

**INVESTIGATION OF HIGGS BOSON DECAYING TO DI-MUON, DARK MATTER  
PRODUCED IN ASSOCIATION WITH A HIGGS BOSON DECAYING TO  $b$ -QUARKS  
AND UNBINNED PROFILED UNFOLDING**

by

Chen-Hsun (Jay) Chan

A dissertation submitted in partial fulfillment of  
the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2023

Date of final oral examination: 05/08/2023

The dissertation is approved by the following members of the Final Oral Committee:

Sau Lan Wu, Professor, Physics

Kevin Black, Professor, Physics

Tulika Bose, Professor, Physics

Snezana Stanimirovic, Professor, Astronomy



## ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Professor Sau Lan Wu. Sau Lan's guidance and support have been instrumental in my success throughout my PhD. I am grateful to have joined her research group and had the opportunity to work on experiments such as the  $H \rightarrow \mu\mu$ , mono- $h(b\bar{b})$ , Higgs combination, 2HDM+ $a$  combination and ITk upgrade. Sau Lan's mentorship and leadership have enabled me to make important contributions to these projects and stand out in the analysis teams. In addition, she connected me to the Lawrence Berkeley National Lab, where I had an amazing experience towards the end of my PhD. Sau Lan is a great mentor who truly cares about students' lives and careers. She not only taught me how to become a great physicist but also inspired me to become a hard-working and self-motivated person.

I am grateful to the Department of Physics at the University of Wisconsin-Madison for providing me with the opportunity to pursue my graduate studies. I would also like to acknowledge my thesis defense committee members, including my advisor Sau Lan Wu, and Kevin Black, Tulika Bose, and Snezana Stanimirovic, for reading my thesis and providing valuable feedback. Their comments and suggestions helped me to improve the quality of my work and make it a more comprehensive and coherent piece of scientific writing.

The Wisconsin ATLAS group is like a big family. I am grateful to Chen Zhou for instructing me and giving me advice for all aspects of my research projects, including  $H \rightarrow \mu\mu$ , mono- $h(b\bar{b})$ , Higgs combination, 2HDM+ $a$  combination, and my ATLAS qualification task on ITk simulation. Chen also helped me with all my presentations, funding proposals and even job applications. He has been super helpful throughout my PhD. I cannot be more grateful for the everyday help and support from Sylvie Padlewski as well. I also enjoyed the active discussions and company from Alex Wang, Rui Zhang, Alkaid Cheng, Tuan Pham, Wasikul Islam and Amy Tee. Furthermore, the technical support from Shaojun Sun and Wen Guan has been incredible.

I would like to thank my ATLAS analysis collaborators, including Giacomo Artoni, Jan Kretzschmar, Fabio Cerutti, Hongtao Tang, Andrea Gabrielli, Gaetano Barone, Miha Zgubic, Yanlin Liu, Miha Muskinja, Alice Alfonsi, Brendon Bullard, and Hanna Borecka-Bielska in  $H \rightarrow \mu\mu$ , Spyridon Argyropoulos, Dan Guest, James Frost, Philipp Gadow, Andrea Matic, Anindya Ghosh, Eleni Skorda, Jon Burr, and Samuel Meehan in  $\text{mono-}h(b\bar{b})$ , Zirui Wang and Lailin Xu in the  $2\text{HDM}+a$  combination, Nicolas Morange, Kunlin Ran and Rahul Balasubramanian in the Higgs combination, and Ben Smart and Chen Zhou for supervising my ATLAS qualification task. It was a pleasure working with all these incredible people, and I learned a lot from each individual.

I also want to acknowledge my fellow PhD students Manuel Silva and Maria Veronica Prado for all the joyful activities in Wisconsin and Switzerland. While I was at CERN, I had lots of fun hanging out with Yaun-Tang Chou, Otto Lau, Yvonne Ng, and Bobby Tse.

The days I was based at Berkeley are among the happiest in my life. I would like to express my sincere appreciation to Maurice Garcia-Sciveres and Kevin Einsweiler for providing me the opportunity to work on the ITk pixel upgrade and to Elisabetta Pianori and Timon Heim for giving me day-to-day training and guidance. I am also extremely grateful to have had the opportunity to work with Ben Nachman on the Machine Learning projects. Ben taught me a lot about Machine Learning, and his innovative ideas continue to inspire and motivate me. I also appreciate the support and company from other lab members, including Aleksandra Dimitrievska, Marija Marjanovic, Daniel Joseph Antrim, Hongtao Yang, Xiangyang Ju, Taisai Kumakura, Haoran Zhao, Miha Muskinja, Karol Krizka, Elodie Resseguie, Juerg Beringer, Marko Stamenkovic, Rebecca Carney, Zhicai Zhang, Mariel Pettee, Shuo Han, Emily Anne Thompson, Elham E Khoda, Ian Dyckes, Maria Giovanna Foti, Carl Haber, Elliot Reynolds, Vinicius Mikuni, Alessandra Ciocio, and Angira Rastogi.

I would also like to thank my mom Li-Chin Huang and my partner Gabriel Alcaraz for their love and support. Their unwavering encouragement and belief in me have been crucial to my success. I am grateful for their presence in my life and their unwavering support throughout my PhD journey.

Finally, I want to express my heartfelt appreciation to all my friends who have been a part of my life in Wisconsin, Switzerland, and Berkeley. I am grateful for the joy and laughter that they brought into my life, and their unwavering support and encouragement have been invaluable to me.

# TABLE OF CONTENTS

	Page
<b>LIST OF TABLES</b> . . . . .	v
<b>LIST OF FIGURES</b> . . . . .	vii
<b>ABSTRACT</b> . . . . .	xii
<b>1 Introduction</b> . . . . .	1
<b>2 SM Higgs boson and the extended Higgs sector</b> . . . . .	4
2.1 SM Higgs boson production at the LHC . . . . .	4
2.2 SM Higgs boson decay at the LHC . . . . .	7
2.3 Extended Higgs sector — type II Two-Higgs-Doublet Model . . . . .	8
2.3.1 $Z'$ -2HDM . . . . .	10
2.3.2 2HDM+ $a$ . . . . .	11
<b>3 The Large Hadron Collider and the ATLAS detector</b> . . . . .	13
3.1 The Large Hadron Collider . . . . .	13
3.2 The ATLAS detector . . . . .	15
<b>4 Data and simulated event samples</b> . . . . .	18
4.1 Data samples . . . . .	18
4.2 Simulated event samples . . . . .	18
4.2.1 Simulated event samples for $H \rightarrow \mu\mu$ . . . . .	19
4.2.2 Simulated event samples for mono- $h(b\bar{b})$ . . . . .	21
<b>5 Physics object definitions</b> . . . . .	22
5.1 Primary vertex . . . . .	22
5.2 Jets . . . . .	22
5.3 Leptons . . . . .	23
5.4 Missing transverse momentum . . . . .	25
5.5 Overlap removal . . . . .	25
<b>6 Search for <math>H \rightarrow \mu\mu</math></b> . . . . .	26
6.1 Event selections and categorization . . . . .	26

	Page
6.1.1	General BDT training strategy . . . . . 27
6.1.2	General strategy of BDT boundary optimization . . . . . 29
6.1.3	$t\bar{t}H$ category . . . . . 30
6.1.4	$VH$ categories . . . . . 31
6.1.5	$ggF$ and VBF categories . . . . . 36
6.1.6	Summary of categorization . . . . . 38
6.2	Statistical analysis . . . . . 38
6.2.1	Signal modeling . . . . . 39
6.2.2	Background modeling . . . . . 41
6.2.3	Results . . . . . 43
<b>7</b>	<b>Search for mono-<math>h(b\bar{b})</math> . . . . . 57</b>
7.1	Event selections and categorization . . . . . 57
7.1.1	Common selections . . . . . 58
7.1.2	Signal regions . . . . . 58
7.1.3	Control regions . . . . . 61
7.2	Statistical analysis . . . . . 62
7.2.1	Signal modeling . . . . . 63
7.2.2	Background modeling . . . . . 63
7.2.3	Systematic uncertainties . . . . . 64
7.2.4	Results . . . . . 66
<b>8</b>	<b>Unbinned profiled unfolding . . . . . 76</b>
8.1	Motivation . . . . . 76
8.2	Unbinned Profiled Unfolding . . . . . 78
8.2.1	Statistical Setup . . . . . 78
8.2.2	Machine Learning Approach . . . . . 79
8.2.3	Machine Learning Implementation . . . . . 80
8.3	Gaussian Examples . . . . . 81
8.3.1	One-dimension in both particle and detector level . . . . . 81
8.3.2	One-dimension in particle level and two-dimension in detector level . . . . . 86
8.4	Higgs Boson Cross Section . . . . . 88
8.5	Summary . . . . . 89
<b>9</b>	<b>Conclusion . . . . . 92</b>
<b>APPENDIX</b>	<b>Binned maximum likelihood unfolding with Gaussian examples . . . . . 94</b>
<b>LIST OF REFERENCES</b>	<b>. . . . . 97</b>

## LIST OF TABLES

Table	Page
2.1	The branching ratios of the SM Higgs decays at $m_H = 125.09$ GeV. . . . . 9
6.1	Summary of the main event selection criteria applied to all events as well as the the selection of jets. The bottom sections give the basic requirements on leptons and $b$ -tagged jets for the analysis categories targeting different Higgs boson production modes. The subleading muon momentum threshold is 15 GeV in all categories except the $VH$ 3-lepton categories, where it is lowered to 10 GeV. . . . . 27
6.2	Summary of four-fold training method. . . . . 29
6.3	List of tested empirical functional forms for the background modelling. . . . . 42
6.4	Number of events observed in the $m_{\mu\mu} = 120$ – $130$ GeV window in data, the number of signal events expected in the SM ( $S_{\text{SM}}$ ), and events from signal ( $S = \mu \times S_{\text{SM}}$ ) and background ( $B$ ) as derived from the combined fit to the data with a signal strength parameter of $\mu = 1.2$ . The uncertainties in $S_{\text{SM}}$ correspond to the systematic uncertainty of the SM prediction, the uncertainty in $S$ is given by that in $\mu$ , and the uncertainty in $B$ is given by the sum in quadrature of the statistical uncertainty from the fit and the spurious signal uncertainty. In addition the observed number of signal events divided by the square root of the number of background events ( $S/\sqrt{B}$ ) and the signal-to-background ratio ( $S/B$ ) in % for each of the 20 categories described in the text are displayed. . . . . 44
7.1	Summary of selections used to define the signal regions used in the analysis. The kinematic variables are defined in the text. . . . . 59
7.2	Relative importance of the different sources of uncertainty for different $Z'$ -2HDMs, with the masses of the $Z'$ boson and the $A$ boson given in the second row, expressed as fractional impact on the signal strength parameter. The fractional impact is calculated by considering the square of the uncertainty in the signal strength parameter arising from a given group of uncertainties (as listed in the left column of the table), divided by the square of the total uncertainty in the signal strength parameter. Due to correlations, the sum of the different impacts of systematic uncertainties might not add up to the total impact of all systematic uncertainties. . . . . 67

Table	Page
7.3 Background yields in comparison with data in the 2 $b$ -tag signal regions for different $E_T^{\text{miss}}$ ranges after a background-only fit to data. Statistical and systematic uncertainties are reported together. . . . .	70
7.4 Background yields in comparison with data in the $\geq 3$ $b$ -tag signal regions for different $E_T^{\text{miss}}$ ranges after a background-only fit to data. Statistical and systematic uncertainties are reported together. . . . .	71

## LIST OF FIGURES

Figure	Page
2.1 The production cross sections of the SM Higgs boson (a) as a function of $m_H$ at $\sqrt{s} = 13\text{TeV}$ and (b) as a function of $\sqrt{s}$ with the experimental measured $m_H = 125.09\text{ GeV}$ . The central values are depicted by solid lines, while the colored bands illustrate the theoretical uncertainties. . . . .	5
2.2 Feynman diagrams showing the LO production of Higgs via (a) $ggF$ and (b) VBF processes. . . . .	6
2.3 Examples of LO Feynman diagrams for Higgs production via $VH$ , where the Higgs boson is produced in association with either a $W$ boson ( $WH$ ) or a $Z$ boson ( $ZH$ ). . .	7
2.4 Examples of LO Feynman diagrams for Higgs boson production via the $t\bar{t}H$ and $b\bar{b}H$ processes. . . . .	7
2.5 Examples of LO Feynman diagrams for Higgs boson production via the $tH$ process. . .	7
2.6 The branching ratios of the primary decay channels of the Higgs boson as a function of $m_H$ . . . . .	8
2.7 An example Feynman diagram of $Z'-2\text{HDM}$ in the mono- $h(b\bar{b})$ search. . . . .	10
2.8 Example diagrams for $2\text{HDM}+a$ with (a) $ggF$ and (b) $b\bar{b}A$ productions. . . . .	12
3.1 The large hadron collider illustrated in the map of the Geneva area in Switzerland. . .	14
3.2 Computer generated image of the entire ATLAS detector. . . . .	17
6.1 Dimuon invariant mass $m_{\mu\mu}$ in the range 76–160 GeV for all events passing the event selections. Data in points are compared to a full set of fully simulated background processes and the total background prediction is scaled to the integrated data yield. The shaded region in the bottom panel shows the impact of the systematic uncertainty on the muon momentum scale, resolution and muon trigger and reconstruction efficiencies. The $H \rightarrow \mu\mu$ signal shown is the sum of the $ggF$ , VBF, VH and $t\bar{t}H$ as open line, normalised to one hundred times the SM prediction for visibility. . . . .	28
6.2 Distributions of each training variable used for the $t\bar{t}H$ category (part 1). . . . .	32

Figure	Page
6.3 Distributions of each training variable used for the $t\bar{t}H$ category (part 2). . . . .	33
6.4 The $t\bar{t}H$ BDT score distributions for signal and background. The vertical grey dashed line represents the BDT score threshold of 0.35 used to select events in the $t\bar{t}H$ category.	34
6.5 Distributions of each training variable used for the $VH$ 3-lepton channel (part 1). . . .	45
6.6 Distributions of each training variable used for the $VH$ 3-lepton channel (part 2). . . .	46
6.7 Distributions of each training variable used for the $VH$ 4-lepton channel (part 1). . . .	47
6.8 Distributions of each training variable used for the $VH$ 4-lepton channel (part 2). . . .	48
6.9 The $VH$ BDT score distributions for signal and background in (a) 3-lepton channel and (b) 4-lepton channel. The grey vertical dashed lines indicate the BDT score boundary define each $VH$ categories. . . . .	48
6.10 Distributions of each training variable in the 0-jet channel. . . . .	49
6.11 Distributions of each training variable in the 1-jet channel (part 1). . . . .	50
6.12 Distributions of each training variable in the 1-jet channel (part 2). . . . .	51
6.13 Distributions of each training variable in the 2-jet channel (part 1). . . . .	52
6.14 Distributions of each training variable in the 2-jet channel (part 2). . . . .	53
6.15 The VBF (a) and $ggF$ (b)-(d) BDT score distributions for signal and background. The grey vertical dashed lines indicate the BDT score boundary define each categories. . .	54
6.16 Summary of the signal and background composition in the twenty categories in the mass region $120\text{GeV} < \mu^+\mu^- < 130\text{GeV}$ . The top panel shows the number of background events ( $B$ , multiplied by a factor of $10^{-5}$ ), the expected signal-to-background ratio for a SM $H \rightarrow \mu\mu$ signal ( $S/B$ ), as well as the number counting significance $S/\sqrt{B}$ . The middle panel shows the expected contributions of different Higgs boson production modes to the total signal in each category. The bottom panel shows the background composition from DY, diboson and top-quark processes in simulation. . .	55
6.17 Examples of signal parametrization in the categories “VBF Very High” and “2-jet Low”. . . . .	56

Figure	Page
6.18 Dimuon invariant mass spectrum in all the analysis categories observed in data. In (a) the unweighted sum of all events and signal plus background probability density functions (pdf) are shown, while in (b) events and pdfs are weighted by $\ln(1 + S/B)$ , where S are the observed signal yields and B are the background yields derived from the fit to data in the $m_{\mu\mu} = 120\text{--}130$ GeV window. The background and signal pdf are derived from the fit to the data, with S normalised to its best-fit value. The lower panels compare the fitted signal pdf, normalised to the signal best-fit value, to the difference between the data and the background model. The error bars represent the data statistical uncertainties. . . . .	56
7.1 Yields in the resolved and merged (a) 1-muon control regions and (b) 2-lepton control regions. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds. The different control region bins included in the fit are indicated on the $x$ -axis by first giving the range in $E_T^{\text{miss}}$ and then the sign of the muon charge (where applicable). . . . .	68
7.2 Distributions of the Higgs boson candidate mass in the 2 $b$ -tag signal regions for different $E_T^{\text{miss}}$ ranges. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds. . . . .	72
7.3 Distributions of the Higgs boson candidate mass in the $\geq 3$ $b$ -tag signal regions for different $E_T^{\text{miss}}$ ranges. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds. . . . .	73
7.4 $E_T^{\text{miss}}$ distributions in the (a) 2 $b$ -tag and (b) $\geq 3$ $b$ -tag signal regions for different $E_T^{\text{miss}}$ ranges. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds. . . . .	74
7.5 Exclusion limits in the $Z'$ -2HDM model. The solid black line shows the observed limit at 95% CL, the dashed black line the expected limit. The green band gives the $\pm 1\sigma$ uncertainties of the expected limit, the yellow band the $\pm 2\sigma$ uncertainties. . . .	74
7.6 Exclusion limits for the 2HDM+ $a$ signal with (a) $\tan\beta = 1$ and $ggF$ production and with (b) $\tan\beta = 10$ and $bbA$ production. The solid black line shows the observed limit at 95% CL, the dashed black line the expected limit. The green band gives the $\pm 1\sigma$ uncertainties of the expected limit, the yellow band the $\pm 2\sigma$ uncertainties. . . .	75
7.7 Model-independent upper limits on the visible cross-section $\sigma_{\text{vis},h(bb)+\text{DM}} \equiv \sigma_{h+\text{DM}} \times \mathcal{B}(h \rightarrow b\bar{b}) \times \mathcal{A} \times \varepsilon$ in the different signal regions. . . . .	75

Figure	Page
8.1 Gaussian 1D example: the nominal $R$ distribution ( $\epsilon = 1$ ) in the reweighted by the trained $w_1$ conditioned at $\epsilon = 1.2$ and compared to $R$ distribution with $\epsilon = 1.2$ . . . . .	82
8.2 Gaussian 1D example: results of the $w_0$ optimization. The nuisance parameter $\epsilon$ is fixed to 1.2, and the the penalty term in Eq. 8.7 is set to 0. (Top) The detector-level spectrum $R$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0 \times w_1$ , compared to the $R$ spectrum of the observed data $D_{\text{obs}}$ . (Bottom) The particle-level spectrum $T$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0$ , compared to the $T$ spectrum of the observed data $D_{\text{obs}}$ . . . . .	84
8.3 Gaussian 1D example: results of the $w_0$ optimization. The nuisance parameter $\epsilon$ is optimized simultaneously with $w_0$ and the best-fit value is $\hat{\epsilon} = 1.03 \pm 0.016$ . (Left) The detector-level spectrum $R$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0 \times w_1$ , compared to the $R$ spectrum of the observed data $D_{\text{obs}}$ . (Right) The particle-level spectrum $T$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0$ , compared to the $T$ spectrum of the observed data $D_{\text{obs}}$ . . . . .	85
8.4 Gaussian 2D example: the nominal detector-level spectra $R$ (left) and $R^*$ (right) with $\epsilon = 1$ reweighted by the trained $w_1$ conditioned at $\epsilon = 1.2$ and compared to the spectra with $\epsilon = 1.2$ . . . . .	85
8.5 Gaussian 2D example: results of the $w_0$ optimization. The nuisance parameter $\epsilon$ is optimized simultaneously with $w_0$ with the prior constraint set to 80%. The fitted $\epsilon$ is $1.20 \pm 0.004$ . (Top-left) The detector-level spectrum $R$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0 \times w_1$ , compared to the $R$ spectrum of the observed data $D_{\text{obs}}$ . (Top-right) The detector-level spectrum $R'$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0 \times w_1$ , compared to the $R^*$ spectrum of the observed data $D_{\text{obs}}$ . (Bottom) The particle-level spectrum $T$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0$ , compared to the $T$ spectrum of the observed data $D_{\text{obs}}$ . . . . .	87
8.6 Higgs boson cross section: the nominal detector-level spectra $m_{\gamma\gamma}$ (left) and $p_{\gamma\gamma}^{\text{T}}$ (right) with $\epsilon_\gamma = 1$ reweighted by the trained $w_1$ conditioned at $\epsilon_\gamma = 1.2$ and compared to the spectra with $\epsilon_\gamma = 1.2$ . . . . .	89
8.7 Higgs boson cross section: results of the $w_0$ optimization. The nuisance parameter $\epsilon_\gamma$ is optimized simultaneously with $w_0$ with the prior constraint set to 50% (orange) or fixed to 1 for comparison (red). The fitted $\epsilon_\gamma$ is $1.19 \pm 0.007$ . (Top-left) The detector-level spectrum $m_{\gamma\gamma}$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0 \times w_1$ , compared to the $m_{\gamma\gamma}$ spectrum of the observed data $D_{\text{obs}}$ . (Top-right) The detector-level spectrum $p_{\gamma\gamma}^{\text{T}}$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0 \times w_1$ , compared to the $p_{\gamma\gamma}^{\text{T}}$ spectrum of the observed data $D_{\text{obs}}$ . (Bottom) The particle-level spectrum $p_{\gamma\gamma}^{\text{T}}$ of the simulation template $D_{\text{sim}}$ reweighted by the trained $w_0$ , compared to the $p_{\gamma\gamma}^{\text{T}}$ spectrum of the observed data $D_{\text{obs}}$ . . . . .	90

Figure	Page
A.1 Gaussian 1D example: results of the binned maximum likelihood unfolding. The prior constraint for $\epsilon$ is set to 80% and the fitted $\epsilon$ is $1.08 \pm 0.02$ . (Left) The fitted detector-level spectrum $R$ of the simulation template $D_{\text{sim}}$ , compared to the $R$ spectrum of the observed data $D_{\text{obs}}$ . (Right) The unfolded particle-level spectrum $T$ of the simulation template $D_{\text{sim}}$ , compared to the $T$ spectrum of the observed data $D_{\text{obs}}$ . . . . .	95
A.2 Gaussian 2D example: results of the binned maximum likelihood unfolding. The prior constraint for $\epsilon$ is set to 80% and the fitted $\epsilon$ is $1.19 \pm 0.003$ . (Top-left) The fitted detector-level spectrum $R$ of the simulation template $D_{\text{sim}}$ , compared to the $R$ spectrum of the observed data $D_{\text{obs}}$ . (Top-right) The fitted detector-level spectrum $R^*$ of the simulation template $D_{\text{sim}}$ , compared to the $R'$ spectrum of the observed data $D_{\text{obs}}$ . (Bottom) The unfolded particle-level spectrum $T$ of the simulation template $D_{\text{sim}}$ , compared to the $T$ spectrum of the observed data $D_{\text{obs}}$ . . . . .	96

## ABSTRACT

The discovery of the Standard Model (SM) Higgs boson by ATLAS and CMS at the LHC in 2012 marked a major milestone in particle physics. However, many questions remain unanswered, which has led to an active research program to search for either rare SM phenomena or Beyond Standard Model (BSM) physics that involve the Higgs boson. In this dissertation, I present two example searches involving the Higgs boson, using proton-proton ( $pp$ ) collision data collected by the ATLAS detector.

The first search tackles the problem of how the SM Higgs couples to the second-generation fermions. It searches for the dimuon decay of the SM Higgs boson ( $H \rightarrow \mu\mu$ ) using data corresponding to an integrated luminosity of  $139 \text{ fb}^{-1}$  collected by the ATLAS detector in  $pp$  collisions at  $\sqrt{s} = 13 \text{ TeV}$  at the LHC. To identify this rare decay, we train boosted decision trees to separate signal and background. We obtain an observed (expected) significance over the background-only hypothesis for a Higgs boson with a mass of  $125.09 \text{ GeV}$  of  $2.0\sigma$  ( $1.7\sigma$ ). The observed upper limit on the cross section times branching ratio for  $pp \rightarrow H \rightarrow \mu\mu$  is 2.2 times the SM prediction at 95% confidence level, while the expected limit on a  $H \rightarrow \mu\mu$  signal assuming the absence (presence) of a SM signal is 1.1 (2.0). The best-fit value of the signal strength parameter, defined as the ratio of the observed signal yield to the one expected in the SM, is  $\mu = 1.2 \pm 0.6$ .

In the second search, we look for Dark Matter produced in association with a Higgs boson decaying to  $b$ -quarks. This search uses the same dataset as the  $H \rightarrow \mu\mu$  search and targets events that contain large missing transverse momentum and either two  $b$ -tagged small-radius jets or a single large-radius jet associated with two  $b$ -tagged subjets. We split events into multiple categories that target different phase spaces of the Dark Matter signals. We do not observe a significant excess from the SM prediction. We interpret the results using two benchmark models with two Higgs

doublets extended by either a heavy vector boson  $Z'$  (referred to as  $Z'-2\text{HDM}$ ) or a pseudoscalar singlet  $a$  (referred to as  $2\text{HDM}+a$ ) that provide a Dark Matter candidate  $\chi$ . For  $Z'-2\text{HDM}$ , the observed limits extend up to a  $Z'$  mass of 3.1 TeV at 95% confidence level for a mass of 100 GeV for the Dark Matter candidate. For  $2\text{HDM}+a$ , we exclude masses of  $a$  up to 520 GeV and 240 GeV for  $\tan\beta = 1$  and  $\tan\beta = 10$ , respectively, and for a Dark Matter mass of 10 GeV. Additionally, we set limits on the visible cross sections, which range from 0.05 fb to 3.26 fb, depending on the regions of missing transverse momentum and  $b$ -quark jet multiplicity.

In addition to the two physics analyses, unfolding is a key procedure in the high energy experiments, which corrects data for the detector effect. I present a new unfolding method that allows to unfold data without having any artificial binning and is also able to profile nuisance parameters simultaneously. This new method provides much higher flexibility and increases the reusability for different downstream tasks compared to the traditional approach. It will benefit any future analyses including Higgs physics and Dark Matter searches.

# Chapter 1

## Introduction

The Higgs boson [1–6] has played a crucial role in the Standard Model (SM) and the Higgs mechanism. It is responsible for particles acquiring mass through their interactions with the Higgs boson. According to the Higgs mechanism, the W and Z bosons acquire mass through the electroweak symmetry breaking, while fermions' mass is generated through the Yukawa coupling between fermions and the Higgs boson. The SM Higgs boson was discovered by the ATLAS and CMS experiments at the LHC in 2012 [7–9]. This achievement, along with subsequent observations [10–14], has resolved a number of puzzles in high energy physics. However, many questions related to the Higgs boson remain unanswered, such as whether it interacts with fermions in the first and second generations and whether it can be a portal to Beyond Standard Model (BSM) particles such as extra Higgs bosons arising from the Two-Higgs-Doublet Model (2HDM) [15, 16], and Dark Matter [17–21]. To answer these remaining questions, a broad program of searches for either rare SM phenomena or BSM that involve the Higgs boson have been proposed and initiated. In this dissertation, I will give a detailed review of two example searches that involve the Higgs boson and use the proton-proton (p-p) collision data collected by the ATLAS detector [22, 23]. The focus is on strategies that enhance the sensitivity to the searched signals.

The first search is a search for the dimuon decay of the SM Higgs boson ( $H \rightarrow \mu\mu$ ) [24]. It is the most promising channel for probing the couplings between the SM Higgs Boson and the second generation fermions at the LHC, thanks to excellent muon identification and reconstruction [25]. However, due to the small decay branching ratio and a large amount of background dominated by the Drell-Yan process, it is a very challenging analysis with a signal-to-background ratio typically at the level of 0.1%. The small signal to background ratio leads to two major challenges. Firstly,

it is difficult to establish the signal without good separation between signal and background. Secondly, measurements can be easily biased by statistical fluctuation and background mismodeling. We thus designed special techniques and analysis strategies to address these two challenges, including a categorization using dedicated boosted decision trees to separate signal and background and enhance the signal sensitivity, and a background modeling procedure based on spurious signal tests which ensure that background distributions can be modeled by the chosen functions without significant bias.

The second search is to look for Dark Matter produced in association with a Higgs boson decaying to  $b$ -quarks [26]. In contrast with the  $H \rightarrow \mu\mu$  search,  $H \rightarrow b\bar{b}$  has a much larger decay branching ratio and has been established prior to these analyses. Such a dominant decay channel is particularly useful to search for BSM, such as Dark Matter, which can be produced in association with the Higgs boson. This channel, which connects the Dark sector with the Higgs sector, has great sensitivity to DM models with the extended Higgs sector [15, 16], such as  $Z'-2\text{HDM}$  [27] and  $2\text{HDM}+a$  [28, 29]. It targets events that contain large missing transverse momentum and a Higgs boson which decays to  $b$ -quarks (referred as mono- $h(b\bar{b})$ ). The Higgs boson can be in low energy or highly boosted. We thus designed different strategies to better reconstruct the Higgs boson. Events are split by different kinematic variables into multiple categories which target different phase spaces of the DM signals. Upper limits on signal cross-section in a model-independent context, which will be useful for future reinterpretation with new DM models, are also calculated.

It is interesting to point out the main differences between these two analyses. Firstly,  $H \rightarrow \mu\mu$  search is a search for a single rare SM process, while the search of mono- $h(b\bar{b})$  looks for multiple possible BSM models with indefinite signal parameters. Secondly,  $H \rightarrow \mu\mu$  search uses a multivariate-variable-analysis (MVA) approach with machine learning to categorize the events and maximize the search sensitivity, while the search of mono- $h(b\bar{b})$  adopts the traditional cut-based approach, which reduces model dependency. Finally,  $H \rightarrow \mu\mu$  uses a data-driven approach with analytical functions to model the signal and background, while mono- $h(b\bar{b})$  estimates the signal and background with Monte Carlo (MC) simulation templates. These two analyses together serve as great examples that cover different perspectives of searches in high-energy particle experiments.

In addition to physics analyses, unfolding is an important procedure in particle physics experiments which corrects for detector effects and provides differential cross section measurements

that can be used for a number of downstream tasks, such as extracting fundamental physics parameters. Traditionally, unfolding is done by discretizing the target phase space into a finite number of bins and is limited in the number of unfolded variables. Recently, there have been a number of proposals to perform unbinned unfolding with machine learning. However, none of these methods (like most unfolding methods) allow for simultaneously constraining (profiling) nuisance parameters. We thus propose a new machine learning-based unfolding method, referred to as unbinned profiled unfolding (UPU) that results in an unbinned differential cross section and can profile nuisance parameters. The machine learning loss function is the full likelihood function, based on binned inputs at detector-level. We first demonstrate the method with simple Gaussian examples and then show the impact on a simulated Higgs boson cross section measurement.

The rest of the thesis is organized as follows: Chapter 2 briefly introduces the properties of the SM Higgs boson, its interactions with other particles and the extended Higgs sector associated with DM; Chapter 3 introduces the LHC and describes the ATLAS detector; Chapter 4 describes the data and simulation samples used for  $H \rightarrow \mu\mu$  and mono- $h(b\bar{b})$  analyses; Chapter 5 summarizes the physics object definitions used for  $H \rightarrow \mu\mu$  and mono- $h(b\bar{b})$  analyses; Chapter 6 details the analysis strategies for  $H \rightarrow \mu\mu$  while Chapter 7 describes the analysis strategies for mono- $h(b\bar{b})$ ; Chapter 8 presents the study of the newly proposed method of unbinned profiled unfolding; Chapter 9 provides a summary and outlook.

## Chapter 2

### SM Higgs boson and the extended Higgs sector

The SM Higgs boson [30–34] is a massive scalar boson with spin-0, CP-even, charge-neutral, and colorless properties. Its mass has been measured to be 125.09 GeV [10]. The Higgs boson ( $H$ ) can interact with quarks ( $q$ ), charged-leptons ( $\ell^\pm$ ),  $Z$ ,  $W^\pm$  as well as itself (self-coupling). The strength of these interactions is determined by the mass of the interacting particle. The Lagrangian of the Higgs couplings can be expressed as:

$$\mathcal{L} \supset -g_{Hff} f \bar{f} H + \delta_V V_\mu V^\mu \left( g_{HVV} H + \frac{g_{HHVV}}{2} H^2 \right) + \frac{g_{HHH}}{6} H^3 + \frac{g_{HHHH}}{24} H^4, \quad (2.1)$$

where  $f \in \{q, \ell^\pm\}$ ,  $V \in \{Z, W^\pm\}$ , and  $\delta_W = 1$ ,  $\delta_Z = \frac{1}{2}$ . The  $g$ 's in Eq. 2.1 represent the coupling strength to each particle:

$$g_{Hff} = \frac{m_f}{\nu}, \quad g_{HVV} = \frac{2m_V^2}{\nu}, \quad g_{HHVV} = \frac{2m_V^2}{\nu^2}, \quad g_{HHH} = \frac{3m_H^2}{\nu}, \quad g_{HHHH} = \frac{3m_H^2}{\nu^2}. \quad (2.2)$$

Specifically,  $g_{Hff}$  is proportional to the mass of the fermion, while  $g_{HVV}$  and  $g_{HHVV}$  are proportional to the square of the mass of the  $Z$  and  $W^\pm$  bosons, respectively. The coupling strength to the Higgs boson itself is represented by  $g_{HHH}$  and  $g_{HHHH}$ , both of which are proportional to the square of the Higgs boson mass,  $m_H$ . The vacuum expectation value of the Higgs field,  $\nu = (\sqrt{2}G_F)^{-\frac{1}{2}} \approx 246$  GeV, is used in Eq. 2.2 to calculate the coupling strengths.

In the following sections, I will briefly discuss the production and decay of the Higgs boson at the LHC, as well as the extended Higgs sector that includes additional particles that bridge the Dark sector.

#### 2.1 SM Higgs boson production at the LHC

The SM Higgs boson can be produced through several processes via  $pp$  collision at the LHC. The main production modes are as follows:

- Gluon fusion production ( $ggF$ )  $pp \rightarrow H$
- Vector boson fusion production (VBF)  $pp \rightarrow qqH$
- Associated production with a  $W$  boson ( $pp \rightarrow WH$ ) or a  $Z$  boson ( $ZH$ )  $pp \rightarrow ZH$
- Associated production with a pair of top quarks ( $t\bar{t}H$ )  $pp \rightarrow t\bar{t}H$ , or a pair of bottom quarks ( $b\bar{b}H$ )  $pp \rightarrow b\bar{b}H$
- Associated production with a single top quark ( $tH$ )  $pp \rightarrow tH$

The production cross-sections for each mode depend on the Higgs boson mass  $m_H$  and the center-of-mass energy ( $\sqrt{s}$ ) of the  $pp$  collision. Fig. 2.1 (a) summarizes the production cross-sections as a function of  $m_H$  at  $\sqrt{s} = 13$  TeV, while Fig. 2.1 (b) shows the cross-sections as a function of  $\sqrt{s}$  with the experimental measured  $m_H = 125.09$  GeV.

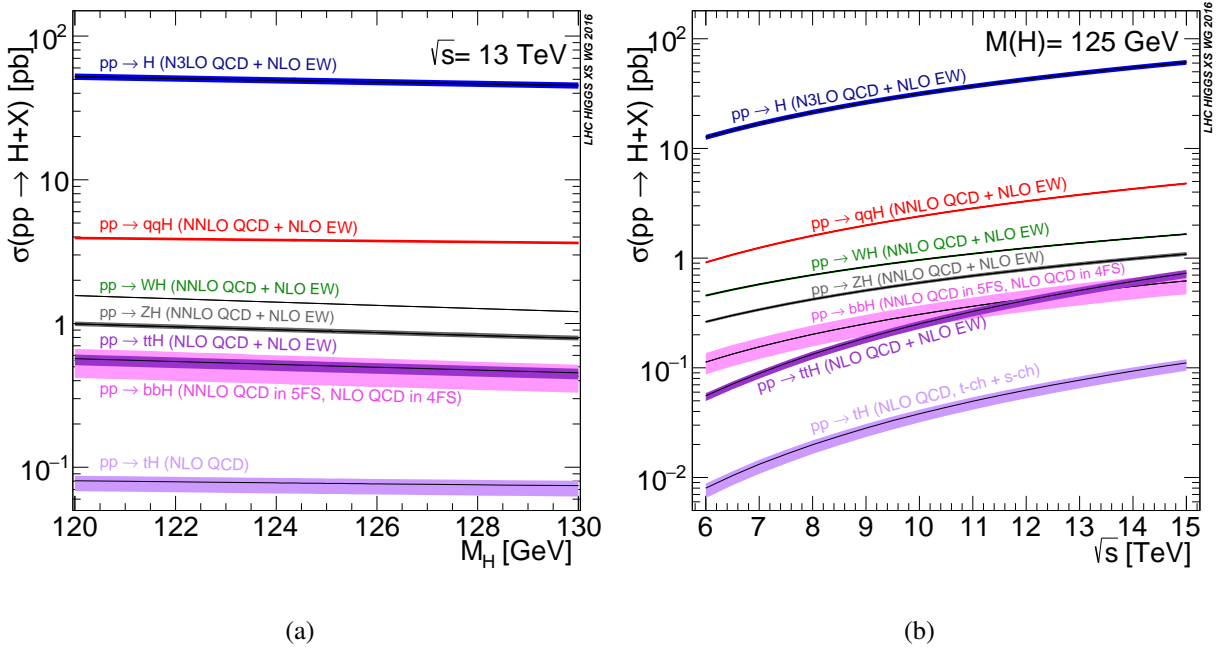


Figure 2.1: The production cross sections of the SM Higgs boson (a) as a function of  $m_H$  at  $\sqrt{s} = 13$  TeV and (b) as a function of  $\sqrt{s}$  with the experimental measured  $m_H = 125.09$  GeV. The central values are depicted by solid lines, while the colored bands illustrate the theoretical uncertainties.

The  $ggF$  production is expected to dominate as it involves the strong interaction between gluons, which is very strong at the LHC energy scales. As gluons are massless, their interaction with the Higgs boson must be mediated through a heavy quark loop. Fig. 2.2 (a) shows the leading-order (LO) Feynman diagram of the  $ggF$  production.

The VBF production has the second largest production cross-section. This process is characterized by two forward jets, which are crucial for its identification. Since the Higgs boson is produced through interactions with vector bosons, this process can be used to measure the Higgs couplings to vector bosons. Fig. 2.2 (b) shows the LO Feynman diagram of the VBF production.

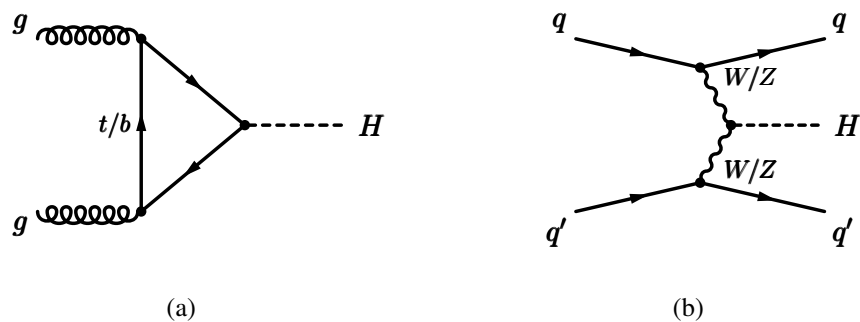


Figure 2.2: Feynman diagrams showing the LO production of Higgs via (a)  $ggF$  and (b) VBF processes.

The next main production mode is the  $VH$  production. The Higgs boson is produced in association with either a  $W$  boson ( $WH$ ) or a  $Z$  boson ( $ZH$ ). The vector boson can decay leptonically or hadronically, but the leptonic decay channels are usually easier to identify due to good reconstruction and resolution of the leptons. The LO Feynman diagrams of the  $VH$  production are shown in Fig. 2.3.

Although the  $t\bar{t}H$ ,  $b\bar{b}H$  and  $tH$  productions have relatively small cross-sections, they are important for direct measurements of the Yukawa couplings to the third-generation fermions as the Higgs boson is produced via the interactions with these fermions. Among them,  $t\bar{t}H$  has been extensively explored to probe the large top-Higgs Yukawa coupling. The LO Feynman diagrams for  $t\bar{t}H$  and  $b\bar{b}H$  productions are shown in Fig. 2.4, while the LO Feynman diagrams of  $tH$  are shown in Fig. 2.5.

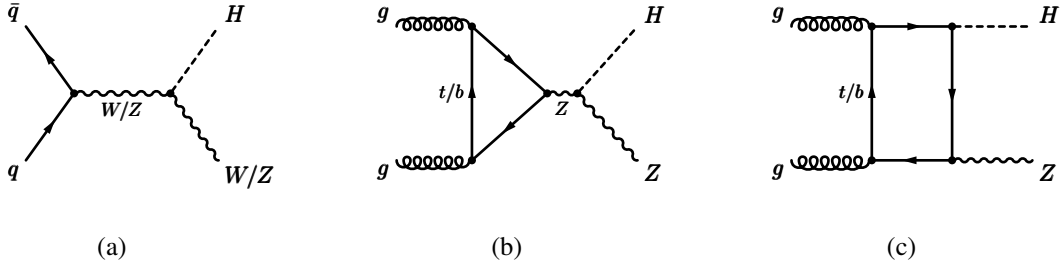


Figure 2.3: Examples of LO Feynman diagrams for Higgs production via  $VH$ , where the Higgs boson is produced in association with either a  $W$  boson ( $WH$ ) or a  $Z$  boson ( $ZH$ ).

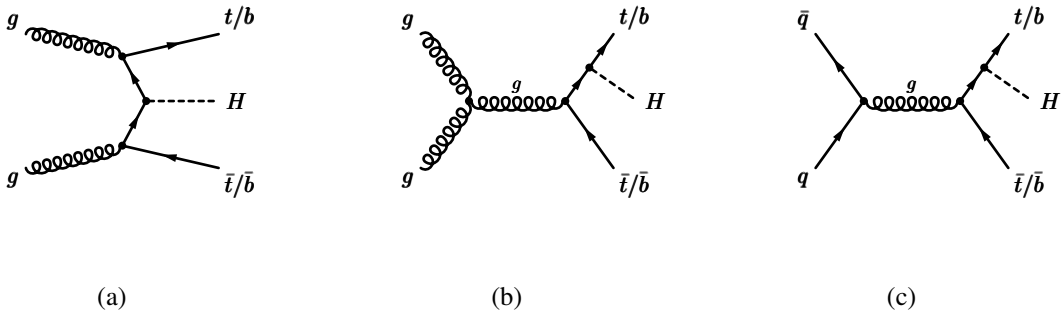


Figure 2.4: Examples of LO Feynman diagrams for Higgs boson production via the  $t\bar{t}H$  and  $b\bar{b}H$  processes.

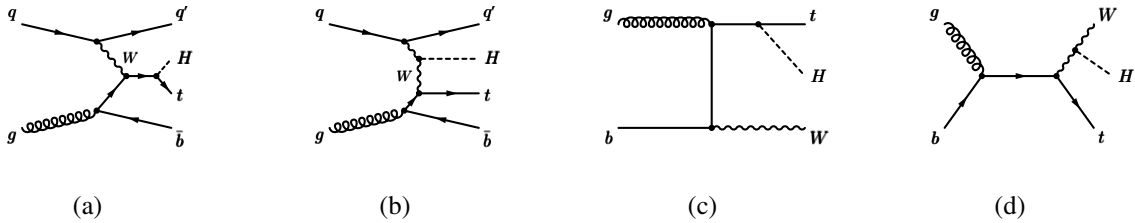


Figure 2.5: Examples of LO Feynman diagrams for Higgs boson production via the  $tH$  process.

## 2.2 SM Higgs boson decay at the LHC

The Higgs boson is a short-lived particle with a lifetime predicted to be approximately  $10^{-22}$  s, corresponding to a width of approximately 4 MeV. Therefore, the Higgs boson can only be observed indirectly through its decay products at the LHC. Fig. 2.6 displays the branching ratios of the primary decay channels of the Higgs boson as a function of  $m_H$ . The branching ratios of the Higgs decays at  $m_H = 125.09$  GeV are listed in Tab. 2.1. At around  $m_H = 125$  GeV,  $H \rightarrow b\bar{b}$  is

the dominant decay channel, with a branching ratio of 57.5%, whereas  $H \rightarrow \mu\mu$  has a relatively small branching ratios due to the very small value of the muon mass.

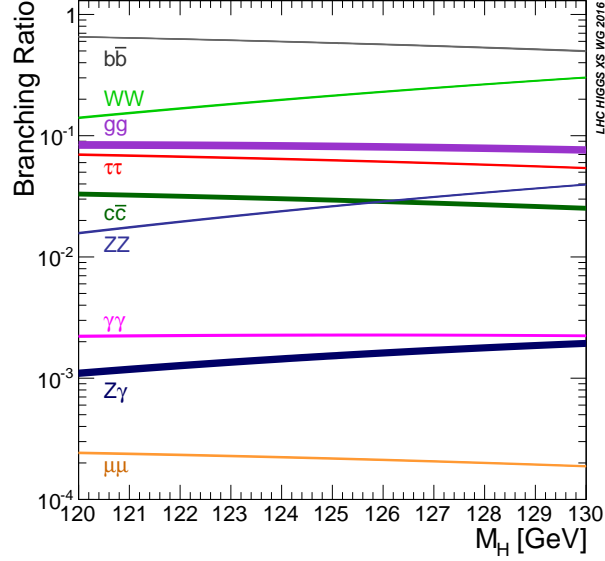


Figure 2.6: The branching ratios of the primary decay channels of the Higgs boson as a function of  $m_H$ .

### 2.3 Extended Higgs sector — type II Two-Higgs-Doublet Model

The SM assumes that the Higgs sector is an  $SU(2)$  doublet, which can be extended with a second Higgs doublet to create the Two-Higgs-Doublet Model (2HDM) [15, 16]. The 2HDM is a key component of many well-motivated BSM theories, such as supersymmetry [35] and axion models [36, 37], and is motivated by the flexibility it provides in the scalar mass spectrum and the existence of additional sources of CP violation [38]. It also enables the connection to the Dark sector through extra particles. Here, the focus will be on the type II 2HDM (2HDM II), where one of the Higgs doublets couples to up-type quarks, while the other couples to down-type quarks and leptons. The Yukawa couplings are expressed by:

$$\mathcal{L} \supset - \left( \bar{Q} Y_u \tilde{\Phi}_u u_R + \bar{Q} Y_d \tilde{\Phi}_d d_R + \bar{L} Y_\ell \Phi_d \ell_R + \text{h.c.} \right), \quad (2.3)$$

where  $Y_f$  are the Yukawa couplings,  $Q$  and  $L$  are the left-handed quark and lepton doublets, and  $f_R$  are the right-handed fermions.  $\Phi_u$  and  $\Phi_d$  are the two Higgs doublets, which are parametrized

Decay channel	Branching ratio [%]
$H \rightarrow b\bar{b}$	$57.5 \pm 1.9$
$H \rightarrow WW^{(*)}$	$21.6 \pm 0.9$
$H \rightarrow gg$	$8.56 \pm 0.86$
$H \rightarrow \tau\tau$	$6.30 \pm 0.36$
$H \rightarrow c\bar{c}$	$2.90 \pm 0.35$
$H \rightarrow ZZ^{(*)}$	$2.67 \pm 0.11$
$H \rightarrow \gamma\gamma$	$0.228 \pm 0.011$
$H \rightarrow Z\gamma$	$0.155 \pm 0.014$
$H \rightarrow \mu\mu$	$0.022 \pm 0.001$

Table 2.1: The branching ratios of the SM Higgs decays at  $m_H = 125.09$  GeV.

as:

$$\begin{aligned}\Phi_u &= \frac{1}{\sqrt{2}} \begin{pmatrix} \cos \beta H^+ \\ \nu_u + \cos \alpha h + \sin \alpha H + i \cos \beta A \end{pmatrix}, \\ \Phi_d &= \frac{1}{\sqrt{2}} \begin{pmatrix} -\sin \beta H^+ \\ \nu_d - \sin \alpha h + \cos \alpha H - i \sin \beta A \end{pmatrix},\end{aligned}\tag{2.4}$$

where  $h, H$  are neutral CP-even Higgs bosons,  $H^\pm$  are charged Higgs bosons, and  $A$  is a neutral CP-odd Higgs boson (pseudo-scalar).  $\nu_f$  are the vacuum expectation values of the two Higgs doublets, which satisfy  $\nu = \sqrt{\nu_u^2 + \nu_d^2} \approx 246$  GeV.

The electroweak (EW) symmetry breaking leads to interactions between the CP-even mass eigenstates and the EW gauge bosons, which can be written as:

$$\mathcal{L} \supset \delta_V V_\mu V^\mu g_{HVV} (\sin(\beta - \alpha) h + \cos(\beta - \alpha) H),\tag{2.5}$$

where  $\alpha$  is the mixing angle between  $h$  and  $H$ , and  $\beta$  is defined with  $\tan \beta \equiv \frac{\nu_u}{\nu_d}$ . To simplify the model, it is well-motivated to enforce the alignment limit  $\alpha = \beta - \pi/2$ , such that  $h$  is consistent with the SM Higgs boson (and therefore  $m_h \approx 125$  GeV). Note that  $H$  here denotes the other Higgs boson (as opposed to the SM Higgs boson), which is likely to be heavier under experimental constraints. This notation will be used throughout the context of this model.

### 2.3.1 $Z'$ -2HDM

The  $Z'$ -2HDM [27] assumes that the pseudo-scalar  $A$  in the 2HDM II is the sole mediator of the interaction between SM and DM. It extends the 2HDM II with an additional vector boson,  $Z'$ , which mixes with the  $Z$  boson. By coupling the DM to the pseudo-scalar,  $A$ , constraints from DM coupling to SM Higgs boson or vector bosons are avoided. The additional vector boson  $Z'$ , on the other hand, allows for resonant production of  $Z' \rightarrow h + A$  (and  $A \rightarrow \chi\chi$ , where  $\chi$  represents the DM particle). This makes it an interesting signal process to look for in the mono- $h(b\bar{b})$  search, as the Feynman diagram shown in Fig. 2.7. Furthermore, the model assumes that  $Z'$  couples only to right-handed quarks and the down-type Higgs doublet  $\Phi_d$ , which avoids potentially stringent limits from dilepton resonance searches. It is assumed that the coupling between  $A$  and  $\chi$  is strong, such that the  $A \rightarrow \chi\chi$  decay has a very large branching ratio, as long as the  $A \rightarrow t\bar{t}$  decay is kinematically forbidden.

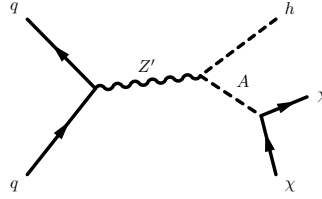


Figure 2.7: An example Feynman diagram of  $Z'$ -2HDM in the mono- $h(b\bar{b})$  search.

The  $Z'$ -2HDM involves eight parameters, namely  $\tan \beta$ ,  $m_H$ ,  $m_{H^\pm}$  and  $m_A$  from the 2HDM II;  $g_Z$  (the gauge coupling of  $Z'$ ),  $m_{Z'}$  (the mass of  $Z'$ ),  $m_\chi$  (the mass of the DM particle  $\chi$ ), and  $g_\chi$  (the coupling between  $A$  and  $\chi$ ). To simplify the analysis, we assume  $m_H = m_{H^\pm} = m_A$  and set  $\tan \beta$  and  $g_\chi$  to unity, and  $g_Z$  to 0.8. Since these parameters have negligible impact on the kinematic distributions, but only on the production cross-sections and decay branching ratios, the results can be easily scaled for different values of these parameters. The choices of these values are made to have the most interesting phenomenology as recommended by the LHC Dark Matter Working Group [28]. Consequently, in the interpretation of the mono- $h(b\bar{b})$  analysis, only  $m_{Z'}$  and  $m_A$  are the free parameters .

### 2.3.2 2HDM+ $a$

The 2HDM+ $a$  [28,29] introduces an extra pseudo-scalar, denoted by  $a$ , to the 2HDM II, which mixes with the pseudo-scalar  $A$  and serves as a mediator between the DM and SM sectors. The interaction between  $a$  and the DM particle  $\chi$  is expressed by:

$$\mathcal{L} \supset -ig_\chi a \bar{\chi} \gamma_5 \chi, \quad (2.6)$$

where  $g_\chi$  is the coupling strength between  $a$  and  $\chi$ . Meanwhile, the relevant couplings between Higgs bosons and SM fermions are given by:

$$\begin{aligned} \mathcal{L} \supset & -g_{htt} \bar{t} [h + \xi_t H - i\xi_t (\cos \theta A - \sin \theta a) \gamma_5] t \\ & - \sum_{f=b,\tau} g_{hff} \bar{f} [h + \xi_f H + i\xi_f (\cos \theta A - \sin \theta a) \gamma_5] f \\ & - g_{htt} V_{tb} \xi_t H^+ \bar{t}_R b_L + g_{hbb} V_{tb} \xi_b H^+ \bar{t}_L b_R + \text{h.c.}, \end{aligned} \quad (2.7)$$

where  $g_{hff}$  are the SM Yukawa couplings defined in Eq. 2.2,  $V_{ij}$  are the elements of the Cabibbo-Kobayashi-Maskawa matrix elements, and  $\xi_f$  are given by:

$$\xi_t = -\cot \beta, \quad \xi_b = \xi_\tau = \tan \beta. \quad (2.8)$$

Notably, large values of  $\cot \beta$  result in stronger couplings between  $t$  and the additional Higgs bosons, while larger values of  $\tan \beta$  lead to stronger couplings between  $b(\tau)$  and the additional Higgs bosons. The model yields rich phenomenology with various signatures and complementarity among different experimental searches. As depicted in Fig. 2.8, there are two primary ways of producing the  $A \rightarrow ha \rightarrow h\chi\chi$  process, namely, gluon fusion production ( $ggF$ ) and associated production with a pair of bottom quarks ( $b\bar{b}A$ ), which give rise to the mono- $h(b\bar{b})$  signature of interest.

The 2HDM+ $a$  can be fully characterized by eleven parameters:  $\tan \beta$ ,  $m_H$ ,  $m_{H^\pm}$  and  $m_A$  from the 2HDM II; the mixing angle between the two pseudo-scalars  $\theta$ ; the quartic couplings between the Higgs bosons  $\lambda_3$ ,  $\lambda_{a1}$ ,  $\lambda_{a2}$ ; the mass of the additional pseudo-scalar  $m_a$ ; the coupling to the Dark sector  $g_\chi$ ; and the mass of the DM particle  $m_\chi$ . To simplify the analysis, we set  $m_A = m_{H^\pm} = m_H$  to avoid constraints from EW precision measurements. Furthermore, we set  $\lambda_3 = \lambda_{a1} = \lambda_{a2} = 3$  to ensure the stability of the Higgs potential and maximize the trilinear

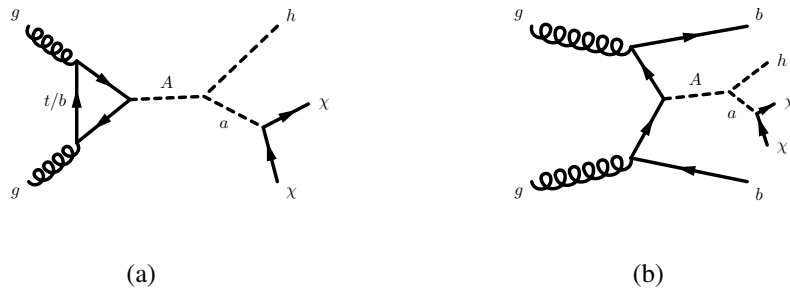


Figure 2.8: Example diagrams for 2HDM+ $a$  with (a)  $ggF$  and (b)  $b\bar{b}A$  productions.

couplings between CP-even and CP-odd Higgs bosons. Lastly, we set  $g_\chi = 1$  to obtain a large branching ratio of  $a \rightarrow \chi\chi$ , provided that the decay is kinematically allowed. Like  $Z'$ -2HDM, these values are chosen to have the most interesting phenomenology as recommended by the LHC Dark Matter Working Group [28]. These considerations reduce the number of free parameters to five, namely  $m_A$ ,  $m_a$ ,  $\tan\beta$ ,  $\theta$  and  $m_\chi$ .

## Chapter 3

# The Large Hadron Collider and the ATLAS detector

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [39] is a particle accelerator designed to study proton-proton collisions at unprecedented energy levels. It is located at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland and is the largest and most powerful particle accelerator in the world.

As illustrated in Fig. 3.1, the LHC is a circular accelerator with a circumference of 27 kilometers (17 miles) and is situated underground, at a depth of 100 meters (328 feet). The protons are accelerated to nearly the speed of light and then made to collide at four different points around the LHC, where large detectors such as ATLAS [22, 23] and CMS [40] are located.

The energy scale of the LHC is immense, with proton-proton collisions occurring at center-of-mass energies up to 13 TeV. This is over 100 times the energy achieved by the previous largest accelerator, the Tevatron [41] at Fermilab in the United States.

The LHC has undergone several upgrades since it began operations in 2008. The second run of the LHC, known as Run 2, began in 2015 and lasted until the end of 2018. During this run, the LHC collided protons at higher energies and intensities than ever before, allowing scientists to study the properties of the Higgs boson in greater detail and search for new particles beyond the standard model of particle physics.

The current phase of the LHC is Run 3, which began in 2022. The main goal of Run 3 is to collect data on proton-proton collisions at a higher luminosity, which is the number of collisions per second. This will enable scientists to study rare processes and phenomena that occur at higher energy scales than previously explored. The Run 3 is expected to continue until 2024, followed by a period of maintenance and upgrade in preparation for the High-Luminosity LHC project.

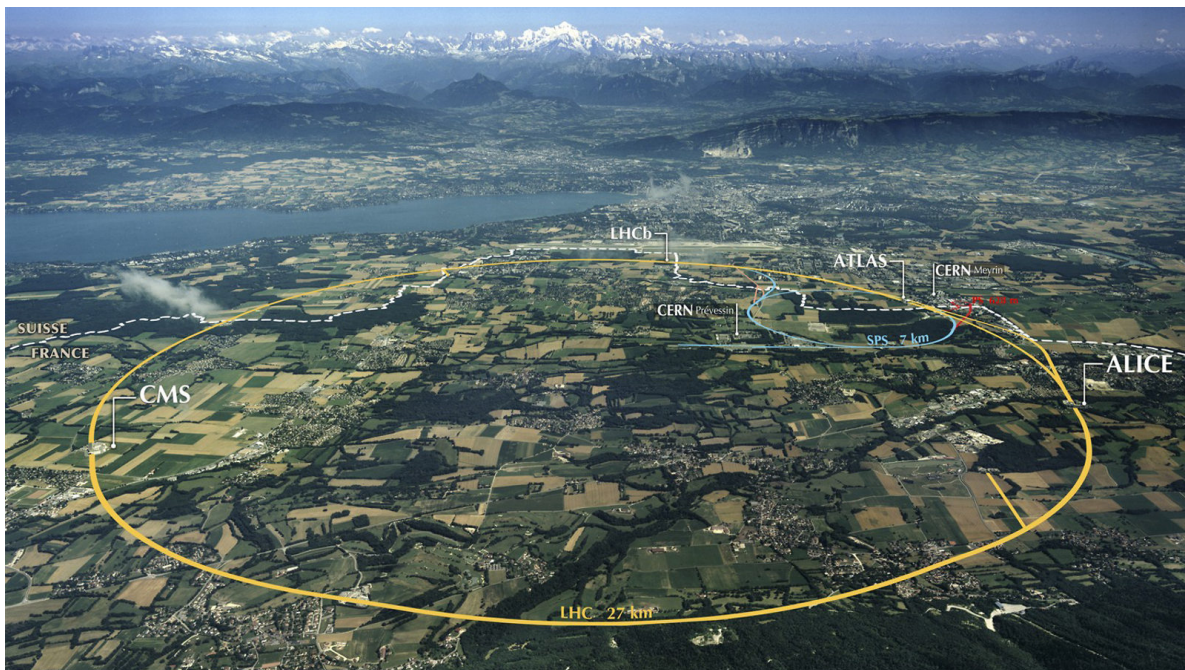


Figure 3.1: The large hadron collider illustrated in the map of the Geneva area in Switzerland.

The High-Luminosity LHC (HL-LHC) project involves upgrading the accelerator and detectors to increase the collision rate and improve the accuracy of the measurements. The HL-LHC is currently planned to start operation in 2028 and will further increase the collision rate and luminosity of the LHC, opening new avenues of research in particle physics.

### 3.2 The ATLAS detector

The ATLAS experiment [22, 23] is a versatile detector that features a forward-backward symmetric cylindrical geometry and nearly complete solid angle coverage of  $4\pi$ . As shown in Fig. 3.2, the ATLAS detector consists of an inner tracking detector, a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer that includes three large superconducting toroidal magnet systems.

The inner-detector system (ID) is subject to a 2 T axial magnetic field that bends charged particles in the  $r - \phi$  plane and provides tracking capabilities within the pseudorapidity range  $|\eta| < 2.5$ . The high-granularity silicon pixel detector covers the vertex region and typically provides four position measurements (hits) per track, with the first hit usually in the insertable B-layer installed before Run 2. It is followed by the silicon microstrip tracker (SCT), which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to  $|\eta| = 2.0$ .

The calorimeter system has approximately 188,000 cells and covers the range  $|\eta| < 4.9$ . Within the region  $|\eta| < 3.2$ , electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) sampling calorimeters (ECAL), with an additional thin LAr presampler covering  $|\eta| < 1.8$  to correct for energy loss in material upstream of the calorimeters. The ECAL has a depth between 24 and 27 radiation lengths ( $X_0$ ), and its granularity in the barrel in terms of  $\Delta\eta \times \Delta\phi$  is typically  $0.025 \times \pi/128$ , with variations in segmentation with layer and  $|\eta|$  as described in Ref. [42].

Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter (HCAL), which is segmented into three barrel structures within  $|\eta| < 1.7$ , and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules (FCAL) optimized for electromagnetic (FCAL1) and hadronic (FCAL2 and FCAL3) measurements, respectively. The combined depth of the calorimeters for hadronic energy

measurements is more than 10 nuclear interaction lengths nearly everywhere across the full detector acceptance ( $|\eta| < 4.9$ ). The granularity is as fine as  $0.1 \times \pi/32$ , again with variations in segmentation with layer and  $|\eta|$  as described in Ref. [42].

The muon spectrometer (MS) includes two types of chambers: trigger and high-precision tracking chambers. These chambers measure the deflection of muons in the  $r$ - $z$  plane due to a magnetic field produced by superconducting air-core toroids. The field integral of the toroids ranges from 2.0 to 6.0 T·m throughout most of the detector. Precision chambers cover the region  $|\eta| < 2.7$  with three stations of monitored drift tube (MDT) chambers. In the  $|\eta| > 2.0$  region, where the background is higher, the innermost MDT station is replaced with cathode-strip chambers (CSCs). Each MDT chamber provides six to eight  $\eta$  measurements along the muon track, while the CSCs provide four simultaneous measurements of  $\eta$  and  $\phi$ . The nominal single-hit resolution of the MDTs and CSCs is about 80  $\mu\text{m}$  and 60  $\mu\text{m}$ , respectively, in the bending plane. The chambers are accurately aligned using optical sensors to achieve a 10% transverse momentum resolution for 1 TeV muons. The muon trigger system covers the range  $|\eta| < 2.4$  with resistive-plate chambers (RPCs) in the barrel, consisting of three doublet stations for  $|\eta| < 1.05$ , and thin-gap chambers (TGCs) in the endcap regions, consisting of one triplet station followed by two doublets for  $1.0 < |\eta| < 2.4$ . The RPCs and TGCs provide tracking information complementary to the precision chambers, especially improving the determination of the track coordinate in the non-bending direction, referred to as the second coordinate. The typical spatial resolution for the position measurements in the RPCs and TGCs is 5-10 mm in both the bending plane and non-bending direction.

To select interesting events, a two-level trigger system [43] is used, consisting of a hardware-based first-level (Level-1, L1) and a software-based high-level trigger (HLT). The L1 trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the HLT reduces in order to record events to disk at about 1 kHz.

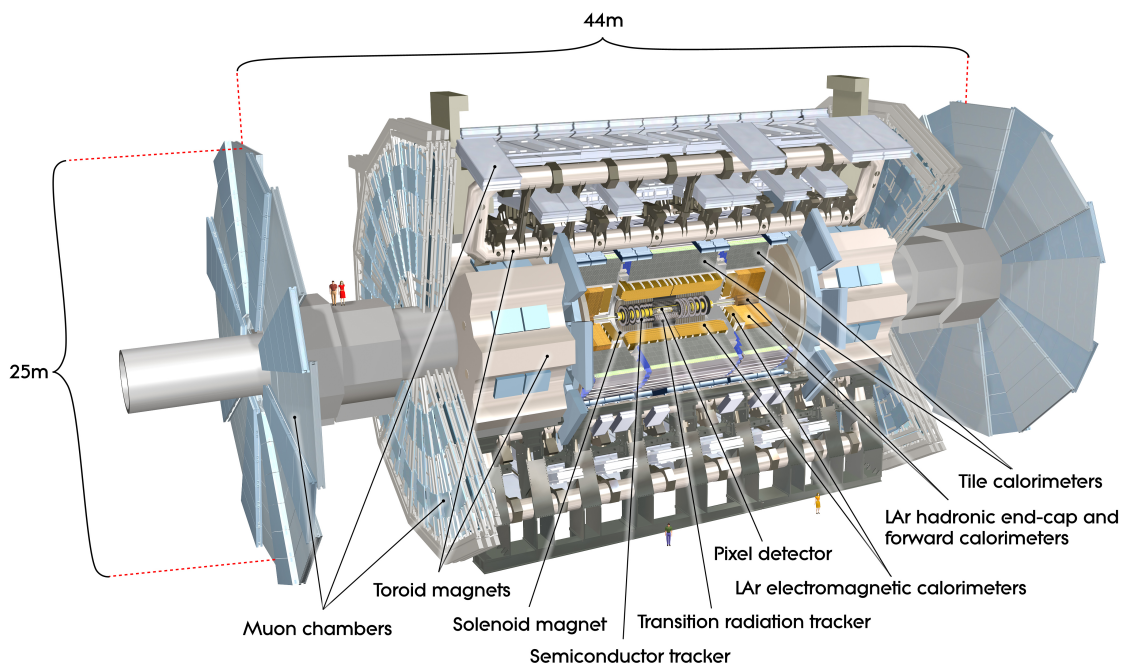


Figure 3.2: Computer generated image of the entire ATLAS detector.

## Chapter 4

### Data and simulated event samples

#### 4.1 Data samples

Both the  $H \rightarrow \mu\mu$  and mono- $h(b\bar{b})$  analyses utilize  $pp$  collision data collected by the ATLAS detector at  $\sqrt{s} = 13$  TeV from 2015 to 2018. To ensure the proper functioning of all components of the ATLAS detector, basic data quality requirements were imposed on all events. The full dataset corresponds to an integrated luminosity of  $139 \text{ fb}^{-1}$ , with an uncertainty of 1.7%. For the  $H \rightarrow \mu\mu$  analysis, events must satisfy a combination of single-muon triggers with transverse momentum ( $p_T$ ) thresholds up to 26 GeV for isolated muons and 50 GeV for muons without any isolation requirement imposed, allowing for the recovery of some inefficiency introduced by the isolation requirement at the trigger level for high momentum muons. For the mono- $h(b\bar{b})$  analysis, events must pass the most efficient available  $E_T^{\text{miss}}$  trigger, which reaches full efficiency by an offline  $E_T^{\text{miss}}$  value of approximately 200 GeV. In addition, mono- $h(b\bar{b})$  also employs events passing single-electron or single-muon triggers [44, 45] to construct 2-lepton control regions (described in section 7.1.3).

#### 4.2 Simulated event samples

Simulated Monte Carlo (MC) samples are generated to optimize event selections and model signal and background processes. The simulation procedure involves generating parton-level events, processing them for parton showering, hadronization, underlying events, etc., and processing them through the ATLAS detector simulation. For all samples, the detector simulation is based on GEANT4 [46, 47] (referred to as full-simulation samples), except for an additional sample for the Drell-Yan process ( $Z/\gamma^* \rightarrow \mu\mu$ ), where experimental effects are approximated using parametrizations (referred to as fast-simulation samples) to speed up the process.

The simulation also includes the effects of multiple  $pp$  collisions in the same or neighboring bunch crossings (pile-up) by overlaying inelastic  $pp$  interactions produced using PYTHIA 8 [48] with the NNLPDF2.3LO set of parton distribution functions (PDFs) [49] and the A3 set of tuned parameters [50]. Events are reweighted to match the distribution of the average number of interactions per bunch crossing observed in data. Simulated events are corrected to reflect the momentum scales and resolutions, as well as the triggers, reconstruction, identification, and isolation efficiencies measured in data for all physics objects used in the analyses.

### 4.2.1 Simulated event samples for $H \rightarrow \mu\mu$

The signal for the  $H \rightarrow \mu\mu$  analysis is the  $H \rightarrow \mu\mu$  process. The  $H \rightarrow \mu\mu$  samples are generated for the main Higgs boson production modes, including  $ggF$ , VBF,  $VH$  and  $t\bar{t}H$ . The Higgs boson mass is set to 125 GeV and the corresponding decay width is  $\Gamma_H = 4.07$  MeV [51]. The samples are normalized using the latest available theoretical calculations of the corresponding SM production cross sections [52], as well as the  $H \rightarrow \mu\mu$  branching ratio of  $2.17 \times 10^{-4}$  calculated with HDECAY [53–56] and PROPHECY4F [57–59]. The  $ggF$  sample is simulated using the POWHEG NNLOPS program [60–67] with the PDF4LHC15 set of PDFs [68]. The simulation achieves next-to-next-to-leading-order (NNLO) accuracy in QCD for inclusive observables after reweighting the Higgs boson rapidity spectrum [69]. The parton-level events are processed by PYTHIA 8 for the Higgs decay, parton showering, final-state radiation (QED FSR), hadronization, and underlying event using the AZNLO set of tuned parameters [70]. The sample is normalized to a next-to-next-to-next-to-leading-order QCD calculation with next-to-leading-order (NLO) electroweak corrections [71–82]. The VBF and  $q\bar{q}/gg \rightarrow VH$  samples are generated at NLO accuracy in QCD using the POWHEG-BOX program [83–85]. The loop-induced  $gg \rightarrow ZH$  sample is generated at leading order (LO) using POWHEG-BOX. The same settings for the PDF set and PYTHON 8 as for the  $ggF$  sample are adopted for the VBF and  $VH$  samples. The VBF sample is then normalized to an approximate-NNLO QCD cross-section with NLO EW corrections [86–88]. The  $VH$  samples are normalized to NNLO QCD cross-section with NLO electroweak corrections for  $q\bar{q}/gg \rightarrow VH$  and NLO + next-to-leading-logarithm (NNLL) accuracy QCD for  $gg \rightarrow ZH$  [89–96]. The  $t\bar{t}H$  sample is generated by MADGRAPH5\_aMC@NLO [97,98] at NLO accuracy in QCD with the NNPDF3.0NLO PDF set [99] and interfaced to PYTHIA 8 using

the A14 set of tuned parameters [100]. The sample is normalized to NLO QCD cross-section with NLO EW corrections [101–104].

The background processes comprise Drell-Yan (DY)  $Z/\gamma^* \rightarrow \mu\mu$  process, diboson ( $WW$ ,  $WZ$  and  $ZZ$ ) process,  $t\bar{t}$ , single-top-quark productions, as well as  $t\bar{t}V$  process. The full-simulated DY and diboson samples are simulated by SHERPA 2.2.1 [105]. Specifically, the DY events are produced using NLO-accurate matrix elements for up to two partons and LO-accurate matrix elements for up to four partons calculated with the Comix [106] and OpenLoops [107, 108] libraries and the NNPDF3.0 NNLO set. They are matched to the SHERPA parton shower [109] using the MEPS@NLO prescription [110–113]. The  $t\bar{t}$  and single-top-quark samples are generated at NLO accuracy with POWHEG-BOX [114, 115] using the NNPDF3.0NLO PDF set interfaced to PYTHIA 8 for parton showering and hadronization using the A14 parameter set. Additionally, the  $Wt$  process undergoes the diagram removal scheme [116] to remove the overlap with the  $t\bar{t}$  production. The  $t\bar{t}V$  sample is generated using MADGRAPH5\_aMC@NLO [97] at NLO in a set-up similar to the  $t\bar{t}H$  sample.

To provide enough statistics for background modeling with low statistical uncertainties, an additional fast-simulated DY samples is produced. This sample is generated using SHERPA 2.2.4 [117] with LO matrix elements and up to three additional partons. It uses the CT14 NNLO PDF set [118]. The parton-level events are then processed with PYTHIA 8 for QED and QCD parton showering and hadronization. The CKKW-L algorithm [119] is used to remove the double-counted QCD emissions with a merging scale of 20 GeV. The experimental effects are parametrized based on the full-simulation samples or directly from ATLAS data. This reproduces the reconstruction and selection efficiencies of detector-level objects using event weighting and models the resolution of the ATLAS detector with predetermined probability distributions. Detailed descriptions are used for the muon momentum resolution and muon trigger and selection efficiencies, photons from QED FSR, hadronic jets from primary interaction, pile-up events in terms of kinematics and the number of associated ID tracks, and for the effect of pile-up and the underlying event on the measurement of the missing transverse momentum  $E_{\text{T}}^{\text{miss}}$ .

### 4.2.2 Simulated event samples for mono- $h(b\bar{b})$

The signal for the mono- $h(b\bar{b})$  analysis comprises  $Z'$ -2HDM and 2HDM+ $a$ . Both models are simulated using MADGRAPH5\_aMC@NLO v2.6.5 [97] at LO in QCD, interfaced with PYTHIA 8 using the A14 set of tuned parameters.  $Z'$ -2HDM and the  $ggF$  production for 2HDM+ $a$  are simulated in the 5-flavor scheme, while the  $b\bar{b}A$  production of 2HDM+ $a$  is simulated in the 4-flavor scheme. Multiple parameters values are used for the parameter scans of the interpretations.

The background consists of single vector-boson productions ( $Z$  and  $W$ ), diboson productions,  $t\bar{t}$ , single top-quark,  $t\bar{t}V$ ,  $t\bar{t}H$  and  $VH$  production. The  $Z$  and  $W$  samples are generated using SHERPA 2.2.1 with the same settings as for the  $DY$  sample used for the  $H \rightarrow \mu\mu$  analysis. The diboson samples are simulated using SHERPA 2.2.2 with NLO-accurate matrix elements for up to one parton and LO-accurate matrix elements for up to three partons for the  $q\bar{q}$ -initiated process and LO-accurate matrix elements for up to one parton for the  $gg$ -initiated process.  $Z$ ,  $W$  and diboson samples are matched with the SHERPA parton shower using the MEPS@NLO prescription. The  $t\bar{t}$ , single top-quark and  $t\bar{t}H$  samples are generated using POWHEGBOX with the same settings as in  $H \rightarrow \mu\mu$  analysis. The  $VH$  samples are generated by POWHEGBOX with the same settings as  $H \rightarrow \mu\mu$ , and the  $t\bar{t}V$  sample is generated using MADGRAPH5\_aMC@NLO at NLO with the same settings as in  $H \rightarrow \mu\mu$ . All samples are interfaced with PYTHIA 8 for the parton shower and hadronization with the same corresponding settings as used for  $H \rightarrow \mu\mu$ .

## Chapter 5

### Physics object definitions

In this chapter, the physics objects utilized to characterize and select events for the two analyses will be described. These physics objects are reconstructed from the ID tracks, calorimeter energy deposits, and MS tracks for each event.

#### 5.1 Primary vertex

To enhance the quality of the event reconstruction, the vertices must meet the requirement of having at least two ID tracks with  $p_T$  greater than 500 MeV. The primary vertex of the hard interaction is identified as the vertex with the highest sum of  $p_T^2$  of tracks.

#### 5.2 Jets

Different event topologies require various approaches to reconstruct jets using the anti-kt algorithm [120, 121]. The reconstruction methods include small-radius (small- $R$ ) jets for both the  $H \rightarrow \mu\mu$  and mono- $h(b\bar{b})$  analyses, large-radius (large- $R$ ), and variable-radius (variable- $R$ ) track-jets for mono- $h(b\bar{b})$ .

Small- $R$  jets are reconstructed with  $R = 0.4$  from ‘particle flow’ objects formed by combining ID tracks and calorimeter energy clusters [122]. These jets must have  $|\eta| < 4.5$  and  $p_T > 25$  GeV (for  $H \rightarrow \mu\mu$ ) or  $p_T > 20$  GeV (for mono- $h(b\bar{b})$ ) for  $|\eta| < 2.5$  and  $p_T > 30$  GeV for  $2.5 < |\eta| < 4.5$ . For small- $R$  jets with  $|\eta| < 2.5$  and  $p_T < 60$  GeV, the Jet Vertex Tagger (JVT) algorithm [123] is used to confirm their origin from the primary vertex, by computing a multivariate likelihood from tracking information.

Large- $R$  jets are reconstructed with  $R = 1.0$  from calorimeter energy clusters calibrated using the local hadronic cell weighting (LCW) scheme [42]. This choice of  $R$  captures all jets produced

in the decay of a boosted heavy object, such as a Higgs boson. To mitigate the impact of pile-up, large- $R$  jets are trimmed by removing any  $R = 0.2$  subjets that have less than 5% of the original jet energy [124].

Variable- $R$  track-jets are used to identify subjets that originate from  $b$ -hadrons within large- $R$  jets. These are reconstructed from ID tracks using an  $R$  that shrinks as the proto-jet's  $p_T$  increases [125] and are matched to the large- $R$  jets by ghost association [126]. The  $R$  of variable- $R$  track-jets is set to  $R = 30 \text{ GeV}/p_T$ , with minimum and maximum values of 0.02 and 0.4, respectively. The algorithm can reconstruct separate jets from closely spaced  $b$ -hadrons, such as in highly boosted  $H \rightarrow b\bar{b}$  decays.

Central small- $R$  jets ( $|\eta| < 2.5$ ) and variable- $R$  track-jets containing  $b$ -hadrons ( $b$ -tagged jets) are identified using the ‘MV2c10’ [127, 128] (used in  $H \rightarrow \mu\mu$  for central small- $R$  jets) or ‘DL1’ tagger (used in mono- $h(b\bar{b})$  for both central small- $R$  jets and variable- $R$  track-jets) [129]. In mono- $h(b\bar{b})$ , the  $b$ -tagging working point (WP), which corresponds to a 77% efficiency (77% WP) in  $t\bar{t}$  events, is used for both central small- $R$  jets and variable- $R$  track-jets. In  $H \rightarrow \mu\mu$ , two  $b$ -tagging WPs, which correspond to a 60% (60% WP) and a 85% (85% WP) efficiency in  $t\bar{t}$  events, are used for central small- $R$  jets.

Since the decays of  $b$ -hadrons can produce muons, which are vetoed when building particle-flow objects and excluded from the energies of either the small- $R$  or large- $R$  jets, in mono- $h(b\bar{b})$  the four-momenta of non-isolated muons falling inside the  $b$ -tagged jet cones can be added to correct for the total jet momentum. The correction is done for the muon (two muons) closest to the jet axis of small- $R$  (large- $R$ ) jets. This correction is used in mono- $h(b\bar{b})$  to calculate the mass of the Higgs boson candidate ( $m_h$ )

### 5.3 Leptons

Muon reconstruction is performed by combining tracks in the ID and MS. For  $H \rightarrow \mu\mu$ , additional muon candidates are considered to improve the efficiency. In the region of  $|\eta| < 0.1$ , muon candidates are identified by matching a reconstructed ID track to either an MS track segment or a calorimetric energy deposit consistent with a minimum-ionizing particle. In the region  $2.5 < |\eta| < 2.7$ , where the ID does not cover, additional muons are reconstructed from an MS track with hits in the three MS layers and combined with forward ID hits.

Muon candidates must satisfy the ‘Loose’ or ‘Medium’ criteria defined in Ref. [130], depending on the analysis. The criteria require a transverse momentum  $p_T > 6$  GeV and  $|\eta| < 2.7$ . Muons with an associated ID track must also be matched to the primary vertex with  $z_0 \sin \theta < 0.5$  mm and  $|d_0|/\sigma(d_0) < 3$ , where  $z_0$  is the longitudinal impact parameter,  $\theta$  is the polar angle of the track,  $d_0$  is the transverse impact parameter calculated relative to the measured beam-line position, and  $\sigma(d_0)$  is the uncertainty on  $d_0$ . Isolation criteria [131] are imposed to suppress non-prompt muons produced from hadron decays. The isolation criteria utilize ID track and calorimeter energy deposit information in a range of  $\Delta R < 0.2$  around the muon.

Muons can lose a significant amount of energy through QED final-state-radiation (FSR), which can reduce the signal dimuon mass and deteriorate the  $m_{\mu\mu}$  resolution. To improve the resolution of the signal kinematic spectrum for  $H \rightarrow \mu\mu$ , up to one final-state photon candidate is included as part of the dimuon system for each event. Photon candidates are reconstructed [131] from energy cluster found in the EM calorimeter with the requirement that  $f_1 > 0.2$ , where  $f_1$  is the energy fraction of the cluster in the first layer of the EM calorimeter. Additionally, only photons with  $\Delta R(\gamma, \mu) < 0.2$  are considered. To reduce background from pile-up interactions, photons are also required to pass a  $p_T^\gamma$  threshold, which increases linearly from 3 GeV at  $\Delta R = 0$  to 8 GeV at  $\Delta R = 0.2$ . If there are more than one photon candidates, the photon with the highest  $p_T^\gamma$  is selected. About 5% of the events include a QED FSR photon, and the total width of the signal  $m_{\mu\mu}$  spectrum is reduced by approximately 3%.

Electrons are reconstructed by matching clusters of energy in the EM calorimeter to ID tracks. The identification criteria for electrons include the ‘Medium’ criterion used in  $H \rightarrow \mu\mu$  and the ‘Loose’ criterion used in mono- $h(b\bar{b})$  [132]. Additionally, electrons must have  $p_T > 7$  GeV (used in both  $H \rightarrow \mu\mu$  and mono- $h(b\bar{b})$ ), or 27 GeV (used only in mono- $h(b\bar{b})$ ),  $|\eta| < 2.47$  and not  $1.37 < |\eta| < 1.52$ . Like muons, electrons must be isolated and be matched to the primary vertex with  $z_0 \sin \theta < 0.5$  mm and  $|d_0|/\sigma(d_0) < 5$ .

The reconstruction of hadronically decaying  $\tau$ -leptons is initiated from small- $R$  jets created using LCW-calibrated clusters [133]. As hadronic  $\tau$ -lepton decays can produce either one or three charged pions, the jets are required to have one or three tracks within  $\Delta R = 0.2$  of the jet axis. A recurrent neural network (RNN) classifier is employed to identify the  $\tau$ -lepton [134], with inputs

constructed from the clusters and tracks associated with the  $\tau$ -lepton. All  $\tau$ -leptons are required to pass the ‘VeryLoose’ WP [134] and have  $|\eta| < 2.4$  and  $p_T > 20$  GeV.

## 5.4 Missing transverse momentum

Particles that escape from the detector, such as neutrinos and Dark Matter, result in missing transverse momentum  $E_T^{\text{miss}}$ .  $E_T^{\text{miss}}$  is the magnitude of the negative vector sum of the transverse momenta of all physics objects, including muons, electrons, and jets, as well as the ID tracks not linked to any physics objects (soft term).

## 5.5 Overlap removal

To prevent the same detector signals from being interpreted as different objects, an overlap removal procedure is implemented. If any object is rejected at one stage, it is not considered in subsequent stages. The procedure proceeds as follows: first, if two electrons share a track, the electron with lower  $p_T$  is discarded. Next, any  $\tau$ -lepton within  $\Delta R = 0.2$  of an electron or muon is removed. Then, any electron that shares a track with a muon is removed. Any small- $R$  jet is within  $\Delta R = 0.2$  of an electron is then discarded, followed by any electron within a cone of  $p_T$ -dependent size around a small- $R$  jet. If a small- $R$  jet with less than three tracks is linked to a muon or is within  $\Delta R = 0.2$  of one, it is removed, and any muon within a cone of  $p_T$ -dependent size around a small- $R$  jet is discarded. Next, any small- $R$  jet within  $\Delta R = 0.2$  of a  $\tau$ -lepton is removed. Finally, any large- $R$  jet within  $\Delta R = 1.0$  of an electron is removed. Track-jets are not included in the overlap removal process as they are solely employed for  $b$ -tagging.

## Chapter 6

### Search for $H \rightarrow \mu\mu$

#### 6.1 Event selections and categorization

To select events in the interesting kinematic phase space for  $H \rightarrow \mu\mu$  signals, a number of event selections are applied. Firstly, events must contain a primary vertex and at least a pair of opposite-charge muons. The leading muon must have  $p_T > 27$  GeV, while the subleading muon must have  $p_T > 15$  GeV in all categories except for the  $VH$  3-lepton category (defined in the following subsections). Further selections are required in regions targeting different production modes, as detailed in the following subsections. Tab. 6.1 summarizes the main event selections, and Fig. 6.1 shows the distribution of selected events in the  $m_{\mu\mu}$  spectrum. The final statistical fitting is performed in the range of  $110 < m_{\mu\mu} < 160$  GeV, where  $m_{\mu\mu}$  is the invariant mass of the dimuon system. Approximately 450,000 data events are selected in this  $m_{\mu\mu}$  range, and the mass window  $120 < m_{\mu\mu} < 130$  GeV is expected to contain about 868  $H \rightarrow \mu\mu$  signal events, which corresponds to about 85% of the total selected signals. The total efficiency times acceptance of about 52% is with respect to all  $H \rightarrow \mu\mu$  with  $ggF$ , VBF,  $VH$ , and  $t\bar{t}H$  productions.

The selected events are classified into 20 categories. The  $t\bar{t}H$  category targets the  $t\bar{t}H$  production mode of the  $H \rightarrow \mu\mu$  signals, three  $VH$  categories target  $VH$ , four VBF categories target VBF, and twelve  $ggF$  categories target  $ggF$ . The categories are defined by various kinematic variables and boosted decision trees (BDT) using the XGBoost package. The categorization procedure separates events into regions with different signal-to-background ratios, enhancing the total signal sensitivity.

The remainder of the section describes the general strategy of BDT training and optimization of the BDT score boundaries between each category, as well as details the categorization for  $t\bar{t}H$ ,  $VH$ , VBF, and  $ggF$  categories.

Selection	
	Primary vertex
Common preselection	Two opposite-charge muons Muons: $ \eta  < 2.7, p_T^{\text{lead}} > 27 \text{ GeV}, p_T^{\text{sublead}} > 15 \text{ GeV}$ (except $VH$ 3-lepton)
Fit Region	$110 < m_{\mu\mu} < 160 \text{ GeV}$
Jets	$p_T > 25 \text{ GeV}$ and $ \eta  < 2.4$ or with $p_T > 30 \text{ GeV}$ and $2.4 <  \eta  < 4.5$
$t\bar{t}H$ Category	at least one additional $e$ or $\mu$ with $p_T > 15 \text{ GeV}$ , at least one $b$ -jet (85% WP)
$VH$ 3-lepton Categories	$p_T^{\text{sublead}} > 10 \text{ GeV}$ , one additional $e$ ( $\mu$ ) with $p_T > 15(10) \text{ GeV}$ , no $b$ -jets (85% WP)
$VH$ 4-lepton Category	at least two additional $e$ or $\mu$ with $p_T > 8, 6 \text{ GeV}$ , no $b$ -jets (85% WP)
$ggF+VBF$ Categories	no additional $\mu$ , no $b$ -jets (60% WP)

Table 6.1: Summary of the main event selection criteria applied to all events as well as the the selection of jets. The bottom sections give the basic requirements on leptons and  $b$ -tagged jets for the analysis categories targeting different Higgs boson production modes. The subleading muon momentum threshold is 15 GeV in all categories except the  $VH$  3-lepton categories, where it is lowered to 10 GeV.

### 6.1.1 General BDT training strategy

The overall approach is to employ a BDT as a classifier to differentiate between the targeted signal and the background. These BDTs take a reduced set of kinematic variables as input to minimize unnecessary systematic uncertainties while maintaining a comparable performance.

All samples are initially weighted by their cross-sections and then normalized such that the mean weights of the training signal and background samples are equal to 1. To prevent bias from tuning on the test samples, we adopt a four-fold training method (outlined in Tab. 6.2). The signal and background samples are first divided into four sections based on the remainder of event numbers divided by 4. In each fold of the training, two of the four sections (50% of the full samples) are used for the training set, one of the four sections (25% of the full samples) is utilized as the validation set, and the remaining section (25% of the full samples) is employed for the test set. The

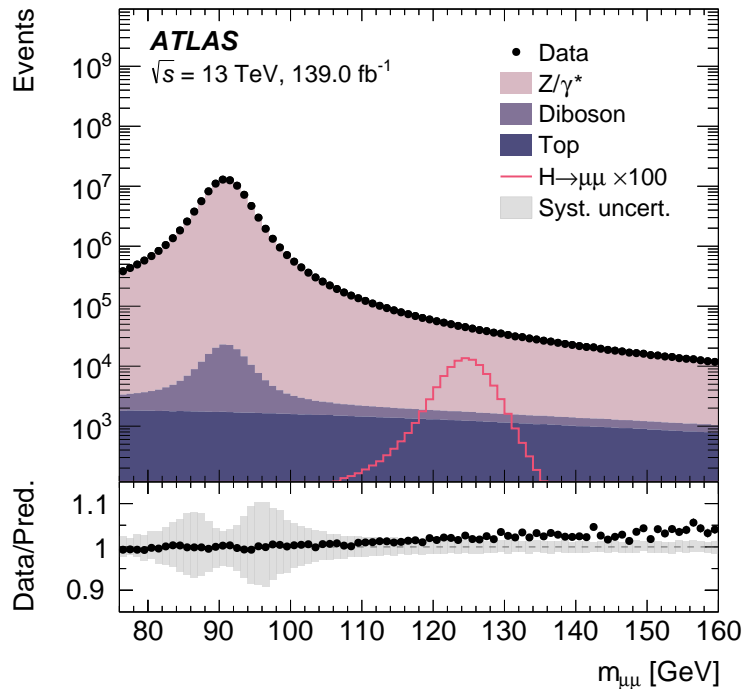


Figure 6.1: Dimuon invariant mass  $m_{\mu\mu}$  in the range 76–160 GeV for all events passing the event selections. Data in points are compared to a full set of fully simulated background processes and the total background prediction is scaled to the integrated data yield. The shaded region in the bottom panel shows the impact of the systematic uncertainty on the muon momentum scale, resolution and muon trigger and reconstruction efficiencies. The  $H \rightarrow \mu\mu$  signal shown is the sum of the ggF, VBF, VH and ttH as open line, normalised to one hundred times the SM prediction for visibility.

training, validation, and test sets are rotated among the four folds of the training. The hyperparameters of the BDT are optimized by maximizing the ROC AUC on the validation set using Bayesian Optimization [135]. After the training, the BDT for each fold is used to classify the corresponding test set.

To ensure a similar shape of the BDT distribution in each fold of the training, the BDT outputs in each fold are transformed separately such that the unweighted signal sample in each fold has a uniform BDT distribution. The QUANTILETRANSFORMER function of the scikit-learn package [136] is used to perform this transformation of the BDT outputs.

Event number % 4	0	1	2	3
Fold 0	Test set	Validation set	Trainin set	Training set
Fold 1	Training set	Test set	Validation set	Training set
Fold 2	Trainin set	Training set	Test set	Validation set
Fold 3	Validation set	Trainin set	Training set	Test set

Table 6.2: Summary of four-fold training method.

### 6.1.2 General strategy of BDT boundary optimization

The BDT boundaries are optimized simultaneously to maximize the total number counting significance, denoted as  $Z_{\text{tot}}$ :

$$Z_{\text{tot}} = \sqrt{\sum_i Z_i^2}, \quad (6.1)$$

where  $Z_i$  is the number counting significance in each BDT category and is calculated as:

$$Z_i = \sqrt{2 \left( (s_i + b_i) \log \left( \frac{s_i + b_i}{b_i} \right) - s_i \right)}, \quad (6.2)$$

where  $s_i$  is the signal yield inside the dimuon mass window,  $120 \leq m_{\mu\mu} \leq 130$  GeV, and  $b_i$  the background yield inside the dimuon mass window extrapolated from the yield of the data sideband,  $110 \leq m_{\mu\mu} \leq 120$  GeV or  $130 \leq m_{\mu\mu} \leq 180$  GeV, by a scale factor of 0.2723. The scale factor is estimated from the inclusive observed dataset by taking the ratio between the center (dimuon mass 120-130 GeV) and the sideband (dimuon mass 110-180 excluding 120-130 GeV) data in the inclusive region of the dimuon selections. To ensure sufficient statistics for the fitting in the later stage, each BDT category requires at least five background events within the dimuon mass window  $120 \leq m_{\mu\mu} \leq 130$  GeV.

During the stage of optimizing the categorization, only the “test set score” is used. Specifically, the BDT score of each event used at this stage is obtained from a BDT model that is completely independent of the BDT models that have taken that event as an input (training event). Therefore, any bias from the training or hyperparameter tuning will not show up in the samples used for the optimization. The same samples used for optimizing the BDT boundaries are also used for the

final sensitivity evaluation. A validation using a 4-fold categorization method has been performed and showed a negligible difference in the sensitivity.

### 6.1.3 $t\bar{t}H$ category

The  $t\bar{t}H$  category aims to identify the  $t\bar{t}H$  production mode of the Higgs boson, where the  $t\bar{t}$  pair undergoes dileptonical or semileptonic decay. To be included in this category, events must have at least 1 lepton (either  $e$  or  $\mu$ ) with  $p_T > 15$  GeV, in addition to the opposite-charge muon pair, and at least 1  $b$ -jet tagged at the 85% WP. The two highest- $p_T$  opposite-charge muons are used to reconstruct the Higgs boson and to calculate the dimuon mass  $m_{\mu\mu}$ .

A BDT is trained using simulated  $t\bar{t}H \rightarrow \mu\mu$  events as signal and simulated SM background processes as background, with a preselection of  $100 < m_{\mu\mu} < 200$  GeV. The BDT uses 12 kinematic variables as input features, which are listed below:

- $p_T^{\mu\mu}$ : transverse momentum of the reconstructed Higgs boson (dimuon system).
- $\cos\theta^*$ : cosine of the lepton decay angle in the Collins-Soper frame.
- $p_T^{\ell_{1(2)}}$ : transverse momenta of the additional leptons.
- $n_{j_c}$ : number of central jets ( $|\eta| < 2.5$ ).
- $n_b$ : number of  $b$ -jets.
- $H_T$ : scalar sum of the transverse momenta of all jets.
- $m_{\text{Lep-}t}$ : transverse mass of the leptonic top-quark system, consisting of the third lepton,  $E_T^{\text{miss}}$  and a  $b$ -jet. If an event contains more than two  $b$ -jets, the  $b$ -jet is selected to minimize the difference between the reconstructed value and the SM value for  $m_t$  and  $m_W$ .
- $m_{\text{Lep-}W}$ : transverse mass of the leptonic  $W$ -boson system, consisting of the third lepton and  $E_T^{\text{miss}}$ .
- $m_{\text{Had-}t}$ : invariant mass of the hadronic top-quark system, consisting of a  $b$ -jet and two non- $b$ -jets. If an event contains more than one possible combination of the hadronic top-quark

system, the combination is chosen by minimizing the  $\chi$ -square differences between the reconstructed value and the SM value for  $m_t$  and  $m_W$ , where the reconstructed  $m_W$  value is calculated from the system composed of two non- $b$ -jets.

- $m_{\ell\ell}$ : invariant mass of the two additional leptons, where the two leptons must have opposite signs and the same flavor.
- $m_{\mu\mu_3}$ : invariant mass of the system composed of the third-highest  $p_T$  muon and one of the two leading muons with opposite charge to the third-highest  $p_T$  muon.

Note that if an event does not contain enough objects to define the variables described above, the variables will be assigned with an unphysical arbitrary value. Fig. 6.2 and 6.3 show the distributions of the training variables used in the  $t\bar{t}H$  category. The distributions of the  $t\bar{t}H$  BDT score for signal and background are shown in Fig. 6.4.

The  $t\bar{t}H$  category is defined based on the BDT score, with the dominant backgrounds including  $t\bar{t}Z$  process,  $t\bar{t}$  and dibosons. Assuming the SM, the expected signal yield in this category is 1.2, with a high purity of 98% for the  $t\bar{t}H$  production relative to the other Higgs boson production modes. In the mass window  $120 < m_{\mu\mu} < 130$  GeV, the signal-to-background ratio is 8% .

#### 6.1.4 $VH$ categories

The  $VH$  categories aim to select events in the  $VH$  production mode, where the vector boson decays leptonically ( $W \rightarrow \ell\nu$  or  $Z \rightarrow \ell\ell$ ). These events must not have any  $b$ -jets tagged at the 85% WP and should have at least one lepton (muon or electron) in addition to the opposite-charge muon pair. The events are further categorized into the 3-lepton channel and 4-lepton channel.

In the 3-lepton channel, events are required to have exactly three leptons, including the sub-leading muon with  $p_T > 10$  GeV and the additional muon (electron) with  $p_T > 10$  (15) GeV. To reduce the main background  $Z \rightarrow \mu\mu$  events, events containing an opposite-charge muon pair with an invariant mass between 80 and 105 GeV are vetoed.

In the 4-lepton channel, events must have at least four leptons, and two additional muons or electrons are required to have  $p_T > 8$  and 6 GeV, respectively. There should be at most one opposite-charge muon pair with an invariant mass between 80 and 105 GeV.

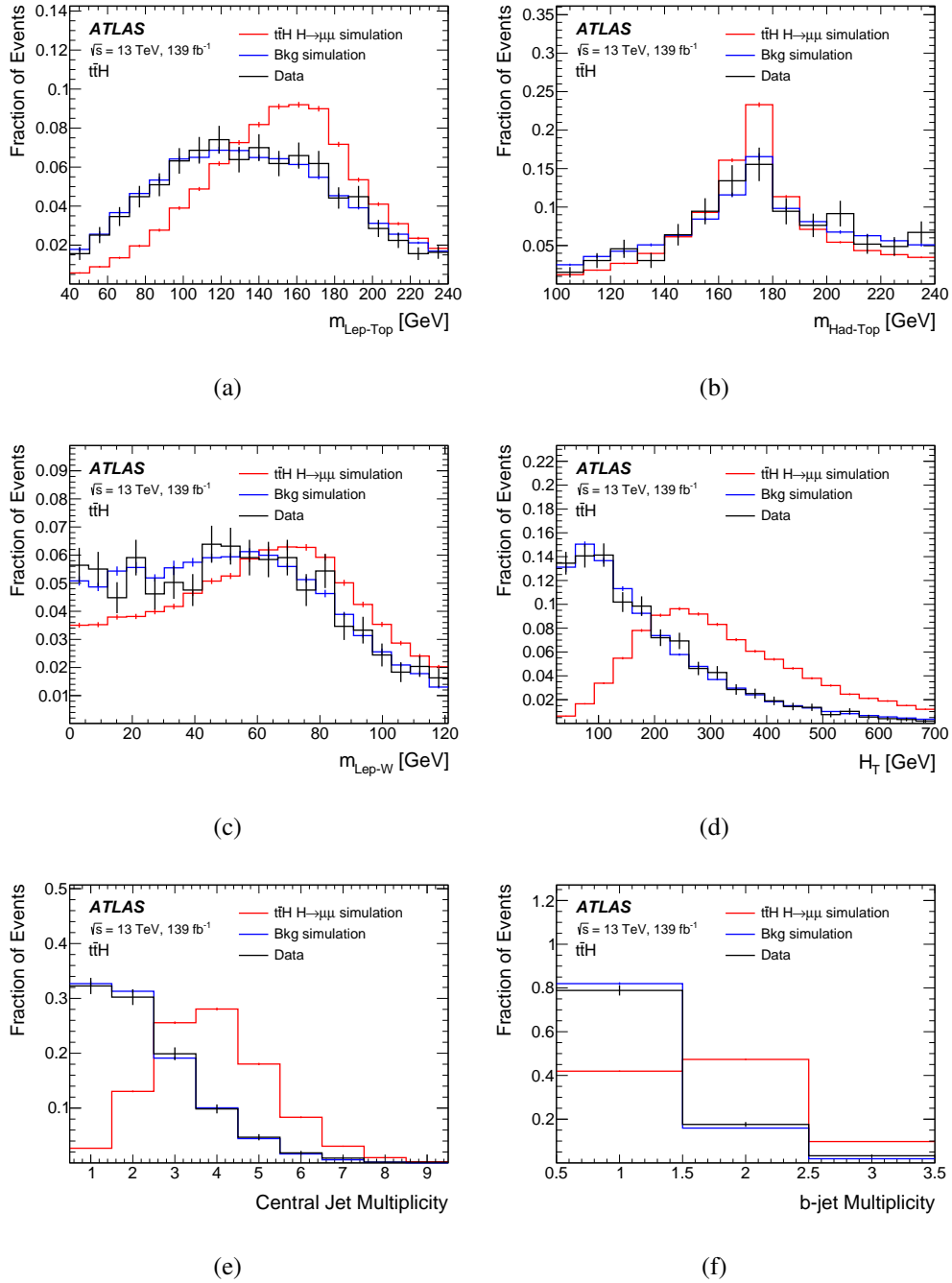


Figure 6.2: Distributions of each training variable used for the  $t\bar{t}H$  category (part 1).

If there are more than two possible combinations of opposite-charge muon pairs, the assignments of the muons to either  $H \rightarrow \mu\mu$  or  $Z \rightarrow \mu\mu$  are made based on the minimization of the  $\chi^2$  criterion. The criterion considers the difference between the reconstructed and the expected

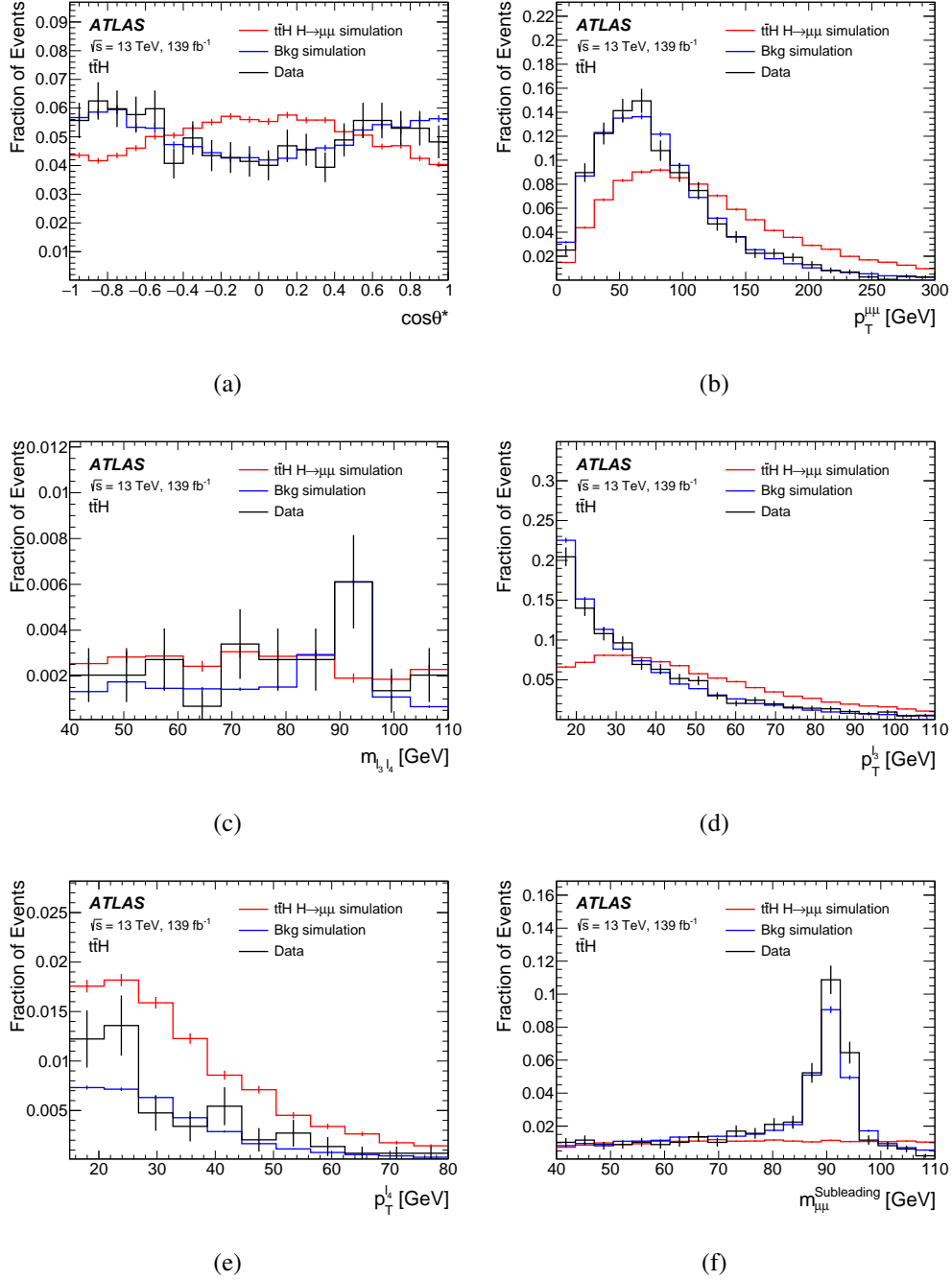


Figure 6.3: Distributions of each training variable used for the  $t\bar{t}H$  category (part 2).

masses of  $H$  and  $Z$  with respect to the expected experimental resolutions. The formula for  $\chi^2$  is shown below:

$$\chi^2 = \begin{cases} \frac{(m_{\mu_1\mu_2} - m_H)^2}{\sigma(m_H)^2} + \frac{(m_{\mu_3, MET} - m_W)^2}{\sigma(m_W)^2}, & \text{if } n_\mu = 3 \\ \frac{(m_{\mu_1\mu_2} - m_H)^2}{\sigma(m_H)^2} + \frac{(m_{\mu_3\mu_4} - m_Z)^2}{\sigma(m_Z)^2}, & \text{if } n_\mu \geq 4 \end{cases}, \quad (6.3)$$

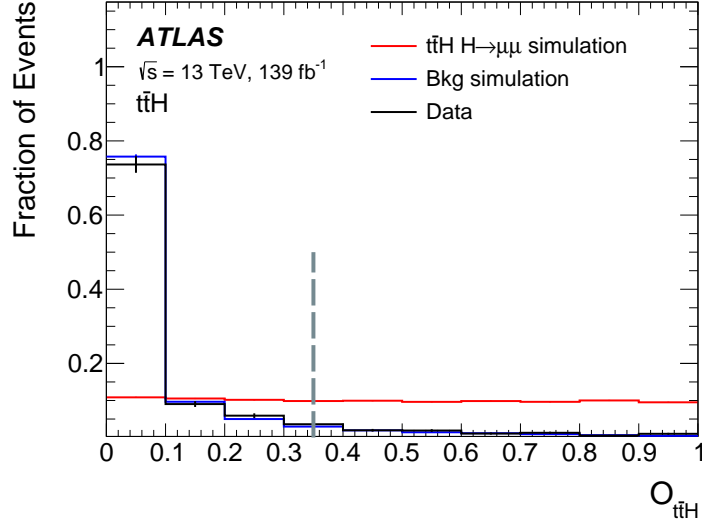


Figure 6.4: The  $t\bar{t}H$  BDT score distributions for signal and background. The vertical grey dashed line represents the BDT score threshold of 0.35 used to select events in the  $t\bar{t}H$  category.

where  $m_{\mu_1\mu_2}$  is the invariant mass of the opposite-charge dimuon system forming  $H \rightarrow \mu\mu$  (also referred to as  $m_{\mu\mu}$ ),  $m_{\text{T}}^{(\mu_3, MET)}$  is the transverse mass of the system consisting of the additional muon and  $E_{\text{T}}^{\text{miss}}$ , and  $m_{\mu_3\mu_4}$  is the invariant mass of the additional opposite-charge dimuon system forming  $Z \rightarrow \mu\mu$ . Using this pairing algorithm achieves an accuracy of about 93% and 97% in the 3-lepton and 4-lepton channels, respectively. If the additional leptons are electrons, they are matched to either  $W \rightarrow e\mu$  or  $Z \rightarrow ee$  without ambiguity.

A BDT is trained in the 3-lepton (4-lepton) channel in the range of  $110 < m_{\mu\mu} < 160$  GeV, using simulated  $WH \rightarrow \mu\mu$  ( $ZH \rightarrow \mu\mu$ ) as the training signal and simulated Standard Model (SM) background as the training background. In the 3-lepton channel, the nine training variables used for the BDT are:

- $\Delta\phi_{\mu\mu, MET}$ : azimuthal separation between  $H \rightarrow \mu\mu$  and  $E_{\text{T}}^{\text{miss}}$ .
- $p_{\text{T}}^W$ : transverse momentum of the  $W$  boson.
- $p_{\text{T}}^{\ell W}$ : transverse momentum of the  $W$  lepton.
- $m_{\text{T}}^W$ : transverse mass of the  $W$  boson system.
- $\Delta\phi_{\mu\mu, W}$ : azimuthal separation between  $H \rightarrow \mu\mu$  and  $W$ .

- $\Delta\eta_{\mu\mu,W}$ : separation in pseudorapidity between  $H \rightarrow \mu\mu$  and  $W$ .
- $E_T^{\text{miss}}$ : missing transverse momentum.
- $p_T^j$ : transverse momentum of the leading jet if present.
- $n_j$ : number of jets.

In the 4-lepton channel, the eight training variables used for the BDT are:

- $\Delta\phi_{\ell_Z,\ell_Z}$ : azimuthal separation between the two leptons from  $Z \rightarrow \ell\ell$ .
- $m_T^Z$ : invariant mass of the  $Z$  boson system.
- $\Delta\phi_{\mu\mu,Z}$ : azimuthal separation between  $H \rightarrow \mu\mu$  and  $Z$ .
- $\Delta\eta_{\mu\mu,Z}$ : separation in pseudorapidity between  $H \rightarrow \mu\mu$  and  $Z$ .
- $E_T^{\text{miss}}$ : missing transverse momentum.
- $p_T^{j1}$ : transverse momentum of the leading jet if present.
- $p_T^{j2}$ : transverse momentum of the subleading jet if present.
- $n_j$ : number of jets.

Like the  $t\bar{t}H$  category, to handle events that do not have enough objects to define certain variables, an unphysical arbitrary value will be assigned to those variables in the BDT for the  $VH$  category. Fig. 6.5 and 6.6 show the distributions of the training variables used for the  $VH$  3 lepton channels, and Fig 6.7 and 6.8 show the distributions of the training variables used for the  $VH$  4 lepton channels. The score distributions of the BDT for both signal and background in the  $VH$  3-lepton and 4-lepton channels are illustrated in Fig. 6.9.

The BDT scores define three  $VH$  categories, with two categories, “VH3LH” and “VH3LM”, defined in the 3-lepton channel using the 3-lepton BDT and one category, “VH4L”, defined in the 4-lepton channel using the 4-lepton BDT. The dominant background in the  $VH$  categories includes diboson,  $t\bar{t}$ , and DY processes, with the diboson process contributing about 70% (55%) of the total background in the VH3LH (VH3LM) category, while in the VH4L category, about 98% of the

background is from the  $ZZ$  process. The expected signal yield based on the SM assumption is 1.4, 2.8, and 0.5 in the VH3LH, VH3LM, and VH4L categories, respectively. The  $VH$  production is expected to have a signal purity of 89% in the VH3LM category and more than 99% in the VH3LH and VH4L categories relative to other Higgs boson production modes.

### 6.1.5 $ggF$ and VBF categories

The remaining events that are not selected by the  $t\bar{t}H$  and  $VH$  categories are classified into the  $ggF$  and VBF categories. The events are divided into three channels based on jet multiplicities: 0-jet, 1-jet, and 2-jet, where 2-jet includes events with the number of jets  $n_j \geq 2$ . Events containing  $b$ -jets tagged at the 60% efficiency working point or a third muon with  $p_T > 15$  GeV are rejected.

In total four BDTs are trained for the  $ggF$  and VBF categories. One of them (VBF classifier) focuses on differentiating VBF production in 2-jet channel, while the other three ( $ggF$  classifiers) target all  $H \rightarrow \mu\mu$  signals in 0-jet, 1-jet, 2-jet each. For the VBF classifier, VBF  $H \rightarrow \mu\mu$  events with  $n_j \geq 2$  are used as the training signal, while the SM background with  $n_j \geq 2$  is used as the training background. For the  $ggF$  classifiers, all  $H \rightarrow \mu\mu$  signal events with  $n_j = 0, 1, \geq 2$  are used as the training signal, and the SM background with  $n_j = 0, 1, \geq 2$  is used as the training background, for 0-jet, 1-jet, 2-jet, respectively. All of the training events are required to be within the range of  $120 < m_{\mu\mu} < 130$  GeV. Seventeen training variables used for the BDTs are listed below:

In total, four BDTs are trained for the  $ggF$  and VBF categories. One of them (the VBF classifier) focuses on differentiating VBF production in the 2-jet channel, while the other three ( $ggF$  classifiers) target all  $H \rightarrow \mu\mu$  signals in the 0-jet, 1-jet, and 2-jet channels, respectively. For the VBF classifier, VBF  $H \rightarrow \mu\mu$  events with  $n_j \geq 2$  are used as the training signal, while the SM background with  $n_j \geq 2$  is used as the training background. For the  $ggF$  classifiers, all  $H \rightarrow \mu\mu$  signal events with  $n_j = 0, 1$  and  $\geq 2$  are used as the training signal, and the SM background with  $n_j = 0, 1$  and  $\geq 2$  is used as the training background for the 0-jet, 1-jet, and 2-jet channels, respectively. All of the training events are required to be within the range of  $120 < m_{\mu\mu} < 130$  GeV. Seventeen training variables are used for the BDTs and are listed below:

- $p_T^{\mu\mu}$ : transverse momentum of the dimuon system.

- $y_{\mu\mu}$ : rapidity of the dimuon system.
- $\cos\theta^*$ : cosine of the lepton decay angle in the Collins-Soper frame.
- $p_T^{j_{1(2)}}$ : transverse momentum of the leading (subleading) jet if present.
- $\eta_{j_{1(2)}}$ : pseudorapidity of the leading (subleading) jet if present.
- $\Delta\phi_{\mu\mu,j_{1(2)}}$ : azimuthal separation between  $H \rightarrow \mu\mu$  and the leading (subleading) jet if present.
- $N_{\text{track}}^{j_{1(2)}}$ : number of ID tracks with  $p_T > 0.5$  GeV associated with the leading (subleading) jet if present with  $p_T > 50$  GeV and  $|\eta| < 2.1$ .
- $p_T^{jj}$ : transverse momentum of the dijet system formed by the two leading jets if present.
- $y_{jj}$ : rapidity of the dijet system if present.
- $\Delta\phi_{\mu\mu,jj}$ : azimuthal separation between  $H \rightarrow \mu\mu$  and the dijet system if present.
- $m_{jj}$ : invariant mass of the dijet system if present.
- $E_T^{\text{miss}}$ : missing transverse momentum.
- $H_T$ : scalar sum of the transverse momenta of all jets.

It is worth noting again that variables will be assigned an unphysical arbitrary value if an event does not contain enough objects to define such variables. Fig. 6.10, 6.11, 6.12, 6.13 and 6.14 show the distributions of each training variable in the 0-jet, 1-jet and 2-jet channels, respectively. Fig. 6.15 shows the BDT score distributions in the  $ggF$  and VBF categories for both signal and background.

Due to low statistics of sideband data with high VBF classifier output score  $O_{\text{VBF}}$ , a function is fit to the VBF classifier output score distribution of sideband data for  $O_{\text{VBF}} \geq 0.5$  in order to reduce the statistical fluctuation when optimizing the BDT boundaries. The function used for the fit is Epoly2, i.e.  $\exp(a_1 O_{\text{VBF}} + a_2 O_{\text{VBF}}^2)$ . The BDT boundaries for the VBF categories in  $n_j \geq 2$  are then optimized by maximizing the number counting significance based on the fit function. Note that the raw event yields are being used outside the fit range.

In order to reduce statistical fluctuations when optimizing BDT boundaries for the VBF categories in  $n_j \geq 2$ , a function is fit to the VBF classifier output score distribution of sideband data with high VBF classifier output score  $O_{\text{VBF}} \geq 0.5$ . The function used for the fit is Epoly2, i.e.  $\exp(a_1 O_{\text{VBF}} + a_2 O_{\text{VBF}}^2)$ . Raw event yields are used outside the fit range. The BDT boundaries for the VBF categories are then optimized by maximizing the number counting significance based on the fit function.

The events in the 2-jet channel are sorted into four VBF categories (“VBF Very High”, “VBF High”, “VBF Medium” and “VBF Low”) based on the VBF classifier output score. The remaining events in the 2-jet channel are sorted into four 2-jet  $ggF$  categories (“2-jet Very High”, “2-jet High”, “2-jet Medium” and “2-jet Low”) based on the 2-jet  $ggF$  classifier output score. The events in the 1-jet channel are sorted into four 1-jet  $ggF$  categories based on the 1-jet  $ggF$  classifier output score. Similarly, the events in the 0-jet channel are sorted into four 0-jet  $ggF$  categories based on the 0-jet  $ggF$  classifier output score.

### 6.1.6 Summary of categorization

Fig. 6.16 provides a summary of all twenty categories, each of which exhibits a signal-to-background ratio ranging from  $10^{-3}$  to 0.2. The DY process serves as the main background in all  $ggF$  and VBF categories, whereas the diboson (top) process dominates the background in the  $VH$  ( $t\bar{t}H$ ) categories. Notably, each production mode is clearly separated within its corresponding category.

## 6.2 Statistical analysis

To extract the  $H \rightarrow \mu\mu$  signal from the observed data, a binned maximum profile likelihood fit is performed simultaneously in all twenty categories using the  $m_{\mu\mu}$  spectrum [137, 138]. The likelihood function is constructed as a product of Poisson probability functions based on the expected signal and background yields in every category and bin considered in the fit, multiplied by Gaussian or log-normal priors which control the nuisance parameters (NP)  $\theta$ :

$$\prod_{i \in \text{bins}} \text{Pois}(n_i | \mu S_i + B_i) \prod_{\theta_j} P_j(\theta_j), \quad (6.4)$$

where  $n_i$  is the observed number of events,  $S_i$  is the expected signal yield and  $B_i$  is the expected background yield in bin  $i$ .  $\mu$  is the signal strength, representing the ratio of the observed signal yield to the expected signal yield. In other words, the observed signal yield in bin  $i$  is equal to  $\mu S_i$ .  $S_i$  and  $B_i$  are controlled by a set of NPs  $\theta_j$ , which can be correlated or uncorrelated with one another. The prior of the nuisance parameters  $P_j$  is either a Gaussian or a log-normal distribution for the constrained case, and a uniform distribution for the unconstrained case. Both  $S_i$  and  $B_i$  are modeled by analytical functions (functions of  $m_{\mu\mu}$ ) in each category.

where  $n_i$  is the observed number of events,  $S_i$  is the expected signal yield, and  $B_i$  is the expected background yield in bin  $i$ . The signal strength  $\mu$  represents the ratio of the observed signal yield to the expected signal yield. In other words, the observed signal yield in bin  $i$  is equal to  $\mu S_i$ .  $S_i$  and  $B_i$  are controlled by a set of NPs  $\theta_j$ , which can be either correlated or uncorrelated with one another. The prior of the nuisance parameters  $P_j$  is either a Gaussian or a log-normal distribution for the constrained case, and a uniform distribution for the unconstrained case. Both  $S_i$  and  $B_i$  are modeled by analytical functions, which are functions of  $m_{\mu\mu}$ , in each category.

### 6.2.1 Signal modeling

In SM, the Higgs boson is predicted to be a narrow resonance with a width of 4.1 MeV for  $m_H = 125.09$  GeV. The observed signal shape is thus determined by detector resolution effects on the muon momentum measurement. Given the asymmetric feature of the detector resolution effects, we parametrize the signal  $m_{\mu\mu}$  distribution as a double-sided crystal-ball function [139] of  $m_{\mu\mu}$  in each category:

In SM, the Higgs boson is predicted to have a narrow resonance with a width of 4.1 MeV for  $m_H = 125.09$  GeV. As a result of detector resolution effects on the muon momentum measurement, the observed signal shape is parametrized as a double-sided crystal-ball function of  $m_{\mu\mu}$  in

each category. This parametrization is necessary due to the asymmetric nature of detector resolution effects.

$$f_s(m_{\mu\mu}; \mu, \sigma, \alpha_L, n_L, \alpha_R, n_R) = \begin{cases} A_L \cdot (B_L - \frac{x-\mu}{\sigma})^{-n}, & \text{for } \frac{x-\mu}{\sigma} < -\alpha_L \\ \exp(-\frac{(x-\mu)^2}{2\sigma^2}), & -\alpha_L \leq \text{for } \frac{x-\mu}{\sigma} \leq \alpha_R, \\ A_R \cdot (B_R - \frac{x-\mu}{\sigma})^{-n}, & \text{for } \frac{x-\mu}{\sigma} > \alpha_R \end{cases}, \quad (6.5)$$

where

$$A_{L/R} = \left( \frac{n_{L/R}}{|\alpha_{L/R}|} \right)_{L/R}^n \cdot \exp\left(-\frac{|\alpha_{L/R}|^2}{2}\right), \quad (6.6)$$

$$B_{L/R} = \frac{n_{L/R}}{|\alpha_{L/R}|} - |\alpha_{L/R}|,$$

where the signal shape parameters  $\mu$ ,  $\sigma$ ,  $\alpha_L$ ,  $n_L$ ,  $\alpha_R$  and  $n_R$  are determined and fixed to a constant when statistical fitting is performed.

The parametrization is done inclusively for all production models since no significant difference is found between the signal shapes of different production modes. To determine the signal shape parameters,  $f_s$  is fitted to the signal MC spectra summed over all production modes, assuming the relative normalizations as predicted by the SM. Examples of signal parametrization in ‘‘VBF Very High’’ and ‘‘2-jet Low’’ are shown in Fig. 6.17. The Gaussian width varies between 2.6 and 3.2 GeV in all categories. To test for potential biases in the extracted signal yields due to the analytic parameterizations, a signal injection procedure is used. In a signal-plus-background fit to pseudo-data constructed from the expected signal and background distributions, the extracted signal yields agree with those injected within the statistical accuracy of about 0.3%.

The uncertainties in signal modeling are represented by constrained nuisance parameters  $\theta_j$ . Both theoretical and experimental sources of uncertainties are considered. Theoretical uncertainties in signal production impact the number of expected signal events in each category. For the main production modes,  $ggF$  and VBF, uncertainties include missing higher-order QCD corrections, PDFs, underlying event, and hadronization. Uncertainties in  $ggF$  signal are derived using an approach described in Ref. [140] that considers QCD scales, renormalization, resummation, and migration between jet-multiplicity regions [141–148]. Uncertainties in  $ggF$  Higgs boson transverse momentum, migration between different kinematic regions, and treatment of the top-quark

mass in loop corrections are also accounted for. Additionally, uncertainties are assigned for  $ggF$  signal acceptance in VBF topologies. Uncertainties in predicted SM branching ratio and Higgs boson production cross sections are included according to Ref. [52]. Modeling uncertainties related to underlying event and parton showering are estimated by PYTHIA 8 systematic eigentune variations and comparison with HERWIG 7 [149, 150]. The impact of theoretical uncertainties on predicted signal acceptance ranges between a few per mill and 15% for  $ggF$  production and a few per mill and 7% for VBF production. For  $VH$  and  $t\bar{t}H$  categories, the impact ranges between a few per mill and about 18%.

Experimental uncertainties arise from systematic uncertainties related to reconstructed physics objects used in the analysis, affecting expected signal yields in each category. Systematic uncertainties in muon momentum scale and resolution also affect signal  $m_{\mu\mu}$  distribution and shape parameters. Experimental uncertainties include muon reconstruction and identification efficiencies, efficiencies due to trigger, isolation, and impact parameter requirements, muon momentum scale and resolution [44, 130],  $E_T^{\text{miss}}$  soft term determination [151], b-tagging efficiency [128], uncertainty in number of tracks associated with jets [152], pile-up modeling [50], uncertainties in electron reconstruction and identification efficiency [132], jet reconstruction efficiency, energy scale and resolution [153]. The impact of experimental uncertainties on predicted signal yields and modeling in different categories is mainly due to uncertainties in jet energy scale and resolution and muon momentum resolution, affecting signal yields by up to about 10% in some 2-jet categories. Muon momentum resolution uncertainty has an impact on fitted yields ranging between 1% and 6% depending on the category. The uncertainty of 240 MeV in the assumed Higgs mass from [10] is also considered, shifting the signal  $m_{\mu\mu}$  distribution and changing  $\mu$  in Eq. 6.5. Additionally, the uncertainty of 1.7% in the combined 2015-2018 integrated luminosity is accounted for, derived from calibration of the luminosity scale using  $x$ - $y$  beam-separation scans [154], obtained using the LUCID-2 detector [155] for primary luminosity measurements.

### 6.2.2 Background modeling

The background shape is modeled using a core function multiplied by an empirical function. The core function is a physics-motivated rigid shape that focuses on the non-trivial background shape, while the empirical function is a flexible function with free parameters that are determined

by the observed data spectrum. For all categories, the core function is the LO DY line-shape, as described in the appendix of Ref. [24], convolved with the muon resolution, which is parametrized as a Gaussian based on the Drell-Yan full simulation sample. The empirical functional form is selected from power-law function and exponential of polynomials families, as defined in Tab. 6.3. The selection is performed separately for each category based on the ‘‘Spurious signal test’’ and background-only fits.

Function	Expression
PowerN	$m_{\mu\mu}^{(a_0+a_1m_{\mu\mu}+a_2m_{\mu\mu}^2+\dots+a_Nm_{\mu\mu}^N)}$
EpolyN	$\exp(a_1m_{\mu\mu} + a_2m_{\mu\mu}^2 + \dots + a_Nm_{\mu\mu}^N)$

Table 6.3: List of tested empirical functional forms for the background modelling.

The spurious signal test assesses the level of modeling bias by fitting a signal-plus-background model, as defined in Eq. 6.4, to a background-only template. For  $t\bar{t}H$  and  $VH$  categories, the background-only template is the spectrum from the full simulation of all SM background, while for VBF and  $ggF$  categories, the DY fast simulation sample is used instead. The resulting signal from the signal+background fit is the spurious signal, which should be less than 20% of the expected data statistical uncertainties. Background-only fits (Eq. 6.4 with  $\mu$  fixed to 0) are performed on the background-only template and the sideband ( $m_{\mu\mu} < 120$  GeV or  $m_{\mu\mu} > 130$  GeV) of the observed data. To pass the criteria, each  $\chi^2$  p-value from the background-only fits is required to be greater than 1%. To determine the empirical function for each category, we select the function with the least degree of freedom that passes all the criteria. The spurious signal yield is also taken as a modeling systematic uncertainty for each category. We implement it as an additional signal (with the same shape as the actual signal shape) with yield associated with the spurious signal yield. The spurious signal uncertainties are uncorrelated across categories, ranging from a few percent up to about 20% of the expected data statistical uncertainties in the VBF and  $ggF$  categories and up to about 30% in the  $VH$  and  $t\bar{t}H$  categories, which have less statistical precision in their background simulated samples.

### 6.2.3 Results

The analysis performs a simultaneous binned maximum profile likelihood fit over the range of 110-160 GeV for  $m_{\mu\mu}$ , using a bin size of 0.1 GeV. The confidence intervals are determined using the profile-likelihood-ratio test statistic [156]. The best-fit signal strength parameter is  $\mu = 1.2 \pm 0.6$ , corresponding to an observed (expected) significance of  $2.0\sigma$  ( $1.7\sigma$ ) with respect to the null hypothesis of no  $H \rightarrow \mu\mu$  signal. The spectra of the dimuon invariant mass for all analysis categories after the signal-plus-background fit are presented in Fig. 6.18. In Fig. 6.18(b), the events are weighted by  $\ln(1 + S/B)$  to better reflect the total sensitivity to the  $H \rightarrow \mu\mu$  signal, where  $S$  and  $B$  are the observed signal yields and background yields derived from the fit to data in the  $m_{\mu\mu} = 120$ -130 GeV window. The values for  $S$ ,  $B$ , and other key quantities are listed in Tab. 6.4. A goodness-of-fit test is performed using the saturated model technique [157], which returns a probability of 10%. The uncertainty in the signal strength is dominated by the statistical error of  $\pm 0.58$ . The systematic uncertainties on the signal strength have an impact of  $^{+0.18}_{-0.13}$ , with contributions from signal theory uncertainties accounting for  $^{+0.13}_{-0.08}$ , signal experimental uncertainties accounting for  $^{+0.07}_{-0.03}$ , and spurious-signal uncertainties accounting for  $\pm 0.10$ .

A modified frequentist CLs method [156, 158] is used to compute an upper limit on the signal strength parameter  $\mu$  at 95% confidence level (CL). The observed upper limit on  $\mu$  is found to be 2.2, with an expected limit of 1.1 for the absence of an  $H \rightarrow \mu\mu$  signal and an expected limit of 2.0 for an  $H \rightarrow \mu\mu$  signal at SM strength. Assuming the SM cross section for Higgs boson production, the corresponding upper limit on the branching ratio is  $B(H \rightarrow \mu\mu) < 4.7 \times 10^{-4}$  at 95% CL. This result is an improvement of approximately 2.5 times in expected sensitivity compared to the previous ATLAS publication [159]. The increase in sensitivity is mainly due to a larger dataset and more advanced analysis techniques, resulting in an additional 25% improvement.

Category	Data	$S_{\text{SM}}$	$S$	$B$	$S/\sqrt{B}$	$S/B$ [%]	$\sigma$ [ GeV]
VBF Very High	15	$2.81 \pm 0.27$	$3.3 \pm 1.7$	$14.5 \pm 2.1$	0.86	22.6	3.0
VBF High	39	$3.46 \pm 0.36$	$4.0 \pm 2.1$	$32.5 \pm 2.9$	0.71	12.4	3.0
VBF Medium	112	$4.8 \pm 0.5$	$5.6 \pm 2.8$	$85 \pm 4$	0.61	6.6	2.9
VBF Low	284	$7.5 \pm 0.9$	$9 \pm 4$	$273 \pm 8$	0.53	3.2	3.0
2-jet Very High	1030	$17.6 \pm 3.3$	$21 \pm 10$	$1024 \pm 22$	0.63	2.0	3.1
2-jet High	5433	$50 \pm 8$	$58 \pm 30$	$5440 \pm 50$	0.77	1.0	2.9
2-jet Medium	18 311	$79 \pm 15$	$90 \pm 50$	$18 320 \pm 90$	0.66	0.5	2.9
2-jet Low	36 409	$63 \pm 17$	$70 \pm 40$	$36 340 \pm 140$	0.37	0.2	2.9
1-jet Very High	1097	$16.5 \pm 2.4$	$19 \pm 10$	$1071 \pm 22$	0.59	1.8	2.9
1-jet High	6413	$46 \pm 7$	$54 \pm 28$	$6320 \pm 50$	0.69	0.9	2.8
1-jet Medium	24 576	$90 \pm 11$	$100 \pm 50$	$24 290 \pm 100$	0.67	0.4	2.7
1-jet Low	73 459	$125 \pm 17$	$150 \pm 70$	$73 480 \pm 190$	0.53	0.2	2.8
0-jet Very High	15 986	$59 \pm 11$	$70 \pm 40$	$16 090 \pm 90$	0.55	0.4	2.6
0-jet High	46 523	$99 \pm 13$	$120 \pm 60$	$46 190 \pm 150$	0.54	0.3	2.6
0-jet Medium	91 392	$119 \pm 14$	$140 \pm 70$	$91 310 \pm 210$	0.46	0.2	2.7
0-jet Low	121 354	$79 \pm 10$	$90 \pm 50$	$121 310 \pm 280$	0.26	0.1	2.7
VH4L	34	$0.53 \pm 0.05$	$0.6 \pm 0.3$	$24 \pm 4$	0.13	2.6	2.9
VH3LH	41	$1.45 \pm 0.14$	$1.7 \pm 0.9$	$41 \pm 5$	0.27	4.2	3.1
VH3LM	358	$2.76 \pm 0.24$	$3.2 \pm 1.6$	$347 \pm 15$	0.17	0.9	3.0
ttH	17	$1.19 \pm 0.13$	$1.4 \pm 0.7$	$15.1 \pm 2.2$	0.36	9.2	3.2

Table 6.4: Number of events observed in the  $m_{\mu\mu} = 120\text{--}130$  GeV window in data, the number of signal events expected in the SM ( $S_{\text{SM}}$ ), and events from signal ( $S = \mu \times S_{\text{SM}}$ ) and background ( $B$ ) as derived from the combined fit to the data with a signal strength parameter of  $\mu = 1.2$ . The uncertainties in  $S_{\text{SM}}$  correspond to the systematic uncertainty of the SM prediction, the uncertainty in  $S$  is given by that in  $\mu$ , and the uncertainty in  $B$  is given by the sum in quadrature of the statistical uncertainty from the fit and the spurious signal uncertainty. In addition the observed number of signal events divided by the square root of the number of background events ( $S/\sqrt{B}$ ) and the signal-to-background ratio ( $S/B$ ) in % for each of the 20 categories described in the text are displayed.

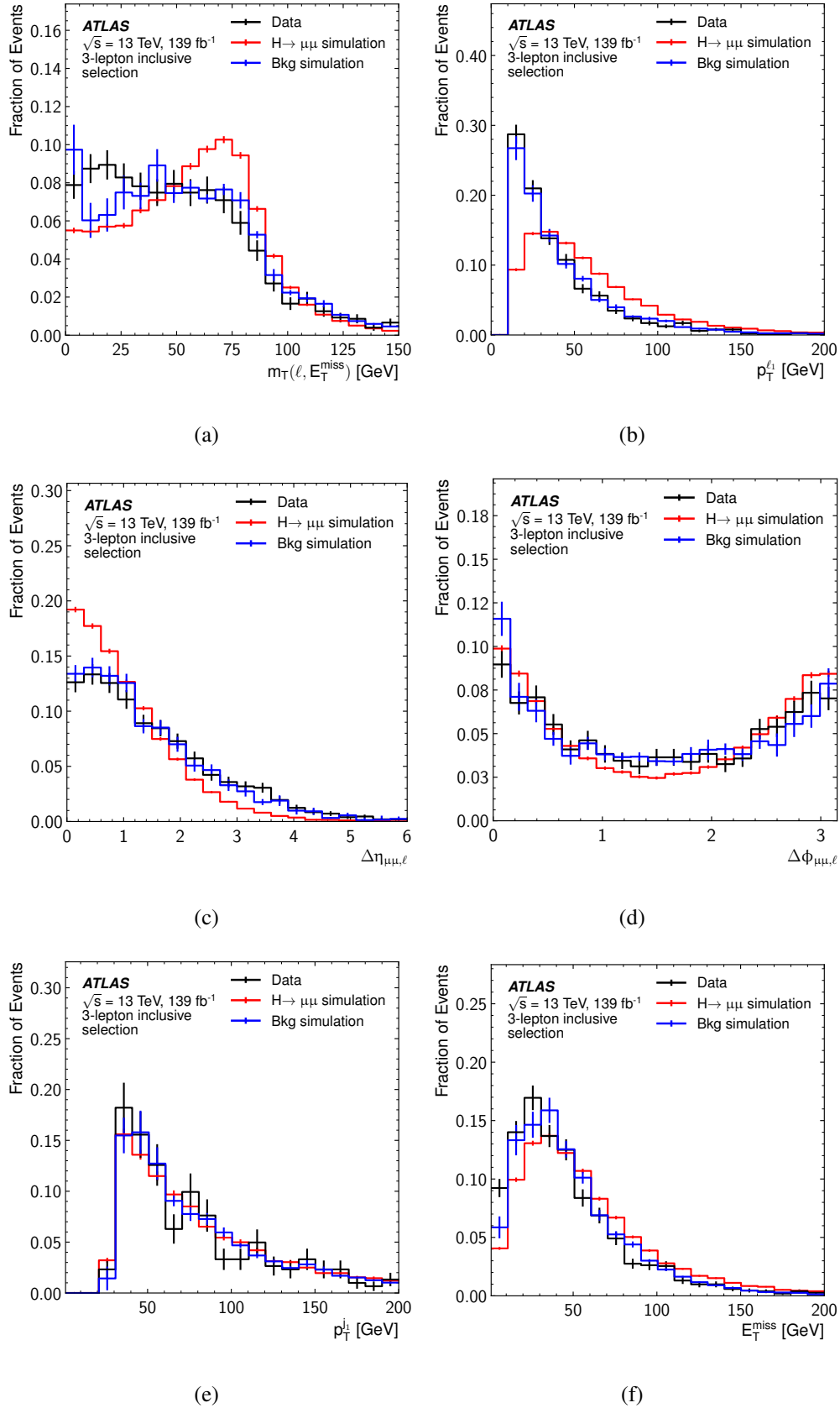
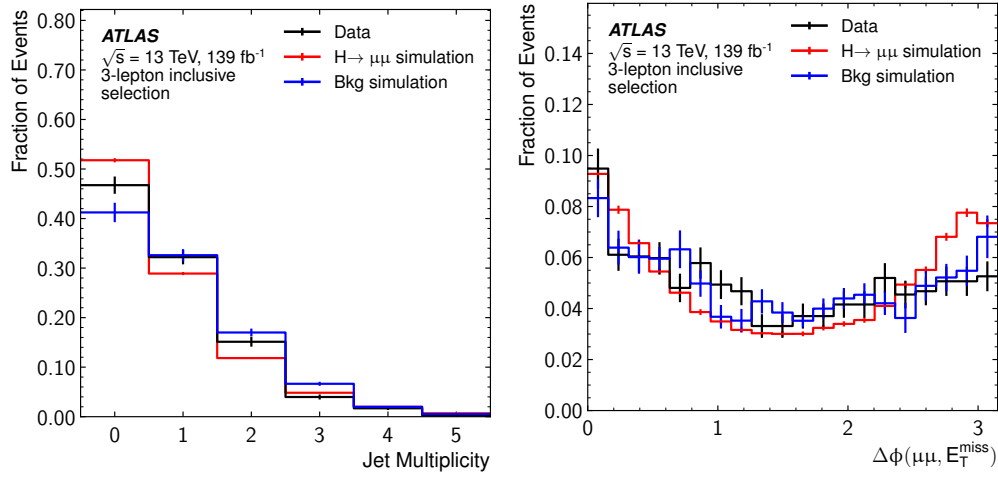
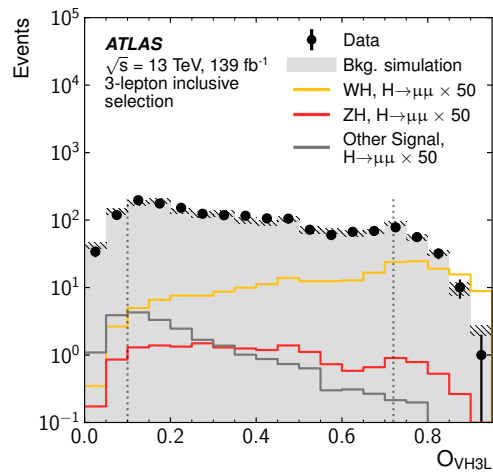


Figure 6.5: Distributions of each training variable used for the  $VH$  3-lepton channel (part 1).



(a)

(b)



(c)

Figure 6.6: Distributions of each training variable used for the  $VH$  3-lepton channel (part 2).

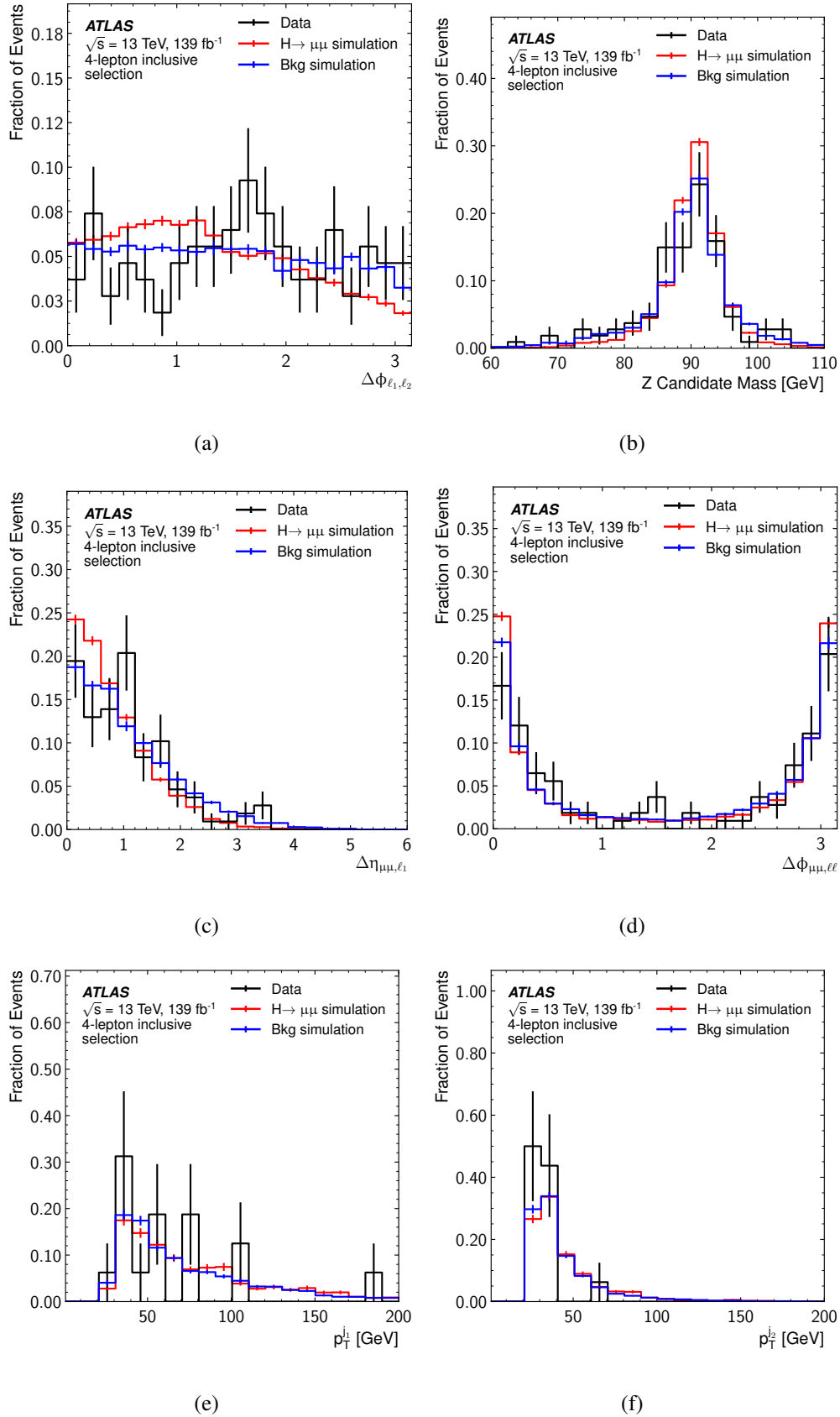


Figure 6.7: Distributions of each training variable used for the  $VH$  4-lepton channel (part 1).

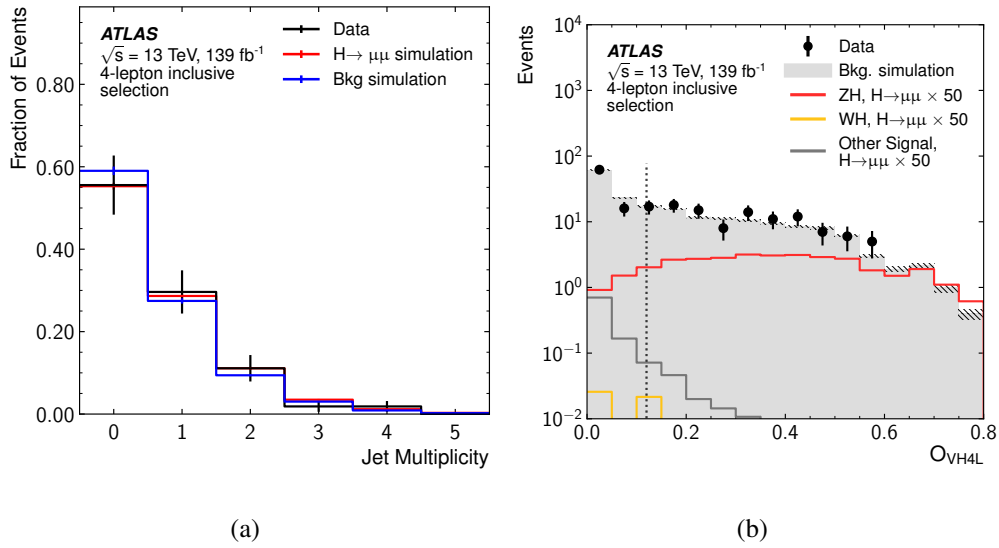


Figure 6.8: Distributions of each training variable used for the  $VH$  4-lepton channel (part 2).

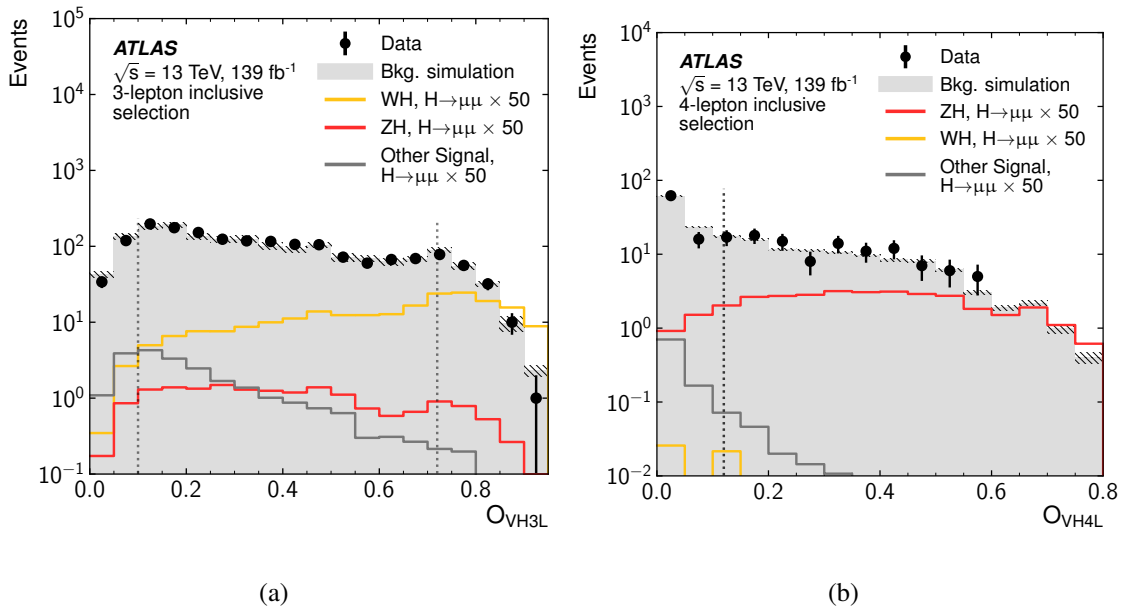


Figure 6.9: The  $VH$  BDT score distributions for signal and background in (a) 3-lepton channel and (b) 4-lepton channel. The grey vertical dashed lines indicate the BDT score boundary define each  $VH$  categories.

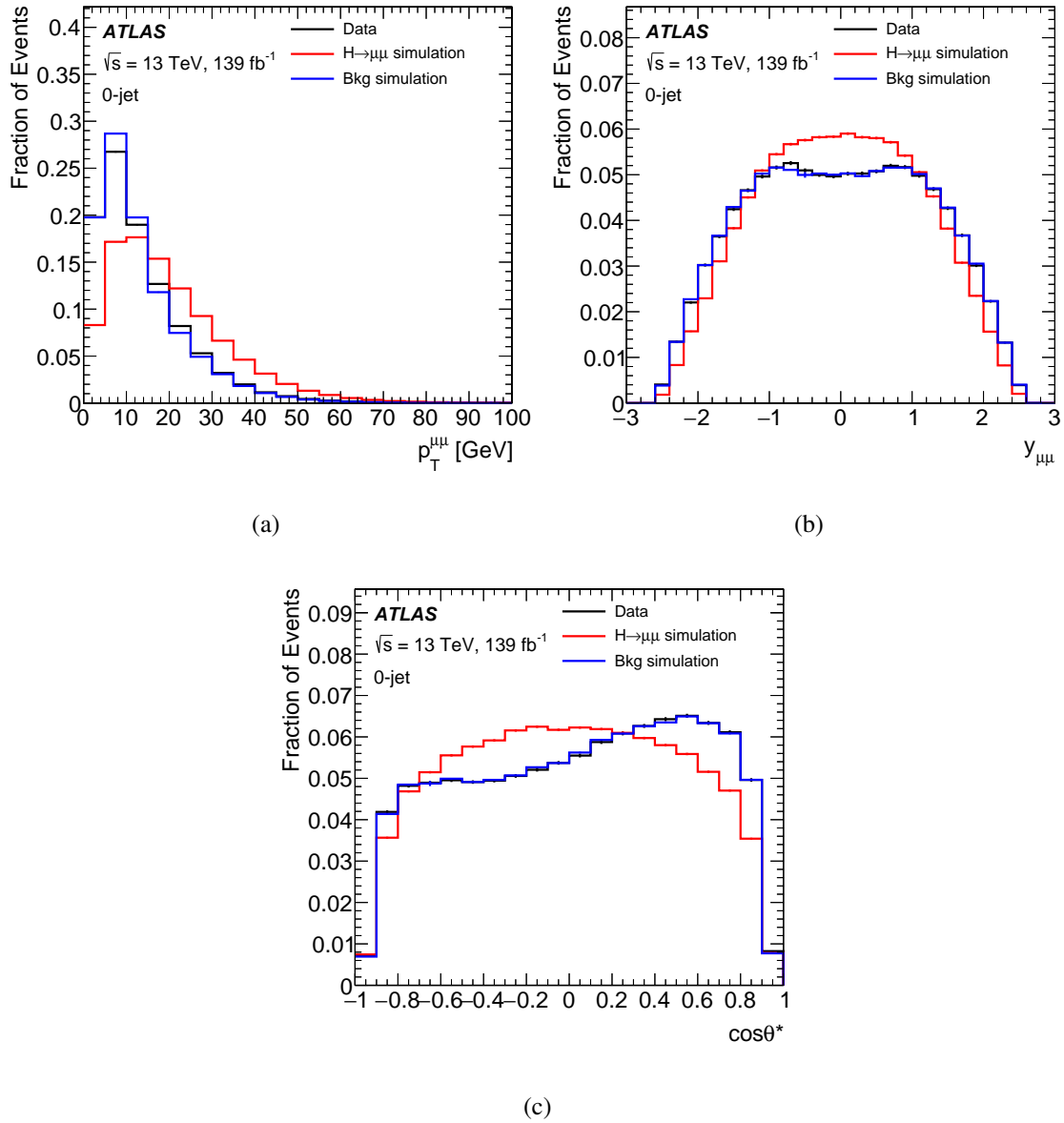


Figure 6.10: Distributions of each training variable in the 0-jet channel.

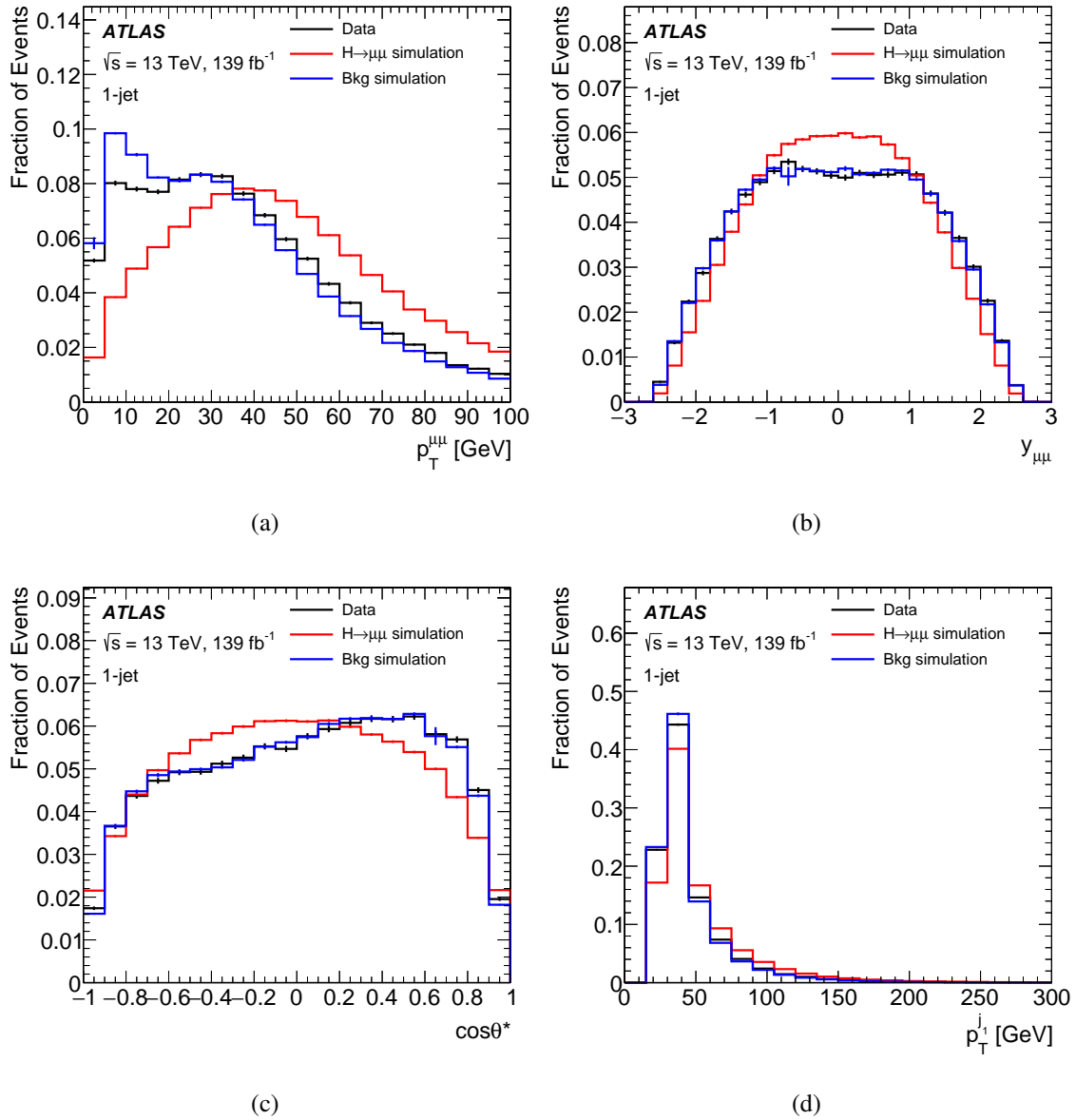


Figure 6.11: Distributions of each training variable in the 1-jet channel (part 1).

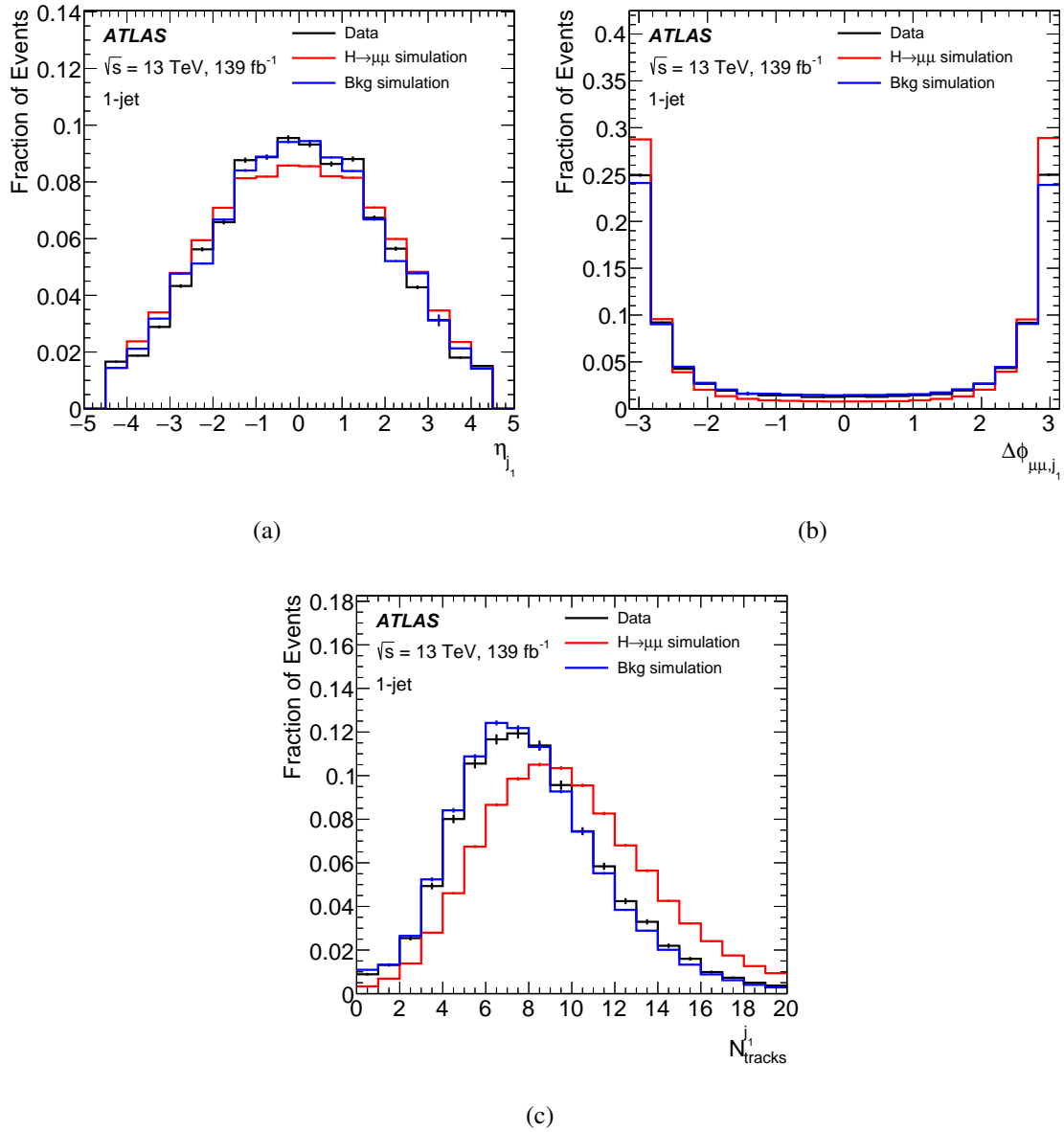


Figure 6.12: Distributions of each training variable in the 1-jet channel (part 2).

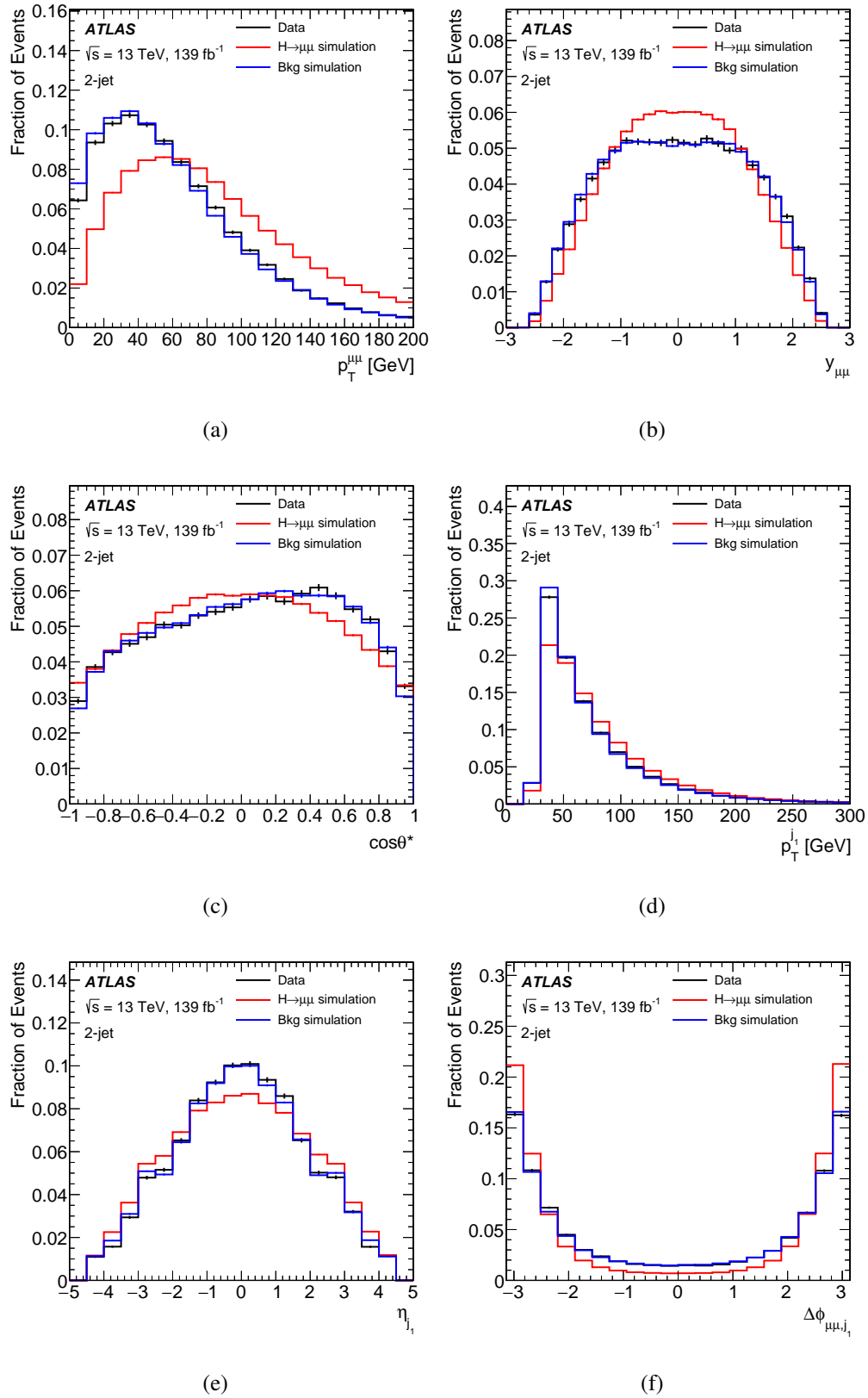


Figure 6.13: Distributions of each training variable in the 2-jet channel (part 1).

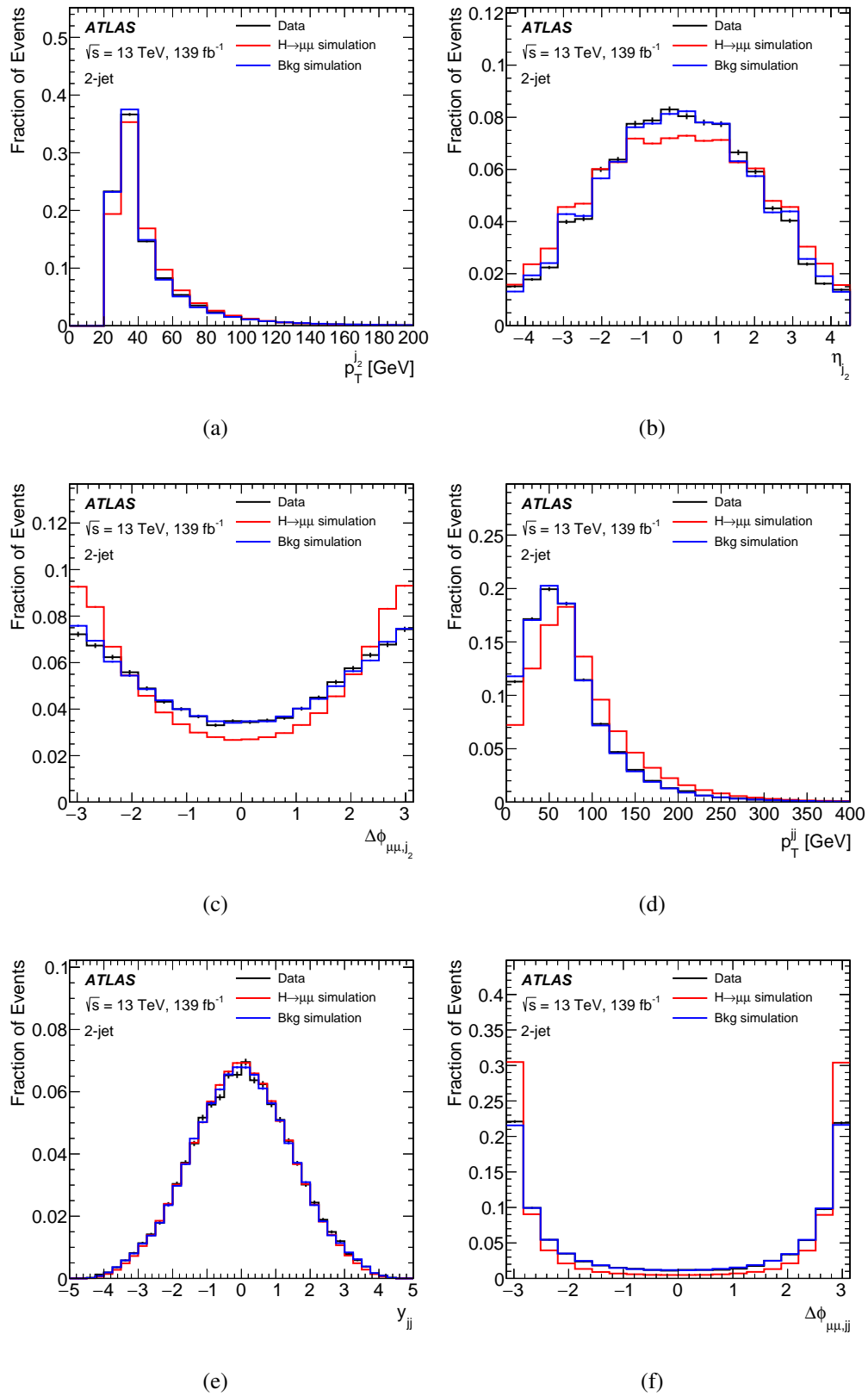


Figure 6.14: Distributions of each training variable in the 2-jet channel (part 2).

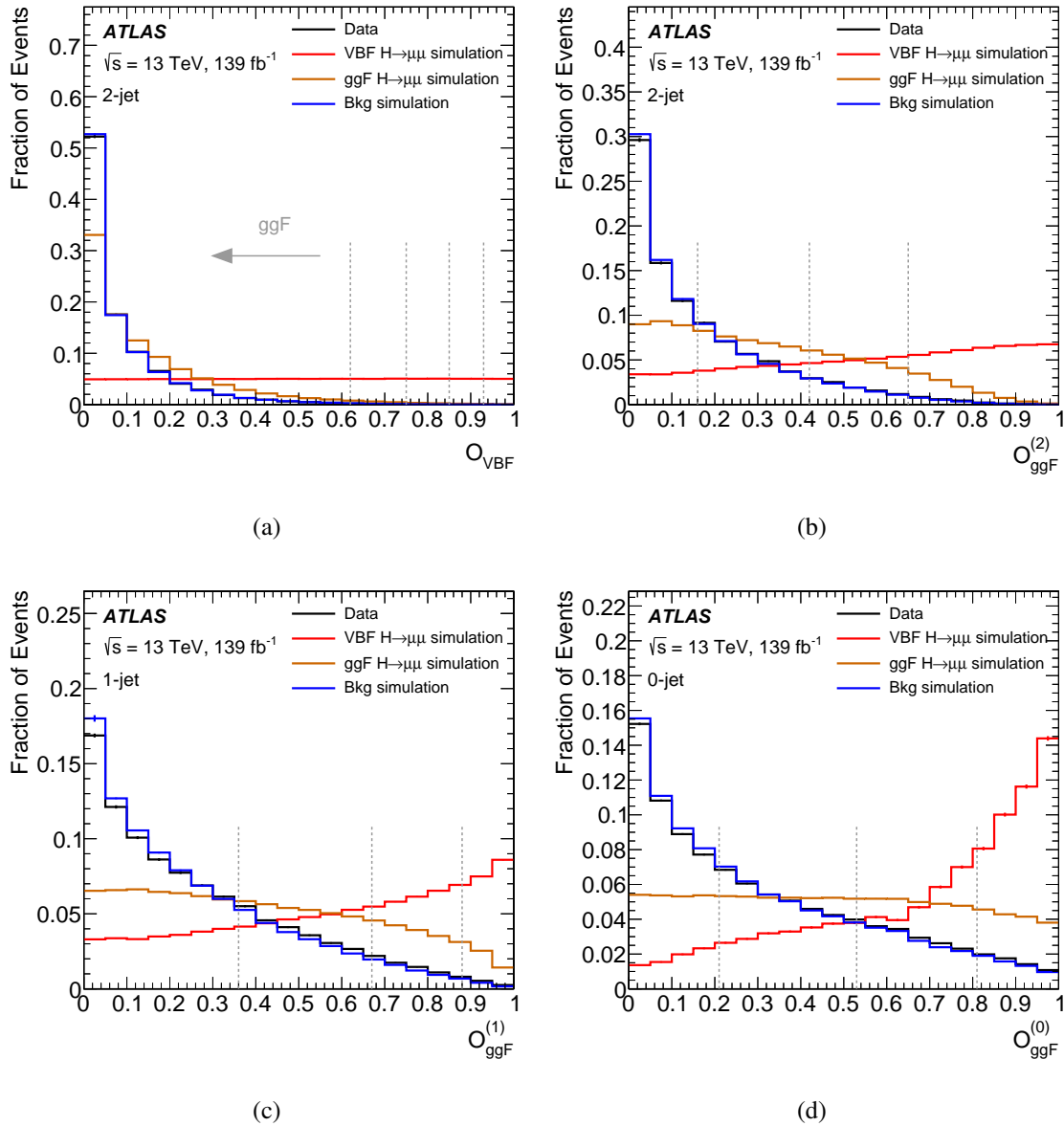


Figure 6.15: The VBF (a) and  $ggF$  (b)-(d) BDT score distributions for signal and background. The grey vertical dashed lines indicate the BDT score boundary define each categories.

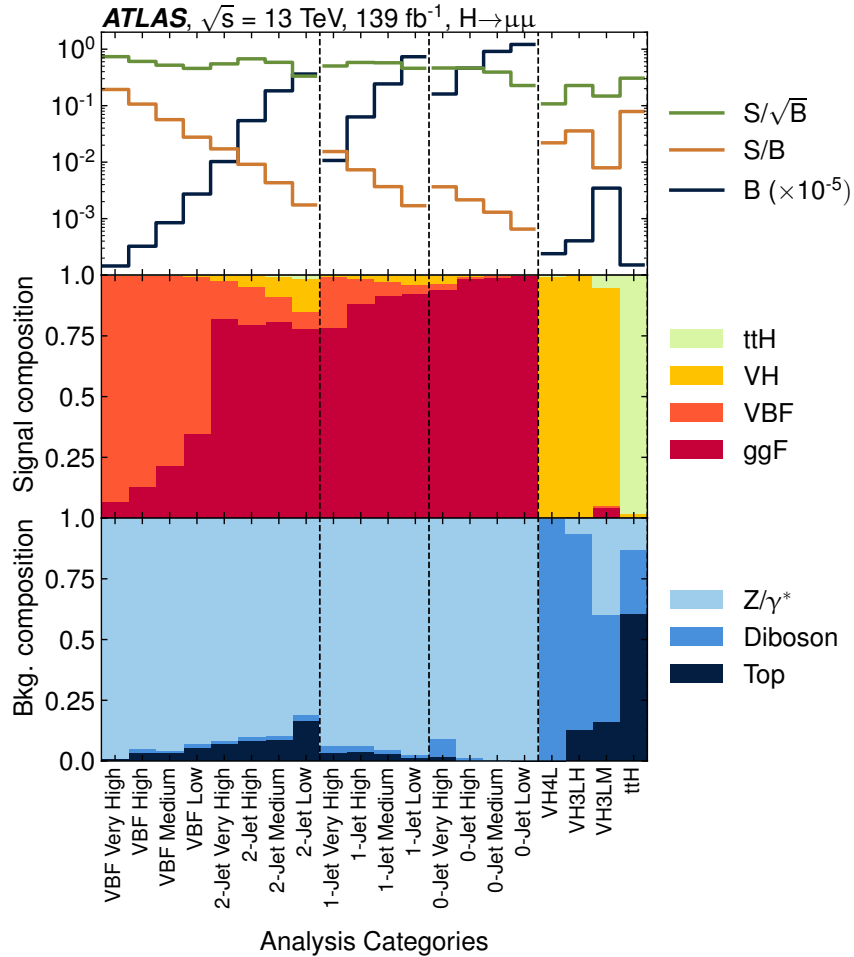


Figure 6.16: Summary of the signal and background composition in the twenty categories in the mass region  $120 \text{ GeV} < \mu^+\mu^- < 130 \text{ GeV}$ . The top panel shows the number of background events ( $B$ , multiplied by a factor of  $10^{-5}$ ), the expected signal-to-background ratio for a SM  $H \rightarrow \mu\mu$  signal ( $S/B$ ), as well as the number counting significance  $S/\sqrt{B}$ . The middle panel shows the expected contributions of different Higgs boson production modes to the total signal in each category. The bottom panel shows the background composition from DY, diboson and top-quark processes in simulation.

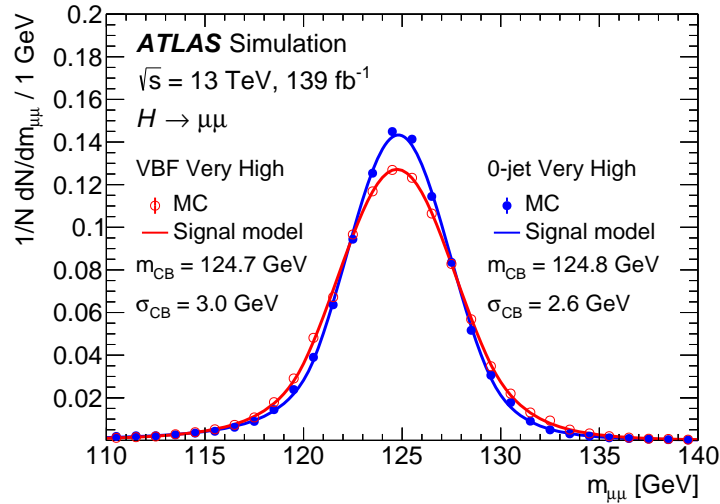


Figure 6.17: Examples of signal parametrization in the categories “VBF Very High” and “2-jet Low”.

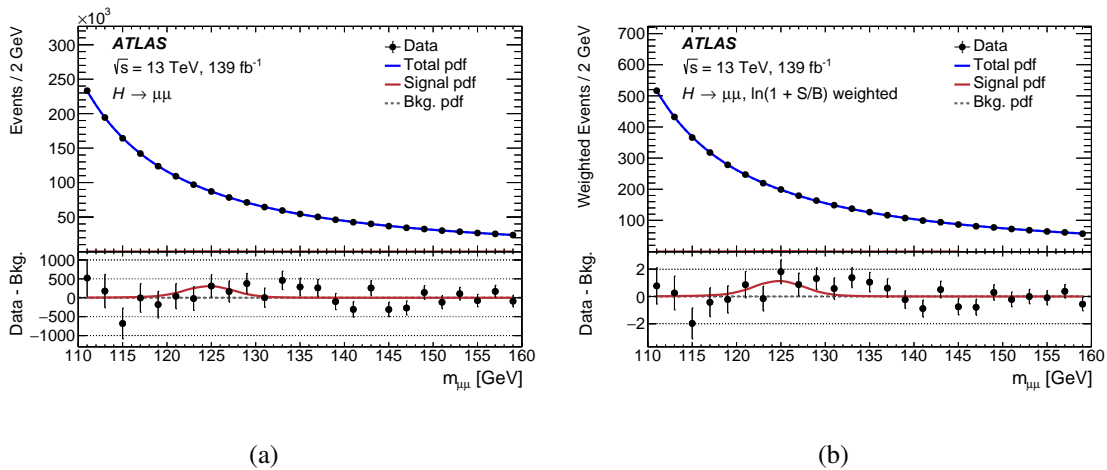


Figure 6.18: Dimuon invariant mass spectrum in all the analysis categories observed in data. In (a) the unweighted sum of all events and signal plus background probability density functions (pdf) are shown, while in (b) events and pdfs are weighted by  $\ln(1 + S/B)$ , where  $S$  are the observed signal yields and  $B$  are the background yields derived from the fit to data in the  $m_{\mu\mu} = 120\text{--}130$  GeV window. The background and signal pdf are derived from the fit to the data, with  $S$  normalised to its best-fit value. The lower panels compare the fitted signal pdf, normalised to the signal best-fit value, to the difference between the data and the background model. The error bars represent the data statistical uncertainties.

## Chapter 7

### Search for mono- $h(bb)$

#### 7.1 Event selections and categorization

In the mono- $h(bb)$  analysis, we aim to detect a Higgs boson decaying into two  $b$ -quarks along with a large missing transverse momentum  $E_T^{\text{miss}}$ . To achieve this, we categorize events into non-overlapping regions, namely signal regions and control regions, based on their expected signal or background content. The control regions require the presence of one or two leptons, while the signal regions veto events containing loose leptons.

Reconstructing the two  $b$ -quarks as separate jets becomes challenging when the Higgs boson is significantly boosted, as the angle between the two  $b$ -jets is inversely proportional to the  $p_T$  of the Higgs boson. To overcome this challenge, we split the analysis into “resolved” regions, where the decay products of the Higgs boson are reconstructed as two separate jets, and “merged” regions, where the entire Higgs boson decay is reconstructed as a single jet.

In the case of  $b$ -associated production within the 2HDM+ $a$  benchmark model, the Higgs boson and DM particles are produced with an additional pair of  $b$ -quarks from gluon splitting. Hence, to enhance sensitivity to these models, we further divide all regions into those requiring exactly two  $b$ -jets and those requiring three or more  $b$ -jets, which we refer to as “2  $b$ -tag” and “ $\geq 3$   $b$ -tag” regions, respectively. As  $E_T^{\text{miss}}$  is often highly correlated with the mediator mass (e.g.  $m_{Z'}$  in the  $Z'$ -2HDM and  $m_A$  in the 2HDM+ $a$ ),  $E_T^{\text{miss}}$  is very sensitive to different signal models and parameters. All regions are thus also divided by  $E_T^{\text{miss}}$  in both resolved and merged regions. The binning of the  $E_T^{\text{miss}}$  splitting varies between 2  $b$ -tag,  $\geq 3$   $b$ -tag, signal regions and control regions, depending on the sample size available.

### 7.1.1 Common selections

To ensure the reliability of events, we require the presence of a reconstructed primary vertex and exclude fake jets originating from beam-induced backgrounds, cosmic-ray showers, or noisy calorimeter cells [160]. Additionally, events must satisfy  $E_T^{\text{miss}} > 150$  GeV and exclude  $\tau$ -lepton. To further suppress background from  $\tau$ -lepton decays, we apply the “extended  $\tau$ -lepton veto”, which involves vetoing events with small- $R$  jets that have a track multiplicity between 1 and 4 and  $\Delta\phi(\text{jet}, E_T^{\text{miss}}) < 22.5^\circ$ . To reject events with  $E_T^{\text{miss}}$  arising from leptonic heavy-flavour decays or severely mismeasured jets, we reject events where any of the leading up to three small- $R$  jets have  $\Delta\phi(\text{jet}, E_T^{\text{miss}}) < 20^\circ$ . Control regions replace  $E_T^{\text{miss}}$  with  $E_{T,\text{lep.invis.}}^{\text{miss}}$  to mimic the signal regions’ effect, where leptons are considered invisible in the calculation of missing transverse momentum. For the final fit, we use the mass of the Higgs boson candidate ( $m_h$ ) as the discriminating variable, where  $50 \text{ GeV} < m_h < 280 \text{ GeV}$  for the resolved regions and  $50 \text{ GeV} < m_h < 270$  for the merged regions. The lower limit is the lowest calibrated large- $R$  jet mass, and the upper limit is set to be significantly larger than the Higgs boson mass, determined by the  $m_h$  binning used in the fit, which depends on the sample size available.

### 7.1.2 Signal regions

Tab. 7.1 summarizes the signal region selections for both the merged and resolved regions. To reduce the contribution of SM processes producing  $E_T^{\text{miss}}$  through the decay  $W \rightarrow \ell\nu$ , events containing any loose electron or muon are rejected.

#### 7.1.2.1 Resolved regions

The resolved regions select events with  $E_T^{\text{miss}} < 500$  GeV and at least two  $b$ -tagged small- $R$  jets, where the two with the highest  $p_T$  form the Higgs boson candidate. The combined  $p_T$  of this two-jet system ( $p_{T_h}$ ) must be greater than 100 GeV, and its mass is corrected for nearby muons, as described in Chap 5, to form  $m_h$ .

The dominant background in the resolved region is  $t\bar{t}$  production, where one top quark decays leptonically, but the lepton is either not reconstructed or not correctly identified. In such cases, all the  $E_T^{\text{miss}}$  in the event (beyond that from mismeasurement) comes from the decay of one of the

Resolved	Merged
Primary $E_T^{\text{miss}}$ trigger	
Data quality selections	
$E_T^{\text{miss}} > 150 \text{ GeV}$	
Lepton veto & extended $\tau$ -lepton veto	
$\Delta\phi(\text{jet}_{1,2,3}, E_T^{\text{miss}}) > 20$	
$E_T^{\text{miss}} < 500 \text{ GeV}$	$E_T^{\text{miss}} > 500 \text{ GeV}$
At least 2 small- $R$ jets	At least 1 large- $R$ jet
At least 2 $b$ -tagged small- $R$ jets	At least 2 $b$ -tagged associated variable- $R$ track-jets
$p_{Th} > 100 \text{ GeV}$ if $E_T^{\text{miss}} < 350 \text{ GeV}$ $p_{Th} > 300 \text{ GeV}$ if $E_T^{\text{miss}} > 350 \text{ GeV}$	—
$m_T^{b,\text{min}} > 170 \text{ GeV}$	—
$m_T^{b,\text{max}} > 200 \text{ GeV}$	—
$S > 12$	—
$N_{\text{small-}R\text{jets}} \leq 4$ if 2 $b$ -tag $N_{\text{small-}R\text{jets}} \leq 5$ if $\geq 3$ $b$ -tag	—
$50 \text{ GeV} < m_h < 280 \text{ GeV}$	$50 \text{ GeV} < m_h < 270 \text{ GeV}$

Table 7.1: Summary of selections used to define the signal regions used in the analysis. The kinematic variables are defined in the text.

two  $W$  bosons, and therefore the transverse mass of the  $E_T^{\text{miss}}$  and the corresponding  $b$ -jet should be approximately bounded from above by the top-quark mass. We define the transverse mass as:

$$m_T^{b,\text{min/max}} = \sqrt{2p_T^{b,\text{min/max}} E_T^{\text{miss}} (1 - \cos\Delta\phi(p_T^{b,\text{min/max}}, E_T^{\text{miss}}))} \quad (7.1)$$

where  $p_T^{b,\text{min}}$  and  $p_T^{b,\text{max}}$  are defined as the  $p_T$  of the  $b$ -jet closest to (min) or furthest from (max) the  $E_T^{\text{miss}}$  in  $\phi$ . Events must satisfy  $m_T^{b,\text{min}} > 170$  GeV and  $m_T^{b,\text{max}} > 200$  GeV.

To reduce contributions from multijet backgrounds, the signal region selections require an object-based  $E_T^{\text{miss}}$  significance [161]  $S > 12$ , which assesses the likelihood that  $E_T^{\text{miss}}$  is due to invisible particles rather than mismeasurements. Data-driven estimates of the remaining multijet contributions were found to be smaller than the expected statistical uncertainty of the data, so the impact of multijet processes is not included in the background estimation.

Events are split to 2  $b$ -tag and  $\geq 3$   $b$ -tag regions. The 2  $b$ -tag regions require events with exactly two  $b$ -tagged jets and at most four small- $R$  jets, counted only for central jets. For the  $\geq 3$   $b$ -tag regions, at most five small- $R$  jets are allowed to ensure a sufficient sample size in the corresponding control regions. The 2  $b$ -tag and  $\geq 3$   $b$ -tag regions are split into three  $E_T^{\text{miss}}$  bins:  $150 \text{ GeV} < E_T^{\text{miss}} < 200 \text{ GeV}$ ,  $200 \text{ GeV} < E_T^{\text{miss}} < 350 \text{ GeV}$  and  $350 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$ . This leads to six resolved signal regions: one for each combination of EmissT bin and number of  $b$ -tagged jets. In the highest  $E_T^{\text{miss}}$  bin, the requirement on  $p_{T_h}$  is tightened to be greater than 300 GeV.

The  $E_T^{\text{miss}}$  triggers become fully efficient at an offline  $E_T^{\text{miss}}$  value close to 200 GeV but are also used for events in the range  $150 \text{ GeV} < E_T^{\text{miss}} < 200 \text{ GeV}$ . To correct for MC mismodelling of the  $E_T^{\text{miss}}$  trigger response, trigger efficiencies are measured in both data and simulation, and scale factors are calculated to correct the simulation. The scale factors are calculated as a function of  $E_{T, \text{lep, invis}}^{\text{miss}}$  in a region whose selection matches the  $150 \text{ GeV} < E_T^{\text{miss}} < 200 \text{ GeV}$  and 2  $b$ -tag region except that all  $E_T^{\text{miss}}$  selections are dropped, exactly one  $b$ -tagged jet is required, and exactly one signal muon is required. Events containing electrons are still vetoed. The scale factors have values in the range 0.95–1.0.

### 7.1.2.2 Merged regions

The merged regions are defined by selecting events with  $E_T^{\text{miss}} > 500$  GeV. At least one large- $R$  jet is required, and the two leading variable- $R$  track-jets associated with the leading large- $R$  jet

are required to be  $b$ -tagged. This large- $R$  jet, defined to be the Higgs boson candidate, has its mass corrected for nearby muons, as described in Chapter 5, to form  $m_h$ . Events are separated into two categories: those with no additional  $b$ -tagged variable- $R$  track-jets (referred to as 2  $b$ -tag region) and those with at least one such  $b$ -tagged variable- $R$  track-jet not associated with the Higgs boson candidate (referred to as  $\geq 3$   $b$ -tag region). The merged 2  $b$ -tag region is divided into two  $E_T^{\text{miss}}$  bins,  $500 \text{ GeV} < E_T^{\text{miss}} < 750 \text{ GeV}$  and  $E_T^{\text{miss}} > 750 \text{ GeV}$ . There is no further binning for the  $\geq 3$   $b$ -tag region.

### 7.1.3 Control regions

We define control regions to better constrain the normalization of main background processes where specific background processes dominate and negligible signal is expected. 1-muon control regions are defined in order to constrain the  $W$  (produced in association with jets) and  $t\bar{t}$  processes.  $W$  and  $t\bar{t}$  processes contribute in the signal regions if leptons in the decays are either not identified or outside the kinematic acceptance. The main contribution arises from hadronically decaying  $\tau$ -leptons. As the shape of event variables is the same for all lepton flavors, the kinematic phase space of the signal region can be approximated closely by control regions requiring a medium muon (1-muon control regions). In this case,  $E_{T, \text{lep, invis}}^{\text{miss}}$  is used as a proxy for  $E_T^{\text{miss}}$  to better approximate the signal regions, which veto the presence of any leptons. Additionally, any other variable using  $E_T^{\text{miss}}$  in its calculation, such as  $E_T^{\text{miss}}$  significance and  $m_T^{\text{b,min/max}}$ , is constructed using  $E_{T, \text{lep, invis}}^{\text{miss}}$ . This ensures that the  $E_T^{\text{miss}}$ -related quantities in the control regions correspond to those in the signal regions. Otherwise, the 1-muon control regions are defined by the same criteria as the signal regions. In 1-muon 2  $b$ -tag control regions, the sign of the muon charge is used as the fit discriminant, which separates  $W$  and  $t\bar{t}$ . In 1-muon  $\geq 3$   $b$ -tag control regions, no further binning other than  $E_T^{\text{miss}}$  is considered due to limited data statistics.

On the other hand, 2-lepton control regions are defined to constrain the  $Z \rightarrow \nu\nu$  process produced in association with jets. As the momentum of the  $Z$  boson does not depend on its decay mode,  $Z \rightarrow \nu\nu$  in the signal region can be closely modeled by  $Z \rightarrow \ell^+\ell^-$  events by requiring exactly two loose electrons or muons with opposite charge. These events are collected using single-electron or single-muon triggers. Furthermore, one of the electrons or muons is required to be a medium electron or muon with  $p_T > 27 \text{ GeV}$  or  $p_T > 25 \text{ GeV}$ , respectively. The invariant mass

of the di-lepton system is required to be consistent with the mass of the  $Z$  boson within 10 GeV. While keeping all other criteria of the signal regions, an additional criterion of  $S < 5$  is imposed to suppress a remaining contribution of  $t\bar{t}$  processes. To be similar to the signal regions,  $E_{T, \text{lep, invis}}^{\text{miss}}$  is used as proxy for  $E_T^{\text{miss}}$  and in the calculation of any other variable using  $E_T^{\text{miss}}$ . In 2-lepton control regions, no further binning other than  $E_T^{\text{miss}}$  is considered.

All control regions are split to four  $E_T^{\text{miss}}$  bins, including the resolved regions:  $150 \text{ GeV} < E_T^{\text{miss}} < 200 \text{ GeV}$ ,  $200 \text{ GeV} < E_T^{\text{miss}} < 350 \text{ GeV}$ ,  $350 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$ , and merged region:  $E_T^{\text{miss}} > 500 \text{ GeV}$ .

## 7.2 Statistical analysis

The same binned profile likelihood function equation 6.4 is used in mono- $h(bb)$  to perform statistical fitting in all signal regions and control regions simultaneously and extract signal yields.  $S_i$  and  $B_i$  are modeled by histograms derived from MC samples.

The analysis is split to a model-specific scenario and a model-independent scenario. In the model-specific scenario, one global signal strength  $\mu$  is used for all signal and control regions. The invariant mass of the Higgs boson candidate  $m_h$  is used as the fit discriminant in the signal regions. The bin size of  $m_h$  is 5 GeV in 2  $b$ -tag  $150 \text{ GeV} < E_T^{\text{miss}} < 200 \text{ GeV}$ ,  $200 \text{ GeV} < E_T^{\text{miss}} < 350 \text{ GeV}$ , 10 GeV in 2  $b$ -tag  $350 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$ ,  $\geq 3$   $b$ -tag  $150 \text{ GeV} < E_T^{\text{miss}} < 200 \text{ GeV}$ ,  $200 \text{ GeV} < E_T^{\text{miss}} < 350 \text{ GeV}$ , 20 GeV in 2  $b$ -tag  $500 \text{ GeV} < E_T^{\text{miss}} < 750 \text{ GeV}$ . In 2  $b$ -tag  $E_T^{\text{miss}} > 750 \text{ GeV}$  and  $\geq 3$   $b$ -tag  $E_T^{\text{miss}} > 500 \text{ GeV}$ , the  $m_h$  binning is customized as [50, 90, 150, 270) GeV, while in  $\geq 3$   $b$ -tag  $350 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$ , the  $m_h$  binning is customized as [50, 110, 150, 280) GeV.

In the model-independent scenario, each signal region is assigned a separate signal strength  $\mu_i$  in order to measure the visible cross-section, defined as:

$$\sigma_{\text{vis}, h(bb)+\text{DM}} \equiv \sigma_{h+\text{DM}} \times \mathcal{B}(h \rightarrow b\bar{b}) \times (\mathcal{A} \times \varepsilon) \quad (7.2)$$

where  $(\mathcal{A} \times \varepsilon)$  with the acceptance  $\mathcal{A}$  and the reconstruction efficiency  $\varepsilon$  quantifies the probability for a certain event to be reconstructed within a window around the Higgs boson mass in a given signal region. The visible cross-section is obtained from the number of signal events in each

signal region extracted from a fit to the  $m(b\bar{b})$  distribution as described below, divided by the integrated luminosity. In contrast to the model-specific scenario, we assume that a signal resonance is produced with a mass close to 125 GeV and decays into a pair of  $b$  quarks in association with  $E_{\text{T}}^{\text{miss}}$ . For this purpose, the binning in  $m_h$  in the signal regions is modified from the model-specific scenario such that all bins under the Higgs boson peak in a range from 90 GeV to 150 GeV are merged into one bin (Higgs window). This Higgs window includes the Higgs boson peak, but excludes important parts of the  $Z$  boson peak.

### 7.2.1 Signal modeling

In the model-specific scenario, the signal models considered for the statistical interpretation include  $Z'$ -2HDM and 2HDM+ $a$  as described in section 2.3. The MC samples for each of the models are used to model the signal shape in the spectra of the fit discriminants. In the model-independent scenario, the signal is assumed to be present only in the Higgs window.

### 7.2.2 Background modeling

The dominant backgrounds in the signal regions are  $t\bar{t}$  and  $W/Z$  processes. The  $W/Z$  backgrounds are subdivided according to the true flavour of the associated jets that constitute the Higgs boson candidate during reconstruction. Simulated jets are labelled according to which hadrons with  $p_{\text{T}} > 5$  GeV are found within a cone of size  $\Delta R = 0.3$  around the jet axis. If a  $b$ -hadron is found the jet is labelled as a  $b$ -jet. If no  $b$ -hadron is found, but a  $c$ -hadron is present, then the jet is labelled as a  $c$ -jet. Otherwise the jet is labelled as a light jet. The flavour of the two leading  $b$ -tagged track-jets is used in the merged region. If the flavour of either (or both) those two jets is a  $b$ -quark, the event is considered to be  $W/Z$ +HF background, where HF stands for Heavy Flavour. In the 2  $b$ -tag resolved regions the dominant backgrounds are  $t\bar{t}$  and  $Z$ +HF, with the latter becoming more important as the  $E_{\text{T}}^{\text{miss}}$  increases. The 2  $b$ -tag merged regions are dominated by  $Z$ +HF. Both the resolved and merged  $\geq 3$   $b$ -tag regions are dominated by  $t\bar{t}$ , where the extra  $b$ -jet typically is a mis-tagged jet originating from a hadronic  $W$  boson decay. At higher  $E_{\text{T}}^{\text{miss}}$  values the  $Z$ +HF background becomes important again.

The kinematic shapes of all background processes are modeled using simulation samples. We use free parameters (i.e. without prior constraint in the likelihood model) to model the normalizations of the dominant backgrounds. The normalizations of  $t\bar{t}$ ,  $W$ +HF are determined by a single free parameter each, while the normalization of  $Z$ +HF is determined by two free parameters where one is assigned to the events in 2  $b$ -tag regions and the other one is assigned to the events in  $\geq 3$   $b$ -tag regions.

The normalizations of smaller backgrounds are taken directly from simulation. These include the production of a  $W$  or  $Z$  boson in association with light jets or at most one jet containing a  $c$ -hadron, single top-quark production (dominated by production in association with a  $W$  boson) and diboson processes. Small contributions also arise from  $t\bar{t}$  processes in association with vector bosons or a Higgs boson. Another background contribution stems from vector-boson production in association with a Higgs boson ( $Vh$ ), which mimics the signal due to the presence of a Higgs boson peak in association with jets. In the case where the vector boson is a  $Z$  boson decaying to neutrinos this is an irreducible background. Similarly, the diboson decay  $ZZ \rightarrow b\bar{b}\nu\nu$  is nearly irreducible due to small difference in  $Z$  and  $h$  mass peaks (compared to the Higgs candidate mass resolution). The multijet background is negligible in all regions after the requirements on the object-based  $E_T^{\text{miss}}$  significance  $S$  and on  $\Delta\phi(\text{jet}, E_T^{\text{miss}})$ , and thus not further considered.

### 7.2.3 Systematic uncertainties

Signal and background expectations are subject to statistical, detector-related and theoretical uncertainties, which are all included in the likelihood as nuisance parameters. Detector-related and theoretical uncertainties may affect the overall normalisation and/or shape of the simulated background and signal event distributions.

Detector-related uncertainties are dominated by contributions from the jet reconstruction. Uncertainties in the jet energy scale (JES) for small- $R$  jets [162] arise from the calibration of the scale of the jet and are derived as function of the jet  $p_T$ , and also  $\eta$ . Further contributions emerge from the jet flavour composition and the pile-up conditions. The ‘category reduction’ scheme as described in Ref. [162] with 29 nuisance parameters is used. Uncertainties in the jet energy resolution (JER) depend on the jet  $p_T$  and  $\eta$  and arise both from the method used to derive the jet resolution and from the difference between simulation and data [162], and they are included with

eight nuisance parameters. Similarly, uncertainties in the jet energy resolution for large- $R$  jets arise from the calibration, the flavour composition and the topology dependence [163]. Further uncertainties are considered for the large- $R$  jet mass scale [163] and resolution [164].

Uncertainties due to the  $b$ -tagging efficiency for heavy-flavour jets, including  $c$ -flavour jets, are derived from  $t\bar{t}$  data [128, 165] and are represented by four nuisance parameters. Uncertainties are also considered for mistakenly  $b$ -tagging a light-flavour jet, with nine nuisance parameters. These are estimated using a method similar to that in Ref. [166].

Uncertainties in the modeling of  $E_T^{\text{miss}}$  are evaluated by considering the uncertainties affecting the jets included in the calculation and the uncertainties in soft term's scale and resolution [151]. The pile-up in simulation is matched to the conditions in data by a reweighting factor. An uncertainty of 4% is assigned to this reweighting factor. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [154], obtained using the LUCID-2 detector [155] for the primary luminosity measurements.

Scale factors, including their uncertainties, are calculated specifically for this analysis to correct the efficiency of  $E_T^{\text{miss}}$  triggers in simulation to that in data. The uncertainties in the scale factors are at most 1%–2% for low  $E_T^{\text{miss}}$  values.

In the regions requiring the presence of leptons, uncertainties in the lepton identification and lepton energy/momentum scale and resolution are included. These are derived using simulated and measured events with  $Z \rightarrow \ell^+\ell^-$ ,  $J/\psi \rightarrow \ell^+\ell^-$  and  $W \rightarrow \ell\nu$  decays [25, 132].

Modelling uncertainties impact the shape of the  $m_h$  distribution, the relative acceptance between different  $E_T^{\text{miss}}$  and  $b$ -tag multiplicity bins and between signal and control regions as well as the overall normalisation of the samples that are not freely floating in the fit.

The theoretical uncertainties are dominated by modelling uncertainties in the  $t\bar{t}$  and  $Z$ +HF backgrounds. For the  $t\bar{t}$  and  $Wt$  processes, the impact of the choice of parton shower and hadronisation model is evaluated by comparing the sample from the nominal generator set-up with a sample interfaced to HERWIG7.04 [149, 150]. To assess the uncertainty in the matching of NLO matrix elements to the parton shower, the POWHEG-BOX sample is compared with a sample of events generated with MADGRAPH5\_aMC@NLO v2.6.2. For the  $Wt$  process, the nominal sample is compared with an alternative sample generated using the diagram subtraction scheme [116, 167]

instead of the diagram removal scheme to estimate the uncertainty arising from the interference with  $t\bar{t}$  production.

For the  $V$ +jet processes, uncertainties arising from the modelling of the parton shower and the matching scheme are evaluated by comparing the nominal samples with samples generated with MADGRAPH5\_aMC@NLO v2.2.2. For the diboson processes, the uncertainties associated with the modelling of the parton shower, the hadronisation and the underlying event are derived using alternative samples generated with POWHEG-BOX [60–62] and interfaced to PYTHIA8.186 [168] or HERWIG++.

For all MC samples, the uncertainties due to missing higher orders are estimated by a variation of the renormalisation and factorisation scales by a factor of two, while the PDF and  $\alpha_s$  uncertainties are calculated using the PDF4LHC prescription [68].

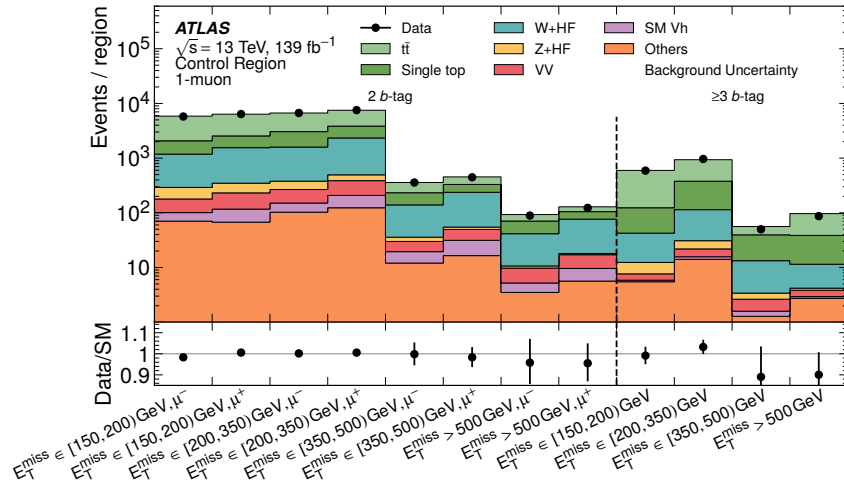
Tab. 7.2 gives the impact of the different sources of systematic uncertainties for selected signal models as evaluated in different model-dependent fits. The signal models with lower masses illustrate the impact of the systematic uncertainties in the resolved regions, while the models with larger mediator masses are more impacted by the merged regions. The theoretical uncertainties in the modelling of the  $t\bar{t}$  background, the experimental uncertainties in the calibration of jets and the limited MC sample size show the largest impact.

## 7.2.4 Results

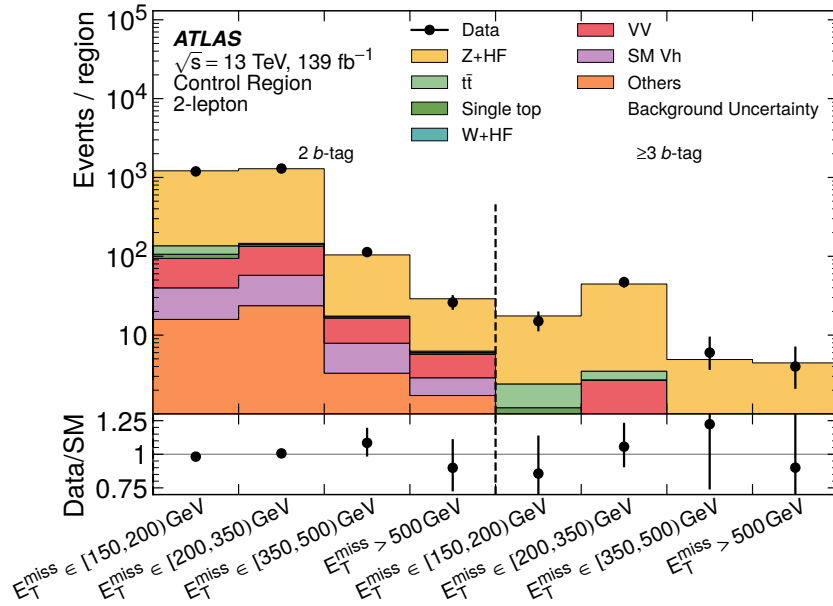
The post-fit background yields are determined in a background-only profile likelihood fit to data in all regions. Fig. 7.1 shows the yields in the 1-muon and 2-lepton control regions. The post-fit normalisation factors for  $t\bar{t}$  and for  $W$ +HF are found to be  $0.93 \pm 0.08$  and  $0.95 \pm 0.14$ , respectively. For  $Z$  boson production in association with two (at least three) heavy-flavour jets the normalisation factors are determined to  $1.41 \pm 0.09$  ( $1.85 \pm 0.24$ ). An upward scaling of the  $Z$ +HF background relative to the simulation was also observed in other studies [169], and was attributed to an underestimation of the  $g \rightarrow b\bar{b}$  rate in SHERPA. A larger scaling is observed in the region with  $\geq 3$   $b$ -tagged jets, dominated by processes with more  $g \rightarrow b\bar{b}$  splittings. The uncertainty on the  $Z$ +HF normalisation factor increases in the  $\geq 3$   $b$ -tag region due to lower statistical precision and the smaller contribution of the  $Z$ +HF background.

Source of uncertainty	Fractional squared uncertainty in $\mu$		
	$Z'$ -2HDM signals, $(m'_Z, m_A)$ [GeV]		
	(800, 500)	(1400, 1000)	(2800, 300)
$Z$ +HF normalisation	0.11	0.03	< 0.01
$W$ +HF normalisation	0.02	0.01	< 0.01
$t\bar{t}$ normalisation	0.16	0.04	< 0.01
$Z$ modelling uncertainties	0.02	0.07	< 0.01
$W$ modelling uncertainties	< 0.01	0.01	< 0.01
$t\bar{t}$ modelling uncertainties	0.13	0.05	< 0.01
Single- $t$ modelling uncertainties	0.18	0.02	< 0.01
Other modelling uncertainties	0.05	0.01	< 0.01
Jets	0.20	0.06	0.01
$b$ -tagging	0.01	0.01	0.04
$E_T^{\text{miss}}$ soft term and pile-up	< 0.01	< 0.01	< 0.01
Other experimental systematic uncertainties	0.01	< 0.01	< 0.01
Signal systematic uncertainties	< 0.01	< 0.01	< 0.01
MC sample size	0.08	0.07	0.11
Statistical uncertainty	0.27	0.61	0.79
Total systematic uncertainties	0.73	0.39	0.21

Table 7.2: Relative importance of the different sources of uncertainty for different  $Z'$ -2HDMs, with the masses of the  $Z'$  boson and the  $A$  boson given in the second row, expressed as fractional impact on the signal strength parameter. The fractional impact is calculated by considering the square of the uncertainty in the signal strength parameter arising from a given group of uncertainties (as listed in the left column of the table), divided by the square of the total uncertainty in the signal strength parameter. Due to correlations, the sum of the different impacts of systematic uncertainties might not add up to the total impact of all systematic uncertainties.



(a)



(b)

Figure 7.1: Yields in the resolved and merged (a) 1-muon control regions and (b) 2-lepton control regions. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds. The different control region bins included in the fit are indicated on the  $x$ -axis by first giving the range in  $E_T^{\text{miss}}$  and then the sign of the muon charge (where applicable).

The distributions of the Higgs boson candidate mass  $m_h$  after the background-only fit are shown in Fig. 7.2 and 7.3. The signal to background ratio is higher in the  $\geq 3$   $b$ -tag signal regions, because the 2HDM+ $a$  signal model is shown with  $\tan \beta = 10$ , where  $b$ -associated production dominates. Tab. 7.3 and 7.4 present the background estimates in comparison with the observed data. No significant deviation from SM expectations is observed, with the largest deficit corresponding to a local significance of  $2.3\sigma$ , and the largest excess amounting to  $1.6\sigma$ . Fig. 7.4 summarises the total yields in the signal regions as a function of  $E_T^{\text{miss}}$ . The background prediction from simulation is scaled upwards in the fit for lower  $E_T^{\text{miss}}$  values, while the simulation agrees better with the data for large  $E_T^{\text{miss}}$  values.

The results are interpreted as exclusion limits at 95% CL in the  $Z'$ -2HDM and the 2HDM+ $a$  scenarios in Fig. 7.5 and 7.6. Considering the  $Z'$ -2HDM case,  $Z'$  masses up to 3 TeV are excluded for  $A$  masses of 300 GeV at 95% CL. The exclusion boundaries for the 2HDM+ $a$  scenario extend up to  $m_a = 520$  GeV for  $m_A = 1.25$  TeV for  $ggF$  production and  $\tan \beta = 1$ . This is an improvement of about 200 GeV in  $m_a$  on previous results [170], which reinterpreted the earlier  $h(\rightarrow b\bar{b}) + E_T^{\text{miss}}$  analysis using  $36.1 \text{ fb}^{-1}$  [171].

The higher exclusion limit at high  $m_A$ , low  $m_a$ , is due to an increase of the cross-section of the  $a \rightarrow ah$  process, without resonant  $A$  production. It should be noted that with the exact parameter choices adopted in this analysis, the  $aah$  coupling becomes larger than  $4\pi$  for  $m_A \gtrsim 1750$  GeV. Moreover, as discussed in Refs. [28, 29] values of  $m_A \gtrsim 1250$  GeV (for  $\tan \beta = 1$ ) or  $m_A \gtrsim 2150$  GeV (for  $\tan \beta = 10$ ) would not be consistent with the requirement of having a bounded-from-below scalar potential, given the parameter choices discussed in this paper. These constraints can be relaxed substantially if the quartic couplings assume a value closer to the perturbativity limit and also in more general 2HDMs containing additional couplings as discussed in Refs. [172, 173]. Therefore, the above should not be considered as a strong requirement for the validity of the model predictions. At high  $m_A$  the width of the additional Higgs bosons grows substantially and the theoretical predictions are subject to additional theoretical uncertainties associated with the treatment of the width. Exclusion limits are therefore not shown in the region of very large widths ( $m_A > 2200$  GeV).

In the case of  $bbA$  production and  $\tan \beta = 10$ , the exclusion limits extend up to  $m_a = 240$  GeV for  $m_A = 900$  GeV. The inclusion of the  $\geq 3$   $b$ -tag regions helps to increase the sensitivity relative

2 $b$ -tag signal regions $E_T^{\text{miss}}$ range	[150, 200) GeV	[200, 350) GeV	[350, 500) GeV	[500, 750) GeV	> 750 GeV
$Z$ +HF	$6470 \pm 310$	$7200 \pm 310$	$507 \pm 26$	$94 \pm 7$	$9.2 \pm 1.8$
$Z$ +light jets	$72 \pm 15$	$137 \pm 29$	$18 \pm 4$	$4.5 \pm 1.0$	$1.17 \pm 0.30$
$W$ +HF	$1590 \pm 210$	$1760 \pm 230$	$106 \pm 14$	$25 \pm 4$	$3.1 \pm 0.6$
$W$ +light jets	$86 \pm 35$	$92 \pm 35$	$14 \pm 5$	$1.6 \pm 0.6$	$0.21 \pm 0.09$
Single top-quark	$570 \pm 260$	$570 \pm 260$	$21 \pm 10$	$2.6 \pm 1.9$	$0.10 \pm 0.16$
$t\bar{t}$	$4680 \pm 290$	$3280 \pm 240$	$76 \pm 9$	$11.4 \pm 1.6$	$0.38 \pm 0.08$
Diboson	$450 \pm 50$	$600 \pm 60$	$56 \pm 7$	$15.2 \pm 1.9$	$1.61 \pm 0.29$
$Vh$	$151 \pm 10$	$202 \pm 12$	$26.6 \pm 1.8$	$5.6 \pm 0.5$	$0.68 \pm 0.12$
$t\bar{t}+V/h$	$7.6 \pm 0.4$	$11.8 \pm 0.5$	$0.45 \pm 0.06$	$0.286 \pm 0.029$	$0.035 \pm 0.006$
Total background	$14070 \pm 110$	$13860 \pm 100$	$825 \pm 19$	$160 \pm 8$	$16.7 \pm 1.9$
Data	14259	13724	799	168	19

Table 7.3: Background yields in comparison with data in the 2  $b$ -tag signal regions for different  $E_T^{\text{miss}}$  ranges after a background-only fit to data. Statistical and systematic uncertainties are reported together.

$\geq 3$ $b$ -tag signal regions $E_T^{\text{miss}}$ range	[150, 200) GeV	[200, 350) GeV	[350, 500) GeV	> 500 GeV
$Z$ +HF	$102 \pm 15$	$278 \pm 28$	$26.4 \pm 3.5$	$15.6 \pm 1.9$
$Z$ +light jets	$0.6 \pm 0.4$	$2.9 \pm 0.8$	$0.34 \pm 0.12$	$0.46 \pm 0.12$
$W$ +HF	$21 \pm 4$	$47 \pm 9$	$4.2 \pm 0.9$	$2.4 \pm 0.4$
$W$ +light jets	$0.01 \pm 0.04$	$1.7 \pm 0.9$	$0.8 \pm 0.4$	$0.031 \pm 0.026$
$t\bar{t}$	$276 \pm 19$	$252 \pm 22$	$5.1 \pm 0.7$	$17.9 \pm 1.8$
Single top-quark	$23 \pm 11$	$55 \pm 25$	$2.9 \pm 1.4$	$3.4 \pm 1.7$
Diboson	$4.8 \pm 1.4$	$12.9 \pm 2.2$	$1.8 \pm 0.4$	$1.26 \pm 0.31$
$Vh$	$0.65 \pm 0.28$	$2.9 \pm 0.5$	$0.40 \pm 0.08$	$0.230 \pm 0.025$
$t\bar{t}+V/h$	$1.78 \pm 0.17$	$3.89 \pm 0.26$	$0.371 \pm 0.035$	$0.78 \pm 0.08$
Total background	$430 \pm 15$	$656 \pm 21$	$42 \pm 4$	$42.0 \pm 2.8$
Data	408	658	42	46

Table 7.4: Background yields in comparison with data in the  $\geq 3$   $b$ -tag signal regions for different  $E_T^{\text{miss}}$  ranges after a background-only fit to data. Statistical and systematic uncertainties are reported together.

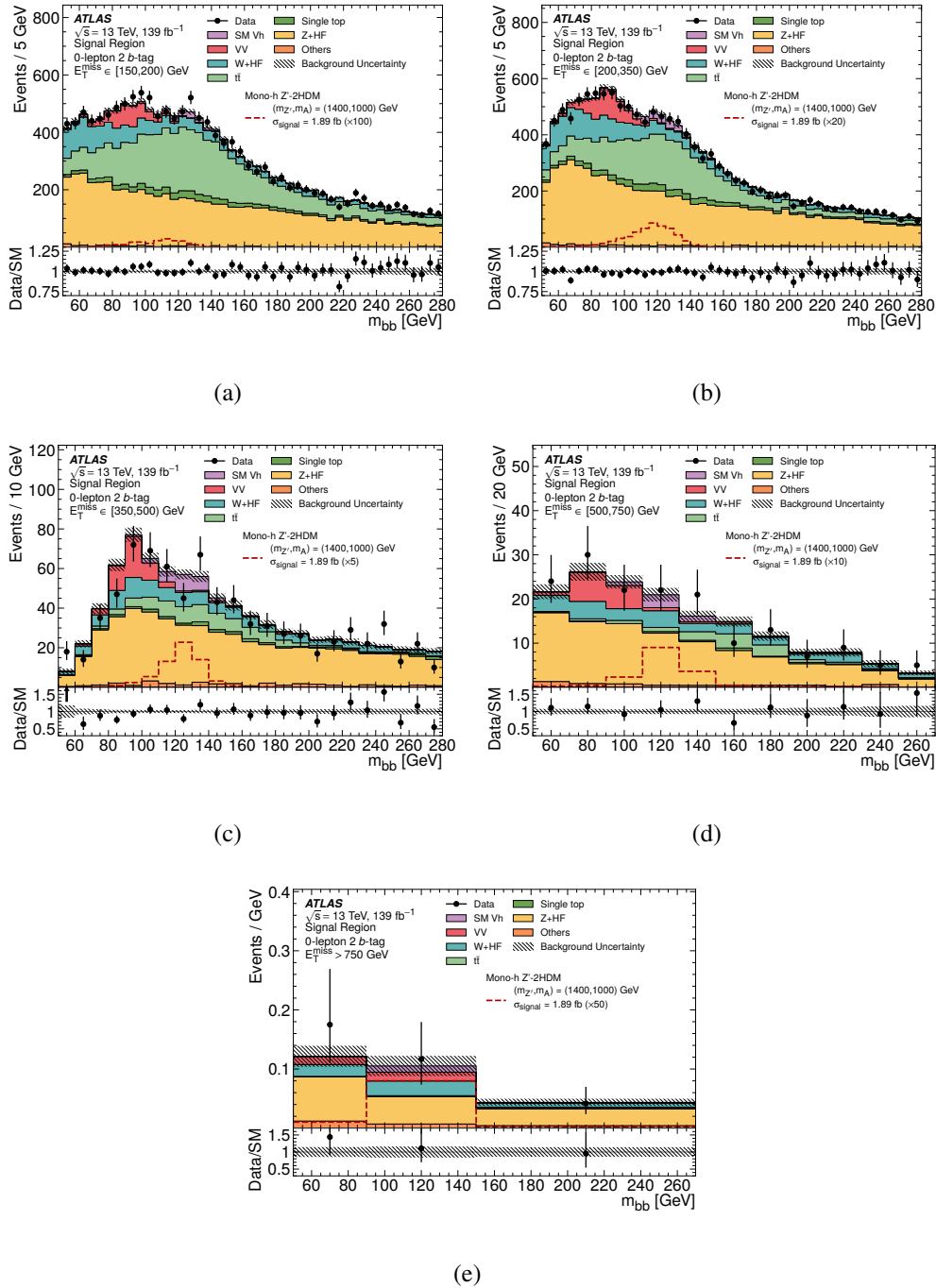


Figure 7.2: Distributions of the Higgs boson candidate mass in the 2  $b$ -tag signal regions for different  $E_T^{\text{miss}}$  ranges. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds.

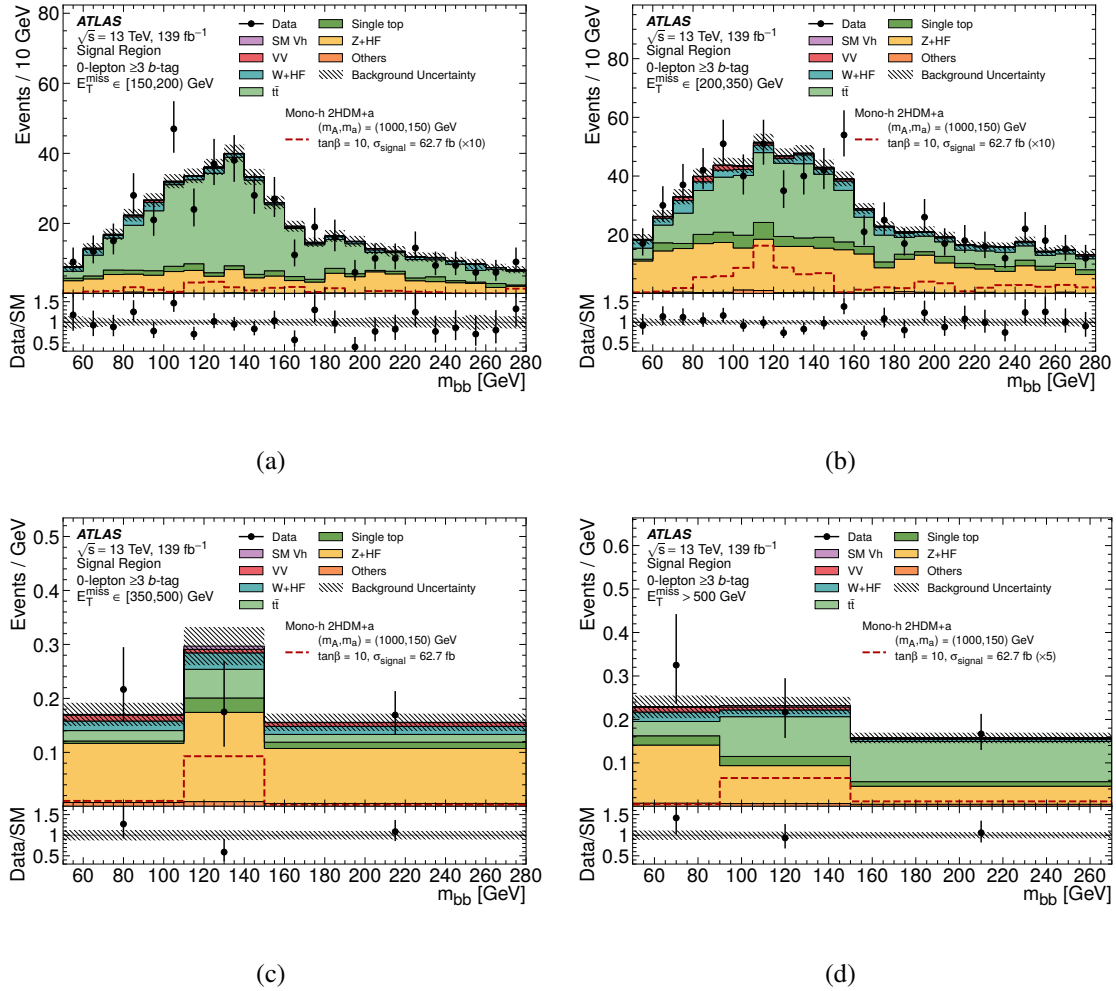


Figure 7.3: Distributions of the Higgs boson candidate mass in the  $\geq 3$   $b$ -tag signal regions for different  $E_T^{\text{miss}}$  ranges. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds.

to the 2  $b$ -tag region by about 30–70%. The difference between observed and expected limits arises from data deficits in the  $\geq 3$   $b$ -tag region, especially the deficit around the Higgs boson peak in the  $E_T^{\text{miss}} \in [350, 500)$  GeV region, as shown in Fig. 7.3. The 2HDM+ $a$  scenario with  $bbA$  production and  $\tan\beta = 10$  is considered for the signatures discussed in this paper for the first time.

Fig. 7.7 displays the upper limits on the visible cross-section, which are derived from the fit with the the model-independent setup. These model-independent limits are particularly useful for the reinterpretation with any additional interesting signal models.

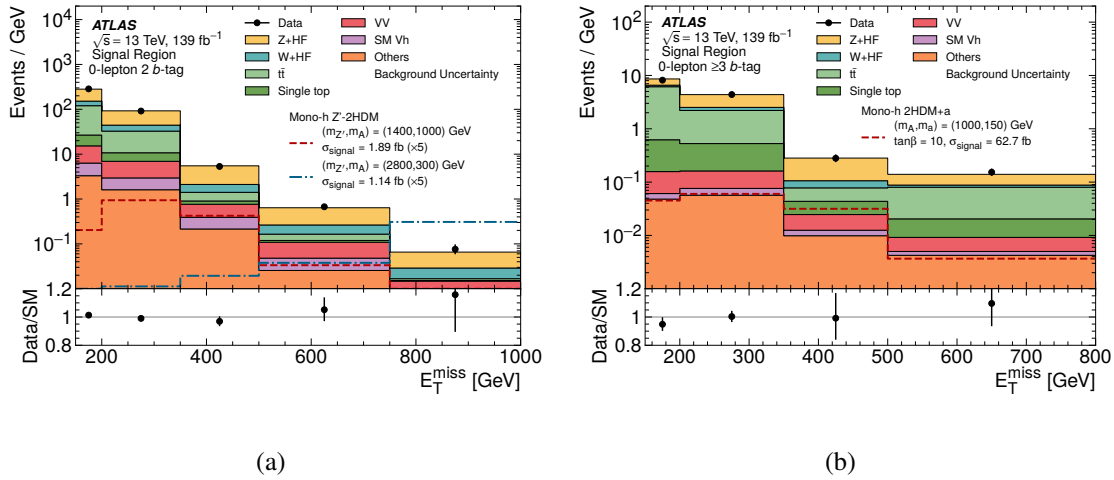


Figure 7.4:  $E_T^{\text{miss}}$  distributions in the (a) 2  $b$ -tag and (b)  $\geq 3$   $b$ -tag signal regions for different  $E_T^{\text{miss}}$  ranges. The top panel compares the fitted background yields with data, while the bottom panel indicates the ratio of the observed data to the predicted Standard Model backgrounds.

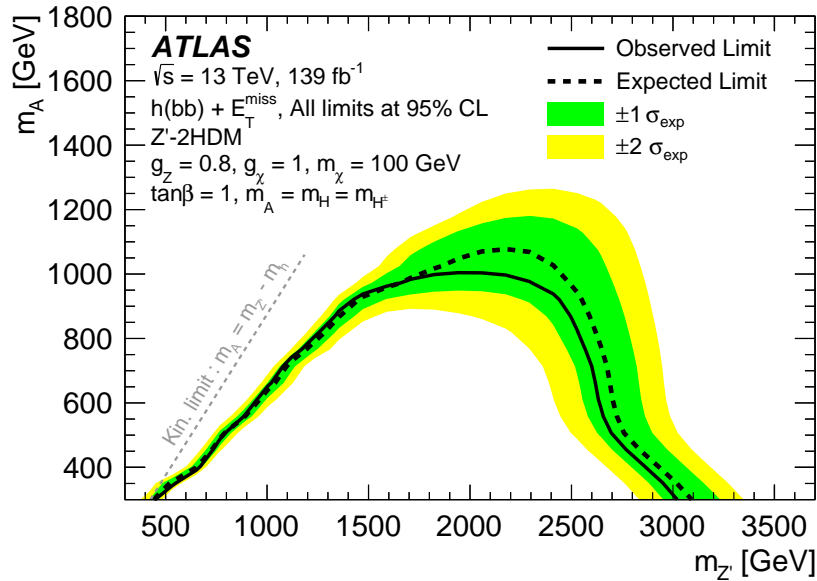


Figure 7.5: Exclusion limits in the  $Z'$ -2HDM model. The solid black line shows the observed limit at 95% CL, the dashed black line the expected limit. The green band gives the  $\pm 1\sigma$  uncertainties of the expected limit, the yellow band the  $\pm 2\sigma$  uncertainties.

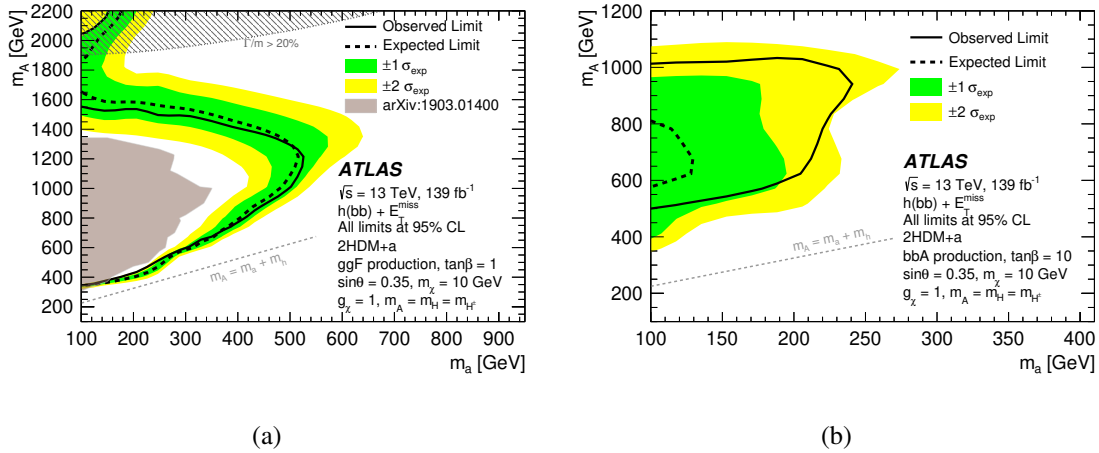


Figure 7.6: Exclusion limits for the 2HDM+a signal with (a)  $\tan\beta = 1$  and  $ggF$  production and with (b)  $\tan\beta = 10$  and  $bbA$  production. The solid black line shows the observed limit at 95% CL, the dashed black line the expected limit. The green band gives the  $\pm 1\sigma$  uncertainties of the expected limit, the yellow band the  $\pm 2\sigma$  uncertainties.

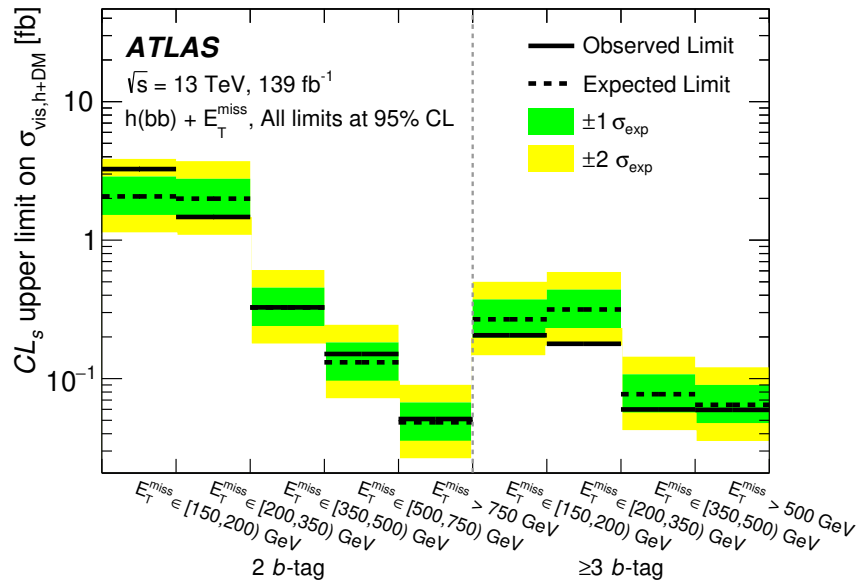


Figure 7.7: Model-independent upper limits on the visible cross-section  $\sigma_{\text{vis},h(bb)+DM} \equiv \sigma_{h+DM} \times \mathcal{B}(h \rightarrow b\bar{b}) \times \mathcal{A} \times \varepsilon$  in the different signal regions.

## Chapter 8

### Unbinned profiled unfolding

#### 8.1 Motivation

One of the most common analysis goals in particle and nuclear physics is the measurement of differential cross sections. These quantities encode the rate at which a particular process occurs as a function of certain observables of interest. From measured cross sections, a number of downstream inference tasks can be performed, including the estimation of fundamental parameters, tuning simulations, and searching for physics beyond the Standard Model. The key challenge of cross section measurements is correcting the data for detector distortions, a process called deconvolution or *unfolding*. See Refs. [174–177] for recent reviews on unfolding and Refs. [178–180] for the most widely-used unfolding algorithms.

Until recently, all cross section measurements were performed with histograms. In particular, the target spectra and experimental observations were binned and the unfolding problem is recast in the language of linear algebra. That is, one would like to determine the signal strength, defined as the ratio of the observed signal yield to the theoretical prediction, for each bin based on the measurements from experimental observations. This approach comes with the limitation that the binning must be determined beforehand. This makes it difficult to compare measurements with different binning. Furthermore, the optimal binning depends on the downstream inference task.

Modern machine learning (ML) has enabled the creation of unfolding methods that can process unbinned data [181]. Deep generative models such as Generative Adversarial Networks (GAN) [182–184] and Variational Autoencoders (VAE) [185, 186] produce implicit models that represents the probability density of the unfolded result and allow to sample from the probability density. Methods based on Normalizing Flows (NF) [187–190] allow for both sampling and density estimation. In contrast, the classifier-based method OmniFold Refs. [191, 192] iteratively

reweights a simulated dataset. A summary of machine learning-based unfolding methods can be found in Ref. [181] and recent applications of these techniques (in particular, of OmniFold) to experimental data are presented in Refs. [193–196]. While powerful, none of these approaches can simultaneously estimate cross sections and fit (nuisance) parameters. This can be a significant shortcoming when the phase space region being probed has non-trivial constraining power for systematic uncertainties.

Unfolding methods that can also profile have been proposed. One possibility is to treat the cross section in each region of particle-level phase space (i.e. in a histogram bin) as a free parameter and then perform a likelihood fit as for any set of parameters of interest and nuisance parameters. For example, this is the setup of the the Simplified Template Cross Section (STXS) (e.g. Refs. [52,197–199]) measurements for Higgs boson kinematic properties. Another possibility is Fully Bayesian Unfolding (FBU) [200], which samples from the posterior probability over the cross section in each bin of the particle-level phase space and over the nuisance parameters. All of these methods require binning.

We are thus motivated to propose a new machine learning-based unfolding method that is both unbinned at particle level and can profile, referred to as Unbinned Profiled Unfolding (UPU). UPU reuses all the standard techniques used in binned maximum likelihood unfolding and combines them with ML methods that allow for unbinned unfolding. Specifically, we use the binned maximum likelihood at detector level as the metric to optimize the unfolding, while the unfolding takes unbinned particle-level simulations as inputs.

The rest of this chapter is organized as follows. In Sec. 8.2, we describe the procedure and implementation details of UPU. We then present simple Gaussian examples to demonstrate the usage of UPU in Sec. 8.3. In Sec. 8.4, we apply UPU to a simulated Higgs boson cross section measurement at the Large Hadron Collider (LHC). The conclusions and outlook are then given in Sec. 8.5.

## 8.2 Unbinned Profiled Unfolding

### 8.2.1 Statistical Setup

UPU generalizes binned maximum likelihood unfolding to the unbinned case. Binned maximum likelihood unfolding can be described by the following optimization setup:

$$(\hat{k}, \hat{\theta}) = \operatorname{argmax}_{(k, \theta)} \Pr(m|k, \theta) p_0(\theta), \quad (8.1)$$

where  $m \in \mathbb{R}^{N_m}$  is a vector representing the counts in each of the  $N_m$  bins at detector level,  $k \in \mathbb{R}^{N_k}$  is a vector representing the counts in each of the  $N_k$  bins at particle level (usually  $N_m \geq N_k$ ),  $\theta$  are the nuisance parameters, and  $p_0$  is the prior on  $\theta$ . Our idea is to keep the structure of Eq. 8.1, but replace  $k$  with an unbinned estimator of the particle-level spectrum. Suppose that the particle-level phase space is<sup>1</sup>  $\mathbb{R}^N$  and let<sup>2</sup>  $\tau[\omega] \in \mathbb{R}^{\mathbb{R}^N}$  parameterize the probability density over this space for parameters  $\omega$ . The goal of UPU is then to optimize

$$(\hat{\omega}, \hat{\theta}) = \operatorname{argmax}_{(\omega, \theta)} \Pr(m|\tau[\omega], \theta) p_0(\theta), \quad (8.2)$$

where the final result would be given by  $\tau[\hat{\omega}]$ . The challenge with the construction in Eq. 8.2 is that for a given truth spectrum  $\tau[\omega]$ , we need to know the predicted detector-level distribution. In the binned case, this is readily computed by multiplying  $k$  by the response matrix  $R_{ij} = \Pr(\text{measure in bin } i | \text{truth is bin } j)$ . When the truth are unbinned, we need the full detector response. This is never known analytically and would be challenging to estimate numerically with a surrogate density estimator<sup>3</sup>. To address this challenge, we make use of the fact that simulated events come in pairs, with a matching between particle-level and detector-level events. Instead of estimating  $\tau$  directly, we use a fixed simulation (with particle-level spectrum  $\tau[\omega_0]$ ) and then learn a reweighting function  $w_0[\lambda]$  to estimate the likelihood ratio between the unfolded result and the fixed simulation at particle level. Schematically:

$$(\hat{\lambda}, \hat{\theta}) = \operatorname{argmax}_{(\lambda, \theta)} \Pr(m|\tau[\omega_0]w_0[\lambda]), \theta) p_0(\theta), \quad (8.3)$$

---

<sup>1</sup> Assuming the space is suitably standardized to remove units. <sup>2</sup> We will use  $[\cdot]$  to denote the parameters of the function and  $(\cdot)$  to denote the inputs of the function, e.g.  $f[\theta](x)$  is a functional in  $\theta$  and a function in  $x$ . <sup>3</sup> Note that Eq. 8.2 is a probability distribution over probability distributions so building it from the per-event detector response is non-trivial.

where in practice, we only have samples from  $\tau[\omega_0]$  and  $w_0$  is a surrogate model. The number of predicted events in a given bin  $i$  is then a sum over weights  $w_0[\hat{\lambda}]$  (evaluated at particle-level) for simulated events with a detector-level value in bin  $i$ . The probability over values  $m$  is then a product over Poisson probability mass functions, since the bins are statistically independent. The fact that the probability mass is known is crucial and means that UPU does not readily generalize the case where the detector-level phase space is also unbinned.

## 8.2.2 Machine Learning Approach

For particle-level features  $T$  and detector-level features  $R$ , the main goal is to train the likelihood ratio estimator  $w_0(T)$ , which reweights the simulated particle-level spectrum. In the absence of profiling, this corresponds to the following loss function:

$$L = \prod_{i=1}^{n_{\text{bins}}} \Pr(n_i \mid \sum_{j=1}^{n_{\text{MC}}} w_0(T_j) \mathbb{I}_i(R_j)), \quad (8.4)$$

where  $n_i$  is the number of observed events in bin  $i$ ,  $n_{\text{MC}}$  is the number of simulated events, and  $\mathbb{I}_i(\cdot)$  is the indicator function that is one when  $\cdot$  is in bin  $i$  and zero otherwise. When  $w_0$  is parameterized as a neural network (see Sec. 8.2.3), then the logarithm of Eq. 8.4 is used for training:

$$\log L = \sum_{i=1}^{n_{\text{bins}}} \left[ n_i \log \left( \sum_{j=1}^{n_{\text{MC}}} w_0(T_j) \mathbb{I}_i(R_j) \right) - \sum_{i=1}^{n_{\text{MC}}} w_0(T_j) \mathbb{I}_i(R_j) \right], \quad (8.5)$$

where we have dropped constants that do not affect the optimization. Experimental nuisance parameters modify the predicted counts in a particular bin given the particle-level counts. We account for these effects with a second reweighting function:

$$w_1(R|T, \theta) = \frac{p_\theta(R|T)}{p_{\theta_0}(R|T)}, \quad (8.6)$$

where  $p_\theta(R|T)$  is the conditional probability density of  $R$  given  $T$  with nuisance parameters  $\theta$ . Importantly,  $w_1$  does not modify the target particle level distribution. Incorporating  $w_1$  into the log likelihood results in the full loss function:

$$\log L = \sum_{i=1}^{n_{\text{bins}}} \left[ n_i \log \left( \sum_{j=1}^{n_{\text{MC}}} w_0(T_j) w_1(R_j|T_j, \theta) \mathbb{I}_i(R_j) \right) - \sum_{j=1}^{n_{\text{MC}}} w_0(T_j) w_1(R_j|T_j, \theta) \mathbb{I}_i(R_j) \right] + \log p_0(\theta). \quad (8.7)$$

Since  $w_1$  does not depend on the particle-level spectrum, it can be estimated prior to the final fit and only the parameters of  $w_0$  and the value(s) of  $\theta$  are allowed to float when optimizing Eq. 8.7.

### 8.2.3 Machine Learning Implementation

In our subsequent case studies, the reweighting functions  $w_0$  and  $w_1$  are parametrized with neural networks. The  $w_0$  function is only constrained to be non-negative and so we choose it to be the exponential of a neural network.

The pre-training of  $w_1$  requires neural conditional reweighting [201], as a likelihood ratio in  $R$  conditioned on  $T$  and parameterized in  $\theta$ . While there are multiple ways of approximating conditional likelihood ratios, the one we found to be the most stable for the examples we have studied for UPU is the product approach:

$$w_1(R|T, \theta) = \left( \frac{p_\theta(R, T)}{p_{\theta_0}(R, T)} \right) \left( \frac{p_{\theta_0}(T)}{p_\theta(T)} \right), \quad (8.8)$$

where the two terms on the righthand side are separately estimated and then their product is  $w_1$ . For a single feature  $T$ , a likelihood ratio between samples drawn from a probability density  $p$  and samples drawn from a probability density  $q$  is estimated using the fact that machine learning-classifiers approximate monotonic transformations of likelihood ratios (see e.g. Ref. [202, 203]). In particular, we use the standard binary cross entropy loss function

$$L_{\text{BCE}}[f] = - \sum_{Y \sim p} \log(f(Y)) - \sum_{Y \sim q} \log(1 - f(Y)), \quad (8.9)$$

and then the likelihood ratio is estimated as  $f/(1-f)$ . The last layer of the  $f$  networks are sigmoids in order to constrain their range to be between 0 and 1. The function  $f$  is additionally trained to be parameterized in  $\theta$  by training with pairs  $(Y, \Theta)$  instead of just  $Y$ , where  $\Theta$  is a random variable corresponding to values  $\theta$  sampled from a prior. We will use a uniform prior when training the parameterized classifiers.

All neural networks are implemented using PyTorch [204] and optimized with Adam [205] with a learning rate of 0.001 and consist of three hidden layers with 50 nodes per layer. All intermediate layers use ReLU activation functions. Each network is trained for 10000 epochs with early stopping using a patience of 10. The  $w_1$  training uses a batch size of 100,000. The  $w_0$  network is simultaneously optimized with  $\theta$  and uses a batch size that is the full dataset, which corresponds to performing the fit in Eq. 8.7 over all the data.

## 8.3 Gaussian Examples

### 8.3.1 One-dimension in both particle and detector level

We now demonstrate the proposed method with simple numerical examples. Here, each data set represents one-dimensional Gaussian random variables in both the particle and detector level. The particle-level random variable  $T$  is described by mean  $\mu$  and standard deviation  $\sigma$ , while the detector-level variable is given by

$$R = T + Z, \quad (8.10)$$

where  $Z$  is a Gaussian random variable with mean  $\beta$  and standard deviation  $\epsilon$ .

In a first example,  $\epsilon$  is considered to be the only nuisance parameter, and  $\beta$  is fixed to 0. Three data sets are prepared for the full training procedure. The first data set  $D_{\text{sim}}^{1.0}$  is used as the nominal simulation sample, which contains 200,000 events with  $\mu = 0$ ,  $\sigma = 1$  and  $\epsilon = 1$ . The second data set  $D_{\text{obs}}$  is used as the observed data, which contains 100,000 events with  $\mu = 0.2$ ,  $\sigma = 1$  and  $\epsilon = 1.2$ . To train the  $w_1$  reweighter, the third data set  $D_{\text{sim}}^*$ , which contains 200,000 events with  $\mu = 0$ ,  $\sigma = 1$  and  $\epsilon$  uniformly distributed from 0.2 to 1.8, is prepared. In addition, another data set  $D_{\text{sim}}^{1.2}$  of 100,000 events with  $\mu = 0$ ,  $\sigma = 1$  and  $\epsilon = 1.2$  is produced for validating the  $w_1$  reweighter. All data sets used in the training procedure are split to 50% for training and 50% for validating.

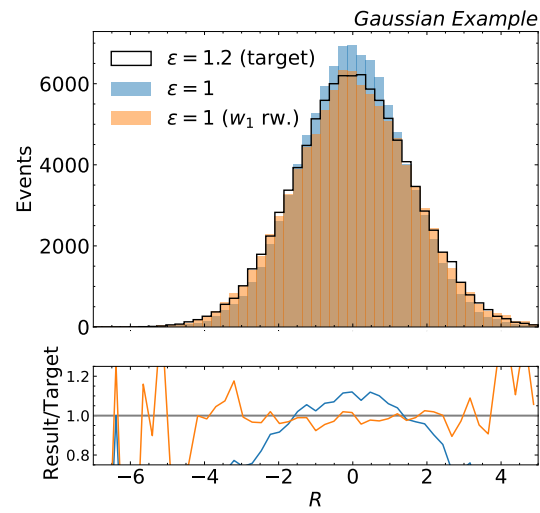


Figure 8.1: Gaussian 1D example: the nominal  $R$  distribution ( $\epsilon = 1$ ) is reweighted by the trained  $w_1$  conditioned at  $\epsilon = 1.2$  and compared to  $R$  distribution with  $\epsilon = 1.2$ .

A  $w_1$  reweighter is trained to reweight  $D_{\text{sim}}^{1.0}$  to  $D_{\text{sim}}^*$ . The trained  $w_1$  is then tested with the nominal  $R$  distribution ( $D_{\text{sim}}^{1.0}$ ) reweighted to  $\epsilon = 1.2$  ( $w_1(R|T, \epsilon = 1.2)$ ) and compared to the  $R$  spectrum with  $\epsilon = 1.2$  ( $D_{\text{sim}}^{1.2}$ ). As shown in Fig. 8.1, the trained  $w_1$  reweighter has learned to reweight the nominal  $R$  spectrum to match the  $R$  spectrum with  $\epsilon$  at 1.2.

With this trained  $w_1$  reweighter, a  $w_0$  reweighter is trained using  $D_{\text{sim}}^{1.0}$  as the simulation template with  $D_{\text{obs}}$  as the observed data used in Eq. 8.7. In the first scenario, the nuisance parameter  $\epsilon$  for the  $w_1$  reweighter is fixed to 1.2, and the penalty term in Eq. 8.7  $\log(\theta)$  is set to 0 (no constraint). As shown in Fig. 8.2, the  $w_0$  reweighter is able to learn to reweight the particle-level spectrum  $T$  by matching the detector-level spectrum  $R$  to the observed spectrum. In the second scenario, the nuisance parameter  $\epsilon$  is trained together with the  $w_0$  reweighter. The prior in the penalty term in Eq. 8.7 is set to be a Gaussian probability density with a 80% uncertainty. As shown in Fig. 8.3, the trained  $w_0$  and optimized  $\epsilon$  are tested. The fitted  $\epsilon$  is  $1.03 \pm 0.016$ <sup>4</sup> (true value is 1.2). The reweighted distribution matches well with observed data in the detector-level spectrum but the particle-level spectrum has a large non-closure. This is because of the degeneracy between the  $w_0$  and  $w_1$  reweighters in the effect on the detector-level spectrum. In other words, detector effects can mimic changes in the particle-level cross section, so the data cannot distinguish between these two scenarios. This is a common issue which also exists in the standard binned maximum likelihood unfolding. For comparison, we also perform the standard binned maximum likelihood unfolding. As shown in App. A, the unfolded  $T$  spectrum in this case also fails to represent the true  $T$  spectrum. An 80% uncertainty is highly exaggerated from typical scenarios, but it clearly illustrates the challenge of profiling and unfolding at the same time (see Sec. 8.5 for a discussion about regularization).

---

<sup>4</sup> The fitted value is averaged over five different  $w_0$  reweighters which are trained in the same way, but with different random initializations. The standard deviation of the fitted values is taken as the error.

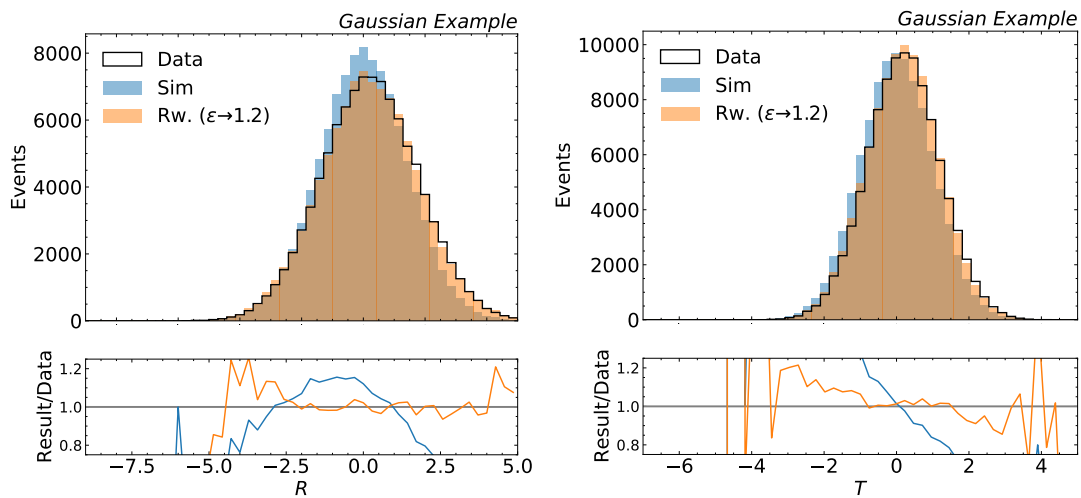


Figure 8.2: Gaussian 1D example: results of the  $w_0$  optimization. The nuisance parameter  $\epsilon$  is fixed to 1.2, and the the penalty term in Eq. 8.7 is set to 0. (Top) The detector-level spectrum  $R$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0 \times w_1$ , compared to the  $R$  spectrum of the observed data  $D_{\text{obs}}$ . (Bottom) The particle-level spectrum  $T$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0$ , compared to the  $T$  spectrum of the observed data  $D_{\text{obs}}$ .

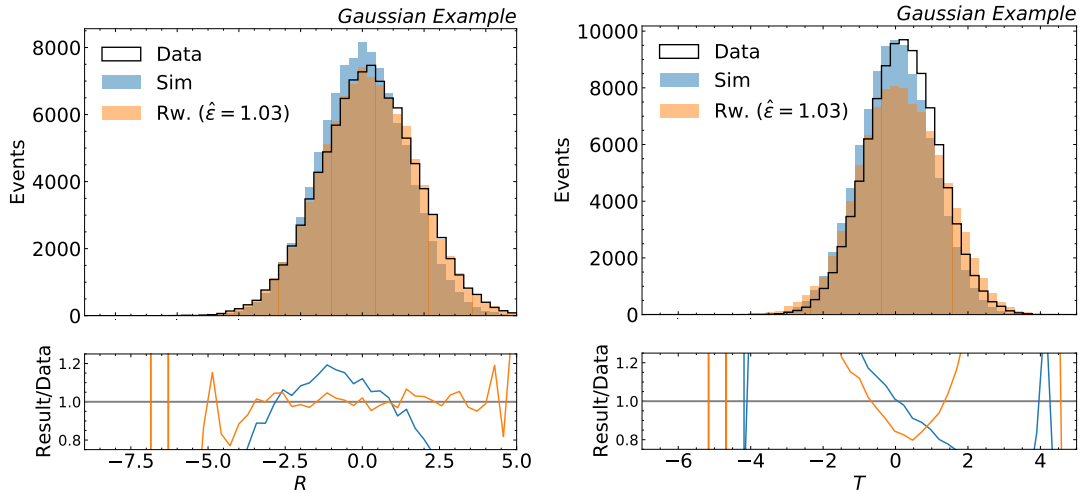


Figure 8.3: Gaussian 1D example: results of the  $w_0$  optimization. The nuisance parameter  $\epsilon$  is optimized simultaneously with  $w_0$  and the best-fit value is  $\hat{\epsilon} = 1.03 \pm 0.016$ . (Left) The detector-level spectrum  $R$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0 \times w_1$ , compared to the  $R$  spectrum of the observed data  $D_{\text{obs}}$ . (Right) The particle-level spectrum  $T$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0$ , compared to the  $T$  spectrum of the observed data  $D_{\text{obs}}$ .

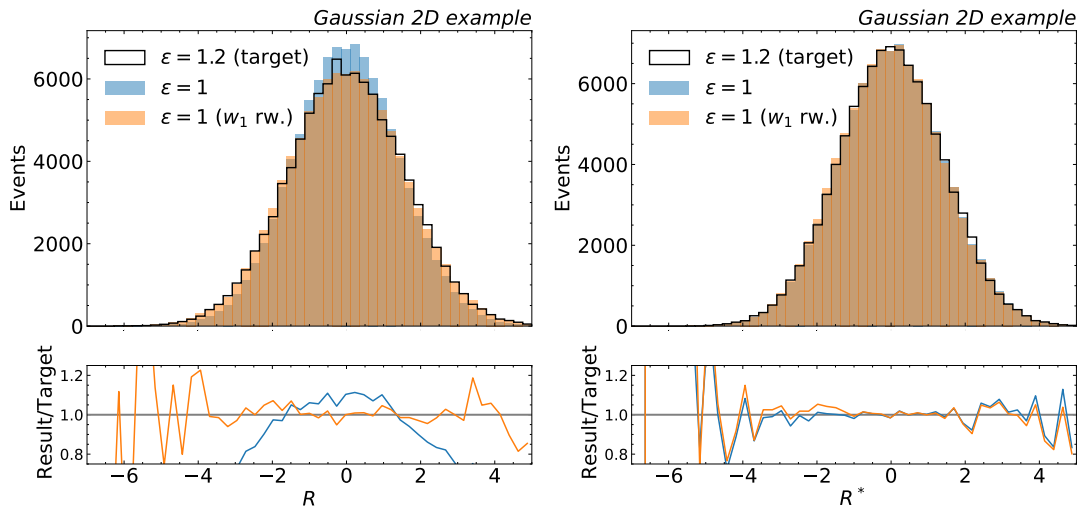


Figure 8.4: Gaussian 2D example: the nominal detector-level spectra  $R$  (left) and  $R^*$  (right) with  $\epsilon = 1$  reweighted by the trained  $w_1$  conditioned at  $\epsilon = 1.2$  and compared to the spectra with  $\epsilon = 1.2$ .

### 8.3.2 One-dimension in particle level and two-dimension in detector level

To break the degeneracy between the  $w_0$  and  $w_1$  reweighters, we now consider a two-dimension distribution in the detector level, which is given by

$$R = T + Z, \quad (8.11)$$

$$R^* = T + Z^*, \quad (8.12)$$

where  $Z$  ( $Z^*$ ) is a Gaussian random variable with mean  $\beta$  ( $\beta^*$ ) and standard deviation  $\epsilon$  ( $\epsilon^*$ ).  $\epsilon$  is considered to be the only nuisance parameter, and  $\beta$ ,  $\beta^*$  are fixed to 0, and  $\epsilon^*$  is fixed to 1. In this case, the nuisance parameter  $\epsilon$  only has effect on the  $R$  spectrum and the  $R^*$  spectrum depends purely on the particle-level spectrum  $T$ .

Similar to the previous example,  $D_{\text{sim}}^{1.0}$  is used as the nominal simulation sample, which contains 200,000 events with  $\mu = 0$ ,  $\sigma = 1$  and  $\epsilon = 1$ .  $D_{\text{obs}}$  is used as the observed data, which contains 100,000 events with  $\mu = 0.8$ ,  $\sigma = 1$  and  $\epsilon = 1.2$ . To train the  $w_1$  reweighter,  $D_{\text{sim}}^*$ , which contains 200,000 events with  $\mu = 0$ ,  $\sigma = 1$  and  $\epsilon$  uniformly distributed from 0.2 to 1.8, is prepared. In addition, another data set  $D_{\text{sim}}^{1.2}$  of 100,000 events with  $\mu = 0$ ,  $\sigma = 1$  and  $\epsilon = 1.2$  is produced for validating the  $w_1$  reweighter. All data sets used in the training procedure are split to 50% for training and 50% for validating.

A  $w_1$  reweighter is trained to reweight  $D_{\text{sim}}^{1.0}$  to  $D_{\text{sim}}^*$ . The trained  $w_1$  is tested with the nominal  $R$  and  $R^*$  spectra ( $D_{\text{sim}}^{1.0}$ ) reweighted to  $\epsilon = 1.2$  and compared to the  $R$  and  $R^*$  spectra with  $\epsilon = 1.2$ . As shown in Fig. 8.4, the trained  $w_1$  reweighter has learned to reweight the nominal  $R$  spectrum to match the  $R$  spectrum with  $\epsilon$  at 1.2, and  $R^*$  is independent of the  $w_1$  reweighter.

Based on the trained  $w_1$  reweighter, a  $w_0$  reweighter and the nuisance parameter  $\epsilon$  are optimized simultaneously using  $D_{\text{sim}}$  as the simulation template with  $D_{\text{obs}}$  as the observed data used in Eq. 8.7. As before, the prior in the penalty term in Eq. 8.7 is configured with an uncertainty of 80%. The fitted  $\epsilon$  is  $1.20 \pm 0.004$  (correct value is 1.2). As shown in Fig. 8.5, the reweighted spectra match well with observed data in both detector and particle level. For more realistic uncertainties (so long as the simulation is close to the right answer), the fidelity is even better.

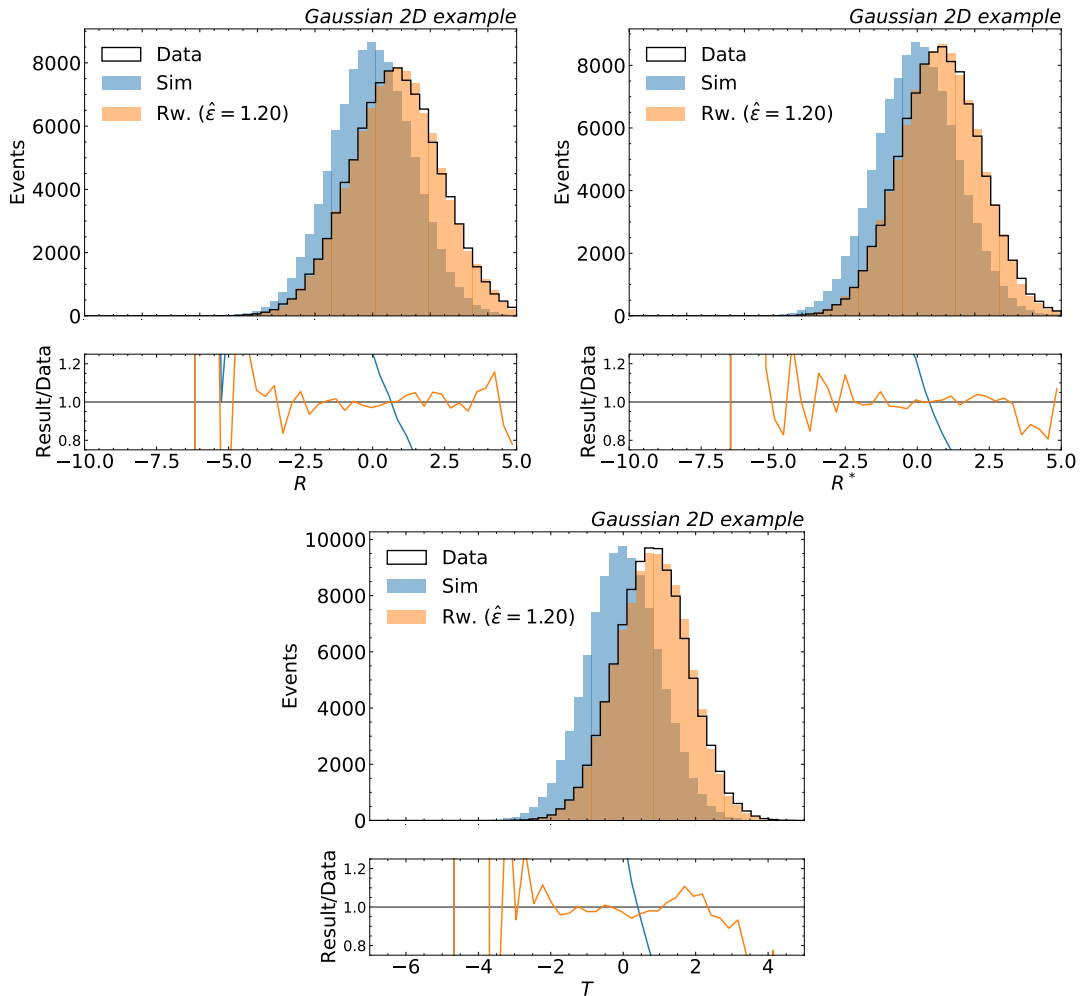


Figure 8.5: Gaussian 2D example: results of the  $w_0$  optimization. The nuisance parameter  $\epsilon$  is optimized simultaneously with  $w_0$  with the prior constraint set to 80%. The fitted  $\epsilon$  is  $1.20 \pm 0.004$ . (Top-left) The detector-level spectrum  $R$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0 \times w_1$ , compared to the  $R$  spectrum of the observed data  $D_{\text{obs}}$ . (Top-right) The detector-level spectrum  $R'$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0 \times w_1$ , compared to the  $R^*$  spectrum of the observed data  $D_{\text{obs}}$ . (Bottom) The particle-level spectrum  $T$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0$ , compared to the  $T$  spectrum of the observed data  $D_{\text{obs}}$ .

## 8.4 Higgs Boson Cross Section

We now demonstrate the unfolding method in a physics case — a Higgs boson cross section measurement. Here, we focus on the di-photon decay channel of the Higgs boson. The goal is then to measure the transverse momentum spectrum of the Higgs boson  $p_H^T$  using the transverse momentum of the di-photon system  $p_{\gamma\gamma}^T$  at detector level. The photon resolution  $\epsilon_\gamma$  is considered as a nuisance parameter. In this case, the  $p_{\gamma\gamma}^T$  spectrum is minimally affected by  $\epsilon_\gamma$ . Therefore, we also consider the invariant mass spectrum of the di-photon system  $m_{\gamma\gamma}$  at detector level, which is highly sensitive to  $\epsilon_\gamma$ . In addition, In order to have a large spectrum difference between different data sets for demonstration purpose, we consider only events that contain at least two reconstructed jets, where the leading-order (LO) calculation would significantly differ from next-to-leading-order calculation (NLO)

Similar to the Gaussian examples, we prepare the following data sets:

- $D_{\text{obs}}$ : used as the observed data.
- $D_{\text{sim}}^{1.0}$ : used as the nominal simulation sample.
- $D_{\text{sim}}^{1.2}$ : used as the simulation sample with a systematic variation.
- $D_{\text{sim}}^*$ : simulation sample with various  $\epsilon_\gamma$  values for training the  $w_1$  reweighter.

$D_{\text{obs}}$  is generated at NLO using the POWHEGBOX program [206, 207], while the rest are generated at LO using MADGRAPH5\_aMC@LO v2.6.5 [97]. For all samples, the parton-level events are processed by PYTHIA 8.235 [48, 208] for the Higgs decay, the parton shower, hadronization, and the underlying event. The detector simulation is based on DELPHES 3.5.0 [209] with detector response modified from the default ATLAS detector card. For both  $D_{\text{obs}}$  and  $D_{\text{sim}}^{1.2}$ , the photon resolution  $\epsilon$  is multiplied by a factor of 1.2. For  $D_{\text{sim}}^*$ , the multiplier of  $\epsilon$  is uniformly scanned between 0.5 and 1.5 with a step size of 0.01.  $D_{\text{sim}}^{1.0}$  uses the default ATLAS detector card.

Each of the spectra of particle-level  $p_{\gamma\gamma}^T$ , detector-level  $p_{\gamma\gamma}^T$  and detector-level  $m_{\gamma\gamma}$  is standardized to the spectrum with a mean of 0 and a standard deviation of 1 before being passed to the neural networks. A  $w_1$  reweighter is then trained to reweight  $D_{\text{sim}}^{1.0}$  to  $D_{\text{sim}}^*$ . The trained  $w_1$  is tested with the nominal detector level  $p_{\gamma\gamma}^T$  and  $m_{\gamma\gamma}$  spectra ( $D_{\text{sim}}^{1.0}$ ) reweighted to  $\epsilon_\gamma = 1.2$  and compared

to the detector level  $p_{\gamma\gamma}^T$  and  $m_{\gamma\gamma}$  spectra with  $\epsilon_\gamma = 1.2$ . As shown in Fig. 8.6, the trained  $w_1$  reweighter has learned to reweight the nominal detector level  $m_{\gamma\gamma}$  spectrum to match the detector level  $m_{\gamma\gamma}$  spectrum with  $\epsilon_\gamma$  at 1.2, and the detector level  $p_{\gamma\gamma}^T$  variable is independent of the  $w_1$  reweighter.

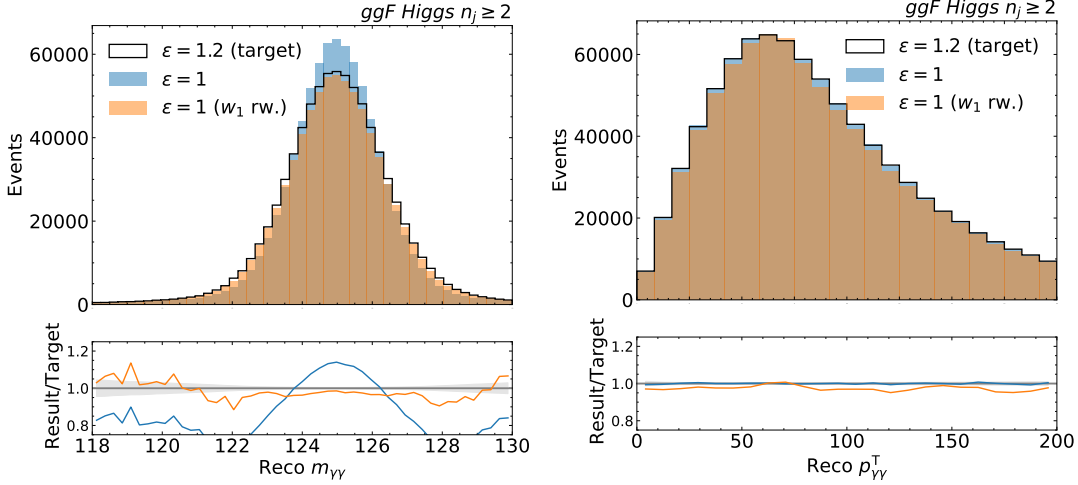


Figure 8.6: Higgs boson cross section: the nominal detector-level spectra  $m_{\gamma\gamma}$  (left) and  $p_{\gamma\gamma}^T$  (right) with  $\epsilon_\gamma = 1$  reweighted by the trained  $w_1$  conditioned at  $\epsilon_\gamma = 1.2$  and compared to the spectra with  $\epsilon_\gamma = 1.2$ .

The  $w_0$  reweighter and  $\epsilon$  are optimized simultaneously based on the pre-trained  $w_1$  reweighter. The prior of  $\epsilon_\gamma$  is 50%. The fitted  $\epsilon_\gamma$  is  $1.19 \pm 0.007$ . As shown in Fig. 8.7, the reweighted spectra match well with observed data in both detector and particle level. This means that the observed data  $p_H^T$  spectrum is successfully unfolded with nuisance parameter  $\epsilon_\gamma$  properly profiled. For comparison, we also perform UPU with  $\epsilon_\gamma$  fixed at 1. As shown in Fig. 8.7, the unfolded  $p_H^T$  spectrum in this case has a larger non-closure with the observed data due to the lack of profiling.

## 8.5 Summary

UPU uses the binned maximum likelihood as the figure of merit to optimize the unfolding reweighting function  $w_0(t)$ , which takes unbinned particle-level spectra as inputs.  $w_0(t)$  and the nuisance parameters  $\theta$  are optimized simultaneously, which also requires to learn a conditional

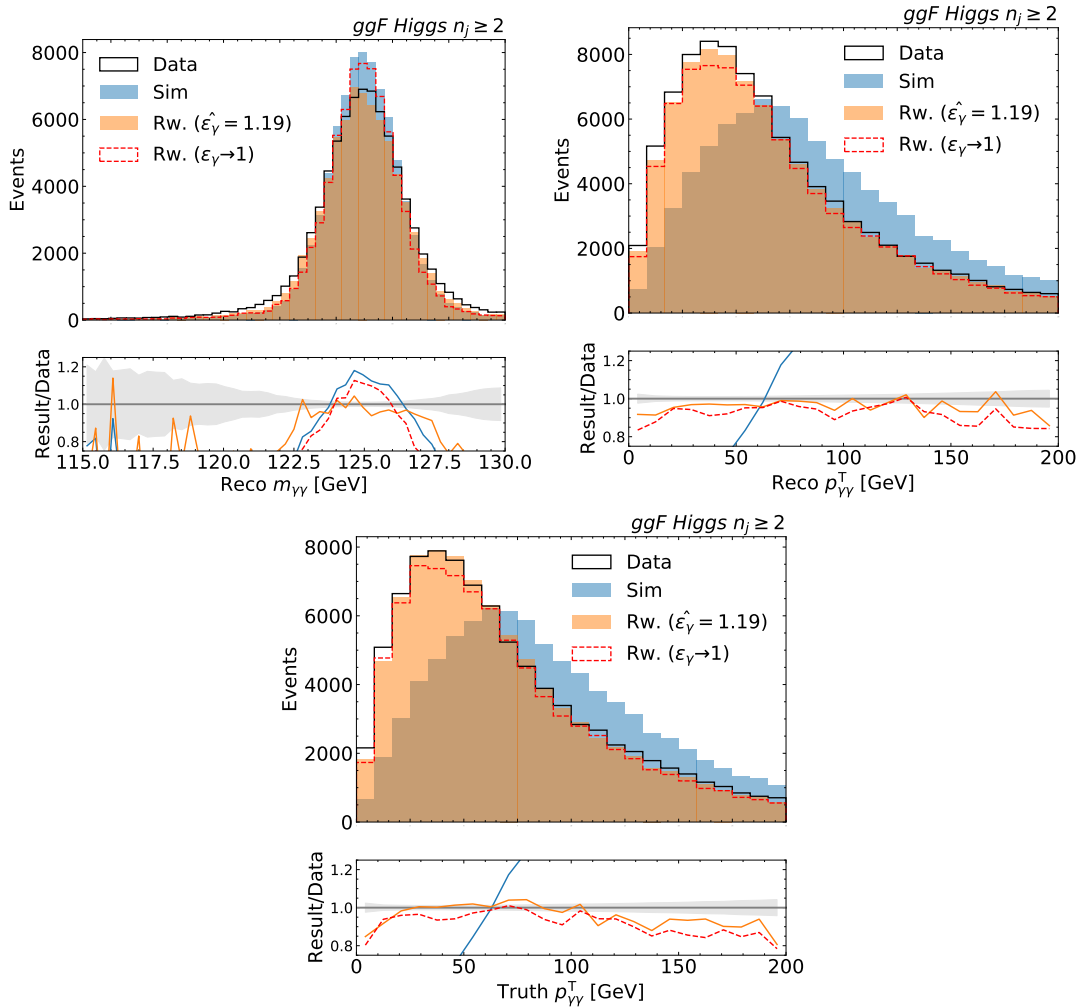


Figure 8.7: Higgs boson cross section: results of the  $w_0$  optimization. The nuisance parameter  $\epsilon_\gamma$  is optimized simultaneously with  $w_0$  with the prior constraint set to 50% (orange) or fixed to 1 for comparison (red). The fitted  $\epsilon_\gamma$  is  $1.19 \pm 0.007$ . (Top-left) The detector-level spectrum  $m_{\gamma\gamma}$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0 \times w_1$ , compared to the  $m_{\gamma\gamma}$  spectrum of the observed data  $D_{\text{obs}}$ . (Top-right) The detector-level spectrum  $p_{\gamma\gamma}^T$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0 \times w_1$ , compared to the  $p_{\gamma\gamma}^T$  spectrum of the observed data  $D_{\text{obs}}$ . (Bottom) The particle-level spectrum  $p_{\gamma\gamma}^T$  of the simulation template  $D_{\text{sim}}$  reweighted by the trained  $w_0$ , compared to the  $p_{\gamma\gamma}^T$  spectrum of the observed data  $D_{\text{obs}}$ .

likelihood ratio  $w_1(t, r|\theta)$  that reweights the detector-level spectra based on the profiled values of nuisance parameters and is taken as an input for the optimization of  $w_0(t)$  and  $\theta$ .

In Gaussian examples, we demonstrated the optimization of  $w_1$  and the optimization of  $w_0$  and  $\theta$ . A limitation of this method can be seen when the considered detector-level observable, a one-dimension Gaussian distribution, is not able to distinguish between effects from particle level and effects from detector level with  $\theta$ . This limitation also exists in the standard binned maximum likelihood unfolding, and the problem can be resolved when we consider another observable which does not depend on  $\theta$  and thus breaks the degeneracy between particle-level and detector-level effects. Regularization can also help mitigate these effects. Our approach has implicit regularization from the smoothness of neural networks. We leave additional studies of regularization to future work.

We also applied UPU to the Higgs boson cross section measurement. We considered one dimension at particle level and two dimensions at detector level. With one detector-level variable sensitive to the target particle-level observable and one sensitive to the effect of nuisance parameters, the data are successfully unfolded and profiled. The impact of profiling is also demonstrated by comparing with the result of nuisance parameter fixed to the nominal value. This can be readily extended to higher dimensions in either particle level or detector level, provided all particle-level and detector-level effects are distinguishable in the considered detector-level spectra. In the case of more than one nuisance parameters, one can either train multiple  $w_1$  for each nuisance parameter separately or train a single  $w_1$  which takes all nuisance parameters as inputs. As the effects of multiple nuisance parameters are usually assumed independent, one could take a product of individually trained reweighters.

As with any measurement, quantifying the uncertainty is critical to interpret UPU results. Just as in the binned case, one can calculate the uncertainty on the nuisance parameters which can be determined by fixing a given parameter to target values and then simultaneously re-optimizing  $w_0$  and the rest of the nuisance parameters. A new feature of UPU is that the likelihood (ratio) itself is only an approximation, using neural networks as surrogate models. This is a challenge for all machine learning-based unfolding, and uncertainties can be probed by comparing the results with different simulations. Future extensions of UPU may be able to also use machine learning to quantify these model uncertainties as well as process unbinned data also at detector level.

## Chapter 9

### Conclusion

With  $139 \text{ fb}^{-1}$  of  $pp$  collision data collected at 13 TeV from 2015 to 2018 by the ATLAS detector at the LHC, two searches involving Higgs boson which use the proton-proton ( $p - p$ ) collision data collected by the ATLAS detector are conducted. The first search is a search for dimuon decay of the SM Higgs boson ( $H \rightarrow \mu\mu$ ), which is the most promising channel for probing the couplings between SM Higgs Boson and the second generation fermions at the LHC. Boosted decision trees are trained to separate signal and background. The observed (expected) significance over the background-only hypothesis for a Higgs boson with a mass of 125.09 GeV is  $2.0\sigma$  ( $1.7\sigma$ ), and the observed upper limit on the cross section times branching ratio for  $pp \rightarrow H \rightarrow \mu\mu$  is 2.2 times the SM prediction at 95% confidence level, while the expected limit on a  $H \rightarrow \mu\mu$  signal assuming the absence (presence) of a SM signal is 1.1 (2.0). The best-fit value of the signal strength parameter, defined as the ratio of the observed signal yield to the one expected in the SM, is  $\mu = 1.2 \pm 0.6$ . This result represents an improvement of about a factor of 2.5 in expected sensitivity compared with the previous ATLAS publication. A factor of two from the improvement is due to the larger analysed dataset and the additional 25% improvement can be attributed to more advanced analysis techniques.

The second search targets events that contain large missing transverse momentum and either two  $b$ -tagged small-radius jets or a single large-radius jet associated with two  $b$ -tagged subjects, which is sensitive to Dark Matter models in the extended Higgs sector. Events are split into multiple categories which target different phase space of the DM signals. No significant excess from the SM prediction is observed. The results are interpreted with two benchmark models with two Higgs doublets extended by either a heavy vector boson  $Z'$  ( $Z'$ -2HDM) or a pseudoscalar singlet  $a$  (2HDM+ $a$ ) and which provide a Dark Matter candidate  $\chi$ . In the case of  $Z'$ -2HDM, the observed

limits extend up to a  $Z'$  mass of 3.1 TeV at 95 % confidence level for a mass of 100 GeV for the Dark Matter candidate. For 2HDM+ $a$ , masses of  $a$  are excluded up to 520 GeV and 240 GeV for  $\tan\beta = 1$  and  $\tan\beta = 10$  and a Dark Matter mass of 10 GeV, respectively. This analysis achieved a large improvement in the sensitivity for  $Z'$ -2HDM as well as 2HDM+ $a$  with the  $ggF$  production mode due to the  $E_T^{\text{miss}}$  binning in the merged region. This analysis also established the exclusion limits for  $b\bar{b}A$  production of 2HDM+ $a$  for the first time thanks to the inclusion of the  $\geq 3$   $b$ -tag regions. In addition, limits on the visible cross sections in the model-independent context are set, which will be useful for the future reinterpretation with additional signal models.

Both searches will benefit from the additional data that will be collected in the next run of LHC, which could increase statistical power and enhance the sensitivity. In addition to the higher statistics, there are many developments and potential tasks which will bring further improvements in the future analyses. For example, there have been a significant improvement in the LAr calorimeter, which increases the granularity of information used by the trigger and will increase the  $E_T^{\text{miss}}$  trigger efficiency. This will enhance the signal efficiency of the mono- $h(b\bar{b})$  search and thus increase the chance of finding signals. mono- $h(b\bar{b})$  can also benefit from adopting an ML-based categorization. In particular, it is worth applying a parametrized neural network to accommodate to different signals or signals in different parameter phase space. On the other hand, we expect further improvements in  $H \rightarrow \mu\mu$  replacing the BDT with deep neural network for the categorization. In particular, deep sets architecture can preserve jet symmetry during the classifier training, and adversarial neural network can allow fully exploiting powerful training variables (e.g. kinematics of individual muons) while maintaining mass independence. All of these items will be crucial for pushing the analyses towards potential discovery.

In addition to the two searches presented, the unbinned profiled unfolding (UPU), a newly proposed unfolding method, is introduced. This method combines the standard binned maximum likelihood unfolding and the machine learning-based methods, resulting in unbinned differential cross-section and is also able to profile nuisance parameters. This method significantly increases the flexibility and reusability of the unfolding result and will be particularly useful for model interpretations with any physics search channel in the future.

## APPENDIX

### Binned maximum likelihood unfolding with Gaussian examples

In this appendix, we present results of the standard binned maximum likelihood unfolding (BMLU) with Gaussian examples. The scenarios are:

- One-dimension in both particle and detector level: this is the same example as described in Sec. 8.3.1. The prior constraint for  $\epsilon$  is set to 80%. The result is shown in Fig. A.1 with  $\epsilon$  fitted to  $1.08 \pm 0.02$ , which also indicates a degeneracy problem between particle and detector levels.
- One-dimension in particle level and two-dimension in detector level: this is the same example as described in Sec. 8.3.2. The prior constraint for  $\epsilon$  is set to 80%. The result is shown in Fig. A.2 with  $\epsilon$  fitted to  $1.19 \pm 0.003$ . The degeneracy problem is resolved after considering an additional spectrum in the detector level.

All the maximum likelihood fittings are performed using pyhf [210,211].

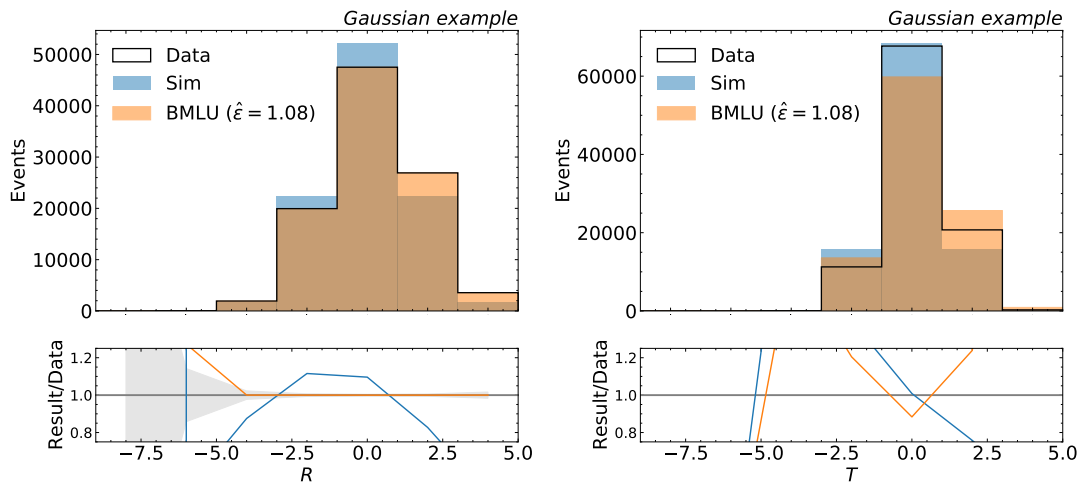


Figure A.1: Gaussian 1D example: results of the binned maximum likelihood unfolding. The prior constraint for  $\epsilon$  is set to 80% and the fitted  $\epsilon$  is  $1.08 \pm 0.02$ . (Left) The fitted detector-level spectrum  $R$  of the simulation template  $D_{\text{sim}}$ , compared to the  $R$  spectrum of the observed data  $D_{\text{obs}}$ . (Right) The unfolded particle-level spectrum  $T$  of the simulation template  $D_{\text{sim}}$ , compared to the  $T$  spectrum of the observed data  $D_{\text{obs}}$ .

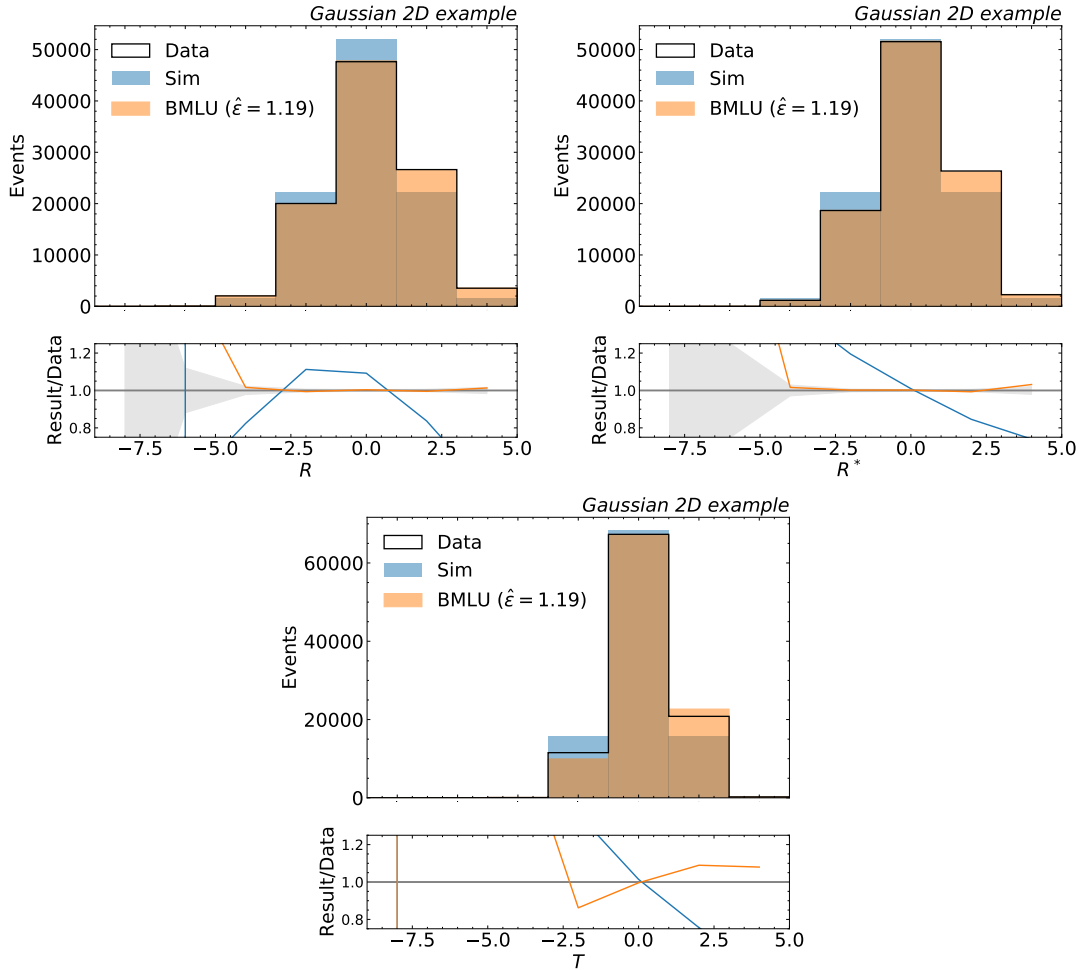


Figure A.2: Gaussian 2D example: results of the binned maximum likelihood unfolding. The prior constraint for  $\epsilon$  is set to 80% and the fitted  $\epsilon$  is  $1.19 \pm 0.003$ . (Top-left) The fitted detector-level spectrum  $R$  of the simulation template  $D_{\text{sim}}$ , compared to the  $R$  spectrum of the observed data  $D_{\text{obs}}$ . (Top-right) The fitted detector-level spectrum  $R^*$  of the simulation template  $D_{\text{sim}}$ , compared to the  $R'$  spectrum of the observed data  $D_{\text{obs}}$ . (Bottom) The unfolded particle-level spectrum  $T$  of the simulation template  $D_{\text{sim}}$ , compared to the  $T$  spectrum of the observed data  $D_{\text{obs}}$ .

## LIST OF REFERENCES

- [1] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323, doi:10.1103/PhysRevLett.13.321.
- [2] P. Higgs, “Broken symmetries, massless particles and gauge fields”, *Physics Letters* **12** (1964), no. 2, 132–133, doi:https://doi.org/10.1016/0031-9163(64)91136-9.
- [3] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, doi:10.1103/PhysRevLett.13.508.
- [4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587, doi:10.1103/PhysRevLett.13.585.
- [5] P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons”, *Phys. Rev.* **145** (May, 1966) 1156–1163, doi:10.1103/PhysRev.145.1156.
- [6] T. W. B. Kibble, “Symmetry Breaking in Non-Abelian Gauge Theories”, *Phys. Rev.* **155** (Mar, 1967) 1554–1561, doi:10.1103/PhysRev.155.1554.
- [7] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B* **716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [8] CMS Collaboration, “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”, *Phys. Lett. B* **716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [9] ATLAS Collaboration, “A particle consistent with the Higgs Boson observed with the ATLAS Detector at the Large Hadron Collider”, *Science* **338** (2012) 1576–1582, doi:10.1126/science.1232005.
- [10] ATLAS, CMS Collaboration, “Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments”, *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.
- [11] ATLAS Collaboration, “Observation and measurement of Higgs boson decays to  $WW^*$  with the ATLAS detector”, *Phys. Rev. D* **92** (2015), no. 1, 012006, doi:10.1103/PhysRevD.92.012006, arXiv:1412.2641.

- [12] ATLAS Collaboration, “Observation of  $H \rightarrow b\bar{b}$  decays and  $VH$  production with the ATLAS detector”, *Phys. Lett. B* **786** (2018) 59–86, doi:10.1016/j.physletb.2018.09.013, arXiv:1808.08238.
- [13] ATLAS Collaboration, “Cross-section measurements of the Higgs boson decaying into a pair of  $\tau$ -leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, *Phys. Rev. D* **99** (2019) 072001, doi:10.1103/PhysRevD.99.072001, arXiv:1811.08856.
- [14] ATLAS Collaboration, “Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector”, *Phys. Lett. B* **784** (2018) 173–191, doi:10.1016/j.physletb.2018.07.035, arXiv:1806.00425.
- [15] T. D. Lee, “A Theory of Spontaneous T Violation”, *Phys. Rev. D* **8** (1973) 1226–1239, doi:10.1103/PhysRevD.8.1226.
- [16] G. C. Branco et al., “Theory and phenomenology of two-Higgs-doublet models”, *Phys. Rept.* **516** (2012) 1–102, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.
- [17] G. Hinshaw et al., “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results”, *ApJS* **208** (October, 2013) 19, doi:10.1088/0067-0049/208/2/19, arXiv:1212.5226.
- [18] Planck Collaboration, “Planck 2018 results. I. Overview and the cosmological legacy of Planck”, *Astron. Astrophys.* **641** (2020) A1, doi:10.1051/0004-6361/201833880, arXiv:1807.06205.
- [19] V. Trimble, “Existence and Nature of Dark Matter in the Universe”, *Annual Review of Astronomy and Astrophysics* **25** (1987), no. 1, 425–472, doi:10.1146/annurev.aa.25.090187.002233, arXiv:https://doi.org/10.1146/annurev.aa.25.090187.002233.
- [20] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: Evidence, candidates and constraints”, *Phys. Rept.* **405** (2005) 279–390, doi:10.1016/j.physrep.2004.08.031, arXiv:hep-ph/0404175.
- [21] J. L. Feng, “Dark Matter Candidates from Particle Physics and Methods of Detection”, *Ann. Rev. Astron. Astrophys.* **48** (2010) 495–545, doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- [22] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation* **3** (aug, 2008) S08003–S08003, doi:10.1088/1748-0221/3/08/s08003.
- [23] ATLAS Collaboration Collaboration, “ATLAS Insertable B-Layer Technical Design Report Addendum”, technical report, (2012). Addendum to CERN-LHCC-2010-013, ATLAS-TDR-019.

- [24] ATLAS Collaboration, “A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector”, *Phys. Lett. B* **812** (2021) 135980, doi:10.1016/j.physletb.2020.135980, arXiv:2007.07830.
- [25] ATLAS Collaboration, “Muon reconstruction and identification efficiency in ATLAS using the full Run 2  $pp$  collision data set at  $\sqrt{s} = 13$  TeV”, *Eur. Phys. J. C* **81** (2021), no. 7, 578, doi:10.1140/epjc/s10052-021-09233-2, arXiv:2012.00578.
- [26] ATLAS Collaboration, “Search for dark matter produced in association with a Standard Model Higgs boson decaying into b-quarks using the full Run 2 dataset from the ATLAS detector”, *JHEP* **11** (2021) 209, doi:10.1007/JHEP11(2021)209, arXiv:2108.13391.
- [27] A. Berlin, T. Lin, and L.-T. Wang, “Mono-Higgs detection of dark matter at the LHC”, *Journal of High Energy Physics* **2014** (2014), no. 6, 78, doi:10.1007/JHEP06(2014)078.
- [28] LHC Dark Matter Working Group Collaboration, “LHC Dark Matter Working Group: Next-generation spin-0 dark matter models”, *Phys. Dark Univ.* **27** (2020) 100351, doi:10.1016/j.dark.2019.100351, arXiv:1810.09420.
- [29] M. Bauer, U. Haisch, and F. Kahlhoefer, “Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators”, *Journal of High Energy Physics* **2017** (2017), no. 5, 138, doi:10.1007/JHEP05(2017)138.
- [30] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508.
- [31] P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons”, *Phys. Rev.* **145** (1966) 1156–1163, doi:10.1103/PhysRev.145.1156.
- [32] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321.
- [33] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585.
- [34] T. W. B. Kibble, “Symmetry breaking in nonAbelian gauge theories”, *Phys. Rev.* **155** (1967) 1554–1561, doi:10.1103/PhysRev.155.1554.
- [35] H. E. Haber and G. L. Kane, “The Search for Supersymmetry: Probing Physics Beyond the Standard Model”, *Phys. Rept.* **117** (1985) 75–263, doi:10.1016/0370-1573(85)90051-1.
- [36] J. E. Kim, “Light Pseudoscalars, Particle Physics and Cosmology”, *Phys. Rept.* **150** (1987) 1–177, doi:10.1016/0370-1573(87)90017-2.
- [37] R. D. Peccei and H. R. Quinn, “CP Conservation in the Presence of Instantons”, *Phys. Rev. Lett.* **38** (1977) 1440–1443, doi:10.1103/PhysRevLett.38.1440.

- [38] M. Trodden, “Electroweak baryogenesis: A Brief review”, in *33rd Rencontres de Moriond: Electroweak Interactions and Unified Theories*, pp. 471–480, 1998. arXiv:hep-ph/9805252.
- [39] “LHC Machine”, *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.
- [40] CMS Collaboration, “The CMS Experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [41] S. Holmes, R. S. Moore, and V. Shiltsev, “Overview of the Tevatron Collider Complex: Goals, Operations and Performance”, *JINST* **6** (2011) T08001, doi:10.1088/1748-0221/6/08/T08001, arXiv:1106.0909.
- [42] ATLAS Collaboration, “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”, *Eur. Phys. J. C* **77** (2017) 490, doi:10.1140/epjc/s10052-017-5004-5, arXiv:1603.02934.
- [43] ATLAS Collaboration, “Performance of the ATLAS Trigger System in 2015”, *Eur. Phys. J. C* **77** (2017), no. 5, 317, doi:10.1140/epjc/s10052-017-4852-3, arXiv:1611.09661.
- [44] ATLAS Collaboration, “Performance of the ATLAS muon triggers in Run 2”, *JINST* **15** (2020), no. 09, P09015, doi:10.1088/1748-0221/15/09/p09015, arXiv:2004.13447.
- [45] ATLAS Collaboration, “Performance of electron and photon triggers in ATLAS during LHC Run 2”, *Eur. Phys. J. C* **80** (2020), no. 1, 47, doi:10.1140/epjc/s10052-019-7500-2, arXiv:1909.00761.
- [46] ATLAS Collaboration, “The ATLAS Simulation Infrastructure”, *Eur. Phys. J. C* **70** (2010) 823–874, doi:10.1140/epjc/s10052-010-1429-9, arXiv:1005.4568.
- [47] S. A. et al., “GEANT4: a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003), no. 3, 250–303, doi:https://doi.org/10.1016/S0168-9002(03)01368-8.
- [48] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [49] R. D. Ball et al., “Parton distributions with LHC data”, *Nucl. Phys. B* **867** (2013) 244–289, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.
- [50] ATLAS Collaboration Collaboration, “The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model”, technical report, CERN, Geneva, (2016). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-017>.
- [51] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”, doi:10.5170/CERN-2013-004, arXiv:1307.1347.

- [52] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”, doi:10.23731/CYRM-2017-002, arXiv:1610.07922.
- [53] A. Djouadi, J. Kalinowski, and M. Spira, “HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension”, *Comput. Phys. Commun.* **108** (1998) 56–74, doi:10.1016/S0010-4655(97)00123-9, arXiv:hep-ph/9704448.
- [54] M. Spira, “QCD effects in Higgs physics”, *Fortsch. Phys.* **46** (1998) 203–284, doi:10.1002/(SICI)1521-3978(199804)46:3<203::AID-PROP203>3.0.CO;2-4, arXiv:hep-ph/9705337.
- [55] A. Djouadi, M. M. Muhlleitner, and M. Spira, “Decays of supersymmetric particles: The Program SUSY-HIT (SUSpect-SdecaY-Hdecay-InTerface)”, *Acta Phys. Polon. B* **38** (2007) 635–644, arXiv:hep-ph/0609292.
- [56] A. Djouadi, J. Kalinowski, M. Mühlleitner, and M. Spira, “HDECAY: Twenty++ years after”, *Computer Physics Communications* **238** (2019) 214–231, doi:https://doi.org/10.1016/j.cpc.2018.12.010.
- [57] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, “Radiative corrections to the semileptonic and hadronic Higgs-boson decays  $H \rightarrow W W / Z Z \rightarrow 4$  fermions”, *JHEP* **02** (2007) 080, doi:10.1088/1126-6708/2007/02/080, arXiv:hep-ph/0611234.
- [58] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, “Precise predictions for the Higgs-boson decay  $H \rightarrow WW/ZZ \rightarrow 4$  leptons”, *Phys. Rev. D* **74** (2006) 013004, doi:10.1103/PhysRevD.74.013004, arXiv:hep-ph/0604011.
- [59] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, “Precision calculations for the Higgs decays  $H \rightarrow ZZ/WW \rightarrow 4$  leptons”, *Nucl. Phys. B Proc. Suppl.* **160** (2006) 131–135, doi:10.1016/j.nuclphysbps.2006.09.104, arXiv:hep-ph/0607060.
- [60] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [61] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [62] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [63] J. M. Campbell et al., “NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM”, *JHEP* **07** (2012) 092, doi:10.1007/JHEP07(2012)092, arXiv:1202.5475.

- [64] K. Hamilton, P. Nason, and G. Zanderighi, “MINLO: Multi-Scale Improved NLO”, *JHEP* **10** (2012) 155, doi:10.1007/JHEP10(2012)155, arXiv:1206.3572.
- [65] K. Hamilton, P. Nason, C. Oleari, and G. Zanderighi, “Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching”, *JHEP* **05** (2013) 082, doi:10.1007/JHEP05(2013)082, arXiv:1212.4504.
- [66] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, “NNLOPS simulation of Higgs boson production”, *JHEP* **10** (2013) 222, doi:10.1007/JHEP10(2013)222, arXiv:1309.0017.
- [67] K. Hamilton, P. Nason, and G. Zanderighi, “Finite quark-mass effects in the NNLOPS POWHEG+MinLO Higgs generator”, *JHEP* **05** (2015) 140, doi:10.1007/JHEP05(2015)140, arXiv:1501.04637.
- [68] J. Butterworth et al., “PDF4LHC recommendations for LHC Run II”, *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.
- [69] S. Catani and M. Grazzini, “An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC”, *Phys. Rev. Lett.* **98** (2007) 222002, doi:10.1103/PhysRevLett.98.222002, arXiv:hep-ph/0703012.
- [70] ATLAS Collaboration, “Measurement of the  $Z/\gamma^*$  boson transverse momentum distribution in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector”, *JHEP* **09** (2014) 145, doi:10.1007/JHEP09(2014)145, arXiv:1406.3660.
- [71] U. Aglietti, R. Bonciani, G. Degrossi, and A. Vicini, “Two loop light fermion contribution to Higgs production and decays”, *Phys. Lett. B* **595** (2004) 432–441, doi:10.1016/j.physletb.2004.06.063, arXiv:hep-ph/0404071.
- [72] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, “NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders”, *Phys. Lett. B* **670** (2008) 12–17, doi:10.1016/j.physletb.2008.10.018, arXiv:0809.1301.
- [73] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, “NNLO Computational Techniques: The Cases  $H \rightarrow \gamma\gamma$  and  $H \rightarrow gg$ ”, *Nucl. Phys. B* **811** (2009) 182–273, doi:10.1016/j.nuclphysb.2008.11.024, arXiv:0809.3667.
- [74] C. Anastasiou, R. Boughezal, and F. Petriello, “Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion”, *JHEP* **04** (2009) 003, doi:10.1088/1126-6708/2009/04/003, arXiv:0811.3458.
- [75] A. Pak, M. Rogal, and M. Steinhauser, “Finite top quark mass effects in NNLO Higgs boson production at LHC”, *JHEP* **02** (2010) 025, doi:10.1007/JHEP02(2010)025, arXiv:0911.4662.
- [76] R. V. Harlander and K. J. Ozeren, “Top mass effects in Higgs production at next-to-next-to-leading order QCD: Virtual corrections”, *Phys. Lett. B* **679** (2009) 467–472, doi:10.1016/j.physletb.2009.08.012, arXiv:0907.2997.

- [77] R. V. Harlander and K. J. Ozeren, “Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order”, *JHEP* **11** (2009) 088, doi:10.1088/1126-6708/2009/11/088, arXiv:0909.3420.
- [78] R. V. Harlander, H. Mantler, S. Marzani, and K. J. Ozeren, “Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass”, *Eur. Phys. J. C* **66** (2010) 359–372, doi:10.1140/epjc/s10052-010-1258-x, arXiv:0912.2104.
- [79] C. Anastasiou et al., “Higgs Boson Gluon-Fusion Production in QCD at Three Loops”, *Phys. Rev. Lett.* **114** (2015) 212001, doi:10.1103/PhysRevLett.114.212001, arXiv:1503.06056.
- [80] C. Anastasiou et al., “High precision determination of the gluon fusion Higgs boson cross-section at the LHC”, *JHEP* **05** (2016) 058, doi:10.1007/JHEP05(2016)058, arXiv:1602.00695.
- [81] F. Dulat, A. Lazopoulos, and B. Mistlberger, “iHixs 2 — Inclusive Higgs cross sections”, *Comput. Phys. Commun.* **233** (2018) 243–260, doi:10.1016/j.cpc.2018.06.025, arXiv:1802.00827.
- [82] M. Bonetti, K. Melnikov, and L. Tancredi, “Higher order corrections to mixed QCD-EW contributions to Higgs boson production in gluon fusion”, *Phys. Rev. D* **97** (2018), no. 5, 056017, doi:10.1103/PhysRevD.97.056017, arXiv:1801.10403. [Erratum: *Phys.Rev.D* 97, 099906 (2018)].
- [83] P. Nason and C. Oleari, “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG”, *JHEP* **02** (2010) 037, doi:10.1007/JHEP02(2010)037, arXiv:0911.5299.
- [84] G. Cullen et al., “Automated One-Loop Calculations with GoSam”, *Eur. Phys. J. C* **72** (2012) 1889, doi:10.1140/epjc/s10052-012-1889-1, arXiv:1111.2034.
- [85] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, “ $HW^\pm/HZ + 0$  and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO”, *JHEP* **10** (2013) 083, doi:10.1007/JHEP10(2013)083, arXiv:1306.2542.
- [86] M. Ciccolini, A. Denner, and S. Dittmaier, “Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC”, *Phys. Rev. Lett.* **99** (2007) 161803, doi:10.1103/PhysRevLett.99.161803, arXiv:0707.0381.
- [87] M. Ciccolini, A. Denner, and S. Dittmaier, “Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC”, *Phys. Rev. D* **77** (2008) 013002, doi:10.1103/PhysRevD.77.013002, arXiv:0710.4749.
- [88] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, “Higgs production via vector-boson fusion at NNLO in QCD”, *Phys. Rev. Lett.* **105** (2010) 011801, doi:10.1103/PhysRevLett.105.011801, arXiv:1003.4451.

- [89] M. L. Ciccolini, S. Dittmaier, and M. Kramer, “Electroweak radiative corrections to associated WH and ZH production at hadron colliders”, *Phys. Rev. D* **68** (2003) 073003, doi:10.1103/PhysRevD.68.073003, arXiv:hep-ph/0306234.
- [90] O. Brein, A. Djouadi, and R. Harlander, “NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders”, *Phys. Lett. B* **579** (2004) 149–156, doi:10.1016/j.physletb.2003.10.112, arXiv:hep-ph/0307206.
- [91] O. Brein, R. Harlander, M. Wiesemann, and T. Zirke, “Top-Quark Mediated Effects in Hadronic Higgs-Strahlung”, *Eur. Phys. J. C* **72** (2012) 1868, doi:10.1140/epjc/s10052-012-1868-6, arXiv:1111.0761.
- [92] L. Altenkamp et al., “Gluon-induced Higgs-strahlung at next-to-leading order QCD”, *JHEP* **02** (2013) 078, doi:10.1007/JHEP02(2013)078, arXiv:1211.5015.
- [93] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, “HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders”, *Comput. Phys. Commun.* **195** (2015) 161–171, doi:10.1016/j.cpc.2015.04.021, arXiv:1412.5390.
- [94] O. Brein, R. V. Harlander, and T. J. E. Zirke, “vh@nnlo - Higgs Strahlung at hadron colliders”, *Comput. Phys. Commun.* **184** (2013) 998–1003, doi:10.1016/j.cpc.2012.11.002, arXiv:1210.5347.
- [95] R. V. Harlander, A. Kulesza, V. Theeuwes, and T. Zirke, “Soft gluon resummation for gluon-induced Higgs Strahlung”, *JHEP* **11** (2014) 082, doi:10.1007/JHEP11(2014)082, arXiv:1410.0217.
- [96] R. V. Harlander, J. Klappert, S. Liebler, and L. Simon, “vh@nnlo-v2: New physics in Higgs Strahlung”, *JHEP* **05** (2018) 089, doi:10.1007/JHEP05(2018)089, arXiv:1802.04817.
- [97] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [98] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”, *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.
- [99] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [100] “ATLAS Pythia 8 tunes to 7 TeV data”, technical report, CERN, Geneva, (2014). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-021>.

- [101] W. Beenakker et al., “NLO QCD corrections to  $t$  anti- $t$   $H$  production in hadron collisions”, *Nucl. Phys. B* **653** (2003) 151–203, doi:10.1016/S0550-3213(03)00044-0, arXiv:hep-ph/0211352.
- [102] S. Dawson et al., “Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections”, *Phys. Rev. D* **68** (2003) 034022, doi:10.1103/PhysRevD.68.034022, arXiv:hep-ph/0305087.
- [103] Y. Zhang et al., “QCD NLO and EW NLO corrections to  $t\bar{t}H$  production with top quark decays at hadron collider”, *Phys. Lett. B* **738** (2014) 1–5, doi:10.1016/j.physletb.2014.09.022, arXiv:1407.1110.
- [104] S. Frixione et al., “Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons”, *JHEP* **06** (2015) 184, doi:10.1007/JHEP06(2015)184, arXiv:1504.03446.
- [105] Sherpa Collaboration, “Event Generation with Sherpa 2.2”, *SciPost Phys.* **7** (2019), no. 3, 034, doi:10.21468/SciPostPhys.7.3.034, arXiv:1905.09127.
- [106] T. Gleisberg and S. Hoeche, “Comix, a new matrix element generator”, *JHEP* **12** (2008) 039, doi:10.1088/1126-6708/2008/12/039, arXiv:0808.3674.
- [107] F. Cascioli, P. Maierhofer, and S. Pozzorini, “Scattering Amplitudes with Open Loops”, *Phys. Rev. Lett.* **108** (2012) 111601, doi:10.1103/PhysRevLett.108.111601, arXiv:1111.5206.
- [108] A. Denner, S. Dittmaier, and L. Hofer, “Collier: a fortran-based Complex One-Loop Library in Extended Regularizations”, *Comput. Phys. Commun.* **212** (2017) 220–238, doi:10.1016/j.cpc.2016.10.013, arXiv:1604.06792.
- [109] S. Schumann and F. Krauss, “A Parton shower algorithm based on Catani-Seymour dipole factorisation”, *JHEP* **03** (2008) 038, doi:10.1088/1126-6708/2008/03/038, arXiv:0709.1027.
- [110] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, “A critical appraisal of NLO+PS matching methods”, *JHEP* **09** (2012) 049, doi:10.1007/JHEP09(2012)049, arXiv:1111.1220.
- [111] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, “QCD matrix elements + parton showers: The NLO case”, *JHEP* **04** (2013) 027, doi:10.1007/JHEP04(2013)027, arXiv:1207.5030.
- [112] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, “QCD matrix elements + parton showers”, *JHEP* **11** (2001) 063, doi:10.1088/1126-6708/2001/11/063, arXiv:hep-ph/0109231.
- [113] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert, “QCD matrix elements and truncated showers”, *JHEP* **05** (2009) 053, doi:10.1088/1126-6708/2009/05/053, arXiv:0903.1219.

- [114] S. Frixione, P. Nason, and G. Ridolfi, “A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”, *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.
- [115] E. Re, “Single-top  $Wt$ -channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [116] S. Frixione et al., “Single-top hadroproduction in association with a  $W$  boson”, *JHEP* **07** (2008) 029, doi:10.1088/1126-6708/2008/07/029, arXiv:0805.3067.
- [117] S. Höche, S. Prestel, and H. Schulz, “Simulation of Vector Boson Plus Many Jet Final States at the High Luminosity LHC”, *Phys. Rev. D* **100** (2019), no. 1, 014024, doi:10.1103/PhysRevD.100.014024, arXiv:1905.05120.
- [118] S. Dulat et al., “New parton distribution functions from a global analysis of quantum chromodynamics”, *Phys. Rev. D* **93** (2016), no. 3, 033006, doi:10.1103/PhysRevD.93.033006, arXiv:1506.07443.
- [119] L. Lonnblad, “Correcting the color dipole cascade model with fixed order matrix elements”, *JHEP* **05** (2002) 046, doi:10.1088/1126-6708/2002/05/046, arXiv:hep-ph/0112284.
- [120] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [121] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- [122] ATLAS Collaboration, “Jet reconstruction and performance using particle flow with the ATLAS Detector”, *Eur. Phys. J. C* **77** (2017), no. 7, 466, doi:10.1140/epjc/s10052-017-5031-2, arXiv:1703.10485.
- [123] ATLAS Collaboration, “Performance of pile-up mitigation techniques for jets in  $pp$  collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector”, *Eur. Phys. J. C* **76** (2016), no. 11, 581, doi:10.1140/epjc/s10052-016-4395-z, arXiv:1510.03823.
- [124] D. Krohn, J. Thaler, and L.-T. Wang, “Jet Trimming”, *JHEP* **02** (2010) 084, doi:10.1007/JHEP02(2010)084, arXiv:0912.1342.
- [125] D. Krohn, J. Thaler, and L.-T. Wang, “Jets with Variable  $R$ ”, *JHEP* **06** (2009) 059, doi:10.1088/1126-6708/2009/06/059, arXiv:0903.0392.
- [126] M. Cacciari, G. P. Salam, and G. Soyez, “The Catchment Area of Jets”, *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.
- [127] ATLAS Collaboration, “Measurements of b-jet tagging efficiency with the ATLAS detector using  $t\bar{t}$  events at  $\sqrt{s} = 13$  TeV”, *JHEP* **08** (2018) 089, doi:10.1007/JHEP08(2018)089, arXiv:1805.01845.

- [128] ATLAS Collaboration, “ATLAS  $b$ -jet identification performance and efficiency measurement with  $t\bar{t}$  events in pp collisions at  $\sqrt{s} = 13$  TeV”, *Eur. Phys. J. C* **79** (2019), no. 11, 970, doi:10.1140/epjc/s10052-019-7450-8, arXiv:1907.05120.
- [129] ATLAS Collaboration Collaboration, “Optimisation and performance studies of the ATLAS  $b$ -tagging algorithms for the 2017-18 LHC run”, technical report, CERN, Geneva, (2017). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-013>.
- [130] ATLAS Collaboration, “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s}=13$  TeV”, *Eur. Phys. J. C* **76** (2016), no. 5, 292, doi:10.1140/epjc/s10052-016-4120-y, arXiv:1603.05598.
- [131] ATLAS Collaboration, “Higgs boson production cross-section measurements and their EFT interpretation in the  $4\ell$  decay channel at  $\sqrt{s}=13$  TeV with the ATLAS detector”, *Eur. Phys. J. C* **80** (2020), no. 10, 957, doi:10.1140/epjc/s10052-020-8227-9, arXiv:2004.03447. [Erratum: *Eur.Phys.J.C* 81, 29 (2021), Erratum: *Eur.Phys.J.C* 81, 398 (2021)].
- [132] ATLAS Collaboration, “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”, *JINST* **14** (2019), no. 12, P12006, doi:10.1088/1748-0221/14/12/P12006, arXiv:1908.00005.
- [133] ATLAS Collaboration Collaboration, “Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC”, technical report, CERN, Geneva, (2015). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045>.
- [134] ATLAS Collaboration Collaboration, “Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment”, technical report, CERN, Geneva, (2019). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2019-033>.
- [135] J. Snoek, H. Larochelle, and R. P. Adams, “Practical Bayesian Optimization of Machine Learning Algorithms”, in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, eds., volume 25, Curran Associates, Inc., 2012.
- [136] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python”, *Journal of Machine Learning Research* **12** (2011) 2825–2830.
- [137] W. Verkerke and D. Kirkby, “The RooFit toolkit for data modeling”, 2003. doi:10.48550/ARXIV.PHYSICS/0306116.
- [138] L. Moneta et al., “The RooStats Project”, 2010. doi:10.48550/ARXIV.1009.1003.

- [139] ATLAS Collaboration, “Search for Scalar Diphoton Resonances in the Mass Range 65 – 600 GeV with the ATLAS Detector in  $pp$  Collision Data at  $\sqrt{s} = 8 \text{ TeV}$ ”, *Phys. Rev. Lett.* **113** (2014), no. 17, 171801, doi:10.1103/PhysRevLett.113.171801, arXiv:1407.6583.
- [140] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”, doi:10.23731/CYRM-2017-002, arXiv:1610.07922.
- [141] X. Liu and F. Petriello, “Reducing theoretical uncertainties for exclusive Higgs-boson plus one-jet production at the LHC”, *Phys. Rev. D* **87** (2013), no. 9, 094027, doi:10.1103/PhysRevD.87.094027, arXiv:1303.4405.
- [142] R. Boughezal et al., “Combining Resummed Higgs Predictions Across Jet Bins”, *Phys. Rev. D* **89** (2014), no. 7, 074044, doi:10.1103/PhysRevD.89.074044, arXiv:1312.4535.
- [143] J. M. Campbell and R. K. Ellis, “MCFM for the Tevatron and the LHC”, *Nucl. Phys. B Proc. Suppl.* **205-206** (2010) 10–15, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.
- [144] I. W. Stewart and F. J. Tackmann, “Theory Uncertainties for Higgs and Other Searches Using Jet Bins”, *Phys. Rev. D* **85** (2012) 034011, doi:10.1103/PhysRevD.85.034011, arXiv:1107.2117.
- [145] S. Gangal and F. J. Tackmann, “Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion”, *Phys. Rev. D* **87** (2013), no. 9, 093008, doi:10.1103/PhysRevD.87.093008, arXiv:1302.5437.
- [146] R. Frederix, S. Frixione, E. Vryonidou, and M. Wiesemann, “Heavy-quark mass effects in Higgs plus jets production”, *JHEP* **08** (2016) 006, doi:10.1007/JHEP08(2016)006, arXiv:1604.03017.
- [147] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, “Higgs boson production at the LHC: transverse momentum resummation effects in the  $H \rightarrow 2\gamma$ ,  $H \rightarrow WW \rightarrow l\nu l\nu$  and  $H \rightarrow ZZ \rightarrow 4l$  decay modes”, *JHEP* **06** (2012) 132, doi:10.1007/JHEP06(2012)132, arXiv:1203.6321.
- [148] M. Grazzini and H. Sargsyan, “Heavy-quark mass effects in Higgs boson production at the LHC”, *JHEP* **09** (2013) 129, doi:10.1007/JHEP09(2013)129, arXiv:1306.4581.
- [149] M. Bahr et al., “Herwig++ Physics and Manual”, *Eur. Phys. J. C* **58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.
- [150] J. Bellm et al., “Herwig 7.0/Herwig++ 3.0 release note”, *Eur. Phys. J. C* **76** (2016), no. 4, 196, doi:10.1140/epjc/s10052-016-4018-8, arXiv:1512.01178.

- [151] ATLAS Collaboration, “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at  $\sqrt{s} = 13$  TeV”, *Eur. Phys. J. C* **78** (2018), no. 11, 903, doi:10.1140/epjc/s10052-018-6288-9, arXiv:1802.08168.
- [152] ATLAS Collaboration, “Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector”, technical report, CERN, Geneva, (2017). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-009>.
- [153] ATLAS Collaboration, “Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, *Phys. Rev. D* **96** (2017), no. 7, 072002, doi:10.1103/PhysRevD.96.072002, arXiv:1703.09665.
- [154] ATLAS Collaboration, “Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC”, technical report, CERN, Geneva, (2019). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-021>.
- [155] G. Avoni et al., “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”, *Journal of Instrumentation* **13** (jul, 2018) P07017, doi:10.1088/1748-0221/13/07/P07017.
- [156] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: *Eur.Phys.J.C* **73**, 2501 (2013)].
- [157] R. D. Cousins, “Generalization of Chisquare Goodness-ofFit Test for Binned Data Using Saturated Models , with Application to Histograms”, 2013.
- [158] A. L. Read, “Presentation of search results: the CLs technique”, *Journal of Physics G: Nuclear and Particle Physics* **28** (sep, 2002) 2693, doi:10.1088/0954-3899/28/10/313.
- [159] ATLAS Collaboration, “Search for the dimuon decay of the Higgs boson in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, *Phys. Rev. Lett.* **119** (2017), no. 5, 051802, doi:10.1103/PhysRevLett.119.051802, arXiv:1705.04582.
- [160] “Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector”, technical report, CERN, Geneva, (2015). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2015-029>.

- [161] ATLAS Collaboration, “Object-based missing transverse momentum significance in the ATLAS detector”, technical report, CERN, Geneva, (2018). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-038>.
- [162] ATLAS Collaboration, “Jet energy scale and resolution measured in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, *Eur. Phys. J. C* **81** (2021), no. 8, 689, doi:10.1140/epjc/s10052-021-09402-3, arXiv:2007.02645.
- [163] ATLAS Collaboration, “In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector”, *Eur. Phys. J. C* **79** (2019), no. 2, 135, doi:10.1140/epjc/s10052-019-6632-8, arXiv:1807.09477.
- [164] ATLAS Collaboration, “In-situ measurements of the ATLAS large-radius jet response in 13 TeV pp collisions”, technical report, CERN, Geneva, (2017). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2017-063>.
- [165] ATLAS Collaboration, “Measurement of  $b$ -tagging Efficiency of  $c$ -jets in  $t\bar{t}$  Events Using a Likelihood Approach with the ATLAS Detector”, technical report, CERN, Geneva, (2018). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-001>.
- [166] ATLAS Collaboration, “Calibration of light-flavour  $b$ -jet mistagging rates using ATLAS proton-proton collision data at  $\sqrt{s} = 13$  TeV”, technical report, CERN, Geneva, (2018). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-006>.
- [167] ATLAS Collaboration, “Studies on top-quark Monte Carlo modelling for Top2016”, technical report, CERN, Geneva, (2016). All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-020>.
- [168] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [169] ATLAS Collaboration, “Measurements of the production cross-section for a  $Z$  boson in association with  $b$ -jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, *JHEP* **07** (2020) 044, doi:10.1007/JHEP07(2020)044, arXiv:2003.11960.
- [170] ATLAS Collaboration, “Constraints on mediator-based dark matter and scalar dark energy models using  $\sqrt{s} = 13$  TeV  $pp$  collision data collected by the ATLAS detector”, *JHEP* **05** (2019) 142, doi:10.1007/JHEP05(2019)142, arXiv:1903.01400.

- [171] ATLAS Collaboration, “Search for Dark Matter Produced in Association with a Higgs Boson Decaying to  $b\bar{b}$  using  $36 \text{ fb}^{-1}$  of  $pp$  collisions at  $\sqrt{s} = 13 \text{ TeV}$  with the ATLAS Detector”, *Phys. Rev. Lett.* **119** (2017), no. 18, 181804, doi:10.1103/PhysRevLett.119.181804, arXiv:1707.01302.
- [172] M. Bauer, M. Klassen, and V. Tenorth, “Universal properties of pseudoscalar mediators in dark matter extensions of 2HDMs”, *JHEP* **07** (2018) 107, doi:10.1007/JHEP07(2018)107, arXiv:1712.06597.
- [173] U. Haisch and G. Polesello, “Searching for heavy Higgs bosons in the  $t\bar{t}Z$  and  $tbW$  final states”, *JHEP* **09** (2018) 151, doi:10.1007/JHEP09(2018)151, arXiv:1807.07734.
- [174] G. Cowan, “A survey of unfolding methods for particle physics”, *Conf. Proc.* **C0203181** (2002) 248.
- [175] V. Blobel, “Unfolding Methods in Particle Physics”, *PHYSTAT2011 Proceedings* (2011) 240, doi:10.5170/CERN-2011-006.
- [176] V. Blobel, “Unfolding”, *Data Analysis in High Energy Physics* (2013) 187, doi:10.1002/9783527653416.ch6.
- [177] R. Balasubramanian et al., “Statistical method and comparison of different unfolding techniques using RooFit”, arXiv:1910.14654.
- [178] G. D’Agostini, “A Multidimensional unfolding method based on Bayes’ theorem”, *Nucl. Instrum. Meth.* **A362** (1995) 487–498, doi:10.1016/0168-9002(95)00274-X.
- [179] A. Hocker and V. Kartvelishvili, “SVD approach to data unfolding”, *Nucl. Instrum. Meth.* **A372** (1996) 469–481, doi:10.1016/0168-9002(95)01478-0, arXiv:hep-ph/9509307.
- [180] S. Schmitt, “TUnfold: an algorithm for correcting migration effects in high energy physics”, *JINST* **7** (2012) T10003, doi:10.1088/1748-0221/7/10/T10003, arXiv:1205.6201.
- [181] M. Arratia et al., “Publishing unbinned differential cross section results”, *JINST* **17** (2022), no. 01, P01024, doi:10.1088/1748-0221/17/01/P01024, arXiv:2109.13243.
- [182] I. J. Goodfellow et al., “Generative Adversarial Nets”, in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS’14*, pp. 2672–2680, MIT Press, Cambridge, MA, USA, 2014.
- [183] K. Datta, D. Kar, and D. Roy, “Unfolding with Generative Adversarial Networks”, arXiv:1806.00433.
- [184] M. Bellagente et al., “How to GAN away Detector Effects”, arXiv:1912.00477.
- [185] D. P. Kingma and M. Welling, “Auto-Encoding Variational Bayes”, arXiv:1312.6114.

- [186] J. N. Howard, S. Mandt, D. Whiteson, and Y. Yang, “Foundations of a Fast, Data-Driven, Machine-Learned Simulator”, arXiv:2101.08944.
- [187] D. J. Rezende and S. Mohamed, “Variational Inference with Normalizing Flows”, *International Conference on Machine Learning* **37** (2015) 1530.
- [188] M. Bellagente et al., “Invertible Networks or Partons to Detector and Back Again”, *SciPost Phys.* **9** (2020) 074, doi:10.21468/SciPostPhys.9.5.074, arXiv:2006.06685.
- [189] M. Vandegar, M. Kagan, A. Wehenkel, and G. Louppe, “Neural Empirical Bayes: Source Distribution Estimation and its Applications to Simulation-Based Inference”, in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, A. Banerjee and K. Fukumizu, eds., volume 130 of *Proceedings of Machine Learning Research*, pp. 2107–2115, PMLR, 11, 2021. arXiv:2011.05836.
- [190] M. Backes, A. Butter, M. Dunford, and B. Malaescu, “An unfolding method based on conditional Invertible Neural Networks (cINN) using iterative training”, arXiv:2212.08674.
- [191] A. Andreassen et al., “OmniFold: A Method to Simultaneously Unfold All Observables”, *Phys. Rev. Lett.* **124** (2020), no. 18, 182001, doi:10.1103/PhysRevLett.124.182001, arXiv:1911.09107.
- [192] A. Andreassen et al., “Scaffolding Simulations with Deep Learning for High-dimensional Deconvolution”, in *9th International Conference on Learning Representations*, 5, 2021. arXiv:2105.04448.
- [193] H1 Collaboration, “Measurement of Lepton-Jet Correlation in Deep-Inelastic Scattering with the H1 Detector Using Machine Learning for Unfolding”, *Phys. Rev. Lett.* **128** (2022), no. 13, 132002, doi:10.1103/PhysRevLett.128.132002, arXiv:2108.12376.
- [194] H1 Collaboration, “Machine learning-assisted measurement of multi-differential lepton-jet correlations in deep-inelastic scattering with the H1 detector”, *H1prelim-22-031* (2022).
- [195] H1 Collaboration, “Multi-differential Jet Substructure Measurement in High  $Q^2$  DIS Events with HERA-II Data”, *H1prelim-22-034* (2022).
- [196] LHCb Collaboration, “Multidifferential study of identified charged hadron distributions in  $Z$ -tagged jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, arXiv:2208.11691 (8, 2022).
- [197] J. R. Andersen et al., “Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report”, in *9th Les Houches Workshop on Physics at TeV Colliders*, 5, 2016. arXiv:1605.04692.
- [198] N. Berger et al., “Simplified Template Cross Sections - Stage 1.1”, arXiv:1906.02754.
- [199] S. Amoroso et al., “Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report”, in *11th Les Houches Workshop on Physics at TeV Colliders: PhysTeV Les Houches*, 3, 2020. arXiv:2003.01700.

- [200] G. Choudalakis, “Fully Bayesian Unfolding”, 2012. doi:10.48550/ARXIV.1201.4612.
- [201] B. Nachman and J. Thaler, “Neural Conditional Reweighting”, *Phys. Rev. D* **105** (2022) 076015, doi:10.1103/PhysRevD.105.076015, arXiv:2107.08979.
- [202] T. Hastie, R. Tibshirani, and J. Friedman, “The Elements of Statistical Learning”. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.
- [203] M. Sugiyama, T. Suzuki, and T. Kanamori, “Density Ratio Estimation in Machine Learning”. Cambridge University Press, 2012.
- [204] A. Paszke et al., “PyTorch: An Imperative Style, High-Performance Deep Learning Library”, in *Advances in Neural Information Processing Systems 32*, pp. 8024–8035. Curran Associates, Inc., 2019.
- [205] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization”, 2014. doi:10.48550/ARXIV.1412.6980.
- [206] C. Oleari, “The POWHEG-BOX”, *Nucl. Phys. B Proc. Suppl.* **205-206** (2010) 36–41, doi:10.1016/j.nuclphysbps.2010.08.016, arXiv:1007.3893.
- [207] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG”, *JHEP* **04** (2009) 002, doi:10.1088/1126-6708/2009/04/002, arXiv:0812.0578.
- [208] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [209] DELPHES 3 Collaboration, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment”, *JHEP* **02** (2014) 057, doi:10.1007/JHEP02(2014)057, arXiv:1307.6346.
- [210] L. Heinrich, M. Feickert, and G. Stark, “pyhf: v0.7.0”. <https://github.com/scikit-hep/pyhf/releases/tag/v0.7.0>. doi:10.5281/zenodo.1169739.
- [211] L. Heinrich, M. Feickert, G. Stark, and K. Cranmer, “pyhf: pure-Python implementation of HistFactory statistical models”, *Journal of Open Source Software* **6** (2021), no. 58, 2823, doi:10.21105/joss.02823.