from the calorimeter with tracking information provides superior angular resolution and  
1003 energy resolution. The particle flow (PFlow) algorithm is one of a handful of algorithms which  
1004 can perform this task.

1005 An overview of the process is given in Figure 5.6. Tracks from the ID which are selected for the  
1006 PFlow algorithm are required to have at least 9 hits in the silicon detector, and missing pixel hits in  
1007 places where a hit would be expected. Additionally, the tracks have  $p_T > 0.5$  GeV, and  $|\eta| < 2.5$ .

1008 The algorithm then attempts to match these tracks to EM scale calorimeter topo-clusters. This  
1009 matching is performed using the distance metric

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

1010 where  $\sigma_\eta$  and  $\sigma_\phi$  represent the angular widths of the topo-clusters, and  $\Delta\eta$  and  $\Delta\phi$  represent the  
1011 distance between the track, extrapolated to the second layer of the EM calorimeter, and barycenter  
1012 of the topo-cluster [54]. The topo-cluster closest to the track as measured by  $\Delta R'$  is considered  
1013 matched to the track. If no topo-cluster is found within the cone size of  $\Delta R' = 1.64$ , it is assumed  
1014 that particle which left the track did not form a topo-cluster in the calorimeter.

1015 The PFlow algorithm predicts the expected single topo-cluster energy for a given track, based  
1016 on the track momentum and topo-cluster position. This value is then compared to the observed  
1017 energy of the topo-cluster, and the probability that the particle energy was deposited in more than  
1018 one topo-cluster is evaluated. If necessary, the algorithm adds more topo-clusters to the track/topo-  
1019 cluster system, in order to account of the full shower energy of the track particle.

1020 To reduce the impact of double counting the energy of a given particle by including both its  
1021 tracker and calorimeter energy measurements, the calorimeter energy measurements associated to a  
1022 given track are subtracted from the total calorimeter measurement. If the expected energy deposited  
1023 by the particle exceeds the topo-cluster energy, the full topo-cluster is removed. If the expected  
1024 energy is less than the EM scale energy of all the considered topo-clusters, topo-cluster cells are  
1025 removed one by one, until the full expected energy deposit of the particle has been removed from  
1026 the calorimeter information. The resulting set of tracks and topo-clusters represent the event with  
1027 no double-counting of energy between subdetectors [54]. This information is passed to the jet-  
1028 finding algorithm.

#### 1029 5.4.3 Jet Clustering

1030 When a parton decays in the detector, its energy deposits often result in multiple calorimeter  
1031 clusters. For physics purposes, it is useful to combine clusters that likely resulted from an individ-

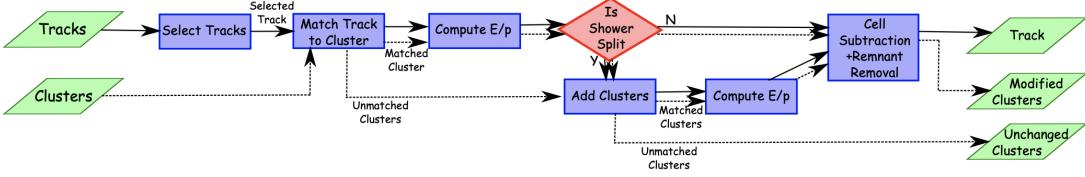


Figure 5.6: A flow chart illustrating the particle flow algorithm progression [54].

1032 ual parton decay, in order to reconstruct the parton. The process of grouping topo-clusters which  
 1033 were produced by the same parton decay is *jet clustering*.

1034 The anti- $k_t$  algorithm is the most commonly used algorithm for jet clustering in ATLAS. The  
 1035 anti- $k_t$  algorithm is based on sequential recombination algorithms [55]. A sequential recombina-  
 1036 tion considers the distance  $d_{ij}$  between objects  $i$  and  $j$  (particles or pseudojets), and the distance  
 1037  $d_{iB}$  between an object  $i$  and the beam line  $B$ . If  $d_{ij}$  between two objects is the smallest distance  
 1038 among those considered,  $i$  and  $j$  are combined into a pseudojet. The process continues until the  
 1039 smallest distance is  $d_{iB}$  at which point the object  $i$  is determined to be a jet and removed from the  
 1040 objects in consideration. The procedure is repeated with the remaining objects until there are none  
 1041 remaining [50].

1042 The anti- $k_t$  algorithm adopts this procedure, but modifies the distance measurements  $d_{ij}$  and  
 1043  $d_{iB}$  to consider the transverse momentum  $k_t$ :

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

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

1044 The addition of the term  $p$  allows adjustments to algorithm which govern the relative power of  
 1045 the momentum versus the geometrical scale  $\Delta_{i,j}$ , which is defined as  $\Delta_{i,j} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$   
 1046 where  $y_i$  and  $\phi_i$  are respectively the rapidity and azimuth of particle  $i$  [50]. The radius parameter  
 1047  $R$  is chosen and determines the geometric cone size [55].

1048 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.

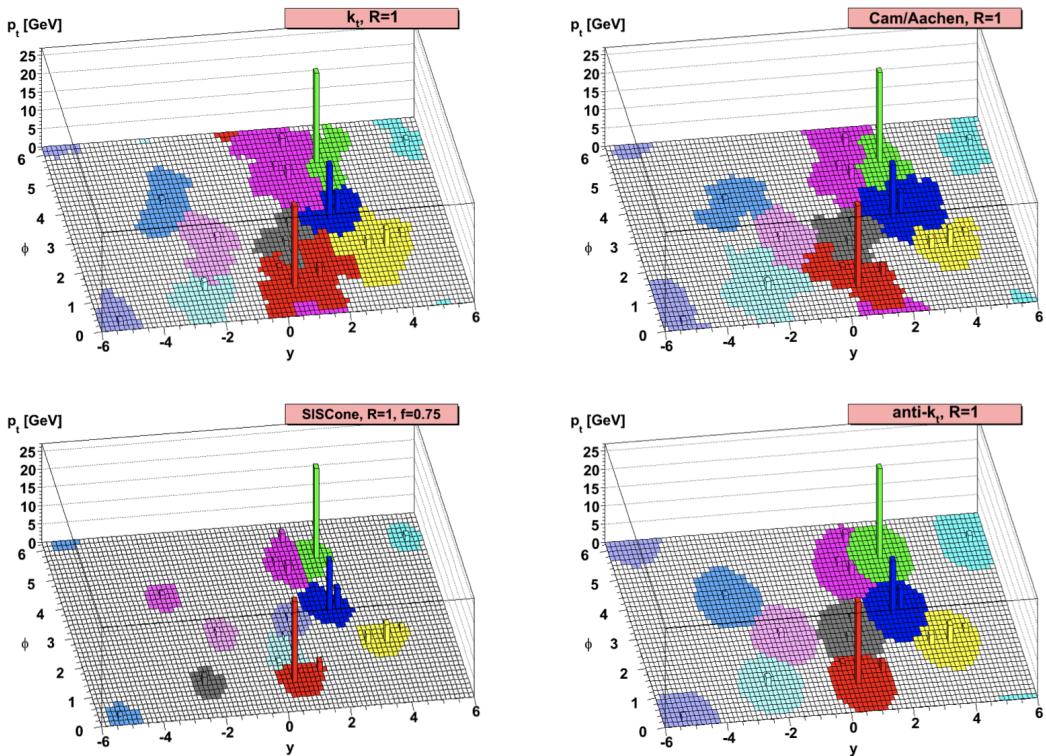


Figure 5.7: A comparison of jet clustering with four different jet algorithms. The anti- $k_t$  algorithm is observed to create the most conical jets, where the shape of the jet is immune to the presence of soft radiation [50].

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

1060 (IRC) safety. Infrared safety implies that the resulting set of jets is unaltered by the presence of  
1061 additional soft particles in the list of seed clusters. As explained above, the anti- $k_t$  algorithm is  
1062 natural infrared safe. Collinear safety requires that the final set of jets is not impacted by collinear  
1063 splitting of one of the jets. If the hardest particle  $p_1$  is split into a collinear pair ( $p_{1a}, p_{1b}$ ) (as  
1064 is common in the fragmentation process for a hard parton), the jet clustering algorithm must  
1065 still recognize ( $p_{1a}, p_{1b}$ ) as the hardest jet in the collision. If another softer particle  $p_2$  with  
1066  $p_{t,1a}, p_{t,1b} < p_{t,2} < p_{t,1}$  is instead considered the hardest particle in the event, a different final  
1067 set of jets would be returned. Collinear safety is a requirement of perturbative QCD, to ensure  
1068 non-divergent higher-order calculations [58]. The anti- $k_t$  algorithm's tendency to cluster hard par-  
1069 ticles first ensures its collinear safety. By satisfying the IRC safety requirement, anti- $k_t$  jets can be  
1070 calculated using perturbative QCD, which improves comparisons with theory.

#### 1071 5.4.4 Ghost Track Association

1072 Once a collection of jets has been created, the jet objects can be studied at both the event-level  
1073 and the jet-level. In the event-level picture, the momentum, energy, and geometric orientation  
1074 of the jets within an event are considered. This yields important information about decay of any  
1075 resonant heavy objects, the total energy in the event, and the distribution of energy amongst the jets.  
1076 At the jet-level picture, the particles composing the jet are considered. The momentum, energy,  
1077 and geometric orientation of the particle tracks provides a low-level picture of the jet, which can  
1078 help determine if the properties of the jet are consistent with standard QCD, or if new physics  
1079 processes might be represented within the patterns found in the constituent particles. Jet-level  
1080 analysis is also widely used in flavor tagging.

1081 For anti- $k_t$  jets with a radius parameter  $R = 0.4$ , one way of studying the jet-level picture  
1082 is through considering the ghost-associated tracks. Track association is process of determining  
1083 which tracks should be considered associated with a given jet. In the ghost association algorithm,  
1084 the anti- $k_t$  clustering algorithm is used for the collection of tracks and calorimeter clusters [59].  
1085 However, the tracks are considered to have infinitesimal momentum (*ghosts*), so their addition to

1086 a jet object does not alter the four-momentum of the jet. This ensures the final jet collection is not  
1087 altered by the presence of the ghost tracks in the reclustering, but information about the associated  
1088 tracks for each reconstructed jet becomes available [60].

1089 Ghost tracks are of particular importance to this analysis, as a means of providing a low-level  
1090 picture of the shape of  $R = 0.4$  jets, and discriminating Standard Model QCD-like jets from dark  
1091 QCD-like jets.

## 1092 5.5 Missing Transverse Energy

1093 A simple principle leveraged in ATLAS physics analyses is checking for conservation of mo-  
1094 mentum among the products of any  $pp$  collisions. The initial state transverse momentum of any  
1095  $pp$  collision is always zero, so the transverse momentum of all final state particles should likewise  
1096 be zero. The missing transverse energy,  $E_T^{\text{miss}}$ , is determined by the magnitude of the negative  
1097 momentum vector sum of all final state objects resulting from the  $pp$  collision.

1098 Specifically, the objects considered in the  $E_T^{\text{miss}}$  calculation are photons, electrons, muons, jets,  
1099 and soft terms. The first four items comprise the hard components of the  $E_T^{\text{miss}}$  calculation, and  
1100 have been discussed previously in this chapter. The final item represents a collection of *soft terms*,  
1101 comprising any detector signals not associated to hard detector objects. These can be based on  
1102 unassociated tracks, or unassociated soft calorimeter clusters. Both are generally not used in the  
1103 same calculation to avoid double counting of soft terms. In this analysis the calorimeter cluster  
1104 soft terms are considered in the  $E_T^{\text{miss}}$  calculation.

1105  $E_T^{\text{miss}}$  can arise due to non-interacting Standard Model objects such as neutrinos, fake sources  
1106 such as mis-reconstructed objects and dead detector regions, or in some theories, non-interacting  
1107 BSM objects such as a dark matter candidate particles. To understand the amount of  $E_T^{\text{miss}}$  at-  
1108 tributable to detector noise and mis-reconstruction,  $E_T^{\text{miss}}$  is studied in  $Z \rightarrow \mu\mu$  where little real  
1109  $E_T^{\text{miss}}$  is expected [61]. As Figure 5.8 illustrates, the resolution of  $E_T^{\text{miss}}$  generally decreases as  
1110  $E_T^{\text{miss}}$  increases, due to detector resolution effects. As  $E_T^{\text{miss}}$  is an important quantity for most dark  
1111 QCD analyses, limitations in the accuracy of the  $E_T^{\text{miss}}$  calculation must be considered.

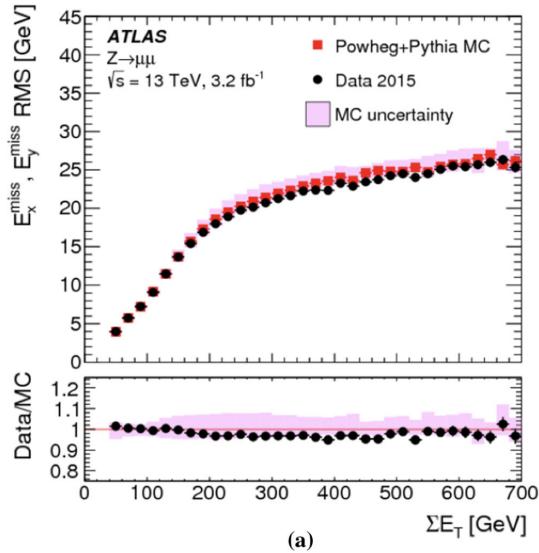


Figure 5.8: A comparison of MC simulation and data for  $Z \rightarrow \mu\mu$  events where real  $E_T^{\text{miss}} = 0$  [61]. The resolution of the missing energy in the transverse ( $x - y$ ) plane is observed to increase with increasing total  $\sum E_T$ .

1112

## **Part III**

1113

## **Search**

1114

1115

## Chapter 6: Monte Carlo and Data

1116        The search for semi-visible jets via s-channel production presented in the following chapters  
1117      is performed with  $139 \text{ fb}^{-1}$  of proton-proton collision data collected by the ATLAS detector dur-  
1118      ing Run 2 (2015 - 2018). The full Run-2 dataset is used for the final interpretation, but Monte  
1119      Carlo simulations of background processes and the semi-visible jet signal process are used in the  
1120      development of the analysis strategy, and to set limits on the observed cross-section of the signal  
1121      model.

1122

1123

## Chapter 7: Machine Learning Tools

### 1124 7.1 Introduction

1125 The search for semi-visible jets presents an opportunity to use novel machine learning (ML)  
1126 tools to uncover patterns in the behavior of dark QCD. The subtlety of the shower differences be-  
1127 tween dark and SM QCD motivates a complex model that can accept high-dimensional low-level  
1128 inputs to best understand key differences between signal and background correlations. Addition-  
1129 ally, the large number of theory parameters which can be chosen arbitrarily and affect the shape of  
1130 the dark QCD shower motivate exploring a data-driven machine learning approach, which could  
1131 be sensitive to a wider variety of dark QCD behavior.

1132 To this end, two machine learning approaches are developed for this search, which are used  
1133 in tandem. The first is a supervised ML method where the ML algorithm is built to maximize  
1134 exclusion sensitivity to the specific generated SVJ signal models used in this analysis. Here, su-  
1135 pervised refers to the use of full and correct labels for all events considered during model training,  
1136 which necessitates training over simulated data. The second is a semi-supervised method, where  
1137 training of the model is data-driven and labels are only partially provided during training. The  
1138 semi-supervised ML algorithm broadens the discovery sensitivity of the search, and reduces the  
1139 dependence on the exact theory parameters chosen for signal model simulation.

1140 The two different ML algorithms used in this approach will be explained in the following  
1141 sections, along with their application in the SVJ analysis strategy.

1142 7.1.1 Particle Flow Network (Supervised)

1143 **Architecture Fundamentals**

1144 A Particle Flow Network (PFN) [62] architecture is selected for two reasons: *permutation in-*  
1145 *variant input modeling* to best describe the events consisting of an unordered set of particles, and a  
1146 *low-level input modeling* using tracks to take advantage of the available high-dimensional informa-  
1147 tion to best exploit available correlations within the event. Permutation invariant input modeling  
1148 is an architecture priority as ordered input modeling has been observed to bias the performance of  
1149 low-level modeling tools as in [63]. Low-level input modeling is an architecture priority to capture  
1150 the intricacies of dark QCD showers which may not express themselves in higher level variables,  
1151 as explored in [64]. A comparison to a high-level *boosted decision tree* (BDT) is available in  
1152 Appendix A.3.

1153 The PFN is used to model input events as an unordered set of tracks. Given the inherently  
1154 unordered and variable-length nature of particles in an event, this choice of modeling as a *set*  
1155 can enable the model to better learn the salient features of the dataset that enable a signal-to-  
1156 background classification. Constructing the PFN involves the creation of new basis variables  $\oplus$   
1157 for each particle in the event. Permutation invariance is enforced by summing over the  $\oplus$  basis for  
1158 every particle in the event to create a new permutation invariant latent space basis  $O$ . Finally the  
1159 classifier  $F$  is a function of the sum over this new basis. The creation of the latent space basis  $O$   
1160 from  $M$  particles  $\vec{p}$  with  $d$  features each can be expressed as:

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

1161 where  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^l$  is a per particle mapping, with  $l$  being the dimension of the new basis  $O$ .  
1162 Figure 7.1 gives a graphical representation of the use of summation in the PFN over per-particle  
1163 information to create a permutation-invariant event representation.

1164 Figure 7.2 provides an annotated diagram of the PFN architecture as used in this analysis.

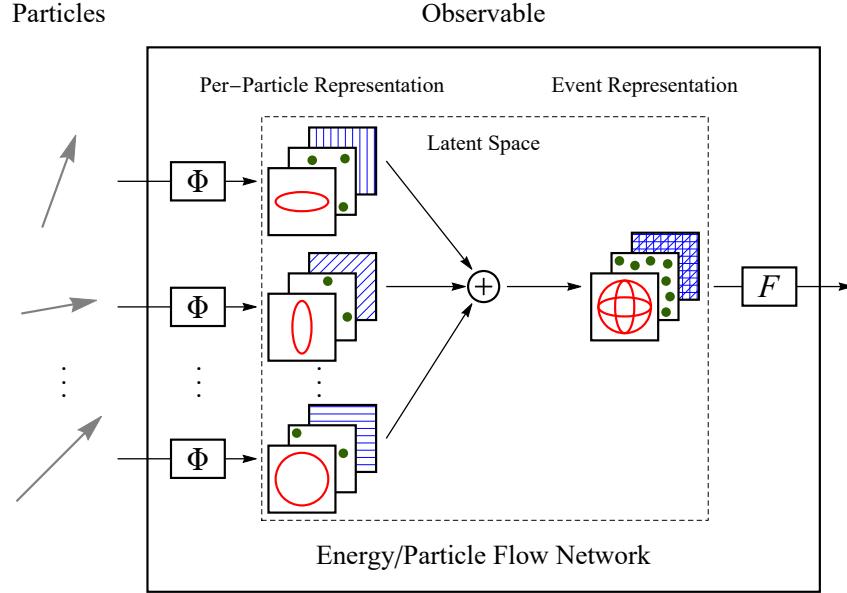


Figure 7.1: The Energy/Particle Flow Network concept, from Ref. [62].

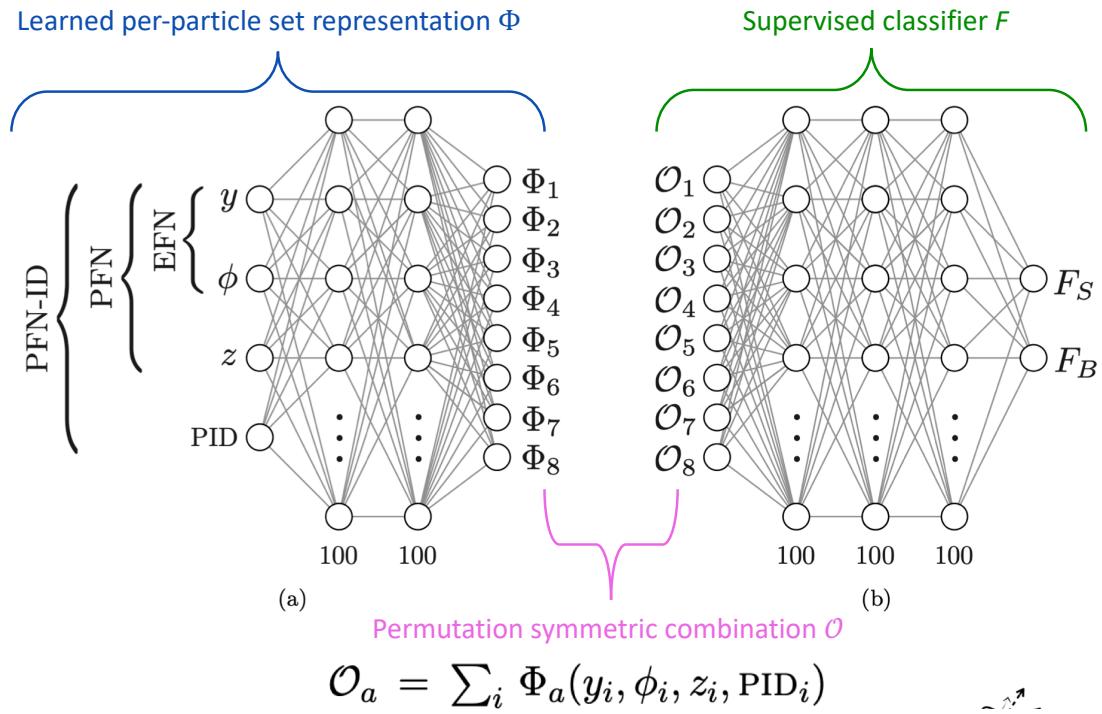


Figure 7.2: An annotated diagram of the PFN architecture.  $y$  and  $\phi$  represent geometric information for the input particles,  $z$  represents energy information, and PID encompasses any other particle ID information in the input.

1165 **Input Modeling, Scaling, and Rotation**

1166 In this implementation, the particle input information comes from all tracks associated to the  
1167 leading and subleading jets. The track association method is Ghost association, as discussed in  
1168 Section 5.4.4. A single jet tagger strategy was also considered, but utilizing tracks from both  
1169 leading jets creates a complete low-level picture of the event, which both focuses on the objects  
1170 most likely to be associated to the decay of the dark quark (as will be justified in Chapter 8)  
1171 and the relationship between those objects. If we consider the dijet topology of semi-visible jets  
1172 as illustrated in Figure 7.3, the advantage of modeling both leading jets simultaneously becomes  
1173 clear. In the semi-visible jet model presented in [64],  $E_T^{\text{miss}}$  in the event is expected to arise due to  
1174 an imbalance in the number of visible tracks of the two jets associated to the dark quark decay.

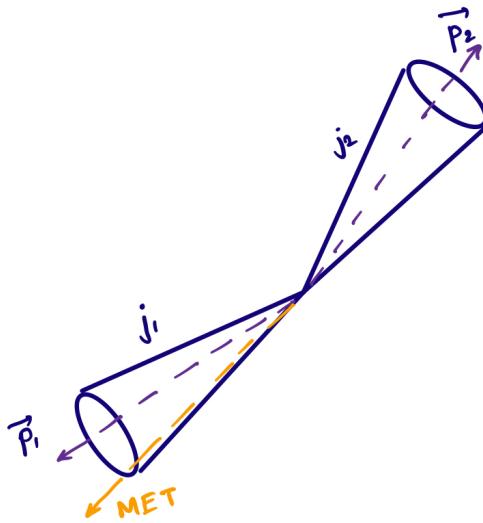


Figure 7.3: A illustration of the expected dijet behavior of semi-visible jets, where one jet is closely aligned with  $E_T^{\text{miss}}$ .

1175 Each track is described using six variables: the four-vector of the track ( $p_T, \eta, \phi, E$ ), and the  
1176 track displacement parameters  $d_0$  and  $z_0$ , where  $d_0$  measures displacement in the radial direction  
1177 from the beamline and  $z_0$  measures displacement along the beamline from the primary interac-  
1178 tion point. Figure 7.4 illustrates these coordinates. Up to 80 tracks per jet are allowed, which is  
1179 a threshold chosen to generally include all the tracks in the jet, which leads to maximal perfor-

1180 mance. Figure 7.5 shows the track multiplicity in the leading and subleading jet for the signal and  
 1181 background samples used in training.

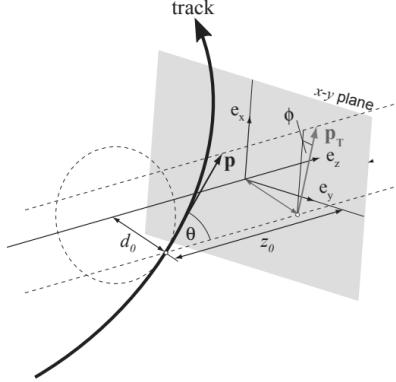


Figure 7.4: Illustration of track coordinates  $d_0$  and  $z_0$ .

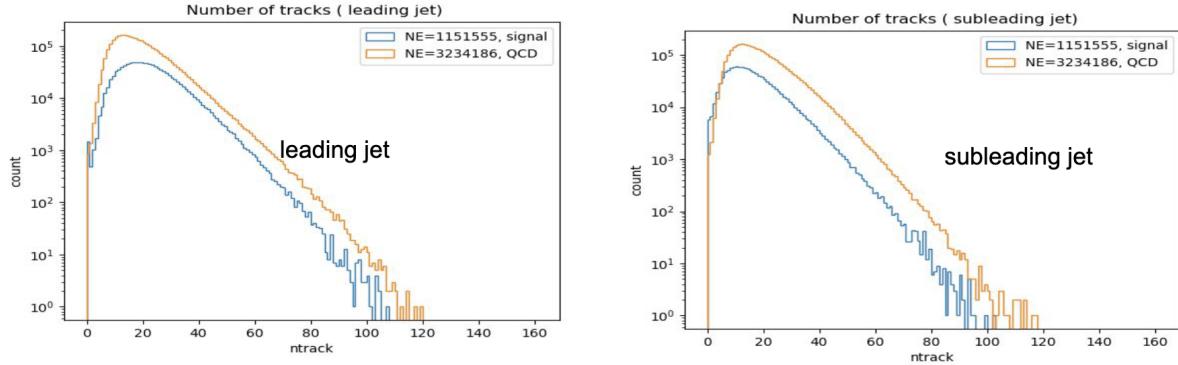


Figure 7.5: Distributions of the track multiplicity in the leading and subleading jets, comparing signal and background PFN training samples.

1182 These tracks (up to 160 total) are the input to the PFN. Referencing Equation 7.1, this corre-  
 1183 sponds to  $M = 160$  and  $d = 6$ . The two leading jets and their associated tracks are rotated so  
 1184 that the center of the system is aligned with  $(\eta, \phi) = (0, 0)$ . Each track is normalized to its relative  
 1185 fraction of the total dijet system energy and transverse momentum- this enforces agnosticism to the  
 1186 total energy and transverse momentum of the event. The rotation and scaling are motivated by the  
 1187 procedures described in [62] to improve the optimality of the PFN learning. Figure 7.6 illustrates  
 1188 the rotation process.

1189 Finally, each of the 6 track variables is scaled so that its range is  $[0,1]$ . This is a common

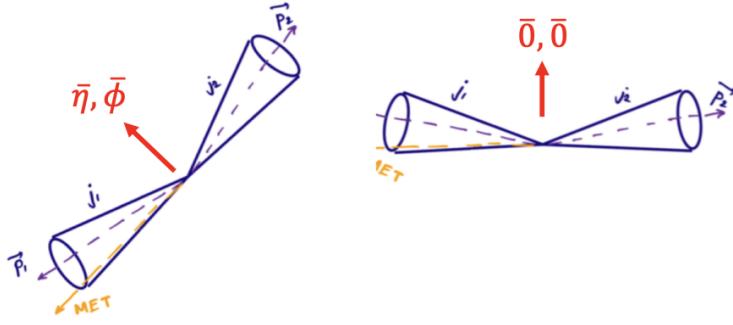


Figure 7.6: A diagram demonstrating how the two jet system is rotated in  $(\phi, \eta)$ .

1190 preprocessing step that ensures the input data is bounded over a similar range, so that arbitrarily  
 1191 large values don't develop an outsized impact on the model. Figure 7.7 show each of 6 track  
 1192 variables before and after scaling and rotation have been applied, demonstrating the impact of  
 1193 these procedures, as well as the track level similarities differences between the background SM  
 1194 QCD processes and the signal SVJ processes. Figure 7.8 illustrates that the data is well modeled  
 1195 by the MC at track level.

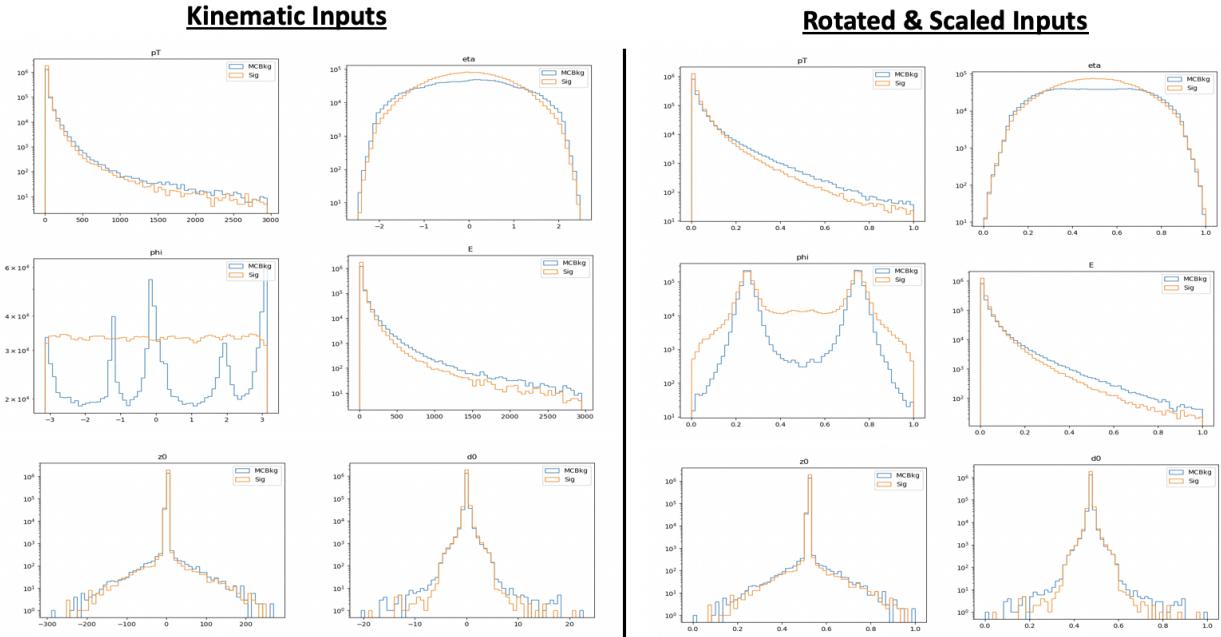


Figure 7.7: The 6 PFN track variables in background MC and signal MC. There are some differences between signal and background, but the track kinematics are largely similar.

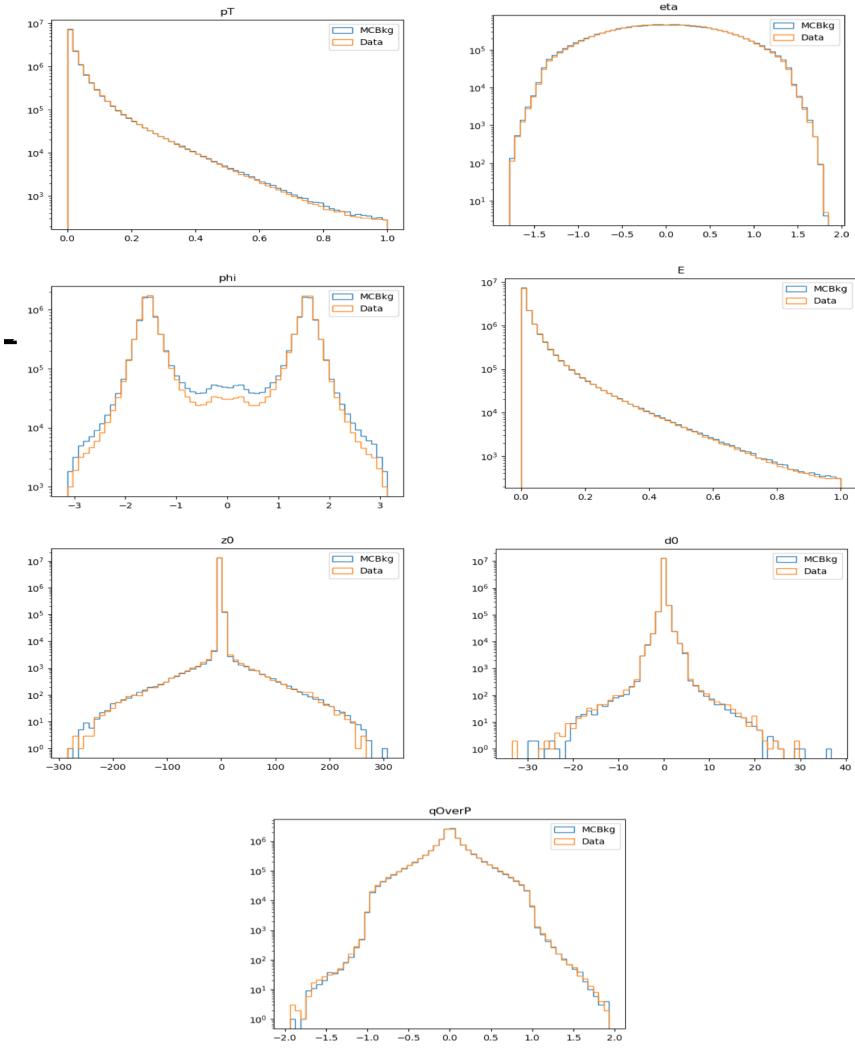


Figure 7.8: The 6 PFN track variables in data and background MC, after the scaling and rotation procedure is applied. There is excellent modeling of the data by the MC within the track variables. The slight discrepancy in the  $\phi$  distribution is due to the modeling of dead TileCal cells by the QCD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

1196 **Training**

1197 As seen in Figure 7.2, two separate architectures are defined and combined to do the super-  
1198 vised training. The PFN uses a masking layer to suppress any zero-padded inputs, making the  
1199 architecture length agnostic. The masking layer ignores any all-zero inputs in the summation layer.  
1200 Additionally, The summation layer in the PFN enforces permutation invariance, so the network is  
1201 unordered. The  $\Phi$  network has 3 dense layers of dimensionality 75 with RELU activation, with  
1202 27.5k trainable parameters and an output  $\Phi$  latent space dimension of 64.

1203 The classifier  $F$  network similarly has 3 dense layers with 75 nodes with RELU activation, and  
1204 a final softmax layer to determine the event-level classification with a categorical cross-entropy  
1205 loss. The Adam optimizer is used with an initial learning rate of 0.001.

1206 The PFN is trained in a fully supervised way using SVJ signal MC and QCD MC events. Al-  
1207 though several SM processes are expected to contaminate the SR (see Chapter 8), QCD is the dom-  
1208 inant background. Training against a QCD-only sample is determined to produced better results  
1209 than training on a more complete background - when training with a background which repre-  
1210 sents samples that are more enriched in  $E_T^{\text{miss}}$ , the ability of the PFN to identify high  $E_T^{\text{miss}}$  signals  
1211 is reduced. When training with a QCD-only background, there is greater contamination from  
1212  $E_T^{\text{miss}}$  enhanced backgrounds in the final SR - however the increased signal acceptance means that  
1213 overall sensitivity is still higher with a QCD-only training. This can be seen in the comparison of  
1214 output classifier distributions in Figure 7.9.

1215 500k events from both background and signal are used in training, where the signal is a com-  
1216 bined file of all simulated signal points and the full QCD background which is sampled according  
1217 to it's MC weights to produce the proper  $p_T$  input shape. A study was done to check the optimality  
1218 of the inclusive signal model PFN as compared to one trained on high and low  $R_{inv}$  points sepa-  
1219 rately, to better capture the differences in high and low  $E_T^{\text{miss}}$  across signals and backgrounds, but  
1220 a small effect is found and the decision is taken to keep the inclusive model (Appendix A.2).

1221 The network is trained for 100 epochs. A train/test/validation split of 78%, 20%, and 2% is used  
1222 for the final PFN training. Figure 7.10 shows the loss during training, which is stable and flattens

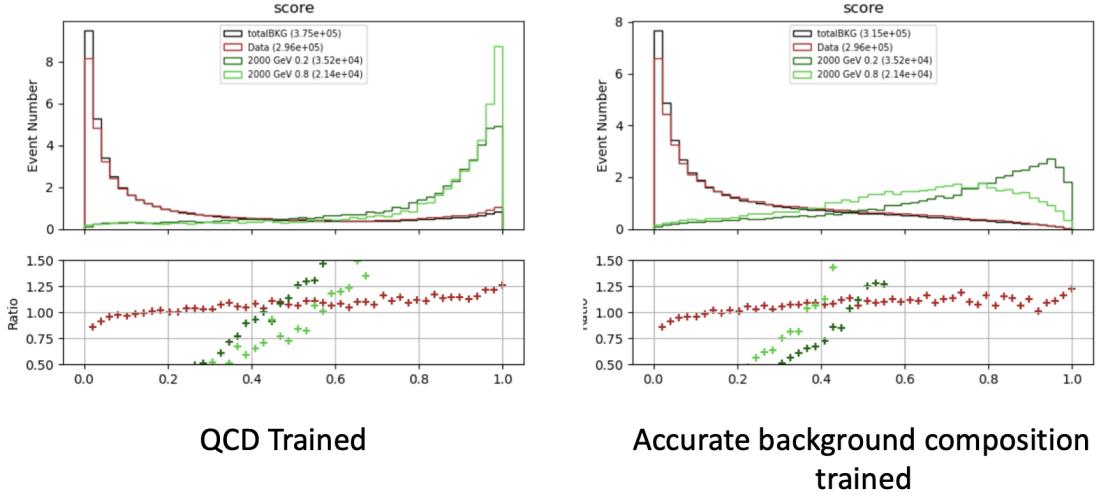


Figure 7.9: PFN score for background MC, data, and signal, comparing a PFN training on QCD-only vs all-background MC samples. The average AUC for the QCD-only training (left) is 0.93, while the average AUC for the mixed background training (right) is 0.84. The sensitivity estimate across the grid is 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.

1223 by the end of training, and the final evaluated losses that provide signal-background discrimination  
 1224 over the test set.

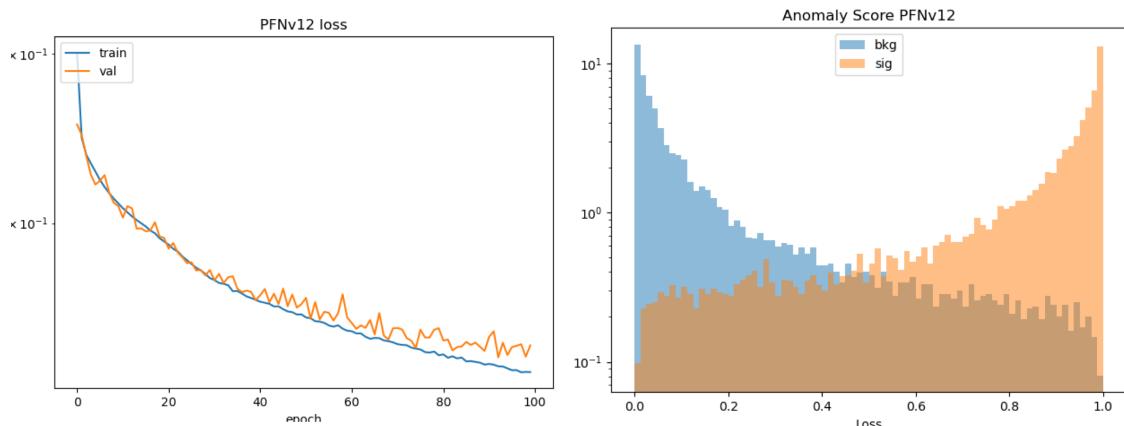


Figure 7.10: PFN architecture loss during training as a function of epoch (left) and the evaluated loss over the signal and background (right).

1225 Optimization studies were performed on the PFN, varying the number of training epochs, num-  
 1226 ber of training events, batch size, learning rate, number of neurons, and dimension of the  $\Phi$  space.  
 1227 A summary of these studies is presented in Appendix A.2. The model presented here represents an  
 1228 optimal choice across these parameters.

1229 **Performance**

1230 The performance of the PFN can be assessed via the area-under-curve (AUC) of the receiver  
1231 operating characteristic (ROC) associated to evaluating the PFN on the test set of signal and back-  
1232 ground events. Figure 7.11 shows the ROC curve of the PFN when classifying the QCD back-  
1233 ground from the combined signal, with an AUC of 0.93. Figure 7.12 shows the AUC of the PFN  
1234 across the SVJ signal grid, demonstrating strong discrimination capability even in the varying cor-  
1235 ners of phase space. Figure 7.13 shows the output score distribution in two signals, data, and the  
1236 total background MC. A selection of PFN score  $> 0.6$  for all SR events is chosen to maximize  
signal sensitivity across the grid.

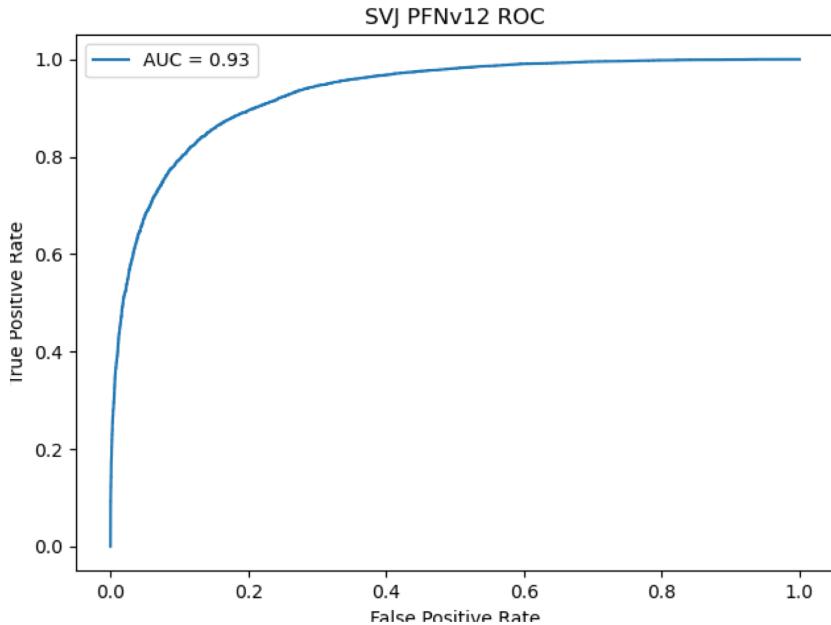


Figure 7.11: ROC the PFN score for combined signal (true positive) and QCD background (false positive).

1237

1238 Another supervised approach was studied using a BDT as the primary selection tool, trained  
1239 over high-level variables describing each event. Studies comparing the PFN and BDT approaches  
1240 are provided in Appendix A.3. Ultimately the low-level high-dimensional approach offered by the  
1241 PFN was selected for its increased performance and lessened kinematic dependence.

1242 Appendix A shows more studies on the ML methods and comparisons of varying approaches.

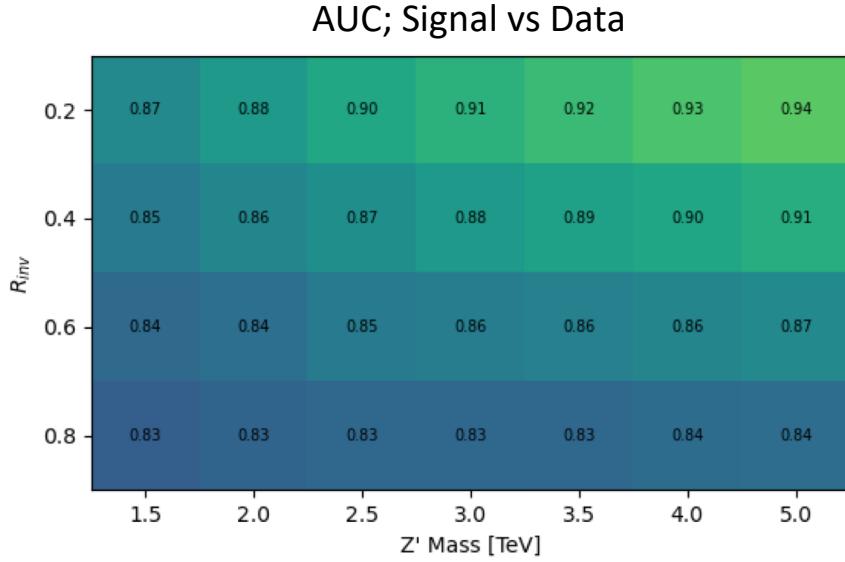


Figure 7.12: AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD-only training sample.

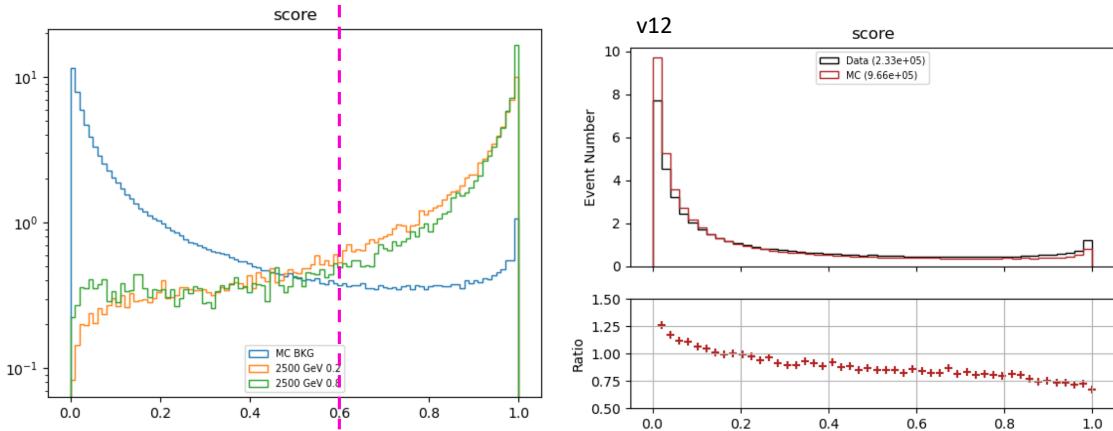


Figure 7.13: PFN score for two signals and the total background MC (top), and between data and MC (bottom). The difference between data and MC efficiency is minimal (< 5%).

1243 7.1.2 ANTELOPE (Semi-supervised)

1244 The semi-supervised analysis approach broadens the discovery sensitivity of the search through  
1245 the use of semi-supervised ML, where training of the model is data-driven and labels are only  
1246 partially provided during training. While broad sensitivity is a general key goal of LHC searches,  
1247 it is particularly motivated in the case of dark QCD models, which can lead to widely varying  
1248 topologies depending on the values of model parameters. In the case of SVJs, the  $R_{inv}$  fraction in  
1249 the jet can dramatically vary the  $E_T^{\text{miss}}$ , shower shape, and other key features, making it difficult to  
1250 find a single standard analysis variable that can distinguish all signal topologies from QCD.

1251 **Architecture Fundamentals**

1252 The model-independent search region of this analysis is implemented with a novel ML ap-  
1253 proach that builds on the ANTELOPE architecture to construct a tool that is capable of performing  
1254 low-level anomaly detection with permutation-invariant inputs. This tool, referred to as **ANomaly**  
1255 **deTEction on particLe flOw latent sPacE (ANTELOPE)**, is a custom solution designed for this  
1256 analysis.

1257 ANTELOPE uses the supervised signal vs. background training of the PFN network described  
1258 in the previous section to generate a permutation invariant latent space that is representative of the  
1259 original input variables, encodes the input events into these latent space variables  $O$ , and trains a  
1260 *variational autoencoder* (VAE) over the events modeled as PFN latent space variables. A VAE  
1261 is a common architecture used for anomaly detection and data-driven ML training. It has been  
1262 used in previous ATLAS searched to model jet level information, such as the search presented  
1263 in [65] using the recurrent architecture described in [63]. One of the limitations of a recurrent  
1264 architecture is the need to order the low level inputs, which affects the performance of the tool. Jet  
1265 constituent information is intrinsically unordered, and therefore a permutation invariant approach  
1266 removes this element of arbitrary decision making from the modeling process. A visual example  
1267 of the ANTELOPE inputs is given in Figure 7.14.

1268 The input to the model is the same 6 track variables for the leading 160 tracks of the leading

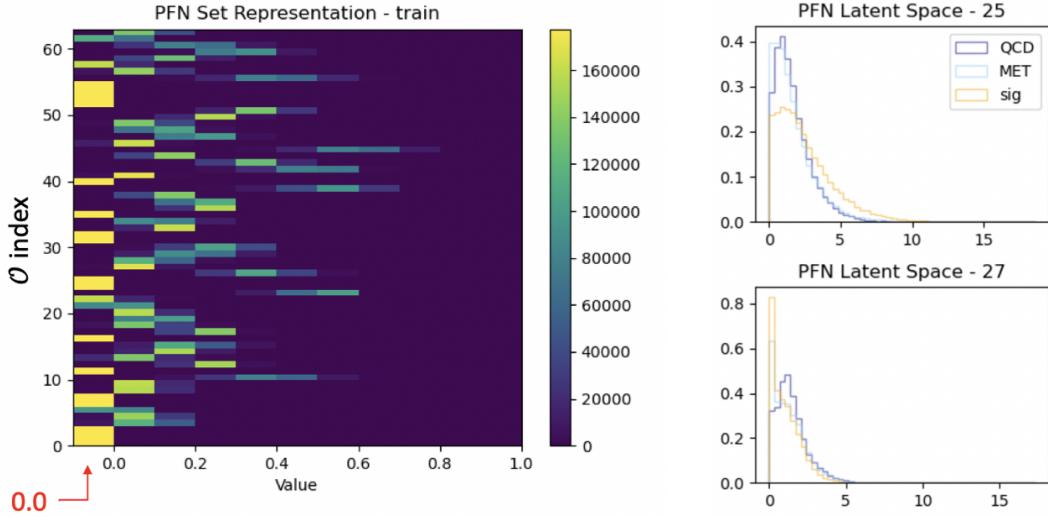


Figure 7.14: A visual representation of the 64 PFN latent space variables which create the input of the VAE component of ANTELOPE. The left shows a 2D histogram of the PFN latent space index (0-63) versus the value assumed by that index. The right shows 1D histograms of two particular PFN latent space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

1269 two jets, as presented in Section 7.1.1. The track information is encoded to the PFN  $\Phi$  latent  
 1270 space using the pre-trained  $\Phi$  network (trained according to the steps outline in Section 7.1.1. The  
 1271 resulting  $\Phi$  basis is summed to created the fixed length symmetric representation  $O$ . The VAE is  
 1272 then trained in an unsupervised way using inputs encoded to  $O$  from data events only. The VAE is  
 1273 given no knowledge of the signal model during training. It is able to perform anomaly detection  
 1274 through an encoder stage which does a lossy compression on the input to a lower-dimensional  
 1275 latent space, and a decoder stage that samples from that latent space and generates an output of the  
 1276 original dimensionality. By using the reconstruction error as a loss, this process enables the VAE  
 1277 to develop a knowledge of the underlying data structure, thereby isolating new out-of-distribution  
 1278 events by their high reconstruction error. This strategy is semi-supervised because the tool has  
 1279 some knowledge of correct labels (eg. through the PFN latent space embedding) but is followed  
 1280 by a data-driven unsupervised stage.

1281 Figure 7.15 provides a diagram of the ANTELOPE architecture.

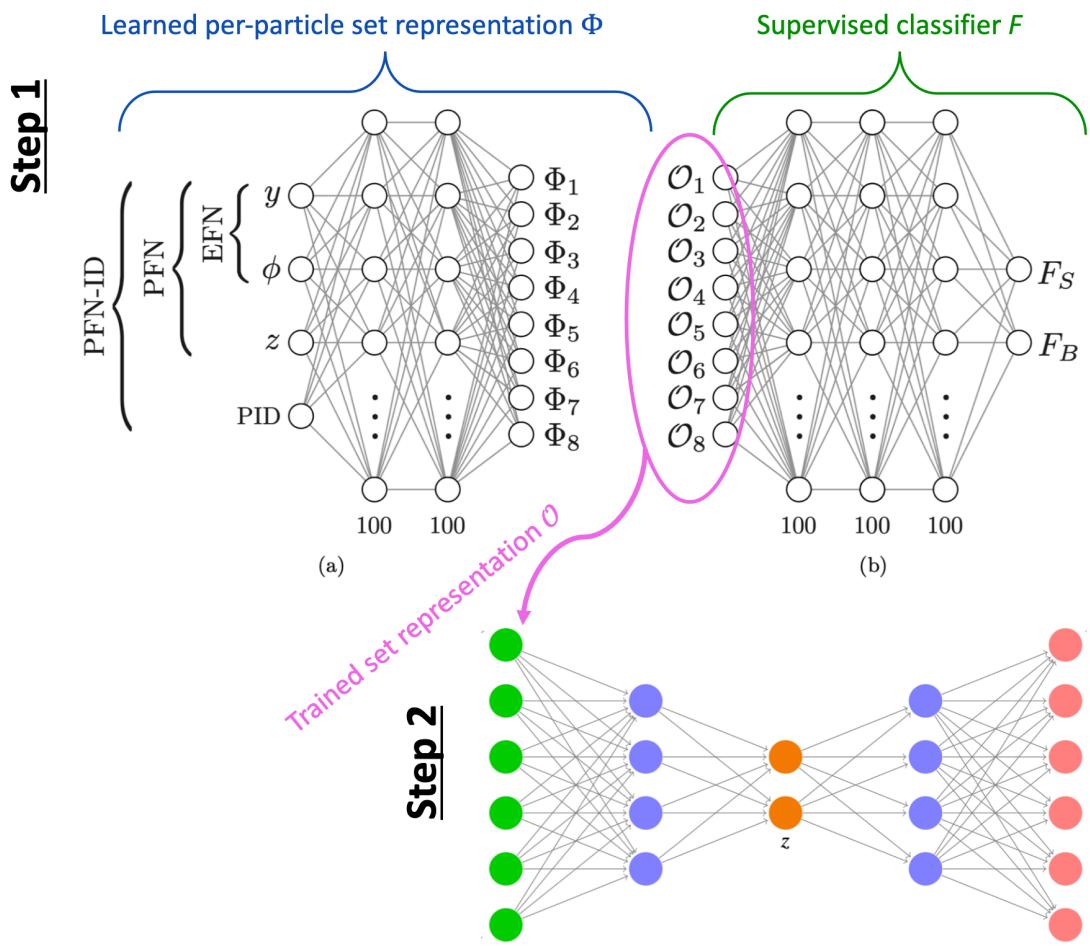


Figure 7.15: An annotated diagram of the ANTELOPE architecture.

1282 **Training**

1283 The VAE stage of the ANTELOPE network is trained directly over a subset of data events  
1284 at preselection (6.7 million available, 500,000 used, with a 80% / 20% training/test split). The  
1285 input dimensionality of the VAE has to match the encoded  $\Phi$  dimension of the PFN, in this case  
1286 64. The encoder has an encoding layer that brings the dimensionality to 32, and a final layer that  
1287 compresses to the latent space dimension of 12. The network is trained for 50 epochs, with a  
1288 learning rate of 0.00001. The loss  $\mathcal{L}$  is the sum of two terms, the mean-squared error (MSE) of  
1289 input-output reconstruction, and the Kullback-Leibler divergence (KLD).

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

1290 As the PFN inputs are sufficiently normalized to remove any spurious information from train-  
1291 ing, no additional normalization is applied to the PFN encoded inputs. The final ANTELOPE score  
1292 used in the analysis is produced by applying a log + sigmoid transformation function to the total  
1293 evaluated loss  $\mathcal{L}$ .

Figure 7.16 shows the loss during training.

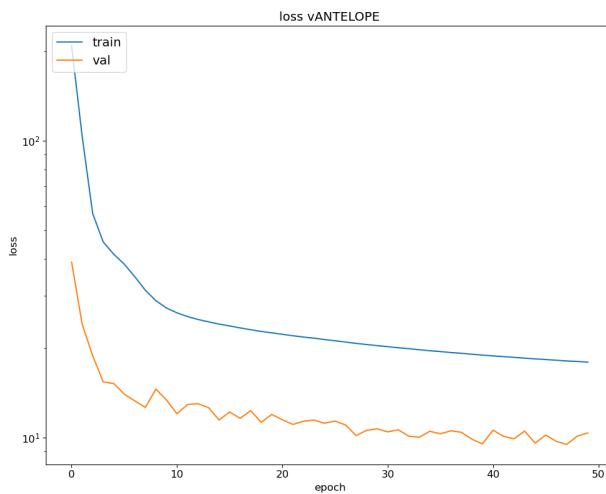


Figure 7.16: ANTELOPE architecture loss during training as a function of epoch.

1294

1295 **Performance**

1296 As with the PFN, the ANTELOPE performance is assessed via the area-under-curve (AUC) of  
1297 the receiver operating characteristic (ROC) associated to evaluating the ANTELOPE on the test  
1298 set of signal and background events. Figure 7.17 shows the output score distribution in data and  
1299 total background MC, showing a very flat ratio and motivating the use of MC for studies of the  
1300 ANTELOPE score.

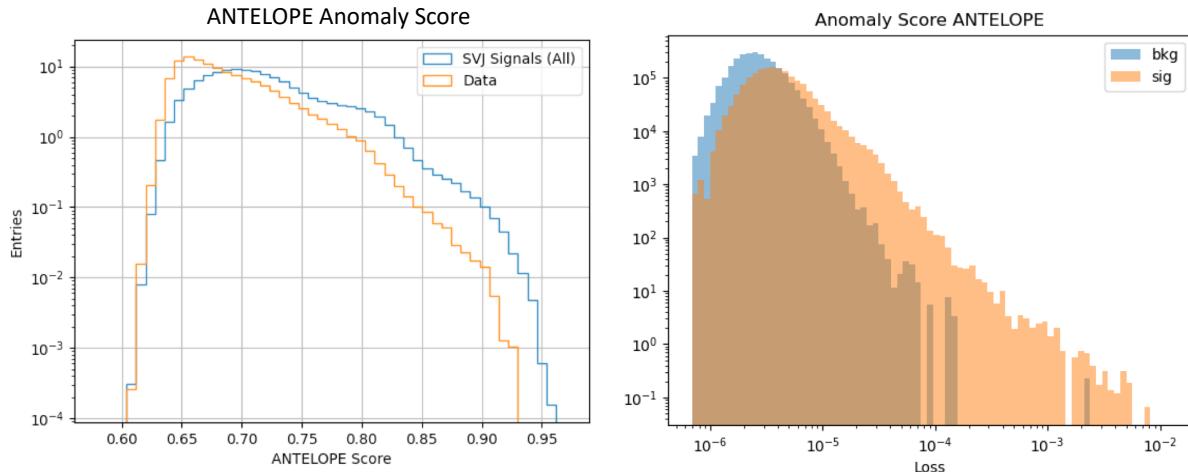


Figure 7.17: ANTELOPE score distribution comparing data and the total background MC (left), with good agreement observed between data and simulated background, and comparing all background MC to signals (right), revealing good discrimination power.

1301 Figure 7.18 shows the AUC of the ANTELOPE across the SVJ signal grid, demonstrating  
1302 strong discrimination capability even in the varying corners of phase space. Compared to the  
1303 supervised PFN method, the ANTELOPE is not as performant (as expected due to the absence of  
1304 signal model in training). However, a selection on events with high ANTELOPE score nonetheless  
1305 provides a 10-40% increase in signal significance by removing background and isolating the long  
1306 tail of anomalous events.

1307 **Model Independence** The unsupervised component of training the ANTELOPE network is ex-  
1308 pected to give it a more generalized sensitivity to new physics with  $E_T^{\text{miss}}$  and jet activity, beyond  
1309 the scope of the SVJ grid. To assess this, alternative signal models are evaluated with the trained

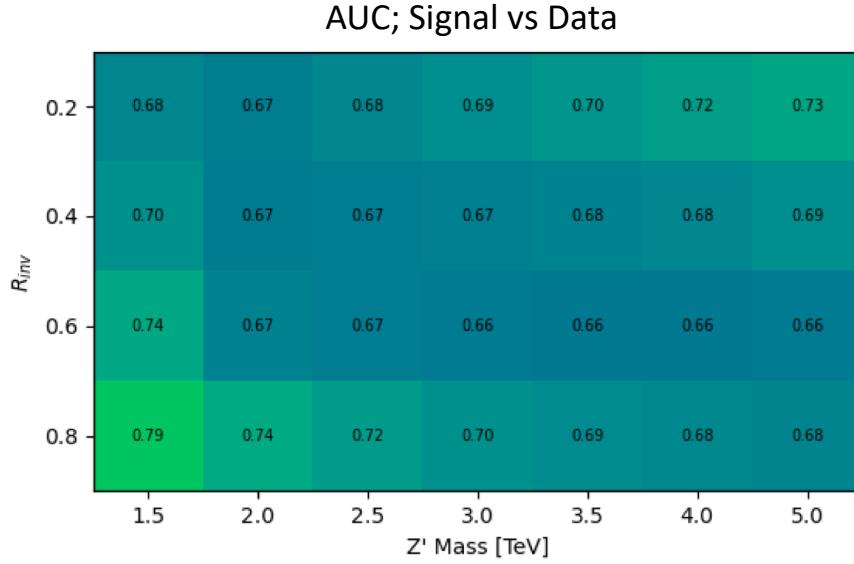


Figure 7.18: AUC from the ANTELOPE score for each signal in the SVJ grid.

1310 ANTELOPE network, as optimized for the SVJ grid, and their sensitivity in the analysis selection  
1311 is evaluated.

1312 The following alternate signal models were considered:

- 1313 •  $Z' \rightarrow t\bar{t}$
- 1314 •  $W' \rightarrow WZ$
- 1315 • Gluino pair production  $\rightarrow R\text{-hadron} + \text{LSP}$  ( $E_T^{\text{miss}}$ ) with gluino masses 2000/3000 GeV, LSP  
1316 mass 100 GeV, and lifetime 0.03 ns (LSP = *lightest supersymmetric particle*)
- 1317 • Emerging jets s-channel with mass 1000 GeV and lifetime 1ns

1318 Figure 7.19 shows the distribution of these signals in several key analysis variables, namely  
1319  $E_T^{\text{miss}}$ , the PFN score, and the ANTELOPE score. This comparison reveals that ANTELOPE is  
1320 sensitive to  $E_T^{\text{miss}}$  in the event; it classifies signals with no real  $E_T^{\text{miss}}$ , like the all-hadronic  $Z'$   
1321 and  $W'$  decays (given our imposed lepton veto) as data-like, but the distributions for signals with  
1322  $E_T^{\text{miss}}$  such as SVJs, R-hadrons, and emerging jets have distributions with higher anomaly score  
1323 tails.

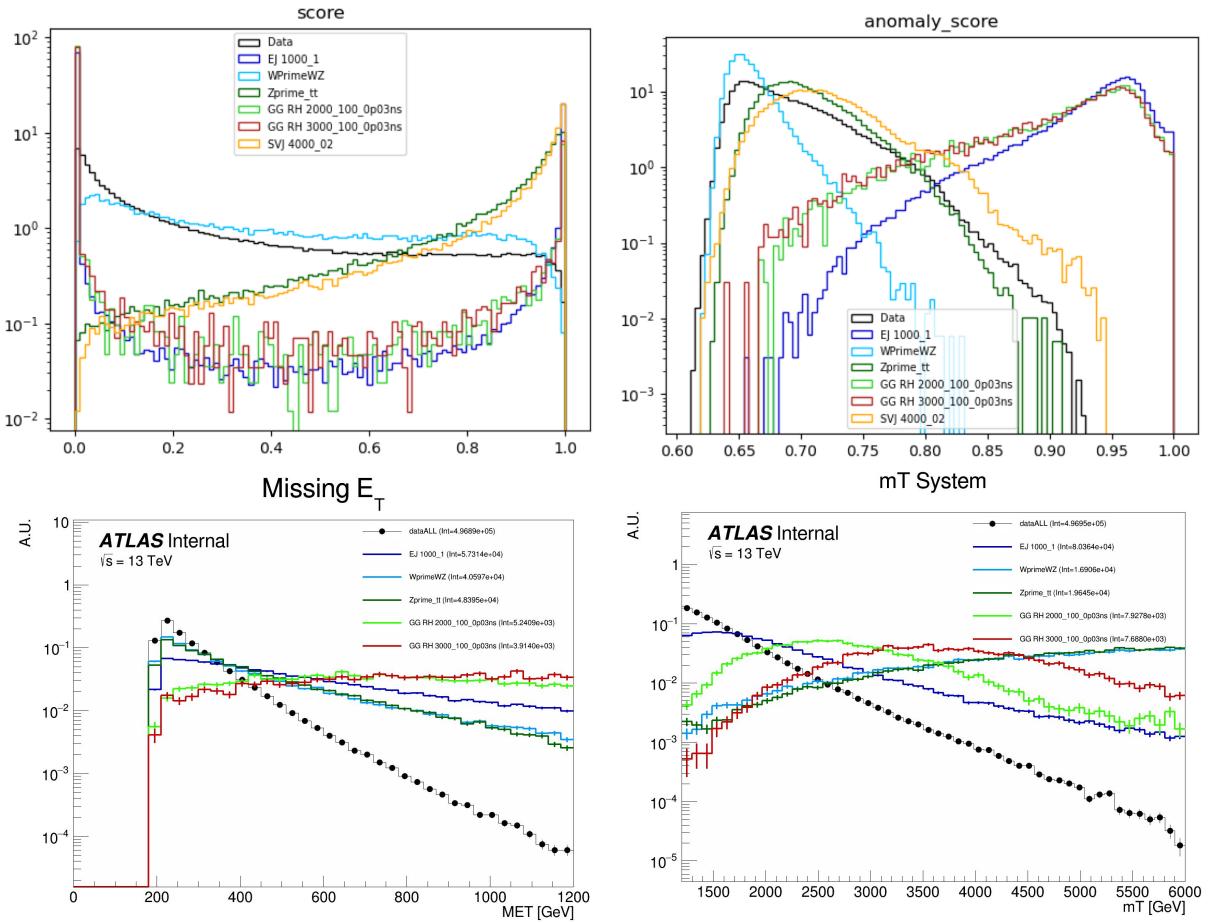


Figure 7.19: Comparing data and the alternate signal models for the PFN score (top left), ANTELOPE score (top right),  $E_T^{\text{miss}}$  (bottom left), and  $m_T$  (bottom right). The emerging jet signal is an example of the gain of the model-independent ANTELOPE approach, where it has a bimodal shape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

1324      Figure 7.20 shows a comparison of the sensitivity of the PFN and ANTELOPE regions across  
 1325     a variety of signals, including the combined SVJ signal used to train the PFN. The benefit of the  
 1326     unsupervised stage of ANTELOPE in enhancing model independence is clearly seen through the  
 1327     boost in performance for other signal models, namely the gluino and emerging jet signals, which  
 1328     have more  $E_T^{\text{miss}}$  than the W' and Z' signals (all-hadronic) that were also tested. As commented  
 1329     above, the PFN outperforms ANTELOPE as expected, because it was designed explicitly for the  
 1330     task of classifying SVJs from background, demonstrating the power of supervised learning for the  
 1331     model-specific approach.

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

	<b>EJ 1000</b>	<b>WprimeWZ</b>	<b>Zprime tt</b>	<b>GG RH 2000</b>	<b>GG RH 3000</b>	<b>SVJ</b>
<b>PFN</b>	0.57	0.65	1.92	0.30	0.32	1.97
<b>ANTELOPE</b>	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  $E_T^{\text{miss}}$  than the SVJs.

1332      Studies on the ANTELOPE architecture and comparisons to other methods can be found in  
 1333     Appendix A.1.

1334

1335

## **Chapter 8: Analysis Strategy**

1336

## Conclusion or Epilogue

1337        Use this page for your epilogue or conclusion if applicable; please use only one of the titles  
1338      for this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1339      applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1340      page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1341      Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1342      use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1343      epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1344      you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1345      of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1346      conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1347      it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1348      this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1349      applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1350      page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1351      Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1352      use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1353      epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1354      you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1355      of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1356      conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete

1357 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1358 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1359 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1360 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1361 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1362 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1363 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1364 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1365 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1366 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1367 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1368 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1369 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1370 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1371 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1372 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1373 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1374 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1375 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1376 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1377 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1378 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1379 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1380 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1381 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1382 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1383 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,

1384 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1385 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1386 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1387 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1388 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1389 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1390 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1391 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1392 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1393 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1394 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1395 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1396 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1397 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1398 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1399 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1400 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1401 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1402 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1403 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1404 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1405 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1406 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1407 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1408 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1409 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1410 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.

1411 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1412 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1413 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1414 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1415 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1416 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1417 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1418 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1419 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1420 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1421 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1422 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1423 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1424 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1425 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1426 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1427 it. Use this page for your epilogue or conclusion if applicable; please use only one of the titles for  
1428 this page. Otherwise, you may delete it. Use this page for your epilogue or conclusion if  
1429 applicable; please use only one of the titles for this page. Otherwise, you may delete it. Use this  
1430 page for your epilogue or conclusion if applicable; please use only one of the titles for this page.  
1431 Otherwise, you may delete it. Use this page for your epilogue or conclusion if applicable; please  
1432 use only one of the titles for this page. Otherwise, you may delete it. Use this page for your  
1433 epilogue or conclusion if applicable; please use only one of the titles for this page. Otherwise,  
1434 you may delete it. Use this page for your epilogue or conclusion if applicable; please use only one  
1435 of the titles for this page. Otherwise, you may delete it. Use this page for your epilogue or  
1436 conclusion if applicable; please use only one of the titles for this page. Otherwise, you may delete  
1437 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  $2\pi$  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.

- 1465 [14] P. Bagnaia et al. “Evidence for  $Z^0 \rightarrow e^+e^-$  at the CERN  $\bar{p}p$  Collider”. In: *Phys. Lett. B* 129  
1466 (1983), pp. 130–140.
- 1467 [15] Serguei Chatrchyan et al. “Observation of a New Boson at a Mass of 125 GeV with the  
1468 CMS Experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 30–61. arXiv: 1207.7235  
1469 [hep-ex].
- 1470 [16] Georges Aad et al. “Observation of a new particle in the search for the Standard Model  
1471 Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29.  
1472 arXiv: 1207.7214 [hep-ex].
- 1473 [17] K. G. Begeman, A. H. Broeils, and R. H. Sanders. “Extended rotation curves of spiral galax-  
1474 ies: Dark haloes and modified dynamics”. In: *Mon. Not. Roy. Astron. Soc.* 249 (1991), p. 523.
- 1475 [18] Y. Ashie et al. “Evidence for an oscillatory signature in atmospheric neutrino oscillation”.  
1476 In: *Phys. Rev. Lett.* 93 (2004), p. 101801. arXiv: hep-ex/0404034.
- 1477 [19] C. Abel et al. “Measurement of the Permanent Electric Dipole Moment of the Neutron”. In:  
1478 *Phys. Rev. Lett.* 124.8 (2020), p. 081803. arXiv: 2001.11966 [hep-ex].
- 1479 [20] Guillaume Albouy et al. “Theory, phenomenology, and experimental avenues for dark show-  
1480 ers: a Snowmass 2021 report”. In: *The European Physical Journal C* 82.12 (Dec. 2022).
- 1481 [21] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08  
1482 (2008), S08001.
- 1483 [22] “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008). Also  
1484 published by CERN Geneva in 2010, S08003.
- 1485 [23] “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008),  
1486 S08004.
- 1487 [24] “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008),  
1488 S08002.
- 1489 [25] “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005.
- 1490 [26] Ana Lopes and Melissa Loyse Perrey. *FAQ-LHC The guide*. 2022.
- 1491 [27] Esma Mobs. “The CERN accelerator complex in 2019. Complexe des accélérateurs du  
1492 CERN en 2019”. In: (2019). General Photo.
- 1493 [28] *Pulling together: Super Conducting electromagnets*. <https://home.cern/science/engineering/pulling-together-superconducting-electromagnets>.  
1494 Accessed: 2024-01-05.

- 1496 [29] *The High-Luminosity LHC*. <https://voisins.web.cern.ch/en/high-luminosity-lhc-hl-lhc>. Accessed: 2024-01-05.
- 1497
- 1498 [30] Aad G., et al. (ATLAS Collaboration and CMS Collaboration). “Combined Measurement of  
1499 the Higgs Boson Mass in pp Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS  
1500 Experiments”. In: *Phys. Rev. Lett.* 114 (19 2015), p. 191803.
- 1501 [31] O. Aberle et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design re-*  
1502 *port*. CERN Yellow Reports: Monographs. Geneva: CERN, 2020.
- 1503 [32] The ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”.  
1504 In: *Journal of Instrumentation* 3.08 (2008), S08003.
- 1505 [33] G Aad, B Abbott, and ATLAS Collaboration. “Performance of the reconstruction of large  
1506 impact parameter tracks in the inner detector of ATLAS”. In: *Eur. Phys. J. C Part. Fields*  
1507 83.11 (Nov. 2023).
- 1508 [34] Joao Pequenao. *Computer Generated image of the ATLAS calorimeter*. 2008.
- 1509 [35] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. AT-  
1510 LAS. Geneva: CERN, 1996.
- 1511 [36] H A Gordon. “Liquid argon calorimetry for the SSC”. In: () .
- 1512 [37] Henric Wilkens and (on behalf of the ATLAS LArg Collaboration). “The ATLAS Liquid  
1513 Argon calorimeter: An overview”. In: *Journal of Physics: Conference Series* 160.1 (2009),  
1514 p. 012043.
- 1515 [38] *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*. Tech.  
1516 rep. Geneva: CERN, 2017.
- 1517 [39] “Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer”. In:  
1518 () .
- 1519 [40] L Pontecorvo. “The ATLAS Muon Spectrometer”. In: (2004). revised version number 1  
1520 submitted on 2003-07-27 16:31:16.
- 1521 [41] *ATLAS magnet system: Technical Design Report, 1*. Technical design report. ATLAS. Geneva:  
1522 CERN, 1997.
- 1523 [42] Joao Pequenao. “Event Cross Section in a computer generated image of the ATLAS detec-  
1524 tor.” 2008.
- 1525 [43] ATLAS Collaboration. “ATLAS Experiment Implements Heterogeneous Particle Recon-  
1526 struction with Intel oneAPI Tools”. General Photo. 2023.

- [44] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.12 (2019), P12006.

[45] Chiara Deponte. “Studies on the properties of non-prompt photons at the ATLAS experiment”. Presented 16 Aug 2022. Technische Universitaet Dortmund (DE), 2022.

[46] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13 \text{ TeV}$ ”. In: *The European Physical Journal C* 76.5 (2016).

[47] Sebastien Rettie. *Muon identification and performance in the ATLAS experiment*. Tech. rep. Geneva: CERN, 2018.

[48] B. R. Webber. *Fragmentation and Hadronization*. 1999. arXiv: hep-ph/9912292 [hep-ph].

[49] Eric M. Metodiev. *The Fractal Lives of Jets | Eric M. Metodiev — ericmetodiev.com*. <https://www.ericmetodiev.com/post/jetformation/>. 2019, note = [Accessed 18-05-2024],

[50] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti-ktjet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), 063–063.

[51] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet user manual: (for version 3.0.2)”. In: *The European Physical Journal C* 72.3 (Mar. 2012).

[52] Steven Schramm. *ATLAS Jet Reconstruction, Calibration, and Tagging of Lorentz-boosted Objects*. Tech. rep. Geneva: CERN, 2017.

[53] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *The European Physical Journal C* 77.7 (July 2017).

[54] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *The European Physical Journal C* 77.7 (July 2017).

[55] Stephen D. Ellis and Davison E. Soper. “Successive combination jet algorithm for hadron collisions”. In: *Physical Review D* 48.7 (Oct. 1993), 3160–3166.

[56] M. Wobisch and T. Wengler. *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering*. 1999. arXiv: hep-ph/9907280 [hep-ph].

[57] Gavin P Salam and Gr  gory Soyez. “A practical seedless infrared-safe cone jet algorithm”. In: *Journal of High Energy Physics* 2007.05 (May 2007), 086–086.

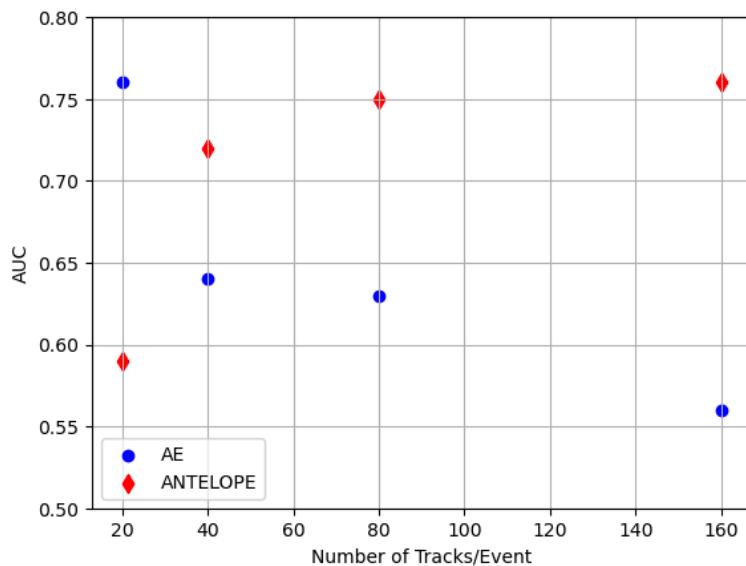
- 1555 [58] Gavin P. Salam. “Towards jetography”. In: *The European Physical Journal C* 67.3–4 (May  
1556 2010), 637–686.
- 1557 [59] *A Monte Carlo study of track association to jets for b-tagging*. Tech. rep. Geneva: CERN,  
1558 2021.
- 1559 [60] *Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector*. Tech. rep.  
1560 All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/P>  
1561 PHYS-PUB-2014-013. Geneva: CERN, 2014.
- 1562 [61] ATLAS Collaboration. “Performance of missing transverse momentum reconstruction with  
1563 the ATLAS detector using proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *Eur. Phys. J. C*  
1564 78.11 (2018), p. 903. arXiv: 1802 . 08168.
- 1565 [62] Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler. “Energy flow networks: deep sets  
1566 for particle jets”. In: *Journal of High Energy Physics* 2019.1 (2019).
- 1567 [63] A. Kahn et al. “Anomalous jet identification via sequence modeling”. In: *Journal of Instru-*  
1568 *mentation* 16.08 (Aug. 2021), P08012.
- 1569 [64] Timothy Cohen et al. “LHC searches for dark sector showers”. In: *Journal of High Energy  
1570 Physics* 2017.11 (Nov. 2017).
- 1571 [65] Georges Aad et al. “Anomaly detection search for new resonances decaying into a Higgs  
1572 boson and a generic new particle  $X$  in hadronic final states using  $\sqrt{s} = 13$  TeV  $pp$  collisions  
1573 with the ATLAS detector”. In: *Phys. Rev. D* 108 (2023), p. 052009. arXiv: 2306 . 03637  
1574 [`hep-ex`].

## Appendix A: Machine Learning Approaches

### 1577 A.1 Unsupervised: AE vs. ANTELOPE

1578 To understand the benefits of the semi-supervised ANTELOPE approach, we study the AN-  
 1579 TELOPE in comparison to a traditional anomaly detection architecture like an autoencoder. The  
 1580 autoencoder cannot accommodate variable length or permutation invariant inputs.

1581 Figure A.1 shows the AUC determined by each of these two tools as a function of number  
 1582 of tracks. The trend is that the AE suffers when more information is given, due to the presence  
 1583 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 A.1

1585 **Signal Contamination**

1586 To understand the effect of signal contamination in training on the ANTELOPE score, we inject  
1587 a percent of signal events into the data used to train the ANTELOPE autoencoder stage and look  
1588 at the AUC on signals. Figure A.2 shows no variation in AUC with 1% contamination in training  
data, but a few % drop going up to 10%.

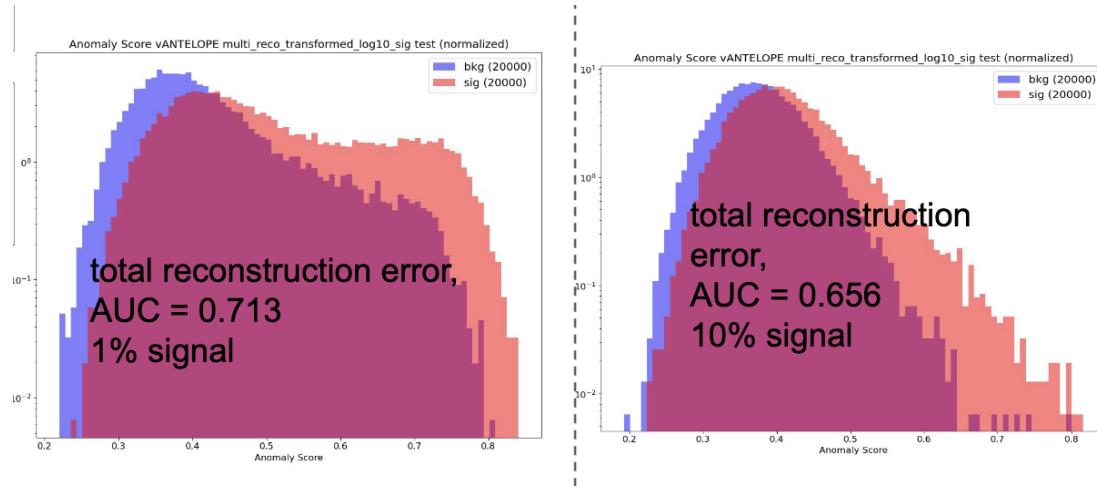


Figure A.2

1589

1590 **A.2 PFN Optimality Checks**

1591 The PFN is trained using QCD as the background. A study was done to compare the perfor-  
1592 mance of the tool in the analysis context if it trains against QCD or a representative MC background  
1593 considering the small fractions of other processes ( $V+jets$ , top) that would contribute at preselec-  
1594 tion. Figure A.3 shows the AUC across the grid for both training approaches, revealing better  
1595 performance if the tool focuses on learning QCD differences.

1596 Further studies were done to ensure the relatively optimality of the single PFN model, trained  
1597 over combined signals, across the grid. As the grid spans signals with a large range of  $E_T^{\text{miss}}$ , their  
1598 varying input features and background composition may be conducive to separate PFN models  
1599 trained on high and low  $E_T^{\text{miss}}$  signal points to better capture the signal-background differences.  
1600 Figure A.4 shows a comparison of the signal-inclusive PFN model performance and the perfor-

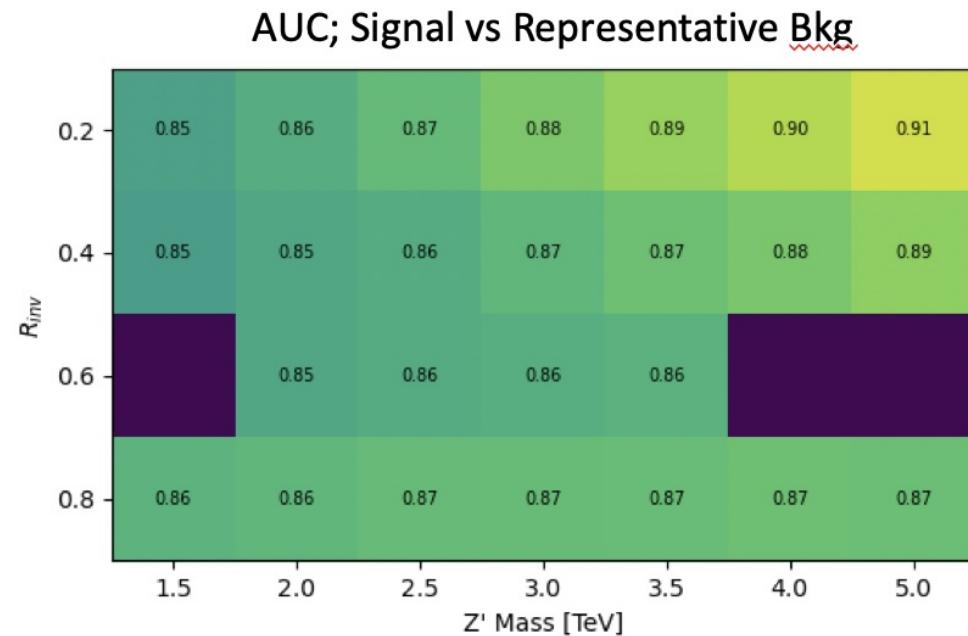
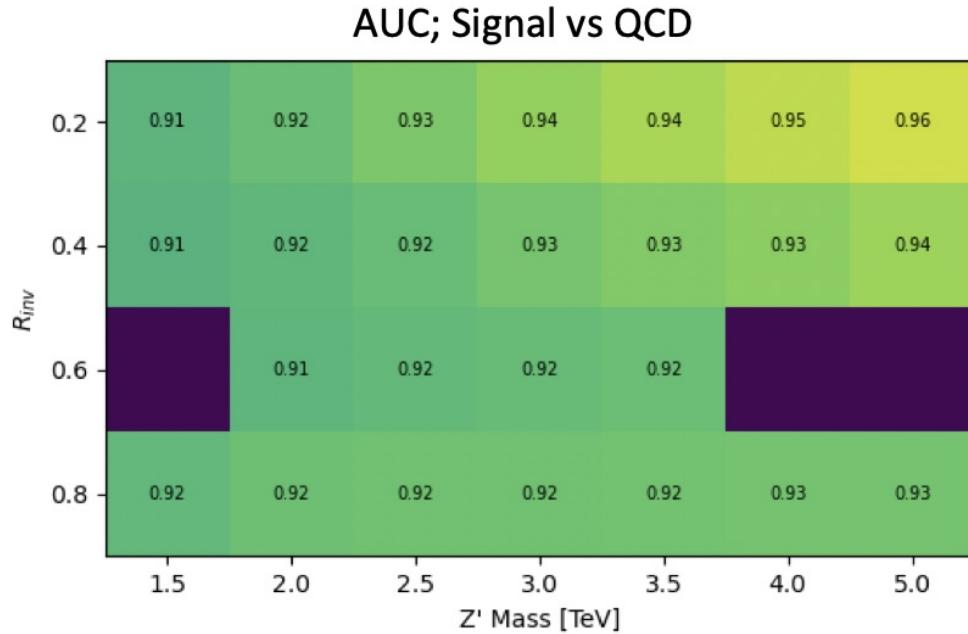


Figure A.3: AUC from the PFN score for each signal in the SVJ grid, shown versus the QCD-only training sample (top) and the total MC background (bottom). Note the three missing points will be added shortly - they were delayed due to a DAOD production mistake.

1601 mance of models separated into high and low  $R_{inv}$  signals in training. The most notable impact is  
 1602 found for the low  $R_{inv}$  and high mass points, indicating that the signal-inclusive PFN is learning  
 1603 morning about the distinction between high- $E_T^{\text{miss}}$  signals and backgrounds. However, these high  
 1604 mass points are also the most challenging to find due to their very wide resonance on top of  $m_T$ ,  
 1605 and in the final projected sensitivity in the  $m_T$  window the differences are  $< 10\%$  across the grid.  
 1606 To maintain a harmonized strategy with the ANTELOPE region we keep the inclusive PFN model  
 1607 as the final version.

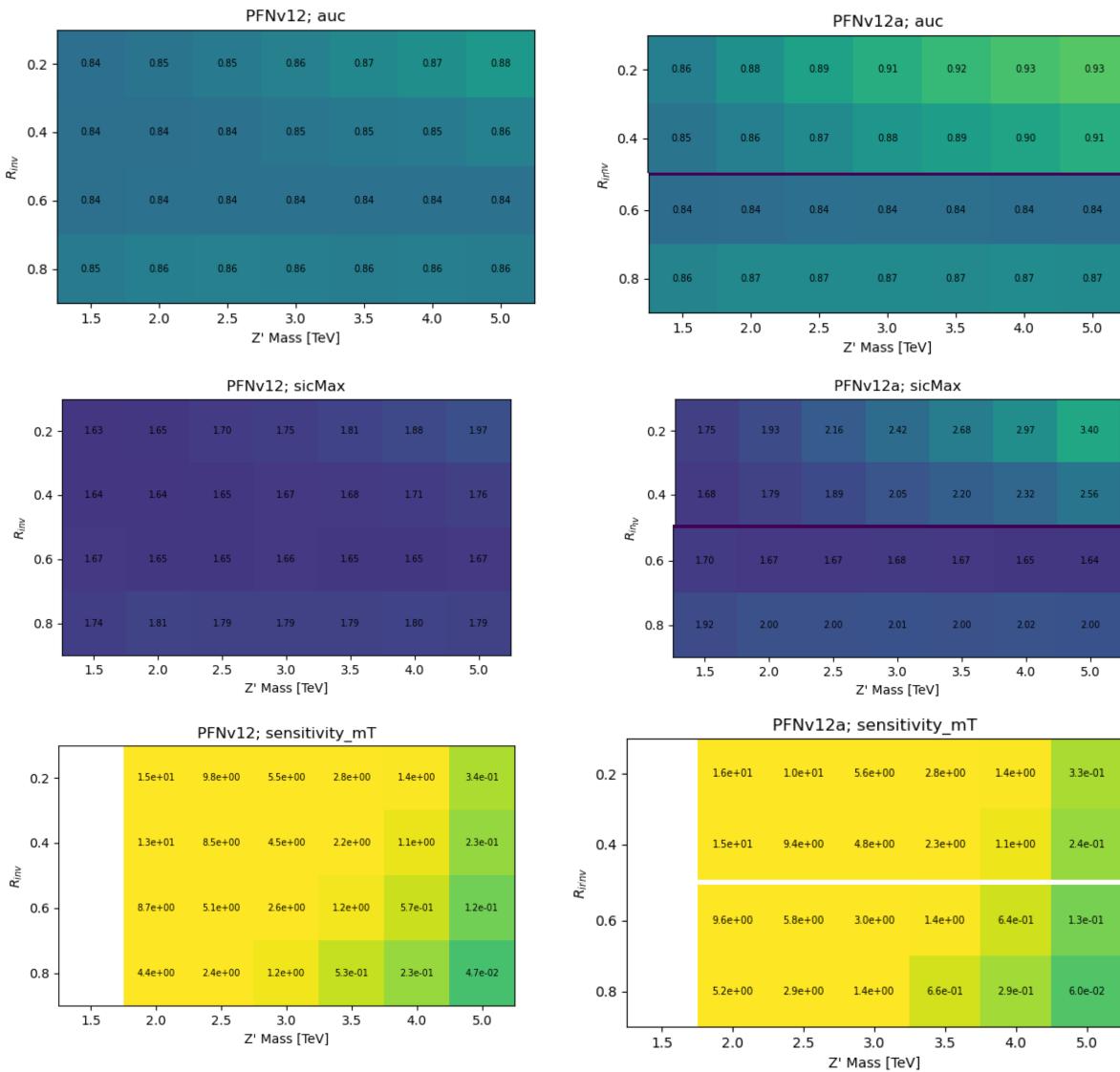


Figure A.4: Comparison of PFN AUC (top), SIC (middle), and sensitivity in the  $m_T$  mass window (bottom) for a single PFN model (left) vs. two models, trained on  $R_{inv} < 0.5$  and  $> 0.5$  separately.

1608      Figure A.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.

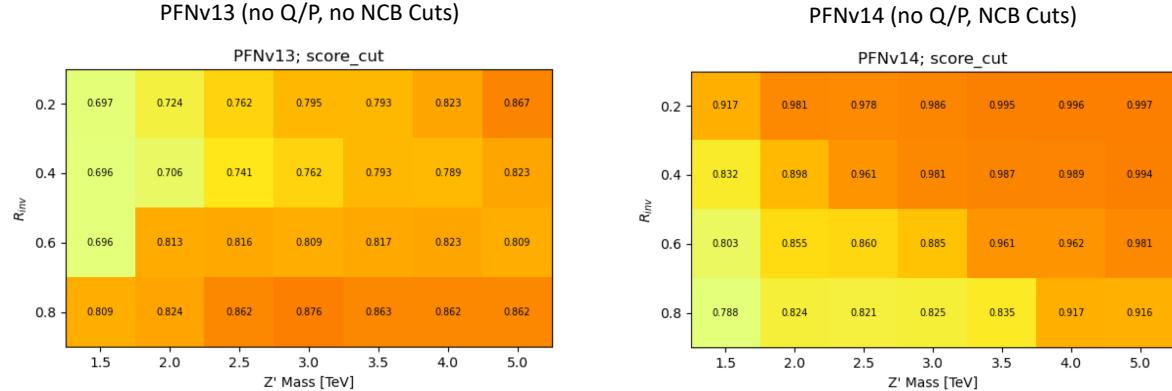


Figure A.5: Preferred cuts on the PFN score for each point in the grid, comparing the effect of adding the NCB preselection.

1609  
1610      Grid cans for optimality were also performed on the number of training epochs, number of  
1611      training events, batch size, learning rate, number of neurons, and dimension of the  $\Phi$  space. The  
1612      results of these scans are summarized in the tables in Figure A.6. The selected or default parameters  
1613      were found to be optimal, or close enough to optimal to justify not increasing the training time or  
1614      complexity of the network for negligible increases in performance.

### 1615      A.3 Supervised: BDT vs. PFN

1616      Studies of the BDT compared to the PFN performance, where training over events modeled  
1617      with jet-related HLVs (high-level track variables, etas, angles, etc.) are compared to events mod-  
1618      eled by the tracks of the 2 leading jets. Figure A.7 shows the performance of the BDT with and  
1619      without explicit use of energy scale variables.

### 1620      A.4 Supervised: Variable Correlations

1621      The correlation between the PFN score and analysis variables known to be unique to semi-  
1622      visible jets are studied, for the sake of better understanding the topology of the tool.  $\Delta\phi(j, E_T^{\text{miss}})$

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

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

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

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

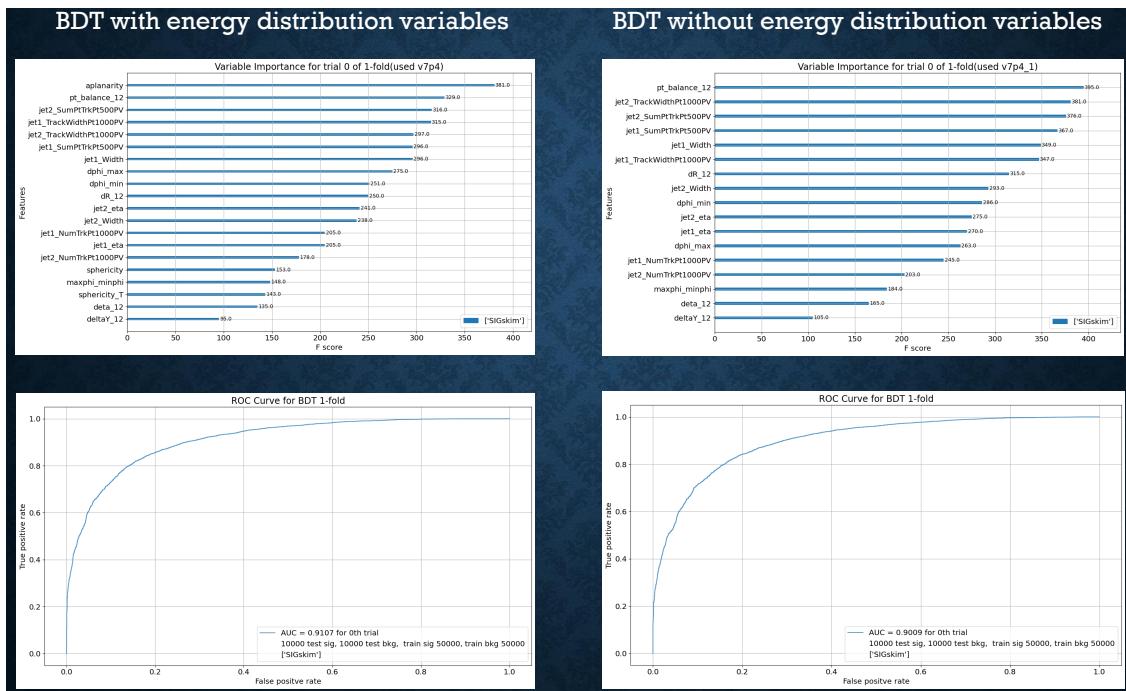


Figure A.7

1623 and  $\Delta\phi(j_1, j_2)$  in the CR and VR are highlighted in Figure A.8 (eg. in different bins of the PFN  
 1624 score), and 2D plots of the score with respect to the orientation variables are in Figure A.9.

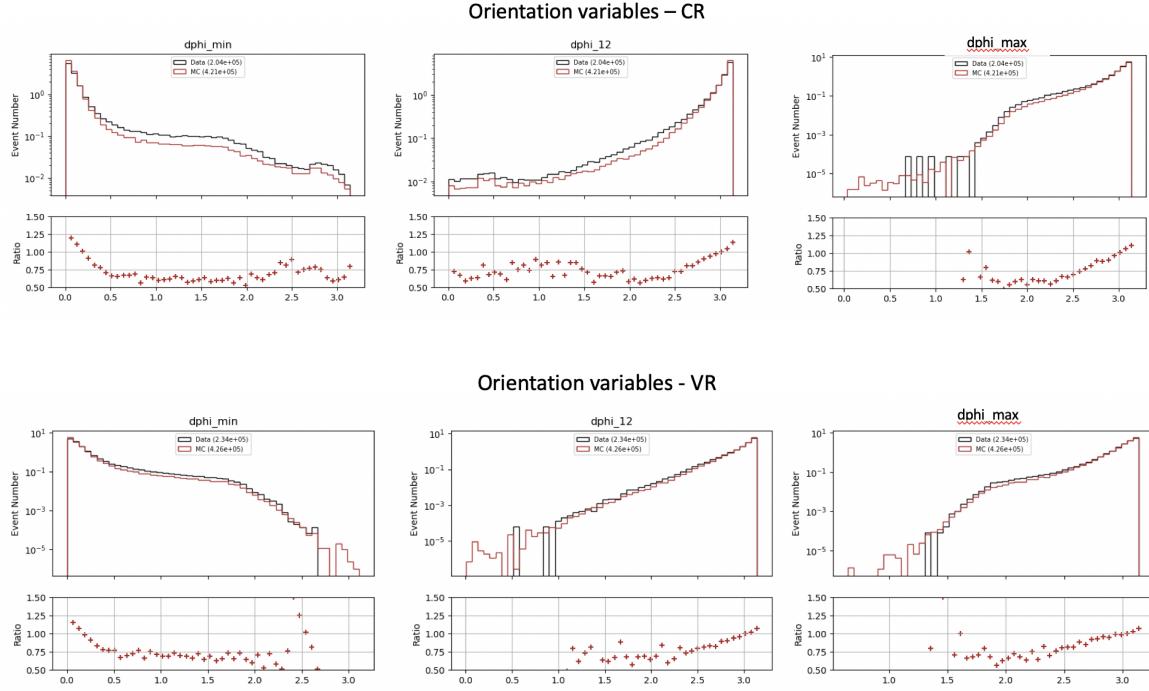


Figure A.8:  $\phi$  orientation variables in the CR and VR

## 1625 A.5 Single Jet vs Jet System ML Approach

1626 The analysis considered both a single jet and jet system ML approach. A jet system approach,  
 1627 where the leading two jets and their orientation with respect to each other was selected for a variety  
 1628 of reasons. The jet system approach captures the MET information which is crucial to identifying  
 1629 SVJs. In the topology where the dark quarks come from a heavy  $Z'$  decay and are back to back,  
 1630 the measurable MET will have to be aligned with one or the other.

1631 Additionally, the performance of both a supervised PFN approach and an unsupervised AE ap-  
 1632 proach was studied in the case of a single jet tagger. While the PFN approach was still performant  
 1633 on a single jet case, the unsupervised approach was significantly improved by using both jets. This  
 1634 is shown in Figure A.10.

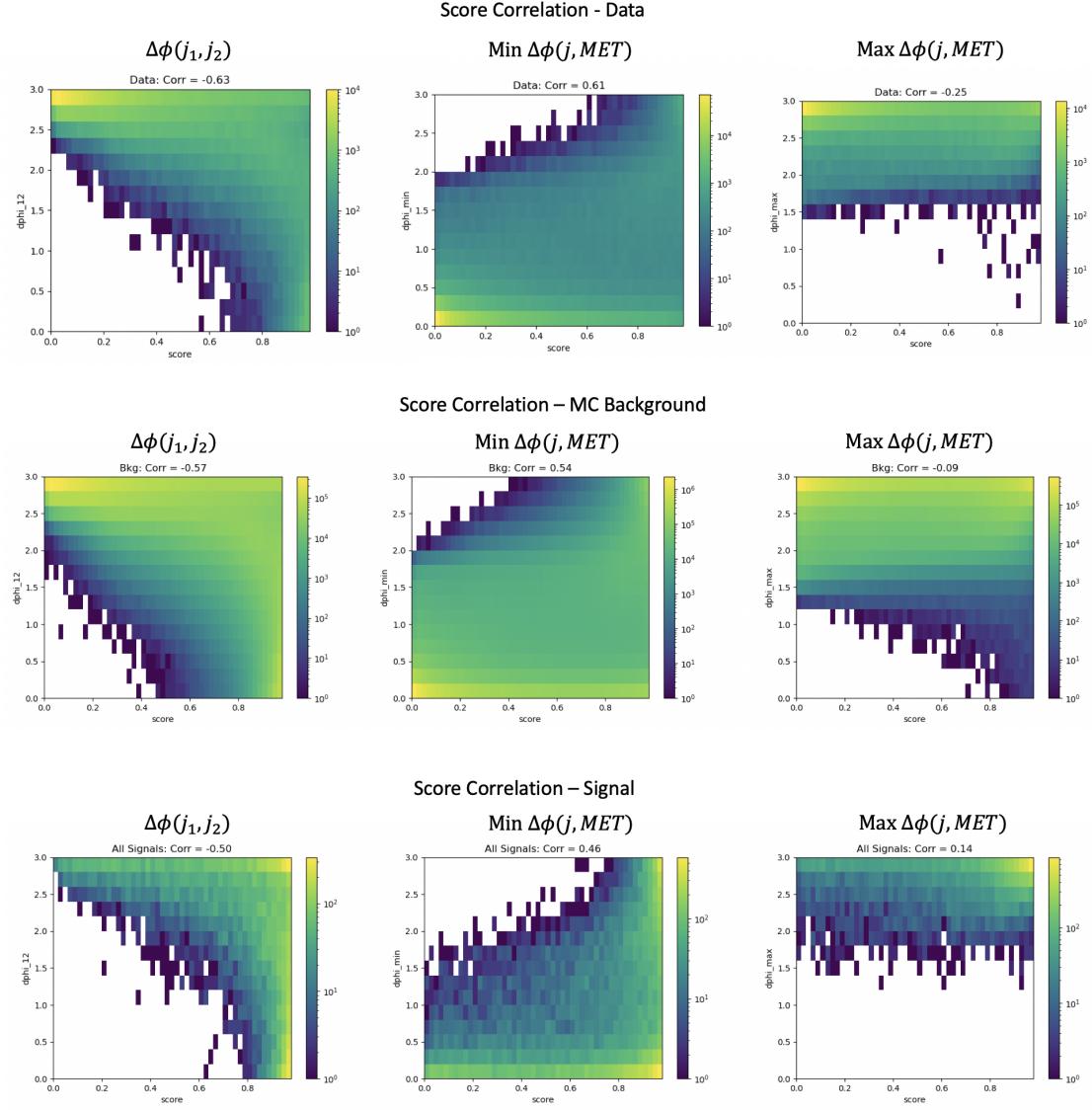


Figure A.9: Correlation between  $\phi$  orientation variables and PFN score

## 1635 A.6 PFN Training Composition

1636 The overall sensitivity and stability across the signal grid is observed to benefit by training  
 1637 the ML tool to reject only the dominant background, QCD. This is evidenced by the PFN re-  
 1638 sponse plots shown in Section 7.1.1 and the following AUC and sensitivity comparison plots in  
 1639 Figure A.11.

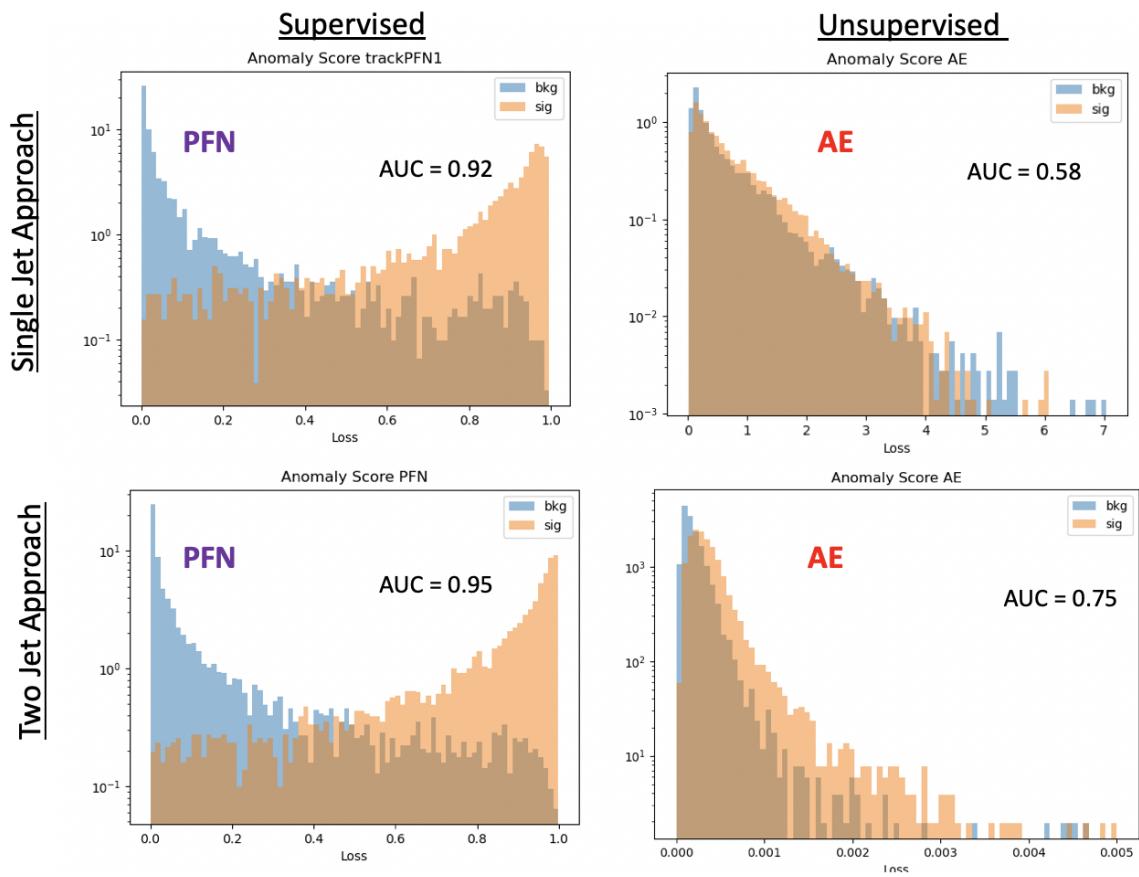


Figure A.10:  $\phi$  Performance comparison between single jet and jet system ML approach

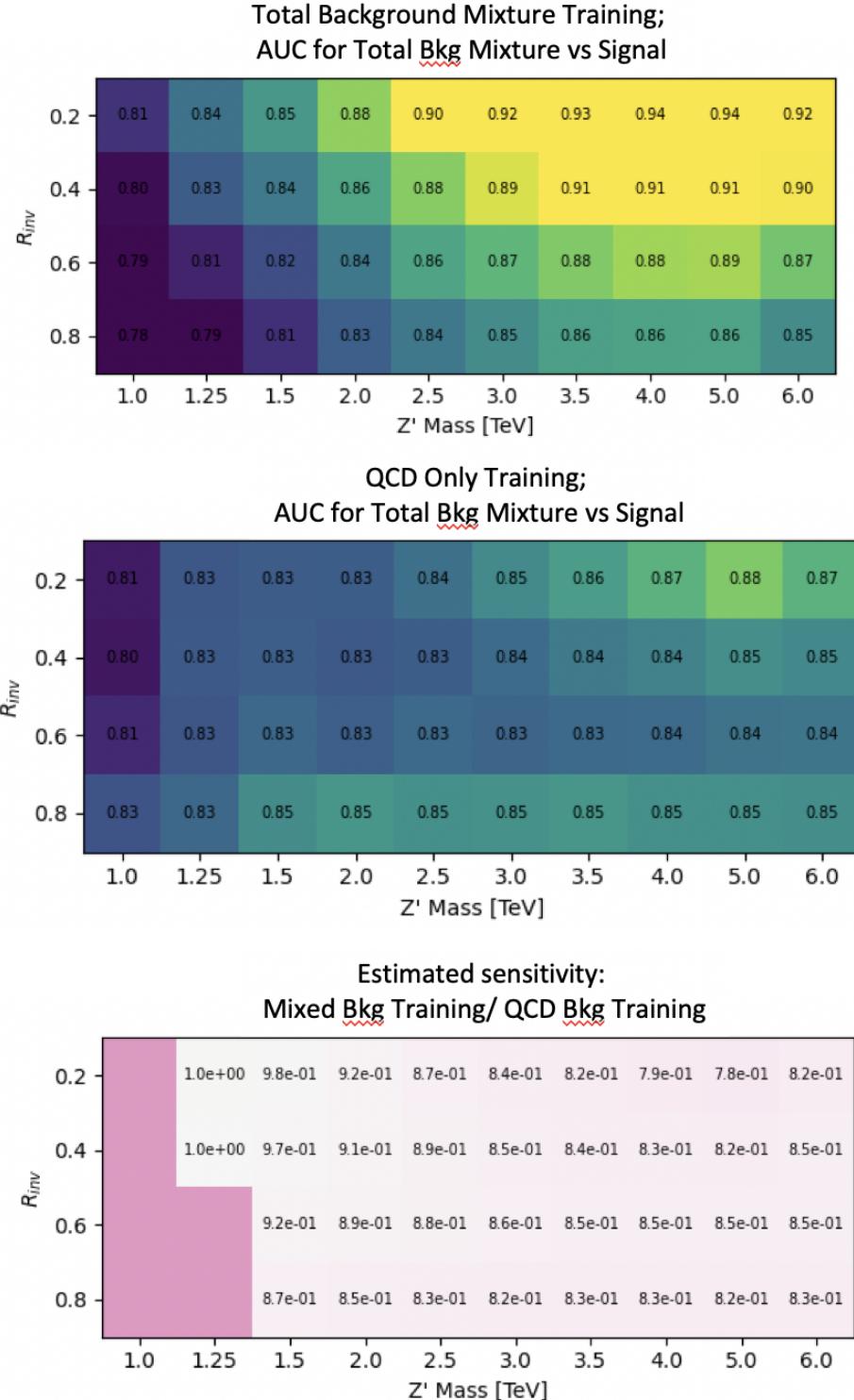


Figure A.11:  $\phi$  Comparison in the AUC score across the grid for the mixed background strategy vs the QCD only strategy. The bottom table highlights that the QCD only strategy gives superior sensitivity across the signal grid.