

**SEARCH FOR EXOTIC HIGGS BOSON DECAY TO
MULTIPLE B-QUARKS WITH THE ATLAS DETECTOR
AT LHC USING MACHINE LEARNING METHODS**

A Dissertation Presented

by

YUAN-TANG CHOU

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2023

Physics

© Copyright by Yuan-Tang Chou 2023

All Rights Reserved

SEARCH FOR EXOTIC HIGGS BOSON DECAY TO MULTIPLE B-QUARKS WITH THE ATLAS DETECTOR AT LHC USING MACHINE LEARNING METHODS

A Dissertation Presented

by

YUAN-TANG CHOU

Approved as to style and content by:

DocuSigned by:

Verena Martinez Outschoorn

E94CB4F9BD5343E...

Verena Martinez Outschoorn, Primary Adviser

DocuSigned by:

Rafael Coelho Lopes de Sa

3306B20E50A74DD...

Rafael Coelho Lopes de Sa, Co-Adviser

DocuSigned by:

Lorenzo Sorbo

87146B90EB54426...

Lorenzo Sorbo, Member

DocuSigned by:

Grant Wilson

CFE9E7EDAEE14A2...

Grant Wilson, Member

DocuSigned by:

Anthony Dinsmore

8DFF3ED1837A4D5...

Anthony Dinsmore, Department Head
Physics

DEDICATION

*To my beloved wife, Yi-Jen Lee,
and my parents, Yi-Hsiung Chou and Shu-Hua Chen,
whose unwavering support is the foundation of this journey.*

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisors Verena Martinez Outschoorn and Rafael Coelho Lopes de Sá. Their relentless support and boundless creativity make this a wonderful adventure beyond my imagination. Verena's exceptional communication skills, coupled with her profound knowledge of muon detectors, continue to impress me. It's always a pleasure to discuss with Rafael on interesting physics topics, and his enthusiasm for physics is truly impressive and encouraging. I also give my deep thanks to Lorenzo Sorbo and Grant Wilson for their valuable comments and thought-provoking questions that have allowed me to delve deeper into the underlying physics questions.

I would like to thank the also other faculty at the ATLAS group at the University of Massachusettes, particularly Stephane Willocq, for invaluable career advice, and Carlo Dallapiccola for engaging in intriguing conversations and discussions on various subjects even beyond physics.

During my two-year stay at UMass, I cherished the wonderful moments shared with my dear friends Sebastián Urrutia Quiroga, Samyukta Krishnamurthy, Matthew DeCapua, Siao-Fong Li, and Mingzhu Cui. Your companionship made my time truly enjoyable and memorable.

A special thank goes to my fellow grad students spanning many years, including Huacheng Cai, Mazin Khader, Cooper Wagner, Makayla Vessella, Dale Abbott, Jackson Burzynski, Jay Sandesara, and Margaret Lutz. Especially, Mazin's foundational work on the search conducted in this analysis deserves particular recognition. Also, I want to express my gratitude to the remarkable postdocs in our group, Peter Tor-

nambe, Edward Moyse, Attilio Picazio, Nora Pettersson, Martina Javurkova, William Leight, and Michiel Jan Veen, for generously responding to my inquisitive questions.

The development of flavor tagging plays a key role in this thesis, and I am indebted to Rafael Teixeira de Lima for his invaluable assistance during the early stages of its development. Furthermore, I am grateful for the support and feedback received from the ATLAS Flavor-tagging combined performance group conveners, including Valerio Dao, Bingxuan Liu, Philipp Gadow, and Manuel Guth, as their contributions have been instrumental in completing this work.

I wish to express my appreciation to Guillermo Loustau De Linares, Thiago Costa De Paiva, and Davide Cieri for their support in helping me understand the simulation and design of new trigger systems during my involvement in the HL-LHC upgrade for the LOMDT trigger project.

I also would like to thank all my friends inside and outside the collaboration, Jay Chan, Meng-Ju Tsai, Ya-Feng Lo, Yvonne Ng, and my wonderful undergrad pals, who also embarked on (and all awarded!) a Ph.D. with me.

Lastly, this thesis would not have been possible without the remarkable collaboration of my fantastic colleagues, Judith Höfer, Paula Martinez Suarez, and Rickard Ström, who have been not only thoughtful but also delightful to work with. Additionally, the insights gained from senior colleagues in the analysis team have been immensely beneficial throughout my studies. I would like to extend my thanks to Claudia Seitz, Imma Riu, and Aurelio Juste Rozas for their valuable input in this regard.

There are many more names that I, unfortunately, could not write all them here, but this adventure would not be possible without their help. I would like to thank all the amazing scientists and friends I met along my journey. Thank you.

ABSTRACT

SEARCH FOR EXOTIC HIGGS BOSON DECAY TO MULTIPLE B-QUARKS WITH THE ATLAS DETECTOR AT LHC USING MACHINE LEARNING METHODS

SEPTEMBER 2023

YUAN-TANG CHOU

B.Sc., NATIONAL TSING HUA UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Verena Martinez Outschoorn and Professor Rafael Coelho
Lopes de Sa

The discovery of the Higgs boson has opened up new possibilities for investigating physics beyond the Standard Model (SM). New particles may interact with the SM through the Higgs boson, and deviations from SM predictions can indicate the presence of new physics. This dissertation focuses on the search for exotic Higgs decay, $H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b})$ where a is a new scalar boson and focuses on the case where the Higgs boson is produced in association with a Z boson. The data were collected by the ATLAS detector at the Large Hadron Collider at center-of-mass energy $\sqrt{s} = 13$ TeV from 2015 to 2018 with a total integrated luminosity of 140 fb^{-1} . A new general-purpose low-mass double-b tagging algorithm was developed and applied in the search. With new machine-learning reconstruction techniques, the search is able to set a strong limit on the branching ratio of $\text{Br}(H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b}))$. Furthermore,

the search also considers models with two new scalars of different masses, a_1 and a_2 , and sets the limits on the branching ratio of $\text{Br}(H \rightarrow a_1 a_2 \rightarrow (b\bar{b})(b\bar{b}))$

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	v
ABSTRACT	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
 CHAPTER	
INTRODUCTION	1
1. THE STANDARD MODEL OF PARTICLE PHYSICS AND BEYOND	3
1.1 The Standard Model	3
1.1.1 Electroweak Symmetry Breaking and the Higgs Mechanism	7
1.1.2 The Higgs boson during the LHC era	9
1.2 Unsolved question in the nature	13
1.3 Physics beyond Standard Model	13
1.3.1 Two-Higgs-doublet model	15
1.3.2 Two-Higgs-doublet plus scalar model	16
1.4 Experimental landscape	17
2. THE ATLAS EXPERIMENT AT LHC	20
2.1 The Large Hadron Collider	20
2.1.1 The CERN accelerator complex	21
2.1.2 Luminosity	22
2.1.3 Pileup	24

2.2	The ATLAS detector	24
2.2.1	Inner Detector	26
2.2.2	Calorimeters	29
2.2.3	Muon Spectrometer	31
2.2.4	Magnetic systems	33
2.2.5	Trigger System	35
2.3	High-Luminosity LHC and ATLAS detector upgrade	36
2.3.1	Upgrade for Level-0 MDT trigger processor	38
3.	PHYSICS OBJECT RECONSTRUCTION	43
3.1	Electrons	43
3.2	Muons	44
3.3	Jets	46
3.4	Soft secondary vertex	49
3.5	Missing Transverse Energy	50
4.	DATA AND MONTE CARLO SIMULATION	52
4.1	Data	52
4.2	Monte Carlo Simulation	53
4.2.1	Simulation of exotic Higgs boson decay $H \rightarrow aa/a_1a_2 \rightarrow b\bar{b}b\bar{b}$	53
4.2.2	Truth categorization of simulated events	58
4.2.3	Simulation of Z + jets background	61
4.2.4	Simulation of $t\bar{t}$ background	61
4.2.5	Simulation for minor background	63
5.	DEEP SETS BASED NEURAL NETWORKS FOR LOW-PT $X \rightarrow b\bar{b}$ IDENTIFICATION	64
5.1	Boosted jet tagging in ATLAS	64
5.2	Special object reconstruction	66
5.2.1	Track-jet reconstruction	66
5.2.2	Multiple secondary vertex reconstruction	67
5.2.3	Track subjet reconstruction	68
5.2.4	Jet flavor labeling	69
5.3	Algorithm architecture	69
5.3.1	Track neural network	71

5.3.2	Secondary vertex neural network	71
5.3.3	Color charge adversarial neural network	72
5.4	Algorithm performance	74
5.4.1	Performance dependency on resonance color charge	76
5.4.2	Performance dependency on resonance mass	76
5.5	Efficiency measurement in data and simulation	80
5.5.1	Event selection	81
5.5.2	Methodology	84
5.5.3	Systematic uncertainties	86
5.5.4	Results	89
5.6	Summary	96
6.	SEARCH FOR EXOTIC HIGGS BOSON DECAY TO MULTIPLE B-QUARKS	97
6.1	Objects selection	98
6.1.1	DNN energy-momentum estimate regression	98
6.1.2	mu-in-jet p_T correction	104
6.2	Event selection	104
6.3	Analysis strategy	105
6.3.1	Event reconstruction using a deep neural network	106
6.3.2	Event selection	113
6.4	Systematic uncertainties	123
6.4.1	Experimental uncertainties	124
6.4.2	Modeling uncertainties	128
6.4.3	Summary of systematic uncertainties	131
6.5	Statistical analysis	133
6.6	Results	135
6.6.1	Fits to the Asimov dataset	136
6.6.2	Fits to blinded data	138
6.6.3	Fits to unblinded data	141
6.6.4	Upper limits on $H \rightarrow 2a \rightarrow 4b$ production	145
6.6.5	Upper limits on $H \rightarrow a_1 a_2 \rightarrow 4b$	147
6.7	Discussion and ideas for future searches	149

7. CONCLUSION	153
---------------------	-----

APPENDICES

A. DETAILS OF THE NEURAL NETWORK TRAINING	155
---	-----

B. RESULTS FOR OTHER SIGNALS	158
------------------------------------	-----

BIBLIOGRAPHY	203
--------------------	-----

LIST OF TABLES

Table	Page
3.1 Selection applied for the vertex reconstruction of the TC-LVT Loose working point.	51
4.1 The un-prescaled single-lepton trigger used for 2015, 1016, 2017, and 2018 data selection.	53
4.2 List of mass hypotheses considered for this analysis.	54
4.3 Summary of MC generators used for the rare process with small contributions in the studies.	63
5.1 Jet flavor labeling based on number of ghost-associated b -hadrons or c -hadrons.	69
5.2 List of features used as input for the DEXTER algorithm.	73
5.3 Ratio of the efficiency measured in data to MC for different jet labels, p_T ranges, and tagging intervals.	89
5.4 Breakdown of the SF uncertainties for each summary category in different tagging intervals and p_T range for B -labeled jets.	94
5.5 Breakdown of the SF uncertainties for each summary category in different tagging intervals and p_T range for b -labeled jets.	95
6.1 List of features used as input for the DNN mass estimate regression.	100
6.2 List of features used as input to the hypothesis testing NN.	111
6.3 Overview of the signal regions targeting the 4b final state. Each signal region exists for each mass hypothesis.	120
6.4 Overview of the Z + jets control regions targeting the 4b final state.	122

6.5	Overview of the $t\bar{t}$ control regions targeting the 4b final state.	123
6.6	List of tracking systematic uncertainties included in the efficiency measurement.	128
6.7	Relative variation in the fraction of each component of $Z + \text{heavy flavor}$	131
6.8	Summary of all systematic uncertainties considered in the measurement, grouped categories.	132
6.9	Summary of the background normalization factors that are free-floating in the fit and the corresponding background components.	135
A.1	List of hyperparameters for DNN Energy-momentum estimate regression.	156
A.2	List of hyperparameters for pairing hypothesis testing NN.	157

LIST OF FIGURES

Figure	Page
1.1 A infographic for the Standard Model	5
1.2 Standard Model Higgs boson production cross sections as a function of the center of mass energy [10].	10
1.3 Standard Model Higgs boson production mechanisms (a) gluon fusion (b) Vector boson fusion (c) Higgs-strahlung (d) associated production with a pair of top quarks	11
1.4 Standard Model Higgs boson decay branching ratio [10].	12
1.5 Ratio of the production cross section times branching ratio that best fits available data relative to the SM expectation for individual decay modes [3].	12
1.6 The decay branching ratio of a pseudoscalar particle, a , with Yukawa couplings similar to the Higgs boson as a function of the mass m_a in a model in which the new CP-odd scalar couples with the same strength to up and down-type fermions [4].	17
1.7 Upper limit at 95% CL on $\sigma(ZH) \times \text{Br}$, where $\text{Br} = \text{Br}(H \rightarrow 2a) \times \text{Br}(a \rightarrow bb)^2$, versus m_a	18
2.1 The LHC accelerator complex and the four major experiments, ATLAS, CMS, ALICE, and LHCb.	21
2.2 The CERN accelerator complex.	22
2.3 Total Integrated Luminosity and Data Quality during Run-2 data-taking in 2015-2018.	23
2.4 Number of Interactions per Crossing.	24
2.5 The ATLAS detector with a cutaway diagram of the sub-detector systems.	25

2.6	Cutaway diagram of ATLAS inner detector systems.....	27
2.7	Cutaway diagram of ATLAS inner detector systems.....	27
2.8	Cutaway diagram of ATLAS calorimeter detector.....	30
2.9	The cumulative amount of material, in units of interaction length, as a function of $ \eta $, in front of the electromagnetic calorimeters, in each electromagnetic and hadronic compartment, and the total amount at the end of the active calorimetry.....	31
2.10	Cutaway diagram for muon sub-detector system.	32
2.11	Schematics diagram for the ATLAS magnetic system.	34
2.12	The ATLAS trigger systems, the detector read-out, and data flow.	35
2.13	LHC/HL-LHC Plan updated in February 2022	36
2.14	Cross-sectional view of the Phase-II ATLAS muon spectrometer.	39
2.15	The transverse momentum of muon reconstructed with offline reconstruction algorithm for muon candidate select by the Level-1 trigger with 20 GeV threshold.....	40
2.16	The trigger efficiency in $ \eta < 2.4$ as a function of p_T measure by offline reconstruction.	40
2.17	Schematic drawing of two momentum estimation methods using MDT track segments.	41
3.1	Illustration of how ATLAS detector detects particles.	44
3.2	Stages of jet energy scale calibrations.	47
4.1	p_T distribution of the Higgs boson and the Z boson for the $ZH + ggZH$ production mode using POWHEG under the $H \rightarrow aa \rightarrow 4b$ signal hypothesis.	55
4.2	Kinematic distributions for the a -bosons under the $H \rightarrow aa \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode.	56
4.3	Kinematic distributions for the a -bosons under the $H \rightarrow a_1 a_2 \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode.	56

4.4	Angular separation of b -pairs from the same parent under the $H \rightarrow aa \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode.	57
4.5	Angular separation of b -pairs from the same parent under the $H \rightarrow a_1a_2 \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode.	57
4.6	p_T distributions for the b -quarks coming from the $H \rightarrow aa \rightarrow 4b$ decay for the $ZH + ggZH$ production mode.	58
4.7	p_T distributions for the b -quarks coming from the $H \rightarrow a_1a_2 \rightarrow 4b$ decay for the $ZH + ggZH$ production mode.	59
5.1	Architecture of the DEXTER algorithm.	70
5.2	Architecture of the DEXTER algorithm with an adversarial NN.	74
5.3	DEXTER discriminant distribution	75
5.4	The saliency map of D_B with respect to the track inputs.	77
5.5	The saliency map of D_B with respect to the SV inputs.	78
5.6	Comparison of the B -labeled jet tagging efficiency vs. the b -labeled jet rejection for B -jets originating from color singlets (red) and color octets (blue).	78
5.7	Performance for different a -boson masses in $H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b})$ events.	79
5.8	DEXTER efficiency in simulation for (a) B - and (b) b -labeled jets as a function of the jet p_T	81
5.9	Diagram of the event selection and channel classification in the top-region.	83
5.10	Diagram of the event selection and channel classification in the Z -region.	85
5.11	Data-to-MC scale factors obtained for B -labeled jets in the (a) 40 – 60% and (b) 0 – 40% tagging intervals.	90
5.12	Data-to-MC scale factors obtained for b -labeled jets in the (a) 40 – 60% and (b) 0 – 40% tagging intervals.	90

5.13	Comparison between data and prediction for the event yields in the top-region channels after the efficiency correction.	91
5.14	Comparison between data and prediction for the event yields in the Z -region channels after the efficiency correction.	92
5.15	Comparison the muon $\text{Ex}k_t^{(2)} \langle S_{d_0} \rangle$ and muon $\text{Ex}k_t^{(2)} m_{\text{SV}}^{\text{max}}$ distribution between data and prediction in the Z -region, after efficiency correction	93
6.1	DNN mass estimate regression architecture.....	102
6.2	Comparison of the target and the predicted mass resolution distribution from the DNN energy-momentum estimation regression.	102
6.3	The large- R track jet mass comparison before and after the DNN Energy-momentum estimation regression.	103
6.4	The SM background compositions after applying the pre-selection.....	105
6.5	Diagram depicting how the NN performs event reconstruction.....	107
6.6	Diagram of NN used for event identification.	110
6.7	The pairing hypothesis score $\hat{y}_{\text{pred}}^{m_a}$ distribution assuming signal hypothesis $m_a = 25\text{GeV}$ for all the predicted categories.	112
6.8	Confusion matrix for fully merged categories where signal events contain two DeXTer jets.	113
6.9	Confusion matrix for semi-merged categories where signal events (a) contain one DeXTer jet and two b-tagged jets or (b) one DeXTer jet, one b-tagged jet, and one secondary vertex.....	114
6.10	Confusion matrix for fully resolved categories where signal events are reconstructed as (a) four b-tagged jets in the detector (b) three b-tagged jets and one secondary vertex.....	114
6.11	Normalized distribution of m_{red}^H in different event reconstruction channels.	116

6.12	Signal and background distributions of the input variables, per channel, for the mass-spin discriminator. Shown for the $m_a = 16$ GeV mass hypothesis.	118
6.13	Signal and background distributions of the input variables, per channel, for the mass-spin discriminator. Shown for the $m_a = 50$ GeV mass hypothesis.	119
6.14	Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset.	136
6.15	Validation regions showing the predicted yields in simulation before and after the fit to the Asimov dataset.	137
6.16	Nuisance parameters for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics.	138
6.17	Correlation matrix of all nuisance parameters and scale factors for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics.	139
6.18	Rankings based on the impact on the signal strength for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics.	140
6.19	Normalization factors for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics.	141
6.20	Predicted yields for the signal and control regions in the blinded fit to data before and after the fit.	142
6.21	Predicted yields for validation regions in the blinded fit to data before and after the fit.	142
6.22	Nuisance parameters and normalization factors for the blinded data fit using the full set of systematics.	143
6.23	Predicted yields for the signal and control regions in the blinded fit to data before and after the fit.	144
6.24	Predicted yields for validation regions in the blinded fit to data before and after the fit.	144

6.25	Nuisance parameters and normalization factors for the blinded data fit using the full set of systematics.	145
6.26	Post-fit predicted yields for the signal regions defined in the analysis (a) 2B (b) 1B2b (c) 1B1b1v (d) 4b (e) 3b1v in the unblinded fit to data with background-only hypothesis using the full set of systematics.	146
6.27	Expected (dashed line) 95% CL _{ij} upper limits on the $H \rightarrow 2a \rightarrow 4b$ cross-section times branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV.	147
6.28	Breakdown of expected (dashed line) 95% CL upper limits on the $H \rightarrow 2a \rightarrow 4b$ cross-section times branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV from different signal regions.	148
6.29	Expected (dashed line) 95% CL upper limits on the $H \rightarrow 2a \rightarrow 4b$ branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV assuming the SM cross section of ZH	148
6.30	Observed (dashed line) 95% CL upper limits on the $H \rightarrow 2a \rightarrow 4b$ branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV assuming the SM cross section of ZH	149
6.31	Expected 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ cross-section times branching fraction for different (m_{a_1}, m_{a_2}) values.	150
6.32	Breakdown of expected (dashed line) 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ cross-section times branching fraction for different (m_{a_1}, m_{a_2}) values from different signal regions.	150
6.33	Expected 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ branching fraction for different (m_{a_1}, m_{a_2}) values assuming the SM cross section of ZH	151
6.34	Observed 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ branching fraction for different (m_{a_1}, m_{a_2}) values assuming the SM cross section of ZH	151
B.1	Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_a = 12$ GeV mass hypothesis.	158

B.2	Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset for the $m_a = 12$ GeV mass hypothesis.	159
B.3	Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a = 12$ GeV mass hypothesis is considered.	160
B.4	Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_a = 16$ GeV mass hypothesis.	161
B.5	Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset for the $m_a = 16$ GeV mass hypothesis.	162
B.6	Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a = 16$ GeV mass hypothesis is considered.	163
B.7	Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_a = 20$ GeV mass hypothesis.	164
B.8	Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset for the $m_a = 20$ GeV mass hypothesis.	165
B.9	Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a = 20$ GeV mass hypothesis is considered.	166
B.10	Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_a = 30$ GeV mass hypothesis.	167
B.11	Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset for the $m_a = 30$ GeV mass hypothesis.	168

B.12 Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a = 30$ GeV mass hypothesis is considered.	169
B.13 Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_a = 50$ GeV mass hypothesis.	170
B.14 Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset for the $m_a = 50$ GeV mass hypothesis.	171
B.15 Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a = 50$ GeV mass hypothesis is considered.	172
B.16 Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_a = 60$ GeV mass hypothesis.	173
B.17 Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset for the $m_a = 60$ GeV mass hypothesis.	174
B.18 Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a = 60$ GeV mass hypothesis is considered.	175
B.19 Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV mass hypothesis.	176
B.20 Validation regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV mass hypothesis.	177
B.21 Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV mass hypothesis is considered.	178

B.22	Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_{a_1}=40$ GeV, $m_{a_2}=60$ GeVmass hypothesis.	179
B.23	Validation regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_{a_1}=40$ GeV, $m_{a_2}=60$ GeVmass hypothesis.....	180
B.24	Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_{a_1}=40$ GeV, $m_{a_2}=60$ GeVmass hypothesis is considered.	181
B.25	Signal and control regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_{a_1}=50$ GeV, $m_{a_2}=70$ GeVmass hypothesis.	182
B.26	Validation regions showing the predicted yields in simulation before and after the fit to the Asimov dataset for the $m_{a_1}=50$ GeV, $m_{a_2}=70$ GeVmass hypothesis.....	183
B.27	Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_{a_1}=50$ GeV, $m_{a_2}=70$ GeVmass hypothesis is considered.	184
B.28	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_a = 12$ GeVmass hypothesis.	185
B.29	Nuisance parameters (top) and normalization factors (bottom)for the blinded data fit using the full set of systematics for the $m_a = 12$ GeVmass hypothesis.	186
B.30	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_a = 16$ GeVmass hypothesis.	187
B.31	Nuisance parameters (top) and normalization factors (bottom)for the blinded data fit using the full set of systematics for the $m_a = 16$ GeVmass hypothesis.	188
B.32	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_a = 20$ GeVmass hypothesis.	189

B.33	Nuisance parameters (top) and normalization factors (bottom right)for the blinded data fit using the full set of systematics, the $m_a = 20$ GeVmass hypothesis.	190
B.34	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_a =$ 30 GeVmass hypothesis.	191
B.35	Nuisance parameters (top) and normalization factors (bottomfor the blinded data fit using the full set of systematics for the $m_a = 30$ GeVmass hypothesis.	192
B.36	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_a =$ 50 GeVmass hypothesis.	193
B.37	Nuisance parameters (top) and normalization factors (bottom)for the blinded data fit using the full set of systematics for the $m_a = 50$ GeVmass hypothesis.	194
B.38	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_a =$ 60 GeVmass hypothesis.	195
B.39	Nuisance parameters (top) and normalization factors (bottom right)for the blinded data fit using the full set of systematics for the $m_a = 60$ GeVmass hypothesis.	196
B.40	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeVmass hypothesis is considered.	197
B.41	Here the $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeVmass hypothesis is considered. Nuisance parameters (top) and normalization factors (bottom right)for the blinded data fit using the full set of systematics.	198
B.42	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeVmass hypothesis is considered.	199
B.43	Here the $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeVmass hypothesis is considered. Nuisance parameters (top) and normalization factors (bottom)for the blinded data fit using the full set of systematics.	200

B.44	Predicted yields for the signal, control, and validation regions in the blinded fit to data before and after the fit for the $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV mass hypothesis is considered.	201
B.45	Here the $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV mass hypothesis is considered. Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics.	202

INTRODUCTION

The observation of the Higgs boson by the ATLAS and CMS collaborations in 2012 [1, 2] opened new doors for physicists to study the nature of the universe. With years of scrutiny, the Standard Model still holds an extraordinary level of agreement in most of the experimental measurements. Nevertheless, several unsolved problems indicate the need for a theory beyond the Standard Model.

The Higgs boson was discovered by observing a small number of decay modes [1, 2]. With several years of studies, the available measurements are only able to constrain non-SM decays to $\lesssim 12\%$ of all decays at 95% confidence level (C.L.) [3] under several assumptions. Searches for exotic Higgs decays provide a direct way to find a hint of beyond the SM physics and shed light on the road to complete the puzzle. This thesis is structured as follows:

In chapter 1, I start by laying the foundation and summarizing our current understanding of the Standard Model. This leads us to unsolved questions and observations still not explained by the Standard Model. These intriguing phenomena motivate us to search for exotic Higgs boson decays using particle colliders.

Chapter 2 presents an overview of the experimental apparatus for high-energy particle physics experiments, the Large Hadron Collider, and the ATLAS detector. I discuss the ATLAS detector as it was used to collect the collision data used in this thesis and talk briefly about the planned future upgrade of the detector system.

Chapter 3 summarizes the process of physics object reconstruction, and Chapter 4 presents an overview of data and Monte Carlo simulated samples used in the following analysis.

The following two chapters focus on the search for exotic Higgs decays to multiple b-quarks with data collected from 2015 to 2018 by the ATLAS detector. The advancement of new identification techniques plays a pivotal role in enhancing the sensitivity of searches for final states involving b-quarks.

In Chapter 5, I detail the development of novel techniques to identify jets containing two b-quarks and calibrations using data.

Chapter 6 showcases the use of this new identification technique along with other modern machine learning methods to search for exotic Higgs decay to b-quarks. Finally, Chapter 7 provides a summary and discusses the results of this thesis.

CHAPTER 1

THE STANDARD MODEL OF PARTICLE PHYSICS AND BEYOND

The goal of particle physics is to study the fundamental mechanisms governing our universe. As far as we know, these mechanisms are described by four interactions: electromagnetic, strong, weak, and gravitational forces. Over several decades of theoretical and experimental advancements, our current understanding of nature is summarized in the Standard Model, which serves as a theoretical framework to describe the interaction of all known elementary particles except gravity.

In this chapter, I provide an overview of the Standard Model, briefly describing its underlying theoretical framework, its successes, and its limitations. This chapter aims to provide the theoretical foundation and motivation for the analysis performed in this thesis.

1.1 The Standard Model

The Standard Model (SM) is a theoretical model describing the properties and interactions of elementary particles. It incorporates three out of the four fundamental forces: electromagnetic, strong, and weak forces. It does not include the gravitational force. Various attempts have been made to have a theory of quantum gravity, such as string theory. In the context of high-energy physics experiments and the scope of this thesis, the gravitational force is too weak compared with the other three forces considered. Hence it's ignored in the following discussion.

The Standard Model is a quantum gauge field theory which encodes the description of the interactions between particles and three fundamental forces in the internal local

symmetries of the $SU(3)_C \times SU(2)_L \times U(1)_Y$. The elementary particles in the Standard Model are shown in Figure 1.1. For each particle, there is a corresponding quantum field where the particle is interpreted as the excitation of that field. Based on their spin, we could categorize them into two types: fermions and bosons.

Particles with half-integer spin are called fermions, which account for the building block of matter. There are two classes of fermions: leptons and quarks, and they are further categorized into three generations. Both are described by the Fermi-Dirac statistics. Each generation of lepton has a charged lepton and corresponding neutrinos. The neutrinos have no charge and hence do not interact via electromagnetic force. On the other hand, all the quarks have fractional electric charges and also color charges.

The other class of elementary particles with integer spin is called boson, which is described by the Bose-Einstein statistics. The force carriers and the Higgs boson in SM fall into this category. The gluon is the mediator of the strong force, which also carries color charges. The photon is the mediator for the electromagnetic force, which interact with all the particle with electric charges. Finally, the W-and Z-boson is responsible for the weak interaction.

The total Lagrangian of the SM can be summarized into four terms:

$$\mathcal{L}_{SM} = \mathcal{L}_{\text{Strong}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \quad (1.1)$$

where the first term describes the strong interaction, and the second term describes the electroweak interaction. The third term represents the kinetic and potential terms of the scalar field and the last term is the Yukawa interaction of fermions with the scalar field. In the following, a short discussion is given to explain the role of a particular term in SM.

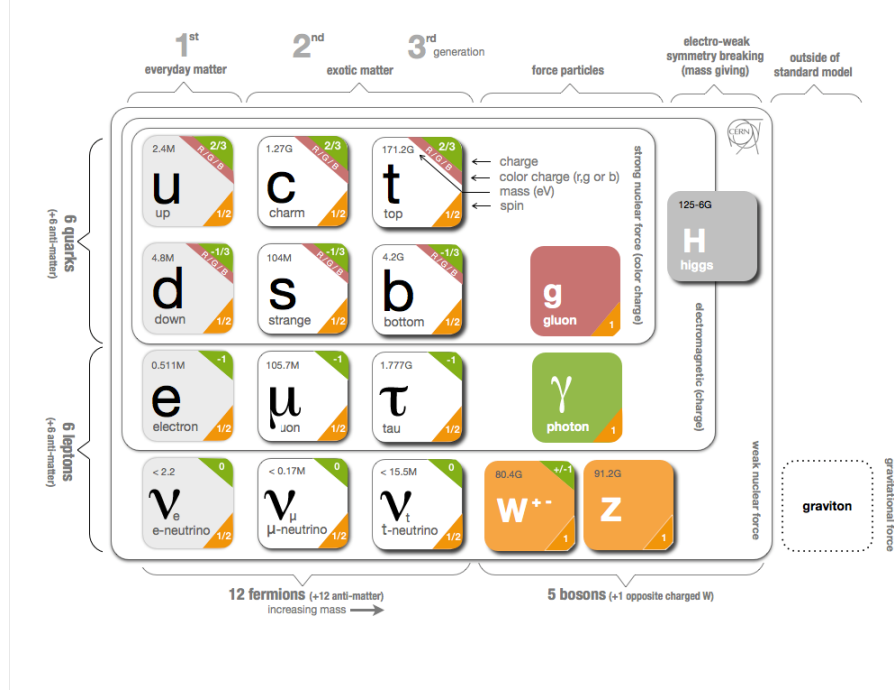


Figure 1.1: A infographic for the Standard Model [4].

Quantum Chromodynamics

The underlying theory for the strong force is called Quantum chromodynamics (QCD), which is a non-abelian gauge theory with $SU(3)_C$ symmetry. The QCD Lagrangian, $\mathcal{L}_{\text{Strong}}$, can be expanded as follow:

$$\mathcal{L}_{\text{Strong}} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + \sum_f \bar{\Psi}_f (i\mathcal{D})\Psi_f, \quad (1.2)$$

where Ψ_f is the Dirac spinor of the quark field, where f is the quark flavor and the $F_{\mu\nu}^a$ is the field strength tensor defined by

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - ig_s f^{abc} A_\mu^b A_\nu^c, \quad (1.3)$$

where A_μ^a is the components of gluon field given by $\mathbf{A}_\mu = A_\mu^a t^a$ with $a = 1, \dots, 8$, f^{abc} is the structure constants of the $SU(3)_C$ group and $\mathcal{D} \equiv \gamma_\mu \mathbf{D}^\mu$ where $\mathbf{D}_\mu$ is the covariant

derivative defined by $\mathbf{D}_\mu = (\partial_\mu - ig_s \mathbf{A}_\mu)$. The g_s is the QCD coupling constant, and t^a is the generator of $SU(3)_C$ in the fundamental representation.

Despite its simple form, the QCD describes the interaction between quarks and gluons, which is essential to understand the physics of a hadron collider. One of the interesting features of the QCD is that the strong coupling change as a function of energy. This property is known as asymptotic freedom [5, 6]. At low energies, the coupling becomes stronger. As a result, there are no free quarks observed in nature. At very high energies, the interactions between quarks and gluons become weak, and perturbation theory can be applied to calculate their interaction.

Electroweak Theory

The next part of the Lagrangian describes the electroweak theory,

$$\begin{aligned} \mathcal{L}_{\text{EW}} = & \bar{Q}_i i\gamma^\mu D_\mu Q_i + \bar{L}_i i\gamma^\mu D_\mu L_i + \bar{e}_R^i i\gamma^\mu D_\mu e_R^i + \bar{\nu}_R^i i\gamma^\mu D_\mu \nu_R^i \\ & + \bar{u}_R^i i\gamma^\mu D_\mu u_R^i + \bar{d}_R^i i\gamma^\mu D_\mu d_R^i - \frac{1}{4} W_a^{\mu\nu} W_{\nu\mu}^a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \end{aligned} \quad (1.4)$$

where Q_i and L_i is the left-handed doublet for each generation of leptons and quarks field, u_R and d_R are right-handed singlet up and down quark fields, e_R and ν_R are vectors of right-handed lepton fields. The covariant derivative, D_μ , is defined as $D_\mu = (\partial_\mu + ig\sigma_a W_\mu^a/2 + ig'Y B_\mu/2)$, where g and g' are the $SU(2)_L$ and $U(1)_Y$ gauge couplings, respectively. The $W_{\mu\nu}^a$ are field strength tensors of a W_μ gauge field indexed in weak isospin space a and the $B_{\mu\nu}$ is field strength tensors associated with B_μ gauge field with the $U(1)$ symmetry. The Lagrangian describes the interaction between the gauge boson and fermions. The first six describes the interaction between fermions and gauge bosons. The last two terms are the interactions between the gauge boson. However, there is no mass term for the gauge boson, which could not explain the observed massive boson in measurement. The introduction of the complex scalar

field later fixed this problem via a mechanism called spontaneous symmetry breaking of electroweak symmetry.

1.1.1 Electroweak Symmetry Breaking and the Higgs Mechanism

To introduce the mass of the gauge boson while maintaining the gauge invariance, the Brout–Englert–Higgs mechanism was introduced back in 1964 [7–9], which is also referred to as the Higgs mechanism. This mechanism introduces a new scalar field as follows:

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \quad (1.5)$$

The Lagrangian for the Higgs field is defined as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) + \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2. \quad (1.6)$$

where the covariant derivative acting on the Higgs doublet is

$$D_\mu \Phi = (\partial_\mu + igW_\mu^a \sigma^a + \frac{ig'}{2} B_\mu) \Phi. \quad (1.7)$$

The last two terms in equation 1.6 are the Higgs potential, where the λ is the self-coupling and μ is related to the mass term of the field. The minimum of the potential occurs at

$$\Phi^\dagger \Phi = \frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \quad (1.8)$$

where we require the $\lambda > 0$ and ν is the vacuum expectation value (VEV). If $\mu^2 < 0$, it will behave like a normal scalar field and have a minimum at zero. However, if $\mu^2 > 0$, the minimum potential energy for the scalar field does not occur for $|\phi| = 0$. There will be an infinite set of degenerate minima.

Similar to the other field in QFT, the excitation of a field corresponds to an elementary particle. One can be studied the scalar field by expanding around the minimum, giving us two terms. One corresponds to the Higgs boson, and others are associated with the Goldstone bosons. We can take the unitarity gauge, which removes the Goldstone boson terms and becomes the longitudinal degree of freedom of the massive W and Z boson.

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \quad (1.9)$$

The remaining term, $H(x)$, can be identified as the Higgs boson. After plugging the number into the Higgs potential, we could get the relation between the Higgs boson mass, λ , and μ to be

$$m_H = \sqrt{2\mu^2} = \sqrt{2\lambda v^2}. \quad (1.10)$$

Using the Higgs boson defined in equation 1.9 and 1.6, one could identify the mass term for W, Z boson, and photon from the expansion term with some algebra. The photon is massless, as expected. As for the mass for W, Z boson is given by

$$m_{W^\pm} = \frac{1}{2}gv, \quad (1.11)$$

and

$$m_Z = \frac{1}{2}\sqrt{g^2 + g'^2}, \quad (1.12)$$

where g and g' are not predicted by the theory and hence can only be determined with experimental measurement.

Similarly, fermions also gains their mass by interacting with the Higgs field, which is described in the Yukawa term in equation 1.1, which is defined as:

$$\mathcal{L}_{\text{Yukawa}} = -\sum_f m_f (\bar{f}_L f_R + \bar{f}_R f_L), \quad (1.13)$$

where f are fermions and the mass of a fermion is

$$m_f = \frac{v}{\sqrt{2}} y_f, \quad (1.14)$$

where y_f is the Yukawa coupling for the fermion. The fermion masses are not predicted in the SM, which gives an additional nine free parameters and are determined solely from the experiment. In the next section, I will discuss the relevant aspect of the Higgs boson in the collider experiment.

1.1.2 The Higgs boson during the LHC era

The observation of a Higgs boson by the ATLAS and CMS collaborations [1, 2] has started a new phase in the LHC physics program aimed at studying the properties of the Higgs boson and search for new physics. In the following, I will briefly describe the property of the Higgs boson and how we study them with the LHC.

Higgs boson production There are four main Higgs boson production mechanisms to produce Higgs boson in a hadron collider. The cross-section predicted per production mode as a function of the center of mass energy can be seen in Figure 1.2

- **Gluon fusion (ggF):** This is the production mechanism with the largest cross-section. The two gluons interact within a loop to produce the Higgs boson, as shown in Figure 1.3 (a). The dominant contribution comes from the top quark. Contributions from lighter quarks are suppressed due to the smaller Yukawa coupling.
- **Weak boson fusion (VBF)** This is the production mechanism with the second-largest cross-section. The Higgs boson is produced via $W^\pm/Z$ radiation

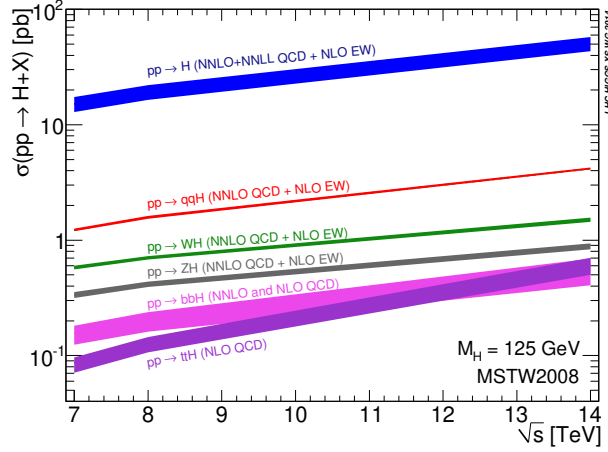


Figure 1.2: Standard Model Higgs boson production cross sections as a function of the center of mass energy [10].

from the quark of the proton, as shown in Figure 1.3 (b). The VBF production has its unique feature of two high-energy jets at the forward regions of the detector.

- **Associated production with a gauge boson (VH):** Higgs-strahlung as shown in Figure 1.3 (c) or in a loop from gluon-gluon interaction. This is the most relevant Higgs boson production mechanism for this thesis, given that the leptonic decay of vector bosons can be used to select the events and help suppress large QCD backgrounds. This is crucial for searching with the hadronic final state of the Higgs boson, such as $H \rightarrow b\bar{b}$
- **Associated production with a pair of top quarks ($t\bar{t}H$):** The Higgs boson can be produced by exchanging a top quark between two gluons and radiating a Higgs boson. This production mechanism allows the direct probe of top-Higgs Yukawa coupling. There is also a similar production for the associated production with a pair of bottom quarks with smaller cross-sections.

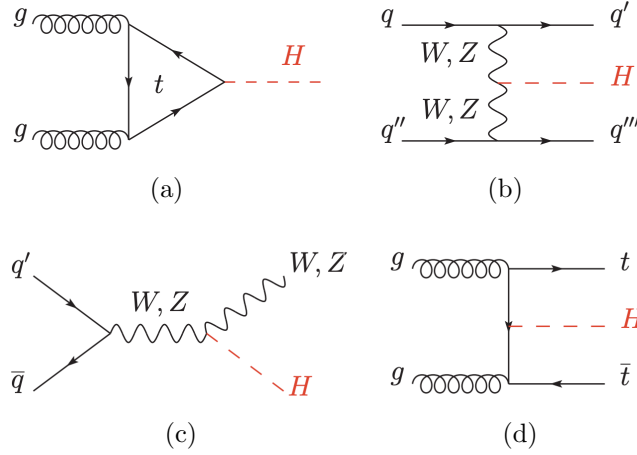


Figure 1.3: Standard Model Higgs boson production mechanisms (a) gluon fusion (b) Vector boson fusion (c) Higgs-strahlung (d) associated production with a pair of top quarks

Decay mode of Higgs boson As discussed in Section 1.1.1, the coupling of the SM Higgs boson is scaled as a function of the mass of the particles. As a result, the Higgs boson predominately decays to heavier particles. The decay branching ratio as a function of Higgs boson mass predicted by the SM is shown in Figure 1.4. The largest decay mode is the $H \rightarrow b\bar{b}$, which accounts for around 57.9 % of the branching ratio. This is followed by the $H \rightarrow WW^*$ channel with a branching ratio of 21.7%.

Several decay modes have been observed and measured at the LHC. The latest result is shown in Figure 1.5, consistent with the SM predictions [3]. However, available measurements are only able to constrain non-SM or *exotic* decays to $\lesssim 12\%$ of all decays at 95% confidence level (C.L.) [3] on the corresponding branching. These decays are interesting because they may be a window to new physics, including scenarios called *Higgs portals*, where the Higgs boson is the leading or only mediator between the SM and new physics sectors.

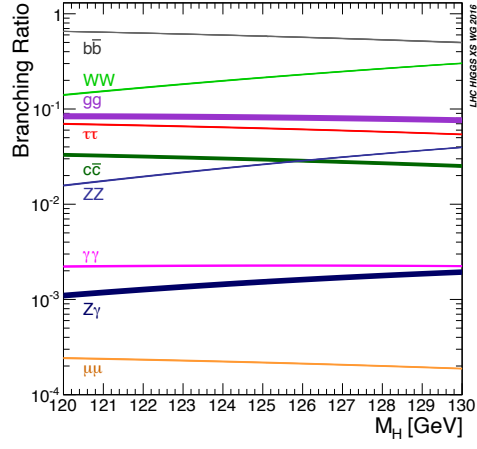


Figure 1.4: Standard Model Higgs boson decay branching ratio [10].

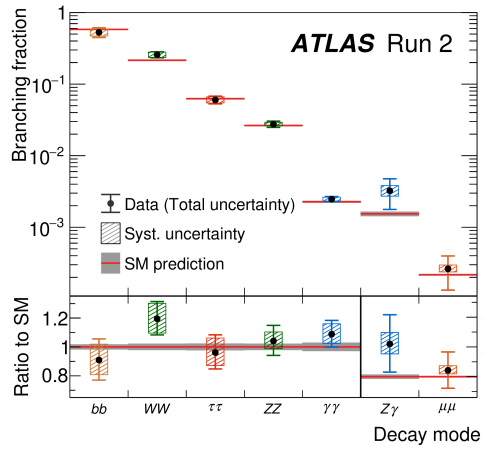


Figure 1.5: Ratio of the production cross section times branching ratio that best fits available data relative to the SM expectation for individual decay modes [3].

1.2 Unsolved question in the nature

Even after several years of studies, the SM has been overwhelmingly accurate in predicting experimental results across all kinds of experiments. However, some questions remain unanswered by the SM, such as dark matter and baryon asymmetry.

Dark matter Existing astrophysics results provide compelling evidence that the ordinary matter observed today only consists of 5% of the total mass-energy content of the universe [11–13]. Galaxy rotation curve, weak lensing of the Bullet Cluster [14], and measurements on cosmological scales of anisotropies in the cosmic microwave background [15, 16] are evidence of the existence of dark matter. However, no particles that exist in the SM can explain all observations simultaneously. One plausible explanation is introducing dark matter (DM), which does not interact electromagnetically.

Matter-antimatter asymmetry Matter and antimatter are believed to be produced in equal amounts at the end of inflation. However, the abundance of baryonic matter and almost no existence of antibaryonic matter is observed in the current universe. No mechanism in the SM can explain the current observed asymmetry.

Hierarchy problem The current measurement of the Higgs boson mass is too small and requires substantial quantum corrections from the virtual loop correction. The SM requires finetuning of the Higgs boson bare mass to agree with the current observed mass. This is considered uncomfortable from a theoretical point of view because there are no known physical mechanisms to produce a fine-tuned theory.

1.3 Physics beyond Standard Model

Numerous theory models try to incorporate phenomena not explained by considering new internal symmetry or adding new quantum fields to the SM. One particularly

interesting place to look for signs of beyond Standard Model is the Higgs boson. As mentioned early in Section 1.1.2, many decay modes are measured with data at the LHC. There is still plenty of room for *exotic* Higgs decay not predicted by the SM. These decays are interesting because they may be a window to new physics, including scenarios called *Higgs portals*, where the Higgs boson is the leading or only mediator between the SM and new physics sectors.

One of the simplest possibilities is that the Higgs boson decays into a pair of new spin-zero particles. This new spin-zero particle, a , will not exhibit electromagnetic, weak, or strong interactions but interact through the observed Higgs boson. The new light particle with even small couplings to the Higgs boson can result in sizeable exotic branching fractions because of the small total width of the Higgs boson. The new spin-0 particle can be either a scalar or a pseudoscalar under parity transformations depending on the model construction and how it mixes with the other scalar to acquire coupling to fermion.

Theories that introduce a new spin-zero scalar coupled to the Higgs Boson could provide a possible explanation for electroweak baryogenesis [17, 18]. The light scalar could also serve as the mediator between SM particles and new physics sectors, such as a mediator in dark matter models [19, 20] that avoid the strong constraint from direct detection.

Exotic decays are predicted by many theories of physics beyond the SM [4], including those with an extended Higgs sector (such as the Next-to-Minimal Supersymmetric Standard Model, NMSSM [21–25]), several models of dark matter [26–30], models with a first-order electroweak phase transition [18, 31], and theories with neutral naturalness [32–34].

In Section 1.3.1, I will discuss the two-Higgs-doublet model and later discuss how we can add additional scalar(s) that motivates the search for exotic Higgs decay in four b-quark final states.

1.3.1 Two-Higgs-doublet model

As discussed in the Section 1.1.2, the SM was completed by adding one complex scalar doublet to the theory. However, the number of doublets one could add to the theory is not limited by nature. One of the simplest extensions based on this idea is the two-Higgs-doublet model (2HDM) [35].

This type of model considers two SU(2) doublet:

$$\Phi_1 \equiv \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix}, \quad \Phi_2 \equiv \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix}, \quad (1.15)$$

where ϕ_1^+ , ϕ_1^0 , ϕ_2^+ , ϕ_2^0 are complex fields. Assuming that CP is conserved in the Higgs sector and all the parameters are real scalar potential of such fields can be written as:

$$\begin{aligned} V = & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1) + \frac{\lambda_1}{2} (\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2} (\Phi_2^\dagger \Phi_2)^2 \\ & + \lambda_3 \Phi_1^\dagger \Phi_1 \Phi_2^\dagger \Phi_2 + \lambda_4 \Phi_1^\dagger \Phi_2 \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2} [(\Phi_1^\dagger \Phi_2)^2 + (\Phi_2^\dagger \Phi_1)^2], \end{aligned} \quad (1.16)$$

Like the Higgs mechanism as SM, the new scalar field will result in two vacuum expectation values, v_1 and v_2 . And there are eight new degrees of freedom in the 2HDM model, resulting in five physical Higgs particles and three Goldstone bosons absorbed by the W and Z boson. The five physical Higgs states are two CP-even h and H , CP-odd A , and Charged Higgs boson $H^\pm$. The observed Higgs boson can be the mix of two CP-even states.

The model can be fully described by setting six of the free parameters in the model: the mass of the four Higgs bosons, the mixing angle between the two CP-even Higgs bosons α , and the free parameter describing the relation between the two vacua given by

$$\tan \beta = \frac{v_1}{v_2}. \quad (1.17)$$

However, the 2HDM models can introduce tree-level flavor-changing neutral currents, which have not been observed so far, which is avoided by introducing the Glashow-Weinberg condition [36]. The condition requires that each group of fermions couples exactly to one of the two doublets. Based on how the doublet interacts with fermion, the 2HDM model can be categorized into the following types:

- **Type I:** This type of model is called fermiophobic because charged fermions only couple to one doublet and provide all the mass to the fermions.
- **Type II:** This type of model is similar to the Minimal Supersymmetric Standard model. Only the up-type quark couple to the one of the doublet. The remaining doublet couple to both the down-type quark and charged leptons
- **Type III:** This type of model is commonly called flipped type, where one doublet couples to up-type quarks and charged leptons. The remaining doublet couples to the down-type quarks.
- **Type IV:** This type of model is called the lepton-specific type. In these models, one doublet couple to the quark and charged leptons couple to the other doublet.

1.3.2 Two-Higgs-doublet plus scalar model

The 2HDM models have been extensively tested at the LHC. Extensive parameter space has been probed, but no additional Higgs bosons predicted have been seen at the LHC. Further extensions of the 2HDM model have been proposed to introduce the possibility of exotic Higgs decay in these models. One of the simplest is adding an additional scalar field that mixes with the doublet state in 2HDM, given the null result from heavy Higgs searches at LHC. This extension allows the new particle can decay directly into fermions. In this case, depending on the type of 2HDM model

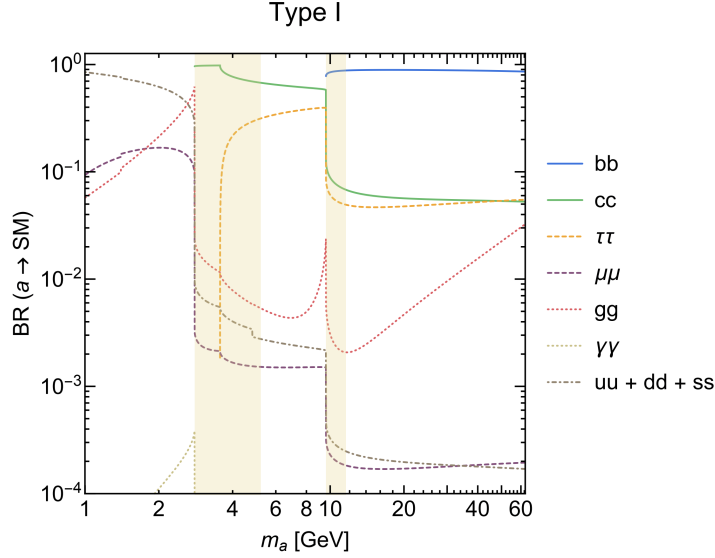


Figure 1.6: The decay branching ratio of a pseudoscalar particle, a , with Yukawa couplings similar to the Higgs boson as a function of the mass m_a in a model in which the new CP-odd scalar couples with the same strength to up and down-type fermions [4]. The shaded regions correspond to cases where the calculation is considered unreliable.

considered, the new light pseudoscalar a will inherit the Yukawa-like couple through mixing with the pseudoscalar A or Higgs boson [4].

In many scenarios where the Higgs boson decays to a pair of new spin-zero particles a , such as the model shown in Fig. 1.6, the dominant decay mode is to a pair of b -quarks ($a \rightarrow bb$) for $m_a > 10$ GeV, above the b -pair production threshold. Despite the experimental challenges associated with such a $4b$ signature ($H \rightarrow 4b$), this is expected to be one of the most sensitive decay modes [4, 37–39].

1.4 Experimental landscape

The analysis presented in this thesis is an update of two results published in 2018 and 2020 [40, 41]. Fig. 1.7 shows the upper limit at 95% CL on $\sigma(ZH) \times \text{Br}$, where $\text{Br} = \text{Br}(H \rightarrow 2a) \times \text{Br}(a \rightarrow bb)^2$, versus m_a from these publications. The analysis trying to reconstruct four fully resolved b -tagged jets is referred to as resolved

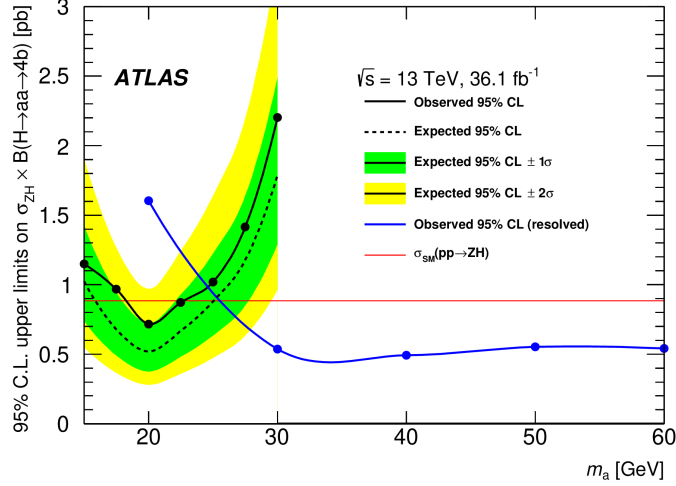


Figure 1.7: Upper limit at 95% CL on $\sigma(ZH) \times \text{Br}$, where $\text{Br} = \text{Br}(H \rightarrow 2a) \times \text{Br}(a \rightarrow bb)^2$, versus m_a . The observed (CLs) values (solid black line) are compared to the expected (median) (CLs) values under the background-only hypothesis (dotted black line). The surrounding shaded bands correspond to the 68% and 95% CL intervals around the expected (CLs) values, denoted by $\pm 1\sigma$ and $\pm 2\sigma$, respectively. The solid blue line indicated the same observed limit from the resolved channel analysis. The solid red line indicates the SM $pp \rightarrow ZH$ cross-section, assuming $\text{Br}(H \rightarrow 2a) \times \text{Br}(a \rightarrow bb)^2 = 1$ [41].

analysis. In the case where the $a \rightarrow bb$ decay is merged into a single reconstructed jet, the analysis is referred to as merged analysis. I contributed to the first merged analysis published in 2020 and gained experience in designing strategies for analysis in this thesis.

Several searches for a Higgs boson decaying to electrons, muons, taus, photons, or b -jets via two pseudoscalars have been performed at both the LHC and the Tevatron [40, 42–56]. These searches have led to limits on the branching ratio of the Higgs boson decaying to $2a$, scaled by the ratio of the production cross-section of the Higgs boson that is searched for to that predicted by the SM, $\sigma(H)/\sigma_{\text{SM}} \times \text{Br}(H \rightarrow 2a)$, between 1% and 3% for pseudoscalar masses between 1 GeV and 3 GeV and between 10% and 100% for masses larger than 5 GeV, in the SM+ a scenario where a is a new

Higgs-mixed scalar which inherits its coupling with SM fields via mixing with the Higgs boson as described in [57].

The signature of a decay of the Higgs boson to two light scalars of different mass $H \rightarrow a_1 a_2$ is yet uncovered. Hence the analysis presented in this thesis also explores the $H \rightarrow a_1 a_2$ decay for the first time. The phenomenology of a model where the SM is extended by two real scalar singlets is given in [58].

CHAPTER 2

THE ATLAS EXPERIMENT AT LHC

This chapter provides an overview of the experimental apparatus and aims to provide readers with background knowledge to understand the techniques and results described in the following chapters. An overview of the Large Hadron Collider (LHC) [59] is given in Section 2.1, along with a brief description of the accompanying accelerator complex and relevant terminology. The thesis used collision data recorded by the ATLAS (A Toroidal LHC ApparatuS) detector, comprised of various detector subsystems described in Section 2.2. Several planned upgrades aim to improve the capability of the LHC machine and ATLAS detector to enhance discovery and physics potential. Section 2.3 outlined these activities with a focus on developing the Level-0 MDT trigger processor upgrade for ATLAS, which is also part of this thesis work.

2.1 The Large Hadron Collider

The LHC is located at the European Organization for Nuclear Research, known as CERN (Conseil européen pour la recherche nucléaire), close to the border of Switzerland and France near Geneva. The LHC is a particle accelerator built in a 27 km circumference tunnel underground which is able to accelerate protons and heavy ions for physics studies. It inherits the same tunnel hosting the LEP (Large Electron-Positron Collider). There are four major experiments installed around the accelerator: ATLAS, CMS, LHCb, and ALICE. The ATLAS and CMS experiments are general-purpose particle detectors designed to study various physical final states. On the other hand, LHCb (Large Hadron Collider beauty experiment) and ALICE

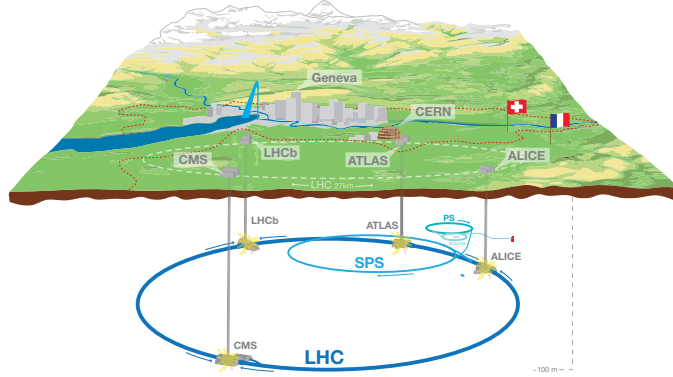


Figure 2.1: Overview of LHC, which resides under France–Switzerland border. The LHC accelerator complex and the four major experiments, ATLAS, CMS, ALICE, and LHCb, are shown. Image from Ref [60]

(A Large Ion Collider Experiment) experiments are specialized detectors focused on studying CP violation and heavy ion physics, respectively, which are beyond the scope of and not covered in this thesis. I will focus on describing the ATLAS experiment, which is used to collect data for this thesis. A cutaway diagram of the LHC and four major experiments are shown in Figure 2.1

2.1.1 The CERN accelerator complex

The LHC is the last component of the CERN accelerator complex, as shown in Figure 2.2. Many machines are chained to reach the targeted energy for the LHC. The first accelerator is the Linac2 to accelerate protons to 50 MeV, which feeds into the Proton Synchrotron Booster (PSB). Protons are accelerated to 1.4 GeV in PSB. The proton beam is sent to the PS (Proton Synchrotron) and further pushed to 25 GeV. The last step before injecting it into LHC is done by the SPS (Super Proton Synchrotron) to 450 GeV. The injected beam circulates into two pipes, one clockwise and the other anticlockwise. Both are accelerating to reach the targeted energy of

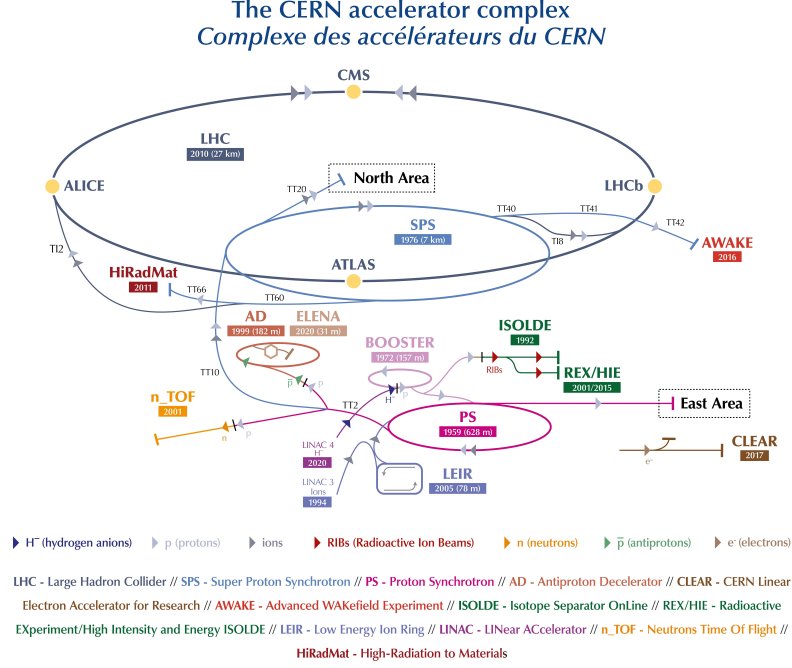


Figure 2.2: The CERN accelerator complex. The LHC is the last ring (dark blue line) in a complex chain of particle accelerators. The smaller machines are used in a chain to boost the particles to final energies and provide beams to other smaller experiments, which are not discussed. Illustrations credit Ref [61]

6.5 TeV per beam. After the beam reaches a stable condition for physics, it will be brought to collide at the center of ATLAS and CMS detector with the total center of mass energy of 13 TeV. The LHC machine can also accelerate heavy nuclei such as Pb or Xe for the LHC heavy ion program.

2.1.2 Luminosity

The events rate of the collision is proportional to the interaction cross-section and instantaneous luminosity $\mathcal{L}(t)$

$$\frac{dN}{dt} = \mathcal{L}(t)\sigma. \quad (2.1)$$

The instantaneous luminosity $\mathcal{L}(t)$ have the unit of $\text{cm}^{-2}\text{s}^{-1}$ which is proportional to machine paramters

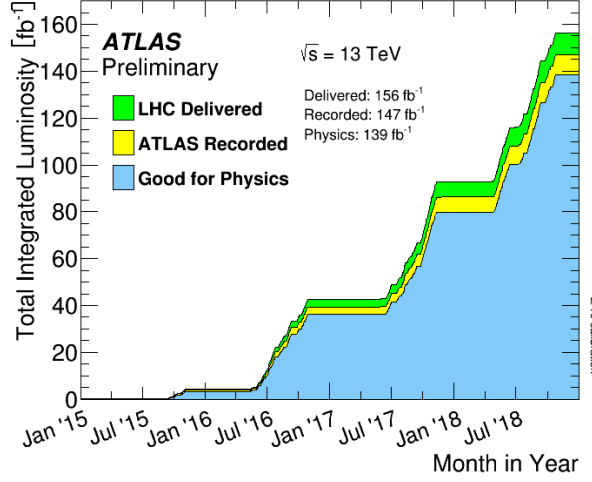


Figure 2.3: Total Integrated Luminosity and Data Quality during Run-2 data-taking in 2015-2018. Summary plot from Ref [62]

$$\mathcal{L} \propto \frac{f N_1 N_2}{4\pi\varepsilon\beta^*}, \quad (2.2)$$

where f is the bunches crossing frequency, N_1 and N_2 are the number of protons in a bunch, ε is the emittance of the beams, and β^* is the beta function of the beam which quantifies how narrow the beam. The nominal value is $\mathcal{L}(t) \approx 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ for the LHC. The total luminosity delivered by the LHC can be calculated by the integral of the instantaneous luminosity with respect to time

$$\mathcal{L}_{\text{integrated}} = \int \mathcal{L}(t) dt, \quad (2.3)$$

which is expressed in the inverse of cross-section, fb^{-1} . The total integrated luminosity delivered and collected by the ATLAS detector in 2015-2018 is shown in Figure 2.3. Only data with all the detector systems functioning are good for physics and are used in this thesis. More detail on the data will be described in Section 4.1.

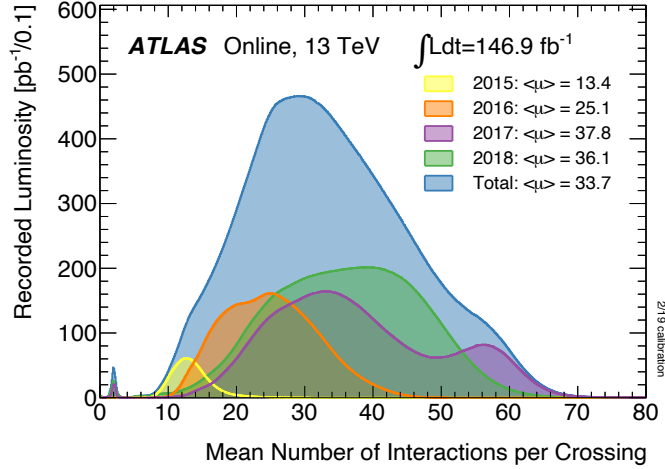


Figure 2.4: Number of Interactions per Crossing. Summary plot from Ref [62]

2.1.3 Pileup

The collision of bunches of protons usually result in multiple interactions per bunch crossing. The mean number of interactions per bunch crossing is shown in Figure 2.4. However, not all collisions are interesting for physics analysis. As a result, the interaction with the largest energy transfer during the collision is selected as the hard-scatter vertex using the associated tracks. The remaining interactions are called pileup, which is from the low-energy QCD process, which is less interesting. Pile-up will also deposit energy and spoil the measurement of the detector. Hence dedicated correction procedure must be carried out to mitigate this allowing us to properly reconstruct the physics objects for analysis, which will be discussed in Chapter 3.

2.2 The ATLAS detector

The ATLAS detector [63] covers nearly the entire solid angle around the collision point, which is 44 m long and 25 m in diameter. A cutaway view of the ATLAS detector is shown in Figure 2.5. The ATLAS detector consists of many sub-detector systems layers targeting to identify and reconstruct individual particles. The innermost tracking detector is surrounded by a superconducting solenoid magnet, electromag-

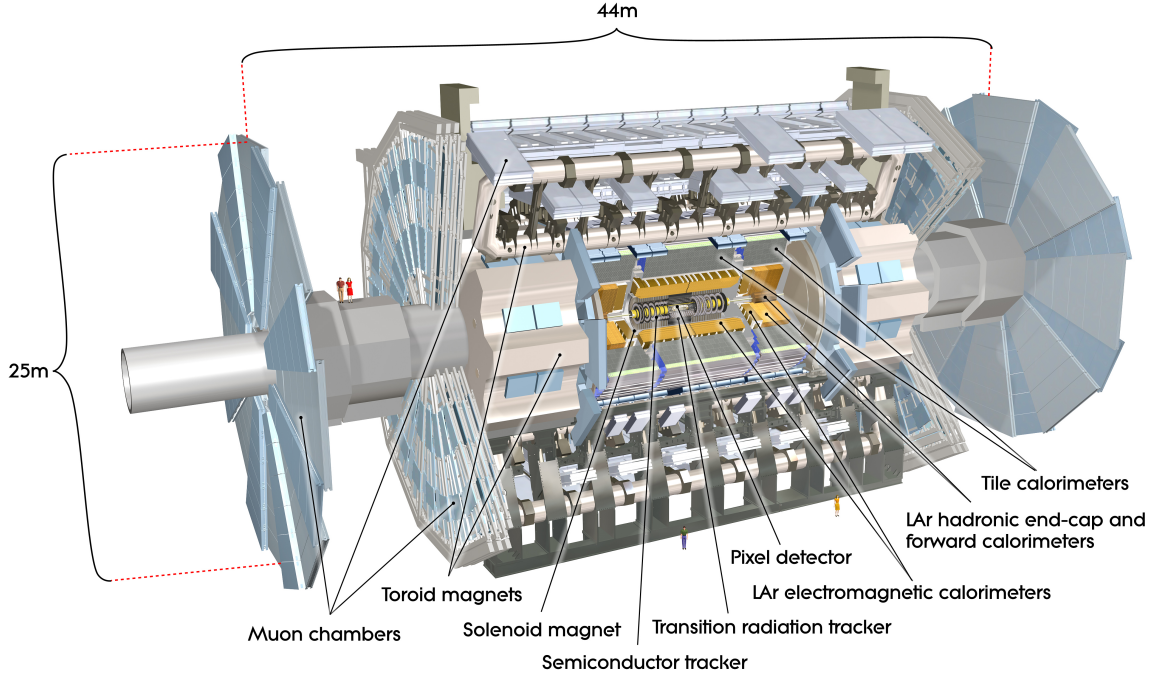


Figure 2.5: The ATLAS detector with a cutaway diagram of the sub-detector systems. Visualization credit to Ref. [64]

netic and hadronic calorimeters, and an external muon spectrometer incorporating three large toroid magnet assemblies. The ATLAS detector can be divided into two parts, barrel and endcap region. In this section, an overview of each subdetector is provided, along with conventions used in the following chapters.

Coordinate systems ATLAS uses a cartesian right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z -axis coinciding with the axis of the beam pipe. The x -axis points from the IP to the center of the LHC ring, and the z -axis along the tunnel. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. For a collider experiment, a significant unknown fraction of energy could escape down to the beam pipe. Defining kinematic variables in the transverse plane perpendicular to the beam pipe is more useful, as the sum of transverse momentum

is known to be zero. The transverse energy is defined as $E_T = E \sin \theta$. Similarly, the transverse momentum is defined as $p_T = \sqrt{p_x^2 + p_y^2}$. The rapidity of an object is defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \quad (2.4)$$

Another more common angular definition is pseudorapidity, defined in terms of the polar angle θ as

$$\eta = -\ln \tan(\theta/2). \quad (2.5)$$

In the ultrarelativistic limit, the pseudorapidity is approximately equal to rapidity. These allow the definition of commonly used angular distance defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (2.6)$$

2.2.1 Inner Detector

The inner detector system is the innermost component of the ATLAS detector close to the beam pipe. The goal is to provide high-precision measurements of the charged particles for the momentum and vertex reconstruction. Figure 2.6 shows the cutaway view of the inner detector systems. The barrel inner detector is arranged on the concentric cylinder around the beam and consists of three layers of detectors. In contrast, the end-cap systems are disk perpendicular to the beam pipe. A high-granularity silicon pixel detector (Pixel), along with including the insertable B-layer (IBL) [66], and a silicon microstrip tracker (SCT), together provide precision tracking in the pseudorapidity range $|\eta| < 2.5$, complemented by a transition radiation tracker (TRT) providing tracking and electron identification information for $|\eta| < 2.0$. All the inner detectors are immersed in the 2 Tesla magnetic field from the solenoidal magnet.

Pixel The Pixel Detector is a compact system that provides high-precision and high-granularity measurements of charge particles as close as the interaction point. It has four concentric layers in the barrel and three disks on each end cap, with 80.4 million readout channels. It uses 250 μm thick silicon as detecting material. Detecting charged particles in silicon detectors relies on the electron-hole pairs produced when they pass through. The electrons then drift to an electrode under the influence of an electric field, creating an electric signal which can be measured. The resolution can reach almost 10 μm in the $R - \phi$ plane and 11 μm in the z-direction. During the LHC shutdown in 2013-2014 (LS1), a new insertable B-Layer was inserted even closer to the beam pipe (3.3 cm) thanks to a new smaller radius beam pipe installed during the same period. This layer provides a measurement closer to the IP, critical for vertex reconstruction and the identified b-quark-initiated jets.

Silicon microstrip tracker (SCT) The SCT [68] is used to detect and reconstruct the tracks of charged particles produced during collisions. It consists of 4088 modules of silicon covering a surface of about 62 m^2 and has over 6 million readout channels. Its layout has been optimized so that each particle crosses at least four silicon layers in the intermediate radial range, yielding an excellent spatial resolution of 25 μm .

Transition radiation tracker (TRT) The TRT is the third layer of the Inner Detector and consists of about 300,000 thin-walled 4 mm diameter straw tubes. Each straw contains a 30 μm gold-plated tungsten wire in the center. The straw is filled with a Xe-based gas mixture (70% $\text{Xe} + 27\% \text{CO}_2 + 3\% \text{O}_2$). The technique is intrinsically radiation hard with a modest cost. The barrel TRT straws are parallel to the beam pipe and perpendicular to the beam axis. Charged particles passing through will ionize the gas mixture and be read out by the wire. The TRT system provides around 30 space points with $\approx 130 \mu\text{m}$ resolution for charged tracks with $|\eta| < 2$ and $p_T > 0.5 \text{ GeV}$. The TRT also exploits the emission of soft photons via

transition radiation by electrons in the polymer fibers interleaved with the straws to discriminate electrons from other particles.

2.2.2 Calorimeters

The ATLAS calorimeter subsector system aims to measure the energy of both charge and neutral particles from the collision, except muons and neutrinos. The calorimeters cover $|\eta| < 4.9$ and consist of electromagnetic and hadronic calorimeter subsystems. Calorimeters can stop most known particles except muons and neutrinos. Both are sampling calorimeters with active and passive material layers interleaved to increase the interaction length but keep the cost within budget. Figure 2.9 shows the interaction length of the ATLAS calorimeters. The passive material layers use high-density material that stops incoming particles. The active material provides signals that can be read out to sample the energy deposited on the detector material. The electromagnetic (EM) sampling calorimeter uses lead as the absorber material and liquid-argon (LAr) as the active medium and is divided into barrel ($|\eta| < 1.475$) and end-cap ($1.375 < |\eta| < 3.2$) regions. The hadron calorimetry uses scintillator tiles or LAr as the active medium and steel, copper, or tungsten as the absorber material.

Liquid Argon Calorimeters The Liquid Argon Calorimeter is designed to identify electrons, photons, and hadrons surrounding the Inner Detector subsystem and solenoid magnet. It used layers of metal (tungsten, copper, or lead) as passive absorbing material and liquid argon as active material. The charged particle passing through the LAr and absorber material will create secondary particles traveling in LAr and creating ionized charges that can be measured. In the barrel, the LAr calorimeter is designed to identify the electrons and photons. In the endcap regions, the LAr calorimeter consists of three parts: electromagnetic endcap (EMEC), hadronic endcap (HEC), and the LAr forward calorimeter (FCal).

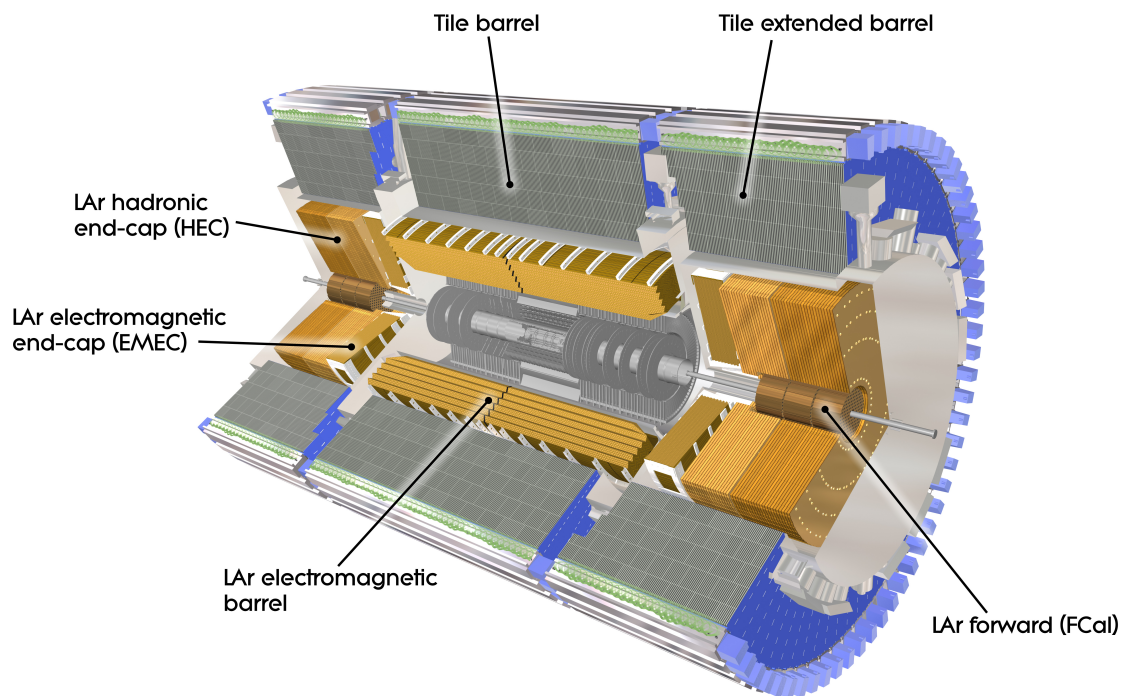


Figure 2.8: Cutaway diagram of ATLAS calorimeter detector. Visualization credit to Ref. [69]

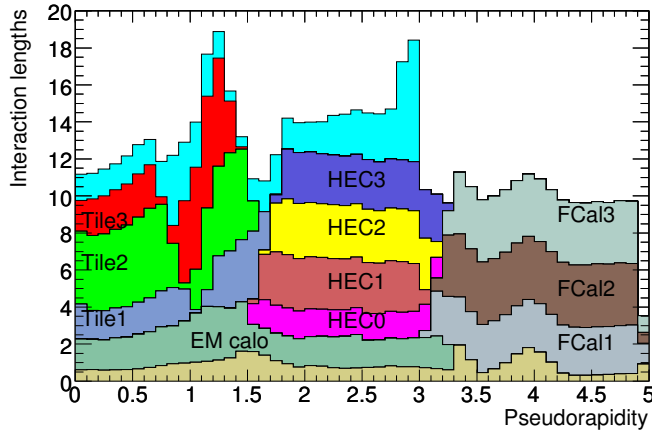


Figure 2.9: The cumulative amount of material, in units of interaction length, as a function of $|\eta|$, in front of the electromagnetic calorimeters, in each electromagnetic and hadronic compartment, and the total amount at the end of the active calorimetry. For completeness, the total amount of material in front of the first active layer of the muon spectrometer is also shown (up to $|\eta| < 3.0$). Plots from Ref. [63]

Hadronic Calorimeters The Tile Calorimeter is the next detector subsystem surrounding the LAr calorimeter. It measures hadronic particles that do not deposit all their energy in the electromagnetic calorimeter. The Tile Calorimeter is made of layers of steel and plastic scintillating tiles. Steel generates showers of new particles when hit by hadrons, producing photons when passing the plastic scintillator tiles. The photons are later covered into an electric current proportional to the original energy of the particle.

2.2.3 Muon Spectrometer

The muon spectrometer measures the deflection of muons with $|\eta| < 2.7$ using multiple layers of high-precision tracking chambers located in a toroidal field of approximately 0.5 T and 1 T in the central and end-cap regions of ATLAS, respectively. The muon spectrometer is also instrumented with separate trigger chambers covering $|\eta| < 2.4$. Figure 2.10 show the cutaway diagram for the muon spectrometer system.

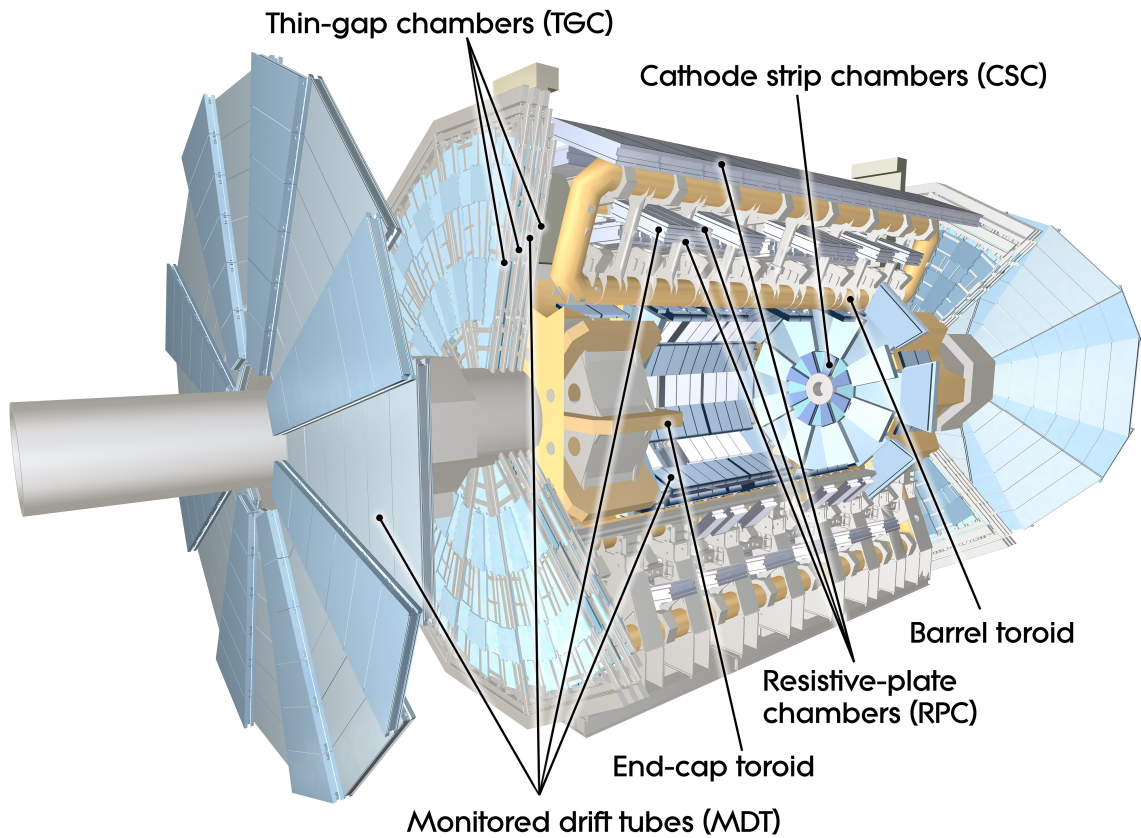


Figure 2.10: Cutaway diagram for muon sub-detector system. Visualization credit to Ref. [70]

Resistive-plate chambers (RPC) The RPC are gaseous detectors. It measures in both η and ϕ coordinates, providing up to six position measurements for a muon with a space-time resolution of $2\text{ cm} \times 2\text{ ns}$. The RPC is constructed with two parallel resistive electrodes with a gas gap of 2 mm. The gas mixture used is $\text{C}_2\text{H}_2\text{F}_4(94.7\%)\text{--}\text{C}_4\text{H}_{10}(5\%)\text{--}\text{SF}_6(0.3\%)$, which is a greenhouse gas that needs a substitute for future operation [71].

Monitored Drift Tube (MDT) The MDT detectors are composed of 3 cm wide aluminum tubes filled with a gas mixture. MDT chambers provide precision charged-particle tracking in the z-r coordinates. It can provide a single-hit spatial resolution of $81.7\text{ }\mu\text{m}$ [72].

Cathode strip chambers (CSC) The CSC are employed in the very forward pseudo-rapidity range, $2.0 < \eta < 2.7$ in the first end-cap layer. There are 32 CSCs in total, 16 chambers on either detector side. CSCs are multiwire proportional chambers. The chambers contain the same gas composition as the MDTs for precision tracking. However, it has a much faster drift time of $40\text{ }\mu\text{s}$, which is more suitable to deal with the high rate close to the beam pipe.

The Thin Gap Chamber (TGC) The TGC consists of parallel $30\text{ }\mu\text{m}$ wires in a gas mixture of $\text{CO}_2/\text{n-pentane}$. Pentane increases ionization when charged particles pass through. It's used to measure the curvature of the muon tracks in the end-cap regions before and after forward toroids. It is designed to provide a fast end-cap muon trigger to select collision events that are potentially interesting for physics quickly.

2.2.4 Magnetic systems

The magnetic system is the heart of the ATLAS detector, allowing the determination of the momenta of charged particles. The ATLAS used two superconducting magnet systems: solenoidal and toroidal [73]. Both are cooled down by liquid helium

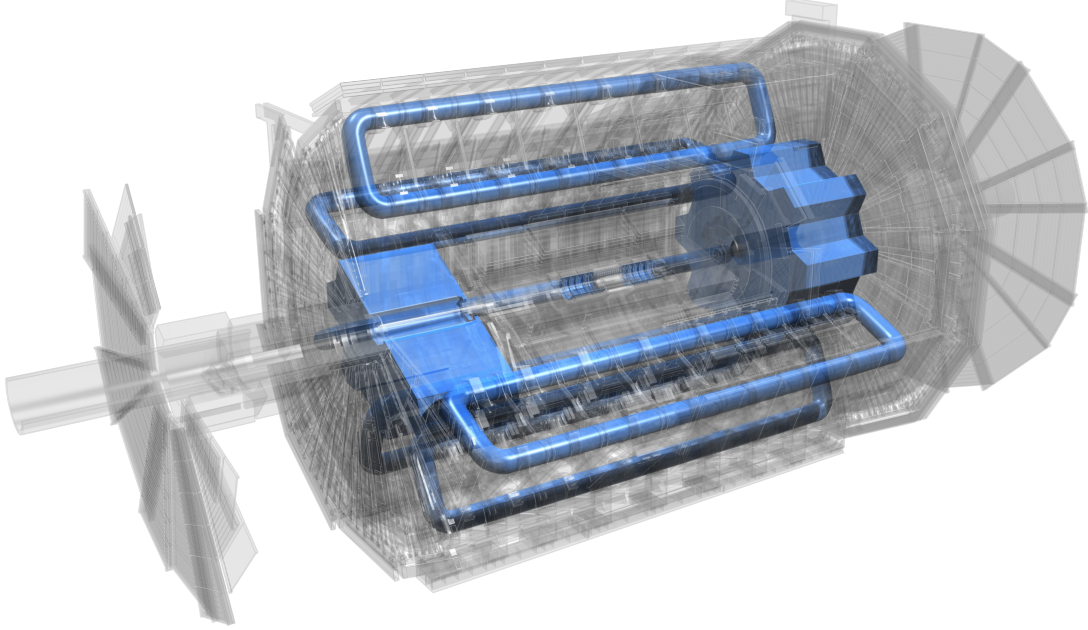


Figure 2.11: Schematics diagram for the ATLAS magnetic system highlighted in blue. Visualization credit to Ref. [75]

to around 4.5 K (-268°C) to provide a strong magnetic field to bend the trajectories of charged particles. Figure. 2.11 shows the visualization for ATLAS magnetic systems.

The central solenoid magnet is 5.6 m long and 2.56 m in diameter. It provides a 2 Tesla axial magnetic field, which encloses the inner detector for precision determination of charged particles momenta. To minimize interactions between the material in the magnet and the particles, the thickness is merely 4.5 cm, which results in 0.66 radiation length before reaching the electromagnetic calorimeter [74].

The forward toroids are composed of three toroid magnets, which primarily provide magnetic fields up to 3.5 Tesla for the determination of muon momenta. The barrel toroidal magnet surrounding the center of the ATLAS detector consists of eight air-core coils, with 25.3 m in length and an outer diameter of 20.1 m. Two end-cap toroids extend the magnetic field to particles leaving the detector close to the beam pipe. Each end cap is 10.7 m in outer diameter and 5.0 m in length.

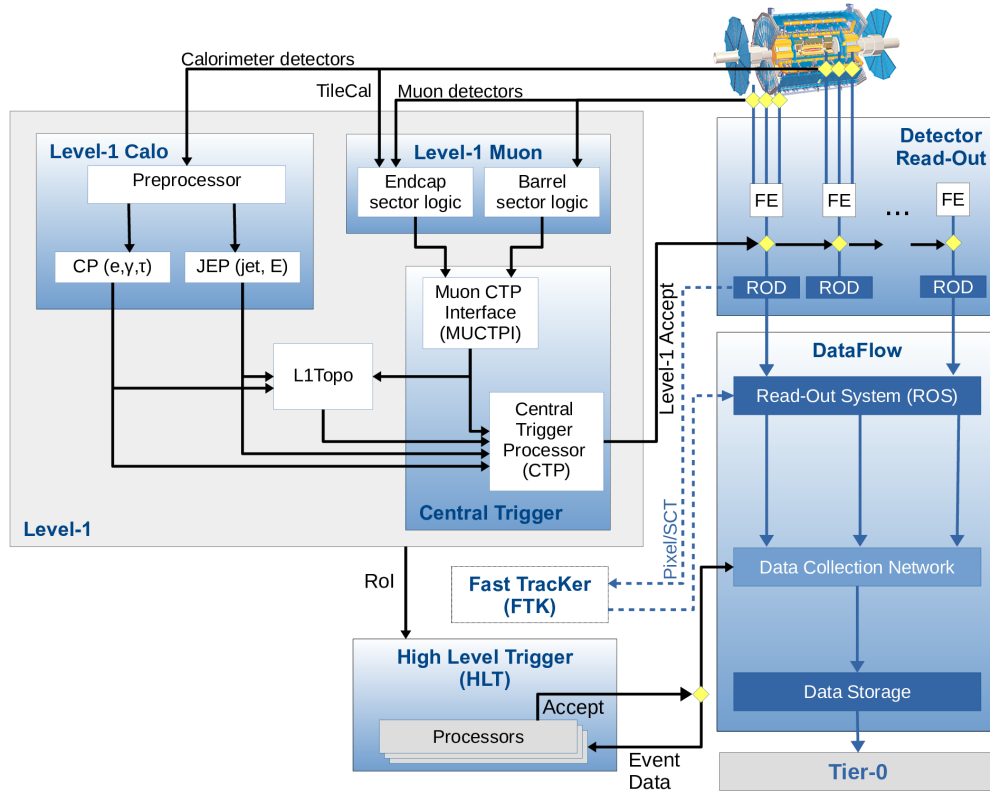


Figure 2.12: The ATLAS trigger systems, the detector read-out, and data flow. The Fast Tracker systems shown in the diagram were not installed and were canceled. Image from Ref [76]

2.2.5 Trigger System

The ATLAS has a two-level trigger system during the Run 2 data-taking, using custom hardware followed by a software-based trigger [76]. This allows the reduction of the events rate from an initial bunch crossing rate of 40 MHz to 1 kHz for offline storage. ATLAS sees around 1.7 billion collisions every second. However, it is unfeasible to record all the data, and only some of the collisions are interesting. Roughly 1500 individual event selections were included in a trigger menu to record physically interesting events and keep the rate within the bandwidth limitation, which the Data Acquisition system (DAQ) can handle. The Level-1 (L1) trigger used custom electronics to trigger coarse information from the calorimeter and muon detectors to

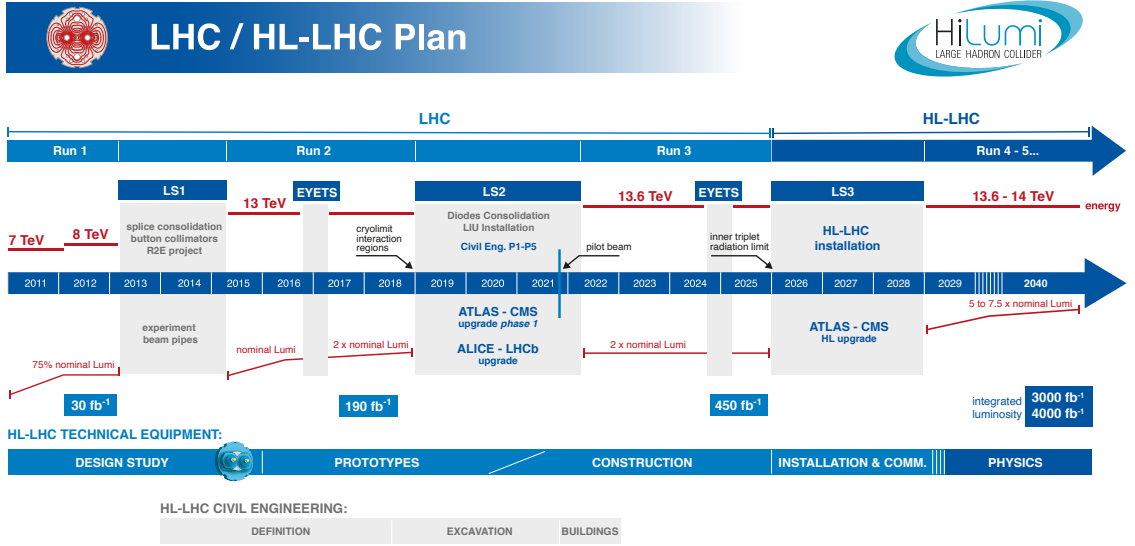


Figure 2.13: LHC/HL-LHC Plan updated in February 2022

achieve such low latency. The latency for L1 to accept an event must be within the 2.5 μ s. This is followed by the next software-based high-level trigger (HLT). After the trigger decision from the L1 is confirmed, the events are buffered in the Read-Out System and further processed by the HLT. The HLT is a computer farm that applies more sophisticated algorithms to make the final decision. Once an event is selected, it will be read and written to the storage for offline physics analysis. Otherwise, it is lost permanently.

2.3 High-Luminosity LHC and ATLAS detector upgrade

The High Luminosity Large Hadron Collider (HL-LHC) is an upgrade of the LHC. The timeline is shown in Figure 2.13. The goal is to enable the machine to deliver a factor of 5 to 7.5 luminosities than the LHC archived with the potential to reach the initially designed center of energy, $\sqrt{s} = 14$ GeV. At the end of its operation, it's expected to reach a total integrated luminosity of 3000 to 4000 fb⁻¹.

Extensive activity to upgrade the machine is ongoing, such as new superconducting magnets and compact superconducting crab cavities for beam rotation [77]. The HL-LHC is expected to start operation in 2029. The upgrade will allow the HL-LHC to produce at least 15 million Higgs bosons annually. This will allow us to have more statistics to probe rare processes and exotic Higgs decay not predicted by the SM. However, the increase in instantaneous luminosity will also increase the number of interactions per bunch crossing. This poses significant challenges to the current ATLAS detector and data acquisition systems. With the ultimate configuration of HL-LHC, the instantaneous luminosity is leveled at $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. This is equivalent to around a pile-up of up to 200 inelastic collisions per bunch crossing. As a result, several upgrades are planned for the ATLAS detector during the Long Shutdown 3 (LS3) to cope with the increase in expected instantaneous luminosity.

Muon chambers Significant upgrades are carried out in the detector [78] and on-detector electronics [79]. The muon system coverage is improved by an additional RPC layer in the inner station and new SMDT chambers. The New Small Wheel detectors are installed in the forward regions using new gaseous detector technologies: Micromegas (MM) and small-strip thin-gap chambers (sTGC). The on-detector electronics for the RPCs, TGCs, and MDTs will be replaced. The MDTs will be used for the first time for the trigger to improve the p_T resolution with its precision tracking of muon curvature. A more detailed description of this work is given in Section 2.3.1

Inner Tracker Detector The ATLAS Inner Tracker (ITk) is a new all-silicon-based inner detector that will replace the current inner detector, including the TRT [80]. The ITk will provide better impact parameter resolution to identify hadrons containing b- and c-quarks. Furthermore, The new detector will extend to pseudorapidity coverage in the forward region from $|\eta| < 2.5$ to $|\eta| < 4$, increasing the acceptance of

important physics processes in forward regions, such as vector boson fusion production of Higgs boson.

High Granularity Timing Detector The High-Granularity Timing Detector (HGTD) is a new detector that can provide a time resolution of 30 to 50 ps [81]. The detector comprises two double-side silicon layers on both sides of the ATLAS detector, covering the pseudorapidity range of $2.4 < |\eta| < 4$. The precise timing of this detector allows us to disentangle pile-ups and achieve 4D tracking using time information.

Trigger and Data Acquisition system Several significant upgrades to the trigger and data acquisition allow an efficient selection of events. A new Level-0 hardware trigger system will be added to replace the existing trigger to reach a readout rate of 1 MHz and a maximum latency of 10 ms. An Event filter system using a large processor farm is used to further reduce the rate with a more sophisticated offline-line algorithm. The system is designed to accommodate an output rate of 10 kHz for permanent storage.

2.3.1 Upgrade for Level-0 MDT trigger processor

The Level-1 muon trigger system will be updated to a new Level-0 Muon trigger system for the HL-LHC [79]. This will allow the inclusion of data from the newly installed detector during the Phase-II upgrade and LS2. The expected muon spectrometer layout is shown in Figure 2.14.

The main goal of this upgrade is to increase the latencies and the rates of the trigger and readout system. Additionally, improving trigger efficiency is a crucial aspect of this upgrade. The reason can be seen in Figure 2.15, where the transverse momentum of the muon candidate selected by the Level-1 trigger using offline reconstruction is shown. As shown in the plot, a substantial contribution of fake muons is

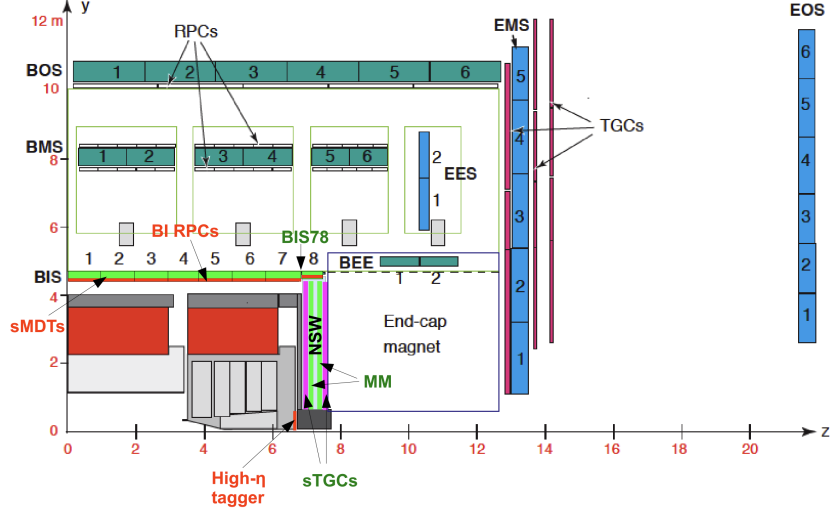


Figure 2.14: Cross-sectional view of the Phase-II ATLAS muon spectrometer. Those installed during the Phase-II upgrade and LS2 are labeled in red and green, respectively.

selected by the current muon trigger. Many selected muon candidates have p_T lower than the threshold due to limited p_T resolution at the trigger level.

The high rate from fake muons will significantly deteriorate the already limited trigger bandwidth for physics analysis during the HL-LHC era. Consequently, it is proposed to include the MDT chambers to provide precision tracking capabilities, which will sharpen the muon trigger efficiency turn-on curves as shown in Figure 2.16.

Dedicated algorithms are developed and implemented on custom off-detector hardware with FPGA (Field-programmable gate array). The MDT trigger algorithms can be divided into three stages as follow:

Hit extraction The selection of MDT hits relies on the information RPC/TGC candidates provided. This process involves filtering MDT hits based on their position and time information. The calculation of drift radius and position is also performed for subsequent stages.

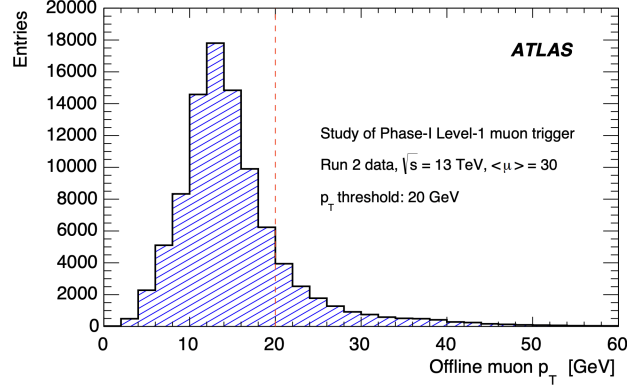


Figure 2.15: The transverse momentum of muon reconstructed with offline reconstruction algorithm for muon candidate select by the Level-1 trigger with 20 GeV threshold.

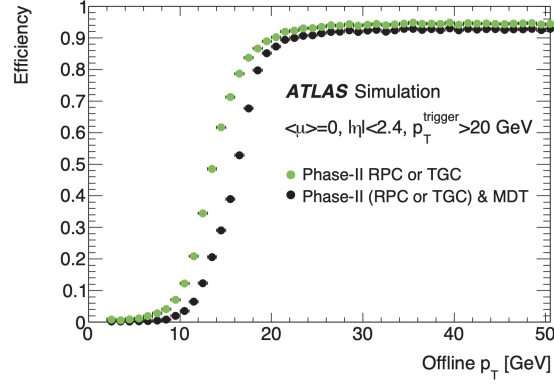


Figure 2.16: The trigger efficiency in $|\eta| < 2.4$ as a function of p_T measure by offline reconstruction.

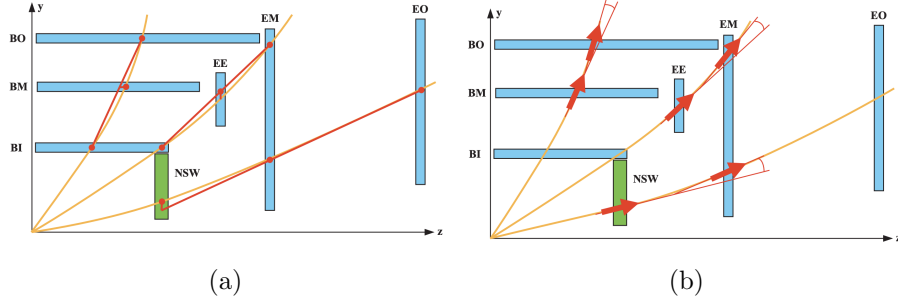


Figure 2.17: Schematic drawing of two momentum estimation methods using MDT track segments. (a) Schematic of the polar angle difference between the two segments. (b) Schematic of the concept of the distance between the position of a segment and the position extrapolated from two segments. For $1.05 < |\eta| < 1.3$, the endcap inner station is assumed alternatively to the barrel inner station.

Segment Finding Following the previous stage, using the MDT hits, the segment-finding algorithm is employed to identify track segments at each station. Two algorithms have been proposed: the Legendre Transform Segment Finder (LSF) and the Compact Segment Finder (CSF).

p_T estimation The track segments are connected across stations to determine the momentum of the muon candidate track. This allows for the estimation of the transverse momentum (p_T).

For this project, I contribute to the simulation and performance studies to support firmware design. The focus is on the hit extraction and DAQ functionalities. In particular, the task involves defining the hit extraction inputs and formats for playback testing. The VHSIC Hardware Description Language (VHDL) is used to design the L0MDT trigger processor on FPGA. A validation framework is developed to verify the VHDL implementation by comparing the VHDL and software implementation results. To simulate the data stream from the detector systems, input files are generated by concatenating the simulated events and mimicking the bunch crossing data stream.

As mentioned above, hit extraction is the first step in the trigger algorithms. The algorithms need to select and calibrate MDT hits under very tight latency constraints and limited resources that could fit in the FPGA. This is achieved using various LUTs to avoid complicated algorithms and reduce required resources and clock times.

CHAPTER 3

PHYSICS OBJECT RECONSTRUCTION

This chapter provides a summary of the object reconstruction algorithm used in this thesis. In a particle collider experiment, events recorded by the detector need further processing before being used in physics analysis. After being written to permanent storage, the offline reconstruction software will process the events accepted by triggers. A serial algorithm is applied to reconstruct physics objects from the detector signals. Figure 3.1 illustrates how particles interact with the detector subsystems. The reconstruction and selection of electrons and muons are described in Sections 3.1 and 3.2, respectively. In Section 3.3, the jet reconstruction algorithm is described. The flavor tagging algorithm applied on the jet to identify jets initiated from heavy flavor quarks is also discussed. Furthermore, the search also uses a soft secondary vertex reconstruction technique to recover soft b-hadrons not reconstructed as jets, as described in Section 3.4. Lastly, the missing transverse momentum is defined in Section 3.5.

3.1 Electrons

Electron candidates are reconstructed from energy deposits (superclusters) in the electromagnetic calorimeter and a matched track in the inner detector [83]. Candidates are selected to satisfy $p_T > 10\text{GeV}$ and with $|\eta| < 2.47$, excluding the calorimeter transition region $1.37 < |\eta| < 1.52$. Electrons must pass the **TightLH** likelihood-based identification criterion. This criterion aims to discriminate genuine electrons from prompt-isolated electrons from energy deposits from hadronic jets, con-

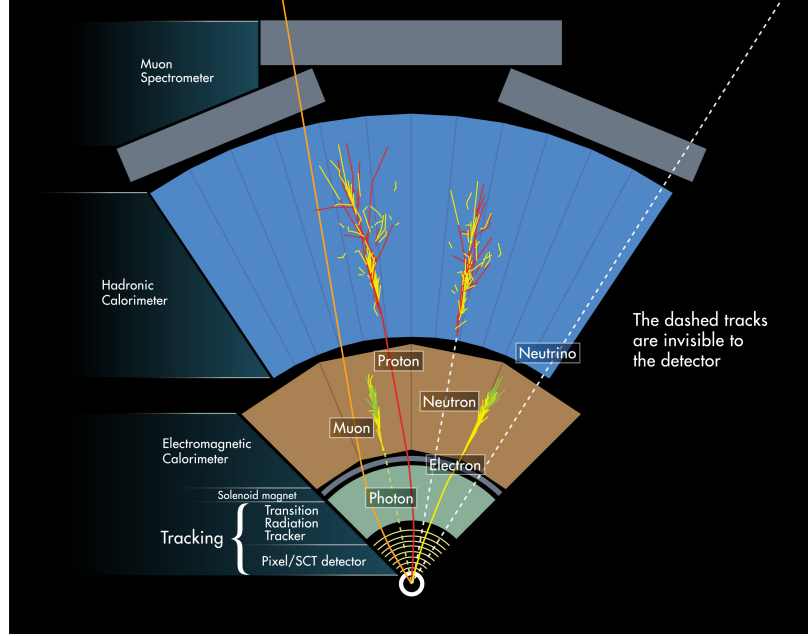


Figure 3.1: Illustration on how ATLAS detector detects particles. Illustration credit to Ref [82]

verted photons, and the decays of heavy-flavor hadrons. Electron candidates are also required to satisfy the working point defined by PromptLeptonTagger to reject fake electrons from heavy-flavor decays or light hadrons misidentified as electrons. The PLVLoose isolation working point is also required to suppress fake electrons further. Lastly, Electron candidates are required to have $|z_0 \sin \theta| < 0.5 \text{ mm}$ and $|\frac{d_0}{\sigma(d_0)}| < 5$, where the longitudinal and transverse impact parameters are computed with respect to the beam-line.

3.2 Muons

As described in the Section 2.2.3, the muon spectrometer is designed to provide a precision measurement of muons with a pseudorapidity coverage up to $|\eta| = 2.7$. The muon reconstruction algorithm uses primary measurements from the muon spectrometer and the ID, which provide coverage only up to $|\eta| = 2.5$. Other detectors are also used but only serve as complimentary [84].

Muon reconstruction starts by reconstructing the muon track independently in the inner detector and the muon spectrometer. The reconstruction is similar to that usually used for charged particle reconstruction in the inner detector. For the muon spectrometer, the reconstruction starts by searching for hit patterns inside each muon chamber to form muon track segments. These muon tracks are later combined with the information from the inner detector for the combined reconstruction. There are five muon types defined depending on the information used.

- Combined (CB) muon: The algorithm uses tracks reconstructed independently from the inner detector and muon spectrometer as input. A global refit is performed, which could add or remove muon spectrometer hits to improve the fit quality. This is the primary method of muon reconstruction for physics analysis in ATLAS.
- Inside-out combined (IO) muons: The IO muons are reconstructed using a complementary inside-out algorithm. It extrapolates the ID track to find at least three aligned muon spectrometer hits for a combined track fit. This algorithm does not rely on an independently reconstructed muon spectrometer track and recovers some efficiency loss in CB muons. For example, in regions of limited muon spectrometer coverage and for low- p_T muons, which may not reach the middle station of the muon spectrometer.
- Segment-tagged (ST) muons: The algorithm extrapolates the inner detector track to the muon spectrometer and associates it to at least one track segment in the precision chambers (MDT or CSC). This method captures cases with low- p_T muons or those muons that fall in regions with reduced muon spectrometer acceptance.
- Calorimeter-tagged (CT) muons: The CT muons are identified by a track that matches the calorimeter energy deposit. Only those energy deposits compatible

with a minimum-ionizing particle are considered. The purity of this method is very low, but it recovers muons in the muon spectrometer corresponding to cables and services.

- Muon-spectrometer extrapolated (ME) muons: The ME muons are mainly used to recover acceptance in regions not covered by the ID ($2.5 < |\eta| < 2.7$). The muons are reconstructed only by muon spectrometer track and require it to extrapolate back from the interaction point.

For the analysis performed in this thesis, muons candidates are required to satisfy $p_T > 10$ GeV and $|\eta| < 2.5$. Further requirements must be applied to reconstructed muon to suppress fake muon from the background, such as pion and kaon decay. Muons are required to pass the Medium quality requirements. Additional vertex association criteria are imposed on muon tracks to reject muon from hadron decays and pile-up. The absolute value of muon d_0 significance must be less than 3, and the value of $|z_0 \sin \theta|$ must be less than 0.5 mm. An isolation criteria `PLVLoose` is required to further suppress non-prompt muons from hadron decay. The detailed definition of identification and isolation criteria can be found in Ref [84]

3.3 Jets

Jet reconstruction

A jet collection can be defined by applying the jet algorithm to a collection of four-vector objects. The inputs can be stable particles from the MC generator, and the resulting jets are usually referred to as truth jets. They can also be charged particle tracks, calorimeter energy deposits, or an algorithmic combination of the two. The reconstruction of the hadronic jet can be improved by using both inner detector and calorimeter systems. This is done by the particle flow (PFlow) algorithm [85]. The particle flow method combines measurements from the inner detector and calorimeter

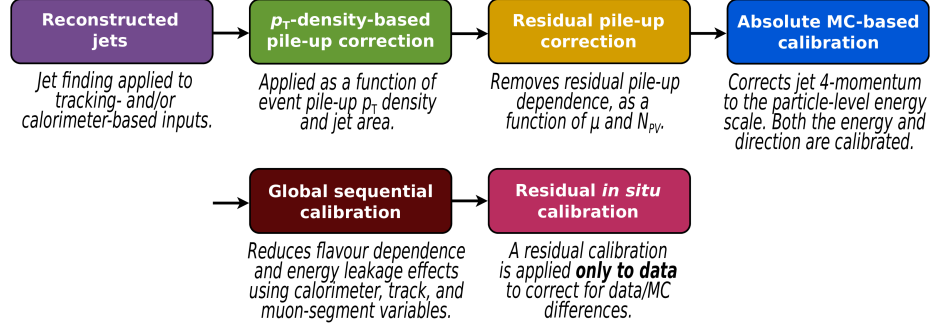


Figure 3.2: Stages of jet energy scale calibrations. Flowchart taken from Ref [86]

systems to form Particle Flow Objects for jet reconstruction. In this procedure, the energy deposited in the calorimeter by charged particles is removed from the topocluster measurements and replaced by the momenta of tracks matched to those topoclusters.

The jets used in this thesis are reconstructed by clustering Particle Flow Objects [85, 86] with the anti- k_t algorithm [87] with a radius parameter of $R = 0.4$ with a four-momentum recombination scheme implemented in the FastJet package [88].

Jet calibration

The jet energy is corrected to the particle level by applying a jet energy scale calibration derived from 13 TeV data and simulation [89]. The calibration procedure is briefly summarized in Figure 3.2. The jet energy scale calibration aims to restore the jet energy to that of jets reconstructed at the particle level.

This thesis uses new corrections on the jet energy delivered centrally by the ATLAS collaboration to cover the low jet p_T range of $15 \text{ GeV} \leq p_T \leq 20 \text{ GeV}$. Additional studies are performed for the Jet Vertex Tagger (JVT) and b-tagging efficiency measurements down to this transverse momentum range. The jets used in this thesis is required to have $p_T > 15 \text{ GeV}$, $|\eta| < 2.5$, and satisfy the LooseBad jet cleaning working point criteria. They are also required to pass a Tight pileup rejection based on the JVT score, w_{JVT} : for jet $p_T \in [15, 60] \text{ GeV}$ and $|\eta| < 2.4$, $w_{\text{JVT}} > 0.5$ is

required [90]. The MC-to-data corrections for the JVT and b-tagging efficiencies are consistent with one, and the corresponding uncertainties are included in the analysis.

Overlap removal

To prevent double-counting of electron energy deposits as jets, the closest jet within $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta \phi)^2} = 0.2$ of a selected electron is removed. If the nearest jet surviving that selection is within $\Delta R_y = 0.4$ of the electron, the electron is discarded. Muons are removed if they are separated from the nearest jet by $\Delta R_y < 0.4$, which reduces the background from heavy-flavour decays inside jets. However, if this jet has fewer than three associated tracks, the muon is kept, and the jet is removed instead. In that case, this avoids an inefficiency for high-energy muons undergoing significant energy loss in the calorimeter.

Flavor tagging

Identifying jets containing b- and c-hadrons against those from neither of them is one of the essential parts of the studies performed in this thesis. Conventionally, the jets containing at least one or more b- and c-hadrons are called heavy-flavor jets. Conversely, those that do not contain b- and c-hadrons are denoted as light jets. Several algorithms aim to identify heavy-flavor jets and are referred to as flavor-tagging algorithms. These algorithms exploit the lifetime, high mass, and high decay multiplicity of b-hadrons to distinguish them from their counterparts. The lifetime of b-hadrons is typically around 1.5 ps which gives rise to a decay vertex displaced from the hard-scatter interaction point.

In this thesis, two algorithms are applied sequentially to categorize jets. First, the double-b tagging algorithm, DeXTer, is applied on the jet to identify boosted double-b quark decays from the light scalar. The details of this algorithm are described in Chapter 5. If the jet is not tagged by DeXTer, the standard b-tagging algorithm,

DL1r, is applied. The algorithms aim to identify $a \rightarrow bb$ decays resolved into two separate b-jets. The detail of this algorithm are given below.

DL1r b-tagging b -jets are identified using the DL1r algorithm [91], a high-level tagger based on a deep NN. The NN is trained on both MC $t\bar{t}$ and Z' events and uses the results of specialized low-level taggers (based on the identification of the interaction point and track parameters and secondary vertices) to obtain the probability for each jet to be a b -, c - or light-jet (p_b , p_c and p_{light} , respectively). To reduce the output dimensionality, these probabilities are combined in one variable as in Eq. 3.1:

$$D_{\text{DL1r}} = \ln \left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\text{light}}} \right), \quad (3.1)$$

where f_c is the charm fraction and the threshold on the DL1r itself are variable parameters used to define WPs with different b -jet efficiency and c -/light-jet rejection rates. In the analysis, an 85% WP is used, using the pseudo-continuous calibrations and corresponding to an 85% b -jet efficiency [92–94].

The corrections to the simulated DL1r tagging efficiency delivered centrally do not cover the low jet p_T range of $15 \text{ GeV} \leq p_T \leq 20 \text{ GeV}$. Hence, a separate calibration of the 85% WP was carried out. The b-tag efficiency correction factors derived for both b- and light-labeled jets agree with one within uncertainty. For c-labeled jets, no additional correction scale factors are derived. The uncertainties from the lowest bin from the nominal c-jet mistagging calibration are inflated by an additional 5% and used.

3.4 Soft secondary vertex

The use of soft Secondary Vertices (SV) is motivated by the case where the low- p_T b hadron is not reconstructed as a jet. In that case, the Track-Cluster-based Low- p_T Vertex Tagger (TC-LVT) [95] algorithm is applied. It relies on reconstructed inner

detector tracks to find secondary vertices that can originate from the decay of b-hadrons. The algorithm starts by finding a set of seed tracks that have $p_T > 1.5 \text{ GeV}$, and transverse impact parameter significance, $\frac{d_0}{\sigma(d_0)} > 0.5$. Clusters are built around these seed tracks by adding additional displaced tracks within an angular distance $\Delta R_{\text{seed}}^{\text{track}} < 0.75$ and seed-to-track distance $r_{\text{seed}}^{\text{track}} < 0.25 \text{ mm}$. The Single Secondary Vertex Finder (SSVF) algorithm [96] is run on these clusters to produce candidate vertices. The four-momentum of each vertex is defined as the invariant sum of the associated track momenta. The analysis selects and uses only the secondary vertex candidates with $p_T > 3 \text{ GeV}$ and a mass $m_{vtx} > 600 \text{ MeV}$. The selections on seed tracks, cluster tracks, and the resulting vertices applied here correspond to the Loose working point defined in [95] and are summarised in Table 3.1.

In the overlap removal procedure, any soft SV within $\Delta R < 0.2$ of a lepton, within $\Delta R < 0.6$ of a jet, or within $\Delta R < 0.9$ of an DEXTER-tagged jet is removed. Any two vertices within an angular distance of $\Delta R < 0.3$ are merged. This is motivated by the assumption that most of these cases are not vertices from two b-hadron decays but more likely a secondary or tertiary decay in the same hadron decay chain. For each vertex, the decay lengths are defined relative to the primary vertex. L_{xy} is the decay length projected on the transverse plane and L_{3d} is the magnitude of decay length relative to the primary vertex. The four-vector of the merged vertices are added, and the resulting vertex retains the decay length values of the merged vertex with the smallest L_{xy} .

The efficiency of the TC-LVT tagger is measured in the data, which determines correction factors for the simulated efficiency and fake rate of the tagger.

3.5 Missing Transverse Energy

The missing transverse momentum E_T^{miss} in the event is defined as the magnitude of the negative vector sum of p_T for all selected and calibrated physics objects in the

Seed track		Cluster tracks		Vertex	
$\frac{d_0}{\sigma(d_0)}$	> 0.5	$\frac{d_0}{\sigma(d_0)}$	> 1.5	m_{vtx}	$> 600 \text{ MeV}$
p_T	$> 1.5 \text{ GeV}$	ΔR_{seed}^{track}	< 0.75	p_T^{vtx}	$> 3 \text{ GeV}$
		r_{seed}^{track}	$< 0.25 \text{ mm}$		

Table 3.1: Selection applied for the vertex reconstruction of the TC-LVT **Loose** working point.

event, with a soft term added to account for soft energy that is not associated with any of the selected objects [97].

$$\sum E_T^{\text{miss}} = - \sum_{i=\mu,e,\gamma,\tau} p_T^i - \sum p_T^{\text{jets}} - \sum p_T^{\text{soft}}. \quad (3.2)$$

This soft term is calculated from inner detector tracks matched to the primary vertex to make it more resilient to pile-up contaminations [98, 99].

Two different transverse momenta are defined and used in the analysis. The variable H_T^{had} is defined as the sum of the jet transverse momentum p_T . In addition, the variable H_T also includes the transverse momentum of leptons.

CHAPTER 4

DATA AND MONTE CARLO SIMULATION

This chapter summarizes data and Monte Carlo simulations used in this thesis. I focus on the generic information relevant to the studies performed in the following chapters. Details only applicable to specific studies are mentioned in the corresponding chapter. The collision data are described in Section 4.1 along with the trigger used. Section 4.2 detail the Monte-Carlo simulation used in the analysis, including both $H \rightarrow aa/a_1a_2 \rightarrow (b\bar{b})(b\bar{b})$ signals and SM backgrounds.

4.1 Data

As described in Chapter 2, a tremendous amount of work is needed to accelerate and collide protons to study the structure of subatomic particles. The search described in the following chapters uses data collected from the LHC at $\sqrt{s} = 13$ TeV center-of-mass energy recorded with the ATLAS detector between 2015 and 2018. This dataset corresponds to an integrated luminosity of 140 fb^{-1} [100]. Only events recorded with a single electron [101] or single muon trigger [102] under stable beam conditions are considered. This requires all detector subsystems to be operated under nominal conditions. Single lepton triggers with different p_T thresholds are combined in a logical OR to increase the overall efficiency. Table 4.1 lists all the trigger chains used for each of the four years of data-taking.

The name of the HLT trigger corresponds to the p_T threshold and isolation requirement. The single-muon trigger chain HLT_mu26_ivarmedium selects an isolated muon with $p_T > 26$ GeV with medium isolation working point.

Trigger name	2015	2016	2017	2018
Electron trigger				
HLT_e24_lhmedium_L1EM20VH	✓			
HLT_e26_lhtight_nod0_ivarloose		✓	✓	✓
HLT_e60_lhmedium	✓			
HLT_e60_lhmedium_nod0		✓	✓	✓
HLT_e120_lhloose	✓			
HLT_e140_lhloose_nod0		✓	✓	✓
Muon trigger				
HLT_mu20_iloose_L1MU15	✓			
HLT_mu26_ivarmedium		✓	✓	✓
HLT_mu50	✓	✓	✓	✓

Table 4.1: The un-prescaled single-lepton trigger used for 2015, 2016, 2017, and 2018 data selection. A logical OR is used for a given year

The triggers with the lower p_T threshold include isolation requirements on the candidate lepton, resulting in inefficiencies at high p_T . These are recovered by the triggers with higher p_T threshold that do not apply isolation requirements. The trigger efficiencies for the combined single electron and combined single muon triggers are above 90%.

4.2 Monte Carlo Simulation

All simulated samples are processed using either GEANT4 description [103] or ATLFast2 of the ATLAS detector. The effect of multiple interactions in the same and neighboring bunch crossings (pileup) is modeled by overlaying the simulated hard-scattering event with inelastic proton-proton (pp) events generated with PYTHIA 8.186 [104] using the NNPDF 2.3LO set of parton distribution functions (PDF) [105] and the A3 set of tuned parameters [106].

4.2.1 Simulation of exotic Higgs boson decay $H \rightarrow aa/a_1 a_2 \rightarrow b\bar{b}b\bar{b}$

Signal samples of associated Higgs boson production with a Z boson, $pp \rightarrow ZH$ are generated with POWHEG using the NNPDF3.0NLO [107] parton distribution

function (PDF). A Higgs boson mass of $m_H = 125$ GeV is assumed. The $pp \rightarrow ZH$ process includes $gg \rightarrow ZH$, which contributes around 10% of the total cross-section. Two separate signal samples are produced for the $qq \rightarrow ZH$ and the $gg \rightarrow ZH$ production processes. The events are interfaced with the PYTHIA 8.244 generator that models the decay of the Higgs boson into a pair of new light scalars with the masses shown in Table 4.2. These new scalars are then decayed into a pair of b -quarks. Events are subsequently showered and hadronized with PYTHIA 8.244 using the AZNLO tune [108] to model the underlying event. Events from minimum-bias interactions are simulated with the PYTHIA 8.244 generators. They are overlaid on the simulated signal events according to the luminosity profile of the recorded data. The contributions from these pile-up interactions are modeled within the same bunch crossing as the hard-scattering process and in neighboring bunch crossings. Finally, the generated samples are processed through a simulation of the detector geometry and response using ATLFast2. All samples are processed through the same reconstruction software as the data. Simulated events are corrected so that the object identification efficiencies, energy scales, and energy resolutions match those determined from data control samples. Table 4.2 lists all mass hypotheses considered in this thesis.

Signal kinematics

This section describes studies of the kinematics of the signal samples generated at the truth level based on the information from the hard interaction before showering.

Signal hypothesis	Parameters	Mass (GeV)
$H \rightarrow aa \rightarrow b\bar{b}b\bar{b}$	m_a	12, 16, 20, 25, 30, 50, 60
$H \rightarrow a_1 a_2 \rightarrow b\bar{b}b\bar{b}$	(m_{a_1}, m_{a_2})	(20,30), (40,60), (50,70)

Table 4.2: List of mass hypotheses considered for this analysis.

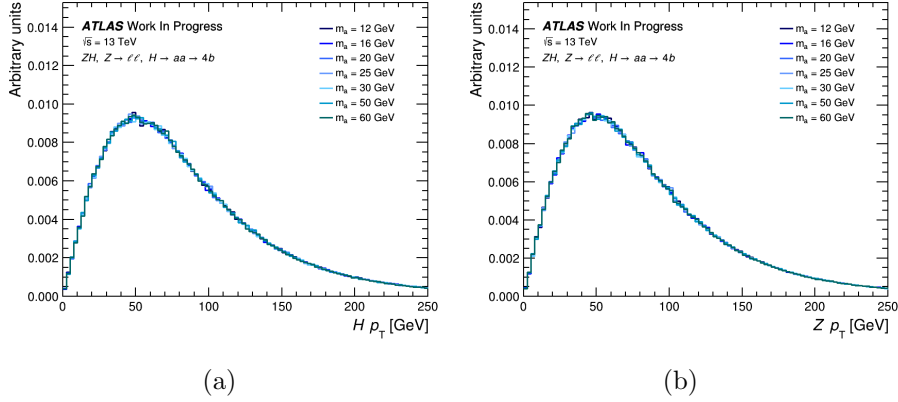


Figure 4.1: p_T distribution of the Higgs boson and the Z boson for the $ZH + ggZH$ production mode using POWHEG under the $H \rightarrow aa \rightarrow 4b$ signal hypothesis. The $H \rightarrow a_1 a_2 \rightarrow 4b$ signal is expected to have the same kinetics

Figures 4.1 show the truth-level Higgs and Z boson p_T distribution, for $pp \rightarrow ZH$ production at $\sqrt{s} = 13$ TeV and all the signal hypotheses. The p_T of the Higgs boson corresponds to that of the Z boson since these particles are recoiling off each other. In addition, the kinematics of H and Z are independent of the mass of the a boson.

Figures 4.2 show kinematic distributions for the a -boson for the $ZH + ggZH$ production under the different signal hypothesis. The angular separations between the heavier a bosons tend to be smaller than between the lighter ones since the p_T over mass ratio is lower. In the case of the separation between b -quarks originating from the same parent, for smaller a masses, the quarks are more likely to be boosted, although this may not be the case when the second a -boson is much heavier than the first. This can be seen in Figure 4.4. If an event has only one of the $a \rightarrow b\bar{b}$ merged into one jet, the other is resolved into two jets or one jet and a secondary vertex. These are called semi-merged channels.

The p_T of the b -quarks originating from a decays is displayed in Figure 4.6. From this set of histograms, it is observed how in many cases, these b -quarks have a very low p_T , below the typical jet calibration in ATLAS $p_T > 20$ GeV. This motivates

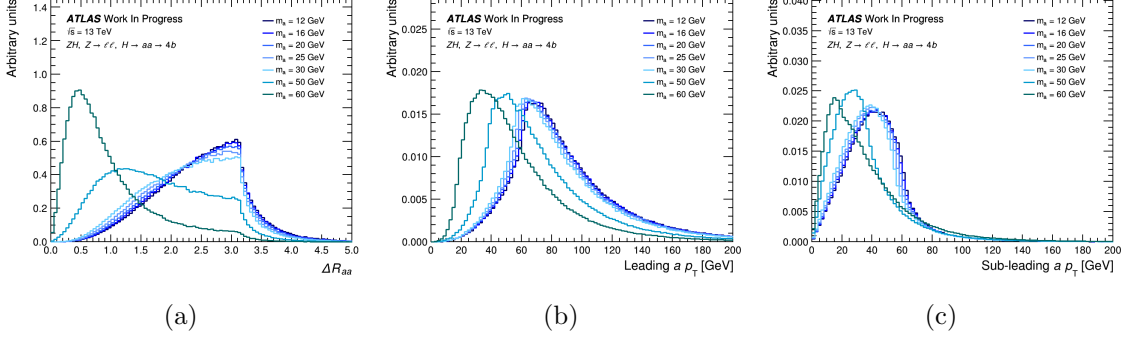


Figure 4.2: Kinematic distributions for the a -bosons under the $H \rightarrow aa \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode. The figures (a) show the angular separation between particles, (b) the p_T distribution of the leading a , and (c) the p_T distribution of the sub-leading a .

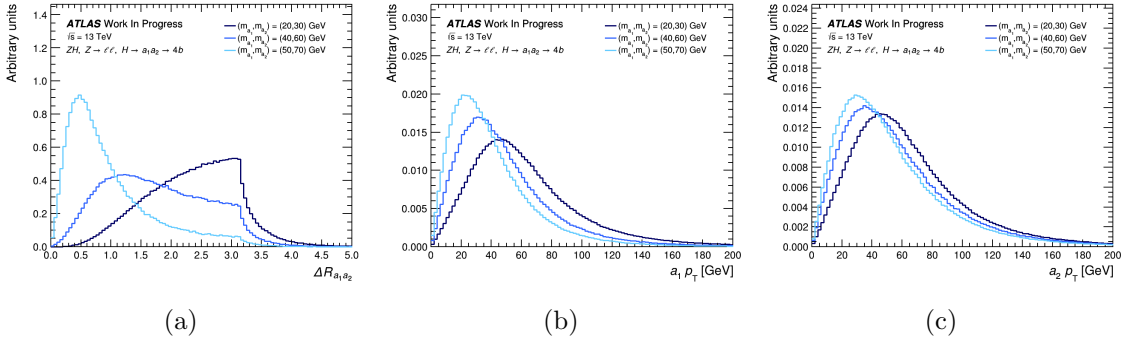


Figure 4.3: Kinematic distributions for the a -bosons under the $H \rightarrow a_1 a_2 \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode. The figures show (a) the angular separation between particles, (b) the p_T distribution of the a_1 , (c) and the p_T distribution of the a_2 .

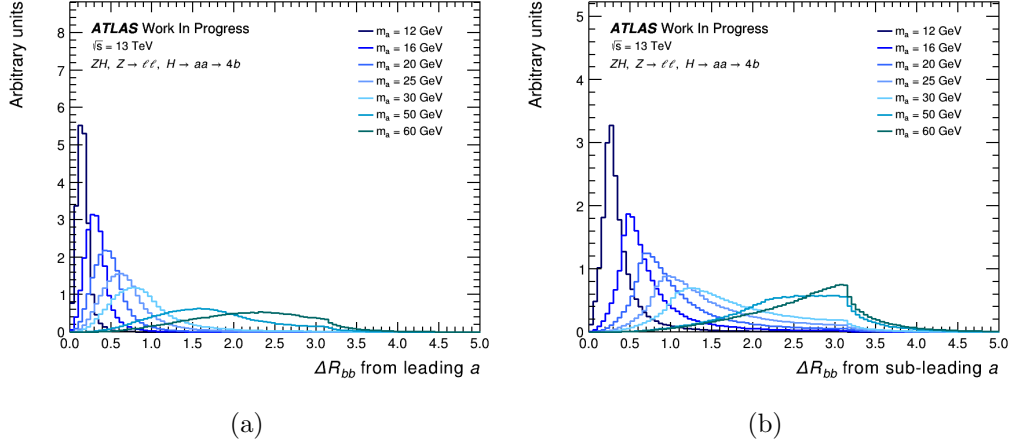


Figure 4.4: Angular separation of b -pairs from the same parent under the $H \rightarrow aa \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode.

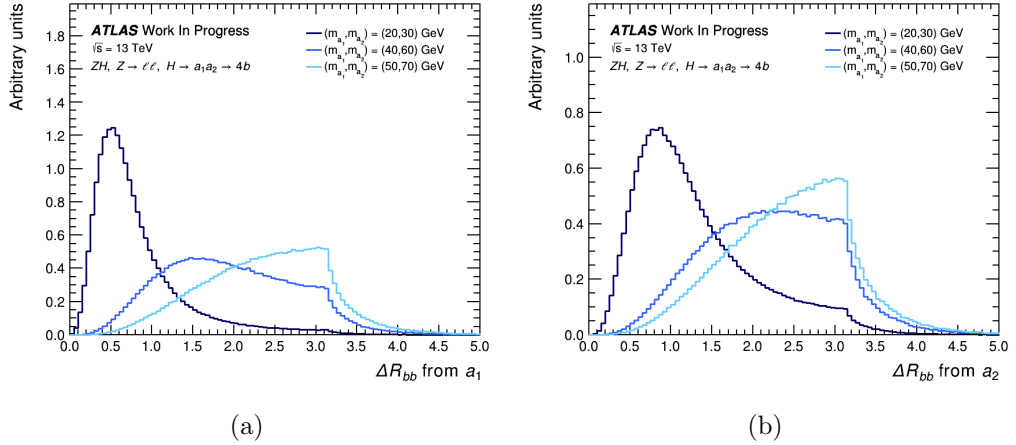


Figure 4.5: Angular separation of b -pairs from the same parent under the $H \rightarrow a_1 a_2 \rightarrow 4b$ signal hypothesis for the $ZH + ggZH$ production mode.

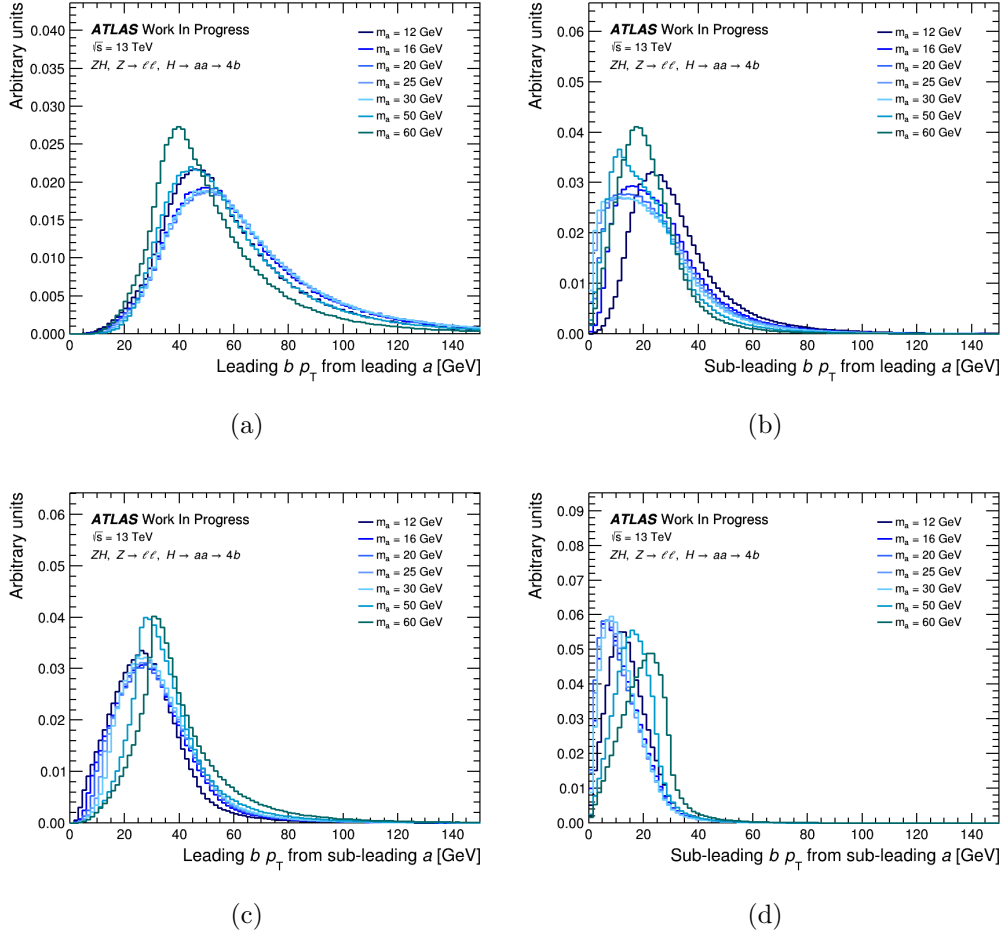
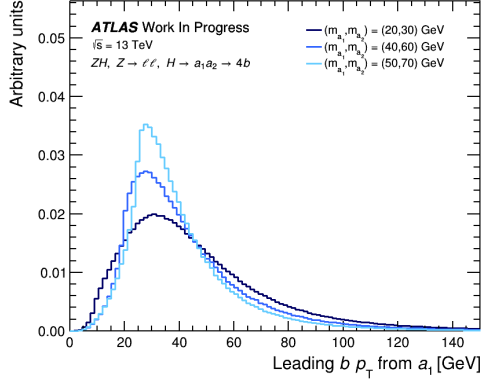


Figure 4.6: p_T distributions for the b -quarks coming from the $H \rightarrow aa \rightarrow 4b$ decay for the $ZH + ggZH$ production mode.

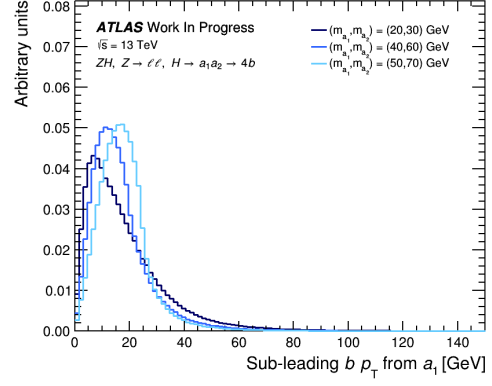
extending jet reconstruction and calibration down to $p_T > 15$ GeV and adding the soft secondary vertex reconstruction to reduce acceptance loss.

4.2.2 Truth categorization of simulated events

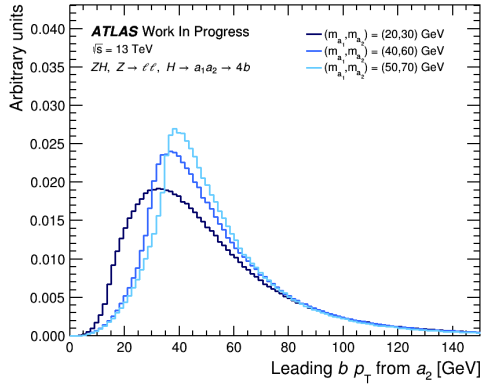
Both the $t\bar{t}$ +jets and Z +jets samples are categorized depending on the flavor of the partons that are matched to particle jets that do not originate from the decay of the $t\bar{t}$ or Z system, respectively. This follows a similar procedure as described in Ref [109]. These generator-level particle jets are reconstructed from stable particles with the anti- k_t algorithm with a radius parameter $R = 0.4$ and are required to have



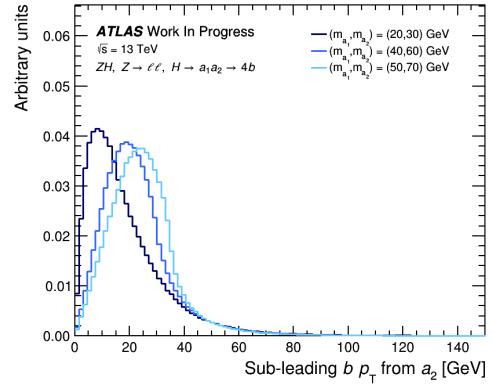
(a)



(b)



(c)



(d)

Figure 4.7: p_T distributions for the b -quarks coming from the $H \rightarrow a_1 a_2 \rightarrow 4b$ decay for the $ZH + ggZH$ production mode.

$p_T > 15$ GeV and $|\eta| < 2.5$. The flavor of a jet is determined by counting the number of b - or c -hadrons within $\Delta R < 0.4$ of the jet axis. The naming convention for these jets is the following:

- b : the jet is matched to a b -hadron.
- B : the jet is matched to at least two b -hadrons.
- c : the jet is matched to a c -hadron, but no b -hadron.
- C : the jet is matched to at least two c -hadrons, but no b -hadron.
- light: the jet is not matched to any b - or c -hadrons.

This allows to classify the $t\bar{t}$ +jets in four orthogonal categories:

- $t\bar{t}+\geq 1B$: events with at least one B jet and any number of b jets not originating from the decay of the $t\bar{t}$ system.
- $t\bar{t}+\geq 1b$: events with at least one b jet, and no B jets, not originating from the decay of the $t\bar{t}$ system.
- $t\bar{t}+\geq 1c$: events with at least one c or C jet, and no b or B jet, not originating from the decay of the $t\bar{t}$ system.
- $t\bar{t}$ +light: events not included in the other three categories.

where $t\bar{t}+\geq 1B$, $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ can also be referred to as $t\bar{t}$ +HF (heavy flavor).

The Z +jets backgrounds follow the same categories counting definitions.

As the contribution of $t\bar{t}+\geq 1c$ and $Z+\geq 1c$ events are small in the signal and control regions selected for the analysis, it is merged into the $t\bar{t}$ +light and Z +light category respectively in the following sections unless otherwise specified.

4.2.3 Simulation of $Z + \text{jets}$ background

The $Z + \text{jets}$ sample was simulated with the SHERPA 2.2.11 generator using next-to-leading-order (NLO) matrix elements for up to two partons in addition to the Z boson and LO matrix elements for up to five partons calculated with the Comix and OPENLOOPS libraries [110, 111]. Virtual electroweak NLO corrections were applied to the ME calculation. The NNPDF3.1NNLO set of PDFs was used. The MC sampling is biased in $\max(p_T^V, H_T)^2$ to enhance statistics in the tails of distributions.

An alternative sample is simulated with the MADGRAPH5_AMC@NLO 2.6.5 program to generate matrix elements for $V + 0, 1, 2$, and 3 additional partons in the final state to NLO accuracy. The ME calculation employed the NNPDF3.0NLO set of PDFs [107]. Events were interfaced to PYTHIA 8.240 [104] for modeling the parton shower, hadronization, and underlying event. The different jet multiplicities are merged using the FxFx prescription [112]. The A14 tune [113] of PYTHIA 8 was used with the NNPDF2.3LO PDF set [105].

4.2.4 Simulation of $t\bar{t}$ background

To enhance the statistics in the phase space relevant to this analysis, dedicated filtered samples were produced for all the $t\bar{t}$ MC samples described below. Based on categorization defined in Section 4.2.2, the simulated events in the inclusive sample are replaced with simulated events from the corresponding filtered samples to avoid duplication and enhanced MC statistics.

$t\bar{t}$ (5FS) samples A suite of MC samples with different accuracy in the Matrix Element (ME) generation and different Parton Showers (PSs) are used for the $t\bar{t}$ (+jets) process. These samples can be used for the nominal $t\bar{t}$ model, the modeling of systematic uncertainties, and cross-checks.

The production of $t\bar{t}$ events is modeled using the POWHEGBOX [114] v2 generator at the next-to-leading order (NLO) in QCD in the five flavor scheme (FS). It uses

PYTHIA 8.230 with similar settings to the POWHEGBOX+PYTHIA8 samples. The h_{damp} parameter is set to $3/4 \cdot (m_t + m_{\bar{t}}) = 1.5 m_{\text{top}}$ [115], and the functional form of the renormalization and factorization scale is set to the default scale $\sqrt{m_{\text{top}}^2 + p_{\text{T}}^2}$.

This sample is used as the nominal prediction for $t\bar{t}$ +light and $t\bar{t}+ \geq 1c$ events. The uncertainties due to ISR and FSR are estimated using weights in the ME and the PS. The impact of using a different parton shower and hadronization model was evaluated by comparing the nominal $t\bar{t}$ sample with another event sample produced with the POWHEG BOX v2 [114, 116–118] generator using the NNPDF3.0NLO [107] parton distribution function (PDF). Events in the latter sample were interfaced with HERWIG 7.13 [119, 120], using the HERWIG 7.1 default set of tuned parameters [120, 121] and the MMHT2014LO PDF set [122]. The decays of the bottom and charm hadrons were simulated using the EVTGEN 1.6.0 program [123].

$t\bar{t}b\bar{b}$ (4FS) samples A dedicated MC simulation is used instead for the top-quark pair production in association with a b -quark pair $t\bar{t}b\bar{b}$ [124]. The predictions are calculated using POWHEGBOX-Res framework at NLO with massive b -quarks [125], $m_b = 4.75\text{GeV}$, with the "4FS NLO as 0118" PDF sets [124]. The POWHEGBOX internal parameter h_{bzd} is set to 5. The renormalization scale is set to the geometric average of the transverse mass of top- and b -quarks defined as $m_{T,i} = \sqrt{m_i^2 + p_{T,i}^2}$, where m_i refer to the mass, $p_{T,i}$ to the transverse momentum and i to the top or b -quark. The factorization scale is related to the average of the transverse mass of the outgoing partons in the matrix element calculation. The PYTHIA 8 parameters for parton shower and hadronization modeling are set to the A14 [113] tune.

This sample is used as nominal $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1B$ prediction. The relevant uncertainties are evaluated following the same recipes as the nominal $t\bar{t}$ samples.

Process	MC generator	Accuracy
Single-top	POWHEG BOX v2 +PYTHIA 8.230	NLO
W + jets	SHERPA 2.2.11	NLO
$t\bar{t}V$	MADGRAPH5_AMC@NLO 2.3.3 +PYTHIA 8.210	NLO
$t\bar{t}H$	POWHEG BOX v2 +PYTHIA 8.230	NLO
VV	SHERPA 2.2.1 or 2.2.2	NLO

Table 4.3: Summary of MC generators used for the rare process with small contributions in the studies.

4.2.5 Simulation for minor background

Other minor background processes could contribute to the studies described in this thesis. A summary of the generator used for these samples can be found in Table 4.3.

CHAPTER 5

DEEP SETS BASED NEURAL NETWORKS FOR LOW-PT $X \rightarrow bb$ IDENTIFICATION

This chapter describes the development of a flavor tagging algorithm, DEXTER, for merged reconstructed jets of low transverse momentum, $p_T < 200$ GeV. It is the first general-purpose low- p_T double-b tagger in ATLAS. It is explicitly designed for the $H \rightarrow aa \rightarrow 4b$ search performed in this thesis with a goal to also apply it to other light resonance searches with a merged double-b jet final state. A brief overview of the existing double-b tagger is given in Section 5.1, which serves as the motivation for the development of DeXTer. Chapter 5.2 describes the special object reconstruction used as input to the algorithm. The DEXTER algorithm development and performance are extensively discussed in Sections 5.3 and 5.4, respectively. To use the tagger in a physics analysis, it is essential to measure the algorithm’s efficiency in data, $t\bar{t}$ and $Z + \text{jets}$ events selected from data are used for this measurement. Results are summarized in Section 5.5.

5.1 Boosted jet tagging in ATLAS

Several SM and BSM processes produce final states with two or more collimated particle jets reconstructed as a single, merged jet in the detector. The excellent spatial resolution of the ATLAS tracker and calorimeter allows different particle jets merged into the single reconstructed jet to be partially resolved [126]. Resolving individual particle jets relies on substructure patterns formed when the jets are clustered together in a single reconstructed jet. In order to properly identify the substructure of

merged jets, it is crucial to have information about all particles from all constituent particle jets, which can be challenging since the standard jet clustering was designed for reconstructed jets associated to a single particle jet. Therefore, substructure identification algorithms are usually applied to jets clustered with large radius (large- R) parameters. One example of such large- R clustering are the calorimeter jets clustered with the anti- k_t algorithm and radius parameter $R = 1.0$ used in several ATLAS analyses [127]. Existing algorithms in ATLAS use large- R jets for boosted heavy resonance, such as W, Z, and Higgs boson. For example, the high- p_T $X \rightarrow b\bar{b}$ tagger uses variable-radius (VR) track-jet reconstruction [128] to identify the large- R jet substructure and apply b -tagging algorithms on them for double- b tagging [129].

However, using a large- R clustering algorithm can possibly cluster a large number of particles that come either from long-range phenomena or pile-up. These additional particles usually have low transverse momentum but can have a non-negligible contribution to the jet energy scale when integrated over a large jet area. Several dedicated techniques have been developed to address these challenges [130], relying on the expected behavior of QCD showers to trim off regions of the jet with a small relative contribution to the jet transverse momentum. The trimming of low transverse momentum particles only works well if the particles associated to the particle jets are of relatively large transverse momentum. Thus, large- R reconstructed jets are only defined and calibrated for $p_T > 200$ GeV, rendering most substructure techniques inaccessible to analyses with final states comprising low- p_T , merged jets. This is not suitable for the jets we are interested in from our signals. The only calibrated jets in this transverse momentum range are the particle-flow jets reconstructed using the anti- k_t algorithm radius parameter $R = 0.4$.

The development of the DEXTER algorithm has been motivated by several searches of BSM physics where the transverse momentum of a $b\bar{b}$ pair is produced by an intermediate light resonance. In particular, this thesis, focuses on exotic Higgs boson de-

cays to a pair of new low-mass scalar particles. The new scalar can subsequently decay to one or more $b\bar{b}$ pairs which require a specialized identification method. Also, new low-mass scalars and pseudoscalars are present in models where they are produced in association with top-quark pairs [131]. Developing a specialized flavor tagging algorithm for light resonances can benefit many searches with such signatures.

5.2 Special object reconstruction

DEXTER uses standard ATLAS particle-flow jets described in Section 3.3. DEXTER is optimized to apply on only a subset of these jets that satisfy the requirements described in this section. We start by reclustering particle-flow jets to find isolated jets that allow us to build a $R = 0.8$ track-jet to capture leakage outside the jets. This is described in Section 5.2.1. Furthermore, track subjets and secondary vertices are reconstructed from the tracks associated with the $R = 0.8$ track jet. These provide additional information for the flavor tagging algorithm.

The DEXTER algorithm uses tracks that satisfy a *loose* selection criteria, defined as $p_T > 0.5 \text{ GeV}$, $|d_0| < 3.5 \text{ mm}$, and $|z_0 \sin \theta| < 5 \text{ mm}$. Only jets with $p_T > 20 \text{ GeV}$ and $|\eta| < 2.5$ are used. To reduce the background from muons from heavy-flavor decays inside jets, muons are required to be separated by $\Delta R > 0.4$ from the nearest jet. A muon with $p_T > 4 \text{ GeV}$ is within $\Delta R < 0.4$ of a jet is called a *soft-muon* if the jet has at least three associated tracks. The jet is removed if a jet has less than three tracks and a nearby muon. Soft-muons are not used to identify events but to aid jet classification in data events.

5.2.1 Track-jet reconstruction

The radius parameter $R = 0.4$ provides good containment for particles created in the fragmentation of a jet originating from the hadronization of a single parton. To capture the particles from the fragmentation or decay of heavy-flavored hadrons from

multiple partons which may escape the $R = 0.4$ jet, DEXTER associates an extended collection of tracks to a reconstructed jet by clustering all PFlow jets and ID tracks matched to the jets using the ghost-association method [132]. This method treats the tracks as four vectors of infinitesimal magnitude during the jet reconstruction and assigns them to the jet with which it is clustered using an anti- k_t algorithm with radius parameter $R = 0.8$. This process is called reclustering [133], and its purpose is two-fold here. First, it creates large- R track-jets around each PFlow reconstructed jet. Second, it creates a technical definition of jet isolation. An isolated jet is defined here as a reclustered jet with a single PFlow jet constituent. The radius parameter $R = 0.8$ is chosen as a compromise between capturing most tracks from heavy-flavored hadron decays and not including unnecessary uncorrelated tracks that cannot be removed with trimming procedures at this momentum range. The same radius parameter has been used successfully in a previous version of this work [41]. The $R = 0.8$ track jets are the primary input to the DEXTER algorithm. The associated $R = 0.8$ track jet is required to satisfy $|\eta| < 2.0$ to account for the extended radius and the acceptance of the ID.

5.2.2 Multiple secondary vertex reconstruction

Many flavor tagging algorithms have explored secondary vertices (SV) in jets in ATLAS [96, 134]. These algorithms either reconstruct the cascade decay of detached b -hadron vertex and subsequent decay vertex to c -hadron or only try to find the displaced vertex of a b -hadron.

To explore the multiple heavy-flavor decays with DEXTER, an algorithm to reconstruct multiple decay vertices was developed by combining two existing algorithms: the track-cluster-based low- p_T vertex tagger (TC-LVT) [95], and the multiple secondary vertex finder algorithms (MSVF) [135].

The TC-LVT algorithm has been developed for soft b -hadron tagging and optimized to reconstruct low- p_T b -hadron decays. This work uses the clustering algorithm from TC-LVT to identify the collection of tracks that may have at least one displaced secondary vertex. The MSVF algorithm is used to identify multiple SVs in the track cluster. The algorithm builds all two-track proto-vertices consistent with displaced tracks incompatible with a hadronic material interaction, a photon conversion, or the decay of long-lived light-flavored hadrons. All displaced tracks reconstructed in the ID are used to build proto-vertices. Proto-vertices define relations between tracks. A single track can be part of more than one proto-vertex and therefore be related to more than one other track. Each set of tracks that are mutually connected to each other form a secondary vertex. After secondary vertices are formed, tracks with small compatibility with the vertex are removed, and the ambiguity caused by distant vertices sharing common tracks is resolved. The MSVF algorithm also merges nearby vertices. Finally, reconstructed SVs are required to be ΔR -matched to a reclustered $R = 0.8$ track-jet.

5.2.3 Track subjet reconstruction

One of the main goals of substructure methods is to identify the flight axis of the several particle jets merged in a single reconstructed jet. Several algorithms have been developed for that purpose [136] with different performances and algorithmic complexity. A good compromise is obtained with the exclusive- k_t clustering of the jet constituents [137]. The exclusive- k_t algorithm is a sequential clustering algorithm that compares the relative k_t -distance $\min(p_{T,i}, p_{T,j}) \times \Delta R_{ij}$ between pairs of components (i, j) and the so-called beam distance $p_{T,i}$. The component is removed if the smallest value in the set is the beam distance. On the other hand, if the smallest value is the k_t -distance, components i and j are clustered together in a pseudo-jet. The algorithm then iterates over the merged pseudo-jet and stops when a well-defined

Jet flavor label	Number of b -hadrons	Number of c -hadrons
B	≥ 2	-
b	$= 1$	$= 0$
c	$= 0$	$= 1$
light	$= 0$	$= 0$

Table 5.1: Jet flavor labeling based on number of ghost-associated b -hadrons or c -hadrons

number of pseudo-jets remain, the subjects used to estimate the flight direction of the two particle-jets merged together.

DEXTER uses as inputs the direction of the subjects formed using the exclusive- k_t algorithm on the tracks from the reclustered $R = 0.8$ track jet, stopping when two pseudo-jets are left. In this note, this particular algorithm instance is referred to as $\text{Ex}k_t^{(2)}$. Each $\text{Ex}k_t^{(2)}$ subject is required to satisfy $p_T > 5 \text{ GeV}$ where the track-subject p_T is estimated from the sum of the tracks' four-momenta.

5.2.4 Jet flavor labeling

In the MC, the true flavor of reconstructed jets is determined by the number of ghost-associated [138] b - and c -hadrons with $p_T > 5 \text{ GeV}$. b -hadrons (c -hadrons) in the decay chain of another b -hadron (c -hadron, respectively) are not ghost-associated to avoid double-counting. Jets are categorized as a B -labeled jet if at least two b -hadrons are ghost-associated to the jet. If exactly one b -hadron and no c -hadron are ghost-associated to the jet, it is categorized as a b -labeled jet. The reconstructed jet is categorized as light-labeled if no b - or c -hadron is ghost-associated.

5.3 Algorithm architecture

DEXTER is inspired by the recent developments using the Deep Sets architecture to develop a permutation-invariant neural network (NN) for flavor tagging in ATLAS. The set-based impact-parameter b -tagging algorithm (DIPS) [139] uses track

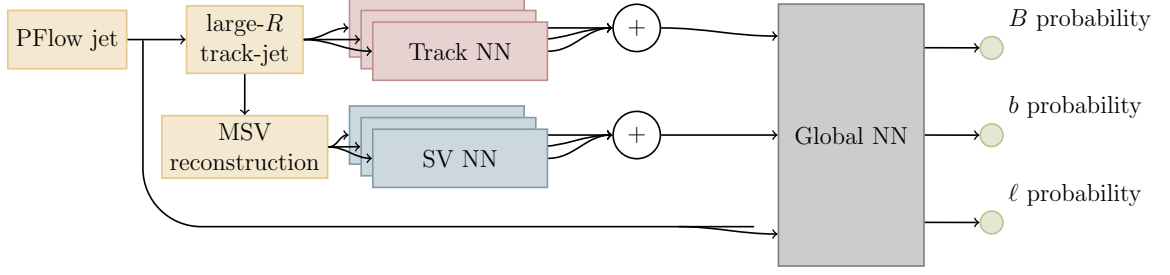


Figure 5.1: Architecture of the DEXTER algorithm. Two sets of feature-extracting NN are used for track and SV observables. A final global NN with a softmax output layer is used to interpret each output as a probability for each flavor.

information as input for jet flavor tagging. This approach minimizes the empirical ordering typically introduced in the prior algorithm on tracks associated to jets. The same architecture has also been used for energy-flow and particle-flow event classification [140, 141].

The DEXTER algorithm extends this architecture to incorporate extra information from reconstructed secondary vertices and kinematical variables from the PFlow jet. The architecture comprises two separate feed-forward NNs which serve as the feature extractor for the tracks in the $R = 0.8$ track jet and for the SVs. An additional global feed-forward NN combines the output of the feature-extraction NNs with the jet kinematics to learn correlations between them. The calibrated energy-momentum of the single PFlow jet in the isolated jet is used as estimate of the jet kinematics for the NN. A final layer with a softmax activation function [142] is used to predict the probability for each flavor label. Figure 5.1 summarizes the structure of the complete algorithm.

The feature-extraction NNs contain 2 hidden layers with 100 neurons and an output layer with 128 features. The 256 features are used as input to the global NN which has 3 hidden layers. All NNs use the ELU activation function [143]. The global NN also receives the PFlow jet p_T and η as inputs.

5.3.1 Track neural network

The tracks clustered in the $R = 0.8$ track-jet are ordered in decreasing value of transverse impact parameter significance S_{d_0} [144]. Properties of the 25 first tracks are used as input to the DEXTER algorithm. The track NN algorithm input variables are presented in Table 5.2. Variables that depend on both the track and the reconstructed jet, namely $\Delta\eta(\text{track}, \text{jet})$ and $\Delta\phi(\text{track}, \text{jet})$, are calculated with respect to the axis of the $\text{Ex}k_t^{(2)}$ track-subjet, since it is a better estimate of the original b -hadron flight direction than the PFlow jet. The sign of the impact parameter significance is not altered because the PFlow jet axis still represents the best estimate for reconstructed jets that do not come from merged particle jets.

5.3.2 Secondary vertex neural network

The reconstructed secondary vertices matched to the $R = 0.8$ track-jet are ordered in decreasing value of the transverse decay length significance $S_{L_{xy}}$ and the properties of the 12 first vertices are used as input to the secondary vertex NN. The properties used are listed in Table 5.2. The definition of the variables is identical to the one used in other SV-based b -taggers [96], with the only differences being the choice of the $\text{Ex}k_t^{(2)}$ jet axes as the reference for $\Delta\eta(\text{track}, \text{jet})$ and $\Delta\phi(\text{track}, \text{jet})$. Decay lengths are always calculated with respect to the event PV:

$$L_{xy} = |(\vec{p}^{\text{SV}} - \vec{p}^{\text{PV}}) \times \hat{z}| \quad L_z = |(\vec{p}^{\text{SV}} - \vec{p}^{\text{PV}}) \cdot \hat{z}|.$$

The decay length significance values are calculated taking into account the covariance matrices of both primary and secondary vertices:

$$S_{L_{xy}} = (p^{\text{SV}} - p^{\text{PV}})_i [P_{xy}^{-1} (C^{\text{SV}} + C^{\text{PV}})^{-1} P_{xy}]_{ij} (p^{\text{SV}} - p^{\text{PV}})_j,$$

$$S_{L_z} = (p^{\text{SV}} - p^{\text{PV}})_i [P_z^{-1} (C^{\text{SV}} + C^{\text{PV}})^{-1} P_z]_{ij} (p^{\text{SV}} - p^{\text{PV}})_j.$$

where P_{xy} and P_z are projectors onto the transverse plane and longitudinal direction, respectively.

Even though an ordering is imposed to select the maximum number of input tracks and secondary vertices, the NN itself is permutation invariant.

5.3.3 Color charge adversarial neural network

The algorithm obtained with the NN depicted in Figure 5.1 will exhibit significant differences in the response for color singlet and octet $b\bar{b}$ states. The difference stems from the larger track multiplicity observed in the shower from octet states and the track distribution from the color interaction between two octet states [145]. This difference could be used to create optimized taggers for the singlet and octet cases. However, since the only calibration sample available comes from $g \rightarrow b\bar{b}$, a choice is made to minimize the differences by penalizing the feature extracting NNs using a domain-adversarial training of the NN (DANN) [146].

The B -labeled jet sample used for training is an ensemble mixture of $a \rightarrow b\bar{b}$ and $g \rightarrow b\bar{b}$ events. The adversarial NN receives the same inputs as the global NN and is built with a categorical loss L_A that discriminates between the two categories. The loss function L_D is used to discriminate B -, b -, and light-labeled jets are built from the three output neurons on the global NN. While the backpropagation of the global NN is performed only with $\partial L_D / \partial \theta_g$ and the backpropagation of the adversarial NN is performed only with $\partial L_A / \partial \theta_a$, the backpropagation of the feature-extraction NNs is done via the gradient:

$$\frac{\partial L_D}{\partial \theta_f} - \lambda \frac{\partial L_A}{\partial \theta_f},$$

where $\theta_g, \theta_a, \theta_f$ are the global, adversarial, and feature-extracting NN weights, respectively. The optimized value $\lambda = 10$ is chosen for the regulator when training

Feature	Description
Jet	
p_T	Jet transverse momentum
η	Jet pseudorapidity
Track	
$\log p_T^{\text{frac}}(\text{track, PFlow jet})$	$\log p_T^{\text{track}}/p_T^{\text{PFlow jet}}$ of the p_T fraction between track and PFlow jet p_T
$\Delta\eta$ (track, $\text{Ex}k_t^{(2)}$ track jet)	Pseudorapidity difference between track and $\text{Ex}k_t^{(2)}$ jet
$\Delta\phi$ (track, $\text{Ex}k_t^{(2)}$ track jet)	Angular difference between between track and $\text{Ex}k_t^{(2)}$ jet
d_0	Transverse impact parameter
$z_0 \sin \theta$	Longitudinal impact parameter
S_{d_0}	d_0/σ_0 : transverse IP significance
$S_{z_0 \sin \theta}$	$z_0 \sin \theta/\sigma_{z_0 \sin \theta}$: longitudinal IP significance
PIX1 hits	Number of hits in the first pixel layer
IBL hits	Number of hits in the IBL
Shared IBL Hits	Number of shared hits in the IBL
Split IBL Hits	Number of split hits in the IBL
Shared pixel hits	Number of shared hits in the pixel layers
Split pixel hits	Number of split hits in the pixel layers
Shared SCT hits	Number of shared hits in the SCT
nPixHits	Number of hits in the pixel layers
nSCTHits	Number of hits in the SCT layers
Secondary Vertex	
$\log(m)$	Track mass of the secondary vertex
$\log p_T^{\text{frac}}(\text{vertex, PFlow jet})$	$\log p_T^{\text{SV}}/p_T^{\text{PFlow jet}}$ of the p_T fraction between the secondary vertex and PFlow p_T jet
$\Delta\eta$ (vertex, $\text{Ex}k_t^{(2)}$ track jet)	Pseudorapidity difference between the secondary vertex and the $\text{Ex}k_t^{(2)}$ jet
$\Delta\phi$ (vertex, $\text{Ex}k_t^{(2)}$ track jet)	Angular difference between between the secondary vertex and the $\text{Ex}k_t^{(2)}$ jet
L_{xy}	Transverse decay length relative to primary vertex
L_z	Longitudinal decay length relative to primary vertex
$S_{L_{xy}}$	Transverse decay length significance
S_{L_z}	Longitudinal decay length significance

Table 5.2: List of features used as input for the DEXTER algorithm. In the table, $\text{Ex}k_t^{(2)}$ jet refers to track subjects reconstructed as described in Section 5.2.3.

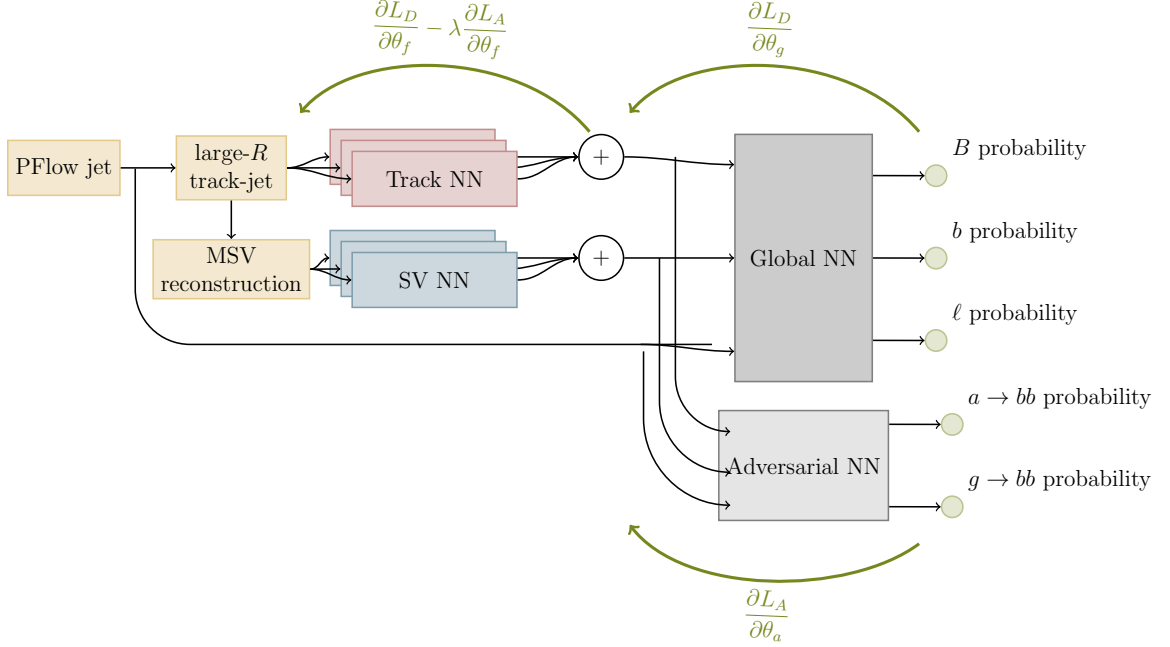


Figure 5.2: Architecture of the DEXTER algorithm with an adversarial NN. A DANN strategy is used to train the feature-extraction NNs. The adversarial NN discriminates $a \rightarrow b\bar{b}$ and $g \rightarrow b\bar{b}$ events included in the B -labeled jet ensemble mixture.

the DEXTER algorithm to minimize the difference in the response between color singlets and octets.

DEXTER was trained with 3.4 million B -, b -, and light-labeled jets. From this dataset, 20% of jets are held out as a validation set. The architecture uses Keras [147] with the TensorFlow backend [148]. An early stopping method that monitors the improvement of the validation set loss is used during the training. The Adam optimizer [149] is used with a learning rate of 0.001 and batch size of 512. Another independent set of 0.8 million jets is kept out for the performance studies shown in Section 5.4.

5.4 Algorithm performance

As described in Section 5.3, the DEXTER algorithm is trained to classify reconstructed jets into three categories based on the extracted input features. The

three-class probabilities predicted by the model (p_B, p_b, p_l) are combined in a tagging discriminant, D_B , defined as:

$$D_B = \ln \frac{p_B}{(1 - f_b)p_l + f_bp_b}, \quad (5.1)$$

where f_b is a free parameter that balances between the rejection of light-labeled jets versus b -labeled jets. The value $f_b = 0.4$ is used in this work based on the flavor composition of the $t\bar{t}$ sample. Different applications may optimize f_b for background rejection. A clear separation between different categories can be observed in Figure 5.3.

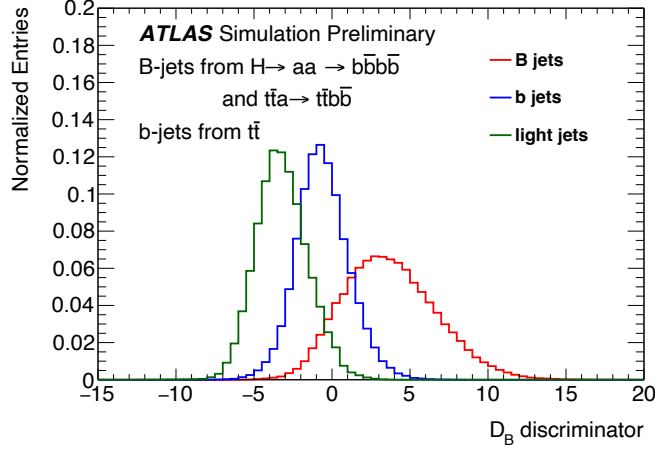


Figure 5.3: DEXTER discriminant distribution evaluated using B -labeled jets from $H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b})$ and $t\bar{t}a$ with $a \rightarrow (b\bar{b})$ and b -, and light-labeled jets from $t\bar{t}$ samples with $f_b = 0.4$. The D_B score distributions are normalized to unity.

The model can be further analyzed by examining saliency maps [150]. The saliency maps are defined as follows:

$$\mathcal{S}^i = \frac{\partial D_B}{\partial z_i}(D_B^{-1}(\tau)) \quad (5.2)$$

is computed by taking the derivative of discriminate D_B with respect to the normalized value:

$$z_i = \frac{x_i - \langle x_i \rangle}{\sigma_{x_i}} \quad (5.3)$$

of the input parameter x_i . The derivative is estimated by taking the average over all tracks (SV) of jets for which $|D_B - \tau| < 0.2$. The saliency map is calculated separately for B and b jets. Moreover, to highlight the importance of highly displaced tracks (SV), the average is calculated separately for tracks (SV) with S_{d_0} ($S_{L_{xy}}$) below 2, between 2 and 5, and above 5.

These values provide the relative importance of each input variable at the working point (WP) defined by a specific value of $D_B = \tau$. Figure 5.4 shows the saliency maps for each track input. The map is estimated for highly displaced tracks ($S_{d_0} > 5$), partially displaced tracks ($2 < S_{d_0} < 5$), and non-displaced ($S_{d_0} < 2$) tracks. Two different WPs are shown: an inclusive tight WP, corresponding to an average 40% efficiency, and an inclusive loose WP, corresponding to an average 60% efficiency. Figure 5.5 shows an equivalent map for SV inputs.

5.4.1 Performance dependency on resonance color charge

The exact response of the DEXTER algorithm depends on the adversarial penalty introduced to minimize the difference between the decay of color singlets and octets. Figure 5.6 compares the difference in response between the two types of decay to a $b\bar{b}$ pair with and without the DANN. The value $\lambda = 10$ was chosen to minimize the difference in the area under the receiver operating characteristic (ROC) curve for $a \rightarrow b\bar{b}$ and $g \rightarrow b\bar{b}$ events.

5.4.2 Performance dependency on resonance mass

Another desirable feature of the tagger is to be independent of the parent particle mass. Not only is it desirable for the tagger performance to be independent of unknown model parameters, but it is also fundamental when mass sidebands are used.

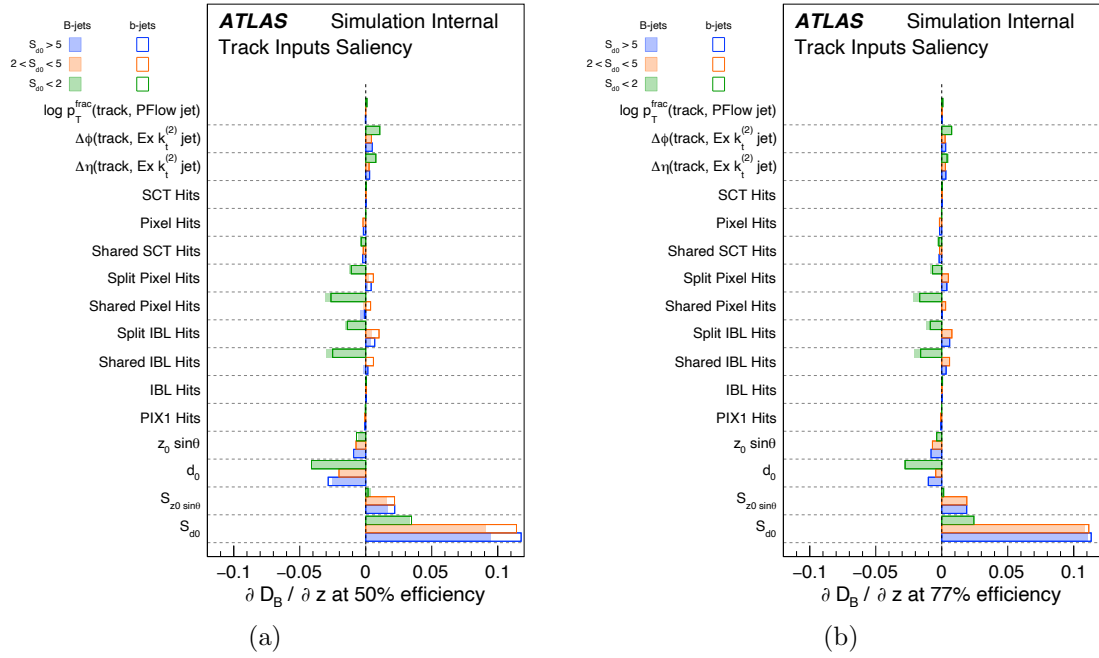


Figure 5.4: The saliency map of D_B with respect to the track inputs. The maps are shown for (a) an inclusive tight WP corresponding to $D_B = 3.2$ and (b) an inclusive loose WP corresponding to $D_B = 1.8$.

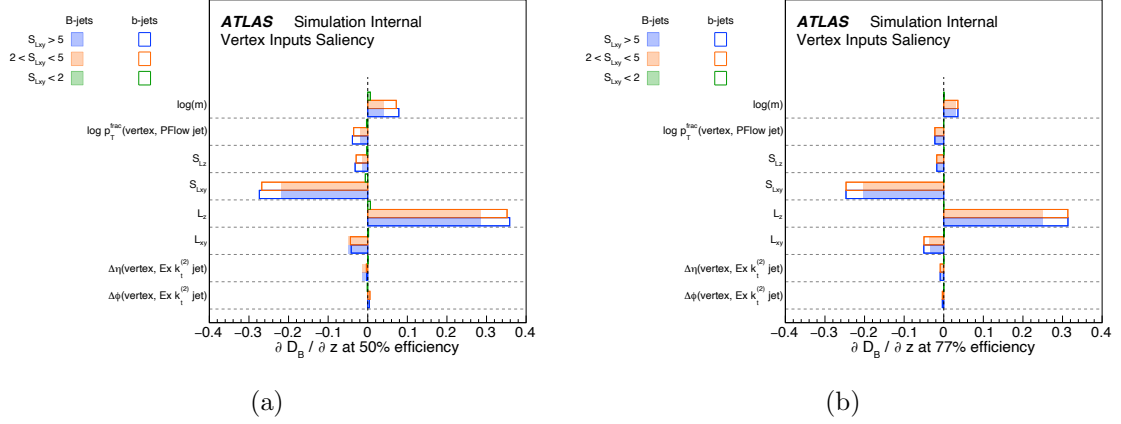


Figure 5.5: The saliency map of D_B with respect to the SV inputs. The maps are shown for (a) an inclusive tight WP corresponding to $D_B = 3.2$ and (b) an inclusive loose WP corresponding to $D_B = 1.8$.

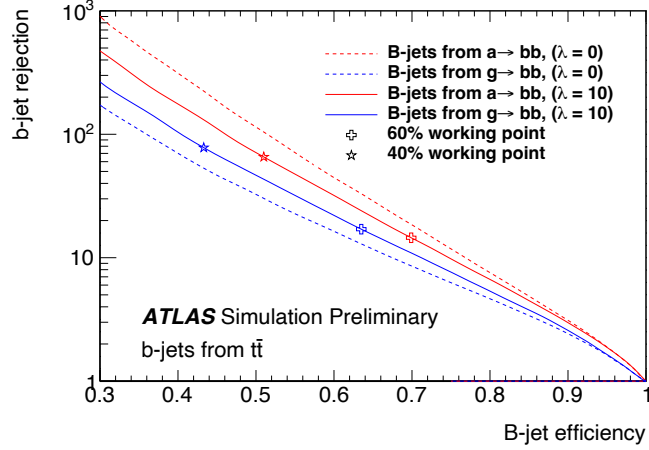


Figure 5.6: Comparison of the B -labeled jet tagging efficiency vs. the b -labeled jet rejection for B -jets originating from color singlets (red) and color octets (blue). The ROC curves are shown for architectures of DEXTER with (solid) and without (dashed) DANN. Two example working points are marked with labels corresponding to the approximate B -labeled jet efficiency for $g \rightarrow b\bar{b}$. The two working points are the boundaries of the intervals for which the efficiencies are measured in data. The difference in efficiency for B -labeled jets between $a \rightarrow b\bar{b}$ and $g \rightarrow b\bar{b}$ is approximately 6 – 7% in the two working points.

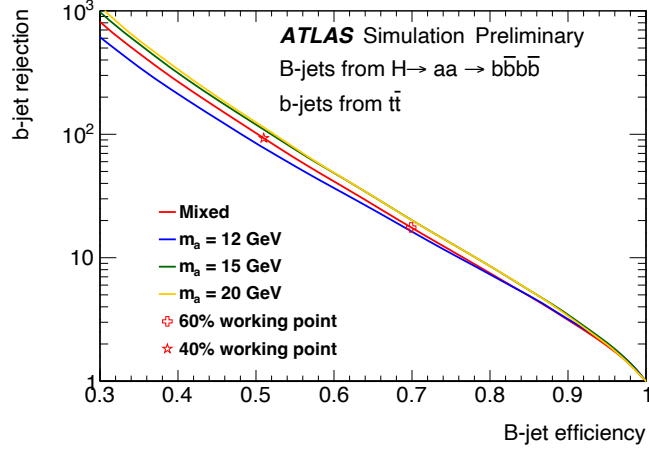


Figure 5.7: Performance for different a -boson masses in $H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b})$ events. The same working points defined in Figure 5.6 are marked. The B -labeled jet efficiency on the two reference working points differs by approximately 2 – 3% across the examined values of m_a .

The PFlow jet mass is not a good proxy for the B -labeled jet particle-level mass since it doesn't provide good containment for multiple particle jets merged in the same reconstructed jet. Because of this, mass correlation is assessed using the true mass of the a -boson in decays $a \rightarrow b\bar{b}$.

Two strategies are adopted to minimize the mass correlation. First, the B -enriched sample used to train DEXTER is composed of an ensemble mix of $H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b})$ and $t\bar{t}a, a \rightarrow b\bar{b}$ samples with different values of the a -boson mass. Second, the impact of each input variable with the mass correlation of the response is studied. When a variable is identified as the cause of a large difference in performance, it is either redefined to be less mass-sensitive or removed when a suitable redefinition is not possible. The performance difference is examined by comparing the ROC curves across different m_a values in $H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b})$ decays, as shown in Figure 5.7.

5.5 Efficiency measurement in data and simulation

The development of DEXTER relies solely on the MC simulated sample. There is no guarantee the tagger will behave similarly when applied to real data. It's essential to perform efficiency measurements with data to correct imperfect modeling in simulation for the practical usage of DEXTER in physics analysis.

However, measuring the efficiency of DEXTER is challenging since selecting merged B -flavored particle jets unbiasedly can be difficult. A particularly clear decay topology that can be explored is $Z(\rightarrow b\bar{b})\gamma$ [151]. However, this decay is only merged for $p_{\text{T}}^Z \gtrsim 400$ GeV, beyond the p_{T} range targeted by DEXTER. A previous version of such tagger [41] used $g \rightarrow b\bar{b}$ processes in multijet events, but those events require prescaled triggers, which limit the number of available events and are challenging to simulate due to the large cross-section. As a result, in this iteration of calibration for jets tagged with the DEXTER algorithm using $Z(\rightarrow \ell^+\ell^-) + g(\rightarrow b\bar{b})$ events as a source of B -labeled jets and $t\bar{t}$ events as a source of b -labeled jets [92].

Efficiencies are measured in two tagging intervals. The first tagging interval is defined by $D_B > 3.2$ and corresponds to efficiencies in the range 0 – 40%. The second tagging interval is defined as $1.8 < D_B < 3.2$ and corresponds to average efficiencies in the range 40 – 60%. Jets with $D_B < 1.8$ belong to the 60 – 100% tagging interval. The representative tagging efficiencies of each interval are determined using a $g \rightarrow b\bar{b}$ sample and would be slightly higher if a $a \rightarrow b\bar{b}$ sample were used, as shown in Figure 5.6. The 0 – 40% tagging interval has an average expected light-labeled jet efficiency of $1/(2.4 \times 10^4)$. For each tagging interval, the efficiency is measured in three exclusive p_{T} bins: $20 < p_{\text{T}}^{\text{jet}} < 90$ GeV, $90 \leq p_{\text{T}}^{\text{jet}} < 140$ GeV, and $140 \leq p_{\text{T}}^{\text{jet}} < 200$ GeV, for a total of 6 different measurements per jet label.

Figure 5.8 shows the efficiency observed in simulation for the 0 – 40% and 40 – 60% tagging intervals. Unless otherwise specified, all efficiencies for simulated samples are determined through truth matching of jets produced by POWHEGBOX +PYTHIA8.

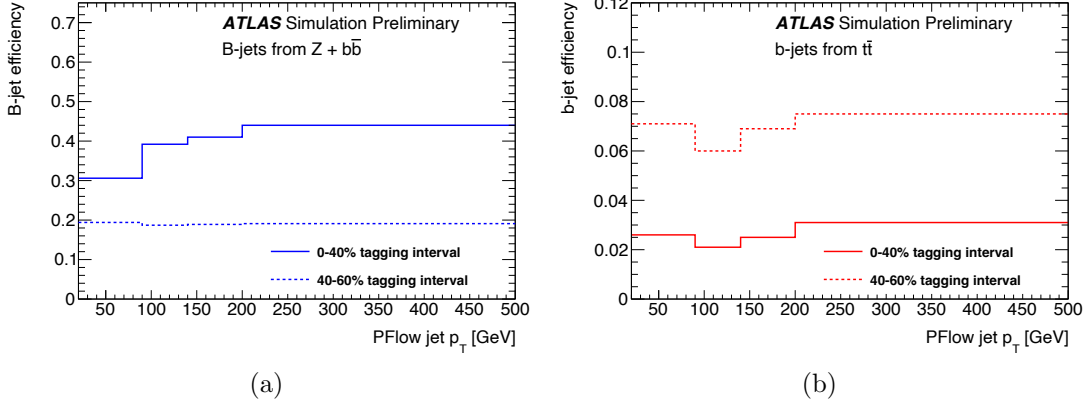


Figure 5.8: DEXTER efficiency in simulation for (a) B - and (b) b -labeled jets as a function of the jet p_T .

The efficiencies for B - and b -labeled jets in all p_T^{jet} ranges and tagging intervals are measured simultaneously. This allows taking into account the full correlation model in the propagation of uncertainties, which is not possible in other calibration strategies where tagging and mis-tagging efficiencies are measured separately [152].

5.5.1 Event selection

Two regions are defined, targeting different processes and jet flavor compositions. A top-region is defined by the presence of a high- p_T electron and a high- p_T muon produced in association with multiple jets to target top quark pair events. On the contrary, Z -region is defined by the presence of an opposite-charge ee or $\mu\mu$ pair with an invariant mass around the Z pole mass. In both regions, the tight lepton identification suppresses any significant non-prompt background. Due to the use of single lepton triggers, the leading lepton is required to satisfy $p_T > 27$ GeV in both top- and Z -region.

Top-region event selection and channels

The definition of the top-region is inspired by the measurement of the b -jets identification efficiency with $t\bar{t}$ events in ATLAS [92]. Candidate events are required to have

exactly two leptons with opposite sign charges. To suppress the contamination from backgrounds with Z -bosons, events are required to have one electron and one muon. A selection on the mass of the two leptons $m_{e\mu} > 50$ GeV is applied to reject events with low dilepton invariant mass that the simulation does not well-model. Events are also required to have exactly two reconstructed jets to suppress light-flavor jets from initial- or final-state radiation. Furthermore, a $\Delta R(\text{jet}, \text{leptons}) > 0.8$ is applied to avoid leptons close to jets used to measure the efficiency.

A simple top-quark pair reconstruction is also adopted to achieve higher purity of two b -jets events from $t\bar{t}$ decays. The electron and the muon are paired with the selected jets in the event using the following criterion:

$$\underset{\ell_1, \ell_2 \in \{e, \mu\}}{\operatorname{argmin}} (m_{j_1, \ell_1}^2 + m_{j_2, \ell_2}^2),$$

where j_1 (j_2) is the jet with the highest (lowest) p_T . The invariant masses of the pairs formed are required to satisfy $m_{j_1, \ell_1} > 20$ GeV, and $m_{j_2, \ell_2} > 20$ GeV to avoid regions populated with low- p_T jets.

The masses of the pairs (j_1, ℓ_1) and (j_2, ℓ_2) are used to define signal regions (SR) and control regions (CR):

- $m_{j_1, \ell_1} < 175$ GeV and $m_{j_2, \ell_2} < 175$ GeV (SR bb),
- $m_{j_1, \ell_1} < 175$ GeV and $m_{j_2, \ell_2} > 175$ GeV (CR $b\ell$),
- $m_{j_1, \ell_1} > 175$ GeV and $m_{j_2, \ell_2} < 175$ GeV (CR ℓb),
- $m_{j_1, \ell_1} > 175$ GeV and $m_{j_2, \ell_2} > 175$ GeV (CR $\ell\ell$).

Jets from events in the signal regions are further categorized depending on their p_T and their DEXTER tagging interval to measure the different efficiencies. Jets from events in the control regions are further categorized only based on their p_T . These regions serve primarily to constrain the flavor composition of the selected jets

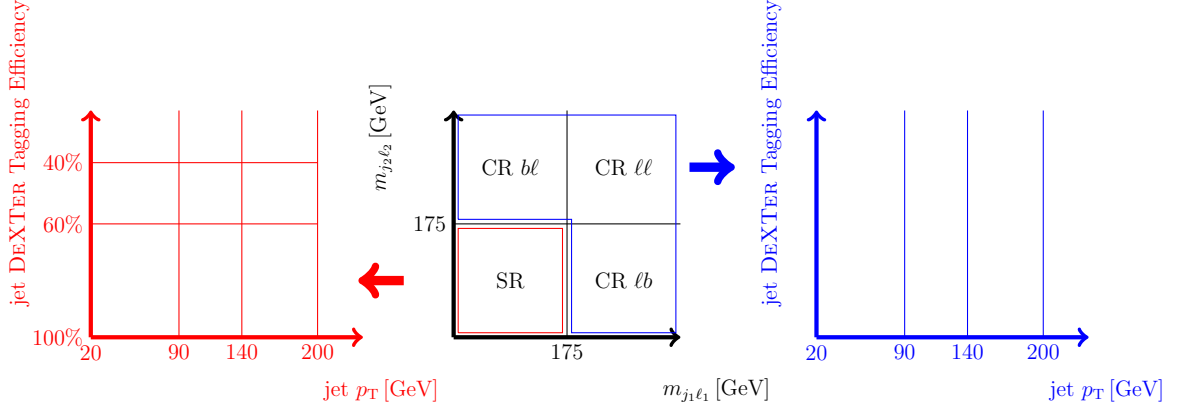


Figure 5.9: Diagram of the event selection and channel classification in the top-region.

by adding channels that, despite not being sensitive to the efficiency values, provide independent degrees of freedom. Figure 5.9 shows a diagrammatic representation of the event selection.

Z-region event selection and channels

Events in the Z -region are required to have exactly two leptons, electrons or muons, with the same flavor and opposite electric charge. The dilepton invariant mass is required to be in the Z -boson mass-pole region $81 < m_{\ell\ell} < 101$ GeV. Events with exactly one *probe* jet, defined as the single jet selected in addition to the leptons and used to measure the efficiency of the DEXTER algorithm, are selected. A condition $\Delta R(\text{jet}, \ell\ell) > 1.0$ is imposed on the angular separation between the probe jet and the dilepton pair to avoid mismodeling in the production of Z bosons with multiple jets. Another similar requirement $\Delta R(\text{jet}, \text{lepton}) > 1.0$ is applied to avoid leptons too close to the probe that could impact the performance of the DEXTER algorithm.

To enrich the fraction of heavy-flavored jets in this region, the probe jet is required to have a soft-muon within $\Delta R(\text{soft-muon}, \text{Ex}k_t^{(2)}) < 0.3$ of one of the two $\text{Ex}k_t^{(2)}$ axes. The track-subjet closest to the soft-muon is called the *muon-Ex* $k_t^{(2)}$ track-subjet while the other is called the *non-muon-Ex* $k_t^{(2)}$ track-subjet.

Jets in the Z -region which fail the 40 – 60% tagging interval condition are further categorized in channels depending on the $\langle S_{d_0} \rangle$ of the muon- $\text{Ex}k_t^{(2)}$ and non-muon- $\text{Ex}k_t^{(2)}$ track-subjets. The observable $\langle S_{d_0} \rangle$ is defined as:

$$\langle S_{d_0} \rangle = \frac{1}{3}(S_{d_0}^1 + S_{d_0}^2 + S_{d_0}^3),$$

where $S_{d_0}^{1,2,3}$ are the signed transverse impact parameter significances of the three tracks ghost-associated to the track-subjet with highest value of S_{d_0} .

The non-muon- $\text{Ex}k_t^{(2)}$ is binned in two regions $\langle S_{d_0} \rangle < 0$ and $\langle S_{d_0} \rangle \geq 0$ to create channels enriched in light- and b -labeled jet components. The number of B -labeled jets that fail the 40 – 60% tagging interval is small compared to these two other components. The muon- $\text{Ex}k_t^{(2)}$ is binned in 6 different bins. A total of $3(p_T \text{ bins}) \times 12(\langle S_{d_0} \rangle \text{ bins})$ channels are used in the 60 – 100% tagging interval.

For jets in the 0 – 40% and 40 – 60% tagging intervals, an additional categorization is performed depending on the largest secondary-vertex mass $m_{\text{SV}}^{\text{max}}$ within the non-muon- $\text{Ex}k_t^{(2)}$ track-subjet $\Delta R(\text{SV}, \text{Ex}k_t^{(2)}) < 0.3$. Two channels are defined by $m_{\text{SV}}^{\text{max}} < 2 \text{ GeV}$ and $m_{\text{SV}}^{\text{max}} \geq 2 \text{ GeV}$ which are enriched in b - and B -labeled jets, respectively. Jets without a reconstructed SV are assigned $m_{\text{SV}}^{\text{max}} = 0 \text{ GeV}$. An identical binning is made for the muon- $\text{Ex}k_t^{(2)}$ track-subjet. A total of $3(p_T \text{ bins}) \times 2(\text{tagging interval}) \times 4(m_{\text{SV}}^{\text{max}} \text{ bins})$ channels are used in the 40 – 60% and 0 – 40% tagging intervals. The total number of channels in the Z -region is 60. Figure 5.10 shows a diagram of the classification described.

5.5.2 Methodology

A binned Poisson likelihood is built based on the number of jets observed in the channels and regions described in Section 5.5.1. Each event from the top-region contributes with two jets, while each event from the Z -region contributes with a single

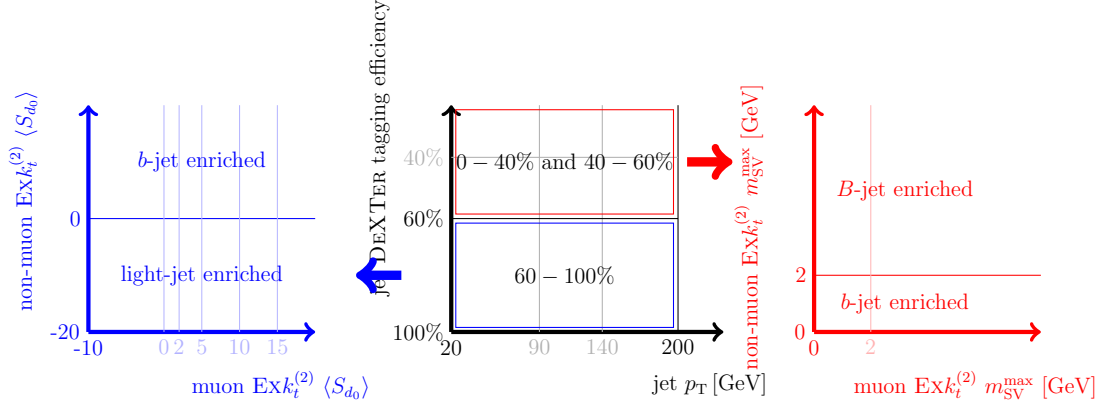


Figure 5.10: Diagram of the event selection and channel classification in the Z -region.

jet. The expected value for the Poisson probability distributions is built from all the simulated samples described in Section 4.2.

A series of multipliers are introduced:

Flavor fraction (μ) Different multipliers are introduced for b - and light-flavored jets, separately for the top-region and the Z -region, and for each p_T -bin. These multipliers provide a data-driven normalization of the corresponding flavor components in the phase space used for the efficiency measurement.

Efficiency scale factors (SF) Scale factors represent the ratio between the efficiency in data to the one in simulation. Different multipliers are introduced for b - and B -flavored jets, separately for each p_T bin and each tagging interval. The SF multipliers are defined so that the total number of predicted events is unchanged by efficiency scale factors.

Therefore, the expected value in each region r ($r = \text{top or } Z$) and channel c can be written as:

$$\begin{aligned}
\lambda_{rc} = & \mu_r^b \text{SF}_{0-40}^b n_{rc,0-40}^b + \mu_r^b \text{SF}_{40-60}^b n_{rc,40-60}^b + \mu_r^b \frac{(1 - \text{SF}_{0-40}^b \varepsilon_{0-40,\text{MC}}^b - \text{SF}_{40-60}^b \varepsilon_{40-60,\text{MC}}^b)}{(1 - \varepsilon_{0-40,\text{MC}}^b - \varepsilon_{40-60,\text{MC}}^b)} n_{rc,60-100}^b \\
& + \text{SF}_{0-40}^B n_{rc,0-40}^B + \text{SF}_{40-60}^B n_{rc,40-60}^B + \frac{(1 - \text{SF}_{0-40}^B \varepsilon_{0-40,\text{MC}}^B - \text{SF}_{40-60}^B \varepsilon_{40-60,\text{MC}}^B)}{(1 - \varepsilon_{0-40,\text{MC}}^B - \varepsilon_{40-60,\text{MC}}^B)} n_{rc,60-100}^B \\
& + \mu_r^{\text{light}} n_{rc}^{\text{light}} + \sum_{i \notin \{b,B,\text{light}\}} n_{rc}^i.
\end{aligned} \tag{5.4}$$

All the parameters in Eq. 5.4 depend on the jet p_T bin, even though no additional index was introduced to simplify the notation. The MC yields n are corrected with dedicated MC-MC scale factors when simulated samples are generated with different programs. The tagging efficiency in simulation depends on the detailed description of the hadronization of heavy-flavored partons and the resonant decays of hadrons inside jets [153]. The results in this note use the efficiency observed in PYTHIA 8 as reference as the scale factors.

Systematic uncertainties are introduced with constrained nuisance parameters. The nuisance parameters representing experimental and modeling uncertainties are accompanied by Gaussian constraint terms. In contrast, those representing statistical variations arising from the limited number of simulated events are accompanied by Poisson constraint terms. The values of all nuisance parameters and flavor fractions μ are determined by building a profile likelihood. The measured value and covariance for the several scale factors SF are determined by their maximum likelihood estimators.

5.5.3 Systematic uncertainties

Several sources of uncertainties are considered in the efficiency measurement. They can be loosely divided into four categories: experimental uncertainties from object reconstruction and identification, uncertainties from simulation modeling, statistical uncertainties from Monte Carlo simulation, and extrapolation uncertainties due to the use of soft-muons to select jets used in the efficiency measurement.

Experimental uncertainties

For each object, a set of reconstruction uncertainties are included. Uncertainties associated with leptons arise from the trigger, reconstruction, identification, and isolation efficiencies [84, 154, 155], as well as the lepton momentum scale and resolution [83, 155]. Uncertainties related to jet reconstruction stem from the efficiency of pileup rejection by the jet vertex tagger [156], the jet energy scale [157] and resolution [158].

Since DEXTER uses detailed information from tracks in the feature-extracting NNs and to define the channels in the Z -region, tracking reconstruction uncertainties are included. Uncertainties related to the track selection efficiency and the number of fake tracks are considered. Finally, dedicated systematic uncertainties are considered for the tracking parameters, including the transverse and longitudinal impact parameters and the track sagitta. The uncertainty on the integrated luminosity for the full Run-2 dataset is 1.7% [100], as obtained using the LUCID-2 detector [159]. A variation in the pileup modeling based on different estimates of the total inelastic cross section is included as an uncertainty.

Modeling uncertainties

Variations in the renormalization and factorization scales are used to estimate the uncertainty due to missing higher-order corrections. The combined PDF and α_s uncertainties follow the PDF4LHC prescription [160]. Additional modeling systematic uncertainties are applied to specific samples and are described below.

$Z + \text{jets}$ modeling uncertainties Additional modeling uncertainties on the $Z + \text{jets}$ MC prediction are considered, related to the description of the p_T distribution, the modeling of associated heavy-flavor production, and the choice of generator. The modeling of the Z boson p_T distribution is improved by reweighting the MC prediction to data using an inclusive selection. The full difference between the original MC prediction and the data is used as a systematic uncertainty. The uncertainty in

the heavy-flavor components is evaluated by comparing the nominal $Z + \text{jets}$ sample with alternatives with varied settings for the overlap between ME and Parton shower emissions and the resummation scale [161].

A generator uncertainty is also considered by comparing the nominal SHERPA 2.2.1 sample to the MADGRAPH5_AMC@NLO sample, as described in Section 4.2. The prediction difference is taken as a systematic uncertainty in the modeling of $Z + \text{jets}$.

$t\bar{t}$ modeling uncertainties For shower uncertainties, the settings of the nominal POWHEGBOX+PYTHIA8 $t\bar{t}$ (5FS) sample are varied, resulting in different event weights; the uncertainty due to initial shower radiation is estimated by simultaneously changing the renormalization and factorization scale in the ME and in the parton shower, while the uncertainty due to final state shower radiation is estimated by changing the associated scale in the simulation. For the uncertainties due to hadronization and NLO matching, the nominal POWHEGBOX+PYTHIA8 $t\bar{t}$ (5FS) sample is compared to the POWHEGBOX+HERWIG7 $t\bar{t}$ (5FS) sample mentioned in Section 4.2.

Single-top modeling uncertainties Systematic uncertainties for the single top MC modeling follow the prescription in $t\bar{t}$ modeling uncertainties. On top of these uncertainties, the interference between the single top and $t\bar{t}$ production is considered by comparing the diagram removal and diagram subtraction schemes.

Extrapolation systematic uncertainties

Systematic variations associated to the use of a soft-muon to select jets in the Z -region are estimated by comparing, for each tagging interval, the efficiency using SHERPA 2.2.1 and MADGRAPH5_AMC@NLO. This additional uncertainty is added to account for any mis-modeling in the efficiency of selecting jets with soft-muons and in the correlation of this efficiency with the DEXTER efficiency. The relative differ-

Jet label	Jet p_T range (GeV)	Tagging Interval	Scale Factor
B -label	[20, 90]	0 – 40%	1.15 ± 0.05 (stat.) ± 0.26 (syst.)
B -label	[90, 140]	0 – 40%	1.00 ± 0.10 (stat.) ± 0.14 (syst.)
B -label	[140, 200]	0 – 40%	1.12 ± 0.16 (stat.) ± 0.16 (syst.)
B -label	[20, 90]	40 – 60%	1.77 ± 0.12 (stat.) ± 0.46 (syst.)
B -label	[90, 140]	40 – 60%	1.13 ± 0.24 (stat.) ± 0.31 (syst.)
B -label	[140, 200]	40 – 60%	1.78 ± 0.34 (stat.) ± 0.34 (syst.)
b -label	[20, 90]	0 – 40%	1.33 ± 0.04 (stat.) ± 0.18 (syst.)
b -label	[90, 140]	0 – 40%	1.91 ± 0.12 (stat.) ± 0.13 (syst.)
b -label	[140, 200]	0 – 40%	1.88 ± 0.11 (stat.) ± 0.08 (syst.)
b -label	[20, 90]	40 – 60%	1.17 ± 0.03 (stat.) ± 0.13 (syst.)
b -label	[90, 140]	40 – 60%	1.53 ± 0.09 (stat.) ± 0.13 (syst.)
b -label	[140, 200]	40 – 60%	1.43 ± 0.07 (stat.) ± 0.09 (syst.)

Table 5.3: Ratio of the efficiency measured in data to MC for different jet labels, p_T ranges, and tagging intervals.

ence between the efficiency with and without soft-muon tagging in the two generators is used as an additional source of systematic uncertainty. This relative difference varies between 1 – 12% depending on the tagging interval and p_T range.

5.5.4 Results

The measured DEXTER efficiencies for B - and b -labeled jets in the two tagging intervals are shown in Figure 5.11 and Figure 5.12. The values are also summarized in Table 5.3. The statistical and total uncertainty are shown separately. The SFs for B -labeled jets in the 0 – 40% tagging interval are consistent with unity and uncertainties less than 20% showing the robustness of the choices made during the algorithm development. At low p_T , the uncertainty in this measurement is dominated by modeling uncertainties.

Figure 5.13 compares the simulated and observed data in the top-regions. Figure 5.14 shows a summary comparison in the Z -region. Figure 5.15 shows an example of $\langle S_{d_0} \rangle$ and $m_{SV}^{\max}$ distributions for two different bins of Figure 5.14. Ta-

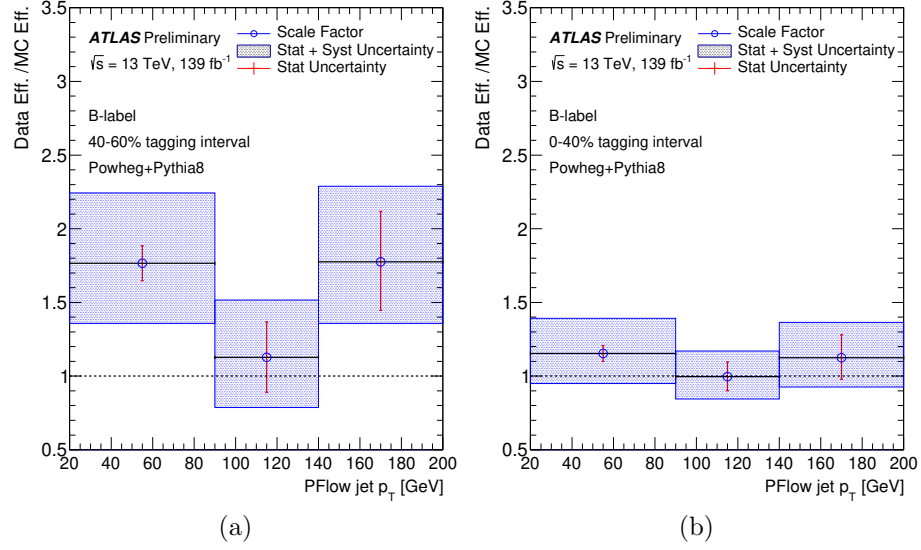


Figure 5.11: Data-to-MC scale factors obtained for B -labeled jets in the (a) 40 – 60% and (b) 0 – 40% tagging intervals. The blue error band includes systematic and statistical uncertainties. The red error bar represents the statistical uncertainty only.

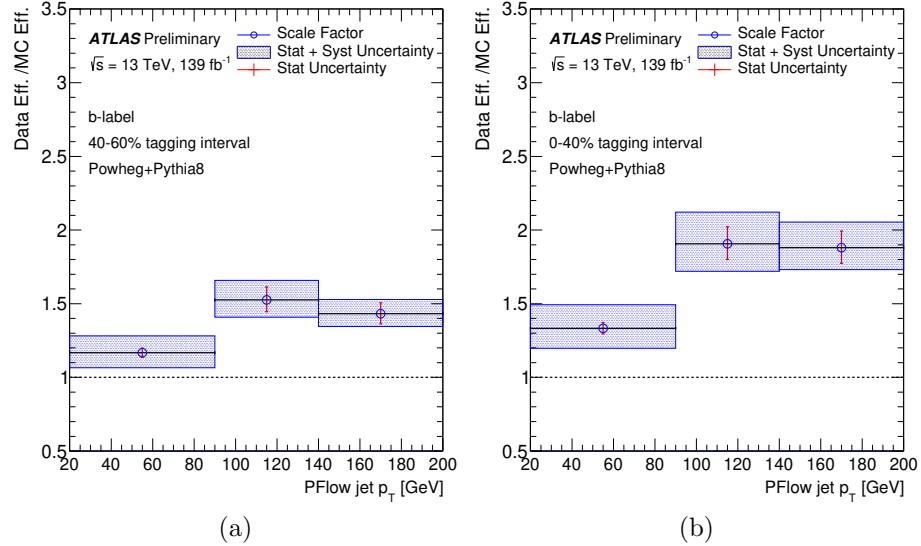


Figure 5.12: Data-to-MC scale factors obtained for b -labeled jets in the (a) 40 – 60% and (b) 0 – 40% tagging intervals. The blue error band includes systematic and statistical uncertainties. The red error bar represents the statistical uncertainty only.

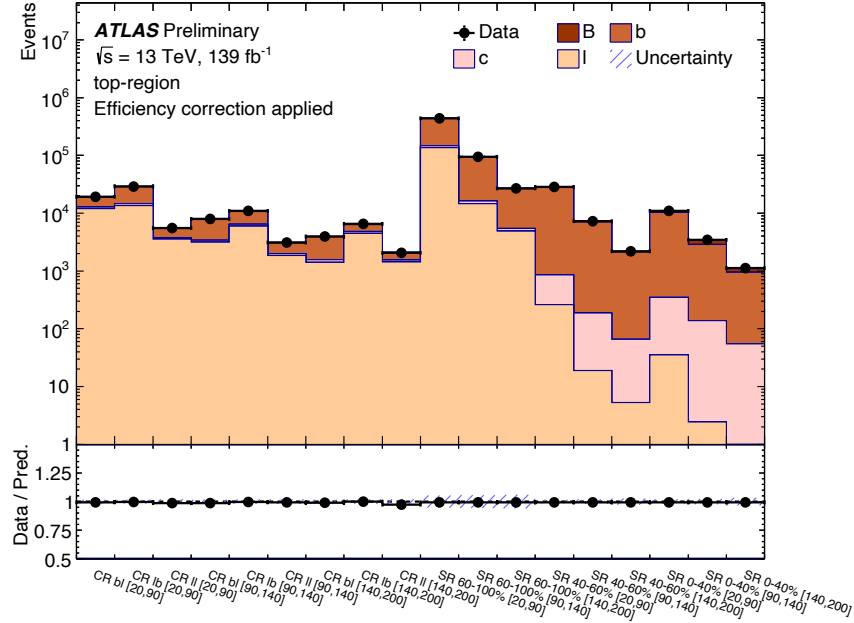


Figure 5.13: Comparison between data and prediction for the event yields in the top-region channels after the efficiency correction. The hatched band represents the systematic uncertainty in each channel.

bles 5.4 and 5.5 show a breakdown of the systematic uncertainties in grouped categories for B- and b-labeled jets, respectively.

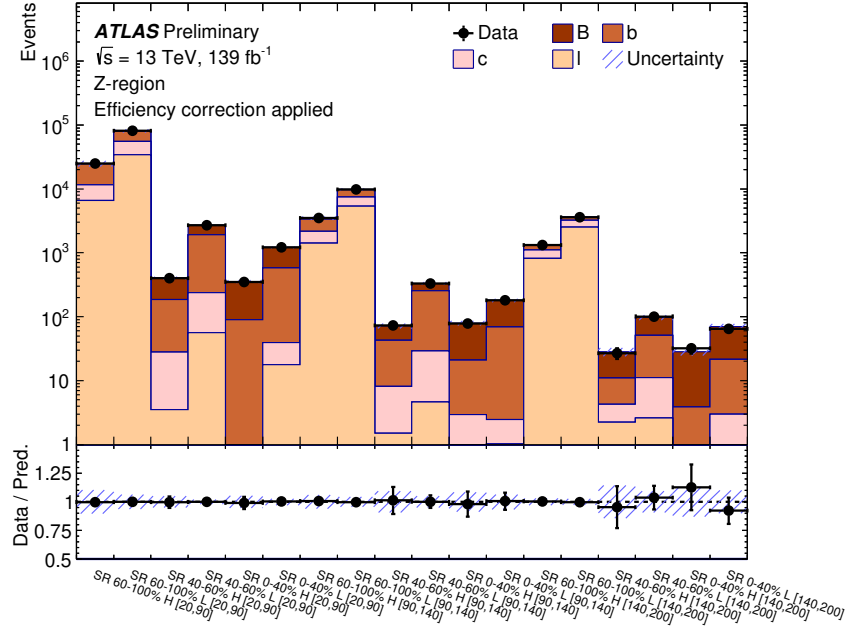


Figure 5.14: Comparison between data and prediction for the event yields in the Z-region channels after the efficiency correction. The labels H and L indicate the two different channels depending on the value of the non-muon- $\text{Ex}k_t^{(2)}$ track-subjet $m_{\text{SV}}^{\text{max}}$ (H : $m_{\text{SV}}^{\text{max}} \geq 2 \text{ GeV}$, L : $m_{\text{SV}}^{\text{max}} < 2 \text{ GeV}$) or $\langle S_{d_0} \rangle$ (H : $\langle S_{d_0} \rangle \geq 0$, L : $\langle S_{d_0} \rangle < 0$). The hatched band represents the systematic uncertainty in each channel.

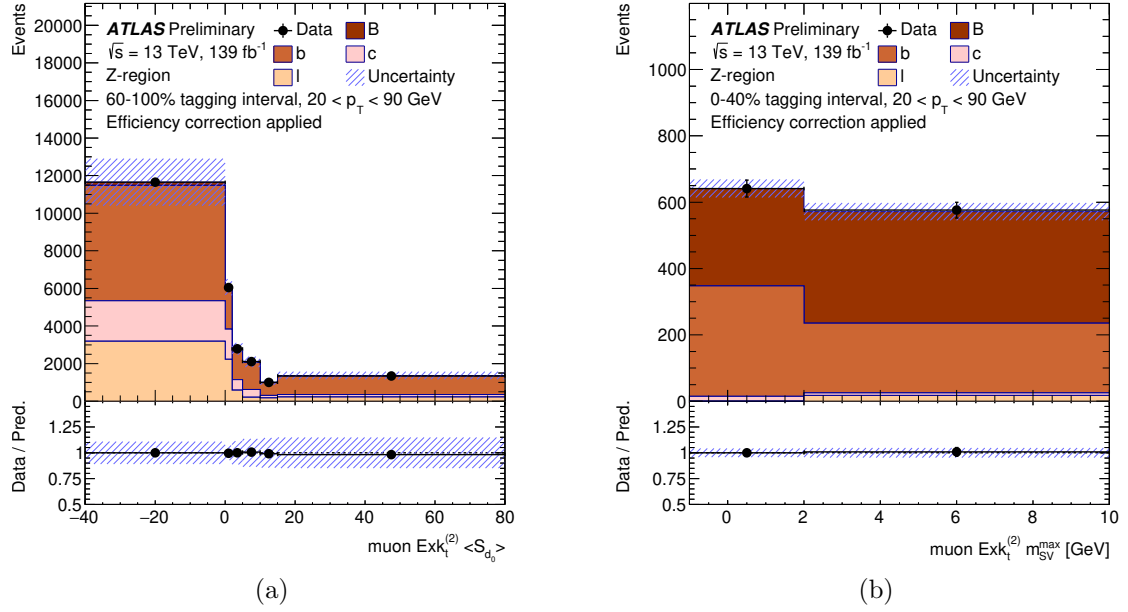


Figure 5.15: Comparison between data and prediction in the Z-region, after efficiency correction, for (a) the muon $\text{Ex}k_t^{(2)} \langle S_{d_0} \rangle$ distribution in the 60 – 100% tagging interval when requiring the non-muon $\text{Ex}k_t^{(2)} \langle S_{d_0} \rangle > 0$ and (b) the muon $\text{Ex}k_t^{(2)} m_{SV}^{\text{max}}$ distribution in the 0 – 40% tagging interval when requiring the non-muon $\text{Ex}k_t^{(2)} m_{SV}^{\text{max}} < 2 \text{ GeV}$.

Source of uncertainties	0 – 40% tagging interval			40 – 60% tagging interval		
	p_T^{jet} range (GeV)			p_T^{jet} range (GeV)		
	[20, 90]	[90,140]	[140, 200]	[20, 90]	[90,140]	[140, 200]
Total uncertainty	0.26	0.17	0.23	0.48	0.39	0.48
Statistical	0.05	0.10	0.16	0.12	0.24	0.34
Systematic	0.26	0.14	0.16	0.46	0.31	0.34
Extrapolation	0.21	0.11	0.13	0.38	0.25	0.23
MC statistics	0.21	0.14	0.10	0.36	0.20	0.25
Jets	0.14	0.05	0.04	0.16	0.09	0.11
Leptons	0.02	0.01	0.01	0.05	0.02	0.01
Luminosity	0.01	0.02	0.02	0.06	0.03	0.02
Flavor fraction	0.19	0.10	0.12	0.30	0.20	0.26
Pileup	0.04	0.01	0.01	0.05	0.01	0.04
Tracking	0.16	0.04	0.05	0.19	0.11	0.09
Single top modeling	0.01	0.02	0.01	0.12	0.05	0.03
$t\bar{t}$ modeling	0.12	0.02	0.02	0.19	0.11	0.06
Z + jets modeling	0.22	0.12	0.14	0.39	0.22	0.25

Table 5.4: Breakdown of the SF uncertainties for each summary category in different tagging intervals and p_T range for B -labeled jets. The statistical uncertainties are also included for comparison. Due to correlations, the total systematic uncertainty is not equal to the sum in quadrature of the individual components.

Source of uncertainties	0 – 40% tagging interval			40 – 60% tagging interval		
	p_T^{jet} range (GeV)			p_T^{jet} range (GeV)		
	[20, 90]	[90,140]	[140, 200]	[20, 90]	[90,140]	[140, 200]
Total uncertainty	0.18	0.18	0.14	0.13	0.16	0.11
Statistical	0.04	0.12	0.11	0.03	0.09	0.07
Systematic	0.18	0.13	0.09	0.13	0.13	0.09
Extrapolation	0.18	0.12	0.07	0.13	0.12	0.07
MC statistics	0.13	0.04	0.02	0.10	0.04	0.02
Jets	0.07	0.04	0.03	0.06	0.04	0.03
Leptons	0.04	0.01	0.00	0.03	0.01	0.00
Luminosity	0.04	0.01	0.01	0.03	0.01	0.01
Flavor fraction	0.17	0.11	0.06	0.12	0.11	0.06
Pileup	0.05	0.02	0.00	0.04	0.02	0.00
Tracking	0.12	0.04	0.02	0.09	0.04	0.02
Single top modeling	0.14	0.03	0.04	0.09	0.03	0.04
$t\bar{t}$ modeling	0.17	0.07	0.03	0.13	0.07	0.03
Z + jets modeling	0.18	0.12	0.07	0.13	0.12	0.07

Table 5.5: Breakdown of the SF uncertainties for each summary category in different tagging intervals and p_T range for b -labeled jets. The statistical uncertainties are also included for comparison. Due to correlations, the total systematic uncertainty is not equal to the sum in quadrature of the individual components.

5.6 Summary

This chapter presents the development of DEXTER, an algorithm that performs flavor tagging of low- p_T jets arising from the hadronization of one or two heavy-flavor hadrons. The algorithm uses a neural network with two permutation-invariant deep-set feature extractors, one for displaced tracks and one for secondary vertices. Displaced vertices are selected from a large- R track jet reconstructed around PFlow jets. Secondary vertices are reconstructed with the TC-LVT clustering algorithm and the MSVF vertex finding algorithm.

The performance is studied in detail, ensuring a reduced dependency of the tagger performance with the parent particle color charge and mass. A measurement of the efficiency for two tagging intervals and three p_T^{jet} ranges using $t\bar{t}$ and $Z + \text{jets}$ events are reported for B - and b -labeled jets. These measurements make applying this algorithm in physics analyses possible, enabling and improving several new BSM searches and SM measurements. In the next chapter, I will demonstrate a use case where DEXTER greatly improves the search sensitivity of $H \rightarrow aa \rightarrow 4b$, specifically for low-mass $a \rightarrow bb$.

CHAPTER 6

SEARCH FOR EXOTIC HIGGS BOSON DECAY TO MULTIPLE B-QUARKS

In this chapter, I will focus on a search for exotic Higgs decay to multiple b-quarks using the full Run-2 data collected from 2015 to 2018 at $\sqrt{s} = 13$ TeV with the ATLAS detector at LHC corresponding to an integrated luminosity of 140 fb^{-1} . Similar searches were done in 2017 and 2020 using the data collected with the ATLAS detector from 2015 and 2016 corresponding to an integrated luminosity of 36 fb^{-1} . The search presented in this chapter is based on both prior experiences. The strategy is redesigned in the last iteration of the analysis. Using state-of-art machine-learning techniques and an almost fourfold increase in the dataset, we can set a stronger upper limit at 95% confidence level on the branching ratio of exotic Higgs decays $\text{BR}(H \rightarrow aa \rightarrow (b\bar{b})(b\bar{b}))$. Furthermore, for the first time, we also consider benchmark models assuming two new scalars and set limits on $\text{BR}(H \rightarrow a_1 a_2 \rightarrow (b\bar{b})(b\bar{b}))$. The chapter is arranged as follows: Section 6.1 gives an overview of additional improvements in object reconstruction not mentioned in the previous chapters. Section 6.2 and Section 6.3 detail the event selection and analysis strategy with machine learning. The systematics considered are summarized in Section 6.4. The statistical method is described in Section 6.5. Section 6.6 presents and discusses the results of the search. This chapter concludes with Section 6.7 which discusses the opportunities for future searches.

6.1 Objects selection

The standard object reconstruction and selection have already been described in Chapter 3. This section focuses on the additional technique developed to improve object reconstruction for the search. One particularly important area is the kinematic correction for heavy-flavor jets. Different methods are applied on jets tagged by DeXTer or DL1r. Jets tagged by DeXTer are referred to as **B-jets** in the following sections. Jets tagged by DL1r are referred to as **b-jets**. In the case of B-jets, the energy-momentum estimation is derived solely from tracks that only consider the contribution from charged particle tracks. In Section 6.1.1, a dedicated NN regression is designed to infer the energy of the jet at the particle level, including the neutral particle contribution and soft muon.

Charged muons and neutrinos from the leptonic decays of the b-quark will escape from detection that solely relies on the calorimeter measurements. A straightforward approach is pursued for b-jets to account for the underestimation of b-jet energy from leptonic decays. This method is described in Section 6.1.2.

6.1.1 DNN energy-momentum estimate regression

As shown in Ref. [162], the large- R track jet provides improved containment when the jet originates from the hadronization of two particle jets. Any B-jet has a single PFlow jet as a seed, and the standard ATLAS jet calibration of that PFlow jet provides an unbiased estimate of the jet’s momentum at the particle level. However, this calibration procedure ignores the increased leakage from the poor containment when two particle jets are present.

The tracks in the region $0.4 < R < 0.8$ provide information that can be used to create a higher-resolution estimate of the B-jet momentum, closer to the momentum of the two-particle b -jets. The procedure developed in this section is an energy-momentum estimate used to perform searches. It provides improved resolution for

jets similar to the ones produced in $VH \rightarrow aa \rightarrow 4b$ signals but should never be interpreted as the real energy momentum of the particle jet.

The large- R track jet mass is defined by the four-vector sum of tracks associated to the large- R track jet. The energy-momentum estimate uses the information from the seed PFlow jet, the tracks of the large- R track jet, and non-isolated muons inside the large- R track jet. The estimate is built as a regression that targets the momentum of truth **large-R trimmed particle jets** that is ΔR -matched to the B-jets. The large- R particle jets are built with the anti- k_t algorithm [87] with a radius parameter of $R = 0.8$ to cluster truth particles into large- R particle jets and apply jet trimming [163] on the resulting truth particle jet collection. These jets are required to be isolated ($\Delta R > 1.0$) from any other truth jet or lepton in the event to avoid bias in the momentum determination.

The NN topology and inputs are very similar to the ones used for flavor tagging with DeXTer [162], and a similar method is applied in b-jet energy and resolution estimates in CMS [164]. Some additional features provide information about the energy distribution inside the B-jet. Kinematic information related to the two $\text{Ex}k_t^{(2)}$ track subjets and non-isolated muons inside the jet are added using the same deep-set topology [165] used for tracks and secondary vertices. A complete set of variables used for the training can be found in Table. 6.1. The training, validation, and testing samples are produced by selecting B-jets from $a \rightarrow bb$ signals and B-jets from Z+jets with $g \rightarrow bb$ MC samples. A more detailed description of the datasets and hyperparameters can be found in Appendix A.1.

The regression has one output neuron, the estimation correction factor $\hat{s}$ of the mass :

$$\hat{s} = \frac{m_{\text{gen}}}{m_{\text{reco}}} \quad (6.1)$$

Feature	Description
Consituent R=0.4 PFlow Jet	
p_T	Jet transverse momentum
η	Jet pseudorapidity
ϕ	Jet azimuthal angle
$\Delta R(\text{Ex}k_t^{(2)} \text{ jet 1}, \text{Ex}k_t^{(2)} \text{ jet 2})$	Angular separation between $\text{Ex}k_t^{(2)}$ jets
Exclusive- k_t jet	
$\log p_T^{\text{frac}}(\text{Ex}k_t^{(2)} \text{ jet}, \text{PFlow jet})$	$\log p_T^{\text{Ex}k_t^{(2)}} / p_T^{\text{PFlow jet}}$: transverse momentum fraction between $\text{Ex}k_t^{(2)}$ jet and jet
$\Delta\eta(\text{Ex}k_t^{(2)} \text{ jet}, \text{PFlow jet})$	Pseudorapidity difference between $\text{Ex}k_t^{(2)}$ jet and PFlow jet
$\Delta\phi(\text{Ex}k_t^{(2)} \text{ jet}, \text{PFlow jet})$	Angular difference between $\text{Ex}k_t^{(2)}$ jet and PFlow jet
$\log(m)$	Mass of the track jet
$n_{\text{tracks}}^{\text{Ex}k_t^{(2)}}$	Number of tracks ghost-associated to the $\text{Ex}k_t^{(2)}$ jet
Track	
$\log p_T^{\text{frac}}(\text{track}, \text{PFlow jet})$	$\log p_T^{\text{track}} / p_T^{\text{PFlow jet}}$: transverse momentum fraction between track and R=0.4 PFlow jet
$\Delta\eta(\text{track}, \text{Ex}k_t^{(2)} \text{ jet})$	Pseudorapidity difference between track and $\text{Ex}k_t^{(2)}$ jet
$\Delta\phi(\text{track}, \text{Ex}k_t^{(2)} \text{ jet})$	Angular difference between between track and $\text{Ex}k_t^{(2)}$ jet
d_0	Transverse impact parameter
$z_0 \sin \theta$	Longitudinal impact parameter
S_{d_0}	d_0/σ_0 : transverse IP significance
$S_{z_0 \sin \theta}$	$z_0 \sin \theta / \sigma_{\sin \theta}$: longitudinal IP significance
PIX1 Hits	Number of hits in first Pixels layer
IBL Hits	Number of hits in IBL
Shared IBL Hits	Shared hits in IBL
Split IBL Hits	Number of split hits in IBL
Shared pixel Hits	Shared hits in Pixels layers
Split pixel Hits	Shared hits in Pixels layers
Shared SCT Hits	Number of shared hits in SCT
nPixHits	Number of hits in Pixel layers
nSCTHits	Number of hits in SCT layers
Secondary Vertex	
$\log(m)$	Track mass of the secondary vertex
$\log p_T^{\text{frac}}(\text{vertex}, \text{PFlow jet})$	$\log p_T^{\text{SecVtx}} / p_T^{\text{PFlow jet}}$: transverse momentum fraction between secondary vertex and R=0.4 PFlow jet
$\Delta\eta(\text{vertex}, \text{Ex}k_t^{(2)} \text{ jet})$	Pseudorapidity difference between secondary vertex and $\text{Ex}k_t^{(2)}$ jet
$\Delta\phi(\text{vertex}, \text{Ex}k_t^{(2)} \text{ jet})$	Angular difference between between secondary vertex and $\text{Ex}k_t^{(2)}$ jet
L_{xy}	Transverse decay length relative to primary vertex
L_z	Longitudinal decay length relative to primary vertex
$S_{L_{xy}}$	Transverse decay length significance
S_{L_z}	Longitudinal decay length significance
Muons	
$\log p_T^\mu$	Transverse momentum of the muon
$\Delta\eta(\text{muon}, \text{Ex}k_t^{(2)} \text{ jet})$	Pseudorapidity difference between muon and $\text{Ex}k_t^{(2)}$ jet
$\Delta\phi(\text{muon}, \text{Ex}k_t^{(2)} \text{ jet})$	Angular difference between between muon and $\text{Ex}k_t^{(2)}$ jet
Quality	Integer representing muon quality

Table 6.1: List of features used as input for the DNN mass estimate regression. In

Here m_{gen} is truth large-R trimmed particle jets mass, and m_{reco} is the original B-jet mass. The target correction factor s can be calculated from MC simulation.

A Huber loss function is defined as

$$H_{\delta}(z) = \begin{cases} \frac{1}{2}z^2, & \text{for } |z| \leq \delta \\ \delta \cdot |z| - \frac{1}{2}\delta^2, & \text{otherwise} \end{cases} \quad (6.2)$$

where $z = s - \hat{s}$.

The complete loss function uses the Huber loss with $\delta = 1$ and is written as:

$$\text{loss}(\hat{s}) = E_{(x,s) \sim p(x,s)}[H_1(s - \hat{s})], \quad (6.3)$$

where $E_{(x,s) \sim p(x,s)}$ is the expectation value, x denotes the set of input features, and $p(x, s)$ is the joint distribution of the input features and the target variables $\hat{s}$ in the training sample.

The resulting resolution can be examined using the reciprocal of $\hat{s}$ shown in Figure. 6.2. The distribution peak is slightly below one because m_{reco} only contains charged particles, whereas the m_{gen} also includes energy-momentum from neutral particles. Figure. 6.3 shows the comparison of B-jet mass before and after applying the DNN energy-momentum estimate regression. The corrected B-jet mass can be obtained by multiplying the original B-jet mass with the estimation factor $\hat{s}$ for each jet, $m_{\text{DNN}} = m_{\text{reco}} \times \hat{s}$. The $a \rightarrow b\bar{b}$ jet mass can be predicted, which is used to improve the signal-background discrimination. When comparing the simulated track jet mass scale to the one observed in data, a small mis-modeling is observed. A correction to the simulated track jet mass scale is derived from data. The resulting track jet mass correction scale is 1.01 ± 0.05 .

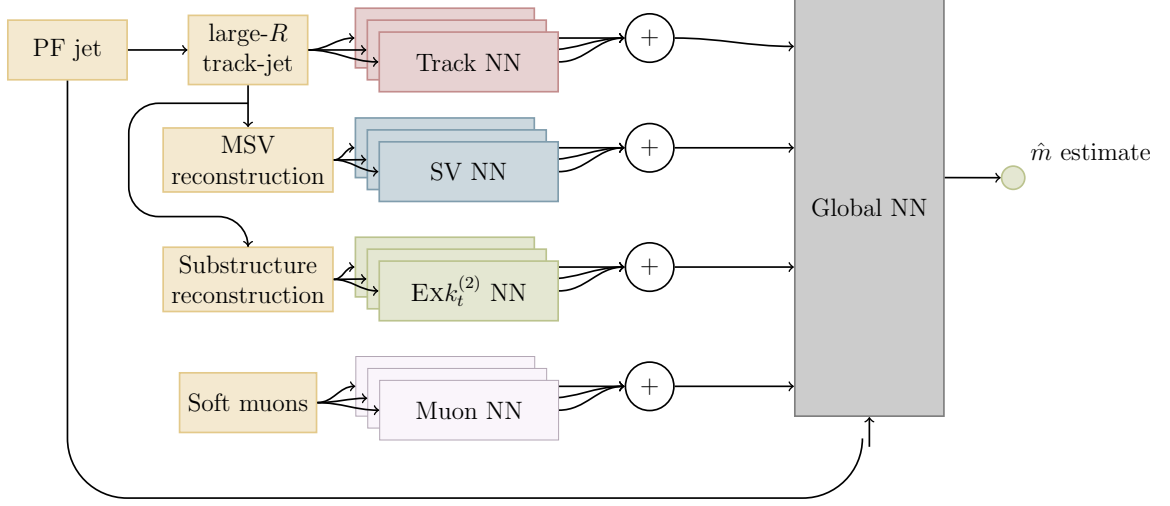


Figure 6.1: DNN mass estimate regression architecture

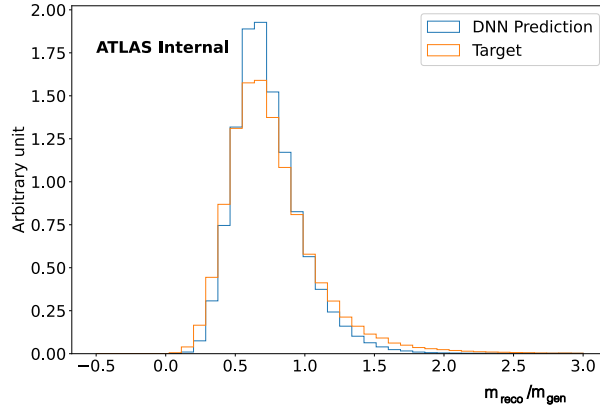


Figure 6.2: Comparison of the target and the predicted mass resolution distribution from the DNN energy-momentum estimation regression. The target distribution in orange represents the target of the DNN during training. The DNN prediction is the predicted $1/\hat{s}$ by the trained DNN.

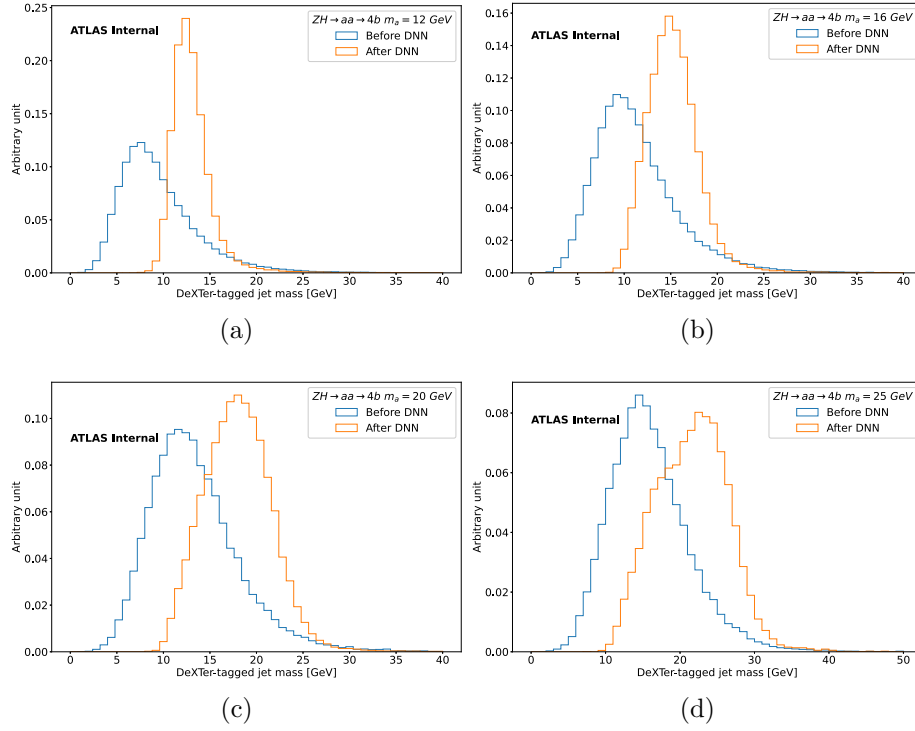


Figure 6.3: The large- R track jet mass comparison before and after the DNN Energy-momentum estimation regression.

6.1.2 mu-in-jet p_T correction

For DL1r b-tagged jets, the mu-in-jet p_T correction is applied. The soft muon requirements are defined as $p_T > 4$ GeV, $|\eta| < 2.5$, and the Medium muon quality requirement. The jet matching procedure, however, is slightly different: muons are added to the b-jet four-momentum if $\Delta R(\text{jet}, \text{soft muon}) < 0.4$. If the soft muons satisfy the requirements, at most, the two leading ones are considered.

6.2 Event selection

Event Preselection

Events satisfying the trigger selection are required to have exactly two reconstructed electrons or muons with leading and subleading lepton satisfying $p_T > 27$ GeV and $p_T > 10$ GeV, respectively. Since single lepton triggers are used, one of the two reconstructed leptons is required to match the lepton reconstructed at trigger level, i.e., they have to be within $\Delta R < 0.15$ of each other. The invariant mass of the two leptons is required to be larger than 50 GeV to reduce the contribution of the $t\bar{t}$ and $Z \rightarrow \tau\tau$ background.

To improve the purity in dilepton events, leptons are further required to satisfy additional identification and isolation criteria, as described in Sections 3.1 and 3.2.

The following pre-selection is applied to further suppress the background based on a combination of b -objects:

$$(2 \times N_{B\text{-jets}} + N_{b\text{-jets}} + N_{\text{SV}}) \geq 3, \quad (6.4)$$

where $N_{B\text{-jets}}$ is the number of DeXTer-tagged B-jets, $N_{b\text{-jets}}$ is the number of DL1r-tagged b-jets, and N_{SV} is the number of soft secondary vertices reconstructed by the TC-LVT algorithm.

Figure 6.4 shows the contribution of different SM processes after applying the pre-selection. The events are categorized based on the categorization defined in Sec-

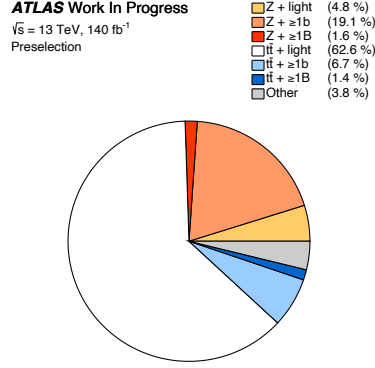


Figure 6.4: The SM background compositions after applying the pre-selection.

tion 4.2.2. The dominant SM backgrounds are from $Z + \text{jets}$ and $t\bar{t} + \text{jets}$, which have larger cross-sections compared to other small SM backgrounds, such as the $t\bar{t}V$ and $t\bar{t}H$ backgrounds. This pre-selection is introduced as many studies were conducted at the pre-selection level. The event selection and definition of signal and control regions of the analysis are described in Section 6.3. Note that all events selected for the signal and control regions will be required to have at least 4 b -objects. A set of data-driven corrections for the two main backgrounds is derived and described in Section 4.2. Event selections for the regions where the corrections are derived are described there.

6.3 Analysis strategy

With the object and event selection defined, this section describes the analysis strategies deployed to search for exotic Higgs decays $H \rightarrow aa/a_1a_2$ in the $4b$ final states. For the signal event with four b hadrons in the final state, a large fraction can be fully reconstructed. This allows the reconstruction of the full $H \rightarrow aa/a_1a_2 \rightarrow 4b$ decay chain. Here, the signal reconstruction and background discrimination are done in two steps. In the first step, a parameterized NN is adapted to identify a pairing hypothesis, to pair the b -objects from the same parent particle. This allows us to

define discriminating variables, e.g., invariant masses, based on the best pairing. The detailed are given in Section 6.3.1. Those variables are then used to train a boosted decision tree (BDT) as the final signal-background discriminator. The signal reconstruction and background discrimination of the $4b$ final state is described in Section 6.3.2.

6.3.1 Event reconstruction using a deep neural network

The heterogeneous nature of the object reconstructions used in this analysis required a novel method for event reconstruction. The kinematics constraint from the reconstructed Higgs boson and a -boson candidates are critical variables in selecting the signal and differentiating it from the SM backgrounds. To properly reconstruct the two a -bosons, a method is needed to predict the pairing of the reconstructed b-objects to form the $a \rightarrow b\bar{b}$ decays. For this, a NN will be used. In the case of $H \rightarrow aa$, the two a -boson decays from the Higgs boson are indistinguishable particles, and a similar argument applies to two b-hadrons in the $a \rightarrow b\bar{b}$ decays. This domain knowledge will be incorporated into the design of the NN architecture to predict the correct pairing hypothesis. In the case for $H \rightarrow a_1a_2$, where the mass of the new scalars is different, the design of the following NN was found to be robust in reconstructing the $H \rightarrow a_1a_2 \rightarrow 4b$ decay chain where the same property still holds in the $a_1/a_2 \rightarrow b\bar{b}$ decays.

The idea of the pairing hypothesis can be demonstrated with one example event shown in Figure 6.5. This event has one b-jet, one jet that failed b-tagging, one B-tagged jet by DeXTer, and one soft-b hadron reconstructed as a secondary vertex. In this case, two possible ways exist for such a signal event to pair them into the correct a-bosons. The DeXTer jet could either capture the $a \rightarrow b\bar{b}$ decay or the two b-hadrons from different a-boson, which can be studied by examining the truth record from MC simulation. Considering cases where the DeXTer jet captures two

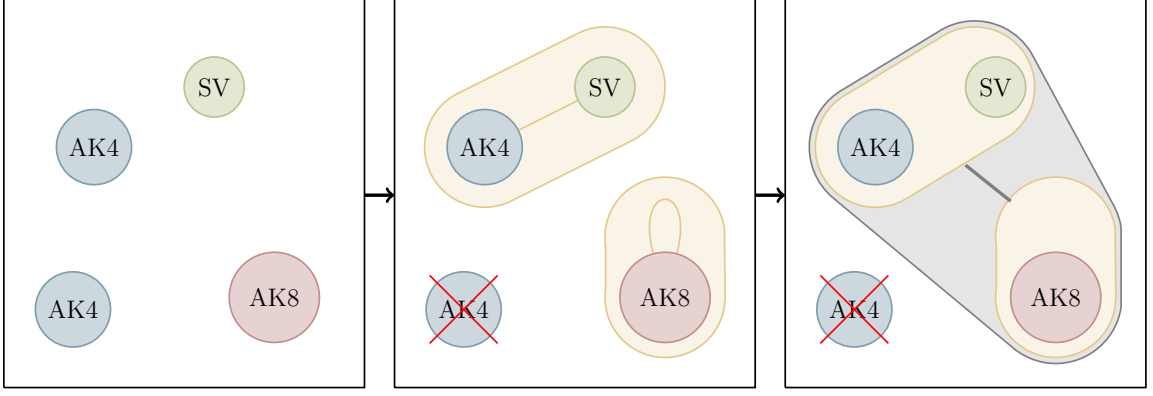


Figure 6.5: Diagram depicting how the NN performs event reconstruction. Each b candidate is considered a node in a graph for which an abstract representation is used based on a node MLP. The a -boson is represented by an edge MLP between nodes. All physically motivated edges are considered, including loop edges representing merged reconstructed objects. The a -bosons themselves form new nodes whose edge MLP represents the fully reconstructed Higgs boson. The value of this edge is used for event reconstruction. The middle diagram shows the two reconstructed a -boson candidates and the rightmost diagram shows the final Higgs candidate.

b -hadrons from different a -boson is especially important for high signal mass points. The random crossing of two b hadrons from different a bosons becomes more likely with increasing angular separation of $a \rightarrow b\bar{b}$ decays.

As mentioned above, the $4b$ final state can be reconstructed in many possible combinations of b -object in the detector, with varying numbers of the reconstructed b -objects (e.g. 1 B -jet and two b -jets vs. four b -jets). This causes an issue with the conventional NN techniques, which require fixed-sized inputs. DeepSet [165] is an advanced NN architecture that can handle variable-length input vectors similar to the recursive neural network (RNN). Furthermore, this allows the abstraction of different b -object reconstruction information into a common latent space vector that can handle different types of reconstructed b -objects. Using this architecture, a unified strategy can be deployed for all possible combinations of reconstructed objects.

The hypothesis testing NN architecture is shown in Figure 6.6. It can be divided into three parts, the object encoder, the a -boson reconstruction, and the Higgs recon-

struction. The architecture stacked two DeepSet architectures to enforce permutation invariance of identical particles in the $H \rightarrow aa \rightarrow b\bar{b}b\bar{b}$ decay chain. All input features to the object encoders (and the Higgs reconstruction network) are summarized in Table 6.2. The objects are encoded into common-size latent space vectors with dimension N , which are later propagated to a permutation invariant sum pooling layer. The aggregated information goes through a parametrized shared weight a -boson NN to learn the latent space representation of the a -boson. The a -boson NN is parametrized by an extra parameter, m_a , representing the a -boson mass of the tested hypothesis. The last sum pooling layer enforces the permutation invariance between the two a -bosons. The summed a -boson vector is concatenated with extra information from the Z kinematics as input to the final Higgs NN. The Z kinematics provide additional information on the Lorentz-boosted Higgs decay. A sigmoid function is used in the output layer to predict the score of a given pairing hypothesis presented to the hypothesis testing NN. A correct truth hypothesis will have a pairing hypothesis score $\hat{y}_{\text{truth}} = 1$, whereas an incorrect pairing hypothesis based on the truth record will have a pairing hypothesis score of $\hat{y}_{\text{truth}} = 0$.

The training dataset is generated using signal and SM background MC simulation. For each event in the simulation, the correct pairing hypothesis can be known from the MC simulation truth information. All possible pairing hypotheses for a given event will be presented to the NN. Only the correct pairing hypothesis is labeled as $\hat{y}_{\text{truth}} = 1$, while all other pairing hypotheses will be assigned a $\hat{y}_{\text{truth}} = 0$ label. The same method generates synthesis datasets from the SM background sample. However, given that these events are not from the signal $ZH, H \rightarrow 2a \rightarrow 4b$ process, all the synthesis datasets from the SM background will be labeled as $\hat{y}_{\text{truth}} = 0$. This was found to be helpful in increasing the overall dataset size and also improving the performance by forcing the NN to use information related to the Higgs boson kinematics.

The output of the NN is two-fold: the subsequent parts of the analysis take into account the hypothesis for the best pairing and a score value for this pairing (the pairing hypothesis score $\hat{y}_{\text{pred}}^{m_a}$). During inference time, the NN is applied to a selected list of possible pairing hypotheses event by event. The hypothesis with the highest pairing hypothesis score is the predicted pairing hypothesis for that event. The predicted pairing hypothesis allows us to identify how to pair the jet to parton and compute the relevant kinematical variables for the BDTs described in Section 6.3.2. The pairing hypothesis score provides an estimate of how confident we are a given event looks like a $ZH \rightarrow 4b$ event based on all the information from input features. It is used in the signal-background discrimination BDTs, as well as signal and control region definitions, as described in Sections 6.3.2 and 6.3.2. The score distribution for the SM background is shown in Figure. 6.7 assuming the $m_a = 25$ GeV signal hypothesis. The majority of the SM background events have a score $\hat{y}_{\text{pred}}^{m_a} < 0.05$, which motivates the choice of signal and control region definitions.

Performance per category

The performance for each truth category is shown in the following confusion matrixes. This is checked using a mixed signal sample containing all the m_a mass points. The following notation defines the corresponding pairing hypothesis predicted by the NN or from the truth record, *Reco object(ture or predicted pairing)*. The confusion matrix for fully merged events is shown in Figure 6.8, where the pairing task is relatively easy for the NN, with the true positive rate for a given truth category reaching 99%. Similar results can also be seen in the cases of semi-merged events. This category can be further divided into events with two b-tagged jets or one b-tagged jet and one secondary vertex with a true positive rate reaching 99%, shown in Figure. 6.9.

As for the fully resolved cases, the possible combination to form the a -boson candidate increases. The NN can still predict the correct pairing hypothesis with a

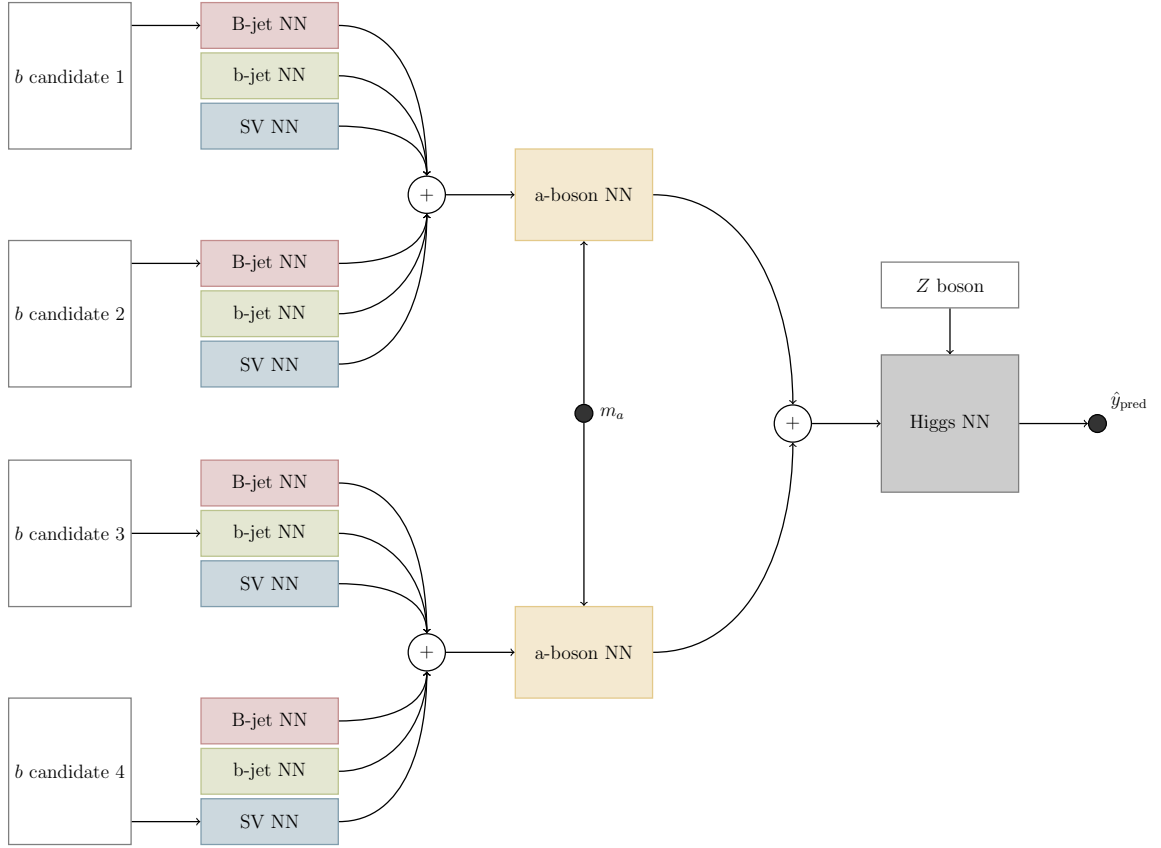


Figure 6.6: Diagram of NN used for event identification. All filled blocks are dense multilayer perceptions (MLP) that share common weights between MLPs of the same color. The a -boson MLPs are parametrized as a function of the mass hypothesis m_a . Unfilled rectangles are inputs from the four b candidates and the Z boson. The value of $\hat{y}_{\text{pred}}$ is used for graph classification. This particular example shows the case of a hypothesis with B-jet, b-jet, and soft-SV b -candidates.

Feature	Description
DeXTer-tagged B-jet	
$\log p_T$	Jet transverse momentum with μ -in-jet correction
$\log m$	DNN track jet mass
η	Jet pseudorapidity
ϕ	Jet azimuthal angle
isDeXTer60WP	Flag indicate if the jet tagged by DeXTer with 60 % WP
isDeXTer40WP	Flag indicate if the jet tagged by DeXTer with 40 % WP
DL1r-tagged b-jet	
$\log p_T$	Jet transverse momentum
$\log m$	Jet mass
η	Jet pseudorapidity
ϕ	Jet azimuthal angle
DL1r Continuous tag weight bin	Flag indicate DL1r tag weight bins
TC-LVT Secondary Vertex	
$\log(m)$	Track mass of the secondary vertex
$\log p_T$	Secondary vertex transverse momentum
η	Secondary vertex pseudorapidity
ϕ	Secondary vertex azimuthal angle
L_{3D}	Decay length relative to primary vertex
$S_{L_{3D}}$	Decay length significance
Z-boson candidate	
p_T	Z-boson candidate transverse momentum
η	Z-boson candidate pseudorapidity
ϕ	Z-boson candidate azimuthal angle
m	Z-boson candidate mass

Table 6.2: List of features used as input for the hypothesis testing NN. The variable reconstruction is described in Chapters 3 and 5. The Z-boson candidate is reconstructed as the 4-momentum sum of the two leptons.

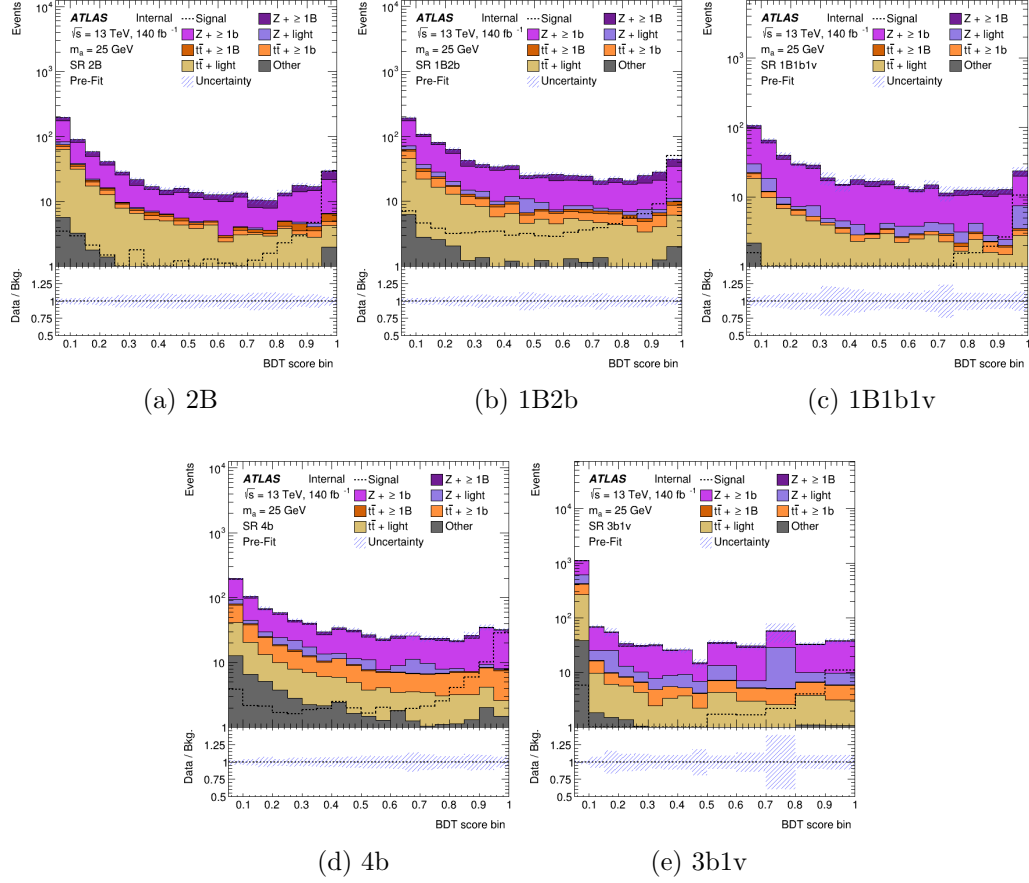


Figure 6.7: The pairing hypothesis score $\hat{y}_{\text{pred}}^{m_a}$ distribution assuming signal hypothesis $m_a = 25$ GeV with all the predicted categories. The SM background is compared to the expected $m_a = 25$ GeV signal simulated events assuming $\text{Br}(H \rightarrow aa \rightarrow 4b) = 0.1$.

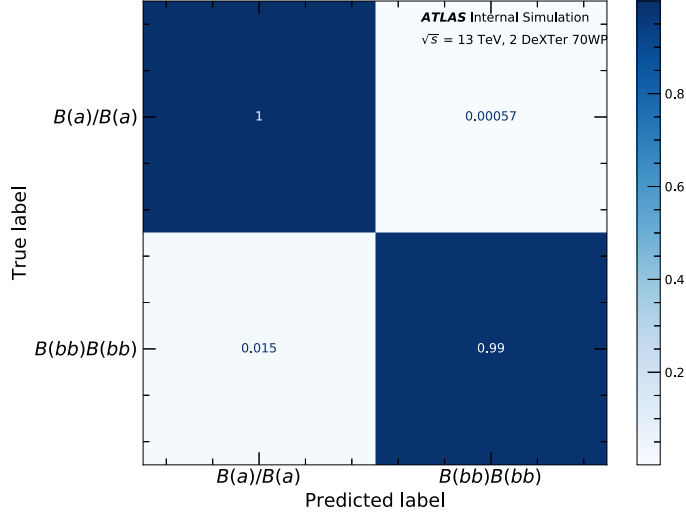


Figure 6.8: Confusion matrix for fully merged categories where signal events contain two DeXTer jets. The matrix is normalized per row to show the fraction of the true positive rate and false positive rate per truth-label category

true positive rate of around 70-80% as shown in Figure. 6.10 for four b-tagged jets and for three b-tagged jets and one secondary vertex.

A detailed description of the training detail is provide in Appendix A.2.

6.3.2 Event selection

Signal-background discriminator

The objects selected by the event reconstruction as coming from the $H \rightarrow aa/a_1a_2$ decay are used to build a representation of the four final-state particles in the decay of the two a -bosons. In the case of objects reconstructed from the hadronization of a single b -quark, like a DL1R-tagger b-jet or a TC-LVT-tagged SV, the association is trivial. In the case of DEXTER-tagged B-jets, the four-momentum of the particle-flow jet is decomposed in two components using projections onto the $\text{Ex}k_t^{(2)}$ track subjets.

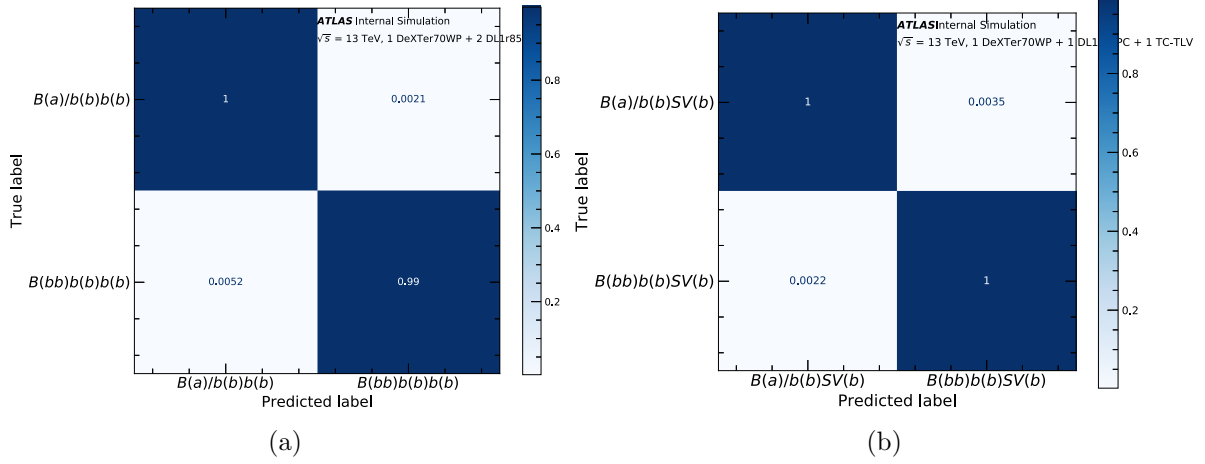


Figure 6.9: Confusion matrix for semi-merged categories where signal events (a) contain one DeXTer jet and two b-tagged jets or (b) one DeXTer jet, one b-tagged jet, and one secondary vertex. The matrix is normalized per row to show the fraction of truth positive rate versus false positive rate per truth-label category

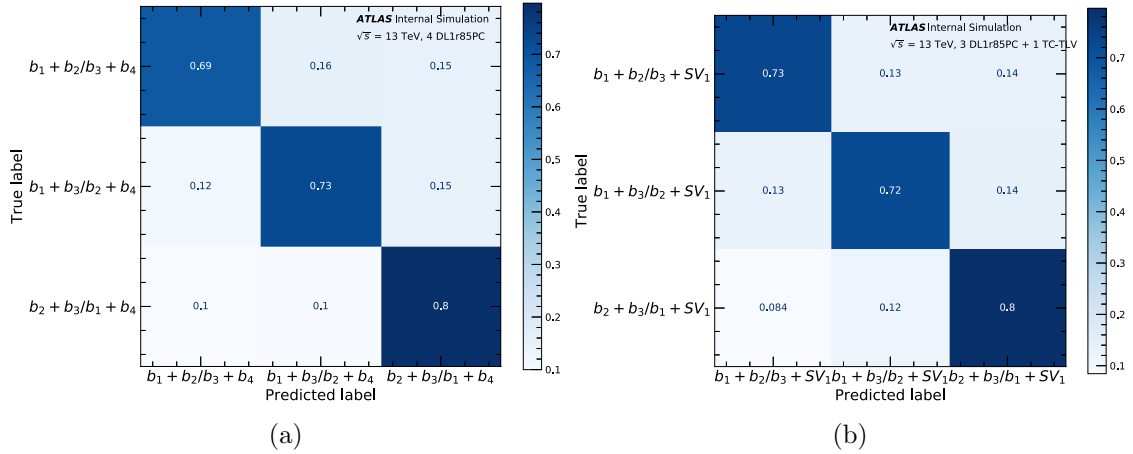


Figure 6.10: Confusion matrix for fully resolved categories where signal events are reconstructed as (a) four b-tagged jets in the detector (b) three b-tagged jets and one secondary vertex. The matrix is normalized per row to show the fraction of truth positive rate versus false positive rate per truth-label category

$$\begin{aligned}
p_{\text{PFlow}} &= c_1 p_{\text{Ex}k_t^{(2)}}^1 + c_2 p_{\text{Ex}k_t^{(2)}}^2 \\
\Rightarrow p_{\text{PFlow}} \cdot p_{\text{Ex}k_t^{(2)}}^1 &= c_1 m_{\text{Ex}k_t^{(2)}}^1 + c_2 p_{\text{Ex}k_t^{(2)}}^1 \cdot p_{\text{Ex}k_t^{(2)}}^2 \\
p_{\text{PFlow}} \cdot p_{\text{Ex}k_t^{(2)}}^2 &= c_1 p_{\text{Ex}k_t^{(2)}}^1 \cdot p_{\text{Ex}k_t^{(2)}}^2 + c_2 m_{\text{Ex}k_t^{(2)}}^2,
\end{aligned} \tag{6.5}$$

where the track subject energy-momentum is estimated by the sum of their tracks' four-momenta and does not include any calorimeter or particle-flow information. A p_T -ordering is used. The system above can be solved for c_1 and c_2 uniquely defining the two final-state particles.

The final-particles will be represented by $(b_{11}, b_{12}, b_{21}, b_{22})$, where the first index distinguishes the two a -bosons and the second distinguishes the two b -quarks in a single a -boson decay. Once the four final-state particles are determined, discriminating variables depending on the mass and spin of the decay $H \rightarrow aa/a_1 a_2 \rightarrow (b\bar{b})(b\bar{b})$ can be defined. The transverse momentum of the reconstructed a -boson candidates is used, and the following two-body reduced masses are built:

$$\begin{aligned}
m_{\text{red}}^{a_1} &= \sqrt{(p_{b_{11}} + p_{b_{12}})^2} - m_a \\
m_{\text{red}}^{a_2} &= \sqrt{(p_{b_{21}} + p_{b_{22}})^2} - m_a,
\end{aligned} \tag{6.6}$$

where m_a is the mass-hypothesis being tested. Subtracting m_a renders the variable comparable for different hypotheses. Similarly, the following four-body reduced mass is used:

$$m_{\text{red}}^H = \sqrt{(p_{11} + p_{12} + p_{21} + p_{22})^2} - m_H - m_{\text{red}}^{a_1} - m_{\text{red}}^{a_2}, \tag{6.7}$$

where a Higgs boson mass of $m_H = 125 \text{ GeV}$ is assumed to build this variable. The use of reduced mass in cascade decays provides a substantial improvement in mass resolution. Figure 6.11 shows the distribution of the reduced four-body mass m_{red}^H for difference event reconstruction channels.

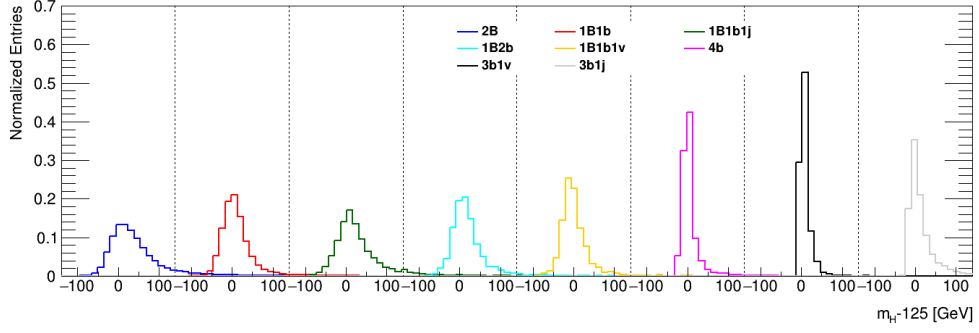


Figure 6.11: Normalized distribution of m_{red}^H in different event reconstruction channels.

The spin-0 nature of the Higgs boson is explored by the variable $\cos \theta^*$ defined as the cosine of the polar angle of $p_{a1}^{\text{CM}} = p_{b11}^{\text{CM}} + p_{b12}^{\text{CM}}$, where the superscript center-of-mass (CM) indicates that the vectors have been boosted to the rest frame of $p_H = (p_{11} + p_{12} + p_{21} + p_{22})$.

The spin-1 nature of the Z boson is explored by the cosine of the Collins-Soper (CS) angle [166], $\cos \theta_{\text{CS}}$ built from the two well-identified leptons $\ell_{(1)}$ and $\ell_{(2)}$ in the event and defined as:

$$\cos \theta_{\text{CS}} = 2 \frac{\ell_{(1)}^+ \ell_{(2)}^- - \ell_{(1)}^- \ell_{(2)}^+}{m_{\ell_{(1)}+\ell_{(2)}} \sqrt{m_{\ell_{(1)}+\ell_{(2)}}^2 + \left(p_T^{\ell_{(1)}+\ell_{(2)}}\right)^2}}, \quad (6.8)$$

where $\ell_{(i)}^\pm = (E_{(i)} \pm p_{z,(i)})/\sqrt{2}$. The $E_{(i)}$ and $p_{z,(i)}$ are corresponding energy and momentum along the z-direction of the lepton i . All the variables defined above, even the Lorentz invariant ones, are modulated by acceptance effects. To capture some of these effects, the transverse momentum of reconstructed Higgs candidates, p_T^H , is included. This allows the BDT to extract information based on the different Higgs boson candidate decay directions in the detector. The azimuthal angle, denoted as $|\Delta\phi(Z, H)|$, and the pseudorapidity separation, denoted as $|\Delta\eta(Z, H)|$, between the Higgs boson and the Z-boson candidate is also included and used in the BDT for the

event topology. Finally, the pairing hypothesis score described in 6.3.1 and the E_T^{miss} are included.

The E_T^{miss} is the sole background-driven variable, included to discriminate background events with neutrinos. The pairing hypothesis score in a summary statistic captures all the information already used to perform the object-to-parton assignment. The nature of the predominant information used depends on which objects have been selected and the mass hypothesis used for the parameterized NN. Therefore, it becomes necessary to train a different discriminator for each mass and reconstructed hypothesis. Figures 6.12 and 6.13 show examples of inputs for two different masses. As can be seen, the mass of the reconstructed a -boson is more prominently used for object-to-parton assignment at low m_a hypotheses. In contrast, the reconstructed H -boson mass becomes more important at high m_a hypotheses. The background distributions are sculpted after applying the $\hat{y}_{\text{pred}}^{m_a}$ selection.

The training of the BDT has been performed with the XGBOOST [167] package. The hyperparameters are tuned with RAY TUNE [168] using the HYPERBAND scheduler [169]. The overtraining is assessed by splitting the full dataset into a training dataset and a hold-out validation sample (with a 60 – 40% split). Signs of overtraining are tested by comparing the performance on the training and the hold-out validation sample. Gradient boosting using between 150 and 600 trees improves the performance of weak classifiers.

Signal region definition

The events in the Signal Regions (SRs) targeting the 4b final state fulfill the pre-selection requirements described in Section 6.2. To select signal events with the leptonic decay of the Z boson, the two leptons in the events are required to be of two electrons or two muons of opposite sign charge, called same flavor (**SF**), and the invariant mass is required to be in $71 \text{ GeV} \leq m_{\ell\ell} \leq 111 \text{ GeV}$ (loose Z mass window).

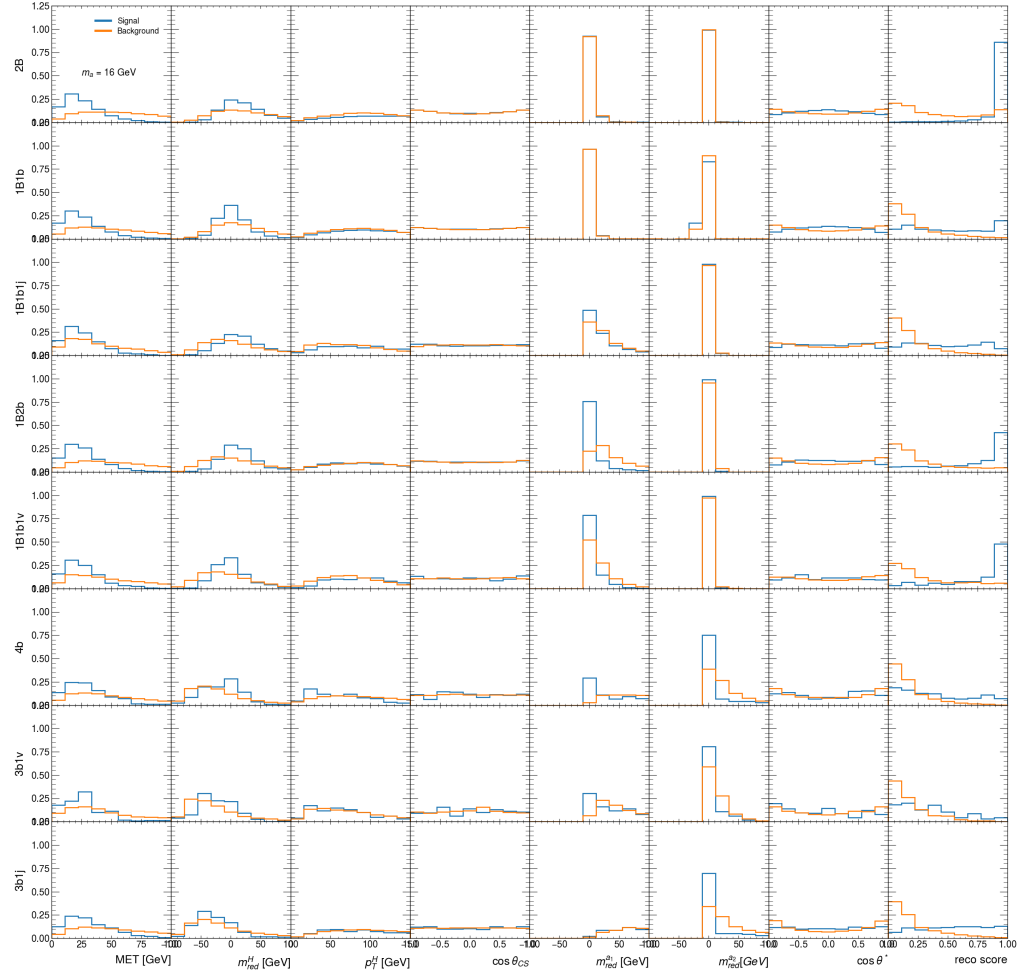


Figure 6.12: Signal and background distributions of the input variables, per channel, for the mass-spin discriminator. Shown for the $m_a = 16$ GeV mass hypothesis.

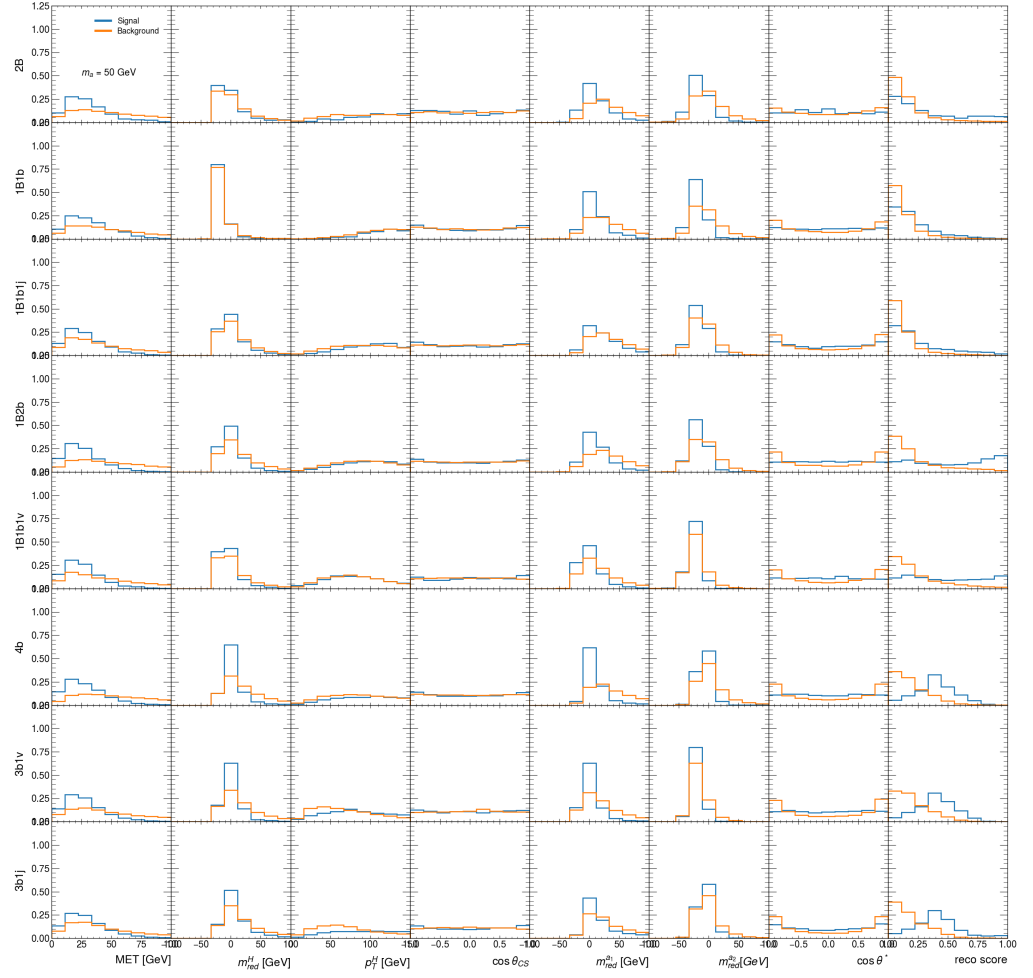


Figure 6.13: Signal and background distributions of the input variables, per channel, for the mass-spin discriminator. Shown for the $m_a = 50$ GeV mass hypothesis.

name	lep comb	Z mass	pairing hypothesis score $\hat{y}_{\text{pred}}^{m_a}$	BDT binning
2B	SF	loose	> 0.5	Loose,Medium,Tight
1B2b	SF	loose	> 0.05	Loose,Medium,Tight
1B1b1v	SF	loose	> 0.05	Loose,Medium,Tight
4b	SF	loose	> 0.05	Loose,Medium,Tight
3b1v	SF	loose	> 0.05	Loose,Medium,Tight

Table 6.3: Overview of the signal regions targeting the 4b final state. Each signal region exists for each mass hypothesis.

In addition, a requirement is applied on the pairing hypothesis score predicted by the hypothesis testing NN; it is required to be $\hat{y}_{\text{pred}}^{m_a} > 0.05$. Only in the 2B region is the pairing hypothesis score tightened to $\hat{y}_{\text{pred}}^{m_a} > 0.5$ due to the shape of the $\hat{y}_{\text{pred}}^{m_a}$ output of the hypothesis testing NN, allowing for more statistics in the CRs defined below. The different regions are then defined using the predicted hypothesis given by the event hypothesis NN. Five SRs are defined and classified by the multiplicity of b-objects in the predicted hypothesis. An overview is shown in Table 6.3. The five SRs are binned in the output discriminant of the BDT. Three bins (Loose, Medium, and Tight) are obtained by performing a bin optimization on the BDT output. The optimization is carried out using the expected limit obtained with a counting experiment of the signal and background yields given by the bin selection under study, assuming a flat 20% background uncertainty for this test. First, the Tight bin is determined by finding the bin boundary that gives the highest sensitivity. Then, the bin boundary between the Loose and Medium bins is found by a second scan.

$Z + \text{jets}$ control regions

Events for the $Z + \text{jets}$ Control Regions (CRs) are selected to be enriched in $Z + \text{HF}$ events and close in phase space to the SRs. They fulfill the pre-selection requirements in Section 6.2. The two leptons in the events are required to be of the same flavor and opposite sign charge, and the invariant mass, $m_{\ell\ell}$, is required to satisfy 81 GeV

$\leq m_{\ell\ell} \leq 101 \text{ GeV}$ (tight Z mass window). To further increase the purity of $Z + \text{jets}$ events, a selection on the missing transverse momentum $E_{\text{T}}^{\text{miss}} < 60 \text{ GeV}$ is added, as it suppresses the di-leptonic $t\bar{t}$ events which have neutrinos in their final state. To ensure orthogonality to the SRs defined above, the selection on the pairing hypothesis score, $\hat{y}_{\text{pred}}^{m_a}$, is inverted.

In addition to this selection, the CRs are split by the number of B- and b-jets in the events. For B-jets, the selection depends on the number of Tight (40% WP) and the number of Loose, but not Tight (70% - 40% WP) tagged jets. For b-jets, only one b-tagging working point is used, as in the low p_{T} range $15 \text{ GeV} \leq p_{\text{T}} < 20 \text{ GeV}$ only the 85% WP is calibrated. The selected regions are chosen to be close to the SRs in the number of b-objects, and an overview of the nine regions is given in Table 6.4.

As $\hat{y}_{\text{pred}}^{m_a}$ depends on the mass hypothesis, the set of $Z + \text{jets}$ CRs exists for each mass hypothesis. They are orthogonal to each other and the SRs within one mass hypothesis but not necessarily across the mass hypotheses.

$t\bar{t}$ control regions

Events for the $t\bar{t}$ CRs are selected to be enriched in $t\bar{t}$ events. They also fulfill the pre-selection requirements described in Section 6.2. The lepton selection here, however, requires the leptons to be of opposite sign charge. Selecting events with one electron and one muon, called different flavor (**DF**), successfully suppresses $Z + \text{jets}$ events. No selection on the lepton invariant mass (in addition to the one in the pre-selection) and no selection on the $\hat{y}_{\text{pred}}^{m_a}$ is applied. The $t\bar{t}$ CRs are also split by the number of B- and b-jets in the events, using the same scheme and definitions as in the $Z + \text{jets}$ CRs. The $t\bar{t}$ CR definition does not depend on the predicted mass hypothesis ($\hat{y}_{\text{pred}}^{m_a}$), so they are the same for each analysis testing a mass hypothesis. An overview is given in Table 6.5.

name	pairing hypothesis score $\hat{y}_{\text{pred}}^{m_a}$	nBJets Tight	nBJets Loose	nbJets (Loose)	nSoftSV	nLightJets
2iB 0ib						
SFonZ_1iBT						
_2iBTplusBL_0ibL	≤ 0.5	1+	2+ (BT+BL)	0+	any	any
SFonZ_0BT_2iBL_0ibL	≤ 0.5	0	2+	0+	any	any
1B 2ib						
SFonZ_1BT_0BL_2ibL	≤ 0.05	1	0	2+	any	any
SFonZ_0BT_1BL_2ibL	≤ 0.05	0	1	2+	any	any
0B 4ib						
SFonZ_0BT_0BL_4ibL	≤ 0.05	0	0	4+	any	any
0B 3b 1isv or 1ij						
SFonZ_0BT_0BL_3bL						
_1isvorj	≤ 0.05	0	0	3	1+ or:	1+

Table 6.4: Overview of the $Z + \text{jets}$ control regions targeting the 4b final state. Each $Z + \text{jets}$ control region exists for each mass hypothesis, as the selection depends on $\hat{y}_{\text{pred}}^{m_a}$. The naming scheme is defined as follows. **SFonZ** indicate that the events must have the same flavor opposite sign charge leptons and the dilepton mass in the Z mass windows. The subsequent suffix indicates the number of b-tagged objects. The **i** indicates if the category is an inclusive category that only requires the minimum number of objects, e.g., **1i** requires events to have at least one without an upper bound. The b-tagging object is defined as follows: **BT**: Tight DeXTer-tagged jet, **BL**: Loose DeXTer-tagged jet, **bL**: DL1r-tagged jet, **sv**: soft secondary vertex, and **j**: jet not tagged by the DL1r 85% WP.

Name	nBJets Tight	nBJets Loose	nbJets	nSoftSV	nLightJets
2iB 0ib					
DF_1iBT_2iBTplusBL_0ibL	1+	2+BT+BL	0+	any	any
DF_0BT_2iBL_0ibL	0	2+	0+	any	any
1B 2ib					
DF_1BT_0BL_2ibL	1	0	2+	any	any
DF_0BT_1BL_2ibL	0	1	2+	any	any
0B 4ib					
DF_0BT_0BL_4ibL	0	0	4+	any	any
0B 3b 1isv or 1ij					
DF_0BT_0BL_3bL_1isvorj	0	0	3	1+ or:	1+

Table 6.5: Overview of the $t\bar{t}$ control regions targeting the 4b final state. The naming scheme is defined as follows. **DF** indicates the events required to have the different flavor, opposite sign charge leptons. The subsequent suffix indicates the number of b-tagged objects. The **i** indicates if the category is an inclusive category that only requires the minimum number of objects, e.g., **1i** requires events to have at least one without an upper bound. The b-tagging object is defined as follows: **BT**: Tight DeXTer-tagged jet, **BL**: Loose DeXTer-tagged jet, **bL**: DL1r-tagged jet, **sv**: soft secondary vertex, and **j**: jet not tagged by the DL1r 85% WP.

Validation regions

To validate the normalization factors outside of control and signal regions, Validation Regions (VRs) are introduced. For a VR enriched in Z + jets, the 1B1b multiplicity region, with precisely 1 B-jet and exactly one b-jet, is used without any requirement on $\hat{y}_{\text{pred}}^{m_a}$. For VRs enriched in $t\bar{t}$, the same flavor lepton selection is chosen, but $m_{\ell\ell}$ is required to be outside the loose Z mass window $m_{\ell\ell} \notin (71, 111) \text{ GeV}$ (off Z mass peak). Otherwise, the same definitions for the different flavor $t\bar{t}$ CRs are used.

6.4 Systematic uncertainties

Various sources of systematic uncertainties are considered for the analysis. A summary is given at the end of this section in Table 6.8. The systematic uncertainties are introduced to the likelihood function in the form of constrained Nuisance Parameters

(NPs) α_i . Each yield n is parametrized as a function of α_i using standard vertical interpolation [170]. Each constraint term is described by a Gaussian distribution with mean 0 and width 1, representing auxiliary measurements or models that provide external information about that source of variation. The normalized Gaussian with mean 0 and width 1 can always be obtained by shifting $\alpha_i \rightarrow (\alpha_i - a_i)/\sigma_i$, where a_i is the maximum likelihood estimator of the best value of α_i obtained from auxiliary measurements and σ_i is the corresponding maximum likelihood estimator of its uncertainty.

6.4.1 Experimental uncertainties

The experimental uncertainties are related to luminosity, pile-up, trigger efficiency, lepton reconstruction and identification, jet, and flavor tagging calibration, and track reconstruction. As the soft SVs are built directly from tracks, the low-level tracking-related uncertainties are also included in the analysis. These systematic uncertainties are included in this result either as an overall event weight or as a data-MC scale factor on a set of objects. The effects of the various systematic variations are evaluated assuming a positive and negative 1σ variation around the nominal value of the quantity of interest to evaluate the effects on the yields used in the fit. Experimental uncertainties are applied to all simulated processes.

Luminosity The uncertainty on the integrated luminosity for the full Run-2 dataset is 0.83% [171], obtained using the LUCID-2 detector [159] for the primary luminosity measurements.

Pile-up modeling A variation in the pile-up reweighting of simulated events is included to cover the uncertainty in the ratio of the predicted and measured inelastic cross-sections in the fiducial volume defined by $M_X > 13$ GeV where M_X is the mass of the hadronic system [172]. In practice, the nominal scale factor applied on the

data pile-up distribution when performing the reweighting is changed into 1.0/0.99 or 1.0/1.07 to derive the up and down systematic uncertainty instead of its nominal value 1.0/1.03.

Lepton triggers, reconstruction, identification, and isolation The reconstruction, identification, and isolation efficiency of electrons and muons, as well as the efficiency of the trigger used to record the events, differ slightly between data and simulation, which is compensated for by dedicated SFs. Efficiency SFs are measured using tag-and-probe techniques on $Z \rightarrow \ell^+ \ell^-$ data and simulated samples [154, 155], and are applied to the simulation to correct for differences. The effect of these SFs and their uncertainties are propagated as corrections to the MC event weight. In total, 4 independent components are considered for electrons and 10 for muons.

Lepton momentum scale and resolution Additional sources of uncertainty originate from the corrections applied to adjust the lepton momentum scale and resolution in the simulation to match those in data, measured using reconstructed distributions of the $Z \rightarrow \ell^+ \ell^-$ and $J/\psi \rightarrow \ell^+ \ell^-$ masses, as well as the E/p ratio measured in $W \rightarrow e\nu$ events where E and p are the electron energy and momentum measured by the calorimeter and the tracker, respectively [83, 155]. To evaluate the effect of momentum scale uncertainties, the event selection is redone with the lepton energy or momentum varied by $\pm 1\sigma$. The event selection is redone for the momentum resolution uncertainties by smearing the lepton energy or momentum. In total, three independent components are considered for electrons and five for muons.

Jet Vertex Tagging efficiency SFs are applied to correct discrepancies between Data and MC for JVT efficiencies. These SFs are estimated using events with a jet recoiling against $Z \rightarrow \mu^+ \mu^-$ [156], and the effect of these SFs, as well as of their uncertainties are propagated as corrections to the MC event weight. Similar studies are performed to extend below the current limit of $p_T > 20\text{GeV}$ down to 15 GeV.

One additional component is considered for low- p_T JVT [173]. The total relative uncertainty of the low- p_T JVT SF is below 3%.

Jet energy scale and resolution The jet energy scale and its uncertainty are derived by combining information from test-beam data, LHC collision data, and simulation [157]. The uncertainties from these measurements are factorized into 8 independent sources. The jet energy resolution was measured in Run 2 data and simulation as a function of jet p_T and rapidity using dijet events, following a similar method to as in Ref. [158]. The combined uncertainty is propagated by smearing the jet p_T in MC, yielding eight independent sources.

Additional jet uncertainties Additional uncertainties are considered related to jet flavor (using the conservative default value of $\pm 50\%$ for the quark/gluon fraction for all MC samples), pile-up corrections, η dependence, high- p_T jets, yielding fourteen additional independent sources.

DL1r flavor tagging To correct flavor tagging efficiencies in simulated samples to match efficiencies in data, SFs are derived as a function of the p_T for b -jets, c -jets, and light jets separately in dedicated calibration analyses. For b -jet efficiencies, $t\bar{t}$ events in the dilepton topology were used, exploiting the very pure sample of b -jets arising from the decays of the top quarks [92]. For c -jet mistag rates, $t\bar{t}$ events in single-lepton topology were used, exploiting the c -jets from the hadronically decaying W bosons [93]. The so-called negative-tag method is used in Z +jets events [174] for light-jets mistag rates. In the three calibration analyses, a large number of uncertainty components are considered, and a principal component analysis is performed, yielding 45, 20, and 20 eigen variations for b -, c - and light-jets, respectively, which are taken as uncorrelated sources of uncertainties. These eigen variations correspond to the number of p_T bins (9, 4, and 4 respectively for b -, c - and light-jets) multiplied by the number of DL1r bins (5).

Dedicated calibrations were performed to extend jets down to $p_T > 15\text{GeV}$ for b - and *light*-jets. The studies are described in [173], and the resulting uncertainty is propagated to the fit. The uncertainty for the b -jets SF is 11.5% and 21.9% for the light jet SF. For the c -jets, an inflated uncertainty from the closest, i.e., lowest, p_T bin of the standard c -jet calibration is considered, which is 11%.

DeXTer flavor tagging The use of DeXTer in this analysis introduced additional SFs to correct the efficiency difference between simulated samples and data. This paragraph provides a brief summary of key information; more extended descriptions of the methodology can be found in the relevant publications in [162, 175].

The SFs for DeXTer are derived as a function of p_T for B and b -jets. The calibration measurements with data are performed using both $t\bar{t}$ and $Z + \text{jets}$ events simultaneously to measure B -tagging and b -jet mistagging efficiency in data. The remaining flavors are found to be negligible in this analysis.

Soft secondary vertices To correct for potential mis-modeling in the reconstruction efficiency and fake rate of the TC-LVT tagger, a calibration is carried out in [173]. Efficiency and a fake rate SF are derived simultaneously in $t\bar{t}$ events with 1 or 2 b -tagged jets. The fake rate SF is determined in two bins of μ_{actual} . Hence, three independent uncertainty components are derived in a principal component analysis and taken as uncorrelated sources of uncertainty.

Missing transverse momentum The performance and modeling of the E_T^{miss} reconstruction are studied in events containing a Z boson decaying to two charged leptons (electrons or muons) [97]. The systematic uncertainties give four independent parameters. A scale uncertainty (up and down) is applied, corresponding to scaling the soft term magnitude up or down in the direction of p_T^{Hard} , the sum of the transverse momentum of reconstructed and calibrated “hard objects”: electrons, photons, muons, hadronically decaying τ -leptons, and jets. Two independent uncertainties

Source	ATLAS Nomenclature
Track Efficiency	TRACK EFFICIENCY GLOBAL (LOOSE)
	TRACK EFFICIENCY IBL (LOOSE)
	TRACK EFFICIENCY PP0 (LOOSE)
	TRACK EFFICIENCY PHYSICS MODEL (LOOSE)
Fake Rate	TRACK FAKE RATE ROBUST (LOOSE)
Impact Parameters	TRACK BIAS Z0 WM
	TRACK RESOLUTION Z0 DEAD
	TRACK RESOLUTION Z0 MEAS
	TRACK BIAS D0 WM
	TRACK RESOLUTION D0 DEAD
	TRACK RESOLUTION D0 MEAS
Track Q/P Sagitta	TRACK BIAS Q/P SAGITTA WM

Table 6.6: List of tracking systematic uncertainties included in the efficiency measurement. Each line represents an independent source of uncertainty considered in the measurement.

for the resolution correspond to smearing the soft term magnitude in the direction parallel or perpendicular to the direction of p_T^{Hard} .

Tracking efficiency and parameters Systematic uncertainties related to the track selection efficiency are determined by changing the amount of tracker material and the physical models in the GEANT4 simulation. Those systematic variations on the number of fake tracks are applied based on the recommendation for the Loose track selection working point. Dedicated systematic uncertainties are considered for the track parameters, including the transverse and longitudinal impact parameters and the track Sagitta. The event selection is redone for each systematics variation. The complete list of track-related systematic uncertainties is shown in Table 6.6.

6.4.2 Modeling uncertainties

For each sample considered, there are modeling uncertainties associated with the missing higher orders in the perturbative expansion of the partonic cross-section and the PDF uncertainties. Modeling uncertainties are considered for signal, Z + jets, and

$t\bar{t}$ simulated events, as the other processes are negligible. For signal, the modeling uncertainties are considered separately for events of the $qq \rightarrow ZH$ and $gg \rightarrow ZH$ production processes. Most modeling uncertainties are treated similarly in signal, $Z + \text{jets}$, and $t\bar{t}$ simulated events; process-specific variations are mentioned below.

Renormalization and Factorization Scale Variations in the renormalization and factorization scales are used to estimate the uncertainty due to missing higher-order corrections. The samples used include on-the-fly variations corresponding to 7 variations of the renormalization (μ_r) and factorization (μ_f) scales in pairs: $\{\mu_r, \mu_f\} \times \{0.5, 0.5\}, \{1, 0.5\}, \{0.5, 1\}, \{1, 1\}, \{2, 1\}, \{1, 2\}, \{2, 2\}$. The uncertainties are combined by taking an envelope of all seven variations.

PDF Uncertainties The PDF uncertainties follow the PDF4LHC recommendations [160] and are available as sample weights. The PDF uncertainty is evaluated for the PDF set used in the matrix element calculation. The α_s uncertainty is derived using the same PDF set evaluated with two different α_s values. The two uncertainties from the PDF and α_s are added in quadrature.

Parton Shower and Hadronization For the uncertainties due to parton shower (PS) and hadronization and due to the NLO matching, the nominal POWHEGBOX+PYTHIA8 $t\bar{t}$ (5FS) sample is compared to the POWHEGBOX+HERWIG7 $t\bar{t}$ (5FS) sample for $t\bar{t}$ simulated event. In the case of the $Z + \text{jets}$ simulated event, the nominal SHERPA 2.2.11 sample is compared to the MADGRAPH5_AMC@NLO 2.6.5 + PYTHIA 8.240 FxFx sample. Finally, the nominal POWHEGBOX+PYTHIA8 samples are compared to POWHEGBOX+HERWIG7 signal simulated event.

ISR and FSR Modeling For ISR and FSR uncertainties, the settings of the nominal POWHEGBOX+PYTHIA8 $t\bar{t}$ (5FS) sample are varied, resulting in different event weights; the uncertainty due to ISR is estimated by simultaneously changing

μ_r and μ_f in the ME and μ_R^{ISR} in the PS. In contrast, the uncertainty due to FSR is estimated by changing μ_R^{FSR} in the PS. For the ISR, the amount of radiation is increased (decreased) by scaling μ_r and μ_f by a factor of 0.5 (2.0) and by varying the renormalization scale for QCD emission in the ISR by a factor of 0.549 (1.960), corresponding to the VAR3CUP (VAR3CDOWN) variation from the A14 tune [113] which sets α_S^{ISR} to 0.140 (0.115) instead of the nominal 0.127. For the FSR, the amount of radiation is increased (decreased), varying the renormalization scale for QCD emission in the FSR by a factor of 0.5 (2.0), corresponding to $\alpha_S^{\text{FSR}} = 0.1423$ (0.1147) instead of the nominal $\alpha_S^{\text{FSR}} = 0.127$.

Z+HF Modeling Uncertainties The uncertainty in the heavy flavor components is derived by comparing the nominal Z + jets sample with alternatives with varied settings for the overlap between matrix element and Parton shower emissions and the resummation scale [161].

The matrix element matching uncertainty (CKKW) is estimated by varying the scale used to calculate the overlap between jets from the matrix element and the parton shower. The nominal value for this parameter is 20 GeV. The up variation increases this to 30 GeV (CKKW30), while the down variation decreases the nominal value to 15 GeV (CKKW15).

The resummation scale (QSF) uncertainty is estimated by varying the scale used for the resummation of soft gluon emission. μ_{QSF} is varied by 2 and $\frac{1}{2}$ with respect to the nominal.

The resulting changes in the relative contributions for each heavy flavor type are summarized in Table 6.7 for each modeling variation. The total uncertainty is derived from the sum in quadrature of the two uncertainties.

$t\bar{t}$ Matrix Element and NLO Matching Uncertainty The impact of the matrix element corrections applied to the top decay in Powheg and the NLO match-

Variation	$Z + b$	$Z + B$	$Z + c$
CKKW30	-10.68%	-7.84%	-4.59%
CKKW15	-10.2%	-14.46%	-5.15%
QSF25	-7.02%	-8.59%	-4.62%
QSF4	-8.99%	-2.46%	-4.58%
Total	12.38%	8.22%	6.48%

Table 6.7: Relative variation in the fraction of each component of $Z + \text{heavy flavor}$

ing uncertainty in $t\bar{t}$ are evaluated by comparing the nominal sample to an alternative sample of events generated with MADGRAPH5_AMC@NLO 2.6.0 interfaced with PYTHIA 8.230 [176]. This method compares NLO matching and the matrix element corrections at the same time.

Background kinematics reweighting uncertainties Additional uncertainties from kinematics reweight determined from data. This includes the uncertainties on the reweighing factor for $t\bar{t} + \text{HF}$ and $Z + \text{jet}$ kinematics reweighting.

B-Jet Mass Scale correction uncertainties To correct the mis-modeling in Ak8 track jet mass, additional correction mass scale corrections are estimated from the data with minimized χ^2 method. Uncertainty is included to account for the uncertainties on the correction scale with this method. The B-jet mass scale varies by $\pm 5\%$ and compares the result with the nominal results with a mass scale of 1.01.

6.4.3 Summary of systematic uncertainties

Table 6.8 lists all the nuisance parameters used in the measurement. They are grouped into types used to summarize the uncertainties later on in the note in the presentation of the results. Note that the renormalization and factorization scale and the parton shower and hadronization uncertainties for $t\bar{t}$ and $Z + \text{jets}$, and the matrix element uncertainty (from AMC@NLO) for $t\bar{t}$ are considered separately, i.e., the NPs are de-correlated, for the different heavy flavor components of the backgrounds.

Systematic Group	Nuisance Parameters
Luminosity	LUMINOSITY
Pileup	PILE-UP MODELING
Lepton	ELECTRON TRIGGER, ELECTRON RECO, ELECTRON ID, ELECTRON ISOLATION, EG RESOLUTION, EG SCALE, MUON TRIGGER (STAT. & SYST.), MUON TTVA (STAT. & SYST.), MUON ID (STAT. & SYST.), MUON ISOLATION (STAT. & SYST.), MUON ID, MUON MS p_T RES., MUON SAGITTA REBIAS, MUON SAGITTA RHO, MUON SCALE
Jet	JVT, JVT LOW- p_T , JER DATA v.s MC MC16, JER NP1, JER NP2, JER NP3, JER NP4, JER NP5, JER NP6, JER NP7RESTTERM, JES NP1, JES NP2, JES NP3, JES NP4, JES NP5, JES NP6, JES NP7, JES NP8RESTTERM, JET FLAVOR COMPOSITION, JET FLAVOR RESPONSE, JET B RESPONSE, PILE-UP OFFSETUP p_T TERM, PILE-UP OFFSET MU TERM, PILE-UP OFFSET NPV TERM, ρ TOPOLOGY, η INTERCALIB (STAT.), η INTERCALIB (MODEL), η INTERCALIB (NON-CLOSURE NEGETA), η INTERCALIB (NON-CLOSURE POSETA), η INTERCALIB (NON-CLOSURE HIGH E), PUNCH-THROUGH CORRECTION, JET HIGH- p_T CORRECTION
DL1r tagging	DL1R CONTINUOUS EIGENVARS - B 0-44, DL1R CONTINUOUS EIGENVARS - C 0-19, DL1R CONTINUOUS EIGENVARS - L 0-19, DL1R LOW- p_T B/C/LIGHT
DeXTer tagging	DEXTER EIGENVARS 0-11
Soft SVs	TC-LVT EIGENVARS 0-2
E_T^{miss}	SOFTTRK SCALE (UP), SOFTTRK SCALE (DOWN), SOFTTRK RESOPARA, SOFTTRK RESOPERP
Tracking	TRACK EFFICIENCY GLOBAL (LOOSE), TRACK EFFICIENCY IBL (LOOSE), TRACK EFFICIENCY PP0 (LOOSE), TRACK EFFICIENCY PHYSIC MODEL (LOOSE), TRACK FAKE RATE ROBUST (LOOSE), TRACK BIAS Z0 WM, TRACK RESOLUTION Z0 DEAD, TRACK RESOLUTION Z0 MEAS, TRACK BIAS D0 WM, TRACK RESOLUTION D0 DEAD, TRACK RESOLUTION D0 MEAS, TRACK BIAS Q/P SAGITTA WM
$t\bar{t}$ modeling	$t\bar{t}$ μ_R and μ_F , $t\bar{t}$ ISR α_S , $t\bar{t}$ FSR, $t\bar{t}$ PYTHIA 8 vs. HERWIG 7, $t\bar{t}$ POWHEG vs. AMC@NLO, $t\bar{t}$ PDF PDF4LHC
Z+jets modeling	Z+jets μ_R and μ_F , Z+jets SHERPA 2.2.11 vs. MADGRAPH+PYTHIA8, Z+jets PDF NNPDF30, Z+B modeling, Z+b modeling, Z+c modeling
Signal modeling	Signal μ_R and μ_F , Signal ISR α_S , Signal FSR, Signal PYTHIA 8 vs. HERWIG 7, Signal PDF PDF4LHC, Signal ggZH μ_R and μ_F , Signal ggZH ISR α_S , Signal ggZH FSR, Signal ggZH PYTHIA 8 vs. HERWIG 7, Signal ggZH PDF PDF4LHC

Table 6.8: Summary of all systematic uncertainties considered in the measurement, grouped categories.

6.5 Statistical analysis

A combination of different distributions from signal and control regions is employed to test for the potential presence of a signal. The statistical analysis relies on a binned likelihood function denoted as $\mathcal{L}(\mu, \theta)$. This function is constructed by taking the product of Poisson probability terms over all bins considered in the analysis. The likelihood function is dependent on two key parameters: the signal-strength parameter μ , which serves as a multiplicative factor to the theoretical signal production cross section for ZH production, and the nuisance parameters θ , which capture the impact of systematic uncertainties on the signal and background expectations. The likelihood function incorporates Gaussian priors to represent these uncertainties.

Hence, the expected number of events in each bin depends on the values of both μ and θ . The nuisance parameters θ allow for variations in the signal and background expectations, accounting for the corresponding systematic uncertainties. Their fitted values correspond to the deviations from the pre-fit expectations that globally provide the best fit to the data. This procedure reduces the impact of systematic uncertainties on the search sensitivity by taking advantage of the highly populated background-dominated channels included in the likelihood fit. It is crucial to possess a good understanding of the systematic effects impacting the shapes of the discriminant distributions. Therefore, extensive validation studies have been conducted using simulations to verify the accuracy of the fitting procedure. Additionally, the fit incorporates dedicated parameters to account for the limited number of MC events.

To enhance the efficiency of the fitting process and ensure a stable minimization, systematic uncertainties are pruned and smoothed. This pruning and smoothing procedure is carried out separately for each region and sample, addressing both the shape and normalization aspects of the systematics.

The pruning algorithm identifies and removes systematic uncertainties that have a minimal impact on normalization, specifically those with an impact of $\leq 1\%$. Like-

wise, systematics that negligibly affects the shape of the final discriminant, where all bins deviate by $\leq 1\%$ from the nominal shape, are also pruned. This pruning step helps reduce the computational burden and streamline the analysis.

Additionally, to mitigate the effects of statistical fluctuations, a smoothing technique is employed. Systematic variations are subjected to smoothing, resulting in a more stable representation of their impact throughout the analysis.

The test statistic $\tilde{q}_\mu$ is defined as the profile likelihood ratio,

$$\tilde{q}_\mu = \begin{cases} -2\ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(0, \hat{\theta}(0))} & \hat{\mu} < 0, \\ -2\ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \quad (6.9)$$

where $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximize the likelihood function with $\mu \geq 0$, and $\hat{\theta}_\mu$ are the values of the nuisance parameters that maximize the likelihood function for a given value of μ . The test statistic $\tilde{q}_\mu$ is implemented in the ROOSTATS package [177] and is used to measure the compatibility of the observed data with the signal plus background hypothesis. The p -value representing the compatibility of the data with the signal plus background hypothesis is estimated by integrating the distribution of $\tilde{q}_\mu$ from signal plus background pseudo-experiments approximated using the asymptotic formulae given in ref. [178], above the observed value of $\tilde{q}_\mu$. The observed p_μ -value is checked for each explored signal scenario.

In the absence of any significant excess above the background expectation, upper limits on the signal production cross section for each signal scenario are derived using q_μ in the CLs method. For a given signal hypothesis, values of the branching fractions of exotic Higgs decays (parameterized by μ) yielding $\text{CLs} < 0.05$, where CLs is computed using the asymptotic approximation, are excluded at $\geq 95\%$ CL.

The normalization of each background is determined from the fit simultaneously with μ , constrained by the uncertainties of the respective theoretical calculations, the

Notation	Description	SM background categories	Applied regions
$\mu_{Z+\geq 1b, \geq 1B}$	Z+ $\geq 1b$ and Z+ $\geq 1B$ normalisation	Z+ $\geq 1b$ and Z+ $\geq 1B$	All
$\mu_{t\bar{t}}$	$t\bar{t}$ cross-section	All $t\bar{t}$	All
$\mu_{t\bar{t}+\geq 1b}$	$t\bar{t}+\geq 1b$ fraction	$t\bar{t}+\geq 1b$	All
$\mu_{t\bar{t}+\geq 1B}$	$t\bar{t}+\geq 1B$ fraction	$t\bar{t}+\geq 1B$	All

Table 6.9: Summary of the background normalization factors that are free-floating in the fit and the corresponding background components.

uncertainty on the luminosity, and the data themselves. Statistical uncertainties in each region are taken into account by dedicated parameters in the fit.

A summary of the free-floating background normalization factors in the fit is given in Table 6.9. For the Z+jet background, only the main Z+ $\geq 1b$ and $\geq 1B$ components are normalized by one common free-floating factor. This Z+ $\geq 1b, \geq 1B$ component is constrained using the dedicated Z+jets CRs. Extensive studies have been carried out on the possibility of floating differently defined Z+HF categories, e.g., separate normalizations for Z+ $\geq 1b$ and $\geq 1B$.

However, due to the lack of available statistics, the fit cannot constrain additional Z+HF categories separately. This is accounted for by the Z+HF modeling uncertainties instead. For the $t\bar{t}$ +jets background, the overall $t\bar{t}$ normalization, and the $t\bar{t}+\geq b$ and $t\bar{t}+\geq B$ normalizations are left free to float separately. They are constrained by the dedicated $t\bar{t}$ control regions that are split by the reconstructed b-object multiplicity described in Section 6.3.

6.6 Results

Section 6.5 describes how a binned likelihood fit is used to search for the signal. As explained in Sec. 6.3, the dilepton channel was examined by looking at possible combinations of signal hypotheses and applying BDTs to maximize the search sensitivity. The total yields in both the signal and control regions are fit. The fit is performed separately for each signal hypothesis, corresponding to a different mass for

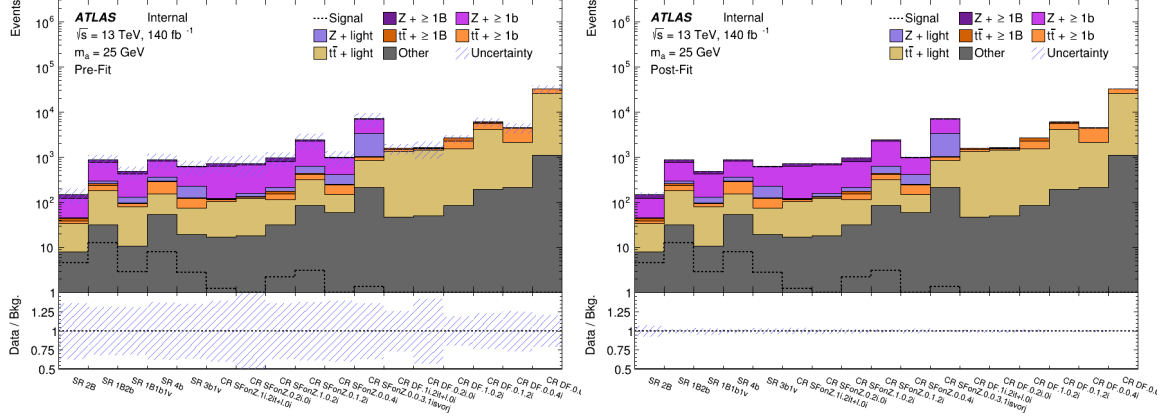


Figure 6.14: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 25\text{ GeV}$ mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line. Fit results for the other mass hypotheses can be found in Appendix B.1.

the new spin-zero particle, m_a . The results for the $m_a = 25\text{ GeV}$ signal are described in this section, while the fit results for the other masses considered are included in Appendix B.

6.6.1 Fits to the Asimov dataset

This section shows the signal plus background fits to the Asimov dataset for the regions described previously in Sec. 6.3. The expected background is fitted to the Asimov dataset under the signal plus background hypothesis. The region summary is shown in Figure 6.14 and Figure 6.15 with the $m_a = 25\text{ GeV}$ as an example.

Figure 6.16 shows the fitted nuisance parameters, and Figure 6.19 shows the normalization factors and gamma parameters to the Asimov dataset. Nearly all systematics remain unconstrained except for the renormalization and factorization scale of $t\bar{t}$ and $Z + \text{jets}$, alternative generators, and DeXTer eigenvector. The renormalization and factorization scales of $t\bar{t}$ and $Z + \text{jets}$ are estimated by varying the corresponding factorization and renormalization scales, as described in Section 6.4. These uncer-

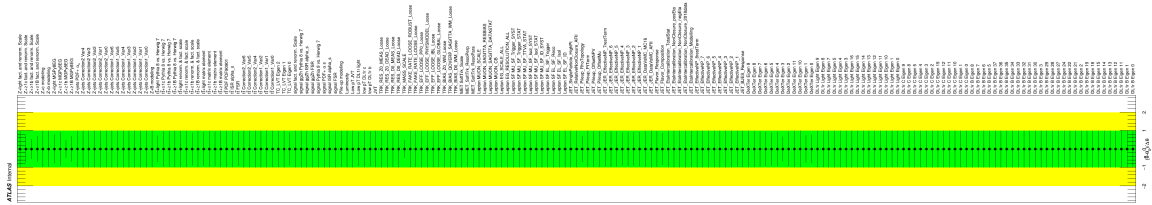


Figure 6.16: Nuisance parameters for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics. The $m_a = 25$ GeV mass hypothesis is considered. Fit results for the other mass hypotheses can be found in Appendix B.1.

to the high correlation between the free-floating normalization factors and alternative samples constrained by the CRs.

The systematics impacts on the signal strength are shown in Figure 6.18. The dominant contributions are from the theory systematics from signal modeling of the parton shower and scale variation followed by the flavor tagging systematics. The ranking of the flavor tagging systematics depends mainly on the mass hypothesis considered. The DeXTer tagging systematics dominate the lower mass hypotheses, whereas the higher mass hypotheses have DLr systematics with a higher impact on the signal strength.

6.6.2 Fits to blinded data

This section shows results for the analysis using full background-only fits to data with the blinded signal regions. The fit is first done in the control regions to estimate the normalization factors and nuisance parameters with data. The fitted values are later used to construct the Asimov dataset in the signal regions and repeat the fit with combined signal and control regions. Figures 6.20 and 6.21 show a comparison of the yields in data and simulation for the control and validation regions before and after the blinded fit to data. Figure 6.21 shows a comparison of data and simulation in the validation regions before and after performing the blinded fit to data. Since

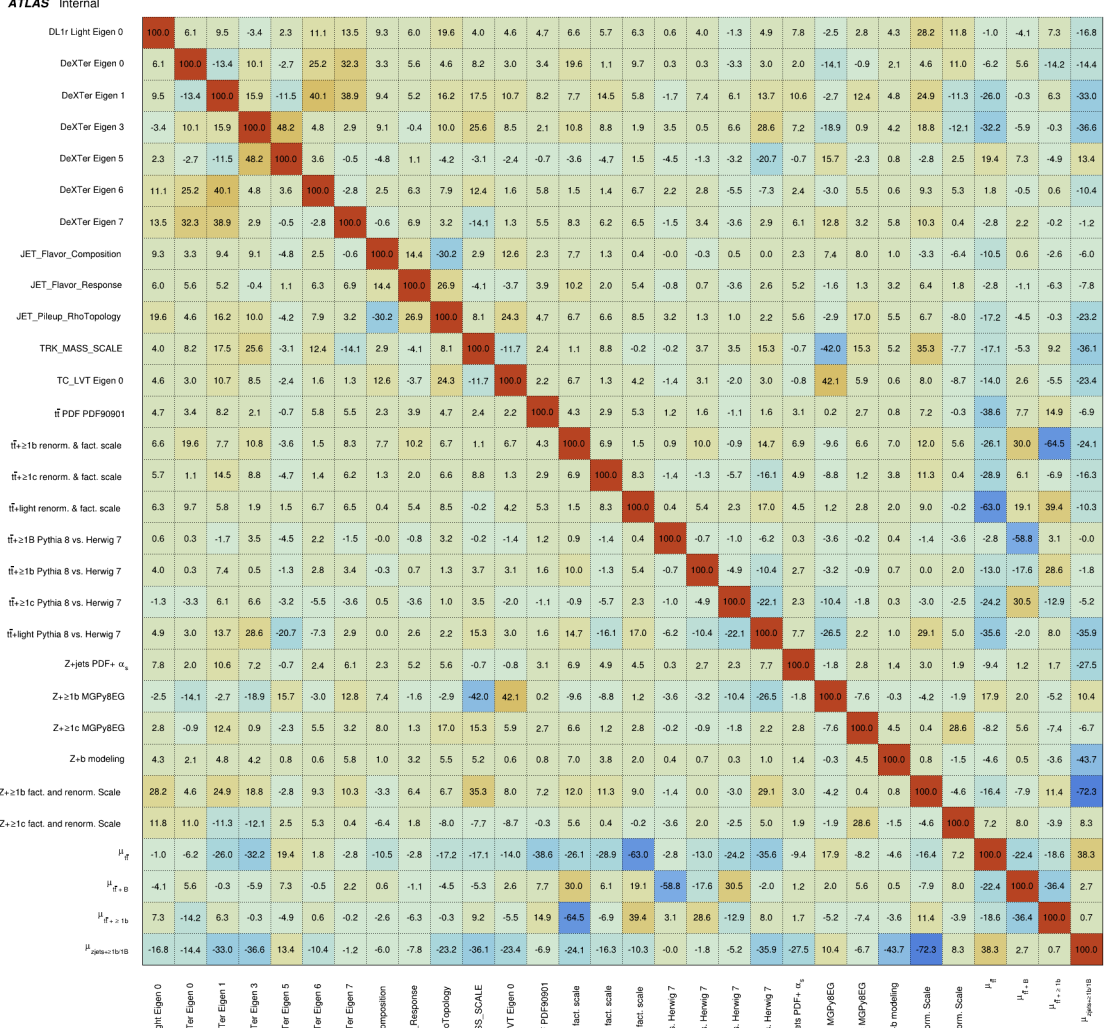


Figure 6.17: Correlation matrix of all nuisance parameters and scale factors for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics. The $m_a = 25$ GeV mass hypothesis is considered. Fit results for the other mass hypotheses can be found in Appendix B.1. Only NPs with at least one correlation larger than 25% are shown.

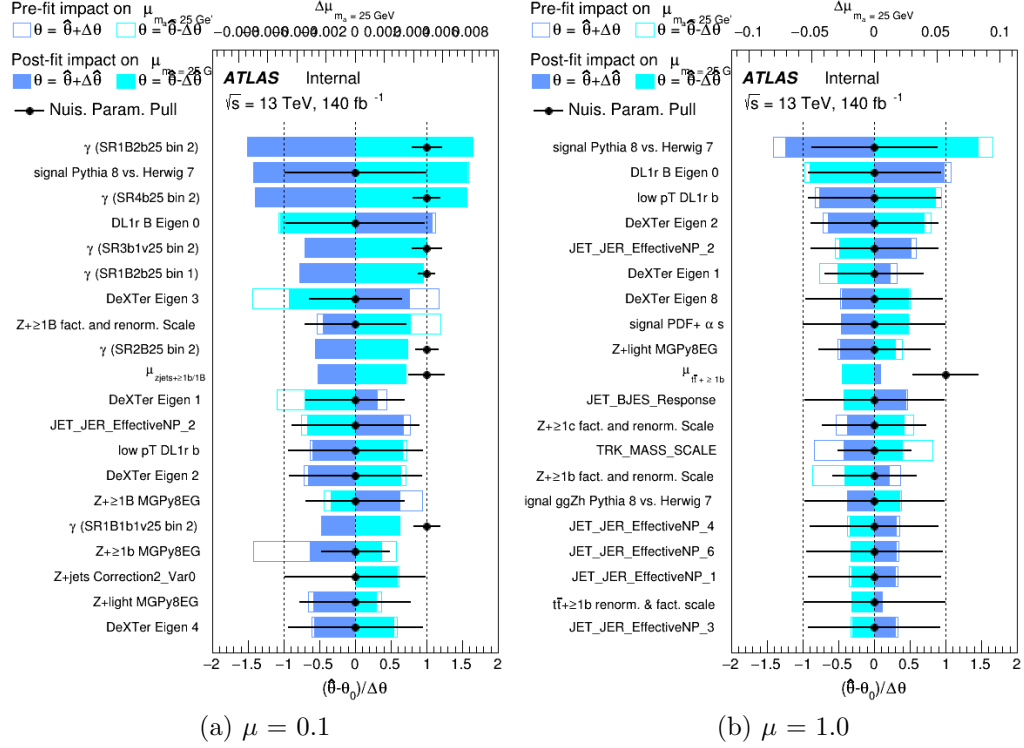


Figure 6.18: Rankings based on the impact on the signal strength for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics. The $m_a = 25$ GeV mass hypothesis is considered. The results are shown in two different levels of signal strength injected. (a) $\mu = 0.1$ (b) $\mu = 1.0$. Fit results for the other mass hypotheses can be found in Appendix B.1.

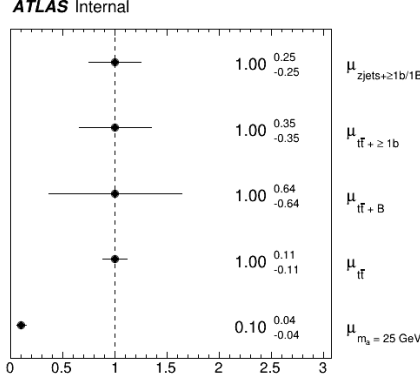


Figure 6.19: Normalization factors for the signal-plus-background fit to the Asimov data fit with signal and control regions using the full set of systematics. The $m_a = 25$ GeV mass hypothesis is considered. Fit results for the other mass hypotheses can be found in Appendix B.1.

the validation regions are not included in the fit, they serve as a cross-check of the background modeling and serve as additional validation of the fit. As can be seen, reasonable modeling of the backgrounds is observed before and after the blinded fit to data. Finally, the yields from the simulation in the signal regions are also shown. These plots show how the shaded area represents the total statistical and systematic uncertainties that are reduced in the control and validation regions due to the profiling of the fit and the improvement of the data/MC agreement.

The nuisance parameters, normalization factors, and gamma parameters for the background-only fit, are shown in Figures 6.22. The constraints on the different nuisance parameters are consistent with those found in the Asimov fit, but since these fits are with real data, some pulls are observed. All pulls are within 1σ .

6.6.3 Fits to unblinded data

This section shows results for the analysis using full background-only fits to data with the unblinded signal regions. Figures 6.23 and 6.24 show a comparison of the yields in data and simulation for the control and validation regions before and after

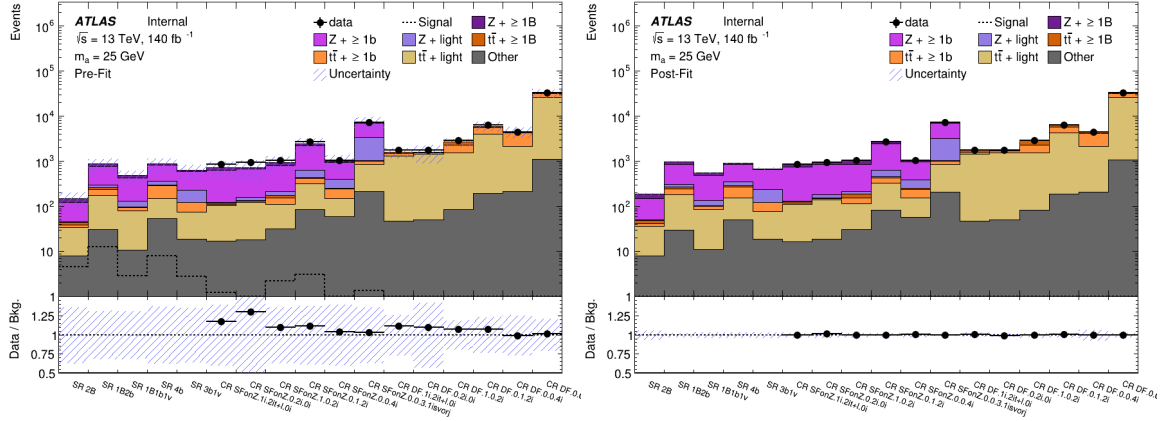


Figure 6.20: Predicted yields for the signal and control regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control and validation regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 25$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 1$ is shown before the fit. In contrast, the signal yield is set to 0 in the fit (background-only hypothesis). Fit results for the other mass hypotheses can be found in Appendix B.

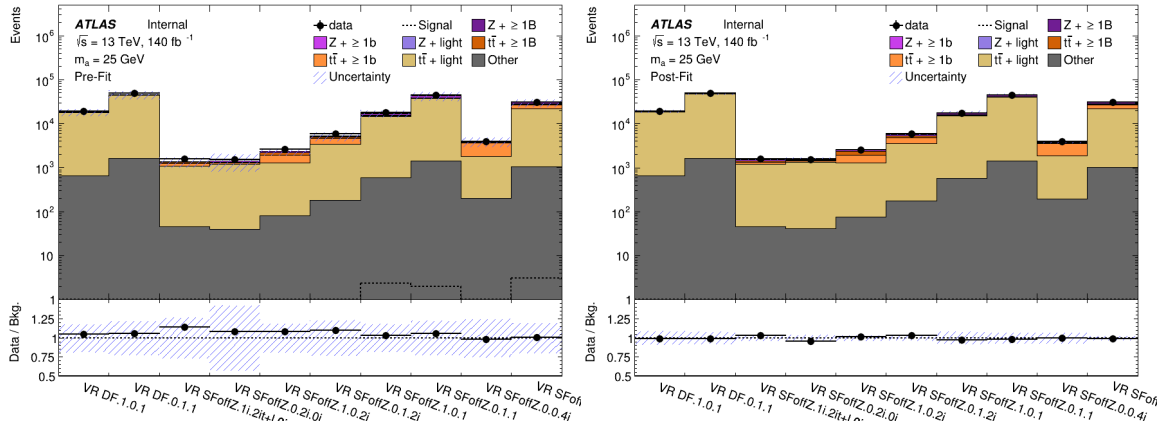


Figure 6.21: Predicted yields for validation regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the validation regions. The expected signal yield for the $m_a = 25$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(a \rightarrow bb) = 1$ is shown before the fit. In contrast, the signal yield is set to 0 in the fit (background-only hypothesis). Fit results for the other mass hypotheses can be found in Appendix B.

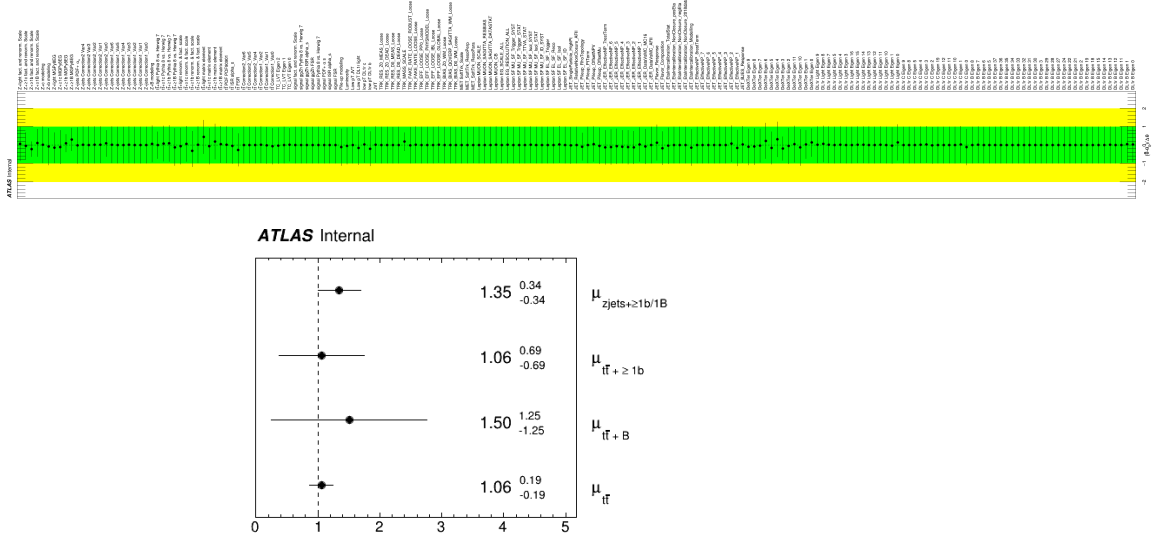


Figure 6.22: Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics. The $m_a = 25$ GeV mass hypothesis is considered. Fit results for the other mass hypotheses can be found in Appendix B.

the unblinded fit to data. Figure 6.24 shows a comparison of data and simulation in the validation regions before and after performing the unblinded fit to data. As can be seen, reasonable modeling of the backgrounds is observed before and after the unblinded fit to data. Finally, the yields from the simulation in the signal regions are also shown.

The nuisance parameters and normalization factors for the background-only fit are shown in Figure 6.25. The constraints on the different nuisance parameters are consistent with those found in the blinded fit. All pulls are within 1σ .

The detailed binning of the signal regions can be seen in Figure 6.26. In the Loose bin across all signal regions, the MC prediction agrees with the observed data. As for the signal-enriched bins, Tight and Medium, the data is slightly higher than the expected MC prediction assuming SM only hypothesis. Overall, the observed data are still consistent with the MC prediction and within the total uncertainties.

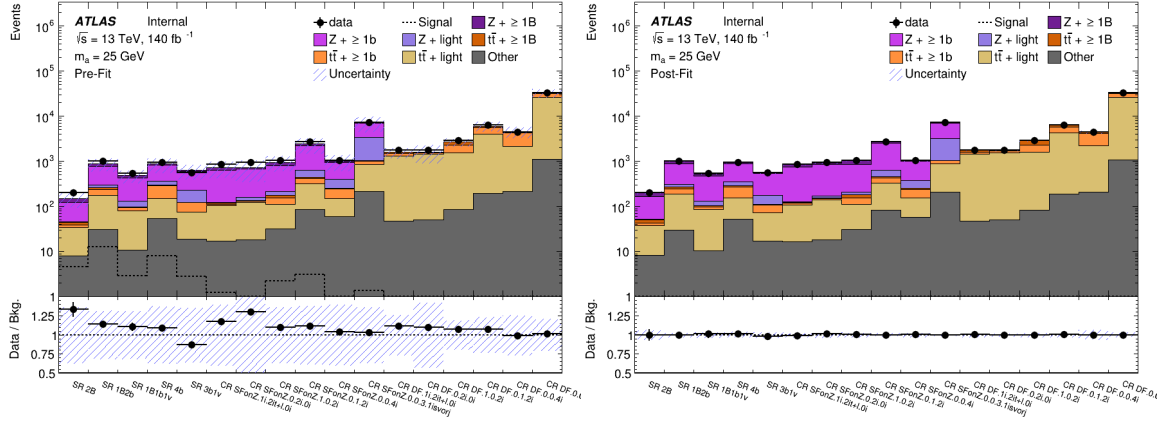


Figure 6.23: Predicted yields for the signal and control regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control and validation regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 25$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 1$ is shown before the fit. In contrast, the signal yield is set to 0 in the fit (background-only hypothesis). Fit results for the other mass hypotheses can be found in Appendix B.

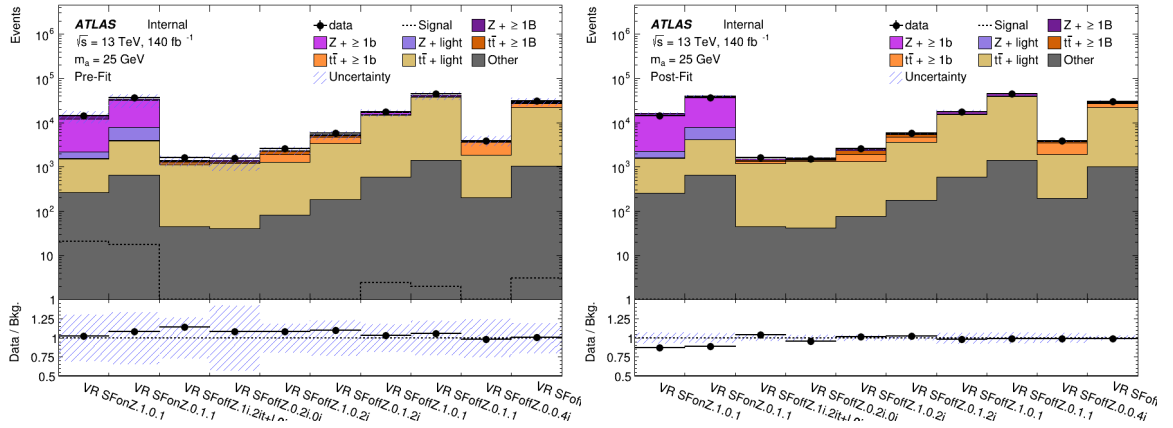


Figure 6.24: Predicted yields for validation regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the validation regions. The expected signal yield for the $m_a = 25$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(a \rightarrow bb) = 1$ is shown before the fit. In contrast, the signal yield is set to 0 in the fit (background-only hypothesis). Fit results for the other mass hypotheses can be found in Appendix B.

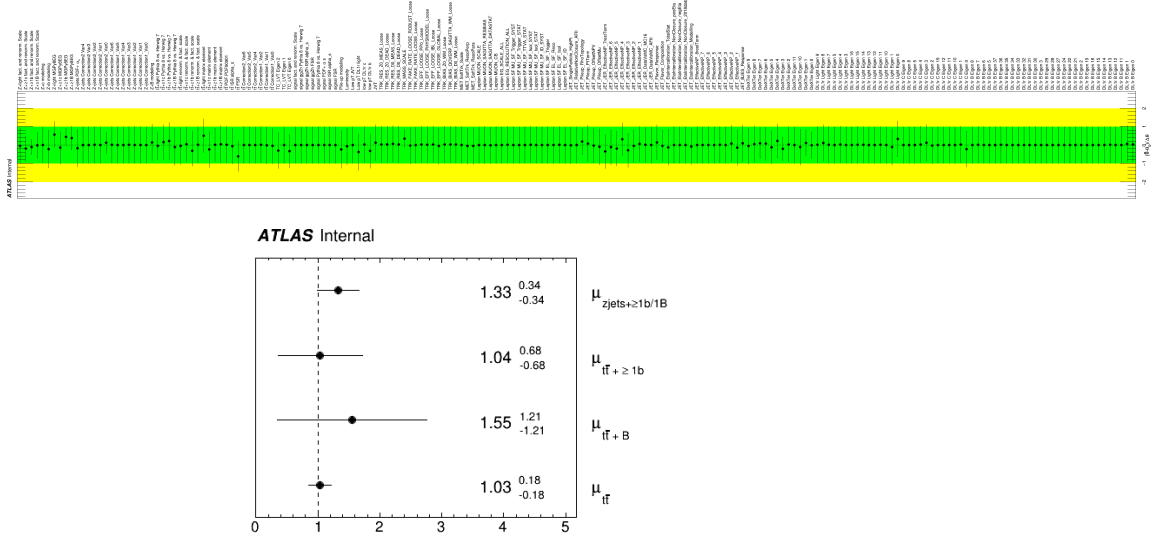


Figure 6.25: Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics. The $m_a = 25$ GeV mass hypothesis is considered. Fit results for the other mass hypotheses can be found in Appendix B.

6.6.4 Upper limits on $H \rightarrow 2a \rightarrow 4b$ production

The resulting expected upper limits on the $H \rightarrow 2a \rightarrow 4b$ production cross section times branching fraction are shown in Figure 6.27 for the mass range $12 < m_a < 60$ GeV. The total signal cross-section times branching ratio is parametrized as a signal strength modifying the SM total ZH cross-section:

$$\sigma_{ZH} \times BR(H \rightarrow aa) \times BR(a \rightarrow b\bar{b})^2 = \mu_{aa} \times \sigma_{ZH}^{SM} \quad (6.10)$$

A breakdown of the expected limit that separates the contribution from different SRs is shown in Figure 6.28. The fully merged 2B SR provides the strongest limit in the lower scalar mass phase space. The sensitivity of 2B SR decreases as the $a \rightarrow b\bar{b}$ decay starts to resolve into two b-jets as scalar mass increase. However, the sensitivity of 2B SR starts to increase when the chance of random crossing of b-hadrons from different scalars reconstructed as B-jets increases in the high scalar mass range. The mass range between 20 to 30 GeV is the transition region from fully

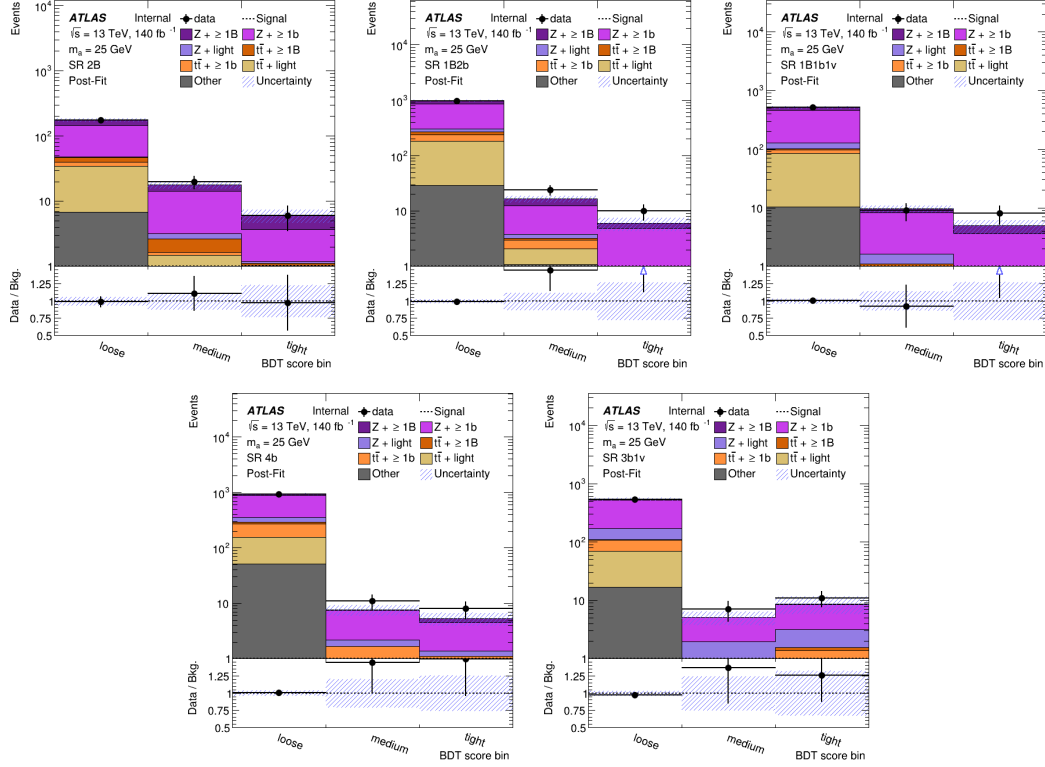


Figure 6.26: Post-fit predicted yields for the signal regions defined in the analysis (a) 2B (b) 1B2b (c) 1B1b1v (d) 4b (e) 3b1v in the unblinded fit to data with background-only hypothesis using the full set of systematics. The $m_a = 25$ GeV mass hypothesis is considered. Fit results for the other mass hypotheses can be found in Appendix B.

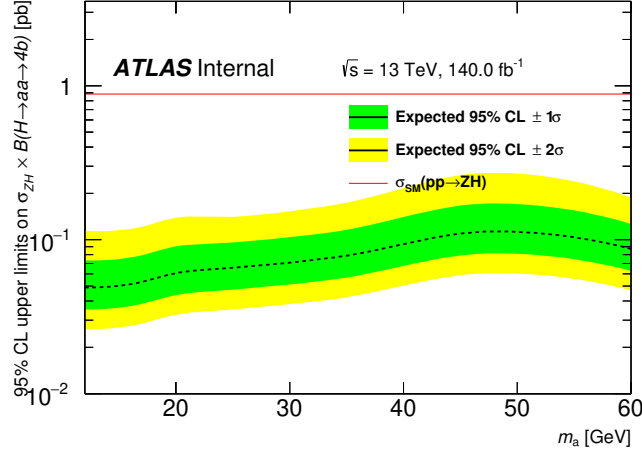


Figure 6.27: Expected (dashed line) 95% CL_{nj} upper limits on the $H \rightarrow 2a \rightarrow 4b$ cross-section times branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV. The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction and its ± 1 standard deviation uncertainty.

merged to fully resolved topology. Therefore, the semi-merged 1B2b and 1B1b1v SRs are the dominant contributions in this range. Finally, 4b and 3b1v SRs provide the best sensitivity in the higher scalar mass ranges.

The expected upper limits can also be set on the $H \rightarrow 2a \rightarrow 4b$ branching fraction, assuming the SM cross section of ZH shown in Figure 6.29

6.6.5 Upper limits on $H \rightarrow a_1 a_2 \rightarrow 4b$

As discussed in Section. 4.2.1, this search considers benchmark models with two additional scalars. The expected upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ production cross section times branching fraction are shown in Figure 6.31, for the benchmark mass points $(m_{a_1}, m_{a_2}) = (20, 30), (40, 60), \text{ and } (50, 70)$ GeV. The total signal cross-section times branching ratio is parametrized as a signal strength modifying the SM total ZH cross-section:

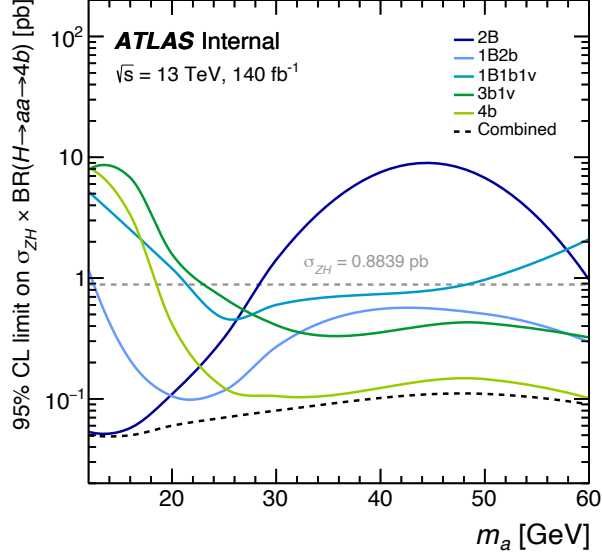


Figure 6.28: Breakdown of expected (dashed line) 95% CL upper limits on the $H \rightarrow 2a \rightarrow 4b$ cross-section times branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV from different signal regions.

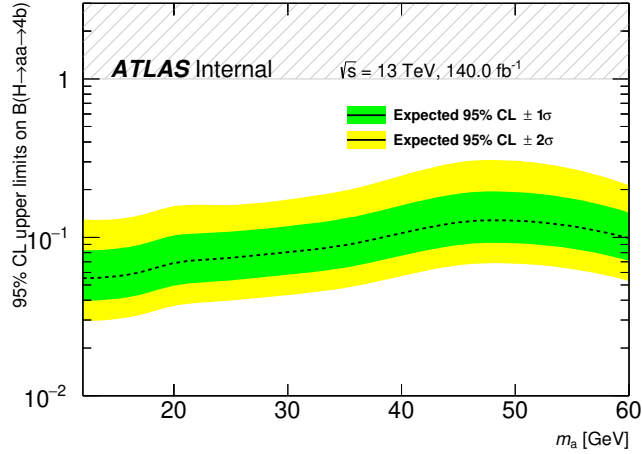


Figure 6.29: Expected (dashed line) 95% CL upper limits on the $H \rightarrow 2a \rightarrow 4b$ branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV assuming the SM cross section of ZH . The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction and its ± 1 standard deviation uncertainty.

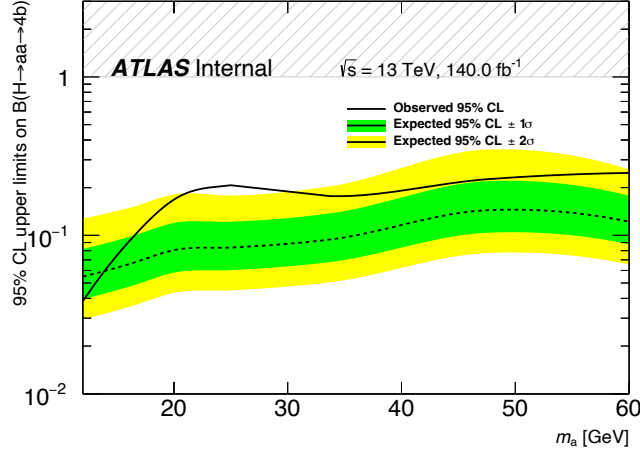


Figure 6.30: Observed (dashed line) 95% CL upper limits on the $H \rightarrow 2a \rightarrow 4b$ branching fraction as a function of the a mass for the mass range $12 < m_a < 60$ GeV assuming the SM cross section of ZH . The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction and its ± 1 standard deviation uncertainty.

$$\sigma_{ZH} \times BR(H \rightarrow a_1 a_2) \times BR(a_1 \rightarrow b\bar{b}) \times BR(a_2 \rightarrow b\bar{b}) = \mu_{a_1 a_2} \times \sigma_{ZH}^{SM} \quad (6.11)$$

A breakdown of expected limit contribution from different SRs is shown in Figure 6.32. Similar to the $H \rightarrow 2a \rightarrow 4b$, the contribution of various signal regions largely depends on the event topology.

6.7 Discussion and ideas for future searches

In this section, I will present some ideas for the next iteration of this analysis. Object reconstruction and identification are crucial factors that significantly impact the sensitivity of the search.

The identification of heavy-flavor jets is a crucial element in this search, as shown throughout this thesis. The use of tracking information directly in the algorithm has proven valuable. One aspect that holds potential for further improvement is the reconstruction of secondary vertices. Recent developments have shown promising

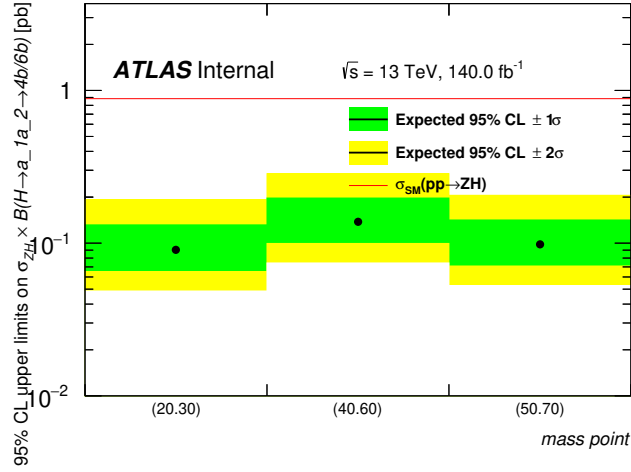


Figure 6.31: Expected 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ cross-section times branching fraction for different (m_{a_1}, m_{a_2}) values. The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit.

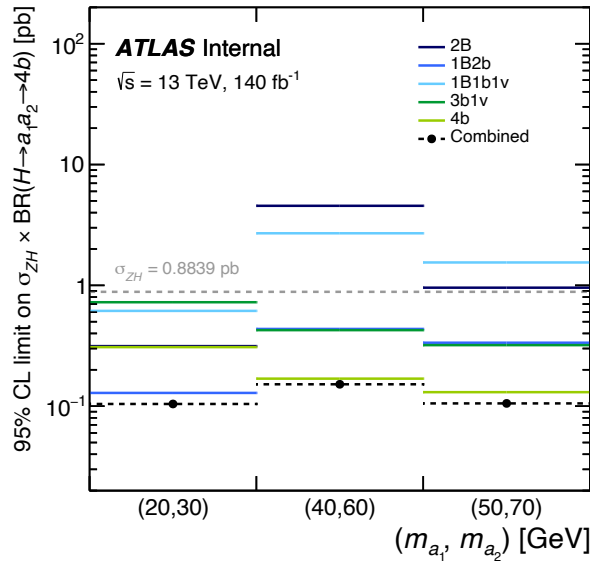


Figure 6.32: Breakdown of expected (dashed line) 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ cross-section times branching fraction for different (m_{a_1}, m_{a_2}) values from different signal regions.

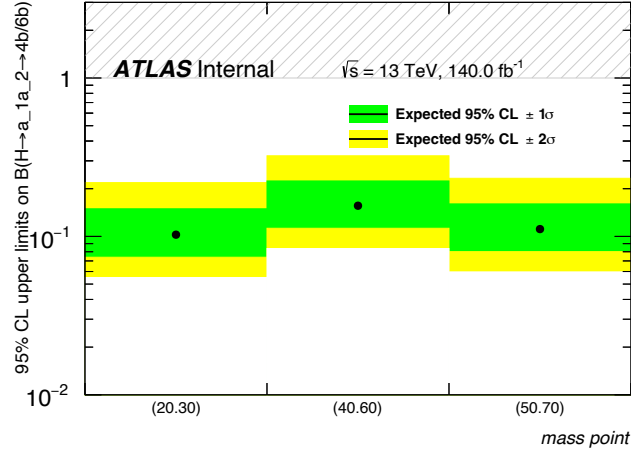


Figure 6.33: Expected 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ branching fraction for different (m_{a_1}, m_{a_2}) values assuming the SM cross section of ZH . The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit.

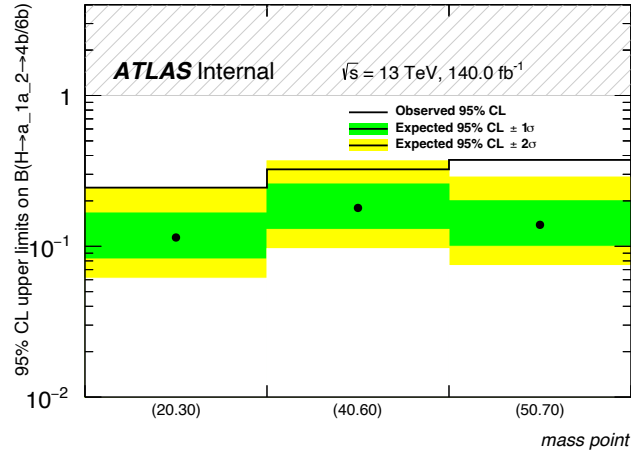


Figure 6.34: Observed 95% CL upper limits on the $H \rightarrow a_1 a_2 \rightarrow 4b$ branching fraction for different (m_{a_1}, m_{a_2}) values assuming the SM cross section of ZH . The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit.

results by incorporating the secondary vertex as an auxiliary task in flavor tagging algorithm training. This enables an end-to-end training approach for flavor tagging identification from tracks. Advanced machine learning architectures, such as transformers [179] and graph neural networks [180], have demonstrated impressive performance gains in various scientific domains, including natural language processing and image recognition.

Another challenging aspect is the calibration of the flavor tagging algorithm. Typically, calibrations are performed in a finite number of bins in kinematic distributions, resulting in correction factors that may exhibit discontinuities. This discontinuity is reflected as steps in the kinematics distribution used for calibration. Optimal transport techniques, such as those described in [181], could offer a method for achieving continuous calibration within the high-dimensional space. The same approach could also be explored to enhance the modeling of Monte Carlo backgrounds using data in the analysis.

Regarding the analysis strategy, one could consider streamlining the approach by adopting an end-to-end ML method. The end-to-end ML approach has outperformed manual feature engineering in various problems. However, the black-box nature of this approach often requires a deeper understanding of what the NN has learned. Incorporating auxiliary tasks can serve to interpret the results or decisions made by the complex NN model.

The ongoing paradigm shift towards integrating machine learning methods into the analysis will continue. Nevertheless, it is essential to recognize that incorporating domain knowledge remains crucial in wielding these tools effectively and wisely.

CHAPTER 7

CONCLUSION

This thesis presents the development of a general-purpose flavor tagging algorithm and searches for exotic Higgs decay to two spin-zero scalars and further decay to four b-quarks, $H \rightarrow aa/a_1a_2 \rightarrow b\bar{b}b\bar{b}$. The search is performed using pp collision at $\sqrt{s} = 13$ GeV collected by the ATLAS detector at LHC from 2015 to 2018 with a total integrated luminosity of 140 fb^{-1} . The main challenges of the search are to reconstruct and identify heavy flavor quarks from the decay of the spin-zero scalar by either jet or secondary vertex and event reconstruction.

A novel general-purpose flavor tagging algorithm for low-mass resonance search, DeXTer, is developed and calibrated with collision data. This algorithm is an end-to-end deep neural network using low-level tracking information and reconstructed secondary vertices to identify merged heavy flavor jets from boosted light resonance decays. The algorithm is designed to be agnostic to the resonance mass using the ensemble mixture of simulations from different signal masses as training datasets. Another feature is using adversarial neural network techniques during training to avoid algorithm learning to distinguish the underlying color charge of the particles. The calibration is carried out by selecting the $Z + \text{jets}$ and $t\bar{t}$ events in the data. The scale factor for both B- and b-jets are estimated. The resulting scale factors are between 1 to 1.91 as a function of jet p_T , tagging interval, and flavor of the jets with total uncertainties between 4 to 25%.

Equipped with special identification developed above and isolated secondary vertex reconstruction, the searches for $H \rightarrow aa/a_1a_2 \rightarrow b\bar{b}b\bar{b}$ analysis is designed by

considering all possible combinations of the final state. This allows us to have a unified strategy to consider resolved, semi-merged, and merged final states of four b-quark. The analysis strategy begins by using a deep neural network to predict the correct pairing of the reconstruction objects. This allows the reconstruction of the Higgs and new scalar candidates and defines the event topology. The reconstructed Higgs and new scalar candidates and other event kinematics are used in boosted decision trees to build signal-to-background discrimination in the signal regions. Control regions for the two main backgrounds, $t\bar{t}$ and $Z + \text{jets}$, are defined to constrain the overall normalization and flavor fractions with data. All regions defined are used in a binned profile likelihood fit. The fit model is verified with the Asimov dataset and blinded fit with data blinded in signal regions. The results are dominated by statistical uncertainty, and the largest systematics are from the limited number of MC events and flavor tagging systematics. The observed limits are obtained with unblinded fit to all the regions defined. Two small excesses are observed in two signal mass points: $m_a = 25$ GeV and signal hypothesis with two new scalars with $(m_{a_1}, m_{a_2}) = (50, 70)$ GeV with a local significance of 2.8σ and 3.2σ , respectively. The results are still statistically limited; hence more data are needed to verify the further observed excess. The limits are set on the cross-section times branching ratio of $\sigma_{ZH} \times Br(H \rightarrow aa \rightarrow b\bar{b}b\bar{b})$ and $\sigma_{ZH} \times Br(H \rightarrow a_1a_2 \rightarrow b\bar{b}b\bar{b})$.

APPENDIX A

DETAILS OF THE NEURAL NETWORK TRAINING

A.1 DNN energy-momentum estimate regression

Hyperparameters and training dataset preparation

The training is done with all the available signal MC samples and part of the background MC with a total of 0.97 million jets, where 16% of jets are held out as validation set. The architecture uses Keras [147] with the TensorFlow backend [148]. An early stopping that monitors the improvement of the validation set loss is used during the training. The training stopped after 20 consecutive epochs without any improvement in the validation set loss. The Adam optimizer [149] is used with a learning rate of 0.001 and batch size of 512. Another independent set of 0.25 million jets is kept out for performance studies.

A grid search over the hyperparameters, including the layer and number of the node in each of the NN blocks, is performed. The number is summarized in Table A.1

A.2 Hypothesis testing neural network

This section presents the detailed performance studies for the pairing hypothesis NN. The versatility of the pairing hypothesis NN can be best demonstrated by examining the performance in different truth categories of the signal events. Based on the number of reconstructed objects, the signal events can be classified into three types, fully merged, semi-merged, and fully resolved. Firstly, the fully merged category is a reconstructed event with two DeXTer jets that reconstructed the $4b$ final state from the $ZH, H \rightarrow aa \rightarrow 4b$ signal sample. Secondly, the semi-merged events

NN block name	List of nodes per layer	Activation function
Track NN	[128, 128, 64, 32, 32, 32]	ELU
SV NN	[128, 128, 64, 32, 32, 32]	ELU
$\text{Ex}k_t^{(2)}$ NN	[32, 32, 16, 16, 16]	ELU
Muon NN	[32, 32, 32, 16, 16]	ELU
Globel NN	[256, 128, 64, 32, 16, 16, 8, 8, 8, 1]	ReLU

Table A.1: List of hyperparameters for DNN Energy-momentum estimate regression. The number of nodes per layer is provided in the order from the input layer to the output layer, where the first and last numbers in the list represent the input and output nodes, respectively. The ELU activation function [143] is used for all the NNs except the last nodes in the Globel NN, in which the ReLU activation function is used [182]

are reconstructed as one DeXTer jet and two b-tagged jets, one b-tagged jet, and one secondary vertex. Finally, the events with four b-objects reconstructed separately in the detector are classified as fully resolved events.

Hyperparameters and training dataset preparation

The training is done with 3.7 million events where 20% of events are held out as validation set. More events are available from the SM background MC simulation, which are kept as hold-out samples for testing. The architecture is implemented using Keras [147] with the TensorFlow backend [148]. An early stopping that monitors the improvement of the validation set loss is used during the training. The training stopped after 20 consecutive epochs without any improvement in the validation set loss. The Adam optimizer [149] is used with a learning rate of 0.001 and batch size of 1024. Another independent set of 3 million events is kept out for performance studies shown in the next section.

A grid search over the hyperparameters, including the layer and number of the node in each of the NN blocks, is performed. The number is summarized in Table A.2

NN block name	List of nodes per layer	Activation function
AK8 NN	[4, 64, 64, 32, 32, 32]	ReLU
AK4 NN	[6, 64, 64, 32, 32, 32]	ReLU
SV NN	[6, 64, 64, 32, 32, 32]	ReLU
a -boson NN	[33, 128, 128, 128, 64]	ReLU
Higgs NN	[71, 128, 128, 128, 64, 8]	Sigmoid

Table A.2: List of hyperparameters for pairing hypothesis testing NN. The number of nodes per layer is provided in the order from the input layer to the output layer, where the first and last numbers in the list represent the input and output nodes, respectively. The ReLU activation function [182] is used for all the NN except the last output node from the Higgs NN, in which the Sigmoid activation function is used

APPENDIX B

RESULTS FOR OTHER SIGNALS

RESULTS FOR OTHER SIGNALS

B.1 Fits to the Asimov dataset with signal injected

B.1.1 $m_a = 12$ GeV

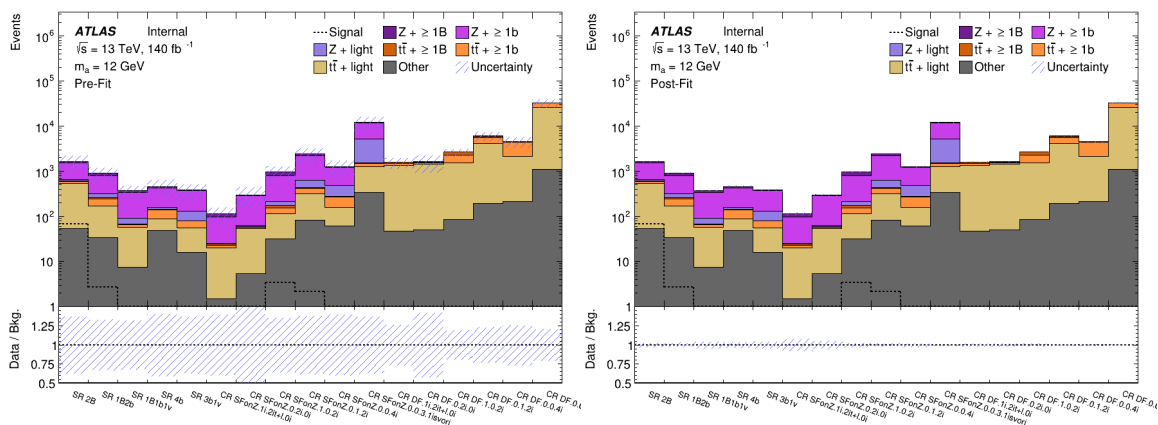


Figure B.1: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 12$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

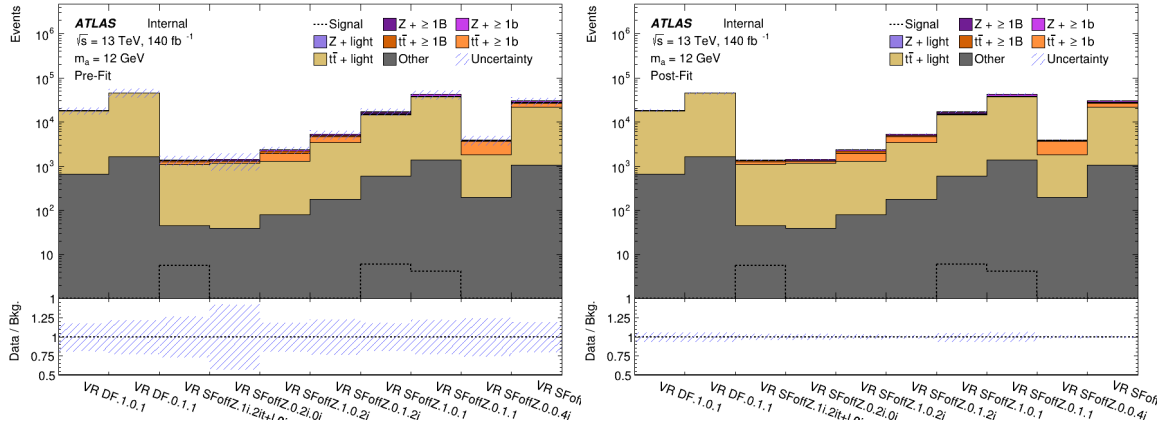


Figure B.2: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 12$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

B.1.2 $m_a = 16$ GeV

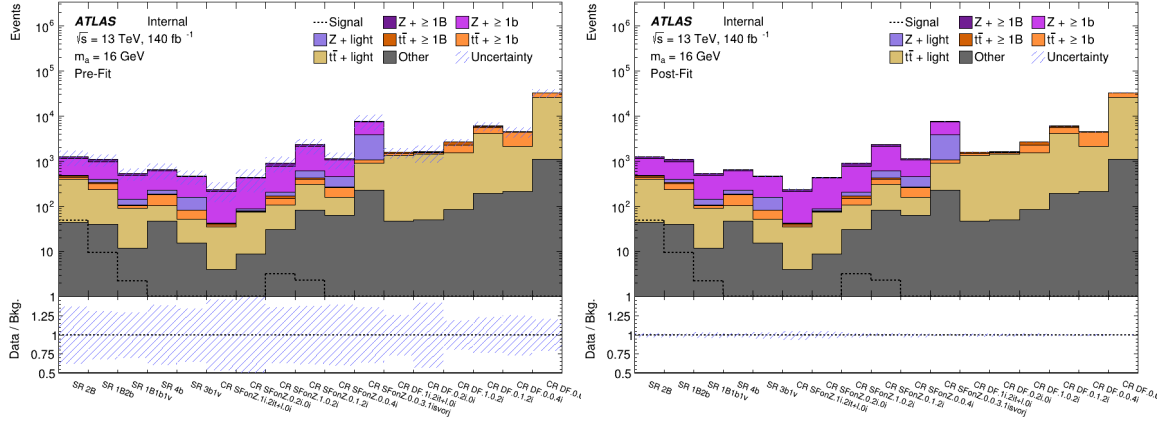


Figure B.4: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 16$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

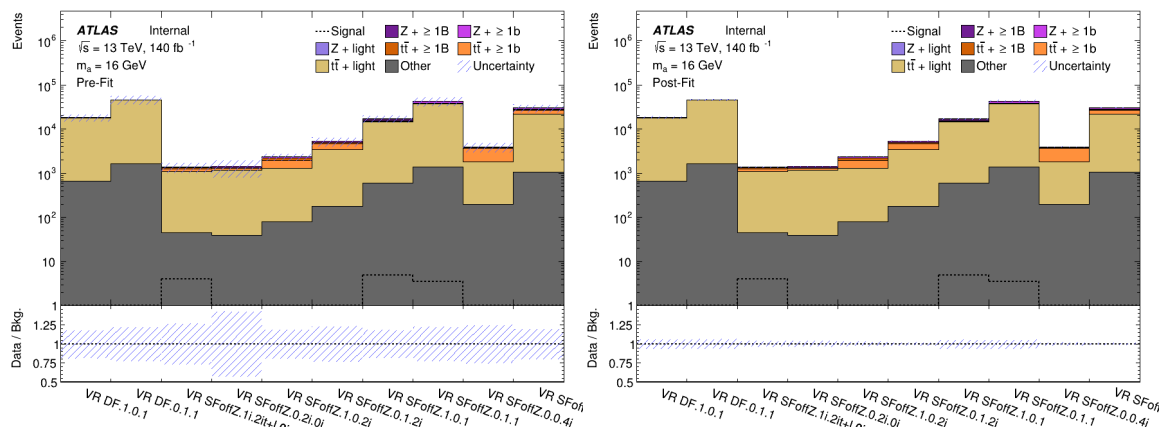


Figure B.5: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 16$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

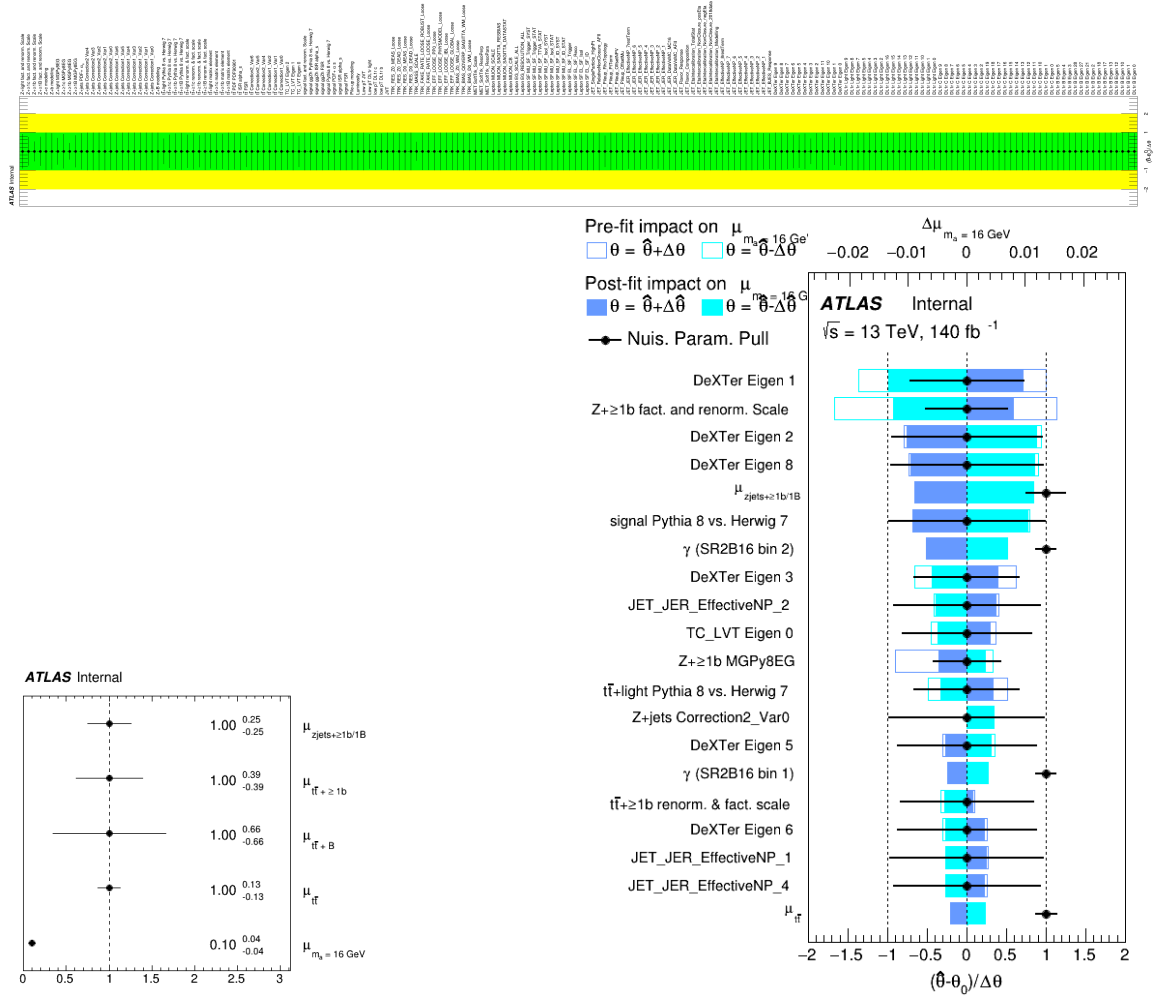


Figure B.6: Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a=16$ GeV mass hypothesis is considered.

B.1.3 $m_a = 20 \text{ GeV}$

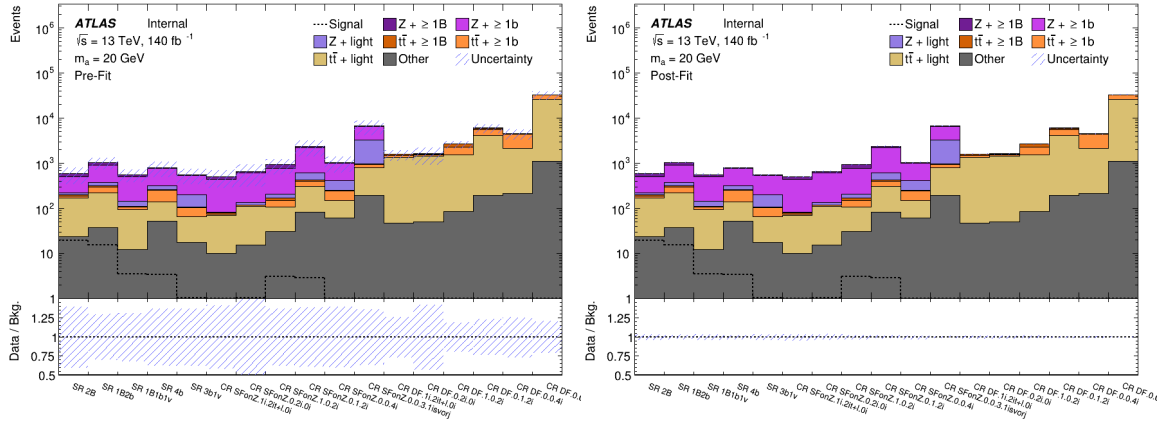


Figure B.7: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

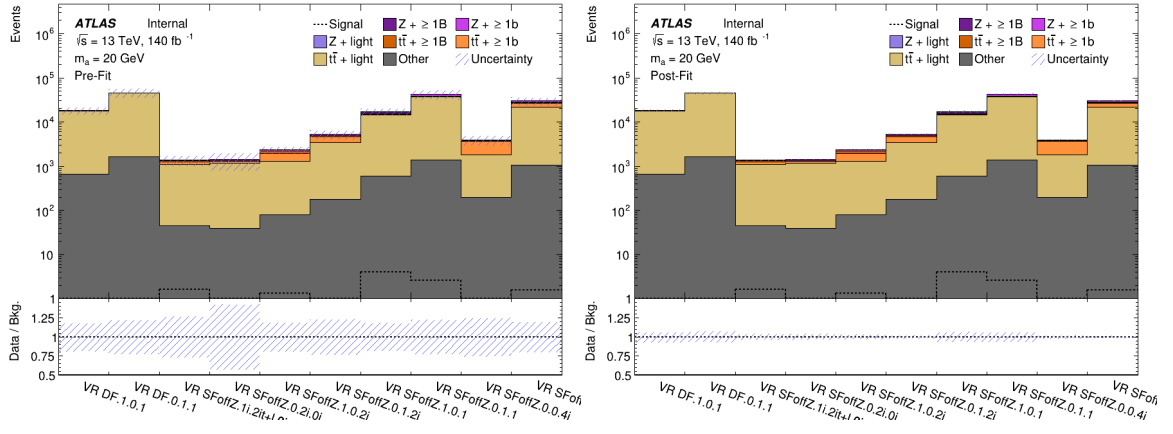


Figure B.8: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

B.1.4 $m_a = 30 \text{ GeV}$

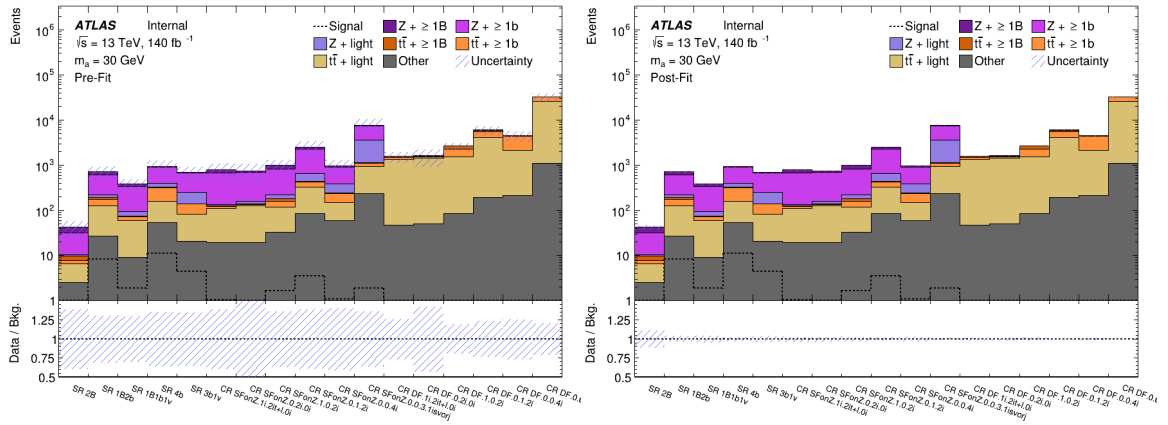


Figure B.10: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 30$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

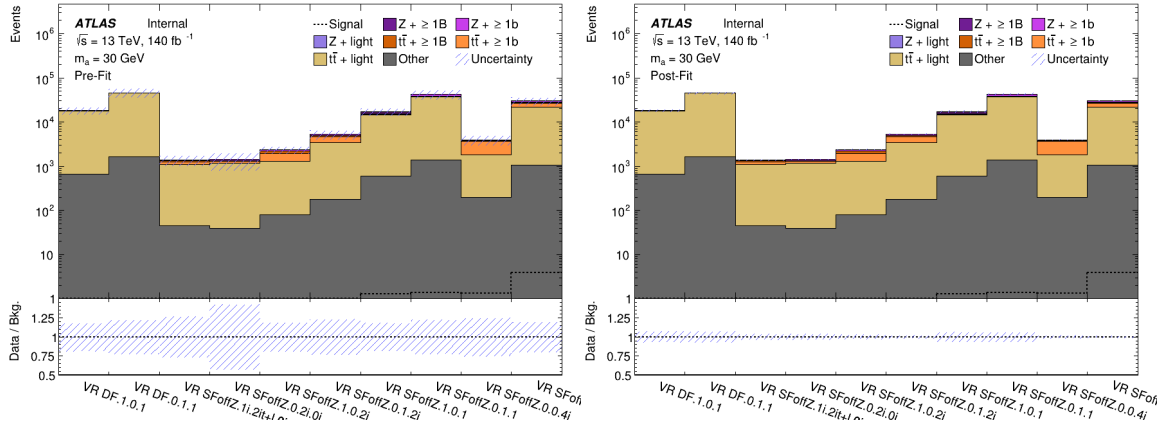


Figure B.11: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 30$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

B.1.5 $m_a = 50$ GeV

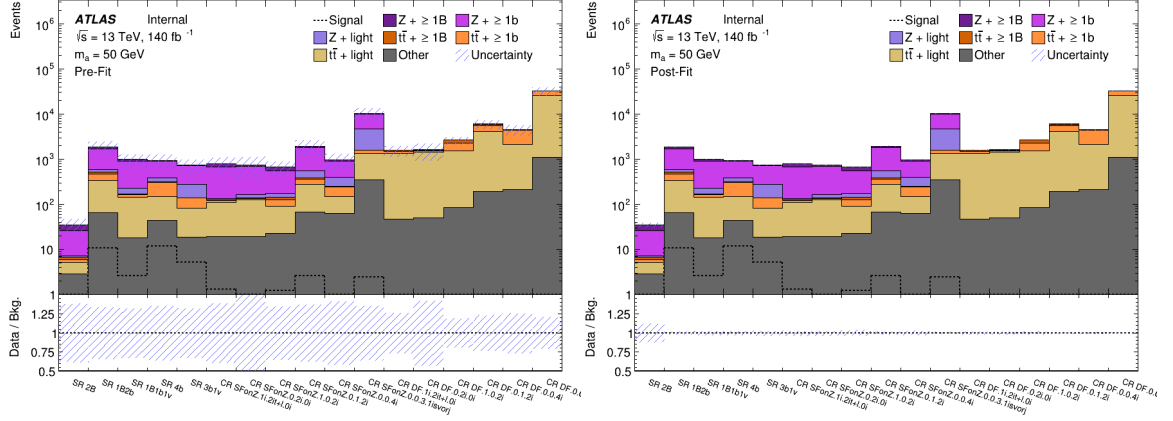


Figure B.13: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 50$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

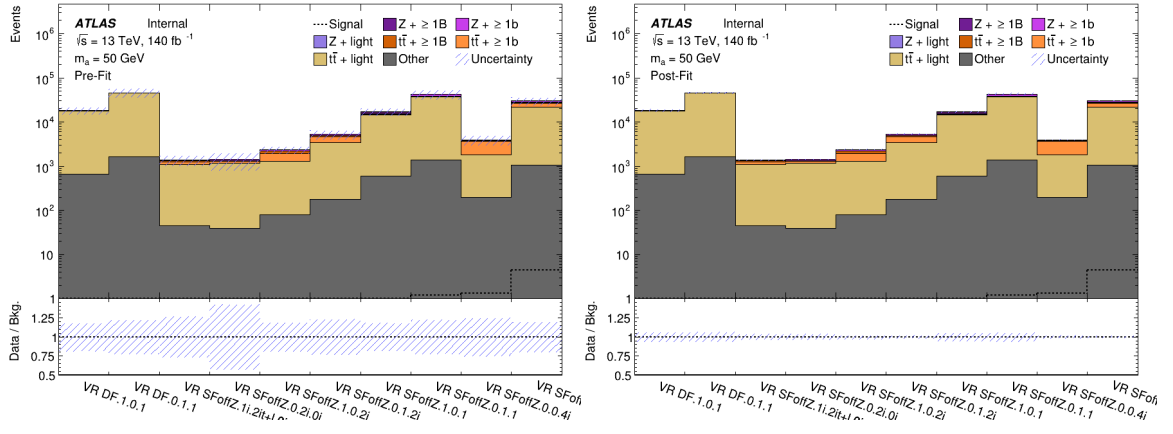


Figure B.14: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 50$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

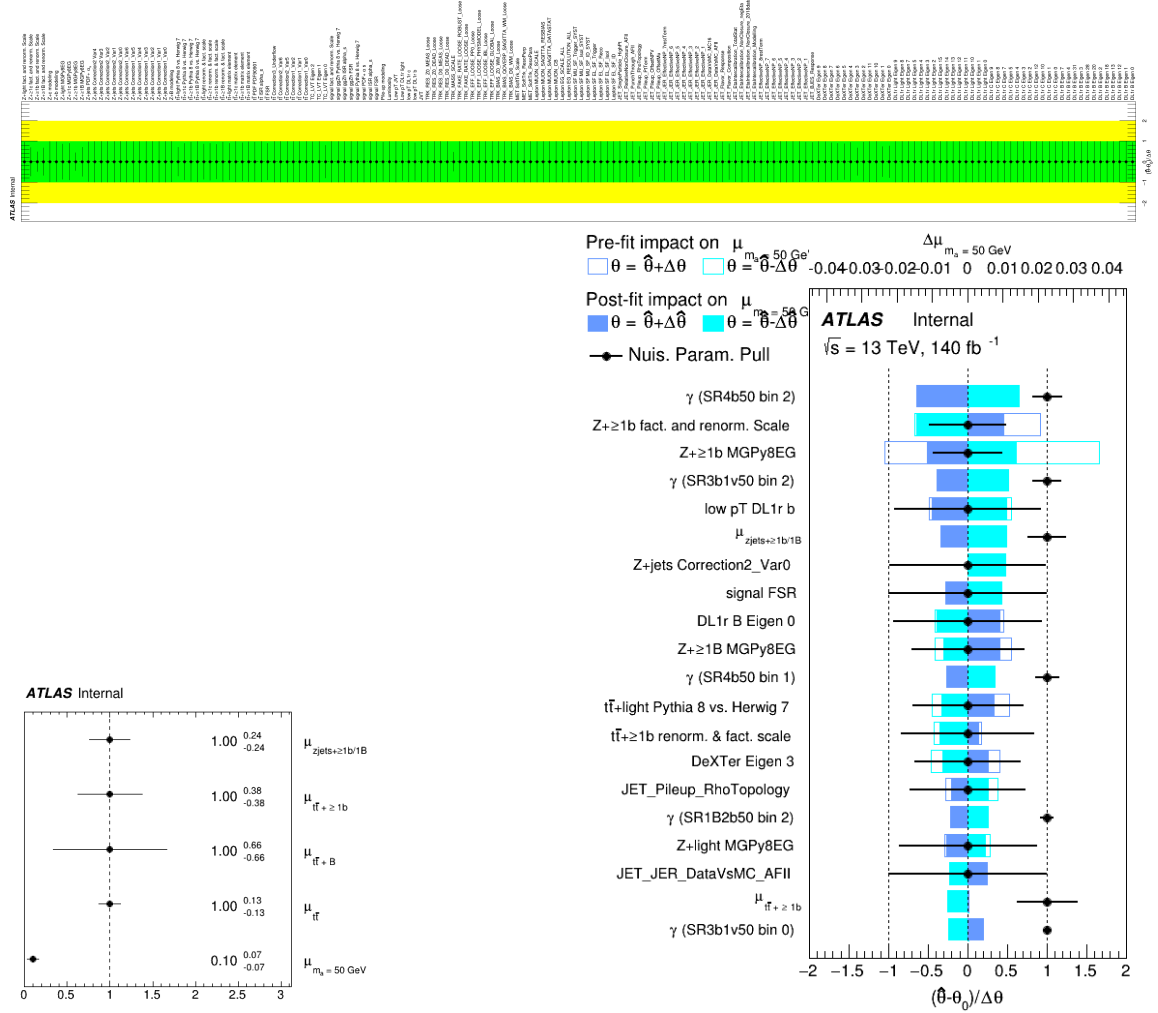


Figure B.15: Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a=50$ GeV mass hypothesis is considered.

B.1.6 $m_a = 60$ GeV

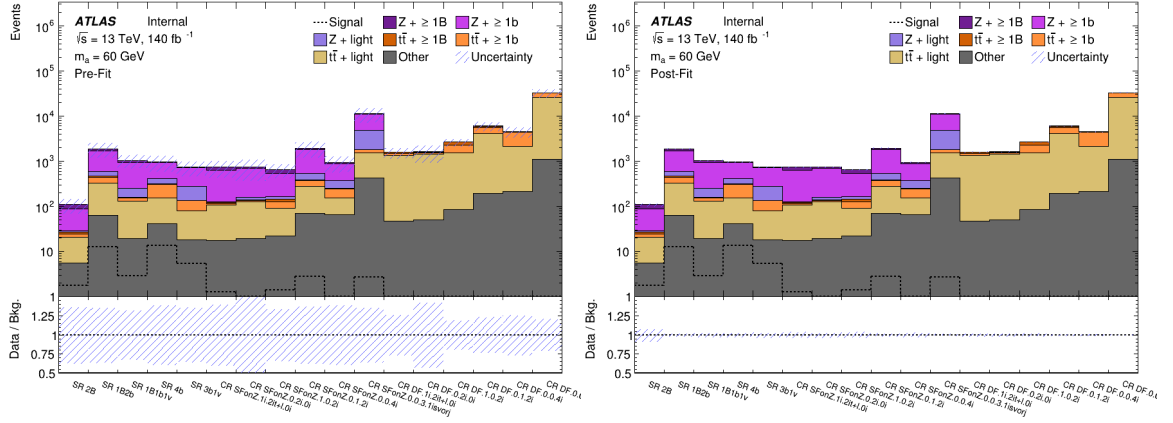


Figure B.16: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 60$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

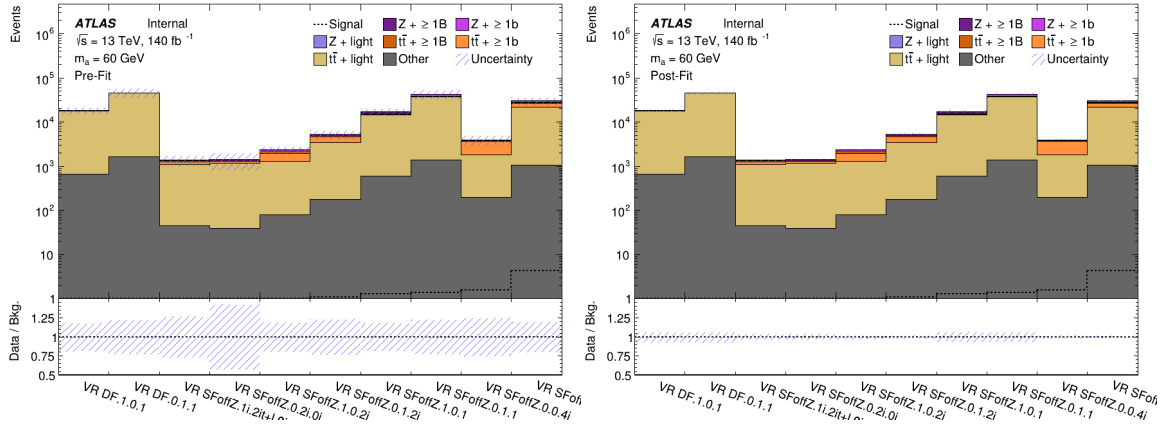


Figure B.17: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 60 \text{ GeV}$ mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow 2a \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

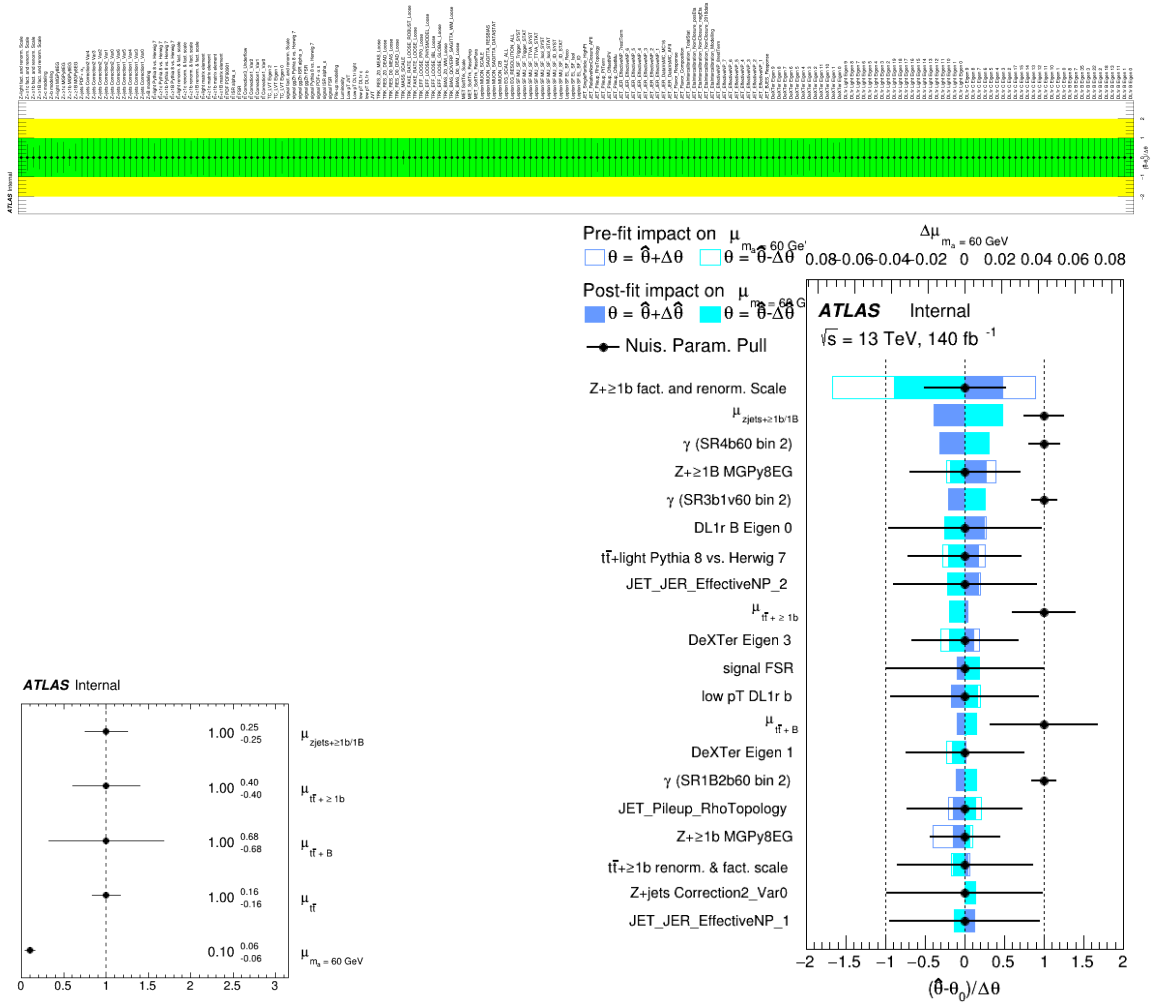


Figure B.18: Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_a=60$ GeV mass hypothesis is considered.

B.1.7 $m_{a_1} = 20 \text{ GeV}$, $m_{a_2} = 30 \text{ GeV}$

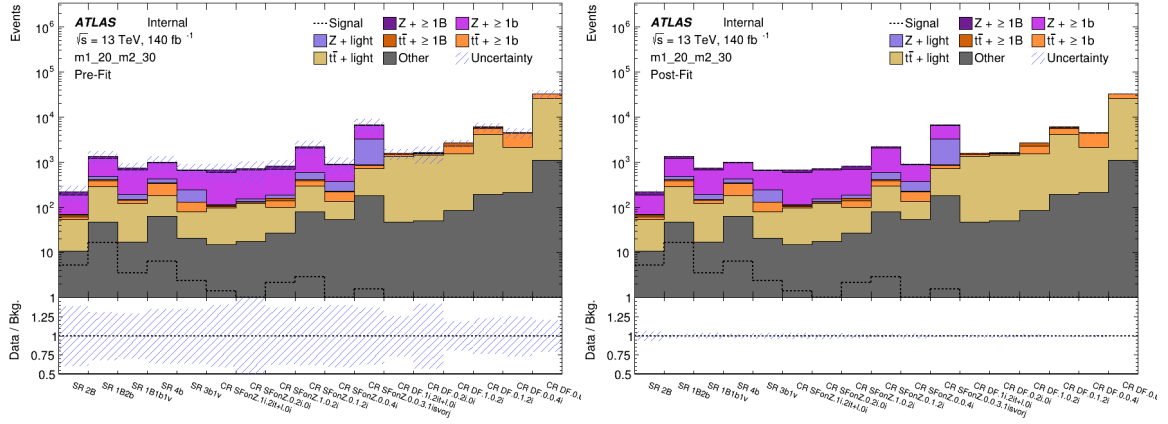


Figure B.19: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_{a_1}=20$ GeV, $m_{a_2}=30$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

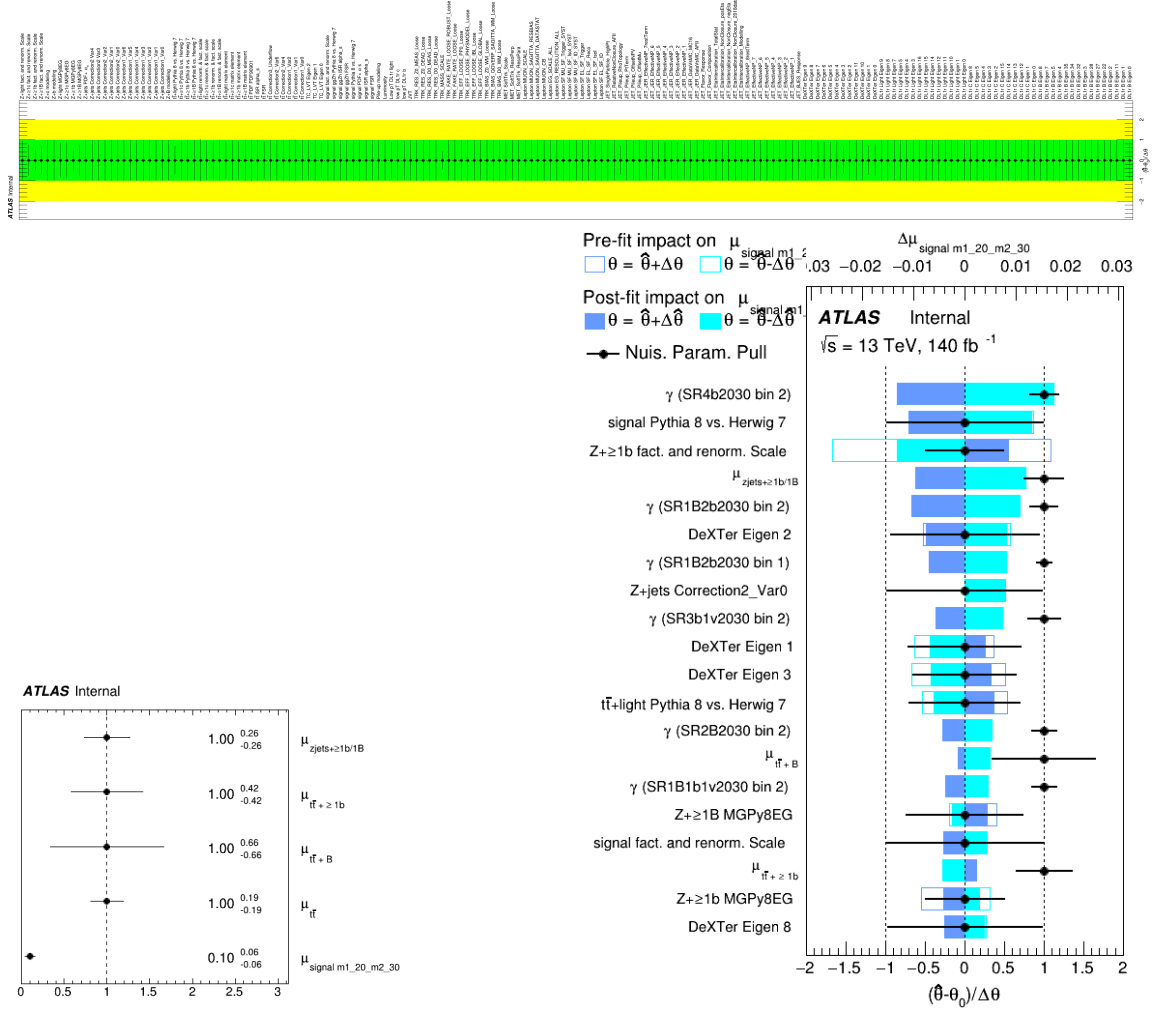


Figure B.21: Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV mass hypothesis is considered.

B.1.8 $m_{a_1} = 40 \text{ GeV}$, $m_{a_2} = 60 \text{ GeV}$

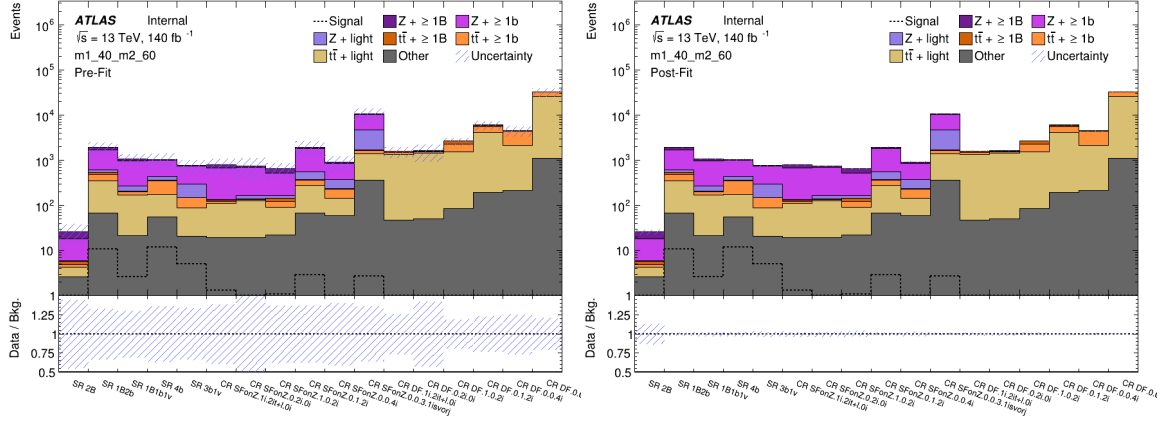


Figure B.22: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_{a_1} = 40 \text{ GeV}$, $m_{a_2} = 60 \text{ GeV}$ mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

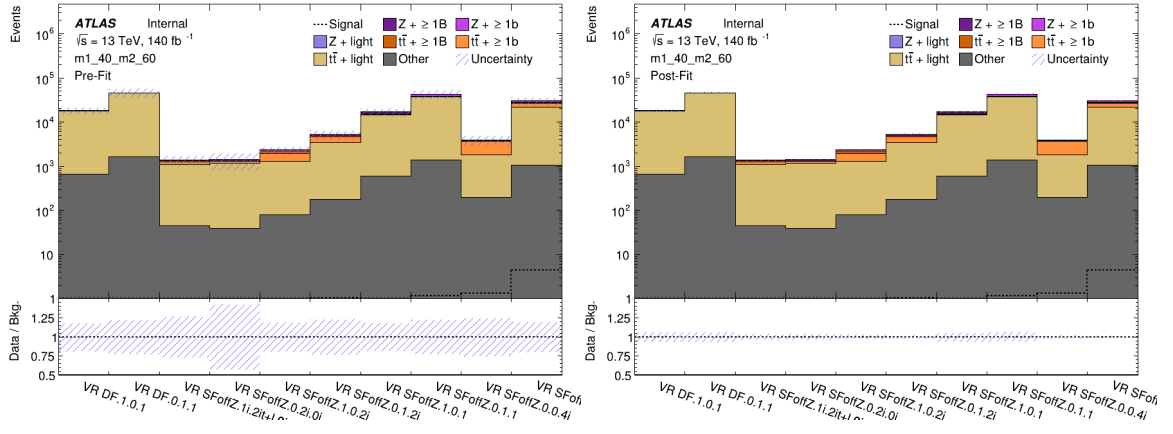


Figure B.23: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

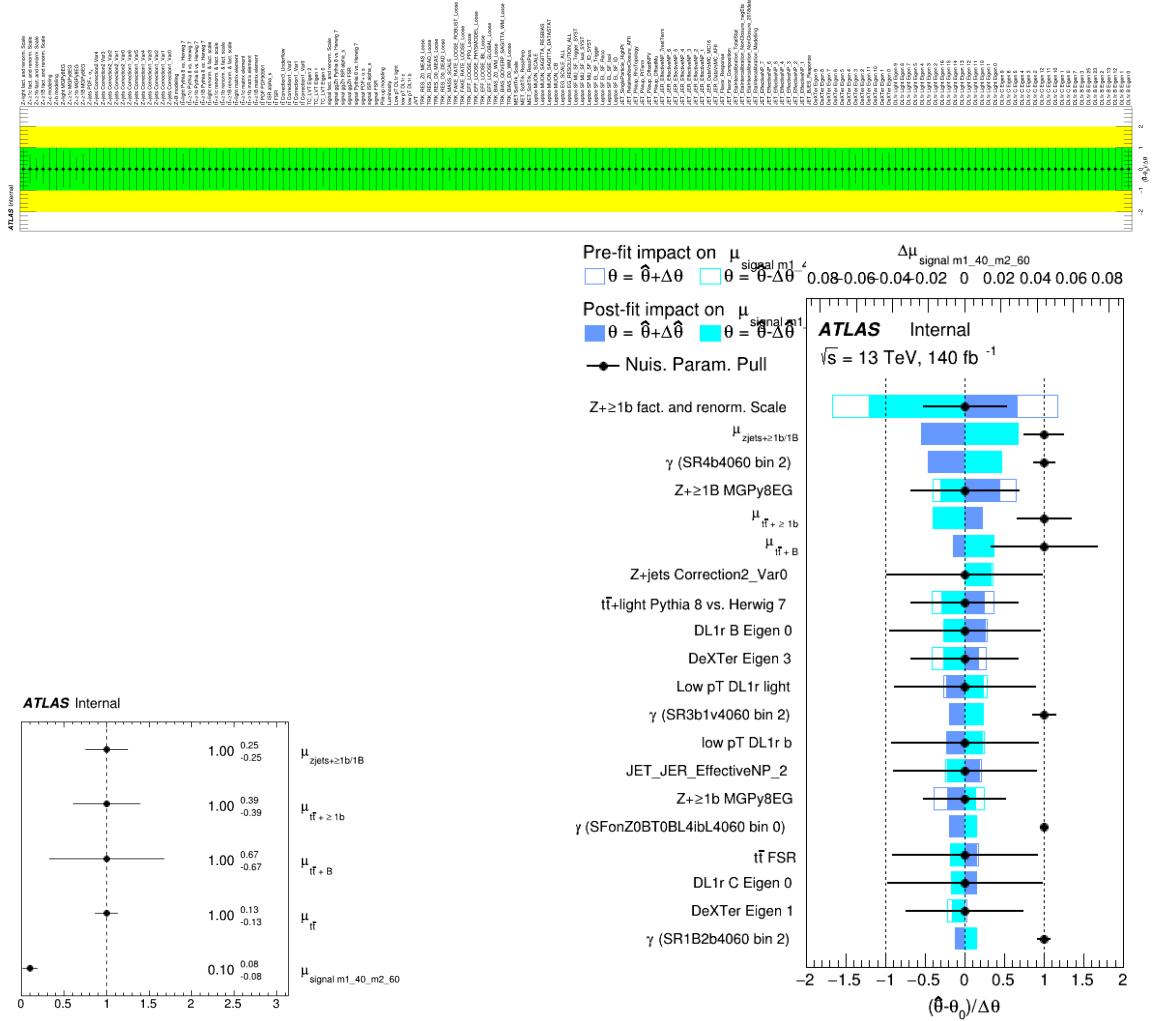


Figure B.24: Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeV mass hypothesis is considered.

B.1.9 $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV

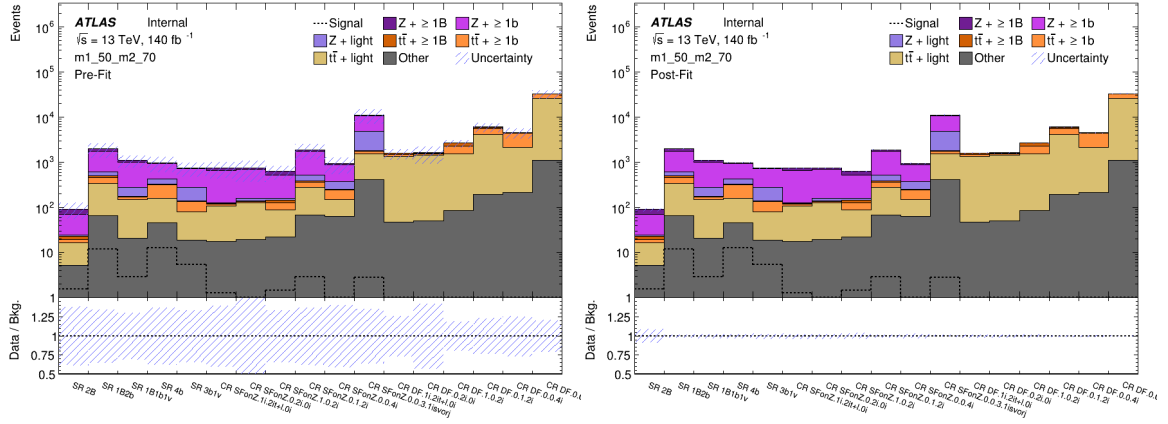


Figure B.25: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_{a_1}=50$ GeV, $m_{a_2}=70$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

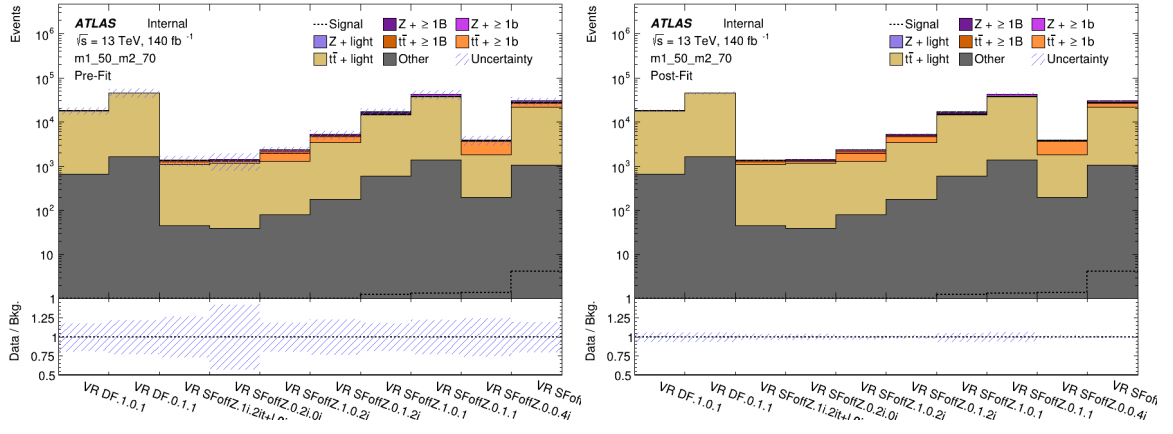


Figure B.26: Validation regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is included in the Asimov dataset and shown with the dashed line.

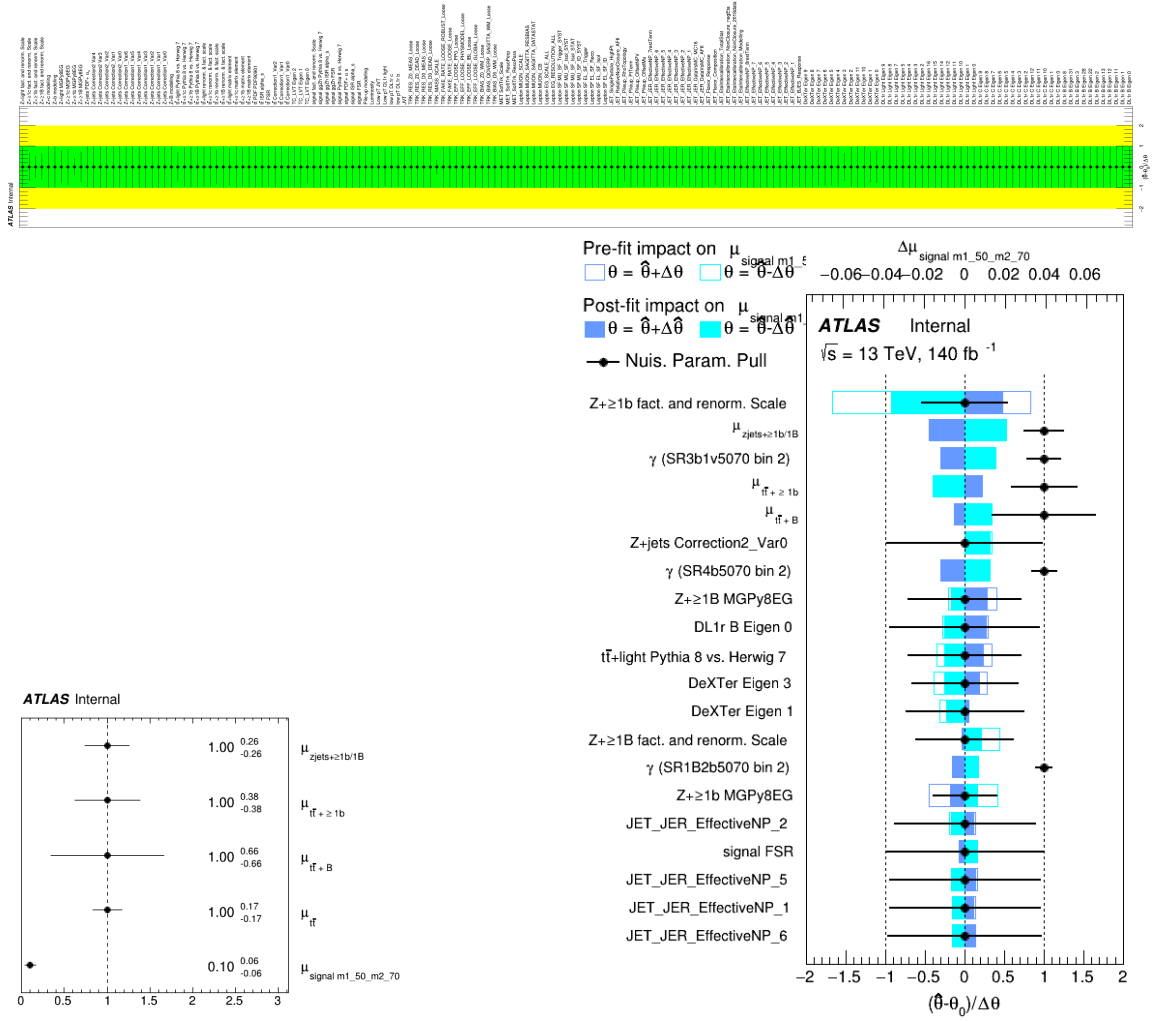


Figure B.27: Nuisance parameters, normalization factors, and rankings based on the impact on the signal strength for the Asimov data fit using the full set of systematics. The $m_{a_1}=50$ GeV, $m_{a_2}=70$ GeV mass hypothesis is considered.

B.2 Fits to blinded data

B.2.1 $m_a = 12$ GeV

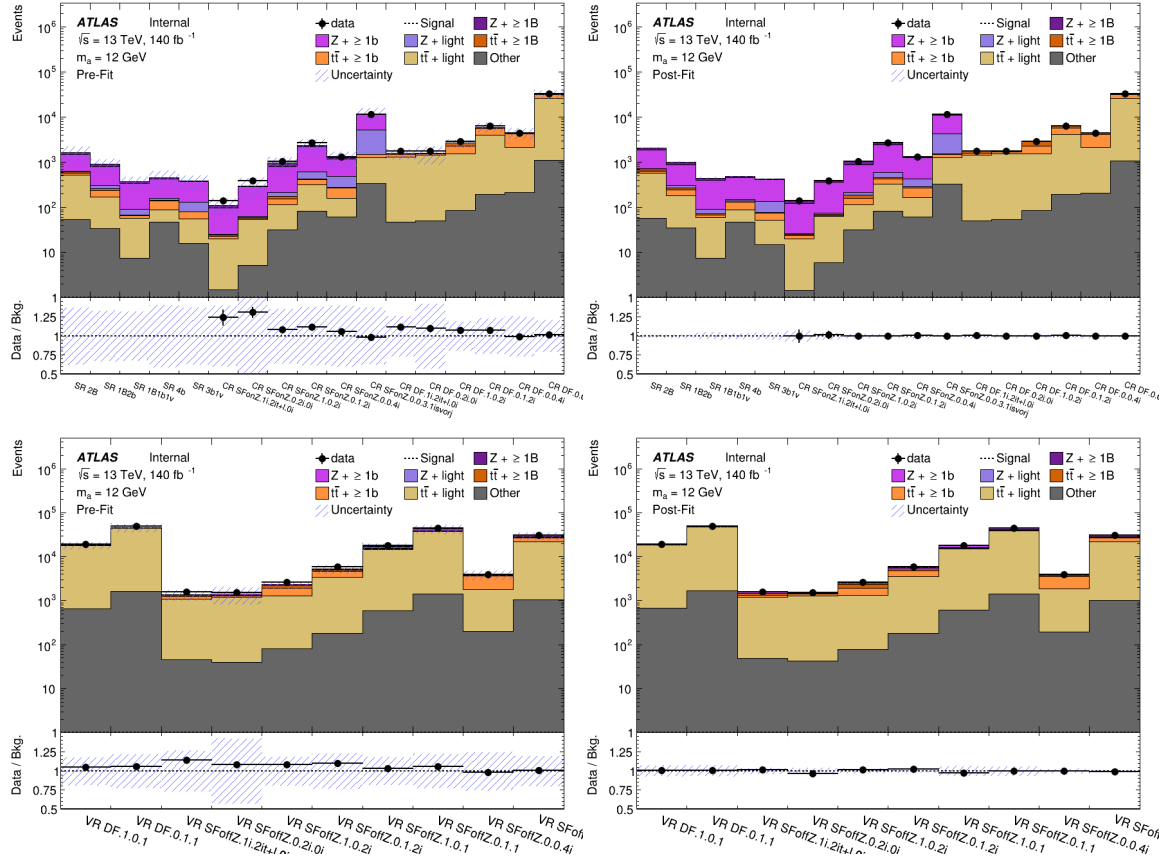


Figure B.28: Here the $m_a = 12$ GeVmass hypothesis is considered. Predicted yields for the signal, control (top), and validation (bottom) regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 12$ GeVmass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow aa \rightarrow 4b) = 0.1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

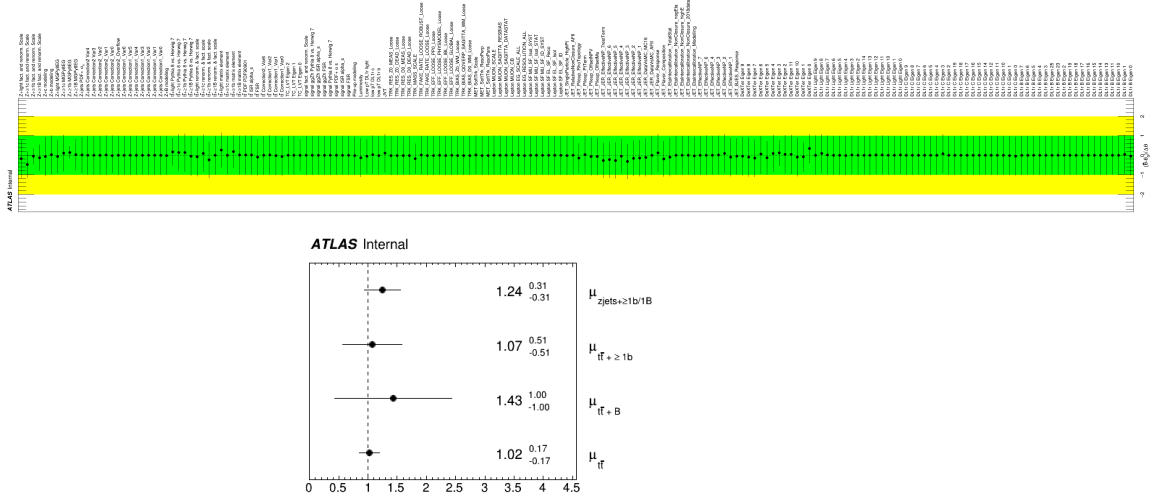


Figure B.29: Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics for the $m_a = 12$ GeV mass hypothesis.

B.2.2 $m_a = 16$ GeV

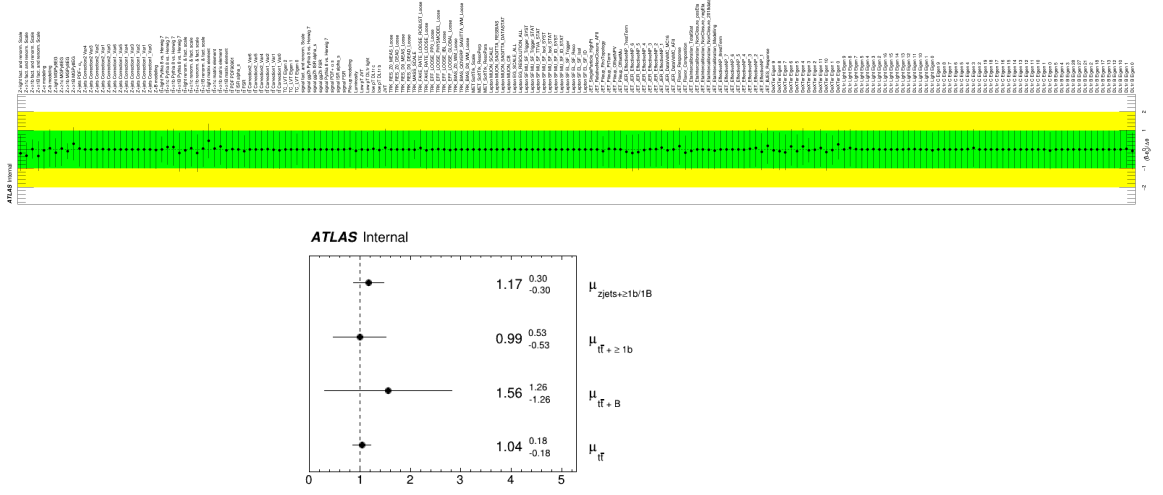


Figure B.31: Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics for the $m_a = 16$ GeV mass hypothesis.

B.2.3 $m_a = 20$ GeV

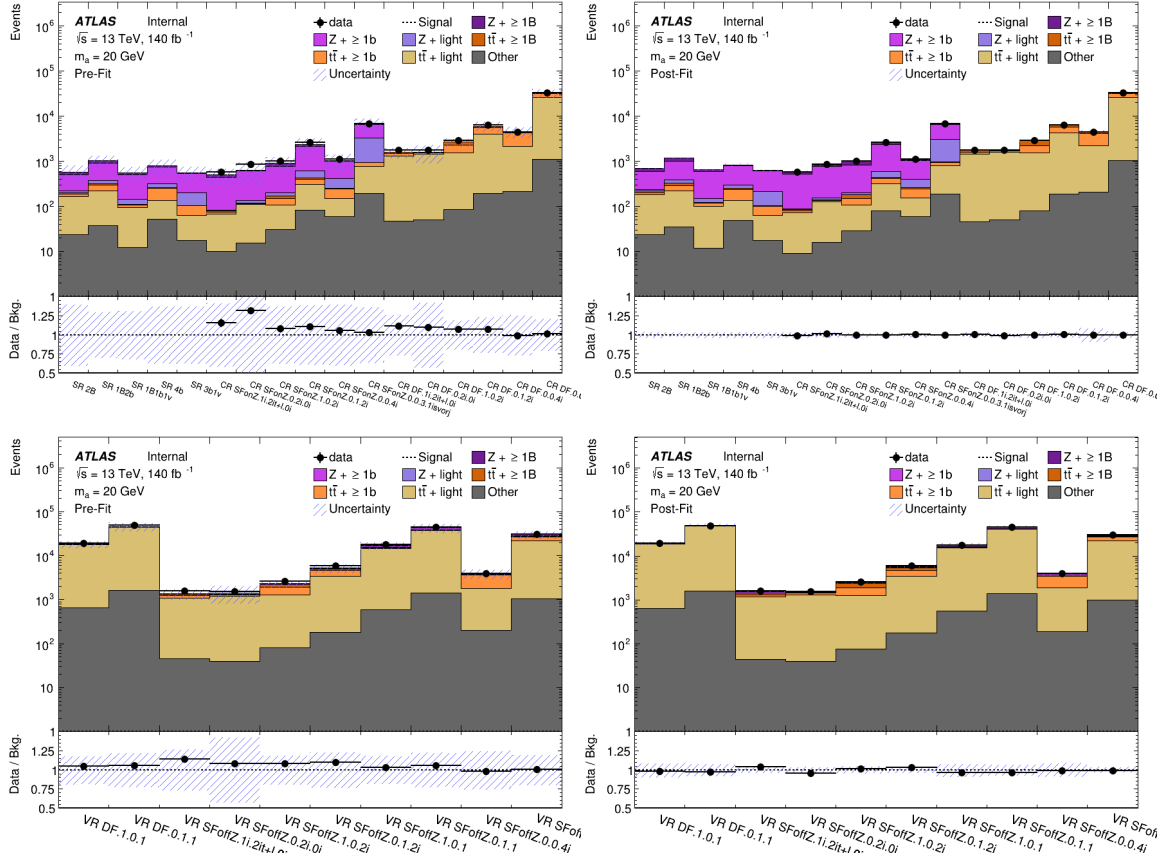


Figure B.32: Here the $m_a = 20$ GeV mass hypothesis is considered. Predicted yields for the signal, control (top), and validation (bottom) regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow aa \rightarrow 4b) = 0.1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

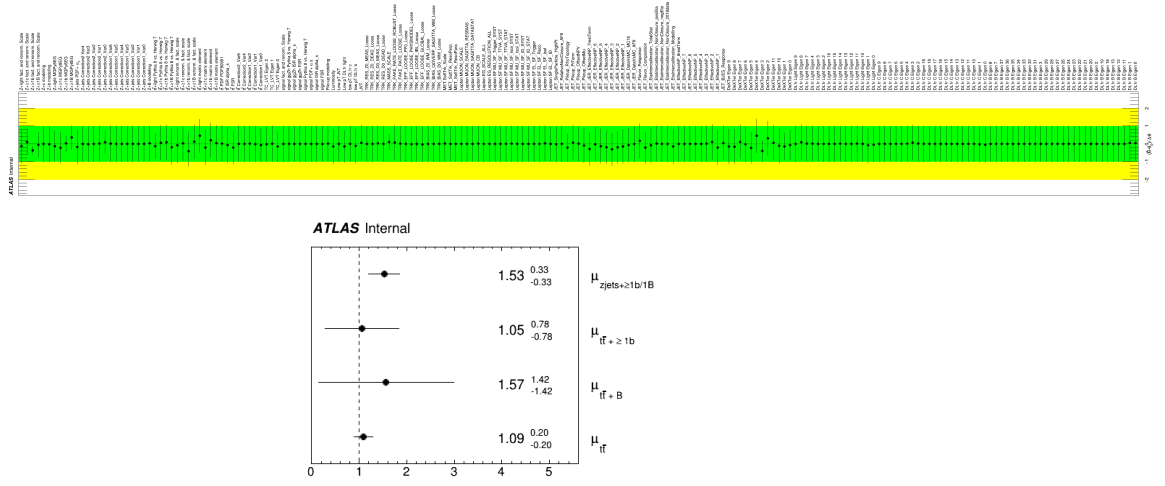


Figure B.33: Nuisance parameters (top) and normalization factors (bottom right) for the blinded data fit using the full set of systematics, the $m_a = 20$ GeV mass hypothesis.

B.2.4 $m_a = 30$ GeV

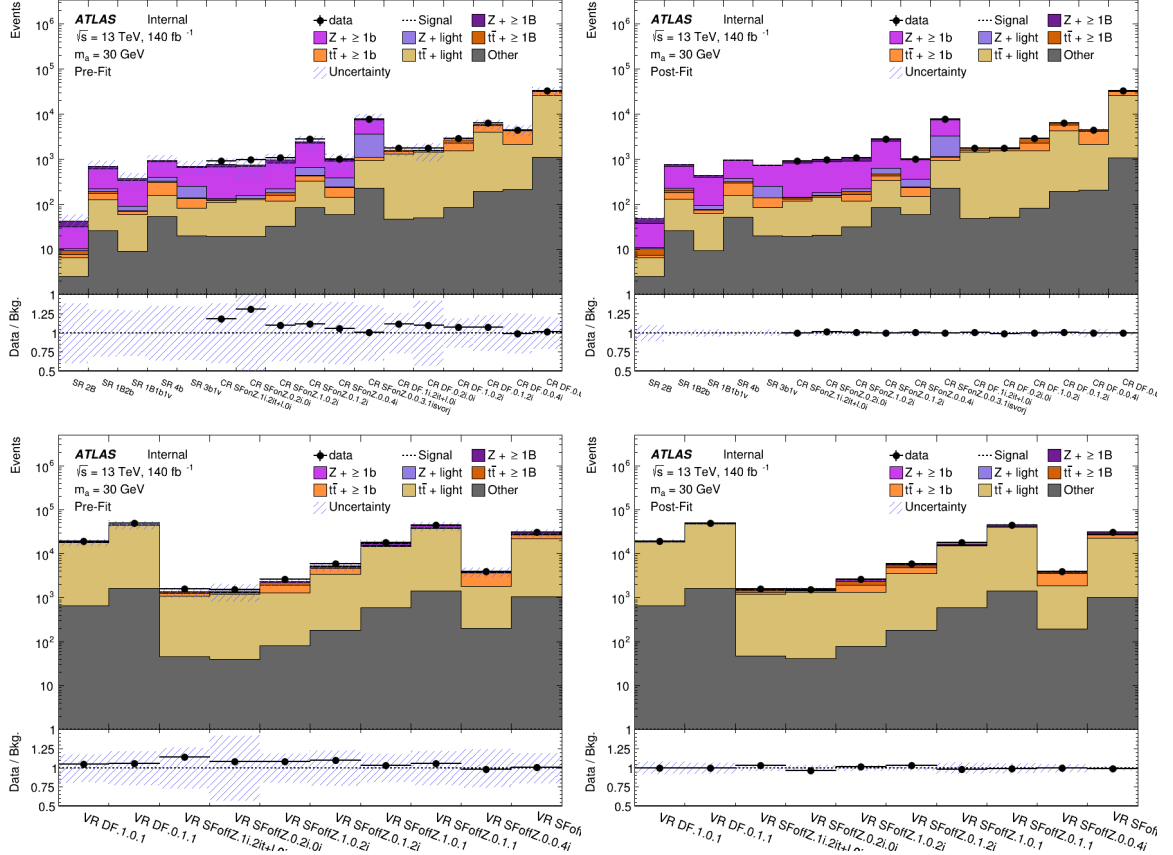


Figure B.34: Here the $m_a = 30$ GeV mass hypothesis is considered. Predicted yields for the signal, control (top), and validation (bottom) regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 30$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow aa \rightarrow 4b) = 0.1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

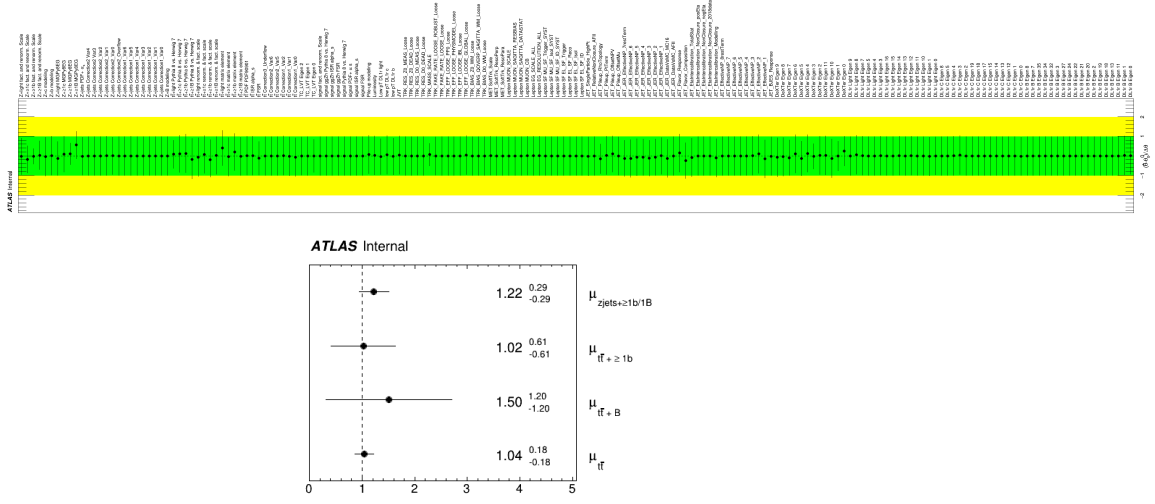


Figure B.35: Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics for the $m_a = 30$ GeV mass hypothesis.

B.2.5 $m_a = 50$ GeV

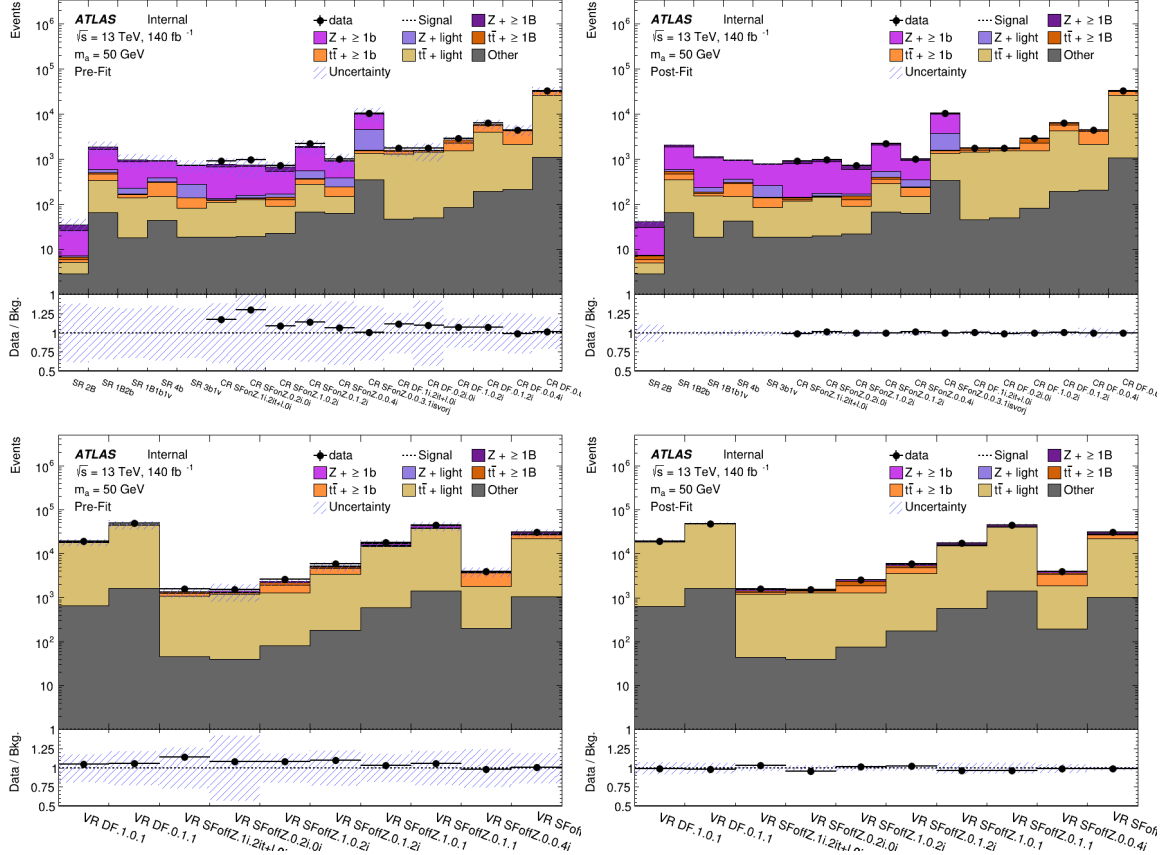


Figure B.36: Here the $m_a = 50$ GeV mass hypothesis is considered. Predicted yields for the signal, control (top), and validation (bottom) regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 50$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow aa \rightarrow 4b) = 0.1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

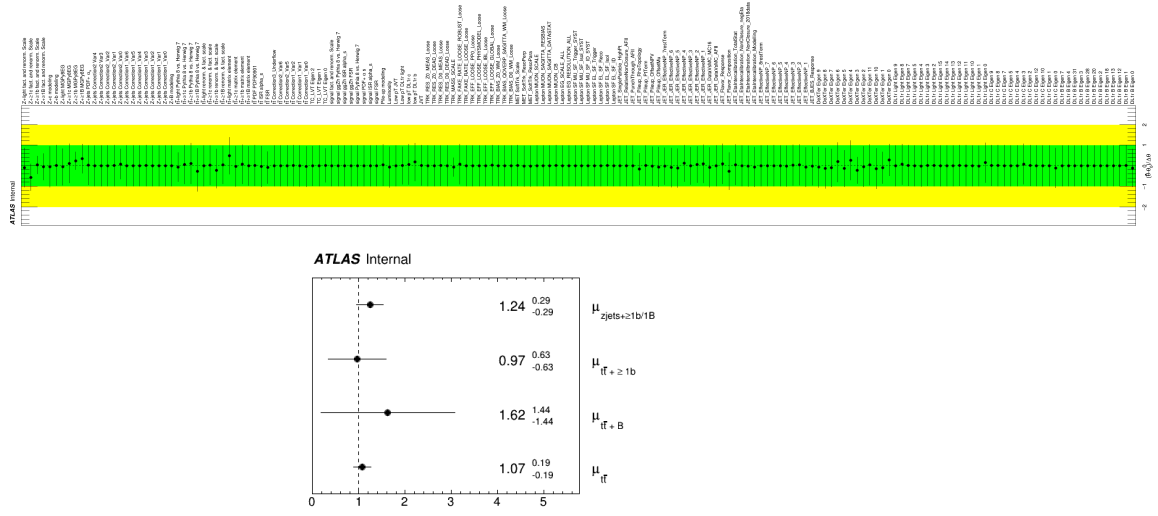


Figure B.37: Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics for the $m_a = 50$ GeV mass hypothesis.

B.2.6 $m_a = 60$ GeV

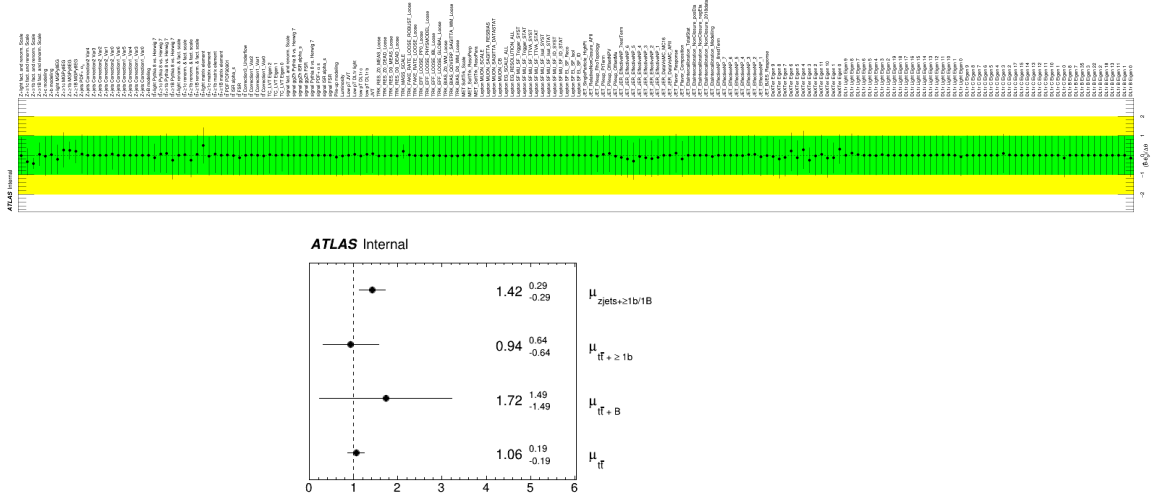


Figure B.39: Nuisance parameters (top) and normalization factors (bottom right) for the blinded data fit using the full set of systematics for the $m_a = 60$ GeV mass hypothesis.

B.2.7 $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV

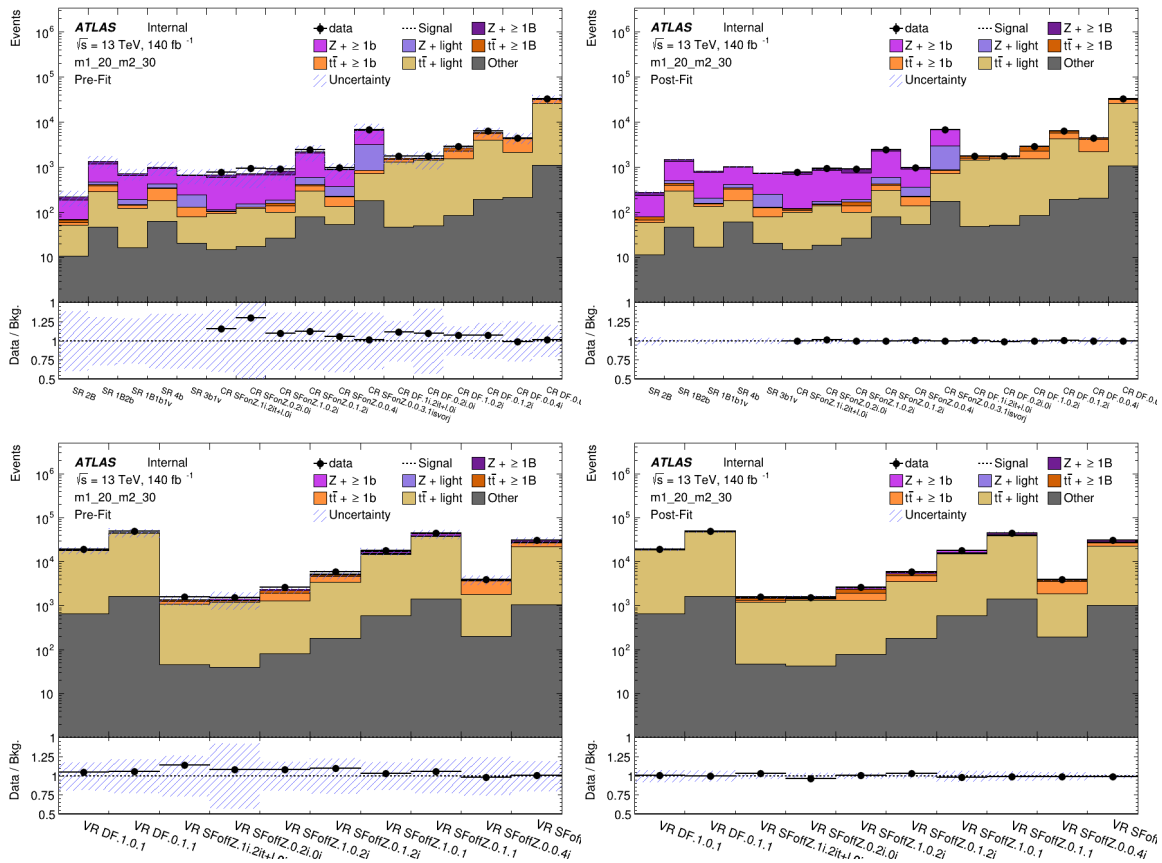


Figure B.40: Here the $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV mass hypothesis is considered. Predicted yields for the signal, control (top), and validation (bottom) regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

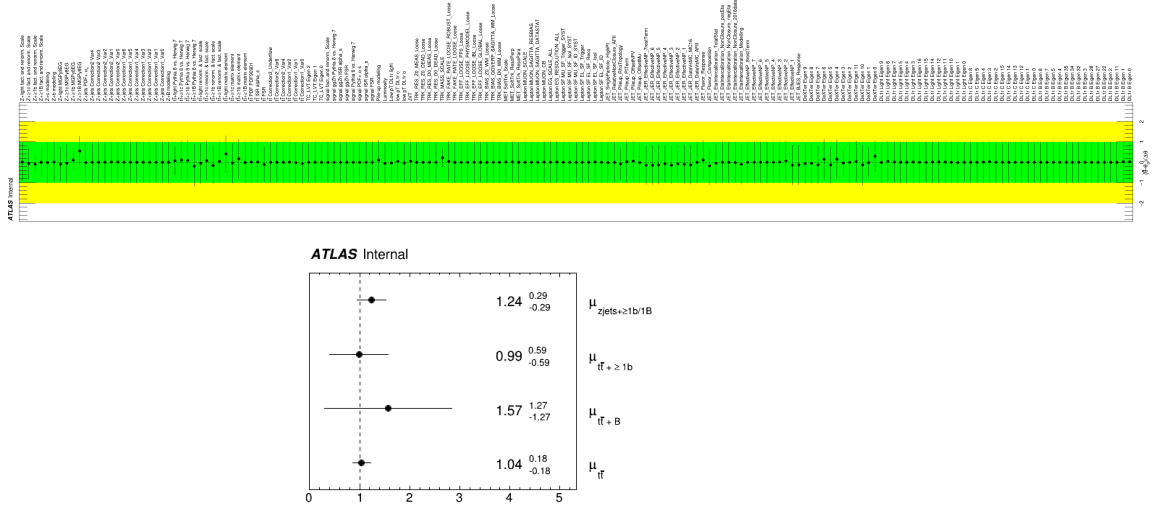


Figure B.41: Here the $m_{a_1} = 20$ GeV, $m_{a_2} = 30$ GeV mass hypothesis is considered. Nuisance parameters (top) and normalization factors (bottom right) for the blinded data fit using the full set of systematics.

B.2.8 $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeV

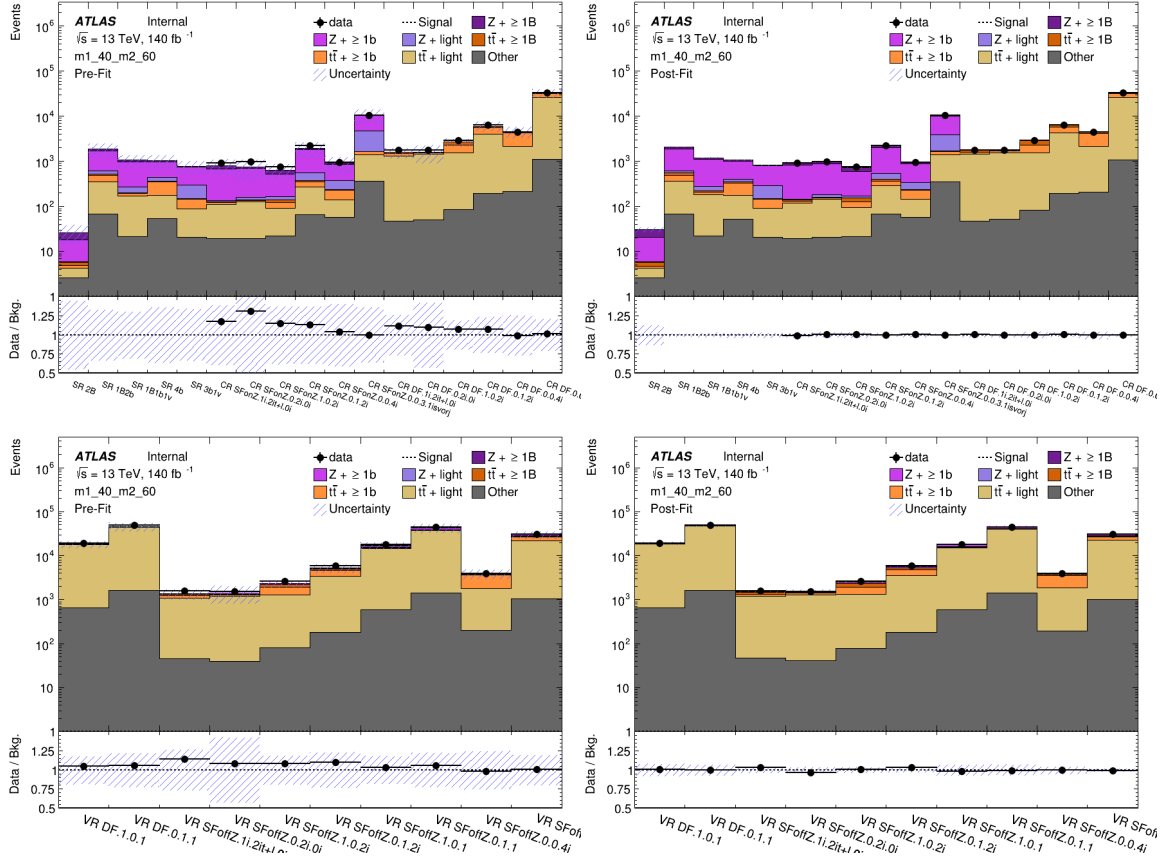


Figure B.42: Here the $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeV mass hypothesis is considered. Predicted yields for the signal, control (top), and validation (bottom) regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

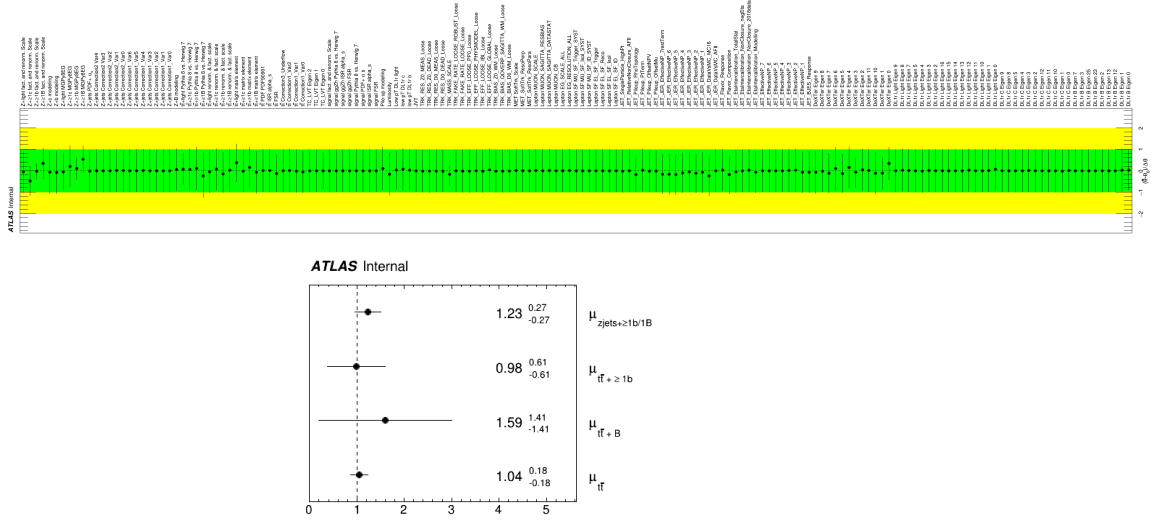


Figure B.43: Here the $m_{a_1} = 40$ GeV, $m_{a_2} = 60$ GeV mass hypothesis is considered. Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics.

B.2.9 $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV

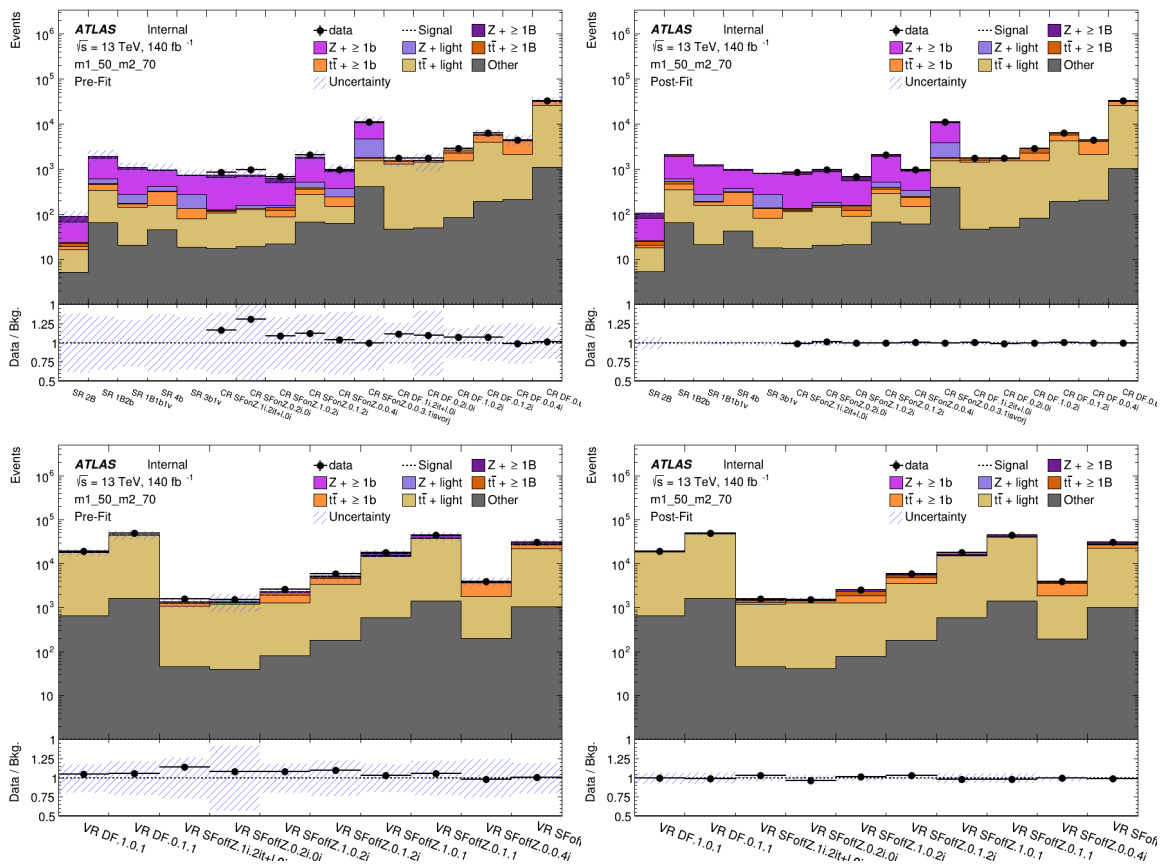


Figure B.44: Here the $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV mass hypothesis is considered. Predicted yields for the signal, control (top), and validation (bottom) regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\text{BR}(H \rightarrow a_1 a_2 \rightarrow 4b) = 0.1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

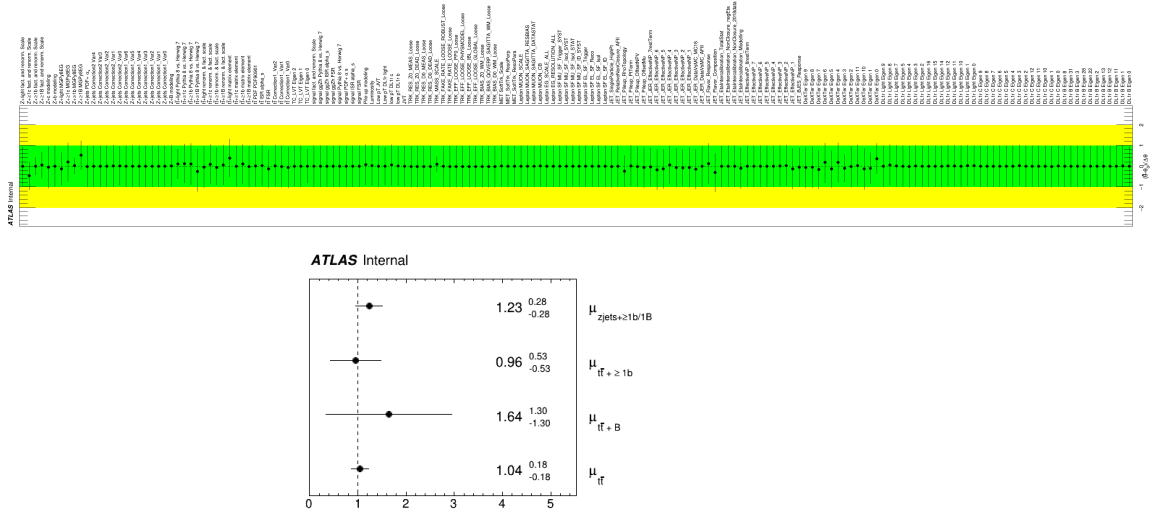


Figure B.45: Here the $m_{a_1} = 50$ GeV, $m_{a_2} = 70$ GeV mass hypothesis is considered. Nuisance parameters (top) and normalization factors (bottom) for the blinded data fit using the full set of systematics.

BIBLIOGRAPHY

- [1] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].
- [2] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].
- [3] ATLAS Collaboration. “A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery”. In: *Nature* 607 (2022), pp. 52–59. DOI: 10.1038/s41586-022-04893-w. arXiv: 2207.00092 [hep-ex].
- [4] David Curtin et al. “Exotic decays of the 125 GeV Higgs boson”. In: *Phys. Rev. D* 90.7 (2014), p. 075004. DOI: 10.1103/PhysRevD.90.075004. arXiv: arXiv:1312.4992[hep-ph] [hep-ph].
- [5] David J. Gross and Frank Wilczek. “Ultraviolet Behavior of Non-Abelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [6] H. David Politzer. “Reliable Perturbative Results for Strong Interactions?” In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [7] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 Nov. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [8] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.

- [9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.585>.
- [10] *LHC Higgs Cross Section Working Group*. Accessed: 2023-07-09. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>.
- [11] J. R. Brownstein and J. W. Moffat. “The Bullet Cluster 1E0657-558 evidence shows modified gravity in the absence of dark matter”. In: *Monthly Notices of the Royal Astronomical Society* 382.1 (Oct. 2007), pp. 29–47. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.12275.x. eprint: <https://academic.oup.com/mnras/article-pdf/382/1/29/3057374/mnras0382-0029.pdf>. URL: <https://doi.org/10.1111/j.1365-2966.2007.12275.x>.
- [12] F. Zwicky. “On the Masses of Nebulae and of Clusters of Nebulae”. In: *Astrophysical Journal* 86 (Nov. 1937), p. 217. DOI: 10.1086/143864.
- [13] N. Aghanim et al. “Planck 2018 results. VI. Cosmological parameters”. In: *Astron. Astrophys.* 641 (2020). [Erratum: *Astron. Astrophys.* 652, C4 (2021)], A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].
- [14] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. “Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter”. In: *Astrophys. J.* 604 (2004), pp. 596–603. DOI: 10.1086/381970. arXiv: astro-ph/0312273.
- [15] P. A. R. Ade et al. “Planck 2015 results. XIII. Cosmological parameters”. In: *Astron. Astrophys.* 594 (2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].
- [16] E. Komatsu et al. “Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation”. In: *Astrophys. J. Suppl.* 180 (2009), pp. 330–376. DOI: 10.1088/0067-0049/180/2/330. arXiv: 0803.0547 [astro-ph].
- [17] Stefano Profumo, Michael J Ramsey-Musolf, and Gabe Shaughnessy. “Singlet Higgs phenomenology and the electroweak phase transition”. In: *Journal of High Energy Physics* 2007.08 (Aug. 2007), pp. 010–010. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2007/08/010. URL: <http://dx.doi.org/10.1088/1126-6708/2007/08/010>.
- [18] Nikita Blinov et al. “Electroweak Baryogenesis from Exotic Electroweak Symmetry Breaking”. In: (2015). arXiv: arXiv:1504.05195 [hep-ph] [hep-ph].

- [19] Maxim Pospelov, Adam Ritz, and Mikhail Voloshin. “Secluded WIMP dark matter”. In: *Physics Letters B* 662.1 (Apr. 2008), pp. 53–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2008.02.052. URL: <http://dx.doi.org/10.1016/j.physletb.2008.02.052>.
- [20] Patrick Draper et al. “Dark Light-Higgs Bosons”. In: *Physical Review Letters* 106.12 (Mar. 2011). ISSN: 1079-7114. DOI: 10.1103/PhysRevLett.106.121805. URL: <http://dx.doi.org/10.1103/PhysRevLett.106.121805>.
- [21] Bogdan A. Dobrescu and Konstantin T. Matchev. “Light axion within the next-to-minimal supersymmetric standard model”. In: *JHEP* 09 (2000), p. 031. DOI: 10.1088/1126-6708/2000/09/031. arXiv: hep-ph/0008192 [hep-ph].
- [22] Ulrich Ellwanger et al. “Towards a no lose theorem for NMSSM Higgs discovery at the LHC”. In: (2003). arXiv: hep-ph/0305109 [hep-ph].
- [23] Radovan Dermisek and John F. Gunion. “Escaping the large fine tuning and little hierarchy problems in the next to minimal supersymmetric model and $h \rightarrow aa$ decays”. In: *Phys. Rev. Lett.* 95 (2005), p. 041801. DOI: 10.1103/PhysRevLett.95.041801. arXiv: hep-ph/0502105 [hep-ph].
- [24] Spencer Chang et al. “Nonstandard Higgs Boson Decays”. In: *Ann. Rev. Nucl. Part. Sci.* 58 (2008), pp. 75–98. DOI: 10.1146/annurev.nucl.58.110707.171200. arXiv: 0801.4554 [hep-ph].
- [25] David E. Morrissey and Aaron Pierce. “Modified Higgs Boson Phenomenology from Gauge or Gaugino Mediation in the NMSSM”. In: *Phys. Rev. D* 78 (2008), p. 075029. DOI: 10.1103/PhysRevD.78.075029. arXiv: 0807.2259 [hep-ph].
- [26] Vanda Silveira and A. Zee. “SCALAR PHANTOMS”. In: *Phys. Lett.* 161B (1985), pp. 136–140. DOI: 10.1016/0370-2693(85)90624-0.
- [27] Maxim Pospelov, Adam Ritz, and Mikhail B. Voloshin. “Secluded WIMP Dark Matter”. In: *Phys. Lett.* **B662** (2008), pp. 53–61. DOI: 10.1016/j.physletb.2008.02.052. arXiv: arXiv:0711.4866 [hep-ph] [hep-ph].
- [28] Patrick Draper et al. “Dark Light Higgs”. In: *Phys. Rev. Lett.* **106** (2011), p. 121805. DOI: 10.1103/PhysRevLett.106.121805. arXiv: arXiv:1009.3963 [hep-ph] [hep-ph].
- [29] Seyda Ipek, David McKeen, and Ann E. Nelson. “A Renormalizable Model for the Galactic Center Gamma Ray Excess from Dark Matter Annihilation”. In: *Phys. Rev. D* **90.5** (2014), p. 055021. DOI: 10.1103/PhysRevD.90.055021. arXiv: arXiv:1404.3716 [hep-ph] [hep-ph].

- [30] Adam Martin, Jessie Shelton, and James Unwin. “Fitting the Galactic Center Gamma-Ray Excess with Cascade Annihilations”. In: *Phys. Rev.* **D90**.10 (2014), p. 103513. DOI: 10.1103/PhysRevD.90.103513. arXiv: arXiv:1405.0272[hep-ph] [hep-ph].
- [31] Stefano Profumo, Michael J. Ramsey-Musolf, and Gabe Shaughnessy. “Singlet Higgs phenomenology and the electroweak phase transition”. In: *JHEP* **08** (2007), p. 010. DOI: 10.1088/1126-6708/2007/08/010. arXiv: arXiv:0705.2425[hep-ph] [hep-ph].
- [32] Gustavo Burdman et al. “Folded supersymmetry and the LEP paradox”. In: *JHEP* **02** (2007), p. 009. DOI: 10.1088/1126-6708/2007/02/009. arXiv: hep-ph/0609152 [hep-ph].
- [33] Nathaniel Craig et al. “Naturalness in the Dark at the LHC”. In: *JHEP* **07** (2015), p. 105. DOI: 10.1007/JHEP07(2015)105. arXiv: arXiv:1501.05310[hep-ph] [hep-ph].
- [34] David Curtin and Christopher B. Verhaaren. “Discovering Uncolored Naturalness in Exotic Higgs Decays”. In: (2015). arXiv: arXiv:1506.06141[hep-ph] [hep-ph].
- [35] T. D. Lee. “A Theory of Spontaneous T Violation”. In: *Phys. Rev. D* **8** (4 Aug. 1973), pp. 1226–1239. DOI: 10.1103/PhysRevD.8.1226. URL: <https://link.aps.org/doi/10.1103/PhysRevD.8.1226>.
- [36] Sheldon L. Glashow and Steven Weinberg. “Natural conservation laws for neutral currents”. In: *Phys. Rev. D* **15** (7 Apr. 1977), pp. 1958–1965. DOI: 10.1103/PhysRevD.15.1958. URL: <https://link.aps.org/doi/10.1103/PhysRevD.15.1958>.
- [37] Marcela Carena et al. “Higgs Signal for $H \rightarrow aa$ at Hadron Colliders”. In: *JHEP* **04** (2008), p. 092. DOI: 10.1088/1126-6708/2008/04/092. arXiv: arXiv:0712.2466[hep-ph] [hep-ph].
- [38] Kingman Cheung, Jeonghyeon Song, and Qi-Shu Yan. “Role of $H \rightarrow \eta\eta$ in Intermediate-Mass Higgs Boson Searches at the Large Hadron Collider”. In: *Phys. Rev. Lett.* **99** (2007), p. 031801. DOI: 10.1103/PhysRevLett.99.031801. arXiv: arXiv:0703149[hep-ph] [hep-ph].
- [39] David E. Kaplan and Matthew McEvoy. “Associated Production of Non-Standard Higgs Bosons at the LHC”. In: *Phys. Rev.* **D83** (2011), p. 115004. DOI: 10.1103/PhysRevD.83.115004. arXiv: arXiv:1102.0704[hep-ph] [hep-ph].

- [40] ATLAS Collaboration. “Search for the Higgs boson produced in association with a vector boson and decaying into two spin-zero particles in the $H \rightarrow aa \rightarrow 4b$ channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 10 (2018), p. 031. DOI: 10.1007/JHEP10(2018)031. arXiv: 1806.07355 [hep-ex].
- [41] ATLAS Collaboration. “Search for Higgs boson decays into two new low-mass spin-0 particles in the $4b$ channel with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Rev. D* 102 (2020), p. 112006. DOI: 10.1103/PhysRevD.102.112006. arXiv: 2005.12236 [hep-ex].
- [42] “Search for the exotic decay of the Higgs boson into two light pseudoscalars with four photons in the final state in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: (Aug. 2022). arXiv: 2208.01469 [hep-ex].
- [43] CMS Collaboration. “Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two muons and two b quarks in pp collisions at 13 TeV”. In: *Phys. Lett. B* 795 (2019), p. 398. DOI: 10.1016/j.physletb.2019.06.021. arXiv: 1812.06359 [hep-ex].
- [44] CMS Collaboration. “Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two b quarks and two τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Lett. B* 785 (2018), p. 462. DOI: 10.1016/j.physletb.2018.08.057. arXiv: 1805.10191 [hep-ex].
- [45] ATLAS Collaboration. “Search for new phenomena in events with at least three photons collected in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 76 (2016), p. 210. DOI: 10.1140/epjc/s10052-016-4034-8. arXiv: 1509.05051 [hep-ex].
- [46] ATLAS Collaboration. “Search for Higgs boson decays into pairs of light (pseudo)scalar particles in the $\gamma\gamma jj$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 782 (2018), p. 750. DOI: 10.1016/j.physletb.2018.06.011. arXiv: 1803.11145 [hep-ex].
- [47] V. M. Abazov et al. “Search for NMSSM Higgs bosons in the $h \rightarrow aa \rightarrow \mu\mu\mu\mu, \mu\mu\tau\tau$ channels using p anti-p collisions at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Rev. Lett.* 103 (2009), p. 061801. DOI: 10.1103/PhysRevLett.103.061801. arXiv: 0905.3381 [hep-ex].
- [48] Albert M Sirunyan et al. “Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two muons and two b quarks in pp collisions at 13 TeV”. In: *Phys. Lett. B* 795 (2019), pp. 398–423. DOI: 10.1016/j.physletb.2019.06.021. arXiv: 1812.06359 [hep-ex].

- [49] CMS Collaboration. “Search for light pseudoscalar boson pairs produced from decays of the 125 GeV Higgs boson in final states with two muons and two nearby tracks in pp collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Lett. B* 800 (2020), p. 135087. DOI: 10.1016/j.physletb.2019.135087. arXiv: 1907.07235 [hep-ex].
- [50] CMS Collaboration. “A search for pair production of new light bosons decaying into muons in proton–proton collisions at 13 TeV”. In: *Phys. Lett. B* 796 (2019), p. 131. DOI: 10.1016/j.physletb.2019.07.013. arXiv: 1812.00380 [hep-ex].
- [51] CMS Collaboration. “Search for low-mass dilepton resonances in Higgs boson decays to four-lepton final states in proton–proton collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 82 (2021), p. 290. DOI: 10.1140/epjc/s10052-022-10127-0. arXiv: 2111.01299 [hep-ex].
- [52] ATLAS Collaboration. “Search for Higgs bosons decaying into new spin-0 or spin-1 particles in four-lepton final states with the ATLAS detector with 139 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV”. In: *JHEP* 03 (2021), p. 041. DOI: 10.1007/JHEP03(2022)041. arXiv: 2110.13673 [hep-ex].
- [53] ATLAS Collaboration. “Search for Higgs bosons decaying to aa in the $\mu\mu\tau\tau$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS experiment”. In: *Phys. Rev. D* 92 (2015), p. 052002. DOI: 10.1103/PhysRevD.92.052002. arXiv: 1505.01609 [hep-ex].
- [54] CMS Collaboration. “Search for a light pseudoscalar Higgs boson in the boosted $\mu\mu\tau\tau$ final state in proton–proton collisions at $\sqrt{s} = 13$ TeV”. In: *JHEP* 08 (2020), p. 139. DOI: 10.1007/JHEP08(2020)139. arXiv: 2005.08694 [hep-ex].
- [55] CMS Collaboration. “Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state of two muons and two τ leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV”. In: *JHEP* 11 (2018), p. 018. DOI: 10.1007/JHEP11(2018)018. arXiv: 1805.04865 [hep-ex].
- [56] ATLAS Collaboration. “Search for Higgs boson decays into a pair of pseudoscalar particles in the $b\bar{b}\mu\mu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Rev. D* 105 (2021), p. 012006. DOI: 10.1103/PhysRevD.105.012006. arXiv: 2110.00313 [hep-ex].
- [57] Marié a Cepeda et al. “Exotic Higgs Decays”. In: *Annual Review of Nuclear and Particle Science* 72.1 (Nov. 2022), pp. 119–149. DOI: 10.1146/annurev-nucl-102319-024147. URL: <https://doi.org/10.1146%5C%2Fannurev-nucl-102319-024147>.

- [58] Tania Robens, Tim Stefaniak, and Jonas Wittbrodt. “Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios”. In: *The European Physical Journal C* 80.2 (Feb. 2020), p. 151. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-7655-x. URL: <https://doi.org/10.1140/epjc/s10052-020-7655-x>.
- [59] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.
- [60] CERN Service graphique. “Overall view of the LHC. Vue d’ensemble du LHC”. In: (2014). General Photo. URL: <https://cds.cern.ch/record/1708849>.
- [61] Esma Mobs. “The CERN accelerator complex in 2019. Complexe des accélérateurs du CERN en 2019”. In: (2019). General Photo. URL: <https://cds.cern.ch/record/2684277>.
- [62] *Public ATLAS Luminosity Results for Run-2 of the LHC*. Accessed: 2023-06-30. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Publications_and_Conference_Resu.
- [63] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.
- [64] Joao Pequeno. “Computer generated image of the whole ATLAS detector”. 2008. URL: <https://cds.cern.ch/record/1095924>.
- [65] Joao Pequeno. “Computer generated image of the ATLAS inner detector”. 2008. URL: <https://cds.cern.ch/record/1095926>.
- [66] B. Abbott et al. “Production and integration of the ATLAS Insertable B-Layer”. In: *JINST* 13 (2018), T05008. DOI: 10.1088/1748-0221/13/05/T05008. arXiv: 1803.00844 [physics.ins-det].
- [67] ATLAS Collaboration. “Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment”. General Photo. 2020. URL: <https://cds.cern.ch/record/2723878>.
- [68] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, Volume 1*. ATLAS-TDR-4; CERN-LHCC-97-016. 1997. URL: <https://cds.cern.ch/record/331063>.
- [69] Joao Pequeno. “Computer Generated image of the ATLAS calorimeter”. 2008. URL: <https://cds.cern.ch/record/1095927>.

- [70] Joao Pequeno. “Computer generated image of the ATLAS Muons subsystem”. 2008. URL: <https://cds.cern.ch/record/1095929>.
- [71] ATLAS Collaboration. “Performance of the ATLAS RPC detector and Level-1 muon barrel trigger at $\sqrt{s} = 13$ TeV”. In: *JINST* 16 (2021), P07029. DOI: 10.1088/1748-0221/16/07/P07029. arXiv: 2103.01029 [hep-ex].
- [72] ATLAS Collaboration. “Resolution of the ATLAS muon spectrometer monitored drift tubes in LHC Run 2”. In: *JINST* 14 (2019), P09011. DOI: 10.1088/1748-0221/14/09/P09011. arXiv: 1906.12226 [hep-ex].
- [73] *ATLAS magnet system: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. DOI: 10.17181/CERN.905C.VDTM. URL: <https://cds.cern.ch/record/338080>.
- [74] A. Yamamoto et al. “Progress in ATLAS central solenoid magnet”. In: *IEEE Transactions on Applied Superconductivity* 10.1 (2000), pp. 353–356. DOI: 10.1109/77.828246.
- [75] Daniel Dominguez, Kevin Patrice Moles, and Sascha Mehlhase. “ATLAS detector schematics”. In: (2021). URL: <https://cds.cern.ch/record/2777214>.
- [76] ATLAS Collaboration. “Operation of the ATLAS trigger system in Run 2”. In: *JINST* 15 (2020), P10004. DOI: 10.1088/1748-0221/15/10/P10004. arXiv: 2007.12539 [hep-ex].
- [77] O. Aberle et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2020. DOI: 10.23731/CYRM-2020-0010. URL: <https://cds.cern.ch/record/2749422>.
- [78] ATLAS Collaboration. *ATLAS Muon Spectrometer Phase-II Upgrade: Technical Design Report*. ATLAS-TDR-026; CERN-LHCC-2017-017. 2017. URL: <https://cds.cern.ch/record/2285580>.
- [79] ATLAS Collaboration. *ATLAS TDAQ Phase-II Upgrade: Technical Design Report*. ATLAS-TDR-029; CERN-LHCC-2017-020. 2017. URL: <https://cds.cern.ch/record/2285584>.
- [80] ATLAS Collaboration. *ATLAS Inner Tracker Pixel Detector: Technical Design Report*. ATLAS-TDR-030; CERN-LHCC-2017-021. 2017. URL: <https://cds.cern.ch/record/2285585>.
- [81] ATLAS Collaboration. *A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade: Technical Design Report*. ATLAS-TDR-031; CERN-LHCC-2020-007. 2020. URL: <https://cds.cern.ch/record/2719855>.

- [82] Joao Pequeno and Paul Schaffner. “How ATLAS detects particles: diagram of particle paths in the detector”. 2013. URL: <https://cds.cern.ch/record/1505342>.
- [83] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data”. In: *JINST* 14 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].
- [84] ATLAS Collaboration. “Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 81 (2021), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex].
- [85] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].
- [86] ATLAS Collaboration. “Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 81 (2020), p. 689. DOI: 10.1140/epjc/s10052-021-09402-3. arXiv: 2007.02645 [hep-ex].
- [87] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti- k_t jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].
- [88] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet user manual”. In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].
- [89] ATLAS Collaboration. *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*. ATLAS-CONF-2015-037. 2015. URL: <https://cds.cern.ch/record/2044941>.
- [90] ATLAS Collaboration. “Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector”. In: *Eur. Phys. J. C* 77 (2017), p. 580. DOI: 10.1140/epjc/s10052-017-5081-5. arXiv: 1705.02211 [hep-ex]. Erratum: in: *Eur. Phys. J. C* 77 (2017), p. 712. DOI: 10.1140/epjc/s10052-017-5245-3.
- [91] ATLAS Collaboration. “ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset”. In: (2022). arXiv: 2211.16345 [physics.data-an].

- [92] ATLAS Collaboration. “ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\text{ TeV}$ ”. In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].
- [93] ATLAS Collaboration. “Measurement of the c -jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13\text{ TeV}$ collected with the ATLAS detector”. In: *Eur. Phys. J. C* 82 (2021), p. 95. DOI: 10.1140/epjc/s10052-021-09843-w. arXiv: 2109.10627 [hep-ex].
- [94] “Calibration of the light-flavour jet mistagging efficiency of the b -tagging algorithms with Z +jets events using 139 fb^{-1} of ATLAS proton-proton collision data at $\sqrt{s} = 13\text{ TeV}$ ”. In: (Jan. 2023). arXiv: 2301.06319 [hep-ex].
- [95] ATLAS Collaboration. *Soft b -hadron tagging for compressed SUSY scenarios*. ATLAS-CONF-2019-027. 2019. URL: <https://cds.cern.ch/record/2682131>.
- [96] ATLAS Collaboration. *Secondary vertex finding for jet flavour identification with the ATLAS detector*. ATL-PHYS-PUB-2017-011. 2017. URL: <https://cds.cern.ch/record/2270366>.
- [97] ATLAS Collaboration. *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*. ATLAS-CONF-2018-023. 2018. URL: <https://cds.cern.ch/record/2625233>.
- [98] ATLAS Collaboration. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13\text{ TeV}$ ”. In: *Eur. Phys. J. C* 78 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].
- [99] ATLAS Collaboration. “Performance of missing transverse momentum reconstruction in proton–proton collisions at $\sqrt{s} = 7\text{ TeV}$ with ATLAS”. In: *Eur. Phys. J. C* 72 (2012), p. 1844. DOI: 10.1140/epjc/s10052-011-1844-6. arXiv: 1108.5602 [hep-ex].
- [100] ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13\text{ TeV}$ using the ATLAS detector at the LHC*. ATLAS-CONF-2019-021. 2019. URL: <https://cds.cern.ch/record/2677054>.
- [101] ATLAS Collaboration. “Performance of electron and photon triggers in ATLAS during LHC Run 2”. In: *Eur. Phys. J. C* 80 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

- [102] ATLAS Collaboration. “Performance of the ATLAS muon triggers in Run 2”. In: *JINST* 15 (2020), P09015. DOI: 10.1088/1748-0221/15/09/p09015. arXiv: 2004.13447 [hep-ex].
- [103] GEANT4 Collaboration, S. Agostinelli, et al. “GEANT4 – a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.
- [104] T. Sjöstrand, S. Mrenna, and P. Skands. “A brief introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].
- [105] Richard D. Ball et al. “Parton distributions with LHC data”. In: *Nucl. Phys. B* 867 (2013), p. 244. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].
- [106] 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>.
- [107] Richard D. Ball et al. “Parton distributions for the LHC run II”. In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].
- [108] ATLAS Collaboration. “Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”. In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].
- [109] ATLAS Collaboration. “Measurement of Higgs boson decay into b -quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 06 (2021), p. 097. DOI: 10.1007/JHEP06(2022)097. arXiv: 2111.06712 [hep-ex].
- [110] Tanju Gleisberg and Stefan Höche. “Comix, a new matrix element generator”. In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].
- [111] Federico Buccioni et al. “OpenLoops 2”. In: *Eur. Phys. J. C* 79.10 (2019), p. 866. DOI: 10.1140/epjc/s10052-019-7306-2. arXiv: 1907.13071 [hep-ph].
- [112] Rikkert Frederix and Stefano Frixione. “Merging meets matching in MC@NLO”. In: *JHEP* 12 (2012), p. 061. DOI: 10.1007/JHEP12(2012)061. arXiv: 1209.6215 [hep-ph].

- [113] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: <https://cds.cern.ch/record/1966419>.
- [114] Stefano Frixione, Giovanni Ridolfi, and Paolo Nason. “A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].
- [115] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. ATL-PHYS-PUB-2016-020. 2016. URL: <https://cds.cern.ch/record/2216168>.
- [116] Paolo Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.
- [117] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method”. In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].
- [118] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].
- [119] M. Bähr et al. “Herwig++ physics and manual”. In: *Eur. Phys. J. C* 58 (2008), p. 639. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].
- [120] Johannes Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].
- [121] Johannes Bellm et al. “Herwig 7.1 Release Note”. In: (2017). arXiv: 1705.06919 [hep-ph].
- [122] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph].
- [123] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152. DOI: 10.1016/S0168-9002(01)00089-4.

- [124] ATLAS Collaboration. *Studies of Monte Carlo predictions for the $t\bar{t}b\bar{b}$ process*. ATL-PHYS-PUB-2022-006. 2022. URL: <https://cds.cern.ch/record/2802806>.
- [125] Fabio Cascioli et al. “NLO matching for $t\bar{t}b\bar{b}$ production with massive b -quarks”. In: *Phys. Lett. B* 734 (2014), pp. 210–214. DOI: 10.1016/j.physletb.2014.05.040. arXiv: 1309.5912 [hep-ph].
- [126] Roman Kogler et al. “Jet Substructure at the Large Hadron Collider: Experimental Review”. In: *Rev. Mod. Phys.* 91.4 (2019), p. 045003. DOI: 10.1103/RevModPhys.91.045003. arXiv: 1803.06991 [hep-ex].
- [127] Georges Aad et al. “Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions”. In: (Sept. 2020). arXiv: 2009.04986 [hep-ex].
- [128] David Krohn, Jesse Thaler, and Lian-Tao Wang. “Jets with Variable R”. In: *JHEP* 06 (2009), p. 059. DOI: 10.1088/1126-6708/2009/06/059. arXiv: 0903.0392 [hep-ph].
- [129] ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run*. ATL-PHYS-PUB-2017-013. 2017. URL: <https://cds.cern.ch/record/2273281>.
- [130] ATLAS Collaboration. “In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector”. In: *Eur. Phys. J. C* 79 (2019), p. 135. DOI: 10.1140/epjc/s10052-019-6632-8. arXiv: 1807.09477 [hep-ex].
- [131] Mirkoantonio Casolino et al. “Probing a light CP-odd scalar in di-top-associated production at the LHC”. In: *Eur. Phys. J. C* 75 (2015), p. 498. DOI: 10.1140/epjc/s10052-015-3708-y. arXiv: 1507.07004 [hep-ph].
- [132] Matteo Cacciari and Gavin P. Salam. “Pileup subtraction using jet areas”. In: *Phys. Lett. B* 659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].
- [133] Benjamin Nachman et al. “Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC”. In: *JHEP* 02 (2015), p. 075. DOI: 10.1007/JHEP02(2015)075. arXiv: 1407.2922 [hep-ph].
- [134] ATLAS Collaboration. *Topological b -hadron decay reconstruction and identification of b -jets with the JetFitter package in the ATLAS experiment at the LHC*. ATL-PHYS-PUB-2018-025. 2018. URL: <https://cds.cern.ch/record/2645405>.

- [135] V Kostyukhin. *Secondary vertex based b-tagging*. Tech. rep. ATL-PHYS-2003-033, revised version number 1 submitted on 2003-09-22 11:21:01. Geneva: CERN, Aug. 2003. URL: <https://cds.cern.ch/record/685550>.
- [136] Jesse Thaler and Ken Van Tilburg. “Identifying Boosted Objects with N-subjettiness”. In: *JHEP* 03 (2011), p. 015. DOI: 10.1007/JHEP03(2011)015. arXiv: 1011.2268 [hep-ph].
- [137] Stephen D. Ellis and Davison E. Soper. “Successive combination jet algorithm for hadron collisions”. In: *Phys. Rev. D* 48 (1993), pp. 3160–3166. DOI: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266.
- [138] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The catchment area of jets”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 005–005. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/005. URL: <http://dx.doi.org/10.1088/1126-6708/2008/04/005>.
- [139] ATLAS Collaboration. *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*. ATL-PHYS-PUB-2020-014. 2020. URL: <https://cds.cern.ch/record/2718948>.
- [140] Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler. “Energy Flow Networks: Deep Sets for Particle Jets”. In: *JHEP* 01 (2019), p. 121. DOI: 10.1007/JHEP01(2019)121. arXiv: 1810.05165 [hep-ph].
- [141] ATLAS Collaboration. *Digluon Tagging using $\sqrt{s} = 13$ TeV pp Collisions in the ATLAS Detector*. ATL-PHYS-PUB-2021-027. 2021. URL: <https://cds.cern.ch/record/2776780>.
- [142] John S Bridle. “Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition”. In: (1990). Ed. by Françoise Fogelman Soulié and Jeanny Hérault.
- [143] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. *Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)*. 2015. DOI: 10.48550/ARXIV.1511.07289. URL: <https://arxiv.org/abs/1511.07289>.
- [144] Georges Aad et al. “Search for Higgs boson decays into two new low-mass spin-0 particles in the $4b$ channel with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV”. In: (May 2020). DOI: 10.3204/PUBDB-2020-02514. arXiv: 2005.12236 [hep-ex].
- [145] ATLAS Collaboration. “Measurement of colour flow using jet-pull observables in $t\bar{t}$ events with the ATLAS experiment at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys.*

- J. C* 78 (2018), p. 847. DOI: 10.1140/epjc/s10052-018-6290-2. arXiv: 1805.02935 [hep-ex].
- [146] Yaroslav Ganin et al. “Domain-Adversarial Training of Neural Networks”. In: *JMLR* 17.59 (2016). arXiv: 1505.07818 [stat.ML].
 - [147] François Chollet et al. *Keras*. 2015. URL: <https://keras.io>.
 - [148] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: <https://www.tensorflow.org/>.
 - [149] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2014. DOI: 10.48550/ARXIV.1412.6980. URL: <https://arxiv.org/abs/1412.6980>.
 - [150] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. *Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps*. 2013. arXiv: 1312.6034 [cs.CV].
 - [151] ATLAS Collaboration. “Identification of boosted Higgs bosons decaying into b -quark pairs with the ATLAS detector at 13 TeV”. In: *Eur. Phys. J. C* 79 (2019), p. 836. DOI: 10.1140/epjc/s10052-019-7335-x. arXiv: 1906.11005 [hep-ex].
 - [152] ATLAS Collaboration. *Efficiency corrections for a tagger for boosted $H \rightarrow b\bar{b}$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. ATL-PHYS-PUB-2021-035. 2021. URL: <https://cds.cern.ch/record/2777811>.
 - [153] ATLAS Collaboration. *Monte Carlo to Monte Carlo scale factors for flavour tagging efficiency calibration*. ATL-PHYS-PUB-2020-009. 2020. URL: <https://cds.cern.ch/record/2718610>.
 - [154] ATLAS Collaboration. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [hep-ex].
 - [155] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 76 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].
 - [156] ATLAS Collaboration. “Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector”. In: *Eur. Phys.*

- J. C* 76 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].
- [157] ATLAS Collaboration. “Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Rev. D* 96 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].
 - [158] ATLAS Collaboration. “Determination of jet calibration and energy resolution in proton–proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector”. In: *Eur. Phys. J. C* 80 (2020), p. 1104. DOI: 10.1140/epjc/s10052-020-08477-8. arXiv: 1910.04482 [hep-ex].
 - [159] G. Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.
 - [160] Jon Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].
 - [161] F. Cascioli et al. “Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production”. In: *JHEP* 01 (2014), p. 046. DOI: 10.1007/JHEP01(2014)046. arXiv: 1309.0500 [hep-ph].
 - [162] ATLAS Collaboration. *DeXTer: Deep Sets based Neural Networks for Low- p_T $X \rightarrow b\bar{b}$ Identification in ATLAS*. ATL-PHYS-PUB-2022-042. 2022. URL: <https://cds.cern.ch/record/2825434>.
 - [163] David Krohn, Jesse Thaler, and Lian-Tao Wang. “Jet Trimming”. In: *JHEP* 02 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].
 - [164] CMS Collaboration. “A deep neural network for simultaneous estimation of b quark energy and resolution”. In: *Comput. Softw. Big Sci.* 4 (2020), p. 10. DOI: 10.1007/s41781-020-00041-z. arXiv: 1912.06046 [hep-ex].
 - [165] Manzil Zaheer et al. *Deep Sets*. 2017. DOI: 10.48550/ARXIV.1703.06114. URL: <https://arxiv.org/abs/1703.06114>.
 - [166] John C. Collins and Davison E. Soper. “Angular distribution of dileptons in high-energy hadron collisions”. In: *Phys. Rev. D* 16 (7 Nov. 1977), pp. 2219–2225. DOI: 10.1103/PhysRevD.16.2219. URL: <https://link.aps.org/doi/10.1103/PhysRevD.16.2219>.

- [167] Tianqi Chen and Carlos Guestrin. “XGBoost: A Scalable Tree Boosting System”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD ’16. San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794. ISBN: 9781450342322. DOI: 10.1145/2939672.2939785. URL: <https://doi.org/10.1145/2939672.2939785>.
- [168] Richard Liaw et al. “Tune: A Research Platform for Distributed Model Selection and Training”. In: *arXiv preprint arXiv:1807.05118* (2018).
- [169] Lisha Li et al. “Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization”. In: *arXiv e-prints*, arXiv:1603.06560 (Mar. 2016), arXiv:1603.06560. DOI: 10.48550/arXiv.1603.06560. arXiv: 1603.06560 [cs.LG].
- [170] Kyle Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., Jan. 2012. URL: <https://cds.cern.ch/record/1456844>.
- [171] ATLAS Collaboration. “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”. In: (2022). arXiv: 2212.09379 [hep-ex].
- [172] ATLAS Collaboration. “Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC”. In: *Phys. Rev. Lett.* 117 (2016), p. 182002. DOI: 10.1103/PhysRevLett.117.182002. arXiv: 1606.02625 [hep-ex].
- [173] *Efficiency Measurement of Low- p_T b -tagging Algorithms*. Tech. rep. Geneva, 2023. URL: <https://cds.cern.ch/record/2841283>.
- [174] ATLAS Collaboration. *Calibration of light-flavour b -jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*. ATLAS-CONF-2018-006. 2018. URL: <https://cds.cern.ch/record/2314418>.
- [175] *DeXTer: Deep Sets based Neural Networks for Low- p_T $X \rightarrow b\bar{b}$ Identification in ATLAS (Internal Note)*. Tech. rep. Geneva: CERN, 2022. URL: <https://cds.cern.ch/record/2741503>.
- [176] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].
- [177] Lorenzo Moneta et al. “The RooStats Project”. In: *PoS ACAT2010* (2010). Ed. by T. Speer et al., p. 057. DOI: 10.22323/1.093.0057. arXiv: 1009.1003 [physics.data-an].

- [178] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an]. Erratum: in: *Eur. Phys. J. C* 73 (2013), p. 2501. DOI: 10.1140/epjc/s10052-013-2501-z.
- [179] Ashish Vaswani et al. *Attention Is All You Need*. 2017. arXiv: 1706.03762 [cs.CL].
- [180] Savannah Thais et al. “Graph Neural Networks in Particle Physics: Implementations, Innovations, and Challenges”. In: *Snowmass 2021*. Mar. 2022. arXiv: 2203.12852 [hep-ex].
- [181] Chris Pollard and Philipp Windischhofer. “Transport away your problems: Calibrating stochastic simulations with optimal transport”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1027 (Mar. 2022), p. 166119. DOI: 10.1016/j.nima.2021.166119. URL: <https://doi.org/10.1016%2Fj.nima.2021.166119>.
- [182] Abien Fred Agarap. “Deep learning using rectified linear units (relu)”. In: *arXiv preprint arXiv:1803.08375* (2018).