

Measuring the mass of the Higgs boson in the four-lepton final state with the ATLAS
detector

A Dissertation

Presented to

The Faculty of the Graduate School of Arts and Sciences
Brandeis University

Department of Physics

Gabriella Sciolla, Advisor

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
Hannah Herde

August 2020



This dissertation, directed and approved by Hannah Herde's Committee, has been accepted and approved by the Faculty of Brandeis University in partial fulfillment of the requirements for the degree of:

DOCTOR OF PHILOSOPHY

Eric Chasalow, Dean
Graduate School of Arts and Sciences

Dissertation Committee:
Gabriella Sciolla, Physics
Craig Blocker, Physics
Alessandro Tricoli, Brookhaven National Laboratory

Copyright by
Hannah Herde

2020

ABSTRACT

Measuring the mass of the Higgs boson in the four-lepton final state with the ATLAS detector

A dissertation presented to the Faculty of the
Graduate School of Arts and Sciences of Brandeis University
Waltham, Massachusetts

By Hannah Herde

We measure the mass of the Standard Model-like Higgs boson in the four-lepton decay channel using 140 fb^{-1} of proton-proton collisions at the Large Hadron Collider at $\sqrt{s}=13$ TeV collected by the ATLAS detector. This corresponds to the full Run 2 dataset, spanning 2015-2018. For a Standard Model Higgs boson with $m_H = 125 \text{ GeV}$, the expected statistical uncertainty is 177 MeV using a per-event strategy. The observed mass of the Higgs boson is $124.92 \pm 0.19 \text{ (stat)} \pm 0.08 \text{ (sys)} \text{ GeV}$ with the same strategy.

Acknowledgements

To my dissertation advisor, **Gabriella Sciolla**, we did it! I have learned and grown so much working with you for the last six years. Thank you for your guidance and support, both in good times (like the first time we programmed the XYZ stage at BNL) and challenging times (like when we were teaching and it would just not stop snowing whenever we scheduled lecture). Thank you for listening when I was frustrated, and helping me find the path again. I feel blessed that you took me on as a student. I will do my best to be a mentor for others as you were for me. Thanks for being my advisor.

To **Craig Blocker** and **Alessandro Tricoli**, thank you for reading this dissertation and providing your thoughtful and detailed feedback. I particularly appreciated your comments intended to help me think more deeply about physics processes.

Several postdocs dedicated substantial time to training and guiding me. Thank you, **Giacomo Artoni**, **Gaetano Barone**, **Max Goblirsch**, and **Stefano Zambito**. Thank you to every teacher I have had during my education. Your encouragement and lessons mean a great deal to me.

To my labmates in the Sciolla Group, thanks for all the lunches, game and movie nights, tea breaks, cooking, ranting sessions, and debugging. **Laura Bergsten**, you taught me about letting go and letting be. Thank you for teaching me, “Who’s to say?” and going on this adventure with me. Thank you for being my friend.

To my colleagues in the ATLAS $H \rightarrow 4\ell$ group (especially **Tom Powell** and **Bianca Ciungu**) and the ITk Strips upgrade community (especially **Jens Dopke**, **Luise Poley**, **Dave Lynn**, **Stefania Stucci**, and **Phil Kuczewski**), thank you for your counsel and your welcome. Thank you for your faith in me and your patience.

To **Graham Armstrong** and the **Large Improv Collider**, y’all got me out of my head when I desperately needed it. You trained me to be a better communicator and a better listener. I promise to carry improv’s lessons with me and treat them as life lessons,

as **Christian Riegel**, **Salvatore Fabbiano**, and **Sam Lagier** trained us.

To the **Wellesley Alumnae Network**, thank you for your love and support. I am honored to be among you.

To my friends, especially **Raji Nagalla**, there are not words good enough. Thank you for everything. Ditto to my family. This accomplishment belongs to you, **Nancy**, **Michael**, **Isabelle**, **George**, and **Sarah**. I would like to give a special shout-out to every **Herde** pet. Your absurd behavior brought so much warmth and joy.

Martin Mörtzell, one day I am going to marry you every day.^a In true improv-fashion, we work together to make each other look good. Thank you for accepting me and letting me accept you. Thank you for coming into my life.

Finally, my work benefits from all the physicists and engineers who came before me, especially my Black colleagues. I performed this work in part on land stolen from the Massa-adchu-es-et (Massachusetts) tribe and the Setauket tribe.^b

^aOK, only one day. But still.

^bNative Land Digital (native-land.ca) is a work-in-progress website from a Canadian Indigenous-led non-profit. It displays the approximate boundaries of Indigenous tribes from whom we stole land to build the white-dominated countries we know today. I have named the tribes indicated on the map for the locations of Brandeis University and Brookhaven National Lab. The named list is not comprehensive.

Table of Contents

Abstract	iv
Acknowledgements	v
Table of Contents	vii
Introduction	1
I Theoretical motivation	2
1 Role of the Higgs boson in the Standard Model of Particle Physics	2
1.1 The Standard Model of Particle Physics	2
1.2 Motivation for the Higgs boson	6
2 The value in measuring the Higgs boson's mass	14
II Apparatus	18
3 Large Hadron Collider	18
4 The ATLAS Detector	21
5 Data collection	24
III Experimental method	26
6 Event selection	26
6.1 Triggers	27
6.2 Lepton reconstruction and calibration	28
6.3 Muon reconstruction and calibration	29
6.4 Electron reconstruction and calibration	38
6.5 Event-level selection	45
6.6 Defining the four-lepton invariant mass, $m_{4\ell}$	46
7 Analyzing the $m_{4\ell}$ spectrum	49

7.1	Event categorization	51
7.2	Signal model development	54
7.3	Defining event resolution	63
7.4	Signal model conditional on per-event resolution	75
7.5	Background model	75
7.6	Creating the full probability density function (PDF)	81
8	Systematic uncertainties and building the likelihood	85
8.1	Lepton uncertainties	86
8.2	$qq \rightarrow ZZ^{(*)}$ shape uncertainty within BDT_{ZZ} bins	90
8.3	Defining the likelihood function and test statistic	92
IV	Results	94
9	Expected results	96
9.1	Full MC	96
9.2	Sampled MC	98
9.3	Summary	104
10	Observed results	105
10.1	Comparing key distributions in data and simulation	106
10.2	Observed m_H	114
10.3	Observed impact of accounting for per-event resolution	123
10.4	Summary	125
V	Outlook	126
11	ATLAS' future: Phase II Inner Tracker Upgrade and the High-Luminosity LHC (HL-LHC)	126
12	The Higgs mass at the HL-LHC	129

Bibliography	132
---------------------	------------

Appendices	142
-------------------	------------

A	Contributions to this work	142
A.1	The story of my doctoral work	143
A.2	Contributions to this dissertation and measurement	145
B	The mathematics of quantum electrodynamics	149
C	$H \rightarrow 4\ell$ ATLAS triggers during Run 2	150
D	Simulating $H \rightarrow ZZ^* \rightarrow 4\ell$ events	153
E	Calibrating the QRNN target percentile in each decay channel	156
F	Event-level resolution	159
F.1	$4e$ Channel	159
F.2	$2\mu 2e$ Channel	160
F.3	$2e 2\mu$ Channel	161
F.4	QRNN resolution data-simulation agreement	162
G	Background PDF shapes	163
H	Template method	169
H.1	Signal model construction	169
H.2	Full template model	191

List of Tables

1	Standard Model quark properties	3
2	Standard Model charged lepton properties	4
3	Rotation groups in the Standard Model	5
4	Muon and electron selection requirements for $H \rightarrow 4\ell$ analyses	28
5	Impact of the supercluster electron reconstruction on the Higgs resonance in the 4ℓ final state	44
6	Event selection criteria for $H \rightarrow 4\ell$ analyses	45
7	Impact of final state radiation recovery and the Z-mass constraint on the expected Higgs mass precision	47
8	Per-event methods	56
9	First and third quantiles of $m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}$ in each decay channel . . .	67
10	Calibrated target percentile for each decay channel	70
11	KEYS smoothing parameter for each background process	78
12	$Z + \mu\mu$ control region definitions	84
13	Impact of MUON_SCALE nuisance parameter	87
14	Impact of largest EG scale systematic uncertainty coming from the $Z \rightarrow ee$ non-closure	90
15	Expected m_H central value and uncertainty	104
16	Expected and observed yields for events with $105 < m_{4\ell}^{\text{constrained}} < 160$ GeV .	105
17	Expected and observed yields for events with $115 < m_{4\ell}^{\text{constrained}} < 130$ GeV .	105
18	Observed m_H central value and uncertainty	114
19	Observed fit results with 2D template method (cross check)	115
20	Single-channel and combined measurements of m_H compatibility using the average resolution model	115
21	Observed fit results with average resolution method	116

22	Single-channel and combined measurements of m_H compatibility using the per-event resolution model	121
23	Observed fit results with the per-event resolution method	121
24	Impact of systematic uncertainties for the per-event resolution method . . .	121
25	One-sided p-value for observed uncertainty	124
26	Expected uncertainty on the measured mass of the Higgs boson for the current and upgraded detector scenarios with 3000 fb ⁻¹ of HL-LHC data	131
27	Summary of the 2015 HLT triggers	151
28	Summary of the 2016 HLT triggers	151
29	Summary of the 2017 HLT triggers	152
30	Summary of the 2018 HLT triggers	152
31	List of m_H values used in the production of simulated Higgs signal samples .	154
32	Signal sample Monte Carlo generators and parton distribution functions . . .	154
33	Cross section predictions for normalizing simulated signal samples	155
34	Description of the production of simulated background samples	155
35	Cross check expected results using the 2D template method	192

List of Figures

1	A timeline of major discoveries in particle physics	1
2	Fundamental particles of the Standard Model	2
3	Beta decay	7
4	Example of spontaneous symmetry breaking	9
5	Standard Model Higgs boson production at the Large Hadron Collider	14
6	Standard Model Higgs boson branching ratios	15
7	Global electroweak fit	16
8	Standard Model electroweak vacuum stability	16
9	Indirect determination of the Higgs boson’s mass	19
10	CERN accelerator complex	19
11	ATLAS detector schematic	22
12	Particle detection in a sector of the ATLAS detector	24
13	ATLAS’ collected integrated luminosity	25
14	Run 2 profile for the number of interactions per bunch crossing	25
15	Muon’s journey in ATLAS	29
16	Muon-sensitive subdetector coverage	29
17	Muon types	31
18	Muon expected resolution and reconstruction efficiency	32
19	Muon momentum corrections η binning	33
20	Roles of muon momentum calibration parameters	35
21	Impact of applying muon momentum corrections at the J/ψ and Z resonances	36
22	Dimuon invariant mass scale at the J/ψ and Z resonances	36
23	Dimuon invariant mass resolution at the J/ψ and Z resonances	37
24	Relative dimuon momentum resolution	37
25	Electron’s journey in ATLAS	38

26	Electron-sensitive subdetector coverage	38
27	Schematic definition of a supercluster for electron reconstruction	40
28	Calibrated electron energy response resolution without pileup	41
29	Calibrated electron energy response resolution accounting for pileup	41
30	Electron energy scale and resolution at the Z resonance	42
31	Z resonance in the dielectron invariant mass spectrum	42
32	Electron reconstruction efficiency	43
33	Impact of the supercluster electron reconstruction on the Higgs resonance in the 4ℓ final state	44
34	Impact of final state radiation and the Z -mass constraint on the four-lepton resonance and leading dilepton invariant mass spectra	46
35	Flowchart for final state radiation recovery procedure	48
36	Extrapolated expected uncertainty as a function of integrated luminosity . .	49
37	$m_{4\ell}^{\text{constrained}}$ distribution for ggH signal events in each decay channel	52
38	BDT_{ZZ} discriminant distribution	54
39	Example of template morphing from the full Run 2 template method cross check	56
40	Example binning palette for formulating the lepton response lookup table for the per-event response method	57
41	Justification for selecting Double-sided Crystal Ball function to parameterize the $m_{4\ell}^{\text{constrained}}$ distribution	62
42	Momentum and energy pulls for individual muons and electrons, given the lepton errors	64
43	Quantile regression neural network's (QRNN) architecture	68
44	QRNN validation in a single bin of phase space	71
45	Summary estimated resolution (r_{QRNN}) performance plots as a function of generated mass point and BDT_{ZZ} for the 4μ channel	72

46	Summary estimated resolution (r_{QRNN}) performance plot as a function of detector region for the 4μ channel	73
47	Summaries comparing the target percentile and QRNN score in slices of $p_T^{4\ell}$, lepton η , and $\sigma_{m_{4\ell}}$ for the 4μ channel	74
48	Expected background composition in each analysis category	77
49	Distributions of event-level resolution in the 4μ final state in signal and back- ground simulated samples	79
50	Average resolution in signal and background simulated samples as a function BDT_{ZZ} score	80
51	Impact of MUON_SCALE nuisance parameter in the 4μ channel	88
52	Impact of muon resolution nuisance parameters in the 4μ channel	89
53	Impact of EG_SCALE_ZEESYST in the $4e$ channel	89
54	$m_{4\ell}^{\text{constrained}}$ distributions in the $qq \rightarrow ZZ^{(*)}$ non-resonant background in the 4μ final state	91
55	Expected m_H^{measured} in the full MC dataset	97
56	Correlation matrix of the parameters of interest obtained when fitting the full MC dataset with the average resolution model	98
57	Expected m_H^{measured} in the sampled MC dataset	99
58	Symmetrized expected uncertainty distributions	99
59	Distribution of best-fit values of m_H using the average resolution model . . .	100
60	Lower and upper statistical uncertainties on m_H from each toy fit using the average resolution method	101
61	Pull on m_H^{measured} using the average resolution method	101
62	Distribution of best-fit values of m_H using the per-event resolution model . .	102
63	Lower and upper statistical uncertainties on m_H from each toy fit using the per-event resolution method	103
64	Pull on m_H^{measured} using the per-event resolution method	103

65	Observed four-lepton invariant mass distribution, inclusive of decay channel	106
66	Observed four-lepton invariant mass distribution in each decay channel . . .	107
67	Observed BDT_{ZZ} distribution, inclusive of decay channel	108
68	Observed BDT_{ZZ} distribution in each decay channel	109
69	Observed QRNN score distribution, inclusive of decay channel	110
70	Observed QRNN score distribution in each decay channel	111
71	Distribution of events in the QRNN score- $m_{4\ell}^{\text{constrained}}$ plane, inclusive of decay channel	112
72	Distribution of events in the QRNN score- $m_{4\ell}^{\text{constrained}}$ plane in each decay channel	113
73	Observed NLL scan and best-fit values of m_H with the average resolution method	116
74	Impact of systematic uncertainties and correlation of parameters of interest for the average resolution method	117
75	Projected per-event resolution fit result on observed dataset, inclusive of decay channel	118
76	Projected per-event resolution fit result on observed dataset in each final state	119
77	Observed NLL scan and best-fit values of m_H with the per-event resolution method	120
78	Impact of systematic uncertainties and correlation of parameters of interest for the per-event resolution method	122
79	Toy-by-toy comparisons of the m_H central value and expected uncertainties .	124
80	Distribution of the total expected m_H uncertainty indicating observed uncer- tainty	125
81	Event displays for LHC and HL-LHC	127
82	Eventual Inner Tracker detector volume	128
83	Anatomy of the ITk Strips barrel	128
84	Thermomechanical prototype Inner Tracker stave	129

85	Extrapolated expected uncertainty on m_H as a function of integrated luminosity in the HL-LHC era	130
86	Simulated combined muon momentum resolution using the Inner Tracker . .	130
87	PhD on a page	142
88	QRNN target percentile calibration	158
89	Summary estimated resolution (r_{QRNN}) performance plots as a function of generated mass point, BDT_{ZZ} , and detector region for the $4e$ channel	159
90	Summaries comparing the target percentile and QRNN score in slices of $p_T^{4\ell}$, lepton η , and $\sigma_{m_{4\ell}}$ for the $4e$ channel	159
91	Summary estimated resolution (r_{QRNN}) performance plots as a function of generated mass point, BDT_{ZZ} , and detector region for the $2\mu 2e$ channel . . .	160
92	Summaries comparing the target percentile and QRNN score in slices of $p_T^{4\ell}$, lepton η , and $\sigma_{m_{4\ell}}$ for the $2\mu 2e$ channel	160
93	Summary estimated resolution (r_{QRNN}) performance plots as a function of generated mass point, BDT_{ZZ} , and detector region for the $2e 2\mu$ channel . . .	161
94	Summaries comparing the target percentile and QRNN score in slices of $p_T^{4\ell}$, lepton η , and $\sigma_{m_{4\ell}}$ for the $2e 2\mu$ channel	161
95	Comparison of the QRNN score in data and simulation using side bands . .	162
96	$qq \rightarrow ZZ^*$ PDF	164
97	$gg \rightarrow ZZ^*$ PDF	165
98	$EWZZ$ PDF	166
99	tXX and VVV PDF	167
100	$Z + Jets, t\bar{t}, WZ$ background PDF	168
101	Smoothed distribution in 4μ in the ggH production mode	170
102	Smoothed distribution in $4e$ in the ggH production mode	171
103	Smoothed distribution in $2\mu 2e$ in the ggH production mode	172
104	Smoothed distribution in $2e 2\mu$ in the ggH production mode	173

105	Smoothed distribution in 4μ in the VBF production mode	174
106	Smoothed distribution in $4e$ in the VBF production mode	175
107	Smoothed distribution in $2\mu 2e$ in the VBF production mode	176
108	Smoothed distribution in $2e 2\mu$ in the VBF production mode	177
109	Smoothed distribution in 4μ in the W^-H production mode	178
110	Smoothed distribution in $4e$ in the W^-H production mode	179
111	Smoothed distribution in $2\mu 2e$ in the W^-H production mode	180
112	Smoothed distribution in $2e 2\mu$ in the W^-H production mode	181
113	Smoothed distribution in 4μ in the W^+H production mode	182
114	Smoothed distribution in $4e$ in the W^+H production mode	183
115	Smoothed distribution in $2\mu 2e$ in the W^+H production mode	184
116	Smoothed distribution in $2e 2\mu$ in the W^+H production mode	185
117	Smoothed distribution in 4μ in the ZH production mode	186
118	Smoothed distribution in $4e$ in the ZH production mode	187
119	Smoothed distribution in $2\mu 2e$ in the ZH production mode	188
120	Smoothed distribution in $2e 2\mu$ in the ZH production mode	189
121	Demonstration of template morphing using RooBSpline with the signal PDF in each analysis category for the Full Run 2 template method cross check . .	190
122	Demonstration of template morphing using RooBSpline with the full PDF in each analysis category for the Full Run 2 template method cross check . . .	191

Introduction

τῶν ἀμόθεν γε, θεά, θύγατερ Διός, εἰπὲ καὶ ἡμῖν.

So tell us, goddess, daughter of heaven, starting from
the middle.

Hom. Od. 1.10, translated by H. Herde

Nature is perpetually “in the middle.” Energy becomes matter which becomes energy again. The miasma of the known universe seems to simultaneously hold and abandon a form, to be something and nothing at all. Specks loose in this paradox, we have learned to observe. From the chaos of these observations, we (try to) discover fundamental order.

So far, we have puzzled out “the rules” for the behavior of about 5% of the energy in our known universe, arguably quite the accomplishment for several thousand years of study. Figure 1 displays a timeline of major discoveries in particle physics, dating back to Ancient India and Ancient Greece. This 5% is matter, the key ingredient of our daily lives. As far as we can tell, it obeys the Standard Model of Particle Physics. We do not know what the rest of the energy budget of the universe *is* or how it *works*. This dissertation is about matter and the hope that perhaps a deeper understanding of it will illuminate something about the rest. In other words, we are starting somewhere in the middle.

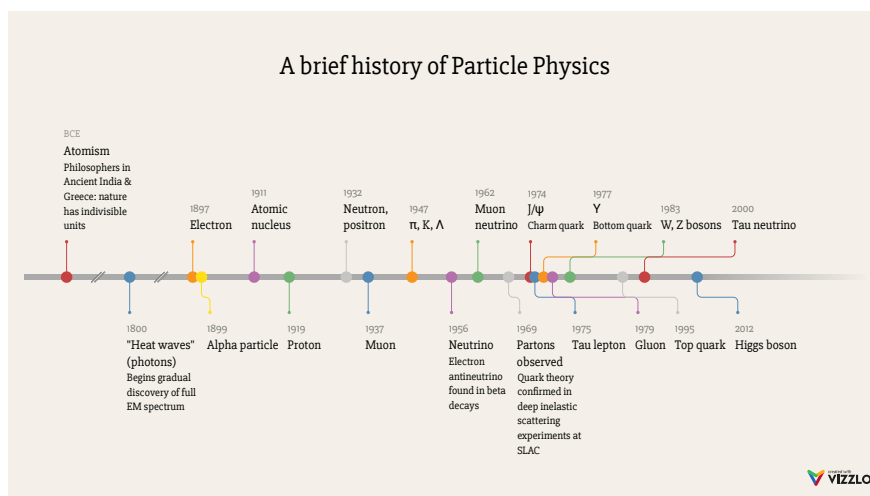


Figure 1: A timeline of major discoveries in particle physics

Source: HH (2017)

Chapter I

Theoretical motivation

1 Role of the Higgs boson in the Standard Model of Particle Physics

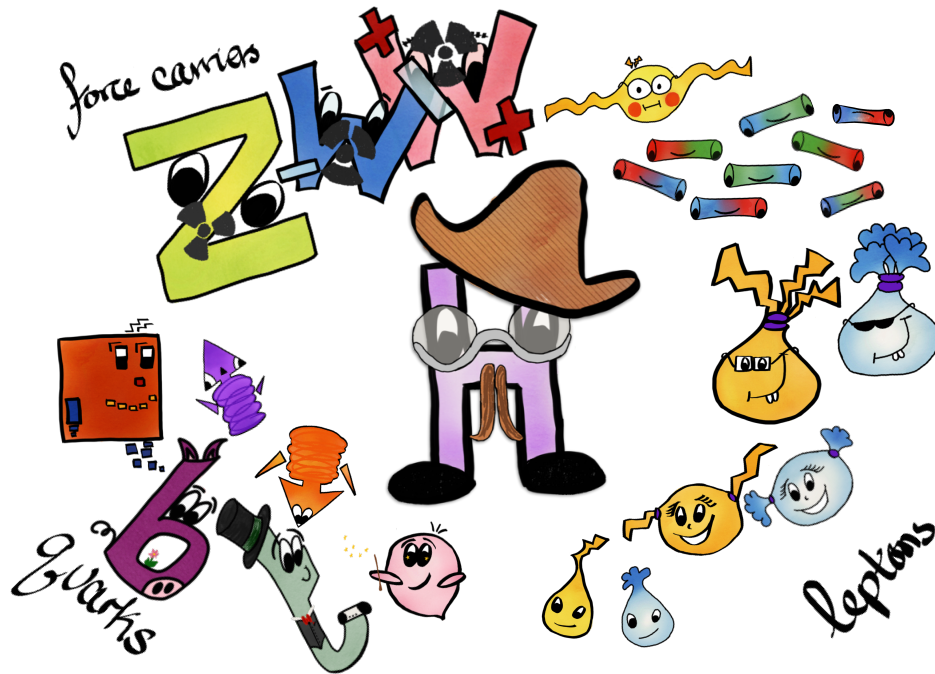


Figure 2: Artistic presentation of the fundamental particles predicted under the Standard Model of Particle Physics. Anti-particles are not pictured. *Source: HH (2018-2020)*

1.1 The Standard Model of Particle Physics

The Standard Model (SM) is a specific quantum field theory equation describing how “matter works.” The SM divides the building blocks of matter into three distinct groups of apparently indivisible components - quarks, leptons, and force carriers - and enumerates all the ways

that these particles might interact with each other. Frustratingly, the SM does not explain how matter in any form might interact with gravity. Quarks and leptons are fermions. They have intrinsic spins of $\frac{1}{2}$ and collections of them follow Fermi-Dirac statistics. Critically, two fermions cannot occupy identical quantum states. The force carriers are bosons. They have integer intrinsic spins and follow Bose-Einstein statistics. In fact, the four force carriers of the SM have spin=1, specifically. Unlike fermions, an unlimited number of bosons can occupy the same quantum state. Finally, every SM particle has an antiparticle with the same mass and opposite charge.

One SM-predicted particle is quite different from the rest of the menagerie. The Higgs boson is the only particle that truly couples to itself.¹ At first glance, the Higgs boson has almost no “personality” - no electric charge, no spin, no structure - but the SM falls apart without it.

1.1.1 Quarks

Quarks carry both electrical and color charges, indicating that they interact with the electromagnetic and strong forces. There are six flavors of quarks - up, down, strange, charm, bottom, and top. Each flavor of quark is progressively heavier than the last, notable in table 1. They change flavor through weak force interactions.

Table 1: Properties of SM quarks. All masses are approximate since quarks are confined [1].
Source: [PDG 2019](#)

Quark	Mass [GeV/c ²]	Electrical charge [in units of proton charge, e]
Up	$\sim 2.2 \times 10^{-3}$	$\frac{2}{3}$
Down	$\sim 4.7 \times 10^{-3}$	$-\frac{1}{3}$
Strange	$\sim 9.3 \times 10^{-2}$	$-\frac{1}{3}$
Charm	~ 1.27	$\frac{2}{3}$
Bottom	~ 4.18	$-\frac{1}{3}$
Top	~ 172.9	$\frac{2}{3}$

¹While gluons couple to other gluons, a single gluon does not interact with itself, while the Higgs boson can.

1.1.2 Leptons

Leptons come in two “species” - electrically charged and neutral. No SM leptons participate in strong interactions. . The charged leptons have electric charge of -1 (in units of e) and come in three flavors: the electron, muon, and tau. Each flavor is successively heavier, as table 2 shows. Only the electron is stable. The neutral leptons, called neutrinos, also come in those three flavors. In the SM framework, neutrinos are massless. Neutrino oscillation shows that neutrinos must in fact have masses, requiring us to modify the Standard Model. Neutrino oscillations are a critical area of study and beyond the scope of this dissertation.² All leptons participate in weak interactions.

Table 2: Properties of SM charged leptons [1].

Source: *PDG 2019*

Charged lepton	Mass [MeV/c ²]	Lifetime [s]
Electron	0.511	Stable
Muon	105.66	2.2×10^{-6}
Tau	1776.86	290.3×10^{-15}

1.1.3 Force carriers

The force carriers are the messengers, or **mediators**, of the SM forces. Figure 2 illustrates all 12 vector boson mediators - one photon, three weak bosons ($W^\pm$, Z), and eight gluons.³

We can reimagine the forces themselves as **fields**. The fields in mathematical spaces have many possible **representations** - scalars, vectors, tensors, *et cetera*. For example, we can think of each of the fermions as excitations of spinor fields. Now, the fermionic fields are matter fields. The forces act on the matter fields, causing them to change.

²Under the SM, right-handed neutrinos do not exist. The weak interaction obeys $SU(2)_L$ after SSB so it is inherently left-handed. Without right-handed neutrinos, we cannot build neutrino Yukawa mechanism terms. The SM contains all dimension-4 operators. One family of extensions to the SM adds dimension-5 Yukawa-style operators, postulating that right-handed neutrinos exist so neutrinos acquire their masses through the Higgs mechanism. Currently, we do not understand the origin of neutrino mass. Adventure awaits!

³Alternatively, we can think of a single gluon particle with eight possible linearly-independent states, forming the “color octet.” A gluon cannot take the color singlet state since long-range strong interactions do not exist. A gluon cannot truly interact with itself, though, as it inherently has one of the color states.

In order to make sense of the electromagnetic, weak, and strong force fields, we need to consider the structure of the SM itself. **The SM is a locally gauge-invariant quantum field theory based on the symmetry group $SU(3) \times SU(2) \times U(1)$.**⁴ In other words, the SM's mathematical representation is a Lagrangian invariant under a particular set of symmetries. Therefore, it must be composed of combinations of fields that are singlets under the prescribed set of transformations. Further, each symmetry establishes a conservation law.

The $U(1)$ symmetry group consists of rotations in one-dimensional complex space (or two-dimensional real space). We can articulate our rotation by an angle θ as:

$$\phi \rightarrow e^{iq\theta} \phi \qquad \phi_1 \phi_2 \rightarrow e^{i(q_1+q_2)\theta} \phi_1 \phi_2$$

It is clear that if $\sum q_i = 0$, then the rotation seems to have no effect – we are **invariant** under the transformation and $\sum q_i = 0$ constitutes a **singlet** of the $U(1)$ symmetry. If we interpret q as a charge, then we see that **charge conservation** is a symmetry under rotation in the mathematical $U(1)$ space.

We could arrive at similar conclusions for the $SU(2)$ and $SU(3)$ groups, keeping in mind that $SU(n)$ represents a rotation in an n -dimensional complex space. Table 3 provides the resulting representations and invariants under the different groups, keeping in mind that $SU(2)$ and $SU(3)$ are non-Abelian, non-commutative groups, so our choice of representation (scalar, spinor, vector, *et cetera*) matters.

Table 3: Invariant states permitted for each rotation group in the SM and their mathematical representations.

Symmetry	Mathematical representation	States
$U(1)$	Sum of charges is 0	Singlet
$SU(2)$	$2 \times 2 = 1 + 3$	Singlet + triplet
$SU(3)$	$3 \times 3 = \bar{3} + 6$ and $3 \times \bar{3} = 1 + 8$	Triples; singlet + octet

Next, we require that the symmetry is **local**, meaning that we could apply different transformations at different points on our field. Imposing locality automatically requires

⁴At least, it is until we need to mathematically account for the experimentally observed masses of the weak bosons and the fermions.

a derivative because we are effectively requiring a kinetic term to appear. That kinetic term allows us to have a dynamic system, but it is not invariant “out of the box” under our symmetry.⁵ The forces themselves save us here: we can introduce fields for each force through a **gauge symmetry**. The gauge symmetry couples the new field via the kinetic term, adding it to our model and making it dynamic.

This is crux of the SM: **Local symmetries imply force fields, otherwise the symmetry would hold throughout the entire mathematical space. Force fields are gauge fields. Couplings with the gauge fields create dynamic behavior – interactions.** We need to couple a gauge field to each kinetic term for each symmetry group present in the SM. The new fields must be massless and spin 1, with adjoint representations: Their adjoint representations must be: $q = 0$ (singlet) for $U(1)$, a triplet for $SU(2)$, and an octet for $SU(3)$. Looking back at figure 2, we see that we have 1 photon, 3 weak bosons ($W^\pm, Z$), and 8 gluons. It’s tempting to claim we have the full picture. Our picture, however, has a glaring flaw: all our vector bosons must be massless [2].

1.2 Motivation for the Higgs boson

Prior to the experimental observation of the weak bosons in the 1980s at CERN, we had observed neutral and charged weak currents. Beta decay (figure 3) provides an example of a reaction with a flavor-changing charged current. In beta decays, a neutron becomes a proton, an electron, and an anti-electron neutrino. We understood that one of the valence down quarks in the neutron must have decayed into an up quark, transforming the neutron into a proton and expelling the leptons in the process. The down quark could emit a mediator with electric charge -1 so that it could gain +1 electric charge to become a $+\frac{2}{3}$ -charged up quark. That mediator would also decay to the lepton pair. We had no measurements of this mediator, though, and the weak force appeared to have a limited range. So we concluded that any such mediator must be heavy. We were ultimately able to observe and measure the

⁵Appendix B reviews the mathematics of introducing the electromagnetic term to the Dirac equation in QED.

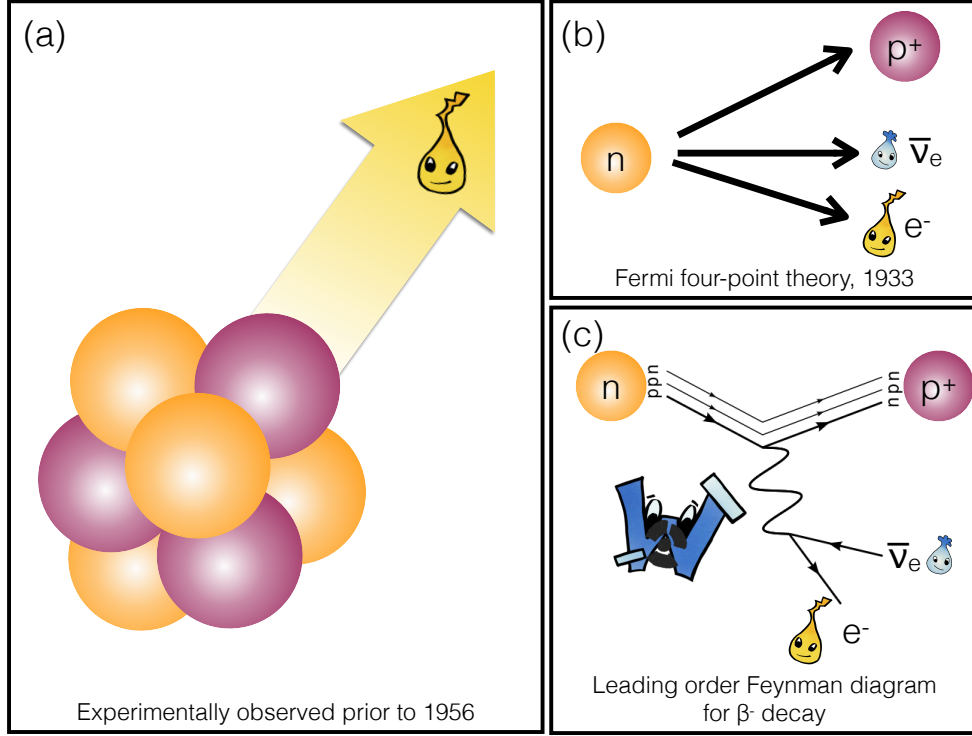


Figure 3: Illustrations of our historical understanding of the beta decay process. (a) Initially, we could only detect the nucleus and its emitted electron. The emitted electron's momentum measurements appeared to violate conservation laws, so Wolfgang Pauli and Enrico Fermi predicted another light particle called the neutrino. George Cowan and Fred Reines experimentally discovered the neutrino in 1956. Chien-Shiung Wu used the beta decay of cobalt-60 to demonstrate that the weak interaction violated parity in 1957. (b) Enrico Fermi's four-point interaction explained beta decay well, but its cross section grew unbounded with the square of the interaction energy. (c) We now understand that a valence down quark within a neutron decays to an up quark, through emission of a W^- boson. The valence quark identity of the new hadron is up-down-up, that of a proton. The W^- boson's decay produces the electron-antineutrino pair.

Source: HH (2020)

W^- , W^+ , and Z bosons and their masses with the Super Proton Synchrotron.

We could also consider this story for the theoretical perspective. Physicists in the 1800s unified the electric and magnetic forces into a single electromagnetic force. Over the next century, many theorists updated Maxwell's classical electromagnetic theory to quantum electrodynamics (QED) to explain behavior at tiny scales. QED was the first quantum field theory built on a **local gauge symmetry**. The local gauge symmetry cast the photon as messenger of the electromagnetic force between interacting charged particles. The theorists developed

renormalization to address cross sections growing unbounded as energy increased (and eventually violating the unitarity limit). Without renormalization, the electron's self energy and the electron and photon fields' vacuum zero-point energy would inevitably become infinite at higher orders in perturbation theory. Renormalization introduced counter-terms to the Lagrangian to convert the bare parameters like mass and electric charge to their finite experimental values.

Meanwhile, Enrico Fermi could explain beta decay using a **four-fermion interaction**. If he postulated that all four fermions - neutron, proton, electron, and anti-electron neutrino - met and interacted directly at a single point in space-time, then he could predict the behavior of beta decay. Later, theorists attempted to generalize his theory to energies above the observed weak scale (approximately the mass of the W boson) and found divergences to infinity, and they could not be removed with renormalization techniques.⁶ Chen Ning Yang and Robert L Mills took inspiration from the local gauge symmetry of QED to establish the **Yang-Mills Theory** in 1954 [3]. They turned to more complicated gauge symmetry groups to provide three force-carrying bosons with the electric charges of -1, 0, and +1, the roots of the SU(2) group previously discussed. Sadly, their bosons were still massless and adding mass terms for them *ad hoc* poisoned the theory's renormalizability.

In the summer of 1964, three different theory groups⁷ published papers in *Physical Review Letters* using Yoichiro Nambu's idea of **spontaneous symmetry breaking, SSB**, [4, 5] to give masses to the vector bosons of the Yang-Mills Theory. By the end of the 1960s, their ideas matured into the **Brout-Englert-Higgs (BEH) mechanism** [6–11].

When SSB occurs, a symmetric system falls into an asymmetric state. Tom Whyntie provides an intuitive example of SSB (figure 4) in *Introducing Particle Physics: A Graphic Guide* [12]. We can also think of SSB mathematically as interrupting the symmetry when

⁶Fermi's theory of the weak interaction is in other words an **effective field theory** in agreement with the true theory up to a given energy scale. Above that scale, we must use the true theory of electroweak interactions to obtain reasonable predictions.

⁷(1) Peter Higgs, (2) Robert Brout and François Englert, (3) Gerald Guralnik, Carl Hagen, and Tom Kibble.



Think of a marble in a wine bottle. If the marble is at the top of the “frog” (the rounded glass bump at the bottom), then the bottle and the marble will look the same from any direction (technically speaking, it exhibits a rotational symmetry about the vertical axis).

However, the system is unstable – the marble will spontaneously roll down into the trough running around the edge, breaking the symmetry.

The system has an underlying symmetry that is hidden – broken by the impracticalities of Nature (in this case, just how tricky it is to balance a marble on a rounded bump).

Figure 4: Whyntie’s example of spontaneous symmetry breaking. A marble positioned on the raised frog of a wine glass must eventually fall and settle around the edge [12].

choosing a ground state (or basis) for our calculations. SSB implies that parameters we *thought* were independent are in fact related. The unified electroweak force of the SM spontaneously breaks from $SU(2) \times U(1)$ to a new symmetry $SU(2)_L \times U(1)_Y$, separating the electromagnetic and weak forces in the process.⁸ To accomplish **electroweak symmetry breaking**, we have to add one new, scalar field (and its associated boson) to the SM Lagrangian.

Meanwhile, Hideki Yukawa had proposed an interaction mediated by a scalar boson to describe the nuclear force between fermions [13]. While quantum chromodynamics (QCD, home of the $SU(3)$ symmetry in the SM) eventually replaced his interaction for nucleons mediated by pions, his coupling for fermions to scalars found new life in assigning mass to leptons and quarks (the **Yukawa mechanism**). Each charged fermion must have its own coupling to the scalar field, since they each have different masses. While there is no reason *a priori* to assume that one scalar field accomplishes both electroweak symmetry breaking and the Yukawa mechanism, theorists added only one field - the Higgs field - to keep things simple. Altogether, the SM has essentially three sets of terms:

$$\mathcal{L}_{SM} = \mathcal{L}_{kin} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \quad (1)$$

The kinetic terms produce our dynamics through the universal gauge interactions. The Yukawa terms account for flavor, providing both the fermion masses and their mixing. The Higgs term's potential looks like:

$$V(\phi) = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^4. \quad (2)$$

⁸The subscript Y indicates the weak hypercharge. The electric charge q can be expressed as a sum of the of third component of the weak isospin, I , and the hypercharge: $q = I_3 + \frac{Y}{2}$. Additionally, we have experimentally observed that the weak force violates parity symmetry. Fermions divide into lefthanded and righthanded groups under the mathematical Poincaré group. Only lefthanded fermions transform under $SU(2)_L$ (L stands for left). In the SM, the righthanded fermions form a singlet, invariant under weak isospin transformations and there are no right-handed neutrinos.

When $\mu^2 < 0$, as in the SM, we have multiple minima such that

$$\phi^* \phi = \frac{-\mu^2}{2\lambda}, \quad (3)$$

forcing us to re-parameterize the theory. In other words, we must choose just one of four possible degrees of freedom as the minimum, called the **vacuum expectation value**, vev , and set the rest to 0, thereby provoking SSB. SSB produces a Nambu-Goldstone boson. The BEH mechanism replaces the Goldstone boson in the name of gauge invariance with a massive vector field and a real scalar field - the Higgs field. We can think of the weak bosons as absorbing the extra degrees of freedom to gain their masses. After electroweak symmetry breaking, we can express the masses of the weak and Higgs bosons in terms of the SU(2) gauge coupling, g , the U(1) gauge coupling g' , and the Higgs vev in the unitary gauge,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ vev + H(x) \end{pmatrix}. \quad (4)$$

The mass of the $W^\pm$ boson is

$$M_W = \frac{vev}{2} g. \quad (5)$$

The mass of the Z boson is

$$M_Z = \frac{vev}{2} \sqrt{g^2 + g'^2}. \quad (6)$$

The Higgs mass is

$$m_H = \sqrt{2\lambda(vev)^2}. \quad (7)$$

Each charged fermion has associated Yukawa Lagrangian terms. After electroweak symmetry breaking, we can write the fermion Yukawa Lagrangian for each charged fermion as

$$\mathcal{L}_{Yukawa} = -m_f \bar{\psi} \psi - \frac{m_f}{vev} \bar{\psi} \psi H, \quad (8)$$

where m_f is the mass of the fermion, ψ is the fermionic field, and H is the Higgs field in the unitary gauge.

Finally, every SM particle that should have a mass has one, either through electroweak symmetry breaking or a Yukawa coupling. We can also list the free parameters of the SM:

- Fine structure constant, $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}$ (where e is the elementary charge of the electron, ϵ_0 is the permittivity of free space, the $\hbar$ is Plank's constant, and c is the speed of light), which sets the coupling strength of electromagnetism;
- Weak mixing angle, $\theta_W = \arctan(g/g')$ (also called the Weinberg mixing angle);
- Strong coupling constant, $\alpha_3(M_Z)$;
- Higgs vev ;
- Higgs mass, m_H (or the Higgs potential coupling constant, λ);
- Four constants required to compute quark flavor changing currents (three mixing angles and the charge-parity violating phase of the Cabibbo-Kobayashi-Maskawa matrix, obtained after electroweak symmetry breaking when obtaining the quark mass states⁹);
- Nine charged fermion masses (six quarks and three charged leptons) to calculate the Yukawa coupling constants.

In all, we must experimentally determine these 18 parameters. Of those 18 parameters, 15 (indicated in red) have a relationship with the Higgs sector.

We need to experimentally determine at least two of the parameters vev , m_H , and λ to calculate all SM predictions in the Higgs sector. We can rearrange equation 5 to solve for the vev in terms of M_W and g . Substituting for g using its relationship with electric charge and the weak mixing angle, $g = \frac{e}{\sin\theta_W}$, gives

$$vev = \frac{2M_W \sin\theta_W}{e}. \quad (9)$$

⁹PDG Chapter 12, [1]

Alternatively, we can use the relationship between the vev and the reduced Fermi constant,

$$G_F^0 = \frac{\sqrt{2}}{8} \frac{g^2}{M_W^2 c^4},$$
$$vev = \left(\sqrt{2} G_F^0 \right)^{-1/2}. \quad (10)$$

Muon lifetime measurements allow us to extract G_F^0 and therefore the vev . The vev is about 246 GeV. We still need either m_H or λ , though.

In summary, the SM must have this scalar, charge-less, color-less boson and its field for three reasons:

1. The Higgs field keeps the SM calculable and renormalizable.
2. If we postulate that the electromagnetic and weak forces unify at a higher energy, we can add a scalar field to create masses for the weak bosons through electroweak symmetry breaking, changing $SU(2) \times U(1) \rightarrow SU(2)_L \times U(1)_Y$ and getting a Higgs boson.
3. Coupling fermionic fields to the Higgs field through the Yukawa mechanism gives masses to the quarks and charged leptons.

2 The value in measuring the Higgs boson's mass

The Higgs boson was discovered in 2012 following the first run of the Large Hadron Collider. The combined mass measured with both general purpose experiments, ATLAS and CMS, was 125 ± 0.24 GeV [14–17]. We have continued to measure the mass of the Higgs boson to improve our precision on this key constant of the SM [18–20]. The goal of this section is to contextualize the impact of improved measurements of the Higgs mass.

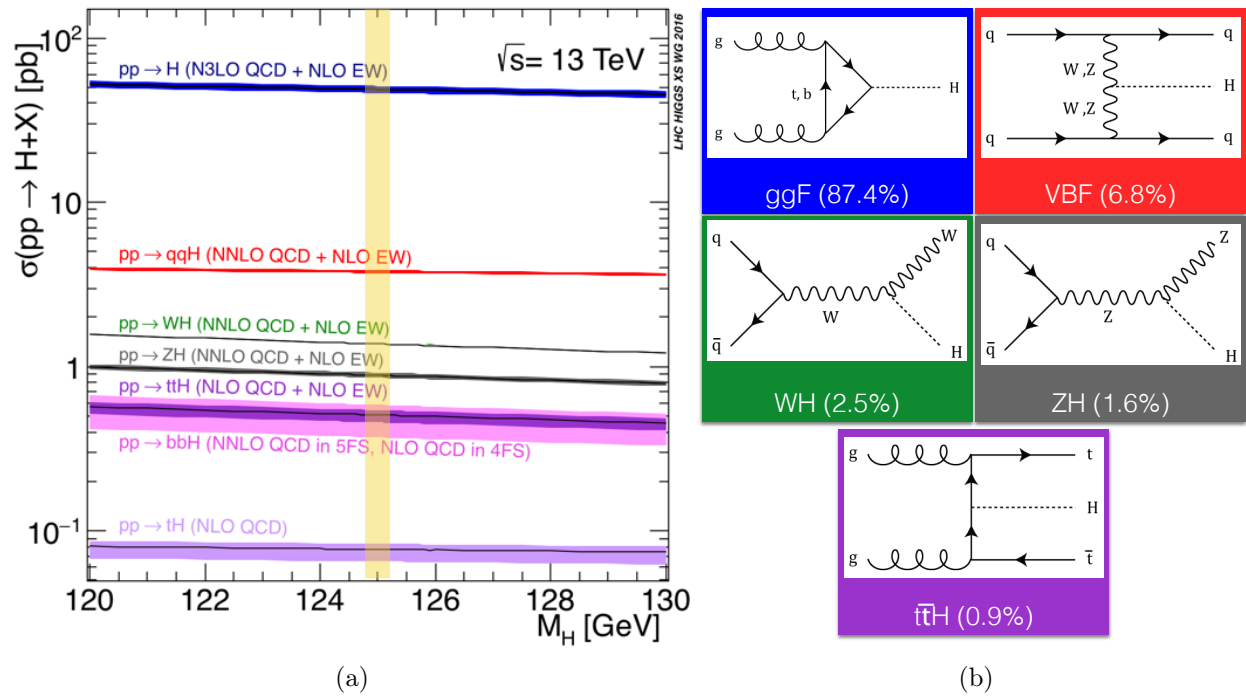


Figure 5: (a) Higgs boson production cross sections via the five leading mechanisms at the LHC as a function of m_H . (b) Feynman diagrams for the leading production mechanisms. The percentages indicate the fraction of expected events produced by that mode at the Large Hadron Collider. The background colors match the contours in (a) [21]. *Source: HH, adapted from the Handbook of LHC Cross sections: 4. Deciphering the nature of the Higgs sector*

Every mass-bearing particle of the SM interacts with the Higgs boson.¹⁰ The vev shapes the process of electroweak symmetry breaking, and therefore the vector boson couplings.

¹⁰Again, neutrinos must have masses, but we do not yet understand the mechanism by which they acquire mass.

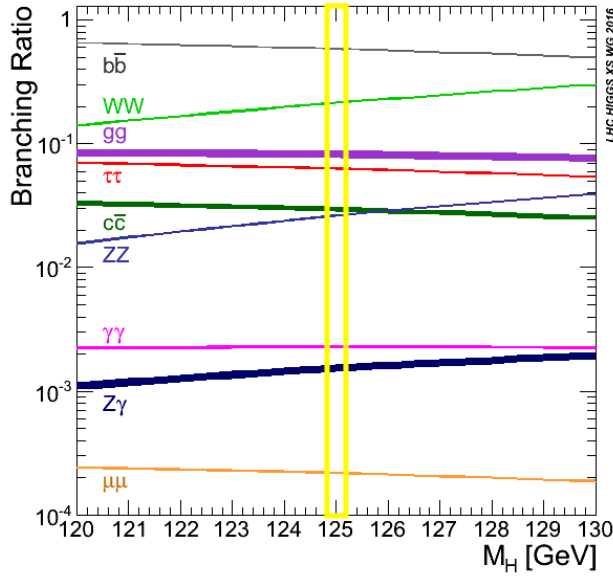


Figure 6: Branching ratios for the Higgs boson decay pathways as a function of m_H [21].
Source: *adapted from the Handbook of LHC Cross sections: 4. Deciphering the nature of the Higgs sector*

The Yukawa couplings are also proportional to the vev . The vev is proportional to the Higgs mass, so therefore all these couplings depend on the Higgs mass, meaning that **the Higgs mass appears in every Higgs production cross section or branching ratio prediction**. Figure 5 shows how the Higgs production cross sections change as a function of the Higgs mass. Figure 6 illustrates how the branching ratio depends on m_H . Around a mass of 125 GeV, the change in branching ratio for $H \rightarrow ZZ^{(*)}$ is greater than 1% per O(100 MeV) change in m_H .

The Higgs mass is a key parameter in defining the electroweak sector. It improves the precision of global electroweak fits to constrain possible processes beyond the SM, like those from the GFitter collaboration (figure 7) [22]. The stability of the electroweak vacuum also depends on m_H . In fact, theoretical predictions using the Run 1 ATLAS-CMS combination for Δm_H of 240 MeV indicate that the vacuum is meta-stable and near a phase transition (figure 8) [23].

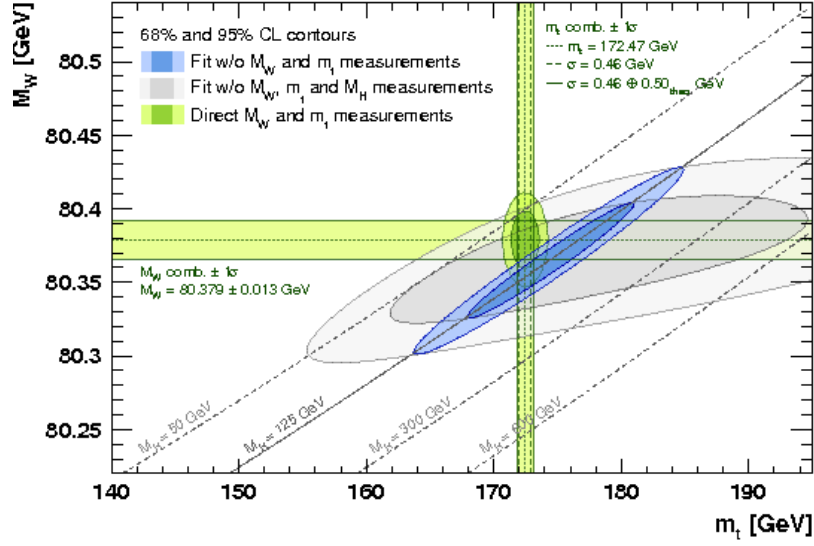


Figure 7: Contours at 68% and 95% CL obtained from scans of M_W versus M_t for the fit including (blue) and excluding the M_H measurement (grey), as compared to the direct measurements (green vertical and horizontal 1σ bands, and two-dimensional 1σ and 2σ ellipses). The direct measurements of M_W and M_t are excluded from the fits. *Source: Eur. Phys. J. C78, 675 (2018)*

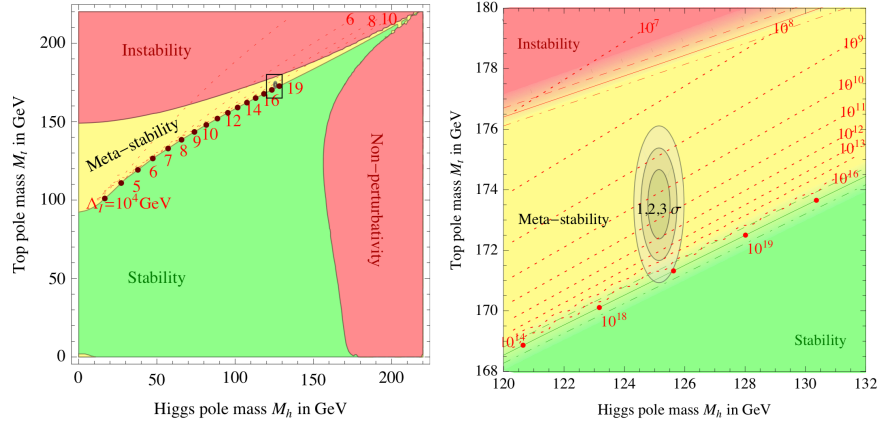


Figure 8: The Higgs mass is M_h in this figure. Left: SM phase diagram in terms of Higgs and top pole masses. The plane is divided into regions of absolute stability, meta-stability, instability of the SM vacuum, and non-perturbativity of the Higgs quartic coupling. The top Yukawa coupling becomes non-perturbative for top mass $M_t > 230$ GeV. The dotted contour-lines show the instability scale Λ_I in GeV assuming $\alpha_3(M_Z) = 0.1184$. Right: Zoom in the region of the preferred experimental range of M_h and M_t (the grey areas denote the allowed region at 1, 2, and 3 σ). The three boundary lines correspond to 1- σ variations of $\alpha_3(M_Z) = 0.1184 \pm 0.0007$, and the grading of the colors indicates the size of the theoretical error. These calculations used $\Delta M_h = 240$ MeV from the LHC Run 1 combinations. *Source: JHEP 12 (2013) 089*

Looking ahead

This dissertation describes a measurement of the Higgs boson mass in the four-lepton final state using all LHC Run 2 data collected by the ATLAS experiment. The LHC and the ATLAS detector are described in chapter II.

Chapter III comprises the bulk of the dissertation, describing this mass measurement in detail. Section 6 describes how we select relevant events for further analysis and construct their four-lepton invariant masses to produce our invariant mass spectrum. Once we have our invariant mass spectrum, section 7 describes how we build models to interpret it. The section will review a traditional approach to creating a probability density function (PDF) called the **template method** and then describes a **novel approach using per-event methods**. Section 8 describes the sources of systematic uncertainty and how we account for them. It also reviews how we construct the likelihood function and test statistic for measuring the Higgs mass.

Chapter IV reports the results. Section 9 summarizes our full expected Higgs mass result, including the impacts of systematic uncertainties. Section 10 introduces actual data collected with the ATLAS experiment and presents our measurement of the Higgs boson's mass as observed in the full Run 2 dataset.

Finally, chapter V considers the future of ATLAS and Higgs mass measurements.

Chapter II

Apparatus

Particle physics experiments began with the table-top-sized Wilson Cloud Chamber in 1911. Over time, we realized that the interconnectedness of nature’s structure required combining detector technologies to gather better snapshots of the evolution of particle creation, destruction, and transformation.

Historically, we designed detectors for a specific signature, like a kaon decay. These days, general purpose detectors surrounded our collision sources, layering different technologies to extract as much information from a collision as possible. General purpose experiments generate massive datasets encompassing many possible signatures: using one detector, we can explore thousands of ideas and opportunities, increasing our investigative power as a community and reducing our overall costs.

This chapter reviews the **Large Hadron Collider (LHC)** at CERN and one of its general purpose experiments, **ATLAS**.

3 Large Hadron Collider

The LHC is the most recent in a long line of particle accelerators based at the European particle physics laboratory, CERN. Built between 1998 and 2008 in the same tunnel that was home to the Large Electron-Positron Collider (LEP) under the Swiss-French border near Geneva, Switzerland, the LHC is currently the world’s largest and highest-energy collider, designed to accelerate protons to nearly the speed of light before colliding them in specific locations. A complete account of the machine is available in references [24] and [25]. The discovery of the SM Higgs boson was somewhat expected at the LHC, given the tight limits (figure 9) on the SM Higgs mass from experiments with LEP and the Tevatron and the physics reach of the machine.

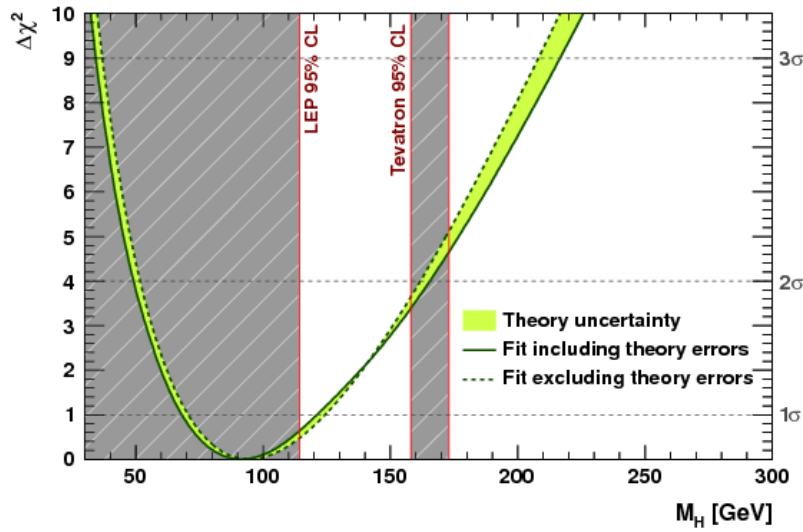


Figure 9: Indirect determination of the Higgs boson mass, excluding results of direct searches for the Higgs boson, from the Gfitter Collaboration in 2011. They used experimental results from LEP, Tevatron, and the early days of the LHC as inputs to their global electroweak fit. [26] It is interesting that LEP measurements had already excluded the central value of the theory-favored Higgs mass. A lighter Higgs boson would have made the electroweak vacuum more stable (figure 8). The LHC was designed to explore the remaining phase space, up to several TeV, ensuring it would find the Higgs boson or some other particle. *Source: Eur. Phys. J. C72 (2012) 2003*

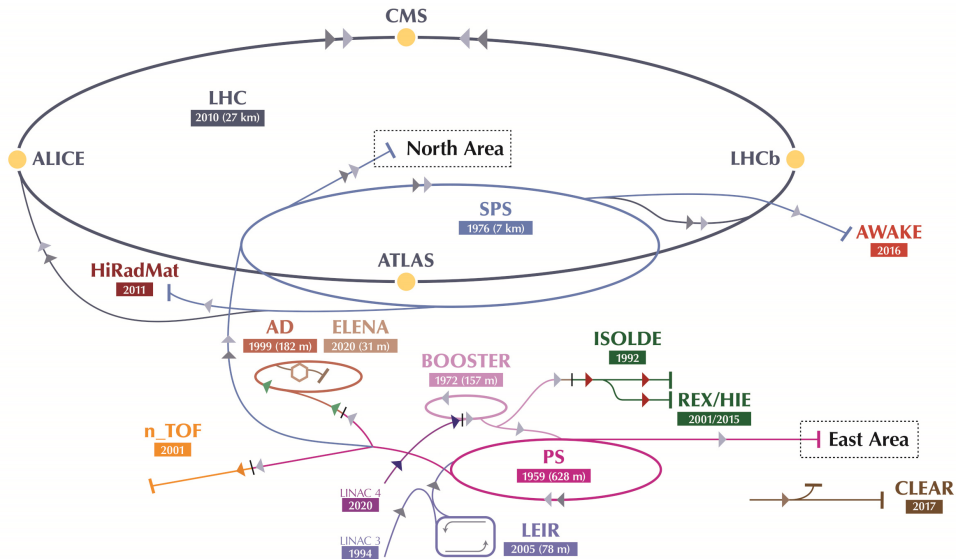


Figure 10: The accelerator complex at CERN. Each accelerator became a feeder for its successor so that we could achieve higher center-of-mass energies. The LHC may itself one day become a feeder for a larger collider. *Source: CERN, “Immersive tour of the accelerator complex,” online*

Figure 10 shows how we repurposed each historical collider in CERN’s accelerator complex to feed protons into its successor and eventually the LHC. A proton’s journey to an energy of 6.5 TeV and its eventual collision point begins in a **duoplasmatron**. It converts hydrogen gas into an ion beam. The device consists of a cathode filament, vacuum chamber, a minimum of two charged grids, and an aperture. The cathode emits free electrons into the small amount of hydrogen released in the vacuum chamber, forming a plasma of protons (otherwise known as ionized hydrogen atoms) and electrons. The charged grids generate a strong electric field to accelerate the plasma so that it leaves the aperture as a proton beam. Then, it enters the accelerator complex to begin its journey through CERN’s history and to 6.5 TeV.

The beam enters a **linear accelerator, linac**, first. LINAC2, built in 1978, provided this introduction to the complex during LHC Run 1 and Run 2. Its first acceleration stage uses a **radio frequency quadrupole** to focus the beam. It accelerated the beam’s constituent particles from 90 keV to 50 MeV over 33 m using its three-tank drift tubes wrapped in 120 quadrupoles. LINAC4 will replace LINAC2 during the LHC’s second long shutdown (2019-present) since it can accelerate protons to up to 160 MeV to meet the needs of the High Luminosity LHC (chapter V).

The beam leaves LINAC2 for the **PS Booster**, the first synchrotron in the complex originally built in 1972. It consists of four superimposed rings with a circumference of 157 m to raise the proton energy from 50 MeV to 1.4 GeV in 1.2 seconds. We then inject the 1.4-GeV beam into the oldest operating synchrotron at CERN, the **Protons Synchrotron, PS**, built in 1959 with a circumference of 628 m. Not all beams at CERN end up in the LHC. Some leave after LINAC2 or the PS Booster. The PS determines the energetic fate of the rest. It accelerates beams destined for the LHC to 26 GeV using a variety of radio frequency cavities, parceled out in **bunches** with high densities of protons. The PS prepares batches of 72 bunches, ultimately determining the LHC filling structure.

The PS then fills the **Super Proton Synchrotron, SPS**, with a circumference of

6.9 km.¹¹ It accelerates the beam from 26 GeV to 450 GeV with a nominal intensity of 1.5×10^{11} protons per bunch.

Finally, the beams are injected to the LHC in either a clockwise or counter-clockwise direction. The LHC holds two beams circulating in opposite directions. It accelerates them from 450 GeV to 6.5 TeV each, for a total energy in the center-of-mass frame of the collisions of 13 TeV. The LHC squeezes the two beams together, forcing them to collide at one of four different interaction points around its 27-km ring. The ATLAS Experiment, an eight-story detector, observes collisions at Point 1, where each crossed bunch typically produces 20 to 30 hard-scatter interactions.

4 The ATLAS Detector

The ATLAS detector (figure 11) consists of four nested layers: the Inner Detector (ID), the Electromagnetic Calorimeter (ECAL), the Hadronic Calorimeter (HCAL), and the Muon Spectrometer (MS). A full description of the ATLAS detector can be found in references [27, 28]. ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector, with the z -axis aligned with the beam pipe. The x -axis points from the IP to the center of the LHC ring, and the y -axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. Pseudorapidity is given by $\eta = -\ln \tan(\theta/2)$, where θ is the polar angle.

The **ID** is the main tracking detector, operating with acceptance $|\eta| < 2.5$ in a 2 T axial magnetic field. It consists of three subdetectors, sensitive to charged particles: four layers of silicon pixels ($5 < r < 12$ cm) and four layers of silicon strips ($30 < r < 51$ cm)

¹¹It is critical that each synchrotron's circumference exceeds its predecessor in CERN's complex. The circular collider tunnels typically host electron-positron colliders, where radius is the limiting factor, and then give homes to stronger dipole bending magnets for hadron colliders. Whenever we accelerate a charged particle perpendicular to its velocity vector, it will emit synchrotron radiation inversely proportional to its mass raised to the fourth power. Electrons have over 10^{13} times the energy loss to synchrotron radiation compared to protons because they are so much lighter. Tighter turns force greater energy losses so we must keep increasing our synchrotrons to increase the arc length of any given turn.

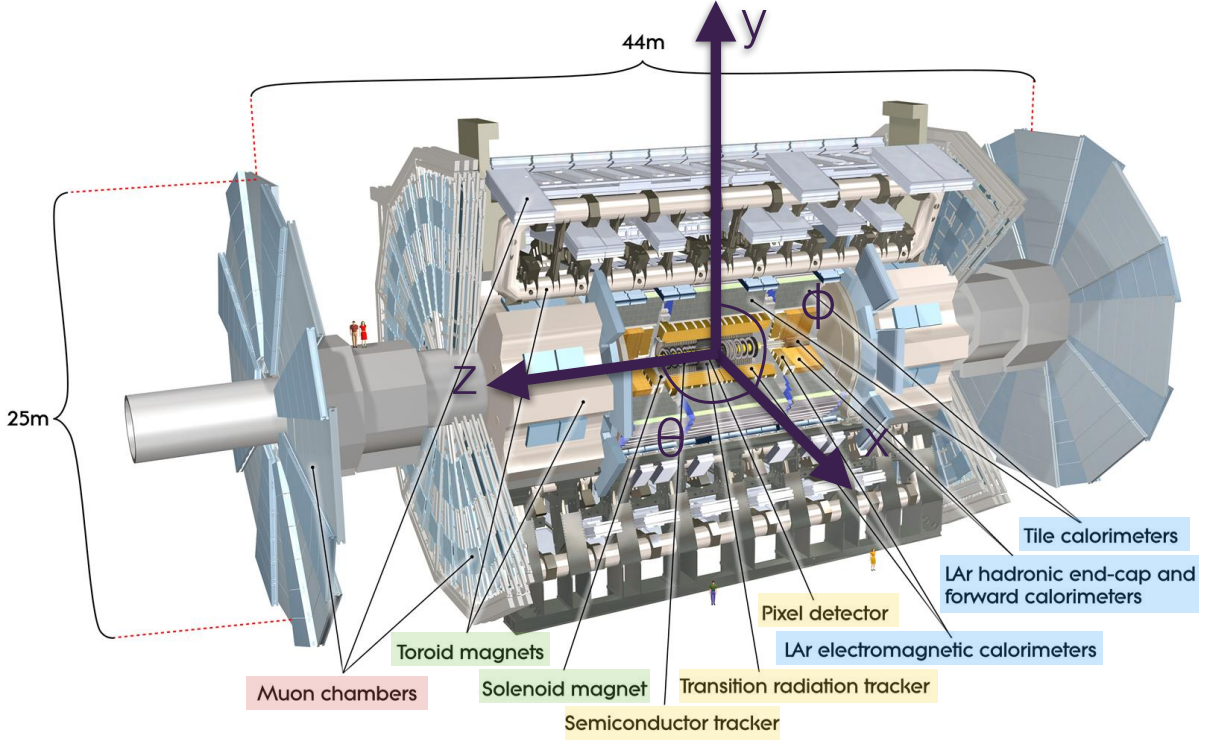


Figure 11: The major subdetectors of the ATLAS detector are the Inner Detector (labels highlighted in yellow), the Electromagnetic and Hadronic Calorimeters (blue), and the Muon Spectrometer (red). The magnet system (green), particularly the toroids, dominates ATLAS' appearance. The Trigger and Data Acquisition (TDAQ) system is integrated into the detector's components. The TDAQ system is effectively a sophisticated on-board computer determining which collisions are interesting enough for ATLAS to record for "offline" analysis. The TDAQ system received some maintenance and speed updates for Run 2. Section 6.1 provides more information on the trigger. The coordinate system is superimposed on the image in purple. *Source: adapted from the ATLAS Experiment "Detector and technology" webpages*

span the full ID acceptance, while 72 straw layers of transition radiation tracker modules ($55 < r < 108$ cm) have a pseudorapidity coverage $|\eta| < 2.0$. The innermost layer is a new addition to the ID for LHC Run 2. The "Insertable B-Layer" was designed to boost ATLAS' tracking efficiency at higher luminosities.

Two calorimeters span the distance between the ID and MS. The **ECAL** covers $|\eta| < 3.2$, employing lead as its absorber and liquid argon (LAr) as the active medium. Charged particles ionize the argon atoms, generating a current proportional to the energy deposit at a precise location. The **HCAL** uses two different technologies. Within $|\eta| < 1.7$, a

tile calorimeter detects hadrons using steel as the absorber and scintillator as the active medium. When an ionizing particle crosses a tile, it produces scintillation light in the polystyrene base material [27]. The light is then collected by photomultiplier tubes. A LAr calorimeter occupies the remaining HCAL acceptance, up to $|\eta| < 3.2$. Finally, LAr forward calorimeters extend the acceptance for both electromagnetic and hadronic measurements to $|\eta| = 4.9$.

The **MS** is the outermost ATLAS subdetector. This muon tracking detector provides independent momentum measurements of charged particles with a relative resolution better than 3% over a wide p_T range and up to 10% at $p_T \sim 1$ TeV. It operates in the pseudorapidity region $|\eta| < 2.7$. The MS consists of one barrel ($|\eta| < 1.05$) and two endcap sections. A system of three large superconducting air-core toroidal magnets provides a magnetic field of 0.5 T. Three layers of resistive plate chambers (RPCs) ($|\eta| < 1.05$) and three layers of thin gap chambers ($1.05 < |\eta| < 2.4$) serve as trigger chambers for the detector. Three layers, called stations, of monitored drift tube (MDTs) chambers ($|\eta| < 2.7$) provide precise muon momentum measurements. Each chamber provides six to eight η measurements along the muon track. The innermost layer of MDTs is replaced by cathode strip chambers in $|\eta| > 2.0$. During the shutdown preceding Run 2, the last missing chambers in the transition region between the barrel and endcaps ($1.0 < |\eta| < 1.4$) were added, completing the MS to its initial design specifications [29].

Figure 12 summarizes particle detection in ATLAS. The ID is sensitive to all charged particles. Electrons and photons are captured by the ECAL while hadrons like protons and neutrons leave deposits in the HCAL. Muons are primarily identified and reconstructed by combining information from the ID and MS detectors. As minimum-ionizing particles, they leave trace signals in the calorimeters. Neutrinos interact so weakly that they are not directly visible to ATLAS, but can be reconstructed as missing momentum.

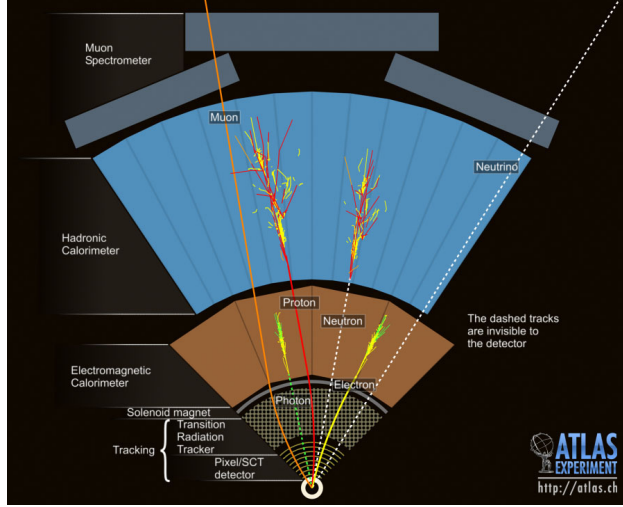


Figure 12: Illustration of the cross section of a sector of the ATLAS detector. Colored lines trace the paths of different particle species. Each nested layer of ATLAS contributes to particle identification and reconstruction.
Source: [CERN-GE-0803022](http://cern-ge-0803022)

5 Data collection

The LHC delivers proton-proton (pp) collisions of a set center-of-mass energy, $\sqrt{s}$, to ATLAS during the warmer months of the year when electricity in central Europe is cheaper. Figure 13 notes the total integrated luminosity delivered to ATLAS during each year of LHC operation. LHC Run 1 consisted of pp collisions at $\sqrt{s} = 7$ and 8 TeV in 2011 and 2012 for a total of 25 fb^{-1} . This dissertation primarily concerns LHC Run 2 from 2015 to 2018 at about twice the center-of-mass energy, $\sqrt{s} = 13$ TeV. ATLAS collected a total of 139 fb^{-1} of pp collisions during Run 2 with 25 ns spacing between bunch crossings. Figure 14 shows the profile for the mean number of interactions per bunch crossing, $\langle\mu\rangle$, during each data-taking period. The mean number $\langle\mu\rangle$ increased over the course of Run 2 as the LHC increased its instantaneous luminosity to provide many more events for analysis. It eventually reached twice its design luminosity by October 2018 for a peak instantaneous luminosity of $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. We only analyzed events if the ID, MS, and calorimeters were all operational at the time of collection. We also required that both magnet systems were active to ensure quality measurements of the momentum of charged particles.

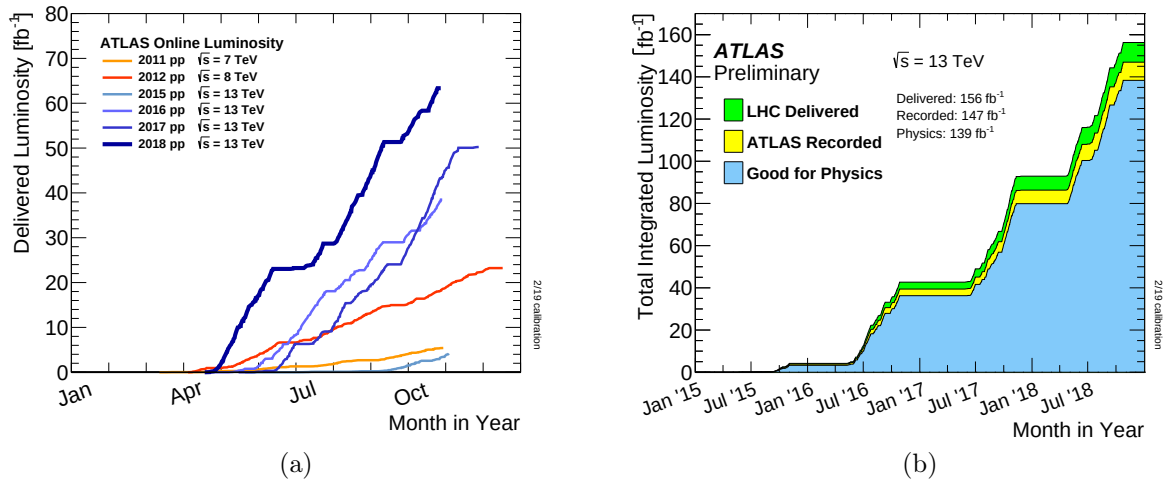


Figure 13: Total integrated luminosity delivered from the LHC and collected by ATLAS during (a) by year and (b) during Run 2. *Source: [ATLAS public luminosity results webpage - Multiple year collision plots](#)*

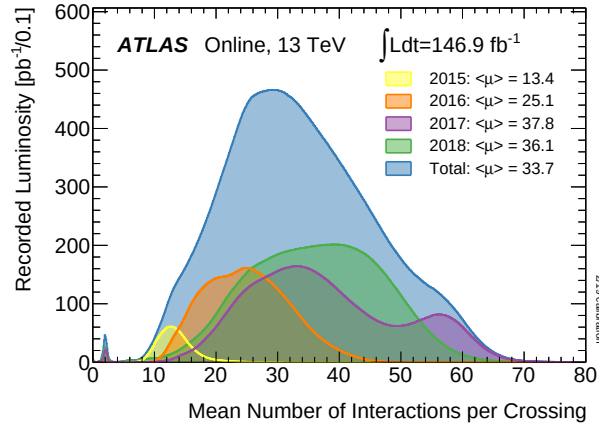


Figure 14: Number of interactions per bunch crossing (or **pileup**) in each year during Run 2. *Source: [ATLAS public luminosity results webpage - 2018 pileup interactions and data taking efficiency](#)*

.....

History suggests that the LHC will become a feeder ring for an even larger collider someday, and we will create more powerful detectors. In the meantime, we can measure the Higgs mass in the full LHC Run 2 dataset.

Chapter III

Experimental method

This chapter describes methods for measuring the Higgs boson's mass, m_H , in the four-lepton (4ℓ ; $\ell = e, \mu$) final state, defined as collision signatures with four muons (4μ), four electrons ($4e$), or two muons and two electrons (with a leading¹² dimuon pair, $2\mu 2e$, or a leading dielectron pair, $2e 2\mu$). The cross-flavor channels are treated separately since the quadruplet's invariant mass resolution depends on the identity of the leading dilepton pair. Third generation charged leptons, taus (τ), are excluded from this analysis because they cannot be fully reconstructed. The four-lepton final state has several advantages:

1. Combinations of electrons and muons give a precisely reconstructed final state (no missing energy). Their invariant mass spectrum provides a clean signal – a narrow peak on a smooth background.
2. Few processes at the LHC produce four leptons in the final state. Consequently,
 - (a) Statistics dominate the total uncertainty on m_H .
 - (b) This channel has few background processes and a good signal-to-background ratio.

For every background event, we see typically two signal events.

6 Event selection

This section describes how we extract candidate 4ℓ events from all collisions collected by ATLAS, using an **event selection** algorithm, and calculate the 4ℓ invariant mass. In ATLAS, all $H \rightarrow 4\ell$ analyses share a common event selection. The event selection's goal is to capture

¹²The **leading** dilepton pair is identified with the on-shell Z boson, and therefore is the dilepton pair with invariant mass closest to the Z pole mass. We require the invariant mass of this dilepton pair to fall between 50 and 106 GeV.

all possible events with a 4ℓ final state. It focuses, therefore, on maximizing acceptance and efficiency. This strategy drives our design choices for our event selection criteria, from our triggers to qualifying muons, electrons, and quadruplets of leptons.

6.1 Triggers

The LHC produces about 600 million collisions per second; it is impossible to record all of them. ATLAS is a **trigger-based** experiment, meaning that it features online and offline algorithms deciding which events to record based on their fastest-detected properties, like large energy deposits in the calorimeter or hits in the RPCs of the MS.

During Run 1, the ATLAS Trigger and Data Acquisition (TDAQ) system had three levels: (1) an online, hardware trigger (“Level 1 trigger”); (2) a software-based trigger (“Level 2 trigger”) capable of reading out data from the detector on the fly; and (3) another software-based trigger called the “Event Filter” composed of algorithmic code to isolate events for later analysis [27, 28]. The TDAQ system was fully updated for Run 2 [30]. The Level-1 Trigger received major hardware updates, raising the maximum Level-1 Trigger rate from 70 kHz in Run 1 to 100 kHz in Run 2. The Level-2 Trigger and Event Filter were merged to form a single, offline, software-based trigger, the “High level trigger (HLT).” The HLT is more flexible and has a faster average output rate of 1 kHz (up from 400 Hz in Run 1).

Normally, trigger design and selection is a major driver in analysis design, since only triggered events are recorded. As previously mentioned, comparatively few proton-proton collisions produce at least four leptons in the final state. Consequently, our emphasis on maximizing efficiency also motivates our trigger strategy, leading us to be expansive in our lists of triggers. We consider events collected using any un-prescaled high-level single-, double-, or triple-lepton trigger. Descriptions of these triggers by data-taking period are available in references [30–33]. Essentially, we are as inclusive as possible at the trigger level. Appendix C provides tables of all the selected triggers, organized by data collection period. Ultimately, our combined trigger efficiency is about 98% [34].

6.2 Lepton reconstruction and calibration

Table 4 provides criteria for the selection of muons and electrons. These criteria are very important for the mass measurement, since they determine our scale and resolution. Sections 6.3 and 6.4 discuss reconstruction and calibration for muons and electrons, respectively.

Table 4: Summary of the event selection requirements for muons and electrons. All $H \rightarrow 4\ell$ analyses in ATLAS use a common event selection, provided here. Jets, or showers of particles in our detector from strong particles, matter more for other measurements like differential cross sections [35] and Simplified Template Cross Sections [36], since they are typically features of the subleading production modes. They are not discussed here in the interest of brevity. The parameters d_0 and z_0 are the transverse and longitudinal impact parameters of tracks with respect to the center of the beam spot. The angle θ is defined in the ATLAS coordinate system, illustrated in figure 11. The distance ΔR lies in the $\eta - \phi$ plane, defined as $\Delta R = \sqrt{\eta^2 + \phi^2}$. The transverse momentum, p_T , is the component of a particle’s momentum perpendicular to the beamline, $p_T = p \sin \theta$. Transverse energy, E_T , is defined as $E_T = \sqrt{m^2 + p_T^2}$, where m is the particle’s invariant mass [36]. *Source: submitted to EPJC*

LEPTONS AND JETS	
MUONS	
Combined: $p_T > 5$ GeV and $ \eta < 2.7$	
Calorimeter-tagged: $p_T > 15$ GeV and $ \eta < 0.1$	
Segment-tagged: $p_T > 5$ GeV and $ \eta < 0.1$	
Stand-alone and silicon-associated forward: $p_T > 5$ GeV and $2.5 < \eta < 2.7$	
ELECTRONS	
$E_T > 7$ GeV and $ \eta < 2.47$	
Track, if present, must have a hit in the innermost pixel layer	
JETS	$p_T > 30$ GeV and $ \eta < 4.5$
IMPACT PARAMETER	- Muons: $ d_0 /\sigma(d_0) < 3$ (if ID track available)
SIGNIFICANCE	- Electrons: $ d_0 /\sigma(d_0) < 5$ (if ID track available)
LEPTON ISOLATION	- The amount of isolation E_T after summing the track-based and 40% of the calorimeter-based contribution must be smaller than 16% of the lepton p_T

6.3 Muon reconstruction and calibration

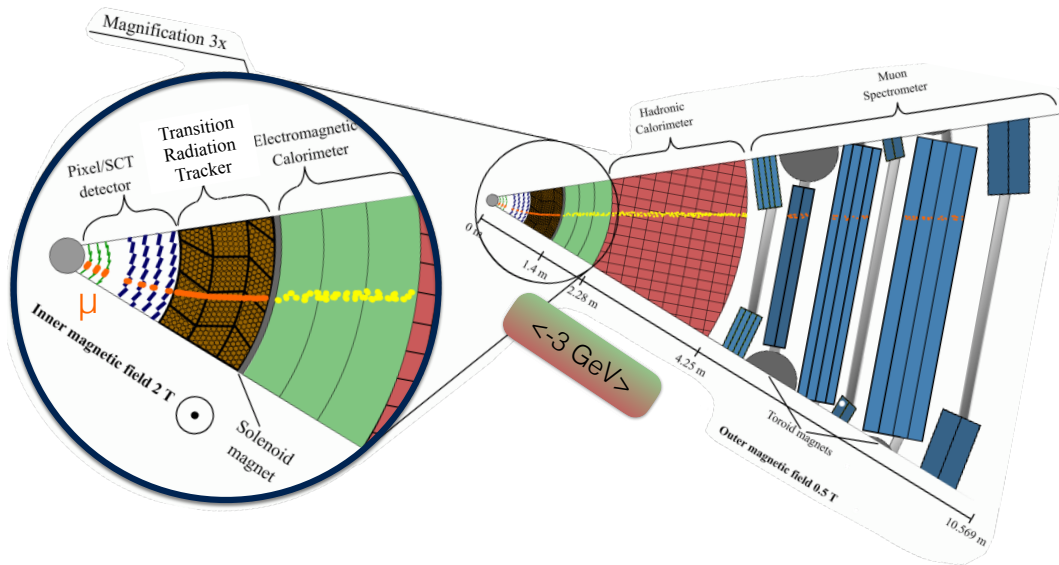


Figure 15: A muon's journey through ATLAS. Muons are minimum-ionizing particles, so they usually pass through all layers of the detector and lose about 3 GeV of energy traversing the calorimeters. *Source: HH (2019), adapted from T Herrmann, O Jařábek, K Jende, M Kobel*

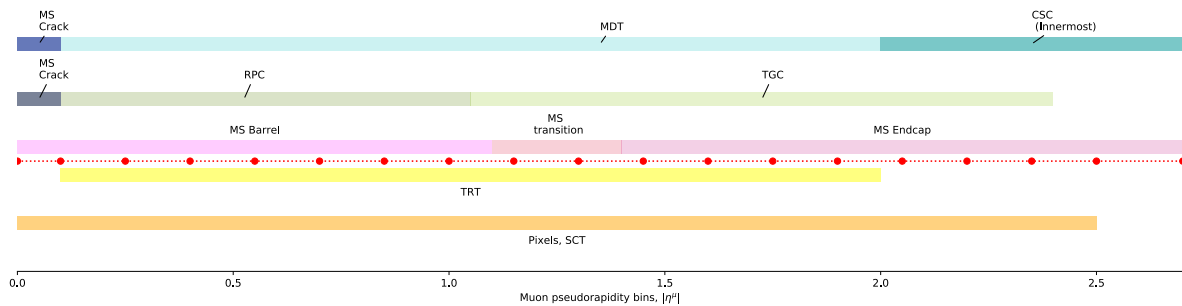


Figure 16: Coverage as a function of the absolute value of pseudorapidity, $|\eta|$, for subdetectors sensitive to muons. The red dots indicate the $|\eta|$ binning for the predicted individual muon response in the per-event response method (discussed in section 7.2.2) [18]. *Source: HH (2018)*

Muon performance dominates the measured Higgs mass, m_H^{measured} , precision, on account of their resolution and abundance. The 4μ channel is expected to be the most sensitive channel since it has the best resolution and most expected events (81 ± 5 out of 216 ± 13 expected signal events).

Figure 15 depicts a typical muon’s journey through the ATLAS detector, noting that, as a **minimum-ionizing particle**, it loses on average 3 GeV of energy in the calorimeters. We reconstruct muons based on the subdetectors they encounter as they traverse ATLAS, producing five “muon types” (section 6.3.1). Figure 16 provides the η coverage of each subdetector. We require selected muons to fall within $|\eta| < 2.7$ so that they are within the full MS volume.

6.3.1 Types of muons

Figure 17 is an (η, ϕ) map of the coverage of different muon types selected for the Run 2 $H \rightarrow 4\ell$ analyses. We use all five muon types:

1. **Combined (CB)** muons result from global refits to hits in the ID and MS. Track reconstruction occurs independently in the ID and MS. Most muons selected for any ATLAS analysis are combined. In $H \rightarrow 4\ell$, we select other muon types to ensure complete coverage.
2. **Muon spectrometer-extrapolated (ME)** muons only have MS tracks. $H \rightarrow 4\ell$ analyses only use these muons outside the ID coverage.
3. **Silicon-associated forward (SAF)** muons have an MS track and a “tracklet” in the ID. They lie on the very edge of the ID volume.
4. **Calorimeter-tagged (CT)** muons have a track in the ID matched to an energy deposit in the calorimeter consistent with a minimum-ionizing particle. Their identification criteria are optimized to cover the services access hole in the MS volume

($|\eta| < 0.1$) and the momentum range $15 < p_T^\mu < 100$ GeV, driving our choices for the CT-selected muons for $H \rightarrow 4\ell$ analyses (table 4).

5. **Segment-tagged (ST)** muons have an ID track matched to an MS local track segment in one layer of the precision chambers (MDT, CSC). ST and CT muons cover the acceptance in the MS services “crack.”

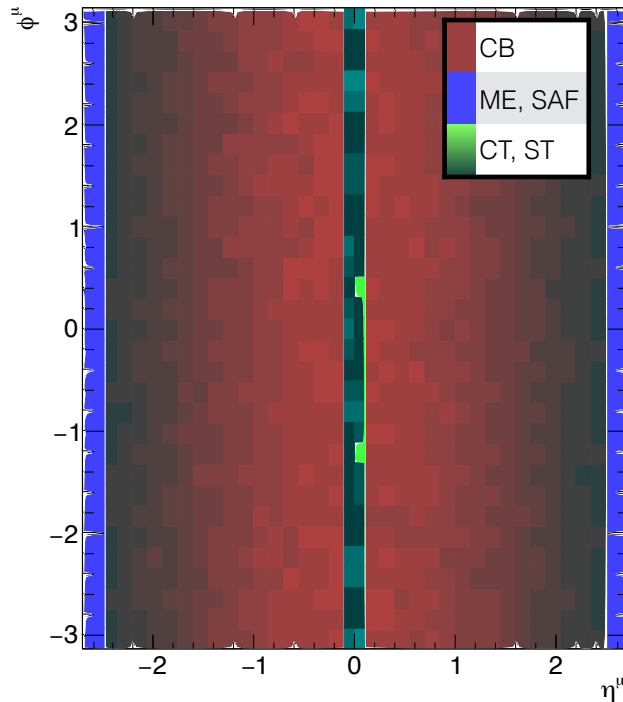


Figure 17: Map of the coverage of each muon type, as selected for Run 2 $H \rightarrow 4\ell$ analyses.
Source: HH (2020)

6.3.2 Highlights

Figure 18 displays two key metrics of muon performance. Figure 18a demonstrates that the ID measurement dominates the muon resolution at the relatively low energy scales relevant for $H \rightarrow 4\ell$ events. Further, we measure tracks better at low momentum since they bend more in the magnetic field. Muons have a better resolution at this momentum scale than electrons, which must rely on calorimeter measurements.¹³ Resolution effects (for example,

¹³Calorimeter resolution improves as incident particle energy increases.

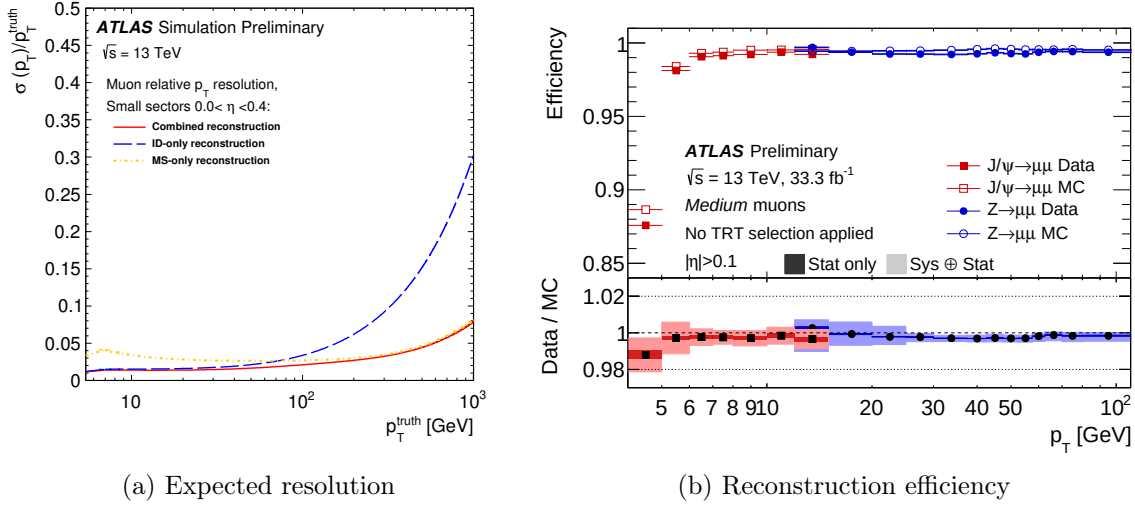


Figure 18: Expected resolution and reconstruction efficiency as a function of p_T^μ . (a) Expected transverse momentum resolution as function of the true transverse momentum (p_T^{truth}) from simulated muon candidates. The resolution is estimated from a Gaussian fit from the core of the relative true-to-measured q/p difference. The resolution of the combined muon momentum (labelled as “combined reconstruction”) is compared to that of the Inner Detector only (labelled as “ID-only reconstruction”) and to that of the Muon Spectrometer-only extrapolated to the interaction vertex (labelled as “MS-only reconstruction”), for muon candidates in $|\eta| < 0.4$ falling in small sectors of the Muon Spectrometer. (b) Reconstruction efficiencies for combined muons (Medium identification algorithm) measured in $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events as a function of the muon momentum. The prediction by the detector simulation is depicted as empty circles (squares), while the full circles (squares) indicate the observation in collision data for $Z \rightarrow \mu\mu$ ($J/\psi \rightarrow \mu\mu$) events. Only statistical errors are shown in the top panel. The bottom panel reports the efficiency scale factors. The darker error bands indicate the statistical uncertainty, while the lighter bands indicate the quadratic sum of statistical and systematic uncertainties. *Sources:* [MUON-2017-01](#), [MUON-2018-003](#)

from multiple scattering or detector misalignments) smear the reconstructed $m_{4\ell}$ shape, and therefore the Higgs resonance’s shape. Energy losses and charge-dependent alignment effects impact the reconstructed muon scale, directly impacting the m_H central value. Reference [37] describes muon reconstruction and calibration for Run 2 in detail.

Figure 18b shows the muon reconstruction efficiency for combined muons. We calibrate muon momentum using combined tracks (see section 6.3), for which we have a high reconstruction efficiency for $p_T^\mu > 5$ GeV – justifying our selection criterion for $p_T^\mu > 5$ GeV in table 4.

6.3.3 Calibrating muon momentum

MC simulation with GEANT4 [38] and realistic ATLAS geometry, resolution, and alignment allow us to model interactions within the detector. Simulation, however, is limited. While more detail is essential to achieving our precision, we must choose between detail and reasonable computing expenses. For example, we do not model the pull of gravity on each MDT's wire. ATLAS is also a dynamic experiment, living in the cavern, enduring radiation damage and repair interventions. Rather than relying on a “perfect” simulation to guide us, we account for the small discrepancies (for example, from imperfect detector material descriptions or residual misalignments in the ID and MS) between simulation and observed data using a calibration procedure.

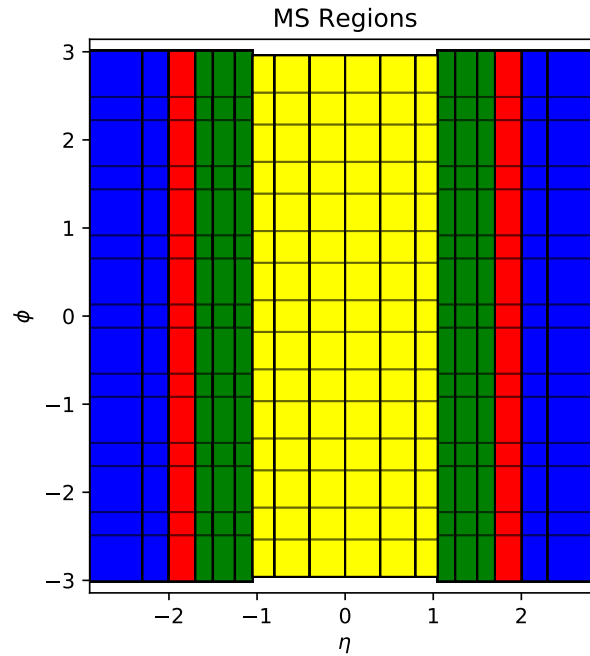


Figure 19: We extract the correction constants using simultaneous unbinned maximum likelihood fits to J/ψ and Z candidates in (p_T^μ, η^μ) bins in data and simulation. We apply the corrections to simulation before we build $m_{4\ell}$. This map shows the η bins, or regions, in the Muon Spectrometer. The yellow area represents the barrel. The green represents the transition region between the barrel and endcap, including the region where the toriodal field changes sign. The red and blue areas together signify the endcap. The blue area is instrumented with cathode strips instead of MDTs. The MS is further subdivided along ϕ into small and large sectors, indicated with the black lines. The ID map is only divided in η since the ID is hermetic in ϕ . *Source: HH (2018)*

The momentum calibration procedure seeks to adjust the shift (through scale corrections) and shape (through smearing corrections) in MC to better reflect the behavior in data, ultimately producing an accurate description of the muon momentum scale and resolution. We derive the corrections in ID and MS regions (figure 19) as a function of p_T using simultaneous unbinned maximum likelihood fits to $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ candidates. We require that both muons in an opposite-sign pair used during the calibration are combined muons. The correction procedure assumes that the relative transverse momentum resolution of a muon in ATLAS can be parameterized as,

$$\sigma(p_T^\mu) = \frac{r_0}{p_T^\mu} \oplus r_1 \oplus r_2 \cdot p_T^\mu, \quad (11)$$

where r_i are the resolution parameters. The calibration itself re-parameterizes the combined momentum as,

$$p_T^{\text{Cor,Det}} = \frac{p_T^{\text{MC,Det}} + \sum_{n=0}^1 \Delta s_n^{\text{Det}}(\eta, \phi) \left(p_T^{\text{MC,Det}}\right)^n}{1 + \sum_{m=0}^2 \Delta r_m^{\text{Det}}(\eta, \phi) \left(p_T^{\text{MC,Det}}\right)^{m-1} g_m}, \quad (12)$$

with correction parameters, Δs_i and Δr_i , modifying the reconstruction's scale and resolution parameters. The term g_m indicates a Gaussian distribution. Figure 20 labels the correction parameters with their roles. Generally, constant parameters account for energy losses, linear parameters adjust for multiple scattering and magnetic field effects, and quadratic terms handle intrinsic resolution effects and misalignments.

The ID and MS tracks are each extrapolated back to the interaction point before deriving the correction. The MS track also includes a parameterized correction for energy losses within the calorimeter volume, typically 3 GeV. Finally, we also assume that this procedure does not change the relative fraction, f , of the relative subdetector contributions to the total combined momentum. The final calibrated combined momentum is

$$p_T^{\text{Cor,CB}} = f \cdot p_T^{\text{Cor,ID}} + (1 - f) \cdot p_T^{\text{Cor,MS}}, \quad (13)$$

We can also derive corrections for charge-dependent alignment effects, described in section 4 of reference [18].

6.3.4 Muon momentum calibration performance

The J/ψ and Z resonances serve as calibration sources at two different momentum scales. Figure 21 compares the J/ψ and Z resonant distributions before and after applying the corrections to simulated data. Figures 22 and 23 consider the dimuon scale and resolution, respectively, as a function of η , where η serves as a proxy for detector region. Figure 24 displays the relative momentum resolution, highlighting the excellent momentum resolution at lower momenta.

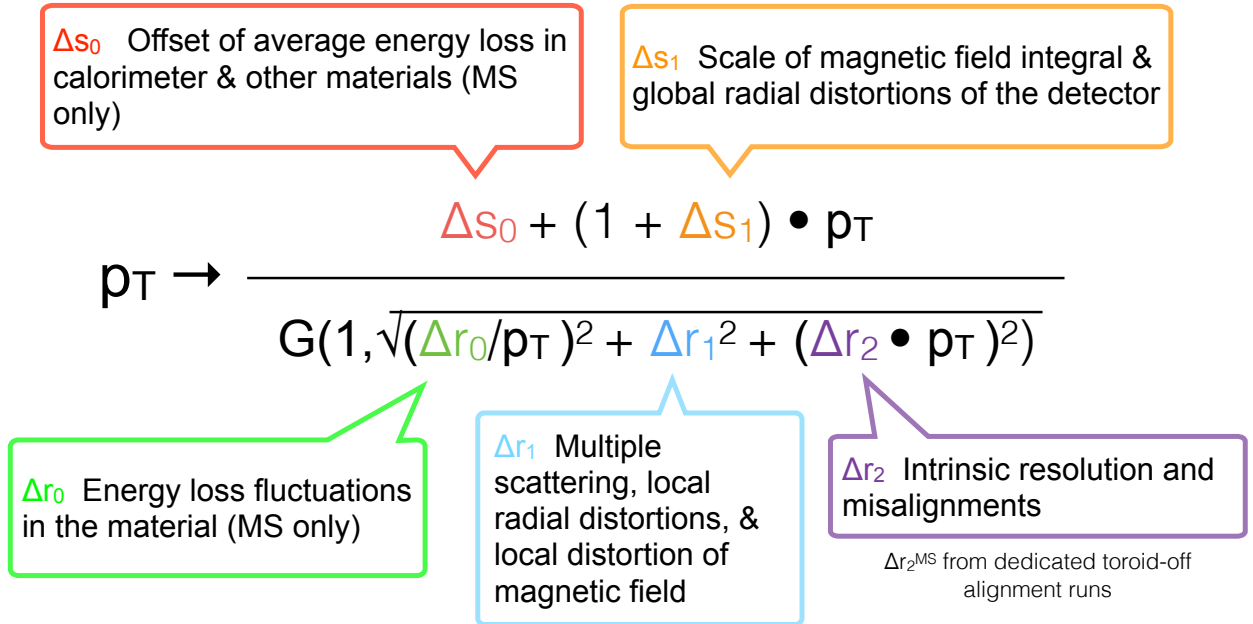
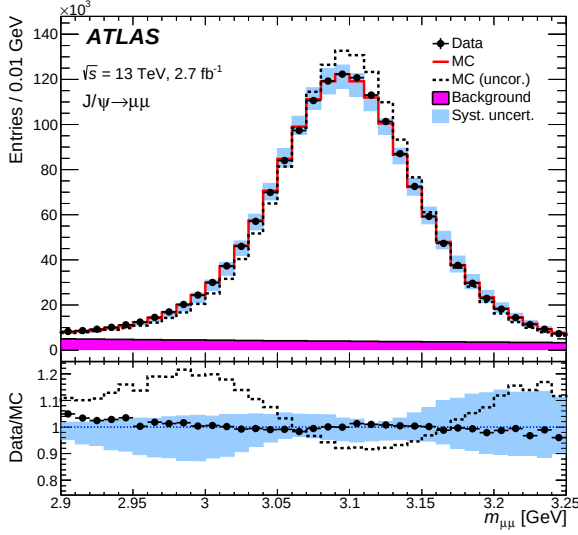
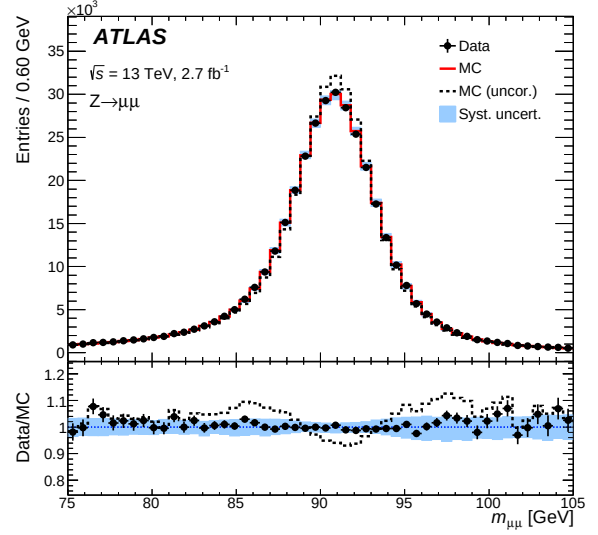


Figure 20: Each calibration parameter ($\Delta s_i, \Delta r_i$) addresses specific detector effects. Δs_i modify the momentum scale. The Δr_i parameters apply Gaussian smearing (indicated with G) to the momentum resolution. *Source: HH (2015)*

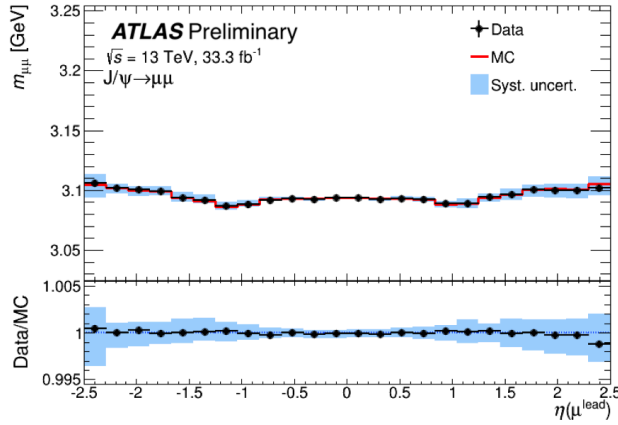


(a) $J/\psi \rightarrow \mu\mu$

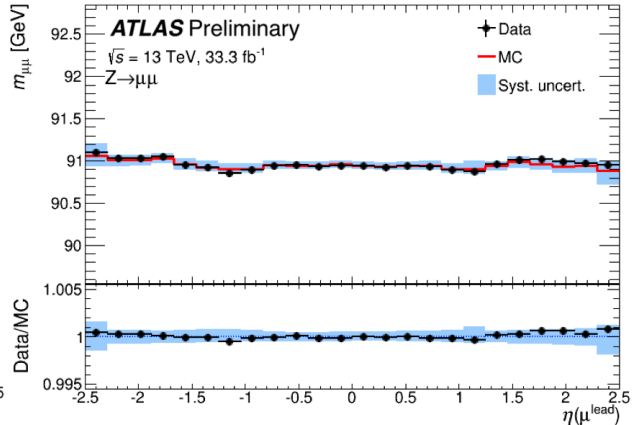


(b) $Z \rightarrow \mu\mu$

Figure 21: Data-simulation comparison before (dotted black line) and after (solid red line) applying muon momentum calibration corrections to simulation. Source: *Eur. Phys. J. C* 76 (2016) 292



(a) $J/\psi \rightarrow \mu\mu$



(b) $Z \rightarrow \mu\mu$

Figure 22: Dimuon invariant mass scale in data (black dots) and corrected MC simulation (solid red line) as a function of pseudorapidity, η , for the J/ψ and Z resonances. Source: *MUON-2017-01*

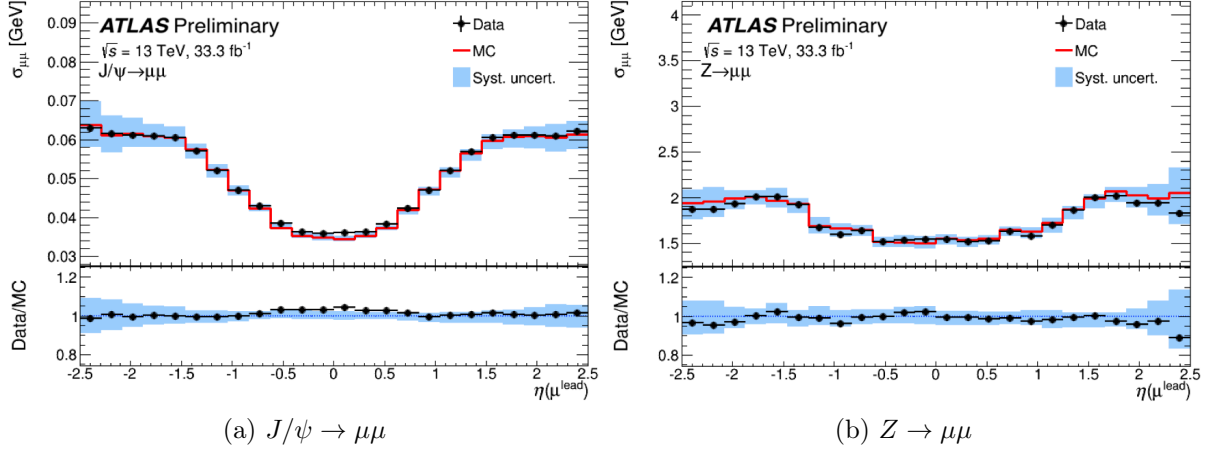


Figure 23: Dimuon invariant mass resolution in data (black dots) and corrected MC simulation (solid red line) as a function of pseudorapidity, η , for the J/ψ and Z resonances. *Source: MUON-2017-01*

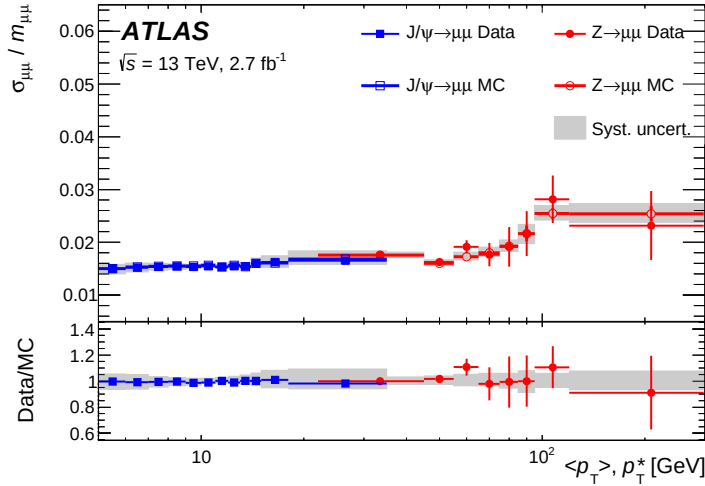


Figure 24: Relative dimuon momentum resolution divided by the dimuon invariant mass for CB muons measured from $J\psi \rightarrow \mu\mu$ ($Z \rightarrow \mu\mu$) events as a function of the average transverse momentum variable $\langle p_T \rangle = \frac{1}{2} (p_T^1 + p_T^2)$ ($p_T^* = m_Z \sqrt{\frac{\sin \theta_1 \sin \theta_2}{2(1 - \cos \alpha_{12})}}$). Both muons are required to be in the same $|\eta|$ range. The error bars represent the statistical uncertainty while the bands show the systematic uncertainties. *Source: Eur. Phys. J. C 76 (2016) 292*

6.4 Electron reconstruction and calibration

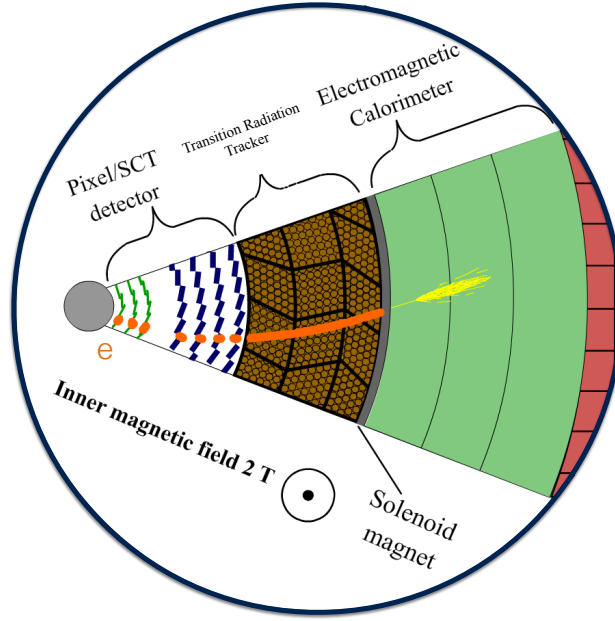


Figure 25: An electron’s journey through ATLAS. The electromagnetic calorimeter absorbs the electron’s energy, so its journey ends there. *Source: HH (2019), adapted from T Herrmann, O Jařábek, K Jende, M Kobel*

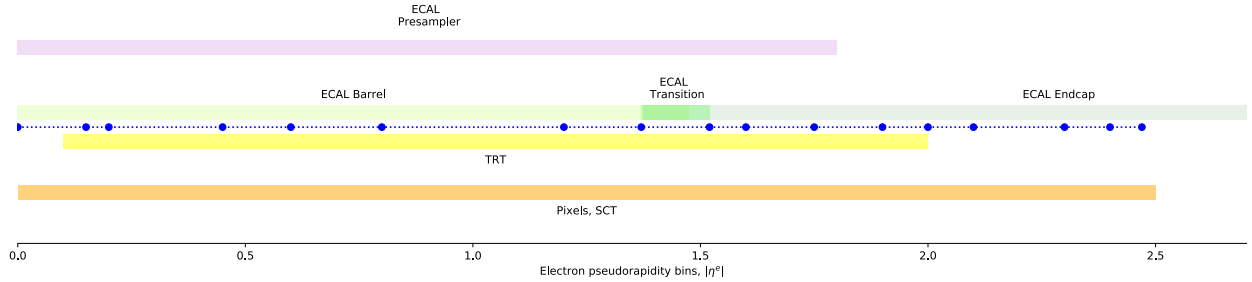


Figure 26: Coverage as a function of the absolute value of pseudorapidity, $|\eta|$, for subdetectors sensitive to electrons. The blue dots indicate the $|\eta|$ binning for the predicted individual electron response in the per-event response method (discussed in section 7.2.2) [18]. *Source: HH (2018)*

Figure 25 depicts a typical electron’s journey through the ATLAS detector, noting that the electromagnetic calorimeter (ECAL) fully absorbs it. Figure 26 provides the η coverage of each subdetector. We require selected electrons to fall within $|\eta| < 2.47$ so that their tracks fall within the full ID volume and are matched successfully with deposits in the ECAL.

6.4.1 Highlights

Prior to 2017, electrons were reconstructed using a fixed-size cluster (**sliding window**) algorithm to group energy deposits in the ECAL [39]. Full Run 2 measurements use the new dynamic topological cell clustering approach to electron reconstruction, described at length in reference [40]. This **supercluster**-based energy calibration (figure 27) mimics an electromagnetic shower’s development in the calorimeter itself. With this approach, we can reconstruct electromagnetic showers down to the order of 100 MeV. The algorithm’s adaptive nature allows it to better recover low-energy bremsstrahlung radiation and reduce lateral leakage energy, so the final corrections are smaller overall. When the lateral leakage is constant, the energy response is also more linear. Finally, this method has better resolution (figure 28) and is more robust to pileup than the sliding window method (figure 29). The difference in energy scale between data and simulation is defined as α_i , where i corresponds to different regions in η . Energy resolution mismodeling is parameterized as an η -dependent additional constant term, c_i . The electron energy calibration assumes that the energy scale correction applied to data takes the form $E^{\text{data,corr}} = \frac{E^{\text{data}}}{1+\alpha_i}$. The calibration also assumes that the resolution correction applied to simulation takes the form $\left(\frac{\sigma_E}{E}\right)^{\text{MC,corr}} = \left(\frac{\sigma_E}{E}\right)^{\text{MC}} \oplus c_i$. Figure 30 displays the α_i and c_i parameters as a function of η for the supercluster method. Figure 31a displays the shape of the calibrated Z resonance in the dielectron invariant mass distribution. Figure 31b explains the impact of pileup on the dielectron scale. The importance of robust responses in the face of mounting pileup increased in the later half of Run 2 (recall figure 14).

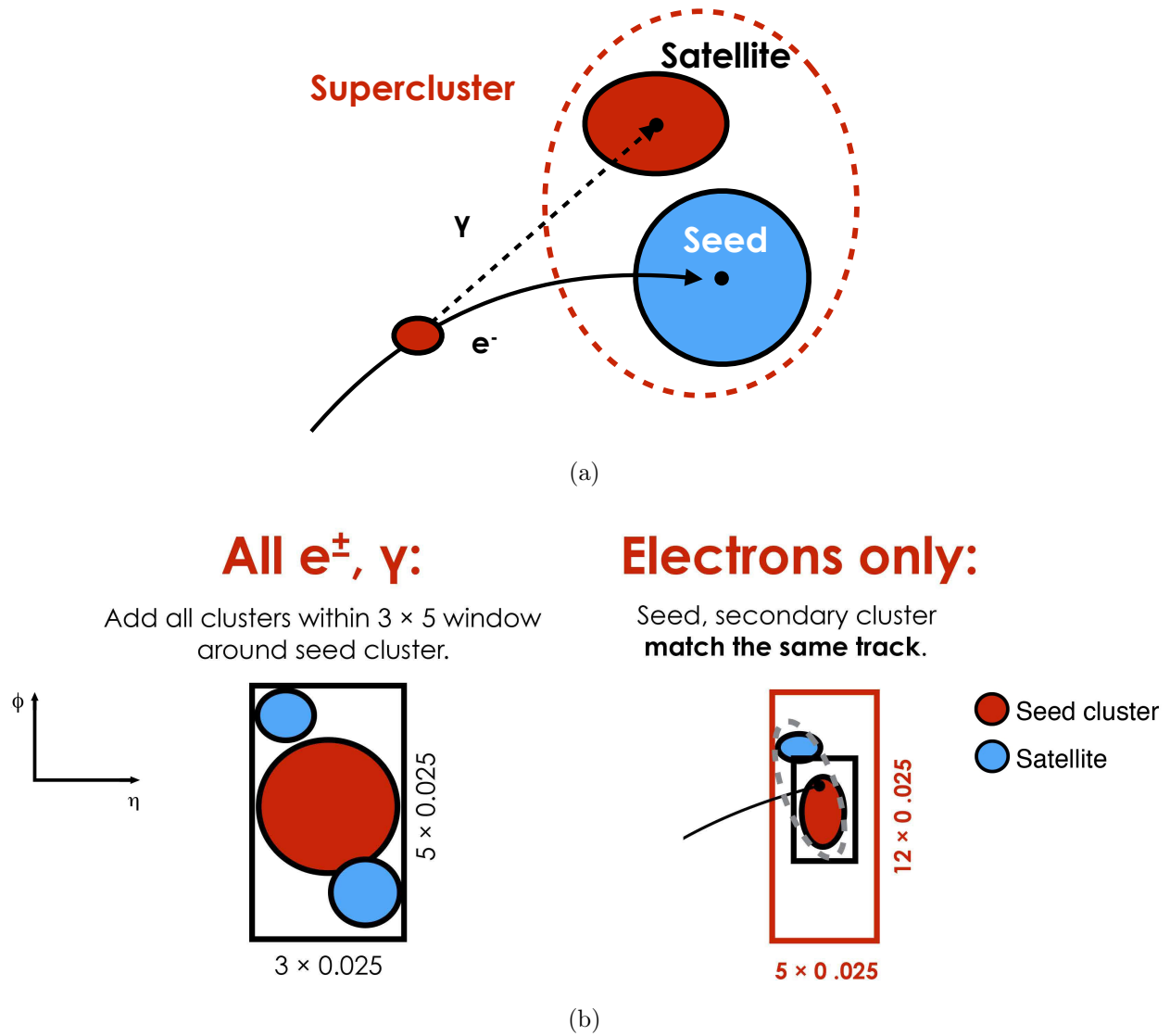


Figure 27: (a) Schematic definition of a supercluster, intended to emulate Bremsstrahlung. (b) Diagram of the superclustering algorithm implementation for electrons and photons. Seed clusters are in red while satellite clusters are in blue. *Source: [ATL-PHYS-PUB-2017-022](#), JINST 14 (2019) P12006*

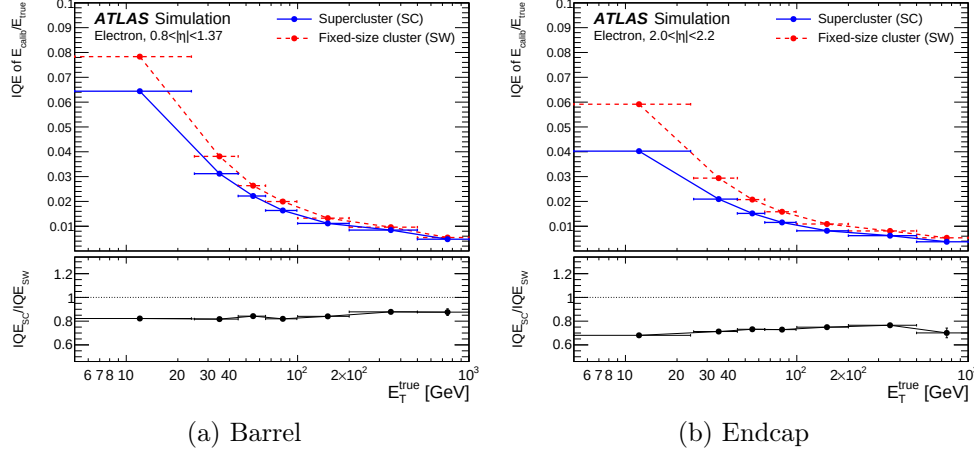


Figure 28: Calibrated single-electron energy response resolution, expressed in terms of the middle 50% of the resolution distribution (interquartile range IQE), when simulated without pileup ($\langle\mu\rangle = 0$). The resolution estimate improves in both the (a) barrel and (b) endcap with the superclustering method. The ratio plot compares the IQE of the superclustering method and its predecessor, the sliding window method. *Source: JINST 14 (2019) P12006*

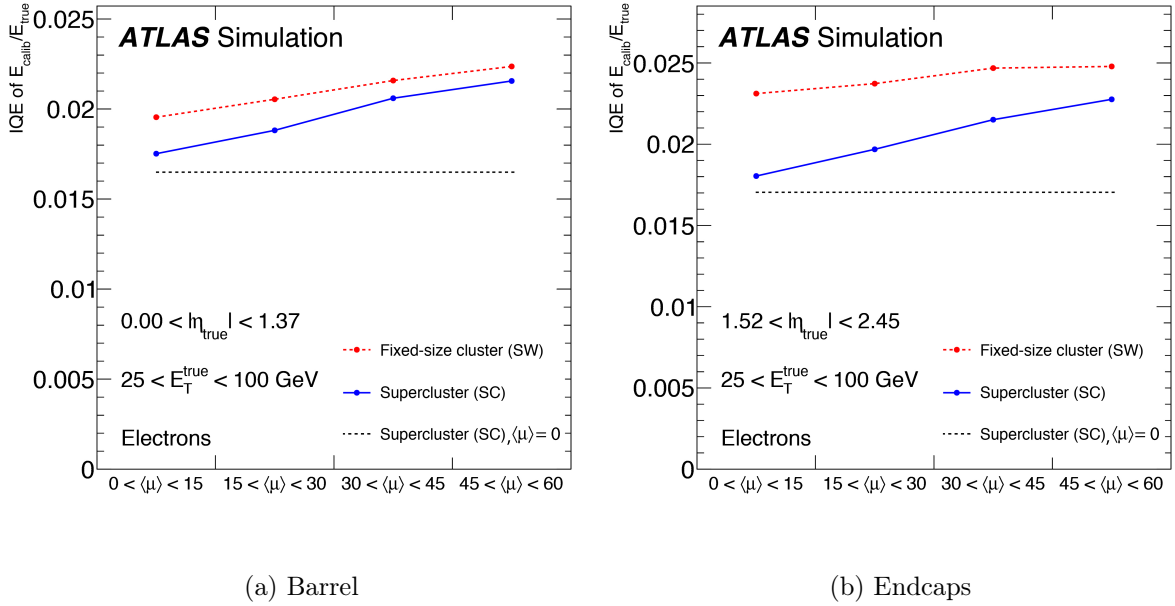


Figure 29: Calibrated energy response resolution, expressed in terms of IQE, for simulated electrons at different $\langle\mu\rangle$ levels in (a) the central calorimeter and (b) for the endcaps. A dashed red line indicates the response for fixed-size clusters, while the supercluster-based response is shown in solid blue. The supercluster-based energy response resolution for $\langle\mu\rangle = 0$ is also given as a black dashed line for comparison. *Source: JINST 14 (2019) P12006*

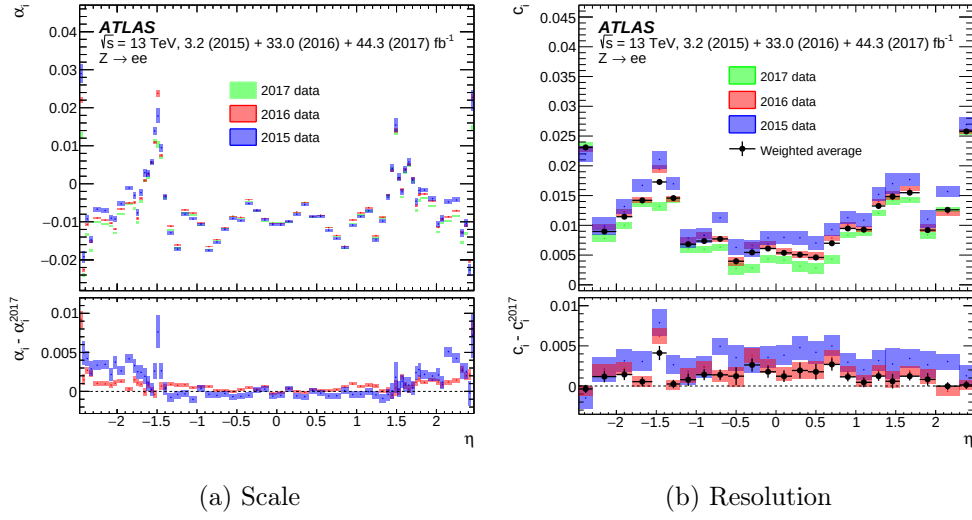


Figure 30: (a) Energy scale factors α_i and (b) additional constant term c_i , as a function of η . The shaded areas correspond to the statistical uncertainties. The bottom panels show the differences between (a) α_i and (b) c_i measured in a given data-taking period and the measurements using 2017 data.
Source: JINST 14 (2019) P12006

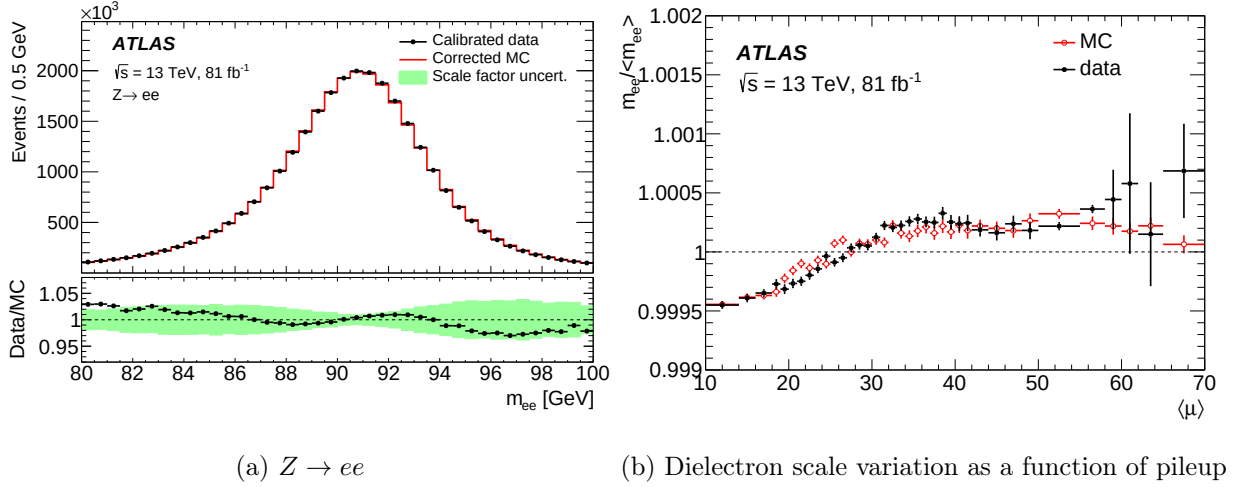


Figure 31: (a) Comparison between data and simulation of the invariant mass distribution of the two electrons in the selected $Z \rightarrow ee$ candidates, after the calibration and resolution corrections are applied. The total number of events in the simulation is normalized to the data. The uncertainty band of the bottom plot represents the impact of the uncertainties in the calibration and resolution correction factors. (b) Relative variation of the peak position of the reconstructed dielectron mass distribution in $Z \rightarrow ee$ events as a function of the average number of interactions per bunch crossing. The error bars represent the statistical uncertainties.
Source: JINST 14 (2019) P12006

Following reconstruction, ATLAS uses a likelihood-based identification for electrons, described in reference [40]. The reconstruction efficiency threshold in matching ID tracks and ECAL clusters, and subsequently identifying them as electron candidates (figure 32), motivates the minimum transverse energy requirement in table 4.

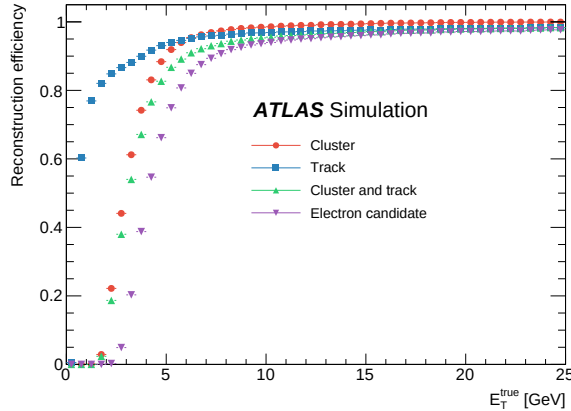


Figure 32: Four reconstruction efficiencies as a function of generated electron transverse momentum, E_T^{true} . The electron reconstruction efficiency threshold (purple triangles) motivates the electron energy selection in table 4. *Source: JINST 14 (2019) P12006*

6.4.2 Superclustering and the Higgs mass

Reference [41] demonstrated the expected performance for this improved electron reconstruction in the context of the $H \rightarrow 4\ell$ resonance. VBF-mediated Higgs boson signal samples were produced with Monte Carlo generator POWHEG [42] at $m_H = 125$ GeV. Monte Carlo generator PYTHIA8 [43] handled their showering at parton-level and the $H \rightarrow 4\ell$ decay.¹⁴ These events then passed through the full simulation chain and were reconstructed twice (once each with the sliding window and supercluster algorithms). Candidate reconstruction and selection otherwise followed previous $H \rightarrow 4\ell$ analyses of the time (as in reference [44]). The resulting $m_{4\ell}$ distributions (figure 33) were fit with a double-sided Crystal Ball function to quantify the supercluster algorithm’s improvement (table 5). The supercluster algorithm’s ability to compensate for Bremsstrahlung ultimately produces narrower resonances with

¹⁴Event simulation has changed since this 2017 study. Please see appendix D for the simulation details for the full Run 2 Higgs mass measurement.

shorter radiative tails compared to those of the sliding window algorithm. Additionally, the supercluster algorithm better recovers the energy scale.

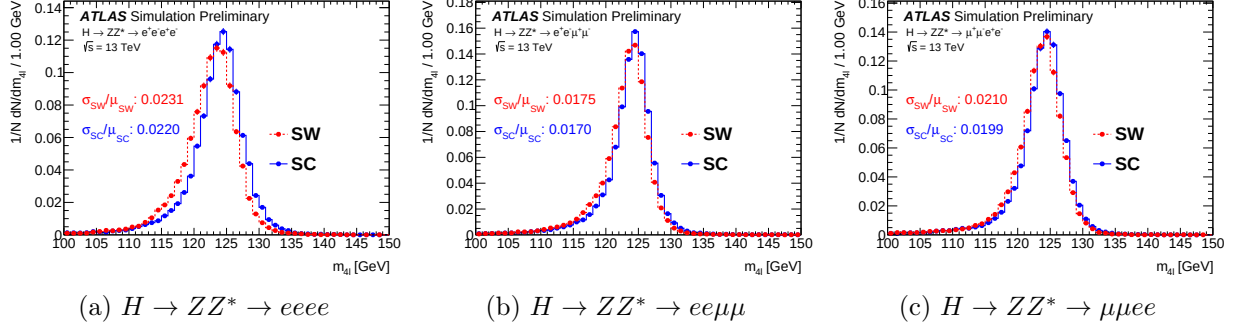


Figure 33: Invariant mass distributions of simulated reconstructed four-lepton candidates with $\langle\mu\rangle \leq 40$ using both supercluster and sliding window algorithms, built using the $H \rightarrow 4\ell$ analysis selection. The resolution and mean of each distribution are computed using a double-sided Crystal Ball fit. *Source:*

[ATL-PHYS-PUB-2017-022](#)

Table 5: Fitted mean, μ , and resolution, σ , for the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ resonance candidates with $\langle\mu\rangle \leq 40$ using both sliding window (SW) and supercluster (SC) algorithms. *Source:*

[ATL-PHYS-PUB-2017-022](#)

Process	μ_{SW} [GeV]	σ_{SW} [GeV]	σ_{SC} [GeV]	μ_{SC} [GeV]	Change in σ/μ
4e	123.30	2.85	124.37	2.73	-4.8%
2 μ 2e	123.90	2.60	124.46	2.48	-5.2%
2e2 μ	124.03	2.17	124.47	2.12	-2.9%

6.5 Event-level selection

Having qualified our muons and electrons individually, we now need to consider properties of the entire event. Table 6 summarizes the selection criteria we impose on the full quadruplet.

The isolation and vertex selections were updated for Run 2 with respect to Run 1.

Table 6: Summary of the criteria applied to the selected Higgs boson candidate in each event. The mass threshold $m_{\min}$ is 12 GeV for all $m_{4\ell} < 140$ GeV. Between 140 and 190 GeV, the threshold value rises by 0.76 GeV for each 1-GeV-increment in $m_{4\ell}$ until reaching a 50-GeV threshold value [36]. The distance ΔR lies in the $\eta - \phi$ plane, defined as $\Delta R = \sqrt{\eta^2 + \phi^2}$.

Source: *submitted to EPJC*

QUADRUPLETS	
All combinations of two same-flavour and opposite-charge lepton pairs	
- Leading lepton pair: lepton pair with invariant mass m_{12} closest to the Z boson mass m_Z	
- Subleading lepton pair: lepton pair with invariant mass m_{34} second closest to the Z boson mass m_Z	
Classification according to the decay final state: 4μ , $2e2\mu$, $2\mu2e$, $4e$	
REQUIREMENTS ON EACH QUADRUPLLET	
LEPTON	- Three highest- p_T leptons must have p_T greater than 20, 15 and 10 GeV
RECONSTRUCTION	- At most one calorimeter-tagged or stand-alone muon
LEPTON PAIRS	- Leading lepton pair: $50 < m_{12} < 106$ GeV
	- Subleading lepton pair: $m_{\min} < m_{34} < 115$ GeV
	- Alternative same-flavour opposite-charge lepton pair: $m_{\ell\ell} > 5$ GeV
	- $\Delta R(\ell, \ell') > 0.10$ for all lepton pairs
COMMON VERTEX	- χ^2 -requirement on the fit of the four lepton tracks to their common vertex
SELECTION OF THE BEST QUADRUPLLET	
- Select quadruplet with m_{12} closest to m_Z from one decay final state	
in decreasing order of priority: 4μ , $2e2\mu$, $2\mu2e$ and $4e$	
- If at least one additional (fifth) lepton with $p_T > 12$ GeV meets the isolation, impact parameter and angular separation criteria, select the quadruplet with the highest matrix-element value	
HIGGS BOSON MASS WINDOW	
- Correction of the four-lepton invariant mass due to the FSR photons in Z boson decays	
- Four-lepton invariant mass window in the signal region: $115 < m_{4\ell} < 130$ GeV	
- Four-lepton invariant mass window in the sideband region:	
$105 < m_{4\ell} < 115$ GeV or $130 < m_{4\ell} < 160$ (350) GeV	

6.6 Defining the four-lepton invariant mass, $m_{4\ell}$

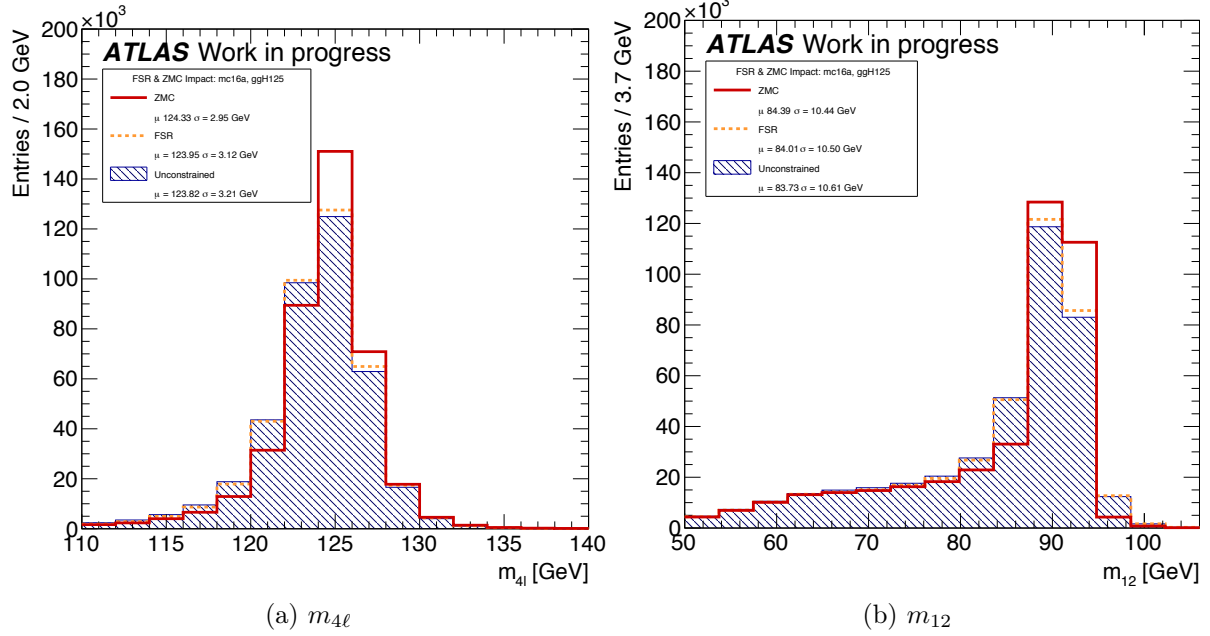


Figure 34: These histograms show the impact of adding final state radiation, FSR, and the Z-mass constraint, ZMC, on the reconstructed (a) $m_{4\ell}$ and (b) m_{12} shapes. Notice that the ZMC histogram (solid red line) has a greater peak population as the ZMC migrates some events back into the peak. *Source: HH (2020)*

We have three approaches for building the 4ℓ invariant mass, described in this section. Figure 34 compares the resulting distributions under the Higgs and on-shell Z peaks.

We calculate the four-lepton four-momentum from the four-momentum of the two intermediate Z bosons, Z_1 and Z_2 , expressing each Lorentz vector in $(p_T, \eta, \phi, M(t))$ coordinates.¹⁵ When the four-vector sum of the members of the leading dilepton pair form Z_1 and the four-

¹⁵Specifically, we build Lorentz vectors for each lepton according to the ROOT class `TLorentzVector`. We initialize each as a blank `TLorentzVector`, then use the specialized class method `SetPtEtaPhiM(pt, eta, phi, m)` to populate it. The class handles summing the Lorentz vectors internally. We sum the Lorentz vectors of the leading (subleading) leptons to calculate the Lorentz vector for Z_1 (Z_2). When we account for final state radiation, FSR, we include the `TLorentzVector` for the FSR candidate in the sum for Z_1 . When we apply the Z-mass constraint, ZMC, we recalculate the components of the four-momentum of each leading lepton (taking into account the FSR candidate, if present) and use the recalculated components in new `TLorentzVectors` to find the sum for Z_1 . Once we have the Z_1 and Z_2 `TLorentzVectors` in hand, we sum them to produce the four-lepton `TLorentzVector`. We can then access its properties through the class' methods. For example, we retrieve the four-lepton invariant mass using `TLorentzVector::M()`.

vector sum of the subleading pair produces Z_2 , the resulting four-lepton system is considered to be **unconstrained**, with four-lepton invariant mass $m_{4\ell}^{\text{unconstrained}}$. This is the simplest definition possible.

When the intermediate Z bosons decay, they may produce **final state radiation (FSR)** in addition to the two charged leptons. We account for FSR in Z_1 decays only, restoring a single FSR photon to about 4% of events [34]. The sum of the four-vectors of both leptons and the FSR photon then gives the Z_1 four-vector. Figure 35 displays a flowchart of the event-level search for FSR candidates. FSR candidates pushing $m_{Z1} > 100$ GeV are always rejected. If we form the four-lepton system while taking into account FSR, we refer to the system as **FSR** and designate the four-lepton invariant mass as $m_{4\ell}^{\text{FSR}}$.

Finally, after applying the FSR recovery procedure, we apply the **Z-mass constraint (ZMC)** algorithm. We recompute the lepton four-momenta of the leading lepton pair (before their Lorentz summation) using this kinematically-constrained fit. The fit assumes that the leading pair constitutes an on-shell Z boson leading to some sculpting of the m_{12} shape (figure 34b). The Z's lineshape is taken as a Breit-Wigner and the expected momentum response of each lepton is treated as a single Gaussian. The ZMC improves the $m_{4\ell}$ resolution by about 17%. If we form the four-lepton system taking into account FSR and ZMC, we consider it **constrained** with invariant mass $m_{4\ell}^{\text{constrained}}$. Going forward, $m_{4\ell}$ refers to $m_{4\ell}^{\text{constrained}}$, unless specifically noted.¹⁶

Table 7: This table summarizes the impact of accounting for final state radiation, FSR, and applying the Z-mass constraint, ZMC, on m_H^{measured} derived with signal-only simulation. The improvement is estimated with respect to $m_{4\ell}^{\text{unconstrained}}$. *Source: $H \rightarrow 4\ell$ mass measurement group*

Four-lepton invariant mass	FSR recovery	ZMC applied	Resolution improvement (%)
$m_{4\ell}^{\text{unconstrained}}$	no	no	—
$m_{4\ell}^{\text{FSR}}$	yes	no	1%
$m_{4\ell}^{\text{constrained}}$	yes	yes	17%

¹⁶Further information on both the FSR scheme and ZMC are available in references [14, 15, 17].

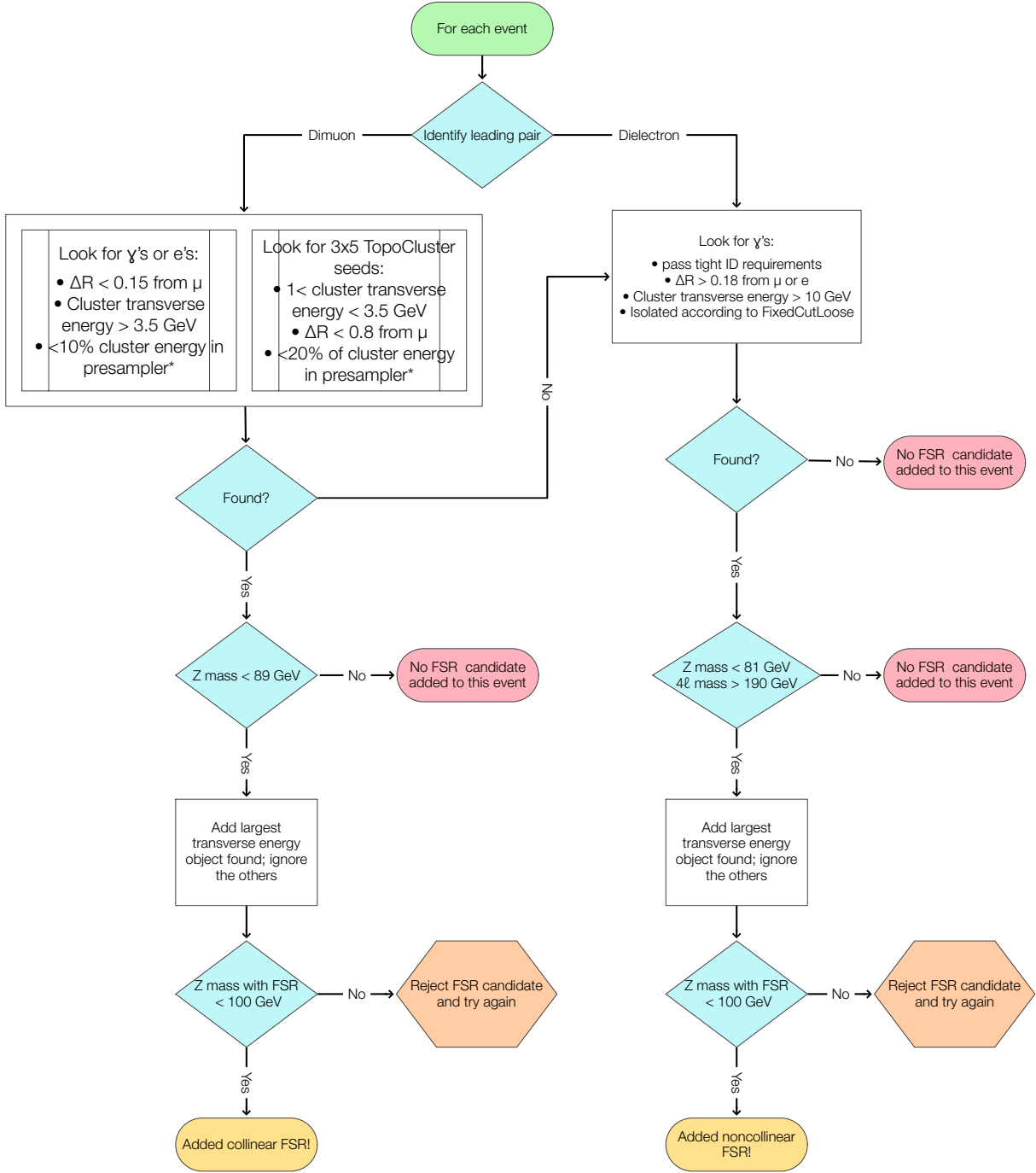


Figure 35: Final state radiation, FSR, search flowchart, detailing the selection and showing that we prioritize collinear FSR over non-collinear (also called “far”) FSR. The E_T requirements are more demanding for non-collinear FSR to reduce the hadronic background. (*) The presampler energy fraction requirement protects against muon-induced ionization. *Source: HH (2020)*

7 Analyzing the $m_{4\ell}$ spectrum

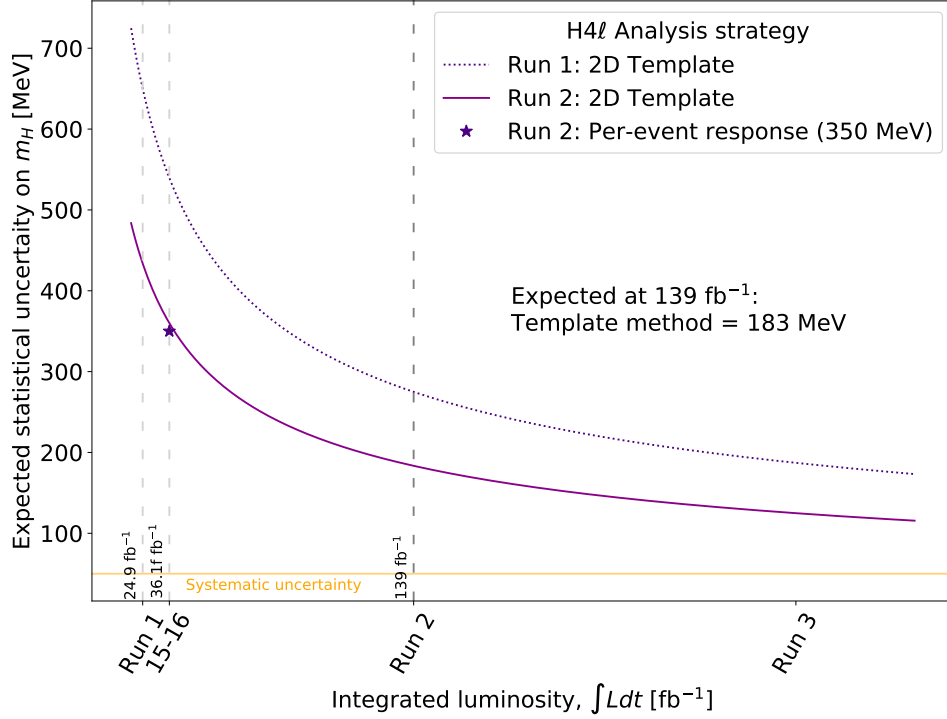


Figure 36: Extrapolating Run 1 [15] and mid-Run 2 [18] expected uncertainty as a function of integrated luminosity. The mid-Run 2 measurement benefitted from improved muon and electron calibrations and clever analysis choices such as improvements in 4ℓ -vertexing so that its expected uncertainty was less than that of the projected Run 1 measurement even while using fundamentally the same 2D-template approach. *Source: HH (2019)*

Thus far, we have selected relevant events and decided how to calculate our discriminating variable, the four-lepton invariant mass, $m_{4\ell}$, chosen for its maximum sensitivity to the Higgs boson’s mass, m_H . This section concerns how we interpret the resulting $m_{4\ell}$ spectrum to extract m_H . We must build a complete model describing our expected signal and background contributions (this section), and then evaluate the observed dataset in this model’s context (section 8.3).

Lepton energy resolution and analysis design drive our uncertainty on m_H since this analysis is statistically limited. Critically, uncertainty on m_H decreases *linearly* with resolution improvements, including improved lepton reconstruction performance and analysis-level

design choices. Meanwhile, the uncertainty on m_H decreases only with the square root of the number of observed events, so gains in luminosity or acceptance are less impactful (figure 36). While systematic uncertainties exist, they do not dominate this measurement. In the past, we have observed about 50 MeV of systematic uncertainty in this channel, while figure 36 indicates that we should expect over three times that in statistical uncertainty - the two uncertainty sources sum in quadrature. The sources of systematic uncertainty will be discussed in section 8.

Improvements in reconstructing muons and electrons were already reviewed in sections 6.3 and 6.4, respectively. This section concentrates on the analysis design, reviewing event categorization first and then discussing the model construction.

A note on machine learning

Traditional analysis choices, like those in section 6, rely on distributions of single variables and cannot consider higher-order relationships between variables in a straight-forward way. **Machine learning (ML)** is a field of study that enables computers to learn without being explicitly programmed. A computer can access those higher-order connections using machine learning techniques. Guest, Cranmer, and Whiteson provide an good overview of machine learning techniques for LHC physics [45].

In a particular class of ML, **supervised learning**, we feed the correct answer to the learning algorithm; we already know the input and output behavior. **Classification** and **regression** present two examples of supervised learning. Classification predicts a discrete-valued output while regression predicts a continuous-valued output. Both techniques depend on having a “training” sample encoding the “correct” relationship between the input and output. We used two supervised learning techniques in this analysis:

1. A classifier designed to separate signal-like and background-like events (section 7.1.1) and
2. A regressor predicting our resolution in each event (section 7.3).

7.1 Event categorization

Discrete characteristics, like the final state, lend themselves to distinct groups while continuous properties, like spectra, can be parameterized. Grouping events according to their discrete, common properties allows us to tune models for each class of events. The final model, then, is a composite over all tuned models, resulting in a more powerful description for the whole group. In this analysis, we separate events into 16 distinct categories based on their decay channel and score from a multivariate classifier algorithm, BDT_{ZZ} (section 7.1.1), designed to separate signal-like and background-like events.

The four final states (4μ , $4e$, $2\mu 2e$, and $2e 2\mu$) already approximately separate events by their resolution, since muons and electrons have different resolutions themselves (figure 37). Resolution effects have more impact on the signal shape where the expected distribution is a narrow resonance, while the dominant background shape from the non-resonant $ZZ^{(*)}$ production processes is relatively flat. Additionally, we categorize events according to a **classifier** designed to separate signal-like and background-like events based on their production and decay properties. The BDT_{ZZ} enables us to build signal-background discrimination into our categorization, allowing us to further specialize our probability density functions in each category.

7.1.1 BDT_{ZZ} classifier

This classifier is an example of supervised machine learning and it is an updated version of the classifier developed for previous ATLAS m_H measurements [18, 46].

We trained this classifier using Gradient-Boosted Decision Trees with TMVA [47], meaning that we used an ensemble of decision trees to predict a target's class - signal event or background event. A **decision tree** consists of a series of nodes, each a single classification test, much like a personality quiz in a teenage magazine or on Buzzfeed. Typically, the tests have a binary result, producing two paths to the next level of nodes. Eventually,

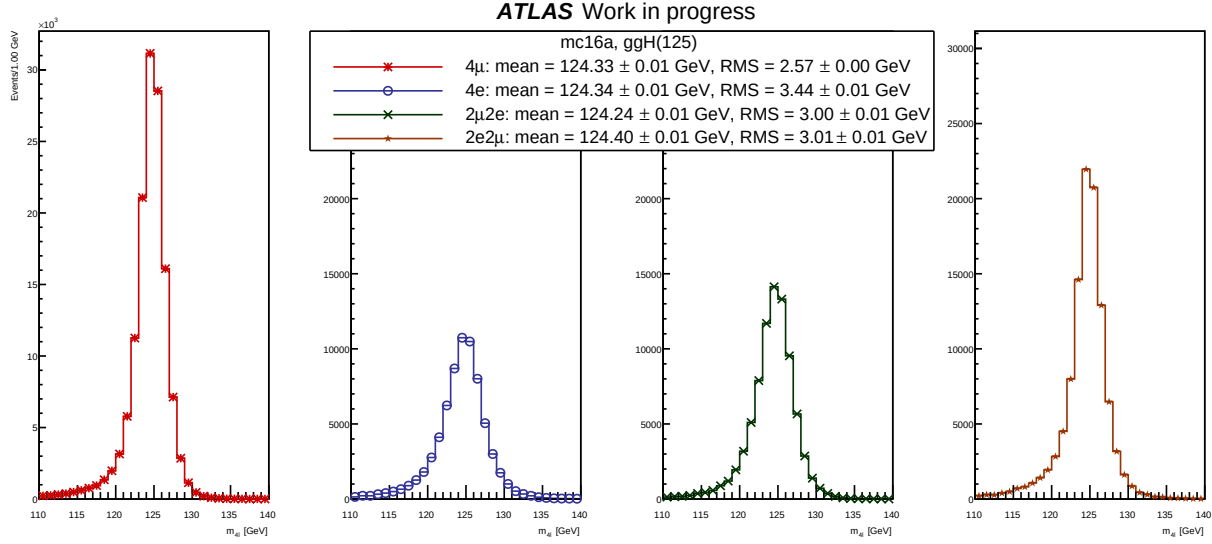


Figure 37: $m_{4\ell}^{\text{constrained}}$ distribution for ggH signal events in each decay channel. The given mean and RMS are for each whole histogram. Treating the RMS as a proxy for the resolution, we see that 4μ has the best resolution, followed by $2e2\mu$, $2\mu 2e$, and finally, $4e$. The ZMC benefits the $2e2\mu$ final state the most since the correction is strongest for electrons. *Source: HH (2020)*

the series of nodes terminates in different groups, assigning a label (the animal that best describes the person taking the quiz in the teenage magazine; signal or background event in particle physics). Decision trees are quick and intuitive. They expose first-order correlations among observables relatively easily. Unfortunately, a single, generated decision tree is quite unlikely to be the best classifier for our scenario. Rather, we can generate an entire random *forest* of decision trees and let their statistical ensemble assign our events' labels. In other words, many trees in agreement strengthens our faith in that event's assignment. *Boosted* decision tree forests increase the weight of misclassified events so that those events receive more attention from the next learners.

As a supervised learning algorithm, we must provide a sample with the “correct” event classifications for training and testing. We composed the training sample from two groups labeled signal and background. The signal consisted of ggH events generated under the Standard Model Higgs hypothesis at a variety of Higgs mass points (123, 124, 124.5, 125,

125.5, 126, and 127 GeV) to protect the trained algorithm from bias as a function of m_H . The background sample consisted of Monte Carlo generator SHERPA [48–51] samples for $qq \rightarrow ZZ^{(*)}$ and $gg \rightarrow ZZ^{(*)}$ processes. The training sample included events passing the standard selection (section 6) in a reduced invariant mass range, $115 < m_{4\ell} < 130$ GeV, so that BDT_{ZZ} would be more sensitive to characteristics of the signal peak. The training occurs separately for same-flavor (4μ and $4e$) and opposite-flavor ($2\mu 2e$ and $2e 2\mu$) final states, on account of the interference effects in final states with identical leptons.

We also choose a set of “features,” corresponding to observables characterizing our events, for the algorithm to use to distinguish ggH-like events from non-resonant $ZZ^{(*)}$ -like events. The features include:

1. Four-lepton transverse momentum, $p_T^{4\ell}$;
2. Four-lepton pseudorapidity, $\eta_{4\ell}$;
3. A kinematic discriminant based on the calculated ggH and $ZZ^{(*)}$ matrix elements:
 - $KD(ZZ^{(*)}) = \ln \left(\frac{|M_{\text{ggH}}|^2}{|M_{ZZ}|^2} \right)$,
 - Accurate to leading order,
 - ggH production and Standard Model $ZZ^{(*)}$ production,
 - Calculated with MADGRAPH5_aMC@NLO [52].

Both four-lepton properties include FSR recovery and exclude the effects of the ZMC. Overall, this discriminant enhances the signal significance in the m_H measurement because it separates signal and background events based on decay properties (primarily) and the decay properties are comparable across all Higgs production modes.

We can evaluate the resulting algorithm on any arbitrary sample as long as the sample’s events have the same features as the training sample. Each event in the evaluated sample receives a score in the interval $[-1, 1]$. Events with a BDT_{ZZ} score closer to 1 are more ggH-like and those with a score closer to -1 are more background-like.

Figure 38 shows the resulting distributions of BDT_{ZZ} score for the signal and background samples, on which the dashed gray lines represent the bin boundaries for defining the analysis categories.

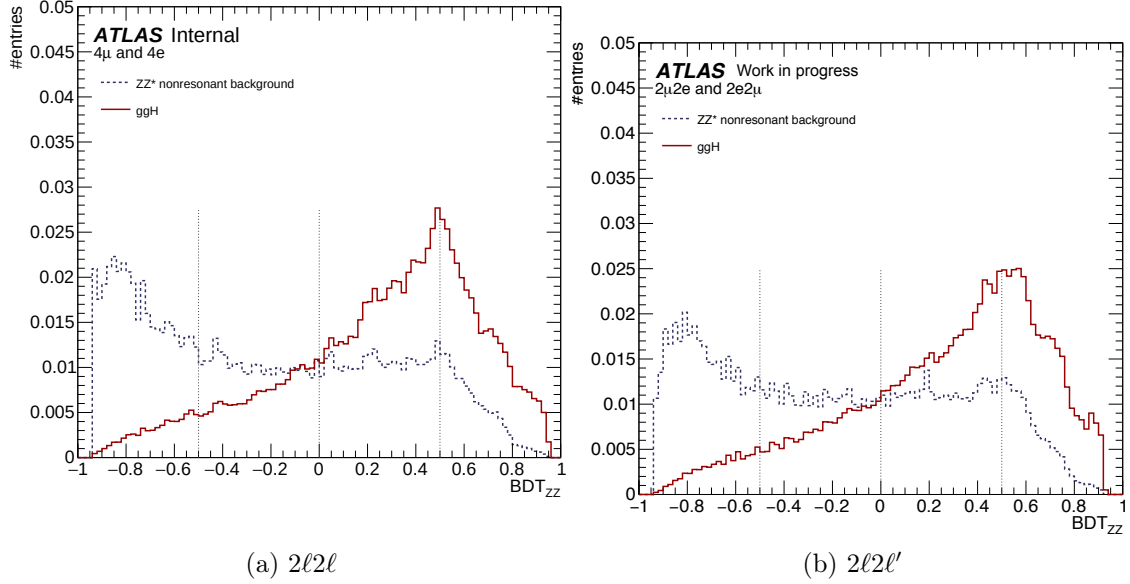


Figure 38: BDT_{ZZ} discriminant used for defining analysis categories. Distributions for signal and background are given in red and blue, respectively. The dashed grey lines mark the bin boundaries used in defining the analysis categories. *Source: $H \rightarrow 4\ell$ mass measurement group*

7.2 Signal model development

As our signal and background physics processes are distinct, our final probability density function (PDF) can be a composite of separate models for the signal and background processes. This subsection addresses the design of the signal model. We must consider that our signal's shape, the decay of a Higgs boson via two intermediate Z bosons to four leptons, is predicted to be a very thin line consistent with a Breit-Wigner with a width of just 4.1 MeV, about 400 times smaller than ATLAS's best detector resolutions. Our signal model then must account for both the theoretical lineshape and the detector response. Comparing our detector to a camera, we can perceive our detector as a **filter** over the decay's true lineshape. In other words, our resulting image (our reconstructed collision) is a **convolution** of the

true lineshape and the detector’s response.

Simulation provides a natural route to recovering that full convolution. First, we can simulate the decay itself, modeling the theoretical lineshape. Then, we can pass the generated decay’s products through a simulation of the ATLAS detector itself, therefore convolving the theoretical lineshape with the detector response. We can take the full simulation’s output distribution as our model, an approach termed a **template method** and discussed in more detail in section 7.2.1. Such methods must reflect the **average** behavior of our generated events. In particular, we are consigned to the **average response** of our events: we lose any sensitivity to the fluctuations in our detector response at the event level.

Per-event methods account for detector response to *each* event at the price of complication, both in the model and its computation. Large datasets prohibit per-event methods simply because they take too long to evaluate. $H \rightarrow ZZ^* \rightarrow 4\ell$ has few expected events so per-event methods are computationally viable. Section 7.2.2 summarizes ATLAS’ per-event approach for the mid-Run 2 dataset, while section 7.2.3 describes the novel per-event approach at the core of this thesis and the latest measurement. CMS’ $H \rightarrow ZZ^* \rightarrow 4\ell$ mass measurement analyses [16, 19, 53] also employed a per-event method, taking the propagated lepton uncertainty in each event as a proxy for the event’s resolution, an approach known as **per-event errors**. All three approaches hinge on defining each event’s detector response reliably. Table 8 summarizes the resolution model for each method.

7.2.1 Template method

In **template methods**, the model is a smoothed probability density function (PDF) of a set of observables and free parameters. Historically, the template method was the primary signal model for the ATLAS Run 1 measurements [14, 15, 17] and a cross check for the ATLAS mid-Run 2 measurement [18], using observable $m_{4\ell}^{\text{constrained}}$ with one free parameter, m_H . It serves as a cross check in this edition of the analysis, described in appendix H.

Template morphing creates a continuous PDF as a function of parameters of interest from

Table 8: Per-event methods hinge on defining each event’s resolution reliably. The table briefly describes three methods. *Source: HH (2020)*

Technique	Detector response description	Experiment	Reference
Per-event errors	Assume an analytic PDF with a resolution-related width parameter (i.e. Gaussian, Crystal Ball). Width parameter proportional to uncertainty on each event’s invariant mass	CMS	[16, 19, 53]
Per-event response	No analytic PDF. Predict full event’s response distribution based on individual physics objects’ energy responses	ATLAS	[18]
Per-event resolution	Assume an analytic PDF with a resolution-related width parameter. Width equal to predicted resolution for each event	ATLAS	<i>Current</i>

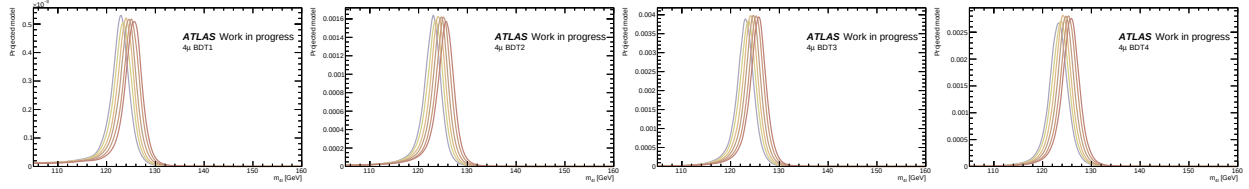


Figure 39: Demonstration of template morphing using RooBSpline [54] in each BDT_{ZZ} bin (increasingly more signal-like when moving left to right) in the 4μ channel. The signal PDF $P(m_{4\ell}^{\text{constrained}}|m_H)$ is drawn in 600 MeV steps from 123 to 126 GeV. *Source: $H \rightarrow 4\ell$ mass measurement group*

a set of discrete (probability) distributions. Each discrete distribution (shape) provides a single point in the parameter space. Interpolating those shapes to intermediate points of the parameter space, or “morphing,” then converts that set of shapes to a continuous function of the unknown parameter(s). In this particular analysis, we interpolate between the extracted shapes using RooBSpline [54]. For the mass measurement, we take the following steps:

1. Extract each discrete $m_{4\ell}^{\text{constrained}}$ distribution associated with one value of m_H , or “template,” from the set of $m_H^{\text{generated}}$ samples;
2. Interpolate them to produce a continuous function with a single free parameter, m_H :

$$P(m_{4\ell}^{\text{constrained}}|m_H).$$

Figure 39 displays the morphed template PDF drawn at several values of m_H in the 4μ

channel. Template morphing traditionally occurs along a single dimension, like $m_{4\ell}^{\text{constrained}}$. It is possible (albeit complicated) to morph along several dimensions simultaneously.

7.2.2 Per-event response

Per-event methods, as previously mentioned, hinge on defining the detector response to *each* event and taking that information into account when constructing the model. The ATLAS mid-Run 2 method, **per-event response** [18], sought to estimate the *entire* detector response function directly from the kinematics of the four primary leptons¹⁷ in the event. The method posits that the full detector response is a repeated convolution of ATLAS' response to each constituent lepton. This method separated the detector response component from the Higgs boson's physical lineshape, reducing the impact of theoretical assumptions about the true nature of the Higgs on the measurement. This description is intended to be conceptual and mathematical details are omitted as it is not the endgame of this dissertation. Rather, this method is included to provide contrast with the per-event resolution method selected for the full Run 2 measurement. For the full description, please see reference [18]. Per-event

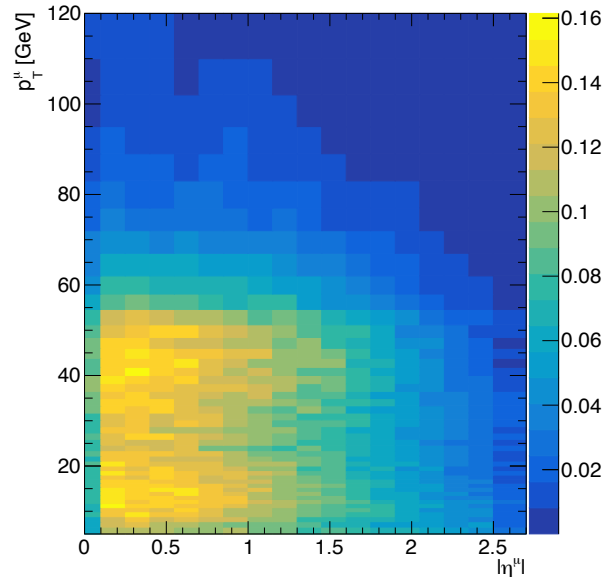


Figure 40: Example binning palette for formulating the lepton response lookup table. *Source: HH (2019)*

¹⁷Some events, like those from $t\bar{t}H$ -associated production, may contain more than four leptons.

response supposes that we can approximate ATLAS’ response to each lepton based on the *expected* behavior of a lepton with a comparable kinematic profile. We built four “palettes” for binning the response (see figure 40 for an example), splitting the leptons according to the following:

1. Lepton is a muon or an electron:

- Electrons and muons respond to material differently. Electrons are much more likely to produce bremsstrahlung, for example, since they are substantially lighter than muons. We may not recover all radiated energy (if it is lost within the tracker volume, for example), smearing the detector’s response.
- Muons and electrons pass through different ATLAS subdetectors and follow different chains of reconstruction.

2. Lepton originates either from the leading or subleading dilepton pair:

- Z_1 and Z_2 in the $m_{4\ell}$ spectrum play different roles.
- Z_1 is more closely associated with an on-shell Z boson, so its decay products are harder leptons.
- Z_2 is almost always off-shell in our mass range, so its decay products tend to be softer leptons.
- We subject leptons in the leading dilepton pair to more stringent selection criteria so they tend to be better reconstructed.
- Taken together, we cannot reasonably expect the leading leptons to have the same response profile as the subleading leptons.

3. Binning in lepton transverse momentum (muons) or energy (electrons):

- As a general rule of thumb, trackers have better resolution at lower particle energies while calorimeters improve at higher energies.

- Splitting along momentum or energy provides sensitivity to these fundamental detector properties.

4. Binning based on lepton centrality, approximated by $|\eta|$:

- ATLAS’ barrel tends to be more hermetic, meaning that it is more fully instrumented, both at the triggering and precision-measuring levels.
- Consequently, we expect cleaner leptonic detector responses at lower $|\eta|$.¹⁸

In this method, we modeled the detector response in each box of each palette as the sum of three Gaussians, the minimum number required to describe the energy response in each box. We chose Gaussians for several reasons. First, the set of all Gaussian functions forms a complete, orthogonal set of basis functions. Any other function, like our individual lepton response, can therefore be expressed as a sum of Gaussian functions. The sum of Gaussian functions is always another Gaussian function, and their convolutions are analytical. We can usually associate properties of Gaussian functions with physical properties, like resolution cores or radiative tails.¹⁹ Ultimately, each lepton’s identity, status as a Z_1 or Z_2 lepton, p_T , and $|\eta|$ provide “coordinates” to look up the appropriate expected energy response function for that lepton (in other words, to supply the three Gaussian widths needed to build our normalized response PDF for that individual lepton). Note that we have effectively **binned** our detector response within this method.

Each event contains four leptons, and we expect the full response model to be the **convolution** of all four lepton models. A total of 81 Gaussians would contribute to a single event’s response model! We must reduce the number of Gaussians in the final description to maintain manageable single-event computation times. **Gaussian Sum Filter (GSF)** [55]

¹⁸Recall figures 16 and 26 which showed sensitive detector coverages for muons and electrons, respectively. The dotted lines in the middles of those figures represent the selected binning for lepton pseudorapidity.

¹⁹As a rule of thumb, the Gaussian with the largest weight in the superposition corresponded to the core of the $m_{4\ell}$ distribution, while the others accounted for tail behavior. Muons adhered closely to this paradigm and the Gaussian widths could even be directly compared with known muon resolutions. Electrons followed this rule less closely.

is a technique based on Kalman Filters²⁰ used in electron reconstruction to reduce these 81 Gaussians to a more manageable four. By design, this process ensures that we do not change the first and second moments of the distribution when we apply GSF. This is critical: the first moment (mean) of the distribution is the energy scale of the event, and the second moment (variance) gives that event’s resolution.

Finally, the complete signal PDF convolves the true lineshape (a Breit-Wigner with 4.1 MeV width) with this $m_{4\ell}$ response, providing a PDF with single observable, $m_{4\ell}^{\text{constrained}}$, and several *conditional* observables based on the response palette binning. The model has two parameters of interest - the Higgs mass and its width. It works – at the price of binning our detector response (effectively creating a resolution look up table) and a fair amount of complication.

7.2.3 Per-event resolution

Per-event resolution maintains the advantages of a per-event method, while providing a continuous parameterization of the detector response and (hopefully) being more intuitive. We choose some function describing analytically our discriminating variable’s distribution and then modify the width of that function based on event-specific properties. The goal of this subsection is to establish the analytic function chosen to represent the $m_{4\ell}$ distribution. Section 7.3 will supply our definition for the event resolution in this analysis and upgrade the analytic model explored here to its per-event description.

To develop an appropriate analytic description, we must return again to the role our detector plays in measuring the Higgs boson’s true lineshape. **The ATLAS detector serves as a filter over the decay’s true lineshape.** The observed $m_{4\ell}$ distribution is a convolution of the detector response and the very narrow Breit-Wigner characteristic of the SM Higgs boson.²¹ Essentially, the objective when developing the analytic description is to

²⁰Also known as linear quadratic estimation

²¹The SM Higgs width is 4.1 MeV, tiny relative to ATLAS’ resolution (typically 2 GeV, in order to be competitive with the Z boson’s natural width). The Higgs width is so small relative to our resolution that we can neglect the convolution. It is reasonable to assume that the template distributions are overwhelmingly

choose a function representative of the **average detector response** and convolve it with the Higgs' own Breit-Wigner distribution. Recall that when creating a template method, we allowed the full simulation to execute this entire process and then took the shape from the simulation. Using an analytic description for $m_{4\ell}$ with one free parameter of interest mimics the template method, with greater flexibility. We could update our description to include more parameters of interest (without requiring additional dimensions of interpolation as in the case of the template method) or convert it to a per-event method as we will see.

Template distributions guide the analytic function development, telling us that the detector response is essentially Gaussian in the core of the resonance, with an extended tail on the lefthand side, indicating particles' energy losses as they travel through ATLAS. The righthand tail falls quickly. In the 1980s, the Crystal Ball Collaboration created a function for exactly this use case. They were the first to graft exponential tails onto a Gaussian core, producing the Crystal Ball function, precisely to handle energy losses within the Crystal Ball detector while studying resonances [56–58]. As we have exponential behavior in both the lefthand and righthand tails, we take the **Double-sided Crystal Ball function (DCB)** to parameterize the $m_{4\ell}$ distribution:

$$DCB(m_{4\ell}; \mu_{CB}, \sigma_{CB}, \alpha_1, n_1, \alpha_2, n_2) = \begin{cases} \left(\frac{n_1}{\alpha_1}\right)^{n_1} \cdot e^{-\frac{\alpha_1^2}{2}} \cdot \left(\frac{n_1}{\alpha_1} - \alpha_1 - \frac{m_{4\ell} - \mu_{CB}}{\sigma_{CB}}\right)^{-n_1} & \text{for } \left(\frac{m_{4\ell} - \mu_{CB}}{\sigma_{CB}}\right) < -\alpha_1 \\ e^{-\frac{1}{2}\left(\frac{m_{4\ell} - \mu_{CB}}{\sigma_{CB}}\right)^2} & \text{for } -\alpha_1 \leq \left(\frac{m_{4\ell} - \mu_{CB}}{\sigma_{CB}}\right) \leq \alpha_2 \\ \left(\frac{n_2}{\alpha_2}\right)^{n_2} \cdot e^{-\frac{\alpha_2^2}{2}} \cdot \left(\frac{n_2}{\alpha_2} - \alpha_2 + \frac{m_{4\ell} - \mu_{CB}}{\sigma_{CB}}\right)^{-n_2} & \text{for } \alpha_2 < \left(\frac{m_{4\ell} - \mu_{CB}}{\sigma_{CB}}\right) \end{cases} \quad (14)$$

where μ_{CB} and σ_{CB} define the mean and width, respectively, of the Gaussian distribution in the core. The parameter μ_{CB} is closely related to m_H ; ideally, they would coincide. We can interpret σ_{CB} as a proxy for our detector resolution. The α_i and n_i parameterize the tails. The parameters α_i determine when the exponential tails begin, while the n_i parameters control how quickly each tail falls. We should expect $n_1 < n_2$ since the lefthand tail, related to

dominated by the detector response.

energy loss, falls more slowly than the righthand tail. Figure 41 overlays the $m_{4\ell}$ distribution in the 4μ channel with its DCB and highlights the exponential tails' relevance.

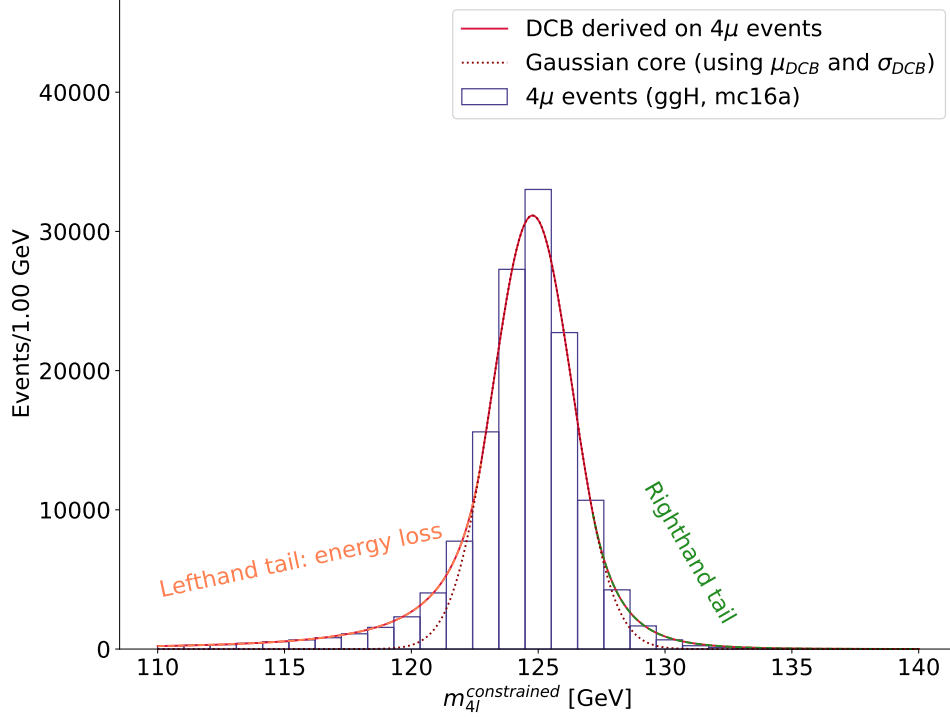


Figure 41: The blue histogram shows the distribution of $m_{4\ell}^{\text{constrained}}$ in the 4μ final state, effectively a template. The red solid line provides the DCB derived for this channel. Note that the DCB is not fit to the distribution, rather it is scaled to the maximum of the histogram. The dotted dark red line reflects the Gaussian core, while the orange and green dotted lines show where the exponential tails take effect in this example. Notice that the Gaussian diverges badly from the template distribution in the tails. *Source: HH (2020) using DCB parameters derived by T. Powell*

The DCB is still designed to describe an ensemble of events at this stage. It has a resolution parameter, σ_{CB} , related to the average resolution for the events from which we derive it. This resolution will impact the best-fit values of the other parameters as well. We can form an average resolution model based on this DCB if we make an unbinned maximum likelihood simultaneous fit to a selection of simulated m_H mass points²², sharing most of the parameters of the DCB - σ_{CB} , α_i , and n_i - across those mass points. We cannot, however,

²²We calibrated the DCB using simulated samples of the ggH and VBF production processes, encompassing about 95% of Standard Model Higgs boson production.

share the mean, μ_{CB} , across mass points. While each mass point carries within its set of events a DCB, they are deliberately generated at different values of m_H , shifting the DCB spectrum of each generated mass point with respect to another. We can re-parameterize the mean in terms of our parameter of interest, m_H , and then share that linearization across mass points, so that μ_{CB} becomes

$$\mu_{CB} = a_\mu \cdot (m_H - 125 \text{ GeV}) + b_\mu + 125 \text{ GeV}, \quad (15)$$

where a_μ and b_μ are shared across mass points. This particular form serves two purposes:

1. The uncertainty of a_μ has a smaller impact on m_H than the uncertainty of b_μ ;
2. Uncertainties of a_μ and b_μ are uncorrelated.

7.3 Defining event resolution

Rather than deriving an *average* resolution model, we can instead replace σ_{CB} with the **resolution in each event** and derive the other DCB parameters **conditional** on that per-event resolution. The goal of this subsection is to define the event resolution and demonstrate the complete (conditional) per-event resolution signal model as used for the current analysis.

7.3.1 “Why don’t you just use the lepton errors?”

Defining the per-event resolution by propagating the single-lepton uncertainties to $m_{4\ell}$ (calculating the **event error**) is a natural option, explored by CMS’ **per-event error method** in references [16, 19, 20, 53]. Unfortunately, the ATLAS lepton uncertainties are underestimated by 10-20% in simulation, so we cannot propagate them directly (figure 42, noting particularly the widths of the pulls).

The next option is to calibrate the individual muon and electron uncertainties, or “errors,” so that the pull plots of figure 42 show widths of 1, indicating that the errors are estimated correctly. The errors themselves depend on a multidimensional parameter space,

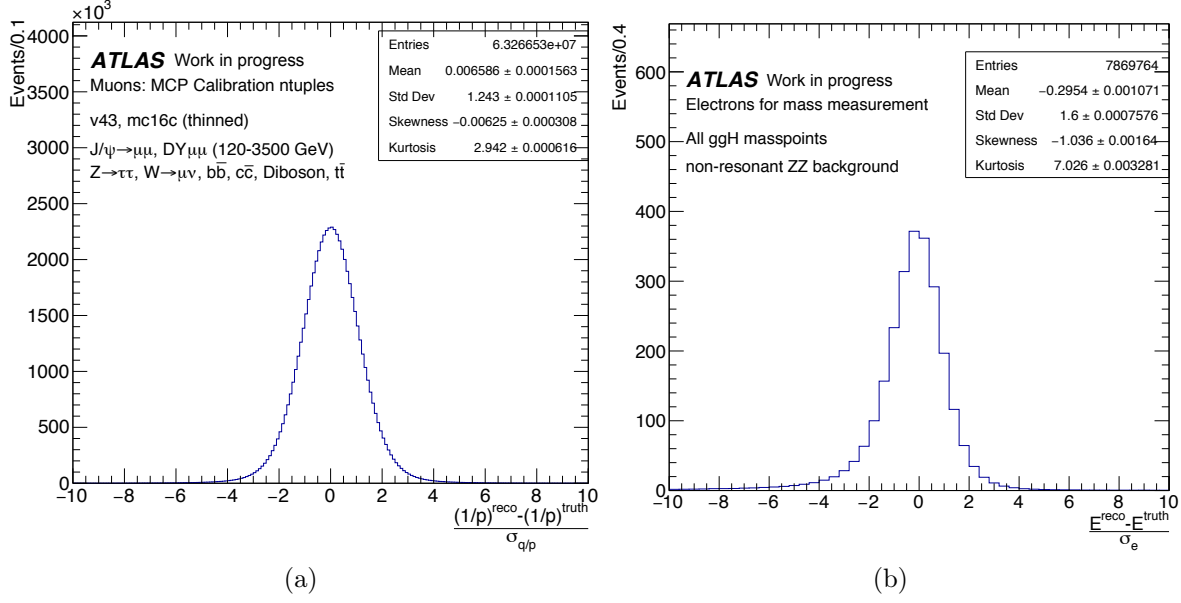


Figure 42: Momentum and energy pulls for individual muons and electrons, given the lepton errors. (a) Pull on inverse of muon momentum, p , with muon-level error $\sigma_{q/p}$, the combined track covariance matrix element. The pull width indicates that the muon error is underestimated by 20%. (b) Pull on electron energy with electron-level error σ_e , the expected resolution predicted for 95% of electrons as a function of transverse energy and detector region. The pull width indicates the electron error is also underestimated. *Source: $H \rightarrow 4\ell$ mass measurement group*

complicating calibrating them analytically. For example, a muon's error is represented by the $\sigma_{q/p}$ term of its track covariance matrix. Attempts to calibrate muon error analytically showed that it depends on the muon's kinematics and normalized χ^2 of its track. It is likely that other dimensions play a role as well, such as the number of hits and holes in the ID track. Worse, these dimensions are all correlated. Multivariate regression approaches using machine learning could help us find a suitable parameterization, mapping $\sigma_{q/p}$ in simulation to a calibrated value, since ML can address hidden correlations as well as the dimensionality. This option should certainly be explored with Combined Performance groups in the future.

The lepton errors also enter the ZMC kinematic fit, as discussed in section 6.6. The constraint correlates the first pair of leptons, along with any selected FSR photon.²³ This fit introduces its own uncertainty which must be reflected in the final event-level error on $m_{4\ell}$.

²³Recall that FSR photons appear in 4% of events so the photon uncertainty calibration is a relatively small concern.

We chose to use ML techniques to **parameterize the resolution directly** for each event based on its observables (rather than develop calibrations for muon and electron errors and then propagate them through the ZMC). The event resolution also depends on a multidimensional, highly-correlated parameter space, so it is also a good candidate for an ML-based regression approach. Recall that a regressor is appropriate because we seek a continuous parameterization of the resolution. **Deep learning** algorithms are particularly appropriate because they employ multiple layers of nodes (nonlinear processing units) for regression, increasing the algorithm’s sensitivity to higher-order correlations. We used two tools to develop the event resolution parameterization:

1. Tensorflow [59]: open-source library initially developed by Google for machine learning applications;
2. Keras [60]: Python API interfacing with Tensorflow. Facilitates quick experimentation with assorted network architectures and hyperparameters.

Directly calibrating lepton errors presents another challenge. We can certainly use a regressor from the deep learning algorithm class – it requires, however, that we define a suitable **target**. In other words, we must tell the regressor its goal - what feature of our dataset we want to compute in terms of the inputs. For the lepton errors, we would try to define our target so that the *width* of the pull distribution would be 1. In other words, we would need the variance of the output calibrated distribution to equal a particular value. Choosing instead to parameterize the resolution simplifies this problem. We can target the width of a distribution directly, rather than try to calibrate that width to some value for all events. We just have to define what we mean by “event resolution” in simulated events. We know the generated $m_{4\ell}^{\text{truth,born}}$ and reconstructed $m_{4\ell}^{\text{constrained}}$ in our simulated datasets, so we can calculate for each event the difference between the two. The width of this distribution represents our resolution. We can use generator-level information in our target during training, since the regressor is essentially a function mapping our input observables to

output event resolutions. The output event resolutions correspond to the regressor’s estimate of our target. As long as our input observables are all available at reconstruction-level, we can *evaluate* our regressor on ATLAS-collected data, though we trained it on a simulated dataset. Our set of input observables included:

- Lepton kinematics for each of the four leptons:
 - Transverse momentum, p_T^ℓ ,
 - Pseudorapidity, η_ℓ ,
 - Azimuthal angle, ϕ_ℓ ;
- Combined transverse momentum of the four-lepton system, $p_T^{4\ell}$ (ZMC applied);
- Four-lepton error output from the Z-mass constraint, $\sigma_{m_{4\ell}}^{\text{constrained}}$ (ZMC applied).²⁴

Future studies on the resolution should include more information about the ID tracks, such as the number of silicon tracker hits or holes, as input features to improve the resolution estimate.

7.3.2 Targeting the a distribution’s width: quantile regression neural networks

The most common regression applications target the **mean** of a distribution. The mean of $m_{4\ell}^{\text{truth,born}} - m_{4\ell}^{\text{constrained}}$ is the **scale** of our Higgs mass measurement. It refers to the absolute displacement in the position of the reconstructed resonance’s peak. The resolution, on the other hand, corresponds to the width of this distribution. We can use a neural network (NN) to predict the width of a distribution – a certain class of NNs modifies their loss functions to predict a designated percentile of a target expression. Traditionally, the percentile matches a quantile of the distribution, 31.7% or 68.3%, corresponding to the standard deviation either side of the mean for a Gaussian distribution, so these kinds of NNs are calls **quantile regression neural networks, QRNNs**.

²⁴Recall that the ZMC takes the covariance matrices (for muons) and expected energy estimates (for electrons) as inputs.

For this analysis, we ask the QRNN to predict a percentile of $|m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}|$.²⁵ Energy losses induce a slight asymmetry in the $m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}$ distribution. The first quantiles (capturing 31.7% of events) of the $m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}$ distribution in each decay channel are slightly lower than those of the third quantile (68.3%) of the same distribution (table 9). The third quantile (68.3%) of $|m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}|$ falls between the first and third quantiles of the unfolded distribution, showing that it is a reasonable approximation. The residual asymmetry amounts to 3 MeV in the 4μ channel and it is absorbed by the tails of the DCB parameterization of $m_{4\ell}^{\text{constrained}}$. As we will see in section 9, there is no bias even though we derived the resolution parameterization from the folded spectrum.

Table 9: First and third quantiles of $m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}$ in each decay channel. The bottom line provides the third quantile of the folded distribution, selected as the target. *Source: $H \rightarrow 4\ell$ mass measurement group*

Quantile of $m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}$	4μ [GeV]	$4e$ [GeV]	$2\mu 2e$ [GeV]	$2e 2\mu$ [GeV]
31.7%	-1.85	-2.84	-2.54	-2.13
68.3%	1.92	2.78	2.52	2.15
68.3% of $ m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}} $	1.88	2.81	2.54	2.12

7.3.3 QRNN architecture

Every neural network has an **architecture** describing its assembly pattern and regularization. Figure 43 visualizes our NN architecture. It contains both a *feed-forward* NN and a *recurrent* NN. The nodes in a feed-forward NN, also called a *multilayer perceptron (MLP)*, do not form a cycle. Nodes in a recurrent NN (rNN) form a directed graph evolving along a sequence. MLPs require input features to have the same length across all events, while rNNs can accept arrays of varying lengths. rNNs are also by design more sensitive to context - the same input element in a different sequence would produce a different output. We took advantage of an rNN to evaluate leptonic behavior to prepare a “summary value” for the

²⁵Folding the distribution in half simplifies the percentile calculation and technical implementation. Future analyses should use $m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}$.

lepton-level kinematics of the event. We can then combine that summary value along with the first layer MLP through a second MLP.

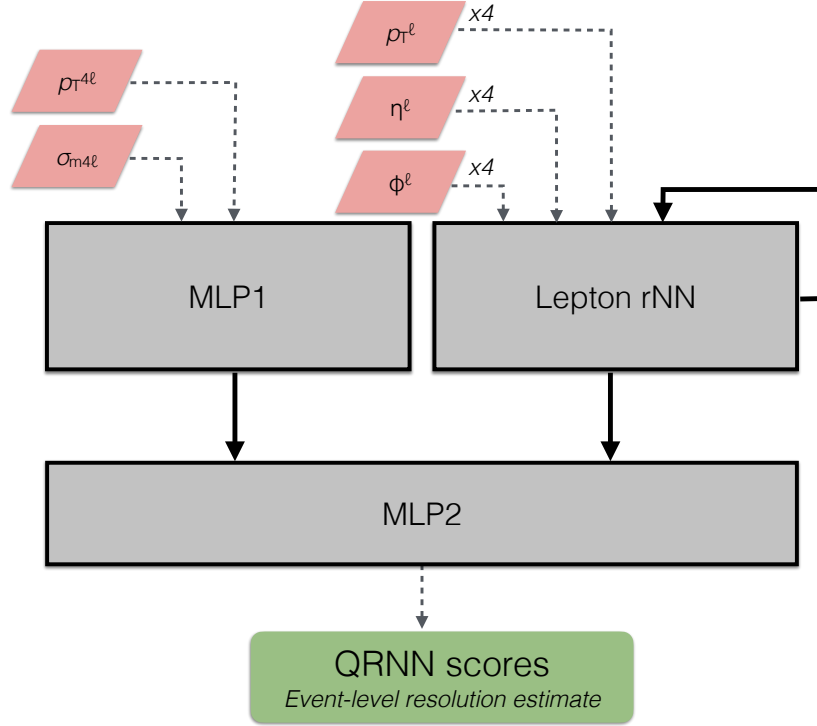


Figure 43: Quantile regression neural network’s (QRNN) architecture

Quantile regression neural network’s architecture consists of two MLPs and one rNN, outlined in black. Solid black arrows indicate the information flow in the neural network. Red rhombuses represent the input features. Dashed gray arrows indicate that $p_T^{4\ell, \text{constrained}}$ and $\sigma_{4m\ell}$ feed MLP1, while the lepton kinematics are processed separately in a dedicated rNN for technical reasons. rNNs can handle arrays of varying lengths and some events, such as $t\bar{t}H$ candidates, may have more than four leptons. MLP2 processes the outputs of MLP1 and the rNN, producing the final QRNN scores (green box). *Source: HH (2020), adapted from the $H \rightarrow 4\ell$ mass measurement group*

7.3.4 QRNN training

We already know that the four final states of the Higgs boson’s decay approximate a resolution categorization: muons and electrons do not have the same resolution in the ATLAS detector. Therefore, we chose to derive our continuous resolution parameterization with a dedicated training for each final state.

We train our regressor on simulated ggH samples generated across a range of Higgs mass

points (123, 124, 124.5, 125, 125.5, 126, and 127 GeV) to protect it from assumptions about the true mass of the Higgs boson.

By default, deep quantile regressors would aim to replicate the 68th percentile of target distribution. Developers made this design choice because they hypothesized an underlying Gaussian distribution (a sensible guess in the context of the Central Limit Theorem) and so 68% captures the third quantile – the RMS of the distribution. The regressor itself does not make any fundamental assumptions about the distribution’s underlying shape. The default choice for the percentile alone reflects the Gaussian hypothesis.

We have chosen to use the regressor to parameterize the event resolution, and to use the event resolution as the width of our DCB, defined as the σ_{CB} parameter of the DCB described in equation 14. Yet, we have no reason *a priori* to believe that the 68th percentile of $|m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}|$ corresponds to the width of our DCB. In fact, we should expect the width of the DCB to be **less than** the width of a Gaussian distribution by construction, as DCB replaces a Gaussian’s tails with exponentials. We repeatedly sampled simulated events to create bootstrapped datasets and fit each set with the per-event resolution model described in 7.2.3, using event resolutions based on estimating the 68th percentile as the width of the DCB. Chapter IV includes a description of the sampling technique for creating these pseudoexperiments. We calculated the pull on m_H^{measured} for this ensemble of pseudodata:

$$\text{pull} = \frac{m_H^{\text{measured}} - m_H^{\text{true}}}{\sigma_{m_H}}, \quad (16)$$

where m_H^{measured} is the measured value of the Higgs mass for each event ensemble, m_H^{true} is the true value of the Higgs mass in the simulated dataset from which we sampled events, and σ_{m_H} is the uncertainty on the measured value of the Higgs mass for each event ensemble. A perfectly-calibrated event resolution (equivalently, QRNN scores produced when targeting the percentile consistent with the DCB width) would result in an RMS of this pull distribution equal to 1.0. We found that the RMS of the pull distribution was 0.92 when we used the

event resolutions based on targeting the 68th percentile, indicating that we overestimated the event resolution. In other words, the QRNN score typically exceeded the DCB width.²⁶ We developed a scan to determine the appropriate target percentile for each decay channel, listed in table 10. Appendix E describes the calibration procedure.

Table 10: Calibrated target percentile for each decay channel. When the QRNN targets the calibrated percentile, the resulting scores represent the width of the DCB, changing for each event. The exponential tails of the DCB model behavior outside the resolution core.

Channel	Calibrated target percentile [%]
4μ	60.22
$4e$	59.41
$2\mu 2e$	61.64
$2e 2\mu$	53.46

7.3.5 Event resolution performance

The generator-level (target) percentile is a property of a distribution of events, while the network itself produces a continuous parameterization of the resolution along the set of input variables. Validating the network performance requires defining specific regions of phase space for which we can calculate the target percentile of $|m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}|$. We can compare the calculated target percentile with the mean QRNN score, $\langle r_{QRNN} \rangle$, for events falling in the same region of phase space. The RMS of the QRNN score provides the spread of the parameterized event resolution in that bin of phase space.

Figure 44 is an example of such a single bin of phase space. For this example, all four leptons were required to have $|\eta| < 0.6$. The green histograms provide the distributions of $|m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}|$ in the training and testing samples. The agreement in the green histograms represents that the network avoids overtraining. The red dotted line indicates the position of the target quantile. The network’s goal is to recover the red dotted line. The blue histograms are the QRNN scores in the training and testing samples. Notice that

²⁶Figure 41 demonstrated that the $m_{4\ell}^{\text{constrained}}$ distribution is not Gaussian. The target distribution is just a folded, shifted version of the $m_{4\ell}^{\text{constrained}}$ distribution, so we anticipate it follows the same behavior.

they are distributed about the target quantile (red dotted line). The black line marks the average QRNN score. Each of the plots in this figure can be summarized with three values:

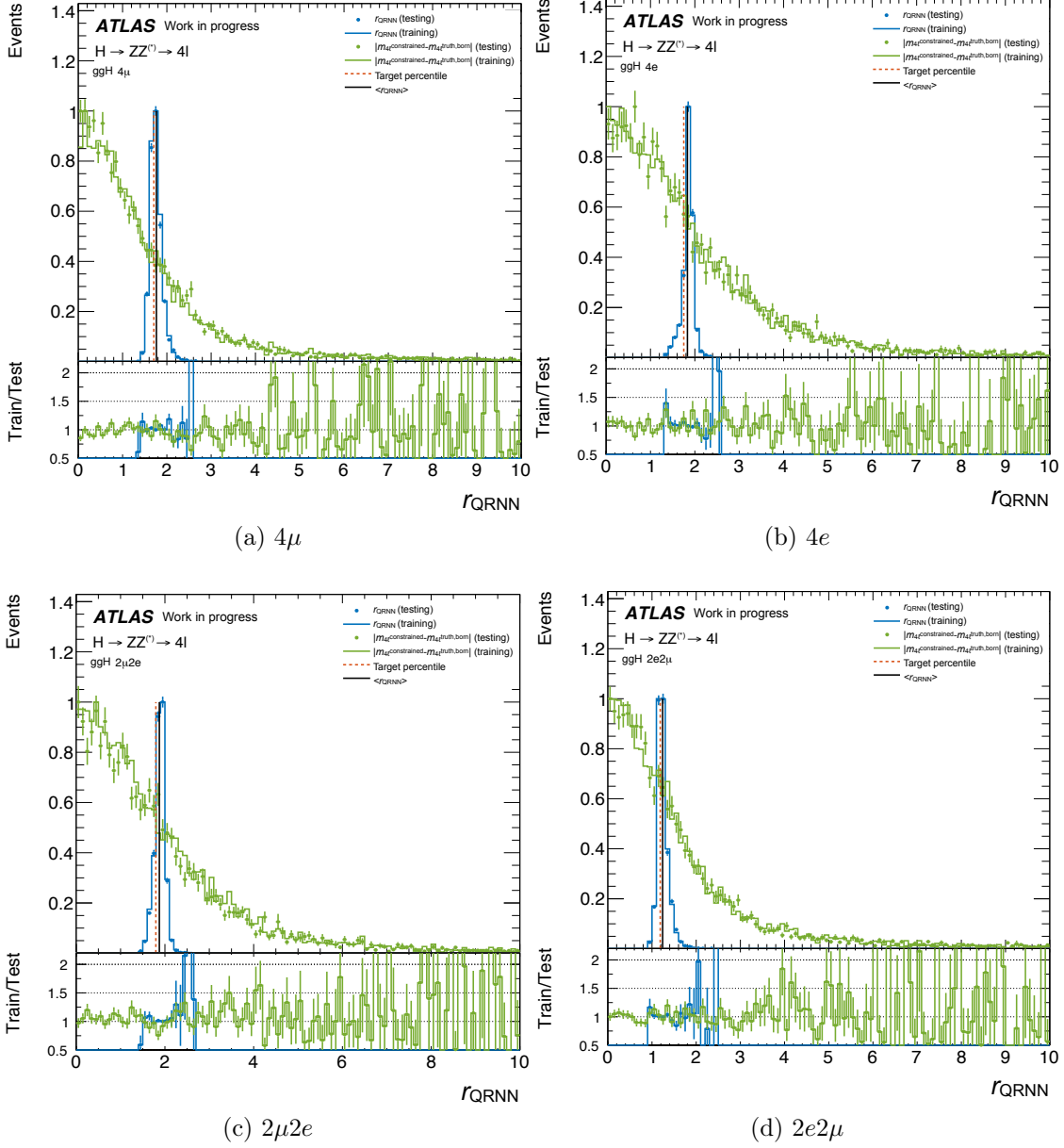
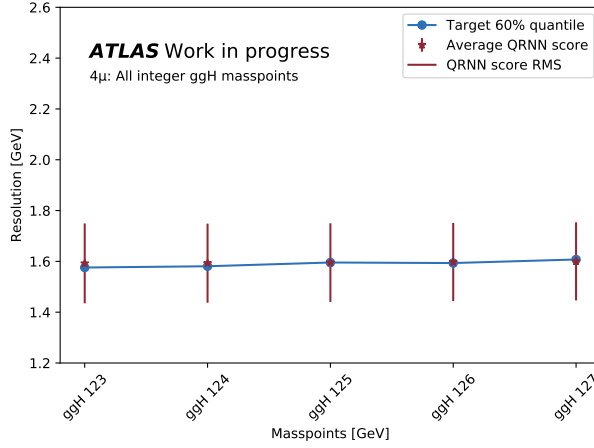


Figure 44: Validating QRNN performance in a single bin of phase space where all four leptons were required to have $|\eta| < 0.6$. The green histograms provide the distributions of $|m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}|$ in the training and testing samples. The agreement in the green histograms represents that the network avoids overtraining. The red dotted line indicates the position of the target quantile. The network's goal is to recover the red dotted line. The blue histograms are the QRNN scores in the training and testing samples. Notice that they are distributed about the target quantile (red dotted line). The black line marks the average QRNN score.

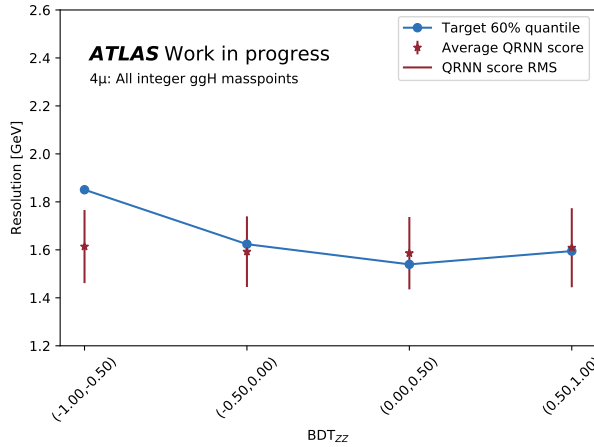
Source: $H \rightarrow 4\ell$ mass measurement group

the target quantile (red dotted line's position), the mean QRNN score (solid black line's position), and the RMS of the QRNN score (RMS of the blue histogram). Extracting those summary values in sequential bins of phase space – in slices of a given variable – permits creating summary plots (figures 45–47) displaying trends in replicating the target quantile.

The summary reports for the performance of the 4μ QRNN follows here, as an example. The rest are available in appendix F. The QRNN cannot model the percentile perfectly in some places as a function of all variables, likely a result of limited training events in certain phase spaces, ultimately limiting the depth of the network architecture.



Predicted resolution does not demonstrate a dependence on generated m_H value, indicating that the resolution estimate will not bias m_H^{measured} .



Very few signal events have low BDT_{ZZ} scores, limiting the training depth and causing the disagreement at low BDT_{ZZ} scores. In general, the predicted resolution has a some limited dependence on BDT_{ZZ} , especially for the dominant $ZZ^{(*)}$ background. A systematic uncertainty was added to cover this variation.

Figure 45: Summary estimated resolution (r_{QRNN}) performance plots as a function of generated mass point and BDT_{ZZ} for the 4μ channel. Each horizontal-axis position corresponds to a summary extracted from a figure 44-like plot. *Source: $H \rightarrow 4\ell$ mass measurement group*

Comparing the resolutions based on the locations of the four leptons in the ATLAS detector shows that we observe better resolution when all muons are in the barrel and worse resolution when they are not – as we would expect. Most muons in $H \rightarrow 4\ell$ events do fall in the barrel. The disagreement in regions with more muons falling in the endcap stems from few events with multiple muons in the endcap. We could improve the agreement in the regions where most muons fall in the barrel in the future using more input variables (for example, the number of hits in the SCT attributed to the muon’s track).

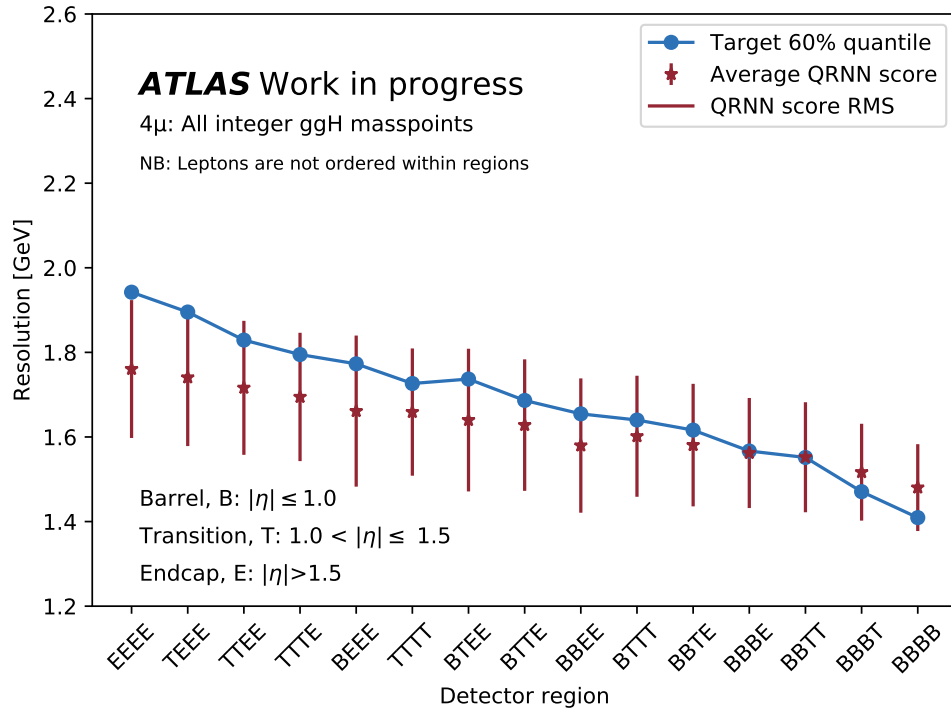
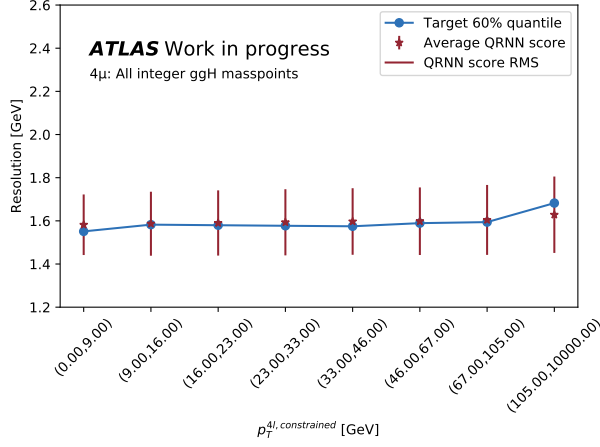
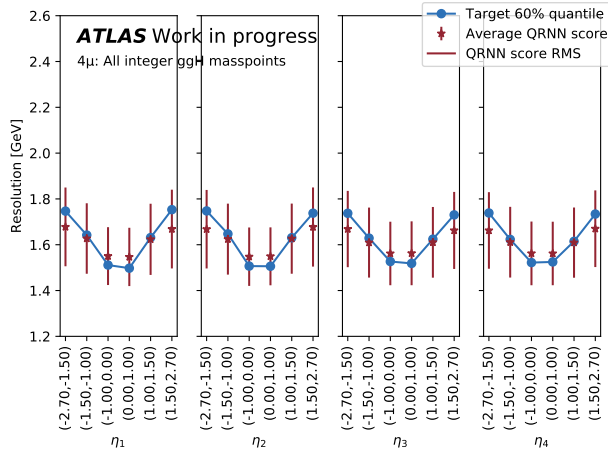


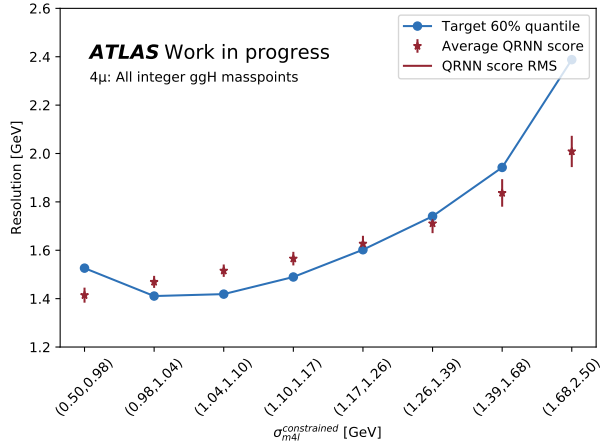
Figure 46: Summary estimated resolution (r_{QRNN}) performance plot as a function of detector region for the 4μ channel. The horizontal-axis position corresponds to a summary extracted from a figure 44-like plot. *Source: $H \rightarrow 4\ell$ mass measurement group*



Estimated resolution sliced in $p_T^{4\ell, \text{constrained}}$ shows good agreement. Further, the resolution is relatively flat, consistent with the single-muon expected resolution performance in this momentum regime (figure 18a).



The resolution along lepton pseudorapidity, η is broken into four subplots, one for each lepton. The estimated resolution improves when leptons are more central, in agreement with our detector performance knowledge.



Some disagreement is expected along $\sigma_{m_{4\ell}}^{\text{constrained}}$ since the error from the ZMC is not calibrated (and in some sense, the QRNN is calibrating it). The trends in both graphs coordinate, validating our performance.

Figure 47: Summaries comparing the target percentile (blue) and QRNN score (red) in slices of $p_T^{4\ell}$, lepton η , and $\sigma_{m_{4\ell}}$ for the 4 μ channel. Each horizontal-axis position corresponds to a summary extracted from a figure 44-like plot. *Source: $H \rightarrow 4\ell$ mass measurement group*

7.4 Signal model conditional on per-event resolution

Now we can generalize the DCB with an average resolution defined in section 7.2.3 to account for per-event resolution. We replace σ_{CB} with the per-event resolution, r_{QRNN}^i , derived at the calibrated target percentile for each event i . We then calibrate the rest of the DCB parameters conditional to the value of r_{QRNN}^i . Equation 14 becomes:

$$DCB(m_{4\ell}; \mu_{CB}, \sigma_{CB}, \alpha_1, n_1, \alpha_2, n_2) \rightarrow DCB(m_{4\ell}, r_{QRNN}^i; \mu_{CB}, \alpha_1, n_1, \alpha_2, n_2) \quad (17)$$

We still use a likelihood fit to extract m_H^{measured} from an ensemble of events with one change – we now take the value of r_{QRNN}^i from each event in the place of the average resolution, σ_{CB} . This approach constitutes the signal model for the full Run 2 measurement.

7.5 Background model

Unfortunately, not all of our events are signal events. About a third of our selected $m_{4\ell}$ candidate events originate from one of the background processes. Broadly speaking, there are two kinds of backgrounds in any analysis:

1. **Irreducible** backgrounds mimic our signal’s signature so exactly that we cannot design a traditional²⁷ selection to avoid them. We *can* reduce their abundance in a particular analysis category if we use multivariate techniques to train a classifier, as discussed in Section 7.1.1. Irreducible background sources for this analysis contain at least four prompt, isolated leptons in the final state. They include:

- Non-resonant $ZZ^{(*)}$ production at leading order ($qq \rightarrow ZZ^{(*)}$, $gg \rightarrow ZZ^{(*)}$),
- Electroweak $ZZ^{(*)}$ production (for example, vector boson scattering, $qq \rightarrow VV \rightarrow ZZ + jj$, where both Z bosons in the final state decay to e or μ),

²⁷A “traditional” or “classical” selection does not use machine learning techniques.

- tXX ($t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}t$, $t\bar{t}t\bar{t}t$, $t\bar{t}WZ$, $t\bar{t}ZZ$, $t\bar{t}WW$, $t\bar{t}Z\gamma$), VVV (triboson production: ZZZ , WZZ , WWZ).

2. **Reducible** backgrounds can be almost entirely eliminated using a traditional selection.

Reducible background sources in the 4ℓ final state include:

- $Z + jets$,
- $t\bar{t}$,
- WZ .

Their final states may feature non-isolated leptons, mis-identified leptons, or leptons produced during in-flight decays.

This section reviews how we construct a PDF for each background contribution. Sections 7.6.2 and 7.6.3 will discuss how we handle their normalizations. Figure 48 displays the expected background composition in each of the 16 analysis categories. Appendix G provides all plots of the extracted PDF shapes. References [35, 36] discuss the background samples in detail. Most of the background shape PDFs are $m_{4\ell}^{\text{constrained}}$ distributions extracted from simulations of each background process in each analysis category. We retrieve these templates using KEYS smoothing [61] with adaptive-width Gaussians to extract non-parametric, unbinned shapes.

The KEYS input parameter ρ controls the level of smoothing. Higher values of ρ produce smoother shapes at the cost of feature preservation. Lower values of ρ keep features while being much more susceptible to statistical fluctuations in the simulated dataset. The smoothing algorithm is also internally density-weighted. Regions with many events retain finer shape variations, while regions with lower statistics are protected from fluctuations. We tuned the specific values of ρ (table 11) using scans over ρ and visually comparing the output smoothed shapes for each background process.

KEYS smoothing also needs to know how to handle boundaries in the observable of the input dataset. Ordinarily, the observable’s value outside the requested smoothing range

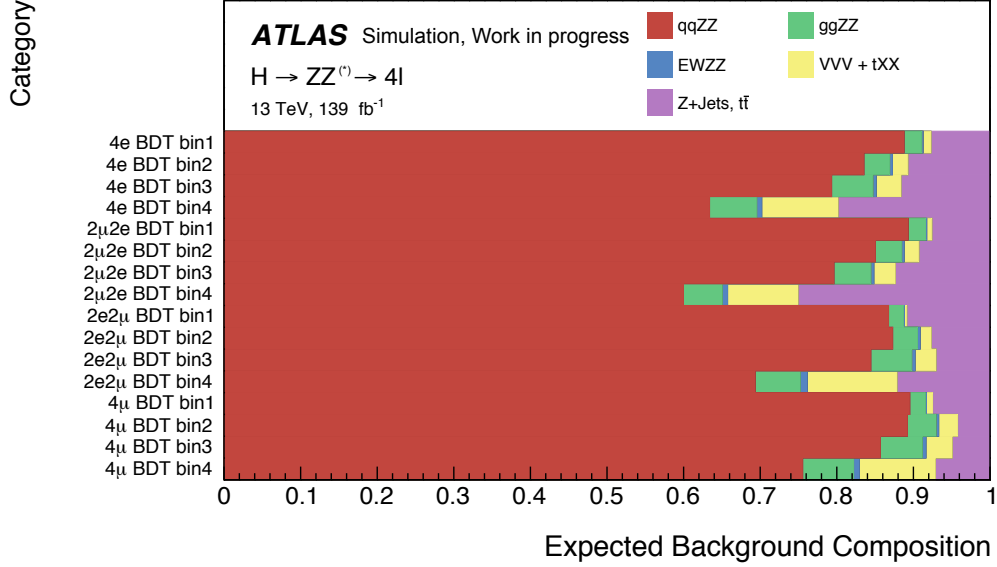


Figure 48: Expected background composition in each analysis category. $qq \rightarrow ZZ^{(*)}$ is the most dominant background in every category. Notably, its share of the composition increases in the lower-scoring BDT_{ZZ} bins. We expect this because we specifically train the classifier to distinguish between ggH and $qq \rightarrow ZZ^{(*)}$. *Source: $H \rightarrow 4\ell$ mass measurement group*

is treated as zero. The resulting PDF could have unphysical edge effects if the process it describes does not already fall to zero when approaching the boundary. **Mirroring** reflects the observable’s behavior over the boundary to mitigate the edge behavior.

In the future, we could improve this analysis if we used a per-event approach for the background model as well. The gains would be very small – the non-resonant $ZZ^{(*)}$ continuum backgrounds in $H \rightarrow 4\ell$ have a similar resolution to that of signal processes (figure 49). We begin to discern stronger difference in resolution when we compare the distributions in BDT_{ZZ} bins, subdividing further than the four bins used for this analysis (figure 50). If future analyzers wish to implement a per-event resolution background, it would be prudent to invest in a continuous treatment of BDT_{ZZ} at the same time.

7.5.1 $qq \rightarrow ZZ^{(*)}$ shape PDF

Three simulated samples contribute to modeling the $qq \rightarrow ZZ^{(*)}$ process in the range $105 < m_{4\ell}^{\text{constrained}} < 160$ GeV. One sample filters for events in a low-mass region, another in a high-

Table 11: KEYS smoothing parameter, ρ , for each background process.
mass measurement group

Source: $H \rightarrow 4\ell$

Process	Decay channel	BDT_{ZZ} bin	ρ
$qqZZ$	4μ	-1,-0.5	1.5
	$4e$	-1,-0.5	1.5
	$2\mu 2e$	-1,-0.5	1
	$2e 2\mu$	-1,-0.5	1.5
	4μ	-0.5,0	2
	$4e$	-0.5,0	2
	$2\mu 2e$	-0.5,0	1.5
	$2e 2\mu$	-0.5,0	2
	4μ	0,0.5	2
	$4e$	0,0.5	2
	$2\mu 2e$	0,0.5	2
	$2e 2\mu$	0,0.5	2
	4μ	0.5,1	1.5
	$4e$	0.5,1	2
	$2\mu 2e$	0.5,1	2
	$2e 2\mu$	0.5,1	1.5
$ggZZ$	All	All	1.5
$EWZZ$	All	All	1.5
tXX, VVV	All	All	2
Reducible	All	All	2

mass region. The third sample is generated inclusively. Duplicate events shared between the three samples are removed from two samples to retain a single copy of the event. Each sample has different statistics, so we cannot smooth the samples with the same ρ simultaneously. Instead, we create a KEYS PDF in each category for each sample from the $m_{4\ell}^{\text{constrained}}$ values for all events passing the event and category selection. We then add the three resulting PDFs together according to their relative expected event yields, taken from simulation, to produce the $qq \rightarrow ZZ^{(*)}$ PDF. We tuned ρ in each analysis category (decay channel and BDT_{ZZ} bin) in the smoothing range and enabled mirroring.

7.5.2 $gg \rightarrow ZZ^{(*)}$ shape PDF

We built the $gg \rightarrow ZZ^{(*)}$ PDF from two simulated samples using KEYS smoothing. We created a KEYS PDF in each category from the $m_{4\ell}^{\text{constrained}}$ values for all events passing

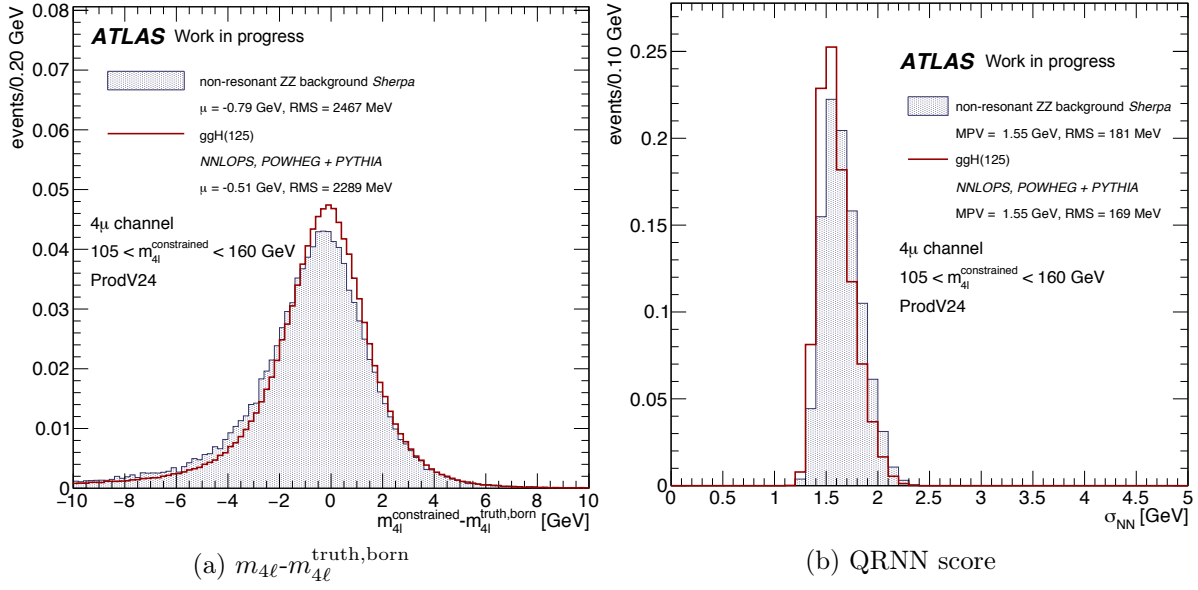


Figure 49: Distributions of event-level resolution in the 4μ final state, inclusive of BDT_{ZZ} , in ggH signal and $ZZ^{(*)}$ background samples. (a) relies on generator-level information while (b) uses the QRNN resolution estimate described in section 7.3. *Source: $H \rightarrow 4\ell$ mass measurement group*

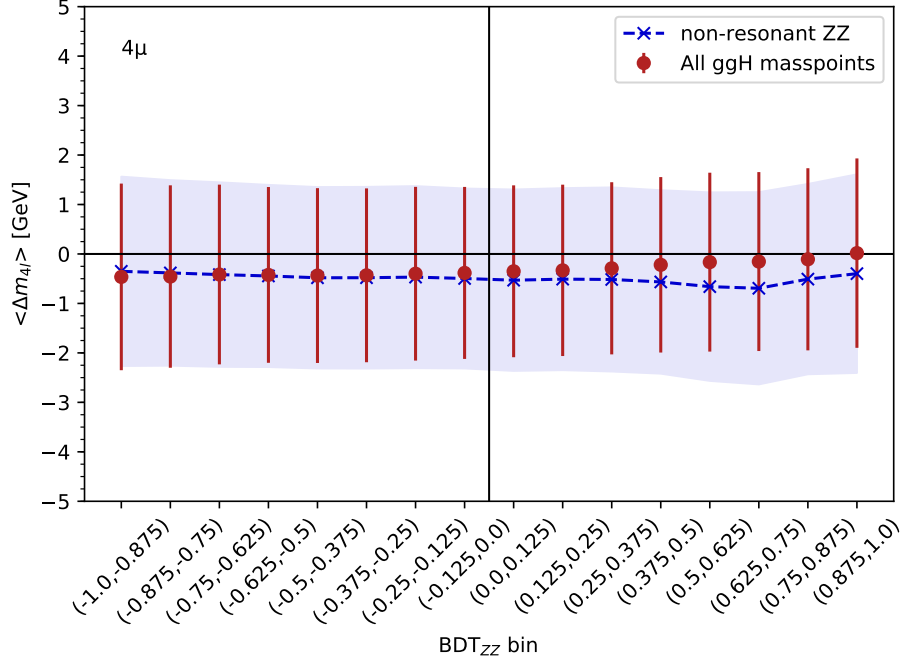
the event and category selection. We tuned ρ in each analysis category (decay channel and BDT_{ZZ} bin) in the smoothing range 105-160 GeV and enabled mirroring.

7.5.3 Electroweak $ZZ^{(*)}$ shape PDF

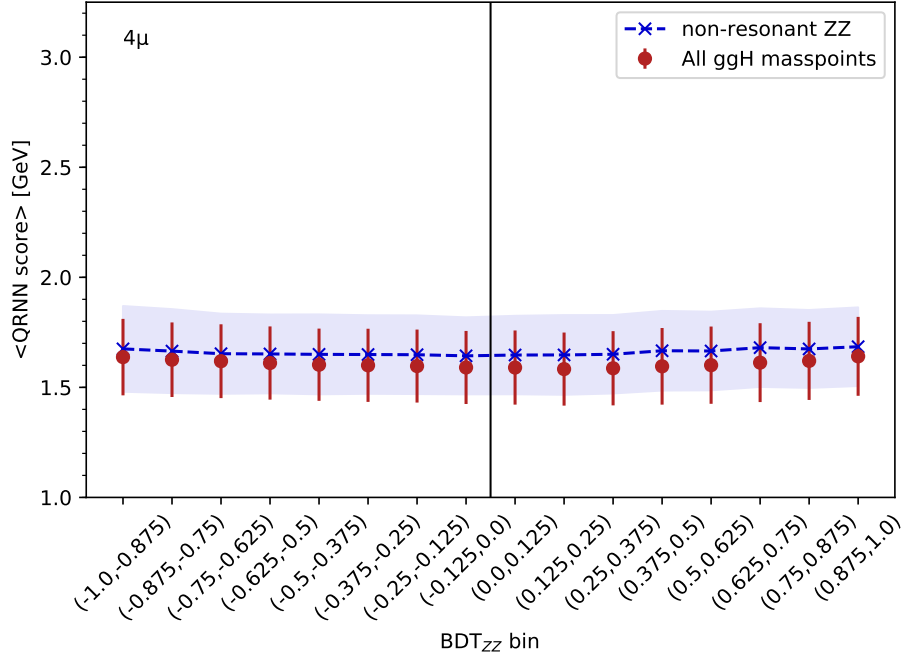
We built this PDF using KEYS smoothing simultaneously across all contributing simulated samples. We created a KEYS PDF in each category from the $m_{4\ell}^{\text{constrained}}$ values for all events passing the event and category selection. We tuned ρ in each analysis category (decay channel and BDT_{ZZ} bin) in the smoothing range and enabled mirroring.

7.5.4 tXX , VVV shape PDF

We built this PDF using KEYS smoothing simultaneously across all contributing simulated samples. We created a KEYS PDF in each category from the $m_{4\ell}^{\text{constrained}}$ values for all events passing the event and category selection. We tuned ρ in each analysis category (decay channel and BDT_{ZZ} bin) in the smoothing range and enabled mirroring.



(a) Average $m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth, born}}$ as a function of BDT_{ZZ} score



(b) Average QRNN score as a function of BDT_{ZZ} score

Figure 50: We divided each BDT_{ZZ} analysis bin into four, then determined the resolution estimate for each bin for ggH (red dots) and $ZZ^{(*)}$ (blue crosses) as described in section 7.3.5. The red lines (blue shaded area) represent the RMS for the ggH ($ZZ^{(*)}$) resolution estimate in a given BDT_{ZZ} quarter bin. As the QRNN modeling improves in future analyses, it will become much more sensitive to the apparent divergence at high BDT_{ZZ} score observed in the top plot.

Source: *HH* (2020)

7.5.5 Reducible backgrounds shape PDF

We create a single shape summarizing the $Z + jets$, $t\bar{t}$, and WZ processes. Section 7.6.3 describes how we estimate the reducible background from observed data. Further information is available in references [35, 36]. Briefly, we extract normalizations for each component from observed data using control regions enriched in heavy-flavor jets, light-flavor jets, $t\bar{t}$, and light-flavor jets misidentified as electrons. We combine the simulated events passing the event and category selection criteria for the heavy-flavor jets, light-flavor jets, and $t\bar{t}$ components according to their data-driven estimates. Then, we add the light-flavor electron backgrounds using data from same-sign control regions, normalized to its data-driven estimate. Once we have combined all components, we use KEYS smoothing to build a PDF in each analysis category. We do not apply mirroring. Additionally, we remove simulated events with large Monte Carlo event weights to avoid biasing the resulting shapes.

7.6 Creating the full probability density function (PDF)

The PDF in a given analysis category is the sum of the signal PDF, weighted by its normalization, and all background PDFs, each weighted by their own normalizations. The full PDF is a simultaneous PDF over all analysis categories, comprising each analysis category's PDF.

7.6.1 Signal normalization

We treated the overall signal normalization in each analysis category as a multiplier on the expected yield per inverse femtobarn. The expected yield is the number of simulated events passing the event selection, mass window, and category selection for a given category. Categories within the same BDT_{ZZ} bin shared a common multiplier, so that the multiplier was inclusive across final states - totaling four multipliers, one per BDT_{ZZ} bin. The resulting model has four floating signal normalization multipliers in the final fit.

The per- BDT_{ZZ} -bin multipliers reduce dependence on theoretical assumptions about the Higgs boson’s decay in several ways. First, the expected number of signal events depends on the branching ratio for a Higgs boson decaying to ZZ^* , BR_{ZZ} . Near $m_H = 125$ GeV, BR_{ZZ} changes quickly. Second, BDT_{ZZ} takes the combined transverse momentum of the four leptons, $p_T^{4\ell}$ as an input. $p_T^{4\ell}$ is closely related to the transverse momentum of the intermediate Higgs boson, p_T^H . Current simulations mis-model the p_T^H spectrum, potentially inducing a bias up to 25 MeV on m_H . Allowing this multiplier to float in each BDT_{ZZ} bin breaks the dependence on p_T^H , while sharing the multiplier across final states within a common BDT_{ZZ} bin preserves the relative expected composition of the events.

7.6.2 Irreducible Background normalizations

We treated the overall normalization for the non-resonant $ZZ^{(*)}$ production ($qq \rightarrow ZZ^{(*)}$ and $gg \rightarrow ZZ^{(*)}$) in each analysis category as a multiplier on the expected yield per inverse femtobarn. We shared the multiplier across analysis categories within the same BDT_{ZZ} bin, so that we have four floating background normalization multipliers in the final fit. These shared multipliers complement the signal normalization multipliers.

We took the normalizations for the EWZZ background component and the $tXX + VVV$ background component from the expected yields per inverse femtobarn.

7.6.3 Reducible background normalization

The selection efficiencies for the reducible background (characterized by mis-identified leptons, leptons from in-flight decays, or non-isolated leptons) show poor agreement between simulation and observed data. These processes also tend to have larger theoretical uncertainties [35, 36]. For these two reasons, we estimated the normalizations for the PDF accounting for $Z + jets$, $t\bar{t}$, and WZ from observed data, as mentioned in 7.5.5.

We studied the background composition and shapes in control regions (CR) designed to be orthogonal to the signal region (SR) defined by the analysis selection. We then extrapolated

the expected background from the CR to the SR using the transfer factor method. We made separate estimations for $Z + \mu\mu$ and $Z + ee$ because the dominant background changes based on the flavor of the subleading dilepton pair. The full description is available in references [35, 36] and summarized here.

Z production associated with leptons from jets is the dominant background source in the $Z + \mu\mu$ final states. Jets may be heavy or light flavor in this scenario. Most $Z + jets$ events in this final state come from heavy hadrons decaying semi-leptonically (“heavy-flavor jets”). Additionally, in-flight decays of pions or kaons from light-flavor jets can also produce the Z ’s associated leptons (“light-flavor jets”). Top quark pair production ($t\bar{t}$) and diboson production (WZ) also create the $Z + \mu\mu$ final state.

We prepared four control regions (table 12) orthogonal to the signal region to estimate their contributions to the $Z + \mu\mu$ final state. We built a PDF of the distribution of the leading dilepton mass, m_{12} , from the simulated shapes for each process with normalization factors for each component. m_{12} is useful here because $Z + jets$ forms a resonance while $t\bar{t}$ does not. Then, we performed a simultaneous unbinned maximum likelihood fit of this PDF to observed data to extract the normalizations. We validate the normalizations of the $Z + \mu\mu$ background components after the fit using a fifth, higher-statistics CR. This CR is not orthogonal to the SR. We obtained this CR using the standard analysis section *without* the standard d_0 and isolation requirements on the subleading dimuon.²⁸ We then extrapolated the validated normalizations to the SR using transfer factors.

The $Z + ee$ background primarily consists of light-flavor jets misidentified as electrons. We formed one CR, $3\ell + X$, to estimate this background and another, $Z + X$, to help estimate the transfer factors. In this CR, the lower E_T electron in the subleading pair (nicknamed “ X ”) has relaxed selection requirements, while the rest must satisfy the usual 4ℓ requirements. The electrons in the subleading pair must have the same sign, so that this CR is orthogonal

²⁸The vertexing requirement is also not applied. We are looking for muons originating from in-flight decays of hadrons. Such muons are by necessity displaced from the primary vertex, so the vertexing requirement is inappropriate for this control region.

Table 12: $Z + \mu\mu$ control regions; matrix element selection is never applied when building control regions.

Control region	Enhances	Description
Inverted d_0	heavy-flavor jets	Leading dilepton: standard 4ℓ selection Subleading dimuon: inverted d_0 selection for at least one muon in the pair, no isolation selection No vertex requirement on quadruplet
Inverted isolation	light-flavor jets	Leading dilepton: standard 4ℓ selection Subleading dimuon: both muons pass d_0 selection; at least one fails standard isolation selection Vertex requirement applied to quadruplet
$e\mu + \mu\mu$	$t\bar{t}$	Leading dilepton: opposite charge, different flavor, passes standard 4ℓ requirements (prevents selecting Z decay) Subleading dimuon: No d_0 or isolation requirements, Same-charge and opposite-charge muons accepted No vertex requirement on quadruplet
Same-sign	Mixture of all sources	Leading dilepton: standard 4ℓ selection Subleading dimuon: no d_0 or isolation requirements, both muons must have the same charge

to the SR. Mis-identified electrons can originate in several ways. The most common are (1) light jets making deposits in the ECAL consistent with an electron and (2) semi-leptonic decays of heavy quarks. Photons may also change into electron pairs (“photon conversions”).

We created templates of the number of innermost pixel layer hits, n_{InnerPix} ²⁹, from simulation for each process in the $Z + X$ CR since it has more statistics. The variable n_{InnerPix} distinguishes between photon conversions and sources 1 and 2, as photons do not leave hits in the pixel detector. We then fit the observed distribution of n_{InnerPix} to the templates to extract each component’s yield.

²⁹Usually, this variable counts the number of hits on the IBL. If the particle passed a dead area of the IBL, then this variable counts the number of hits in the next-to-innermost pixel layer.

8 Systematic uncertainties and building the likelihood

Any measurement in high-energy physics has an inherent statistical uncertainty, resulting from our finite set of observations. Recalling figure 36, we can compare the Run 1 and mid-Run 2 2D template method results, extrapolated with increasing integrated luminosity, assuming no changes to the methods or calibrations of those papers. While both were two-dimensional fits in $m_{4\ell}$ and a signal-background classifier, the statistical uncertainty for the mid-Run 2 template method is significantly smaller than the projection of the Run 1 method based on increased luminosity anticipated. This demonstrates that we can reduce the statistical uncertainty when we improve our muon and electron calibrations and develop more clever approaches to our analysis, as we did by mid-Run 2.

Systematic uncertainties further reduce the precision of our measurements. They reflect the quality of our data collection, calibrations, and our assumptions. In this analysis, where our goal is to determine the central value of a resonant mass, we anticipate the uncertainty of our decay products' energy scales to play the largest role. That is, muon momentum scale and electron energy scale uncertainties should dominate the systematic component of our total measurement uncertainty. ATLAS Combined Performance groups estimate these uncertainties when calibrating muon momentum and electron energy, and their derivation is available in references [37] and [40] respectively. We must also consider the uncertainties of the muon momentum and electron energy resolutions. Other possible sources of systematic uncertainty for the $H \rightarrow 4\ell$ mass measurement include any uncertainty in lepton identification, reconstruction, and isolation efficiencies; luminosity; modeling the number of interactions per bunch crossing (“pileup reweighting”); simulation statistics; and theoretical assumptions or modeling. Additionally, our choices introduced two additional systematic uncertainty sources. First, figure 50 shows that $m_{4\ell}$ resolution of $ZZ^{(*)}$ processes changes as a function BDT_{ZZ} score, hinting that the $ZZ^{(*)}$ distribution shape may actually change within a single BDT_{ZZ} bin. We introduce a systematic uncertainty for any such shape

variation. Second, we must account for any bias, called *non-closure*³⁰, in our DCB model, estimated using pseudoexperiments (see the description for sampled Monte Carlo datasets in chapter IV) to be 6 MeV. Ultimately, we represent each systematic uncertainty in the likelihood function with a nuisance parameter and PDFs describing its effect on the signal and background models.

Most systematic uncertainties were negligible. For example, the luminosity uncertainty only impacts backgrounds with normalizations estimated from simulation like tXX and VVV (as the others take their normalizations from data) so its total impact was less than 3 MeV. This section reviews the implementation scheme for the dominant systematic uncertainties: the lepton-related uncertainties and the $ZZ^{(*)}$ shape uncertainty as a function of BDT_{ZZ} bin.

8.1 Lepton uncertainties

When generating Monte Carlo simulated samples to determine the expected behavior of our models, we built a set of nominal samples from the central values of the lepton calibrations, as well as two additional sets with the calibration values set at their $\pm 1\sigma$ variations. We estimated the impact on the signal model by calibrating DCB models to these up and down variation samples for the ggH and VBF processes. We only allowed b_μ and the scale factor on the QRNN score, f , (the DCB width σ_{CB}) to vary when calibrating the per-event (average) resolution up and down variation DCB models.³¹ We then added the variations of each DCB parameter to the fit as Gaussian constraints on those parameters.³² For background PDFs, we produced sets of smoothed templates for the up and down variations for

³⁰Discrepancies discovered when testing a model with the validation procedure are referred to as non-closure

³¹We cannot reasonably expect the nominal, calibrated QRNN score, designed so that $f = 1$, to apply as well for the systematically shifted samples. Instead, we permit the floating scale factor, f , to avoid over-constraining the other parameters.

³²Each DCB parameter has a nominal value and the ability to take a value lying in the n-dimensional hyperspace defined by all up and down variations considered. We kept only those systematic uncertainties changing the total shape by more than 0.25%, as retaining them only changed the uncertainty by 1 MeV. Each systematic uncertainty adds another degree of freedom to the model, increasing the convergence time to successfully fit the model to data, so it is important to prune away systematic uncertainties without much influence.

each systematic uncertainty. Then, we morphed the templates across the full phase space defined by the systematic variations [62].

8.1.1 Muon momentum scale and resolution

The Muon Combined Performance group provided five distinct nuisance parameters to summarize the systematic uncertainty on the muon momentum calibration described in section 6.3. `MUON_SCALE` covers the variation in the muon momentum scale. `MUON_ID` and `MUON_MS` cover the variation in the momentum resolution in the Inner Detector and Muon Spectrometer, respectively. The last two parameters cover variations in the corrections for bias in the muon track curvature measurement from geometrical detector distortions that preserve the charged particle’s helical path, known as sagitta biases.³³ We considered these systematics for the shapes of both the signal and background models, as well as the background composition. Figure 51 gives an idea of the sizes of the up and down variations induced by `MUON_SCALE` on $m_{4\ell}^{\text{constrained}}$ and the QRNN score in the 4μ channel. `MUON_SCALE` behaves as a shift, imposing a variation of up to 10% on the signal samples with respect to the nominal sample.

Table 13 shows variation on the mean and RMS of $m_{4\ell}^{\text{constrained}}$ with `MUON_SCALE`. Considered inclusively, the difference between the nominal and down mean of $m_{4\ell}^{\text{constrained}}$ is 53 MeV and 76 MeV between the up mean and the nominal. There is no variation in the $4e$ channel, affirming that our systematics implementation was valid.

Table 13: Summary of the effects of the up and down systematic variations for muon momentum scale on the mean and RMS of distributions of $m_{4\ell}^{\text{constrained}}$ in each of the four channels in GeV.
Source: $H \rightarrow 4\ell$ mass measurement group

Channel	Nominal Mean	$\Delta(\text{Up} - \text{Nom})_{\text{mean}}$	$\Delta(\text{Down} - \text{Nom})_{\text{mean}}$	Nominal RMS	$\Delta(\text{Up} - \text{Nom})_{\text{RMS}}$	$\Delta(\text{Down} - \text{Nom})_{\text{RMS}}$
Inclusive	124.538	-0.076	0.053	2.595	-0.002	0.003
4μ	124.516	-0.129	0.091	2.275	-0.006	0.007
$2e2\mu$	124.610	-0.051	0.033	2.588	-0.005	0.003
$2\mu 2e$	124.450	-0.073	0.054	2.749	-0.002	0.001
$4e$	124.579	0.000	0.000	3.037	0.000	0.000

We expected the resolution-related nuisance parameters to show the largest effect on the

³³A description of the correction procedure is available in reference [18].

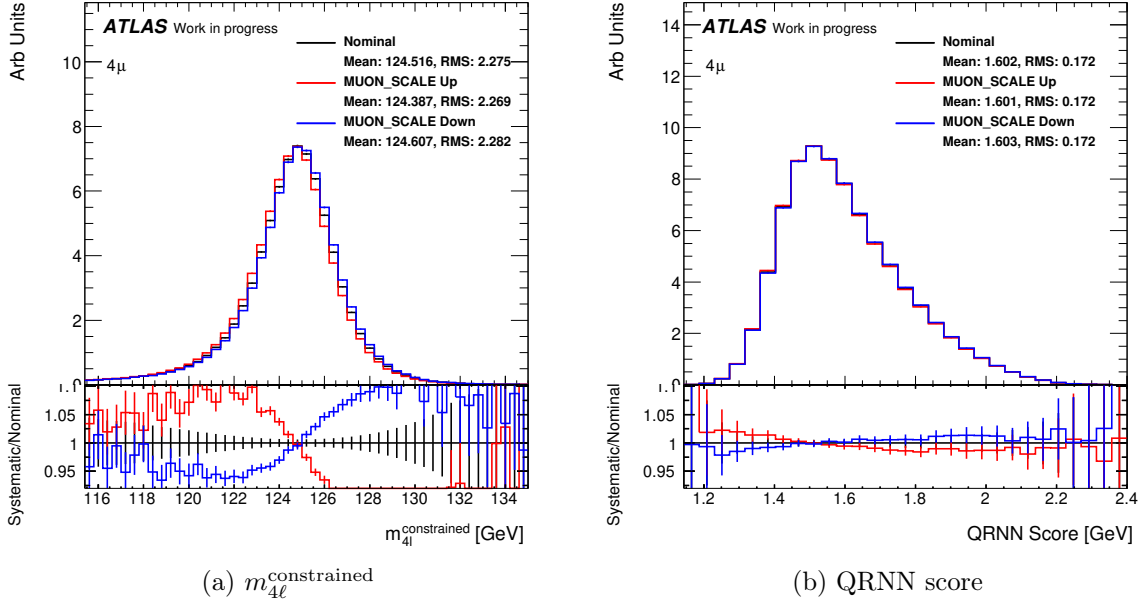


Figure 51: Impact of MUON_SCALE in the 4μ channel, inclusive of BDT_{ZZ} bin. *Source: $H \rightarrow 4\ell$ mass measurement group*

QRNN score, since it served as our estimate of the event resolution, but figure 52 shows that they induce variations of less than 1% in the core the QRNN distribution in the 4μ channel.

8.1.2 Electron energy scale and resolution

The E/Gamma Combined Performance group provided 74 distinct nuisance parameters to summarize the systematic uncertainty on the electron energy calibration described in section 6.4. The largest of these is the scale systematic coming from the non-closure in the $Z \rightarrow ee$ calibration, EG_SCALE_ZEESYST. Figure 31a visualizes an example of this non-closure. We considered all these systematics for the shapes of both the signal and background models, as well as the background composition. Figure 53 gives an idea of the sizes of the up and down variations induced by EG_SCALE_ZEESYST on $m_{4\ell}^{\text{constrained}}$ and the QRNN score in the $4e$ channel. EG_SCALE_ZEESYST also behaves as a shift, imposing a variation of up to 10% on the signal samples with respect to the nominal sample. Table 14 shows variation on the mean and RMS of $m_{4\ell}^{\text{constrained}}$ induced by EG_SCALE_ZEESYST. There is a ~ 22 MeV difference

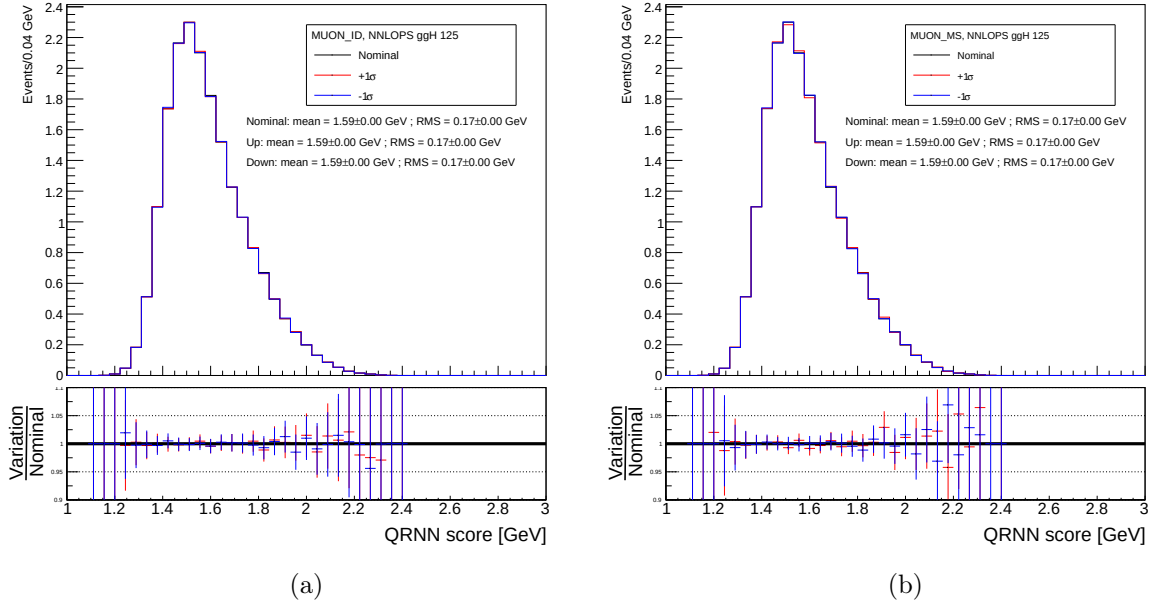


Figure 52: Impacts of resolution nuisance parameters MUON_ID and MUON_MS on the QRNN score. *Source: HH (2020)*

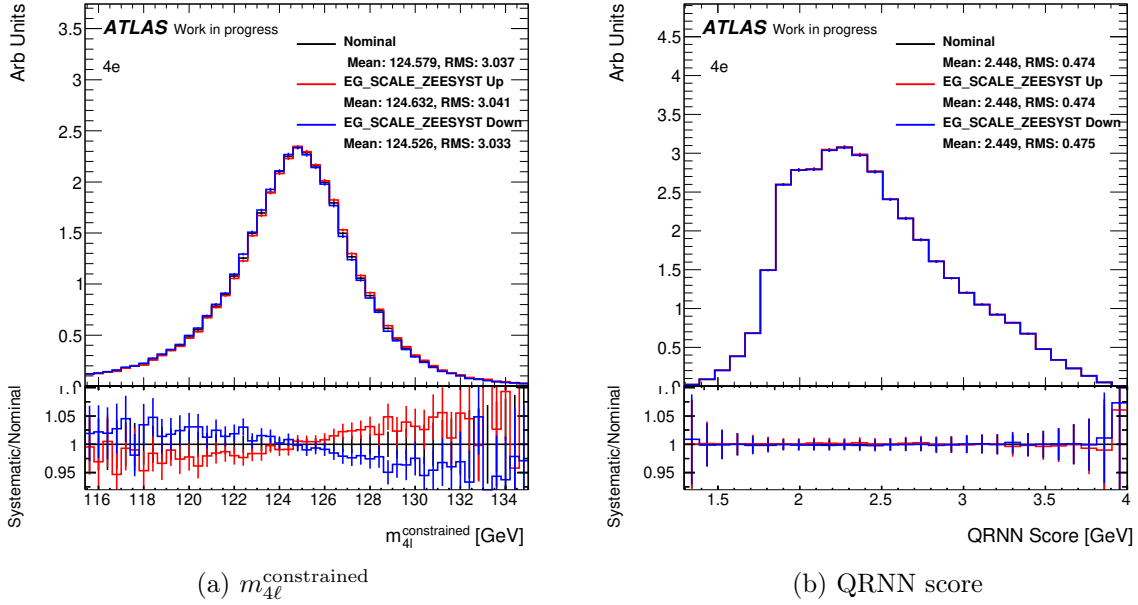


Figure 53: Impact of EG_SCALE_ZEESYST in the $4e$ channel, inclusive of BDT_{ZZ} bin. *Source: $H \rightarrow 4\ell$ mass measurement group*

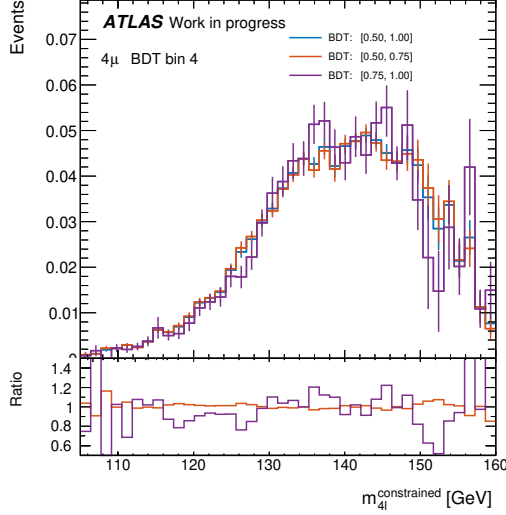
on the inclusive mean of $m_{4\ell}^{\text{constrained}}$. There is no variation in the 4μ channel, affirming that our systematics implementation was valid.

Table 14: Impact of largest EG scale systematic uncertainty coming from the $Z \rightarrow ee$ non-closure
Source: $H \rightarrow 4\ell$ mass measurement group

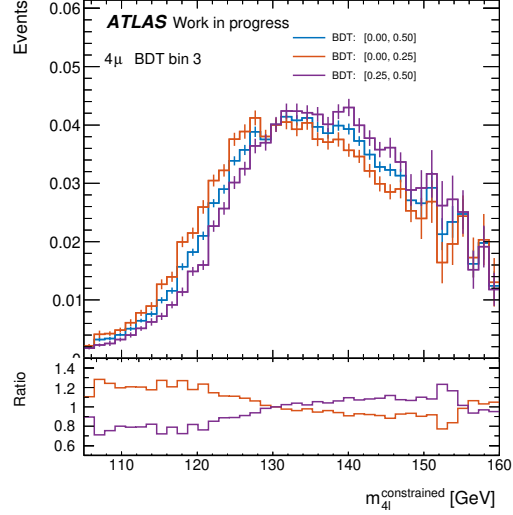
Channel	Nominal Mean	$\Delta(\text{Up} - \text{Nom})_{\text{mean}}$	$\Delta(\text{Down} - \text{Nom})_{\text{mean}}$	Nominal RMS	$\Delta(\text{Up} - \text{Nom})_{\text{RMS}}$	$\Delta(\text{Down} - \text{Nom})_{\text{RMS}}$
Inclusive	124.538	0.022	-0.022	2.595	0.003	-0.002
4μ	124.516	0.000	0.000	2.275	0.000	0.000
$2e2\mu$	124.610	0.032	-0.032	2.588	0.003	-0.001
$2\mu2e$	124.450	0.022	-0.023	2.749	0.003	-0.003
$4e$	124.579	0.053	-0.053	3.037	0.005	-0.004

8.2 $qq \rightarrow ZZ^{(*)}$ shape uncertainty within BDT_{ZZ} bins

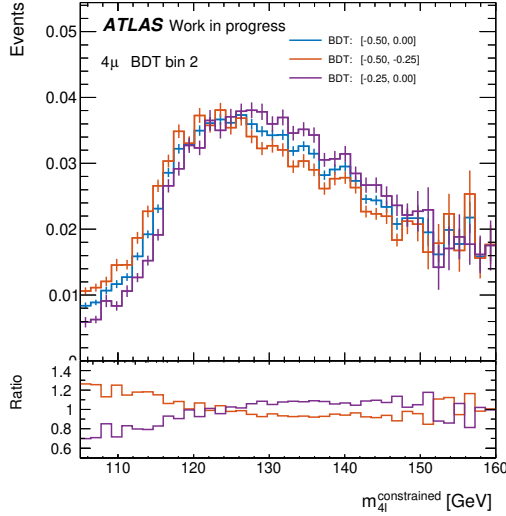
Figure 54 confirms our suspicions that the $qqZZ$ shape is indeed changing within the BDT_{ZZ} bins selected for the analysis. We created a PDF of the variation in each analysis bin by extracting templates of the events that pass the first and last quarter cut of each BDT_{ZZ} bin. The figure includes those distributions in red and purple. We smoothed these templates and separated their individual effects, then associated a nuisance parameter with each BDT_{ZZ} bin.



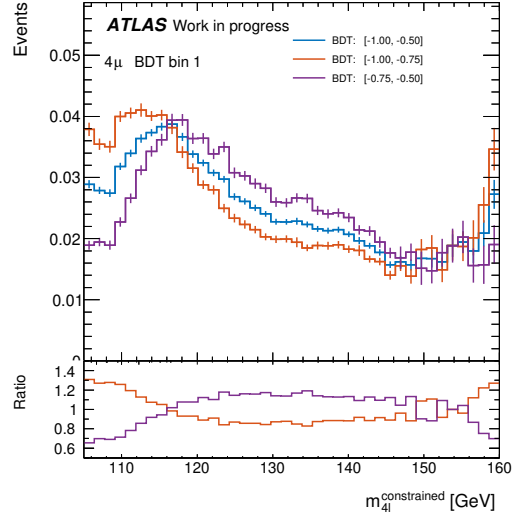
(a) BDT_{ZZ} bin 4 (most signal-like)



(b) BDT_{ZZ} bin 3



(c) BDT_{ZZ} bin 2



(d) BDT_{ZZ} bin 1 (most background-like)

Figure 54: $m_{4\ell}^{\text{constrained}}$ distributions in the $qq \rightarrow ZZ^{(*)}$ non-resonant background in the 4μ final state. Each plot shows the shape within the BDT_{ZZ} bin boundaries used in the analysis (blue) and in half-bins (purple and red). The red and purple histograms reflect how the $qqZZ$ shape changes within a single BDT_{ZZ} bin, and the variation is covered with a systematic uncertainty in this analysis. Plots for all analysis categories are available in appendix A of the source.

Source: $H \rightarrow 4\ell$ mass measurement group

8.3 Defining the likelihood function and test statistic

This section describes constructing the likelihood function and test statistic for extracting m_H from the datasets. The likelihood construction for the average and per-event resolution models is very similar. This section presents the likelihood function in the context of the per-event resolution model with the event resolution, r_{QRNN}^i , as the conditional observable. Converting to the average resolution model requires removing the conditional observable, r_{QRNN}^i , and replacing it with σ_{CB} produced when calibrating the DCB signal model unconditionally. Detailed references on these methods include [14] and [63].

8.3.1 Likelihood

We built an extended likelihood function using the expected yields from source s in category c , identified as $N_{s,c}^{\text{exp}}$, and the PDFs along $m_{4\ell}^{\text{constrained}}$ for each source and category as described in section 7. A subscript H indicates a (Higgs) signal model in the likelihood function expression. We assume that the total number of events is Poisson-distributed, indicated with $\text{Pois}()$ in equation 18. The full likelihood function also includes the systematic uncertainties represented with nuisance parameters. In the “stat-only” likelihood, these terms are omitted.

The extended likelihood function $P_{s,c}(m_{4\ell}^{\text{constrained}})$ depends on our parameters of interest (m_H , the signal and $ZZ^{(*)}$ overall normalizations) and the conditional observable identified with the resolution in each event, r_{QRNN}^i . The observed number of events in each category is N_c^{obs} , and the nuisance parameters describing the systematic uncertainties are denoted as θ . Using these components, the constructed likelihood function is:

$$\begin{aligned}
 L(m_H, \theta) = & \prod_c \text{Pois}(N_c^{\text{obs}}; N_c^{\text{exp}}(m_H, \theta)) \quad \times \\
 & \left\{ \prod_i^{events} N_{H,c}^{\text{exp}}(m_H, \theta) P_{H,c}(m_{4\ell,i}^{\text{constrained}}; m_H, r_{QRNN}^i, \theta) \quad + \right. \\
 & \left. \sum_c^{bkg} N_c^{\text{exp}}(\theta) P_{s,c}(m_{4\ell,i}^{\text{constrained}}; \theta) \right\} \quad (18)
 \end{aligned}$$

where the expected number of events in each category is the sum of the number of events in each of the individual sources,

$$N_c^{exp}(m_H, \theta) = \mu_{\text{signal}}^{BDTb} \cdot N_{H,c}^{exp}(m_H, \theta) + \mu_{ZZ^{(*)}}^{BDTb} \cdot (N_{qqZZ,c}^{exp}(\theta) + N_{ggZZ,c}^{exp}(\theta)) + \sum_s^{bkg} N_{s,c}^{exp}(\theta) \quad (19)$$

with the relevant parameters of interest $\mu_{\text{signal}}^{BDTb}$ for the overall signal normalization and $\mu_{ZZ^{(*)}}^{BDTb}$ for the $ZZ^{(*)}$ overall normalization in a BDT_{ZZ} bin b . The label “bkg” designates all background sources except $ZZ^{(*)}$.

8.3.2 Defining the test statistic

The test statistic used for fitting m_H is $q = -2 \ln \lambda(m_H)$, using the profile likelihood ratio

$$\lambda(m_H) = \frac{L\left(m_H, \mu_{\text{signal}}^{\hat{BDTi}}(m_H), \mu_{ZZ^{(*)}}^{\hat{BDTi}}(m_H), \hat{\theta}(m_H)\right)}{L\left(\hat{m}_H, \mu_{\text{signal}}^{\hat{BDTi}}, \mu_{ZZ^{(*)}}^{\hat{BDTi}}, \hat{\theta}\right)} \quad (20)$$

where the circumflex indicates the best fit value with all other parameters left free. The parameter $\hat{\theta}(m_H)$ is the best fit value of the nuisance parameters at a given m_H . The parameters $\mu_{\text{signal}}^{\hat{BDTi}}(m_H)$ and $\mu_{ZZ^{(*)}}^{\hat{BDTi}}(m_H)$ are the best fit values of the signal and $ZZ^{(*)}$ overall normalizations in each BDT_{ZZ} bin i at a given m_H . Uncertainties are set in the asymptotic limit, with $\pm 1\sigma$ ($\pm 2\sigma$) set where $q = 1.00$ (3.84). All plots of the negative-log-likelihood show the test statistic, $q = -2 \ln \lambda(m_H)$, on the vertical axis and $m_{4\ell}^{\text{constrained}}$ on the horizontal axis. We identify the measured value of m_H by minimizing the test statistic.

Chapter IV

Results

This chapter describes the results of measuring the Higgs boson’s mass, m_H , using the experimental method described in chapter III. We processed four types of datasets to extract these results:

Asimov dataset: A pseudoexperiment generated from the model where all observed quantities (parameters of interest, nuisance parameters) are set equal to their expected values. In this case, we generated an Asimov dataset for the template method with $m_H = 125$ GeV, $\mu_{signal}^{BDT*} = 1$, and $\mu_{ZZ(*)}^{BDT*} = 1$. If we had included systematic uncertainties, then all of their associated nuisance parameters would have been set to 1. Asimov datasets essentially have infinite statistics - when we fit them, their parameters’ returned best-fit values should match their input values at the time of the dataset’s generation.

Sampled Monte Carlo (MC) dataset: It is not possible to build traditional Asimov datasets from conditional PDFs, like the per-event resolution PDF. By definition, Asimov sets reveal the behavior of the model if sampled for infinite statistics. In that limit, the distributions of conditional observables are well-represented by their average values alone, and we lose the value of the conditional approach. Instead, we create pseudoexperiments with randomized values of the conditional observable(s) and finite statistics reflective of our accumulated luminosity. Each set of pseudoexperiments is one “toy” dataset. We generate and fit an ensemble of toys to study the expected behavior of our PDF, producing distributions of best-fit parameter values. The distribution of fit errors on m_H^{measured} quantifies our expected uncertainty. There are multiple ways to sample the values of the conditional observable for each pseudoexperiment. For example, we could draw a value from a Gaussian distribution or simulated template

of our parameter. In this analysis, we draw the events directly from the nominal MC sample. We Poisson-fluctuate the number of events sampled per toy around the number of expected events. Single MC events have generator weights indicating how many “real” events they represent. To ensure proper sampling, we accept or reject each event selected for the toy dataset based on its position on the Cumulative Distribution Function of weights for that MC sample. When considering systematic uncertainties, we fluctuate the value of each nuisance parameter as well, assuming that each parameter can be modeled as a Gaussian centered on 1 with a width of the systematic variation.

Full MC dataset: Instead of sampling the simulated data, we can treat the complete ensemble of events as a single dataset.

Observed dataset: Data collected with the ATLAS Experiment from pp collisions with a center-of-mass energy of 13 TeV during LHC Run 2 (2015-2018). The integrated luminosity of this dataset is 139 fb^{-1} , accounting for trigger and data-quality requirements. The mean number of pp interactions per bunch crossing during this period was 33.7.

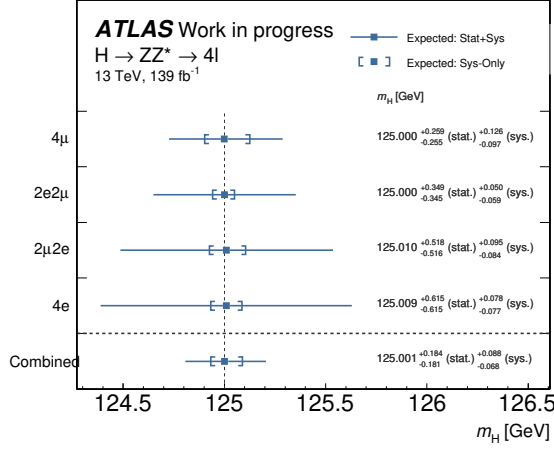
9 Expected results

This section reports the expected results using the average and per-event resolution methods for measuring the mass of the Higgs boson in ensembles of sampled MC datasets. The template method implementation and its results on an Asimov dataset are available in appendix H. Recall that the template method takes its PDF of $m_{4\ell}$ from simulation, while the average resolution method uses a parameterization for the $m_{4\ell}$ lineshape.

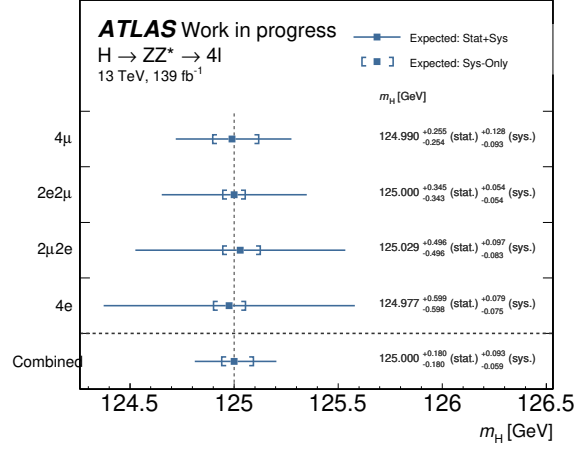
9.1 Full MC

Analyzing the full MC dataset allows us to interpret the influence of different systematic uncertainties. Ranking plots visualize the impact of each systematic uncertainty on m_H in a single fit, as well as the pull on each systematic uncertainty’s nuisance parameter. We cannot make ranking plots for sampled MC (toys) because they are ensembles - each its own fit (and so each one would have its own ranking plot). We fit all available MC statistics with the average resolution model. Unfortunately, it is very computationally intensive to fit the entire MC sample with per-event methods, so we only processed 10% of the MC statistics for the per-event resolution method. These per-event results should therefore be considered conceptually – noting whether the relative influence of key systematic uncertainties changes compared to the ranking when using an average event resolution. It is not appropriate to draw conclusions about the possible improvement from the per-event approach from this dataset.

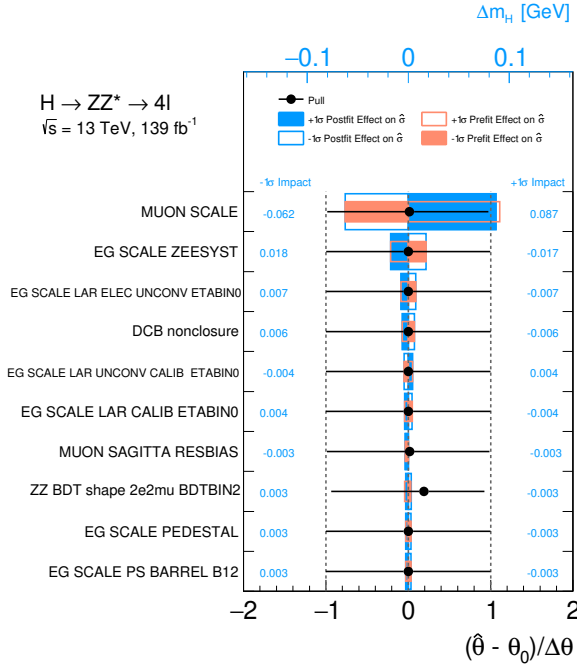
Figure 55 gives the expected Higgs mass in each final state and the combined measurement for the (a) average and (b) per-event resolution method. The bottom half displays the ranking plots for the (c) average and (d) per-event methods. The muon momentum scale dominates the systematic uncertainty for both methods, followed by the electron energy scale. Systematic uncertainties impacting m_H by less than 1 MeV are not included in the ranking plots, including the muon resolution systematic uncertainty.



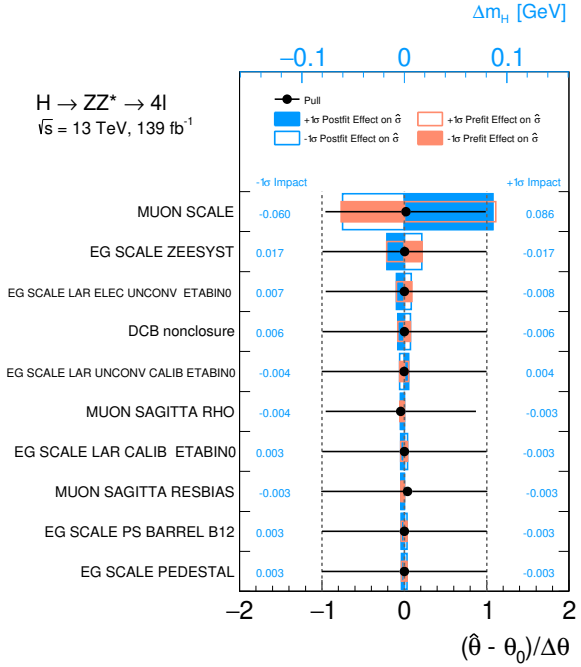
(a) Average resolution



(b) Per-event resolution



(c) Average resolution



(d) Per-event resolution

Figure 55: Expected m_H^{measured} in the full MC dataset for the (a) average resolution and (b) per-event resolution methods. Systematic uncertainty rankings for the (c) average resolution and (d) per-event resolution methods. It is too computationally intensive to process the entire MC dataset with the per-event resolution method so the results in (b) and (d) use 10% of the generated MC available.

Source: $H \rightarrow 4\ell$ mass measurement group

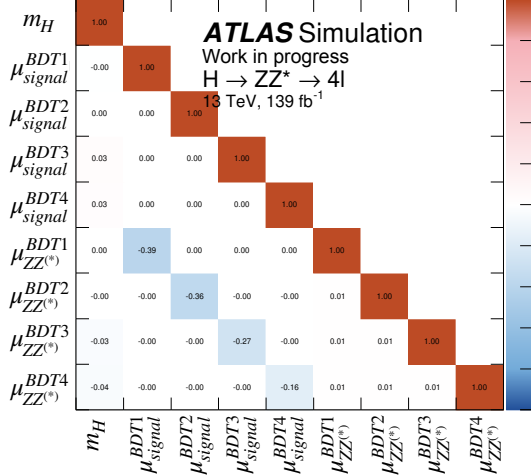


Figure 56: Correlation matrix of the parameters of interest obtained when fitting the full MC dataset with the average resolution model. *Source: $H \rightarrow 4\ell$ mass measurement group*

We can also use the full MC dataset to explore correlations (figure 56) among our parameters of interest. There are no notable correlations between m_H and the overall normalizations per BDT_{ZZ} bin. The μ_{signal}^{BDTi} and $\mu_{ZZ^{(*)}}^{BDTi}$ parameters are moderately inversely correlated because each event in each BDT_{ZZ} bin is either a signal or background event.

9.2 Sampled MC

We created sampled MC datasets following the description at the beginning of the section to probe the behavior of the average and per-event resolution methods. For each ensemble of toys, we obtain distributions of the parameters of interest. Figure 57 provides the summary results for the sampled MC fits for the (a) average and (b) per-event resolution methods.

The per-event resolution improves the expected statistical uncertainty on m_H by 1% (figure 58).

Sections 9.2.1 and 9.2.2 show the distributions of parameters of interest from fitting ensembles of toys sampled from MC. Including systematic uncertainties changes the likelihood by adding the associated Gaussian nuisance terms. Therefore results are presented in each figure (a) excluding the systematic uncertainties from the likelihood, colloquially known as a “stat-only” fit, and (b) including them.

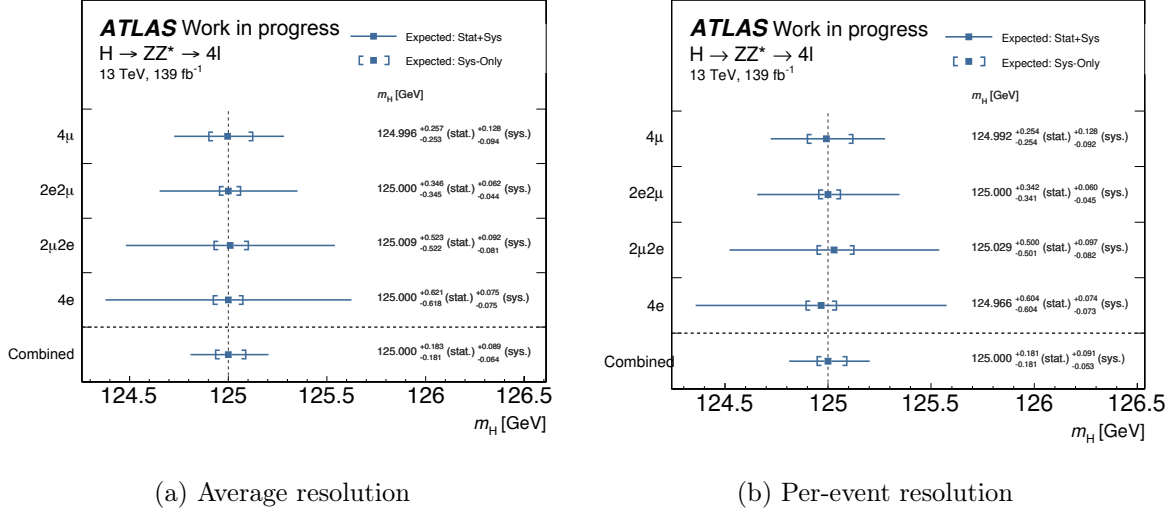


Figure 57: Expected value of m_H in the (a) average and (b) per-event resolution models when fitting sampled MC toys. *Source: $H \rightarrow 4\ell$ mass measurement group*

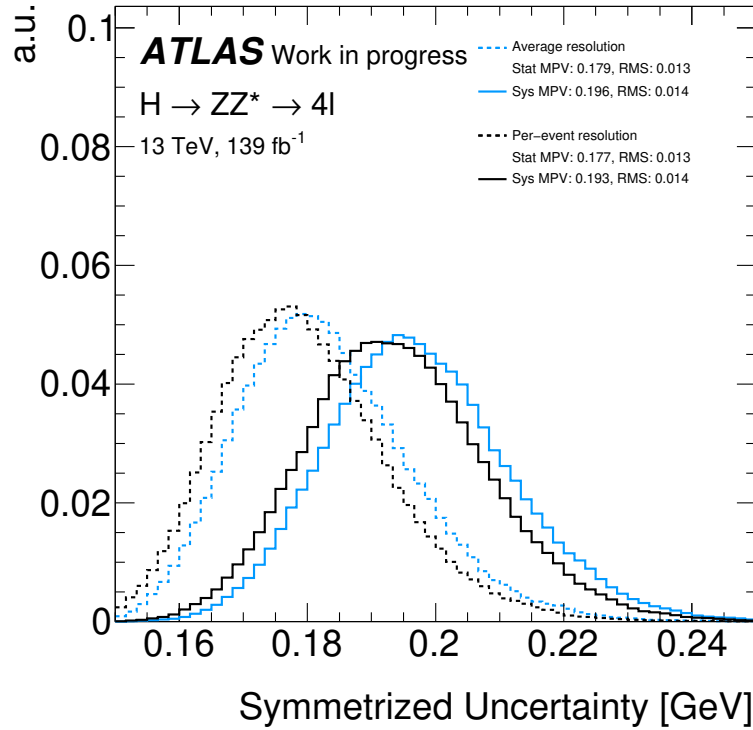


Figure 58: Distributions of the symmetrized statistical (black) and total (blue) uncertainty on m_H when fitting the sampled MC toys with the average and per-event resolution methods. Dashed lines indicate the distributions of the expected statistical uncertainty in the sampled MC toys, while the solid lines represent the total uncertainty. *Source: $H \rightarrow 4\ell$ mass measurement group*

9.2.1 Average resolution

The results from fitting an ensemble of toys sampled from MC with the average resolution model are included here. Figure 59 shows the distributions of best-fit values for m_H using the average resolution method. The fitted mean, identified with a Gaussian fit to the smoothed distribution, indicates any bias in the method. Figure 60 displays the distribution of upper and lower errors on m_H^{measured} from each toy and provides context for the quoted most-probable-value (MPV) errors in figure 59. Figure 61 shows the distribution of pulls, calculated in each event as the difference in generated m_H and the fitted m_H^{measured} of each toy divided by the symmetrized error on m_H^{measured} of that toy. Gaussian pulls, indicated by means = 0 and RMS = 1 with little-to-no skew or kurtosis, indicate that the method performs well.

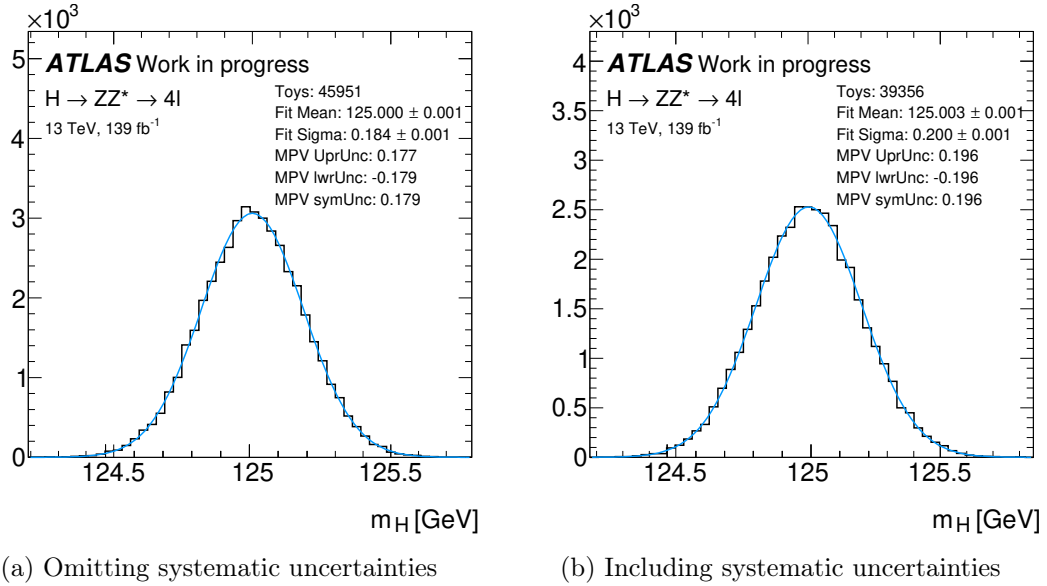


Figure 59: Distribution of best-fit values of m_H using the average resolution model. The expected value of m_H is extracted using a Gaussian fit to the smoothed distribution. *Source: $H \rightarrow 4\ell$ mass measurement group*

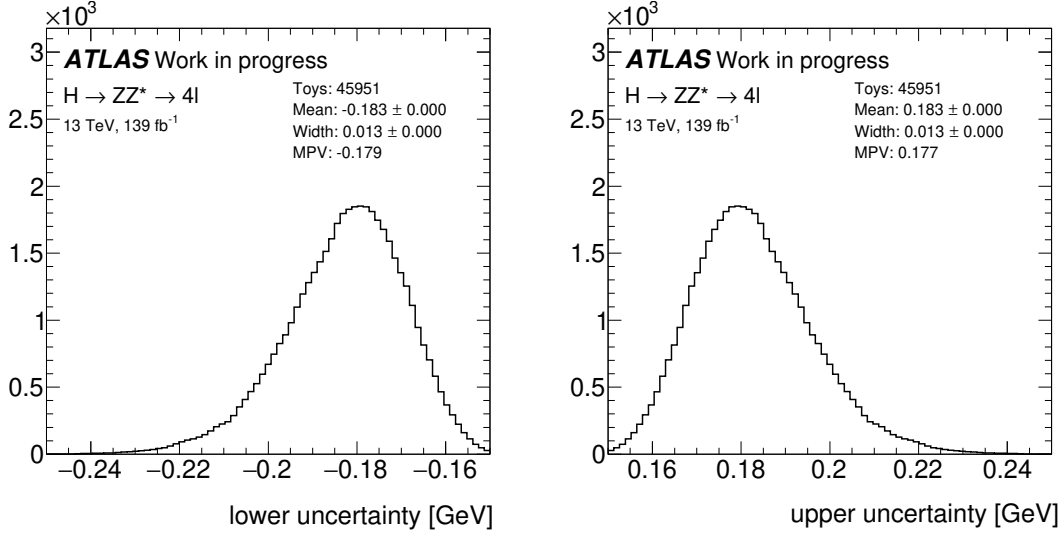


Figure 60: Lower and upper statistical uncertainties on m_H from each toy fit using the average resolution method, derived when omitting the systematics from the likelihood. The upper and lower errors are taken as the most probable value of these distributions when smoothed.
Source: $H \rightarrow 4\ell$ mass measurement group

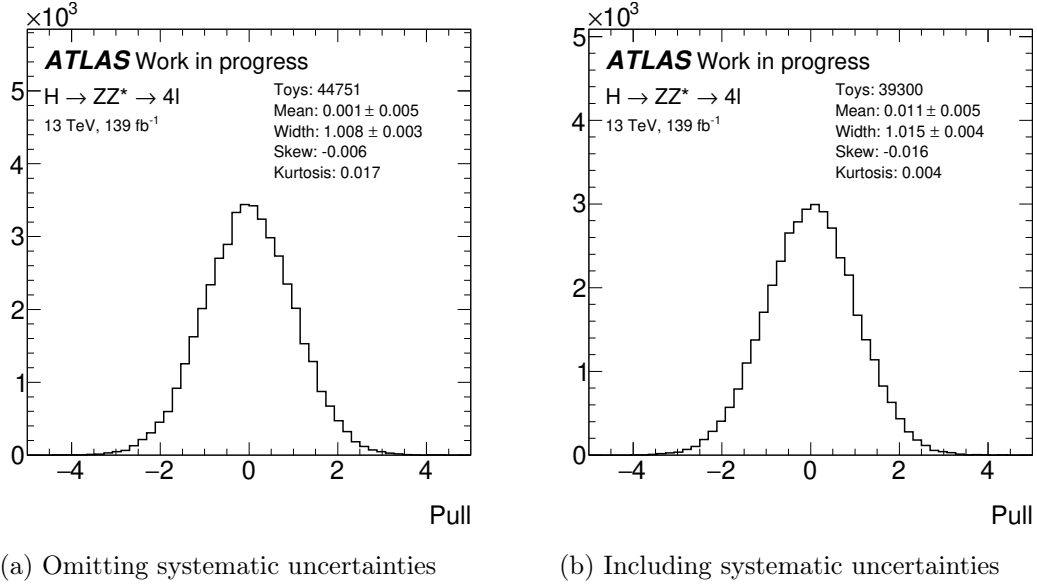


Figure 61: Pull on m_H^{measured} using the average resolution method
Source: $H \rightarrow 4\ell$ mass measurement group

9.2.2 Per-event resolution

The results from fitting an ensemble of toys sampled from MC with the per-event resolution model are included here. Figure 62 shows the distributions of best-fit values for m_H using the per-event resolution method. The fitted mean, identified with a Gaussian fit to the smoothed distribution, indicates any bias in the method. Figure 63 displays the distribution of upper and lower errors on m_H^{measured} from each toy and provides context for the quoted most-probable-value (MPV) errors in figure 62. Figure 64 shows the distribution of pulls on m_H^{measured} . Gaussian pulls, indicated by means = 0 and RMS = 1 with little-to-no skew or kurtosis, indicate that the method performs well.

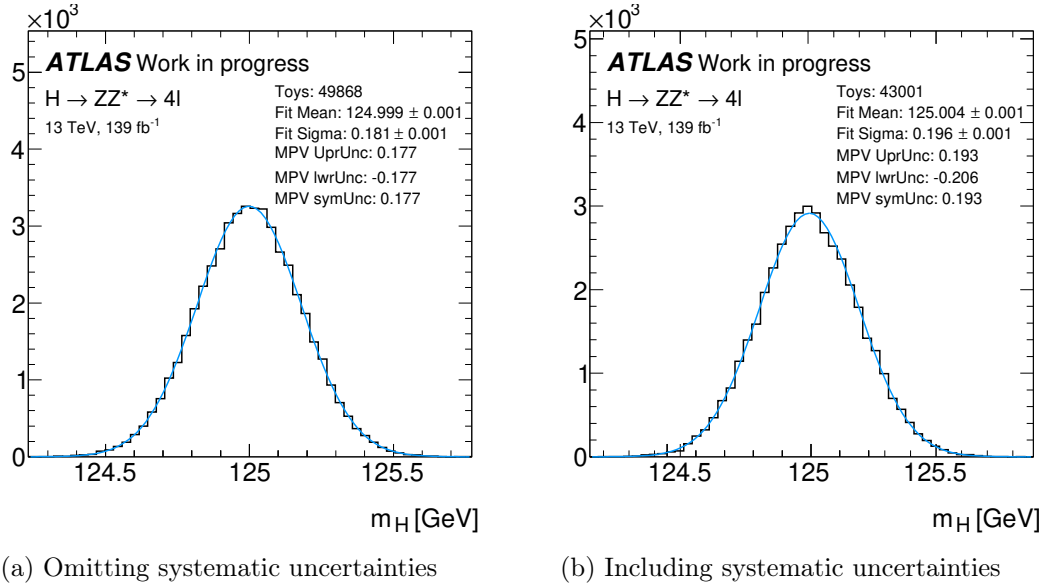


Figure 62: Distribution of best-fit values of m_H using the per-event resolution model. The expected value of m_H is extracted using a Gaussian fit to the smoothed distribution. *Source: $H \rightarrow 4\ell$ mass measurement group*

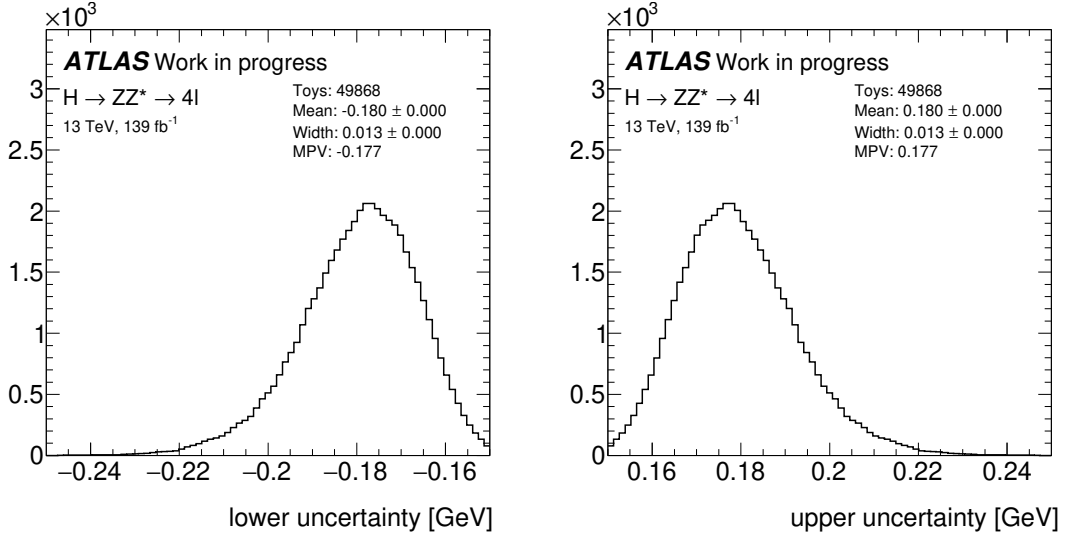


Figure 63: Lower and upper statistical uncertainties on m_H from each toy fit using the per-event resolution method, derived when omitting the systematics from the likelihood. The upper and lower errors are taken as the most probable value of these distributions when smoothed. *Source: $H \rightarrow 4\ell$ mass measurement group*

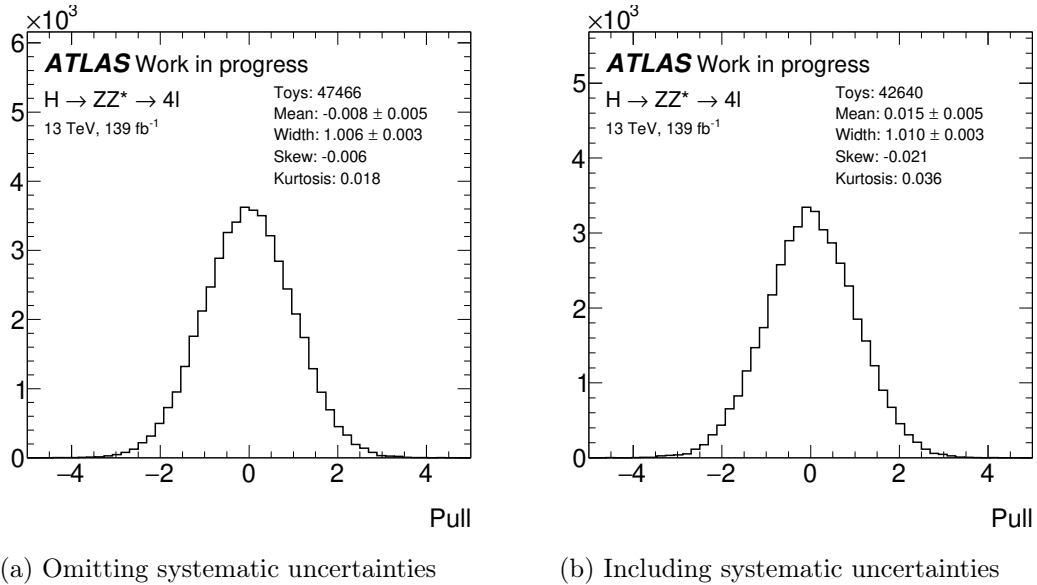


Figure 64: Pull on m_H^{measured} using the per-event resolution method. The widths of pulls ~ 1 , indicating that our QRNN-based resolution estimate is reasonable. *Source: $H \rightarrow 4\ell$ mass measurement group*

9.3 Summary

The expected value for m_H using the per-event resolution method is

$$m_H = 125.005 \pm 0.177 \text{ (stat.)} \pm 0.080 \text{ (sys.) GeV},$$

in agreement with a simulated Higgs mass of 125 GeV, demonstrating the method’s viability. The expected variation in the total (statistical) uncertainty for both methods is 0.014 (0.013) GeV, depicted in figure 58. Table 15 compares summarizes the impact of accounting for the per-event resolution with and without systematics.

Table 15: m_H value and uncertainty. Note that for per-event resolution, only 10% of the full MC dataset is used in the “Monte Carlo Dataset” rows. *Source: $H \rightarrow 4\ell$ mass measurement group*

Dataset	Method	m_H [GeV]	Statistical uncertainty [GeV] (No systematics)	Total uncertainty [GeV] (Including systematics)
Monte Carlo dataset	Average	125.001	0.183	0.199
	Per-event	125.000	0.180	0.196
Sampled MC toys	Average	125	0.179	0.196
	Per-event	125.005	0.177	0.193

10 Observed results

This section reports our observed ATLAS data. First, we review key data-simulation comparisons, notably for $m_{4\ell}^{\text{constrained}}$, BDT_{ZZ} , and the QRNN score distributions. Then, we provide the observed value of m_H . Finally, we determine the probability of the observed uncertainty and the compatibility of our results given different m_H fit parameter schemes.

Table 16 (table 17) shows the expected and observed yields for events with $m_{4\ell}^{\text{constrained}}$ on the fit interval [105, 160] GeV (on the reduced interval [115, 130] GeV). We observe good agreement in overall number of events.

Table 16: The expected and observed yields for events with $105 < m_{4\ell}^{\text{constrained}} < 160$ GeV. *Source: $H \rightarrow 4\ell$ mass measurement group*

Final state	signal	qqZZ	ggZZ	$Z + jets \ \& \ t\bar{t}$	tXX	Total	Observed
4μ	81 ± 5	117 ± 6	5.1 ± 3.3	7.4 ± 0.5	3.3 ± 0.4	214 ± 9	218
$2e2\mu$	55.7 ± 3.3	80 ± 4	3.5 ± 2.2	8.4 ± 0.6	2.40 ± 0.22	150 ± 6	170
$2\mu 2e$	42.0 ± 3.0	60 ± 4	2.8 ± 1.8	9.5 ± 1.5	2.20 ± 0.28	116 ± 6	112
$4e$	37.9 ± 2.8	51 ± 5	2.6 ± 1.6	7.2 ± 1.0	2.07 ± 0.24	101 ± 6	105
Total	216 ± 13	309 ± 18	14 ± 9	32.6 ± 2.8	10.0 ± 1.0	581 ± 24	605

Table 17: The expected and observed yields for with $115 < m_{4\ell}^{\text{constrained}} < 130$ GeV. *Source: $H \rightarrow 4\ell$ mass measurement group*

Final state	signal	qqZZ	ggZZ	$Z + jets \ \& \ t\bar{t}$	tXX	Total	Observed
4μ	78 ± 5	37.0 ± 2.0	1.2 ± 0.8	2.20 ± 0.16	0.60 ± 0.07	119 ± 5	114
$2e2\mu$	53.1 ± 3.1	25.7 ± 1.3	0.7 ± 0.5	2.50 ± 0.18	0.42 ± 0.04	82.5 ± 3.4	94
$2\mu 2e$	40.2 ± 2.9	16.8 ± 1.2	0.56 ± 0.35	3.1 ± 0.5	0.33 ± 0.04	61.1 ± 3.2	59
$4e$	35.3 ± 2.6	14.6 ± 1.5	0.54 ± 0.34	2.37 ± 0.33	0.37 ± 0.04	53.2 ± 3.1	47
Total	207 ± 13	94 ± 5	3.1 ± 2.0	10.2 ± 0.9	1.72 ± 0.17	316 ± 14	314

10.1 Comparing key distributions in data and simulation

Figures 65 and 66 show the $m_{4\ell}^{\text{constrained}}$ distribution before any fit in data and simulation inclusively and per final state, respectively.

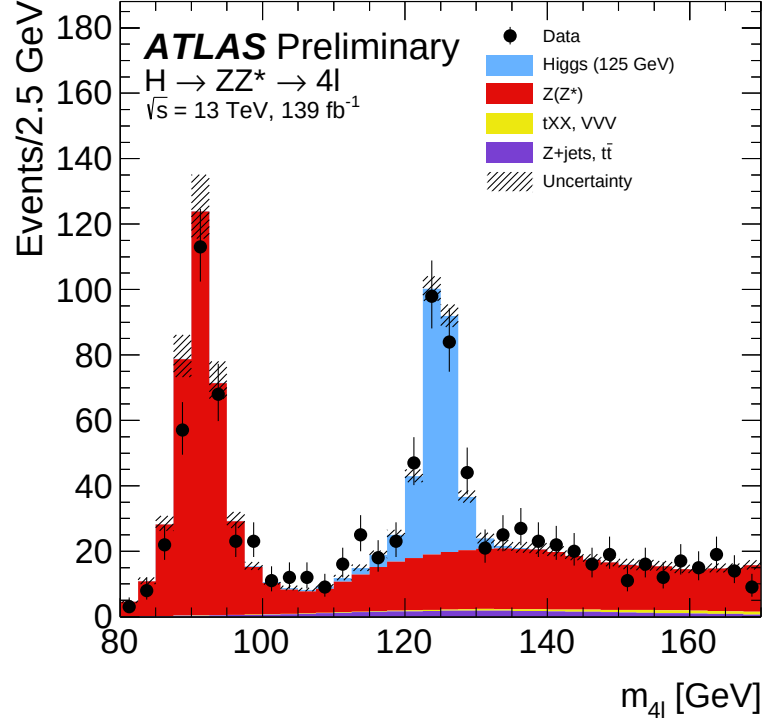


Figure 65: Four-lepton invariant mass distribution in simulation (colored bars) and data (black dots) collected with ATLAS during LHC Run 2. *Source: [ATLAS-CONF-2020-005](#)*

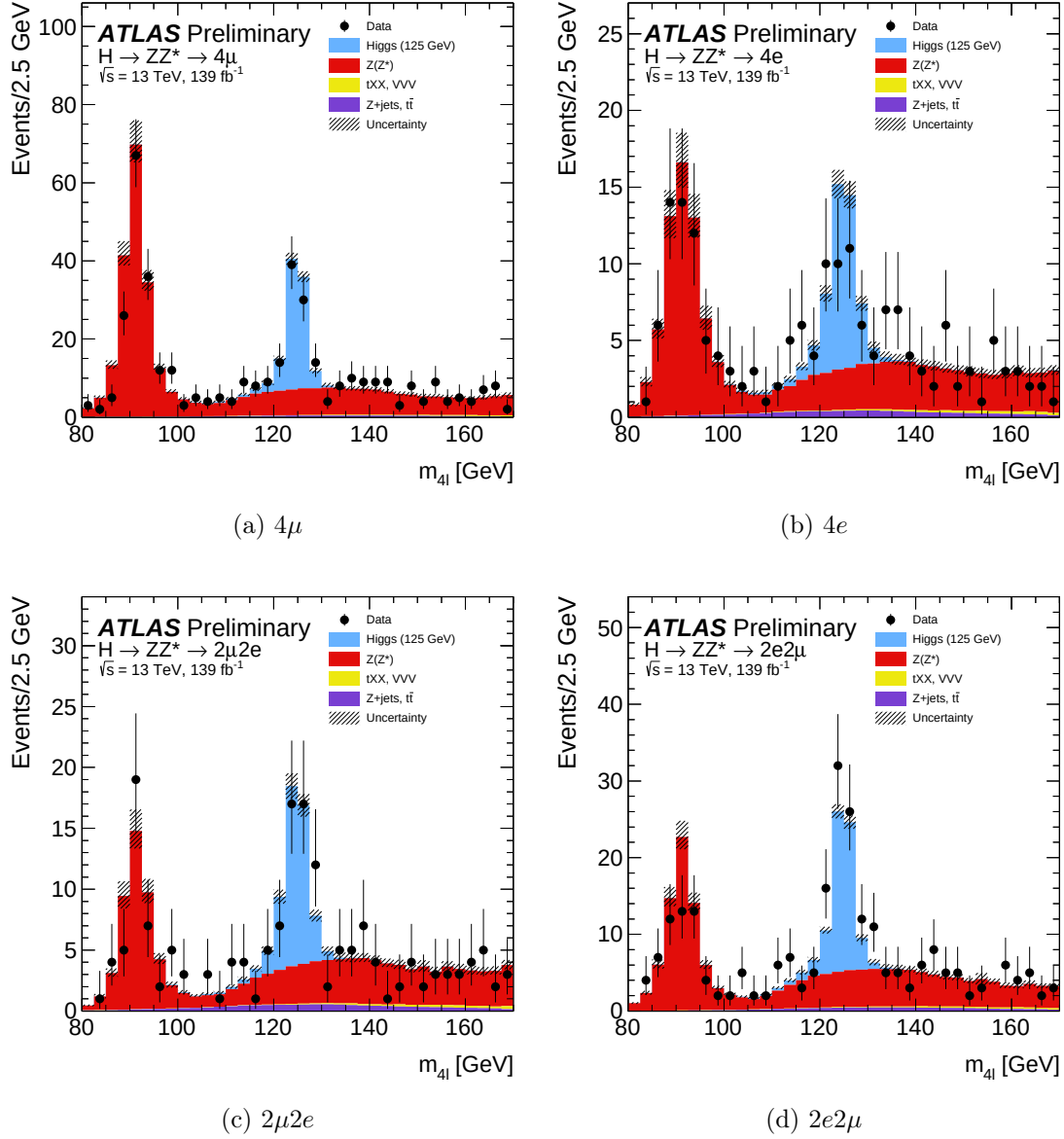


Figure 66: Four-lepton invariant mass distribution in each decay channel, inclusive of BDT_{ZZ} score, in simulation (colored bars) and data (black dots) collected with ATLAS during LHC Run 2.

Source: [ATLAS-CONF-2020-005](#)

Figures 67 and 68 show the BDT_{ZZ} distribution before any fit in data and simulation inclusively and per final state, respectively. The plots have double the number of BDT_{ZZ} bins used in the analysis to better elucidate the agreement. Notice the downward fluctuation at low BDT_{ZZ} score and the upward fluctuation at high BDT_{ZZ} score. The observed overall signal and $ZZ^{(*)}$ normalization parameters reflect this.

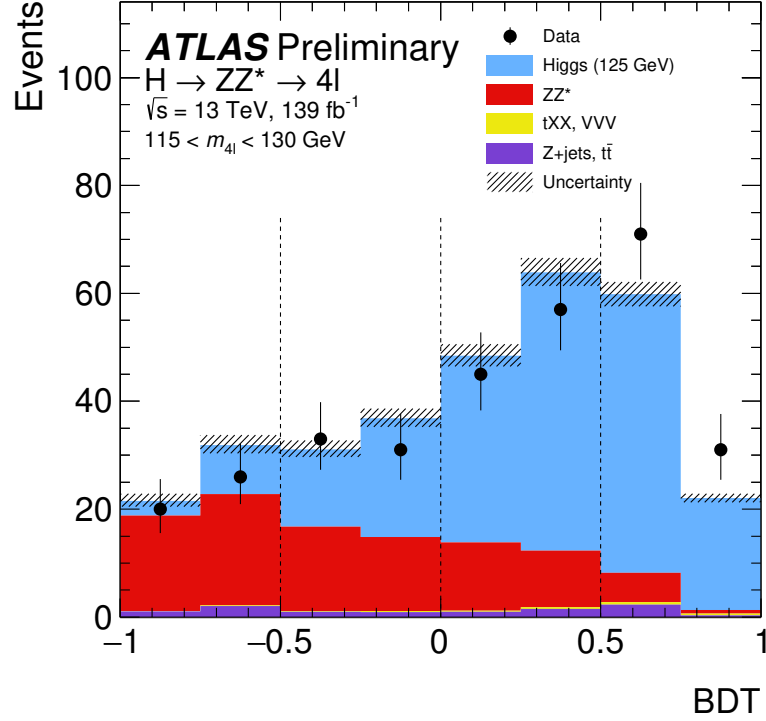
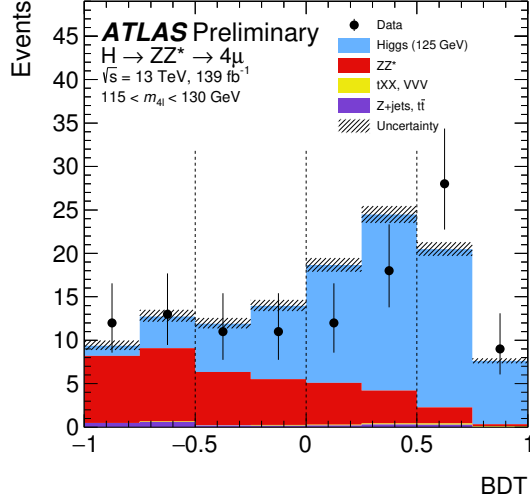
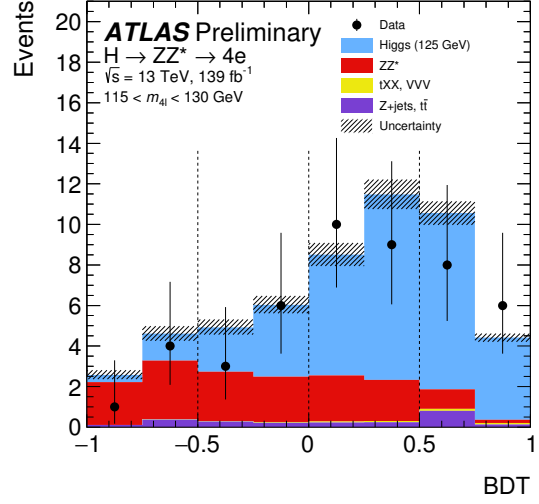


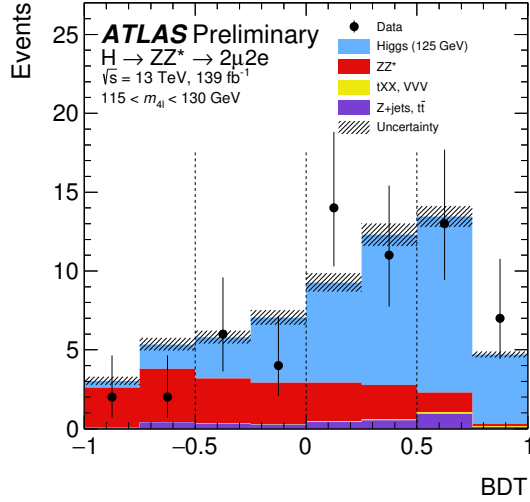
Figure 67: BDT_{ZZ} distribution in simulation (colored bars) and data (black dots) collected with ATLAS during LHC Run 2. *Source: [ATLAS-CONF-2020-005](#)*



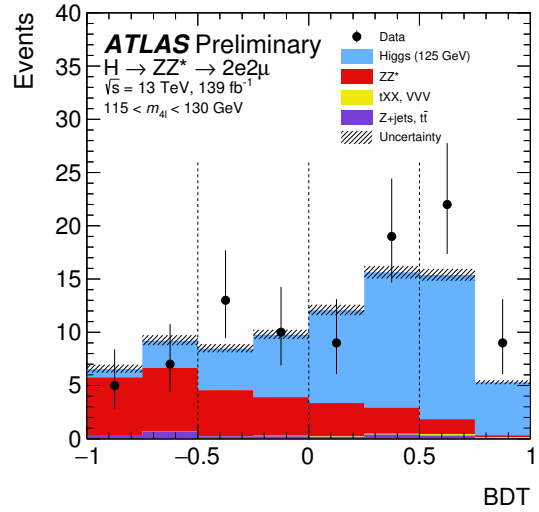
(a) 4μ



(b) $4e$



(c) $2\mu 2e$



(d) $2e 2\mu$

Figure 68: BDT_{ZZ} distributions in each decay channel in simulation (colored bars) and data (black dots) collected with ATLAS during LHC Run 2. Source: [ATLAS-CONF-2020-005](#)

Figures 69 and 70 show the QRNN score distribution in the signal region in data and simulation inclusively and per final state, respectively.

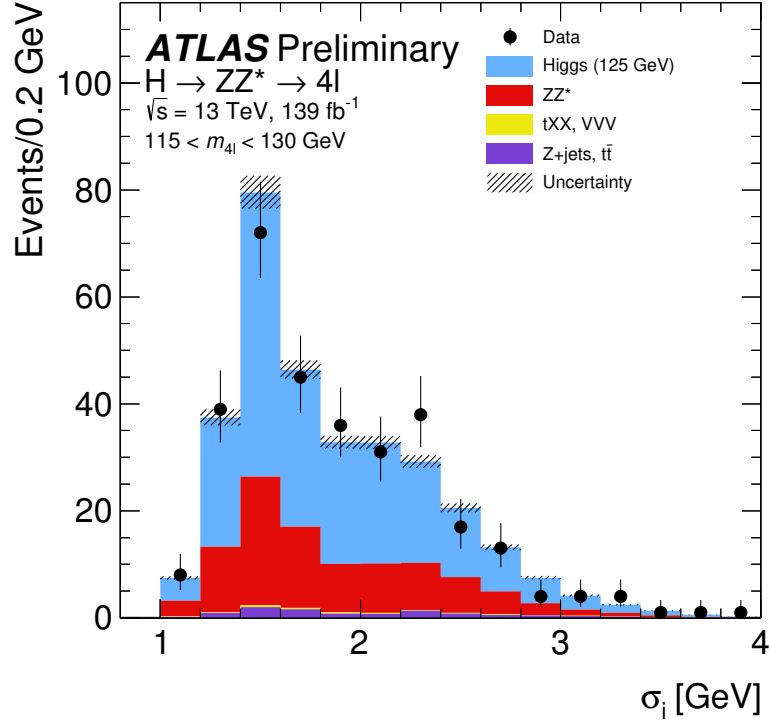
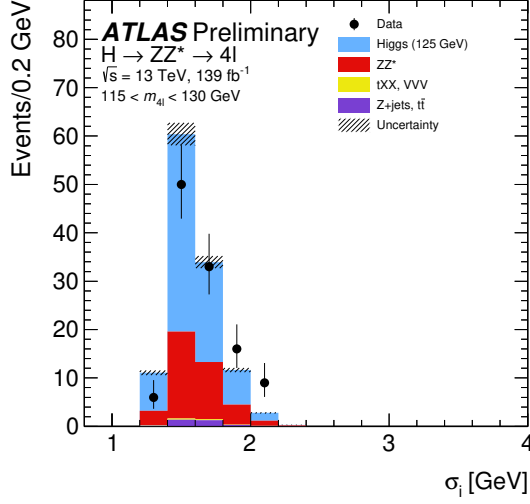
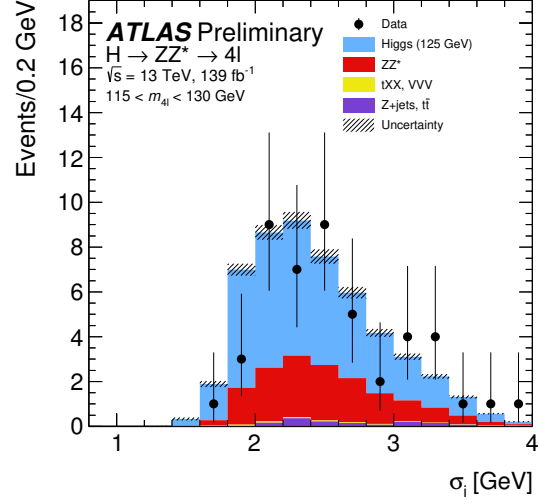


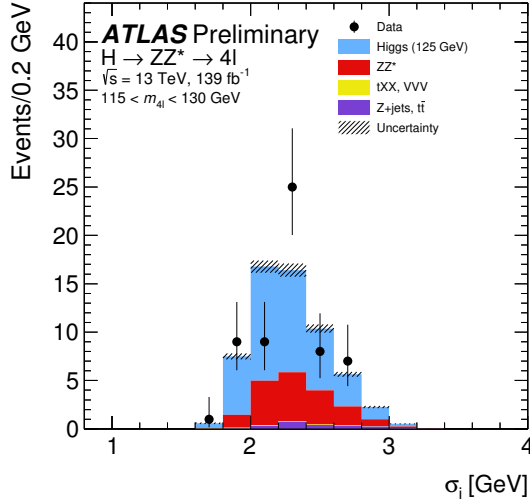
Figure 69: QRNN score, σ_i , in the resonance's core inclusive of final state in simulation (colored bars) and data (black dots) collected with ATLAS during LHC Run 2. *Source:* [ATLAS-CONF-2020-005](#)



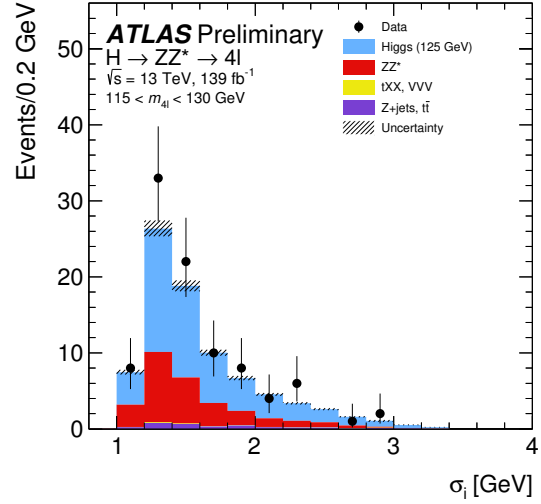
(a) 4μ



(b) $4e$



(c) $2\mu 2e$



(d) $2e2\mu$

Figure 70: The σ_i values predicted by the neural network for data (black points) in the 115–130 GeV $m_{4\ell}^{\text{constrained}}$ range are shown compared to the expectation in each final state. The expectation comes from MC simulation for signal, ZZ^* , tXX , and VVV ; for $Z + jets$ and $t\bar{t}$ the shape from MC simulation is normalized using the data-driven estimate. The uncertainty band includes all systematic uncertainties. *Source: [ATLAS-CONF-2020-005](#)*

Figures 71 and 72 alternatively present the QRNN score distributions in a 2D plane with $m_{4\ell}^{\text{constrained}}$. This representation highlights the collection of events with good resolution between 120 and 130 GeV.

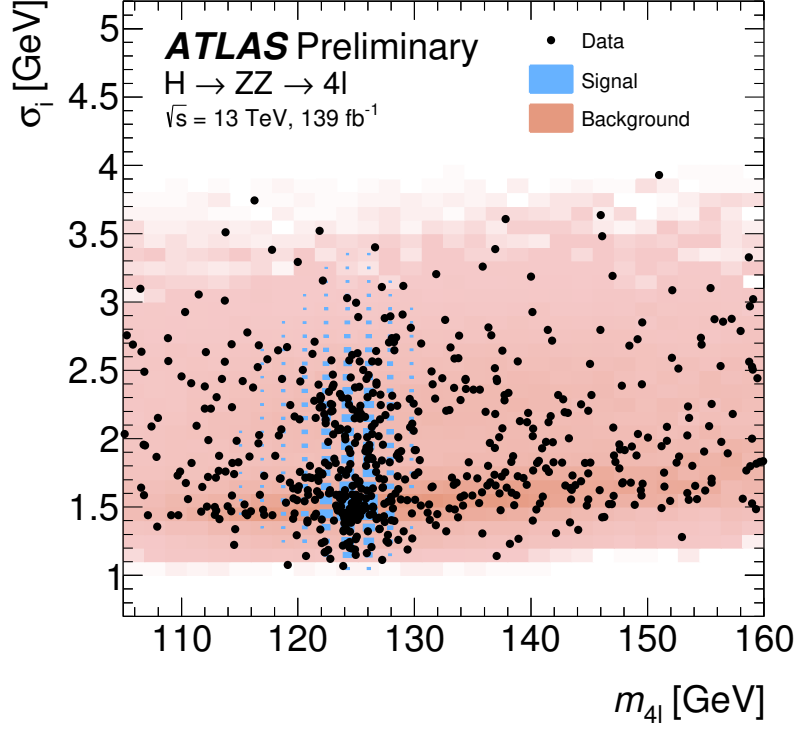


Figure 71: The $m_{4\ell}^{\text{constrained}}$ vs. σ_i in all final states combined. The black points are the data, while the signal (background) prediction is in blue (red), with larger boxes (darker colors) indicating a higher abundance. *Source: [ATLAS-CONF-2020-005](#)*

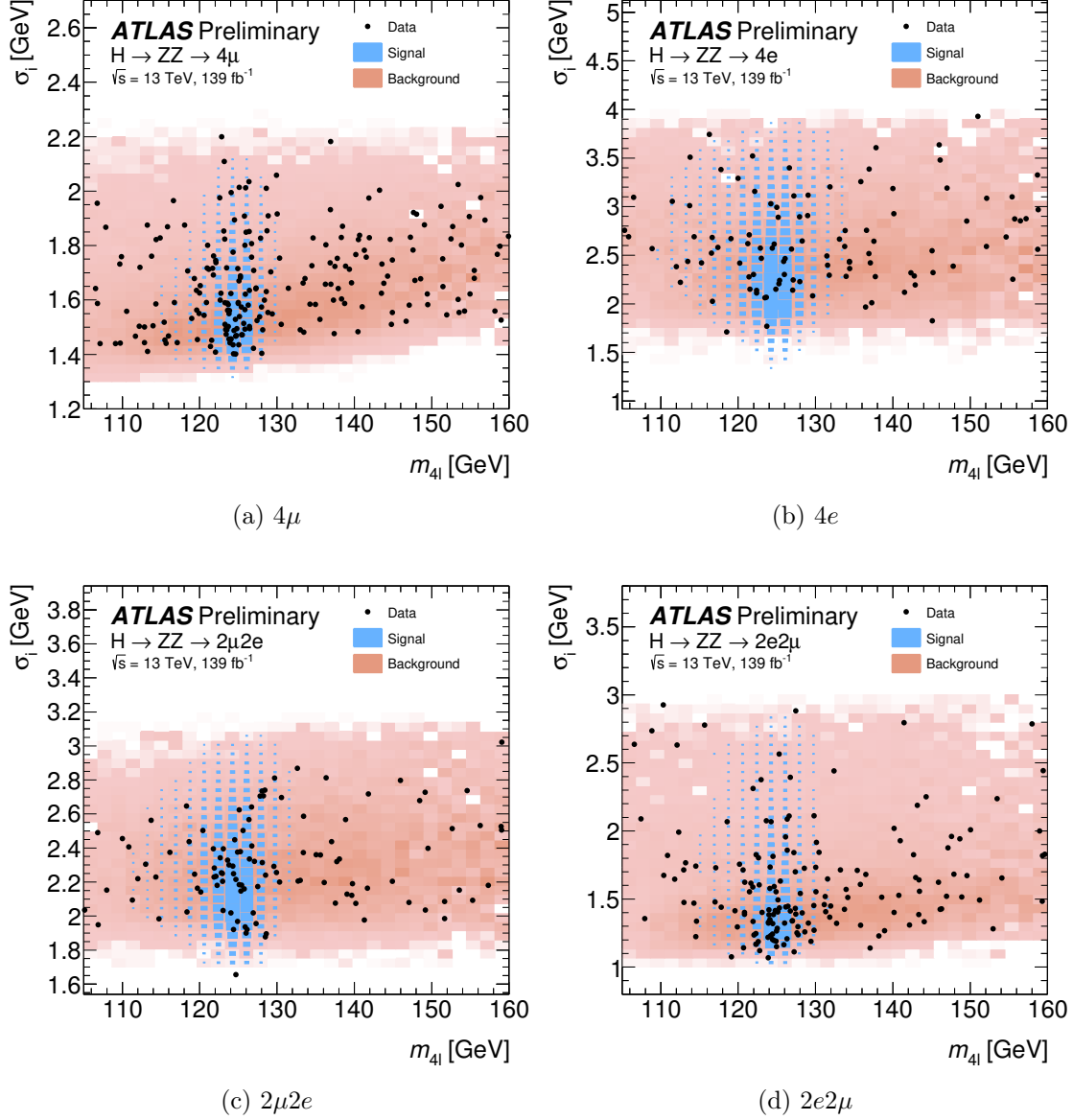


Figure 72: The $m_{4\ell}^{\text{constrained}}$ vs. σ_i distribution of events in each decay channel. The black points are the data, while the signal (background) prediction is in blue (red), with larger boxes (darker colors) indicating a higher abundance. Note that the vertical scale is different between the different final states and the inclusive plot, due to the resolution differences.
Source: [ATLAS-CONF-2020-005](#)

10.2 Observed m_H

Table 18: The m_H value from the fit to data using the 2D template, average resolution, and per-event resolution models.

Method	m_H [GeV]	Statistical uncertainty [GeV]	Total uncertainty [GeV] <i>Including systematics</i>
2D template method	124.927	0.197	—
Average resolution	124.971	0.197	+0.216 −0.211
Per-event resolution	124.922	0.187	+0.207 −0.197

This subsection reports the observed results for m_H . For each of the three methods, we extracted the estimate of m_H using a simultaneous profile likelihood fit in the 16 analysis categories. The Higgs mass, m_H , and the overall normalization parameters were allowed to vary to find their best-fit values, as were the nuisance parameters representing the systematic uncertainties. Table 18 provides the extracted m_H for all three methods.

10.2.1 2D template method

The 2D template method provided a crosscheck for this round of the analysis (described in full in appendix H). We performed a stat-only fit with the template method, observing a Higgs mass

$$m_H^{\text{template}} = 124.927 \pm 0.197 \text{ (stat.) GeV.}$$

Table 19 documents the observed overall signal and $ZZ^{(*)}$ normalizations from the template method fit. As we anticipated from the BDT_{ZZ} data-simulation distributions (figures 67 and 68), the normalizations compensate for the fluctuations in data.

10.2.2 Average resolution

The observed value of m_H using the average resolution method was

$$m_H^{\text{average resolution}} = 124.971 \pm 0.197 \text{ (stat.)}_{-0.074}^{+0.092} \text{ (syst.) GeV.}$$

Table 19: Fit results using the 2D Template method on the observed Run 2 dataset (no systematic uncertainties considered) *Source: HH (2020)*

Observed results	Template 2D
m_H^{measured}	$124.927 \pm 0.197 \text{ GeV}$
$\mu_{\text{signal}}^{BDT1}$	0.34 ± 0.49
$\mu_{\text{signal}}^{BDT2}$	0.82 ± 0.22
$\mu_{\text{signal}}^{BDT3}$	0.87 ± 0.12
$\mu_{\text{signal}}^{BDT4}$	1.36 ± 0.15
μ_{ZZ*}^{BDT1}	1.17 ± 0.11
μ_{ZZ*}^{BDT2}	1.14 ± 0.14
μ_{ZZ*}^{BDT3}	1.23 ± 0.14
μ_{ZZ*}^{BDT4}	1.25 ± 0.22

The one-sided p-value for observing this total uncertainty on m_H or greater with this method is 14.6% (statistical uncertainty only: 14.9%), found by integrating figure 58 from the observed uncertainty for the average resolution method to infinity. Figure 73a shows the negative-log-likelihood (NLL) scan for the combined fit. Alternatively, we can perform the fit independently in each decay channel while fitting all BDT_{ZZ} categories simultaneously. Figure 73b displays the fit results for all channels in addition to the combined. The results obtained when measuring the Higgs mass separately in each decay channel (or BDT_{ZZ} bin) are compatible with the nominal result obtained when we assume a single, common Higgs mass parameter and combine information across all 16 analysis categories. Table 20 report this compatibility. The probability of obtaining results in agreement with distinct or shared m_H parameters is at least 90%. The observed overall signal and $ZZ^{(*)}$ normalizations are given in table 21.

Table 20: Compatibility of sharing one m_H parameter across all analysis categories with having dedicated m_H parameters for each decay channel or each BDT_{ZZ} bin using the average resolution model. *Source: $H \rightarrow 4\ell$ mass measurement group*

m_H Scheme	Number m_H Parameters	Compatibility
Per Decay Channel	4	93.6%
Per BDT bin	4	90.1%

Table 21: Fit results using the average resolution method on the observed Run 2 dataset (including systematics) *Source: $H \rightarrow 4\ell$ mass measurement group*

Observed results	Average resolution
m_H^{measured}	$124.971 \pm 0.197 \text{ (stat.) } {}^{+0.092}_{-0.074} \text{ (syst.) GeV}$
$\mu_{\text{signal}}^{BDT1}$	$0.33 {}^{+0.50}_{-0.47}$
$\mu_{\text{signal}}^{BDT2}$	$0.77 {}^{+0.23}_{-0.21}$
$\mu_{\text{signal}}^{BDT3}$	$0.84 {}^{+0.12}_{-0.12}$
$\mu_{\text{signal}}^{BDT4}$	$1.30 {}^{+0.15}_{-0.14}$
$\mu_{\text{signal}}^{BDT1}$	$1.09 {}^{+0.12}_{-0.11}$
$\mu_{\text{signal}}^{BDT2}$	$1.07 {}^{+0.14}_{-0.13}$
$\mu_{\text{signal}}^{BDT3}$	$1.15 {}^{+0.15}_{-0.14}$
$\mu_{\text{signal}}^{BDT4}$	$1.09 {}^{+0.23}_{-0.21}$

Figure 74 provides the ranking plot and correlation matrix for the fit to data. The muon momentum resolution systematics, MUON_MS and MUON_ID, appear in the ranking on data, while they did not appear in the expected results ranking. Muon and electron scale systematic uncertainties remain the most significant. The correlation matrix agrees with our expectations.

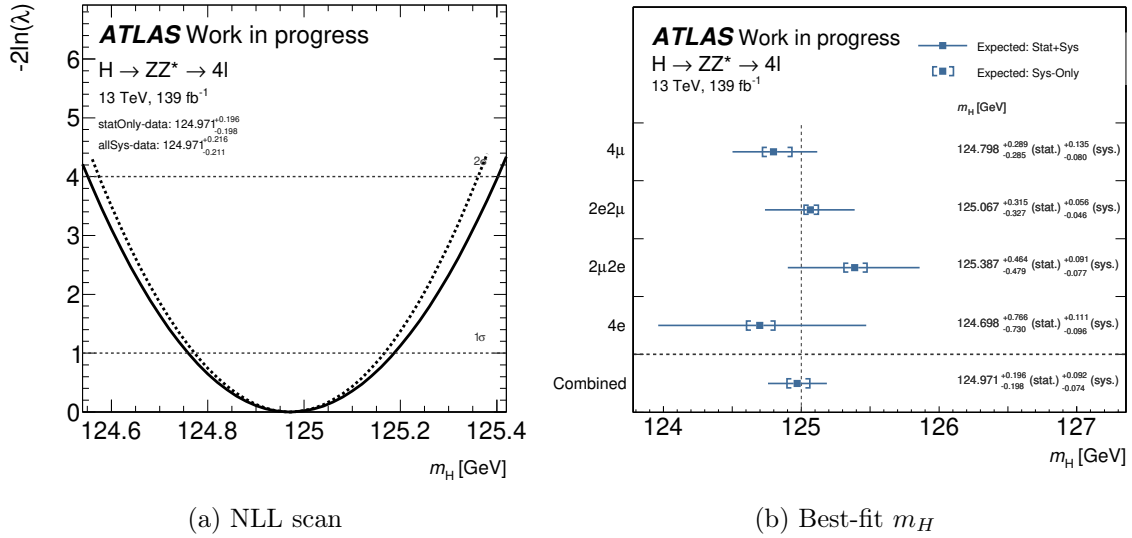


Figure 73: (a) NLL scan with and without systematics for fitting the average resolution model to observed data. (b) The best-fit values of m_H in the individual decay channels in data using the average resolution method. *Source: $H \rightarrow 4\ell$ mass measurement group*

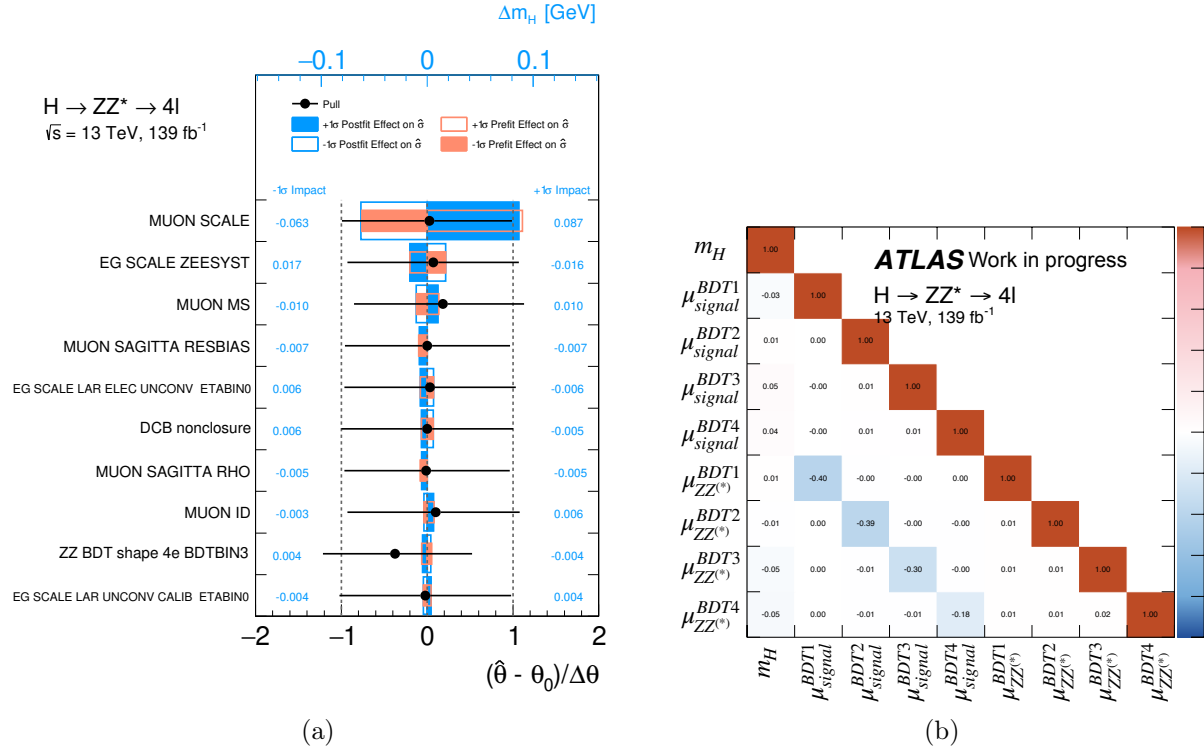


Figure 74: (a) Impact of systematic uncertainties on m_H in data using the average resolution method. (b) Correlation matrix for parameters of interest extracted with the average resolution method.

Source: $H \rightarrow 4\ell$ mass measurement group

10.2.3 Per-event resolution

The observed value of m_H using the per-event resolution method was

$$m_H^{\text{per-event resolution}} = 124.922 \pm 0.187 \text{ (stat.) } {}^{+0.090}_{-0.062} \text{ (syst.) GeV},$$

using the combined fit on all 16 analysis categories. Therefore, the per-event resolution method improves the precision on m_H by 10 MeV. The one-sided p-value for observing this total uncertainty on m_H or greater with this method is 27.3% (statistical uncertainty: 28.9%), found by integrating figure 58 from the observed uncertainty for the per-event resolution method to infinity. Figure 75 shows the fit to the $H \rightarrow 4\ell$ candidates in all categories combined. Figure 76 projects the fitted $m_{4\ell}^{\text{constrained}}$ distribution into each decay channel.

Figure 77a shows the negative-log-likelihood (NLL) scan for the combined fit in black. Alternatively, we can perform the fit independently in each decay channel while fitting all

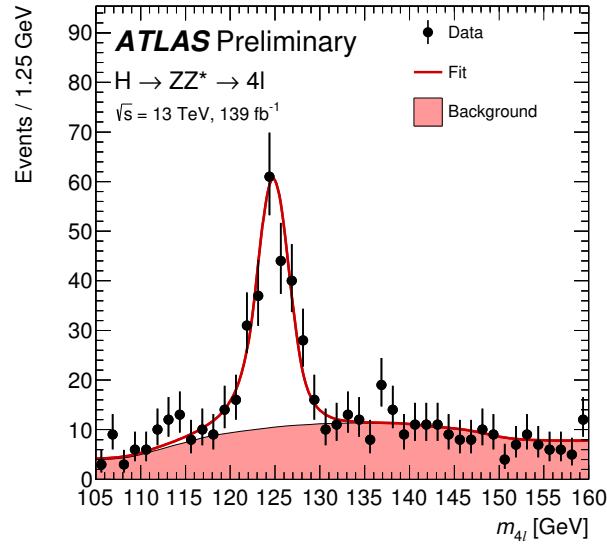


Figure 75: The $m_{4\ell}^{\text{constrained}}$ data distribution from all categories combined (black points) is shown along with the fit result when accounting for the per-event resolution (red line). The background component of the fit result is shown separately in solid red.

Source:

[ATLAS-CONF-2020-005](#)

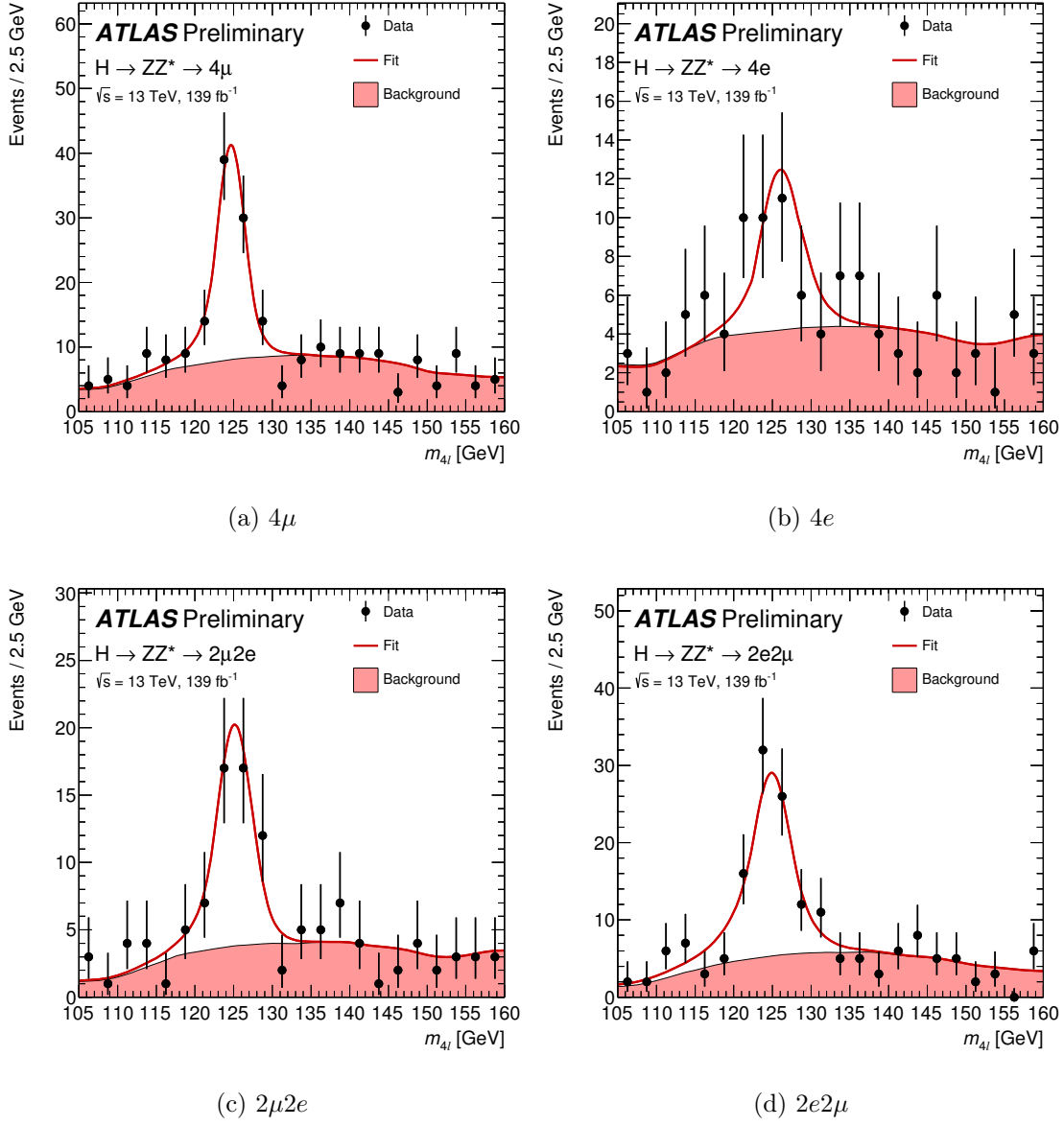


Figure 76: The $m_{4\ell}^{\text{constrained}}$ data distribution (black points) in each final state, integrated over BDT_{ZZ} bin, is shown along with the fit result when accounting for the per-event resolution (red line). The background component of the fit result is shown separately in solid red. *Source: [ATLAS-CONF-2020-005](#)*

BDT_{ZZ} categories simultaneously, displayed on the plot with colored lines. Figure 77b summarizes the fit results for all channels in addition to the combined. The results obtained when measuring the Higgs mass separately in each decay channel (or BDT_{ZZ} bin) are compatible with the nominal result obtained when we assume a single, common Higgs mass parameter

and combine information across all 16 analysis categories. Table 22 report this compatibility. The probability of obtaining results in agreement with distinct or shared m_H parameters is at least 90%. Table 23 provides the observed overall signal and $ZZ^{(*)}$ normalizations when fitting with the per-event resolution model.

Table 24 summarizes the leading systematic uncertainties in the per-event resolution measurement. The muon momentum resolution systematic uncertainty quoted at 10 MeV comprises both the ID and MS momentum resolution uncertainties. The leading systematic uncertainties are the muon momentum scale and electron energy resolution.

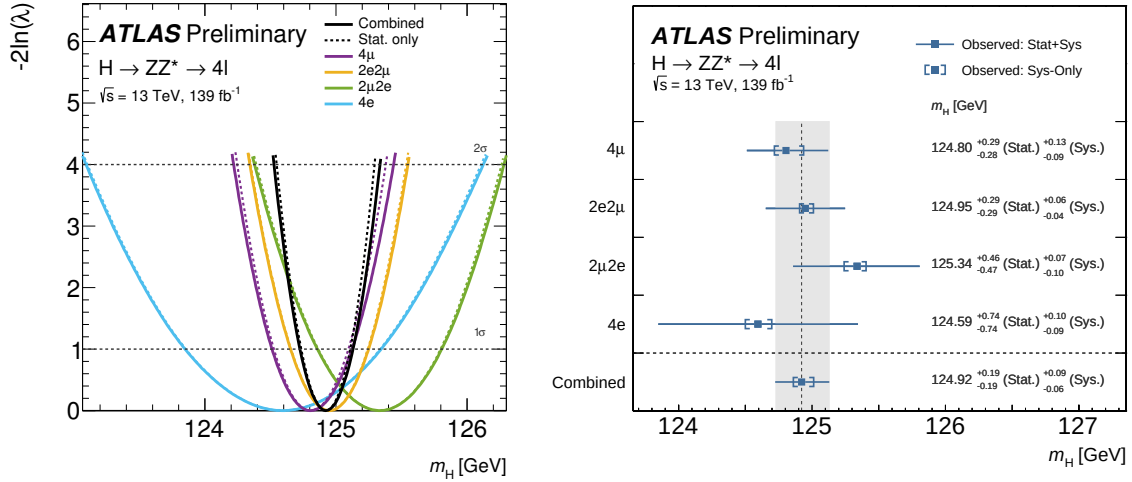


Figure 77: On the left, likelihood scans are shown for the fit in each of the final states 4μ (purple), $2e2\mu$ (orange), $2\mu2e$ (green), and $4e$ (blue), and combined (black), both with (solid lines) and without (dashed lines) taking systematic uncertainties into account. The horizontal dashed lines indicate the location of the one and two σ uncertainties. On the right, the fit results obtained in each final state are shown together with the combined result. The uncertainty on each point is the total statistical and systematic uncertainty; the brackets show the systematic uncertainty only. The vertical dashed line indicates the combined result, and the grey band its total uncertainty.

Source: [ATLAS-CONF-2020-005](#)

Table 22: Compatibility of sharing one m_H parameter across all analysis categories with having dedicated m_H parameters for each decay channel or each BDT_{ZZ} bin using the per-event resolution model.

Source: $H \rightarrow 4\ell$ mass measurement

group

m_H Scheme	Number m_H Parameters	Compatibility
Per Decay Channel	4	95.4%
Per BDT bin	4	90.1%

Table 23: Fit results using the per-event method on the observed Run 2 dataset (including systematics)

Source: HH (2020)

Observed results	Per-event resolution	
m_H^{measured}	124.922	$^{+0.186}_{-0.187}$ (stat.) $^{+0.092}_{-0.074}$ (syst.) GeV
μ_{signal}^{BDT1}	0.35	$^{+0.51}_{-0.47}$
μ_{signal}^{BDT2}	0.82	$^{+0.23}_{-0.21}$
μ_{signal}^{BDT3}	0.86	$^{+0.12}_{-0.12}$
μ_{signal}^{BDT4}	1.30	$^{+0.15}_{-0.14}$
μ_{signal}^{BDT1}	1.08	$^{+0.12}_{-0.11}$
μ_{ZZ*}^{BDT2}	1.06	$^{+0.14}_{-0.13}$
μ_{ZZ*}^{BDT3}	1.14	$^{+0.14}_{-0.13}$
μ_{ZZ*}^{BDT4}	1.09	$^{+0.23}_{-0.21}$

Table 24: Impact of systematic uncertainties for the per-event resolution method

Impact of the leading systematic uncertainties on m_H , using the per-event resolution

	Systematic Uncertainty	Impact (GeV)
method.	Muon momentum scale	+0.08, −0.06
	Electron energy scale	±0.02
	Muon momentum resolution	±0.01
	Muon sagitta bias correction	±0.01

Source: [ATLAS-CONF-2020-005](#)

Figure 78 provides the ranking plot and correlation matrix for the fit to data. The muon momentum resolution systematics, MUON_MS and MUON_ID, appear in the ranking on data, while they did not appear in the expected results ranking. Muon and electron scale systematic uncertainties remain the most significant. The correlation matrix agrees with our expectations.

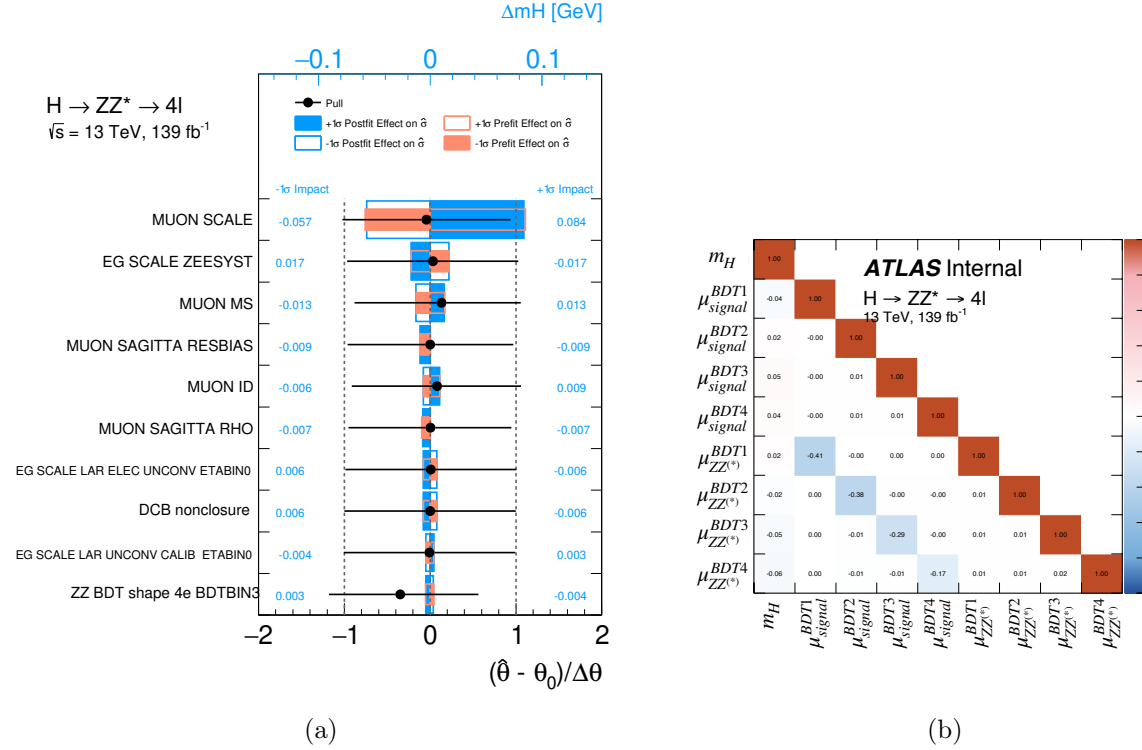


Figure 78: (a) Impact of systematic uncertainties on m_H in data using the per-event resolution method. (b) Correlation matrix for parameters of interest extracted with the per-event resolution method.

Source: $H \rightarrow 4\ell$ mass measurement group

10.3 Observed impact of accounting for per-event resolution

To quantify the impact of accounting for the per-event resolution, we had fit identical ensembles of sampled MC toys, assuming $m_H = 125$ GeV, with each method. We can integrate our comparison distributions of those pseudoexperiments from the observed quantity to infinity to establish the **compatibility** of our observations with our expectations and between methods.

First, we checked the toy-by-toy difference in fitted m_H (figure 79a). We observed no bias in the distribution of differences in fitted m_H (mean = 0 GeV). We do observe a difference in the m_H estimate in data of 50 MeV, if we neglect the per-event resolution and instead use the average resolution. Integrating from 50 MeV to infinity on figure 79a shows that there is a 26% chance of observing this change or larger in m_H central value when replacing the per-event resolution with the average. Then, we checked the change in m_H uncertainty on a toy-by-toy basis (figure 79b). The mean difference shows that accounting for the per-event resolution improves our expected precision on m_H by about 4 MeV, or 2%, when evaluated on identical events. We observed a difference in data in m_H uncertainty of 10 MeV when replacing the per-event resolution with the average. Integrating from 10 MeV to infinity on figure 79b calculates the one-sided p-value of the compatibility of this observed difference of 10 MeV with our expected difference. It is 0.17.

Figure 80 reminds us of the distributions of the expected total uncertainty predicted with fits to sampled MC and adds dotted lines indicating the observed uncertainty. Table 25 summarizes the one-sided p-values for observing an m_H uncertainty at least as high as the one observed for each method.

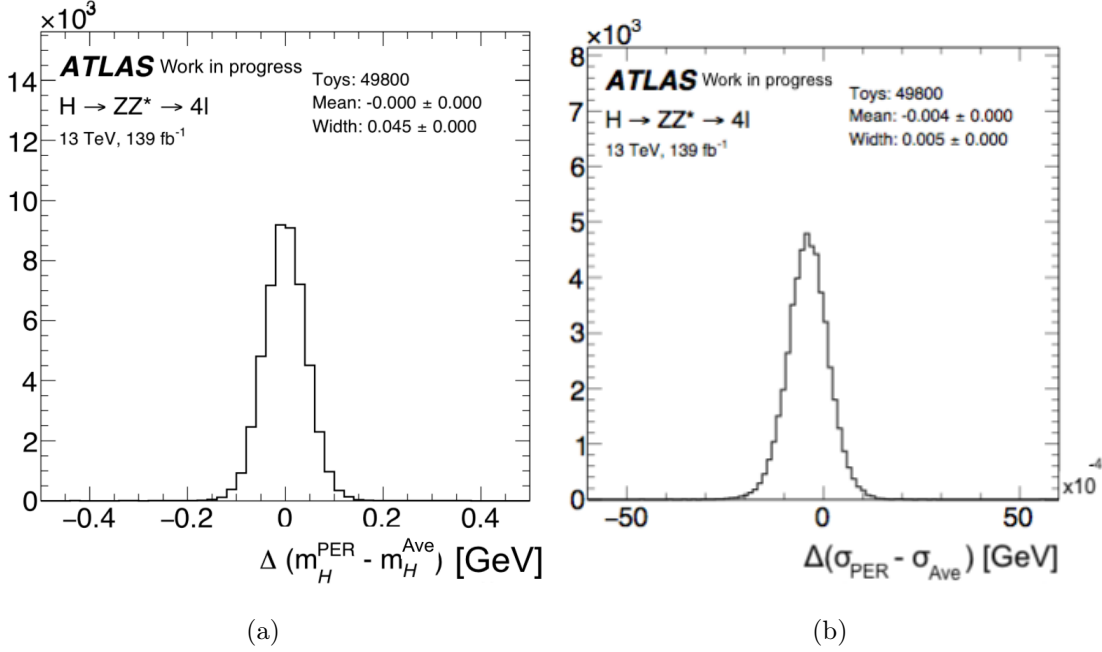


Figure 79: Toy-by-toy comparisons of the (a) m_H central value and (b) uncertainty results from fitting identical ensembles with both methods. Here, “analytic” indicates the average resolution approach.
Source: $H \rightarrow 4\ell$ mass measurement group

Table 25: One-sided p-value for observed uncertainty

The one sided p-value for observing an uncertainty on m_H as high or higher than the one observed for each method. We integrate from the total observed uncertainty (observed statistical uncertainty) to infinity on figure 80 (figure 80) to extract these p-values.

Method	p-value	p-value
	<i>Statistical uncertainty</i>	<i>Total uncertainty</i>
Average resolution	14.9%	14.6%
Per-event resolution	27.3%	28.9%

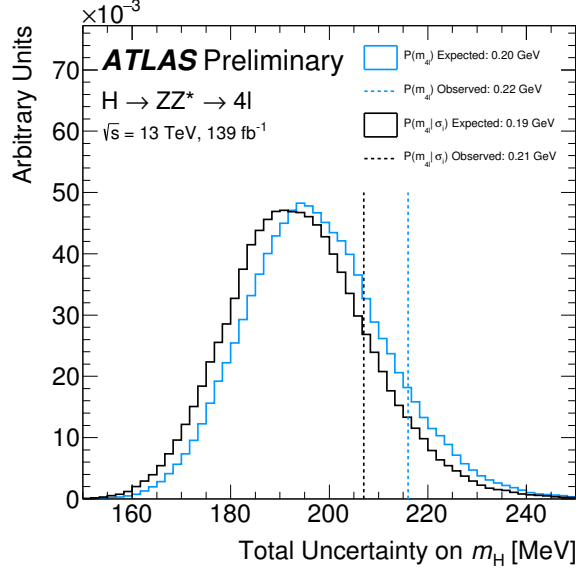


Figure 80: The distributions of the total m_H uncertainty from pseudo-experiments assuming $m_H = 125$ GeV are shown, for both when the fit does (black) and does not (blue) account for the per-event resolution. The solid lines correspond to the expected uncertainty distribution from the pseudo-experiments while the vertical dashed lines indicate the observed values of the uncertainties. *Source: [ATLAS-CONF-2020-005](#)*

10.4 Summary

Combining the statistical and systematics in quadrature, the measured value of the Higgs boson mass in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channel using the full LHC Run 2 dataset collected with ATLAS interpreted with the per-event resolution method is

$$m_H = 124.92_{-0.20}^{+0.21} \text{ GeV},$$

in good agreement with previous ATLAS and CMS measurements. The measurement is statistically dominated and its total uncertainty is in agreement with the expected uncertainty. The leading systematic uncertainties are the muon momentum scale and electron energy scale.

Chapter V

Outlook

11 ATLAS’ future: Phase II Inner Tracker Upgrade and the High-Luminosity LHC (HL-LHC)

The Large Hadron Collider (LHC) will be incrementally upgraded over the next decade to the High Luminosity LHC (HL-LHC), delivering a five-fold increase in instantaneous luminosity in the late 2020s. HL-LHC is designed to achieve that luminosity in part by dramatically increasing the number of interactions per bunch crossing. Figure 81 displays simulations of the LHC and HL-LHC pileup conditions, illustrating the substantial tracking and vertexing challenges the HL-LHC poses. Some of the current ATLAS subdetectors cannot accommodate the anticipated harsher radiation environment and increased detector occupancies. The current Inner Detector is particularly vulnerable given its close proximity to the beamline. By the end of Run 3, the pixel and SCT subdetectors are expected to have suffered severe radiation damage to their sensors and electronics. Predicted detector occupancies for the instantaneous luminosity at the HL-LHC also exceed the design specifications of the transition radiation tracker. Consequently, the ATLAS community is developing an all-silicon tracker, known as the Inner Tracker (ITk), to replace the entire Inner Detector (figure 82).

The ITk must provide higher granularity to cope with increased detector occupancies and feature an improved material budget. Its sensors and readout components must also be more radiation-resistant [64]. The new ITk will employ both hybrid silicon pixel modules (pixels) at smaller radii and silicon micro-strip modules (strips) at larger radii. It will consist of five cylindrical pixel layers and four cylindrical strip layers in the barrel region. Figure 83 summarizes the ITk Strip detector anatomy in the barrel. Each cylindrical strip layer consists of “staves” mounted on carbon fiber cylinders. Each stave consists of silicon modules mounted



Figure 81: HL-LHC is designed to produce 10 times the number of interactions per bunch crossing (pileup) compared to the design of the LHC. (a) ATLAS event display from the first interactions during Run 2, demonstrating LHC’s design pileup, $\langle\mu\rangle \sim 20 - 30$. (b) ATLAS event display for simulated $t\bar{t}$ events at the HL-LHC design average pileup $\langle\mu\rangle = 200$. *Source: ATLAS event at 13 TeV - First stable beam, 3 June 2015 - run: 266904, ATLAS Upgrade Event Displays*

on a thermo-mechanical carbon composite support structure called a stave core. The core provides cooling and mechanical support. A silicon sensor module consists of a silicon strip sensor mounted with hybrid readout electronics (see left image, figure 83). A stave has 14 modules mounted on each side of the stave core (28 modules total) and has approximate dimensions of 140 cm x 11 cm. The entire ITk detector requires nearly 400 staves [64]. All staves constructed in the USA, half of the global production, will be assembled and tested at Brookhaven National Lab (BNL) prior to their shipment to CERN for mechanical integration. Figure 84 shows the first stave loaded in the US using the procedure developed by the Brandeis team stationed at BNL, a prototype with thermomechanical properties comparable to production staves.

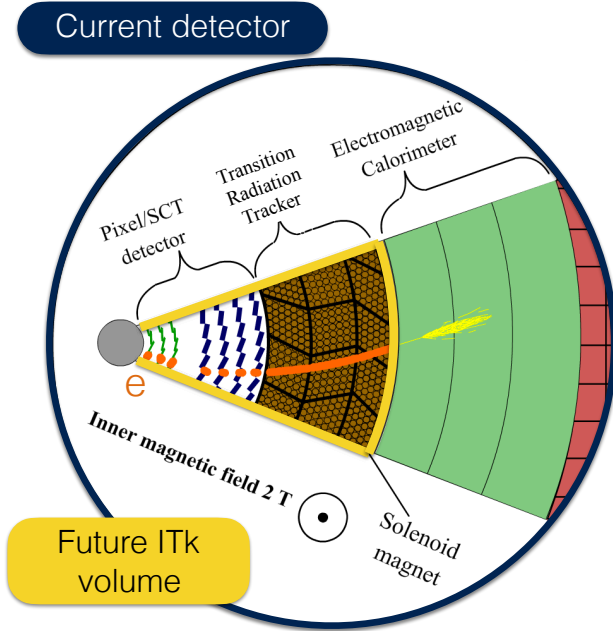


Figure 82: The Pixel, SCT, and TRT compose the current Inner Detector volume. The ITk system, outlined in yellow, will occupy this entire space with its active area and services. *Source: adapted from T Herrmann, O Jařábek, K Jende, M Kobel*

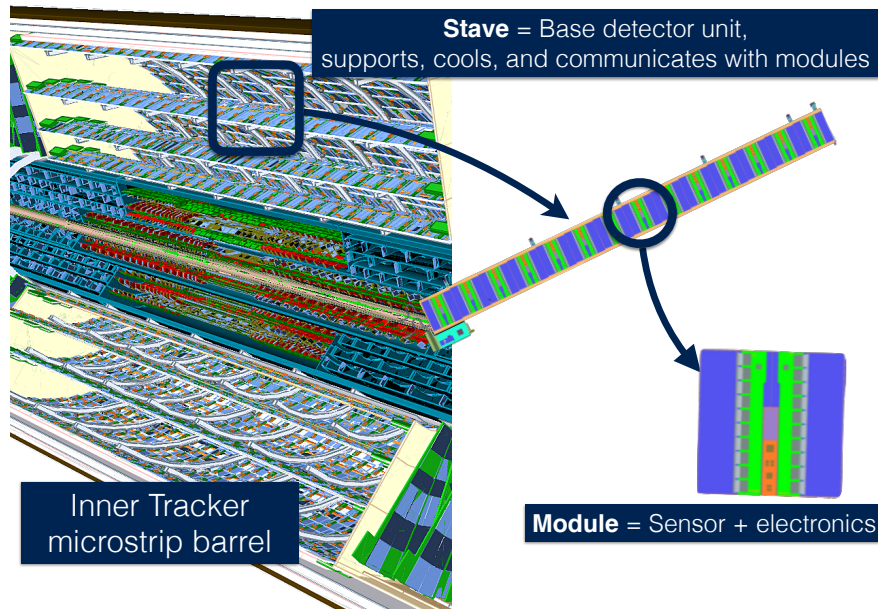


Figure 83: Anatomy of the ITk Strips barrel. The endcap is similarly modular, composed of petals mounted with wedge-shaped modules. The petals are arranged into disks. As the base detector unit, the stave supports, cools, and provides timing, trigger, and control (TTC) communication to and from modules. *Source: adapted from the ATLAS ITk Strips TDR*



Figure 84: The first stave loaded in the USA, the thermomechanical prototype, loaded at BNL in 2017. *Source: HH (2017)*

12 The Higgs mass at the HL-LHC

Figure 85 revisits figure 36 from section 7, projecting the 2D-template method from reference [18] to 3000 fb^{-1} , the total integrated luminosity we expect to collect with ATLAS and the HL-LHC. The measurement finally becomes systematically-limited at about 2000 fb^{-1} , assuming we do not reduce our systematic uncertainties further in the meantime. By the end of the HL-LHC run, we could expect a statistical uncertainty of 48 MeV with the 2D template method. If we did not reduce our systematic uncertainty from its mid-Run 2 level of 50 MeV, then we expect our total precision on m_H to fall to just 70 MeV.

This projection assumes that we had our current ATLAS detector during the HL-LHC run. Dedicated performance studies using simulations of the ITk and ATLAS in HL-LHC conditions enable us to produce plots like figure 86, forecasting the performance of our reconstructed particles. Regarding single muons, this figure tells us that we could expect a resolution improvement of as much as 30% using the ITk at $p_T^\mu = 45 \text{ GeV}$ [65]. The Strips system will be particularly important to the Higgs mass measurement, since most of the hits sampling the muon track's curvature (from which we measure transverse momentum) will fall in the strip layers.

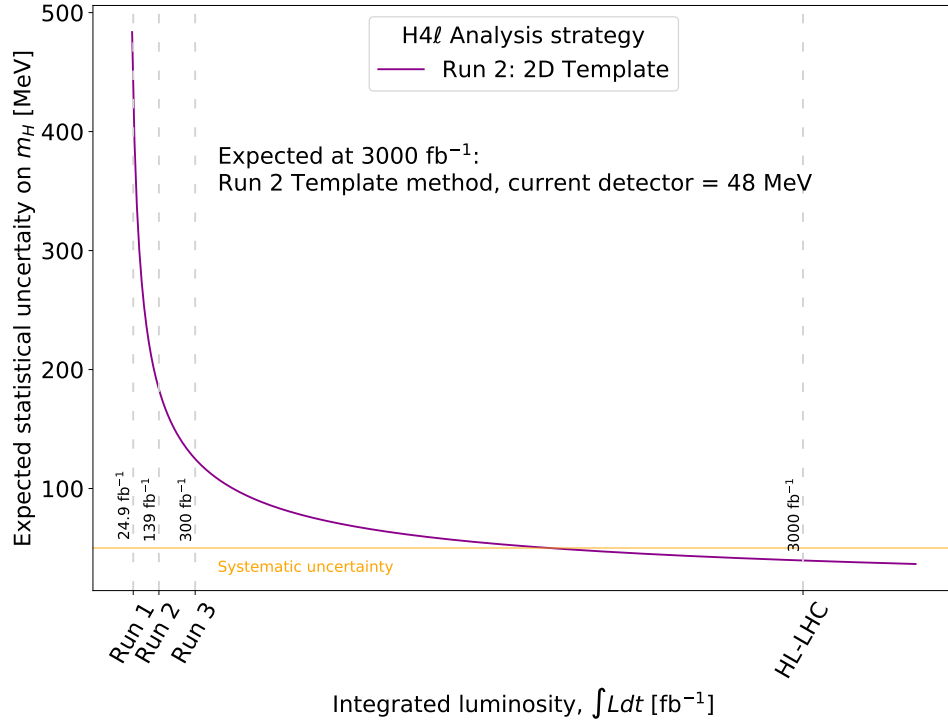


Figure 85: Projecting the expected statistical uncertainty of the $H \rightarrow 4\ell$ mass measurement to HL-LHC luminosities compared with current systematic uncertainties, assuming the Run 2 ATLAS detector and no changes in reconstruction or analysis approach with respect to reference [18].
Source: HH (2020)

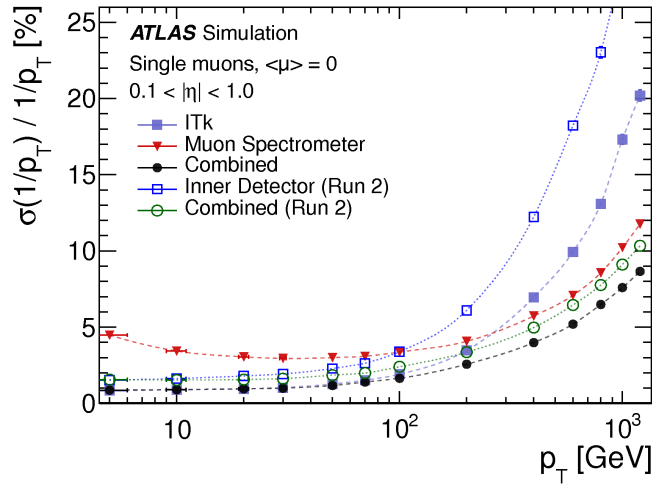


Figure 86: Combined muon momentum resolution and the individual contributions from the ITk and the upgraded Muon Spectrometer
Source: ATLAS ITk Pixel technical design report

Reference [66] projected the expected statistical uncertainty on m_H when we have accumulated 3000 fb^{-1} with HL-LHC using the per-event response approach found in reference [18]. The note took into account different scenarios for improvements to the muon momentum calibration. Table 26 provides the m_H precision given four scenarios. The first scenario mimics figure 85 and anticipates that the contributions of electrons to the mass measurement essentially vanish, reducing the systematic uncertainty to the muon momentum scale uncertainty quoted in the mid-Run 2 measurement [18]. The other scenarios anticipate the ITk’s improved tracking resolution relative to the Inner Detector and reducing the systematic uncertainties. Ultimately, we could reach a precision on m_H of 33 MeV, almost 5.5 times smaller than what we can currently achieve with the ATLAS experiment alone. This accomplishment would in turn allow us to better explore the Higgs sector and the SM.

Table 26: Expected uncertainty on the measured mass of the Higgs boson for the current and upgraded detector scenarios with 3000 fb^{-1} of HL-LHC data. *Source:* [ATL-PHYS-PUB-2018-054](#)

Improvement condition	Δ_{tot} (MeV)	Δ_{stat} (MeV)	Δ_{sys} (MeV)
Current detector (no improvement)	52	39	35
μ momentum resolution by 30% or similar	47	30	37
μ momentum resolution/scale by 30%/50%	38	30	24
μ momentum resolution/scale by 30%/80%	33	30	14

References

- [1] M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev. D* **98** (3 2018) 030001, URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001> (cit. on pp. 3, 4, 12).
- [2] Y. Grossman, *Standard Model. CERN Summer Student Lecture Series*, (2014), Provided url is for the first recording of the five-part lecture series. All are available using the “Similar records” feature on CDS. Lecture 3 was particularly influential in my thinking during my first summer at CERN, URL: <https://cds.cern.ch/record/1728275> (cit. on pp. 6, 145).
- [3] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191, URL: <https://link.aps.org/doi/10.1103/PhysRev.96.191> (cit. on p. 8).
- [4] Y. Nambu and G. Jona-Lasinio, *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I*, *Phys. Rev.* **122** (1961) 345 (cit. on p. 8).
- [5] Y. Nambu and G. Jona-Lasinio, *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. II*, *Phys. Rev.* **124** (1961) 246 (cit. on p. 8).
- [6] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, *Phys. Rev. Lett.* **13** (1964) 321 (cit. on p. 8).
- [7] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132 (cit. on p. 8).
- [8] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508 (cit. on p. 8).
- [9] G. Guralnik, C. Hagen, and T. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (1964) 585 (cit. on p. 8).
- [10] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, *Phys. Rev.* **145** (1966) 1156 (cit. on p. 8).
- [11] T. Kibble, *Symmetry breaking in non-Abelian gauge theories*, *Phys. Rev.* **155** (1967) 1554 (cit. on p. 8).
- [12] T. Whyntie, *Introducing particle physics*, Icon Books, 2013, ISBN: 978-1848315891 (cit. on pp. 8, 9).
- [13] H. Yukawa, *On the Interaction of Elementary Particles. I*, *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series* **17** (1935) 48 (cit. on p. 10).
- [14] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in di-boson final states with the ATLAS detector at the LHC*, *Phys. Lett. B* **726** (2013) 88, arXiv: [1307.1427 \[hep-ex\]](https://arxiv.org/abs/1307.1427) (cit. on pp. 14, 47, 55, 92, 169), Erratum: *Phys. Lett. B* **734** (2014) 406.

- [15] ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, *Phys. Rev. D* **90** (2014) 052004, arXiv: [1406.3827 \[hep-ex\]](#) (cit. on pp. 14, 47, 49, 55, 169).
- [16] CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, *Eur. Phys. J. C* **75** (2015) 212, arXiv: [1412.8662 \[hep-ex\]](#) (cit. on pp. 14, 55, 56, 63).
- [17] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803, arXiv: [1503.07589 \[hep-ex\]](#) (cit. on pp. 14, 47, 55, 169).
- [18] ATLAS Collaboration, *Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, *Phys. Lett. B* **784** (2018) 345, arXiv: [1806.00242 \[hep-ex\]](#) (cit. on pp. 14, 29, 35, 38, 49, 51, 55–57, 87, 129–131, 143, 147, 169).
- [19] CMS Collaboration, *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **11** (2017) 047, arXiv: [1706.09936 \[hep-ex\]](#) (cit. on pp. 14, 55, 56, 63).
- [20] *A measurement of the Higgs boson mass in the diphoton decay channel*, tech. rep. CMS-HIG-19-004, CMS yy and combination Run 2 – supersedes CMS-PAS-HIG-19-004: CERN, 2019, URL: <https://cds.cern.ch/record/2709657> (cit. on pp. 14, 63).
- [21] LHC Higgs Cross Section Working Group, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the nature of the Higgs sector*, *CERN-2017-002-M* (CERN, Geneva, 2016), arXiv: [1610.07922 \[hep-ph\]](#) (cit. on pp. 14, 15).
- [22] J. Haller et al., *Update of the global electroweak fit and constraints on two-Higgs-doublet models*, *The European Physical Journal C* **78** (2018) 493, ISSN: 1434-6044, eprint: [1803.01853](#) (cit. on p. 15).
- [23] D. Buttazzo et al., *Investigating the near-criticality of the Higgs boson*, *Journal of High Energy Physics* **2013** (2013), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP12\(2013\)089](http://dx.doi.org/10.1007/JHEP12(2013)089) (cit. on p. 15).
- [24] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001 (cit. on p. 18).
- [25] E. Todesco and J. Wenninger, *Large Hadron Collider momentum calibration and accuracy*, *Phys. Rev. Accel. Beams* **20** (2017) 081003, URL: <https://link.aps.org/doi/10.1103/PhysRevAccelBeams.20.081003> (cit. on p. 18).
- [26] M. Baak et al., *Updated status of the global electroweak fit and constraints on new physics*, *The European Physical Journal C* **72** (2012), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-012-2003-4> (cit. on p. 19).
- [27] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003 (cit. on pp. 21, 23, 27).

- [28] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19, 2010, URL: <https://cds.cern.ch/record/1291633> (cit. on pp. 21, 27), Addendum: ATLAS-TDR-19-ADD-1, 2012, URL: <https://cds.cern.ch/record/1451888>.
- [29] ATLAS Collaboration, *ATLAS Muon Spectrometer: Technical Design Report*, ATLAS-TDR-10, 1997, URL: <https://cds.cern.ch/record/331068> (cit. on p. 23).
- [30] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, 2016, URL: <https://cds.cern.ch/record/2136007> (cit. on pp. 27, 150).
- [31] ATLAS Collaboration, *Trigger Menu in 2016*, ATL-DAQ-PUB-2017-001, 2017, URL: <https://cds.cern.ch/record/2242069> (cit. on pp. 27, 150).
- [32] ATLAS Collaboration, *Trigger Menu in 2017*, ATL-DAQ-PUB-2018-002, 2018, URL: <https://cds.cern.ch/record/2625986> (cit. on pp. 27, 150).
- [33] *Trigger menu in 2018*, tech. rep. ATL-DAQ-PUB-2019-001, CERN, 2019, URL: <https://cds.cern.ch/record/2693402> (cit. on pp. 27, 150).
- [34] *Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector at the LHC*, tech. rep. ATLAS-CONF-2020-005, CERN, 2020, URL: <https://cds.cern.ch/record/2714883> (cit. on pp. 27, 47, 143, 153).
- [35] ATLAS Collaboration, *Measurements of the Higgs boson inclusive and differential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV*, 2020, arXiv: [2004.03969](https://arxiv.org/abs/2004.03969) [[hep-ex](#)] (cit. on pp. 28, 76, 81–83, 153).
- [36] ATLAS Collaboration, *Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2020, arXiv: [2004.03447](https://arxiv.org/abs/2004.03447) [[hep-ex](#)] (cit. on pp. 28, 45, 76, 81–83, 153).
- [37] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C **76** \(2016\) 292](#), arXiv: [1603.05598](https://arxiv.org/abs/1603.05598) [[hep-ex](#)] (cit. on pp. 32, 85, 144).
- [38] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, [Nucl. Instrum. Meth. A **506** \(2003\) 250](#) (cit. on pp. 33, 154).
- [39] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, [JINST **14** \(2019\) P03017](#), arXiv: [1812.03848](https://arxiv.org/abs/1812.03848) [[hep-ex](#)] (cit. on p. 39).
- [40] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, (2019), arXiv: [1908.00005](https://arxiv.org/abs/1908.00005) [[hep-ex](#)] (cit. on pp. 39, 43, 85).
- [41] ATLAS Collaboration, *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, ATL-PHYS-PUB-2017-022, 2017, URL: <https://cds.cern.ch/record/2298955> (cit. on p. 43).

- [42] C. Oleari, *The POWHEG BOX*, Nuclear Physics B - Proceedings Supplements **205-206** (2010) 36, ISSN: 0920-5632, URL: <http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.016> (cit. on p. 43).
- [43] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Computer Physics Communications **178** (2008) 852, ISSN: 0010-4655, URL: <http://dx.doi.org/10.1016/j.cpc.2008.01.036> (cit. on p. 43).
- [44] ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with 36.1 fb^{-1} pp collision data at the center-of-mass energy of 13 TeV with the ATLAS detector*, ATLAS-CONF-2017-045, 2017, URL: <https://cds.cern.ch/record/2273852> (cit. on p. 43).
- [45] D. Guest, K. Cranmer, and D. Whiteson, *Deep Learning and its Application to LHC Physics*, Ann. Rev. Nucl. Part. Sci. **68** (2018) 161, arXiv: [1806.11484 \[hep-ex\]](#) (cit. on p. 50).
- [46] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. D **91** (2015) 012006, arXiv: [1408.5191 \[hep-ex\]](#) (cit. on p. 51).
- [47] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007, arXiv: [physics/0703039 \[physics.data-an\]](#) (cit. on pp. 51, 146, 147).
- [48] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: [0811.4622 \[hep-ph\]](#) (cit. on pp. 53, 155).
- [49] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: [0808.3674 \[hep-ph\]](#) (cit. on pp. 53, 155).
- [50] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: [1111.5206 \[hep-ph\]](#) (cit. on pp. 53, 155).
- [51] E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: [1905.09127 \[hep-ph\]](#) (cit. on pp. 53, 155).
- [52] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: [1405.0301 \[hep-ph\]](#) (cit. on p. 53).
- [53] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, arXiv: [1503.07589 \[hep-ex\]](#) (cit. on pp. 55, 56, 63).
- [54] L. A. Piegl and W. Tiller, “The NURBS book: Monographs in visual communication”, 2nd ed., Springer, 1997, chap. B-Spline basis function (cit. on p. 56).
- [55] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, URL: <https://cds.cern.ch/record/1449796> (cit. on p. 59).

- [56] M. J. Oreglia, “*A Study of the Reactions $\psi\prime \rightarrow \gamma\gamma\psi$* ”, Crystal Ball function discussed in appendix D., PhD thesis: SLAC, 1980 (cit. on p. 61).
- [57] J. E. Glaiser, “Charmonium spectroscopy from radiative decays of the J/ψ and ψ ”, Crystal Ball function defined on page 178., PhD thesis: Stanford University, 1982, URL: <https://www.slac.stanford.edu/cgi-bin/getdoc/slac-r-255.pdf> (cit. on p. 61).
- [58] T. Skwarnicki, “A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances”, Crystal Ball function discussed in appendix E., PhD thesis: Cracow, INP, 1986 (cit. on p. 61).
- [59] Martin Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, Software available from tensorflow.org, 2015, URL: <https://www.tensorflow.org/> (cit. on pp. 65, 147).
- [60] F. Chollet et al., *Keras*, <https://keras.io>, 2015 (cit. on pp. 65, 147).
- [61] K. S. Cranmer, *Kernel estimation in high-energy physics*, *Comput. Phys. Commun.* **136** (2001) 198, arXiv: [hep-ex/0011057](https://arxiv.org/abs/hep-ex/0011057) [[hep-ex](#)] (cit. on pp. 76, 169).
- [62] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, *Interpolation between multi-dimensional histograms using a new non-linear moment morphing method*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **771** (2015) 39, ISSN: 0168-9002, URL: <http://www.sciencedirect.com/science/article/pii/S0168900214011814> (cit. on p. 87).
- [63] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554, arXiv: [1007.1727](https://arxiv.org/abs/1007.1727) [[physics.data-an](#)], Erratum: *Eur. Phys. J. C* **73** (2013) 2501 (cit. on p. 92).
- [64] ATLAS Collaboration, *Technical Design Report for the ATLAS ITk Strip Detector*, tech. rep. ATL-COM-UPGRADE-2017-006, CERN, 2017, URL: <https://cds.cern.ch/record/2256817> (cit. on pp. 126, 127).
- [65] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, tech. rep. CERN-LHCC-2017-021. ATLAS-TDR-030, CERN, 2017, URL: <https://cds.cern.ch/record/2285585> (cit. on p. 129).
- [66] ATLAS Collaboration, *Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC*, ATL-PHYS-PUB-2018-054, 2018, URL: <https://cds.cern.ch/record/2652762> (cit. on pp. 131, 148).
- [67] ATLAS Collaboration, *Measurement of inclusive and differential cross sections in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **10** (2017) 132, arXiv: [1708.02810](https://arxiv.org/abs/1708.02810) [[hep-ex](#)] (cit. on pp. 142, 143).
- [68] ATLAS Collaboration, *Search for high-mass new phenomena in the dilepton final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **761** (2016) 372, arXiv: [1607.03669](https://arxiv.org/abs/1607.03669) [[hep-ex](#)] (cit. on pp. 142, 144).

- [69] S. H. Abidi et al., *Event selection and background estimation in the $H \rightarrow ZZ \rightarrow 4l$ channel at $\sqrt{s}=13\text{TeV}$ - Final Run 2 Analysis*, Internal note ATL-COM-PHYS-2018-1710, H4l full run 2 common internal note: CERN, 2018, URL: <https://cds.cern.ch/record/2651267> (cit. on p. 146).
- [70] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: [physics/0306116](https://arxiv.org/abs/physics/0306116) [[physics.data-an](#)] (cit. on p. 147).
- [71] L. Adamek et al., *Measurement of the Higgs boson mass in the $H \rightarrow ZZ(*) \rightarrow 4l$ channel at $\sqrt{s}=13\text{ TeV}$ with the 139 fb^{-1} with the ATLAS detector [Internal documentation]*, tech. rep. ATL-COM-PHYS-2019-1364, Full run 2 $H \rightarrow 4l$ mass measurement ATLAS internal note: CERN, 2020, URL: <https://cds.cern.ch/record/2697141> (cit. on p. 147).
- [72] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852, arXiv: [0710.3820](https://arxiv.org/abs/0710.3820) [[hep-ph](#)] (cit. on p. 153).
- [73] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 (cit. on p. 153).
- [74] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7\text{ TeV}$ with the ATLAS detector*, *JHEP* **09** (2014) 145, arXiv: [1406.3660](https://arxiv.org/abs/1406.3660) [[hep-ex](#)] (cit. on p. 153).
- [75] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: <https://cds.cern.ch/record/1966419> (cit. on p. 153).
- [76] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: [1005.4568](https://arxiv.org/abs/1005.4568) [[physics.ins-det](#)] (cit. on p. 153).
- [77] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, URL: <https://cds.cern.ch/record/2206965> (cit. on p. 154).
- [78] R. D. Ball et al., *Parton distributions with LHC data*, *Nuclear Physics B* **867** (2013) 244–289, ISSN: 0550-3213, URL: <http://dx.doi.org/10.1016/j.nuclphysb.2012.10.003> (cit. on p. 154).
- [79] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, *JHEP* **10** (2013) 222, arXiv: [1309.0017](https://arxiv.org/abs/1309.0017) [[hep-ph](#)] (cit. on p. 154).
- [80] K. Hamilton, P. Nason, and G. Zanderighi, *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, *JHEP* **05** (2015) 140, arXiv: [1501.04637](https://arxiv.org/abs/1501.04637) [[hep-ph](#)] (cit. on p. 154).
- [81] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, arXiv: [1002.2581](https://arxiv.org/abs/1002.2581) [[hep-ph](#)] (cit. on p. 154).
- [82] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: [hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146) (cit. on p. 154).

- [83] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, [*JHEP* **11** \(2007\) 070](#), arXiv: [0709.2092 \[hep-ph\]](#) (cit. on p. 154).
- [84] K. Hamilton, P. Nason, and G. Zanderighi, *MINLO: multi-scale improved NLO*, [*JHEP* **10** \(2012\) 155](#), arXiv: [1206.3572 \[hep-ph\]](#) (cit. on p. 154).
- [85] J. M. Campbell et al., *NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM*, [*JHEP* **07** \(2012\) 092](#), arXiv: [1202.5475 \[hep-ph\]](#) (cit. on p. 154).
- [86] K. Hamilton, P. Nason, C. Oleari, and G. Zanderighi, *Merging $H/W/Z + 0$ and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*, [*JHEP* **05** \(2013\) 082](#), arXiv: [1212.4504 \[hep-ph\]](#) (cit. on p. 154).
- [87] S. Catani and M. Grazzini, *An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC*, [*Phys. Rev. Lett.* **98** \(2007\) 222002](#), arXiv: [hep-ph/0703012](#) (cit. on p. 154).
- [88] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, [*J. Phys. G* **43** \(2016\) 023001](#), arXiv: [1510.03865 \[hep-ph\]](#) (cit. on p. 154).
- [89] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, [*JHEP* **02** \(2010\) 037](#), arXiv: [0911.5299 \[hep-ph\]](#) (cit. on p. 154).
- [90] H. B. Hartanto, B. Jager, L. Reina, and D. Wackeroth, *Higgs boson production in association with top quarks in the POWHEG BOX*, [*Phys. Rev. D* **91** \(2015\) 094003](#), arXiv: [1501.04498 \[hep-ph\]](#) (cit. on p. 154).
- [91] M. Wiesemann et al., *Higgs production in association with bottom quarks*, [*JHEP* **02** \(2015\) 132](#), arXiv: [1409.5301 \[hep-ph\]](#) (cit. on p. 154).
- [92] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, [*Phys. Rev. D* **82** \(2010\) 074024](#), arXiv: [1007.2241 \[hep-ph\]](#) (cit. on p. 154).
- [93] R. D. Ball et al., *Parton distributions for the LHC Run II*, [*JHEP* **04** \(2015\) 040](#), arXiv: [1410.8849 \[hep-ph\]](#) (cit. on p. 154).
- [94] U. Aglietti, R. Bonciani, G. Degrossi, and A. Vicini, *Two loop light fermion contribution to Higgs production and decays*, [*Phys. Lett.* **B595** \(2004\) 432](#), arXiv: [hep-ph/0404071 \[hep-ph\]](#) (cit. on p. 155).
- [95] M. Bonetti, K. Melnikov, and L. Tancredi, *Higher order corrections to mixed QCD-EW contributions to Higgs boson production in gluon fusion*, [*Phys. Rev. D* **97** \(2018\) 056017](#), [Erratum: *Phys. Rev. D* **97**, no. 9, 099906 (2018)], arXiv: [1801.10403 \[hep-ph\]](#) (cit. on p. 155).
- [96] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NNLO computational techniques: The cases $H \rightarrow \gamma\gamma$ and $H \rightarrow gg$* , [*Nucl. Phys. B* **811** \(2009\) 182](#), arXiv: [0809.3667 \[hep-ph\]](#) (cit. on p. 155).

- [97] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, [*Phys. Lett. B* **670** \(2008\) 12](#), arXiv: [0809.1301 \[hep-ph\]](#) (cit. on p. 155).
- [98] A. Pak, M. Rogal, and M. Steinhauser, *Finite top quark mass effects in NNLO Higgs boson production at LHC*, [*JHEP* **02** \(2010\) 025](#), arXiv: [0911.4662 \[hep-ph\]](#) (cit. on p. 155).
- [99] R. V. Harlander, H. Mantler, S. Marzani, and K. J. Ozeren, *Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass*, [*Eur. Phys. J.* **C66** \(2010\) 359](#), arXiv: [0912.2104 \[hep-ph\]](#) (cit. on p. 155).
- [100] R. V. Harlander and K. J. Ozeren, *Top mass effects in Higgs production at next-to-next-to-leading order QCD: Virtual corrections*, [*Phys. Lett.* **B679** \(2009\) 467](#), arXiv: [0907.2997 \[hep-ph\]](#) (cit. on p. 155).
- [101] R. V. Harlander and K. J. Ozeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, [*JHEP* **11** \(2009\) 088](#), arXiv: [0909.3420 \[hep-ph\]](#) (cit. on p. 155).
- [102] F. Dulat, A. Lazopoulos, and B. Mistlberger, *iHigs 2 - Inclusive Higgs cross sections*, [*Comput. Phys. Commun.* **233** \(2018\) 243](#), arXiv: [1802.00827 \[hep-ph\]](#) (cit. on p. 155).
- [103] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, [*Phys. Rev. Lett.* **114** \(2015\) 212001](#), arXiv: [1503.06056 \[hep-ph\]](#) (cit. on p. 155).
- [104] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, [*JHEP* **05** \(2016\) 058](#), arXiv: [1602.00695 \[hep-ph\]](#) (cit. on p. 155).
- [105] M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC*, [*Phys. Rev. Lett.* **99** \(2007\) 161803](#), arXiv: [0707.0381 \[hep-ph\]](#) (cit. on p. 155).
- [106] M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, [*Phys. Rev. D* **77** \(2008\) 013002](#), arXiv: [0710.4749 \[hep-ph\]](#) (cit. on p. 155).
- [107] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Higgs production via vector-boson fusion at NNLO in QCD*, [*Phys. Rev. Lett.* **105** \(2010\) 011801](#), arXiv: [1003.4451 \[hep-ph\]](#) (cit. on p. 155).
- [108] O. Brein, R. V. Harlander, and T. J. E. Zirke, *vh@nnlo - Higgs Strahlung at hadron colliders*, [*Comput. Phys. Commun.* **184** \(2013\) 998](#), arXiv: [1210.5347 \[hep-ph\]](#) (cit. on p. 155).
- [109] R. V. Harlander, J. Klappert, S. Liebler, and L. Simon, *vh@nnlo-v2: New physics in Higgs Strahlung*, [*JHEP* **05** \(2018\) 089](#), arXiv: [1802.04817 \[hep-ph\]](#) (cit. on p. 155).
- [110] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, [*Phys. Lett. B* **579** \(2004\) 149](#), arXiv: [hep-ph/0307206](#) (cit. on p. 155).

- [111] O. Brein, R. Harlander, M. Wiesemann, and T. Zirke, *Top-Quark Mediated Effects in Hadronic Higgs-Strahlung*, [*Eur. Phys. J.* **C72** \(2012\) 1868](#), arXiv: [1111.0761 \[hep-ph\]](#) (cit. on p. 155).
- [112] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, *Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK*, [*JHEP* **03** \(2012\) 075](#), arXiv: [1112.5142 \[hep-ph\]](#) (cit. on p. 155).
- [113] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, [*Comput. Phys. Commun.* **195** \(2015\) 161](#), arXiv: [1412.5390 \[hep-ph\]](#) (cit. on p. 155).
- [114] M. Ciccolini, S. Dittmaier, and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, [*Phys. Rev. D* **68** \(2003\) 073003](#), arXiv: [hep-ph/0306234 \[hep-ph\]](#) (cit. on p. 155).
- [115] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak, and T. J. E. Zirke, *Gluon-induced Higgs-strahlung at next-to-leading order QCD*, [*JHEP* **02** \(2013\) 078](#), arXiv: [1211.5015 \[hep-ph\]](#) (cit. on p. 155).
- [116] R. V. Harlander, A. Kulesza, V. Theeuwes, and T. Zirke, *Soft gluon resummation for gluon-induced Higgs Strahlung*, [*JHEP* **11** \(2014\) 082](#), arXiv: [1410.0217 \[hep-ph\]](#) (cit. on p. 155).
- [117] W. Beenakker et al., *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions*, [*Nucl. Phys. B* **653** \(2003\) 151](#), arXiv: [hep-ph/0211352](#) (cit. on p. 155).
- [118] S. Dawson, C. Jackson, L. Orr, L. Reina, and D. Wackeroth, *Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections*, [*Phys. Rev. D* **68** \(2003\) 034022](#), arXiv: [hep-ph/0305087](#) (cit. on p. 155).
- [119] Y. Zhang, W.-G. Ma, R.-Y. Zhang, C. Chen, and L. Guo, *QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider*, [*Phys. Lett. B* **738** \(2014\) 1](#), arXiv: [1407.1110 \[hep-ph\]](#) (cit. on p. 155).
- [120] S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao, and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, [*JHEP* **06** \(2015\) 184](#), arXiv: [1504.03446 \[hep-ph\]](#) (cit. on p. 155).
- [121] S. Dawson, C. Jackson, L. Reina, and D. Wackeroth, *Exclusive Higgs boson production with bottom quarks at hadron colliders*, [*Phys. Rev. D* **69** \(2004\) 074027](#), arXiv: [hep-ph/0311067](#) (cit. on p. 155).
- [122] S. Dittmaier, M. Krämer, and M. Spira, *Higgs radiation off bottom quarks at the Tevatron and the CERN LHC*, [*Phys. Rev. D* **70** \(2004\) 074010](#), arXiv: [hep-ph/0309204](#) (cit. on p. 155).
- [123] R. Harlander, M. Krämer, and M. Schumacher, *Bottom-quark associated Higgs-boson production: Reconciling the four- and five-flavour scheme approach*, (2011), arXiv: [1112.3478 \[hep-ph\]](#) (cit. on p. 155).

- [124] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, *Higgs production in association with a single top quark at the LHC*, [Eur. Phys. J. C **75** \(2015\) 267](#), arXiv: [1504.00611 \[hep-ph\]](#) (cit. on p. 155).
- [125] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, *tWH associated production at the LHC*, (2016), arXiv: [1607.05862 \[hep-ph\]](#) (cit. on p. 155).

Appendices

A Contributions to this work

This appendix summarizes my journey through this doctorate and explains the division of labor to produce the Full Run 2 Higgs mass measurement with ATLAS.

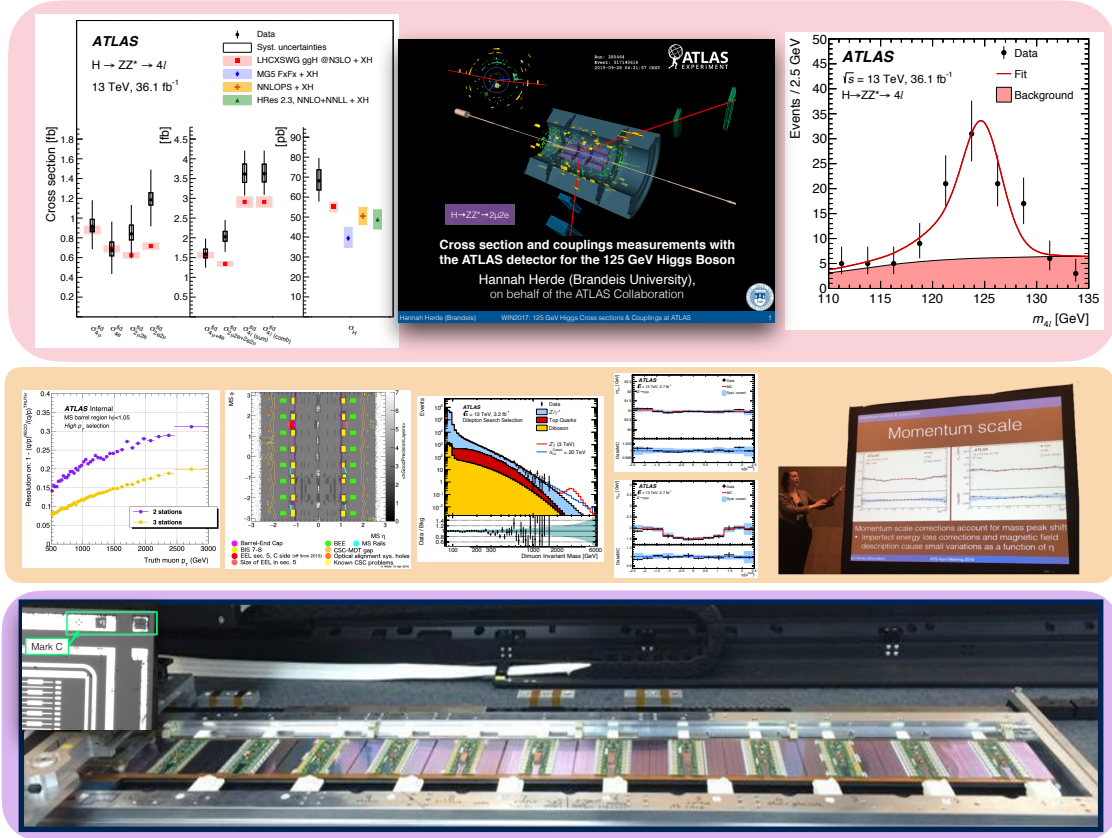


Figure 87: Thematic grouping of my projects during my doctoral work. **Higgs boson studies** (purple): I contributed to the mid-Run 2 fiducial and differential cross section measurements for a Higgs boson decaying to the four-lepton final state [67]. I also measured the mass of the Higgs boson in the same channel using the full Run 2 dataset (this work). **Performance work** (orange background): I developed an event selection algorithm to identify muons appropriate for analyses at high p_T , contributing to the first search at 13 TeV for a dilepton resonance decaying to two leptons [68]. I then transitioned to calibrating muon momentum to complement my studies of the Higgs boson. **Hardware** (purple background): Module loading for the ATLAS Phase II Inner Tracker upgrade project at Brookhaven National Laboratory and further developed consolidated the quality control criteria and strategy when I returned to CERN.

Source: HH (2014-2020)

A.1 The story of my doctoral work

Figure 87 displays images from the various projects spanning my six years pursuing this degree. I have grouped the projects thematically spanning three major areas of study: (1) Higgs boson studies in the four-lepton final state, (2) performance work, and (3) hardware.

A.1.1 Higgs boson studies in the four-lepton final state

I joined the ATLAS SM $H \rightarrow 4\ell$ group (HSG2) in January 2016.

In 2016, I measured the fiducial and total Higgs boson cross sections at 13 TeV using the 2015-2016 dataset [67]. My specific contributions included studying the impact of event migrations between fiducial phase spaces and observable bins on unfolding methods, the impact of “truth-level” isolation definitions in simulated datasets when defining the fiducial volume, and performing the profile likelihood fits to extract the fiducial and total cross sections. I presented 2015-2016 cross section and coupling measurements at WIN2017 on behalf of the Collaboration.

In late 2017, I joined the effort to measure the Higgs mass. I studied uncertainty decomposition with pseudoexperiments for then-outgoing paper with the 2015-2016 dataset (36 fb⁻¹) [18]. Then, I stayed on to develop improvements for the Full Run 2 measurement, eventually becoming the lead analyzer (or, “analysis contact” in ATLAS-speak) in July 2019. We recently published a conference note [34] and a paper with the per-event resolution strategy is forthcoming. I will tell that story in section A.2.

Additionally, on behalf of the $H \rightarrow ZZ^{(*)}$ analysis group, I maintained a common cross sections tool for simulated datasets (2016-2020). This work contributes to all $H \rightarrow ZZ^{(*)}$ publications. Starting in 2019, I also maintained the common statistical framework for likelihood construction (built on RooFit workspaces) on behalf of HSG2.

A.1.2 Performance

Prof. Sciolla and I chose each combined performance project to complement a concurrent physics analysis effort.

From 2015 to 2017, I designed and maintained an ATLAS-wide event selection algorithm to maximize the momentum resolution for high transverse momentum muon tracks with Dr. Artoni and Dr. Zambito. I showed that momentum resolution improves by a factor of two when requiring muon candidate’s track to contain all three muon spectrometer stations, compared to the ATLAS standard. Our work contributed to a journal article [37]. I also presented the our “first look” at ATLAS muon performance at 13 TeV at EPS 2015.

Dr. Artoni and I then used the algorithm to search for exotic decays of a Z' narrow resonance to the dilepton final states in 13 TeV events, resulting in another journal article [68]. We developed the sole event selection algorithm monitoring both the electron and muon channels with ATLAS analysis tools, C++, Python, and ROOT.

In 2017, I switched my performance focus to lower-momentum muons to again harmonize my performance studies with my analysis (at that point, measuring the Higgs mass). I measured the muon expected resolution in ATLAS Monte Carlo simulation. I also calibrated momentum resolution and scale in 2015-2017 data for the ATLAS Experiment with Dr. Barone. Finally, I compared the muon spectrometer alignments in 2015, 2016, and 2017 ATLAS data and their impact on the expected Higgs mass precision.

A.1.3 Hardware

I applied to the US Department of Energy Office (DOE) and received an SCGSR fellowship to spend a year in-residence (2016-2017) at Brookhaven National Laboratory (BNL) working with their Inner Tracker (ITk) Strips community, specifically with Dr. Lynn and Dr. van Nieuwenhuizen. The fellowship served two primary purposes. First, I designed a system and procedure for mounting strip modules on their local support and service structures (“staves”) in the central ITk section, working closely with physicists, engineers, and technicians. The

system is a custom coordinate measurement machine, using a camera mounted on an XYZ stage controlled and monitored by LabVIEW programming. It includes a full system calibration sequence and machine-vision-assisted module positioning routine, achieving final module positions well within their $25\text{ }\mu\text{m}$ tolerances. Second, as the first Brandeis student at BNL, I trained subsequent students to operate and maintain the mounting system. We mounted the US thermomechanical stave prototype while I was at BNL. I led our team of four Brandeis students through updating our system based on the experience.

ITk design and construction is a global endeavor. While the US will produce half of the staves for the strips system, the UK will build the other half. Back at CERN in autumn 2017, I worked closely with Dr. Dopke at Rutherford Lab (UK) to harmonize our approaches to module mounting. Soon, I became the *de facto* link between the US and UK module loading teams. Prior to the June 2018, I prepared and led weekly meetings between both stave loading teams to review our quality assurance and quality control (QA/QC) plans. I organized lists of tests and required equipment according to the Technical Design Report and US reports to the DOE. Dr. Dopke and I then worked with Prof. Stelzer at Simon Fraser University (Canada) to generalize to the wedge-shaped local support structures in the ITk endcaps. I presented our combined international strategy at the Oxford QA/QC meeting in 2018, and later at the module loading Preliminary Design Review.

A.2 Contributions to this dissertation and measurement

I would like to acknowledge Prof. Blocker for his particle physics class at Brandeis University, Prof. Lawrence for his quantum field theory class at Brandeis University, and Prof. Grossman for his set of Standard Model lectures in the 2014 CERN Summer Student program [2]. All were influential in my thinking and have left their marks on chapter I. Prof. Grossman's lectures were particularly impactful on my conceptualization of the Standard Model as, in the summer of 2014, they were my first encounter with subject matter.

I credit the CERN-FermiLab 2019 summer school with my understanding of the LHC. I

originally wrote the description of the ATLAS detector found in section 4 for my Advanced Exam, delivered at Brandeis University in January 2016.

Chapters III and IV are the bulk of this thesis. I started conceptualizing this particular way of telling the story in 2018 for the HZZ Workshop at Oxford University. I gave a talk there on the importance of muon and electron calibrations to the measurement, which strongly influenced my later thinking and my writing of sections 6.3 and 6.4. Many figures in that chapter began in some form in that talk. The rest of section 6 on the event selection describes in my own words work from the collective HSG2, detailed in reference [69].

Section 7 is the culmination of the $H \rightarrow 4\ell$ mass measurement analysis team from 2017 to 2020. Thomas Powell at the University of Sheffield developed the DCB model for his doctoral thesis (in preparation). Dr. Barone and I started thinking about a multivariate approach to deriving the event resolution in 2017.

Initially, I tried to calibrate the individual muon momentum errors, $\sigma_{q/p}$, analytically as a function of p_T , η , and the normalized χ^2 of the track. I tried something similar for the electron error, σ_e , as well. I hoped to propagate the calibrated errors of all four leptons through the ZMC to define my event error, intending to make a classic per-event errors measurement. Unfortunately, the parameter space is too complex for an explicitly analytic approach, so we turned to machine learning. We concentrated our machine-learning efforts on calibrating the four-lepton invariant mass error, $\sigma_{m_{4\ell}}$, directly to reduce complication.

Dr. Barone, who was the analysis contact at the time, and I spent most of 2018 and the first half of 2019 understanding how to put together a sound per-event model using Powell’s DCB as a foundation. Our original goal was to make the PDF conditional on both BDT_{ZZ} and the event resolution. We intended to express our event resolution as a scale factor f multiplied by $\sigma_{m_{4\ell}}$, $f \cdot \sigma_{m_{4\ell}}$. We tried to derive f using variously a BDT or NN regressor implemented in TMVA [47]. During this time, our regression targeted different permutations of $f = 1 + \frac{m_{4\ell}^{\text{truth,born}} - m_{4\ell}^{\text{reconstructed}}}{m_{4\ell}^{\text{truth,born}} \cdot \sigma_{m_{4\ell}}}$. In principle, this expression was designed to produce a calibrated event error equal to the measured error plus an estimate of the

“missing” error. In practice, it added a statistically-invalid contribution to each event error. The resulting pull on m_H was strongly shifted from zero and almost bimodal.

In June 2019, we changed approaches to try to derive the event resolution directly using a NN. At first, we targeted the mean of $m_{4\ell}^{\text{truth,born}} - m_{4\ell}^{\text{constrained}}$. As explained in section 7.3.2, the mean of that particular target corresponds to the *scale* of this measurement, not its *resolution*. I remembered a talk in a Muon Combined Performance meeting from Dr. Artoni in July 2018 about using a quantile regression to estimate the resolution of individual muons.³⁴ When I became analysis contact a month later, I recruited two more students, Bianca Ciungu and Lukas Adamek, both from the University of Toronto, to join the analysis. I shifted to coordinating the analysis, communicating the big picture, and helping the other students find their conceptual footing. Ciungu and Adamek took over the resolution regression. I shared with them my idea to switch to a quantile regressor. Ciungu implemented it as an NN using Keras [60] and Tensorflow [59]. At the end of September 2019, Ciungu showed the first results using the quantile regressor in an HZZ weekly meeting, using sampled MC results I produced with her resolution estimate, demonstrating the method’s viability. For the next six weeks, Powell, Ciungu, and I all produced workspaces with the average resolution approach as cross checks for each other using different frameworks based on RooFit [70]. After we received permission to form an editorial board in December 2019, I stopped producing average and per-event resolution workspaces myself.

Additionally, I produced the BDT_{ZZ} discriminant, originally defined for the 2015-2016 measurement [18], using updated ATLAS calibrations with TMVA [47]. I also produced the full template cross check.

I intended sections 7, 8, and chapter IV to summarize the full documentation for the measurement, available in our internal note [71]. Along with Powell and Ciungu, I served as an editor of this note. Where Powell or Ciungu was primary author in the original, I have tried to rephrase in my own words. Where I was the primary author in the original, the text

³⁴“Muon momentum resolution studies,” <https://indico.cern.ch/event/733703/>

is largely unchanged, most notably for sections 7.3, 7.5, and appendix H.

Finally, chapter V is in part a replication of the background section of my application for the SCGSR fellowship submitted in November 2015. I prepared the extrapolated expected statistical uncertainty on m_H myself, displayed in figure 85, while Dr. Barone and Dr. Gabrielli prepared the Higgs mass prospects work quoted in the 2018 PUB note [66].

B The mathematics of quantum electrodynamics

Quantum electrodynamics (QED) describes the interactions of electrically charged particles with the electromagnetic field using the language of quantum field theory. This appendix uses QED as a demonstration of gauge invariance.

We can describe any free charged fermion (like an electron) as an excitation of a fermionic Dirac field, $\psi(x)$. The field is a function Minkowski spacetime coordinates (the three spatial coordinates and a time coordinate).³⁵ The Dirac equation provides the Lagrangian of the free fermion's field:

$$\mathcal{L}_{Dirac} = \bar{\psi} (i\partial\!\!\!/ - m) \psi, \quad (21)$$

where $\partial\!\!\!/ = (\nabla^2 - \frac{\partial^2}{\partial t^2})$ and m is the excitation's mass.

In section 1.1.3, we described the U(1) symmetry group of rotations in one-dimensional complex space. These gauge transformations take the form:

$$\psi \rightarrow \psi' = U\psi \quad (22)$$

$$= e^{i\alpha}\psi. \quad (23)$$

Now, when α does not depend on spacetime, it is clear that the Lagrangian of the Dirac equation is symmetric under this transformation. When we allow α to depend on spacetime, we are inherently *localizing* these transformations. When $\alpha \rightarrow \alpha(x)$, we can apply different transformations at different points of the fermionic field. Enforcing this **local gauge invariance**, we have to introduce a kinetic term and a new vector field, A_μ ,

$$\mathcal{L} = \mathcal{L}_{Dirac} + \text{kinetic term}, \quad (24)$$

$$\mathcal{L} = \bar{\psi} (i\partial\!\!\!/ - m) \psi + e\bar{\psi}\gamma^\mu A_\mu\psi. \quad (25)$$

³⁵Lorentz transformations are rotations in Minkowski space such that $x^2 + y^2 + z^2 + (ict)^2 = \text{constant}$.

The new vector field’s transformation provides the kinetic term mentioned in section 1.1.3:

$$A_\mu \rightarrow U A_\mu U^\dagger = A_\mu + \frac{1}{e} \partial_\mu \alpha(x). \quad (26)$$

At this point, the charged fermion can move. It generates a current, $e\bar{\psi}\gamma^\mu\psi$, that interacts with the new vector field, which has a spin-1 excitation identified with the photon. We also know the structure of this interaction – the elementary charge e . Yet, as mentioned in section 1.1.3, equation 25 is not invariant “out of the box” under the $U(1)$ symmetry.

Completely integrating the electromagnetic field, A_μ , and therefore satisfying local gauge invariance, means that we need to allow the electromagnetic field itself to propagate through the vacuum, leading us to the familiar form of the QED Lagrangian:

$$\mathcal{L} = \bar{\psi} (i\not{D} - m) \psi + \frac{1}{4} F^{\mu\nu} F_{\mu\nu}, \quad (27)$$

with covariant derivative $\not{D} = \not{\partial} - ie\not{A}$ and the electromagnetic field tensor, $F^{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. In other words, the electromagnetic force saves itself here. The full implementation of its field preserving local gauge invariance completes the dynamics of our model. There is no quadratic term for A_μ because the photon is massless. To actually make QED fully renormalizable, we have to add the Faddeev-Popov Lagrangian to equation 27, omitted from this discussion.

C $H \rightarrow 4\ell$ ATLAS triggers during Run 2

This appendix lists all triggers used for full Run 2 $H \rightarrow 4\ell$ analyses in ATLAS. Tables 27-30 list the selected triggers by data collection period. Please note that these tables rely heavily on ATLAS jargon. Full public descriptions of the trigger menus in each data taking year are available in references [30–33]. We accept any events triggered with any unprescaled single-, double-, or triple-lepton trigger.

Table 27: Summary of the HLT triggers that are used during the 2015 data taking for the three analysis channels. When multiple chains are indicated, it is intended that the OR among them is requested. *Source: $H \rightarrow 4\ell$ event selection and background estimation common internal note (ATL-COM-PHYS-2018-1710)*

Channel	Single-lepton	Di-lepton	Tri-lepton
$4e$	e24_lhmedium_L1EM18VH (Period D) e24_lhmedium_L1EM20VH (Period E-J) e60_lhmedium e120_lhloose	2e12_lhloose_L12EM10VH	e17_lhloose_2e9_lhloose
4μ	mu20_iloose_L1MU15 mu40 mu60_0eta105_msonly	2mu10 mu18_mu8noL1	3mu6 3mu6_msonly mu18_2mu4noL1
$2e2\mu$	$4e$ OR 4μ	$4e$ OR 4μ OR e17_lhloose_mu14 e24_medium_L1EM20VHI_mu8noL1 e7_medium_mu24	$4e$ OR 4μ OR 2e12_lhloose_mu10 e12_lhloose_2mu10

Table 28: Summary of the HLT triggers that are used during the 2016 data taking for the three analysis channels. When multiple chains are indicated, it is intended that the OR among them is requested. *Source: $H \rightarrow 4\ell$ event selection and background estimation common internal note (ATL-COM-PHYS-2018-1710)*

Channel	Single-lepton	Di-lepton	Tri-lepton
$4e$	e24_lhmedium_ivarloose (Period A-D3) e26_lhtight_nod0_ivarloose (Period D4-F) e60_lhmedium_nod0 e140_lhloose_nod0 e300_etcut (Period A-D3)	2e15_lhvloose_nod0_L12EM13VH (Period A-D3) 2e17_lhvloose_nod0 (Period D4-F)	e17_lhloose_nod0_2e9_lhloose_nod0 (Period A-D3) replaced by HLT_e17_lhloose_nod0_2e10_lhloose_nod0_L1EM15VH_3EM8VH when instantaneous luminosity $> 1.2 \cdot 10^{34} \text{cm}^{-2}\text{s}^{-1}$
4μ	mu24_ivarloose_L1MU15 (Period A) mu24_iloose_L1MU15 (Period A) mu24_ivarmedium (Period B-E) mu24_imedium (Period B-E) mu26_ivarmedium (Period D4-) mu26_imedium (Period D4-E2) mu40 (Period A) mu50	2mu10 (Period A) 2mu10_nomucomb (Period A) 2mu14 (Period B-) 2mu14_nomucomb (B-D3) mu20_mu8noL1 (Period A-E) mu20_nomucomb_mu6noL1_nscan03 (Period A-D3) mu22_mu8noL1 (Period D4-F)	3mu4 (Period A, 14-) 3mu6 (Period B-D3) 3mu6_msonly (Period D4-F) mu11_nomucomb_2mu4noL1_nscan03_L1MU11_2MU6 (Period A-D3) mu20_2mu4noL1 mu20_nomucomb_mu6noL1_nscan03 mu20_msonly_mu10noL1_msonly_nscan05_noComb (Period A-D3)
$2e2\mu$	$4e$ OR 4μ	$4e$ OR 4μ OR e17_lhloose_nod0_mu14 (Period A-D3) e24_lhmedium_nod0_L1EM20VHI_mu8noL1 (Period A-D3) e7_lhmedium_nod0_mu24 (Period A-D3) e17_lhloose_mu14 (Period D4-F) e24_lhmedium_L1EM20VHI_mu8noL1 (Period D4-F) e7_lhmedium_mu24 (Period D4-F)	$4e$ OR 4μ OR HLT_e17_lhloose_nod0_mu14 HLT_e24_lhmedium_nod0_L1EM20VHI_mu8noL1 (Period A-D3) HLT_e26_lhmedium_nod0_L1EM22VHI_mu8noL1 (Period D4-) HLT_e7_lhmedium_nod0_mu24 HLT_e12_lhloose_nod0_2mu10 HLT_2e12_lhloose_nod0_mu10

Table 29: Summary of the HLT triggers that are used during the 2017 data taking for the three analysis channels. When multiple chains are indicated, it is intended that the OR among them is requested. *Source: $H \rightarrow 4\ell$ event selection and background estimation common internal note (ATL-COM-PHYS-2018-1710)*

Channel	Single-lepton	Di-lepton	Tri-lepton
$4e$	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0 HLT_e300_etcut	HLT_2e17_lhvloose_nod0_L12EM15VHI ³⁶ HLT_2e24_lhvloose_nod0	HLT_e24_lhvloose_nod0_2e12_lhvloose_nod0_L1EM20VH_3EM10VH
4μ	HLT_mu26_ivarmedium HLT_mu50 HLT_mu60_0eta105_msonly	HLT_2mu14 HLT_mu22_mu8noL1 HLT_mu22_mu8noL1_calotag_0eta010	HLT_mu20_2mu4noL1 HLT_3mu6_msonly HLT_3mu4 (below 1.2e34) HLT_3mu6 HLT_4mu4
$2e2\mu$	$4e$ OR 4μ	$4e$ OR 4μ OR HLT_e17_lhloose_nod0_mu14 HLT_e26_lhmedium_nod0_mu8noL1 HLT_e7_lhmedium_nod0_mu24	$4e$ OR 4μ OR HLT_e12_lhloose_nod0_2mu10 HLT_2e12_lhloose_nod0_mu10

Table 30: Summary of the HLT triggers that are used during the 2018 data taking for the three analysis channels. When multiple chains are indicated, it is intended that the OR among them is requested. *Source: $H \rightarrow 4\ell$ event selection and background estimation common internal note (ATL-COM-PHYS-2018-1710)*

Channel	Single-lepton	Di-lepton	Tri-lepton
$4e$	HLT_e26_lhtight_nod0_ivarloose HLT_e26_lhtight_nod0 ³⁷ HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0 HLT_e300_etcut	HLT_2e17_lhvloose_nod0_L12EM15VHI HLT_2e24_lhvloose_nod0 (L1_2EM20VH)	HLT_e24_lhvloose_nod0_2e12_lhvloose_nod0_L1EM20VH_3EM10VH
4μ	HLT_mu26_ivarmedium HLT_mu50 HLT_mu60_0eta105_msonly	HLT_2mu14 (L1_2MU10) HLT_mu22_mu8noL1 (L1_MU20)	HLT_mu20_2mu4noL1 (L1_MU20) HLT_3mu6 (L1_3MU6)
$2e2\mu$	$4e$ OR 4μ	$4e$ OR 4μ OR HLT_e17_lhloose_nod0_mu14 HLT_e26_lhmedium_nod0_mu8noL1 HLT_e7_lhmedium_nod0_mu24	$4e$ OR 4μ OR HLT_e12_lhloose_nod0_2mu10 HLT_2e12_lhloose_nod0_mu10

³⁶L1_2EM15VHI was accidentally prescaled in periods B5-B8 (runs 326834-328393) with an effective reduction of $0.6fb^{-1}$

³⁷Unprescaled since run 349169

D Simulating $H \rightarrow ZZ^* \rightarrow 4\ell$ events

This appendix provides information on the Monte Carlo simulation of events used to study the expected performance of the mass measurement. The original description, quoted here, appears in the conference note [34], edited by Dr. Leight and Dr. Anastopoulos.

We simulated the production of the Standard Model (SM) Higgs boson via most production modes at minimum next-to-leading-order, including gluon fusion (ggH), vector boson fusion (VBF), with an associated vector boson (VH , where V is a W or a Z boson), with a top quark pair (ttH), in association with a bottom quark pair (bbH), and in association with a single top quark ($tH+X$ where X is either jb or W , defined in the following as tH). We exceptionally simulated $gg \rightarrow ZH$ at leading order. The recent cross sections papers [35, 36] provide the full description.

Tables 32 and 33 below list the generators and parton distribution function sets used and the cross-section calculations used to normalize the predictions, respectively. For all signal processes, PYTHIA 8 [72] was used for the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ decay as well as for parton showering, hadronization, and the underlying event. PYTHIA 8 was interfaced to EVTGEN v1.2.0 [73] for simulation of the bottom and charm hadron decays. For the ggH , VBF , and VH processes, the AZNLO [74] set of tuned parameters was used, while the A14 [75] set was used for ttH , bbH and tH processes. For ggH , VBF , and, VH production, we simulated samples with a range of m_H values from 123 to 127 GeV, while other production modes are simulated assuming $m_H=125$ GeV.

The modeling of the $ZZ^{(*)}$ continuum background, as well as secondary backgrounds from WZ production, Z bosons produced in association with jets, and $t\bar{t}$ events, and also the small backgrounds from VVV or tXX production, are also described in reference [36]. Some details are given in Table 34. The alternate $qqZZ$ samples are used for systematic uncertainty studies.

Generated events are processed through the ATLAS detector simulation [76] within the

GEANT4 framework [38] and reconstructed in the same way as collision data. Additional pp interactions in the same and nearby bunch crossings, referred to as pile-up, are included in the simulation. The pile-up events are generated using PYTHIA 8 with the A3 set of tuned parameters [77] and the NNPDF2.3LO PDF set [78]. The simulation samples are weighted to reproduce the observed distribution of the mean number of interactions per bunch crossing in the data.

Table 31: List of m_H values used in the production of simulated ggH, VBF, and VH samples.
Source: [ATLAS-CONF-2020-005](#)

Production Mode	m_H values simulated (GeV)
ggH	123.0, 124.0, 124.5, 125.0, 125.5, 126.0, 127.0
VBF	123.0, 124.0, 125.0, 125.5, 126.0, 127.0
VH excepting $gg \rightarrow ZH$	124.0, 124.5, 125.0, 125.5, 126.0

Table 32: The Monte Carlo generators and parton distribution functions (PDFs) used to generate the signal samples.
Source: [ATLAS-CONF-2020-005](#)

Process	Generator	PDF
ggF	NNLOPS [79–83], which reweights the Higgs y spectrum from HJ-MINLO [84–86] to HNNLO [87]	PDF4LHC [88] (NNLO)
VBF	POWHEG-BOX v2 [81–83, 89]	PDF4LHC (NLO)
VH	POWHEG-BOX v2	PDF4LHC (NLO)
ttH	POWHEG-BOX v2 [90]	PDF4LHC (NLO)
bbH	MADGRAPH5_aMC@NLO v.2.3.3 [91]	CT10 NLO [92]
tH	MADGRAPH5_aMC@NLO v2.6.0	NNPDF30 [93]

Table 33: Description of cross section predictions used to normalize the MC signal samples, the accuracy of the calculations (in QCD if not noted otherwise), and the composition of the production modes in the SM. N³LO is the abbreviation for next-to-next-to-next-to-leading order and EW for electroweak. *Source:* [ATLAS-CONF-2020-005](#)

Process	Accuracy	Fraction [%]
ggF	N ³ LO, NLO EW corrections [94–104]	87.4
VBF	NLO, NLO EW corrections [105, 106] with approximate NNLO QCD corrections [107]	6.8
VH (excluding $gg \rightarrow ZH$)	NNLO [108–111], NLO EW corrections [112–114]	3.5
$gg \rightarrow ZH$	NLO [108, 109, 115] with NLL QCD corrections [116]	0.6
$t\bar{t}H$	NLO [117–120]	0.9
$b\bar{b}H$	five-flavour: NNLO, four-flavour: NLO [121–123]	0.9
tH	NLO [124, 125]	0.2

Table 34: Description of the production of simulated background samples. *Source:* [ATLAS-CONF-2020-005](#)

Process	Generator	PDF
$qqZZ$	SHERPA v2.2.2 [48–51]	NNPDF3.0nnlo
$ggZZ$	SHERPA v2.2.2	NNPDF3.0nnlo
EW ZZ	SHERPA v2.2.2	NNPDF3.0nnlo
$qqZZ$ (alternate)	POWHEG-BOX v2	CT10
$qqZZ$ (alternate)	MADGRAPH5_aMC@NLO	PDF4LHC (NLO)
WZ	POWHEG-BOX v2	CT10
Z +jets	SHERPA v2.2.1	NNPDF3.0nnlo
$t\bar{t}$	POWHEG-BOX v2	NNPDF3.0nnlo
tXX	MADGRAPH5_aMC@NLO	NNPDF3.0nnlo
VVV	SHERPA v2.2.2	NNPDF3.0nnlo

E Calibrating the QRNN target percentile in each decay channel

This pull distribution tells us that the QRNN score derived when targeting the 68th percentile of the $|m_{4\ell}^{\text{constrained}} - m_{4\ell}^{\text{truth,born}}|$ distribution is in the neighborhood of the percentile corresponding to the DCB width. The width of the pull is close to 1.0. We developed a scan to determine the appropriate target percentile for each decay channel:

1. Train 10 separate QRNNs.
 - Each targets a different percentile between 56% and 68%.
 - The scan interval is designed to exclude the tails of the $m_{4\ell}$ distribution, as we will model them with the DCB's exponential tails.
2. For each QRNN:
 - (a) Build a DCB (equation 14) where $\sigma_{CB} = f \cdot r_{\text{QRNN}}$ -
 - f is a scale factor,
 - r_{QRNN} is the event resolution, defined as the QRNN score for a network targeting a given percentile;
 - (b) Fit the DCB to simulated events scaled to our expected luminosity -
 - Perform the fit *conditional* to r_{QRNN} ,
 - Allow all other DCB parameters to vary so that they find their best-fit values conditional on r_{QRNN} ;
 - (c) This process associates one set of DCB parameters plus some scale factor f with each target percentile.
3. Plot f as a function of the target percentile.

4. $f = 1$ indicates the calibrated target percentile for the tested decay channel.

5. Extract calibrated target percentile with a linear regression such that

$$f(\text{targetquantile}) = 1.$$

This procedure absorbs the variations from changes in the DCB parameters into f *by design*. Therefore, we cannot impose any assumptions on the tail parameters while choosing our target quantile.

Figure 88 displays f as a function of the target percentile for each decay channel. In each plot, the x-axis indicates the target quantile, between 56% and 68%. Each plot also has a left and a right y-axis. The left y-axis provides the value of f at each percentile step (black points), where $f = 1$ indicates the calibrated target percentile for a given decay channel. The DCB parameters are also highly correlated, particularly with the width. To check that the parameters do not vary drastically during the percentile scan, we took the ratio of each DCB parameter (μ_{CB} , α_i , n_i where $i = 1, 2$) at a given percentile step with respect to its value under the Gaussian Hypothesis (QRNN targeting 68th percentile). The right y-axis reports these ratios each DCB parameter represented with a colored line on the plot. The most sensitive parameter, μ_{CB} , is flat across the full range of the scan in all decay channels. Most parameters change less than 1% over the full range of the scan, except n_1 and n_2 in the $2e2\mu$ channel. Fortunately, n_1 and n_2 variations affect m_H by less than 1 MeV. The plots report the extracted target percentile for each decay channel based on the linear regression. Finally, we trained new QRNNs for each decay channel using their calibrated target percentiles. This final set of QRNNs represents our event resolution.

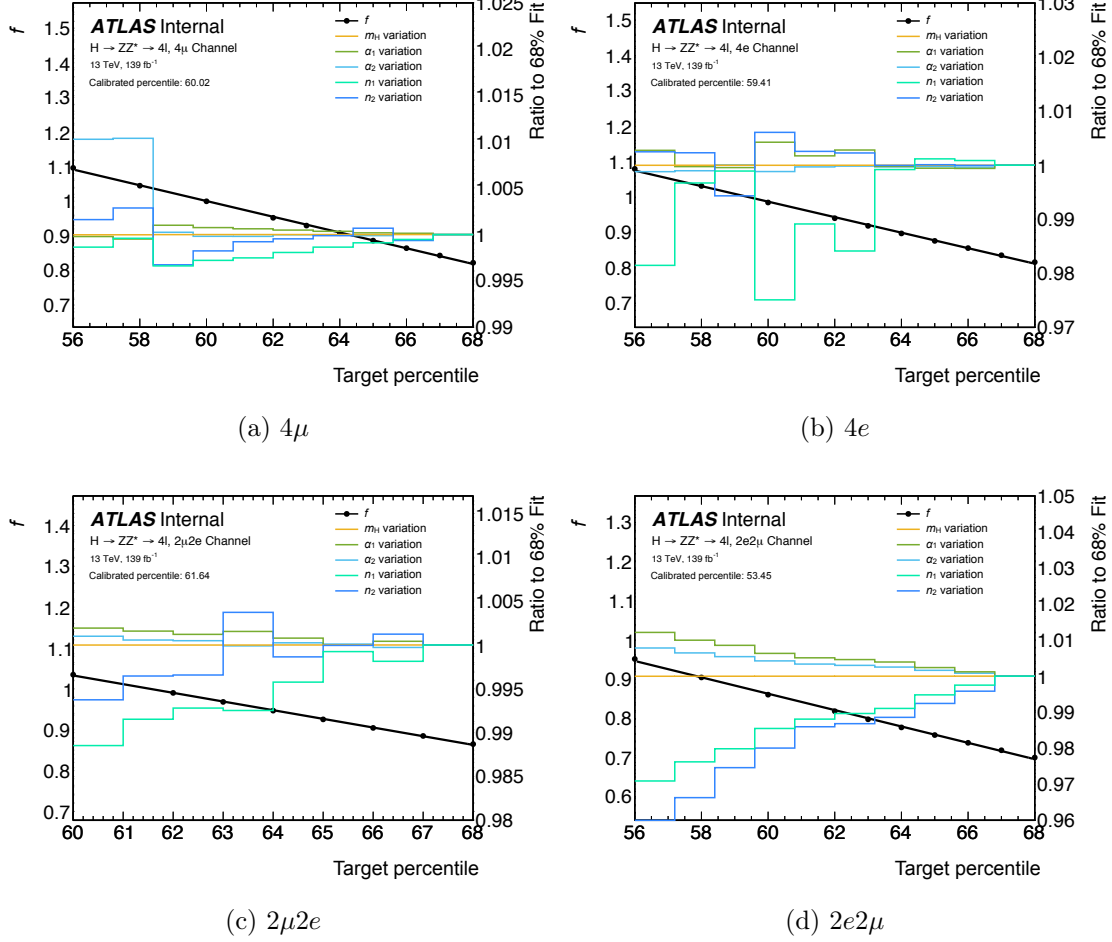


Figure 88: 10 NNs targeting percentiles between 56% and 68% in 1% steps were trained in each decay channel. A DCB defined by equation 14 with $\sigma = f \cdot r_{\text{QRNN}}$ is fit conditional to each output QRNN score, r_{QRNN} , producing 10 sets of DCB parameters. All other parameters are allowed to vary to find their best-fit values during the fit. This figure plots the DCB fit results at each 1% step in target percentile. The left axis shows the value of f , indicated with black points. The right axis provides the ratio of each DCB parameter (m_H , α_1 , n_1 , α_2 , and n_2 ; indicated with solid colored lines in each plot) at a given step with respect to its value under the Gaussian hypothesis. The calibrated target percentile is extracted with a linear fit requiring f (target quantile) = 1. *Source: $H \rightarrow 4\ell$ mass measurement group*

F Event-level resolution

This appendix includes the performance distributions for the quantile regression neural network in the $4e$, $2\mu 2e$, and $2e2\mu$ channels.

F.1 $4e$ Channel

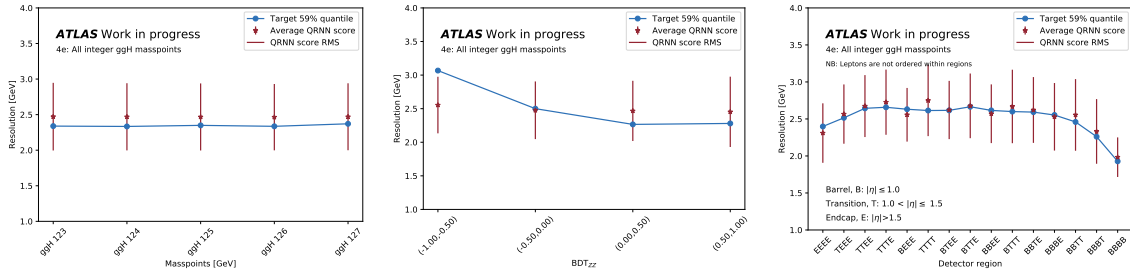


Figure 89: Summaries comparing the target quantile (blue) and QRNN score (red) in different phase spaces using all ggH mass points. The left plot compares resolution for each ggH mass point. The middle plot compares resolution in each BDT_{ZZ} bin. The right plot compares resolution based on the locations of the four leptons in the ATLAS detector. Each horizontal-axis position corresponds to a summary extracted from a Figure 44-like plot.

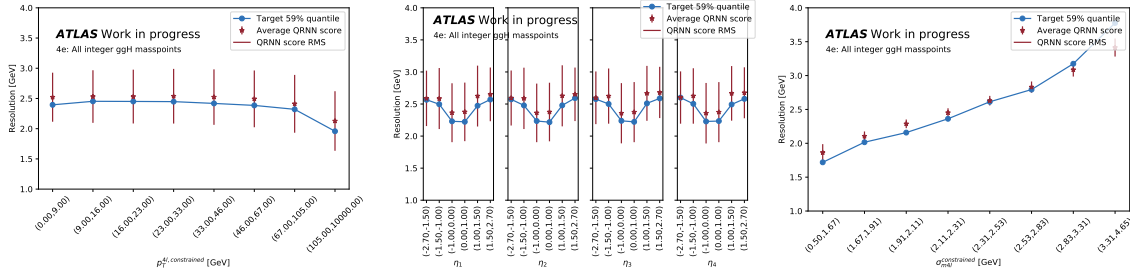


Figure 90: Summaries comparing the target quantile (blue) and QRNN score (red) in slices of several variables used in the training. The left plot considers the resolution along $p_{T,4\ell}^{\text{constrained}}$. The middle plot gives resolution along lepton pseudorapidity, η . It is broken into four subplots, one for each lepton. The right plot provides the resolution along $\sigma_{m_{4\ell}}^{\text{constrained}}$. Note that some disagreement is expected along $\sigma_{m_{4\ell}}^{\text{constrained}}$ since the error from the ZMC is not calibrated.

F.2 $2\mu 2e$ Channel

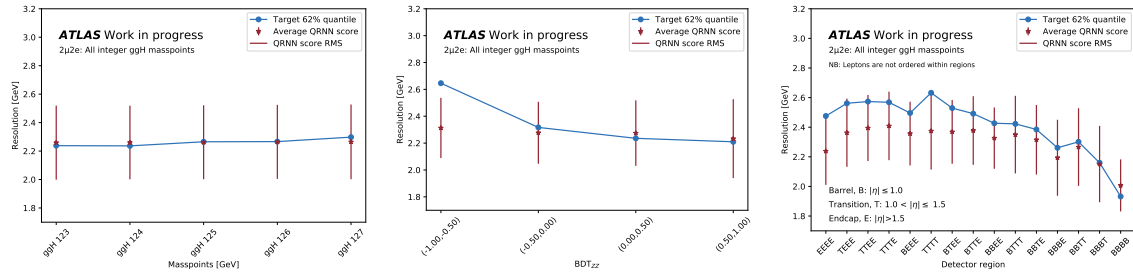


Figure 91: Summaries comparing the target quantile (blue) and QRNN score (red) in different phase spaces using all ggH mass points. The left plot compares resolution for each ggH mass point. The middle plot compares resolution in each BDT_{ZZ} bin. The right plot compares resolution based on the locations of the four leptons in the ATLAS detector. Each horizontal-axis position corresponds to a summary extracted from a Figure 44-like plot.

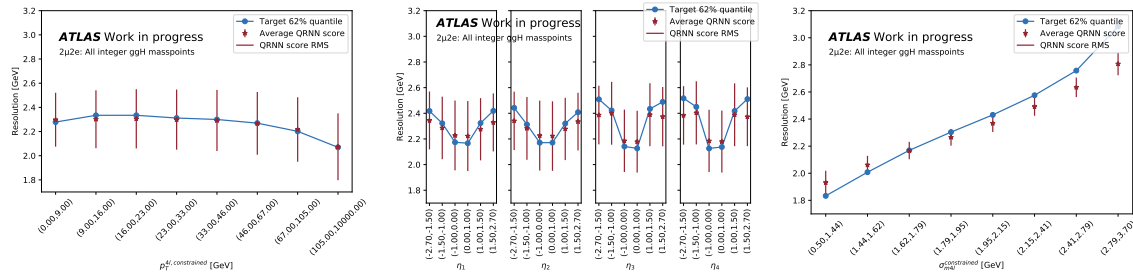


Figure 92: Summaries comparing the target quantile (blue) and QRNN score (red) in slices of several variables used in the training. The left plot considers the resolution along $p_{T,4\ell}^{\text{constrained}}$. The middle plot gives resolution along lepton pseudorapidity, η . It is broken into four subplots, one for each lepton. The right plot provides the resolution along $\sigma_{m_{4\ell}}^{\text{constrained}}$. Note that some disagreement is expected along $\sigma_{m_{4\ell}}^{\text{constrained}}$ since the error from the ZMC is not calibrated.

F.3 $2e2\mu$ Channel

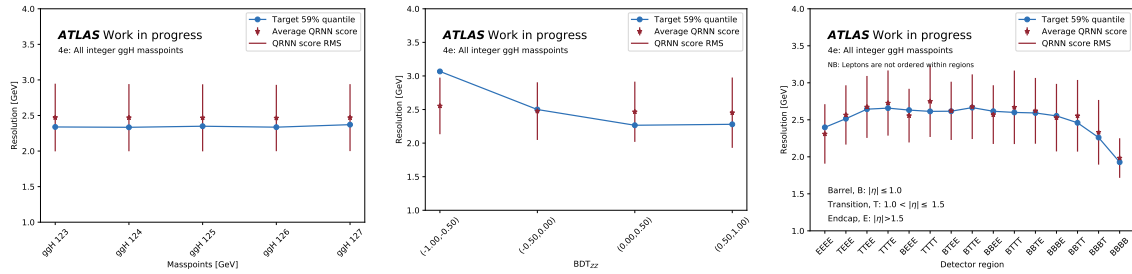


Figure 93: Summaries comparing the target quantile (blue) and QRNN score (red) in different phase spaces using all ggH mass points. The left plot compares resolution for each ggH mass point. The middle plot compares resolution in each BDT_{ZZ} bin. The right plot compares resolution based on the locations of the four leptons in the ATLAS detector. Each horizontal-axis position corresponds to a summary extracted from a Figure 44-like plot.

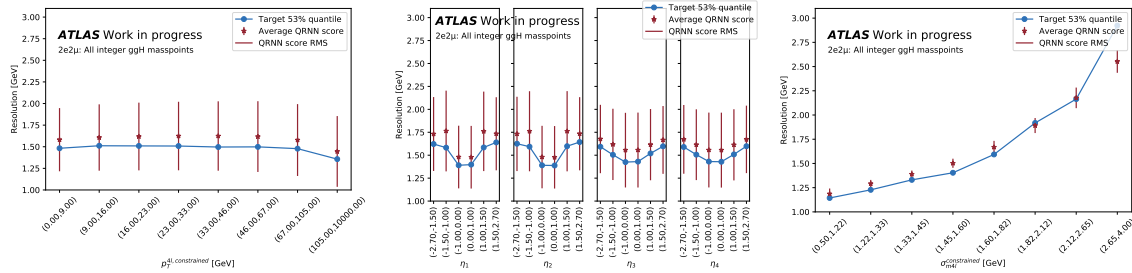


Figure 94: Summaries comparing the target quantile (blue) and QRNN score (red) in slices of several variables used in the training. The left plot considers the resolution along $p_{T,4\ell}^{\text{constrained}}$. The middle plot gives resolution along lepton pseudorapidity, η . It is broken into four subplots, one for each lepton. The right plot provides the resolution along $\sigma_{m_{4\ell}}^{\text{constrained}}$. Note that some disagreement is expected along $\sigma_{m_{4\ell}}^{\text{constrained}}$ since the error from the ZMC is not calibrated.

F.4 QRNN resolution data-simulation agreement

Figure 95 compares the agreement in QRNN score for data and MC simulation using sidebands around the expected m_H peak. The sidebands are $105 < m_{4\ell}^{\text{constrained}} < 115$ GeV and $130 < m_{4\ell}^{\text{constrained}} < 160$ GeV, comprising the same regions driving the fit of the non-resonant $ZZ^{(*)}$ background. Based on this comparison, we expect consistent network performance in data and simulation.

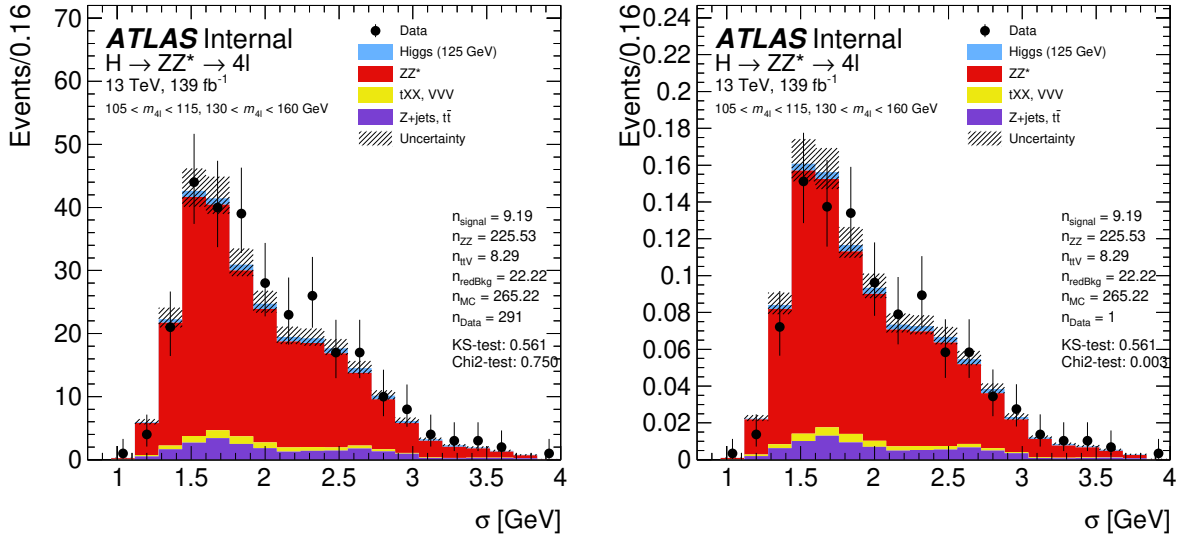


Figure 95: Comparison of the QRNN score in data and simulation using side bands. The normalized version of the same plot is shown to the right.

G Background PDF shapes

This appendix supplements section 7.5. Figures 96-100 display the shapes of the background PDFs for each background process. All of these figures are available in the [ATLAS-CONF-2020-005 internal documentation](#).

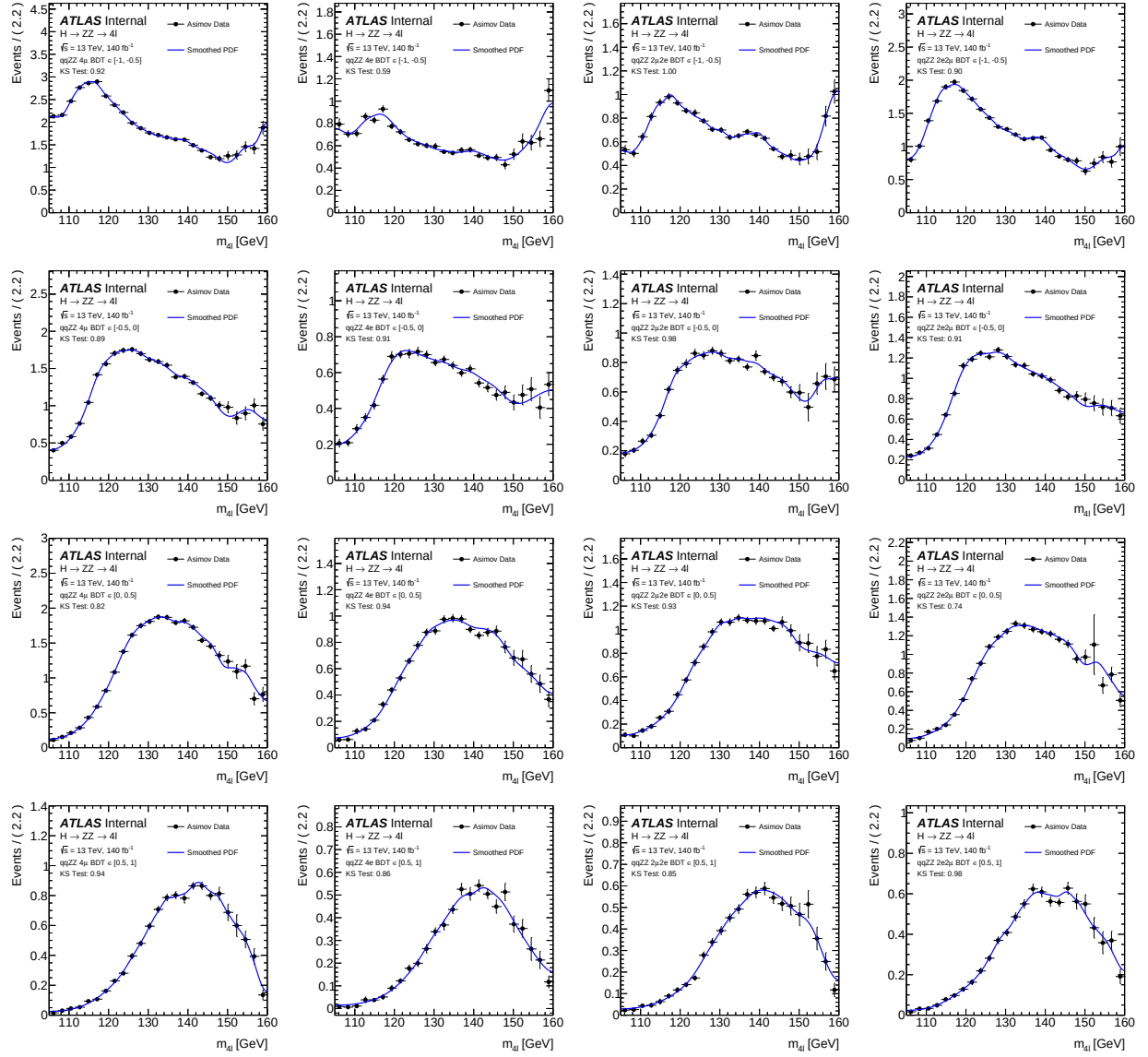


Figure 96: Result of KEYS smoothing of $qq \rightarrow ZZ^*$ used as the PDF in the final mass fit. Shown for the full 16 categories, with columns separating decay channels ($4\mu, 4e, 2\mu 2e, 2e 2\mu$), and rows separating BDT_{ZZ} bins (background-like to signal-like from top to bottom).

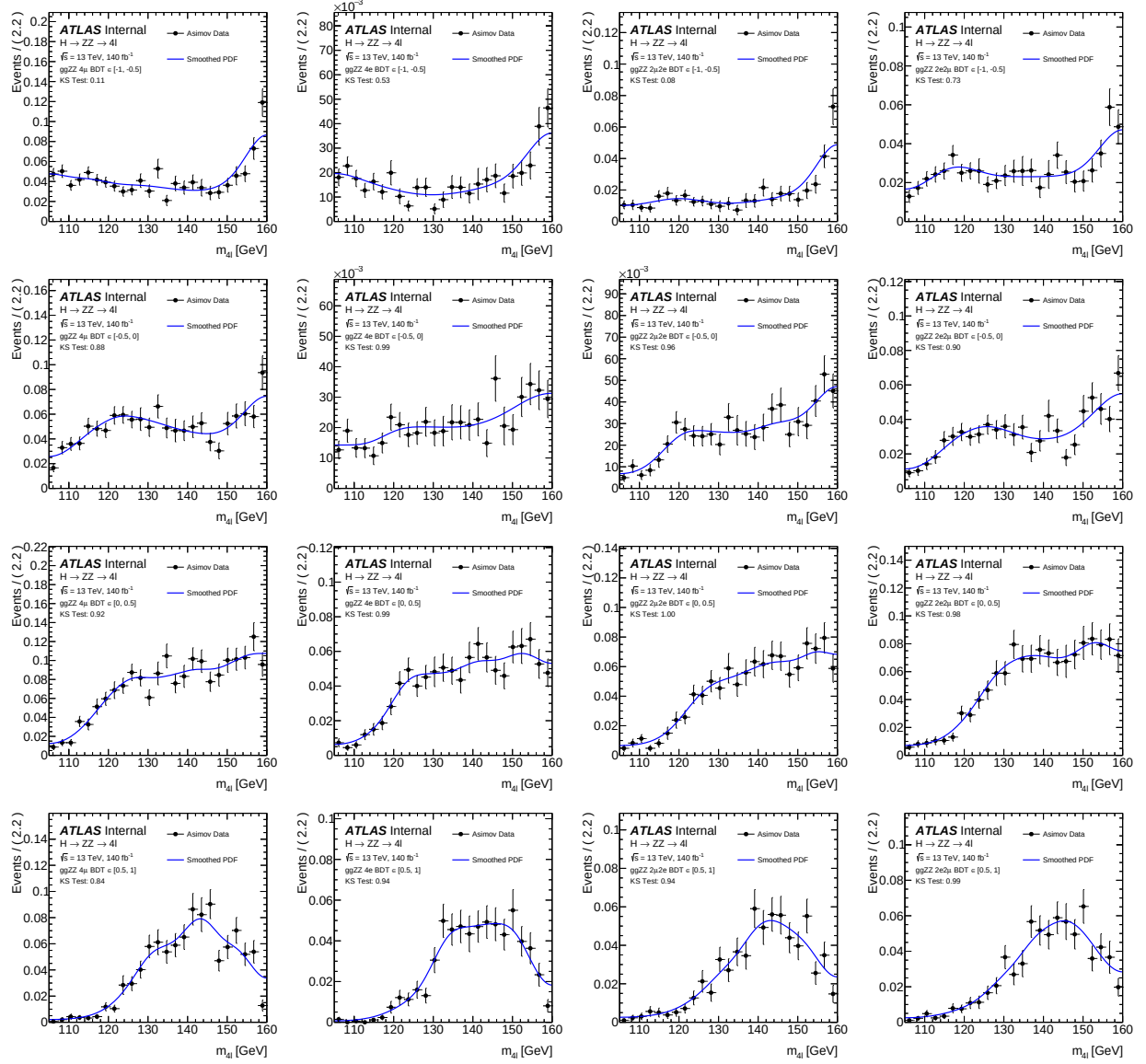


Figure 97: Result of KEYS smoothing of $gg \rightarrow ZZ^*$ used as the PDF in the final mass fit. Shown for the full 16 categories, with columns separating decay channels (4μ , $4e$, $2\mu 2e$, $2e 2\mu$), and rows separating BDT_{ZZ} bins (background-like to signal-like from top to bottom).

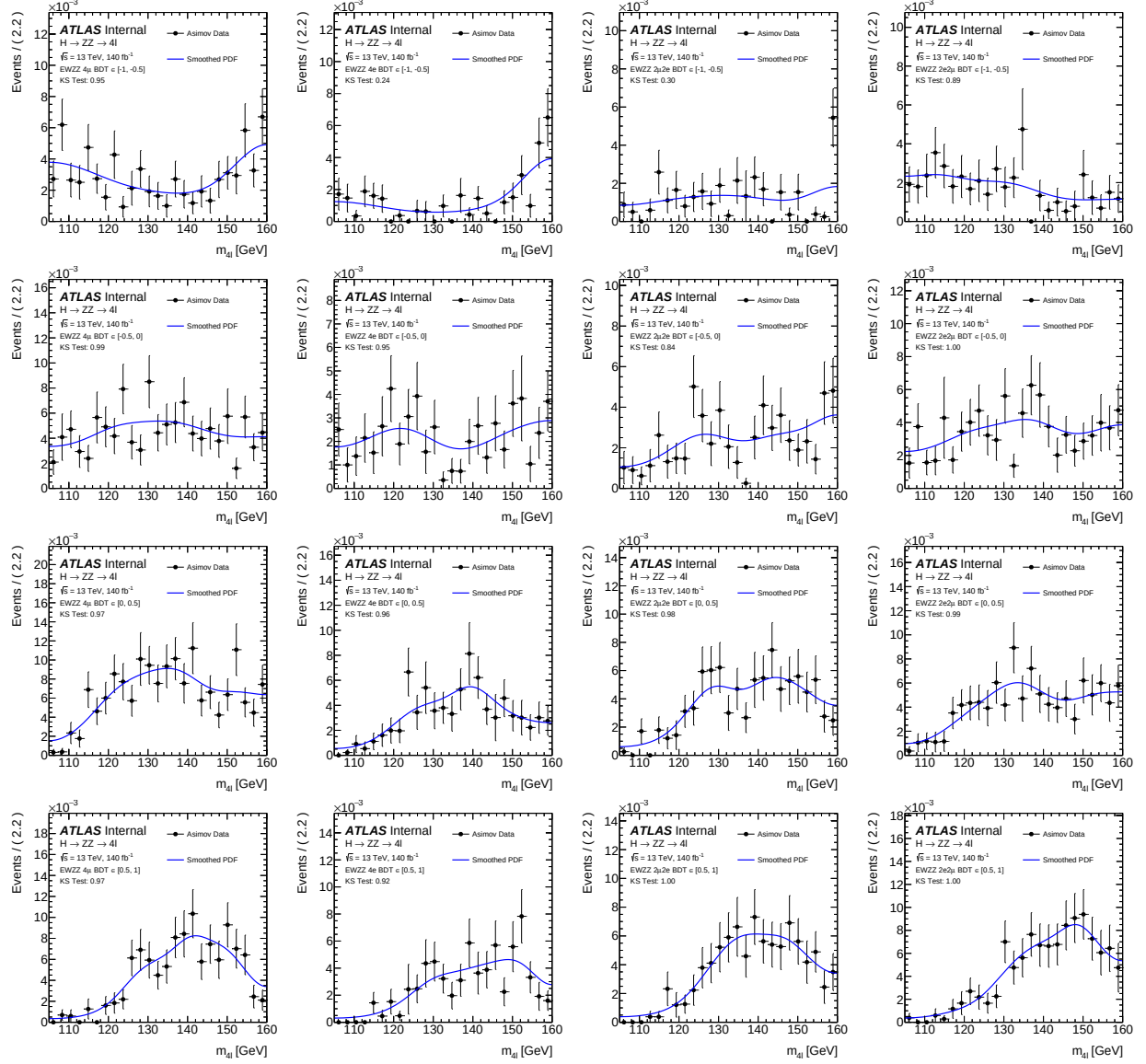


Figure 98: Result of KEYS smoothing of $EWZZ$ used as the PDF in the final mass fit. Shown for the full 16 categories, with columns separating decay channels (4μ , $4e$, $2\mu 2e$, $2e 2\mu$), and rows separating BDT_{ZZ} bins (background-like to signal-like from top to bottom).

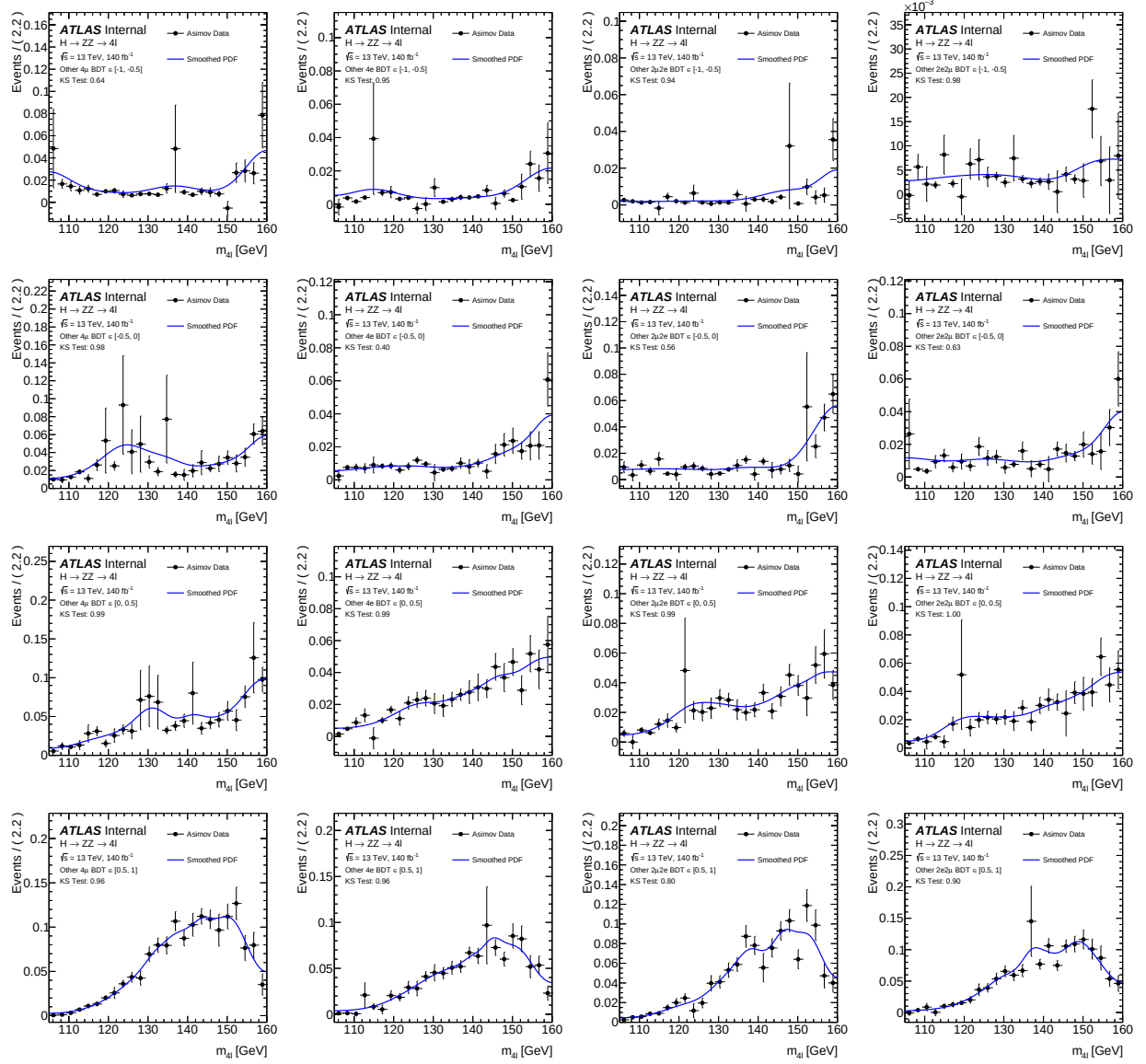


Figure 99: Result of KEYS smoothing of tXX and VVV used as the PDF in the final mass fit. Shown for the full 16 categories, with columns separating decay channels (4μ , $4e$, $2\mu 2e$, $2e 2\mu$), and rows separating BDT_{ZZ} bins (background-like to signal-like from top to bottom).

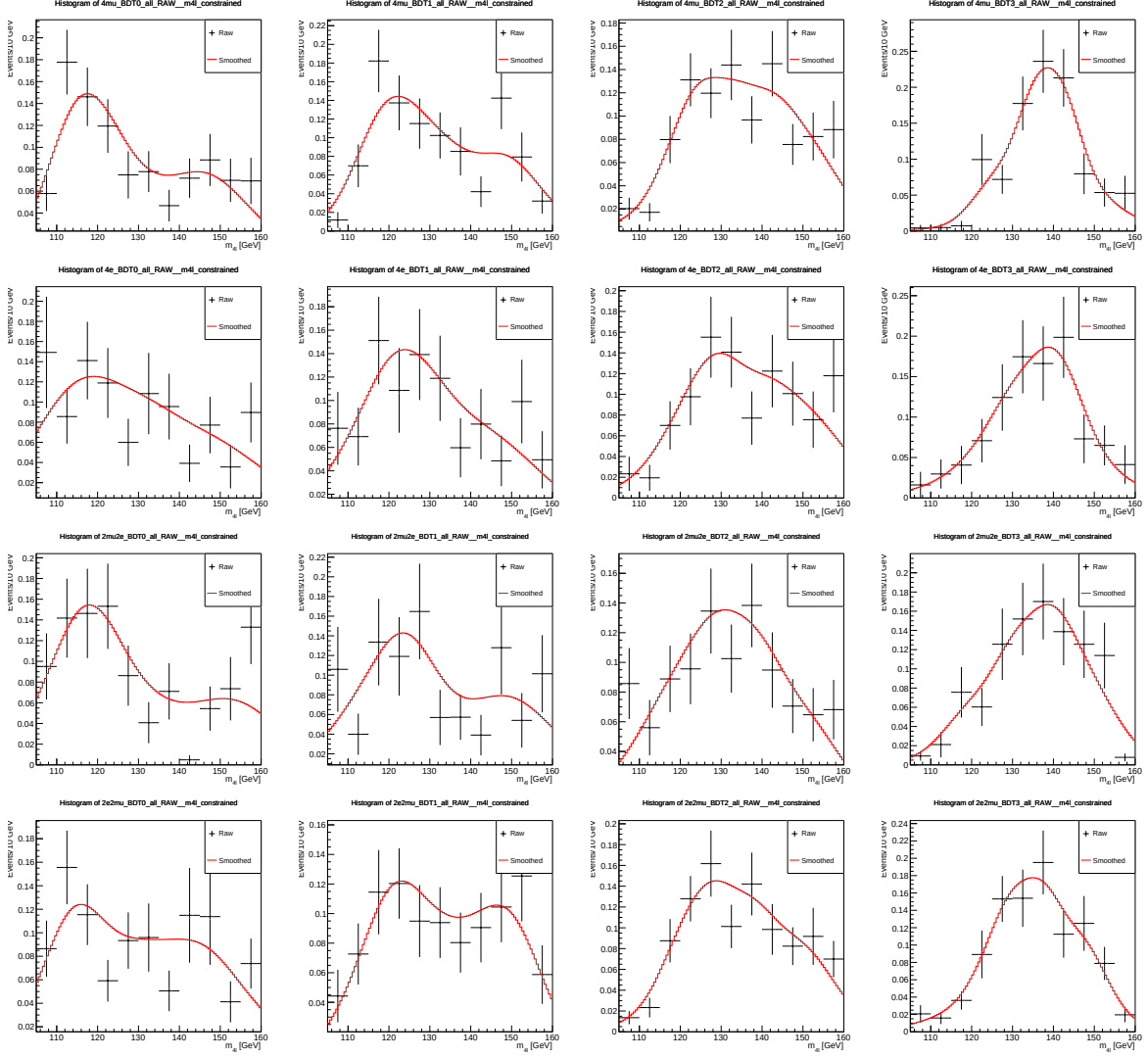


Figure 100: $Z + Jets, t\bar{t}, WZ$ background PDF
Smoothed shapes used as PDF of $Z + Jets, t\bar{t}, WZ$ background. The first row shows 4μ ,
the second $4e$, the third $2\mu 2e$, and the last $2e 2\mu$.

H Template method

The **template method** constitutes an alternative signal model and cross check for this analysis. In this method, the model is a smoothed probability density function (PDF) of observable $m_{4\ell}^{\text{constrained}}$ with one free parameter, m_H . Historically, the template method was the primary signal model for the ATLAS Run 1 measurements [14, 15, 17] and a cross check for the ATLAS mid-Run 2 measurement [18].

H.1 Signal model construction

Template morphing creates a continuous PDF as a function of given unknown parameters (in this case, m_H) from a set of discrete (probability) distributions. Each discrete distribution provides a single point in the parameter space. In this check, each discrete distribution, or “template,” is taken from simulation at different generated m_H masses. Interpolating those shapes to intermediate mass points, or “morphing,” then converts that set of shapes to a continuous function of the unknown parameter(s). In other words, this process takes shapes from a set of $m_H^{\text{generated}}$ samples and interpolates them to produce a continuous function $P(m_{4\ell}^{\text{constrained}}|m_H)$.

H.1.1 Individual signal templates

At each mass point, the generated ggH, VBF, W^-H , W^+H , and ZH contributions are treated individually. For each production mode, separate template distributions of $P(m_{4\ell}^{\text{constrained}})$ are built for $m_H^{\text{generated}} = 124, 125, 126$ GeV for each of the 16 analysis categories. Each template is built using KEYS smoothing [61] with smoothing parameter $\rho = 1$ in all categories for in a single production mode. Figures 101 - 120 show the smoothed templates in each category for each production mode.

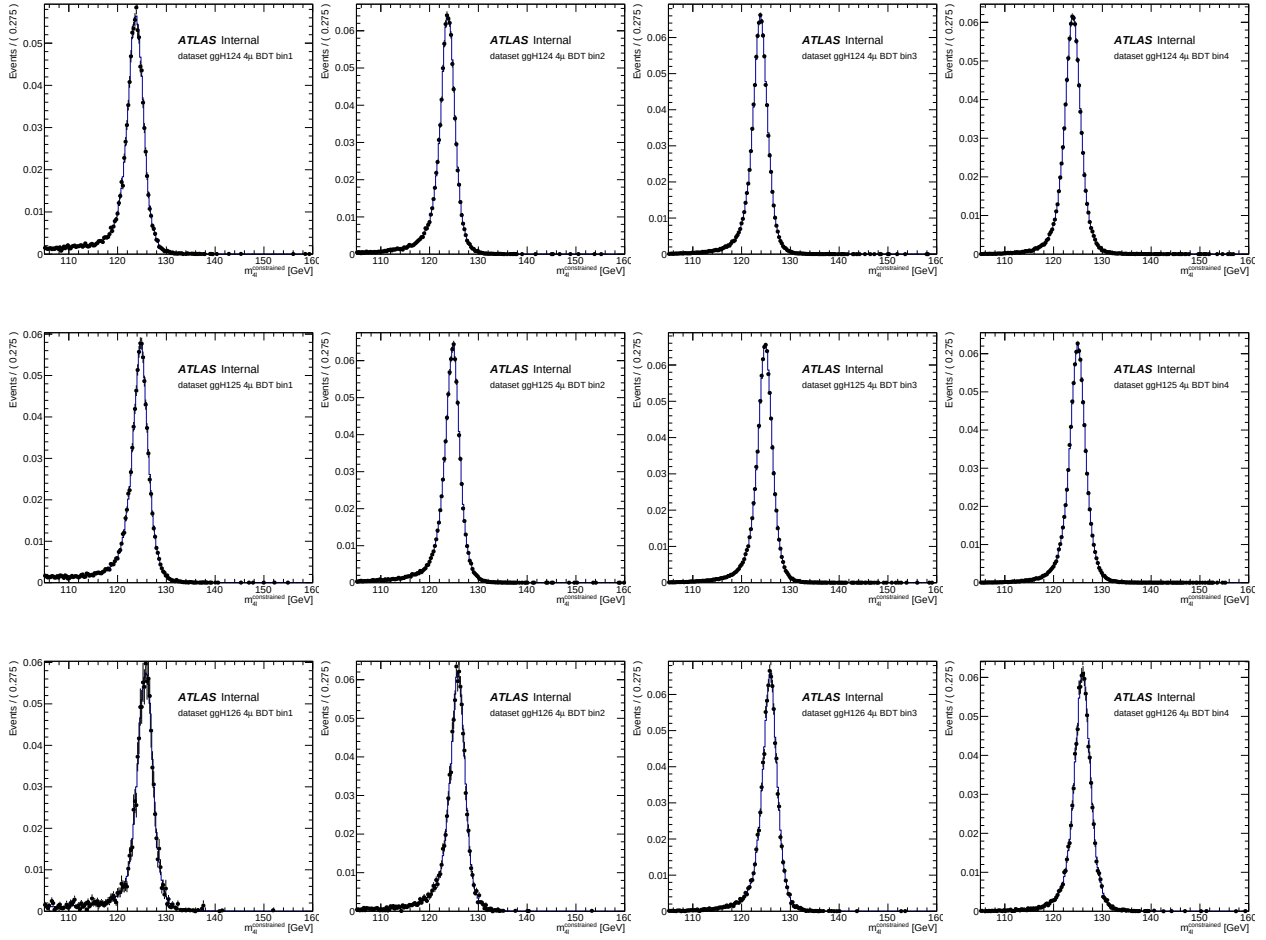


Figure 101: Smoothing result for template signal model in 4μ in the ggH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

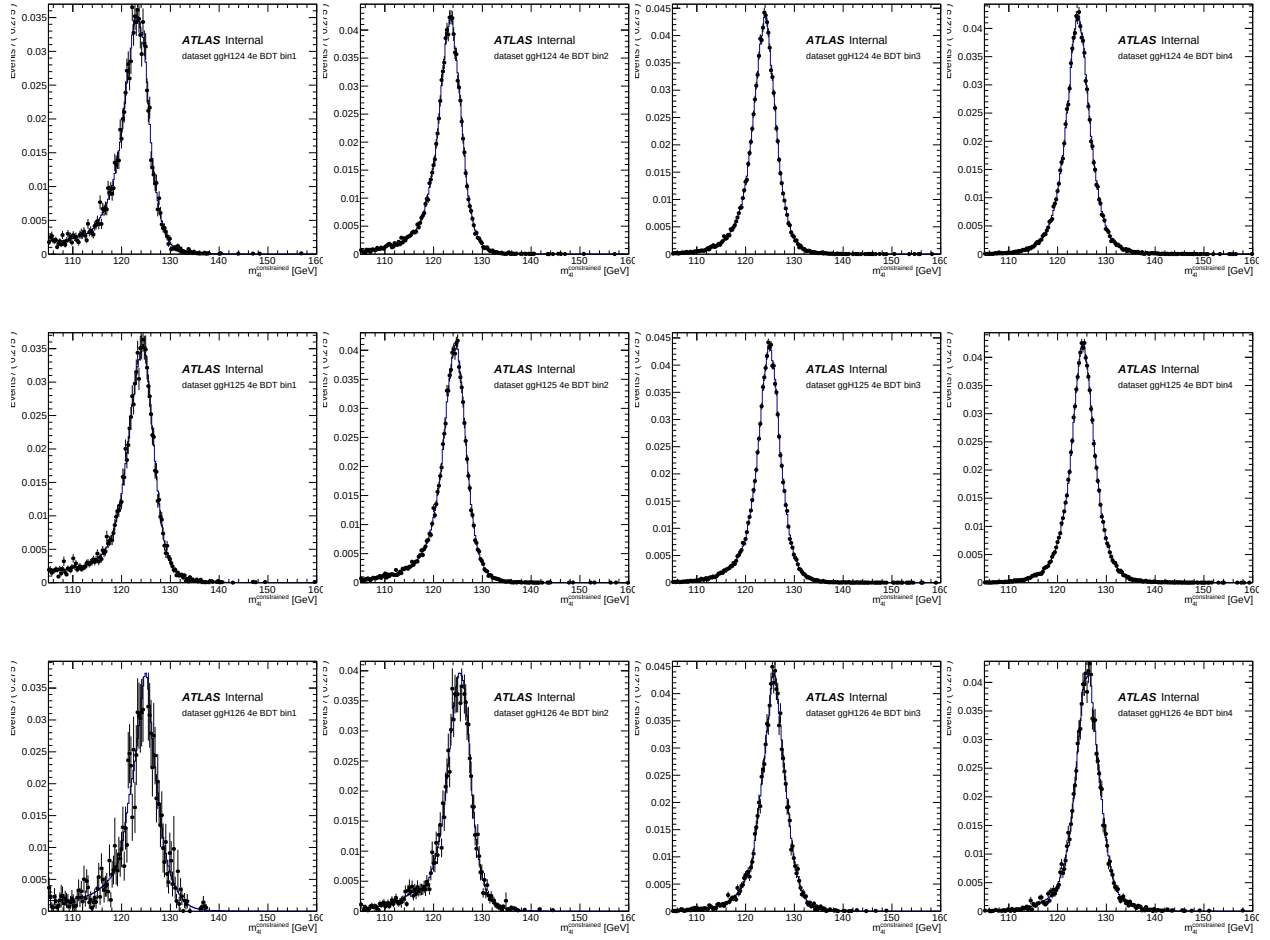


Figure 102: Smoothing result for template signal model in 4e in the ggH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

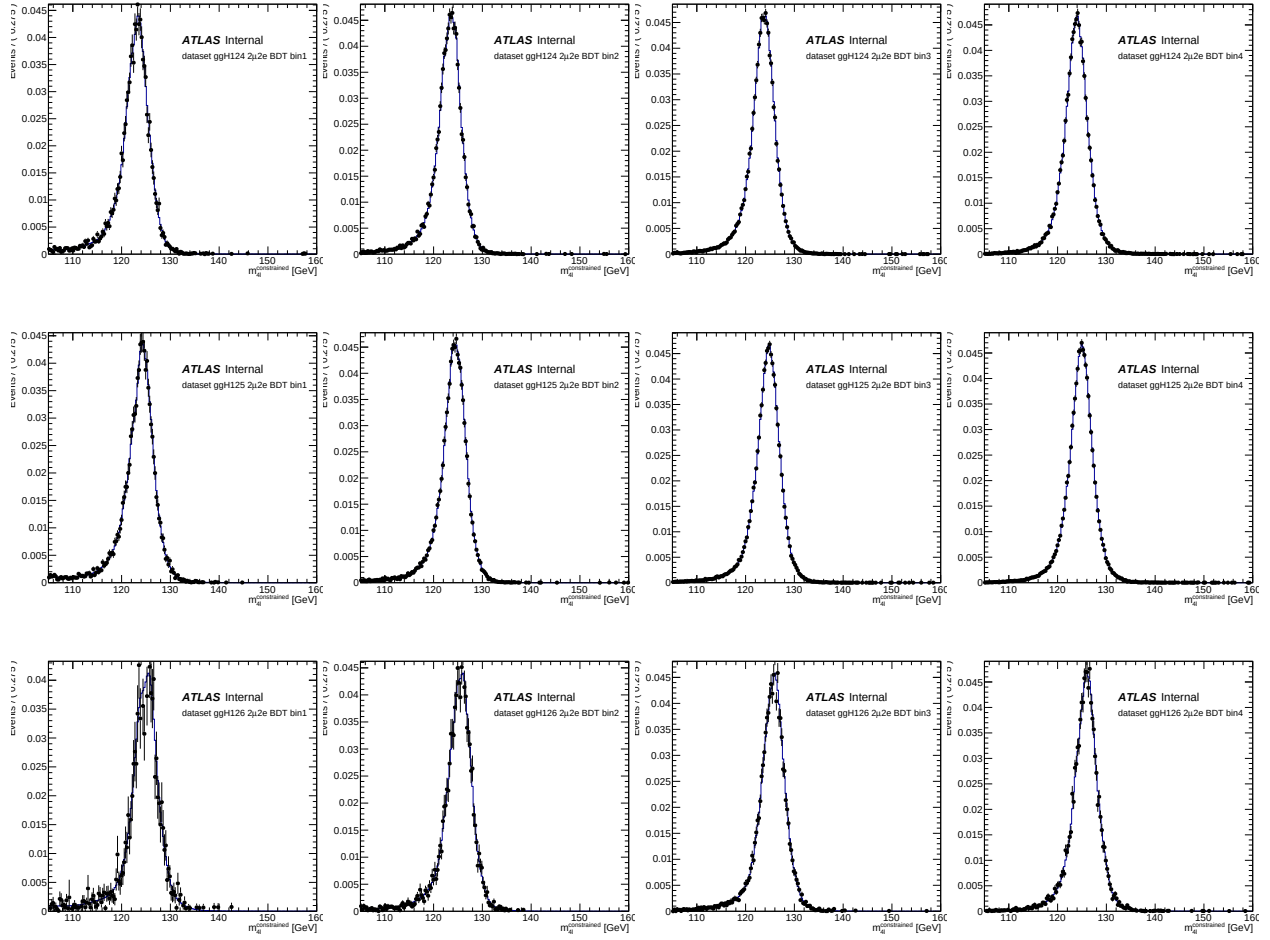


Figure 103: Smoothing result for template signal model in $2\mu 2e$ in the ggH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

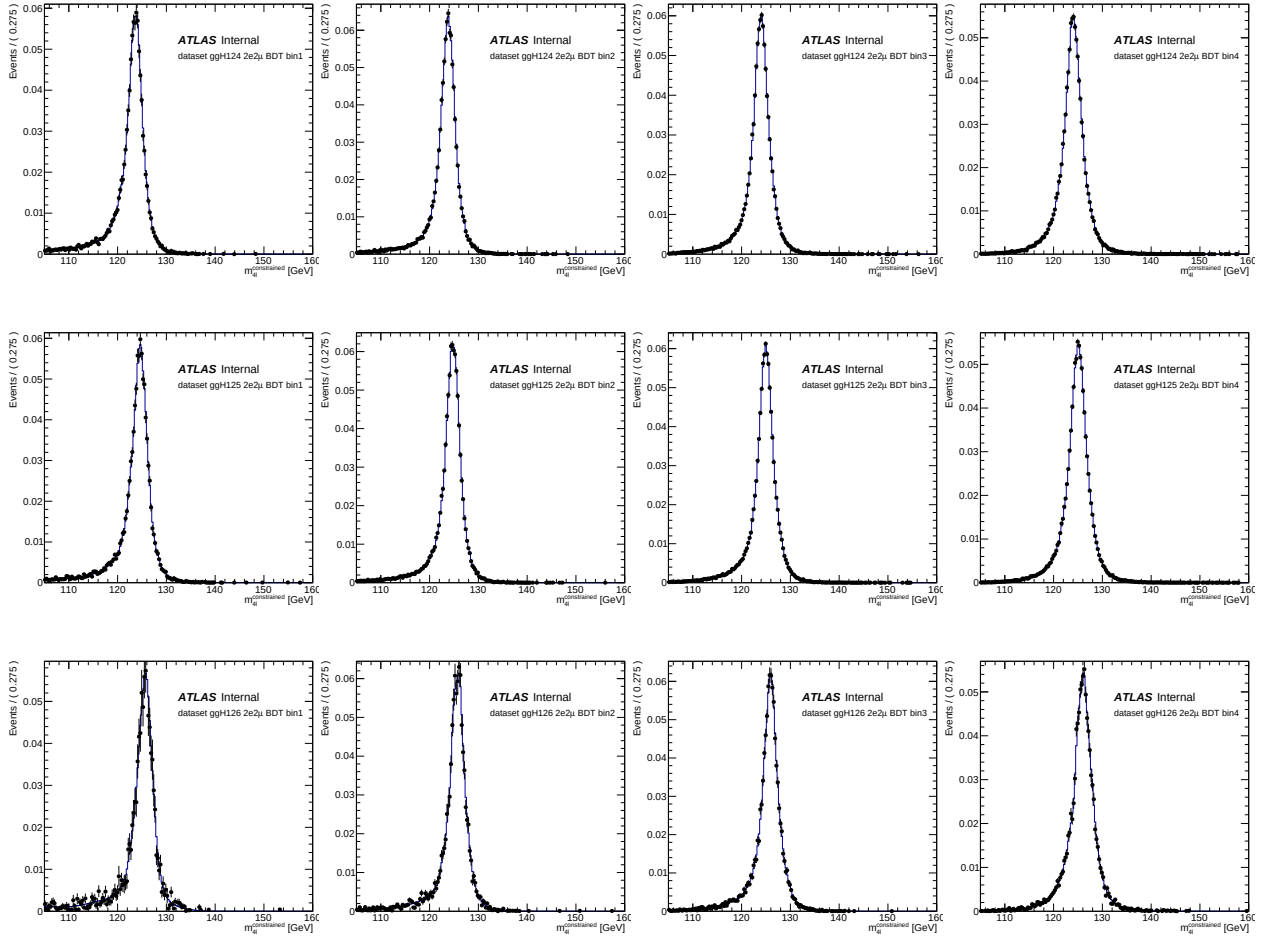


Figure 104: Smoothing result for template signal model in $2e2\mu$ in the ggH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

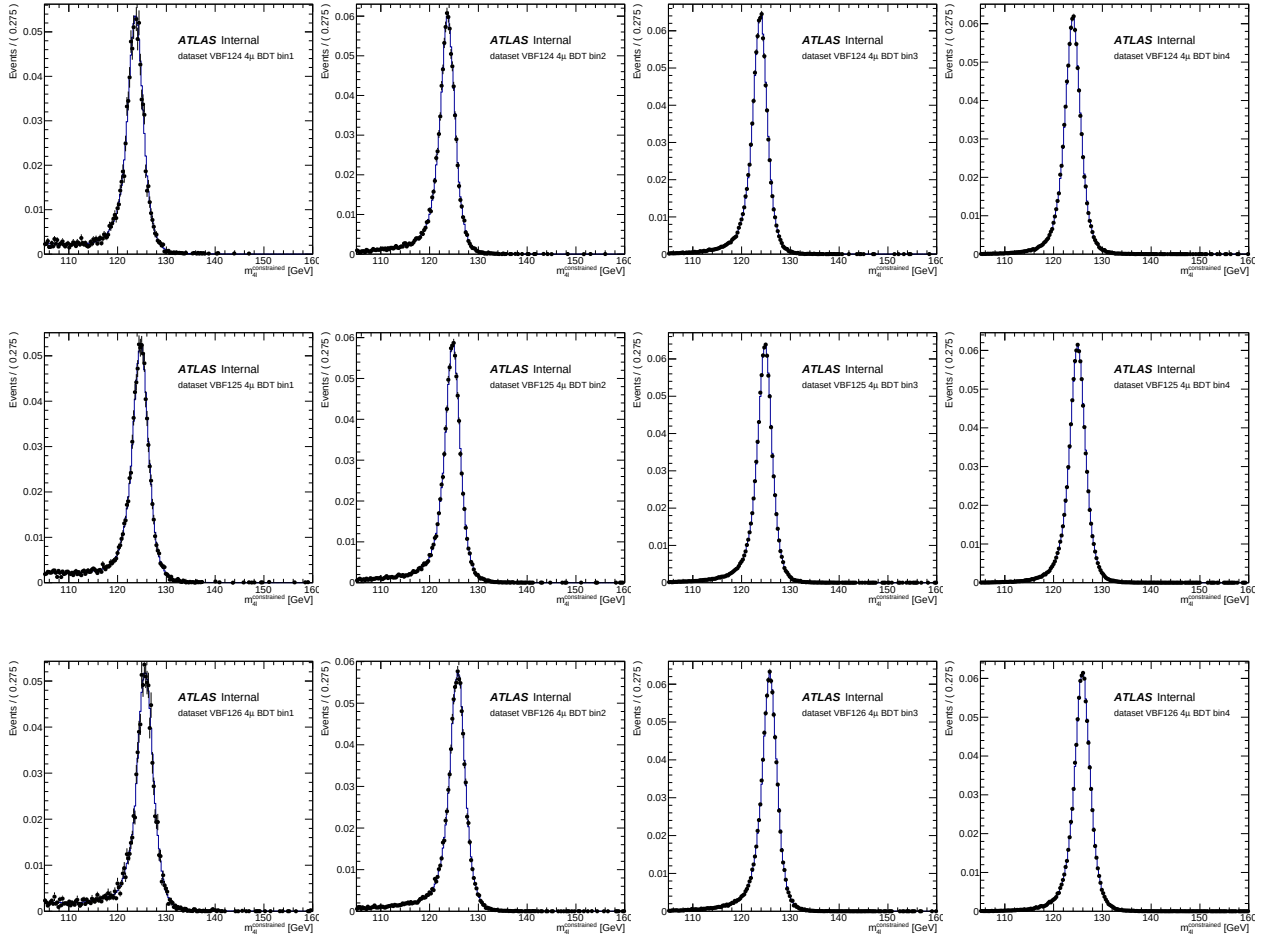


Figure 105: Smoothing result for template signal model in 4μ in the VBF production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

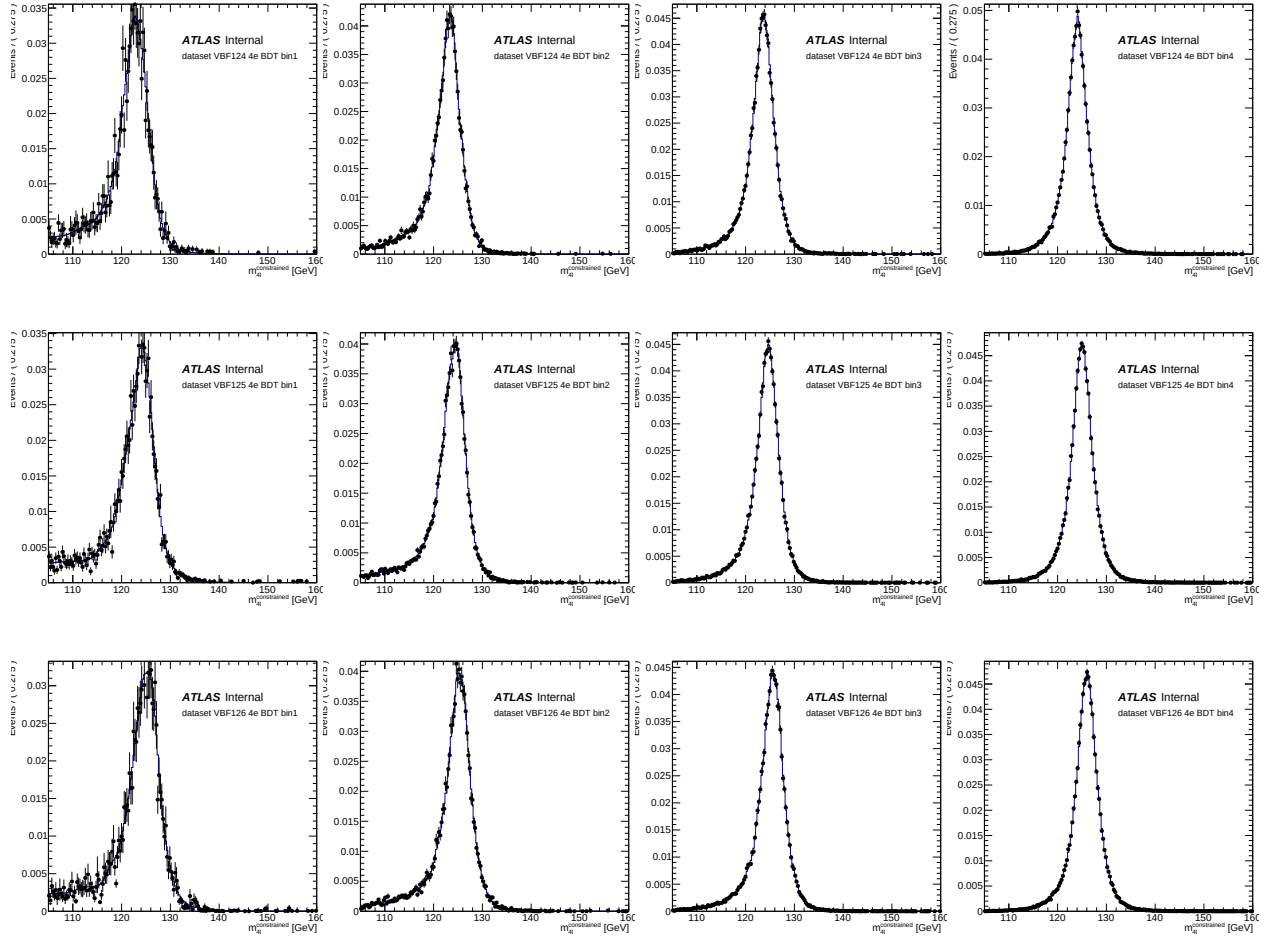


Figure 106: Smoothing result for template signal model in 4e in the VBF production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

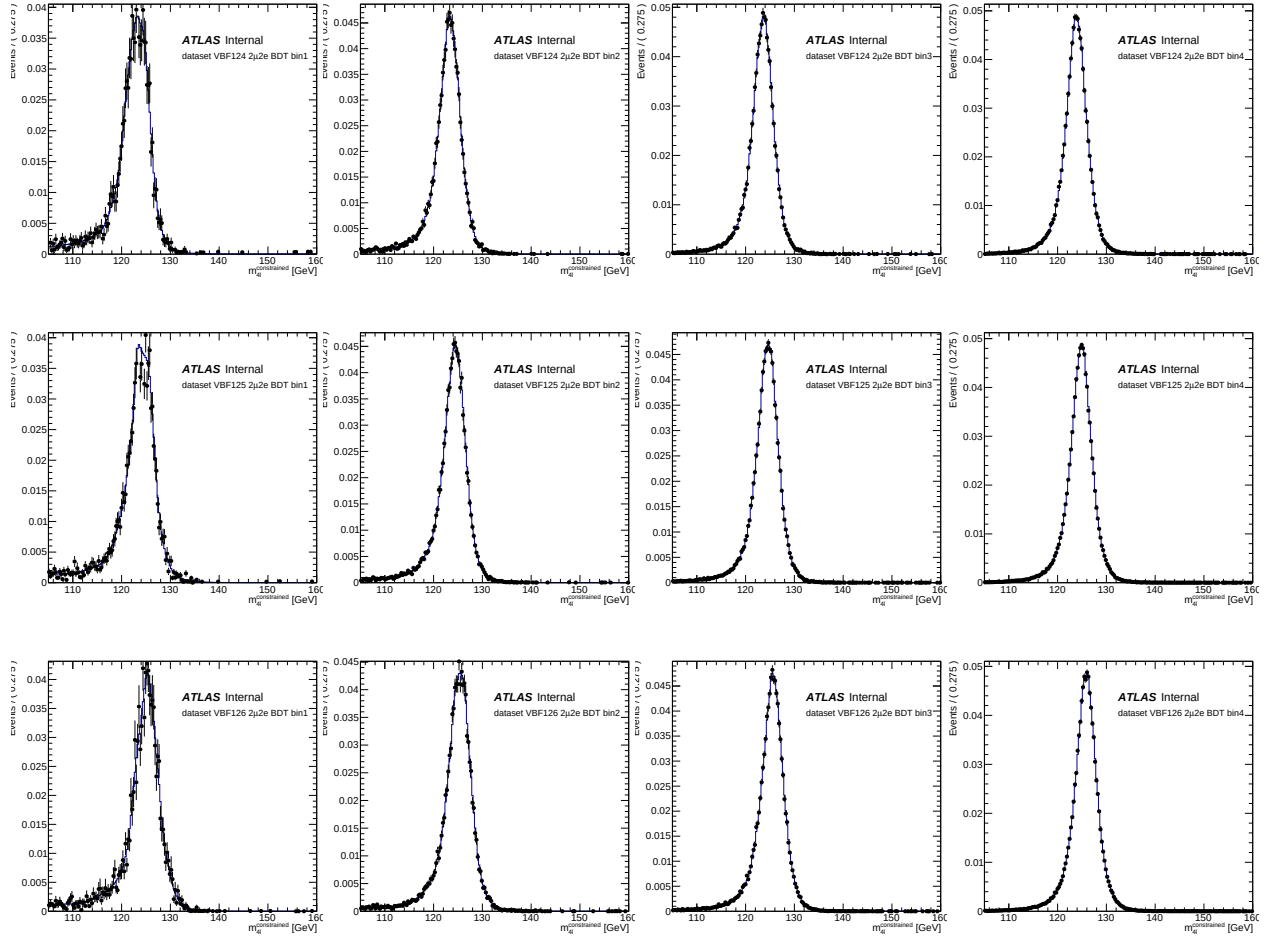


Figure 107: Smoothing result for template signal model in $2\mu 2e$ in the VBF production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

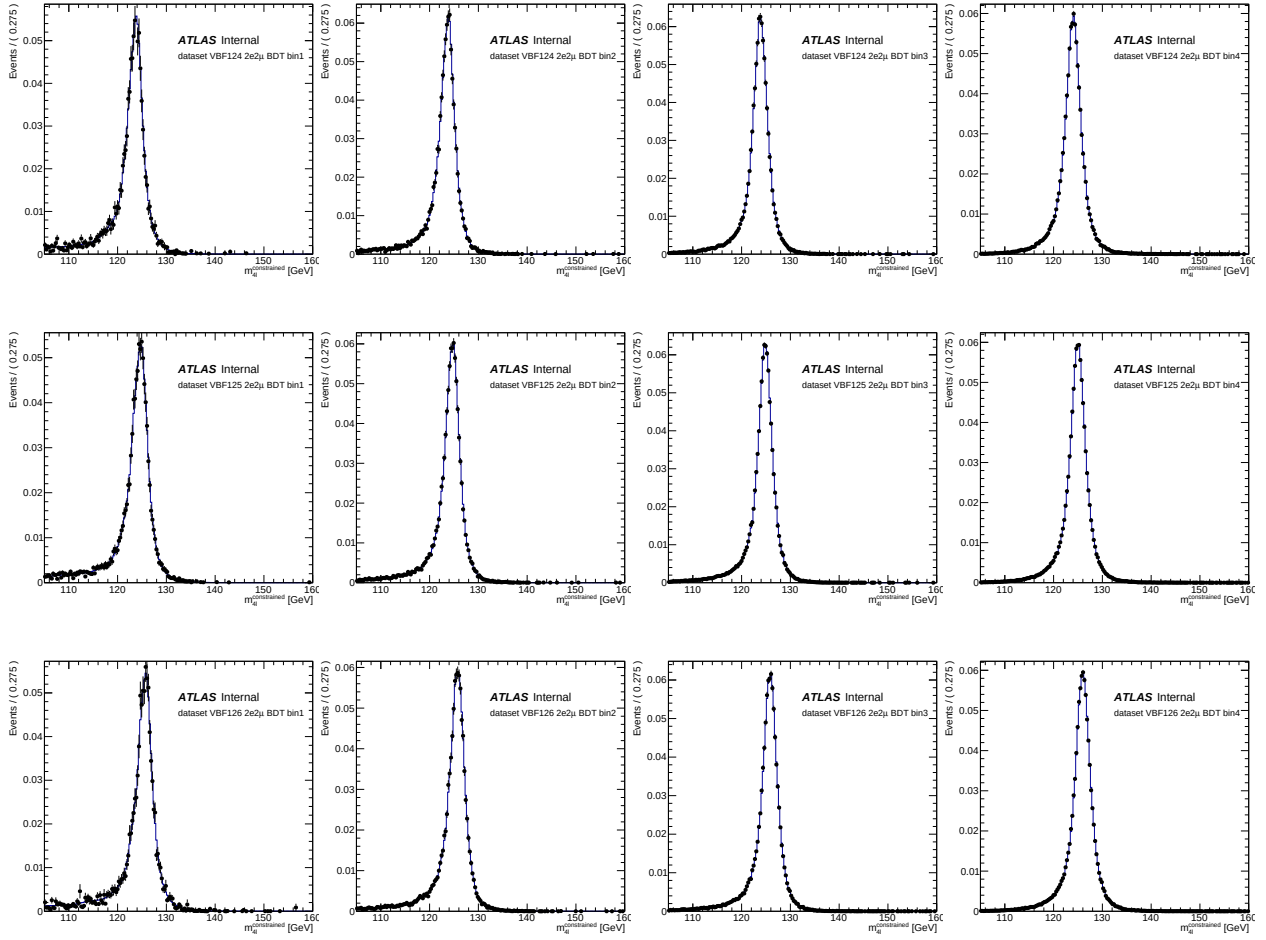


Figure 108: Smoothing result for template signal model in $2e2\mu$ in the VBF production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

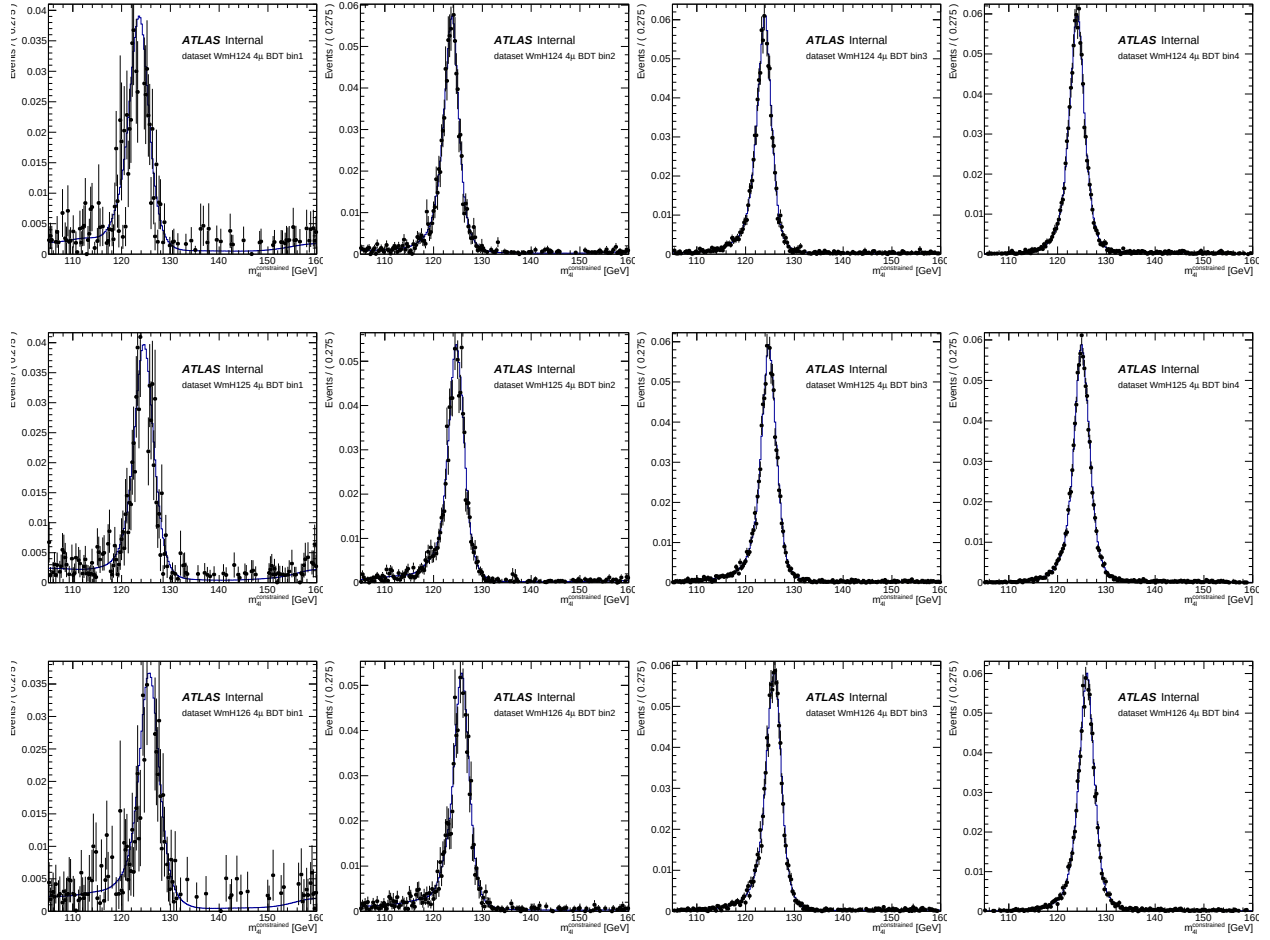


Figure 109: Smoothing result for template signal model in 4μ in the W^-H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

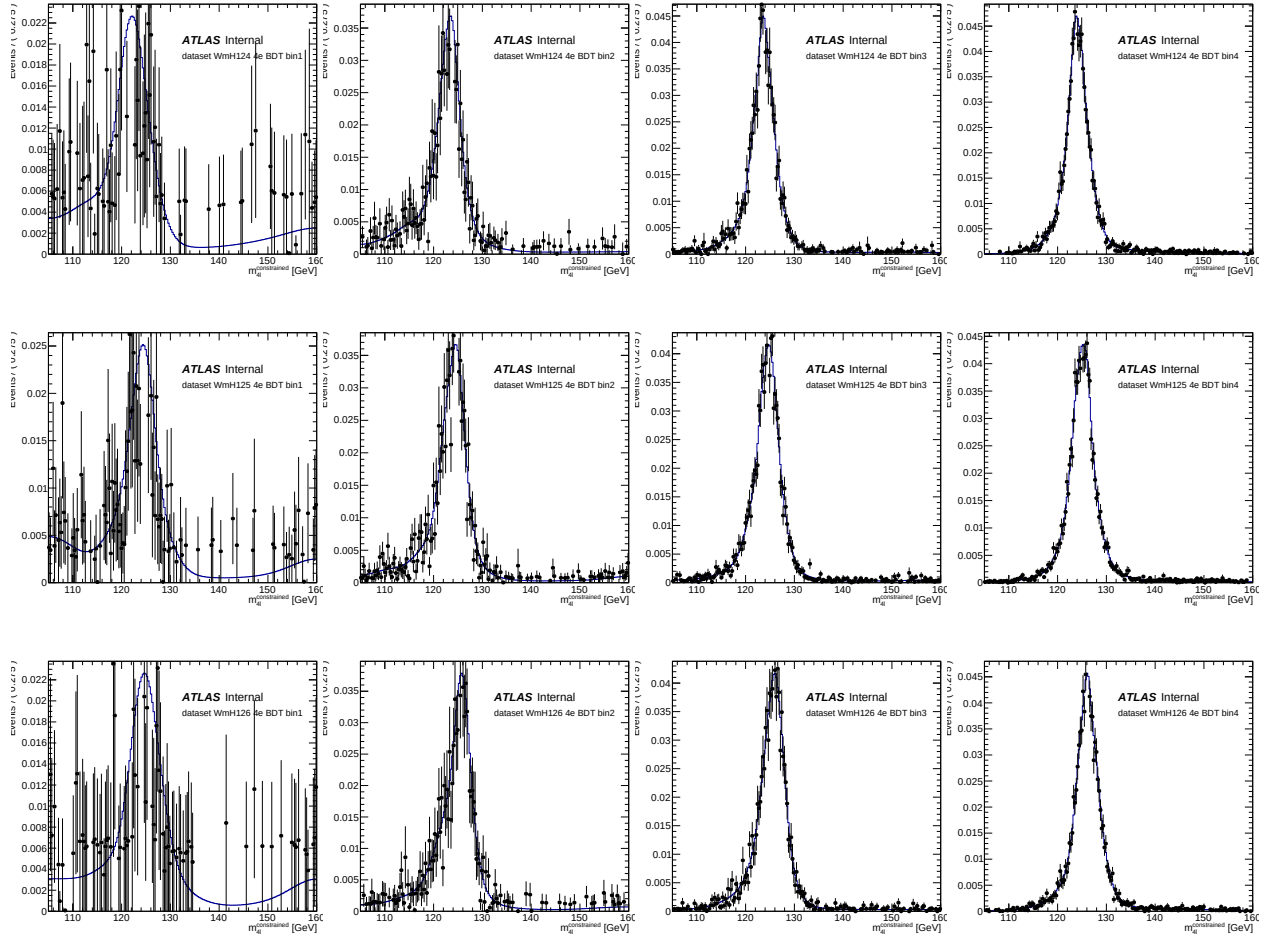


Figure 110: Smoothing result for template signal model in 4e in the W-H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

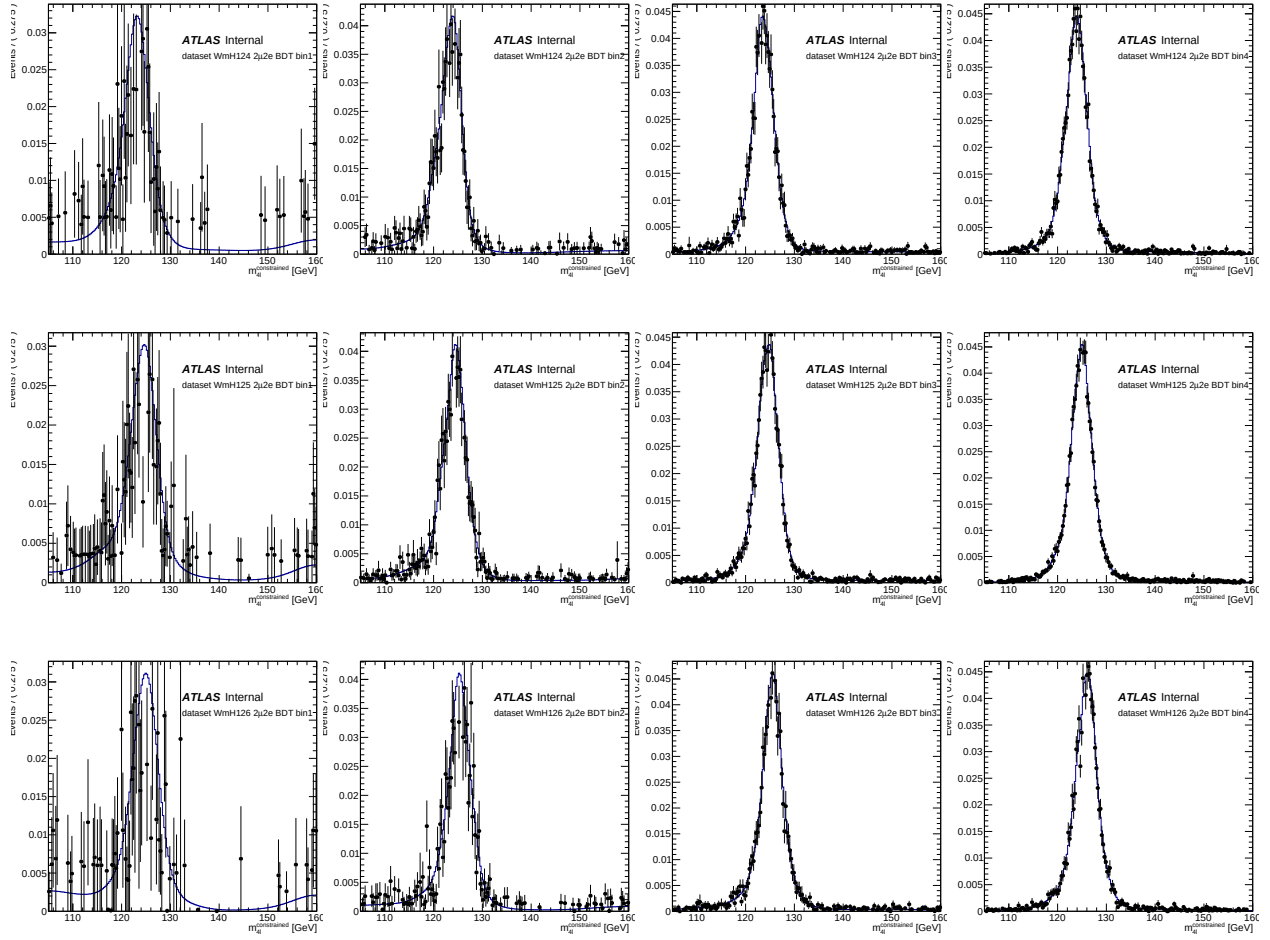


Figure 111: Smoothing result for template signal model in $2\mu 2e$ in the W - H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

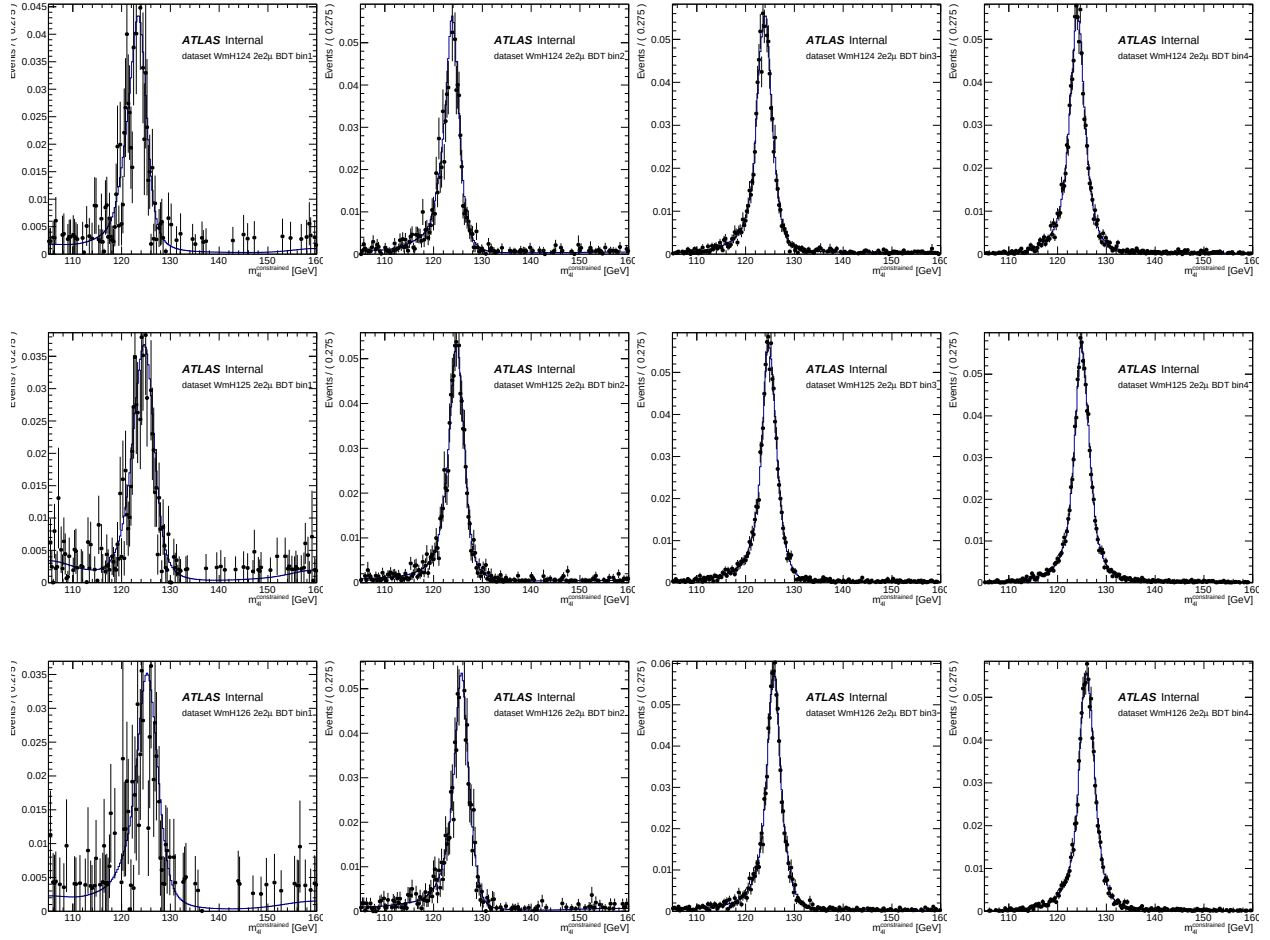


Figure 112: Smoothing result for template signal model in $2e2\mu$ in the W - H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

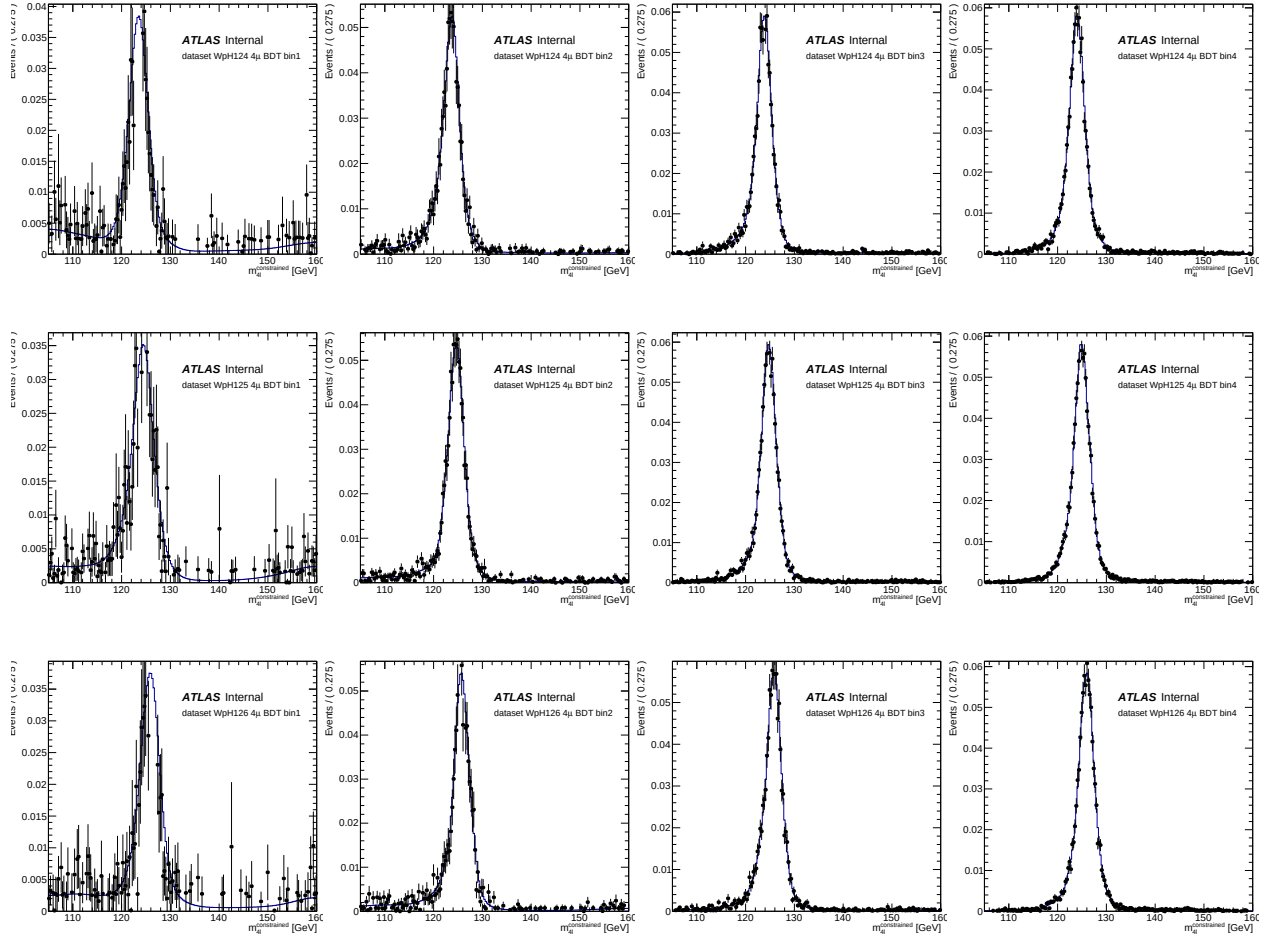


Figure 113: Smoothing result for template signal model in 4μ in the W^+H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

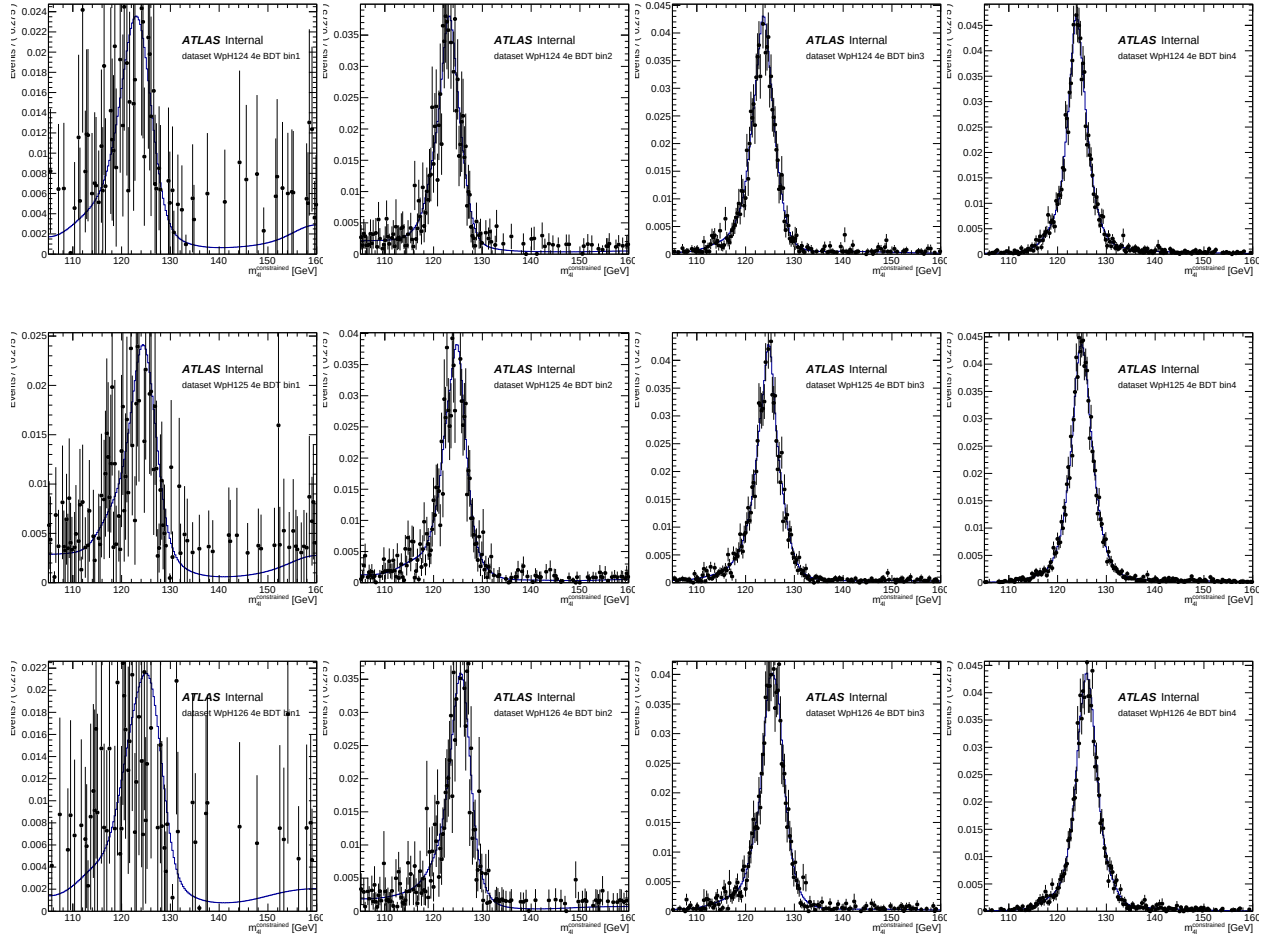


Figure 114: Smoothing result for template signal model in 4e in the W^+H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

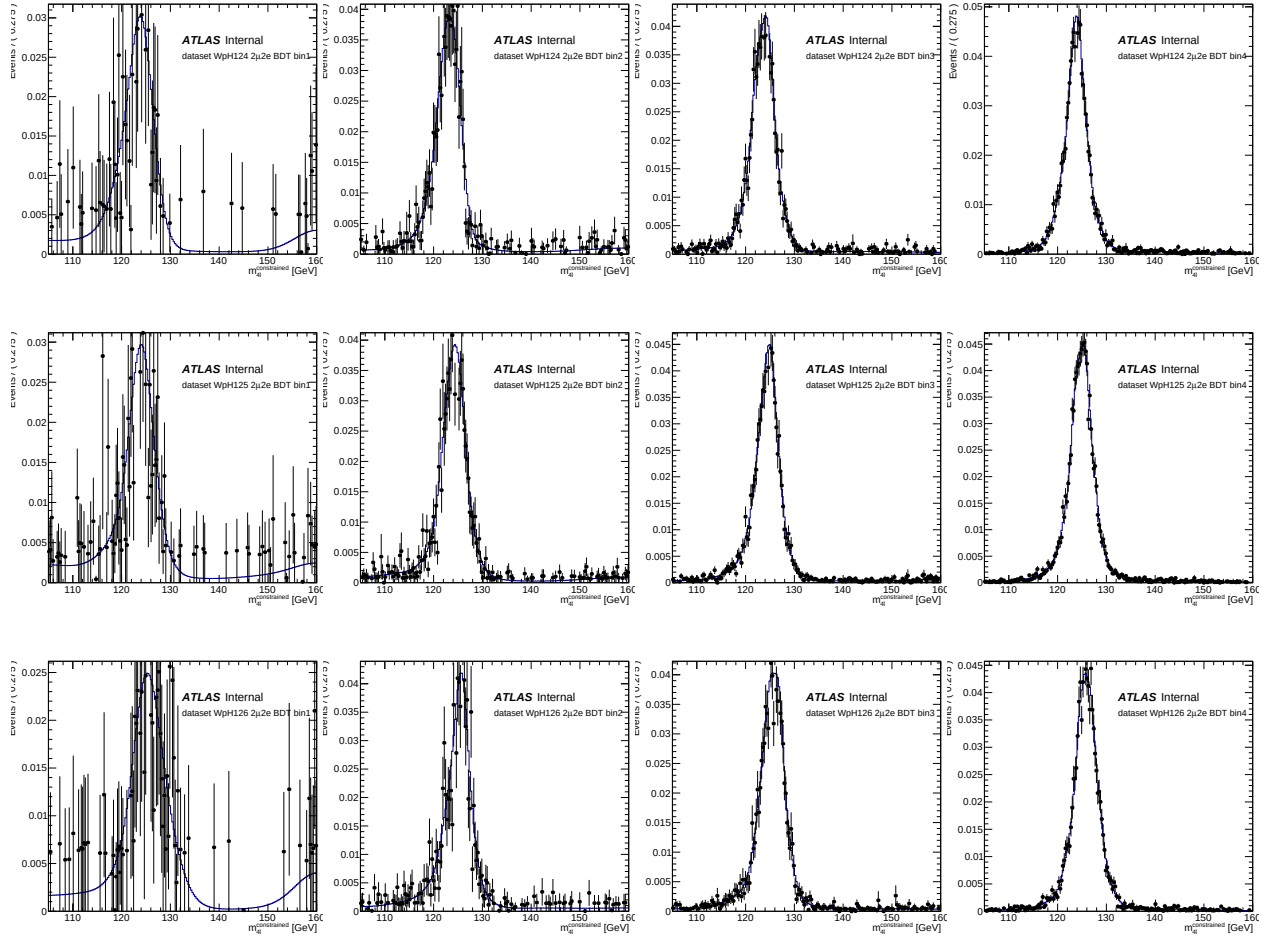


Figure 115: Smoothing result for template signal model in $2\mu 2e$ in the W^+H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

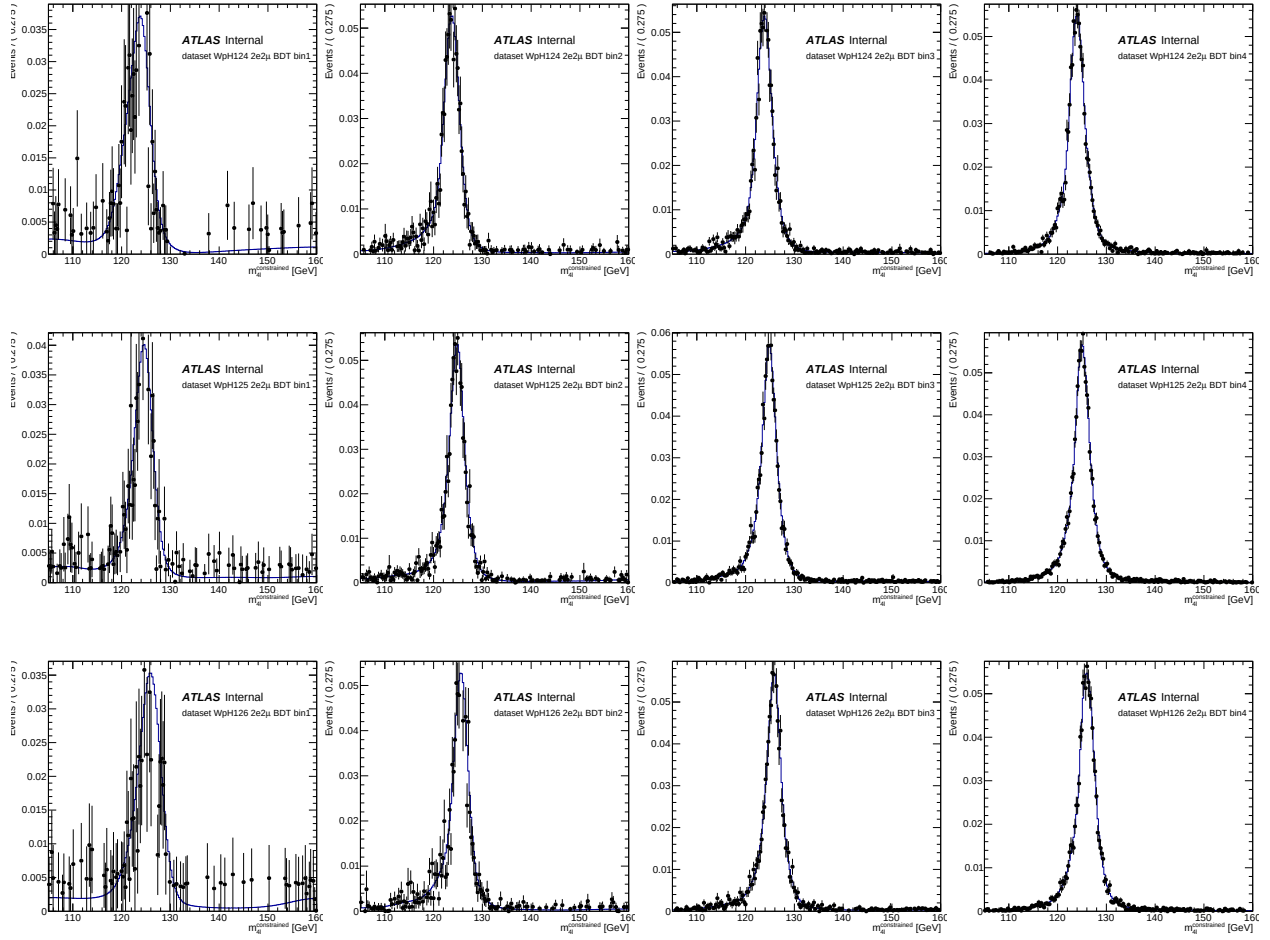


Figure 116: Smoothing result for template signal model in $2e2\mu$ in the W^+H production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

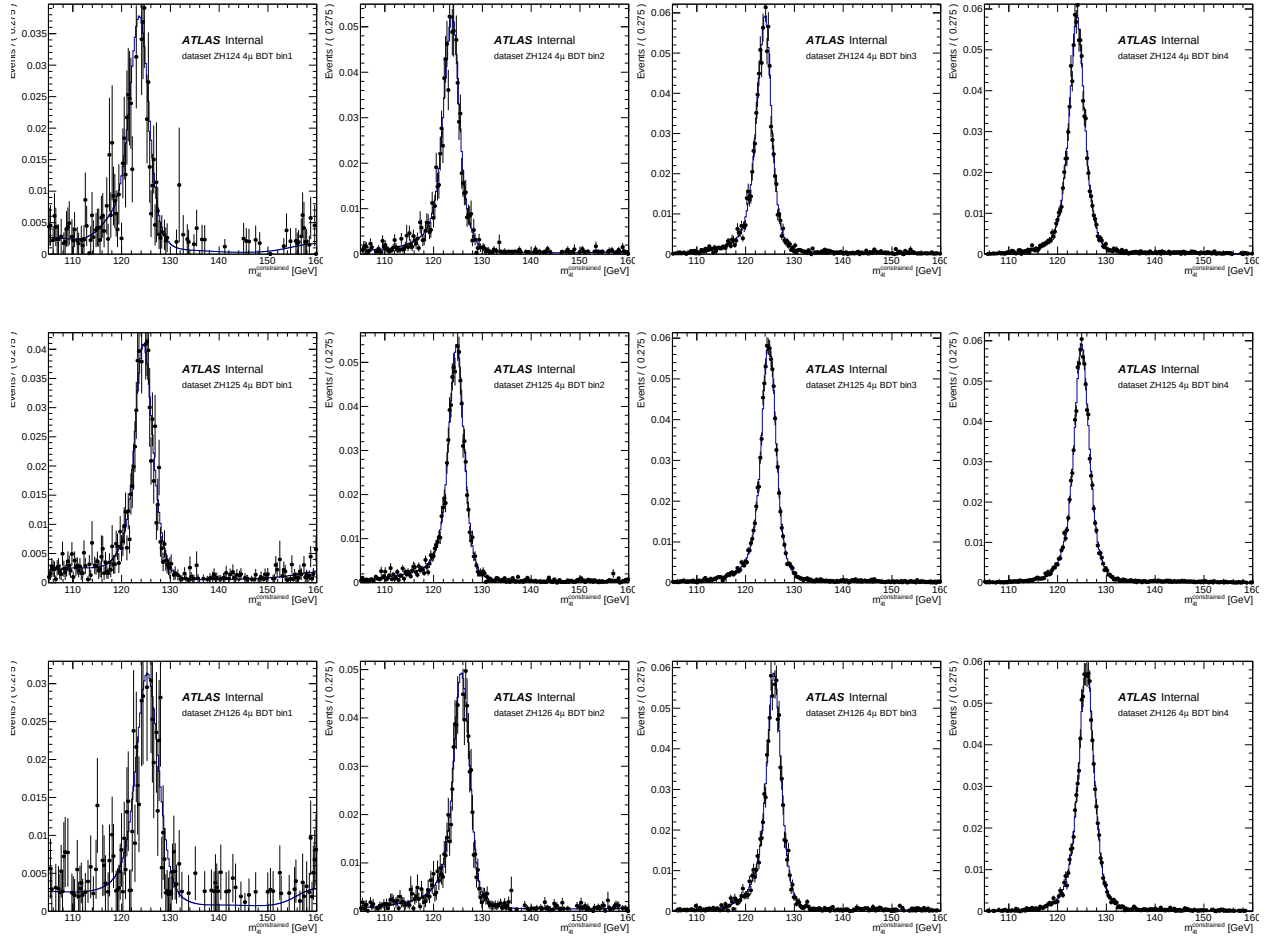


Figure 117: Smoothing result for template signal model in 4μ in the ZH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

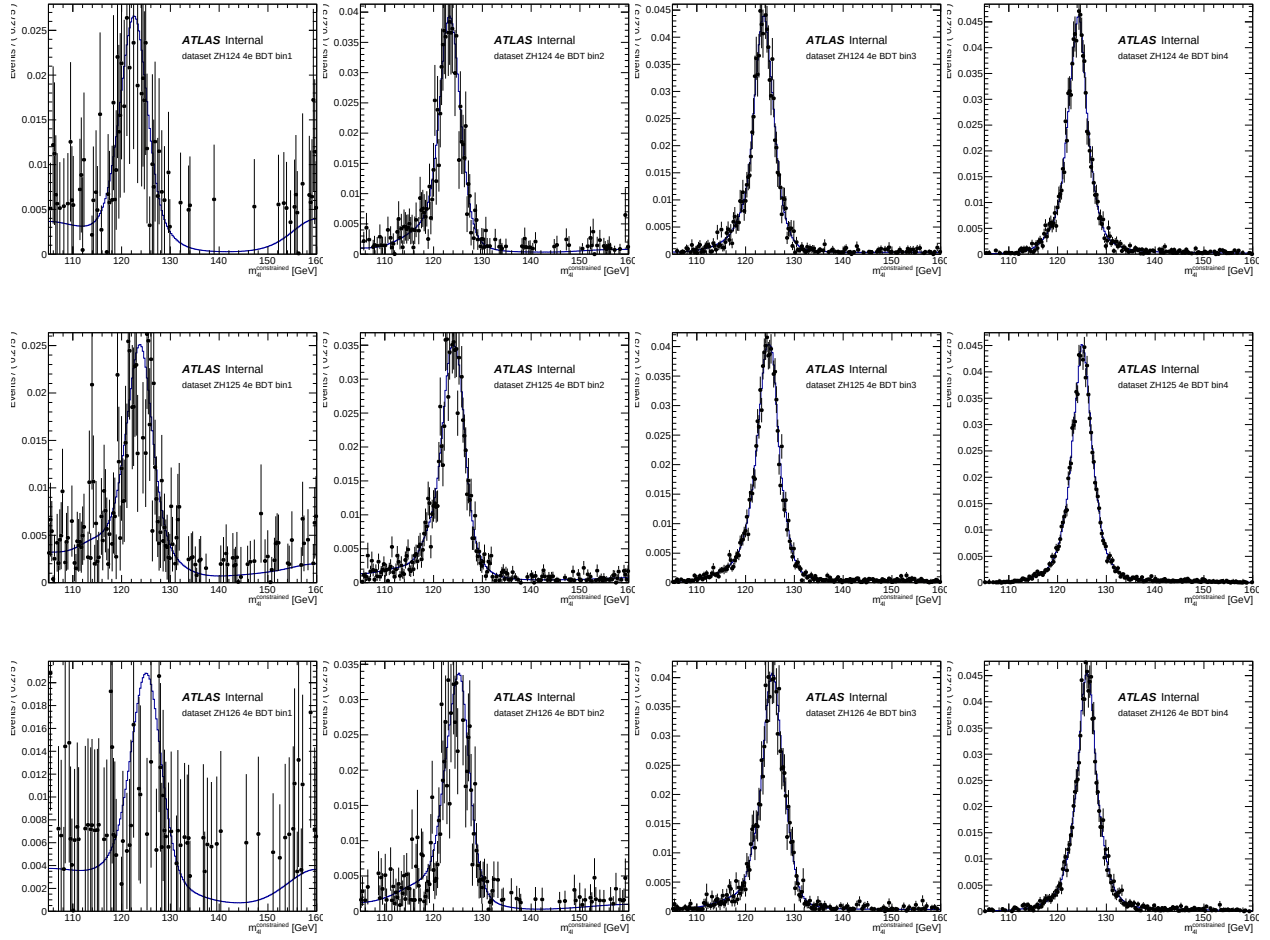


Figure 118: Smoothing result for template signal model in 4e in the ZH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

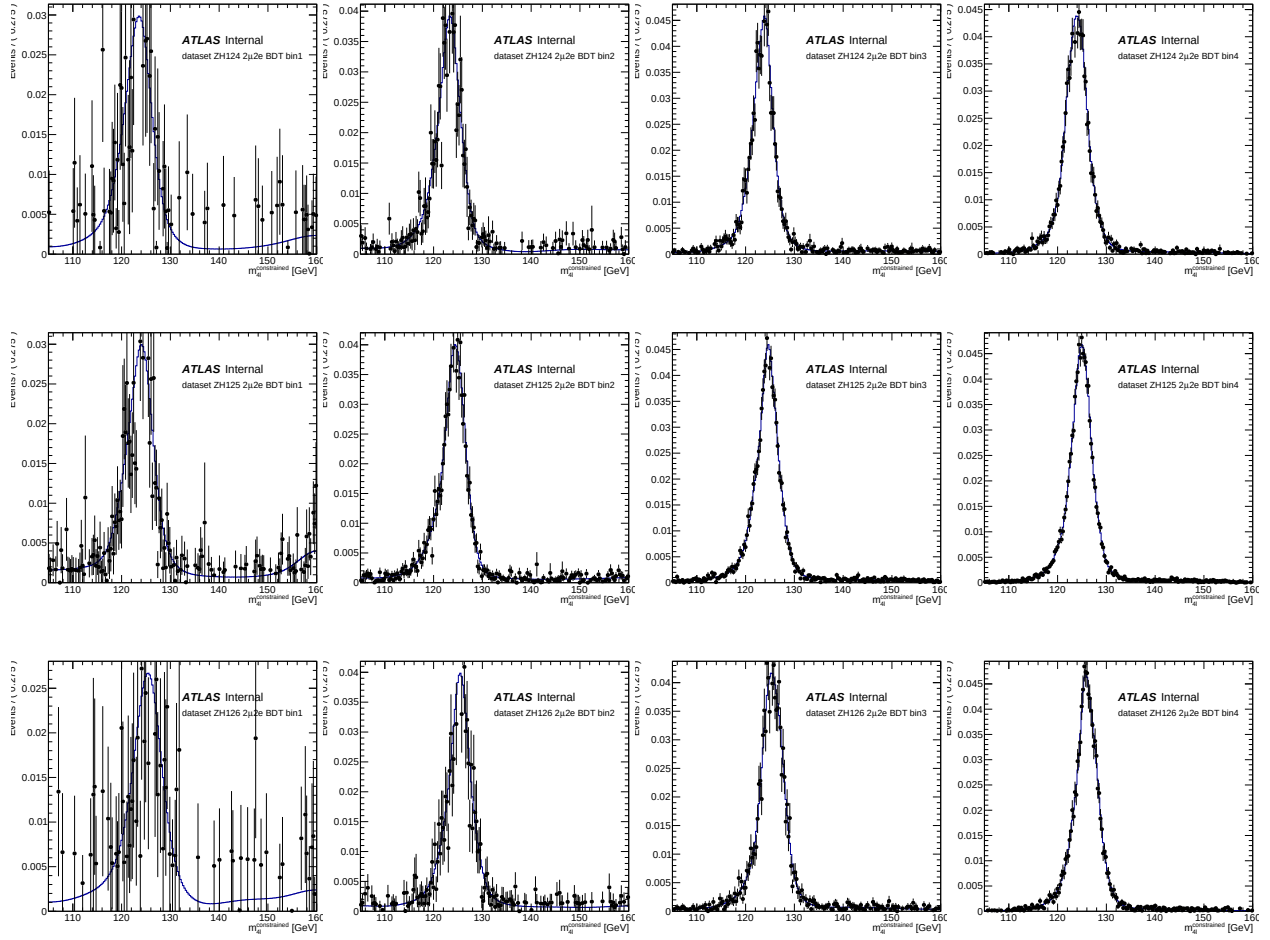


Figure 119: Smoothing result for template signal model in $2\mu 2e$ in the ZH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

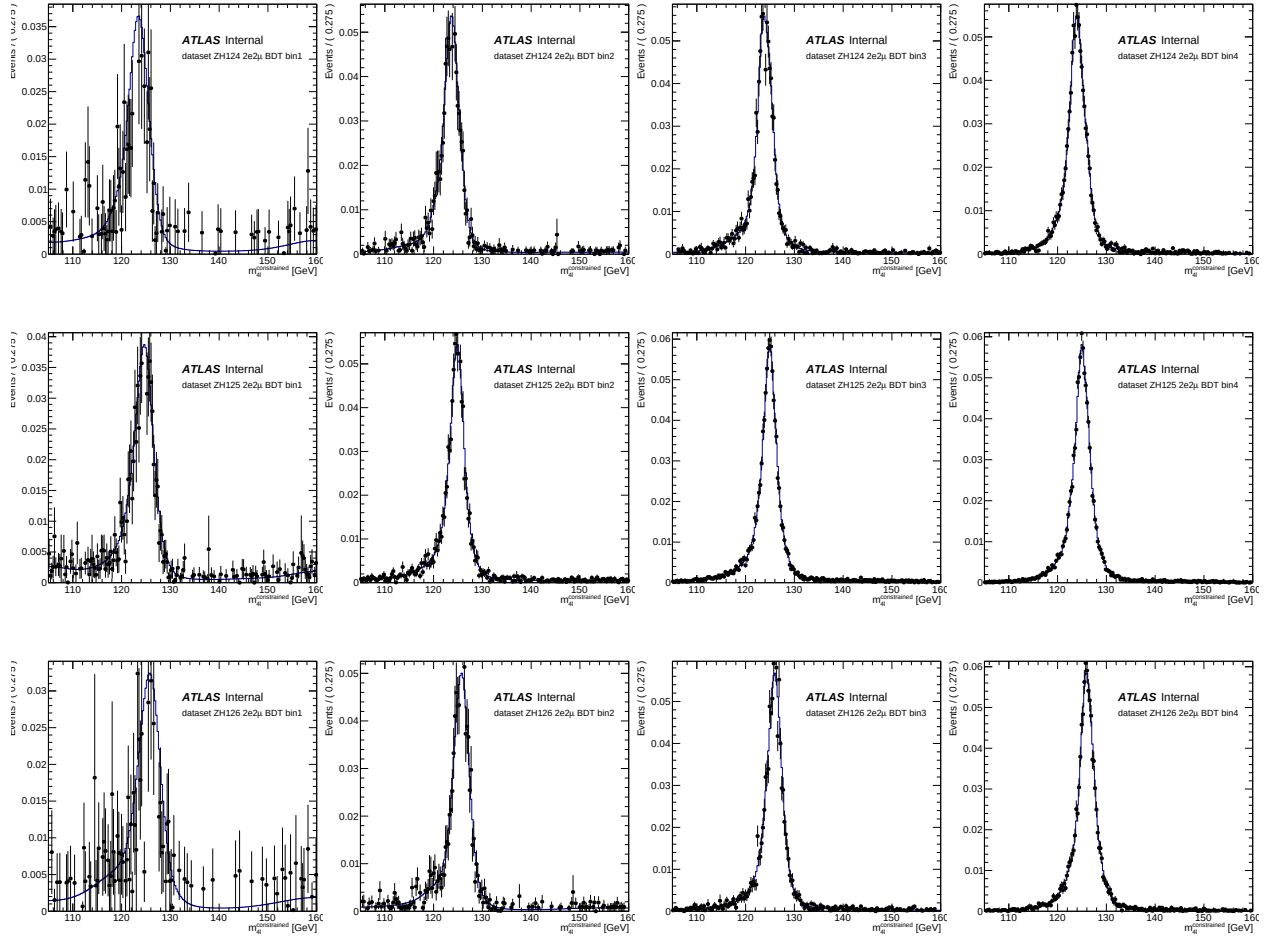


Figure 120: Smoothing result for template signal model in $2e2\mu$ in the ZH production mode. The $m_{4\ell}^{\text{constrained}}$ distribution is shown for MC-simulated mass points 124, 125, and 126 in each row. The columns correspond to different classifier bins becoming more signal-like from left to right.

H.1.2 Template morphing for the signal PDF

Once created, the templates must be morphed into a single, continuous parameterization of $m_{4\ell}^{\text{constrained}}$ with free parameter m_H in each production mode. The templates are morphed using RooBSpline [PieglerAndTiller]. The morphing occurs separately along each production mode to account for shape variations from the different kinematics of each mode. Producing a single monolithic PDF jointly across all production modes worsens the resolution on m_H^{measured} by 6 MeV. The morphed templates are then summed together according to their expected yields in $105 < m_{4\ell}^{\text{constrained}} < 160$ GeV given their Standard Model cross sections. The resulting signal PDF is drawn at several m_H values and shown in Figure 121.

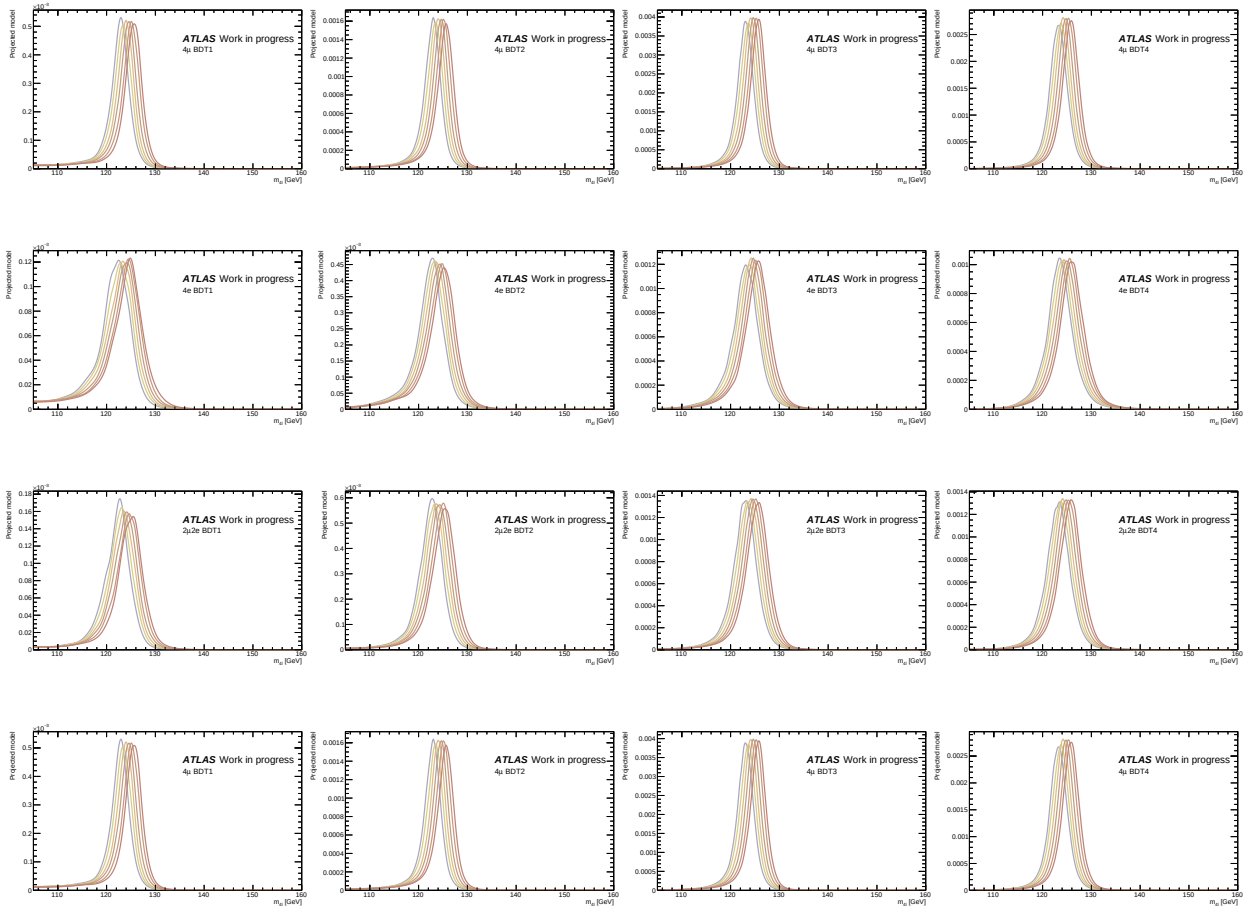


Figure 121: Demonstration of template morphing using RooBSpline in each analysis category. The signal PDF $P(m_{4\ell}^{\text{constrained}} | m_H)$ is drawn in 600 MeV steps from 123 to 126 GeV

H.2 Full template model

The signal model is then summed with the same smoothed PDFs for the $qqZZ$, $ggZZ$, $EWZZ$, and tXX/VVV ³⁸ backgrounds described in Section 7.5. The overall normalizations for signal and $qqZZ + ggZZ$ in each BDT_{ZZ} bin are allowed to float as in the main analysis. The other backgrounds are fixed to their estimated yields from simulation. The reducible background was omitted from this test. The full PDF in each category for several values of m_H is shown in Figure 122.

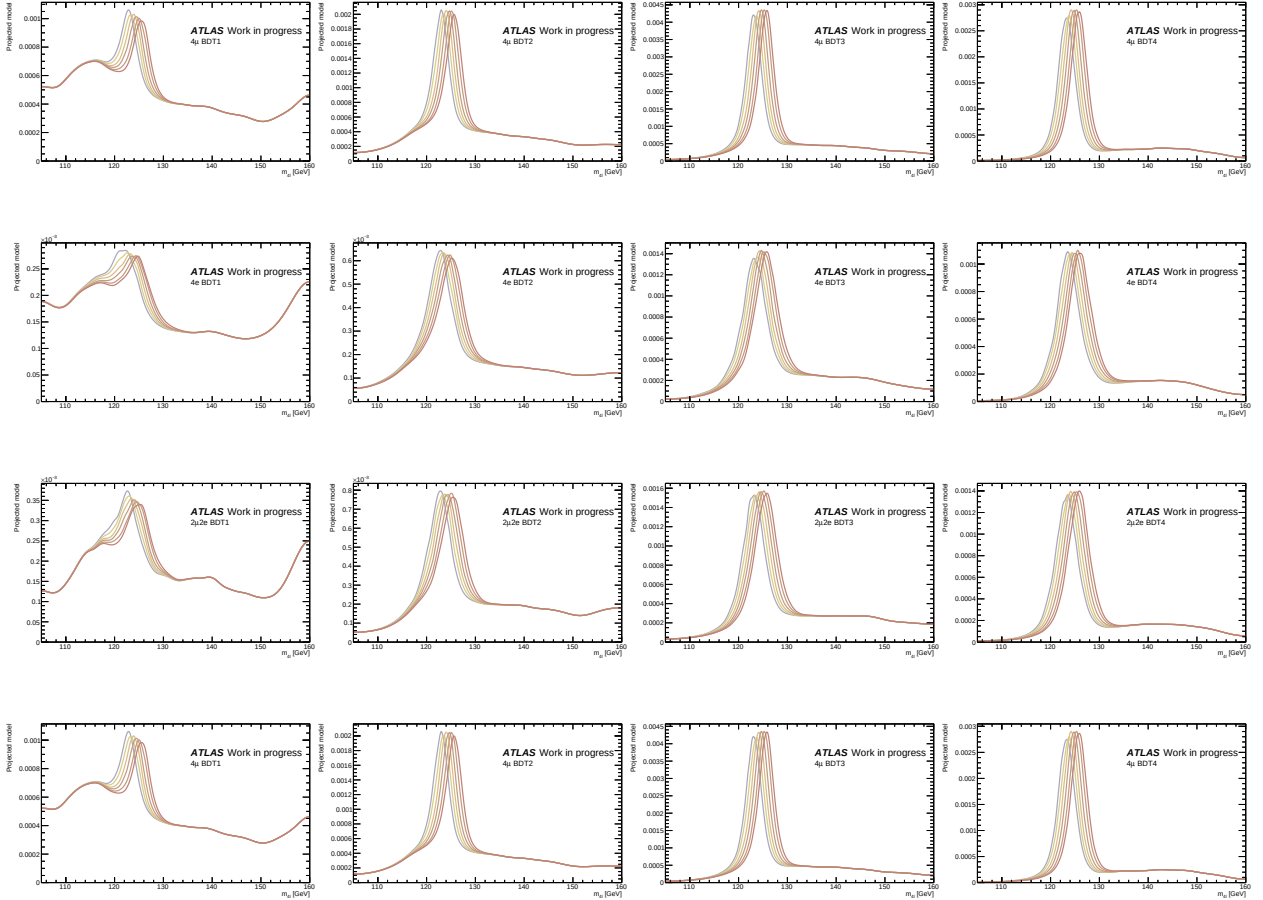


Figure 122: Demonstration of template morphing using RooBSpline in each analysis category. The signal and background PDF $P(m_{4\ell}^{\text{constrained}}|m_H)$ is drawn in 600 MeV steps from 123 to 126 GeV

³⁸ tXX : tZ , ttW , ttt , $tttt$, $ttWZ$, $ttZZ$, $ttWW$, ttZ

Table 35: Template 2D cross check results for fitting an Asimov dataset generated with $m_H = 125$, $\mu_{signal}^{BDT*} = 1$, and $\mu_{ZZ*}^{BDT*} = 1$ from the model

Asimov results	Template 2D
m_H^{measured}	$125 \pm 0.185 \text{ GeV}$
μ_{signal}^{BDT1}	1.00 ± 0.539
μ_{signal}^{BDT2}	1.00 ± 0.225
μ_{signal}^{BDT3}	1.00 ± 0.124
μ_{signal}^{BDT4}	1.00 ± 0.125
μ_{ZZ*}^{BDT1}	1.00 ± 0.107
μ_{ZZ*}^{BDT2}	1.00 ± 0.131
μ_{ZZ*}^{BDT3}	1.00 ± 0.127
μ_{ZZ*}^{BDT4}	1.00 ± 0.199

The full template model constitutes a simultaneous PDF over all analysis categories where the PDF in each category is given from the above procedure. For this test, an Asimov dataset was generated at $m_H = 125 \text{ GeV}$ and with all normalizations = 1 from the simultaneous PDF. This dataset was then fit using an unbinned maximum likelihood fit to the simultaneous PDF to estimate the expected uncertainty on m_H^{measured} with the Template method. m_H , the signal normalization in each BDT_{ZZ} bin, and the non-resonant $ZZ^{(*)}$ normalization in each BDT_{ZZ} bin were all free during the fit. The model shows no bias and an expected uncertainty on m_H^{measured} of 185 MeV. The results are summarized in Table 35.