

¹ Generically Orthogonal Decompositions of
² Collision Events and Measurement
³ Combinations in Standard Model $VH(b\bar{b})$
⁴ Searches with the ATLAS Detector

⁵ A DISSERTATION PRESENTED
⁶ BY
⁷ STEPHEN K. CHAN
⁸ TO
⁹ THE DEPARTMENT OF PHYSICS

¹⁰ IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
¹¹ FOR THE DEGREE OF
¹² DOCTOR OF PHILOSOPHY
¹³ IN THE SUBJECT OF
¹⁴ PHYSICS

¹⁵ HARVARD UNIVERSITY
¹⁶ CAMBRIDGE, MASSACHUSETTS
¹⁷ 2018

¹⁸ ©2018 – STEPHEN K. CHAN

¹⁹ ALL RIGHTS RESERVED.

20 **Generically Orthogonal Decompositions of Collision Events**
21 **and Measurement Combinations in Standard Model $VH(b\bar{b})$**
22 **Searches with the ATLAS Detector**

23 **ABSTRACT**

24 This thesis describes variations on the two lepton channel of the Run-2 search for the SM Higgs
25 boson produced in association with a vector boson using different variable sets for multivariate anal-
26 ysis (MVA) training. The three variable sets in question are the set of variables from the fiducial anal-
27 ysis, a set based on the Lorentz Invariants (LI) concept, and a set based on a combination of masses
28 and decay angles derived using the RestFrames (RF) package. Aside from the variable sets used for
29 MVA training and discriminant distributions, the analysis is otherwise identical to the fiducial an-
30 lysis. Both the LI and RF sets perform competitively on the basis of significances, with the RF set
31 showing a $\sim 3.5\%$ improvement in expected fits to Asimov and data, though neither set boosts ob-
32 served significance. Both sets also reduce the observed error on $\hat{\mu}$, with the LI set reducing the error
33 due to systematics by 7.5% and the RF set doing so by 16%.

34 The issue of combining multiple results from different channels and datasets is also examined
35 through the combination of the fiducial Run 1 and Run 2 ATLAS $VH(b\bar{b})$ results, which results in
36 an observed signal strength of $0.90^{+0.18}_{-0.18}(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$ and an observed (expected) significance
37 of 3.6 (4.0) standard deviations, the first ever evidence of this process.

Contents

38

39	o	INTRODUCTION	I
40	I	THE STANDARD MODEL HIGGS AND COLLIDER EVENT VARIABLES	4
41	1.1	The Standard Model Higgs Boson	5
42	1.2	Higgs Boson Production and Decay at the Large Hadron Collider	8
43	1.3	Collider Events and Event Level Variables	II
44	1.4	Characterization with Event-Level Variables	13
45	1.5	Lorentz Invariants	16
46	1.6	RestFrames Variables	18
47	1.7	Extensions to the ℓ and τ Lepton Channels	20
48	2	THE LARGE HADRON COLLIDER AND THE ATLAS DETECTOR	24
49	2.1	The CERN Accelerator Complex	25
50	2.2	The Large Hadron Collider	25
51	2.3	ATLAS at a Glance	31
52	2.4	The Inner Detector	37
53	2.5	The ATLAS Calorimeters	41
54	2.6	The Muon Spectrometer	49
55	3	DATA AND SIMULATED SAMPLES	55
56	4	SIGNAL AND BACKGROUND MODELING	58
57	4.1	Event Generation In a Nutshell	59
58	4.2	Description of Modeling Uncertainty Categories	61
59	4.3	Process Specific Systematic Summaries	72
60	5	OBJECT AND EVENT RECONSTRUCTION AND SELECTION	82
61	5.1	Triggers	84
62	5.2	Electrons	86
63	5.3	Muons	88
64	5.4	Missing Transverse Energy	90
65	5.5	Jets	91
66	5.6	Flavor Tagging	100
67	5.7	Miscellania and Systematics Summary	III
68	5.8	Event Selection and Analysis Regions	III

69	6 MULTIVARIATE ANALYSIS CONFIGURATION	II4
70	6.1 Training Samples and Variable Selection	II5
71	6.2 MVA Training	126
72	6.3 Statistics Only BDT Performance	133
73	7 STATISTICAL FIT MODEL AND VALIDATION	I41
74	7.1 The Fit Model	I42
75	7.2 Fit Inputs	I45
76	7.3 Systematic Uncertainties Review	I46
77	7.4 The VZ Validation Fit	I48
78	7.5 Nuisance Parameter Pulls	I52
79	7.6 Postfit Distributions	I63
80	7.7 VH Fit Model Validation	I70
81	8 FIT RESULTS	I86
82	9 MEASUREMENT COMBINATIONS	I92
83	9.1 Lepton Channel Combinations	I93
84	9.2 The Combined Fit Model	I97
85	9.3 Combined Fit Results	I212
86	10 CLOSING THOUGHTS	I220
87	APPENDIX A MICROMEGAS TRIGGER PROCESSOR SIMULATION	I223
88	A.1 Algorithm Overview	I224
89	A.2 Monte Carlo Samples	I229
90	A.3 Nominal Performance	I229
91	A.4 Fit Quantities	I230
92	A.5 Efficiencies	I236
93	A.6 Incoherent Background	I239
94	A.7 BCID	I244
95	A.8 Charge Threshold	I244
96	A.9 Misalignments and Corrections	I248
97	A.10 Individual Cases	I253
98	A.11 $ds \neq 0$	I254
99	A.12 $dz \neq 0$	I254
100	A.13 $dt \neq 0$	I254
101	A.14 $\alpha \neq 0$	I255
102	A.15 $\beta \neq 0$	I255
103	A.16 $\gamma \neq 0$	I255
104	A.17 Simulation Correction of the Algorithm Under Nominal Conditions	I258
105	A.18 Translation Misalignments Along the Horizontal Strip Direction (Δs)	I261

106	A.19 Translation Misalignments Orthogonal to the Beamline and Horizontal Strip Direction (Δz)	263
108	A.20 Translation Misalignments Parallel to the Beamline (Δt)	265
109	A.21 Chamber Tilts Towards and Away from the IP (γ_s Rotation)	266
110	A.22 Rotation Misalignments Around the Wedge Vertical Axis (β_z)	268
111	A.23 Rotation Misalignments Around the Axis Parallel to the Beamline (α_t)	269
112	A.24 Conclusion	271
113	APPENDIX B TELESCOPING JETS	273
114	B.1 Monte Carlo Simulation	274
115	B.2 Jet Reconstruction and Calibration	275
116	B.3 Event Reconstruction and Selection	276
117	B.4 Validation of Jet Calibration	278
118	B.5 Truth-Level Analysis	281
119	B.6 Errors on Telescoping Significances	284
120	B.7 Counting	284
121	B.8 Multiple Event Interpretations	285
122	B.9 $t(z) = z$	286
123	B.10 $t(z) = \rho_S(z) / \rho_B(z)$	287
124	B.11 Reconstructed-Level Analysis	292
125	B.12 Conclusions and Prospects	298
126	APPENDIX C HISTORICAL ASIDE	299
127	REFERENCES	313

Acknowledgments

¹³⁰ THIS THESIS WOULD NOT HAVE BEEN POSSIBLE without large amounts of espresso.

吾生也有涯，而知也无涯。以有涯隨无涯，殆已；
已而為知者，殆而已矣。

Zhuangzi, Book 3

0

131

132

Introduction

133 SINCE THE DISCOVERY of a Standard Model (SM) like Higgs boson at the LHC in 2012[1][2], one
134 of the main outstanding physics goals of the LHC has been to observe the primary SM Higgs decay
135 mode, $H \rightarrow b\bar{b}$, with efforts primarily targeted at searching for Higgs bosons produced in associ-
136 ation with a leptonically decaying vector (W or Z , denoted generically as V) boson. This primary

¹³⁷ Higgs decay mode also offers the best opportunity to observe direct Higgs coupling to quarks. As
¹³⁸ the integrated luminosity of data collected at the LHC increases, $H \rightarrow b\bar{b}$ searches will increasingly
¹³⁹ become limited by the ability to constrain systematic uncertainties, with the latest result from AT-
¹⁴⁰ LAS at $\sqrt{s} = 13$ TeV using 36.1 fb^{-1} of pp collision data already approaching this regime, having a
¹⁴¹ $VH(b\bar{b})$ signal strength of $1.20^{+0.24}_{-0.23}(\text{stat.})^{+0.34}_{-0.28}(\text{syst.})$ at $m_H = 125$ GeV [3].

¹⁴² While this effort will likely require a combination of several different methods at various differ-
¹⁴³ ent stages in the analysis chain, one possible avenue forward is to revise the multivariate anlaysis
¹⁴⁴ (MVA) discriminant input variables used, as various schemes offer the promise of reducing system-
¹⁴⁵ atic uncertainties through more efficient use of both actual and simulated collision data. This thesis
¹⁴⁶ discusses two such alternate MVA schemes, the RestFrames (RF) and Lorentz Invarants (LI) vari-
¹⁴⁷ ables, in the context of the 2-lepton channel of the Run 2 analysis in [3] and [4], henceforth referred
¹⁴⁸ to as the “fiducial analysis,” before a brief discussion of combinations across channels and datasets.

¹⁴⁹ Electroweak symmetry breaking, Standard Model Higgs production and decay, and event level
¹⁵⁰ variables are treated in Chapter 1. The Large Hadron Collider and ATLAS detector are the subject
¹⁵¹ of Chapter 2. Data and simulation samples used are described in Chapter 3. Signal and background
¹⁵² modeling with accompanying systematis are defined in Chapter 4. Object and event reconstruction
¹⁵³ definitions and event selection requirements are outlined in Chapter 5. The multivariate analysis, in-
¹⁵⁴ cluding a description of the LI and RF variable sets and a summary of performance in the absence of
¹⁵⁵ systematic uncertainties, is described in Section 6. The statistical fit model and systematic uncertain-
¹⁵⁶ ties are described in Section 7, and the fit results may be found in Chapter 8. Combining channels
¹⁵⁷ and datasets at different \sqrt{s} values is discussed in the context of the Run 1 + Run 2 SM $VH(b\bar{b})$

¹⁵⁸ combination in Chapter 9. Finally, conclusions and closing thoughts are presented in Chapter 10.

¹⁵⁹ Editorial notes:

¹⁶⁰ 1. pdf will be *probability* distribution function

¹⁶¹ 2. PDF will be *parton* distribution function

¹⁶² 3. Unless otherwise stated, ATLAS and LHC/CERN images are from public available material
¹⁶³ from experiment webpages. Copyright terms may be found here <https://atlas.cern/>
¹⁶⁴ **copyright**.

*The relationship between theorists and experimentalists
is like that between a truffle farmer and his pig*

Howard Georgi

1

¹⁶⁵

¹⁶⁶

The Standard Model Higgs and Collider

¹⁶⁷

Event Variables

¹⁶⁸ MUCH HAS BEEN SAID about the so-called Standard Model (SM) of particle physics, so only the

¹⁶⁹ bare essentials of electroweak symmetry breaking and Higgs production relevant to SM $VH(b\bar{b})$ will

¹⁷⁰ be addressed here. This discussion follows [5] Chapter 11 in both content and notation. We then

¹⁷¹ move onto the treatment of kinematic variables in collider events, including the two novel schemes

¹⁷² considered in this thesis, the Lorentz Invariants (LI) and RestFrames (RF) concepts.

¹⁷³ I.I THE STANDARD MODEL HIGGS BOSON

¹⁷⁴ The generic scalar Lagrangian potential (the kinetic term will be addressed later) for a scalar in the

¹⁷⁵ SM is:

$$V(\Phi) = m^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (I.1)$$

¹⁷⁶ where Φ is a complex scalar doublet field under $SU(2)$ from which the physical Higgs emerges after

¹⁷⁷ symmetry breaking. Its four degrees of freedom are typically decomposed as follows:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\phi^+ \\ \phi^0 + i\alpha^0 \end{pmatrix} \quad (I.2)$$

¹⁷⁸ ϕ^+ is the complex charged component of the Higgs doublet, and ϕ^0 and α^0 are the CP-even and

¹⁷⁹ CP-odd neutral components, respectively.

¹⁸⁰ If the sign of $m^2 \Phi^\dagger \Phi$ is negative, Φ acquires a *vacuum expectation value* or VEV:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \sqrt{\frac{2m^2}{\lambda}} \end{pmatrix} \quad (I.3)$$

¹⁸¹ with this value typically denoted $v = \sqrt{2m^2/\lambda} = (\sqrt{2}G_F)^{-1/2} \approx 246$ GeV (with the coupling

¹⁸² of the 4-Fermi effective theory of weak interactions measured through experiments involving muon
¹⁸³ decay), and ϕ^0 is rewritten as $\phi^0 = H + v$.

¹⁸⁴ This non-zero VEV induces spontaneous symmetry breaking in the SM's gauge (local) symme-
¹⁸⁵ try group of $SU(3)_C \times SU(2)_L \times U(1)_Y$ since the VEV does not respect the $SU(2)_L \times U(1)_Y$
¹⁸⁶ symmetry of the Lagrangian (i.e. $\langle \Phi \rangle$ is not invariant under a gauge transformation of this group).

¹⁸⁷ Three of the four generators of this subgroup are spontaneously broken, which implies the existence
¹⁸⁸ of three massless Goldstone bosons, which are in turn “eaten” by linear combinations of the W^a
¹⁸⁹ (with coupling strength g) and B (with coupling strength g') bosons to form the longitudinal polar-
¹⁹⁰ izations of the familiar W^\pm and Z bosons, with the last generator giving rise to the usual, unbroken
¹⁹¹ $U(1)_{EM}$ symmetry and its massless photon, A , as well as the scalar Higgs boson H . To see this, one
¹⁹² starts with the full Higgs SM Lagrangian (kinetic minus potential only)

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi), \quad D_\mu \Phi = (\partial_\mu + ig\sigma^a W_\mu^a + ig' Y B_\mu / 2) \Phi \quad (1.4)$$

¹⁹³ One simply plugs in the reparametrized Φ with $\phi^0 = H + v$, collects the terms involving v together
¹⁹⁴ with the appropriate W and B kinetic terms to extract:

$$M_W^2 = \frac{g^2 v^2}{4}, \quad M_Z^2 = \frac{(g'^2 + g^2) v^2}{4} \quad (1.5)$$

¹⁹⁵ This is left as an exercise for the reader; this exercise also makes manifest that the Higgs couples with
¹⁹⁶ the W^\pm and Z with strength quadratic in the gauge boson masses. Since the Higgs field also respects

¹⁹⁷ the $SU(3)_C$ color symmetry, the eight gluons are also left massless, and the H is left interacting with
¹⁹⁸ photons and gluons primarily through heavy quark loops (i.e. no tree-level interactions).

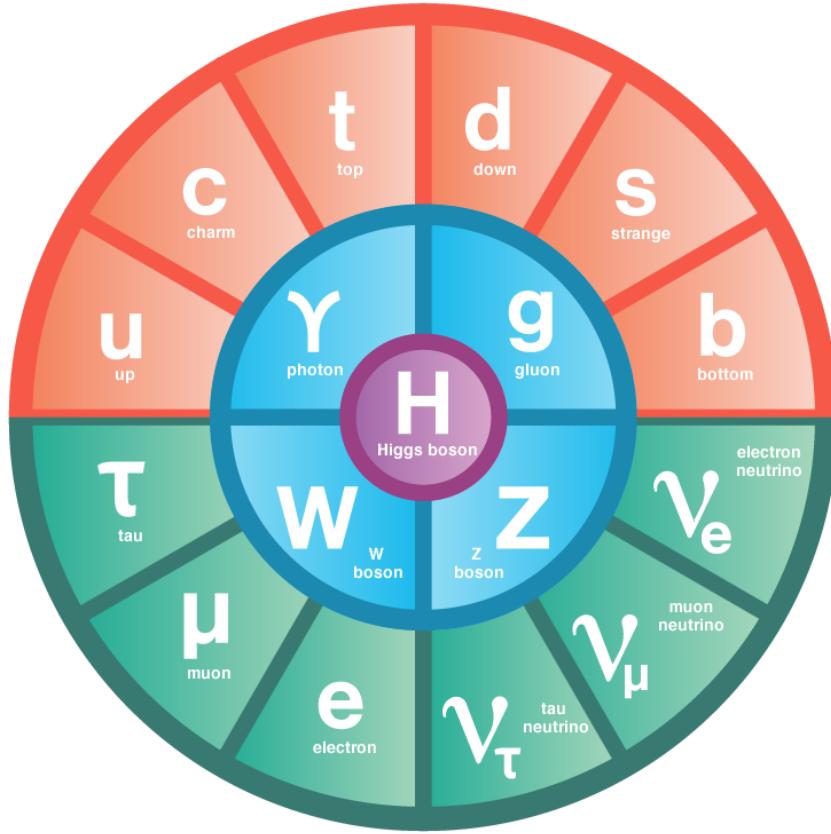


Figure 1.1: The fundamental particles of the Standard Model. IC: [6]

¹⁹⁹ The Higgs is often introduced to the public at large as the mechanism through which fundamen-
²⁰⁰ tal fermions (enumerated in Figure 1.1) acquire mass—this is through the Yukawa interactions of the
²⁰¹ Higgs:

$$\mathcal{L}_{Yukawa} = -\hat{h}_{d_{ij}} \bar{q}_{L_i} \Phi d_{R_j} - \hat{h}_{u_{ij}} \bar{q}_{L_i} \tilde{\Phi} u_{R_j} - \hat{h}_{l_{ij}} \bar{l}_{L_i} \Phi e_{R_j} + h.c. \quad (1.6)$$

²⁰² where $\tilde{\Phi} = i\sigma_2 \Phi^*$, q_L (l_L) and u_R , d_R (e_R) are the quark (lepton) left-handed doublets and right

203 handed singlets of the weak $SU(2)_L$ group, with each term parametrized by a 3×3 matrix in family
204 space (also known as the fermion generations). The neutrinos have been purposely omitted since
205 the mechanism that generates their mass is as of yet unknown, though these Yukawa interactions
206 could have a non-zero contribution to neutrino masses. Once the Higgs VEV value is known and
207 the Yukawa interaction matrices $\hat{b}_{f_i j}$ (with $i, j \in \{1, 2, 3\}$) are diagonalized, the fermion masses
208 can simply be written as $m_{f_i} = b_{f_i} v / \sqrt{2}$. The SM has no motivation for any of these mass values,
209 instead leaving them as empirically determined free parameters.

210 Note that from \mathcal{L}_{Yukawa} , it is easy to see that the Higgs couplings with fermions scale linearly with
211 fermion mass. Higgs self-couplings and beyond the standard model (BSM) Higgs scenarios are be-
212 yond the scope of this thesis.

213 I.2 HIGGS BOSON PRODUCTION AND DECAY AT THE LARGE HADRON COLLIDER

214 The leading order Feynman diagrams for the four dominant modes of Higgs production at the LHC
215 are shown in Figure 1.2, each described briefly in turn. The dominant process, accounting for some
216 87% of Higgs production at the nominal LHC center of mass energy of 14 TeV, is gluon-gluon fu-
217 sion (ggF), shown at top left in Figure 1.2. At high center of mass energies, most of a proton's mo-
218 mentum is predominantly carried by gluons (as opposed to the constituent valence quarks associ-
219 ated with the hadron's identity). This, along with the difficulties associated with high luminosity
220 antiproton beam production, is why the LHC was designed as a proton-proton collider instead of
221 a proton-antiproton collider (like the Tevatron or SppS). As mentioned above, the Higgs does not
222 couple directly to gluons but must instead be produced through the fermion loop shown in the fig-

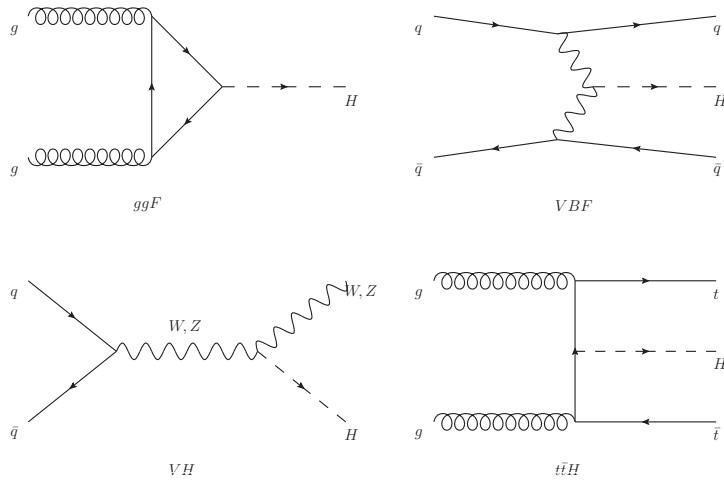


Figure 1.2: Dominant Higgs production modes.

ure. The heaviest fundamental fermion by far is the top quark, with $m_t = 173$ GeV, so top loops dominate this process. While not particularly relevant for this thesis, about 14% of events in the lepton channel of the $H \rightarrow b\bar{b}$ analysis are ggF initiated.

The next most prevalent process is vector boson fusion (VBF), where vector bosons (W or Z , denoted generically as V) from quarks in the colliding protons “fuse” to form a Higgs. These quarks typically form jets in the forward region, which provide a unique signature for this process. This process is not relevant for this thesis.

The third leading process is “Higgsstrahlung” or Higgs production in association with a vector boson, often simply VH production. In this process, a quark-antiquark pair in the colliding protons forms an energetic vector boson, which then radiates a Higgs (this is similar to photon emission of accelerating electrons, called “Bremsstrahlung,” hence the name). Some fraction of the time (about 21% of the time for WH and 6.7% of the time for ZH), the energetic V will decay leptonically (i.e.

235 into a decay involving an electron or a muon), which provides a unique and triggerable signature
 236 for this process. Another 20% of the time for ZH production, the Z will decay to neutrinos, which
 237 are not absorbed by detectors and show up as missing transverse energy (\vec{E}_T^{miss}), another triggerable
 238 signature. This ability to trigger on leptons and \vec{E}_T^{miss} and the requirement that this leptonic signa-
 239 ture be consistent with a V allow one to significantly reduce the impact of multijet background (a
 240 very common generic processes at the LHC) on analysis. Hence, this is the process of primary impor-
 241 tance to this thesis.

242 The final important Higgs production process is $t\bar{t}H$ production, the box diagram in the lower
 243 right of Figure 1.2. Again, the top pair provides a useful signature for analysis. This, like VBF, is also
 244 not considered in this thesis.

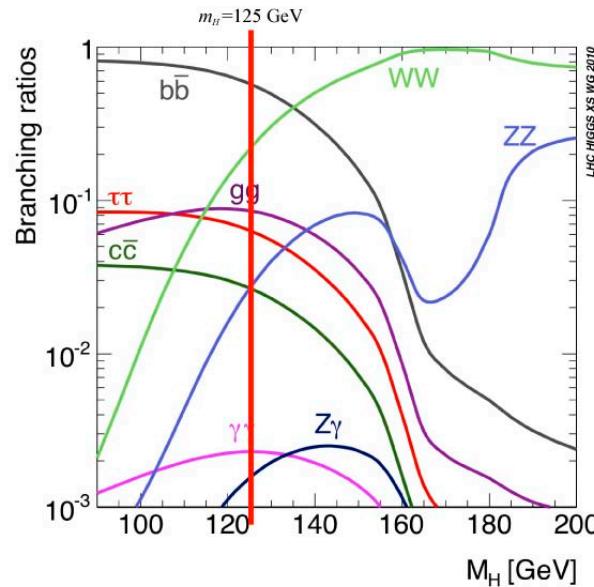


Figure 1.3: Higgs decay mode branching fractions as a function of its mass; a line has been drawn at the observed Higgs mass of 125 GeV.

Once the Higgs has been produced, it can decay in a number of ways, as shown in Figure 1.3. By far the most dominant decay mode of the Higgs is to $b\bar{b}$ with a branching fraction of 58%. This b -quark pair then hadronizes into two b -jets (for a more thorough discussion of jets and b -jets in particular, see Section 5.5). However, many processes at the LHC create pairs of b -jets with invariant masses consistent with the Higgs and have much higher production rates ($t\bar{t}$ production at the LHC is in the neighborhood of hundreds of pb, compared to Higgs cross sections of a few pb), so a clear process signature is necessary to study $H \rightarrow b\bar{b}$ production at the LHC. This is why the bulk of search efforts have focused on VH production. A summary of Higgs production cross sections and simple extrapolations to raw numbers of Higgs bosons produced for VH for leptonically decaying V is shown in Table 1.1

\sqrt{s} (TeV)	ZH	WH	ggF	total σ	$N_{V \rightarrow \ell^+ \nu} H$
7	$0.34^{+4\%}_{-4\%}$	$0.58^{+3\%}_{-3\%}$	$15.3^{+10\%}_{-10\%}$	17.5	$4.7 \text{ fb}^{-1} \rightarrow 589$
8	$0.42^{+5\%}_{-5\%}$	$0.70^{+3\%}_{-3\%}$	$19.5^{+10\%}_{-11\%}$	22.3	$20.3 \text{ fb}^{-1} \rightarrow 3100$
13	$0.88^{+5\%}_{-5\%}$	$1.37^{+2\%}_{-2\%}$	$44.1^{+11\%}_{-11\%}$	50.6	$36.1 \text{ fb}^{-1} \rightarrow 11100$
14	$0.99^{+5\%}_{-5\%}$	$1.51^{+2\%}_{-2\%}$	$49.7^{+11\%}_{-11\%}$	57.1	$1000 \text{ fb}^{-1} \rightarrow 343000$

Table 1.1: Cross sections (in pb) for processes important to the SM VH ($b\bar{b}$) analysis and the total Higgs cross section as a function of center of mass energy. Also given are the total number of Higgs bosons produced for given luminosities through both WH and ZH processes. Uncertainties are theoretical.

1.3 COLLIDER EVENTS AND EVENT LEVEL VARIABLES

Collision data in experiments like ATLAS is structured using what is known as the *event data model*. In this model, one collision corresponds to one event. Since each bunch crossing contains more than one proton, there can be more than one collision per event and more than one hard scatter per col-

lision. For each collision, tracks in an experiment's inner detector are used to identify the most energetic collision, which is taken to be the event. The raw data, the various tracks, energy deposits, and hits in the detector, undergo reconstruction (described at length in Chapter 5) both through automated, experiment-wide, standardized production and through analysis-specific level selections, corrections, and calibrations. The result of this considerable effort is a collection of labeled 4-vectors, representing the final state objects. This is shown in Figure 1.4.

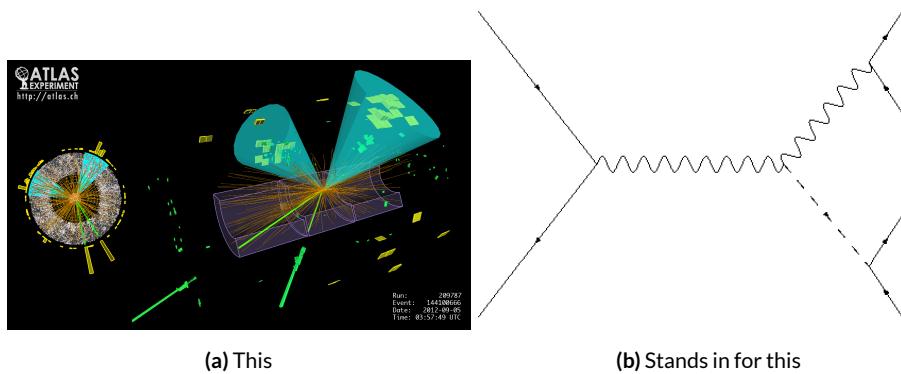


Figure 1.4: Reconstruction in a nutshell

In the process that is the focus of this thesis, every event ultimately is condensed into a lepton pair (two electrons or two muons), two or three jets*, all 4-vectors, and a \vec{E}_T^{miss} vector in the transverse plane. Further selection then takes place to winnow down events into interesting regions of phase space hopefully more rich in signal-like events. Once events are selected in a search like the one in this thesis, one then analyzes the data to test its consistency with some background only hypothesis to produce the usual statistical results. This can be done in various ways, with principal approaches

*Sometimes more, though this is a small fraction of events, and the wisdom of this choice may be questioned

271 being: a simple counting experiment (often referred to as the “cut and count” approach), a func-
272 tional fit for excesses over a falling background spectrum (the so-called “bump hunt” used in anal-
273 yses like the $H \rightarrow \gamma\gamma$ discovery channel), or the use of discriminant distributions as PDF’s in a
274 likelihood fit (the approach of this analysis). These distributions can be simple counts (i.e. single bin
275 distributions) in analysis regions, quantities of interest (the distribution of the invariant mass of the
276 two b -jets in selected events with the greatest transverse momenta, m_{bb} , is used as a validation), or a
277 multivariate analysis (MVA) discriminant.

278 **I.4 CHARACTERIZATION WITH EVENT-LEVEL VARIABLES**

279 Traditionally, particle physicists have favored the approach of using distributions of physical vari-
280 ables since it is easier to develop physical intuition for what these distributions should look like
281 during validation, so it is no surprise that as many LHC analyses have transitioned to using MVA
282 techniques that these variables form the basis of many very robust physics results. These variables
283 do quite well summarize many of the main physics features of an event for the signal topology, cer-
284 tainly much better than feeding all 18–22 4-vector components directly into a machine learning algo-
285 rithm. In $ZH \rightarrow \ell\ell b\bar{b}$ events, for example, one wishes to characterize the ZH system by using the
286 lepton pair as a stand-in for the Z and the b -jet pair as a stand-in for the H , and composite variables
287 like m_{bb} and $m_{\ell\ell}$ can be used to check whether events are consistent with these objects. There are
288 also variables like \vec{p}_T^V that characterize the momentum scale of the event, angles like $\Delta R(b_1, b_2)$ and
289 $\Delta\phi(V, H)$ that can be further used to characterize the overall “shape” of these events, and variables
290 like \vec{E}_T^{miss} that can discriminate against backgrounds like $t\bar{t}$ that do not have a closed topology.

291 Nevertheless, the intuition based approach, with incremental addition of variables as they prove
292 useful in the lifetime of an analysis's iterations, does beg the question of whether there is a more sys-
293 tematic way to treat this information. There are clearly patterns to which variables are useful: these
294 correspond to important information about the hypothesized physics objects and their relation-
295 ships, and there have been many attempts to systematize the way these variables are found. Such
296 systematic, top-down approaches often promise to increase performance in two ways. The first is by
297 having higher descriptive power, often through some sophisticated treatment of the missing trans-
298 verse energy in an event, \vec{E}_T^{miss} . \vec{E}_T^{miss} is just a single quantity, and if there is just one invisible object
299 in a desired event topology, using \vec{E}_T^{miss} on its own often provides sufficient sensitivity. In more com-
300 plicated topologies with multiple invisible particles in the final state, for example in many supersym-
301 metry searches, a more careful treatment of the missing energy is often necessary.

302 The second means of improvement is through using a more orthogonal basis of description,
303 which allows one to more efficiently use data and simulation samples. A more orthogonal basis im-
304 plies that variables contain less overlapping information with each other and so allow for a more
305 efficient exploration of parameter space. This means one can gain higher sensitivity from equivalent
306 datasets using a more orthogonal basis. To see why this might be the case, take an MVA discrimi-
307 nant for $ZH \rightarrow \ell\ell b\bar{b}$ formed using only the classic variables $\Delta R(b_1, b_2)$ and p_T^V . In the $ZH \rightarrow$
308 $\ell\ell b\bar{b}$ topology, the transverse mass of the Z and H (and hence the lepton pair and jet pair) are equiv-
309 alent. This means that at higher p_T^V the p_T of b -jets will also be higher, which in turn implies that
310 they will have a smaller angle of separation and hence a smaller $\Delta R(b_1, b_2)$. This correlation is not
311 unity—each variable still does have information the other does not—but is still very high. Hence,

312 when training an MVA, which in principle knows nothing about these variables other than some
313 set limits, an undue number of training events will be wasted converging upon relations that could
314 be known *a priori*, and while this might be easy to hard code in for a two variable toy example, the
315 dimensionality of any real discriminant makes this prohibitive. An MVA that uses data (both ac-
316 tual and simulated) more efficiently will also tend to be have lower variance, offering a potential av-
317 enue for reduction in the error on quantities of interest due to systematic uncertainties. Details of
318 how this plays out in a likelihood fit will be deferred to the discussion of the fit model used in the
319 $VH(b\bar{b})$ search in Chapter 7.

320 Many of these novel schemes are designed to explicitly address the first issue of invisibles in the
321 final state in channels where it is of paramount importance while having the second issue as some-
322 thing of a fringe benefit. However, as the amount of data taken at the LHC grows, analyses will in-
323 creasingly become systematics limited, so an exploration to the veracity of the second claim has great
324 potential for the high luminosity era of the LHC. The $ZH \rightarrow \ell\ell b\bar{b}$ process offers a great setting for
325 investigating this issue on its own since its closed topology largely mitigates any improvement from
326 more sophisticated treatments of \vec{E}_T^{miss} . We introduce two of these more top-down approaches to
327 event-level variables below: the “Lorentz Invariant” (LI) [7] and “RestFrames inspired” (RF) [8]
328 variable schemes. A broad overview of the concepts behind these schemes will be given here, with a
329 more in-depth discussion of their implementation deferred until Chapter 6.

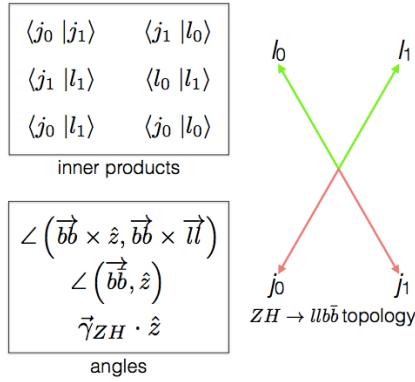


Figure 1.5: Summary of LI variables in the $ZH \rightarrow \ell\ell b\bar{b}$ topology.

330 **I.5 LORENTZ INVARIANTS**

331 The LI variables, first put forth by S. Hagebeck and others [7], are based upon the concept that
 332 once the 4-vectors of an event are determined, all of the information in an event are encoded into
 333 their inner products (Lorentz invariant quantities, hence the name) and the angles between them.
 334 This makes for 16 quantities in all: the ten inner products of the 4-vectors, the three Euler angles,
 335 and the three parameters specifying the boost of the ZH system. The masses of the four final state
 336 objects are not considered very useful and so can be removed to leave six meaningful inner products
 337 (the ${}_4C_2$ combinations[†] between distinct final state 4-vectors). Since these inner products can have
 338 an ill-defined physical interpretation and in order to help MVA training, each inner product is scaled

[†] ${}_nC_r = \frac{n!}{r!(n-r)!}$, read “ n choose r ” and known as the binomial coefficient, is the number of unique possible ways to choose combinations of r objects from a total set of n without regard to ordering within combinations.

³³⁹ by:

$$x \rightarrow \frac{x}{x + c} \quad (1.7)$$

³⁴⁰ where c is the mean of the distribution in the signal MC distribution. These inner products are de-
³⁴¹ noted $x_i y_j$, where x and y are either j (for jet) or l (for lepton) and the indices are either o (i) for
³⁴² the leading (subleading) object by p_T in the event.

³⁴³ The number of useful angles can be reduced by recognizing some symmetries inherent in the fi-
³⁴⁴ nal state. The symmetry around the beam axis eliminates one angle. Furthermore, the boost of the
³⁴⁵ VH system is primarily in the beam direction (z) direction, marginalizing the utility of the trans-
³⁴⁶ verse boost angles. This leaves the boost in the z direction, denoted `gamma_ZHz`, and two angles
³⁴⁷ chosen to be the angle between the $b\bar{b}$ system and the beam (`angle_bb_z`) and the angle between
³⁴⁸ $(\vec{b}_1 + \vec{b}_2) \times \hat{z}$ and $(\vec{b}_1 + \vec{b}_2) \times (\vec{l}_1 + \vec{l}_2)$ (`angle_bbz_bbll`).

³⁴⁹ These variables do contain a lot of information similar to the usual set: there are mass equivalents
³⁵⁰ ($j_0 \leftrightarrow m_{bb}$, and $l_0 \leftrightarrow m_{\ell\ell}$) and angles. Instead of individual final state object scales, there
³⁵¹ are the four jet/lepton inner products, though this correspondence (and indeed any physical inter-
³⁵² pretation) is far from clear. An important advantage of the LI variable set is that all of the variables
³⁵³ are in it are orthogonal in the signal case by construction. A drawback of this framework in a com-
³⁵⁴ pletely closed final state is that there is no way to treat E_T^{miss} in a Lorentz invariant way.

³⁵⁵ There is also no prescription for any additional jets in the event beyond the two b -tagged jets.
³⁵⁶ They are simply ignored in these variable calculations since the fiducial analysis requirement of ex-

357 actly two b -tagged jets eliminates any combinatoric ambiguity, and additional, untagged jets are as-
 358 sumed (not entirely rigorously) to be unrelated to the signal-like hard scatter.

359 **1.6 RESTFRAMES VARIABLES**

360 The RestFrames variables [8], calculated using the software package of the same name, is based upon
 361 the idea that the most natural frame in which to analyze objects of the signal decay tree is in their in-
 dividual production (rest) frames. The signal decay tree for $ZH \rightarrow \ell\ell b\bar{b}$ is show in Figure 1.6. Gen-

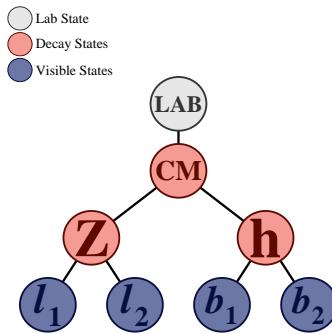


Figure 1.6: The $ZH \rightarrow \ell\ell b\bar{b}$ decay tree.

362
 363 erally, one does not typically have enough information to determine exactly each of the intermediate
 364 rest frames or the boosts between the frames, but in a completely closed final state like $ZH \rightarrow \ell\ell b\bar{b}$,
 365 this can be done in the usual way by adding the 4-vectors of the final state objects and solving the
 366 usual equations from special relativity (RestFrames does this automatically for each event).

367 Each frame has associated with it the boost from its immediate parent and a mass scale; that mass
 368 (in this case the correspondence between RF mass variables and standard mass variables is exact) and
 369 the angles between the Euclidean three vector associated with boost and the axis of the decay prod-
 370 ucts provide useful variables. In general, the polar angle (typically given as a cosine) is considered

371 more useful than the azimuthal angle (typically just a $\Delta\phi$), though this is dependent on the candi-
 372 date decay tree. The Z frame, for example, has M_Z , which is just the usual $m_{\ell\ell}$, $\cos Z$, the cosine of
 373 the polar angle between the lepton momentum axis in their production frame and the boost from
 374 the ZH center of mass (CM) frame, and the angle dphiCMZ .

375 In addition to the masses and angles attached to individual object rest frames, energy scales associ-
 376 ated with the CM frame can be used to contextualize other event level quantities. In particular, one
 377 can use the mass of the CM frame as a natural scale to evaluate the momentum of the CM frame,
 378 and the p_T of the CM frame as a natural scale for the event's E_T^{miss} , yielding the variables:

$$R_{p_T} = \frac{p_{T,\text{CM}}}{p_{T,\text{CM}} + M_{\text{CM}}}, \quad R_{p_z} = \frac{p_{z,\text{CM}}}{p_{z,\text{CM}} + M_{\text{CM}}}, \quad R_{\text{met}} = \frac{E_T^{\text{miss}}}{E_T^{\text{miss}} + p_{T,\text{CM}}} \quad (\text{I.8})$$

379 denoted R_{p_T} , R_{p_z} , and R_{met} . These can be thought of as behaving like significance based variables
 380 in particle physics, like METHT or impact parameter significances, or event level defined versions
 381 of the scalings applied to the LI inner products. These are used instead of the final state object scales
 382 and standard E_T^{miss} of the standard variable set.

383 Unlike the LI variables, the physical interpretation of RF variables is very clear. Everything has
 384 physical units, and these are variables one might have introduced in the usual process of develop-
 385 ing an MVA with the traditional mindset. The solution to the issue of additional jets in an event is
 386 not immediately clear. In order to keep the two non-standard MVA's on as equal footing as possi-
 387 ble, the approach of simply ignoring additional jets is taken in this thesis. Nevertheless, it would be
 388 easy enough to redefine the H intermediate frame to have, for example, the two b -tagged jets and the

389 highest p_T untagged jet for any subset of events. This flexibility is not a feature of the Lorentz Invari-
 390 ants framework. Of course, `RestFrames` cannot tell you what approach to take, but it is capable of
 391 handling more flexible topologies once optimization studies have been completed.

392 1.7 EXTENSIONS TO THE I AND O LEPTON CHANNELS

393 Both the LI and RF variable concepts are readily extendable to the i-lepton channel. In this topol-
 394 ogy, one of the leptons in the $ZH \rightarrow \ell\ell b\bar{b}$ diagram is replaced by a neutrino, the lone invisible
 395 particle in this final state. We can assume that the neutrino has zero mass and transverse momentum
 396 equal to the \vec{E}_T^{miss} in the event, leaving one undetermined degree of freedom, the longitudinal mo-
 397 mentum of the neutrino, p_z^ν .

398 The LI concept was in fact initially formulated to improve sensitivity in the i-lepton channel,
 399 with the same orthogonality of variables described in the 2-lepton case being the main draw. The
 400 LI approach to estimating the neutrino longitudinal momentum is outlined in [7], which we repro-
 401 duce here. We first guess the neutrino energy in its rest frame and then boost to the lab frame:

$$\langle E_\nu \rangle = \frac{1}{4} m_{WH} \implies \langle p_z^\nu \rangle = \beta \gamma \langle E_\nu \rangle = \frac{p_z^{WH}}{m_{WH}} \langle E_\nu \rangle = \frac{1}{4} p_z^{WH} \quad (1.9)$$

402 Finally, assuming energy and momentum in aggregate are equally shared among final state con-
 403 stituents, we arrive at

$$\langle p_z^\nu \rangle = \frac{1}{4} \times \frac{4}{3} (p_z^l + p_z^{j0} + p_z^{j1}) \quad (1.10)$$

404 The RF approach for the i-lepton case amounts to replacing the $Z \rightarrow \ell\ell$ in 1.6 with $W \rightarrow \ell\nu$.

405 As alluded to in the 2-lepton discussion, when there is missing information in the final state from
 406 invisible particles and/or combinatoric ambiguities, recursive jigsaw reconstruction (RJR) offers a
 407 standard toolkit for deriving estimated boosts between rest frames by analytically minimizing on
 408 unknown quantities. While in more exotic final states with multiple invisible particles and com-
 409 binatoric ambiguities the choice of jigsaw rule can be subjective, the case of W is well-studied and
 410 outlined in detail in Section V.A. of [8]. It reproduces the usual transverse mass of the W in place of
 411 MZ in the 2-lepton case. Not surprisingly, the underlying calculation is also much the same as the LI
 412 case (where rest frames and boost were explicitly invoked); again, information is the same, only its
 413 decomposition is different.

414 The 0-lepton channel would appear to present some difficulty as two neutrinos in the final state
 415 introduce extra degrees of freedom, but both concepts may be extended by treating the invisibly
 416 decaying Z as a single invisible particle and requiring the Z to be on-shell, as shown schematically in
 417 Figure 1.7. Both of these requirements may be folded into the 1-lepton framework to produce similar
 sets of variables.

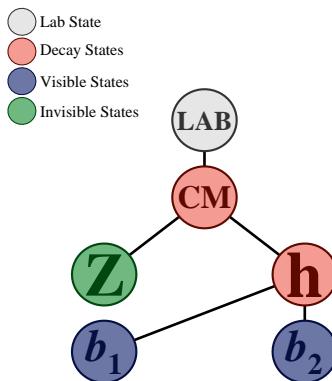


Figure 1.7: The $ZH \rightarrow \nu\nu b\bar{b}$ decay tree.

418

Variable	Name	0-lepton	1-lepton	2-lepton
\vec{p}_T^V	pTV		✓	✓
$\vec{E}_{\text{T}}^{\text{miss}}$	MET	✓	✓	✓
$\vec{p}_T^{\text{jet}1}$	pTB1	✓	✓	✓
$\vec{p}_T^{\text{jet}2}$	pTB2	✓	✓	✓
$\text{MV}_{2\text{C10}}(\text{jet}_1)^*$	$\text{MV}_{2\text{C10B1}}$	✓	✓	✓
$\text{MV}_{2\text{C10}}(\text{jet}_2)^*$	$\text{MV}_{2\text{C10B2}}$	✓	✓	✓
m_{jj}	mBB	✓	✓	✓
$\Delta R(\text{jet}_1, \text{jet}_2)$	dRBB	✓	✓	✓
$ \Delta\eta(\text{jet}_1, \text{jet}_2) $	dEtaBB	✓		
$\Delta\phi(V, H)$	dPhiVBB	✓	✓	✓
$\Delta\eta(V, H)$	dEtaVBB			✓
$M_{\text{eff}}(M_{\text{eff}3})$	HT	✓		
$\min(\Delta\phi(\ell, \text{jet}))$	dPhiLBmin		✓	
m_{T}^W	mTW		✓	
m_{ll}	mLL			✓
$\Delta Y(W, H)$	dYWH		✓	
m_{top}	mTop		✓	
Only in 3 Jet Events				
\vec{p}_T^{jet3}	pTJ3	✓	✓	✓
$\text{MV}_{2\text{C10}}(\text{jet}_3)^*$	$\text{MV}_{2\text{C10B3}}$	✓	✓	✓
m_{jjj}	mBBJ	✓	✓	✓

Table 1.2: Variables used to train the multivariate discriminant. Starred variables (b -tag scores) are not included in current versions of the standard discriminants, but have traditionally been included and most likely will be reintroduced as soon as their accompanying systematics are available.

⁴¹⁹ While the precise variables that would be included in 0- and 1-lepton LI and RF MVA discrimi-
⁴²⁰ nants is beyond the scope of this thesis, looking at Table 9.1, we can see the dimensionality and in-
⁴²¹ puts of the discriminants of the fiducial analysis. The correspondence for LI/RF variables and stan-
⁴²² dard variables extends nicely to the other lepton channels. The reduction in multiplicity of variables
⁴²³ owing the lower number of degrees of freedom provided by treating the Z as a single invisible par-
⁴²⁴ ticle in the 0-lepton channel would likely not be an issue, as one would just be able to use a greater
⁴²⁵ fraction of available variables in the MVA discriminant.

Noli turbare circulos meos

Archimedes

2

426

427 The Large Hadron Collider and the ATLAS

Detector

428

429 THE CERN ACCELERATOR COMPLEX AND ITS EXPERIMENTS stand as a testament to human in-
430 genuity and its commitment to the pursuit of fundamental knowledge. In this chapter, we give a

⁴³¹ cursory overview of the CERN accelerator complex, including the Large Hadron Collider (LHC),
⁴³² before moving on to a more detailed review of the ATLAS detector.

⁴³³ **2.1 THE CERN ACCELERATOR COMPLEX**

⁴³⁴ The journey of protons from hydrogen canister to high energy collisions through the CERN ac-
⁴³⁵ celerator complex, illustrated in Figure 2.1, is also one through the history of CERN’s accelerator
⁴³⁶ program. After being ionized in an electric field, protons are first accelerated in a linear accelera-
⁴³⁷ tor, LINAC 2*, to a kinetic energy of 50 MeV. From there, they are fed into the Proton Synchotron
⁴³⁸ Booster†, which further accelerates them to 1.4 GeV and, as its name implies, feeds them to the 628
⁴³⁹ m Proton Synchotron (PS, 1959[10]) and up to 25 GeV. The penultimate stage is the 7 km Super
⁴⁴⁰ Proton Synchotron (SPS, 1976; responsible for the discovery of the W and Z bosons and the 1983
⁴⁴¹ Nobel Prize [11]), which accelerates the protons to a kinetic energy of 450 GeV. Finally, these 450
⁴⁴² GeV protons are injected into the LHC[12], a proton-proton collider housed in the 27 km circumfer-
⁴⁴³ ence tunnel that housed the Large Electron Positron Collider (LEP) before its operations ceased in
⁴⁴⁴ 2000.

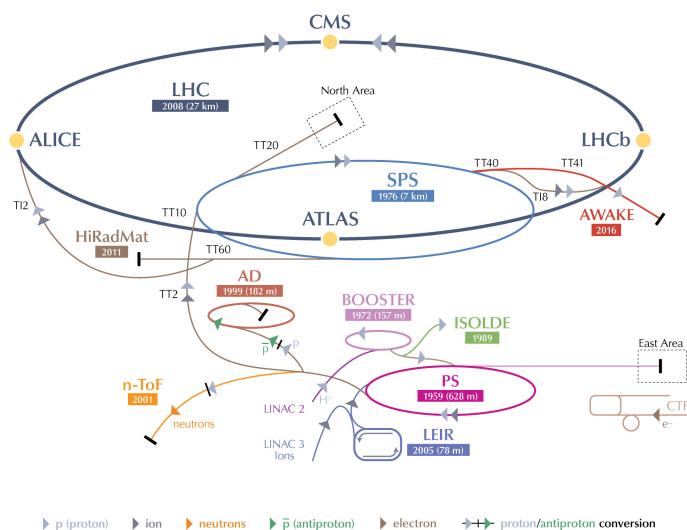
⁴⁴⁵ **2.2 THE LARGE HADRON COLLIDER**

⁴⁴⁶ The LHC was designed to function primarily as a proton-proton collider with a center of mass en-
⁴⁴⁷ ergy $\sqrt{s} = 14$ TeV and an instantaneous luminosity of $1 \times 10^{34} \text{ cm}^{-2} \cdot \text{s}^{-1}$, though it is also capable

*1978’s LINAC 2 is the successor to 1959’s LINAC 1; it will be replaced in 2020 by LINAC 4; LINAC 3 is responsible for ion production.

†Protons can be directly from a LINAC into the PS, but the higher injection energy allows for approximately 100 times more protons to be used at once[9], 1972.

CERN's Accelerator Complex



LHC Large Hadron Collider SPS Super Proton Synchrotron PS Proton Synchrotron

AD Antiproton Decelerator CTF3 Clic Test Facility AWAKE Advanced WAKEfield Experiment ISOLDE Isotope Separator OnLine Device
LEIR Low Energy Ion Ring LINAC LINear ACcelerator n-ToF Neutrons Time Of Flight HiRadMat High-Radiation Materials

©CERN 2013

Figure 2.1: The CERN Accelerator Complex [13]

⁴⁴⁸ of producing heavy ion (Pb-Pb) collisions, which it does for approximately one month in a typical
⁴⁴⁹ year of physics collisions. Owing to an accident at the beginning of the LHC's initial run, the acceler-
⁴⁵⁰ ator has operated at center of mass energies of 7, 8, and now 13 TeV.

⁴⁵¹ One of the major cost-saving features of the he LHC is that, unlike the defunct Superconducting
⁴⁵² Supercollider (SSC), its construction did not call for a purpose built tunnel, with the LHC instead
⁴⁵³ being housed in the old LEP tunnel. LEP, however, like the Tevatron, was a particle-antiparticle
⁴⁵⁴ collider, which meant that both beams could circulate within the same beam pipe, so the LEP tun-
⁴⁵⁵ nel was never built to house two separate storage rings and magnet systems (as the SSC would have
⁴⁵⁶ had). To accomplish the technically challenging task of housing two storage rings and sets of mag-
⁴⁵⁷ nets in one system, the LHC magnets feature a "twin bore" design. The magnets themselves make
⁴⁵⁸ use of superconducting NbTi cables and are cooled using superfluid helium to a temperature of 2
⁴⁵⁹ K, which allows for operational field strengths in excess of 8 T. A stable design is achieved by having
⁴⁶⁰ the magnets share a common cold mass (a 27.5 ton iron yoke for each dipole kept at 1.9 K in which
⁴⁶¹ the magnets and beam pipes are embedded) and cryostat and by arranging the superconductor wind-
⁴⁶² ings so that the magnetic fluxes of the two systems rotate in opposite directions. This results in an
⁴⁶³ extremely complicated magnetic structure. The design layout of an LHC dipole magnet is shown
⁴⁶⁴ in Figure 2.2. These dipole magnets are responsible for bending the LHC's proton beams, and their
⁴⁶⁵ strength is the principal limiting factor in the center of mass energy achievable at a circular collider.

⁴⁶⁶ The ideal version of a proton beam in the LHC consists of infinitely small bunches of protons
⁴⁶⁷ of equal momentum equally spaced in the LHC ring (itself not a perfect circle). In reality, the pro-
⁴⁶⁸ tons in the beam deviate from each of these assumptions, with dispersion in both physical space

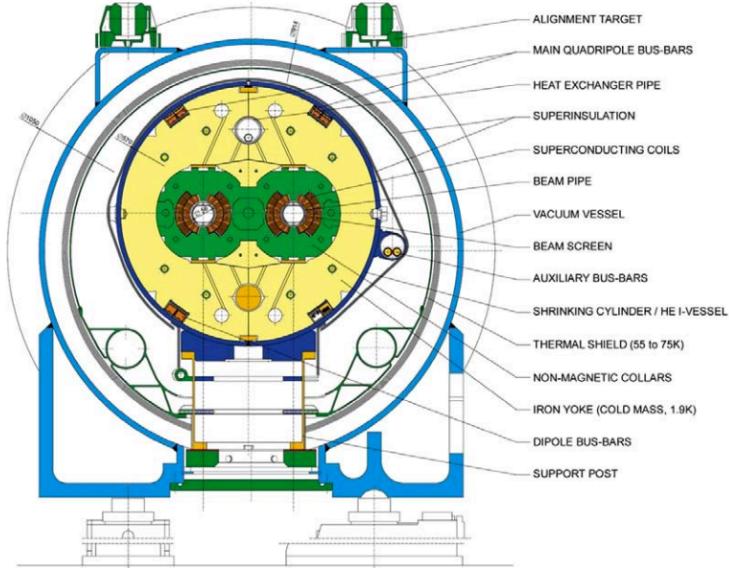
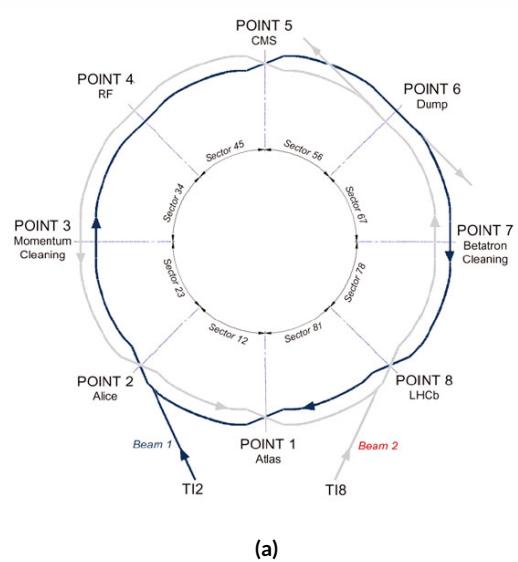


Figure 2.2: Schematic drawing of an LHC dipole magnet and cryogenics system.

and momentum space. In general, charged particles in an accelerator ring will demonstrate pseudo-harmonic “betatron” oscillations about the ideal orbit, the amplitude of which gives a characteristic of the beam’s size. In order to get high energy protons to actually collide, different magnets are used to focus the beam and help nudge deviating particles back into more ideal behavior. There are quadrupole magnet assemblies in the short straight sections to accomplish this, as well as quadrupole, octupole, and sextupole magnets interspersed throughout the length of the LHC ring for beam stabilization and other higher order corrections. The interior of the LHC beam pipe operates at a nominal pressure of $\sim 10^{-7}$ Pa, famously more rarefied than outer space.

The LHC ring itself is between 45 m and 170 m below ground and has a 1.4% incline towards Lac Léman with eight arcs and eight straight sections. In the middle of each of the eight straight sections,

479 there are potential interaction points (each colloquially referred to by its number as “Point N ”),
480 with each point housing either accelerator infrastructure or an experiment. A schematic of the con-
481 tents of each component, as well as a more detailed view of the infrastructure in the LHC ring, can
482 be found in Figure 2.4.



(a)

Figure 2.3: Schematic and detailed views of the LHC ring. IC: [14], [15]

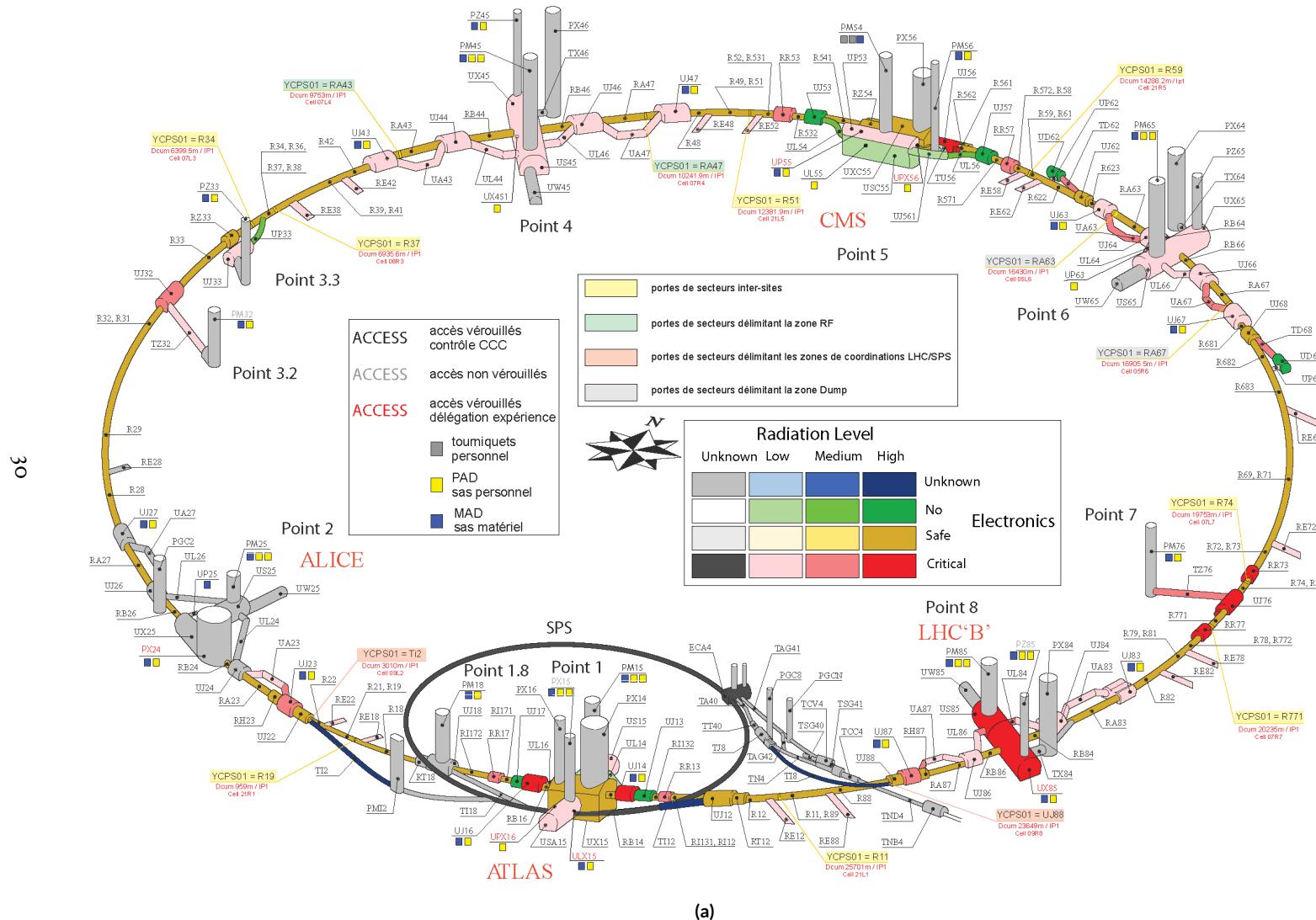


Figure 2.4: Schematic and detailed views of the LHC ring. IC:[14], [15]

483 Points 1, 2, 5, and 8 house the LHC’s experiments, ATLAS (*A Toroidal LHC ApparatuS*, one
484 of the two general purpose detectors, discussed in detail below), ALICE (A Large Ion Collider Ex-
485 periment, a dedicated heavy ion experiment), CMS (Compact Muon Solenoid, the other general
486 purpose detector), and LHCb (LHC beauty, a *B* physics experiment), respectively. Point 3 houses a
487 series of collimators that scatter and absorb particles in the beam with a large momentum deviation
488 (which will have different orbital radii) from other particles in the beam (“momentum cleaning”),
489 while Point 7 has a similar setup to remove particles with large betatron amplitudes (“betatron clean-
490 ing”). Both of these dedicated cleaning assemblies are in addition to the magnetic focusing assem-
491 blies discussed above and address the same issues. Point 4 contains the LHC’s RF (radio frequency;
492 400 MHz) acceleration system, responsible for taking protons from their injection energy of 450
493 GeV to their collision energy of 3.5, 4, 6.5, or 7 TeV. Point 6 is where the energetic ionizing radiation
494 of circulating beams can be safely taken out of the collider into a block of absorbing material, either
495 at the end of a data-taking run or in the event of an emergency (in the event of irregular behavior,
496 it is essential to do this as quickly as possible to minimize damage to the accelerator and to experi-
497 ments); this is known as a “beam dump.”

498 2.3 ATLAS AT A GLANCE

499 2.3.1 COORDINATES AND DISTANCES IN THE ATLAS DETECTOR

500 *A Toroidal LHC ApparatuS* is one of the two (the other being CMS) general purpose, high lumi-
501 nosity detectors at the LHC, located at Interaction Point 1, as described above. With a length of 44

502 m and a height of 25 m, it is the detector with largest physical dimensions at the LHC.[‡]. While pri-
 503 marily a high luminosity proton-proton collision detector, ATLAS does collect heavy ion collision
 504 data, typically for one month during a year of typical operation.

505 The ATLAS coordinate system is shown in Figure 2.5. It is a right-handed coordinate system cen-
 506 tered at the nominal collision point, with the x axis pointing towards the center of the LHC ring,
 507 the z axis pointing up, and the y axis completing the right-handed coordinate system.

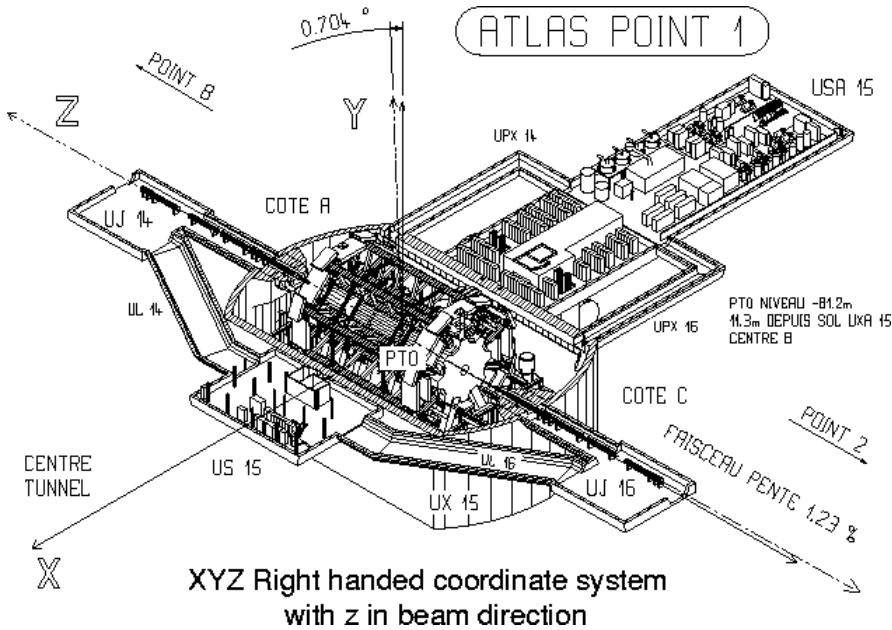


Figure 2.5: The ATLAS coordinate system. "A" side is the airport, and "C" side is "Charlie's," a pub in Saint-Genis, France.

508 While the Cartesian coordinates are useful for specifying the locations of things like detector com-
 509 ponents and activated calorimeter cells, cylindrical polar coordinates with the same origin, z axis, and
 510 handedness are often more suitable, with a point in 3-space expressed as (r, ϕ, η) . r is the perpen-

[‡]This is the only reason CMS can call itself “compact.”

511 dicular distance from the beam axis. This differs from the usual spherical ρ , the distance of a point
 512 from the origin, because the ATLAS detector is cylindrical[§], and so detector components are more
 513 easily located using r instead of ρ . In some contexts, the latter is used, though this is (or should be)
 514 made clear. ϕ is the usual (right-handed) azimuthal angle around the beam axis, with o at the $+x$
 515 axis.

516 In a lepton collider where total momentum is conserved, a useful coordinate is the relativistic
 517 rapidity of a particle:

$$y = \frac{1}{2} \ln \left[\frac{E + p_z}{E - p_z} \right] \quad (2.1)$$

518 with E and p_z as the energy and longitudinal momentum of the particle, respectively. The rapidity
 519 is the relativistic analog of a rotation angle; boosts can be added in a manner similar to rotations[¶],
 520 and differences in rapidity are invariant under boosts. In a hadronic collider, where the participants
 521 in the hard scatter are partons inside of the proton of unknown momentum fraction, longitudinal
 522 momentum is not conserved. Nevertheless, since the incident momentum is entirely longitudinal,
 523 momentum is still conserved in the transverse plane, so quantities like transverse momentum \vec{p}_T
 524 or energy (E_T)^{||} are often very useful in analysis. However, in the massless limit^{**}, we can take $E =$

[§]“toroidal;” the hole is the beam pipe

[¶]Generally, one need only insert the appropriate factor of i , the square root of -1 ; this introduces differences in sign and changes all of the trigonometric functions associated with rotations into hyperbolic trigonometric functions.

^{||}Energy is not a vector quantity, but one can take the scalar or vectorial sum of vectors formed from energy deposits with their location as the direction and energy value as magnitude. In practice, primitives are almost always assumed to be massless, so transverse energy and momentum may loosely be thought of as equivalent, with $E_T = |\vec{p}_T| = p_T$

^{**}not a terrible one for most particles depositing energy in the calorimeter; pions have masses of ~ 130 MeV, and typical energies of calorimeter objects are $\sim 10^3$ GeV, making for a boost of roughly 100.

525 $\sqrt{p_T^2 + p_z^2}$. Hence, with θ taken as the zenith angle and o corresponding to the $+z$ direction, for a
 526 massless particle, $p_z = E \cos \theta$. Using the usual half angle formula $\cos \theta = (1 - \tan^2 \theta) / (1 + \tan^2 \theta)$

527

$$\gamma = \frac{1}{2} \ln \left[\frac{1 + \cos \theta}{1 - \cos \theta} \right] = \frac{1}{2} \ln \left[\frac{(1 + \tan^2(\theta/2)) + (1 - \tan^2(\theta/2))}{(1 + \tan^2(\theta/2)) - (1 - \tan^2(\theta/2))} \right] = -\ln \left(\tan \frac{\theta}{2} \right) \quad (2.2)$$

528 This last expression, denoted η , is known as the pseudorapidity and is used instead of the polar
 529 angle as a coordinate in hadron colliders. Moreover, pion production (the most common hadronic
 530 process) is constant as a function of η in $p\bar{p}$ collisions.

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \quad (2.3)$$

531 Lower values of $|\eta|$ ($\lesssim 1.3$) correspond to more central areas of the detector known as the “barrel,”
 532 with the typical layout here being concentric, cylindrical layers. Larger values of $|\eta|$ (to ~ 2.5 for
 533 some systems and up to as much as $\sim 4.5 - 5$ for others) are known as the “end caps,” where ma-
 534 terial is typically arranged as disks of equal radius centered on the beam pipe stacked to ever greater
 535 values of $|z|$. This terminology will be useful when discussing the various subsystems of the ATLAS
 536 detector. Since decay products from a collision propagate radially (in the calorimeter portions of
 537 the detector with no magnetic field), the radial coordinate is not so important for composite physics
 538 objects like electrons or jets, which are typically expressed as momentum 4-vectors. Hence, η and ϕ
 539 are often the only useful spatial coordinates. Distances between objects are often expressed not as a

540 difference in solid angle, but as a distance, ΔR , in the $\eta - \phi$ plane, where

$$\Delta R_{12} = (\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2 \quad (2.4)$$

541 Two important concepts when discussing particles traveling through matter (e.g. particle detec-
542 tors) are radiation lengths and (nuclear) interaction lengths, which characterize typical lengths for
543 the energy loss of energetic particles traveling through materials. In general, the energy loss is mod-
544 eled as an exponential

$$E = E_0 e^{-l/L} \quad (2.5)$$

545 where E_0 is the initial energy, and L is a characteristic length. These lengths depend both on the in-
546 cident particle and the material through which they pass. In the case of uniform, composite mate-
547 rials, the length may be found by calculating the reciprocal of the sum of mass fraction weighted
548 reciprocal characteristic lengths of the components. This formula works quite well for modeling the
549 very regular behavior of electromagnetic showers (energetic photons convert into electron/positron
550 pairs, which emit photons...). In this case, L is denoted X_0 ; this is the radiation length. Hadronic
551 showers are far more complicated, with shower multiplicity and makeup being much more vari-
552 able^{††}. Nevertheless, a characteristic length can be tabulated for a standard particle type, typically
553 pions, and is called the nuclear interaction length.

^{††}Different initial hadrons will shower very differently, and hadronic showers will have phenomena like neutral pions converting to photons (which then shower electromagnetically), making them much trickier to deal with.

554 2.3.2 GENERAL LAYOUT OF ATLAS

555 The ATLAS detector and its main components are shown in Figure 2.6. ATLAS is designed as a
556 largely hermetic detector, offering full coverage in ϕ and coverage in $|\eta|$ up to 4.7. The multiple sub-
557 systems allow for good characterization of the decay products from collisions in the LHC. The in-
558 nermost system is the inner detector (ID); composed primarily of silicon pixels and strips immersed
559 in a magnetic field, it is designed to reconstruct the curved trajectories of charged particles produced
560 in collisions while taking up as little material as possible.

561 Surrounding the ID is the liquid argon based electromagnetic calorimeter (ECAL), which is de-
562 signed to capture all of the energy of the electromagnetic showers produced by electrons and pho-
563 tons coming from particle collisions. The ECAL is in turn encapsulated by a scintillating tile and
564 liquid argon based hadronic calorimeter (HCAL) that captures any remaining energy from the jets
565 produced by hadronizing quarks and gluons.

566 The outermost layer of ATLAS is the muon spectrometer (MS), which has its own magnetic field
567 produced by toroidal magnets. Muons are highly penetrating particles that escape the calorimeters
568 with most of their initial momentum, so the MS and its magnets are designed to curve these charged
569 particles and measure their trajectories to measure their outgoing momenta. Each of these detector
570 systems has several principal subsystems and performance characteristics, which will be described in
571 turn below.

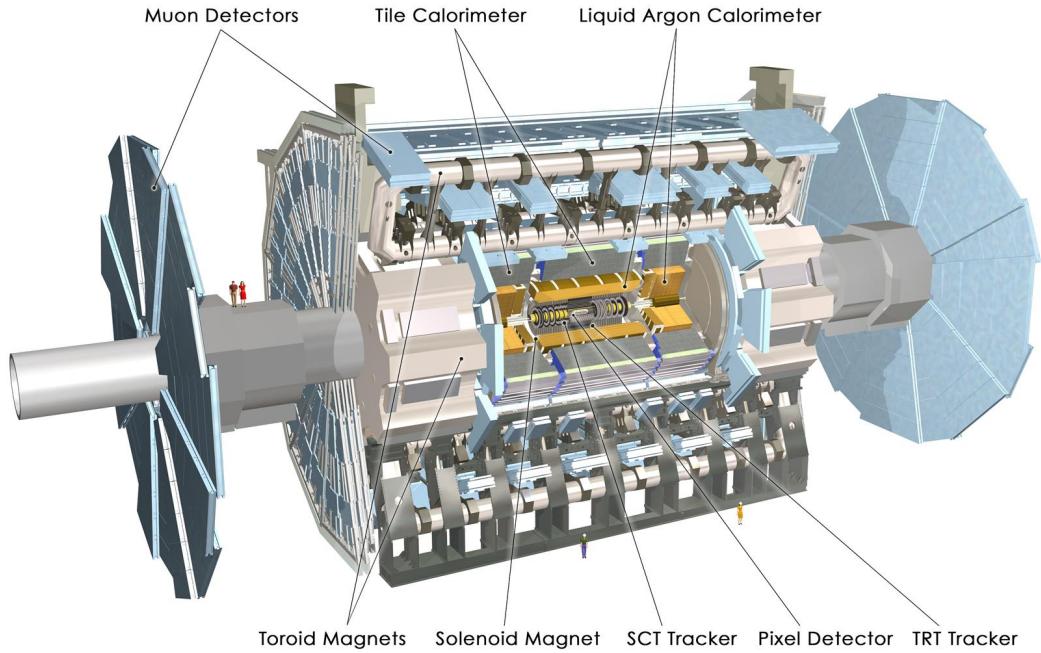


Figure 2.6: The ATLAS detector with principal subsystems shown.

572 2.4 THE INNER DETECTOR

573 ATLAS's inner detector (ID) is surrounded by a 2 T superconducting solenoid that is cryogenically
 574 cooled to a temperature of 4.5 K. The ID uses two silicon detector subsystems (the Pixel and Semi-
 575 Conductor (strip) Tracker (SCT)) to track the curved trajectories of charged particles emanating
 576 from particle collisions and a Transition Radiation Tracker (TRT) composed of gas straw detectors
 577 with filaments for e/π discrimination, as shown in Figure 2.7. The ID offers full coverage in ϕ and
 578 extends to an $|\eta|$ of 2.5.

579 Since the components of the ID do not provide an energy measurement, it is desirable for a track-
 580 ing system to have as small a material budget as possible so that more accurate energy measurements

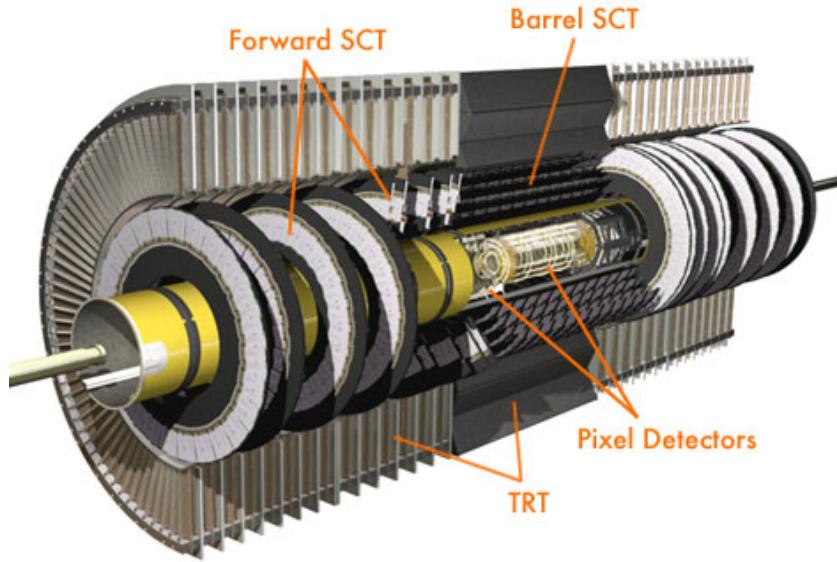


Figure 2.7: The ATLAS inner detector. IC: [16]

581 may be done in the calorimeters. Generally, there are two radiation lengths in the inner detector (the
 582 precise figure varies with η); the full material budget, with the layout of the individual layers in each
 583 subsystem, can be seen in Figure 2.8.

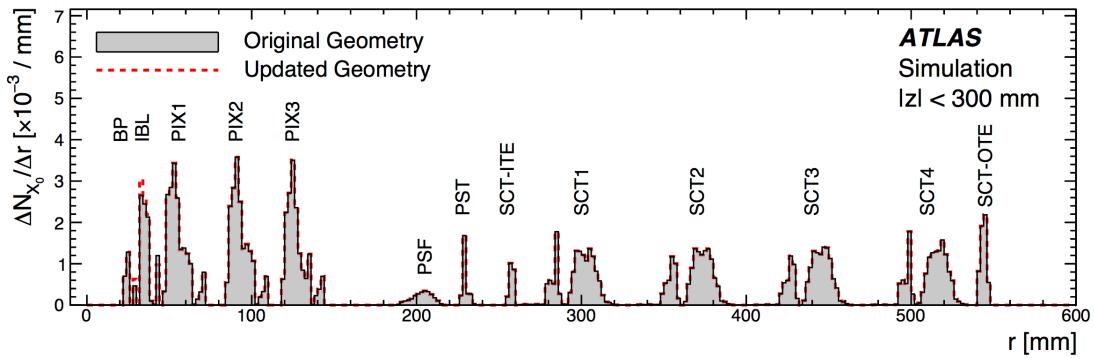


Figure 2.8: The ID material budget. IC: [17]

584 2.4.I THE PIXEL DETECTOR

585 The innermost part of ATLAS is the Pixel Detector, which, as the name suggests, is comprised of
586 four layers of silicon pixels in the barrel at 32, 51, 89, and 123 mm from the beam pipe, and three lay-
587 ers in the end caps at 495, 580, and 650 mm from the beam pipe, with over 80 million channels total.

588 The innermost layer of pixels, the insertable *B* layer (IBL) was installed during the 2013–14 LHC
589 shutdown. The pixels are cooled to a temperature of $\sim -5^\circ\text{C}$, with N_2 gas and operate at 150–600
590 V. The pixels themselves come in two sizes $50 \times 400(600) \times 250 \mu\text{m}$, with the larger pixels in the
591 outer layers. They provide nominal resolution of $10(115) \mu\text{m}$ resolution in $r - \phi$ direction.

592 In order to improve total coverage in the detector and prevent any gaps, pixels are not installed
593 flush with each other. Pixels in the barrel are tilted at about 20° , with an overlap in $r - \phi$, as shown
594 in Figure 2.9. The disks of the ID end caps are rotated with respect to each other by 3.75° .

595 2.4.2 THE SILICON MICROSTRIP DETECTOR (SCT)

596 The layout of the SCT is similar to that of the Pixel detector, except that, for cost considerations, the
597 SCT uses silicon strips. These strips are also cooled to $\sim -5^\circ\text{C}$ with N_2 gas and operate from 150–
598 350 V. Strip dimensions are $80 \times 6000 \times 285 \mu\text{m}$, and provide nominal $17(580) \mu\text{m}$ resolution in
599 $r - \phi(z)$. Barrel strips feature an 11° tilt and come in four layers at 299, 371, 443, and 514 mm. There
600 are nine end cap disks on each side at z values varying from 934–2720 mm.

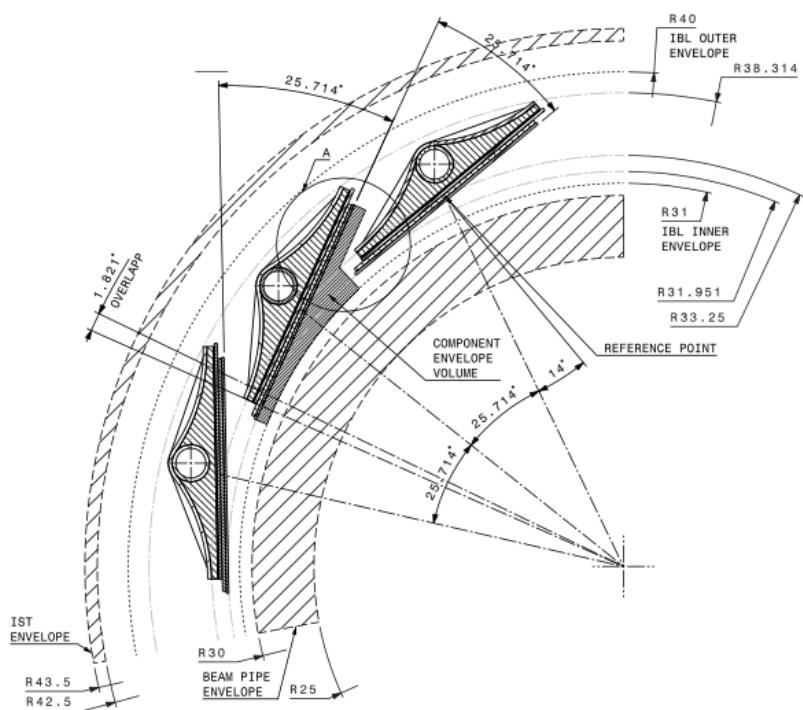


Figure 2.9: Arrangement of pixels in the barrel. IC: [18]

601 2.4.3 TRANSITION RADIATION TRACKER (TRT)

602 The final and outermost subsystem in the ID is the Transition Radiation Tracker (TRT). It provides
603 coverage for $|\eta|$ up to 2.0 and is composed of straw detectors with a 4 mm diameter that run the
604 length of the detector module. The straws provide $130 \mu\text{m}$ resolution, are filled with a Xe-CO₂-O₂
605 (70-27-3) gas combination, and operate at -1500 V. The filaments and foil lining inside the straws
606 induce X-ray emission in electrons and pions passing through the TRT as they move from a dielec-
607 tric to a gas; this “transition radiation” is the source of the TRT’s name. Since the energy deposited
608 due to transition radiation is proportional to the relativistic boost γ , for constant momentum, this
609 is inversely proportional to mass. Thus, electrons will have $\sim 130/0.5 = 260\times$ more transition
610 radiation than pions, in principle enabling excellent electron/pion discrimination. The TRT will be
611 replaced by silicon strips in the Phase II upgrade.

612 2.5 THE ATLAS CALORIMETERS

613 ATLAS has four main calorimeter systems: the liquid argon based Electromagnetic Calorimeter
614 (ECAL), the Hadronic End Cap (HEC), the Forward Calorimeters (FCAL), and the scintillating
615 tile based hadronic Tile Calorimeter in the barrel. Their layout and material budget in interaction
616 lengths can be seen in Figure 2.II.

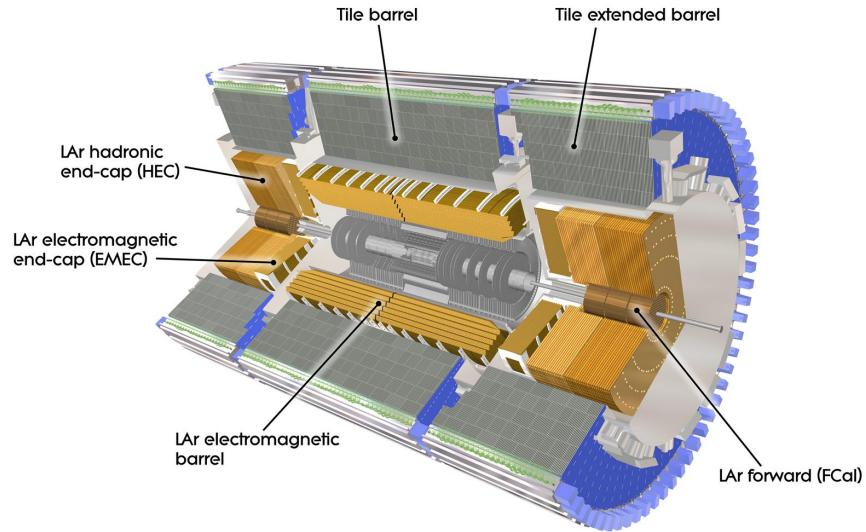


Figure 2.10: The ATLAS calorimeters.

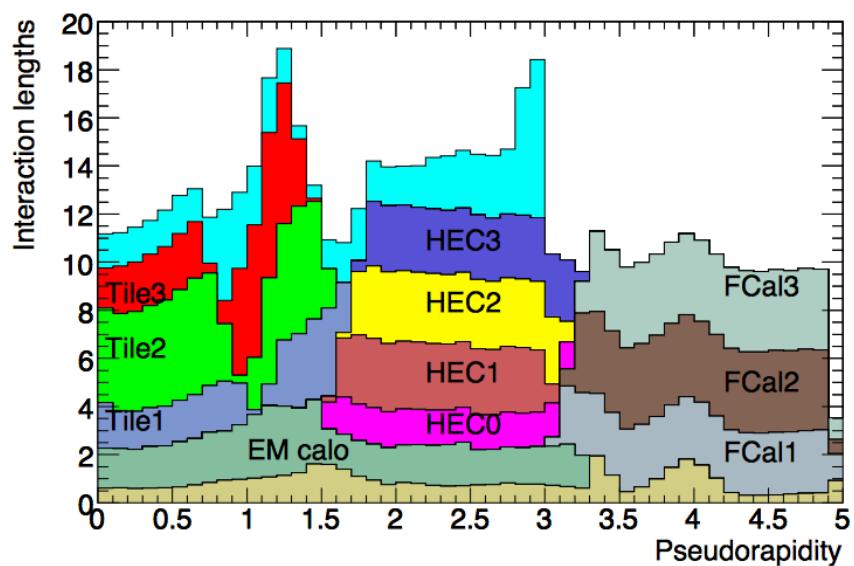


Figure 2.11: Material depth of the ATLAS calorimeters. IC: [16]

617 2.5.1 CALORIMETER RESOLUTION

618 Before diving into the specifics of each of the ATLAS calorimeters, we review some aspects of calorime-
619 ter energy resolution performance. A calorimeter’s relative energy resolution (a ratio) can be broken
620 up into three orthogonal components, as shown in Equation 2.6.

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \quad (2.6)$$

621 S is the photoelectron statistics or stochastic term and represents the coefficient to the usual count-
622 ing term (assuming Gaussian statistics); N is a noise term, which is constant per channel (and hence
623 comes in as $1/E$ in the relative energy resolution); and C is a constant “calibration” term, which re-
624 flects how well one intrinsically understands a detector (i.e. mismodelling introduces an irreducible
625 component to the energy resolution). If any detector were perfectly modeled/understood, it’s C
626 term would be zero. $N \sim 0.1 - 0.5$ GeV for a typical calorimeter regardless of type, so S and C
627 are typically quoted.

628 A typical stochastic term scales as $S \sim \text{few\%} \sqrt{d_{\text{active}} [\text{mm}] / f_{\text{samp}}}$, where f_{samp} is the sampling
629 fraction or the ratio of a calorimeter by mass is composed of an active material (i.e. one that regis-
630 ters energy deposits). The tile calorimeter, for example, has a sampling fraction of about 1/36. There
631 are several reasons that this fraction is so low. First, many active volumes have insufficient stopping
632 power; one wants to capture as much energy as possible from electromagnetic and hadronic showers
633 inside the calorimeter, and this simply is not possible for most active media (one notable exception

634 to this is the CMS crystal-based calorimeter; ATLAS is a more conservative design), so well-behaved
635 absorbers like lead or iron are necessary to ensure all the energy is contained within a calorimeter.
636 Another factor is cost; things like liquid argon are expensive. Finally, most active media are unsuit-
637 able for structural support, so sturdy absorbing materials help relieve engineering constraints.

638 **2.5.2 THE ELECTROMAGNETIC CALORIMETER (ECAL)**

639 The ECAL has liquid argon (LAr) as an active material and lead as an absorber. The ECAL barrel
640 extends to $|\eta|$ of 1.475, with three layers at 1150, 1250, and 2050 mm, and its end cap, comprised of
641 two wheels, covers $1.375 < |\eta| < 2.5$, (3.2) for the inner (outer) wheel, with 3 (2) layers out to
642 3100 mm. There is a 1.1 (0.5) cm thick layer of LAr pre-sampler up to $|\eta|$ of 1.8 in the barrel (end cap)
643 of the ECAL, which is designed to aid in correcting for electron and photon energy loss in the ID.

644 The LAr and lead absorber are arranged in alternating, beveled, sawtooth layers in what is known
645 as an “accordion” geometry, shown in Figure 2.12, which shows the layout of a barrel module in the
646 ECAL. The absorber thickness is 1.53 (1.13) mm for $|\eta|$ less (more) than 0.8 to ensure a constant sam-
647 pling fraction. This arrangement helps provide greater coverage in ϕ .

648 The ECAL overall typically covers 2–4 interaction lengths or about 20–40 radiation lengths. Its
649 performance corresponds to resolution coefficients $S = 0.1 \text{ GeV}^{-1/2}$ and $C = 0.002$ with a 450
650 ns drift time. In order to optimize the material budget and overall detector construction, the ECAL
651 barrel infrastructure is integrated with that of the ID’s solenoid. The granularity of the ECAL barrel
652 middle layer, $\Delta\eta \times \Delta\phi$ cells of size 0.025×0.025 , are used to define the granularity of calorimeter
653 cluster reconstruction in ATLAS.

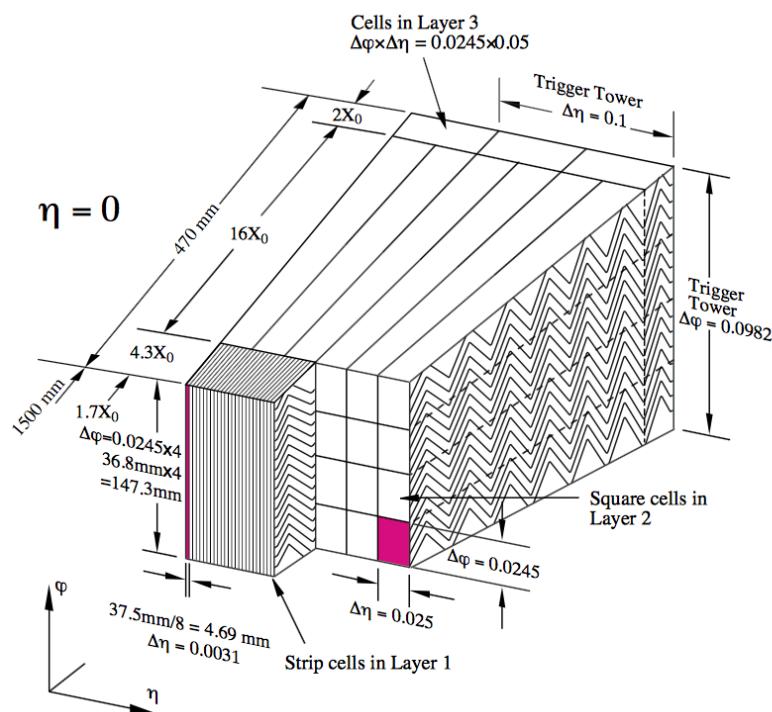


Figure 2.12: The accordion geometry of the LAr electromagnetic calorimeters is prominently shown in this illustration of an ECAL barrel module. IC: [16]

654 2.5.3 HADRONIC END CAPS (HEC)

655 The HEC covers an $|\eta|$ range of 1.5 to 3.2. Like the ECAL end caps, the HEC consists of two identi-
656 cal wheels out to a distance from the beam axis of 2030 mm; its layout is shown in Figure 2.13. The
657 HEC also has LAr as the active material, but instead has flat copper plates as absorbers for sampling
658 fraction of 4.4% and 2.2% in the first and second wheels, respectively. Its granularity in $\eta - \phi$ is
659 0.1×0.1 for $|\eta|$ up to 2.5 and 0.2×0.2 in the more forward regions.

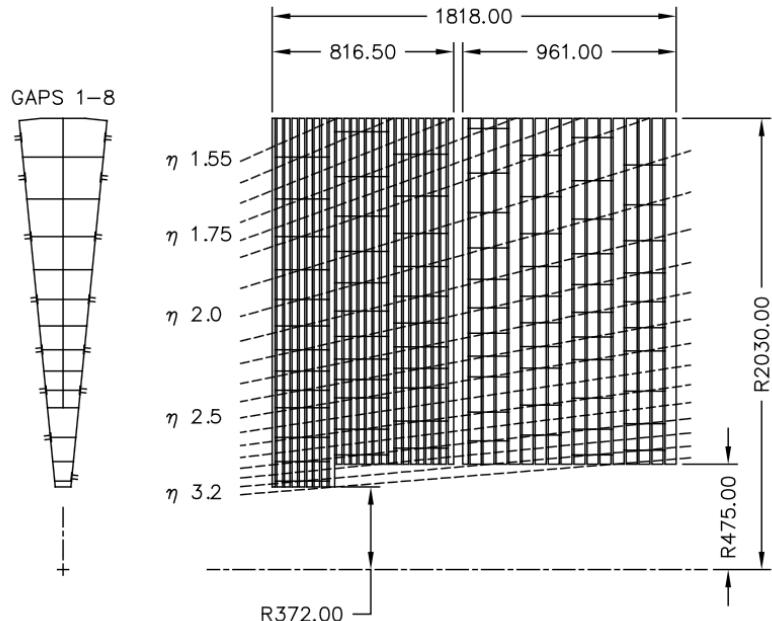


Figure 2.13: The layout of the HEC in $r - \phi$ and $r - z$; dimensions are in millimeters. IC: [16]

660 2.5.4 THE FORWARD CALORIMETER (FCAL)

661 The FCAL covers an $|\eta|$ range from 3.1 to 4.9, again using LAr as the active material in gaps between
662 rods and tubes in a copper-tungsten matrix, as shown in Figure 2.14. These system has characteris-
663 tic performance corresponding to stochastic term of $S \approx 1 \text{ GeV}^{-1/2}$. There are three modules in
664 the FCAL: one electromagnetic and two hadronic, with the latter two featuring a higher tungsten
665 content for a larger absorption length.

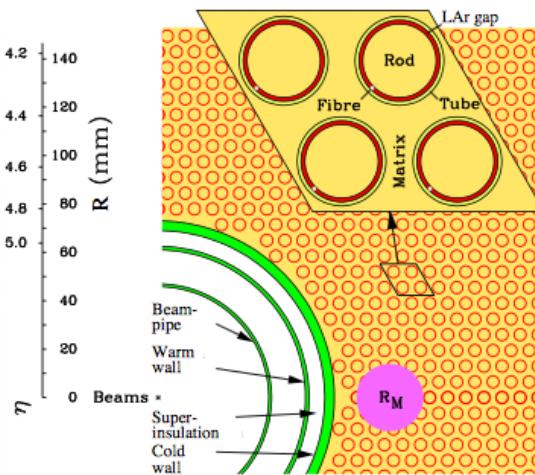


Figure 2.14: The material layout for a typical section of the FCAL in the transverse plane. IC: [16]

666 2.5.5 THE HADRONIC TILE CALORIMETER

667 The tile calorimeter, covering an $|\eta|$ of up to 1.7 is made up of 64 modules in the barrel (each cover-
668 ing $\Delta\phi$ of $360/64 = 5.625^\circ$), each with a layout as in Figure 2.15. It is designed to be self-supporting
669 for structural reasons, and so is the only calorimeter without LAr as a an active medium, with a stag-
670 gered matrix of active scintillating polystyrene and supporting steel. It operates at 1800 V with a 400

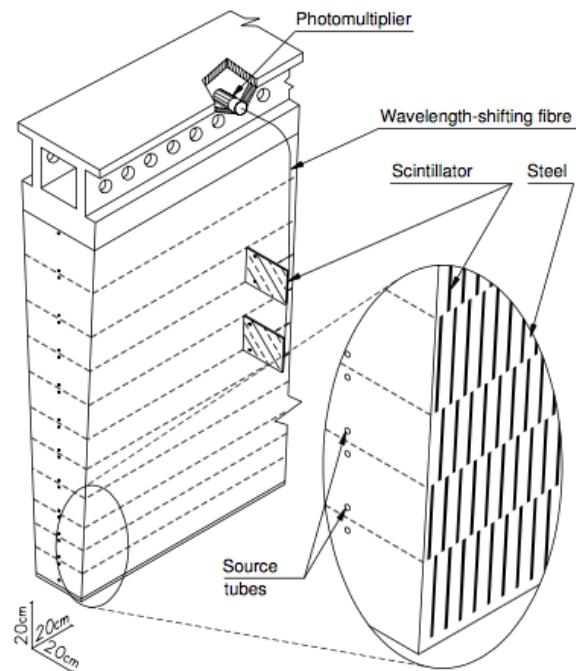


Figure 2.15: The material layout for a typical section of the hadronic tile calorimeter. IC: [16]

671 ns dead time and has a thickness corresponding to 10–20 interaction lengths (2.28–4.25 m). Its cells
672 have a $\Delta\eta \times \Delta\phi$ granularity of 0.1×0.1 in the first two layers and 0.2×0.1 in the last layer. Its
673 performance corresponds to $S = 0.5 \text{ GeV}^{-1/2}$ and $C = 0.05$ (0.03 after calibration).

674 2.6 THE MUON SPECTROMETER

675 Since the energy of muons is not captured within the calorimeters, the stations of the ATLAS MS
676 surround the entire detector and provide tracks of outgoing muons that can be matched to tracks in
677 the ID. The ATLAS toroids, which provide field strengths of up to 2.5 (3.5) T in the barrel (end cap)
678 with typical strengths of 0.5–1.0 T, bend the muons, which allows for a muon momentum measure-
679 ment since the muon mass is known. The relative momentum resolution of a tracker (assuming, as
680 in ATLAS, that bending primarily happens in the ϕ direction) may be expressed as

$$\frac{\sigma_{p_T}}{p_T} = c_0 \oplus c_1 \cdot p_T \quad (2.7)$$

681 The c_0 term represents a degradation in resolution due to multiple scattering, and is typically 0.5–
682 2% [19]. The c_1 term describes the phenomenon of, holding magnetic field constant, higher momen-
683 tum muons curving less. This term has typical values of $10^{-3} - 10^{-4} \text{ GeV}^{-1}$. At very high p_T val-
684 ues, this is of particular concern since a very small curvature can result in charge misidentification.

685 A cross-sectional view (in $r-z$) of the muon spectrometer with station names, detector types, and
686 layouts is shown in Figure 2.16. There are three layers of muon detectors in both the barrel (at 5 000,
687 7 500, and 10 000 mm) and end cap (at 7 000 (11 000), 13 500, and 21 000 mm), with the innermost

688 end cap layer split in two due to the end cap toroid. This corresponds to an $|\eta|$ range up to 2.4 for
 both precision and trigger coverage, and up to 2.7 for precision detection only.^{†‡}

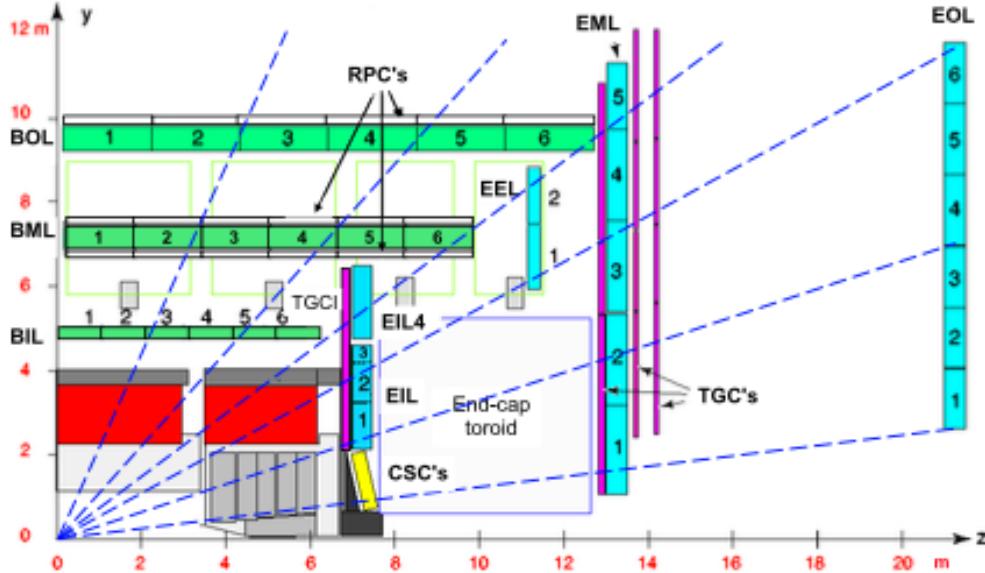


Figure 2.16: The ATLAS muon spectrometer. Naming of the MDT stations obeys the following convention [BE] (barrel or end cap) [IEMO] (inner, inner extended (end cap only), middle, or outer layer) [1-6] (increasing in z (r) for the barrel (end cap)), so E11 is the station in the inner most end cap layer closest to the beam pipe. IC: [16]

689
 690 The MS can reconstruct muons with transverse momenta from 5 GeV up to 3 TeV (with 10% res-
 691 olution at 1 TeV (3% at 100 GeV)). Detectors in the MS fall into two broad headings, precision detec-
 692 tors and trigger detectors, both described below. Nominal performance of the current detector types
 693 in the MS is summarized in Figure 2.17, a table taken from [16]. It should be noted that $|\eta|$ ranges
 694 quoted below, where applicable, do not include the range 0-0.1, where this a gap in the MS to allow
 695 for cabling and other services to the ATLAS detector; for a discussion of compensatory measures in

^{†‡}This will change with the New Small Wheel Phase I Upgrade. cf. Appendix A

⁶⁹⁶ muon reconstruction, see Chapter 5.

Type	Function	Chamber resolution (RMS) in			Measurements/track		Number of	
		z/R	ϕ	time	barrel	end-cap	chambers	channels
MDT	tracking	35 μm (z)	—	—	20	20	1088 (1150)	339k (354k)
CSC	tracking	40 μm (R)	5 mm	7 ns	—	4	32	30.7k
RPC	trigger	10 mm (z)	10 mm	1.5 ns	6	—	544 (606)	359k (373k)
TGC	trigger	2–6 mm (R)	3–7 mm	4 ns	—	9	3588	318k

Figure 2.17: ATLAS MS detector performance. IC: [16]

⁶⁹⁷ 2.6.1 PRECISION DETECTORS

⁶⁹⁸ The ATLAS MS has two types of precision detectors: Monitored Drift Tubes (MDT's) and Cathode Strip Chambers (CSC's). An MDT is a tube with a 3 cm diameter with length depending on
⁶⁹⁹ the station in which the tube is located. The tube is filled with an Ar/CO₂ gas mixture and has a
⁷⁰⁰ tungsten-rhenium wire at its center that is kept at 3 000 V when operational. The MDT's provide 35
⁷⁰¹ μm resolution (per chamber) in their cross-sectional dimension (there is no sensitivity along the axis
⁷⁰² of the wire). Resolution of this magnitude requires very precise knowledge of the location of the
⁷⁰³ wires within the MDT's; this is generally true for detectors in the MS (trigger as well as precision);
⁷⁰⁴ to this end, stations of the MS are aligned using an optical laser system. For a detailed description
⁷⁰⁵ of how misalignment can affect performance, see Appendix A for a detailed discussion of misalign-
⁷⁰⁶ ment's simulated effects on the performance of the proposed Micromegas trigger processor in the
⁷⁰⁷ New Small Wheel (NSW) of the Phase I upgrade. Their 700 ns dead time, however, precludes their
⁷⁰⁸ use as trigger detectors and also in the region of the small wheel (innermost endcap) closest to the
⁷⁰⁹ beam pipe ($|\eta|$ from 2.0 to 2.7), where rates are highest.
⁷¹⁰

711 In this region, the precision detectors are the CSC's, which have a much lower dead time of ~ 40
712 ns. These are multiwire proportional chambers with cathode planes that have orthogonal sets of
713 strips, allowing for a measurement in both principal directions. CSC detector sizes also vary by sta-
714 tion, coming in both small and large chambers. The CSC strip pitch is 5.31 (5.56) mm for the large
715 (small) chambers, with position determined from the induced charge distribution in the strips. This
716 corresponds to a nominal resolution of 60 (5 000) μm per plane in the bending (non-bending) direc-
717 tion. These are slated to be replaced by Micromegas detectors in the NSW.

718 **2.6.2 TRIGGER DETECTORS**

719 Trigger detectors have a fundamentally different role than the precision detectors, instead needing to
720 deliver “good enough” approximate values of muon track positions and p_T values. The MS has two
721 types of trigger detectors: Resistive Plate Chambers (RPC’s) in the barrel and Thin Gap Chambers
722 (TGC’s) in the end caps. They collectively cover an $|\eta|$ range to 2.4, and their arrangement is shown
723 in Figure 2.18.

724 The RPC’s are parallel plate detectors with a dead time of 5 ns and a thickness of 2 mm, kept at
725 a potential of 9 800 V; they are deployed in three layers. RPC’s, too, feature strips with orthogonal
726 arrangements on the top and bottom planes, with a strip pitch of 23–35 mm.

727 The TGC’s are multiwire proportional chambers with a dead time of 25 ns. Also, featuring or-
728 thogonal strips, the TGC’s also provide a ϕ measurement to compensate for the lack of MDT sensi-
729 tivity in this direction. There are four layers of TGC’s in the end cap. TGC’s will be supplanted by
730 sTGC’s (small thin gap chambers) in the NSW.

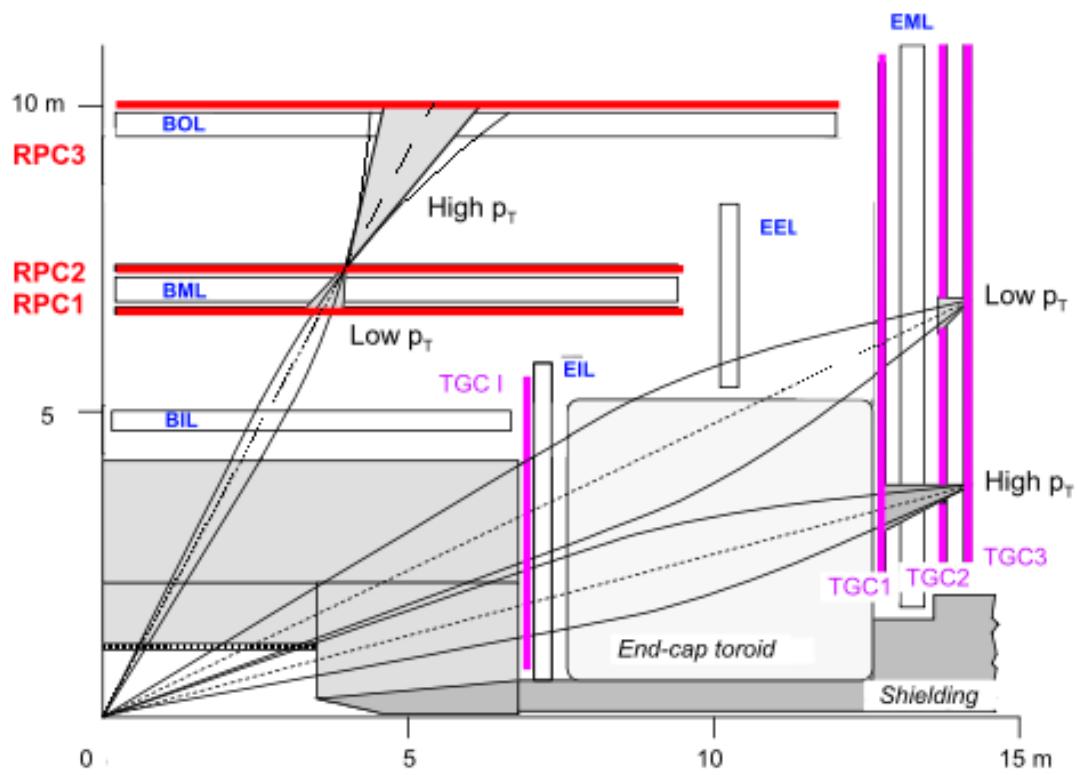


Figure 2.18: ATLAS MS trigger detector arrangement. IC: [16]

⁷³¹ For more details on how detector level trigger objects work in the ATLAS MS, see Appendix A
⁷³² for details on the Micromegas trigger processor algorithm.

What do you read, my lord?

Words, words, words.

Hamlet, 2:2

3

733

734

Data and Simulated Samples

735 THE DATA AND Monte Carlo simulation (MC) samples used in this thesis are the same as in the fidu-
736 cial analysis. The data corresponds to 36.1 fb^{-1} of pp collision data collected in 2015+16 at the AT-
737 LAS detector at $\sqrt{s} = 13 \text{ TeV}$. Details of the Run 1 analysis referenced in Chapter 9, may be found
738 in [20]. Only events recorded with all systems in ATLAS in good working order and passing certain

739 quality requirements, according to a Good Run List (GRL), are analyzed.

740 Details about MC samples may be found in [21], and signal and background modeling are dis-
741 cussed in the next. The $ZH \rightarrow \ell\ell b\bar{b}$ process is considered for both multivariate analysis (MVA)
742 optimization and the final statistical analysis, while $WH \rightarrow \ell\nu b\bar{b}$ and $ZH \rightarrow \nu\nu b\bar{b}$ production
743 are included in the final statistical analysis only. Signal MC samples were generated separately for qq
744 and gg initiated VH processes. $qqVH$ samples were generated with Powheg MiNLO + Pythia8
745 [22, 23] with the AZNLO [24] tune set and NNPDF3.0 PDF [25], with alternate samples gener-
746 ated using MadGraph5_AMC@NLO [26] for the hard scatter generation and Pythia8 for the
747 hardronization, parton shower (PS), underlying event (UE), and multiple parton interactions (MPI).
748 Nominal $ggZH$ samples were generated using Powheg for the matrix element (ME) and Pythia8
749 for the parton shower (PS), underlying event (UE), and multiple parton interactions (MPI), again
750 applying the AZNLO tune and NNPDF3.0 PDF set. [27]

751 The background processes considered in these studies are $Z+jets$, $t\bar{t}$, and diboson production for
752 both MVA optimization and the final statistical analysis with single top production and $W+jets$
753 only considered in the final statistical analysis. $V+jets$ samples are generated using Sherpa 2.2.1 [28]
754 for both the ME and PS. These samples are generated in different groups, according to the identity
755 of the V , the max (H_T, p_T^V) of events, with further subdivisions according to the flavor of the two
756 leading jets in an event, b , c , or l , for a total of six categories. $t\bar{t}$ samples are generated using Powheg
757 with the NNPDF3.0 PDF set interfaced with Pythia8 using the NNPDF2.3 PDF's and the Al4
758 tune [29]. Single top samples use Powheg with the CT10 PDF's interfaced with Pythia6 using
759 the CTEQ6L1 PDF's [30, 31]. Diboson samples are generated with Sherpa 2.2.1 interfaced with the

⁷⁶⁰ NNPDF_{3.0} NNLO PDF set normalized to NLO cross sections [[32](#)].

*There are certain calculations one simply doesn't do in
public.*

Alan Blaer

4

761

762

Signal and Background Modeling

763 THIS CHAPTER summarizes the modeling of the dominant signal and background processes in
764 this analysis, including corrections and systematic uncertainties (systematic uncertainty, also called
765 nuisance parameter (NP), titles are set in this font) related to each process. Further details on the
766 specifics of these topics, including in-depth studies for the derivation and definitions of some of the

767 quantities cited, may be found in [21]. We start with a general discussion of modeling and associated
768 major categories of uncertainties before addressing each of the physics processes.

769 **4.1 EVENT GENERATION IN A NUTSHELL**

770 Before diving into the specifics of modeling and systematic uncertainties associated with each ma-
771 jor set of physics processes considered in this analysis, we review at a schematic level* the problem
772 of simulation event generation. Once a physics processes of interest has been determined, how one
773 simulates an ensemble of particle collisions to model the process in question. This is illustrated in
774 Figure 4.1. Note that the scope of this problem does not include how these generated collision prod-
775 ucts propagate through one's detector. This problem is left for Chapter 5.

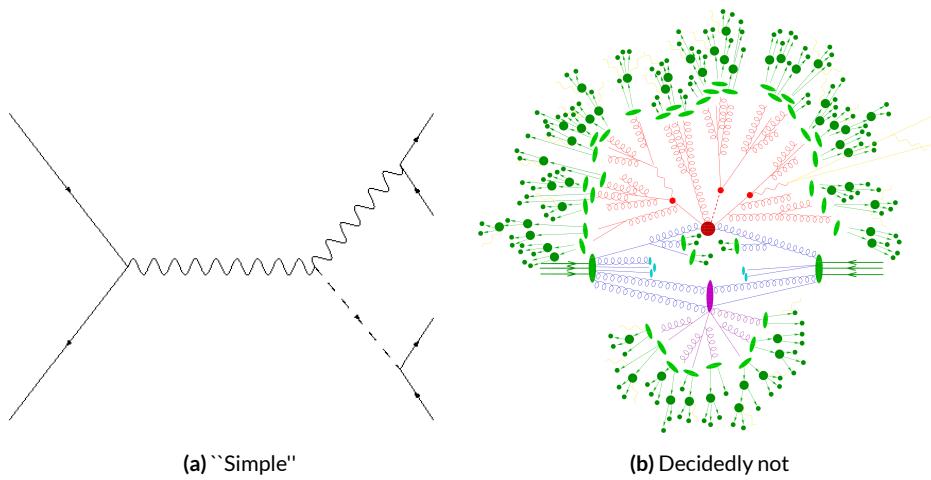


Figure 4.1: The problem here is how to get from (a) to (b).

776 The primary source of complication in event generation comes from dealing with hadronic ob-

*i.e. this will not be a technically rigorous discussion. For a more thorough treatment, the reader is di-
rected to the usual references.

jects both in the initial state (the lefthand side of Figure 4.1 (a); the LHC is a hadron collider) and the final state (this analysis searches for Higgs decays to b -jets, the lower righthand side of Figure 4.1).
Common to all hadronic objects, by definition, are the many considerations that go into calculations in quantum chromodynamics (QCD). In calculating the hard scatter process itself, one must make a variety of choices, such as the parton distribution function (PDF) set to use and to what order in perturbation theory to do the calculation (common choices are leading order (LO), (next to) next to leading order ((N)NLO), and (next to) next to leading log (NNLL)) [33]. Similar considerations often need to be made for the electroweak parts of an event. These considerations and others in event generation (broadly called event generator “tuning”) will be discussed in more detail below.

The initial state includes not only the hard scatter partons that generate the physics process of interest, but also the rest of the partons in the colliding protons, known as the underlying event (UE). Moreover, the hard scatter partons may not be the only interacting partons in an event, further complicating matters; this phenomenon is known as multiple parton interactions (MPI). Specific to the final state are the kinematic distributions of the final state objects—what their energies and angular distributions will be—in addition to the overall cross section of the process that is measurable by the detector (acceptance effects). Furthermore, one has to model hadronization, the process by which any free (colored) partons in an event transform into colorless hadrons.

Typically, it takes several steps and tools to accomplish this. The hard scatter itself is often modelled with a dedicated event generator like PowHEG [22] or MADGRAPH [26], with events generated then interfaced with a tool like PYTHIA [23] for the PS, UE, and MPI, though there are exceptions (SHERPA [34], for example, can do both the hard scatter and hadronization/ for some pro-

798 cesses).

799 4.2 DESCRIPTION OF MODELING UNCERTAINTY CATEGORIES

800 Each of the steps in event generation described above has associated uncertainties. Some uncertain-
801 ties are inherent in the calculations themselves. The choice of which order in perturbation theory
802 to do a calculation, for example, comes with it an implicitly defined level of precision[†]. Extrapolat-
803 ing from one energy/momentum scale to another also introduces uncertainty. Furthermore, there
804 is no *a priori* correct choice to make at each step in event generation, so each choice (the choice of
805 generator, PDF set, parton shower calculator, all of their configurable parameters, etc.) implies an
806 additional layer of uncertainty.

807 In order to quantify these choices, each source of systematic uncertainty is treated separately and
808 given a unique name. To make this more concrete, take the specific example of the uncertainty asso-
809 ciated with the $H \rightarrow b\bar{b}$ branching ratio of 58%, called `ATLAS_BR_bb`, which encapsulates a num-
810 ber of effects (higher order terms, the mass of the b quark, and choice of α_S). The quoted (in prin-
811 ciple asymmetric) uncertainty on the Higgs BR is not itself a direct input into the analysis model.
812 Instead, the effect of varying the branching ratio up and down by one standard deviation is propa-
813 gated to simulated collision events and recorded (i.e. the analysis is run with the Higgs branching
814 ratio at $\pm 1\sigma$, and the results are recorded alongside the nominal result). The nominal and “up” and
815 “down” variations are then typically taken to define a normally distributed, freely floating param-

[†]though this is less well-defined in QCD calculations than for electroweak calculations since they don’t converge

816 eter in the statistical fit model. Since these parameters associated with systematic uncertainties are
817 not typically considered interesting quantities, they are often referred to as “nuisance parameters”
818 (NP’s). The terms “systematic,” “systematic uncertainty,” and “nuisance parameter” are often used
819 interchangeably.

820 The specifics of exactly how the effects of variations are saved and propagated to the full fit model
821 are deferred to Chapter 7. The discussion here is confined to how systematic uncertainties for signal
822 and background modeling and their accompanying variations are defined. Modeling systematics are
823 derived separately for each physics process (simulation sample). Sometimes, all of the variation for
824 a given process is encapsulated in a single systematic, but oftentimes the variations from multiple
825 considerations are distinct enough to be treated separately. Furthermore, each of these separate sys-
826 tematics for a given sample/process may be treated in a number of ways (e.g. 0-lepton events may
827 be treated differently from 2-lepton events). An additional subtlety is that a continuous parameter
828 like a branching ratio lends itself quite naturally to defining Gaussian $\pm 1\sigma$ variations, while for dis-
829 crete variations, like choice of PDF set for parton showers, how to proceed is less obvious. This is
830 addressed on a case-by-case basis, as described below.

831 Before enumerating each of the principal physics processes and their systematics, we begin by
832 describing considerations and choices that must be addressed for every physics process in order to
833 make the discussion of individual samples and systematics both clearer and less repetitive.

834 4.2.1 PHYSICS CONSIDERATIONS

835 In general, evaluating the uncertainties arising from the many choices in event generation entails
836 producing alternate samples of events, which practically means tuning parameters in the various soft-
837 ware packages and/or using alternate packages/libraries to make new samples. Once these samples
838 have been created, they are compared at truth-level (particle level) using a package called Rivet [35]
839 instead of using the full ATLAS detector reconstruction for computational considerations. Given
840 the nature of the problem and the tools, there are generally three main categories of physics issues,
841 each described below.

842 UNDERLYING EVENT AND PARTON SHOWER

843 The modeling of the underlying event (UE) and parton shower (PS) are usually handled by the same
844 package and so are usually treated together. The typical nominal choice in the fiducial analysis is
845 PYTHIA8. One approach to modeling these uncertainties is simply to see what happens when a
846 different model is used and then compare this alternate set of events to the nominal set, taking the
847 difference as the (implicitly one standard deviation) scale of variation. Another approach is to vary
848 some parameter within a given model, for example, using different tunes in the A14 set for PYTHIA8
849 with their accompanying variations, to characterize the scale of variation.

850 A natural question is how to treat these two approaches on the same footing. When examining
851 a set of potential variations related to the same process or effect, oftentimes the largest single varia-
852 tion in a set is picked as defining the scale for the systematic uncertainty. Another approach is to use

853 the average over a set of variations.[‡] The `ATLAS_UEPS_VH_hbb` systematic, for example, uses the
854 `Pythia8 + A14` tunes approach to determine the scale of UE variation and compares `Pythia8` with
855 `Herwig7` to characterize the PS variation. Each of the `A14` tunes comes with an up and down varia-
856 tion, and the difference between each of these variations and a nominal setup may be expressed as a
857 ratio, R , of total events.

858 As is often done when a physical argument can be made for combining related, but ultimately
859 orthogonal categories/measurements/uncertainties/systematics, the combined UE+PS systematic is
860 taken to be the sum in quadrature of these two effects:

$$\sum_{tunes} \max_{tune} (|R_{up} - R_{down}|) \oplus \sigma_{PS} \quad (4.1)$$

861 QCD SCALE

862 The term “QCD scale” in the context of modeling uncertainties refers to the choice of renormal-
863 ization (μ_R) and factorization (μ_F) scales used in QCD calculations. These are typically treated to-
864 gether. Usually, some multiplicative scale factor, f , is chosen, and each scale is varied in concert with
865 the other scale by $1, f$, and $1/f$ (nine total combinations), sometimes with a cap on how large the
866 combined variation can be (so ignoring the (f, f) and $(1/f, 1/f)$ cases). Just as in the UE+PS, the
867 largest variation is usually taken as the systematic uncertainty.

868 Another important choice in the context of renormalization, to deal with infinities inherent in

[‡]Generally, the maximum is used if it is much larger than other variations, and the average is used if scales are comparable. In general, the historical preference is to be conservative.

869 quantum field theory calculations with loops, is the choice of regularization (to keep track of the
870 infinities) and subtraction schemes (to eliminate them). A common combination is dimensional reg-
871 ularization (deforming the four dimensional integral to $4 - \varepsilon$ dimensions and then taking $\varepsilon \rightarrow 0$)
872 with $\overline{\text{MS}}$ (“MS-bar” (MS for “minimal subtraction”) subtracts off only the infinities plus some com-
873 mon factors that always appear in dimensional regularization). For more details, see, for example,
874 [33]. While each scheme is internally consistent, each can yield different answers at a given order of
875 calculation (and different subtraction schemes may be used for calculations at different orders).

876 PARTON DISTRIBUTION FUNCTIONS AND α_s

877 Finally, separate uncertainties are often made for the choice of parton distribution function (PDF)
878 set and associated choice of strong coupling for $\text{QCD}(\alpha_s)$. Much as in the previous two cases, one
879 can vary the parameter α_s and study what samples of simulation events made using different PDF
880 sets relative some nominal setup look like. Similarly, one can take the maximum, average, or sum in
881 quadrature of different variations to characterize a systematic uncertainty.

882 4.2.2 MODELING SYSTEMATIC TYPES

883 With the concept of what type of effect is taken as a single systematic uncertainty and how its varia-
884 tions are generally evaluated, it is now time to turn to the issue of what exactly is being varied.

885 ACCEPTANCE/NORMALIZATION

886 The most basic type of modeling uncertainty is a normalization uncertainty, often called an accep-
887 tance uncertainty. This simply denotes the uncertainty on the number of predicted events for a
888 given process in a given region of phase space (usually delineated by the number of leptons in the
889 final state and sometimes also by the number of and jets the p_T^V [§] of an event) and is usually expressed
890 as a percent.

891 As an example, the uncertainty on the theoretical prediction of the $H \rightarrow b\bar{b}$ branching ratio,
892 denoted ATLAS_BR_bb (it is an ATLAS-wide systematic), is expressed as a normalization system-
893 atic with a value of 1.7%, affecting all VH processes. Now imagine we have an event in a VH sample
894 with weight 1.0. The nominal histograms for this region gets filled with this event's relevant informa-
895 tion with weight 1.0, while the ATLAS_BR_bb__1up (__1do) histograms get filled with weight 1.017
896 (0.983).

897 SHAPE SYSTEMATICS

898 In addition to normalization systematics expressed as single numbers attached to different processes
899 in different regions, there are also the so-called “shape systematics” and “shape corrections,” heuristi-
900 cally differences in distributions that exist in distributions even after correcting for normalizations,
901 which gives distributions a different “shape” even if their integrals are the same (c.f. Figure 4.3).

[§]This is the transverse mass of the lepton pair for 2-lepton events, the vectorial sum of the single lepton
and \vec{E}_T^{miss} for 1-lepton events, and the \vec{E}_T^{miss} for 0-lepton events.

902 These have the schematic form

$$w_{event} = A_{region} \times f_{region}(event) \quad (4.2)$$

903 where w_{event} is the simulated event's weight, A_{region} is the overall normalization (in principle includ-
904 ing any systematics), and $f_{region}(event)$ is some function of event-level variables, usually a single vari-
905 able, like p_T^V or m_{bb} . The purpose of these systematics is to take into account (in the case of a system-
906 atic) or correct (in the case of a correction applied to the event weight) the non-trivial dependence
907 of a normalization on one of these quantities. Some of these are taken from histograms while others
908 are parametric functions (in this analysis, usually linear ones).

909 An example of the former case is the quantity δ_{EW} , the difference between the nominal $qqVH$
910 cross section and the differential cross section as a function of p_T^V at next to leading order (NLO). As
a correction, this term is simply used as a correction factor $k_{EW}^{NLO} = (1 + \delta_{EW})$.

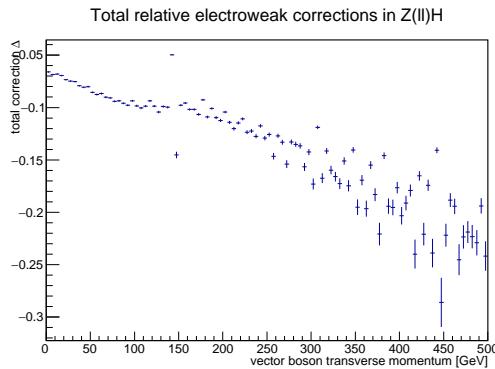


Figure 4.2: The δ_{EW} correction term for 2-lepton $qqZH$.

911

912 An example of the latter case is the systematic associated with the m_{bb} dependence of the the

913 $t\bar{t}$ normalization for 2 jet, $p_T^V \in [75, 150]$ GeV, 2 lepton events. In this case, a variety of effects are
 914 studied (ME, PS, UE), as shown in Figure 4.3. The top half of the plot is the m_{bb} plot for this re-
 915 gion, with the black bars representing the nominal spectrum and spectra generated with different
 916 ME, PS, and UE choices. The ratio plot in the bottom half of the figure shows the scale of varia-
 917 tion normalized to bin content. From this ratio plot, it is clear that the choice of ME (pink points)
 918 was seen to have the largest effect on normalization. The linear fit in the plot reasonably envelopes
 919 this maximum variation was done, and so is taken as the systematic variation. Hence, in this case,
 920 $f_{region}(event)$ is a linear function of m_{bb} , with positive (negative) slope for the up (down) variation.

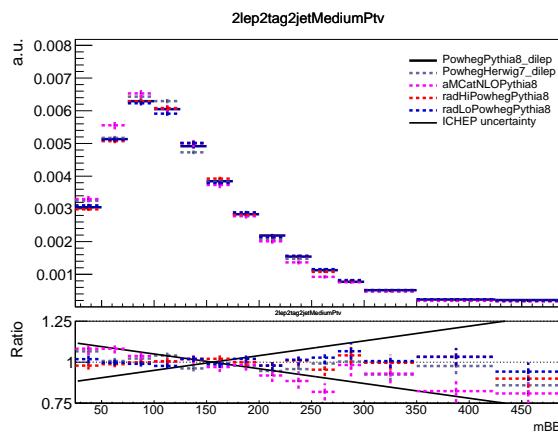


Figure 4.3: The derivation of the 2-lepton $t\bar{t} m_{bb}$ shape systematic.

921 DIVIDING MODELING UNCERTAINTIES: ACCEPTANCE RATIOS

922 In addition to uncertainties on absolute normalizations (both inclusive and region specific), mod-
 923 eling uncertainties are sometimes introduced for the ratio of normalizations between different re-
 924 gions. While these can be simple ratios, evaluating a systematic's effect between regions means eval-

925 uating nominal and alternate choices between regions, so the so-called “double ratio” is often taken
 926 as the scale of variation (plus one). The ATLAS_UEPS_VH_hbb systematic mentioned above, for ex-
 927 ample, has associated with it, ATLAS_UEPS_VH_hbb_32JR. This systematic is evaluated by dividing
 928 the 3 jet to 2 jet ratio in the nominal setup by the same ratio in an alternate setup. These ratios gener-
 929 ically look like:

$$\frac{\text{Acceptance}[\text{Category}_A(\text{nominalMC})]}{\text{Acceptance}[\text{Category}_B(\text{nominalMC})]} \Big/ \frac{\text{Acceptance}[\text{Category}_A(\text{alternativeMC})]}{\text{Acceptance}[\text{Category}_B(\text{alternativeMC})]} \quad (4.3)$$

930 Double ratio systematics are often included in addition to single systematics when a single system-
 931 atic could potentially overestimate the total variation due to a single effect. In the above example,
 932 the choice of UE+PS in signal events may cause different overall variation in 2 jet events as compared
 933 to events with 3 or more jets due to QCD considerations. However, extrapolating between jet mul-
 934 tiplicities can be a non-trivial exercise entirely regardless of the choice of generator for UE+PS (cf.
 935 the discussion of the Stewart-Tackman approach 4.2.2 below), so to include this variation in a single
 936 systematic would lead to overly conservative systematic uncertainties. Hence, using previous knowl-
 937 edge of this separate variation between jet multiplicity regions, one can define a double ratio system-
 938 atic for a more accurate fit model.

939 The three main categories are ratios between different flavor regions, ratios between different lep-
 940 ton channels[¶], and ratios between regions with different numbers of jets, n_{jet} . The first category is
 941 only relevant for $V+jets$ systematics and will be treated in that process’s dedicated section below. As

[¶]e.g. $Z+$ heavy flavor jets (at least one b -jet in the event; often denoted “hf” normalizations in 0- and 2-lepton events

942 this thesis is primarily concerned with the 2-lepton channel, the second category will not be treated
 943 in detail, though the treatment is much the same as other ratio systematics.^{||} In order to discuss the
 944 n_{jet} ratios in systematics (e.g. the ratios in the double ratio example), we must first describe how ex-
 945 clusive n_{jet} cross section calculations are done.

946 THEORETICAL ASIDE: STEWART-TACKMANN A way to calculate uncertainties on processes in re-
 947 gions with different numbers of jets was developed by Stewart and Tackmann and is implicitly used
 948 for most n_{jet} ratio systematics [36]. The problem is how to calculate the cross section and associated
 949 uncertainty for a process with exclusively N jets in the final state. Generically:

$$\sigma_{\geq N} = \sigma_N + \sigma_{\geq N+1} \quad (4.4)$$

950 The physical interpretation of one parton to one jet is an idealized case. In order to demarcate
 951 between jets, one has some quantity that is used as a cutoff in an integral that defines the border be-
 952 tween jet regions.

$$\sigma_{\geq N} = \int_0^{p_{cut}} \frac{d\sigma_N}{dp} + \int_{p_{cut}} \frac{d\sigma_{\geq N+1}}{dp} \quad (4.5)$$

953 Since these cutoffs (not necessarily constant, etc.) can make calculations more complicated, inclu-
 954 sive cross sections tend to be easier to calculate. Hence, it is usually much easier to evaluate the two
 955 inclusive cross sections and find the uncertainties on these by varying α_S in the usual way (cf. Sec-

||Such ratios allow for information in one channel to help constrain other channels, particularly for hard to model processes like $Z+hf$. This helps to reduce final overall uncertainties in combined fits. For a discussion of the interplay of nuisance parameters in combined fits, cf. Chapter 9.

⁹⁵⁶ tion 4.2.1). One then assumes the inclusive uncertainties are uncorrelated, for a covariance matrix for
⁹⁵⁷ $\{\sigma_{\geq N}, \sigma_N, \sigma_{\geq N+1}\}$ of (with Δ_x^2 as the variance associated with x):

$$\Sigma = \begin{pmatrix} \Delta_{\geq N}^2 & \Delta_{\geq N}^2 & 0 \\ \Delta_{\geq N}^2 & \Delta_{\geq N}^2 + \Delta_{\geq N+1}^2 & -\Delta_{\geq N+1}^2 \\ 0 & -\Delta_{\geq N+1}^2 & \Delta_{\geq N+1}^2 \end{pmatrix} \quad (4.6)$$

⁹⁵⁸ These calculations contain Sudakov double logs of $\ln^2(p/Q)$, where Q corresponds to the scale
⁹⁵⁹ of the hard scatter process (m_H), and p_{cut} is usually something like a p_T cutoff. When integrating
⁹⁶⁰ over all of the phase space, these terms can come to dominate calculations when $p \gg Q$. The $N+1$
⁹⁶¹ term in the covariance matrix is an uncertainty associated with the cutoff, but the Sudakov double
⁹⁶² logs will dominate any higher order terms. Stewart and Tackmann give the following reasoning:

⁹⁶³ “In the limit $\alpha_S^2 \approx 1$, the fixed-order perturbative expansion breaks down and the logarithmic
⁹⁶⁴ terms must be resummed to all orders in α_S to obtain a meaningful result. For typical experimental
⁹⁶⁵ values of p_{cut} fixed-order perturbation theory can still be considered, but the logarithms cause large
⁹⁶⁶ corrections at each order and dominate the series. This means varying the scale in α_S in Eq. (9) di-
⁹⁶⁷ rectly tracks the size of the large logarithms and therefore allows one to get some estimate of the size
⁹⁶⁸ of missing higher-order terms caused by p_{cut} , that correspond to Δ_{cut} . Therefore, we can approxi-
⁹⁶⁹ mate $\Delta_{cut} = \Delta_{\geq 1}$, where $\Delta_{\geq 1}$ is obtained from the scale variation for $\sigma_{\geq 1}$.”

⁹⁷⁰ The above considerations are important for this analysis since phase space is separated into 2 and
⁹⁷¹ ≥ 3 jet regions, and the uncertainties for these regions are anti-correlated.

972 4.3 PROCESS SPECIFIC SYSTEMATIC SUMMARIES

973 Brief descriptions of modeling systematics, including recapitulations of nominal sample generation,
974 are given in the following sections. The general approach here is to copy the relevant summary tables
975 and describe any major deviations from the general procedures described in the previous section.
976 The dominant backgrounds for the 2-lepton channel are $Z+hf$ and $t\bar{t}$, accounting for well over 90%
977 of all background events. Diboson samples are the next-leading background and are an important
978 validation sample; others are included for completeness. A summary of all the modeling systematics
in this analysis are given in Table 4.1.

Process	Systematics
Signal	$H \rightarrow b\bar{b}$ decay, QCD scale, PDF+ α_s scale, UE+PS (acc, p_T^V , m_{bb} , 3/2 jet ratio)
$Z+jets$	Acc, flavor composition, $p_T^V+m_{bb}$ shape
$t\bar{t}$	Acc, $p_T^V+m_{bb}$ shape
Diboson	Overall acc, UE+PS (acc, p_T^V , m_{bb} , 3/2 jet ratio), QCD scale (acc (2, 3 jet, jet veto), p_T^V , m_{bb})
Single top	Acc, $p_T^V+m_{bb}$ shape

Table 4.1: Summary of modeling systematic uncertainties, with background samples listed in order of importance.

979

980 4.3.1 SIGNAL PROCESSES

981 Nominal signal $qqVH$ samples are generated using PowHEG with the MINLO (multiscale improved
982 NLO) [37] procedure applied interfaced with PYTHIA8 using the AZNLO tune [24] and NNPDF3.0
983 PDF set [25]. For the 2-lepton case, gluon fusion initiated Higgs production is also considered (ac-
984 counting for $\sim 14\%$ of the total cross section in this channel), with samples generated with PowHEG

985 interfaced with PYTHIA8 using the AZNLO tune. The NNPDF2.3 set [38] is used for both the ME
986 and UE+PS.

987 Alternate samples $qqVH$ samples are generated using MADGRAPH5_aMC@NLO [39] for the
988 ME and PYTHIA8 for the UE+PS, hadronization and MPI. The NNPDF2.3_5f FFN PDF sets and
989 the Al4 tune [40]; the latter has variations included. PowHEG+MinLO+HERWIG7 were samples
990 were also used for systematics.

991 The signal systematics categories are $H \rightarrow bb$ decay cross section, QCD scale, PDF+ α_s scale, and
992 UE+PS. Additionally, there is the NLOEWK correction described above. The correction scale factor
993 is derived using the HAWK MC software. To encapsulate NNLOEW effects the maximum of 1%,
994 the square of the correction factor, and the photon induced cross section is used as a systematic.

995 Table 4.2, reproduced from [21], summarizes the signal cross section systematics, which are ap-
996 plied uniformly across the analysis channels (as applicable).

Sys Name	source	Norm. effect	applied to
ATLAS_BR_bb	$H \rightarrow bb$ dec. unc. (HO effects, m_b , α_s)	1.7%	all VH
ATLAS_QCDscale_VH	QCD scale uncertainty	0.7%	$qq \rightarrow VH$
ATLAS_QCDscale_ggZH	QCD scale uncertainty	27%	$gg \rightarrow ZH$
ATLAS_pdf_Higgs_VH	PDF+ α_s uncertainty	1.9% 1.6%	$qq \rightarrow WH$ $qq \rightarrow ZH$
ATLAS_pdf_Higgs_ggZH	PDF+ α_s uncertainty	5.0%	$gg \rightarrow ZH$

Table 4.2: Summary of all systematic uncertainties on the VH cross section including their value, source and the corresponding nuisance parameter name.

997 The remaining signal systematics are analysis channel specific and are summarized in Table 4.3.

998 The methodologies match those described in Section 4.2. The UE+PS systematics were derived us-
999 ing the alternate samples mentioned above; QCD scale uncertainties were derived by varying scales

1000 by 1/3 and 3; and PDF uncertainties were derived by comparing the nominal set with the PDF4-
 1001 LHC15_30 PDF set [41].

NP name	oL		iL		zL	
	2j	3j	2j	3j	2j	$\geq 3j$
ATLAS_UEPS_VH_hbb	10.0%	10.0%	12.1%	12.1%	13.9%	13.9%
ATLAS_UEPS_VH_hbb_32JR	–	13.0%	–	12.9%	–	13.4%
ATLAS_UEPS_VH_hbb_VPT	shape only		shape+norm			
ATLAS_UEPS_VH_hbb_MBB	shape only					
QCDscale_VH_ANA_hbb_J2	6.9%	–	8.8%	–	3.3%	–
QCDscale_VH_ANA_hbb_J3	-7%	+5%	-8.6%	+6.8%	-3.2%	+3.9%
QCDscale_VH_ANA_hbb_JVeto	–	-2.5%	–	3.8%	–	–
QCDscale_VH_ANA_hbb_VPT	shape only		shape+norm			
QCDscale_VH_ANA_hbb_MBB	shape only					
pdf_HIGGS_VH_ANA_hbb	1.1%	1.1%	1.3%	1.3%	0.5%	0.5%
pdf_VH_ANA_hbb_VPT	shape only		shape+norm			
pdf_VH_ANA_hbb_MBB	shape only					

Table 4.3: Summary of all systematic uncertainties on the VH acceptance and shapes originating from altering the PDF and α_S uncertainties, including their corresponding nuisance parameter name.

1002 4.3.2 $V+$ JETS

1003 Nominal $V+$ jets samples are generated using SHERPA 2.2.1@NLO** [42] for both the ME and PS,
 1004 interfaced with the NNPDF’s and using a five quark flavor scheme, and alternative samples are de-
 1005 rived using MADGRAPH5 interfaced with PYTHIA8. In order to increase statistics in important re-
 1006 gions of phase space, these samples were separated into kinematic slices based on p_T^V and into bins of
 1007 jet flavor. The kinematic slices were in the quantity $\max(H_T, P_T^V)$ and had the intervals [0–70, 70–
 1008 140, 140–280, 280–500, 500–1000, > 1000] GeV. The jet flavor slices were made using flavor vetoes
 1009 and filters:

**SHERPA 2.1 is used for some variations not available in SHERPA 2.2.1.

- BFilter: at least 1 b-hadron with $|\eta| < 4, p_T > 0$ GeV
- CFilterBVeto: at least 1 c-hadron with $|\eta| < 3, p_T > 4$ GeV; veto events which pass the BFilter
- CVetoBVeto: veto events which pass the BFilter and/or the CFilterBVeto

These in turn are related to the main flavor regions used in the analysis, based on the flavor of the two leading jets in an event (based on p_T). These five flavors (with up, down, and strange collectively known as “light”) yield six different flavor combinations: bb, bc, bl (these first three collectively known as “heavy flavor” or $V+hf$), cc, cl, ll (or just “light” or l). Ratio systematics are often made with respect to the acceptance in the bb region.

$V+jet$ systematics are derived in several steps. The first is to use double ratios of acceptances between analysis regions and nominal versus alternative MC’s (so $(\text{Region1-nominal}/\text{Region2-nominal}) / (\text{Region1-alternate}/\text{Region2-alternate})$). The main region comparisons are 2 jet versus 3 jet ($3+jet$ for 2-lepton) and then 0-lepton versus 2-lepton (1-lepton) for $Z+hf$ ($W+hf^{\dagger\dagger}$). The final uncertainty contains the sum in quadrature of four effects:

1. Variation of 0.5 and 2 of QCD scales in the SHERPA sample
2. Sum in quadrature of half the variation from different resummation and CKKW merging scales ^{††}
3. Maximal variation between nominal setup and SHERPA 2.2.1 with the MMHT2014nnlo68cl and CT14nnlo PDF sets
4. Difference between the SHERPA and MADGRAPH₅ sets

^{††}The $W+hf$ CR versus the SR is also considered for $W+hf$

^{‡‡}cf. [43], Section 2 for a summary of the CKKW method for different parton multiplicities used in SHERPA

1030 Summaries of the Z +jets uncertainties are provided here; the reader is referred to [21] for the
 1031 W +jets systematics, as these events are virtually non-existent in the 2-lepton case with which this
 1032 thesis is almost exclusively concerned. In Table 4.4, from [21] are the normalization systematics.

Process	Name	prior in region					
		2jet			(\geq)3jets		
		$_{\text{2L: lo}}$	$_{\text{2L: hi}}$	$_{\text{oL}}$	$_{\text{2L: lo}}$	$_{\text{2L hi}}$	$_{\text{oL}}$
$Z+l$	SysZclNorm	18%					
$Z+cl$	SysZlNorm	23%					
$Z+hf$	norm_Zbb	Floating Normalization					
$Z+hf$	SysZbbNorm_L2_J3	-	-	-	30%	30%	-
$Z+hf$	SysZbbNorm_J3	-	-	-	-	-	17%
$Z+hf$	SysZbbNorm_0L	-	-	7%	-	-	7%
$Z+hf$	SysZbbPTV	effect on each region obtained from shape rw					

Table 4.4: Effect of modeling systematics on Z +jets normalization in the 2lepton regions. For systematic uncertainties implemented with a prior the effect of $1-\sigma$ variation is reported. The uncertainties labelled as Zbb act on the entire $Z+hf$ background. Region labels ``lo'' and ``hi'' refer to event p_T^V .

1033 The flavor composition ratio systematics are in Table 4.5, also from [21].
 1034 Finally, the p_T^V and m_{bb} shape systematics are derived using control regions in data. The func-

1035 tional form for the p_T^V systematic is $\pm 0.2 \log 10(p_T^V/50\text{GeV})$, and that of the m_{bb} systematic is $\pm 0.0005 \times$
 1036 $(m_{jj} - 100\text{ GeV})$.

1037 4.3.3 TOP-PAIR PRODUCTION

1038 Nominal $t\bar{t}$ samples are produced with POWHEG at NLO for the ME calculation using the NNPDF3.0
 1039 PDF set interfaced with PYTHIA8.210 using the A14 tune and the NNPDF2.3 PDF set at LO. The

These use the same selections as the signal regions except for b -tags (0, 1, and 2 tags are studied), with the added requirement in 2tag regions that m_{bb} not be in the range of 110–140GeV.

Category	Nuisance Parameter Name	Prior	Applied to
$Z+bc/Z+bb$	SysZbcZbbRatio	40%	$Z+bc$ events (0-Lepton)
		40%	$Z+bc$ events (2-Lepton 2jet)
		30%	$Z+bc$ events (2-Lepton ≥ 3 jet)
$Z+bl/Z+bb$	SysZblZbbRatio	25%	$Z+bl$ events (0-Lepton)
		28%	$Z+bl$ events (2-Lepton 2jet)
		20%	$Z+bl$ events (2-Lepton ≥ 3 jet)
$Z+cc/Z+bb$	SysZccZbbRatio	15%	$Z+cc$ events (0-Lepton)
		16%	$Z+cc$ events (2-Lepton 2jet)
		13%	$Z+cc$ events (2-Lepton ≥ 3 jet)

Table 4.5: The priors on the relative acceptance variations for $Z+hf$. The first column details the flavor components across which the acceptance variation is being considered, the second column lists the names of the corresponding nuisance parameter in the Profile Likelihood Fit, the third contains the value of the prior and the fourth column the processes and categories to which this nuisance parameter is applied.

parameters `hdamp` (nominal value $1.5m_{top}$, a resummation damping factor for ME/PS matching that can heuristically thought of as tuning high p_T radiation) in PowHEG and `pThard` (nominal value 0) and `pTdef` (nominal value 2) in PyTHIA (both control merging with PowHEG) are varied to evaluate certain systematics. Alternative $t\bar{t}$ samples use PowHEG+HERWIG7, MADGRAPH5_aMC-@NLO+PyTHIA8.2, and the nominal setup with varied tunes and parameter values. Uncertainties are taken to cover the largest difference between the nominal and any of these alternate configurations.

The overall $t\bar{t}$ normalization is a floating normalization, and further systematics attached to the ratio of acceptances between regions (3-to-2 jet, SR-to-WhfCR, and 1-to-0 lepton) are defined using double ratios; these are summarized in Tables 4.6 and 4.7, taken from [21].

Shape systematics for p_T^V and m_{bb} are linear and taken to cover the largest difference reasonably. The use of a top $e - \mu$ control region helps constrain this.

	0-lepton		1-lepton			
Systematic	2j	3j	WCR 2j	SR 2j	WCR 3j	SR 3j
norm_ttbar	floating normalization					
SysttbarNorm_L0	8%	8%	-	-	-	-
SysttbarNorm_J2	9%	-	9%	9%	-	-
SysttbarNorm_DWhfCR_L1	-	-	25%	-	25%	-

Table 4.6: Effect of modeling systematics on $t\bar{t}$ normalization in the 0 and 1-lepton analysis region.

	2jet		≥ 3 jets	
	lo	hi	lo	hi
norm_ttbar_J2_L2	floating normalization	-	-	-
norm_ttbar_J3_L2	-	-	floating normalization	-
SystTbarPTV_L2_L2	effect on each region obtained from shape rw			

Table 4.7: Effect of modeling systematics on $t\bar{t}$ normalization in the 2lepton regions. The SystTbarPTV_L2_L2 systematic is implemented as a shape systematic over the full $VpT > 75$ GeV range, and as a result has different acceptance effects in the low and high VpT regions. Systematics are treated the same in both signal and control regions, and “lo” and “hi” refer to the p_T^V split in events.

¹⁰⁵¹ well, as described above in 4.2.2. These are summarized in Table 4.8, again taken from [21].

Analysis region	Uncertainty	Value	Source	NP
o,1 lepton	p_T^V shape	shape	fit through largest deviation (aMC@NLO + PYTHIA8)	TTbarPTV
2 lepton	p_T^V shape	norm	fit through largest + shape	deviation (aMC@NLO + PYTHIA8)
TTbarPTV_L2				
o,1 lepton	$m_{b\bar{b}}$ shape	shape	fit through largest only	deviation (aMC@NLO + PYTHIA8)
TTbarMBB				
2 lepton	$m_{b\bar{b}}$ shape	shape	fit through largest only	deviation (aMC@NLO + PYTHIA8)
TTbarMBB_L2				

Table 4.8: Summary of all shape uncertainties for the $t\bar{t}$ process with short descriptions and the name of the corresponding nuisance parameters.

¹⁰⁵² 4.3.4 DIBOSON PRODUCTION

¹⁰⁵³ Three diboson production processes (collectively denoted VV) are important for these analyses: ZZ ,

¹⁰⁵⁴ WZ , and WW . Nominal samples are created using SHERPA 2.2.1 using the NNPDF3.0 PDF set. Al-

¹⁰⁵⁵ ternative samples use Powheg+PYTHIA8 and Powheg+HERWIG++. The methodology here is

¹⁰⁵⁶ similar to that of the $t\bar{t}$ systematics, with both overall acceptance and lepton channel specific uncer-

¹⁰⁵⁷ tainties, with the exception that UE+PS and QCD scale are treated separately (PDF+ α_S was found

¹⁰⁵⁸ to be negligible). p_T^V shape systematics are described using linear fits, while $m_{b\bar{b}}$ shape systematics

¹⁰⁵⁹ are described using hyperbolic tangents (third degree polynomials) in the 2 jet (3 jet) regions. Once

¹⁰⁶⁰ again, summary tables from [21] are reproduced here.

Sys Name	source	Norm. effect	applied to
SysWWNorm	overall cross section uncertainty	25%	WW in all regions
SysWZNorm	overall cross section uncertainty	26%	WZ in all regions
SysZZNorm	overall cross section uncertainty	20%	ZZ in all regions

Table 4.9: Summary of all systematic uncertainties on the diboson cross section including their value, source and the corresponding nuisance parameter name.

NP name	oL: $ZZ \rightarrow \nu\bar{\nu} b\bar{b}$		1L: $WZ \rightarrow \ell\nu b\bar{b}$		2L: $ZZ \rightarrow \ell^+ \ell^- b\bar{b}$	
	2j	3j	2j	3j	2j	$\geq 3j$
SysVZ_UEPS_Acc	5.6%	5.6%	3.9%	3.9%	5.8%	5.8%
SysVZ_UEPS_32JR	–	7.3%	–	10.8%	–	3.1%
SysVZ_UEPS_VPT	shape+norm			shape only		shape+norm
SysVZ_UEPS_MBB	shape only					
SysVZ_QCDscale_J2	10.3%	–	12.7%	–	11.9%	–
SysVZ_QCDscale_J3	-15.2%	+17.4%	-17.7%	+21.2%	-16.4%	+10.1%
SysVZ_QCDscale_JVeto	–	+18.2%	–	+19.0%	–	–
SysVZ_QCDscale_VPT	shape+norm			shape only		shape+norm
SysVZ_QCDscale_MBB	shape only					

Table 4.10: Summary of the systematic uncertainties on the VH acceptance in each analysis region and on the p_T^V and $m_{b\bar{b}}$ shapes originating from altering the QCD scale, including their nuisance parameter name.

4.3.5 SINGLE TOP PRODUCTION

- 1061 Single top sample are generated separately for the different production channels (s , t , and Wt) using
- 1062 **POWHEG** with the CT10 NLO PDF's interfaced with **PYTHIA6** using the PERUGIA2012 PS tune
- 1063 and the corresponding CTEQ6L1 LO PDF's and PHOTOS (TAUOLA) for QED final state (τ) de-
- 1064 cays.

Production	Uncertainty	Value	Source	Nuisance Parameter
<i>s</i> -channel	overall normalization	4.6%	sum in quadrature of μ_R , μ_F , α_S and PDF uncertainties	<code>stopNorm</code>
<i>t</i> -channel	overall normalization	4.4%	sum in quadrature of μ_R , μ_F , α_S and PDF uncertainties	<code>stoptNorm</code> correlated with 2 jet and 3 jet case
<i>t</i> -channel	2 jet region acceptance of deviations in alternative generators	17%	sum in quadrature	<code>stoptAcc</code> correlated with overall and 3 jet case
<i>t</i> -channel	3 jet region acceptance	20%	sum in quadrature of deviations in alternative generators	<code>stoptAcc</code> correlated with overall and 2 jet case
<i>Wt</i> channel	overall normalization	6.2%	sum in quadrature of μ_R , μ_F , α_S and PDF uncertainties	<code>stopWtNorm</code> correlated with 2 jet and 3 jet case
<i>Wt</i> channel	2 jet region normalization	35%	sum in quadrature of deviations in alternative generators	<code>stopWtAcc</code> correlated with overall and 3 jet case
<i>Wt</i> channel	3 jet region normalization	41%	sum in quadrature of deviations in alternative generators	<code>stopWtAcc</code> correlated with overall and 2 jet case
<i>t</i> -channel	p_T^V shape	shape	fit through largest deviation (POWHEG+ HERWIG++) $\pm 0.001 \times p_T^V \mp 0.17 + 1$	<code>StopPTV</code>
<i>t</i> -channel	$m_{b\bar{b}}$ shape	shape	fit through largest deviation (POWHEG+ PYTHIA6 radHi-radLo) $\pm 0.0008 \times m_{b\bar{b}} \mp 0.12 + 1$	<code>StopMBB</code>
<i>Wt</i> channel	p_T^V shape	shape	fit through largest deviation (POWHEG+ PYTHIA6 with diagram subtraction) $\pm 0.003 \times p_T^V \mp 0.69 + 1$	<code>StopWtPTV</code>
<i>Wt</i> channel	$m_{b\bar{b}}$ shape	shape	fit through largest deviation (POWHEG+ PYTHIA6 with diagram subtraction) $\pm 0.0036 \times m_{b\bar{b}} \mp 0.52 + 1$ ($m_{b\bar{b}} < 275$ GeV) $\mp 0.47 + 1$ ($m_{b\bar{b}} \geq 275$ GeV)	<code>StopWtMBB</code>

Table 4.11: Summary of all uncertainties for the single top process with short descriptions and the name of the corresponding nuisance parameters, updated for the winter baseline analysis.

子路曰：衛君待子而為政，子將奚？

子曰：必也正名乎！

Confucius, *The Analects*

5

1066

1067

Object and Event Reconstruction and

1068

Selection

1069 IN BREAKING WITH THE STANDARD CONVENTION both object definitions and their associated

1070 experimental systematic uncertainties will be defined in this chapter: the hope is that the proximity

1071 of these descriptions will allow them to elucidate each other. Summary tables are almost exclusively
1072 taken from [44] or [4]. This analysis, like most analyses in ATLAS, use central object definitions
1073 from collaboration combined performance (CP) groups^{*} using standard analysis tools and recom-
1074 mendations from these groups for the various objects and their accompanying systematic uncertain-
1075 ties.

1076 Before proceeding to the objects used in this analysis, we begin with a few remarks on uncertain-
1077 ties associated with object reconstruction. Event-level variables and selections are discussed more in
1078 depth in Chapters 1 and 6. As described in Section 4.2, systematics quantify the uncertainty asso-
1079 ciated with certain effects, and are generally treated in an analysis by saving histograms of discrimi-
1080 nating distributions corresponding to the nominal analysis except with the systematic in question
1081 varied by plus and minus one standard deviation each (one histogram each). While for modeling
1082 systematics this only corresponds to different event weights, for experimental systematics like those
1083 described in this chapter (with the exception of flavor tagging and certain trigger systematics), this is
1084 done by varying the parameter in question and re-running reconstruction with the systematic varied
1085 before recomputing all event level quantities and then saving discriminant values in their appropri-
1086 ate distributions. This is, in general, a much more computationally intensive process in the analysis,
1087 which is why an entire software framework, the `CxAODFramework`, was created for this analysis (see
1088 Section 3 of [44] for more details).

*Teams of physicists within ATLAS dedicated to studying different aspects of reconstruction general to very many analyses. An example is the ATLAS Muon CP group, whose duties include providing definitions for how muons are reconstructed in ATLAS, different quality requirements and thresholds, and muon related systematic uncertainties, as well as software packages to make these definitions and guidelines easier to implement.

1089 5.1 TRIGGERS

1090 Tables of the triggers used with the 2015 and 2016 datasets are given in Tables 5.1 and 5.2.

o lep	1 lep	2 lep
HLT_xe70	HLT_xe70 HLT_e24_lhmedium_L1EM20VH OR HLT_e60_lhmedium OR HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 OR HLT_mu40 HLT_e24_lhmedium_L1EM20VH OR HLT_e60_lhmedium OR HLT_e120_lhloose

Table 5.1: Summary table of triggers used in 2015 Data.

1091 The o-lepton channel uses a \vec{E}_T^{miss} trigger, while the 2-lepton channel uses single lepton triggers,
1092 with the 1-lepton analysis using both. Since the o- and 1-lepton channels are largely beyond the scope
1093 of this thesis, the discussion here will be limited to the single lepton triggers; the interested reader is
1094 directed towards [44] and its cited sources for an in-depth discussion of the use of the \vec{E}_T^{miss} trigger.

1095 The efficiency of triggers is in general different on simulated datasets than in actual data collected
1096 in ATLAS, so a scale factor to correct for this difference in efficiency must be applied to simulation
1097 events. This scale factor is given by the muon CP group for muons for both the 1- and 2-lepton cases
1098 and from the electron CP group for the 1-lepton case. For the two electron case, this was calculated
1099 by the analysis team as (details in [44]):

$$\frac{1 - (1 - \epsilon_{\text{MC}}^{e1} \times \text{SF}^{e1}) \times (1 - \epsilon_{\text{MC}}^{e2} \times \text{SF}^{e2})}{1 - (1 - \epsilon_{\text{MC}}^{e1}) \times (1 - \epsilon_{\text{MC}}^{e2})} \quad (5.1)$$

1100 There are also systematic uncertainties associated with these trigger efficiencies. The single elec-
1101 tron trigger efficiency systematic uncertainty is encapsulated in a single systematic, `EL_EFF_Trigger_-`

period	o lep	1 lep	2 lep
A	HLT_xe90_mht_L1XE50	HLT_xe90_mht_-L1XE50, HLT_e26_-lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0	HLT_mu24_-ilosse(data), HLT_-mu24_ilosse_-L1MU15(MC) OR HLT_mu40, HLT_e26_-lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0
B-D ₃	HLT_xe90_mht_L1XE50	HLT_xe90_mht_-L1XE50, HLT_e26_-lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0	HLT_mu24_ivarmedium OR HLT_mu50, HLT_-e26_lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0
D ₄ -E ₃	HLT_xe110_mht_-L1XE50	HLT_xe110_mht_-L1XE50, HLT_e26_-lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0	HLT_mu24_ivarmedium OR HLT_mu50, HLT_-e26_lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0
$\geq F_1$	HLT_xe110_mht_-L1XE50	HLT_xe110_mht_-L1XE50, HLT_e26_-lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0	HLT_mu26_ivarmedium OR HLT_mu50, HLT_-e26_lhtight_nod0_-ivarloose OR HLT_-e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_-lhloose_nod0

Table 5.2: Summary table of triggers used in 2016 Data.

1102 **Total_1NPCOR_PLUS_UNCOR**, while the single muon trigger efficiency has two components, one
1103 each for the sample statistics, **MUON_EFF_TrigStatUncertainty**, and systematic uncertainties
1104 **MUON_EFF_TrigSystUncertainty** associated with that efficiency’s measurement.

1105 While the momentum associated with the lowest un-prescaled single lepton triggers changes de-
1106 pending on data-taking conditions (the numbers associated with the triggers in the tables can be
1107 thought of as nominal p_T values for trigger level objects), the lowest typical value is ~ 25 GeV. In
1108 order to maintain this triggering capability on low p_T muons in the higher luminosity environment
1109 of the Run 3 LHC and beyond, trigger-capable detectors will be installed in upgraded New Small
1110 Wheels (NSW) of the ATLAS muon detector during the Phase I upgrade. Detailed studies in sim-
1111 ulation of the trigger algorithm performance under nominal and misaligned conditions for the Mi-
1112 cromegas detectors to be installed in the NSW may be found in Appendix A.

1113 **5.2 ELECTRONS**

1114 Electrons in ATLAS are reconstructed using a combination of the ATLAS electromagnetic calorime-
1115 ter (ECAL) and inner detector (ID). Reconstruction begins by searching for so-called “seed clusters”
1116 in the ECAL. The ECAL is divided into a 200×256 tower grid in the $\eta - \phi$ plane, with each tower
1117 having a size of 0.025 square in η and ϕ , corresponding to the granularity of the ECAL in its middle
1118 layer, with all energy in a tower summed longitudinally. A “sliding window” of 3×5 cells in the
1119 $\eta - \phi$ plane is then used to identify EM clusters associated with electrons based on criteria detailed
1120 in [45]. This comparatively simple algorithm (in contrast to jet reconstruction detailed below) is
1121 effective since electromagnetic showers have a well defined behavior and shape.

Once seed clusters have been formed, they are associated with tracks in the inner detector. Combined cluster-tracks pairs form electron candidates. In order for a electron candidate to be considered a suitable electron for analysis, it must pass certain quality requirements, based on a cut on the value of a likelihood-based (LH) discriminant (cf. [46] for details). This discriminant is given by:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \quad \mathcal{L}_S(\vec{x}) = \prod_{i=1}^n P_{s,i}(x_i) \quad (5.2)$$

where the s and S (b and B) subscripts refer to distributions in fiducial signal (background) distributions in bins of $|\eta|$ and E_T . The $P(x_i)$ are probability distributions functions (pdf)'s for input variables. Several sets of input variables exist for increasingly stringent quality requirements on electrons; this analysis uses Loose LH electrons as the base for electron selection, with the input variables relating to leakage into the hadronic calorimeter (HCAL), shower and energy deposits in each of the ECAL layers, track quality requirements, TRT hits, and track-cluster matching. This analysis adds a LooseTrackOnly isolation requirement (the p_T sum of tracks within a certain $\eta - \phi$ distance of the candidate track must be below a certain value), impact parameter significance cuts, and an explicit B-layer hit requirement. The ZH -signal electrons must further pass a $27 \text{ GeV } p_T$ cut ($1.05 \times p_T^{\text{trigger}}$). These requirements are summarized in Table 5.3.

e Selection	p_T	η	ID	d_0^{sig}	$ \Delta z_0^{\text{BL}} \sin \theta $	Isolation
$VH - \text{loose}$	$> 7 \text{ GeV}$	$ \eta < 2.47$	LH Loose + B-layer cut	< 5	$< 0.5 \text{ mm}$	LooseTrackOnly
$ZH - \text{signal}$	$> 27 \text{ GeV}$	$ \eta < 2.47$	LH Loose + B-layer cut	< 5	$< 0.5 \text{ mm}$	LooseTrackOnly
$WH - \text{signal}$	$> 27 \text{ GeV}$	$ \eta < 2.47$	LH Tight	< 5	$< 0.5 \text{ mm}$	FixedCutHighPtCaloOnly

Table 5.3: Electron selection requirements.

1136 5.2.1 ELECTRON SYSTEMATICS

1137 The electron CP group has tabulated standard systematic uncertainties to be associated with the use
1138 of reconstructed electrons in ATLAS analyses in two main categories. The first category is related
1139 to efficiency corrections and is broken into three components: identification (`EL_EFF_ID_Total-`
1140 `CorrUncertainty`), reconstruction (`EL_EFF_Reco_TotalCorrUncertainty`), and isolation
1141 (`EL_EFF_Iso_TotalCorrUncertainty`). The second category deals with electron energy scale
1142 (roughly, the uncertainty in taking the energy deposits in an EM cluster and turning them into an
1143 electron energy) and energy resolution (the width associated with this). This is in practice a very
1144 complicated procedure, with over 60 systematics associated, but this analysis is not at all sensitive
1145 to these effects and so a simplified model of two systematics, `EG_RESOLUTION_ALL` and `EG_SCALE-`
1146 `_ALL`, is used.

1147 5.3 MUONS

1148 This analysis uses the standard CP muon collection in an event, though these muons in ATLAS are
1149 constructed in a variety of ways; for full details see [47] and [48]. Most muons are constructed us-
1150 ing tracks in the chambers of the muon spectrometer (MS), with a variety of algorithms available.
1151 MS tracks are sufficient to reconstruct a muon (a fit on these tracks can be used to point back to an
1152 interaction point for vertex matching, for example) and, in the $|\eta| \in (2.5, 2.7)$ interval where there
1153 is no tracking, these standalone (SA) muons are the default. The most common and robust form
1154 of muon reconstruction combines tracks in the MS with tracks in the ID (more precisely, a global

refit with hits from both subsystems is typically done) to form combined (CB) muons. CB and SA
 muons automatically pass the loose reconstruction requirements for the Loose muons used in this
 analysis. Additionally, since there is a gap in the $|\eta| < 0.1$ range in the MS to make room for cabling
 and other detector services, there are two further muon types used in this range: the segment tagged
 (ST) muons that match ID tracks to segments in the MDT or CSC chambers and the calorimeter
 tagged (CT) muons that match ID tracks to calorimeter clusters consistent with minimum ionizing
 particles (which muons in ATLAS generally are).

Further quality requirements are imposed on Loose muons for the different muon categories
 used in this analysis. Isolation requirements similar to the electrons in corresponding categories are
 imposed, and impact parameter requirements are also imposed. The ZH signal muons also have a
 p_T cut at 27 GeV and a requirement that the muon fall within the $|\eta|$ range of the ID.

μ Selection	p_T	η	ID	d_0^{sig}	$ \Delta z_0^{\text{BL}} \sin \theta $	Isolation
$VH - \text{loose}$	$> 7 \text{ GeV}$	$ \eta < 2.7$	Loose quality	< 3	$< 0.5 \text{ mm}$	LooseTrackOnly
$ZH - \text{signal}$	$> 27 \text{ GeV}$	$ \eta < 2.5$	Loose quality	< 3	$< 0.5 \text{ mm}$	LooseTrackOnly
$WH - \text{signal}$	$> 25 \text{ GeV}$	$ \eta < 2.5$	Medium quality	< 3	$< 0.5 \text{ mm}$	FixedCutHighPtTrackOnly

Table 5.4: Muon selection requirements.

5.3.1 MUON SYSTEMATICS

Similar to the treatment of systematic uncertainties associated with the electrons, muons have CP de-
 fined systematics. The muon momentum scale and resolution systematics are divided into three cat-
 egories associated one for uncertainties related to ID tracks (`MUONS_ID`), one for MS tracks (`MUONS-
_MS`), one for the overall scale (`MUONS_SCALE`), and two for charge dependent momentum scales

1171 (`MUON_SAGITTA_RHO` and `MUON_SAGITTA_RESBIAS`). The remaining systematics have a `STAT` and
1172 `SYS` component corresponding to the sample statistics and systematic uncertainties for their individ-
1173 ual components. Efficiency scale factors use different standard candles in different p_T ranges (J/ψ 's
1174 (Z 's) below (above) 15 GeV), and so these systematics are broken up into two categories (`MUON_EFF-`
1175 `_STAT` and `MUON_EFF_SYS`; `MUON_EFF_STAT_LOWPT` and `MUON_EFF_SYS_LOWPT`). There are also
1176 isolation systematics (`MUON_ISO_STAT`, `MUON_ISO_SYS`) and track to vertex association systematics
1177 (`MUON_TTVA_STAT`, `MUON_TTVA_SYS`).

1178 **5.4 MISSING TRANSVERSE ENERGY**

1179 High precision performance of \vec{E}_T^{miss} is not so crucial to the 2-lepton analysis (though it is very im-
1180 portant to the other channels), so the interested reader is referred to [49]. \vec{E}_T^{miss} in ATLAS is the neg-
1181 ative vectoral sum of physics objects (in this analysis just jets and leptons, though in principle also
1182 including τ 's and γ 's) and a so-called track based soft term (TST). The TST is comprised of valid
1183 ID tracks not associated with any physics objects in an event. These tracks must be associated to an
1184 event's primary vertex, have a $p_T > 0.4$ GeV, and pass other quality requirements.

1185 The \vec{E}_T^{miss} systematic uncertainties relevant to this analysis are related to track based energy scale
1186 and resolutions in both the soft term and in the jets and are: `MET_SoftTrk_ResoPara`, `MET_Soft-`
1187 `Trk_ResoPerp`, `MET_SoftTrk_ScaleDown`, `MET_SoftTrk_ScaleUp`, `MET_JetTrk_Scale-`
1188 `Down`, and `MET_JetTrk_ScaleUp`.

1189 5.5 JETS

1190 Unlike leptons, all analyses considered in this thesis are sensitive to factors regarding jet reconstruc-
1191 tion and associated systematic uncertainties. A general discussion of jets precedes jet reconstruction
1192 in ATLAS and associated systematics relevant to this thesis.

1193 5.5.1 JET ALGORITHMS

1194 The hadronic nature of jets makes jet reconstruction a lot more complicated than electron or photon
1195 reconstruction, where a regular shower shape and track matching (or lack thereof in the case of the
1196 chargeless photon) provide a fairly straightforward and robust approach. The interested reader is
1197 referred to [50] for an excellent survey, from which this discussion is greatly abbreviated.

1198 Looking at an event like the one in Figure 5.1, unambiguous individual jets are particularly easy to
1199 identify, more or less popping out of the $\eta - \phi$ plane plot, but this is not always the case.

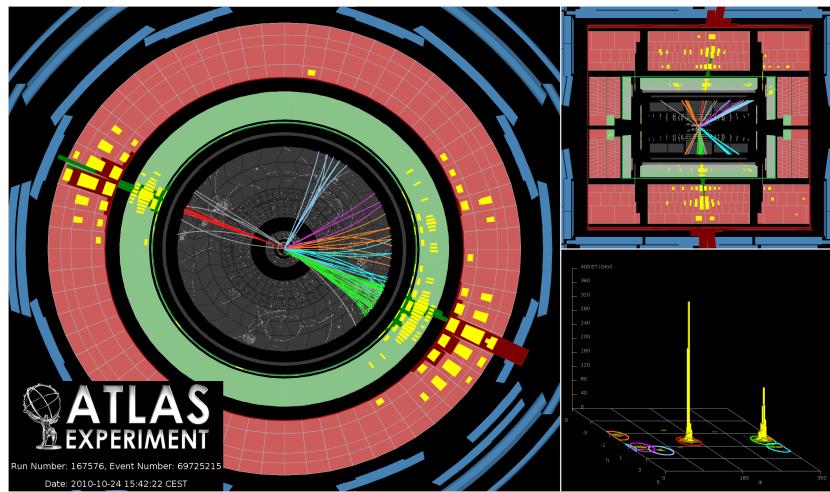


Figure 5.1: A clean ATLAS dijet event.

1200 Two general methods of turning particles/calorimeter towers into jets exist: cone-based and se-
 1201 quential recombination. The general theme of the former is to find a hard (energetic) particle and
 1202 draw a circle around it in the $\eta - \phi$ plane in an intelligent manner, while the theme of the latter
 1203 is to find some metric of distance between particles and then to cluster pairs based on this distance
 1204 into jets in an intelligent way. Cone-based algorithms are simple (and therefore generally quite fast)
 1205 but generally lack some properties of the sequentially recombined jets (though there are notable ex-
 1206 ceptions like SISCone). Cone algorithm reconstructed jets are important for trigger level objects in
 1207 ATLAS, though since no jet triggers are used in this analysis, they will not be discussed any further
 1208 here.

1209 The general drawback of cone-based algorithms is that they are not infrared and collinear (IRC)
 1210 safe. That is, neither the emission of a soft (IR) quark or gluon during hadronization nor the collinear
 1211 splitting of hard particles during hadronization should not change the final jet collection in an event.
 1212 These are fairly common edge cases and can lead to certain pathologies in QCD calculations. In-
 1213frared and collinear safety are diagrammed schematically in Figures 5.2 (a) and (b), taken from [50].

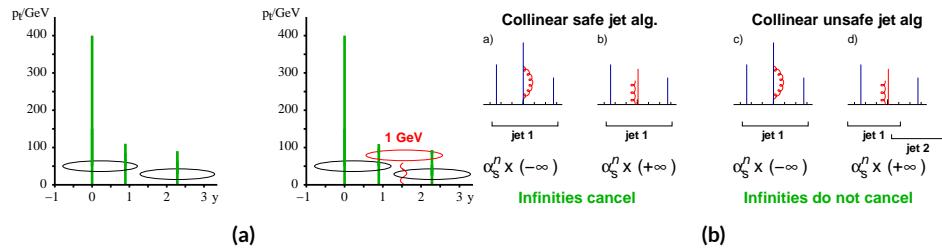


Figure 5.2: Infrared (a) and collinear (b) safety.

1214 Sequential recombination algorithms are generally safe from these effects, as these edge cases are
 1215 very “close” to each other by construction. A sequential recombination algorithm proceeds as fol-

1216 lows

- 1217 i. Evaluate the set of distances d_{ij} (for pairs of objects) and d_{iB} (the “beam distance” for each
1218 individual object)

$$d_{ij} = \min \left(p_{Ti}^{2p}, p_{Tj}^{2p} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = p_{Ti}^{2p} \quad (5.3)$$

- 1219 2. Find the minimum distance

- 1220 3. If the minimum distance is:

- 1221 • A d_{ij} : cluster these objects together, and go to step 1
- 1222 • A d_{iB} : call the i^{th} object a jet, remove it from the set of objects to be clustered, and go to
1223 step 1

- 1224 4. Repeat until all objects are clustered into jets

1225 The choices one must make in sequential recombination are the size parameter R , akin to a cone
1226 radius in cone-based algorithms, and the momentum power p . Common choices and their trade-offs
1227 are:

- 1228 • +i: the k_t algorithm; favors the softer particles in an event, so the cluster sequence gives a history
1229 of hadronization, but jet shapes are irregular (i.e. not circular in the $\eta - \phi$ plane)
- 1230 • o: the Cambridge-Aachen algorithm: a pure distance metric; less substructure but jets tend to
1231 be more circular
- 1232 • -i: the anti- k_t algorithm: clustering begins with hardest particles in an event; regular, localized
1233 jet shapes, but virtually no substructure in clustering history

1234 Jet reconstruction using all three algorithms on the same event, as well as SISCone, are shown in
1235 Figure 5.3.

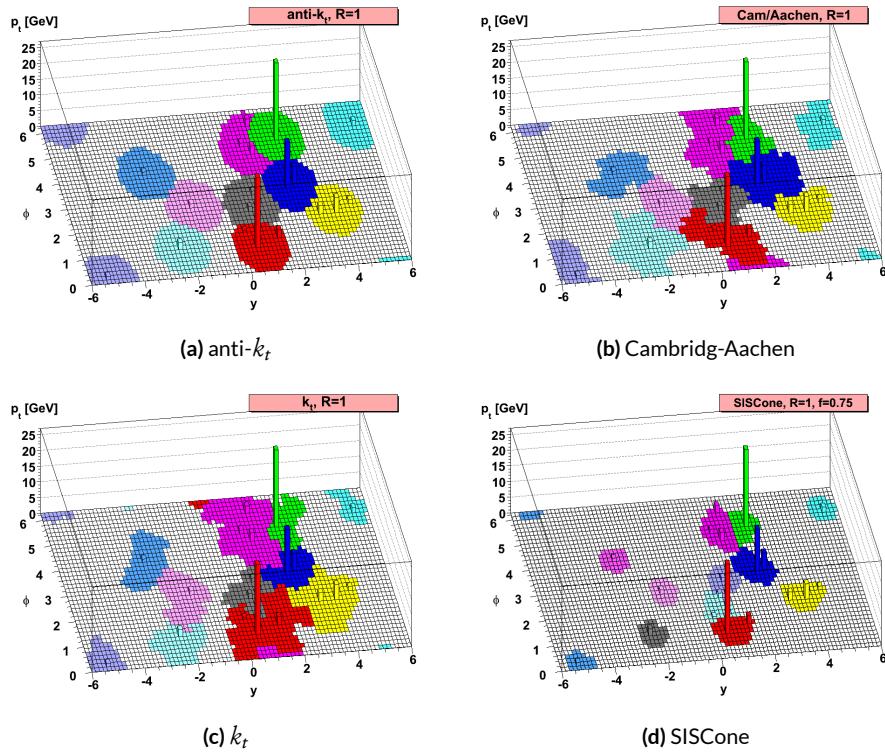


Figure 5.3: Different jet algorithms used on the same event. IC:[50]

1236 All three algorithms have uses for different applications in ATLAS, with anti- k_t $R = 0.4$ jets
1237 being the default jet collection.[†] These are the jets used in this analysis.
1238 If the choice of jet algorithm seems a little arbitrary, it is. There is no one-size-fits-all jet collection
1239 perfect for every application, and analyzers have to make these choices for themselves. One interest-
1240 ing choice is the jet size parameter, R . A large R jet will contain more of the radiation coming from a
1241 final state object, but its large size makes it susceptible to contamination from the underlying event
1242 and pileup (as well as other analysis objects if R is sufficiently large or objects sufficiently boosted),
1243 with small R jets having the opposite features. $R = 0.4$ is a fairly middle-of-the-road choice. A natu-
1244 ral question to ask is whether there needs to be just one jet collection in an analysis. Might there not
1245 be more information to be gained from looking at more jet sizes or clusterings? Preliminary studies
1246 point to this answer being yes and are addressed in Appendix B.

1247 5.5.2 STANDARD ATLAS HBB JETS

1248 There are a few considerations that arise with jets in physical detectors. The first is what type of ob-
1249 ject to use as the stand in for “particles” since ATLAS measures energy deposits, not pions. The ap-
1250 proach ATLAS has settled upon are calorimeter topological clusters (or CaloTopoClusters for short)
1251 [S1]. Unlike the sliding window algorithm used for electron clusters, CaloTopoClusters use a noise
1252 significance based approach in the “4-2-o” algorithm. Each cell in the electromagnetic and hadronic
1253 calorimeters (ECAL and HCAL, respectively) has associated with it a characteristic noise level (N in
1254 Equation 2.6), with this noise level in each channel, it is possible to construct a “significance” for the

†The other collections find their primary uses in jet substructure techniques. For an example, cf. the discussion of jet trimming in Appendix B.

registered energy deposit in a given channel for a given event by dividing the measured value by its characteristic noise. Groups of cells having a significance of 4 are taken as the centers of clusters in the $\eta - \phi$ plane. The second layer in a cluster includes all neighboring cells to the central layer with significance of at least 2, and the final layer includes all the nearest neighbors to the second layer.
 This is described in Figure 5.4 from [52].

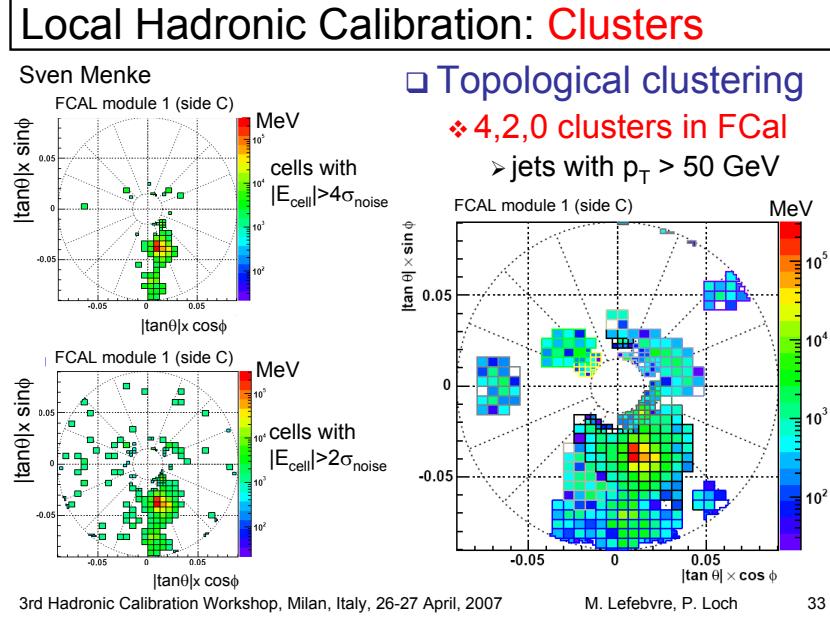


Figure 5.4: A description of the 4-2-0 clustering algorithm.

Once CaloTopoClusters have been formed and clustered into jets, they are calibrated using the electromagnetic (EM) scale (the scale for clusters coming from EM showers). Further details may be found in [53].

Jets in this analysis fall into two categories, “signal” and “forward” jets, and are required to pass certain quality requirements, described in Table 5.5. All jets must pass a series of jet cleaning require-

ments using calorimeter level variables to eliminate jets coming from problematic calorimeter cells
 and certain backgrounds. Some signal jet candidates also make use of a Jet Vertex Tagger (JVT) that
 uses primary vertex and jet and track p_T information to decide whether certain soft jets are likely to
 have come from the the primary (hard scatter) vertex in an event or are to be considered pileup. Fur-
 ther details on JVT may be found in [54]. Jets are further corrected using standard CP tools and a
 dedicated PtReco correction, all outlined in Section 7.3 of [44].

Jet Category	Selection Requirements
Forward Jets	jet cleaning $p_T > 30 \text{ GeV}$ $2.5 \leq \eta < 4.5$
Signal Jets	$p_T > 20 \text{ GeV}$ and $ \eta < 2.5$ jet cleaning $\text{JVT} \geq 0.59$ if ($p_T < 60 \text{ GeV}$ and $ \eta < 2.4$)

Table 5.5: AntiKt4EMTopoJets selection requirements. The jet cleaning is applied via the JetCleaningTool, that removes events in regions corresponding to hot calorimeter cells.

Overlap removal in this analysis is done according to the following precedence, taken from [44]
 with further steps only taken into account if an object survives previous steps:
 • tau-electron: If $\Delta R(\tau, e) < 0.2$, the τ lepton is removed.
 • tau-muon: If $\Delta R(\tau, \mu) < 0.2$, the τ lepton is removed, with the exception that if the τ lepton
 has $p_T > 50 \text{ GeV}$ and the muon is not a combined muon, then the τ lepton is not removed.
 • electron-muon: If a combined muon shares an ID track with an electron, the electron is re-
 moved.
 If a calo-tagged muon shares an ID track with an electron, the muon is removed.

1279 • electron-jet: If $\Delta R(\text{jet}, e) < 0.2$ the jet is removed.

1280 For any surviving jets, if $\Delta R(\text{jet}, e) < \min(0.4, 0.04 + 10 \text{ GeV}/p_T^e)$, the electron is removed.

1281 • muon-jet If $\Delta R(\text{jet}, \mu) < 0.2$ or the muon ID track is ghost associated to the jet, then the jet is

1282 removed if the jet has less than three associated tracks with $p_T > 500 \text{ MeV}$ ($\text{NumTrkPt}500\text{PV}^{\text{jet}} < 3$)

1283 or both of the following conditions are met: the p_T ratio of the muon and jet is larger than 0.5 ($p_T^\mu/p_T^{\text{jet}} >$

1284 0.5) and the ratio of the muon p_T to the sum of p_T of tracks with $p_T > 500 \text{ MeV}$ associated to the

1285 jet is larger than 0.7 ($p_T^{\text{muon}}/\text{SumPtTrkPt}500\text{PV}^{\text{jet}} > 0.7$).

1286 For any surviving jets, if $\Delta R(\text{jet}, \mu) < \min(0.4, 0.04 + 10 \text{ GeV}/p_T^\mu)$, the muon is removed.

1287 • tau-jet: If $\Delta R(\tau, \text{jet}) < 0.2$, the jet is removed.

1288 • electron-fat jet: If $\Delta R(e, \text{fat jet}) < 1.2$, the fat jet is removed.

1289 Jets are corrected using a muon-in-jet correction and then a kinematic fitter (Appendix D of [4])

1290 for the 2-lepton case (PtReco correction for the 0- and 1- lepton case). The muon-in-jet correction

1291 is designed for b -jets. Since the decay of a b -quark to a c -quark and finally to a light quark (these are

1292 the multiple vertices for which JetFitter in Section 5.6.1 searches) involves two weak decays, there are

1293 two W -bosons involved in the decay. Some of these will decay semileptonically, and, while electron

1294 and τ energy will be captured by the calorimeters, semileptonic μ 's will only be registered in the MS,

1295 which occurs in some 44% of all decays from a theoretical standpoint, which amounts to about 12%

1296 in practice (due to track isolation requirements for the leptons). This value is about 1–2% for elec-

1297 trons, which deposit their energy in the calorimeter and so require no correction; any jet with a valid

1298 lepton associated to it is deemed semileptonic (all others are called hadronic). Any jet with muons

1299 associated with it has the closest muon's 4-vector (in the $\eta - \phi$ plane) added to it.

1300 The PtReco correction is a scale factor on the muon-in-jet corrected jet's 4-vector based on the
 1301 jet's p_T and whether the jet is hadronic or semileptonic. This correction factor is based on particle
 1302 level studies done on a TruthWZ sample. As the o- and i-lepton cases are not the focus of this thesis,
 1303 the interested reader is directed to Section 7.3 of [44].

1304 The kinematic fitter used in 2-lepton events with two or three jets takes as its input 12 fit parame-
 1305 ters,

- 1306 • energies of 2 electron or p_T of 2 muons
- 1307 • energies of 2 b -jets
- 1308 • η, ϕ of 2 leptons and 2 jets
- 1309 • p_X and p_Y of $\ell\ell b\bar{b}$ system.
- 1310 • $m_{\ell\ell}$

1311 and 3 constraints for the variation of these parameters,

- 1312 • parameters : Gaussian (b -jet energy : Transfer Functions (TF); these are denoted L , with an
 1313 L_{truth} as a prior) (the ϕ parameters)
- 1314 • p_X and p_Y of $\ell\ell b\bar{b}$ system : zero with a width of 9 GeV obtained from ZH signal MC.
- 1315 • m_{ll} : Breit-Wigner (BW) distribution of Z boson (final term, leptons denoted Ω)

1316 which leads to test statistic from the usual likelihood formalism to be minimized in each event:

$$\begin{aligned}
 -2 \ln \mathcal{L} = & \sum_{i=j} \left(\frac{(\phi_i^n - \phi_i^0)^2}{\sigma_\phi^2} \right) + \left(\frac{(\Omega_l^n - \Omega_l^0)^2}{\sigma_\Omega^2} \right) - 2 \ln(L^j) - 2 \ln(L_{truth}^j) \\
 & + \sum_{i=x,y} \frac{(\sum p_i^n - \sum P_i)^2}{\sigma_{\sum p_i}^2} + 2 \ln((m_{\ell\ell} - M_X^2)^2 + M_X^2 \Gamma^2)
 \end{aligned} \tag{5.4}$$

1317 5.5.3 JET SYSTEMATICS

1318 As with the electron systematics, jet energy scale (JES) and resolution (JER) are the two principal
1319 considerations for systematic uncertainties, with even more standard. JER, as with the electron en-
1320 ergy resolution, is a single systematic uncertainty, `JET_JER_SINGLE_NP`. There is also a single JVT
1321 efficiency `JET_JvtEfficiency` systematic uncertainty. There are 88 nominal JES systematics, and
1322 this analysis is sufficiently sensitive to these variations that a single systematic is grossly insufficient.
1323 Nevertheless, some simplification is possible, with the 75 of these nuisance parameters (mostly statis-
1324 tical uncertainties related to the Z +jet and γ +jet calibrations) being reduced to 8, and several explic-
1325 itly named nuisance parameter. These remaining named NP's are: 3 NP's related to the η intercali-
1326 bration used to extrapolate standard calibrations to other jet η regions, 4 NP's related to the flavor
1327 composition of principle background samples (W/Z +jets, top, and diboson), 4 pileup systematics, a
1328 single NP for the b -jet energy scale, a high p_T jet energy scale systematic, and one for jets that punch
1329 through the HCAL to leave energy deposits in the MS. These are listed explicitly in Table 5.7.

1330 5.6 FLAVOR TAGGING

1331 Given that the final state in this analysis involves pairs of jets originating from b -quarks, deploying
1332 effective flavor tagging algorithms is imperative. While flavor tagging in general can be used to isolate
1333 any flavor (b , c , or light (u , d , s , or gluon-initiated jets)), this analysis exclusively looks for b -jets, so
1334 this disucssion will focus on b 's. At truth-level in sumlation, this is fairly straightforward: one need
1335 only look at the particles contained within a jet and seeing if any include a b -quark (sometimes a B

₁₃₃₆ hadron) in the decay chain.

₁₃₃₇ **5.6.1 DIRECT TAGGING**

₁₃₃₈ One of the most distinctive features of b -jets is the presence of secondary vertices, as illustrated in Fig-
₁₃₃₉ ure 5.5. While most partons created in particle collisions will hadronize promptly, b -quarks will first
₁₃₄₀ hadronize into B -hadrons, which have lifetimes of about a picosecond. This small but finite lifetime
₁₃₄₁ means that these particles will travel about half a millimeter or so before decaying into a jet in much
₁₃₄₂ the usual way, and the tracks from this decay will point back to this displaced, secondary vertex.

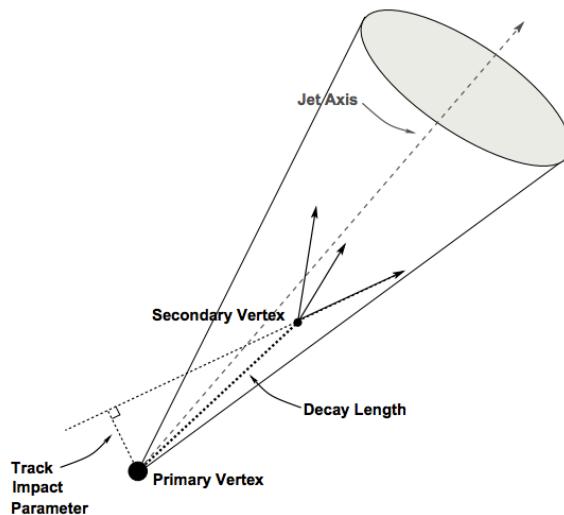


Figure 5.5: An illustration of a secondary vertex in a b -jet. Image credit: [55]

₁₃₄₃ There are various secondary vertex algorithms used as inputs to the nominal b -tagging algorithm

₁₃₄₄ [56], with three main types of algorithms used as inputs

- ₁₃₄₅ 1. Track impact parameter based algorithms: I₂PD (signed transverse only; more pileup ro-
₁₃₄₆ bust), I₃PD (signed transverse and longitudinal)

- 1347 2. Inclusive secondary vertex reconstruction: SV1 (start with two track vertex pairs and con-
 1348 struct a secondary vertex)

 1349 3. Multiple vertex reconstruction (decay chain): JetFitter ($PV \rightarrow b \rightarrow c$ decay chain using Kalman
 1350 filter)
- 1351 All of these are combined into a boosted decision tree (BDT) and trained on five million $t\bar{t}$ events
 1352 with an 90%/10% light/ c jet background to form the MV2c10 algorithm, with 10 referring to the per-
 1353 centage of charm events in the training background. The 10% charm ratio was found to be a good
 1354 balance between increased charm rejection capability (as opposed to MV2c00, which has no charm
 1355 in the background training samples) and loss in light jet rejection (compared to MV2c20, which has
 1356 20% charm events in background training samples).

1357 ANALYSIS SPECIFIC CONCERNS AND SYSTEMATIC UNCERTAINTIES In addition to specifying
 1358 the tagging algorithm, the working point efficiency must be specified. As with selection algorithms
 1359 in general, there is a trade off between efficiency/recall (identifying all the b -jets, minimizing type
 1360 II error) and purity/precision (making sure all jets positively identified are in fact b -jets, minimiz-
 1361 ing type I error). Nominal efficiency working points have been calibrated by the flavor tagging CP
 1362 group and are outlined in Table 5.6.

name	MV2c10 weight cut	b -tag eff, [%]	c RR	light RR
FixedCutBEff_60	0.9349	60.03	34.54	1538.78
FixedCutBEff_70	0.8244	69.97	12.17	381.32
FixedCutBEff_77	0.6459	76.97	6.21	134.34
FixedCutBEff_85	0.1758	84.95	3.10	33.53

Table 5.6: b -tagging working points available for MV2c10 for AntiKt4EMTopoJets. RR is the rejection rate (the inverse of efficiency).

1363 These values are aggregate figures, as both the jet's p_T and η are inputs to the MV2c10 discrim-
1364 inant. The working point chosen for this analysis is the 70% `FixedCutBEff_70` working point,
1365 with "fixed cut" referring to the fact that this particular usage of the MV2c10 BDT value is a simple
1366 cut value.

1367 Just as with the trigger and lepton identification efficiencies, flavor tagging efficiencies differ from
1368 their nominal values somewhat depending on what simulation or data sample is being used. To
1369 account for this difference, just as in the other cases, scale factors are applied to simulation event
1370 weights. It is through these event weights, as with the modeling systematics, that the flavor tagging
1371 systematic uncertainties are applied. Given that there are 24 input variables to MV2c10 and that
1372 flavor tagging is in general a very difficult problem, it is not surprising that, as with the JES, there
1373 are very many systematic uncertainties associated with flavor tagging. However, as with JES, the
1374 CP group has compacted the full systematic set into a reduced set of 13 systematic uncertainties: 3
1375 each associated with c and light jets, 5 for b -jets (with the naming convention `FT_EFF_Eigen_-`
1376 ($B|C|Light$) N), one for the extrapolation of scale factors to different jet p_T regimes (`FT_EFF_-`
1377 `Eigen_extrapolation`), and one for the charm to bottom extrapolation (`FT_EFF_Eigen_-`
1378 `extrapolation_from_charm`) [57]. This schematic is a middle-of-the-road "Medium" set of sys-
1379 tematics.

1380 5.6.2 TRUTH TAGGING

1381 Since imposing a 2 b -tag requirement overwhelmingly rejects events dominated by c - and light jets,
1382 statistics in such MC samples are very low. In order to circumvent this problem and restore full MC

1383 statistics, the tag rate function, or “truth-tagging” procedure (in contrast to the standard or “direct
1384 tagging” procedure) is applied, in which all events are kept but given a weight that preserves the over-
1385 all shape and normalization of underlying distributions. Intuitively, this is done by giving events with
1386 real b -jets in MC a much higher weight than events having only c - or light jets. Truth-tagging is ap-
1387 plied to all samples when conducting MVA training in order to maximize statistics and reduce the
1388 risk of overtraining. Truth-tagging is also used for data-MC comparison plots in 2-tag regions and
1389 for $V + cc$, $V + c\ell$, $V + \ell$, and WW samples used in the final likelihood fit. A detailed description of
1390 the truth-tagging process is provided below.

1391 Each jet in a given event has associated with it a b -tagging efficiency, denoted ε , that is a function
1392 of its p_T , η , and real flavor (b , c , or light) from truth-level information in MC. Intuitively, this effi-
1393 ciency can be thought of as the likelihood that a given jet will be b -tagged. Hence, b -jets have a much
1394 higher b -tagging efficiency than c -jets, which in turn have a higher b -tagging efficiency than light jets.
1395 We define a truth-tag weight for a given combination of tagged and untagged jets as the product of
1396 the efficiencies of the tagged jets times the product of the complement of the efficiencies of the un-
1397 tagged jets. For example, for an event with three jets, labeled 1, 2, and 3, if jets 1 and 2 are tagged, and
1398 jet 3 is untagged, the truth-tag weight associated with this combination is

$$\varepsilon_1 \varepsilon_2 (1 - \varepsilon_3) \quad (5.5)$$

1399 In order to obtain a truth-tag weight for an event, one takes the sum of the weights for each pos-
1400 sible tag combination. The current analysis requires that all events have exactly two b -tagged jets, so

1401 the truth-tag weight is the sum of all the weights of all possible pairs of tagged jets (events with fewer
 1402 than two jets are discarded). Going back to the three jet example, one has the possible combinations:
 1403 jets 1+2 as tagged and jet 3 as untagged; jets 1+3 as tagged and jet 2 as untagged; and finally jets 2+3 as
 1404 tagged and jet 1 as untagged, which yields a total event weight of

$$w_{tot} = \varepsilon_1 \varepsilon_2 (1 - \varepsilon_3) + \varepsilon_1 \varepsilon_3 (1 - \varepsilon_2) + \varepsilon_2 \varepsilon_3 (1 - \varepsilon_1) \quad (5.6)$$

1405 For some applications (e.g. in order to use variables like pTB1, the p_T of the harder b -tagged jet in
 1406 an event, in MVA training), it is necessary to choose a combination of jets in an event as “tagged.”
 1407 This combination is chosen randomly, with the probability for a given combination to be chosen
 1408 being proportional to its truth-tag weight. In the three jet example, the probability of tagging jets
 1409 1+2 is:

$$\frac{\varepsilon_1 \varepsilon_2 (1 - \varepsilon_3)}{w_{tot}} = \frac{\varepsilon_1 \varepsilon_2 (1 - \varepsilon_3)}{\varepsilon_1 \varepsilon_2 (1 - \varepsilon_3) + \varepsilon_1 \varepsilon_3 (1 - \varepsilon_2) + \varepsilon_2 \varepsilon_3 (1 - \varepsilon_1)} \quad (5.7)$$

1410 Though not used in the current analysis, functionality exists for generic truth-tagging require-
 1411 ments (i.e. an arbitrary number of tags on an arbitrary number of jets) through the logical combina-
 1412 toric extension and for so-called “pseudo-continuous tagging,” where a b -tag score is generated for
 1413 each jet in a given event. Since a random combination of jets is set by hand to pass the b -tagging cuts
 1414 regardless of its b -tag score, a new score must be generated if this information is to be used in further
 1415 analysis. Under current settings, jets that are tagged are assigned a random b -tag score that is sampled
 1416 from the MV2c10 cumulative distribution above the 70% efficiency working point cut. All other

¹⁴¹⁷ jets in the event are assigned a random b -tag score below the 70% working point cut. Since these dis-
¹⁴¹⁸ tributions are discrete, the scores are not truly continuous (cf. example distributions in Figure 5.6), hence the “pseudo-continuous” nomenclature.

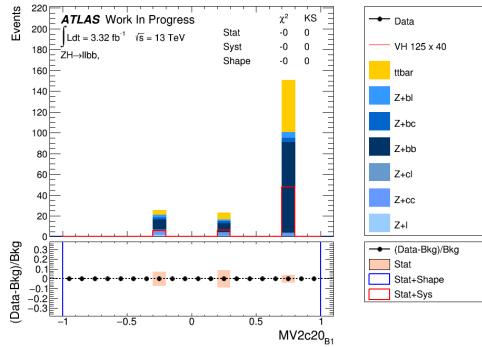


Figure 5.6: An example of a pseudo-continuous b -tagging distribution

¹⁴¹⁹

¹⁴²⁰ A number of closure tests were performed on both the nominal and several systematics cases. In
¹⁴²¹ the plots that follow, truth (solid) and direct (dashed) tagging distributions for m_{bb} and $\Delta R(b_1, b_2)$
¹⁴²² in different p_T^V regimes for 2 lepton, 2 jet events. Agreement between the truth and direct tagging
¹⁴²³ cases is generally very good, an example of which can be seen in Figure 5.7 for a signal qqZllH125
¹⁴²⁴ sample, and the overall benefit of truth-tagging can be somewhat dramatically seen in the corre-
¹⁴²⁵ sponding plots $Z + \ell$ samples in Figure 5.8. At high p_T^V ($p_T^V > 200$ GeV), however, in events with two
¹⁴²⁶ real b -jets, there is a much greater likelihood that the b -jets will merge into a single jet, which render
¹⁴²⁷ the naïve assumption that jets remain discrete invalid. While this does not appear to be a problem in
¹⁴²⁸ most samples (cf. $t\bar{t}$ in Figure 5.9), there is a mismodelling effect at low m_{bb} and low $\Delta R(b_1, b_2)$ at
¹⁴²⁹ $p_T^V > 200$ GeV for $W/Z + bb$ samples where truth-tagging overestimates the number of events in
¹⁴³⁰ this merged regime, as can be seen in Figure 5.10

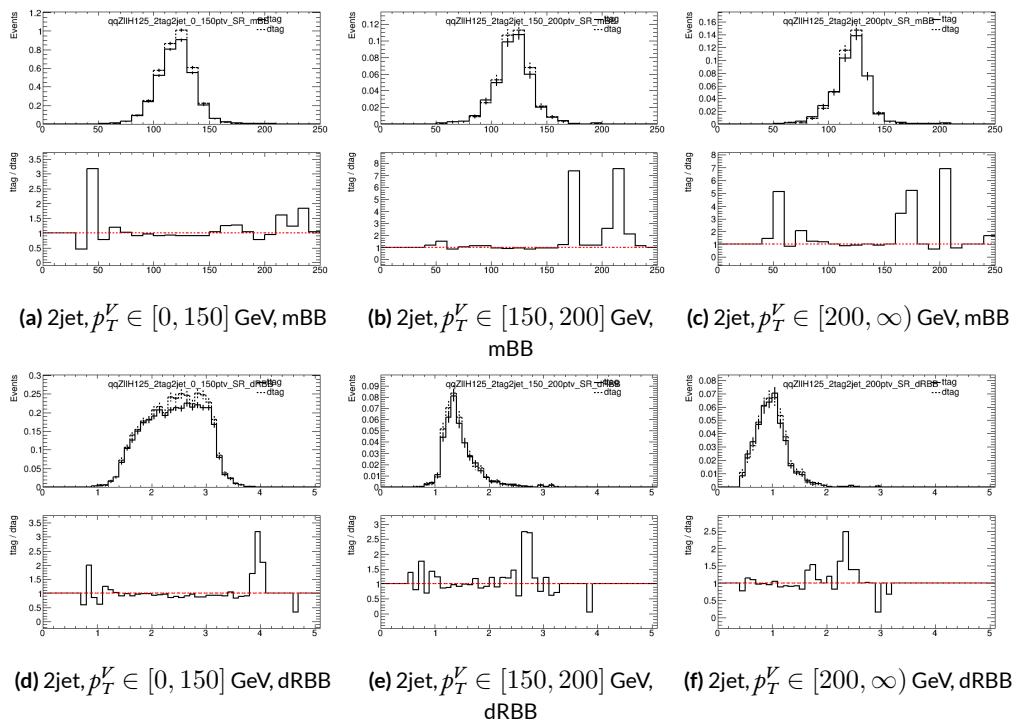


Figure 5.7: Truth-tagging closure tests for 2 lepton, 2 jet qqZIIH125 samples in three different p_T^V regions.

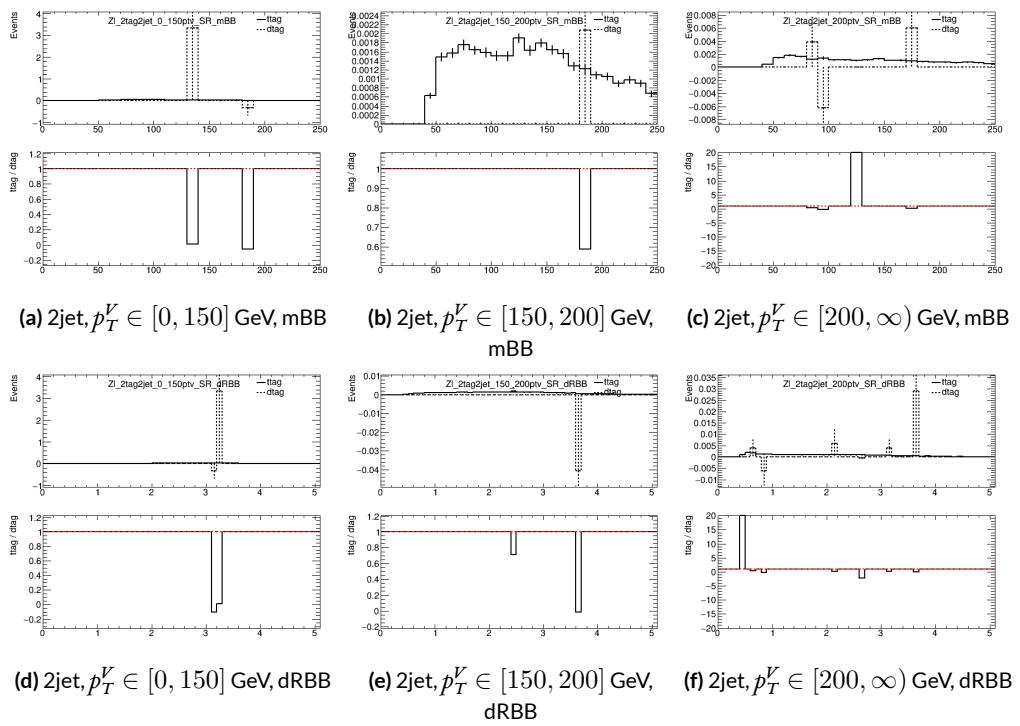


Figure 5.8: Truth-tagging closure tests for 2 lepton, 2 jet $Z + \ell$ samples in three different p_T^V regions.

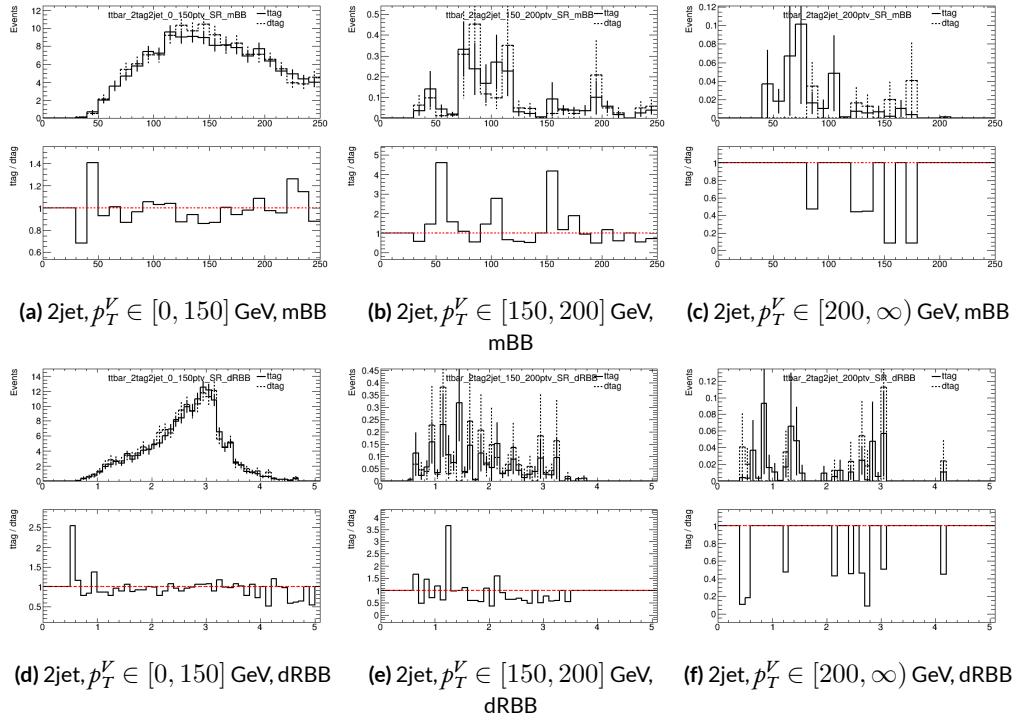


Figure 5.9: Truth-tagging closure tests for 2 lepton, 2 jet $t\bar{t}$ samples in three different p_T^V regions.

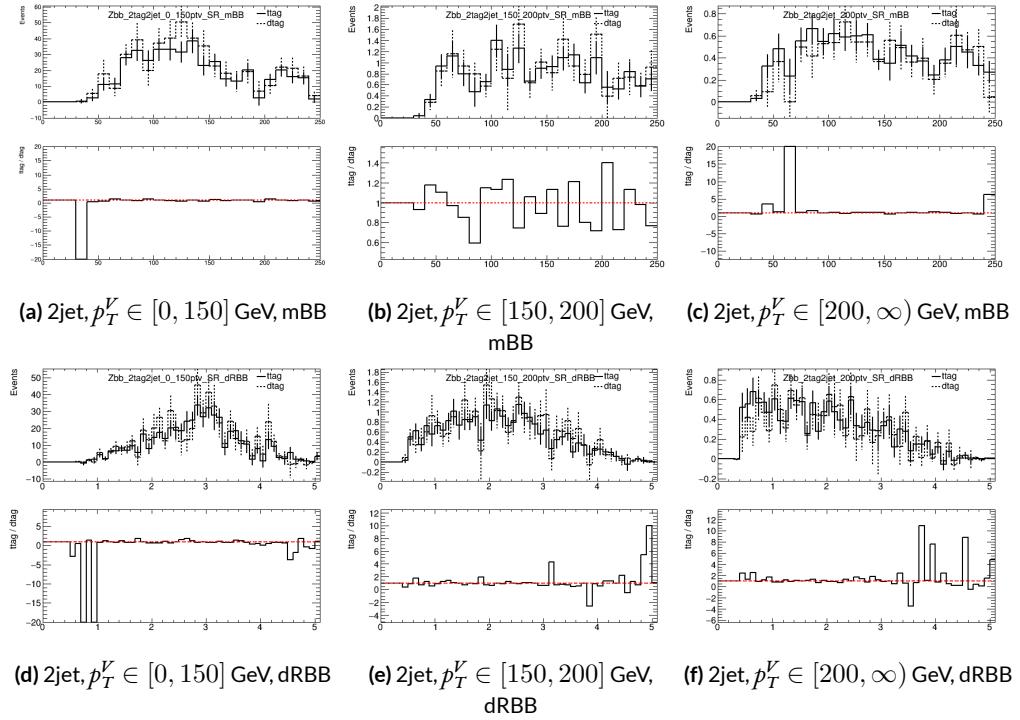


Figure 5.10: Truth-tagging closure tests for 2 lepton, 2 jet $Z + bb$ samples in three different p_T^V regions.

1431 5.7 MISCELLANIA AND SYSTEMATICS SUMMARY

1432 A summary of all experimental systematics, taken from [4], may be found below. In addition to the
1433 systematics discussed above, there are also two further systematics, on the total integrated luminosity
1434 and on the event reweighting factor used to account for pileup, both included in Table 5.7.

1435 5.8 EVENT SELECTION AND ANALYSIS REGIONS

1436 With object and event reconstruction described, it is now time to address which events are actually
1437 selected for use in analysis. This analysis focuses specifically on the 2-lepton channel of the fiducial
1438 analysis, with the event selection and analysis region definitions being identical. Common to all
1439 lepton channels in the fiducial analysis is the set of requirements on the jets in a given event. There
1440 must be at least two central jets and exactly two signal jets that have been “*b*-tagged” according to the
1441 MV2c10 algorithm [56], with at least one of these *b*-jets having $p_T > 45$ GeV. For MVA training and
1442 certain background samples, a process known as “truth-tagging” is applied instead of the standard
1443 *b*-tagging to boost sample statistics and stabilize training/fits (cf. [4] Section 4.2 for details). After
1444 event selection, the *muon-in-jet* and *PtReco* corrections, described in [44] 6.3.3-4, are applied to the
1445 *b*-jets.

1446 In addition to the common selections, there are 2-lepton specific selections. All events are re-
1447 quired to pass an un-prescaled single lepton trigger, a full list of which may be found in Tables 5 and
1448 6 of [44] with the requirement that one of the two selected leptons in the event must have fired the
1449 trigger. There must be 2 VH-loose leptons, and at least one of these must be a ZH-signal lepton (cf.

Systematic uncertainty	Short description	Reference
Luminosity	uncertainty on total integrated luminosity	Section 11.1 in Ref. [44]
Pileup Reweighting	uncertainty on pileup reweighting	Section 11.1 in Ref. [44]
	Electrons	
EL_EFF_Trigger_Total_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty	Section 11.2.2. in Ref. [44]
EL_EFF_Reco_Total_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty	Section 11.3.1. in Ref. [44]
EL_EFF_ID_Total_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty	Section 11.3.1. in Ref. [44]
EL_EFF_Iso_Total_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty	Section 11.3.1. in Ref. [44]
EG_SCALE_ALL	energy scale uncertainty	Section 11.3.2. in Ref. [44]
EG_RESOLUTION_ALL	energy resolution uncertainty	Section 11.3.2. in Ref. [44]
	Muons	
MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainty	Section 11.2.2. in Ref. [44]
MUON_EFF_TrigSystUncertainty	reconstruction and ID efficiency uncertainty for muons with $p_T > 15$ GeV	Section 11.4.1. in Ref. [44]
MUON_EFF_STAT	reconstruction and ID efficiency uncertainty for muons with $p_T < 15$ GeV	Section 11.4.1. in Ref. [44]
MUON_EFF_SYS	isolation efficiency uncertainty	Section 11.4.1. in Ref. [44]
MUON_EFF_STAT_LOWPT	track-to-vertex association efficiency uncertainty	Section 11.4.1. in Ref. [44]
MUON_EFF_SYST_LOWPT	momentum resolution uncertainty from inner detector	Section 11.4.2. in Ref. [44]
MUON_ISO_STAT	momentum resolution uncertainty from muon system	Section 11.4.2. in Ref. [44]
MUON_ISO_SYS	momentum scale uncertainty	Section 11.4.2. in Ref. [44]
MUON_TTVA_STAT	charge dependent momentum scale uncertainty	Section 11.4.2 in Ref. [44]
MUON_TTVA_SYS		
MUON_ID		
MUON_MS		
MUON_SCALE		
MUON_SAGITTA_RHO		
MUON_SAGITTA_RESBIAS		
	Jets	
JET_21NP_JET_EffectiveNP_1	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EffectiveNP_2	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EffectiveNP_3	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EffectiveNP_4	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EffectiveNP_5	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EffectiveNP_6	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EffectiveNP_7	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EffectiveNP_8restTerm	energy scale uncertainty from the in situ analyses splits into 8 components	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EtaIntercalibration_Modeling	energy scale uncertainty on eta-intercalibration (modeling)	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EtaIntercalibration_TotalStat	energy scale uncertainty on eta-intercalibrations (statistics/method)	Section 11.5.1. in Ref. [44]
JET_21NP_JET_EtaIntercalibration_NonClosure	energy scale uncertainty on eta-intercalibrations (non-closure)	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Pileup_OffsetMu	energy scale uncertainty on pile-up (mu dependent)	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Pileup_OffsetNPV	energy scale uncertainty on pile-up (NPV dependent)	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Pileup_PtTerm	energy scale uncertainty on pile-up (pt term)	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Pileup_RhoTopology	energy scale uncertainty on pile-up (density ρ)	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Flavor_Composition_Zjets	energy scale uncertainty on $Z+jets$ sample's flavour composition	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Flavor_Composition_Wjets	energy scale uncertainty on $W+jets$ sample's flavour composition	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Flavor_Composition_top	energy scale uncertainty on top sample's flavour composition	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Flavor_Composition	energy scale uncertainty on VV and VH sample's flavour composition	Section 11.5.1. in Ref. [44]
JET_21NP_JET_Flavor_Response	energy scale uncertainty on samples' flavour response	Section 11.5.1. in Ref. [44]
JET_21NP_JET_BJES_Response	energy scale uncertainty on b-jets	Section 11.5.1. in Ref. [44]
JET_21NP_JET_PunchThrough_MC15	energy scale uncertainty for punch-through jets	Section 11.5.1. in Ref. [44]
JET_21NP_JET_SingleParticle_HighPt	energy scale uncertainty from the behaviour of high- p_T jets	Section 11.5.1. in Ref. [44]
JET_JER_SINGLE_NP	energy resolution uncertainty	Section 11.5.1. in Ref. [44]
JET_JvtEfficiency	JVT efficiency uncertainty	Section 11.5.1. in Ref. [44]
FT_EFF_Eigen_B	b -tagging efficiency uncertainties (“BTAG_MEDIUM”): 3 components for b jets, 3 for c jets and 5 for light jets	Section 11.7. in Ref. [44]
FT_EFF_Eigen_C		
FT_EFF_Eigen_L		
FT_EFF_Eigen_extrapolation	b -tagging efficiency uncertainty on the extrapolation to high- p_T jets	Section 11.7. in Ref. [44]
FT_EFF_Eigen_extrapolation_from_charm	b -tagging efficiency uncertainty on tau jets	Section 11.7. in Ref. [44]
	MET	
METTrigStat	trigger efficiency uncertainty	Section 11.2.1. in Ref. [44]
METTrigTop/Z	track-based soft term related longitudinal resolution uncertainty	Section 11.6. in Ref. [44]
MET_SoftTrk_ResoPara	track-based soft term related transverse resolution uncertainty	Section 11.6. in Ref. [44]
MET_SoftTrk_ResoPerp	track-based soft term related longitudinal scale uncertainty	Section 11.6. in Ref. [44]
MET_SoftTrk_Scale	track MET scale uncertainty due to tracks in jets	Section 11.6. in Ref. [44]

Table 5.7: Summary of the experimental systematic uncertainties considered. Details on the individual systematic uncertainties can be found in the given Sections of Ref. [44].

¹⁴⁵⁰ Tables 5.3 and 5.4 for definitions). This lepton pair must have an invariant mass between 81 and 101
¹⁴⁵¹ GeV. In addition to the jet corrections described above, a kinematic fitter is applied to the leptons
¹⁴⁵² and two leading corrected jets in an event with three or fewer jets[‡] to take advantage of the fact that
¹⁴⁵³ the 2-lepton final state is closed (cf. [20]); these objects are only used for MVA training/fit inputs.

¹⁴⁵⁴ In order to increase analysis sensitivity, the analysis is split into orthogonal regions based on the
¹⁴⁵⁵ number of jets and the transverse momentum of the Z candidate (the vectoral sum of the lepton
¹⁴⁵⁶ pair; this p_T is denoted p_T^V): 2 and ≥ 3 jets; p_T^V in $[75, 150), [150, \infty)$ GeV. In addition to the signal
¹⁴⁵⁷ regions where the leptons are required to be the same flavor (e or μ), there are top $e - \mu$ control
¹⁴⁵⁸ regions used to constrain the top backgrounds.

¹⁴⁵⁹ All of these requirements are summarized in 5.8.

Category	Requirement
Trigger	un-prescaled, single lepton
Jets	≥ 2 central jets; 2 b -tagged signal jets, harder jet with $p_T > 45$ GeV
Leptons	2 VH-loose leptons (≥ 1 ZH-signal lepton); same (opp) flavor for SR (CR)
$m_{\ell\ell}$	$m_{\ell\ell} \in (81, 101)$ GeV
p_T^V regions (GeV)	$[75, 150), [150, \infty)$

Table 5.8: Event selection requirements

¹⁴⁶⁰ It should be noted that the use of ≥ 3 jet events is a 2-lepton specific selection. These regions are
¹⁴⁶¹ exclusive 3 jet regions in the 0- and 1-lepton channels, but the fiducial 2-lepton analysis was found to
¹⁴⁶² see a $\sim 4\%$ gain in sensitivity in studies by including ≥ 4 jet events [4].

[‡]The gain from using the kinematic fitter is found to be smeared out in events with higher jet multiplicities.

猛き者も遂には滅びぬ、

偏に風の前の塵に同じ。

Heike monogatari

6

1463

1464

Multivariate Analysis Configuration

1465 IN ORDER TO fully leverage the descriptive power of the 13 TeV dataset, this analysis makes use of a

1466 multivariate (MVA) discriminant. Where traditionally event counts or single discriminating vari-

1467 ables per region of phase space have been fed to fits, MVA discriminants seek to integrate additional

1468 information not captured in the conventional phase space cuts plus dijet invariant mass distribu-

1469 tions. Formulating the MVA discriminant is an exercise in supervised learning to construct a binary
1470 classifier, where one uses labeled “signal” and “background” MC events to optimize the parameters
1471 of a statistical model—in this case a boosted decision tree (BDT) with some set of physically moti-
1472 vated variables (or “factors”). The interested reader is directed to the standard references on machine
1473 learning for further details. Sample and variable selection, including variables derived using the the
1474 RestFrames and Lorentz Invariants concepts introduced in Sections 1.5–1.7, are discussed in Section
1475 6.1; MVA training is treated in Section 6.2; and the data statistics only (no systematics) performance
1476 of the three MVA discriminants is explored in Section 6.3.

1477 6.1 TRAINING SAMPLES AND VARIABLE SELECTION

1478 A subset of samples described in Chapter 3 is used for multivariate analysis training, with $qqZH \rightarrow$
1479 $\ell\ell b\bar{b}$ and $ggZH \rightarrow \ell\ell b\bar{b}$ used as signal samples and $Z+jets$, $t\bar{t}$, and VV used as background samples.
1480 Truth-tagging (Section 5.6.2) is used on all samples in MVA training to improve training statistics
1481 and stability. All figures quoted in this section scale distributions to a luminosity of 36.1 fb^{-1} .

1482 6.1.1 STANDARD VARIABLES

1483 The standard set of variables taken as a baseline is the same as used in the fiducial analysis. The vari-
1484 ables fall into several main categories: energy/momenta scales of composite objects (m_{bb} , m_{bbj} ,
1485 p_T^V , $m_{\ell\ell}$), angles ($\Delta R(b_1, b_2)$, $\Delta\phi(V, H)$, $\Delta\eta(V, H)$), transverse momenta of the jets in the event
1486 ($p_T^{b_1}$, $p_T^{b_2}$, $p_T^{j_3}$), and E_T^{miss} . Input distributions for these variables in all the 2 (≥ 3 jet) analysis signal
1487 regions may be found in Figure 6.1 (6.2). The “kf” at the end of variable names denotes that these

are derived using 4-vectors that are the result of the kinematic fitter. The distributions in the figure

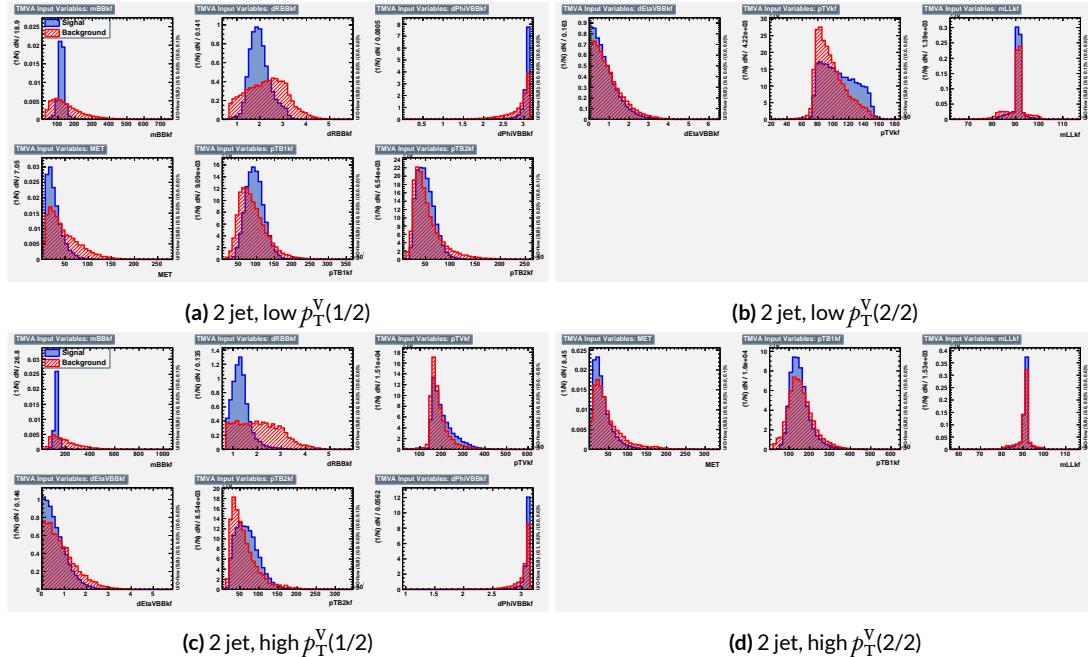


Figure 6.1: Input variables in 2 jet signal regions for the standard variable set. Signal distributions are in red, and background distributions are in blue.

1488

1489 are used as inputs for one of the two k-folded final discriminants, and the order of the distributions
 1490 is the hyperparameter optimized order for feeding into the BDT; what precisely this means will be
 1491 discussed in following sections. While variables in the analysis regions are generally similar, there are
 1492 some notable exceptions. p_T^V and the correlated $\Delta R(b_1, b_2)$ have different shapes, by construction
 1493 for the former and by correlation for the latter, at low and high p_T^V . * The ≥ 3 jet regions also have
 1494 variables that are not applicable to the 2 jet regions; the inclusion of m_{BBJ} (the invariant mass of the
 1495 two b -jets and leading untagged jet) in particular is of note and suggests a potential avenue forward

*Recall that higher p_T^V means, in a balanced final state like $ZH \rightarrow \ell\ell b\bar{b}$, the b -jet pair will have higher p_T and hence be more collimated (lower $\Delta R(b_1, b_2)$); this is not necessarily the case for background events, as the distributions show.

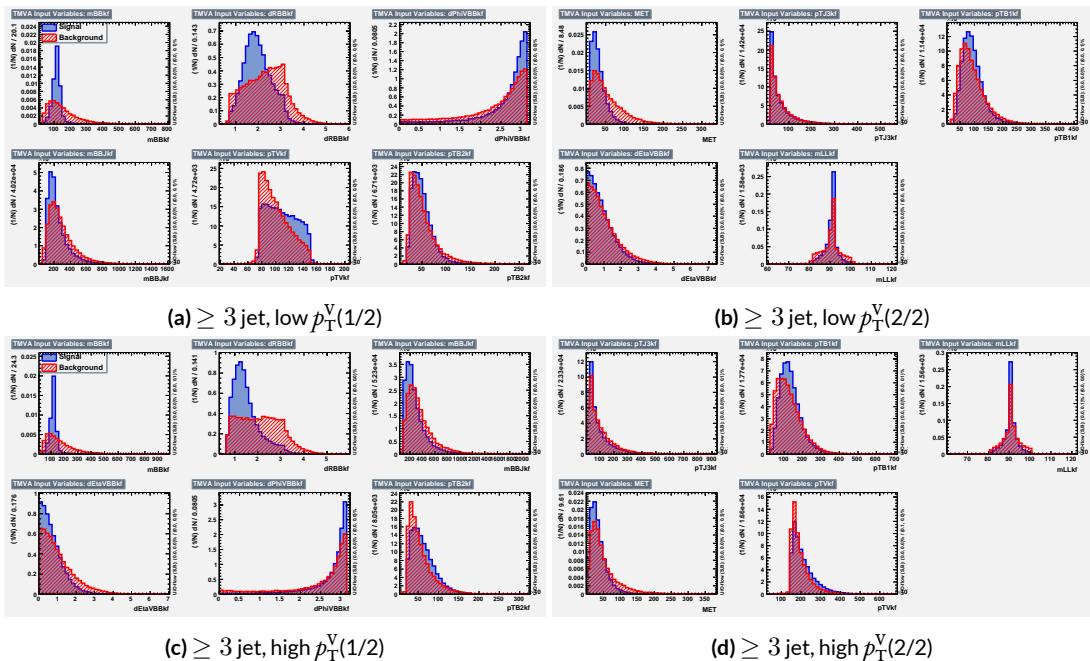


Figure 6.2: Input variables in ≥ 3 jet signal regions for the standard variable set. Signal distributions are in red, and background distributions are in blue.

1496 for refinements of the non-standard variables.

1497 Looking at the correlation matrices for the standard variables in Figure 6.3, it is easy to see that

1498 there are large number of non-trivial correlations (when looking at these correlation matrices, keep

1499 in mind that purple is perfect anti-correlation, red is perfect correlation (as with all variables with

themselves—hence diagonals of red), and green is roughly zero correlation).

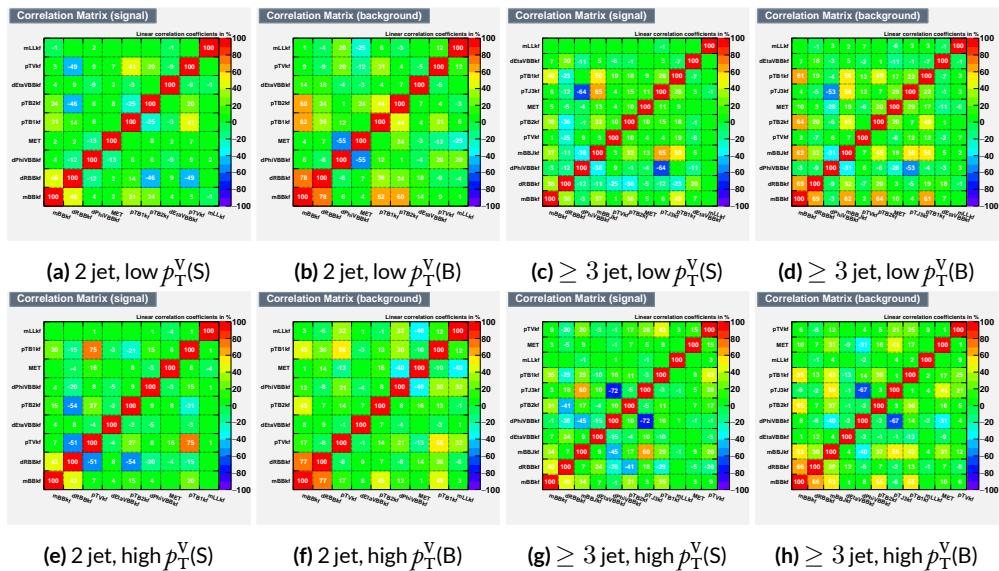


Figure 6.3: Signal and background variable correlations for the standard variable set.

1500

1501 6.1.2 LORENTZ INVARIANTS

1502 In choosing the set of variables used for a set of Lorentz Invariants based discriminants, we decided
 1503 to use S. Hagebeck's set from [7] and related studies. Distributions of these variables in the same
 arrangement as with the standard variables may be seen in Figures 6.4 and 6.5. One thing to note

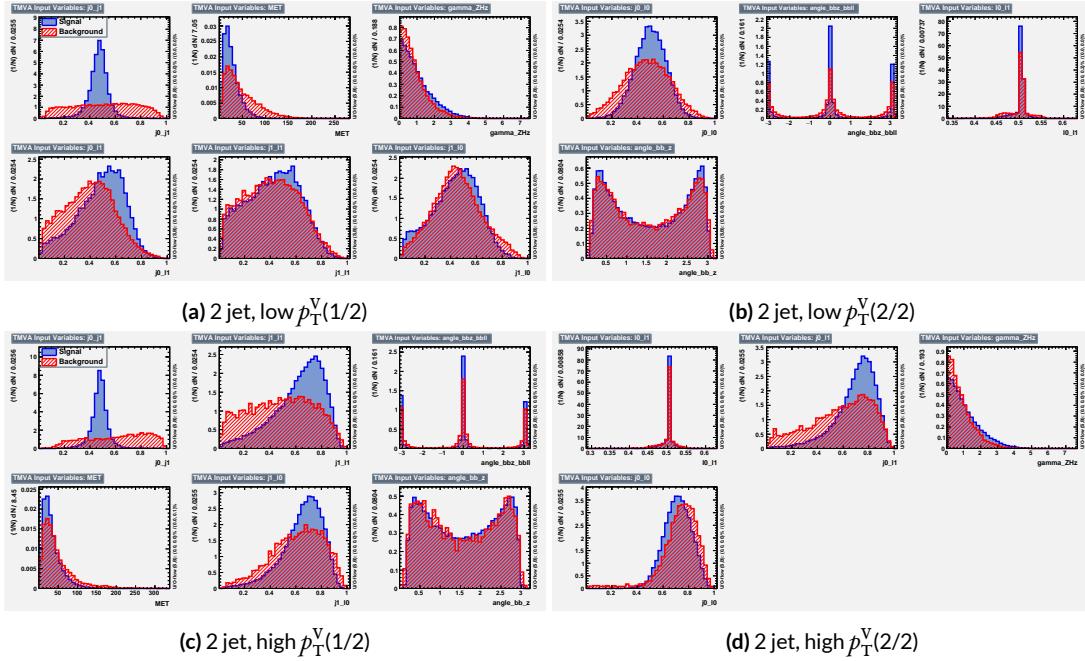


Figure 6.4: Input variables in 2 jet signal regions for the LI variable set. Signal distributions are in red, and background distributions are in blue.

1504

1505 about the variable set chosen here is that \vec{E}_T^{miss} has been added to the standard LI set. Since the LI
 1506 construction assumes that this quantity is zero, there is no obvious way to include it. Nevertheless,
 1507 as the correlation matrices for the LI variables show in Figure 6.6, there is actually very little correla-
 1508 tion between \vec{E}_T^{miss} and the other variables (with this being slightly less the case for the background
 1509 correlations, as to be expected since $t\bar{t}$, a principal background, is \vec{E}_T^{miss} -rich). Hence, if including

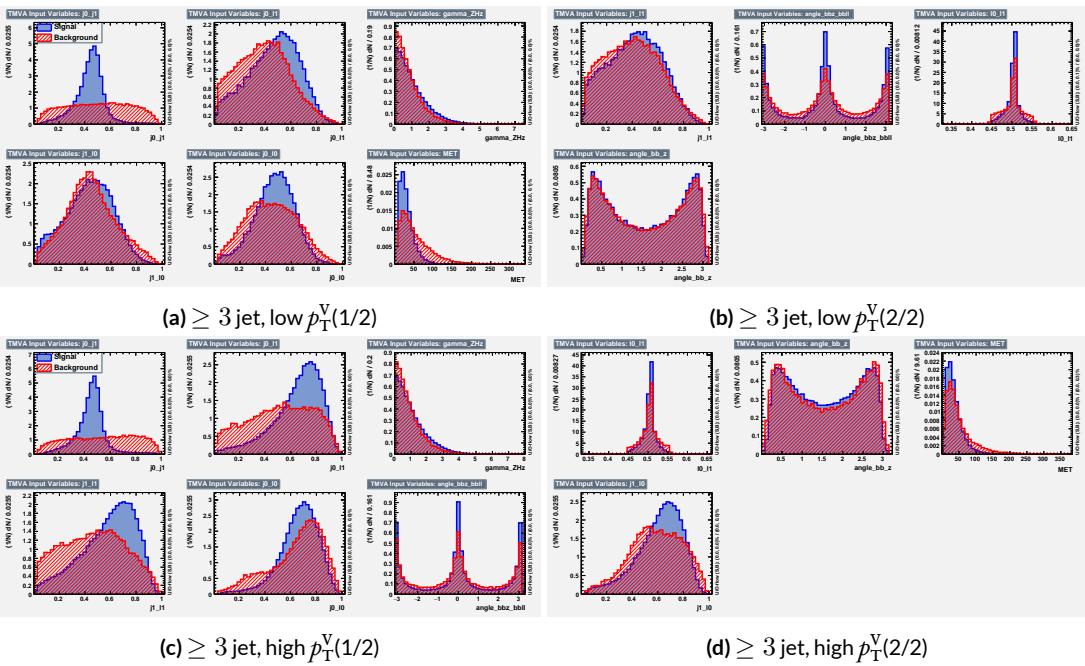


Figure 6.5: Input variables in ≥ 3 jet signal regions for the LI variable set. Signal distributions are in red, and background distributions are in blue.

1510 \vec{E}_T^{miss} violates the spirit somewhat of the LI variables, it does not break terribly much with the aim of
having a more orthogonal set.

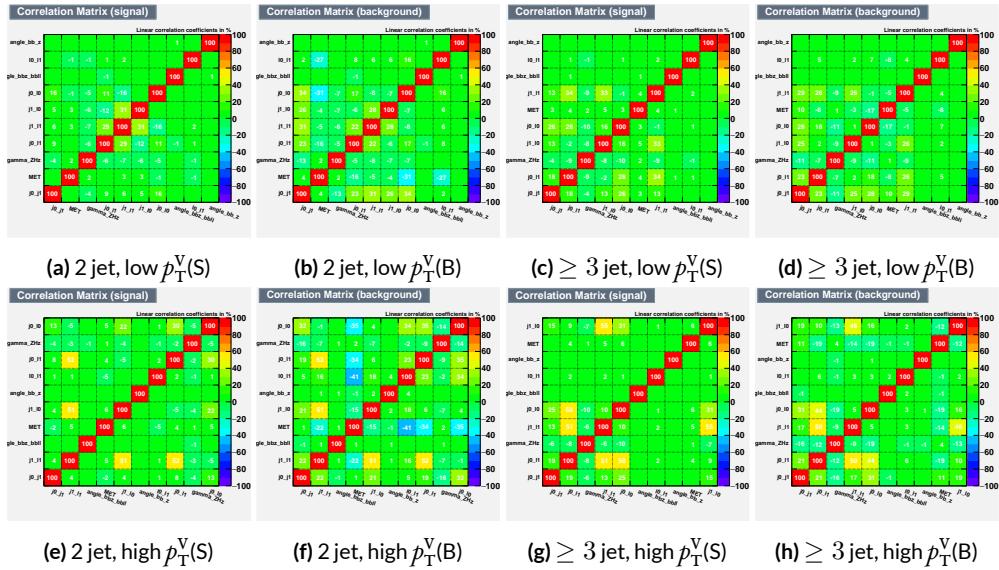


Figure 6.6: Signal and background variable correlations for the LI variable set.

1512 6.1.3 RESTFRAMES VARIABLES

1513 There is no precedent for using the RestFrames variables in the $ZH \rightarrow \ell\ell b\bar{b}$ analysis, so a subset
1514 of possible RF variables had to be selected as the basis of a discriminant. The masses and cosines of
1515 boost angles from parent frames for the CM, Z , and H frames gives six variables, and it was decided
1516 that it would be good to match the LI in terms of variable number and treatment (i.e. no special
1517 treatment of the third jet), which leaves four more variables. In addition to the cosines, there are
1518 also the $\Delta\phi$ angles. Furthermore, there are the event-by-event scaled momentum ratios, both lon-
1519 gitudinal and transverse. There is also both a $\Delta\phi$ and an CM-scaled ratio for the \vec{E}_T^{miss} . All of these
1520 variables were included in a ranking using slightly different training settings as the main hyperpa-
1521 rameter optimization variable ranking described below. The goal of this study was not to develop a
1522 discriminant, as the number of variables is too high, but rather to see which ones are generally use-
1523 ful. Table 6.1 shows the results of this study. Percent gains (losses) at each step by adding the variable
1524 with biggest gain (smallest loss) are shown in green (red). The final row shows an aggregate rank-
1525 ing, calculated simply by adding up a variables ranks in all bins and ordering the variables smallest
1526 to greatest. This simple aggregation does not take into account which regions are potentially more
1527 sensitive and so where taken simply to give an idea of how variables generally performed. With this
1528 in mind, the RF variables were chosen to be the masses M_{CM} , M_H , and M_Z , the angles $\cos CM$, \cosh ,
1529 $\cos Z$, $\cos \phi CMH$, and the ratios R_{pt} , R_{pz} , and R_{met} . Their distributions may be seen in Figures 6.7
1530 and 6.8.

1531 Correlations for the chosen RF variables are shown in Figure 6.9. These correlations are much

Region	Variable Chain
2jet pTVbin1	Rpt (65.8%), Rpz (29.0%), cosZ (11.4%), MZ (-1.75%), dphiCMH (7.26%), cosCM (3.95%), cosH (0.142%), MCM (2.18%), dphiCMZ (-2.3%), dphiCMMet (-0.236%), dphiLABCM (0.404%), Rmet (-4.04%)
3jet pTVbin1	Rpt (50.8%), Rpz (15.6%), MZ (14.8%), cosZ (3.08%), MCM (3.79%), dphiCMH (3.24%), cosH (0.755%), dphiCMMet (1.04%), Rmet (-1.03%), cosCM (5.31%), dphiCMZ (-1.27%), dphiLABCM (-2.88%), pTJ3 (-1.27%)
2jet pTVbin2	Rpt (52.0%), Rpz (13.8%), cosZ (16.9%), cosH (6.49%), MCM (1.71%), cosCM (6.21%), Rmet (4.25%), dphiCMMet (-1.53%), dphiLABCM (-0.757%), dphiCMH (0.213%), MZ (-0.788%), dphiCMZ (-2.39%)
3jet pTVbin2	Rpt (31.5%), Rpz (21.6%), cosH (8.97%), cosZ (1.42%), cosCM (11.3%), dphiCMZ (-2.84%), MCM (8.17%), dphiCMH (-0.841%), dphiLABCM (-0.00318%), dphiCMMet (-2.6%), pTJ3 (-3.21%), MZ (-1.8%), Rmet (-6.29%)
Aggregate	Rpt (o,o,o,o), Rpz (i,i,i,i), cosZ (2,3,2,3), cosH (6,6,3,2), MCM (7,4,4,6), MZ (3,2,10,11), dphiCMH (4,5,9,7), cosCM (5,9,5,4), dphiCMMet (9,7,7,9), dphiCMZ (8,10,11,5), Rmet (11,8,6,12), dphiLABCM (10,11,8,8)

Table 6.1: Full RF variable ranking study summary. Green (red) percentages represent gains (losses) in a validation significance at each step.

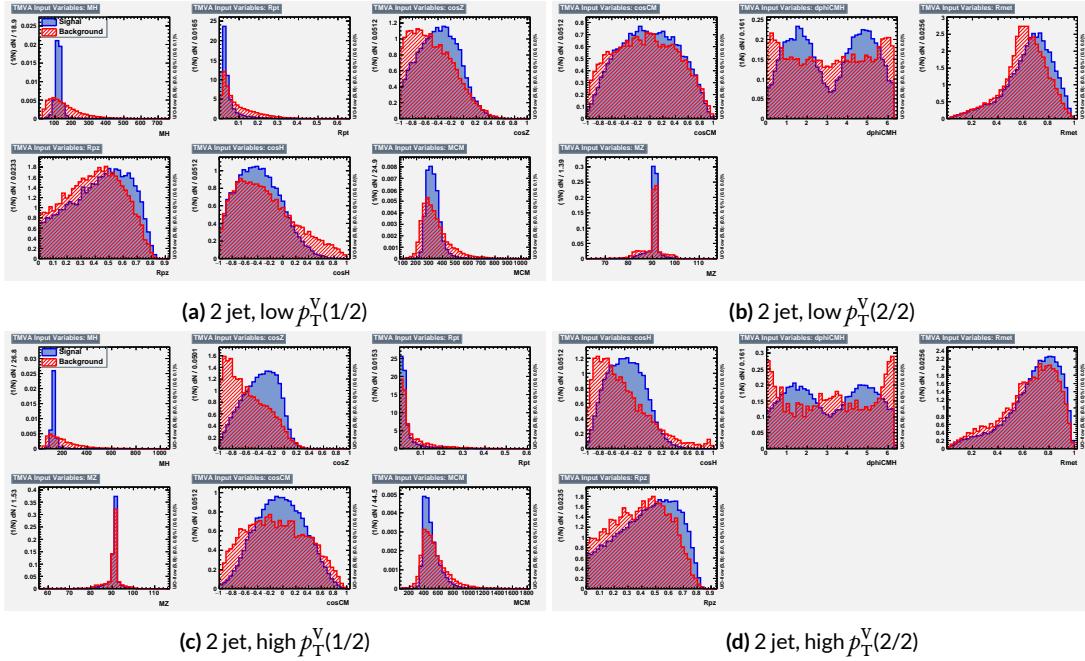


Figure 6.7: Input variables in 2 jet signal regions for the RF variable set. Signal distributions are in red, and background distributions are in blue.

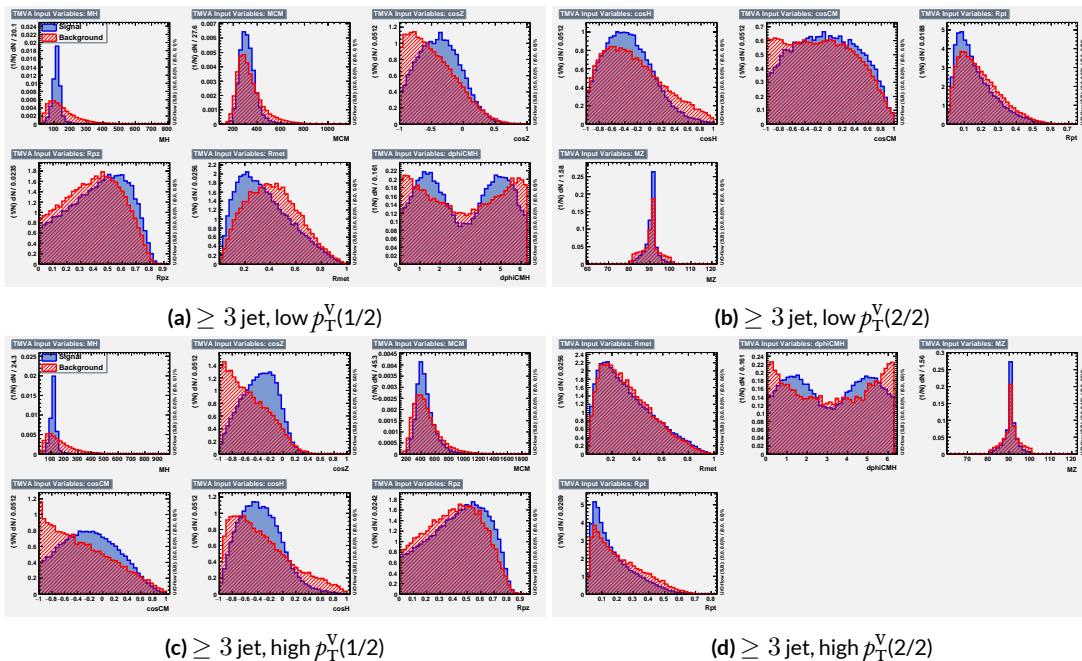


Figure 6.8: Input variables in ≥ 3 jet signal regions for the RF variable set. Signal distributions are in red, and background distributions are in blue.

lower than for the standard case but still slightly higher than for the LI case. Notably, many strong correlations that exist for signal events do not exist in background events and vice versa, so what is lost in orthogonality may very well be recuperated in greater separation[†]. Given the generally better performance of the RF sets, as we shall see in following sections and chapters, this slight tradeoff is likely an aesthetic one, with the main benefits of a more orthogonal basis likely realized at this level of correlation.

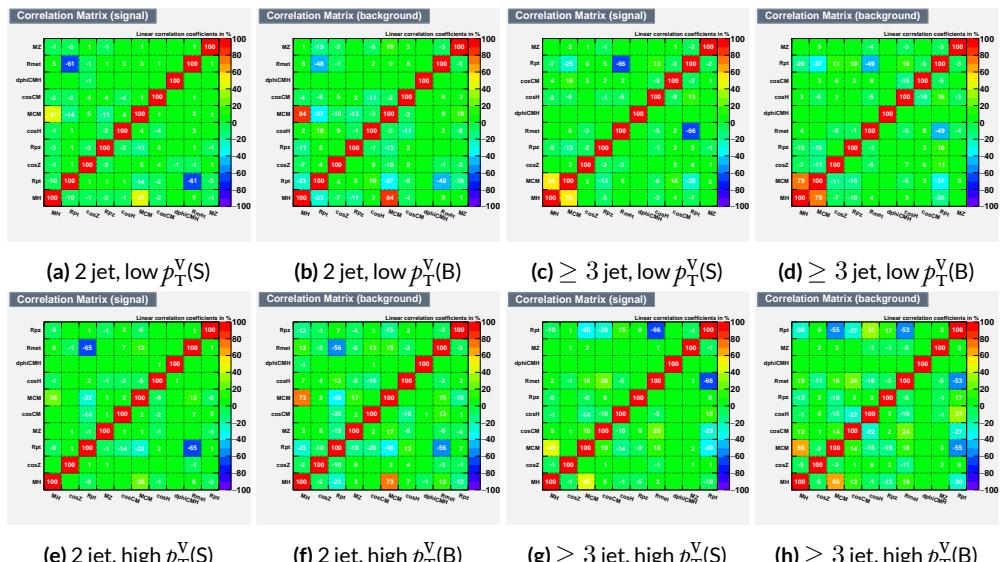


Figure 6.9: Signal and background variable correlations for the RF variable set.

A summary of the variables used in the three cases is given in 6.2.

[†]It is very hard to say for certain whether this is the case for MVA discriminants, and such dedicated studies might make worthwhile future studies.

Variable Set	Variables
Standard	mBB, mLL, (mBBJ), pTV, pTB1, pTB2, (pTJ3), dRBB, dPhiVBB, dEtaVBB, MET 9(11) vars
Lorentz Invariants	j0_j1, j0_l1, l0_l1, j1_l1, j0_l0, j1_l0, gamma_ZHz, angle_bbz_bbll angle_bb_z, MET 10 vars
RestFrames	MH, MCM, MZ, cosH, cosCM, cosZ, Rpz, Rpt, dphiCMH, Rmet 10 vars

Table 6.2: Variables used in MVA training. Variables in parentheses are only used in the ≥ 3 jet regions.

1539 6.2 MVA TRAINING

1540 With variables chosen, the MVA discriminants must be trained and optimized. MVA training and
 1541 hyperparameter optimization (in this case, just the order in which variables are fed into the MVA) is
 1542 conducted using the “holdout” method. In this scheme, events are divided into three equal portions
 1543 (in this case using `EventNumber%3`), with the first third (the “training” set) being used for the initial
 1544 training, the second third (the “validation” set) being used for hyperparameter optimization, and
 1545 the final third (the “testing” set) used to evaluate the performance of the final discriminants in each
 1546 analysis region.

1547 The MVA discriminant used is a boosted decision tree (BDT). Training is done in TMVA using
 1548 the training settings of the fiducial analysis [4][‡]. For the purposes of hyperparameterization and test-
 1549 ing, transformation D with $z_s = z_b = 10$ is applied to the BDT distributions, and the cumulative
 1550 sum of the significance $S/\sqrt{S + B}$ in each bin is calculated for each pair of distributions.

1551 Transformation D is a histogram transformation, developed during the Run 1 SM $VH(b\bar{b})$ search,

[‡]Namely, !H:!V:BoostType=AdaBoost:AdaBoostBeta=0.15:SeparationType=GiniIndex:-PruneMethod=NoPruning:NTrees=200:MaxDepth=4:nCuts=100:nEventsMin=5%

1552 designed to reduce the number of bins in final BDT distributions and thereby mitigate the effect of
 1553 statistical fluctuations in data while also maintaining sensitivity. Such an arbitrary transformation
 1554 may be expressed as:

$$Z(I[k, l]) = Z(z_s, n_s(I[k, l]), N_s, z_b, n_b(I[k, l]), N_b) \quad (6.1)$$

1555 where

- 1556 • $I[k, l]$ is an interval of the histograms, containing the bins between bin k and bin l ;
- 1557 • N_s is the total number of signal events in the histogram;
- 1558 • N_b is the total number of background events in the histogram;
- 1559 • $n_s(I[k, l])$ is the total number of signal events in the interval $I[k, l]$;
- 1560 • $n_b(I[k, l])$ is the total number of background events in the interval $I[k, l]$;
- 1561 • z_s and z_b are parameters used to tune the algorithm.

1562 Transformation D uses:

$$Z = z_s \frac{n_s}{N_s} + z_b \frac{n_b}{N_b} \quad (6.2)$$

1563 Rebinning occurs as follow:

- 1564 1. Begin with the highest valued bin in the original pair of distributions. Call this the “last” bin
1565 and use it as l , and have k be this bin as well.
- 1566 2. Calculate $Z(I[k, l])$
- 1567 3. If $Z \leq 1$, set $k \rightarrow k - 1$ and return to step 2. If not, rebin bins $k-l$ into a single bin and name
1568 $k - 1$ the new “last” bin l .
- 1569 4. Continue until all bins have been iterated through; if $Z \leq 1$ for any remaining n of the
1570 lowest-valued bins (as is often the case), simply rebin these as a single bin.

1571 Variable ranking is done iteratively (greedily) in each analysis region. In each set, the validation
1572 significance of a BDT using an initial subset of variables is calculated (`dRBB` and `mBB` for the standard
1573 set; `j0_j1` for the LI set; and `MH` for the RF set). Each of the remaining unranked variables are then
1574 added separately, one at a time, to the BDT. The variable yielding the highest validation significance
1575 is then added to the set list of ranked variables and removed from the list of unranked variables. This
1576 process is repeated until no variables remain.

1577 This training process is visualized using “ranking plots.” A prototypical example is given in Fig-
1578 ure 6.10, the ranking for the RF variable set in the 2 jet, high p_T^V region. Each step is shown in a dis-
1579 tinct color. In this example, the first step begins by training nine BDT’s with the variable `MH` and one
1580 of the other variables in this region. Graphically, we arrange the results ascending left to right based
1581 on their validation $S/\sqrt{S+B}$. At this step, this yields `cosh`, `cosCM`, `Rpz`, `MZ`, `dphiCMH`, `Rmet`, `MCM`,
1582 `Rpt`, and the best performing `cosZ`. Hence, this step is labeled “2 vars (+`cosZ`)” to denote that the
1583 two variable `MH-cosZ` BDT is used as the basis for the three variable step. Note that while each case
1584 is arranged left to right, each of these iteration is in practice computed in parallel. When examining a
1585 plot of this type, even though there are no errors given, one can use the smoothness of overall curve
1586 as an indicator for ranking stability (sometimes low statistics for certain samples can make jittery
1587 rankings inevitable, particularly near the end of rankings).

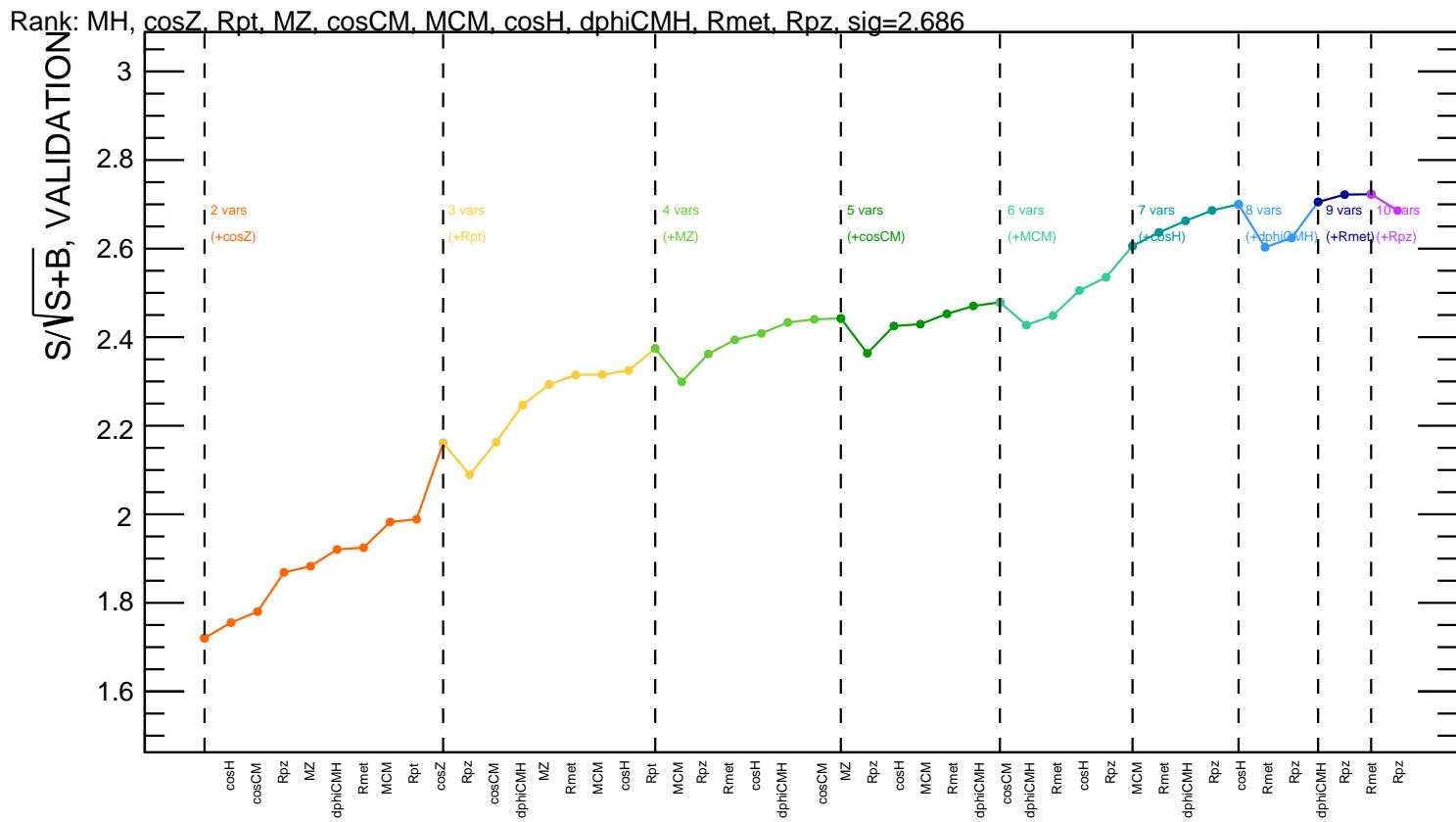


Figure 6.10: An example ranking, the RF variable set in the 2 jet, high p_T^V region.

1588

These rankings are shown in Figures 6.11–6.13. Rankings tend to be fairly stable.

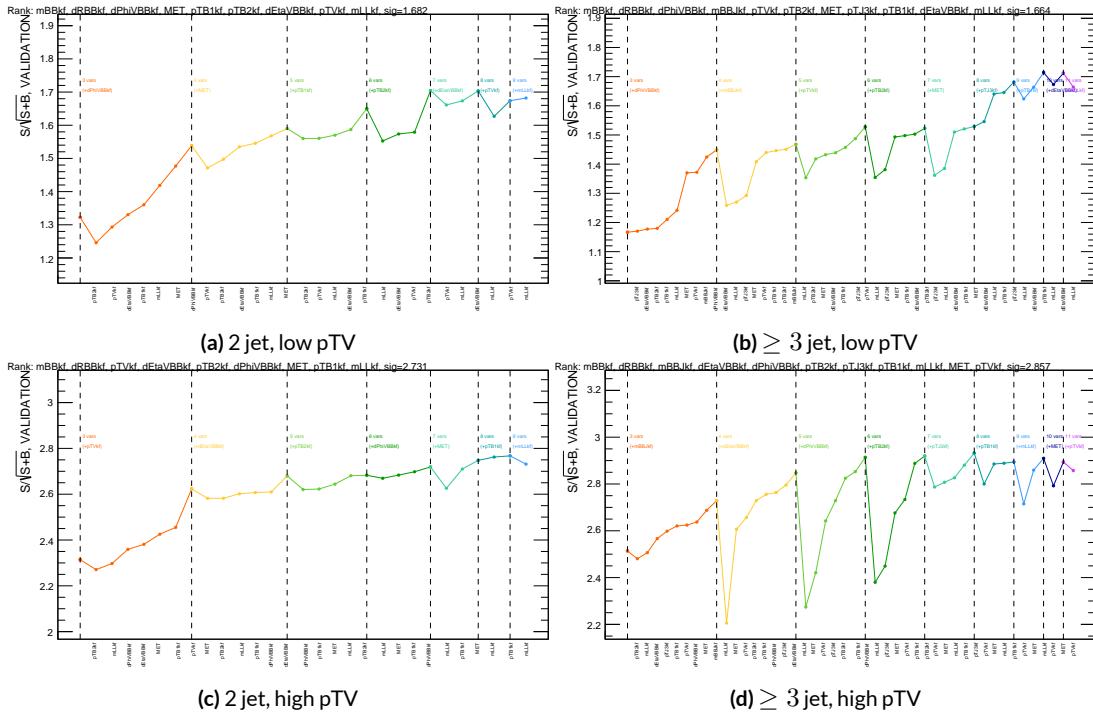


Figure 6.11: Rankings for the standard variable set.

1589

Once variables have been ranked, the BDT may be used both to evaluate performance in a simplified analysis scenario in the absence of systematic uncertainties (described below in Section 6.3) and

1590

to create xml files for the production of fit inputs for an analysis including systematics. Following

1591

the approach taken in the fiducial analysis, BDT discriminants using two “k-folds” are produced to

1592

prevent overtraining, since the samples used for training are the same as those used to produce in-

1593

puts for the full profile likelihood fit. In this scheme, a BDT trained on events with an even (odd)

1594

`EventNumber` are used to evaluate events with an odd (even) `EventNumber`.

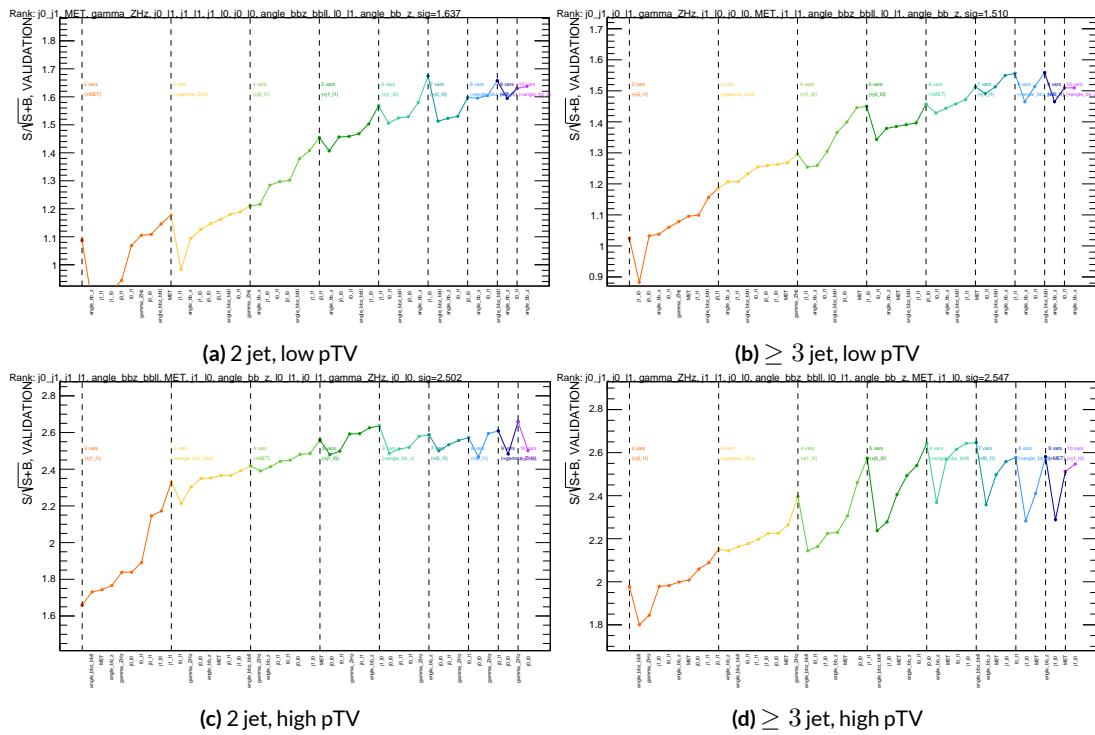


Figure 6.12: Rankings for the L1 variable set.

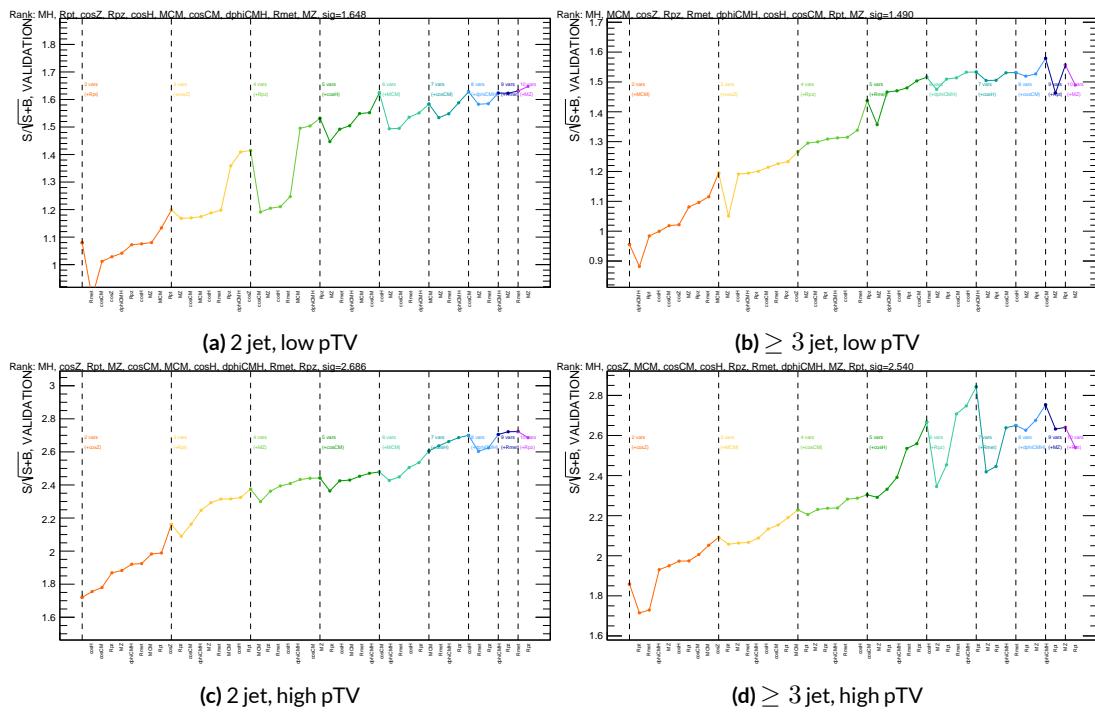


Figure 6.13: Rankings for the RF variable set.

1596 6.3 STATISTICS ONLY BDT PERFORMANCE

1597 As described above, cumulative significances can be extracted from pairs of signal and background
1598 BDT output distributions in a given region. In order to evaluate performance of variable sets in the
1599 absence of systematic uncertainties, such pairs can be constructed by evaluating BDT score on the
1600 testing set of events using the optimal variable rankings in each region. We show two versions of
1601 each testing distribution for each variable set in each signal region in Figures 6.15–6.17. The training
1602 distribution is always shown as points. The plots with block histograms with numbers of bins that
1603 match (do not match) the training distribution do not (do) have transformation D applied. Trans-
1604 formation D histograms are included to show the distributions actually used for significance eval-
1605 uation, while the untransformed histograms are included to illustrate that the level of overtraining is
1606 not too terrible[§]. For better comparison of the distributions, all histograms have been scaled to have
1607 the same normalization.

1608 An example of training/(untransformed) testing distributions is given in 6.14. Visually, one ex-
1609 pects the training and testing distributions to look similar, but not too similar (an indication of over-
1610 training). This qualitative intuition can be quantified via the two sample Kolmogorov-Smirnov test,
1611 which uses the maximum difference in distribution CDF's to obtain an “overtraining probability”
1612 given by:

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)| = \sqrt{\frac{n+m}{nm} \times \left(-\frac{1}{2} \ln\left(\frac{\alpha}{2}\right)\right)} \quad (6.3)$$

[§]The raw distributions include a K-S test statistic for signal (background) distributions.

₁₆₁₃ In the $n = m$ case, $\alpha = 2e^{-nD_n^2}$, which is the usual null hypothesis (same distribution) probabil-

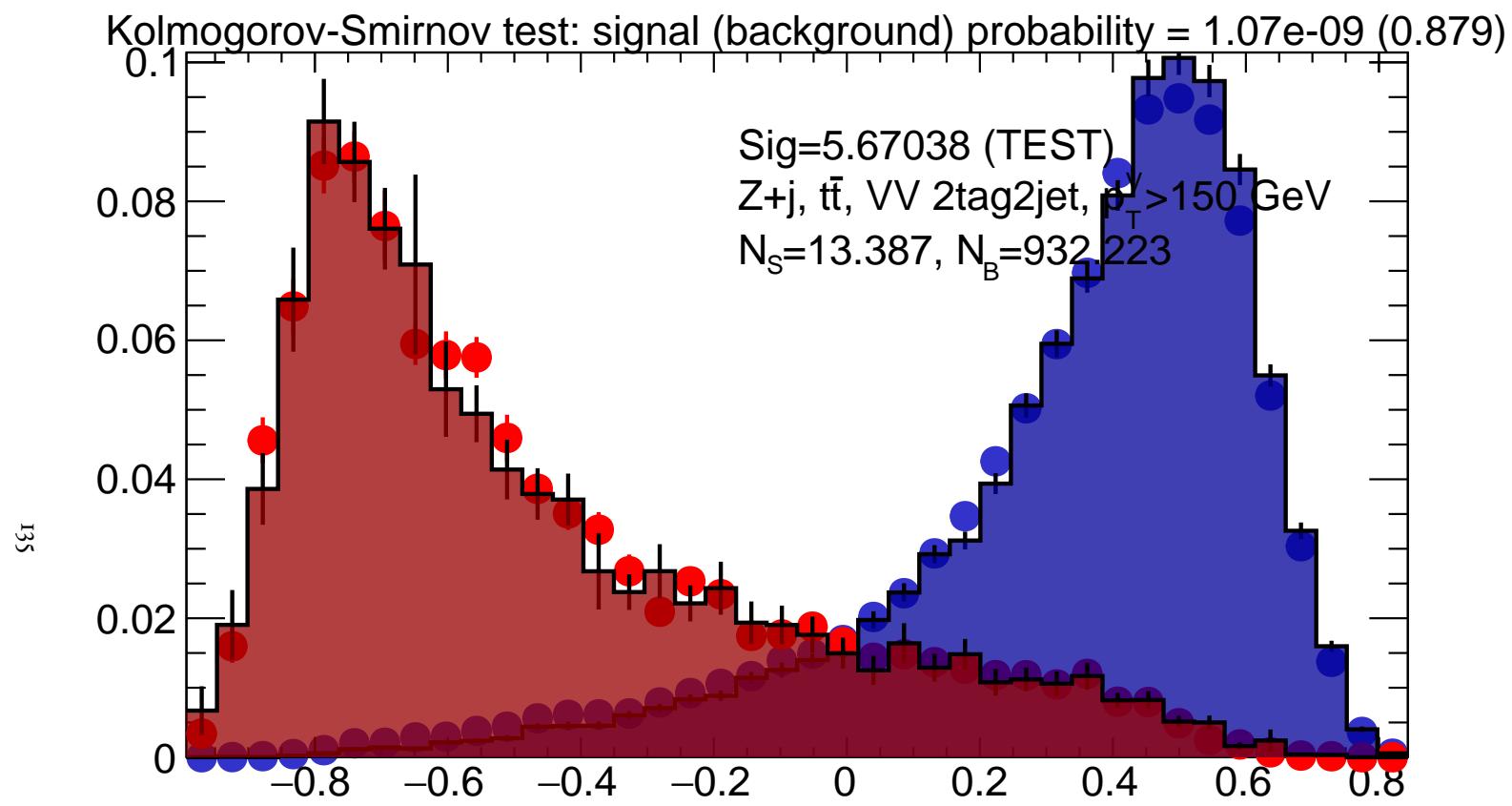
₁₆₁₄ ity. Generally, the background probabilities are higher because the background distributions rely on

₁₆₁₅ some relatively low statistics MC samples (e.g. VV , which has low yields with analysis event selection

₁₆₁₆ requirements). The signal distribution, on the other hand, only relies on high statistics signal MC

₁₆₁₇ samples (with most events passing selection requirements and making it into the discriminants by

₁₆₁₈ design).



1619 As can be seen in the summary of cumulative significances for each of these analysis regions and
1620 variable sets in Figure 6.18, the performance of each of the variable sets is quite similar. The standard
1621 set performs best, with the LI (RF) set having a cumulative significance that is 7.9% (6.9%) lower.
1622 This suggests that the LI and RF variables, in the $ZH \rightarrow \ell\ell b\bar{b}$ closed final state, have no more in-
1623 trinsic descriptive power than the standard set. That these figures are all relatively high (~ 4.5) is
1624 due largely to the absence of systematics and possibly in part due to the fact that many of the most
1625 significant bins occur at high values of the BDT output, which, as can be seen in any of the testing
1626 distributions, contain a small fraction of background events. An interesting feature to note in Fig-
1627 ure 6.18 is that while the standard set does perform better in all regions, the gap is larger in the ≥ 3
1628 jet regions, suggesting that further optimization in the ≥ 3 jet case could be useful. Moreover, as
1629 discussed at the end of Chapter 5, the choice of ≥ 3 jet and not exclusive 3 jet regions is a 2-lepton
1630 specific choice and may not be justified for the non-standard variable sets.

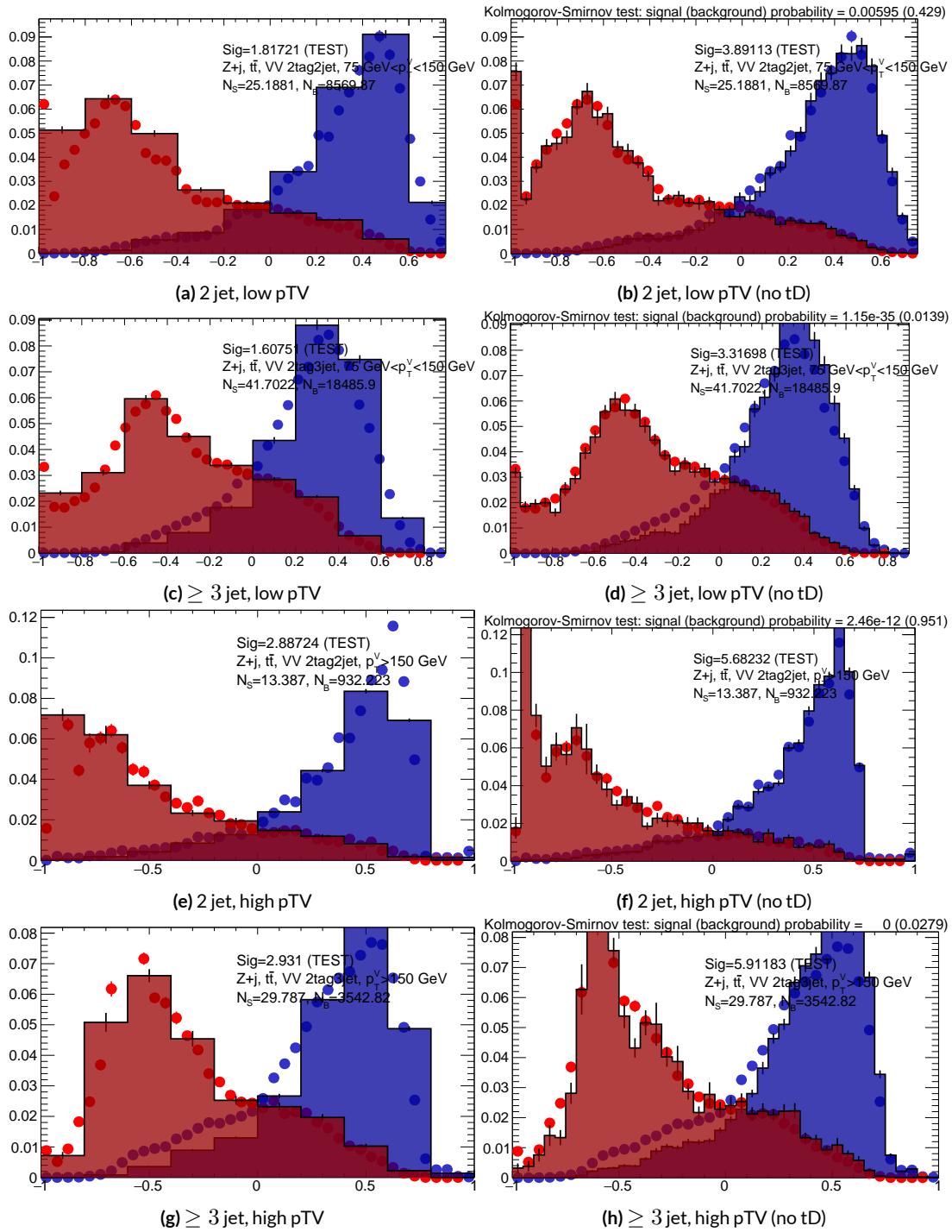


Figure 6.15: Training (points) and testing (block histogram) MVA distributions used for stat only testing for the standard variable set.

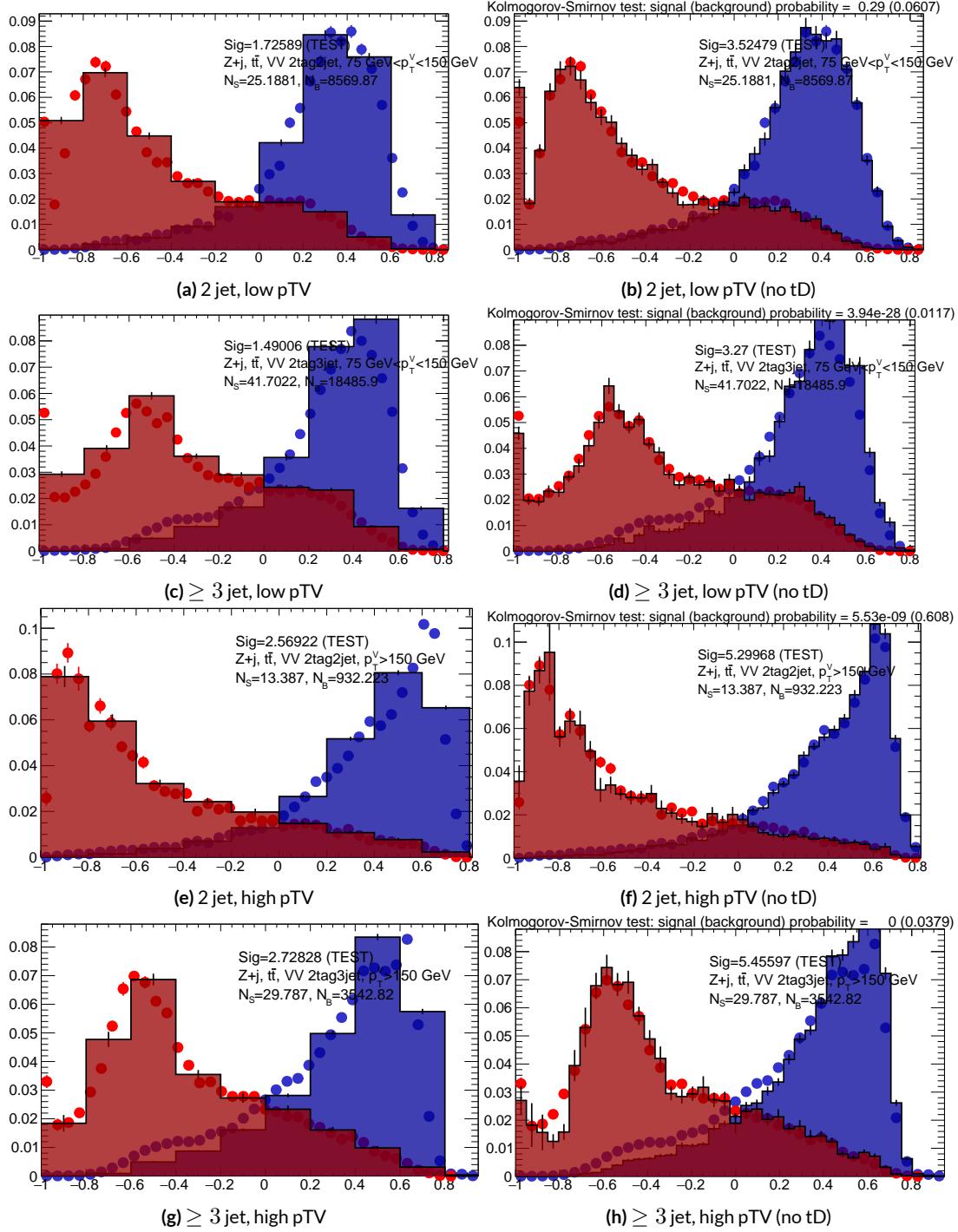


Figure 6.16: Training (points) and testing (block histogram) MVA distributions used for stat only testing for the LI variable set.

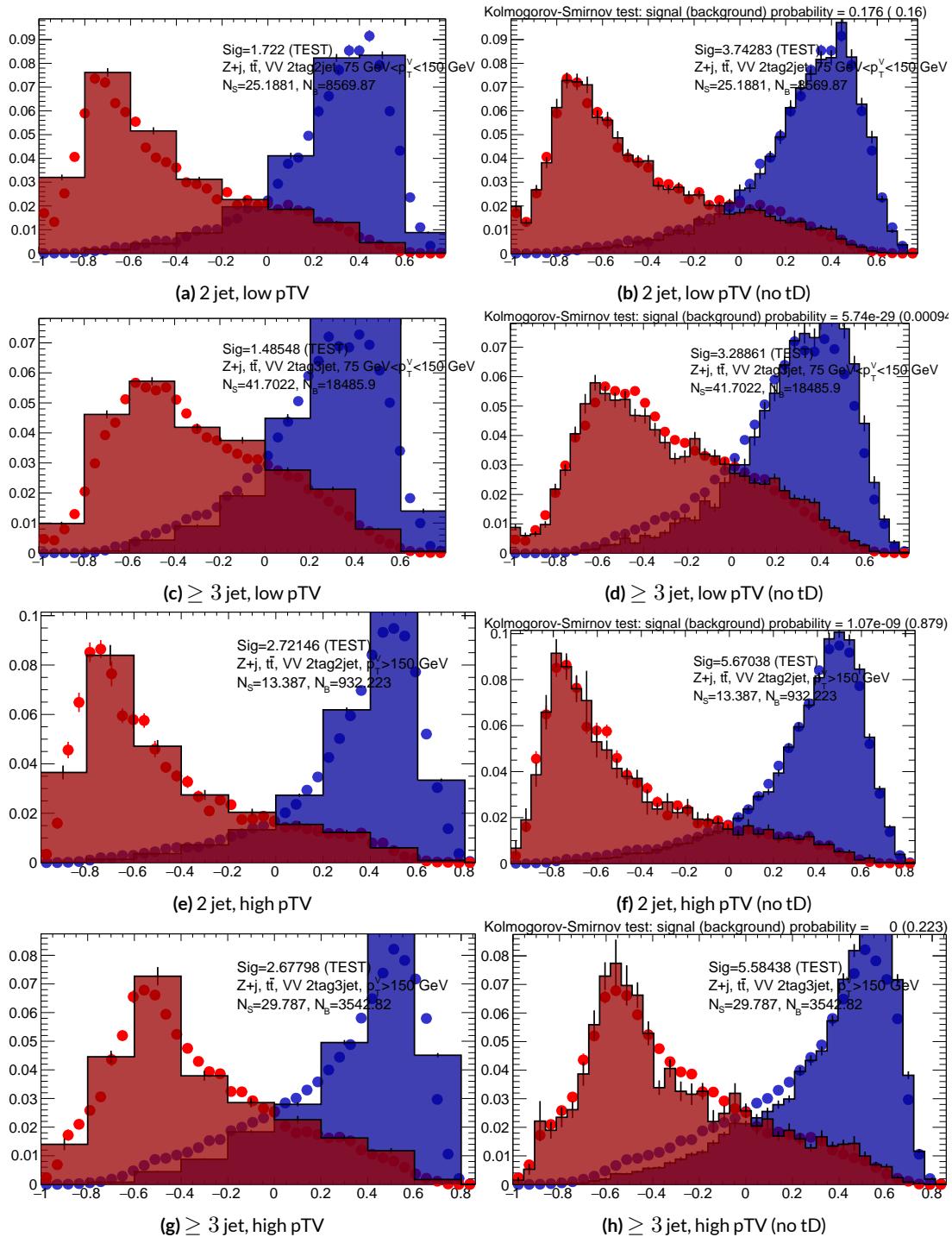


Figure 6.17: Training (points) and testing (block histogram) MVA distributions used for stat only testing for the RF variable set.

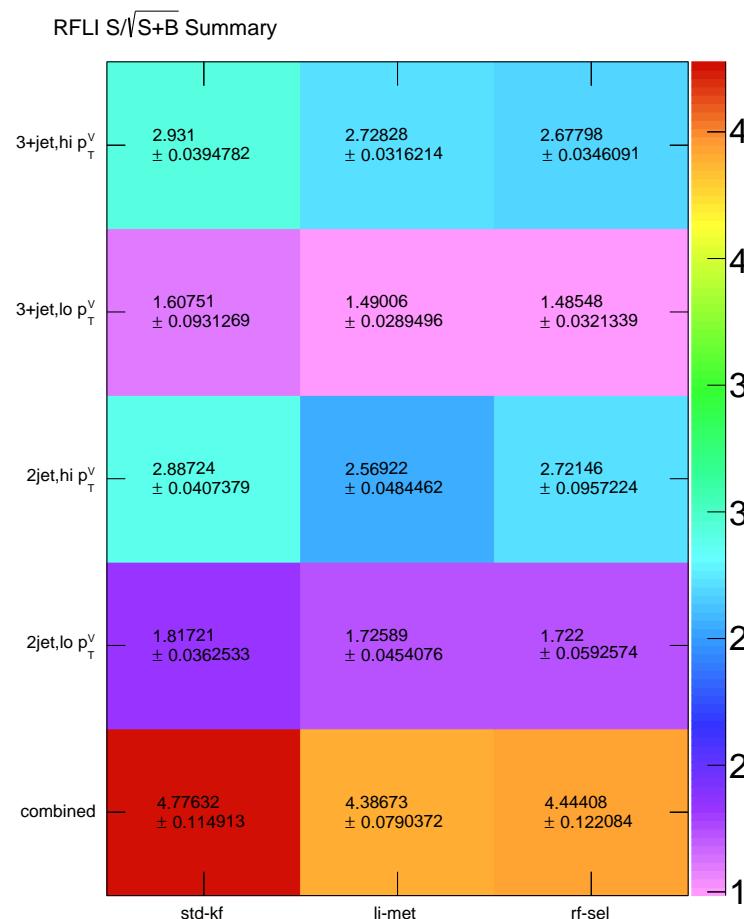


Figure 6.18: Results of testing significances sorted by analysis region and variable set.

Multivac picked you as most representative this year.

Not the smartest, or the strongest, or the luckiest, but

just the most representative. Now we don't question

Multivac, do we?

Isaac Asimov, "Franchise"

1631

7

1632

Statistical Fit Model and Validation

1633 THE ULTIMATE GOAL of an analysis like the search for SM $VH(b\bar{b})$ decay is to say with as much
1634 justified precision as possible with the ATLAS collision data whether or not the SM-like Higgs ob-
1635 served in other decay modes also decays to b -quarks and, if so, whether this rate is consistent with
1636 the SM prediction. In the limit of perfect modeling of both background processes and detector/reconstruction,

1637 the only free parameter is this production rate, referred to typically as a “signal strength,” denoted μ ,
1638 with $\mu = 1$ corresponding to the SM prediction and $\mu = 0$ corresponding to the SM with no
1639 Higgs.

1640 To get a better sense of what this might look like, take a look at the example discriminant distri-
1641 bution in Figure 7.1. The black points are data (with statistical error bars), and the colored block
1642 histograms have size corresponding to the number of predicted events for each process in each bin of
1643 the final BDT. In the limit of perfect understanding, a fit would correspond to a constant scale fac-
1644 tor on the red, signal histogram, where one would choose a best fit μ value, denoted $\hat{\mu}$, that would
1645 minimize the sum in quadrature of differences between the number of observed data events and
1646 $\mu s_i + b_i$, where s_i and b_i are the predicted number of signal and background events in each bin.

1647 The only source of uncertainty would be due to data statistics, so for an infinitely large dataset with
1648 perfect understanding, μ could be fitted to arbitrary precision. This, of course, is not the case since
1649 there is a finite amount of data and very many sources of systematic uncertainty, discussed in pre-
1650 vious chapters. This chapter will first describe how systematic uncertainties are integrated into the
1651 statistical fit of this analysis before describing two sets of cross checks on both a validation VZ fit and
1652 on the fit for the VH fit of interest.

1653 7.1 THE FIT MODEL

1654 In order to derive the strength of the signal process $ZH \rightarrow \ell\ell b\bar{b}$ and other quantities of interest
1655 while taking into account systematic uncertainties or nuisance parameters (NP’s, collectively de-
1656 noted θ), a binned likelihood function is constructed as the product over bins of Poisson distribu-

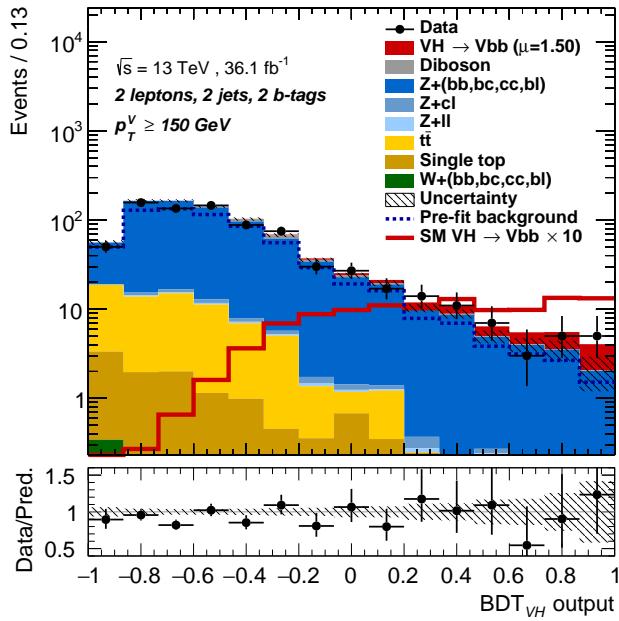


Figure 7.1: An example postfit distribution. The reason this looks different from postfit distributions later in this chapter is that this is a log plot. It shows rare samples ($Z+c\bar{l}$, $Z+l\bar{l}$, $W+jets$), but obscures Data/MC agreement, which can only be seen via the ratio plot.

1657 tions:

$$\mathcal{L}(\mu, \theta) = \text{Pois}(n | \mu S + B) \left[\prod_{i \in \text{bins}} \frac{\mu s_i + b_i}{\mu S + B} \right] \prod_{j \in \text{NP's}} \mathcal{N}_{\theta_j}(\theta_j, \sigma_j^2 | 0, 1) \quad (7.1)$$

1658 where n is the total number of events observed, s_i and b_i are the number of expected signal and back-
 1659 ground events in each bin, and S and B are the total expected signal and background events. The
 1660 signal and background expectations generally are functions of the NP's θ . NP's related to the nor-
 1661 malization of signal and background processes fall into two categories. The first set is left to float
 1662 freely like μ while the second set are parametrized as log-normally distributed to prevent negative
 1663 predicted values. All other NP's are parametrized with Gaussian priors. This results in a “penalty”
 1664 on the NLL discussed below of $(\hat{\alpha} - \mu_\alpha)^2 / \sigma_\alpha^2$, for NP α , normally parametrized with mean μ_α

1665 (corresponding to the nominal prediction) and variance σ_α^2 (derived as discussed in Chapters 4 and
1666 5) for an MLE of $\hat{\alpha}$.

1667 One can maximize^{*} the likelihood in Equation 7.1 for a fixed value of μ to derive estimators for
1668 the NP's θ ; values of θ so derived are denoted $\hat{\theta}_\mu$ to emphasize that these are likelihood maximizing
1669 for a given μ . The profile likelihood technique finds the likelihood function's maximum by compar-
1670 ing the values of the likelihood over all possible values of μ using these "profiles" and picking the
1671 one with the greatest $\mathcal{L}(\mu, \hat{\theta}_\mu)$ value; these values of μ and θ are denoted $\hat{\mu}$ and $\hat{\theta}$. The profile like-
1672 lihood can further be used to construct a test statistic[†]

$$q_\mu = -2 \left(\log \mathcal{L}(\mu, \hat{\theta}_\mu) - \log \mathcal{L}(\hat{\mu}, \hat{\theta}) \right) \quad (7.2)$$

1673 This statistic can be used to derive the usual significance (p value), by setting $\mu = 0$ to find the com-
1674 patibility with the background-only hypothesis [58]. If there is insufficient evidence for the signal
1675 hypothesis, the CL_s method can be used to set limits [59].

1676 In order to both validate the fit model and study the behavior of fits independent of a given dataset,
1677 a so-called "Asimov"[‡] dataset can be constructed for a given fit model; this dataset has each bin equal
1678 to its expectation value for assumed values of the NP's and a given μ value (in this case, $\mu = 1$, the
1679 SM prediction).

^{*}Maximization is mathematically identical to finding the minimum of the negative logarithm of the likelihood, which is numerically an easier problem.

[†]The factor of -2 is added so that this statistic gives, in the asymptotic limit of large N , a χ^2 distribution.

[‡]A reference to the short story quoted at the beginning of this chapter in which a computer picks a single voter to stand in for the views of the entire American electorate.

1680 7.2 FIT INPUTS

1681 Inputs to the binned likelihood are distributions of the BDT outputs described in Chapter 6 for the
1682 signal regions and of m_{bb} for the top $e - \mu$ control regions. These regions split events according
1683 to their p_T^V and number of jets. All events are required to have two b -tagged jets, as well as pass the
1684 other event selection requirements summarized in Table 5.8; the only difference between the signal
1685 and control region selections is that the same flavor requirement (i.e. leptons both be electrons or
1686 muons) is flipped so that events in the control region have exactly one electron and one muon. The
1687 BDT outputs are binned using transformation D, while the m_{bb} distributions have 50 GeV bins,
1688 with the exception of the 2 jet, high p_T^V region, where a single bin is used due to low statistics.

1689 Input distributions in MC are further divided according to their physics process. The signal pro-
1690 cesses are divided based on both the identity of associated V and the number of leptons in the final
1691 state; $ZH \rightarrow \ell\ell b\bar{b}$ events are further separated into distributions for qq and gg initiated processes.
1692 $V+jets$ events are split according to V identity and into the jet flavor bins described in Chapter 3.
1693 Due to the effectiveness of the 2 b -tag requirement suppressing the presence of both c and l jets,
1694 truth-tagging is used to boost MC statistics in the cc , cl , and ll distributions.[§] For top backgrounds,
1695 single top production is split according to production mode (s , t , and Wt), with $t\bar{t}$ as single category.
1696 Diboson background distributions are also split according to the identity of the V 's (ZZ , WZ , and
1697 WW). Fit input segmentation is summarized in Table 7.1.

§Since WW is not an important contribution to the already small total diboson background, no truth-
tagging was applied here, in contrast to the fiducial analysis.

Category	Bins
# of Jets	2, 3+
p_T^V Regions (GeV)	$[75, 150], [150, \infty)$
Sample	data, signal $[(W, qqZ, ggZ)] \times n_{lep}$, $V+jets [(W, Z) \times (bb, bc, bl, cc, cl, ll)]$, $t\bar{t}$, diboson (ZZ, WW, WZ), single top (s, t, Wt)

Table 7.1: Fit input segmentation.

1698 7.3 SYSTEMATIC UNCERTAINTIES REVIEW

1699 Tables 7.2 and 7.3 summarize modeling (Chapter 4) and experimental (Chapter 5) systematic uncer-
 1700 tainties considered in this analysis, respectively. In addition to these, simulation statistics uncertain-
 1701 ties (“MC stat errors”) are also included in the fit model. There is one distribution per systematic
 1702 (one each for up and down) per sample per region. The $\pm 1\sigma$ variation for a systematic is calculated
 1703 as the difference in the integrals between the nominal and up/down varied distributions.

Process	Systematics
Signal	$H \rightarrow bb$ decay, QCD scale, PDF+ α_S scale, UE+PS (acc, p_T^V , m_{bb} , 3/2 jet ratio)
$Z+jets$	Acc, flavor composition, $p_T^V+m_{bb}$ shape
$t\bar{t}$	Acc, $p_T^V+m_{bb}$ shape
Single top	Acc, $p_T^V+m_{bb}$ shape
Diboson	Overall acc, UE+PS (acc, p_T^V , m_{bb} , 3/2 jet ratio), QCD scale (acc (2, 3 jet, jet veto), p_T^V , m_{bb})

Table 7.2: Summary of modeling systematic uncertainties.

1704 The systematics distributions undergo processes known as “smoothing” and “pruning” before
 1705 being combined into the final likelihood used in minimization.
 1706 The difference between systematics varied distributions and nominal distributions approaches

Process	Systematics
Jets	21 NP scheme for JES, JER as single NP
E_T^{miss}	trigger efficiency, track-based soft terms, scale uncertainty due to jet tracks
Flavor Tagging	Eigen parameter scheme (CDI File: 2016-20_7-13TeV-MC15-CDI-2017-06-07_v2)
Electrons	trigger eff, reco/ID eff, isolation eff, energy scale/resolution
Muons	trigger eff, reco/ID eff, isolation eff, track to vertex association, momentum resolution/scale
Event	total luminosity, pileup reweighting

Table 7.3: Summary of experimental systematic uncertainties.

1707 some stable value in the limit of large simulation statistics, but if the fluctuations due to simulation
 1708 statistics in a distribution are large compared to the actual physical effect (whether this is because
 1709 the actual effect is small or if the actual distribution is derived from a small number of simulation
 1710 events), then systematic uncertainty will be overestimated by, in effect, counting the MC stat error
 1711 multiple times. Smoothing is designed to mitigate these effects by merging adjacent bins in some
 1712 input distributions. Smoothing happens in two steps (the full details of smoothing algorithms may
 1713 be found in [4] and in the `WSMaker` code):

- 1714 1. Merge bins iteratively where bin differences are smallest in input distributions until no local
 1715 extrema remain (obviously, a single peak or valley is allowed to remain)
- 1716 2. Sequentially merge bins (highest to lowest, like transformation D) until the statistical uncer-
 1717 tainty in a given bin is smaller than 5% of merged bin content

1718 Not all systematic uncertainties defined are included in the final fit. Systematics are subject to
 1719 “pruning” (individually in each region/sample: there are two histograms per systematic (up/down)
 1720 per region per sample, so pruning just consists of removing the histograms from the set of distribu-
 1721 tions included in the likelihood) if they are do not have a significant impact, defined as follows:

- Normalization/acceptance systematics are pruned away if either:
 - The variation is less than 0.5%
 - Both up and down variations have the same sign

- Shape systematics pruned away if either:
 - Not one single bin has a deviation over 0.5% after the overall normalization is removed
 - If only the up or the down variation is non-zero

- Shape+Normalization systematics are pruned away if the associated sample is less than 2% of the total background and either:
 - If the predicted signal is < 2% of the total background in all bins and the shape and normalization error are each < 0.5% of the total background
 - If instead at least one bin has a signal contribution > 2% of the total background, and only in each of these bins, the shape and normalization error are each < 2% of the signal yield

7.4 THE VZ VALIDATION FIT

One of the primary validation cross-checks for the fiducial analysis was a *VZ* fit—that is, conducting the entire analysis but looking for $Z \rightarrow b\bar{b}$ decays instead of the Higgs. The idea here is that the Z is very well understood and so “rediscovering” Z decay to b 's is taken as a benchmark of analysis reliability since the complexity of the fit model precludes the use of orthogonal control regions for validation as is done in other analyses (generally, if there is a good control region, one prefers to use it

1741 to constrain backgrounds and improve the fit model). To do this, a new MVA discriminant is made
 1742 by keeping all hyperparameter configurations the same (e.g. variable ranking) but using diboson
 1743 samples as signal. For the 2-lepton case, this means using $ZZ \rightarrow \ell\ell b\bar{b}$ as the signal sample. This
 1744 new MVA is used to make the inputs described in Section 7.2, and the fit is then run as for the VH
 1745 fit (again, with ZZ as signal). VH samples are considered background in these diboson fits.

1746 The VZ fit sensitivities for the standard, LI, and RF fits are summarized in Table 7.4. The ex-
 1747 pected significances are all fairly comparable and about what was the case in the fiducial analysis.
 1748 The observed significance for the standard set matches fairly well with the expected value on data,
 1749 but the LI and RF observed significances are quite a bit lower.

	Standard	LI	RF
Expected (Asimov)	3.83	3.67	3.72
Expected (data)	3.00	2.95	3.11
Observed (data)	3.17	1.80	2.09

Table 7.4: Expected (for both data and Asimov) and observed $VZ \rightarrow \ell\ell b\bar{b}$ sensitivities for the standard, LI, and RF variable sets.

1750 These values, however, are consistent with the observed signal strength values, which can be seen
 1751 in Figure 7.2 (b), with both the LI and RF fits showing a deficit of signal events with respect to the
 1752 SM expectation, though not by much more than one standard deviation (a possible explanation is
 1753 explored in the following section). Just as in the VH fits, errors arising systematic uncertainties are
 1754 lower in the fits to the observed dataset. That the effect is not noticeable in Asimov fits is not too
 1755 surprising, since this analysis (and these variable configurations in particular), is not optimized for
 1756 VZ .

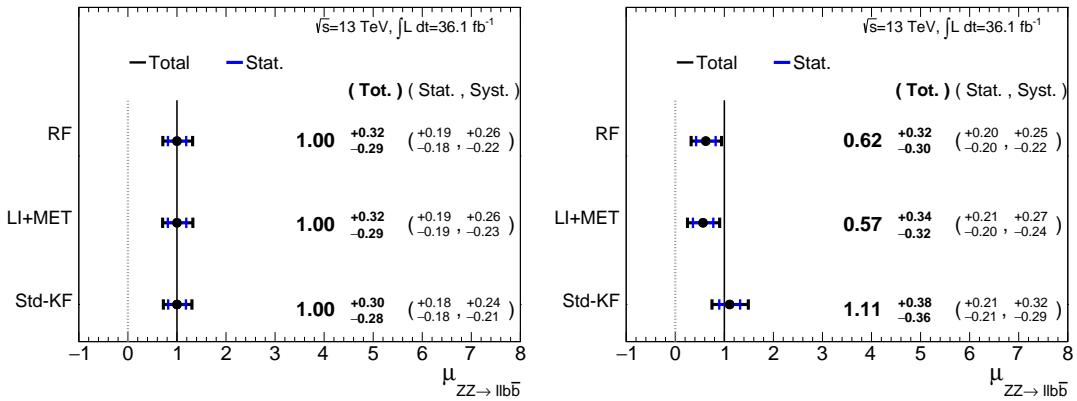


Figure 7.2: μ summary plots for the standard, LI, and RF variable sets. The Asimov case (with $\mu = 1$ by construction) is in (a), and $\hat{\mu}$ best fit values and error summary are in (b).

1757 7.4.I 2 AND ≥ 3 JET FITS

1758 While the treatment of simply ignoring any additional jets in the event seems adequate for the VH
 1759 analysis (discussed below), the potential shortcoming of this treatment appears in the VZ analysis
 1760 when the 2 and ≥ 3 jet cases are fit separately[¶], as can be seen in Figure 7.3. Compared to the stan-
 1761 dard fit, the LI and RF fits have lower $\hat{\mu}_{\geq 3 \text{ jet}}$ values, consistent with the interpretation that the ad-
 1762 ditional information from the third jet in the ≥ 3 jet regions for the standard case is important for
 1763 characterizing events in these regions for VZ fits.

1764 A natural question to ask is why this would be an issue for the VZ but not the VH case. One
 1765 potential answer is that at high transverse boosts, there is a greater probability for final state
 1766 radiation in the hadronically decaying Z , so there are more events where the third jet should be in-
 1767 cluded in the calculation of variables like $m_{b\bar{b}}$ or for angles involving the $b\bar{b}$ system (e.g. $\cos\theta$ in the
 1768 RF case). While the absolute scale at which the low and high p_T^V regions are separated remains the

[¶]standalone fits, with half the regions each, not 2 POI fits

same does not change from the VH to the VZ analysis, 150 GeV, the implicit cutoff on the transverse boost of the hadronically decaying boson does. For the Higgs, with a mass of 125 GeV, the p_T^V cutoff corresponds to $\gamma \sim 1.56 - 6.74$, but for the Z , with a mass of 91 GeV, this is $\gamma \sim 1.93 - 9.21$, about 23–37% higher.

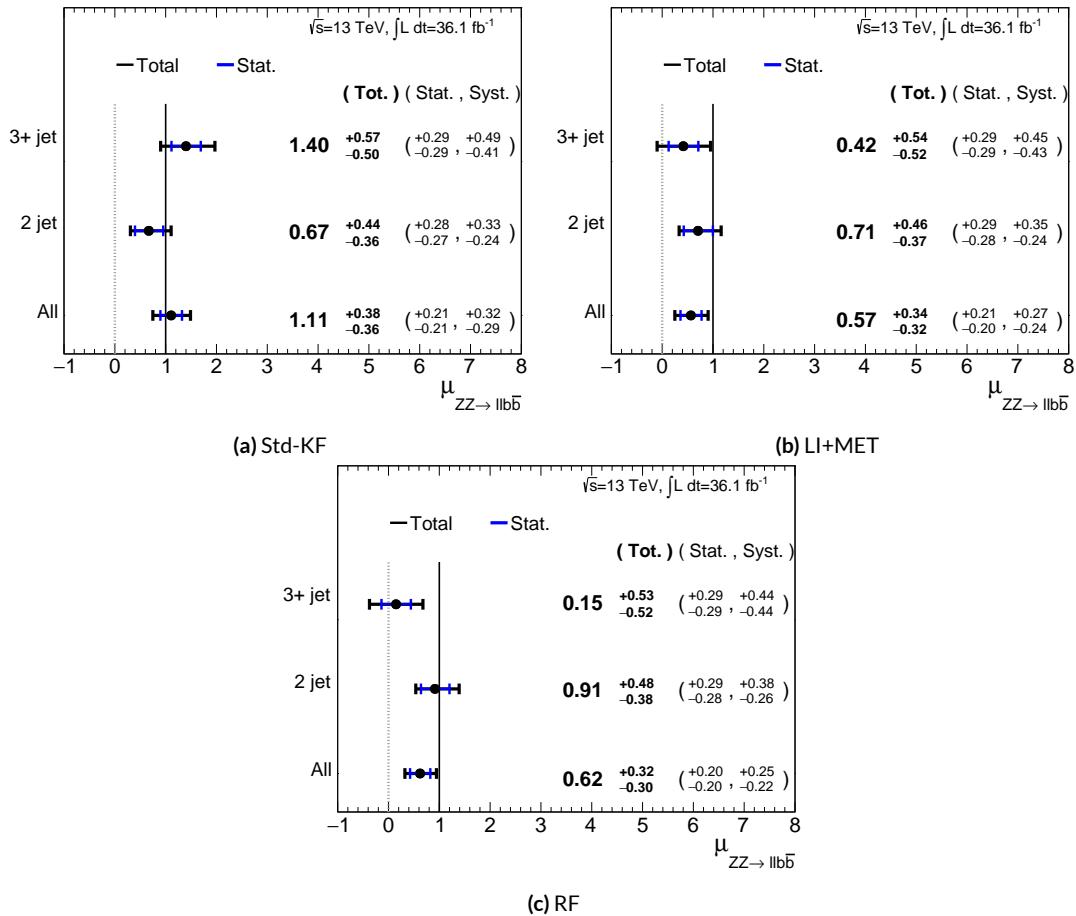


Figure 7.3: $\hat{\mu}$ summary plots with standalone fits for the different n_{jet} regions for the standard, LI, and RF variable sets.

If either the LI or RF schemes were to be used in a mainstream analysis, these validation fits suggest that the third jet ought to be included in variable schemes (e.g. by adding the third jet to the

1775 Higgs in the high p_T^V case). On the issue of whether or not ≥ 4 jet events should be included, the RF
1776 set shows very little sensitivity to this change (a 2 jet and 3 jet only fit moves $\hat{\mu}$ to 0.64, while doing so
1777 for the LI set moves it to 0.40), so this, like the addition of the third jet into the variable sets, would
1778 have to be addressed individually. Nevertheless, this optimization is beyond the scope of this thesis,
1779 which aims to preserve as much of the fiducial analysis as possible for as straightforward a compari-
1780 son as possible.

1781 For completeness, we include the full set of fit validation results for the VZ fit, explaining them in
1782 turn.

1783 7.5 NUISANCE PARAMETER PULLS

1784 The first set of plots statistical fit experts will want to look at are the “pulls” and “pull comparisons.”
1785 In these plots, the best fit (nominal) values and one standard deviation error bars are shown for ob-
1786 served (Asimov) pull plots, with the green and yellow bands corresponding to $\pm 1, 2\sigma$, respectively.
1787 These plots are divided by NP category for readability. [¶] In pull comparisons, these pulls are over-
1788 layed and color-coded. Pull comparisons here have the following color code: black is the standard
1789 variable set, red is the LI set, and blue is the RF set.

1790 We give an example in Figure 7.4, which shows the Z -jets modeling systematics. The green (yel-
1791 low) band represents ± 1 (2) σ . NP’s with norm are the freely floating parameters mentioned above
1792 initialized to 1 ± 0 , while all other NP’s are the Gaussian (or log-normal for non-freely floating nor-

¶Over 100 non-MC stat NP’s survive pruning in these 2-lepton only Run 2 fits; well over 500 survive in the Run 1+Run 2 combined fit.

₁₇₉₃ malizations) are initialized^{**} in the fit to 0 ± 1 . Both types of NP's can be seen in the figure. Gen-
₁₇₉₄ erally, some deviation is expected in a fit, but several "strongly pulled" (i.e. the absolute value of an
₁₇₉₅ NP central value divided by the error bar size is greater than one) NP's is a sign of a "bad" fit. Here,
₁₇₉₆ ZPtV is strongly pulled for all three fits, and ZMbb is strongly pulled for the LI fits; this is unsurpris-
₁₇₉₇ ing since $Z+jets$ modeling has traditionally been hard. Such plots let the analyzer look for problems
₁₇₉₈ in fit models and compare different analyses and are the first fit diagnostic plot used in the fit work-
₁₇₉₉ flow.

^{**}with the usual mean subtraction/standard deviation division applied

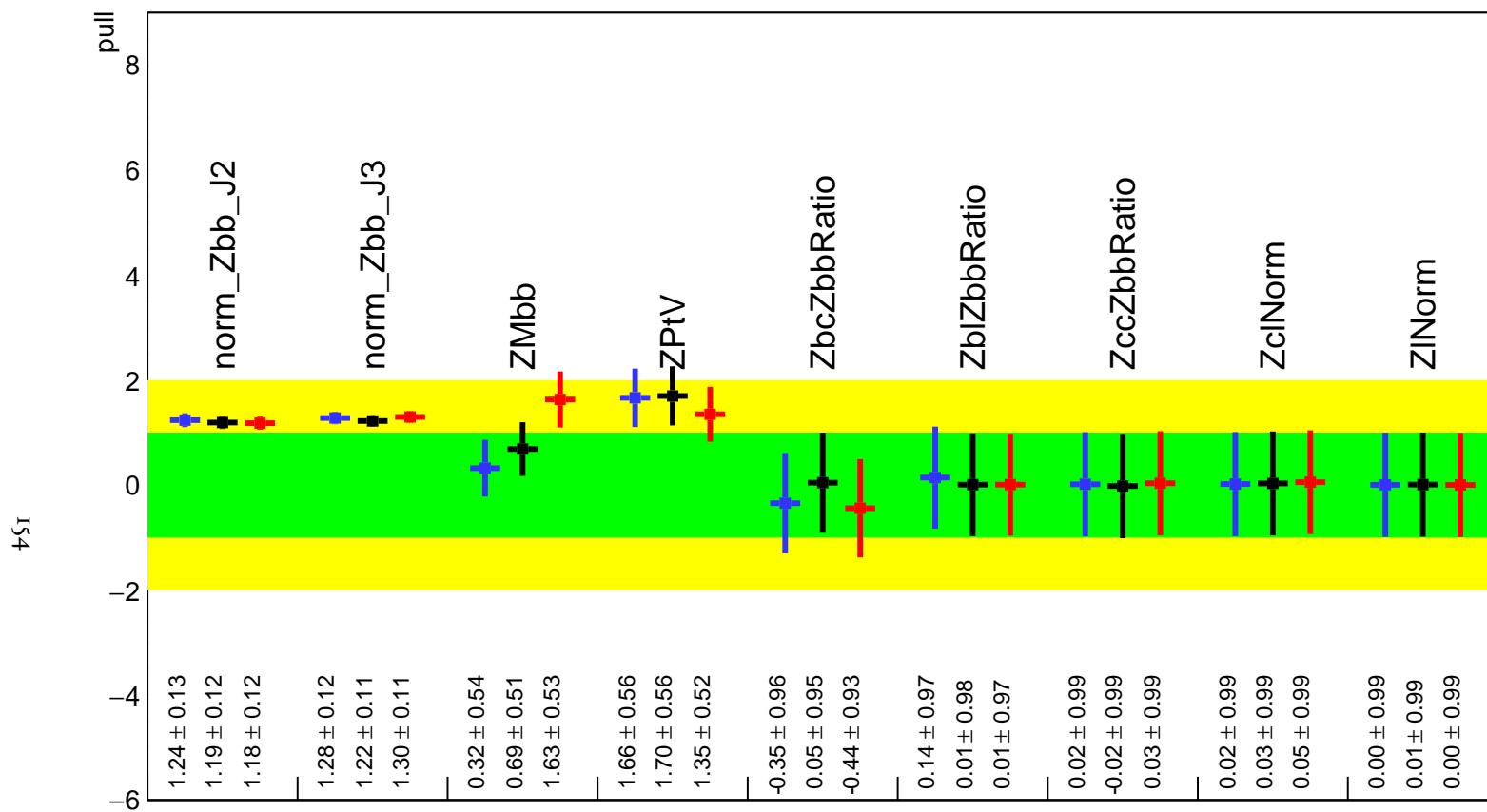


Figure 7.4: Pull comparison for signal process modeling NP's.

1800 A well-behaved fit has pulls close to nominal values (“closeness” should be interpreted in the
 1801 context of pull value divided by pull error). As can be seen in Figures 7.5–7.9, the fits for the three
 1802 different variable sets are fairly similar from a NP pull perspective, though the $Z+jets$ m_{bb} and p_T^V
 1803 NP’s and the jet energy resolution NP are heavily pulled (a handful of poorly behaved pulls is not
 1804 uncommon, though typically warrants further investigation). As a general note, these pull plots cal-
 1805 culate pulls using a simultaneous HESSE matrix inversion, which is fine for relatively small fits, but
 1806 the more reliable MINOS result, which calculates the impact of each NP on its own, should be cross-
 1807 checked for significant pulls^{††}. The ranking plots below do this.

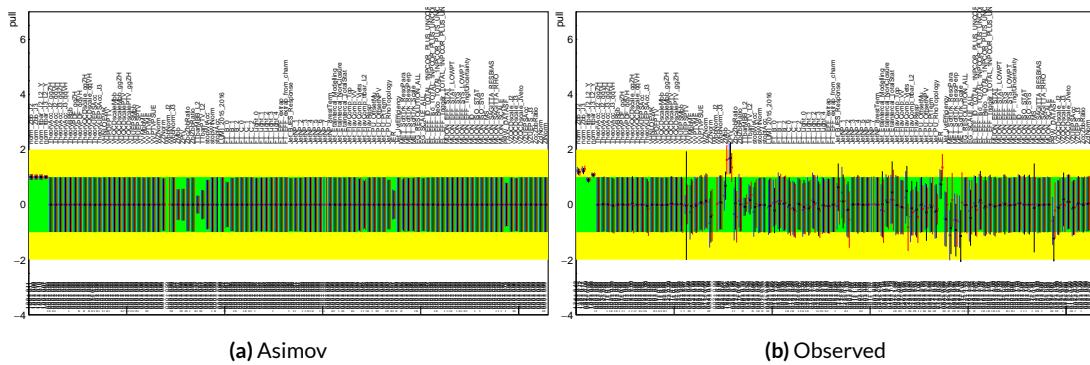


Figure 7.5: Pull comparison for all NP’s but MC stats.

1808 Nuisance parameter correlation matrices (for correlations with magnitude at least 0.25) for all
 1809 three variable set fits can be found in Figures 7.11–7.13. These are useful for seeing which NP’s move
 1810 together (if there is no physical argument for them to do so, this is a potential indicator that further
 1811 investigation is warranted). These are similar to the correlation matrices in e.g. Figure 6.9 except
 1812 with ordering inverted (so the diagonal is oriented the other way).

^{††}This becomes more of an issue for very large fits, like the full Run 1 + Run 2 combined fits in Chapter 9.

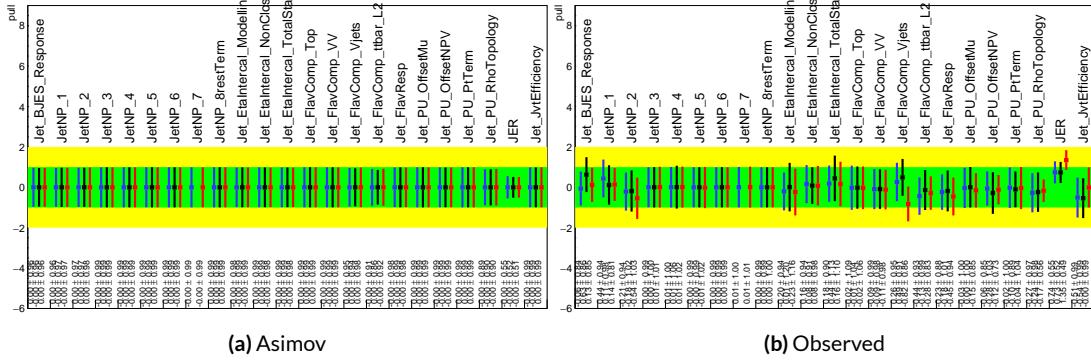


Figure 7.6: Pull comparison for jet NP's.

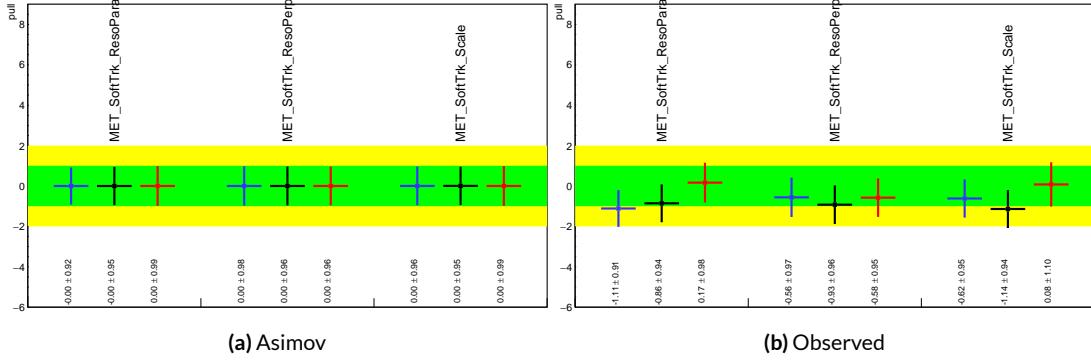


Figure 7.7: Pull comparison for MET NP's.

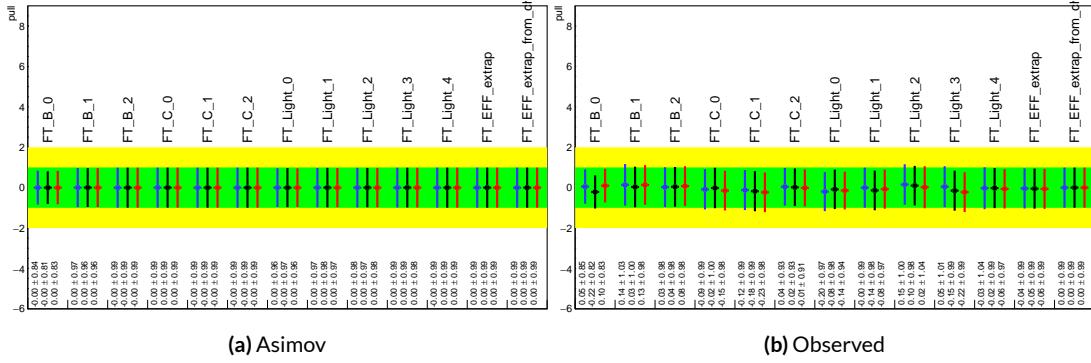


Figure 7.8: Pull comparison for Flavour Tagging NP's.

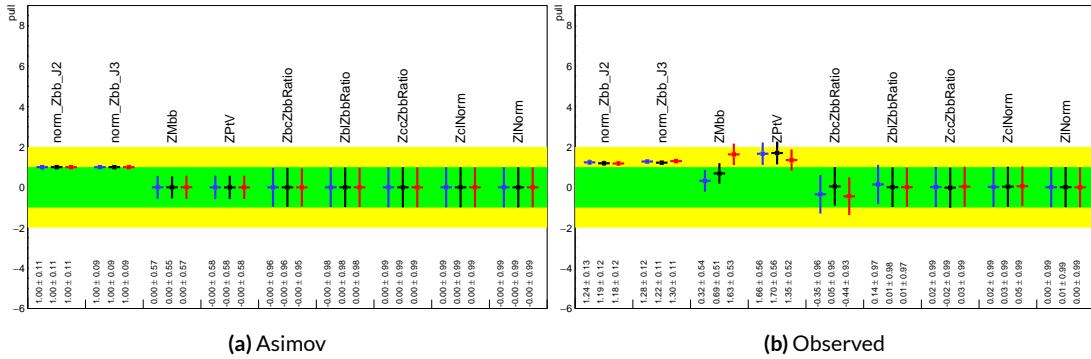


Figure 7.9: Pull comparison for $Z + \text{jets}$ NP's.

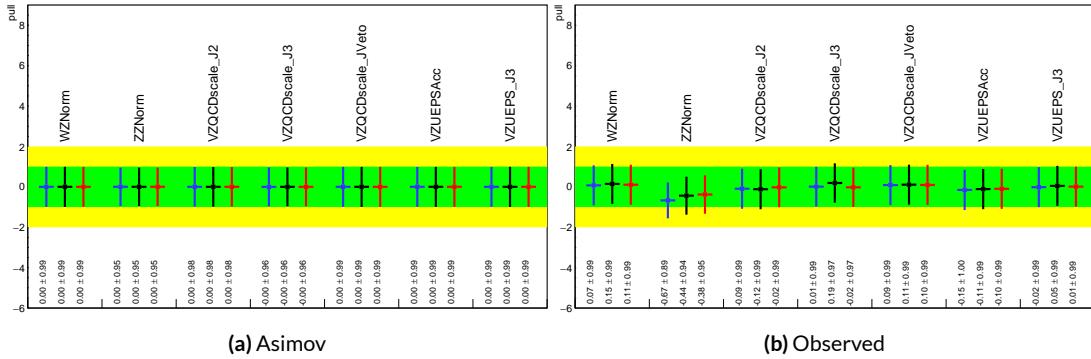


Figure 7.10: Pull comparison for diboson process modeling NP's.

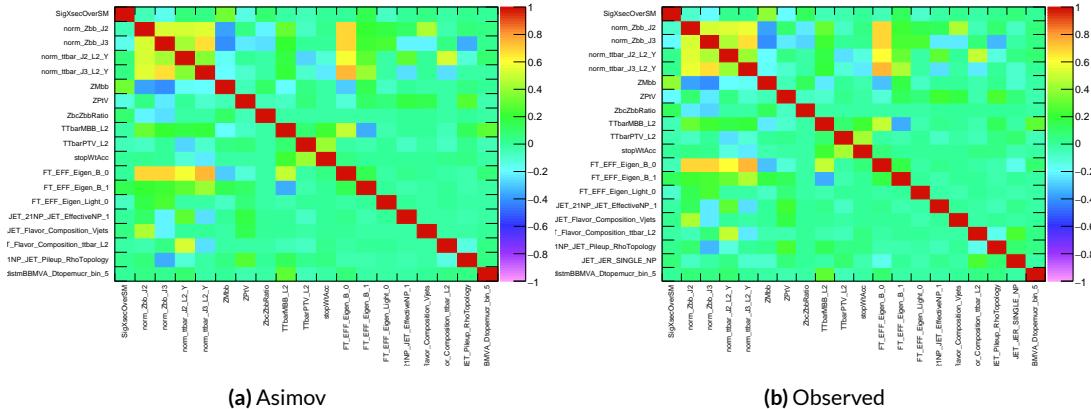


Figure 7.11: NP correlations for standard variable fits.

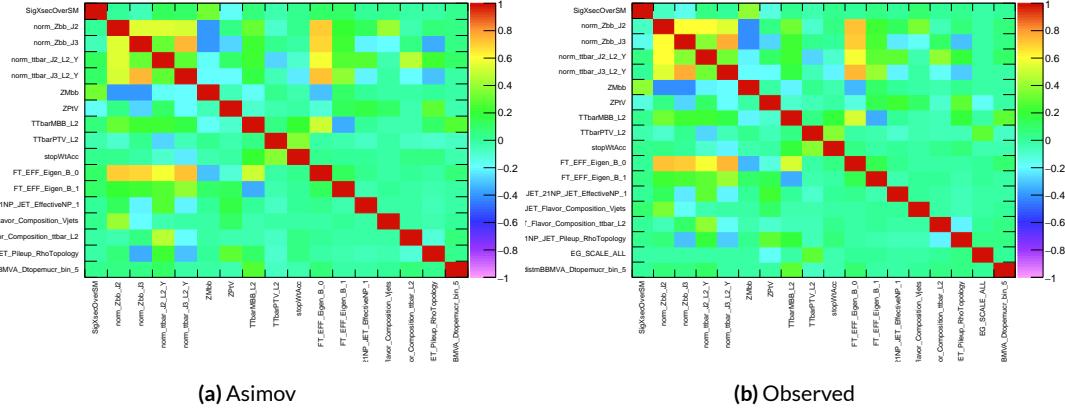


Figure 7.12: NP correlations for LI variable fits.

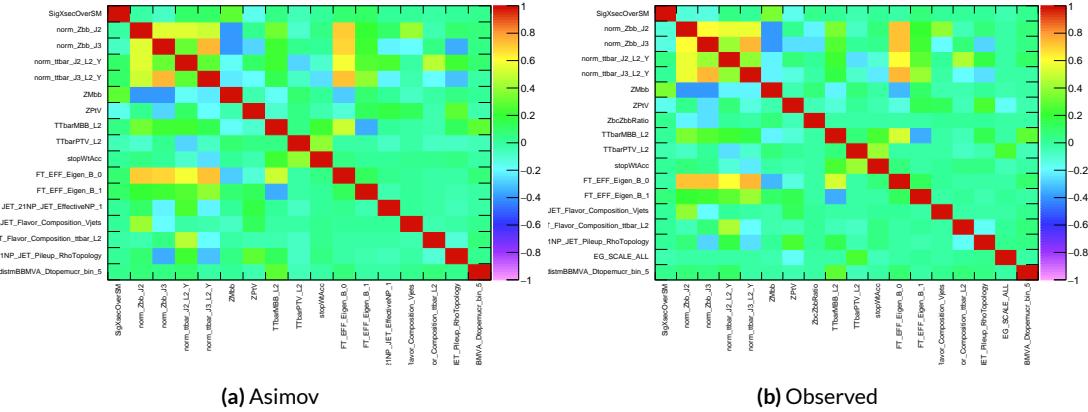


Figure 7.13: NP correlations for RF variable fits.

1813 7.5.1 NUISANCE PARAMETER RANKING PLOTS AND BREAKDOWNS

1814 The next set of fit results that is used to diagnose the quality of a fit is the impact of different nui-
1815 sance parameters on the total error on μ , both individually and as categories. Figure 7.14 shows the
1816 top 25 nuisance parameters ranked by their postfit impact on $\hat{\mu}$; these plots use the aforementioned,
1817 more reliable MINOS approach. This set of rankings is fairly similar, with $Z+jets$ systematics being
1818 particularly prominent. The advantage of seeing individual nuisance parameter rankings, as op-
1819 posed to impacts of categories in aggregate, is that particularly pathological NP's are easier to see; in
1820 particular, jet energy resolution and $Z+jets p_T^V$ systematic from the pull comparison plots show up
1821 with high rankings. Blue (Yellow) bands are post-(pre-) fit impact on μ , and the black (red) points
1822 with error bars are the same as for the standard (floating normalization) NP's in the pull comparison
1823 plots.

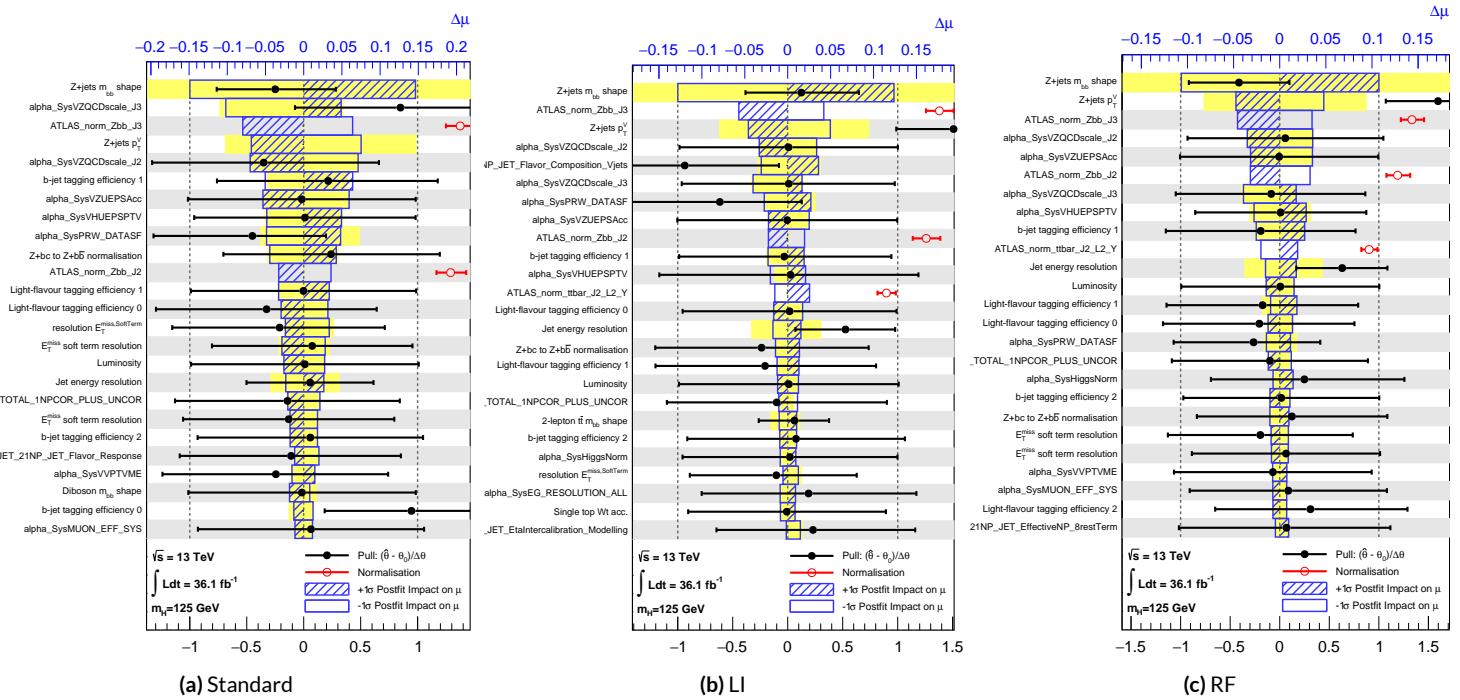


Figure 7.14: Plots for the top 25 nuisance parameters according to their postfit impact on $\hat{\mu}$ for the standard (a), LI (b), and RF (c) variable sets.

¹⁸²⁴ This is consistent with the picture of NP's taken in aggregate categories in Tables 7.5 and 7.6,
¹⁸²⁵ known as “breakdowns,” with $Z+jets$ in particular featuring prominently. Of particular interest
¹⁸²⁶ is also the lower impact of MC stats in the observed fit.

	Std-KF	LI+MET	RF
Total	+0.305 / -0.277	+0.324 / -0.292	+0.319 / -0.288
DataStat	+0.183 / -0.179	+0.190 / -0.186	+0.188 / -0.184
FullSyst	+0.244 / -0.212	+0.262 / -0.226	+0.258 / -0.221
Floating normalizations	+0.092 / -0.084	+0.098 / -0.079	+0.094 / -0.076
All normalizations	+0.093 / -0.084	+0.098 / -0.079	+0.094 / -0.076
All but normalizations	+0.214 / -0.179	+0.229 / -0.188	+0.224 / -0.182
Jets, MET	+0.052 / -0.043	+0.041 / -0.034	+0.047 / -0.037
Jets	+0.034 / -0.029	+0.033 / -0.028	+0.032 / -0.026
MET	+0.035 / -0.027	+0.015 / -0.012	+0.020 / -0.016
BTag	+0.064 / -0.051	+0.063 / -0.031	+0.059 / -0.032
BTag b	+0.053 / -0.041	+0.061 / -0.028	+0.055 / -0.025
BTag c	+0.011 / -0.010	+0.006 / -0.005	+0.007 / -0.006
BTag light	+0.030 / -0.027	+0.016 / -0.013	+0.022 / -0.019
Leptons	+0.021 / -0.012	+0.022 / -0.014	+0.023 / -0.014
Luminosity	+0.039 / -0.022	+0.039 / -0.022	+0.040 / -0.022
Diboson	+0.049 / -0.028	+0.047 / -0.026	+0.047 / -0.026
Model Zjets	+0.106 / -0.105	+0.113 / -0.110	+0.102 / -0.099
Zjets flt. norm.	+0.039 / -0.053	+0.024 / -0.029	+0.021 / -0.031
Model Wjets	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Wjets flt. norm.	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Model ttbar	+0.015 / -0.013	+0.032 / -0.017	+0.030 / -0.016
Model Single Top	+0.004 / -0.003	+0.009 / -0.008	+0.005 / -0.004
Model Multi Jet	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Signal Systematics	+0.003 / -0.003	+0.003 / -0.003	+0.003 / -0.003
MC stat	+0.097 / -0.094	+0.108 / -0.103	+0.107 / -0.104

Table 7.5: Summary of impact of various nuisance parameter categories on the error on μ for Asimov fits for the standard, LI, and RF variable sets.

	Std-KF	LI+MET	RF
$\hat{\mu}$	1.1079	0.5651	0.6218
Total	+0.381 / -0.360	+0.339 / -0.316	+0.322 / -0.299
DataStat	+0.214 / -0.211	+0.210 / -0.205	+0.201 / -0.197
FullSyst	+0.315 / -0.292	+0.267 / -0.241	+0.252 / -0.225
Floating normalizations	+0.120 / -0.122	+0.095 / -0.089	+0.082 / -0.079
All normalizations	+0.121 / -0.123	+0.095 / -0.090	+0.082 / -0.079
All but normalizations	+0.279 / -0.254	+0.228 / -0.200	+0.213 / -0.184
Jets, MET	+0.076 / -0.065	+0.045 / -0.043	+0.038 / -0.033
Jets	+0.047 / -0.040	+0.044 / -0.041	+0.027 / -0.024
MET	+0.055 / -0.046	+0.015 / -0.015	+0.012 / -0.010
BTag	+0.083 / -0.079	+0.041 / -0.031	+0.041 / -0.035
BTag b	+0.063 / -0.059	+0.032 / -0.022	+0.031 / -0.026
BTag c	+0.018 / -0.017	+0.008 / -0.007	+0.010 / -0.009
BTag light	+0.051 / -0.046	+0.024 / -0.021	+0.025 / -0.022
Leptons	+0.022 / -0.011	+0.015 / -0.008	+0.019 / -0.008
Luminosity	+0.044 / -0.022	+0.026 / -0.006	+0.027 / -0.008
Diboson	+0.049 / -0.026	+0.025 / -0.013	+0.027 / -0.017
Model Zjets	+0.156 / -0.162	+0.133 / -0.133	+0.115 / -0.117
Zjets flt. norm.	+0.061 / -0.089	+0.041 / -0.064	+0.028 / -0.056
Model Wjets	+0.000 / -0.001	+0.000 / -0.001	+0.000 / -0.001
Wjets flt. norm.	+0.000 / -0.001	+0.000 / -0.001	+0.000 / -0.001
Model ttbar	+0.015 / -0.024	+0.018 / -0.005	+0.017 / -0.009
Model Single Top	+0.005 / -0.003	+0.010 / -0.008	+0.007 / -0.004
Model Multi Jet	+0.000 / -0.001	+0.000 / -0.001	+0.000 / -0.001
Signal Systematics	+0.005 / -0.004	+0.009 / -0.006	+0.005 / -0.006
MC stat	+0.140 / -0.143	+0.132 / -0.131	+0.128 / -0.129

Table 7.6: Summary of impact of various nuisance parameter categories on the error on $\hat{\mu}$ for observed fits for the standard, LI, and RF variable sets.

1827 7.6 POSTFIT DISTRIBUTIONS

1828 Finally, postfit distributions for the MVA discriminant (m_{bb}) distribution in the signal (top $e - \mu$
1829 control) region for the standard, Lorentz Invariant, and RestFrames variable sets are shown. It is
1830 generally considered good practice to check the actual postfit distributions of discriminating quan-
1831 tities used to make sure there is good agreement. ^{‡‡} It should be noted that agreement is not always
1832 great when “eyeballing” a distribution, as fits are messy and $V+hf$ modeling is notoriously hard.
1833 This is particularly true in the VZ fit since normalizations for $Z+hf$ in particular are derived using
1834 VH optimized sidebands. This is also why a lot of these plots are presented as log plots (which hide
1835 disagreement better; the general argument goes that one has the ratio plots on the bottom and log
1836 plots allow one to see rare backgrounds in plots).

^{‡‡}Sometimes distributions of input variables (MC histograms scaled by their postfit normalizations) are also used.

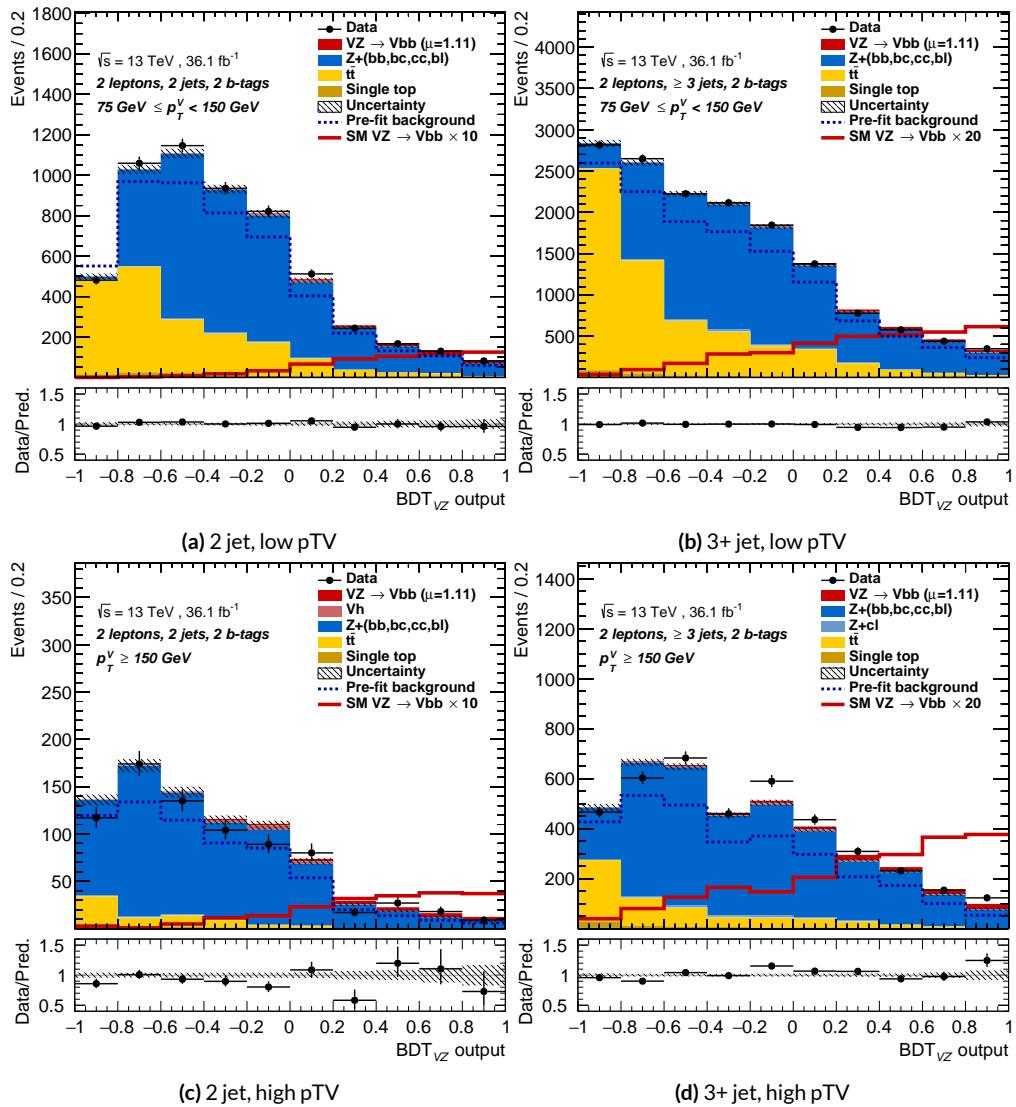
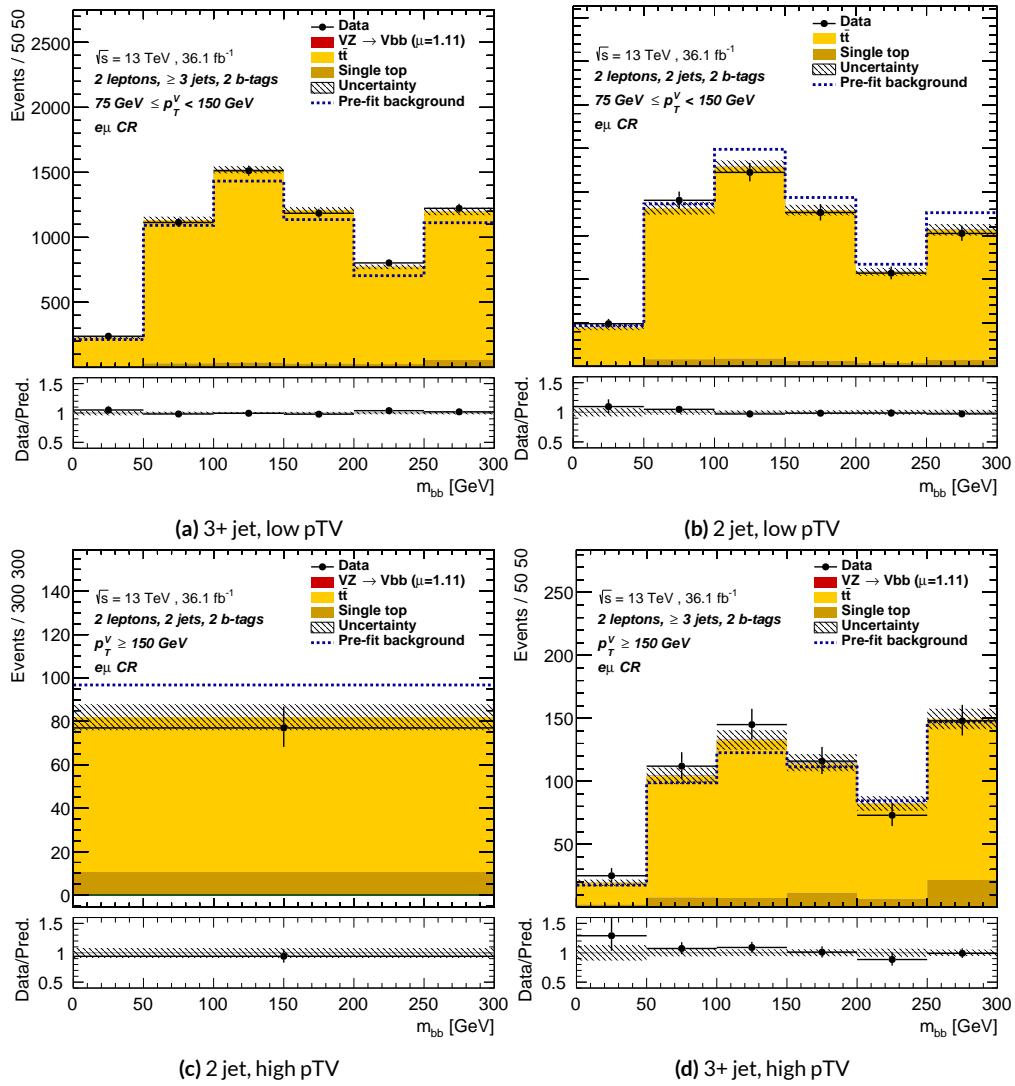


Figure 7.15: Postfit BDT_{VZ} plots in the signal region for the standard variable set.



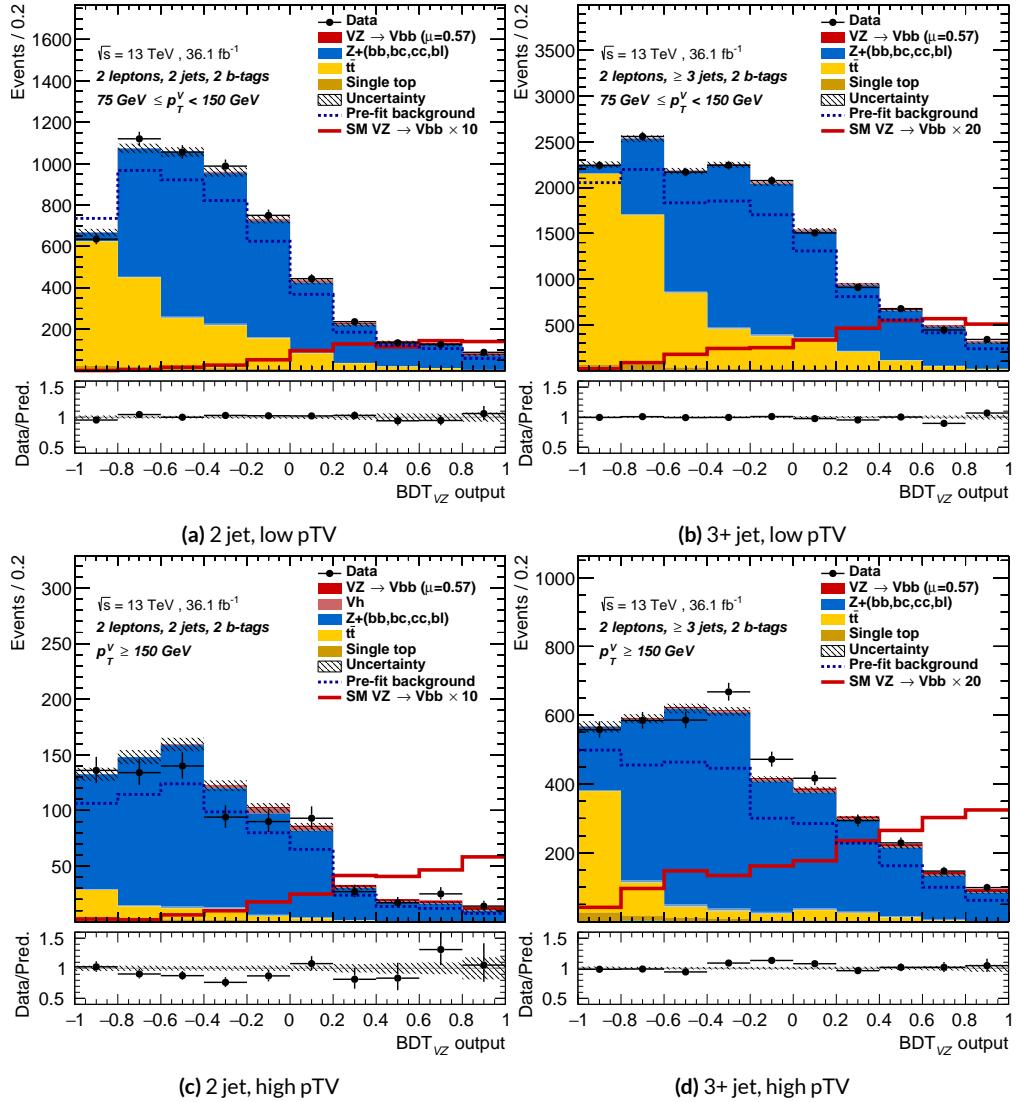


Figure 7.17: Postfit BDT_{VZ} plots in the signal region for the LI variable set.

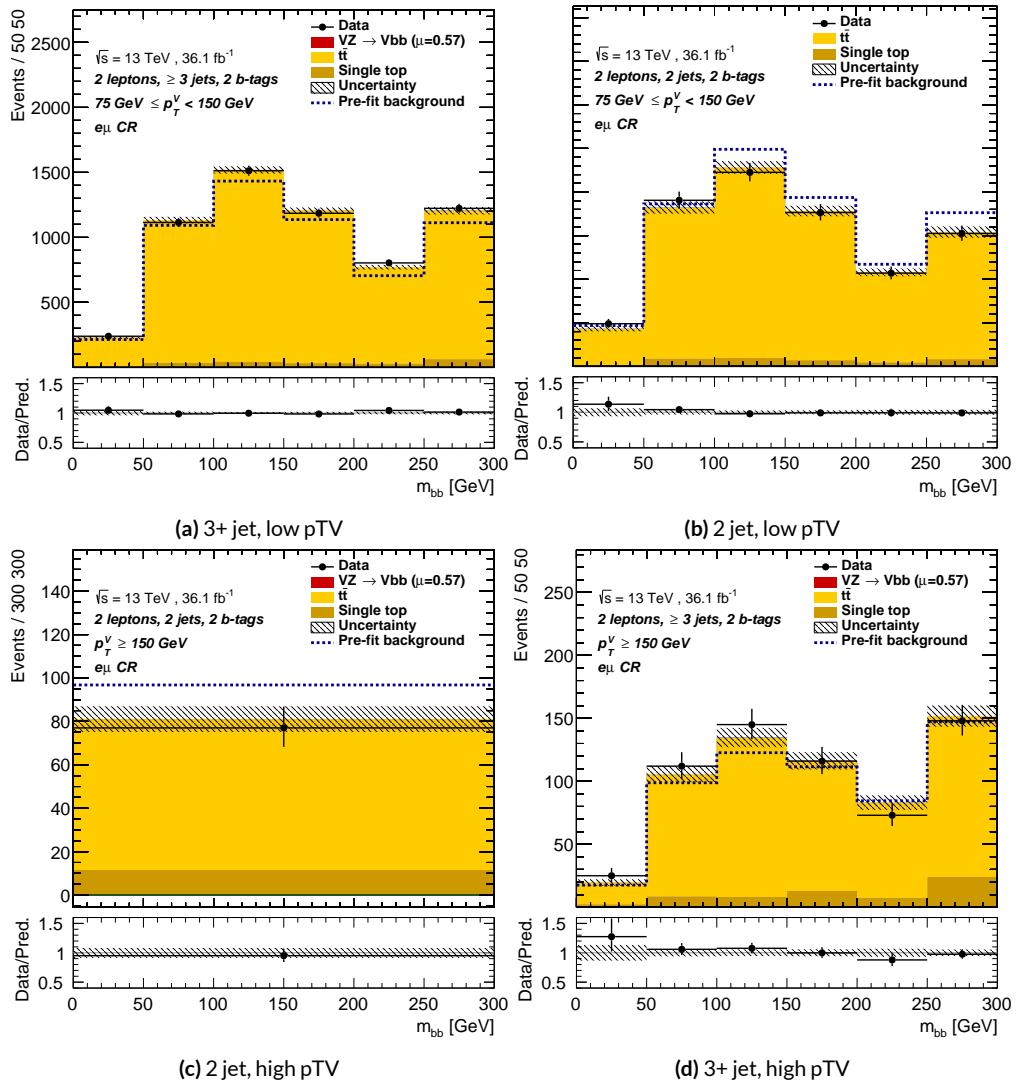


Figure 7.18: Postfit m_{bb} plots in the top $e - \mu$ CR for the LI variable set.

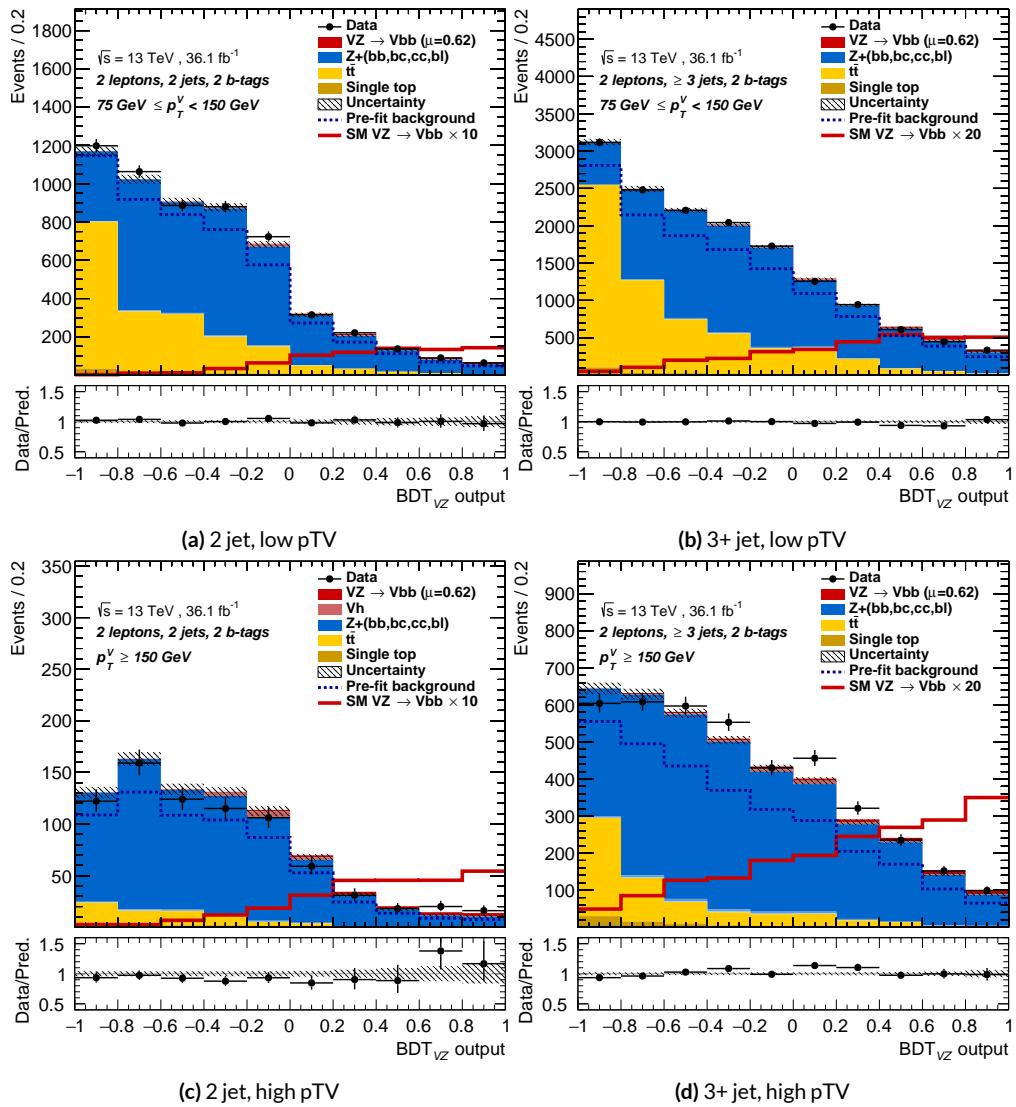


Figure 7.19: Postfit BDT_{VZ} plots in the signal region for the RF variable set.

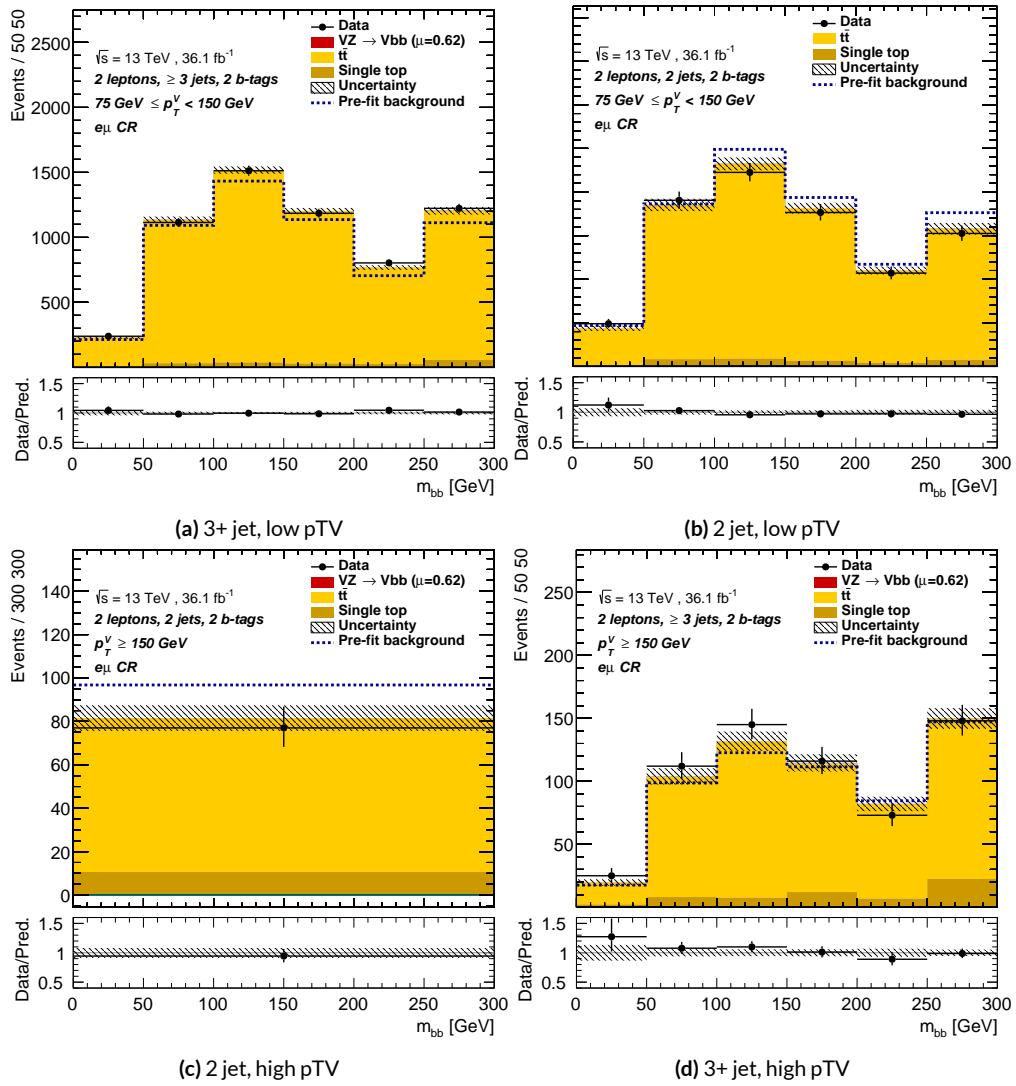


Figure 7.20: Postfit m_{bb} plots in the top $e - \mu$ CR for the RF variable set.

1837 7.7 VH FIT MODEL VALIDATION

1838 We now move onto the fit validation distributions and numbers for the VH fit of interest.

1839 7.7.1 NUISANCE PARAMETER PULLS

1840 As can be seen in Figures 7.21–7.25, the fits for the three different variable sets are fairly similar from
1841 a NP pull perspective. Again, black is the standard variable set, red is the LI set, and blue is the RF
1842 set. The possible exception is the signal $UE+PS p_T^V$ systematic, which looks very different for all three
1843 cases (underconstrained for the standard, but overconstrained for the novel variable cases), though
1844 this difference goes away in the ranking plot, meaning this is almost certainly an unphysical artifice
1845 of the faster HESSE inversion used to produce the pull comparison plots.

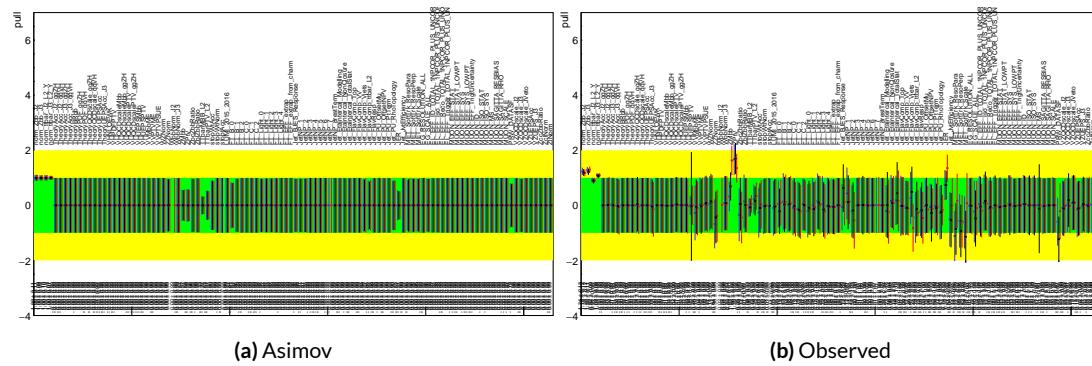


Figure 7.21: Pull comparison for all NP's but MC stats.

1846 Nuisance parameter correlation matrices (for correlations with magnitude at least 0.25) for all
1847 three variable set fits can be found in Figures 7.27–7.29.

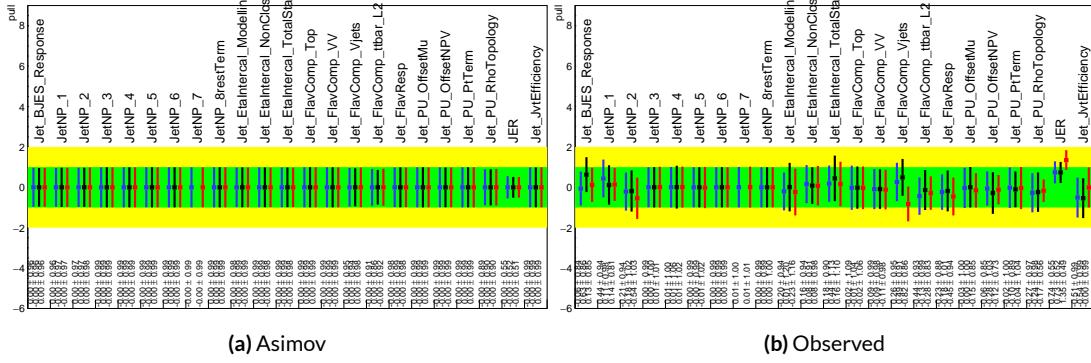


Figure 7.22: Pull comparison for jet NP's.

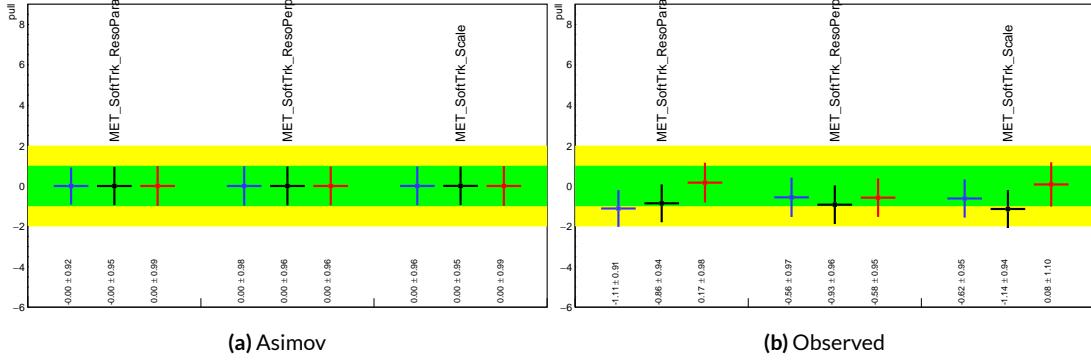


Figure 7.23: Pull comparison for MET NP's.

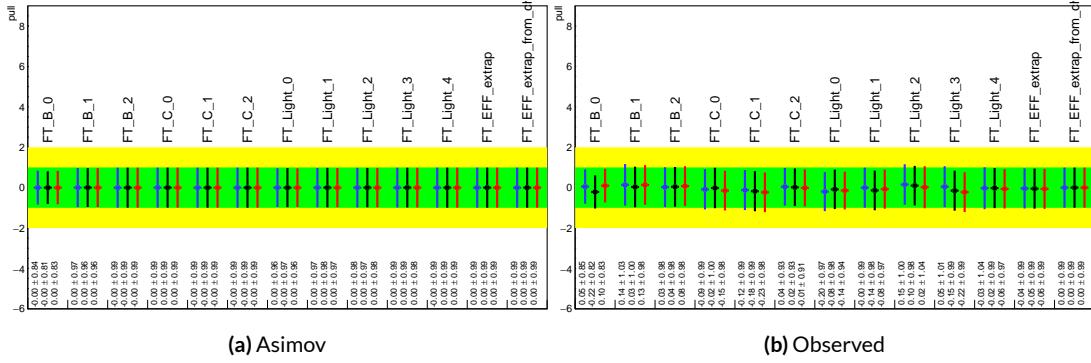


Figure 7.24: Pull comparison for Flavour Tagging NP's.

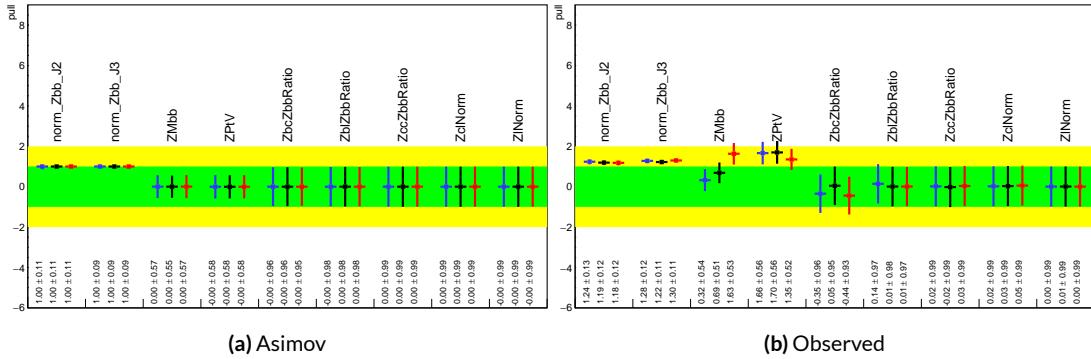


Figure 7.25: Pull comparison for $Z + \text{jets}$ NP's.

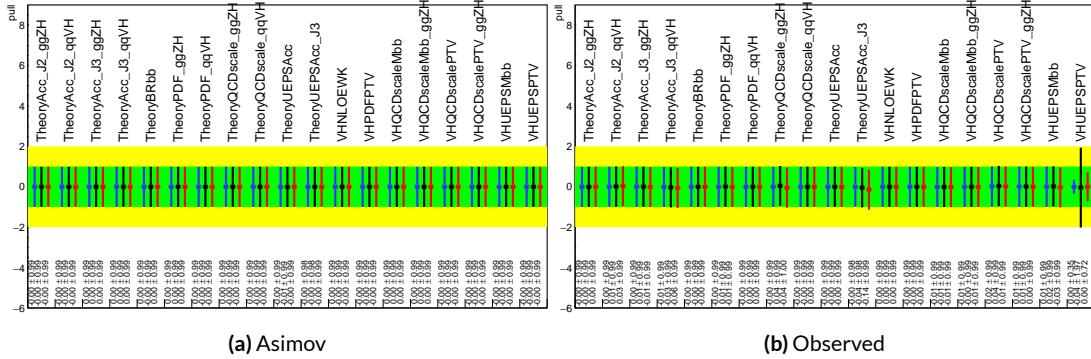


Figure 7.26: Pull comparison for signal process modeling NP's.

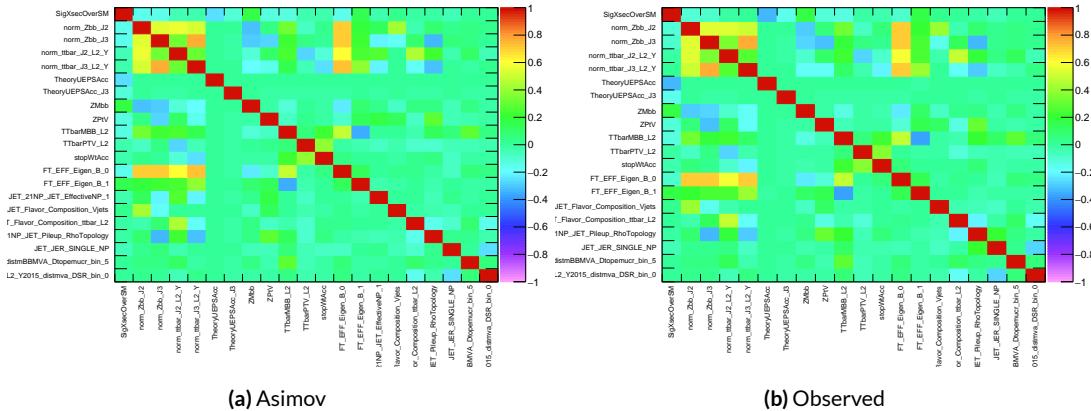


Figure 7.27: NP correlations for standard variable fits.

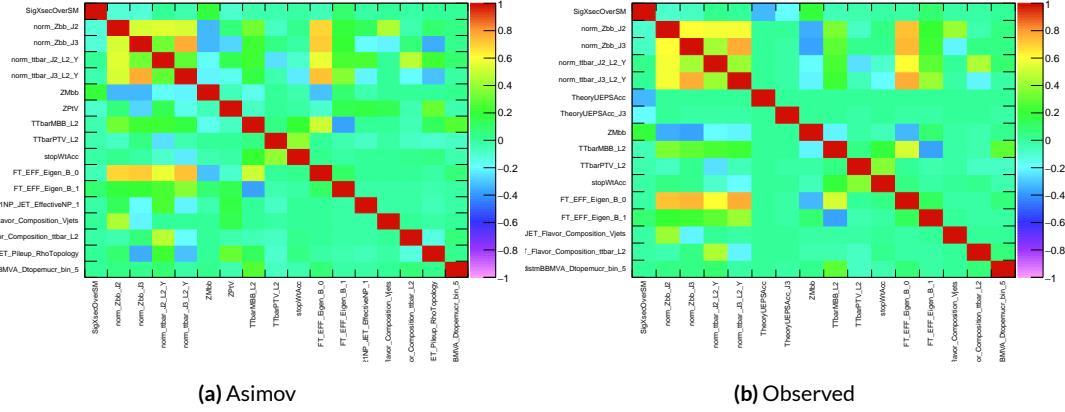


Figure 7.28: NP correlations for LI variable fits.

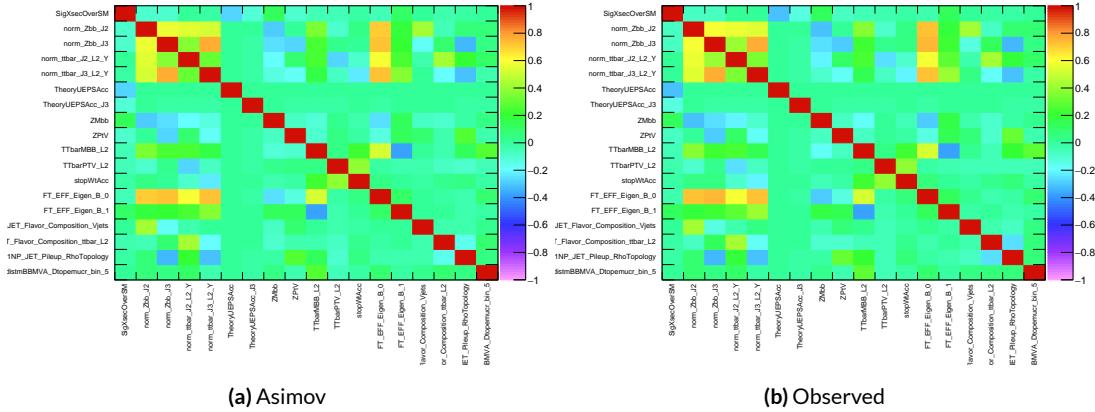


Figure 7.29: NP correlations for RF variable fits.

1848 7.7.2 FULL BREAKDOWN OF ERRORS

1849 A postfit ranking of nuisance parameters according to their impact on $\hat{\mu}$ for the different variable
1850 sets may be found in Figure 7.30, with rankings being fairly similar. In particular, the signal UE+PS
1851 p_T^V systematic is top-ranked for all three variable sets and also looks very similar, unlike in the pull
1852 comparison plot, reiterating the importance of evaluating individually the impact of highly ranked
1853 NP's. The $Z+jets p_T^V$ is highly pulled in all three cases, though this is less severe for the non-standard
1854 set (it is off the scale for the standard variable ranking). The RF discriminant mitigates the effect of
1855 poorly modeled jet energy resolution better than the other sets.

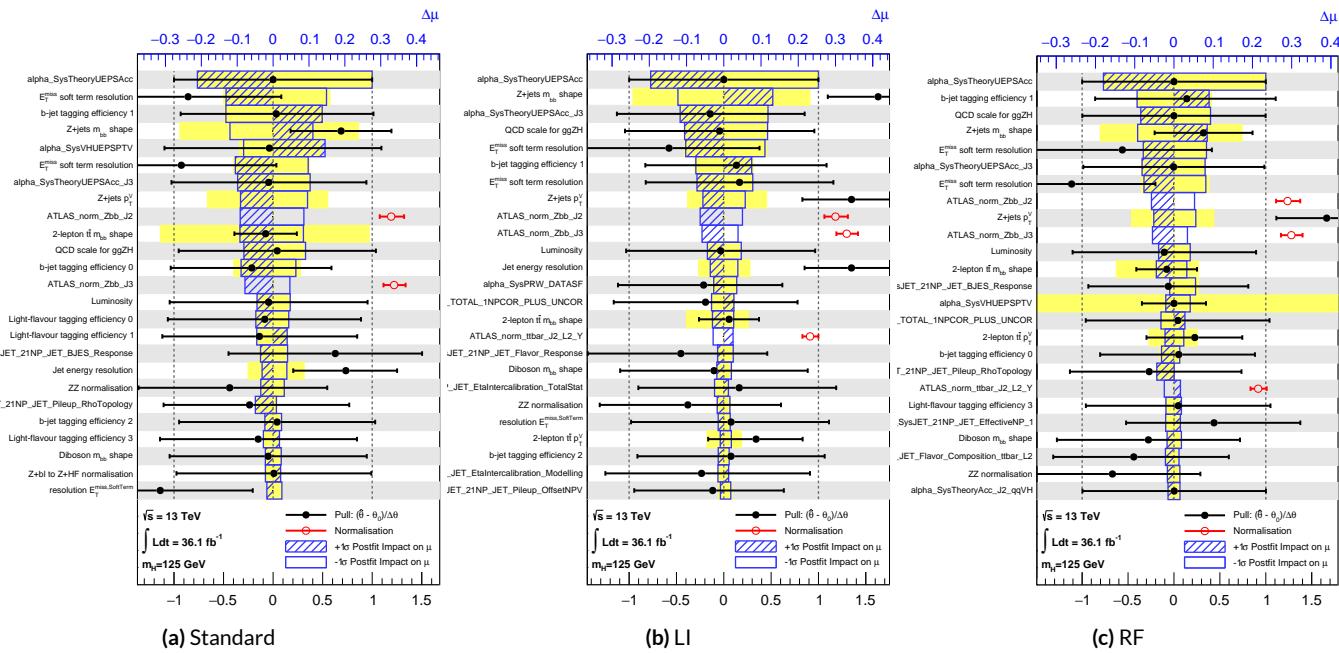


Figure 7.30: Plots for the top 25 nuisance parameters according to their postfit impact on $\hat{\mu}$ for the standard (a), LI (b), and RF (c) variable sets.

1856 The Asimov (Table 7.7) and observed (Table 7.8) breakdowns both consistently suggest that the
 1857 LI variable set does a better job of constraining systematic uncertainties than the standard set and
 1858 that the RF set does better still. It is also not surprising that the gain is more substantial in the ob-
 1859 served fit than in the Asimov fits, as in the latter there are the “penalty” terms from pulls in addition
 1860 to the overall broadening in the likelihood.

	Std-KF	LI+MET	RF
Total	+0.608 / -0.511	+0.632 / -0.539	+0.600 / -0.494
DataStat	+0.420 / -0.401	+0.453 / -0.434	+0.424 / -0.404
FullSyst	+0.440 / -0.318	+0.441 / -0.319	+0.425 / -0.284
Floating normalizations	+0.122 / -0.125	+0.110 / -0.111	+0.093 / -0.089
All normalizations	+0.128 / -0.129	+0.112 / -0.112	+0.099 / -0.092
All but normalizations	+0.403 / -0.274	+0.387 / -0.250	+0.382 / -0.227
Jets, MET	+0.180 / -0.097	+0.146 / -0.079	+0.122 / -0.083
Jets	+0.051 / -0.030	+0.044 / -0.035	+0.025 / -0.042
MET	+0.173 / -0.091	+0.140 / -0.074	+0.117 / -0.063
BTag	+0.138 / -0.136	+0.069 / -0.071	+0.076 / -0.078
BTag b	+0.125 / -0.125	+0.067 / -0.070	+0.073 / -0.075
BTag c	+0.018 / -0.016	+0.004 / -0.004	+0.005 / -0.005
BTag light	+0.057 / -0.051	+0.020 / -0.014	+0.009 / -0.018
Leptons	+0.013 / -0.012	+0.029 / -0.026	+0.012 / -0.023
Luminosity	+0.052 / -0.020	+0.050 / -0.016	+0.050 / -0.019
Diboson	+0.043 / -0.039	+0.035 / -0.031	+0.038 / -0.029
Model Zjets	+0.119 / -0.117	+0.124 / -0.127	+0.095 / -0.086
Zjets flt. norm.	+0.080 / -0.106	+0.052 / -0.092	+0.026 / -0.072
Model Wjets	+0.001 / -0.001	+0.001 / -0.001	+0.000 / -0.001
Wjets flt. norm.	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Model ttbar	+0.076 / -0.080	+0.025 / -0.035	+0.025 / -0.040
Model Single Top	+0.015 / -0.015	+0.002 / -0.004	+0.021 / -0.007
Model Multi Jet	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Signal Systematics	+0.262 / -0.087	+0.272 / -0.082	+0.290 / -0.088
MC stat	+0.149 / -0.136	+0.168 / -0.154	+0.153 / -0.136

Table 7.7: Expected error breakdowns for the standard, LI, and RF variable sets

	Std-KF	LI+MET	RF
$\hat{\mu}$	1.7458	1.6467	1.5019
Total	+0.811 / -0.662	+0.778 / -0.641	+0.731 / -0.612
DataStat	+0.502 / -0.484	+0.507 / -0.489	+0.500 / -0.481
FullSyst	+0.637 / -0.451	+0.591 / -0.415	+0.533 / -0.378
Floating normalizations	+0.153 / -0.143	+0.128 / -0.118	+0.110 / -0.109
All normalizations	+0.158 / -0.147	+0.130 / -0.119	+0.112 / -0.110
All but normalizations	+0.599 / -0.402	+0.544 / -0.354	+0.486 / -0.318
Jets, MET	+0.218 / -0.145	+0.198 / -0.113	+0.167 / -0.106
Jets	+0.071 / -0.059	+0.065 / -0.047	+0.036 / -0.051
MET	+0.209 / -0.130	+0.190 / -0.102	+0.152 / -0.077
BTag	+0.162 / -0.166	+0.093 / -0.070	+0.115 / -0.099
BTag b	+0.142 / -0.147	+0.090 / -0.066	+0.110 / -0.094
BTag c	+0.022 / -0.021	+0.006 / -0.006	+0.007 / -0.007
BTag light	+0.074 / -0.072	+0.025 / -0.022	+0.031 / -0.029
Leptons	+0.039 / -0.029	+0.035 / -0.031	+0.034 / -0.030
Luminosity	+0.079 / -0.039	+0.073 / -0.034	+0.069 / -0.032
Diboson	+0.047 / -0.043	+0.031 / -0.028	+0.029 / -0.028
Model Zjets	+0.164 / -0.152	+0.141 / -0.143	+0.101 / -0.105
Zjets flt. norm.	+0.070 / -0.109	+0.041 / -0.086	+0.033 / -0.083
Model Wjets	+0.001 / -0.001	+0.001 / -0.000	+0.001 / -0.001
Wjets flt. norm.	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Model ttbar	+0.067 / -0.102	+0.029 / -0.040	+0.040 / -0.048
Model Single Top	+0.015 / -0.020	+0.001 / -0.005	+0.004 / -0.006
Model Multi Jet	+0.000 / -0.000	+0.000 / -0.000	+0.000 / -0.000
Signal Systematics	+0.434 / -0.183	+0.418 / -0.190	+0.364 / -0.152
MC stat	+0.226 / -0.201	+0.221 / -0.200	+0.212 / -0.189

Table 7.8: Observed signal strengths, and error breakdowns for the standard, LI, and RF variable sets

1861 7.7.3 POSTFIT DISTRIBUTIONS AND S/B PLOTS

1862 Postfit distributions for the MVA discriminant (m_{bb}) distribution in the signal (top $e - \mu$ control)
1863 region for the standard, Lorentz Invariant, and RestFrames variable sets are found in Figures 7.31–
1864 7.36. Here, as in the VZ fit, agreement is reasonable. In a combined fit with all three channels, $Z+hf$
1865 normalizations in particular would be correlated across the 0- and 2-lepton channels, which might
1866 help to better constrain this mismodeling (and perhaps as a result some of the $Z+jets$ systematics as
1867 well).

1868 One final type of plot presented as a result is the binned $\log_{10} (S/B)$ in signal regions distribu-
1869 tions may be found in Figure 7.37. For these plots, one fills a histogram with the $\log_{10} (S/B)$ ratio in
1870 each postfit distribution bin weighted by the total number of events. In this case, a log plot is help-
1871 ful because the highest bins would be invisible on a linear plot. These distributions are allegedly use-
1872 ful for seeing where most of one's sensitivity lies. Practically, it is problematic if the pull (from the
1873 null hypothesis) is higher at lower S/B values, which may indicate a poorly optimized discriminant.

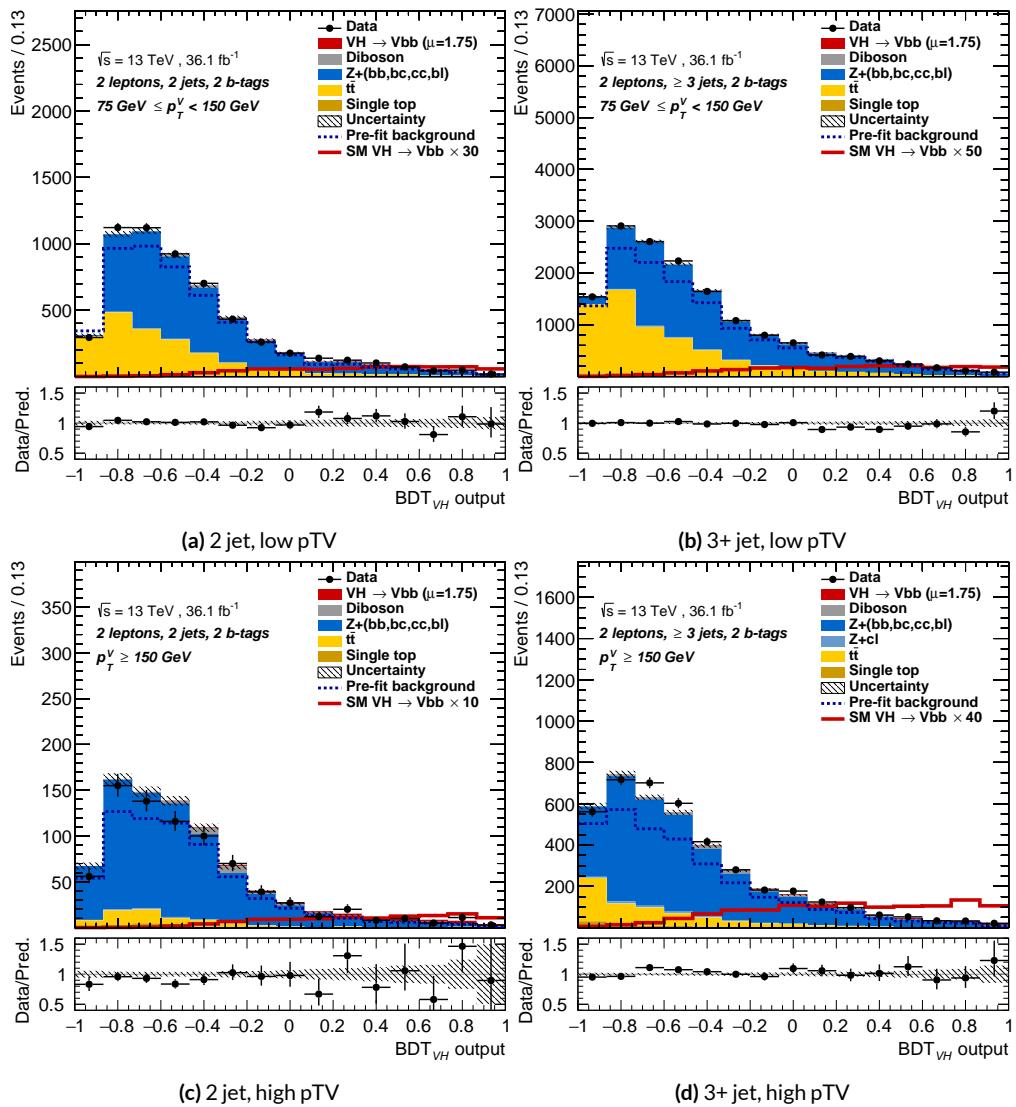


Figure 7.31: Postfit BDT_{VH} plots in the signal region for the standard variable set.

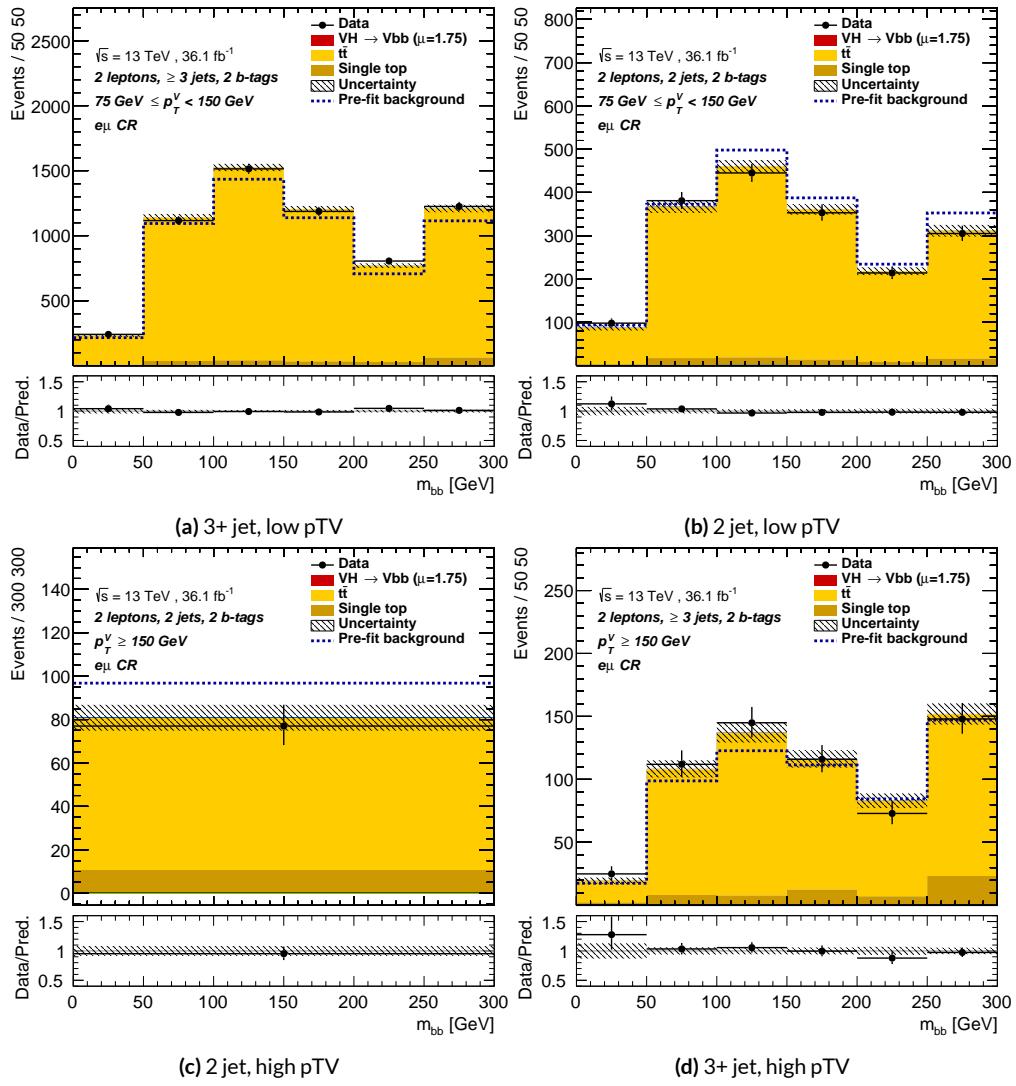


Figure 7.32: Postfit m_{bb} plots in the top $e - \mu$ CR for the standard variable set.

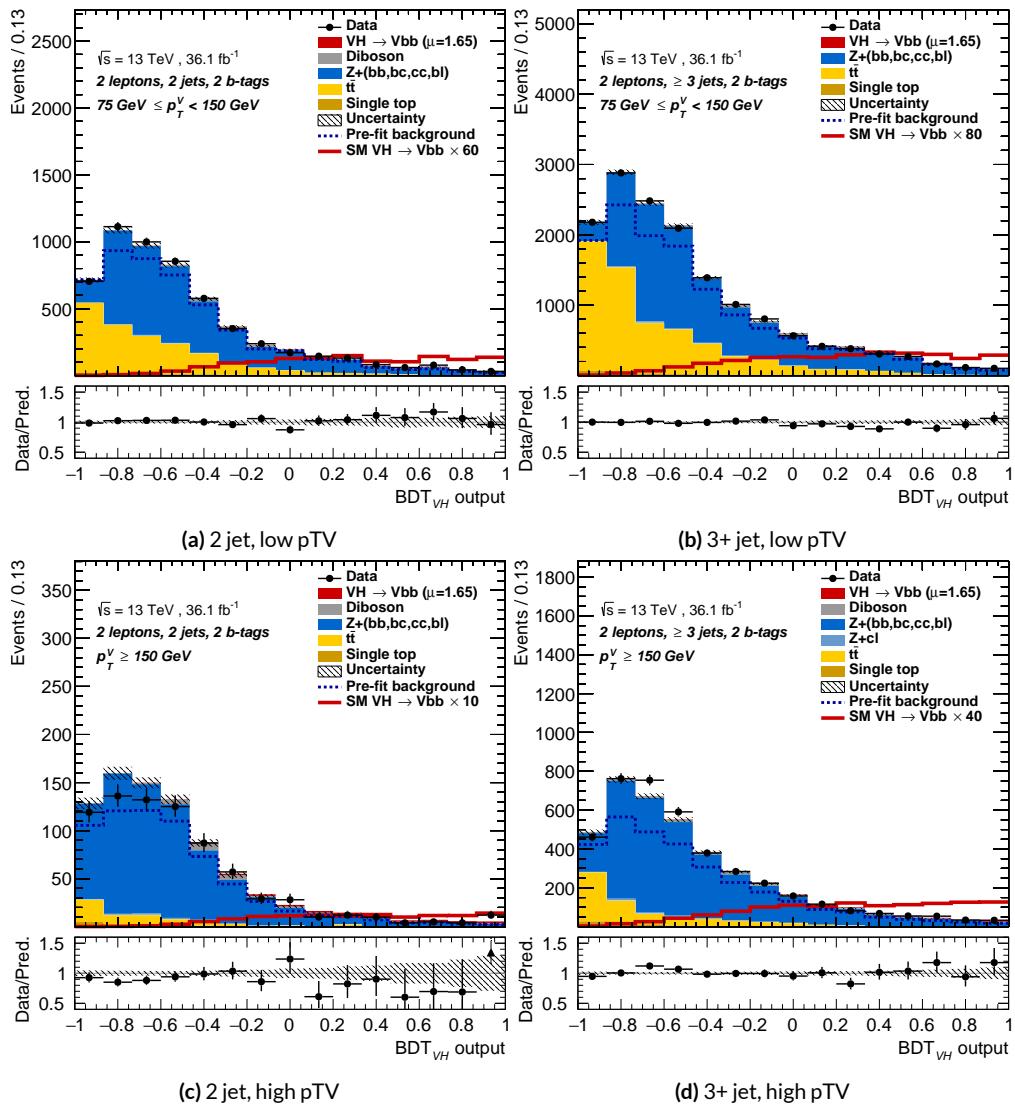


Figure 7.33: Postfit BDT_{VH} plots in the signal region for the LI variable set.

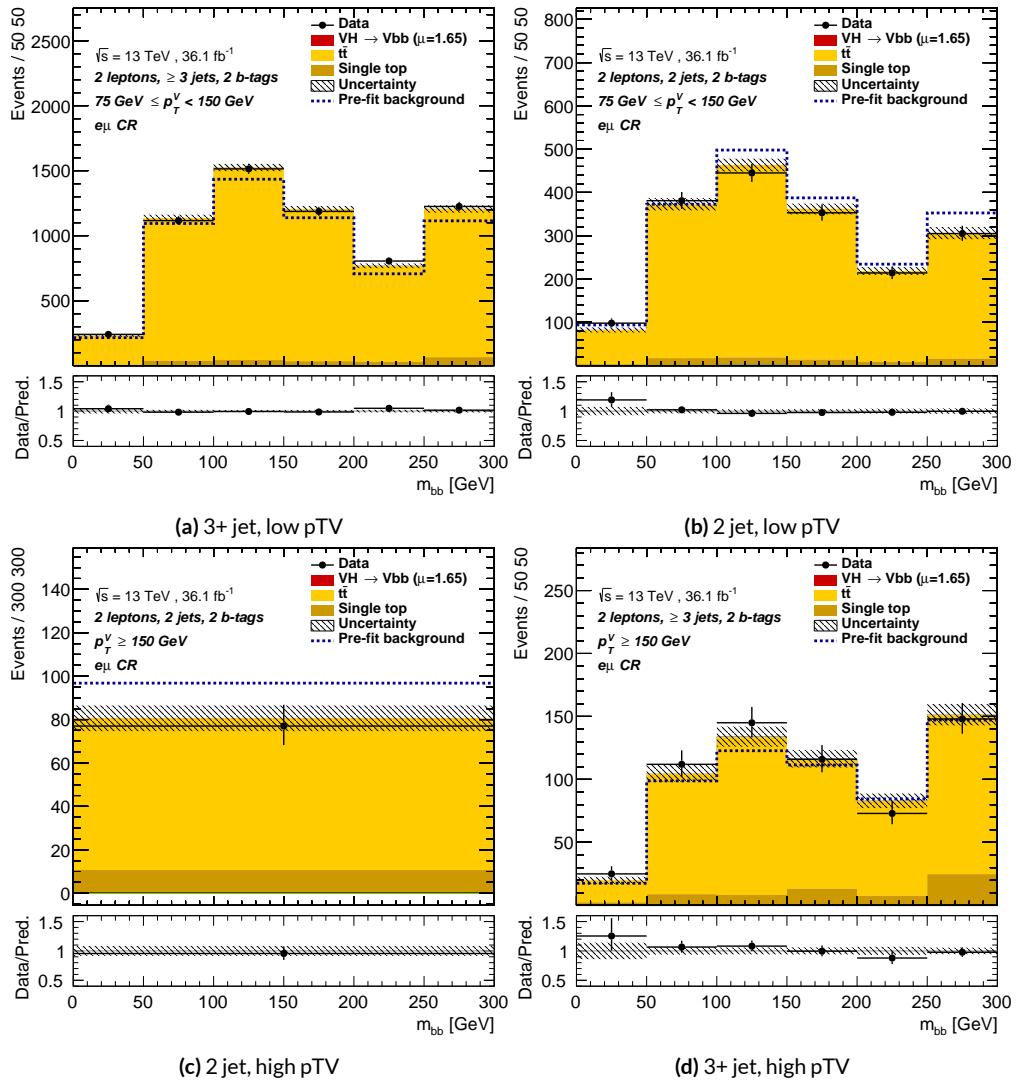


Figure 7.34: Postfit m_{bb} plots in the top $e - \mu$ CR for the LI variable set.

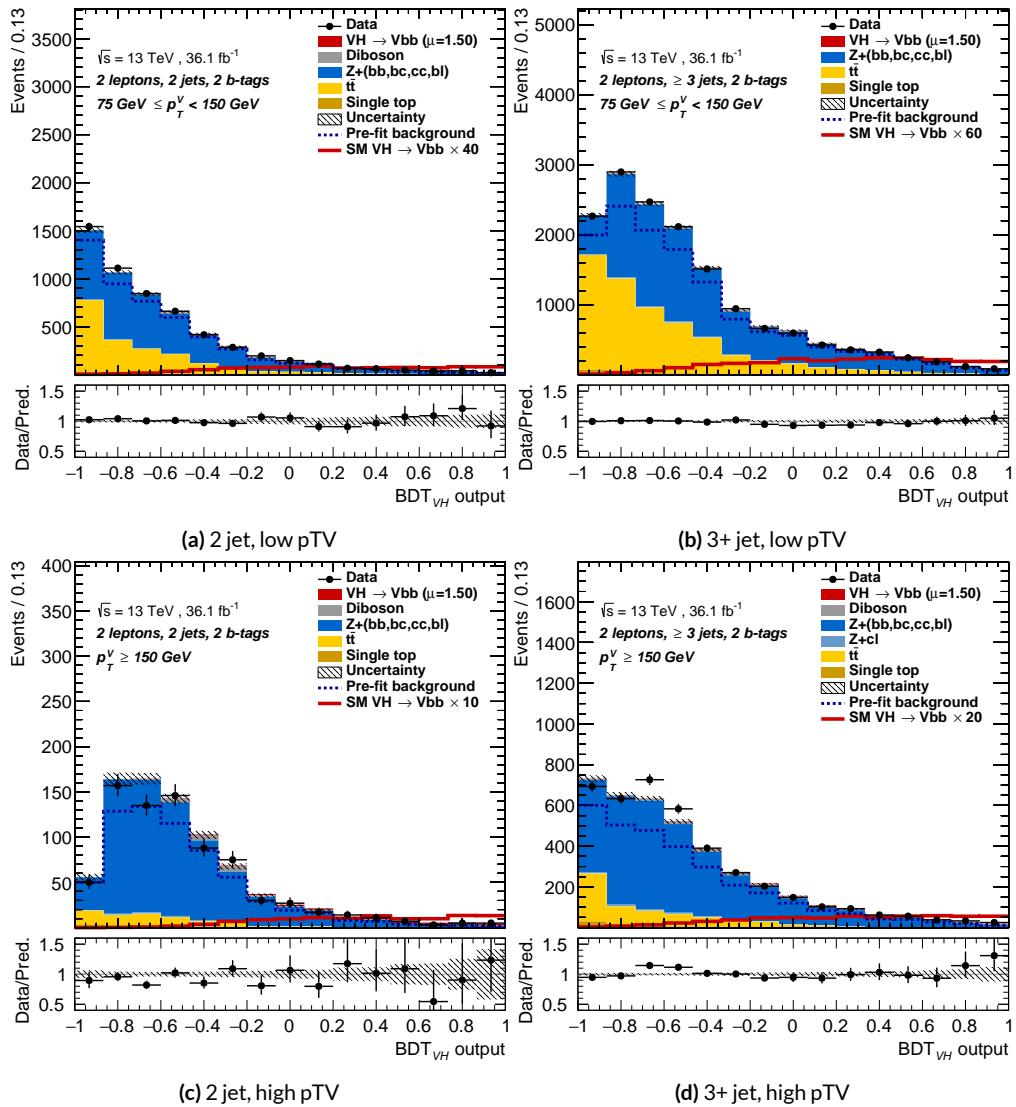


Figure 7.35: Postfit BDT_{VH} plots in the signal region for the RF variable set.

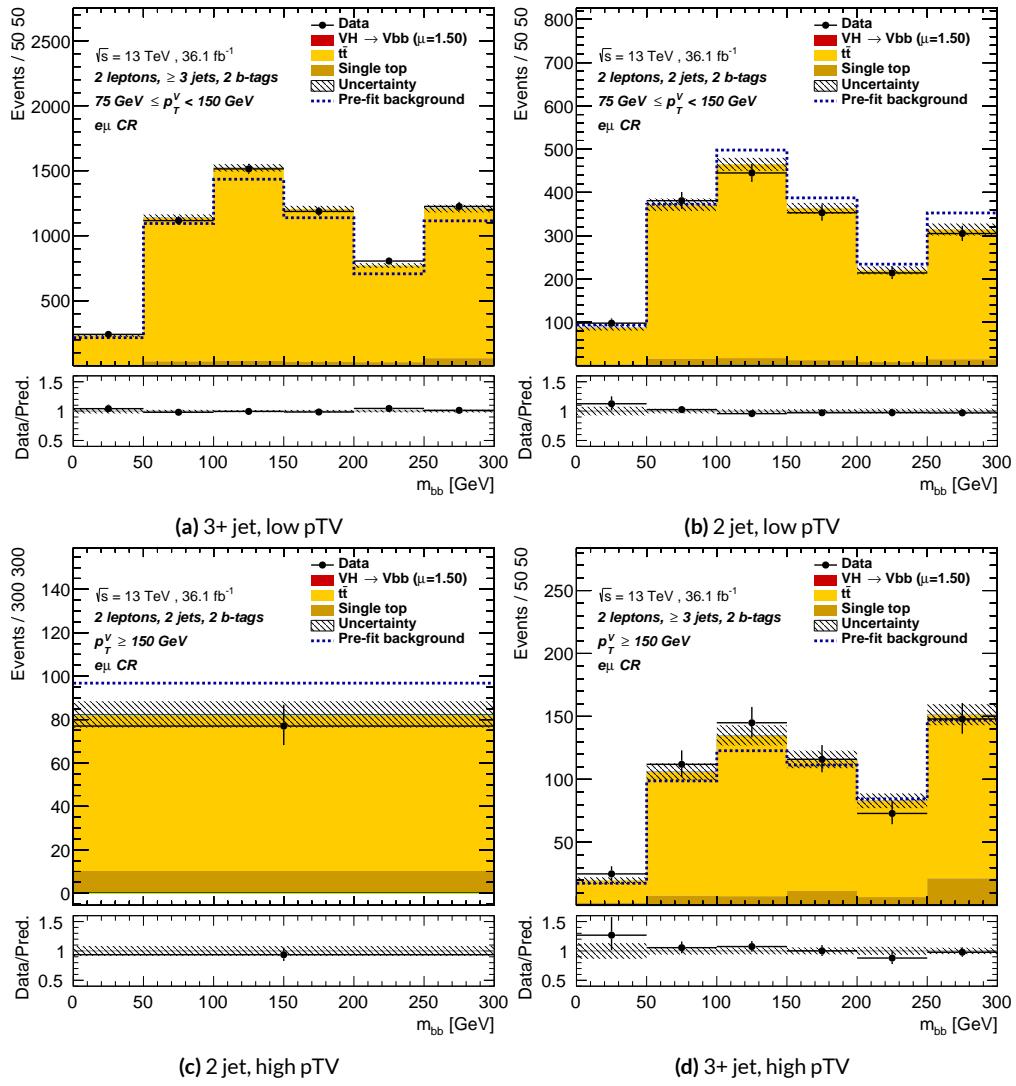


Figure 7.36: Postfit m_{bb} plots in the top $e - \mu$ CR for the RF variable set.

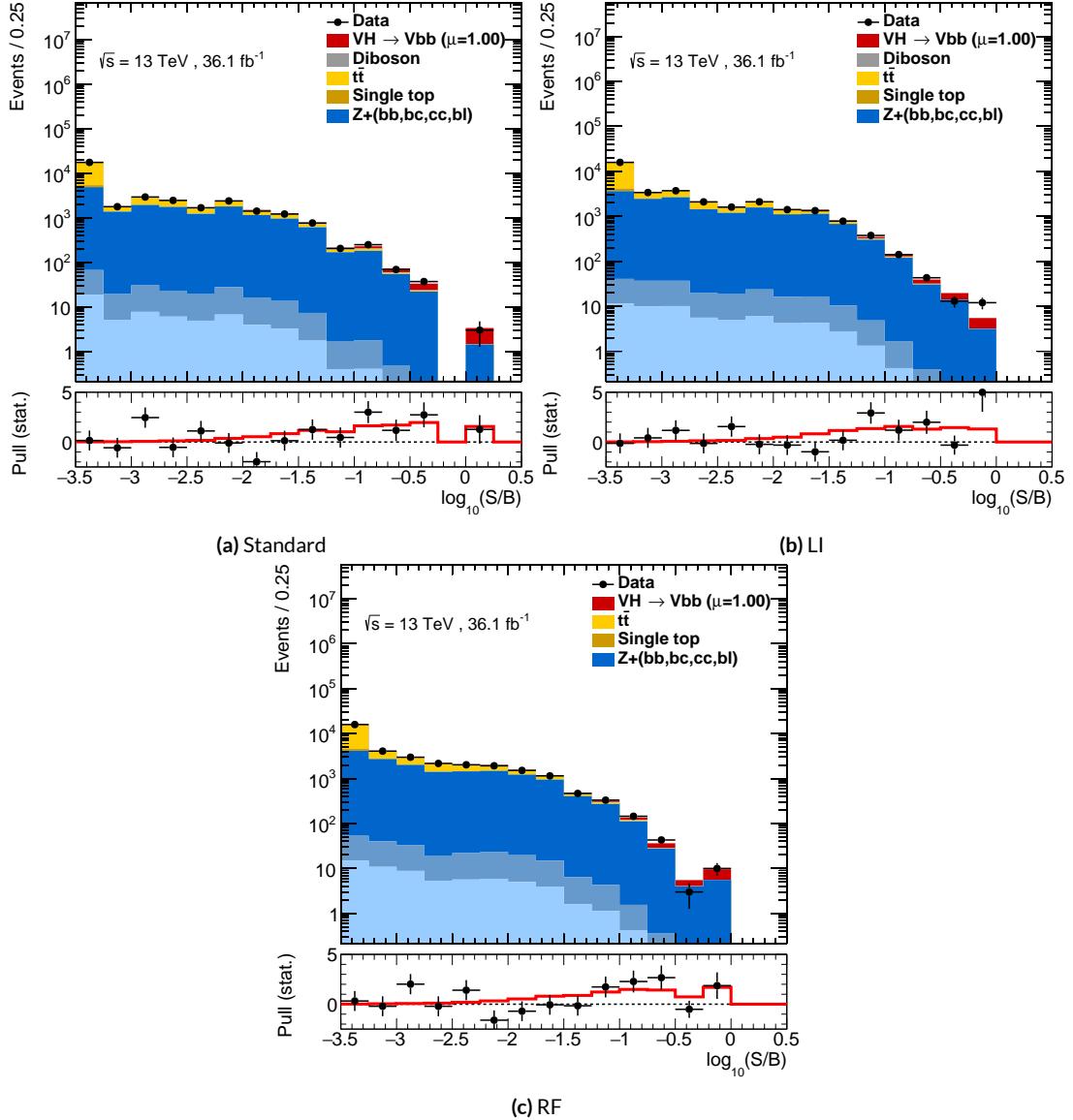


Figure 7.37: Binned S/B plots for the standard (a), LI (b), and RF (c) variable sets. Signal is weighted to $\mu = 1$ for comparison to the SM prediction.

*Kein Operationsplan reicht mit einiger Sicherheit
über das erste Zusammentreffen mit der feindlichen
Hauptmacht hinaus.*

Helmuth von Moltke

8

1874

1875

Fit Results

1876 THE RESULTS IN THIS CHAPTER were first reported in [60] and describe how the three different
1877 fit models detailed and validated Chapter 7, corresponding to the standard, RF, and LI variable sets
1878 described in Chapter 6 perform on actual VH fits. In particular sensitivities, nuisance parameter
1879 impacts, and signal strengths on expected fits to Asimov datasets and both expected and observed

1880 fits on the actual 36.1 fb^{-1} dataset are compared.

1881 Expected and observed sensitivities for the different variable sets may be found in Table 8.1. The
1882 RF fits feature the highest expected sensitivities, outperforming the standard set by 3.5% and 3.4%
1883 for fits to Asimov and observed datasets, respectively. The LI variable has a lower significance than
1884 both for expected fits to both Asimov and data with a 6.7% (1.7%) significance than the standard set
1885 for the Asimov (observed) dataset. While the fit using standard variables does have a higher observed
1886 significance than both the LI and RF fits, by 2.8% and 8.6%, respectively, these numbers should be
1887 viewed in the context of the best fit $\hat{\mu}$ values, discussed below. That is, the standard set may yield the
1888 highest sensitivity for this particular dataset, but this is not necessarily (and likely is not) the case for
1889 any given dataset.

	Standard	LI	RF
Expected (Asimov)	2.06	1.92	2.13
Expected (data)	1.76	1.73	1.80
Observed (data)	2.87	2.79	2.62

Table 8.1: Expected (for both data and Asimov) and observed significances for the standard, LI, and RF variable sets.

1890 A summary of fitted signal strengths and errors for both the Asimov (a) and observed (b) datasets
1891 are shown in Figure 8.1.* A summary of error breakdowns is given in Tables 8.2 (Asimov) and 8.3
1892 (observed) for total error, data statistics contributions, total systematic error contributions, and cat-
1893 egories for which the total impact is ≥ 0.1 for the standard fit. As is to be expected for both the
1894 Asimov and observed dataset fits, the contribution to the total error on μ arising from data statistics

*For reference, the standalone 2-lepton fit from the fiducial analysis is $2.11^{+0.50}_{-0.48}(\text{stat.})^{+0.64}_{-0.47}(\text{syst.})$

¹⁸⁹⁵ is nearly identical, since each set of fits uses the same selections and data.[†]

	Std-KF	LI+MET	RF
Total	+0.608 / -0.511	+0.632 / -0.539	+0.600 / -0.494
DataStat	+0.420 / -0.401	+0.453 / -0.434	+0.424 / -0.404
FullSyst	+0.440 / -0.318	+0.441 / -0.319	+0.425 / -0.284
Signal Systematics	+0.262 / -0.087	+0.272 / -0.082	+0.290 / -0.088
MET	+0.173 / -0.091	+0.140 / -0.074	+0.117 / -0.063
Flavor Tagging	+0.138 / -0.136	+0.069 / -0.071	+0.076 / -0.078
Model Zjets	+0.119 / -0.117	+0.124 / -0.127	+0.095 / -0.086

Table 8.2: Summary of error impacts on total μ error for principal categories in the Asimov standard, LI, and RF fits.

	Std-KF	LI+MET	RF
Total	+0.811 / -0.662	+0.778 / -0.641	+0.731 / -0.612
DataStat	+0.502 / -0.484	+0.507 / -0.489	+0.500 / -0.481
FullSyst	+0.637 / -0.451	+0.591 / -0.415	+0.533 / -0.378
Signal Systematics	+0.434 / -0.183	+0.418 / -0.190	+0.364 / -0.152
MET	+0.209 / -0.130	+0.190 / -0.102	+0.152 / -0.077
Flavor Tagging	+0.162 / -0.166	+0.093 / -0.070	+0.115 / -0.099
Model Zjets	+0.164 / -0.152	+0.141 / -0.143	+0.101 / -0.105

Table 8.3: Summary of error impacts on total $\hat{\mu}$ error for principal categories in the observed standard, LI, and RF fits.

¹⁸⁹⁶ The contribution from systematic uncertainties, however, does vary considerably across the vari-
¹⁸⁹⁷ able sets. The Asimov fits are a best case scenario in the sense that, by construction, all NP's are equal
¹⁸⁹⁸ to their predicted values (and so no "penalty" is paid for pulls on Gaussian NP's). The systematics er-
¹⁸⁹⁹ ror from the LI fit is slightly higher (subpercent) than that from the standard fit, and 4.6% higher er-
¹⁹⁰⁰ ror overall due to differences in data stats. The RF Asimov fit, however, has a 6.5% lower total error

[†]Though not exactly identical. Since the BDT's are different for the different variable sets, the binning (as determined by transformation D) and bin contents in each set are generally different, leading to slightly different data statistics errors.

1901 from systematics than the standard Asimov fit (and a 2.2% lower error overall). Moreover, for both
 1902 the LI and RF sets, errors are markedly smaller for the MET and Flavor Tagging categories, with the
 1903 RF fit also featuring a smaller errors on Z +jets modeling; the only notable exception to this trend in
 1904 Asimov fits are the signal systematics.

1905 These trends are more pronounced in the observed fits. As can be seen in Table 8.3, both the LI
 1906 and RF fits have smaller errors from systematic uncertainties, both overall and in all principal cate-
 1907 gories, with the LI and RF fits having 7.5% (3.7%) and 16% (8.8%) lower systematics (total) error on
 1908 $\hat{\mu}$, respectively.

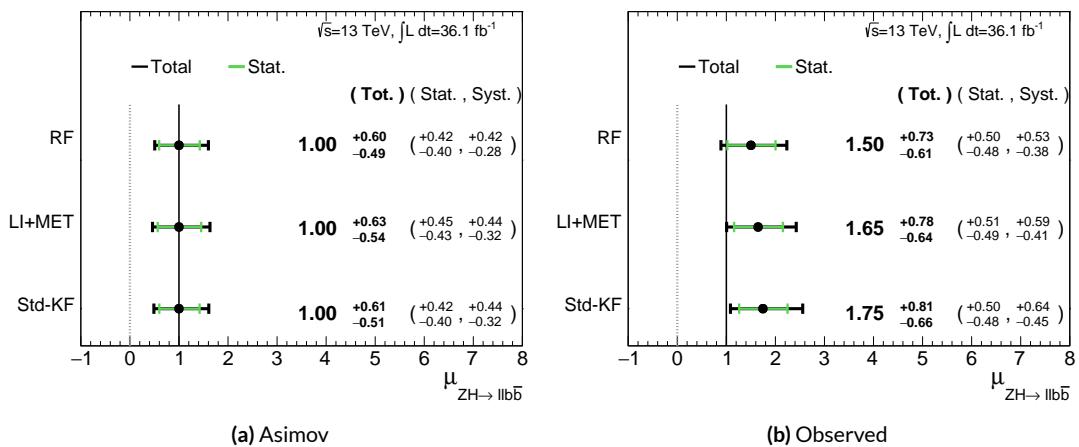


Figure 8.1: μ summary plots for the standard, LI, and RF variable sets. The Asimov case (with $\mu = 1$ by construction) is in (a), and $\hat{\mu}$ best fit values and error summary are in (b).

1909 Studying the performance of the Lorentz Invariant and RestFrames variable sets at both a data
 1910 statistics only context and with the full fit model in the $ZH \rightarrow \ell\ell b\bar{b}$ channel of the $VH(b\bar{b})$ anal-
 1911 ysis suggests that these variables may offer a potential method for better constraining systematic un-
 1912 certainties in $VH(b\bar{b})$ searches as more orthogonal bases in describing the information in collision

₁₉₁₃ events.

₁₉₁₄ The marginally worse performance of the LI and RF variables (7.9% and 6.9%, respectively) with
₁₉₁₅ respect to the standard variable at a stats only level illustrates that neither variable set has greater
₁₉₁₆ intrinsic descriptive power in the absence of systematics in this closed final state. Hence, any gains
₁₉₁₇ from either of these variable sets in a full fit come from improved treatment of systematic uncertain-
₁₉₁₈ ties.

₁₉₁₉ With full systematics, the LI variable set narrows the sensitivity gap somewhat, with lower signif-
₁₉₂₀ icances by 6.7% (1.7%) on expected fits to Asimov (data) and by 2.8% on observed significances. The
₁₉₂₁ RF variable set outperforms the standard set in expected fits with 3.5% (3.4%) higher significance
₁₉₂₂ on Asimov (data), but has an 8.6% lower observed significance, though the observed significances
₁₉₂₃ should be viewed in the context of observed $\hat{\mu}$ values.

₁₉₂₄ Moreover, the LI and RF variable sets generally perform better in the context of the error on μ .
₁₉₂₅ The LI fit is comparable to the standard set on Asimov data and has a 7.5% lower total systematics er-
₁₉₂₆ ror on $\hat{\mu}$ on observed data, while the RF fit is lower in both cases, with systematics error being 6.5%
₁₉₂₇ (16%) lower on Asimov (observed) data.

₁₉₂₈ These figures of merit suggest that both the LI and RF variables are more orthogonal than the
₁₉₂₉ standard variable set used in the fiducial analysis. Moreover, the RF variable set does seem to con-
₁₉₃₀ sistently perform better than the LI set. Furthermore, both variable sets have straightforward exten-
₁₉₃₁ sions to the other lepton channels in the $VH(b\bar{b})$ analysis. The magnitude of any gain from the
₁₉₃₂ more sophisticated treatment of E_T^{miss} in these extensions is beyond the scope of these studies, but
₁₉₃₃ the performance in this closed final state do suggest that there is some value to be had in these non-

¹⁹³⁴ standard descriptions independent of these considerations.

*If I have seen further, it is by standing on ye shoulders of
giants.*

Isaac Newton

9

1935

1936

Measurement Combinations

¹⁹³⁷ WHILE THE DISCUSSION thus far has focused on improvements looking towards future in just the

¹⁹³⁸ $ZH \rightarrow \ell\ell b\bar{b}$ channel, any actual result for SM $VH(b\bar{b})$ combines all channels and all available

¹⁹³⁹ datasets. Using additional channels at a given center of mass energy is straightforward since the fit

¹⁹⁴⁰ model is designed with this combination in mind. This will be described in the context of the 36.1

1941 fb^{-1} 13 TeV result in Section 9.1.

1942 Combining dataset results (known as “workspaces”) from different center of mass energies is not
1943 so simple an exercise since both the underlying physics (and its associated modeling) and the treat-
1944 ment of key experimental considerations, like flavor tagging, and their associated systematics change
1945 from dataset to dataset. A combined fit model must take these considerations into account, and the
1946 formulation of the fit model combining the Run 1 ($\sqrt{s} = 7 \text{ TeV}$ with 4.7 fb^{-1} of data, and $\sqrt{s} = 8$
1947 TeV with 20.3 fb^{-1} of data) and Run 2 ($\sqrt{s} = 13 \text{ TeV}$ with 36.1 fb^{-1}) SM $VH(b\bar{b})$ results is the topic
1948 of Section 9.2. Its results, as reported in [3], are given in 9.3.

1949 9.1 LEPTON CHANNEL COMBINATIONS

1950 Preparation of results for the o- and i-lepton channels is functionally very similar to above discus-
1951 sions in Chapters 4, 5, 6, and 7. From a modeling standpoint, each channel comes in with different
1952 dominating background processes and dedicated simulation, described at length in [21], though
1953 there is a lot of overlap. In particular, $t\bar{t}$, , and diboson production is important for all three chan-
1954 nels. The only important process not discussed here is contribution from multijet background,
1955 which is a small but important background in the i-lepton case.

1956 With respect to object definitions, no new objects are defined in the o- and i-lepton analyses,
1957 though the treatment of \vec{E}_T^{miss} is of greater concern in these channels, as \vec{E}_T^{miss} is a part of the signal
1958 final states in these channels. Triggers and event selection requirements are optimized by channel. A
1959 full list of requirements is given in Table 9.1 from [4].

1960 The mechanics of MVA training and implementation is very much the same across analysis chan-

Common Selections	
Jets	≥ 2 central jets
b -jets	2 b -tagged signal jets
Leading jet p_T	> 45 GeV
$ \Delta R(\text{jet1}, \text{jet2}) $ (cut-based only)	$\leq 1.8 (p_T^V < 200 \text{ GeV}), \leq 1.2 (p_T^V > 200 \text{ GeV})$
o Lepton	
Trigger	HLT_xe70, xe90_mht, and xe110_mht
Leptons	o VH-loose lepton
\vec{E}_T^{miss}	> 150 GeV
S_T	> 120 (2 jets), > 150 GeV (3 jets)
$ \min \Delta\phi(\vec{E}_T^{\text{miss}}, \text{jet}) $	$> 20^\circ$ (2jet), $> 30^\circ$ (3jet)
$ \Delta\phi(\vec{E}_T^{\text{miss}}, b) $	$> 120^\circ$
$ \Delta\phi(\text{jet1}, \text{jet2}) $	$< 140^\circ$
$ \Delta\phi(\vec{E}_T^{\text{miss}}, E_{T, \text{trk}}^{\text{miss}}) $	$< 90^\circ$
p_T^V regions (BDT)	> 150 GeV
p_T^V regions (cut-based)	$[150, 200]$ GeV, $[200, \infty]$ GeV
1 Lepton	
Trigger	e channel: un-prescaled single electron Tables 5 and 6 of Ref. [44]
Leptons	μ channel: see o-lepton triggers 1 WH-signal lepton
\vec{E}_T^{miss}	> 1 VH-loose lepton veto
m_{top}	> 30 GeV (e channel)
m_T^W (cut-based only)	< 225 GeV or $m_{bb} > 75$ GeV
p_T^V regions (BDT)	< 120 GeV
p_T^V regions (cut-based)	> 150 GeV $[150, 200]$ GeV, $[200, \infty]$ GeV
2 Lepton	
Trigger	un-prescaled single lepton Tables 5 and 6 of Ref. [44]
Leptons	2 VH-loose leptons (≥ 1 ZH-signal lepton)
$m_{\ell\ell}$	Same flavor, opposite-charge for $\mu\mu$ $81 < m_{\ell\ell} < 101$ GeV
\vec{E}_T^{miss} significance (cut-based)	$\vec{E}_T^{\text{miss}}/\sqrt{H_T} < 3.5\sqrt{\text{GeV}}$
p_T^V regions (BDT)	$[75, 150], [150, \infty]$ GeV
p_T^V regions (cut-based)	$[75, 150], [150, 200], [200, \infty]$ GeV

Table 9.1: Summary of the signal event selection in the 0-, 1- and 2-lepton analyses.

1961 nels, with the major difference being the selection of input variables to the BDT discriminants. For
1962 a discussion of how the different final states affect variable selection see the discussion in Section 1.7
1963 and in particular Table for the input variables used in the final analysis.

1964 As previously mentioned, the fit model is flexible enough to seamlessly integrate combined results
1965 for the three separate lepton channels for a given dataset. Most nuisance parameters are treated as
1966 common across all fit regions. Some regions will have greater bearing on certain nuisance parameters—
1967 2-lepton regions, virtually free of $W+jets$ events, will have virtually no effect on $W+jets$ modeling
1968 systematics, for example. One notable exception are NP's with `_L[012]` suffixes, which are pre-
1969 dominantly the double ratio systematics discussed in Section 4.2.2 and function similar to the 2 vs.
1970 greater than 3 jet event double ratio systematics.

1971 Adding different channels has great potential to constrain certain systematic uncertainties. Look-
1972 ing at the breakdown of systematic uncertainties in 2-lepton fits in Table 7.8, for example, multijet
1973 and $W+jets$ NP's contribute virtually no uncertainty, while \vec{E}_T^{miss} and $Z+jets$ have very high impacts
1974 on the uncertainty on $\hat{\mu}$. 1-lepton events will bring up the multijet and $W+jets$ uncertainties (and
1975 justify their inclusion in the combined fit model; their inclusion in the 2-lepton fit is for portabil-
1976 ity and a sanity check). Since the 2-lepton final state is by construction \vec{E}_T^{miss} free, it is not surpris-
1977 ing that uncertainty due to \vec{E}_T^{miss} is high. Single channel standalone fits are never final results in this
1978 analysis, so the result relies on the other channels to better and more accurately constrain this uncer-
1979 tainty since these other final states do have \vec{E}_T^{miss} in their final states and are the channels for which
1980 the \vec{E}_T^{miss} treatment in the analysis has been optimized. Furthermore, something like $Z+hf$ modeling
1981 is difficult to do given how constrained the analysis signal region is. Combining the information in

¹⁹⁸² the o- and 2-lepton results (and introducing double ratio NP's to recognize that these channels do
¹⁹⁸³ have important differences) also helps to constrain this difficult systematic uncertainty.

¹⁹⁸⁴ Once the fit inputs in each channel have been prepared and validated, a combined workspace can
¹⁹⁸⁵ be directly constructed using the combined fit model. Significances are given in Table 9.2, and $\hat{\mu}$
¹⁹⁸⁶ summaries for 2 and 3 POI fits are given in Figure 9.1. This combined workspace with observed sig-
¹⁹⁸⁷ nificant strength of $1.20^{+0.24}_{-0.23}$ (stat.) $^{+0.34}_{-0.28}$ (syst.) is the Run 2 input for the Run 1 + Run 2 combination
¹⁹⁸⁸ discussed below.

Channel	Exp. sig. (Asimov)	Exp. sig. (data)	Obs. sig.
o-lepton (SR)	1.99	1.73	0.53
1-lepton (SR+CR)	1.81	1.81	2.30
2-lepton (SR+CR)	1.95	1.86	3.55
o,1,2-lepton (SR+CR)	3.19	3.03	3.54

Table 9.2: Observed significance from an unconditional fit to the data corresponding to 36.1 fb^{-1} and expected significances from a fit to an Asimov dataset and from a fit to the data. Expected significances from individual regions are estimated separately.

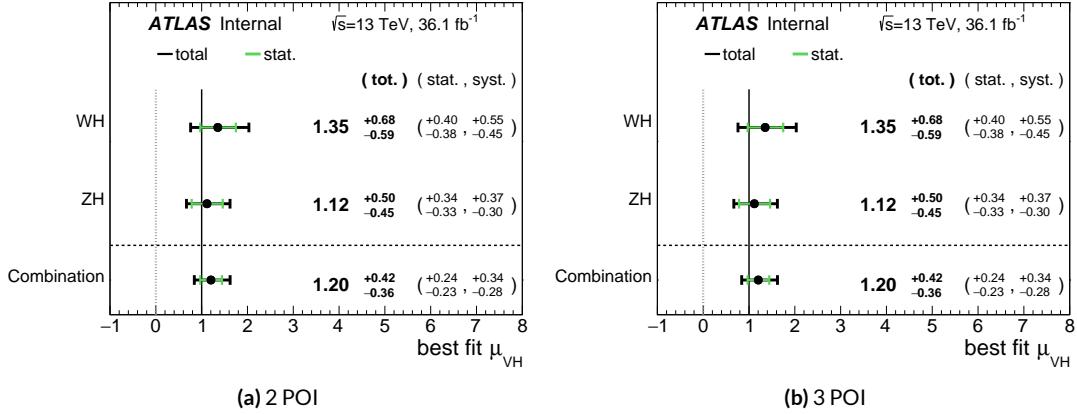


Figure 9.1: Run 2 signal strength summary plots for 2 (WH/ZH, (a)) and 3 (0, 1, and 2 lepton, (b)) POI fits.

1989 9.2 THE COMBINED FIT MODEL

1990 It is clear the signal strength parameter of interest should be fully correlated among the different
1991 datasets. Some signal modeling systematics were left unchanged from Run 1 through Run 2 and/or
1992 were designed to be explicitly correlated. Beyond these two special cases, it is not immediately clear
1993 what level of correlation should be imposed. The general methodology for settling upon a correla-
1994 tion scheme is as follows:

- 1995 1. Identify which NP categories have significant impacts on μ
- 1996 2. Of these NP's, identify which have one-to-one correspondences or established correlation
1997 schemes among \sqrt{s} values
- 1998 3. Test whether correlation has a sizable impact on expected fit quantities

1999 The only two sizable experimental NP categories are jet energy scale (JES) and flavor tagging sys-
2000 tematics. Correlation schemes of varying degrees of completeness exist for these categories, so ex-
2001 plicit NP correlations can be tested for these two categories. As these studies were conducted before
2002 unblinding, “sizable impact” was judged by comparing fit results (sensitivities, pull comparisons,
2003 and breakdowns) on combined workspaces using the unblinded and public $\mu = 0.51$ result for
2004 Run 1 and Asimov data for the Run 2 result. These are treated in Sections 9.2.1 and 9.2.2. Modeling
2005 systematics require a slightly different treatment, and are explored in 9.2.3.

2006 As noted in Chapter 7 when looking at pull comparison plots for combined workspaces, the error
2007 bars in these plots are calculated using a simultaneous HESSE matrix inversion, which can fail to give
2008 sensible values for high dimensional models (the combined workspaces have well over 500 NP's).
2009 This is not true of the nuisance parameter ranking plots, which use a MINOS based approach to test

2010 the effect of each NP individually. This is much slower but much more rigorous, which is why only
2011 ranking plots appear outside of supporting material and pull comparisons are considered “diagnos-
2012 tic” plots.

2013 9.2.1 JET ENERGY SCALE SYSTEMATICS

2014 Fortunately for the case of jet energy scale systematics, the JetEtMiss group provides two recom-
2015 mended “strong” and “weak” correlation schemes between Run 1 and Run 2. These were used as
2016 a point of departure for the JES combination correlation scheme. However, the JES NP’s in both
2017 the Run 1 and Run 2 workspaces are a reduced set of NP’s, with some 56 (75) NP’s reduced to 6 (8)
2018 for Run 1 (2). In order to restore the full set of JES NP’s, the effective NP’s in each workspace are un-
2019 folded using maps detailing the linear combinations of unfolded NP’s that form the effective NP’s.

2020 The linear combinations used to unfold the effective JES NP’s were calculated as follows:

$$NP_{i,eff} = \frac{\sum_j A_{ij} |NP_{j,unf}| NP_{j,unf}}{\sqrt{\sum_j A_{ij}^2 |NP_{j,unf}|^2}} \quad (9.1)$$

2021 where *eff* and *unf* are for effective and unfolded NP’s, respectively, the A_{ij} ’s are scalar coefficients
2022 taken from raw maps, and $|NP_{j,unf}|$ are the amplitudes of the unfolded NP’s. The raw A_{ij} and scaled
2023 maps for Run 1 and Run 2 may be found in Figure 9.2
2024 Unfolding was found to have very little effect on both expected sensitivities and errors, as can be
2025 seen in Tables 9.3– 9.6.

2026 It was also found that fit sensitivities and breakdowns were similarly indifferent to the use of ei-

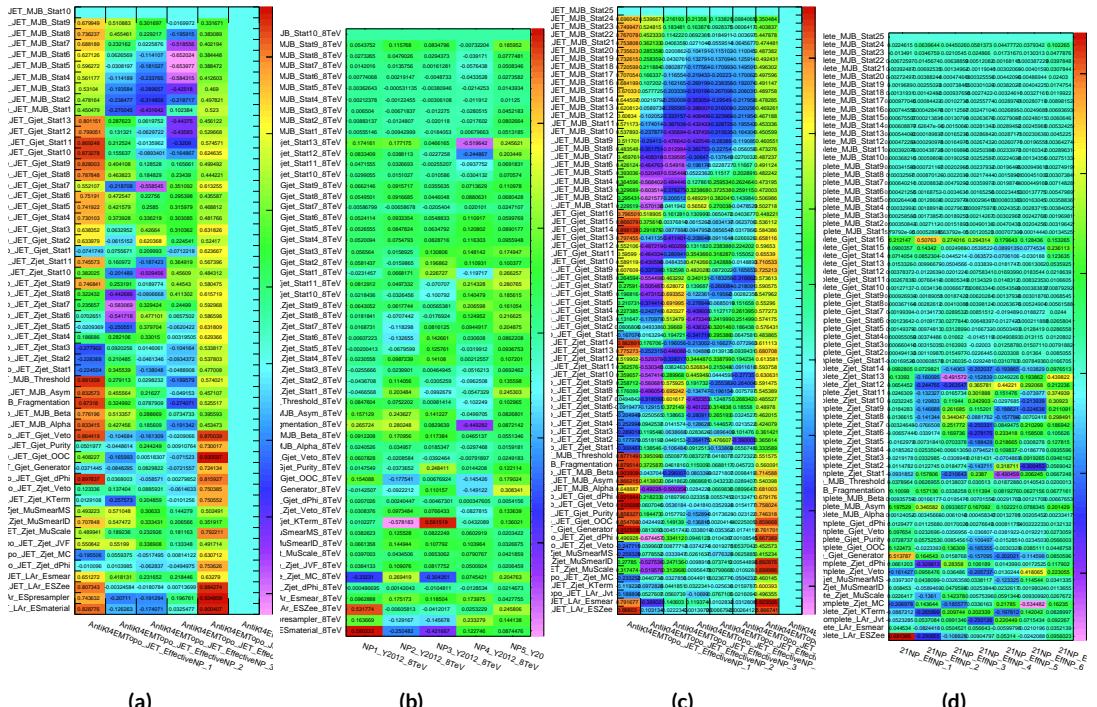


Figure 9.2: The raw and scaled coefficients for unfolding Run 1 (a and b) and Run 2 (c and d), respectively

	R ₁ Unf	R ₁ Eff	R ₂ Unf	R ₂ Eff	Comb Unf	Comb Eff
Exp. Sig.	2.604	2.606	3.014	3.014	4.005	3.998
Obs. Sig.	1.369	1.374	3.53	3.53	3.581	3.571
Exp. Limit	0.76 ^{+0.30} _{-0.21}	0.76 ^{+0.30} _{-0.21}	0.73 ^{+0.29} _{-0.21}	0.73 ^{+0.29} _{-0.21}	0.51 ^{+0.20} _{-0.14}	0.51 ^{+0.20} _{-0.14}
Obs. Limit	1.21	1.21	1.94	1.94	1.36	1.37

Table 9.3: Expected and observed sensitivities for Run 1, Run 2, and combined workspaces with effective and unfolded JES NP's.

	R ₁ Unfold	R ₁ Eff
$ \Delta\hat{\mu} $		0.0018
$\hat{\mu}$	0.5064	0.5082
Total	+0.400 / -0.373	+0.401 / -0.373
DataStat	+0.312 / -0.301	+0.312 / -0.301
FullSyst	+0.250 / -0.220	+0.251 / -0.220
Jets	+0.060 / -0.051	+0.060 / -0.052
BTag	+0.094 / -0.079	+0.095 / -0.079

Table 9.4: Error on signal strength breakdowns for Run 1 workspaces with effective and unfolded JES NP's.

	R ₂ Unfold	R ₂ Eff
$ \Delta\hat{\mu} $		0.0
$\hat{\mu}$	1.2051	1.2052
Total	+0.421 / -0.366	+0.421 / -0.366
DataStat	+0.239 / -0.234	+0.239 / -0.234
FullSyst	+0.346 / -0.282	+0.346 / -0.282
Jets	+0.066 / -0.047	+0.066 / -0.047
BTag	+0.119 / -0.106	+0.119 / -0.106

Table 9.5: Error on signal strength breakdowns for Run 2 workspaces with effective and unfolded JES NP's.

	Comb Unfold	Comb Eff
$ \Delta\hat{\mu} $		0.0006
$\hat{\mu}$	0.8992	0.8985
Total	+0.278 / -0.261	+0.278 / -0.261
DataStat	+0.185 / -0.181	+0.185 / -0.181
FullSyst	+0.208 / -0.187	+0.208 / -0.188
Jets	+0.040 / -0.044	+0.041 / -0.036
BTag	+0.076 / -0.076	+0.077 / -0.076

Table 9.6: Error on signal strength breakdowns for combined workspaces with effective and unfolded JES NP's.

2027 ther the strong or weak JES correlation schemes, as shown in Tables 9.7 and 9.8.

	JES Weak Unfold	JES Weak Eff	JES Strong Unfold	JES Strong Eff
Exp. Sig.	3.57	3.57	3.59	3.59
Exp. Limit	$0.493^{+0.193}_{-0.138}$	$0.494^{+0.193}_{-0.138}$	$0.493^{+0.193}_{-0.138}$	$0.493^{+0.193}_{-0.138}$

Table 9.7: Expected sensitivities for both effective and unfolded combined workspaces using the strong and weak JES correlation schemes.

	Comb Unfold	Comb Eff	Strong Unfold	Strong Eff
$\Delta\hat{\mu}$	0.0009		0.0025	
Total	+0.269 -0.254	+0.27 -0.255	+0.27 -0.255	+0.27 -0.255
DataStat	+0.181 -0.177	+0.181 -0.177	+0.181 -0.177	+0.181 -0.178
FullSyst	+0.199 -0.183	+0.2 -0.183	+0.2 -0.183	+0.201 -0.183
Jets	+0.0387 -0.032	+0.041 -0.0337	+0.0425 -0.0329	+0.0432 -0.0338
BTag	+0.0975 -0.0933	+0.098 -0.0936	+0.0979 -0.0935	+0.098 -0.0936

Table 9.8: Error on signal strength breakdowns for both effective and unfolded combined workspaces using the strong and weak JES correlation schemes.

2028 Comparisons of top ranked nuisance parameters in Figures 9.3–9.5 and for the complete JES pull
 2029 comparisons in Figures 9.6–9.9 also show very little difference with respect to correlation scheme
 2030 (except obviously for the number of JES NP’s). Constrained pulls in pull comparisons should once
 2031 again be taken as a shortcoming of HESSE and not the fit model.

2032 As a result of these studies, the weak JES correlation scheme with uncorrelated effective JES NP’s
 2033 (i.e. just the b -jet energy scale NP) has been chosen as the treatment of JES in the Run 1 + Run 2
 2034 combined fit.

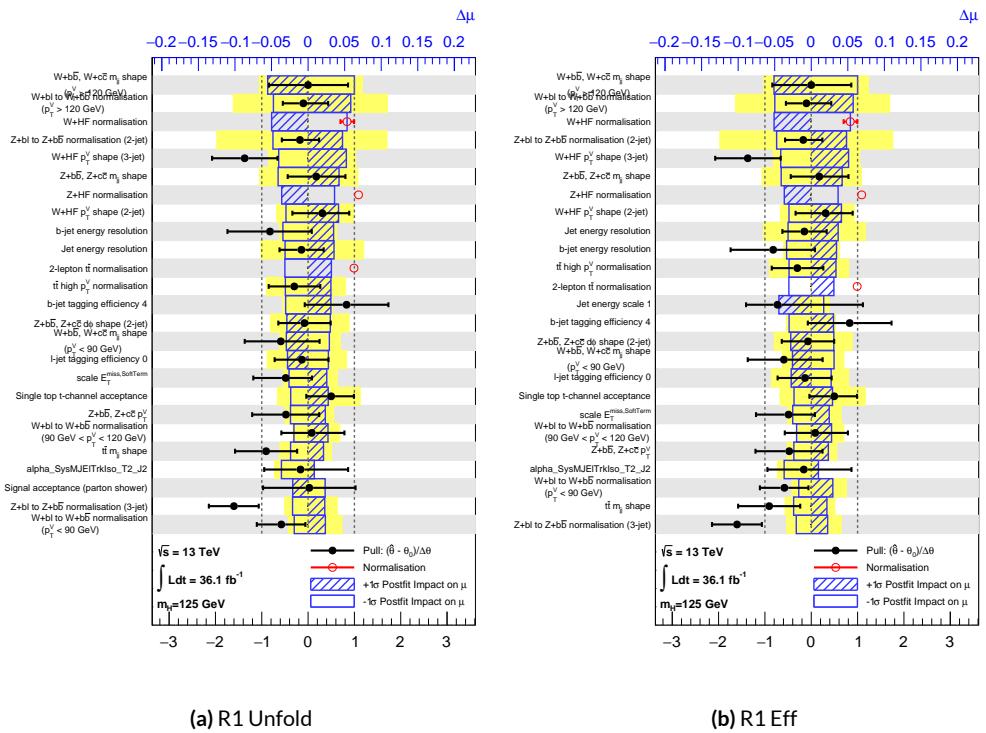


Figure 9.3: Ranks for the effective and unfolded JES NP Run1 combined workspaces.

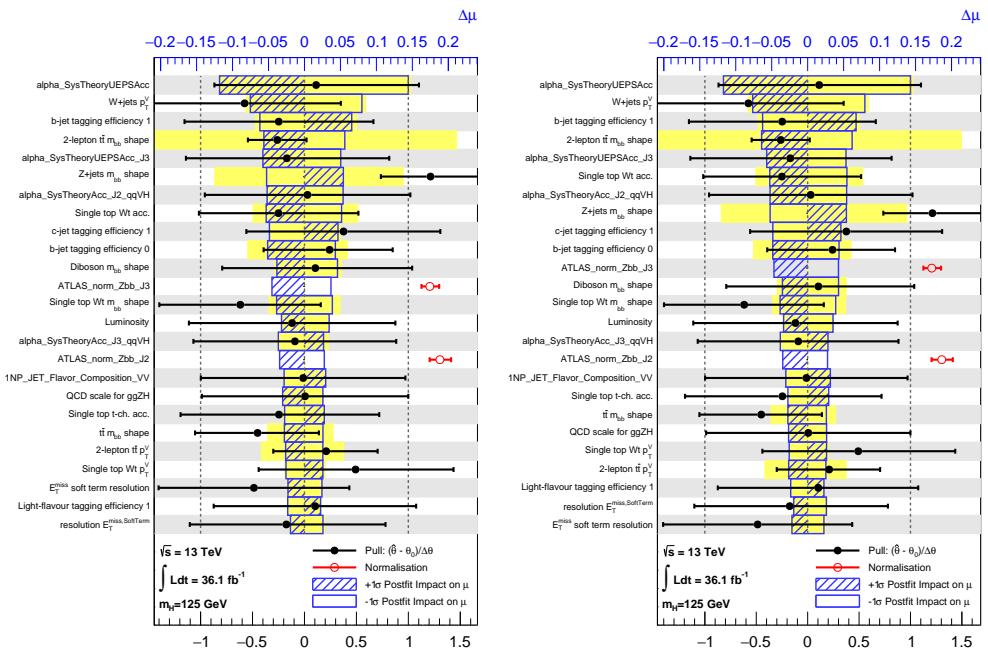


Figure 9.4: Ranks for the effective and unfolded JES NP Run2 combined workspaces.

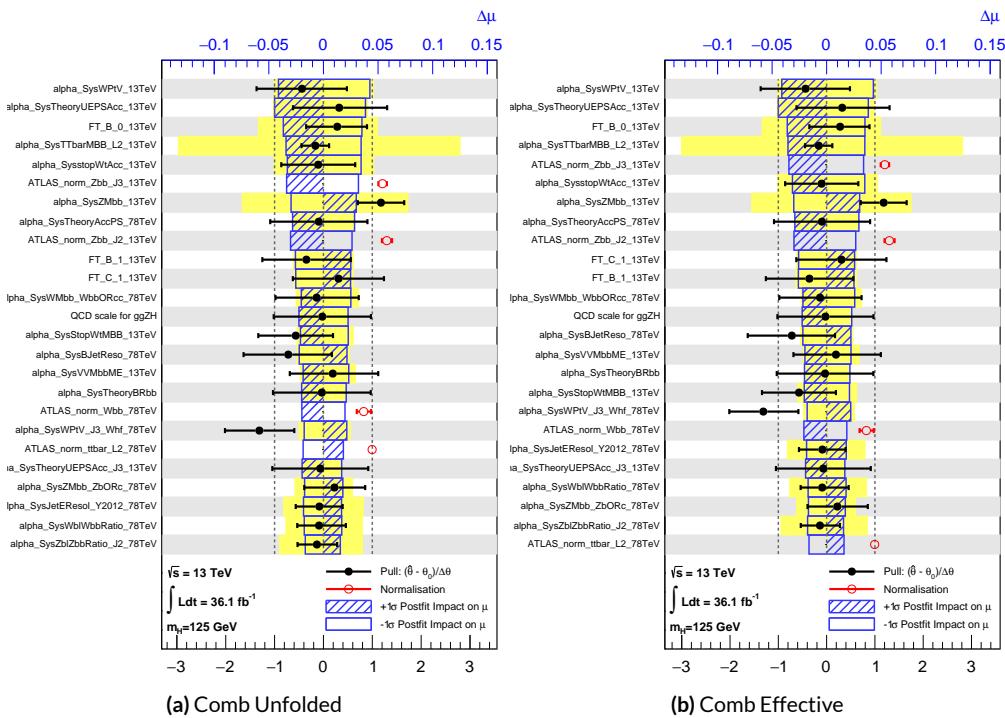


Figure 9.5: Ranks for the effective and unfolded JES NP Run1+Run2 combined workspaces.

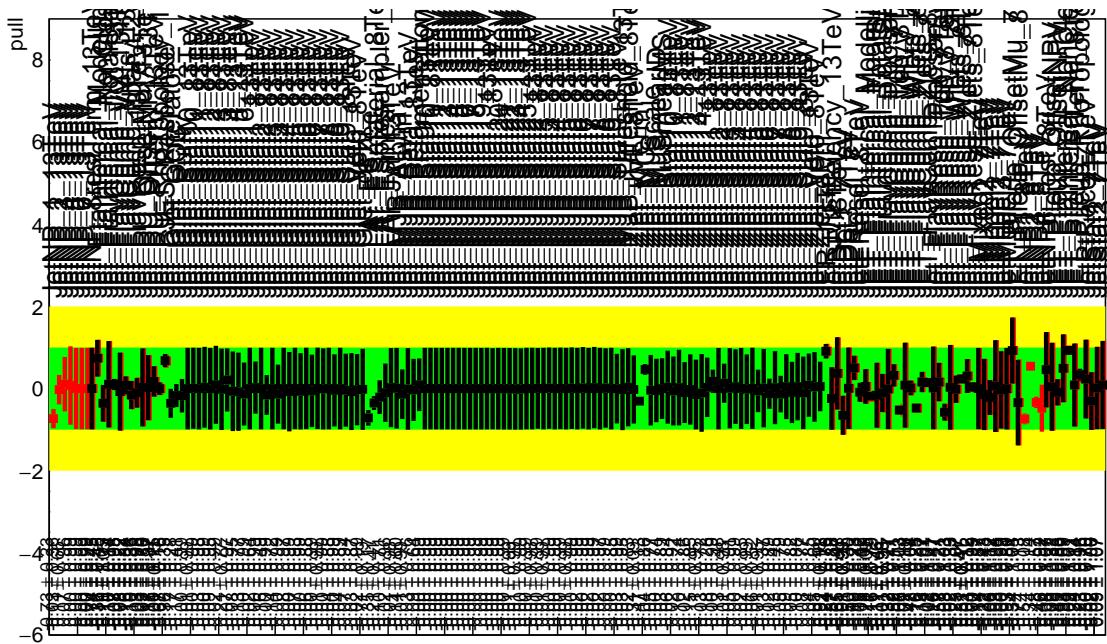


Figure 9.6: Pull Comparisons: jesu---Jet Comb Unfold, Comb Eff, Strong Unfold, Strong Eff

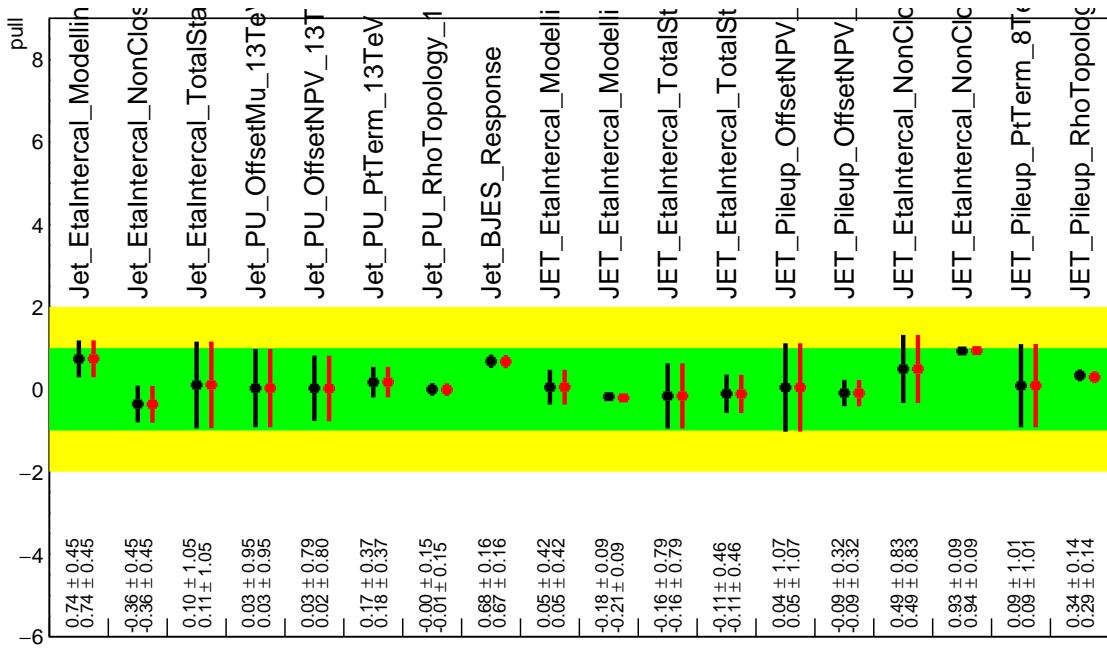


Figure 9.7: Pull Comparisons: jesu---JetMatched Comb Unfold, Comb Eff, Strong Unfold, Strong Eff

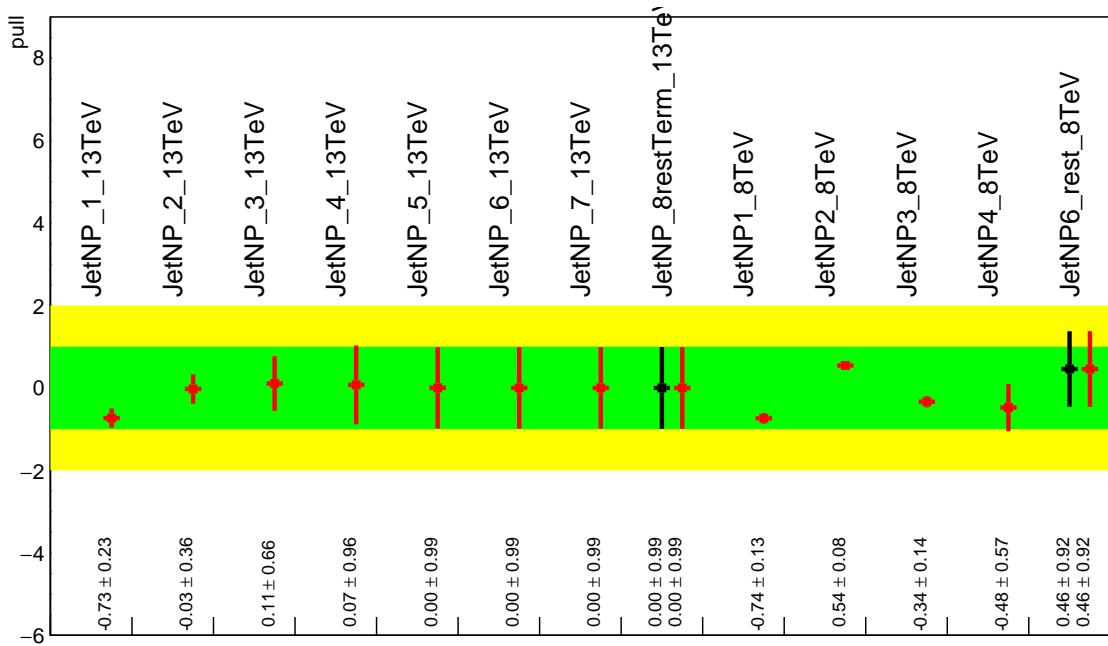


Figure 9.8: Pull Comparisons: jesu---JetEff Comb Unfold, **Comb Eff**, Strong Unfold, **Strong Eff**

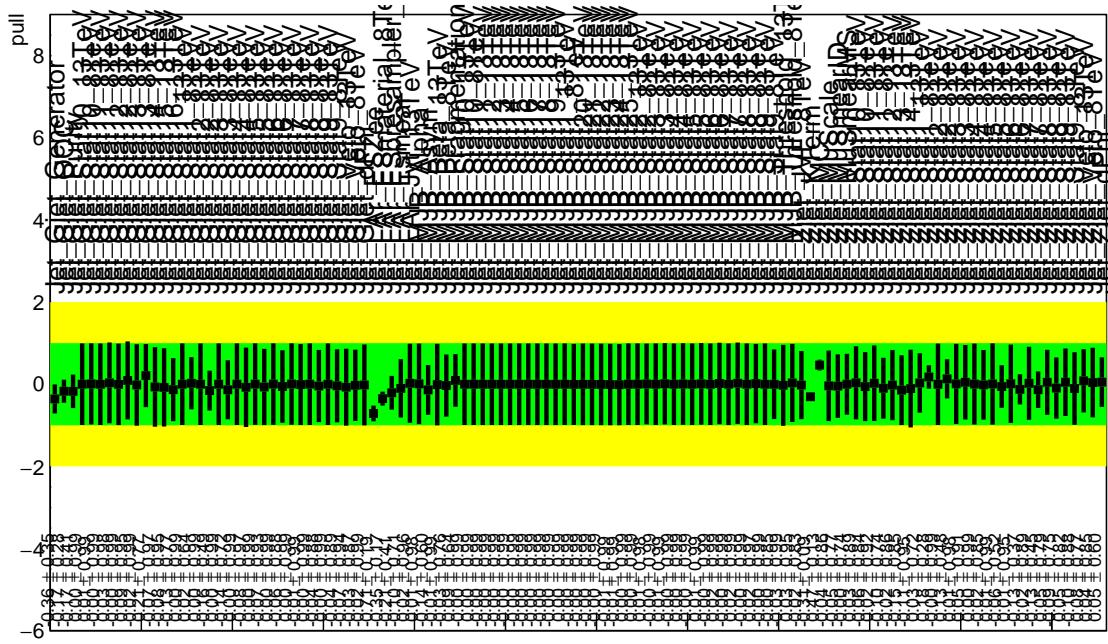


Figure 9.9: Pull Comparisons: jesu---JetUnfold Comb Unfold, **Comb Eff**, Strong Unfold, **Strong Eff**

2035 9.2.2 FLAVOR TAGGING

2036 Unfortunately, the ATLAS Flavor Tagging group did not provide any recommendations for corre-
2037 lating Run 1 and Run 2 NP's, though given the high ranking of these NP's in the Run 2, result, per-
2038 forming at least some studies was deemed crucial. Nevertheless, great improvements and changes to
2039 the treatment of flavor tagging between Run 1 and Run 2 does weaken the argument for any strong
2040 flavor tagging correlation scheme.

2041 Given that c -tagging changed significantly between Run 1 and Run 2 and that light tagging NP's
2042 are very lowly ranked, these sets of NP's are left uncorrelated. Moreover, the change in the physical
2043 meaning of the effective b -tagging NP's means a full correlation of such NP's (insomuch as they exist
2044 in each result) is one of limited utility. Hence, it was decided to leave flavor tagging NP's uncorre-
2045 lated. However, since the meaning of the leading b -tagging NP's is approximately constant across
2046 years and since Run 2 b -tagging NP's are very highly ranked in both the Run 2 only and combined
2047 fits, tests correlating these NP's were conducted, the results of which can be seen below. It should be
2048 noted that the leading B NP at 8 TeV, SysBTagB0Effic_Y2012_8TeV, has an opposite effect on $t\bar{t}$
2049 normalization than the 7 and 13 TeV NP's, and so must be flipped using a similar strategy as for JES
2050 unfolding. Initial studies of flavor tagging correlations did not flip this NP, and so results for this
2051 scheme (labeled "Bo 8TeV Not Flipped") have also been included for comparison.

2052 It is clear from these results that correlating the leading effective Eigen NP associated with b 's can
2053 have a noticeable effect on final fit results and that the 8 TeV Bo NP is the most important compo-
2054 nent of a combined Bo NP. It is also not so surprising that the 8 TeV result should drive the com-

	Comb Eff	BTag Bo	Bo 8TeV Not Flipped
Exp. Sig.	3.998	4.127	3.921
Obs. Sig.	3.571	3.859	3.418
Exp. Limit	0.51 ^{+0.2} _{-0.143}	0.5 ^{+0.196} _{-0.14}	0.517 ^{+0.202} _{-0.144}
Obs. Limit	1.37	1.41	1.35

Table 9.9: Expected and observed sensitivities for a combination featuring the weak JES scheme, combination with the weak JES scheme + leading b NP's correlated, and the b correlation with the 8 TeV NP with sign unflipped.

	Comb Eff	BTag Bo	Bo 8TeV Not Flipped
$ \Delta\hat{\mu} $	—	0.0446	0.0268
$\hat{\mu}$	0.8985	0.9431	0.8717
Total	+0.278 / -0.261	+0.275 / -0.256	+0.282 / -0.263
DataStat	+0.185 / -0.181	+0.180 / -0.177	+0.189 / -0.186
FullSyst	+0.208 / -0.188	+0.207 / -0.186	+0.209 / -0.186
BTag	+0.077 / -0.076	+0.071 / -0.068	+0.079 / -0.075
BTag b	+0.062 / -0.059	+0.055 / -0.049	+0.064 / -0.060

Table 9.10: Breakdowns of the impact of different NP sets on total error on $\hat{m}\hat{\mu}$ for a combination featuring the weak JES scheme and a combination with the weak JES scheme + leading b NP's correlated.

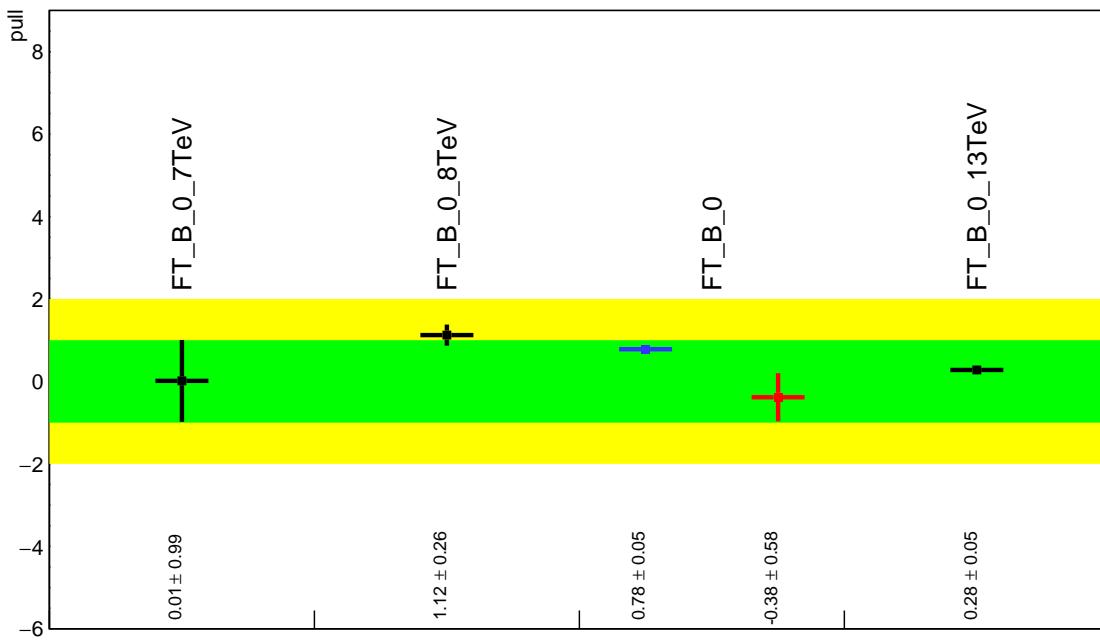


Figure 9.10: Pull Comparisons: btag-b---BTagB0 Comb Eff, BTag BO

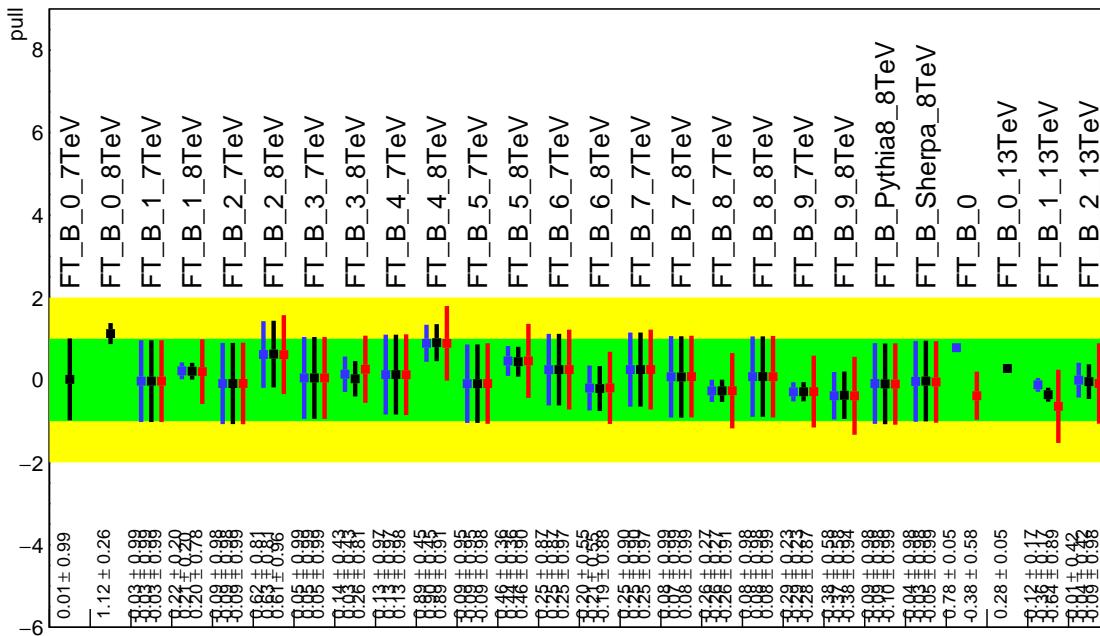


Figure 9.11: Pull Comparisons: btag-b---BTagB Comb Eff, BTag BO

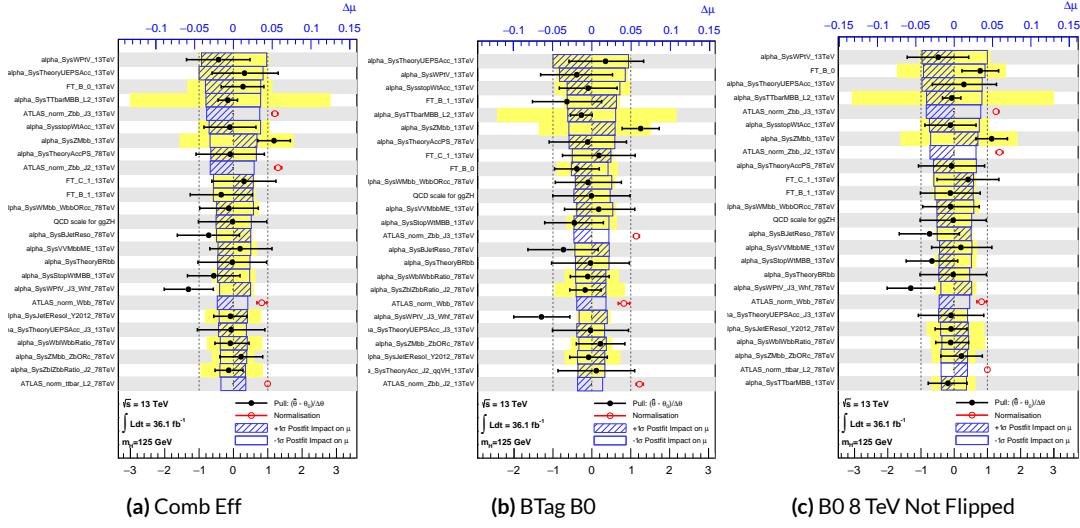


Figure 9.12: NP rankings for a combination featuring the weak JES scheme and a combination with the weak JES scheme + leading b NP's correlated.

2055 bined nuisance parameter since it is the only result to make use of both pseudocontinuous tagging-
 2056 based and b -tag regions into the final fit, implicitly yielding much more information about b 's. The
 2057 13 TeV fit has neither of these regions. What is less clear is whether there are sufficient grounds for
 2058 implementing this correlation (i.e. does the correspondence of these NP's across years warrant a full
 2059 correlation). While there are no current plans to do so, this matter warrants careful scrutiny if Run 1
 2060 is to be combined with future results.

2061 9.2.3 MODELING SYSTEMATICS

2062 Another principal systematic category is modeling uncertainties. The effect of correlating groups
 2063 of systematics was estimated using the same strategy employed by the ATLAS/CMS SM $VH(b\bar{b})$
 2064 combination for Run 1. This extrapolation can be used to estimate the impact of correlations on
 2065 the estimated signal strength, the total error on the signal strength, and the χ^2 of the result. The

2066 impact of such correlations is no more than a few percent effect, as the following tables demonstrate,
 2067 beginning with the category with the greatest shift, W+jets modeling, in Table 9.11.

	$ \Delta\mu $	σ	$ \Delta\sigma $	χ^2
$\rho=-1$	0.0024	0.2448	0.011 (4.3%)	0.95
$\rho=-0.6$	0.0015	0.2493	0.00654 (2.55%)	0.9804
$\rho=-0.3$	0.0008	0.2526	0.00325 (1.27%)	1.0045
$\rho=0$	—	0.2558	—	1.0298
$\rho=0.3$	0.0008	0.259	0.0032 (1.25%)	1.0564
$\rho=0.6$	0.0017	0.2622	0.00636 (2.49%)	1.0844
$\rho=1$	0.0029	0.2664	0.0105 (4.11%)	1.1242

Table 9.11: Run 1 + Run 2 W+jets modeling correlation projections

2068 9.2.4 FINAL CORRELATION SCHEME

2069 The final Run 1 + Run 2 correlation scheme is shown in Table 9.12. As detailed above, neither JES
 2070 nor modeling systematics had any demonstrable effect on combined fit results. Hence, only signal
 2071 NP's and the b -jet energy scale are correlated (the weak JES scheme without unfolding). While the
 2072 effect of flavor tagging correlations is less clear, the result physical arguments for correlation are less
 2073 strong; the size of effect was discovered rather late in the analysis process; and no nuisance parameter
 2074 unfolding maps exist for flavor tagging as they do for JES, so it was decided to leave these uncorre-
 2075 lated as well.

7 TeV NP	8 TeV NP	13 TeV NP
	ATLAS_BR_bb	SysTheoryBRbb
	SysTheoryQCDscale_ggZH	SysTheoryQCDscale_ggZH
	SysTheoryQCDscale_qqVH	SysTheoryQCDscale_qqVH
—	SysTheoryPDF_ggZH_8TeV	SysTheoryPDF_ggZH
—	SysTheoryPDF_qqVH_8TeV	SysTheoryPDF_qqVH
—	SysTheoryVHPt_8TeV	SysVHNLOEWK
SysJetFlavB_7TeV	SysJetFlavB_8TeV	SysJET_21NP_JET_BJES_Response

Table 9.12: A summary of correlated nuisance parameters among the 7, 8, and 13 TeV datasets.

2076 **9.3 COMBINED FIT RESULTS**

2077 **9.3.1 COMBINED FIT MODEL VALIDATION**

2078 Before moving onto the final results, we present the rest of the validations for the Run 1 + Run 2
 2079 combined fits, beginning with impacts of ranked individual nuisance parameters in Figure 9.13 and
 2080 for all nuisance parameter categories in Table 9.13. Both of these sets of results point to the most im-
 2081 portant nuisance parameters being signal systematics, b -tagging, and V +jets modeling systematics,
 2082 with Run 2 NP's generally being higher ranked. That some NP's are strongly pulled is not unusual
 2083 as the fit model has so many NP's; V +jets modeling in particular has been historically difficult.

2084 In addition to looking at the behaviors of nuisance parameters to gauge fit model performance
 2085 and stability, fits are conducted using multiple parameters of interest. Typical divisions are Run 1
 2086 vs. Run 2, lepton channels, and WH vs ZH . As mentioned in Chapter 7, the profile likelihood test
 2087 statistic given in Equation 7.2 is, in the limit of large sample statistics, a χ^2 distribution with degrees
 2088 of freedom equal to the number of parameters of interest plus number of nuisance parameters.
 2089 Thus, changing the number of interest parameters and leaving the rest of the fit model unchanged

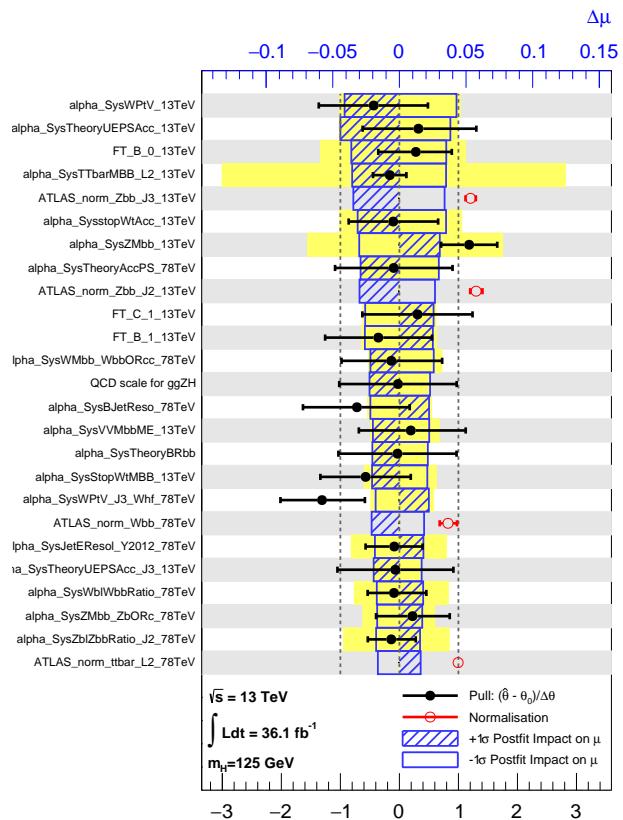


Figure 9.13: Ranked nuisance parameters for the Run1+Run2 combination.

Total	+0.278 / -0.261
DataStat	+0.185 / -0.181
FullSyst	+0.208 / -0.188
Floating normalizations	+0.055 / -0.056
All normalizations	+0.068 / -0.069
All but normalizations	+0.192 / -0.172
Jets, MET	+0.046 / -0.040
Jets	+0.041 / -0.036
MET	+0.023 / -0.018
BTag	+0.077 / -0.076
BTag b	+0.062 / -0.059
BTag c	+0.033 / -0.032
BTag light	+0.028 / -0.028
Leptons	+0.008 / -0.008
Luminosity	+0.026 / -0.014
Diboson	+0.030 / -0.027
Model Zjets	+0.049 / -0.050
Zjets flt. norm.	+0.032 / -0.040
Model Wjets	+0.082 / -0.083
Wjets flt. norm.	+0.031 / -0.027
Model ttbar	+0.047 / -0.046
ttbar flt. norm.	+0.025 / -0.026
Model Single Top	+0.047 / -0.045
Model Multi Jet	+0.027 / -0.038
Signal Systematics	+0.098 / -0.052
MC stat	+0.080 / -0.084

Table 9.13: Summary of the impact of different nuisance parameter categories on the total error on $\hat{\mu}$ for the combined Run1+Run2 fit.

2090 means that the difference between the nominal fit and a fit with more parameters of interest ought
 2091 to also be distributed as a χ^2 distribution with degrees of freedom equivalent to the number of extra
 2092 parameters of interest. This difference can then be interpreted as a compatibility between the two
 2093 results using the standard tables for this distribution, giving another gauge of fit performance. These
 2094 are shown in Table 9.14.

Fit	Compatibility
Leptons (3 POI)	1.49%
WH/ZH (2 POI)	34.2%
Run 1/Run 2 (2 POI)	20.8%
Run 1/Run 2 \times Leptons (6 POI)	7.10%
Run 1/Run 2 \times WH/ZH (4 POI)	34.6%

Table 9.14: Summary of multiple POI compatibilities. The well-known Run 1 7 TeV 0-lepton deficit is responsible for the low compatibility with the 6 and 3 POI fits.

2095 The low compatibilities associated with treating the lepton channels as separate parameters of
 2096 interest are a symptom of the low signal strengths associated with the Run 1 0-lepton channel, in par-
 2097 ticular the 7 TeV result. Given the relatively small amount of data associated with the 7 TeV result,
 2098 this should not be a cause for alarm. Signal strength summary plots for the fits treating Run 1 and
 2099 Run 2 separately are shown in Figures 9.14-9.16, where the effect of the Run 1 parameters can be seen
 2100 graphically.

2101 9.3.2 FINAL RESULTS

2102 The combined results yields an observed (expected) significance of 3.57 (4.00) and an observed (ex-
 2103 pected) limit of 1.37 ($0.510^{+0.200}_{-0.143}$), with a signal strength of $\hat{\mu} = 0.898^{+0.278}_{-0.261}$.

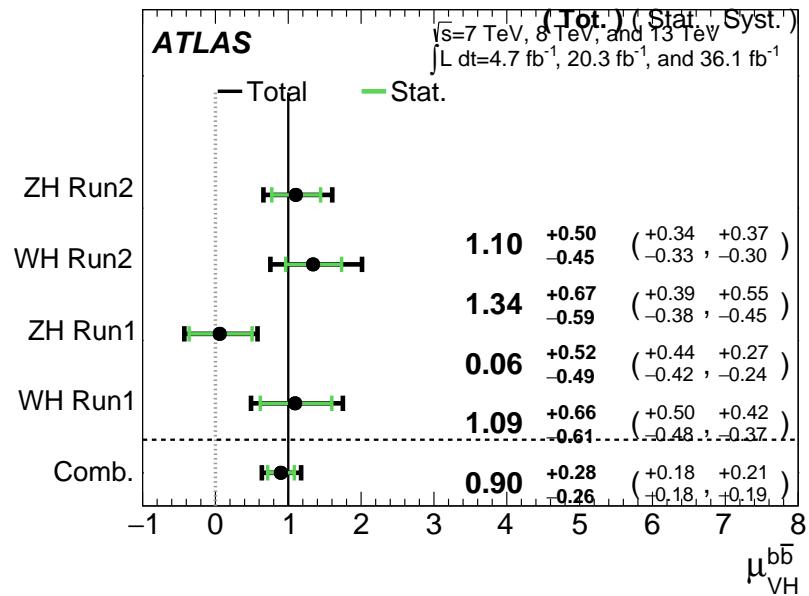


Figure 9.14: $\hat{\mu}$ summary plot for a four parameter of interest fit.

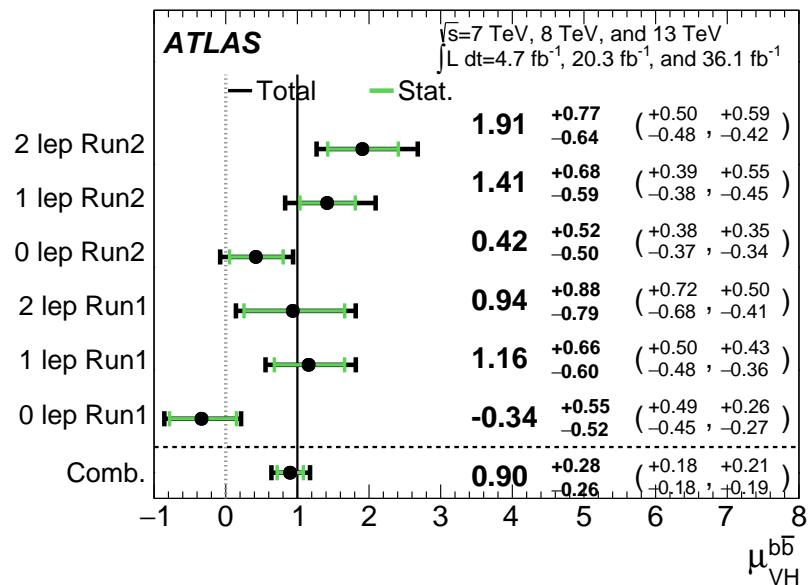


Figure 9.15: $\hat{\mu}$ summary plot for a six parameter of interest fit.

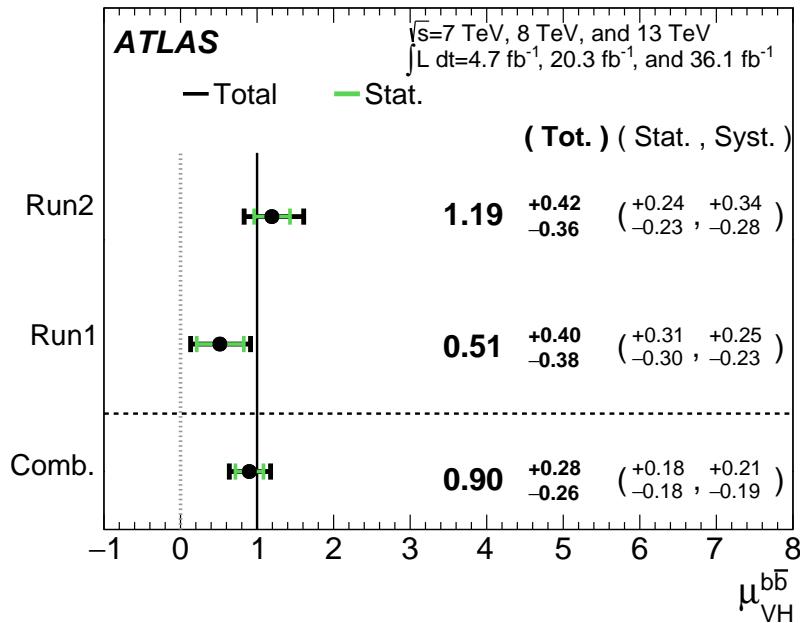


Figure 9.16: $\hat{\mu}$ summary plot for a two parameter of interest (Run 1 and Run 2) values.

2104 The two and three parameter of interest fit signal strength summary plots, as well as a summary
 2105 of the historical values of the 7, 8, and 13 TeV results may be found in Figures 9.17-9.19. The main
 2106 results for Run 1, Run 2, and the combination may be found in Table 9.15. These results were collec-
 2107 tively noted as the first ever experimental evidence for SM $VH(b\bar{b})$ in [3].

Dataset	$\hat{\mu}$	Total Error in $\hat{\mu}$	Obs. (Exp.) Significance
Run 1	0.51	+0.40 / -0.37	1.4 (2.6)
Run 2	1.20	+0.42 / -0.36	3.54 (3.03)
Combined	0.90	+0.28 / -0.26	3.57 (4.00)

Table 9.15: A summary of main results for the Run 1, Run 2, and combined fits.

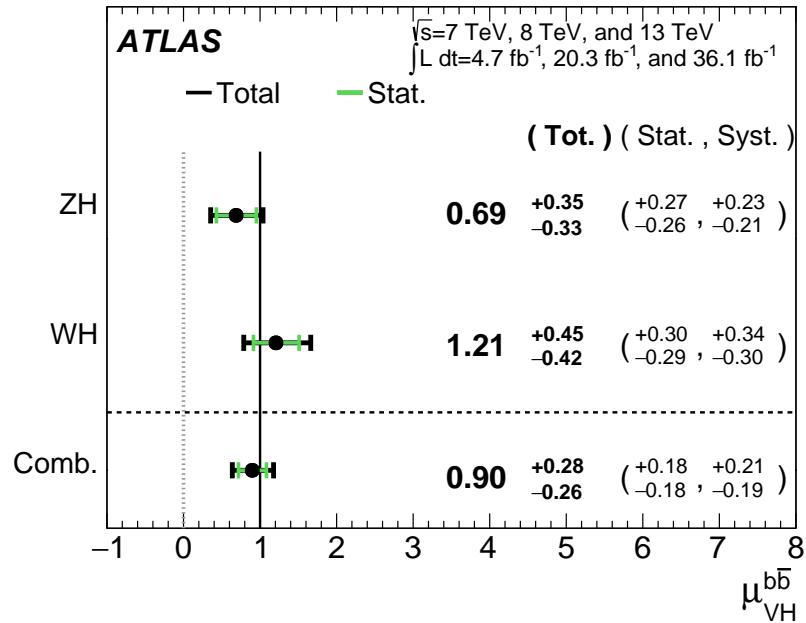


Figure 9.17: $\hat{\mu}$ summary plot for a two parameter of interest fit.

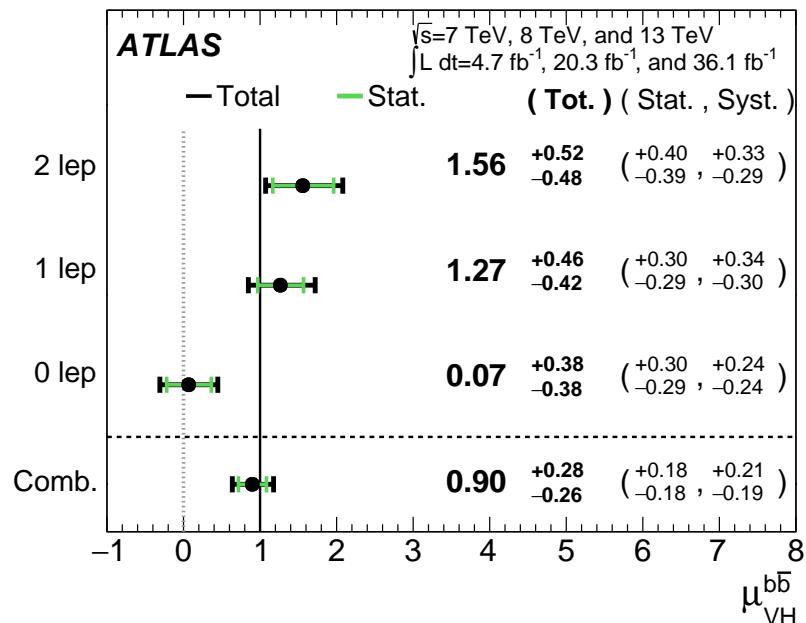


Figure 9.18: $\hat{\mu}$ summary plot for a three parameter of interest fit.

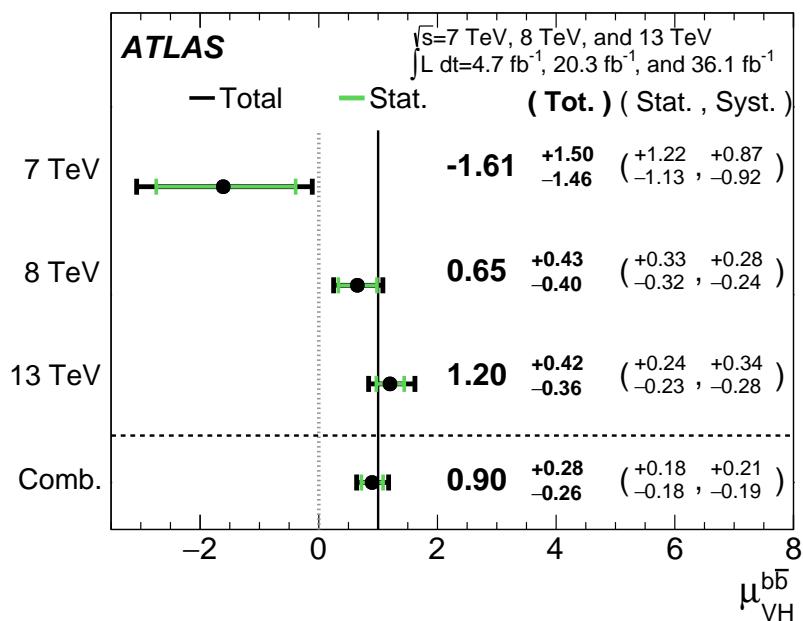


Figure 9.19: $\hat{\mu}$ summary plot for different \sqrt{s} values.

Vanitas vanitatum, omnis vanitas

Ecclesiastes 1:2

10

2108

2109

Closing Thoughts

2110 SINCE BOTH THE LHC and ATLAS are performing very well, it is only a matter of time before the
2111 evidence for SM $VH(b\bar{b})$ passes the 5 Gaussian standard deviation threshold necessary for discovery.
2112 Official discovery may come less than a year after reports of first evidence and may not even require
2113 a combination with the Run 1 result, depending on the latter two years of ATLAS Run 2 data (2017

220

2114 and 2018).

2115 It is entirely natural to ask, then, how essential the techniques and results described in this thesis
2116 will prove to be moving forward. Neither the LI/RF multivariate techniques nor combination with
2117 Run 1 datasets and their accompanying low signal strength values are necessary for discovery, and
2118 the latter may not even be essential to timely^{*} discovery of SM $VH(b\bar{b})$. Nevertheless, both sets of
2119 results hold great potential as key parts of a concerted ensemble of efforts towards precision Higgs
2120 physics.

2121 With the perhaps final major center of mass energy increase at the energy frontier ever complete,
2122 analyses must rely on increased integrated luminosity. Hence, it is becoming increasingly likely that
2123 any new fundamental physics at colliders will require the use of results of systematics limited analy-
2124 ses. This is the regime where the techniques described in this thesis will be most useful.

2125 As the LHC and its experiments undergo successive stages of upgrades and operate in evermore
2126 extreme environments, the statistical fit models used to describe LHC data will continue to evolve in
2127 complexity and diverge from their predecessors. The techniques described in Chapter 9 will become
2128 increasingly more vital to producing the best physics results possible. The improvement in precision
2129 from $\hat{\mu}_{VH} = 1.20^{+0.24}_{-0.23}(\text{stat.})^{+0.34}_{-0.28}(\text{syst.})$ to $\hat{\mu}_{VH} = 0.90^{+0.18}_{-0.18}(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$ is just the begin-
2130 ning.

2131 The best methods for reduction of systematic uncertainties will naturally depend in part on the
2132 state of the art for both fundamental physics process and detector modeling, but techniques that
2133 can reduce systematic uncertainties independent of fit model, dataset, and physics process provide

*i.e. before or coincident with CMS

2134 a promising avenue forward. The improvements in systematic uncertainties using the Lorentz In-
 2135 variant and RestFrames variable techniques in the $ZH \rightarrow \ell\ell b\bar{b}$ analysis, summarized in Table 10.1,
 2136 show that a smarter and more orthogonal decomposition of information in a collision event pro-
 2137 vides benefits independent of any clever treatment of \vec{E}_T^{miss} (which both schemes also provide). Both
 2138 techniques are readily extendable to other analysis channels, with the RestFrames concept demon-
 2139 strating stronger performance and greater flexibility to nearly completely generic final states.

	Standard	LI	RF
$\hat{\mu}$	$1.75^{+0.50, 0.64}_{-0.48, 0.45})$	$1.65^{+0.51, 0.59}_{-0.49, 0.41}$	$1.50^{+0.50, 0.53}_{-0.48, 0.36}$
Asi. $\Delta err(\mu)$	—	< 1%, +4.6%	-6.5%, -2.2%
Obs. $\Delta err(\hat{\mu})$	—	-7.5%, -3.7%	-16%, -8.8%
Stat only sig.	4.78	4.39 (-7.9%)	4.44 (-6.9%)
Exp. (Asi.) sig.	2.06	1.92 (-6.7%)	2.13 (+3.5%)
Exp. (data) sig.	1.76	1.73 (-1.7%)	1.80 (+3.4%)
Obs. (data) sig.	2.87	2.79 (-2.8%)	2.62 (-8.6%)

Table 10.1: Summary of performance figures for the standard, LI, and RF variable sets. Uncertainties on $\hat{\mu}$ are quoted stat., syst. In the case of the latter two, % differences are given where relevant. Differences in errors on μ are on full systematics and total error, respectively.

2140 Critical work remains to be done refining and extending the treatment of both the LI and RF
 2141 techniques in $VH(b\bar{b})$ analyses and their fit models, and completely independent techniques, like

2142 the use of multiple event interpretations addressed in Appendix B promise further improvements
 2143 still.

2144 No one can say for certain what the future of the energy frontier of experimental particle physics
 2145 may hold, but more nuanced treatments of the information in collision events born of meaningful
 2146 physical insight are sure to light the way.

If it's stupid but it works, it isn't stupid.

Conventional Wisdom

A

2147

2148 Micromegas Trigger Processor Simulation

2149 IN ORDER TO PRESERVE key physics functionality by maintaining the ability to trigger on low p_T
2150 muons, the Phase I Upgrade to ATLAS includes a New Small Wheel (NSW) that will supply muon
2151 track segments to the Level 1 trigger. These NSW trigger segments will combine segments from the
2152 sTGC and Micromegas (MM) trigger processors (TP). This note will focus in particular on the al-

223

2153 gorithm for the MMTP, described in detail with initial studies in [61]. The goal of this note is to de-
2154 scribe the MMTP algorithm performance under a variety of algorithm settings with both nominal
2155 and misaligned chamber positions, as well as addressing a number of performance issues.

2156 This note is organized as follows: the algorithm and its outputs are briefly described in Section
2157 A.1; Monte Carlo samples used are in Section B.1; nominal algorithm performance and certain quan-
2158 tities of interest are described in Section A.3; algorithm performance under misalignment, misalign-
2159 ment corrections, and corrected performance are shown in Section A.9; and conclusions are pre-
2160 sented in Section A.24.

2161 A.1 ALGORITHM OVERVIEW

2162 The MMTP algorithm is shown schematically in Figure A.1, taken from [61], where a more detailed
2163 description may be found. The algorithm begins by reading in hits, which are converted to slopes.
2164 These slopes are calculated under the assumption that the hit originates from the IP; slopes calcu-
2165 lated under this assumption are denoted by a superscript g for global in order to distinguish them
2166 from local slopes calculated using only hits in the wedge. In the algorithm simulation, events are
2167 screened at truth level to make sure they pass certain requirements. The track's truth-level coor-
2168 dinates must place it with the wedge since some generated tracks do not reach the wedge. These
2169 hits are stored in a buffer two bunch crossings (BCs) in time deep that separates the wedge into so-
2170 called "slope-roads." If any given slope-road has sufficient hits to pass what is known as a coinci-
2171 dence threshold, a fit proceeds. A coincidence threshold is a requirement for an event expressed as
2172 $aX+bUV$, which means that an slope-road must have at least a hits in horizontal (X) planes and at

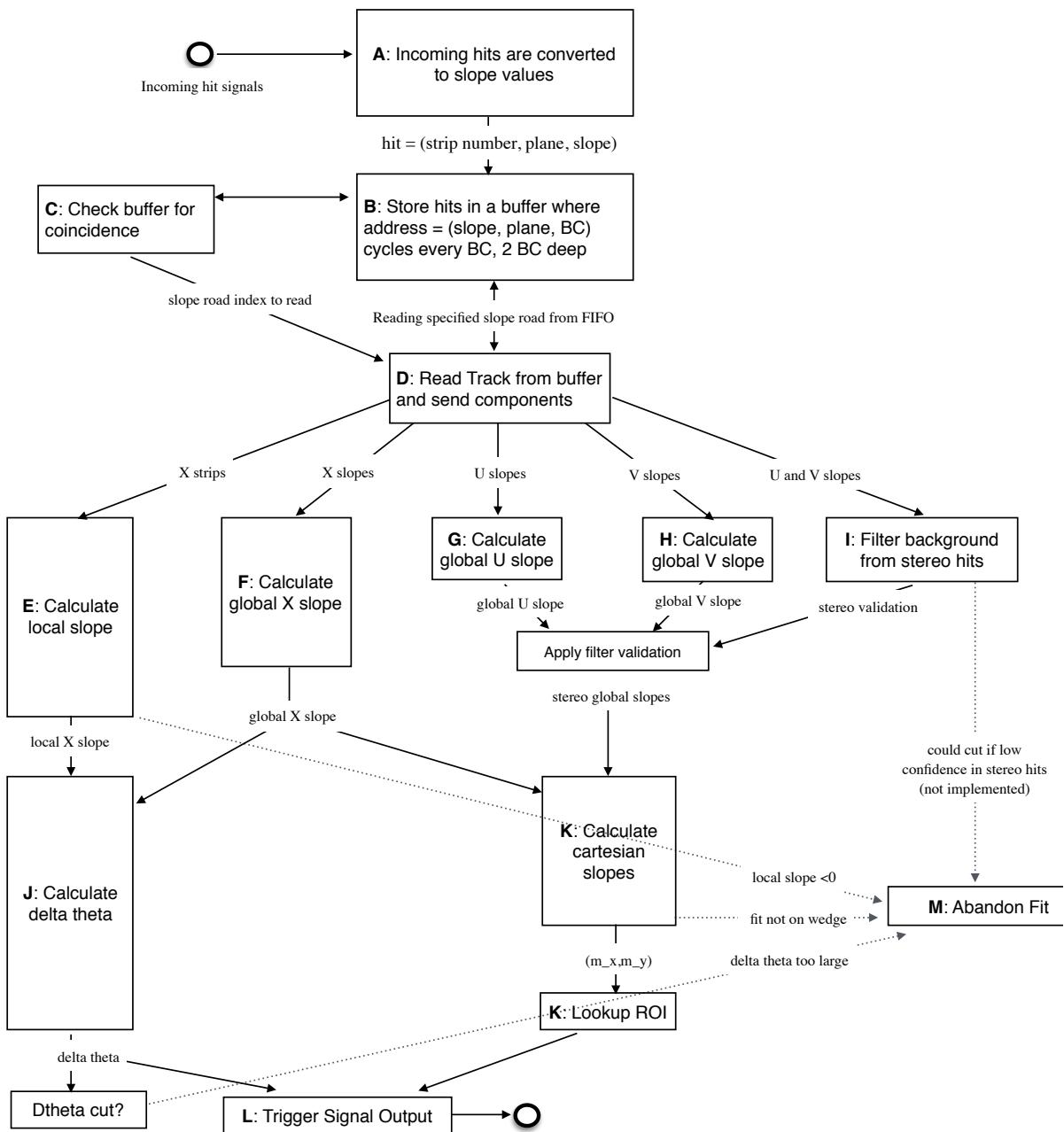


Figure A.1: A flow chart describing the algorithm steps, taken from [61].

2173 least b hits in stereo (U or V (corresponding to positive and negative stereo rotations)) planes. For
 2174 coincidence thresholds with a $2X$ hit requirement there is the extra requirement that, in the case of
 2175 only $2X$ hits, one be on each quadruplet in order to ensure an adequate lever arm for the $\Delta\theta$ calcu-
 2176 lation. Note that less stringent (lower hit) coincidence thresholds are inclusive; i.e. a slope-road pass-
 2177 ing a $4X+4UV$ cut automatically passes $2X+1UV$. The coincidence threshold, size of the slope-roads
 2178 (denoted b), and the number of slope-roads into which each horizontal and stereo hits get written
 2179 centered upon their nominal value are configurable parameters of the algorithm.

2180 An individual hit's slope is calculated as shown in Equation A.1, where y_{base} is the local y coordi-
 2181 nate (orthogonal to the beamline and direction of the horizontal strips) of a station's base, w_{str} is the
 2182 strip pitch, n_{str} is the hit's strip number, and z_{plane} is the location of the hit's plane along the beam-
 2183 line.

$$M_{hit} = \frac{y}{z} = \frac{y_{base}}{z_{plane}} + \frac{w_{str}}{z_{plane}} \times n_{str} \quad (\text{A.1})$$

2184 In the fit, individual hit slopes in a slope-road are used to calculate global slopes associated with each
 2185 plane type, which are averages (e.g. M_X^{ℓ} for the average slope of horizontal planes). These in turn are
 2186 used to calculate the three composite slopes: slopes associated with the horizontal (m_x) and vertical
 2187 coordinates (m_y) and the local slope of hits in the horizontal planes (M_X^l), all of which are shown in
 2188 Equation A.4. Note that the expression for M_X^l differs but is equivalent to the expression given in
 2189 [61]. This is due to a procedural change in the algorithm. The local X slope is expressed in [61] as:

$$M_X^{local} = A_k \sum_i y_i z_i - B_k \sum_i y_i, \quad B_k = \frac{1}{n} \sum_i z_i A_k = \bar{z} A_k \quad (\text{A.2})$$

2190 Procedurally, this entails doing the sums over y_i and $y_i z_i$, multiplying the sums by A_k , B_k , and then
 2191 subtracting both of these numbers, $\mathcal{O}(10^3)$, to get local slopes, $\mathcal{O}(10^{-1})$, while requiring preci-
 2192 sion on these numbers on the order of $\mathcal{O}(10^{-3})$. This requires precision in the sums $\mathcal{O}(10^{-7})$,
 2193 and with 32 bit fixed point numbers, there are deviations with respect to the floating point calcula-
 2194 tions at the level of $\mathcal{O}(10^{-5})$, which is enough to introduce a significant bias in the $\Delta\theta$ calculation.

2195 In order to prevent these errors, we do the subtraction first

$$M_X^{local} = A_k \sum_i y_i z_i - B_k \sum_i y_i = A_k \sum_i (y_i z_i - y_i \bar{z}) = B_k \sum_i y_i \left(\frac{z_i}{\bar{z}} - 1 \right) \quad (\text{A.3})$$

2196 Thus, we change the order of operations and store $1/\bar{z}$ instead of A_k in addition to B_k . We also
 2197 change the units of y_i and z_i in the calculation by dividing the millimeter lengths by 8192.* With
 2198 these changes, a 32 bit fixed point based algorithm has essentially identical performance to that of an
 2199 algorithm based on the usual C++ 32 floating point numbers. Future work includes converting the
 2200 32 bit fixed point arithmetic to 16 bit where possible in the algorithm. While introducing 16 bit num-
 bers uniformly might seem preferable, since simple 16-bit operations in the firmware can be done in
 2201 a single clock tick, and a larger number of bits increases the algorithm latency, some numbers in the
 2202 algorithm will require a larger number of bits, in particular in the local slope calculation, which is
 2203 the single calculation in the algorithm requiring the largest numeric range.

2204 In Equation A.4, θ_{st} is the stereo angle of 1.5 degrees; the sums are over relevant planes; \bar{z} is the
 2205 average position in z of the horizontal planes; and y_i and z_i in the local slope expression refer to the y

*Chosen since it is a perfect power of 2 and of order the length scale of z in millimeters

2207 and z coordinates of hits in X planes.

$$m_x = \frac{1}{2} \cot \theta_{st} (\mathcal{M}_U^g - \mathcal{M}_V^g), \quad m_y = \mathcal{M}_X^g, \quad \mathcal{M}_X^l = \frac{\bar{z}}{\sum_i z_i^2 - 1/n (\sum_i z_i)^2} \sum_i y_i \left(\frac{z_i}{\bar{z}} - 1 \right) \quad (\text{A.4})$$

2208 From these composite slopes, the familiar expressions for the fit quantities θ (the zenith), ϕ (the az-
2209 imuth[†]), and $\Delta\theta$ (the difference in θ between the direction of the segment extrapolated back to the
2210 interaction point and its direction when entering the detector region; the following is an approxima-
2211 tion) may be calculated, as noted in [61]:

$$\theta = \arctan \left(\sqrt{m_x^2 + m_y^2} \right), \quad \phi = \arctan \left(\frac{m_x}{m_y} \right), \quad \Delta\theta = \frac{\mathcal{M}_X^l - \mathcal{M}_X^g}{1 + \mathcal{M}_X^l \mathcal{M}_X^g} \quad (\text{A.5})$$

2212 Looking at Equations A.4 and A.5, the dependence of fit quantities on input hit information be-
2213 comes clear. $\Delta\theta$ relies exclusively on information from the horizontal (X) planes. Both θ and ϕ rely
2214 on both horizontal and stereo slope information. However, the sum in quadrature of m_x and m_y in
2215 the arctangent for θ means that θ is less sensitive to errors in stereo hit information than ϕ . Given
2216 that θ_{st} is small, $\cot \theta_{st}$ is large (~ 38), so m_x multiplies small differences in \mathcal{M}_U and \mathcal{M}_V , where m_y
2217 is simply an average over slopes. This means that while errors in horizontal hit information will af-
2218 fect all three fit quantities, comparable errors in stereo hits will have a proportionately larger effect
2219 on θ and particularly on ϕ . The $\Delta\theta$ cut after step J in Figure A.1 has been implemented, requiring
2220 all fits to have $|\Delta\theta| < 16$ mrad. This requirement ensures good quality fits but also slightly reduces

[†]Defined with respect to the center (y) axis and *not* the axis of the strips (x) as is sometimes typical, so a hit along the center of the wedge has $\phi = 0$

2221 algorithm efficiency.

2222 A.2 MONTE CARLO SAMPLES

2223 The Monte Carlo (MC) samples used for these studies were generated in Athena release 20.1.0.2 us-
2224 ing simulation layout ATLAS-R2-2015-01-01-00 with muon GeoModel override version MuonSpectrometer-
2225 R.07.00-NSW and modifications to have two modules per multiplet and xxuvuvxx geometry with a
2226 stereo angle of 1.5 degrees. Muons of a single p_T were generated around the nominal IP with a smear-
2227 ing of 50 mm along the beam line and 0.015 mm orthogonal to it; these muons were pointed toward
2228 a single, large sector of the NSW. Each event consists of one muon fired towards the single NSW
2229 wedge separated by effectively infinite time from other events.

2230 A.3 NOMINAL PERFORMANCE

2231 In order to evaluate algorithm performance, a number of quantities are evaluated, including the fit
2232 quantities θ , ϕ , and $\Delta\theta$ as well as algorithm efficiency. Unless otherwise stated, that algorithm is
2233 run with a 4X+4UV coincidence threshold, slope-road size of 0.0009, an X tolerance of two slope-
2234 roads (i.e. hits in horizontal planes are written into the two slope-roads closest to the hits' value),
2235 a UV tolerance of four slope-roads[†], and a charge threshold requirement on hits of 1 (measured in
2236 units of electron charge) for a sample of 30 000 events with a muon p_T of 100 GeV. Samples were
2237 also generated for p_T values of 10 GeV, 20 GeV, 30 GeV, 50 GeV, and 200 GeV, which were used in

2238 [†]The larger tolerance on stereo hits takes into account the particulars of the m_x calculation mentioned in
Section A.1.

2238 some of the following studies.

2239 **A.4 FIT QUANTITIES**

2240 In order to evaluate the performance of the algorithm’s fit quantities θ , ϕ , and $\Delta\theta$, fit values are com-
2241 pared to truth-level MC values. The residual of the three fit quantities, $\theta_{fit} - \theta_{tru}$, $\phi_{fit} - \phi_{tru}$, and
2242 $\Delta\theta_{fit} - \Delta\theta_{tru}$, are recorded for every fitted track. The distributions of these quantities, in particular
2243 their biases and standard deviations, are then used to evaluate performance. In most cases, follow-
2244 ing [61], the mean and standard deviation of a 3σ Gaussian fit are quoted, as they capture the main
2245 features of the algorithm and generally behave like the raw mean and rms. Nevertheless, discussion
2246 of the raw quantities will be included when their behavior deviates markedly from that of the 3σ fit
2247 quantities.

2248 The truth-level quantities used in residual distribution are taken from information in the MC.

2249 These come directly from the MC for θ , ϕ , and $\Delta\theta$. These quantities, along with the geometry of
2250 the (large) wedge, are then in turn used to calculate truth-level values for any intermediate quantities
2251 used in the algorithm. $m_{x,tru}$, for instance, is given by $\tan \theta_{tru} \sin \phi_{tru}$.

2252 Residual distributions for fit quantities under the previously described default settings of the al-
2253 gorithm are shown in Figure A.2. Both the $\theta_{fit} - \theta_{tru}$ and $\Delta\theta_{fit} - \Delta\theta_{tru}$ distributions feature a
2254 mostly Gaussian shape with more pronounced tails. The mean bias for these distributions is negligi-
2255 ble at under one tenth of a milliradian, and the fitted (raw) rms values are 0.349 (0.614) mrad for θ
2256 and 1.03 (2.55) mrad for $\Delta\theta$. The case of the $\phi_{fit} - \phi_{tru}$ distribution is less straightforward, with both
2257 the shape and bias arising from the xxuvuvxx geometry and relatively large extent of one of the two

2258 η -stations, as explained in Appendix B of [62]. The fitted (raw) rms for the ϕ distribution is 8.67
 2259 (16.6) mrad.

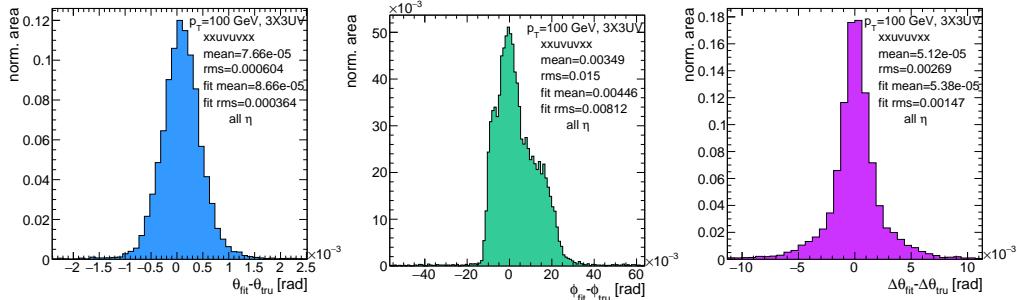


Figure A.2: Nominal residual plots; $\theta, \phi, \Delta\theta$ for $p_T = 100 \text{ GeV}$ muons

2260 Both increasing muon p_T and increasing muon η for a fixed p_T imply increasing muon energy. As
 2261 muons become more energetic, two effects compete in affecting the quality of fit. On the one hand,
 2262 higher energy muons are deflected less by the ATLAS magnetic field, which should tend to improve
 2263 the quality of the fit, since the fitted θ (upon which $\Delta\theta$ also relies) and ϕ values are calculated under
 2264 the infinite momentum muon (straight track) assumption. However, as muon energy increases, the
 2265 likelihood that the muon will create additional secondaries increases, which creates extra hits that
 2266 degrade the quality of the fit. While the geometry of the multiplet is such that there is very good res-
 2267 olution in the direction orthogonal to the horizontal strip direction, the shallow stereo angle of 1.5
 2268 degrees means that early hits caused by secondaries can have an outsize impact on m_x . $\Delta\theta$, which
 2269 does not rely upon stereo information should feel the effect of secondaries the least and benefit from
 2270 straighter tracks the most and hence benefit from higher muon energies; ϕ , relying upon stereo in-
 2271 formation the most, would be most susceptible to secondaries and benefit the least from straighter

tracks and hence least likely to benefit from higher muon energy; θ relies upon both horizontal and vertical slope information, though small errors are less likely to seriously affect the calculation, so the two effects are most likely to be in conflict for this fit quantity.

The interplay of these effects on the residual standard deviations can be seen in their dependences on η (Figure A.3; note that the final point in each of these plots is the rms of the distribution overall η) and p_T (Figure A.4). For $p_T = 100$ GeV muons, $\Delta\theta$ performance increases with η (energy), and ϕ performance decreases, as expected;[§] for θ , the two effects appear to compete, with performance first increasing with η until the effects of secondaries begins to dominate. Integrated over all η , the effects are less clearly delineated. Both $\Delta\theta$ and θ performance increases with increasing p_T , suggesting straighter tracks with increasing energy are the dominant effect for these quantities, while ϕ performance appears to improve and then deteriorate (the slight improvement at high p_T is due to the addition of the $\Delta\theta$ cut into the algorithm, which filters out very poor quality fits).

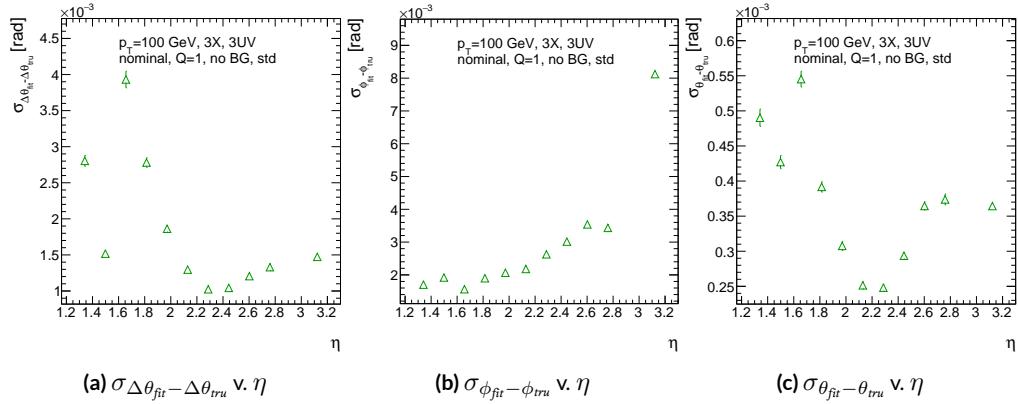


Figure A.3: The rms distributions of $\Delta\theta$, ϕ , and θ as a function of η for $p_T = 100$ GeV; the final point in each plot is the rms obtained from a fit to the full distribution including all η bins.

[§]The much worse overall performance for ϕ is due to the η dependent bias and other effects

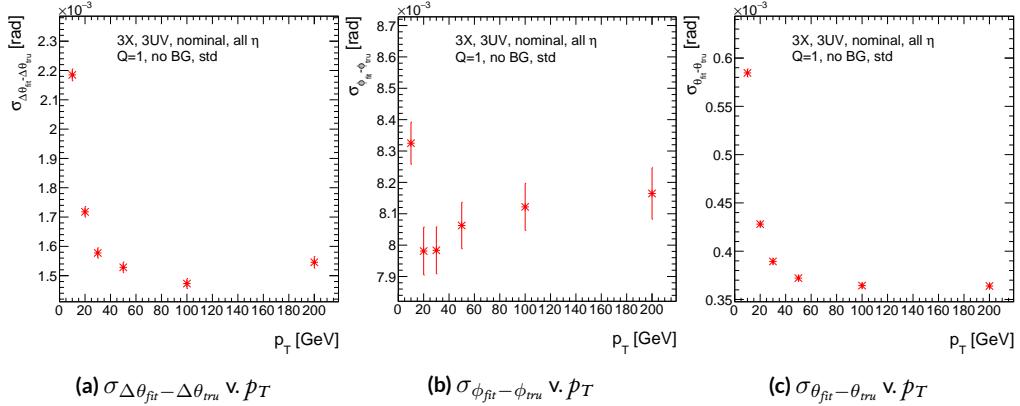


Figure A.4: The rms distributions of $\Delta\theta$, ϕ , and θ as a function of p_T .

2284 The rms of the three benchmark quantities as a function of algorithm set (i.e. slope-road) coinci-
 2285 dence threshold are shown in Figure A.5 using Gaussian fits and in Figure A.6 for the raw quantities.
 2286 The fitted σ 's for θ and ϕ are fairly stable across coincidence threshold. $\Delta\theta$, on the other hand, per-
 2287 forms better particularly for the most stringent coincidence threshold; this is a result of the fact that
 2288 additional information for more hits greatly improves the quality of the local slope fit calculation.
 2289 The raw rms is a different story. Naïvely, one would expect the performance to get better with more
 2290 stringent coincidence threshold, but this is not the case in Figure A.6. As the coincidence thresh-
 2291 old gets more stringent, fewer and fewer tracks are allowed to be fit. When moving from 2X hits to
 2292 3X hits, the tracks that get vetoed populate the tails of the distribution outside the 3σ fit range but
 2293 are not in the very extremes of the distribution. While tracks with 2X hits are of lower quality than
 2294 those with 3 and 4 X hits, tracks with the very worst fit values pass even the most stringent coinci-
 2295 dence threshold requirements (e.g. as a result of many hits arising from a shower of secondaries).
 2296 This is best illustrated when comparing the 2X+1UV $\Delta\theta$ residual distribution with the 4X+4UV

distribution in Figure A.7. As both the overlayed normalized curves and ratio distribution show,
 while the most central regions are fairly similar, the $\omega X + 1$ UV distribution is much more prominent
 in the tails but not the extreme tails, which means that, though the overall $\omega X + 1$ UV raw rms goes
 down, the overall quality of algorithm fits is worse.

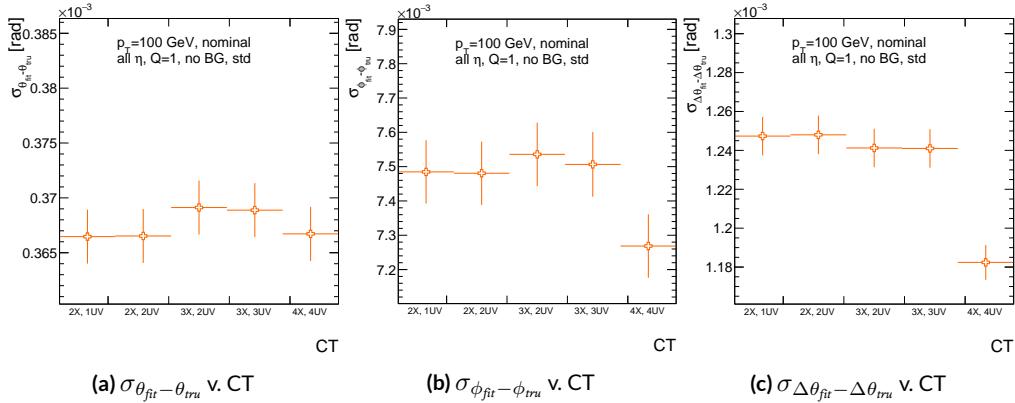


Figure A.5: The fitted rms of residual distributions for θ , ϕ , and $\Delta\theta$ as a function of coincidence threshold for $p_T = 100$ GeV.

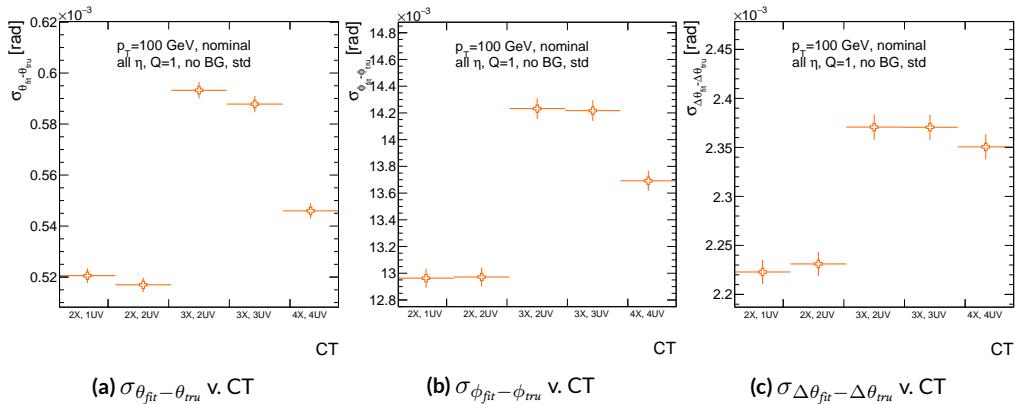


Figure A.6: The raw rms of residual distributions for θ , ϕ , and $\Delta\theta$ as a function of coincidence threshold for $p_T = 100$ GeV.

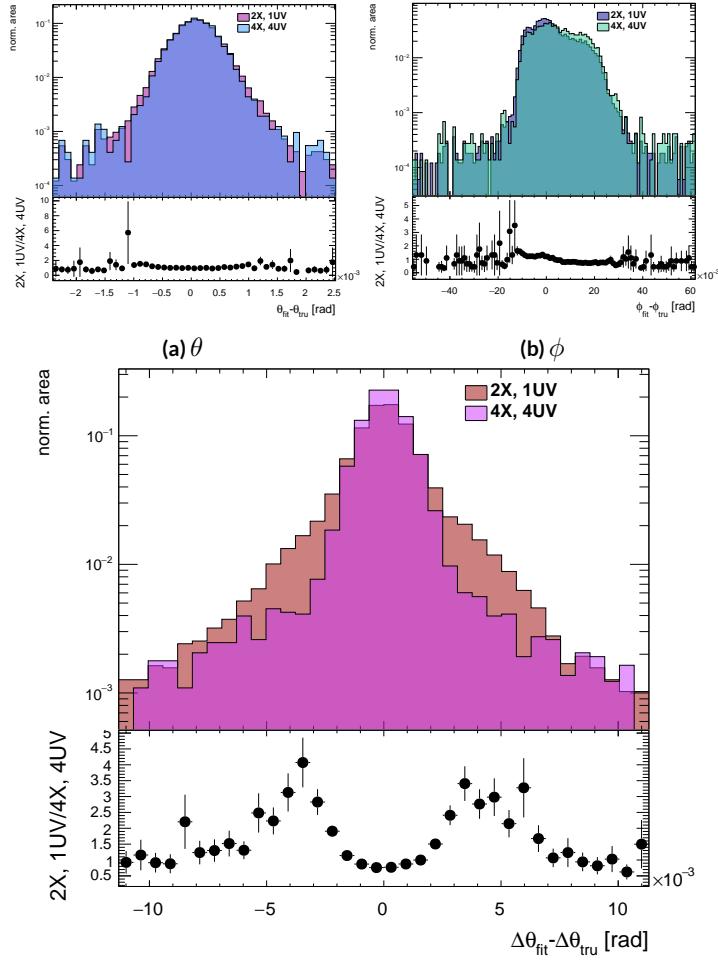


Figure A.7: Nominal $\Delta\theta$ residual distribution for $p_T = 100$ GeV muons with coincidence thresholds 2X+1UV and 4X+4UV normalized to the same area and plotted together (top) as well as the ratio of the 2X+1UV distribution and the 4X+4UV per bin.

2301 A.5 EFFICIENCIES

2302 Two general efficiencies have been formulated to study the performance of the MMTP algorithm.

2303 The first, denoted ε_{alg} , is the fraction of tracks that pass some (slope-road) coincidence threshold

2304 configuration that are successfully fit. An event that passes a slope-road coincidence but does not fit

2305 fails because some of the hits included are of sufficiently poor quality to throw off the fit. This effi-

2306 ciency answers the question of how often the algorithm performs fits when technically possible, giv-

2307 ing a measure of overall algorithm performance for a given configuration. For example, $\varepsilon = 95\%$ for

2308 $3X+2UV$ means that 95% of tracks that produce at least $3X$ hits and $2UV$ hits in at least one slope-

2309 road will be successfully fitted 95% of the time. The performance of this efficiency as a function of

2310 coincidence threshold, η (with the final point once again being the efficiency integrated over all η),

2311 and p_T is shown in Figure A.8. ε_{alg} is fairly constant in η and decreases with increased p_T , which can

2312 be attributed to the increased likelihood of secondaries introducing lower quality hits that cause the

2313 fit to fail.

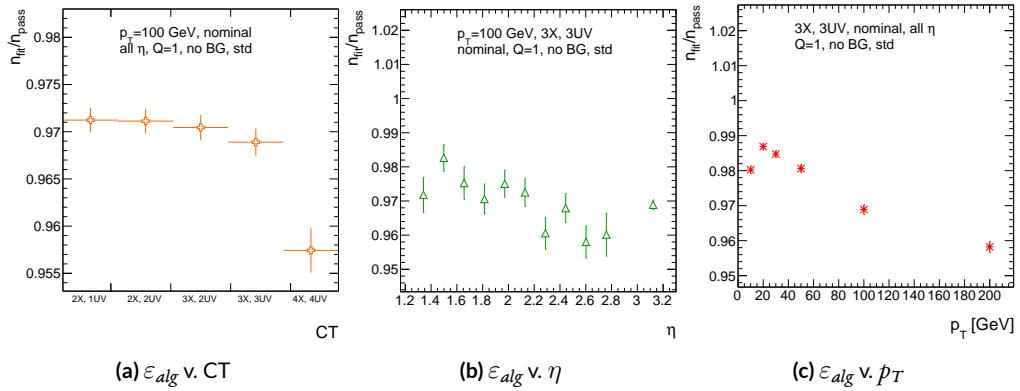


Figure A.8: ε_{alg} and as a function of coincidence threshold, η (final point is ε_{alg} integrated over all η), and p_T .

2314 The second efficiency type, denoted ε_{fit} , is the fraction of tracks that enter the wedge whose fits
 2315 (if any) satisfy a given coincidence threshold. This efficiency can be used to help establish an optimal
 2316 coincidence threshold setting in the algorithm, balancing the improved overall fit quality of higher
 2317 thresholds with the greater number of fits for lower thresholds. Hence, an ε_{fit} of 95% at 3X+2UV
 2318 means that 95% of tracks entering the wedge are fit and that these fits include at least 3X and 2UV
 2319 hits. ε_{fit} as a function of coincidence threshold is shown in Figure A.9 (a), which shows that the ma-
 2320 jority of fits having at most 3X+3UV hits. That there is a marked drop to 4X+4UV is not surpris-
 2321 ing, as there is a substantial population outside the 4X+4UV bin in Figure A.10. The behavior of
 2322 ε_{fit} with η in Figure A.9 (b) (with the final point once again being the efficiency integrated over all
 2323 η) is much more varied, with geometric effects of detector acceptance coming into play. The per-
 2324 formance of ε_{fit} as a function of p_T , shown in Figure A.9 (c), is similar to that of ε_{alg} coincidence
 2325 threshold, again consistent with the effects of secondaries at higher energies.

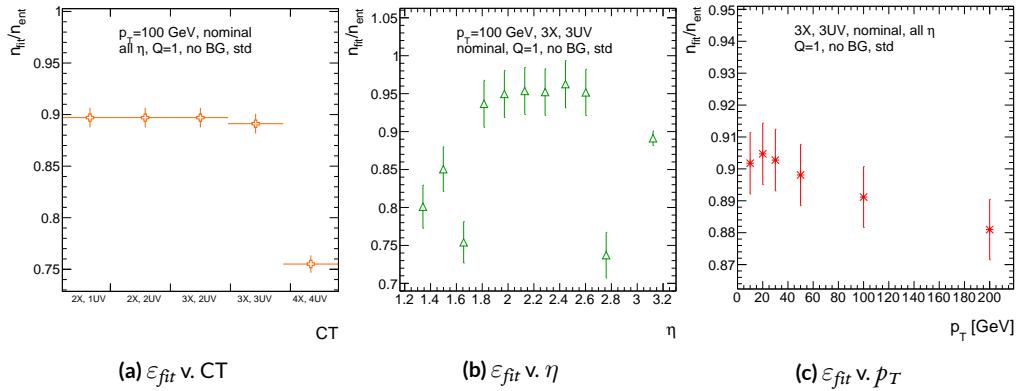


Figure A.9: ε_{fit} and as a function of coincidence threshold, η (final point is ε_{fit} integrated over all η), and p_T .

2326 In order to better understand efficiency behavior with coincidence threshold, the distribution of

2327 highest slope-road coincidence thresholds in events is shown in Figure A.10, with the o,o bin con-
 2328 taining events that did not meet requirements for the minimum $2X+1UV$ coincidence threshold for
 2329 a fit to occur. That the efficiency is lower at higher coincidence threshold suggests that most of the
 2330 fits that fail have high hit multiplicity (i.e. a similar number fails in each of the coincidence thresh-
 2331 old bins in Figure A.8 (a)), which is consistent with the interpretation that the primary source of fit
 2332 failures is bad hits originating from secondaries created by higher energy muons.

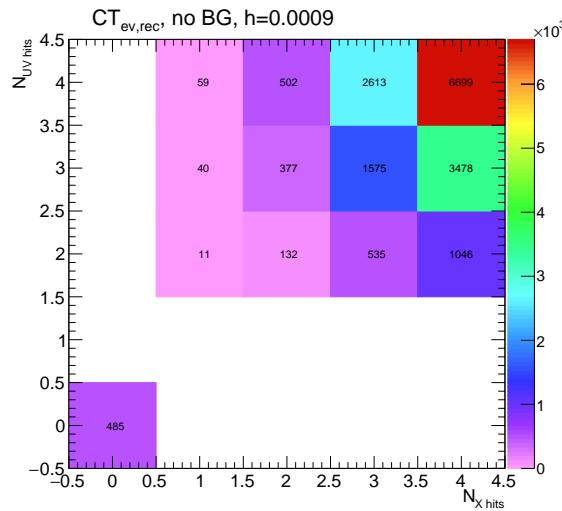


Figure A.10: The distribution of highest slope-road coincidence thresholds in events; the 0,0 bin is the number of events passing selection requirements that fail to form the minimum $2X+1UV$ coincidence threshold necessary for a fit.

2333 A.6 INCOHERENT BACKGROUND

2334 The default slope-road size and tolerances associated with horizontal and stereo hits used in the
2335 above studies were configured to optimize algorithm performance, similar to studies in [61]. In or-
2336 der to evaluate algorithm performance under conditions with more limited resources, as might be
2337 expected at run-time, additional studies were conducted with the slope-road size and hit tolerances
2338 set equivalent to the sensitive area of a single VMM chip[¶] both with and without generation of inco-
2339 herent background.

2340 Incoherent background is generated based on the assumption that the intensity only varies as a
2341 function of the distance from a point to the beamline, r . The number of hits per unit area per unit
2342 time as a function of r is given in Equation A.6 and taken from [61].

$$I = I_0 (r/r_0)^{-2.125} \quad (\text{A.6})$$

2343 where $r_0 = 1000$ mm and $I_0 = 0.141$ kHz/mm²

2344 Background generation happens per event as follows:

- 2345 1. Determine the total number of hits to be generated in this event according to a Poisson distri-
2346 bution
- 2347 2. Assign a time to hits uniformly in $[t_{\text{start}} - t_{\text{VMM}}, t_{\text{end}}]$ where start and end are for the event
2348 clock and t_{VMM} is the VMM chip deadtime (100 ns)
- 2349 3. Assign a plane to hits uniformly
- 2350 4. Assign a ϕ value to hits uniformly

[¶]One VMM is assumed to cover 64 MM strips at 0.445 mm each.

2351 5. Assign an r to hits according to Equation A.6

2352 6. Calculate hit information according to these values.

2353 The expectation value for the Poisson distribution is determined by integrating Equation A.6

2354 over the surface area of the wedge to get the total hit rate for the wedge, Γ , and then multiplying this

2355 by the length of the time window over which hits may be generated. With $H = 982$ mm, $b_1 =$

2356 3665 mm, and $\theta_w = 33\pi/180$, we find^{||}:

$$\Gamma = 2I_0 r_0^{2.125} \int_0^{\theta_w/2} d\phi \int_{H \sec \phi}^{(H+b_1) \sec \phi} r dr r^{-2.125} = 98.6657 \text{ MHz} \quad (\text{A.7})$$

2357 In this case, we have taken the nominal values of the MM sector geometry for H (wedge base), b_1

2358 (the wedge height), and θ_w (the wedge opening angle).

2359 The effects of incoherent background and larger slope road size are summarized in Figure A.11 for

2360 efficiencies and in Figure A.13 and Table A.1 for residual of fit quantities.

2361 Figure A.11 show the effect of both wider slope-roads and the introduction of background on ef-

2362 ficiencies. The introduction of wider slope-roads increases the chance that an early errant hit (either

2363 from secondaries/ionization or background) will be introduced into the fit, and the presence of in-

2364 coherent background greatly increases the number of such errant hits. Both wider slope-roads and

2365 background drive down the number of fits (numerator) in both efficiencies, and background can

2366 artificially inflate the denominator of ε_{alg} , a reco-level, slope-road coincidence threshold. The shape

2367 of the ε_{fit} versus coincidence threshold distributions remains fairly constant with each complicat-

^{||}Using Mathematica and the extra factor of r from the volume element

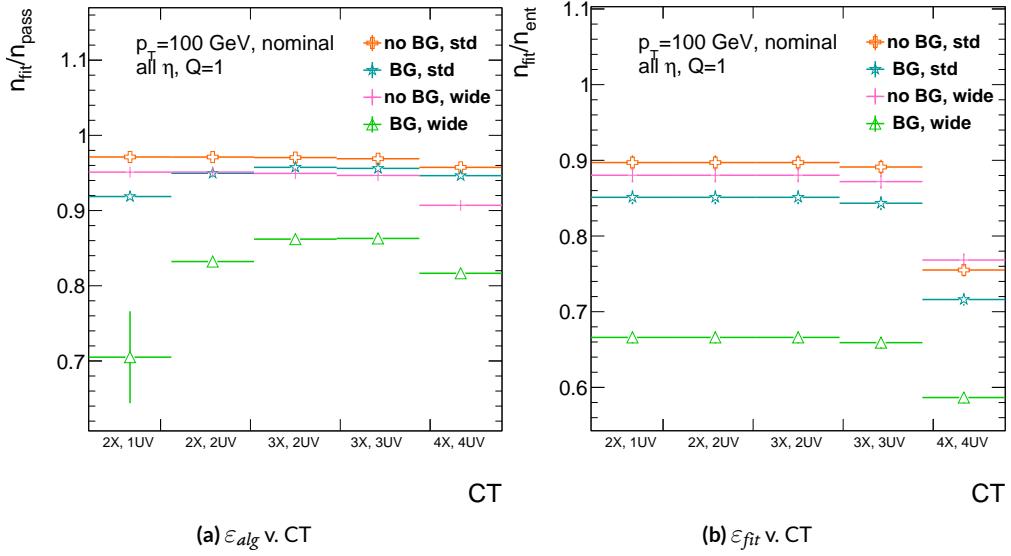


Figure A.11: The algorithm and total efficiencies as a function of coincidence threshold for different background settings and slope-road sizes (standard and wide (one slope road as 1 VMM chip)).

ing factor (standard, wider slope-roads, background, both wider slope-roads and background), suggesting many muons will simply not be fit with any number of hits; ε_{fit} does not take into account the coincidence threshold of tracks that are not fit, so the effect appears uniform across coincidence threshold. The effects seen for ε_{alg} , which are not uniform across coincidence threshold can be better understood when examining the distribution of event highest coincidence thresholds, shown for wide slope-roads both without and with background in Figure A.12. Take, for example the 2X+1UV case. The 2X+1UV bin in particular has a marked increase when background is introduced. No new good tracks are introduced between the no background and background cases, so the increase is entirely due to bad, background hits; hence, these events do not (and should not) fit, causing the particularly pronounced drop in this bin between these two cases in Figure A.11.

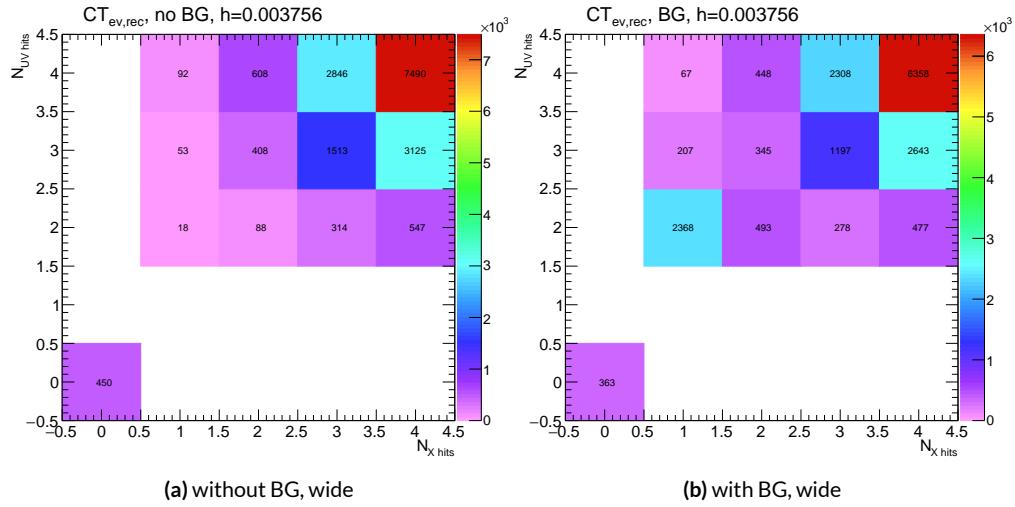


Figure A.12: The distribution of highest slope-road coincidence thresholds in events for the algorithm with wide slope-roads (width of 1 VMM) both without (a) and with (b) incoherent background; the 0,0 bin is the number of events passing selection requirements that fail to form the minimum $2X+1UV$ coincidence threshold necessary for a fit.

The effect of increasing slope-road size and incoherent background on fit quantity residual rms values as a function of p_T is shown in Figure A.13. As the figure shows, the fitted rms values are fairly robust against increased slope-road size and background. This does not hold for all of the raw rms values, however, as shown in Table A.1. Just as with the efficiencies, the introduction of background has a larger effect than that of increased slope-road size, which does not seem to have an overly large impact on any of the fit quantities on its own. While $\Delta\theta$ remains robust to both increased slope-road size and background (likely due to the $\Delta\theta$ cut of 16 mrad built into the algorithm), θ shows some degradation in performance, and the ϕ residual raw rms shows a very large increase upon the introduction of background. Nevertheless, the contrasting behavior of the fitted and raw rms values suggests that tracks that drive up the raw rms values already had very poor fit quality even before the introduction of background, so the impact on fit quantities should remain fairly limited.

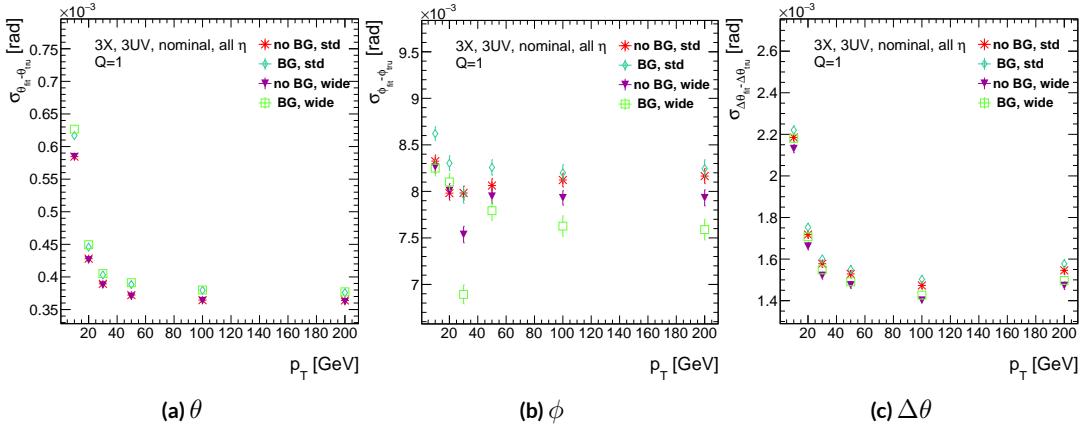


Figure A.13: The three fit quantity residual rms values as a function of p_T for different background settings and slope-road sizes (standard and wide (one slope road as 1 VMM chip)).

	No BG, std	No BG, wide	BG, std	BG, wide
θ	0.364 (0.604)	0.363 (0.542)	0.379 (0.886)	0.380 (1.07)
ϕ	8.12 (15.0)	7.93 (13.2)	8.20 (24.6)	7.63 (24.8)
$\Delta\theta$	1.47 (2.69)	1.40 (2.66)	1.50 (2.89)	1.43 (2.90)

Table A.1: The fitted (absolute) σ of fit quantity residuals in mrad under different algorithm settings.

2389 As Table A.1 shows, rms values appear to be robust to an increase in slope-road size. Neverthe-
2390 less, though the fitted σ residual values are also fairly robust to the introduction of background, the
2391 raw rms values are not. While the raw $\Delta\theta$ rms stays stable, both θ and ϕ suffer noticeable degra-
2392 dation, which suggests that the introduction of background has a detrimental effect on horizontal
2393 slope residual (i.e. on stereo strips in particular). This level of degradation is likely acceptable for θ ,
2394 though further steps may need to be taken to address ϕ .

2395 A.7 BCID

2396 A fitted track's BCID is determined by the most common BCID associated with its hits. Concerns
2397 were raised that this might cause incorrect BCID association for fitted tracks. In order to address
2398 this, the rate of successful BCID association for fitted tracks was recorded. Figure A.14 shows the
2399 dependence of this success rate as a function of p_T and coincidence threshold in the different back-
2400 ground and resource conditions used in the previous section. The successful BCID identification
2401 rate is always over 99.5%, demonstrating that this issue is not a concern with the state-of-the-art de-
2402 tector simulation.

2403 A.8 CHARGE THRESHOLD

2404 The MMTP uses the first hits registered passing a charge threshold requirement given in units of
2405 electron charge. In principle, it would be beneficial to be able to use any hits that are registered re-
2406 gardless of deposited charge, but in the high rate environment envisioned for the NSW, this require-
2407 ment might need to be raised. Nominal algorithm settings have this charge threshold requirement

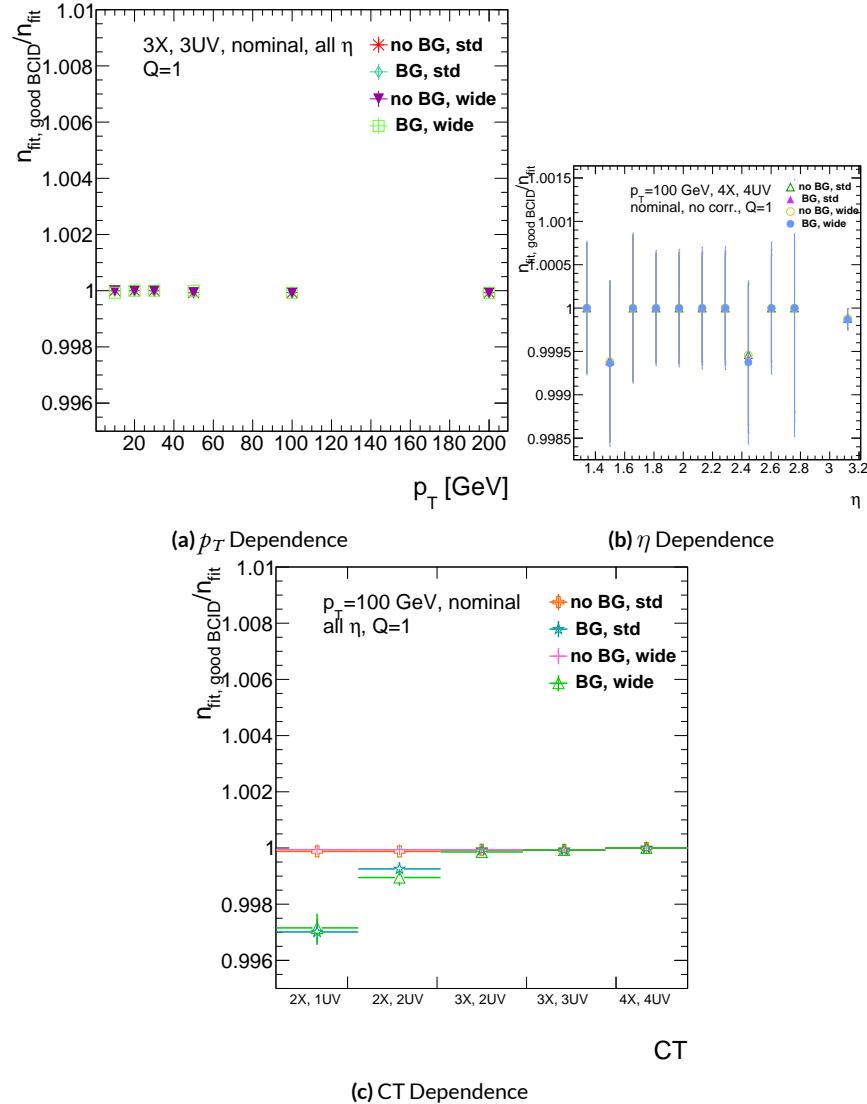


Figure A.14: The rate of good BCID association based majority hit BCID as a function of p_T and coincidence threshold.

2408 set to 1, and studies were conducted on algorithm performance for charge threshold values of 0, 1,
 2409 and 2. Efficiencies as a function of coincidence threshold for different charge thresholds are shown
 2410 in Figure A.15. Increasing the charge threshold lowers both efficiencies, particularly at high coinci-
 2411 dence threshold, which suggests that energetic muons with secondaries create both very many hits
 2412 and hits with higher charge. While the shapes of the fit quantity distributions as a function of p_T in
 2413 Figure A.16 are fairly constant across charge threshold, performance is not. θ and $\Delta\theta$ show some im-
 2414 provement with higher charge threshold, particularly at low p_T , suggesting that resolution improves
 2415 in the vertical direction, but ϕ shows degradation at higher charge threshold, which is a symptom
 2416 of more highly charged particles experiencing greater bending in the ATLAS magnetic field in the ϕ
 2417 direction.

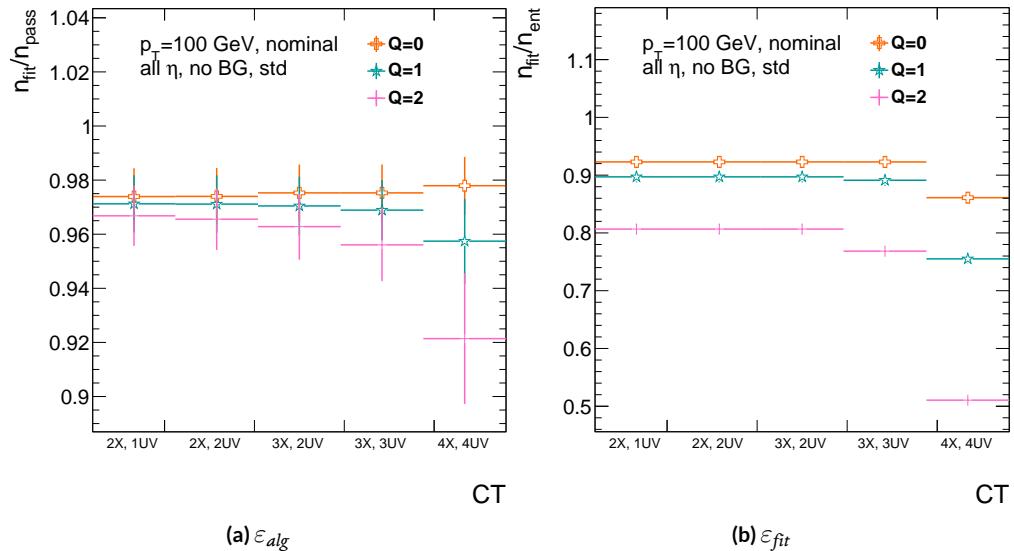


Figure A.15: The efficiencies as a function of coincidence threshold for charge thresholds of 0, 1, and 2.

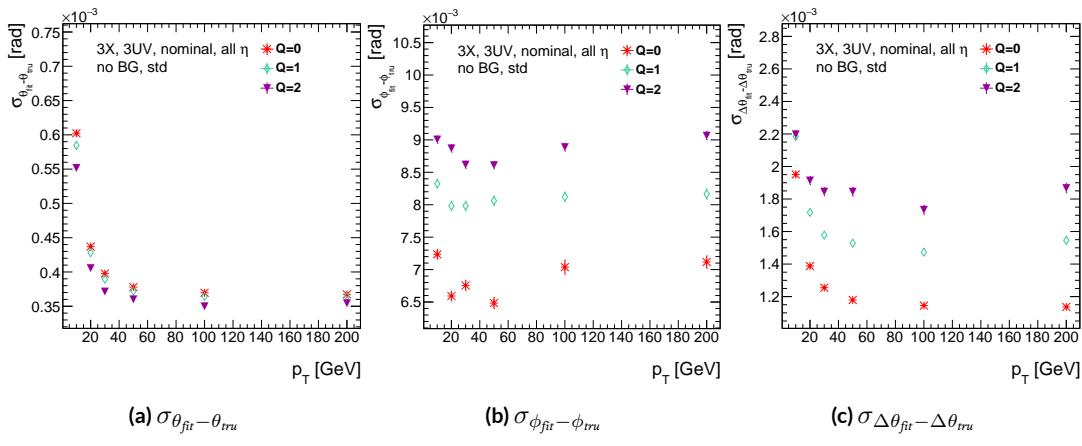


Figure A.16: The fit quantity residual rms values as a function of p_T for charge thresholds of 0, 1, and 2.

2418 A.9 MISALIGNMENTS AND CORRECTIONS

2419 The performance of the trigger algorithm under misalignment has been studied for each of the six
2420 alignment quantities (three translations and three rotations all along the principal axes) described
2421 in [?] and [?], whose convention we will follow here. For the simulated wedge studied here the
2422 local coordinates described in [?] are taken to be centered at the center of the base of the wedge^{**},
2423 the local t axis corresponds to the axis of the beam line, the local z axis corresponds to the direction
2424 orthogonal to both the beam line and the horizontal strips, and the local s axis completes the right-
2425 handed coordinate system. The rotation angles α , β , and γ correspond to rotations around the local
2426 t , z , and s axes, respectively. Note that the local s , z , and $-t$, axes correspond to the usual global x , y ,
2427 and z axes. Misalignments were studied in twenty evenly spaced increments from nominal positions
2428 to misalignments of 1.5 mrad for the rotations (-1.5 mrad to +1.5 mrad for the γ case), and of 5 mm
2429 (a roughly corresponding linear shift) for the translations. In all cases, the front quadruplet is mis-
2430 aligned while the rear quadruplet remains in its nominal position. While only the front quadruplet
2431 of a single wedge is misaligned, the framework for misalignment presented below could be used to
2432 study generic local and global misalignments. The six misalignments are schematically represented
2433 in Figure A.17.

2434 Chamber misalignments manifest themselves as altered strips in algorithm input. In order to sim-
2435 ulate the effects of misalignment, the change in the local y coordinate—the distance from the bot-

2436 ^{**}Not, as is sometimes the case, the centroid position for simplicity's sake, as the agreed upon geometry
of the detector changed several times while studies were in progress; any transformation in a centroid-origin
coordinate system can of course be formed by a combination of the six transformations examined.

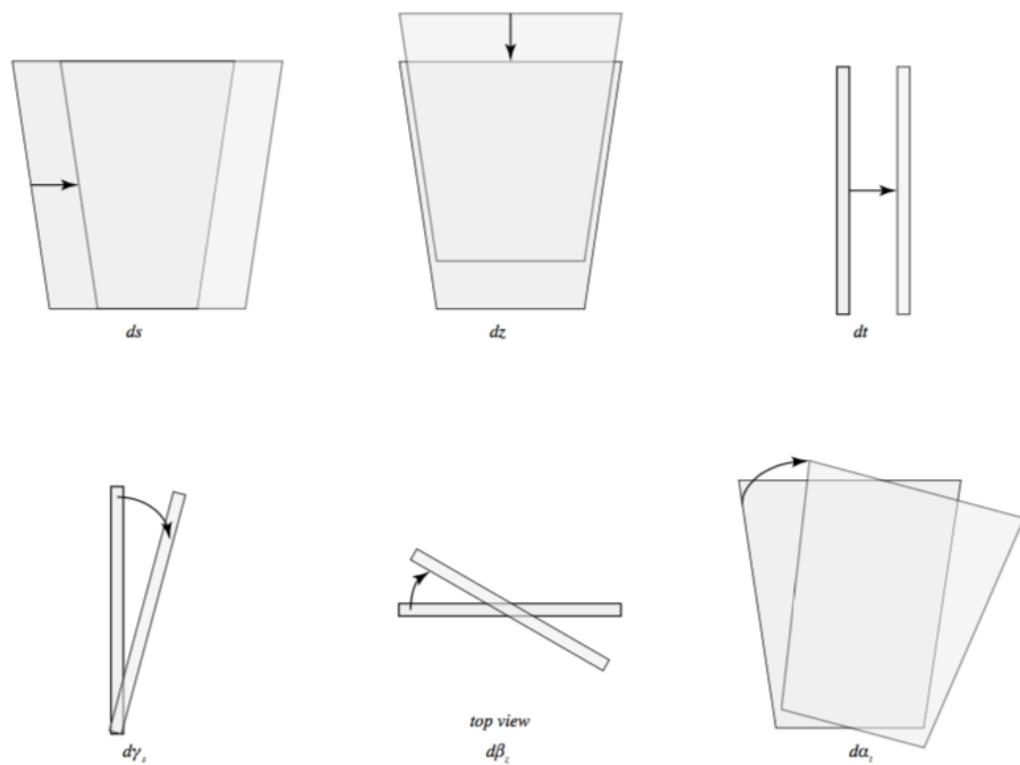


Figure A.17: The different misalignment cases as defined in the AMDB manual.

2436 tom wedge center in the direction perpendicular to both the beamline and the strip direction—is
 2437 calculated for a track coming straight from the interaction point defined by the truth-level θ and ϕ
 2438 angles for generic misalignment. This displacement in y is then added to input hit information and
 2439 the algorithm is then run normally.

2440 To understand how this displacement is calculated, some notation first needs to be described.

Table A.2: A summary of notation used in this section: note that non-AMDB notation is used in this section.

Symbol	Definition
s_x, s_y, s_z, \vec{s}	Position of the muon hit in ATLAS global coordinates; the infinite momentum muon track
\hat{n}	Vector normal to the plane; taken to be \hat{z} (the beamline) in the nominal case
$\vec{\mathcal{O}}_{IP}^{g,l}$	Position of the interaction point in ATLAS global (g) or wedge local (l) coordinates
$\vec{\mathcal{O}}_{base}^{g,l}$	Position of the plane base in ATLAS global (g) or wedge local (l) coordinates; $(0, y_{base}, z_{pl})$ ($(0, 0, 0)$) for the nominal case in global (local) coordinates
$\vec{\zeta}$	$\vec{s} - \vec{\mathcal{O}}_{base}$
primed quant.	quantities after misalignment

2441 Generically speaking, a hit is the intersection of a line (the muon track) with a plane (the individual plane in the multiplet). We assume the muon moves in a straight line defined by the origin and
 2442 the truth-level θ_{pos} and ϕ_{pos} (i.e. the infinite momentum limit) and that the MM plane is rigid and
 2443 defined by a point, which we take to be the center of the bottom edge of the plane, and a normal
 2444 vector, which we take to be the z axis in the nominal case.

2446 The coordinate axes x, y, z axes used here correspond to the usual AMDB $s, z, -t$ axes. Since the
 2447 direction does not really matter when studying misalignment or corrections thereof, the major dif-

²⁴⁴⁸ ference is the choice of origin.

²⁴⁴⁹ The muon track we denote^{††} \vec{s} , the bottom point of the plane $\vec{\mathcal{O}}_{base}$, and the normal vector \hat{n} .

²⁴⁵⁰ The muon track will always be given as (the wedge gets moved, not the muon):

$$\vec{s} = \mathcal{O}_{IP} + k\hat{s} \quad (\text{A.8})$$

$$\hat{s} = \sin \theta_{pos} \sin \phi_{pos} \hat{x} + \sin \theta_{pos} \cos \phi_{pos} \hat{y} + \cos \theta_{pos} \hat{z} \quad (\text{A.9})$$

$$\vec{s} = k\hat{s} = \frac{z_{pl}}{\cos \theta_{pos}} \hat{s} = z_{pl} (\tan \theta \sin \phi \hat{x} + \tan \theta \cos \phi \hat{y} + 1) \quad (\text{A.10})$$

²⁴⁵¹ where $k \in \mathbb{R}$, along with the unit vector \hat{s} , defines the point where the track intersects the wedge.

²⁴⁵² Rotations are done before translations, according to the order prescribed in the AMDB guide for

²⁴⁵³ chamber alignment, so the axes the principal axes of the plane are rotated according to the following

²⁴⁵⁴ matrix (where s , c , and t are the obvious trigonometric substitutions)

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & c\gamma & -s\gamma \\ 0 & s\gamma & c\gamma \end{pmatrix} \begin{pmatrix} c\beta & 0 & s\beta \\ 0 & 1 & 0 \\ -s\beta & 0 & c\beta \end{pmatrix} \begin{pmatrix} c\alpha & -s\alpha & 0 \\ s\alpha & c\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c\gamma & -s\gamma \\ 0 & s\gamma & c\gamma \end{pmatrix} \begin{pmatrix} c\alpha c\beta & -s\alpha c\beta & s\beta \\ s\alpha & c\alpha & 0 \\ -c\alpha s\beta & s\alpha s\beta & c\beta \end{pmatrix} \\ & = \boxed{\begin{pmatrix} c\alpha c\beta & -s\alpha c\beta & s\beta \\ s\alpha c\gamma + c\alpha s\beta s\gamma & c\alpha c\gamma - s\alpha s\beta s\gamma & -c\beta s\gamma \\ s\alpha s\gamma - c\alpha s\beta c\gamma & c\alpha s\gamma + s\alpha s\beta c\gamma & c\beta c\gamma \end{pmatrix}} = \mathcal{A} \end{aligned} \quad (\text{A.11})$$

^{††}Recall ϕ_{pos} is defined with respect to the y axis instead of the x axis, as might otherwise be typical.

2455 The thing that matters is what the new strip hit is—i.e. what the new y value is since this, along
 2456 with a plane number, is all that is fed into the algorithm. To find this, we must solve for the new
 2457 point of intersection with the rotated plane and then apply the effects of translations. The path con-
 2458 necting the base of the wedge with the intersection of the muon track will always be orthogonal to
 2459 the normal vector of the plane. Our quantities after misalignment, denoted by primed quantities,
 2460 will look like

$$\mathcal{O}_{base} \rightarrow \mathcal{O}_{base} + ds\hat{x} + dz\hat{y} + dt\hat{z} = \mathcal{O}'_{base}, \hat{n} \rightarrow A\hat{n} = A\hat{z} = \hat{z}', \vec{s} \rightarrow k'\hat{s} + \mathcal{O}_{IP} = \vec{s}' \quad (\text{A.12})$$

2461 so, moving to explicit, global coordinates in the last line so we can do the computation (relying on
 2462 the fact that any vector in the wedge, namely $\vec{\zeta} = \vec{s} - \mathcal{O}$ the local coordinates of the interaction
 2463 point, is necessarily orthogonal to \hat{n}):

$$0 = \hat{n} \cdot (\vec{\mathcal{O}}_{base} - \vec{s}) \rightarrow 0 = A\hat{z}' \cdot (\vec{\mathcal{O}}'_{base} - (k'\hat{s} + \vec{\mathcal{O}}_{IP})) \quad (\text{A.13})$$

$$\rightarrow k' = \frac{s\beta\vec{\mathcal{O}}'_{base-IP,x} - c\beta s\gamma\vec{\mathcal{O}}'_{base-IP,y} + c\beta c\gamma\vec{\mathcal{O}}'_{base-IP,z}}{\hat{s} \cdot \hat{z}'} \quad (\text{A.14})$$

$$= \frac{s\beta ds - c\beta s\gamma(y_{base} + dz) + c\beta c\gamma(z_{pl} + dt)}{s\beta s\theta s\phi - c\beta s\gamma s\theta c\phi + c\beta c\gamma c\theta} \quad (\text{A.15})$$

2464 To find our new y coordinate, we need to evaluate $s'_y = \hat{y}' \cdot k'\vec{s}$ to find the final correction of:

$$\Delta y = \vec{\zeta}' \cdot \hat{y}' - \vec{\zeta} \cdot \hat{y} = (k'\hat{s} - \vec{\mathcal{O}}'_{base}) \cdot \hat{y}' - (s_y - y_{base}) \quad (\text{A.16})$$

²⁴⁶⁵ The correction will be plane dependent since (denoting the stereo angle ω):

$$\hat{y}_x = \hat{y} \rightarrow \hat{y}'_x = -s\alpha c\beta \hat{x} + (c\alpha c\gamma - s\alpha s\beta s\gamma) \hat{y} + (c\alpha s\gamma + s\alpha s\beta c\gamma) \hat{z} \quad (\text{A.17})$$

²⁴⁶⁶ and

$$\begin{aligned} \hat{y}_{U,V} = & \pm s\omega \hat{x}' + c\omega \hat{y}'_{U,V} = [\pm c\alpha c\beta s\omega - s\alpha c\beta c\omega] \hat{x} + [\pm (s\alpha c\gamma + c\alpha s\beta s\gamma) s\omega \\ & + (c\alpha c\gamma - s\alpha s\beta s\gamma) c\omega] \hat{y} + [\pm (s\alpha s\gamma - c\alpha s\beta c\gamma) s\omega + (c\alpha s\gamma + s\alpha s\beta c\gamma) c\omega] \hat{z} \end{aligned} \quad (\text{A.18})$$

²⁴⁶⁷ A.10 INDIVIDUAL CASES

²⁴⁶⁸ Currently we only study the cases where one misalignment parameter is not zero. We examine these
²⁴⁶⁹ in detail below, calculating the most pertinent quantities in the misalignment calculation, k'/k and
²⁴⁷⁰ the new horizontal and stereo y axes. Before setting out, we simplify the expressions for the trans-
²⁴⁷¹ formed \hat{y}' 's, removing any terms with the product of two sines of misalignment angles, which will be
²⁴⁷² zero.^{††}

$$\hat{y}'_x = -s\alpha c\beta \hat{x} + c\alpha c\gamma \hat{y} + c\alpha s\gamma \hat{z} \quad (\text{A.19})$$

²⁴⁷³

$$\hat{y}'_{U,V} = [\pm c\alpha c\beta s\omega - s\alpha c\beta c\omega] \hat{x} + [\pm s\alpha c\gamma s\omega + c\alpha c\gamma c\omega] \hat{y} + [\mp c\alpha s\beta c\gamma s\omega + c\alpha s\gamma c\omega] \hat{z} \quad (\text{A.20})$$

^{††}If only one misalignment parameter is non-zero, then two or more sines will contain at least one term will contain $\sin 0 = 0$.

²⁴⁷⁴ If the translations are zero,

$$k' = \frac{-c\beta s\gamma y_{base} + c\beta c\gamma z_{pl}}{s\beta s\theta s\phi - c\beta s\gamma s\theta c\phi + c\beta c\gamma c\theta}, \quad k'/k = \frac{-c\beta s\gamma y_{base}/z_{pl} + c\beta c\gamma}{s\beta t\theta s\phi - c\beta s\gamma t\theta c\phi + c\beta c\gamma} \quad (\text{A.21})$$

²⁴⁷⁵ A.II $ds \neq 0$

²⁴⁷⁶ $k'/k = 1$ (the point of intersection does not move closer or further from the IP), and only the stereo planes are affected. Note that only relevant term in Equation A.16, for the stereo strip \hat{y} for $\vec{\mathcal{O}}'_{base} =$
²⁴⁷⁷ $ds\hat{x}$ is:

$$\pm \sin \omega ds \approx \pm 0.0261 ds \quad (\text{A.22})$$

²⁴⁷⁹ meaning that a displacement in x of 17 mm, more than three times the range of misalignments studied, would be necessary for a shift in the stereo planes corresponding to one strip width.

²⁴⁸¹ A.12 $dz \neq 0$

²⁴⁸² $k'/k = 1$ (the point of intersection does not move closer or further from the IP). This case is the trivial one (cf. Equation A.16 with $\vec{\mathcal{O}}'_{base} = dz\hat{y}$). y just gets moved in the opposite direction as the wedge. Correction is an additive constant.

²⁴⁸⁵ A.13 $dt \neq 0$

²⁴⁸⁶ $k'/k = (z_{pl} + dt) / z_{pl}$. y gets modified by a simple scale factor. Correct by storing changing definitions of plane positions in algorithm to match the misaligned values.

²⁴⁸⁸ A.14 $\alpha \neq 0$

²⁴⁸⁹ $k'/k = 1$ and

$$\hat{y}'_x = -s\alpha\hat{x} + c\alpha\hat{y} \quad (\text{A.23})$$

$$\hat{y}'_{U,V} = [\pm c\alpha s\omega - s\alpha c\omega] \hat{x} + [\pm s\alpha s\omega + c\omega] \hat{y} \quad (\text{A.24})$$

²⁴⁹⁰ A.15 $\beta \neq 0$

²⁴⁹¹ We have $k'/k = (1 + \tan \beta \tan \theta \sin \phi)^{-1}$, and

$$\hat{y}'_x = \hat{y} \quad (\text{A.25})$$

$$\hat{y}'_{U,V} = \hat{y} \pm (c\beta\hat{x} - s\beta\hat{z}) s\omega \quad (\text{A.26})$$

²⁴⁹² A.16 $\gamma \neq 0$

$$k'/k = \frac{1 - \tan \gamma \frac{y_{base}}{z_{pl}}}{1 - \tan \gamma \tan \theta \cos \phi} \quad (\text{A.27})$$

$$\hat{y}'_x = c\gamma\hat{y} + s\gamma\hat{z} \quad (\text{A.28})$$

$$\hat{y}'_{U,V} = \pm s\omega\hat{x} + c\omega\hat{y} - s\gamma c\omega\hat{z} \quad (\text{A.29})$$

²⁴⁹³ In order to evaluate algorithm performance under misalignment and corrections for misalign-

²⁴⁹⁴ ment, the absolute means and relative resolutions of the fit quantities θ , ϕ , and $\Delta\theta$ are measured as

2495 a function of misalignment. In the following, results will only be shown for which the effects of mis-
2496 alignment are significant. “Significant,” for misalignments of 1 mm (0.3 mrad) for translations (ro-
2497 tations) means more than a 5% degradation in rms and/or bias shifts in θ , ϕ , and $\Delta\theta$ of 0.01 mrad, 1
2498 mrad, and 0.1 mrad, respectively.

2499 While corrections are typically done on a case-by-base basis, they fall under two general cate-
2500 gories, analytic and simulation based. Analytic corrections rely upon specific knowledge of the mis-
2501 alignment, with each case being handled separately; as such, the additional resources required, both
2502 extra constants and operations, if any, vary accordingly. Simulation based corrections are all done in
2503 the same manner. The algorithm is run over a training MC sample (same setup but with $p_T = 200$
2504 GeV instead of the normal 100 GeV sample so as not to overtrain the corrections), and the mean bi-
2505 ases for θ , ϕ , and $\Delta\theta$ are saved for different, equally spaced regions in the $\eta - \phi$ plane over the wedge
2506 based on the fitted θ and ϕ values. Currently, these values are saved for 10 η and 10 ϕ bins (100 η, ϕ
2507 bins total), with the number of bins in each direction being a configurable parameter. When the al-
2508 gorithm runs with simulation based correction, this table of constant corrections is saved in a LUT
2509 before runtime, and corrections are added to final fit quantities based on the (uncorrected) θ and
2510 ϕ fit values. With the settings mentioned, this is 300 extra constants ($10\eta\text{-bins} \times 10\phi\text{-bins} \times 3$ fit
2511 quantities) and two extra operations (a lookup and addition for each quantity done in parallel). The
2512 simulation correction can, in principle, also be applied to the algorithm in nominal conditions with
2513 non-trivial improvements, as detailed below in Section A.17. Depending on the misalignment case in
2514 question, different approaches work better. A summary of correction methods, including resources
2515 necessary for the individual analytic cases, is shown in Table A.3.

	Δ_s	Δz	Δt	γ_s	β_z	α_t
Analytic Resources	yes+ 11c/2op	yes+ oc/oop	yes+ oc/oop	yes 56c/1op	no —	yes 400c/2n _X op, 32c/12n _X op
Simulation	yes+	no	no	no	yes+	yes+

Table A.3: A summary of corrections with additional constants/operations (written as $n_{const}c/n_{ops}op$; n_X is the number of X hits in a fit) necessary for analytic corrections. Yes means a correction exists but might not entirely remove misalignment effects, while yes+ means a quality of correction is only limited by knowledge of misalignment and memory

2516 A.17 SIMULATION CORRECTION OF THE ALGORITHM UNDER NOMINAL CONDITIONS

2517 In addition to using simulation based correction to counter the effects of several classes of misalign-
 2518 ment, the correction can be applied at to the algorithm under nominal conditions. The main effect
 2519 of this correction is to mitigate the effects of the bias in stereo strips. As such, the correction has a
 2520 larger effect on quantities that rely on the aggregate slope m_y , as can be seen in in Figure A.18, im-
 2521 proving $\sigma_{\theta_{fit} - \theta_{true}}$ resolution by about 25%, and reducing $\sigma_{\phi_{fit} - \phi_{true}}$ by over 50% and restoring a largely
 2522 Gaussian shape. The slight, apparent degradation in $\Delta\theta$ is due to a more mild version of the effect
 2523 seen in Figure A.7.

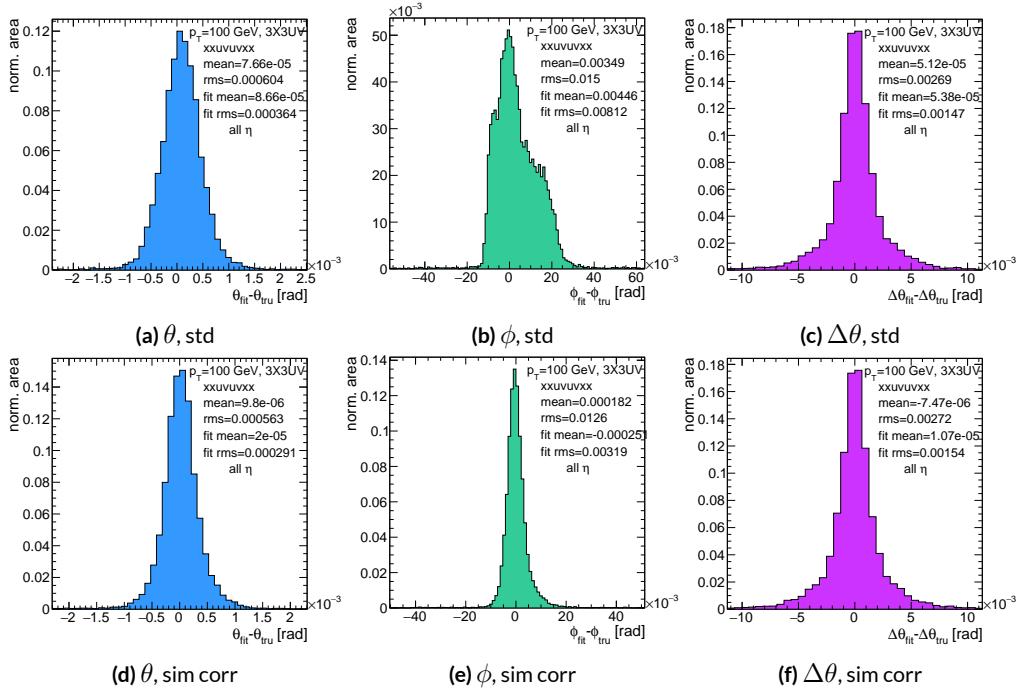


Figure A.18: Nominal residual plots for both uncorrected and simulation corrected cases; $\theta, \phi, \Delta\theta$ for $p_T = 100$ GeV muons

2524 As can be seen in Figure A.19, the simulation based correction also removes the η dependence to
 2525 fit quantity resolution distributions, as expected. One consequence of this is that simulation-based
 2526 corrections applied to the misalignment cases below will restore performance to the “sim” and not
 2527 the “std” distributions of Figure A.18. Hence, when making comparisons between simulation cor-
 2528 rected curves and the nominal performance point, simulation-corrected distributions of benchmark
 quantities versus misalignment will often look generally better.

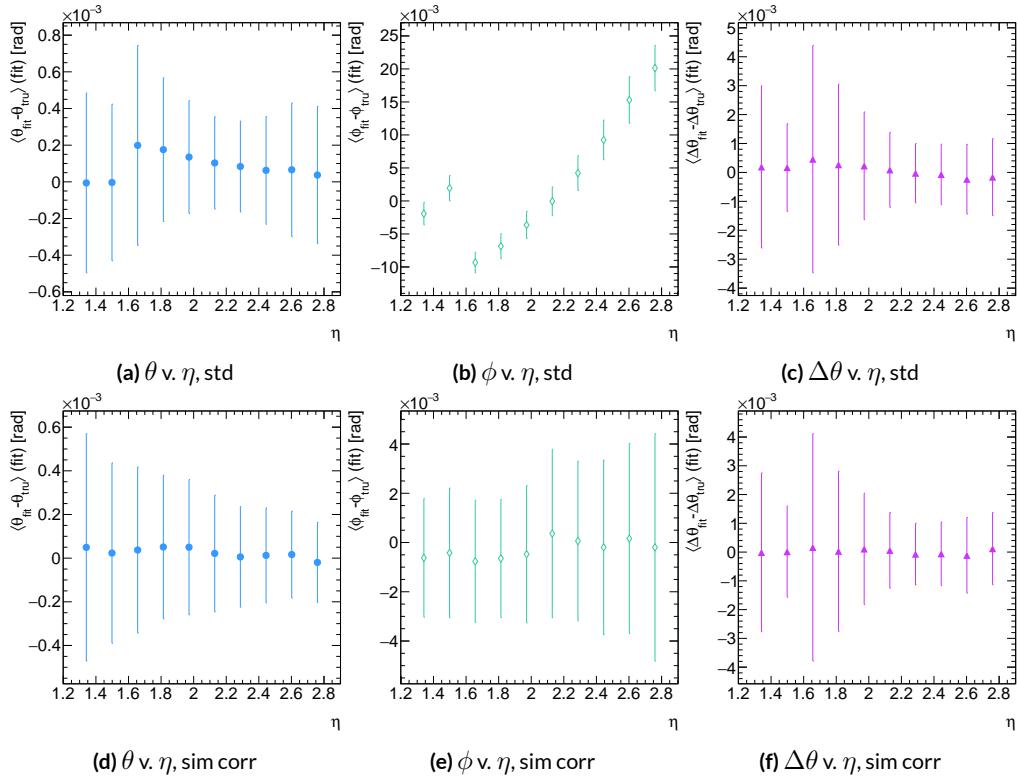


Figure A.19: Nominal residual plots as a function of η with points as means and error bars as rms values in each η bin for the angles $\theta, \phi, \Delta\theta$ for $p_T = 100$ GeV muons in the uncorrected and simulation corrected cases.

2529
 2530 That the improvements from a simulation-based correction improve performance of the algo-

2531 rithm in nominal conditions most for the quantities that depend most on stereo information (ϕ and
 2532 θ) and remove the η dependence of fit quantity resolutions suggests that there could, in principle, be
 2533 analytic corrections that could be applied to the nominal algorithm. One possible solution is to in-
 2534 troduce an additional set of constants, having the y_{base} depend on the strip number, similar to the γ_s
 2535 correction for z_{plane} described in Section A.21, which would add a lookup per hit and $8 \times n_{bins,y}$ extra
 2536 constants that would be optimized as the γ_s correction was.

$$M_{hit} = \frac{\gamma}{z} = \frac{y_{base}}{z_{plane}} (n_{str}) + \frac{w_{str}}{z_{plane}} n_{str} \quad (\text{A.30})$$

2537 The simulation correction residual rms values suggest a limit on the quality of such correction
 2538 and could perhaps be implemented generically on their own regardless of misalignment for rms val-
 2539 ues on fit quantities of 0.291 mrad for θ , 3.19 mrad for ϕ , and 1.54 for $\Delta\theta$, which represent a 20%
 2540 improvement for θ , a 62% improvement for ϕ , and a slight degradation in $\Delta\theta$ of 4.7%, again owing
 2541 to an effect similar to the one in A.7.

2542 A.18 TRANSLATION MISALIGNMENTS ALONG THE HORIZONTAL STRIP DIRECTION (Δs)

2543 A translation in s (i.e. along the direction of a horizontal strip) only affects the stereo strips, and,
2544 since the stereo angle is small, a very large misalignment is necessary for effects to be noticeable (a
2545 misalignment of roughly 17 mm corresponds to one strip's misalignment in the stereo planes). The
2546 only quantity to show any meaningful deviation with misalignments with translations in s is the ϕ
2547 residual bias (a change of 0.4 mrad at $\Delta s = 1$ mm), as can be seen in the uncorrected curve of Figure
2548 A.20.

2549 A translation in s induces a constant shift in the calculated horizontal slope, m_x in Equation A.4.
2550 This constant shift should only depend on which stereo planes included in a fit are misaligned and
2551 how misaligned they are. Hence, the correction to m_x , for a sum over misaligned stereo planes i ,
2552 with their individual misalignments in s and plane positions in z is:

$$\Delta m_x = \frac{1}{N_{\text{stereo}}} \sum_{i, \text{misal stereo}} \frac{\Delta s_i}{z_{i, \text{plane}}} \quad (\text{A.31})$$

2553 Given prior knowledge of misalignment, these corrections to m_x can be performed ahead of time
2554 and saved in a lookup table (LUT), similar to the LUT used for constants in the X local slope (M_x^l)
2555 calculation. The added overhead of this analytic correction is hence eleven constants in memory, a
2556 lookup, and one addition. The correction perfectly corrects the effects of misalignment, as can be
2557 seen in Figure A.20. The simulation based correction described above can also be used to correct
2558 for Δs misalignments, with the results of that correction also shown in Figure A.20. The apparent

2559 discrepancy between the simulated and analytic correction is a natural consequence of the fact that
 2560 the simulation correction, as previously mentioned, restores the ϕ residual distribution to an overall
 2561 more Gaussian shape.

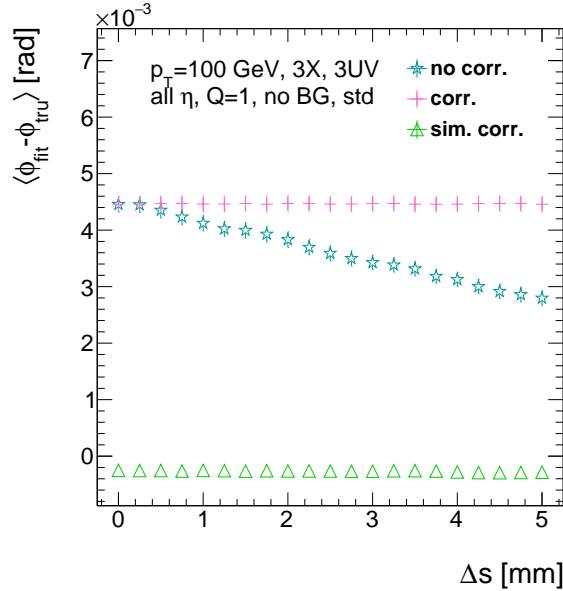


Figure A.20: The mean of the ϕ residual as a function of misalignment for the uncorrected case and the analytic and simulation correction cases.

2562 A.19 TRANSLATION MISALIGNMENTS ORTHOGONAL TO THE BEAMLINE AND HORIZON-
2563 TAL STRIP DIRECTION (Δz)

2564 A translation in AMDB z , the direction orthogonal to both the beamline and the horizontal strip
2565 direction, corresponds to a translation in the y of Equation A.1, affecting all slope calculations. This
2566 has a large impact on the θ residual bias and both the bias and rms of $\Delta\theta$ residual, as can be seen in
2567 Figures A.21 (a)–(c). The marked degradation and non-linear behavior in performance at very high
2568 levels of misalignments is a result of low statistics; there are fewer fits at high level of misalignments
2569 since for $\Delta z \gtrsim 3$ mm, most fits will fail the $\Delta\theta$ cut. The θ bias shifts by about 0.075 mrad at $\Delta z =$
2570 1 mm, and $\Delta\theta$ shifts by about 5 mrad for the same level of misalignment. While the fitted rms of the
2571 $\Delta\theta$ residual remains fairly stable for $\Delta z < 1$ mm or so, between $\Delta z = 2$ mm and $\Delta z = 3$ mm, the
2572 rms increases by 15% before the $\Delta\theta$ cut issue mentioned above intervenes.

2573 Fortunately, these misalignments are straightforward to correct with knowledge of the misalign-
2574 ment. The only modification necessary for this correction is to change the definitions of y_{base} in
2575 Equation A.1 for the individual hit slope addressing. This is done before runtime and adds no over-
2576 head to the algorithm, and the correction quality is only limited by knowledge of the misalignment.
2577 The results of this correction are also shown in Figures A.21 (a)–(c) and restore nominal perfor-
2578 mance.

Since $\Delta\theta = \frac{M_x^l - M_x^e}{1 + M_x^l M_x^e}$ and $M_x^l = B_k \sum y_i (z/\bar{z} - 1)$, a shift Δy translates (with typical slope values of ~ 0.3) to $5B_k (z_1 + z_2)/\bar{z}$ (with B_k in units of inverse mm); set equal to 16 mrad ($\Delta\theta$ is centered at zero), this corresponds to $\Delta y = 2.7$ mm

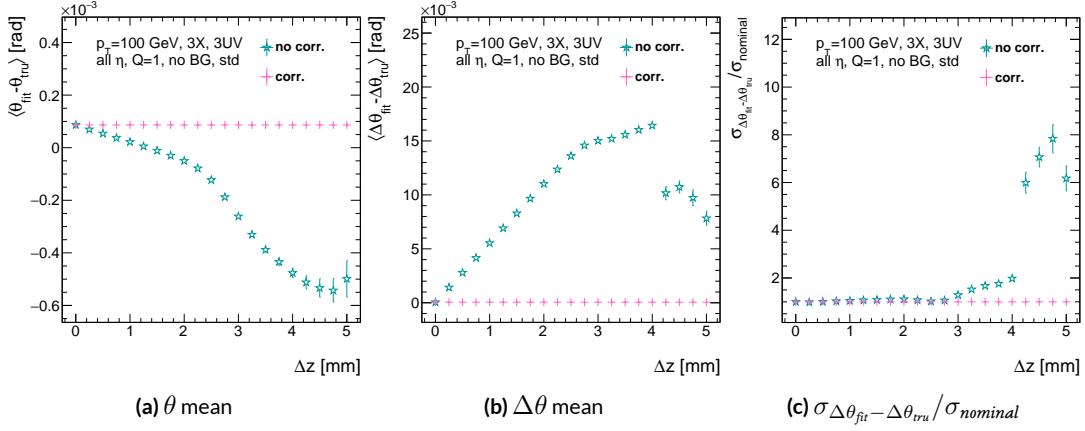


Figure A.21: The affected quantities of Δz misalignments: θ bias, $\Delta\theta$ bias, and $\sigma_{\Delta\theta_{\text{fit}} - \Delta\theta_{\text{tru}}} / \sigma_{\text{nominal}}$ for both the misaligned and corrected cases.

2579 A.20 TRANSLATION MISALIGNMENTS PARALLEL TO THE BEAMLINE (Δt)

2580 The effects of misalignment due to translations in t are very similar to those due to translations in
 2581 z without the complication of the $\Delta\theta$ cut, affecting the z instead of the y coordinate that enters
 2582 into hit slope calculations. Again, θ bias, $\Delta\theta$ bias, and $\sigma_{\Delta\theta_{fit}-\Delta\theta_{true}}$ are the primarily affected quan-
 2583 tities. For $\Delta t = 1$ mm, θ bias shifts by about 0.02 mrad, $\Delta\theta$ bias shifts by just under 2 mrad, and
 2584 $\sigma_{\Delta\theta_{fit}-\Delta\theta_{true}}$ degrades by about 20%. The correction for this misalignment once again costs no over-
 2585 head and consists of changing stored constants in the algorithm, in this case the positions along the
 2586 beamline of the misaligned planes, with results similarly limited by knowledge of the misalignment.

2587 The slight improvement with correction to $\Delta\theta$ rms is due to the real effect of a larger lever arm.

2588 Both the misaligned and corrected distributions of affected quantities of interest are shown in Fig-
 ure A.22.

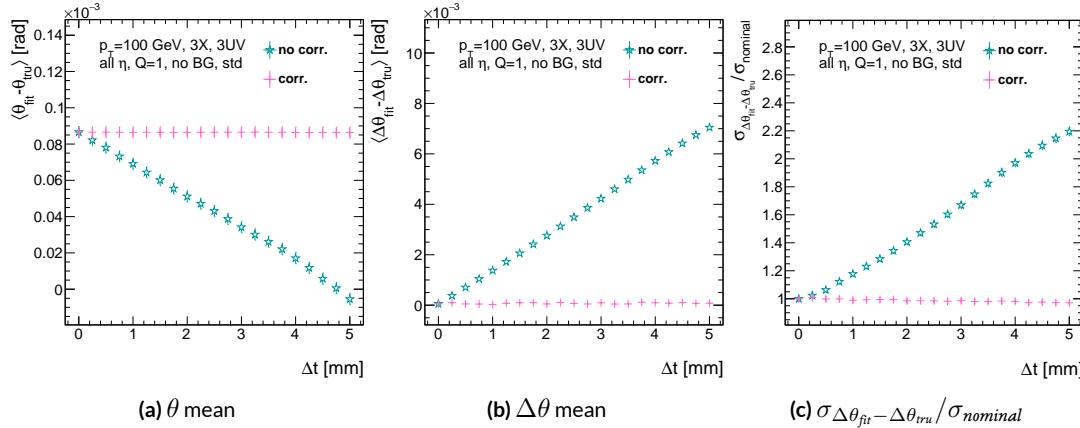


Figure A.22: The affected quantities of Δt misalignments: θ bias, $\Delta\theta$ bias, and $\sigma_{\Delta\theta_{fit}-\Delta\theta_{true}} / \sigma_{nominal}$ for both the misaligned and corrected cases.

2590 A.21 CHAMBER TILTS TOWARDS AND AWAY FROM THE IP (γ_s ROTATION)

2591 Chamber misalignment due to rotations around the s axis act effectively like a translation in t that
2592 depends on strip number. These rotations tilt misaligned chambers away from (towards) the IP for
2593 positive (negative) values of γ_s . Since, unlike for the other two rotation cases that will be studied,
2594 positive and negative rotation values are not symmetric, this misalignment is studied for both posi-
2595 tive and negative γ_s values. The divergent effect at the tails is a result of a large population of fits not
2596 having fit quantities within the cores, and so not appearing in the fit rms. Once again, affected quan-
2597 tities of interest θ bias, $\Delta\theta$ bias, and $\sigma_{\Delta\theta_{fit}-\Delta\theta_{tru}}$. The effects of misalignment can be seen in Figures
2598 A.23 (a)–(c). The relationship between biases and γ_s is roughly linear with $\Delta\gamma_s = 0.3$ mrad (the an-
2599 gular scale corresponding to linear shifts of ~ 1 mm) corresponding to 0.005 mrad (0.12 mrad) for θ
2600 ($\Delta\theta$). For $\sigma_{\Delta\theta_{fit}-\Delta\theta_{tru}}$, degradation is not symmetric. For negative (positive) γ_s , with the quadruplet
2601 tilted towards (away from) the IP, slope-roads are artificially expanded (shrunk), decreasing (increas-
2602 ing) the granularity of the trigger, explaining the asymmetry in Figure A.23 (c), with the degradation
2603 being a 10% (25%) effect for γ_s of $+(-)0.3$ mrad.

2604 Corrections are less simple in this case. In principle, corrections of the same accuracy of the trans-
2605 lations could be calculated per strip, but the overhead of one correction per strip (many thousands
2606 of constants) is prohibitive. Instead, each plane was divided into eight equal segments with a t value
2607 (z in the slope calculation) assigned to strips in each region to correct for the misalignment. This
2608 amounts to 56 extra constants and a 2D instead of a 1D LUT for z positions while the algorithm
2609 runs. The corrected distributions can also be seen in Figures A.23 (a)–(c). The corrections, while not

as effective as for the simple translation cases, are still very effective with the quoted misalignment values for bias shifts down to 0.001 mrad (0.25 mrad) for θ ($\Delta\theta$) and no more than a 2% degradation in $\sigma_{\Delta\theta_{fit}-\Delta\theta_{true}}$ for $|\gamma_s| = 0.3$ mrad.

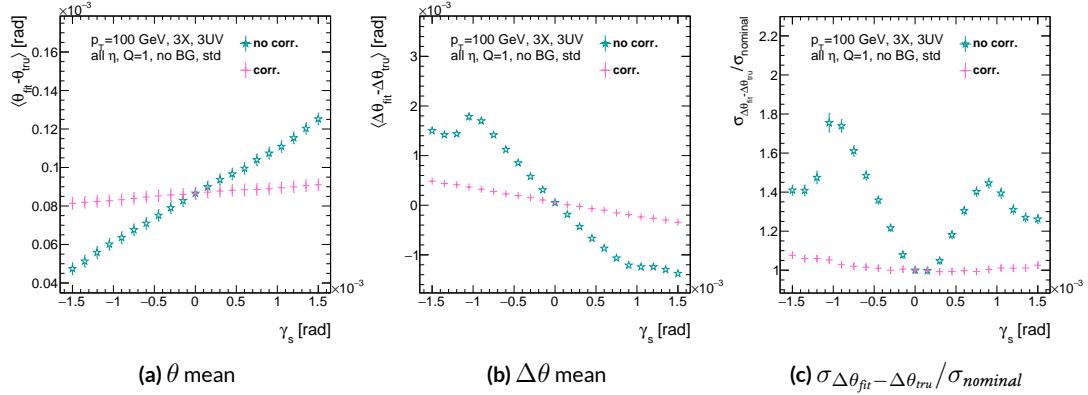


Figure A.23: The noticeable effects of rotations in the s axis and the behavior of these quantities (θ and $\Delta\theta$ bias shifts and $\sigma_{\Delta\theta_{fit}-\Delta\theta_{true}}/\sigma_{nominal}$) with and without misalignment correction.

2612

2613 A.22 ROTATION MISALIGNMENTS AROUND THE WEDGE VERTICAL AXIS (β_z)

2614 While misalignments coming from rotations around the z axis (the direction orthogonal to both
2615 the beamline and the horizontal strip direction) foreshorten the strips as seen from the IP and add
2616 a deviation in t , the long lever arm largely washes out any effects of this misalignment. Only the
2617 $\sigma_{\Delta\theta_{fit}-\Delta\theta_{tru}}$ is noticeably affected, though only at severe misalignments, with only about a 1% degra-
2618 dation in performance at $\beta_z = 0.3$ mrad (corresponding to a linear shift of ~ 1 mm). A simulation
2619 based correction works well to cancel out the effects of this misalignment, and the $\sigma_{\Delta\theta_{fit}-\Delta\theta_{tru}}$ as a
2620 function of misalignment with and without corrections are shown in Figure A.24. The apparent
2621 2% effect in the simulation corrected curve is a result of a more mild version of the effect shown in
2622 Figure A.7.

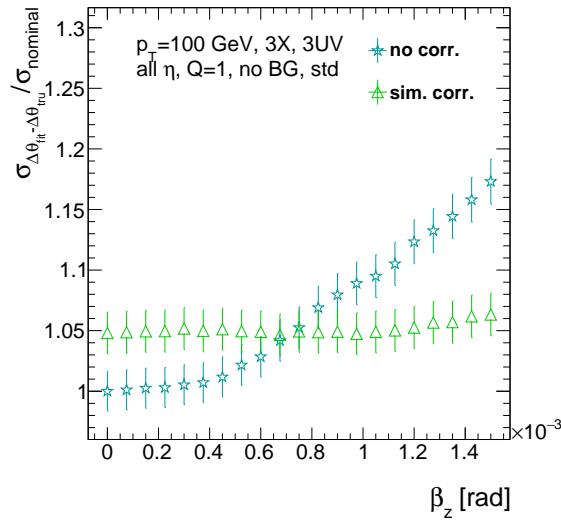


Figure A.24: The effects of rotations in the z axis on $\sigma_{\Delta\theta_{fit}-\Delta\theta_{tru}} / \sigma_{nominal}$ a function of β_z both with and without misalignment corrections.

2623 A.23 ROTATION MISALIGNMENTS AROUND THE AXIS PARALLEL TO THE BEAMLINE (α_t)

2624 Misalignments arising from rotations around the t axis (parallel to the beamline at the center of
2625 the base of the wedge) are essentially rotations in the ϕ direction. The quantities of interest most
2626 affected are the ϕ bias and $\sigma_{\Delta\theta_{fit}-\Delta\theta_{tru}}$, as shown in Figures A.25 (a) and (b), respectively, and cor-
2627 respond to a shift in ϕ bias of 0.2 mrad and a 10% degradation in $\sigma_{\Delta\theta_{fit}-\Delta\theta_{tru}}$ for $\alpha_t = 0.3$ mrad
2628 (corresponding to a linear shift of ~ 1 mm). The raw instead of fitted mean ϕ biases is used in Fig-
2629 ure A.25 (a) to better illustrate the effect of misalignment.

2630 Since the effect of misalignment is dependent on horizontal (along the strip direction, \hat{s}) in addi-
2631 tion to vertical information, corrections cannot be applied before a fit takes place. The ϕ bias shift is
2632 uniform over the entire wedge, so a constant additive correction to ϕ based on the level of misalign-
2633 ment can be applied to all fits depending on how many misaligned stereo planes enter in the fit. $\Delta\theta$
2634 is less straightforward, but corrections to the y and z information used in the local slope calculation
2635 in Equation A.4 can be applied once θ_{fit} and ϕ_{fit} are known. These corrections are calculated ahead
2636 of time in bins of uniform η and ϕ as with the simulation corrections using the same framework
2637 as the misalignment calculation. The results of both types of correction can be seen in Figure A.22.
2638 The apparent discrepancy between the simulation and analytic corrections in the ϕ bias happens for
2639 the same reason as in the Δs misalignment correction cases, as simulation correction restores a more
2640 Gaussian shape to the ϕ residual distribution opposed to the uncorrected nominal case, as discussed
2641 in Section A.17.

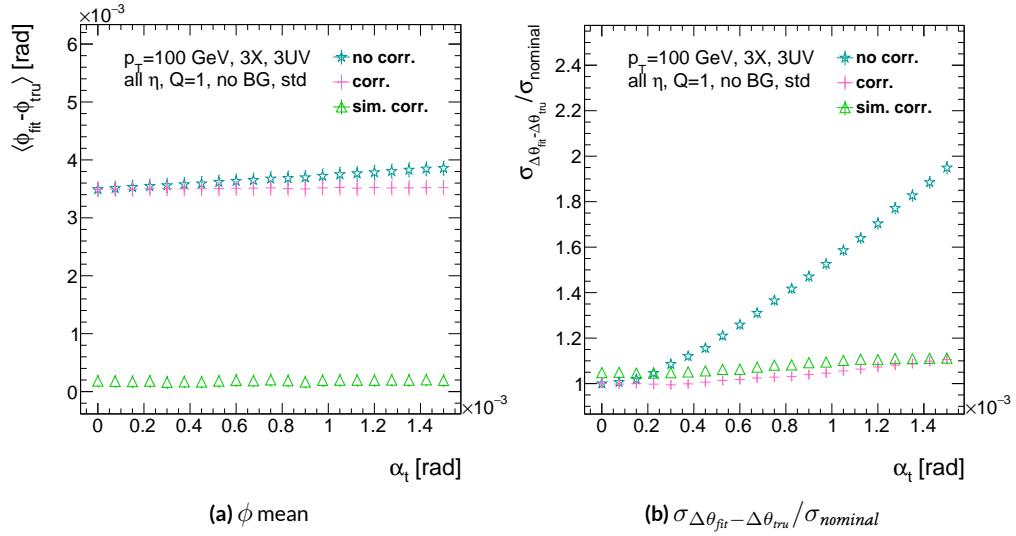


Figure A.25: The effects of rotation misalignments around the t axis for ϕ bias and $\sigma_{\Delta\theta_{\text{fit}} - \Delta\theta_{\text{true}}} / \sigma_{\text{nominal}}$ as a function of misalignment. The uncorrected and both the analytic and simulation correction cases are shown.

2642 A.24 CONCLUSION

2643 The algorithm for Micromegas detectors in the NSW Trigger Processor performs well in a variety of
2644 conditions and has proven robust to a number of effects to deliver measurements on muon tracks
2645 of the three angles θ , ϕ , $\Delta\theta$. Under nominal conditions, the rms values for the residuals of these
2646 quantities are 0.364 mrad for θ , 8.12 mrad for ϕ , and 1.47 mrad for $\Delta\theta$. Algorithm performance was
2647 found to be largely independent of the charge threshold setting, and a hit majority BCID associa-
2648 tion was found to provide proper timing information over 99.7% even in the most relaxed settings
2649 (2X+1UV coincidence threshold requirement+wide slope-road+background). The introduction of
2650 wide slope-roads to better mimic potentially limited algorithm resources at run time and the intro-
2651 duction of incoherent background was found to have a manageable effect on fit quantity residual
2652 rms values and on total algorithm efficiency for sufficiently stringent coincidence threshold. The ef-
2653 ffects of the three translation and three rotation misalignments specified by AMDB convention were
2654 studied, and correction methods for each of the six cases was developed. Simulation-based correc-
2655 tions were found to improve nominal algorithm performance to residual rms value of 0.291 mrad for
2656 θ , 3.19 mrad for ϕ , and 1.54 for $\Delta\theta$, which represent improvements of 20%, 62%, and -4.7%, respec-
2657 tively. Misalignment corrections were found to restore nominal performance for all but the rotation
2658 around the s axis, and a summary of tolerances may be found in Table A.4.

	No Correction	Correction
Δs	4 mm (ϕ bias)	> 5 mm
Δz	0.25 mm ($\Delta\theta$)	> 5 mm
Δt	0.25 mm ($\Delta\theta$)	> 5 mm
γ_s	0.15 mrad ($\Delta\theta$ bias)	0.75 mrad
β_z	0.9 mrad ($\Delta\theta$ rms)	> 1.5 mrad
α_t	0.375 mrad ($\Delta\theta$ rms)	> 1.5 mrad

Table A.4: A summary of levels of misalignment corresponding to a 10% degradation in any residual rms or, for biases shifts of 0.01 mrad for θ , 1 mrad for ϕ , and 0.25 mrad for $\Delta\theta$ for both the uncorrected and corrected cases; > 5 mm and > 1.5 mrad mean that such a degradation does not occur for the range of misalignment studied. Most affected quantity in parentheses.

Tod-Not-Brot

Old German Proverb

2659

B

2660

Telescoping Jets

2661 ANOTHER APPROACH TO IMPROVING $ZH \rightarrow \ell\ell b\bar{b}$ is the use of telescoping jets [63], which har-
2662 nesses the power of multiple event interpretations. The use of multiple event interpretations was
2663 originally developed with non-deterministic jet algorithms like the Q-jets (“quantum” jets) algo-
2664 rithm [64]. When a traditional or “classical” algorithm, such as the Cambridge-Aachen[65] and

2665 anti- k_t [66] algorithms, is applied to an event, it produces one set of jets for that event, i.e. a single
2666 interpretation of that event. With multiple event interpretations, each event is instead given an en-
2667 semble of interpretations. In the case of Q-jets, this ensemble is created through a non-deterministic
2668 clustering process for an anti- k_t jet algorithm. With telescoping jets, multiple jet cone radii (the char-
2669 acteristic size parameter, R) around a set of points in the pseudorapidity-azimuth ($\eta - \phi$) plane are
2670 used to generate a series of jet collections. Instead of an event passing or not-passing a given set of
2671 cuts, a fraction (called the cut-weight, z) of interpretations will pass these cuts. This cut-weight al-
2672 lows for enhanced background suppression and increased significance of observed quantities for a
2673 given data set, as detailed in Ref. [67]. The telescoping jets algorithm provides the benefits of mul-
2674 tiple event interpretations without the significant computational overhead of a non-deterministic
2675 algorithm like the Q-jets algorithm, and its multiple cone sizes are particularly suited to studying
2676 processes like associated production, which suffers from a pronounced low tail in the dijet invariant
2677 mass distribution due to final state radiation (FSR) “leaking” outside the relatively narrow jets used
2678 for object reconstruction.

2679 B.I MONTE CARLO SIMULATION

2680 The MC simulated samples used in this study are the same as in Ref. [?]. The signal sample used
2681 is generated in PYTHIA8 [68] with the CTEQ6L1 parton distributions functions (PDFs) and AU2
2682 tune[69, 70, 71] for the ZH process with $m_H = 125$ GeV (henceforth, $ZH125$). The primary back-
2683 ground processes examined in this study were Z +jets with massive b and c quarks. These samples
2684 are generated with version 1.4.1 of the SHERPA generator [72].

2685 B.2 JET RECONSTRUCTION AND CALIBRATION

2686 In order to construct telescoping jets, jet axes must first be found around which to “telescope.” In
2687 the reconstructed-level analysis, the anti- k_t algorithm with $R = 0.4$ is used to reconstruct jets from
2688 topological clusters in the calorimeters. The four vectors of these anti- k_t algorithm with $R = 0.4$ jets
2689 are calibrated to match truth information obtained from simulation and validated in data. To take
2690 into account the effect of pile-up interactions, jet energies are corrected using a jet-area based tech-
2691 nique [73], and each jet with $p_T < 50$ GeV and $|\eta| < 2.4$ is subject to a requirement that at least 50%
2692 of the scalar sum of the p_T of tracks matched to this jet be composed of tracks also associated with
2693 the primary vertex. Jet energies are also calibrated using p_T and η -dependent correction factors [74].
2694 Furthermore, at least two jets must have $|\eta| < 2.5$ in order to be b -tagged. The MV1 algorithm [75]
2695 [? ? ?] is used for b -tagging. Once jets are reconstructed and b -tag weights have been calculated, the
2696 two hardest, b -tagged jets are used as the telescoping jet axes. Additional details can be found in Ref.
2697 [76].

2698 After the telescoping jet axes have been established, telescoping jets are constructed using topo-
2699 logical clusters in the calorimeters at a variety of jet cone sizes. Including the original anti- k_t jets
2700 used for the $R = 0.4$ case, twelve total sets of jets of cone sizes ranging from $R = 0.4\text{--}1.5$ are
2701 constructed, with each successive size having a radius 0.1 larger than the preceding set. For each jet
2702 axis, telescoping jets consist of any topological cluster lying within R of the axis. In the event of over-
2703 lap, clusters are assigned to the closer jet axis. If a given cluster is equidistant from the two jet axes,
2704 the cluster is assigned to whichever jet axis is associated with the anti- k_t jet with higher p_T . Calibra-

2705 tion for the telescoping jets is conducted using corrections for anti- k_t calorimeter topological cluster
2706 jets; the $R = 0.4$ corrections are used for telescoping $R = 0.5$, and the $R = 0.6$ corrections are
2707 used for telescoping $R \geq 0.6$ (cf. Sec. B.4). The telescoping cone jets ($R \geq 0.5$) at reconstructed
2708 level are trimmed using Cambridge-Aachen jets with $R = 0.3$ and $f_{cut} = 0.05$ with respect to the
2709 untrimmed jet p_T [77]. Since these jets are trimmed, the active area correction is not applied. In the
2710 event a Z candidate electron falls within R of the axis of a telescoping jet, its 4-momentum is sub-
2711 tracted from that of the jet vectorially.

2712 A similar process is used to construct telescoping jets in the truth-level analysis below. Instead of
2713 the two hardest b -tagged anti- k_t with $R = 0.4$ jets reconstructed with calorimeter topological clus-
2714 ters, the two hardest truth b -jets in an event are used. Instead of making a cut on b -tagging weight
2715 to b -tag, truth jets are examined to see whether a b -hadron with $p_T > 5$ GeV is contained within
2716 $\Delta R < 0.4$ of the jet axis; the presence of a b -hadron is used to b -tag truth-level jets. These two jets
2717 again provide the jets for the $R = 0.4$ case and the axes around which telescoping takes place. Stable
2718 truth particles, not including muons and neutrinos, are used in place of calorimeter topological clus-
2719 ters. Z candidate electron-telescoping jet overlap removal is performed at truth level, too. Missing E_T
2720 is calculated using the vector sum of the four momenta of stable truth-level neutrinos. Since there
2721 are no pileup particles stored at truth level, truth-level telescoping jets are not trimmed.

2722 B.3 EVENT RECONSTRUCTION AND SELECTION

2723 Events are selected on the basis of a combination of leptonic, jet, and missing E_T requirements,
2724 which are outlined in Table B.1. Leptons are categorized by three sets of increasingly stringent qual-

2725 ity requirements, which include lower limits on E_T , upper limits on $|\eta|$, impact-parameter require-
2726 ments, and track-based isolation criteria. The requirements differ for electrons [78] and muons
2727 [79]. Events are selected with a combination of single lepton, dielectron, and dimuon requirements.
2728 Each event must contain at least one lepton passing medium requirements and at least one other
2729 lepton passing loose requirements. These leptons are used to create a dilepton invariant mass cut to
2730 ensure the presence of a Z boson and suppress multijet backgrounds.

2731 Event selection requirements are also imposed on the anti- k_t with $R = 0.4$ jets. There must be at
2732 least two b -tagged jets in a given event. The p_T of the harder b -tagged jet must be at least 45 GeV, and
2733 the second b -tagged jet must have p_T of at least 20 GeV. There are further topological cuts on the
2734 separation of the two jets $\Delta R(b, \bar{b})$, the distance between the two jets in the (η, ϕ) plane, according
2735 to the transverse momentum of the Z boson, p_T^Z . These are shown in Table B.2.

2736 The truth-level analysis has the same missing E_T , jet p_T , m_{ll} , and additional topological selection
2737 criteria, but the use of truth-level information simplifies the other requirements. Instead of lepton
2738 quality requirements, Z boson candidate leptons' statuses and MC record barcodes are checked to
2739 ensure the leptons are stable.

2740 In the jet calibration validation, the reconstructed level analysis lepton and m_{ll} requirements are
2741 imposed, but neither the missing E_T nor the jet selection requirements are applied so as not to bias
2742 the validation.

Requirement	Reconstructed	Truth	Validation
Leptons	1 med. + 1 loose	2 produced by Z boson	1 med. + 1 loose
b -jet	2 b -tags	2 b -jets	—
p_T jet 1 (jet 2)		$> 45 \text{ GeV} (> 20) \text{ GeV}$	—
Missing E_T		$E_T^{\text{miss}} < 60 \text{ GeV}$	—
Z boson		$83 < m_{ll} < 99 \text{ GeV}$	

Table B.1: A summary of basic event selection requirements. Truth-level b -tagging is done with truth-level information.

$p_T^Z [\text{GeV}]$	$\Delta R(b, b)$
0–90	0.7–3.4
90–120	0.7–3.0
120–160	0.7–2.3
160–200	0.7–1.8
> 200	< 1.4

Table B.2: Topological requirements of the event selection.

2743 B.4 VALIDATION OF JET CALIBRATION

2744 In order to validate the jet energy scale and resolution of jets constructed with this telescoping-jets
 2745 algorithm, values of $p_T^{\text{rec}}/p_T^{\text{tru}}$ are studied for each value of R for the Z +jets MC sample. In a given
 2746 event, all jets, not just the two hardest b -tagged jets, are telescoped. These jets are constructed in the
 2747 same way as in the reconstructed and truth-level analyses; reconstructed-level jets are made from
 2748 calorimeter topological clusters within R of the anti- k_t with $R = 0.4$ jet axes and then trimmed,
 2749 and truth-level jets are made from stable truth particles within R of the anti- k_t with $R = 0.4$ jet
 2750 axes. The reconstructed and truth-level telescoping jet ensembles are matched according to the sep-
 2751 aration in the (η, ϕ) plane of their corresponding anti- k_t with $R = 0.4$ jets used as seeds. Only jets

2752 with $|\eta| < 1.2$ are examined here, and the results of studies on the $ZH125$, ZZ , and $t\bar{t}$ samples, as
 2753 well as over other $|\eta|$ ranges, are outlined in [63]. Any reconstructed jets not within $\Delta R = 0.3$ of a
 2754 truth jet are discarded. In the event that multiple reconstructed jets are the same distance away from
 2755 a given truth jet, the reconstructed jet with the highest p_T gets matched. Matching is retained for all
 2756 R values (i.e. telescoping seeds are matched, and telescoping jets are assumed to match if the anti- k_t
 2757 jets from which their seeds are derived match).

2758 Once anti- k_t with $R = 0.4$ reconstructed and truth jets are matched, response functions are cre-
 2759 ated by generating a series of distributions of p_T^{rec}/p_T^{tru} in 20 GeV bins of p_T^{tru} from 20–200 GeV, one
 2760 bin for 200–300 GeV, and one bin for 300–500 GeV for each R , with bins chosen for purposes of
 2761 statistics. Ensembles with $p_T^{tru} < 20$ GeV are ignored since no calibration exists for jets with trans-
 2762 verse momentum below this value. The values of $\langle p_T^{rec}/p_T^{tru} \rangle$ in each p_T^{tru} bin are calculated by doing
 2763 a two sigma gaussian fit on the distribution of p_T^{rec}/p_T^{tru} in that bin and taking the mean of that fit,
 2764 and the error on the mean is taken from the error of this parameter in the fit. The resolutions are the
 2765 values of the square root of the variance on this fit. As the total response distributions in Figure B.1
 2766 show, performance is best for low R values and high values of p_T^{tru} . Figure B.1 shows the $R = 0.4$
 2767 (anti k_t) case to show a baseline for performance, $R = 0.6$ to show the deviations with “correct”
 2768 calibrations, and $R = 1.0, 1.5$ to show how big those deviations get with larger R jets. The resolu-
 2769 tions, $\sigma_{p_T^{rec}/p_T^{tru}}$, as a function of p_T^{tru} are shown in Figure B.1(b). For $p_T^{tru} > 60$ GeV, response is fairly
 2770 consistent over various R values. Resolution, as might naively be expected, is worse for increasingly
 2771 larger values of R . For $p_T^{tru} < 60$ GeV, resolution degrades, and response degrades in particular for
 2772 increasing R ; this is likely a result from residual pileup effects.

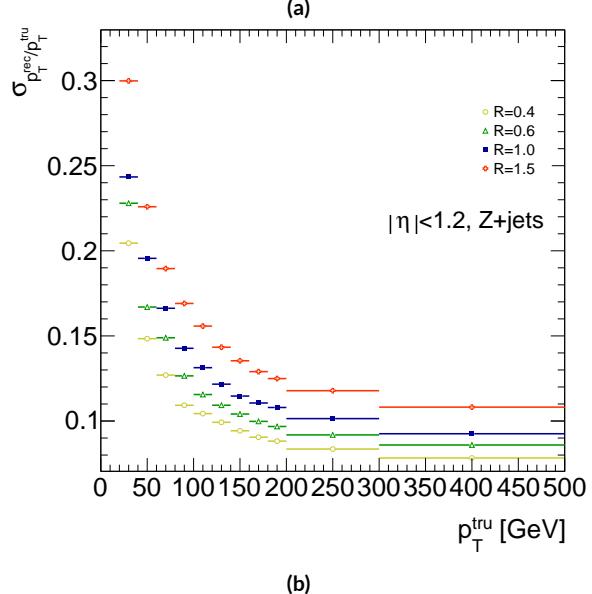
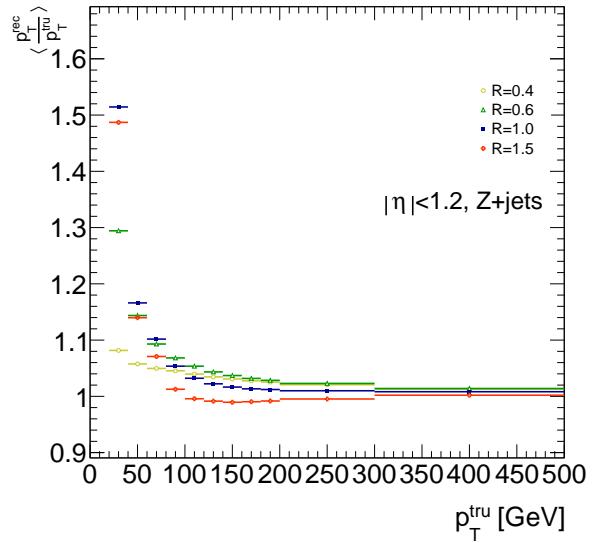


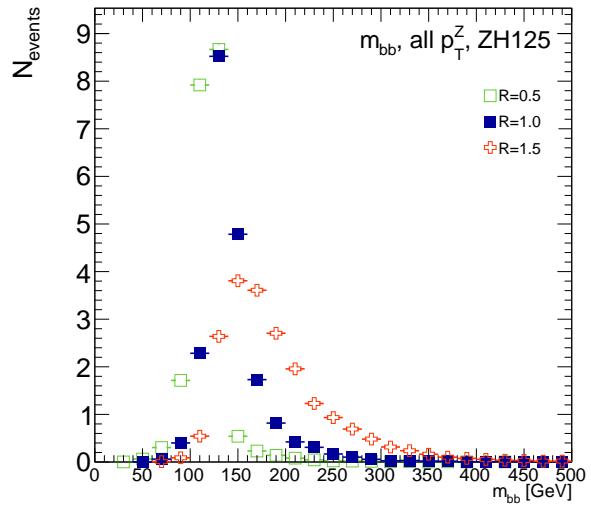
Figure B.1: The mean and resolution of p_T^{rec}/p_T^{tru} for the background $Z+jets$ sample for $|\eta| < 1.2$ and for $R = 0.4, 0.6, 1.0$, and 1.5 in 20 GeV bins of p_T^{tru} for $20-200 \text{ GeV}$, one bin for $200-300 \text{ GeV}$, and one bin for $300-500 \text{ GeV}$, with bins chosen for purposes of statistics.

2773 B.5 TRUTH-LEVEL ANALYSIS

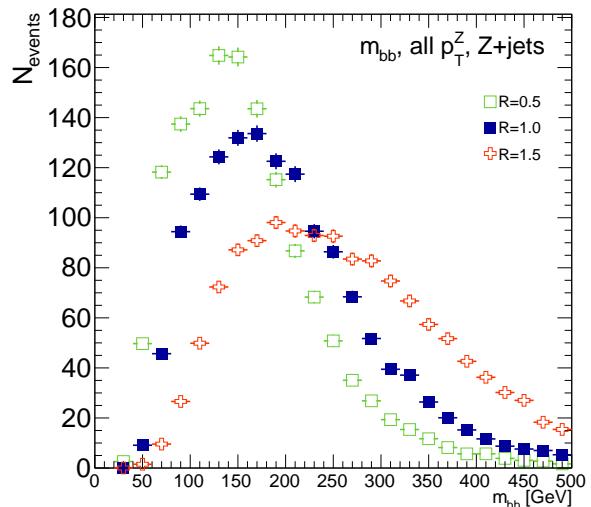
2774 To understand the limits and sources of any potential improvements, a truth-level analysis was con-
2775 ducted on MC samples with a ZH_{125} signal sample and a $Z+jets$ background sample. Distributions
2776 for the dijet invariant mass, m_{bb} , were made for each telescoping radius. Both signal and background
2777 samples develop more pronounced tails in the high m_{bb} region as R increases, as shown in Figure
2778 B.2. N_{events} is normalized to expected values in data.

2779 One way to take advantage of this information is to make a cut on m_{bb} for two different radii.
2780 This is graphically depicted in Figure B.3 for the optimized combination of $m_{bb,R=0.9}$ (telescoping
2781 cone jets constructed as outlined in Sec. B.2) vs. $m_{bb,R=0.4}$ (anti- k_t jets). At truth-level, the majority
2782 of events in the signal ZH_{125} sample are concentrated in relatively narrow region of parameter space,
2783 where this is certainly not the case for the more diffuse $Z+jets$ background sample.

2784 Another way to take advantage of multiple event interpretations is to make use of an event's cut-
2785 weight, denoted z and defined as the fraction of interpretations in a given event that pass a certain set
2786 of cuts (in this note, a cut on m_{bb}). The distribution of cut-weights for a sample of events is denoted
2787 $\rho(z)$. To enhance the significance of a cut-based analysis, events can be weighted by the cut-weight
2788 or any function $t(z)$ of the cut-weight. Weighting events by $t(z)$ modifies the usual $S/\delta B$ formula
2789 used to calculate significances. In this note, δB is based on Poissonian statistics and is taken as $0.5 +$
2790 $\sqrt{0.25 + N_B}$, where N_B is the number of background events.

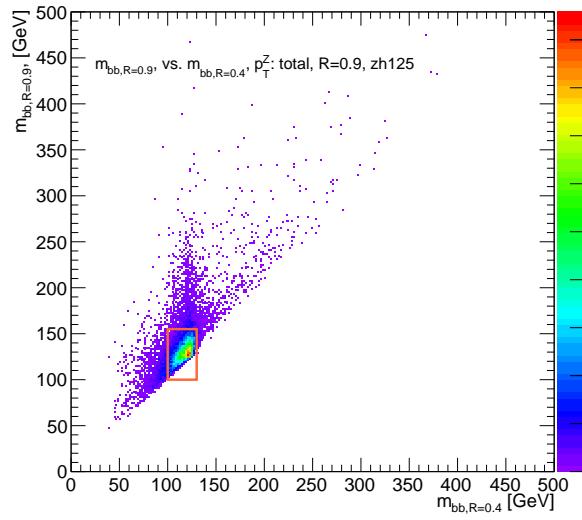


(a)

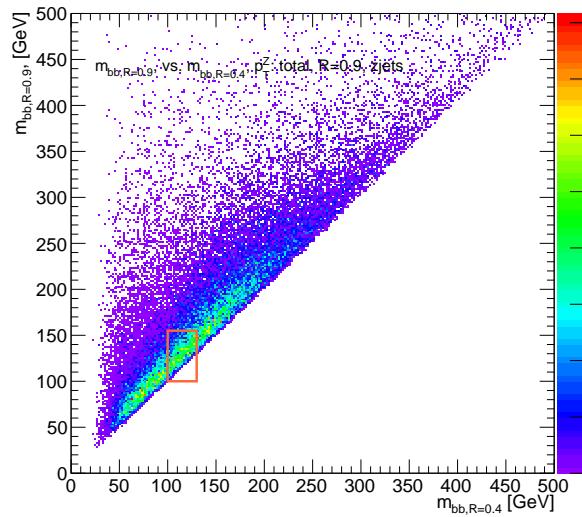


(b)

Figure B.2: The m_{bb} distribution for the telescoping jets with $R = 0.5, 1.0$, and 1.5 truth-level jets is shown for the signal and background samples in (a) and (b), respectively.



(a)



(b)

Figure B.3: The 2D distribution of $m_{bb,R=0.9}$ vs. $m_{bb,R=0.4}$ is shown for signal and background truth-level samples in (a) and (b), respectively. The region chosen for the double m_{bb} cut is outlined in orange.

2791 B.6 ERRORS ON TELESCOPING SIGNIFICANCES

2792 Significances of measurements are quoted in units of expected background fluctuations, schemati-
2793 cally, $S/\delta B$. For counting experiments with high numbers of events, we can use Gaussian statistics
2794 and express this as S/\sqrt{B} , which we here denote as \mathcal{S} . However, with lower statistics, it becomes
2795 more appropriate to use Poissonian statistics, and

$$\mathcal{S}_{gaus} = \frac{S}{\sqrt{B}} \rightarrow \mathcal{S}_{pois} = \frac{S}{0.5 + \sqrt{0.25 + B}} \quad (\text{B.1})$$

2796 where $0.5 + \sqrt{0.25 + B}$ is the characteristic upward fluctuation expected in a Poissonian data set
2797 using the Pearson chi-square test[80].

2798 B.7 COUNTING

2799 The significance is given as above, where $S = N_S$ and $B = N_B$. That is, the signal and background
2800 are just the number of events in signal and background that pass some cuts. The error for the Guas-
2801 sian case is the standard:

$$\Delta \mathcal{S}_{gaus} = \frac{1}{\sqrt{B}} \Delta S \oplus \frac{S}{2B^{3/2}} \Delta B \quad (\text{B.2})$$

2802 The error for the Poissonian case is:

$$\Delta \mathcal{S}_{pois} = \frac{1}{0.5 + \sqrt{0.25 + B}} \Delta S \oplus \frac{S}{2(0.5 + \sqrt{0.25 + B})^2 \sqrt{0.25 + B}} \Delta B \quad (\text{B.3})$$

2803 where \oplus denotes addition in quadrature, and $\Delta S(B)$ is the error on signal (background).

2804 B.8 MULTIPLE EVENT INTERPRETATIONS

2805 Using multiple event interpretations changes the formulae used in with simple counting. That is, S

2806 is not necessarily merely N_S , the number of events passing some signal cuts, and similarly for B and

2807 N_B . Using an event weighting by some function of the cut-weight, z , denoted $t(z)$, $S = N_S \langle t \rangle_{\rho_S}$

2808 and $B = N_B \langle t^2 \rangle_{\rho_B}$. So

$$\mathcal{S}_{t,gaus} = \frac{N_S \langle t \rangle_{\rho_S}}{\sqrt{N_B \langle t^2 \rangle_{\rho_B}}} \rightarrow \mathcal{S}_{t,pois} = \frac{N_S \int_0^1 dz t(z) \rho_S(z)}{0.5 + \sqrt{0.25 + N_B \int_0^1 dz t^2(z) \rho_B(z)}} \quad (\text{B.4})$$

For histograms, everything is done bin-wise. The notation used below is as follows: ρ_i is the value of $\rho(z)$ at bin i (where the bins run from 0 to n_{tel} , where n_{tel} is the total number of telescoping radii). $t_i = t_i(\rho_{S,i}, \rho_{B,i}, i/n_{tel})$ is the value of $t(z)$ at bin i , which can depend, in principle, on $\rho_{S,i}$, $\rho_{B,i}$, and i/n_{tel} (the last of which is z in bin i). Explicitly,

$$N = \sum_{i=0}^{n_{tel}} \rho_i, \quad \int_0^1 dz t(z) \rho_S(z) = \sum_{i=0}^{n_{tel}} t_i \rho_{S,i}, \quad \int_0^1 dz t^2(z) \rho_B(z) = \sum_{i=0}^{n_{tel}} t_i^2 \rho_{B,i}$$

2809 For the calculations that follow, let $\xi = \sum_{i=0}^{n_{tel}} t_i \rho_{S,i}$, $\psi = 0.5 + \sqrt{0.25 + N_B \sum_{i=0}^{n_{tel}} t_i^2 \rho_{B,i}}$,

2810 $\partial_S = \frac{\partial}{\partial \rho_{S,i}}$ (and similarly for B), so $\mathcal{S}_t = N_S \xi / \psi$

²⁸¹¹ Some partial derivatives:

$$\begin{aligned}
 \partial_S N_S &= 1, & \partial_{B,i} N_B &= 1 \\
 \partial_S \xi &= t_i + (\partial_S t_i) \rho_{S,i}, & \partial_B \xi &= (\partial_B t_i) \rho_{B,i} \\
 \partial_S \psi &= \frac{N_B t_i (\partial_S t_i) \rho_{B,i}}{\sqrt{0.25 + N_B \sum_{i=0}^{n_{tel}} t_i^2 \rho_{B,i}}}, & \partial_B \psi &= \frac{\sum_{i=0}^{n_{tel}} t_i^2 \rho_{B,i} + N_B (t_i^2 + 2t_i (\partial_B t_i) \rho_{B,i})}{2\sqrt{0.25 + N_B \sum_{i=0}^{n_{tel}} t_i^2 \rho_{B,i}}} \\
 \partial_S \mathcal{S}_t &= \frac{\xi}{\psi} + \frac{N_S}{\psi} \partial_S \xi - \frac{N_S \xi}{\psi^2} \partial_S \psi, & \partial_B \mathcal{S}_t &= N_S \left(\frac{1}{\psi} \partial_B \xi - \frac{\xi}{\psi^2} \partial_B \psi \right)
 \end{aligned}$$

²⁸¹² Thus,

$$\Delta \mathcal{S}_{t,i} = \left[\frac{\xi}{\psi} + \frac{N_S}{\psi} \partial_S \xi - \frac{N_S \xi}{\psi^2} \partial_S \psi \right] \Delta \rho_{S,i} \oplus N_S \left[\frac{1}{\psi} \partial_B \xi - \frac{\xi}{\psi^2} \partial_B \psi \right] \Delta \rho_{B,i} \quad (\text{B.5})$$

²⁸¹³ and the total error is given by the sum in quadrature over all bins i of $\Delta \mathcal{S}_{t,i}$.

²⁸¹⁴ B.9 $t(z) = z$

²⁸¹⁵ With $t(z) = z$, $t_i = i/n_{tel}$, so $\partial_S t_i = \partial_B t_i = 0$. So:

$$\begin{aligned}
 \partial_S \psi &= \partial_B \xi = 0 \\
 \partial_S \xi &= \frac{i}{n_{tel}} \\
 \partial_B \psi &= \frac{\sum_i i^2 \rho_{B,i} + N_B t^2}{n_{tel} \sqrt{n_{tel}^2 + N_B \sum_i i^2 \rho_{B,i}}}
 \end{aligned}$$

²⁸¹⁶ so $\Delta\mathcal{S}_{z,i}$ reduces to

$$\Delta\mathcal{S}_{t,i} = \left[\frac{\xi + N_S t_i}{\psi} \right] \Delta\rho_{S,i} \oplus \left[\frac{N_S \xi}{\psi^2} \partial_B \psi \right] \Delta\rho_{B,i} \quad (\text{B.6})$$

²⁸¹⁷ B.10 $t(z) = \rho_S(z) / \rho_B(z)$

²⁸¹⁸ With the likelihood optimized* $t^*(z) = \rho_S(z) / \rho_B(z)$, $t_i = \rho_{S,i} / \rho_{B,i}$, so $\partial_S t_i = 1 / \rho_{B,i}$ and $\partial_B t_i =$

²⁸¹⁹ $-\rho_{S,i} / \rho_{B,i}^2$. So:

$$\begin{aligned} \partial_S \xi &= 2 \frac{\rho_{S,i}}{\rho_{B,i}} = 2t_i \\ \partial_B \xi &= -\frac{\rho_{S,i}}{\rho_{B,i}} = -t_i \\ \partial_S \psi &= \frac{N_B t_i}{\sqrt{0.25 + N_B \sum_i \rho_{S,i}^2 / \rho_{B,i}}} \\ \partial_B \psi &= \frac{\sum_i \rho_{S,i}^2 / \rho_{B,i} - N_B (\rho_{S,i} / \rho_{B,i})^2}{\sqrt{1 + 4N_B \sum_i \rho_{S,i}^2 / \rho_{B,i}}} \end{aligned}$$

²⁸²⁰ simplifying somewhat the terms in the per bin error in Equation B.6.

²⁸²¹ The new significance figure using multiple event interpretations becomes, with ρ_S and ρ_B denot-
²⁸²² ing the cut-weight distributions in signal and background, respectively

$$\frac{S}{\delta B} = \frac{N_S \langle t \rangle_{\rho_S}}{0.5 + \sqrt{0.25 + N_B \langle t^2 \rangle_{\rho_B}}} \quad (\text{B.7})$$

*for the Gaussian statistics case

2823 Of particular interest is the likelihood optimized $t(z)$,[†] $t^*(z) = \rho_S(z)/\rho_B(z)$. m_{bb} windows are
 2824 chosen separately for each scheme studied to maximize total significances and are summarized in
 2825 Table B.3.

$$\left(\frac{S}{\delta B}\right)_z = \frac{N_S \epsilon_S}{0.5 + \sqrt{0.25 + N_B (\epsilon_B^2 + \sigma_B^2)}} \quad (B.8)$$

2826

$$\left(\frac{S}{\delta B}\right)_{t^*(z)} = \frac{N_S \int_0^1 dz \frac{\rho_S^2(z)}{\rho_B(z)}}{0.5 + \sqrt{0.25 + N_B \int_0^1 dz \frac{\rho_S^2(z)}{\rho_B(z)}}} \quad (B.9)$$

2827 where $\epsilon_{S,B}$ are the means of $\rho_{S,B}(z)$ and σ_B^2 is the variance of $\rho_B(z)$. Further details can be found in
 2828 Refs. [63, 67] and Appendix B.6.

Analysis Type	$S/\delta B$ Type	Optimal m_{bb} Window
Reconstructed	$anti-k_t R = 0.4$ $t(z) = z$ $t(z) = \rho_S(z)/\rho_B(z)$ $anti-k_t R = 0.4$, telescoping $R = 0.6$	$90-140$ GeV $110-155$ GeV $110-155$ GeV $95-140$ GeV ($R = 0.4$), $105-160$ GeV ($R = 0.6$)
Truth	$anti-k_t R = 0.4$ $t(z) = z$ $t(z) = \rho_S(z)/\rho_B(z)$ $anti-k_t R = 0.4$, telescoping $R = 0.9$	$100-130$ GeV $115-140$ GeV $120-135$ GeV $100-130$ GeV ($R = 0.4$), $100-155$ GeV ($R = 0.9$)

Table B.3: m_{bb} windows studied. These windows were chosen to optimize significances over all p_T^Z .

2829 The truth-level distributions $\rho_S(z)$, $\rho_B(z)$, and $\rho_S(z)/\rho_B(z)$ are shown for the m_{bb} window
 2830 that optimizes $(S/\delta B)_{t^*(z)}$ in Figure B.4, and significance improvements as a function of p_T^Z are
 2831 summarized in Figure B.5. Uncertainties in Figures B.5 and B.9 are statistical uncertainties. JES sys-

[†]Derived under the assumption of Gaussian statistics in Ref [67]

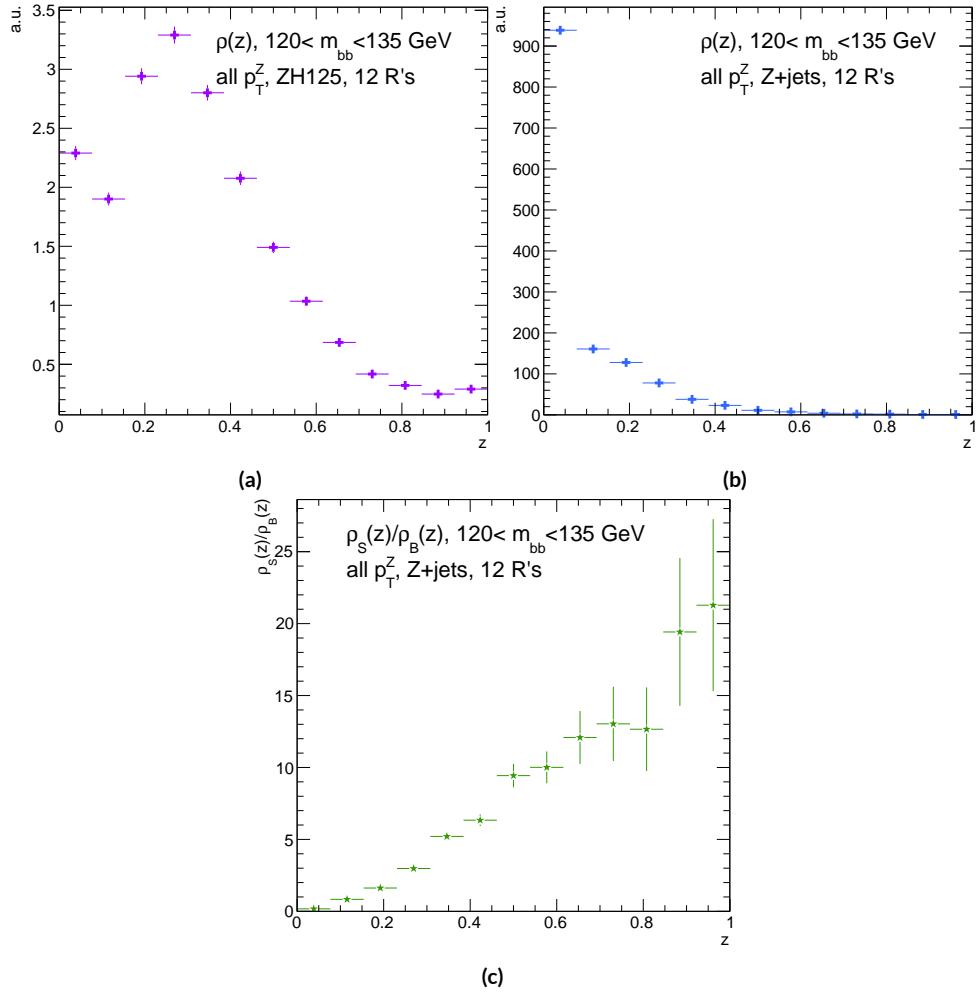


Figure B.4: Truth-level $\rho(z)$ distributions for the m_{bb} window optimizing $(S/\delta B)_{r^*(z)}$. $\rho_S(z)$ for the signal ZH125 sample is shown in (a), and $\rho_B(z)$ for the background Z+jets sample is shown in (b). The distribution of $\rho_S(z)/\rho_B(z)$ for these samples is shown in (c).

tematics will need to be evaluated for different R 's, as modeling uncertainties is an outstanding issue, but these systematics will likely be strongly correlated for the different R 's and are not anticipated to be a very large contribution to total uncertainties. While the two dimensional m_{bb} cut and $t(z) = z$ schemes only showed marginal improvement at truth level at 2.87%[‡] and 1.45%, respectively, the likelihood optimized $t^*(z)$ showed a more substantial 40.7% improvement overall, with a steady increase in improvement with increasing p_T^Z . Figure B.5 (d) summarizes the improvements with respect to p_T^Z for the $t^*(z)$ event weight for five, seven, and twelve telescoping radii (interpretations) per event. Improvements increase with a greater number of interpretations and are more pronounced at higher p_T^Z for this scheme. The optimal $120 < m_{bb} < 135$ GeV window for $t^*(z)$ case is among the smallest studied. The benefits of this window's narrowness are suggested in Figure B.4. While the background cut-weight distribution, $\rho_B(z)$ in Figure B.4 (b) behaves as one might with a marked peak at $z = 0$, the signal $\rho_S(z)$ distribution peaks at a relatively modest $z = 0.3$, which indicates that much of the gain at truth level comes from background rejection. This is possible at truth level since there is both truth-level information available and no smearing and since ρ_S/ρ_B is the relevant quantity (as shown in Figure B.4 (c)).

[‡]The limited improvement is provably due to the simplified treatment of the 2D case; better performance with a more sophisticated treatment has been observed in Ref. [81].

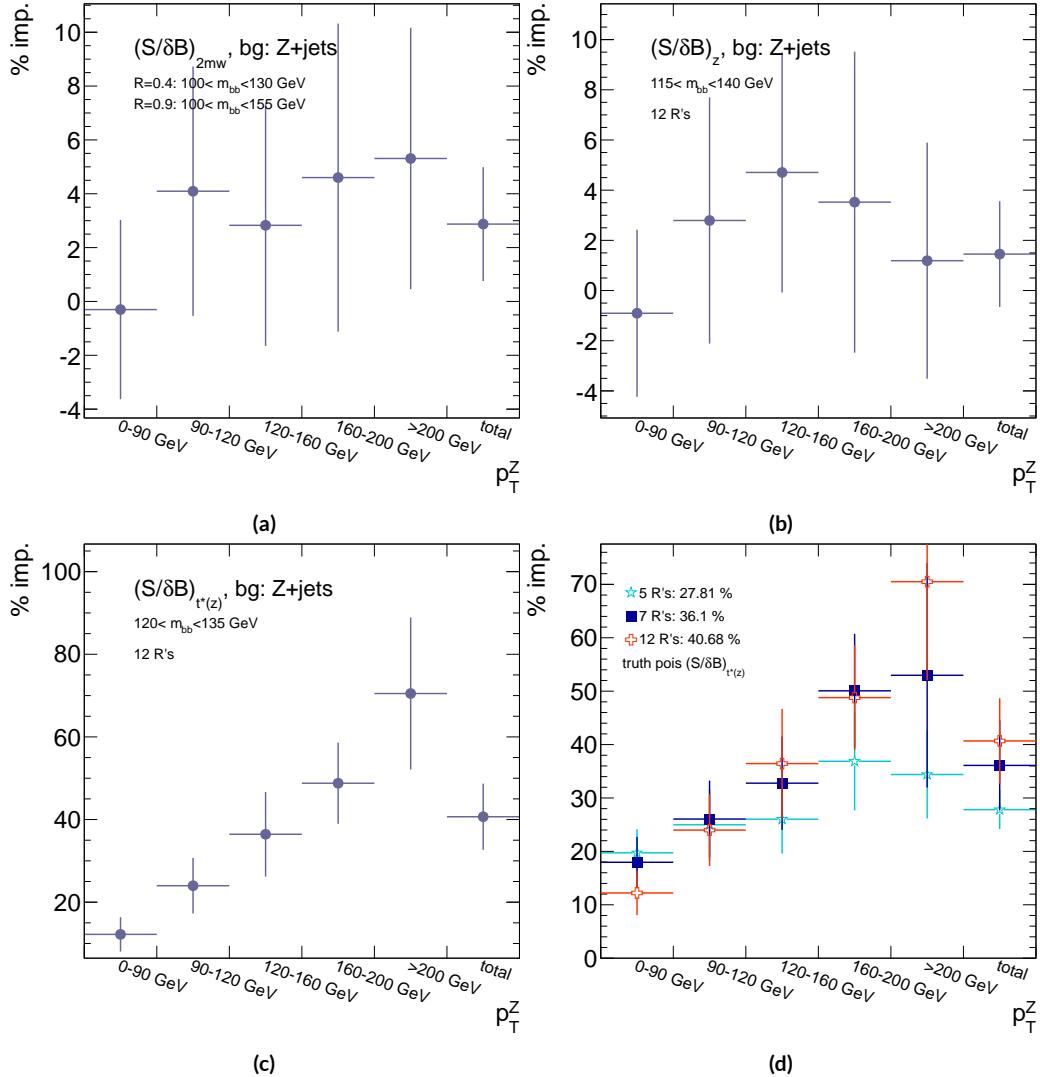


Figure B.5: A summary of the improvements for different truth-level telescoping jet cuts and weights is shown in bins of p_T^Z . The final bin is the total improvement over all p_T^Z . Shown are improvements for the 2D m_{bb} cut (a), $t(z) = z$ (b), $t(z) = t^*(z)$ with 12 radii (c), and $t(z) = t^*(z)$ for various radii (d).

2847 B.II RECONSTRUCTED-LEVEL ANALYSIS

2848 At reconstructed level, the same overall effect of introducing a high tail in m_{bb} distributions with
2849 increasing R is evident in comparing Figures B.2 and B.6. The optimal m_{bb} windows, however, grow
2850 larger, due to the lack of truth-level information.

2851 Total significance gains at reconstructed level for the two dimensional m_{bb} cut and the $t(z) = z$
2852 case are similar, at 2.87% and 1.45%, respectively. The optimal two-dimensional m_{bb} cut at recon-
2853 structed level is $95 < m_{bb,R=0.4} < 140 \text{ GeV}$, $105 < m_{bb,R=0.6} < 160 \text{ GeV}$. Just as at truth level,
2854 the $R = 0.4$ m_{bb} cut is comparable to the optimal single $R = 0.4$ m_{bb} cut, and the second m_{bb} cut is
2855 at similar values (cf. Table B.3 and Figures B.3 and B.7). However, the optimal second telescoping ra-
2856 dius is markedly smaller at $R = 0.6$ versus the optimal truth-level second radius of $R = 0.9$, which
2857 suggests that effects like pileup at reconstructed level obscure correlations between the $R = 0.4$
2858 interpretations and limit the usefulness of larger R interpretations in this particular scheme. The
2859 $t(z) = z$ case has a wider optimal window and yields about half the improvement it does at truth
2860 level.

2861 The optimal m_{bb} window for the $t^*(z)$ case is also markedly wider at reconstructed level, at $110 <$
2862 $m_{bb} < 155 \text{ GeV}$ in comparison to the truth-level optimal $120 < m_{bb} < 135 \text{ GeV}$. The $\rho(z)$ dis-
2863 tributions for the signal $ZH125$ and background $Z+\text{jets}$ as well as the $\rho_S(z) / \rho_B(z)$ in this window
2864 are shown in Figure B.8. Compared with the truth-level distributions in Figure B.4, both the sig-
2865 nal and background optimal $\rho(z)$ distributions have higher values at higher z . The peak in $\rho_S(z)$ at
2866 $z = 1$ suggests that at reconstructed level, maximizing the number of more “signal-like” events is

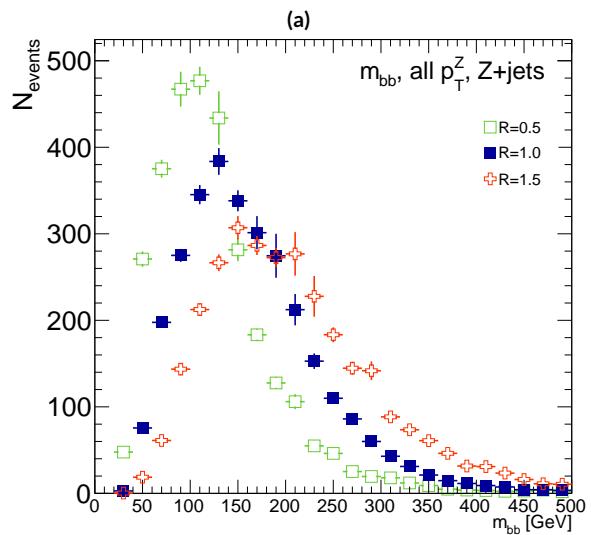
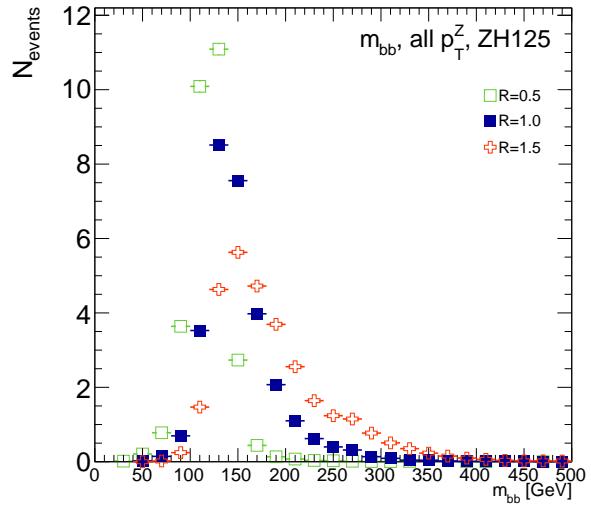


Figure B.6: The m_{bb} distribution for the telescoping jets with $R = 0.5$, $R = 1.0$, and $R = 1.5$ reconstructed-level jets is shown for the signal and background samples in (a) and (b), respectively.

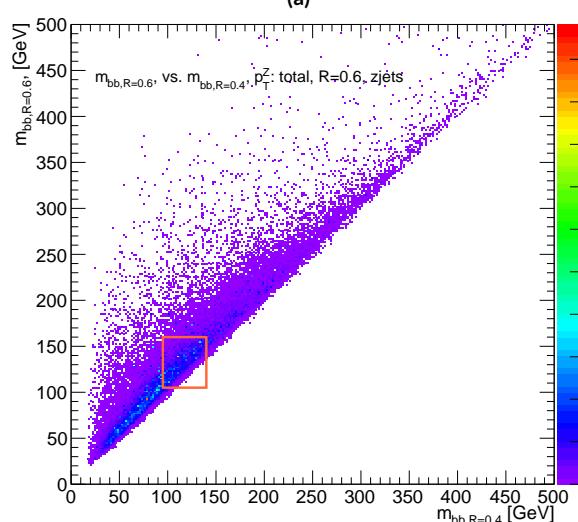
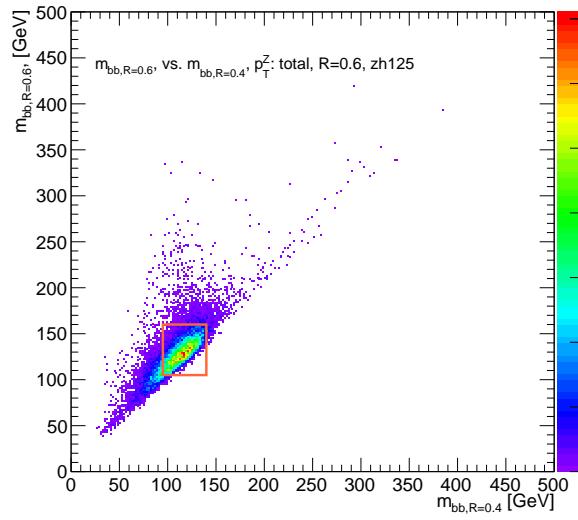


Figure B.7: The 2D distribution of $m_{bb,R=0.8}$ vs. $m_{bb,R=0.4}$ is shown for signal and background reconstructed-level samples in (a) and (b), respectively. The region chosen for the double m_{bb} cut is outlined in orange.

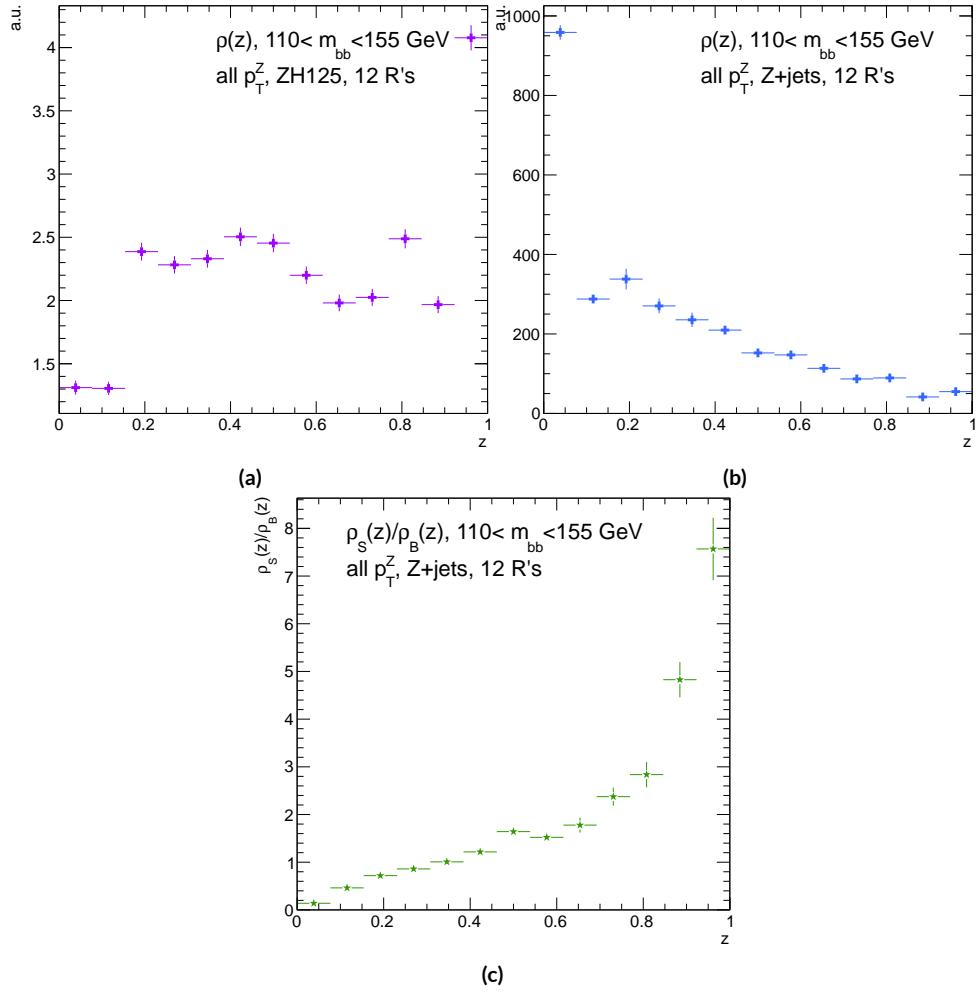


Figure B.8: Reconstructed-level $\rho(z)$ distributions for the m_{bb} window optimizing $(S/\delta B)_{t^*(z)} \cdot \rho_S(z)$ for the signal $ZH125$ sample is shown in (a), and $\rho_B(z)$ for the background $Z+jets$ sample is shown in (b). The distribution of $\rho_S(z)/\rho_B(z)$ for these samples is shown in (c).

2867 the key to optimizing significances, as opposed to the optimal, background suppressing $\rho(z)$ distri-
 2868 butions at truth level. The use of a greater number of interpretations per event (telescoping radii)
 2869 does appear to result in overall greater improvement as at truth level, as twelve radii performed bet-
 2870 ter than five, but this is less clear at reconstructed level, as shown in Figure B.9 (d). The improve-
 2871 ment at reconstructed level using an event weight of $t^*(z)$ is 20.5%, just over half the improvement
 2872 at truth level but still quite significant. Summaries of improvements as a function of p_T^Z for all three
 2873 cases studied and for the $t^*(z)$ case for different numbers of telescoping radii are shown in Figure
 2874 B.9.

Type	0–90	90–120	120–160	160–200	> 200	total
akt4 _{rec}	0.47492	0.28214	0.28339	0.25748	0.37337	0.76887
akt4 _{tru}	0.57414	0.30655	0.37309	0.35042	0.53569	0.98619
$2 m_{bb,rec}$	0.48903	0.2858	0.28812	0.25972	0.38297	0.78611
$2 m_{bb,tru}$	0.5724	0.3191	0.38364	0.36655	0.56414	1.0145
z_{rec}	0.50698	0.27962	0.29937	0.25688	0.36846	0.79158
z_{tru}	0.56894	0.31511	0.39065	0.36277	0.54206	1.0005
$t^*(z)_{rec}$	0.55085	0.29931	0.33367	0.30107	0.51321	0.92649
$t^*(z)_{tru}$	0.64425	0.38008	0.50904	0.5214	0.91337	1.3873

Table B.4: A summary of significances for different weighting schemes and cuts and for reconstructed and truth jets for a luminosity of 20.3 fb^{-1} . akt4 refers to the standard cut-based analysis using anti- k_t with $R = 0.4$ jets. Column titles are p_T^Z regions in GeV.

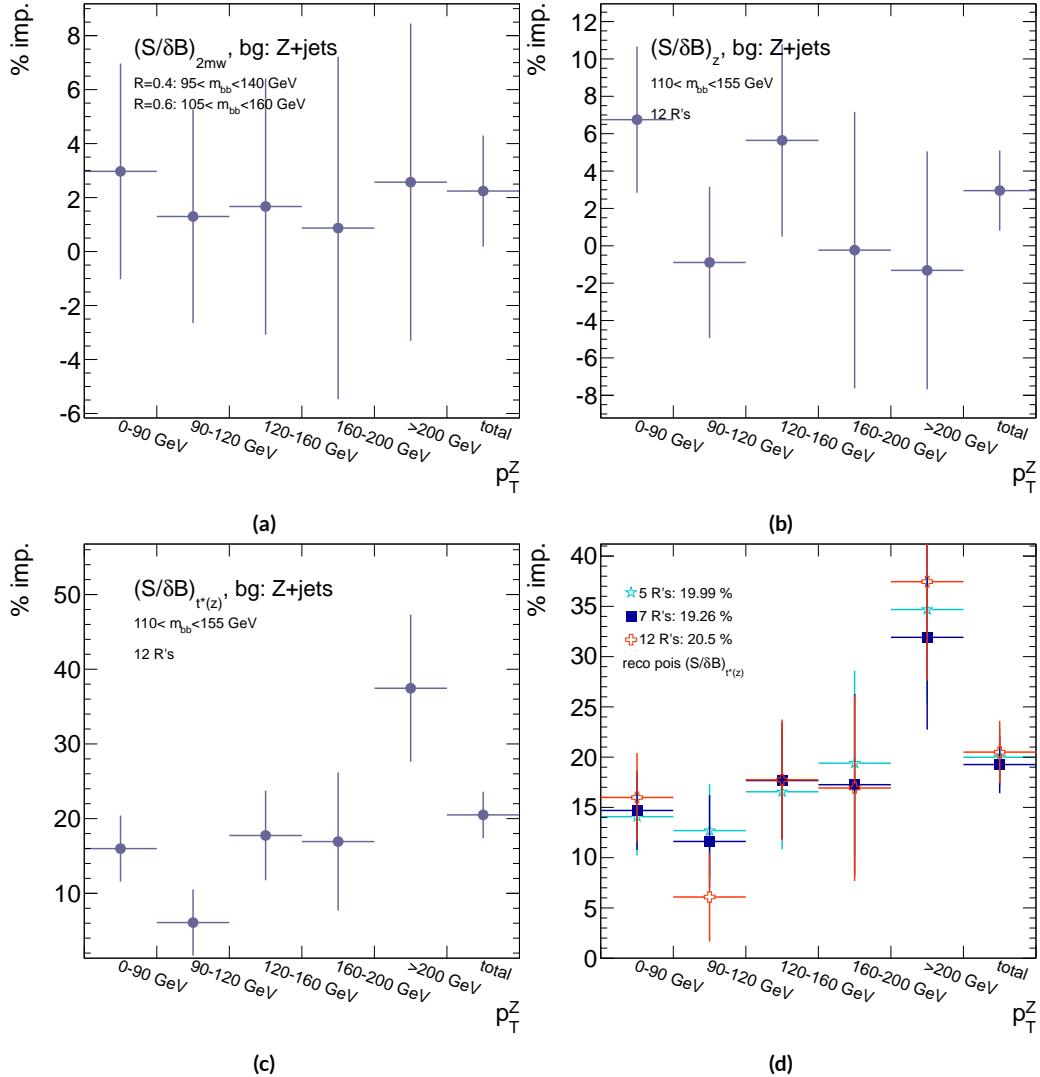


Figure B.9: A summary of the improvements for different reconstructed-level telescoping jet cuts and weights is shown in bins of p_T^Z . The final bin is the total improvement over all p_T^Z . Shown are improvements for the 2D m_{bb} cut (a), $t(z) = z$ (b), $t(z) = t^*(z)$ with 12 radii (c), and $t(z) = t^*(z)$ for various radii (d).

2875 **B.12 CONCLUSIONS AND PROSPECTS**

2876 The use of telescoping jets to provide multiple event interpretations shows promise as an avenue to
2877 increase significances in the $H \rightarrow b\bar{b}$ search in ATLAS and make an observation in the systematics-
2878 limited environment of early Run 2. A preliminary study using the telescoping jets algorithm with
2879 12 telescoping radii to build 12 event interpretations on 2012 Monte Carlo based on the full, cut-
2880 based Run 1 analysis yielded a 20.5% improvement in $S/\delta B$ over using anti- k_t with $R = 0.4$ alone
2881 at reconstructed level using a likelihood maximized event weighting to study the $ZH \rightarrow llb\bar{b}$ pro-
2882 cess. The jets used in this note at reconstructed level were trimmed in order to guarantee reasonable
2883 resolution in the large- R interpretations. The algorithm, in particular, showed discriminating power
2884 at high p_T^Z , so better performance can be expected in Run 2 with a higher \sqrt{s} and higher numbers of
2885 events with large p_T^Z . Additionally, the many simplifying assumptions regarding jet calibration and
2886 the relatively basic use of information[§] from multiple invariant masses in this note suggest that even
2887 further improvements than those quoted are possible. While this note did not explore the correla-
2888 tions between multiple event interpretations and the variables used in the BDT of the latest multi-
2889 variate version of the $H \rightarrow b\bar{b}$ analysis[20], new phenomenological studies suggest that such corre-
2890 lations are not strong[81]. The corresponding reconstructed-level study, using a BDT, is left for future
2891 work. Also left for future work are better understanding the effects of jet trimming and which inter-
2892 pretations are the most useful.

[§]For examples of more sophisticated treatments compared to the treatment in this note, see Ref [81].

Ah, peut on être heureux?

Quand on forme des autres voeux?

J. P. Rameau, *Les Indes Galantes*

2893

C

2894

Historical Aside

2895 INTERNATIONAL COLLABORATIONS with thousands of scientists like those at CERN's (the Euro-
2896 pean Organisation for Nuclear Research) 27 km circumference Large Hadron Collider (LHC) are
2897 fast becoming the norm in many fields of science, making the past seven decades of discovery in par-
2898 ticle physics seem a natural part of history's long march of progress. Seemingly arcane terms like dark

299

2899 matter and the Higgs boson (the infamous “God particle”) even pop up in blockbuster movies and
2900 primetime television. All of this, however, would have been impossible without the fascist and then
2901 communist existential threats to the American world order throughout the 20th century.

2902 The Manhattan Project and its atomic arsenal were a direct response to the threat of global fas-
2903 cism in World War II. They both secured the United States’ position as a world power at the end of
2904 the war and laid the foundations for many particle physics developments for the following three
2905 decades. High ranking American officials were well aware that this would have been impossible
2906 without the contributions of particle physicists. Some of these physicists, like Enrico Fermi and
2907 Arthur Compton, were already Nobel laureates and luminaries in the field. Others, like Richard
2908 Feynman and Owen Chamberlain, would go on to make their marks in the decades following the
2909 war. Though a few of these physicists, most notably Edward Teller, would continue their work on
2910 nuclear weapons, most of these physicists would return to basic science research as the nation turned
2911 towards the uneasy peace time of the Cold War.

2912 Particle physicists’ service and connections made during the war would serve them well in the
2913 decades to come as the military-controlled Manhattan Project transitioned to the civilian-led Atomic
2914 Energy Commission (AEC). The AEC was founded in 1946 to oversee the nuclear arsenal, the devel-
2915 opment of atomic power, and related fundamental research in the United States. Many of those on
2916 AEC board were former Manhattan Project particle physicists, including Glenn Seaborg, the AEC
2917 chairman from 1961-1971. Congressional oversight for AEC funding consisted of a single committee,
2918 the Joint Committee on Atomic Energy (JCAE), whose deliberations often took place behind closed
2919 doors owing to the AEC’s sensitive national security mission. Elementary particle physics research

²⁹²⁰ was clearly central to the AEC mission at its founding, as nuclear fission was the bleeding edge of particle physics at the beginning of World War II and represented the culmination of decades of highly specialized research that had no immediately obvious practical application. Furthermore, particle accelerator technology, the main workhorse then as now for basic science research in particle physics and the most expensive item on any particle physicist's wish list, had been crucial to many of these discoveries. The anticipation of future windfalls as momentous as the power of the atom and the exemplary performance of particle physicists during the war ensured that experimental particle physics and particle accelerators would remain the crown jewel of AEC research throughout the organization's existence.

²⁹²⁹ The AEC's sizable budget (thanks to its crucial mission of securing the nation's nuclear arsenal) and lavish support were the biggest contributing factors to the development of particle physics in the mid 20th century through its funding of accelerator facilities. Particle accelerators use powerful electromagnetic fields to take beams of subatomic particles, usually protons or electrons, as close to the speed of light as possible before colliding them into either fixed targets or other beams to produce high energy collisions. Physicists use these collisions to test models of the universe that predict behavior in these extreme regimes. Without more energetic collisions, progress becomes functionally impossible. While the first such accelerator was smaller than the average human hand, studying more complete models of the universe called for more energetic collisions and hence bigger, more powerful, and more expensive accelerators.

²⁹³⁹ Soon, these experiments became too big and expensive for individual universities to operate on their own. Progress in American particle physics became entirely dependent on the AEC, and hence

2941 on the continued threat of nuclear annihilation. National laboratories, all under AEC stewardship,
2942 became regional centers of research for particle physicists. By the late 1960's, Brookhaven National
2943 Laboratory, Lawrence Berkeley National Laboratory, and the Stanford Linear Accelerator Center
2944 hosted the majority of cutting edge accelerator facilities in the country alongside a dwindling num-
2945 ber of single university accelerators. By the decade's end Cornell hosted the only such university op-
2946 erated facility. The competitive rivalry among these different institutions fostered American success
2947 and dominance in experimental particle physics through the 1970's. The culmination of AEC pa-
2948 tronage was the National Accelerator Laboratory (now Fermilab), which began operations in 1967.
2949 Fermilab's construction was not a foregone conclusion given the economically challenging backdrop
2950 of the Vietnam War, but an emphasis on cost effective plans for both the laboratory and accelerator
2951 backed by the full support of the AEC secured Fermilab's funding. Fermilab would ultimately be-
2952 come home to the Tevatron, the final particle accelerator in the United States to claim the title of the
2953 world's most powerful.

2954 Particle physics only became more dependent on the existence of a Soviet threat with the end of
2955 the AEC. Due to budgetary pressures, the AEC was abolished in 1975, and its duties were eventually
2956 reorganized into the Department of Energy (DOE). Under DOE administration, proposed parti-
2957 cle physics experiments now had to compete against research projects from the entire range of fields
2958 germane to American energy instead of only other nuclear and particle physics projects. Moreover,
2959 DOE leadership had far fewer officials with track records of supporting particle physics research
2960 projects above all others. Nevertheless, there remained one last, great effort to promote collider
2961 physics in the United States, the Superconducting Supercollider (SSC). The SSC was an incredibly

ambitious design: a 50 mile ring under the Waxahachie desert with superconducting magnets to accelerate protons and antiprotons to energies more than three times higher than the LHC's current world record. The project was conceived during the Reagan administration and billed as a megaproject to reassert American dominance as the president took a more aggressive approach to the Soviet threat. Unfortunately, the project was perhaps too ambitious and suffered from management problems. It is not surprising, then, that the end of the Cold War spelled the end of the SSC. With no external threat to American global dominance, there was little impetus to continue funding such an expensive and over-budget project. There has not been a single initiative since for the United States to recapture its once commanding lead over efforts in Western Europe.

Western Europe was the only other serious center of 20th century particle physics, and successes there also depended upon five decades of existential threat to the United States, though in a less direct fashion. Most obviously, American institutions and physicists have been essential to the development of European particle physics, just as European physicists were crucial to the success of the Manhattan Project. Seven of the ten Cold War era CERN Directors General were either educated or did research at American universities, and every major particle physics discovery since the end of World War II has relied on both American and European talent and infrastructure. Furthermore, the European approach to experimental particle physics, epitomized by CERN, emphasized consensus and cooperation and was emblematic of larger geopolitical currents on the European continent in the latter half of the 20th century. Such a culture and its success would have been impossible without the same threats that facilitated American success in particle physics. While limited resources of member states were no doubt contributing factors in CERN's genesis, the collaborative culture of

2983 CERN and other pan-European organizations was a reaction to centuries of competition for continental dominance. After the total destruction of the world wars, enough was enough. The relatively
2984 peaceful prosperity on the Western side of the Iron Curtain made European cooperation possible,
2985 while the threat at Western Europe's doorstep only heightened the urgency of pan-European desires.
2986 Hence, the symbolic importance of European unity during the Cold War is hard to underestimate, and
2987 CERN-facilitated European cooperation made it a forerunner to organizations like the European
2988 Union and a model to the world. Every major achievement in particle physics after 1940 therefore
2989 relies on facilities and institutions on both sides of the Atlantic that would never have been formed
2990 without the back to back threats of global fascism and Soviet Communism.

2992 With the cancellation of the SSC in 1993 and the closing of Fermilab's Tevatron in 2011, CERN
2993 and its LHC remain the lone laboratory and experiment at the energy frontier. The United States
2994 is now a mere "observer state" at CERN: American talent and funding are essential to CERN and
2995 its mission, but the United States does not have a seat on CERN's governing council. It remains
2996 to be seen whether a legacy of over six decades of international cooperation will provide sufficient
2997 motivation for particle physics to continue at CERN after the LHC without guarantee of any dis-
2998 covery at the next experiment. Current nuclear threats, while attention grabbing, are far from exis-
2999 tential and unlikely to reignite any initiative for distinctly American science megaprojects. The only
3000 other prospect for a future collider at the energy frontier is China, whose nationalistic desire for su-
3001 perpower status may prove a sufficiently powerful and lasting motivator for the next generation of
3002 collider.

References

- [1] Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- [2] Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.
- [3] ATLAS Collaboration. Evidence for the $b \rightarrow b\bar{b}$ decay with the atlas detector, 2017.
- [4] Tatsuya Masubuchi, Jose Benitez, Andrew Stuart Bell, Spyridon Argyropoulos, Hannah Arnold, Yara Amaral Coutinho, Arturo Rodolfo Sanchez Pineda, Adrian Buzatu, Giovanni Calderini, and et al. Chan, Stephen Kam-wah. Search for a Standard Model Higgs boson produced in association with a vector boson and decaying to a pair of b -quarks. Technical Report ATL-COM-PHYS-2016-1724, CERN, Geneva, Nov 2016. URL <https://cds.cern.ch/record/2235437>.
- [5] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016. doi: 10.1088/1674-1137/40/10/100001.
- [6] Symmetry Magazine. The Standard Model of Particle Physics, 2015. URL <http://www.symmetrymagazine.org/standard-model/>.
- [7] Stephan Hagebock. Lorentz Invariant Observables for Measurements of Hbb Decays with ATLAS, CERN, Geneva, 2017. URL https://indico.cern.ch/event/611919/contributions/2471459/attachments/1425444/2205142/17.03.28_DPG.pdf.
- [8] Paul Jackson and Christopher Rogan. Recursive jigsaw reconstruction: Hep event analysis in the presence of kinematic and combinatoric ambiguities. 2017.
- [9] The Proton Synchrotron Booster. Jul 2012. URL <http://cds.cern.ch/record/1997372>.
- [10] The Proton Synchrotron. Jan 2012. URL <http://cds.cern.ch/record/1997189>.

- 3027 [ii] The Super Proton Synchrotron. Jan 2012. URL <http://cds.cern.ch/record/1997188>.
- 3028
- 3029 [12] Lyndon Evans and Philip Bryant. Lhc machine. *Journal of Instrumentation*, 3(08):S08001, 2008. URL <http://stacks.iop.org/1748-0221/3/i=08/a=S08001>.
- 3030
- 3031 [13] Fabienne Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. Oct 2013. URL <https://cds.cern.ch/record/1621583>. General Photo.
- 3032
- 3033 [14] CERN. LHC first beam: a day to remember, 2008. URL <http://cerncourier.com/cws/article/cern/36296>.
- 3034
- 3035 [15] Giulio Stancari, Valentina Previtali, Alexander Valishev, Roderik Bruce, Stefano Redaelli, Adriana Rossi, and Belen Salvachua Ferrando. Conceptual design of hollow electron lenses 3036 for beam halo control in the Large Hadron Collider. Technical Report FERMILAB-TM- 3037 2572-APC. FERMILAB-TM-2572-APC, CERN, Geneva, Oct 2014. URL <https://cds.cern.ch/record/1700455>. Comments: 23 pages, 1 table, 10 figures.
- 3038
- 3039
- 3040 [16] The ATLAS Collaboration, G Aad, E Abat, J Abdallah, and et al. A A Abdelalim. The atlas 3041 experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, 2008. 3042 URL <http://stacks.iop.org/1748-0221/3/i=08/a=S08003>.
- 3043
- 3044 [17] ATLAS Collaboration. Study of the material of the atlas inner detector for run 2 of the lhc, 2017.
- 3045
- 3046 [18] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design 3047 Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010. URL 3048 <https://cds.cern.ch/record/1291633>.
- 3049
- 3050 [19] Christopher C Tully. *Elementary particle physics in a nutshell*. Princeton Univ. Press, Princeton, NJ, 2011. URL <https://cds.cern.ch/record/1417476>.
- 3051
- 3052 [20] F Ahmadov, L Alio, BMM Allbrooke, T Bristow, D Buescher, A Buzatu, Y Coadou, C Debenedetti, Y Enari, G Facini, W Fisher, P Francavilla, G Gaycken, J Gentil, R Goncalo, 3053 G Gonzalez Parra, JF Grivaz, C Gwilliam, S Hageboeck, G Halladjian, M Jackson, D Jamin, R Jansky, K Kiuchi, V Kostyukhin, K Lohwasser, and et al. Lopez Mateos, D. Supporting 3054 Document for the Search for the bb decay of the Standard Model Higgs boson in associated 3055

- 3056 (W/Z)H production with the ATLAS detector. Technical Report ATL-COM-PHYS-2014-
 3057 051, CERN, Geneva, Jan 2014. URL <https://cds.cern.ch/record/1645654>.
- 3058 [21] Aidan Robson, Giacinto Piacquadio, and Elisabeth Schopf. Signal and Background Mod-
 3059 elling Studies for the Standard Model $VH, H \rightarrow b\bar{b}$ and Related Searches: Modelling sup-
 3060 port note for VH(bb) 2015+2016 dataset publication. Technical Report ATL-COM-PHYS-
 3061 2016-1747, CERN, Geneva, Dec 2016. URL <https://cds.cern.ch/record/2235887>.
 3062 This is a support note for the VH(bb) SM publication using the 2015+2016 datasets.
- 3063 [22] S. Alioli et al. NLO Higgs boson production via gluon fusion matched with shower in
 3064 POWHEG. *JHEP*, 0904:002, 2009.
- 3065 [23] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to
 3066 PYTHIA 8.1. *Comput.Phys.Commun.*, 178:852–867, 2008. doi: [10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036).
- 3067 [24] Georges Aad et al. Measurement of the Z/γ^* boson transverse momentum distribution in
 3068 pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 09:145, 2014. doi: [10.1007/JHEP09\(2014\)145](https://doi.org/10.1007/JHEP09(2014)145).
- 3069 [25] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. doi:
 3070 [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- 3071 [26] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. Madgraph
 3072 $\text{\texttt{5}}$: Going beyond. 2011. doi: [10.1007/JHEP06\(2011\)128](https://doi.org/10.1007/JHEP06(2011)128).
- 3073 [27] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and
 3074 R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-*
 3075 *2011-002*, CERN, Geneva, 2011.
- 3076 [28] M.L. Ciccolini, S. Dittmaier, and M. Kramer. Electroweak radiative corrections to associated
 3077 WH and ZH production at hadron colliders. *Phys.Rev.*, D68:073003, 2003. doi: [10.1103/PhysRevD.68.073003](https://doi.org/10.1103/PhysRevD.68.073003).
- 3078 [29] Michał Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross
 3079 Section at Hadron Colliders Through $\alpha(\frac{4}{3})$. *Phys. Rev. Lett.*, 110:252004, 2013. doi: [10.1103/PhysRevLett.110.252004](https://doi.org/10.1103/PhysRevLett.110.252004).
- 3080 [30] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, et al. HATHOR: Hadronic Top and
 3081 Heavy quarks cross section calculator. *Comput.Phys.Commun.*, 182:1034–1046, 2011. doi:
 3082 [10.1016/j.cpc.2010.12.040](https://doi.org/10.1016/j.cpc.2010.12.040).
- 3083

- 3086 [31] P. Kant, O.M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Molbitz, P. Rieck, and P. Uwer.
 3087 Hathor for single top-quark production: Updated predictions and uncertainty estimates for
 3088 single top-quark production in hadronic collisions. *Computer Physics Communications*, 191:
 3089 74 – 89, 2015. ISSN 0010-4655. doi: <http://dx.doi.org/10.1016/j.cpc.2015.02.001>. URL <http://www.sciencedirect.com/science/article/pii/S0010465515000454>.
- 3090
 3091 [32] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc.*
 3092 *Suppl.*, 205–206:10–15, 2010. doi: <10.1016/j.nuclphysbps.2010.08.011>.
- 3093 [33] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*.
 3094 Cambridge University Press, 2014. ISBN 1107034736, 9781107034730.
 3095 URL <http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model>.
- 3096
 3097 [34] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009. doi: <10.1088/1126-6708/2009/02/007>.
- 3100 [35] Andy Buckley, Jonathan Butterworth, David Grellscheid, Hendrik Hoeth, Leif Lonnblad,
 3101 James Monk, Holger Schulz, and Frank Siegert. Rivet user manual, 2010.
- 3102 [36] Iain W. Stewart and Frank J. Tackmann. Theory uncertainties for higgs and other searches
 3103 using jet bins. 2011. doi: <10.1103/PhysRevD.85.034011>.
- 3104 [37] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. Hw ±/hz + o and
 3105 1 jet at nlo with the powheg box interfaced to gosam and their merging within minlo. *Journal*
 3106 *of High Energy Physics*, 2013(10):83, 2013. ISSN 1029-8479. doi: [10.1007/JHEP10\(2013\)083](10.1007/JHEP10(2013)083).
 3107 URL [http://dx.doi.org/10.1007/JHEP10\(2013\)083](http://dx.doi.org/10.1007/JHEP10(2013)083).
- 3108 [38] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013.
 3109 doi: <10.1016/j.nuclphysb.2012.10.003>.
- 3110 [39] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer,
 3111 P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading or-
 3112 der differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:
 3113 079, 2014. doi: [10.1007/JHEP07\(2014\)079](10.1007/JHEP07(2014)079).
- 3114 [40] ATLAS Run 1 Pythia8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva,
 3115 Nov 2014. URL <https://cds.cern.ch/record/1966419>.

- 3116 [41] Michiel Botje et al. The PDF4LHC Working Group Interim Recommendations. 2011.
- 3117 [42] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter.
3118 Event generation with sherpa 1.1. *Journal of High Energy Physics*, 2009(02):007, 2009. URL
3119 <http://stacks.iop.org/1126-6708/2009/i=02/a=007>.
- 3120 [43] Nils Lavesson and Leif Lonnblad. W+jets matrix elements and the dipole cascade. 2005. doi:
3121 [10.1088/1126-6708/2005/07/054](https://doi.org/10.1088/1126-6708/2005/07/054).
- 3122 [44] Adrian Buzatu and Wei Wang. Object selections for SM Higgs boson produced in association
3123 with a vector boson in which $H \rightarrow b\bar{b}$ and V decays leptonically with Run-2 data: Object
3124 support note for VH(bb) 2015+2016 dataset publication. Technical Report ATL-COM-
3125 PHYS-2016-1674, CERN, Geneva, Nov 2016. URL <https://cds.cern.ch/record/2233686>. This is a support note for the VH(bb) SM publication using the 2015+2016
3126 datasets.
- 3127 [45] M Delmastro, S Gleyzer, C Hengler, M Jimenez, T Koffas, M Kuna, K Liu, Y Liu, G Mar-
3128 chiori, E Petit, M Pitt, E Soldatov, and K Tackmann. Photon identification efficiency mea-
3129 surements with the ATLAS detector using LHC Run 1 data. Technical Report ATL-COM-
3130 PHYS-2014-949, CERN, Geneva, Aug 2014. URL <https://cds.cern.ch/record/1747242>.
- 3131 [46] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-
3132 proton collision data. Technical Report ATLAS-CONF-2014-032, CERN, Geneva, Jun 2014.
3133 URL <https://cds.cern.ch/record/1706245>.
- 3134 [47] ATLAS Collaboration. MCPAnalysisGuidelinesMC15, 2015. URL <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC15>.
- 3135 [48] Muon reconstruction performance in early $\sqrt{s}=13$ TeV data. Technical Report ATL-PHYS-
3136 PUB-2015-037, CERN, Geneva, Aug 2015. URL <https://cds.cern.ch/record/2047831>.
- 3137 [49] ATLAS Collaboration. Performance of missing transverse momentum reconstruction for
3138 the ATLAS detector in the first proton-proton collisions at $\sqrt{s}=13$ TeV. Technical Re-
3139 port ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul 2015. URL <https://cds.cern.ch/record/2037904>.
- 3140
- 3141
- 3142
- 3143
- 3144
- 3145

- 3146 [50] Gavin P. Salam. Towards jetography. 2009. doi: 10.1140/epjc/s10052-010-1314-6.
- 3147 [51] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder,
3148 and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical
3149 Report ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, CERN, Geneva, Apr
3150 2008. URL <https://cds.cern.ch/record/1099735>.
- 3151 [52] Loch, Peter and Lefebvre, Michel . Introduction to Hadronic Calibration in AT-
3152 LAS, 2007. URL <https://indico.cern.ch/event/13514/contributions/138580/subcontributions/12940/attachments/107156/152645/IntroLectureDraft2-8.pdf>.
- 3155 [53] Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector
3156 at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul 2015.
3157 URL <https://cds.cern.ch/record/2037613>.
- 3158 [54] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector.
3159 Technical Report ATLAS-CONF-2014-018, CERN, Geneva, May 2014. URL <https://cds.cern.ch/record/1700870>.
- 3161 [55] B Heinemann, F Hirsch, and S Strandberg. Performance of the ATLAS Secondary Vertex
3162 b-tagging Algorithm in 7 TeV Collision Data. Technical Report ATLAS-COM-CONF-2010-
3163 042, CERN, Geneva, May 2010. URL <https://cds.cern.ch/record/1266283>. (Was
3164 originally 'ATL-COM-PHYS-2010-274').
- 3165 [56] Expected performance of the ATLAS b -tagging algorithms in Run-2. Technical Report ATL-
3166 PHYS-PUB-2015-022, CERN, Geneva, Jul 2015. URL <https://cds.cern.ch/record/2037697>.
- 3168 [57] Gordon Watts, Frank Filthaut, and Giacinto Piacquadio. Extrapolating Errors for b -tagging.
3169 Technical Report ATL-COM-PHYS-2015-7II, CERN, Geneva, Jul 2015. URL <https://cds.cern.ch/record/2034234>. This is for internal information only, no approval to ever be
3170 seen outside of ATLAS.
- 3172 [58] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for
3173 likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. doi: 10.1140/epjc/
3174 s10052-011-1554-0.

- 3175 [59] A. L. Read. Presentation of search results: the CL_s technique. *J. Phys. G*, 28:2693–2704,
 3176 2002.
- 3177 [60] Stephen Kam-wah Chan and John Huth. Variations on MVA Variables in the SM $ZH \rightarrow$
 3178 $\ell\ell b\bar{b}$ Search. Technical Report ATL-COM-PHYS-2017-1318, CERN, Geneva, Aug 2017. URL
 3179 <https://cds.cern.ch/record/2281412>.
- 3180 [61] B Clark, D Lopez Mateos, N Felt, J Huth, and J Oliver. An Algorithm for Micromegas Seg-
 3181 ment Reconstruction in the Level-1 Trigger of the New Small Wheel. Technical Report ATL-
 3182 UPGRADE-INT-2014-001, CERN, Geneva, Sep 2014. URL [https://cds.cern.ch/](https://cds.cern.ch/record/1753329)
 3183 [record/1753329](https://cds.cern.ch/record/1753329).
- 3184 [62] Stephen Kam-wah Chan, David Lopez Mateos, and John Huth. Micromegas Trigger Pro-
 3185 cessor Algorithm Performance in Nominal, Misaligned, and Corrected Misalignment Condi-
 3186 tions. Technical Report ATL-COM-UPGRADE-2015-033, CERN, Geneva, Dec 2015. URL
 3187 <https://cds.cern.ch/record/2113121>.
- 3188 [63] S Chan, J Huth, D Lopez Mateos, and K Mercurio. $ZH \rightarrow ll b\bar{b}$ Analysis with Telescoping
 3189 Jets. Technical Report ATL-PHYS-INT-2015-002, CERN, Geneva, Jan 2015. URL <https://cds.cern.ch/record/1982439>.
- 3190 [64] A. Hornig T. S. Roy D. Krohn S. D. Ellis, S. D. Ellis and M. D. Schwartz. Qjets: A non-
 3191 deterministic approach to tree-based jet substructure. *Phys. Rev. Lett.*, 108, 2012. URL
 3192 <http://arxiv.org/abs/1201.1914>.
- 3193 [65] G. D. Leder S. Moretti Y. L. Dokshitzer, Y. L. Dokshitzer and B. R. Webber. Better jet clus-
 3194 tering algorithms. *JHEP*, 9708, 1997. URL <http://arxiv.org/abs/hep-ph/9707323>.
- 3195 [66] G. P. Salam M. Cacciari, M. Cacciari and G. Soyez. The anti- k_t jet clustering algorithm.
 3196 *JHEP*, 0804, 2008. URL <http://arxiv.org/abs/0802.1189>.
- 3197 [67] D. Krohn D. Kahawala, D. Kahawala and M. D. Schwartz. Jet sampling: Improving event
 3198 reconstruction through multiple interpretations. 2013. URL <http://arxiv.org/abs/1304.2394>.
- 3199 [68] S. Mrenna T. Sjöstrand, T. Sjöstrand and P. Z. Sands. A brief introduction to pythia 8.1.
 3200 *Comput. Phys. Commun.*, 178, 2008. URL <http://arxiv.org/abs/0710.3820>.

- 3203 [69] D. Stum J. Huston H. Lai Pm M. Nadolsky et al. J. Pumplin, J. Pumplin. New generation of
 3204 parton distributions with unvertainties from global qcd analysis. *JEGP*, 0207, 2002. URL
 3205 <http://arxiv.org/abs/hep-ph/0201195>.
- 3206 [70] ATLAS tunes of PYTHIA 6 and Pythia 8 for MCII. Technical Report ATL-PHYS-PUB-2011-
 3207 009, CERN, Geneva, Jul 2011. URL <http://cds.cern.ch/record/1363300>.
- 3208 [71] New ATLAS event generator tunes to 2010 data. Technical Report ATL-PHYS-PUB-2011-
 3209 008, CERN, Geneva, Apr 2011. URL <https://cds.cern.ch/record/1345343>.
- 3210 [72] T. Gleisberg et al. T. Gleisberg et al. Event generation with sherpa 1.1. *JHEP*, 02, 2009. URL
 3211 <http://arxiv.org/abs/0811.4622>.
- 3212 [73] M. Cacciari, M. Cacciari G. P. Salam, and G. P. Salam. Pileup subtraction using jet areas.
 3213 *Phys. Lett.*, B659, 2008. URL <http://arxiv.org/abs/0707.1378>.
- 3214 [74] Jet energy measurement with the atlas detector in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$
 3215 *Eur. Phys. J.*, C73, 2013. URL <http://arxiv.org/abs/1112.6426>.
- 3216 [75] b -jet tagging calibration on c -jets containing D^{*+} mesons. Technical Report ATLAS-CONF-
 3217 2012-039, CERN, Geneva, Mar 2012. URL <https://cds.cern.ch/record/1435193>.
- 3218 [76] Search for the bb decay of the Standard Model Higgs boson in associated W/ZH production
 3219 with the ATLAS detector. Technical Report ATLAS-CONF-2013-079, CERN, Geneva, Jul
 3220 2013. URL <https://cds.cern.ch/record/1563235>.
- 3221 [77] J. Thaler L. Wang D. Krohn, D. Krohn. Jet trimming. 2009. URL <http://arxiv.org/abs/0912.1342>.
- 3223 [78] Electron performance measurements with the atlas detector using the 2010 lhc proton-proton
 3224 collision data. *Eur. Phys. J.*, C72, 2012. URL <http://arxiv.org/abs/1110.3174v2>.
- 3225 [79] Measurement of the $w \rightarrow l\nu$ and $z/\gamma^* \rightarrow ll$ production corss sections in proton-proton
 3226 collisions at $\sqrt{s} = 7 \text{ TeV}$ with the atlas detector. *JHEP*, 1012, 2010. URL <http://arxiv.org/abs/1010.2130>.
- 3228 [80] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling. In *2003 Computing in
 3229 High Energy and Nuclear Physics, CHEP03*, 2003.

- 3230 [81] Yang-Ting Chien, David Farhi, David Krohn, Andrew Marantan, David Lopez Mateos, and
3231 Matthew Schwartz. Quantifying the power of multiple event interpretations. 2014. doi:
3232 [10.1007/JHEP12\(2014\)140](https://doi.org/10.1007/JHEP12(2014)140).