

Searching for direct and indirect signatures of new physics with measurements of the Higgs boson

Charlotte Rose Knight

Imperial College London
Department of Physics

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy

Abstract

This thesis describes two searches for new physics using measurements of the Higgs boson at the Compact Muon Solenoid (CMS) experiment. The first search looks for the direct production of a new boson, X , which decays to two Higgs bosons, or a Higgs boson and a new scalar boson, Y . This search is performed in final states including two photons and two tau leptons, and considers m_X between 260 and 1000 GeV, and m_Y between 50 and 800 GeV. The second search considers a scenario where new particles have masses too large to be produced on-shell at the LHC but where their existence can be probed instead by their indirect effects on measurements of single Higgs boson production. In this search, these effects are interpreted in the Standard Model Effective Field Theory (SMEFT).

The searches use data collected by the CMS experiment at the LHC between 2016 and 2018 in proton-proton collisions at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 138 fb^{-1} . In the direct search, an excess consistent with others seen by the CMS experiment is observed at $(m_X, m_Y) = (650, 95) \text{ GeV}$ with a local significance of 2.3 standard deviations, and other local excesses are observed up to 3.2 standard deviations, depending on m_X and m_Y . Upper limits are also reported on the production cross sections, and ranges of m_X and m_Y are excluded assuming different theoretical scenarios. In the indirect search, individual constraints on up to 43 SMEFT Wilson coefficients (WCs) are reported, where a discrepancy with respect to the SM prediction is observed for the $C_{Hq}^{(3)}$ WC, corresponding to a p-value of 0.01. Additionally, the simultaneous constraints on up to 17 linear combinations of WCs are reported.

Statement of Originality

The work presented in this thesis is described in my words and references are cited where appropriate. The originality of the content is dependent on the chapter. Chapters 1 and 7 provide an introduction and conclusion to the thesis and are entirely original.

Chapters 2 to 4 provide background information that supports the work presented in the rest of the thesis. The content in these chapters is mostly not original, but is presented in my own words.

Chapter 5 presents a resonant search for two scalar bosons performed by the CMS collaboration in the $\gamma\gamma\tau\tau$ final state. This work has been partly documented in Ref. [1] and a more complete description is provided in this thesis. I was predominantly responsible for this analysis, performing all work described in Chapter 5 except for the triggering and preselection described in Section 5.4.

Chapter 6 presents a SMEFT interpretation of single Higgs boson production measurements made by the CMS experiment. This work has been partly documented in Ref. [2]. A more detailed description of the derivation of the EFT parameterization, which I was responsible for, is given in Section 6.2. The measurements themselves, the combination of them, and the extraction of the final results, were performed by other members of the CMS collaboration.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

I would like to thank the Imperial College London High Energy Physics group and the Science and Technology Facilities Council for giving me the opportunity to carry out this research. Thank you to my supervisor, Nick, for your support and invaluable insight into the world of particle physics. To Jon, thank you for your care and sustained and plentiful help throughout the PhD. To everyone else in the group and the wider CMS collaboration who has contributed along the way, thank you.

To my friends, thank you for the love and happiness that you brought and continue to bring into my life. In particular, I'd like to acknowledge Alie, Alina, Flax, and Victoria. During the more difficult times, our relationships kept me going, and I could not have completed this journey without you. To Hanae, I am so incredibly grateful for our friendship. You have been a constant source of support and joy, and meeting you has been one of the best things to come out of this PhD. To Mum, Dad and Laura, thank you for your life-long support and encouragement. I love you all.

Contents

1	Introduction	19
2	Theory	21
2.1	The Standard Model of Particle Physics	22
2.1.1	Particle Content	22
2.1.2	Interactions	24
2.1.3	Quantum Field Theory	26
2.1.4	The Standard Model Lagrangian	33
2.2	Higgs Boson Phenomenology	42
2.2.1	Higgs Boson Production and Decay Modes	42
2.2.2	Simplified Template Cross Sections	44
2.3	Theories Beyond the Standard Model	52
2.3.1	Warped Extra Dimensions	52
2.3.2	The Next-to-Minimal-Supersymmetric Standard Model	54
2.3.3	Effective Field Theory	57
3	The CMS Experiment	69
3.1	The Large Hadron Collider	69
3.2	The CMS Detector	72
3.2.1	Introduction	72
3.2.2	Coordinate System	74
3.2.3	Inner Tracker	74
3.2.4	Electromagnetic Calorimeter	76
3.2.5	Hadronic Calorimeter	78
3.2.6	Muon System	80
3.2.7	Trigger	82
3.3	Object Reconstruction	83
3.3.1	Tracking	83
3.3.2	Calorimeter Clustering	84

3.3.3	Link Algorithm	85
3.3.4	Muons	85
3.3.5	Electrons and Photons	88
3.3.6	Jets	94
3.3.7	Tau Leptons	95
3.3.8	Summary	102
4	Statistical Procedures	103
4.1	Likelihood Function	104
4.1.1	Counting Analyses	104
4.1.2	Shape Analyses	105
4.1.3	Systematic Uncertainties	106
4.2	Inference	108
4.2.1	Significances	108
4.2.2	Limit Setting	109
4.2.3	Maximum Likelihood Estimates and Confidence Intervals	110
4.2.4	Binned χ^2 fits	110
5	Search for Di-Higgs in Final States with Photons and Tau Leptons	113
5.1	Introduction	113
5.2	Analysis Strategy	117
5.3	Samples	120
5.4	Trigger and Preselection	122
5.5	Event Selection	127
5.5.1	Training Features	127
5.5.2	Parametric Neural Networks	140
5.5.3	Training and Performance	143
5.5.4	Category Optimization	157
5.5.5	Sculpting of the Diphoton Mass Distribution	162
5.6	Search Granularity	167
5.7	Sources of Systematic Uncertainty	174
5.7.1	Theoretical Uncertainties	174
5.7.2	Experimental Uncertainties	174
5.8	Modelling	176
5.8.1	Single and Di-Higgs Production	176
5.8.2	Nonresonant Background	187
5.8.3	Resonant Background from Electron Misidentification	192
5.9	Results	198

5.9.1	$X \rightarrow HH$ Searches	198
5.9.2	$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Search	202
5.9.3	Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	208
5.9.4	High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	215
5.10	Summary	220
6	Effective Field Theory Interpretation	223
6.1	Introduction	223
6.2	SMEFT Parameterization	227
6.2.1	Overview and Assumptions	227
6.2.2	General Methodology	228
6.2.3	Acceptance Corrections	231
6.2.4	Summary	234
6.3	Results in the Nominal Basis	239
6.4	Results in the Rotated Basis	241
6.5	Summary	248
7	Conclusion	249

List of Figures

2.1	SM Particle Content	23
2.2	Electroweak Feynman Diagram Vertices	24
2.3	Three-Point Electroweak Boson Feynman Diagram Vertices	25
2.4	Four-Point Electroweak Boson Feynman Diagram Vertices	25
2.5	Strong Feynman Diagram Vertices	25
2.6	Higgs Field Potential	31
2.7	Measurements of α_s as a Function of Energy Scale	35
2.8	LO Feynman Diagrams for the Dominant Higgs Boson Production Modes	43
2.9	LO Feynman Diagrams for the Subdominant Higgs Boson Production Modes	43
2.10	LO Feynman Diagrams for $H \rightarrow gg$ and $H \rightarrow \gamma\gamma$ Decays	44
2.11	Stage 0 Binning in the STXS	46
2.12	Measurements of the Stage 0 STXS	46
2.13	Stage 1.2 ggH Binning in the STXS	47
2.14	Stage 1.2 VBF Binning in the STXS	47
2.15	Stage 1.2 VH Binning in the STXS	48
2.16	Stage 1.2 ttH Binning in the STXS	48
2.17	Measurements of the Stage 1.2 STXS	50
2.18	Measurements of the Stage 1.2 STXS Split by Decay Channel	51
2.19	Graviton and Radion Branching Fractions	53
2.20	Graviton and Radion Production Cross Sections	54
2.21	Feynman Diagram for a CP-Odd Higgs Boson Decaying to Two Photons Via a Chargino Loop	55
2.22	Maximally-Allowed Cross Sections for $pp \rightarrow A \rightarrow H_{125}(\rightarrow \tau\tau)a(\rightarrow \gamma\gamma)$ in the NMSSM	56
2.23	LO Feynman Diagrams for Muon Decay in the SM and in Fermi Theory .	57
2.24	LO Feynman Diagrams for ggH Production in the SM and the SMEFT .	65
2.25	LO Feynman Diagram for the $H \rightarrow 4l$ Decay in the SM and a New Diagram from \mathcal{L}_6 Contributions	67

2.26	LO Feynman Diagrams for the $H \rightarrow 4l$ Decay that Include $hZ\gamma$ and $h\gamma\gamma$ Vertices	67
3.1	The CERN Accelerator Complex	70
3.2	Delivered and Recorded Luminosity at CMS, Cumulative Over 2015 to 2018	71
3.3	A Schematic of the CMS Detector	73
3.4	The CMS Inner Tracker	75
3.5	The CMS ECAL	77
3.6	The CMS HCAL	79
3.7	The CMS Muon System	81
3.8	Medium Muon ID Efficiency in 2018	87
3.9	Muon Momentum Resolution as a Function of p_T	88
3.10	Electron and Photon Identification Performance	90
3.11	Signal Efficiencies of Electron and Photon Identification Algorithms for the WP90 Working Point, Measured in Data and Simulation	91
3.12	Electron Energy Corrections	93
3.13	Electron Energy Scale Agreement Between Simulation and Data	94
3.14	τ_h Reconstruction and Identification Efficiency	98
3.15	Performance of the DeepTau ID Algorithm	99
3.16	τ_h Energy Scale Corrections	100
3.17	τ_h Energy Scale Agreement Between Simulation and Data	101
5.1	LO Feynman Diagrams for Nonresonant and Resonant Di-Higgs Production	113
5.2	SM Di-Higgs Branching Fractions	114
5.3	Signal and Background $m_{\gamma\gamma}$ Distributions in the Di-Higgs Search	118
5.4	Nominal Mass Point Granularity	121
5.5	Preselection Efficiency as a Function of m_X in the $X \rightarrow HH$ Searches	125
5.6	Preselection Efficiency as a Function of m_X and m_Y in the $X \rightarrow YH$ Searches	126
5.7	Correlation Between the Most Important Training Features	131
5.8	Distributions of Training Features (1)	133
5.9	Distributions of Training Features (2)	134
5.10	Distributions of Training Features (3)	135
5.11	Distributions of Training Features (4)	136
5.12	Distributions of Training Features (5)	137
5.13	Distributions of Training Features (6)	138
5.14	Distributions of Training Features (7)	139
5.15	Illustration of pNN Interpolation Ability	142
5.16	Architecture for the pNN	144

5.17	Evolution of BCE Loss for the $X^{(2)} \rightarrow HH$ pNN Split by Mass Point . . .	147
5.18	Evolution of BCE Loss for the $X^{(0)} \rightarrow HH$, and $X \rightarrow YH$ pNNs	148
5.19	ROC Curves for the $X^{(2)} \rightarrow HH$ pNN	149
5.20	AUC Scores for the $X \rightarrow HH$ pNNs	150
5.21	AUC Scores for the $X \rightarrow YH$ pNNs	151
5.22	pNN Validation Tests in the $X^{(2)} \rightarrow HH$ Search	153
5.23	pNN Validation Tests in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Search	154
5.24	pNN Validation Tests in the Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	155
5.25	pNN Validation Tests in the High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	156
5.26	Transformed pNN Output Score Distribution in $X^{(2)} \rightarrow HH$ Search	158
5.27	Transformed pNN Output Scores in the $X^{(0)} \rightarrow HH$ and $X \rightarrow YH$ Searches	159
5.28	Performance of a Grid Search of Category Boundaries at Every Mass Point Compared to a Single Point	160
5.29	Performance of a Grid Search of Category Boundaries at Every Mass Point Compared to a New Categorization Approach	161
5.30	Distribution of $m_{\gamma\gamma}$ in Nonresonant Background Samples After Preselection	163
5.31	Evidence of No $m_{\gamma\gamma}$ Sculpting in $X^{(2)} \rightarrow HH$ at $m_X = 550$ GeV	164
5.32	Evidence of $m_{\gamma\gamma}$ Sculpting in High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search When Including $m(\gamma\gamma\tau\tau)$ and $m(\tau\tau\gamma_1)$ as Training Features	165
5.33	Evidence of No $m_{\gamma\gamma}$ Sculpting in High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search When Excluding $m(\gamma\gamma\tau\tau)$ and $m(\tau\tau\gamma_1)$ as Training Features	166
5.34	Relative Changes in Expected Upper Limits When Using Categories Not Designed for the Targeted Mass Point	168
5.35	Search Granularity for the $X \rightarrow HH$ Searches	169
5.36	Search Granularity for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Search	171
5.37	Search Granularity for the Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	172
5.38	Search Granularity for the High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	173
5.39	DCB Signal Fits in $X \rightarrow YH$ Searches	178
5.40	Signal Efficiencies as a Function of m_X in $X^{(2)} \rightarrow HH$ Search	180
5.41	Signal Efficiency Splines for $m_X = 375$ GeV in the $X^{(2)} \rightarrow HH$ Search . . .	181
5.42	Signal Efficiency Splines for $m_X = 433$ GeV in the $X^{(2)} \rightarrow HH$ Search . . .	182
5.43	DCB Shape Parameters as a Function of m_X in the $X^{(2)} \rightarrow HH$ Search . .	183
5.44	Validation of Shape Parameter Interpolation	184
5.45	Toy Examples of Linear and Cubic Interpolation in 2D	185
5.46	Signal Efficiency as a Function of m_Y in the $X \rightarrow YH$ Searches	185
5.47	DCB Shape Parameters as a Function of m_Y in the $X \rightarrow YH$ Searches . .	186

5.48	Distribution of $m_{\gamma\gamma}$ in Data and Background MC After Preselection for the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Searches	188
5.49	Envelope Construction for $(m_X, m_Y) = (300, 50)$ GeV in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Search	190
5.50	Envelope Construction for $(m_X, m_Y) = (300, 70)$ GeV in the Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	191
5.51	Envelope Construction for $(m_X, m_Y) = (1000, 700)$ GeV in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search	192
5.52	Distributions of $m_{\gamma\gamma}$ in Different Regions of the DY ABCD Method	194
5.53	Correlation Check for the DY ABCD Method	196
5.54	Expected DY Events Extracted Using ABCD Method	197
5.55	Signal-Plus-Background Fits to Data for $X^{(2)} \rightarrow HH$ Search at $m_X = 375$ GeV	199
5.56	$X \rightarrow HH$ Upper Limits	201
5.57	$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Observed Local Significances	203
5.58	Signal-Plus-Background Fits to Data for $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Search at $(m_X, m_Y) = (320, 60)$ GeV	204
5.59	$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Upper Limits in the 2D (m_X, m_Y) Plane	205
5.60	$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Upper Limits in as Function of m_Y in Slices of m_X	206
5.61	$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Upper Limits in as Function of m_X in Slices of m_Y	207
5.62	Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Observed Local Significances	209
5.63	Signal-Plus-Background Fits to Data for Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search at $(m_X, m_Y) = (525, 115)$ GeV	210
5.64	Signal-Plus-Background Fits to Data for Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search at $(m_X, m_Y) = (650, 95)$ GeV	211
5.65	Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Upper Limits in the 2D (m_X, m_Y) Plane	212
5.66	Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Upper Limits in as Function of m_Y in Slices of m_X	213
5.67	Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Upper Limits in as Function of m_X in Slices of m_Y	214
5.68	High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Observed Local Significances	215
5.69	Signal-Plus-Background Fits to Data for High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search at $(m_X, m_Y) = (462, 161)$ GeV	216
5.70	High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Upper Limits in the 2D (m_X, m_Y) Plane	217
5.71	High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Upper Limits in as Function of m_Y in Slices of m_X	218
5.72	High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Upper Limits in as Function of m_X in Slices of m_Y	219

6.1	Illustration of EFT Effects in High-Energy Tails of Distributions	224
6.2	Distribution of m_{34} in the SM and in the SMEFT with $C_{HB} = 0.5$	232
6.3	Impact of Acceptance Corrections on the $H \rightarrow 4l$ Partial Width	233
6.4	Impact of SMEFT Operators in the SMEFT Basis (1)	236
6.5	Impact of SMEFT Operators in the SMEFT Basis (2)	237
6.6	Impact of SMEFT Operators in the SMEFT Basis (3)	238
6.7	Individual Constraints on the SMEFT Wilson Coefficients	240
6.8	PCA Illustration	242
6.9	Truncated PCA Rotation Matrix	243
6.10	Impact of SMEFT Operators in the Rotated Basis (1)	244
6.11	Impact of SMEFT Operators in the Rotated Basis (2)	245
6.12	Simultaneous Constraints on the SMEFT Eigenvectors	247

List of Tables

2.1	Free Parameters in the SM	41
2.2	Cross Sections for Higgs Boson Production Modes	42
2.3	Branching Fractions for Higgs Boson Decay Modes	44
2.4	Warsaw Basis of Operators	59
2.5	topU31 Fermionic Operator Basis	61
3.1	Tau Lepton Decays and Branching Fractions	96
3.2	Identification Efficiencies at Different DeepTau ID Working Points	97
5.1	η -dependent Preselection Requirements	123
5.2	Preselection Requirements for Electrons, Muons, τ_h , and Jets	123
5.3	Training Features Considered for pNN Training Before Feature Selection	128
5.4	Most Important pNN Training Features	129
5.5	Performance Comparison Between Full Set of Training Features and Subset of Most Important Features	130
5.6	Final pNN Training Features	132
5.7	Numbering Scheme Used to Describe the Channels in the Di-Higgs Analysis	140
5.8	Values of Hyperparameters Used in a Grid Search to Optimize the pNN Architecture	143
5.9	Sum of the Number of Background and Signal Events Across All Simulated Datasets	145
5.10	Analysis Category Definitions in Di-Higgs Search	162
5.11	Number of Mass Points in Each Di-Higgs Search	170
5.12	Function Families Used to Model the Nonresonant Background in the $X \rightarrow$ HH and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Searches	188
5.13	Summary of Results from $X \rightarrow HH$ and $X \rightarrow YH$ Searches	221
6.1	Summary of the Analyses in the Combination Used for the SMEFT Inter- pretation	225

6.2	SMEFT Interpretation Methodology Overview	229
6.3	Contributing Wilson Coefficients to SMEFT Interpretation	235

Chapter 1

Introduction

It is the goal of particle physics to formulate a theory that accurately predicts the fundamental constituents of matter and their interactions. The Standard Model (SM) of particle physics [3–5] is the most successful theory to date, describing the electromagnetic, weak, and strong interactions. At its inception, the SM predicted the existence of a number of unseen fundamental particles, all of which have since been discovered. The last of these particles, the Higgs boson, was discovered in 2012 at the Large Hadron Collider (LHC) at CERN [6, 7] and marked the completion of the SM.

Despite its many successes, the SM fails to describe several key observations including dark matter [8], neutrino oscillations [9], and the matter-antimatter asymmetry of the universe [10]. Therefore, the search for a theory beyond the Standard Model (BSM) that can also explain these observations is a key goal of modern particle physics. There are already a great number of theories that have been proposed, but it is not clear which theory is the correct one, so particle physics experiments continue to look for signatures of new physics that could point towards a particular theory or family of theories.

At the LHC, data continues to be taken in the search for new physics, now at higher collision energies and more intense luminosities than were needed for the Higgs boson discovery. Searches can be generally grouped into two categories: direct and indirect. Direct searches look for the on-shell production of a BSM particle and, if successful, would provide a measurement of the mass *and* couplings of the new particle. Indirect searches, on the other hand, look for deviations from the SM predictions in the measurements of SM particles and processes, and would only provide information about the ratio of the new particle’s couplings to its mass. Therefore, a direct observation of new physics is usually preferred, and often considered more convincing evidence. However, if the mass of the new particle is so large that it cannot be produced on-shell at the LHC, then indirect searches are the only option, making direct and indirect searches complementary to each other.

This thesis describes both a direct and an indirect search for new physics involving measurements of the Higgs boson at the LHC, specifically at the Compact Muon Solenoid (CMS) experiment, which is one of four major experiments situated around the LHC ring. The direct search is for a new boson, X , which decays to two SM Higgs bosons ($X \rightarrow HH$), or to a new scalar boson, Y , and to a SM Higgs boson ($X \rightarrow YH$). In this thesis, this search will be referred to as the di-Higgs search. The indirect search uses measurements of Higgs boson cross sections, combined across Higgs boson decay channels, and interprets them in an Effective Field Theory (EFT) framework, which provides constraints on a variety of BSM scenarios.

In Chapter 2, an introduction to the SM is provided, with a focus on the Higgs boson and its phenomenology. This chapter also introduces EFTs, and the BSM theories that motivate the direct search for $X \rightarrow HH$ and $X \rightarrow YH$. In Chapter 3, the LHC and CMS experiment are described, and in Chapter 4, the statistical methods used in the searches are discussed. Finally, Chapters 5 and 6 describe the direct and indirect searches respectively, and a conclusion is drawn in Chapter 7.

Chapter 2

Theory

The Standard Model (SM) of particle physics describes interactions between all known elementary particles via the strong, weak, and electromagnetic forces. It is capable of making astonishingly precise and accurate predictions. For example, the prediction and measurement of the anomalous magnetic dipole moment of the electron agree to one part in 10^{12} [11]. Furthermore, measurements made of high energy collisions, like those at the LHC, remain consistent with SM predictions. However, the SM cannot explain several key observations, including dark matter [8], neutrino oscillations [9], and the matter-antimatter asymmetry of the universe [10]. It is anomalies like these that motivate the formulation of theories that go beyond the SM.

Anomalous observations are not the only source of motivation when developing new theories. There are also theoretical problems with the SM such as the hierarchy problem [12] and that the SM does not include a description of gravity. From a scientific perspective, a BSM theory does not need to solve these problems, it only needs to provide predictions that accurately describe our measurements. However, history teaches us that when pursuing solutions to theoretical problems, we can also make scientific breakthroughs. The prediction of the charm quark in 1970 [13] before its discovery in 1975 [14] is a great example of this.

This chapter begins with an explanation of the SM, starting with the particle content and the allowed interactions before moving onto a mathematical description of the theory. Following this, the content is tailored towards the results discussed later in this thesis, those being the di-Higgs search in Chapter 5 and the Effective Field Theory (EFT) interpretation in Chapter 6. This content includes Higgs boson phenomenology at the LHC and relevant extensions to the SM, including EFTs and two BSM theories that motivate the search in Chapter 5, namely Warped Extra Dimensions (WED) and the Next-to-Minimal Supersymmetric Standard Model (NMSSM).

2.1 The Standard Model of Particle Physics

The SM is formulated as a relativistic quantum field theory (QFT) where each particle in the theory is represented by its own field and excitations of those fields are then the physical particles that we observe in nature. Particles can interact with each other via the exchange of a subset of SM particles called *force carriers*. The types of allowed interactions and their corresponding strengths are encoded in a Lagrangian density, \mathcal{L}_{SM} , together with the particle properties (mass and spin).

2.1.1 Particle Content

A schematic of the SM particle content, including every SM particle with its properties and categorizations is provided in Fig. 2.1. The particle content can be initially categorized into two groups: spin- $\frac{1}{2}$ fermions which are matter constituents, and spin-integer bosons which are force carriers and are spin-1 except for the Higgs boson which is spin-0.

There are four spin-1 bosons in the SM: the photon (γ) which mediates the electromagnetic force, the W^\pm and Z bosons which mediate the weak force, and the gluon (g) which mediates the strong force. Particles that interact with photons carry *electric charge*; particles that interact with gluons carry a *colour charge*, which comes in three possible states: r , g and b , and particles that interact with W^\pm and Z bosons carry charges called *weak isospin* and/or *weak hypercharge*. More details on these charges are given in Section 2.1.4.

The fermions are split by those that interact with the strong force, called *quarks*, and those that do not, called *leptons*. Leptons are further split into those which are electrically charged (l), and those which are not, called *neutrinos* (ν). Quarks are similarly split by electric charge into up-type quarks and down-type quarks, with charge $\frac{2}{3}$ and $-\frac{1}{3}$ respectively (in units of elementary charge, e).

The fermions can be also categorized into three generations based on a mass hierarchy, where the first generation is the least massive. Each generation contains an up-type quark, a down-type quark, a charged lepton, and a neutrino. The first generation contains the main constituents of the visible matter in the universe, those being the up and down quarks, and the electron. Additionally, all charged fermions have a corresponding antiparticle which has the same mass but opposite charge and parity.

The remaining particle is the Higgs boson which plays a special role in the SM. We will see later on that its introduction to the model is necessary to correctly describe the distribution of masses for particles in the theory.

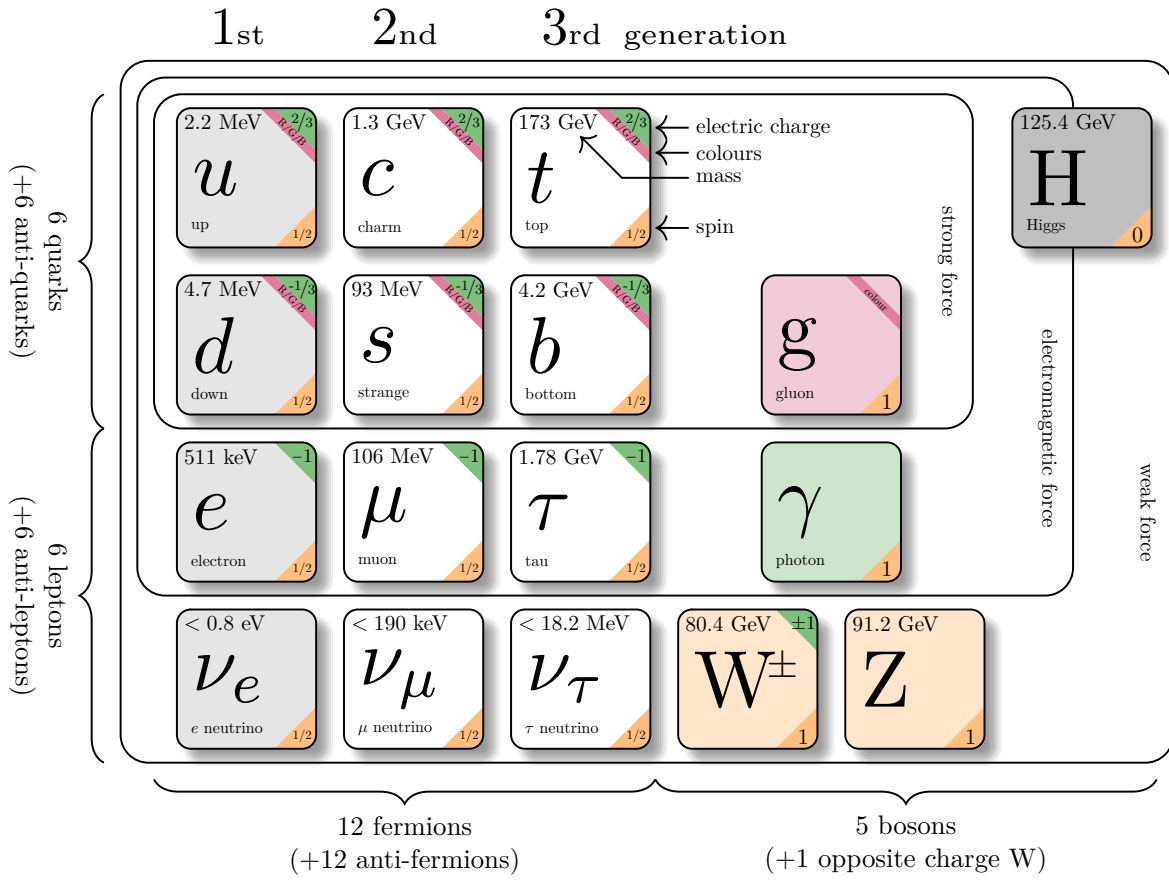


Figure 2.1: Particle content of the SM. The mass, electric charge, colour and spin are given for each particle. All masses are taken from Ref. [15] except the Higgs boson mass which is taken from Ref. [16].

2.1.2 Interactions

The electromagnetic and weak forces are described together by a single force: the *electroweak* force. Particles that interact via the electroweak force include all fermions and the electroweak bosons: Z , W , γ and H (all bosons except the gluon). Feynman diagram vertices for these interactions are given in Fig. 2.2. All fermions can interact via the exchange of Z boson and all left-handed particles/right-handed anti-particles can interact via the exchange of a W boson. All electrically charged fermions can interact via the exchange of a photon and all massive fermions can interact with the Higgs boson. The three and four-point interactions involving only electroweak bosons, shown in Figs. 2.3 and 2.4, are also allowed. In the strong force, quarks and gluons interact with each other according to the Feynman diagram vertices shown in Fig. 2.5.

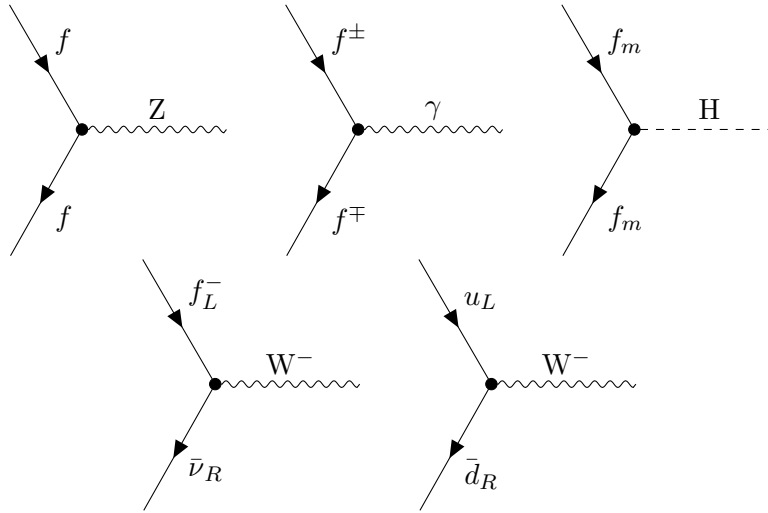


Figure 2.2: Feynman diagram vertices allowed in the SM involving electroweak interactions with fermions. Interactions with a photon, Z boson, or a Higgs boson, require the fermion (f) flavour to be the same. Massive fermions are denoted by f_m and exclude neutrinos. Interactions with W bosons can involve fermions from different generations and must involve either a lepton (l) and a neutrino (ν) or an up-type quark (u) and a down-type quark (d). The L and R subscripts denote left-handed and right-handed fermions respectively. Charged-conjugated versions of the W -interaction diagrams are also possible if the handedness is reversed.

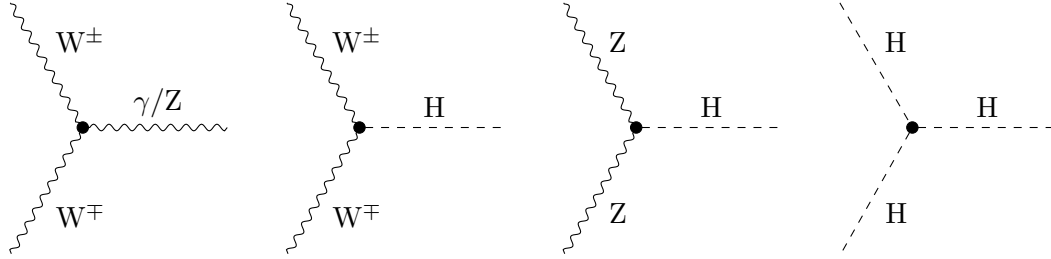


Figure 2.3: Feynman diagram vertices allowed in the SM involving three electroweak bosons.

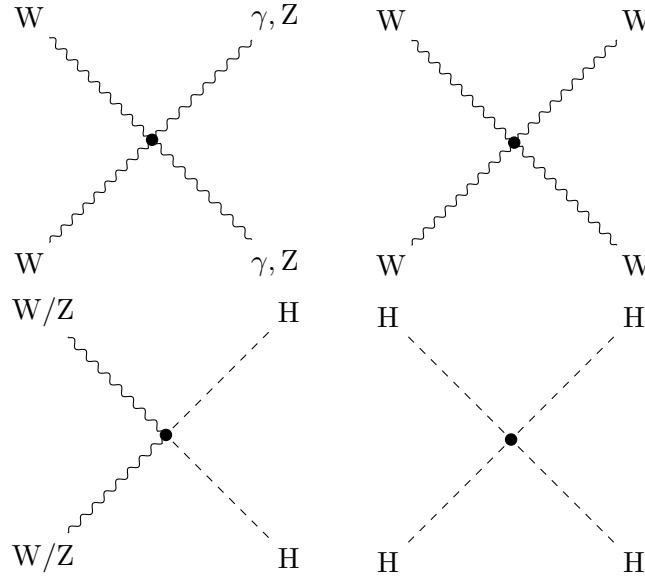


Figure 2.4: Feynman diagram vertices allowed in the SM involving four electroweak bosons.

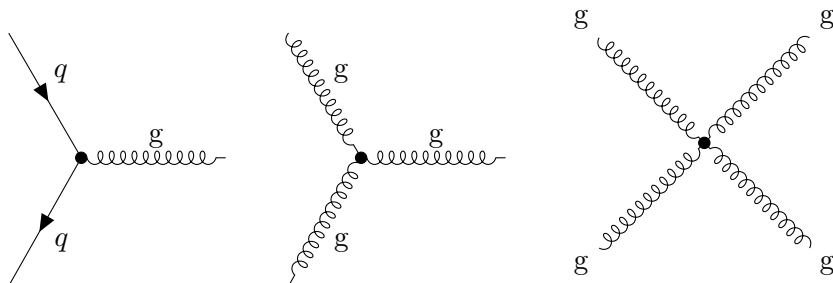


Figure 2.5: Feynman diagram vertices allowed in the SM that involve a strong interaction. On the left, q refers to a quark (or antiquark) of the same flavour.

2.1.3 Quantum Field Theory

The SM can be described by a Lagrangian density, \mathcal{L}_{SM} , which is a function of the particle fields and their derivatives. By applying the Euler-Lagrange equations to \mathcal{L}_{SM} , we can then derive equations of motion that describe the dynamics and interactions of these fields. In this section, concepts related to QFT that will help to understand the construction of \mathcal{L}_{SM} will be described. In the rest of this section, Lagrangian densities will be represented by \mathcal{L} and referred to as “Lagrangians” (omitting “density”).

Real Scalar Fields

One of the simplest particle theories one could write down consists of a single Lorentz real scalar field, $\phi(x, y, z, t)$, with a Lagrangian:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 \quad (2.1)$$

and after applying the Euler-Lagrange equations:

$$0 = \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \quad (2.2)$$

this becomes:

$$0 = \partial_\mu \partial^\mu \phi + m^2 \phi \quad (2.3)$$

which can be identified as a relativistic wave equation, namely the Klein-Gordon equation, which describes a particle with mass m . In this theory, the particles are non-interacting. To introduce interactions, we need to add terms that are higher orders of ϕ , for example:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 + \lambda_3 \phi^3 + \lambda_4 \phi^4 \quad (2.4)$$

where $\lambda_3 \phi^3$ and $\lambda_4 \phi^4$ correspond to three-point and four-point interactions respectively. Generally, a term containing greater than two fields, including powers of the same field, corresponds to an interaction involving that combination of fields. Adding terms that have powers of ϕ greater than four are not allowed if one wants a renormalizable theory, i.e. a theory that has finite predictions at arbitrarily high energy scales [17]. Strictly speaking, a renormalizable Lagrangian must have dimensions $[\mathcal{L}] = \text{GeV}^4$. Given that $[\partial_\mu \phi] = \text{GeV}^2$, $[\phi] = \text{GeV}$ and $[m] = \text{GeV}$, this is satisfied in Eq. 2.4.

Regardless of renormalizability, a Lagrangian must additionally:

1. be real since the action, S , which is an integral over the Lagrangian, must be real;
2. be a local function of the fields and their derivatives calculated at the same spacetime point;
3. be invariant under any symmetries of the theory.

Equation 2.4 is invariant under Poincaré transformations, which include translations and rotations in space, as well as Lorentz boosts. According to Noether's theorem, a continuous symmetry has a corresponding conservation law and in the case of Poincaré transformations, the conserved quantities are energy, momentum, and angular momentum. The full SM Lagrangian has additional symmetries which will be highlighted throughout the rest of this section.

Lorentz invariance is an assumed requirement for all Lagrangians in this chapter. The condition on renormalization is also required except for effective field theories which will be discussed in Section 2.3.3.

Vector Fields

In the SM, the spin-1 force carriers are represented by vector fields. The most general Lagrangian that is Lorentz invariant and contains a real vector field is:

$$\mathcal{L} = aS^2 + bF_{\mu\nu}F^{\mu\nu} + cG_{\mu\nu}G^{\mu\nu} + dA_\mu A^\mu \quad (2.5)$$

where:

$$S = \partial_\mu A^\mu, \quad \text{is a Lorentz scalar;} \quad (2.6)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad \text{is an antisymmetric rank 2 tensor;} \quad (2.7)$$

$$G_{\mu\nu} = \partial_\mu A_\nu + \partial_\nu A_\mu - \frac{1}{2}\eta_{\mu\nu}S, \quad \text{is a symmetric and traceless rank 2 tensor;} \quad (2.8)$$

and a, b, c and d are arbitrary constants. If we want the Lagrangian to also be gauge invariant, meaning that it is invariant under transformations of the form:

$$A_\mu \rightarrow A_\mu - \partial_\mu \lambda \quad (2.9)$$

where λ is an arbitrary scalar function of spacetime, only one of the terms is viable and the most general Lagrangian is:

$$\mathcal{L} = F_{\mu\nu}F^{\mu\nu}. \quad (2.10)$$

This is the Maxwell Lagrangian for the free electromagnetic field.

Complex Scalar Fields

Returning to scalar fields but now considering the field to be complex, the most general Lagrangian becomes:

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - V(\phi^* \phi) \quad (2.11)$$

where the potential, $V(\phi^* \phi)$, is a polynomial of order two or less. The three-point interaction term is now missing due to the requirement that \mathcal{L} be real, and the Lagrangian has a new symmetry compared to the real scalar field Lagrangian. The Lagrangian is now invariant under the rotation of a complex phase:

$$\phi(x) \rightarrow e^{i\theta} \phi(x) \quad (2.12)$$

where θ is a real number. The symmetry is called *Abelian* because the transformations commute:

$$e^{-i\theta_1} e^{-i\theta_2} = e^{-i\theta_2} e^{-i\theta_1}, \quad (2.13)$$

and is called a *global symmetry* because the same transformation is applied at all points in spacetime, i.e. θ is a constant.

The Lagrangian of Equation 2.11 is not invariant under a local transformation of the type:

$$\phi(x) \rightarrow e^{i\theta(x)} \phi(x) \quad (2.14)$$

but an invariant Lagrangian can be created if a new vector field, A_μ , is introduced which simultaneously transforms as

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{e} \partial_\mu \theta(x) \quad (2.15)$$

which is the same as the gauge transformation in Eq. 2.9 where $\lambda = \theta(x)/e$. The simultaneous transformation of ϕ and A_μ is referred to as a *local gauge* transformation and the new vector field is known as a *gauge field*. The Lagrangian is given by:

$$\mathcal{L} = (D_\mu \phi)^* D^\mu \phi - V(\phi^* \phi) \quad (2.16)$$

where the covariant derivative, D_μ , is defined as:

$$D_\mu \phi \equiv \partial_\mu \phi + ie A_\mu \phi. \quad (2.17)$$

Now adding the $F_{\mu\nu} F^{\mu\nu}$ term (Eq. 2.7), we arrive at the most general gauge invariant

Lagrangian for a complex scalar field:

$$\mathcal{L} = F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^* D^\mu\phi - V(\phi^*\phi). \quad (2.18)$$

In the SM, local gauge symmetries are imposed on the Lagrangian which give rise to the existence of vector bosons, which are the force carriers: the gluons, photon, and the W^\pm and Z bosons. Given that they are related to gauge transformations, the force carriers are also referred to as *gauge bosons*.

Non-Abelian Gauge Fields

The simultaneous transformation of ϕ and A_μ in Eqs. 2.14 and 2.15 is a $U(1)$ transformation. When generalizing to $U(N)$, the Lagrangian in Eq. 2.16 becomes:

$$\mathcal{L} = (D_\mu\phi)^\dagger D^\mu\phi - V(\phi^\dagger\phi) \quad (2.19)$$

where ϕ and A_μ transform as

$$\phi \rightarrow M\phi \quad (2.20)$$

$$A_\mu \rightarrow MA_\mu M^\dagger + \frac{i}{g}(\partial_\mu M)M^\dagger \quad (2.21)$$

where $M(x)$ is an element of the $U(N)$ group. Equation 2.21 holds if A_μ is an element of the Lie algebra, meaning that it can be written as $A_\mu = A_\mu^a T^a$ where A_μ^a are real constants, T^a are the generators of some representation of the Lie algebra, and $a \in \{1, \dots, D\}$ where $D = N^2$ is the dimensionality of $U(N)$. Therefore, imposing a local $U(N)$ symmetry has lead to the introduction of D gauge bosons, where each gauge boson is represented by a component, A_μ^a .

In the non-Abelian case, the $F_{\mu\nu}F^{\mu\nu}$ term is no longer gauge invariant. In an attempt to find a similar term that *is* gauge invariant, we redefine $F_{\mu\nu}$ as:

$$F_{\mu\nu} = -\frac{i}{g}[D_\mu, D_\nu] \quad (2.22)$$

$$= \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu] \quad (2.23)$$

which is still consistent with the Abelian case since $[A_\nu, A_\nu]$ would be zero. With this definition, $F_{\mu\nu}F^{\mu\nu}$ is still not invariant, but its trace is. Therefore, the most general scalar field Lagrangian with a $U(N)$ gauge symmetry is:

$$\mathcal{L} = -\frac{1}{2} \text{Tr } F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^\dagger D^\mu\phi - V(\phi^\dagger\phi). \quad (2.24)$$

Writing the first term alone and in terms of the gauge field components, A_μ^a , we find:

$$\mathcal{L} = -\frac{1}{2} \text{Tr} F_{\mu\nu} F^{\mu\nu} = -\frac{1}{4} F_{\mu\nu}^a F^{\mu\nu a} \quad (2.25)$$

$$= -\frac{1}{4} (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) (\partial_\nu A_\mu^a - \partial_\mu A_\nu^a) \quad (2.26)$$

$$+ \frac{g}{2} f^{abc} (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) A^{\mu b} A^{\nu c} - \frac{g^2}{4} f^{abc} f^{ade} A_\mu^b A_\nu^c A^{\mu d} A^{\nu e} \quad (2.27)$$

where the last two terms represent three-point and four-point interactions of the gauge bosons, in an analogous way to Eq. 2.4.

The conclusions reached here about non-abelian gauge fields apply to any subset of $U(N)$, including $SU(N)$ groups which are seen in the SM.

Spinors

In addition to the Klein-Gordon equation (Eq. 2.3), particles that have spin $\frac{1}{2}$ must also satisfy the Dirac equation:

$$i\gamma^\mu \partial_\mu \psi - m\psi = 0 \quad (2.28)$$

where m is the mass of the particle, ψ is the particle field, and γ^μ are the 4×4 gamma matrices [12].

A Lorentz scalar that can be constructed out of spinor fields is $\bar{\psi}\psi$ where we have introduced the Dirac adjoint, $\bar{\psi}$, as:

$$\bar{\psi} \equiv \psi^\dagger \gamma^0. \quad (2.29)$$

For any pair of spinors, ψ and χ , $\bar{\psi}\gamma^\mu\chi$ is a Lorentz vector, and for any Lorentz vector, a_μ , $\bar{\psi}\not{a}\chi$ is a Lorentz scalar where we have introduced the Dirac slash notation:

$$\not{a} \equiv \gamma^\mu a_\mu. \quad (2.30)$$

This also applies when the vector is a derivative so $\bar{\psi}\not{\partial}\psi$ is also a Lorentz scalar.

In the SM, spinors are decomposed into their left and right-handed components. To define these components, we first introduce a fifth gamma matrix:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \quad (2.31)$$

The left/right-handed component of a spinor, $\psi_{L/R}$, is given by $P_{L/R}\psi$ where:

$$P_L = \frac{1}{2}(\mathbb{1} - \gamma^5) \text{ and } P_R = \frac{1}{2}(\mathbb{1} + \gamma^5). \quad (2.32)$$

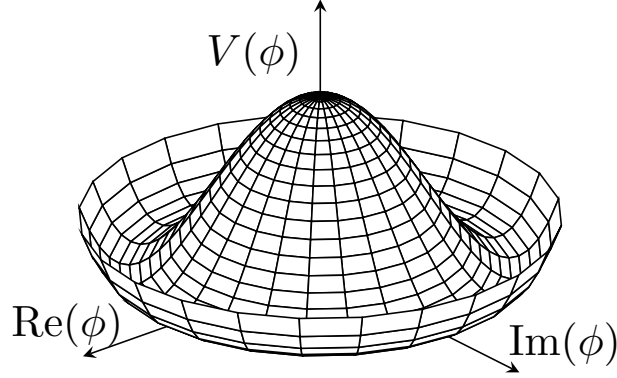


Figure 2.6: The potential of a complex scalar field, $V(\phi) = m^2\phi^*\phi + \frac{1}{2}\lambda(\phi^*\phi)^2$, where $m^2 < 0$ and $\lambda > 0$.

Under a parity transformation, a left/right-handed spinor transforms as:

$$\psi_{L/R} \rightarrow \gamma^0 \psi_{L/R} \quad (2.33)$$

and it can be shown that $P_{L/R}(\gamma^0 \psi_{L/R}) = 0$, i.e. a parity transformation changes a left-handed spinor into a right-handed spinor and a right-handed spinor into a left-handed spinor. Given this transformation property, we can create a parity-violating theory by writing a Lagrangian that has different terms for ψ_L than for ψ_R . This will be essential in describing the weak interaction that behaves differently with left and right-handed fermions.

Spontaneous Symmetry Breaking

So far, we have interpreted a field, ϕ , in a Lagrangian as the field of a physical particle. This has the implicit assumption that a field value of zero corresponds to the vacuum state of the Lagrangian, i.e. the state with the lowest energy. However, this need not be the case. Consider the complex scalar Lagrangian of Eq. 2.11, which has a global U(1) symmetry, with the potential:

$$V(\phi^*, \phi) = m^2\phi^*\phi + \frac{1}{2}\lambda(\phi^*\phi)^2. \quad (2.34)$$

If $m^2 < 0$ and $\lambda > 0$, then the potential will look like that shown in Fig. 2.6 and have a circle of minima in the complex plane at:

$$\phi = \frac{v}{\sqrt{2}}e^{i\theta}, \quad \theta \in [0, 2\pi] \quad (2.35)$$

where we have used the conventions, $\mu^2 = -m^2$ and $v = \mu/\sqrt{\lambda}$, where v is real and referred to as the *vacuum expectation value* (vev). We can choose a particular minimum, $\phi = \phi_0 = v/\sqrt{2}$, and expanding around this vacuum state we get:

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \rho(x)), \quad \rho \in \mathbb{C} \quad (2.36)$$

$$= \frac{1}{\sqrt{2}}(v + \varphi(x) + i\chi(x)), \quad \varphi, \chi \in \mathbb{R} \quad (2.37)$$

and the potential becomes:

$$V(\phi^*, \phi) = -\frac{\mu^4}{2\lambda} + \mu^2\varphi^2 + \frac{1}{2}\lambda v\varphi(\varphi^2 + \chi^2) + \frac{1}{8}\lambda(\varphi^2 + \chi^2)^2 \quad (2.38)$$

where the second term is a mass term for φ and there are no mass terms for χ . By expanding around the vacuum state, we have revealed the physical particle spectrum for this Lagrangian, which is two scalar particles, φ and χ , where $m_\varphi = \sqrt{2}\mu$ and χ is massless. The massless scalar particle is known as a Goldstone boson.

This Lagrangian is invariant under a global U(1) transformation: $\phi \rightarrow e^{i\theta}\phi$, but since $e^{i\theta}\phi_0 \neq \phi_0$, this does not correspond to the analogous transformation: $\rho \rightarrow e^{i\theta}\rho$. We therefore say that the global U(1) symmetry is *spontaneously broken*. The original symmetry is still present, but not apparent now that the Lagrangian is written in terms of the physical particle fields.

We can now generalize to theories that are invariant under any symmetry group, G , whereby the field transforms as $\phi \rightarrow M\phi$, where M is an element of the group. The symmetry is spontaneously broken if $M\phi_0 \neq \phi_0$ for any M , or unbroken if $M\phi_0 = \phi_0$ for all M . Considering an infinitesimal transformation of ϕ_0 :

$$\phi_0 \rightarrow \phi_0 + i\theta^a T^a \phi_0 \quad (2.39)$$

where T^a are the generators of the group, and θ^a are infinitesimally-small real constants, we identify broken generators as ones where $T\phi_0 \neq \phi_0$ and unbroken generators as ones where $T\phi_0 = \phi_0$. The set of unbroken generators corresponds to a *residual symmetry group*, H , which is a subgroup of the original group.

When $T\phi \neq 0$ for all T , there may still exist linear combinations of the generators, $\hat{T} = c^a T^a$ where $\hat{T}\phi_0 = 0$. To determine these linear combinations, we first define the symmetry breaking matrix:

$$S^{ab} = \phi_0^\dagger \{T^a, T^b\} \phi_0. \quad (2.40)$$

For U(N) and SU(N) symmetry groups where T^a are Hermitian, S^{ab} is real and symmetric and therefore has D real eigenvectors, c_i , with eigenvalues λ_i . It can be shown that a

new generator defined as $\hat{T}^i = c_i^a T^a$, is an unbroken generator when $\lambda_i = 0$, and a broken generator when $\lambda_i \neq 0$. It can be further shown that every broken generator gives rise to a massless Goldstone particle. This is known as Goldstone's theorem [18] and holds for theories with global symmetries. To determine the particle spectra of a gauge theory, we turn to the Higgs mechanism.

The Higgs Mechanism

Consider the Lagrangian of Eq. 2.16, which has a local U(1) symmetry, with the same potential as Eq. 2.34. Now that the vector field, A^μ , is introduced, the vacuum state is specified by values of A^μ and ϕ simultaneously. In the temporal gauge ($A^0 = 0$), this is given by Eq. 2.35 and $A_\mu = 0$.

Once again expanding the scalar field as $\phi = (v + \varphi + i\chi)/\sqrt{2}$, we regain the same potential as Eq. 2.38 and new terms from the expansion of $F_{\mu\nu}F^{\mu\nu}$ and $(D_\mu\phi)^*D^\mu\phi$. Writing only the quadratic terms and transforming to the unitary gauge ($\theta(x) = -\chi/v$) we find:

$$\mathcal{L} = F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2v^2A_\mu A^\mu + \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \mu^2\varphi^2 + \dots \quad (2.41)$$

where we can identify two distinct fields, A_μ and ϕ , which have masses $\sqrt{2}\mu$ and ev respectively. As in the global case, the U(1) symmetry is broken, but unlike the global case, the symmetry breaking has not led to any massless particles. Instead, the originally-massless vector boson has been given a mass.

When generalizing to theories that are invariant under any symmetry group, a similar conclusion is found. If there are D generators of the group, where D' are broken, then there are D' massive vector particles and $D - D'$ massless vector bosons. In the SM, this is the mechanism by which the gauge bosons of the weak force, W^\pm and Z , acquire mass.

2.1.4 The Standard Model Lagrangian

Quantum Electrodynamics

The QFT of electromagnetism is known as quantum electrodynamics (QED). It describes the interaction between charged fermions and photons which are represented by a spinor, ψ , and a vector field, A^μ , respectively. The QED Lagrangian is:

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi, \quad (2.42)$$

and is invariant under global and local U(1) transformations which lead to the conservation of particle number and electric charge respectively. It also is symmetrical under parity, charge conjugation, and time reversal transformations. However, experiments have shown

that parity symmetry is only approximate, and can be violated [19]. Therefore, the QED Lagrangian does not sufficiently describe the behaviour of charged fermions in nature. In the SM, parity violation is introduced via the weak force which couples only to left-handed fermions.

We can start to consider how we can introduce parity violation by writing the fermionic components of the QED Lagrangian in terms of the left and right-handed components of the fermion:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}_L i \not{\partial} \psi_L + \bar{\psi}_R i \not{\partial} \psi_R - m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L). \quad (2.43)$$

If the fermion is massless, then the left-handed and right-handed components are decoupled from each other and we can write a Lagrangian containing only a left-handed spinor:

$$\mathcal{L} = \bar{\psi}_L i \not{\partial} \psi_L + \dots \quad (2.44)$$

which would be parity-violating. Therefore, parity violation can be introduced by treating left-handed and right-handed fermions differently in the SM, and this is done in electroweak theory which supersedes QED as a description of electromagnetism and describes the weak interaction at the same time.

Quantum Chromodynamics

The QFT of the strong interaction is known as quantum chromodynamics (QCD). The Lagrangian is required to be invariant under local SU(3) transformations which implies the existence of eight gauge bosons called gluons. The Lagrangian is given by:

$$\mathcal{L}_{\text{QCD}} = \sum_f i \bar{\psi}_i^f (\not{D}_{ij} - m^f \delta_{ij}) \psi_j^f - \frac{1}{4} G^{c\mu\nu} G_{\mu\nu}^c \quad (2.45)$$

and the covariant derivative is written as:

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} + i g_s G_\mu^c \frac{\lambda_{ij}^c}{2} \quad (2.46)$$

where the strength of the interaction is characterized by g_s , the gluons are represented by G_μ^a , and λ^c are the Gell-Mann matrices which are a set of 3×3 matrices that generate SU(3) [12]. The quark spinors are given by ψ_i^f where $f \in \{u, d, c, s, t, b\}$ corresponds to a quark *flavour* and $i \in \{r, g, b\}$ corresponds to a quark *colour*.

Since the SU(3) symmetry is non-abelian, the gluons are self-interacting and this has profound consequences for the strong interaction. Firstly, g_s increases as the energy scale of an interaction decreases (see Fig. 2.7) which is opposite to the behaviour seen in the electromagnetic interaction where there are no self-interactions. This means that

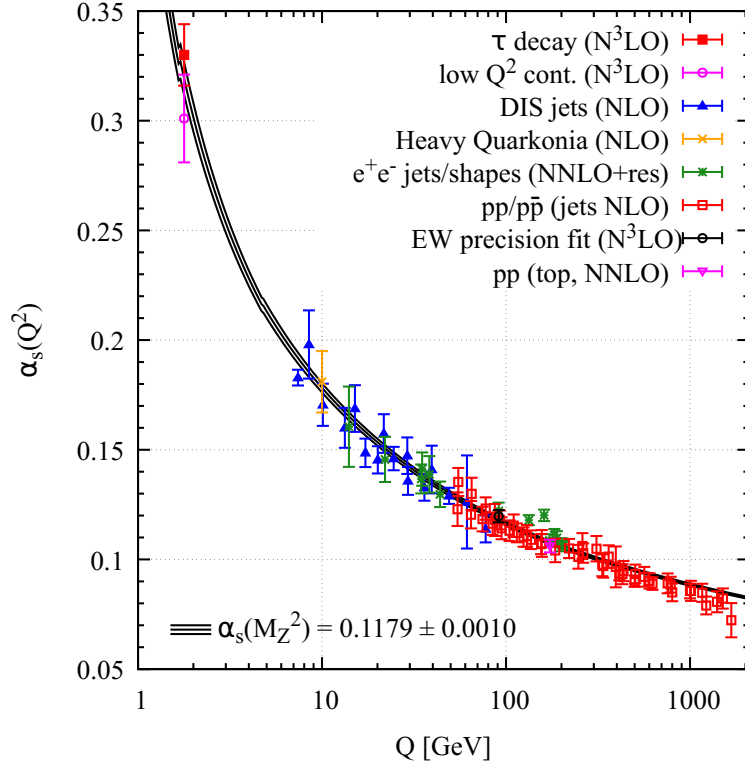


Figure 2.7: Summary of measurements of α_s ($\alpha_s = g_s^2/4\pi$) as a function of the energy scale, Q , taken from Ref. [15]. Each set of coloured points corresponds to a different set of measurements and the degree of QCD perturbation theory used to extract α_s is indicated by the brackets in the legend. Further details about the measurements and α_s calculations can be found in Ref. [15].

at energy scales below $O(1 \text{ GeV})$ where $g_s \sim 1$, calculations using perturbation theory are no longer possible, making predictions from QCD more challenging. Fortunately, for energies $O(> 100 \text{ GeV})$ such as those probed by the LHC, $g_s \sim 0.1$ and perturbation theory is applicable.

A further consequence of gluon self-interactions is the concept of colour confinement which means that quarks and gluons cannot be observed as isolated free particles. Instead, they must exist within colourless bound states called hadrons. States with even and odd numbers of quarks are called mesons and baryons respectively. The two-quark and three-quark variants of these states are commonly produced by QCD interactions at the LHC, and states with higher than three quarks such as tetraquarks and pentaquarks have also been observed [20].

Electroweak Gauge Theory

In the SM, the electroweak bosons arise from an $SU(2) \times U(1)$ gauge symmetry of \mathcal{L}_{SM} . The representation of some field, χ , under $SU(2)$ is given by generators $\{T^a\}$, $a \in \{1, 2, 3\}$, and the generator of $U(1)$ can be any real number, Y , such that a gauge transformation is given by:

$$\chi \rightarrow e^{i\theta^a T^a + i\eta Y} \chi \quad (2.47)$$

where θ^a and η are real numbers. The generators T^a and Y are called *weak isospin* and *weak hypercharge* respectively. The covariant derivative is given by:

$$D_\mu = \partial_\mu + ig_2 A_\mu^a T^a + ig_1 Y B_\mu \quad (2.48)$$

where A_μ^a and B_μ are the $SU(2)$ and $U(1)$ gauge fields respectively.

Experiments have measured the mass of the W and Z bosons to be non-zero [15] and in the SM, these masses are introduced via the Higgs mechanism. This necessitates electroweak theory to include a complex scalar field, ϕ , which is the *Higgs field*, and is in the fundamental representation of $SU(2)$ and $Y = 1/2$. By defining a fourth generator:

$$t^4 = \frac{g_1}{2g_2} \mathbb{1}, \quad (2.49)$$

and $A_\mu^4 = B_\mu$, the gauge fields can be represented by a single vector field in the covariant derivative of the Higgs field:

$$D_\mu \phi = \partial_\mu \phi + ig_2 A_\mu^{a'} t^{a'} \phi, \quad (2.50)$$

which is useful for applying the ideas of the symmetry breaking matrix introduced in Eq. 2.40. The Lagrangian of the theory is:

$$\mathcal{L} = -\frac{1}{2} \text{Tr} W_{\mu\nu} W^{\mu\nu} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (D_\mu \phi)^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \quad (2.51)$$

where $W_{\mu\nu}$ and $F_{\mu\nu}$ are the field strength tensors of the $SU(2)$ and $U(1)$ gauge fields respectively. After electroweak symmetry breaking, the residual symmetry group is $U(1)$ and in terms of the original gauge fields, the physical gauge boson spectrum is:

$$A_\mu = \cos \theta_W B_\mu + \sin \theta_W A_\mu^3, \quad \text{with mass } m_\gamma = 0 \quad (2.52)$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (A_\mu^1 \mp i A_\mu^2), \quad \text{with mass } m_W = \frac{g_2 v}{2} \quad (2.53)$$

$$Z_\mu = \cos \theta_W A_\mu^3 - \sin \theta_W B_\mu, \quad \text{with mass } m_Z = \frac{v}{2} \sqrt{g_1^2 + g_2^2} \quad (2.54)$$

which are the photon, W^\pm bosons, and Z boson respectively where θ_W is the Weinberg angle, and is related to g_1 and g_2 by:

$$\sin \theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \quad \text{and} \quad \cos \theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}}. \quad (2.55)$$

In addition, the Higgs field gives rise to a physical scalar boson, the Higgs boson, with mass $m_H = \sqrt{2}\mu$.

In the physical mass basis, the covariant derivative is written as:

$$D_\mu = \partial_\mu + i \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}} (T^3 + Y) \mathcal{A}_\mu + i \frac{g_2^2 T^3 - g_1^2 Y}{\sqrt{g_1^2 + g_2^2}} Z_\mu + \frac{i g_2}{\sqrt{2}} (W_\mu^+ T^+ + W_\mu^- T^-) \quad (2.56)$$

where $T^\pm = T^1 \pm iT^2$. From the second term, we see that the electric charge, e , which characterizes the strength of interactions with photons, is given by:

$$e = \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}} = g_2 \sin \theta_W. \quad (2.57)$$

The Weinberg angle can be determined by the ratio of the W and Z masses:

$$\cos \theta_W = \frac{m_W}{m_Z} \approx \frac{80.377 \text{ GeV}}{91.188 \text{ GeV}} = 0.88145 \pm 0.00013. \quad (2.58)$$

The parameter, e , can be determined from the fine structure constant, α :

$$e = \sqrt{4\pi\alpha} \approx \sqrt{\frac{4\pi}{137}} \approx 0.30 \quad (2.59)$$

and the $SU(2)$ and $U(1)$ gauge couplings are:

$$g_2 = \frac{e}{\sin \theta_W} \approx 0.64, \quad g_1 = \frac{e}{\cos \theta_W} \approx 0.34 \quad (2.60)$$

The Higgs vev, v , can be determined by the W boson mass:

$$v = \frac{2m_W}{g_2} \approx 246 \text{ GeV} \quad (2.61)$$

and given a measurement of the Higgs boson mass, $m_H = 125.38 \pm 0.14$ [16], the Higgs self-coupling is determined:

$$\lambda = \frac{m_H^2}{2v^2} \approx 0.13. \quad (2.62)$$

Electroweak Interactions with Fermions

The interactions of the fermions with the electroweak force are determined by the representation that the fermions take under the SU(2) and U(1) gauge transformations. Initially, we will consider only the first generation of fermions: the electron (e) and electron neutrino (ν), and the up (u) and down (d) quarks. Since the weak interaction can change leptons into neutrinos and up into down quarks, left-handed fermions are written in doublets which are in the fundamental representation of SU(2):

$$l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \quad q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}. \quad (2.63)$$

The generators of U(1), Y , take the value of $-1/2$ and $1/6$ for the left-handed leptons and quarks respectively, where these values are determined by requiring that the fermions have the correct electric charges. Given that the right-handed fermions do not interact with the weak force, they are said to be neutral under SU(2) meaning they do not transform under SU(2) and the corresponding generator is zero. Additionally, the right-handed neutrino is neutral under U(1) since it has no electric charge. Therefore, it does not interact with any SM forces, and consequently, it is left out of the theory. The rest of the right-handed fermions are written as singlets, denoted by e_R , u_R and d_R and have $Y = -1, 2/3$, and $-1/3$ respectively.

Terms involving fermions and the covariant derivative, referred to as kinetic terms, encode electroweak interactions in the Lagrangian and are:

$$\mathcal{L}_{\text{EW}} = \bar{l}_L i \not{D} l_L + \bar{e}_R i \not{D} e_R + \bar{q}_L i \not{D} q_L + \bar{u}_R i \not{D} u_R + \bar{d}_R i \not{D} d_R + \dots \quad (2.64)$$

where D is the covariant derivative defined in Eq. 2.48. We must also add terms that encode the mass of the fermions. Since the left-handed and right-handed fermions transform differently under SU(2) and U(1), mass terms like:

$$-m(\bar{e}_L e_R + \bar{e}_R e_L) \quad (2.65)$$

are not gauge invariant and cannot be included in \mathcal{L}_{SM} . Instead, the fermion masses are introduced via the Higgs field and spontaneous symmetry breaking similarly to how the gauge bosons acquired their masses. This is achieved with *Yukawa* terms:

$$\mathcal{L}_{\text{EW}} = -y_e(\bar{l}_L \phi l_R + \bar{l}_R \phi^\dagger l_L) - y_u(\bar{q}_L \tilde{\phi} u_R + \bar{u}_R \tilde{\phi}^\dagger q_L) - y_d(\bar{q}_L \phi d_R + \bar{q}_R \phi^\dagger d_L) + \dots \quad (2.66)$$

where y are real constants called *Yukawa* couplings and $\tilde{\phi} \equiv i\sigma_2 \phi^*$. After spontaneous

symmetry breaking, these terms become:

$$\mathcal{L}_{\text{EW}} = -\frac{y_e v}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{y_u v}{\sqrt{2}}(\bar{u}_L u_R + \bar{u}_R u_L) - \frac{y_d v}{\sqrt{2}}(\bar{d}_L d_R + \bar{d}_R d_L) + \dots \quad (2.67)$$

which corresponds to masses of $y_e v/\sqrt{2}$, $y_u v/\sqrt{2}$ and $y_d v/\sqrt{2}$ for the electron and up and down quarks respectively. No mass term is given to the neutrino and it is therefore massless in the SM. Given the observation of neutrino oscillations [9], we know that neutrinos are, in fact, not massless in nature, however, for the results discussed in the rest of this thesis, the neutrino masses are small enough that $m_\nu = 0$ is a safe assumption. Therefore, neutrino masses will not be discussed further.

Generalizing to three generations, we define three-component column vectors:

$$L_L = \begin{pmatrix} l_L^1 \\ l_L^2 \\ l_L^3 \end{pmatrix}, \quad L_R = \begin{pmatrix} l_R^1 \\ l_R^2 \\ l_R^3 \end{pmatrix}, \quad Q_L = \begin{pmatrix} q_L^1 \\ q_L^2 \\ q_L^3 \end{pmatrix}, \quad D_R = \begin{pmatrix} d_R^1 \\ d_R^2 \\ d_R^3 \end{pmatrix}, \quad U_R = \begin{pmatrix} u_R^1 \\ u_R^2 \\ u_R^3 \end{pmatrix} \quad (2.68)$$

and then write the kinetic and Yukawa terms as:

$$\begin{aligned} \mathcal{L}_{\text{EW}} = & \bar{L}_L i \not{D} L_L + \bar{L}_R i \not{D} L_R + \bar{Q}_L i \not{D} Q_L + \bar{D}_R i \not{D} D_R + \bar{U}_R i \not{D} U_R \\ & - (\bar{L}_L Y_l \phi L_R + \bar{Q}_L Y_d \phi D_R + \bar{Q}_L Y_u \tilde{\phi} U_R + \text{h.c.}) + \dots \end{aligned} \quad (2.69)$$

where $Y_{l,d,e}$ are complex 3×3 Yukawa matrices and h.c. corresponds to the Hermitian conjugate. By redefining the fermion fields such that the Yukawa matrices are diagonal, we can find the physical mass basis for the fermions. The fields would be redefined according to:

$$\begin{aligned} L_L &\rightarrow V_{lL} L_L, \quad L_R \rightarrow V_{lR} L_R, \quad U_R \rightarrow V_{uR} U_R, \quad D_R \rightarrow V_{dR} D_R, \\ Q_L &= \begin{pmatrix} U_L \\ D_L \end{pmatrix} \rightarrow \begin{pmatrix} V_{uL} U_L \\ V_{dL} D_L \end{pmatrix} \end{aligned} \quad (2.70)$$

where the V matrices are determined by singular value decomposition of the Yukawa matrices. For the leptons, this leaves the kinetic terms:

$$\mathcal{L}_{\text{EW}} = \sum_f \bar{l}_L^f i \not{D} l_L^f + \bar{l}_R^f i \not{D} l_R^f + \dots \quad (2.71)$$

where $f \in \{e, \mu, \tau\}$, unchanged. Since the terms are the same across flavour, the leptons interact identically in the SM, and this is known as *lepton universality*. Furthermore, the

terms are invariant under transformations of the type:

$$l_L^f \rightarrow e^{i\theta^f} l_L^f, \quad l_R^f \rightarrow e^{i\theta^f} l_R^f, \quad \theta^f \in \mathbb{R} \quad (2.72)$$

meaning that there is a global $U(1)_e \times U(1)_\mu \times U(1)_\tau$ symmetry. This symmetry corresponds to a conserved lepton number for each generation.

After the field redefinitions of Eq. 2.70, the kinetic terms for the right-handed quarks remain the same in an analogous way to the leptons. However, the kinetic terms for the left-handed quarks do not if $V_{uL} \neq V_{dL}$. When writing the covariant derivative in the physical mass basis, we find that the photon and Z boson terms are the same as before. Therefore, interactions with photons and Z bosons are identical across quark generation. On the other hand, the W boson terms are changed. In terms of the redefined quark fields, the kinetic W boson term is:

$$\mathcal{L}_{EW} = -\frac{g_2}{\sqrt{2}}(V_{CKM}^{fg} \bar{u}_L^f \gamma^\mu W_\mu^+ d_L^g + \text{h.c.}) + \dots \quad (2.73)$$

where $V_{CKM} \equiv V_{uL}^\dagger V_{dL}$ is the Cabibbo-Kobayashi-Maskawa (CKM) matrix. If V_{CKM} is not diagonal, it allows the interaction between left-handed quarks and the W boson to change quark flavour. There are four free parameters of the CKM matrix which are three mixing angles: θ_{12} , θ_{13} , θ_{23} , and a complex phase: δ . These parameters have been measured and the corresponding CKM matrix is given by:

$$V_{CKM} \approx \begin{pmatrix} 0.974 & 0.227 & 0.004 \\ 0.226 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix} \quad (2.74)$$

where the magnitudes of each element, $|V_{CKM}^{fg}|$, are shown. This indicates that there are small levels of mixing between neighbouring generations, and even smaller mixing between the first and third generations.

The full electroweak Lagrangian is given by the combination of the gauge terms in Eq. 2.51 with the fermion terms in Eq. 2.69:

$$\begin{aligned} \mathcal{L}_{EW} = & -\frac{1}{2} \text{Tr} W_{\mu\nu} W^{\mu\nu} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\ & + (D_\mu \phi)^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2 \\ & + \bar{L}_L i \not{D} L_L + \bar{L}_R i \not{D} L_R + \bar{Q}_L i \not{D} Q_L + \bar{D}_R i \not{D} D_R + \bar{U}_R i \not{D} U_R \\ & - (\bar{L}_L Y_l \phi L_R + \bar{Q}_L Y_d \phi D_R + \bar{Q}_L Y_u \tilde{\phi} U_R + \text{h.c.}). \end{aligned} \quad (2.75)$$

Fermion masses	$m_u, m_c, m_t, m_d, m_s, m_b, m_e, m_\mu, m_\tau$
Gauge couplings	g_s, g_1, g_2
Higgs	λ, μ^2
CKM	$\theta_{12}, \theta_{13}, \theta_{23}, \delta$
Strong CP phase	$\theta^{\text{CP}} (\approx 0)$

Table 2.1: The 19 free parameters of the SM.

The Standard Model Lagrangian

To write the full SM Lagrangian, we combine \mathcal{L}_{EW} with \mathcal{L}_{QCD} to get:

$$\begin{aligned}
\mathcal{L}_{\text{SM}} = & -\frac{1}{4}G^{c\mu\nu}G_{\mu\nu}^c - \frac{1}{2}\text{Tr} W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
& + (D_\mu\phi)^\dagger D^\mu\phi + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \\
& + \bar{L}_L i \not{D} L_L + \bar{L}_R i \not{D} L_R + \bar{Q}_L i \not{D} Q_L + \bar{D}_R i \not{D} D_R + \bar{U}_R i \not{D} U_R \\
& - (\bar{L}_L Y_l \phi L_R + \bar{Q}_L Y_d \phi D_R + \bar{Q}_L Y_u \tilde{\phi} U_R + \text{h.c.})
\end{aligned} \tag{2.76}$$

where the covariant derivative is:

$$D_\mu = \partial_\mu + ig_s G_\mu^c \frac{\lambda^c}{2} + ig_2 A_\mu^a T^a + ig_1 Y B_\mu, \tag{2.77}$$

and where the quark mass terms in \mathcal{L}_{QCD} have been dropped since they are now accounted for with the Yukawa terms. In addition to the 18 free parameters mentioned already, there is also the strong CP phase, θ^{CP} , that can lead to CP violation in the strong interaction. Experimentally, θ^{CP} , is known to be very small, $\theta^{\text{CP}} \simeq 0$. More information about this parameter can be found in Ref. [21]. All 19 free parameters of the SM are listed in Table 2.1.

2.2 Higgs Boson Phenomenology

2.2.1 Higgs Boson Production and Decay Modes

In the SM, the Higgs boson directly couples to all particles except the gluon and photon, meaning that the Higgs boson can be produced in many different ways at the LHC, and can decay into a variety of final states. In this section, the main production and decay modes of the Higgs boson at the LHC are catalogued, and a discussion about the types of new physics that can be probed via each of them is begun. This is particularly relevant for the EFT interpretation in Chapter 6 where that discussion will continue, and will also be useful for the di-Higgs search in Chapter 5.

For proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV, the dominant production modes of the Higgs boson (those with the largest cross section) are gluon-gluon fusion (ggH), vector boson fusion (VBF), associated production with a vector boson (VH), and associated production with a top quark-antiquark pair (ttH). Further, subdominant production modes include single-top associated production (tH), gluon-initiated associated production with a Z boson (ggZH), and associated production with a bottom quark-antiquark pair (bbH). Leading order diagrams for these processes are shown in Figs. 2.8 and 2.9 and their cross sections are provided in Table 2.2.

The Higgs boson decay modes can be categorized by whether they lead to a two-body final state, or a four-body final state. The two-body final states include decays to massive fermions, and decays to gluons or photons. Since the Higgs boson does not couple directly to gluons or photons, the decays to these particles proceed via loops of predominantly top quarks, where in $H \rightarrow \gamma\gamma$ decays, loops of W^\pm bosons are also allowed. The LO Feynman diagrams for these decays are shown in Fig. 2.10. The four-body decays arise from the decay of the Higgs boson to WW or ZZ, where the vector bosons further decay into leptons or quarks. The most common decay channels are $H \rightarrow b\bar{b}$, $H \rightarrow W\bar{W}$, $H \rightarrow g\bar{g}$, $H \rightarrow \tau\bar{\tau}$, $H \rightarrow c\bar{c}$, $H \rightarrow Z\bar{Z}$, and $H \rightarrow \gamma\gamma$, and their branching fractions are shown in Table 2.3.

Table 2.2: Standard Model predictions of the cross sections for different Higgs boson production modes in proton-proton collisions at $\sqrt{s} = 13$ TeV for $m_H = 125$ GeV [22].

Production mode	ggH	VBF	WH	ZH	bbH	ttH	ggZH	tHq	tHW
Cross section [pb]	48.6	3.78	1.37	0.761	0.528	0.507	0.123	0.074	0.015

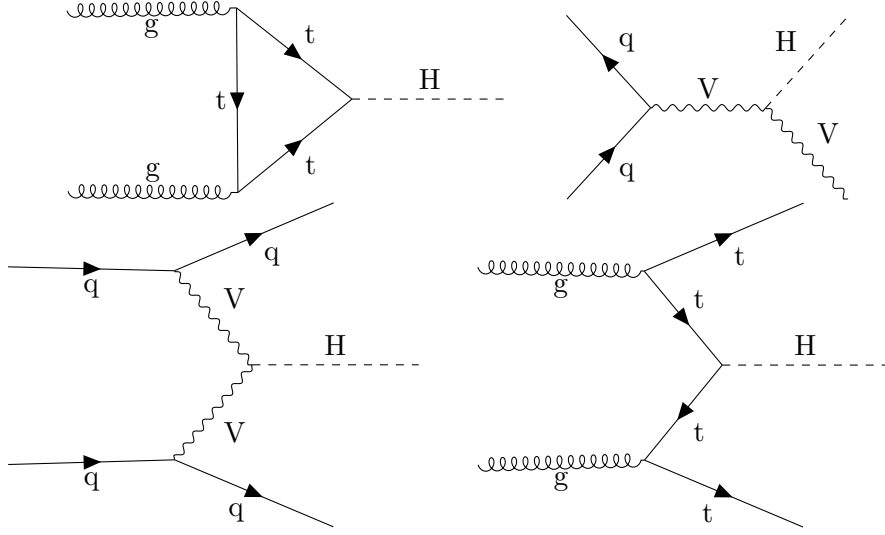


Figure 2.8: LO Feynman diagrams for the dominant production modes of the Higgs boson in proton-proton collisions at $\sqrt{s} = 13$ TeV. From left to right: gluon-gluon fusion (ggH), associated production with a vector boson (VH), vector boson fusion (VBF), and associated production with a top quark-antiquark pair (ttH).

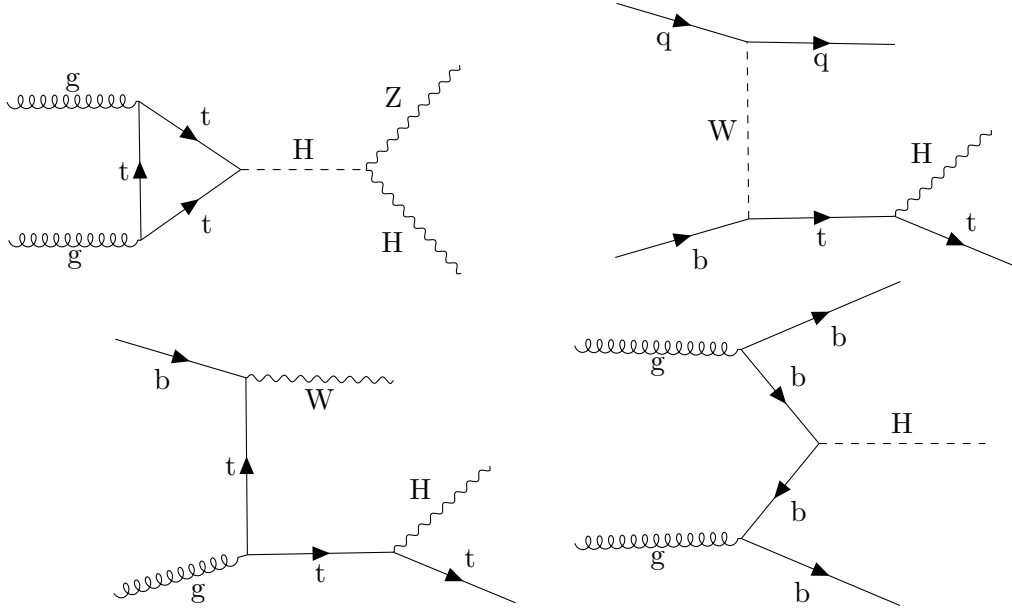
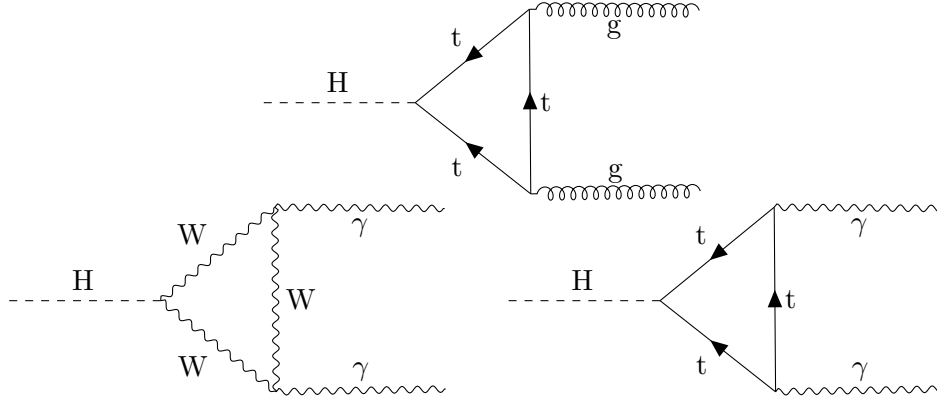


Figure 2.9: LO Feynman diagrams for the subdominant production modes of the Higgs boson in proton-proton collisions at $\sqrt{s} = 13$ TeV. From left to right: gluon-initiated associated production with a Z boson (ggZH), single-top associated production with a quark (tHq), single-top associated production with a W boson (tHW), and associated production with a bottom quark-antiquark pair (bbH).

Table 2.3: Branching fractions for the dominant Higgs boson decay modes at $m_H = 125$ GeV [22].

Decay mode	bb	WW	gg	$\tau\tau$	cc	ZZ	$\gamma\gamma$	Other
Branching fraction [%]	58.2	21.4	8.19	6.27	2.89	2.62	0.227	0.194

**Figure 2.10:** LO Feynman diagrams for the decay of the Higgs boson to two gluons (top) or two photons (bottom).

2.2.2 Simplified Template Cross Sections

Measurements of the Higgs boson take several forms where the simplest measurement is perhaps of the inclusive production cross section of the Higgs boson. This is often presented as a signal strength, μ , which is the ratio of the measured cross section to the SM prediction. A combination of CMS analyses of Higgs boson events from proton-proton collisions at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 138 fb^{-1} , measured this signal strength to be $\mu = 1.014^{+0.055}_{-0.053}$ [2]. This measurement is consistent with the SM ($\mu = 1$) and therefore does not suggest any strong presence of new physics. However, if there was a large deviation, this measurement alone would not provide great insight into the *type* of new physics that may be present. The discrepancy could be explained by a difference in any one of the production or decay modes.

To gain a better understanding of the type of new physics that may be present, the Higgs boson is also studied in a more differential way. This is facilitated by the Simplified Template Cross Section (STXS) framework [22], which provides a common binning scheme to be used across decay channels. The binning scheme is designed to maximize sensitivity to BSM effects whilst remaining as model-independent as possible, allowing interpretations in a variety of new physics models such as an Effective Field Theory (EFT), which is the topic of Chapter 6.

The STXS framework is divided into stages of progressively greater granularity. It begins with stage 0, with bins according to production mode only, where the grouping is

slightly different to the production modes previously mentioned. This binning is shown in Fig. 2.11. There are bins for the ggH, VBF, ttH and bbH modes, a VH hadronic bin which is the VH mode where the V boson decays hadronically, WH and ZH leptonic bins for where V boson decays leptonically, and a tH bin which includes both the tHq and tHW modes. The ZH leptonic bin is further split by whether the event was gg or qq initiated (ggZH or qqZH). Finally, an analysis can choose to merge any of the bins if there is not enough data to measure them separately.

Results under the STXS framework can be presented in a number of ways. When making measurements in a single Higgs boson decay channel, the results are often presented as $\sigma_i \cdot \mathcal{B}^f$, where σ_i is a cross section corresponding to a particular STXS bin, and \mathcal{B}^f is the branching fraction for the given decay channel. In a combination of decay channels, the results may be presented as several sets of $\sigma_i \cdot \mathcal{B}^f$, one for each decay channel, or as $\sigma_i \cdot \mathcal{B}^{ZZ}$ and ratios of branching fractions, $\mathcal{B}^f/\mathcal{B}^{ZZ}$, where the $H \rightarrow ZZ$ decay channel is arbitrarily chosen as the reference channel.

In the CMS combination mentioned previously [2], results are given in terms of $\sigma_i \cdot \mathcal{B}^{ZZ}$ and $\mathcal{B}^f/\mathcal{B}^{ZZ}$ for the stage 0 STXS where the bbH bin is merged into the ggH bin, and the ggZH and qqZH bins are also merged to form a single ZH leptonic bin. These results, shown in Fig. 2.12, allow for deviations in specific production and/or decay modes to be highlighted. In this case, the measured cross sections in the WH and ZH leptonic modes are higher than their predictions, and there is a notable deviation for $\mathcal{B}^{Z\gamma}/\mathcal{B}^{ZZ}$, which together, suggest enhanced Higgs boson couplings to W and Z bosons. There is also an enhancement for tH production, and a discrepancy for $\mathcal{B}^{bb}/\mathcal{B}^{ZZ}$. Overall, there is poor compatibility with the SM, corresponding to a p-value of 0.01.

Additional information can be extracted by splitting the stage 0 bins further by e.g. the Higgs boson p_T , and this is where the later stages of the STXS come in. Schematics for the stage 1.2 binning are shown in Figs. 2.13 to 2.16. The variables used to split the bins, and the number of bins depends on the production mode. The stage 0 ggH bin is split according to the Higgs boson transverse momentum, p_T^H , the number of jets in the event, and in cases of ≥ 2 jets, the invariant mass of the two leading jets, m_{jj} , and the transverse momentum of the Higgs boson and the two jets, p_T^{Hjj} . For $p_T^H > 200$ GeV, there is further splitting according to p_T^{Hj}/p_T^H , where p_T^{Hj} is the transverse momentum of the Higgs boson and the leading jet. Excepting p_T^{Hj} , the stage 0 qqH bin is split according to the same variables. The VH bins are split by the transverse momentum of the vector boson, p_T^V , and the number of jets in the event. The ttH bin is split by p_T^H only. Finally, the bbH and tH bins are not split further in stage 1.2 as there is limited experimental sensitivity to them.

The number of bins defined per production mode depends on the expected sensitivity,

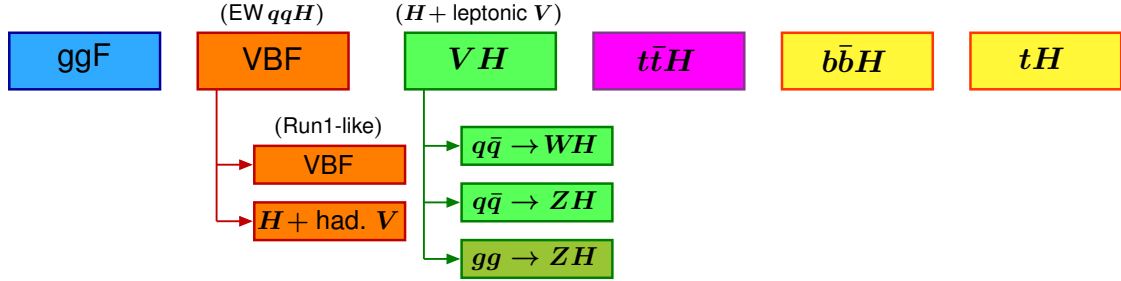


Figure 2.11: Stage 0 binning in the STXS.

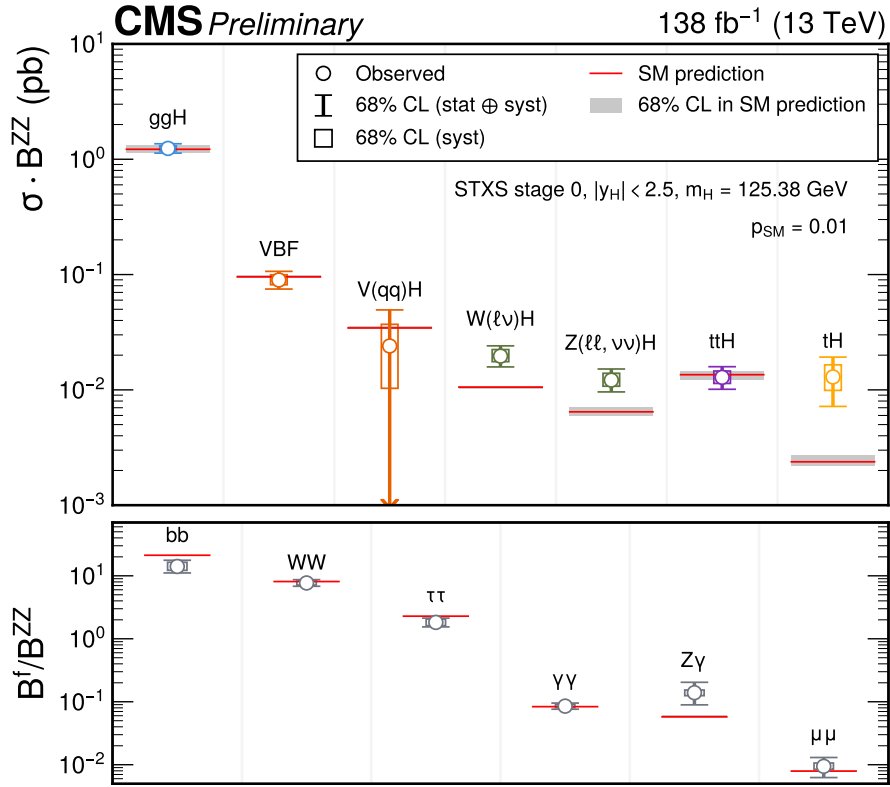
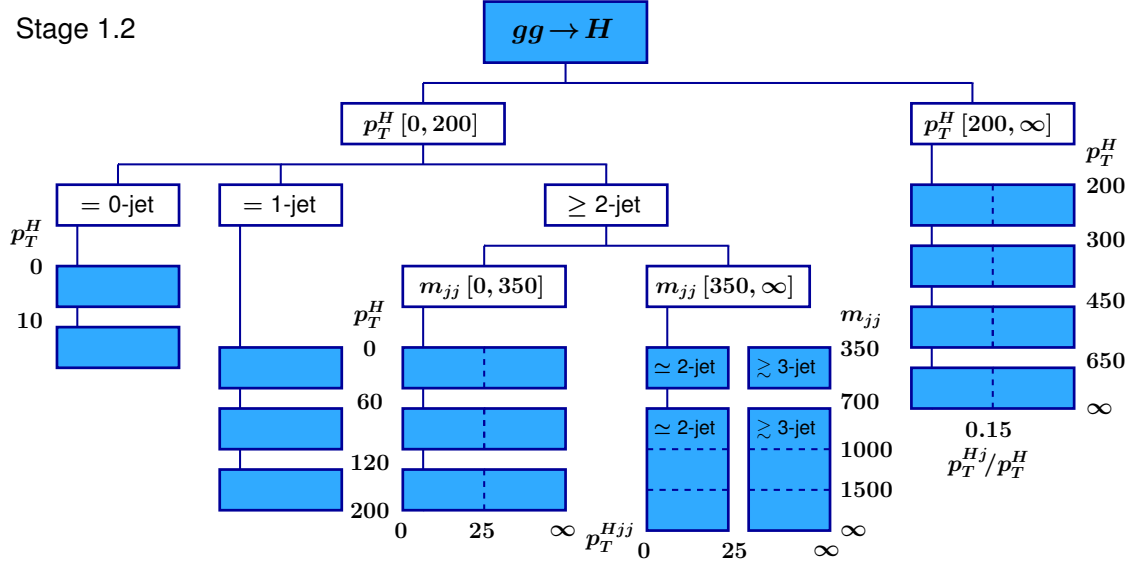
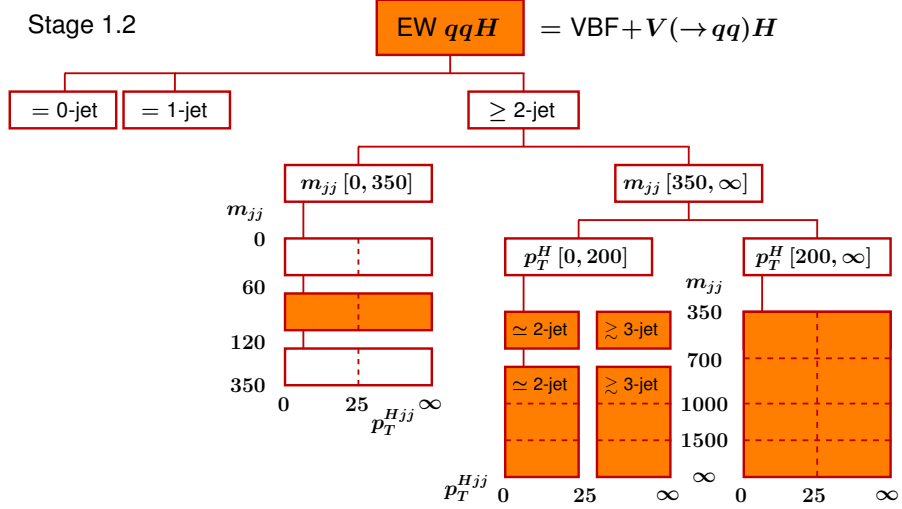


Figure 2.12: STXS stage 0 cross sections and branching fraction ratios measured by the CMS experiment in a combination of Higgs boson decay channels in proton-proton collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 138 fb^{-1} [2]. Theoretical uncertainties which affect the normalizations of the measured parameters are not included in the fit.

Stage 1.2

**Figure 2.13:** Stage 1.2 binning for the ggH production mode in the STXS.**Figure 2.14:** Stage 1.2 binning for the qqH production mode in the STXS.

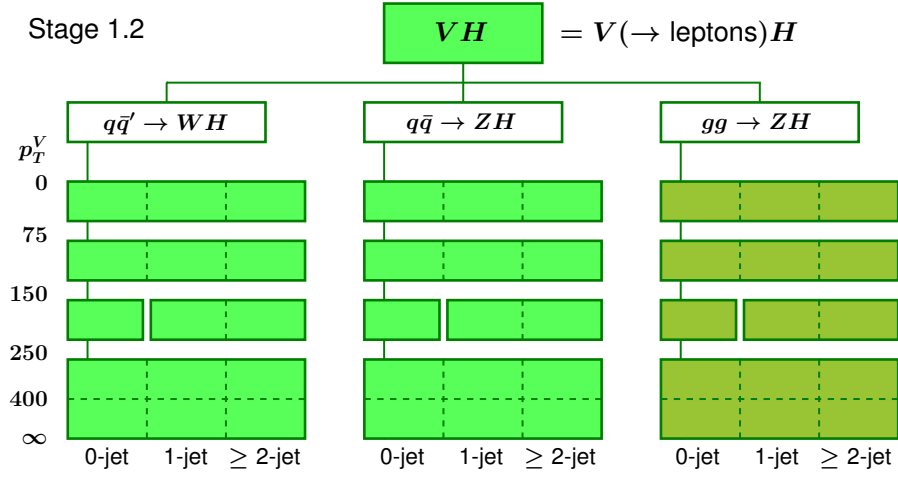


Figure 2.15: Stage 1.2 binning for the VH leptonic production mode in the STXS.

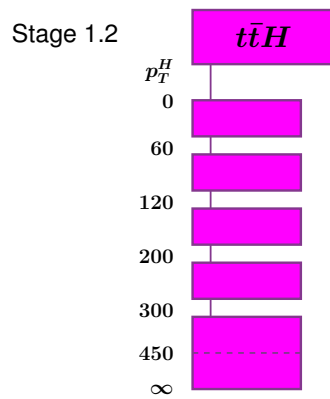


Figure 2.16: Stage 1.2 binning for the $t\bar{t}H$ production mode in the STXS.

where ggH, qqH, VH, and ttH have 27, 24, 45, and 6 bins respectively where in Figs. 2.13 to 2.16, dashed lines indicate suggested places for bins to be merged. In the CMS combination [2], a total of 32 bins were measured simultaneously: 13 for ggH, 5 for qqH, 4 for WH leptonic, 4 for ZH leptonic, 5 for ttH, and a single bin for tH. The results, in terms of $\sigma_i \cdot \mathcal{B}^{ZZ}$ and \mathcal{B}^f/B^{ZZ} , are shown in Fig. 2.17. The compatibility with the SM is slightly better than the stage 0 results, with a p-value of 0.06. Now, in this stage 1.2 splitting, the WH and ZH disagreement can be identified as originating primarily by the $p_T^V > 250$ GeV bins.

Finally, the results are also presented in terms of $\sigma_i \cdot \mathcal{B}^f$ for every input decay channel as shown in Fig. 2.18. Now, the origins of the disagreements in the cross sections can be identified from the individual decay channels and consistency checks can be performed. For example, enhancements in the high p_T^V WH leptonic bins are driven by the $H \rightarrow b\bar{b}$, $H \rightarrow WW$ and $H \rightarrow \tau\tau$ channels, and the excess for tH is driven by the $H \rightarrow \gamma\gamma$ channel which is the only channel with enough sensitivity to measure it separately from ttH.

A particular BSM theory can be confronted with these results by parameterizing the cross sections and branching fractions in terms of the theory's parameters, and then performing a fit to the data to place constraints on the theory parameters. This CMS combination does this with an Effective Field Theory (EFT), specifically the Standard Model EFT (SMEFT). The theoretical details underpinning this interpretation is discussed in Section 2.3.3 and the derivation of the SMEFT parameterization and the final results are discussed in Chapter 6.

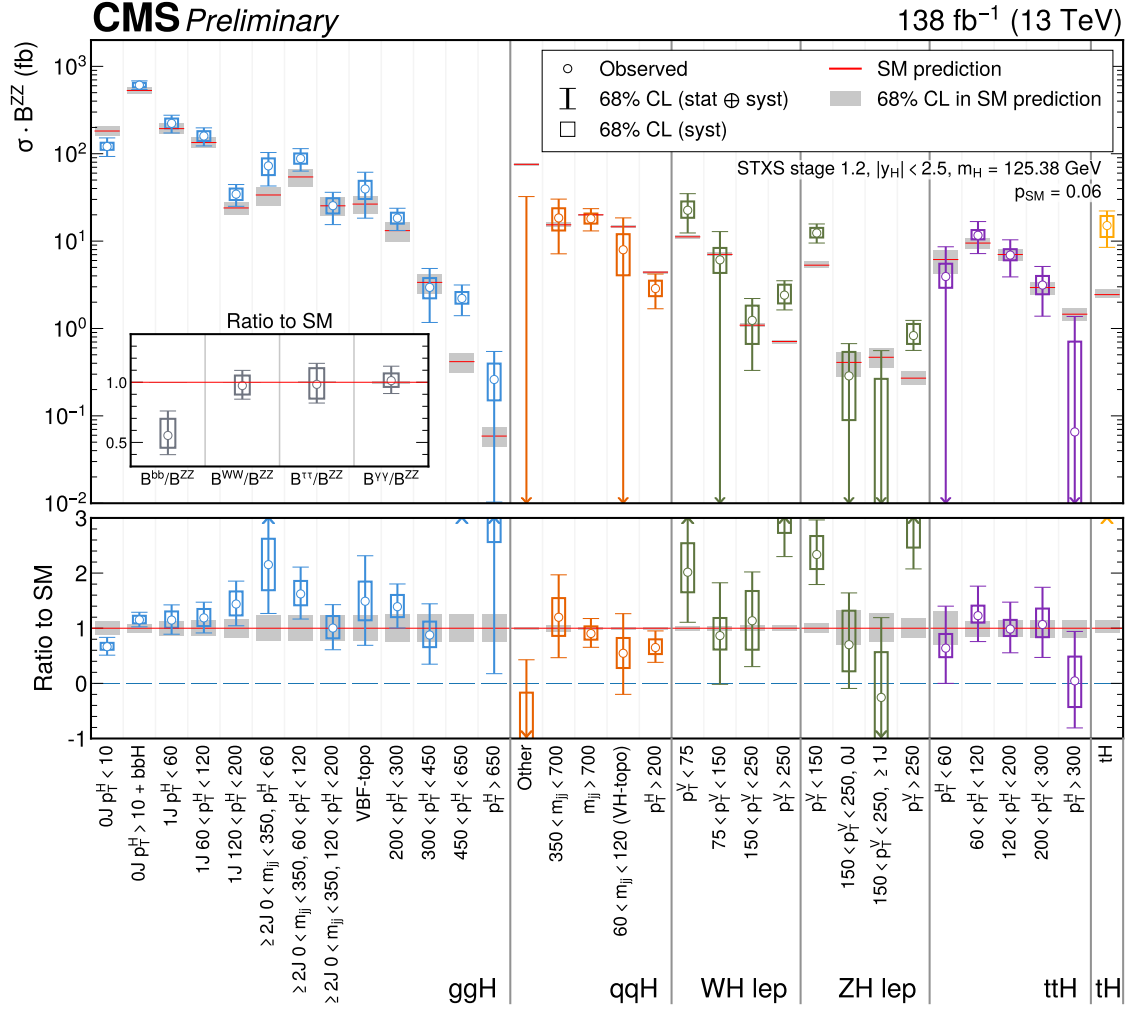


Figure 2.17: STXS stage 1.2 cross sections and branching fraction ratios measured by the CMS experiment in a combination of Higgs boson decay channels in proton-proton collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 138 fb⁻¹ [2]. Theoretical uncertainties which affect the normalizations of the measured parameters are not included in the fit. In cases where the best-fit values and/or 68% CL intervals lie outside the range of the plot, their location (above or below) is indicated by arrows.

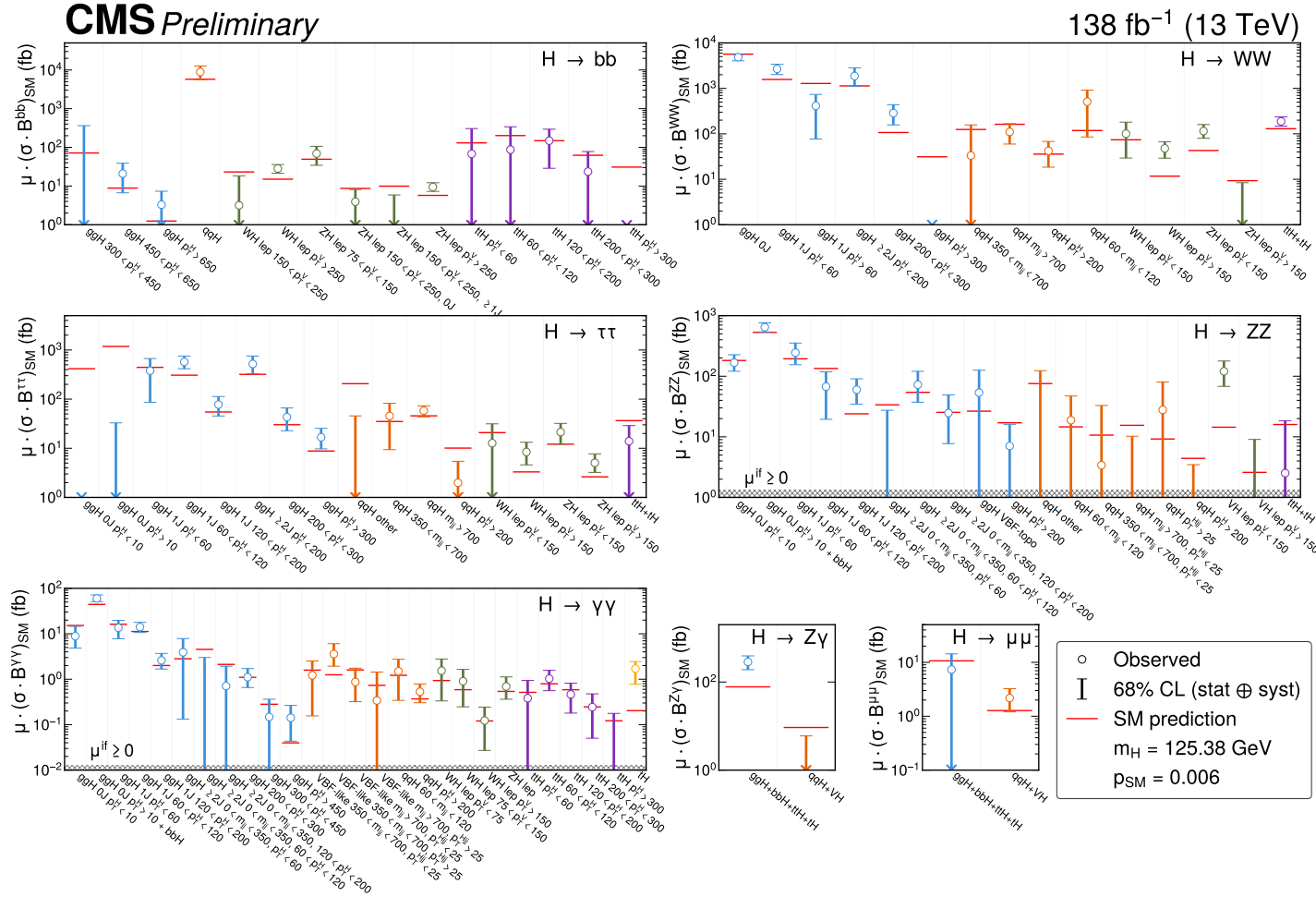


Figure 2.18: STXS stage 1.2 cross section times branching fractions measured by the CMS experiment in a combination of Higgs boson decay channels in proton-proton collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 138 fb⁻¹ [2]. Theoretical uncertainties which affect the normalizations of the measured parameters are not included in the fit. In cases where the best-fit values and/or 68% CL intervals lie outside the range of the plot, their location (above or below) is indicated by arrows. In the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ decay channels, the results are constrained to be positive as indicated by the hatched grey lines.

2.3 Theories Beyond the Standard Model

Theories beyond the Standard Model (BSM) can be used to explain anomalous observations such as dark matter [8], and to solve theoretical problems like the hierarchy problem [12]. In this section, three theories (or families of theories) relevant to the experimental work presented in this thesis are discussed: Warped Extra Dimensions (WED), the Next-to-Minimal-Supersymmetric Standard Model (NMSSM), and Effective Field Theories (EFTs).

2.3.1 Warped Extra Dimensions

In the Warped Extra Dimensions (WED) model [23, 24], a five-dimensional geometry is proposed where a small spatial dimension is added to the traditional 4D spacetime. This theory alleviates the hierarchy problem and introduces two new *gravity particles*: a spin-0 boson called the Radion which we denote $X^{(0)}$, and a spin-2 boson called the Graviton which we denote $X^{(2)}$. Two theoretical scenarios are described in Ref. [24], one where the SM particles are not allowed to propagate along the extra dimension and another where they are allowed, and these scenarios are referred to as the RS1 and Bulk scenarios respectively.

The decay channels and branching fractions for the Radion and Graviton are shown in Fig. 2.19. In the Bulk scenario, the branching fraction to two SM Higgs bosons is about 30% and about 10% for the Radion and Graviton respectively for masses above 300 GeV. There are higher branching fractions to other decay channels such as WW but more sensitive searches can be achieved with HH if the right Higgs boson decay channels are chosen. One of the most competitive HH decay channels is $b\bar{b}\gamma\gamma$ which has lower backgrounds and better mass resolution compared to $X \rightarrow WW$.

Even better constraints can be achieved by performing searches with additional Higgs boson decay channels and then combining these searches, and this motivates the $X \rightarrow HH$ search in the $\gamma\gamma\tau\tau$ final state presented in Chapter 5. In the RS1 scenario, the $X^{(2)} \rightarrow HH$ decay channel has a significantly lower branching fraction of around $\sim 0.5\%$. Therefore, the search in Chapter 5 only considers the Bulk scenario. Production cross sections for the Radion and Graviton at $\sqrt{s} = 13$ TeV in the Bulk scenario are shown in Fig. 2.20 for particular values of kl , Λ_R and \tilde{k} which are free parameters of the theory and described in Ref. [24]. The dominant production mode for both the Radion and Graviton is gluon-gluon fusion.

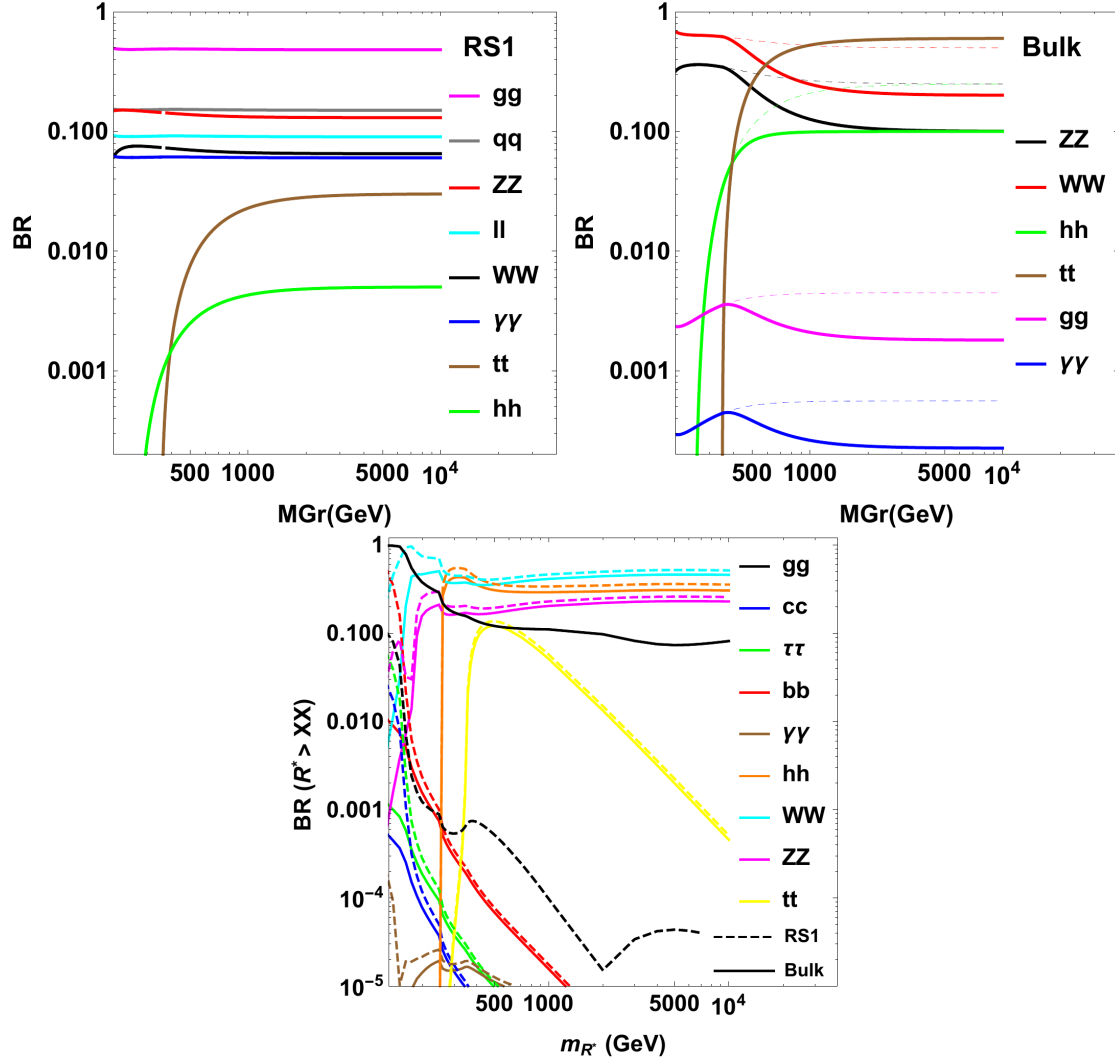


Figure 2.19: Branching fractions for the Graviton (top) and Radion (bottom) as functions of the particle masses. The RS1 and Bulk scenarios for the Graviton are shown in the top-left and top-right respectively whereas for the Radion, they are shown on the same plot and are differentiated by dashed and solid lines respectively. Figures are taken from Ref. [24].

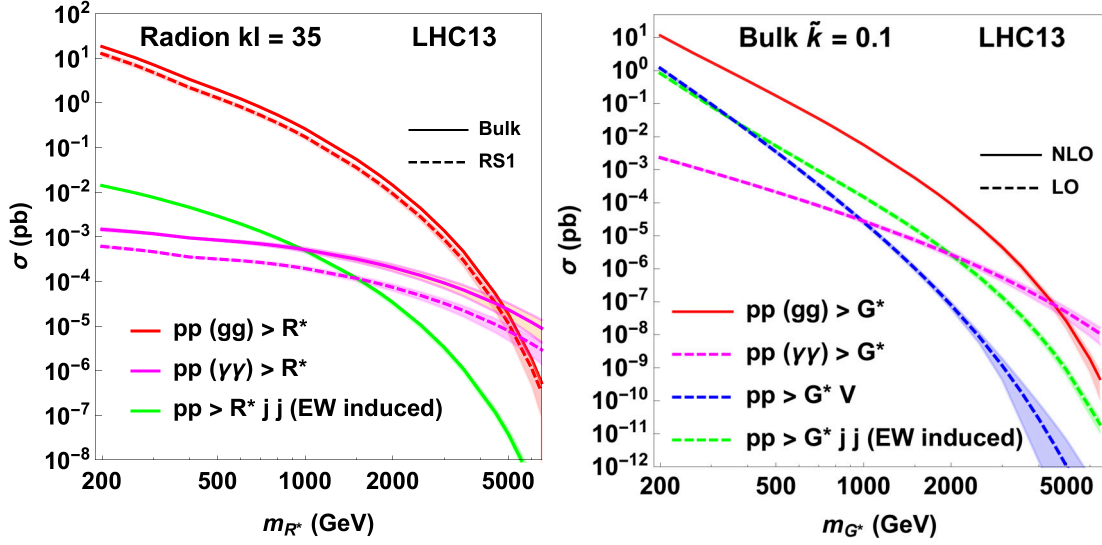


Figure 2.20: Production cross sections at the LHC with $\sqrt{s} = 13$ TeV for the Radion (R^*) on the left, and the Graviton (G^*) on the right, shown as functions of the resonance masses. The Radion cross sections are shown for $kl = 35$ and $\Lambda_R = 3$ TeV and the Graviton cross sections are shown for $\tilde{k} = 0.1$. Figures are taken from Ref. [24].

2.3.2 The Next-to-Minimal-Supersymmetric Standard Model

In supersymmetric theories, every SM particle has a *superpartner* which differs in spin by $1/2$. In other words, every SM boson has a fermionic superpartner, and every SM fermion has a bosonic superpartner. Supersymmetric extensions of the SM are motivated by a solution to the hierarchy problem, an automatic unification of the running gauge couplings at a Grand Unified (GUT) scale M_{GUT} , and the introduction of a stable neutral particle which can be identified as a dark matter candidate [25].

In the Minimal Supersymmetric Standard Model (MSSM), there are two Higgs SU(2)-doublets, Φ_1 and Φ_2 , which lead to three neutral and two charged Higgs bosons [26, 27]. An unattractive property of the MSSM is that the Lagrangian must contain a supersymmetric mass term for the Higgs doublets, which has to be of the order of the SUSY breaking scale, M_{SUSY} , for phenomenological reasons. Ideally, the electroweak scale generated by the Higgs vevs would depend only on M_{SUSY} , which would be the only remaining scale asking for an explanation as to why it is far below M_{GUT} or the Planck scale M_{Planck} . This issue with the MSSM, denoted as the “ μ -problem”, is rectified in the Next-to-Minimal Supersymmetric Standard Model (NMSSM) [25] and the solution requires the introduction of a singlet field, S . Models with this scalar structure are referred to as two-Higgs-doublet + singlet models (2HDM+S).

In 2DMH+S models, there are 3 CP-even and 2 CP-odd neutral scalars, and in the gauge eigenbasis these are denoted H^{SM} , H^{NSM} , H^{S} and A^{NSM} , A^{S} for the CP-even and

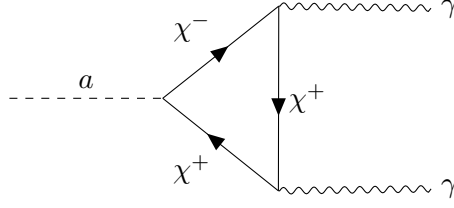


Figure 2.21: Feynman diagram for the decay of the CP-odd Higgs boson, a , to two photons mediated by a chargino (χ^\pm) loop.

CP-odd bosons respectively. The couplings of these states to SM particles are:

$$\begin{aligned}
 H^{\text{SM}}(f_1, f_2, VV) &= (g_{\text{SM}}, g_{\text{SM}}, g_{\text{SM}}), \\
 H^{\text{NSM}}(f_1, f_2, VV) &= (g_{\text{SM}}/\tan\beta, -g_{\text{SM}}\tan\beta, 0), \\
 H^{\text{S}}(f_1, f_2, VV) &= (0, 0, 0), \\
 A^{\text{NSM}}(f_1, f_2, VV) &= (g_{\text{SM}}/\tan\beta, -g_{\text{SM}}\tan\beta, 0), \\
 A^{\text{S}}(f_1, f_2, VV) &= (0, 0, 0)
 \end{aligned}$$

where f_1 (f_2) are SM fermions that couple to Φ_1 (Φ_2), VV corresponds to pairs of W or Z bosons, g_{SM} is the coupling of a SM Higgs boson to such particles, and $\tan\beta$ is the ratio of vacuum expectation values: $\tan\beta = v_1/v_2$ where $v_1 \equiv \langle\Phi_1\rangle$ and $v_2 \equiv \langle\Phi_2\rangle$ [28].

The CP-even and CP-odd mass eigenstates obtained after mixing of gauge eigenstates are denoted $h_i = \{h_{125}, H, h\}$ and $a_i = \{A, a\}$ respectively where h_{125} is identified as the 125 GeV SM-like state observed at the LHC. The mass ordering is such that $m_H > m_h$ and $m_A > m_a$. In the alignment limit where $h_{125} = H^{\text{SM}}$, couplings for interactions involving $h_{125}h_{125}$ and one of the other scalars go to zero. On the other hand, couplings for Hh_{125} and Ah_{125} are non-zero, which allows for so-called *cascade decays*: $H \rightarrow hh_{125}$ and $A \rightarrow ah_{125}$. This motivates the $X \rightarrow YH$ search in Chapter 5 where X and Y are new scalars, and H is the SM Higgs boson.

$Y \rightarrow \gamma\gamma$ in the NMSSM at Low m_Y

If the lightest CP-odd Higgs in the mass basis, a , is very singlet-like, i.e. its main component is A^{S} , then its couplings to SM particles are heavily suppressed, and therefore, the typical hierarchy of Higgs decays: $bb, WW, \tau\tau, ZZ, \dots$ becomes irrelevant. It is still possible for a to decay to SM particles through loop interactions (see Fig. 2.21). The Higgs decay to two photons, which is suppressed in the SM, can now become the dominant decay mode for a , with a branching fraction up to 85% depending on the theoretical scenario [29]. Since a couples weakly to SM particles, its direct production, $pp \rightarrow a$ is also

suppressed. The couplings of a to BSM particles are however, not suppressed. Therefore, a search for $pp \rightarrow A \rightarrow ah_{125}$, where $a \rightarrow \gamma\gamma$, is uniquely placed to study this region of the NMSSM phase space. This motivates the inclusion of $Y \rightarrow \gamma\gamma$ in the $X \rightarrow YH$ search in Chapter 5.

In Ref. [30], maximally-allowed cross sections for this process are calculated given constraints from a set of relevant measurements. These cross sections are shown in Fig. 2.22. If in the search for this process, we can set upper limits for the cross sections below the maximally-allowed, it means that our search will provide tighter constraints on the NMSSM than available at the time that Ref. [30] was published (May 2022).

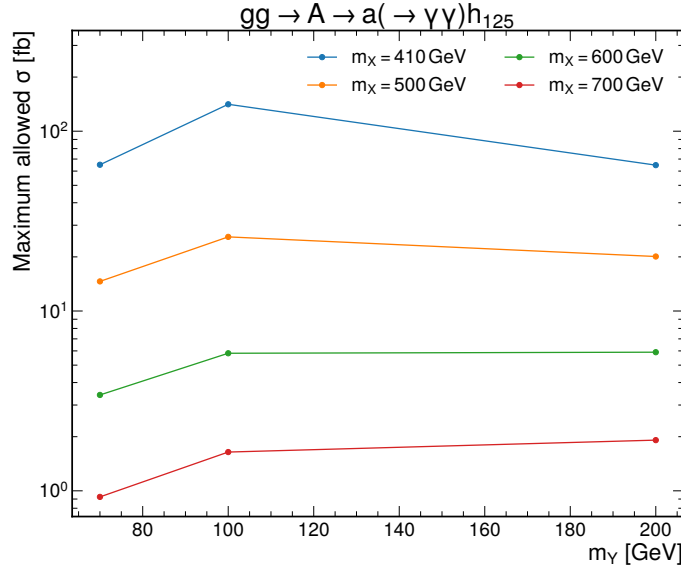


Figure 2.22: Maximally-allowed cross sections for $pp \rightarrow A \rightarrow h_{125}a(\rightarrow \gamma\gamma)$ in the NMSSM provided experimental constraints [30].

2.3.3 Effective Field Theory

Fermi Theory

An effective field theory (EFT) is a low-energy approximation of another theory that is capable of making accurate predictions up to a particular energy scale. For example, the weak interaction can be approximated by an EFT called Fermi theory [31] which contains no description of the W or Z bosons. In the SM, muon decay is mediated by the exchange of a W boson but in Fermi theory, this is calculated using a four-point interaction (see Fig. 2.23).

In the SM diagram, the mediating W boson, which is also referred to as a *propagator*, introduces a term to the matrix element:

$$\mathcal{M} = -i \frac{g_{\mu\nu} - \frac{q_\mu q_\nu}{m_W^2}}{q^2 - m_W^2 + im_W \Gamma_W} \times \dots \quad (2.78)$$

where q is the four-vector of the momentum transferred by the W boson, $\Gamma_W = 2.09 \text{ GeV}$ is the decay width of the W boson, and $g_{\mu\nu}$ is the Minkowski metric. In muon decay, $\sqrt{q^2} \ll m_W$ since $q^2 = m_\mu^2$ and given that $\Gamma_W \ll m_W$ as well, the propagator term is approximately:

$$\frac{ig_{\mu\nu}}{m_W^2} \quad (2.79)$$

which, ignoring the vector indices, is a constant and can be absorbed into the coupling strength parameter of the Fermi theory, i.e. the theories are the same except for a factor of $1/m_W^2$ in the coupling strengths. The full calculation in Fermi theory predicts the decay width to be:

$$\Gamma = \frac{G_F^2}{(\hbar c)^6} \frac{(m_\mu c^2)^5}{192\pi^3} \quad (2.80)$$

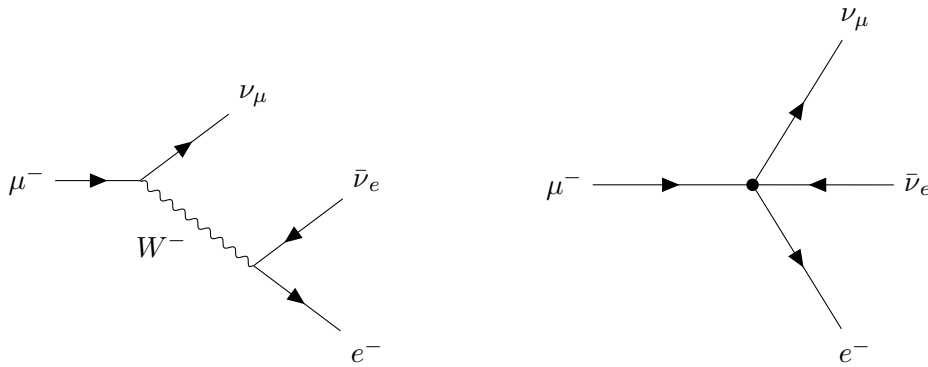


Figure 2.23: LO Feynman diagrams for muon decay in the SM (left) and in Fermi theory (right).

where G_F is the *Fermi constant* which characterizes the strength of the interaction. This can be related to the electroweak SU(2) coupling, g_2 , with:

$$G_F = \frac{1}{4\sqrt{2}} \frac{g_2^2}{m_W^2} \quad (2.81)$$

and therefore, a measurement of the muon decay width provides a relationship between the couplings of the weak interaction and the W boson mass.

More generally, an interaction involving a propagator can be approximated by a point interaction if $q^2 \ll \Lambda^2$ where Λ is the mass of the propagator. Then, measurements of interactions in that energy regime can be used to place constraints on the relationship between g and Λ , where g is a coupling parameter for these interactions. This extends to propagators outside the SM and therefore, EFT can be used to place constraints on new physics when using measurements at an energy scale much less than the mass of the new particle.

Standard Model Effective Field Theory

The Standard Model Effective Field Theory (SMEFT) extends the SM Lagrangian by adding terms of higher dimension:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda} \sum_i C_i^{(5)} O_i^{(5)} + \frac{1}{\Lambda^2} \sum_i C_i^{(6)} O_i^{(6)} + \dots \quad (2.82)$$

where $O_i^{(d)}$ are dimension- d terms called *operators* that contain the SM fields and are invariant under the same gauge group as the SM, and $C_i^{(d)}$ are complex numbers called *Wilson coefficients* that characterize the magnitude of each operator's contribution to $\mathcal{L}_{\text{SMEFT}}$. Since $\mathcal{L}_{\text{SMEFT}}$ is a perturbative expansion, it requires that $C_i/\Lambda^d < \mathcal{O}(1)$ for all C_i .

We neglect all operators that violate lepton-number or baryon-number conservation and any operators that are dimension-7 or higher since they are suppressed by higher orders of $1/\Lambda$. At dimension five, one operator remains which generates a Majorana mass term for neutrinos, and we also neglect this. The rest of this section will focus on the remaining dimension-6 terms which we denote by \mathcal{L}_6 .

We use a non-redundant basis for the operators called the Warsaw basis [32]. The basis definition is written in Table 2.4, using the following notation:

$$\tilde{X}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} X_{\rho\sigma}, \quad H^\dagger i \overleftrightarrow{D} H = H^\dagger (iD_\mu H) - (iD_\mu H^\dagger) H, \quad (2.83)$$

$$\sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu], \quad H^\dagger i \overleftrightarrow{D}^i H = H^\dagger \sigma^i (iD_\mu H) - (iD_\mu H^\dagger) \sigma^i H. \quad (2.84)$$

$\mathcal{L}_6^{(1)} - X^3$		$\mathcal{L}_6^{(6)} - \psi^2 XH$		$\mathcal{L}_6^{(8b)} - (\bar{R}R)(\bar{R}R)$	
Q_G	$f^{abc} G_\mu^{a\nu} G_\nu^{b\rho} G_\rho^{c\mu}$	Q_{eW}	$(\bar{l}_p \sigma^{\mu\nu} e_r) \sigma^i H W_{\mu\nu}^i$	Q_{ee}	$(\bar{e}_p \gamma_\mu e_r) (\bar{e}_s \gamma^\mu e_t)$
$Q_{\tilde{G}}$	$f^{abc} \tilde{G}_\mu^{a\nu} G_\nu^{b\rho} G_\rho^{c\mu}$	Q_{eB}	$(\bar{l}_p \sigma^{\mu\nu} e_r) H B_{\mu\nu}$	Q_{uu}	$(\bar{u}_p \gamma_\mu u_r) (\bar{u}_s \gamma^\mu u_t)$
Q_W	$\epsilon^{ijk} W_\mu^{i\nu} W_\nu^{j\rho} W_\rho^{k\mu}$	Q_{uG}	$(\bar{q}_p \sigma^{\mu\nu} T^a u_r) \tilde{H} G_{\mu\nu}^a$	Q_{dd}	$(\bar{d}_p \gamma_\mu d_r) (\bar{d}_s \gamma^\mu d_t)$
$Q_{\tilde{W}}$	$\epsilon^{ijk} \tilde{W}_\mu^{i\nu} W_\nu^{j\rho} W_\rho^{k\mu}$	Q_{uW}	$(\bar{q}_p \sigma^{\mu\nu} u_r) \sigma^i \tilde{H} W_{\mu\nu}^i$	Q_{eu}	$(\bar{e}_p \gamma_\mu e_r) (\bar{u}_s \gamma^\mu u_t)$
$\mathcal{L}_6^{(2)} - H^6$		Q_{uB}	$(\bar{q}_p \sigma^{\mu\nu} u_r) \tilde{H} B_{\mu\nu}$	Q_{ed}	$(\bar{e}_p \gamma_\mu e_r) (\bar{d}_s \gamma^\mu d_t)$
Q_H	$(H^\dagger H)^3$	Q_{dG}	$(\bar{q}_p \sigma^{\mu\nu} T^a d_r) H G_{\mu\nu}^a$	$Q_{ud}^{(1)}$	$(\bar{u}_p \gamma_\mu u_r) (\bar{d}_s \gamma^\mu d_t)$
$\mathcal{L}_6^{(3)} - H^4 D^2$		Q_{dW}	$(\bar{q}_p \sigma^{\mu\nu} d_r) \sigma^i H W_{\mu\nu}^i$	$Q_{ud}^{(8)}$	$(\bar{u}_p \gamma_\mu T^a u_r) (\bar{d}_s \gamma^\mu T^a d_t)$
$Q_{H\Box}$	$(H^\dagger H) \Box (H^\dagger H)$	Q_{dB}	$(\bar{q}_p \sigma^{\mu\nu} d_r) H B_{\mu\nu}$		
Q_{HD}	$(D^\mu H^\dagger H) (H^\dagger D_\mu H)$				
$\mathcal{L}_6^{(4)} - X^2 H^2$		$\mathcal{L}_6^{(7)} - \psi^2 H^2 D$		$\mathcal{L}_6^{(8c)} - (\bar{L}L)(\bar{R}R)$	
Q_{HG}	$H^\dagger H G_{\mu\nu}^a G^{a\mu\nu}$	$Q_{Hl}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H) (\bar{l}_p \gamma^\mu l_r)$	Q_{le}	$(\bar{l}_p \gamma_\mu l_r) (\bar{e}_s \gamma^\mu e_t)$
$Q_{H\tilde{G}}$	$H^\dagger H \tilde{G}_{\mu\nu}^a G^{a\mu\nu}$	$Q_{Hl}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H) (\bar{l}_p \sigma^i \gamma^\mu l_r)$	Q_{lu}	$(\bar{l}_p \gamma_\mu l_r) (\bar{u}_s \gamma^\mu u_t)$
Q_{HW}	$H^\dagger H W_{\mu\nu}^i W^{i\mu\nu}$	Q_{He}	$(H^\dagger i \overleftrightarrow{D}_\mu H) (\bar{e}_p \gamma^\mu e_r)$	Q_{ld}	$(\bar{l}_p \gamma_\mu l_r) (\bar{d}_s \gamma^\mu d_t)$
$Q_{H\tilde{W}}$	$H^\dagger H \tilde{W}_{\mu\nu}^i W^{i\mu\nu}$	$Q_{Hq}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H) (\bar{q}_p \gamma^\mu q_r)$	Q_{qe}	$(\bar{q}_p \gamma_\mu q_r) (\bar{e}_s \gamma^\mu e_t)$
Q_{HB}	$H^\dagger H B_{\mu\nu} B^{\mu\nu}$	$Q_{Hq}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H) (\bar{q}_p \sigma^i \gamma^\mu q_r)$	$Q_{qu}^{(1)}$	$(\bar{q}_p \gamma_\mu q_r) (\bar{u}_s \gamma^\mu u_t)$
$Q_{H\tilde{B}}$	$H^\dagger H \tilde{B}_{\mu\nu} B^{\mu\nu}$	Q_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H) (\bar{u}_p \gamma^\mu u_r)$	$Q_{qu}^{(8)}$	$(\bar{q}_p \gamma_\mu T^a q_r) (\bar{u}_s \gamma^\mu T^a u_t)$
Q_{HWB}	$H^\dagger \sigma^i H W_{\mu\nu}^i B^{\mu\nu}$	Q_{Hd}	$(H^\dagger i \overleftrightarrow{D}_\mu H) (\bar{d}_p \gamma^\mu d_r)$	$Q_{qd}^{(1)}$	$(\bar{q}_p \gamma_\mu q_r) (\bar{d}_s \gamma^\mu d_t)$
$Q_{H\tilde{W}B}$	$H^\dagger \sigma^i H \tilde{W}_{\mu\nu}^i B^{\mu\nu}$	$Q_{Hud} + \text{h.c.}$	$i(\tilde{H}^\dagger D_\mu H) (\bar{u}_p \gamma^\mu d_r)$	$Q_{qd}^{(8)}$	$(\bar{q}_p \gamma_\mu T^a q_r) (\bar{d}_s \gamma^\mu T^a d_t)$
$\mathcal{L}_6^{(5)} - \psi^2 H^3$		$\mathcal{L}_6^{(8a)} - (\bar{L}L)(\bar{L}L)$		$\mathcal{L}_6^{(8d)} - (\bar{L}R)(\bar{R}L), (\bar{L}R)(\bar{L}R)$	
Q_{eH}	$(H^\dagger H) (\bar{l}_p e_r H)$	Q_{ll}	$(\bar{l}_p \gamma_\mu l_r) (\bar{l}_s \gamma^\mu l_t)$	Q_{ledq}	$(\bar{l}_p^j e_r) (\bar{d}_s q_{tj})$
Q_{uH}	$(H^\dagger H) (\bar{q}_p u_r \tilde{H})$	$Q_{qq}^{(1)}$	$(\bar{q}_p \gamma_\mu q_r) (\bar{q}_s \gamma^\mu q_t)$	$Q_{quqd}^{(1)}$	$(\bar{q}_p^j u_r) \epsilon_{jk} (\bar{q}_s^k d_t)$
Q_{dH}	$(H^\dagger H) (\bar{q}_p d_r H)$	$Q_{qq}^{(3)}$	$(\bar{q}_p \gamma_\mu \sigma^i q_r) (\bar{q}_s \gamma^\mu \sigma^i q_t)$	$Q_{quqd}^{(8)}$	$(\bar{q}_p^j T^a u_r) \epsilon_{jk} (\bar{q}_s^k T^a d_t)$
		$Q_{lq}^{(1)}$	$(\bar{l}_p \gamma_\mu l_r) (\bar{q}_s \gamma^\mu q_t)$	$Q_{lequ}^{(1)}$	$(\bar{l}_p^j e_r) \epsilon_{jk} (\bar{q}_s^k u_t)$
		$Q_{lq}^{(3)}$	$(\bar{l}_p \gamma_\mu \sigma^i l_r) (\bar{q}_s \gamma^\mu \sigma^i q_t)$	$Q_{lequ}^{(3)}$	$(\bar{l}_p^j \sigma_{\mu\nu} e_r) \epsilon_{jk} (\bar{q}_s^k \sigma^{\mu\nu} u_t)$

Table 2.4: \mathcal{L}_6 operators in the Warsaw basis [32], categorized into eight classes $\mathcal{L}_6^{(n)}$ as in [33]. Only baryon number-conserving invariants are retained. The flavor indices p, r, s, t are suppressed in the operators' labels.

In Table 2.4, there are 59 independent operators and naively, one might expect therefore, that there are only 59 Wilson coefficients that need to be measured to specify the theory. However, some operators carry flavour indices of which all combinations need to be summed over:

$$\mathcal{L}_6 = \frac{1}{\Lambda^2} \sum_i \sum_{p,r=1}^3 C_{i,pr} Q_{i,pr} + \dots \quad (2.85)$$

and this leads to a higher number of Wilson coefficients. In total, there are 2599 free parameters in \mathcal{L}_6 , counting real and imaginary components of the Wilson coefficients separately. It is not currently possible to experimentally constrain all of these parameters simultaneously and nor will it be in the short-term future. Therefore, we use flavour assumptions to reduce the number of free parameters to a more reasonable level.

Flavour Assumptions

The most restrictive flavour assumption we can make is the symmetry of the kinetic terms: $U(3)^5 = U(3)_q \times U(3)_u \times U(3)_d \times U(3)_l \times U(3)_e$, where each field is assigned to a three-component representation of the associated group. In this assumption, the terms in Eq. 2.85 become:

$$\mathcal{L}_6 = \frac{1}{\Lambda^2} \sum_i \sum_{p,r=1}^3 C_i X_{i,pr} Q_{i,pr} + \dots \quad (2.86)$$

where the flavour structure of each operator is factored out into $X_{i,pr}$ leaving a single Wilson coefficient per operator. Under the $U(3)^5$ flavour assumption, there are a total of 85 free parameters.

If a set of measurements can distinguish an operator's effect on one flavour of fermion from another, e.g. by combining measurements of top-quark production and light-jet production, then we need not be so restrictive with our flavour assumption. A less restrictive option compared to $U(3)^5$ is the so-called **topU31** assumption [34] where the quarks of the first two generations and quarks of the 3rd are described by independent fields, denoted (q_p, u_p, d_p) and (Q, t, b) respectively. The fermionic operators in this basis are provided in Table 2.5. A $U(2)^3 = U(2)_q \times U(2)_u \times U(2)_d$ symmetry is imposed in the quark sector and paired with a $U(3)^2 = U(3)_l \times U(3)_e$ symmetry in the lepton sector. With this flavour assumption, an operator can contribute differently to processes involving the first two generations from processes involving the third. Therefore, we can probe new physics effects that have a hierarchical structure in the quark sector. In the SMEFT interpretation described in Chapter 6, the **topU31** assumption is used.

$\mathcal{L}_6^{(5)} - \psi^2 H^3$							
Q_{uH}	$(H^\dagger H)(\bar{q} Y_u^\dagger u \tilde{H})$	Q_{dH}	$(H^\dagger H)(\bar{q} Y_d^\dagger d H)$	Q_{eH}	$(H^\dagger H)(\bar{l}_p e_r H)$		
Q_{tH}	$(H^\dagger H)(\bar{Q} \tilde{H} t)$	Q_{bH}	$(H^\dagger H)(\bar{Q} H b)$				
$\mathcal{L}_6^{(6)} - \psi^2 X H$							
Q_{eW}	$(\bar{l}_p \sigma^{\mu\nu} e_r) \sigma^i H W_{\mu\nu}^i$	Q_{uW}	$(\bar{q} Y_u^\dagger \sigma^{\mu\nu} u) \sigma^i \tilde{H} W_{\mu\nu}^i$	Q_{uB}	$(\bar{q} Y_u^\dagger \sigma^{\mu\nu} u) \tilde{H} B_{\mu\nu}$	Q_{uG}	$(\bar{q} Y_u^\dagger \sigma^{\mu\nu} T^a u) \tilde{H} G_{\mu\nu}^a$
Q_{eB}	$(\bar{l}_p \sigma^{\mu\nu} e_r) H B_{\mu\nu}$	Q_{tW}	$(\bar{Q} \sigma^{\mu\nu} t) \sigma^i \tilde{H} W_{\mu\nu}^i$	Q_{tB}	$(\bar{Q} \sigma^{\mu\nu} t) \tilde{H} B_{\mu\nu}$	Q_{tG}	$(\bar{Q} \sigma^{\mu\nu} T^a t) \tilde{H} G_{\mu\nu}^a$
Q_{dW}	$(\bar{q} Y_d^\dagger \sigma^{\mu\nu} d) \sigma^i H W_{\mu\nu}^i$	Q_{dB}	$(\bar{q} Y_d^\dagger \sigma^{\mu\nu} d) H B_{\mu\nu}$	Q_{dG}	$(\bar{q} Y_d^\dagger \sigma^{\mu\nu} T^a d) H G_{\mu\nu}^a$		
Q_{bW}	$(\bar{Q} \sigma^{\mu\nu} b) \sigma^i H W_{\mu\nu}^i$	Q_{bB}	$(\bar{Q} \sigma^{\mu\nu} b) H B_{\mu\nu}$	Q_{bG}	$(\bar{Q} \sigma^{\mu\nu} T^a b) H G_{\mu\nu}^a$		
$\mathcal{L}_6^{(7)} - \psi^2 H^2 D$							
$Q_{Hl}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{l}_p \gamma^\mu l_r)$	$Q_{Hl}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H)(\bar{l}_p \sigma^i \gamma^\mu l_r)$	Q_{He}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{e}_p \gamma^\mu e_r)$		
$Q_{Hq}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{q} \gamma^\mu q)$	$Q_{Hq}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H)(\bar{q} \sigma^i \gamma^\mu q)$	Q_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{u} \gamma^\mu u)$	Q_{Hd}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{d} \gamma^\mu d)$
$Q_{HQ}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{Q} \gamma^\mu Q)$	$Q_{HQ}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H)(\bar{Q} \sigma^i \gamma^\mu Q)$	Q_{Ht}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{t} \gamma^\mu t)$	Q_{Hb}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{b} \gamma^\mu b)$
Q_{Hud}	$i(\tilde{H}^\dagger D_\mu H)(\bar{u} Y_u Y_d^\dagger \gamma^\mu d)$	Q_{Htb}	$i(\tilde{H}^\dagger D_\mu H)(\bar{t} \gamma^\mu b)$				
$\mathcal{L}_6^{(8a)} - (\bar{L} L)(\bar{L} L)$							
$Q_{lq}^{(1)}$	$(\bar{l}_p \gamma_\mu l_r)(\bar{q} \gamma^\mu q)$	$Q_{lq}^{(3)}$	$(\bar{l}_p \sigma^i \gamma_\mu l_r)(\bar{q} \sigma^i \gamma^\mu q)$	Q_{ll}	$(\bar{l}_p \gamma_\mu l_r)(\bar{l}_s \gamma^\mu l_t)$		
$Q_{lQ}^{(1)}$	$(\bar{l}_p \gamma_\mu l_r)(\bar{Q} \gamma^\mu Q)$	$Q_{lQ}^{(3)}$	$(\bar{l}_p \sigma^i \gamma_\mu l_r)(\bar{Q} \sigma^i \gamma^\mu Q)$	$Q_{QQ}^{(1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{Q} \gamma^\mu Q)$	$Q_{QQ}^{(8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{Q} T^a \gamma^\mu Q)$
$Q_{qq}^{(1,1)}$	$(\bar{q} \gamma_\mu q)(\bar{q} \gamma^\mu q)$	$Q_{qq}^{(1,8)}$	$(\bar{q} T^a \gamma_\mu q)(\bar{q} T^a \gamma^\mu q)$	$Q_{qq}^{(3,1)}$	$(\bar{q} \sigma^i \gamma_\mu q)(\bar{q} \sigma^i \gamma^\mu q)$	$Q_{qq}^{(3,8)}$	$(\bar{q} \sigma^i T^a \gamma_\mu q)(\bar{q} \sigma^i T^a \gamma^\mu q)$
$Q_{Qq}^{(1,1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{q} \gamma^\mu q)$	$Q_{Qq}^{(1,8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{q} T^a \gamma^\mu q)$	$Q_{Qq}^{(3,1)}$	$(\bar{Q} \sigma^i \gamma_\mu Q)(\bar{q} \sigma^i \gamma^\mu q)$	$Q_{Qq}^{(3,8)}$	$(\bar{Q} \sigma^i T^a \gamma_\mu Q)(\bar{q} \sigma^i T^a \gamma^\mu q)$
$\mathcal{L}_6^{(8b)} - (\bar{R} R)(\bar{R} R)$							
Q_{eu}	$(\bar{e}_p \gamma_\mu e_r)(\bar{u} \gamma^\mu u)$	Q_{ed}	$(\bar{e}_p \gamma_\mu e_r)(\bar{d} \gamma^\mu d)$	Q_{ee}	$(\bar{e}_p \gamma_\mu e_r)(\bar{e}_s \gamma^\mu e_t)$		
Q_{et}	$(\bar{e}_p \gamma_\mu e_r)(\bar{t} \gamma^\mu t)$	Q_{eb}	$(\bar{e}_p \gamma_\mu e_r)(\bar{b} \gamma^\mu b)$	Q_{tt}	$(\bar{t} \gamma_\mu t)(\bar{t} \gamma^\mu t)$	Q_{bb}	$(\bar{b} \gamma_\mu b)(\bar{b} \gamma^\mu b)$
$Q_{uu}^{(1)}$	$(\bar{u} \gamma_\mu u)(\bar{u} \gamma^\mu u)$	$Q_{uu}^{(8)}$	$(\bar{u} T^a \gamma_\mu u)(\bar{u} T^a \gamma^\mu u)$	$Q_{tu}^{(1)}$	$(\bar{t} \gamma_\mu t)(\bar{u} \gamma^\mu u)$	$Q_{tu}^{(8)}$	$(\bar{t} T^a \gamma_\mu t)(\bar{u} T^a \gamma^\mu u)$
$Q_{dd}^{(1)}$	$(\bar{d} \gamma_\mu d)(\bar{d} \gamma^\mu d)$	$Q_{dd}^{(8)}$	$(\bar{d} T^a \gamma_\mu d)(\bar{d} T^a \gamma^\mu d)$	$Q_{bd}^{(1)}$	$(\bar{b} \gamma_\mu b)(\bar{d} \gamma^\mu d)$	$Q_{bd}^{(8)}$	$(\bar{b} T^a \gamma_\mu b)(\bar{d} T^a \gamma^\mu d)$
$Q_{ud}^{(1)}$	$(\bar{u} \gamma_\mu u)(\bar{d} \gamma^\mu d)$	$Q_{ud}^{(8)}$	$(\bar{u} T^a \gamma_\mu u)(\bar{d} T^a \gamma^\mu d)$	$Q_{td}^{(1)}$	$(\bar{t} \gamma_\mu t)(\bar{d} \gamma^\mu d)$	$Q_{td}^{(8)}$	$(\bar{t} T^a \gamma_\mu t)(\bar{d} T^a \gamma^\mu d)$
$Q_{ub}^{(1)}$	$(\bar{u} \gamma_\mu u)(\bar{b} \gamma^\mu b)$	$Q_{ub}^{(8)}$	$(\bar{u} T^a \gamma_\mu u)(\bar{b} T^a \gamma^\mu b)$	$Q_{tb}^{(1)}$	$(\bar{t} \gamma_\mu t)(\bar{b} \gamma^\mu b)$	$Q_{tb}^{(8)}$	$(\bar{t} T^a \gamma_\mu t)(\bar{b} T^a \gamma^\mu b)$
$Q_{utbd}^{(1)}$	$(Y_u Y_d^\dagger)_{pr}(\bar{u}_p \gamma_\mu t)(\bar{b} \gamma^\mu d_r)$	$Q_{utbd}^{(8)}$	$(Y_u Y_d^\dagger)_{pr}(\bar{u}_p T^a \gamma_\mu t)(\bar{b} T^a \gamma^\mu d_r)$				
$\mathcal{L}_6^{(8c)} - (\bar{L} L)(\bar{R} R)$							
Q_{lu}	$(\bar{l}_p \gamma_\mu l_r)(\bar{u} \gamma^\mu u)$	Q_{ld}	$(\bar{l}_p \gamma_\mu l_r)(\bar{d} \gamma^\mu d)$	Q_{qe}	$(\bar{q} \gamma_\mu q)(\bar{e}_p \gamma^\mu e_r)$	Q_{le}	$(\bar{l}_p \gamma_\mu l_r)(\bar{e}_s \gamma^\mu e_t)$
Q_{lt}	$(\bar{l}_p \gamma_\mu l_r)(\bar{t} \gamma^\mu t)$	Q_{lb}	$(\bar{l}_p \gamma_\mu l_r)(\bar{b} \gamma^\mu b)$	Q_{Qe}	$(\bar{Q} \gamma_\mu Q)(\bar{e}_p \gamma^\mu e_r)$		
$Q_{qu}^{(1)}$	$(\bar{q} \gamma_\mu q)(\bar{u} \gamma^\mu u)$	$Q_{Qu}^{(1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{u} \gamma^\mu u)$	$Q_{qt}^{(1)}$	$(\bar{q} \gamma_\mu q)(\bar{t} \gamma^\mu t)$	$Q_{Qt}^{(1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{t} \gamma^\mu t)$
$Q_{qu}^{(8)}$	$(\bar{q} T^a \gamma_\mu q)(\bar{u} T^a \gamma^\mu u)$	$Q_{Qu}^{(8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{u} T^a \gamma^\mu u)$	$Q_{qt}^{(8)}$	$(\bar{q} T^a \gamma_\mu q)(\bar{t} T^a \gamma^\mu t)$	$Q_{Qt}^{(8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{t} T^a \gamma^\mu t)$
$Q_{qd}^{(1)}$	$(\bar{q} \gamma_\mu q)(\bar{d} \gamma^\mu d)$	$Q_{Qd}^{(1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{d} \gamma^\mu d)$	$Q_{qb}^{(1)}$	$(\bar{q} \gamma_\mu q)(\bar{b} \gamma^\mu b)$	$Q_{Qb}^{(1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{b} \gamma^\mu b)$
$Q_{qd}^{(8)}$	$(\bar{q} T^a \gamma_\mu q)(\bar{d} T^a \gamma^\mu d)$	$Q_{Qd}^{(8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{d} T^a \gamma^\mu d)$	$Q_{qb}^{(8)}$	$(\bar{q} T^a \gamma_\mu q)(\bar{b} T^a \gamma^\mu b)$	$Q_{Qb}^{(8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{b} T^a \gamma^\mu b)$
$Q_{qQtu}^{(1)}$	$(Y_u^\dagger)_{pr}(\bar{q}_p \gamma_\mu Q)(\bar{t} \gamma^\mu u_r)$	$Q_{qQtu}^{(8)}$	$(Y_u^\dagger)_{pr}(\bar{q}_p T^a \gamma_\mu Q)(\bar{t} T^a \gamma^\mu u_r)$	$Q_{qQbd}^{(1)}$	$(Y_d^\dagger)_{pr}(\bar{q}_p \gamma_\mu Q)(\bar{b} \gamma^\mu d_r)$	$Q_{qQbd}^{(8)}$	$(Y_d^\dagger)_{pr}(\bar{q}_p T^a \gamma_\mu Q)(\bar{b} T^a \gamma^\mu d_r)$
$\mathcal{L}_6^{(8d)} - (\bar{L} R)(\bar{R} L), (\bar{L} R)(\bar{L} R)$							
Q_{ledq}	$(\bar{l}_p^j e_r)(\bar{d} Y_d q_j)$	Q_{lebQ}	$(\bar{l}_p^j e_r)(\bar{b} Q_j)$	$Q_{leQt}^{(1)}$	$(\bar{l}_p^j e_r) \epsilon_{jk} (\bar{Q}^k t)$	$Q_{leQt}^{(3)}$	$(\bar{l}_p^j \sigma_{\mu\nu} e_r) \epsilon_{jk} (\bar{Q}^k \sigma^{\mu\nu} t)$
$Q_{lequ}^{(1)}$	$(\bar{l}_p^j e_r) \epsilon_{jk} (\bar{q}^k Y_u^\dagger u)$	$Q_{lequ}^{(3)}$	$(\bar{l}_p^j \sigma_{\mu\nu} e_r) \epsilon_{jk} (\bar{q}^k Y_u^\dagger \sigma^{\mu\nu} u)$	$Q_{leQb}^{(1)}$	$(\bar{Q}^j t) \epsilon_{jk} (\bar{Q}^k b)$	$Q_{leQb}^{(8)}$	$(\bar{Q}^j T^a t) \epsilon_{jk} (\bar{Q}^k T^a b)$
$Q_{quqd}^{(1)}$	$(\bar{q}^j Y_u^\dagger u) \epsilon_{jk} (\bar{q}^k Y_d^\dagger d)$	$Q_{quqd}^{(8)}$	$(\bar{q}^j Y_u^\dagger T^a u) \epsilon_{jk} (\bar{q}^k Y_d^\dagger T^a d)$	$Q_{quqd}^{(1)'} $	$(Y_u^\dagger)_{sr} (Y_d^\dagger)_{pt} (\bar{q}_p^j u_r) \epsilon_{jk} (\bar{q}_s^k d_t)$	$Q_{quqd}^{(8)'}$	$(Y_u^\dagger)_{sr} (Y_d^\dagger)_{pt} (\bar{q}_p^j T^a u_r) \epsilon_{jk} (\bar{q}_s^k T^a d_t)$
$Q_{Qtqd}^{(1)}$	$(\bar{Q}^j t) \epsilon_{jk} (\bar{q}^k Y_d^\dagger d)$	$Q_{Qtqd}^{(8)}$	$(\bar{Q}^j T^a t) \epsilon_{jk} (\bar{q}^k Y_d^\dagger T^a d)$	$Q_{quQb}^{(1)}$	$(\bar{q}^j Y_u^\dagger u) \epsilon_{jk} (\bar{Q}^k b)$	$Q_{quQb}^{(8)}$	$(\bar{q}^j Y_u^\dagger T^a u) \epsilon_{jk} (\bar{Q}^k T^a b)$
$Q_{Quqb}^{(1)}$	$(Y_u^\dagger)_{pr} (\bar{Q}^j u_r) \epsilon_{jk} (\bar{q}_p^k b)$	$Q_{Quqb}^{(8)}$	$(Y_u^\dagger)_{pr} (\bar{Q}^j T^a u_r) \epsilon_{jk} (\bar{q}_p^k T^a b)$	$Q_{qtQd}^{(1)}$	$(Y_d^\dagger)_{pr} (\bar{q}_p^j t) \epsilon_{jk} (\bar{Q}^k d_r)$	$Q_{qtQd}^{(8)}$	$(Y_d^\dagger)_{pr} (\bar{q}_p^j T^a t) \epsilon_{jk} (\bar{Q}^k T^a d_r)$

Table 2.5: Basis of fermionic operators for the topU31 flavor assumptions. Here (q, u, d) , Y_u, Y_d denote quarks of the first 2 generations and their 2×2 Yukawa matrices. Quark fields of the 3rd generation are (Q, t, b) . Flavor indices p, r, s, t run over $\{1, 2\}$ for light quarks and $\{1, 2, 3\}$ for leptons. Whenever flavor indices are not specified, they are implicitly contracted within each current.

Field Redefinitions

Some operators in the SMEFT generate terms that are the same form as those in \mathcal{L}_{SM} and act as a simple scaling of the SM terms. For example, the kinetic terms of the physical Higgs boson field become:

$$\mathcal{L}_{\text{SM}} + \mathcal{L}_6 = \frac{1}{2} \partial_\mu \partial^\mu h [1 - 2\Delta\kappa_H] + \dots \quad (2.87)$$

where

$$\Delta\kappa_H = \bar{C}_{H\Box} - \frac{\bar{C}_{HD}}{4}, \quad \text{and we define } \bar{C} \equiv \frac{v^2}{\Lambda^2} C_i. \quad (2.88)$$

The original particle basis can be recovered if we redefine the Higgs boson field as:

$$h \rightarrow [1 + \Delta\kappa_H] h \quad (2.89)$$

and the effect of this is an overall scaling of the SM Higgs couplings by factors of $1 + \Delta\kappa_H$. Therefore, the Wilson coefficients, $C_{H\Box}$ and C_{HD} should appear in the SMEFT parameterization of any Higgs boson process.

Similarly, the Q_{HWB} operator introduces a kinetic mixing between the B and W^3 fields of the form:

$$\mathcal{L}_6 = -\frac{C_{HWB}}{2} \frac{v^2}{\Lambda^2} W_{\mu\nu}^3 B^{\mu\nu} + \dots \quad (2.90)$$

which can be removed if we redefine the fields with a rotation:

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -\bar{C}_{HWB}/2 \\ -\bar{C}_{HWB}/2 & 1 \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}. \quad (2.91)$$

which leads to a shift in the Weinberg angle which is now:

$$\tan \theta_W = \frac{g_1}{g_W} + \frac{1}{2} \bar{C}_{HWB} \left(1 - \frac{g_1^2}{g_W^2} \right) \quad (2.92)$$

Therefore, C_{HWB} introduces modifications to all couplings involving a photon or a Z boson.

Input Parameters

To determine the free parameters of the SM (Table 2.1), a sufficiently large set of independent observables, \mathcal{O} , must be chosen, measured, and then compared with SM predictions for the observables that are functions of the free parameters. In the SMEFT, the observables, \mathcal{O} , can receive contributions from higher-dimension operators, and therefore, the corresponding shifts in the SM free parameters must be calculated. For example, with

the `topU31` flavour assumption, the shift in the Fermi constant is given by:

$$\Delta G_F = 2\bar{C}_{Hl}^{(3)} - \bar{C}'_{ll} \quad (2.93)$$

and therefore, $C_{Hl}^{(3)}$ and C'_{ll} lead to a scaling for electroweak processes in a similar way to how field redefinitions lead to a scaling in Higgs processes by $C_{H\Box}$ and C_{HD} .

In the electroweak gauge sector, there are four free parameters: g_1, g_2, v, λ which are usually determined by using 4 observables from the set:

$$\{\alpha_{em}, G_F, m_Z, m_W, m_H\}. \quad (2.94)$$

The Higgs mass, m_H , is always used to fix λ , but the choice of the other three observables is free. This choice will have consequences for which SM parameters obtain shifts and this can affect the validity of an EFT parameterization. If shifts are applied to m_Z and/or m_W , the Wilson coefficient dependence can become highly non-linear in diagrams containing the Z or W bosons as a propagator [35] since the relevant term in the matrix element is proportional to:

$$\frac{1}{q^2 - (m_V + \Delta m_V)^2 + im_V \Gamma_V}, \quad m_V \in \{m_Z, m_W\}. \quad (2.95)$$

For this reason, the $\{G_F, m_Z, m_W\}$ scheme is preferred where m_Z and m_W are fixed to their measured values. In this scheme, the electroweak parameter shifts are:

$$g_1^2 = \hat{g}_1^2 \left[1 + 2 \frac{\delta g_1}{\hat{g}_1} \right], \quad \frac{\delta g_1}{\hat{g}_1} = -\frac{1}{2} \left[\Delta G_F + \frac{\Delta m_Z^2}{\sin^2 \theta_W} \right], \quad (2.96)$$

$$g_W^2 = \hat{g}_W^2 \left[1 + 2 \frac{\delta g_W}{\hat{g}_W} \right], \quad \frac{\delta g_W}{\hat{g}_W} = -\frac{\Delta G_F}{2}, \quad (2.97)$$

$$v_T^2 = \hat{v}^2 \left[1 + 2 \frac{\delta v}{\hat{v}} \right], \quad \frac{\delta v}{\hat{v}} = \frac{\Delta G_F}{2}, \quad (2.98)$$

$$\lambda = \hat{\lambda} \left[1 - \frac{\delta \lambda}{\hat{\lambda}} \right], \quad \frac{\delta \lambda}{\hat{\lambda}} = -\Delta G_F - \Delta m_h^2, \quad (2.99)$$

where the \hat{g} notation refers to the SM (unshifted) values of these parameters and Δm_Z^2 and Δm_h^2 are given by:

$$\Delta m_Z^2 = \frac{\bar{C}_{HD}}{2} + \frac{2g_1 g_W}{g_1^2 + g_W^2} \bar{C}_{HWB} \quad (2.100)$$

$$\Delta m_h^2 = 2\Delta\kappa_H - \frac{3}{2\lambda} \bar{C}_H. \quad (2.101)$$

The $\{G_F, m_Z, m_W\}$ scheme is adopted in the EFT interpretation in Section 2.3.3.

In the Yukawa sector, the input parameters are the fermion masses, from which the Yukawa couplings are inferred. Therefore, operators that lead to fermion mass terms, propagate into shifts in the Yukawa couplings where:

$$Y_\psi \rightarrow \hat{Y}_\psi + \delta Y_\psi, \quad \delta Y_\psi = -\frac{\Delta G_F}{2} \hat{Y}_\psi + \Delta M_\psi \quad (2.102)$$

and with the `topU31` flavour assumption:

$$\Delta M_l = \frac{1}{2} \bar{C}_{eH}^* Y_l, \quad \Delta M_{u,c} = \frac{1}{2} \bar{C}_{uH}^* Y_u, \quad \Delta M_t = \frac{1}{2} \bar{C}_{tH}^* \quad (2.103)$$

$$\Delta M_{d,s} = \frac{1}{2} \bar{C}_{dH}^* Y_d, \quad \Delta M_b = \frac{1}{2} \bar{C}_{bH}^*. \quad (2.104)$$

Therefore, all $h\psi\psi$ couplings will be affected by one of C_{eH} , C_{uH} , C_{tH} , C_{dH} or C_{bH} .

General Form for SMEFT Parameterizations

To constrain the Wilson coefficients, we need to first parameterize some observables in terms of the coefficients. Consider the matrix element, \mathcal{M} , for a generic process which can be expressed as:

$$\mathcal{M} = \mathcal{M}_{SM} + \mathcal{M}_{\mathcal{L}_6} \quad (2.105)$$

where \mathcal{M}_{SM} is the SM matrix element, and $\mathcal{M}_{\mathcal{L}_6}$ is the matrix element resulting from \mathcal{L}_6 contributions, where the contributions can generally be split into those that generate new vertices, e.g. the four-point vertex for muon decay (Fig. 2.23), and those that modify existing vertices, e.g. through field redefinitions and input parameter shifts. The latter can also be interpreted as generating new vertices if the modified SM vertices are considered separately as the original SM vertex and a new vertex that represents the shift in the vertex coupling strength [34].

If we restrict ourselves to Feynman diagrams with single insertions of SMEFT vertices, the SMEFT matrix element becomes:

$$\mathcal{M}_{\mathcal{L}_6} = \frac{1}{\Lambda^2} \sum_i C_i \mathcal{M}_i \quad (2.106)$$

and squaring the total matrix element and taking the ratio to $|\mathcal{M}_{SM}|^2$ gives us:

$$\frac{|\mathcal{M}|^2}{|\mathcal{M}_{SM}|^2} = 1 + \frac{1}{\Lambda^2} \sum_i \frac{2 \operatorname{Re}(\mathcal{M}_{SM}^* \mathcal{M}_i)}{|\mathcal{M}_{SM}|^2} C_i + \frac{1}{\Lambda^4} \sum_{ij} \frac{2 \operatorname{Re}(\mathcal{M}_i^* \mathcal{M}_j)}{|\mathcal{M}_{SM}|^2} C_i C_j. \quad (2.107)$$

$$= 1 + \frac{1}{\Lambda^2} \sum_i \alpha_i C_i + \frac{1}{\Lambda^4} \sum_{ij} \beta_{ij} C_i C_j. \quad (2.108)$$

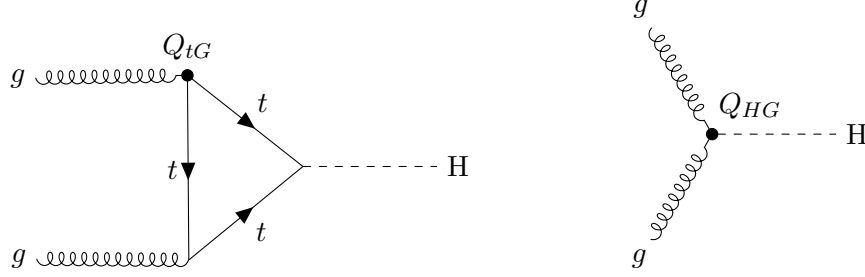


Figure 2.24: LO Feynman diagrams for ggH production in the SM (left) and the SMEFT (left and right). In the SM, gluons do not directly couple to the Higgs boson so the LO diagram is loop-induced, mediated predominately by top quarks. In the SMEFT, the SM diagram is affected by contributions to the ttG vertex by the Q_{tG} operator. There also exists a new diagram (right) where the gluons couple directly to the Higgs boson which is possible with the introduction of Q_{HG} .

For observables, \mathcal{O} , like cross sections and decay widths that can be expressed as integrals of $|\mathcal{M}|^2$ over phase space, the SMEFT parameterization then takes the general quadratic form:

$$\mu(\vec{C}) \equiv \frac{\mathcal{O}(\vec{C})}{\mathcal{O}_{SM}} = 1 + \frac{1}{\Lambda^2} \sum_i A_i C_i + \frac{1}{\Lambda^4} \sum_{ij} B_{ij} C_i C_j \quad (2.109)$$

where A_i and B_{ij} are real constants, the sum over i and j are sums over all contributing Wilson coefficients, and \mathcal{O}_{SM} is the SM prediction for the observable. In this thesis, Λ is arbitrarily taken to be 1 TeV and the parameterizations and constraints presented can be obtained at a different value, $\Lambda = X$, by scaling the values by a factor of $(X/1 \text{ TeV})^2$.

Gluon-Gluon Fusion Higgs Boson Production

In the SM, gluon-gluon fusion Higgs boson production (ggH) is the dominant production mode for the Higgs boson (see Table 2.2). Furthermore, in BSM theories where the top-quark loop (see left of Fig. 2.24) can be replaced by a loop containing a new particle, BSM contributions enter at the same order as the SM, meaning that ggH is particularly sensitive to new physics. This process is therefore important for the SMEFT interpretation of Higgs boson measurements in Chapter 6.

In ggH, the \mathcal{L}_6 contributions from field redefinitions and input parameter shifts are given, to linear order, by:

$$\mu_{ggH} = 1 + 2\Delta\kappa_H + \Delta G_F - 2|\Delta M_t/Y_t| \quad (2.110)$$

$$= 1 + \frac{1}{4} \frac{v^2}{\Lambda^2} \left(4C_{H\Box} - C_{HD} - 4C_{Hl}^{(3)} + 2C'_l - 4|C_{tH}| \right) \quad (2.111)$$

$$\approx 1 + 0.12C_{H\Box} - 0.03C_{HD} - 0.12C_{Hl}^{(3)} + 0.06C'_l - 0.12|C_{tH}| \quad (2.112)$$

where we have used $v = 246 \text{ GeV}$ and $\Lambda = 1 \text{ TeV}$. In the second group, there are contributions from Q_{tG} and Q_{HG} :

$$Q_{tG} = (\bar{Q}\sigma^{\mu\nu}T^a t)\tilde{H}G_{\mu\nu}^a, \quad Q_{HG} = H^\dagger H G_{\mu\nu}^a G^{a\mu\nu} \quad (2.113)$$

which after electroweak symmetry breaking, lead to terms including:

$$Q_{tG} = \frac{\tilde{v}}{\sqrt{2}}(\bar{Q}\sigma^{\mu\nu}T^a t)G_{\mu\nu}^a + \dots, \quad Q_{HG} = 12vhG_{\mu\nu}^a G^{a\mu\nu} + \dots \quad (2.114)$$

which in turn lead to ttg and ggh vertices. Therefore, contributions from Q_{tG} appear in the loop-induced diagram at the ttg vertices, and contributions from Q_{HG} arise via the introduction of a new diagram, shown in Fig. 2.24 (right) containing only a three-point vertex, ggh . The magnitude of these contributions depend on the energy of an interaction and therefore, are more easily calculated with MC simulation. The relevant details for this are left for Chapter 6.

Higgs Boson to Four Leptons Decay

The $H \rightarrow l^+ l^- l^+ l^-$ ($H \rightarrow 4l$) decay channel, is one of the most sensitive decay channels when making measurements of the Higgs boson. Therefore, like ggH production, it is important in the EFT interpretation of Chapter 6. At LO in the SM, this process proceeds via mediating Z bosons: $H \rightarrow ZZ^* \rightarrow 4l$, where at least one Z boson must be off-shell, and the other is predominately on-shell. The Feynman diagram for this is shown in Fig. 2.25. As with ggH production, the global scaling of Higgs couplings leads to the following terms:

$$\mu_{H \rightarrow 4l} = 1 + 0.12C_{H\Box} - 0.03C_{HD} + \dots \quad (2.115)$$

but with $H \rightarrow 4l$, there are no terms relating to scaling of the Yukawa couplings, and the contributions of Q_{Hl}^3 and Q'_{ll} are non-trivial to calculate because, as well as global scalings of the electroweak couplings, they also lead to new diagrams involving a $hZll$ vertex (see Fig. 2.25).

Further contributions arise from the Q_{HW} , Q_{HB} and Q_{HWB} operators:

$$Q_{HW} = H^\dagger H W_{\mu\nu}^i W^{I\mu\nu}, \quad Q_{HB} = H^\dagger H B_{\mu\nu} B^{\mu\nu}, \quad Q_{HWB} = H^\dagger \sigma^i H W_{\mu\nu}^i B^{\mu\nu} \quad (2.116)$$

which after electroweak symmetry breaking, leads to terms involving $hZ\gamma$ and $h\gamma\gamma$ vertices. These lead to the Feynman diagrams shown in Fig. 2.26 which lead to significant enhancements of the $H \rightarrow 4l$ process where the invariant mass of an opposite-sign lepton pair, m_{ll} , is small. This is because the photon propagator term (Eq. 2.95) is proportional

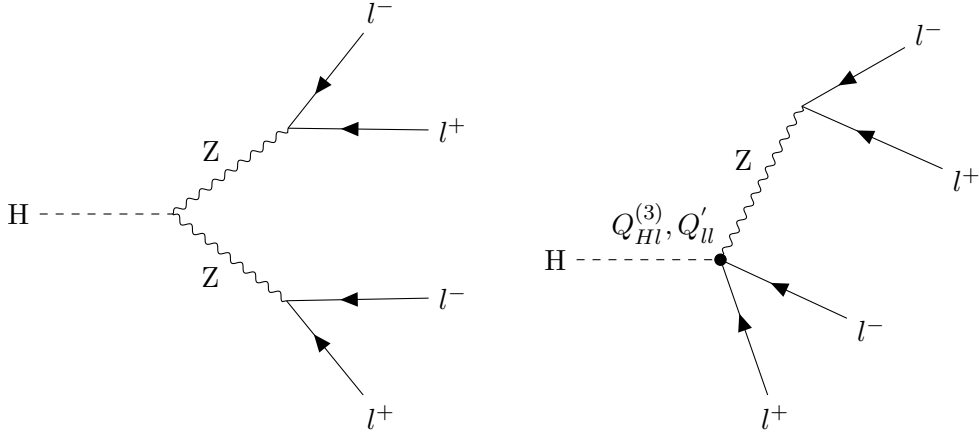


Figure 2.25: LO Feynman diagram for the $H \rightarrow 4l$ decay in the SM (left) and a new diagram from \mathcal{L}_6 (right) with a $hZll$ vertex that is generated by the $Q_{Hl}^{(3)}$ and Q'_{ll} operators.

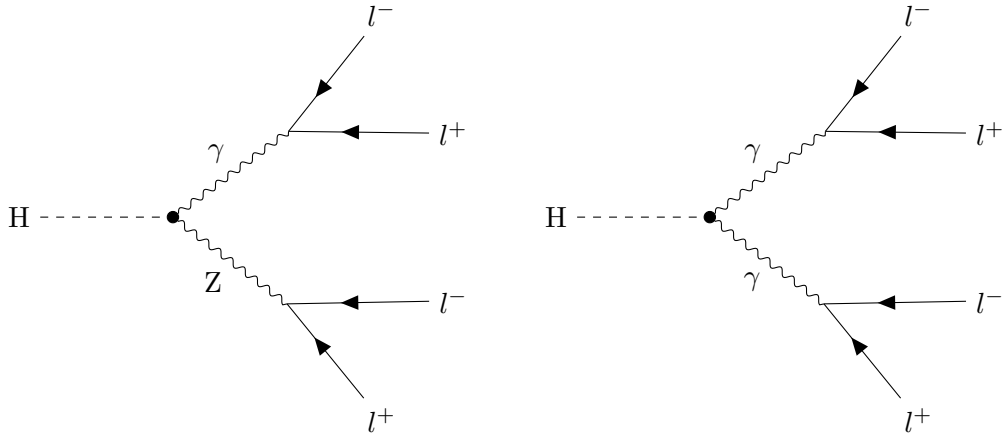


Figure 2.26: LO Feynman diagrams for the $H \rightarrow 4l$ decay that include $hZ\gamma$ (left) and $h\gamma\gamma$ (right) vertices that are generated from the Q_{HW} , Q_{HB} and Q_{HWB} operators.

to $1/q^2$ since the photon is massless. This can cause issues for reweighting-based approaches to calculating the Q_{HW} , Q_{HB} and Q_{HWB} contributions since these techniques rely on the assumption that only small changes to the distributions of kinematic variables occur. Furthermore, the parameterization is highly dependent on any selection placed on m_{ll} , meaning that consideration from the analysis selection is important. Both of these issues are discussed further in Chapter 6.

A further consideration in the $H \rightarrow 4l$ process, and others containing a W or Z boson propagators, is the total width, Γ , of the bosons. These acquire their own \mathcal{L}_6 corrections and since the propagator term (Eq. 2.95) contains the total width, this leads to another correction to these types of processes. These so-called *propagator corrections* are dominant in the on-shell regime where $q^2 \sim m_V^2$. Therefore, we can approximate their impact with

the on-shell expression. In the limit $q^2 = m_V^2$, the propagator term becomes:

$$\frac{1}{im_V \Gamma_V} = \frac{1}{im_V \Gamma_V^{\text{SM}}} \left(1 + \frac{\delta \Gamma_V}{\Gamma_V^{\text{SM}}} \right)^{-1} \approx \frac{1}{im_V \Gamma_V^{\text{SM}}} \left(1 - \frac{\delta \Gamma_V}{\Gamma_V^{\text{SM}}} \right) \quad (2.117)$$

where the expansion is fair if $\delta \Gamma_V \ll \Gamma_V$. In the SMEFT interpretation of Chapter 6, these corrections are applied with the SMEFTSIM tool [34] which parameterizes $\delta \Gamma_V$ to linear order in the Wilson coefficients and then propagates this effect to the parameterization of cross sections and partial widths using the expansion in Eq. 2.117.

Chapter 3

The CMS Experiment

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [36] is the world’s largest and most powerful accelerator. It is situated at CERN, the European Organization for Nuclear Research, near Geneva, Switzerland, in a circular tunnel of 27 km circumference about 100 m underground. It is designed to accelerate hadrons to high enough energies to study fundamental physics at the electroweak scale. The hadrons collide at four interaction points (IPs), where the ATLAS [37], CMS [38], ALICE [39], and LHCb [40] experiments are located. More description about the CMS experiment, which the work presented in this thesis is based on, is given in the next section. The rest of this section will describe the LHC operation with protons as it was during the 2015–2018 data-taking period, which is the period that measurements presented in this thesis were taken in.

The LHC does not operate as a single, independent machine, rather it is the final destination in a chain of accelerators that prepare protons for collision. This series of accelerators are part of the CERN accelerator complex, shown in Fig. 3.1. Hydrogen gas is ionized to create a source of protons, which is then accelerated by the Linear Accelerator 2 (LINAC 2) to 50 MeV, then to 1.4 GeV in the BOOSTER, and then to 25 GeV by the Proton Synchrotron (PS). Here, the protons are collected into bunches, with about 10^{11} protons per bunch and a spacing of 25 ns, before being injected into the Super Proton Synchrotron (SPS) where they are accelerated to 450 GeV. Finally, up to 2556 bunches of protons are injected into the LHC in two counter-rotating beams, and accelerated to their final energy of 6.5 TeV per beam before they are allowed to collide at the four interaction points at a centre-of-mass energy of $\sqrt{s} = 13$ TeV.

The number of events produced at an LHC experiment for a particular physics process

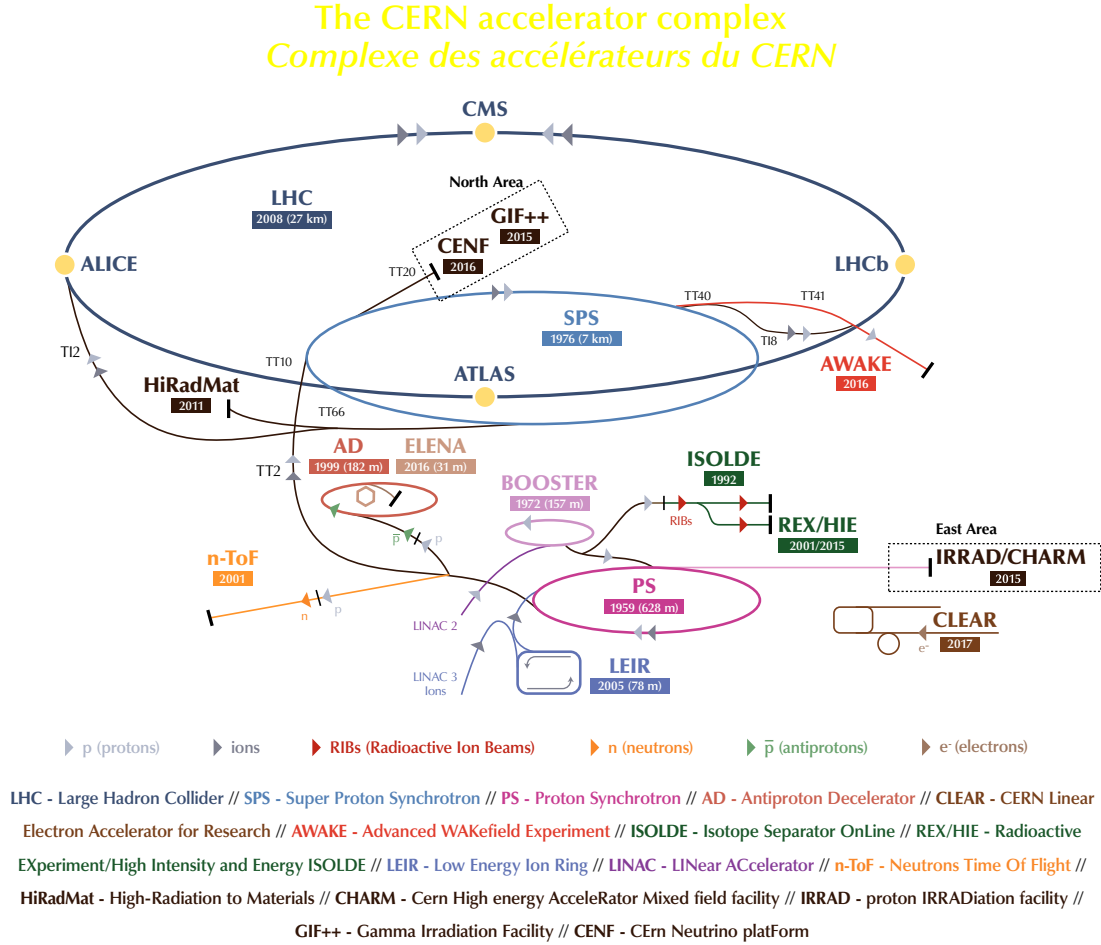


Figure 3.1: An illustration of CERN accelerator complex as of August 2018. The LHC, with its four interaction points and associated experiments, is shown together with the series of preceding accelerators, those being LINAC 2, BOOSTER, PS and SPS. Additional accelerators and beam destinations related to other CERN activities are also shown. The figure is taken from Ref. [41].

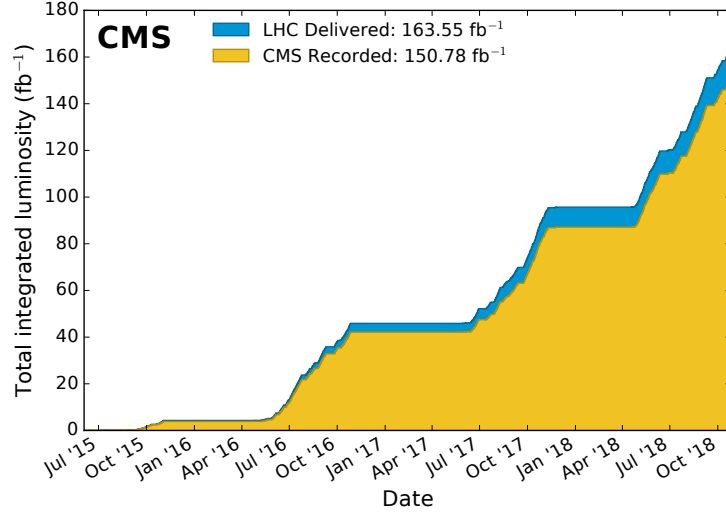


Figure 3.2: Delivered and recorded luminosity at CMS, cumulative over 2015 to 2018. Figure taken from Ref. [45].

is given by:

$$\sigma(\sqrt{s}) \int_{\text{time}} \mathcal{L}_{\text{int}} \quad (3.1)$$

where σ is the cross section for the process and \mathcal{L}_{int} is the integrated luminosity. During the 2015–2018 data-taking period, the LHC delivered a total integrated luminosity of 164 fb^{-1} to the CMS experiment. However, the experiment was not recording data all the time, meaning that the recorded integrated luminosity was instead 151 fb^{-1} . Figure 3.2 shows how this luminosity was cumulated from 2015 to 2018. A further reduction in the luminosity occurs when only considering the data that is certified for physics analysis, leaving 2.3 fb^{-1} , 36.3 fb^{-1} , 41.5 fb^{-1} , and 59.8 fb^{-1} for 2015, 2016, 2017, and 2018 respectively [42–44]. In the majority of CMS analyses using data collected during this period, the 2015 dataset is discarded since it is not big enough to justify the additional work required to account for year-dependent operating conditions. The work in this thesis follows suit, using only the 2016–2018 dataset, corresponding to 138 fb^{-1} in total.

3.2 The CMS Detector

3.2.1 Introduction

The Compact Muon Solenoid (CMS) [38] is one of two general-purpose detectors at the LHC. One of the main goals for CMS was the discovery of the Higgs boson, which since its observation in 2012 [6, 7], has evolved into a goal of understanding the Higgs boson's properties and assessing their compatibility with the SM. Further motivation for the CMS experiment is the search for physics beyond the Standard Model, like supersymmetry, extra dimensions, and dark matter. These physics goals led to the following priorities for the design of the CMS detector:

1. good muon identification and momentum resolution, good dimuon mass resolution, and the ability to unambiguously determine the charge of muons;
2. good charged-particle momentum resolution and reconstruction efficiency, and efficient triggering and tagging of tau leptons and b-jets, requiring pixel detectors close to the interaction point;
3. good electromagnetic energy resolution and good diphoton and dielectron mass resolution, as well as efficient π^0 rejection;
4. good dijet-mass resolution, and good missing momentum resolution, requiring hermetic geometric coverage.

The centre-of-mass energy provided by the LHC allows CMS to probe physics up to the TeV scale. Therefore, the detector is designed such that the above requirements should hold for particles with momenta up to 1 TeV.

Besides its collision energy, the LHC is also unprecedented in its luminosity, with bunch crossings occurring every 25 ns, and multiple interactions per bunch crossing. The number of coincident interactions per bunch crossing is referred to as *pileup*, and during the 2015–2018 data taking period, had an average value of 34 at CMS¹ [45]. This high luminosity environment requires a high level of spatial and timing granularity in the detector, and a fast readout system. Furthermore, high radiation levels are expected, which requires radiation-hard detectors and front-end electronics.

A schematic of the CMS detector is given in Fig. 3.3. The structure takes a cylindrical shape, with the beam pipe running through the central axis. The detector is divided into a barrel region, which covers the central region of the detector, and two endcaps, which cover

¹An experiment can have a reduced luminosity if desired by choosing to cross the LHC beams with a smaller overlap. Typically, the CMS and ATLAS experiments operate at the maximum possible luminosity whilst the LHCb experiment chooses a lower luminosity that is preferable for measuring b decays.

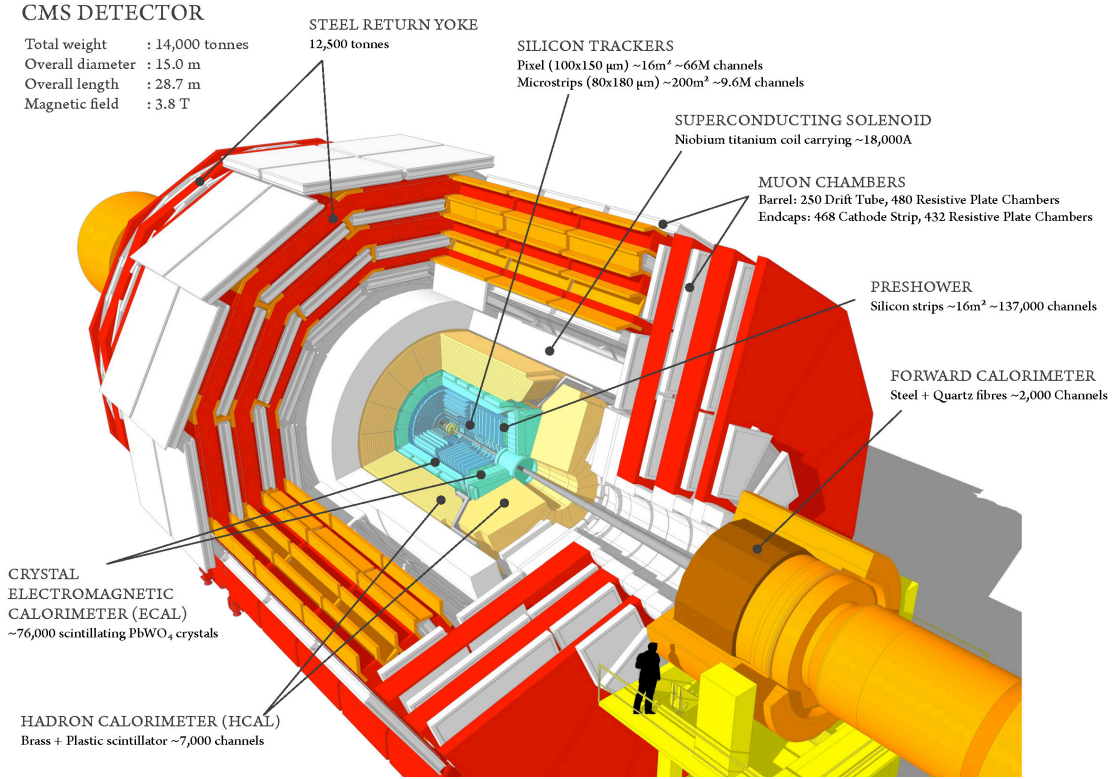


Figure 3.3: A schematic of the CMS detector. Taken from Ref. [46].

each end of the detector. Key to the design of CMS, is a 13-m-long, 6-m-inner-diameter, 3.8-T superconducting solenoid magnet centred around the beam pipe which provides a high bending power for charged particles which in turn allows for precise momentum measurements. The solenoid envelopes the inner tracker and calorimetry while the muon system is located outside the solenoid, and is embedded in the return yoke of the magnet.

The di-Higgs search of Chapter 5 uses all key aspects of the CMS detector since it involves the measurements of photons, and the measurement of tau leptons, whose decay products include charged and neutral hadrons, electrons, and muons. In the di-Higgs analysis, good resolution for the diphoton and ditau invariant masses is crucial to separate signal from background, and identify the mass of a new particle, if found. Therefore, the electromagnetic calorimeter (ECAL) and the inner tracker are of particular importance because they determine the energy and momentum of photons and charged particles (majority of tau lepton decay products) respectively.

3.2.2 Coordinate System

The CMS coordinate system has the origin centred at the nominal collision point, with the y-axis pointing vertically upwards, the x-axis pointing radially towards the centre of the LHC ring, and the z-axis pointing along the beam line in the direction of the counter-clockwise beam. An alternative, cylindrical coordinate system is also adopted, with the azimuthal angle ϕ measured from the x-axis in the x-y plane, the polar angle θ measured from the z-axis, and the radial coordinate r measured from the beam line. Instead of θ , pseudorapidity η is often used, defined as:

$$\eta = -\ln[\tan(\theta/2)]. \quad (3.2)$$

Particles with high values of η correspond to particles that are close to the beam line, and are said to be *forward*. Distance between particles in η - ϕ space is measured by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Transverse momentum, p_T , is defined as the magnitude of the component of momentum perpendicular to the beam line, and is given by $p_T = \sqrt{p_x^2 + p_y^2}$. In many processes of interest, the outgoing particles are produced with a large p_T and therefore this is an important quantity. Furthermore, since the incoming protons have zero p_T , the sum of p_T over all outgoing particles should be zero due to conservation of momentum. In cases where particles, such as neutrinos, are not detected, their total p_T can be inferred from the imbalance in the transverse momentum, which is called missing transverse momentum (p_T^{miss}). Similarly, transverse energy, E_T , is defined as $E_T = E \sin \theta$.

3.2.3 Inner Tracker

The inner tracker is designed to perform precise measurements of charged-particle trajectories and precise reconstruction of secondary vertices originating from the decay of tau leptons and long-lived hadrons. It surrounds the interaction point and has a length of 5.8 m and a diameter of 2.5 m. Silicon-based detector technology was chosen for its great granularity, fast response, and reasonable radiation hardness. The tracker is divided into two main components: the pixel detector, which is closest to the interaction point, and the strip tracker, which surrounds the pixel detector. Figure 3.4 shows a sketch of one quarter of the inner tracker.

Originally, the pixel detector consisted of three barrel layers at $r = 4.4$ cm, 7.3 cm and 10.2 cm, and two forward/backwards disks at $|z| = 34.5$ cm and 46.5 cm, extending in radius from about 6 cm to 15 cm. In total, the pixel detector had 66 million pixels, covering an area of 1.06 m² [38]. To accommodate higher luminosities, the pixel detector was upgraded with increased granularity in the end-of-year technical stop of the LHC

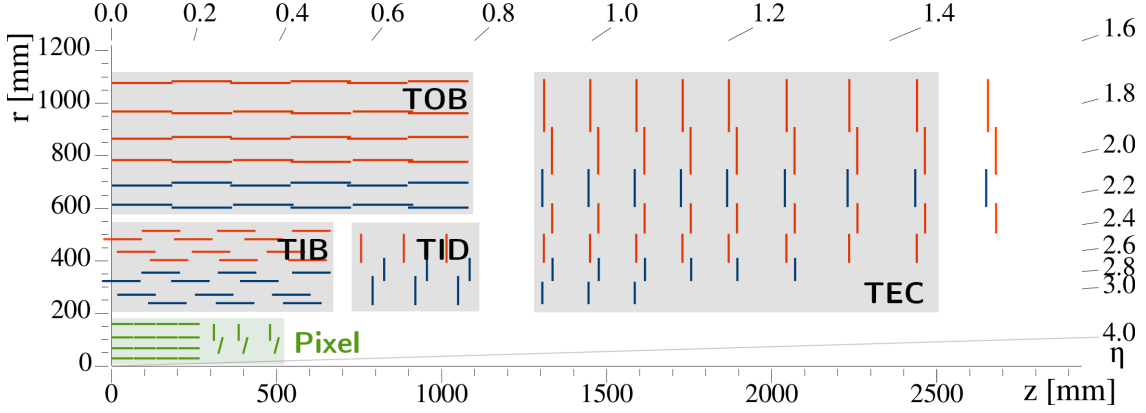


Figure 3.4: Sketch of one quarter of the CMS inner tracker after the 2016/2017 upgrade in the r - z plane. The pixel detector is shown in green, while the single-sided and double-sided strip modules are shown in red and blue respectively. Figure adapted from Ref. [49].

in 2016/2017 [47]. The new pixel detector consists of four barrel layers at $r = 2.9$ cm, 6.8 cm, 11 cm, and 16 cm, and three forward/backwards disks at $|z| = 29$ cm, 40 cm, and 52 cm, extending in radius from 4.5 cm to 16.1 cm. In total, the upgraded pixel detector has 124 million pixels, of size $100 \mu\text{m} \times 150 \mu\text{m}$, covering a total area of 1.9 m^2 [48]. This geometry is designed to measure four space points (hits) along a charged particle's trajectory, in the pseudorapidity range $|\eta| < 2.5$. This results in a spatial resolution of $12 \mu\text{m}$ in the transverse (r - ϕ) plane, and $22 \mu\text{m}$ in the longitudinal (z) direction [48].

At radii between 20 cm and 116 cm, lower occupancies allow for the use of strip detectors. The strip tracker consists of three subsystems. The Tracker Inner Barrel and Disks (TIB/TID) subsystem comprises four barrel layers and three forward/backward disks which extend to a radius of 55 cm and use $320 \mu\text{m}$ thick silicon microstrip sensors. Depending on the layer, the strip pitch² is 80 – $141 \mu\text{m}$ in the TIB/TID. With the strips laying parallel to the beam axis in the TIB, and radially in the TID, the TIB/TID provides up to 4 ϕ measurements per track, leading to a spatial resolution in $r\phi$ of 23 – $35 \mu\text{m}$. Extending to a radius of 116 cm, the Tracker Outer Barrel (TOB) subsystem consists of six barrel layers, which use $500 \mu\text{m}$ thick silicon microstrip sensors with a pitch of 122 – $183 \mu\text{m}$. The TOB provides up to 6 ϕ measurements per track, leading to a spatial resolution in $r\phi$ of 35 – $53 \mu\text{m}$.

The TID/TIB and TOB subsystems extend to $|z| = 118$ cm. Beyond this are the Tracker EndCaps (TEC), which extend to $|z| = 282$ cm and cover $22.5 < |\eta| < 113.5$ cm. Each endcap consists of nine disks, with up to 7 rings of silicon strip detectors of $320 \mu\text{m}$ thickness for the inner 4 rings, and $500 \mu\text{m}$ thickness for the outer 3 rings with radial

²Strip pitch refers to the centre-to-centre distance between adjacent strips

strips of 97–184 μm pitch. Therefore, they provide up to 9 measurements of ϕ per track.

Finally, the first two layers and rings of the TIB, TID and TOB, as well as rings 1,2 and 5 of the TECs are fitted with a second strip detector module which is mounted back to back with the first module. These modules have their strips rotated at an almost perpendicular angle to the first module, and therefore provide a measurement of a second coordinate, that being z in the barrel and r on the disks. The spatial resolution of these measurements are 230 μm and 530 μm in the TIB and TOB respectively, and varies with pitch in the TID and TEC.

This tracker layout ensures at least ≈ 9 hits per trajectory in the silicon strip tracker for $|\eta| < 2.4$, with the ultimate acceptance of the tracker ending at $|\eta| = 2.5$. In total, the strip tracker has 9.3 million strips and 198 m^2 of active silicon area [38].

3.2.4 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is designed to reconstruct the energy of electromagnetic showers originating from electrons and photons for $|\eta| < 3$. It is a homogenous calorimeter composed of scintillating crystals which emit light as a result of the charged particles created by electron and photon showers in the ECAL. By measuring the amount of light emitted, the energy of the particles can be inferred.

Lead-tungstate (PbWO_4) was chosen to be the crystal material due to the following properties:

- short radiation length: $X_0 = 0.89$ cm, allowing for a compact calorimeter;
- narrow Molière radius: $R_M = 2.2$ cm, resulting in a fine granularity;
- fast response time: 80% of the scintillation light is emitted within 25 ns;
- good radiation hardness.

The ECAL is divided into two main components: the barrel (EB) and the endcaps (EE). The EB covers the region $|\eta| < 1.479$, while the EE covers the region $1.479 < |\eta| < 3.0$ [38]. The overall structure of the ECAL is shown in Fig. 3.5.

The EB is composed of 61200 crystals, with a 360-fold granularity in ϕ , and a 170-fold granularity in η . The crystals are tapered, with a front and rear cross-section of 22×22 mm^2 and 26×26 mm^2 respectively. The length of the crystals is 230 mm ($25.8X_0$) and the front faces of the crystals are 1.3 m from the beam line. To avoid particles from the IP aligning with cracks between crystals, the crystals are arranged in a quasi-projective geometry such that they point 3° away from the vector that points towards the nominal IP.

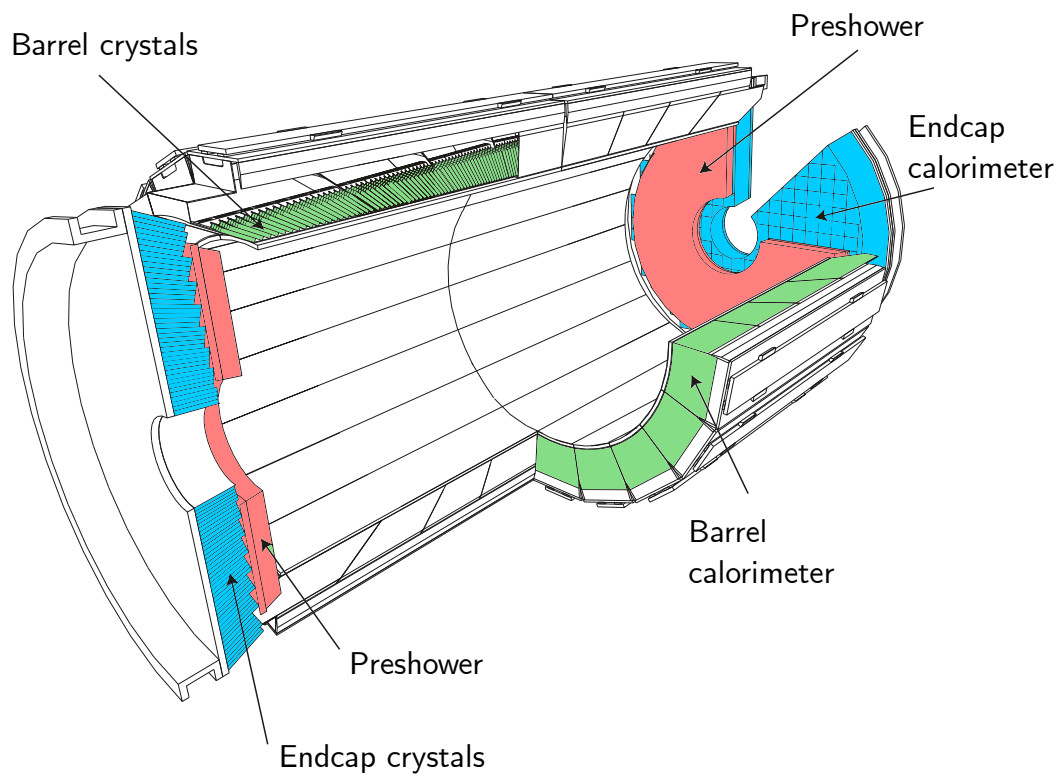


Figure 3.5: A schematic of the ECAL. Figure adapted from Ref. [38].

The EE is composed of 14648 crystals that are arranged 1.3 m away from the nominal IP. The crystals have a front and rear cross-section of $28.6 \times 28.6 \text{ mm}^2$ and $30.0 \times 30.0 \text{ mm}^2$ respectively, with a length of 220 mm ($24.7X_0$). Similar to the EB, the EE crystals are arranged in a quasi-projective geometry, and point towards a focus point 1.3 m away from the nominal IP, pointing between 2° and 8° from the nominal IP.

Additionally, a sampling calorimeter called the Preshower detector is placed in front of the endcap crystals. This detector has one layer of lead as an absorbing material, followed by two layers of orthogonally-placed silicon strip sensors of 1.9 mm pitch. It provides a precise position measurement of incident electrons and photons which is particularly useful for discriminating against photons originating from π^0 decays.

The Lead-tungstate (PbWO_4) crystals emit only 4.5 photoelectrons per MeV, which is relatively low compared to other scintillating materials. To compensate for this, photomultipliers are used to amplify the signal. In the EB, silicon avalanche photodiodes (APDs) are used, with a gain of 50, and in the EE, vacuum phototriodes (VPTs) are used, with a gain of 10. The choice of the respective photomultipliers is due to the different configuration of the magnetic field and expected levels of radiation found in the barrel and endcaps [38].

Finally, the intrinsic energy resolution of the ECAL is modelled by:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (3.3)$$

where E is expressed in units of GeV, and $S = 2.8\%$ is the stochastic term, $N = 12\%$ is the noise term, and $C = 0.3\%$ is the constant term, where the values are determined from test-beam data [38]. For photons with energy $E = \frac{1}{2}m_H \approx 62.5 \text{ GeV}$, the energy resolution is approximately 0.3 GeV (0.5%).

3.2.5 Hadronic Calorimeter

Quarks and gluons produced in proton-proton collisions are not directly detected. They instead hadronize into sprays of hadrons, called jets, which we can detect, and the measured properties of the jet can then be used to infer information about the original quark or gluon. The hadronic calorimeter (HCAL) is designed to measure the energy and position of the constituent hadrons, allowing the jet to be reconstructed. This is particularly important for neutral hadrons which do not leave any hits in the inner tracker and deposit little energy in the ECAL. Furthermore, the HCAL is designed to be as hermetic as possible, covering $|\eta| < 5.2$, and is deep enough such that the majority of the energy from visible particles in a collision are absorbed, except for muons which are measured in the muon system. This is crucial for a precise measurement of p_T^{miss} .

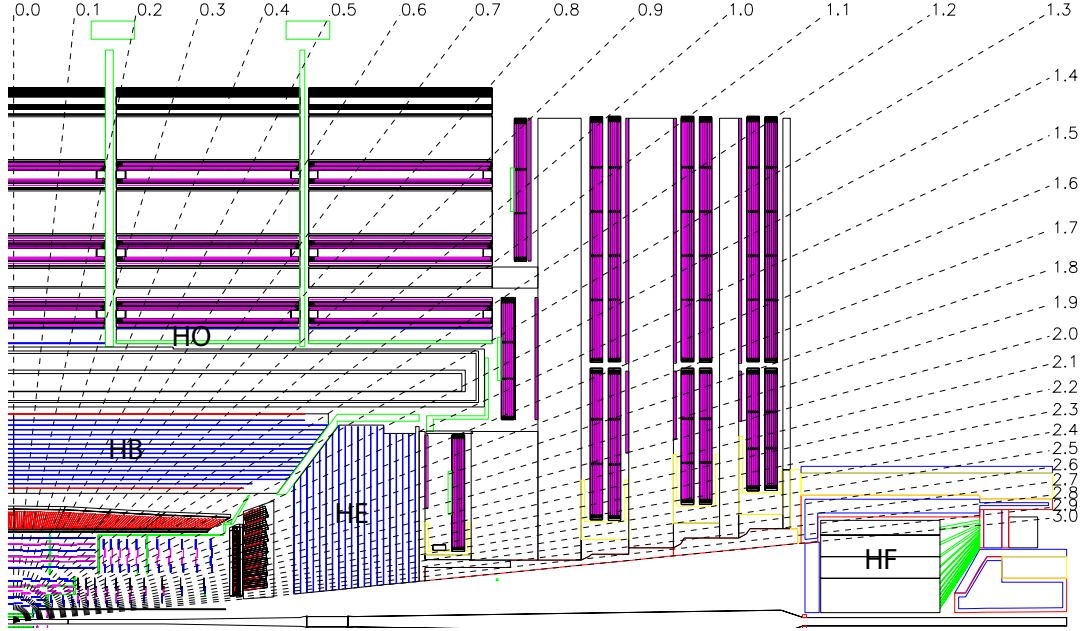


Figure 3.6: A schematic of one quarter of the HCAL in the r - z plane. The interaction point is the lower-left corner of the diagram and the dashed lines represent constant lines of η . Figure taken from Ref. [38].

The HCAL is a sampling calorimeter, with steel and brass used as the absorbing material, and plastic scintillator or quartz fibres as the active material. Hadrons shower via nuclear interactions with the detector material as they travel through the HCAL, and the light produced by this in the active material is read out via fibres. The energy of the hadrons is then inferred from the amount of light measured. The layout for one quarter of the HCAL is shown in Fig. 3.6.

The barrel (HB) extends from the ECAL to the inner coil of the CMS solenoid and covers the region $|\eta| < 1.3$. Plastic scintillator is used as the active material and it is segmented into 16 η sectors, and 72 ϕ sectors, leading to a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The thickness of the absorber corresponds to 5.82 interaction lengths (λ_I) at $\eta = 0$ and grows to 10.6 λ_I at $\eta = 1.3$. Including the ECAL provides about 1.1 λ_I of additional material, but this is still not enough to fully contain the hadronic showers. Therefore, an outer calorimeter (HO), with similar technology to the HB, is placed outside the solenoid and acts as a tail catcher, leading to a minimum depth of 11.8 λ_I in the barrel region.

The hadronic calorimeter endcaps extend the coverage to $|\eta| < 3$. Again, the active material is a plastic scintillator and the granularity is the same as the HB for $1.3 < |\eta| < 1.6$, and reduces to $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for $1.6 < |\eta| < 3$. Including the ECAL, the

total depth of the HE corresponds to about $10\lambda_I$.

The forward calorimeter (HF) extends the HCAL to its ultimate coverage of $|\eta| < 5.2$. This detector will absorb 88% of the energy produced by a proton-proton interaction and is therefore subjected to intense doses of radiation. To withstand this, quartz fibres were chosen as the active material where charged particles generate light via the Cherenkov effect. The fibres are embedded in a steel absorber and run parallel to the beam line. Since the ECAL does not cover this pseudorapidity region, the HF also designed to discriminate electromagnetic showers from hadronic ones. It does this by taking advantage of the fact that electromagnetic showers have a smaller typical depth than hadronic showers. Only half of the fibres run the full length of the HF, whereas the other half start 22 cm from the front, therefore allowing the depth of a shower to be inferred.

3.2.6 Muon System

Muons traverse the CMS detector with minimal interaction, and make up the majority of the visible particles that emerge through the calorimeter systems. This provides an opportunity to improve the identification, momentum resolution, and charge determination of muons by placing a final detector beyond the calorimeters. The CMS muon system is embedded in the return yoke of the solenoid magnet and covers the pseudorapidity $|\eta| < 2.4$. A schematic showing the layout of the muon system is given in Fig. 3.7.

The muon system is composed of layers of gaseous detectors which provide multiple measurements of a muon's trajectory. In the barrel, covering the $|\eta| < 1.2$ region, drift tube (DT) chambers are used. The chambers are made from $42 \times 13 \text{ mm}^2$ drift cells with wires arranged either parallel to the beam line or perpendicular depending on which layer the cells are in, thereby providing a measurement in the r - ϕ or r - z plane respectively [51].

In the endcap regions, cathode strip chambers (CSC) are installed and cover a range $0.9 < |\eta| < 2.4$. The CSCs are multi-wire proportional chambers with cathode strips and anode wires arranged perpendicular to each other which provide measurements in the r - ϕ plane and of η respectively. A third type of gaseous detector, resistive plate chambers (RPC), are used in both the barrel and endcaps in the $|\eta| < 1.9$ range. These detectors have a fast response time and are primarily used to provide timing information to the muon trigger.

Using this system, muons with $p_T > 30 \text{ GeV}$ can be identified with an efficiency greater than 95% with a misidentification rate of less than 1%. Furthermore, for muons with $p_T < 200 \text{ GeV}$, the momentum resolution is better than 3%, and for muons with $200 > p_T > 1000 \text{ GeV}$, where the addition of the muon system is especially beneficial, a resolution of 6% or better is found [50].

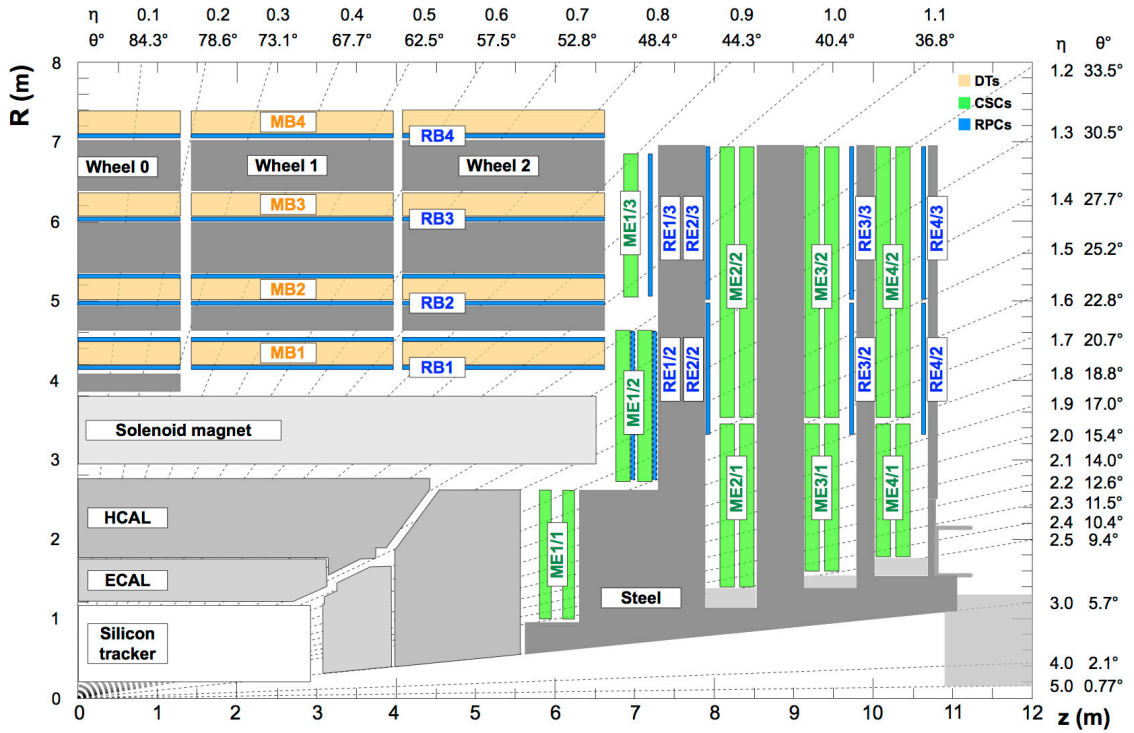


Figure 3.7: A schematic of one quarter of the CMS muon system in the r - z planes. The interaction point is the lower-left corner of the diagram and the dashed lines represent constant lines of η and θ . Figure taken from Ref. [50].

3.2.7 Trigger

The beam crossing interval at the LHC is 25 ns, corresponding to an event rate of 40 MHz. With event sizes of about 1 MB, this corresponds to a data rate of 38 TB/s [38]. This is an unfeasible data rate to store and process, and therefore a trigger system is used to reduce the rate to a manageable level. The CMS trigger system operates in two stages: a level-1 (L1) trigger which brings the rate down to 100 kHz, and a high-level trigger (HLT) which further reduces the rate to 1 kHz.

To provide a decision within the allowed latency of 3.2 μ s, the L1 trigger uses only coarse information about an event provided by the calorimeters and the muon system, and uses fast, custom-made electronics, primarily field-programmable gate arrays (FPGAs). A decision is made based upon the presence of high- p_T muons, electrons, tau leptons, photons, jets, and missing transverse energy. Information about the isolation of these objects can also be used. If an event passes the L1 trigger, the full event data is read out and passed to the HLT, which is software-based and runs on a CPU farm. The HLT is able to perform more complex reconstruction, and apply more sophisticated algorithms to make a trigger decision. There are hundreds of algorithms in the HLT (also referred to as HLT *paths*) which are tuned to provide high signal efficiencies for the wide range of physics processes that the CMS experiment is interested in. The di-Higgs search described in Chapter 5 uses diphoton HLT paths which are described in more detail in Section 5.4.

3.3 Object Reconstruction

The outputs from the CMS detector, after some processing of the raw signals, are essentially, hits in the tracking systems and energy deposits in the calorimeters. To perform a physics analysis, we want to identify what physics objects were produced in the collisions and what their properties were. Physics objects include electrons, muons, taus, photons and jets. An example of an important property is the four-momentum of the object. The process of interpreting the outputs of the detector as physics objects is known as *reconstruction*.

In the CMS experiment, object reconstruction begins with the particle-flow (PF) algorithm and its creation of the basic *elements*: tracks and energy clusters [52]. These elements are *linked* within and across detector systems to form *blocks* of elements, which are then used to create object candidates. By combining information across systems, the particle-flow algorithm is able to create object candidates with great resolution and efficiency and this concept was key to the design of the algorithm. Further details on tracking, clustering and the link algorithm are provided in Sections 3.3.1 to 3.3.3.

PF object candidates are input to specialized algorithms that complete the object reconstruction. In most cases, these algorithms are used to apply final identification criteria and calibration but in some cases, they also create more complex objects such as jets. These algorithms are described in Sections 3.3.4 to 3.3.7 where details are restricted to those which are most relevant to the di-Higgs search in Chapter 5.

3.3.1 Tracking

The algorithm [53–56] for reconstructing tracks in the inner tracker is based on the combinatorial Kalman filter (CKF) [57–60]. Initially, tracking seeds are generated from a few hits consistent with a charged-particle trajectory. The seeds provide a coarse estimate of the track trajectory and from that, the CKF propagates through the tracker, looking for compatible hits in each tracker layer and updating the trajectory estimate as it does so. If multiple compatible hits are found in a layer, the CKF will create multiple track candidates and propagate them all (hence combinatorial). After exhausting all tracker layers, the tracks are refitted with greater precision and then selected based on quality criteria such as the number of hits and the χ^2 of the fit.

Given the combinatorial nature of the algorithm, CKF can be computationally expensive. This could be handled by, for example, only propagating tracks that have a high p_T or that coincide with the interaction point, but this would lead to a loss of efficiency for low p_T tracks or tracks that originate from displaced vertices. To mitigate this, an iterative approach is used where multiple runs of the CKF are performed, with varying types

of seeds and selection criteria which target different types of tracks, for example, high p_T in one run and displaced tracks in the next. After each stage, the hits associated with the selected tracks are removed from the list of hits available to the CKF. This reduces the combinatorial complexity of the problem and allows the CKF to focus on the remaining hits. Ultimately, this iterative approach leads to higher reconstruction efficiency whilst keeping the misconstruction rate low [52].

In the muon systems, a separate Kalman-filter based algorithm is used to reconstruct *standalone-muon* tracks [50]. These muon tracks can then be matched with inner tracks to create *global-muon* tracks. Similarly, inner tracks are propagated to the muon system to create *tracker-muon* tracks. Due to the high reconstruction efficiency in the inner tracker and muon systems, about 99% of muons within the geometrical acceptance of the muon systems are reconstructed as global muons or tracker muons, and often as both [52].

Electrons emit a sizeable fraction of their energy as bremsstrahlung photons before reaching the ECAL. If the momentum of an electron changes enough along its trajectory, the iterative CKF algorithm can fail to reconstruct the track. Therefore, for electrons, a Gaussian Sum Filter (GSF) algorithm is instead used which allows for more sudden and significant energy losses along the particle's path [61].

3.3.2 Calorimeter Clustering

Energy deposits in the calorimeters are grouped into *clusters* where each cluster is hypothesized to originate from a single particle incident on the ECAL [52]. The clustering is performed independently in the ECAL barrel, ECAL endcaps, HCAL barrel, HCAL endcaps, Preshower and the HF. In the HF, the electromagnetic and hadronic components of each cell directly give rise to separate clusters. In the rest of the calorimeters, a more complex algorithm is used which is described below.

First, cluster seeds are identified as cells with energy larger than a given threshold and larger than any neighbouring cells, where neighbouring cells are the four cells that share a side with the seed cell or the eight cells that share a side or corner, depending on the calorimeter system. The seed cells are then grown into *topological clusters* by adding cells that share a corner and have an energy of at least twice the noise level of the cell.

Particles that are close to each other in η - ϕ space can have overlapping energy deposits, which in turn, can lead to a topological cluster having multiple seeds. In these cases, the energies from the cells are split and shared amongst the seeds. This is done using a Gaussian-mixture model that postulates that the energy deposits in the M individual cells come from N different Gaussian energy deposits, one for each seed. After fitting the model, the energy and positions of the N Gaussian deposits are taken as cluster parameters [52].

Electrons and photons have a significant probability of undergoing bremsstrahlung radiation or photon conversion prior to reaching the ECAL leading to several distinct clusters in the ECAL. To measure the energy of the original electron or photon, the clusters are grouped into a *supercluster* which is then used to create the electron or photon candidate. A supercluster (SC) is first formed by including all clusters in an E_T -dependent zone in $\phi - \eta$ space centred around the seed cluster [62]. Then, a conversion-finding algorithm [38] is used to identify tracks and associated ECAL deposits consistent with a photon conversion. Furthermore, at each tracker layer, tangents from an electron's trajectory are extrapolated to look for clusters from bremsstrahlung photons. All of this information is then fed into the PF link algorithm which builds electron and photon candidates.

3.3.3 Link Algorithm

After the creation of the PF elements (tracks and clusters), a linking algorithm is used to form particle candidates by creating associations between elements that are likely to have originated from the same particle [52]. Firstly, inner tracks are linked to calorimeter clusters when the track's extrapolated trajectory overlaps with the cluster. Similarly, clusters in different calorimeters are linked if the cluster's centre from the calorimeter with greater granularity is within the cluster of the calorimeter with coarser granularity. If there are multiple possible links, the link with the smallest distance in ϕ - η space, or x - y space for ECAL-Preshower links, is chosen. Finally, links between tracks can be created if the tracks share a common secondary vertex.

Particle candidates are now formed sequentially, starting with muons, then electrons and photons, then charged and neutral hadrons [52]. Identification of muons and photons are also revisited at later stages to recover candidates that were missed by initially stringent criteria. At each stage, the PF elements related to the particle candidates are removed from the list of elements available to the link algorithm. This ensures that the algorithm does not link the same element to multiple candidates and that the same element is not used to create multiple candidates. Finally, an event post-processing step is performed to correct for artificially-large p_T^{miss} [52].

3.3.4 Muons

Muon Identification

In CMS, there are five main identification (ID) types for muons: loose, medium, tight, soft, and high momentum [50]. Each type is defined by cuts on a number of variables including the track fit χ^2 , number of hits per track, the compatibility of the muon track

with the PV, the compatibility of the inner track with the standalone muon track (for global muons), and a related quantity called muon segment compatibility [63] which takes a value between 0 and 1 where 1 represents high compatibility. Additionally, a kink-finding algorithm is used to discriminate against charged hadrons which are more likely to interact with the tracker material than muons are. This algorithm, at several places along the track, splits the inner track (if present) in two and tests if the two halves are compatible with a single track.

The loose muon ID is designed to have a low efficiency for hadrons, and to select muons originating from the interaction vertex (prompt muons) or from light and heavy flavour decays. It is defined as a muon selected by the PF algorithm which is a tracker or global muon. The medium muon ID, which is used in the di-Higgs search in Chapter 5, is designed for prompt muons and muons from heavy flavour decay. A medium muon is a loose muon which passes the following additional requirements:

- has an inner track with hits from more than 80% of the tracker layers it traverses;
- if the muon is only a tracker muon, the muon segment compatibility must be greater than 0.451;
- if the muon is a global muon, the muon segment compatibility must be greater than 0.303 and the global track fit χ^2/dof must be less than 3;
- the position match between the inner track and the standalone muon track must have $\chi^2 < 12$;
- and the maximum χ^2 computed by the kink-finding algorithm must be less than 20.

These requirements were tuned to provide an identification efficiency of 99.5% for muons in simulated $W \rightarrow \mu\nu$ and $Z \rightarrow \mu\mu$ events [50].

The tight muon ID is designed to suppress muons from decay in flight and from hadronic punch-through, the soft muon ID is designed for low- p_T muons from b-hadron decays, and the high momentum muon ID is designed for muons with p_T greater than 200 GeV. The definition for these IDs can be found in Ref. [50].

Further requirements can also be made on the particle-flow isolation, I_{PF} , of the muon, which is defined as the pileup-corrected sum of the p_T of the charged hadrons, photons, and neutral hadrons in a cone of $\Delta R < 0.3$ around the muon. The pileup correction subtracts the expected contribution originating from pileup interactions. Tight and loose working points are defined to achieve muon ID efficiencies of 95% and 98% respectively in simulated $Z \rightarrow \mu\mu$ events [50].

Reconstruction, identification, and isolation efficiencies are measured in data and compared to simulation. The ratio of the efficiencies, referred to as *scale factors*, are then used to correct the simulation to match data by multiplying the simulation events weights

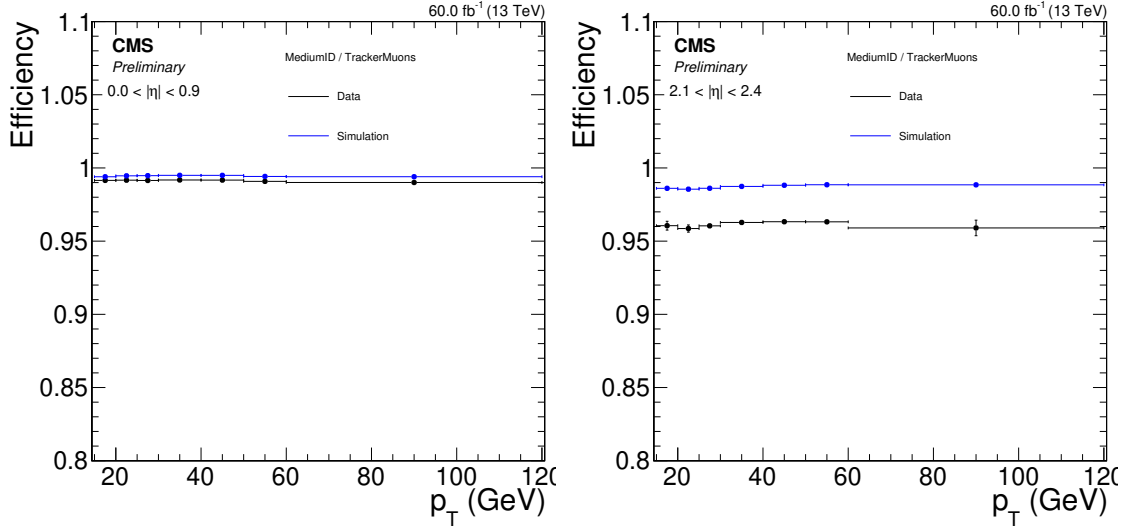


Figure 3.8: Medium muon ID efficiency in 2018 as measured in data and simulation as a function of muon p_T for $|\eta| < 0.9$ (left) and $2.1 < |\eta| < 2.4$ (right).

by these factors. These scale factors are derived separately for each year of data-taking to account for varying detector conditions, and in bins of muon p_T and η to account for the dependence of the muon kinematics and detector geometry on the simulation mismodelling. The full details of the derivation of these scale factors can be found in Refs. [50, 64–68].

As an example, the efficiency of the medium muon ID as measured in data and simulation for 2018 is shown in Fig. 3.8 for two η regions. For $|\eta| < 0.9$, the efficiency in simulation is about 99.5% as expected, and in data is slightly lower, at about 99%. For $2.1 < |\eta| < 2.4$, where fewer inner tracker hits are expected due to the tracker geometry (see Fig. 3.4) the efficiency is lower at about 99% and 96% for simulation and data respectively. In both cases, and especially for $2.1 < |\eta| < 2.4$, there is disagreement between data and simulation, highlighting the need for scale factors.

Muon Momentum

Muon momentum is determined by the Tune-P algorithm [69] which selects the p_T measurement from different refits of the muon track based on the goodness-of-fit information and estimated p_T resolution. The types of refits include tracker-only, tracker and first muon detector plane, global without muon chambers with high occupancies, and a *dynamic-truncation* fit which propagates the track to the muon stations but stops as soon as no compatible hit is found in two consecutive stations.

The momentum scale calibration and resolution estimation are derived from $Z/\gamma^* \rightarrow$

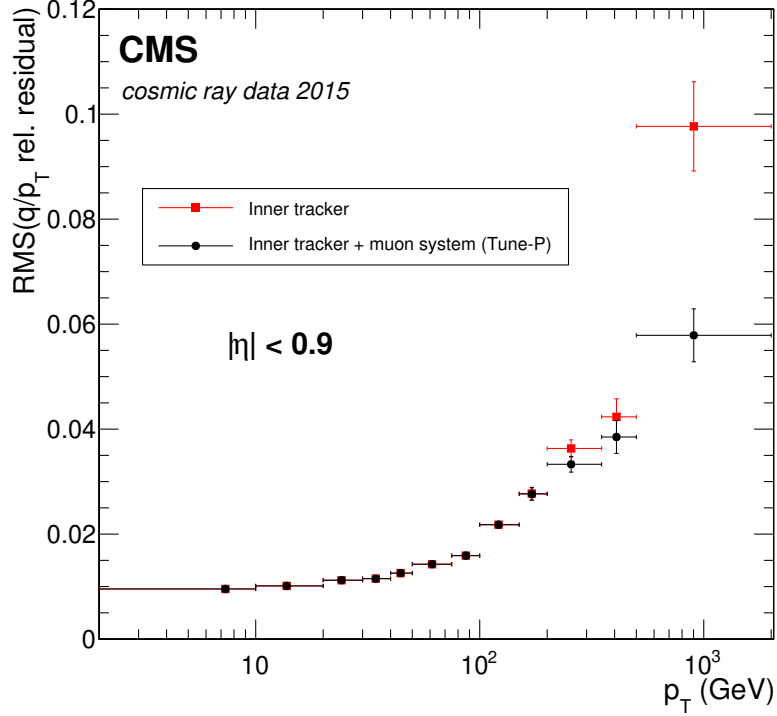


Figure 3.9: Relative muon momentum resolution as a function of p_T for $|\eta| < 0.9$ measured using cosmic ray data collected in 2015. The vertical error bars indicate statistical uncertainties. Figure taken from Ref. [50].

$\mu\mu$, $J/\psi \rightarrow \mu\mu$ and $\Upsilon(1S) \rightarrow \mu\mu$ events in data [70]. The momentum scale corrections are about 0.2% and 0.3% in the barrel and endcap respectively, and the resolution is about 1% in the barrel and 3% in the endcap for muons with p_T up to 100 GeV. The momentum resolution of intermediate and high p_T muons is additionally measured in cosmic ray events by comparing the momentum measurements in the upper and lower halves of the detector [63]. The momentum resolution, as measured with this method in 2015 data, is shown in Fig. 3.9.

3.3.5 Electrons and Photons

Electrons and photons share many of the same reconstruction and identification algorithms because, with the exception of an electron's track, the two particles leave similar signatures in the detector. Both particles are expected to be contained within the ECAL, and have almost identical showering properties. In this section, reconstruction for electron and photon (e/γ) objects will be described together, with differences highlighted where necessary.

Electron and Photon Identification

Background sources for prompt electrons include photon conversions, hadrons misidentified as electrons, and secondary electrons from semileptonic decays of b or c quarks. Background sources for prompt photons are primarily secondary photons from light neutral mesons (π^0 and η). Using variables related to the isolation and shower shape of the electrons and photons, multivariate (MVA) discriminators are trained to separate real electrons and photons from these background sources [62]. Additional tracker-related variables are also used for the identification of electrons.

The isolation variables are constructed by considering a cone of $\Delta R < 0.3$ around the e/γ candidate and summing the p_T of the charged hadrons (I_{ch}), photons (I_γ), and neutral hadrons (I_n) inside the cone. As is done for I_{PF} , a pileup correction is applied in these sums, and these variables are useful for discriminating against backgrounds originating from jets. Another important variable is the ratio of energy of associated HCAL clusters to the supercluster energy (H/E), which is expected to be lower for e/γ objects than for hadrons.

The shower shape variables exploit differences in the showering of prompt e/γ objects from photons from neutral hadron decays, which are expected to have a wider profile on average. One such variable is $\sigma_{\text{in}\eta}$, which is the standard deviation of the shower in η in terms of the absolute number of crystal cells [62]. Another important variable is R_9 , which is a measure of how localized the energy deposit is, and is defined as the ratio of the energy of the 3×3 cell grid surrounding the SC seed to the energy of the SC. For electrons, variables related to the compatibility of the track with the SC are also used which include the differences between the energy, η and ϕ of the electron, measured by the track and the SC.

These discriminating variables are combined into single discriminators using BDTs. This is done separately for electron and photons and the BDTs are trained on simulated $DY + \text{jets}$ and $\gamma + \text{jets}$ events for electrons and photons respectively. In both cases, the dependence on η and E_T is included, either by introducing these variables as inputs to the BDT, or by training several BDTs in different bins of η and E_T . The performance of these algorithms as measured by simulation for 2017 is shown in Fig. 3.10. Two working points are defined: WP90 and WP80, corresponding to signal efficiencies of about 90% and 80% respectively and physics analyses are left to choose which point to use based upon their level of background.

The electron ID efficiencies are measured in data using $Z \rightarrow ee$ events. The photon ID efficiencies are also measured in $Z \rightarrow ee$ events, where the electrons are reconstructed in the same way as photons by omitting the information from the electron track. This is motivated by the fact that electrons and photons shower similarly in the ECAL and is

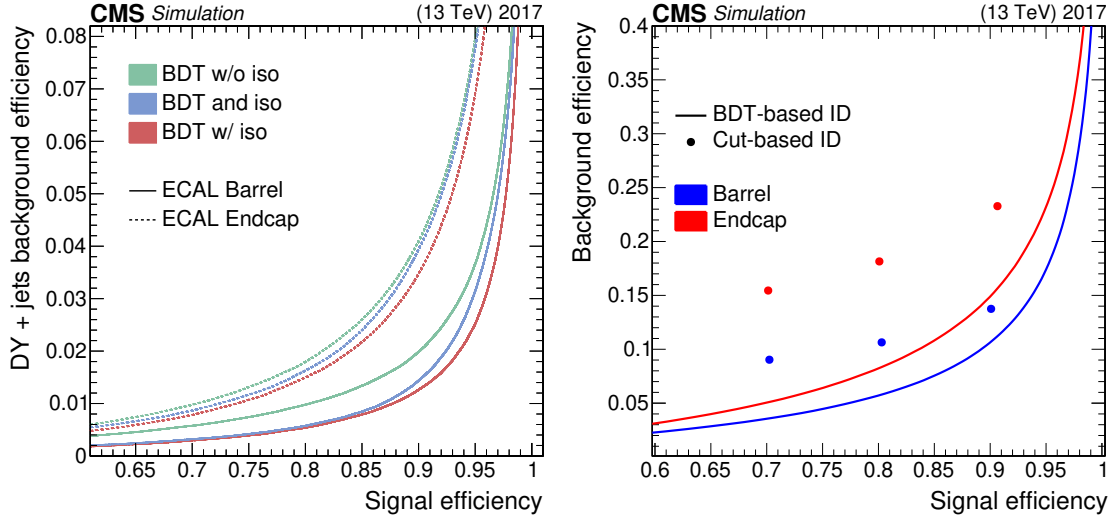


Figure 3.10: Performance of the electron (left) and photon (right) IDs in the ECAL barrel and endcaps evaluated for 2017 with simulated events. Different algorithms are shown for comparison's sake. For electrons, three algorithms are shown: a BDT trained without the isolation variables, the same BDT with isolation cuts applied separately, and a BDT with isolation variables included in training, where the last algorithm is the most performant. For photons, BDT-based and cut-based IDs are shown where the former is the most performant. Figures taken from Ref. [62].

validated using $Z \rightarrow \mu\mu\gamma$ events. The measured electron and photon ID efficiencies and a comparison to simulated efficiencies are shown for the WP90 working points in Fig. 3.11. Scale factors are derived separately for each year and in bins of η and E_T to correct the simulation to match data, and are up to 5% [62].

Finally, when searching for photon candidates, a conversion-safe electron veto (CSEV) can be used [71]. This veto requires that there are no charged-particles tracks pointing to the photon SC, where the tracks are required to have a hit in the first layer of the pixel detector, and are not matched to a reconstructed conversion vertex. This leads to photon and electron efficiencies of about 99% and 5% for the barrel, and 98% and 20% for the endcap respectively. An even more stringent veto, the *pixel veto*, rejects any photon where there exists at least two pixel hits that form a track pointing to the SC. This leads to photon and electron efficiencies of about 95% and 1% for the barrel, and 80% and 5% for the endcap respectively [71]. In a similar fashion to the ID efficiencies, the veto efficiencies are measured in data using $Z \rightarrow \mu\mu\gamma$ events for photons, and $Z \rightarrow ee$ events for electrons and scale factors are used to correct the simulation.

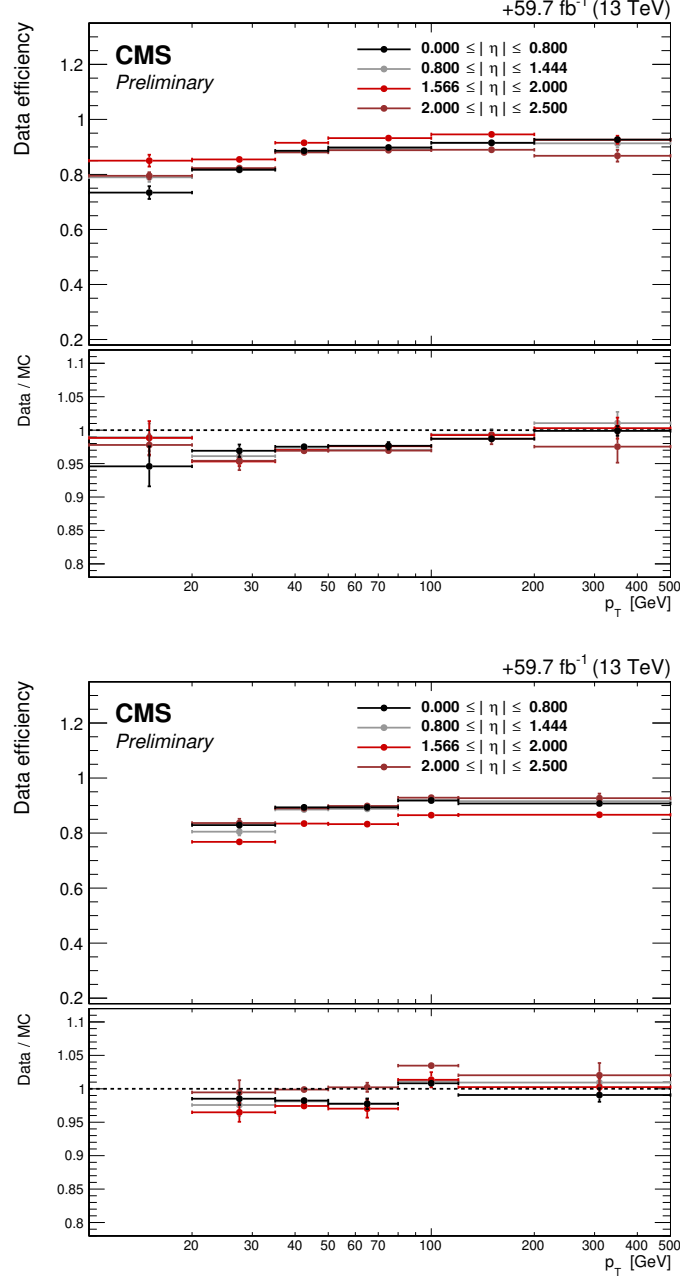


Figure 3.11: Signal efficiencies of the electron (top) and photon (bottom) BDT IDs at the WP90 working points. In the top halves of each figure, the efficiencies as measured in 2018 data with $Z \rightarrow ee$ events are shown. In the bottom halves, the ratio of the efficiencies in data to simulation (scale factors) are shown. The vertical error bars indicate combined statistical and systematic uncertainties. Figures taken from Ref. [62].

Energy Calibration

The energy deposited in an ECAL crystal, E_i , is given by the following equation [72]:

$$E_i = G \cdot LC_i(t) \cdot C_i \cdot A_i \quad (3.4)$$

where:

- A_i is the pulse amplitude in ADC (analogue-to-digital converter) counts,
- G are global factors that convert ADC counts to GeV,
- C_i are intercalibration coefficients that account for differences between individual crystal's light-yield and photodetector response,
- and $LC_i(t)$ are time-dependent corrections due to radiation-induced response changes to the crystals.

The derivation of these corrections is described in Refs. [72, 73]. The reconstructed energy of a supercluster, E_{SC} , is typically lower than the true energy, E_{true} , of the originating electron and photon due to imperfect containment of a shower, energy loss in the tracker material, and the application of thresholds when forming clusters. To correct this, simulated events with photons and electrons are studied and the distribution of $E_{\text{true}}/E_{\text{SC}}$ is parameterized by a Double Crystal Ball (DCB) function, which is an extension of the Crystal Ball function [74] with power law tails on both sides. A multivariate regressor is used to fit the DCB shape parameters as functions of the shower shape variables, H/E, and the supercluster's uncorrected energy and position in the detector [75]. Then, for a given supercluster, the energy is corrected by the mean of the Gaussian core of the DCB function, and the energy resolution is given by the width of the Gaussian core. Separate regressions are trained for electrons and photons to account for the small differences in how they shower.

For electrons with energies less than 200 GeV, the supercluster energy is combined with the momentum of the GSF track to improve the energy resolution. The combined energy measurement, E_{combined} is given by:

$$E_{\text{combined}} = \frac{E_{\text{ECAL}}/\sigma_E^2 + p_{\text{GSF}}/\sigma_p^2}{1/\sigma_E^2 + 1/\sigma_p^2} \quad (3.5)$$

where E_{ECAL} is the regression-corrected supercluster energy, p_{GSF} is the momentum of the GSF track, and σ_E and σ_p are the energy and momentum resolutions respectively. A final regression is applied to correct E_{combined} which uses the inputs to Eq. 3.5 as well as additional tracker quantities [62]. The distribution of $E_{\text{true}}/E_{\text{combined}}$ before and after

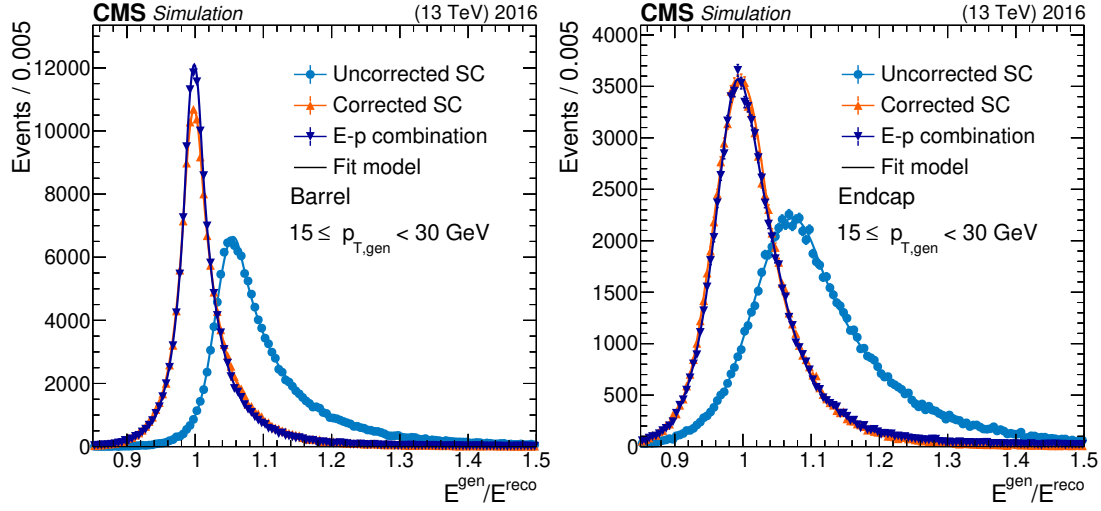


Figure 3.12: Ratio of the true to the reconstructed electron energy for $15 < p_T < 30$ GeV with and without regression corrections, and before and after the combination of the ECAL and tracker measurements, with a DCB function fit overlaid, in 2016 simulated events for barrel (left) and endcap (right) electrons. Vertical bars on the markers represent statistical uncertainties. Figure taken from Ref. [62].

the energy corrections, and after the combination with the track momentum is shown in Fig. 3.12 for electrons in 2016 simulated events. After corrections, the distribution is centred around 1, and the width is reduced, indicating that the energy is corrected, and the resolution is improved. The resolution improves further after the combination with the track momentum. The distributions before the combination with the track momentum are indicative of those for photons.

After these simulation-based corrections, there are differences between data and simulation in the energy scales and resolutions of e/γ objects. Further corrections are applied to correct the energy scale in data to match simulation, and since the resolution in simulation is better than that in data, smearings are applied to the simulation to match the data [62]. These corrections are derived with $Z \rightarrow ee$ events by comparing the distribution of the invariant mass of the Z boson in data and simulation. This is done in several stages.

In the first stage, energy scale corrections are derived in about 18 hour intervals and in bins of η corresponding to $0 < |\eta| < 1$, $1.00 < |\eta| < 1.44$, $1.57 < |\eta| < 2.00$ and $2.00 < |\eta| < 2.50$. The region, $1.44 < |\eta| < 1.57$ represents the transition between the barrel and endcap regions and is not used. These initial corrections account for long-term drifts in the energy scale and accounts for η -dependent radiation damage. In the second stage, the energy scale and resolution is corrected in 50 categories, corresponding to 5 bins of η and 10 bins of R_9 . The magnitude of these corrections is up to 1.5% and the systematic uncertainty associated with them is 0.05–0.1 (0.1–0.3)% in the EB (EE),

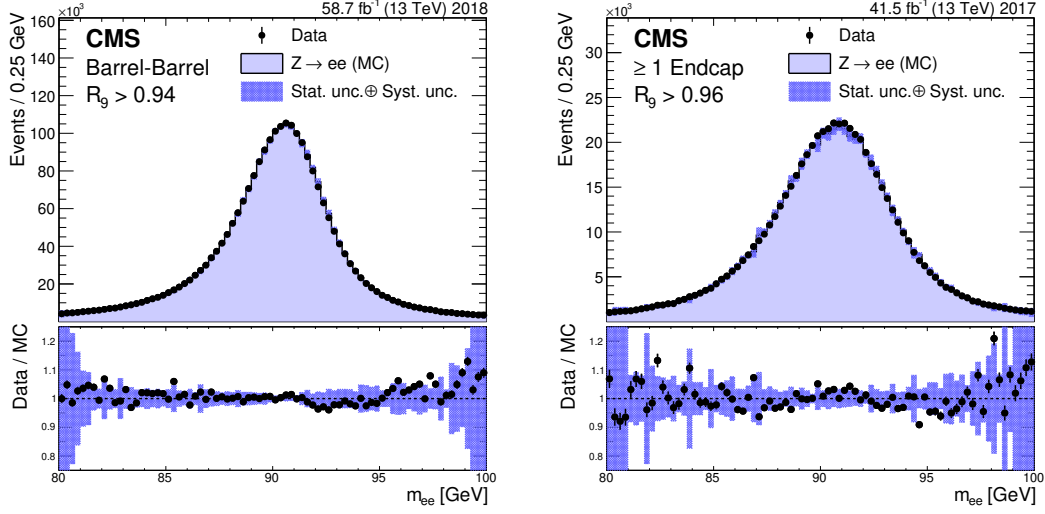


Figure 3.13: Invariant mass distribution of the Z boson in $Z \rightarrow ee$ events, after scale and smearing corrections. Results are shown for barrel (left) and endcap (right) electrons in high R_9 categories. The vertical bars on the markers represent the statistical uncertainties in data and the hatched regions show the combined statistical and systematic uncertainties in the simulation. The lower panels display the ratio of the data to the simulation with the bands representing the uncertainties in the simulation. Figure taken from Ref. [62].

depending on the R_9 bin [62].

The final agreement between data and simulation in $Z \rightarrow ee$ events is shown in Fig. 3.13 in two representative categories. The final energy resolution is found to be 2–5%, depending on the η and R_9 of the electron [62].

3.3.6 Jets

Quarks and gluons (partons) produced in a proton-proton collision shower and hadronize into collimated sprays of particles. To make a measurement of an outgoing parton, these particles are first clustered into *jets* and the parton’s properties are inferred from properties of the jet, such as its energy. In most CMS analyses, including the di-Higgs search in Chapter 5, jets are reconstructed using the anti- k_T algorithm [76, 77] with a distance parameter of 0.4.

Jet energy corrections are derived from simulation studies so that the energy of reconstructed jets match the energy of particle-level jets, where the difference between reconstructed jets and particle-level jets is whether the jet constituents are reconstructed or taken directly from the generator (truth-level). Further corrections are made using dijet, γ +jets, and Z+jet events to account for differences between data and simulation [78, 79]. The jet energy resolution for jets with $p_T = 30$ GeV is about 15–25% depending on the

level of pileup, and improves to 10–15% at $p_T = 100$ GeV and 5% at $p_T = 1$ TeV [79].

The identification of jets originating from b quarks (b jets) is performed using a deep neural network algorithm, DeepJet [80]. The algorithm uses properties of the particle-flow constituents of a jet as well information from associated secondary vertices. Three working points are defined: loose, medium, and tight, corresponding to a rate of misidentifying a light jet as a b jet of about 10%, 1% and 0.1% respectively.

3.3.7 Tau Leptons

Tau leptons decay into a variety of final states, which are summarized in Table 3.1. About one third of the time, tau leptons decay *leptonically* into an electron or muon, and two neutrinos. In such a decay, the electrons and muons are reconstructed as described in Sections 3.3.4 and 3.3.5, and properties of the neutrinos are inferred from the missing transverse momentum. The other two thirds of the time, tau leptons decay *hadronically* into hadrons and a neutrino. These types of tau leptons, denoted τ_h , are reconstructed with the hadrons-plus-strips (HPS) algorithm.

Hadrons-Plus-Strips Algorithm

The HPS algorithm [81, 82] begins by searching the constituents of reconstructed jets for charged hadrons and neutral mesons. A π^0 meson promptly decays into two photons, which are then likely to convert into e^+e^- pairs as they traverse the tracker material. The energy of the π^0 meson is therefore spread over a $\Delta\eta \times \Delta\phi$ region which is referred to as a *strip*. The basic ingredients of the HPS algorithm are charged hadrons and strips, hence the name.

A strip is reconstructed by clustering electrons and photons inside a jet, and begins by seeding the strip with the highest p_T electron or photon in the jet. Then, a $\Delta\eta \times \Delta\phi$ area centred on the seed is defined and the next-highest p_T electron or photon inside the area is added to the strip. The strip position, which was originally defined by the seed, is then recomputed using a p_T -weighted average of all electrons and photons in the strip and the process is repeated until no more electrons or photons can be added.

The size of the $\Delta\eta \times \Delta\phi$ area is dynamic and is dependent on the p_T of the strip and of the candidate electron or photon to account for several effects [82]. Firstly, the likelihood of a photon converting into an e^+e^- pair is higher at lower p_T , so the strip size is expected to be larger for lower p_T strips. Furthermore, the decay products for a τ_h with higher p_T tend to be boosted in the direction of the τ_h , leading to a smaller strip. The functions used to define the strip size are derived from simulation of τ_h decays and are designed such that 95% of all electrons and photons arising from a τ_h decay are included in the strip.

Decay mode	Resonance	\mathcal{B} (%)
Leptonic decays		35.2
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$		17.8
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
Hadronic decays		64.8
$\tau^- \rightarrow h^- \nu_\tau$		11.5
$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	$\rho(770)$	25.9
$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	9.5
$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other		3.3

Table 3.1: Decays of τ leptons and their branching fractions (\mathcal{B}). Where appropriate, the known intermediate resonances for a decay are indicated. Charged hadrons are denoted by the symbol h^\pm and although only τ^- decays are shown, the values are also valid for the charge-conjugated processes.

Based on the set of charged hadrons and strips in a jet, the HPS algorithm identifies the decay mode from Table 3.1. To reduce misidentification with jets, the mass of the sum of the hadron candidates is required to be compatible with a $\rho(770)$ or $a_1(1260)$ resonance, depending on the decay mode. Additionally, τ_h candidates are required to have a charge of ± 1 , and to have no charged particles or strips outside a signal cone, defined by $R_{\text{sig}} = (3.0 \text{ GeV})/p_T$, where the p_T is that of the hadronic system, and the cone size is limited to 0.05–0.10. Finally, a single τ_h candidate is allowed per jet by selecting the candidate with the highest p_T . The τ_h reconstruction efficiency is measured in data with $Z \rightarrow \tau\tau$ events and is found to be about 75% for τ_h with $p_T > 30 \text{ GeV}$ [83].

DeepTau ID

The DeepTau ID algorithm [83] is a deep neural network that is used to identify τ_h candidates and reject jets, electrons and muons misidentified as τ_h candidates. The algorithm uses a combination of higher-level inputs, which are summary variables related to the τ_h candidate, and lower-level inputs, which are variables related to reconstructed particles in the vicinity of the τ_h candidate. Higher-level inputs include the τ_h four momentum, number of charged and neutral particles used to reconstruct the τ_h candidate, and isolation variables. To construct the lower-level inputs, particles within a $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$ area centred around the τ_h candidate are considered, and properties such as the particle type, p_T , η , ϕ , charge, and compatibilities with primary and secondary vertices are included.

The network is a multiclassifier with four outputs: jet, μ , e , and τ_h , and it is trained

	VVTight	VTight	Tight	Medium	Loose	VLoose	VVLoose	VVVLoose
D_e	60%	70%	80%	90%	95%	98%	99%	99.5%
D_μ	—	—	99.5%	99.8%	99.9%	99.95%	—	—
D_{jet}	40%	50%	60%	70%	80%	90%	95%	98%

Table 3.2: Target τ_h identification efficiencies for different working points defined for the three different discriminators. The target efficiencies are evaluated with simulated $H \rightarrow \tau\tau$ events for τ_h with $p_T \in [30, 70]$ GeV.

with a modified cross-entropy loss function on simulated events from the following processes: $Z + \text{jets}$, $W + \text{jets}$, $t\bar{t}$, QCD multijet production, and $Z' \rightarrow \tau\tau$, $Z' \rightarrow ee$ and $Z' \rightarrow \mu\mu$ where $1 < m_{Z'} < 5$ TeV. The final discriminators against jets, muons, and electrons are given by:

$$D_\alpha = \frac{y_\tau}{y_\tau + y_\alpha} \quad (3.6)$$

where $\alpha \in \{\text{jet}, \mu, e\}$ and y_i are the four outputs of the network. Working points are defined for each discriminator based upon the expected τ_h identification efficiencies as measured in $H \rightarrow \tau\tau$ events for τ_h candidates with $30 < p_T < 70$ GeV. These working points are summarized in Table 3.2.

The combined efficiency of τ_h reconstruction and identification as measured in different bins of τ_h p_T with simulated $Z \rightarrow \tau\tau$ decays is shown in Fig. 3.14. After applying the Loose, Medium or Tight D_{jet} working points, the efficiency increases with τ_h p_T and for $p_T > 30$ GeV, the efficiency is above 40%, 50%, and 60% for the respective working points. The corresponding misidentification rates for jets, electrons, and muons as a function of the τ_h identification efficiency are shown in Fig. 3.15. At a τ_h identification efficiency of 80%, the misidentification rate for electrons is about 0.1%, for muons is less than 0.03%, and for jets is between 1 and 4%, depending on τ_h p_T .

The τ_h reconstruction and identification efficiencies are also measured in data with $Z \rightarrow \mu\tau_h$ and $W^* \rightarrow \nu\tau_h$ events [83] and scale factors are derived to correct the simulation to match data. These scale factors are derived in bins of p_T and their values and uncertainties are shown in Fig. 3.14. The corrections are at most 10% and have an uncertainty of 2–5% [83].

τ_h Energy Scale

The τ_h energy scale is measured in data with $\mu\tau_h$ events by fitting the distributions of the $\mu\tau_h$ invariant mass, m_{vis} , and the reconstructed τ_h mass, m_{τ_h} . Simulated templates of the m_{vis} and m_{τ_h} distributions are created for τ_h energy shifts between $\pm 3\%$ in steps of 0.2% and a maximum likelihood fit is performed for each shift. This is performed in bins of the τ_h decay mode and separately for m_{vis} and m_{τ_h} . The corresponding energy scale

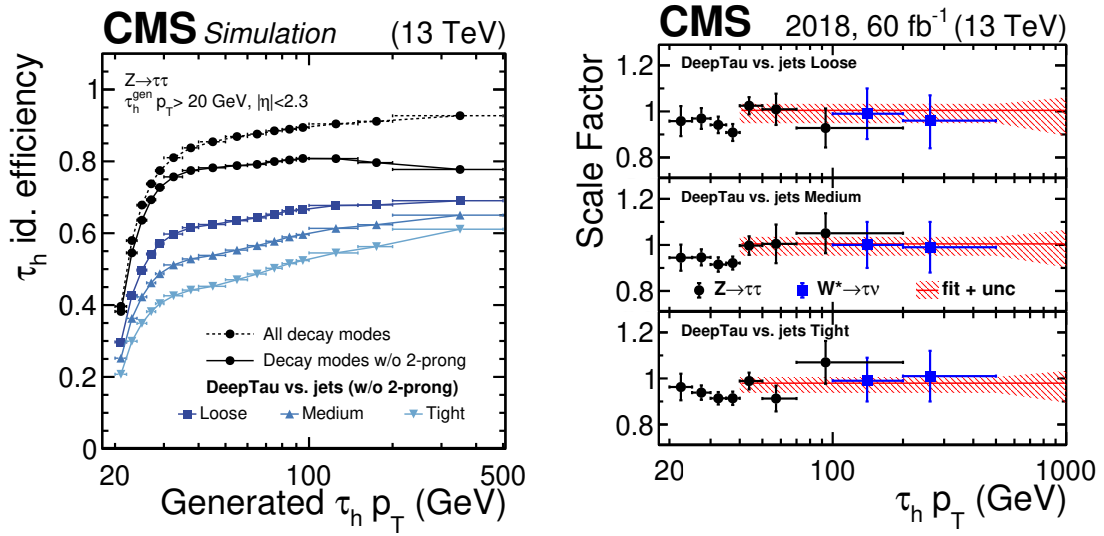


Figure 3.14: Reconstruction and identification efficiency of τ_h candidates as measured in simulated $Z \rightarrow \tau\tau$ events (left) and corresponding scale factors measured with data collected in 2018 (right). In both plots, the results are shown as functions of the true (generated) τ_h p_T , and the efficiencies after applying the Loose, Medium and Tight DeepTau ID working points are calculated after excluding 2-prong decays, which are those containing missing charged hadrons. In the right plot, the vertical bars represent combined statistical and systematic uncertainties in the scale factors, and the red line and hatched region represents a constant fit to the scale factors for $p_T > 40$ GeV and the associated uncertainty. Figures taken from Ref. [83].

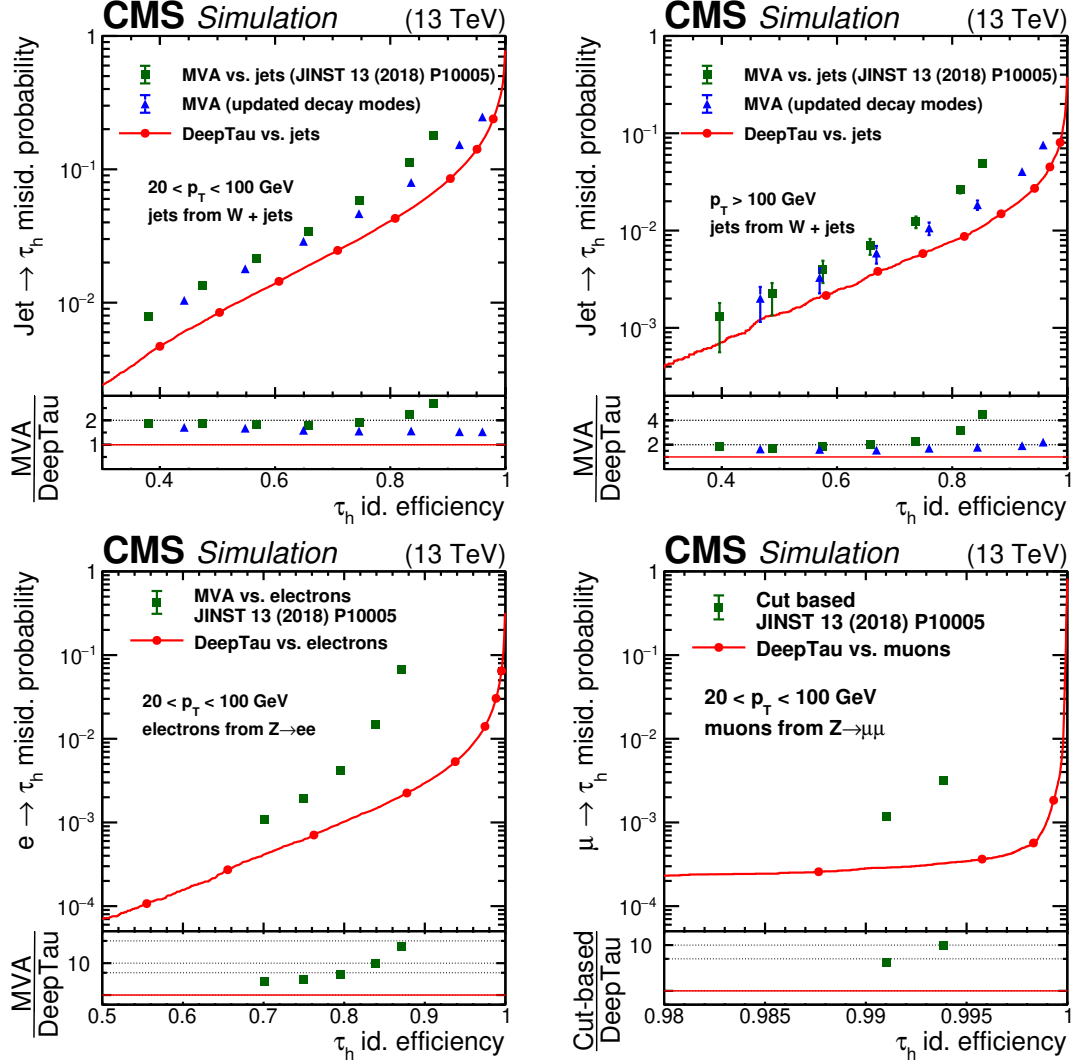


Figure 3.15: Misidentification rates for jets (top), electrons (bottom left), and muons (bottom right) as a function of the τ_h identification efficiency for the D_{jet} , D_e , and D_μ discriminators respectively. Also shown is the performance of previous τ_h identification algorithms and the ratio of misidentification rates between one of the alternate algorithms and DeepTau is shown in the bottom half of each plot. In all cases, DeepTau outperforms the alternate algorithm. For jets, efficiencies are shown for jets with $20 < p_T < 100$ GeV (top left) and $p_T > 100$ GeV (top right) and they indicate that higher p_T jets are less likely to be misidentified. For electrons and muons, the efficiencies are only shown for $20 < p_T < 100$ GeV since the misidentification rate is approximately constant with p_T for these objects. Vertical bars represent statistical uncertainties and the full red circles represent the working points. Figures taken from Ref. [83].

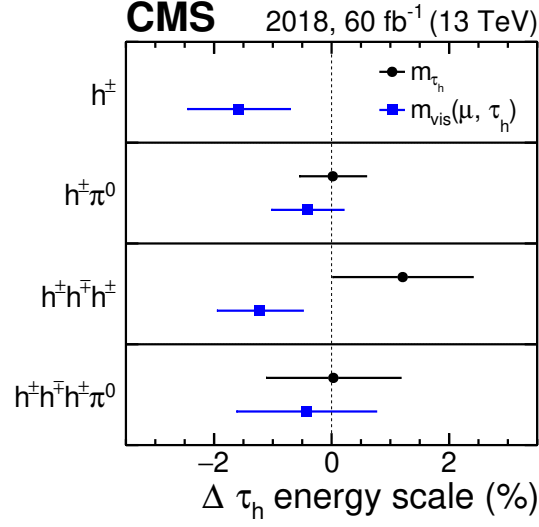


Figure 3.16: Relative difference for the τ_h energy obtained between data and simulated events for 2018 for the four main reconstructed τ_h decay modes. The results are obtained from fits to the distribution of either the reconstructed m_{vis} (blue lines) or m_{τ_h} (black lines). The horizontal bars represent the combined statistical and systematic uncertainties in the measurements. Figure taken from Ref. [83].

corrections are shown in Fig. 3.16 and are found to be up to 2% with uncertainties of 0.6–0.8% depending on decay mode. The results from measuring m_{vis} and m_{τ_h} are found to be consistent with each other and the final energy scale corrections are derived from a combination of the two approaches. After applying the corrections, good agreement is found in the m_{vis} and m_{τ_h} distributions as shown in Fig. 3.17.

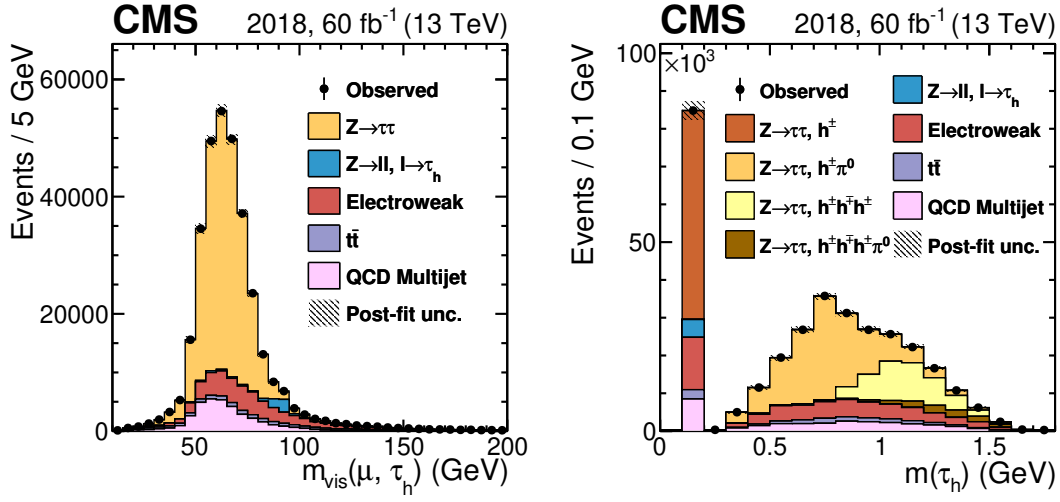


Figure 3.17: Distribution of the reconstructed visible invariant mass of the $\mu\tau_h$ system, m_{vis} (left) and of the visible invariant τ_h mass (right) in 2018 data (black dots) and simulated events (filled histograms). Vertical bars correspond to statistical uncertainties. The τ_h is required to pass the Tight D_{jet} working point, and have $p_T > 20$ GeV and $|\eta| < 2.3$. The distributions incorporate all measured scale factors and energy corrections and are scaled to the outcome of a maximum likelihood to the observed data with the $Z \rightarrow \tau\tau$ contribution freely floating. In the $m(\tau_h)$ distribution, the $Z \rightarrow \tau\tau$ contributions are shown separately for the different τ_h decay modes. Figures taken from Ref. [83].

3.3.8 Summary

Object reconstruction at the CMS experiment is a complex process that involves dedicated algorithms for each type of object, often combining information from different subdetectors to maximize performance. All the reconstruction algorithms described in this Section have some relevance to the di-Higgs search presented in Chapter 5 but the most relevant are those for the photons and τ_h objects. The excellent ECAL energy resolution leads to a resolution of about 1% on the diphoton invariant mass, $m_{\gamma\gamma}$, for values of $m_{\gamma\gamma}$ around 125 GeV. In $H \rightarrow \gamma\gamma$ analyses, this leads to a narrow signal peak in the $m_{\gamma\gamma}$ distribution which is used to identify the signal above a large background. On the τ_h side, signal is separated from background using the Loose D_{jet} WP which identifies τ_h candidates with an 80% efficiency whilst maintaining a jet misidentification rate between 1% and 4% depending on the τ_h p_T .

Chapter 4

Statistical Procedures

In Chapter 5, a search is performed for the $pp \rightarrow X \rightarrow H(H/Y) \rightarrow \gamma\gamma\tau\tau$ (di-Higgs) process, and in Chapter 6, measurements of the STXS (Section 2.2) are interpreted in the SMEFT (Section 2.3.3). Statistical tests are performed to determine the compatibility of the data with both the SM and, in the case of Chapter 5, with the presence of new resonances leading to a $\gamma\gamma\tau\tau$ final state, or in the case of Chapter 6, with the presence of new physics originating from one or more \mathcal{L}_6 operators. Practically, these tests are performed using the COMBINE framework [84], which is a publicly-available software package designed by the CMS collaboration. A full description of this package is found in Ref. [84] and this chapter summarizes the details that are relevant to Chapters 5 and 6.

The construction of likelihood functions is described in Section 4.1. The determination of p-values (and significances) that provide numerical measures of the compatibility of the data is described in Section 4.2.1. The determination of upper limits on the cross section of the di-Higgs process is described in Section 4.2.2, and the estimation of confidence intervals for the Wilson coefficients of the SMEFT is described in Section 4.2.3. Lastly, a short description of χ^2 fits is also given in Section 4.2.4.

4.1 Likelihood Function

To make statistical inferences from some data, first a statistical model, $p(\text{data}; \vec{\Phi})$, is constructed that describes the probability of observing the data given the model parameters, $\vec{\Phi}$. Model parameters are divided into two categories: parameters of interest (POIs), $\vec{\mu}$, and nuisance parameters, $\vec{\nu}$. In this thesis, the primary POIs are the cross section of the di-Higgs process, or the Wilson coefficients of the SMEFT, and the nuisance parameters model systematic uncertainties, such as the uncertainty in the measurement of the LHC luminosity at the CMS experiment. The data is summarized by a set of primary observables, \vec{x} , and a set of auxiliary observables, \vec{y} . The primary observables are those that are intended to constrain $\vec{\mu}$, and the auxiliary observables are those related to the systematic uncertainties where each nuisance parameter is associated with a particular auxiliary observable.

The statistical model is given by:

$$p(\vec{x}, \vec{y}; \vec{\Phi}) = p(\vec{x}; \vec{\mu}, \vec{\nu}) \prod_k p_k(y_k; \nu_k), \quad (4.1)$$

where $p(\vec{x}; \vec{\mu}, \vec{\nu})$ and $p_k(y_k; \nu_k)$ are the probability distributions for the primary and auxiliary observables respectively. Then, for a set of independently distributed observables, $\{\vec{x}_d\}$, a likelihood function is given by:

$$\mathcal{L}(\vec{\Phi}) = \prod_d p(\vec{x}_d; \vec{\mu}, \vec{\nu}) \prod_k p_k(y_k; \nu_k), \quad (4.2)$$

where d labels the entries in the dataset. An element of \vec{x} could be the event count in an analysis category, or the value of a continuous variable. These elements are treated as statistically independent so that $p(\vec{x}; \vec{\mu}, \vec{\nu})$ can be expressed as:

$$p(\vec{x}; \vec{\mu}, \vec{\nu}) = \prod_i p_i(x_i; \vec{\mu}, \vec{\nu}) \quad (4.3)$$

where p_i is the probability density function (pdf) for the observable x_i . Constructing the likelihood then becomes a task of determining $p_k(y_k; \nu_k)$ and $p_i(x_i; \vec{\mu}, \vec{\nu})$, where the form of $p_i(x_i; \vec{\mu}, \vec{\nu})$ depends on the type of analysis being performed.

4.1.1 Counting Analyses

The simplest type of analysis is a counting analysis in a single category, where the primary observable is the observed number of events, n . In this case, $p(x; \vec{\mu}, \vec{\nu})$ is a Poisson

distribution, and the likelihood is given by:

$$\mathcal{L}(\vec{\Phi}) = \lambda(\vec{\mu}, \vec{\nu})^n \frac{e^{-\lambda(\vec{\mu}, \vec{\nu})}}{n!} \prod_k p_k(y_k; \nu_k). \quad (4.4)$$

In an analysis without systematic uncertainties where there is an expected number of signal and background events, s and b , and where the only parameter of interest is a signal strength, μ , that scales the expected number of signal events, the natural log of the likelihood is given by:

$$\ln \mathcal{L}(\mu) = n \ln(\mu s + b) - (\mu s + b) - \ln(n!). \quad (4.5)$$

Expanding this to more than one category with observed counts, n_i , and expected signal and background counts, s_i and b_i , this likelihood becomes:

$$\ln \mathcal{L}(\mu) = \sum_i n_i \ln(\mu s_i + b_i) - (\mu s_i + b_i) - \ln(n_i!) \quad (4.6)$$

where μ scales the expected signal counts in each category equally. Equation 4.6 is the likelihood used for a simplified upper limit calculation that is described in Section 5.5.4 and used in the development of the di-Higgs search in Chapter 5.

4.1.2 Shape Analyses

More sensitive results can be obtained by using choosing primary observables that go beyond the number of events in a category. For this reason, the final results extraction in the di-Higgs search uses the invariant mass of the selected diphoton candidate, $m_{\gamma\gamma}$, of every event in one or more categories as the primary observables. The shape of an $m_{\gamma\gamma}$ distribution is described by an analytical function, $f_{pc}(m_{\gamma\gamma}; \vec{\mu}, \vec{\nu})$, which is the probability density function for $m_{\gamma\gamma}$ for a process, p , in a category, c . The probability density function for the primary observables, $p(\vec{x}; \vec{\mu}, \vec{\nu})$, takes into account the relative contributions of each process in a category, and sums over all categories, and is given by:

$$p(\vec{x}; \vec{\mu}, \vec{\nu}) = \sum_{p,c} \frac{\lambda_{pc}(\vec{\mu}, \vec{\nu}) f_{pc}(m_{\gamma\gamma}; \vec{\mu}, \vec{\nu})}{\sum_{p,c} \lambda_{pc}(\vec{\mu}, \vec{\nu})} \quad (4.7)$$

where $\lambda_{pc}(\vec{\mu}, \vec{\nu})$ is the expected number of events for a process, p , in category, c . Finally, for a dataset of n events, a Poisson term, $\mathcal{P}(n; \sum_p \lambda_p(\vec{\mu}, \vec{\nu}))$, is also added to the likelihood of Eq. 4.2.

4.1.3 Systematic Uncertainties

In the COMBINE framework, the probability densities for the auxiliary observables, $p_k(y_k; \nu_k)$, are given by normal, Poisson, and uniform distributions, depending on the auxiliary observable [84]. In the di-Higgs search in Chapter 5, only normal distributions:

$$p(y; \nu) = N(y; \nu, \sigma_\nu) = \frac{1}{\sigma_\nu \sqrt{2\pi}} \exp\left(-\frac{(y - \nu)^2}{2\sigma_\nu^2}\right) \quad (4.8)$$

are used. By default, y is set to zero so that $\nu = 0$ is the most likely value of the nuisance parameter based upon $p(y; \nu)$ alone. This is also referred to as the nominal value of the nuisance parameter.

For a process p , and category c , the expected number of events for a process, $\lambda_{pc}(\vec{\mu}, \vec{\nu})$, is factorized into a component related to the POIs, and components related to the nuisance parameters:

$$\lambda_{pc}(\vec{\mu}, \vec{\nu}) = g_{pc}(\vec{\mu}) \sum_k h_{pck}(\nu_k) \quad (4.9)$$

where $\vec{\mu}$ and $\vec{\nu}$ are assumed to be statistically independent. For a given set of $\vec{\mu}$ values, the $g(\vec{\mu})$ component, represents the nominal value of the expected number of events, e.g. in an analysis where the LHC luminosity is part of the measurement, the calculation of $g(\vec{\mu})$ assumes the best-fit value for the luminosity from the auxiliary measurement. Then, the $h(\nu)$ components encode how λ should change if a nuisance parameter is shifted from their nominal values, and these components are approximated by:

$$h(\nu) = \begin{cases} \kappa_{\text{up}}^\nu & \text{if } \nu > 0.5 \\ \kappa_{\text{down}}^{-\nu} & \text{if } \nu < -0.5 \\ \exp(\nu K(\kappa_{\text{up}}, \kappa_{\text{down}}, \nu)) & \text{otherwise} \end{cases} \quad (4.10)$$

where K takes a form [84] such that the first and second derivatives of $h(\nu)$ are continuous for all values of ν and that when $\kappa = \kappa_{\text{down}} = 1/\kappa_{\text{up}}$, the form reduces to:

$$h(\nu) = \kappa^\nu. \quad (4.11)$$

The values of κ_{up} and κ_{down} are taken from the impact that the nuisance parameter has on the rate of the process at $\nu = -1$ and $\nu = 1$, where a value of $\nu = 0$ represents no impact at all. For example, if the LHC luminosity measurement is $X_{-5\%}^{+10\%}$, then for a nuisance that encodes this uncertainty, $\kappa_{\text{up}} = 1.1$ and $\kappa_{\text{down}} = 0.95$.

In shape analyses, the dependence of the distribution's shapes on the nuisances parameters must also be encoded. In the di-Higgs search, this is achieved by first fitting an

analytical function, $f(m_{\gamma\gamma}; \vec{\mu}, \vec{\Psi})$, to the simulated events where the nuisance parameters, ν_k , are at their nominal values. The fitted shape parameters, $\vec{\Psi}$, are treated as nominal values, and then the dependence of the shape parameters on the nuisance parameters are encoded by:

$$\Psi_i = \hat{\Psi}_i (1 + \sum_k \alpha_{ik} \nu_k) \quad (4.12)$$

where α_{ik} are determined by studying how the distributions of the simulated events change where the nuisance parameters are shifted. More details on this procedure are provided in Section 5.8.

4.2 Inference

4.2.1 Significances

In the search for the $pp \rightarrow X \rightarrow H(H/Y) \rightarrow \gamma\gamma\tau\tau$ process in Chapter 5, there is a single parameter of interest, μ , which is the cross section for the process. The significance of the observed data with respect to the background-only hypothesis is calculated using a test statistic, q_0 , defined from the likelihood function:

$$q_0 = \begin{cases} -2 \ln \frac{\mathcal{L}(0, \hat{\vec{\nu}}(0))}{\mathcal{L}(\hat{\mu}, \hat{\vec{\nu}})} & \text{if } \hat{\mu} > 0, \\ 0 & \text{otherwise} \end{cases} \quad (4.13)$$

where $\hat{\mu}$ is the maximum likelihood estimator for μ , and $\hat{\vec{\nu}}(\mu)$ and $\hat{\vec{\nu}}$ are the values of $\vec{\nu}$ that maximize the likelihood assuming a specific value of μ and $\mu = \hat{\mu}$ respectively [85]. A measure of the compatibility of the data with the background-only hypothesis is given by the p-value, p_0 , which is calculated as:

$$p_0 = \int_{q_0^{\text{obs}}}^{\infty} f(q_0|0) dq \quad (4.14)$$

where q_0^{obs} is the test statistic for the observed data, and $f(q_0|0)$ is the distribution of the test statistic determined by the generation of pseudo-data sets assuming the background-only hypothesis where the nuisance parameters are set to $\hat{\vec{\nu}}(\mu)$ and the auxiliary observables are sampled from their probability distributions. The significance, Z , is then defined by the p-value according to:

$$Z = \Phi^{-1}(1 - p_0) \quad (4.15)$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard normal distribution.

4.2.2 Limit Setting

To calculate upper limits on the cross section of the di-Higgs process, a different test statistic, $q(\mu)$, is used [85], which is defined by:

$$q(\mu) = \begin{cases} -2 \ln \frac{\mathcal{L}(\mu, \hat{\hat{\nu}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\hat{\nu}})} & \text{if } 0 \leq \hat{\mu} \leq \mu, \\ -2 \ln \frac{\mathcal{L}(\mu, \hat{\hat{\nu}}(\mu))}{\mathcal{L}(0, \hat{\hat{\nu}}(0))} & \text{if } \hat{\mu} < 0, \\ 0 & \text{if } \hat{\mu} > \mu \end{cases} \quad (4.16)$$

and two p-values, p_μ and p_b , are calculated with:

$$p_\mu = \int_{q_{\text{obs}}(\mu)}^{\infty} f(q(\mu)|\mu) dq \quad (4.17)$$

and

$$p_b = \int_{q_{\text{obs}}(\mu)}^{\infty} f(q(\mu)|0) dq \quad (4.18)$$

where $f(q(\mu)|\mu)$ is the distribution of the test statistic determined by the generation of pseudo-data sets like in the calculation of a significance, except now assuming a non-zero value for μ . The p-values, p_μ and p_b , are measures of the compatibility of the data with the signal-plus-background hypothesis, assuming a particular value of μ , or assuming $\mu = 0$ (background-only) respectively.

Any value of μ which has a $p_\mu < \alpha$ can then be said to be excluded at the $100(1 - \alpha)\%$ confidence level. However, this approach is considered too aggressive because a downward fluctuation in the background would lead to smaller values of p_μ , thereby excluding a greater range of μ . To counteract this, μ values are instead excluded according to the $\text{CL}_s = p_\mu/p_b$ criterion [15, 86, 87], and an upper limit at the $100(1 - \alpha)\%$ confidence level is quoted as the value of μ which satisfies $\text{CL}_s = \alpha$.

Expected upper limits are calculated by first generating pseudo-data sets assuming the background-only hypothesis. Then, values of the test statistic that correspond to the 2.5, 16, 50, 85 and 97.5% quantiles are taken from the generated distribution, and the corresponding upper limits are calculated by replacing q_{obs} with these values. The upper limit for the 50% quantile is referred to as the median expected limit.

The upper limits can also be calculated without generating pseudo-data sets by relying on an asymptotic approximation for $f(q(\mu)|\mu)$ according to the prescription in Ref. [85]. This is the approach taken during the category optimization in the di-Higgs search in Chapter 5 because for the sake of an optimization study, the approximation is good

enough, and the asymptotic approach is computationally faster than generating pseudo-data sets, which is useful since the optimization requires the upper limits to be recalculated thousands of times.

4.2.3 Maximum Likelihood Estimates and Confidence Intervals

In the EFT interpretation of Chapter 6, the Wilson coefficients are estimated by maximizing the likelihood function, $\mathcal{L}(\vec{\mu}, \vec{\nu})$, with respect to the Wilson coefficients, and confidence intervals are estimated using the profile likelihood ratio:

$$q(\vec{\mu}) = -2 \ln \frac{\mathcal{L}(\vec{\mu}, \hat{\vec{\nu}}(\vec{\mu}))}{\mathcal{L}(\hat{\vec{\mu}}, \hat{\vec{\nu}})} \quad (4.19)$$

where $\hat{\vec{\mu}}$ are the maximum likelihood estimates for the POIs, which are the Wilson coefficients in this case, and $\hat{\vec{\nu}}(\vec{\mu})$ and $\hat{\vec{\nu}}$ are the values of the nuisance parameters that maximize \mathcal{L} for a particular set of $\vec{\mu}$ and for the maximum likelihood estimates ($\hat{\vec{\mu}}$) respectively. In the absence of large non-Gaussian uncertainties, Wilks' theorem states that the profile likelihood ratio test statistic is approximately χ^2 distributed [88–90] with k degrees of freedom, where k is the number of POIs. In Chapter 6, scans of the likelihood function are performed one Wilson coefficient at a time ($k = 1$) where the other Wilson coefficients are either set to zero, or left floating (are profiled). So, using Wilks' theorem, an approximate 68% (95%) confidence interval is given by the values of μ where $q(\vec{\mu}) = 1$ (3.84).

4.2.4 Binned χ^2 Fits

In the case of a simple binned fit where there are no nuisance parameters, the likelihood function is a product of Poisson terms:

$$\mathcal{L}(\vec{\mu}) = \prod_i \frac{1}{n_i!} \lambda_i(\vec{\mu})^{n_i} e^{-\lambda_i(\vec{\mu})} \quad (4.20)$$

where n_i and $\lambda_i(\vec{\mu})$ are the observed and expected number of events in bin i . Assuming that λ is large ($\gtrsim 10$), a Poisson distribution can be approximated by a normal distribution with a mean and variance equal to λ :

$$\mathcal{L} \approx \prod_i \frac{1}{\sqrt{2\pi\lambda_i(\vec{\mu})}} \exp\left(-\frac{(n_i - \lambda_i(\vec{\mu}))^2}{2\lambda_i(\vec{\mu})}\right) \quad (4.21)$$

and the test statistic (Eq. 4.19) becomes:

$$q(\vec{\mu}) \approx \chi^2(\vec{\mu}) = \sum_i \frac{(n_i - \lambda_i(\vec{\mu}))^2}{\sigma_i^2} \quad (4.22)$$

where $\sigma_i^2 = \lambda_i(\vec{\mu})$ is the variance, and χ^2 is referred to as the chi-squared test statistic. Often, a further approximation is made by estimating the variance from the observed number of events using $\sigma_i = \sqrt{n_i}$, which is a fair approximation when λ is large [15]. This chi-squared test statistic is often used in fits to weighted simulated events where the distribution of events in each bin cannot be described by a Poisson distribution. In these fits, the variance in each bin is given by $\sigma_i^2 = \sum_j w_{ij}^2$. This type of fit is used in the $\gamma\gamma\tau\tau$ search in Chapter 5, primarily to derive models of the $m_{\gamma\gamma}$ distribution for the single and di-Higgs processes (Section 5.8.1).

Chapter 5

Search for Di-Higgs in Final States with Photons and Tau Leptons

5.1 Introduction

The cross section for the simultaneous production of a pair of Higgs bosons (di-Higgs production) at the LHC has a SM prediction of about 35 fb at $\sqrt{s} = 13$ TeV [22]. This is three orders of magnitude lower than the cross section for the production of a single Higgs boson, and as such, makes the search for di-Higgs production exceedingly difficult. Nevertheless, the CMS and ATLAS collaborations have performed these searches [91–93] and while they have not discovered di-Higgs production yet, they still act as important tests of the SM. One of the LO Feynman diagrams for di-Higgs production, shown in Fig. 5.1 (left), includes the HHH vertex, thereby allowing these searches to place direct constraints on the Higgs self-coupling, λ , or alternatively on $\kappa_\lambda = \lambda/\lambda_{\text{SM}}$.

The type of di-Higgs production in the SM is *nonresonant*, meaning that it does not proceed via an intermediate resonance. In some BSM theories, di-Higgs production can proceed via new resonances (*resonant* production), with a rate substantially higher than

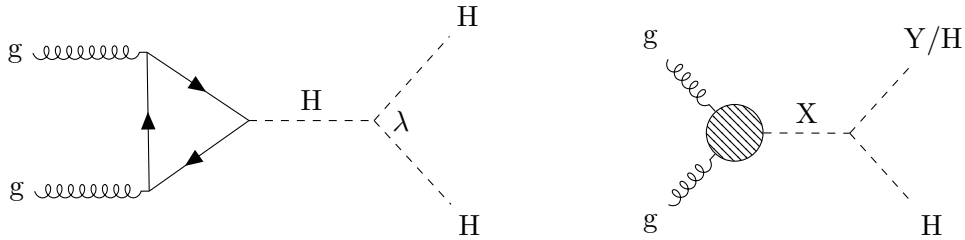


Figure 5.1: Examples of LO Feynman diagrams for nonresonant (left) resonant (right) di-Higgs production. In the resonant diagram, Y is a new scalar.

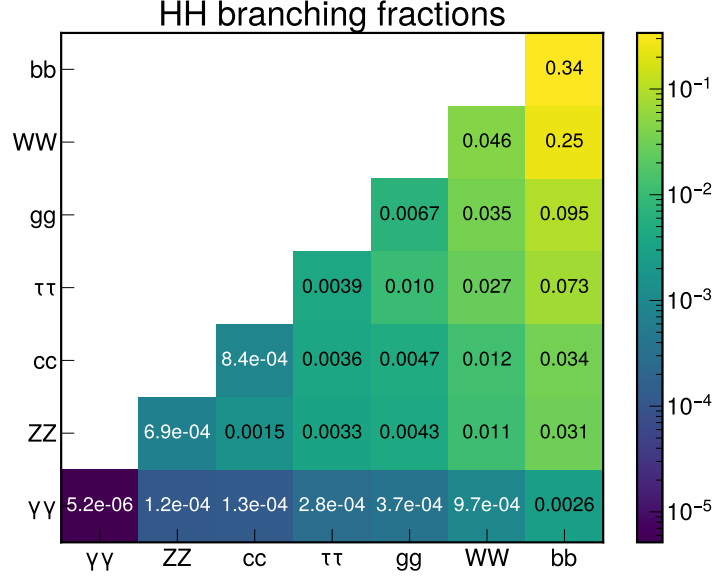


Figure 5.2: Branching fractions of HH to different final states in the SM. For example, $\mathcal{B}(\text{HH} \rightarrow \gamma\gamma\tau\tau) = 2.8 \times 10^{-4}$.

the SM prediction. An example for resonant production is shown in the Feynman diagram in Fig. 5.1 (right). In the WED and NMSSM theories described in Sections 2.3.1 and 2.3.2, the possible resonant processes include $X^{(0)} \rightarrow \text{HH}$ and $X^{(2)} \rightarrow \text{HH}$, where $X^{(0)}$ and $X^{(2)}$ are new spin-0 and spin-2 particles, respectively, and $X \rightarrow \text{YH}$, where X and Y are new spin-0 particles. To gain a complete understanding of di-Higgs production at the LHC, it is important to search for both resonant and nonresonant production, and for a particular final state, the two are often done in parallel.

In designing a search for di-Higgs production, one must select which final states to use. To maximize sensitivity, it is important to choose a final state that has a combination of a high branching fraction, high signal efficiency, and low background contamination. Figure 5.2 shows the HH branching fractions for the most common Higgs boson decays. The $\text{HH} \rightarrow b\bar{b}b\bar{b}$ decay has the highest branching fraction at 33% but suffers from a large QCD multi-jet background, and relatively poor resolution on the reconstructed Higgs boson mass. By substituting one of the $b\bar{b}$ decays with $\gamma\gamma$, background levels are greatly reduced, and the analysis benefits from the good $m_{\gamma\gamma}$ resolution. This makes the $\text{HH} \rightarrow b\bar{b}\gamma\gamma$ decay one of the most sensitive channels for HH production at the LHC, despite its lower branching fraction of 0.26%.

Searches for nonresonant and resonant di-Higgs production in the $b\bar{b}\gamma\gamma$ final state have been performed by the CMS experiment with data from proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ corresponding to an integrated luminosity of 138 fb^{-1} [94, 95]. The

nonresonant search observed an upper limit on $\sigma(pp \rightarrow HH \rightarrow b\bar{b}\gamma\gamma)$ of 5.2 times the SM prediction, and constrained the Higgs self coupling to $-3.3 < \kappa_\lambda < 8.5$ [94]. The resonant search explored $X \rightarrow HH$ and $X \rightarrow YH$ production, with $m_X \in [260, 1000]$ GeV and $m_Y \in [300, 1000]$ GeV respectively. In $X \rightarrow YH$ production, the $Y \rightarrow b\bar{b}$ decay with $m_Y \in [70, 800]$ GeV was considered. In the $X \rightarrow HH$ searches, no significant deviations from the background-only hypothesis were observed, and upper limits on $\sigma(pp \rightarrow X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma)$ were reported as 0.82–0.07 and 0.87–0.06 fb depending on m_X for the spin-0 and spin-2 scenarios respectively. In the $X \rightarrow YH$ search, an excess with a local (global) significance of 3.8 (2.8) standard deviations was observed at $m_X = 650$ GeV and $m_Y = 90$ GeV [95]. All upper limits and confidence intervals were set at 95% CL.

The excess at $m_X = 650$ GeV and $m_Y = 90$ GeV in the $X \rightarrow YH$ search is particularly interesting because it aligns with several others excesses reported by the CMS experiment. Searches for new scalars decaying to WW , $\tau\tau$, and $\gamma\gamma$ final states have reported local (global) significances of 3.8 (2.6), 2.8 (2.2) and 2.9 (1.3) standard deviations for scalar masses of 650 GeV, 100 GeV and 95 GeV respectively [96–98]. Assuming that these excesses are real and hint toward a new particle sector, they imply that $\sigma(X \rightarrow YH)$ and decay rates of the new scalars to $b\bar{b}$, WW , $\tau\tau$, and $\gamma\gamma$ are high enough to be observed (or almost observed) at the LHC.

This motivates a search for $X \rightarrow YH$ production in the $\gamma\gamma\tau\tau$ final state, which has not been explored yet. Furthermore, for low values of m_Y (~ 100 GeV) in the $Y \rightarrow \gamma\gamma$ search, this analysis provides a unique opportunity to constrain regions of the NMSSM phase space, as discussed in Section 2.3.2. In the $X \rightarrow HH$ searches, the results are not expected to be more sensitive than those reported by the $b\bar{b}\gamma\gamma$ search since the $H \rightarrow \tau\tau$ branching fraction is lower than for $H \rightarrow b\bar{b}$, but the search is motivated nonetheless by a future combination of results that will use all available final states to search for $X \rightarrow HH$ production with greater sensitivity than the individual analyses.

This chapter presents a search for resonant di-Higgs production in the $\gamma\gamma\tau\tau$ final state using proton-proton collision data collected by the CMS experiment between 2016 and 2018 at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 138 fb^{-1} . The search consists of four main analyses, each targeting a different process:

1. $X^{(0)} \rightarrow HH$: resonant HH production via a spin-0 particle ($X^{(0)}$)
2. $X^{(2)} \rightarrow HH$: resonant HH production via a spin-2 particle ($X^{(2)}$)
3. $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$: resonant YH production via a scalar, X , where $Y \rightarrow \tau\tau$
4. $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$: resonant YH production via a scalar, X , where $Y \rightarrow \gamma\gamma$

For the $X \rightarrow HH$ searches, $m_X \in [260, 1000]$ GeV, and for the $X \rightarrow YH$ searches, $m_X \in [300, 1000]$ GeV and $m_Y \in [50, 800]$, where in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, the lower

bound on m_χ is instead 70 GeV due to triggering requirements. While these processes are motivated by the WED and NMSSM theories, the searches are designed to be model-independent, except for the assumption that the new particles have a narrow width. The final results are significances in the case of signal excesses, and upper limits on the cross section of the targeted processes.

5.2 Analysis Strategy

Events are selected beginning with one of two different diphoton triggers, one which requires $m_{\gamma\gamma} > 90$ GeV, and another than requires $m_{\gamma\gamma} > 65$ GeV but has a reduced luminosity of 132 fb^{-1} . The former is used for all searches except for the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search when $m_Y < 125$ GeV. The $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search is therefore treated as two different searches, a *low-mass* search for $m_Y < 125$ GeV and a *high-mass* search for $m_Y > 125$ GeV. Following the trigger, preselection criteria are applied which includes the requirement of at least one tau lepton candidate, and a selection on $m_{\gamma\gamma}$ which corresponds to 100–180 GeV for the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches, while in the low and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches, it corresponds to 65–150 GeV and 100–1000 GeV respectively. The triggering and preselection requirements are described in greater detail in Section 5.4.

Toy examples of the $m_{\gamma\gamma}$ distributions for the signal and background processes for each search are shown in Fig. 5.3. In every search, there is a nonresonant background which is predominantly comprised of prompt photon and diphoton production with associated jets (γ +jets and $\gamma\gamma$ +jets), where the jets are misidentified as photons and/or tau leptons. There are also smaller contributions from $V + \gamma$, $t\bar{t}$, $t\bar{t} + \gamma$, and $t\bar{t} + \gamma\gamma$ where the vector bosons (from direct production or the decay of top quarks) can decay into tau leptons and associated jets can be misidentified as photons. This nonresonant background forms a smoothly-falling distribution in $m_{\gamma\gamma}$.

Furthermore, there is a resonant background in every search from single Higgs boson production in the SM, where the Higgs boson decays to two photons. This background leads to a peak in the $m_{\gamma\gamma}$ distribution at the Higgs boson mass (≈ 125 GeV). In the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, there is another resonant background present at about 91 GeV from $Z \rightarrow e^+e^-$ decays where both electrons are misidentified as photons. This background is referred to as the DY background since it predominantly comes from the Drell-Yan (DY) process.

The signal process forms a peak in the $m_{\gamma\gamma}$ distribution at a location depending on the search and m_Y . In the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches, the two photons originate from a Higgs boson and therefore form a peak at about 125 GeV. In the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches, the photons originate from the Y scalar and therefore form a peak at m_Y .

To improve the sensitivity of the search, signal events are separated from background events using a machine learning classifier. Neural networks that are trained to be parametric in m_X , and m_Y in case of the $X \rightarrow YH$ searches, are used for this purpose. A parametric neural network (pNN) is one whose classification behaviour is dependent on the specified m_X (and m_Y) and leads to better signal-to-background separation at each

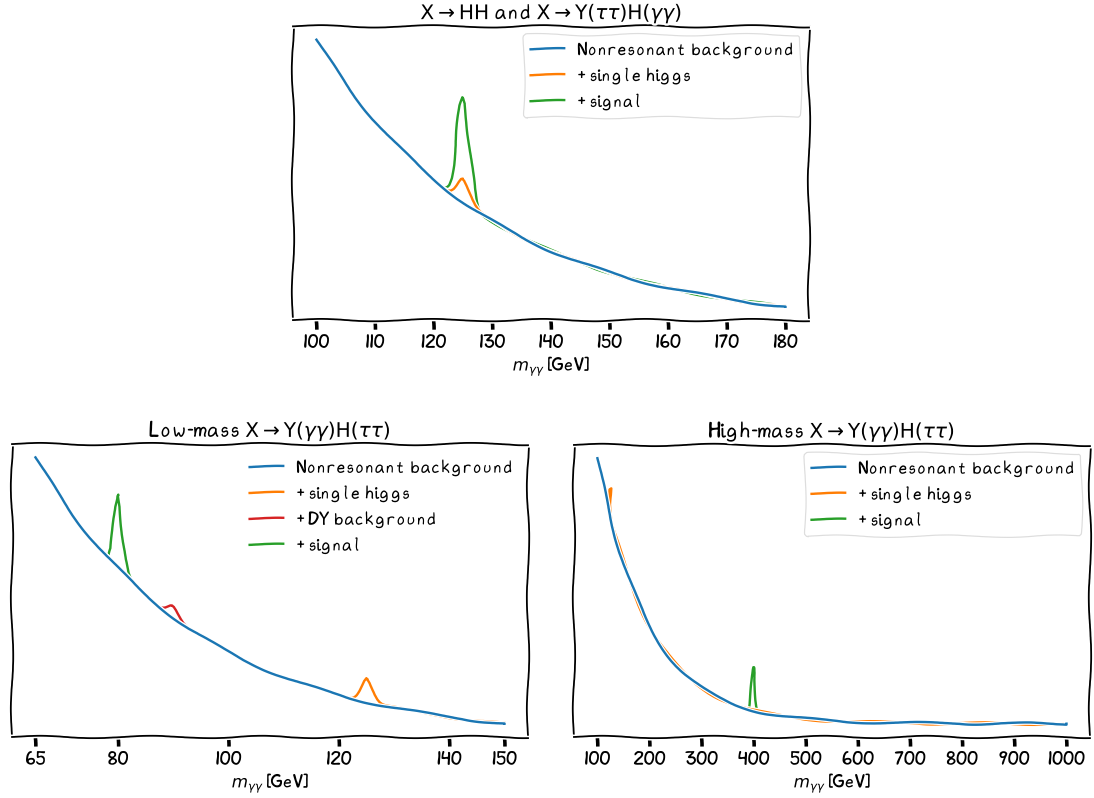


Figure 5.3: Toy examples of the $m_{\gamma\gamma}$ distributions of the background and signal processes for each search. In each search, there is a smoothly-falling nonresonant background (blue), and a background from SM single Higgs production (orange). In the low-mass search, there is also a DY background (red) originating from the misidentification of electrons as photons in $Z \rightarrow e^+e^-$ decays. The signal peaks at about 125 GeV in the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches, and at m_Y in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches, where examples at $m_Y = 80$ and 400 GeV are shown for the low and high-mass searches respectively. The shape and relative rates of every process are not exact and are for illustrative purposes only.

mass point than possible with a single non-parametric neural network. The training and performance of the pNNs is described in Section 5.5.

In every search, and for every mass point, a set of analysis categories are defined using the pNN output that are optimized for the best expected sensitivity. The granularity in m_X (and m_Y) that the results are reported at depends on the behaviour of the pNN selection, and on the $m_{\gamma\gamma}$ resolution in the case of the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches. The granularity is set such that the search is sensitive to any signal within the mass range(s) of interest, and the procedure that specifies this is described in Section 5.6.

Results are extracted through maximum likelihood fits to the $m_{\gamma\gamma}$ distributions in the analysis categories. The signal and background models are defined as analytic functions of $m_{\gamma\gamma}$, where the signal and background models are determined by fits to simulated events or to data, depending on the process. The creation of these models is described in Section 5.8, and the final results are presented in Section 5.9.

5.3 Samples

This analysis uses data collected by the CMS experiment in pp collisions with $\sqrt{s} = 13$ TeV during 2016, 2017 and 2018, corresponding to integrated luminosities of 36.3, 41.5 and 59.8 fb^{-1} for each year respectively, and a total of 138 fb^{-1} [42–44]. In 2018, the HLT path used for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search was not introduced until after the first 5.1 fb^{-1} of data was collected, and therefore the luminosity for this search in 2018 is reduced to 54.7 fb^{-1} , corresponding to a total of 132 fb^{-1} across 2016, 2017, and 2018.

Monte-Carlo (MC) simulation is generated for the signal processes at LO accuracy using MADGRAPH5_aMC@NLO v2.6.5 [99]. The values of m_X (and m_Y for $X \rightarrow YH$ searches) that samples are generated at is shown in Fig. 5.4. In this Chapter, these values are referred to as *nominal* mass points. Results are reported at these values, as well as at *intermediate* mass points, which are defined in Section 5.6.

The γ +jets and $\gamma\gamma$ +jets backgrounds are modelled with MADGRAPH5_aMC@NLO v2.6.5 [99, 100] and SHERPA v2.2.1 [101] respectively, both at LO accuracy. The other nonresonant backgrounds: $V + \gamma$, $t\bar{t}$, $t\bar{t} + \gamma$, and $t\bar{t} + \gamma\gamma$, are modelled at NLO accuracy in perturbative QCD with MADGRAPH5_aMC@NLO v2.6.5 [99, 100, 102] except for $t\bar{t}$ which uses MADGRAPH5_aMC@NLO v2.6.1.

To model the SM single Higgs background, simulated events are generated for the ggH, VBF, VH and ttH production modes using MADGRAPH5_aMC@NLO v2.4.2 [99, 100, 102] at NLO accuracy in perturbative QCD. Further production modes are not simulated because their contributions are negligible. The samples are normalized to the SM cross sections and branching fractions provided by the LHC Higgs Cross Section Working Group (Tables 2.2 and 2.3), which are calculated at higher orders in perturbative QCD and EW theory [22].

In the generation of all simulated samples, parton distribution functions are taken from the NNLO NNPDF 3.1 set [103]. Samples are interfaced with PYTHIA 8.240 [104] with the CP5 tune [105, 106] for parton showering, fragmentation with the standard p_T -ordered parton shower scheme, and the underlying event description. Finally, the detector response is modelled using the GEANT4 package [107] and the events are reconstructed using the same algorithms as for the data.

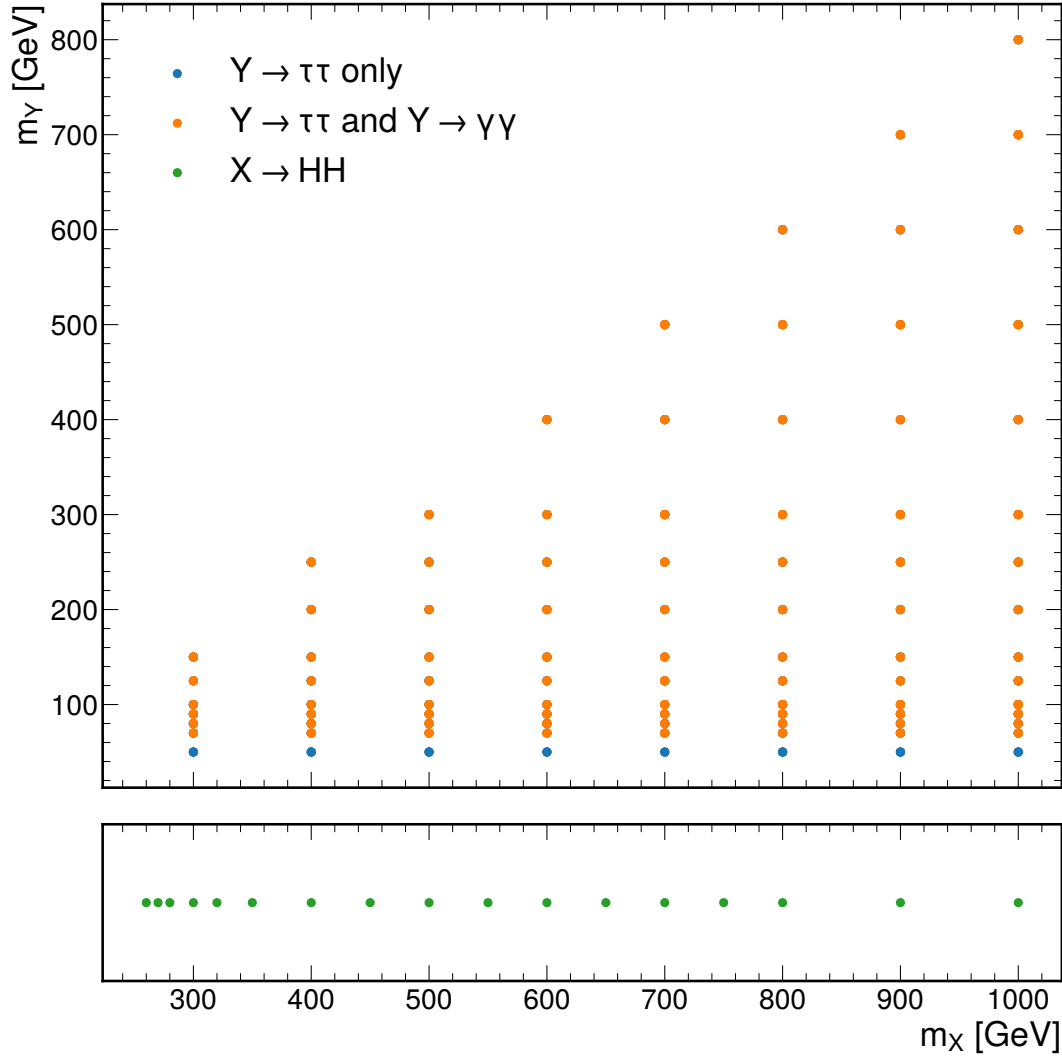


Figure 5.4: Granularity in m_X (and m_Y) that MC samples are generated at for the $X \rightarrow HH$ (green), $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (orange) and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (orange and blue) processes. These points define the so-called nominal mass points.

5.4 Trigger and Preselection

Events from data-taking periods that are certified for physics events are first selected according to HLT paths that require two photons that pass a set of criteria on the p_T of the photons and the invariant mass of the diphoton system. The HLT path used in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search requires $m_{\gamma\gamma} > 55 \text{ GeV}$ and $p_T > 30(18) \text{ GeV}$ for the leading (subleading) photon. The HLT path used for all of the other searches in this analysis requires $m_{\gamma\gamma} > 90 \text{ GeV}$ and has the same photon p_T requirements except in 2017 and 2018 where the subleading photon threshold was increased to 22 GeV . The HLT paths also apply a set of loose requirements on shower shape and isolation variables of the photons. The impact of including tau lepton triggers was studied, but the improvement in signal efficiencies was found to be less than 1% and therefore, these triggers were not used in the final analysis.

Triggered events are then required to pass a set of preselection criteria. This selection requires tau lepton candidates in the event, applies the particle identification algorithms described in Section 3.3, and places loose requirements on kinematic properties of the photons and tau lepton candidates. Selections are also placed on a number of shower shape and isolation variables for the photons: I_{ch} , H/E , R_9 , $\sigma_{i\eta i\eta}$ and I_γ (all defined in Section 3.3.5), and on impact parameters d_{xy} and d_z , which are the shortest distances of the corresponding particle track to the primary vertex in the transverse and longitudinal directions respectively.

The preselection requirements are kept as loose as possible, either aligning with the trigger (HLT) requirements, which were originally tuned to maximize the signal efficiency for SM H production for a particular trigger rate/bandwidth, or aligning with requirements used to derive simulation corrections, e.g. the electron ID efficiency scale factors (Section 3.3.5) were calculated for electrons with $p_T > 10 \text{ GeV}$. Later, the pNN will be used to apply the final selection designed to optimize the sensitivity of the analysis.

Beginning with the photons, the preselection requirements depend on which search is being performed where different requirements are made for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search compared to the rest of the searches. Starting with the requirements for the rest of the searches, all photon candidates are required to have:

- $p_T > 25 \text{ GeV}$,
- $|\eta| < 2.5$ and not $1.44 < |\eta| < 1.57$ (to exclude the barrel-endcap transition region),
- $I_{\text{ch}} < 20 \text{ GeV}$,
- $I_{\text{ch}}/p_T^\gamma < 0.3$,
- $H/E < 0.08$,
- Photon ID WP90

Table 5.1: A subset of the preselection requirements on photon candidates that are based upon η , R_9 , $\sigma_{i\eta i\eta}$, I_γ , and I_{tk} .

	R_9	$\sigma_{i\eta i\eta}$	I_γ (GeV)	I_{tk} (GeV)
Barrel	[0.50, 0.85]	<0.015	<4.0	<6.0
	>0.85	—	—	—
Endcaps	[0.80, 0.90]	<0.035	<4.0	<6.0
	>0.90	—	—	—

Table 5.2: Preselection requirements on electrons, muons, τ_h , and jets. Descriptions about the ID algorithms for electrons, muons and τ_h can be found in Section 3.3. In addition to the listed criteria, selected electrons, muons and τ_h are required to have $\Delta R > 0.2$ with respect to each other and selected photons, and jets are required to have $\Delta R > 0.4$ with respect to all selected objects. Finally, electrons whose trajectories are consistent with the barrel-endcap transition region ($1.44 < |\eta| < 1.57$) are excluded.

	p_T	$ \eta $	$ d_{xy} $	$ d_z $	ID
Electrons	>10 GeV	<2.5	<0.045	<0.2	WP90
Muons	>15 GeV	—	<0.045	<0.2	Medium cut-based
τ_h	>20 GeV	<2.3	—	<0.2	VVLoose D_e , VLoose D_μ , Loose D_{jet}
Jets	>25 GeV	<2.4	—	—	—
IsoTracks	>5 GeV	—	<0.2	<0.1	—

and pass additional η -dependent requirements on R_9 , $\sigma_{i\eta i\eta}$ and I_γ , which are summarized in Table 5.1. These requirements also include a selection on track isolation, I_{tk} , which is the sum of track p_T in a hollow cone with a smaller (larger) annulus of $\Delta R = 0.04$ (0.33). For the $X \rightarrow \text{HH}$ and $X \rightarrow Y(\tau\tau)\text{H}(\gamma\gamma)$ searches, diphoton candidates are then formed by pairing photons that have $100 < m_{\gamma\gamma} < 180$ GeV and where the leading (subleading) photon has $p_T > 35$ (25) GeV. Additionally, the leading (subleading) photon is required to have $p_T/m_{\gamma\gamma} > 0.33$ (0.25). The same requirements are applied in the high-mass $X \rightarrow Y(\gamma\gamma)\text{H}(\tau\tau)$ search except the $m_{\gamma\gamma}$ requirement is changed to $100 < m_{\gamma\gamma} < 1000$ GeV.

In the low-mass search, the p_T requirement on photon candidates is lowered to 18 GeV, and the pixel veto (Section 3.3.5) is applied to reduce contamination from the DY background. All other photon requirements are the same. When forming diphoton candidates, the leading (subleading) photon is required to have $p_T > 30$ (18) GeV, the $p_T/m_{\gamma\gamma}$ requirements are removed, and $65 < m_{\gamma\gamma} < 150$ GeV is required. In all searches, if more than one diphoton candidate is found, the candidate with the highest p_T is chosen.

Electron, muon, τ_h , and jet candidates are selected according to the criteria summarized in Table 5.2. This includes selection on the candidate p_T , η , ID variables, and the impact parameters, d_{xy} and d_z , which are the shortest distances of the track to the

primary vertex in the transverse and longitudinal directions respectively.

Ditau candidates are formed by creating pairs of candidates from the selected τ_h , muon, electron candidates according to the following priority: $\tau_h\tau_h$, $\tau_h\mu$, $\tau_h e$, μe , $\mu\mu$, ee , where a particular pairing is referred to as a channel. If there is one τ_h candidate, and no electron or muon candidates, tracks with $I_{\text{ch}} < 5 \text{ GeV}$ and $p_T > 5 \text{ GeV}$, or $I_{\text{ch}}/p_T < 0.2$ are considered as additional tau lepton candidates. These isolated tracks (IsoTracks) are used to create an additional channel: $\tau_h + \text{IsoTrack}$. Events are selected if they have a ditau candidate, or if there is a single τ_h , which is the final channel: τ_h .

Events with an opposite-sign same flavour ee or $\mu\mu$ pair are rejected if they are consistent with a $Z \rightarrow ll$ or $Z \rightarrow ll\gamma$ ($l = e, \mu$) decay. This is done by rejecting any events where $80 < m_{ll} < 100 \text{ GeV}$ or $86 < m_{ll\gamma} < 96 \text{ GeV}$ where $m_{ll\gamma}$ is calculated with respect to both the leading and subleading photons.

The efficiency of the triggering and preselection requirements for signal events is calculated in simulation with a number of corrections applied. These corrections include the object energy scale, energy resolution, and ID corrections discussed in Section 3.3. They also include corrections for the triggering and photon preselection requirements which are derived from data with the “tag-and-probe” method using $Z \rightarrow e^+e^-$ events [108], and are applied by reweighting the simulated events in bins of p_T , R_9 and η [16, 109, 110].

The signal efficiency after triggering and preselection requirements is shown as a function of the resonance mass, m_X , for the $X \rightarrow \text{HH}$ searches in Fig. 5.5, which is seen to increase as a function of m_X . This is expected since a higher m_X will lead to candidates with higher p_T , which will be more likely to pass the p_T requirements. Furthermore, the efficiency of ID requirements tend to increase with higher p_T (see Fig. 3.14).

The preselection efficiency in the $X \rightarrow \text{YH}$ searches is shown in Fig. 5.6 as functions of m_X and m_Y . As in the $X \rightarrow \text{HH}$ searches, the efficiency increases with m_X . In the high-mass $X \rightarrow \text{Y}(\gamma\gamma)\text{H}(\tau\tau)$ search, the trend in m_Y is such that the highest efficiencies are found at low m_Y , whereas in the low-mass search less of a significant trend is seen, though that is likely due to the smaller range of m_Y being considered. In the $X \rightarrow \text{Y}(\tau\tau)\text{H}(\gamma\gamma)$ search, a more complex trend in m_Y is seen, where the highest efficiencies are seen for median values of m_Y . This is possibly because at lower values of m_Y , the tau lepton candidates will be less separated and the ΔR requirement between tau lepton candidates will reduce the efficiency.

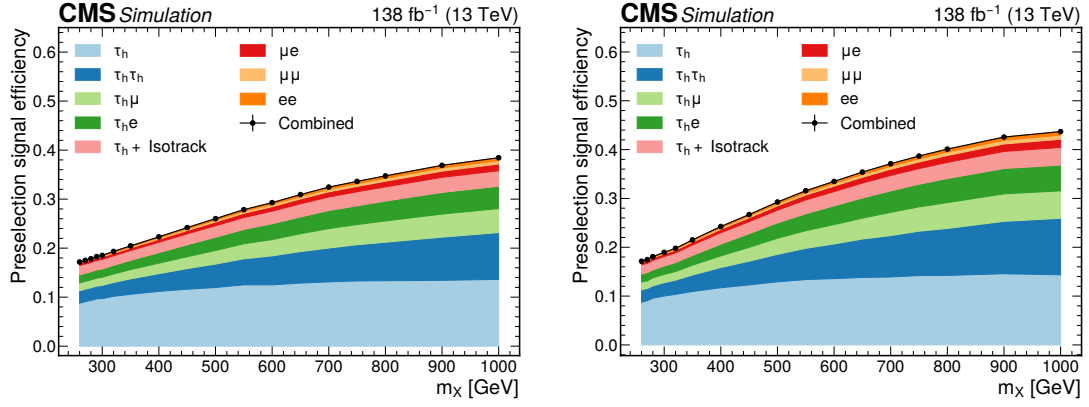


Figure 5.5: Preselection signal efficiency as a function of m_X in the $X^{(0)} \rightarrow HH$ (left) and $X^{(2)} \rightarrow HH$ (right) searches. The contributions from the different channels are shown by the stacked filled histograms and the combined efficiency shown by the black points. The statistical uncertainty in the combined uncertainty is plotted but is too small to be seen.

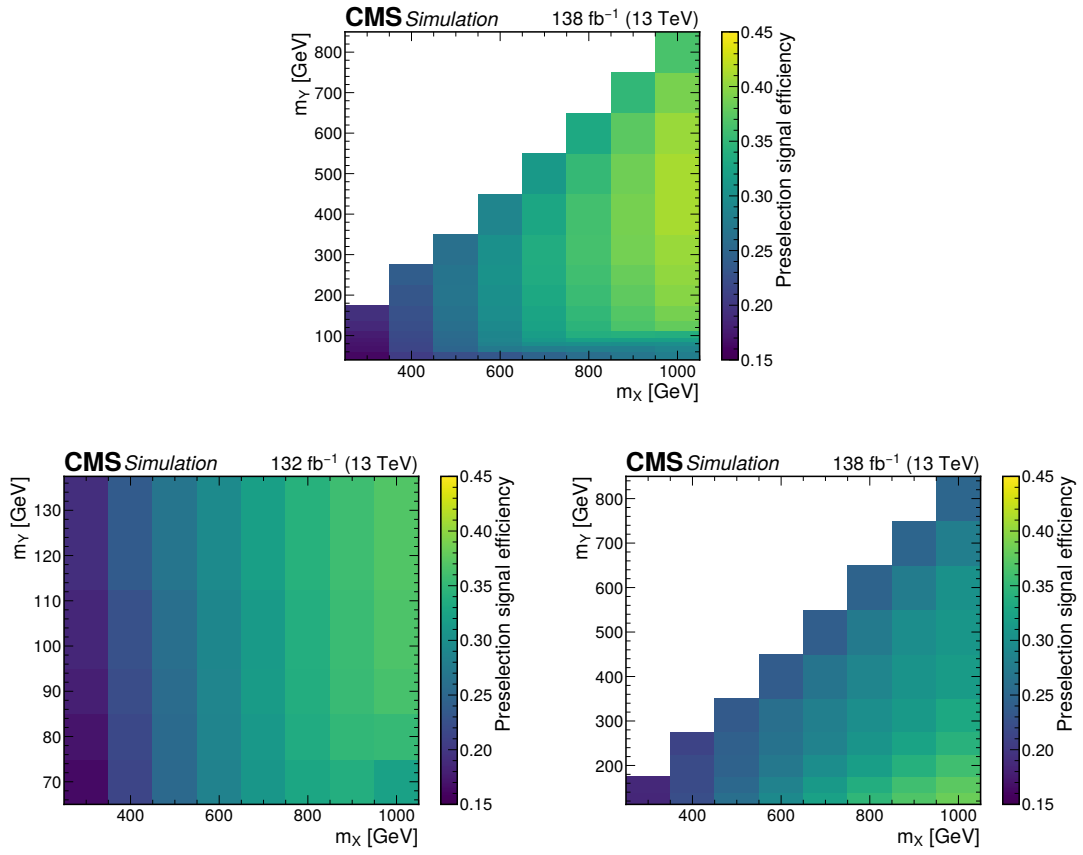


Figure 5.6: Preselection signal efficiency as a function of m_X and m_Y in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (top), low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom-left), and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom-right) searches.

5.5 Event Selection

5.5.1 Training Features

Variables considered as inputs to the pNN (training features) are mostly kinematic variables of the photons, tau lepton candidates, jets, and MET in the event. This includes angular relations ($\Delta\eta$, $\Delta\phi$, ΔR) between these objects, and the invariant masses of combinations of them. Also included are the b jet ID for selected jets, particle multiplicities, the event channel, and the data-taking year. The full set of 67 variables is listed in Table 5.3.

The variables for the leading and subleading tau lepton candidates only include information from the visible decay products of the tau lepton whereas in the ditau system, the variables are computed with the SVFIT algorithm [111] which incorporates a correction from the MET. The invariant mass of the diphoton and ditau systems, $m(\gamma\gamma\tau\tau)$, and the invariant mass of the diphoton system, the leading tau lepton candidate, and the MET, $m(\gamma\gamma\tau_1\text{MET})$, are included as estimators of m_X . The latter variable was calculated neglecting the z -component of the MET, and was created specifically in mind of the τ_h category where there is only a single tau lepton candidate.

When selecting input features for the pNNs, it is important to consider the background modelling. The techniques described in Section 5.8.2 assume that the nonresonant backgrounds are smoothly falling in the $m_{\gamma\gamma}$ distribution. If they are not, a peak in the background could be mistakenly reported as a signal excess. To avoid this, input features that are highly correlated with $m_{\gamma\gamma}$ are removed or transformed. Therefore, $m_{\gamma\gamma}$ is not included as a feature. However, it is not always obvious whether other features will lead to sculpting until after the pNN has been trained and therefore, feature selection and training can be a recursive process. In this section, the initial steps made to reduce sculpting and the subsequent feature selection are first described. Then, at the end of this section, final alterations made to the training features after studies into sculpting are described.

Initial steps to reduce sculpting are to divide the photon p_T by $m_{\gamma\gamma}$ since its scale is partially (completely if the Higgs boson is produced at rest) determined by $m_{\gamma\gamma}$. This is standard practice for $H \rightarrow \gamma\gamma$ analyses in the CMS collaboration [110] and whilst it will not completely remove the correlation, it is found sufficient to remove any significant sculpting. For similar reasons, the $m(\gamma\gamma\tau\tau)$ and $m(\gamma\gamma\tau_1\text{MET})$ variables are also divided by $m_{\gamma\gamma}$.

Not all the 67 considered variables provide good discrimination power and their inclusion can lead to longer training times which in turn slows down studies into the event selection. Therefore, a subset of features which retain the majority of the discrimination power is desired. A feature selection procedure was performed for the $X^{(2)} \rightarrow HH$ search

Leading and subleading photon	$p_T/m_{\gamma\gamma}, \eta, \phi$
Leading and subleading tau candidates	$p_T, \eta, \phi, m, \text{charge, PDG ID}$
Leading and subleading jet	$p_T, \eta, \text{DeepJet ID}$
Highest b-tag scoring jet	DeepJet ID
MET	p_T, ϕ
Diphoton system	$p_T/m_{\gamma\gamma}, \eta, \phi, \Delta\eta, \Delta\phi, \Delta R,$
Ditau system	$p_T, \eta, \Delta\eta, \Delta\phi, \Delta R, m$
Angular variables	$\Delta\phi(\gamma\gamma, \text{MET}), \Delta\phi(\tau\tau, \text{MET}),$ $\Delta\phi(\tau_1, \text{MET}), \Delta\phi(\tau_2, \text{MET}),$ $\Delta[\eta, \phi, R](\gamma\gamma, \tau_1), \Delta[\eta, \phi, R](\gamma\gamma, \tau_2),$ $\Delta[\eta, \phi, R](\gamma\gamma, \tau\tau),$ $\Delta R(\gamma_1, [\tau\tau, \tau_1, \tau_2]), \Delta R(\gamma_2, [\tau\tau, \tau_1, \tau_2])$
Multiplicity	$N_e, N_\mu, N_\tau, N_{\text{IsoTracks}}, N_{\text{jets}}, N_{\text{b-jets}}$
Invariant masses	$m(\gamma\gamma\tau\tau)/m_{\gamma\gamma}, m(\gamma\gamma\tau_1\text{MET})/m_{\gamma\gamma},$ $m(\tau\tau\gamma_1), m(\tau\tau\gamma_2)$
Other	channel, year

Table 5.3: Training features considered for the training of pNNs before feature selection. There are 67 features in total. Leading and subleading particles are denoted by p_1 and p_2 respectively. The $[x, y, z]$ notation means there is a variable for each choice in $[x, y, z]$, e.g. $\Delta[\eta, \phi, R](\gamma\gamma, \tau_1)$ corresponds to $\Delta\eta(\gamma\gamma, \tau_1), \Delta\phi(\gamma\gamma, \tau_1), \Delta R(\gamma\gamma, \tau_1)$.

Photons	$p_T^{\gamma\gamma}/m_{\gamma\gamma}, p_T^{\gamma_1}/m_{\gamma\gamma}, p_T^{\gamma_2}/m_{\gamma\gamma}, \Delta R(\gamma\gamma), \Delta\phi(\gamma\gamma)$
Tau candidates	$p_T^{\tau_1}, m(\tau_1), p_T^{\tau\tau}, m(\tau\tau), \Delta R(\tau\tau), \Delta\eta(\tau\tau), \Delta\phi(\tau\tau)$
Reconstructed m_X	$m(\gamma\gamma\tau\tau)/m_{\gamma\gamma}, m(\gamma\gamma\tau_1\text{MET})/m_{\gamma\gamma}$
Additional angular variables	$\Delta\phi(\gamma\gamma, \text{MET}), \Delta\phi(\tau\tau, \text{MET}),$ $\Delta[\eta, \phi, R](\gamma\gamma, \tau_1), \Delta[\eta, \phi, R](\gamma\gamma, \tau_2),$ $\Delta\phi(\gamma\gamma, \tau\tau), \Delta\eta(\gamma\gamma, \tau\tau),$ $\Delta R(\gamma_1, \tau\tau), \Delta R(\gamma_1, \tau_1), \Delta R(\gamma_2, \tau_1)$
Other	$m(\tau\tau\gamma_1), p_T^{\text{MET}}, p_T^{j_1}, \text{channel}$

Table 5.4: The most important features found after a selection procedure based on the feature importance from BDTs trained on the $X^{(2)} \rightarrow \text{HH}$ signal samples. There are 31 features in total. Leading and subleading particles are denoted by p_1 and p_2 respectively. The $[x, y, z]$ notation means there is a variable for each choice in $[x, y, z]$, e.g. $\Delta[\eta, \phi, R](\gamma\gamma, \tau_1)$ corresponds to $\Delta\eta(\gamma\gamma, \tau_1), \Delta\phi(\gamma\gamma, \tau_1), \Delta R(\gamma\gamma, \tau_1)$.

and the results were used for all searches. The beginning of the procedure is as follows:

1. For every generated mass point (Fig. 5.4), train a Boosted Decision Tree (BDT) using all considered features to separate the $X^{(2)} \rightarrow \text{HH}$ sample from all background samples. The BDT is trained using the XGBOOST package [112] with the default model parameters.
2. For every generated mass point, identify the least important training feature using the ‘gain’ feature importance from XGBOOST and remove it from the list of features.
3. Iteratively train the BDTs, removing features one-by-one until a significant drop in performance is seen. The metric for performance was chosen as the signal efficiency at a 1% background efficiency and the iterations were stopped after 17 features remained in each BDT which corresponded to a maximal loss of 0.5% signal efficiency across all mass points. Removing further features led to $> 1\%$ loss in signal efficiency which was considered too significant.
4. Create a set of features from the union of the remaining features from each BDT. This ensures that the feature set is appropriate for any kinematic regime dictated by the range of m_X .

Following these steps, a set of 31 features is found (see Table 5.4). This size of this set can be further reduced without significant loss of information by removing features which are highly correlated with each other.

m_X	AUC full	AUC subset	Sig. eff. full	Sig. eff. subset	Sig. eff. difference
260	0.9292	0.9279	48.43%	48.13%	-0.30%
300	0.9513	0.9505	50.71%	51.77%	-1.06%
400	0.9928	0.9926	86.05%	85.58%	-0.47%
500	0.9980	0.9980	96.44%	96.48%	+0.04%
800	0.9998	0.9998	99.85%	99.84%	-0.01%
1000	0.9999	0.9999	99.95%	99.95%	0.00%

Table 5.5: Comparison of a BDT's performance to discriminate $X^{(2)} \rightarrow HH$ signal samples from background when using the full set of 67 considered training features and when using the subset of 28 features found after the BDT feature importance procedure and the removal of three highly correlated features. The signal efficiencies are calculated at a background efficiency of 1%.

A dataset is created with an equal number ($\sim 35K$) of unweighted events sampled from the $m_X = 300$ GeV signal sample and the combined background sample. Only events which belong to categories with two tau lepton candidates are used so that all the variables are well-defined. The Spearman's rank correlation between the remaining 31 features in this dataset is calculated and shown in Fig. 5.7. As expected, blocks of high correlation can be found for variables which are derived from the same objects, and particularly high correlation is found for related variables, e.g. $p_T/m_{\gamma\gamma}$ and ΔR for the diphoton system.

When trialling features to remove, the impact on the efficiencies for the $m_X = 260, 300, 400, 500, 800$ and 1000 GeV signal samples (including single tau lepton candidate category) was checked. After removing three features from the most correlated blocks of features, the greatest difference in signal efficiency was found to be 1%, at which point the removal of features was stopped, leaving 28 selected features. These features, which are highlighted in red in Fig. 5.7 are: $\Delta R(\gamma\gamma)$, $\Delta\phi(\gamma\gamma, \tau_1)$ and $\Delta\phi(\gamma\gamma, \tau_2)$. The changes in signal efficiencies for the different mass points are shown in Table 5.5.

When training the pNNs with this selection of 28 features, sculpting of the $m_{\gamma\gamma}$ distribution was found and this lead to the removal of $m(\gamma\gamma\tau_1\text{MET})/m_{\gamma\gamma}$. For all but the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, this was sufficient to eliminate any sculpting. In the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, which has a significantly wider range in $m_{\gamma\gamma}$ than the rest, it was found necessary to also remove $m(\gamma\gamma\tau\tau)/m_{\gamma\gamma}$ and $m(\tau\tau\gamma_1)$. The final set of features used for training the pNNs is shown in Table 5.6.

Distributions of the data and MC given the standard preselection (except for low-mass $Y \rightarrow \gamma\gamma$) in all of the 27 final training features are given in Figs. 5.8 to 5.13. Also shown in these plots are signal distributions for the $X^{(2)} \rightarrow HH$ process for $m_X = 260, 500$ and 1000 GeV. For variables which are undefined in cases where there is a single tau candidate, e.g. $m_{\tau\tau}$, the single tau events are not included in the histograms.

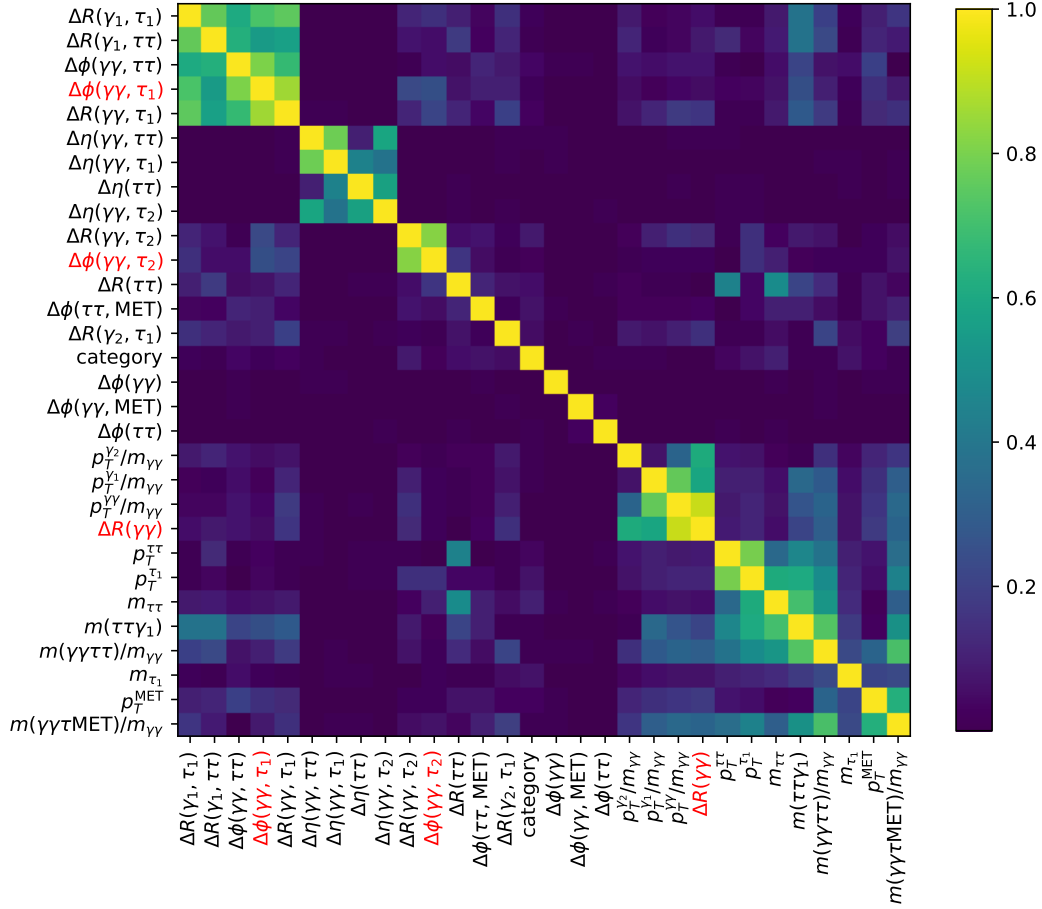


Figure 5.7: The Spearman rank correlation between the 31 features found after a selection procedure based on the feature importance from BDTs trained on the $X^{(2)} \rightarrow HH$ signal samples. The dataset that the correlation is calculated in is a combination of the $X^{(2)} \rightarrow HH$ ($m_X = 300$ GeV) signal sample and the combined background sample. Only events which belong to categories with two tau lepton candidates are used so that all variables are well-defined, and for the same reason, p_T^{j1} is not shown since there is no guarantee that a jet is present in the event. The features highlighted in red are subsequently removed.

Photons	$p_T^{\gamma\gamma}/m_{\gamma\gamma}, p_T^{\gamma_1}/m_{\gamma\gamma}, p_T^{\gamma_2}/m_{\gamma\gamma}, \Delta R(\gamma\gamma), \Delta\phi(\gamma\gamma)$
Tau candidates	$p_T^{\tau_1}, m(\tau_1), p_T^{\tau\tau}, m(\tau\tau), \Delta R(\tau\tau), \Delta\eta(\tau\tau), \Delta\phi(\tau\tau)$
Reconstructed m_X	$m(\gamma\gamma\tau\tau)/m_{\gamma\gamma}^*$
Additional angular variables	$\Delta\phi(\gamma\gamma, \text{MET}), \Delta\phi(\tau\tau, \text{MET}),$ $\Delta[\eta, R](\gamma\gamma, \tau_1), \Delta[\eta, R](\gamma\gamma, \tau_2),$ $\Delta\phi(\gamma\gamma, \tau\tau), \Delta\eta(\gamma\gamma, \tau\tau),$ $\Delta R(\gamma_1, \tau\tau), \Delta R(\gamma_1, \tau_1), \Delta R(\gamma_2, \tau_1)$
Other	$m(\tau\tau\gamma_1)^*, p_T^{\text{MET}}, p_T^{j_1}, \text{category}$

Table 5.6: The 27 final features selected to train the pNNs determined by a feature importance procedure, removing highly correlated variables, and removing variables which lead to sculpting of the $m_{\gamma\gamma}$ distribution. The features, $m(\gamma\gamma\tau\tau)/m_{\gamma\gamma}$ and $m(\tau\tau\gamma_1)$, are marked with an asterisk to indicate that they are not used in the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Leading and subleading particles are denoted by p_1 and p_2 respectively. The $[x, y, z]$ notation means there is a variable for each choice in $[x, y, z]$, e.g. $\Delta[\eta, \phi, R](\gamma\gamma, \tau_1)$ corresponds to $\Delta\eta(\gamma\gamma, \tau_1), \Delta\phi(\gamma\gamma, \tau_1), \Delta R(\gamma\gamma, \tau_1)$.

In general, better separation is seen for the $m_X = 1000 \text{ GeV}$ signal compared to 260 GeV . This is particularly apparent in variables such as $p_T^{\gamma\gamma}/m_{\gamma\gamma}$ (Fig. 5.8), and $m(\gamma\gamma\tau\tau)/m_{\gamma\gamma}$ (Fig. 5.13), but there are also variables such as $\Delta R(\gamma_1, \tau_1)$ (Fig. 5.10) which are more useful at lower m_X . This highlights the need for a selection of training features that considered the whole range of m_X .

The signal distributions of the ΔR variables that represent the angular separation of the diphoton and ditau systems, or components thereof, e.g. $\Delta R(\gamma\gamma, \tau_1)$ in Fig. 5.9, have shapes that might be difficult to understand initially. However, the shapes can be explained by separately considering the $\Delta\phi$ and $\Delta\eta$ components to ΔR . At higher values of m_X , the Higgs bosons from the $X \rightarrow HH$ process are produced with higher p_T and approach the maximum separation in the $x - y$ plane that corresponds to $\Delta\phi = \pi$ (see $\Delta\phi(\gamma\gamma, \tau\tau)$ in Fig. 5.10). Consequently, the ΔR variables peak at $\sim \pi$ and have a spread corresponding to the spread in $\Delta\eta$ (see $\Delta\eta(\gamma\gamma, \tau_1)$ in Fig. 5.9). At lower values of m_X , e.g. $m_X = 260 \text{ GeV}$, there is less separation in the $x - y$ plane and therefore, the ΔR variables spread towards lower values, and still retain a “soft” maximum at $\sim \pi$ due to the $\Delta\phi$ component.

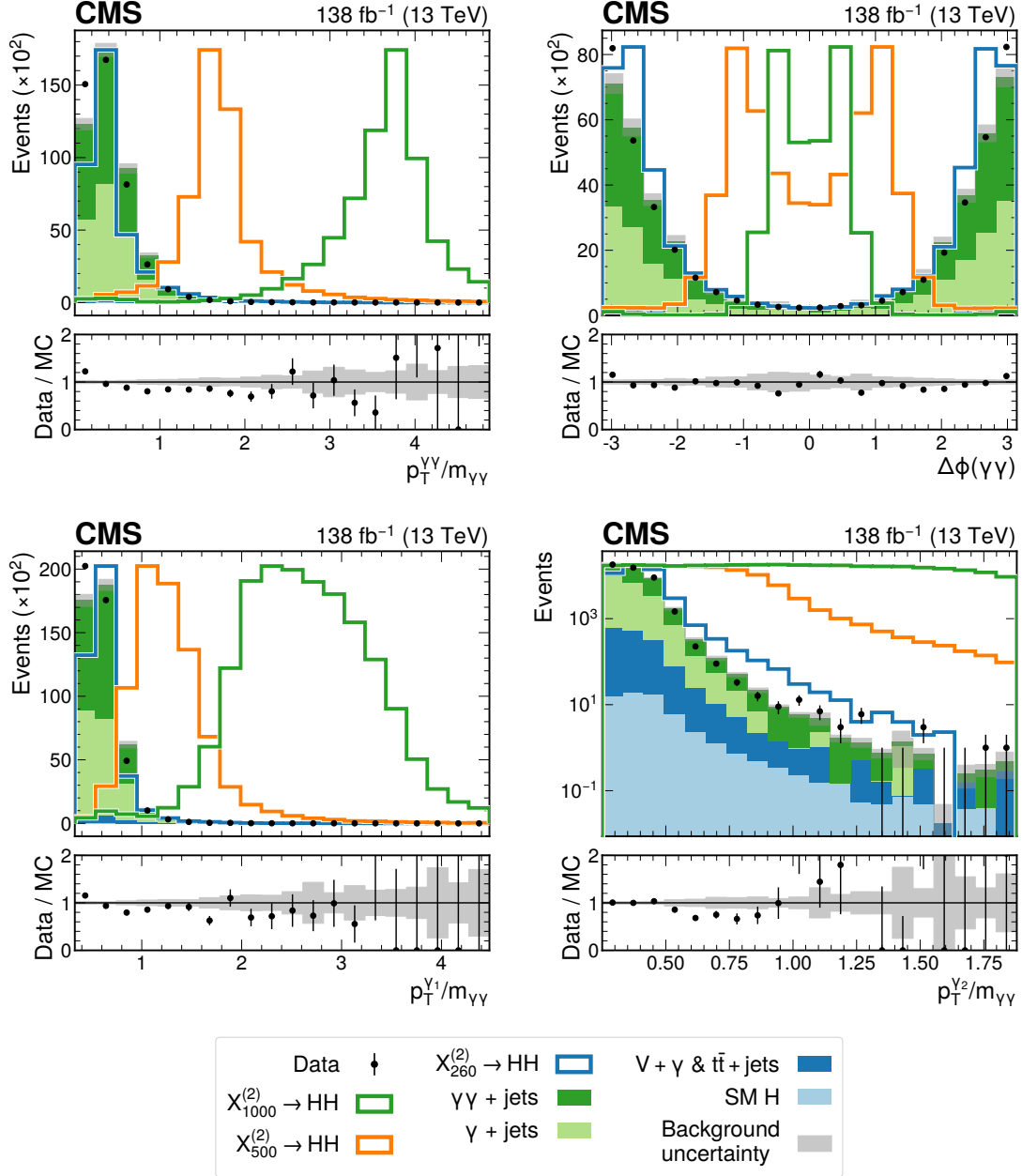


Figure 5.8: Distributions of data, background MC, and $X^{(2)} \rightarrow HH$ signals for $m_X = 260, 700$ and 1000 GeV, in a subset of the variables used to train the pNN. Background MC is normalized to data and the signal's normalization is arbitrary. The statistical uncertainty in the background simulation is shown by the grey shaded bands.

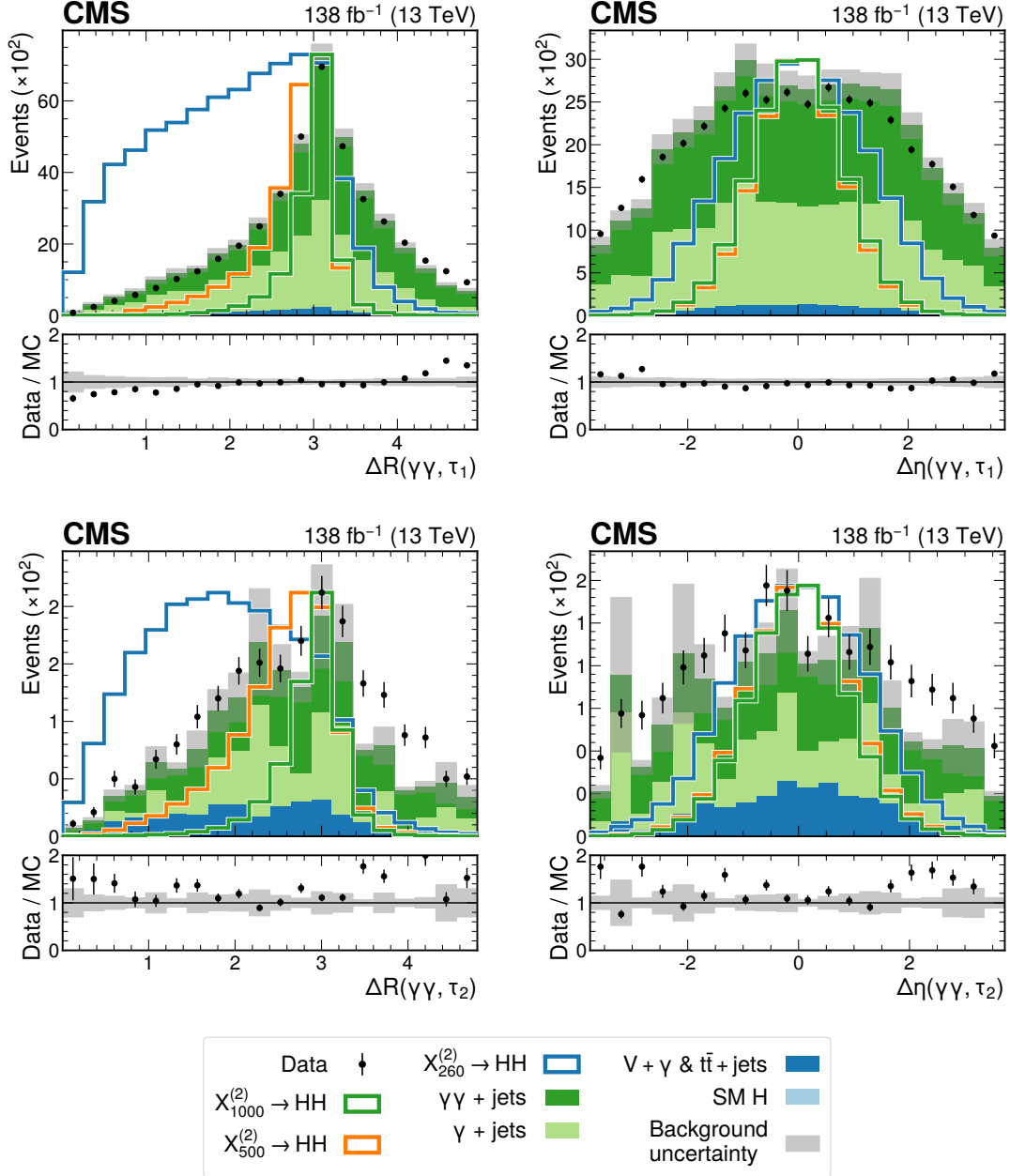


Figure 5.9: Distributions of data, background MC, and $X^{(2)} \rightarrow HH$ signals for $m_X = 260, 700$ and 1000 GeV, in a subset of the variables used to train the pNN. Background MC is normalized to data and the signal's normalization is arbitrary. The statistical uncertainty in the background simulation is shown by the grey shaded bands.

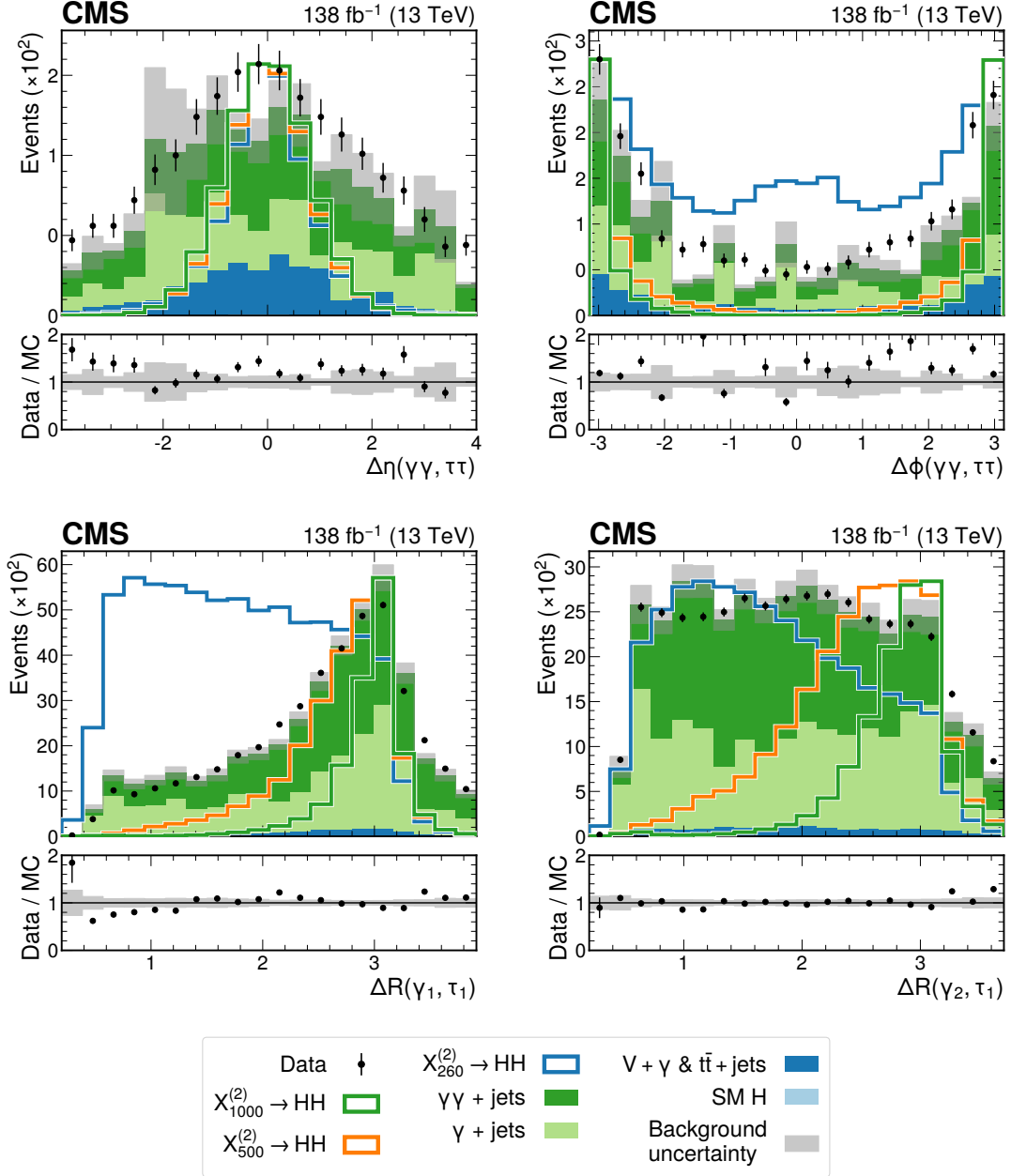


Figure 5.10: Distributions of data, background MC, and $X^{(2)} \rightarrow HH$ signals for $m_X = 260, 700$ and 1000 GeV, in a subset of the variables used to train the pNN. Background MC is normalized to data and the signal's normalization is arbitrary. The statistical uncertainty in the background simulation is shown by the grey shaded bands.

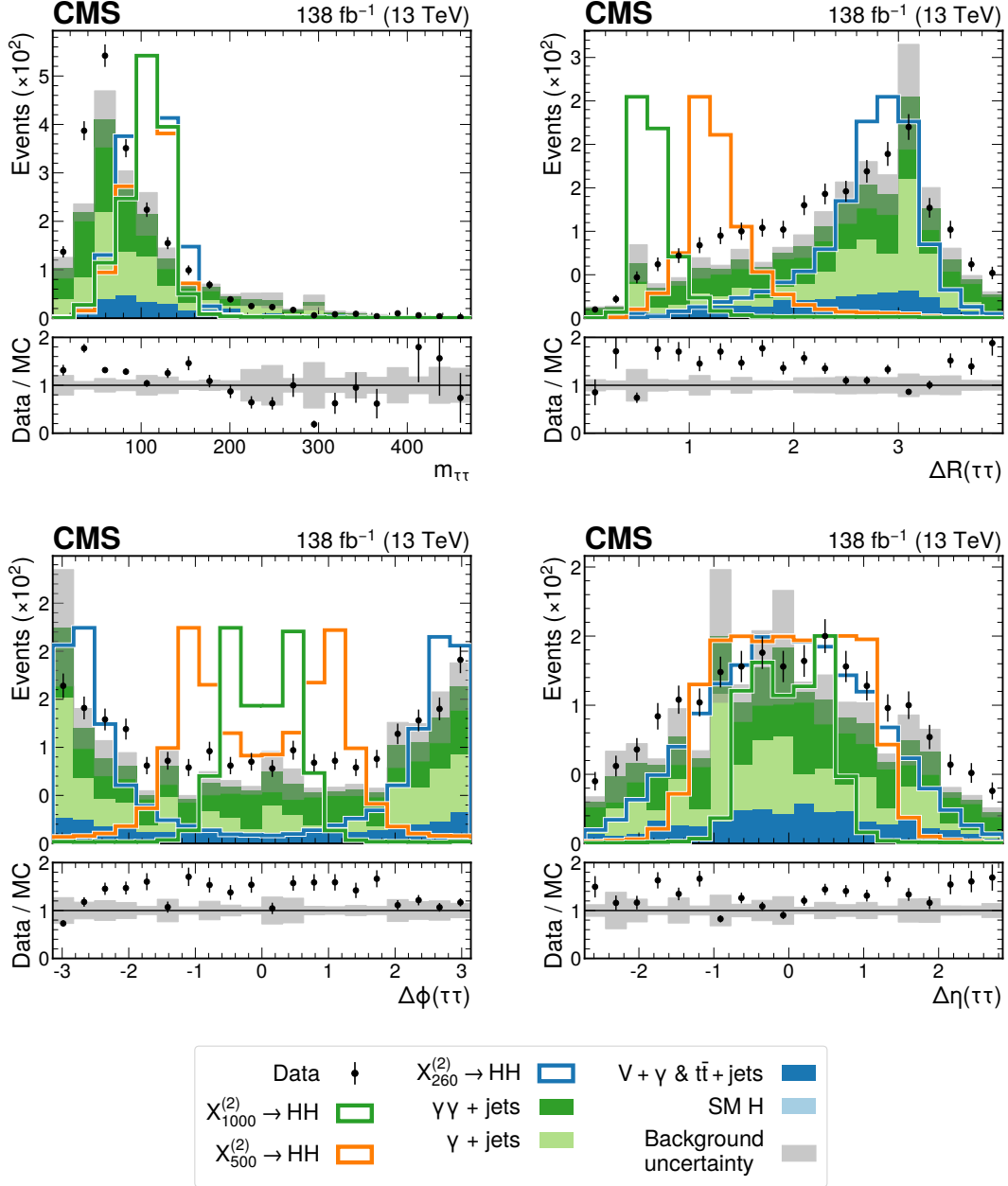


Figure 5.11: Distributions of data, background MC, and $X^{(2)} \rightarrow HH$ signals for $m_X = 260, 700$ and 1000 GeV, in a subset of the variables used to train the pNN. Background MC is normalized to data and the signal's normalization is arbitrary. The statistical uncertainty in the background simulation is shown by the grey shaded bands.

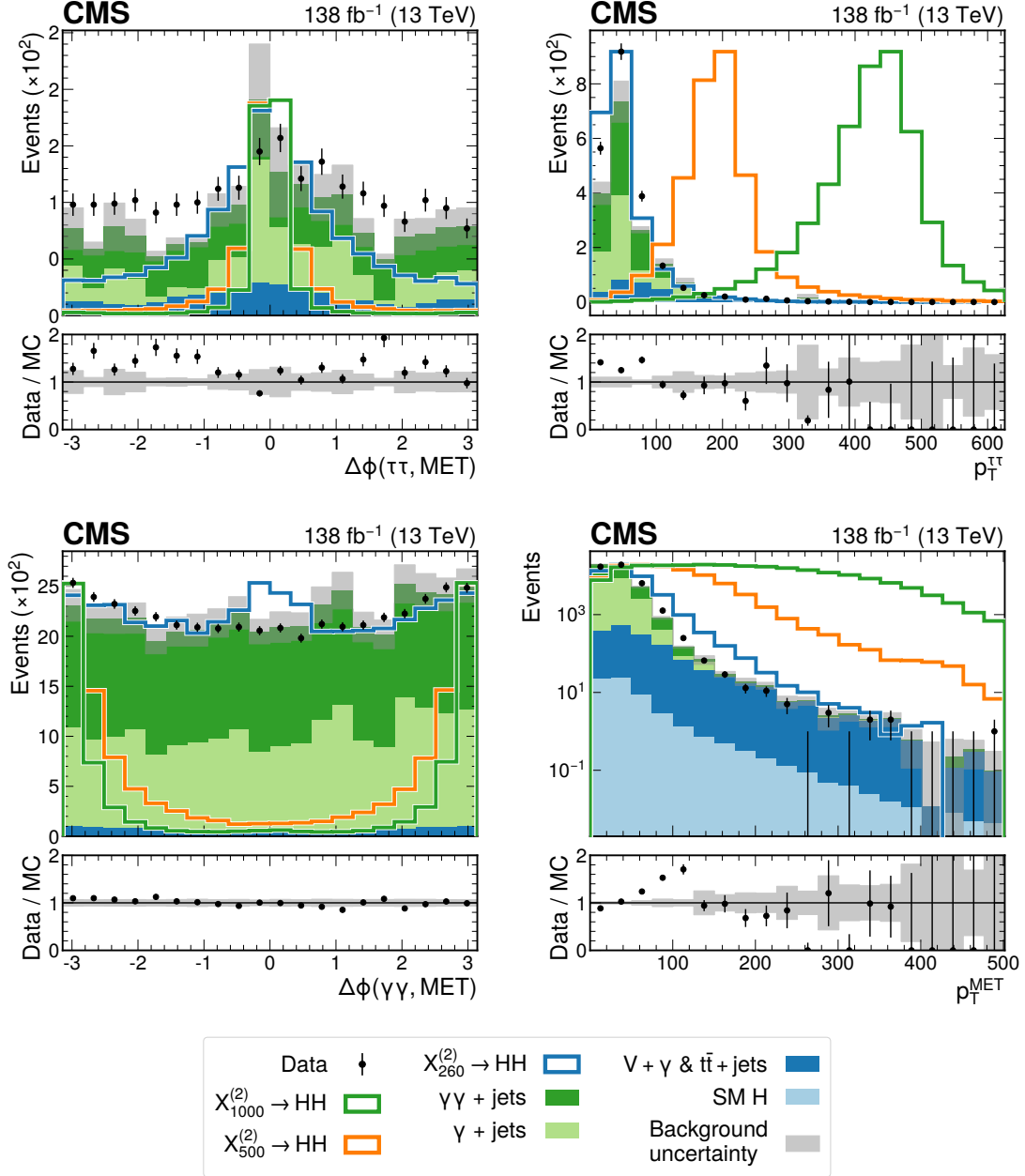


Figure 5.12: Distributions of data, background MC, and $X^{(2)} \rightarrow \text{HH}$ signals for $m_X = 260, 700$ and 1000 GeV , in a subset of the variables used to train the pNN. Background MC is normalized to data and the signal's normalization is arbitrary. The statistical uncertainty in the background simulation is shown by the grey shaded bands.

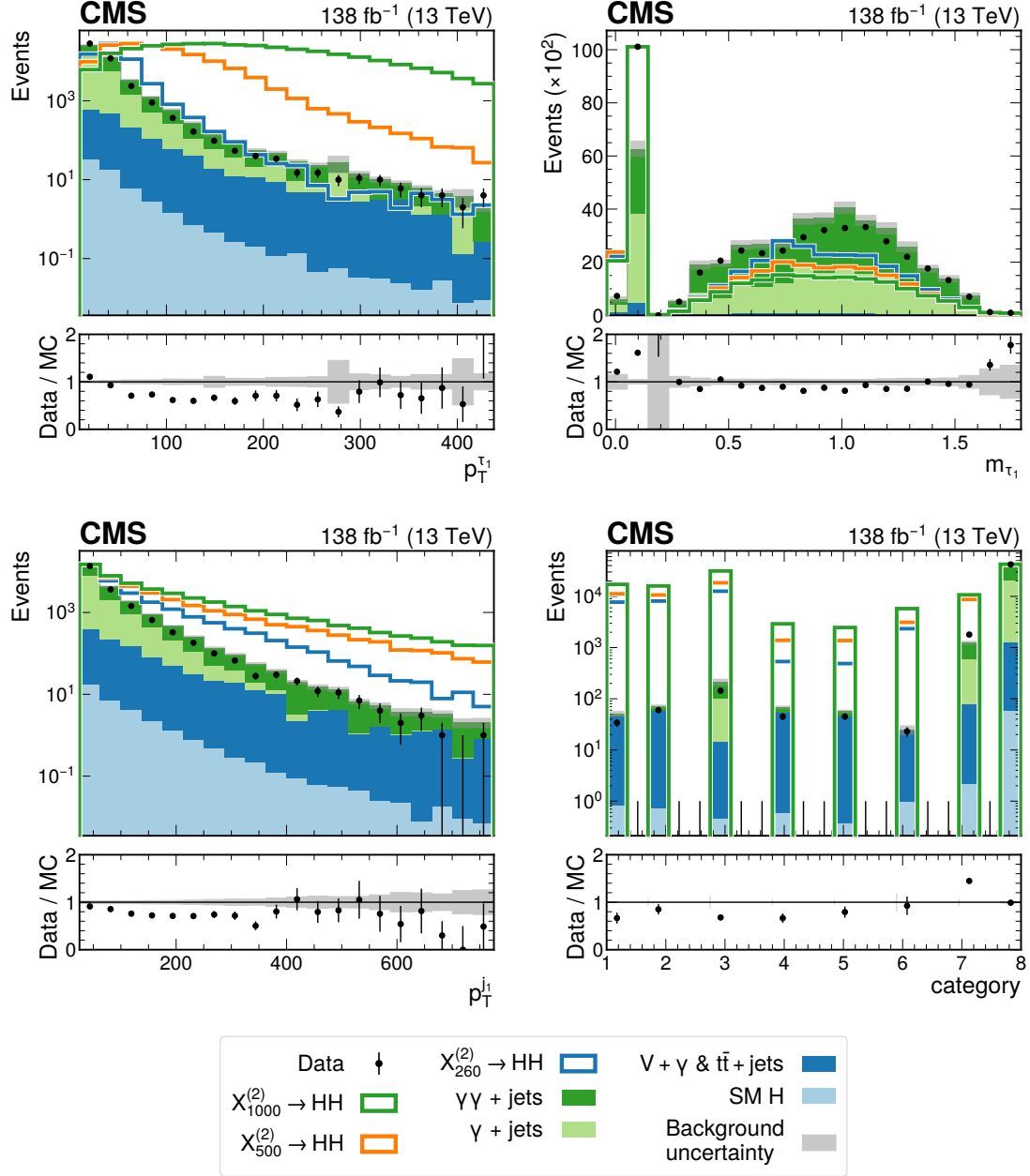


Figure 5.13: Distributions of data, background MC, and $X^{(2)} \rightarrow HH$ signals for $m_X = 260, 700$ and 1000 GeV, in a subset of the variables used to train the pNN. Background MC is normalized to data and the signal's normalization is arbitrary. The statistical uncertainty in the background simulation is shown by the grey shaded bands. The bottom-right plot shows the channel of an event, where the numbering scheme is given in Table 5.7.

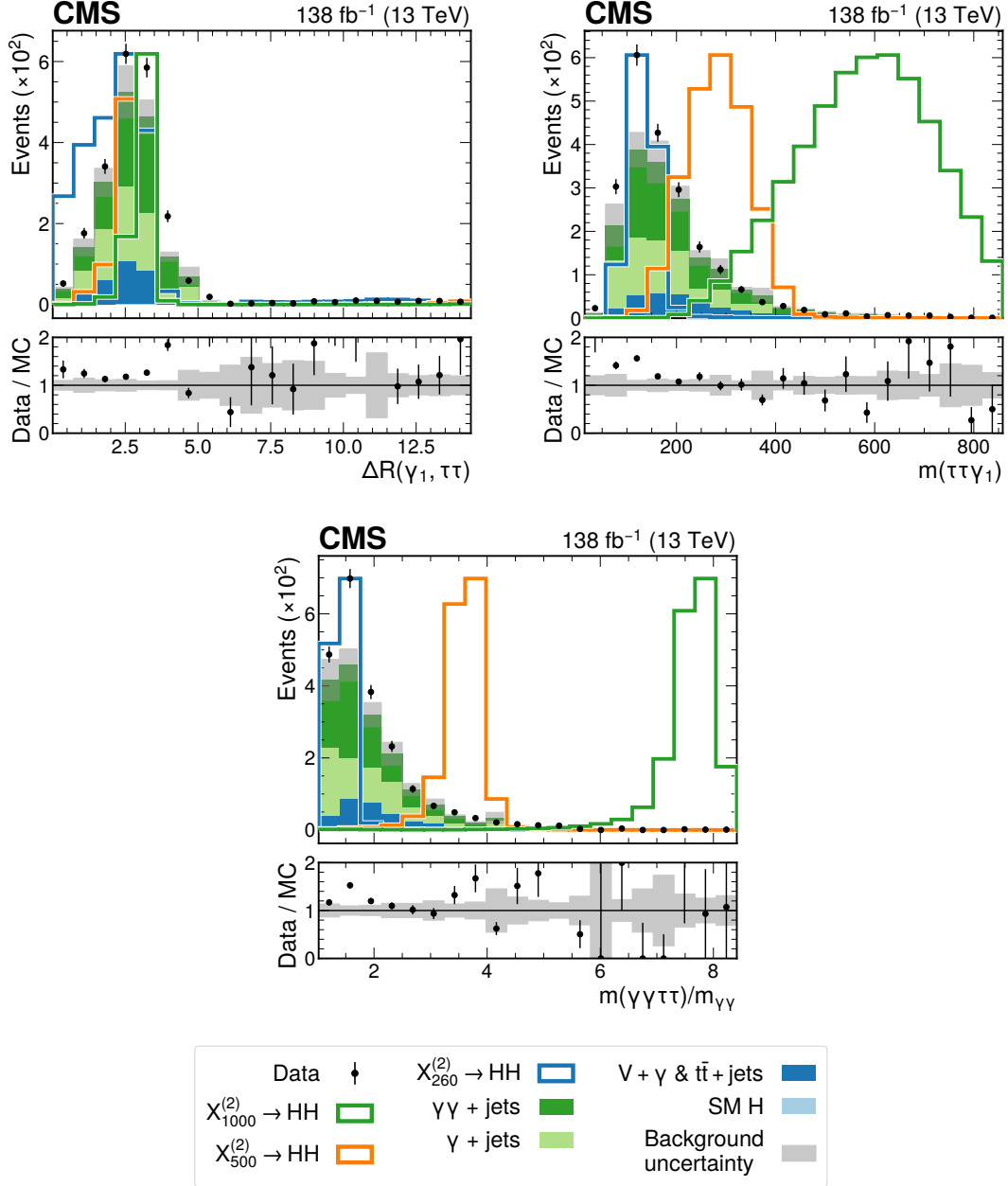


Figure 5.14: Distributions of data, background MC, and $X^{(2)} \rightarrow HH$ signals for $m_X = 260, 700$ and 1000 GeV, in a subset of the variables used to train the pNN. Background MC is normalized to data and the signal's normalization is arbitrary. The statistical uncertainty in the background simulation is shown by the grey shaded bands.

Channel Number	Tau candidates
1	$\tau_h \mu$
2	$\tau_h e$
3	$\tau_h \tau_h$
4	$\mu \mu$
5	ee
6	μe
7	$\tau_h + \text{IsoTrack}$
8	τ_h

Table 5.7: Numbering scheme used to describe the channels in the analysis which differ based upon the types and number of tau lepton candidates reconstructed.

5.5.2 Parametric Neural Networks

In analyses like this one where there are different target signals and many different resonance masses to search for, there are several approaches available to help discriminate signal from background. Examples include:

1. Training a classifier for every target signal and every mass point. In this approach, every combination of process and mass point is treated as a separate analysis whose job is to only target this specific signal. This approach lends itself to good sensitivity since each analysis can be optimized to its signal without compromise introduced by the consideration of other processes and mass points. However, in this analysis, this would lead to the training of 208 classifiers in total and doing this, including hyperparameter optimization, could be a lengthy task.
2. Training a single classifier for all the processes and all the mass points. In this approach, all the signal MC is included in the training at once. Whilst this simplifies the classifier training, it is also non-optimal since it applies the same logic to discriminate against all the different signals. A more sensible approach would be to train a classifier per process with all mass points included, but this would still not be as optimal as approach 1.
3. Grouping together signals which are kinematically similar. Here, multiple classifiers would be trained per process, where each classifier is designed to target a particular kinematic regime. This is the approach taken by a resonant $HH \rightarrow bb\gamma\gamma$ analysis performed by the CMS collaboration where signals are grouped based upon a ‘boost factor’ defined as $m_X/(m_H + m_Y)$ [95]. This represents a compromise between approach 1, which is more sensitive but also more complex, and approach 2, which is more simple but less sensitive.

An alternative to the approaches described above are Parameterized Neural Networks (pNNs) [113]. Per process, a single pNN could be trained that would, in principle, be as optimal as the first approach described above. This is possible because the pNN can change how it uses the discriminating (training) features depending on the value of a mass parameter(s). In more colloquial terms, it changes how to discriminate signal from background depending on the mass point it is asked to discriminate for.

To understand how a pNN works, first consider one of the $X \rightarrow HH$ processes, and a NN trained on a single mass point. This network is a function, $f(\vec{x})$, which can be roughly thought of as the probability that an event, described by training features, \vec{x} , is signal. If following approach 1, one would have a set of such functions $\{f^1, f^2, \dots\}$ where f^i is a NN trained with signal samples corresponding to m_X^i . One could put these functions together to create:

$$f(\vec{x}; m_X) = \begin{cases} f^1(\vec{x}) & \text{if } m_X = m^1 \\ f^2(\vec{x}) & \text{if } m_X = m^2 \\ \vdots & \vdots \end{cases}. \quad (5.1)$$

When using this function, picking a value of m_X is equivalent to picking a particular NN to use. When training a pNN, the target function is $f(\vec{x}; m_X)$, i.e. a single function, which given a value of m_X , will provide a discriminator specific to that value of m_X . In practice, this is achieved by:

1. Including m_X as an additional training feature (this is a different feature to the reconstructed m_X features described in Section 5.5.1).
2. Training the NN on all of the signal MC simultaneously where the events are assigned a value of m_X which corresponds to the dataset they originated from, i.e. events from the $m_X = 300$ GeV dataset will be given a value of 300 for the training feature m_X .
3. Constructing training batches such that there are an equal number of signal events and background events in each batch. Then, randomly pairing each signal event with a background event in the batch and giving the background event the same value of m_X as the signal event. This ensures that on a per-batch basis, that there is no power to discriminate between signal and background using the m_X training feature.
4. Renormalizing the signal datasets ($m_X = 300, 400, \dots, 1000$ GeV) such that the sum of weights in each is the same.
5. Renormalizing the signal (all masses collectively) and background datasets such that the sum of weights is the same between the two.

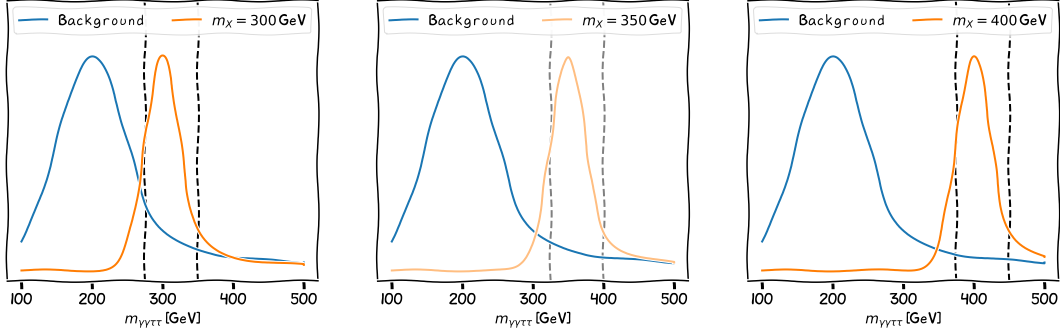


Figure 5.15: Toy examples of distributions of $m(\gamma\gamma\tau\tau)$ in background and in signal for the $m_X = 300, 350$ and 400 GeV $X \rightarrow HH$ signals. In each case, a selection represented by the dashed black lines on $m(\gamma\gamma\tau\tau)$ is proposed to separate signal from background. This selection is used to illustrate the selection that a pNN may effectively apply to this variable. By training the pNN on the $m_X = 300$ and 400 GeV signals, it may learn to apply a selection for the 350 GeV mass point that is somewhere between the selection used for 300 and 400 GeV.

The last two steps are not necessary to make a network parametric in m_X , but they ensure that the network is not biased towards any particular mass point and lead to better overall performance.

It is possible to gain an intuition about what the pNN is doing implicitly. Consider a network trained on a single mass point. At each layer in the network, more complicated features can be made from the outputs of the previous layer. At the end of the network, these features are combined to make a single discriminator, $f(\vec{x})$. If m_X is included as a training feature, then the features can be made to change depending on m_X . For example, the mean value of a feature in signal may scale with m_X , e.g. $m(\gamma\gamma\tau\tau)$ or $p_T^{\gamma\gamma}$. By including m_X in the network, it is possible that features such as $x_1 = m_X - m(\gamma\gamma\tau\tau)$ or $x_2 = m_X - a * p_T^{\gamma\gamma}$ ($a \in \mathbb{R}$) could be created. Then, the network could give high scores to events where x_1 and x_2 are close to zero.

Another advantage to pNNs is their ability to interpolate between mass points. Unlike Eq. 5.1, the function provided by a pNN is defined for all values for m_X . It is therefore possible that a pNN can be used to discriminate for mass points that were not seen during training. Consider training on $m_X = 300$ and 400 GeV. A pNN will learn how to discriminate for those two masses, and it could “guess” that to discriminate at $m_X = 350$ GeV it must apply logic that is somewhere in between that used for 300 and 400 GeV. This concept is illustrated in Fig. 5.15.

To justify their use in this analysis, a pNN must fulfil two criteria:

1. Good performance at every mass point: when evaluating on individual mass points, the pNN performs similarly to using a set of NNs where each NN is trained only on a single mass point (as in approach 1),
2. Good performance at interpolated mass points: the pNN is able to discriminate well at mass points which it has not been trained on and that are in the ranges of masses used during training.

In Section 5.5.3, tests are devised to investigate this and the degree to which the pNNs meet these criteria is presented.

For the NMSSM models, the pNNs must be parametric in m_X and m_Y . This is done by introducing both m_X and m_Y as training features instead of just m_X , and following the same techniques.

5.5.3 Training and Performance

The pNN is implemented in PYTORCH [114] and is a multilayer perceptron network of 3 layers of 50 nodes activated by exponential linear unit (ELU) functions with a dropout probability of 0.05. A schematic is given in Fig. 5.16. This architecture, was chosen after a grid search of the hyperparameters shown in Table 5.8. This optimization procedure was performed for the $X^{(2)} \rightarrow HH$ search and the set of hyperparameters chosen were those that lead to high AUC scores for the $m_X = 260, 500$ and 1000 GeV signals. With this set of hyperparameters as a baseline, further grid searches were performed on the other searches but no significant improvement in AUC scores was found by moving away from the baseline hyperparameters so the same hyperparameters were used for all searches.

In each search: $X^{(0)} \rightarrow HH$, $X^{(2)} \rightarrow HH$, $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$, low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$, and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$, a pNN is trained using the features described in Section 5.5.1. The number of events for the background processes, and for the signal processes in each search are summarized in Table 5.9. From these, 50% of the events are reserved for testing and later for signal modelling, 40% forms a training dataset used for loss minimization, and 10% is used as a validation dataset used to track the pNN performance during training.

Number of layers	2, 3, 4
Number of nodes	25, 40, 50, 60, 75
Dropout probability	0, 0.025, 0.05, 0.075, 0.1

Table 5.8: Values of hyperparameters used in a grid search to optimize the pNN architecture.

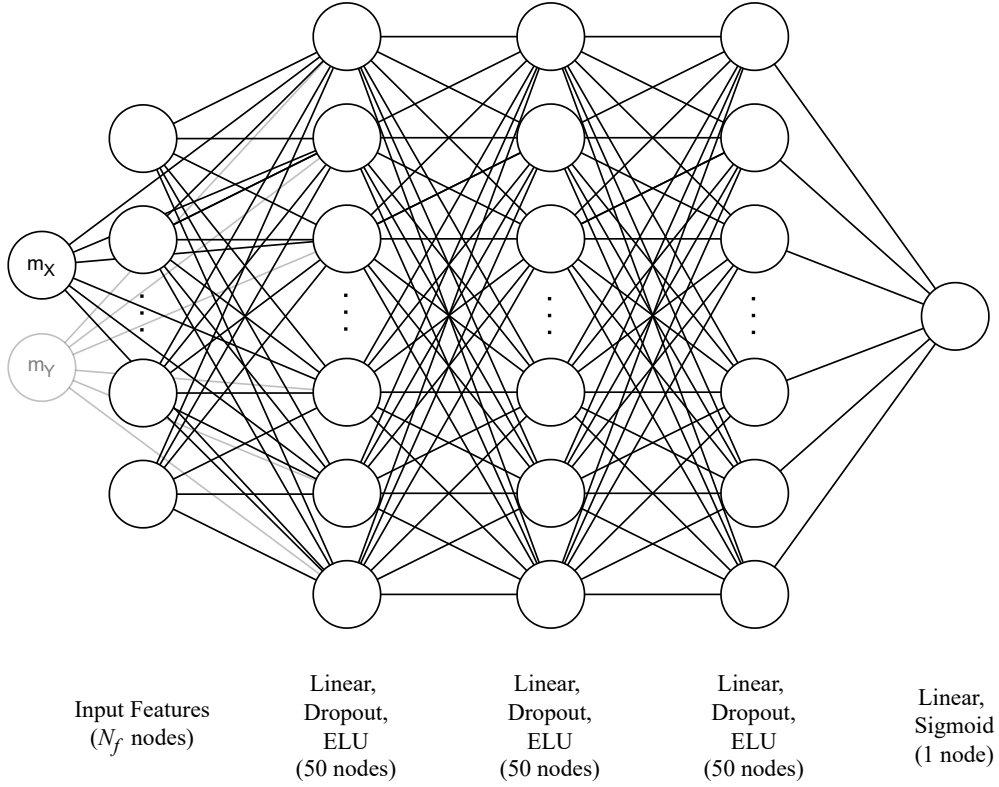


Figure 5.16: Architecture for the pNN. The parametric features, m_X and m_Y , are connected in the same way that the input features are, but are not included in the count of the input features, N_f . In the $X \rightarrow HH$ pNNs, the m_Y node and connections do not exist, which is why they are drawn with a grey colour in the diagram. For the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, $N_f = 25$, and for all other searches, $N_f = 27$.

Process	Number of events	Average per mass point
Background	274K	—
$X^{(0)} \rightarrow HH$	625K	37K
$X^{(2)} \rightarrow HH$	680K	40K
$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$	7.5M	82K
Low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	2.9M	73K
High-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	3.8M	75K

Table 5.9: Sum of the number of background and signal events across all simulated datasets.

During training, two of the background samples, γ +jets and $t\bar{t}$, are not used, because despite having relatively high numbers of expected events in this analysis, the samples have small numbers of events which means that their inclusion reduces the effective statistical power of the dataset. This leads to overtraining and studies into the $X^{(2)} \rightarrow HH$ analysis found a 5–13% worsening to the expected upper limits with their inclusion. Similarly, negatively weighted events are excluded when training because their omission was found to lead to marginal improvements in the expected limits.

The network is trained by minimizing the average binary cross entropy (BCE) loss function:

$$-\frac{1}{N} \sum_i^N w_i [y_i \log(f(\vec{x}_i; m_X^i, m_Y^i)) + (1 - y_i) \cdot \log(1 - f(\vec{x}_i; m_X^i, m_Y^i))] \quad (5.2)$$

where the sum is over N events, $f(\vec{x}; m_X, m_Y)$ is the network output, w_i is the event weight, and y is a truth label that is equal to zero for background events and one for signal events. The minimization is performed using the ADAM optimizer [115] with a batch size of 128 and an initial learning rate of 0.01 which is reduced by a factor of 0.9 at every epoch where the training loss does not decrease. These values of the training hyperparameters were found to lead to good performance and reasonably-fast convergence of the loss function. Changes around these values did not lead to substantially different performance which is why they were not included in the grid search of hyperparameters discussed previously.

Batches are constructed by randomly sampling, without replacement, 64 events from the signal component of the training dataset and 64 events from the background component. When sampling events from the signal or background components, the weights of events are taken into account and therefore, the events are treated as unweighted in the BCE loss calculation for the batch. An epoch is defined as one pass through the background component of the dataset.

The evolution of the loss on the training and validation datasets when training the

$X^{(2)} \rightarrow HH$ pNN is shown in Fig. 5.17. To gain insight into the pNN training, the training and validation datasets are split according to the signal processes where an equal number of background events are given to each split, and the losses over each split are studied. These losses for $m_X = 260, 500$ and 1000 GeV are also shown in Fig. 5.17. As expected, the overall contribution to the loss at the end of training is greatest from the 260 GeV dataset since this signal is more difficult to discriminate for than the 500 or 1000 GeV signals. Furthermore, greater improvements to the loss are seen in the 260 GeV dataset over the training duration which implies that the training was driven more by the 260 GeV signal than the 500 and 1000 GeV signals. Or in other words, there was more that had to be learned for the lower m_X signals than the higher m_X signals. The loss evolution of the pNNs for the other searches are shown in Fig. 5.18.

The training duration is determined by the validation dataset. During training, the pNN parameters are saved to disk after each epoch, and the individual validation losses for every signal process are tracked. If no relative improvement of $> 0.1\%$ is seen in any of the losses over a 15 epoch period, then the training is stopped. This ensures that the training is not stopped whilst there is a signal process for which the discrimination is still improving. Finally, the pNN parameters from the epoch with the lowest total validation loss are used.

To test the pNN's performance, datasets are created for every mass hypothesis with simulated signal events corresponding to that hypothesis and simulated background events which include all processes described in Section 5.3. These events are all taken from the testing dataset so that they are independent of the events used in training. In each dataset, the pNN is evaluated, setting the m_X and m_Y (parametric) training features to the mass hypothesis of the dataset. Equivalent datasets are created using events from the training dataset to look for signs of overtraining.

In each dataset (mass hypothesis), the signal efficiency is calculated as a function of the background efficiency to create a receiver operating characteristic (ROC) curve, where these efficiencies are calculated with respect to the events that pass preselection. The area under the ROC curve (AUC) is then calculated to give a single number that represents the performance of the pNN for that mass hypothesis. For illustrative purposes, the ROC curves are plotted reversed in this section, i.e. background efficiency against signal efficiency. However, AUC scores still refer to the area under the ROC curve as it is normally defined, i.e. signal efficiency against background efficiency.

ROC curves for the $X^{(2)} \rightarrow HH$ pNN at $m_X = 260, 300, 500$ and 1000 GeV are shown in Fig. 5.19 and AUC scores for the rest of the mass hypotheses, and for the other searches are shown in Figs. 5.20 and 5.21. AUC scores for the testing dataset are in the range 0.9 to 0.9999 and are typically higher for signals with higher m_X and lower m_Y , which

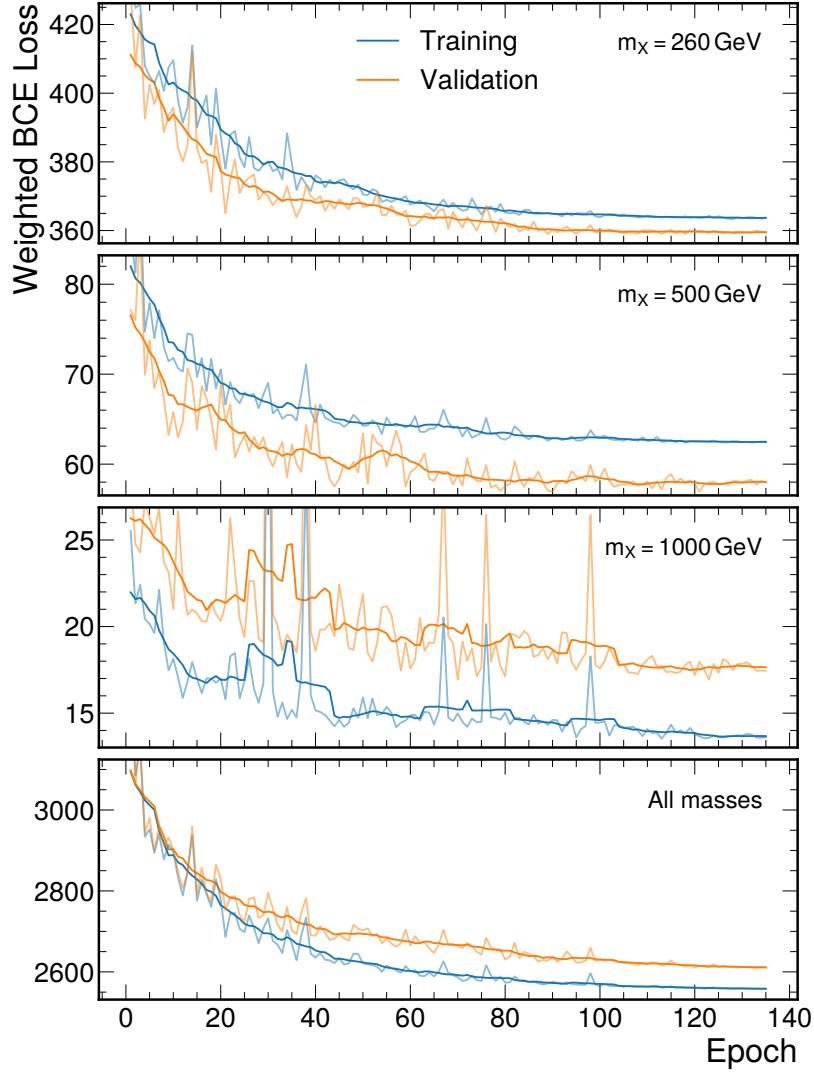


Figure 5.17: Total (not averaged) BCE loss for the $X^{(2)} \rightarrow HH$ pNN after every epoch of training for both the training and validation datasets. The datasets are split into components corresponding to the $m_X = 260, 500$, and 1000 GeV signals and the corresponding losses are shown in the top three plots, with the loss over the whole datasets shown in the bottom plot. The total normalization of the dataset, and therefore, the scale of the loss, is arbitrary, but the normalization of each signal component is equal. The first loss values shown at Epoch 1 correspond to the losses after 1 epoch of training. The opaque lines represent a rolling average in a centred window of 10 epochs of the true loss values, shown by the faint lines of the same colour.

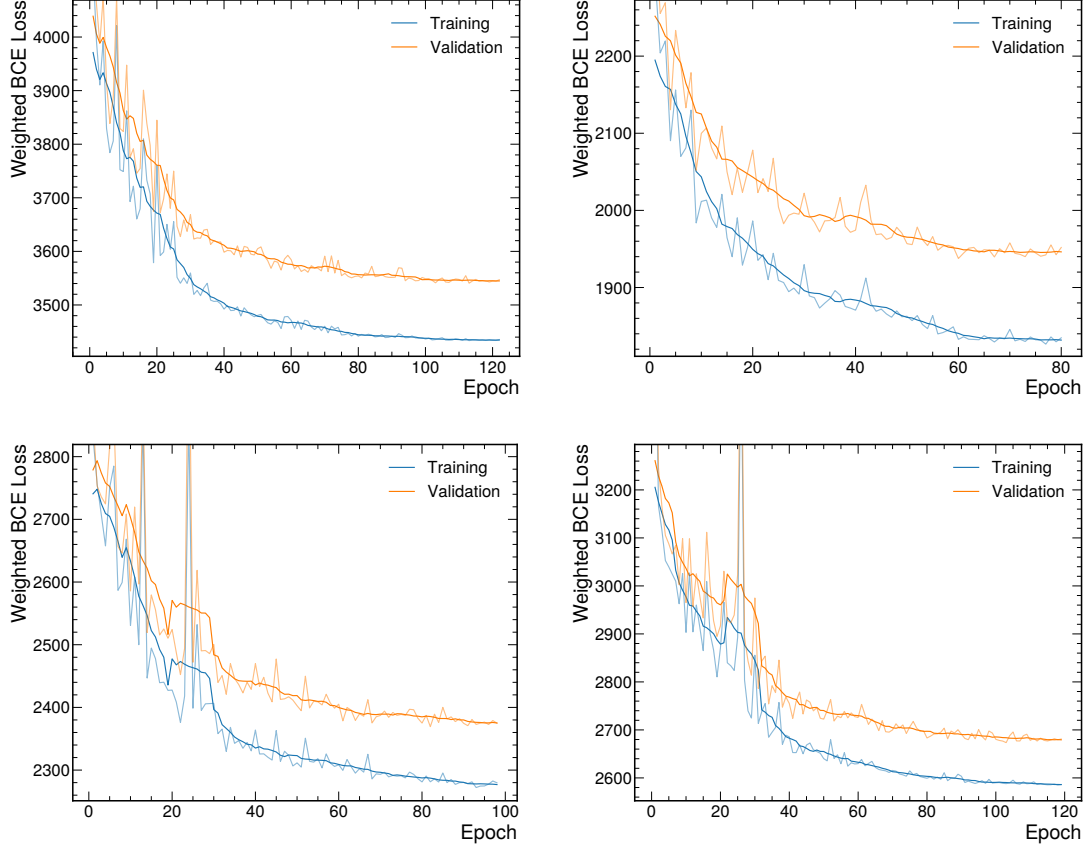


Figure 5.18: Total (not averaged) BCE loss for the $X^{(0)} \rightarrow HH$ (top-left), $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (top-right), low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom-left) and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom-right) pNNs after every epoch of training for both the training and validation datasets. The total normalization of the dataset, and therefore, the scale of the loss, is arbitrary. The first loss values shown at Epoch 1 correspond to the losses after 1 epoch of training. The opaque lines represent a rolling average in a centred window of 10 epochs of the true loss values, shown by the faint lines of the same colour.

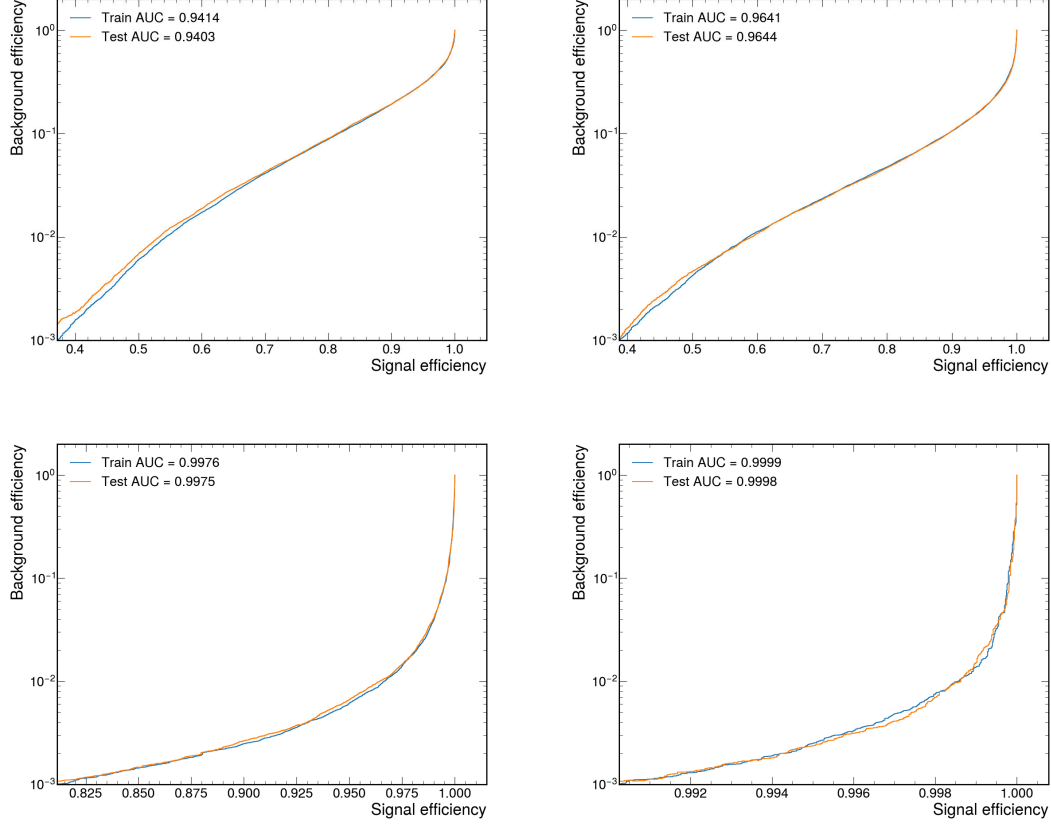


Figure 5.19: ROC curves for the $X^{(2)} \rightarrow \text{HH}$ pNN for $m_X = 260$ (top-left), $m_X = 300$ (top-right), $m_X = 500$ (bottom-left) and $m_X = 1000$ GeV (bottom-right). The curves are shown for the training and testing datasets. The efficiencies are calculated with respect to the events that pass preselection.

correspond to processes with higher boosted H and Y bosons.

In Figs. 5.19 and 5.20, the ROC curves and AUC scores are shown for the training dataset as well. The ROC curves for the $X^{(2)} \rightarrow \text{HH}$ pNN are almost indistinguishable, especially for background efficiencies of $> 1\%$ which is representative of the final categorization (see Section 5.5.4). Similarly, small differences between the train and test AUC scores can be seen across the whole m_X range for $X^{(2)} \rightarrow \text{HH}$, and for the $X^{(0)} \rightarrow \text{HH}$ pNN. This indicates that the pNNs have not been overtrained. This is also observed in the other searches. Regardless, only the events from the testing dataset are used for signal modelling to avoid any bias.

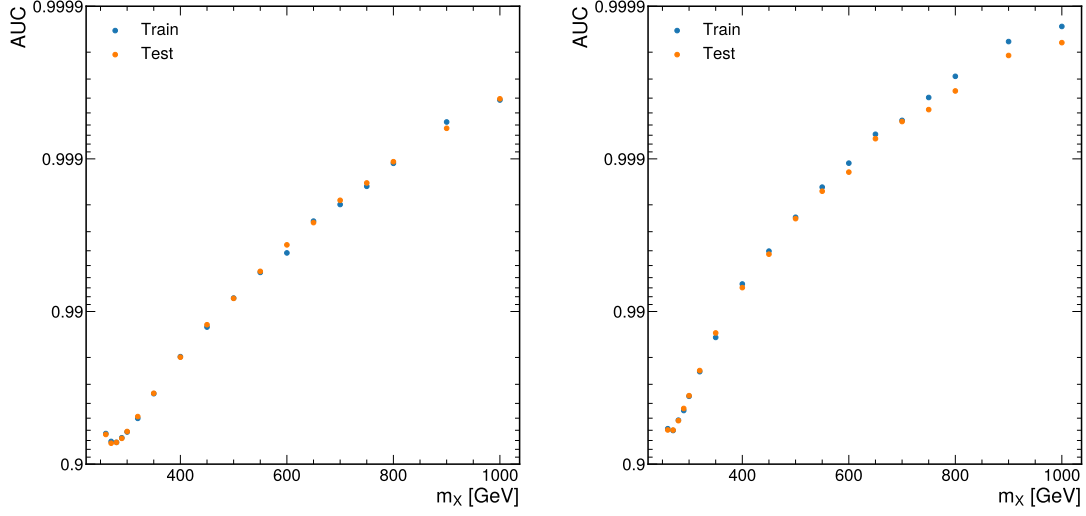


Figure 5.20: AUC scores for the $X^{(0)} \rightarrow HH$ (left) and $X^{(2)} \rightarrow HH$ (right) pNNs for as a function of m_X . The scores are shown for the training and testing datasets. The signal and background efficiencies used to create the underlying ROC curves are calculated with respect to the events that pass preselection.

In Section 5.5.2, it was stated that the pNNs must fulfil two criteria to justify their use in this analysis: 1. there is good performance at every mass point, and 2. there is good performance at interpolated mass points. To study these criteria, the following tests are performed:

1. Good performance at every mass point
 - (a) Train a pNN using all of the mass points ('all' network)
 - (b) Train a separate NN for each mass point ('only' network)
 - (c) Compare the performance between (a) and (b) where the performance is evaluated on the (only) mass point that the NN was trained on
2. Good performance at interpolated mass points
 - (a) Train a pNN using all of the mass points ('all' network)
 - (b) Train a separate pNN on all mass points except one ('skip' network)
 - (c) Compare the performance between (a) and (b) where the performance is evaluated on the excluded (skipped) mass point

The performance metric used is the signal efficiency found at a background efficiency of 1% which is representative of the final categorization in the analysis. In a search, the first test is performed at every mass point, and the second test is performed at every mass point

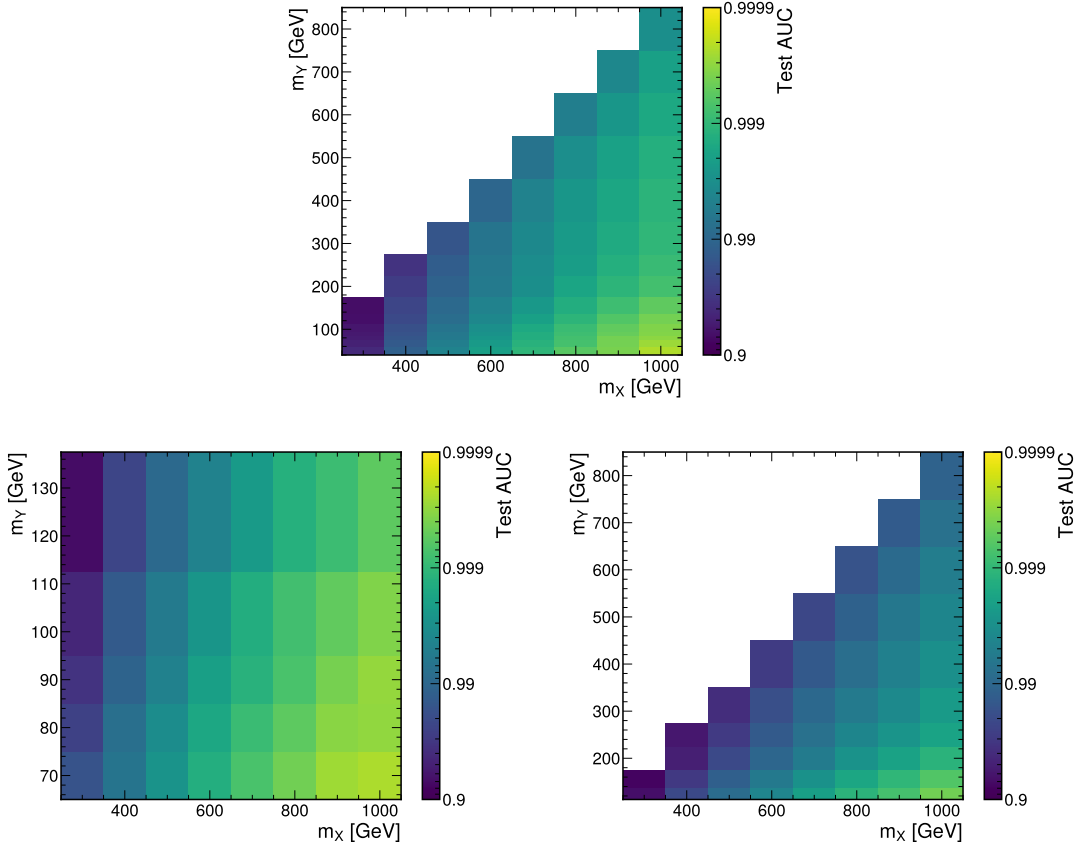


Figure 5.21: AUC scores calculated on the testing datasets for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (top), low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom-left), and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom-right) pNNs as functions of m_X and m_Y . The signal and background efficiencies used to create the underlying ROC curves are calculated with respect to the events that pass preselection.

except those at the boundaries of the mass ranges. This excludes scenarios where the pNN is asked to extrapolate to a mass point that is outside the range of masses used during training. The results of these tests are shown in Figs. 5.22 to 5.25 for the $X^{(2)} \rightarrow HH$, $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$, low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$, and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ pNNs respectively. The $X^{(0)} \rightarrow HH$ pNN was not tested but is expected to perform similarly to the $X^{(2)} \rightarrow HH$ pNN.

Across all searches, the majority of the tests show differences of less than 1%. The ‘only’ tests tend to show larger differences where the biggest differences are typically found at lower m_X and higher m_Y , being up to 5% and 2% for the ‘only’ and ‘skip’ comparisons respectively. This trend in m_X and m_Y is not entirely unexpected. As evident from the shape of the training features (Section 5.5.1), the evolution of the loss for different masses (Fig. 5.17), and the AUC scores (Figs. 5.20 and 5.21), signals at lower m_X and higher m_Y present a more complex discrimination problem. Therefore, it is unsurprising that dedicated networks (the ‘only’ networks) show greater chances for improvement, and that it is more difficult to interpolate the behaviour of the network accurately.

The majority of the tests show differences of less than 1%, with a few losses up to 5%. Given the benefits of using a pNN, the small number of larger losses are considered acceptable and pNNs are chosen as the final discriminators for the analysis. In future analyses, the losses could be minimized by splitting up the kinematic regime and training more specialized networks, or by investigating more complex architectures for the pNN.

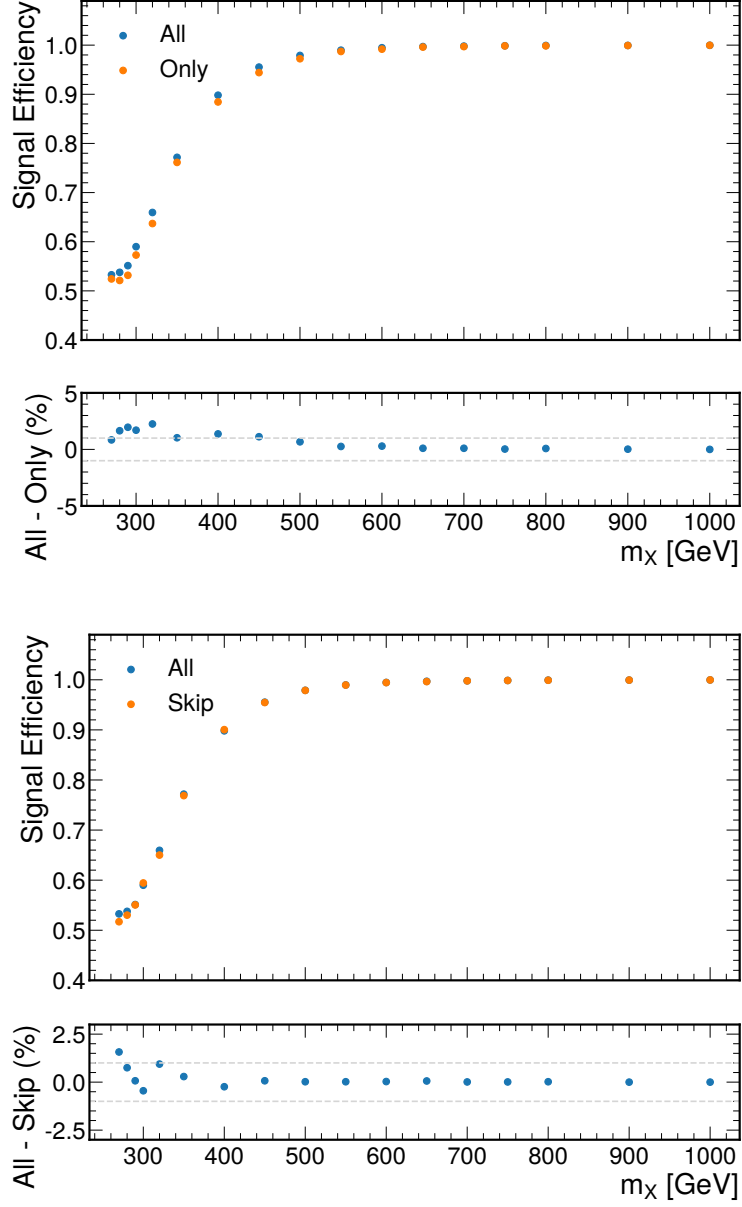


Figure 5.22: For the $X^{(2)} \rightarrow HH$ search, results from the tests designed to check whether the pNN performs well at every mass point (top) and at interpolated mass points (bottom). Signal efficiencies correspond to a background efficiency of 1%. The ‘all’ efficiencies correspond to a pNN trained on all mass points. The ‘only’ efficiencies correspond to a NN trained only on the mass point which the efficiency is quoted for. The ‘skip’ efficiencies correspond to a pNN trained on all mass points except for the mass point where the efficiency is quoted for. The grey dashed lines in the bottom half of each plot correspond to differences of 1%.

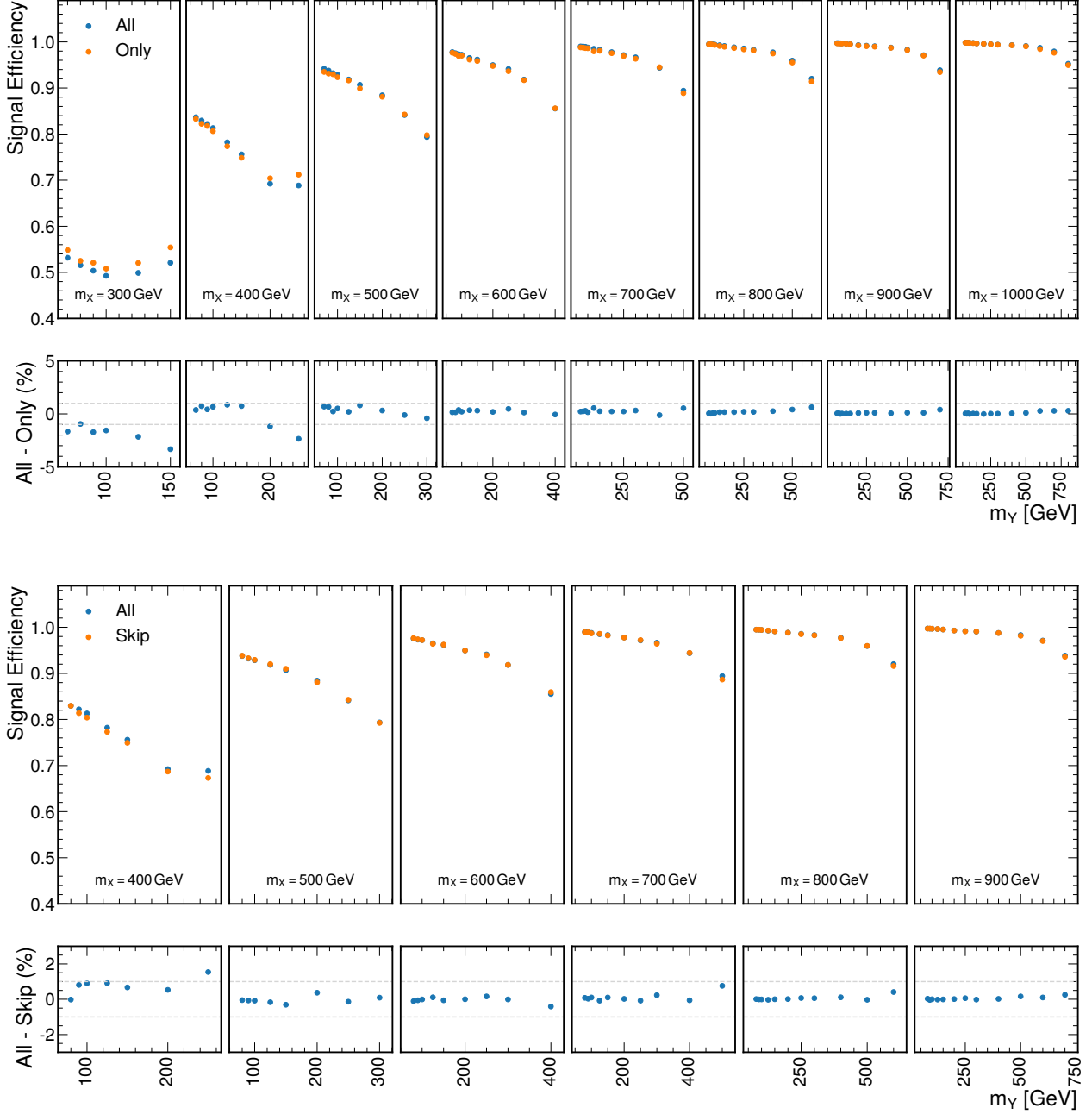


Figure 5.23: For the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search, results from the tests designed to check whether the pNN performs well at every mass point (top) and at interpolated mass points (bottom). Signal efficiencies correspond to a background efficiency of 1%. The ‘all’ efficiencies correspond to a pNN trained on all mass points. The ‘only’ efficiencies correspond to a NN trained only on the mass point which the efficiency is quoted for. The ‘skip’ efficiencies correspond to a pNN trained on all mass points except for the mass point where the efficiency is quoted for. The grey dashed lines in the bottom half of each plot correspond to differences of 1%.

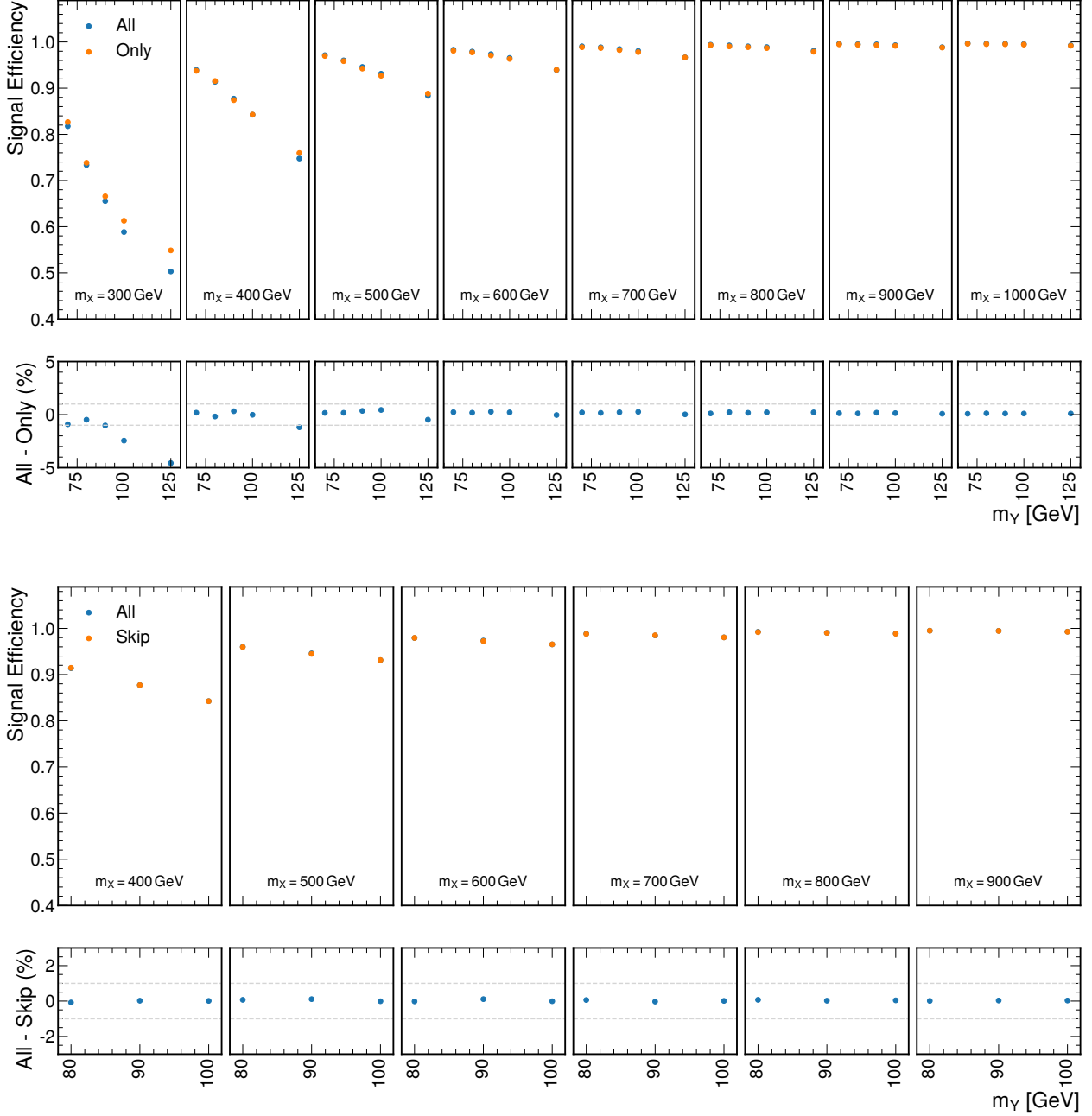


Figure 5.24: For the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, results from the tests designed to check whether the pNN performs well at every mass point (top) and at interpolated mass points (bottom). Signal efficiencies correspond to a background efficiency of 1%. The ‘all’ efficiencies correspond to a pNN trained on all mass points. The ‘only’ efficiencies correspond to a NN trained only on the mass point which the efficiency is quoted for. The ‘skip’ efficiencies correspond to a pNN trained on all mass points except for the mass point where the efficiency is quoted for. The grey dashed lines in the bottom half of each plot correspond to differences of 1%.

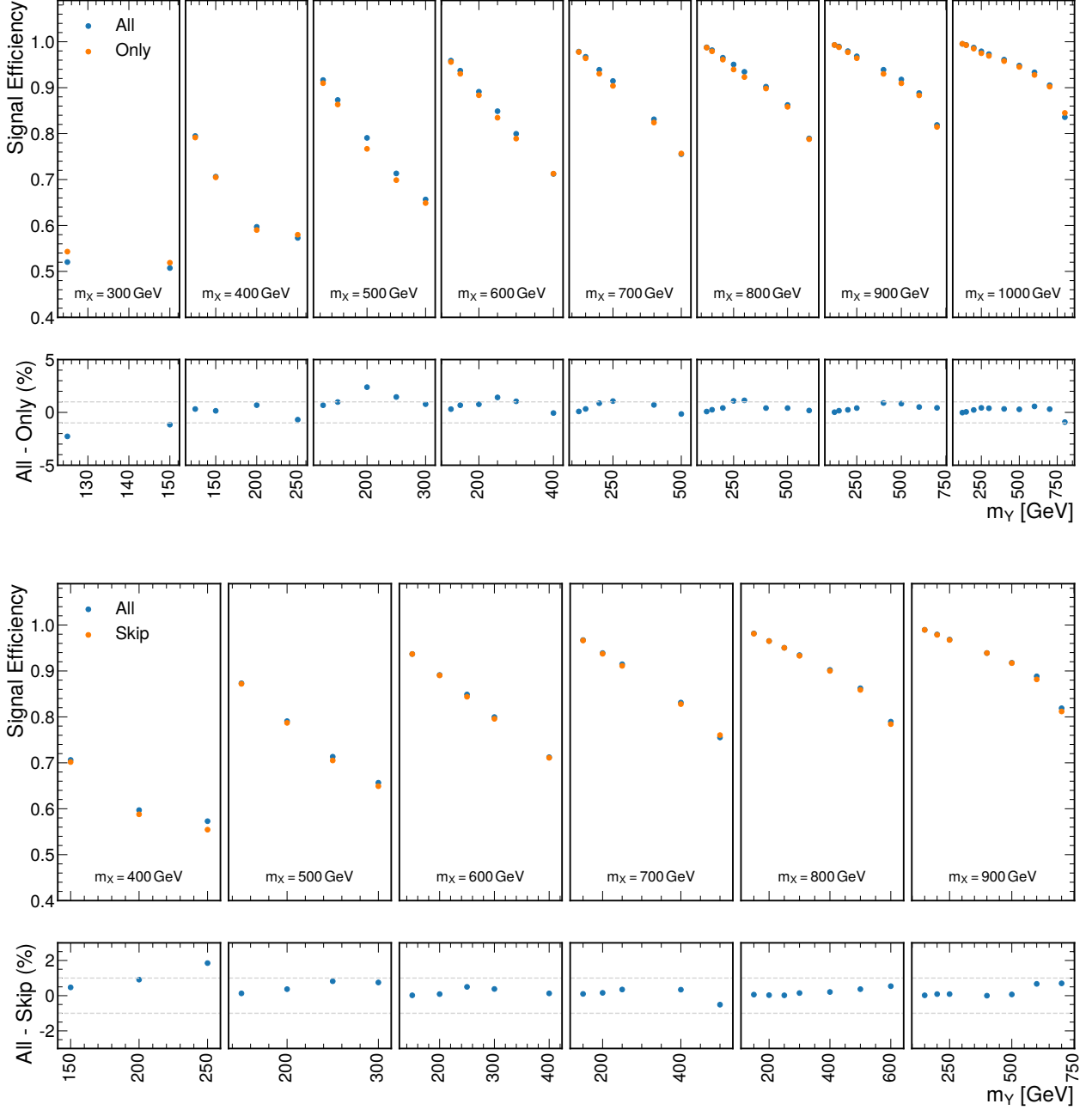


Figure 5.25: For the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, results from the tests designed to check whether the pNN performs well at every mass point (top) and at interpolated mass points (bottom). Signal efficiencies correspond to a background efficiency of 1%. The ‘all’ efficiencies correspond to a pNN trained on all mass points. The ‘only’ efficiencies correspond to a NN trained only on the mass point which the efficiency is quoted for. The ‘skip’ efficiencies correspond to a pNN trained on all mass points except for the mass point where the efficiency is quoted for. The grey dashed lines in the bottom half of each plot correspond to differences of 1%.

5.5.4 Category Optimization

For a pair of masses, $(m_X, m_Y) = (m_X^i, m_Y^j)$, a set of analysis categories are defined based upon $f(\vec{x}; m_X = m^i, m_Y = m^j)$, where $f(\vec{x}; m_X, m_Y)$ is the pNN output score. Before defining these categories, the raw output score from the network is transformed such that the background MC from the test dataset is flat in the transformed score. The transformed score is denoted $\tilde{f}(\vec{x}; m_X, m_Y)$. Like in training, the $\gamma + \text{jets}$ and $t\bar{t} + \text{jets}$ datasets are not used for the transformation. This transformation is motivated by the interpolation that is necessary to create signal models for the intermediate mass points. When placing a cut on $\tilde{f}(\vec{x}; m_X, m_Y)$, it is equivalent to placing a cut on the *training* background efficiency, and at a fixed background efficiency, the signal efficiency is expected to be a smooth function of m_X and m_Y . Figs. 5.20 to 5.25 already suggest that this is indeed the case.

Distributions of the transformed output score for a selection of mass points are shown for the $X \rightarrow \text{HH}$ and $X \rightarrow \text{YH}$ searches in Figs. 5.26 and 5.27. The background distributions are almost flat, but not exactly since the $\gamma + \text{jets}$ and $t\bar{t} + \text{jets}$ datasets, which are not in the transformation, are included in these plots. The signal distributions peak sharply at $\tilde{f}(\vec{x}; m_X, m_Y) = 1$, meaning that the optimized categories can be expected to be defined at high values of $\tilde{f}(\vec{x}; m_X, m_Y)$.

To optimize the category boundaries, a metric for the performance of a set of boundaries must be chosen. Given that the normalization of the signal is not known a priori, a metric independent of the normalization is desired, and here, the expected upper limit at the 95% CL is used. The procedure used to calculate this limit is a simplified and computationally faster version of the final procedure used in the analysis. For a given nominal mass point, the upper limit is calculated as follows:

1. In each analysis category:
 - (a) Fit an exponential function to the $m_{\gamma\gamma}$ distribution in the nonresonant MC samples.
 - (b) Define a signal region as the $\pm 1\sigma$ window centred around $\hat{m}_{\gamma\gamma}$, where σ and $\hat{m}_{\gamma\gamma}$ are the standard deviation and mean values respectively of $m_{\gamma\gamma}$ in the signal MC events.
 - (c) Determine the expected number of background events in the signal region by integrating the fitted exponential over that range.
 - (d) Determine the expected number of signal events in the signal region by summing the weights of the signal MC events that in the region.
2. Use the expected number of signal and background events in the signal regions to perform a counting experiment and determine the upper limit according to the

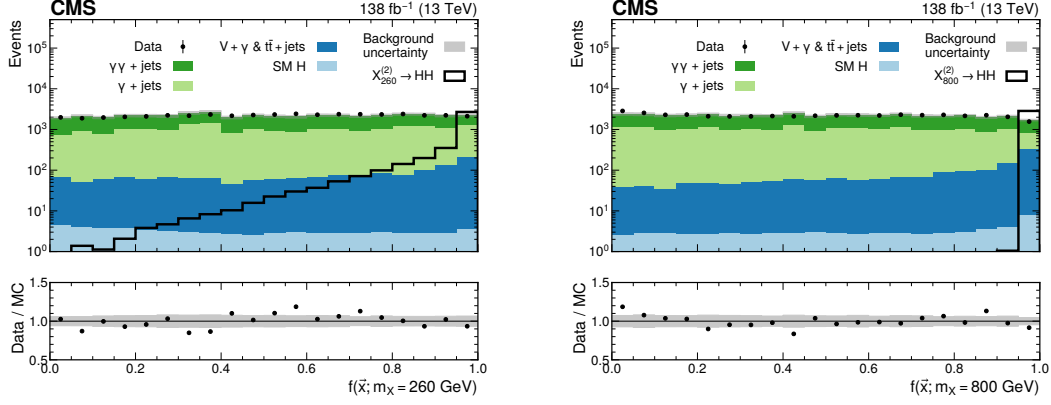


Figure 5.26: Transformed pNN output distribution in the $X^{(2)} \rightarrow HH$ search, evaluated at $m_X = 260$ GeV (left) and $m_X = 800$ GeV (right). The filled histograms represent the background simulation, and the corresponding statistical uncertainty is shown by the grey bands. The data are shown by black points, and the signal is shown by a black line. The background MC simulation is normalized to data and the signal is normalized to an arbitrary cross section for representation purposes.

techniques described in Chapter 4 using an asymptotic approximation for the distribution of the test statistic [85].

The expected upper limit at an intermediate mass point cannot be determined in the same way because there is no signal MC for those points. The signal yields could instead be interpolated from those at the nominal mass points, but this would complicate the categorization optimization procedure. Instead, a procedure which used only the nominal mass points, but is still applicable, and optimal (or close to) at the intermediate mass points was devised.

At first, a categorization procedure was investigated where the same boundaries on $\tilde{f}(\vec{x}; m_X, m_Y)$ were used for every mass point. This was tested on the $X^{(2)} \rightarrow HH$ search, and the boundary values were determined by a grid search of 3 categories that optimized the expected limit at $m_X = 280$ GeV. This strategy may bias itself towards the $m_X = 280$ GeV mass point, so the resulting expected limits at other masses were compared to those found by performing the grid search at those masses instead. The comparison, shown in Fig. 5.28, finds up to 25% differences in the expected upper limit and this was considered an unacceptable loss in sensitivity and therefore, alternative procedures were considered.

During the grid search, it was imposed that the minimum number of expected background events be 10 in each category. This requirement is made to ensure a bias of less than 20% is induced in the observed significances because of the choices of background functions used in the nonresonant background modelling. It was noticed that

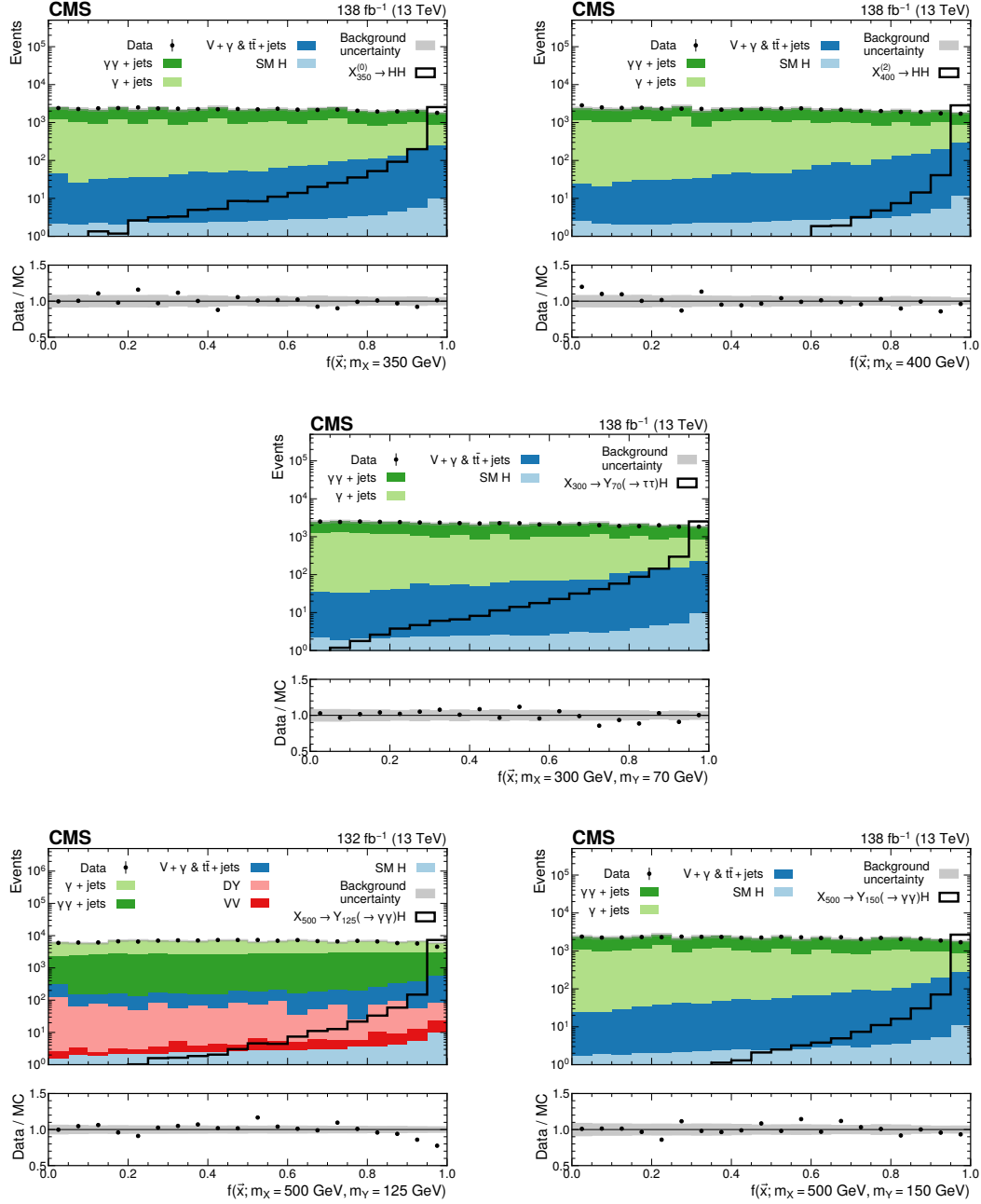


Figure 5.27: Transformed pNN output distribution in the $X^{(0)} \rightarrow HH$ (upper left), $X^{(2)} \rightarrow HH$ (upper right), $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (middle), low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (lower left) and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (lower right) searches. The pNNs are evaluated at the mass points where the largest excess with respect to the background-only hypothesis is observed. The filled histograms represent the background simulation, and the corresponding statistical uncertainty is shown by the grey bands. The data are shown by black points, and the signal is shown by a black line. The background MC simulation is normalized to data and the signal is normalized to an arbitrary cross section for representation purposes.

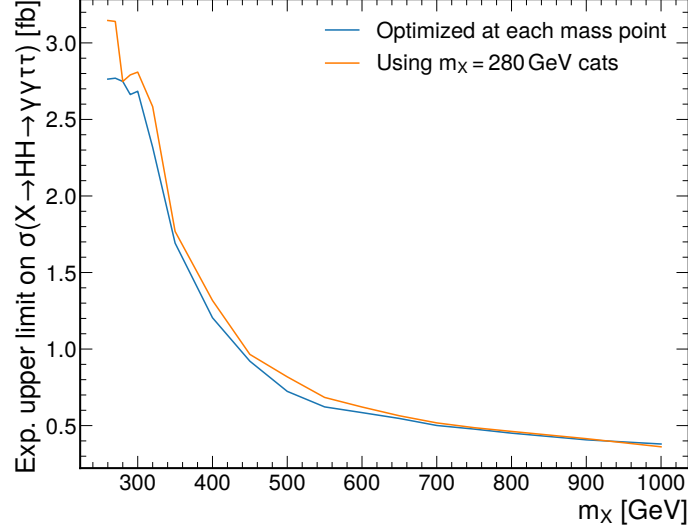


Figure 5.28: Comparison between the 95% CL expected upper limit on $\sigma(\text{pp} \rightarrow \text{X} \rightarrow \text{HH} \rightarrow \gamma\gamma\tau\tau)$ reached when using the boundaries optimized for each mass point and when using the boundaries found for $m_X = 280$ GeV for all mass points.

the grid search would routinely hit this boundary, defining the highest-scoring categories to have 10 expected background events, and then defining the last category such that the remaining signal was accepted. This was observed regardless of which mass point was being optimized for. This motivated defining the analysis categories in terms of the number of expected background events, which can be later converted into boundaries on $\tilde{f}(\vec{x}; m_X, m_Y)$ by using the background MC events, where this conversion is possible for both nominal and intermediate mass points. The procedure is as follows:

1. Order the background MC events by their $\tilde{f}(\vec{x}; m_X, m_Y)$ score.
2. Begin by defining a category with the 10 highest scoring events.
3. For every nominal mass point, find the corresponding boundary in $\tilde{f}(\vec{x}; m_X, m_Y)$, and then calculate the expected upper limit using the simplified procedure described previously.
4. Consider a new category, with the next 10 highest scoring background events, and recalculate the expected limit at every nominal mass point.
5. If adding this category improves any of the limits by $\geq 1\%$, keep this new category and consider further categories of 10 events.
6. If the improvement is not great enough, consider adding 20 events instead. If that is

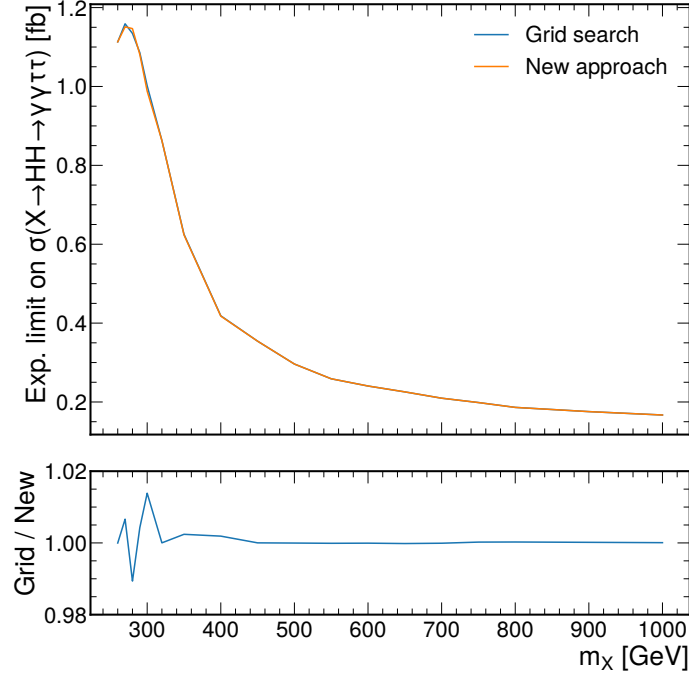


Figure 5.29: Comparison between the 95% CL expected upper limit on $\sigma(pp \rightarrow X \rightarrow HH \rightarrow \gamma\gamma\tau\tau)$ found with the new categorization approach (in orange) and those found via a grid search performed at for each mass point individually.

still not enough, consider adding 40 and keep multiplying the number by two until either a significant improvement is found or all of the background MC is exhausted.

Once again, this procedure may lead to less sensitive limits than a grid search approach that optimizes the limit at every mass point. So, expected limits at the nominal mass points in the $X^{(2)} \rightarrow HH$ search were compared between the new categorization approach and a grid search of 4 categories optimized at every nominal mass point. A grid search of 4 categories was chosen since the addition of extra categories was found to only lead to sub-percent improvements in the expected limits. The results of the comparison are shown in Fig. 5.29. At most a 1% loss in the expected limit is found using the new approach which was adopted for the final results.

The new procedure is applied to all searches individually, and the resulting category definitions are shown in Table 5.10. All searches have 4 categories with 10 expected background events each, followed by a category of 20, and then a category of 80. In the $X \rightarrow YH$ searches, a final category of 320 expected background events is also defined. After the preselection for the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches, there are about 44K data events. Assuming that the majority of these events are background, this means the

Category	0	1	2	3	4	5	6
$X^{(0)} \rightarrow HH$	10	10	10	10	20	80	—
$X^{(2)} \rightarrow HH$	10	10	10	10	20	80	—
$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$	10	10	10	10	20	80	320
Low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	10	10	10	10	20	80	320
High-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	10	10	10	10	20	80	320

Table 5.10: Number of expected background events in each category for each analysis. The categories' boundaries in terms of $\tilde{f}(\vec{x}; m_X)$ are defined such that these numbers are satisfied and the top scoring events belong to category 0. There is one last category in addition to the ones shown here that contains the remainder of the events.

categorization in these searches, summing over all categories, corresponds to a background efficiency of 0.3% or 1.0% for the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches respectively.

5.5.5 Sculpting of the Diphoton Mass Distribution

In this analysis, the nonresonant background model is derived from data assuming that the shape of the $m_{\gamma\gamma}$ distribution is smoothly falling, and can be well described by functions such as exponential and power law functions. In Fig. 5.30, the distribution of the $m_{\gamma\gamma}$ distribution after the standard preselection (except low-mass $Y \rightarrow \gamma\gamma$) is shown for nonresonant background MC with a fit of a power law function. The fit function describes this distribution well, so at this stage, the nonresonant background modelling is valid. However, after applying the pNN selection, there is a risk that the shape of the $m_{\gamma\gamma}$ distribution will change (be *sculpted*) and invalidate the modelling. This is possible because variables that are correlated with the $m_{\gamma\gamma}$, such as photon $p_T/m_{\gamma\gamma}$ are included in the pNN as a training features.

To investigate the presence of such an effect, binned χ^2 fits of power law functions are performed to the $m_{\gamma\gamma}$ distribution of the nonresonant background MC after applying different pNN selections. These tests are applied at every nominal mass point and in every search. Then, the tests that show poor goodness of fit are investigated further.

First, the nonresonant background MC that is fitted includes the γ +jets, $\gamma\gamma$ +jets, $V + \gamma$, $t\bar{t}$, $t\bar{t} + \gamma$ and $t\bar{t} + \gamma\gamma$ samples, and two pNN selections: $f(\vec{x}; m_X, m_Y) > z_i$, are applied, which correspond to a relative statistical uncertainty in the yield of the remaining background of 5% and 10%. The selection is chosen in this way to ensure that there are enough simulated events to study the shape of the $m_{\gamma\gamma}$ distribution.

As an example, these two fits in $X^{(2)} \rightarrow HH$ search for $m_X = 550$ GeV are shown in Fig. 5.31. In both cases, the power law function describes the distribution well, with p-values of 0.47 and 0.29 for the 10% and 5% uncertainty selections respectively. Given that the γ +jets dataset has poor statistics, and its inclusion reduces the overall statistical

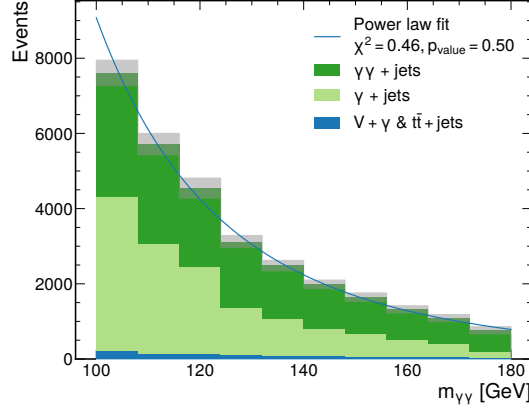


Figure 5.30: Distribution of the diphoton mass ($m_{\gamma\gamma}$) in nonresonant background samples after the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ preselection. A power law function resulting from a χ^2 fit to the MC is shown by the blue line. Statistical uncertainties in the background MC are shown by the shaded bands.

power of the dataset, further fits are performed where the γ +jets dataset is removed. Whilst the background MC will therefore be less representative of the data, it leads to tighter selections of the pNN score. For the same reason, more fits are performed where only the $\gamma\gamma$ +jets dataset is included since it has the greatest statistical power. In all cases, no significant evidence of sculpting is found.

In all the searches, the χ^2 fit is performed with 10 bins. In the $X \rightarrow HH$ searches and the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search, the χ^2 fits are performed in the range 100–180 GeV. In the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, the range is 65–150 GeV. In the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, the range is a 200 GeV window centred around m_Y , with a lower bound of 100 GeV.

Initially, this procedure highlighted an issue with the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, when $m(\gamma\gamma\tau\tau)$ and $m(\tau\tau\gamma_1)$ were included as training features of the pNN. The same fits described above for the $X^{(2)} \rightarrow HH$ search at $m_X = 550$ GeV but for the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search at $(m_X, m_Y) = (400, 200)$ GeV are shown in Fig. 5.32. As the pNN selection becomes tighter, the presence of sculpting is clear, with a peak forming at $m_Y = 125$ GeV. This is also reflected in the higher χ^2 values. After removing the $m(\gamma\gamma\tau\tau)$ and $m(\tau\tau\gamma_1)$ features from training, the same plots, shown in Fig. 5.33, show that the sculpting is no longer present. Sculpting that was present in other mass points was similarly removed. The same procedure was then applied to the other searches and no sculpting was observed.

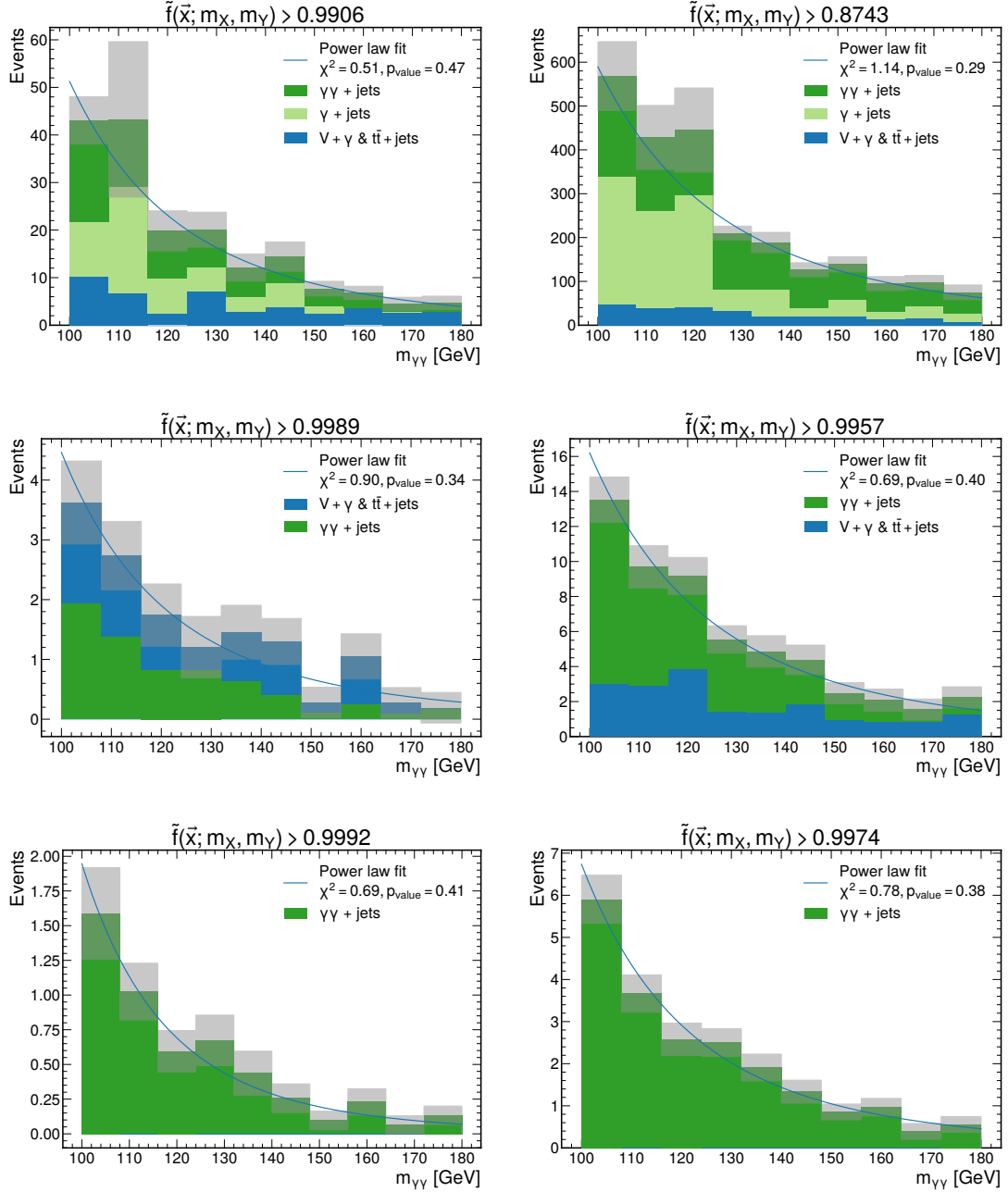


Figure 5.31: Distributions of $m_{\gamma\gamma}$ in nonresonant background samples in the $X^{(2)} \rightarrow HH$ search with different pNN selections applied targeting $m_X = 550$ GeV. In the top plots, all nonresonant background samples are included. In the middle plots, the γ +jets dataset is removed. In the bottom plots, only the $\gamma\gamma$ +jets dataset is included. The pNN boundary is given by the value that corresponds to a relative statistical uncertainty in the yield of the remaining background of 10% and 5% for the left and right plots respectively. Power law functions resulting from χ^2 fits to the MC are shown by the blue lines. Statistical uncertainties in the background MC are shown by the shaded bands.

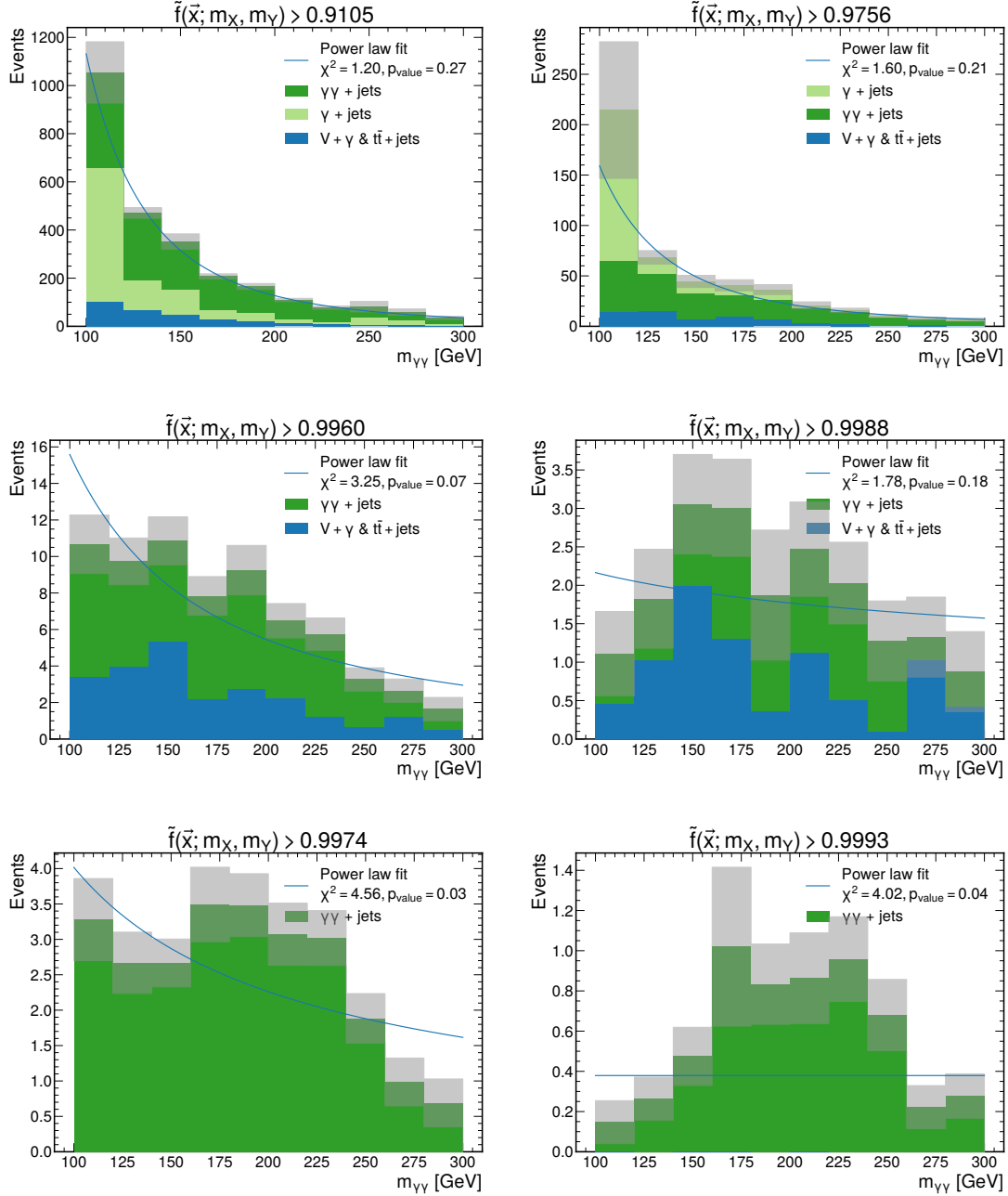


Figure 5.32: When including $m(\gamma\gamma\tau\tau)$ and $m(\tau\tau\gamma_1)$ in the pNN training, distributions of $m_{\gamma\gamma}$ in nonresonant background samples in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search with different pNN selections applied targetting $(m_X, m_Y) = (400, 200)$ GeV. In the top plots, all nonresonant background samples are included. In the middle plots, the γ +jets dataset is removed. In the bottom plots, only the $\gamma\gamma$ +jets dataset is included. The pNN boundary is given by the value that corresponds to a relative statistical uncertainty in the yield of the remaining background of 10% and 5% for the left and right plots respectively. Power law functions resulting from χ^2 fits to the MC are shown by the blue lines. Statistical uncertainties in the background MC are shown by the shaded bands.

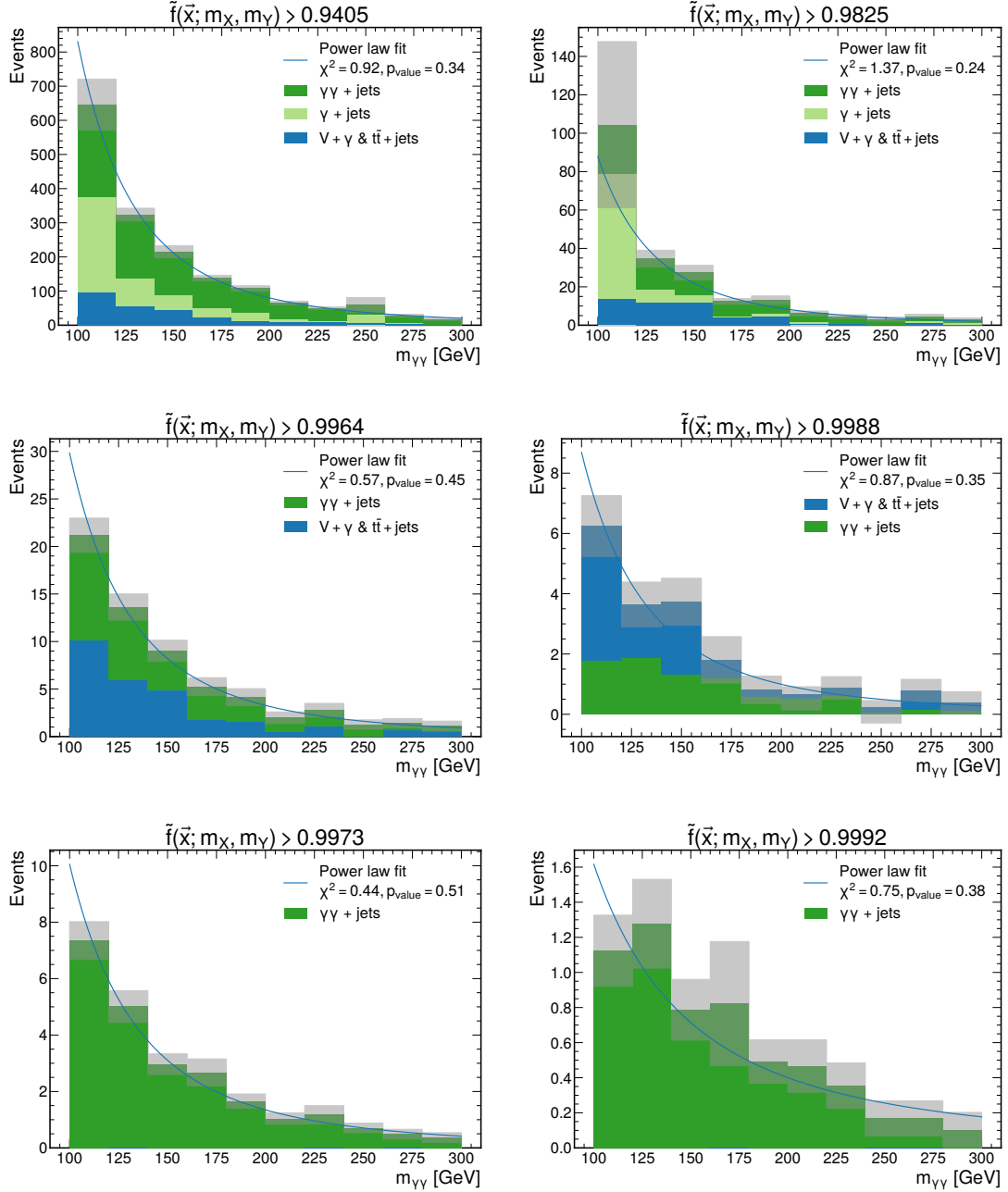


Figure 5.33: When excluding $m(\gamma\gamma\tau\tau)$ and $m(\tau\tau\gamma_1)$ in the pNN training, distributions of $m_{\gamma\gamma}$ in nonresonant background samples in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search with different pNN selections applied targetting $(m_X, m_Y) = (400, 200)$ GeV. In the top plots, all nonresonant background samples are included. In the middle plots, the γ +jets dataset is removed. In the bottom plots, only the $\gamma\gamma$ +jets dataset is included. The pNN boundary is given by the value that corresponds to a relative statistical uncertainty in the yield of the remaining background of 10% and 5% for the left and right plots respectively. Power law functions resulting from χ^2 fits to the MC are shown by the blue lines. Statistical uncertainties in the background MC are shown by the shaded bands.

5.6 Search Granularity

If one searched only at the nominal mass points, i.e. those defined by the generated MC samples (Fig. 5.4), then it is possible that a signal excess at a mass point between nominal mass points could be missed. This is because the analysis categories are designed to target a particular mass point, and could have a selection that is sensitive only to a narrow range of masses around the targeted mass point. The most direct way to understand this is to consider the $m(\gamma\gamma\tau\tau)/m_{\gamma\gamma}$ variable that is an input to the pNN selection. This distribution is shown in Fig. 5.14 for simulated background and signal events in the $X^{(2)} \rightarrow HH$ search. For $m_X = 260$ GeV, a highly discriminating cut would be a window around the peak of the signal distribution, however this would almost completely exclude the 500 GeV and 1000 GeV signals also shown in the plot.

In practice, the pNN selection is more complex than this, and there are more nominal mass points than shown in the plot, but the principle remains the same. To ensure that there is good sensitivity to all mass points in the search region, analysis categories are created at intermediate mass points, and the corresponding signals are also searched for. The granularity at which these intermediate mass points are chosen is determined by a procedure that is described in the rest of this section.

The sensitivity to signals in between nominal mass points is studied by evaluating the expected upper limit for a nominal mass point when using the analysis categories designed for neighbouring mass points. Examples of this in the $X^{(2)} \rightarrow HH$ search are shown in Fig. 5.34. For the $m_X = 350$ GeV signal, the categories that are designed for $m_X = 350$ lead to the lowest expected upper limit, as expected. However, if using the categories designed for $m_X = 320$ GeV, the limit is worse by about 50%. Assuming that the reduction in sensitivity is linear with the distance in m_X between a signal and a nominal mass point, this suggests that a signal between $m_X = 320$ and 350 GeV may have a reduced sensitivity of up to 25%. By introducing an intermediate mass point at 335 GeV, the maximum reduction in sensitivity to signals in the 320–350 GeV range is estimated to be 12.5%, or if two points are added at 330 and 340 GeV, the maximum reduction in sensitivity is estimated to be 8.3%.

In the $X \rightarrow HH$ searches, this is how the granularity of the search is determined, where the number of points added to intervals between nominal mass points is such that the estimated loss in sensitivity is at most 10%. In the 320–350 GeV example, one could have also considered how the 320 GeV signal performed when using the 350 GeV categories, which could have lead to a greater or lesser worsening in sensitivity than seen when testing the 350 GeV signal with the 320 GeV categories. Therefore, for a given interval, the granularity procedure is performed both ways, and the greatest number of

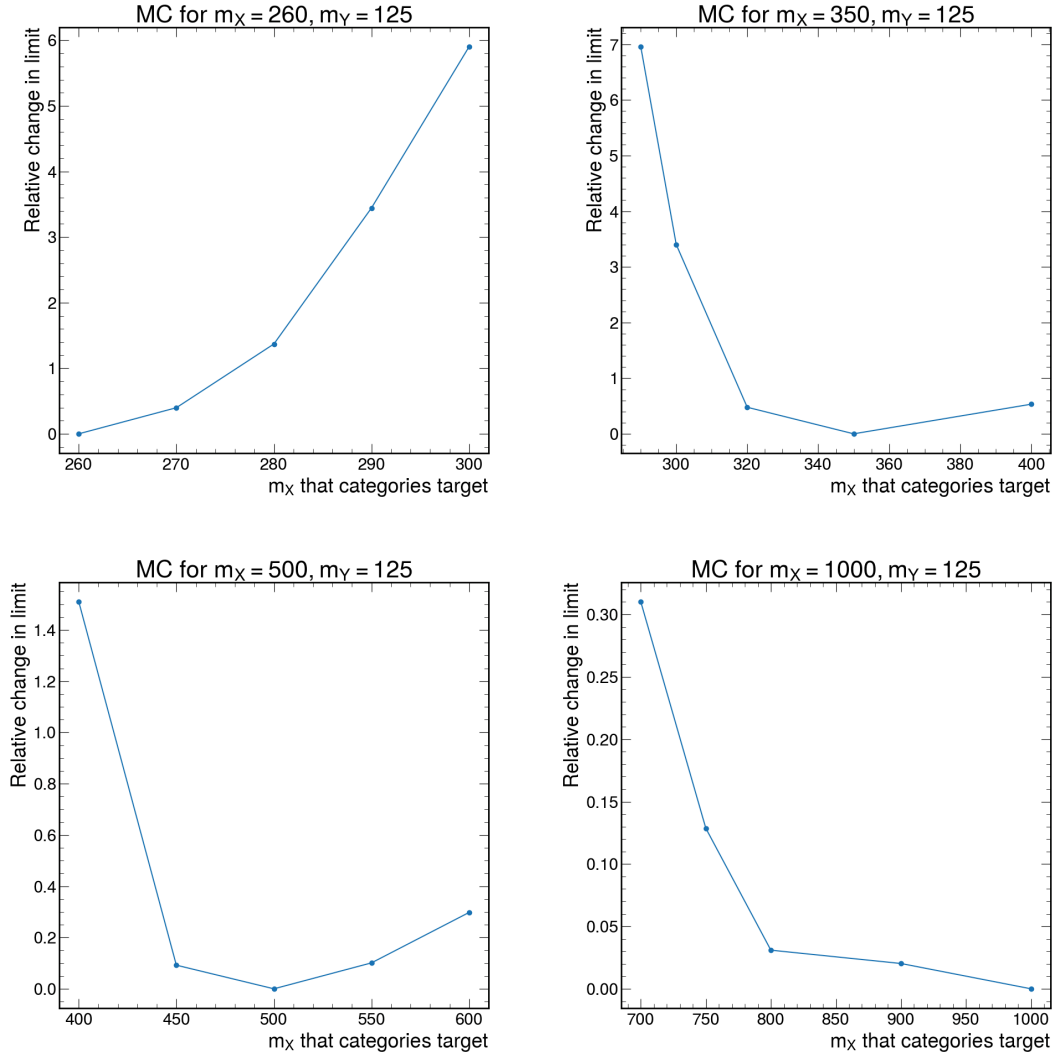


Figure 5.34: Relative changes in expected upper limits found for the $m_X = 260$ (top-left), 350 (top-right), 500 (bottom-left) and 1000 GeV (bottom-right) $X^{(2)} \rightarrow HH$ signals when using categories designed to target different mass points. The changes are shown for mass points close to the signal mass, and the change is with respect to the limit found with the categories targeting the signal mass.

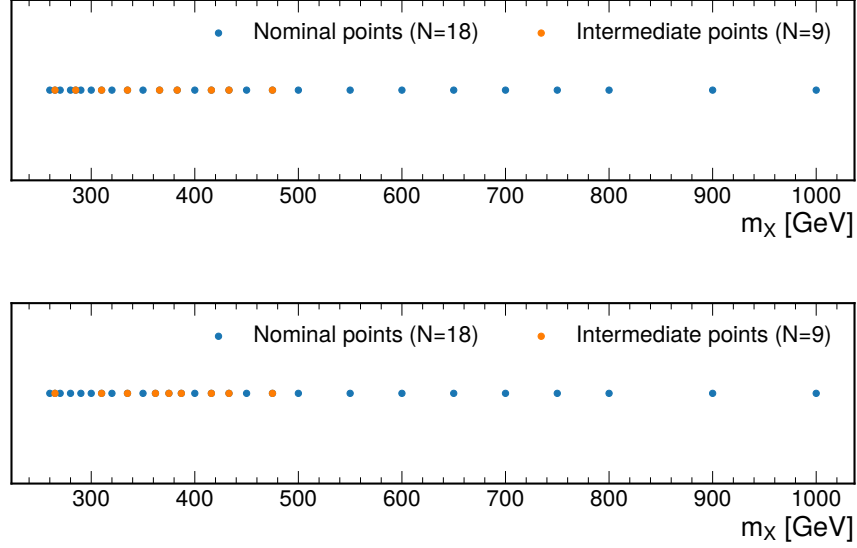


Figure 5.35: Search granularity for the $X \rightarrow HH$ searches. Shown are the nominal mass points (blue) and the intermediate mass points (orange) for the $X^{(0)} \rightarrow HH$ (top) and $X^{(2)} \rightarrow HH$ (bottom) searches.

intermediate points determined to be needed is used. The final granularities for the $X^{(0)} \rightarrow HH$ and $X^{(2)} \rightarrow HH$ searches were determined separately and are shown in Fig. 5.35, where 9 intermediate mass points are added to each search. These points are added primarily at low values of m_X where the largest changes in sensitivity are found when using neighbouring categories. At higher values of m_X , such as 500 and 1000 GeV, smaller changes in expected limits are found when using neighbouring categories (Fig. 5.34) and therefore fewer intermediate mass points are needed.

In the $X \rightarrow YH$ searches, a hollowed-out grid is first formed by applying the same procedure as in the $X \rightarrow HH$ searches along the m_X and m_Y axes independently, in slices of m_Y and m_X respectively, where the slices correspond to nominal mass points. Then, the grid must be filled in. For every mass point, (m_X^i, m_Y^j) , that has an intermediate value of m_Y^j , the granularity across the m_X axis between m_X^i and the next nominal m_X point is given by the granularity at the previous nominal m_Y point. The granularities for the $X \rightarrow YH$ searches are shown in Figs. 5.36 to 5.38. In the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$, low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches, 310, 93, and 197 intermediate mass points are added. The granularity is again finer at lower m_X , and also at higher m_Y where the discrimination between signal and background is more challenging.

In the $Y \rightarrow \gamma\gamma$ analyses, an even finer granularity in m_Y is defined corresponding to the width of the signal peak in the $m_{\gamma\gamma}$ distribution. Given that the pNN selection does

Search	Number of mass points
$X^{(0)} \rightarrow HH$	27
$X^{(2)} \rightarrow HH$	27
$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$	398
Low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	1765
High-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	2547

Table 5.11: The total (nominal plus intermediate) mass points considered in each search.

not change substantially over changes in m_Y corresponding to the $m_{\gamma\gamma}$ resolution, it is not necessary to create dedicated categories for these additional mass points. Instead, the categories defined for the nearest mass point are used. The final number of mass points (nominal plus intermediate) in each search is given in Table 5.11.

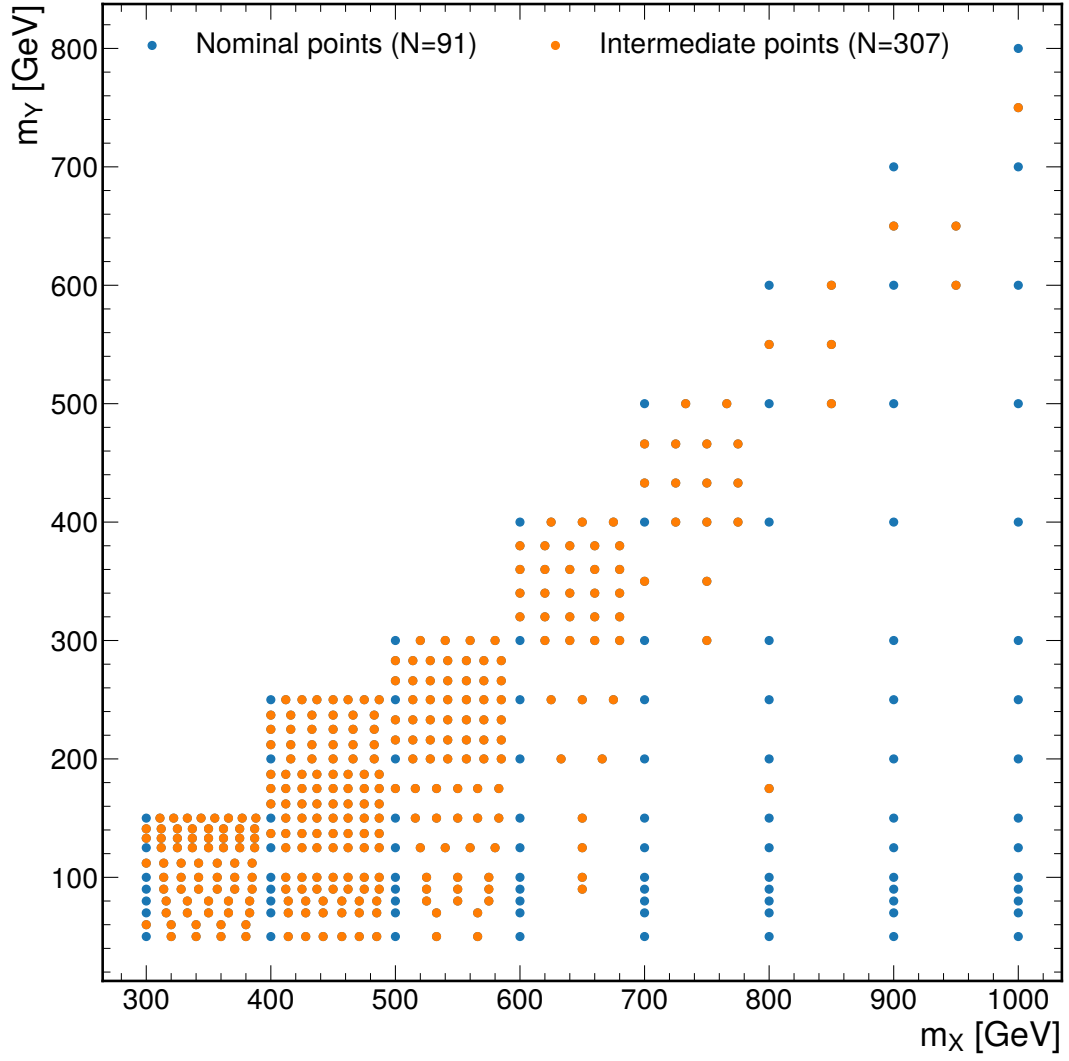


Figure 5.36: Search granularity for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search. Shown are the nominal mass points (blue) and the intermediate mass points (orange).

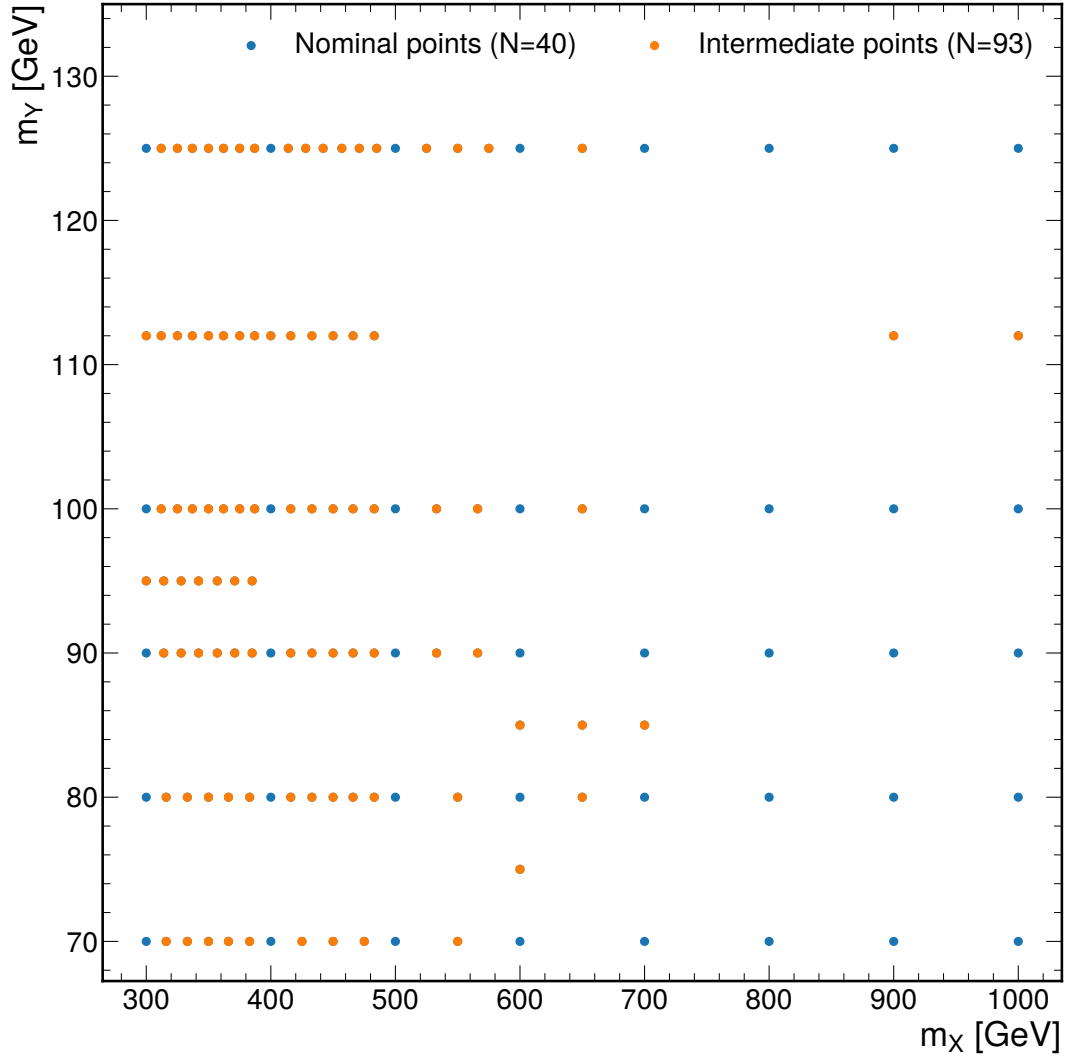


Figure 5.37: Search granularity for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Shown are the nominal mass points (blue) and the intermediate mass points (orange). In addition to the points shown in this plot, searches are also performed at intervals in m_Y corresponding to the width of the signal peak in $m_{\gamma\gamma}$.

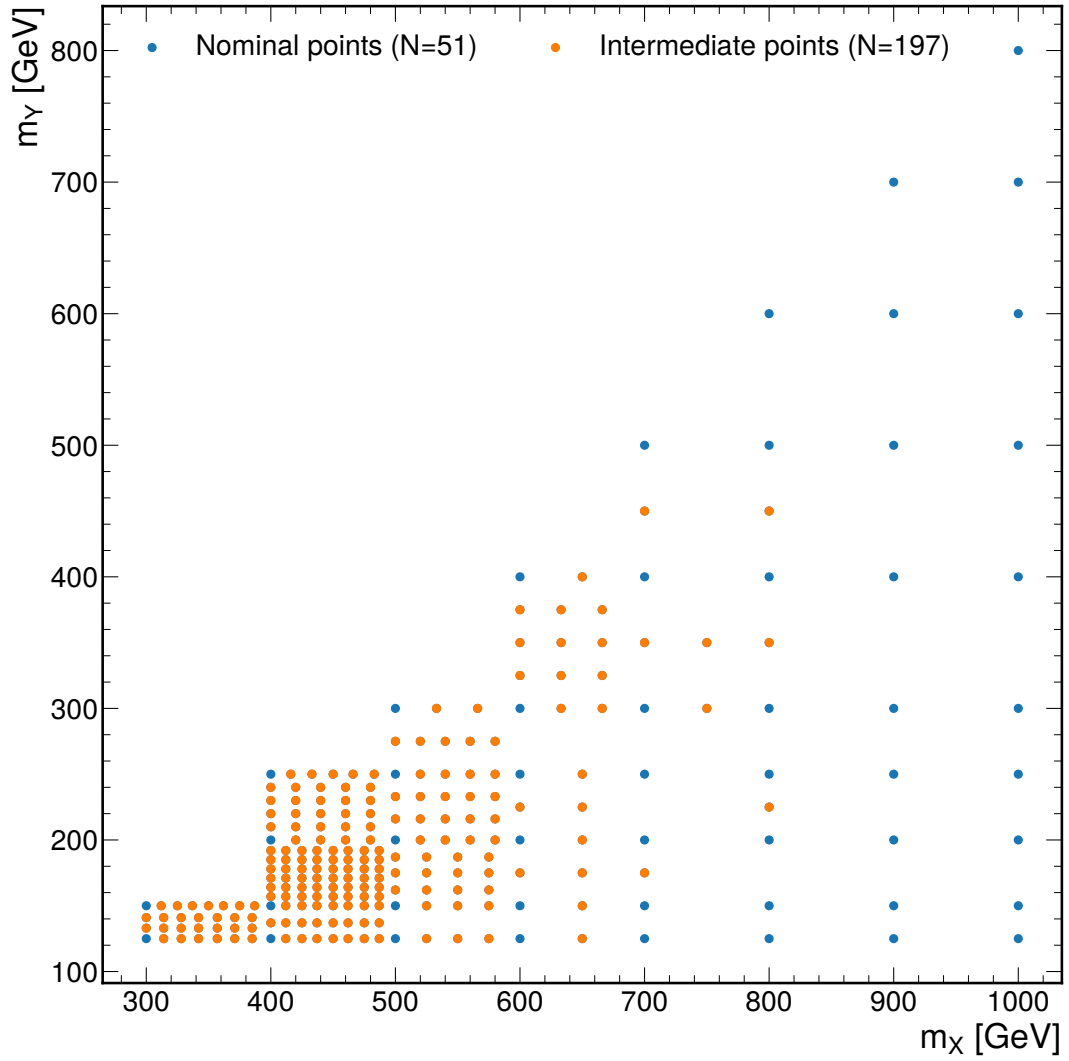


Figure 5.38: Search granularity for the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Shown are the nominal mass points (blue) and the intermediate mass points (orange). In addition to the points shown in this plot, searches are also performed at intervals in m_Y corresponding to the width of the signal peak in $m_{\gamma\gamma}$.

5.7 Sources of Systematic Uncertainty

In this analysis, systematic uncertainties affect the normalization and shape of the $m_{\gamma\gamma}$ distribution for the single and di-Higgs processes. The other contributing processes: the nonresonant background, and in the case of the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ analysis, the DY background, are fit entirely from data and have no associated systematic uncertainties. Normalization and shape changes due to systematic uncertainties are modelled in the likelihood with the techniques described in Section 4.1.3. The majority of systematic uncertainties affect only the normalization of the processes. This section details the sources of systematic uncertainty considered, and where appropriate, values that indicate the size of the uncertainty. Systematic uncertainties can be split into theoretical and experimental uncertainties, which are described in Section 5.7.1 and Section 5.7.2 respectively. Further details about systematic uncertainties and their impact can be found in Section 5.8 and Section 5.9.

5.7.1 Theoretical Uncertainties

Theoretical modelling of the single and di-Higgs processes have uncertainties related to QCD renormalization (μ_R) and factorization (μ_F) scales, the strong coupling constant (α_s), and parton distribution functions (PDFs). These uncertainties enter the calculation of the inclusive cross sections for the single Higgs processes [22], which lead to an overall scaling of the normalization of single Higgs processes in every analysis category. The largest contribution comes from the uncertainty related to the μ_R and μ_F scales and is of the order of 5–10%, depending on H production mode [22]. These theoretical uncertainties also enter the MC simulation used to calculate the efficiency of each single and di-Higgs process in each analysis category, and lead to category-specific scalings of each process. However, the corresponding impacts on the efficiencies were found to be less than 1% and were therefore not included in the final results extraction. There are also uncertainties on the $H \rightarrow \gamma\gamma$ and $H \rightarrow \tau\tau$ branching fractions whose values are around 2% [22].

5.7.2 Experimental Uncertainties

Experimental uncertainties consist mostly of uncertainties related to object calibration and the integrated luminosity measurement. The dominant sources of systematic uncertainty in this analysis are:

- *Photon energy calibration:* calibration of the photon energy scale and resolution that is described in Section 3.3.5 and in more detail in Refs. [62, 72, 73]. The corresponding uncertainties affect both the rate and shape of the processes and the

size of these impacts are described in Section 5.8.

- *Preselection*: scale factors described in Section 5.4 that are used to correct the efficiency of the preselection requirements in simulation to match data. The corresponding uncertainties on the signal efficiencies are between 4 and 5%.
- *Integrated luminosity*: integrated luminosities for the 2016, 2017 and 2018 data-taking periods, which have individual uncertainties between 1.2 and 2.5% [42–44].
- *DeepTau ID scale factors*: scale factors described in Section 3.3.7 that are used to correct the efficiency of selections made using the DeepTau ID algorithm. The corresponding uncertainties on the signal efficiencies are up to 7%.
- *Photon ID scale factors*: scale factors described in Section 3.3.5 that are used to correct the efficiency of selections made using the photon ID. The corresponding uncertainties on the signal efficiencies are up to 2%.
- *τ_h energy scale calibration*: calibration of the τ_h energy scale described in Section 3.3.7 that leads to an uncertainty in the τ_h energy of between 0.6 and 0.8%, depending on τ_h decay mode.
- *Signal model interpolation*: interpolation of the signal efficiency at intermediate mass points. In each analysis category, an uncertainty is derived from the splines used in the interpolation that corresponds to a possible migration of signal events from one category to another. This is described further in Section 5.8.1.

Additional systematic uncertainties are considered but have impacts on the final results less than ten times smaller than the sources described above. These include uncertainties related to trigger efficiencies, pileup simulation, electron identification, and the scale, resolution, and identification of muons, jets, and MET.

5.8 Modelling

To extract the final results in Section 5.9, a statistical model for the data, and a corresponding likelihood function is constructed. There is a single parameter of interest, μ , which is the cross section for the di-Higgs signal process in fb, and the primary observable is chosen as the diphoton mass, $m_{\gamma\gamma}$, in the analysis categories defined in Section 5.5.4. The likelihood function is constructed as described in Chapter 4 for a parametric shape analysis and the inputs to this construction, for every process, and for every analysis category, are:

- $\lambda(\vec{\mu}, \vec{\nu})$ — the expected number of events
- $f(m_{\gamma\gamma}; \mu, \vec{\Psi})$ — the probability density function (pdf) for the diphoton mass

which can be optionally factorized into nominal values and components that encode the impact of systematic uncertainties (Section 4.1.3), in which case the inputs are:

- $g(\vec{\mu})$ — the nominal expected number of events (Eq. 4.9)
- $\vec{\Psi}$ — the nominal shape parameters for $f(m_{\gamma\gamma}; \mu, \vec{\Psi})$
- $(\kappa_{\text{up}})_k$ and $(\kappa_{\text{down}})_k$ — parameters describing the impact of the nuisance parameter ν_k on $\lambda(\mu, \vec{\nu})$ (Eq. 4.10)
- α_{ik} — parameters describing the impact of the nuisance parameter ν_k on Ψ_i (Eq. 4.12).

This section describes the determination of these inputs, where the methods used depend on the type of process. Section 5.8.1 describes the MC-driven methods used for the signal process and single Higgs background process. Sections 5.8.2 and 5.8.3 describe the data-driven methods used for the nonresonant background, and the DY background, respectively.

5.8.1 Single and Di-Higgs Production

At the nominal mass points, the single Higgs and di-Higgs processes are modelled using MC simulation, and the methods used are identical. This section begins by describing these methods, before moving onto the models at the intermediate mass points, which for the di-Higgs processes, are determined using interpolation methods.

For the signal and single Higgs processes, models are created per data-taking year (2016, 2017 and 2018). This is not strictly necessary since the final fit to data is performed with all years combined, but is done for practical reasons because the systematic uncertainties are derived per year.

Single Higgs and Nominal Mass Points

For the di-Higgs process, the nominal expected number of events in analysis category, c , for data-taking year, y , is parameterized as:

$$g_{cy}(\mu) = \mu \cdot \sigma \cdot \epsilon_{cy} \cdot (\mathcal{L}_{\text{int}})_y \quad (5.3)$$

where $(\mathcal{L}_{\text{int}})_y$ is the integrated luminosity in fb^{-1} for data-taking year, y , ϵ_{cy} is an efficiency which is the fraction of events that pass the selection criteria in category c for a particular year, relative to all produced events, and σ is set to $\sigma = 1$ so that μ represents the cross section for the process in fb where the cross section includes the branching fractions of the Higgs boson decays. Similarly, for a single Higgs process, p , g_{pc} is parameterized as:

$$g_{pc} = \sigma_p \cdot \mathcal{B}_{\text{H} \rightarrow \gamma\gamma} \cdot \epsilon_{pcy} \cdot (\mathcal{L}_{\text{int}})_y \quad (5.4)$$

where μ is now omitted and the values of σ_p and $\mathcal{B}_{\text{H} \rightarrow \gamma\gamma}$ are taken from the LHC Higgs Cross Section Working Group [22], which are listed in Tables 2.2 and 2.3. For the single and di-Higgs processes, the efficiencies, ϵ , are calculated using MC simulation.

The nominal pdfs, $f_{pcy}(m_{\gamma\gamma}; \mu, \vec{\Psi})$, are determined by χ^2 fits of a Double Crystal Ball (DCB) function to MC events. A DCB consists of a Gaussian core which switches to a power law function on the left and another power law function on the right of the distribution:

$$f(m_{\gamma\gamma}; \hat{m}_{\gamma\gamma}, \sigma, \beta_l, m_l, \beta_r, m_r) = \begin{cases} \exp(-x^2/2) & \text{for } -\beta_l < x < \beta_r \\ A(\beta_l, m_l) \cdot (B(\beta_l, m_l) - x)^{-m_l} & \text{for } x < -\beta_l \\ A(\beta_r, m_r) \cdot (B(\beta_r, m_r) + x)^{-m_r} & \text{for } x > \beta_r \end{cases} \quad (5.5)$$

where

$$x = \frac{m_{\gamma\gamma} - \hat{m}_{\gamma\gamma}}{\sigma} \quad (5.6)$$

$$A(\beta, m) = \left(\frac{m}{|\beta|} \right)^m \cdot \exp\left(-\frac{|\beta|^2}{2}\right) \quad (5.7)$$

$$B = \frac{m}{\beta} - |\beta| \quad (5.8)$$

The χ^2 fits are performed in 50 bins of the $m_{\gamma\gamma}$ distribution in a 25 GeV window centred at 125 GeV for the single Higgs processes, and the di-Higgs processes for the $X \rightarrow \text{HH}$ searches and the $X \rightarrow Y(\tau\tau)\text{H}(\gamma\gamma)$ searches. In the $X \rightarrow Y(\gamma\gamma)\text{H}(\tau\tau)$ searches, the window is centred at m_Y , and the size of the window is scaled linearly with m_Y to account for the increasing width of the $m_{\gamma\gamma}$ distribution. Examples of these fits are

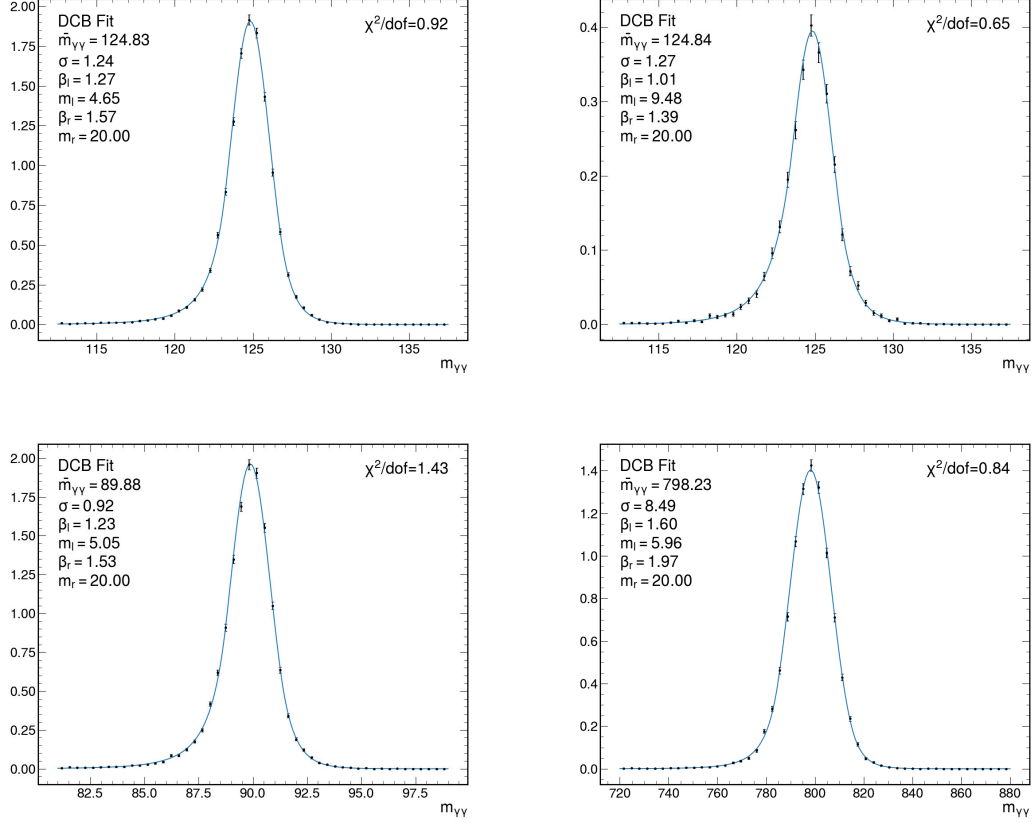


Figure 5.39: DCB fits in category 0 in 2018 for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search (top) and $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search (bottom). The mass points are $(m_X, m_Y) = (600, 90)$ (top-left), $(300, 70)$ (top-right), $(600, 90)$ (bottom-left), and $(1000, 800)$ GeV (bottom-right).

shown in Fig. 5.39. In general, the DCB is a good model for the $m_{\gamma\gamma}$ distribution as indicated by values of χ^2 which are approximately equal to the degrees of freedom.

To study the impact of a nuisance parameter on the modelling of the single and di-Higgs processes, simulated datasets are created when changing the nuisance parameter to $\nu = -1$ or $\nu = 1$, which are referred to as down and up variations respectively. These datasets are created either by reweighting the nominal dataset ($\nu = 0$), or by rerunning the simulation with $\nu = -1$ or $\nu = 1$, where the approach taken depends on the nuisance parameter. Typically, nuisance parameters related to ID scale factors take the former approach, and nuisance parameters related to energy scale and resolution corrections take the latter approach. The κ factors that describe the impact of a nuisance parameter are derived by taking the ratio of the expected number of events in the up and down variations to the expected number of events in the nominal variation.

The derivation of the α parameters is more involved. Firstly, only the effect of a

nuisance parameter on $\hat{m}_{\gamma\gamma}$ and σ is studied. The impact on all other shape parameters of the DCB are assumed to be negligible. Instead of refitting the DCB to the up and down variations, proxy variables are used to describe the changes in $\hat{m}_{\gamma\gamma}$ and σ . For $\hat{m}_{\gamma\gamma}$, this is simply the mean of the $m_{\gamma\gamma}$ distribution, and for σ , it is the smallest $m_{\gamma\gamma}$ interval that contains 68% of the distribution, which for a Gaussian distribution, would be equivalent to the standard deviation. Then the ratio of the proxy variable in the up and down variations to the proxy variable in the nominal variation is taken, and the average of those ratios is taken as the α parameter. In the $X^{(2)} \rightarrow \text{HH}$ search in category 0, the average value of α for the photon energy scale calibration is 0.005 and 0.014 for $\hat{m}_{\gamma\gamma}$ and σ respectively, and in the high-mass $X \rightarrow Y(\gamma\gamma)\text{H}(\tau\tau)$ search, the equivalent numbers are 0.01 and 0.04.

Intermediate Mass Points

For the single Higgs processes, the modelling at the intermediate mass points follows the same methods described above since an analysis at an intermediate mass point just corresponds to different selection on the pNN score.

For the di-Higgs processes at the intermediate mass points, ϵ and $\vec{\Psi}$ are determined from the values found at the closest nominal mass point, or by interpolating the values found with neighbouring nominal mass points, depending on the parameter. The choice of which approach to take is motivated by the level of dependence of the parameter on m_X and m_Y . Figure 5.40 shows ϵ at the nominal mass points as a function of m_X in the $X^{(2)} \rightarrow \text{HH}$ search when selecting $\tilde{f}(\vec{x}; m_X)$ in ranges of 0.98–0.99 or 0.99–1.0. These plots show that the ϵ can change significantly between mass points, motivating the use of interpolation methods.

To determine ϵ at an intermediate mass point, m_X^{int} , in the $X \rightarrow \text{HH}$ searches, for a category defined by $b_1 < \tilde{f}(\vec{x}; m_X^{\text{int}}) < b_2$, linear and cubic splines are created from the ϵ values derived using the nominal MC samples with selections: $b_1 < \tilde{f}(\vec{x}; m_X) < b_2$. Examples of this for category 0 and category 1 for $m_X = 375$ and 433 GeV in the $X^{(2)} \rightarrow \text{HH}$ search are shown in Fig. 5.41. To study the robustness of the splines, they are recreated when removing the nominal mass point closest to m_X^{int} , and the difference between the splines at m_X^{int} is calculated. Then, the value for ϵ at m_X^{int} is given by the linear or cubic spline derived using all mass nominal points, depending on which type minimizes this difference. A systematic uncertainty is then assigned to ϵ using the difference found for the chosen spline.

Figure 5.43 shows how the shape parameters of the DCB change at the nominal mass points as a function of m_X in the $X^{(2)} \rightarrow \text{HH}$ search when selecting $\tilde{f}(\vec{x}; m_X)$ in the range 0.99–1.0. In these plots, the dependence on $\hat{m}_{\gamma\gamma}$ is shown via $\Delta m_{\gamma\gamma} = \hat{m}_{\gamma\gamma} - 125 \text{ GeV}$.

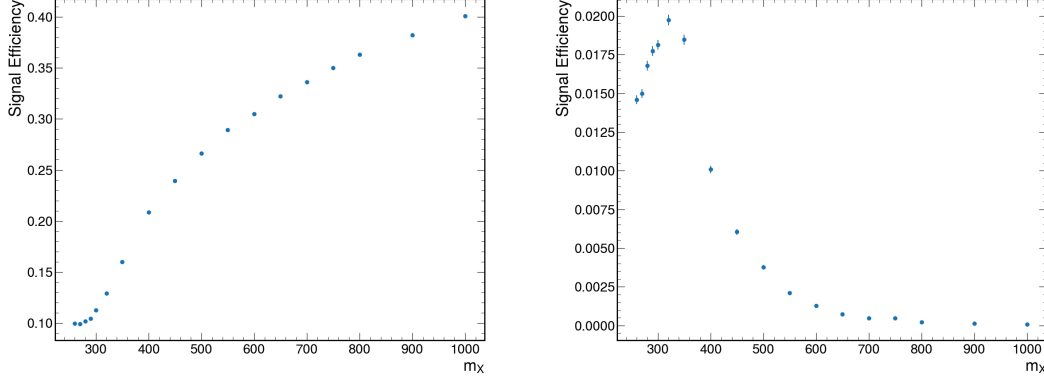


Figure 5.40: Signal efficiency as a function of m_X for the $X^{(2)} \rightarrow HH$ averaged over years. On the left, a requirement of $\tilde{f}(\vec{x}; m_X) > 0.99$ is placed, and on the right, a requirement of $0.98 < \tilde{f}(\vec{x}; m_X) < 0.99$. These requirements illustrate the trends seen in the purest category and a less-pure category respectively.

The $\Delta m_{\gamma\gamma}$, σ , and β_l shape parameters are seen to change significantly enough between mass points to justify the use of interpolation methods. The approach is the same as for ϵ , except that only linear splines are used, and no systematic uncertainty is assigned to the interpolated shape parameters. The values of β_r , m_l , and m_r , show smaller dependence on m_X , and are therefore fixed to the values found at the closest nominal mass point.

To validate the approach taken to determine the shape parameters, chosen nominal mass points are treated as intermediate mass points, and the interpolated shapes are compared to those found from a fit to the MC sample. These comparisons are shown in Fig. 5.44. The goodness-of-fit using the interpolated shapes is still fair, validating the approach.

In the $X \rightarrow YH$ searches, ϵ and $\vec{\Psi}$ must be studied as a function of m_Y as well. Figure 5.46 shows ϵ as a function of m_Y when $m_X = 1000$ GeV for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches, and Fig. 5.47 similarly shows the dependence of the shape parameters on m_Y in the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. These figures lead to the same conclusions as for the $X \rightarrow HH$ searches. The parameters: ϵ , $\Delta m_{\gamma\gamma}$, σ , and β_l must be interpolated, now as functions of m_X and m_Y , and the β_r , m_l and m_r parameters are taken from the closest nominal mass point. The 2D splines are implemented using the SCIPY package [116] where the cubic spline uses techniques described in Ref. [117]. To illustrate how these splines work in 2D, a linear and cubic spline for an arbitrary function is shown in Fig. 5.45. The choice of the type of spline, and the derivation of the systematic uncertainty, is the same as for the 1D splines.

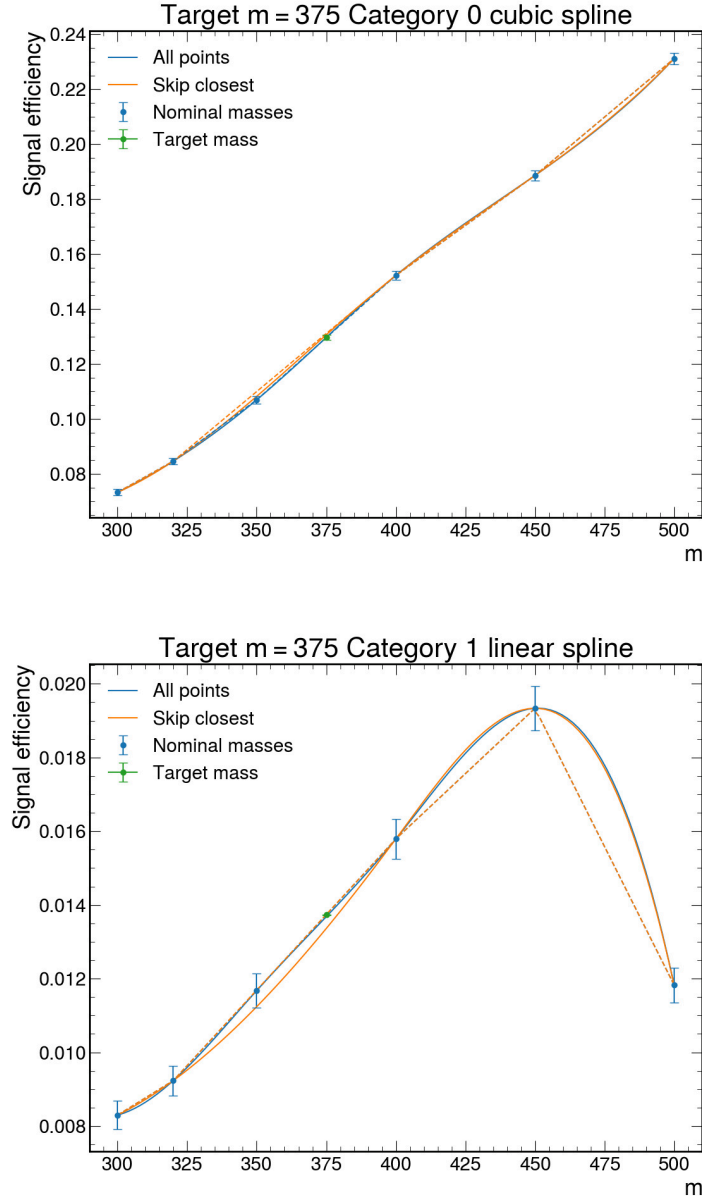


Figure 5.41: Splines used to estimate the signal efficiency for $m_X = 375$ in the $X^{(2)} \rightarrow HH$ search in 2018 for category 0 (top) and category 1 (bottom). In the plots, m refers to m_X . The solid and dashed lines are cubic and linear splines respectively and the spline type chosen to calculate the signal efficiency is specified in the plot title. The uncertainty in the interpolated signal efficiency (at $m_X = 375$) is shown by the green error bar and is taken from the difference between the spline found when using all nominal mass points and the spline found when skipping the closest mass point.

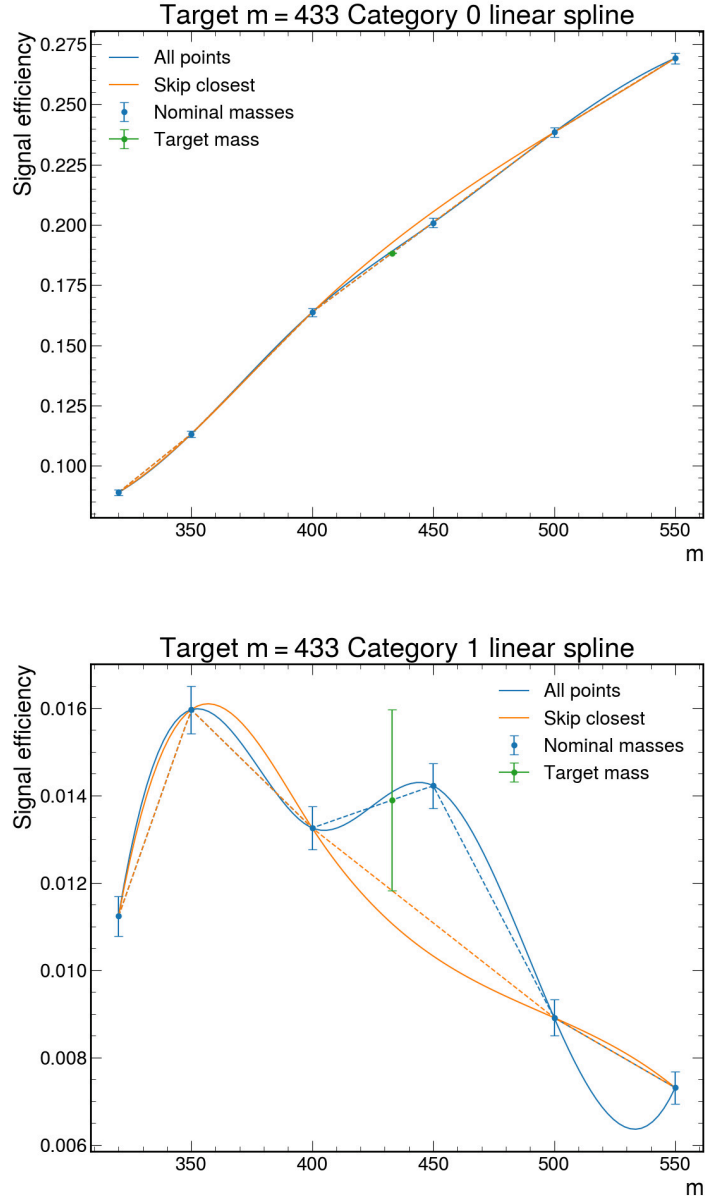


Figure 5.42: Splines used to estimate the signal efficiency for $m_X = 433$ in the $X^{(2)} \rightarrow HH$ search in 2018 for category 0 (top) and category 1 (bottom). In the plots, m refers to m_X . The solid and dashed lines are cubic and linear splines respectively and the spline type chosen to calculate the signal efficiency is specified in the plot title. The uncertainty in the interpolated signal efficiency (at $m_X =$) is shown by the green error bar and is taken from the difference between the spline found when using all nominal mass points and the spline found when skipping the closest mass point.

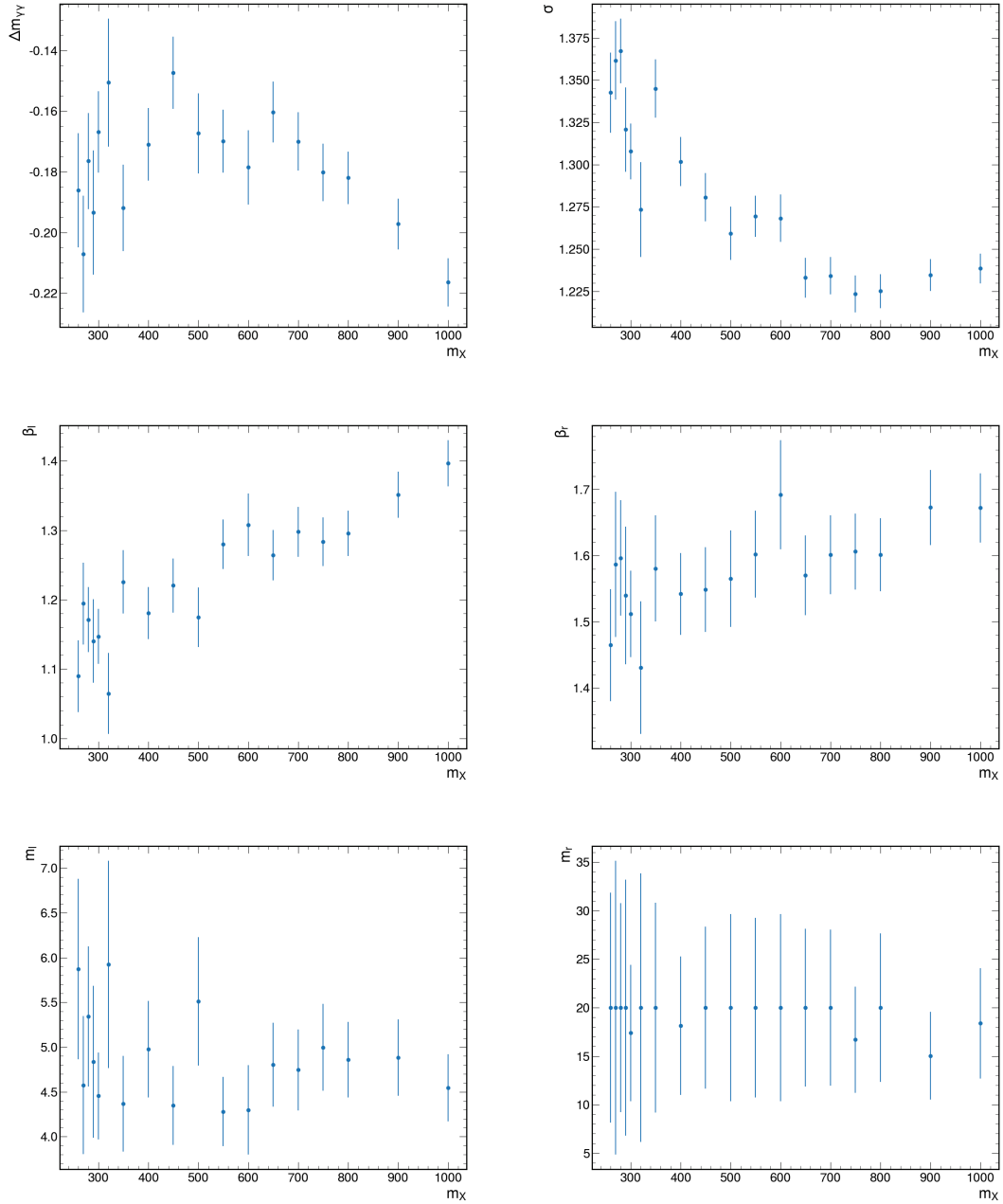


Figure 5.43: Shape parameters from DCB fits to simulated $X^{(2)} \rightarrow HH$ events, as a function of m_X . For each fit, a requirement of $\hat{f}(\vec{x}; m_X) > 0.99$ is placed. The uncertainty in the shape parameters are indicated by the blue error bars.

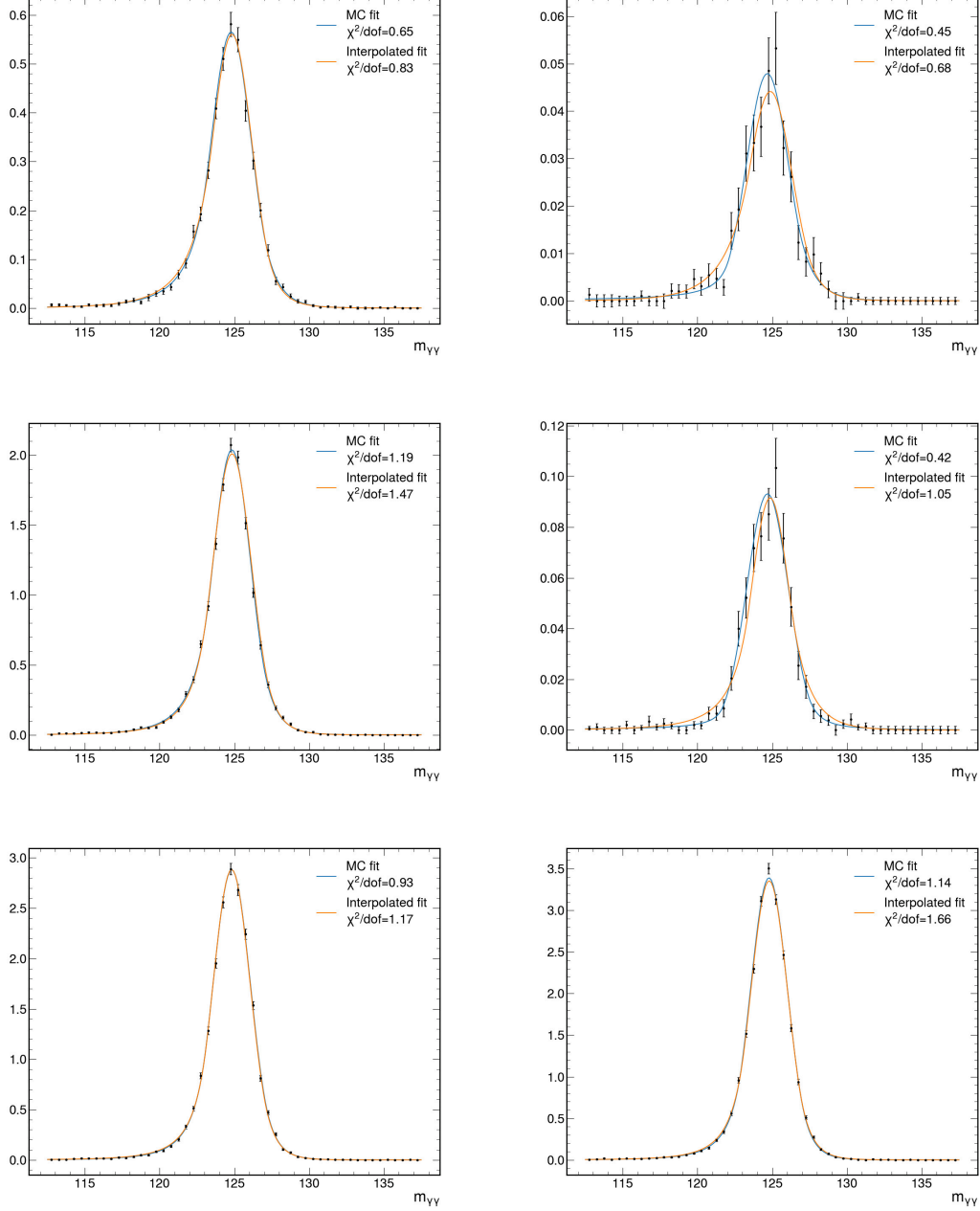


Figure 5.44: Comparisons between signal models derived via direct fits to MC events and derived via the interpolation procedure which uses fits to MC events at neighbouring mass points. All comparisons are for the $X^{(2)} \rightarrow HH$ analysis in 2018. To isolate the comparison to the shape interpolation, the signal efficiency of the interpolated model is set to the value from the direct fit. The top and middle plots show the $m_X = 270$ and 500 GeV mass points respectively where the left and right plots are for category 0 and category 1 respectively. The bottom-left and bottom-right plots are for category 0 for the $m_X = 700$ and 900 GeV mass points respectively.

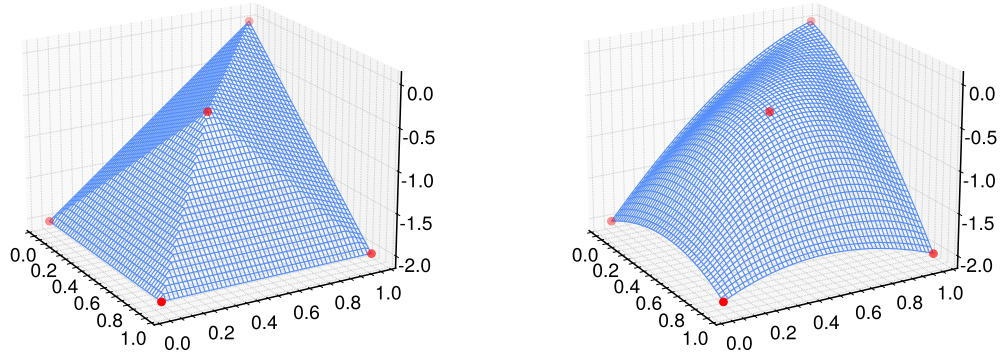


Figure 5.45: Examples of linear (left) and cubic (right) interpolation for an arbitrary function in two dimensions.

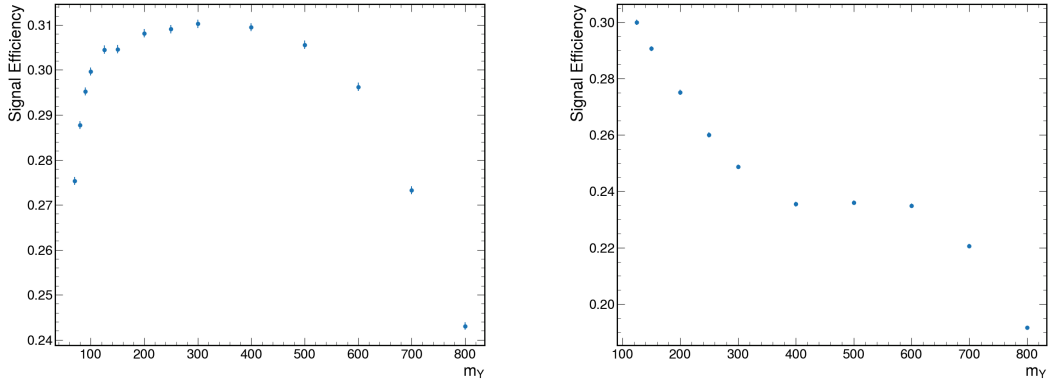


Figure 5.46: Signal efficiency as a function of m_Y where $m_X = 1000 \text{ GeV}$ for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (left) and high-mass $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (right) searches. These models are derived from MC across all years with a requirement on the pNN score of $\tilde{f}(\vec{x}; m_X, m_Y) > 0.99$ placed which is representative of the purest categories in the analysis.

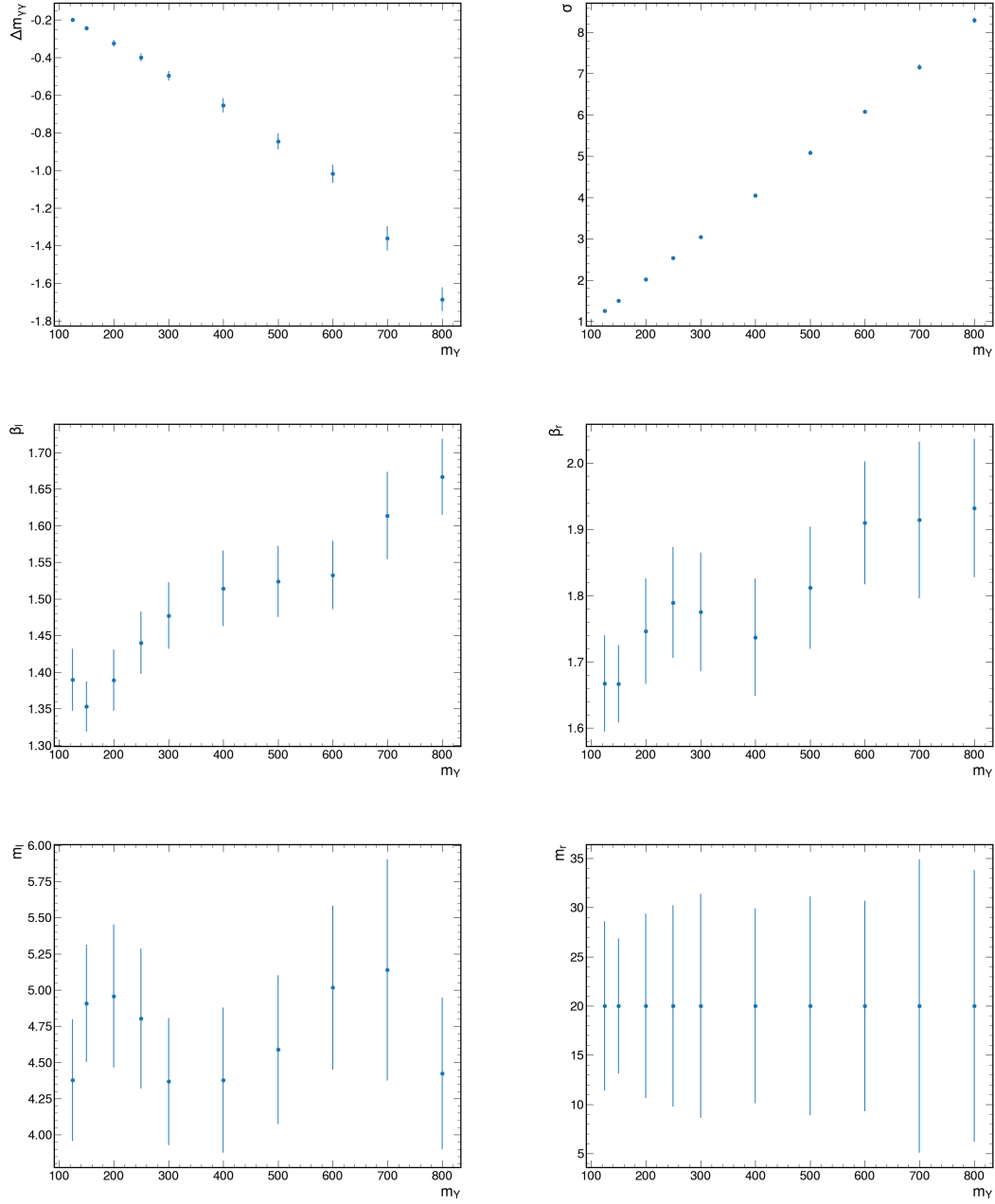


Figure 5.47: Fitted shape parameters as a function of m_Y for the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search where $m_Y = 1000$ GeV. These models are derived from MC across all years with a requirement on the pNN score of $\hat{f}(\vec{x}; m_X, m_Y) > 0.99$ placed which is representative of the purest categories in the analysis.

5.8.2 Nonresonant Background

The nonresonant background models are derived directly from data, where the sidebands regions, i.e. regions of $m_{\gamma\gamma}$ well separated from the signal peak, are able to constrain the rate and shape of the background in the signal region. The methodology is outlined first for the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches and then the differences for the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches are described.

$X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Searches

The distribution of $m_{\gamma\gamma}$ in data and background MC after preselection for the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches is shown in Fig. 5.48. The nonresonant component of the background has a smoothly falling $m_{\gamma\gamma}$ distribution. There is not a single, well-motivated analytical function to describe this distribution, so multiple *families* of functions, listed in Table 5.12, are considered. Within a family, there are functions of varying *orders*, for example, for a sum of exponentials, the order corresponds to the number of exponentials in the sum. A higher order means more parameters in the function, and therefore greater freedom in the shape of the function.

If a function is chosen that does not accurately describe the underlying $m_{\gamma\gamma}$ distribution, that choice could bias the final results. Therefore, an approach must be taken that introduces an uncertainty related to the choice of function. This is done by using the *discrete profiling* method [118], where the choice of function is encoded as a discrete nuisance parameter which is profiled in the likelihood fit to data. This leads to an increased uncertainty on the parameter of interest that corresponds to the size of the impact that the discrete nuisance parameter has on the fit.

Ideally, the discrete nuisance parameter could correspond to the choice between an infinite number of functions (beyond those listed in Table 5.12) but this is practically impossible to achieve. Instead, a subset of functions, called an *envelope*, is chosen that contains functions that fit well to the data, and have sufficiently different shapes such that a fit with the envelope is a good approximation of a fit with infinitely many functions.

Every analysis category has its own envelope of functions. The building of a category's envelope begins by performing a maximum likelihood fit with the lowest-order function of each family in Table 5.12 to the $m_{\gamma\gamma}$ distribution in data in the analysis category. Only the sidebands of the $m_{\gamma\gamma}$ distribution are used, which excludes the $115 < m_{\gamma\gamma} < 135$ GeV region, so that a signal peak at about 125 GeV will not influence the fit. This procedure continues up to an order corresponding to 6 free parameters, or until an order is reached where a loose goodness-of-fit criteria, p -value > 0.01 , is satisfied, in which case the function is added to the envelope.

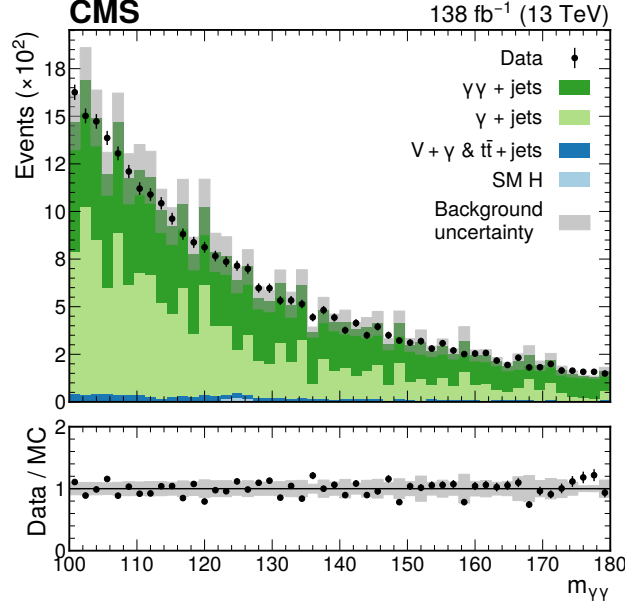


Figure 5.48: Distribution of $m_{\gamma\gamma}$ in data and background MC after preselection for the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches. The background MC is normalized to data and its statistical uncertainty is indicated by the grey shaded bands.

Table 5.12: Function families considered for modelling the nonresonant background in the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches with the discrete profiling method. Each function is shown for order, N , and the free parameters of the functions are denoted p_i .

Name	Formula
Sum of exponentials	$f_N(x) = \sum_{i=0}^N p_{2i} \exp(p_{2i+1}x)$
Sum of power laws	$f_N(x) = \sum_{i=0}^N p_{2i} x^{-p_{2i+1}}$
Bernstein polynomials	$f_N(x) = \sum_{i=0}^N p_i \binom{N}{i} x^i (1-x)^{N-i}$
Laurent series	$f_N(x) = \sum_{i=0}^N p_i x^{-4+g(i)}; \quad g(i) = \sum_{j=0}^i (-1)^j j$
Exponential of a polynomial	$f_N(x) = \exp\left(\sum_{i=0}^N p_i x^i\right)$

Starting from the functions that satisfy the loose goodness-of-fit criteria, further orders are added according to an F-test requirement [119]. When adding a function of a higher order, the difference in $-2\ln\mathcal{L}$ between the two fits, Δ , is used to calculate a p-value defined as:

$$p = \int_{\Delta}^{\infty} \chi^2(x, m) dx \quad (5.9)$$

where $\chi^2(x, m)$ is a χ^2 distribution with m degrees of freedom, and m is the difference in the number of free parameters between the two orders. If the p-value is less than 0.05, the fit quality is deemed to have improved sufficiently enough, and the function is added to the envelope. If it is more than 0.05, the function is not added, and no more functions of that family are considered. This continues until a maximum order corresponding to 6 free parameters is reached.

The pdf of the nonresonant background in an analysis category is then given by:

$$f(m_{\gamma\gamma}; \vec{\Psi}_1, \vec{\Psi}_2 \dots \vec{\Psi}_n, \nu) = \begin{cases} f_1(m_{\gamma\gamma}; \vec{\Psi}_1) & \text{if } \nu = 1 \\ f_2(m_{\gamma\gamma}; \vec{\Psi}_2) & \text{if } \nu = 2 \\ \vdots & \vdots \\ f_n(m_{\gamma\gamma}; \vec{\Psi}_n) & \text{if } \nu = n \end{cases} \quad (5.10)$$

where ν is the discrete nuisance parameter, and f_i is the i th function of the envelope of n functions. The expected number of nonresonant background events in an analysis category is left as a free parameter that is fit in data.

Examples of the envelope construction from the $(m_X, m_Y) = (300, 50)$ GeV mass point in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search, are shown in Fig. 5.49 for categories with 10, 20, 80 and 320 expected background events. In general, the envelopes contain a single function from each family, and do not contain functions with orders greater than 2. This is a consequence of the small number of expected background events in the analysis categories. To illustrate this point, an envelope is constructed for a category with the remaining events that pass preselection but not the pNN selection, and is also shown in Fig. 5.49. There, functions of up to order 4 are included.

$X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Searches

The nonresonant background modelling in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches is done in a similar way to the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches, but with some differences. Firstly, the sidebands definition is changed to exclude a region of $m_Y \pm (125 \text{ GeV}/m_Y)$, except for $m_Y < 72 \text{ GeV}$ where the $m_{\gamma\gamma} < 68 \text{ GeV}$ region is kept so that there are enough events to perform the envelope construction. Only 4% of the events from a $m_Y = 70 \text{ GeV}$

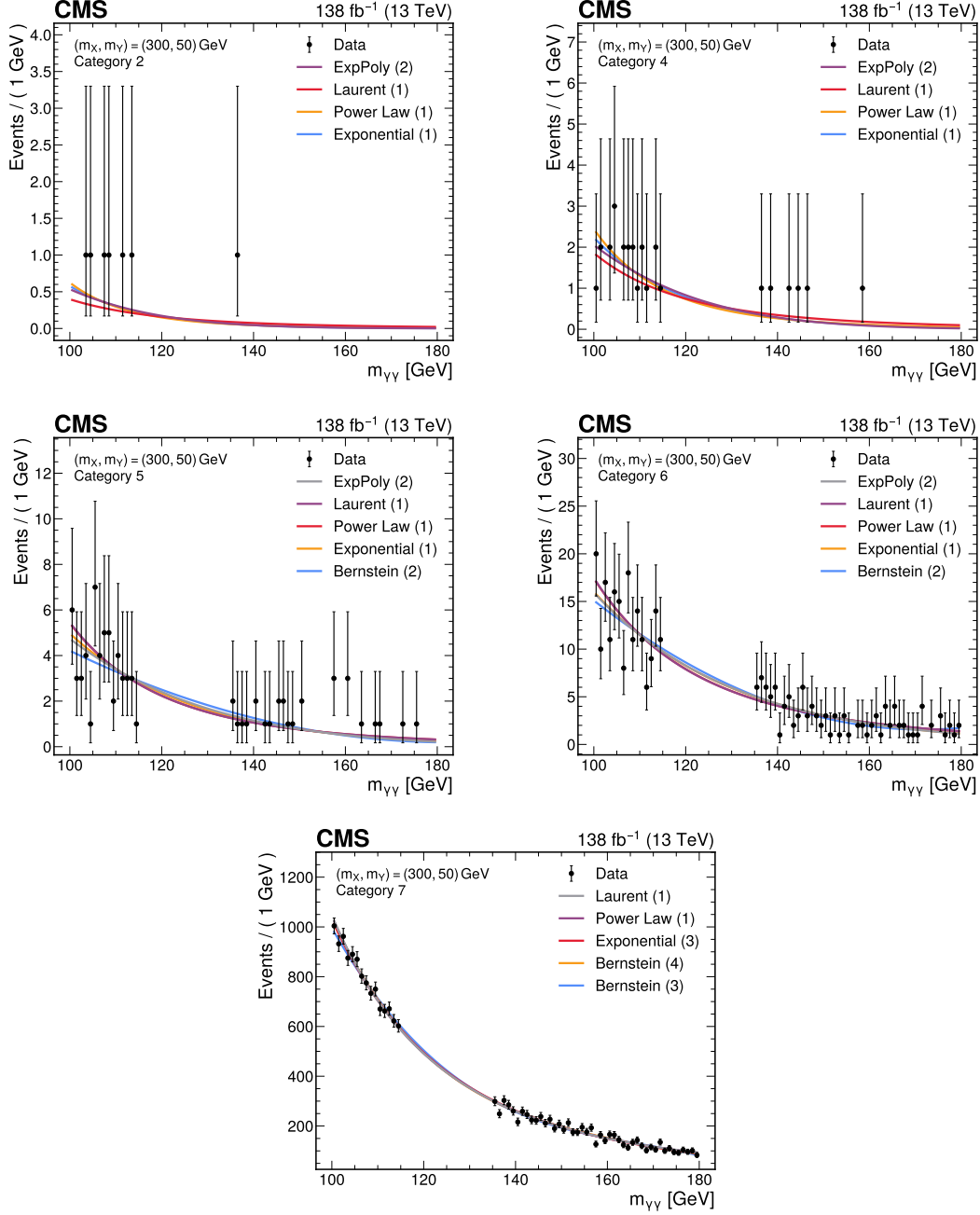


Figure 5.49: Envelope construction for the $(m_X, m_Y) = (300, 50)$ GeV mass point in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search. Shown are the $m_{\gamma\gamma}$ distributions in data in the sideband regions for categories 2, 4, 5 and 6 (top-left to middle-right) which are expected to have 10, 20, 80 and 320 background events respectively. An additional category (bottom) consisting of the events passing preselection but not passing the pNN selection, is also shown to illustrate how the envelope construction works in categories with more events. The functions entering the envelope are indicated in the legend and their fits to the $m_{\gamma\gamma}$ distributions are shown in the plots.

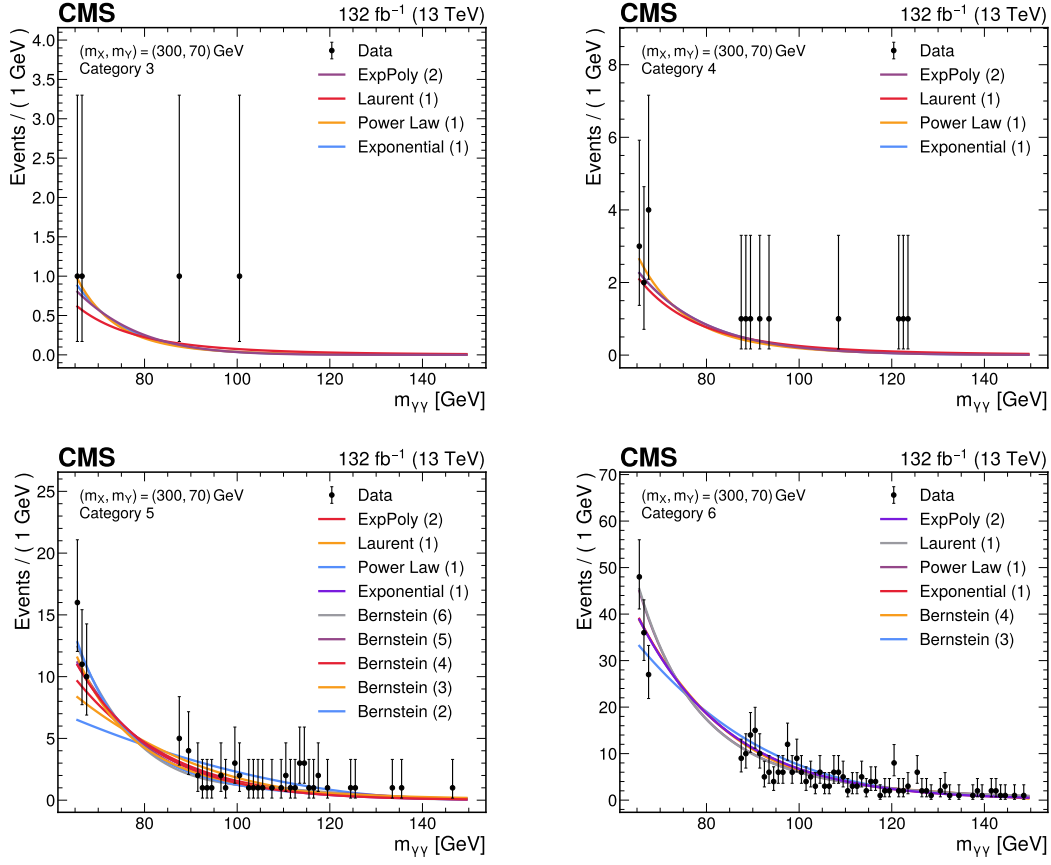


Figure 5.50: Envelope construction for the $(m_X, m_Y) = (300, 70)$ GeV mass point in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Shown are the $m_{\gamma\gamma}$ distributions in data in the sideband regions for categories 3, 4, 5 and 6 (top-left to middle-right) which are expected to have 10, 20, 80 and 320 background events respectively. The functions entering the envelope are indicated in the legend and their fits to the $m_{\gamma\gamma}$ distributions are shown in the plots.

signal fall into this region so the impact from a potential signal in data on the envelope construction is likely to be small. For the low-mass search, the sideband definition is the only difference, and the final envelopes are not substantially different from those in the $X \rightarrow HH$ and $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ searches, as illustrated by examples for $(m_X, m_Y) = (300, 70)$ GeV in Fig. 5.50.

In the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, a further change is made to the function families considered. Over the larger $m_{\gamma\gamma}$ range (up to 1 TeV), the Bernstein polynomials do not pass the loose goodness-of-fit criteria, so they were removed from consideration. To reintroduce some shape flexibility to the envelope, a new family of functions, referred

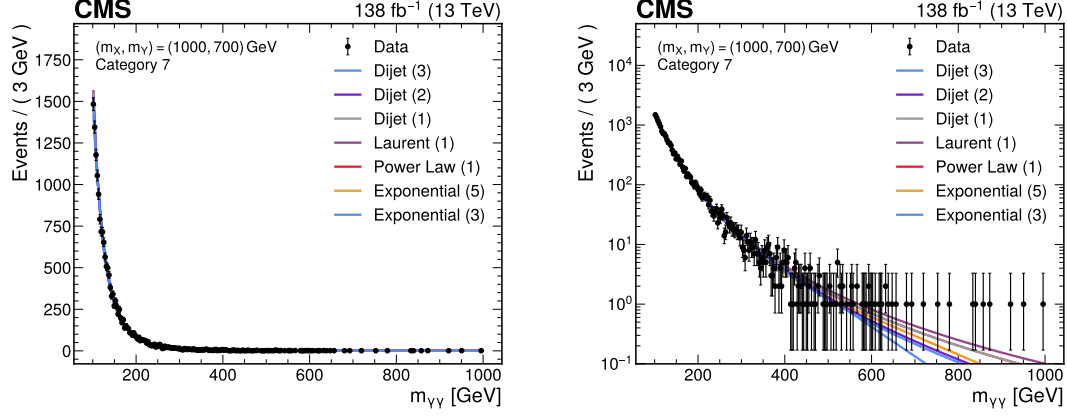


Figure 5.51: Envelope construction for the $(m_X, m_Y) = (1000, 700)$ GeV mass point in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Shown is the $m_{\gamma\gamma}$ distribution in data for a category consisting of the events passing preselection but not passing the pNN selection. The functions entering the envelope are indicated in the legend and their fits to the $m_{\gamma\gamma}$ distributions are shown in the plots. The same plot is shown with a linear (left) and log (right) y-scale.

to as Dijet functions, is added. The Dijet function is defined as:

$$f_n(x) = \begin{cases} x^{-p_1}, & n = 1 \\ (1-x)^{p_1} x^{-p_2}, & n = 2 \\ (1-x)^{p_1} x^{-(p_2+p_3 \ln x)}, & n = 3 \\ (1-x)^{p_1} x^{-(p_2+p_3 \ln x + p_4 \ln^2 x)}, & n = 4 \\ (1-x)^{p_1} x^{-(p_2+p_3 \ln x + p_4 \ln^2 x + p_5 \ln^3 x)}, & n = 5 \end{cases} \quad (5.11)$$

The suitability of this function family is demonstrated in Fig. 5.51 for the category comprised of the remaining events after the pNN selection for the $(m_X, m_Y) = (1000, 700)$ GeV mass point where Dijet functions of up to order 3 are included in the envelope.

5.8.3 Resonant Background from Electron Misidentification

In the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, there is an additional background originating from Drell-Yan ($pp \rightarrow Z$) where the Z boson decays to a pair of electrons which are both misidentified as photons, and where associated jets are misidentified as tau leptons. A smaller contribution also arises from diboson ($pp \rightarrow ZZ/WW/ZW$) events, where a Z boson decays to a pair of electrons which are both misidentified as photons, and the other vector boson decays to tau lepton(s) which are then selected. If both electrons from a $Z \rightarrow e^+e^-$ decay are selected to form a diphoton candidate, this would result in a peak in

the $m_{\gamma\gamma}$ distribution at the Z boson mass. The modelling of this background is therefore crucial to avoid the reporting of a false signal. Given that this background primarily arises from Drell-Yan (DY) events, it is referred to as the DY background.

Initially, simulated DY and diboson events were studied as way to model the DY background. However, the statistics were too small to provide a reliable estimate of the expected number of events in the analysis categories. Therefore, a data-driven approach was instead developed, using an ABCD method with control regions based upon the pNN selection and the CSEV and pixel veto (Section 3.3.5). At a given mass point, the regions defined by the analysis categories are denoted D_i . The objective of the ABCD method is to create a model of the resonant background (from $e \rightarrow \gamma$ misidentification) in each D_i region.

Each D_i region has a corresponding B_i region that has the same pNN selection but uses a dielectron trigger and inverts the CSEV and pixel veto (inverted selection). By inverting the vetoes, electrons from $Z \rightarrow e^+e^-$ decays are likely to be reconstructed and selected as photon candidates, thereby creating a region where the $m_{\gamma\gamma}$ peak from the $Z \rightarrow e^+e^-$ decay is dominant, which is shown in Fig. 5.52 (top-right) for the D_0 region for the $(m_X, m_Y) = (650, 95)$ GeV mass point. In both D_i and B_i , the $m_{\gamma\gamma}$ variable is constructed using information from the ECAL only. Given that the shower shape of electrons and photons is similar, and that the selection on kinematic variables is the same between regions B_i and D_i , the shape of the $m_{\gamma\gamma}$ peak is expected to be the same in both regions. Therefore, the B_i region is used to derive the shape of the DY background in region D_i .

To determine the rate of the DY background in D_i , the extracted rate from B_i is multiplied by a transfer factor that is constrained by regions A and C . Region C is defined by the normal preselection and an inversion of the pNN selection, i.e. contains events that do not pass the pNN selection for any of the analysis categories. Region A has the same pNN selection but uses a dielectron trigger and inverts the CSEV and pixel veto. Examples of the $m_{\gamma\gamma}$ distributions in these regions are also shown for the $(m_X, m_Y) = (650, 95)$ GeV mass point in Fig. 5.52 (left).

In every region, there is a nonresonant background that is treated in the same way as described in Section 5.8.2. The ABCD method is then implemented by adding a DY process to every region where the expected number of events in region D is given by:

$$\lambda_D = \lambda_B \times \frac{\lambda_C}{\lambda_A} \quad (5.12)$$

where λ_j is the number of events in region $j \in \{A, B, C\}$, where these values are left as free parameters in the fit to data. The function chosen to parameterize the shape of

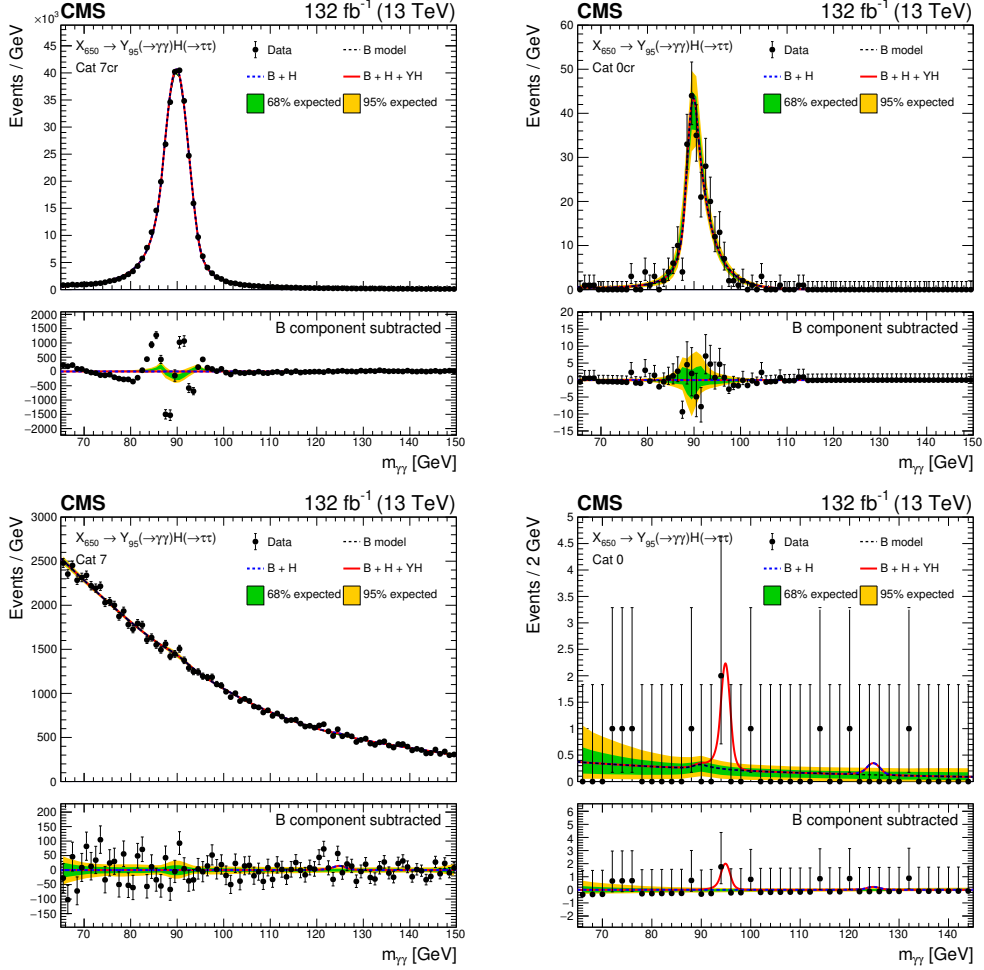


Figure 5.52: Distributions of $m_{\gamma\gamma}$ in data in regions A (top-left), B_0 (top-right), C (bottom-left), and D_0 (bottom-right) for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search and for $(m_X, m_Y) = (650, 95)$ GeV. A signal-plus-background fit (red) is performed with data (black points). The nonresonant background is modelled following the methods described in Section 5.8.2. The DY background is described by a DCB or a Gaussian function depending on goodness-of-fit tests. In this example, DCB functions are used in all regions where the shape parameters are shared between regions A and C , and a separate set of parameters are shared between regions B_0 and D_0 . The one (green) standard deviation and two (yellow) standard deviation bands show the uncertainties in the nonresonant + DY background component of the model (black dashed line).

the DY background is the same for each pairing of D_i and B_i , and can be a Gaussian or a DCB. At first, a Gaussian function is fit to the $m_{\gamma\gamma}$ distribution in region B_i , and a goodness-of-fit test is performed. If the p-value is greater than 0.01, then a Gaussian function is used, otherwise a DCB function is used. In the final fit to data, the shape parameters, $\vec{\Psi}_i$, for the function are the same in regions D_i and B_i for a given i , but change across different i , to account for how the pNN selection may influence the $m_{\gamma\gamma}$ shape. The shape parameters are also left to float in the final fit to data.

This ABCD method assumes that the shape of the pNN score is the same regardless of whether the normal preselection or inverted selection is used. In other words, the variables used to define the ABCD regions are uncorrelated. This was tested in simulated DY and diboson events where the reconstructed photons were matched to electrons from $Z \rightarrow e^+e^-$ decays at the generator level. The shapes of the pNN score for the $m_Y = 90$ GeV and $m_X = 300, 500, 800$ and 1000 GeV mass points, for the normal preselection and inverted selected are shown in Fig. 5.53. Within the statistical uncertainties, the two distributions are compatible, and the same is found for other values of m_X at $m_Y = 90$ GeV.

The expected number of DY events (yield) in the D regions extracted from a background-only maximum likelihood fit to data using this ABCD method are shown in Fig. 5.54. The yields in D_0 are shown as a function of m_X for $m_Y = 90$ and 100 GeV, and are found to range between about 0.03 and 0.4 with an upwards trend in m_X . The numbers are also shown in regions D_0 to D_3 for the $(m_X, m_Y) = (650, 100)$ GeV and $(m_X, m_Y) = (325, 100)$ GeV where yields remain below 0.4, which is also the case for all other mass points. Given the small expected number of events, the impact on the final results is expected to be small. This is discussed further in Section 5.9.

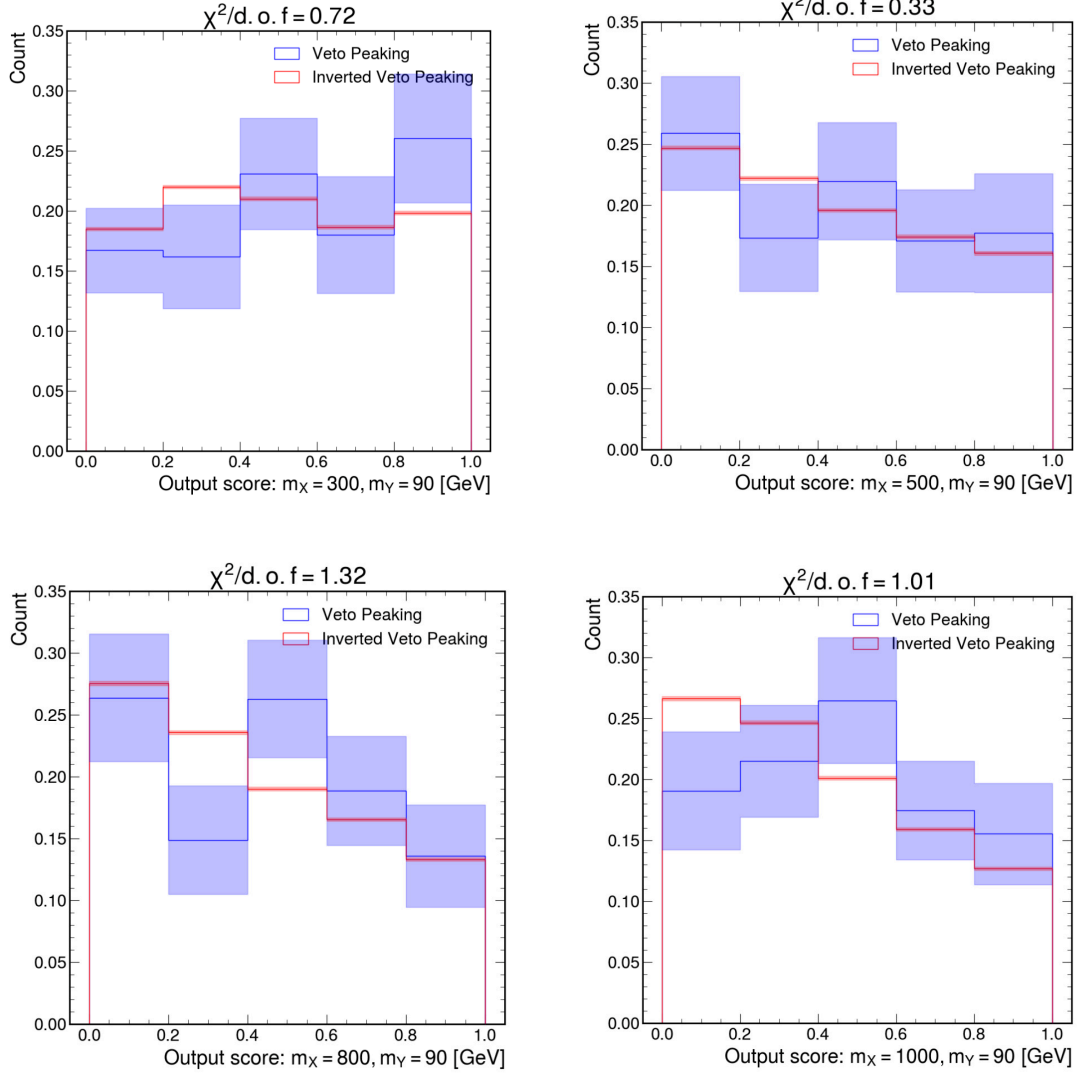


Figure 5.53: Distribution of the transformed pNN output score, $\tilde{f}(\vec{x}; m_X, m_Y)$, in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search for $m_Y = 90$ GeV and $m_X = 300$ (top-left), 500 (top-right), 800 (bottom-left) and 1000 GeV (bottom-right). The distributions are made from simulated DY and diboson events where the reconstructed photons are matched to electrons from $Z \rightarrow e^+e^-$ decays at the generator level, and with the normal preselection (blue) applied, or with the inverted selection (red) applied. The statistical uncertainty is shown by the shaded bands and the compatibility of the two distributions are tested with a χ^2 statistic.

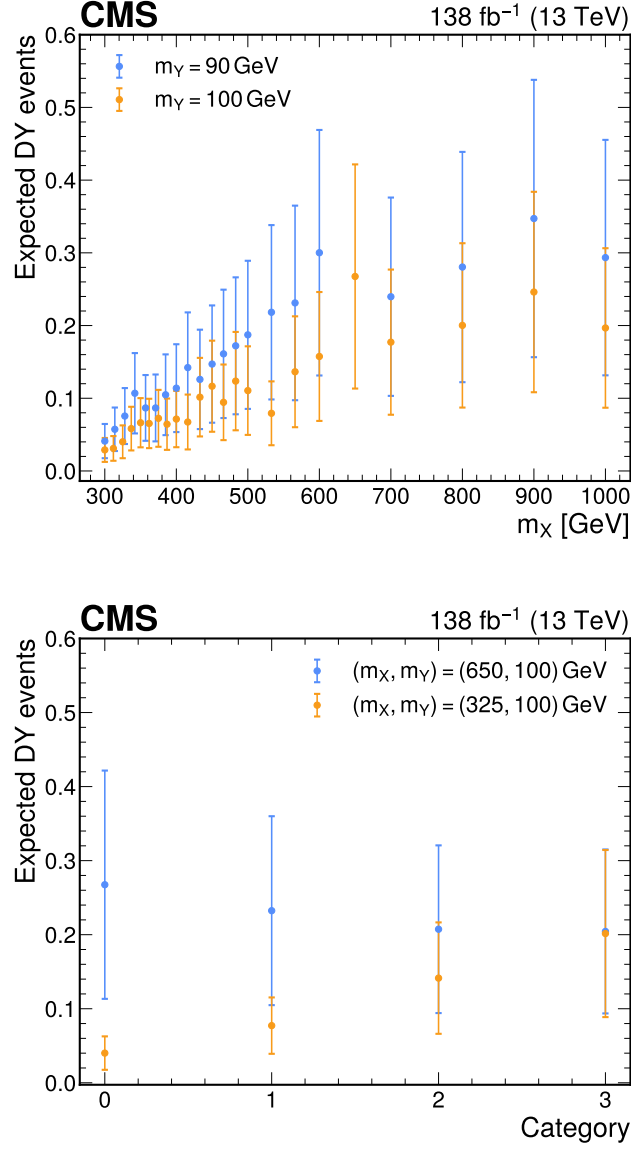


Figure 5.54: Expected number of DY events (yields) in regions D_i (λ_D) extracted from a background-only maximum likelihood fit to data. On the top, the yields are shown for region D_0 as a function of m_X for $m_Y = 90$ and 100 GeV. On the bottom, the yields are shown for regions D_i where i denotes the category, for the $(m_X, m_Y) = (650, 100)$ GeV and $(m_X, m_Y) = (325, 100)$ GeV mass points. The error bars indicate the statistical uncertainty in the yields extracted from the fit.

5.9 Results

In each search, and for every mass point, simultaneous maximum likelihood fits are performed to the $m_{\gamma\gamma}$ distributions in data in the corresponding analysis categories. Goodness-of-fit tests are performed for the background models as part of the procedures described in Sections 5.8.2 and 5.8.3 which ensure a p -value of greater than 0.01 is found in every category.

The calculation of the final results requires the generation and subsequent fitting of $\mathcal{O}(10^9)$ pseudo-datasets. To decrease the fitting time for each dataset, a pruning procedure is applied to the categories defined in Section 5.5.4 to remove categories that do not contribute significantly to the final results. A category's contribution is estimated by $\sigma_i = s_i/b_i$, where s_i and b_i are the signal and background yields in category i in a window centred on the $m_{\gamma\gamma}$ signal peak and of size corresponding to two times the width peak. The pruning procedure keeps categories that have $\sigma_i/\sigma_T > 0.01$ where $\sigma_T = (\sum_i \sigma_i^2)^{1/2}$. The procedure is most impactful for the mass points where the pNN performance is best and the majority of the signal is captured in the first few categories. For these mass points, it can reduce the fit time by up to a factor of 5.

This section describes the results search-by-search, starting with the $X \rightarrow HH$ searches in Section 5.9.1, followed by the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search in Section 5.9.2, and finally the low and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches in Sections 5.9.3 and 5.9.4 respectively.

5.9.1 $X \rightarrow HH$ Searches

In both $X \rightarrow HH$ searches, the observed data is found to be consistent with the background-only hypothesis with no significant excesses found. The largest excess was found in the $X^{(2)} \rightarrow HH$ search at $m_X = 375$ GeV with a local significance of 1.7σ . The $m_{\gamma\gamma}$ distribution in data in every category for this mass point, and the corresponding signal-plus-background model fit, is shown in Fig. 5.55. The $m_{\gamma\gamma}$ distributions appear to be modelled well by the signal-plus-background model and the signal peak due to the $X \rightarrow HH$ process is most noticeable in category 0 which is expected since this is the category designed to have the highest signal efficiency. Beneath the signal peak, the contribution from the single H background can be seen, which is smaller than the $X \rightarrow HH$ signal in category 0, but becomes comparable to the signal in the rest of the categories. Given there are few events in each category, the shape of the background is not well constrained by the data, which can be seen by the uncertainty bands in each plot.

The expected and observed 95% CL upper limits on $\sigma(\text{pp} \rightarrow X \rightarrow HH)$ for the $X^{(0)} \rightarrow HH$ and $X^{(2)} \rightarrow HH$ searches are shown in Fig. 5.56. The limits are found to decrease as a function of m_X , which is expected given the higher preselection signal

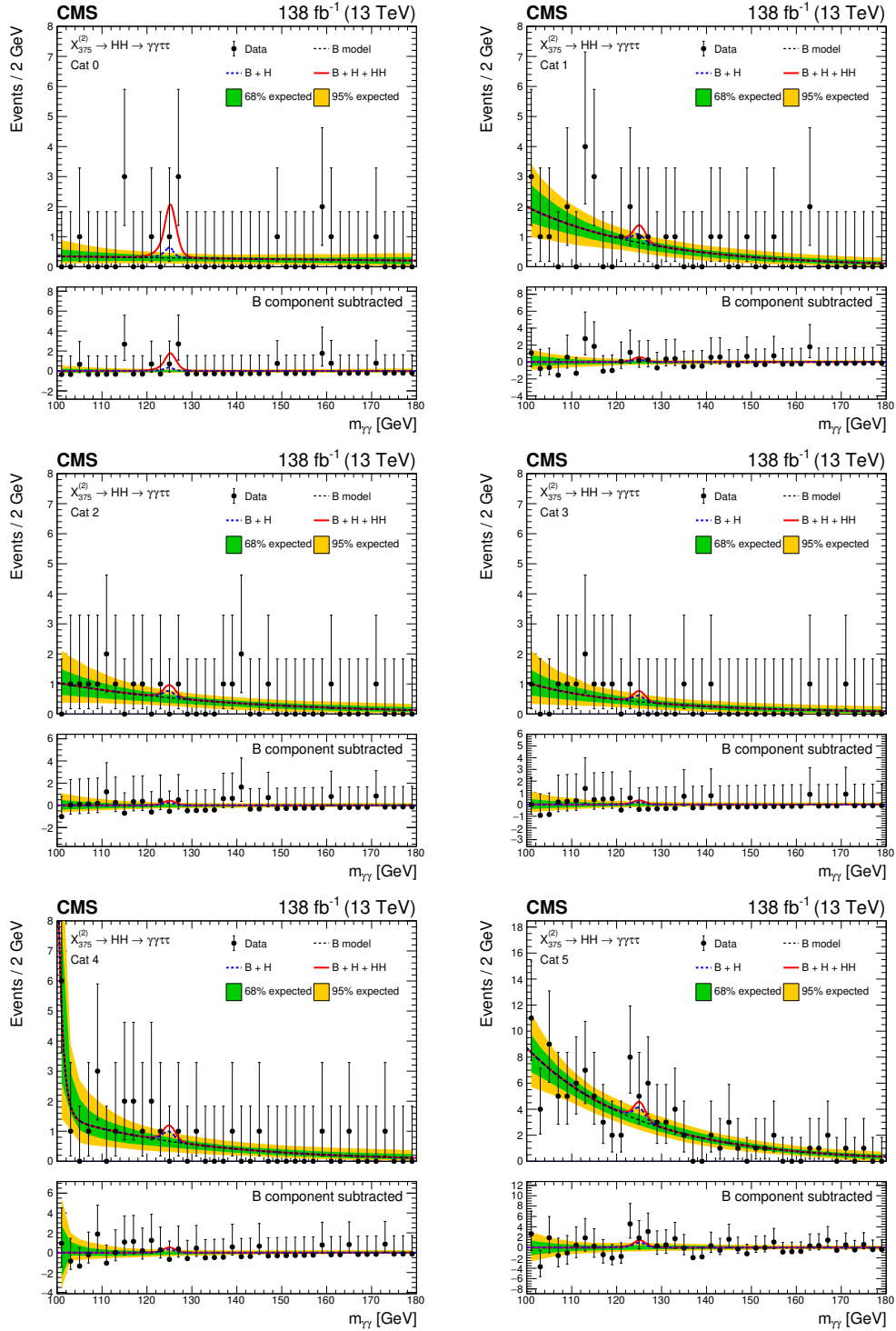


Figure 5.55: Distributions of $m_{\gamma\gamma}$ in data in each analysis category and the signal-plus-background models (red) with data (black points) for the mass hypothesis with the largest excess in the $X^{(2)} \rightarrow HH$ search: $m_X = 375$ GeV. The one (green) standard deviation and two (yellow) standard deviation bands show the uncertainties in the nonresonant background component of the model (black dashed line). The resonant single-Higgs background is plotted separately in blue (blue dashed line). The lower panel shows the residuals after subtraction of the nonresonant background component.

efficiency and better signal-to-background discrimination seen at higher masses. The observed (expected) upper limits vary between 160–2200 fb (240–1800 fb) and 180–1900 fb (200–1800 fb) for the $X^{(0)} \rightarrow HH$ and $X^{(2)} \rightarrow HH$ searches respectively. Interpreting these results in the RS bulk model (Section 2.3.1), a spin-0 Radion resonance with mass of up to 550 GeV is excluded when $\Lambda_R = 3$ TeV, or up to 900 GeV is excluded when $\Lambda_R = 2$ TeV. For the spin-2 KK Graviton, a resonance with a mass between 310 and 700 GeV is excluded when assuming $\tilde{\kappa} = 1$.

In an analogous search performed by the CMS experiment in the $b\bar{b}\gamma\gamma$ final state [95], the observed upper limits on $\sigma(pp \rightarrow X \rightarrow HH)$ were found to be about 26–310 fb and 23–290 fb for the $X^{(0)} \rightarrow HH$ and $X^{(2)} \rightarrow HH$ searches respectively. Therefore, the $\gamma\gamma\tau\tau$ final state does not provide the most competitive results in the $X \rightarrow HH$ searches, but this was expected given the smaller branching fraction of $H \rightarrow \tau\tau$ relative to $H \rightarrow b\bar{b}$. Nonetheless, the $\gamma\gamma\tau\tau$ final state can be used in a future combination of final states that should provide a sensitivity greater than any final state alone. In the $X \rightarrow YH$ searches, which are discussed in the rest of this section, the branching fractions of the Y scalar are unspecified and therefore, these searches can yield interesting standalone results.

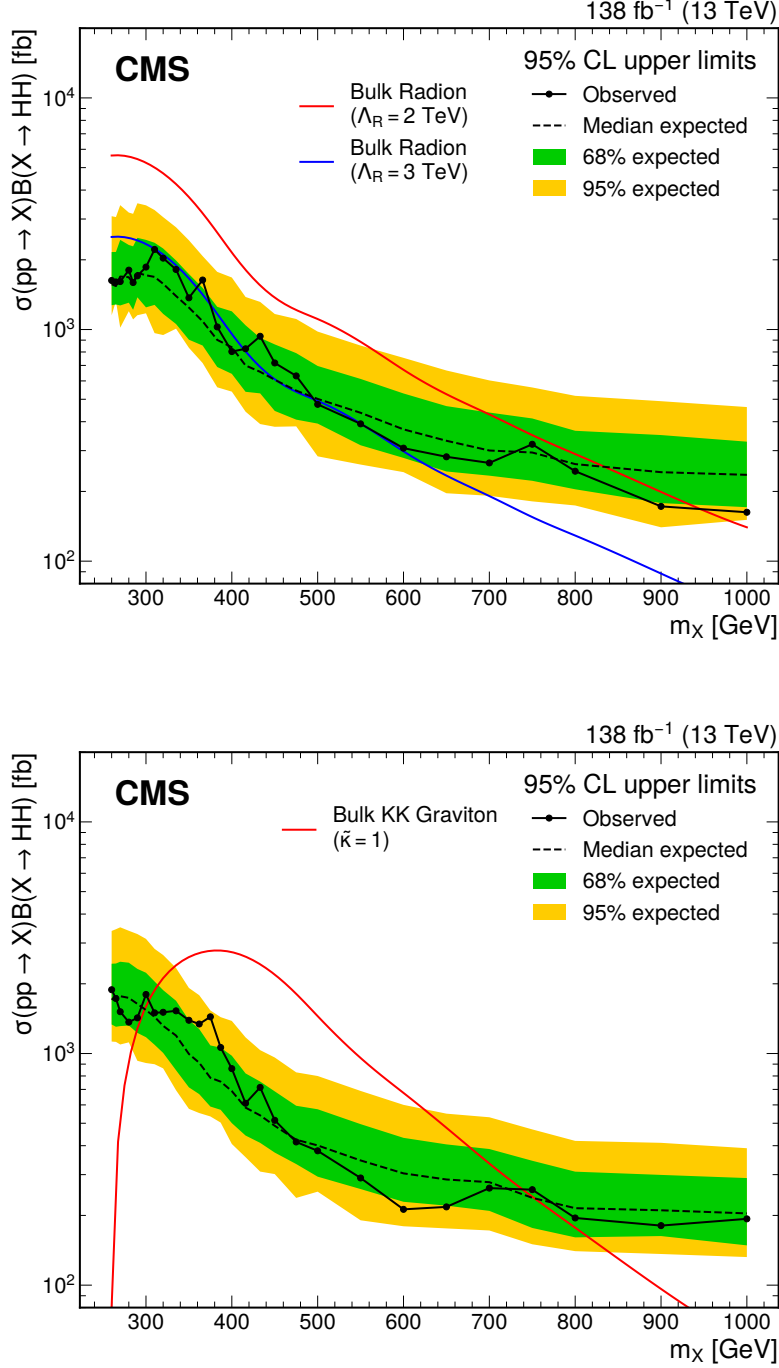


Figure 5.56: Expected and observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow HH)$ for the $X^{(0)} \rightarrow HH$ (top) and $X^{(2)} \rightarrow HH$ (bottom) searches. The solid and dashed black lines represent observed and median expected limits respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68 and 95%, respectively, of the distribution of limits expected under the background-only hypothesis. The red and blue lines show WED theory predictions at values of the theory parameters specified in the legend.

5.9.2 $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ Search

Figure 5.57 shows the local significances for every mass point in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search. Local significances of about 2 standard deviations can be seen for a band in m_X and m_Y that stretches between $m_X = 300$ and 500 GeV and for values of $(m_X - m_Y) \approx 200$ GeV. The largest excess is seen at $(m_X, m_Y) = (320, 60)$ GeV with a local (global) significance of 2.6 (2.2) standard deviations. The $m_{\gamma\gamma}$ distributions in data for this mass point, and the corresponding signal-plus-background model fit, is shown in Fig. 5.58. Looking at these figures, the excess seems to be driven primarily by the 6 events in the interval $120 < m_{\gamma\gamma} < 130$ GeV in category 0, although there are also noticeable, albeit smaller, excesses in categories 3 and 5.

The observed 95% CL upper limits on $\sigma(\text{pp} \rightarrow X \rightarrow YH \rightarrow \gamma\gamma\tau\tau)$ for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search are shown in a 2D heatmap in Fig. 5.59, and the expected and observed limits are shown in slices of m_X and m_Y in Figs. 5.60 and 5.61, respectively, where the slices are taken at the nominal mass points. Like in the $X \rightarrow HH$ searches, the lowest limits are set at higher values of m_X , and the highest limits are found at lower m_X , particularly in the band where the highest significances are found. There is less of a noticeable trend in m_Y , which is reflected in the preselection signal efficiency (Fig. 5.6). The observed (expected) upper limits vary between 0.059–1.2 fb (0.087–0.68 fb), depending on the mass point.

In the $b\bar{b}\gamma\gamma$ final state [95], observed upper limits on the equivalent cross section: $\sigma(\text{pp} \rightarrow X \rightarrow YH \rightarrow b\bar{b}\gamma\gamma)$, were found to be about 0.04–0.90 fb. Therefore, for models where $\mathcal{B}(Y \rightarrow \tau\tau)$ and $\mathcal{B}(Y \rightarrow b\bar{b})$ are similar, the $\gamma\gamma\tau\tau$ final state provides a similar sensitivity to the $b\bar{b}\gamma\gamma$ final state. Furthermore, this search in the $\gamma\gamma\tau\tau$ state searched for Y resonances with masses down to 50 GeV whereas the $b\bar{b}\gamma\gamma$ search only went down to 70 GeV.

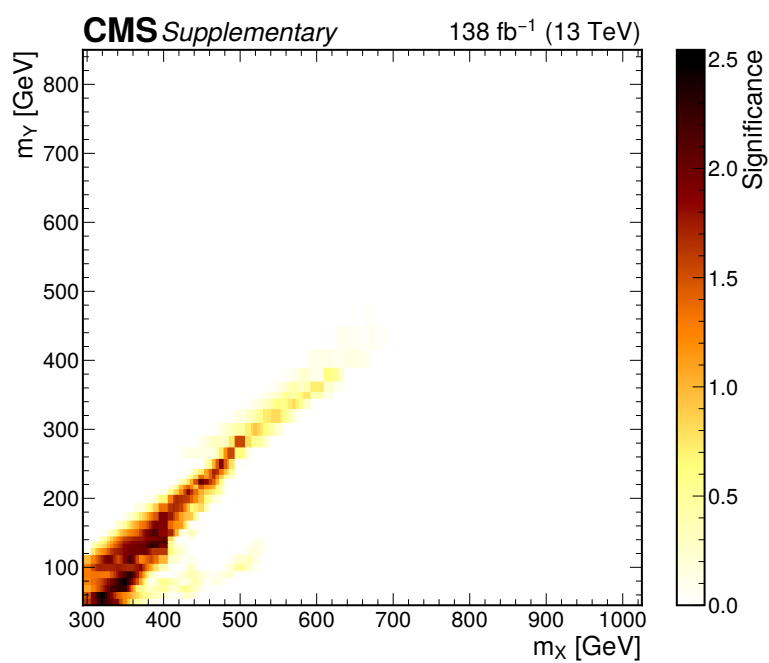


Figure 5.57: Observed local significances in the 2D (m_X, m_Y) plane for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search.

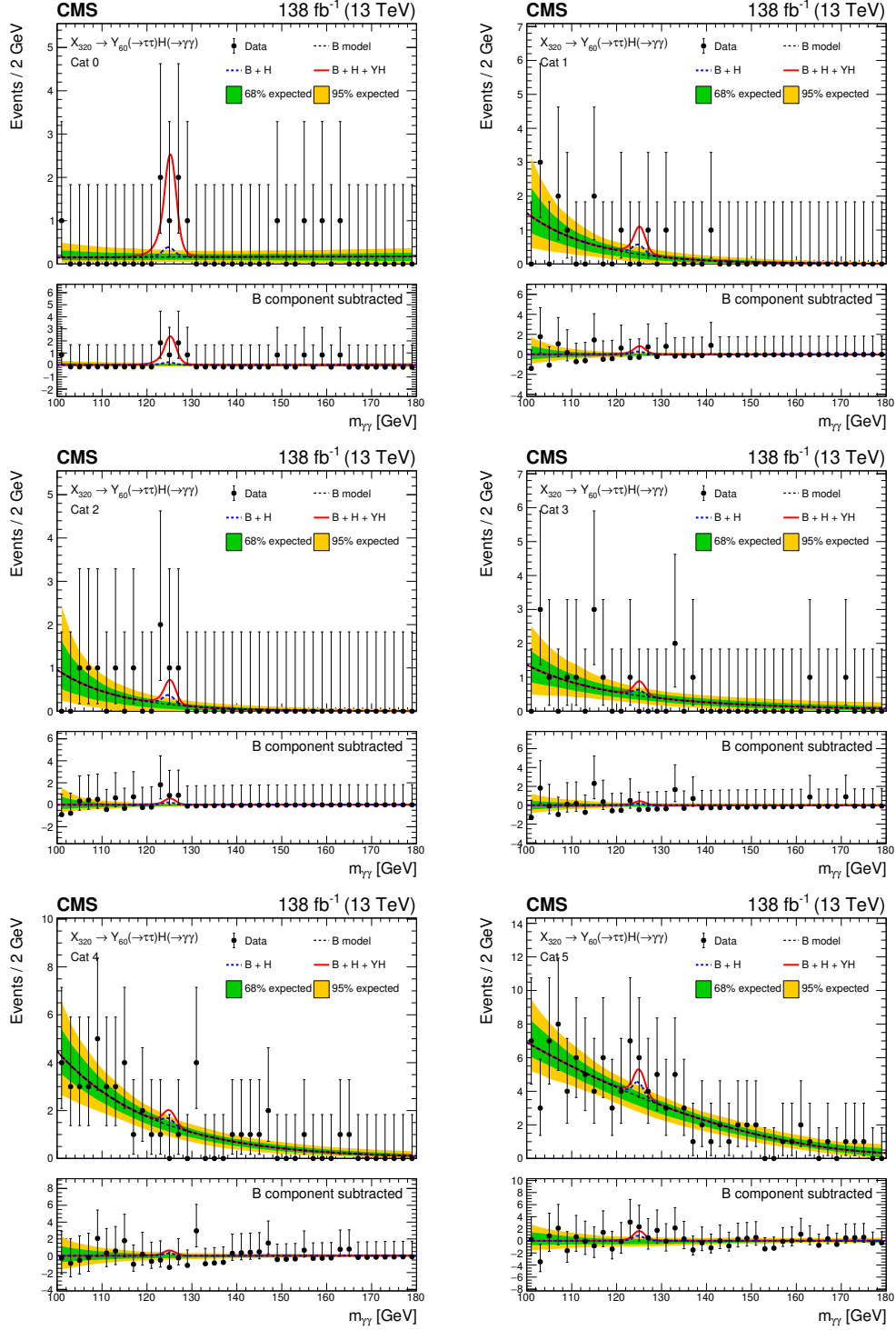


Figure 5.58: Distributions of $m_{\gamma\gamma}$ in data in each analysis category and the signal-plus-background models (red) with data (black points) for the mass hypothesis with the largest excess in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search: $(m_X, m_Y) = (320, 60)$ GeV. The one (green) standard deviation and two (yellow) standard deviation bands show the uncertainties in the nonresonant background component of the model (black dashed line). The resonant single-Higgs background is plotted separately in blue (blue dashed line). The lower panel shows the residuals after subtraction of the nonresonant background component.

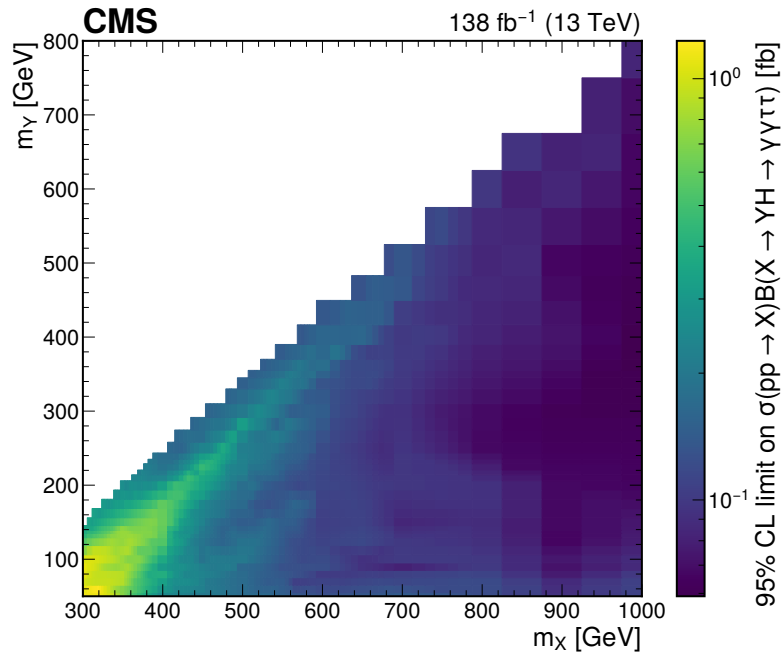


Figure 5.59: Observed 95% CL upper limits on $\sigma(\text{pp} \rightarrow X \rightarrow YH \rightarrow \gamma\gamma\tau\tau)$ in the 2D (m_X, m_Y) plane for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search.

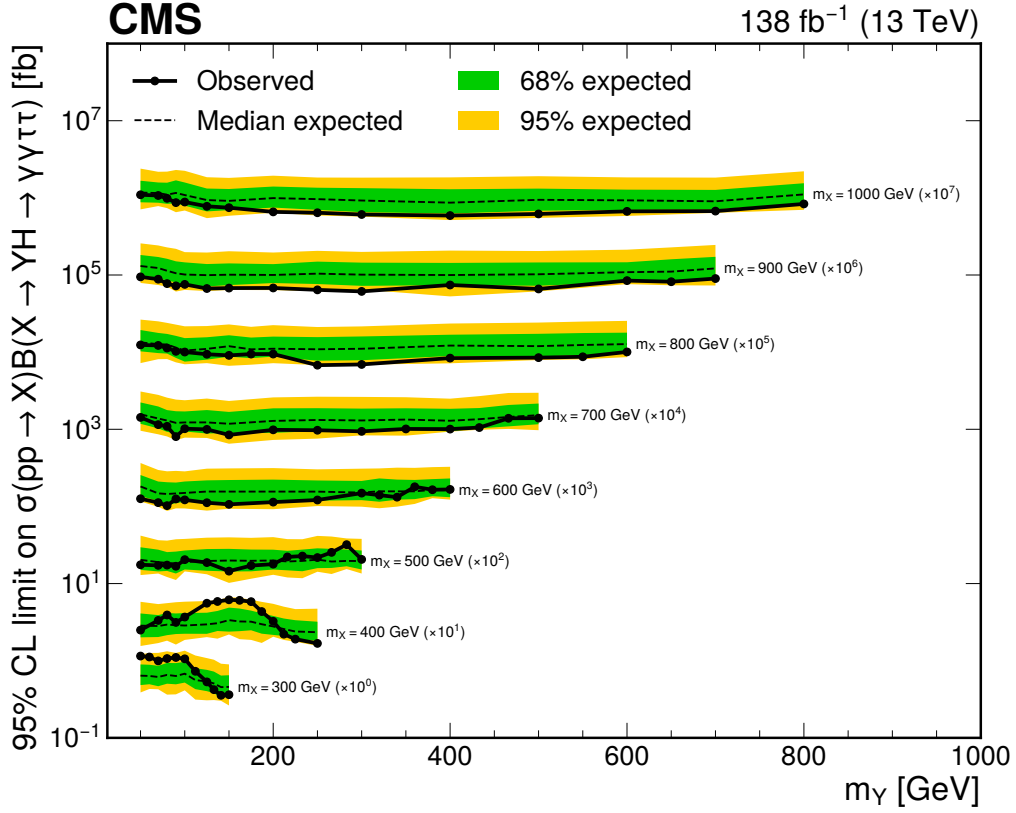


Figure 5.60: Expected and observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH \rightarrow \gamma\gamma\tau\tau)$ for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search. Limits are shown as a function of m_γ for the nominal values of m_X where the limits are scaled by orders of 10 as labelled in the plot. The solid and dashed black lines represent observed and median expected limits respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68 and 95%, respectively, of the distribution of limits expected under the background-only hypothesis.

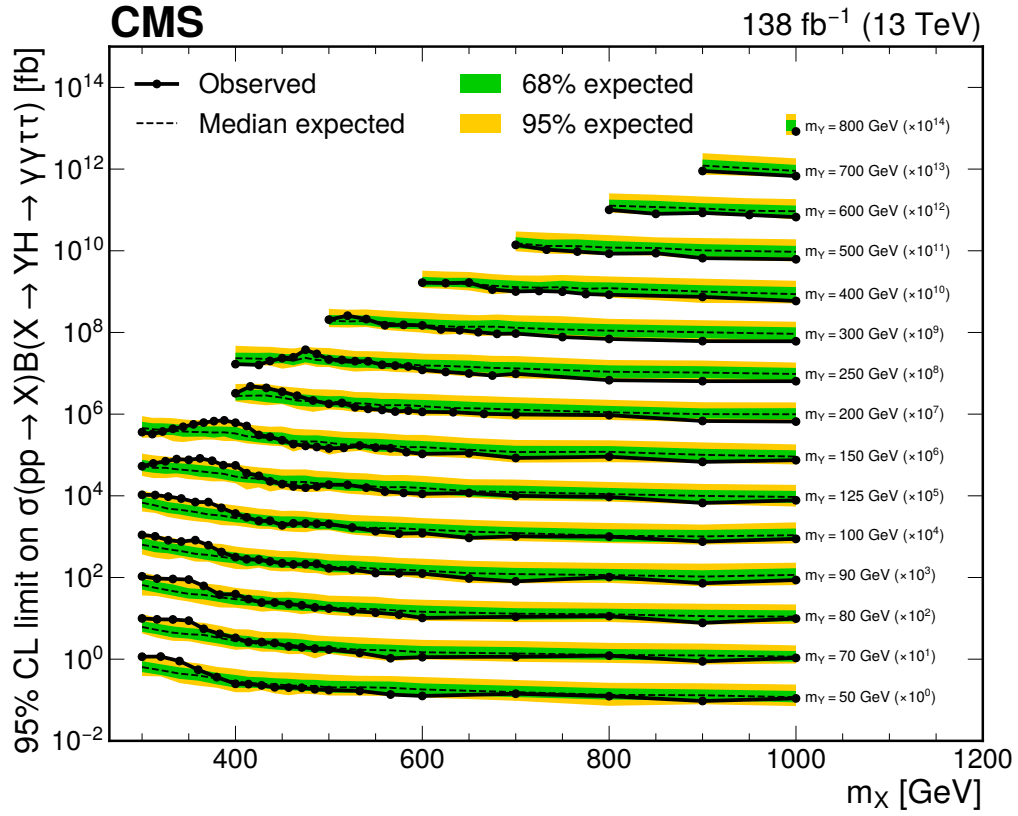


Figure 5.61: Expected and observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH \rightarrow \gamma\gamma\tau\tau)$ for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search. Limits are shown as a function of m_X for the nominal values of m_Y where the limits are scaled by orders of 10 as labelled in the plot. The solid and dashed black lines represent observed and median expected limits respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68 and 95%, respectively, of the distribution of limits expected under the background-only hypothesis.

5.9.3 Low-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search

Figure 5.62 shows the local significances for every mass point in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Compared to the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search, there is less structure in the significance plot, owing to the fact that changing m_Y now changes the location of the $m_{\gamma\gamma}$ peak in the signal model, as well as the pNN selection as before. The largest excess is seen at $(m_X, m_Y) = (525, 115)$ GeV, with a local significance of 3.2 standard deviations. The $m_{\gamma\gamma}$ distributions in data for this mass point, and the corresponding signal-plus-background model fit, is shown in Fig. 5.63. The excess is driven primarily by a low background prediction and 2 events near $m_{\gamma\gamma} = 115$ GeV in category 0.

There is also an excess with a significance of 2.3 standard deviations at $(m_X, m_Y) = (650, 95)$ GeV, which is particularly interesting due to recent excesses seen by the CMS experiment at similar mass points (see Section 5.1). Importantly, from Fig. 5.62, the excess also appears to be fairly localized to $(m_X, m_Y) = (650, 95)$ GeV, which gives the coincidence of the excesses in different final states more weight. The $m_{\gamma\gamma}$ distributions in data for this mass point, and the corresponding signal-plus-background model fit, is shown in Fig. 5.64. The excess is similarly driven by a low background prediction and 2 events near $m_{\gamma\gamma} = 95$ GeV in category 0.

At this mass point, m_Y is closer to m_Z and therefore, the DY background modelling may be more important. However, the expected number of DY events is at most 0.27 events across all categories, and noticeably smaller than the extracted signal peak in category 0 seen in Fig. 5.64 (top-right). Given that the normalization of the DY background is small, and that its peak does not have significant overlap with the signal peak, the impact on the significance is small, where removing the DY background model entirely changes the significance by less than 0.1 standard deviations. At other mass points, where there is greater overlap between the DY background and the signal, the impact on the results is still small. For the $(m_X, m_Y) = (1000, 90)$ GeV mass point, which has one of the largest expected number of DY events (Section 5.8.3), the observed and expected upper limits change by 2% when removing the DY background modelling.

The global significance of the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search is only 0.1 standard deviations. Compared to the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search, the look-elsewhere effect is larger due to the greater (about a factor of five) number of mass points searched at. Therefore, as a standalone result, the excesses at $(m_X, m_Y) = (525, 115)$ GeV and $(650, 95)$ GeV are not particularly interesting. However, the excess at $(m_X, m_Y) = (650, 95)$ GeV is interesting given the recent excesses seen with 650 and ~ 95 GeV resonances (Section 5.1).

The observed 95% CL upper limits on $\sigma(\text{pp} \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search are shown in a 2D heatmap in Fig. 5.65, and the expected and observed limits are shown in slices of m_X and m_Y in Figs. 5.66 and 5.67, respectively,

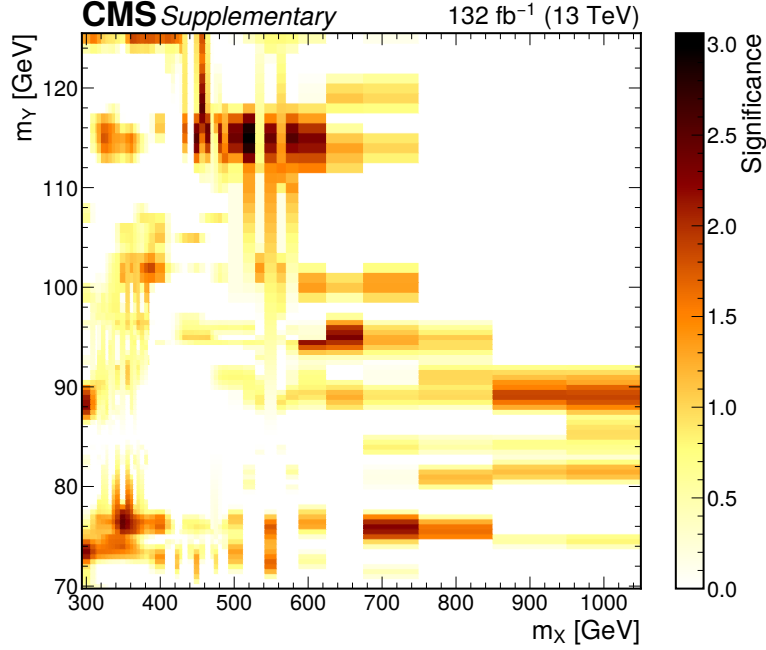


Figure 5.62: Observed local significances in the 2D (m_X, m_Y) plane for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search.

where the slices are taken at the nominal mass points. Similar trends in m_X and m_Y are seen in the limits as in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ search, and the observed (expected) upper limits vary between 0.69–15 fb (0.73–8.3 fb), depending on the mass point.

The observed limits are compared to the maximally-allowed values in the NMSSM given experimental constraints (Section 2.3.2), and the region of masses where the limit is below the maximally-allowed value is shown in Fig. 5.65. For most values of m_Y in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, this region corresponds to $m_X < 650$ GeV, or for all values of m_Y in the search, this region corresponds to $m_X < 620$ GeV. This implies that these results can be used to apply tighter constraints on the NMSSM parameter space than previously possible. Deriving those constraints is beyond the scope of this thesis, but could be done in a future iteration of the work presented in Ref. [30].

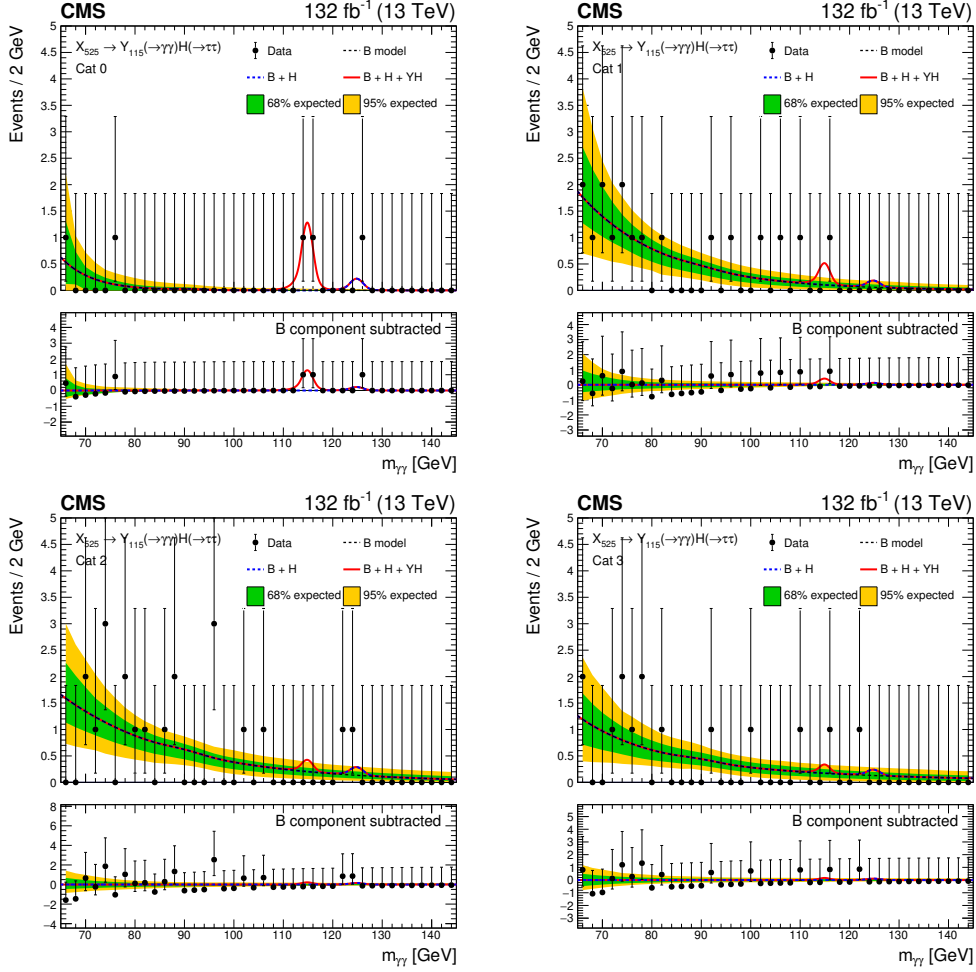


Figure 5.63: Distributions of $m_{\gamma\gamma}$ in data in each analysis category corresponding to a signal region and the signal-plus-background models (red) with data (black points) for the mass hypothesis with the largest excess in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search: $(m_X, m_Y) = (525, 115)$ GeV. The one (green) standard deviation and two (yellow) standard deviation bands show the uncertainties in the nonresonant + DY background component of the model (black dashed line). The resonant single-Higgs background is plotted separately in blue (blue dashed line). The lower panel shows the residuals after subtraction of the nonresonant background component.

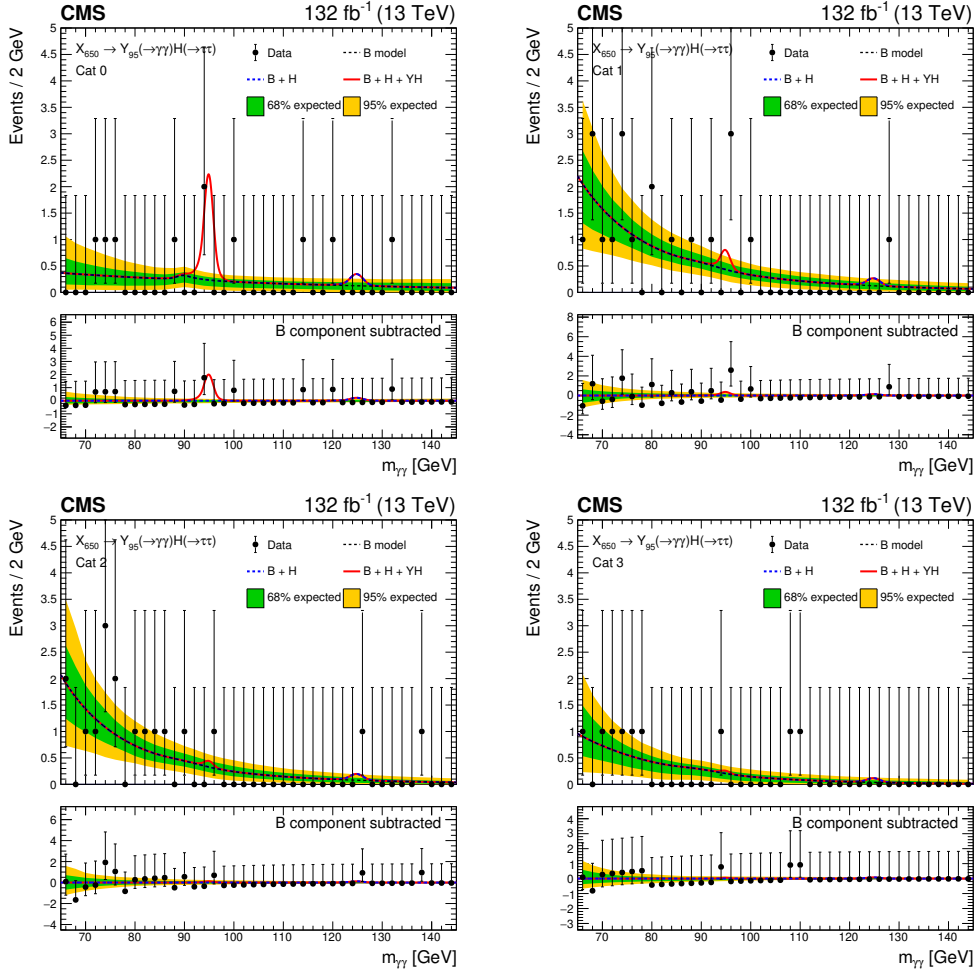


Figure 5.64: Distributions of $m_{\gamma\gamma}$ in data in each analysis category corresponding to a signal region and the signal-plus-background models (red) with data (black points) in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search for $(m_X, m_Y) = (650, 95)$ GeV. The one (green) standard deviation and two (yellow) standard deviation bands show the uncertainties in the nonresonant + DY background component of the model (black dashed line). The resonant single-Higgs background is plotted separately in blue (blue dashed line). The lower panel shows the residuals after subtraction of the nonresonant background component.

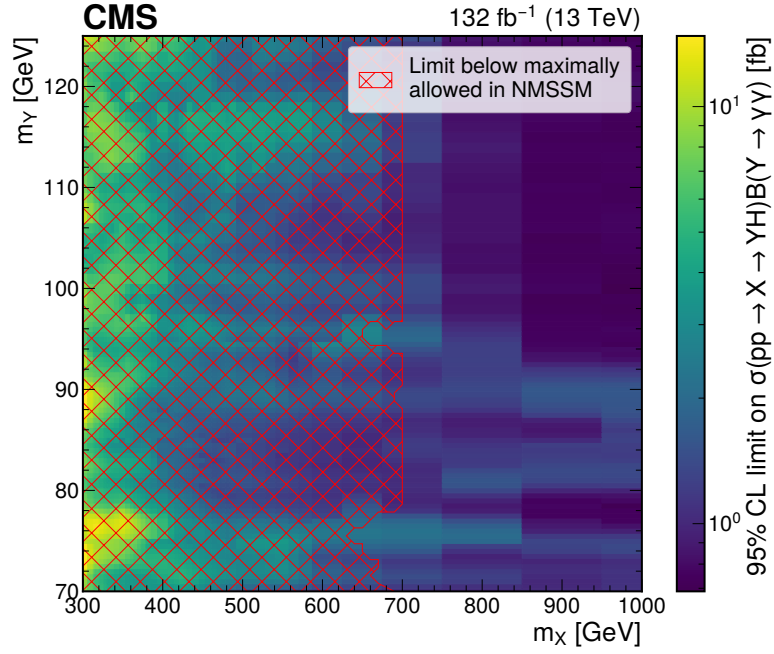


Figure 5.65: Observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ in the 2D (m_X, m_Y) plane for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. The region of points where the observed limit is below the maximally-allowed values in the NMSSM given experimental constraints (Section 2.3.2) is indicated by the red hatching.

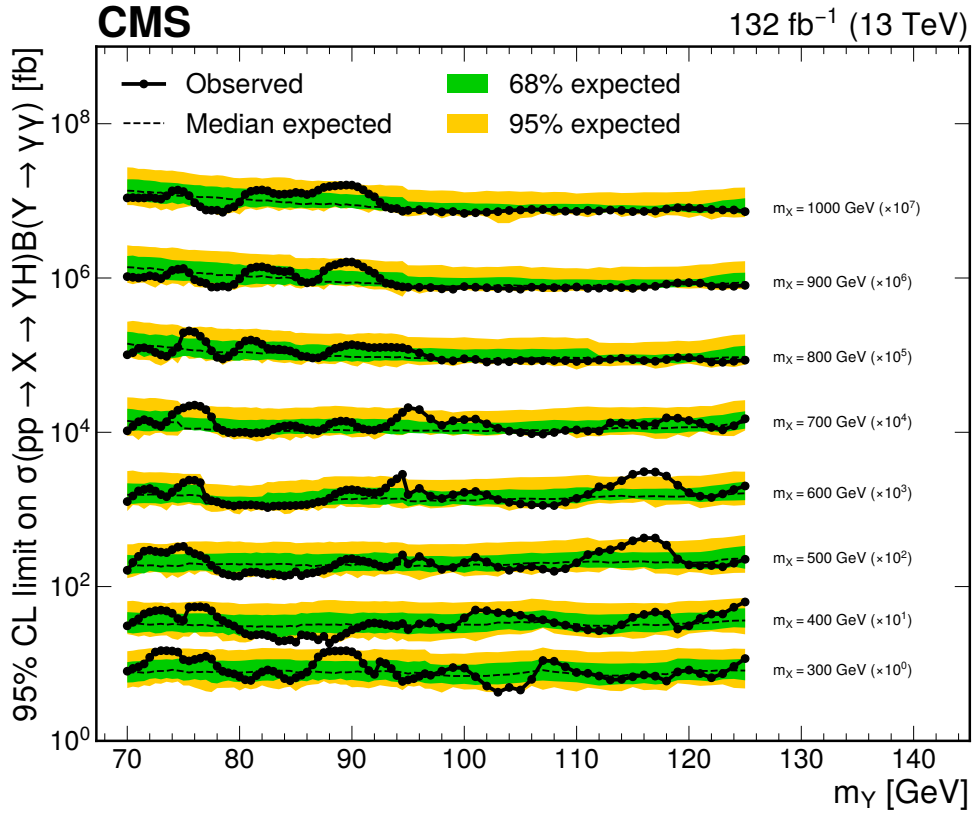


Figure 5.66: Expected and observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Limits are shown as a function of m_Y for the nominal values of m_X where the limits are scaled by orders of 10 as labelled in the plot. The solid and dashed black lines represent observed and median expected limits respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68 and 95%, respectively, of the distribution of limits expected under the background-only hypothesis.

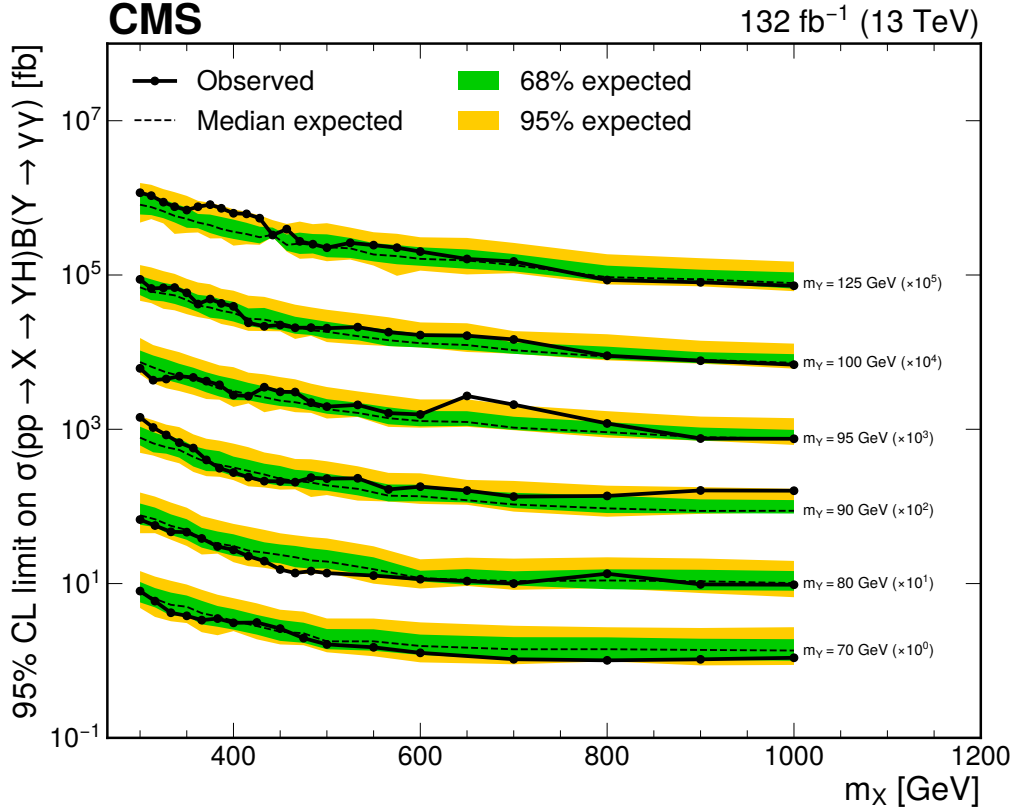


Figure 5.67: Expected and observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ for the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Limits are shown as a function of m_X for the nominal values of m_Y where the limits are scaled by orders of 10 as labelled in the plot. The solid and dashed black lines represent observed and median expected limits respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68 and 95%, respectively, of the distribution of limits expected under the background-only hypothesis.

5.9.4 High-Mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ Search

Figure 5.62 shows the local significances for every mass point in the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. The largest excess is seen at $(m_X, m_Y) = (450, 161)$ GeV, with a local significance of 3.2 standard deviations. As in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search, the large number of mass points searched at leads to a large look-elsewhere effect, and the global significance is 0.3 standard deviations. The $m_{\gamma\gamma}$ distributions in data for this mass point, and the corresponding signal-plus-background model fit, is shown in Fig. 5.69. In every category, there is a small background prediction, and at least one event near the signal peak.

The observed 95% CL upper limits on $\sigma(\text{pp} \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ for the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search are shown in a 2D heatmap in Fig. 5.70, and the expected and observed limits are shown in slices of m_X and m_Y in Figs. 5.71 and 5.72, respectively, where the slices are taken at the nominal mass points. Compared to the other $X \rightarrow YH$ searches, there is a greater dependence of m_Y on the limits, where the expected limits tend to increase with m_Y . This can be explained by the preselection signal efficiencies (Fig. 5.6) which see a similar trend in m_Y . The observed (expected) upper limits vary between 0.64–10 fb (0.70–7.6 fb), depending on the mass point.

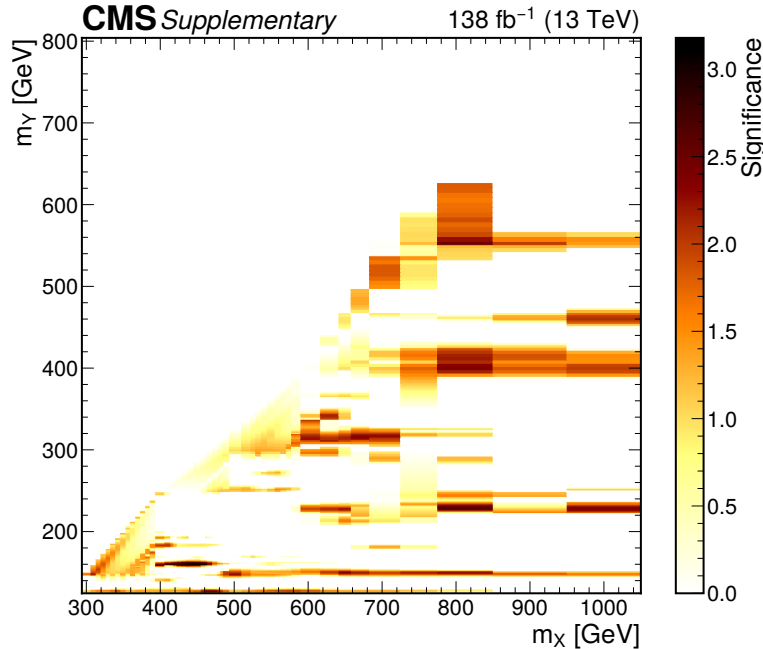


Figure 5.68: Observed local significances in the 2D (m_X, m_Y) plane for the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search.

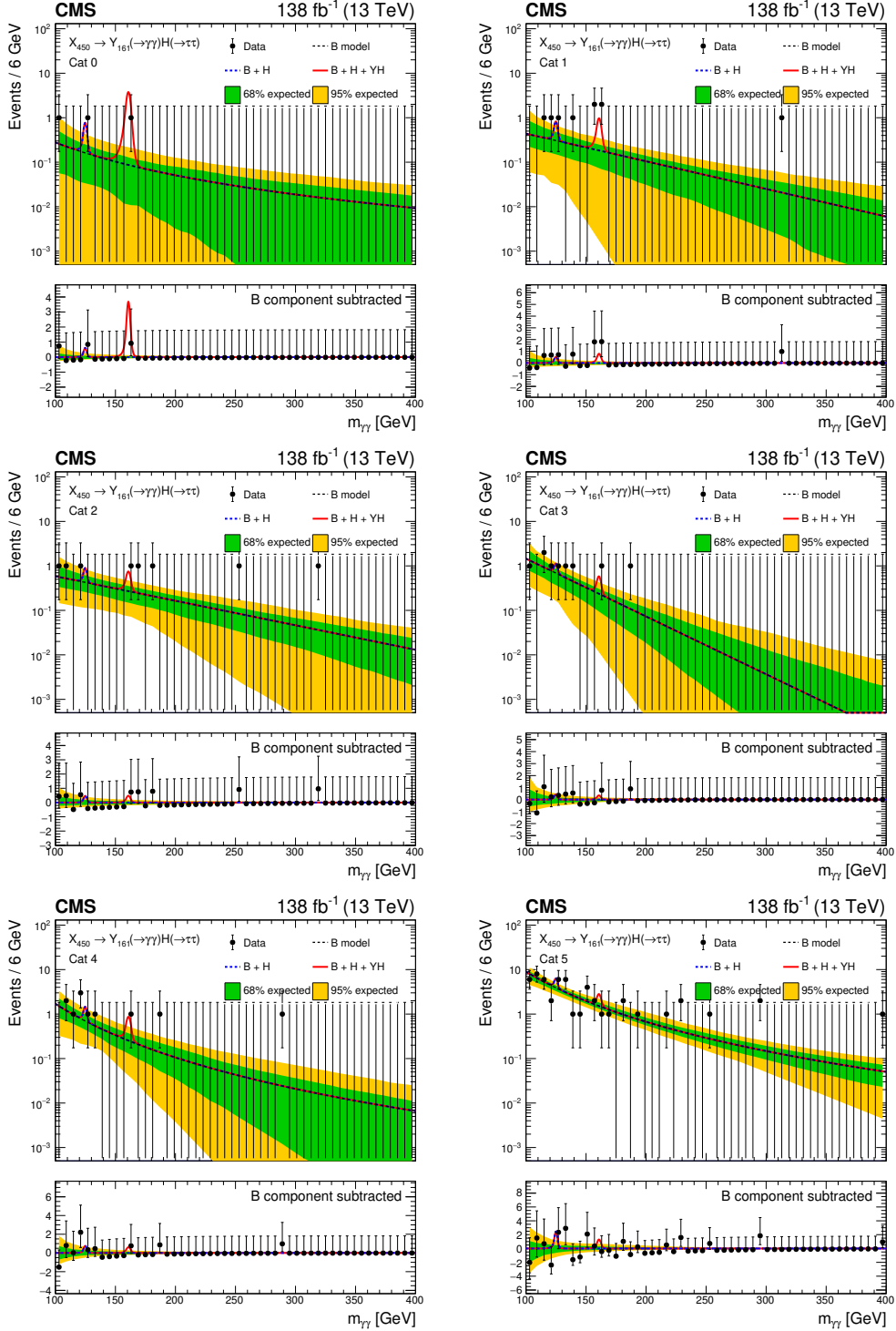


Figure 5.69: Distributions of $m_{\gamma\gamma}$ in data in each analysis category and the signal-plus-background models (red) with data (black points) in the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search for $(m_X, m_Y) = (462, 161)$ GeV. The one (green) standard deviation and two (yellow) standard deviation bands show the uncertainties in the nonresonant background component of the model (black dashed line). The resonant single-Higgs background is plotted separately in blue (blue dashed line). The lower panel shows the residuals after subtraction of the nonresonant background component.

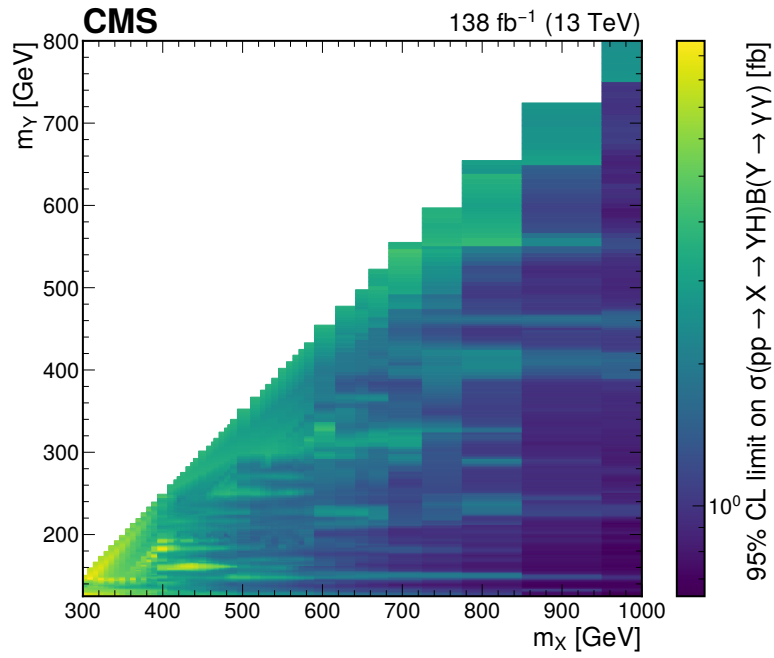


Figure 5.70: Observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ in the 2D (m_X, m_Y) plane for the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search.

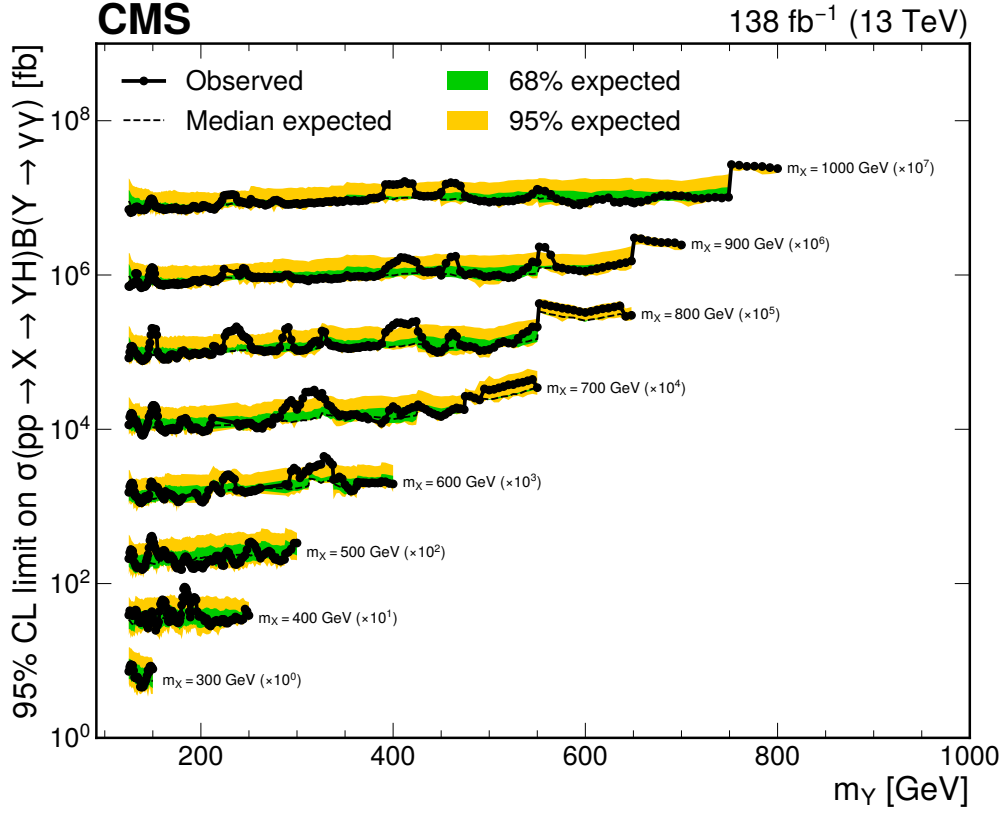


Figure 5.71: Expected and observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ for the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Limits are shown as a function of m_Y for the nominal values of m_X where the limits are scaled by orders of 10 as labelled in the plot. The solid and dashed black lines represent observed and median expected limits respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68 and 95%, respectively, of the distribution of limits expected under the background-only hypothesis.

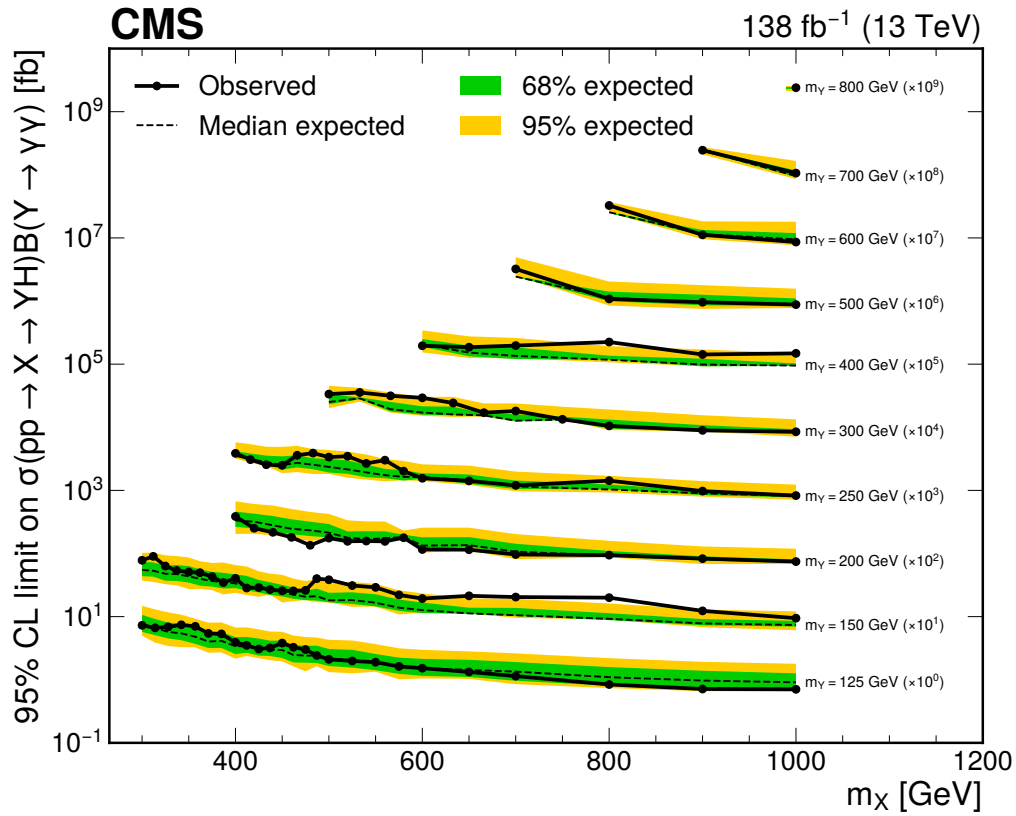


Figure 5.72: Expected and observed 95% CL upper limits on $\sigma(pp \rightarrow X \rightarrow YH)\mathcal{B}(Y \rightarrow \gamma\gamma)$ for the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search. Limits are shown as a function of m_X for the nominal values of m_Y where the limits are scaled by orders of 10 as labelled in the plot. The solid and dashed black lines represent observed and median expected limits respectively. The inner (green) band and the outer (yellow) band indicate the regions containing 68 and 95%, respectively, of the distribution of limits expected under the background-only hypothesis.

5.10 Summary

This chapter presented the search for the resonant production of a pair of scalar bosons in the $\gamma\gamma\tau\tau$ final state via two primary processes: $pp \rightarrow X \rightarrow HH$ where X is a new boson that can be spin-0 or spin-2, and $pp \rightarrow X \rightarrow YH$ where X and Y are new scalar bosons, and Y is considered to decay to two photons or to two tau leptons. These two processes are motivated by the WED and NMSSM theories respectively. The searches were performed using data collected by the CMS experiment between 2016 and 2018 at the LHC at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 138 fb^{-1} .

Events were first selected with diphoton triggers and then a loose preselection on diphoton and ditau candidates was applied. Parametric neural network (pNN) classifiers were trained to separate signal events from background events in each of the $X \rightarrow HH$ and $X \rightarrow YH$ searches. The pNNs are parametric in m_X for the $X \rightarrow HH$ searches, and m_X and m_Y for the $X \rightarrow YH$ searches, meaning that the pNN can specialize its classification behaviour to a chosen mass point, resulting in good performance at every mass point. The granularity in m_X and m_Y that searches were performed at was optimized to reduce the possibility of missing an excess in the mass ranges for the search, which were 260–1000 GeV for m_X , and 50–800 GeV for m_Y . At every mass point, analysis categories were defined based upon the pNN score, where the boundaries on the pNN score were optimized for the best expected sensitivity.

In every analysis category, statistical models were created to describe the $m_{\gamma\gamma}$ distribution of the signal and background processes. The signal models were derived from simulated events, and interpolation methods were used to create models for mass points that were not simulated. The models of the SM single Higgs production background and of the nonresonant background were derived following typical techniques for $H \rightarrow \gamma\gamma$ analyses at the CMS experiment, whereas the DY background originating from the misidentification of electrons as photons was modelled with a data-driven ABCD method developed specifically for this analysis.

Observed significances and upper limits were extracted from maximum likelihood fits to the $m_{\gamma\gamma}$ distributions in the analysis categories and these results are summarized in Table 5.13. The largest excesses were found in the $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches at $(m_X, m_Y) = (525, 155) \text{ GeV}$ and $(m_X, m_Y) = (450, 161) \text{ GeV}$ corresponding to local significances of 3.2 standard deviations for both. The global significances for the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$, and the low and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches were 2.2, 0.1, and 0.3 standard deviations respectively. Therefore, as a standalone result, these searches do not provide any significant evidence for physics beyond the SM. However, given recent excesses reported in other CMS analyses, the excess at $(m_X, m_Y) = (650, 95) \text{ GeV}$ in the

Table 5.13: Summary of results from the $X \rightarrow HH$ and $X \rightarrow YH$ searches. The masses of the largest excesses and the corresponding local significances are shown as well as the global significance of the search where it has been calculated. The ranges of m_X that are excluded for the Radion, $X^{(0)}$, and Graviton, $X^{(2)}$, are shown, as is the range of m_X in the low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search where the observed limits are below their maximally-allowed values given experimental constraints [30].

Search	Largest Excess	Global Sig.	Exclusion Range / Limits Below Maximally Allowed
$X^{(0)} \rightarrow HH$	1.3σ at $m_X = 366$ GeV	—	$m_X < 900$ GeV ($\Lambda_R = 3$ TeV) $m_X < 550$ GeV ($\Lambda_R = 2$ TeV)
$X^{(2)} \rightarrow HH$	1.7σ at $m_X = 375$ GeV	—	$310 < m_X < 900$ GeV ($\tilde{\kappa} = 1$)
$X \rightarrow Y(\tau\tau)H(\gamma\gamma)$	2.6σ at $(m_X, m_Y) = (320, 60)$ GeV	2.2σ	—
Low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	3.2σ at $(m_X, m_Y) = (525, 115)$ GeV	0.1σ	$m_X < 620$ GeV for $70 < m_Y < 125$
High-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$	3.2σ at $(m_X, m_Y) = (450, 161)$ GeV	0.3σ	—

$X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ search is important.

In a similar $X \rightarrow YH$ search in the $b\bar{b}\gamma\gamma$ final state, the CMS experiment reported an excess with a local significance of 3.8 standard deviations at $(m_X, m_Y) = (650, 90)$ GeV [95]. Furthermore, in searches for new scalars decaying to WW , $\tau\tau$, and $\gamma\gamma$ final states the CMS experiment reported local significances of 3.8, 2.8 and 2.9 standard deviations for masses of 650 GeV, 100 GeV and 95 GeV respectively [96–98]. For the search in the $\gamma\gamma\tau\tau$ final state described in this chapter, a similar excess was observed with a local significance of 2.3 standard deviations at $(m_X, m_Y) = (650, 95)$ GeV. Put together, these results form a pattern that is worth investigating further, either from analyses in other final states, or from the analysis of future data collected at the LHC.

Chapter 6

Effective Field Theory Interpretation

6.1 Introduction

So far at the LHC, no direct evidence for new physics has been observed with a significance greater than 5 standard deviations. It could be that the searches are not looking in the right direction, e.g. all that is needed is to search in a particular final state, or consider an exotic signature, like a long-lived particle, which are less commonly explored at the LHC. Alternatively, it could just be that the couplings in the new theory are small such that the new particles are produced at a rate that is too low to be observed with the current data set. Optimistically, one could then hope that the new particles are within reach at the end of the LHC's lifetime.

However, it is also possible that the mass of the new particle is larger than the LHC energy reach, in which case, the only evidence of new physics provided by the LHC would be through indirect effects of new particles, which can be parameterized by the Wilson coefficients in an Effective Field Theory (EFT). The observation of a Wilson coefficient taking a non-zero value would be an indication of new physics, and the compatibility of that observation with BSM theories can be assessed via a procedure called *matching* [120] where the constraints on Wilson coefficients are translated into constraints on the parameters of a BSM theory.

The indirect effects predicted by an EFT can be seen in SM processes as a deviation in the higher-energy tail of a distribution, as illustrated in Fig. 6.1, or as an overall scaling of processes, or even exceptionally, as a deviation in the lower-energy tails of a distribution, as in the case of the $H \rightarrow 4l$ decay discussed in Section 2.3.3. Therefore, indirect hints of new physics could be present in the large variety of SM measurements already made at the

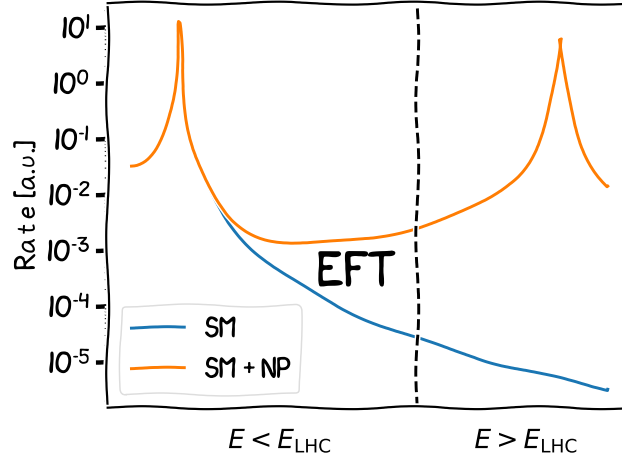


Figure 6.1: Toy examples of invariant mass distributions, one which is from a SM resonance (blue), and another that also includes a new resonance (orange). The dashed line indicates the boundary of the LHC’s energy reach which is $\mathcal{O}(1 \text{ TeV})$. Given that the new resonance is beyond that line, its direct detection would not be possible at the LHC, but its indirect effects, which are visible to the left of the dashed line, could be detected. Such an effect is predicted by EFTs and could be detected in an EFT interpretation.

LHC. Furthermore, under a common EFT framework, measurements can be combined to provide better sensitivity to new physics than possible with a standalone analysis.

This chapter presents an SMEFT (Section 2.3.3) interpretation of a combination of Higgs boson analyses that use data collected by the CMS experiment during proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ between 2016 and 2018, corresponding to an integrated luminosity of 138 fb^{-1} [2]. This is the same combination that reported the STXS results discussed in Section 2.2.2. The input analyses, summarized in Table 6.1, measure the rate of Higgs boson production in an exhaustive variety of decay channels and production modes. Their combination allows for the isolation of new physics effects to specific production or decay modes, and even to specific kinematic regions within those modes, thanks to the inclusion of the STXS (Section 2.2) binning in most of the analyses. In an EFT framework, this corresponds to the possibility to simultaneously constrain a good number of Wilson coefficients.

The parameterization of these measurements in terms of the Wilson coefficients, described in Section 6.2, is determined partly with the provision of analytically-derived equations, and mostly via the use of MC-based techniques. The methodology improves upon the previous iteration of this combination performed by the CMS experiment [131] in a number of ways, most notably, by calculating the ggH and ggZH predictions at loop-level, introducing propagator corrections, and including acceptance corrections for the

Table 6.1: Summary of the analyses in the combination used for the SMEFT interpretation. Analyses target particular final states and one or more production modes which are indicated by the production tags. In some analyses, the measurement of particular production modes is split according to the stage 1.2 STXS. This is indicated by a checkmark in the final column, and the corresponding splitting is described in the production tags. For example, the $H \rightarrow \gamma\gamma$ analysis measures the ttH production mode in p_T bins, as shown in Fig. 2.17, whereas the $H \rightarrow 4l$ analysis measures this mode inclusively.

Input analysis	Final states	Production tags	Stage 1.2
$H \rightarrow \gamma\gamma$ [110]	$\gamma\gamma$	ggH, $p_T^H \times \text{N-jet} + \text{BSM bins}$ VBF, $m_{jj} \times p_T^{Hjj} + \text{BSM bins}$ VH hadronic WH leptonic, p_T^V bins ZH leptonic ttH, p_T^H bins tH	✓
$H \rightarrow 4l$ [121]	$4\mu, 2e2\mu, 4e$	ggH, $p_T^H \times \text{N-jet} + \text{BSM bins}$ VBF, $m_{jj} \times \text{N-jet} + \text{BSM bins}$ VH hadronic VH leptonic, p_T^V bins ttH	✓
$H \rightarrow l\nu l\nu$ [122]	$e\mu/\mu e, ee+\mu\mu$ $e\mu+jj, 3l, 4l$	ggH, $p_T^H \times \text{N-jet} + \text{BSM bins}$ VBF-like $m_{jj} + \text{BSM bins}$ VH hadronic WH leptonic ZH leptonic	✓
$H \rightarrow \tau\tau$ [123]	$e\mu, e\tau_h, \mu\tau_h$ $\tau_h\tau_h$	ggH, $p_T^H \times \text{N-jet} + \text{BSM bins}$ VBF, $m_{jj} \times \text{N-jet} + \text{BSM}$ WH leptonic, p_T^V bins ZH leptonic, p_T^V bins	✓
$H \rightarrow bb$ boosted [124]	bb	ggH, high p_T^H bins VBF, high p_T^H bins	✓
VBF, $H \rightarrow bb$ [125]	bb	VBF, resolved	
VH, $H \rightarrow bb$ [126]	bb	WH leptonic, p_T^V bins ZH leptonic, p_T^V bins	✓
ttH, $H \rightarrow bb$ [127]	bb	ttH, p_T^H bins tH	✓
ttH multilepton [128]	$2l(ss), 3l, 4l$ $1l+2\tau_h, 2l(ss)+1\tau_h$ $3l+1\tau_h$	ttH, tH	✓
$H \rightarrow \mu\mu$ [129]	$\mu\mu$	ggH, VBF, VH, ttH	
$H \rightarrow Z\gamma$ [130]	$l^+l^-\gamma$	ggH, VBF	

$H \rightarrow 4l$ and $H \rightarrow l\nu l\nu$ decay channels.

The final results, provided in Sections 6.3 and 6.4, are presented in two ways. First, 68% and 95% CL constraints on individual Wilson coefficients are shown, where the other coefficients are set to their SM values of zero. These results are sensitive to very specific types of new physics signatures (related to just one Wilson coefficient). However, it is expected that a BSM theory would lead to non-zero values of multiple Wilson coefficients. Therefore, simultaneous constraints on multiple Wilson coefficients are also extracted. In these results, a rotated basis for the Wilson coefficients is used to determine the constraints, where the rotated basis is determined via a principal component analysis (PCA).

6.2 SMEFT Parameterization

6.2.1 Overview and Assumptions

The combination of input analyses is performed at the likelihood level [2], and for the SMEFT interpretation, the probability distribution for the primary observables in each analysis, $p(\vec{x}; \vec{\mu}, \vec{\nu})$ (Eq. 4.1), must be rewritten in terms of the SMEFT Wilson coefficients, \vec{C} .

Firstly, the probability distributions for the background processes are assumed to be independent of \vec{C} . This is a reasonable assumption since most background processes are constrained well by the data, e.g. in control regions. For the signal, every STXS bin (stage 0 or 1.2 depending on the analysis) is treated as a separate process, which for an analysis category, c , has an expected number of events given by:

$$N_{pd}^c(\vec{C}) = \sigma^p(\vec{C}) \cdot \mathcal{B}^d(\vec{C}) \cdot A_{pd}^c(\vec{C}) \cdot \mathcal{L} \quad (6.1)$$

where σ^p is the cross section for the STXS bin, \mathcal{B}^d is the branching fraction for the decay channel, A_{pd}^c is the acceptance, and \mathcal{L} is the integrated luminosity. Here, acceptance is taken to mean the fraction of events from a process that land in a particular analysis category.

In shape-based analyses, $p(\vec{x}; \vec{\mu}, \vec{\nu})$ also includes a description of the shape of the observable. In this interpretation, the shape is assumed to be independent of \vec{C} , and therefore, only the parameterization of $N_{pd}^c(\vec{C})$ is considered. This is a safe assumption for observables like the diphoton invariant mass whose shape is mainly driven by the experimental resolution of the CMS ECAL. For observables like a BDT output score, whose inputs could be e.g. the Higgs boson p_T , the assumption is less safe, but the assumption is kept since the SMEFT impact on the shape is still expected to be a second-order effect compared to the impact on the rate.

Furthermore, for all except the $H \rightarrow 4l$ and $H \rightarrow l\nu l\nu$ decay channels, the acceptance is also assumed to be independent of \vec{C} . The validity of this assumption and the special treatment applied to $H \rightarrow 4l$ and $H \rightarrow l\nu l\nu$ is discussed further in Section 6.2.3. For the rest of the channels, the SMEFT parameterization is reduced to determining $\sigma^p(\vec{C})$ and $\mathcal{B}^d(\vec{C})$ for each STXS bin and decay channel respectively.

Only single insertions of dimension-6 operators are considered, which leads to quadratic dependencies of σ^p and Γ^d on \vec{C} (see Section 2.3.3), where Γ^d is the partial width for decay channel i , and the branching fractions are then expressed as $\Gamma^d(\vec{C})/\Gamma^H(\vec{C})$ where Γ^H is the total width of the Higgs boson. The total width parameterization is created by deriving the parameterization for all decay channels, including those not in Table 6.1, and

summing the widths: $\Gamma_H(\vec{C}) = \sum_d \Gamma^d(\vec{C})$.

The SMEFT contributions to σ^p and Γ^d are assumed to factorize from NLO (and higher order) QCD and EW corrections so that one can write:

$$\sigma^p(\vec{C}) = \sigma_{\text{SM}}^p \cdot \frac{\sigma_{\text{LO}}^p(\vec{C})}{\sigma_{\text{LO,SM}}^p}, \quad \Gamma^d(\vec{C}) = \Gamma_{\text{SM}}^d \cdot \frac{\Gamma_{\text{LO}}^d(\vec{C})}{\Gamma_{\text{LO,SM}}^d} \quad (6.2)$$

where σ_{SM}^p and Γ_{SM}^d are the SM cross sections and decay widths calculated at the highest order available, $\sigma_{\text{LO,SM}}^p$ and $\Gamma_{\text{LO,SM}}^d$ are the same calculated at LO, and $\sigma_{\text{LO}}^p(\vec{C})$ and $\Gamma_{\text{LO}}^d(\vec{C})$ are the SMEFT equivalents, also calculated at LO. Equation 6.2 is usually written in terms of signal strengths, $\mu^p(\vec{C})$ and $\mu^d(\vec{C})$, which are defined as:

$$\mu^p(\vec{C}) = \frac{\sigma_{\text{LO}}^p(\vec{C})}{\sigma_{\text{LO,SM}}^p}, \quad \mu^d(\vec{C}) = \frac{\Gamma_{\text{LO}}^d(\vec{C})}{\Gamma_{\text{LO,SM}}^d} \quad (6.3)$$

where the dependence of μ on \vec{C} is given by:

$$\mu^k(\vec{C}) = 1 + \sum_i A_i^k C_i + \sum_{ij} B_{ij}^k C_i C_j \quad (6.4)$$

where i and j run over the Wilson coefficients. Equations of the form given in Eq. 6.4 are often called *scaling equations*, and collectively, completely characterize the SMEFT parameterization.

Finally, the parameterization is chosen to be derived in the Warsaw basis with the `topU31` flavour assumption and the $\{G_F, m_Z, m_W\}$ input parameter scheme (Section 2.3.3), and only CP-even operators are considered because the STXS stage 1.2 bins are not sensitive to the differences between CP-even and CP-odd operators.

6.2.2 General Methodology

The scaling equations are derived with a mixture of techniques, summarized in Table 6.2, which depend on the production mode and decay channel. The scaling equations are determined entirely via MC simulation, with the exception of the $H \rightarrow \gamma\gamma$ and $H \rightarrow Z\gamma$ partial widths. These equations are instead taken from analytical calculations provided in Refs. [132] and [133] respectively, where a conversion is applied to the `topU31` basis. These processes are the exceptions because they both contain an electroweak loop in their LO Feynman diagrams and this type of diagram cannot be computed in the SMEFT with currently-available MC tools.

To derive a scaling equation using MC simulation, the corresponding cross section or partial width is calculated at LO using different values of the Wilson coefficients, and the

Table 6.2: Summary of techniques used to derive every scaling equation in the SMEFT interpretation. The $H \rightarrow \gamma\gamma$ and $H \rightarrow Z\gamma$ equations are taken from analytical calculations described in Refs. [132, 133]. All other equations are derived via MC simulation, using either the SMEFT_{SIM} or SMEFT@NLO UFO models. The acceptance dependence on \vec{C} is accounted for in the $H \rightarrow 4l$ and $H \rightarrow l\nu l\nu$ equations. The majority of the equations rely on event reweighting, with the exception of the inclusive (without acceptance corrections) $H \rightarrow 4l$ and $H \rightarrow l\nu l\nu$ equations that are used in the total width ($H \rightarrow \text{all}$) parameterization. Those equations are instead derived by generating dedicated samples at different values of \vec{C} .

Method	Mode / Channel	UFO Model	Reweighting	Dedicated	Acceptance corrections
MC	ggH	SMEFT@NLO	✓		
	ggZH		✓		
MC	qqH	SMEFT _{SIM}	✓		
	WH		✓		
	ZH		✓		
	bbH		✓		
	ttH		✓		
	tH		✓		
	$H \rightarrow b\bar{b}$		✓		
	$H \rightarrow \tau\tau$		✓		
	$H \rightarrow \mu\mu$		✓		
	$H \rightarrow l\nu l\nu$		✓		✓
	$H \rightarrow 4l$		✓		✓
Analytical	$H \rightarrow \gamma\gamma$		N/A		
	$H \rightarrow Z\gamma$				
MC & analytical	$H \rightarrow \text{all}$	SMEFT _{SIM}	✓	✓	

A and B terms are inferred. For N Wilson coefficients, the cross section or partial width must be calculated at $2N + (N - 1)/2$ points in the Wilson coefficient space to determine all terms.

A cross section/partial width is calculated by first generating events for the corresponding process using MADGRAPH v2.6.7 [99, 100] where the SM is assumed. For a partial width, the sum of weights from these events corresponds to $\Gamma^d(\vec{C} = 0)$. If a cross section is being calculated, the events are interfaced with PYTHIA 8.306 [104] with the CP5 tune [105, 106] for parton showering and fragmentation before being categorized into the STXS bins by a RIVET [134] routine. Then, the sum of event weights in an STXS bin corresponds to the value of $\sigma^p(\vec{C} = 0)$.

Two UFO (Universal FeynRules Output) [135] models are used to implement the SMEFT contributions in the MADGRAPH simulations. The SMEFT@NLO [136] model is capable of calculating SMEFT contributions at NLO in QCD and is used for the loop-induced ggH and ggZH processes. Alternatively, ggH could be simulated using LO UFO models like SMEFTSIM, but this relies on an EFT (not SMEFT) approximation of the process which breaks down for $p_T^H > m_t$ [34], and neglects SMEFT contributions that enter via the top-quark loop (see left of Fig. 2.24). In all cases apart from ggH and ggZH, the SMEFTSIM model is used, and is preferred over SMEFT@NLO in these cases because it also includes propagator corrections (see Section 2.3.3).

Using these UFO models, the partial widths and cross sections for different values of \vec{C} are mainly determined by reweighting the events [137] and then summing the new weights as before. This is a fair method when only minor changes to the phase space are expected due to the SMEFT contributions, where here, phase space means the components of all outgoing particle four-momenta. If a SMEFT contribution leads to an area of phase space being far more likely than in the SM, then a large SMEFT contribution will be described by a small number of events, leading to a large statistical uncertainty. This ultimately translates into an uncertainty in the A and B terms.

In the final result extraction, the uncertainties on the A and B terms are neglected so it is important that they remain small. However, there are cases, like that described in Section 2.3.3 for $H \rightarrow 4l$, where high uncertainties are expected, and are indeed found. For example, in an equation derived using 10^5 reweighted events for the $H \rightarrow 4e$ partial width, the A term for the Q_{HW} operator had an uncertainty of 740%. This cannot be neglected and generating enough events to reduce this uncertainty to a reasonable level would be too computationally expensive.

A similar level of uncertainty was found in the other $H \rightarrow 4f$ channels: $H \rightarrow l\nu l\nu$, $H \rightarrow llqq$, $H \rightarrow \nu\nu qq$, $H \rightarrow qqqq$, specifically for the Q_{HW} , Q_{HB} and Q_{HWB} operators. Therefore, in these channels, the terms involving these operators are not determined by

reweighting. Instead, the relevant partial widths are calculated by re-generating events (dedicated events) at the appropriate Wilson coefficient values. This leads to $< 10\%$ uncertainty in all the corresponding terms.

In Table 6.2, one might have noticed that the $H \rightarrow 4l$ is not indicated to use dedicated generation, which seems contradictory to the discussion above. This is because two scaling equations are derived for $H \rightarrow 4l$, one which corresponds to the inclusive partial width and one that corresponds to a fiducial partial width. The former is used in a calculation of the total width, and does require dedicated generation, hence why the total width is indicated to use dedicated generation in Table 6.2. The latter is used to parameterize N_{pd}^c for the $H \rightarrow 4l$ channel, and does not require dedicated generation for reasons related to acceptance corrections which are discussed in Section 6.2.3. For the same reasons, $H \rightarrow l\nu l\nu$ is also not indicated to use dedicated generation in Table 6.2.

6.2.3 Acceptance Corrections

In the $H \rightarrow 4l$ analysis [121], opposite-charge, same-flavour lepton pairs are formed and a selection of $12 < m_{ll} < 120 \text{ GeV}$ is placed. For $H \rightarrow 4l$ events, one pair of leptons typically comes from an on-shell Z boson ($m_Z \sim 91 \text{ GeV}$) and is therefore likely to pass this selection. The invariant mass of the pair of leptons whose mass is the furthest away from m_Z is denoted m_{34} , and the distribution of this variable in simulated $H \rightarrow 4l$ events is shown in Fig. 6.2. The distribution is shown for the SM and also for $C_{HB} = 0.5$. A significant enhancement is seen at low values of m_{34} , which is explained by the SMEFT Feynman diagrams shown in Fig. 2.26, that introduce a term like $1/m_{34}^2$ to the matrix element. Given that the shape of the m_{34} distribution depends on the SMEFT, and that the analysis criteria places a cut on this variable, the acceptance, A , is dependent on \vec{C} .

So far, the $H \rightarrow 4l$ scaling equation has been derived by considering the SMEFT contribution to all $H \rightarrow 4l$ events (inclusively). To account for the acceptance dependence, the scaling equation for $H \rightarrow 4l$ is derived using only the events that pass the analysis selection. In this way, the scaling equation for the $H \rightarrow 4l$ branching fraction accounts for the acceptance dependence and a separate equation for the acceptance is not needed.

Practically, this is achieved by reusing the simulated datasets originally used by the analysis, applying the selection criteria, and then reweighting the remaining events using a standalone reweighting tool [137, 138]. Reweighting is a valid approach here because the previously problematic region of phase space ($1/m_{ll}^2 \sim 0$) is removed by the selection criteria.

These datasets were generated with the full CMS simulation chain, including the detector response modelled with the GEANT4 package [107], and the object reconstruction. This is important because repeating a similar derivation by applying the selection criteria

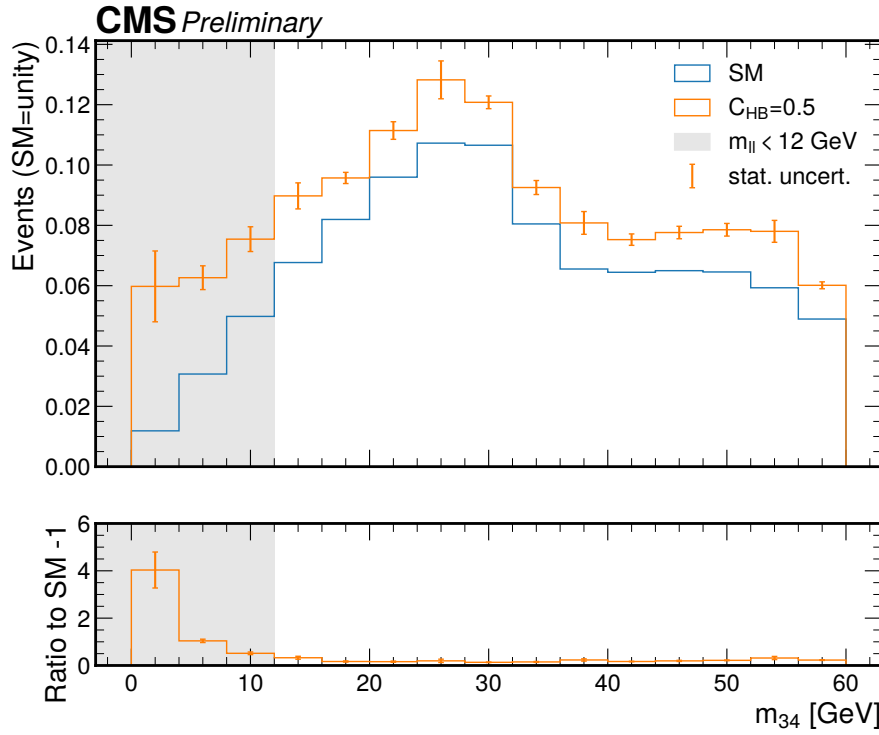


Figure 6.2: Distribution of m_{34} in the SM (blue) and in the SMEFT with $C_{HB} = 0.5$ (orange). The additional contribution from the SMEFT relative to the SM is shown in the bottom plot where the largest contribution is seen at the smallest values of m_{34} .

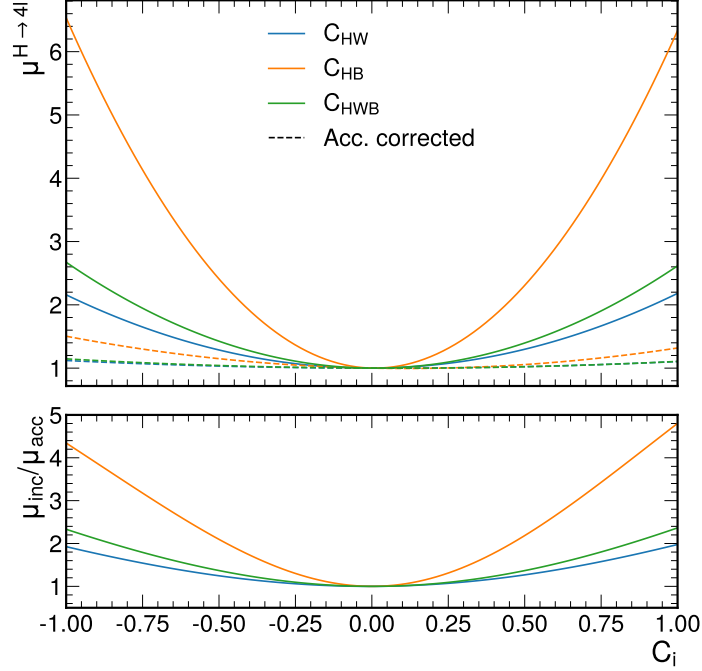


Figure 6.3: Scaling of the $H \rightarrow 4l$ partial width as functions of C_{HW} , C_{HB} , and C_{HWB} . The solid and dashed lines show the scaling equation when not considering acceptance corrections (an inclusive calculation), and when acceptance corrections are applied, respectively. The ratio of these two equations is shown in the bottom plot.

to generator-level quantities could lead to different (less valid) results if, for example, the experimental resolution on m_{34} was particularly poor.

The inclusive and acceptance-corrected $H \rightarrow 4l$ partial width scaling is shown in Fig. 6.3 as functions of the Wilson coefficients whose A and B terms change the most after considering acceptance corrections. These coefficients are C_{HB} , C_{HW} and C_{HWB} , which is expected since they all give rise to Feynman diagrams shown in Fig. 2.26. At values of $C = 1$, the scaling of the partial width is different up to a factor of 5, illustrating the need for acceptance corrections for the $H \rightarrow 4l$ decay. Similarly, selection criteria placed in the $H \rightarrow l\nu l\nu$ analysis [122] can also lead to acceptance dependence and the scaling equation for the $H \rightarrow l\nu l\nu$ partial width is derived using the same methods.

In the two-body decays, the decay products must always decay back-to-back and with energy $m_H/2$ in the Higgs boson rest frame. Therefore, SMEFT contributions cannot change the distribution of any kinematic variable related to the decay and there cannot be acceptance corrections. Hence, deriving the scaling equations with the approaches described in this subsection is unnecessary for the two-body decays.

Considering now the different production modes and their cross sections, there is the possibility that acceptance corrections are required. Selected criteria are often placed, for example, on p_T^H , and this is a variable whose distribution can be changed in the SMEFT. However, the STXS is binned in the variables whose distributions are expected to change the most from new physics contributions, like p_T^H . Therefore, for the scaling equation for a particular STXS bin, the analysis criteria are already well aligned with the events being used to derive the equation, at least in the variables that would lead to acceptance dependence. So, any acceptance dependence related to production is expected to be small and is neglected.

6.2.4 Summary

The parameterization of the Higgs boson cross sections and the decay widths are derived up to quadratic order in the Wilson coefficients. The rate of a process, which is a product of the cross sections and branching fractions, is then Taylor expanded up to quadratic order in the Wilson coefficients to isolate the contributions which are suppressed by $1/\Lambda^2$ (linear terms) and $1/\Lambda^4$ (quadratic terms), and to exclude terms suppressed by higher orders which are expected to be small. Furthermore, only the significant terms from the expansion are kept by removing terms where $|A|$ or $|B|$ is less than 0.01. This is reasonable because none of the measurements are made to greater than a few percent precision, so they cannot be sensitive to these Wilson coefficients unless $C \gg \mathcal{O}(1)$, in which case the EFT expansion is no longer valid anyway. The resulting parameterization has 43 different Wilson coefficients which are listed in Table 6.3.

The impacts of the Wilson coefficients on the STXS and Higgs boson branching fractions are shown in Figs. 6.4 to 6.6. As discussed in Section 2.3.3, there are different classes of contributions. The $C_{H\Box}$ coefficient enters every process and scales them by an equal amount (purple in top quarter of Fig. 6.4). Then, there are contributions that primarily affect a single process, such as C_{HG} for ggH (blue in top-middle quarter of Fig. 6.4), or the four-fermion Wilson coefficients for the $t\bar{t}H$ process (Fig. 6.6). These sorts of contributions highlight the importance of combining as many different processes as possible. Finally, there are contributions that benefit from the STXS binning, such as $C_{Hq}^{(3)}$ for the VH processes (Fig. 6.5) where the high p_T^V bins are more sensitive to this coefficient.

Table 6.3: A list of the 43 Wilson coefficients and their associated operators that the SMEFT interpretation is sensitive to. This table is a subset of Tables 2.4 and 2.5 that provide the full set of operators in the Warsaw basis under the `topU31` flavour assumption. Table taken from Ref. [2].

Group	WC	Operator	Group	WC	Operator
X^3	c_W	$\epsilon^{ijk} W_\mu^{i\nu} W_\nu^{j\rho} W_\rho^{k\mu}$	$\psi^2 H^2 D$	c_{He}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{e}_p \gamma^\mu e_r)$
	c_G	$f^{abc} G_\mu^{a\nu} G_\nu^{b\rho} G_\rho^{c\mu}$		c_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{u} \gamma^\mu u)$
$H^4 D^2$	$c_{H\Box}$	$(H^\dagger H)\Box(H^\dagger H)$		c_{Hd}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{d} \gamma^\mu d)$
	c_{HD}	$(D^\mu H^\dagger H)(H^\dagger D^\mu H)$		$\text{Re}(c_{Htb})$	$i(H^\dagger D_\mu H)(\bar{t} \gamma^\mu b)$
				c_{Ht}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{t} \gamma^\mu t)$
$X^2 H^2$	c_{HG}	$H^\dagger H G_{\mu\nu}^a G^{a,\mu\nu}$		c_{Hb}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{b} \gamma^\mu b)$
	c_{HW}	$H^\dagger H W_{\mu\nu}^i W^{i,\mu\nu}$	$(\bar{L}L)(\bar{L}L)$	$c_{ll}^{(1)}$	$(\bar{l}_p \gamma_\mu l_r)(\bar{l}_s \gamma^\mu l_t)$
	c_{HB}	$H^\dagger H B_{\mu\nu} B^{\mu\nu}$		$c_{Qq}^{(11)}$	$(\bar{Q} \gamma_\mu Q)(\bar{q} \gamma^\mu q)$
	c_{HWB}	$H^\dagger H W_{\mu\nu}^i B^{\mu\nu}$		$c_{Qq}^{(18)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{q} T^a \gamma^\mu q)$
$\psi^2 H^3$	$\text{Re}(c_{eH})$	$(H^\dagger H)(\bar{l}_p e_r H)$		$c_{Qq}^{(31)}$	$(\bar{Q} \sigma^i \gamma_\mu Q)(\bar{q} \sigma^i \gamma^\mu q)$
	$\text{Re}(c_{bH})$	$(H^\dagger H)(\bar{Q} \tilde{H} t)$		$c_{Qq}^{(38)}$	$(\bar{Q} \sigma^i T^a \gamma_\mu Q)(\bar{q} \sigma^i T^a \gamma^\mu q)$
	$\text{Re}(c_{tH})$	$(H^\dagger H)(\bar{Q} H b)$	$(\bar{R}R)(\bar{R}R)$	$c_{tu}^{(1)}$	$(\bar{t} \gamma_\mu t)(\bar{u} \gamma^\mu u)$
$\psi^2 XH$	$\text{Re}(c_{tG})$	$(\bar{Q} \sigma^{\mu\nu} T^a t) \tilde{H} G_{\mu\nu}^a$		$c_{tu}^{(8)}$	$(\bar{t} T^a \gamma_\mu t)(\bar{u} T^a \gamma^\mu u)$
	$\text{Re}(c_{tW})$	$(\bar{Q} \sigma^{\mu\nu} t) \sigma^i \tilde{H} W_{\mu\nu}^i$		$c_{td}^{(1)}$	$(\bar{t} \gamma_\mu t)(\bar{d} \gamma^\mu d)$
	$\text{Re}(c_{tb})$	$(\bar{Q} \sigma^{\mu\nu} t) \tilde{H} B_{\mu\nu}$		$c_{td}^{(8)}$	$(\bar{t} T^a \gamma_\mu t)(\bar{d} T^a \gamma^\mu d)$
	$\text{Re}(c_{bG})$	$(\bar{Q} \sigma^{\mu\nu} T^a b) H G_{\mu\nu}^a$	$(\bar{L}L)(\bar{R}R)$	$c_{qt}^{(1)}$	$(\bar{q} \gamma_\mu q)(\bar{t} \gamma^\mu t)$
	$\text{Re}(c_{bW})$	$(\bar{Q} \sigma^{\mu\nu} b) \sigma^i H W_{\mu\nu}^i$		$c_{qt}^{(8)}$	$(\bar{q} T^a \gamma_\mu q)(\bar{t} T^a \gamma^\mu t)$
$\psi^2 H^2 D$	$c_{Hl}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{l}_p \gamma^\mu l_r)$		$c_{Qu}^{(1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{u} \gamma^\mu u)$
	$c_{Hl}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H)(\bar{l}_p \sigma^i \gamma^\mu l_r)$		$c_{Qu}^{(8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{u} T^a \gamma^\mu u)$
	$c_{Hq}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{q} \gamma^\mu q)$		$c_{Qd}^{(1)}$	$(\bar{Q} \gamma_\mu Q)(\bar{d} \gamma^\mu d)$
	$c_{Hq}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H)(\bar{q} \sigma^i \gamma^\mu q)$		$c_{Qd}^{(8)}$	$(\bar{Q} T^a \gamma_\mu Q)(\bar{d} T^a \gamma^\mu d)$
	$c_{HQ}^{(1)}$	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{Q} \gamma^\mu Q)$			
	$c_{HQ}^{(3)}$	$(H^\dagger i \overleftrightarrow{D}_\mu^i H)(\bar{Q} \sigma^i \gamma^\mu Q)$			

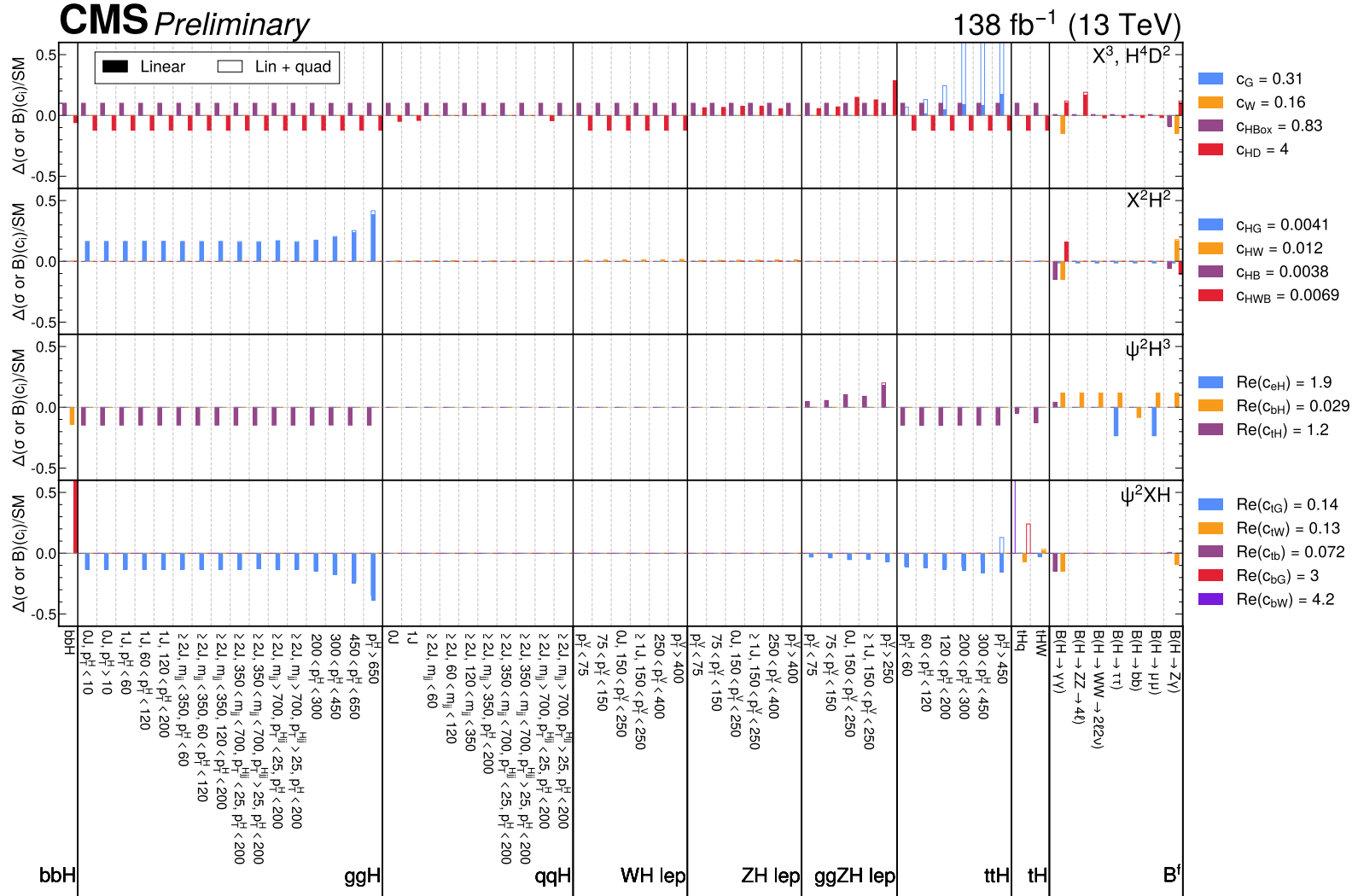
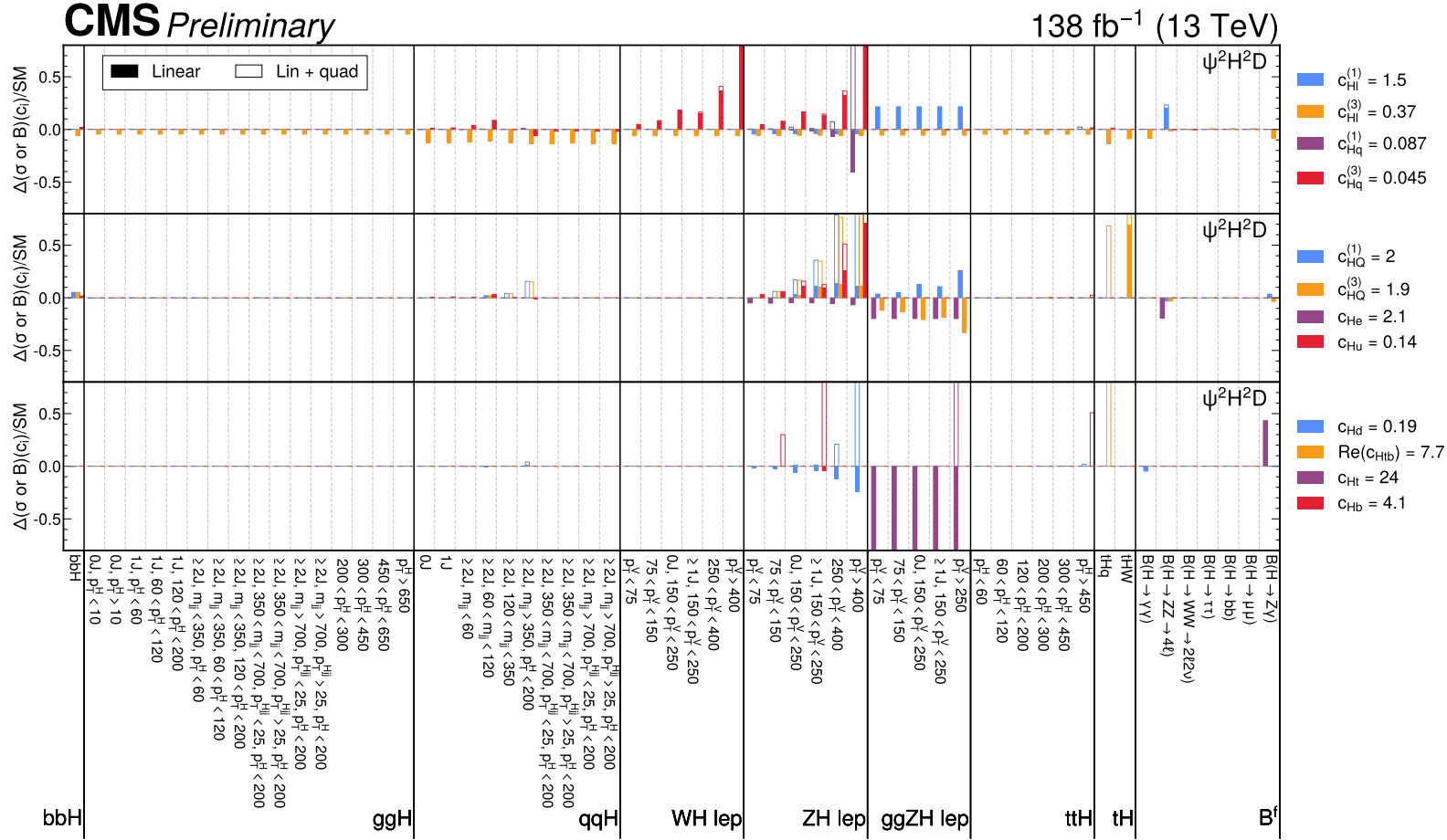


Figure 6.4: Impact of the SMEFT operators on the STXS and Higgs boson branching fractions. The impacts are shown for operators from the following groups: X^3 , $H^4 D^2$, $X^2 H^2$, $\psi^2 H^3$ and $\psi^2 XH$ (see Tables 2.4 and 2.5). The Wilson coefficients are set to the expected symmetrized 95% CL interval value, assuming all other Wilson coefficients are set to zero (SM). The impacts are presented relative to the SM prediction and are shown by the unfilled bars, and by the filled bars when considering only the linear (A) terms.



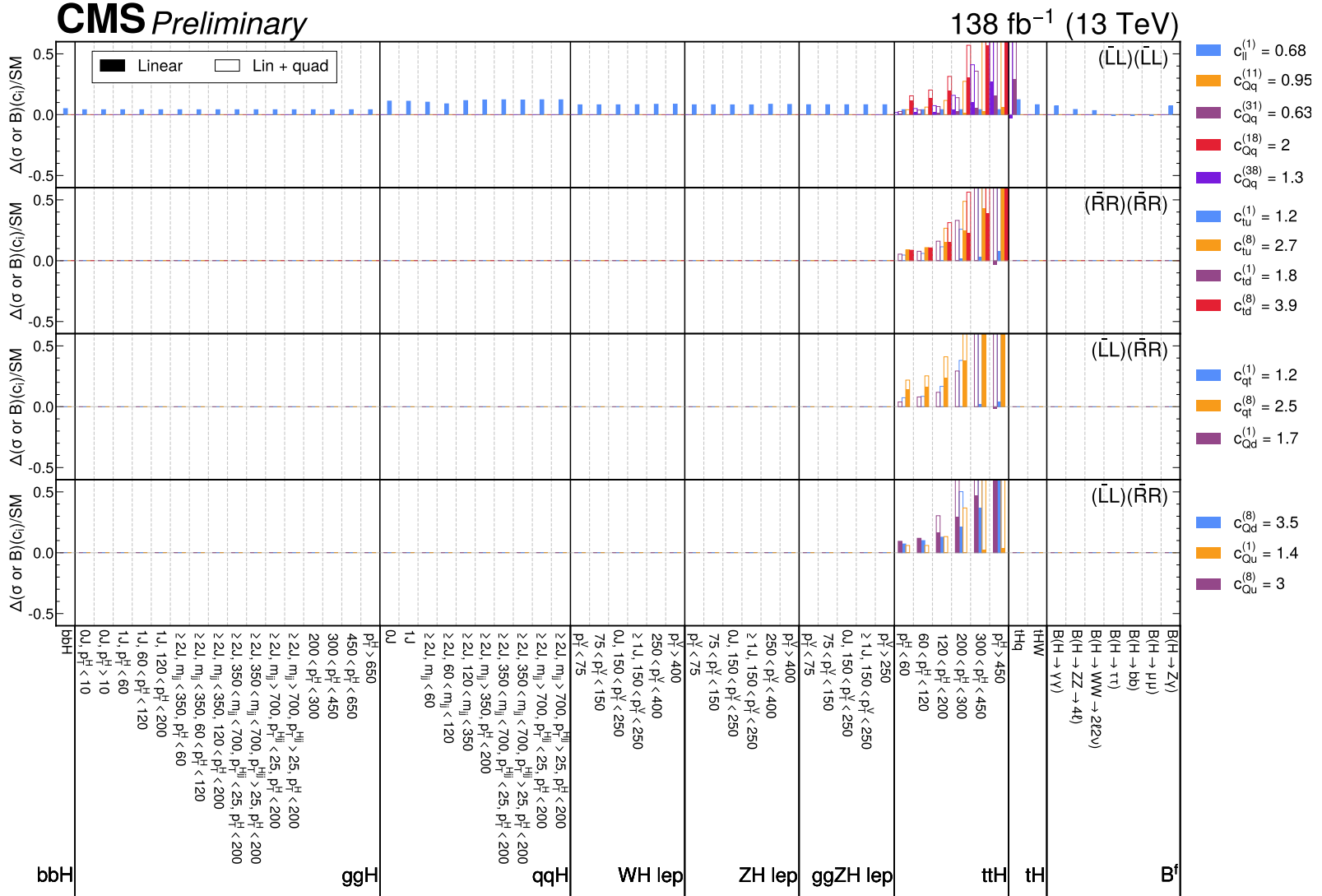


Figure 6.6: Impact of the SMEFT operators on the STXS and Higgs boson branching fractions. The impacts are shown for operators from the four-fermion groups (see Tables 2.4 and 2.5). The Wilson coefficients are set to the expected symmetrized 95% CL interval value, assuming all other Wilson coefficients are set to zero (SM). The impacts are presented relative to the SM prediction and are shown by the unfilled bars, and by the filled bars when considering only the linear (A) terms.

6.3 Results in the Nominal Basis

Maximum likelihood fits (Section 4.2.3) are performed to extract individual constraints on the 43 Wilson coefficients when fixing the other Wilson coefficients to their SM values of zero. The observed best-fit values, and 68% and 95% CL intervals are shown in Fig. 6.7 when using the SMEFT parameterization up to linear order only, or up to quadratic order (linear-plus-quadratic). Also shown are the 95% CL lower limits on the probed energy scale, Λ , when assuming $c_i = 1$.

In general, the results show good agreement with the SM, with the C_{HG} and C_{HB} coefficients showing the tightest constraints corresponding to probed energy scales of up to 15 TeV. The most discrepant result is that for $C_{Hq}^{(3)}$, which has a p-value with respect to the SM hypothesis of 0.01, regardless of whether the linear or up to quadratic parameterization is used. This discrepancy is expected because this coefficient has its highest impact in the high p_T^V WH and ZH leptonic bins, which saw the largest disagreements with the SM in the STXS measurements (Fig. 2.17).

The largest differences between the results when using the linear and linear-plus-quadratic parameterizations is seen for poorly constrained coefficients, where in some more extreme cases, the coefficients cannot be constrained with the linear contributions alone. This trend is expected because when $C \ll 1$, the quadratic terms are suppressed but as C becomes larger, there is less suppression. These differences are important to remember when considering the results in the rotated basis which is restricted to use the linear parameterization.

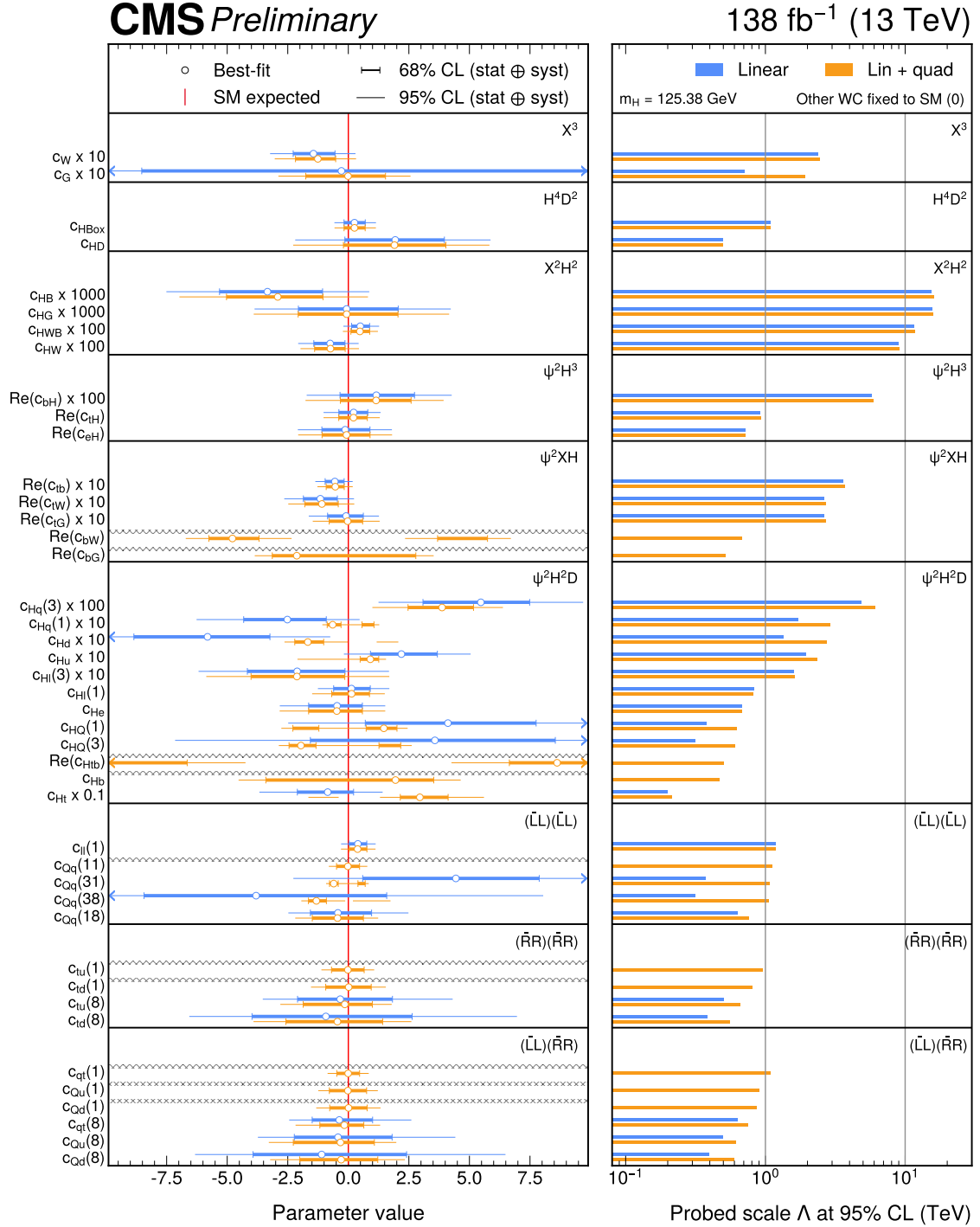


Figure 6.7: Observed constraints on the individual Wilson coefficients when setting the values of the other coefficients to zero. The left panel shows the best-fit values, the SM expectation ($C = 0$), and the 68% and 95% CL intervals. The results from the linear and linear-plus-quadratic parameterizations are shown in blue and orange respectively. For the linear parameterization, some coefficients are not constrained, which is indicated by a hatched line. When a confidence interval extends beyond the range of the plot, it is indicated by an arrow. The right panel shows the 95% CL lower limits on the probed energy scale, Λ , when assuming $c_i = 1$.

6.4 Results in the Rotated Basis

Given that BSM theories usually predict a variety of new physics effects, e.g. more than one new particle and/or interaction, it is expected that multiple Wilson coefficients will be non-zero. If that's the case, then the individual constraints in the nominal basis from Section 6.3 are not valid if one wants to draw a comparison with the predictions from a BSM theory. Instead, one needs to constrain the Wilson coefficients simultaneously. Here, this means providing constraints on Wilson coefficients when leaving the other coefficients freely floating in the fit, i.e. when making no assumption about their values.

The data available in this combination is not sufficient to constrain all 43 Wilson coefficients simultaneously since some Wilson coefficients have nearly degenerate effects on the observables, leading to flat directions in the likelihood. This is illustrated in some 2D toy examples in Fig. 6.8, where in the first example, both coefficients, C_i and C_j , can be simultaneously constrained, but in the second example, there is a flat direction along the line $C_i = -C_j$. In this case, the linear combination of coefficients, $C_0 = C_i + C_j$, can be constrained, but the combination, $C_1 = C_i - C_j$, cannot.

In the 43-dimensional case, the idea is the same, and the curvature of the log-likelihood function is studied at its minimum to determine which directions are flat, and which directions are sensitive to the data (the rotated basis). This is formalized in a principal component analysis technique which derives the directions from $\mathcal{H}_{\text{SMEFT}}$, which is the Hessian matrix corresponding to the second derivatives of the negative log likelihood (NLL) with respect to the Wilson coefficients, evaluated at the NLL minimum. When using the linear-plus-quadratic parameterization, this matrix is a function of the Wilson coefficients. Since the values of the coefficients are not known a priori, they are set to zero for the sake of deriving the rotated basis, which is equivalent to using the linear parameterization only.

An eigenvector decomposition is used to write $\mathcal{H}_{\text{SMEFT}} = \mathcal{R}^T \Lambda \mathcal{R}$ where \mathcal{R} is a rotation matrix that defines the linear combinations of Wilson coefficients (the eigenvectors) via $\text{EV}_j = \sum_k \mathcal{R}_{jk} C_k$, and Λ is a diagonal matrix containing the corresponding eigenvalues, λ_i , which indicate the constraining power for each eigenvector where the expected 68% CL interval is given by $1/\sqrt{\lambda_i}$. In the fit, combinations of Wilson coefficients with $1/\sqrt{\lambda_i} > 10$ are set to zero, therefore removing flat (or approximately flat) directions in the likelihood, and the 17 remaining eigenvectors, shown in Fig. 6.9, are included in the final fit. Furthermore, the impacts of the eigenvectors on the per-channel $\sigma \cdot \mathcal{B}$ STXS measurements presented in Fig. 2.18 are shown in Figs. 6.10 and 6.11.

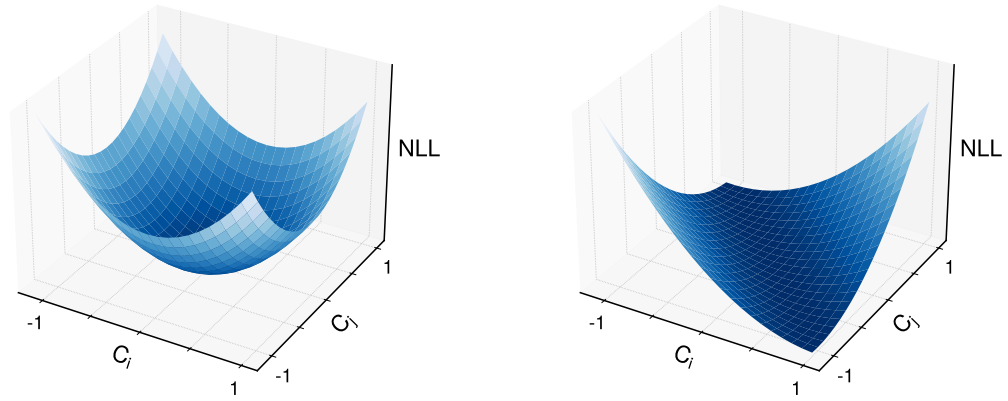


Figure 6.8: Two illustrative examples of a negative log likelihood (NLL) surface in two dimensions. On the left, the surface represents a scenario where it is possible to constrain both Wilson coefficients, C_i and C_j simultaneously. On the right, there is a flat direction along $C_i = -C_j$ meaning neither coefficient can be constrained whilst leaving the other floating. However, a linear combination, $C_0 = C_i + C_j$ is able to be constrained which corresponds to the direction perpendicular to the flat direction.

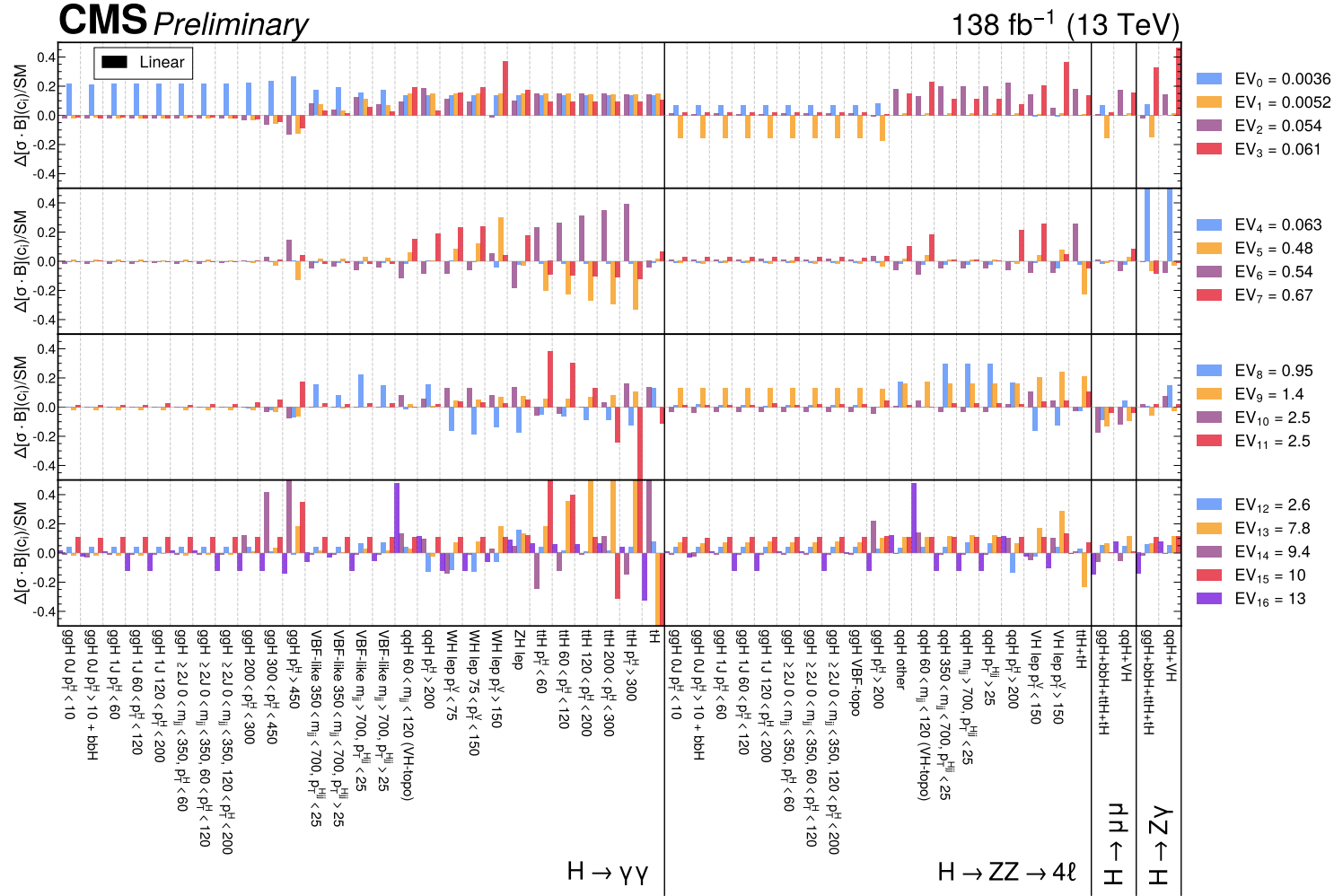


Figure 6.10: Impact of the eigenvectors on the STXS cross sections times branching fractions for the $H \rightarrow \gamma\gamma$, $H \rightarrow 4l$, $H \rightarrow \mu\mu$ and $H \rightarrow Z\gamma$ decay channels. The eigenvectors are set to the expected symmetrized 95% CL interval value in the profiled fit (when other eigenvectors are freely floating). The impacts are presented relative to the SM prediction and shown by filled bars for the linear parameterization only.

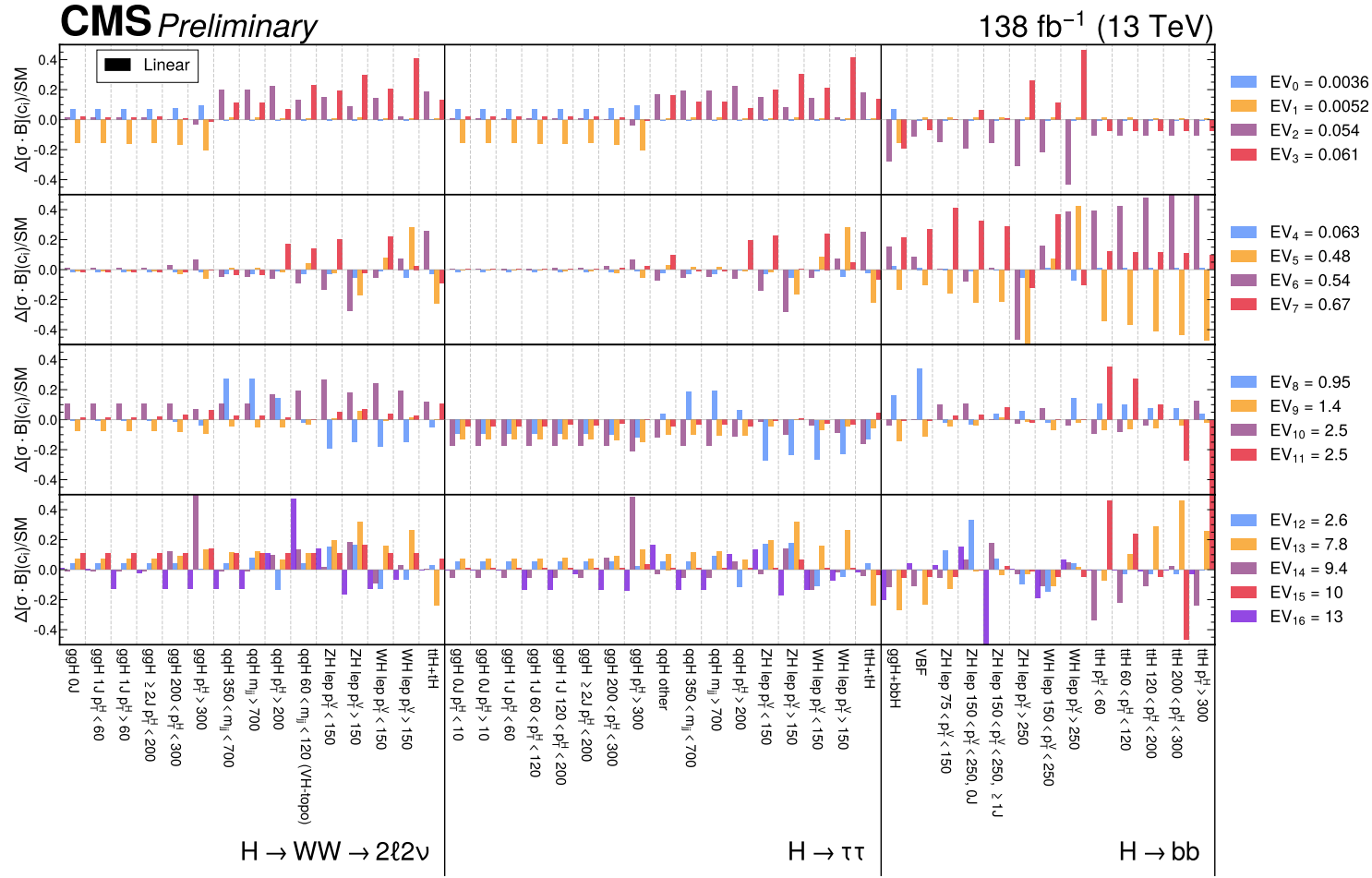


Figure 6.11: Impact of the eigenvectors on the STXS cross sections times branching fractions for the $H \rightarrow l\nu l\nu$, $H \rightarrow \tau\tau$, and $H \rightarrow bb$ decay channels. The eigenvectors are set to the expected symmetrized 95% CL interval value in the profiled fit (when other eigenvectors are freely floating). The impacts are presented relative to the SM prediction and shown by filled bars for the linear parameterization only.

Using these figures and the impacts in the nominal basis (Fig. 6.4), one can understand the construction of the eigenvectors and assign meaning to them. For example, the most sensitive eigenvector, EV_0 , is a combination of C_{HG} , which primarily affects ggH production, and C_{HB} , C_{HW} and C_{HWB} , which have the largest impact on the $H \rightarrow \gamma\gamma$ branching fraction. Therefore, this eigenvector primarily affects ggH production, specifically in the $H \rightarrow \gamma\gamma$ decay channel (see Fig. 6.10). It is unsurprising that this particular combination has the best expected sensitivity since inclusive measurements of ggH and $H \rightarrow \gamma\gamma$ in this combination have the smallest uncertainties when compared to other production and decay modes [2]. Similarly, it follows that the next eigenvector, EV_1 , primarily affects ggH production in the other channels, and has a similar sensitivity to EV_0 .

Maximum likelihood fits are used to extract 1D constraints on each of the 17 eigenvectors when allowing the other 16 eigenvectors to float freely. In these fits, only the linear parameterization is used, and this is done for several reasons. Firstly, the rotated basis is defined for the linear parameterization only and by introducing the quadratic terms, the basis is made no longer orthogonal. Secondly, the quadratic terms lead to more local minima in the NLL, causing the minimizer to struggle to find the true minimum, which in turn can lead to less reliable confidence intervals.

Finally, including the quadratic terms corresponds to an inconsistent SMEFT expansion in powers of $1/\Lambda^2$ because the quadratic terms enter at $1/\Lambda^4$, which is the same order that dimension-8 contributions would begin to enter. This is an issue because the inclusion of quadratic terms may lead to tighter constraints, which may not hold if the dimension-8 contributions were included. Therefore, it is preferred to use the linear parameterization only, which is more conservative and may provide a more realistic indication of the sensitivity.

The expected and observed best-fit values, 68% and 95% CL intervals, and corresponding lower limits on the probed energies scales for the eigenvectors are shown in Fig. 6.12. The 95% CL intervals range from ± 0.005 to ± 20 around the best-fit value depending on the eigenvector, and lower limits on the probed energy scales are up to 11 TeV and as low as 0.2 TeV.

Overall, fair agreement is found with the SM, with a p-value of 0.11. The most discrepant eigenvector is EV_3 , which is primarily a combination of $Re(C_{bH})$ and $C_{Hq}^{(3)}$. This follows from the discrepancy in $C_{Hq}^{(3)}$ seen in the nominal basis results which is explained by disagreements with the SM in measurements of the high p_T^V WH and ZH leptonic bins. Figures 6.10 and 6.11 show that this eigenvector has its largest impact in these particular bins as well.

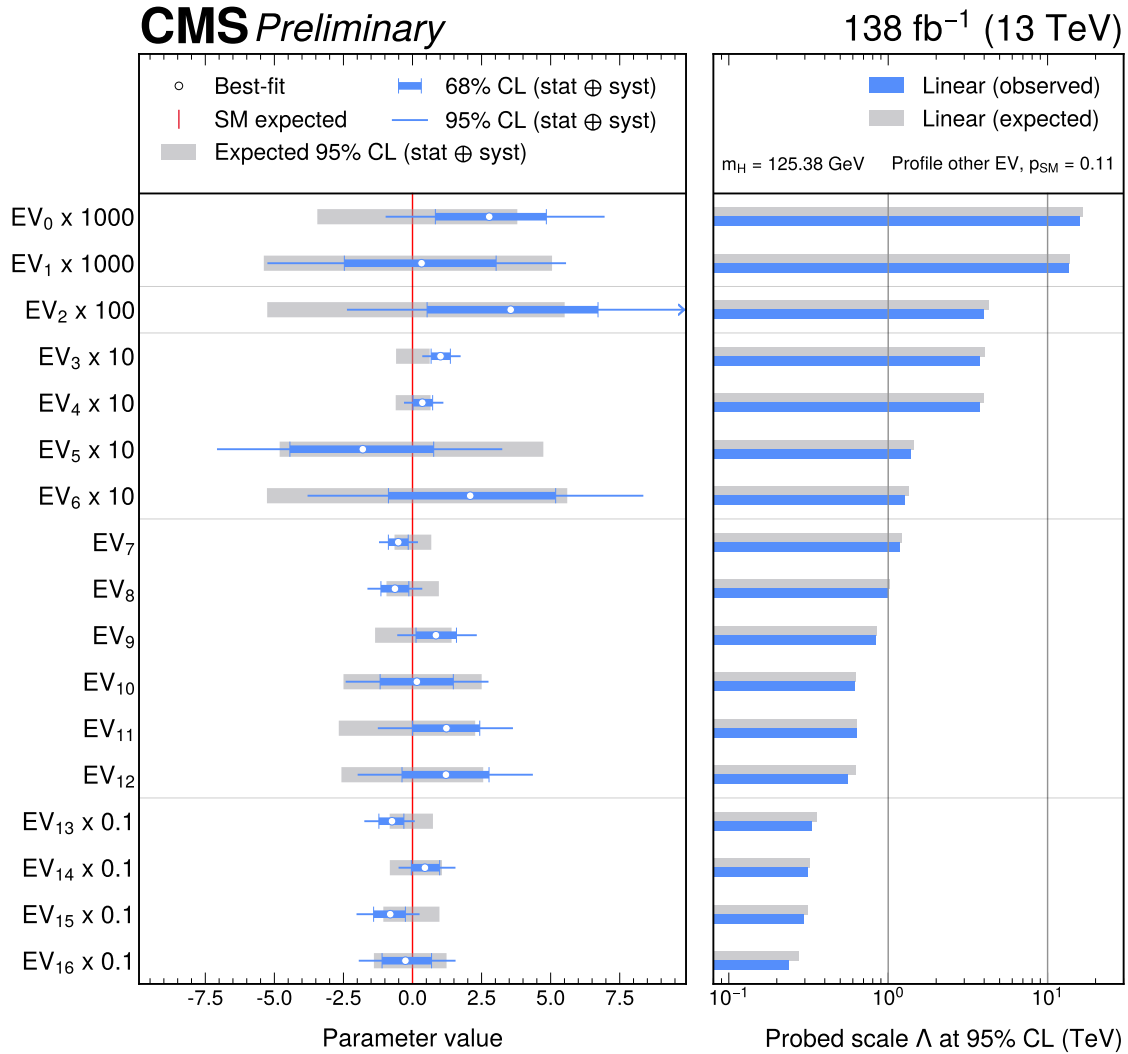


Figure 6.12: Observed and expected constraints on the eigenvectors from a profiled fit where the other eigenvectors are left freely floating. The left panel shows the best-fit values, the SM expectation ($C = 0$), and the 68% and 95% CL intervals. Only the results for the linear parameterization are shown. When a confidence interval extends beyond the range of the plot, it is indicated by an arrow. The right panel shows the 95% CL lower limits on the probed energy scale, Λ , when assuming $c_i = 1$.

6.5 Summary

Effective field theories (EFTs) are a powerful tool in the search for new physics. In the absence of the direct detection of new particles, EFTs provide a framework to parameterize the indirect effects that new physics can have on SM processes. In this Chapter, a SMEFT interpretation of a combination of Higgs boson analyses at the CMS experiment was presented. Most of the analyses in the combination were designed to measure the STXS, which studies the Higgs boson in a differential manner, splitting events according to the production mode, and by kinematic variables like p_T^H .

In Section 6.2, the derivation of the parameterization was described, detailing the different choices required to achieve a valid parameterization for all production and decay modes. In the majority of cases, the scaling equations were derived including propagator corrections by using the SMEFTSIM UFO model to generate and reweight events to different values for the Wilson coefficients. For the ggH and ggZH production modes, the scaling equations were derived similarly, but with the SMEFT@NLO model, allowing loop-level corrections to be included and ensuring that the parameterization is valid for the high p_T^H STXS bins. For the $H \rightarrow 4l$ and $H \rightarrow l\nu l\nu$ decay channels, acceptance corrections were derived with a new standalone reweighting procedure, and dedicated samples were generated to derive their contributions to the total width of the Higgs boson. All possible Wilson coefficients under the `topU31` flavour assumption were considered, and the final scaling equations contained contributions from 43 different Wilson coefficients.

Results were reported on the individual constraints of 43 different Wilson coefficients or the simultaneous constraints of 17 linear combinations of Wilson coefficients. In the individual results, the tightest constraint corresponded to a lower limit on the probed energy scale of 15 TeV when assuming $C_i = 1$, whereas in the simultaneous results, the tightest constraint corresponded to a lower limit of 11 TeV. Poor agreement with the SM was found in the individual constraint on $C_{Hq}^{(3)}$ corresponding to a p-value of 0.01 and a similar discrepancy was also found in the simultaneous fit for a linear combination of coefficients that heavily featured $C_{Hq}^{(3)}$. In both cases, the discrepancy was identified to be driven by excesses of events in high p_T^V WH and ZH leptonic channels.

Chapter 7

Conclusion

The Standard Model (SM) of particle physics is a remarkably successful theory, accurately describing the known fundamental particles and their interactions at the LHC and beyond. However, there are several key observations that the SM cannot explain, including dark matter [8], neutrino oscillations [9], and the matter-antimatter asymmetry of the universe [10]. Additionally, the SM has theoretical issues such as the hierarchy problem and the fact that it does not provide a description of gravity. Theories beyond the Standard Model (BSM) have been proposed that explain these anomalies and/or rectify the theoretical issues, but it is not clear which theory is correct.

Such theories include Warped Extra Dimensions (WED), which alleviates the hierarchy problem, and the Next-to-Minimal Supersymmetric Standard Model (NMSSM), which solves the hierarchy problem and provides a candidate for dark matter. These theories were tested in Chapter 5 by searching for the resonant production of a pair of scalar bosons, a process that both theories predict. More specifically, WED predicts the production of a new boson, X , which can be spin-0 or spin-2 and decays to two SM Higgs bosons ($X \rightarrow HH$), and the NMSSM predicts a new boson, X , which is only spin-0 and decays to a new scalar boson, Y , and a SM Higgs boson ($X \rightarrow YH$).

The searches for $X \rightarrow HH$ and $X \rightarrow YH$ were performed using data collected by the CMS experiment between 2016 and 2018 at the LHC at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 138 fb^{-1} . Mass ranges of 260–1000 GeV and 50–800 GeV were considered for m_X and m_Y respectively and for the first time at the CMS experiment, these searches were performed in final states including two photons and two tau leptons. This final state is particularly interesting given recent excesses reported in these decay channels for resonances around 95 and 650 GeV [96–98]. In the $Y \rightarrow \gamma\gamma$ decay channel, a local excess corresponding to 2.3 standard deviations was observed at $(m_X, m_Y) = (650, 95)$ GeV which is not significant enough to claim a discovery, but does motivate the continued search for new resonances in this mass region, especially with new data collected at the LHC from

2022 and beyond. Other excesses were observed in the $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$, and low and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ searches corresponding to local (global) significances of 2.6 (2.2), 3.2 (0.1), 3.2 (0.1) standard deviations respectively. Additionally, 95% CL upper limits were placed on the cross sections of the $X \rightarrow HH$ and $X \rightarrow YH$ processes as functions of m_X and m_Y , and theoretically excluded ranges of m_X and m_Y were quoted for some of the searches.

Assuming that the WED or NMSSM theories are correct, it is possible that the new particles have masses larger than 1 TeV, and are therefore beyond the range of masses searched for. Similarly, the masses could be so large that the particles could not be produced at the LHC. In this case, a direct search for these particles at the LHC would be unfeasible. However, indirect effects of these particles could still be observed as small deviations in distributions of SM processes, which can be parameterized by the Wilson coefficients in an Effective Field Theory (EFT).

In Chapter 6, these indirect effects were searched for with a SMEFT interpretation of a combination of Higgs boson measurements performed by the CMS experiment using data collected between 2016 and 2018 at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 138 fb^{-1} . The combination included an exhaustive variety of decay channels, and studied the production of the Higgs boson in a differential way using the STXS framework. This allowed for the individual constraint of 43 different Wilson coefficients where the tightest constraint corresponded to a probed energy scale up to 15 TeV when assuming $C_i = 1$, and allowed for the simultaneous constraint of 17 linear combinations of Wilson coefficients where the tightest constraint corresponded to a probed energy scale up to 11 TeV. Furthermore, a discrepancy with respect to the SM was observed in the individual constraint on $C_{Hq}^{(3)}$ corresponding to a p-value of 0.01.

Two fundamentally different types of searches for new physics have been presented in this thesis. In both the direct and indirect searches, no evidence for new physics was observed at the crucial level of five standard deviations, but they have shown hints towards new physics, whether that be the 650 GeV and 95 GeV excess seen in the $X \rightarrow YH$ search, or the discrepancy in the $C_{Hq}^{(3)}$ Wilson coefficient. Looking forward, the CMS experiment will continue to take data, and will do so with unprecedented rates after the High-Luminosity LHC [139] becomes operational in the mid-2030s, resulting in a final integrated luminosity of 3000 fb^{-1} at the end of the LHC's lifetime. This dataset represents a fantastic opportunity, and both direct and indirect searches should be utilized to maximize the chance of discovery new physics using it.

Bibliography

- [1] Aram Hayrapetyan et al. “Search for the nonresonant and resonant production of a Higgs boson in association with an additional scalar boson in the final state in proton-proton collisions at $\sqrt{s} = 13$ TeV” (June 2025). arXiv: 2506.23012 [hep-ex].
- [2] “Combined measurements and interpretations of Higgs boson production and decay at sqrt(s)=13 TeV”. Tech. rep. Geneva: CERN, 2025. URL: <https://cds.cern.ch/record/2929999>.
- [3] S. L. Glashow. “Partial Symmetries of Weak Interactions”. *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.
- [4] Steven Weinberg. “A Model of Leptons”. *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.
- [5] Abdus Salam. “Weak and Electromagnetic Interactions”. *Conf. Proc. C* 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.
- [6] Serguei Chatrchyan et al. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].
- [7] Georges Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].
- [8] Douglas Clowe et al. “A direct empirical proof of the existence of dark matter”. *Astrophys. J. Lett.* 648 (2006), pp. L109–L113. DOI: 10.1086/508162. arXiv: astro-ph/0608407.
- [9] Y. Fukuda et al. “Evidence for oscillation of atmospheric neutrinos”. *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

- [10] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. “Matter and Anti-matter in the Universe”. *New J. Phys.* 14 (2012), p. 095012. DOI: 10.1088/1367-2630/14/9/095012. arXiv: 1204.4186 [hep-ph].
- [11] X. Fan et al. “Measurement of the Electron Magnetic Moment”. *Phys. Rev. Lett.* 130.7 (2023), p. 071801. DOI: 10.1103/PhysRevLett.130.071801. arXiv: 2209.13084 [physics.atom-ph].
- [12] Mark Thomson. “Modern particle physics”. New York: Cambridge University Press, 2013. DOI: 10.1017/CB09781139525367.
- [13] S. L. Glashow, J. Iliopoulos, and L. Maiani. “Weak Interactions with Lepton-Hadron Symmetry”. *Phys. Rev. D* 2 (1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285.
- [14] E. G. Cazzoli et al. “Evidence for $\Delta S = -\Delta Q$ Currents or Charmed Baryon Production by Neutrinos”. *Phys. Rev. Lett.* 34 (1975), pp. 1125–1128. DOI: 10.1103/PhysRevLett.34.1125.
- [15] P. A. Zyla et al. “Review of Particle Physics”. *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.
- [16] Albert M Sirunyan et al. “A measurement of the Higgs boson mass in the diphoton decay channel”. *Phys. Lett. B* 805 (2020), p. 135425. DOI: 10.1016/j.physletb.2020.135425. arXiv: 2002.06398 [hep-ex].
- [17] Michael E. Peskin and Daniel V. Schroeder. “An Introduction to quantum field theory”. Reading, USA: Addison-Wesley, 1995. DOI: 10.1201/9780429503559.
- [18] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. “Broken Symmetries”. *Phys. Rev.* 127 (1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965.
- [19] C. S. Wu et al. “Experimental Test of Parity Conservation in β Decay”. *Phys. Rev.* 105 (1957), pp. 1413–1414. DOI: 10.1103/PhysRev.105.1413.
- [20] Yan-Rui Liu et al. “Pentaquark and Tetraquark states”. *Prog. Part. Nucl. Phys.* 107 (2019), pp. 237–320. DOI: 10.1016/j.ppnp.2019.04.003. arXiv: 1903.11976 [hep-ph].
- [21] Dan-di Wu. “A Brief Introduction to the Strong CP Problem”. *Z. Naturforsch. A* 52 (1997). Ed. by B. G. Englert and G. Sussmann, pp. 179–181.
- [22] D. de Florian et al. “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”. 2/2017 (Oct. 2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

- [23] Lisa Randall and Raman Sundrum. “A Large mass hierarchy from a small extra dimension”. *Phys. Rev. Lett.* 83 (1999), pp. 3370–3373. DOI: 10.1103/PhysRevLett.83.3370. arXiv: hep-ph/9905221.
- [24] Alexandra Carvalho. “Gravity particles from Warped Extra Dimensions, predictions for LHC” (Mar. 2014). arXiv: 1404.0102 [hep-ph].
- [25] Ulrich Ellwanger, Cyril Hugonie, and Ana M. Teixeira. “The Next-to-Minimal Supersymmetric Standard Model”. *Phys. Rept.* 496 (2010), pp. 1–77. DOI: 10.1016/j.physrep.2010.07.001. arXiv: 0910.1785 [hep-ph].
- [26] Pierre Fayet. “Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino”. *Nucl. Phys. B* 90 (1975), pp. 104–124. DOI: 10.1016/0550-3213(75)90636-7.
- [27] Pierre Fayet. “Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions”. *Phys. Lett. B* 69 (1977), p. 489. DOI: 10.1016/0370-2693(77)90852-8.
- [28] Sebastian Baum and Nausheen R. Shah. “Two Higgs Doublets and a Complex Singlet: Disentangling the Decay Topologies and Associated Phenomenology”. *JHEP* 12 (2018), p. 044. DOI: 10.1007/JHEP12(2018)044. arXiv: 1808.02667 [hep-ph].
- [29] S. F. King et al. “Discovery Prospects for NMSSM Higgs Bosons at the High-Energy Large Hadron Collider”. *Phys. Rev. D* 90.9 (2014), p. 095014. DOI: 10.1103/PhysRevD.90.095014. arXiv: 1408.1120 [hep-ph].
- [30] Ulrich Ellwanger and Cyril Hugonie. “Benchmark planes for Higgs-to-Higgs decays in the NMSSM”. *Eur. Phys. J. C* 82.5 (2022), p. 406. DOI: 10.1140/epjc/s10052-022-10364-3. arXiv: 2203.05049 [hep-ph].
- [31] E. Fermi. “An attempt of a theory of beta radiation. 1.” *Z. Phys.* 88 (1934), pp. 161–177. DOI: 10.1007/BF01351864.
- [32] B. Grzadkowski et al. “Dimension-Six Terms in the Standard Model Lagrangian”. *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].
- [33] Rodrigo Alonso et al. “Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology”. *JHEP* 04 (2014), p. 159. DOI: 10.1007/JHEP04(2014)159. arXiv: 1312.2014 [hep-ph].
- [34] Ilaria Brivio. “SMEFTsim 3.0 — a practical guide”. *JHEP* 04 (2021), p. 073. DOI: 10.1007/JHEP04(2021)073. arXiv: 2012.11343 [hep-ph].

- [35] Ilaria Brivio et al. “Electroweak input parameters”. Tech. rep. LHC EFT WG note, 5 pages. Geneva: CERN, 2021. arXiv: 2111.12515. URL: <https://cds.cern.ch/record/2792440>.
- [36] “LHC Machine”. *JINST* 3 (2008). Ed. by Lyndon Evans and Philip Bryant, S08001. DOI: 10.1088/1748-0221/3/08/S08001.
- [37] G. Aad et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.
- [38] S. Chatrchyan et al. “The CMS Experiment at the CERN LHC”. *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.
- [39] K. Aamodt et al. “The ALICE experiment at the CERN LHC”. *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.
- [40] A. Augusto Alves Jr. et al. “The LHCb Detector at the LHC”. *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.
- [41] Esma Mobs. “The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018” (2018). General Photo. URL: <https://cds.cern.ch/record/2636343>.
- [42] CMS Collaboration. “CMS luminosity measurement for the 2017 Data-Taking period at $\sqrt{s} = 13$ TeV”. CMS Physics Analysis Summary CMS-PAS-LUM-17-004. 2018. URL: <https://cds.cern.ch/record/2621960>.
- [43] CMS Collaboration. “CMS luminosity measurement for the 2018 Data-Taking period at $\sqrt{s} = 13$ TeV”. CMS Physics Analysis Summary CMS-PAS-LUM-18-002. 2019. URL: <https://cds.cern.ch/record/2676164>.
- [44] Albert M Sirunyan et al. “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS”. *Eur. Phys. J. C* 81 (2021), p. 800. DOI: 10.1140/epjc/s10052-021-09538-2. arXiv: 2104.01927 [hep-ex].
- [45] “LumiPublicResults < CMSPublic < TWiki”. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>. (Accessed on 13/05/2024).
- [46] Tai Sakuma and Thomas McCauley. “Detector and Event Visualization with SketchUp at the CMS Experiment”. *J. Phys. Conf. Ser.* 513 (2014). Ed. by D. L. Groep and D. Bonacorsi, p. 022032. DOI: 10.1088/1742-6596/513/2/022032. arXiv: 1311.4942 [physics.ins-det].
- [47] “CMS Technical Design Report for the Pixel Detector Upgrade”. Tech. rep. Sept. 2012. DOI: 10.2172/1151650.

- [48] W. Adam et al. “The CMS Phase-1 Pixel Detector Upgrade”. *JINST* 16.02 (2021), P02027. DOI: 10.1088/1748-0221/16/02/P02027. arXiv: 2012.14304 [physics.ins-det].
- [49] Armen Tumasyan et al. “The Phase-2 Upgrade of the CMS Tracker”. Tech. rep. June 2017. DOI: 10.17181/CERN.QZ28.FLHW.
- [50] A. M. Sirunyan et al. “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JINST* 13.06 (2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528 [physics.ins-det].
- [51] Serguei Chatrchyan et al. “The Performance of the CMS Muon Detector in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV at the LHC”. *JINST* 8 (2013), P11002. DOI: 10.1088/1748-0221/8/11/P11002. arXiv: 1306.6905 [physics.ins-det].
- [52] A. M. Sirunyan et al. “Particle-flow reconstruction and global event description with the CMS detector”. *JINST* 12.10 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965 [physics.ins-det].
- [53] Wolfgang Adam et al. “Track Reconstruction in the CMS tracker”. Tech. rep. Geneva: CERN, 2006. URL: <https://cds.cern.ch/record/934067>.
- [54] Susanna Cucciarelli et al. “Track reconstruction, primary vertex finding and seed generation with the Pixel Detector”. Tech. rep. Geneva: CERN, 2006. URL: <https://cds.cern.ch/record/927384>.
- [55] Serguei Chatrchyan et al. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. *JINST* 9.10 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569 [physics.ins-det].
- [56] Walaa Elmetenawee. “CMS tracking performance in Run 2 and early Run 3”. *PoS VERTEX2023* (2024), p. 074. DOI: 10.22323/1.448.0074. arXiv: 2312.08017 [hep-ex].
- [57] R. Fruhwirth. “Application of Kalman filtering to track and vertex fitting”. *Nucl. Instrum. Meth. A* 262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.
- [58] Pierre Billoir. “Progressive track recognition with a Kalman like fitting procedure”. *Comput. Phys. Commun.* 57 (1989), pp. 390–394. DOI: 10.1016/0010-4655(89)90249-X.
- [59] Pierre Billoir and S. Qian. “Simultaneous pattern recognition and track fitting by the Kalman filtering method”. *Nucl. Instrum. Meth. A* 294 (1990), pp. 219–228. DOI: 10.1016/0168-9002(90)91835-Y.

- [60] R. Mankel. “A Concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system”. *Nucl. Instrum. Meth. A* 395 (1997), pp. 169–184. DOI: 10.1016/S0168-9002(97)00705-5.
- [61] W. Adam et al. “Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC”. *eConf* C0303241 (2003), TULT009. DOI: 10.1088/0954-3899/31/9/N01. arXiv: physics/0306087.
- [62] Albert M Sirunyan et al. “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”. *JINST* 16.05 (2021), P05014. DOI: 10.1088/1748-0221/16/05/P05014. arXiv: 2012.06888 [hep-ex].
- [63] S Chatrchyan et al. “Performance of CMS Muon Reconstruction in Cosmic-Ray Events”. *JINST* 5 (2010), T03022. DOI: 10.1088/1748-0221/5/03/T03022. arXiv: 0911.4994 [physics.ins-det].
- [64] Albert M Sirunyan et al. “Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JINST* 15.02 (2020), P02027. DOI: 10.1088/1748-0221/15/02/P02027. arXiv: 1912.03516 [physics.ins-det].
- [65] “Muon Identification and Isolation efficiency on full 2016 dataset” (2017). URL: <https://cds.cern.ch/record/2257968>.
- [66] “Muon identification and isolation efficiencies with 2017 and 2018 data” (2018). URL: <https://cds.cern.ch/record/2629364>.
- [67] “Muon reconstruction performance during Run II” (2019). URL: <https://cds.cern.ch/record/2682902>.
- [68] “Muon Reconstruction and Identification Performance with Run-2 data” (2020). URL: <https://cds.cern.ch/record/2727091>.
- [69] Serguei Chatrchyan et al. “Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV”. *JINST* 7 (2012), P10002. DOI: 10.1088/1748-0221/7/10/P10002. arXiv: 1206.4071 [physics.ins-det].
- [70] A. Bodek et al. “Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments”. *Eur. Phys. J. C* 72 (2012), p. 2194. DOI: 10.1140/epjc/s10052-012-2194-8. arXiv: 1208.3710 [hep-ex].
- [71] Vardan Khachatryan et al. “Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV”. *JINST* 10.08 (2015), P08010. DOI: 10.1088/1748-0221/10/08/P08010. arXiv: 1502.02702 [physics.ins-det].

- [72] Aram Hayrapetyan et al. “Performance of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 13$ TeV”. *JINST* 19.09 (2024), P09004. DOI: 10.1088/1748-0221/19/09/P09004. arXiv: 2403.15518 [physics.ins-det].
- [73] Serguei Chatrchyan et al. “Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV”. *JINST* 8 (2013), P09009. DOI: 10.1088/1748-0221/8/09/P09009. arXiv: 1306.2016 [hep-ex].
- [74] M. Oreglia. “A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$ ”. Other thesis. Dec. 1980.
- [75] Vardan Khachatryan et al. “Observation of the Diphoton Decay of the Higgs Boson and Measurement of Its Properties”. *Eur. Phys. J. C* 74.10 (2014), p. 3076. DOI: 10.1140/epjc/s10052-014-3076-z. arXiv: 1407.0558 [hep-ex].
- [76] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti- k_t jet clustering algorithm”. *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].
- [77] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet User Manual”. *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].
- [78] “Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector”. Tech. rep. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2205284>.
- [79] “Jet energy scale and resolution measurement with Run 2 Legacy Data Collected by CMS at 13 TeV” (2021). URL: <http://cds.cern.ch/record/2792322>.
- [80] “Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13TeV with Phase 1 CMS detector” (2018). URL: <http://cds.cern.ch/record/2646773>.
- [81] Vardan Khachatryan et al. “Reconstruction and identification of τ lepton decays to hadrons and ν at CMS”. *JINST* 11.01 (2016), P01019. DOI: 10.1088/1748-0221/11/01/P01019. arXiv: 1510.07488 [physics.ins-det].
- [82] A. M. Sirunyan et al. “Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV”. *JINST* 13.10 (2018), P10005. DOI: 10.1088/1748-0221/13/10/P10005. arXiv: 1809.02816 [hep-ex].
- [83] Armen Tumasyan et al. “Identification of hadronic tau lepton decays using a deep neural network”. *JINST* 17 (2022), P07023. DOI: 10.1088/1748-0221/17/07/P07023. arXiv: 2201.08458 [hep-ex].

- [84] Aram Hayrapetyan et al. “The CMS Statistical Analysis and Combination Tool: Combine”. *Comput. Softw. Big Sci.* 8.1 (2024), p. 19. DOI: 10.1007/s41781-024-00121-4. arXiv: 2404.06614 [physics.data-an].
- [85] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].
- [86] Thomas Junk. “Confidence level computation for combining searches with small statistics”. *Nucl. Instrum. Meth. A* 434 (1999), pp. 435–443. DOI: 10.1016/S0168-9002(99)00498-2. arXiv: hep-ex/9902006.
- [87] Alexander L. Read. “Presentation of search results: The CL_s technique”. *J. Phys. G* 28 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.
- [88] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. *Ann. Math. Statist.* 9 (1938), p. 60. DOI: 10.1214/aoms/1177732360.
- [89] Abraham Wald. “Tests of statistical hypotheses concerning several parameters when the number of observations is large”. *Trans. Amer. Math. Soc.* 54 (1943), p. 426. DOI: 10.1090/S0002-9947-1943-0012401-3.
- [90] Robert F. Engle. “Chapter 13 Wald, likelihood ratio, and Lagrange multiplier tests in econometrics”. *Handbook of Econometrics*. Vol. 2. Elsevier, 1984, p. 775. DOI: 10.1016/S1573-4412(84)02005-5.
- [91] Aram Hayrapetyan et al. “Constraints on the Higgs boson self-coupling from the combination of single and double Higgs boson production in proton-proton collisions at $s=13\text{TeV}$ ”. *Phys. Lett. B* 861 (2025), p. 139210. DOI: 10.1016/j.physletb.2024.139210. arXiv: 2407.13554 [hep-ex].
- [92] “Combination of searches for nonresonant Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ ”. Tech. rep. Geneva: CERN, 2024. URL: <https://cds.cern.ch/record/2917252>.
- [93] Georges Aad et al. “Combination of Searches for Higgs Boson Pair Production in pp Collisions at $s=13\text{ TeV}$ with the ATLAS Detector”. *Phys. Rev. Lett.* 133.10 (2024), p. 101801. DOI: 10.1103/PhysRevLett.133.101801. arXiv: 2406.09971 [hep-ex].

- [94] Albert M Sirunyan et al. “Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JHEP* 03 (2021), p. 257. DOI: 10.1007/JHEP03(2021)257. arXiv: 2011.12373 [hep-ex].
- [95] Armen Tumasyan et al. “Search for a new resonance decaying into two spin-0 bosons in a final state with two photons and two bottom quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JHEP* 05 (2024), p. 316. DOI: 10.1007/JHEP05(2024)316. arXiv: 2310.01643 [hep-ex].
- [96] “Search for high mass resonances decaying into W^+W^- in the dileptonic final state with 138 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV”. Tech. rep. Geneva: CERN, 2022. URL: <https://cds.cern.ch/record/2803723>.
- [97] Armen Tumasyan et al. “Searches for additional Higgs bosons and for vector leptoquarks in $\tau\tau$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JHEP* 07 (2023), p. 073. DOI: 10.1007/JHEP07(2023)073. arXiv: 2208.02717 [hep-ex].
- [98] Aram Hayrapetyan et al. “Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s}=13\text{TeV}$ ”. *Phys. Lett. B* 860 (2025), p. 139067. DOI: 10.1016/j.physletb.2024.139067. arXiv: 2405.18149 [hep-ex].
- [99] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].
- [100] Johan Alwall et al. “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”. *Eur. Phys. J. C* 53 (2008), p. 473. DOI: 10.1140/epjc/s10052-007-0490-5. arXiv: 0706.2569 [hep-ph].
- [101] Enrico Bothmann et al. “Event generation with SHERPA 2.2”. *SciPost Phys.* 7 (2019), p. 34. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].
- [102] Rikkert Frederix and Stefano Frixione. “Merging meets matching in MC@NLO”. *JHEP* 12 (2012), p. 061. DOI: 10.1007/JHEP12(2012)061. arXiv: 1209.6215 [hep-ph].
- [103] Richard D. Ball et al. “Parton distributions from high-precision collider data”. *Eur. Phys. J. C* 77 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

- [104] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].
- [105] Vardan Khachatryan et al. “Event generator tunes obtained from underlying event and multiparton scattering measurements”. *Eur. Phys. J. C* 76 (2016), p. 155. DOI: 10.1140/epjc/s10052-016-3988-x. arXiv: 1512.00815 [hep-ex].
- [106] Albert M Sirunyan et al. “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements”. *Eur. Phys. J. C* 80 (2020), p. 4. DOI: 10.1140/epjc/s10052-019-7499-4. arXiv: 1903.12179 [hep-ex].
- [107] S. Agostinelli et al. “GEANT4 — a simulation toolkit”. *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.
- [108] Serguei Chatrchyan et al. “Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV”. *JHEP* 10 (2011), p. 132. DOI: 10.1007/JHEP10(2011)132. arXiv: 1107.4789 [hep-ex].
- [109] A. M. Sirunyan et al. “Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JHEP* 11 (2018), p. 185. DOI: 10.1007/JHEP11(2018)185. arXiv: 1804.02716 [hep-ex].
- [110] Albert M Sirunyan et al. “Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s} = 13$ TeV”. *JHEP* 07 (2021), p. 027. DOI: 10.1007/JHEP07(2021)027. arXiv: 2103.06956 [hep-ex].
- [111] Lorenzo Bianchini et al. “Reconstruction of the Higgs mass in $H \rightarrow \tau\tau$ Events by Dynamical Likelihood techniques”. *J. Phys. Conf. Ser.* 513 (2014). Ed. by D. L. Groep and D. Bonacorsi, p. 022035. DOI: 10.1088/1742-6596/513/2/022035.
- [112] Tianqi Chen and Carlos Guestrin. “XGBoost: A Scalable Tree Boosting System” (Mar. 2016). DOI: 10.1145/2939672.2939785. arXiv: 1603.02754 [cs.LG].
- [113] Pierre Baldi et al. “Parameterized neural networks for high-energy physics”. *Eur. Phys. J. C* 76.5 (2016), p. 235. DOI: 10.1140/epjc/s10052-016-4099-4. arXiv: 1601.07913 [hep-ex].
- [114] Adam Paszke et al. “PyTorch: An Imperative Style, High-Performance Deep Learning Library” (Dec. 2019). arXiv: 1912.01703 [cs.LG].
- [115] Diederik P. Kingma and Jimmy Ba. “Adam: A Method for Stochastic Optimization”. Dec. 2014. arXiv: 1412.6980 [cs.LG].
- [116] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

- [117] Peter Alfeld. “A trivariate clough—tocher scheme for tetrahedral data”. *Computer Aided Geometric Design* 1.2 (1984), pp. 169–181. DOI: [https://doi.org/10.1016/0167-8396\(84\)90029-3](https://doi.org/10.1016/0167-8396(84)90029-3).
- [118] P. D. Dauncey et al. “Handling uncertainties in background shapes: the discrete profiling method”. *JINST* 10.04 (2015), P04015. DOI: [10.1088/1748-0221/10/04/P04015](https://doi.org/10.1088/1748-0221/10/04/P04015). arXiv: 1408.6865 [physics.data-an].
- [119] R. A. Fisher. “On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P”. *Journal of the Royal Statistical Society* 85.1 (1922), pp. 87–94. URL: <http://www.jstor.org/stable/2340521>.
- [120] Adrian Carmona et al. “Matchmakereft: automated tree-level and one-loop matching”. *SciPost Phys.* 12.6 (2022), p. 198. DOI: [10.21468/SciPostPhys.12.6.198](https://doi.org/10.21468/SciPostPhys.12.6.198). arXiv: 2112.10787 [hep-ph].
- [121] Albert M Sirunyan et al. “Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV”. *Eur. Phys. J. C* 81.6 (2021), p. 488. DOI: [10.1140/epjc/s10052-021-09200-x](https://doi.org/10.1140/epjc/s10052-021-09200-x). arXiv: 2103.04956 [hep-ex].
- [122] Armen Tumasyan et al. “Measurements of the Higgs boson production cross section and couplings in the W boson pair decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *Eur. Phys. J. C* 83.7 (2023), p. 667. DOI: [10.1140/epjc/s10052-023-11632-6](https://doi.org/10.1140/epjc/s10052-023-11632-6). arXiv: 2206.09466 [hep-ex].
- [123] Armen Tumasyan et al. “Measurements of Higgs boson production in the decay channel with a pair of τ leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV”. *Eur. Phys. J. C* 83.7 (2023), p. 562. DOI: [10.1140/epjc/s10052-023-11452-8](https://doi.org/10.1140/epjc/s10052-023-11452-8). arXiv: 2204.12957 [hep-ex].
- [124] Aram Hayrapetyan et al. “Measurement of boosted Higgs bosons produced via vector boson fusion or gluon fusion in the $H \rightarrow b\bar{b}$ decay mode using LHC proton-proton collision data at $\sqrt{s} = 13$ TeV”. *JHEP* 12 (2024), p. 035. DOI: [10.1007/JHEP12\(2024\)035](https://doi.org/10.1007/JHEP12(2024)035). arXiv: 2407.08012 [hep-ex].
- [125] Aram Hayrapetyan et al. “Measurement of the Higgs boson production via vector boson fusion and its decay into bottom quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JHEP* 01 (2024), p. 173. DOI: [10.1007/JHEP01\(2024\)173](https://doi.org/10.1007/JHEP01(2024)173). arXiv: 2308.01253 [hep-ex].

- [126] Armen Tumasyan et al. “Measurement of simplified template cross sections of the Higgs boson produced in association with W or Z bosons in the $H \rightarrow b\bar{b}$ decay channel in proton-proton collisions at $\sqrt{s}=13$ TeV”. *Phys. Rev. D* 109.9 (2024), p. 092011. DOI: 10.1103/PhysRevD.109.092011. arXiv: 2312.07562 [hep-ex].
- [127] Aram Hayrapetyan et al. “Measurement of the $t\bar{t}H$ and tH production rates in the $H \rightarrow b\bar{b}$ decay channel using proton-proton collision data at $\sqrt{s} = 13$ TeV”. *JHEP* 02 (2025), p. 097. DOI: 10.1007/JHEP02(2025)097. arXiv: 2407.10896 [hep-ex].
- [128] Albert M Sirunyan et al. “Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV”. *Eur. Phys. J. C* 81.4 (2021), p. 378. DOI: 10.1140/epjc/s10052-021-09014-x. arXiv: 2011.03652 [hep-ex].
- [129] Albert M Sirunyan et al. “Evidence for Higgs boson decay to a pair of muons”. *JHEP* 01 (2021), p. 148. DOI: 10.1007/JHEP01(2021)148. arXiv: 2009.04363 [hep-ex].
- [130] Armen Tumasyan et al. “Search for Higgs boson decays to a Z boson and a photon in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JHEP* 05 (2023), p. 233. DOI: 10.1007/JHEP05(2023)233. arXiv: 2204.12945 [hep-ex].
- [131] “Combined Higgs boson production and decay measurements with up to 137 fb⁻¹ of proton-proton collision data at $\sqrt{s} = 13$ TeV”. Tech. rep. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2706103>.
- [132] Sally Dawson and Pier Paolo Giardino. “Electroweak corrections to Higgs boson decays to $\gamma\gamma$ and W^+W^- in standard model EFT”. *Phys. Rev. D* 98.9 (2018), p. 095005. DOI: 10.1103/PhysRevD.98.095005. arXiv: 1807.11504 [hep-ph].
- [133] Sally Dawson and Pier Paolo Giardino. “Higgs decays to ZZ and $Z\gamma$ in the standard model effective field theory: An NLO analysis”. *Phys. Rev. D* 97.9 (2018), p. 093003. DOI: 10.1103/PhysRevD.97.093003. arXiv: 1801.01136 [hep-ph].
- [134] Christian Bierlich et al. “Robust Independent Validation of Experiment and Theory: Rivet version 3”. *SciPost Phys.* 8 (2020), p. 026. DOI: 10.21468/SciPostPhys.8.2.026. arXiv: 1912.05451 [hep-ph].
- [135] Celine Degrande et al. “UFO - The Universal FeynRules Output”. *Comput. Phys. Commun.* 183 (2012), pp. 1201–1214. DOI: 10.1016/j.cpc.2012.01.022. arXiv: 1108.2040 [hep-ph].
- [136] Céline Degrande et al. “Automated one-loop computations in the standard model effective field theory”. *Phys. Rev. D* 103.9 (2021), p. 096024. DOI: 10.1103/PhysRevD.103.096024. arXiv: 2008.11743 [hep-ph].

-
- [137] Olivier Mattelaer. “On the maximal use of Monte Carlo samples: re-weighting events at NLO accuracy”. *Eur. Phys. J. C* 76.12 (2016), p. 674. DOI: 10.1140/epjc/s10052-016-4533-7. arXiv: 1607.00763 [hep-ph].
 - [138] CMS Collaboration. “Standalone reweighting with nanoAOD-tools”. Version 1.0.0. Mar. 2025. DOI: 10.5281/zenodo.15091146. URL: <https://github.com/Charlotte-Knight/nanoAOD-tools>.
 - [139] I. Zurbano Fernandez et al. “High-Luminosity Large Hadron Collider (HL-LHC): Technical design report”. 10/2020 (Dec. 2020). Ed. by I. Béjar Alonso et al. DOI: 10.23731/CYRM-2020-0010.