

Probing new physics in the vector boson fusion production of Higgs bosons

A Dissertation

Presented to

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

Department of Physics

Craig Blocker, Gabriella Sciolla, Advisor

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

by
Jiayi Chen

August 2022



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

DOCTOR OF PHILOSOPHY

Wendy Cadge, Dean
Graduate School of Arts and Sciences

Dissertation Committee:

Gabriella Sciolla, Physics

Craig Blocker, Physics

Sandra Kortner, Physics, Max-Planck-Institut für Physik

Copyright by
Jiayi Chen

2022

Acknowledgements

I have many people to thank for their love and support throughout my Ph.D. All of these incredible people have shaped me to become a better self.

Foremost, I would like to thank my advisors. To **Gabriella Sciolla**, you have inspired me with your great passion about physics since the first day we met. You are my rock for every big moments in my graduate study such as qualifications and important conferences. Thank you for always expressing your confidence in me and applauding even the tiniest achievement I made. To **Craig Blocker**, thank you for your constant presence and participation in my research. The rigor and calmness you bring to our intellectual conversations will follow me for the rest of my career. I feel extremely lucky to have worked with the both of you and I could not have reached my goals without you.

Secondly, I want to thank **Sandra Kortner** for serving on my defense committee. Your comments always challenge me to think about problems from a different perspective.

To my analysis team members, I could not have accomplished this thesis without you. To **Alessandro Tricoli**, thank you for your three-year mentorship full of encouragement and insightful guidance. I credit you for leading me to a research direction that I have grown to be passionate about. I also owe a big thank-you to my postdoctorate supervisor **Gaetano Barone** for teaching me how to develop a physics analysis and how to approach technical problems. Many thanks also to the other students in the team, **Sagar Addepalli**, **Laura Bergsten**, **Chara Kitsaki**, and **Guy Rosin**.

I am extremely grateful for the entire **HWW** team for the strong foundation they have built for students like me to base our own work on. In particular, I would like to thank **Carsten Burgard** and **Bryan Kortman** for teaching me the theory and tools in great details. If it were not for your constant patience and kindness, I could not have absorbed these much knowledge easily.

Next, I would like to express my sincere gratitude to our ITk collaborators. While there are countless people I am grateful for, I want to especially thank senior physicists, **Dave Lynn** and **Stefania Stucci**, from Brookhaven National Lab for mentoring me during my time at the lab, my technical supervisor for my ATLAS qualification task **Andy Blue** for the guidance in my leadership role and **Matt Basso** for teaching me many important coding skills.

Thank you to all my colleagues in the Sciolla group, **Max Goblirsch** for answering my weird coding questions, **Hannah Herde** for convincing me to join Brandeis in the first place and always giving me a hand in work and in life, **Laura Bergsten** for patiently mentoring me on hardware and analysis, **Francesca Capocasa** for the silly laughters we shared and long walks we took, **Prajita Bhattarai** for being such a wonderful friend who cooks amazing food and takes me to zumba, and **Sagar Addepalli** for your joyful companionship in our analysis.

To my friends who share this exciting journey with me **Alyssa Garcia**, **Daichi Hayakawa**, **Huang Fang**, **Zahra Zarei** and **Shiyu Zhou**.

To everyone in my family - my **Mom**, my **Dad**, and my **Grandfather** and **Grandmother** - I could not have accomplish this without knowing you are always there for me. To my parents, you have been my role models on every level. You always prioritize my goals in

life. You would fly to different countries and cities to visit me despite your busy schedules. To my first math teacher, my grandfather, you taught me many lessons not with lecture but by proof with actions.

Finally, to my husband **Zisen**, thank you for patiently waiting for me during the time we are apart. Your faith in myself gave me great confidence to overcome obstacles in life. Your unwavering support is the sole reason for me to stay focused on my life goal. My heartfelt gratitude and love to you.

To my grandfather, Guoxiang Wen.

ABSTRACT

Probing new physics in the vector boson fusion production of Higgs bosons

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

By Jiayi Chen

This thesis reports the vector boson fusion production cross section for the Higgs boson in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ final state measured as a function of kinematic observables sensitive to the Higgs boson production and decay properties as well as integrated in a fiducial phase space. The analysis is performed using the proton-proton collision data collected by the ATLAS detector in the LHC Run-2 at $\sqrt{s} = 13$ TeV center-of-mass energy, corresponding to an integrated luminosity of 139 fb^{-1} . The data are corrected for the effects of detector inefficiency and resolution, and the measurements are compared to different state-of-the-art theoretical predictions. The observed fiducial cross section of $1.68^{+0.40}_{-0.38} \text{ fb}$ is compatible with the Standard Model prediction at a cross section of $2.07 \pm 0.04 \text{ fb}$. The differential cross sections are finally used to constrain anomalous CP-even and CP-odd interactions using a dimension-6 Effective Field Theory formalism.

Table of Contents

Acknowledgements	iv
Abstract	vii
List of Tables	xi
List of Figures	xiv
List of Abbreviations	xxvii
1 Introduction	1
2 Theory	3
2.1 The Standard Model	3
2.1.1 Overview	3
2.1.2 Quantum electrodynamics	6
2.1.3 Quantum chromodynamics	6
2.1.4 The electroweak theory and the Higgs mechanism	8
2.1.5 From the Lagrangian to observables	10
2.2 The Higgs boson measurements	11
2.2.1 Main Higgs production modes	11
2.2.2 Main Higgs decay modes	14
2.3 Effective Field Theory	15
2.3.1 Fermi theory	16
2.3.2 Standard Model Effective Field Theory	17
3 Apparatus	20
3.1 The Large Hadron Collider	20
3.2 The ATLAS detector	22
3.3 Physics object reconstruction in the ATLAS detector	27
3.3.1 Tracks and vertices	27
3.3.2 Electrons	28
3.3.3 Muons	30

	3.3.4 Jets	33
4	Analysis overview	36
5	Data and Monte Carlo samples	39
	5.1 Data samples	39
	5.2 Monte Carlo samples	39
	5.2.1 Signal samples	39
	5.2.2 Background samples	40
6	Signal optimization	42
	6.1 Physics objects	42
	6.2 Observables	45
	6.3 Pre-selection	47
	6.4 Signal region selection	48
	6.5 Signal optimization with boosted decision tree	49
	6.5.1 Signal optimization for fiducial cross-section measurement	51
	6.5.2 Signal optimization for differential cross-section measurement . .	54
7	Background estimation	56
	7.1 Jets mis-identified as leptons	57
	7.2 Top and diboson	61
	7.3 Drell-Yan	61
	7.4 Gluon-gluon fusion plus two jets	62
8	Fiducial phase space and particle-level predictions	66
	8.1 Fiducial phase space	66
	8.1.1 Particle-level objects	66
	8.1.2 Particle-level selection	67
	8.2 VBFNLO predictions	67
	8.3 Theory predictions comparison	71
9	Signal extraction strategy	73

9.1	Systematics	73
9.2	Unfolding	74
9.3	Statistical analysis for the cross-section measurements	76
10	Cross-section measurement results	81
10.1	Fiducial cross section	81
10.2	Differential cross sections	86
11	Effective Field Theory interpretation	92
11.1	EFT simulation and parametrization strategy	94
11.1.1	EFT fiducial parametrization	95
11.1.2	Higgs boson width parametrization	104
11.2	EFT sensitivity tests	105
11.3	EFT results	108
12	Conclusion and outlook	115
References		116
Appendices		123
A	Expected differential cross sections	123
B	Differential cross sections	128
C	EFT sensitivity test	143
D	EFT bias test	144

List of Tables

1	Electron selection requirements.	43
2	Muon selection requirements.	43
3	Pre-selection cuts applied before event selections for different signal and background-enriched regions. The lepton identification criteria is different for the fakes. The fake factor implementation will be detailed in Sec. 7.1.	47
4	VBF signal region selections.	49
5	Ranking of the variables used as input to the training of the BDT signal vs. top and diboson.	51
6	Ranking of the variables used as input to the training of the BDT of top and diboson vs. other processes.	53
7	Variables removed from the BDT input list for several observables to be measured.	55
8	Composition of each signal and background process in the CRZjets, the ggF CR, and the SR. The reported composition values are obtained from MC simulations.	56
9	Requirements for the identified and anti-identified electrons and muons. . . .	57
10	Ranking of the variables used as input to the training of the BDT trained in the ggF control region.	64
11	Definition of fiducial phase space for VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ in the $e - \mu$ decay channel.	67
12	The fiducial cross sections and the theory systematics of all parton-level and particle-level samples.	72
13	Binning optimized for each distribution. Inclusive bins have no upper bound.	76
14	Groups of observable bins which share the same independent top and diboson normalization factor.	79

15	Breakdown of the relative uncertainties $\Delta\sigma^{\text{fid}}/\sigma^{\text{fid}}$ in the inclusive fiducial cross-section measurement. The middle block shows the experimental systematic uncertainty breakdown and the bottom block shows those of the theory systematic uncertainties from different processes.	82
16	P-values for the compatibilities between the measured cross sections and the signal predictions given by POWHEG+PYTHIA8, POWHEG+HERWIG7, MADGRAPH5_AMC@NLO+ HERWIG7, VBFNLO+PYTHIA8, and fixed order calculation of VBFNLO at NLO.	86
17	Summary of EFT operators in the SMEFT formalism that are probed in the VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channel [1]. The corresponding structure in terms of the SM fields in the Warsaw basis (second column) is shown together with the associated Wilson coefficient (first column). The Higgs doublet field and its complex conjugate are denoted as H and $H^\dagger$, respectively. The left-handed quark doublets of flavor p (the right-handed up-type or down-type quarks of flavour r) are denoted as q_p (u_r or d_r), while $V_{\mu\nu}$ and $\tilde{V}_{\mu\nu}$ (with $V = B, W^n$ where $n = \{1, 2, 3\}$) represent a given gauge field of the electroweak interactions, and τ^n are the Pauli matrices. The bosonic operators with (without) a dual field $\tilde{V}_{\mu\nu}$ are CP-odd (CP-even).	92
18	Fiducial EFT parametrization of c_{HW} , c_{HWB} , c_{HB} and their CP odd counterparts for the differential bins of $\Delta\phi_{jj}$	99
19	EFT parametrization of c_{Hq3} , c_{Hq1} , c_{Hu} and c_{Hd} for the differential bins of p_T^{j1}	100
20	Best-fit values to the SM Asimov dataset, and the 2σ intervals obtained from profile likelihood scans of the Wilson coefficients under linear-only and linear plus quadratic parametrization.	106
21	Best-fit values to data and 2σ intervals obtained from profile likelihood scans of the Wilson coefficients under linear-only and linear plus quadratic parametrization.	109

22	Best-fit values to the SM Asimov dataset, and the 2σ intervals obtained from profile likelihood scans of the c_{Hq3} group of Wilson coefficients under linear-only and linear plus quadratic parametrization.	143
----	---	-----

List of Figures

1	The Standard Model of particle physics [2].	4
2	The Higgs potential as a function of the real (Re) and imaginary (Im) part of the field ϕ [3]. The yellow ball at the center represents the electroweak theory before spontaneous symmetry breaking where the gauge bosons are massless; the blue ball represents solution after breaking the electroweak symmetry. . .	9
3	The cross section of Higgs production for the dominant modes as a function of the center-of-mass energy at fixed Higgs mass of 125 GeV [4]. The error band includes all theoretical uncertainties.	12
4	The Feynman diagrams of several Higgs production modes: associated production with a vector boson (upper left); vector boson fusion (upper right); gluon-gluon fusion (lower left); associated production with heavy quarks (lower right) [5]. The vector bosons are denoted as V and virtual particles denoted with a star (*).	13
5	The signal strengths for the Higgs production cross section (top five) and branching-ratio (bottom four) signal strength measurements by the ATLAS collaboration [6]. The measurements of all production and decay channels use the gluon-gluon fusion production and Higgs to Z bosons decay as the baseline. The observed cross sections and branching ratios (black data points) are compatible with the SM predictions (red vertical line) within the uncertainties of the measurements (black error bars) and the theory predictions (gray error bands).	13

6	The signal strengths for the Higgs production cross section (left) and branching-ratio (right) signal strength measurements by the CMS collaboration [7]. The observed cross sections and branching ratios (black data points) are compatible with the SM predictions (black vertical line) within the uncertainties of the measurements (blue error bars), with a slight tension for the WH production mode which has a 2σ deviation.	14
7	The predicted branching ratios of the Higgs boson decay for the dominant decay channels as a function of the Higgs mass [4].	15
8	The electroweak theory reduced to the Fermi theory at low energy [8].	17
9	Schematic layout of the acceleration chain of the LHC and the main detectors at the four interaction points [9].	21
10	The ATLAS detector and its composite sub-detectors [10].	24
11	The ATLAS inner detector and the sub-detectors [10].	25
12	Cut-away view of the ATLAS calorimeter system [10].	26
13	Cut-away view of the ATLAS Muon Spectrometer [10].	26
14	Likelihood-based identification efficiencies for three working points measured from data as functions of E_T and η [11].	29
15	Isolation efficiencies for a few defined working points as functions of E_T and η [11].	31
16	Reconstruction efficiency for the Medium identified muons as a function of the muon transverse momentum p_T [12].	33
17	The jet transverse momentum resolution for calorimeter jets (LC+JES) and particle flow jets as a function of p_T (left) and η (right) [13]. The average pile-up ($\langle\mu\rangle$) is 24.	34
18	Jet-veto efficiency as a function of N_{Vtx} with (open markers) and without (solid markers) the JVT > 0.7 criterion [14].	35

19	Feynman diagram for the signal process: Higgs production via vector boson fusion and decay to two W bosons. The “ V ” stands for vector boson. One of the W bosons originating from the Higgs decay is virtual (off the mass shell), and is labeled as W^*	36
20	The representative Feynman diagrams of the dominant background processes: (a) $t\bar{t}$ process, (b) SM WW QCD process, (c) the Drell-Yan background producing two tau leptons, (d) ggF $H \rightarrow WW^*$ process plus two jets.	38
21	The Higgs p_T and $\Delta\phi_{ll}$ distributions in the pre-selection region from the ATLAS Run-2 data (black data points) and MC simulation (filled histogram). The uncertainty band shows only MC statistical uncertainties, and the error bars on the data points represent data statistical uncertainties. The bottom panel shows the data-to-MC ratio.	48
22	The $\Delta\phi_{ll}$, m_{jj} , m_T and jet multiplicity distributions in the VBF signal region from the ATLAS Run-2 data (black data points) and MC simulations (filled histogram). The uncertainty band includes MC statistical, experimental and theory uncertainties, and the error bars on the data points represent data statistical uncertainties.	50
23	The VBF BDT output. (a) BDT output distribution of normalized VBF (signal) and top and diboson processes (background) for the testing and training samples. The binning corresponds to the one used in the analysis. (b) VBF BDT output distribution of the signal and all backgrounds MC samples in a finer binning. The error band includes the statistical, experimental and theory uncertainties.	52

24	The top and diboson BDT output. (a) BDT output distribution of normalized top and diboson (signal) and other processes (background) for the testing and training samples. The binning corresponds to the one used in the analysis. (b) The top and diboson BDT output distribution of the signal and all backgrounds MC samples in a finer binning. The error band includes the statistical, experimental and theory uncertainties. The bottom panel shows the ratio of the data to the expected distributions.	53
25	The pre-fit MC distribution (filled histogram) of the 2-dimensional $D_{\text{VBF}} - D_{\text{top+VV}}$ discriminant (a) in SR1 onto the D_{VBF} axis and (b) in SR2 onto the $D_{\text{top+VV}}$ axis, compared with the Run-2 dataset (black data points). The error bars indicate the data statistical uncertainties, while the uncertainties in the MC prediction are shown by the hatched band, which includes experimental and theoretical uncertainties. The bottom panel shows the ratio of the data to the expected distributions.	54
26	Fake factors derived in the jet-enriched region as a function of electron p_T . .	58
27	Fake factors derived in the jet-enriched region as a function of muon p_T in the low- η region (left) and in the high- η region (right).	59
28	Kinematics distributions in the W +jets control region, with the data (black markers), electroweak processes in MC (blue histograms), and difference (blue markers) which represents the estimate of the fake background. No fake factors are applied.	60
29	The $m_{\tau\tau}$ and m_{ll} distributions in the Drell-Yan region from the ATLAS Run-2 data (black data points) and MC simulation (filled histogram). The error bars on the data points show statistical uncertainties in data. The uncertainty band includes MC statistical, experimental and theory uncertainties. The bottom panel shows the data-to-MC ratio.	62

30	The m_T and m_{jj} distributions in the ggF region from the ATLAS Run-2 data (black data points) and MC simulation (filled histogram). The error bars on the data points show statistical uncertainties in data. The uncertainty band includes MC statistical, experimental and theory uncertainties. The bottom panel shows the data-to-MC ratio.	63
31	The ggF BDT output. (a) Normalized BDT output distributions of the ggF (signal) and all other processes (background) for the testing and training samples. The binning corresponds to the one used in the analysis. (b) fully weighted BDT output distribution of ggF signal and all background processes in the ggF control region. The error band includes the statistical, experimental and theory uncertainties. The bottom panel shows the data-to-MC ratio. . .	64
32	The pre-fit MC distribution (filled histogram) of the discriminants used in (a) CRZjets and (b) ggF CR compared with the Run-2 dataset (black data points). The error bars on the data points indicate the data statistical uncertainties, while the uncertainties in the prediction are shown by the hatched band, which includes experimental and theoretical uncertainties. The bottom panel shows the data-to-MC ratio.	65
33	Distributions of lepton p_T produced by VBFNLO+PYTHIA 8 with and without tau decay. The ratios in the bottom panels are used as the tau-decay correction to the VBFNLO at NLO sample. The error bands in the ratio panels represent the statistical uncertainties of the MC samples.	69
34	Leading and subleading jet p_T generated by VBFNLO at different orders of calculation in QCD and EW: LO, EW NLO, QCD NLO, and NLO in both EW and QCD. The distributions are not corrected for tau decays.	70
35	Leading and subleading jet p_T generated by VBFNLO-NLO and VBFNLO+PYTHIA 8 compared to VBFNLO-LO.	70

36	Comparison of m_{jj} and m_{ll} differential cross sections by POWHEG + PYTHIA 8, POWHEG + HERWIG 7, VBFNLO-NLO and VBFNLO + PYTHIA 8 predictions, with their total theory uncertainties. The four predictions follow the same binning but the points are offset to make each point visible for the comparison.	72
37	Migration matrices for Higgs p_T , m_{jj} , m_{ll} and $\Delta\phi_{jj}$. The bin values correspond to the number of MC events in the truth and reconstruction bins and the fiducial-fakes category.	77
38	Profiled negative-log-likelihood of σ^{fid} using the ATLAS Run-2 dataset (black) and the Asimov dataset (blue). Dashed curve shows the likelihood profile with only statistical uncertainty, whereas solid curve shows that including all nuisance parameters.	82
39	The measured fiducial cross section at $\sqrt{s} = 13$ TeV in comparison with the theoretical predictions in the fiducial phase space obtained from the parton shower generators POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from the parton-level generator VBFNLO at LO and NLO. The vertical bands around the measurement indicate the statistical uncertainty (yellow) and the sum in quadrature of statistical and systematic uncertainties (green).	84
40	The measured background normalization factors (μ) and the correlations among the measured fiducial cross section (σ_1) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$). The Drell-Yan normalization factors are labeled as “Z+jets” for the parameter used for simultaneous estimation in SR and CRZjets, and “Z+jets CRGGF” or “Z+jets ggF” for the one used in ggF CR. The only ggF normalization factor is labeled “ggF”. Finally, the top plus diboson background is denoted by “VV+Top” for the parameter from SR and “VV+Top ggF” for the one from ggF CR.	85

41	The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for p_T^H in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.	87
42	The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell\ell}$, $\Delta y_{\ell\ell}$, Δy_{jj} and $\cos\theta_\eta^*$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.	89
43	The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell_1}$, $p_T^{\ell_2}$, m_{ll} and $\Delta\phi_{\ell\ell}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.	90

44	The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{j_1}$, $p_T^{j_2}$, m_{jj} , and $\Delta\phi_{jj}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.	91
45	Validation of the SM prediction simulated by MADGRAPH5 with the SMEFTsim package at QCD LO. The sample (blue) is showered with VBFNLO+PYTHIA 8 and is validated against VBFNLO+PYTHIA 8 (gray). The bottom panel shows the ratio with respect to the VBFNLO+PYTHIA 8 prediction. The gray (blue) error band represents the MC statistical uncertainties of the VBFNLO(MADGRAPH5) sample.	96
46	EFT parametrization fit (curve) of c_{HW} in the four $\Delta\phi_{jj}$ bins in comparison with the relative EFT corrections obtained from the EFT fiducial samples (data points). The relative EFT corrections are the ratio of the EFT and SM cross sections simulated at $c_{HW} = -0.8, -0.5, -0.2, 0.1, 0.4, 0.6, 1.0$. The error bars represent the MC statistical uncertainty and are used to perform the parametrization fit.	98
47	The relative EFT correction on $\Delta\phi_{jj}$ by the CP even c_{HW} , c_{HWB} , and c_{HB} at fixed values close to the expected $+1\sigma$ limits. The solid blue lines show the EFT corrections at order of c_i/Λ^2 (linear); the blue dashed lines show the corrections at order of c_i^2/Λ^4 (quadratic); finally the red dashed lines show the full linear plus quadratic corrections.	101

48	The relative EFT correction on $\Delta\phi_{jj}$ by the CP odd c_{HW} , c_{HWB} , and c_{HB} at fixed values close to the expected $+1\sigma$ limits. The solid blue lines show the EFT corrections at order of c_i/Λ^2 (linear); the blue dashed lines show the corrections at order of c_i^2/Λ^4 (quadratic); finally the red dashed lines show the full linear plus quadratic corrections.	102
49	The relative EFT correction on $p_T^{j_1}$ (left panel) and $p_T^{\ell_1}$ (right panel) by c_{Hq1} , c_{Hd} (top), c_{Hq3} and c_{Hu} (bottom) at fixed values close to the expected $+1\sigma$ limits.	103
50	Relative EFT correction by $\tilde{c}_{HW}$ on the $\Delta\phi_{jj}$ distribution with eight equal-sized bins (left) and with four equal sized bins (right).	107
51	The measured differential cross section (black data points) and uncertainty (gray error band) as a function of $\Delta\phi_{jj}$ (top left and top right) and $p_T^{j_1}$ (bottom) is shown together with the predictions without (blue data points) and with EFT corrections (colored histograms). The values of the Wilson coefficients correspond approximately to the observed negative limits set at two standard deviations with the full parametrizations. The contributions from the interference between the SM and EFT amplitudes (lin.) and from the total of the interference and the BSM terms (lin.+quad.) are represented as dotted and dashed lines, respectively.	110
52	Profile likelihood scan of the c_{HW} group with the linear parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).	111
53	Profile likelihood scan of the c_{Hq3} group with the linear parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).	112

54	Profile likelihood scan of the c_{HW} group with the full parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).	113
55	Profile likelihood scan of the c_{Hq3} group with the full parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).	114
56	The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for p_T^H in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.	124
57	The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell\ell}$, $\Delta y_{\ell\ell}$, Δy_{jj} and $\cos\theta_\eta^*$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.	125

58	The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell_1}, p_T^{\ell_2}, m_{ll}$ and $\Delta\phi_{\ell\ell}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.	126
59	The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{j_1}, p_T^{j_2}, m_{jj}$, and $\Delta\phi_{jj}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.	127
60	The measured background normalization factors (μ) and the correlations among the measured cross section in m_{jj} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	130
61	The measured background normalization factors (μ) and the correlations among the measured cross section in Higgs p_T bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	131
62	The measured background normalization factors (μ) and the correlations among the measured cross section in Δy_{jj} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	132

63	The measured background normalization factors (μ) and the correlations among the measured cross section in $\Delta\phi_{jj}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	133
64	The measured background normalization factors (μ) and the correlations among the measured cross section in m_{ll} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	134
65	The measured background normalization factors (μ) and the correlations among the measured cross section in $\Delta y_{\ell\ell}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	135
66	The measured background normalization factors (μ) and the correlations among the measured cross section in $\Delta\phi_{\ell\ell}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	136
67	The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{\ell\ell}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	137
68	The measured background normalization factors (μ) and the correlations among the measured cross section in $\cos\theta_\eta^*$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	138
69	The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{\ell_1}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	139
70	The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{\ell_2}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	140
71	The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{j_1}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	141

72	The measured background normalization factors (μ) and the correlations among the measured cross section in p_T^{j2} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).	142
73	The post-fit result of the nominal fit model with the BSM data is compared with the BSM prediction at particle-level. The Wilson coefficient and its value used to generate the BSM data are: $c_{HW} = 1.5$, $c_{Hq3} = 0.4$. The BSM reweighting variable is $s\Delta y_{jj}$	145
74	The post-fit result of the nominal fit model with the BSM data is compared with the BSM prediction at particle-level. The Wilson coefficient and its value used to generate the BSM data are: $c_{HW} = 1.5$, $c_{Hq3} = 0.4$. The BSM reweighting variables are m_{jj} plus Δy_{jj}	146

List of Abbreviations

- **ATLAS**: A Toroidal LHC Apparatus
- **BDT**: boosted decision tree
- **BR**: branching ratio
- **BSM**: beyond Standard Model
- **CERN**: European Organization for Nuclear Research
- **CJV**: central jet veto
- **CMS**: Compact Muon Solenoid
- **CP**: charge-parity
- **CR**: control region
- **CSC**: cathode strip chamber
- **ECAL**: electromagnetic calorimeter
- **EFT**: Effective Field Theory
- **EW**: electroweak
- **ggF**: gluon-gluon fusion
- **HCAL**: hadronic calorimeter
- **ID**: inner detector
- **JES**: jet energy scale
- **JVT**: Jet Vertex Tagger

- **LEP**: Large Electron-Positron Collider
- **LAr**: lead/liquid-argon
- **LH**: likelihood-based
- **LHA**: Les Houches Accord
- **LO**: leading order
- **MC**: Monte Carlo
- **MDT**: monitored drift tubes
- **MET**: missing transverse energy
- **MS**: Muon Spectrometer
- **MVA**: multi-variant analysis
- **NLO**: next-to-leading order
- **NNLO**: next-to-next-leading order
- **NP**: nuisance parameter
- **OLV**: outside lepton veto
- **PDF**: parton distribution function
- **PS**: Proton Synchrotron
- **PSB**: proton synchrotron booster
- **QCD**: quantum chromodynamics
- **QED**: quantum eletrodynamics
- **SCT**: Semiconductor Tracker

- **SM**: Standard Model
- **SMEFT**: Standard Model Effective Field Theory
- **SPS**: Super Proton Synchrotron
- **SR**: signal region
- **TRT**: transition radiation tracking
- **VBF**: vector boson fusion
- **VH**: associated production with a vector boson

1 Introduction

With 50 years of experimental searches and theoretical postulations, the Standard Model of particle physics is a remarkable triumph - one beautiful equation which encapsulates all known fundamental particles and their interactions. The Standard Model describes three fundamental forces - strong, electromagnetic and weak forces - mediated by the respective force carrier particles. Moreover, the Standard Model needs to permit particle mass while keeping theoretical calculations finite at large energy scales. Here, “particle mass” is the inertial mass (m), which makes up the particle rest energy given by Einstein’s famous formula: $E = mc^2$. The mechanism proposed by Brout and Englert [15], Higgs [16], and Guralnik, Hagen and Kibble [17] manifested, what is now known as, the Higgs boson. While particles were massless and traveled at the speed of light immediately after the Big Bang, about 10^{-12} seconds later, they acquired mass by interacting with the Higgs boson and stayed massive after the universe cooled [18].

To imitate the beginning of the universe, the Large Hadron Collider program was initiated in 1994, designed to accelerate and collide protons at an energy high enough to probe the mass of the Higgs boson. Using only a partial dataset of the first operational run, the ATLAS and CMS experiments both observed an excess at a mass of 125 GeV with a high significance-level of around 5 standard deviations, declaring the discovery of the Higgs boon in 2012 [19, 20]. Its couplings with several heavy particles in the theory were subsequently observed, validating that the new particle was indeed the predicted Higgs boson.

The discovery of the Higgs boson is not only satisfactory but also puzzling. For example, why is the Higgs boson so light? The Higgs mass at 125 GeV is at the boundary of a stable vs. meta-stable universe, where the vacuum state or the ground state can quantum-tunnel to another ground state in a meta-stable system. Furthermore, in the *hierarchy problem*, it is unexplained why the weak force is 10^{24} times stronger than the gravitational force or equivalently, why the Higgs mass is so much lighter than the Plank scale (the energy scale

of the hypothesized quantum gravity). These problems call for determination of the precise nature of the Higgs boson.

This thesis is one of the many “quests” for new physics signature using a collider experiment. We use the data collected by the ATLAS experiment from 2015-2018 to measure the particular production channel of vector boson fusion and decay channel of the Higgs boson to two W bosons ($H \rightarrow WW^*$). While the measurement cannot directly answer the questions posed earlier, we hope to use the reported results to narrow the window for new physics and to guide future experiments deigned for directly searching new physics candidates.

2 Theory

Particle physics studies the questions, ‘what are the most fundamental constituents of the universe’ and ‘how do they interact’ [21]. To date, the Standard Model of particle physics (SM) is the most successful model describing the past 50 years of experimental data. It provides a unified description of the most fundamental particles and forces (except for gravity). However, many remaining questions, such as the nature of dark matter, motivate the search for new physics beyond the Standard Model. In this section, we first give an overview of the SM theory in Sec. 2.1, followed by phenomenology of the Higgs boson at the Large Hadron Collider (LHC) in Sec 2.2. Finally, we review the Effective Field Theory approach to probe for new physics signatures in Sec. 2.3 and its application in the $H \rightarrow WW^*$ analysis in later sections.

2.1 The Standard Model

2.1.1 Overview

The goal of many physics studies is to provide an effective mathematical description of a physical system at a chosen scale [8]. The mathematical framework for particle physics is *quantum field theory*, which incorporates both special relativity and quantum mechanics. Fundamental particles are excited states of their quantum fields and their dynamics and interactions are governed by a *Lagrangian function*, denoted as $\mathcal{L}$ throughout this thesis. In particular, the Standard Model is a *gauge* field theory that is symmetric under various local group transformations. The Lagrangian of the Standard Model is also *renormalizable*, making calculation of observables finite and the evaluation of the model parameters accurate with a perturbative expansion [22]. The postulation of these two properties guided the development of the Standard Model in the past century and led to its extraordinary compatibility with generations of high energy experiments.

The constituents of the Standard Model are categorized into fermions and bosons. Fermions

and their anti-particles have half a unit of intrinsic angular momentum, that is $\text{spin} - \frac{1}{2}$. They make up matter in the universe. For example, nuclei are made of quarks and form atoms with electrons which are leptons. As shown in Fig. 1, there are three generations of quarks (purple) and leptons (green), respectively. Their masses increase from low to high generation (except for neutrinos whose masses are yet to be determined). While the up (u), charm (c) and top (t) quarks have $+\frac{2}{3}$ unit charge, the down (d), strange (s) and bottom (b) quarks have $-\frac{1}{3}$ unit charge. The former are generalized as the *up*-type quarks and the latter as the *down*-type quarks. In the lepton family, each charged lepton - electron, muon, and, tau - is accompanied by a neutrino with the same *lepton flavor*. Anti-leptons have anti-flavors. Lepton flavor is conserved and this is why the weak process produces lepton pairs such as $e^-\bar{\nu}_e$, $\mu^+\nu_\mu$, e^+e^- and $\mu^+\mu^-$, so that the flavor and anti-flavor cancel.

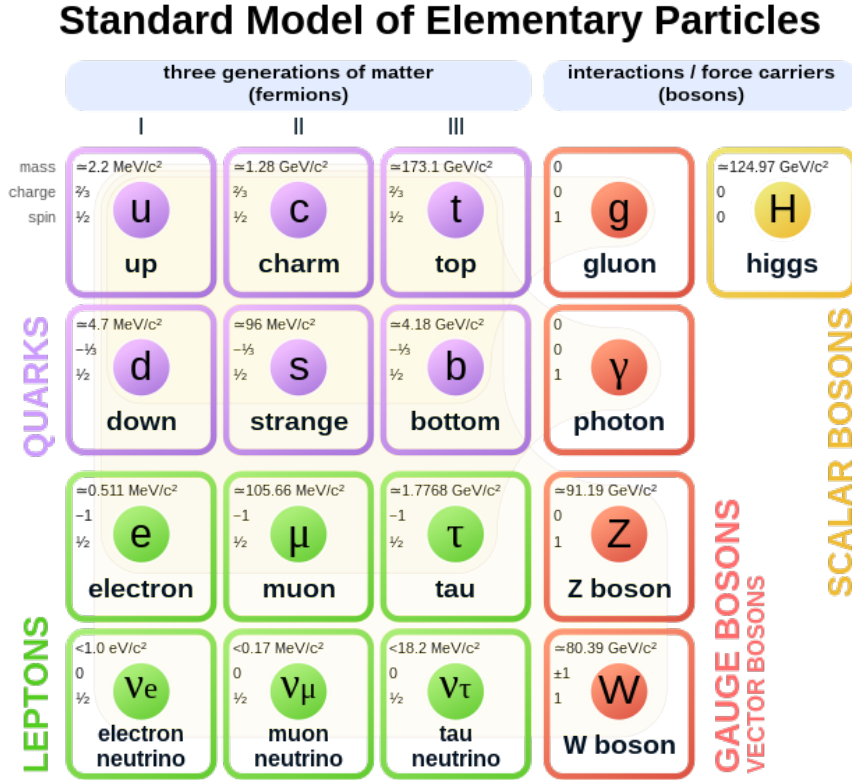


Figure 1: The Standard Model of particle physics [2].

Bosons have integer spins. The spin-1 bosons, also referred to as the gauge bosons,

mediate three fundamental forces. The gluon is responsible for the strong force, holding the nuclei together. At a typical energy scale of 1 GeV, the strong nuclear force has a range of about 10^{-15} m, which is roughly the distance for a free gluon to travel in space. The weak force is mediated by the W and Z bosons with a range of 10^{-17} m and is responsible for β -decays. Light, or electromagnetic waves, can travel infinite distance as its quanta (photons) are massless. Gravity, the fourth force, is too weak to be observed in particle physics experiments. However, the large-scale structures in the universe are formed because of the attractive gravitational force with a long range. A spin-2 boson, the *graviton*, is postulated as the quantum of gravity. Today, the quantum gravity problem remains unsolved and is beyond the scope of this thesis.

The recently discovered Higgs boson is the only spin-0 particle in the theory. Prior to the Higgs mechanism, particle physicists faced two dilemmas. First, the weak bosons were required by gauge invariance to be massless and consequently the weak force to be long-range. However, this contradicted many experimental observations at the time. Second, typically a mass term is presented in the Lagrangian for massive particles: $-m\psi\bar{\psi}$, where m is the mass and ψ and $\bar{\psi}$ are the fields of the particle and anti-particle, respectively. However, the observation of parity (symmetry under a mirror reflection) violation established that the charged-current weak interactions are *left-handed*, meaning the directions of spin and motion are opposite. Since the left-handed and right-handed components of the fields behave differently, the mass terms violate the symmetry of the theory. The introduction of the Higgs mechanism spontaneously breaks the electroweak symmetry and gives rise to the spinless boson, now known as the Higgs boson. The Higgs boson gives masses to the three weak bosons and the fermions, while keeping the theory gauge invariant.

In the following subsections, we divide the discussion into the different gauge theories relevant to the Standard Model, ordered by their appearance in the particle physics chronology.

2.1.2 Quantum electrodynamics

Quantum electrodynamics (QED) unifies the electric and magnetic fields, and the massless quanta of the electromagnetic field are photons (γ). All particles with electric charge can exchange photons. Most macroscopic physics phenomena are governed by the electromagnetic force with an infinite range since the photon is massless. The strength of the interaction is quantified by the fine structure constant α .

According to Noether's theorem, each symmetry implies a conservation law. For example, the translation invariance in space leads to conservation of momentum, and translation invariance in time implies conservation of energy. Both can be derived from the Lagrangian of classical mechanics. Similarly, the QED theory is derived by imposing phase rotation symmetry on the quantum-mechanical wave function of electron, $\psi(x) \rightarrow \psi'(x) = \psi(x)e^{i\theta(x)}$. As a result, the invariance under a global phase rotation (independent of coordinate) proved the conservation of electric charge; the invariance of a local phase rotation implied the existence of a massless field, i.e. the photon, that couples to electric charge. The local gauge invariance is preserved in QED since the photon is massless. This transformation group is denoted as $U(1)_{\text{EM}}$.

The success of QED based on a local phase symmetry became a prototype for constructing gauge theories with conditional symmetries, motivating successive development of gauge theories of the Standard Model.

2.1.3 Quantum chromodynamics

The strong force is governed by the quantum chromodynamics (QCD) gauge theory with a $SU(3)_C$ symmetry, where the subscript C stands for *color*. Analogous to QED, the gluon as the strong force carrier couples to particles with color charge, which is the conserved quantity in QCD. The color charges are named after the three primary colors: red, green and, blue. Free quarks form color-neutral particles predominately in two ways: a three-quark bound state, the *baryon*, with the combination of the three color charges, or a quark-antiquark

bound state, the *meson*, with zero net color charge. Collectively, the two bound states are referred to as *hadrons*. Unlike the photon, which is electrically neutral, gluons are color-charged and hence can couple to other gluons.

The strong coupling constant (α_s) is dependent on the energy scale (μ):

$$\alpha_s(\mu) = \frac{12\pi}{(32 - 2n_f) \cdot \ln(\mu^2/\Lambda_{\text{QCD}})}. \quad (1)$$

Here, n_f is the number of types of quarks whose masses are below the energy scale μ and the QCD scale (Λ_{QCD}) is measured to be approximately 0.2 GeV. The running of strong coupling strength results in two crucial phenomena for strong interaction. First, a color-charged particle cannot be isolated or directly observed at low-energy scale, which is known as *color confinement*. The energy tension while separating a color-neutral state increases so drastically that it produces a quark-antiquark pair, forming new color-neutral composite particles. On the contrary, the strong interaction reaches *asymptotic freedom* at high energy scales (short distance) and the exchange of gluons requires minimal energy. Asymptotic freedom explains the abundant gluon-induced QCD processes at the LHC, while confinement induces the *parton shower* and *hadronization* for an outgoing quark or gluon.

Another important concept of QCD is the parton distribution function (PDF). In this thesis, it describes the momentum fraction of each *parton* that makes up the proton, and a parton can be either a quark or a gluon. At low energy, a proton consists of two u and one d quarks, the valence partons. On the other hand, protons at high energy are made of not only the valence partons but also of nonvalence partons, such as sea quark-antiquark pairs or gluons. The sea partons add up to be flavor-neutral but each can be any type. The appropriate choice of a PDF model improves the accuracy of the predictions from phenomenology. Usually the development of PDFs comes from a combination of simulation and experimental data at a given scale.

2.1.4 The electroweak theory and the Higgs mechanism

The electroweak theory unifies the weak and electromagnetic forces at high energy scale and this is referred to as the electroweak symmetry. Like other gauge theories introduced in the previous sections, electroweak gauge theory needed to choose a symmetry and subsequently obtain gauge bosons that couple to a conserved charge. The correct symmetry group choice turned out to be $SU(2)_L \times U(1)_Y$, where L represents for left-handedness and Y for the *weak hypercharge* which the electroweak symmetry conserves. To solve the mass problem, the electroweak symmetry needs to be broken to obtain the separate weak and electromagnetic forces: $SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}$. This is called *spontaneous symmetry breaking*, a concept originated from ferromagnetism. Ferromagnetic materials undergo a phase transition as the temperature decreases to below the critical temperature and the electron spins will align in the same direction, breaking the rotational symmetry of the system.

With the goal to induce spontaneous symmetry breaking into the electroweak theory, a gauge invariant potential (V) as function of a complex doublet scalar field (Φ) is introduced

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad (2)$$

where the field is given as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (3)$$

A positive μ^2 corresponds to the unbroken symmetry where the potential only has one minimum, which is the ground state or vacuum state. On the other hand, when μ^2 is negative, the ground state has a continuous degeneracy. As a 2-dimensional representation of the potential given in Eq. 2, Fig. 2 illustrates the potential for a singlet complex field $V(\phi) = \mu^2 |\phi|^2 + \lambda |\phi|^4$, shaped like a “Mexican hat”. The vacuum expectation value is $v = |\mu|/\sqrt{\lambda}$. The choice for a specific minimum is arbitrary, just like the orientation of magnetization in ferromagnetism. A conventional choice is to set $\phi_3 = v$ and other degrees

of freedom to zero, and the Higgs field (H) emerges as the excitation from the vacuum state

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

Now, the rotational symmetry is broken (or hidden). Substituting this definition back into the Lagrangian immediately gives the gauge boson mass terms and their interaction with the Higgs boson. The coupling strength is proportional to the square of the boson mass (M_V): $2M_V^2/v$.

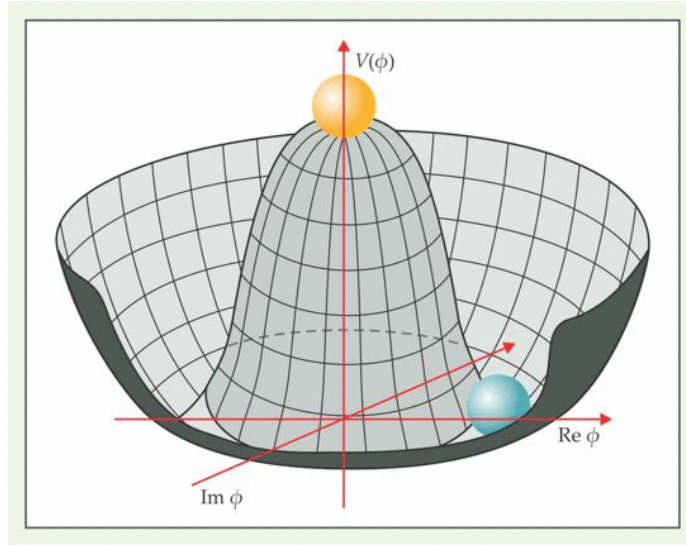


Figure 2: The Higgs potential as a function of the real (Re) and imaginary (Im) part of the field ϕ [3]. The yellow ball at the center represents the electroweak theory before spontaneous symmetry breaking where the gauge bosons are massless; the blue ball represents solution after breaking the electroweak symmetry.

Fermions get their masses through the Yukawa theory which is a renormalizable gauge theory describing the interactions between fermions and spinless particles. As an example, the Yukawa coupling between the electron field (e) and the Higgs field is [23]

$$\mathcal{L}_{\text{Yukawa},e} = -\frac{\zeta_e v}{\sqrt{2}} \bar{e}e - \frac{\zeta_e}{\sqrt{2}} H \bar{e}e, \quad (5)$$

where the Yukawa coupling constant ζ_e is a phenomenological parameter. The first term in Eq. 5 is the electron mass term, where electron mass m_e is $\zeta_e v / \sqrt{2}$. The second term

represents the coupling of Higgs with the electron-positron pair with a coupling constant proportional to the fermion mass $\frac{\zeta_e}{\sqrt{2}} = m_e/v$. Generally, heavier particles have stronger coupling to the Higgs boson. Furthermore, the Higgs mechanism enables quark coupling in the “off-diagonal” directions, allowing heavier generations to weakly decay to lighter ones. Finally, the Higgs mass is not predicted by the theory. We will discuss the experimentally measured Higgs properties and other phenomenologies in the next section.

2.1.5 From the Lagrangian to observables

In summary, the Standard Model Lagrangian encapsulates all particle interactions, which are the gauge boson interactions, fermion interactions with the gauge bosons, the Yukawa coupling of fermions with the Higgs field, and the Higgs self-interaction and coupling to the gauge bosons:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge boson}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{Higgs}}. \quad (6)$$

The *matrix element* method is the standard strategy to calculate *observables*, i.e. the quantities to be observed at an experiment. The *Feynman rules* provide the mathematical representations for different particle fields and their interactions (with a graphical representation called the Feynman diagram). One can construct the matrix element, denoted as $\mathcal{M}$, according to the fields in the initial and final states and sometimes the intermediate propagator fields. The square of the matrix element is directly related to the most common observable in collider physics experiments, the *scattering cross section* denoted as σ . The cross section is a measure of the probability of a particle-interaction process and has units of area, usually specified in barns (b), where $1 \text{ b} = 10^{-28} \text{ m}^2$. The calculation of a cross section depends on the kinematics of the incoming and outgoing particles. Therefore, this multi-dimensional calculation is usually done by Monte Carlo simulation instead of a numerical calculation.

Experimentally, the cross section can be obtained from the number of counted events for

a process (N)

$$\sigma = \frac{N}{L_{\text{int}}}, \quad (7)$$

where L_{int} is the *integrated luminosity*. The luminosity (L), or instantaneous luminosity, gives the incident particle flux per unit time, which is related to the performance of a particle accelerator. The integrated luminosity, with units of inverse area or inverse barn, describes the incident flux provided by the accelerator over a period of time:

$$L_{\text{int}} = \int L dt. \quad (8)$$

2.2 The Higgs boson measurements

In this section, we provide a summary of the Higgs boson phenomenology at the Large Hadron Collider (LHC), focusing on its production and decay modes. The different production and decay modes are generally studied separately as the final state particles differ and therefore need different analysis strategy to maximize the sensitivity.

The predictions for the production and decay of the Higgs boson are dependent on its mass. As mentioned in Sec. 2.1.4, the mass of the Higgs boson is not determined by the theory. The combined measurement by the ATLAS and CMS experiments at the LHC of the Higgs mass yields [24]

$$m_H = 125.09 \pm 0.21(\text{stat}) \pm 0.21(\text{sys}) \text{ GeV}. \quad (9)$$

2.2.1 Main Higgs production modes

The production cross sections of the Higgs boson depend on the energy of the collider. The predicted cross sections of several Higgs production modes as a function of the center-of-mass energy at fixed Higgs mass of 125 GeV are given in Fig. 3. Ordered by the cross section from large to small, the production modes and their representation in Fig. 3 are

- gluon-gluon fusion (ggF): $pp \rightarrow H$, where p stands for proton;
- vector boson fusion (VBF): $pp \rightarrow qqH$, where q stands for quark;
- associated production with a vector boson (VH): $pp \rightarrow WH$ and $pp \rightarrow ZH$;
- associated production with heavy quarks (bbH, ttH): $pp \rightarrow bbH$, and $pp \rightarrow ttH$.

The Feynman diagrams for these processes are shown in Fig. 4.

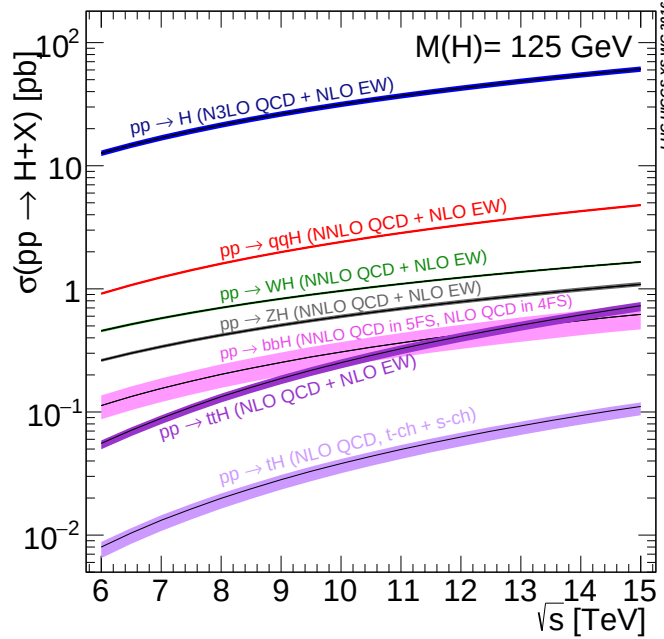


Figure 3: The cross section of Higgs production for the dominant modes as a function of the center-of-mass energy at fixed Higgs mass of 125 GeV [4]. The error band includes all theoretical uncertainties.

These leading production modes are observed by the ATLAS experiment [6] as shown in Fig. 5, and by the CMS experiment [7] as shown in Fig. 6 (left). Both collaborations report the data-over-simulation ratio, called the *signal strength*. The SM prediction is given by the red (black) vertical line in the ATLAS (CMS) results and is centered at one. The cross sections measured by both experiments are compatible with the SM prediction, proving the success of the Higgs modeling.

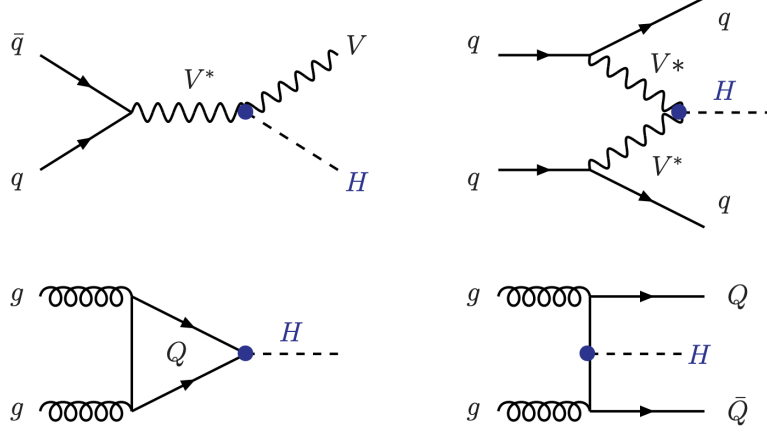


Figure 4: The Feynman diagrams of several Higgs production modes: associated production with a vector boson (upper left); vector boson fusion (upper right); gluon-gluon fusion (lower left); associated production with heavy quarks (lower right) [5]. The vector bosons are denoted as V and virtual particles denoted with a star (*).

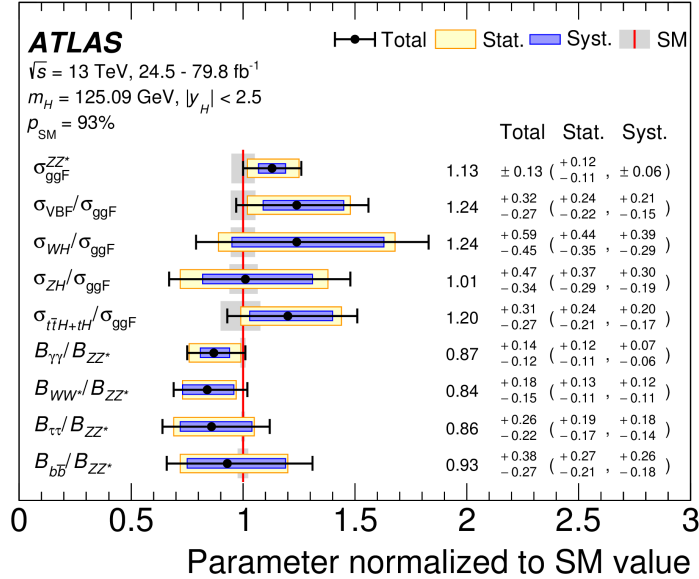


Figure 5: The signal strengths for the Higgs production cross section (top five) and branching-ratio (bottom four) signal strength measurements by the ATLAS collaboration [6]. The measurements of all production and decay channels use the gluon-gluon fusion production and Higgs to Z bosons decay as the baseline. The observed cross sections and branching ratios (black data points) are compatible with the SM predictions (red vertical line) within the uncertainties of the measurements (black error bars) and the theory predictions (gray error bands).

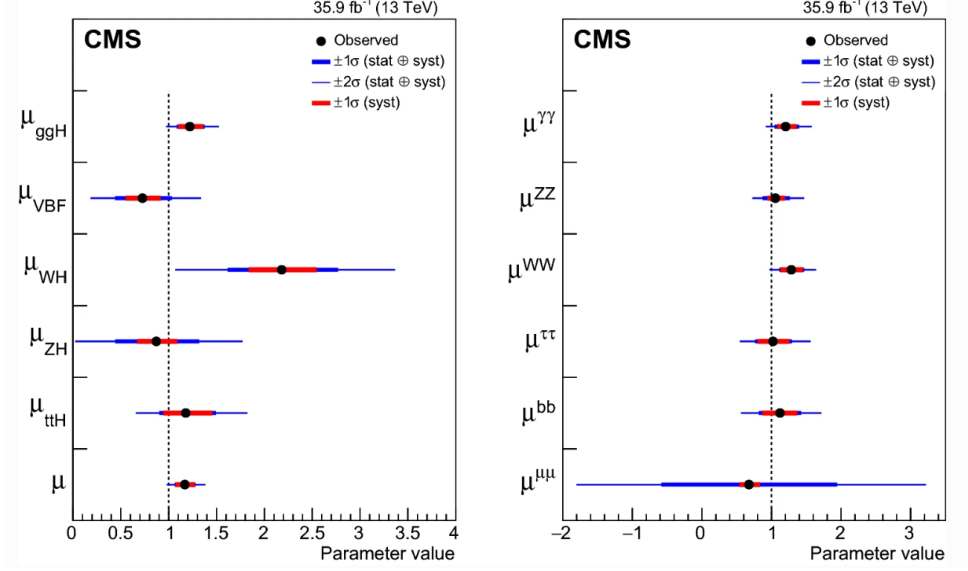


Figure 6: The signal strengths for the Higgs production cross section (left) and branching-ratio (right) signal strength measurements by the CMS collaboration [7]. The observed cross sections and branching ratios (black data points) are compatible with the SM predictions (black vertical line) within the uncertainties of the measurements (blue error bars), with a slight tension for the WH production mode which has a 2σ deviation.

2.2.2 Main Higgs decay modes

The probability for a particle to decay in a particular channel is called the branching ratio (BR). The branching ratio of the Higgs boson is defined as the ratio between the *partial decay width* ($\Gamma_{H \rightarrow X}$) and the *total decay width* (Γ_H) of the Higgs boson:

$$\text{BR} = \frac{\Gamma_{H \rightarrow X}}{\Gamma_H}. \quad (10)$$

Decay widths have units of inverse time, and the total decay width is the inverse of the lifetime of the particle. The partial decay widths of all possible decay channels add up to the total width, and therefore the branching ratios add up to one. Similar to the scattering cross section, the partial width of each decay channel can be derived from the matrix element calculation. The predicted branching ratios of the Higgs boson are presented in Fig. 7.

Several decay modes are measured by both the ATLAS experiment (Fig. 5 bottom) and the CMS experiment (Fig. 6 right), and the results are in agreement with SM predictions.

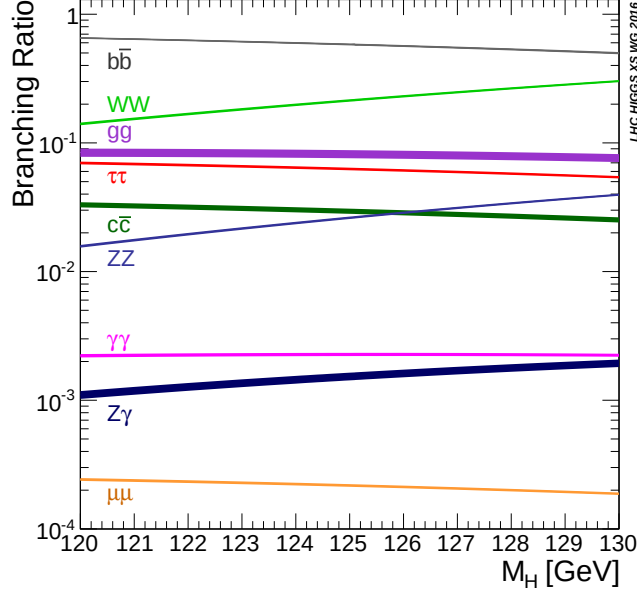


Figure 7: The predicted branching ratios of the Higgs boson decay for the dominant decay channels as a function of the Higgs mass [4].

2.3 Effective Field Theory

In this section, we review a model independent approach to physics beyond the Standard Model: the Effective Field Theory (EFT). This formalism describes the effect of any short-range heavy resonance at low energy scale, without assuming any of its properties - thus a model independent formalism. This is the result of *integrating out* degrees of freedom at given scale. As an example, the study of molecular interactions does not need to consider the QCD interactions of quarks and gluons inside the protons and neutrons, since their effect are integrated out so that the nuclei can be treated as the smallest unit. In the EFT approach, the degree of freedom that is “integrated out” is a potential new resonance which can only be a virtual propagator due to its large mass. According to the Feynman rule, a propagator field with 4-momentum q and mass M contributes to the matrix element a factor proportional to $\frac{1}{q^2 - M^2}$, which can expand as:

$$\frac{1}{q^2 - M^2} = -\frac{1}{M^2} \left(1 + \frac{q^2}{M^2} + \mathcal{O}(M^{-4}) \right). \quad (11)$$

To first order, the heavy particle field is decoupled from the theory. However, it modifies the coupling of the SM particles with a factor of $1/M^2$. This is called the Appelquist-Carazzone theorem [25]. Without having to introduce a new particle field, one can build higher-dimensional *operators* suppressed by powers of M :

$$\frac{1}{M^{n-4}} c_i^{(n)} \mathcal{O}_i^{(n)}, \quad (12)$$

where $\mathcal{O}_i^{(n)}$ is a dimension- n operator (with units of energy to the n^{th} power) and $c_i^{(n)}$ is the corresponding *Wilson coefficient* which is the modification factor coming from the heavy field integration.

In this section, we first discuss a famous example of the EFT application, the Fermi theory, as the low-energy effective theory of the weak interactions [8]. Next, we re-introduce the Standard Model as a low-energy effective theory of the full gauge theory at much higher energy scales beyond the LHC [1].

2.3.1 Fermi theory

Prior to the genesis of the W boson theory, the matrix element of weak interaction was expressed as a 4-particle interaction [8]:

$$\mathcal{M}_{\text{Fermi theo.}} = G_F g_{\mu\nu} [\bar{\psi}_3 \gamma^\mu (1 - \gamma^5) \psi_1] [\bar{\psi}_4 \gamma^\nu (1 - \gamma^5) \psi_2], \quad (13)$$

where G_F is the Fermi constant and describes the weak coupling strength, and the terms in the square bracket are the left-handed charged leptons and neutrino couplings. The Feynman diagram of such a 4-particle interaction is demonstrated on the right of Fig. 8.

On the other hand, the full theory of beta decay is mediated by a W field (Fig. 8 left) and its matrix element is written as:

$$\mathcal{M}_{\text{full}} = - \left[\frac{g_W}{\sqrt{2}} \bar{\psi}_3 \frac{1}{2} \gamma^\mu (1 - \gamma^5) \psi_1 \right] \cdot \left[\frac{g_{\mu\nu} - q_\mu q_\nu / m_W^2}{q^2 - m_W^2} \right] \cdot \left[\frac{g_W}{\sqrt{2}} \bar{\psi}_4 \frac{1}{2} \gamma^\nu (1 - \gamma^5) \psi_2 \right], \quad (14)$$

where g_W is the coupling constant of the weak interaction and m_W is the mass of the W boson. By doing the same Taylor expansion as Eq. 11, we obtain the Fermi theory matrix element in terms of g_W

$$\mathcal{M}_{\text{Fermi theo.}} = \frac{g_W^2}{8m_W^2} g_{\mu\nu} [\bar{\psi}_3 \gamma^\mu (1 - \gamma^5) \psi_1] [\bar{\psi}_4 \gamma^\nu (1 - \gamma^5) \psi_2]. \quad (15)$$

The two weak coupling vertices contracted and formed one interaction point of four light particles. From this effective theory, we obtained an important relationship of the Fermi constant and the weak coupling strength at low energy limit:

$$G_F = \frac{\sqrt{2}g_W^2}{8m_W^2}. \quad (16)$$

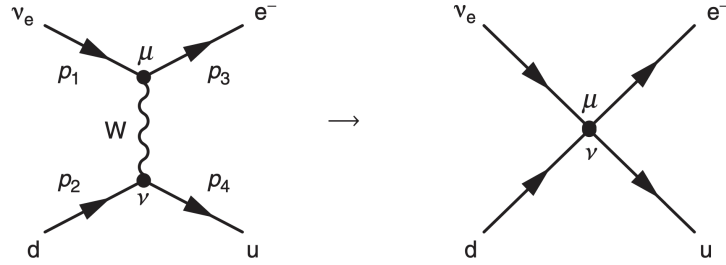


Figure 8: The electroweak theory reduced to the Fermi theory at low energy [8].

By comparing this example with Eq. 12, we can find $G_F m_W^2$ is the Wilson coefficient of this effective theory. Integrating out the W field in Eq. 14, the relation in Eq. 16 is recovered.

With the mass of the W boson known to a great precision today, theoretical models below the electroweak scale can use Fermi theory as a *top-down* EFT application. This type of application correctly describes the physics phenomena at low energy with less complexity.

2.3.2 Standard Model Effective Field Theory

Unlike top-down EFT applications such as the Fermi theory, a *bottom-up* EFT is applied when the mass of the propagator is not known. A cut-off energy scale Λ is usually chosen in the bottom-up application and its value is expected to be between the energy scale of

the experiment and the mass of new heavy particle. The goal of this approach is to access the new physics effect below the scale Λ . Therefore, precision measurements are needed to capture small deviations from the SM that are not hidden by statistical fluctuation. Above Λ , the effective theory is not valid and significantly deviates from the full theory. The EFT approach is an *indirect search*, because this method only “narrows the window” for new physics but cannot make direct discovery of new physics. In a *direct search*, an invariant mass spectrum is typically used to search for a new resonance. In the effective theory, the invariant mass cannot be obtained since the new particle fields are integrated out from the Lagrangian.

The effective theory for the Standard Model is required to obey the following prescriptions [26]:

- its gauge group contains $SU(3)_C \times SU(2)_L \times U(1)_Y$;
- its constituents are the SM fields;
- it reduces to the SM at low energy scales;
- it respects the unitarity bound and analyticity.

The last statement stems from the unitarity of the SM operators which implies the amplitude cannot grow too fast with energy scale.

The EFT formalism used in this analysis is the Standard Model Effective Field Theory (SMEFT) [1]. The operator basis of SMEFT is the first non-redundant operator basis at dimension-6, the Warsaw basis [27]. Dimension-5 and dimension-7 operators are not considered since they violate baryon number conservation, and the complete set of dimension-8 operators is still under theoretical development. The SMEFT Lagrangian is expressed as

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} \mathcal{O}_i. \quad (17)$$

The SMEFT model is assumed to have gauge symmetry and flavor symmetry.

Many particle physics experiments choose SMEFT as the interpretation formalism. Therefore, analyses targeting different processes can be connected through SMEFT. The direct merit of combining measurements is the precision improvement. Additionally, we should approach new physics from a global perspective as the same underlying physics can affect multiple SM processes. There have been examples of combined EFT analysis, such as the combination of the EFT interpretation on both the signal and the dominant background process [28], the combined EFT interpretation in the Higgs sector [29], and finally the combination across collaborations [30]. After collecting the limits on the Wilson coefficients, they can be “translated” into meaningful physical parameters [31], help to exclude disfavored phase space of new physics, and guide future experimental searches.

3 Apparatus

Established in 1954, the European Organization for Nuclear Research (CERN) hosts the world’s largest particle accelerator, the Large Hadron Collider (LHC). There are currently eight experiments investigating a wide range of physics at the LHC. ATLAS and CMS are the largest experiments and co-discovered the Higgs boson in 2012. This thesis uses data collected by the ATLAS detector between 2015-2018.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [32] is a two-ring circular hadron accelerator and collider with a circumference of 26.7 km. It is designed to provide proton-proton (pp) collisions at a peak instantaneous luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The LHC can also collide heavy ions such as lead nuclei at lower luminosities, typically $10^{27} \text{ cm}^{-2}\text{s}^{-1}$.

The LHC project was approved by the CERN Council in December 1994 and became the successor of the Large Electron-Positron Collider (LEP), re-using the LEP tunnels. The LHC tunnel lies 50 to 175 meters underground and spans through the Swiss and French border. The first run of the LHC, referred to as Run-1, started in 2010 at a center-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$ and ramped up to 8 TeV later. The LHC was shut down between 2013-2015 to enhance both the detectors and accelerators. Following the long shutdown, the LHC started its Run-2 from 2015 to 2018 at $\sqrt{s} = 13 \text{ TeV}$. The LHC expects to reach 14 TeV in the next operation.

The proton acceleration is a multi-step process [33] prior to the TeV-scale energy in the LHC as sketched in Fig. 9. The source of the protons are hydrogen molecules, which are separated into protons and electrons in an intense electric field. The protons are injected to the accelerator chain in *bunches*, which are groups of approximately 10^{11} protons. The protons are first accelerated by the LINAC-2 linear accelerator to 50 MeV; next, injected into the proton synchrotron booster (PSB), the protons are accelerated to 1.4 GeV; after the

PSB, the Proton Synchrotron (PS) boosts the proton to 26 GeV; finally, the protons enter the 7-km long Super Proton Synchrotron (SPS) to be accelerated to 450 GeV. From this stage and on-ward, the LHC will carry on the succeeding acceleration to multi-TeV scale using superconducting radio frequency cavities. Inside the tunnel, the protons are guided by dipole magnets and focused by quadrapole magnets.

There are four interaction points where the two beams meet and collide, and the collisions are recorded by the four detectors (ATLAS, CMS, LHCb, and ALICE) surrounding these interaction points. The ATLAS [10] and CMS [34] experiments are designed to perform a large variety of new physics searches and precision measurements of the Standard Model processes. On the other hand, the ALICE experiment [35] is dedicated to heavy-ion collisions and the study of quark-gluon plasma, and the LHCb experiment [36] investigates flavor physics.

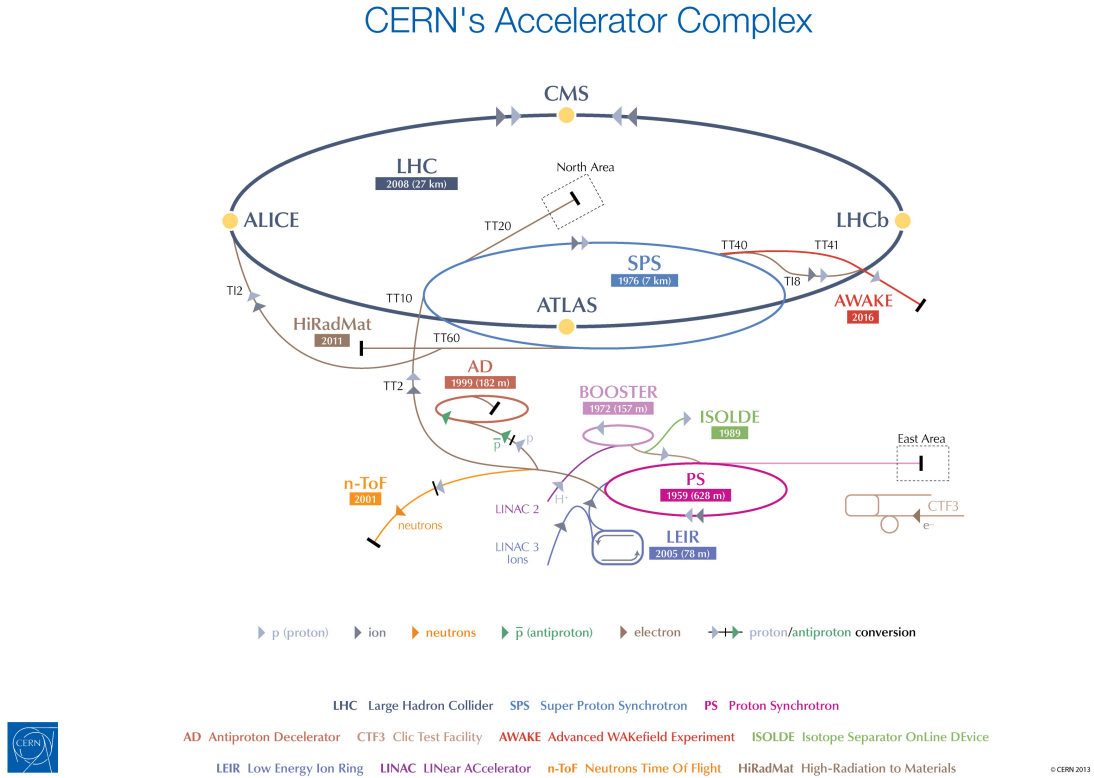


Figure 9: Schematic layout of the acceleration chain of the LHC and the main detectors at the four interaction points [9].

3.2 The ATLAS detector

ATLAS stands for A Toroidal LHC ApparatuS. It is a general purpose detector that can probe both proton-proton and heavy ion collisions [10]. As shown in Fig. 10, the ATLAS detector is a multi-layer cylindrical-shaped detector, 44 m long and 25 m wide, making it the largest detector at the LHC ring. The layers, or sub-detectors of ATLAS are concentric around the beamline. The collisions take place at the center of the detector and the outgoing particles interact with the detector materials, leaving signatures such as tracks or energy measured by the sub-detectors. The innermost detector closest to the interaction point is the inner detector (ID) which tracks charged particles. The tracker is immersed in a 2 T axial magnetic field which bends the particle trajectories to measure their momentum. Further away from the interaction point are the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). Finally, the Muon Spectrometer (MS) is the outermost layer and provides a second momentum measurement of the muons since they are not stopped by the previous detector materials. The momentum measurement in the MS relies on the toroidal magnetic field in this region.

ATLAS coordinate system

ATLAS analysis uses a right-handed coordinate system. The center of the detector is the nominal interaction point and defines the origin of the coordinate system. The x -axis points from the interaction point to the center of the LHC ring, the y -axis points upward, and the beamline direction defines the z -axis. The entire ATLAS detector is symmetric in the forward (positive z) and backward (negative z) directions. The azimuthal angle ϕ is measured around the beamline in the x - y plane and the polar angle θ is measured with respect to the z -axis. However, the pseudorapidity variable η is more commonly used than the polar angle θ :

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right). \quad (18)$$

The central detector region closer to the interaction point has lower absolute pseudorapidity value $|\eta|$, whereas the forward region of the detector has higher value. The angular separation variable (ΔR) is defined using pseudorapidity:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}, \quad (19)$$

where $\Delta\phi$ and $\Delta\eta$ describe the angle difference between two positions inside the detector. While the detector has nearly 2π coverage in ϕ , it is bounded at $|\eta| < 4.5$ ($1.3^\circ < \theta < 178.7^\circ$).

In addition to being used for describing detector dimensions, pseudorapidity originates from the rapidity variable, which is a measurement of whether the outgoing particle is more central (more perpendicular to the z -axis) or more forward (more parallel to the z -axis). The rapidity (y) is calculated as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (20)$$

where E is the total energy and p_z is the z component of the particle momentum. Rapidities are additive under a Lorentz boost in the z -direction. For a massless particle, Eq. 20 reduces to Eq. 18, that is $\eta = y$.

A common measurement on a detected particle is the transverse momentum (p_T) which is directly calculated from the curvature of the particle trajectory in the x - y plane. Mathematically, it is: $p_T = p \sin \theta = \sqrt{p_x^2 + p_y^2}$, where the x and y indices indicate the x and y components of the momentum.

Inner Detector

From the collision point, roughly 1000 particles arrive at the ATLAS detector every 25 ns, imposing significant challenges on the tracking detector. The inner detector (ID) is composed of the Pixel detector [37], the Semiconductor Tracker (SCT) [38] and the transition radiation straw-tube tracking (TRT) detector [39], ordered by distance from the beamline. The entire

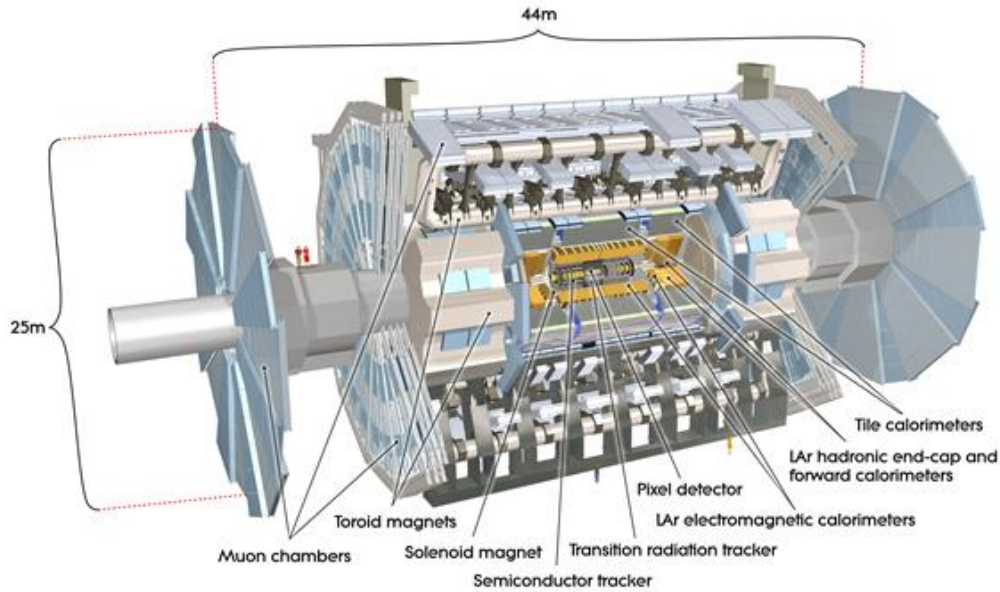


Figure 10: The ATLAS detector and its composite sub-detectors [10].

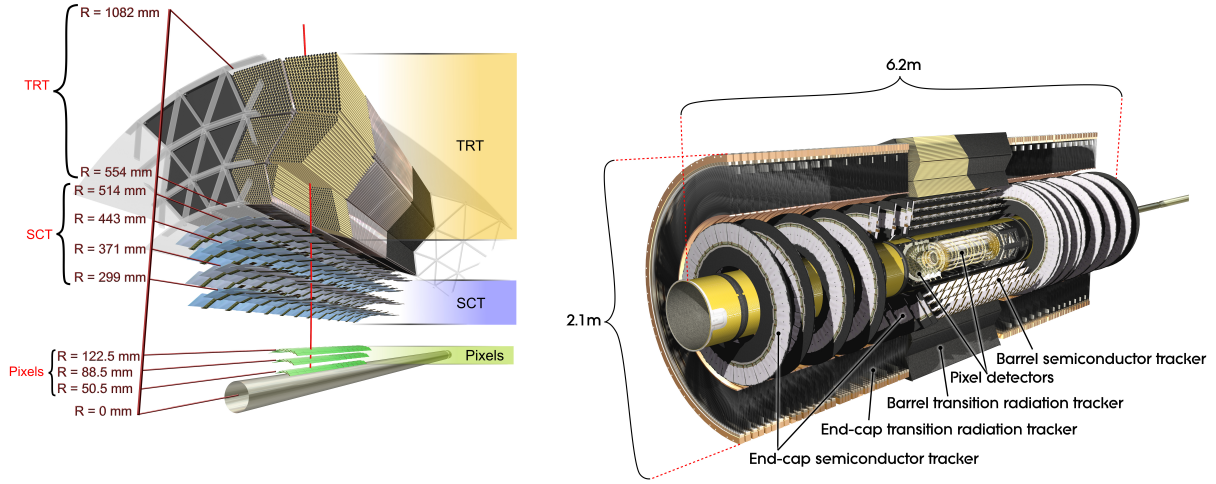
ID extends to $|\eta| = 2.5$. The layout of the three components is presented in Fig. 11.

The Pixel detector has three layers and contains in total 1744 pixel sensors with a minimum pixel size of $50 \times 400 \mu\text{m}^2$ to deliver the finest granularity since the Pixel detector stands the closest to the interaction points. Moving outward, the SCT has silicon microstrip sensors with a mean strip pitch of $80 \mu\text{m}$, providing four additional space points for each track. Finally, the TRT made of 4-mm diameter straw-tubes supplies additional radial (R) and azimuthal (ϕ) measurements up to $|\eta| = 2.0$.

During the long shutdown between Run-1 and Run-2, the Insertable B-Layer [40] made of pixel sensors was installed between the beamline and the Pixel detector. The fourth pixel layer is motivated by the tracking requirements and the high *pile-up* (number of simultaneous collisions). The pixel size is $50 \times 250 \mu\text{m}^2$, even smaller than the Pixel detector.

Calorimeters

The calorimeter system measures energy up to pseudorapidity $|\eta| = 4.9$. The inner electromagnetic calorimeter (ECAL) [41] consists of highly granular lead/liquid-argon (LAr)



(a) Cross-sectional view of the inner detector.

(b) Cut-away view of the inner detector.

Figure 11: The ATLAS inner detector and the sub-detectors [10].

calorimeters in the barrel and endcap region. The ECAL extends to $|\eta| = 3.2$, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in the upstream detector material. The hadronic calorimeter (HCAL) [42] consists of steel/scintillator tiles in pseudorapidity range $|\eta| < 1.7$, copper/LAr endcaps in the range $1.5 < |\eta| < 3.2$, and a tungsten/LAr forward calorimeter ranging from $3.2 < |\eta| < 4.9$. The subsystems of the ATLAS calorimeter are illustrated in Fig. 12.

Muon Spectrometer

The outermost ATLAS sub-detector is the Muon Spectrometer (MS) as muons are able to penetrate all the previous detector layers. Therefore, the MS gives a second momentum measurement for the muons deflected by the magnetic field generated by the superconducting toroids. The MS is a gaseous detector, in which an incident muon will ionize the gas. In the range $|\eta| < 2.0$, the muons will travel through the monitored drift tubes (MDT), each of which consists of straw-tubes with gas inside and a wire in the center to collect ionization electrons. In the pseudorapidity range, $2.0 < |\eta| < 2.7$, the cathode strip chambers (CSC) have segmented cathode into strips to allow higher precision measurement. The MS has a

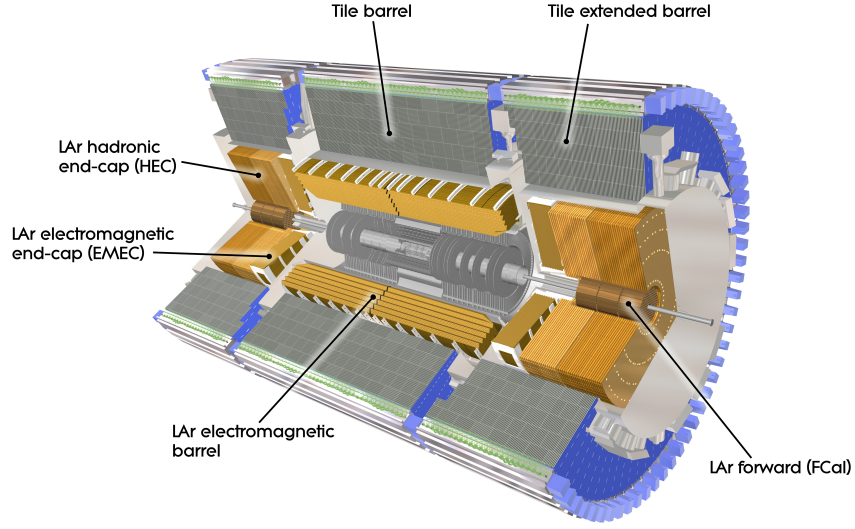


Figure 12: Cut-away view of the ATLAS calorimeter system [10].

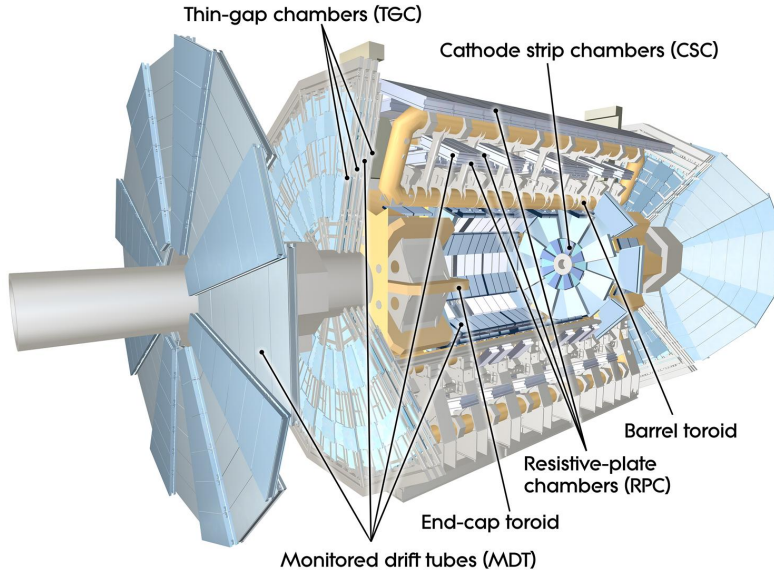


Figure 13: Cut-away view of the ATLAS Muon Spectrometer [10].

trigger system consisting of Resistive Plate Chambers up to $|\eta| < 1.05$ and thin gap chambers beyond $|\eta| = 1.05$. The components of the MS are shown in Fig. 13.

Trigger system

Given the 40 MHz bunch crossing rate and 25 interactions on average per bunch crossing, ATLAS faces the challenge of storing an enormous amount of data. To only keep a manageable amount, the ATLAS trigger system makes quick decisions whether to record the event permanently or not. There are two subsystems for triggering: Level-1 trigger and high-level trigger [43]. The Level-1 trigger is hardware-based, selecting high energy physics objects from the MS and the calorimeters in a defined region of interest. The Level-1 trigger information is then passed to software-based high-level trigger, based on the outcome of its reconstruction algorithms. The trigger system reduces the rate to about 1kHz. The $H \rightarrow WW^*$ analysis uses a combination of single-lepton triggers and a dilepton $e - \mu$ trigger in order to maximize the total trigger efficiency.

3.3 Physics object reconstruction in the ATLAS detector

Different types of particles leave unique signatures in the ATLAS detector. Reconstruction is a process of “translating” signals from one or more sub-detectors into meaningful information about the outgoing particles. In this section, we detail the reconstruction for the objects relevant to the $H \rightarrow WW^*$ analysis.

3.3.1 Tracks and vertices

The ID serves as the primary detector for *track* and *primary vertex* reconstruction, where a primary vertex is expected to be the pp interaction position and its reconstruction follows the track reconstruction [44].

The track of a charged particle is circular in the transverse direction in the x - y plane due to the homogeneous magnetic field. Track reconstruction adopts both pattern recognition and track fitting techniques and follows the main *inside-out* reconstruction and a successive *outside-in* sequence [45]. In the inside-out approach, the track segments are first built using only hits from the silicon trackers. Based on these smoothed and cleaned track segments,

an extension algorithm will look for compatible sets of measurements provided by the TRT. The TRT hits are added as an extension, and the silicon-based track is not modified. This approach is limited when particles from secondary interactions do not interact with the silicon detector. In this case, the outside-in sequence is applied, where the track pattern recognition begins in the TRT and back-tracks into the silicon detector.

Charged particle tracks are extrapolated inward and assigned to vertices [44]. In ATLAS, the vertex with the highest quadrature sum of track p_T ($\sum_{\text{track}} p_T^2$) becomes the primary vertex, whereas the rest of the reconstructed vertices are considered to have originated from pile-up events. Therefore, precise track reconstruction determines its correct association with other tracks and the correct identification of the primary vertex.

In general, a charged particle track is characterized with the following parameters: the curvature represented by the ratio of charge and transverse momentum (q/p_T); the transverse impact parameter (d_0) given by the distance of the closest approach to the primary vertex in the x - y plane and the longitudinal impact parameter (z_0) in the z -direction; the azimuthal angle (ϕ_0), and the polar angle (θ_0) of the particle direction at the closest approach point. Analyses usually impose criteria on these parameters to improve sensitivity to the physics process of interest.

3.3.2 Electrons

Due to its low mass, an electron with energy greater than 10 MeV loses its energy predominantly via Bremsstrahlung, where photons are radiated when the electron is deflected by nuclei in the vicinity. The energetic photon can subsequently convert to electron-positron. While this chain of processes can occur inside the ID, it is the most significant inside the ECAL due to the abundance of heavy atomic nuclei, generating an electromagnetic shower. Experimentally, electron reconstruction is provided by three components: energy deposition clusters in the ECAL, reconstructed track in the ID, and the close matching in $\eta \times \phi$ space between the track and clusters [11].

While the reconstruction efficiency for electrons with transverse energy $E_T > 15$ GeV is consistently high (97-99%), the *identification* and *isolation* efficiencies vary across the E_T spectrum. Electrons from the primary vertex, i.e. the hard-scattering vertex, are *prompt*, as opposed to non-prompt electrons, which originate from hadronic activity considered as background. The identification of prompt electrons in the pseudorapidity range $|\eta| < 2.47$ is likelihood-based (LH), taking into consideration the measurements from the tracking, the calorimetry and their combination. This multi-variant approach constructs separate signal and background probability density functions using the $Z \rightarrow ee$ and $J/\psi \rightarrow ee$ simulated events. To cover various analysis situations for prompt electrons, the identification delivers four *working points* at fixed values of the LH discriminant. Ordered by increasing threshold on the LH discriminant, i.e. increasing purity of the prompt signal over background, the four working points are: VeryLoose, Loose, Medium, and Tight. As shown in Fig. 14, the identification efficiency is measured in bins of η and E_T , and as the simulated signal purity becomes higher the identification becomes less efficient.

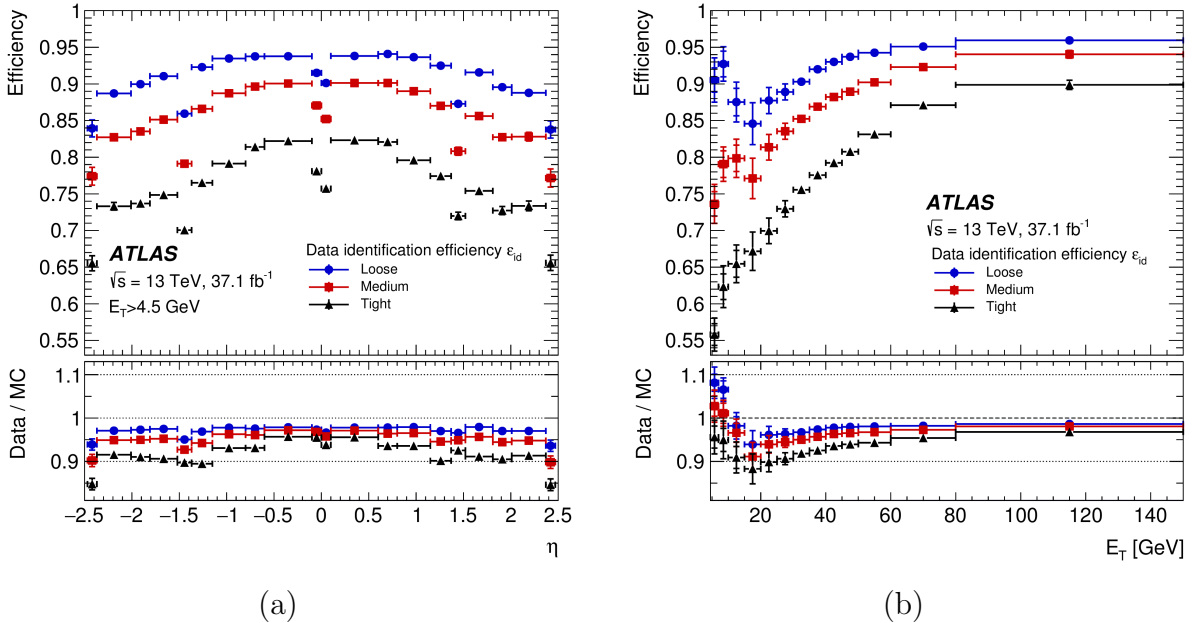


Figure 14: Likelihood-based identification efficiencies for three working points measured from data as functions of E_T and η [11].

As for electron isolation, this process helps to “clean up” the noise and pile-up around the

prompt electron candidate to achieve more precise track and energy measurements. Since the boosted electron from the primary vertex tends to create a collimated electromagnetic shower, there should be low activity in the surrounding area in the η - ϕ plane. There are two types of isolation: calorimeter-based and track-based. Both strategies estimate the amount of activity within an isolation cone around the electron candidate. In the calorimeter-based isolation, the discriminating variable is the sum of transverse energy inside the cone ($E_{T,\text{cone}}^{\text{isol}}$) where the cone size is $\Delta R = 0.2$. On the other hand, track-based isolation sums the transverse momentum of tracks ($p_{T,\text{cone}}^{\text{isol}}$). Compared to the calorimeter, the tracker can afford a smaller cone size thanks to its fine granularity. Therefore, the cone size varies based on the transverse momentum of the electron candidate (p_T):

$$\Delta R = \min \left(\frac{10 \text{ GeV}}{p_T}, \Delta R_{\text{max}} \right), \quad (21)$$

where the maximum cone size is at $\Delta R_{\text{max}} = 0.2$. For a more energetic electron with large p_T , the isolation cone is smaller. Similar to electron identification, various working points are provided and the efficiencies of a few working points are presented in Fig. 15. The ‘Fix (Tight)’ working point, defined by $E_{T,\text{cone}}^{\text{isol}}/p_T < 0.06$ and $p_{T,\text{cone}}^{\text{isol}}/p_T < 0.06$, is applied in the $H \rightarrow WW^*$ analysis.

3.3.3 Muons

Unlike electrons, muons interact with the ATLAS detector mainly through ionization which deposits much less energy. Muon detection relies on both track measurements from the inner detector (ID) and the muon spectrometer (MS). Track reconstruction in the ID is discussed in Sec. 3.3.1. Muon reconstruction in the MS first searches for hit patterns inside each muon chamber and then fits the hits from different layers [12].

Depending on which subdetectors are used in the muon reconstruction, four muon types are defined:

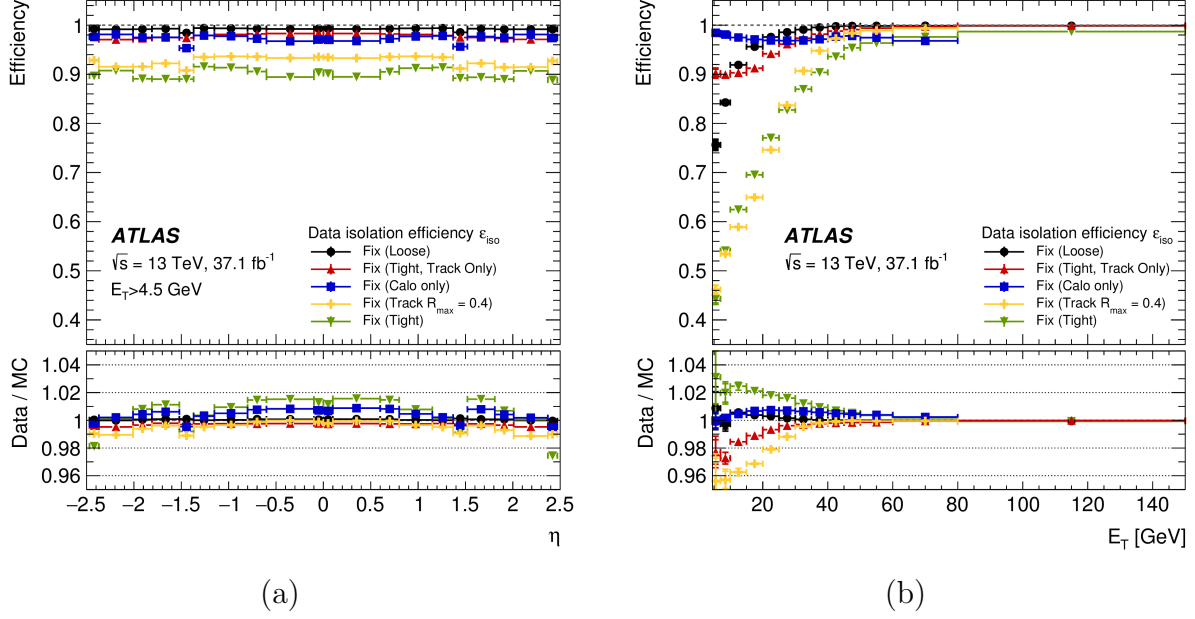


Figure 15: Isolation efficiencies for a few defined working points as functions of E_T and η [11].

- combined muon: combined track by a global refit using hits from tracks that are reconstructed independently in the ID and the MS;
- segment-tagged muon: an ID track extrapolated outward with at least one local track segments in the MS;
- calorimeter-tagged muon: an ID track matched to energy deposit in the calorimeter compatible with a minimum-ionizing particle;
- extrapolated muon: track from MS only, with loose compatibility with the interaction point.

Muon identification and isolation are performed by imposing quality requirements optimized using $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ samples. Working points are defined using simulated $t\bar{t}$ events where the top quarks predominantly decay to a bottom quark and a W boson and the muons from the W decays are considered prompt signal. While there are four muon identification working points and seven for isolation, we focus on the Medium and Tight muon identification working points and the Tight muon isolation working point used in the

$H \rightarrow WW^*$ analysis.

Three variables are used for identification quality for combined muons:

- q/p significance: absolute difference of the measured charge (q) and momentum (p) ratio from the ID and MS divided by the uncertainties summed in quadrature;
- ρ' : absolute difference of the separate transverse momentum measurement between the ID and the MS divided by the combined measurement;
- χ^2 : normalized goodness of the combined fit.

For the Medium muon identification, only combined muons and extrapolated muons are included. The Medium selection minimizes systematic uncertainties related to muon reconstruction. The requirements on the number of hits depend on the pseudorapidity region of the hits, but mostly 3 layers in the monitored drift tubes (MDT) and cathode strip chambers (CSC) are needed. Finally, the q/p significance is required to be less than seven. The above requirements are also imposed on Tight muons, but only combined muons with hits in at least two MS stations are considered. Additionally, Tight muons are required to have a normalized χ^2 smaller than eight and a two-dimensional cut in ρ' and q/p significance as a function of p_T is applied to reject background. The reconstruction efficiency for the Medium muon selection in Fig. 16 shows stable muon reconstruction efficiencies across the muon transverse momentum spectrum.

Analogous to electron isolation, muon isolation also employs track-based and calorimeter-based strategies. The track-based isolation sums ID track p_T with $p_T > 1$ GeV in an isolation cone with varied cone size defined in Eq. 21 at $\Delta R_{\text{max}} = 0.3$. The calorimeter-based isolation evaluates $E_{T,\text{cone}}^{\text{isol}}$ by summing E_T within the fix-sized cone $\Delta R = 0.2$. The Tight isolation working point is defined to have 96% efficiency constant in η and p_T .

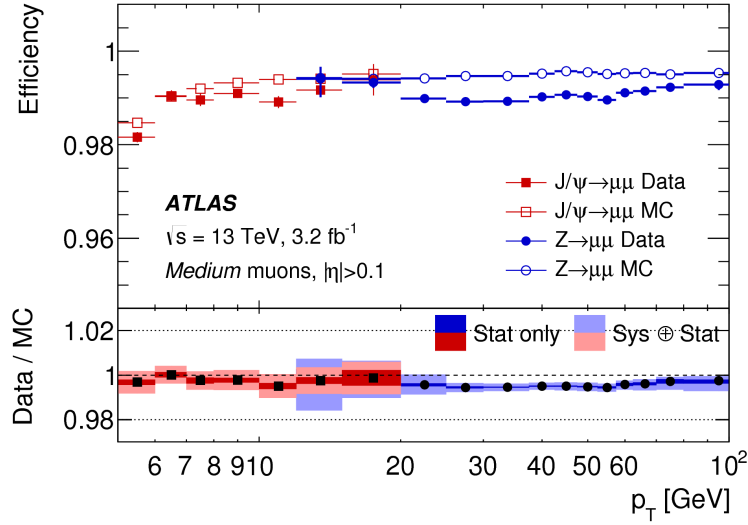


Figure 16: Reconstruction efficiency for the Medium identified muons as a function of the muon transverse momentum p_T [12].

3.3.4 Jets

Due to color confinement, a quark/gluon cannot exist on its own and hadronizes to become a color-neutral stream of particles, referred to as a *jet*. While different components of the jets interact with the entirety of the ATLAS detector, most analyses use jets reconstructed from energy deposited in the calorimeters. From these depositions, topological clusters (topo-clusters) are built from calorimeter cells (topo-clusters), and the topo-clusters are the input to the anti- k_t algorithm [46] to reconstruct jets. For analyses based on the Run-1 dataset, the ATLAS reconstruction algorithm applied a jet energy scale (JES) correction factor which incorporates the tracks associated with the jet to improve resolution [47]. The Particle Flow approach was recently introduced when analyzing Run-2 dataset [13]. In the range $|\eta| < 2.5$, the Particle Flow algorithm takes both the positive energy topo-clusters and the selected tracks matched to the primary vertex. Outside the geometrical acceptance of the tracker, only the calorimeter information is available. Including the tracker information, the particle flow algorithm significantly improves the transverse momentum resolution for low- p_T jets (Fig. 17b) and central jets with low $|\eta|$ (Fig. 17a).

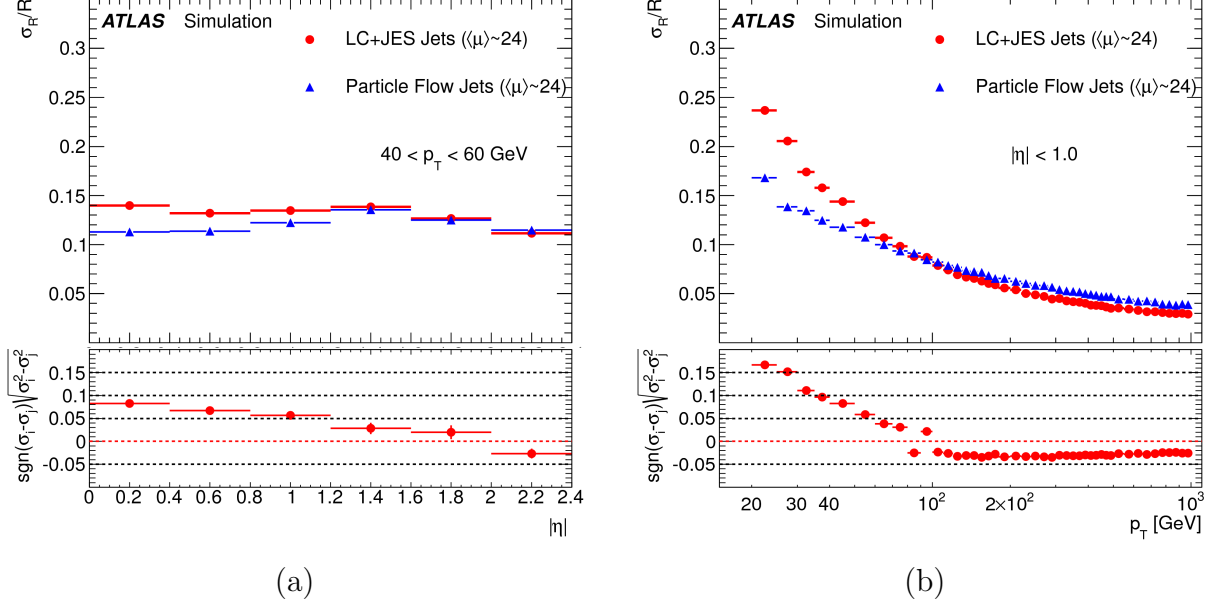


Figure 17: The jet transverse momentum resolution for calorimeter jets (LC+JES) and particle flow jets as a function of p_T (left) and η (right) [13]. The average pile-up ($\langle\mu\rangle$) is 24.

Additionally, a Jet Vertex Tagger (JVT) variable is derived to determine the probability that the jet is from a hard-scattering process or pile-up [14]. This variable is a multivariant combination of the number of tracks originated from the primary vertex (N_{Vtx}) and the jet-vertex-fraction defined as the ratio of p_T sum for jet-associated tracks from the primary vertex versus the total p_T sum of all tracks in the jet. The efficiency of rejecting pile-up jets (jet-veto) is tested on a simulated vector boson fusion (VBF) Higgs to four leptons sample. Defined as fraction of events which do not contain additional jet within the tracker acceptance other than the product from the VBF process, the jet-veto efficiency is consistently high at approximately 94% (85%) for the jet-veto at $p_T > 30$ GeV ($p_T > 20$ GeV) with a JVT > 0.7 criterion. The jet-veto efficiencies with and without the JVT selection are compared in Fig. 18 as a function of N_{Vtx} .

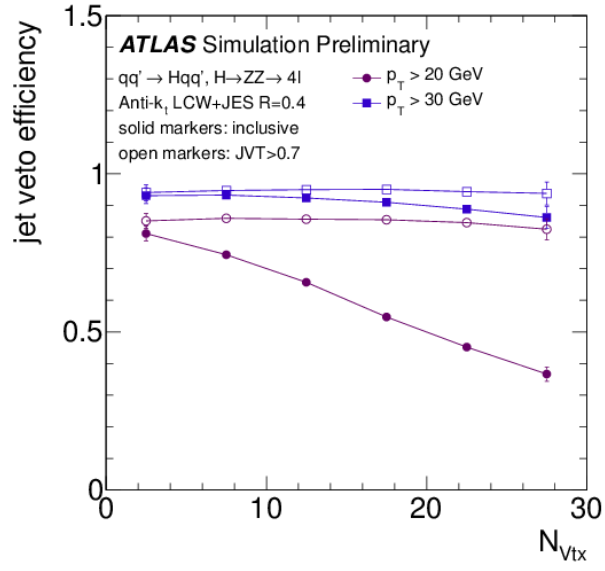


Figure 18: Jet-veto efficiency as a function of N_{Vtx} with (open markers) and without (solid markers) the $JVT > 0.7$ criterion [14].

4 Analysis overview

The vector boson fusion (VBF) production mode of the Higgs boson in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decay channel was previously studied by the CMS [48,49] and ATLAS [50,51] collaborations. These analyses used a partial Run-2 dataset corresponding to an integrated luminosity of 36 fb^{-1} and at center-of-mass energies of 7 TeV and 8 TeV during Run-1 of the LHC. However, the VBF production mode was not observed. Only recently, ATLAS made the first observation of the VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channel using the full Run-2 dataset, corresponding to 139 fb^{-1} [52]. This measurement reported an inclusive VBF $H \rightarrow WW^*$ cross section.

This thesis presents the cross-section measurement of VBF $H \rightarrow WW^*$ (Fig. 19) in a fiducial phase space with two jets, one electron, one muon and two neutrinos in the final state, where the leptons are the decay products of the W bosons. Furthermore, the cross section is measured for the first time as a function of several observables with corrections for the detector resolutions and inefficiencies. The analysis uses the full Run-2 data corresponding to an integrated luminosity of 139 fb^{-1} collected by the ATLAS experiment in 2015-2018 at the center-of-mass energy $\sqrt{s} = 13 \text{ TeV}$.

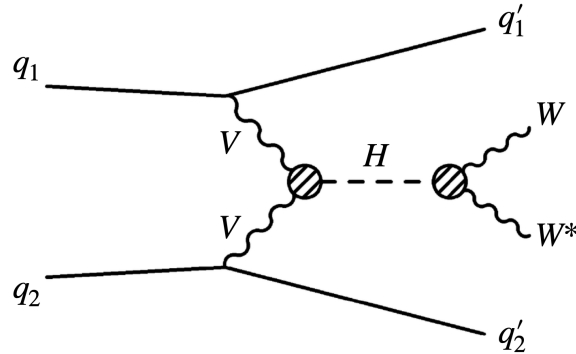


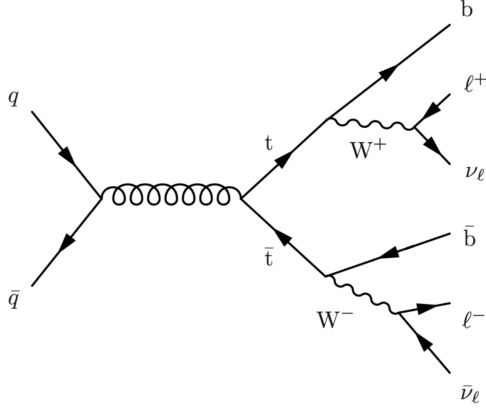
Figure 19: Feynman diagram for the signal process: Higgs production via vector boson fusion and decay to two W bosons. The “ V ” stands for vector boson. One of the W bosons originating from the Higgs decay is virtual (off the mass shell), and is labeled as W^* .

Several SM processes produce the same final state objects and are considered as background in this analysis. First, there are processes which also involve a pair of W bosons: SM

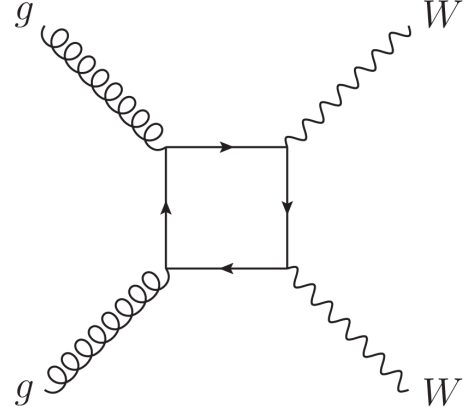
continuum WW process and $t\bar{t}$, which primarily decay to $b\bar{b}W^-W^+$. Next, the other Higgs production modes, such as gluon-gluon fusion (ggF) and associated Higgs production with a vector boson (VH), can also produce events that are signal-like. Additionally, events from the W +jets and multi-jet processes where the jets are misidentified as leptons are significant backgrounds. Finally, the Drell-Yan process, also referred to as Z +jets, with two tau leptons from the Z boson decay can also produce a pair of opposite-flavor leptons and large missing transverse energy which can look like signal. The Feynman diagrams of the dominant background processes are shown in Fig. 20. They add up to 95% of the total background in the signal-enriched region.

A signal-enriched region is defined by kinematic selections optimized by Monte Carlo (MC) simulations to improve the signal-over-background ratio. In this region, the signal yield is extracted from a binned likelihood fit to a 2-dimensional output distributions of a multivariant analysis, while the background contributions are simultaneously estimated using data-to-simulation normalization factors. As discussed in Sec. 3.3, the detector technologies have intrinsic resolutions and inefficiencies. In attempt to reverse the detector distortion, we use *unfolding response matrices* to transform the *reconstruction-level* measurements in the signal region to *particle-level* in the fiducial region. In general, a particle-level or truth-level measurement is considered detector-independent, making it comparable across different experiments and with different theoretical models. In particular, the unfolded distributions are compared with various state-of-the-art theory predictions at different orders of calculations in QCD and EW, and re-interpreted with the Effective Field Theory (EFT) formalism.

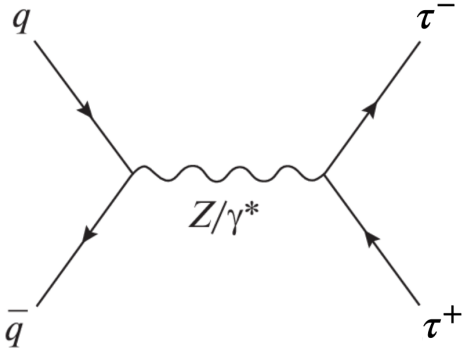
The sections dedicated to analysis strategy and results are organized as the following. First, we describe the data and MC samples in Sec. 5. Based on the MC simulations, the signal optimization strategies in Sec. 6 and the background estimation strategies in Sec. 7 will be discussed in detail. We will provide several theoretical studies in Sec. 8 to outline the fiducial phase space definition and the particle-level predictions to be used for constructing the unfolding response matrices and for comparisons with the unfolded results. In Sec. 9, we



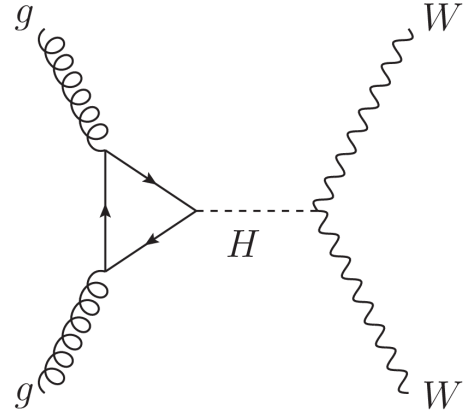
(a) $t\bar{t}$ where both top quarks decay to a bottom quark and a W boson.



(b) SM WW QCD processes.



(c) Drell-Yan process with Z decayed to two tau leptons.



(d) Gluon-gluon fusion and Higgs to two W bosons.

Figure 20: The representative Feynman diagrams of the dominant background processes: (a) $t\bar{t}$ process, (b) SM WW QCD process, (c) the Drell-Yan background producing two tau leptons, (d) ggF $H \rightarrow WW^*$ process plus two jets.

will elaborate the statistical analysis used to extract the cross sections with the evaluation of various systematic uncertainties and the unfolding strategy. Finally, we will report the results using the ATLAS Run-2 dataset in Sec. 10 and the EFT interpretation strategy and results in Sec. 11.

5 Data and Monte Carlo samples

5.1 Data samples

This analysis uses the full ATLAS Run-2 dataset¹, consisting of all proton-proton collision data collected from 2015-2018 at $\sqrt{s} = 13$ TeV. The integrated luminosity is 139 fb^{-1} with an uncertainty of 1.7%, as measured by the LUCID-2 detector [53].

The data are subjected to quality requirements. Events recorded during periods when the relevant detector components were not operating normally are rejected, and only events from the standard *All Good Good Run List* are used in the analysis [54].

5.2 Monte Carlo samples

Monte Carlo (MC) event generators simulate the signal and background events produced in the proton-proton collisions at the center-of-mass energy of $\sqrt{s} = 13$ TeV. In general, the MC events are produced at parton-level and are successively simulated for parton shower and hadronization to particle-level. Next, all MC samples are processed through GEANT-4 [55] to simulate particle interactions with the ATLAS detector. Finally, similar to data, MC events are processed by the standard ATLAS reconstruction programs to reach the reconstruction-level. The reconstruction-level MC samples are used to optimize the signal event selection and the multivariant analysis and to evaluate various systematic uncertainties.

5.2.1 Signal samples

The nominal signal VBF events are generated with POWHEG [56], accurate to next-to-leading order (NLO) in QCD and EW corrections. It is interfaced with PYTHIA 8 [57] to model the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decay, parton shower, hadronization and *underlying events* (activities of the rest of the partons) with the AZNLO tune [58]. The EVTGEN program [59] was used for the properties of the bottom and charm hadron decays.

¹This excludes 0.13 fb^{-1} of data collected in 2015 using the 50 ns bunch spacing configuration, as well as the low pile-up runs.

An alternative sample uses the same POWHEG VBF sample but the Higgs decay and parton shower are simulated by HERWIG 7. This sample is compared to the nominal sample for the evaluation of uncertainties arising from parton shower and underlying event modeling. Additionally, the difference between this sample and a MADGRAPH5_AMC@NLO+HERWIG 7 sample provides the matching uncertainty. All samples use NNPDF30_nlo_as_0118 [60] as the nominal parton distribution function (PDF).

All the inclusive signal MC samples are normalized to a cross section calculated up to the full NLO order in QCD and EW [61, 62] with an additional approximate next-to-next-to-leading (NNLO) QCD correction [63]. The VBF inclusive cross section is [4]

$$\sigma_{\text{VBF}H}^{\text{truth}} = 3781 \pm 79 \text{ fb}, \quad (22)$$

where the uncertainty is the total theory uncertainty. Multiplying the cross section with the branching ratio of the Higgs boson decaying to two W bosons, which is 0.214 ± 0.003 [4], we obtain the VBF $H \rightarrow WW^*$ cross section

$$\sigma_{\text{VBF}H \rightarrow WW^*}^{\text{truth}} = 809 \pm 20 \text{ fb}. \quad (23)$$

The cross section and branching ratios are simulated at a Higgs mass of 125.00 GeV. Including the W boson leptonic decay branching ratio of 0.109 ± 0.001 [64], the expected VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ cross section is derived:

$$\sigma_{\text{VBF}H \rightarrow WW^* \rightarrow \ell\nu\ell\nu}^{\text{truth}} = 86.5 \pm 2.4 \text{ fb}. \quad (24)$$

5.2.2 Background samples

The nominal samples of the most dominant background processes are used in designing the kinematic cuts and the multivariate analysis. First, the SM WW sample is simulated using SHERPA2.2.2 [65] with 0 and 1 jet at NLO accuracy, and 2 and 3 jets at LO accuracy.

Electroweak diboson production in association with two jets ($VVjj$) is generated using MADGRAPH5_AMC@NLO with LO matrix elements and interfaced with PYTHIA 8 for parton shower. The Drell-Yan sample is simulated with SHERPA2.2.1 with a dedicated parton shower tuning [66]. For the top background, the top-quark pair production sample is generated with POWHEG and interfaced with PYTHIA 8. The inclusive cross section of the sample is normalized to the value calculated at NNLO [67]. The single top background, mainly Wt , is simulated by POWHEG and PYTHIA 8.

The $Z\gamma$ and $W\gamma$ contributions in the analysis are estimated from MC samples modeled by SHERPA2.2.8 at the NLO accuracy for 0 and 1 jets. The overlap between these $V\gamma$ samples and Drell-Yan events is removed.

The MC samples of W +jets and Z +jets used to validate and estimate jets misidentified as leptons are simulated by POWHEGMinLO and showered by PYTHIA 8.

Finally, the other Higgs production modes with the Higgs decaying in the same $H \rightarrow WW^*$ channel are also considered. The analysis includes the ggF and associated Higgs boson production VH ($V = W, Z$), whereas the Higgs production modes in association with a heavy quark pair ($t\bar{t}H$, $b\bar{b}H$) are found to have a negligible contribution and are not included in the analysis. The nominal ggF sample is simulated at NNLO accuracy in QCD using the POWHEG NNLOPS program [68–72], interfaced with PYTHIA 8.212 for parton shower. On the other hand, the VH sample is simulated at NLO accuracy with one additional jet using POWHEG+PYTHIA 8.212.

6 Signal optimization

The event selection of this analysis is optimized based on the nominal MC simulation of the signal and backgrounds. The same selections are applied to the data. The VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu + 2j$ event candidates are required to have exactly one electron and one muon of opposite charge, in addition to at least two jets with specific kinematic configurations that enhance the VBF signal. In this section, we first define the final-state physics objects in Sec. 6.1 and the common set of observables used in this analysis in Sec. 6.2. Next, we discuss event pre-selection in Sec. 6.3 prior to introducing the signal region selection in Sec. 6.4. Finally, we discuss the multivariate analysis, the boosted decision tree (BDT) technique, used to further optimize analysis sensitivity in Sec. 6.5.

6.1 Physics objects

The physics objects reconstructed using techniques discussed in Sec. 3.3 are further selected and calibrated to serve the purposes of this analysis. Subsequently, the kinematics of these objects enter the event selection, the BDT training and evaluation, and the calculation of observables that this analysis intends to measure.

Leptons

Lepton identification choices impact the analysis primarily through the rejection of the background from jets mis-identified as leptons, with a trade-off for signal efficiency. Leptons are required to originate from a primary vertex with no less than two associated tracks, each with a transverse momentum $p_T > 500$ MeV. The absolute value of the longitudinal impact parameter of each lepton track, $|z_0 \sin \theta_0|$, is required to be less than 0.5 mm for both electrons and muons. Moreover, the significance of the transverse impact parameter calculated with respect to the beamline ($|d_0|/\sigma_{d_0}$) is required to be less than 5 (3) for electrons (muons).

The electrons are selected from clusters of energy deposits in the calorimeter that match

a track reconstructed in the inner detector. The chosen identification working points are the combination of *Tight* for electrons with $p_T < 25$ GeV and *Medium* for electrons with $p_T > 25$ GeV, where the working point definitions are given in Sec. 3.3.2. Additionally, electrons are required to have a pseudorapidity $|\eta| < 2.47$, excluding the transition region between the barrel and the endcaps in the liquid argon calorimeter in region $1.37 < |\eta| < 1.52$. To reduce the number of misidentified leptons, the *Fix (Tight)* isolation working point is applied. Finally, a requirement (`AUTHOR = 1`) is added to select electrons that are reconstructed exclusively as electrons. All of the electron requirements are summarized in Table 1.

p_T range	Electron ID	η range	Author	Isolation	Impact parameter
< 25 GeV > 25 GeV	Tight Medium	$(\eta < 1.37) \parallel (1.52 < \eta < 2.47)$	1	Fix (Tight)	$ z_0 \sin \theta < 0.5$ mm, $ d_0 /\sigma_{d_0} < 5$

Table 1: Electron selection requirements.

To optimize the muon selection for the signal, both the muon identification and isolation working points are chosen to be *Tight*. The definition of the working points is detailed in Sec. 3.3.3. The muon object requirements are summarized in Table 2.

p_T range	Muon ID	η range	Isolation	Impact parameter
> 15 GeV	Tight	$ \eta < 2.5$	Tight	$ z_0 \sin \theta < 0.5$ mm, $ d_0 /\sigma_{d_0} < 3$

Table 2: Muon selection requirements.

Jets

Jets are reconstructed using the anti- k_t algorithm [46] with a radius parameter $R = 0.4$ implemented in the FastJet package [73]. Jets with $p_T > 30$ GeV are selected. The sub-threshold jets ($20 \text{ GeV} < p_T < 30 \text{ GeV}$) are kept and are used later to reject events with central soft jets (discussed in Sec. 6.4). The pseudorapidity range of jets extends to $|\eta| < 4.5$. In addition, the analysis applies a Jet Vertex Tagger (JVT) requirement defined in Sec. 3.3.4. Jets with $p_T < 60$ GeV are required to have JVT greater than 0.59 to suppress pile-up.

A b -tagged jet contains b -hadrons and is identified via the the DL1r b -tagging algorithm [74]. The b -tagged jets with $p_T \geq 20$ GeV and $|\eta| < 2.5$ are referred to as b -jets in the analysis.

Missing transverse energy

This analysis denotes the magnitude of the missing transverse energy (MET) as E_T^{miss} . By definition, MET is the negative vectorial sum of the transverse momenta of the *hard* objects, and the *soft term* [75].

$$\vec{E}_T^{miss} = - \sum_{\text{hard},i} \vec{p}_{T,i} - \sum_{\text{soft},j} \vec{p}_{T,j}. \quad (25)$$

The former are the reconstructed and calibrated jets and leptons. The later are reconstructed tracks that are not associated with these objects but originate from the primary vertex.

Overlap removal

The above reconstructed objects are identified independently. When two objects of the same type or different types have significant overlap, one of the objects is removed to reduce backgrounds. This process is called the *overlap removal*. We use the official Analysis Software Group tool for such removal, which is done in the following steps:

- electron-electron: if there is an overlapping second layer cluster or a shared track, the lower- p_T electron is removed.
- electron-muon: if a combined muon shares an inner detector track with an electron, the electron is removed; if a calorimeter-tagged muon shares an inner detector track with an electron, the muon is removed.
- electron-jet: the jet is removed if $\Delta R(\text{jet}, \text{electron}) < 0.2$; for the rest of the jets, the electron is removed if $\Delta R(\text{jet}, \text{electron}) < \min(0.4, 0.04 + 10 \text{ GeV}/p_T^e)$.
- muon-jet: when $\Delta R(\text{jet}, \text{muon}) < 0.2$, the jet is removed if the jet has less than three

associated tracks with $p_T > 500$ MeV or both of the following conditions are met - the p_T ratio of the muon and jet is larger than 0.5 ($p_T^\mu/p_T^{jet} > 0.5$) and the ratio of the muon p_T to the sum of p_T of tracks with $p_T > 500$ MeV associated to the jet is larger than 0.7. For any surviving jets, if $\Delta R(\text{jet}, \text{muon}) < \min(0.4, 0.04 + 10 \text{ GeV}/p_T^\mu)$, the muon is removed.

However, a b -tagged jet is always saved from overlap removal.

6.2 Observables

This section defines observables frequently used throughout this thesis. The observables are measured or derived from the physics objects defined in the previous section. The two leptons and two jets are generally ordered by their transverse momentum referred as leading/subleading lepton or jet.

This analysis measures the following 13 observables:

- m_{jj} : the invariant mass of the two leading jets;
- Δy_{jj} : the rapidity difference between the two leading jets;
- signed $\Delta\phi_{jj}$: the difference in azimuthal angle between the forward jet with higher rapidity and the central jet with lower rapidity:

$$\Delta\phi_{jj} = \phi^{\text{forward}} - \phi^{\text{central}}, \quad (26)$$

where the observable is also denoted by $|\Delta\phi_{jj}|$ in this thesis for simplicity and the absolute value is labeled as $|\Delta\phi_{jj}|$;

- p_T^{j1} : leading jet transverse momentum;
- p_T^{j2} : subleading jet transverse momentum;
- m_{ll} : the invariant mass of the two charged leptons;

- $\Delta\phi_{ll}$: the azimuthal angle between the two final state leptons;
- Δy_{ll} : the rapidity difference between the two leptons;
- $p_{T,ll}$: the transverse momentum of the two-lepton system;
- $\cos(\theta_\eta^*)$: calculated from the pseudorapidity difference between the leptons ($\Delta\eta_{ll}$):

$$\cos(\theta_\eta^*) = \tanh\left(\frac{\Delta\eta_{ll}}{2}\right); \quad (27)$$

- $p_T^{\ell_1}$: leading lepton transverse momentum;
- $p_T^{\ell_2}$: subleading lepton transverse momentum;
- p_T^H : the transverse momentum of the Higgs boson calculated as the magnitude of the vectorial p_T sum of charged leptons and MET from the W boson decays.

Furthermore, we define several additional observables used in the event selections and the BDT training. First, there are several observables dedicated to background suppression. The invariant mass of the tau leptons in $Z \rightarrow \tau\tau$ events ($m_{\tau\tau}$) is introduced to reduce the Drell-Yan background. Based on the Collinear Approximation Method [76], neutrinos from the tau decays can be assumed to be the only source of the observed MET in the event and their energy fractions can be computed. The $m_{\tau\tau}$ variable is calculated based on the momentum of the leading lepton (p^{l1}), subleading lepton (p^{l2}), and MET (magnitude denoted as E_T^{miss}):

$$m_{\tau\tau} = \frac{m_{ll}}{\sqrt{x_1 x_2}}, \quad (28)$$

$$x_1 = \frac{p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}}{p_y^{l2} E_{T,x}^{miss} - p_x^{l2} E_{T,y}^{miss} + p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}}, \quad (29)$$

$$x_2 = \frac{p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}}{p_x^{l1} E_{T,y}^{miss} - p_y^{l1} E_{T,x}^{miss} + p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}}, \quad (30)$$

where the indices, x and y , represent the x and y components of the momenta, respectively.

Rejecting events with b -jets significantly suppresses top-quark induced background. The number of jets with $p_T > 20$ GeV identified as b -jets from the b -tagged algorithm (DL1r) at 85% efficiency is defined as N_{b-jet} .

One of the most important variables for the BDT training and the discriminating variable for the Drell-Yan background estimation is the transverse mass (m_T), defined as

$$m_T = \sqrt{(E_{T,l} + E_T^{miss})^2 - |\vec{p}_{T,l} + \vec{E}_T^{miss}|^2} \quad (31)$$

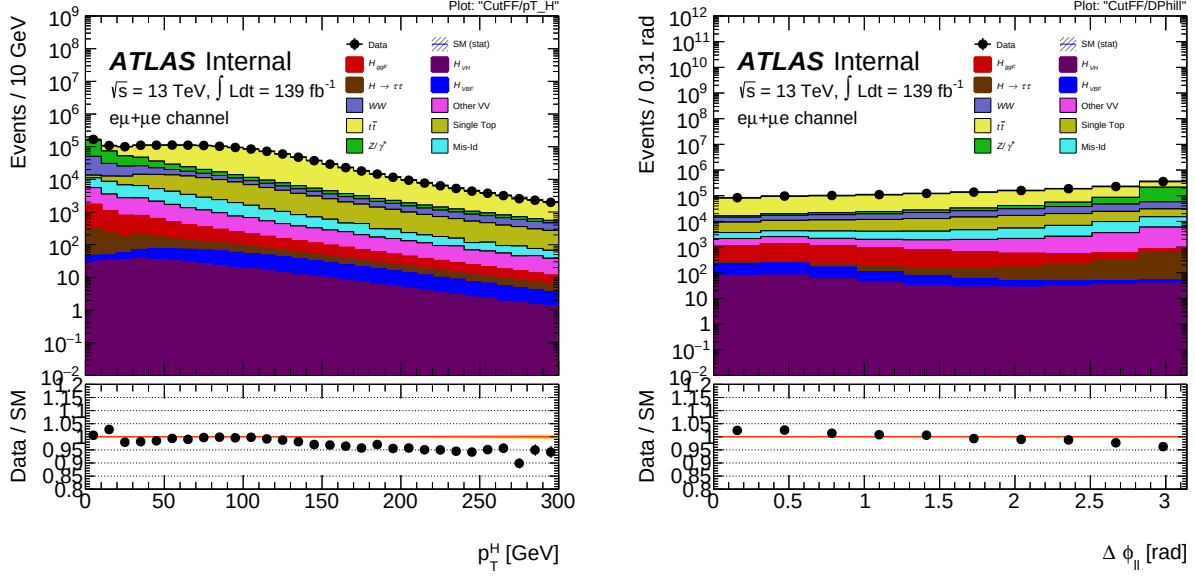
where $E_{T,l} = \sqrt{|\vec{p}_{T,l}|^2 + m_l^2}$ and $\vec{p}_{T,l}$ is the vector sum of the lepton transverse momentum. This variable peaks around the Z boson mass for the Drell-Yan process. On the other hand, the m_T spectrum for the VBF $H \rightarrow WW^*$ signal has a cut-off at the Higgs mass.

6.3 Pre-selection

Prior to the final event selections aiming to maximize sensitivity to the respective signal or background process, a number of common kinematic cuts known as pre-selections are applied. Pre-selection cuts are summarized in Table 3. In the channel selection, events are required to have exactly two opposite-flavor and opposite-sign charged leptons (e and μ) to reduce the background from diboson processes. Events with any additional lepton are vetoed. Fig. 21 shows some examples of kinematic distributions in the pre-selection region. The signal-over-background ratio (s/b) at pre-selection level is approximately 0.0004.

Pre-selection cut	Description
Channel selection	$e\mu/\mu e$ channel and event weights applied
Trigger selection and matching	trigger selected and trigger weight applied
Only two Leptons	Require exactly two identified leptons (except for fakes)
Lead lepton p_T	$p_T^{\ell_1} > 22$ GeV
Subleading lepton p_T	$p_T^{\ell_2} > 15$ GeV
Opposite sign leptons	Require zero dilepton charge
m_{ll} selection cut	$m_{ll} > 10$ GeV

Table 3: Pre-selection cuts applied before event selections for different signal and background-enriched regions. The lepton identification criteria is different for the fakes. The fake factor implementation will be detailed in Sec. 7.1.



(a) Higgs p_T .

(b) $\Delta\phi_u$.

Figure 21: The Higgs p_T and $\Delta\phi_u$ distributions in the pre-selection region from the ATLAS Run-2 data (black data points) and MC simulation (filled histogram). The uncertainty band shows only MC statistical uncertainties, and the error bars on the data points represent data statistical uncertainties. The bottom panel shows the data-to-MC ratio.

6.4 Signal region selection

There are three macro regions defined by different kinematic selections: the VBF signal region (SR), the Drell-Yan region (CRZjets) and the ggF region (ggF CR). We will discuss the definition of the control regions dedicated to estimate the respective background process in Sec. 7.

The SR cuts are listed and described in Table 4. Selected events are required to have at least 2 jets with $p_T > 30$ GeV. To reject the top-quark induced background, events with any b -jet are vetoed. The $Z \rightarrow \tau\tau$ veto suppresses Drell-Yan background. Requiring $m_{jj} > 450$ GeV and $\Delta y_{jj} > 2.1$ further reduces the top background while maintaining the signal efficiency, since jets from the VBF process tend to be more energetic and more separated in rapidity. The $\Delta\phi_u < 1.4$ requirement reduces the SM WW background events in which the leptons are less collimated than the signal. Due to a pronounced color-flow

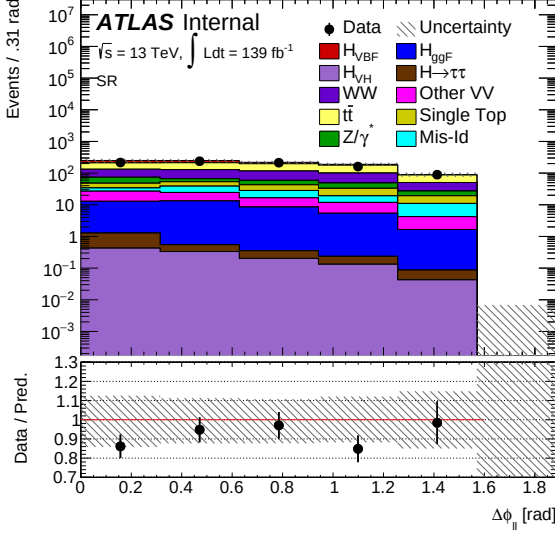
suppression in the VBF events, a central jet veto (CJV) is imposed to select events with low hadronic activities between the rapidity space of the two leading jets. Finally, the outside lepton veto (OLV) is introduced as the Higgs decay products should be between the two leading jets. The two VBF topology cuts, CJV and OLV, successfully reject the majority of VH and ggF background to a few percent level. In Fig. 22, we show the comparison of the ATLAS Run-2 data and MC simulation at reconstruction-level selected in the SR for several observables. The MC simulation is generally compatible with the data within the data statistical uncertainties.

Signal region cut	Description
2-jet (30,30)	Require at least 2 tagged jets ($p_T \geq 30$ GeV)
b-veto	$N_{b-jet} = 0$
CJV (20 GeV)	Reject events with a third central rapidity jet $p_T > 20$ GeV
OLV bool	Lepton η required to be between the two leading jets
$Z \rightarrow \tau\tau$ veto	$m_{\tau\tau} < m_Z - 25$ GeV
Δy_{jj} cut	$\Delta y_{jj} > 2.1$
m_{jj} cut	$m_{jj} > 450$ GeV
$\Delta\phi_{ll}$ cut	$\Delta\phi_{ll} < 1.4$

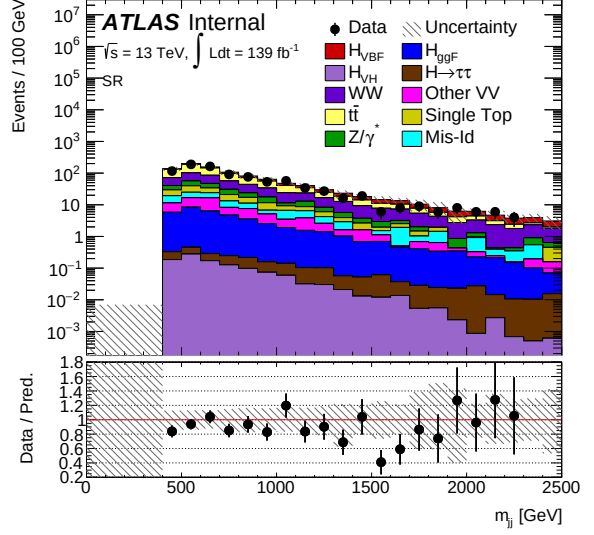
Table 4: VBF signal region selections.

6.5 Signal optimization with boosted decision tree

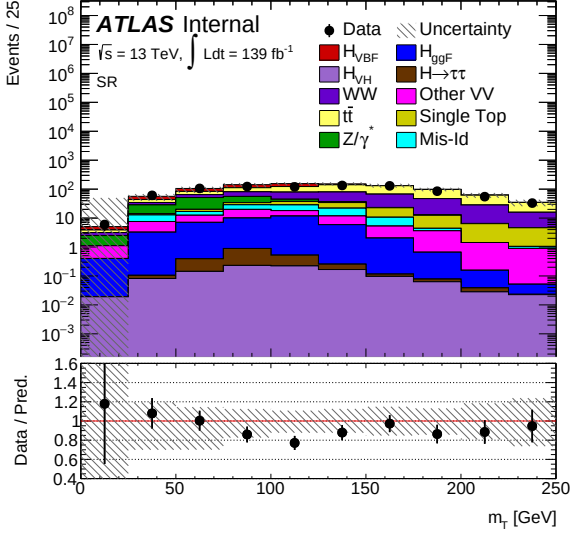
Even after the SR selections, s/b is only 0.13. To further maximize sensitivity, we adopt the boosted decision tree (BDT) algorithm which is a multivariant machine learning technique implemented in the TMVA package [77]. In general, a decision tree is a collection of cuts based on the input variables, designed to classify events as signal-like or background-like. A given signal event is correctly identified if it is placed in a signal-dominated “leaf” and vice-versa for background events. After the initial tree is built another tree is grown to better separate the signal and background events mis-identified by the first tree. This proceeds iteratively until there is a collection of a specified number of trees, a process known as boosting. A weighted average is taken from all the trees to form a BDT output discriminant with values ranging from -1 to 1. Signal-like events are scored higher by the BDT and



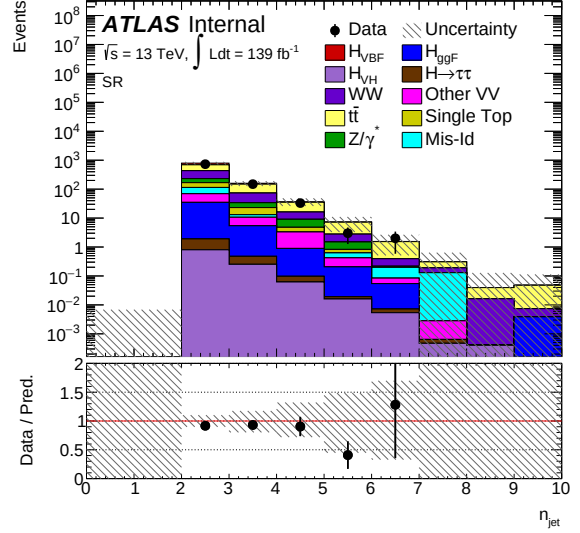
(a) $\Delta\phi_{ll}$.



(b) m_{jj} .



(c) m_T .



(d) Jet multiplicity.

Figure 22: The $\Delta\phi_{ll}$, m_{jj} , m_T and jet multiplicity distributions in the VBF signal region from the ATLAS Run-2 data (black data points) and MC simulations (filled histogram). The uncertainty band includes MC statistical, experimental and theory uncertainties, and the error bars on the data points represent data statistical uncertainties.

background-like events are scored lower.

In contrast to the inclusive VBF $H \rightarrow WW^*$ cross-section measurement [78] which uses a Deep Neural Network technique, our usage of machine learning is less aggressive in signal vs. background separation because we need to minimize any bias in the differential measure-

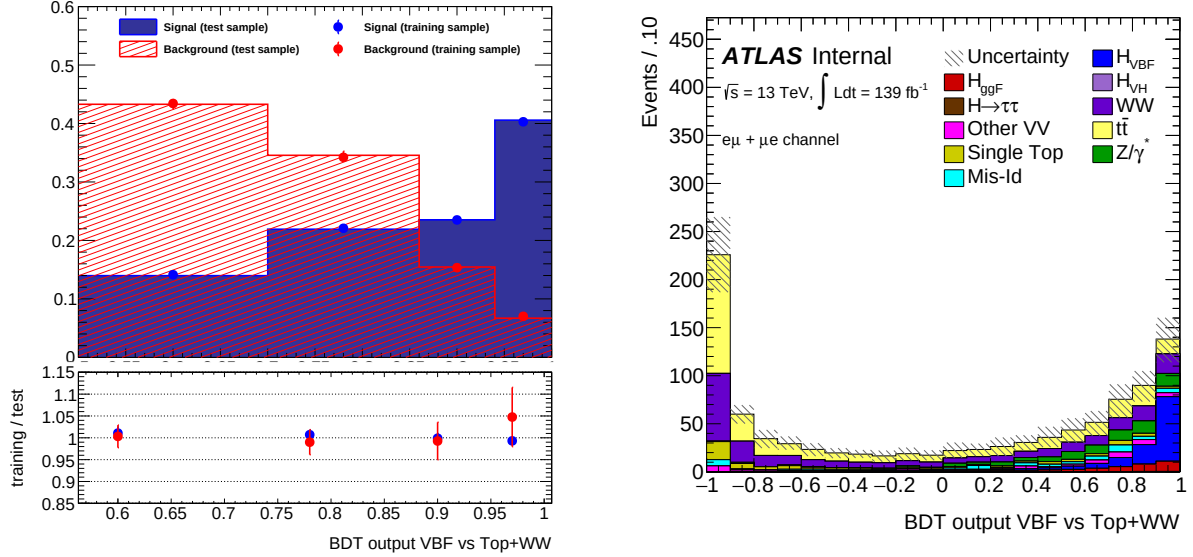
ments for the new physics interpretation. Meanwhile, we choose the minimal set of training variables to provide both high discrimination and good modeling.

6.5.1 Signal optimization for fiducial cross-section measurement

For the fiducial cross section measurement, two BDTs are trained in the SR. First, a BDT discriminating VBF signal against the top and diboson background (D_{VBF}) is trained with 12 variables from a partial set of MC samples of the signal and background nominal predictions. The variables are listed in Table. 5. The variable $\sum C_l$ is the sum of lepton η -centrality which quantifies the positions of the leptons relative to the leading jets in pseudorapidity [79]. None of the training variables has significantly higher importance comparing to the others which is desired for bias minimization. The good data-MC agreement shown in Fig. 22 also ensures good modeling of the BDTs. We do not see evidence of over-training by comparing the BDT output distribution of the normalized signal and background for the training sample and a testing sample in Fig. 23a between $0.5 < D_{\text{VBF}} < 1.0$. The full D_{VBF} distribution projected with fine binning is shown in Fig. 23b. The signal events (blue) are concentrated in the high BDT regime, whereas the top and diboson (yellow and dark purple) events score lower in the BDT.

Variable	Importance
m_T	0.14
$\Delta\phi_{ll}$	0.12
Δy_{jj}	0.10
m_{jj}	0.09
m_{ll}	0.09
p_T^{j1}	0.08
$\sum C_l$	0.08
Δy_{ll}	0.07
η_{j1}	0.07
η_{j2}	0.07
p_T^{j2}	0.06
$ \Delta\phi_{jj} $	0.05

Table 5: Ranking of the variables used as input to the training of the BDT signal vs. top and diboson.



(a) VBF BDT output of the training and test- (b) VBF BDT output of all samples in the ing samples. signal region.

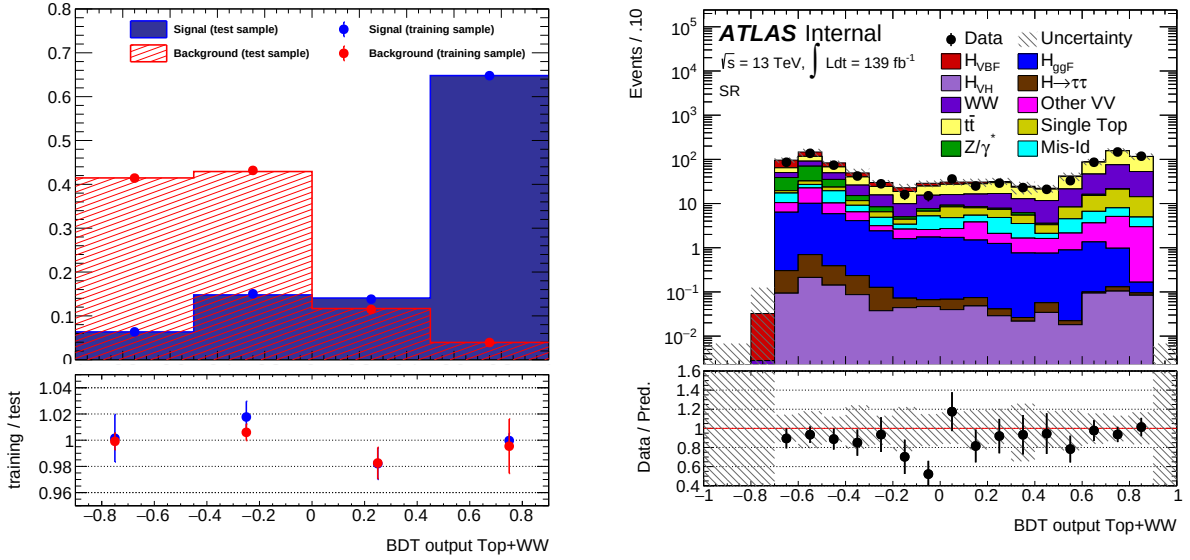
Figure 23: The VBF BDT output. (a) BDT output distribution of normalized VBF (signal) and top and diboson processes (background) for the testing and training samples. The binning corresponds to the one used in the analysis. (b) VBF BDT output distribution of the signal and all backgrounds MC samples in a finer binning. The error band includes the statistical, experimental and theory uncertainties.

The second BDT in the SR is trained to separate top plus diboson from every other processes ($D_{\text{top+VV}}$). The input variables are given in Table 6 and the BDT outputs are shown in Fig. 24. The variable M_{l1j1} is the invariant mass of the leading jet and lepton.

Since the signal is concentrated in the high VBF BDT regime, SR is further divided into two sub-regions: $D_{\text{VBF}} > 0.5$ (SR1) and $D_{\text{VBF}} \leq 0.5$ (SR2). As a combined discriminant used for the signal extraction fit, D_{VBF} is projected in SR1 to the binning $[0.5, 0.7, 0.86, 0.94, 1.0]$, whereas in SR2, $D_{\text{top+VV}}$ is projected to four equal bins between $D_{\text{top+VV}} = -1$ and $D_{\text{top+VV}} = 1$. This is what we call the 2-dimensional $D_{\text{VBF}} - D_{\text{top+VV}}$ discriminant. From MC simulation, the bin with the highest BDT score in SR1 reaches $s/b = 1.30$, ten times larger than the inclusive SR. On the other hand, SR2 is more dominated by top and diboson background, and the purity of top and diboson in the last $D_{\text{top+VV}}$ bin is 96%. The pre-fit 2-dimensional discriminant from simulation (filled histogram) and data (black data point) is

Variable	Importance
m_T	0.64
Δy_{jj}	0.11
$\Delta \phi_{ll}$	0.08
$ \Delta \phi_{jj} $	0.05
$\sum C_i$	0.04
M_{l1j1}	0.04
η_{j2}	0.03
η_{j1}	0.01

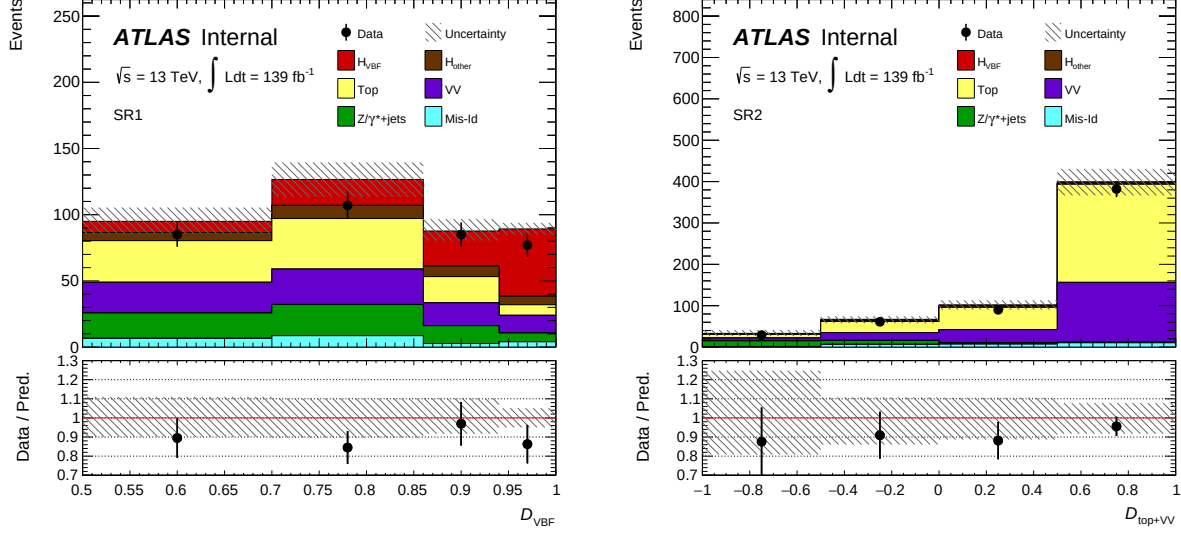
Table 6: Ranking of the variables used as input to the training of the BDT of top and diboson vs. other processes.



(a) Top and diboson BDT output of the train- and testing samples. (b) Top and diboson BDT output of all samples in the signal region.

Figure 24: The top and diboson BDT output. (a) BDT output distribution of normalized top and diboson (signal) and other processes (background) for the testing and training samples. The binning corresponds to the one used in the analysis. (b) The top and diboson BDT output distribution of the signal and all backgrounds MC samples in a finer binning. The error band includes the statistical, experimental and theory uncertainties. The bottom panel shows the ratio of the data to the expected distributions.

shown in Fig. 25.



(a) VBF BDT in SR1.

(b) Top and diboson BDT in SR2.

Figure 25: The pre-fit MC distribution (filled histogram) of the 2-dimensional $D_{\text{VBF}} - D_{\text{top+VV}}$ discriminant (a) in SR1 onto the D_{VBF} axis and (b) in SR2 onto the $D_{\text{top+VV}}$ axis, compared with the Run-2 dataset (black data points). The error bars indicate the data statistical uncertainties, while the uncertainties in the MC prediction are shown by the hatched band, which includes experimental and theoretical uncertainties. The bottom panel shows the ratio of the data to the expected distributions.

6.5.2 Signal optimization for differential cross-section measurement

For the differential cross-section measurements, the inclusive SR is segmented according to the observable binning. In each bin, a similar 2-dimensional discriminant is used, except that the D_{VBF} is re-trained in the SR for each differential measurements with a reduced set of input variables, excluding the observable to be measured and its highly correlated variables ($> 40\%$) to minimize bias. The variables removed from the BDT input list for each observable to be measured are summarized in Table. 7. The other observables not shown in the table use the same D_{VBF} as the inclusive fiducial measurement.

Observable to be measured	Variables removed from BDT input
$\Delta y_{jj}, m_{jj}$	$\Delta y_{jj}, m_{jj}$
$\Delta\phi_u$	$\Delta\phi_u, m_u$
$\Delta\phi_{jj}$	$ \Delta\phi_{jj} $
p_T^{j1}, p_T^{j2}	p_T^{j1}, p_T^{j2}
m_u	$m_u, \Delta y_u, \Delta\phi_u$
Δy_u	$\Delta y_u, m_u$

Table 7: Variables removed from the BDT input list for several observables to be measured.

7 Background estimation

The dominant background processes in the signal region are categorized into mis-identified jet (fakes), Drell-Yan, ggF plus two jets, top, and diboson. The fakes are evaluated with a data-driven fake factor method to be discussed in Sec. 7.1, whereas the other dominant background processes are estimated by the respective data-over-simulation normalization factors. The top-quark pair ($t\bar{t}$) and single-top-quark (Wt) productions are collectively referred to as the top background, and the non-resonant diboson background includes the pair productions of vector bosons such as WW , WZ , ZZ , and $W\gamma$. The similar kinematic features of the top and diboson backgrounds make it hard to estimate their contributions separately. Therefore, the analysis estimates the sum of the two backgrounds and the strategy is explained in Sec. 7.2. The estimation of the Drell-Yan background, including both the QCD and EW processes, and the ggF background are driven by the Drell-Yan control region (CRZjets) defined in Sec. 7.3 and the ggF control region (ggF CR) discussed in Sec. 7.4.

The normalization of other miscellaneous processes, which are expected to be much less significant in the signal and control regions, are fixed to their nominal MC predictions. Table 8 summarizes the composition of different processes in the three macro regions obtained from MC simulations. The VH, $V\gamma$ and Higgs boson decay to two tau leptons ($H\tau\tau$) are fixed to the MC simulation in all the regions.

	Signal and background to be estimated from data (composition %)	Signal and background estimated from MC (composition %)
CRZjets	Z+jets (72%)	top+VV (19%), $V\gamma$ (3%), VBF (3%), $H\tau\tau$ (2%),ggF (1%), VH (< 1%)
ggF CR	top+VV (84%), Z+jets (8%) ggF (2%)	$V\gamma$ (1%), VBF (< 1%), VH (< 1%) $H\tau\tau$ (< 1%)
SR	top+VV (72%), Z+jets (8%) VBF (11%), ggF (2%)	$V\gamma$ (1%), $H\tau\tau$ (< 1%), VH (< 1%)

Table 8: Composition of each signal and background process in the CRZjets, the ggF CR, and the SR. The reported composition values are obtained from MC simulations.

7.1 Jets mis-identified as leptons

A jet mis-identified as a lepton is referred to as a *fake lepton*. The dominant backgrounds which contain fake leptons are the W +jets and the multi-jet processes. The modeling of the fake leptons is known to have significant discrepancies with data. Therefore, this background is estimated by a data-driven *fake factor* method. The fake factor is defined as the ratio of the number of fake leptons being identified and *anti-identified* (anti-ID). The Z +jets and di-jet control data samples in a jet-enriched region are used to derive the fake factors needed by this analysis. The fake factors are applied to the anti-ID leptons counted in the W +jets control region with the same kinematic selections as the signal region except that one lepton is anti-ID. In this way, we extrapolate the amount of identified fake leptons in the signal region from the estimated anti-ID fake leptons.

In this thesis, the anti-ID leptons are defined as the reconstructed leptons passing looser identification criteria than the identified leptons discussed in Sec. 6.1. No isolation requirements are imposed on the anti-ID electrons or muons. Moreover, the upper bound for the transverse impact parameter significance $|d_0|/\sigma_{d_0}$ for the anti-ID muons is increased to 15. The requirements for the identified and anti-ID leptons are compared in Table 9.

Electron		Muon	
identified	anti-identified	identified	anti-identified
$p_T > 15\text{GeV}$		$p_T > 15\text{GeV}$	
$ \eta < 2.47, \text{excluding } 1.37 < \eta < 1.52$		$ \eta < 2.45$	
$ z_0 \sin \theta < 0.5 \text{ mm}$		$ z_0 \sin \theta < 0.5 \text{ mm}$	
$ d_0 /\sigma(d_0) < 5$		$ d_0 /\sigma(d_0) < 3$	$ d_0 /\sigma(d_0) < 15$
Tight identification if $p_T < 25 \text{ GeV}$	Loose identification	Tight identification	Medium identification
Medium identification if $p_T > 25 \text{ GeV}$			
Fix (Tight) isolation		Tight isolation	Veto against identified muon
$\text{AUTHOR} = 1$	Veto against identified electron		

Table 9: Requirements for the identified and anti-identified electrons and muons.

While the analysis uses two types of fake factors, the Z +jets and di-jet fake factors, we

dedicate the discussion to the Z +jets fake factor derivation from the Z +jets control sample. The derivation of Z +jets fake factors uses a sample sufficiently pure in Z +fake lepton. The events containing three leptons with $p_T > 15$ GeV are selected, and two of the leptons need to have the same flavor but opposite charge to be considered as decay products of the Z boson. Invariant-mass windows are defined to reconstruct the Z mass for the two-electron and two-muon decay channels. The lepton-pair whose invariant mass is closest to the Z mass are considered real leptons and the remaining lepton is the fake candidate. The selection requirements have 99% efficiency of correctly assigning the leptons from the Z decay.

After the non- Z +jets MC events are subtracted from the control sample, the fake factor ($F.F.$) is calculated as the probability of the third lepton being identified divided by the probability of being anti-identified:

$$F.F. = \frac{N^i}{N^a} = \frac{N_{\text{total}}^i - N_{\text{non-Z+jets}}^i}{N_{\text{total}}^a - N_{\text{non-Z+jets}}^a}, \quad (32)$$

The fake factors are derived as a function of the fake lepton transverse momentum and pseudorapidity. The fake factor as a function of electron p_T is shown in Fig. 26. The muon fake factors are derived in the region $|\eta| < 1.05$ and $1.05 < |\eta| < 2.50$ and further calculated as a function of the muon p_T as shown in Fig. 27.

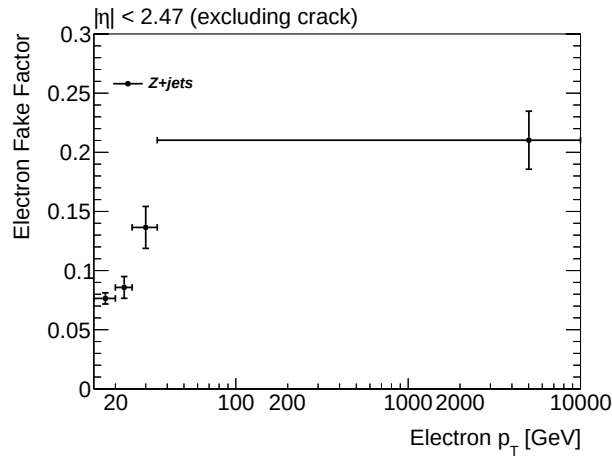


Figure 26: Fake factors derived in the jet-enriched region as a function of electron p_T .

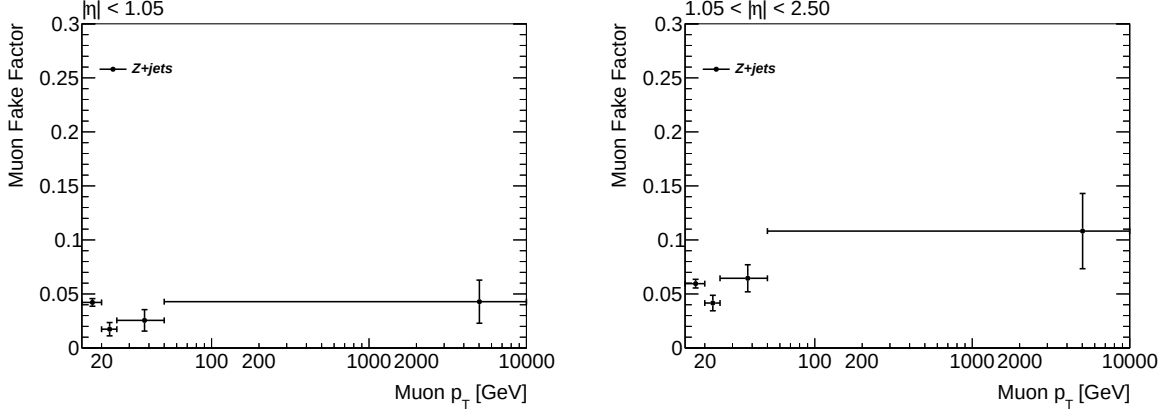


Figure 27: Fake factors derived in the jet-enriched region as a function of muon p_T in the low- η region (left) and in the high- η region (right).

Background events from mis-identified leptons can be classified as “single-fake” or “double-fake”. The single-fake backgrounds have one real lepton and one mis-identified lepton, whereas the double-fake background denotes events with both leptons mis-identified. The former generally come from the W +jets process and the latter from QCD multi-jet processes. The number of events in the signal region can be expressed as

$$N_{i,i} = N_{i,i}^{\text{EW}} + N_{i,i}^{W+\text{jets}} + N_{i,i}^{\text{QCD}}, \quad (33)$$

where the superscript ‘i’ stands for an identified lepton. The $N_{i,i}^{\text{EW}}$ term denotes all the electroweak process with two real leptons, including both signal and backgrounds. The second and third terms denote the W +jets and the QCD multi-jet backgrounds with one and two fake but identified leptons, respectively. In order to estimate the fake background, a W +jets control region is defined with exactly one identified lepton and one anti-ID lepton, and the number of events in this region is expressed as:

$$N_{i,a} = N_{i,a}^{W+\text{jets}} + N_{i,a}^{\text{EW}} + N_{i,a}^{\text{QCD}}, \quad (34)$$

where the superscript ‘a’ stands for an anti-ID lepton. Here, the $N_{i,a}^{\text{EW}}$ term denotes processes with two real leptons but one being anti-ID, which contaminate W +jets control region.

The MC simulated electroweak events are subtracted from the data before the fake factor application. The QCD multi-jet contribution is small compared to the W +jets, but it is explicitly accounted for in the single-fake estimation. Similarly, the double-fake is estimated from a dedicated region with two anti-ID leptons.

The triggering scenarios determine which fake factor to be applied:

- di-lepton trigger: apply Z+jets fake factors;
- single-lepton trigger:
 - single-fake and the anti-ID lepton fires the trigger: apply di-jet fake factors;
 - single-fake but the identified lepton fires the trigger: apply Z+jets fake factors;
 - double-fake: apply di-jet fake factors to the lepton that fires the trigger; apply Z+jets fake factors to the other lepton.

Shown in Fig. 28, the difference between the data and the electroweak processes in the W +jets control region is taken as the fakes. After applying the fake factor, we obtain the contribution of identified fake leptons in the SR.

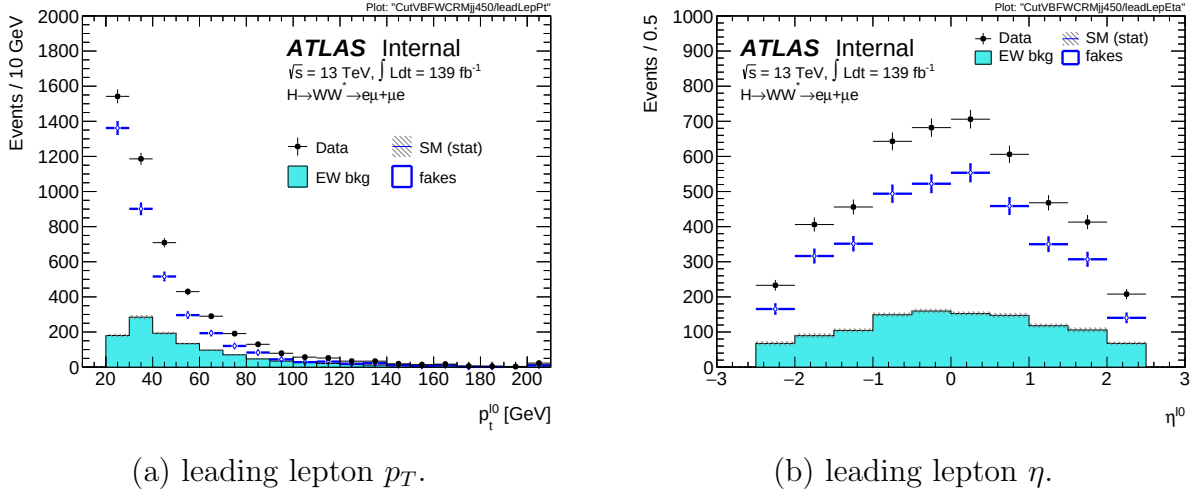


Figure 28: Kinematics distributions in the W +jets control region, with the data (black markers), electroweak processes in MC (blue histograms), and difference (blue markers) which represents the estimate of the fake background. No fake factors are applied.

7.2 Top and diboson

From MC simulation, the top plus diboson background is the dominant background with 72% in composition in the SR. The accurate estimation of this background directly impacts the signal measurements.

For the fiducial cross-section measurement, the top and diboson background is estimated from the inclusive SR by a data-over-simulation normalization factor. The estimation is driven by the high $D_{\text{top+VV}}$ regime in SR2 as it is the most pure in top and diboson background.

On the other hand, top plus diboson background is estimated from every observable bin or every few bins in the differential measurements. The estimation scheme is determined based on two criteria: the statistical uncertainty of each normalization factor needs to be below 20% and the MC simulation of top plus diboson in the region of estimation is relatively flat to reduce large extrapolations. We will specify the normalization scheme in Sec. 9.

7.3 Drell-Yan

The Drell-Yan background (Z +jets) is estimated simultaneously in the SR and the dedicated control region (CRZjets) defined with kinematic selections that are as close as possible to the SR selections while maximizing the purity for the Drell-Yan events.

The Drell-Yan background is dominated by $Z \rightarrow \tau\tau$ events where the two τ leptons decay to one electron, one muon and neutrinos. The Drell-Yan background can be discriminated from the signal as well as the other backgrounds by using the information associated to the mass of the reconstructed $\tau - \tau$ system ($m_{\tau\tau}$). The calculation is given in Eq. 30. By selecting events within a 25 GeV mass window about the Z mass, the Drell-Yan region is very pure in the Z +jets contribution (72%). Meanwhile, this region also has a requirement on the invariant mass of the two leptons ($m_{ll} < 80$ GeV) since the leptons from the Z boson decay tend to have low energy due to the extra tau neutrinos. Fig. 29 reports the $m_{\tau\tau}$ and m_{ll} distributions in the Drell-Yan region. Finally, the $\Delta\phi_{ll}$ and Δy_{jj} SR selections

are not applied in CRZjets to maintain enough statistics. The m_T variable is used as the discriminant in this region with 5 equal-sized bins between 40–150 GeV. The simulation and data are compared in Fig. 32 (left).

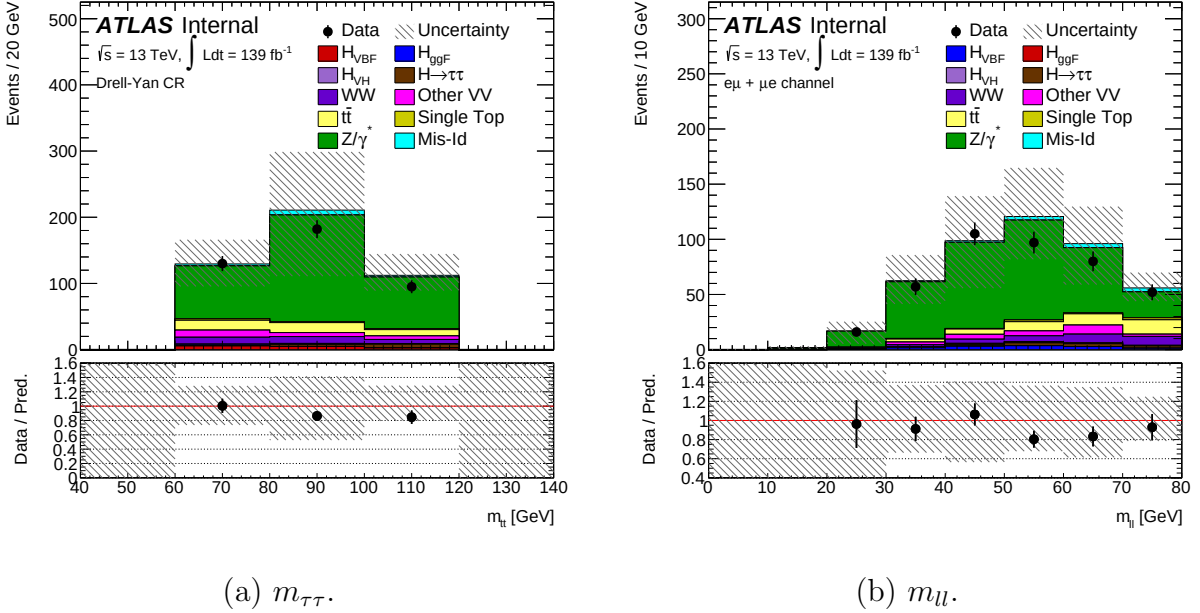


Figure 29: The $m_{\tau\tau}$ and m_{ll} distributions in the Drell-Yan region from the ATLAS Run-2 data (black data points) and MC simulation (filled histogram). The error bars on the data points show statistical uncertainties in data. The uncertainty band includes MC statistical, experimental and theory uncertainties. The bottom panel shows the data-to-MC ratio.

7.4 Gluon-gluon fusion plus two jets

The ggF control region (ggF CR) is designed to be orthogonal with the SR by inverting either the central jet veto (CJV) or the outside lepton veto (OLV) requirement, i.e. an event needs to pass one of these requirements but fail the other one. Since the ggF CR is also required to contain at least two jets with $p_T > 30$ GeV, this region is kinematically close to the signal region however with a trade-off of low ggF purity, as most of the ggF+2jets events arise from higher order QCD effect. Due to the low ggF statistics, the Δy_{jj} and cuts m_{jj} cuts are not included. Other kinematic selections are the same as the SR. A few kinematic distributions in the ggF region are given in Fig. 30.

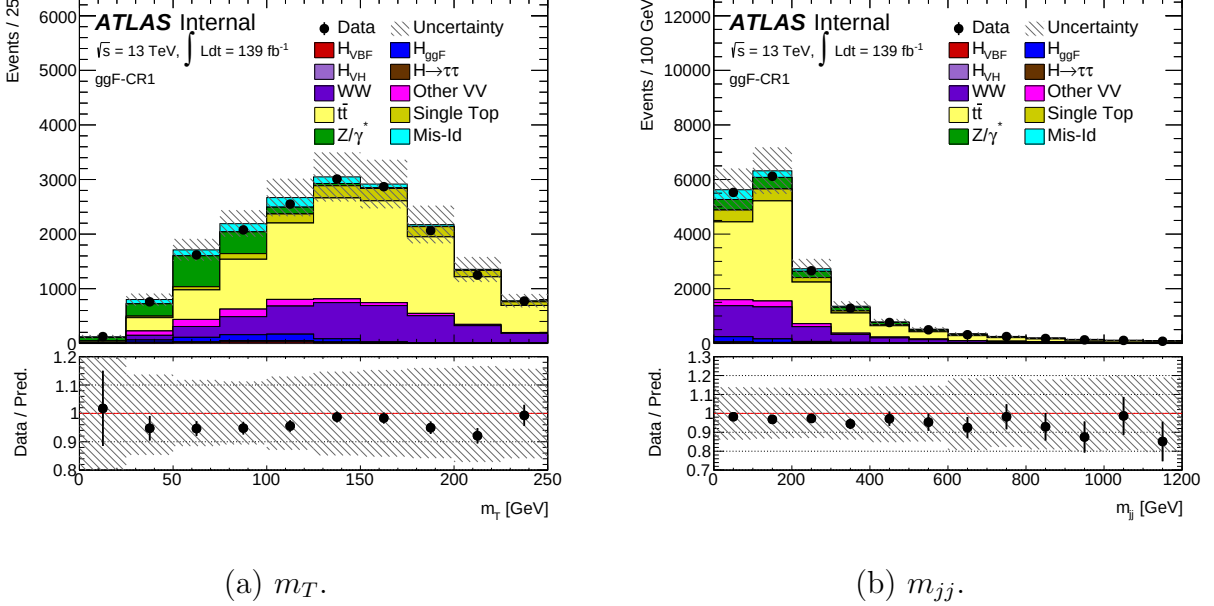


Figure 30: The m_T and m_{jj} distributions in the ggF region from the ATLAS Run-2 data (black data points) and MC simulation (filled histogram). The error bars on the data points show statistical uncertainties in data. The uncertainty band includes MC statistical, experimental and theory uncertainties. The bottom panel shows the data-to-MC ratio.

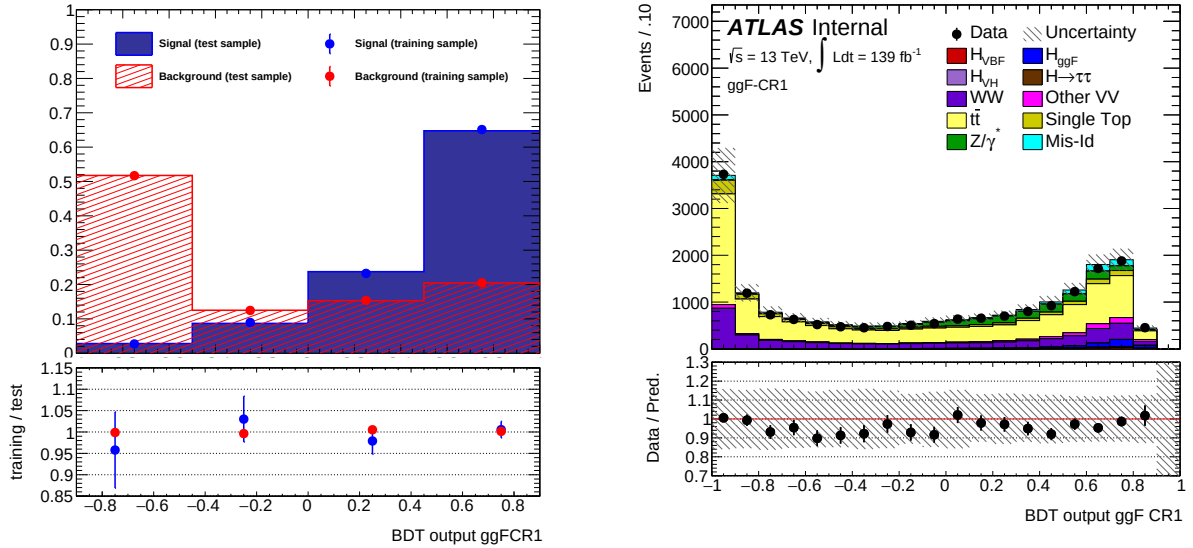
The ggF composition in the ggF CR is 2.4%, whereas the top plus diboson composition reaches 83.6% and the Drell-Yan is 7.8%. The sensitivity to ggF is enhanced by a BDT trained in the ggF CR (D_{ggF}) to distinguish ggF from all other background processes. The training variables are listed in Table 10. The BDT output results are shown in Fig. 31. The agreement between the training and testing samples demonstrates no sign of over-training. The ggF BDT is projected into four equal bins to be fitted to data. The ggF purity is the highest in the last bin at 6.6%. The ggF process is simultaneously estimated in the SR and the ggF CR.

To accurately estimate ggF contribution, the top plus diboson as well as the Drell-Yan backgrounds are estimated from data in this region with two additional data-over-simulation normalization factors. The decision to estimate Drell-Yan independently in ggF CR is also due to the different QCD and EW compositions compared to those in the signal region and Drell-Yan regions. Despite the small contribution from the ggF plus two jets process, estimating ggF directly from data reduces model dependency as our MC simulation is limited

in higher QCD perturbation. In this way, we relaxes the fit by adding the additional degree of freedom. The pre-fit MC fitting template in ggF CR is shown in Fig. 32 (right), and we observe good agreement between the simulation and data.

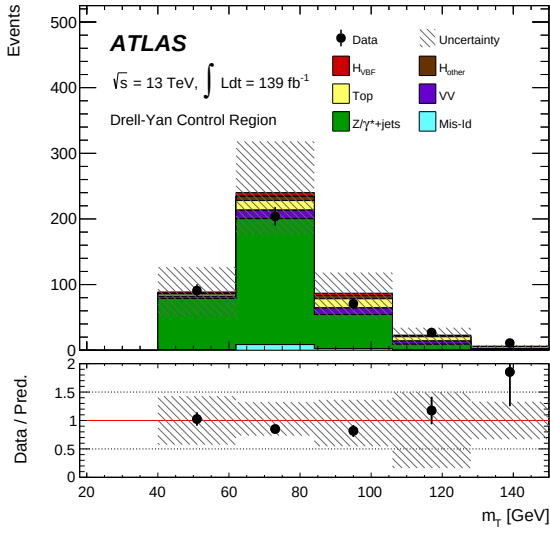
Variable	Importance
m_T	0.25
$\Delta\phi_{ll}$	0.24
Δy_{ll}	0.17
m_{ll}	0.16
$ \Delta\phi_{jj} $	0.12
$E_{T,miss}$	0.04
p_{j1}^T	< 0.01
p_{j2}^T	< 0.01

Table 10: Ranking of the variables used as input to the training of the BDT trained in the ggF control region.

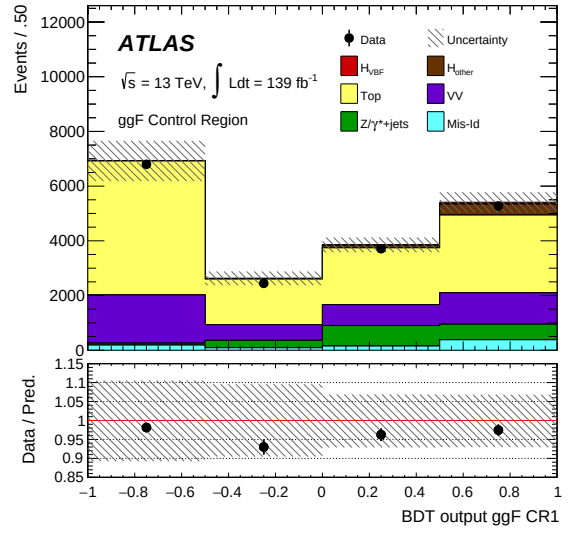


(a) The ggF BDT output of the training and testing samples. (b) The ggF BDT output of all samples in the ggF CR.

Figure 31: The ggF BDT output. (a) Normalized BDT output distributions of the ggF (signal) and all other processes (background) for the testing and training samples. The binning corresponds to the one used in the analysis. (b) fully weighted BDT output distribution of ggF signal and all background processes in the ggF control region. The error band includes the statistical, experimental and theory uncertainties. The bottom panel shows the data-to-MC ratio.



(a) m_T in CRZjets.



(b) D_{ggF} in ggF CR.

Figure 32: The pre-fit MC distribution (filled histogram) of the discriminants used in (a) CRZjets and (b) ggF CR compared with the Run-2 dataset (black data points). The error bars on the data points indicate the data statistical uncertainties, while the uncertainties in the prediction are shown by the hatched band, which includes experimental and theoretical uncertainties. The bottom panel shows the data-to-MC ratio.

8 Fiducial phase space and particle-level predictions

In this section, we present five theoretical predictions that will be compared with the unfolded data: POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7, VBFNLO+PYTHIA 8 and VBFNLO-NLO. All samples use NNPDF30_nlo_as_0118 [60] as the nominal parton distribution function (PDF). Similar to reconstruction-level selection, we define the particle-level objects and the particle-level event selections in Sec. 8.1. While we provide comparison of the cross sections from five different simulations in Sec. 8.3, we only detail the theoretical studies of the signal process using VBFNLO in Sec. 8.2 since the other three samples are the particle-level record of the reconstruction-level signal MC samples described in Sec. 5.2

8.1 Fiducial phase space

This section first introduces the object definitions at particle-level in Sec. 8.1.1. Next, in Sec. 8.1.2, we define the *fiducial phase space* in which we measure the cross sections. Both the object definitions and the fiducial phase space are chosen to be as close as possible to the reconstruction-level to minimize theoretical extrapolation and model dependency.

8.1.1 Particle-level objects

Electrons and muons at particle-level are required to stem from a hard scattering vertex and they are allowed to come from a tau decay. They are “dressed”, which means the momenta of photons emitted within a radius parameter $R=0.1$ around the lepton direction are added to the charged lepton momentum after QED final-state radiation.

The constituents of particle-level jets are those with a lifetime τ such that $c\tau \geq 10$ mm. The anti- k_t clustering algorithm [80] with $R = 0.4$ is used. The selected charged leptons and neutrinos from W boson decays are not included in the jet clustering. We use ghost-matching association [81] between jets and final-state B-hadrons in the event to tag the particle-level b -jets. A jet is removed if it overlaps with an electron within $\Delta R < 0.2$.

Finally, particle-level MET is the transverse component of the vectorial sum of non-interacting particles.

8.1.2 Particle-level selection

The particle-level kinematic selections are summarized in Table 11.

Category	Fiducial Requirement	Cut Value
Leptons	Opposite sign	Yes
	electron and muon channel	Yes
	$ \eta(\ell) $	< 2.5
	$p_T^{\ell_1}$	$> 22 \text{ GeV}$
	$p_T^{\ell_2}$	$> 15 \text{ GeV}$
	N_{leptons}	≥ 2
	$\Delta R(\ell, \ell)$	> 0.1
	m_{ll}	$> 10 \text{ GeV}$
	$m_{\tau\tau}$	$< m_Z - 25 \text{ GeV}$
	$\Delta\phi_{ll}$	< 1.4
Jets (anti- k_t R=0.4)	p_T^{j1}, p_T^{j2}	$> 30 \text{ GeV}$
	$ \eta(\text{jet}) $	< 4.5
	N_{jets}	≥ 2
	$N_{\text{b-jet}} (p_T > 20 \text{ GeV})$	< 1
	m_{jj}	$> 450 \text{ GeV}$
	Δy_{jj}	> 2.1
lepton-jet separation	$\Delta R(\ell, \text{jet})$	> 0.4
VBF	OLV	Yes
	CJV($< 20 \text{ GeV}$)	Yes

Table 11: Definition of fiducial phase space for VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ in the $e - \mu$ decay channel.

8.2 VBFNLO predictions

VBFNLO [62] is an efficient *parton-level* MC program that can simulate VBF processes at NLO in QCD and EW. Parton-level event is the most fundamental record from a MC simulation as the quarks (or partons) in the final state are yet to be showered. Generally, a parton-level generator, such as MADGRAPH5_AMC@NLO, POWHEG and VBFNLO, calculates the matrix element following the Feynman rules of the Standard Model accord-

ing to the desired initial state and final state particles with intermediate propagators and subsequently obtains the amplitude and the probability (cross section) for an event to occur at each “point” in the kinematic phase space. The output events are usually stored in the standard Les Houches Accord (LHA) format [82] to be interfaced with generators such as PYTHIA 8 and HERWIG 7 to simulate the parton shower and hadronization effects and produce the particle-level predictions. Parton-level simulations are less time-consuming, giving us an opportunity to investigate the QCD and EW effects, the use of different scales or PDFs, and the parton-shower effect.

At leading order (LO), VBFNLO can generate LHA events. We interface the program with PYTHIA 8 and obtain another particle-level sample. The showered events are processed by a common analysis framework developed by the ATLAS HWW analysis group to be consistent with other showered samples. In this section, we explore the fixed-order calculations at LO and NLO with VBFNLO and the shower effect using the VBFNLO+PYTHIA 8 sample. Both LO and NLO VBFNLO samples were produced with the same dynamical QCD scale (the momentum transfer of the exchanged gauge boson) and the G_F EW scheme.

VBFNLO and Pythia 8

The parton-level VBFNLO events are generated in the inclusive VBF HWW phase space, in which all generations of leptons are allowed in the final state. The tau leptons from the W boson decay are decayed by PYTHIA 8. Specifically, the leptonic decay products of the tau lepton are part of the fiducial region. While the SM prediction of the fraction of tau decay events in the inclusive phase space is 18%, it is 14% in the fiducial region, since part of the W mass energy transfers to the additional neutrino in the process and the selection cuts remove events with low-energy leptons.

Since the VBFNLO generator at parton-level cannot further decay the final state lepton, VBFNLO sample at NLO does not include events from the tau decays. This causes the VBFNLO NLO sample to have a much smaller fiducial cross section, and the shapes of

several lepton distributions skew relative to the other samples. To correct for the lack of tau-decay events, we calculate correction factors, which are the bin-by-bin ratios between the VBFNLO+PYTHIA 8 differential cross sections with and without tau-decay events, and multiply these factors to the VBFNLO-NLO distributions. Some examples of the correction factors are shown in Fig. 33.

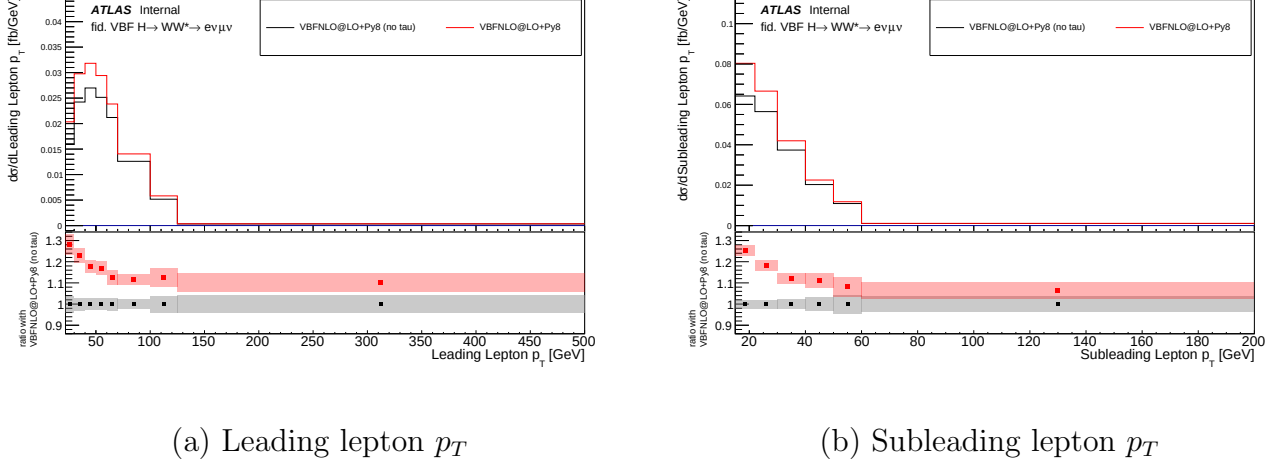


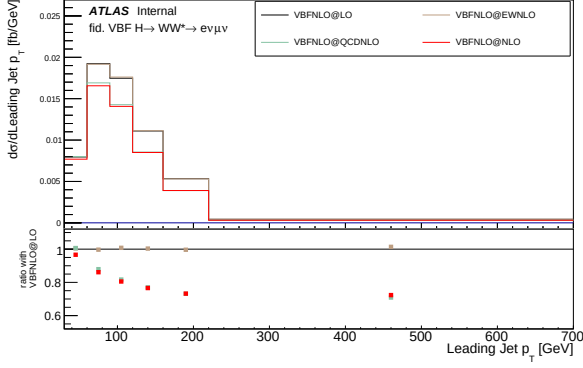
Figure 33: Distributions of lepton p_T produced by VBFNLO+PYTHIA 8 with and without tau decay. The ratios in the bottom panels are used as the tau-decay correction to the VBFNLO at NLO sample. The error bands in the ratio panels represent the statistical uncertainties of the MC samples.

VBFNLO at next-to-leading order

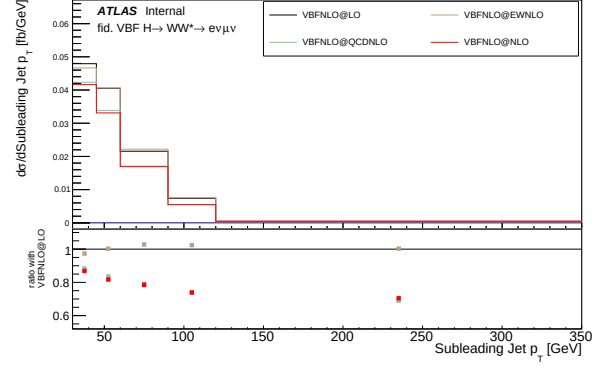
VBFNLO uses the VEGAS algorithm [83] for integrating over the phase space to obtain the total cross section as well as the differential cross sections. Applying EW and QCD NLO corrections separately, we obtain a k-factor of 1.00 and 0.82 in the fiducial phase space respectively. The simultaneous EW and QCD NLO correction gives a k-factor of 0.81. A comparison of the leading and subleading jet p_T with and without EW correction, and with and without QCD correction is shown in Fig. 34. The comparisons of these different VBFNLO samples show that the QCD NLO correction dominate the EW correction in most of the fiducial region. The higher QCD effect is more significant at higher energy scale.

Fig. 35 shows the impact of the parton shower on the LO sample. The VBFNLO

calculation at LO with parton shower results in distributions that are similar in shape to those at NLO, illustrating the higher order effect and the parton shower effect are similar in the defined fiducial region.

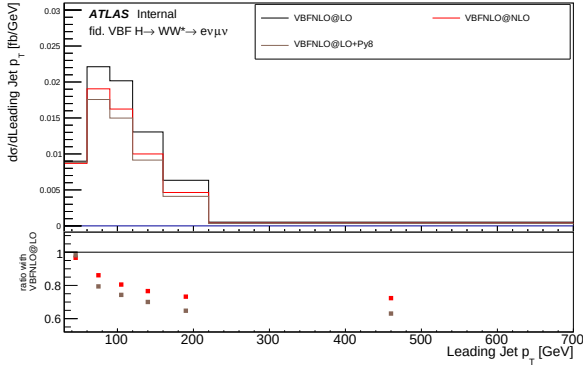


(a) Leading jet p_T

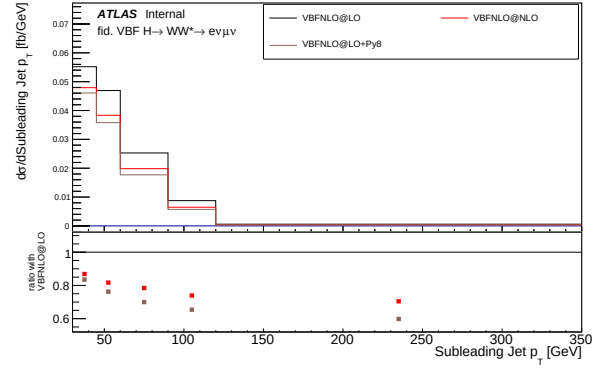


(b) Subleading and subleading jet p_T

Figure 34: Leading and subleading jet p_T generated by VBFNLO at different orders of calculation in QCD and EW: LO, EW NLO, QCD NLO, and NLO in both EW and QCD. The distributions are not corrected for tau decays.



(a) Leading jet p_T



(b) Subleading jet p_T

Figure 35: Leading and subleading jet p_T generated by VBFNLO-NLO and VBFNLO+PYTHIA 8 compared to VBFNLO-LO.

VBFNLO uncertainties

Three sources of theory uncertainties are studied in VBFNLO: PDF, α_s variation, and QCD scale. For each of the systematics variation, a VBFNLO sample is produced, and the fiducial and differential cross sections are calculated. The LHAPDF6.2.3 package [84] is used to include different PDF sets. The PDF4LHC50 PDF set [85] is used to evaluate the PDF uncertainties and α_s uncertainties evaluated at the $\pm 1\sigma$ values $\alpha_s = 0.017$ and 0.019 . The QCD scale uncertainty is evaluated by taking the envelope of the different predictions with renormalization scale and factorization scales multiplied by a factor of 0.5 or 2.0 independently.

8.3 Theory predictions comparison

The fiducial cross sections and associated uncertainties of all five samples are summarized in Table 12. The five NLO samples provide cross sections in agreement at 1 – 11% level. The difference is primarily due to the difference in the modeling of the third jet in the VBF process. Among the NLO samples, VBFNLO-NLO parton-level calculations provides the largest cross sections, as the NLO sample does not contain soft third jets from parton shower. As a results, the central jet veto removes less events from the VBFNLO samples. Overall, the cross section give by VBFNLO-LO is greater than other samples by 19-28% because the sample does not contain a third jet.

Due to the lack of theoretical systematics of the MADGRAPH5_AMC@NLO+HERWIG 7 sample, we only provide its fiducial cross section. The m_{jj} and m_{ll} differential cross sections of the other four samples are compared in Fig. 36 with their respective total theoretical uncertainties. As for the rest of the 11 observables, we compare them directly with the data in Sec. 10. From these results, we observe a significant reduction of theoretical uncertainties from QCD LO to NLO calculations, most notably for those associated to QCD scales and PDFs.

Sample	cross section	scale	PDF	α_s	combined
POWHEG+PYTHIA 8	2.07 fb	+1.0%/ - 0.3%	$\pm 1.8\%$	$\pm 0.4\%$	+2.1%/ - 1.9%
VBFNLO+PYTHIA 8	2.01 fb	+5.4%/ - 4.5%	$\pm 2.3\%$	$\pm 1.3\%$	+6.0%/ - 5.2%
VBFNLO at NLO (with tau correction)	2.18 fb	+2.4%/ - 4.8%	$\pm 4.3\%$	$\pm 1.5\%$	+5.1%/ - 6.6%
VBFNLO at LO (with tau correction)	2.69 fb	+5.4%/ - 4.5%	$\pm 2.3\%$	$\pm 1.3\%$	+6.0%/ - 5.2%
MADGRAPH5_AMC@NLO+ HERWIG 7	2.16 fb	—	—	—	—
POWHEG+HERWIG 7	1.93 fb	+1.0%/ - 0.3%	$\pm 1.8\%$	$\pm 0.4\%$	+2.1%/ - 1.9%

Table 12: The fiducial cross sections and the theory systematics of all parton-level and particle-level samples.

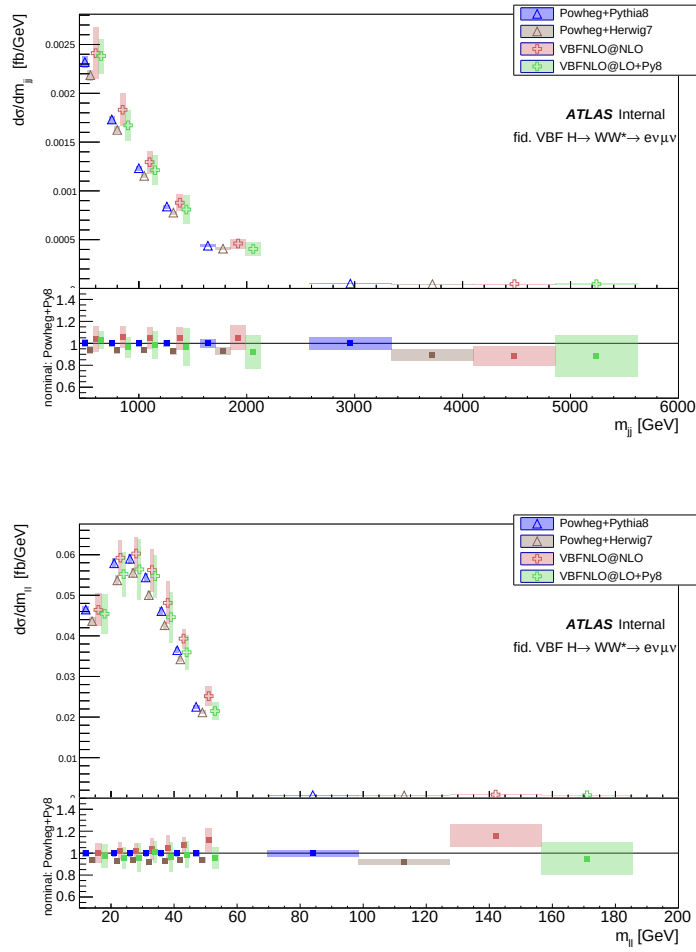


Figure 36: Comparison of m_{jj} and m_{ll} differential cross sections by POWHEG + PYTHIA 8, POWHEG + HERWIG 7, VBFNLO-NLO and VBFNLO + PYTHIA 8 predictions, with their total theory uncertainties. The four predictions follow the same binning but the points are offset to make each point visible for the comparison.

9 Signal extraction strategy

The key ingredients of the statistical analysis are the data collected at the experiment and the predictions simulated by various Monte Carlo programs. The goal is to fit the model with free parameters to data where the free parameters are the cross sections to be measured, background normalization factors, as well as other degrees of randomness originating from detector inefficiency and resolution and from the choices of simulation programs and the input parameters [86].

This section details the fitting method employed to measure the inclusive fiducial and differential cross sections. The strategy extracts signal and estimate background simultaneously from data by minimizing the negative logarithm of a binned profile likelihood which includes the systematic uncertainties as nuisance parameters [87]. First, we discuss the different types of systematic uncertainties considered in this analysis in Sec. 9.1. Next, Sec. 9.2 focuses on the unfolding strategy which is used to reduce detector distortion on the differential cross sections. Finally, we elaborate the fiducial and differential cross-section measurement strategy in Sec. 9.3.

9.1 Systematics

The uncertainties on the results are estimated from both experimental and theoretical sources. They affect both the shape and normalization of the MC templates used in the fit and, where applicable, the acceptance of the signal as well as the bin-to-bin migrations. The uncertainties are symmetrized by taking the maximum of the upward and downward uncertainties. Due to the statistical limitation of the MC simulation, the input systematics which impact the shape of the templates are smoothed with a Gaussian kernel procedure [88]. Each symmetrized and smoothed uncertainty is later included in the likelihood function as a Gaussian-constrained nuisance parameter.

Experimental uncertainties relevant to the leptons originate from the reconstruction,

isolation and identification efficiency, the energy scale and resolution [89, 90] as well as their trigger efficiency [91, 92]. Furthermore, the uncertainties induced by jet energy scale and resolution are obtained from simulations and *in situ* measurements [93]. The uncertainty in the suppression of jets originating from pile-up interactions is studied in various simulated pile-up models [94], in the jet-vertex association [14], and in the b-tagging efficiency [95]. The uncertainties on MET are evaluated by propagating the uncertainties from the jets and leptons and by applying momentum scale and resolution uncertainties to the track-based soft term [75]. Finally, the luminosity uncertainty is $\pm 1.7\%$, measured by the LUCID-2 detector [53, 96].

The theoretical uncertainties for the signal and the major irreducible background processes are included. For evaluation of the fake background using fake factors, three sources of uncertainties are involved: the statistical uncertainty on the extrapolation factor itself, an uncertainty related to the subtraction of electroweak processes from the Z+jets sample used to derive the fake factor, and an uncertainty in the correction factor for the sample composition.

As for the signal and other background processes, the renormalization and factorization scale uncertainty, the PDF uncertainty and the α_s uncertainty are evaluated for each process. By comparing alternative samples which used different parton shower simulators or different matching schemes, we acquire the uncertainties related to different simulation algorithms (referred as a two-point comparison). Signal theory uncertainties also impact the unfolding matrices which are constructed by the particle-level and reconstruction-level MC samples.

Finally, the statistical uncertainties from the MC simulation for each bin of the MC templates are also included as nuisance parameters.

9.2 Unfolding

The unfolding process corrects for detector effects such as inefficiencies and limited resolution on the kinematic observables to be measured. In this thesis, we adopt the *matrix-based*

unfolding technique, taking into account the bin-to-bin migrations. The migration matrices, also referred as the response matrices, are constructed using the truth and reconstruction records of the nominal signal sample. Although the signal region and fiducial region are defined as kinematically close as possible, there are events which fall into one region but not the other. Those events that pass the truth-level selections but fail at reconstruction-level are called *misses*, and those that pass at reconstruction-level but fail at truth-level are called *fiducial fakes*. The migration matrix ($M_{n,m}$) is constructed such that the number of reconstructed events in the n^{th} bin ($\nu_n^{\text{recon.}}$) can be parametrized as a function of the number of events in the truth bins (denoted as $\nu_m^{\text{fid.}}$ for the m^{th} truth bin):

$$\nu_n^{\text{recon.}} = \sum_m^{N_{\text{obs.}}} M_{n,m} \nu_m^{\text{fid.}} + f_n \nu_n^{\text{fid.}}. \quad (35)$$

Each migration matrix element $M_{n,m}$ includes a correction factor to “remove” the misses, and the fiducial fakes are corrected by the last term $f_n \nu_n^{\text{fid.}}$. By substituting $\nu_m^{\text{fid.}}$ with the product of the integrated luminosity L_{int} and the cross section per truth bin $\sigma_m^{\text{fid.}}$, we obtain

$$\nu_n^{\text{recon.}} = \sum_m^{N_{\text{obs.}}} M_{n,m} L_{\text{int}} \sigma_m^{\text{fid.}} + f_n L_{\text{int}} \sigma_n^{\text{fid.}}. \quad (36)$$

The theoretical uncertainties on the migration matrix elements and on the fiducial fakes are included.

The efficiency (ϵ_n) and purity (P_n) are defined to quantify the bin migrations:

$$\epsilon_n = \frac{\nu_n^{\text{fid., recon.}}}{\nu_n^{\text{fid.}}}; \quad (37)$$

$$P_n = \frac{\nu_n^{\text{fid., recon.}}}{\nu_n^{\text{recon.}}}, \quad (38)$$

where $\nu_n^{\text{fid., recon.}}$ is the number of MC events matched to the same n^{th} bin at truth-level and reconstruction-level, and $\nu_n^{\text{recon.}}$ and $\nu_n^{\text{fid.}}$ are the total number of events in the respective

reconstructed and truth bin. The binning scheme for each observable is chosen to maintain a purity and efficiency higher than 60%. Meanwhile, the bin edges are optimized to balance sufficient number of bins to retain the features of the observable shape and sufficient number of signal events per observable bin. Finally, the chosen binning is studied to provide the sensitivity for the Effective Field Theory interpretation. The final binning scheme is summarized in Table 13.

Observable [unit]	Bin Edges
m_{ll} [GeV]	10,30,40 55,Inclusive
m_{jj} [GeV]	450,950,1200,1500,2200,6000
$\cos(\theta^*)$	0,0.0625,0.125,0.1875,0.3125,1.0
Δy_{ll}	0,0.4,0.6,0.8,1.0,Inclusive
Δy_{jj}	2.1,4.0,5,5.5,6.25,9.0
$\Delta\phi_{ll}$ [rad]	0,0.2,0.4,0.6,0.8,1.4
$\Delta\phi_{jj}$ [rad]	$-\pi, -\frac{\pi}{2}, 0, \frac{\pi}{2}, \pi$
Higgs p_T [GeV]	0,80,120,260,850
$p_T^{\ell\ell}$ [GeV]	0,60,100,140,1000
Leading Lepton p_T [GeV]	22,40,50,70,100,500
Subleading Lepton p_T [GeV]	15,30,40,50,200
Leading Jet p_T [GeV]	30,120,160,220,700
Subleading Jet p_T [GeV]	30,60,90,120,350

Table 13: Binning optimized for each distribution. Inclusive bins have no upper bound.

The migration matrices and fiducial fakes for a few observables are shown in Fig. 37 where the matrices for jet and MET-related observables such as Higgs p_T and m_{jj} have larger migrations because they are more susceptible to detector effect, whereas the angular and lepton observables such as m_{ll} and $\Delta\phi_{jj}$ have nearly diagonal migration matrices.

9.3 Statistical analysis for the cross-section measurements

In general, to describe the probability of observing N events with the average expectation of ν events, one uses the Poisson distribution:

$$P(N|\nu) = \frac{\nu^N}{N!} e^{-\nu}. \quad (39)$$

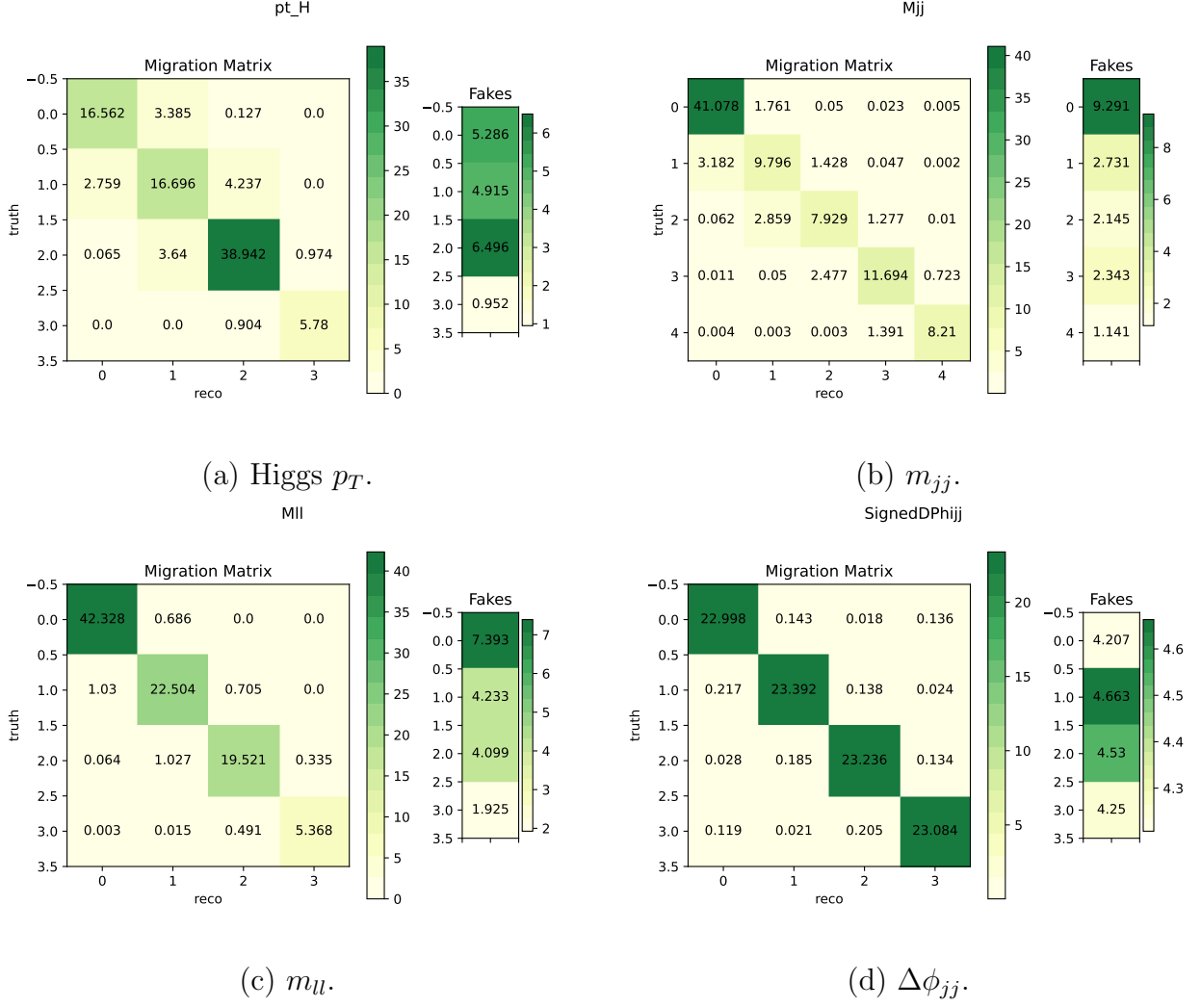


Figure 37: Migration matrices for Higgs p_T , m_{jj} , m_{ll} and $\Delta\phi_{jj}$. The bin values correspond to the number of MC events in the truth and reconstruction bins and the fiducial-fakes category.

We will consistently use N for the observed number of events and ν for the expectation. For each discriminant bin (subscripted with i) in each region (subscripted with “reg”), we denote the probability as $P(N_i^{\text{reg}}|\nu_i^{\text{reg}})$. For the two control regions in this analysis, the expectation value is parametrized as a function of the background normalization factors (μ_{bkg}) in the corresponding region:

$$\nu_i^{\text{CR}_x} = \sum_{\text{bkg}} \mu_{\text{bkg}} \nu_{\text{bkg},i}^{\text{CR}_x} + \nu_{\text{misc},i}^{\text{CR}_x}, \quad (40)$$

where $\nu_{\text{bkg},i}^{\text{CR}_x}$ is the number of simulated events in this bin for the background processes normalized to data, whereas $\nu_{\text{misc},i}^{\text{CR}_x}$ represents the number of events from miscellaneous processes fixed to their nominal MC predictions. The signal events are accounted for in the miscellaneous events in the control regions.

For the signal region, which we denote as SR_n for the n^{th} bin of an observable (there is only one bin for the fiducial cross section measurement), the total number of expected events in the i^{th} bin of the 2-dimensional $D_{\text{VBF}} - D_{\text{top+VV}}$ discriminant can be written as

$$\nu_i^{\text{SR}_n} = \nu_{\text{sig},i}^{\text{SR}_n} + \sum_{\text{bkg}} \mu_{\text{bkg}} \nu_{\text{bkg},i}^{\text{SR}_n} + \nu_{\text{misc},i}^{\text{SR}_n}, \quad (41)$$

where $\nu_{\text{sig},i}^{\text{SR}_n}$ is the number of events from the signal process. In the fiducial cross-section measurement, $\nu_{\text{sig},i}^{\text{SR}_n}$ at reconstruction-level can be parametrized by the fiducial cross section $\sigma^{\text{fid.}}$.

$$\nu_{\text{sig},i}^{\text{SR}} = A_i \cdot (L_{\text{int}} \cdot C \cdot \sigma^{\text{fid.}}). \quad (42)$$

The C -factor is the correction factor for detector effects, and its nominal value is 0.45; the total number of expected signal events in the signal region is the product in the parenthesis; A_i is the acceptance factor for each discriminant bin. The C -factor and acceptance factors are both obtained from the MC simulations, and their systematic uncertainties are incorporated as nuisance parameters.

In the differential measurements, as the fiducial cross section is measured per observable bin, $\sigma^{\text{fid.}}$ in Eq. 42 becomes a vector ($\vec{\sigma}^{\text{fid.}}$ or $\sigma_m^{\text{fid.}}$) and the correction factor evolves into the corresponding response matrix ($M_{n,m}$) and fiducial fakes ($f_n \sigma^{\text{fid.}}$) as discussed in Sec. 9.2. The modified Eq. 36 with the acceptance factor $A_{n,i}$ for the n^{th} observable bin and the i^{th} discriminant bin yields

$$\nu_{\text{sig},i}^{\text{SR}_n} = A_{n,i} \cdot L_{\text{int}} \cdot \left(\sum_m^{N_{\text{obs.}}} M_{n,m} \sigma_m^{\text{fid.}} + f_n \sigma^{\text{fid.}} \right). \quad (43)$$

While the background normalization strategies for ggF and the Drell-Yan are the same in the fiducial and the differential measurements which simultaneously normalize the simulation to data in the respective control region and the inclusive signal region (detailed in Sec. 7), the top and diboson normalization has a finer granularity in the differential measurements. For the fiducial measurement, the top and diboson are estimated in the inclusive signal region with one single normalization factor. On the other hand, the top and diboson are estimated in every observable bin or every few observable bins. The normalization scheme for each observable is summarized in Table 14. The bin numbers in the bracket share the same independent top and diboson normalization factor.

Observable	bin numbers
$m_{jj}, \Delta y_{jj}$	[1], [2, 3], [4, 5]
$m_{ll}, \Delta\phi_{jj}, \Delta y_{ll}, p_T^{j1}, p_T^{j2}, p_T^{\ell2}$	[1, 2], [3, 4]
$\Delta\phi_{ll}, \Delta y_{ll}$	[1, 2, 3], [4, 5]
Higgs p_T	[1], [2, 3], [4]
$\cos(\theta^*)$	[1, 2], [3], [4], [5]
$p_T^{\ell1}$	[1, 2], [3], [4, 5]

Table 14: Groups of observable bins which share the same independent top and diboson normalization factor.

Finally, the systematic uncertainties are included in the likelihood function as nuisance parameters (NPs), denoted as θ_i , and each NP is constrained by a Gaussian function $\mathcal{N}(\theta_i)$.

Now that we have parametrized the number of expected events as a function of the parameters we would like to extract from the fit, the final likelihood function is constructed as the product of the Poisson probability functions in all the signal and control regions and the constraint Gaussian functions for all the NPs:

$$\mathcal{L}(\vec{\sigma}^{\text{fid.}}, \vec{\mu}, \vec{\theta}) = \prod_n^{N_{\text{obs.}}} \prod_i^{N_{\text{disc.}}^{\text{SR}_n}} P(N_i^{\text{SR}_n} | \nu_i^{\text{SR}_n}(\vec{\sigma}^{\text{fid.}}, \vec{\mu})) \times \prod_x^{\text{CR}_x} \prod_i^{N_{\text{disc.}}^{\text{CR}_x}} P(N_i^{\text{CR}_x} | \nu_i^{\text{CR}_x}(\vec{\mu})) \times \prod_i^{N_{\text{NP}}} \mathcal{N}(\theta_i). \quad (44)$$

The cross sections and background normalization factors are extracted at the maximum value of the likelihood function (or the minimum at the negative-log-likelihood $-\ln \mathcal{L}$). To

separately estimate the uncertainty from data statistics and the systematics, the statistical uncertainties are obtained by minimizing the negative-log-likelihood without the Gaussian functions for the nuisance parameters and the systematics uncertainties are calculated by quadratically subtracting the statistical uncertainties from the total uncertainties.

10 Cross-section measurement results

In this section, we report the observed cross sections using the ATLAS Run-2 dataset, corresponding to 139 fb^{-1} of integrated luminosity at a center-of-mass energy $\sqrt{s} = 13 \text{ TeV}$.

10.1 Fiducial cross section

Before using the ATLAS Run-2 dataset, the fitting strategy is tested by a pseudo-dataset (Asimov data) matched to the nominal predictions. By minimizing the negative-log-likelihood as discussed in Sec. 9, we obtain the expected cross section in the fiducial phase space:

$$\sigma_{\text{exp}}^{\text{fid}} = 2.07_{-0.41}^{+0.44} \text{ fb} = 2.07_{-0.34}^{+0.36} (\text{stat})_{-0.23}^{+0.25} (\text{sys}) \text{ fb}. \quad (45)$$

The results verify closure with the nominal signal prediction provided by POWHEG+PYTHIA 8. Furthermore, the uncertainty given by the Asimov fit illustrates the expected sensitivity of the measurement. The statistical (stat) uncertainty and systematic (sys) uncertainty on the cross-section measurement are propagated from the data statistical uncertainties and modeling systematic uncertainties, respectively. The expected precision given by the statistical fluctuation in data is about 17%, whereas the systematic uncertainty is 12%.

The observed fiducial cross section using the ATLAS Run-2 dataset,

$$\sigma_{\text{obs}}^{\text{fid}} = 1.68_{-0.38}^{+0.40} \text{ fb} = 1.68_{-0.32}^{+0.33} (\text{stat})_{-0.21}^{+0.23} (\text{sys}) \text{ fb}, \quad (46)$$

is approximately 1σ deviation below the expected cross section. The negative-log-likelihood profiles of observed and the expected cross sections are compared in Fig. 38.

The overall relative precision of the cross-section measurement is about 23%, dominated by the statistical uncertainty in the data sample. Table 15 summarizes the breakdown of the relative uncertainties in the measured fiducial cross section. The sum in quadrature of the individual sources differs from the total uncertainty due to correlations among the different

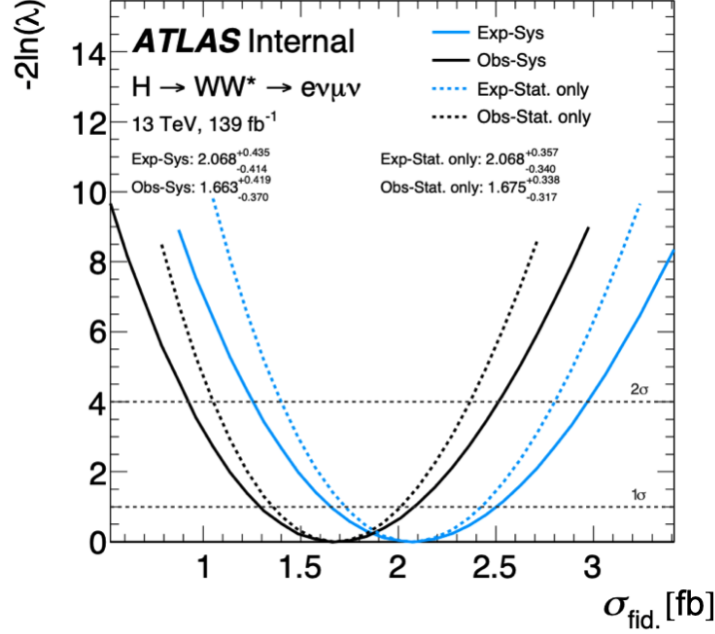


Figure 38: Profiled negative-log-likelihood of σ^{fid} using the ATLAS Run-2 dataset (black) and the Asimov dataset (blue). Dashed curve shows the likelihood profile with only statistical uncertainty, whereas solid curve shows that including all nuisance parameters.

systematic uncertainties. The systematic uncertainty is dominated by the theory uncer-

Source	$\Delta\sigma^{\text{fid}}/\sigma^{\text{fid}}$ [%]
Data statistical uncertainties	19.6
Total systematics uncertainties	12.9
MC statistical uncertainties	5.1
Luminosity	1.5
Mis-Id lepton	<0.1
Flavor tagging	<0.1
Jet and MET	4.9
Leptons	1.5
Drell-Yan modeling	3.4
Diboson modeling	3.4
Top modeling	6.2
ggF modeling	5.4
Signal modeling	4.9

Table 15: Breakdown of the relative uncertainties $\Delta\sigma^{\text{fid}}/\sigma^{\text{fid}}$ in the inclusive fiducial cross-section measurement. The middle block shows the experimental systematic uncertainty breakdown and the bottom block shows those of the theory systematic uncertainties from different processes.

tainties, especially the ones from two-point comparisons of two MC samples using different generators. They tend to have large shape differences in the discriminants which cannot be absorbed by the background normalization uncertainties. Moreover, the jet experimental systematic uncertainties have large impact due to the jet kinematic selections. The MET uncertainties are propagated to the calculation of several observables, including m_T which is one of the most important input variables for the BDTs and is the discriminating variable in CRZjets.

Moreover, the comparison between the measured cross section and the theoretical predictions is shown in Fig. 39. The predictions and theoretical uncertainties are obtained from the POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7, and VBFNLO+PYTHIA 8 at particle-level, as well as from VBFNLO at parton-level in LO and NLO. Except for the predictions given by MG5+HERWIG 7 and VBFNLO-LO, other predictions agree with the observed cross section within the measured uncertainty and the theoretical uncertainty of each prediction. The disagreement with the MG5+HERWIG 7 prediction is small and can be potentially covered by the theoretical uncertainty if calculated. On the other hand, the fixed order calculation by VBFNLO at LO in QCD and EW is higher than the observed cross section by approximately twice the measured uncertainty. Since the LO simulation does not contain a third jet, consequently the central jet veto does not act on events from the VBFNLO-LO sample. Hence, the large difference suggests that the measurement is sensitive to the higher order perturbative QCD effect.

Furthermore, we report the *p-value* which represents the probability of observing the deviation between the SM hypothesis and the measurement given the statistical and systematic uncertainties. The p-value is assessed using a χ^2 test where the χ^2 can be calculated from the likelihood-ratio between the observed value ($\sigma_{\text{obs}}^{\text{fid.}} = 1.68 \text{ fb}$) and the expected value ($\sigma_{\text{exp}}^{\text{fid.}} = 2.07 \text{ fb}$) according to Wilks' theorem [97]

$$\chi^2 = -2\ln \left(\frac{\mathcal{L}(\sigma_{\text{exp}}^{\text{fid.}})}{\mathcal{L}(\sigma_{\text{obs}}^{\text{fid.}})} \right). \quad (47)$$

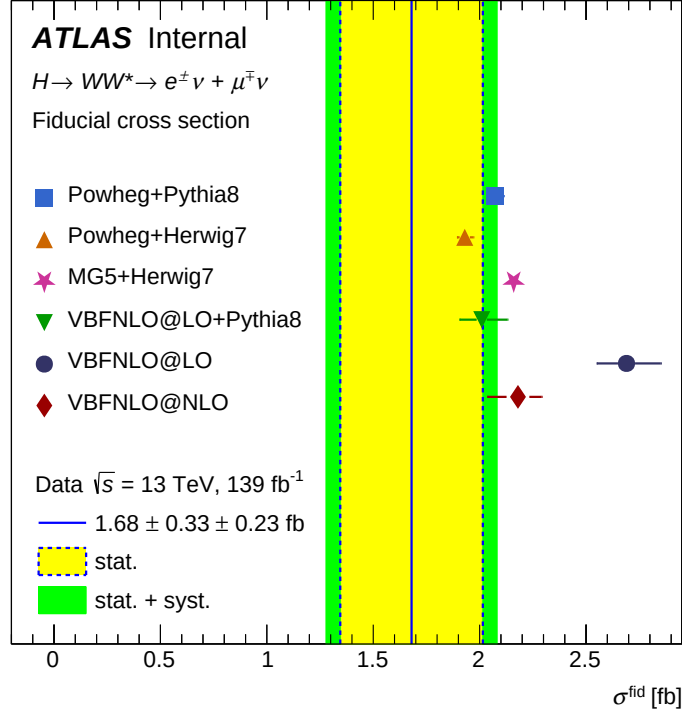
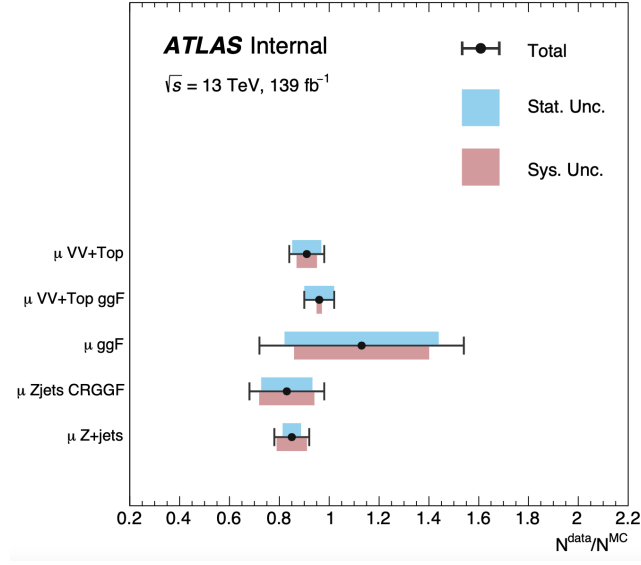


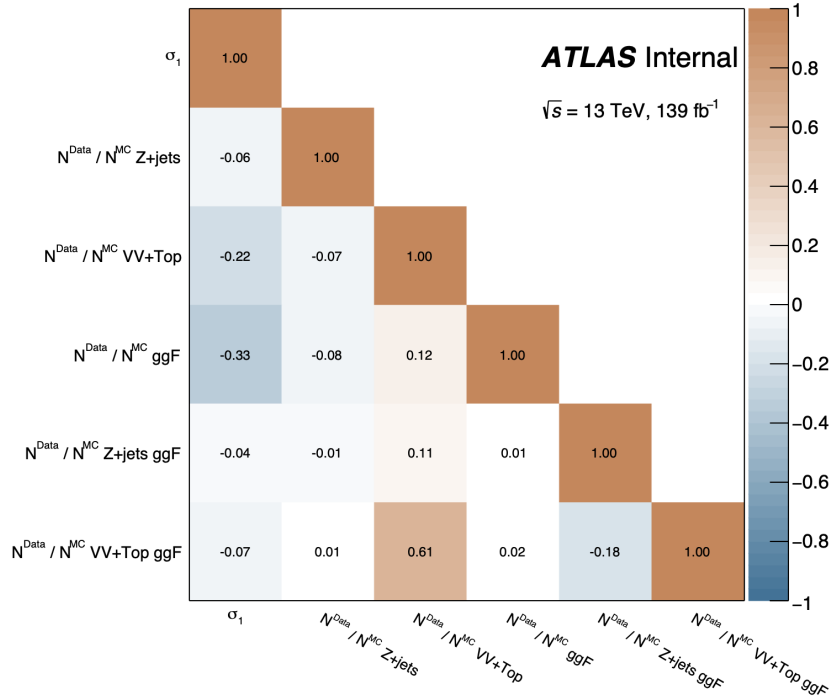
Figure 39: The measured fiducial cross section at $\sqrt{s} = 13 \text{ TeV}$ in comparison with the theoretical predictions in the fiducial phase space obtained from the parton shower generators POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from the parton-level generator VBFNLO at LO and NLO. The vertical bands around the measurement indicate the statistical uncertainty (yellow) and the sum in quadrature of statistical and systematic uncertainties (green).

The p-value for the POWHEG+PYTHIA 8 prediction yields 0.33 and above 0.20 for the other four predictions.

The results of the background normalization factors and correlations among the fiducial cross section and the normalization factors are reported in Fig. 40. The two normalization factors for Drell-Yan in different regions are compatible within their uncertainties. Similar observation is made for the top and diboson normalization as well. Next, there is no extreme correlation among the parameters comparing to the expectation. The high positive correlation between the two normalization factors in the SR and in ggF CR for the top plus diboson background is resulted from the top and diboson theory systematic uncertainties.



(a) Background normalization factors.



(b) Correlations of the measured parameters.

Figure 40: The measured background normalization factors (μ) and the correlations among the measured fiducial cross section (σ_1) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$). The Drell-Yan normalization factors are labeled as “Z+jets” for the parameter used for simultaneous estimation in SR and CRZjets, and “Z+jets CRGGF” or “Z+jets ggF” for the one used in ggF CR. The only ggF normalization factor is labeled “ggF”. Finally, the top plus diboson background is denoted by “VV+Top” for the parameter from SR and “VV+Top ggF” for the one from ggF CR.

10.2 Differential cross sections

The Asimov data is used to check for closure with the SM prediction and for evaluating the expected precision for the differential measurements. The Asimov results are presented in Appendix A. Closure is obtained for all differential distributions. The expected precision ranges from 40-80%, with a few exceptions of bins in low-statistics region with $> 100\%$ uncertainty. We plan to set an upper bound in the cross sections measured in these regions.

The measured differential fiducial cross sections as a function of the transverse momentum of the Higgs boson is shown in Fig. 41, compared with four theoretical predictions. The measured $p_T^{\ell\ell}$, $\Delta y_{\ell\ell}$, Δy_{jj} and $\cos\theta_\eta^*$ distributions are shown in Fig. 42, and several differential cross sections with respect to variables that probe the lepton and jet activities follow in Fig. 43 and Fig. 44, respectively. Generally, the compatibilities across observables with all the theoretical predictions are high. For the nominal POWHEG+PYTHIA8 prediction, the p-values are in the range 0.22-0.98. The p-values from the compatibility tests between the measured cross sections and the five theoretical predictions are summarized in Table 16.

observable	POWHEG PYTHIA8	POWHEG HERWIG7	MADGRAPH5 HERWIG7	VBFNLO at LO PYTHIA8	VBFNLO at NLO
m_{ll}	0.46	0.36	0.55	0.95	0.55
$\Delta\phi_{ll}$	0.97	0.86	1.00	0.56	1.00
m_{jj}	0.48	0.37	0.57	1.00	0.47
Δy_{jj}	0.82	0.70	0.92	0.77	0.91
$\Delta\phi_{jj}$	0.63	0.47	0.74	0.65	0.76
Δy_{ll}	0.33	0.26	0.39	0.92	0.40
$\cos\Theta^*$	0.93	0.81	0.98	0.59	0.99
Higgs p_T	0.65	0.53	0.76	0.69	0.74
p_T^{j1}	0.26	0.19	0.32	0.95	0.32
p_T^{j2}	0.98	0.90	1.00	0.59	1.00
$p_T^{\ell1}$	0.22	0.17	0.26	0.74	0.27
$p_T^{\ell2}$	0.65	0.70	0.93	0.56	0.93
p_T^{ll}	0.86	0.70	0.95	0.72	0.95

Table 16: P-values for the compatibilities between the measured cross sections and the signal predictions given by POWHEG+PYTHIA8, POWHEG+HERWIG7, MADGRAPH5_AMC@NLO+HERWIG7, VBFNLO+PYTHIA8, and fixed order calculation of VBFNLO at NLO.

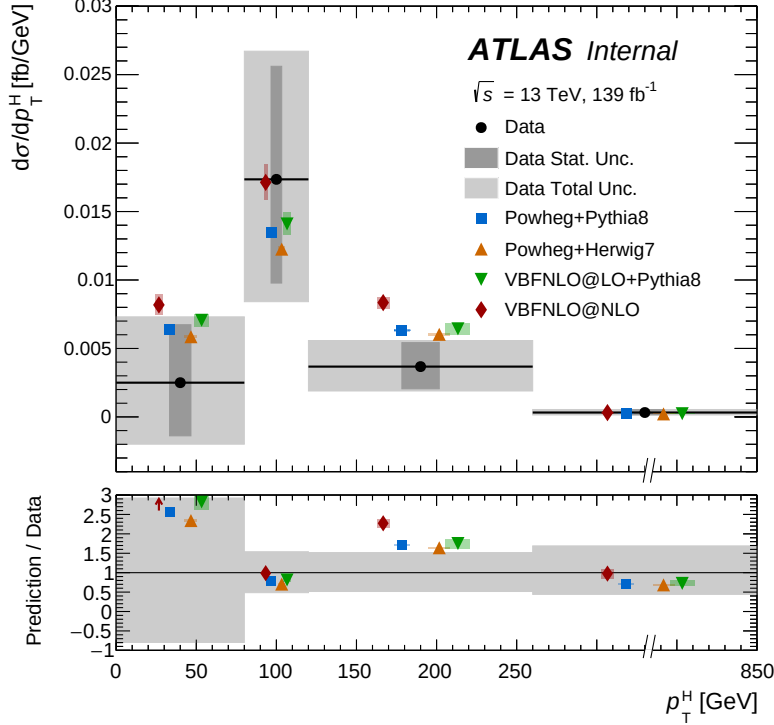


Figure 41: The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for p_T^H in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.

The relative precision of the differential measurements ranges from 40% to above 100%. The data statistical uncertainty dominates the systematics uncertainties for all the measured observable bins. The impact of each systematic source depends on the kinematic region relevant to the measurement. For example, in the measurement of m_{jj} , the jet energy resolution affects the precision at the 9 – 29% level for the second to last bins (the impact on the first bin is greater than 100% due to the significant deficiency of signal), whereas this source of uncertainty only impacts $\Delta y_{\ell\ell}$ at the 3–12% level since the measurement of a lepton distribution is less dependent on jet kinematics. The background modeling uncertainties are largest in regions where the backgrounds are most significant. The signal modeling

uncertainties become dominant in regions where the signal fraction is higher. These regions are usually at the extremity of the fiducial phase space, for example, high m_{jj} and high Higgs p_T . Hence, the modeling from different generators tend to differ and the systematic uncertainties originate from two-point comparisons dominant in these regions. In summary, the impacts from the systematic uncertainties are compatible with the expectations.

The observed differential distributions help us to investigate the origin of the signal deficit with respect to the SM predictions. Several energy spectra, such as the m_{jj} , m_{ll} , and p_T^{j1} , have deficiencies at low energies which are significant compared to the uncertainties of the measurement. A large deficit in low m_{jj} region is also observed in the $H \rightarrow WW^*$ cross-section measurement by ATLAS [78] using the Simplified Template Cross Section method [98, 99].

The measured background normalization parameters for all observables and the correlation matrices among the measured cross-section values in each bin of the distributions and the background normalization factors are shown in Appendix B. The robust estimation of background is crucial for measuring the signal yields. To verify consistency, we compare the background normalization factors across observables and find compatibility. The top plus diboson parameters in the SRs fluctuate most significantly (0.70 – 0.95) since they are estimated in different kinematic regions, some of which are more susceptible to higher order effects. The correlations among parameters are generally less than 20%, with a few exceptions up to 35% correlation. However, the reason of these high correlations are well-understood and discussed in Appendix B.

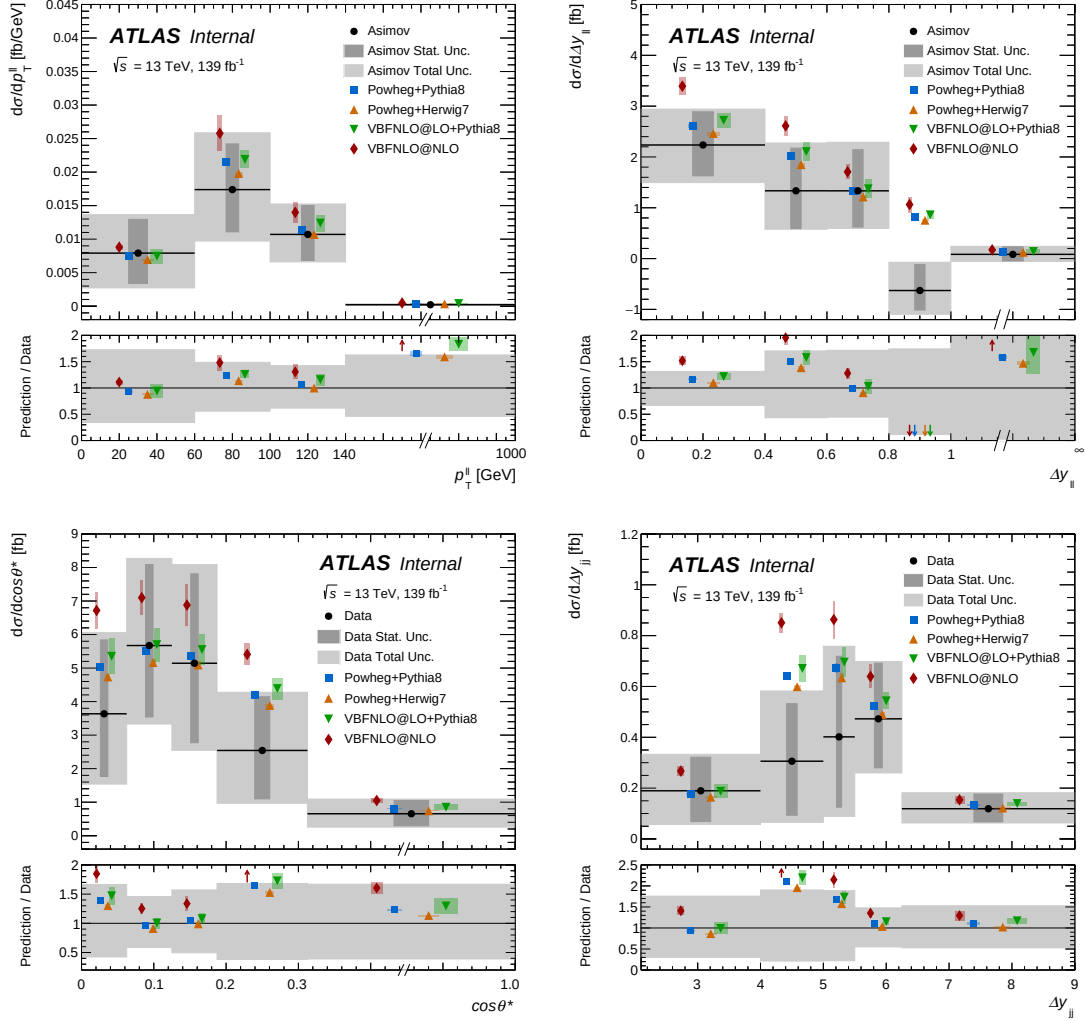


Figure 42: The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell\ell}$, $\Delta y_{\ell\ell}$, Δy_{jj} and $\cos\theta_\eta^*$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+ PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.

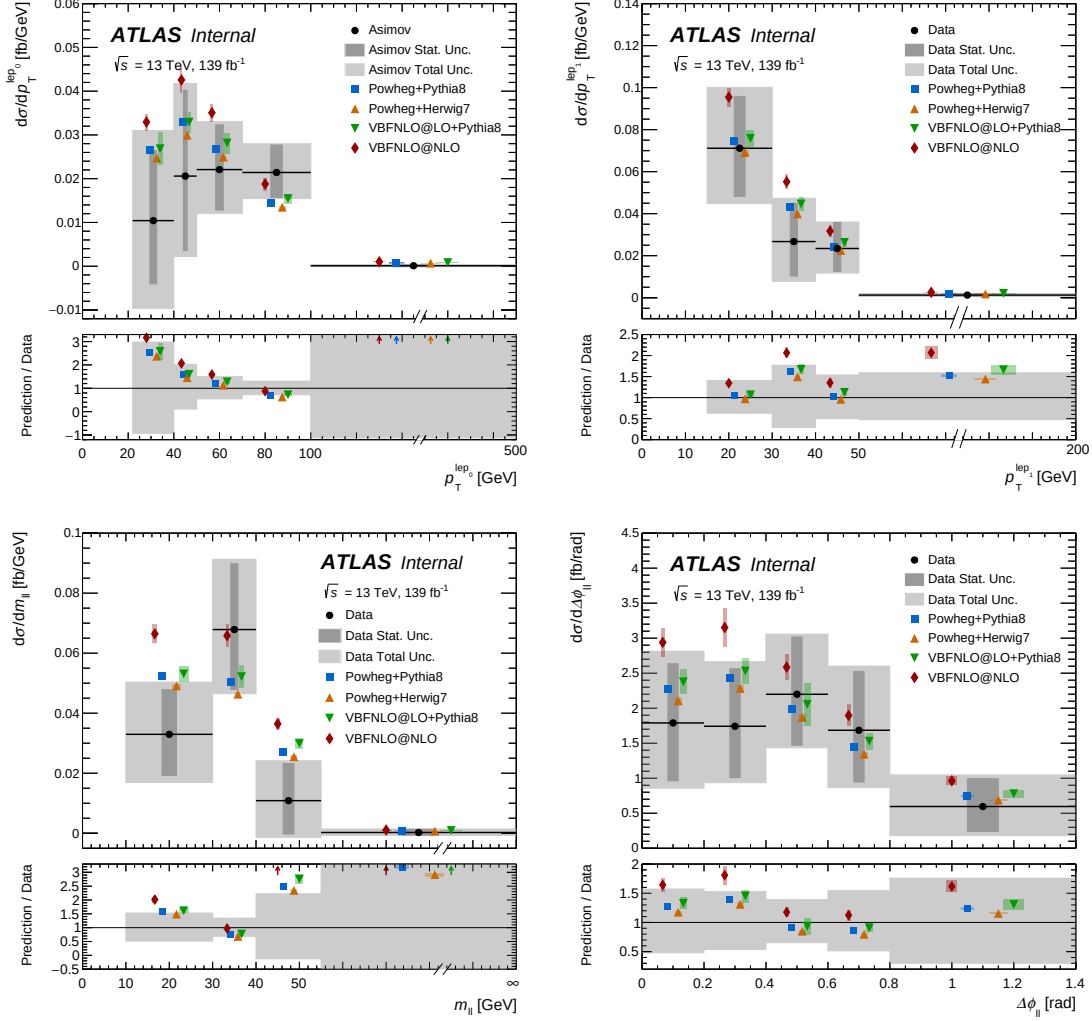


Figure 43: The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell_1}$, $p_T^{\ell_2}$, $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.

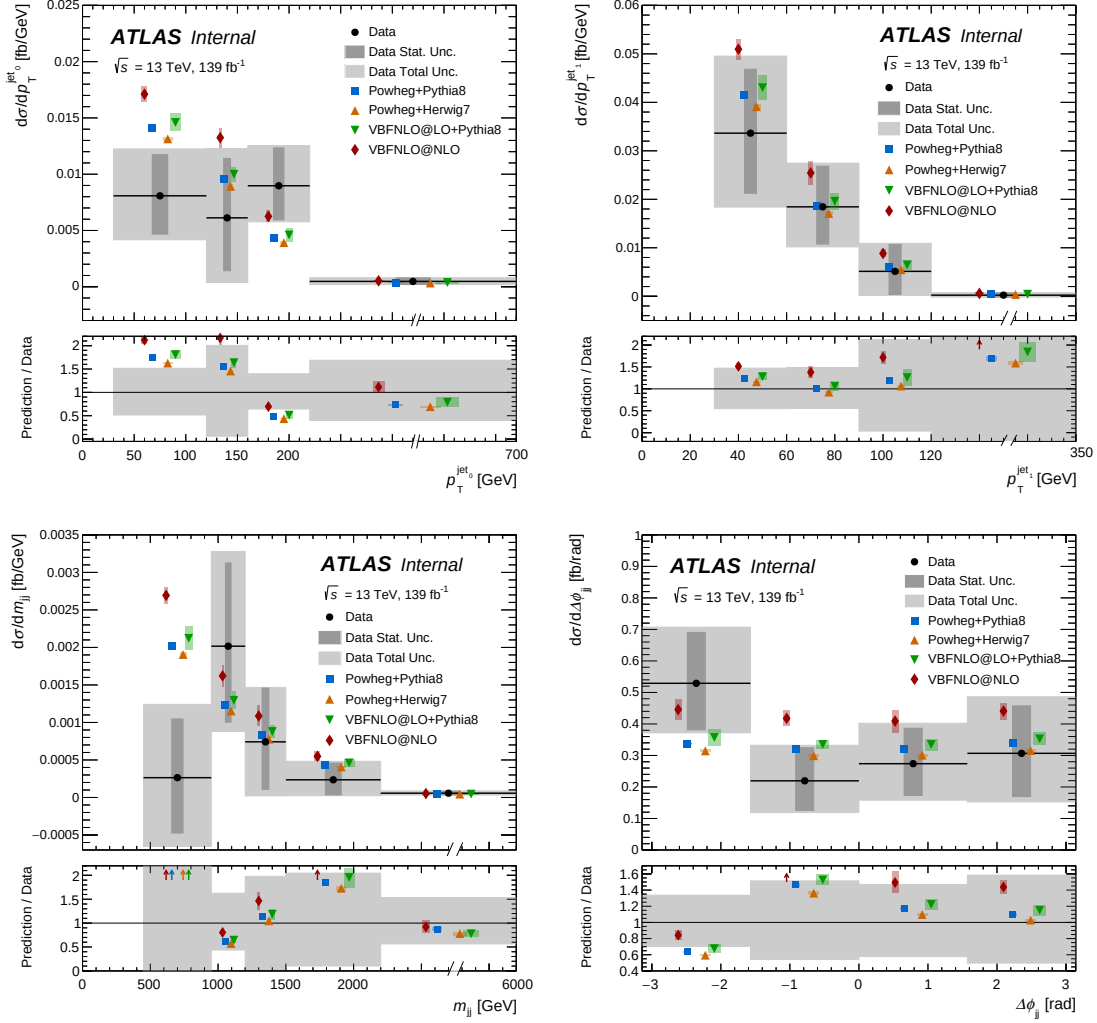


Figure 44: The measured differential fiducial cross sections at $\sqrt{s} = 13$ TeV for p_T^{j1} , p_T^{j2} , m_{jj} , and $\Delta\phi_{jj}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the measured distribution, and the gray area represents the total uncertainty of the measurement.

11 Effective Field Theory interpretation

In this section, we re-interpret the data with the Standard Model Effective Field Theory (SMEFT) formalism in the Warsaw basis to probe physics beyond the Standard Model (BSM). We consider only dimension-6 operators (O_i) that affect our signal process. The results of the reinterpretation are the limits we set on each operator's corresponding Wilson coefficient (c_i), where smaller limits suggest narrower windows for new physics.

This analysis is sensitive to seven charge-parity (CP) even operators and three CP odd operators. The operators and their composition of SM fields are given in Table 17. The six operators on the top, denoted as the c_{HW} group in the following sections, describe anomalous couplings between the gauge bosons and the Higgs boson, whereas the other four operators making up the c_{Hq3} group depict new physics at the coupling vertices of quarks and the Higgs boson.

Wilson coefficient	Operator definition
c_{HW}	$H^\dagger H W_{\mu\nu}^n W^{n\mu\nu}$
$\tilde{c}_{HW}$	$H^\dagger H \tilde{W}_{\mu\nu}^n W^{n\mu\nu}$
c_{HB}	$H^\dagger H B_{\mu\nu} B^{\mu\nu}$
$\tilde{c}_{HB}$	$H^\dagger H \tilde{B}_{\mu\nu} B^{\mu\nu}$
c_{HWB}	$H^\dagger \tau^n H W_{\mu\nu}^n B^{\mu\nu}$
$\tilde{c}_{HWB}$	$H^\dagger \tau^n H \tilde{W}_{\mu\nu}^n B^{\mu\nu}$
c_{Hq1}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{q}_p \gamma^\mu q_r)$
c_{Hq3}	$(H^\dagger i \overleftrightarrow{D}_\mu^n H)(\bar{q}_p \tau^n \gamma^\mu q_r)$
c_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{u}_p \gamma^\mu u_r)$
c_{Hd}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{d}_p \gamma^\mu d_r)$

Table 17: Summary of EFT operators in the SMEFT formalism that are probed in the VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channel [1]. The corresponding structure in terms of the SM fields in the Warsaw basis (second column) is shown together with the associated Wilson coefficient (first column). The Higgs doublet field and its complex conjugate are denoted as H and $H^\dagger$, respectively. The left-handed quark doublets of flavor p (the right-handed up-type or down-type quarks of flavour r) are denoted as q_p (u_r or d_r), while $V_{\mu\nu}$ and $\tilde{V}_{\mu\nu}$ (with $V = B, W^n$ where $n = \{1, 2, 3\}$) represent a given gauge field of the electroweak interactions, and τ^n are the Pauli matrices. The bosonic operators with (without) a dual field $\tilde{V}_{\mu\nu}$ are CP-odd (CP-even).

The Lagrangian of the SMEFT framework takes the form $\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} \mathcal{O}_i$, where Λ is the cut-off energy scale. The linear dependency of the SMEFT Lagrangian on the Wilson coefficients is carried to the matrix element calculation. The SMEFT matrix element $\mathcal{M}_{\text{SMEFT}}$ can be written as

$$\mathcal{M}_{\text{SMEFT}} = \mathcal{M}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} \mathcal{M}_i, \quad (48)$$

where $\mathcal{M}_{\text{SM}}$ is the SM matrix element and $\mathcal{M}_i$ is the matrix element of the EFT operator $\mathcal{O}_i$. As a result, the square of the matrix element becomes

$$|\mathcal{M}_{\text{SMEFT}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + 2 \sum_i \frac{c_i}{\Lambda^2} \text{Re}(\mathcal{M}_{\text{SM}}^* \mathcal{M}_i) + \sum_{i,j} \frac{c_i c_j}{\Lambda^4} \text{Re}(\mathcal{M}_i^* \mathcal{M}_j). \quad (49)$$

Since our EFT interpretation will not consider the interference between two different EFT operators, Eq. 49 can be simplified to

$$|\mathcal{M}_{\text{SMEFT}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + 2 \sum_i \frac{c_i}{\Lambda^2} \text{Re}(\mathcal{M}_{\text{SM}}^* \mathcal{M}_i) + \sum_i \frac{c_i^2}{\Lambda^4} |\mathcal{M}_i|^2. \quad (50)$$

We refer to the second term in Eq. 50 as the interference term or the linear term, because it is the interference of EFT matrix element with that of the SM and it is linearly dependent on the Wilson coefficients. The third term is called the BSM term or the quadratic term as it depends on the Wilson coefficients quadratically. A physical observable X which is proportional to the square of the matrix element such as scattering cross section or decay width, can be parameterized as a quadratic function of c_i [100]

$$X(c_i) = X_{\text{SM}} + X_{\text{int.}} + X_{\text{BSM}} = X_{\text{SM}} \cdot [1 + \sum_i (A_i^X c_i + B_i^X c_i^2)]. \quad (51)$$

This dependency on the Wilson coefficients is verified in the EFT simulation which we discuss in the next section.

11.1 EFT simulation and parametrization strategy

The fiducial cross section for the n^{th} observable bin can be written as the product of the VBF production cross section (σ^{VBF}), the $H \rightarrow l\nu l\nu$ branching ratio, which is the partial width ($\Gamma^{H \rightarrow l\nu l\nu}$) divided by the total width (Γ^H), and the acceptance factor P_n

$$\sigma_n^{\text{fid.}, \Gamma} = \sigma^{\text{VBF}} \cdot \frac{\Gamma^{H \rightarrow l\nu l\nu}}{\Gamma^H} \cdot P_n. \quad (52)$$

While the production cross section and Higgs decay widths are linearly and quadratically dependent on the Wilson coefficients by definition (Eq. 51), there is no theoretical reason to expect the acceptance factors to be related to the Wilson coefficients in the same way. For example, some nontrivial relations between acceptance and the Wilson coefficients were observed in the EFT interpretation of the $H \rightarrow ZZ^* \rightarrow 4l$ channel [101]. Therefore, the EFT strategy of this analysis is developed to parametrize the cross sections as a function of the Wilson coefficient $\sigma_n^{\text{fid.}}(c_i)$ while incorporating the acceptance effect and omitting terms of order higher than c_i^2/Λ^4 . The strategy consists of two components:

- combined parametrization for the production cross section σ^{VBF} , the partial width $\Gamma^{H \rightarrow l\nu l\nu}$, and the acceptance P_n by generating fiducial EFT samples and performing a parametrization fit for each Wilson coefficient;
- parametrization of the total width of the Higgs boson Γ^H incorporated after the small value expansion around c_i/Λ^2 .

The former strategy is detailed in Sec. 11.1.1, and the latter and the combination of the two parametrizations are discussed in Sec. 11.1.2. Later in this section, we elaborate the EFT fit strategy to constrain each Wilson coefficient. We report the limits using the linear-only parametrization, where all c_i^2/Λ^4 terms are dropped from the parametrization functions, and also the limits using the complete linear plus quadratic parametrizations. By comparing the two sets of results, we can probe the higher order effects from EFT.

11.1.1 EFT fiducial parametrization

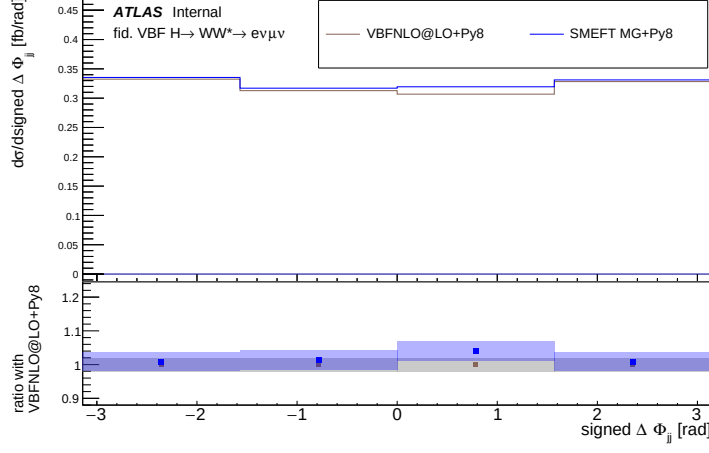
In this section, we discuss the strategy to derive the *fiducial parametrization functions* which describe the EFT impact on the VBF production cross section, the $H \rightarrow \ell\nu\ell\nu$ partial width, and the acceptance factors. These functions are denoted as $\sigma_n^{\text{fid.}}(c_i)$ for the n^{th} observable bin. Since the relation between acceptance factors and the Wilson coefficients is unknown, multiple EFT fiducial samples are generated at different values of each Wilson coefficient. By fitting the cross sections as a function of each Wilson coefficient, we obtain the fiducial parametrization functions.

The EFT effect on the full VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ process, excluding the effect on the total width of the Higgs boson, is simulated by MADGRAPH5 with the SMEFTsim package [102] at leading order in QCD and interfaced with PYTHIA 8. According to Eq. 51, the total SMEFT cross section is the sum of the SM cross section σ_{SM} and a EFT corrective cross section $\Delta\sigma_{\text{EFT}}$

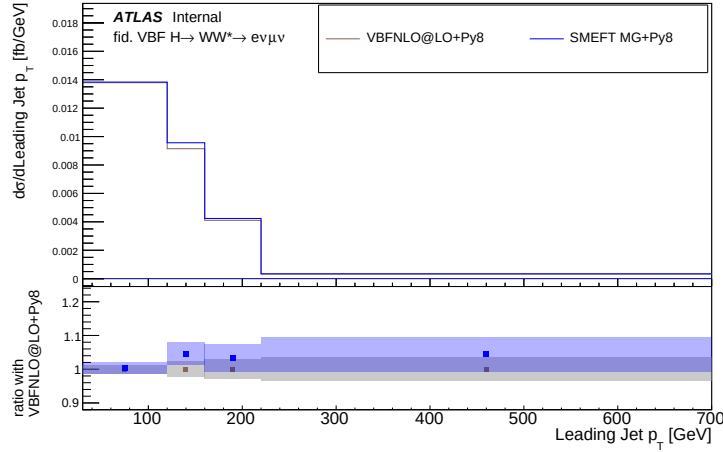
$$\sigma_{\text{SMEFT}} = \sigma_{\text{SM}} + \Delta\sigma_{\text{EFT}}. \quad (53)$$

The SM and EFT cross sections are obtained from the SM and EFT samples produced by MADGRAPH5+PYTHIA 8. The relative EFT correction $\Delta\sigma_{\text{EFT}}/\sigma_{\text{SM}}$ is assumed to be independent of QCD or EW order and hence can be applied to higher order signal simulations. Next, the NLO parton distribution function (PDF), NNPDF30_nlo_as_0118, and the G_F EW scheme are used to be consistent with other higher order SM simulations. The EFT samples are generated in the inclusive phase space of VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ and intermediate tau leptons decays are simulated by PYTHIA 8. The cut-off energy scale Λ is fixed at 1 TeV for all the EFT simulations. Each EFT sample is generated with one and only one Wilson coefficient set to be nonzero. Moreover, the total width of the Higgs boson in the simulations is fixed to the SM value since the total width calculated at LO deviates significantly from the observed value. The SM sample generated by MADGRAPH5+PYTHIA 8 after the fiducial selections is validated by the VBFNLO+PYTHIA 8 sample introduced in Sec. 8.2. The

validation of two EFT-sensitive observables, $\Delta\phi_{jj}$ and $p_T^{j_1}$, are shown in Fig. 45. The two LO samples agree within their MC statistical uncertainties.



(a) $\Delta\phi_{jj}$



(b) $p_T^{j_1}$

Figure 45: Validation of the SM prediction simulated by MADGRAPH5 with the SMEFTsim package at QCD LO. The sample (blue) is showered with VBFNLO+PYTHIA 8 and is validated against VBFNLO+PYTHIA 8 (gray). The bottom panel shows the ratio with respect to the VBFNLO+ PYTHIA 8 prediction. The gray (blue) error band represents the MC statistical uncertainties of the VBFNLO(MADGRAPH5) sample.

The EFT simulations generate separately the interference samples and the BSM samples, where the former samples are required to have cross sections linearly dependent on the Wilson

coefficients and the latter proportional to the square of each Wilson coefficient. In total, seven interference and seven BSM samples per Wilson coefficient are produced at values: $c_{\text{gen.}} = -0.8, -0.5, -0.2, 0.1, 0.4, 0.6, 1.0$. The SM sample is generated with all the Wilson coefficients fixed to zero. Finally, only EFT events in the fiducial phase space are selected and projected into the binning given in Table 13 via a Rivet routine [103].

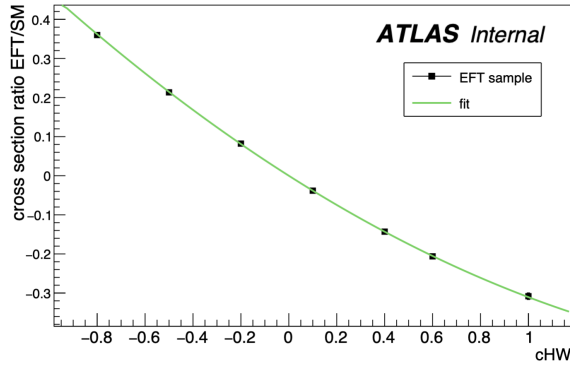
As a result of the parametrization fit using the SM and EFT fiducial samples, a linear plus quadratic function turns out to be an appropriate description for the fiducial cross sections:

$$\begin{aligned}\sigma_n^{\text{fid.}}(c_i) &= \sigma^{\text{VBF}}(c_i) \cdot \frac{\Gamma^{H \rightarrow l\nu l\nu}(c_i)}{\Gamma_{\text{SM}}^H} \cdot P_n(c_i) \\ &= \sigma_{\text{SM},n}^{\text{fid.}} \cdot (1 + A_{i,n}^{\text{fid.}} c_i + B_{i,n}^{\text{fid.}} c_i^2) + \mathcal{O}(1/\Lambda^6).\end{aligned}\tag{54}$$

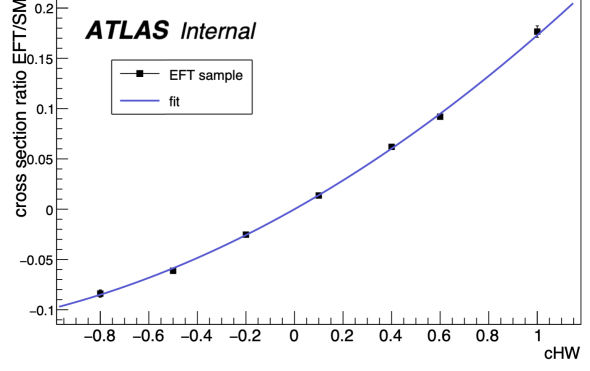
While we cannot separate the three sources of EFT impact, the combined parametrization serves the purpose for this analysis.

The two groups of Wilson coefficients, the c_{HW} group and the c_{Hq3} group, affect different kinematic observables. The observables that are most sensitive to the c_{HW} and c_{Hq3} groups are $\Delta\phi_{jj}$ and the leading jet p_T , respectively. We perform the parametrization fits for each observable bin as illustrated in Fig. 46 for c_{HW} . The seven data points are the ratios of the EFT and SM cross sections per bin (i.e. the relative EFT corrections) obtained from the EFT fiducial samples generated at the seven $c_{\text{gen.}}$ values. The error bars originate from the statistical uncertainties from the SM and EFT samples. The curves are the minimum- χ^2 fit to the parametrization functions $A_{i,n}^{\text{fid.}} c_i + B_{i,n}^{\text{fid.}} c_i^2$. The fitted values for the fiducial parametrization coefficients ($A_{i,n}^{\text{fid.}}$ and $B_{i,n}^{\text{fid.}}$) and χ^2 divided by the numbers of degrees of freedom (χ^2/DoF) for each bin of the respective sensitive observables are summarized in Table 18 for the c_{HW} group and Table 19 for the c_{Hq3} group.

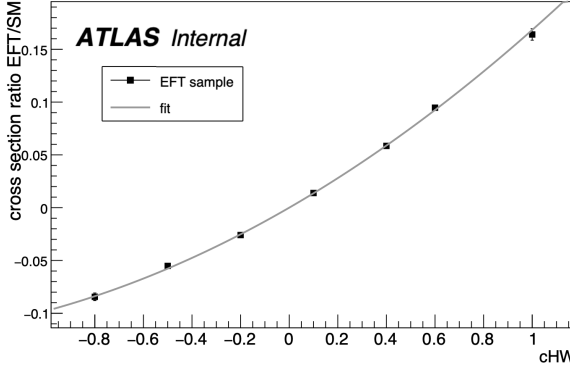
Except for c_{Hu} and $\tilde{c}_{HWB}$, all the parametrization fits converge at approximately $\chi^2/\text{DoF} = 1.0$, proving the linear and quadratic fiducial parametrization is successful in describing the EFT correction on SM distributions. The large χ^2/DoF values of c_{Hu} and $\tilde{c}_{HWB}$ suggest there are functions more suitable to describe their impact in our fiducial region. It is not



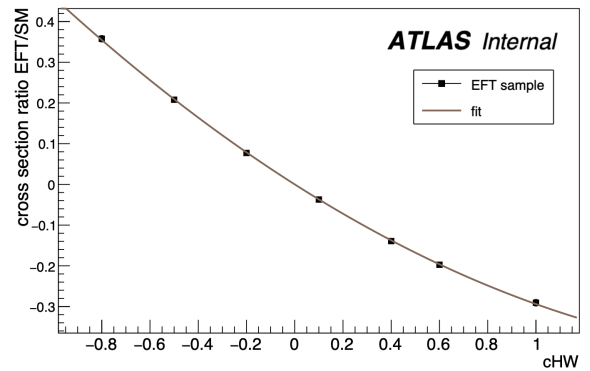
(a) $\Delta\phi_{jj}$ bin1 ($-\pi - -\pi/2$).



(b) $\Delta\phi_{jj}$ bin2 ($-\pi/2 - 0$).



(c) $\Delta\phi_{jj}$ bin3 ($0 - \pi/2$).



(d) $\Delta\phi_{jj}$ bin4 ($\pi/2 - \pi$).

Figure 46: EFT parametrization fit (curve) of c_{HW} in the four $\Delta\phi_{jj}$ bins in comparison with the relative EFT corrections obtained from the EFT fiducial samples (data points). The relative EFT corrections are the ratio of the EFT and SM cross sections simulated at $c_{HW} = -0.8, -0.5, -0.2, 0.1, 0.4, 0.6, 1.0$. The error bars represent the MC statistical uncertainty and are used to perform the parametrization fit.

unexpected that the EFT affects acceptance in a nontrivial way, since the fiducial selections are not independent of each other. For instance, the requirements on the lepton p_T and on the invariant mass of the di-lepton system (m_{ll}) are correlated, i.e. rejecting events with $m_{ll} < 10$ GeV is implicitly scraping away part of the phase space on the lepton p_T scale. Thus, the acceptance of the EFT events reflects such nonlinearity from the fiducial selections. However, we decide to keep the linear plus quadratic form, but propagate the uncertainties arising from choosing a suboptimal parametrization function to the final limits. We inflate the MC statistical uncertainties given by the c_{Hu} and $\tilde{c}_{HWB}$ samples until χ^2/DoF reaches unity. This strategy does not change the values of the parametrization coefficients $A_{i,n}^{\text{fid.}}$ and

Wilson Coefficient	bin	fiducial parametrization	χ^2/DoF
c_{HWB}	1	$-0.0101c + 0.003c^2$	1.141
	2	$0.049c + 0.005c^2$	0.316
	3	$0.048c + 0.004c^2$	0.674
	4	$-0.010c + 0.003c^2$	0.417
c_{HB}	1	$-0.018c + 0.001c^2$	0.356
	2	$0.014c + 0.001c^2$	0.822
	3	$0.014c + 0.00c^2$	0.641
	4	$-0.017c + 0.001c^2$	0.582
c_{HW}	1	$-0.389c + 0.079c^2$	0.123
	2	$0.136c + 0.037c^2$	0.829
	3	$0.133c + 0.035c^2$	0.662
	4	$-0.377c + 0.083c^2$	0.336
$\tilde{c}_{HW}$	1	$-0.286c + 0.059c^2$	0.642
	2	$-0.265c + 0.058c^2$	1.781
	3	$0.264c + 0.047c^2$	0.750
	4	$0.278c + 0.052c^2$	1.184
$\tilde{c}_{HWB}$	1	$-0.032c + 0.001c^2$	7.491
	2	$-0.029c + 0.001c^2$	8.144
	3	$0.029c + 0.003c^2$	8.030
	4	$0.031c + 0.004c^2$	9.900
$\tilde{c}_{HB}$	1	$-0.017c + 0.001c^2$	0.641
	2	$-0.016c + 0.001c^2$	1.082
	3	$0.015c + 0.000c^2$	0.559
	4	$0.016c + 0.000c^2$	0.137

Table 18: Fiducial EFT parametrization of c_{HW} , c_{HWB} , c_{HB} and their CP odd counterparts for the differential bins of $\Delta\phi_{jj}$.

$B_{i,n}^{\text{fid.}}$, but it changes their uncertainties. Later, we find this error inflation has negligible effect on the limits we set on these two Wilson coefficients.

To make the parametrization fit strategy more robust, the uncertainties on $A_{i,n}^{\text{fid.}}$ and $B_{i,n}^{\text{fid.}}$ are incorporated as nuisance parameters in the final EFT fit and are constrained by Gaussian functions. In addition, correlations of $A_{i,n}^{\text{fid.}}$ and $B_{i,n}^{\text{fid.}}$ are taken into account. These EFT-related nuisance parameters have small impact on the limits we set on the Wilson coefficients.

The relative EFT corrections given by the parametrization functions are shown in Fig. 47 and Fig. 48 for the CP even and odd Wilson coefficients in the c_{HW} group, respectively. The

Wilson Coefficient	bin	fiducial parametrization	χ^2/DoF
c_{Hq3}	1	$-0.278c + 0.023c^2$	0.429
	2	$-0.643c + 0.128c^2$	0.306
	3	$-1.010c + 0.317c^2$	1.661
	4	$-2.511c + 2.232c^2$	0.152
c_{Hq1}	1	$0.005c + 0.004c^2$	2.852
	2	$0.016c + 0.021c^2$	1.460
	3	$0.023c + 0.058c^2$	0.943
	4	$0.109c + 0.596c^2$	0.811
c_{Hd}	1	$0.006c + 0.002c^2$	1.130
	2	$0.014c + 0.009c^2$	0.673
	3	$0.023c + 0.026c^2$	0.225
	4	$0.057c + 0.218c^2$	0.103
c_{Hu}	1	$-0.015c + 0.003c^2$	49.812
	2	$-0.039c + 0.013c^2$	24.625
	3	$-0.064c + 0.038c^2$	21.809
	4	$-0.165c + 0.344c^2$	19.197

Table 19: EFT parametrization of c_{Hq3} , c_{Hq1} , c_{Hu} and c_{Hd} for the differential bins of p_T^{j1} .

impact from CP even operators are symmetric around $\Delta\phi_{jj} = 0$, whereas CP odd operators modify the SM prediction asymmetrically. The Wilson coefficients are fixed at values close to their expected $+1\sigma$ limits. The solid blue lines show the EFT impact at order of c_i/Λ^2 . Since the interference between the SM and EFT matrix elements can be destructive, the EFT impact can be negative in part of the phase space. On the other hand, the blue dashed lines illustrate the quadratic impact which is always positive as it originates from the absolute square of the EFT matrix elements. The red dashed lines are the sum of linear and quadratic corrections. While the effects at the first order of c_i/Λ^2 are expected to be more dominant, the second order effect is significant for several Wilson coefficients despite being suppressed by $1/\Lambda^4$. Fig. 49 shows the EFT impact by c_{Hq3} , c_{Hu} , c_{Hq1} and c_{Hd} . Their impacts are more significant in the high p_T^{j1} and high $p_T^{\ell_1}$ regime.

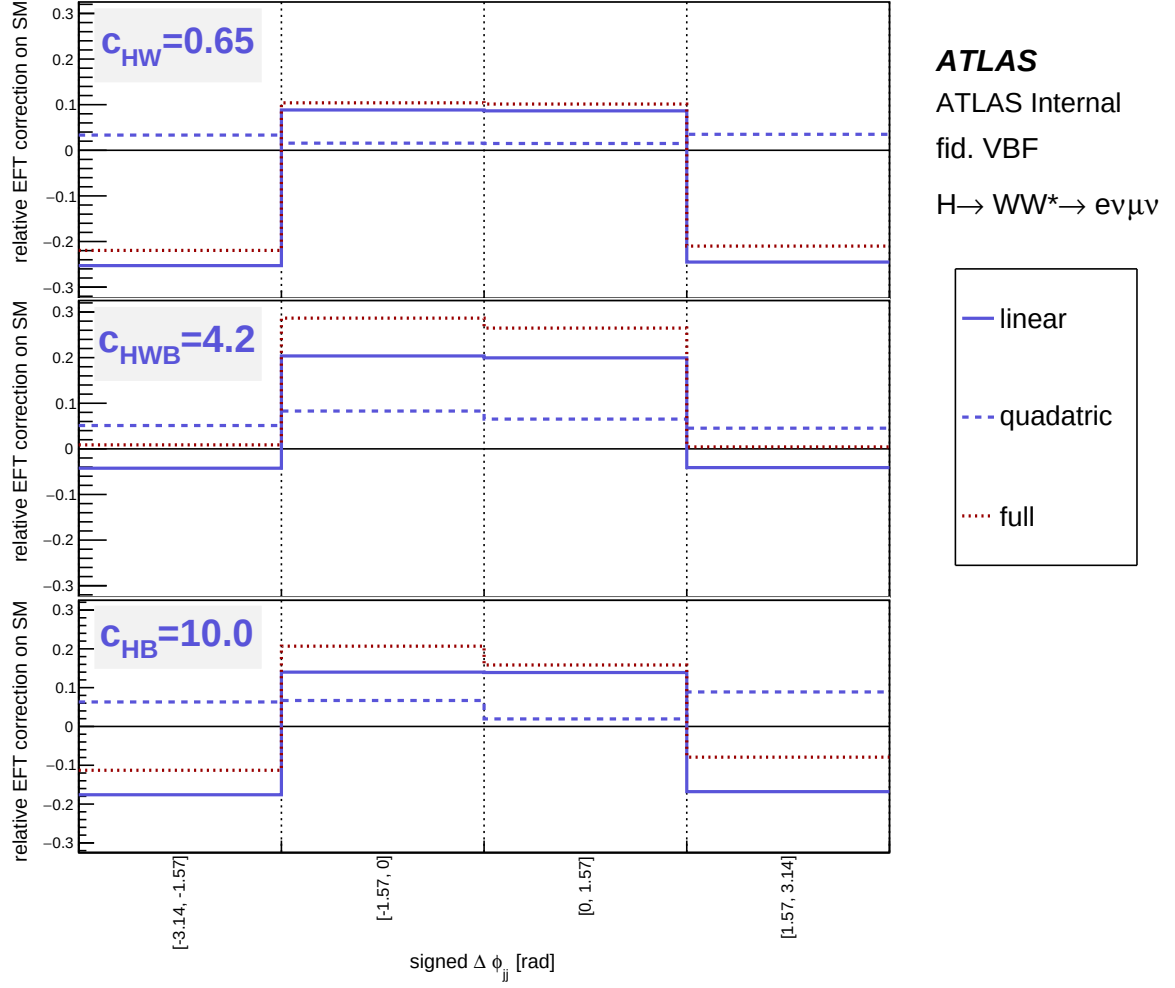


Figure 47: The relative EFT correction on $\Delta\phi_{jj}$ by the CP even c_{HW} , c_{HWB} , and c_{HB} at fixed values close to the expected $+1\sigma$ limits. The solid blue lines show the EFT corrections at order of c_i/Λ^2 (linear); the blue dashed lines show the corrections at order of c_i^2/Λ^4 (quadratic); finally the red dashed lines show the full linear plus quadratic corrections.

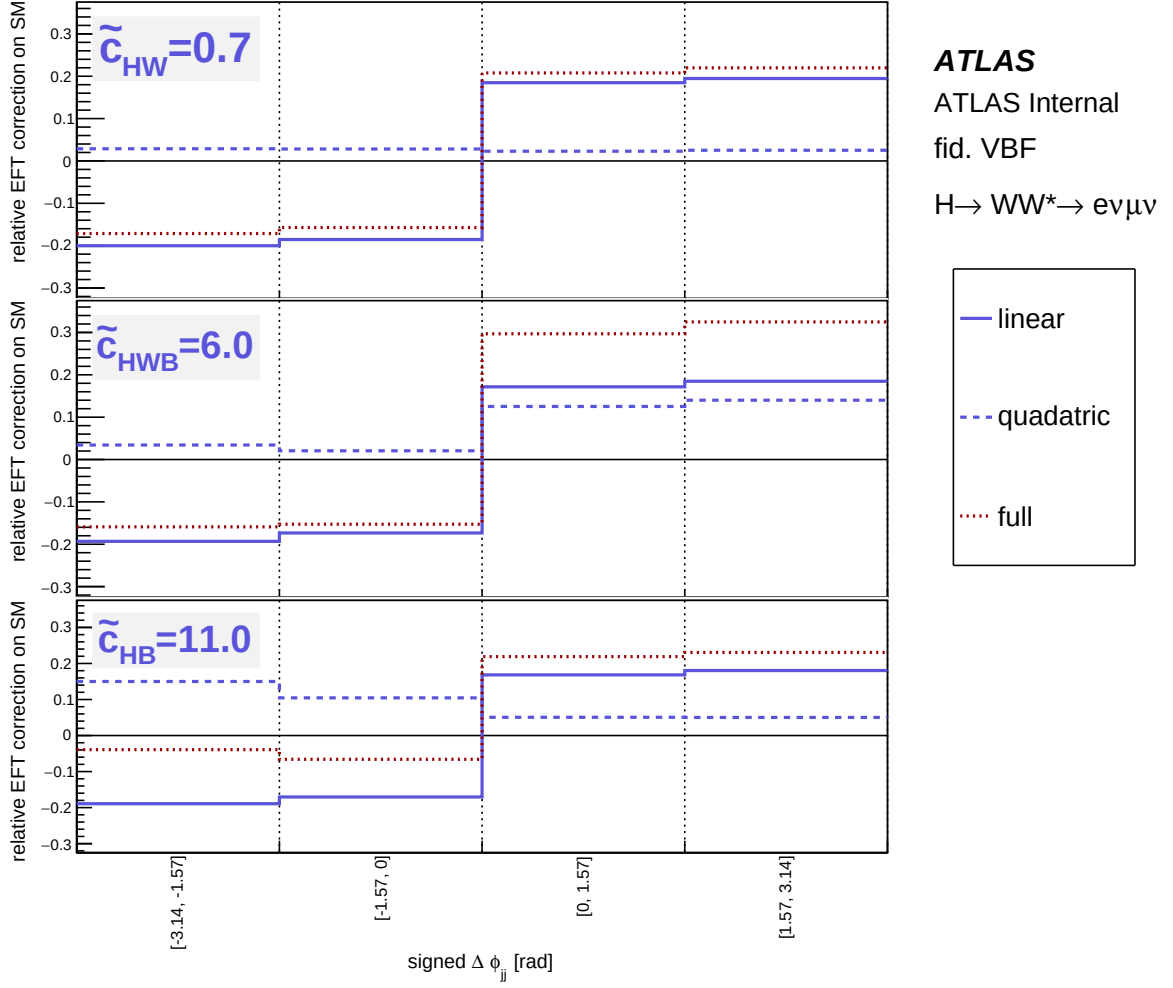


Figure 48: The relative EFT correction on $\Delta\phi_{jj}$ by the CP odd c_{HW} , c_{HWB} , and c_{HB} at fixed values close to the expected $+1\sigma$ limits. The solid blue lines show the EFT corrections at order of c_i/Λ^2 (linear); the blue dashed lines show the corrections at order of c_i^2/Λ^4 (quadratic); finally the red dashed lines show the full linear plus quadratic corrections.

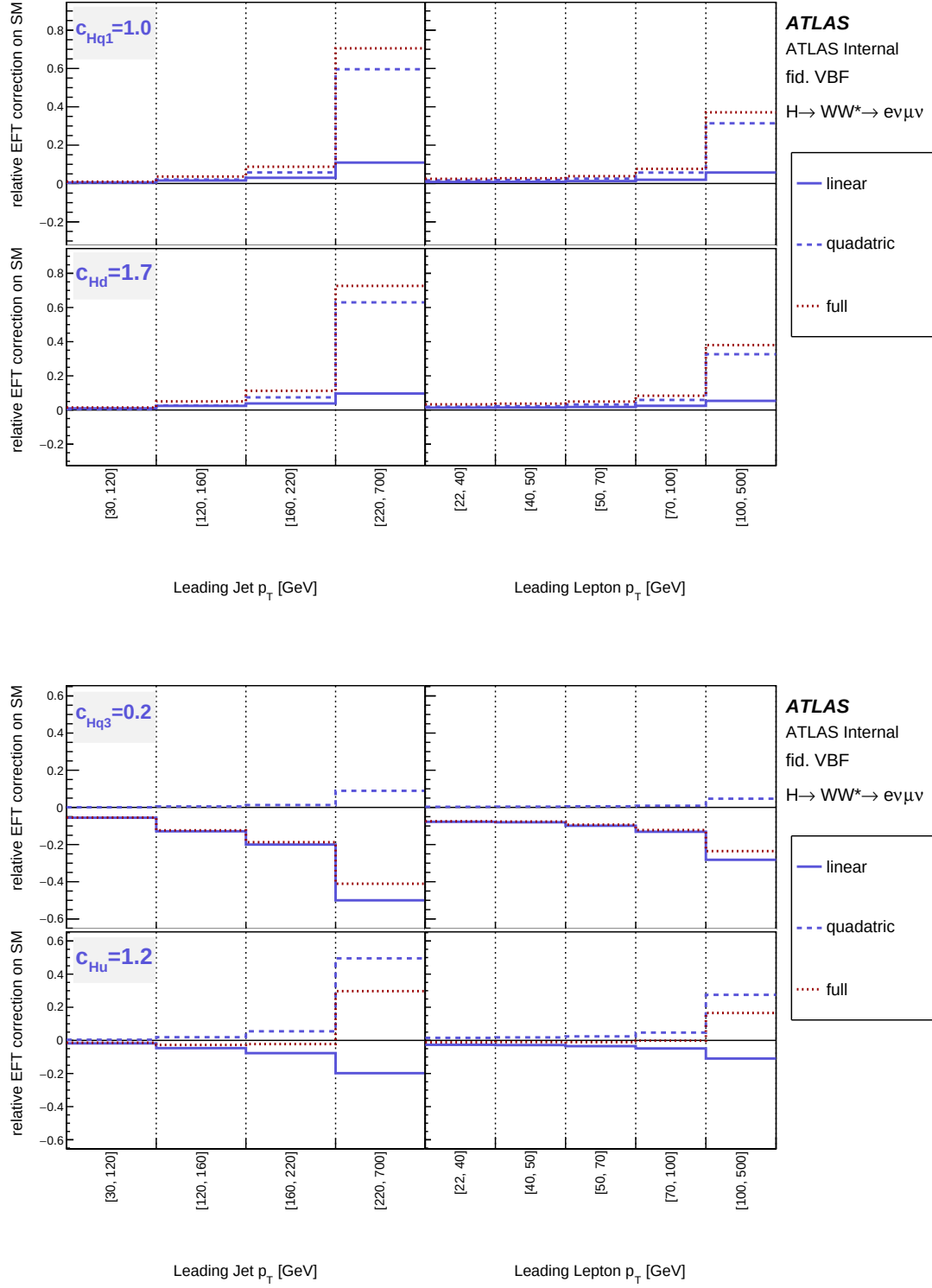


Figure 49: The relative EFT correction on $p_T^{j_1}$ (left panel) and $p_T^{\ell_1}$ (right panel) by c_{Hq1} , c_{Hd} (top), c_{Hq3} and c_{Hu} (bottom) at fixed values close to the expected $+1\sigma$ limits.

11.1.2 Higgs boson width parametrization

The fiducial parametrizations obtained through the above procedures do not include the EFT impact on the total width of the Higgs boson. We use the Higgs-width parametrization reported in the EFT interpretation of the $H \rightarrow ZZ$ channel [101]. Only c_{HW} , c_{HWB} and c_{HB} have significant impact on the total width of the Higgs boson Γ^H :

$$\Gamma^H(c_i) = \Gamma_{\text{SM}}^H \cdot f_{\Gamma_H}(c_i); \quad (55)$$

$$f_{\Gamma_H}(c_i) = 1 - 0.05c_{HW} + 0.16c_{HW}^2 - 0.08c_{HB} + 1.21c_{HB}^2 + 0.06c_{HWB} + 0.36c_{HWB}^2, \quad (56)$$

where $\Gamma^H(c_i)$ is calculated by summing the parametrizations of all dominant Higgs decay channels simulated by MADGRAPH5. Since the Higgs width is a CP even observable, the CP odd operators only have impact at the order of $1/\Lambda^4$ [104]. The quadratic parametrization coefficients are taken to be the same as the CP even parametrization and included in the final EFT fit. Comparing the Higgs-width parametrization given in Eq. 56 and the fiducial parametrization functions in Table 18, we can see the quadratic EFT effect by c_{HB} and c_{HWB} as well as their CP odd partners are many orders greater on the total width. The primary decay mode affected by these EFT operators is the $H \rightarrow \gamma\gamma$ process [105].

Incorporating the EFT impact on Γ^H , the parametrization takes the explicit form:

$$\sigma_n^{\text{fid.},\Gamma}(c_i) = \frac{\sigma_n^{\text{fid.}}(c_i)}{f_{\Gamma_H}(c_i)} = \sigma_n^{\text{fid.},\text{SM}} \frac{1 + A_{i,n}^{\text{fid.}}c_i + B_{i,n}^{\text{fid.}}c_i^2}{1 + A_i^\Gamma c_i + B_i^\Gamma c_i^2}. \quad (57)$$

The final parametrization used in the EFT fit is the Taylor expansion of Eq. 57, assuming $c_i/\Lambda^2 \ll 1$, to the first order for the linear parametrization and to the second order for the linear plus quadratic parametrization:

$$\sigma_n^{\text{fid.},\Gamma}(c_i) = \sigma_{\text{SM},n}^{\text{fid.}}[1 + (A_{i,n}^{\text{fid.}} - A_i^\Gamma)c_i]; \quad (58)$$

$$\sigma_n^{\text{fid.},\Gamma}(c_i) = \sigma_{\text{SM},n}^{\text{fid.}}[1 + (A_{i,n}^{\text{fid.}} - A_i^\Gamma)c_i + (B_{i,n}^{\text{fid.}} - B_i^\Gamma + A_i^{\Gamma^2} - A_{i,n}^{\text{fid.}} \cdot A_i^\Gamma)c_i^2]. \quad (59)$$

In the limits we set for c_{HW} , c_{HWB} and c_{HB} , while the effect is small on the limits of c_{HW} , the limits we set on c_{HWB} and c_{HB} significantly improved by incorporating EFT corrections on the total width of the Higgs boson.

11.2 EFT sensitivity tests

The EFT limits are obtained by minimizing the negative-log-likelihood where the likelihood function is taken from the differential measurement in Eq. 44, with the differential cross sections $\bar{\sigma}^{\text{fid.}}$ replaced by the parametrization functions $\bar{\sigma}^{\text{fid.},\Gamma}(c_i)$ and with the addition of Gaussian-constraint functions for the EFT nuisance parameters. The SM cross sections $\sigma_{\text{SM},n}^{\text{fid.}}$ are taken from the particle-level POWHEG+PYTHIA 8 sample simulated at NLO in QCD and EW.

The EFT sensitivity is tested by fitting to the SM Asimov dataset. Considering both the linearized expanded parametrization (hereafter referred to as linear) and the linear plus quadratic parameterization of each Wilson coefficient, Table 20 reports the expected 2σ limits for each Wilson coefficient using its optimal observable. The Asimov data statistical uncertainties and all the systematic uncertainties are included in the fit. Since the SM cross sections in the parametrization functions are fixed to the nominal prediction, the best-fit values for all Wilson coefficients are zero.

For the c_{Hq3} group, we also tested the sensitivity using $p_T^{\ell_1}$ but found less stringent constraints compared to the limits set by $p_T^{j_1}$. The expected limits for the c_{Hq3} group using $p_T^{\ell_1}$ is reported in Appendix. C.

A few observations can be made from the EFT sensitivity test:

- the analysis has good sensitivity (limits between $[-1,1]$) to c_{Hq3} to the first order using the linearized parametrization;
- the analysis has good sensitivity to c_{Hq3} , c_{HB} and $\tilde{c}_{HB}$ to the second order using the

Parameter	Optimal observable	Best-fit value [TeV ⁻²]	Linear 2σ limits [TeV ⁻²]	Linear + Quadratic 2σ limits [TeV ⁻²]
c_{HW}/Λ^2	$\Delta\phi_{jj}$	0.00	$[-1.68, 1.59]$	$[-1.39, 1.38]$
$\tilde{c}_{HW}/\Lambda^2$	$\Delta\phi_{jj}$	0.00	$[-1.65, 1.66]$	$[-1.40, 1.37]$
c_{HWB}/Λ^2	$\Delta\phi_{jj}$	0.00	$[-10.46, 9.13]$	$[-1.21, 1.10]$
$\tilde{c}_{HWB}/\Lambda^2$	$\Delta\phi_{jj}$	0.00	$[-14.94, 14.99]$	$[-1.15, 1.14]$
c_{HB}/Λ^2	$\Delta\phi_{jj}$	0.00	$[-5.85, 6.40]$	$[-0.59, 0.66]$
$\tilde{c}_{HB}/\Lambda^2$	$\Delta\phi_{jj}$	0.00	$[-28.17, 28.23]$	$[-0.62, 0.62]$
c_{Hq3}/Λ^2	p_T^{j1}	0.00	$[-0.56, 0.47]$	$[-0.43, 1.20]$
c_{Hq1}/Λ^2	p_T^{j1}	0.00	$[-11.92, 14.78]$	$[-1.85, 1.65]$
c_{Hu}/Λ^2	p_T^{j1}	0.00	$[-8.31, 6.94]$	$[-2.03, 2.57]$
c_{Hd}/Λ^2	p_T^{j1}	0.00	$[-20.90, 24.82]$	$[-3.03, 2.73]$

Table 20: Best-fit values to the SM Asimov dataset, and the 2σ intervals obtained from profile likelihood scans of the Wilson coefficients under linear-only and linear plus quadratic parametrization.

linear plus quadratic parametrization;

- the effects from the quadratic terms are significant for c_{HWB} , $\tilde{c}_{HWB}$, c_{HB} , $\tilde{c}_{HB}$, c_{Hq1} , c_{Hu} , and c_{Hd} , as the changes in the limits are greater than 50%.

The significant quadratic effects raise a common question of whether the effective theory violates the unitarity bound as the new physics contribution should not grow significantly with energy scale. However, the unitarity bound should be considered irrelevant in an EFT application [26, 106], as EFT stems from the assumption that the new physics scale is much greater than the energy scale of the processes we measure, and hence we can make the small value expansion for the first and second order around c_i/Λ^2 and omit higher orders. Furthermore, both c_i and Λ are treated as free parameters in this analysis. While the EFT samples are generated at $\Lambda = 1$ TeV, the final limits are set on c_i/Λ^2 .

The limits of c_{HB} and c_{HWB} in both parametrizations are driven by the EFT impact on the total width of the Higgs boson, in particular the partial width of other decay channels. For c_{HWB} (c_{HB}), the limit set with linear plus quadratic parametrization is improved by approximately 86% (96%). On the other hand, the impact on c_{HW} limits is insignificant with the maximal difference at a 20% level in the second order before and after including the

Higgs-width parametrization. Overall, the limits set on c_{HW} are predominantly driven by the fiducial parametrization. This is because the corresponding operator of c_{HW} describes anomalous coupling of the Higgs boson and the vector bosons and our process has two of such vertices. For the CP odd operators, since they only have impact on the Higgs width in the second order, limits set on $\tilde{c}_{HWB}$ and $\tilde{c}_{HB}$ after including the quadratic terms are significantly more stringent compared to the linear-only limits.

The nuisance parameters of the parametrization coefficients have negligible effect on the final limits. The inflation of the MC statistical uncertainties in the fiducial parametrization fits for c_{Hu} and $\tilde{c}_{HW}$ is also found to have sub-percentage effect on their limits.

One potential method to improve the EFT sensitivity is to modify the observable binning. While the nominal binning for $\Delta\phi_{jj}$ is four equal-sized bins between $-\pi$ and π , we have tested using an eight-bin scheme for the c_{HW} group. The EFT correction by $\tilde{c}_{HW}$ for the two binning schemes are compared in Fig. 50. The highest impact for $\tilde{c}_{HW} = 0.7$ is at a $\pm 30\%$ level for the finer binning and at a $\pm 20\%$ level for the coarse binning. However, the statistical uncertainties for the differential $\Delta\phi_{jj}$ measurement are at a 40% level with the four-bin scheme, which is 30% more precise than expected for the eight-bin scheme. Therefore, the EFT limits set by the two binning schemes are in agreement at percentage level. Finally, the four-bin scheme is chosen with a trade-off of averaging out the high EFT impact region below and above $\Delta\phi_{jj} = 0$.

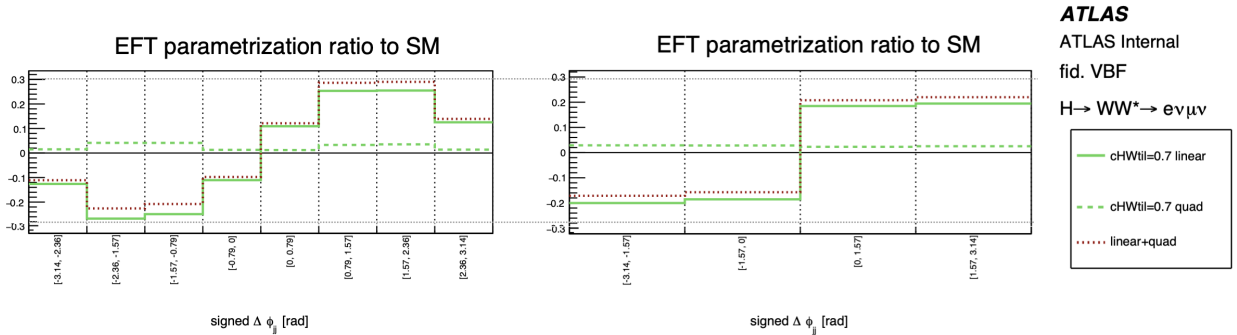


Figure 50: Relative EFT correction by $\tilde{c}_{HW}$ on the $\Delta\phi_{jj}$ distribution with eight equal-sized bins (left) and with four equal sized bins (right).

Finally, the expected EFT constraints given by the sensitivity tests are used to study the potential BSM bias when using the boosted decision tree (BDT) techniques which are trained based on SM simulated events. We show the BSM bias tests results in Appendix D. The BSM test is conducted conservatively with the Wilson coefficients chosen to be much greater than expected limits. As a result, the bias is contained within 5-10%, which is insignificant compared to the expected statistical uncertainties.

11.3 EFT results

This section presents the EFT interpretation results evaluated from the ATLAS Run-2 dataset. We use the same negative-log-likelihood minimization strategy as discussed in Sec. 11.2. The linear-only and linear plus quadratic results are summarized in Table 21. Fig. 51 presents the measured differential cross sections and the SM predictions and the SMEFT predictions with Wilson coefficients fixed to the observed negative limits set at two standard deviations.

The complete profile likelihood scans are compared with the expected scans in Fig. 52 and Fig. 53 for the linear results, and in Fig. 54 and Fig. 55 for the linear plus quadratic results. For the c_{Hq3} group under the full parametrization, as shown on Fig. 55, the double local minima start being noticeable, as expected from the double roots for a quadratic equation. However, their confidence intervals remain contiguous and no disjoint intervals can be extracted. In general, the systematic uncertainties have small impact on the EFT limits. However, a large effect is observed for c_{Hq3} with the linear plus quadratic parametrization. This is due to several nuisance parameters from jet resolution and top background modeling which are known to have large impact on p_T^{j1} . Overall, the coefficients are all compatible with zero, indicating there is no evidence for any contribution from EFT operators.

Finally, the constraints provided by this analysis are consistent with the results reported in the previous ATLAS VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ EFT analysis [28] using a partial Run-2 dataset corresponding to 36 fb^{-1} collected between 2015-2016. Both analysis illustrate

better sensitivity to c_{HW} and c_{Hq3} comparing to other Wilson coefficients. Our analysis improves the constraints of these two coefficients by a factor of 2-3. In comparison with other Higgs decay channels, the $H \rightarrow l\nu l\nu$ channel is more competitive comparing to the ATLAS $H \rightarrow \gamma\gamma$ analysis [105] in its sensitivity to c_{HW} in both the first the second order of c_i/Λ^2 and to the CP odd operators in the first order as they only affect the decay in the second order.

The large impact from quadratic terms is also observed using the Run-2 dataset. This observation motivates the development for the dimension-8 operators, which can potentially cancel the large quadratic effect by the dimension-6 operators, making the theoretical calculations finite at higher order.

Parameter	Optimal observable	Linear		Linear + Quadratic	
		Best-fit value [TeV ⁻²]	2 σ limits [TeV ⁻²]	Best-fit value [TeV ⁻²]	2 σ limits [TeV ⁻²]
c_{HW}/Λ^2	$\Delta\phi_{jj}$	-0.99	[-2.64, 0.60]	-0.87	[-1.80, 0.61]
$\tilde{c}_{HW}/\Lambda^2$	$\Delta\phi_{jj}$	-0.28	[-1.84, 1.31]	-0.30	[-1.09, 1.37]
c_{HWB}/Λ^2	$\Delta\phi_{jj}$	-3.14	[-13.67, 5.90]	-0.50	[-1.24, 1.08]
$\tilde{c}_{HWB}/\Lambda^2$	$\Delta\phi_{jj}$	-2.60	[-16.77, 11.80]	-0.39	[-1.17, 1.14]
c_{HB}/Λ^2	$\Delta\phi_{jj}$	-1.30	[-6.66, 4.56]	-0.18	[-0.60, 0.66]
$\tilde{c}_{HB}/\Lambda^2$	$\Delta\phi_{jj}$	-4.75	[-31.63, 22.22]	-0.21	[-0.63, 0.63]
c_{Hq3}/Λ^2	p_T^{j1}	-0.18	[-0.74, 0.30]	-0.17	[-0.56, 0.43]
c_{Hq1}/Λ^2	p_T^{j1}	6.11	[-6.92, 21.65]	0.98	[-2.24, 2.04]
c_{Hu}/Λ^2	p_T^{j1}	-3.00	[-11.42, 4.19]	-1.18	[-2.54, 3.11]
c_{Hd}/Λ^2	p_T^{j1}	8.47	[-12.86, 33.43]	-1.96	[-3.70, 3.38]

Table 21: Best-fit values to data and 2σ intervals obtained from profile likelihood scans of the Wilson coefficients under linear-only and linear plus quadratic parametrization.

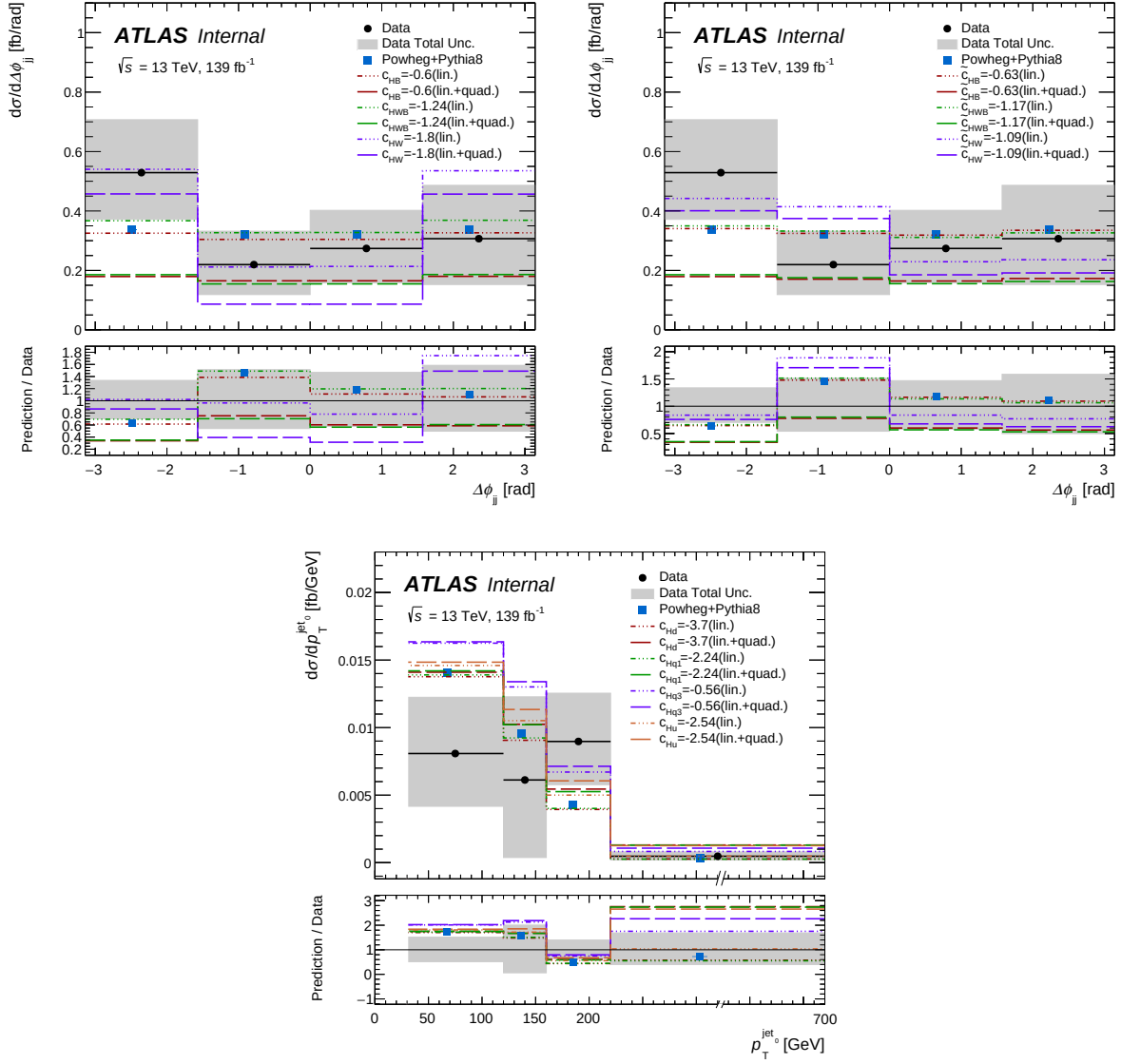


Figure 51: The measured differential cross section (black data points) and uncertainty (gray error band) as a function of $\Delta\phi_{jj}$ (top left and top right) and p_T^{jet} (bottom) is shown together with the predictions without (blue data points) and with EFT corrections (colored histograms). The values of the Wilson coefficients correspond approximately to the observed negative limits set at two standard deviations with the full parametrizations. The contributions from the interference between the SM and EFT amplitudes (lin.) and from the total of the interference and the BSM terms (lin.+quad.) are represented as dotted and dashed lines, respectively.

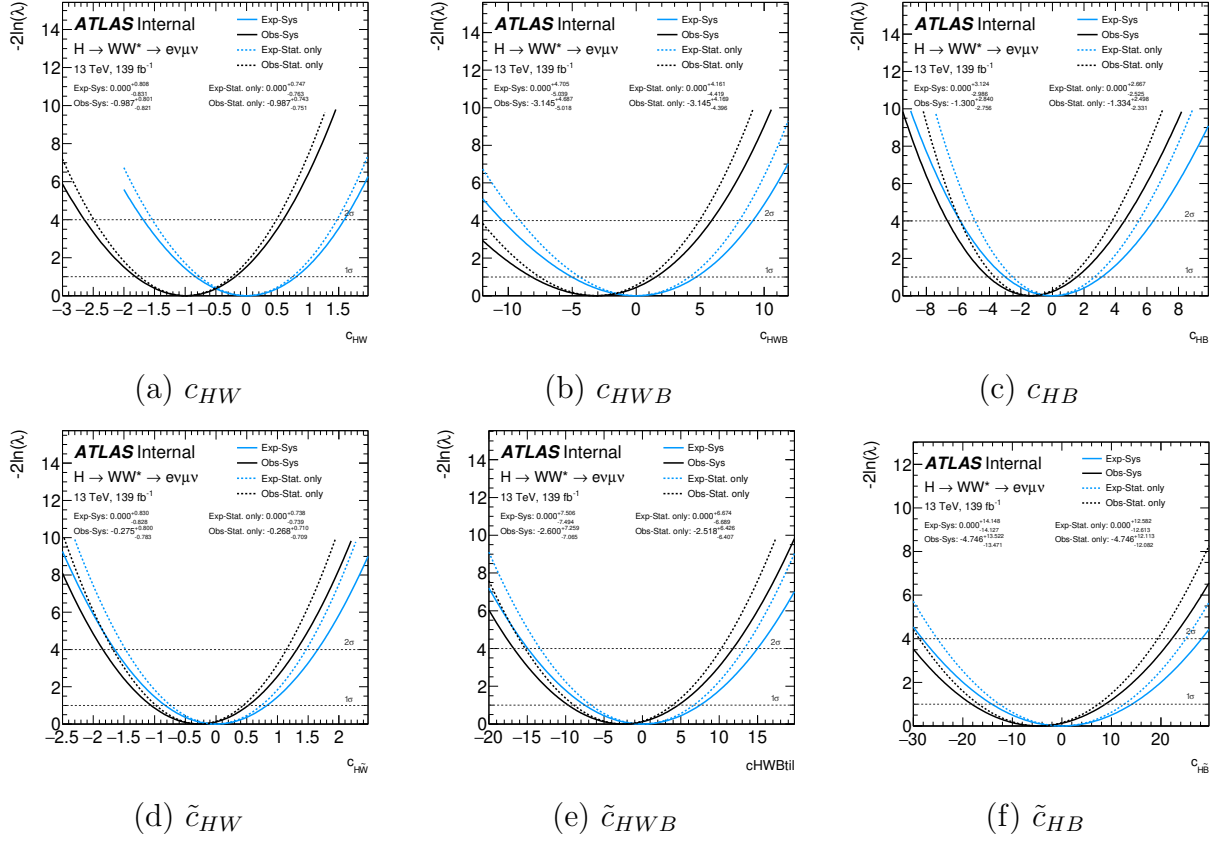
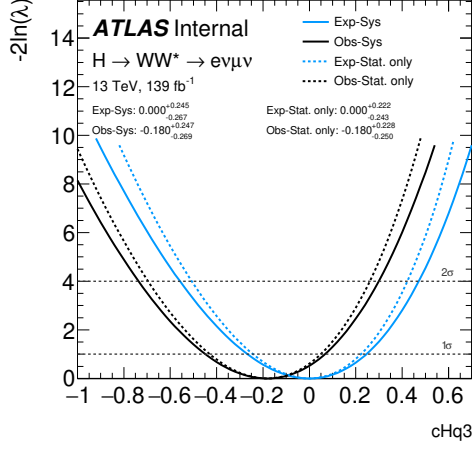
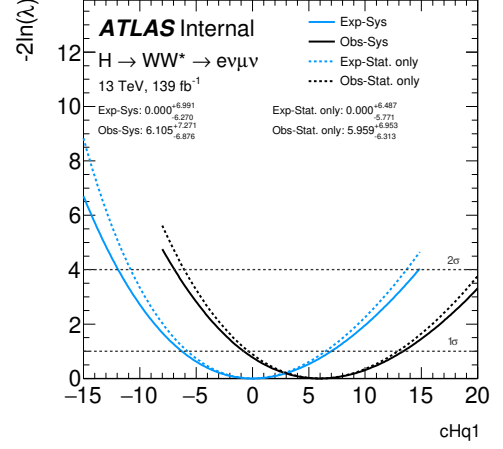


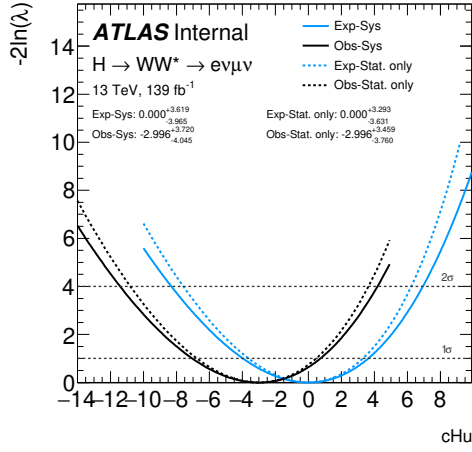
Figure 52: Profile likelihood scan of the c_{HW} group with the linear parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).



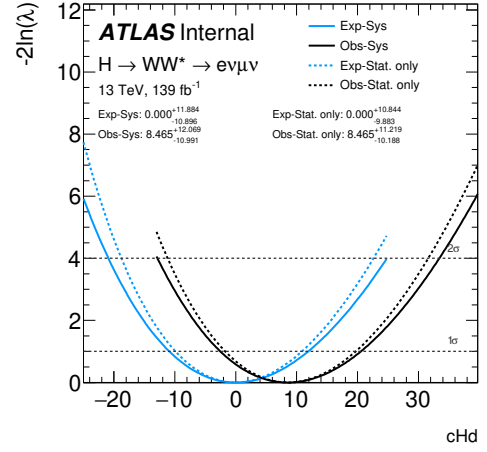
(a) c_{Hq3}



(b) c_{Hq1}



(c) c_{Hu}



(d) c_{Hd}

Figure 53: Profile likelihood scan of the c_{Hq3} group with the linear parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).

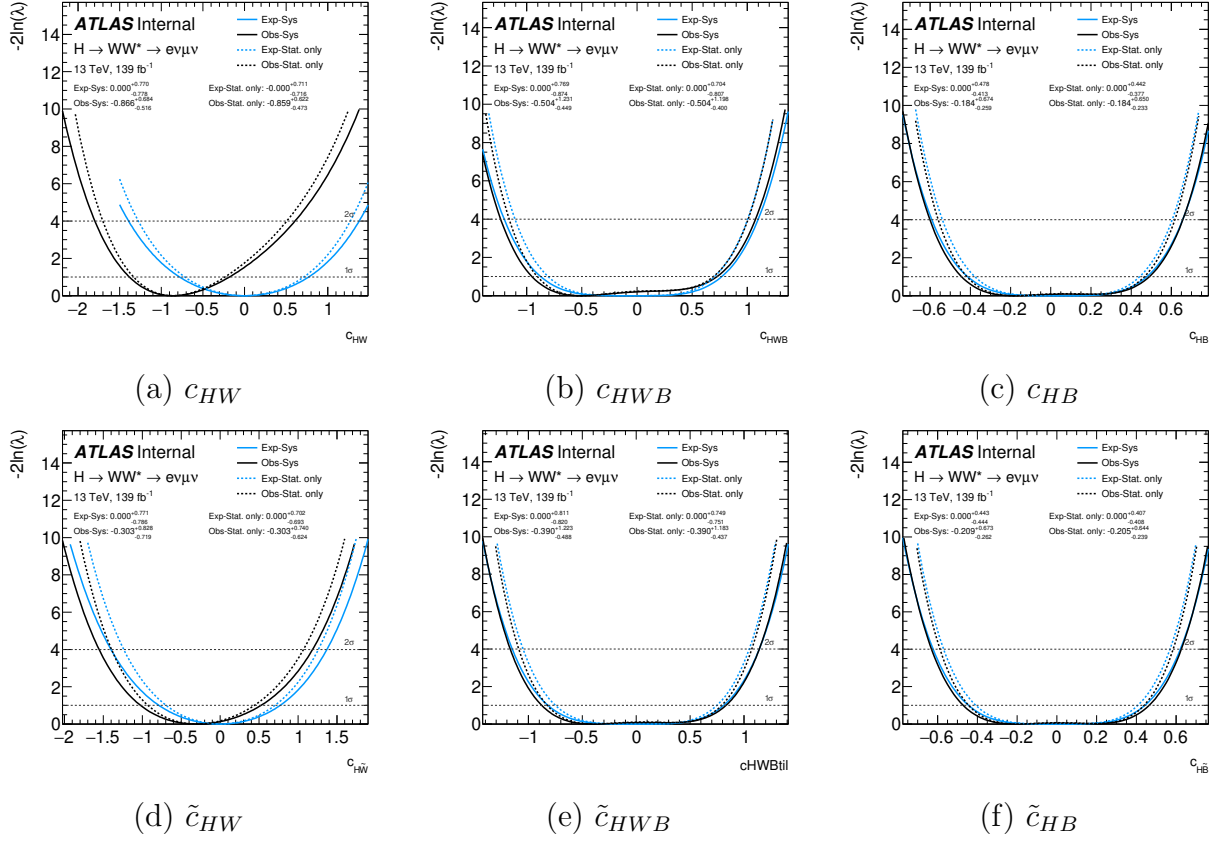
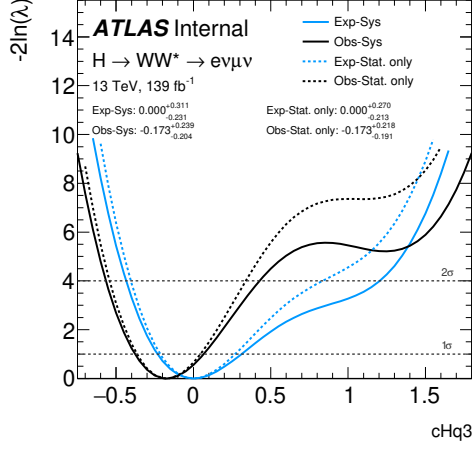
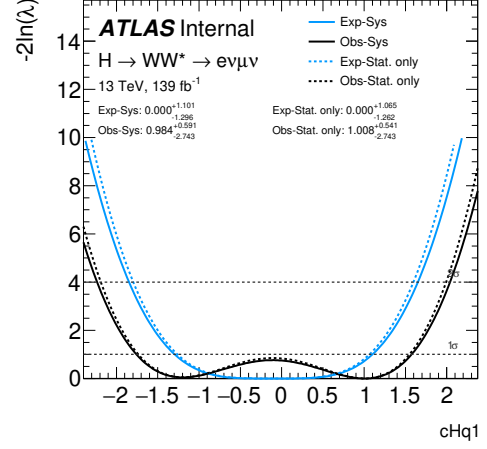


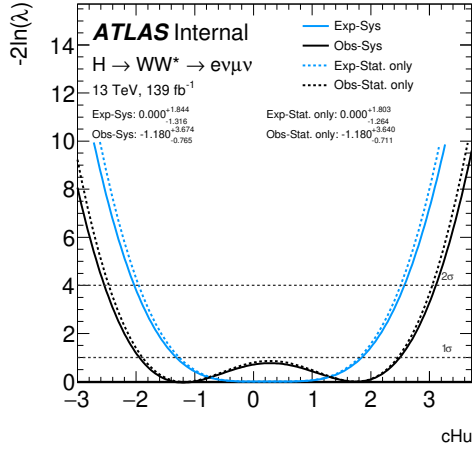
Figure 54: Profile likelihood scan of the c_{HW} group with the full parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).



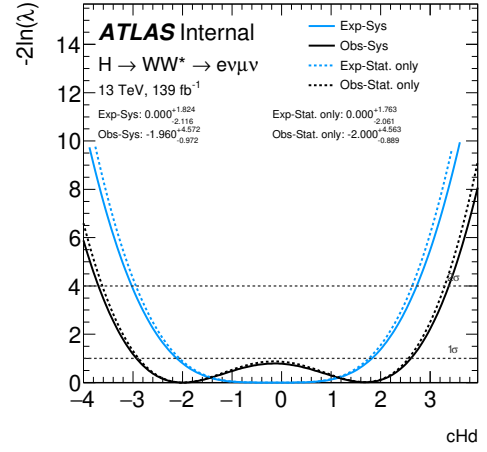
(a) c_{Hq3}



(b) c_{Hq1}



(c) c_{Hu}



(d) c_{Hd}

Figure 55: Profile likelihood scan of the c_{Hq3} group with the full parametrization to the SM Asimov dataset (blue) and to data (black) with statistical uncertainties only (dashed curve), and with statistical and experimental systematics uncertainties (solid curve).

12 Conclusion and outlook

The fiducial and differential cross-section measurements for the vector boson fusion production of the Higgs boson in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ final state are presented for the first time, using proton-proton collision data collected by the ATLAS experiment at a center-of-mass energy of $\sqrt{s} = 13$ TeV and with an integrated luminosity of 139 fb^{-1} . The data are selected with opposite-flavor leptons from the decay of W bosons along with two energetic jets in the final state. The kinematic observables measured from these objects are corrected to particle-level, accounting for detector inefficiencies and resolutions. The measured fiducial cross section is $1.68_{-0.32}^{+0.33} \text{ (stat)}_{-0.21}^{+0.23} \text{ (sys)} \text{ fb}$ and is consistent with several Standard Model predictions including NLO QCD and NLO electroweak corrections. The fiducial cross section is also measured differentially as a function of kinematic variables of the final-state leptons and jets and is compared with several predictions from perturbative QCD calculations and Monte Carlo simulations. The data-theory agreement is reflected by the consistently high p-value ranging from 0.22-0.98 for the nominal signal prediction.

A model-independent new physics framework called Effective Field Theory that includes the interference between the SM and dimension-6 operators as well as pure BSM contribution is used to set limits on 11 Wilson coefficients. The analysis strategy is fine-tuned to provide stringent constraints on several Wilson coefficients using the observables sensitive to the impact from their corresponding operators. Owing to the generality of EFT application, the EFT interpretation provided by this analysis can be prospectively combined with other analyses probing different SM processes to improve the constraining power on the Wilson coefficient. Furthermore, the precision driven by the statistical uncertainty of data will be further enhanced by the next operational run (Run-3) and the High-Luminosity Phase of LHC.

References

- [1] Ilaria Brivio and Michael Trott. The standard model as an effective field theory. *Physics Reports*, 793:1–98, 2019.
- [2] Wikipedia contributors. Plagiarism — Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg, 2004. [Online; accessed 21-April-2022].
- [3] The future of the higgs boson. *PHYSICS TODAY*, 66(12):27, Dec 2013.
- [4] S Heinemeyer et al. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*. CERN Yellow Reports: Monographs. Jul 2013.
- [5] Abdelhak Djouadi. The anatomy of electroweak symmetry breaking: Tome i: The higgs boson in the standard model. *Physics Reports*, 457(1):1–216, 2008.
- [6] ATLAS Collaboration. Combined measurements of Higgs boson production and decay using up to 80 fb⁻¹ of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment. *Phys. Rev. D*, 101(1):012002, 2020.
- [7] CMS Collaboration. Combined measurements of higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 79(5), May 2019.
- [8] Mark Thomson. *Modern particle physics*. Cambridge University Press, New York, 2013.
- [9] Fabienne Marcastel. CERN’s Accelerator Complex. La chaîne des accélérateurs du CERN. Oct 2013. General Photo.
- [10] The ATLAS Collaboration. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, Aug 2008.
- [11] ATLAS Collaboration. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79:639. 40 p, Feb 2019.
- [12] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}=13$ TeV. *Eur. Phys. J. C*, 76:292. 45 p, Mar 2016.
- [13] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS detector. *The European Physical Journal C*, 77(7), Jul 2017.
- [14] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. Technical report, CERN, Geneva, May 2014.
- [15] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

- [16] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132–133, 1964.
- [17] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.
- [18] Heather Gray and Bruno Mansoulié. The higgs boson: the hunt, the discovery, the study and some future perspectives. <https://atlas.cern/updates/feature/higgs-boson>. Accessed: 2022-05-18.
- [19] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lh. *Physics Letters B*, 716(1):1–29, 2012.
- [20] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lh. *Physics Letters B*, 716(1):30–61, 2012.
- [21] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. Wiley, New York, NY, 2008.
- [22] Mary K. Gaillard, Paul D. Grannis, and Frank J. Sciulli. The standard model of particle physics. *Reviews of Modern Physics*, 71(2):S96–S111, Mar 1999.
- [23] Chris Quigg. Electroweak symmetry breaking in historical perspective. *Annual Review of Nuclear and Particle Science*, 65(1):25–42, 2015.
- [24] ATLAS Collaboration and CMS Collaboration. *Phys. Rev. Lett.*, 114:191803, May 2015.
- [25] Thomas Appelquist and J. Carazzone. Infrared singularities and massive fields. *Phys. Rev. D*, 11:2856–2861, May 1975.
- [26] Céline Degrande et al. Effective field theory: A modern approach to anomalous couplings. *Annals of Physics*, 335:21–32, Aug 2013.
- [27] B. Grzadkowski, M. Iskrzyski, M. Misiak, and J. Rosiek. Dimension-six terms in the standard model lagrangian. *Journal of High Energy Physics*, 2010(10), Oct 2010.
- [28] ATLAS Collaboration. Combined effective field theory interpretation of $H \rightarrow WW^*$ and WW measurements using ATLAS data. Technical report, CERN, Geneva, Mar 2021.
- [29] Interpretations of the combined measurement of Higgs boson production and decay. Technical report, CERN, Geneva, Oct 2020.
- [30] Jacob J. Ethier et al. Combined smeft interpretation of higgs, diboson, and top quark data from the lh, 2021.
- [31] Sally Dawson and Christopher W. Murphy. Standard model EFT and extended scalar sectors. *Physical Review D*, 96(1), Jul 2017.

- [32] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001–S08001, Aug 2008.
- [33] CERN. LHC Guide. Brochure. Mar 2017.
- [34] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, Aug 2008.
- [35] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002–S08002, Aug 2008.
- [36] The LHCb Collaboration. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, Aug 2008.
- [37] ATLAS Collaboration. Technical Design Report for the ATLAS Inner Tracker Pixel Detector. Technical report, CERN, Geneva, Sep 2017.
- [38] ATLAS Collaboration. Technical Design Report for the ATLAS Inner Tracker Strip Detector. Technical report, CERN, Geneva, Apr 2017.
- [39] ATLAS Collaboration. Particle Identification Performance of the ATLAS Transition Radiation Tracker. Technical report, CERN, Geneva, Sep 2011.
- [40] Capeans M. et al. ATLAS Insertable B-Layer Technical Design Report. Technical report, Sep 2010.
- [41] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.
- [42] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.
- [43] ATLAS Collaboration. The ATLAS Data Acquisition system in LHC Run 2. Technical report, CERN, Geneva, Feb 2017.
- [44] Federico Meloni. Primary vertex reconstruction with the ATLAS detector. Technical report, CERN, Geneva, Oct 2016.
- [45] T Cornelissen, M Elsing, I Gavrilenko, W Liebig, E Moyse, and A Salzburger. The new ATLAS track reconstruction (NEWT). *Journal of Physics: Conference Series*, 119(3):032014, Jul 2008.
- [46] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [47] ATLAS Collaboration. Jet global sequential corrections with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. Technical report, CERN, Geneva, Mar 2015.

- [48] Serguei Chatrchyan et al. Observation of a New Boson with Mass Near 125 GeV in pp Collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 06:081, 2013.
- [49] Albert M. Sirunyan et al. Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 791:96, 2019.
- [50] ATLAS Collaboration. Measurements of gluon-gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett. B*, 789:508–529, 2019.
- [51] ATLAS Collaboration. Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector. *Phys. Rev. D*, 92(1):012006, 2015.
- [52] ATLAS Collaboration. Observation of vector-boson-fusion production of Higgs bosons in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Aug 2020.
- [53] G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13(07):P07017, 2018.
- [54] Paul Laycock et al. ATLAS data preparation in run 2. Technical report, CERN, Geneva, Feb 2017.
- [55] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [56] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 02:037, 2010.
- [57] Torbjorn Sjostrand et al. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, 2015.
- [58] ATLAS Collaboration. Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 09:145, 2014.
- [59] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth. A*, 462:152, 2001.
- [60] NNPDF Collaboration. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [61] M. Ciccolini, A. Denner and S. Dittmaier. Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC. *Phys. Rev. Lett.*, 99, 2007.
- [62] K. Arnold et al. Vbfno: A parton level monte carlo for processes with electroweak bosons. *Computer Physics Communications*, 180(9):16611670, Sep 2009.

- [63] P. Bolzoni and others. Higgs production via vector-boson fusion at NNLO in QCD. *Phys. Rev. Lett.*, 105, 2010.
- [64] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.
- [65] F. Cascioli, S. Hoehe, F. Krauss, P. Maierhofer, S. Pozzorini, and F. Siegert. Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production. *JHEP*, 01:046, 2014.
- [66] T. Gleisberg, Stefan. Hoche, F. Krauss, M. Schonherr, S. Schumann, et al. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.
- [67] Michal Czakon, Paul Fiedler, and Alexander Mitov. Total top-quark pair-production cross section at hadron colliders through $\mathcal{O}(\alpha_s^4)$. *Physical Review Letters*, 110(25), Jun 2013.
- [68] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [69] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [70] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [71] Keith Hamilton, Paolo Nason, Emanuele Re, and Giulia Zanderighi. NNLOPS simulation of Higgs boson production. *JHEP*, 10:222, 2013.
- [72] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator. *JHEP*, 05:140, 2015.
- [73] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J. C*, 72:1896, 2012.
- [74] ATLAS Collaboration. Deep Learning in Flavour Tagging at the ATLAS experiment. Technical report, CERN, Geneva, Oct 2017.
- [75] 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*, 78:903, 2018.
- [76] T. Plehn, David L. Rainwater, and D. Zeppenfeld. A method for identifying $H \rightarrow \tau\tau \rightarrow e^+\mu^+ + \text{missing p(T)}$ at the CERN LHC. *Phys.Rev.*, D61:093005, 2000.
- [77] A Hoecker et al. Tmva - toolkit for multivariate data analysis, 2007.
- [78] Measurements of gluon fusion and vector-boson-fusion production of the Higgs boson in $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decays using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical report, CERN, Geneva, Mar 2021.

- [79] V. Barger, R.J.N. Phillips, and D. Zeppenfeld. Minijet veto: a tool for the heavy higgs search at the lh. *Physics Letters B*, 346(1):106–114, 1995.
- [80] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, Apr 2008.
- [81] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The catchment area of jets. *Journal of High Energy Physics*, 2008(04):005–005, Apr 2008.
- [82] J. Alwall et al. A standard format for les houches event files. *Computer Physics Communications*, 176(4):300–304, Feb 2007.
- [83] G P Lepage. VEGAS - an adaptive multi-dimensional integration program. Technical Report CLNS-447, Cornell Univ. Lab. Nucl. Stud., Ithaca, NY, Mar 1980.
- [84] J. Butterworth et al. Les houches 2013: Physics at tev colliders: Standard model working group report, 2014.
- [85] J. Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys.*, G43, 2016.
- [86] Luca Lista. Practical statistics for particle physicists. *CERN Yellow Reports: School Proceedings*, pages Vol 5 (2017): Proceedings of the 2016 European School of High–Energy Physics, 2017.
- [87] Eilam Gross Ofer Vitells Glen Cowan, Kyle Cranmer. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), Feb 2011.
- [88] Kyle S. Cranmer. Kernel estimation in high-energy physics. *Comput. Phys. Commun.*, 136:198–207, 2001.
- [89] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data. *JINST*, 14:P12006, 2019.
- [90] ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV. Dec 2020.
- [91] ATLAS Collaboration. Performance of electron and photon triggers in ATLAS during LHC Run 2. *Eur. Phys. J. C*, 80(1):47, 2020.
- [92] ATLAS Collaboration. Performance of the ATLAS muon triggers in Run 2. *JINST*, 15(09):P09015, 2020.
- [93] 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*, 96:072002, 2017.

- [94] 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*, 76:581, 2016.
- [95] 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*, 79:970, 2019.
- [96] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. ATLAS-CONF-2019-021, 2019.
- [97] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals Math. Statist.*, 9(1):60–62, 1938.
- [98] 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*, May 2016.
- [99] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2/2017, Oct 2016.
- [100] Methodology for EFT interpretation of Higgs boson Simplified Template Cross-section results in ATLAS. Technical report, CERN, Geneva, Oct 2019.
- [101] ATLAS Collaboration. Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 80(10):957, 2020.
- [102] Ilaria Brivio, Yun Jiang, and Michael Trott. The smeftsim package, theory and tools. *Journal of High Energy Physics*, 2017(12), Dec 2017.
- [103] Andy Buckley et al. Rivet user manual. *Computer Physics Communications*, 184(12):2803–2819, Dec 2013.
- [104] Aleksandr Azatov et al. Helicity selection rules and noninterference for bsm amplitudes. *Phys. Rev. D*, 95:065014, Mar 2017.
- [105] ATLAS Collaboration. Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 98:052005, 2018.
- [106] José Wudka. Electroweak effective lagrangians. *International Journal of Modern Physics A*, 09(14):2301–2361, Jun 1994.

Appendices

A Expected differential cross sections

This section presents the expected differential cross sections extracted using the Asimov dataset. The results demonstrate closure with the nominal POWHEG+PYTHIA 8 prediction. The expected precision of cross-section measurement per bin is generally at the 40-80% level. A few bins measured at high energy scale are intentionally measured with uncertainty greater than 100%, due to the expected low statistics. We plan to use the measurement to put an upper bound in the cross sections measured in these bins.

The Asimov results for p_T^H is shown in Fig. 56. The $p_T^{\ell\ell}$, $\Delta y_{\ell\ell}$, Δy_{jj} and $\cos\theta_\eta^*$ distributions are shown in Fig. 57. Several lepton variables ($p_T^{\ell_1}$, $p_T^{\ell_2}$, m_{ll} and $\Delta\phi_{\ell\ell}$) and jet variables ($p_T^{j_1}$, $p_T^{j_2}$, m_{jj} , and $\Delta\phi_{jj}$) are shown in Fig. 58 and Fig. 59, respectively

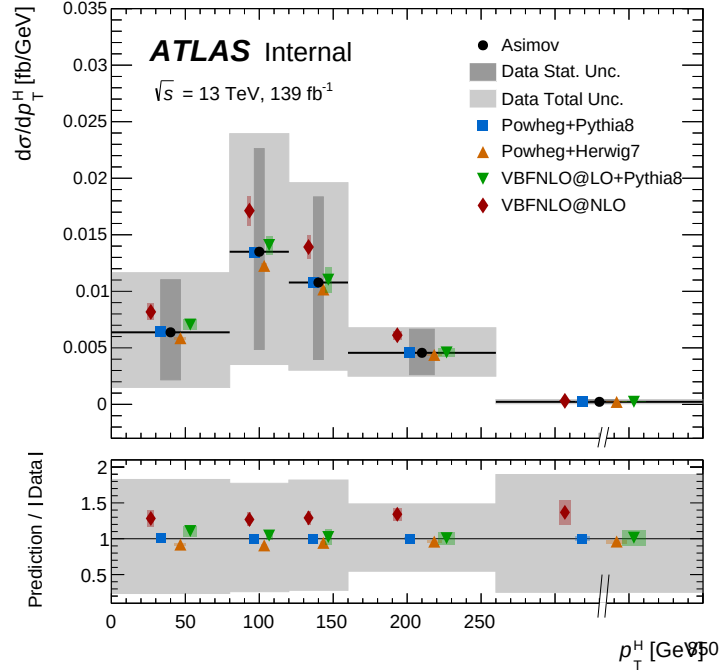


Figure 56: The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for p_T^H in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.

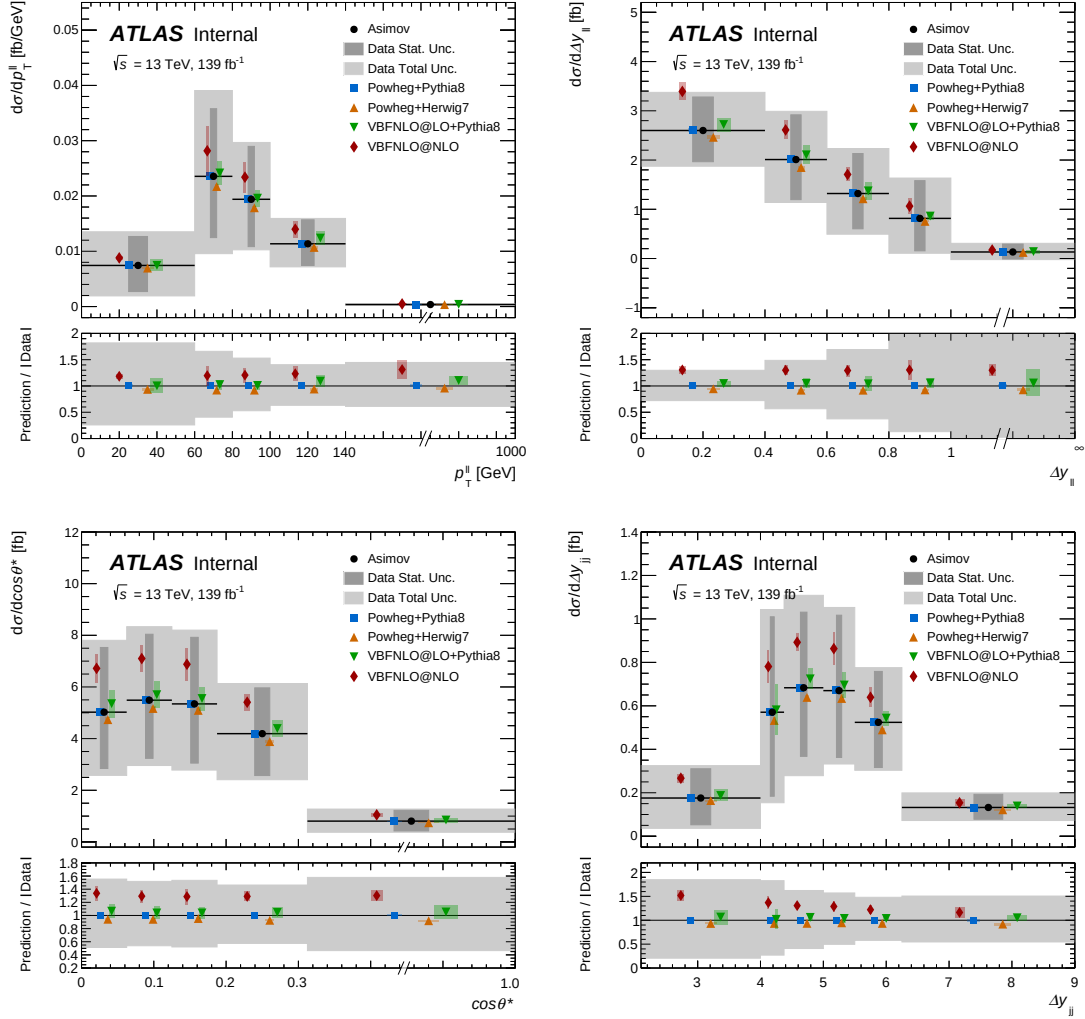


Figure 57: The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell\ell}$, $\Delta y_{\ell\ell}$, Δy_{jj} and $\cos\theta_{\eta}^*$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+ PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.

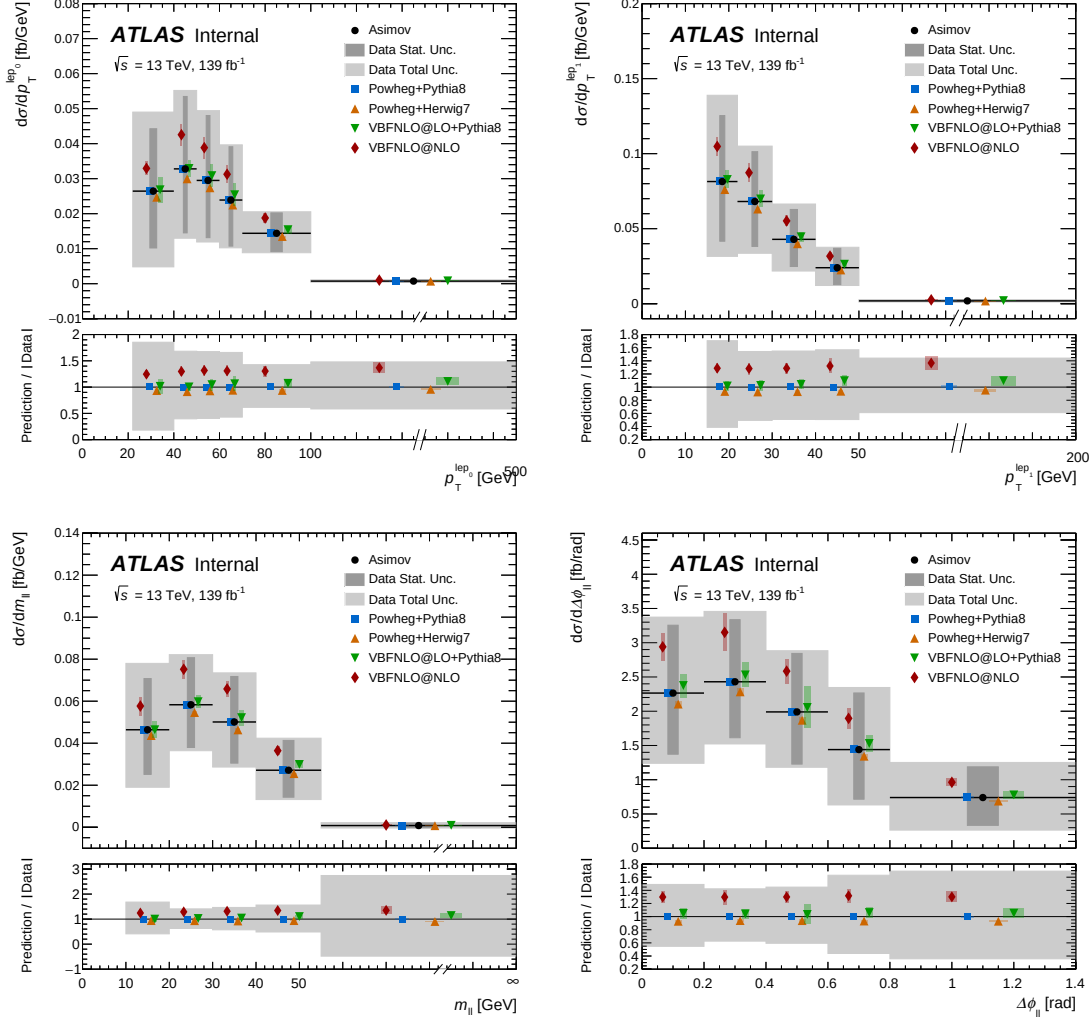


Figure 58: The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for $p_T^{\ell_1}$, $p_T^{\ell_2}$, $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.

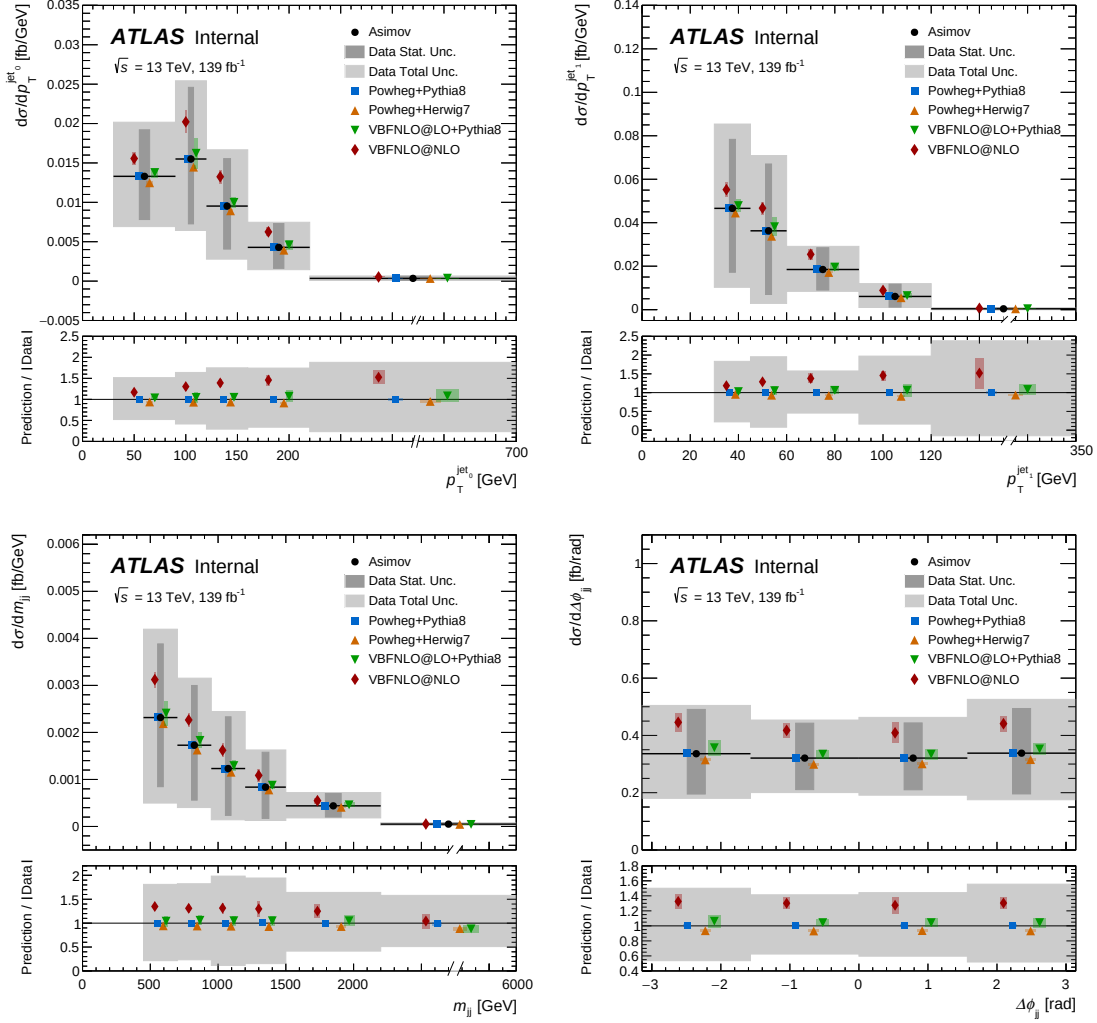


Figure 59: The expected differential fiducial cross sections at $\sqrt{s} = 13$ TeV for p_T^{j1} , p_T^{j2} , m_{jj} , and $\Delta\phi_{jj}$ in comparison with the theoretical predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, MG5+HERWIG 7 and VBFNLO+PYTHIA 8 at particle-level and from VBFNLO-NLO at parton-level. The error bars on the data points show the statistical uncertainties, while the total uncertainties are indicated by the gray boxes. The error bars on the expected cross sections indicate the total theory uncertainty. The bottom panel shows the ratio of different predictions to the data, and the gray area represents the total uncertainty of the measurement.

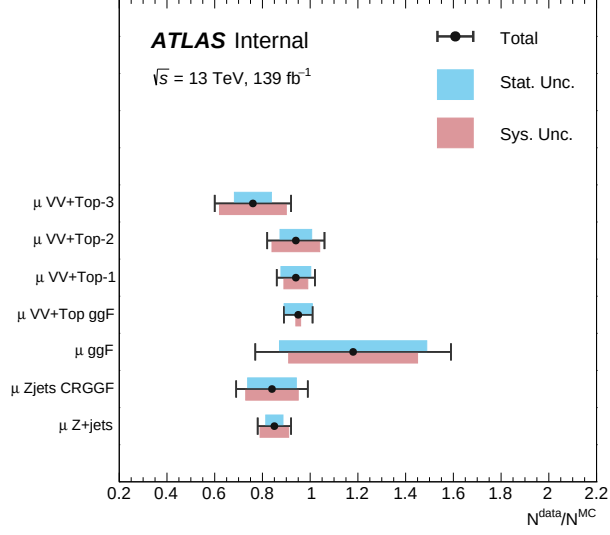
B Differential cross sections

We report measured background normalization factors (labeled as μ or $N^{\text{Data}}/N^{\text{MC}}$) and the correlation among the cross sections (σ_i , where i corresponds to the bin number) and the normalization factors for all the measured differential cross sections in Fig. 60-Fig. 72. The Drell-Yan normalization factors are labeled as “Z+jets” for the parameter used for simultaneous estimation in SR and CRZjets, and “Z+jets CRGGF” or “Z+jets ggF” for the one used in ggF CR. The only ggF normalization factor is labeled “ggF”. Finally, the top plus diboson background is denoted by “VV+Top- j ” (j numerates the number of normalization factors according to Table 14) for the parameters from SRs and “VV+Top ggF” for the one from ggF CR.

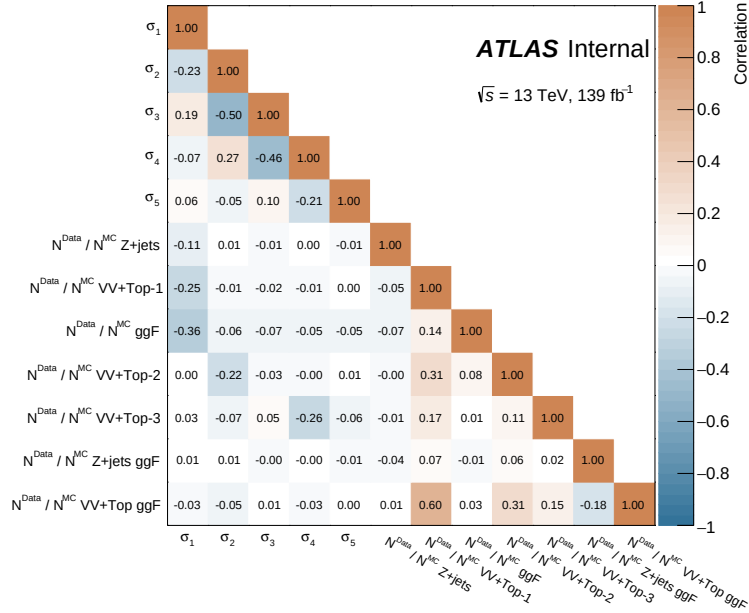
The background normalizations are compatible across measured observables, and comparable to the ones in the inclusive fiducial cross-section measurement. The top and diboson normalization factors in the SRs range from $0.70 - 0.95$, whereas the ones in the ggF CR range from $0.95 - 0.98$. The top and diboson normalization in the SRs fluctuate more significantly since they are estimated in different kinematic regions, some of which are more susceptible to higher order effects. The Z+jets factor from CRZjets and ggF CR ranges from $0.84 - 0.87$ and $0.78 - 0.86$, respectively. The estimation of the Drell-Yan normalization is stable since it is driven by CRZjets which is invariant across measurements of different observables. Finally, the ggF normalization factor spans the largest range from $0.94 - 1.21$ with an uncertainty of 40%.

The correlation matrices among the measured cross-section values in each bin of the distributions and the background normalization factors, as results of the likelihood fits, are obtained independently for each differential cross-section measurement. The correlations among parameters are generally less than 20%, with a few exceptions. Anti-correlations (10% – 50%) are observed between the yields of adjacent fiducial bins due to bin-to-bin migrations. Between the top and diboson normalization factors and the cross section bins

from which the top plus diboson background is estimated, anti-correlation up to about 35% is observed. Up to 35% anti-correlation is measured between cross sections and ggF normalization factors from differential bins with large ggF contamination. Similar to the inclusive fiducial cross section measurement, a high correlation up to 60% among the top plus diboson factors in the SRs and in ggF CR is caused by their modeling uncertainties.

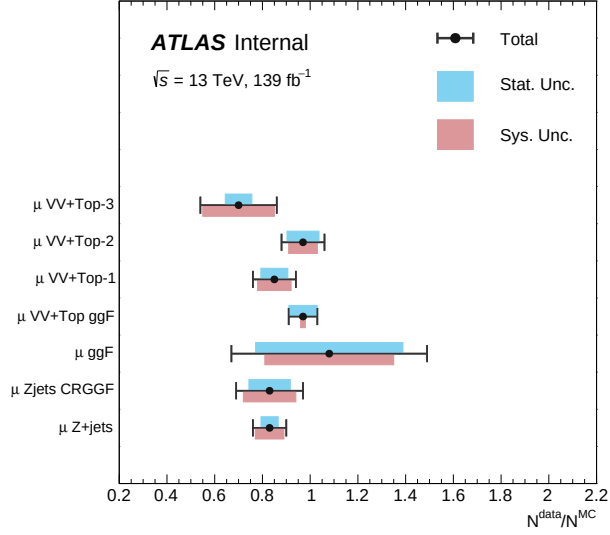


(a) Background normalization factors.

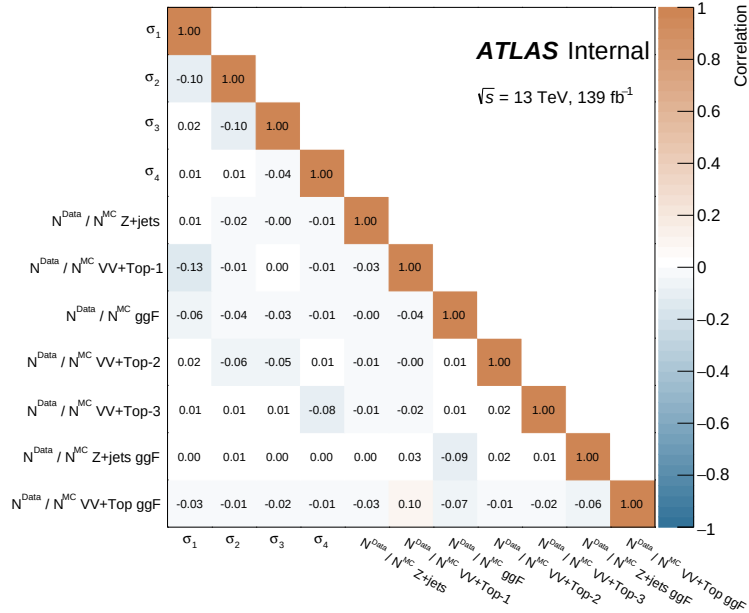


(b) Correlations of the measured parameters.

Figure 60: The measured background normalization factors (μ) and the correlations among the measured cross section in m_{jj} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

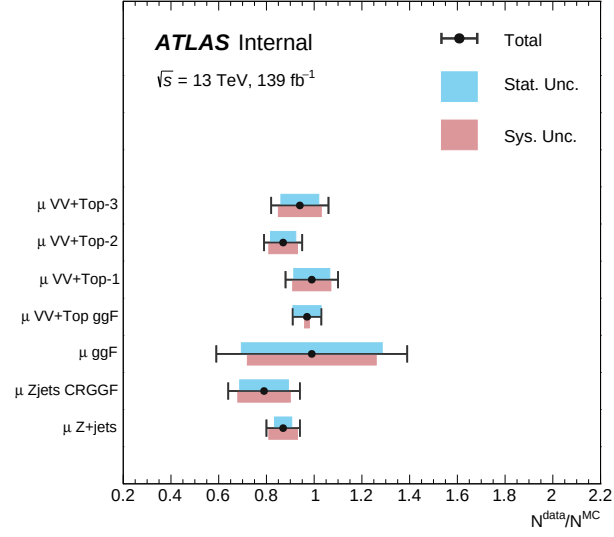


(a) Background normalization factors.

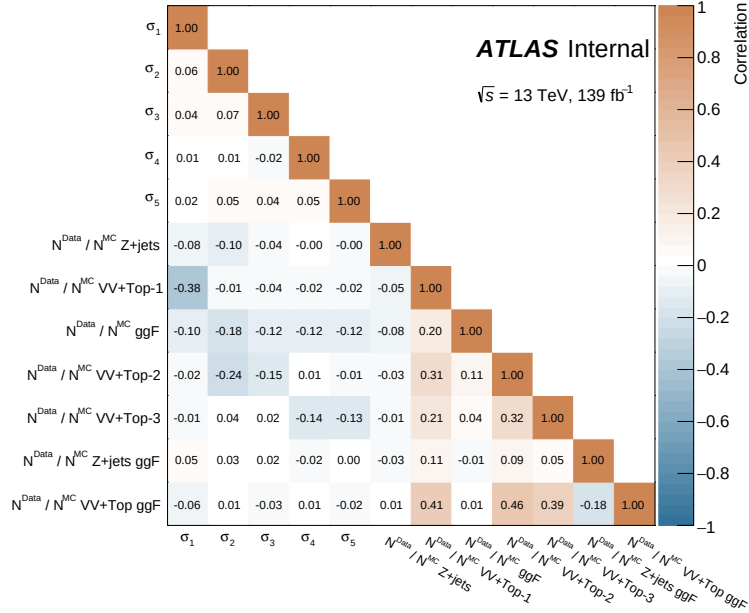


(b) Correlations of the measured parameters.

Figure 61: The measured background normalization factors (μ) and the correlations among the measured cross section in Higgs p_T bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

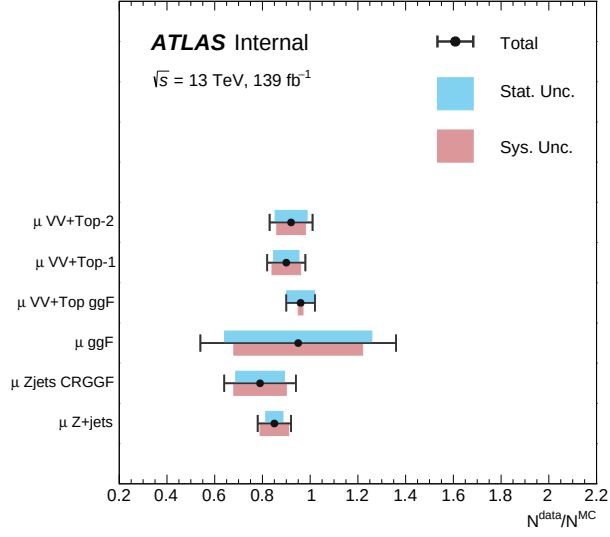


(a) Background normalization factors.

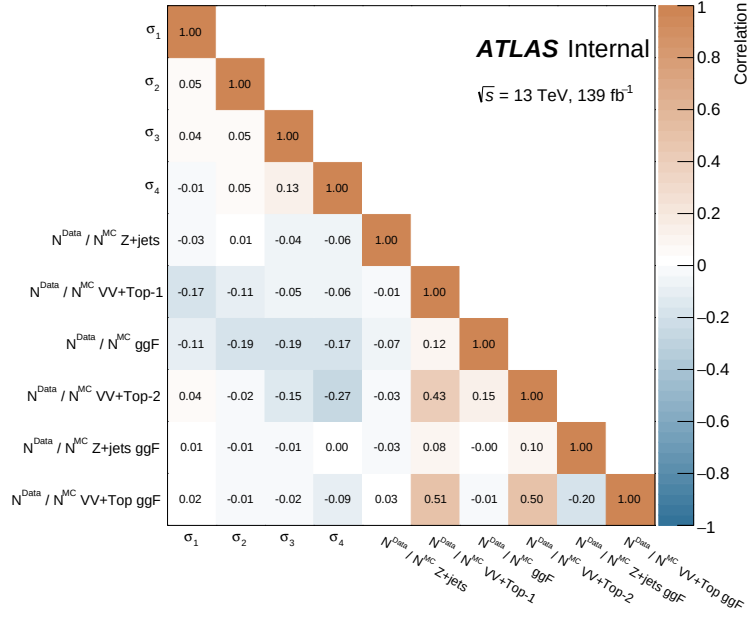


(b) Correlations of the measured parameters.

Figure 62: The measured background normalization factors (μ) and the correlations among the measured cross section in Δy_{jj} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

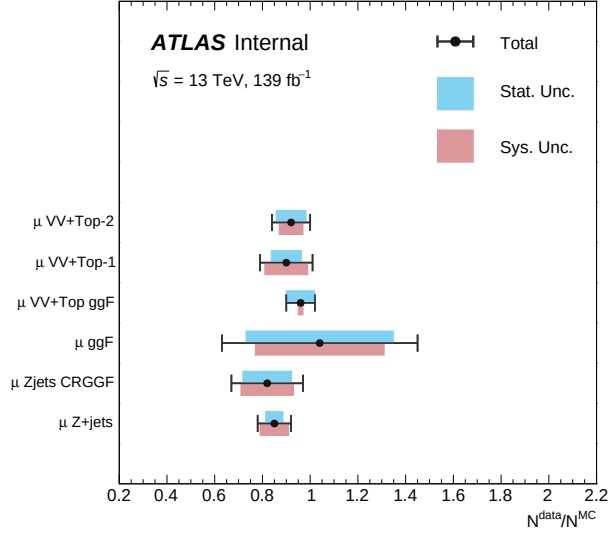


(a) Background normalization factors.

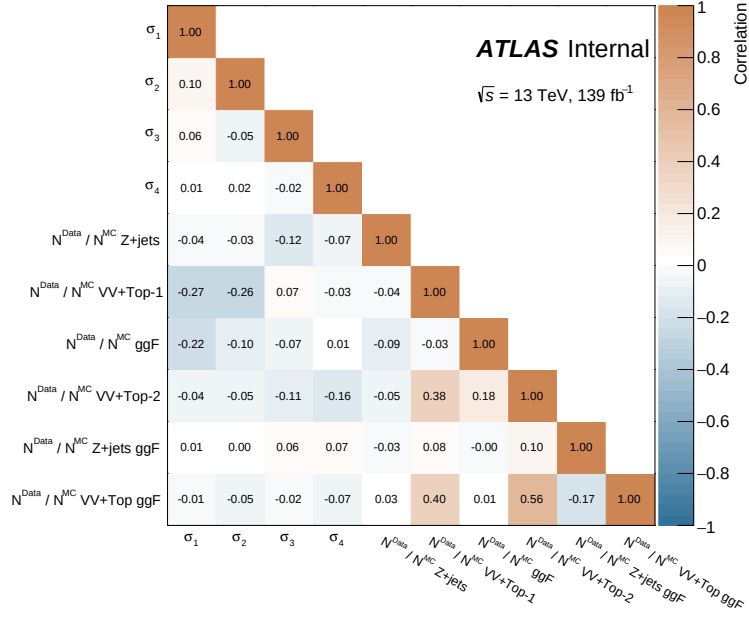


(b) Correlations of the measured parameters.

Figure 63: The measured background normalization factors (μ) and the correlations among the measured cross section in $\Delta\phi_{jj}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

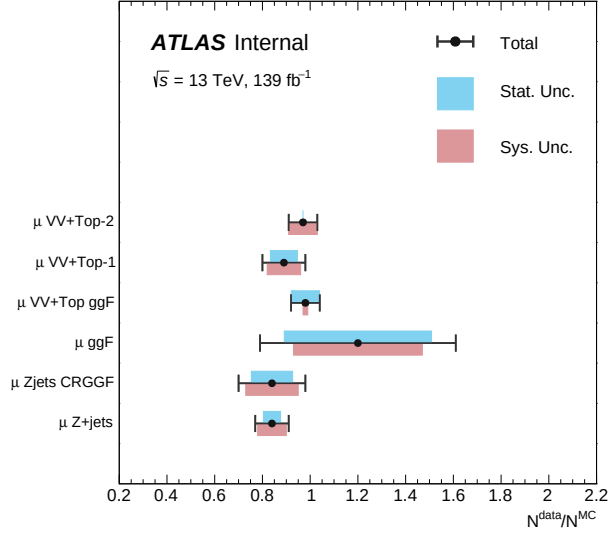


(a) Background normalization factors.

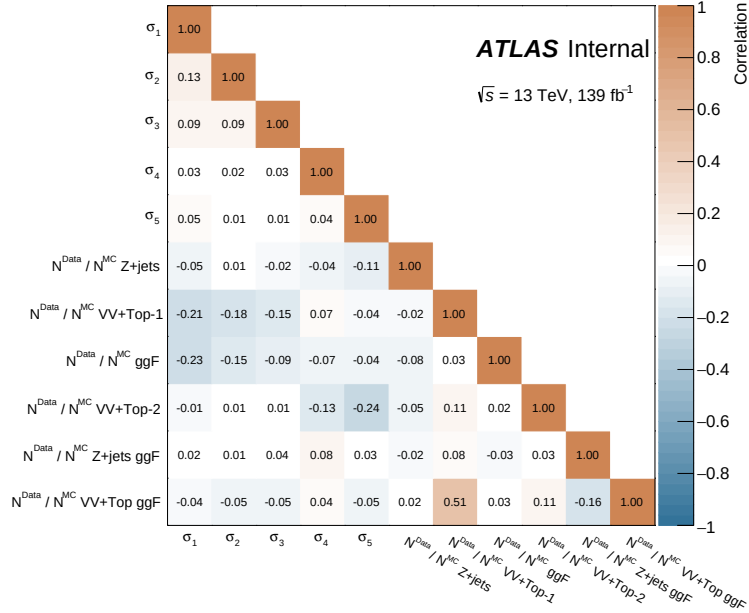


(b) Correlations of the measured parameters.

Figure 64: The measured background normalization factors (μ) and the correlations among the measured cross section in m_{ll} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

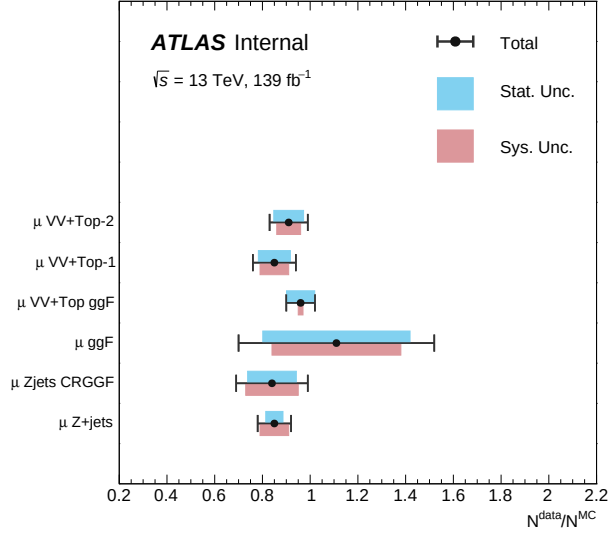


(a) Background normalization factors.

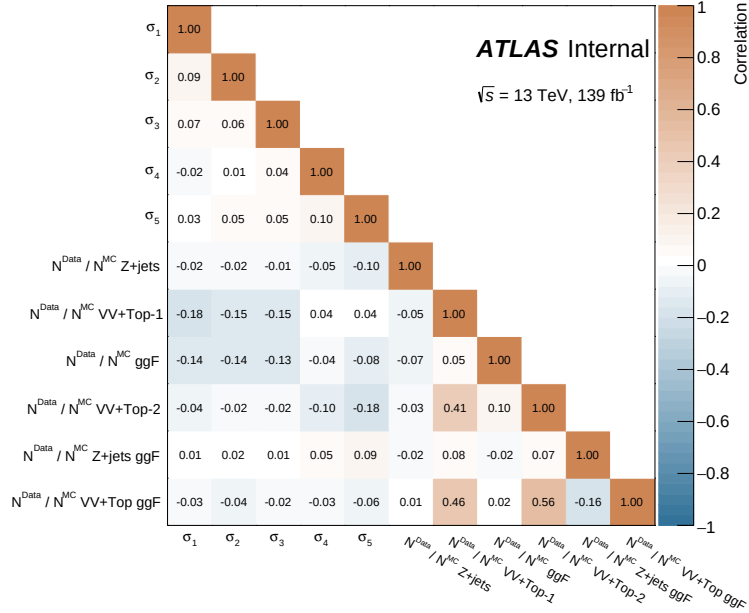


(b) Correlations of the measured parameters.

Figure 65: The measured background normalization factors (μ) and the correlations among the measured cross section in $\Delta y_{\ell\ell}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

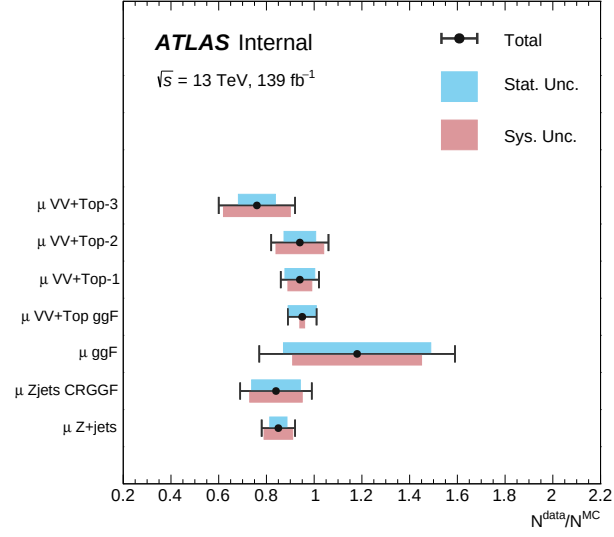


(a) Background normalization factors.

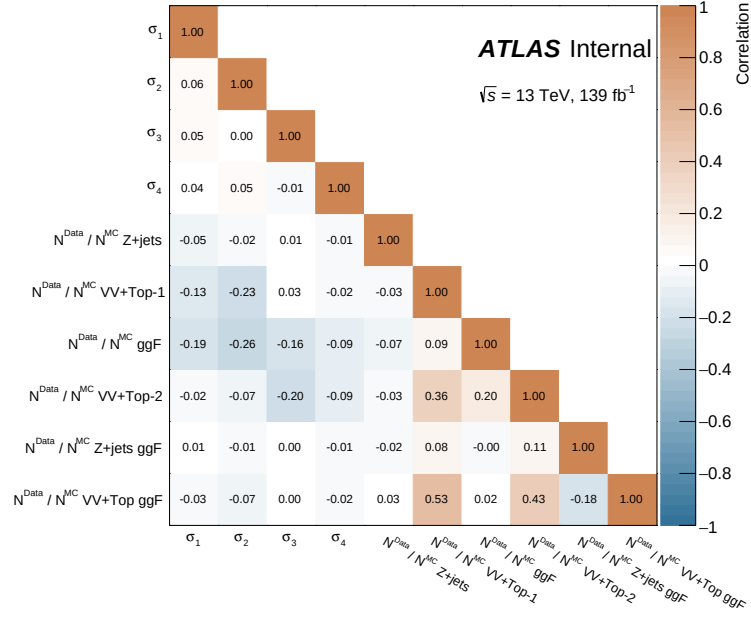


(b) Correlations of the measured parameters.

Figure 66: The measured background normalization factors (μ) and the correlations among the measured cross section in $\Delta\phi_{\ell\ell}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

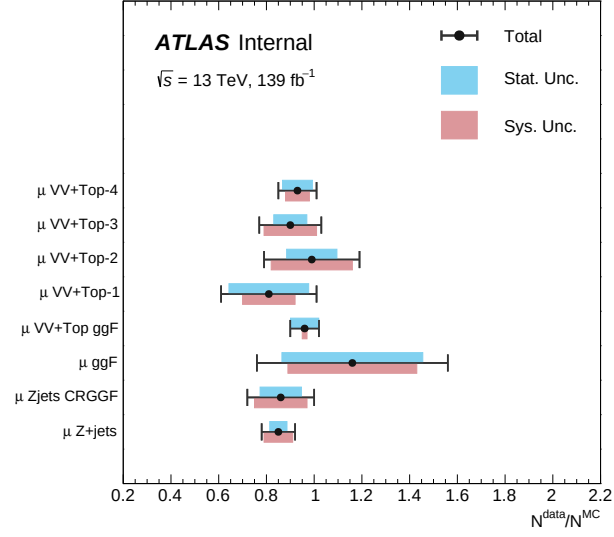


(a) Background normalization factors.

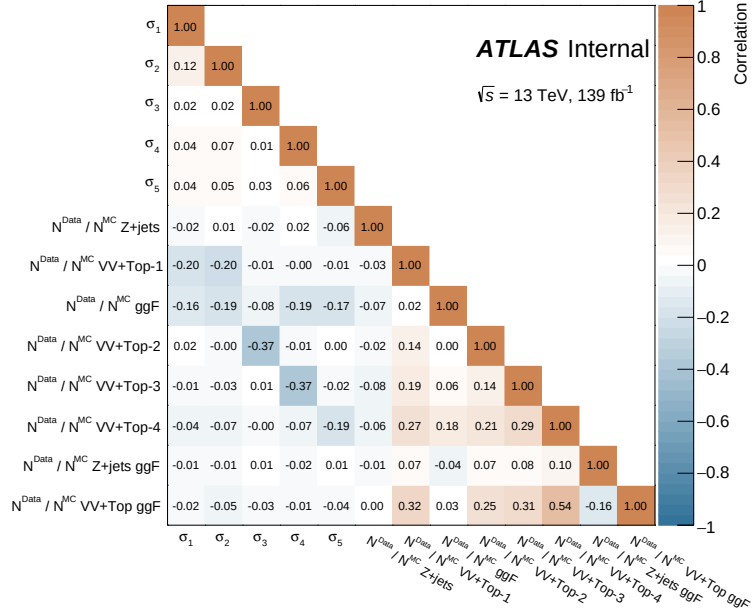


(b) Correlations of the measured parameters.

Figure 67: The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{\ell\ell}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

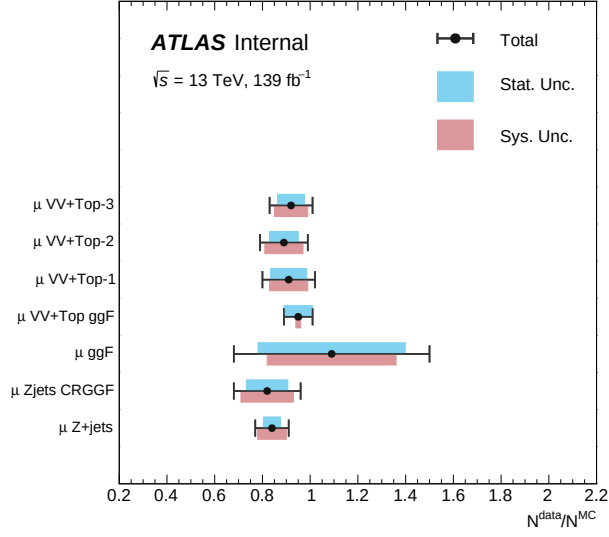


(a) Background normalization factors.

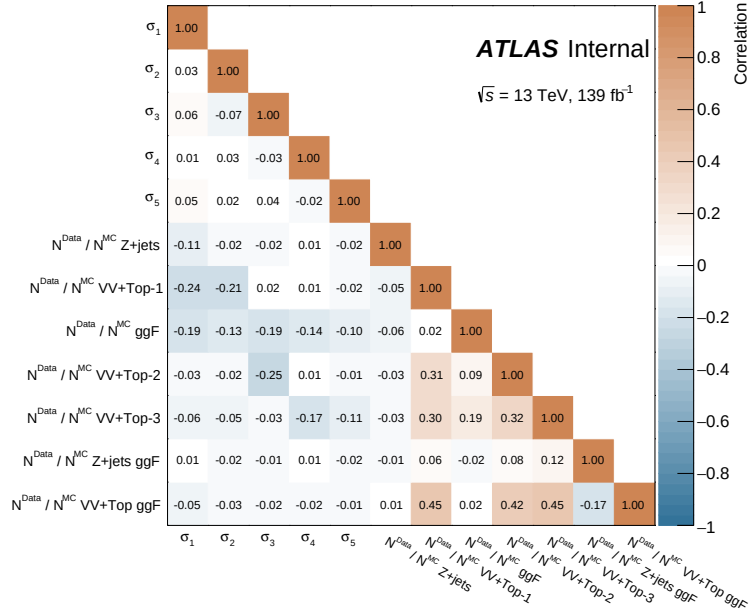


(b) Correlations of the measured parameters.

Figure 68: The measured background normalization factors (μ) and the correlations among the measured cross section in $\cos \theta_\eta^*$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

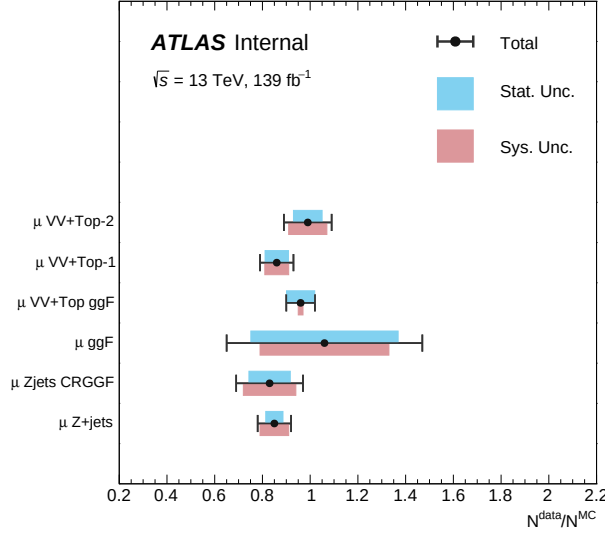


(a) Background normalization factors.

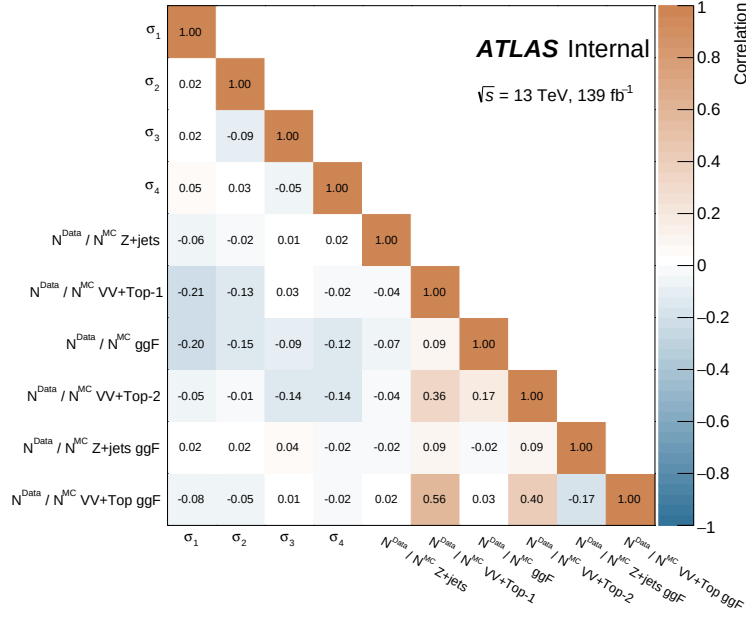


(b) Correlations of the measured parameters.

Figure 69: The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{\ell_1}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

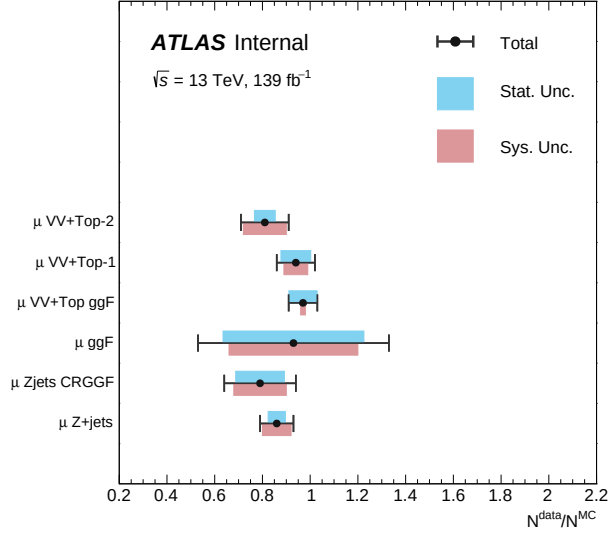


(a) Background normalization factors.

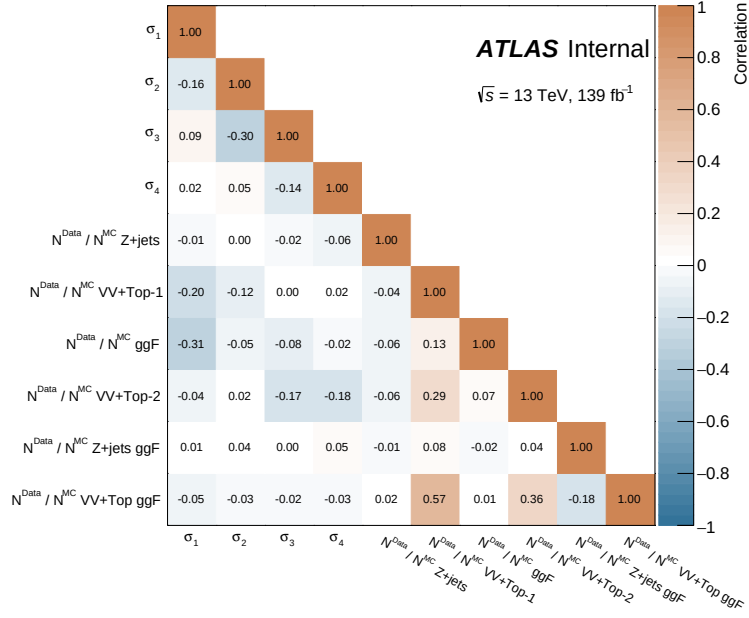


(b) Correlations of the measured parameters.

Figure 70: The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{\ell_2}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

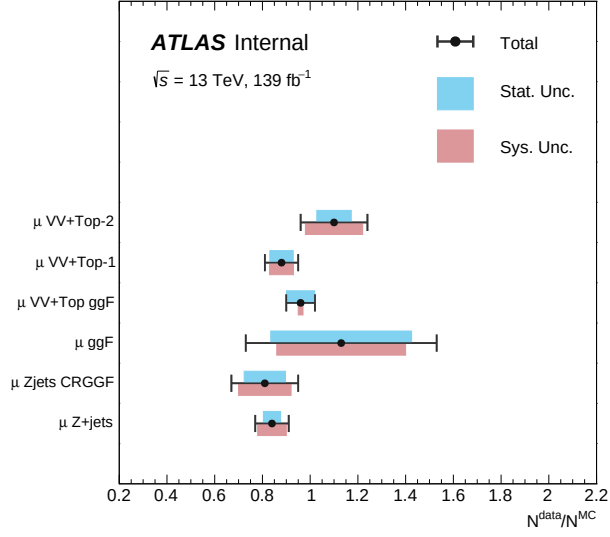


(a) Background normalization factors.

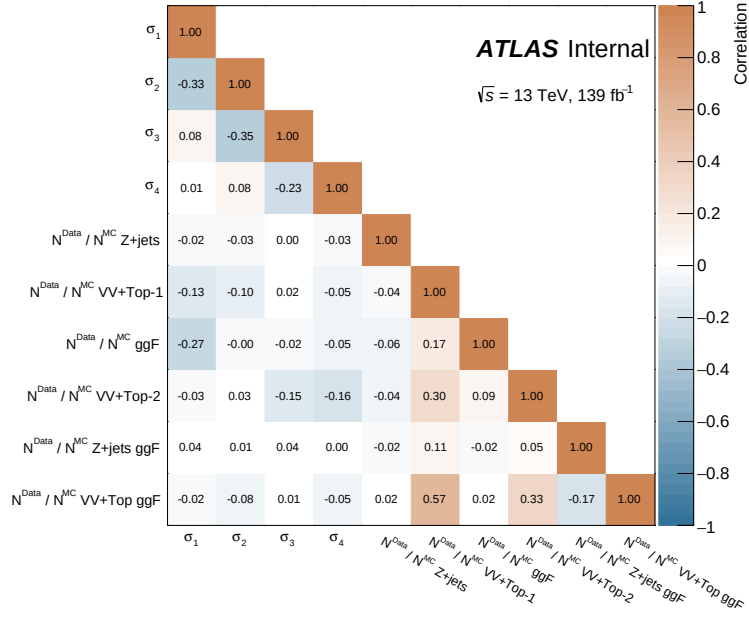


(b) Correlations of the measured parameters.

Figure 71: The measured background normalization factors (μ) and the correlations among the measured cross section in $p_T^{j_1}$ bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).



(a) Background normalization factors.



(b) Correlations of the measured parameters.

Figure 72: The measured background normalization factors (μ) and the correlations among the measured cross section in p_T^{j2} bins (σ_i) and the normalization factors ($N^{\text{Data}}/N^{\text{MC}}$).

C EFT sensitivity test

In the search for optimal observables for setting EFT limits, multiple observables are affected by the Wilson coefficients in the c_{Hq3} group. While the most sensitive observable is $p_T^{j_1}$, the $p_T^{\ell_1}$ distribution shape is also susceptible to these EFT operators. We test the sensitivity using $p_T^{\ell_1}$ and report the limits given by the Asimov data in Table 22.

Wilson coeff	Observable	Best-fit value [TeV ⁻²]	Linear 2 σ limits [TeV ⁻²]	Linear + Quadratic 2 σ limits [TeV ⁻²]
c_{Hq3}/Λ^2	$p_T^{\ell_1}$	0.00	[−0.56, 0.47]	[−0.43, 1.20]
c_{Hq1}/Λ^2	$p_T^{\ell_1}$	0.00	[−11.92, 14.78]	[−1.85, 1.65]
c_{Hu}/Λ^2	$p_T^{\ell_1}$	0.00	[−8.31, 6.94]	[−2.03, 2.57]
c_{Hd}/Λ^2	$p_T^{\ell_1}$	0.00	[−20.90, 24.82]	[−3.03, 2.73]

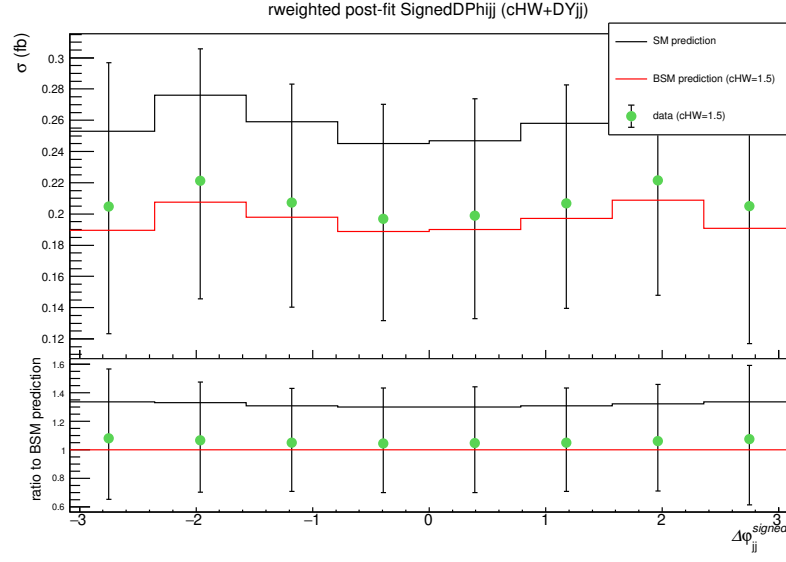
Table 22: Best-fit values to the SM Asimov dataset, and the 2 σ intervals obtained from profile likelihood scans of the c_{Hq3} group of Wilson coefficients under linear-only and linear plus quadratic parametrization.

D EFT bias test

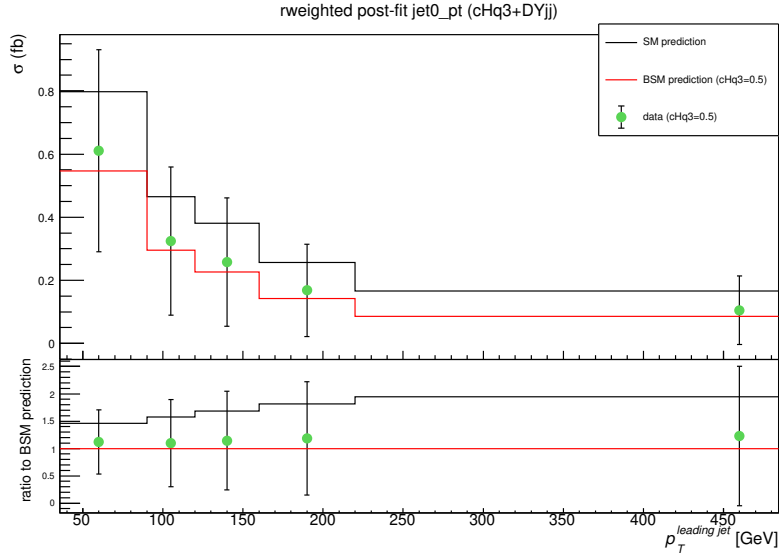
Despite the fact that BDTs are developed to contain the minimal information from MC predictions while retaining the sensitivity to signal, we need to test if the BDT values are biased towards SM prediction. We conducted fits with SM plus BSM data. Due to the lack of EFT samples at reconstruction-level, the BSM data is made by reweighting the signal to SM plus EFT at a large value of Wilson coefficient. The values are chosen to be close to twice as large as the 1σ expected limit for each Wilson coefficient. The reweighting spectrum is chosen to be some of the most important variables in the BDT training. The first BSM test only reweighted MC events based on one variable, which is considered very conservative as it distort the correlation with other observables and will induce significant change in the BSM data. The result for this 1-dimensional reweighting is shown in Fig. 73.

We also reweighted according to two highly correlated BDT input variables at the same time (2-dimensional reweighting), since BSM physics might also modify the correlations of the BDT input variables. The example shown in Fig. 74 is based on a reweighting according to the m_{jj} and Δy_{jj} distributions.

In summary, we observe a maximum 10-20% bias in the case of BSM modeled by EFT where the Wilson coefficients chosen have sufficiently surpassed the values we can put constraints at. Despite the very large values of the Wilson coefficient chosen (which produce distributions that are incompatible with the SM assumption in the template used), the bias is well within the statistical uncertainty of our analysis. Therefore, we conclude for this test that the BDT does not introduce bias in the case of BSM.

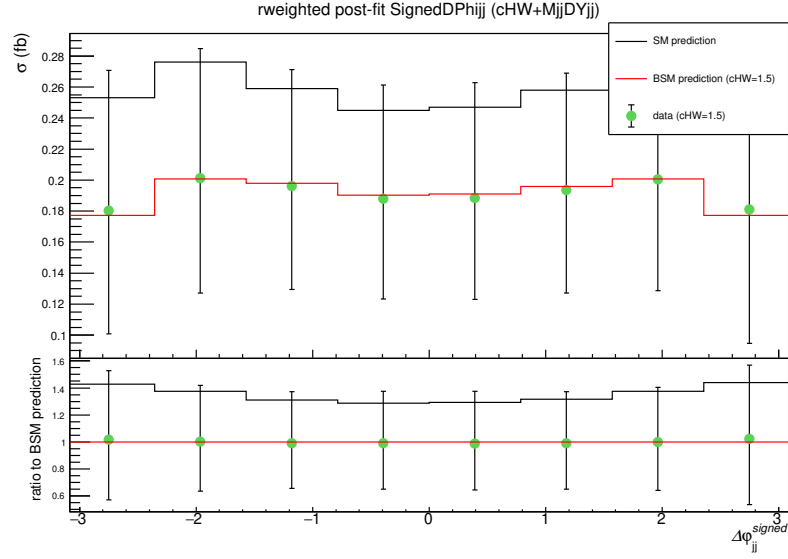


(a)

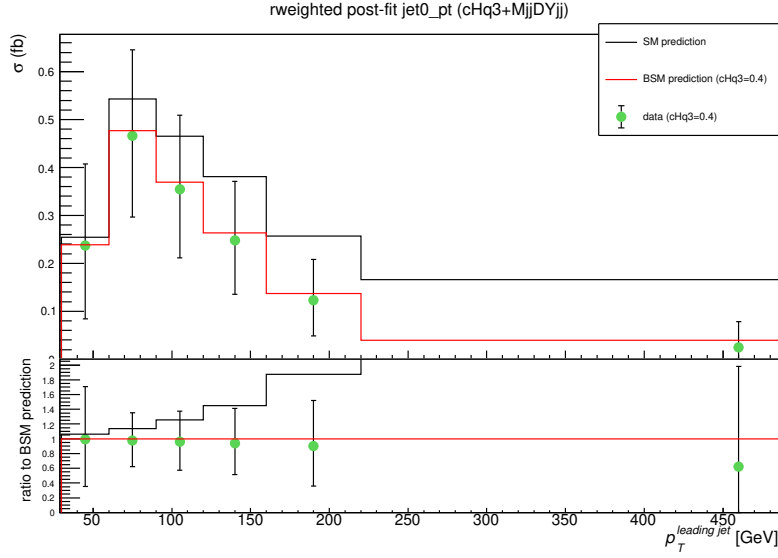


(b)

Figure 73: The post-fit result of the nominal fit model with the BSM data is compared with the BSM prediction at particle-level. The Wilson coefficient and its value used to generate the BSM data are: $c_{HW} = 1.5$, $c_{Hq3} = 0.4$. The BSM reweighting variable is $s\Delta y_{jj}$.



(a)



(b)

Figure 74: The post-fit result of the nominal fit model with the BSM data is compared with the BSM prediction at particle-level. The Wilson coefficient and its value used to generate the BSM data are: $c_{HW} = 1.5$, $c_{Hq3} = 0.4$. The BSM reweighting variables are m_{jj} plus Δy_{jj} .