Measurement of the $t\bar{t}H$ production cross section and search for tH production in a fiducial phase space in the $H \to \gamma\gamma$ decay channel at $\sqrt{s} = 13.6\,\text{TeV}$ with the CMS experiment

Von der Fakultät für Mathematik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Florian Alexander Mausolf, M. Sc.

aus Dortmund

Berichter: Prof. Dr. rer. nat. Johannes Erdmann
apl. Prof. Dr. rer. nat. Malgorzata Worek

Tag der mündlichen Prüfung: 3. August 2026

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

Florian Alexander Mausolf, M. Sc.
Measurement of the $t\bar{t}H$ production cross section and search for tH production in a fiducial phase space in the $H \to \gamma\gamma$ decay channel at $\sqrt{s} = 13.6\,\text{TeV}$ with the CMS experiment

Dissertation in Physik
Angefertigt am III. Physikalischen Institut A der RWTH Aachen University

Eidesstattliche Erklärung

Florian Alexander Mausolf erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, erstellt wurden.
Hiermit erkläre ich an Eides statt:

1. Diese Arbeit wurde vollständig in der Phase als Promovierender dieser Fakultät und Universität angefertigt.

2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt.

3. Soweit eigene Veröffentlichungen oder Veröffentlichungen Dritter herangezogen wurden, sind diese klar kenntlich gemacht.

4. Wenn aus eigenen Veröffentlichungen oder aus Veröffentlichungen Dritter zitiert wurde, wurde die jeweilige Quelle stets angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate.

5. Alle wesentlichen Quellen von Unterstützung wurden benannt.

6. Sofern Teile dieser Dissertation auf der Zusammenarbeit mit anderen basieren, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde. Eine zusammenfassende Übersicht zur Abgrenzung der Beiträge in kollaborativen Arbeiten findet sich in Anhang E.

7. Teile dieser Arbeit wurden vor deren Einreichung veröffentlicht oder als Preprint öffentlich zugänglich gemacht, und zwar in:

   - J. Erdmann, N. K. Kasaraguppe, and F. Mausolf, "Learning to bin: differentiable and Bayesian optimization for multi-dimensional discriminants in high-energy physics", 2026, *arXiv*:2601.07756.
   - J. Erdmann, F. Mausolf, and J. L. Späh, "KAN We Improve on HEP Classification Tasks? Kolmogorov–Arnold Networks Applied to an LHC Physics Example", *Comput. Softw. Big Sci.* **9** (2025) 9, *doi*:10.1007/s41781-025-00138-3, *arXiv*:2408.02743.

   Eine weiterführende Übersicht zu wissenschaftlichen Arbeiten und Beiträgen aus der Promotionszeit, einschließlich weiterer Veröffentlichungen und unveröffentlichter Beiträge, findet sich in Anhang E.

Aachen, den _____ _____

Unterschrift

Erklärung zur Nutzung von Hilfsmitteln auf Basis künstlicher Intelligenz

Bei der Erstellung dieser Dissertation wurden die KI-basierten Anwendungen ChatGPT und Codex von OpenAI unterstützend für sprachliche Überarbeitungen sowie für technische Hilfestellungen bei der Formulierung, Strukturierung und Implementierung einzelner Text- und Codebausteine verwendet. Die wissenschaftliche Konzeption der Arbeit, die Auswahl und Entwicklung der Methoden, die Festlegung der verwendeten Algorithmen, die Durchführung der Analysen, die Auswertung der Daten sowie die Interpretation der Ergebnisse erfolgten eigenständig durch mich. Sämtliche durch diese Anwendungen erzeugten Vorschläge wurden von mir inhaltlich und technisch kritisch geprüft, gegebenenfalls angepasst und nur nach eigener Bewertung übernommen.

Aachen, den _____ _____

Unterschrift

Abstract

This thesis presents a measurement of the production cross section of a Higgs boson in association with a pair of top quarks or with a top quark and a W boson, $t\bar{t}H + tHW$, and a search for its t-channel production together with a single top quark, tH, in the diphoton decay channel of the Higgs boson. The analysed data were taken in proton–proton collisions with the CMS experiment at the CERN Large Hadron Collider at a centre-of-mass energy of 13.6 TeV and correspond to an integrated luminosity of 62.4 fb^{-1}.

Events with a reconstructed diphoton candidate and at least one b-tagged jet are selected and divided into hadronic and leptonic channels according to the absence or presence of a reconstructed electron or muon. To reduce the dependence on assumptions about the top-quark Yukawa coupling, the analysis uses deep-neural-network classifiers trained on signal samples generated in different coupling scenarios and decorrelated adversarially with respect to the CP-even and CP-odd coupling modifiers κ_t and $\tilde{\kappa}_t$. A novel strategy for the optimisation of category boundaries is developed and applied to group events within a three-dimensional classifier-output space, using an approximation in which the signal significances are differentiable functions of the boundary positions.

The measurement is performed in a fiducial phase space defined at the particle level by requirements on the kinematic variables of the photons and the presence of a b jet, which further reduces the model dependence. The signal extraction is performed with a simultaneous profile likelihood fit to the diphoton invariant-mass distributions in 16 categories, for which the expected confidence intervals remain stable under variations of $(\kappa_t, \tilde{\kappa}_t)$ in the signal model.

The measured fiducial cross section is $\sigma_{\text{fid}}^{t\bar{t}H+tHW} = 650^{+310}_{-260}$ ab, in agreement with the standard model (SM) prediction of 602^{+50}_{-64} ab. The observed and expected significances of the $t\bar{t}H + tHW$ signal both correspond to 2.8 standard deviations. No significant tH signal is observed, and an upper limit of $\sigma_{\text{fid}}^{tH} < 755$ ab is set at 95% confidence level (CL), corresponding to 10.2 times the SM prediction; the expected upper limit is 830 ab. Furthermore, constraints are derived on the coupling modifiers $(\kappa_t, \tilde{\kappa}_t)$. Assuming the SM branching ratio for the diphoton decay, the inverted-sign coupling scenario $(\kappa_t, \tilde{\kappa}_t) = (-1, 0)$ is excluded at 95% CL, while the CP-odd scenario $(\kappa_t, \tilde{\kappa}_t) = (0, 1)$ lies close to the observed 68% CL contour.

Kurzfassung

Diese Dissertation präsentiert eine Messung des Produktionswirkungsquerschnitts eines Higgs-Bosons in Assoziation mit einem Top-Quark-Paar oder mit einem Top-Quark und einem W-Boson, $t\bar{t}H + tHW$, sowie eine Suche nach seiner t-Kanal-Produktion zusammen mit einem einzelnen Top-Quark, tH, im Diphoton-Zerfallskanal des Higgs-Bosons. Die analysierten Daten wurden in Proton–Proton-Kollisionen mit dem CMS-Experiment am CERN Large Hadron Collider bei einer Schwerpunktsenergie von 13,6 TeV aufgenommen und entsprechen einer integrierten Luminosität von 62,4 fb^{-1}.
Ereignisse mit einem rekonstruierten Diphoton-Kandidaten und mindestens einem als b-Jet identifizierten Jet werden ausgewählt und entsprechend der Abwesenheit oder Anwesenheit eines rekonstruierten Elektrons oder Myons in hadronische und leptonische Kanäle unterteilt. Um die Abhängigkeit von Annahmen über die Top-Quark-Yukawa-Kopplung zu verringern, verwendet die Analyse tiefe neuronale Netze als Klassifikatoren, die auf Signalereignissen aus verschiedenen Kopplungsszenarien trainiert und mithilfe eines Gegenspieler-Netzwerks so dekorreliert werden, dass ihre Ausgaben möglichst wenig von den CP-geraden und CP-ungeraden Kopplungsmodifikatoren κ_t und $\tilde{\kappa}_t$ abhängen. Es wird eine neuartige Strategie zur Optimierung von Kategoriegrenzen entwickelt und angewendet, um Ereignisse innerhalb eines dreidimensionalen Ausgaberaums der Klassifikatoren zu gruppieren, wobei die Signalsignifikanzen als differenzierbare Funktionen der Positionen der Kategoriegrenzen approximiert werden.
Die Messung wird in einem auf Teilchenebene definierten, experimentell zugänglichen Phasenraum durchgeführt, der durch Anforderungen an die kinematischen Variablen der Photonen und durch die Anwesenheit eines b-Jets festgelegt ist, was die Modellabhängigkeit weiter reduziert. Die Signalextraktion erfolgt mit einem simultanen Profile-Likelihood-Fit an die Verteilungen der invarianten Diphotonmassen in 16 Kategorien, wobei die erwarteten Konfidenzintervalle unter Variationen von $(\kappa_t, \tilde{\kappa}_t)$ im Signalmodell stabil bleiben.
Der gemessene Wirkungsquerschnitt in diesem Phasenraum beträgt $\sigma_{\text{fid}}^{t\bar{t}H+tHW} = 650^{+310}_{-260}$ ab und stimmt mit der Vorhersage des Standardmodells (SM) von 602^{+50}_{-64} ab überein. Die beobachtete und die erwartete Signifikanz des $t\bar{t}H + tHW$-Signals entsprechen 2,8 Standardabweichungen. Es wird kein signifikanter tH-Signalbeitrag beobachtet, und es wird eine obere Grenze von $\sigma_{\text{fid}}^{tH} < 755$ ab bei einem Konfidenzniveau (CL) von 95% gesetzt, entsprechend dem 10,2-fachen der SM-Vorhersage; die erwartete obere Grenze beträgt 830 ab. Darüber hinaus werden Einschränkungen für die Kopplungsmodifikatoren $(\kappa_t, \tilde{\kappa}_t)$ abgeleitet. Unter der Annahme des SM-Verzweigungsverhältnisses für den Diphoton-Zerfall wird das Szenario mit invertiertem Vorzeichen der Yukawa-Kopplung, $(\kappa_t, \tilde{\kappa}_t) = (-1, 0)$, bei 95% CL ausgeschlossen, während das CP-ungerade Szenario $(\kappa_t, \tilde{\kappa}_t) = (0, 1)$ nahe an der beobachteten 68%-CL-Kontur liegt.

Contents

Page

List of abbreviations xi

1. Introduction 1

2. The Higgs boson and the top quark in theory, phenomenology, and experiment 3
- 2.1. The standard model of particle physics . . . 3
  - 2.1.1. Quantum chromodynamics . . . 4
  - 2.1.2. Electroweak theory and symmetry breaking . . . 4
  - 2.1.3. Particle content . . . 7
  - 2.1.4. Open questions . . . 8
- 2.2. Modifications to the top-quark Yukawa coupling beyond the standard model . . . 9
- 2.3. Hadron collider phenomenology . . . 11
  - 2.3.1. Higgs boson production at the Large Hadron Collider . . . 13
  - 2.3.2. Decays of the Higgs boson and the top quark . . . 17
- 2.4. Overview of experimental research on the Higgs boson . . . 19
  - 2.4.1. General properties . . . 19
  - 2.4.2. $t\bar{t}H$ and tH production and the top-quark Yukawa coupling . . . 20

3. The CMS experiment at the CERN Large Hadron Collider 23
- 3.1. The Large Hadron Collider . . . 23
- 3.2. The CMS detector . . . 24
  - 3.2.1. Tracking detector . . . 24
  - 3.2.2. Electromagnetic calorimeter . . . 26
  - 3.2.3. Hadron calorimeter . . . 26
  - 3.2.4. Muon detectors . . . 27
  - 3.2.5. Trigger and data acquisition . . . 28
- 3.3. Reconstruction, identification, and calibration of physics objects . . . 28
  - 3.3.1. Tracks and primary vertices . . . 28
  - 3.3.2. Photons and electrons . . . 29
  - 3.3.3. Jets and b-tagging . . . 34
  - 3.3.4. Muons . . . 35
  - 3.3.5. Missing transverse momentum . . . 36

4. Analysis strategy and workflow 37
- 4.1. Analysis strategy . . . 37
- 4.2. Event processing with the HiggsDNA framework . . . 40
- 4.3. Signal-region event selection . . . 41
- 4.4. Validation region with $Z \to e^+e^-$ decays . . . 43

5. Datasets used in the analysis **45**
5.1. Recorded proton–proton collision data . . . 45
5.2. Simulated datasets . . . 47
5.2.1. Signal processes . . . 47
5.2.2. Resonant background from other $H \to \gamma\gamma$ processes . . . 49
5.2.3. Irreducible continuum background . . . 50
5.2.4. Reducible continuum background . . . 51
5.3. Other samples for validation with $Z \to e^+e^-$ decays . . . 51

6. Kinematic signal properties and modelling of the continuum background for the analysis optimisation **53**
6.1. Kinematic properties of the signal and coupling dependence . . . 53
6.2. Background modelling for classifier training and categorisation optimisation . . . 59
6.2.1. Common modelling strategy . . . 59
6.2.2. Hadronic channel . . . 60
6.2.3. Leptonic channel . . . 65

7. Classification of events **69**
7.1. Brief overview of machine-learning concepts . . . 69
7.2. Study on Kolmogorov–Arnold networks . . . 70
7.2.1. Motivation . . . 70
7.2.2. Simulated dataset and event selection . . . 71
7.2.3. Training strategy and network configurations . . . 72
7.2.4. Results and interpretation . . . 73
7.3. Datasets and preprocessing for the cross-section measurement . . . 76
7.4. Neural-network setup for coupling-robust classification . . . 78
7.5. $t\bar{t}H$–tH classification and comparison of decorrelation methods . . . 80
7.6. Signal–background classification . . . 87
7.7. Validation and robustness checks . . . 91

8. Differentiable optimisation of the event categorisation **95**
8.1. Categorisation approach . . . 96
8.1.1. Differentiable optimisation . . . 97
8.1.2. Optimisation objective and penalties . . . 98
8.2. Toy examples . . . 99
8.3. Application to the $t\bar{t}H$ and tH cross-section measurement . . . 102
8.4. Categorisation penalty for robustness against CP assumptions . . . 106

9. Statistical analysis: methodology and results **111**
9.1. Signal and resonant background model . . . 111
9.2. Continuum background model . . . 117
9.3. Likelihood framework and inference . . . 119
9.4. Robustness against assumptions on the top-quark Yukawa coupling . . . 122
9.5. Results from observed data . . . 129

10. Conclusions **137**

A. Effect of CP assumptions on tH efficiencies **139**

B. Robustness studies with alternative Asimov datasets **141**

C. Continuum-background modelling **147**

D. Observed per-category diphoton mass spectra **151**

E. Overview of personal scientific work and publications **157**

Bibliography **161**

Acknowledgements **179**

List of abbreviations

2HDM two-Higgs-doublet model

AUC area under the curve

BDT Boosted Decision Tree

BCE binary cross-entropy

BSM beyond the standard model

BPix barrel pixel

CCE categorical cross-entropy

CL confidence level

CKM Cabibbo–Kobayashi–Maskawa

CP charge-parity

CSC cathode strip chamber

DNN deep neural network

DisCo distance correlation

DGLAP Dokshitzer–Gribov–Lipatov–Altarelli–Parisi

DOF degrees of freedom

DT drift tube

DY Drell–Yan

EB ECAL barrel

ECAL electromagnetic calorimeter

EDM electric dipole moment

EE ECAL endcap

EW electroweak

FNUF front non-uniformity

FSR final-state radiation

GEM gas electron multiplier

GATO Gradient-based cATegorisation Optimiser

GMM Gaussian mixture model

GOF goodness of fit

HB HCAL barrel

HCAL hadron calorimeter

HE HCAL endcap

HF hadron forward calorimeter

HL-LHC High-Luminosity LHC

HO outer hadron calorimeter

ID identification

ISR initial-state radiation

JER jet-energy resolution

JSD Jensen–Shannon divergence

KAN Kolmogorov–Arnold network

KS Kolmogorov–Smirnov

LHC Large Hadron Collider

LO leading order

MC Monte Carlo

MLP multilayer perceptron

N3LO next-to-next-to-next-to-leading order

NLO next-to-leading order

NP nuisance parameter

NNLL next-to-next-to-leading-logarithmic

NNLO next-to-next-to-leading order

OOA out of acceptance

PDF parton distribution function

pdf probability density function

PF particle flow

POI parameter of interest

PS parton shower

PUPPI Pileup Per Particle Identification

QCD quantum chromodynamics

QED quantum electrodynamics

ReLU rectified linear unit

ROC receiver operating characteristic

RPC resistive plate chamber

SF scale factor

SM standard model

SMEFT standard model effective field theory

STXS simplified template cross section

VBF vector-boson fusion

1. Introduction

The discovery of the Higgs boson [1–3] by the ATLAS [4] and CMS [5] Collaborations at the Large Hadron Collider (LHC) [6] completed the particle content of the standard model (SM) of particle physics [7–9]. At the same time, it opened the way to a systematic experimental study of the quantum numbers and couplings of the only known scalar elementary particle, which has since become one of the central goals of high-energy physics research. Precise measurements of the Higgs boson couplings test whether the observed scalar follows the coupling pattern expected from the mechanism of electroweak (EW) symmetry breaking in the SM [10–15] and offer sensitivity to physics beyond the standard model (BSM). The Yukawa coupling between the Higgs boson and the top quark is of particular importance because the top quark is the most massive known elementary particle and therefore has the largest Yukawa coupling in the SM. Its value impacts the stability of the EW vacuum [16, 17] and plays a special role in many BSM theories, for example in explanations of EW symmetry breaking [18, 19]. A key question is whether the observed Higgs boson corresponds to a charge-parity (CP)-even scalar state, as predicted in the SM, or whether it contains a CP-odd component. This possibility is particularly interesting in view of baryogenesis. Within EW baryogenesis scenarios based on a first-order EW phase transition, the amount of CP violation present in the SM is insufficient to account for the observed baryon asymmetry of the universe [20, 21]. The search for additional sources of CP violation, including possible contributions from the scalar sector, is therefore a central part of the LHC physics programme. A pure CP-odd nature of the observed Higgs boson is already excluded by measurements of its couplings to weak vector bosons with data taken during Run 1 of the LHC at centre-of-mass energies $\sqrt{s} = 7$ and 8 TeV [22, 23]. Direct CP-sensitive measurements were later also pursued for the top-quark and tau-lepton Yukawa couplings, where purely CP-odd hypotheses are likewise disfavoured at the level of three standard deviations by the data. At the same time, the current experimental constraints still allow for sizeable CP-odd admixtures in the Yukawa couplings [24–28].

Direct experimental access to the top-quark Yukawa interaction at tree level can be obtained by measuring the associated production of the Higgs boson with top quarks. However, these processes are rare at the LHC and constitute only a fraction of approximately 1% of the total Higgs boson production cross section in the SM. The relevant modes are the production with a pair of top quarks, denoted by $t\bar{t}H$, and the single-top-quark associated production. In this thesis, the symbol tH refers to the dominant t-channel single-top-quark contribution, while the minor production mode with a W boson, tHW, is treated separately in notation and later combined with $t\bar{t}H$ because of its experimental similarity. The $t\bar{t}H$ production rate depends primarily on the magnitude and only mildly on the CP structure of the coupling. By contrast, tH production is suppressed by destructive interference in the SM. It is rarer by roughly an order of magnitude and highly sensitive to changes in the sign of the CP-even coupling [29–31]. The same interference structure also makes the relative increase of the tH cross section particularly pronounced in the presence of a CP-odd contribution. Thus, the simultaneous measurement of the $t\bar{t}H$ and tH cross sections is a powerful probe of the sign and the CP structure of the top-quark Yukawa coupling. The $t\bar{t}H$ process has been observed including the data taken in Run 2 of the LHC

at $\sqrt{s} = 13\,\text{TeV}$ and the Run-1 dataset [32, 33]. On the other hand, only upper limits could be placed on the tH production so far. Multiple searches have been performed without establishing experimental evidence for its existence, but several analyses reported mildly significant excesses with larger observed signal significances than expected by up to 2.4 standard deviations in the Run-2 datasets [34–36]. This provides further motivation to search for tH production with the Run-3 data recorded at $\sqrt{s} = 13.6\,\text{TeV}$, where Run 3 denotes the current LHC data-taking period that began in 2022.

This thesis presents a measurement of the combined $t\bar{t}H$+tHW production cross section and a search for the tH process. The analysed datasets were collected in proton–proton collisions with the CMS detector in 2022 and 2023 and correspond to an integrated luminosity of $62.4\,\text{fb}^{-1}$. The analysis is carried out in the diphoton decay channel of the Higgs boson ($H \to \gamma\gamma$), which has played a central role in the experimental study of the Higgs boson since its discovery and remains one of the key channels for its precise characterisation [26, 27, 37–47]. Its clean final state and excellent diphoton invariant-mass resolution achieved with the CMS detector provide strong discrimination against backgrounds. Furthermore, the full and precise reconstruction of the Higgs boson decay products allows for precise measurement of its kinematic variables. For top-quark-associated Higgs boson production, the suitability of the $H \to \gamma\gamma$ decay channel was notably demonstrated by the first observations of $t\bar{t}H$ production in a single Higgs boson decay channel [26, 27].

To measure both $t\bar{t}H + tHW$ and tH production, the contributions of the processes must be distinguished experimentally. This separation is challenging due to similar final states and is performed with machine-learning-based classifiers in this thesis. At the same time, the degree of separation between these processes is itself sensitive to the structure of the top-quark Yukawa coupling. In particular, it has been pointed out that the kinematic properties of $t\bar{t}H$ production with a CP-odd coupling become more similar to those of tH production in the SM, and vice versa, which can introduce a significant model dependence in the measurement [48]. This motivates an analysis strategy that is robust against variations in the strength, sign, and CP structure of the top-quark Yukawa coupling, so that the measurement remains as model independent as possible and can be reinterpreted under alternative coupling hypotheses without introducing biases or requiring enlarged uncertainties. The coupling dependence of experimental $t\bar{t}H$–tH classification has received little attention in the literature so far and is therefore investigated in detail in this thesis. It is addressed by decorrelating the machine-learning-based event-classifier scores from the coupling assumptions in the training sample, similarly to Ref. [49]. In addition, a novel event-categorisation strategy based on a differentiable optimisation algorithm is developed to maximise the analysis sensitivity. The measurement is formulated in a fiducial phase space aligned with the experimentally accessible region, thereby avoiding model-dependent extrapolations to the full phase space.

This thesis is structured as follows. Chap. 2 introduces the theoretical foundations of the SM, focusing on the Higgs boson and the top quark, and gives an overview of the experimental research most relevant for this study. Chap. 3 then describes the CMS detector and the reconstruction and calibration of physics objects as used in this thesis. Chaps. 4, 5, and 6 then discuss the analysis strategy, the recorded and simulated datasets used, and the experimental signatures of the signal and background processes. The deep-learning-based event classification and the differentiable categorisation-optimisation approach are described in Chaps. 7 and 8. Finally, Chap. 9 presents the statistical analysis and results obtained, before the thesis concludes in Chap. 10.

2. The Higgs boson and the top quark in theory, phenomenology, and experiment

This chapter first summarises the theoretical framework of the SM with a focus on the Higgs boson and its interactions. It then introduces possible modifications of the top-quark Yukawa coupling beyond the SM and the corresponding effective parameterisation used throughout this thesis. On this basis, the chapter discusses the hadron-collider phenomenology of Higgs boson production and decay at the LHC with a focus on the top-quark-associated processes, and closes with a concise overview of the current experimental status.

2.1. The standard model of particle physics

The notation and conventions in this chapter follow the reviews on quantum chromodynamics (QCD), the EW model, and Higgs boson physics from Refs. [50–52], to which the reader is referred for a more comprehensive introduction to the SM and its phenomenology. Standard Dirac notation is used throughout, with $\bar{\psi} = \psi^\dagger \gamma^0$, γ^μ the Dirac matrices, and repeated Lorentz and internal indices implicitly summed over.

The SM is a renormalisable relativistic quantum field theory that provides the established description of the known elementary particles and their non-gravitational interactions. Renormalisable here means that the ultraviolet divergences appearing in higher-order quantum corrections can be absorbed into a finite set of counterterms corresponding to parameters and fields already present in the Lagrangian. Once these parameters are fixed experimentally, the theory remains predictive order by order in perturbation theory. It is a gauge theory based on the local symmetry group $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ [7–9]. Gauge theory means that the interaction structure is determined by local symmetry principles: matter fields carry charges under the gauge group, and the corresponding gauge fields mediate the interactions. Before EW symmetry breaking, this symmetry describes the strong and EW interactions. Electromagnetism arises only after the breaking of $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ to the residual $\mathrm{U}(1)_{\mathrm{EM}}$ symmetry. The dynamics of the fields and their interactions are encoded in the Lagrangian density $\mathcal{L}$, from which the equations of motion and interaction vertices used for theoretical predictions are derived. A useful schematic decomposition of the SM Lagrangian is

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{gauge}} + \mathcal{L}_{\mathrm{fermion}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}} \,. \tag{2.1}$$

In this notation, $\mathcal{L}_{\mathrm{gauge}}$ contains the gauge-boson kinetic and self-interaction terms, $\mathcal{L}_{\mathrm{fermion}}$ the fermion kinetic terms and gauge interactions, $\mathcal{L}_{\mathrm{Higgs}}$ the kinetic term and scalar potential of the Higgs field, and $\mathcal{L}_{\mathrm{Yukawa}}$ the couplings between the Higgs field and fermions. Gravity is not part of the SM and is fully negligible for the collider phenomenology discussed in this thesis. A central ingredient of the SM is EW symmetry breaking, through which the weak gauge bosons acquire mass and fermion masses arise from Yukawa interactions with the Higgs field. The relevant sectors of the SM are discussed in more detail in the following.

2.1.1. Quantum chromodynamics

QCD is the theory of strong interactions of quarks and gluons. The QCD part of the SM Lagrangian can be written as

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{\psi}_{q,a} \left(i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t^C_{ab} \mathcal{A}^C_\mu - m_q \delta_{ab} \right) \psi_{q,b} - \frac{1}{4} F^A_{\mu\nu} F^{A\mu\nu} , \tag{2.2}$$

where the sum runs over the quark flavours. The field $\psi_{q,a}$ denotes a quark spinor of flavour q, while $a = 1, 2, 3$ denotes the colour index, and m_q is the corresponding quark mass. The matrices t^C_{ab} are the generators of $\text{SU}(3)_C$ acting on the colour-triplet quark fields, and $\mathcal{A}^C_\mu$ denotes the gluon fields, where the colour index $C = 1, \ldots, 8$ labels the eight colour states of the gluon field. The symbol δ_{ab} is the Kronecker delta in colour space. The first term in Eq. (2.2) describes the free propagation of quarks. The second term introduces the interaction of quarks with the gluon field. It is formulated as a rotation of the colour state of the quark in $\text{SU}(3)_C$ space, scaled by the strong gauge coupling g_s, which is conventionally expressed through the strong coupling constant $\alpha_s = g_s^2/(4\pi)$. The third term contains the quark masses. The explicit quark-mass term is written here directly in the broken phase and in the mass basis, as appropriate for collider phenomenology. Its gauge-invariant origin from Yukawa interactions after EW symmetry breaking is discussed below. The gluon field-strength tensor is

$$F^A_{\mu\nu} = \partial_\mu \mathcal{A}^A_\nu - \partial_\nu \mathcal{A}^A_\mu - g_s f^{ABC} \mathcal{A}^B_\mu \mathcal{A}^C_\nu . \tag{2.3}$$

The structure constants f^{ABC} follow the commutation relations $[t^A, t^B] = i f^{ABC} t^C$ of the generators. The last term in $F^A_{\mu\nu}$ is the characteristic non-Abelian contribution arising because the generators of $\text{SU}(3)_C$ do not commute. It implies gluon self-interactions, which are a central qualitative difference between QCD and quantum electrodynamics (QED). More generally, terms of the form $-\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ are gauge-field kinetic terms.

The QCD sector written here is the colour part of $\mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}}$ in the decomposition from Eq. (2.1). Phenomenologically, the interaction terms in the Lagrangian give rise in particular to a $gq\bar{q}$ vertex, a three-gluon vertex, and a four-gluon vertex; the first two are proportional to g_s, while the four-gluon vertex is proportional to g_s^2.

The renormalised strong coupling depends on the renormalisation scale μ_R, which is typically chosen of the order of the hard momentum scale Q of the process. This scale dependence is described by the QCD beta function, conventionally defined as $\beta(\alpha_s) = \mu_R^2 \, \mathrm{d}\alpha_s / \mathrm{d}\mu_R^2$ [50]. The gluon self-interactions make this beta function negative, such that $\alpha_s(\mu_R)$ decreases at large scales, leading to asymptotic freedom [53, 54]. Conversely, $\alpha_s(\mu_R)$ grows at low scales, perturbation theory becomes unreliable at scales of order 1 GeV, and coloured states are confined into colour-neutral hadrons rather than being observed as free particles.

2.1.2. Electroweak theory and symmetry breaking

The EW part of the SM is based on the gauge group $\text{SU}(2)_L \times \text{U}(1)_Y$. The corresponding gauge boson fields are the three W^a_μ with $a = 1, 2, 3$ associated with the $\text{SU}(2)_L$ group and B_μ associated

with $\mathrm{U}(1)_Y$. Their gauge couplings are g and g'. The left-handed lepton and quark fields of generation i transform as weak-isospin doublets under $\mathrm{SU}(2)_L$,

$$l_{Li} = \begin{pmatrix} \nu_i \\ l_i^- \end{pmatrix}_L , \qquad q_{Li} = \begin{pmatrix} u_i \\ d_i' \end{pmatrix}_L , \tag{2.4}$$

while the corresponding right-handed fields transform as $\mathrm{SU}(2)_L$ singlets and therefore do not couple to the weak-isospin gauge fields, although they may still carry weak hypercharge. Here, ν_i and l_i^- denote the neutrino and charged-lepton fields, u_i the up-type quark field, and d_i' the down-type weak-interaction eigenstate of generation i.

In order to explain masses of the weak bosons and the short-ranged behaviour of the weak interaction, the EW symmetry breaking mechanism is introduced by including the complex scalar Higgs doublet [10–15]

$$\Phi = \begin{pmatrix} \phi^+ \\ \frac{1}{\sqrt{2}} \left(\phi^0 + i\chi\right) \end{pmatrix} , \tag{2.5}$$

where ϕ^+, ϕ^0, and χ denote the scalar field components of the $\mathrm{SU}(2)_L$ doublet. The scalar sector is described by the Lagrangian

$$\mathcal{L}_{\text{Higgs}} = \left(D_\mu \Phi\right)^\dagger \left(D^\mu \Phi\right) - V(\Phi) , \qquad \text{with } V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 , \tag{2.6}$$

where $D_\mu = \partial_\mu + ig\sigma^a W_\mu^a/2 + ig' Y B_\mu/2$ denotes the EW covariant derivative using the Pauli matrices σ^a and the weak hypercharge Y. The parameters μ^2 and λ determine the shape of the Higgs potential $V(\Phi)$. For $\mu^2 < 0$, the Higgs field develops a non-zero vacuum expectation value

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} , \qquad v = \sqrt{-\mu^2/\lambda} \simeq 246\,\text{GeV} , \tag{2.7}$$

where v denotes its magnitude. This vacuum expectation value breaks the EW symmetry to the residual $\mathrm{U}(1)_{\text{EM}}$ symmetry of QED, whose generator is the electric-charge operator

$$Q = T_3 + \frac{Y}{2} . \tag{2.8}$$

Here, Q denotes the electric charge, T_3 the third component of weak isospin, and Y the weak hypercharge. The four real degrees of freedom in the Higgs doublet then decompose into three would-be Goldstone fields and one physical neutral scalar. By applying a specific gauge choice, the unitary gauge, these three fields are removed from the explicit scalar spectrum and instead provide the additional longitudinal polarisation states required for the now massive $W^\pm$ and Z bosons. Then, the Higgs doublet can be written in terms of the vacuum expectation value and the remaining physical Higgs field as

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} . \tag{2.9}$$

After electroweak symmetry breaking, the electroweak gauge fields are reorganised into the physical fields

$$\begin{aligned} A_\mu &= B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \,, \\ Z_\mu &= -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W \,, \\ W_\mu^\pm &= \frac{W_\mu^1 \mp i W_\mu^2}{\sqrt{2}} \,. \end{aligned} \tag{2.10}$$

Here, A_μ, Z_μ, and $W_\mu^\pm$ are the fields of the physical photon, Z boson, and charged W bosons, respectively. The weak mixing angle is defined through $\tan\theta_W = g'/g$, and the elementary electric charge satisfies $e = g\sin\theta_W = g'\cos\theta_W$. At tree level, this yields the following relations for the masses of the physical Higgs and gauge bosons:

$$m_H = \sqrt{2\lambda}\, v \,, \qquad m_W = \frac{gv}{2} \,, \qquad m_Z = \frac{\sqrt{g^2+g'^2}\, v}{2} \,, \qquad m_\gamma = 0 \,. \tag{2.11}$$

Inserting the shifted Higgs field $\Phi(x) = \frac{1}{\sqrt{2}}(0,\, v + H(x))^T$ into the Higgs-sector Lagrangian from Eq. (2.6) generates the tree-level Higgs boson couplings to the massive weak bosons,

$$\mathcal{L}^V_{\text{Higgs}} = \left(\frac{H}{v} + \frac{H^2}{2v^2}\right)\left(2m_W^2 W_\mu^+ W^{-\mu} + m_Z^2 Z_\mu Z^\mu\right) \,, \tag{2.12}$$

where the terms linear in H describe the HWW and HZZ interactions and the terms quadratic in H the quartic couplings $HHWW$ and $HHZZ$. Likewise, inserting the shifted field into the scalar potential $V(\Phi)$ yields the Higgs mass term and the Higgs self-couplings,

$$V(H) = \frac{m_H^2}{2}H^2 + \frac{m_H^2}{2v}H^3 + \frac{m_H^2}{8v^2}H^4 \,, \tag{2.13}$$

where the cubic and quartic terms correspond to the trilinear and quartic self-couplings. The EW gauge kinetic terms also generate self-interactions among the non-Abelian weak gauge fields, but they are not written explicitly here because the following discussion focuses on the Higgs-field, fermion, and weak-current interactions relevant for this thesis.

Gauge-invariant Yukawa interactions generate fermion masses after EW symmetry breaking through the Lagrangian

$$\mathcal{L}_Y = -\hat{Y}^d_{ij}\, \bar{q}_{Li}\Phi\, d_{Rj} - \hat{Y}^u_{ij}\, \bar{q}_{Li}\widetilde{\Phi}\, u_{Rj} - \hat{Y}^l_{ij}\, \bar{l}_{Li}\Phi\, e_{Rj} + \text{h.c.} \,, \tag{2.14}$$

where $\hat{Y}^d_{ij}$, $\hat{Y}^u_{ij}$, and $\hat{Y}^l_{ij}$ are the Yukawa matrices for down-type quarks, up-type quarks, and charged leptons, and i, j are generation indices. These matrices are free parameters of the SM in flavour space and are not predicted by the gauge symmetry itself. The abbreviation h.c. denotes the Hermitian conjugate terms required for a real Lagrangian. Here, $\widetilde{\Phi} = i\sigma_2\Phi^*$ is the conjugate Higgs doublet, constructed with the second Pauli matrix σ_2 so that it transforms again as an $\mathrm{SU}(2)_L$ doublet. The Yukawa interactions then generate the fermion mass matrices

$$\left(m_f\right)_{ij} = \frac{v}{\sqrt{2}}\hat{Y}^f_{ij} \,, \qquad m_{f_i} = \frac{y_{f_i} v}{\sqrt{2}} \,, \tag{2.15}$$

where $f = u, d, l$ and $\left(m_f\right)_{ij}$ denotes the fermion mass matrix in flavour space, and y_{f_i} the diagonal Yukawa coupling of fermion mass eigenstate f_i. In the mass basis, the Higgs–fermion cou-

pling is therefore proportional to the corresponding fermion mass. Diagonalising the up- and down-type quark Yukawa matrices requires separate unitary rotations in flavour space. The mismatch of the left-handed quark rotations gives rise to the Cabibbo–Kobayashi–Maskawa (CKM) matrix [55, 56], such that in the quark sector the down-type weak-interaction eigenstates are related to the mass eigenstates by

$$d_i' = \sum_j (V_{\mathrm{CKM}})_{ij} d_j \,. \tag{2.16}$$

In the minimal SM, no right-handed neutrino fields are introduced, so the Yukawa Lagrangian above contains no neutrino mass term.

For later collider phenomenology, the resulting broken-phase fermion couplings can be summarised by the following expression for the fermionic part of the Lagrangian. This form is useful here because it collects the physical interactions that enter scattering amplitudes and decay rates after EW symmetry breaking:

$$\begin{aligned}\mathcal{L}_F = &\sum_i \bar{\psi}_i \left(i\gamma^\mu \partial_\mu - m_i - \frac{m_i}{v} H \right) \psi_i \\ &- \frac{g}{2\sqrt{2}} \sum_i \bar{\Psi}_i \gamma^\mu (1-\gamma_5) \left(T^+ W_\mu^+ + T^- W_\mu^- \right) \Psi_i \\ &- e \sum_i Q_i \bar{\psi}_i \gamma^\mu \psi_i A_\mu - \frac{g}{2\cos\theta_W} \sum_i \bar{\psi}_i \gamma^\mu \left(g_V^i - g_A^i \gamma_5 \right) \psi_i Z_\mu \,,\end{aligned} \tag{2.17}$$

with

$$g_V^i = T_3^i - 2Q_i \sin^2\theta_W \,, \qquad g_A^i = T_3^i \,. \tag{2.18}$$

In this expression, the first line contains the fermion kinetic term together with the fermion mass term and the Higgs boson coupling proportional to m_i/v. The second line describes the charged weak current mediated by the $W^\pm$ bosons and acting only on left-handed fermion doublets. The third line contains the electromagnetic current coupled to the photon and the neutral weak current coupled to the Z boson through vector and axial-vector coefficients g_V^i and g_A^i. Here, Ψ_i denotes a weak-isospin doublet, ψ_i a fermion mass eigenstate, $T^\pm$ the weak-isospin raising and lowering operators, and Q_i and T_3^i the electric charge and third weak-isospin component of fermion i, respectively.

2.1.3. Particle content

After EW symmetry breaking, the gauge structure discussed above leads to the physical particle spectrum and interaction pattern schematically shown in Fig. 2.1. The elementary particles of the SM can be grouped into matter fermions with spin $1/2$ and gauge and scalar bosons with integer spin.

The fermions constitute the matter fields of the SM and are arranged into quarks and leptons, each appearing in three generations with identical gauge quantum numbers and increasing masses. Quarks carry colour charge and therefore participate in the strong interaction; they also carry EW charges and are grouped into up-type quarks with electric charge $+2/3$ and down-type quarks with electric charge $-1/3$. Leptons do not carry colour charge and therefore do not couple to gluons; the charged leptons have electric charge -1, while the neutrinos are electrically neutral. In the minimal SM, neutrinos are massless because no right-handed neutrino fields and hence no neutrino Yukawa mass terms are introduced. However, observations of atmospheric and solar neutrino oscillations [57, 58] establish flavour mixing and non-zero neu-

trino mass-squared differences, demonstrating that the minimal SM must be extended. The first generation contains the up and down quarks, u and d, as well as the electron neutrino ν_e and the electron e. The second generation contains the charm and strange quarks, c and s, as well as the muon neutrino ν_μ and the muon μ. The third generation contains the top and bottom quarks, t and b, as well as the tau neutrino ν_τ and the tau lepton τ. The top quark, which is a central object of this thesis, was discovered in 1995 at the Tevatron collider [59, 60]. Since then, it has become the subject of an extensive experimental programme covering precision measurements of its properties, including cross sections of associated production processes, and searches for possible deviations from the SM predictions.
All fermions have corresponding antiparticles with opposite charges. Unless stated otherwise, process labels in this thesis include the corresponding charge-conjugate processes; for example, tH is used to refer to both top-quark and antitop-quark associated Higgs boson production when no distinction is made. The notation q denotes a generic quark, ℓ denotes an electron or muon, and V denotes a weak vector boson, where no distinction is needed.
The preceding discussion of QCD and EW symmetry breaking has already introduced the bosonic degrees of freedom of the SM. In the broken phase relevant for collider phenomenology, these are the gluon g, the weak bosons $W^\pm$ and Z, the massless photon γ, and the Higgs boson H. The Higgs boson is the remaining physical scalar degree of freedom and in the SM has quantum numbers $J^{PC} = 0^{++}$, where J, P, and C denote the total angular momentum, parity, and charge-conjugation quantum numbers, respectively.

2.1.4. Open questions

The SM has been tested in a wide range of precision measurements and is so far consistent with established collider data. Despite this phenomenological success, the SM is not complete, as it cannot account for several observed phenomena and leaves fundamental questions unanswered. Examples include astrophysical observations from galaxy clusters and galactic rotation curves [62, 63], which point to a non-luminous but gravitationally interacting matter component that is not provided by any SM particle. This is referred to as dark matter, and its nature is one of the most important open questions in particle physics and cosmology. The neutrino mass splittings and flavour mixing discussed above provide another example of observed phenomena that are not accommodated in the minimal SM.
The observed baryon asymmetry, i.e. the fact that matter dominates over antimatter in the universe, provides a further important motivation to search for deviations from SM predictions. Its dynamical generation requires baryon-number violation, C- and CP-violating interactions, and a departure from thermal equilibrium, as summarised by the Sakharov criteria [64]. The SM does not provide a mechanism that can generate the observed asymmetry from a baryon-symmetric early-universe state at the required level. In the SM, CP violation in the flavour sector is encoded in a single physical phase of the CKM matrix. In addition, QCD would allow a CP-violating θ term. However, measurements of electric dipole moments (EDMs) constrain the corresponding parameter to be extremely small, whereas there is no intrinsic explanation for this suppression in the SM, which is one aspect of the strong-CP problem. The SM contains baryon-number violating non-perturbative effects and sources of CP violation, but these are insufficient to explain the observed asymmetry [20, 21]. Searches for additional sources of CP violation are therefore well motivated, including in the Higgs sector as studied in this thesis.
On the theoretical side, the Higgs sector is associated with the hierarchy or naturalness problem. If the SM is interpreted as an effective theory below a much higher physical scale, the mass

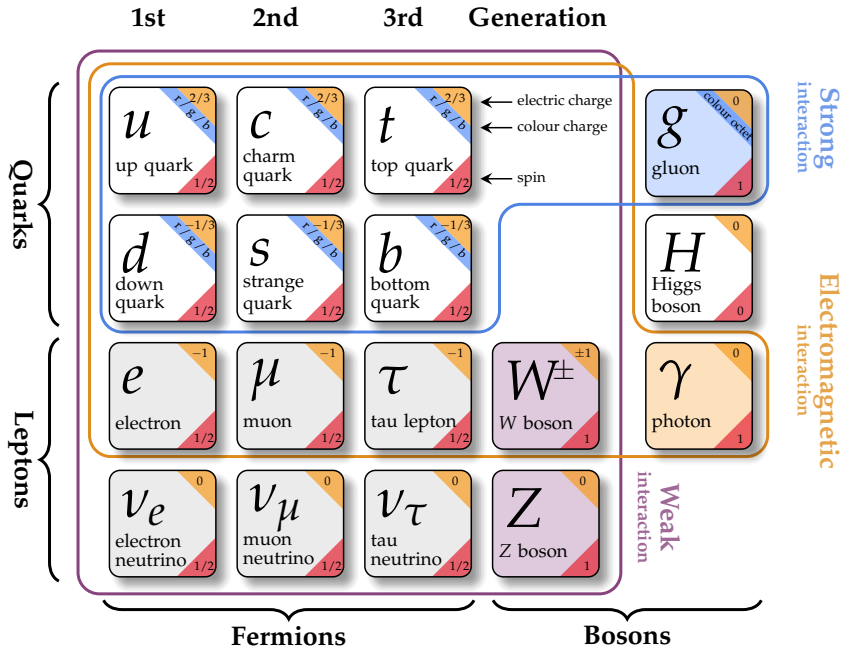


Figure 2.1: Illustration of the elementary particles and interactions in the SM. The electric charge, the spin, and colour charges of each particle are shown. Fermions and bosons as well as the three generations of matter are highlighted. Particles interacting with the strong, the electromagnetic, and the weak interaction are surrounded by coloured boxes. The vertical ordering of the lepton entries is chosen to group electrically charged particles visually and does not represent the ordering within the weak-isospin doublets. Adapted from Ref. [61].

parameter of the scalar Higgs field is generically sensitive to that scale through radiative corrections. Maintaining the EW scale far below such a scale therefore appears to require a tuning of this mass parameter against those corrections [65]. In addition, the measured Higgs boson and top quark masses place the SM vacuum close to the boundary between absolute stability and metastability [16, 17]. These observations do not by themselves imply a breakdown of the SM, but they make the Higgs sector a particularly sensitive place to search for deviations. The Higgs sector and the top quark are of particular interest in this context, since EW symmetry breaking and Yukawa interactions are central to the SM mass-generation mechanism and the top quark has by far the largest Yukawa coupling. Measurements of Higgs boson production and decay in association with top quarks therefore provide a direct way to test this sector and motivate the discussion of modified top-quark Yukawa interactions in the following subsection.

2.2. Modifications to the top-quark Yukawa coupling beyond the standard model

While the SM predicts a purely scalar Yukawa interaction between the Higgs boson and the top quark, several classes of BSM theories allow for deviations in its strength and CP structure. Examples include extended Higgs sectors and composite-Higgs models, in which the Higgs boson arises from new strong dynamics and its couplings can deviate from the SM prediction, often

most prominently in interactions involving the top quark [66]. Among the simplest extensions are two-Higgs-doublet models (2HDMs), in which the SM scalar sector is extended by a second Higgs doublet [67]. These models predict a richer phenomenology with two additional neutral and two charged scalar particles. In a CP-conserving 2HDM, two of the neutral scalars are CP-even, while the third is a CP-odd pseudoscalar, typically denoted by A. The pseudoscalar couplings to fermions are allowed at tree level, while it has no tree-level couplings to ZZ or WW. In a CP-violating version, by contrast, the neutral mass eigenstates are mixtures of CP-even and CP-odd components, so that the observed Higgs boson could simultaneously couple to fermions through scalar and pseudoscalar bilinears. Such a mixture would therefore introduce CP violation in the Yukawa interaction. Concrete studies of the complex 2HDM show that sizeable pseudoscalar Yukawa components of the observed Higgs boson can remain phenomenologically viable in parts of the parameter space [68].

This qualitative difference between fermion and vector-boson couplings is not restricted to the pure pseudoscalar state, but also applies when the observed Higgs boson is not a pure CP-even state and has a CP-odd component. Such a component can generate pseudoscalar Yukawa couplings to fermions already at tree level, while the corresponding CP-odd couplings to massive vector bosons arise only through loop-induced or higher-dimensional interactions and are therefore expected to be suppressed [69]. Higgs-fermion couplings thus provide a particularly sensitive and direct probe of pseudoscalar contributions to the interactions of the observed Higgs boson. Among the fermions, the top quark is especially important because its large mass corresponds to a Yukawa coupling of order unity. The top-quark Yukawa interaction is therefore expected to be particularly sensitive to a pseudoscalar admixture.

Such modifications of the top-quark Yukawa coupling can be formulated in the standard model effective field theory (SMEFT) framework with dimension-six operators that are gauge invariant under $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. In particular, the operator $Q_{u\Phi}^{33} = \left(\Phi^\dagger \Phi\right) \bar{q}_L^3 \widetilde{\Phi}\, t_R$ modifies the top-quark Yukawa interaction [70, 71]. Here, $q_L^3 = (t_L, b_L)^T$ denotes the third-generation left-handed quark doublet in the weak-interaction basis, while t_R denotes the corresponding right-handed up-type quark field.

For the phenomenological interpretation in this thesis, these effects are described with the commonly used κ-framework, in which Higgs boson couplings are rescaled relative to their SM values by coupling modifiers κ_i [72]. For the top-quark Yukawa sector, this is combined with the Higgs Characterisation framework [70], in which the usual CP-even rescaling κ_t is supplemented by an explicit CP-odd component $\tilde{\kappa}_t$. Following Ref. [70], both structures are normalised to m_t/v, such that deviations in the overall strength of the scalar and pseudoscalar top-quark Yukawa interactions are encoded in the dimensionless parameters κ_t and $\tilde{\kappa}_t$. This provides a practical parameterisation of selected Higgs boson interactions with physical particles, rather than a complete gauge-invariant effective field theory [71].

Other dimension-six operators can also affect $t\bar{t}H$ or tH production, for example through modified interactions between top quarks and gauge bosons, such as the Wtb vertex, interactions between gauge bosons and the Higgs boson, or top-quark chromomagnetic interactions; the interplay of these contributions has been studied explicitly in the SMEFT for the production of a Higgs boson or a Z boson in association with a single top quark [73]. Existing global SMEFT fits constrain many of these operator directions [74], but a simultaneous interpretation in the full operator space is beyond the scope of this thesis. Consequently, the $(\kappa_t, \tilde{\kappa}_t)$ interpretation used

here is restricted to modifications of the strength and CP structure of the top-quark Yukawa interaction. The effective parametrisation [70] used throughout this thesis is then given by

$$\mathcal{L}_{Ht\bar{t}} = -\frac{m_t}{v} H\bar{t}\left(\kappa_t + i\tilde{\kappa}_t\gamma_5\right) t\,. \tag{2.19}$$

Here, after EW symmetry breaking and rotation to the mass basis, t is used to denote the top-quark Dirac field. The SM predicts $(\kappa_t, \tilde{\kappa}_t) = (1, 0)$. Although CP transformations include charge conjugation, the distinction between the scalar and pseudoscalar bilinears in Eq. (2.19) is most directly understood through their behaviour under parity. Both $\bar{t}t$ and $i\bar{t}\gamma_5 t$ are even under charge conjugation, but $\bar{t}t$ is parity-even whereas $i\bar{t}\gamma_5 t$ changes sign under spatial reflection. The pseudoscalar contribution therefore introduces a chirality-sensitive structure that can appear in P-odd angular and spin correlations, rather than different coupling strengths to top and antitop quarks.
Other related parameterisations are also commonly used in the literature. In polar coordinates in the coupling-modifier plane, a mixing angle α can be defined as

$$\tan\alpha = \frac{\tilde{\kappa}_t}{\kappa_t}, \tag{2.20}$$

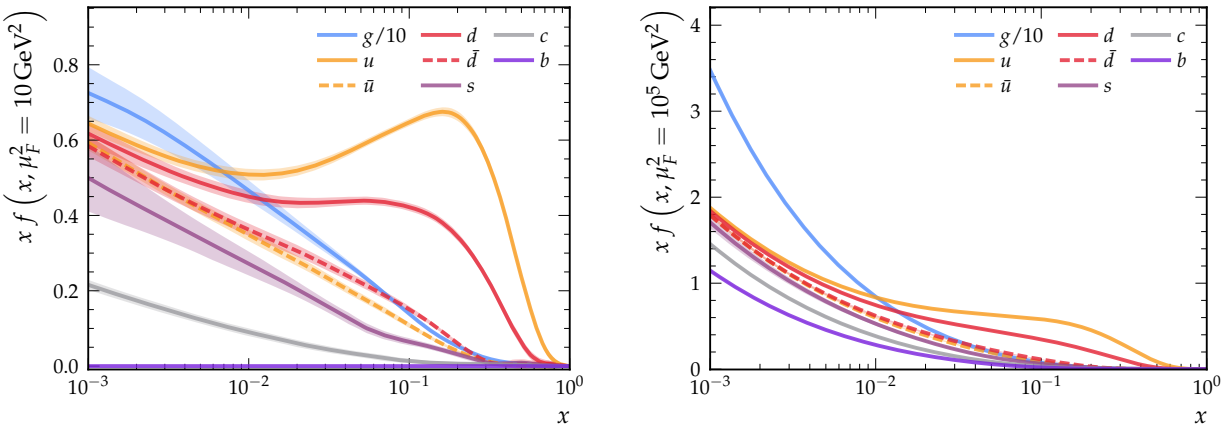
and used with an overall coupling strength modifier related to the above parameterisation as $\sqrt{\kappa_t^2 + \tilde{\kappa}_t^2}$. Furthermore, the parameter

$$f_{CP} = \frac{\tilde{\kappa}_t^2}{\tilde{\kappa}_t^2 + \kappa_t^2}\mathrm{sign}\left(\frac{\tilde{\kappa}_t}{\kappa_t}\right) \tag{2.21}$$

is also commonly used; it is related to the fraction of the cross section induced by the pseudoscalar coupling [75].

2.3. Hadron collider phenomenology

The LHC collides hadrons, which are bound states of quarks held together by the strong force. The hard-scattering processes relevant in this thesis arise in inelastic proton–proton collisions. The full collision occurrence, including the hard process and all accompanying activity, is called an event. The conserved quantum numbers of a hadron are determined by the valence quarks. The proton is a baryon with valence-quark content uud and is overall colour neutral. Its structure, however, also includes virtual gluons and quark-antiquark pairs, which are referred to as sea quarks.
Hadronic interactions at sufficiently high energies can be well described in the parton model, where the term parton denotes the hadron constituents. The parton distribution functions (PDFs) provide the probability density of finding a parton carrying a certain momentum fraction x of the hadron. The PDFs depend on the factorisation scale μ_F associated with the hard interaction, with their evolution naturally expressed in terms of $\ln\mu_F^2$. Modern PDF sets are determined from global fits to multiple measurements, including deep-inelastic scattering data as well as proton–antiproton and proton–proton collider data [76, 77]. The set of PDFs used in the generation of simulation samples in this thesis from Ref. [78] is shown in Fig. 2.2 for two example scales, $\mu_F^2 = 10\,\mathrm{GeV}^2$ and $\mu_F^2 = 10^5\,\mathrm{GeV}^2$. The former is chosen as a low scale for illustrative purposes, whereas the latter is of the order of the interaction scale relevant for the production of $t\bar{t}H$ and



(a) $\mu_F^2 = 10\,\mathrm{GeV}^2$ **(b)** $\mu_F^2 = 10^5\,\mathrm{GeV}^2$

Figure 2.2: PDFs of the proton for the NNPDF3.1 NNLO set [78] with $\alpha_s(m_Z) = 0.118$, shown as a function of the Bjorken scaling variable x at two representative factorisation scales, with $\mu_F^2 = 10\ \mathrm{GeV}^2$ on the left and $\mu_F^2 = 10^5\ \mathrm{GeV}^2$ on the right. The gluon distribution is scaled down by a factor of 10 for visualisation. The displayed uncertainty bands correspond to the PDF uncertainty of the set, including the associated α_s variation.

tH events. The comparison illustrates that this scale evolution can nevertheless produce sizeable changes over the broad range of scales shown: at low scales, the valence quarks are more pronounced at intermediate and large x, while at high scales the gluon and sea-quark densities are enhanced at low x. While the PDFs themselves are non-perturbative objects that are typically determined from data, their dependence on the factorisation scale is perturbatively described by the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) evolution equations [76, 79–82], in which splitting functions encode the probabilities for branchings such as $q \to qg$, $g \to gg$, and $g \to q\bar{q}$. Phenomenologically, increasing the factorisation scale corresponds to probing the proton at finer resolution, such that more collinear parton branchings are resolved. As a result, momentum is redistributed from the valence quarks at larger x towards gluons and sea quarks at lower x.

The physical quantity measured in this thesis is a cross section, denoted by σ, which characterises the probability for a given interaction process to occur. Cross sections are conventionally expressed in units of area with $1\,\mathrm{b} = 10^{-24}\,\mathrm{cm}^2$. In collider physics, the event rate for a process is related to the cross section through

$$\frac{dN}{dt} = L\sigma\,, \tag{2.22}$$

where the instantaneous luminosity L is an experimental quantity that depends on the properties of the colliding beams, such as the number and density of particles per bunch and the bunch crossing frequency. For a data set collected over a finite running period, the corresponding integrated luminosity is denoted by $\mathcal{L}_{\mathrm{int}}$. The expected event yield is assumed to follow a Poisson distribution because the events are independent and occur with a constant average rate. The expectation is then given by

$$N_{\mathrm{exp}} = \mathcal{L}_{\mathrm{int}}\,\sigma\,A\,\varepsilon\,, \tag{2.23}$$

where A denotes the acceptance, i.e. the fraction of produced events falling into the accessible phase space, and ε the efficiency for reconstructing and selecting those events. At sufficiently large scale $Q \gg \Lambda_{\mathrm{QCD}}$, where Λ_{QCD} denotes the characteristic scale at which the strong interac-

tion becomes non-perturbative, the partons in the hadron are asymptotically free as $\alpha_s(Q)$ becomes small. For sufficiently hard scales, the hadronic cross section can therefore be factorised into a convolution of the PDFs, which capture the non-perturbative proton structure, with a partonic hard-scattering cross section, which can be calculated perturbatively [83]. For the production of a final state X in proton–proton collisions with partons collinear with the beam axis, this can be written as

$$\sigma_{pp\to X}(s) = \sum_{i,j} \int_0^1 dx_1 \int_0^1 dx_2 \, f_i(x_1, \mu_F^2) \, f_j(x_2, \mu_F^2) \, \hat{\sigma}_{ij\to X}(\hat{s}, \mu_F, \mu_R) \,, \tag{2.24}$$

where $\hat{s} = x_1 x_2 s$ denotes the centre-of-mass energy squared of the colliding partons, $f_i(x, \mu_F^2)$ is the PDF of parton species i, and $\hat{\sigma}_{ij\to X}$ is the partonic cross section for the hard process. The renormalisation scale μ_R is introduced when ultraviolet divergences in the perturbative calculation are absorbed into redefined parameters and couplings. It enters the prediction through the running coupling and, at finite perturbative order, through a residual scale dependence of the hard-scattering cross section.

In addition to the hard-scattering process, the remaining partons of the interacting protons can interact with each other, typically in softer inelastic scatterings. This is referred to as the underlying event. At a hadron collider, a further important effect is pileup. Because bunches of protons are collided, additional proton–proton collisions contribute to the signature of an event of interest. Furthermore, out-of-time pileup occurs, which denotes contributions from neighbouring bunch crossings that cannot always be distinguished experimentally. After the hard scattering, the colour-charged particles can perturbatively radiate gluons through the parton shower (PS), while quarks can also radiate photons. The PS approximates the dominant higher-order radiation in regions where emissions are soft or collinear to the emitting particle, because such emissions are enhanced in QCD [84]. By generating a sequence of probabilistic branchings, it effectively accounts for the corresponding higher-order logarithmic contributions beyond the fixed-order hard process. In this way, the PS connects the fixed-order calculation to the multiparticle final state observed experimentally, before non-perturbative hadronisation is applied. The modelling of additional radiation is therefore important for jet multiplicities, event kinematics, and the acceptance of the selections used in this thesis. Subsequently, the non-perturbative hadronisation takes place, in which colour-neutral bound states are formed, typically with high particle multiplicity, such that high-energy quarks and gluons become visible in the form of collimated sprays of particles, so-called jets.

Relevant for this thesis is the distinction between a total cross section, describing the interaction rate of a process inclusive over final-state properties, and a fiducial cross section, which is restricted to a subset of allowed final-state momenta at the particle level and is chosen to approximate the experimentally accessible phase space. Here, particle level refers to stable particles after hadronisation and decay of short-lived resonances, to be distinguished from the parton level where partonic states are taken into account without PS and hadronisation. A further distinction is the detector level, which refers to reconstructed objects after including detector effects and event reconstruction. This level is used to describe the experimental data.

2.3.1. Higgs boson production at the Large Hadron Collider

At the LHC, the Higgs boson can be produced through several different production modes. Representative Feynman diagrams at leading order (LO) for the Higgs boson production modes

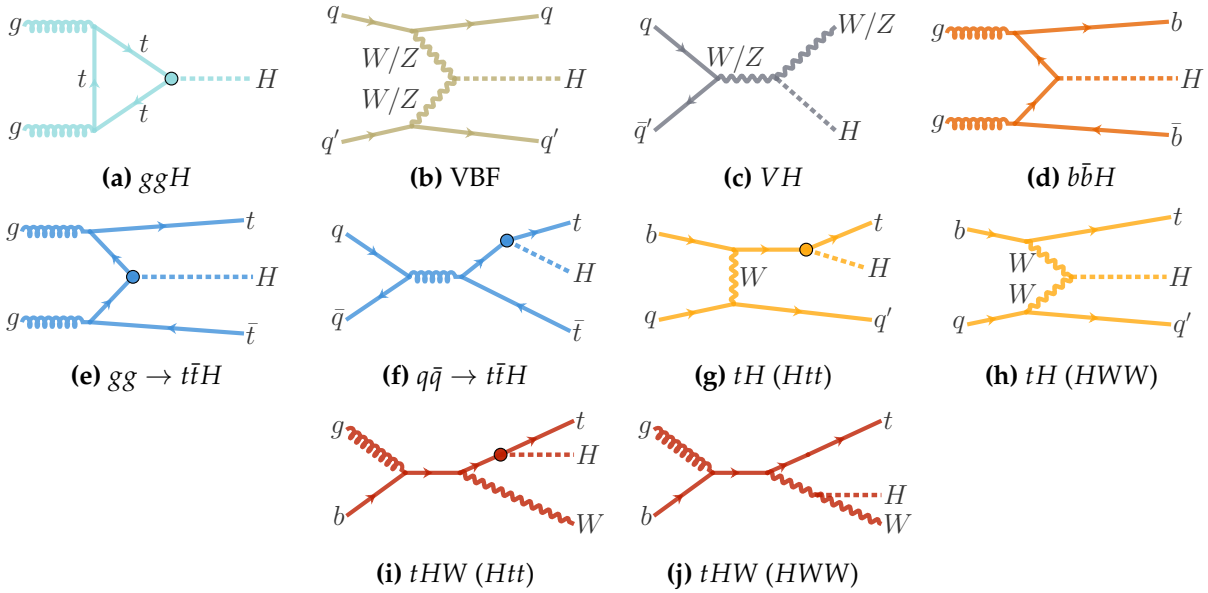


Figure 2.3: Representative LO Feynman diagrams for the most important Higgs boson production modes at the LHC in the SM. The first row shows the production modes which are regarded as part of the background in this study. The second row shows example diagrams for the gluon- and quark-initiated contributions to $t\bar{t}H$ production, and two diagrams for tH production. The third row shows the corresponding tHW production diagrams. For tH and tHW, both the diagram involving the top-quark Yukawa coupling and the diagram involving the HWW coupling are shown, since their destructive interference plays a central role in the phenomenology of single-top-quark-associated Higgs boson production. The top-quark Yukawa coupling is marked with the filled circle.

considered as part of the backgrounds of this analysis and the comparably rare top-quark associated signal processes $t\bar{t}H$, tH, and tHW are shown in Fig. 2.3.

Unless stated otherwise, the inclusive SM production cross sections and quoted theory uncertainties in the following are given for $\sqrt{s} = 13.6\,\text{TeV}$ and a Higgs boson mass of 125 GeV, and are taken from the review in Ref. [52]. These values are quoted to summarise the current theory status of the production modes, and can differ slightly from the cross sections used to normalise the simulated samples in the analysis. The latter follow the LHC Higgs Working Group recommendations adopted for the analysis setup [85] and are listed, after interpolation to the Higgs boson mass hypothesis used here, in Sec. 5.2.1. The reader is referred to that source and references therein for a more comprehensive summary of the individual production modes and the corresponding results of theoretical calculations. Only the basic production picture is summarised here; the resulting fiducial acceptances and reconstructed kinematic properties of the signal modes, both in the SM and for modified Yukawa couplings, are discussed later in Sec. 6.1.

Among all Higgs boson production modes at the LHC, gluon fusion, denoted as ggH, has by far the highest cross section. It proceeds through the loop-induced coupling of the Higgs boson to gluons, mediated predominantly by the top quark due to its large Yukawa coupling. The inclusive SM cross section is predicted to be about 52.1 pb with a relative uncertainty of $^{+6.9\%}_{-9.5\%}$. For the ggH cross section, higher-order corrections have been found to be particularly important. The inclusive prediction includes QCD corrections up to next-to-next-to-next-to-leading order (N3LO) and EW corrections at next-to-leading order (NLO). At Born level, the ggH process

has a $2 \to 1$ kinematic structure, while accompanying jets, recoil, and non-zero Higgs boson transverse momentum arise from additional QCD radiation.
Vector-boson fusion (VBF) has the second-highest inclusive cross section and describes Higgs boson production from the fusion of W or Z bosons radiated off incoming quarks. The inclusive SM cross section prediction is $4.11\,\text{pb} \pm 3.3\%$. The inclusive cross section has been calculated up to N3LO in QCD, while NLO EW corrections are known. Its characteristic experimental signature consists of two energetic forward jets with a large rapidity separation and a large dijet invariant mass, while central hadronic activity is comparatively suppressed.
The next-highest cross sections are those of associated production with a weak vector boson, denoted as $W^{\pm}H$ and ZH or collectively as VH. The leading contributions are quark-induced, while in the ZH case an additional loop-induced $gg \to ZH$ contribution enters the total rate. This loop-induced contribution contains top-quark loops and is therefore in principle sensitive to modifications of the top-quark Yukawa coupling, but this dependence is neglected for the resonant-background treatment in this thesis. The inclusive SM cross section predictions are $1.47\,\text{pb}^{+2.4\%}_{-2.6\%}$ for WH and $0.954\,\text{pb}^{+5.6\%}_{-5.0\%}$ for ZH. The theoretical description of these processes receives several distinct higher-order QCD and EW contributions. QCD corrections up to next-to-next-to-leading order (NNLO) together with NLO EW corrections are available, and the inclusive cross section has more recently also been determined up to N3LO in QCD. Experimentally, the characteristic feature is the presence of the accompanying weak boson, whose leptonic, hadronic, or invisible decays lead to a broad range of final states. Compared with ggH production, the additional hard objects in VBF and VH production lead to a typically harder Higgs boson transverse-momentum spectrum. The associated production with bottom quarks, $b\bar{b}H$, contributes only at the percent level to the total Higgs boson production and is considered as part of the background in this thesis.
All signal processes contain the top-quark Yukawa coupling already at tree level, which makes them directly sensitive to the coupling and reduces the reliance on assumptions about possible loop-level new-physics effects. Among them, the $t\bar{t}H$ process has the highest cross section. At the LHC, the process is dominated by the gluon-initiated channel, while a smaller contribution arises from quark-antiquark annihilation. Representative LO topologies for these two initial states are shown in Figs. 2.3e and 2.3f. A recent NNLO+NNLL+EW prediction reports an inclusive SM cross section of 592 fb [86]. It combines complete NLO corrections, including EW effects, with approximate NNLO QCD corrections and next-to-next-to-leading-logarithmic (NNLL) soft-gluon resummation. The total relative uncertainty is $^{+3.4\%}_{-3.6\%}$, obtained by adding the quoted individual uncertainty components in quadrature. A complementary NNLO event-generator calculation matched to PS has also become available in a different setup, using approximate two-loop amplitudes and providing differential predictions for fiducial final states including $H \to \gamma\gamma$ [87]. Experimentally, the presence of a top-quark pair together with the Higgs boson leads to richer and higher-multiplicity final states than in the other production modes.
Furthermore, $t\bar{t}H$ predictions were provided for scenarios with non-standard Yukawa couplings. In this context, Ref. [29] studied this process at NLO in QCD including PS effects in the Higgs Characterisation framework and showed that NLO corrections substantially reduce the scale uncertainties, while also discussing the impact on kinematic distributions in CP-odd or admixed scenarios. More recently, Ref. [88] extended this discussion to the off-shell leptonic final state with full double-, single-, and non-resonant contributions at NLO in QCD and showed that both NLO corrections and off-shell effects can already be relevant at the level of integrated fiducial cross sections and become even more important in differential distributions. For the present

thesis, these results are particularly relevant because they show that non-standard top-quark Yukawa couplings modify not only the total production rate but also the kinematic properties that enter the fiducial acceptance and the later event classification. More general SMEFT studies with decaying top quarks have shown that higher-order corrections, fiducial requirements, and effective operators in the top-quark decay can also affect the shapes of standard $t\bar{t}H$ observables significantly, so the restricted $(\kappa_t, \tilde{\kappa}_t)$ framework used in this thesis captures only a subset of possible effects from BSM physics [89].

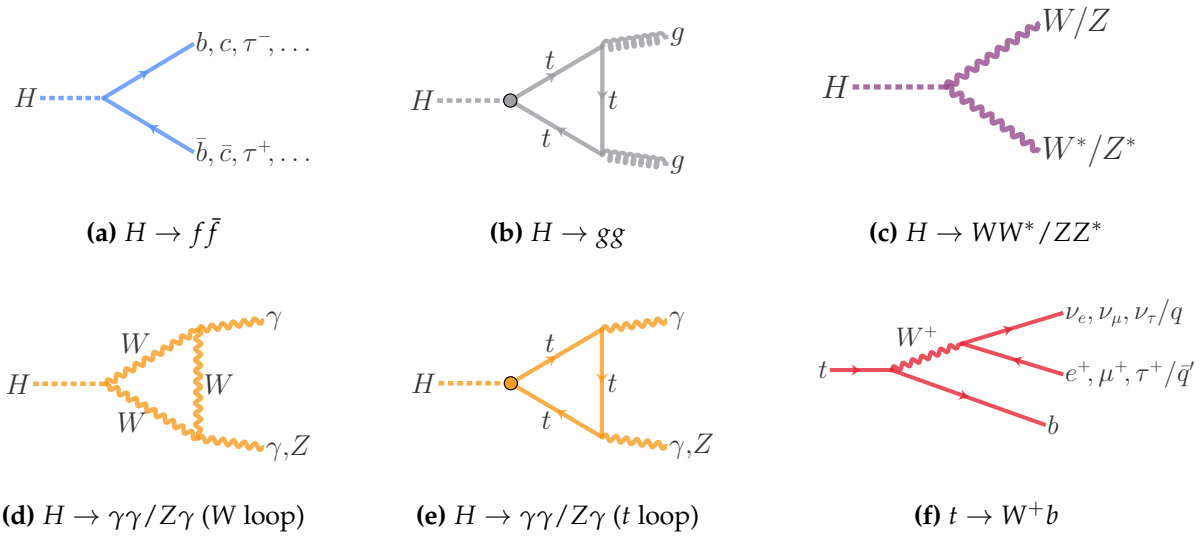
Within the framework used, an important simplification is that the interference between amplitudes from CP-even and CP-odd top-quark Yukawa couplings is itself CP-odd. It may therefore be non-zero locally in phase space, but cancels in inclusive cross sections or in observables integrated over a CP-symmetric phase-space region. This cancellation can be understood as the averaging over CP-conjugate configurations: for each event configuration, the corresponding CP-conjugate configuration contributes with the opposite sign to the interference, and both contributions are summed in an inclusive or CP-symmetric observable. Sensitivity to this interference therefore requires an observable that retains this sign, such as an appropriately defined signed azimuthal angle, a spin-sensitive decay observable, or a triple-product correlation [88, 90]. For $t\bar{t}H$, the scalar–pseudoscalar interference term in the squared matrix element vanishes exactly at LO; at NLO, non-zero local contributions arise, but they cancel for sufficiently inclusive or CP-symmetrically integrated observables [29, 88]. This motivates the omission of terms linear in $\tilde{\kappa}_t$ or proportional to $\kappa_t\tilde{\kappa}_t$ in the inclusive cross-section parameterisations used later in this thesis, unless an observable is explicitly constructed to isolate the CP-odd interference. The standard observables considered later, such as rates, transverse momenta, invariant masses, and jet or lepton multiplicities, are not designed for this purpose. The implementation details of the simulated datasets used in this study are discussed in Sec. 5.2.1. The impact of changes in the Yukawa coupling on fiducial acceptances and reconstructed kinematic observables in the signal processes is discussed in Sec. 6.1.

The Higgs boson can also be produced in association with a single top quark. The dominant contribution is given by the t-channel process, often denoted as tHq or $tHqb$ to emphasise the accompanying light quark and, in four-flavour-scheme descriptions, the initial $g \to b\bar{b}$ splitting. In this thesis, this dominant contribution is denoted simply as tH. The inclusive SM cross section is about 85 fb with a relative uncertainty of $^{+8.5\%}_{-18\%}$ [52].

The inclusive t-channel cross section has been computed at NLO in QCD in both the four- and five-flavour schemes within the Higgs Characterisation Model in Ref. [30]. The tH process is not charge symmetric: tH production has a cross section that is about a factor of 1.9 larger than $\bar{t}H$ at 13 TeV because the quark-initiated contribution benefits from the valence-quark content of the proton [30]. A later NLO QCD+EW calculation considered the related tHj final state in the five-flavour scheme, inclusive over the underlying single-top-quark associated production topologies rather than separating the t-channel from the s-channel and tW-associated contributions introduced below [91]. In that setup, the EW corrections reduce the cross section by about 3% [91].

Closely related but even rarer in the SM is the associated production with a single top quark and a W boson, tHW, for which NLO QCD predictions were presented in Ref. [31]. That calculation yields an inclusive cross section of 17.2 fb with a total relative uncertainty of $^{+8.0\%}_{-9.3\%}$. At NLO, resonant top-quark contributions that overlap with $t\bar{t}H$ production are removed using a diagram-removal scheme including interference effects [31].

Representative Feynman diagrams for the processes tH and tHW are shown in Fig. 2.3 **(g)–(j)** at LO. They share the same characteristic suppression mechanism in the SM, because the full



(a) $H \to f\bar{f}$ **(b)** $H \to gg$ **(c)** $H \to WW^*/ZZ^*$

(d) $H \to \gamma\gamma/Z\gamma$ (W loop) **(e)** $H \to \gamma\gamma/Z\gamma$ (t loop) **(f)** $t \to W^+b$

Figure 2.4: Representative LO Feynman diagrams for the most important Higgs boson and top-quark decay modes relevant for this thesis. The first row shows generic fermionic Higgs boson decays, the loop-induced decay to gluons via a virtual top-quark loop, and the tree-level decay into weak vector bosons. The second row shows two representative loop topologies contributing to the $\gamma\gamma$ and $Z\gamma$ final states, as well as the dominant top-quark decay $t \to W^+b$ together with representative leptonic and hadronic W-boson decays. The top-quark Yukawa coupling is marked with the filled circle.

scattering amplitude receives contributions where the Higgs boson is radiated either from the top quark or from the W boson. In the SM, these contributions interfere destructively in the full scattering amplitude, so that they cancel to a large extent. Allowing for a CP-odd admixture in the Yukawa coupling and a possible rescaling of the weak-vector-boson coupling by κ_V, the amplitudes can be schematically written as

$$\mathcal{A}(tH/tHW) \sim \kappa_t \mathcal{A}_{Htt} + \kappa_V \mathcal{A}_{HWW} + \tilde{\kappa}_t \mathcal{A}_{Att}, \tag{2.25}$$

where Htt, Att, and HWW label the CP-even and CP-odd Yukawa couplings to top quarks and the coupling to the W boson, respectively. This suppression makes both processes very rare in the SM but at the same time renders them exceptionally sensitive to changes in the sign and CP structure of the top-quark Yukawa interaction. Throughout this thesis, $\kappa_V = 1$ is assumed. The s-channel tH contribution amounts to approximately 3% of the total tH cross section and differs substantially kinematically [30], such that it is negligible within the current precision and neglected in this thesis.

2.3.2. Decays of the Higgs boson and the top quark

The Higgs boson and the top quark decay rapidly, and only their decay products interact with the detector material and leave measurable signatures. Representative LO Feynman diagrams for the dominant Higgs boson decay channels at $m_H = 125\,\text{GeV}$ and for the top-quark decay are shown in Fig. 2.4. The SM Higgs boson branching ratios and quoted theory uncertainties in the following are taken from Ref. [52] for $m_H = 125\,\text{GeV}$.

Because the Higgs boson couples at tree level both to fermions through Yukawa interactions and to weak vector bosons through the EW symmetry-breaking sector, it gives rise to the tree-level decays $H \to f\bar{f}$ for all massive fermions with $m_f < m_H/2$, and $H \to WW^*/ZZ^*$, where one weak boson is off shell. Accordingly, the majority of Higgs bosons decay via $H \to b\bar{b}$ with $\mathcal{B}(H \to b\bar{b}) = 58.2\%$ and a relative uncertainty of 1.7%. Other sizeable fermionic decay modes are $H \to \tau^+\tau^-$ and $H \to c\bar{c}$. Their branching ratios are $\mathcal{B}(H \to \tau^+\tau^-) = 6.27\% \pm 2.3\%\,(\text{rel.})$ and $\mathcal{B}(H \to c\bar{c}) = 2.89\%\,^{+6.8\%}_{-2.2\%}\,(\text{rel.})$. On the other hand, the muonic decay is suppressed with $\mathcal{B}(H \to \mu^+\mu^-) = 2.18 \times 10^{-4} \pm 2.4\%\,(\text{rel.})$. The branching ratio of $H \to WW^*$ is the second-highest, with $\mathcal{B}(H \to WW^*) = 21.4\% \pm 2.2\%\,(\text{rel.})$. The decays $H \to gg$ and $H \to ZZ^*$ have branching ratios of $8.18\% \pm 7.2\%\,(\text{rel.})$ and $2.62\% \pm 2.2\%\,(\text{rel.})$, respectively. The decay to gluons is induced by a heavy-quark loop with the top quark as the dominant contribution, but it is experimentally overwhelmed by the QCD multijet background and hence not accessible as a standalone decay channel at the LHC.

The diphoton decay and the related $H \to Z\gamma$ decay mode are much rarer, with branching ratios of $\mathcal{B}(H \to \gamma\gamma) = 2.27 \times 10^{-3} \pm 2.8\%\,(\text{rel.})$ and $\mathcal{B}(H \to Z\gamma) = 1.53 \times 10^{-3} \pm 6.9\%\,(\text{rel.})$. For $H \to \gamma\gamma$, the amplitude is dominated by the W-boson loop and receives a destructively interfering contribution from the top-quark loop, with the two representative LO diagrams shown in Fig. 2.4 **(d)** and **(e)**. The corresponding decay width is known with full NLO QCD corrections, additional three-loop QCD corrections, and NLO EW corrections. For $m_H \simeq 125\,\text{GeV}$, the known radiative effects are well below the percent level [52]. Despite the small branching ratio and the sizeable continuum background relative to the signal, the fully reconstructed two-photon final state profits from an excellent invariant-mass resolution in the CMS detector. The signal therefore appears as a narrow peak on top of a smoothly falling background, making this channel particularly powerful for Higgs boson studies at the LHC.

This interference pattern makes the diphoton branching ratio also sensitive to the relative sign and, more generally, the coupling structure of the top-quark Yukawa interaction. When allowing for a CP-odd contribution and a variable coupling to vector bosons, the induced change in the diphoton branching ratio can be parametrised [69] as

$$\frac{\mathcal{B}(H \to \gamma\gamma)}{\mathcal{B}(H \to \gamma\gamma)_{\text{SM}}} = 1.6 \left((\kappa_V - 0.21\,\kappa_t)^2 + 0.16\,\tilde{\kappa}_t^2 \right) , \tag{2.26}$$

where again $\kappa_V = 1$ and all other Yukawa couplings are fixed to their SM values in this thesis. This illustrates that modifications of the top-quark Yukawa interaction affect measurements in the diphoton decay channel not only through the production processes, but also through the diphoton branching ratio itself. This introduces a model dependence into total Higgs boson cross section measurements in the diphoton final state because the experimentally accessible observables include the product of a cross section and the branching ratio. In the fiducial cross section measurement performed in this thesis, the result is largely independent of assumptions on the branching ratio because the presence of two photons is part of the fiducial definition. A small dependence arises because resonant backgrounds from $H \to \gamma\gamma$ processes other than the signal modes are included in the background contributions, which are scaled with the SM branching ratio without additional flexibility.

The top quark decays almost exclusively through $t \to W^+b$. Due to its large mass, the lifetime is only of the order $10^{-25}\,\text{s}$, which is shorter than the hadronisation time scale $1/\Lambda_{\text{QCD}} \sim 10^{-24}\,\text{s}$ and the spin-decorrelation time scale $m_t/\Lambda_{\text{QCD}}^2 \sim 10^{-22}\,\text{s}$. The top quark therefore almost al-

ways decays before it hadronises and transfers its spin information to the kinematic distributions of the products of its decay chain.
For this thesis, it is useful to distinguish between top-quark decays that yield an electron or muon in the final state and those that do not, because this is used to define the hadronic and leptonic channels of this analysis. Using the W- and τ-lepton branching ratios from Ref. [92], the effective branching ratio for a top quark to yield an electron or muon through $W \to e\nu$, $W \to \mu\nu$, or $W \to \tau\nu$ followed by $\tau \to e/\mu\,\nu\nu$, is about 25.5%. The complementary fraction of about 74.5% therefore contributes to the hadronic category at truth level. For the $t\bar{t}$ and tW processes, which include the decay of two W bosons, this implies branching ratios of about 6.5% for dilepton decays, 38.0% for lepton+jets decays, and 55.5% for all-hadronic decays, where lepton again refers to electrons and muons only.

2.4. Overview of experimental research on the Higgs boson

Many Higgs boson measurements have now reached the precision regime, although several Higgs boson couplings are still only weakly constrained. The following brief overview is inspired by the review in Ref. [52] and is complemented by more recent results where useful. For a broader overview of Higgs boson measurements beyond the points most relevant for this thesis, the reader is referred to that review.

2.4.1. General properties

After the discovery, the Higgs boson has been studied in multiple production and decay modes, and the properties of the observed particle have been found to be compatible with the SM predictions within current experimental uncertainties [34, 37, 38, 93]. Among the earliest key results was the establishment that, in its interactions with weak vector bosons, the observed state is compatible with a spin-zero boson with scalar, SM-like couplings, as inferred from analyses of its decays to WW^*/ZZ^*. In particular, Run-1 analyses by the ATLAS and CMS Collaborations strongly favoured the scalar $J^P = 0^+$ hypothesis over pure pseudoscalar and spin-2 alternatives in these bosonic channels, at high confidence levels [22, 23]. The overall signal strength, defined as a rate modifier relative to the SM expectation, was measured by the CMS Collaboration to be $\mu = 1.014^{+0.055}_{-0.053}$ in a combination of measurements using the Run-2 dataset [34], and with comparable precision in a recent combination by the ATLAS Collaboration [93]. The Higgs boson has been observed in each of the $H \to b\bar{b}$, $H \to ZZ^*$, $H \to WW^*$, $H \to \tau^+\tau^-$, and $H \to \gamma\gamma$ decay channels [37, 38], while evidence for the rare decays $H \to \mu^+\mu^-$ [94, 95] and $H \to Z\gamma$ [96] has also been reported. In terms of production modes, the main processes ggH, VBF, VH, and $t\bar{t}H$ have been observed experimentally, whereas the rare tH process and Higgs boson pair production remain below the level of established evidence.
The Higgs boson mass is known to sub-permille precision, with the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ decay channels providing the strongest sensitivity. Using the combination of Run-1 and Run-2 data, the CMS Collaboration measured $m_H = 125.08 \pm 0.12\,\text{GeV}$ in the $H \to ZZ^* \to 4\ell$ channel [97]. The SM predicts a total Higgs boson width of about 4.1 MeV, which is far below the experimental mass resolution at the LHC. The most stringent constraint that relies only on on-shell Higgs boson final states currently comes from the diphoton channel, where the CMS Collaboration exploits the interference between $gg \to H \to \gamma\gamma$ and the continuum $gg \to \gamma\gamma$ background and obtains $\Gamma_H < 92\,\text{MeV}$ at 95% confidence level (CL) [98]. A significantly stronger, but also more model-dependent, determination is obtained in the $H \to ZZ^* \to 4\ell$ channel when

off-shell information is included, yielding $\Gamma_H = 3.0^{+2.0}_{-1.5}\,\text{MeV}$, consistent with the SM prediction [97]. Comparable off-shell constraints have also been obtained by the ATLAS Collaboration [99]. With the observation of the $t\bar{t}H$ process, the $H \to b\bar{b}$ decay, and the $H \to \tau^+\tau^-$ decay, the Yukawa couplings to the heavy third-generation fermions are established experimentally, while $H \to \mu^+\mu^-$ even provides evidence for a second-generation Yukawa coupling. Since the present thesis uses the diphoton decay channel, the loop-induced interaction of the Higgs boson with photons is also relevant. The corresponding modifier κ_γ describes the net loop-induced $H\gamma\gamma$ amplitude, including the combined SM W-boson and top-quark contributions and possible additional loop effects, rather than a separate tree-level coupling [70, 72]. In the corresponding coupling fits, the ATLAS and CMS Collaborations measure $\kappa_\gamma = 0.97 \pm 0.06$ and $\kappa_\gamma = 1.07 \pm 0.07$, respectively [34, 93].
Measurements sensitive to the CP properties of Higgs boson couplings have also been performed for particles other than the top quark. In the $H \to \tau\tau$ decay channel, angular correlations of the tau-lepton decay products constrain the corresponding CP mixing angle to $\alpha_{H\tau\tau} = 7^\circ \pm 16^\circ$ in the combination of Run-2 and partial Run-3 data collected with the CMS detector, compatible with the SM expectation [100]. Anomalous CP-odd couplings to weak vector bosons are also tightly constrained by Run-2 measurements in $H \to ZZ^*/WW^*$ decays and in vector-boson-mediated production, with no evidence for CP violation observed [101–103]. These interpretations are usually formulated in SMEFT or anomalous-coupling frameworks rather than through a simple $\tilde{\kappa}_V$ modifier analogous to the fermionic pseudoscalar coupling. Dedicated searches for WH production through vector-boson scattering exploit the destructive interference between amplitudes involving the HWW and HZZ couplings in the SM; a negative relative sign makes this interference constructive and strongly enhances the rate, allowing such scenarios to be excluded [104, 105]. These results motivate treating the weak-vector-boson coupling as SM-like in this thesis. With the convention $\kappa_V = 1$, a flipped relative sign between the top-quark Yukawa and vector-boson amplitudes is described by negative values of κ_t.

2.4.2. $t\bar{t}H$ and tH production and the top-quark Yukawa coupling

Among collider observables, the $t\bar{t}H$ and tH processes provide the most direct and least model-dependent access to the top-quark Yukawa coupling, because the coupling enters already at tree level. Experimentally, however, one still measures $\sigma \times \mathcal{B}$ in a given Higgs boson decay channel, such that the extraction of the coupling strength is not entirely independent of assumptions on the Higgs decay couplings and the total width. Current Higgs boson measurements address this model dependence with different levels of generality. Inclusive signal-strength measurements report rates relative to SM expectations and therefore inherit the extrapolation, acceptance, and branching-ratio assumptions entering the reference prediction. The simplified template cross section (STXS) scheme reduces this model dependence by reporting signal strengths in predefined production and kinematic regions, while fiducial cross-section measurements further restrict the quoted result to an experimentally accessible phase space [41, 45, 72, 106]. In practice, however, the event categories used to extract STXS or dedicated signal strengths often rely on multivariate discriminants trained with SM signal simulation, so their optimality and interpretation can still depend on the assumed signal model. Dedicated CP-sensitive $t\bar{t}H$ and tH analyses, in contrast, often maximise sensitivity with categories or classifiers optimised for specific coupling hypotheses or CP-sensitive observables [26–28]. The analysis presented in this thesis follows an intermediate strategy: it measures fiducial cross sections in a $t\bar{t}H$- and tH-enriched $H \to \gamma\gamma$ phase space, while retaining multivariate separation power and explicitly reducing and

testing the dependence of the result on variations of $(\kappa_t, \tilde{\kappa}_t)$ in the signal model. The detailed implementation of this strategy is discussed in Chap. 4.

The $t\bar{t}H$ process was observed at both the ATLAS and CMS experiments in 2018 [32, 33], and it has since been measured in several final states, including $H \to \gamma\gamma$, $H \to b\bar{b}$, and multilepton analyses sensitive to $H \to WW^*$, $H \to \tau\tau$, and $H \to ZZ^*$. A recent combination by the CMS Collaboration of $t\bar{t}H$ measurements in the diphoton, $b\bar{b}$, and multilepton channels yields a signal strength of $\mu_{t\bar{t}H} = 0.77^{+0.18}_{-0.15}$ [34]. The measurement of the tH process remains experimentally much more challenging, because the cross section is lower while the rejection of backgrounds is more difficult. Moreover, $t\bar{t}H$ events constitute an important background for tH searches. A recent combination by the CMS Collaboration measured $\mu_{tH+tHW} = 5.9^{+2.4}_{-2.2}$, corresponding to an excess of 2.2 standard deviations over the SM hypothesis [34]. Within that combination, the fitted tH signal strength also shows the largest deviation from the SM expectation. The ATLAS Collaboration reports a similar upward fluctuation: the Run-2 combination of the $H \to b\bar{b}$ and multilepton channels shows an observed excess of 2.8 standard deviations for $tH + tHW$ production over the background-only hypothesis, compared with 0.4 standard deviations expected for the SM signal [35]. Within a broader combination including the $H \to \tau\tau$ channel, the excess is reduced; the signal strength is determined as $\mu_{tH+tHW} = 4.5^{+2.3}_{-2.1}$, and an upper limit $\mu_{tH+tHW} < 8.5$ is set at 95% CL [93].

Other ways to constrain the top-quark Yukawa coupling are more model dependent. The ggH production rate and the $H \to \gamma\gamma$ branching ratio are sensitive to the top-quark Yukawa coupling through loops, but assumptions regarding the presence of additional heavy coloured or charged particles entering the loops are required. Furthermore, assumptions on the total Higgs boson width and on other Higgs boson couplings are necessary. The strongest model-dependent constraints on a CP-odd admixture in the top-quark Yukawa coupling can be obtained from the EDM of the electron. Under the assumptions that the electron Yukawa coupling is SM-like and that no cancellations with other CP-violating contributions occur, the current bound corresponds to $|\tilde{\kappa}_t| < 0.001$ [52, 107]. Precision measurements of $t\bar{t}$ production provide additional indirect sensitivity, because virtual Higgs boson exchange affects the EW corrections to the $m_{t\bar{t}}$ spectrum near the production threshold. This approach has been used by the CMS Collaboration in the lepton+jets and dilepton final states and more recently by the ATLAS Collaboration in the lepton+jets channel to extract the top-quark Yukawa coupling from $t\bar{t}$ kinematic distributions [108–110]. Measurements of four-top-quark production offer another indirect handle since the cross section receives contributions from diagrams with off-shell Higgs boson exchange between the top quarks, and both the CMS and ATLAS Collaborations have interpreted these measurements to obtain constraints on the top-quark Yukawa coupling [111, 112]. These constraints are currently weaker than those from direct $t\bar{t}H$ and tH measurements.

The current direct constraints in the $(\kappa_t, \tilde{\kappa}_t)$ plane from the CMS experiment are shown in Fig. 2.5. They are obtained from a combination of measurements targeting the $t\bar{t}H$ and tH production modes in the $H \to \gamma\gamma$, $H \to ZZ^*$, and $H \to b\bar{b}$ decay channels and in an orthogonal multilepton phase space, under the assumption of SM Higgs boson branching ratios [28]. There, the $H \to \gamma\gamma$ measurement [26] provides the most sensitive contribution, although that analysis was not optimised to exploit the potential sensitivity from tH events explicitly. The results are compatible with the SM prediction. The measurement with the CMS dataset sets a limit of $f_{CP} < 0.85$ at 95% CL [28]. It leaves a small region in the negative-κ_t half-plane allowed at 95% CL, around $\kappa_t \approx -0.7$ with a small possible $\tilde{\kappa}_t$ admixture. Similarly, the ATLAS measurement in the diphoton decay channel excludes $|\alpha| > 43^\circ$ at 95% CL [27]. Thus, sizeable CP-odd admixtures of the

top-quark Yukawa coupling remain experimentally allowed, motivating continued measurements of $t\bar{t}H$ and tH production cross sections with corresponding coupling interpretations.

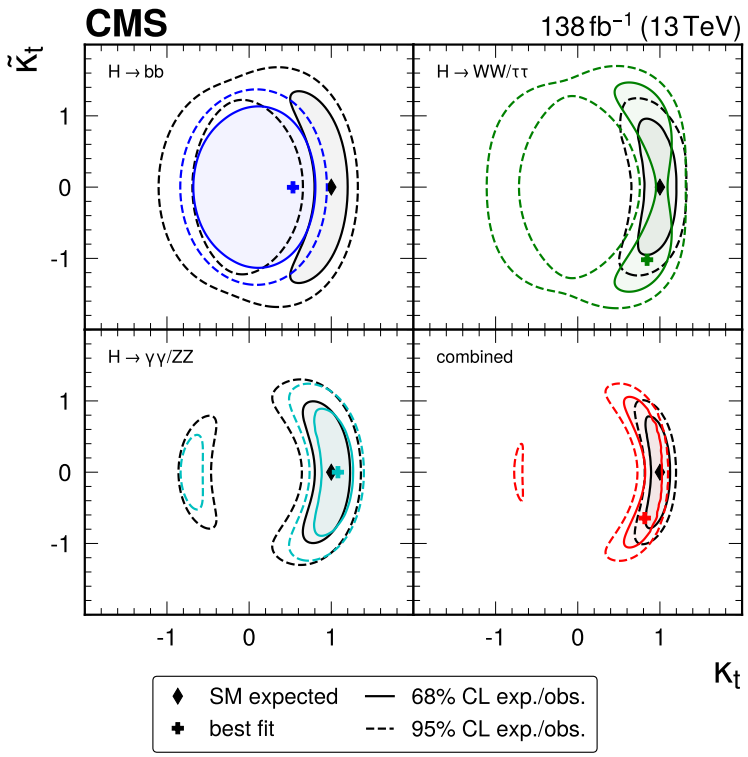


Figure 2.5: Observed (coloured) and expected (black) 68% and 95% confidence regions in the $(\kappa_t, \tilde{\kappa}_t)$ plane from a combination of measurements targeting $t\bar{t}H$, tH, and tHW events using the data collected during Run 2 with the CMS detector. The figure is taken from Ref. [28], which performed a combination of analyses in the $H \to b\bar{b}$, $H \to \gamma\gamma$, $H \to ZZ^*$, and $H \to WW^*$ / $\tau^+\tau^-$ decay channels under the assumption of SM Higgs boson branching ratios.

3. The CMS experiment at the CERN Large Hadron Collider

This chapter introduces the LHC at CERN and the CMS detector. The emphasis is on the detector systems, reconstruction methods, and calibration algorithms most relevant for the analysis presented in this thesis.

3.1. The Large Hadron Collider

The LHC is located in a circular underground tunnel of approximately 27 km circumference near Geneva and is the highest-energy particle collider constructed to date [6]. It can accelerate protons or heavy ions; this thesis focuses only on proton operation. As a circular collider, it allows the beams to pass through the same accelerating structures over many turns and to collide repeatedly at the interaction points. The acceleration is provided by radio-frequency cavities, whose oscillating electric fields increase the beam energy. The main high-energy limitation of circular accelerators is synchrotron radiation emitted by charged particles in the bending magnets. For ultra-relativistic particles of the same charge, the energy loss per turn scales approximately as $E^4/(m^4\rho)$, where E is the beam energy, m is the particle mass, and ρ is the bending radius. The inverse fourth-power dependence on the particle mass makes this loss much smaller for protons than for electrons, which is one reason why a circular TeV-scale proton collider is feasible. The circular storage of many proton bunches then provides repeated collision opportunities and enables the large integrated luminosities required for precision measurements and searches for rare phenomena. The beam acceleration starts in Linac4, a linear accelerator, and proceeds through the Proton Synchrotron Booster, the Proton Synchrotron, and the Super Proton Synchrotron, which are circular synchrotrons, before the beams are transferred into the LHC ring for the final acceleration stage. The beam energy is increased from 160 MeV after Linac4 to 2 GeV in the Booster, to 26 GeV in the Proton Synchrotron, and to 450 GeV in the Super Proton Synchrotron, from where the beams are injected into the LHC. In the LHC ring, superconducting dipole magnets bend the beams along the circular trajectory and quadrupole magnets focus them transversely. The two counter-circulating beams are stored as bunches and brought to collision at four main interaction points hosting the ALICE, ATLAS, CMS, and LHCb experiments [4, 5, 113, 114].

In Run 1, proton–proton collisions were delivered at centre-of-mass energies of 7 TeV in 2010 and 2011 and 8 TeV in 2012. After the first long shutdown, Run 2 operated from 2015 to 2018 at 13 TeV, followed by Long Shutdown 2. Run 3 began in July 2022 and is delivering proton–proton collisions at a centre-of-mass energy of 13.6 TeV. The integrated luminosity delivered to the CMS experiment increased from about 29 fb^{-1} in Run 1 to 164 fb^{-1} in Run 2, while more than 350 fb^{-1} have been delivered in Run 3 by the time of writing this thesis [115]. In standard Run-3 operation, the bunch spacing is 25 ns, corresponding to a nominal bunch-crossing rate of 40 MHz. In high-intensity Run-3 operation, the LHC circulated about 2400 bunches per beam, which was a

key ingredient in reaching peak instantaneous luminosities above $2.5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ at the CMS interaction point [115, 116].

3.2. The CMS detector

The CMS detector, whose name originates from "Compact Muon Solenoid", is a general-purpose detector with almost full solid-angle coverage around the interaction point [5, 117]. It is designed to address a broad physics programme at the LHC, ranging from precision measurements of SM processes and Higgs boson properties to flavour physics, heavy-ion studies, and various searches for new physics effects. Its detector concept is therefore optimised for the precise and efficient reconstruction of electrons, photons, muons, jets, and missing transverse momentum over a wide kinematic range. It has an approximately cylindrical geometry that is azimuthally symmetric around the beam axis and can be divided into a central barrel region and two end-caps on either side. Its central element is a superconducting solenoid of 6 m inner diameter, which provides a magnetic field of 3.8 T oriented along the beam axis and is surrounded by a steel return yoke. The solenoid coil is wound from superconducting niobium-titanium material and is operated at cryogenic temperature. The strong magnetic field bends charged-particle trajectories and thereby enables precise momentum measurements from particle tracks. Around the beam pipe, the subdetectors are arranged in a layered structure designed to reconstruct the trajectories and energies of the particles produced in the collisions. The innermost part is the silicon tracking detector, which is surrounded by the electromagnetic calorimeter (ECAL) and hadron calorimeter (HCAL); the large bore of the solenoid allows both calorimeters to be installed inside the magnetic field volume. Outside the solenoid, the steel return yoke provides the magnetic flux return and houses the muon spectrometer as the outermost subsystem. The overall structure of the CMS detector is shown in Fig. 3.1.
In order to describe the collisions inside the CMS detector and its geometry, a right-handed coordinate system with origin at the nominal interaction point is used. The x-axis points towards the centre of the LHC ring, the y-axis points vertically upwards, and the z-axis points along the anticlockwise beam direction. The polar angle θ is defined with respect to the positive z-axis, and the azimuthal angle ϕ is defined in the x-y plane relative to the x-axis. The transverse momentum p_T denotes the magnitude of the momentum component perpendicular to the beam axis and is invariant under Lorentz boosts along the beam direction. The rapidity is defined as $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$, where p_z is the momentum component along the beam axis; under Lorentz boosts along the beam direction, all rapidities are shifted by a common constant, such that rapidity differences are invariant and are therefore frequently used to describe collider kinematics. For detector geometry, the pseudorapidity $\eta = -\ln\tan(\theta/2)$ is commonly used instead, since it depends only on the polar angle, and for highly relativistic particles one has approximately $y \approx \eta$, so that pseudorapidity differences inherit this boost invariance to a good approximation. The angular separation between two reconstructed objects is typically expressed as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

3.2.1. Tracking detector

The innermost part of the CMS detector consists of an all-silicon tracking detector, which is designed to reconstruct the trajectories of charged particles produced in the collisions. It covers the pseudorapidity range $|\eta| < 2.5$ and is immersed in the 3.8 T magnetic field of the solenoid, which allows for precise momentum measurements through the curvature of charged-particle trajectories [118].

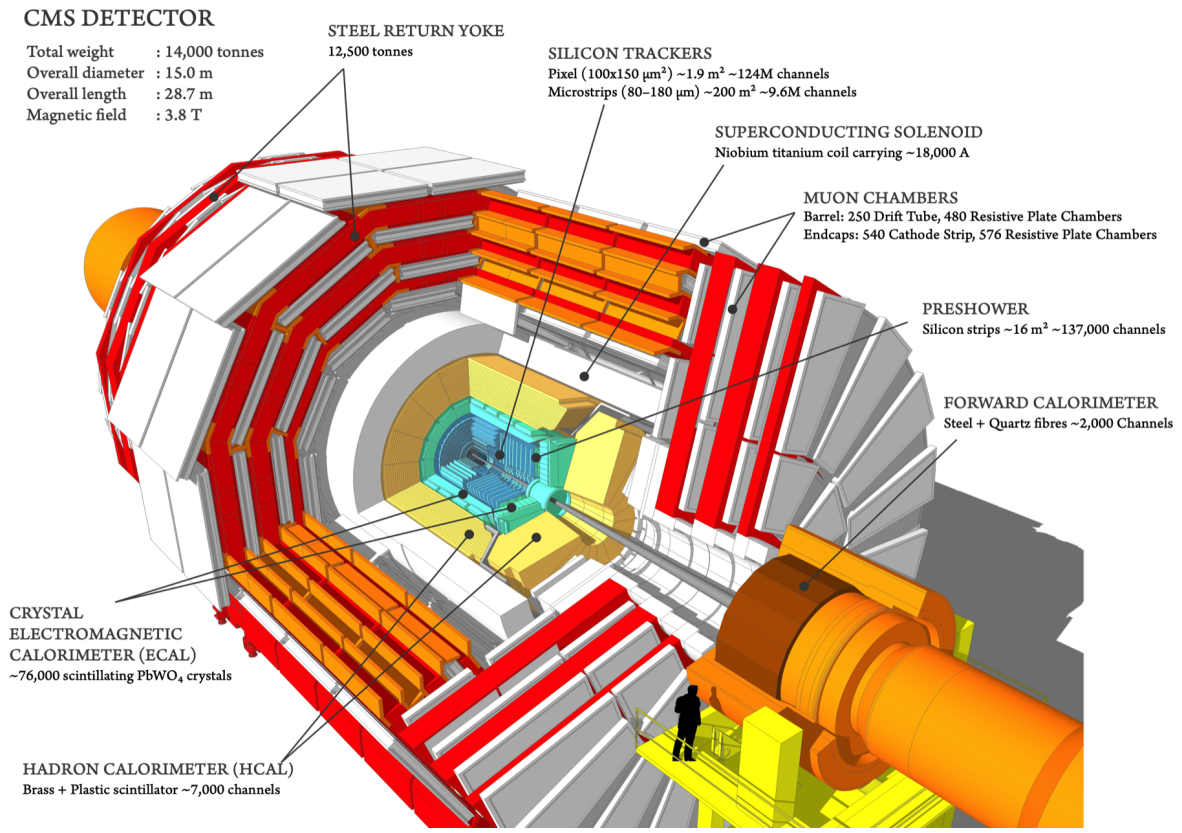


Figure 3.1: Schematic cut-away view of the CMS detector, showing the main subdetectors and the overall structure. The figure is taken from Ref. [117].

The innermost part of the tracker is the silicon pixel detector, which provides the highest spatial granularity close to the interaction point and is therefore crucial for precise impact-parameter measurements, primary-vertex reconstruction, and track seeding. The detector used in this thesis has the geometry of the Phase-1 pixel detector [119], consisting of four barrel layers at radii of about 2.9, 6.8, 10.9, and 16.0 cm, and three endcap disks on each side located at about $z = \pm 29.1$, ± 39.6, and ± 51.6 cm, providing four-hit coverage within $|\eta| < 2.5$. It contains about 124 million pixels of size $100 \times 150\,\mu\text{m}^2$. During Long Shutdown 2, the barrel pixel (BPix) detector was refurbished and its innermost layer was fully replaced before the start of Run 3 [120]. Because the pixel sensors are segmented into small cells in both directions of the sensor plane, a single pixel hit directly provides two local position coordinates of the traversing particle.

The outer part of the tracker consists of the silicon strip detector, which extends to a radius of about 110 cm from the interaction point. It is composed of ten coaxial barrel layers and, in the endcaps, three inner and nine outer disks, with a total of about 9.3 million strips [118]. The inner barrel and inner disks operate in the region $r < 55$ cm and $|z| < 118$ cm and achieve hit resolutions of about 13–38 μm in the r–ϕ plane, while the outer barrel and larger endcap disks reach about 18–47 μm and extend the coverage to approximately $|z| = 282$ cm [118]. In contrast to the pixel detector, a single strip sensor measures only one precise local coordinate, and the second coordinate is obtained in stereo arrangements from modules mounted with a small relative angle.

The material budget of the tracking detector with mechanical support and cooling system is in the range of approximately 0.4 X_0 in the central barrel, increases to approximately 2.0 X_0 at

$|\eta| \approx 1.5$ and decreases to $1.3\,X_0$ at $|\eta| \approx 2.5$ [121]. Here, the radiation length X_0 describes the characteristic decay length for the energy loss of an electron by bremsstrahlung, while the corresponding mean free path for high-energy photon conversion via e^+e^- pair production is about $9/7\,X_0$. The resulting conversion probability is therefore substantial, in particular in the larger-$|\eta|$ regions with increased tracker material budget, which is an important aspect for the photon signature and its reconstruction.

3.2.2. Electromagnetic calorimeter

The ECAL is made of lead tungstate ($PbWO_4$) crystals, a dense scintillating material that is well suited for the precise energy measurement of photons and electrons [122]. Its short radiation length of $X_0 = 0.89\,\text{cm}$ and small Molière radius of $R_M = 2.2\,\text{cm}$ lead to very compact electromagnetic showers, with about 90% of the shower energy contained within a radius of one Molière radius around the shower axis. Furthermore, it provides a fast light emission of around 80% of the produced scintillator light within the 25 ns bunch spacing.
The ECAL barrel (EB) consists of 61 200 trapezoidal crystals. The crystals have a front face of about $2.2 \times 2.2\,\text{cm}^2$ and a length of 23.0 cm, corresponding to $25.8\,X_0$. Each crystal covers a region of approximately $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$, and the barrel covers the pseudorapidity range $|\eta| < 1.48$. It is divided into two half-barrels, each composed of 18 supermodules covering 20° in ϕ and containing 1700 crystals.
Each ECAL endcap (EE) consists of 7324 trapezoidal crystals and covers the pseudorapidity range $1.48 < |\eta| < 3.0$. The crystals have a front face of about $2.86 \times 2.86\,\text{cm}^2$ and a length of 22.0 cm, corresponding to $24.7\,X_0$. They are grouped into 5×5 crystal units, often referred to as supercrystals, and each endcap is mechanically built from two half-disks.
The crystals are arranged in a quasi-projective geometry, such that their axes point close to, but not exactly towards, the nominal interaction point. This reduces cracks aligned with particle trajectories and improves the containment and position measurement of electromagnetic showers. In front of the endcaps, a preshower detector is placed. It consists of two layers of silicon sensors interleaved with lead absorbers of $3\,X_0$ depth and provides additional information for the identification (ID) of photons and electrons in the endcap region.
The scintillation light is detected by avalanche photodiodes in the barrel and by vacuum phototriodes in the endcaps, which convert it into electrical signals. These signals are subsequently amplified, shaped with a multi-gain preamplifier, and digitised. During detector operation, irradiation reduces the transparency of the $PbWO_4$ crystals and therefore the amount of scintillation light reaching the photodetectors. These time-dependent response changes are monitored with the dedicated ECAL laser system and corrected at the channel level in the reconstruction [122, 123]. In addition, a separate intercalibration step is performed to equalise the relative response of the individual ECAL crystals, such that the same deposited energy is reconstructed consistently across the calorimeter [124]. These intercalibration coefficients are derived by combining several complementary data-driven methods, including symmetry-based approaches, the reconstruction of $\pi^0/\eta \to \gamma\gamma$ decays, and electrons from W and Z decays through E/p and invariant-mass constraints [125].

3.2.3. Hadron calorimeter

The HCAL is designed to measure the energy of hadrons and, together with the ECAL, to provide a hermetic measurement of the visible energy flow in the event. Its central parts are sam-

pling calorimeters, in which passive absorber layers are interleaved with active material. In the barrel and endcaps, denoted HCAL barrel (HB) and HCAL endcap (HE), the absorber is primarily brass, while the inner and outer layers of HB are constructed from steel. The active material in HB and HE consists of plastic scintillator tiles. The scintillation light is collected with embedded wavelength-shifting fibres, which shift the blue scintillation light to longer wavelengths and guide it efficiently to the photodetectors.

Hadronic showers are initiated by strong interactions of hadrons with the absorber nuclei and develop over the nuclear interaction length, which makes them much more extended than electromagnetic showers. For the absorber materials used in HB and HE, the characteristic interaction lengths are about $\lambda_I \approx 15\,\text{cm}$ for brass and $\lambda_I \approx 17\,\text{cm}$ for steel. The HCAL is therefore needed to contain and measure the hadronic component of the event and is crucial for the reconstruction of jets and for the measurement of missing transverse momentum.

The barrel region covers the pseudorapidity range $|\eta| < 1.39$. Its cells have a granularity of about 0.087 in both $\Delta\eta$ and $\Delta\phi$, such that one HCAL cell maps approximately onto a 5×5 crystal region in the ECAL and forms, together with the layers behind it, an outward-pointing calorimeter tower. The endcaps cover the range $1.31 < |\eta| < 3.0$, with up to 18 layers of scintillators and a coarser granularity of up to about 0.174 in $\Delta\eta$ and $\Delta\phi$ at high pseudorapidities.

The hadron forward calorimeter (HF) is located in the forward region outside the endcaps and covers the range $3.0 < |\eta| < 5.2$. In contrast to HB and HE, the HF uses quartz fibres embedded in a steel absorber and detects Cherenkov light produced by relativistic charged shower particles. This design provides the radiation hardness required in the very forward region, and the use of long and short fibres allows a partial discrimination between electromagnetic and hadronic shower components. Because the barrel ECAL and HCAL inside the solenoid do not fully contain all hadronic showers, an additional outer hadron calorimeter (HO) is installed outside the magnet coil. There, plastic scintillator layers sample the energy leaking into the steel return yoke and thereby improve the total hadronic energy measurement.

3.2.4. Muon detectors

Muons are measured in the outermost CMS subdetector system, which is embedded in the steel return yoke of the superconducting solenoid and covers the pseudorapidity range $|\eta| < 2.4$. Its large spatial extent provides a long lever arm for the momentum measurement of muons. In the steel return yoke, the magnetic flux returns outside the solenoid, so the field direction is opposite to that inside the solenoid and the bending direction of charged-particle trajectories is reversed.

In the barrel region, the muon system is based on drift tubes (DTs), while in the endcaps cathode strip chambers (CSCs) are used, where the particle rates are higher and the magnetic field is less uniform. Resistive plate chambers (RPCs) complement both regions with fast timing information for triggering and bunch-crossing identification. In the endcap region, gas electron multiplier (GEM) detectors cover the pseudorapidity range $1.55 < |\eta| < 2.18$. They were installed during Long Shutdown 2 and are operated in Run 3 to provide additional muon hit information in the high-rate environment, with the aim of improving the redundancy and trigger performance of the forward muon system.

The DT system is arranged in four barrel stations radially, which are segmented into five wheels along the beam direction and 12 sectors in azimuth, within the region $|\eta| < 1.2$. In a DT chamber, the muon position is reconstructed from the drift time of ionisation electrons towards the anode wire. Chambers in the first three stations contain three superlayers, two of which measure

the bending-plane (r–ϕ) coordinate and one the orthogonal coordinate, while the fourth station contains two superlayers.
The CSC chambers consist of arrays of positively charged anode wires and segmented cathode strips arranged perpendicular to each other and cover the region $0.9 < |\eta| < 2.4$. When a muon traverses the gas volume and ionises it, the liberated electrons drift towards the anode wires and produce an avalanche close to them, while the induced charge is measured on the cathode strips. Together, the wires and strips provide two spatial coordinates of the hit, with the cathode strips yielding the more precise measurement in the bending direction. The fast signal on the anode wires also provides precise timing information, which is important for associating muon hits with the correct bunch crossing in the high-rate endcap region. The CSC system is arranged in four endcap stations and each chamber consists of six gas layers, whose combined hits are used to reconstruct local track segments.

3.2.5. Trigger and data acquisition

The bunch-crossing rate of the LHC is far too high for all events to be recorded permanently. The CMS experiment therefore employs a two-level trigger system [117, 126]. The Level-1 trigger is implemented in custom hardware and uses information from the calorimeters and the muon system to reduce the event rate from 40 MHz to a typical output rate of about 110 kHz in Run 3. The remaining events are processed by the software-based High-Level Trigger, which performs a more detailed event reconstruction and selects a final sample at a rate of a few kHz for storage, complemented by additional parking and scouting streams with reduced or delayed event processing [117].

3.3. Reconstruction, identification, and calibration of physics objects

The subdetectors described above are used to reconstruct the trajectories and energies of the particles produced in the collisions, which are then combined to form physics objects such as photons, electrons, muons, jets, and missing transverse momentum. While multiple established algorithms exist for the reconstruction, ID, and calibration of several objects, this section focuses on the aspects and versions that are most relevant for this thesis. The reconstruction of physics objects is primarily based on the particle flow (PF) algorithm [127] integrated in the software framework CMSSW [128], which aims to optimally combine the measurements from the individual subdetectors in order to provide a complete description of the event in terms of reconstructed particle candidates.

3.3.1. Tracks and primary vertices

Charged-particle tracks are reconstructed from hits in the silicon pixel and strip detectors with an iterative tracking procedure, which fits their trajectories in the magnetic field and provides the momentum, charge, and impact-parameter information used throughout the CMS object reconstruction [118]. Interaction vertices are reconstructed by clustering compatible tracks and fitting them to a common point in space. The primary vertex is then chosen from the reconstructed vertices according to the transverse momenta of the associated tracks and provides the reference point for observables such as impact parameters and pseudorapidity.

In this thesis, the reconstructed vertex with the largest $\sum p_T^2$ [129] is selected as primary vertex, whereas in previous measurements, a dedicated Boosted Decision Tree (BDT) classifier was trained to increase the fraction of correct vertex assignment for ggH events [130]. This can be useful, because the invariant mass determination requires precise measurement of the photon energies as well as the opening angle of the diphoton system. This angle depends on the supercluster positions as well as the primary vertex location. It was found that for $H \to \gamma\gamma$ events with a primary vertex assignment that differs from the true one by less than $\Delta z = 1\,\text{cm}$, the angular resolution contribution to the mass resolution is negligible. Correct primary vertex assignment is particularly useful to increase the diphoton mass resolution of ggH events, which tend to be produced with low jet multiplicity, whereas the photons alone typically do not leave enough tracker signature to identify the correct primary vertex. For the present analysis, this simplified primary-vertex choice is not expected to lead to a significant degradation of the invariant mass resolution. It is found to yield the correct primary vertex within $\Delta z < 1\,\text{cm}$ for around 99% and 95% of the selected simulated $t\bar{t}H$ and tH events, respectively. This treatment follows the preliminary convention for analysis of Run-3 data in the $H \to \gamma\gamma$ decay channel at the CMS experiment [45], where the use of a reduced data format does not allow the evaluation of the vertex-assignment BDT in its original form.

3.3.2. Photons and electrons

The reconstruction of photons and electrons in the CMS detector is closely related because both produce electromagnetic showers in the ECAL. Starting from a seed crystal with energy above a threshold and exceeding the energy measured in nearby crystals, a clustering algorithm adds neighbouring crystals with significant energy deposits. The resulting clusters are merged into superclusters, which are designed to capture most of the energy of the original particle, including contributions from bremsstrahlung photons radiated by electrons and from photons that convert into electron–positron pairs in the tracker material [131]. Each supercluster candidate is then treated under both the electron and the photon hypothesis, while the subsequent object reconstruction and ID differ accordingly.

Reconstruction and electron-photon separation The distinction between photons and electrons is based on the presence of a compatible track in the silicon tracker [131]. For the photon hypothesis, the key criterion is the conversion-safe electron veto, which rejects superclusters matched to tracks with hits in the innermost pixel layers and thus consistent with electrons originating from the interaction point. The reconstructed shower shape and the activity around the candidate are then used both for particle ID and for deriving subsequent corrections. The variables most relevant for this thesis are the following:

- $R9$ is the ratio of the energy contained in the 3×3 crystal matrix centred on the seed crystal to the total uncorrected supercluster energy. It measures the compactness of the shower and helps distinguish unconverted or late-converting photons, which typically have larger $R9$, from earlier conversions.

- $\sigma_{i\eta i\eta}$ is the log-weighted second moment of the shower width along the pseudorapidity direction, which is not broadened by the bending of charged particles in the solenoidal magnetic field, and helps distinguish showers originating from a single electromagnetic particle from broader or overlapping topologies, for example, from boosted $\pi^0 \to \gamma\gamma$ decays.

- H/E is the ratio of the HCAL energy associated with the candidate to the corresponding ECAL energy and is therefore sensitive to hadronic contamination.
- The most important isolation observables used in this thesis are the charged-hadron isolation I_{trk}, defined from the scalar sum of track transverse momenta in a hollow cone around the photon candidate, and the photon isolation I_{γ}^{ρ}, defined from the photon component of the PF isolation within a cone around the candidate. A correction based on the pileup-sensitive event-level energy density ρ is included in I_{γ}^{ρ}. The cone size is $\Delta R = 0.3$ for both variables.
- The presence of a pixel seed provides additional discrimination against electrons that may be misidentified as photons.

Energy determination The raw energy summed in the clustered crystals after crystal-wise calibrations is a suboptimal estimator of the true particle energy. It is affected by detector effects such as longitudinal and lateral leakage, intermodular gaps, unresponsive crystals, noise thresholds, and pileup. Therefore, a multivariate regression based on BDTs implemented in the TMVA [132] package is used to predict an improved energy estimate from the shower properties [133]. The regression is trained separately for photons and electrons on large simulation samples with a uniform energy distribution and targets the ratio of the raw supercluster energy to the true generated energy. Rather than predicting only a single corrected value, it yields a parametric response model whose main outputs include the most probable value, used to correct the energy, and the width of the predicted response distribution. In $H \to \gamma\gamma$ analyses, the latter can be used to estimate the expected diphoton-mass resolution event by event and thereby improve the analysis sensitivity, as done in the first Run-3 measurement of $H \to \gamma\gamma$ cross sections [45].
For electrons, the track momentum provides complementary information and is combined with the ECAL supercluster energy in the so-called E–p combination. This improves the energy resolution, in particular at low energies where the tracker momentum measurement can be more precise than the calorimetric measurement alone.
For unconverted or late-converting photons whose supercluster lies in the EB and whose energy is in the tens-of-GeV range, an energy resolution of about 1% is achieved. For other photons, the resolution is about 1.3% up to $|\eta| = 1$ and about 2.5% at $|\eta| = 1.44$ in the EB. In the EE, about 2.5% is achieved for unconverted or late-converting photons and between 3% and 4% otherwise [131]. For electrons, the combination of the ECAL energy measurement with the reconstructed track momentum yields an energy resolution of about 1.7–4.5% for electrons from Z-boson decays, depending on the pseudorapidity region and the bremsstrahlung pattern along the trajectory [133].

Identification An important task in the reconstruction of photons is the rejection of background mimicking the photon signature. At the particle level, prompt photons comprise photons produced directly in the hard interaction or in prompt resonance decays such as $H \to \gamma\gamma$, as well as fragmentation photons radiated from high-energy partons. Photons not produced in this way are referred to as non-prompt photons. Important examples are photons from neutral meson decays inside jets, in particular $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$. Since these mesons are Lorentz-boosted at the LHC, the two photons are typically highly collimated and deposit their energy in neighbouring, or even the same, ECAL crystals, resulting in a photon-like electromagnetic

shower. At the detector level, the term photon candidate is used for any object reconstructed by the photon algorithms, irrespective of the particle that caused the electromagnetic shower. In this thesis, the notation $j \to \gamma$ is used for the corresponding jet-to-photon misidentification, in which a hadronic object is reconstructed and identified as a photon candidate; analogously, $e \to \gamma$ denotes an electron reconstructed and identified as a photon candidate. For $H \to \gamma\gamma$ analyses, signal events contain two isolated prompt photons from the Higgs boson decay, whereas reducible backgrounds arise when one or both photon candidates originate from such $j \to \gamma$ or $e \to \gamma$ misidentifications. Photon candidates of hadronic origin often produce broader showers and larger nearby hadronic activity than isolated prompt photons, which makes variables such as $R9$, $\sigma_{i\eta i\eta}$, H/E, and the isolation observables powerful discriminants. The final ID is based on BDT classifiers implemented in the XGBoost package [134], following the general approach described in Ref. [133]. Their inputs include the shower-shape and isolation variables discussed above together with related observables characterising the local detector activity. The BDTs are conditioned on the candidate p_{T} and η as well as ρ. In combination with a reweighting of the signal and background training samples to common kinematic spectra, this reduces undesired dependencies of the discriminator on the kinematic properties of the candidate and on pileup conditions.

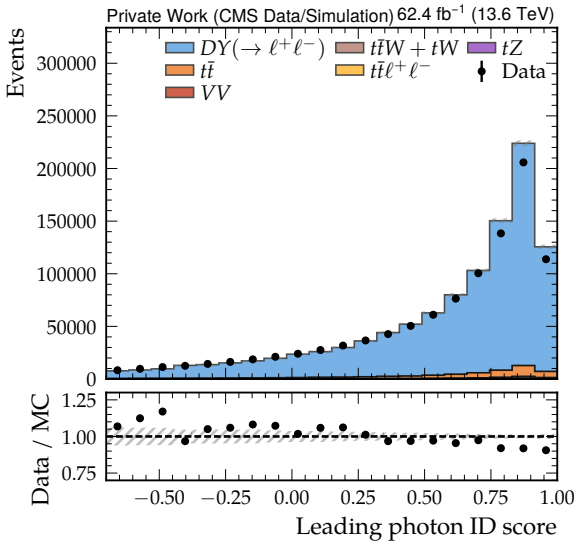
Corrections Significant discrepancies between simulation and data remain in the reconstructed shower shapes, isolation observables, and energy response and are an important source of systematic uncertainty in $H \to \gamma\gamma$ analyses [135]. They arise in particular from imperfect modelling of radiation damage in the ECAL crystals, uncertainties in the material budget upstream of the calorimeter, and changes in detector noise thresholds. To reduce the resulting biases in ID, energy determination, and selection efficiencies, dedicated corrections are applied.

For photons and electrons, most of the efficiency-related corrections are derived from minimally biased electron samples selected using the tag-and-probe technique [136]. In this method, $Z \to e^+e^-$ events are selected with high purity by requiring one electron (the tag) to satisfy stringent ID criteria, while the second electron is the minimally biased probe. The tag-and-probe method provides large, high-purity samples of leptons. This makes it possible to compare data and simulation for the observables relevant to the reconstruction and ID with small residual selection bias. For photons, such electrons are reconstructed with the photon algorithms, since no data sample of isolated prompt photons with comparable purity and event yield is available. Because the shower properties differ between converted and unconverted photons, several of these corrections are derived as functions of $R9$.

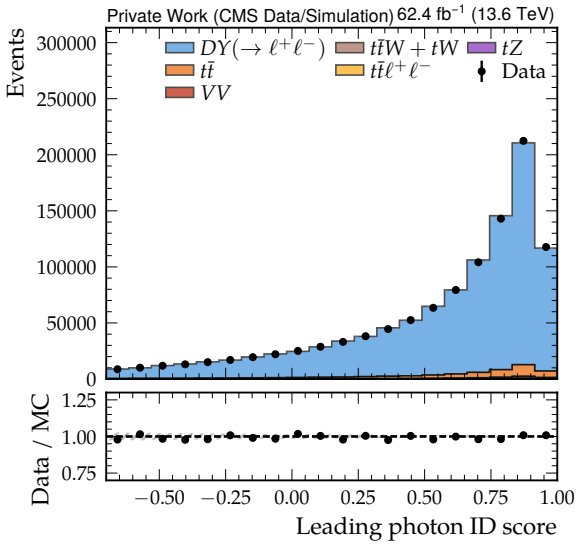
For the photon ID score, the residual mismodelling is reduced with the correction procedure of Ref. [137], which is based on an autoregressive normalising flow [138, 139]. The model is trained on probe electrons reconstructed as photons and learns an invertible transformation from the original multivariate input distribution to a Gaussian base distribution common to data and simulation. This is realised by conditioning the model on a boolean label that indicates whether an event is from data or simulation. Simulated photon candidates are mapped to this base space and then transformed back with the data label, thereby morphing the full set of photon-ID input variables towards the data distribution. The corrected variables are then passed again through the photon-ID BDT to obtain a corrected photon ID score. This procedure is found to align the relevant distributions to the percent level for electrons reconstructed as photons.

The energy response of photons and electrons also requires dedicated corrections after the regression step. The energy resolution in simulation is generally better than in data, in particular

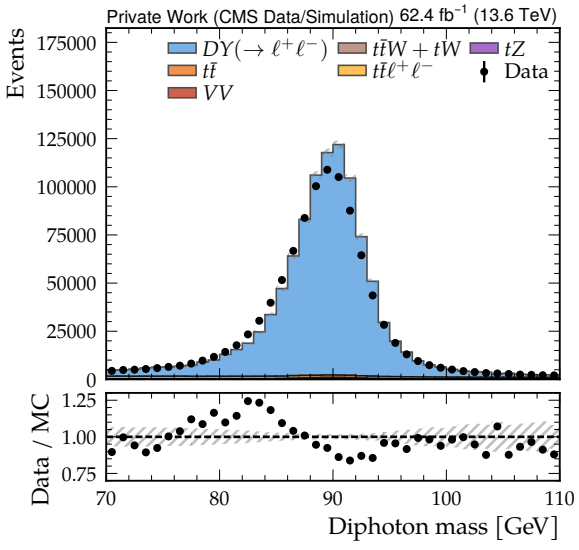
because effects such as radiation damage to the crystals are not modelled perfectly. This is corrected by applying an additional multiplicative smearing to simulated photon energies, together with residual energy-scale corrections applied to data. The scale and resolution corrections are derived from dedicated $Z \to e^+e^-$ calibration samples in which the electrons are reconstructed either as electrons or as photons, based on the method from Ref. [140]. They are obtained from an analytical likelihood fit to the reconstructed mass spectra, avoiding the random-number-based convolution techniques used in previous approaches. This analytical treatment of the smearing leads to a fully differentiable likelihood, which improves the numerical stability and computational efficiency of the minimisation while allowing the simultaneous extraction of correction parameters in multiple kinematic categories. The scale correction is binned in fine periods of data-taking time in order to account for the progressive change in the ECAL crystal transparency and response.
The framework of Ref. [140] also includes an additional layer of photon-energy scale determination with $Z \to \mu^+\mu^-\gamma$ events. The full correction chain has been applied in the recent Higgs boson mass measurement in the diphoton decay channel [43]. In Ref. [43], residual discrepancies between data and simulation are still observed in the $\mu^+\mu^-\gamma$ three-body mass distribution after the electron-based calibration stages, and are reduced by the application of the final photon-energy scale correction derived from the $Z \to \mu^+\mu^-\gamma$ photons. These photon-based corrections were, however, not yet available when the data used in this thesis were processed, and are thus not applied.
The effects of the photon ID and energy corrections are illustrated in Fig. 3.2 with a $Z \to e^+e^-$ sample in which electrons are reconstructed with the photon algorithms; the corresponding validation region is kinematically similar to the signal region of this analysis and defined later in Sec. 4.4. The uncorrected ID score distribution shows that simulated electrons reconstructed as photons tend to appear as more signal-like than the corresponding data objects. At the same time, the invariant mass spectrum in data is significantly shifted towards lower values by approximately 1 GeV, while the resolution mismodelling is visible in a slightly sharper peak in simulation. The correction chain improves the agreement between data and simulation both for the photon ID score and for the reconstructed diphoton-mass distribution, such that the distributions agree at the percent level after including these corrections.
Furthermore, efficiency scale factors (SFs) are applied for the trigger and photon preselection, and for the electron ID requirement used in the analysis. Here, it is sufficient to note that the corresponding object-level efficiencies are corrected in simulation, while the explicit working-point definitions used in the analysis are given in Chap. 4. A further correction specific to electrons accounts for the fact that the matching between a reconstructed supercluster and its associated track is not fully efficient. This effect is corrected with a dedicated electron-reconstruction SF. For photons, the corresponding reconstruction efficiency is reasonably close to 100% in both data and simulation [133]. While the trigger and electron-related SFs are determined from tag-and-probe electrons, the efficiency of the electron-veto requirement is determined from a sample of $Z \to \mu^+\mu^-\gamma$ events, which provide genuine prompt photons.



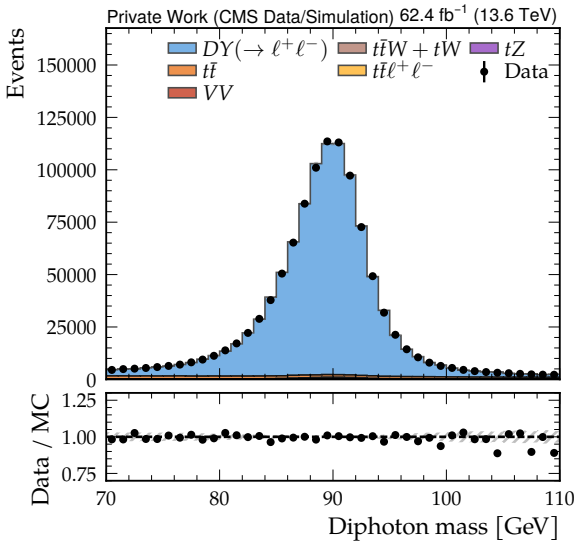
(a) Uncorrected photon ID score



(b) Corrected photon ID score



(c) Uncorrected diphoton mass



(d) Corrected diphoton mass

Figure 3.2: Illustration of the effect of the photon corrections in the $Z \to e^+e^-$ validation region defined in Sec. 4.4, where electrons are reconstructed with the photon algorithms. The figures on the left show the uncorrected photon ID score distribution and the reconstructed diphoton invariant-mass distribution. The figures on the right show the corresponding distributions after including the normalising-flow-based correction to the photon ID input variables and the scale and resolution corrections to the photon energy, respectively.

3.3.3. Jets and b-tagging

Jets used in this thesis are reconstructed from PF candidates with the anti-k_t algorithm [141, 142] and a distance parameter of $R = 0.4$. In Run 3, the default jet collection is based on the Pileup Per Particle Identification (PUPPI) algorithm, which suppresses the contribution from additional proton–proton interactions by assigning weights to PF candidates according to their compatibility with pileup activity [143–145]. The resulting PUPPI jets provide a stable jet response already before the final in-situ calibration, in particular in the high-pileup conditions relevant for the present analysis [144]. The jet energy scale and resolution are calibrated in a sequential procedure. The first step is a simulation-based detector-response correction in bins of p_{T} and η, which aligns the average reconstructed jet energy with the corresponding particle-level energy in simulation. Residual corrections are derived in data with in-situ techniques based on dijet, Z+jet, γ+jet, and multijet events, while the jet-energy resolution (JER) in simulation is adjusted with additional smearing [144]. The achieved jet-energy resolution is of order 15–20% at 30 GeV and improves to about 10% at 100 GeV [146].

Identifying jets initiated by a b quark is a crucial aspect of this analysis in order to identify top quark decays; this task is referred to as b-tagging. Heavy-flavour jets initiated by b or c quarks have characteristic experimental signatures that can be exploited by dedicated tagging algorithms [147]. For this thesis, the relevant case is b-tagging. The b hadrons have a comparatively long lifetime of about 1.5 ps. With typical Lorentz boosts they travel a few millimetres to about a centimetre before decaying, such that their decay products often form a displaced secondary vertex, as illustrated in Fig. 3.3. The displacement of individual decay tracks is quantified by the impact parameter, defined as the distance between the reconstructed track trajectory and the primary interaction vertex at the point of closest approach. Both the transverse impact parameter in the plane perpendicular to the beam axis and the full three-dimensional impact parameter are used. Tracks from displaced b-hadron decays tend to have significantly larger impact parameters than tracks originating directly from the primary vertex, which makes this observable a powerful input for b-tagging. Additional discrimination arises from the comparatively large b-hadron mass and the correspondingly harder fragmentation, such that the b hadron carries a relatively large fraction of the jet momentum and its decay products tend to have sizeable transverse momenta with respect to the jet axis. Finally, semi-leptonic decays provide complementary information: around 20% of b hadrons decay with an electron or muon in the final state, so the presence of a soft lepton inside the jet is another characteristic heavy-flavour signature.

Since each of these signatures is only partially discriminating on its own, b-tagging is a challenging task and benefits strongly from modern multivariate methods. Powerful deep-learning approaches are therefore used to exploit the available information effectively. In order to train these models, a definition of jet flavour in simulated events is required. In simulation, the flavour of a jet is defined by clustering not only the stable generated final-state particles, which determine the physical jet kinematics, but also generated b- and c hadrons that do not have a daughter hadron of the same flavour [147]. This avoids multiple counting of a single heavy-flavour decay chain, for example when an excited b hadron decays to another b hadron before its weak decay. To ensure that these hadrons do not modify the jet four-momentum, they are included in the clustering only with infinitesimal momenta, such that only their directional information enters the jet algorithm. This procedure is known as ghost association [148]. Jets containing at least one ghost-associated b hadron are labelled as b jets, while jets containing at least one ghost-associated c hadron and no b hadron are labelled as c jets; all remaining jets are classified as light-flavour jets [147].

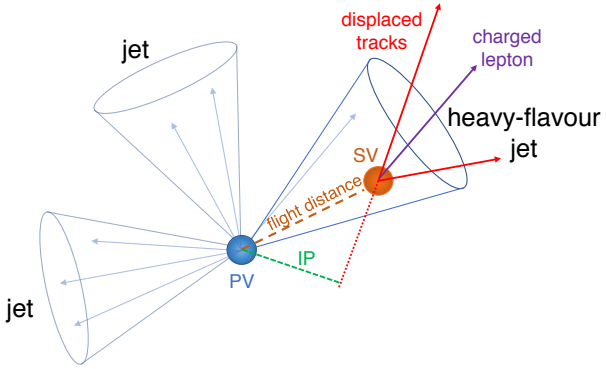


Figure 3.3: Schematic illustration of a jet containing a long-lived hadron decaying at a secondary vertex. The displaced decay topology and the impact parameter of a decay track with respect to the primary vertex are indicated schematically. The figure is taken from Ref. [147].

In this thesis, the *Robust Particle Transformer* algorithm [149] is used as the b-tagger. It is based on the Particle Transformer architecture [150], which uses the attention mechanism [151] to learn from the set of reconstructed jet constituents. Compared with earlier architectures, it includes pairwise interaction information between constituents and the primary vertex directly in the model. The robustness referred to in its name is achieved through an adversarial training procedure [152], in which the input features are distorted during training in the direction that maximally increases the loss, such that the classifier becomes less sensitive to genuine mismodelling effects between data and simulation. For jets from top quark decays with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and a b-jet efficiency of 60%, the algorithm achieves a light-jet rejection of about 2500 [149]. The tagger can be used either at fixed working points, defined for a chosen signal efficiency, or through its continuous output score, which retains the full discriminating information and is the choice made in this thesis.

Although the adversarial training procedure reduces the sensitivity to simulation mismodelling, residual differences between data and simulation remain in the output score distributions. In addition to SFs defined at fixed tagging working points, a calibration of the full b-tagging discriminant shape is therefore provided, which is also used in this thesis. In this approach, per-jet correction factors are derived as functions of the tagger output such that the simulated discriminator distributions agree with those observed in data for the relevant jet flavours. For b-tagging, the calibration is obtained with an iterative-fit procedure using flavour-enriched control samples, in particular dileptonic $t\bar{t}$ events for b jets and Z+jets events for light-flavour jets [147, 149]. This shape calibration is particularly useful when the continuous tagger output is used directly as an input to multivariate discriminants rather than only through fixed working points, although it generally comes with larger uncertainties than fixed-working-point SFs.

3.3.4. Muons

Muons used in this thesis are reconstructed as so-called global muons [153]. The reconstruction starts from a standalone-muon track built from hits in the muon system with a Kalman-filter technique [154], seeded by groups of hits in the DT or CSC systems. This standalone-muon track is then matched to a track reconstructed in the inner tracking detector and refitted with the combined information from both systems [153].

The resolution of muon p_T is as good as 1% for the p_T range 20–100 GeV in the barrel and up to 3% in the endcaps. It degrades for higher p_T, where it increases to almost 10% for 1 TeV. At transverse momenta most relevant for this thesis, the inner tracker provides the dominant contribution to the momentum resolution, while the muon system is crucial for the ID of muons through the matching of tracker tracks to hits and segments reconstructed outside the calorimeters. For muons with transverse momenta below about 200 GeV, the contribution of the muon system to the momentum measurement is small, whereas at higher transverse momentum the additional lever arm of the muon chambers significantly improves the combined momentum resolution [153]. The calibration includes corrections to the momentum scale in data and to the resolution in simulation derived from $Z \to \mu^+\mu^-$ events, with dedicated methods for different momentum regimes [155].
Furthermore, requirements on the ID and isolation of muons are defined in order to suppress non-prompt muons and cosmic muons as backgrounds. Muon ID is based on several quality criteria, including the degree of matching between the inner-track and standalone-muon trajectories, the number of hits contributing to the track reconstruction, the quality of the global fit, and the compatibility with the primary vertex. The isolation of muons is computed from the scalar sum of the transverse momenta of reconstructed particles within a cone of $\Delta R < 0.3$ around the muon direction. As an approximate correction for pileup contributions from neutral particles, half of the p_T sum of charged particles inside the isolation cone that are not originating from the primary vertex is subtracted [153].

3.3.5. Missing transverse momentum

In this thesis, the missing transverse momentum vector, $\vec{p}_T^{\,\text{miss}}$, is reconstructed with the PUPPI algorithm as the negative vector sum of the transverse momenta of the PUPPI-weighted PF candidates in the event [145, 156]. Its magnitude is referred to as missing transverse momentum, p_T^{miss}. Since the initial state has, to a very good approximation, no net transverse momentum, a nonzero $\vec{p}_T^{\,\text{miss}}$ can indicate the presence of neutrinos or other weakly interacting neutral particles that are not reconstructed in the detector. The compatibility of the reconstructed missing transverse momentum with the hypothesis of zero true missing transverse momentum is quantified using the reconstructed per-event resolution and defines the p_T^{miss} significance variable [156]. The PUPPI weighting suppresses the contribution from particles compatible with pileup interactions before the transverse momentum sum is formed and hence provides a more stable performance as a function of pileup. The resolution of $\vec{p}_T^{\,\text{miss}}$ depends strongly on the event topology, in particular on the hadronic recoil and the pileup conditions. The so-called unclustered component denotes the contribution from reconstructed PF candidates that are not associated with the explicitly corrected high-p_T physics objects and is taken into account in the $\vec{p}_T^{\,\text{miss}}$ reconstruction [156]. In this thesis, an additional correction is implemented that propagates all corrections applied to the reconstructed jet, photon, electron, and muon momenta to $\vec{p}_T^{\,\text{miss}}$, which keeps the event description consistent.

4. Analysis strategy and workflow

This chapter describes the overall strategy followed in this analysis and introduces the event-processing framework used to prepare the reduced event samples for the later analysis steps. The selection of the reconstructed events is explained for the two channels of the signal region as well as for a validation region used to investigate the modelling of specific observables.

4.1. Analysis strategy

The first key choice of this analysis is the use of the diphoton decay channel of the Higgs boson. It has several features that are strategically advantageous for the present measurement. Because the intrinsic decay width of the Higgs boson is negligible compared to the experimental photon-energy resolution, the $H \to \gamma\gamma$ signal appears as a narrow resonance in the diphoton invariant-mass ($m_{\gamma\gamma}$) spectrum on top of a continuous background. Requiring two reconstructed isolated photons at a hadron collider experiment rejects the vast majority of the hadronic background and therefore the diphoton channel has a moderate background level. Furthermore, the full and precise reconstruction of the Higgs boson kinematics from the two photons is valuable for the separation between the signal processes $t\bar{t}H$ and tH, which is a central aspect of the present measurement. If a sufficiently wide range of the $m_{\gamma\gamma}$ spectrum is selected, it allows the signal and the continuum background component to be extracted from the observed data. This reduces the dependence of the measurement on the continuum-background modelling in simulation, whereas in other channels such as $H \to b\bar{b}$ or multilepton final states, the modelling of background processes without Higgs bosons is a major source of systematic uncertainty [28, 32, 33, 157, 158]. These properties make the diphoton decay channel one of the most sensitive channels not only for the analysis of the $t\bar{t}H$ process, but also in previous searches for the rare tH production [39, 40], despite the small branching ratio.

A central goal of the present analysis is that the measured cross sections do not strongly depend on assumptions on the top-quark Yukawa coupling. As a result, the measurement can be directly reinterpreted to derive constraints on the coupling parameters. Such a strategy has received only limited attention in previous measurements, although existing analyses already indicate that it matters experimentally. While in most previous cross section measurements of $t\bar{t}H$ and searches for tH production, the dependence of the obtained results on the coupling assumption is not reported, the searches for tH production from Refs. [35, 159] report a significant impact when changing the coupling assumption to the inverted-sign hypothesis. In Ref. [159], the upper limits on tH production differ by a factor of approximately three between the inverted-sign hypothesis and the SM case. Similarly, Ref. [35] discusses a significant loss of sensitivity in the inverted-sign case because the classifier is trained on the SM signal simulation.

This observation motivates the use of event selections and discriminants that remain as stable as possible across coupling hypotheses. In this context, Ref. [48] showed that changes in the coupling modifiers $(\kappa_t, \tilde{\kappa}_t)$ towards a CP-odd configuration induce kinematic changes that can alter the event composition of the analysis categories, in particular the relative contributions of

the different signal processes. That study proposed a cut-based separation using a few carefully designed observables that remain comparatively stable across coupling hypotheses. The present analysis picks up this idea and extends it in several key aspects: first, the actual impact of coupling variations on the output distributions of multivariate classifiers for the $t\bar{t}H$–tH and signal–background classification tasks is analysed and demonstrated; second, a deep-learning-based solution is developed to achieve a less model-dependent separation between these event classes. Compared with a separation based on a small number of hand-crafted observables, this approach retains more of the available event information and can therefore provide higher separation power, leading to improved sensitivity and smaller statistical uncertainties in the final measurement. This is achieved by explicitly decorrelating the classifier outputs from the coupling structure of the signal events used in the training. Additionally, the measurement is formulated in a fiducial phase space in order to avoid a model-dependent extrapolation from the experimentally accessible to the total phase space.

A further central element of the analysis strategy is the separation into a hadronic and a leptonic channel, depending on whether a reconstructed electron or muon is present. The leptonic channel is expected to have a more favourable signal-to-background ratio, because the lepton requirement strongly suppresses the background. For $t\bar{t}H$ production, the leptonic channel is known to be particularly sensitive to the signal, because the presence of two W bosons implies that about 45% of events contain at least one prompt electron or muon in the final state, based on the decay fractions discussed in Chap. 2. The efficiency for selecting at least one lepton is therefore expected to be sizeable. For tH production, however, the situation is more asymmetric as this combinatoric enhancement effect is not present: the fraction of events with a prompt electron or muon is expected to be around one quarter. After detector acceptance, reconstruction, isolation, and ID requirements on the charged lepton, the ratio of tH events selected in the hadronic over the leptonic channel is expected to become even larger. This motivates the inclusion of the hadronic channel especially for the tH process, which was not done in searches for this process in the published $H \to \gamma\gamma$ measurements using the full Run-2 datasets [39, 40]. After requiring a b-tagged jet and exploiting the characteristic topology of tH production, including the presence of energetic forward jets, the dominant background is nevertheless expected to remain manageable with modern deep-learning techniques. Hence, the combination of the hadronic and leptonic channels is expected to provide a more sensitive experimental access to the tH signal than a purely leptonic strategy. The exploration of tH production with all-hadronic top quark decays is an experimental challenge addressed in this thesis.

The resulting workflow of this analysis is illustrated in Fig. 4.1. Starting from the data and simulation provided within the CMS Collaboration in a slimmed data format called NanoAOD, events with two photon candidates and one b-jet candidate are selected. The events are retained in a broad diphoton-mass range, so that the later signal extraction can use both the Higgs boson signal window and the sidebands in data. Up to this point, the event selection is intentionally loose to ensure a high signal efficiency. After grouping events into the two channels, a dedicated high-statistics continuum-background model is developed for classifier training and categorisation optimisation, as discussed in Chap. 6. Throughout the analysis, the much smaller tHW production process is treated as part of the $t\bar{t}H$ signal due to its experimental similarity. Then, deep neural networks (DNNs) are trained to suppress the remaining backgrounds and to enrich regions of the selected phase space in either $t\bar{t}H + tHW$ or tH events. A decorrelation technique is applied with the aim of reducing the dependence of the classifier output distributions on the coupling state of the signal events. The selected events are then grouped into categories that maximise the expected signal significances using a novel differentiable optimisation approach.

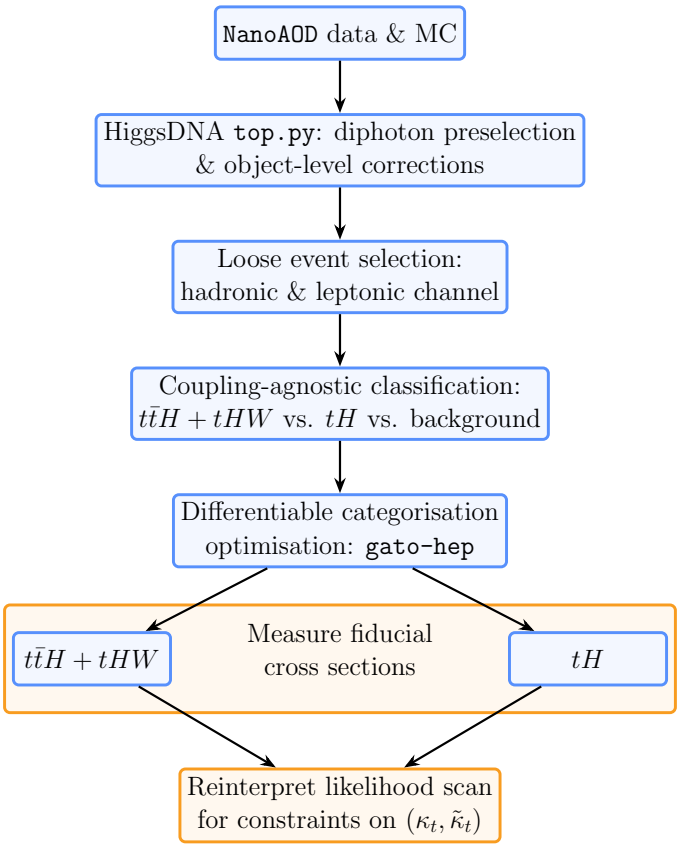


Figure 4.1: Schematic overview of the analysis workflow.

The final signal extraction is performed from the $m_{\gamma\gamma}$ spectrum in each category, following the standard $H \to \gamma\gamma$ strategy used at the CMS experiment. The resulting likelihood scan of the fiducial $t\bar{t}H + tHW$ and tH cross sections is then reinterpreted to derive constraints on the coupling parameters $(\kappa_t, \tilde{\kappa}_t)$.

Fiducial phase-space definition

The fiducial phase space is defined at particle level by the minimal set of object requirements common to both the hadronic and leptonic channels: two isolated photons in the detector acceptance and at least one central b jet. The single-b-jet requirement is motivated by the tH topology, where only one central b jet from the top quark decay is expected in the signal topology.
The fiducial photons are required to satisfy $|\eta| < 2.5$, excluding the barrel-endcap transition region $1.444 < |\eta| < 1.566$. In addition, the geometric mean of the transverse momenta of the two selected photons is required to satisfy $\sqrt{p_{\mathrm{T}}^{\gamma 1} p_{\mathrm{T}}^{\gamma 2}} > \frac{m_{\gamma\gamma}}{3}$, while the subleading photon must fulfil $p_{\mathrm{T}}^{\gamma 2} > \frac{m_{\gamma\gamma}}{4}$. Moreover, the isolation parameter, defined as the sum of p_{T} of all visible stable final state particles within $\Delta R < 0.3$ around each photon, is required to be smaller than 10 GeV. The b jet is required to satisfy $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$, motivated by the tracker coverage needed for b-tagging.
The use of the geometric mean of the photon transverse momenta is now standard in CMS Run-3 measurements in the $H \to \gamma\gamma$ decay channel [45]. Compared with the earlier choice of independent photon-p_{T} thresholds scaled to the diphoton mass, this definition is motivated by

an improved perturbative convergence of fiducial predictions for ggH production in the diphoton final state [160]. In the present measurement, the detector-level preselection is chosen to follow this fiducial definition closely.

4.2. Event processing with the HiggsDNA framework

The event processing is performed with the HiggsDNA framework [161], where the acronym DNA denotes Diphoton NanoAOD. This refers to the reduced CMS NanoAOD data tier [162, 163], for which HiggsDNA is designed. It uses the columnar Python analysis ecosystem, in particular `coffea`, `uproot`, and `awkward` arrays [164–167]. In contrast, the earlier $H \to \gamma\gamma$ workflows were based on the `flashgg` framework [168] inside CMSSW using the MiniAOD data tier. The MiniAOD representation carries much more information per event, for example individual reconstructed per-channel energies from the calorimeters, while NanoAOD mostly contains high-level objects and their properties, such as reconstructed photons and electrons with shower-shape and isolation variables. The NanoAOD format was designed to strongly reduce the storage and data traffic required for researchers who perform the typical analyses based on high-level variables. This reduces the usual event size from tens of kB to the order of 1 kB. As HiggsDNA does not depend on the extensive CMSSW framework, it can be installed in a standard Python environment and does not require the ROOT analysis framework [169] at runtime. This substantially simplifies the integration with modern scientific-Python tools such as `numpy`, `pandas`, and `pyarrow`, and it is one of the reasons why HiggsDNA is being adopted as the common processing framework for CMS $H \to \gamma\gamma$ measurements in Run 3 and, increasingly, also for ongoing analyses of Run-2 data.

The framework follows the processor-based execution model provided by `coffea`. It contains several dedicated analysis workflows targeting different physics analyses, all built on a common infrastructure. The individual processors share common building blocks for the object selections, the application of calibrations and SFs together with their systematic variations, and the bookkeeping and output machinery needed for large-scale distributed processing. The same HiggsDNA processor can be run interactively on a single machine or scaled out on distributed resources such as HTCondor [170] and Slurm clusters [171], for example through orchestration layers such as Parsl [172]. Examples of such workflows in public results built on the HiggsDNA framework include the inclusive and differential fiducial Higgs boson cross-section measurement [45], the search for Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state [173], and the search for Higgs boson production in association with a charm quark [174]. In addition, dedicated processors exist for tag-and-probe selection and for the production of $Z \to e^+e^-$ and $Z \to \mu^+\mu^-\gamma$ calibration and validation samples.

The event content is read directly from ROOT-based NanoAOD files through `uproot` into columnar `awkward` arrays, after which the event selection and the relevant object-level manipulations are performed. The crucial point for computational efficiency is that the file content is not processed through an explicit Python loop over events. Instead, physics observables are represented as arrays, including the variable-length object collections typical of collider events, and selections and transformations are applied to whole arrays from large chunks of events at once. The numerically intensive operations are thereby delegated to compiled kernels, while Python mainly configures the workflow. This is what makes the columnar approach fast in Python: a pure Python event loop would be too slow for such big-data analyses, whereas array-wise operations can reach a performance comparable to traditional compiled workflows while preserving a

much simpler analysis implementation [164, 166]. By contrast, the previous `flashgg` workflow was built around the CMSSW software stack with C++ modules and explicit event loops, which is performant but considerably more rigid from the perspective of the analysis development.
At the analysis level, HiggsDNA applies the trigger selection, the loose photon preselection, object calibrations, and the corresponding correction factors and uncertainties. In practice, these corrections are configured through external JSON files, such that updates to calibration constants or uncertainty prescriptions can be propagated by exchanging the corresponding configuration file rather than regenerating the upstream input datasets as a whole. This is a substantial practical advantage over the earlier MiniAOD-based workflow, where several corrections, for example energy-scale and random-number-based resolution adjustments, were already embedded in centrally produced objects and could not be reverted easily. Later changes were therefore much less flexible and could require an expensive upstream reprocessing of the input samples. The relevant event content is then written to reduced Parquet files in a flat representation without nested event structures. These outputs can be read efficiently in the subsequent analysis code with standard tools such as `awkward`, `pyarrow`, or `pandas`, and they provide the common input to the signal-region selection, the $Z \to e^+e^-$ validation region, the multivariate classification, and the final statistical analysis.
For the present measurement, a dedicated HiggsDNA processor, denoted `top.py`, is used. Beyond the diphoton candidate construction, it selects the jets and saves the associated b-tagging scores and also writes the electron, muon, and missing-transverse-momentum observables needed for the analysis of top quark decays to the output files. This processor was developed during the work presented in this thesis and constitutes the first HiggsDNA processor that consistently extends the diphoton workflow to electrons, muons, missing transverse momentum, and b-tagging information. As a consequence, several corrections and uncertainty treatments for these additional objects in the general HiggsDNA framework were implemented during this work.

4.3. Signal-region event selection

Building on the reconstruction, identification, and calibration of the physics objects described in Chap. 3, the following defines the specific requirements for the reconstructed objects used in the analysis selection.

Object definitions

Photon candidates are required to satisfy $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and to lie within the ECAL and tracker acceptance, i.e. $|\eta_{\mathrm{SC}}| < 2.5$ with the barrel-endcap transition region $1.444 < |\eta_{\mathrm{SC}}| < 1.566$ excluded. Here, the subscript SC refers to the associated ECAL supercluster, while the final photon direction used in the kinematic reconstruction receives a small correction with respect to the reconstructed primary vertex. In addition, they must pass a loose requirement of reaching a photon ID score greater than -0.7, and satisfy a hadronic-leakage requirement of $H/E < 0.08$ together with the conversion-safe electron veto. The remaining preselection follows the standard CMS $H \to \gamma\gamma$ strategy based on separate requirements for high-$R9$ and low-$R9$ candidates. For high-$R9$ photons, the compactness of the shower already provides strong discrimination against backgrounds, whereas for low-$R9$ photons additional requirements on the shower width and on charged-hadron and photon isolation are imposed, with the photon-isolation requirement corrected for the event energy density ρ. In addition, each photon must either satisfy $R9 > 0.8$

Table 4.1: Summary of the single-photon preselection criteria applied in HiggsDNA. The requirements are given separately for EB and EE photon candidates with high and low $R9$ values; dashes indicate that no additional requirement is imposed in that category.

Quantity	EB, high $R9$	EB, low $R9$	EE, high $R9$	EE, low $R9$
$R9$	> 0.85	$0.50 < R9 < 0.85$	> 0.90	$0.80 < R9 < 0.90$
p_T		$> 25\,\text{GeV}$		
$\lvert\eta_{SC}\rvert$		< 1.444	$1.566 < \lvert\eta_{SC}\rvert < 2.5$	
Photon ID score		> -0.7		
H/E		< 0.08		
Electron veto		required		
Charged-hadron isolation		for $R9 \leq 0.8$: $I_{trk} < 20\,\text{GeV}$ or $I_{trk}/p_T < 0.3$		
$\sigma_{i\eta i\eta}$	–	< 0.015	–	< 0.035
I_{trk}	–	$< 6\,\text{GeV}$	–	$< 6\,\text{GeV}$
I_γ^ρ	–	$< 4\,\text{GeV}$	–	$< 4\,\text{GeV}$

or pass a loose charged-hadron isolation requirement, $I_{trk} < 20\,\text{GeV}$ or $I_{trk}/p_T < 0.3$. The photon definition used in this analysis is summarised in Tab. 4.1 and is designed to be slightly tighter than the photon definition used in the trigger selection.

Reconstructed electrons are required to satisfy $p_T > 15\,\text{GeV}$ and $|\eta| < 2.5$, excluding the ECAL transition region, and to pass a multivariate ID working point calibrated for an efficiency of 90% for electrons from $Z \to e^+e^-$ decays. Reconstructed muons are required to satisfy $p_T > 10\,\text{GeV}$ and $|\eta| < 2.4$, to pass the medium cut-based ID working point, and to satisfy a tight PF isolation requirement. Both electrons and muons are required to be separated from the two selected photons by $\Delta R > 0.2$.

Jets are reconstructed over the wider detector acceptance and are required to satisfy $p_T > 20\,\text{GeV}$ and $|\eta| < 4.7$, together with a tight lepton-veto ID requirement. They are cleaned against the selected photons and leptons with a separation requirement of $\Delta R > 0.4$. An additional veto is applied on jets with $p_T < 50\,\text{GeV}$ in the region $2.5 < |\eta| < 3.0$ in order to suppress known detector artefacts in this phase-space region. This reduces the signal efficiency by around 2.5% for tH events with semi-leptonic top quark decays, while the effect on $t\bar{t}H$ and hadronic tH events is below 1%. Hence, this veto is regarded as a useful compromise between signal efficiency and the suppression of spurious background contributions. Within this thesis, a jet is referred to as b-tagged if it passes a loose working point of the Robust Particle Transformer b-tagging discriminator, designed to achieve a light-jet rejection factor of 10. In practice, only jets within the tracker acceptance, $|\eta| < 2.5$, are considered for b-tagging, since the identification relies on displaced-track information.

Diphoton selection criteria

Events are selected online with dedicated diphoton triggers, following the standard CMS Run-3 strategy for analyses in the $H \to \gamma\gamma$ final state [45]. These triggers require two photon candidates and combine transverse-momentum thresholds with additional requirements on shower-shape and isolation-related quantities in order to keep the trigger rate manageable at high instantaneous luminosity. The corresponding trigger-efficiency correction is derived with a tag-and-probe method using $Z \to e^+e^-$ events in data and applied to the simulation.

In HiggsDNA, events are retained if at least two photon candidates satisfy the common preselection described above. In the case of more than two photons, the pair with the highest p_T

Table 4.2: Summary of the event selection criteria of both channels in the signal region and the validation region.

Requirement	Hadronic channel	Leptonic channel	$Z \to e^+e^-$ validation region
Trigger	diphoton	diphoton	dielectron
Photon definition	nominal	nominal	e^+e^- reconstructed as $\gamma\gamma$
Electron veto	required	required	inverted
Diphoton mass	100–180 GeV	100–180 GeV	70–110 GeV
Fiducial selection	required	required	required
Photon ID scores	> -0.7	> -0.7	> -0.7
Loose b-tag	≥ 1	≥ 1	≥ 1
Jet multiplicity	≥ 2	≥ 1	≥ 2
Lepton multiplicity	0	≥ 1	0

of the diphoton system is selected. The diphoton mass is restricted to $100 < m_{\gamma\gamma} < 180\,\text{GeV}$. Within this selected mass range, the interval $120 < m_{\gamma\gamma} < 130\,\text{GeV}$ defines the signal window which is blinded during the optimisation steps of this analysis, while the remaining parts of the spectrum are referred to as sidebands which are not expected to contain a significant signal contribution. Compared with a narrower mass window around the Higgs boson resonance, this wider range substantially increases the event statistics available for background modelling and for the training of the multivariate classifiers, while the final signal extraction is performed with narrow bins in $m_{\gamma\gamma}$ to exploit the full discrimination from the mass. The leading photon is required to satisfy $p_\text{T} > 35\,\text{GeV}$, and the selected diphoton candidate is further required to satisfy the detector-level analogue of the fiducial photon requirements introduced above.

Further criteria for the hadronic and leptonic channel

The event-level requirements for the analysed phase-space regions are summarised in Tab. 4.2. In both channels of the signal region, the fiducial requirement of at least one central b jet is implemented at detector level by demanding at least one b-tagged jet. This already suppresses the dominant continuum background strongly, while keeping a high signal efficiency.
The hadronic channel is defined by the absence of reconstructed electrons and muons and requires at least two jets to suppress low-multiplicity backgrounds. The leptonic channel requires at least one reconstructed electron or muon. Since the jet multiplicity is lower for semi-leptonic top quark decays, no further jet in addition to the b-tagged one is required in this channel. Likewise, no explicit selection on missing transverse momentum is applied, such that the multivariate classifiers can exploit the full event information without introducing additional hard thresholds at the preselection stage.

4.4. Validation region with $Z \to e^+e^-$ decays

In addition to the signal region, a validation region based on $Z \to e^+e^-$ decays is defined to test the modelling of detector-response effects in a control sample with much higher event yields than the signal region. As discussed in Sec. 3.3.2, electron and photon reconstruction at the CMS experiment is closely related because both object types are built from electromagnetic showers in the ECAL, with the main separation provided by tracking information. A region in which electrons are reconstructed as photon candidates is therefore useful for validating the modelling of shower-shape, isolation, and energy-response observables relevant to photons, under the as-

sumption that the dominant residual mismodelling is similar for electrons and photons. After the electron-specific substitutions described below, the remaining selection is aligned with the hadronic signal region. The region can therefore also be used to compare data and simulation for the jet, b-tagging, missing-transverse-momentum, and classifier observables entering the analysis. Observables tied specifically to the explicit lepton requirement of the leptonic channel are not validated with this region, since an analogous $Z \to e^+e^-$ sample with an additional lepton does not provide sufficient event counts for this purpose in data. Within this sample, the processes $t\bar{t}Z$ and t-channel tZ production with $Z \to e^+e^-$ provide the closest analogues to the $t\bar{t}H$ and tH signal topologies, since the associated top-quark production topologies are retained while the Higgs boson is replaced by a Z boson. Although the region is not dominated by these processes, it still probes the modelling of observables in event classes that are structurally close to the targeted signal.
In the event selection, photon candidates in the NanoAOD photon collection are required to be matched to reconstructed electrons, and the electron-veto requirement that is normally used to suppress electrons in the signal region is inverted. The two electron-matched photon candidates are required to have opposite reconstructed electric charge and to satisfy a dedicated dielectron trigger selection. At the event level, the reconstructed diphoton mass is required to satisfy $70 < m_{\gamma\gamma} < 110\,\text{GeV}$, such that the selected sample is centred on the Z-boson resonance. The remaining downstream event selection is identical to that of the hadronic signal region, including the fiducial requirement, the photon ID score threshold on both selected candidates, the loose b-tag requirement, and the hadronic jet and lepton multiplicity cuts. For the efficiency correction of the electrons reconstructed as photons, the electron-reconstruction SF is applied.

5. Datasets used in the analysis

This chapter provides details about the data that are used to measure the cross sections. First, the data that were recorded with the CMS detector are described, where the relevant differences of the analysed data-taking periods are highlighted. The later part of this chapter details the modelling of the $H \to \gamma\gamma$ processes as used in the final measurement, as well as other simulation samples that are used to optimise the analysis chain and for validation purposes.

5.1. Recorded proton–proton collision data

The proton–proton collision data used in this thesis were recorded with the CMS detector in the first two years of the Run-3 data-taking programme, 2022 and 2023. These data were taken at $\sqrt{s} = 13.6\,\text{TeV}$, corresponding to the highest centre-of-mass energy achieved in a particle collider.

The measurements of the integrated luminosity that was delivered to CMS yield the results of $41.5\,\text{fb}^{-1}$ for 2022 [175] and $32.7\,\text{fb}^{-1}$ for 2023 [176], with uncertainties of 1.4% and 1.3%, respectively. The integrated luminosity that was recorded with all relevant subdetectors working properly, as required for the analysis presented in this thesis, is lower, and corresponds to a total of $62.4\,\text{fb}^{-1}$ for the two years combined. Here, the absolute luminosity scale is calibrated with van der Meer scans [177, 178], in which the two LHC beams are displaced with respect to each other in the transverse plane. The interaction rate is then measured as a function of the beam separation, allowing the effective beam overlap and thus the absolute luminosity to be determined.

The two years of data taking are further subdivided according to changes in relevant detector conditions [179]. In 2022, a power-cooling problem affected about 7% of the channels in the positive-z EE, such that they were switched off after the issue appeared. In this part of the EE, the reconstruction efficiency for electrons and photons is reduced. This effect is also taken into account in the simulation, and the 2022 data-taking period is therefore simulated in two parts. These periods are labelled *2022preEE* and *2022postEE* in the following. In 2023, part of the data taking was affected by a failure in the clock-synchronisation system of 28 modules of the BPix detector in layers 3 and 4, which induced a track reconstruction inefficiency within

Table 5.1: Integrated luminosities of the data-taking periods used in this thesis. Only the runs which are certified as good for physics analysis, as used in this thesis, are included.

Data-taking period	Integrated luminosity [fb^{-1}]
2022preEE	7.98
2022postEE	26.7
2023preBPix	18.1
2023postBPix	9.69
Total 2022+2023	62.4

$-1.5 < \eta < 0$ and $-1.2 < \phi < -0.8$. This effect is also included in the simulation and the respective periods are labelled *2023preBPix* and *2023postBPix*. The integrated luminosities of the analysed sub-periods are given in Tab. 5.1.
Properties of the Run-3 data-taking conditions at the CMS experiment are summarised in Fig. 5.1. It shows the luminosity-weighted distribution of the mean number of interactions per bunch crossing, calculated with the reference inelastic proton–proton cross section of $\sigma_{\text{in}}^{pp} = 80.0\,\text{mb}$ used for Run-3 comparisons, and the cumulative delivered and recorded luminosity for the Run-3 proton–proton data taking at the CMS experiment [115]. For the two years used in this thesis, the average pileup is about $\langle\mu\rangle = 46$ in 2022 and $\langle\mu\rangle = 52$ in 2023. The cumulative-luminosity plot illustrates the strong increase in the delivered and recorded data volumes from 2022 to 2025, while the pileup distribution shows that the datasets recorded in 2024 and 2025 are not only substantially larger but also extend to significantly higher pileup values. While these later years are not yet used in the present analysis, they indicate the experimental conditions relevant for future extensions of this measurement to the full Run-3 dataset.

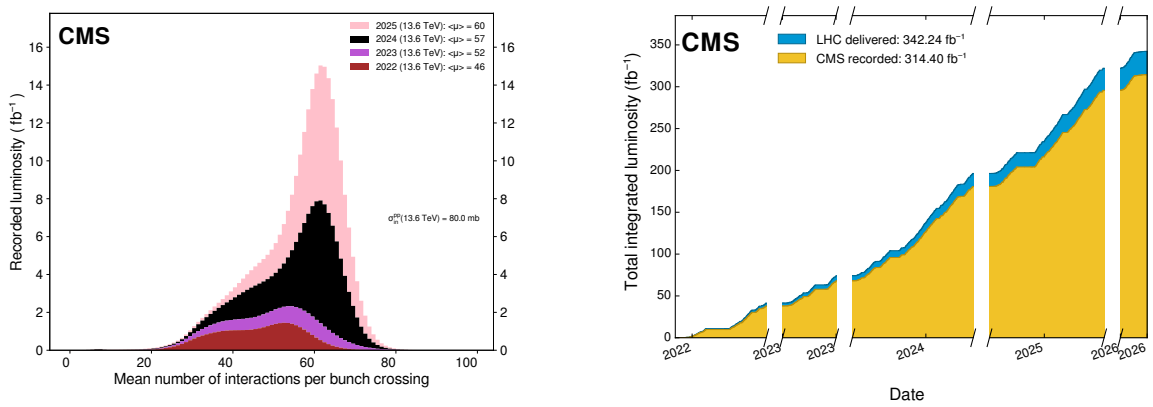


Figure 5.1: Summary plots for the Run-3 proton–proton collision data at $\sqrt{s} = 13.6\,\text{TeV}$ recorded with the CMS detector in Run 3 of the LHC up to this point [115]. Left: luminosity-weighted distribution of the mean number of interactions per bunch crossing, calculated with a reference cross section for inelastic proton–proton scattering of $\sigma_{\text{in}}^{pp} = 80.0\,\text{mb}$. Right: cumulative delivered and recorded luminosity as a function of time, illustrating the substantial increase in the available Run-3 dataset beyond the years used in this thesis.

5.2. Simulated datasets

In order to compare the measured data to theoretical expectations, simulated datasets based on Monte Carlo (MC) methods are required. In all datasets described below, the response of the simulated events in the CMS detector is modelled with a detailed detector simulation based on GEANT4 [180]. Unless stated otherwise, the simulated hard-scattering events are interfaced to PYTHIA 8.306 [84] for PS and hadronisation, using the CP5 set of tuned parameters [181]. For Higgs boson processes, PYTHIA is also used to model the $H \to \gamma\gamma$ decay. To model pileup, additional minimum-bias proton–proton interactions, i.e. generic inelastic collisions dominated by soft QCD processes, are overlaid on the simulated hard-scattering events. Throughout the simulated datasets discussed below, the top-quark mass is set to 172.5 GeV. For the event generation, the Higgs boson mass is set to 125 GeV and its width parameter is set to 4 MeV. The NNPDF3.1 set of PDFs (`NNPDF31_nnlo_as_0118_mc_hessian_pdfas`) [78] is the conventional choice for centrally produced Run-3 simulation within the CMS Collaboration, where applicable. The renormalisation and factorisation scales are chosen according to the default dynamical-scale prescription of `MadGraph5_aMC@NLO` [182] for most of the processes mentioned below. For samples generated at NLO in QCD, this corresponds to one half of the sum of the transverse masses of the final-state particles. For LO samples, the transverse mass of the irreducible $2 \to 2$ system obtained from a k_T clustering of the final-state partons is used.
The selected phase space receives contributions not only from the signal processes $t\bar{t}H$, tH, and tHW, but also from other Higgs boson production modes and from non-resonant background processes. Following the photon-background terminology introduced in Sec. 3.3.2, the latter are separated into irreducible continuum backgrounds with two prompt photons and reducible continuum backgrounds with $j \to \gamma$ or $e \to \gamma$ misidentifications. The simulation samples are grouped accordingly into signal processes, resonant backgrounds containing $H \to \gamma\gamma$ decays from other production modes, and continuum backgrounds with a smoothly falling diphoton-mass spectrum. For the continuum background processes, the cross sections obtained from the generators used are taken as the initial normalisation. As discussed in Chap. 6, the dominant continuum components are subsequently corrected with data-driven normalisation factors for the classifier training and categorisation optimisation to improve the agreement with the data.

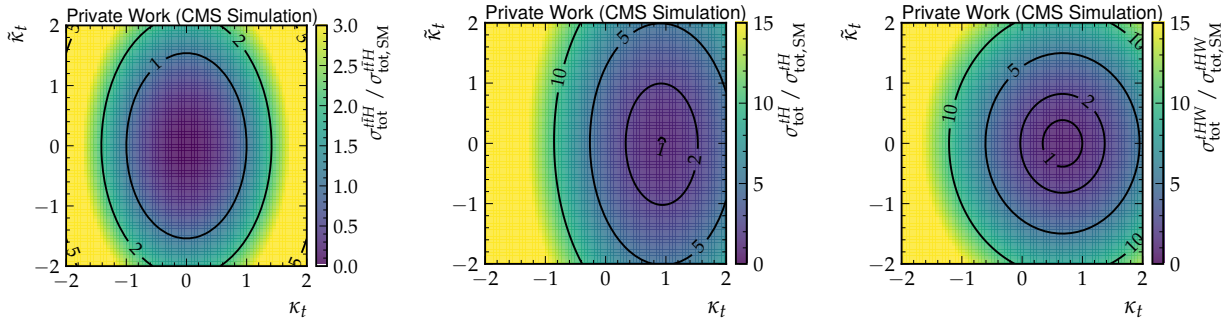
5.2.1. Signal processes

The signal processes $t\bar{t}H$, tH, and tHW are simulated with `MadGraph5_aMC@NLO` in version 2.9.9 [183], interfaced with the Higgs Characterisation Model [70] to customise its couplings. For $t\bar{t}H$, the event generation is performed at NLO accuracy in QCD in the five-flavour scheme with up to one additional jet at matrix-element level and FxFx merging [184]. This follows the convention in the CMS Collaboration for this process. The top quark decays are modelled with `MadSpin` [185], in order to retain the spin correlations in the top-quark decay chain. In addition to the SM sample, a dedicated sample with a purely CP-odd top-quark Yukawa coupling, corresponding to $\kappa_t = 0$ and $\tilde{\kappa}_t = 1$, is produced at NLO accuracy. The $t\bar{t}H$ signal model is constructed from samples with purely CP-even and purely CP-odd top-quark Yukawa couplings only. This treatment follows previous experimental practice in CP-sensitive $t\bar{t}H$ analyses and has been validated with additional samples with mixed Yukawa structures to provide a good approximation for the relevant classifier-output distributions [26, 158]. For the inclusive cross-section normalisation, omitting this interference is also consistent with the discussion in Sec. 2.3.1, since the CP-odd scalar–pseudoscalar interference integrates to zero over the full phase space.

The t-channel tH production process is also simulated at NLO in QCD, using the four-flavour scheme. In this scheme, bottom quarks are treated as massive and the initial $g \to b\bar{b}$ splitting is described explicitly at matrix-element level. While the four- and five-flavour schemes give consistent predictions for inclusive rates and core kinematic distributions within scale uncertainties at NLO, the four-flavour scheme provides a more complete description of observables involving the spectator bottom quark and additional jets [30]. Besides the SM configuration, five additional coupling points in the $(\kappa_t, \tilde{\kappa}_t)$ plane are generated on the unit circle at mixing angles $\alpha = \pm\pi/4$, $\alpha = \pm\pi/2$, and $\alpha = \pi$, together with one sample with the top-quark Yukawa coupling set to zero. These points are chosen to allow for later interpolation within the coupling-modifier plane, including the CP-even interference between amplitudes involving the HWW and scalar top-quark Yukawa couplings, while mostly covering the parameter region compatible with previous measurements [27, 28, 36, 158]. The HWW coupling is kept fixed to its SM value throughout. For improved statistical precision in the relevant final states, the tH simulation is split into samples with semi-leptonic and hadronic top quark decays and generated with equal event counts. The tH samples and the $t\bar{t}H$ sample with a purely CP-odd top-quark Yukawa coupling used here were produced in the course of this thesis and extend the signal modelling beyond what was available in previous CMS analyses, where these configurations were generally treated only at LO accuracy. This is important not only for the inclusive normalisation but also for the kinematic properties relevant for the analysis, as emphasised in Refs. [29, 30], and reduces the scale uncertainties significantly.

The tHW process is simulated at LO in QCD in the five-flavour scheme. Its contribution is subleading for the present analysis, and the available samples contain matrix-element reweights that allow the interpolation to other coupling hypotheses. This interpolation is based on the built-in reweighting machinery of `MadGraph5_aMC@NLO`, for which alternative event weights are obtained from the ratio of matrix elements under different coupling scenarios evaluated for the same phase-space point [186, 187]. These weights include a set of coupling points distributed on an ellipse in the $(\kappa_t, \tilde{\kappa}_t)$ plane with half-axes chosen as $\kappa_t = 1$ and $\tilde{\kappa}_t = 2/3$. In this LO five-flavour-scheme setup, the tHW matrix element does not contain the doubly resonant topologies that overlap with $t\bar{t}H$ production. The corresponding tHW–$t\bar{t}H$ interference is therefore not included in the present signal modelling and would arise only in more complete treatments, for example at NLO in the five-flavour scheme or already at LO in the four-flavour scheme [31, 72].

The generated signal samples are then used to parameterise the dependence of the total inclusive signal cross sections on the Yukawa-coupling modifiers. This dependence is determined from least-squares fits of second-order polynomials. For all three signal processes, terms linear in $\tilde{\kappa}_t$ or proportional to $\kappa_t\tilde{\kappa}_t$ are omitted in the inclusive cross-section parameterisations, following the discussion in Chap. 2. For the single-top quark associated production modes, this choice was additionally validated by testing extended parameterisations including the omitted linear CP-odd terms; the corresponding fitted coefficients were found to be compatible with zero within the fit precision. In the tH and tHW parameterisations, the constant terms represent the contribution from the squared HWW amplitude, with the HWW coupling fixed to its SM value, while the terms linear in κ_t describe the CP-even interference between amplitudes involving the HWW and scalar top-quark Yukawa couplings. For $t\bar{t}H$, no term linear in κ_t and no constant term are included because there is no corresponding tree-level amplitude without a top-quark Yukawa coupling, if off-shell effects introducing a HWW term are neglected. Only the fitted central values are quoted in the equations below and used in the statistical analysis; the uncertainties of



(a) $t\bar{t}H$ **(b)** tH **(c)** tHW

Figure 5.2: Scaling behaviour of the total cross sections for the signal processes as a function of the top-quark Yukawa coupling parameters $(\kappa_t, \tilde{\kappa}_t)$, as obtained from the simulation used in this thesis. The corresponding polynomial parameterisations are given in Eqs. (5.1)–(5.3).

the fit parameters are assumed to be negligible for the present analysis. The resulting ratios of the total inclusive cross sections to the corresponding SM predictions are given by

$$\sigma^{t\bar{t}H}_{\text{tot}}(\kappa_t, \tilde{\kappa}_t) = \left(1.00\,\kappa_t^2 + 0.42\,\tilde{\kappa}_t^2\right) \sigma^{t\bar{t}H}_{\text{tot,SM}}\,, \tag{5.1}$$

$$\sigma^{tH}_{\text{tot}}(\kappa_t, \tilde{\kappa}_t) = \left(2.83\,\kappa_t^2 - 5.28\,\kappa_t + 1.00\,\tilde{\kappa}_t^2 + 3.45\right) \sigma^{tH}_{\text{tot,SM}}\,, \tag{5.2}$$

$$\sigma^{tHW}_{\text{tot}}(\kappa_t, \tilde{\kappa}_t) = \left(2.62\,\kappa_t^2 - 3.51\,\kappa_t + 1.91\,\tilde{\kappa}_t^2 + 1.89\right) \sigma^{tHW}_{\text{tot,SM}}\,. \tag{5.3}$$

This scaling behaviour is visualised in Fig. 5.2. These results are qualitatively consistent with similar studies in the literature, for example from Refs. [28, 48]. As discussed in Sec. 2.3.1, the strong dependence of the single-top quark associated production modes on the coupling parameters originates from the interference between amplitudes involving the top-quark Yukawa coupling and the HWW coupling. In the SM, this interference is destructive, leading to the suppressed tH and tHW production rates. For an inverted-sign CP-even Yukawa coupling, corresponding to $\kappa_t = -1$ and $\tilde{\kappa}_t = 0$, the interference becomes constructive, leading to a strong enhancement of the total rates by factors of about 11.6 for tH and 8.0 for tHW relative to the SM prediction. For a purely CP-odd coupling, $\kappa_t = 0$ and $\tilde{\kappa}_t = 1$, the enhancement is also substantial, with a factor of 4.4 for tH and 3.8 for tHW, whereas the $t\bar{t}H$ rate is reduced to 0.42 times its SM value. The smaller coefficient of $\tilde{\kappa}_t^2$ compared to κ_t^2 in the $t\bar{t}H$ scaling reflects the different radiation pattern of scalar and pseudoscalar Higgs bosons from top quarks. In the collinear high-energy picture, the scalar $t \to tH$ fragmentation function contains an additional soft-enhanced term proportional to $(1-x)/x$, where x is the longitudinal momentum fraction carried by the emitted Higgs boson, which is absent for pseudoscalar emission $t \to tA$ [88, 188–191]. Consequently, pseudoscalar associated production is suppressed relative to scalar production for a light Higgs boson and equal coupling strengths.

5.2.2. Resonant background from other $H \to \gamma\gamma$ processes

The resonant-background samples considered here are ggH, VBF, VH, and $b\bar{b}H$ production with $H \to \gamma\gamma$ decays. The ggH, VBF, and inclusive VH samples are produced in the Higgs Characterisation model with `MadGraph5_aMC@NLO` at NLO in QCD in the five-flavour scheme. The ggH sample is then reweighted in bins of the Higgs boson p_{T} and jet multiplicity to morph

the spectrum to the NNLO+PS prediction from Refs. [192, 193]. In the ggH production, the bottom-quark Yukawa coupling is set to zero, such that diagrams with the Higgs boson radiated from bottom quarks are not included. These contributions are instead covered by a separate $b\bar{b}H$ sample generated with POWHEG [194–196] at NLO in QCD, which also contains the interference terms proportional to the product of the top- and bottom-quark Yukawa couplings.
For the optimisation studies, more granular simulation is additionally used for the VH production modes, with separate samples for W^+H, W^-H, and ZH, and with further splitting according to the vector-boson decay modes. The decays $W \to \ell\nu$, $W \to q\bar{q}'$, $Z \to \ell^+\ell^-$, $Z \to \nu\nu$, and $Z \to q\bar{q}$ are included in individual datasets. These samples are generated with POWHEG at NLO in QCD and are employed for classifier training and categorisation optimisation. For the statistical analysis itself, however, the inclusive VH sample is used to ensure statistical independence of the events used in the classifier training and to avoid introducing unnecessary granularity into the modelling workflow of the $H \to \gamma\gamma$ processes.
For the Higgs boson processes that enter the statistical analysis directly, the SM reference cross sections used in this thesis are summarised in Tab. 5.2. These values are taken from the cross-section recommendations for $\sqrt{s} = 13.6\,\text{TeV}$ [85]. There, the central values were obtained by interpolating predictions between $\sqrt{s} = 13$ and $14\,\text{TeV}$. Following that prescription, an interpolation between nearby Higgs boson mass points is done in this analysis in order to obtain predictions for $m_H = 125.08\,\text{GeV}$, the result from the most precise measurement performed at the CMS experiment [97], which is the mass hypothesis used in the likelihood fits in this analysis. This choice defines the signal normalisation convention used throughout the present statistical model. The values quoted in the table correspond to the production cross sections; the branching ratio for $\mathcal{B}(H \to \gamma\gamma) = 0.227\%$ [72] is additionally included in the normalisation in the analysis.

Table 5.2: SM production cross sections used to normalise the Higgs boson processes in this analysis. The values are obtained from the recommendations at $\sqrt{s} = 13.6\,\text{TeV}$ [85], interpolated to $m_H = 125.08\,\text{GeV}$.

Signal process	σ [fb]	Resonant background process	σ [pb]
$t\bar{t}H$	568.9	ggH	52.2
tH	83.5	VBF	4.08
tHW	17.2	VH	2.40
		$b\bar{b}H$	0.526

5.2.3. Irreducible continuum background

In the irreducible continuum background, the relevant processes contain two isolated prompt photons together with additional jets and can therefore populate the selected phase space with final states very similar to those of the signals. The dominant contribution is diphoton production, denoted by $\gamma\gamma$, which can enter the signal region if events contain a b-tagged jet; it may also populate the leptonic phase space in the presence of misidentified leptons. This process is simulated at LO with Sherpa 2.2.2 [197], where quark- and quark-gluon-induced Born configurations are generated with up to three jets at matrix-element level, and also the gluon-induced box contribution is included without additional partons in the matrix element. The PS, hadronisation, and pileup are modelled with Sherpa as well.
Much smaller cross sections, but final states that are experimentally almost indistinguishable from the signal except for the diphoton resonance structure, arise from top-quark production

with two additional photons. The processes $t\bar{t}\gamma\gamma$ and the t-channel $t\gamma\gamma$ are therefore also included in the simulated background model. They are simulated with MadGraph5_aMC@NLO at LO in QCD, in the five-flavour scheme for $t\bar{t}\gamma\gamma$ and in the four-flavour scheme for $t\gamma\gamma$. The $t\gamma\gamma$ simulation was requested specifically for the study presented in this thesis. Its consideration marks an improvement in the background description of the tH-like phase space over previous searches for tH production in the $H \to \gamma\gamma$ decay channel at the CMS experiment. For $t\bar{t}\gamma\gamma$, a multiplicative k-factor of 1.7 is applied to the LO cross section, as derived in Ref. [198] in a similar phase space. This is taken as the best available higher-order correction for this rare irreducible background in the present analysis.

5.2.4. Reducible continuum background

The reducible continuum background is dominated by processes in which one or two of the reconstructed photon candidates originate from misidentified objects. Single-photon production with additional jets, labelled γ+jet, contributes when one $j \to \gamma$ misidentification is present, while QCD multijet production contributes if two $j \to \gamma$ misidentifications occur. Both processes are simulated at LO in QCD with Pythia 8.306 as $2 \to 2$ scattering processes without additional outgoing partons at matrix-element level. In addition, a generator-level electromagnetic-enrichment filter is applied to enhance the efficiency for events entering the diphoton-like phase space.
Processes with top quarks also give rise to important reducible backgrounds. Top-quark pair production in association with a photon, $t\bar{t}\gamma$, and t-channel single-top quark production with a photon, $t\gamma$, are simulated at NLO in QCD with MadGraph5_aMC@NLO. In these samples, the photons are generated explicitly in the hard process, i.e., radiated from quarks in the matrix element before the top quark decays are modelled with MadSpin. They can enter the selected phase space in the presence of one additional $j \to \gamma$ or $e \to \gamma$ misidentification, or when further isolated photons are produced in top quark decays. The $t\bar{t}\gamma$ sample is generated in five slices of the photon transverse momentum in order to improve the description of the high-p_{T} region. The $t\gamma$ and $\bar{t}\gamma$ processes are generated separately, but inclusive in photon p_{T}. Inclusive $t\bar{t}$ and t-channel single-top quark production are also included and are simulated with MADGRAPH5_AMC@NLO at NLO in QCD. For $t\bar{t}$, the dileptonic, lepton+jets, and all-hadronic decay channels are generated separately.
Vector-boson production in association with a photon $(V\gamma)$ is considered as an additional reducible background source. In particular, the Drell–Yan (DY) process with an additional photon, $\gamma^*/Z^*\gamma \to \ell^+\ell^-\gamma$, and $W\gamma$ production are simulated with MADGRAPH5_AMC@NLO at NLO in QCD with up to one additional matrix-element parton and FxFx merging. Both processes are split into five photon-p_{T} slices. The DY+γ process is especially relevant in the leptonic channel, where it contributes mostly through $Z \to e^+e^-$ decays with an $e \to \gamma$ misidentification.

5.3. Other samples for validation with $Z \to e^+e^-$ decays

The validation region based on $Z \to e^+e^-$ decays was introduced in Sec. 4.4. The simulation samples used for this validation are summarised in the following.
In analogy to the $t\bar{t}H$ and tH signal processes with $H \to \gamma\gamma$ decays, the processes $t\bar{t}\ell^+\ell^-$ and t-channel tZ production are simulated at NLO in QCD with MADGRAPH5_AMC@NLO. The dominant background in this dielectron phase space is ordinary DY production, which is simulated as an $\ell^+\ell^-$+jets sample with up to two jets at matrix-element level and NLO accuracy

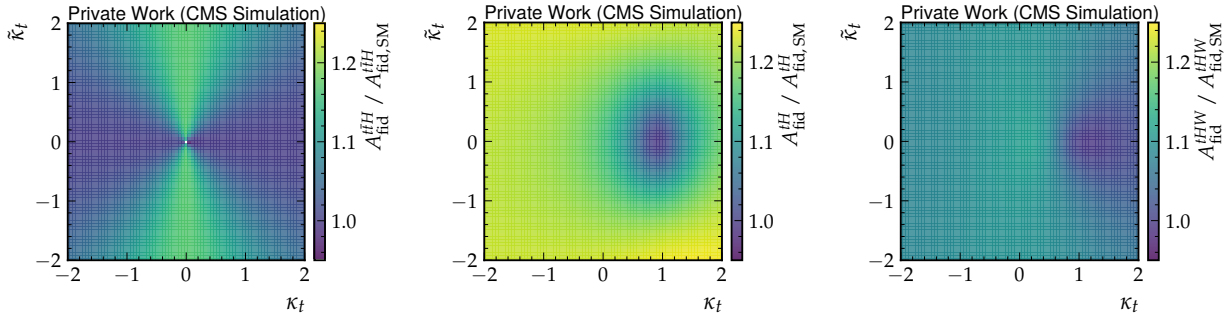
in QCD with MADGRAPH5_AMC@NLO. Top-quark pair production provides the second most important contribution and is described with the same $t\bar{t}$ datasets as used in the signal-region studies. Smaller contributions arise from diboson production and from top-quark production in association with a W boson. The processes $t\bar{t}W$ and tW and most VV production modes are simulated with MADGRAPH5_AMC@NLO at NLO in QCD, while for WW production the POWHEG generator is used at NLO accuracy.

6. Kinematic signal properties and modelling of the continuum background for the analysis optimisation

This chapter studies the kinematic properties of the signal and background processes relevant for the analysis optimisation. A detailed understanding of these properties is necessary to design a sensitive measurement, since the expected signal yields are small whereas the number of background events passing the preselection is much larger. For the signals, the focus is on the features that discriminate between $t\bar{t}H$ and tH production and on how these observables change under different top-quark Yukawa-coupling hypotheses. This discussion motivates the formulation of the measurement in a fiducial phase space and the later design of a classification strategy that remains robust under coupling variations.
The training of powerful deep-learning classifiers for background suppression requires a high-statistics event-level model of the continuum background. The available simulation, however, does not provide a sufficiently reliable description of the continuum background without further modifications. This is especially the case for reducible backgrounds with $j \to \gamma$ misidentifications. The approach developed here combines two aspects to improve the description of the continuum background: an increase of the remaining event count of simulated events with non-prompt photons after the selection, and data-driven corrections to the dominant background normalisations to align the normalisation and composition of the background model more closely with the data. This background model is distinct from the one used in the final statistical analysis, where the continuum contribution is determined directly from the diphoton-mass sidebands in data. Its purpose is therefore limited to classifier training, validation, and categorisation optimisation, while the final statistical analysis to determine the fiducial cross sections remains data-driven based on the observed diphoton-mass spectrum.

6.1. Kinematic properties of the signal and coupling dependence

The discussion in this section is based on events passing the detector-level selection introduced in Chap. 4. In the literature, the dependence of the kinematic properties of the signal processes on the Yukawa couplings is discussed, for example, in Refs. [29–31, 48, 69, 88, 199]. For the present analysis, one immediate consequence is particularly important: the fraction of signal events passing the detector-level selection depends appreciably on the assumed coupling hypothesis. This effect arises mainly from changes in the kinematic properties of the Higgs boson, which in turn modify the probability that the two photons satisfy the acceptance requirements. This is the main acceptance effect relevant for the fiducial formulation of the measurement defined in Sec. 4.1. The corresponding fiducial acceptance as a function of the coupling parameters, $A_{\text{fid}}(\kappa_t, \tilde{\kappa}_t)$, is shown in Fig. 6.1. The processes $t\bar{t}H$, tH, and tHW are shown separately in this figure to illustrate the process-specific dependence, although the tHW contribution is combined with $t\bar{t}H$ in the later analysis. The acceptance is obtained as the ratio of a second-order polyno-



(a) $t\bar{t}H$ **(b)** tH **(c)** tHW

Figure 6.1: Fiducial acceptance as a function of the coupling parameters κ_t and $\tilde{\kappa}_t$ for the three signal processes relative to the SM for a constant branching ratio $\mathrm{BR}(H \to \gamma\gamma)$. The acceptance is defined as the fraction of events that pass the fiducial phase space selection criteria, which are designed to closely match the experimental selection. The three processes are shown separately for interpretation, while the tHW contribution is combined with $t\bar{t}H$ in the later analysis.

mial fit to the fiducial cross section and the corresponding total cross-section parameterisation from Eqs. (5.1)–(5.3). The fiducial prediction for a signal process p is therefore obtained by multiplying its total production cross section by this acceptance and by the $H \to \gamma\gamma$ branching fraction,

$$\sigma^{p}_{\mathrm{fid}}(\kappa_t, \tilde{\kappa}_t) = \sigma^{p}_{\mathrm{tot}}(\kappa_t, \tilde{\kappa}_t)\, A^{p}_{\mathrm{fid}}(\kappa_t, \tilde{\kappa}_t)\, \mathcal{B}(H \to \gamma\gamma)\,. \tag{6.1}$$

Possible effects from variations of the branching ratio $\mathcal{B}(H \to \gamma\gamma)$ are not included in Fig. 6.1 to isolate the effect on the production processes. In the later analysis, the branching ratio variation is taken into account in one of the two interpretations performed to derive the coupling constraints shown in Sec. 9.5. Under the SM hypothesis, the acceptance is about 46% for $t\bar{t}H{+}tHW$ and 39% for tH events. It increases to about 53% and 47% in the purely CP-odd scenario and to about 46% and 48% in the inverted-sign scenario, respectively. These differences correspond to relative changes of up to about 20% with respect to the SM expectation, which is non-negligible compared with the precision already reached in previous $t\bar{t}H$ measurements [34, 93].

The signal processes share the presence of two prompt photons from the Higgs boson decay and at least one top quark with additional jet activity, but their overall topology is nevertheless distinct. In tH production, the Higgs boson and the top quark recoil against each other and are typically produced rather centrally, while the characteristic light quark from the t-channel topology tends to appear as an energetic forward jet. In the four-flavour scheme, an additional b jet which is typically very soft and forward is present due to the initial gluon splitting. In contrast, $t\bar{t}H$ production contains a top-quark pair and is therefore characterised by a larger overall jet multiplicity, while the activity in the forward region is typically smaller than in tH. The different charge structure also leads to an experimentally relevant asymmetry in the leptonic final state: tH events contain a larger fraction of positively charged leptons, whereas $t\bar{t}H$ events yield equal fractions of positive and negative lepton charges.

Among the reconstructed objects discussed below, the selected jet with the highest b-tagging score is referred to as the b-tagged jet, or briefly the b jet. The remaining jets are ordered by transverse momentum and denoted as jet 1, jet 2, and so on. In the leptonic channel, the lepton with the highest transverse momentum is referred to as the lepton for simplicity. Since only a small fraction of selected $t\bar{t}H$ events contains two reconstructed leptons, the kinematic prop-

erties of such subleading leptons are not used explicitly, whereas the lepton multiplicity itself remains informative and is considered.
Representative observables that separate $t\bar{t}H+tHW$ from tH while showing only a comparatively mild dependence on the coupling hypothesis are shown in Fig. 6.2. The larger hadronic activity of $t\bar{t}H$ events together with the higher forward activity for tH motivate a separate consideration of central and forward jet multiplicities as discriminators, which show only mild changes when considering a different coupling scenario. The number of b-tagged jets behaves similarly robustly, since it is mainly determined by the top-quark decay products and is therefore only weakly affected by changes in the Yukawa coupling. The same holds for the reconstructed lepton charge. Even in the hadronic channel, the missing transverse momentum tends to be larger for $t\bar{t}H$ than for tH. This is expected, since the hadronic selection is defined only by the absence of a selected lepton and is therefore not a pure all-hadronic truth category. Events with neutrinos can still enter when, for example, an electron or a muon from a top quark decay is not reconstructed or fails the selection, or when a top quark decay includes a hadronically decaying τ lepton.

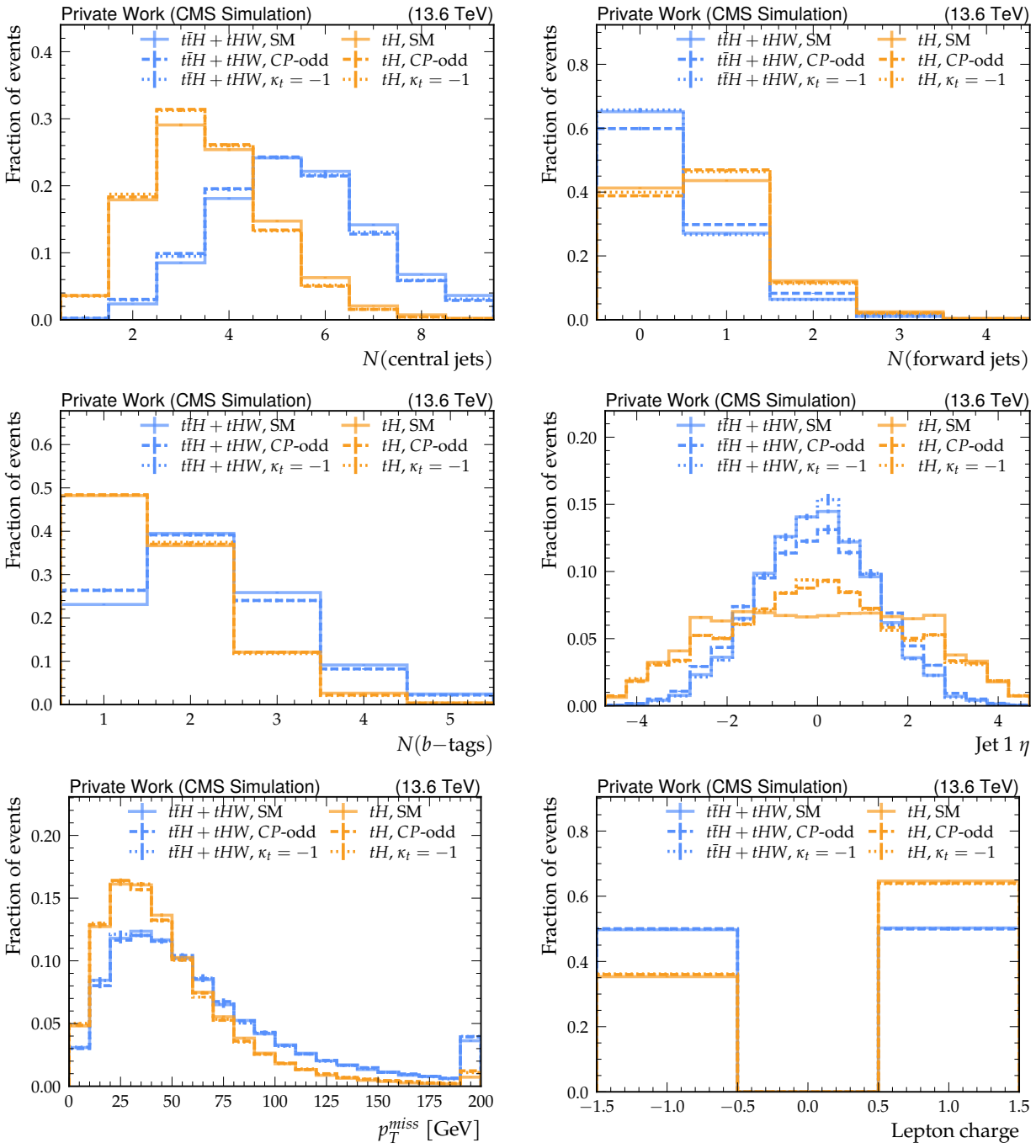


Figure 6.2: Kinematic observables found to be discriminative between the $t\bar{t}H + tHW$ and tH signal classes, but appear rather insensitive to the considered coupling scenarios. Events passing the hadronic selection are shown as an example, except for the distribution of the reconstructed lepton charge, which is only defined for events passing the leptonic selection. The coupling dependence is indicated by comparing the SM (solid), CP-odd (dashed), and inverted-sign (dotted) coupling scenarios. Statistical modelling uncertainties are marked and possible overflow is included in the rightmost bin.

By contrast, a number of observables shows a pronounced dependence on the coupling hypothesis, with representative examples shown in Fig. 6.3. These effects are mainly driven by Higgs boson observables and by angular correlations between reconstructed objects. In particular, the Higgs boson tends to have a harder transverse-momentum spectrum in the purely CP-odd and inverted-sign scenarios than in the SM case. As discussed in Ref. [69], pseudoscalar couplings suppress the near-threshold regime in $t\bar{t}H$ production, which leads to a reduction of events with small Higgs boson p_{T}. At the same time, the Higgs boson is produced more centrally in these alternative coupling scenarios. This explains the increase of the fiducial acceptance seen in Fig. 6.1. Previous studies identified quantities such as the rapidity difference of the two top quarks and invariant masses involving the Higgs boson and top quarks as particularly sensitive to the CP structure of the interaction [29, 69, 199]. In the present analysis, however, the top quarks are not reconstructed explicitly. Instead, the focus is on reconstructed low-level observables that can later be exploited by machine-learning methods. For tH production, Ref. [30] showed that the Higgs boson and the top quark tend to be closer in pseudorapidity in the purely CP-odd and inverted-sign scenarios than in the SM case. In the reconstructed final state, this manifests itself in observables such as the angular separation between the Higgs boson and the b-tagged jet, the angular separation between the b-tagged jet and jet 1, and, after leptonic selection, the rapidity difference between the Higgs boson and the lepton. Here, it is found that for many of the observables discussed below, the CP-odd and inverted-sign coupling scenarios exhibit very similar deviations from the SM prediction in the shapes of kinematic distributions, although both scenarios are known to modify the characteristic SM behaviour of tH production through different amplitude structures. This is also observed for tHW events; however, when combining them with $t\bar{t}H$ as done in this analysis, the inverted-sign scenario is kinematically close to the SM, because $t\bar{t}H$ production is invariant under this sign inversion.
Particularly relevant for this analysis is that several observables which are strongly discriminative for SM $t\bar{t}H$ and SM tH events become much more similar between the two processes in the CP-odd [48] and inverted-sign coupling scenarios. Examples shown here are $\Delta R(\gamma\gamma,\, b\text{-jet})$, $\Delta y(\gamma\gamma,\, \ell)$, and $\Delta R(b\text{-jet},\, \text{jet } 1)$. The quantity $y^{\text{jet}\,1} \oplus y^{\gamma\gamma} = \sqrt{(y^{\text{jet}\,1})^2 + (y^{\gamma\gamma})^2}$ was proposed as a comparatively coupling-robust discriminator between $t\bar{t}H$ and tH events in Ref. [48] and is also considered here. The motivation is that $t\bar{t}H$ events are concentrated closer to the origin in the $\left(y^{\text{jet}\,1}, y^{\gamma\gamma}\right)$ plane, while tH events populate larger radii due to the forward light jet. For tH, a CP-odd component changes the distribution approximately like a rotation in this plane, such that the radial distance remains more stable than rapidity-difference observables. In the leptonic channel, the behaviour described in Ref. [48] is reproduced. In the hadronic channel, however, a more pronounced model dependence is observed for tH events. The reason is that in this selection, the forward jet characteristic of tH production is less likely the leading jet in p_{T}. Instead, the leading jet can also originate from a top quark decay, which weakens the intended cancellation between the more forward jet and the more central Higgs boson in these scenarios.
Overall, the kinematic study suggests that a classifier trained only on SM signal samples will separate $t\bar{t}H$ and tH events most efficiently in the SM case itself, while more sizeable migrations between the two signal classes have to be expected for CP-odd or inverted-sign coupling hypotheses, or in intermediate regions of the parameter space. There, even a classifier trained specifically with the respective simulation is expected to be less efficient due to the stronger overlap in the phase space. This observation is consistent with the discussion in Ref. [48] and motivates the development of a classification and categorisation strategy that remains effective over a broad range of coupling assumptions by using dedicated decorrelation strategies.

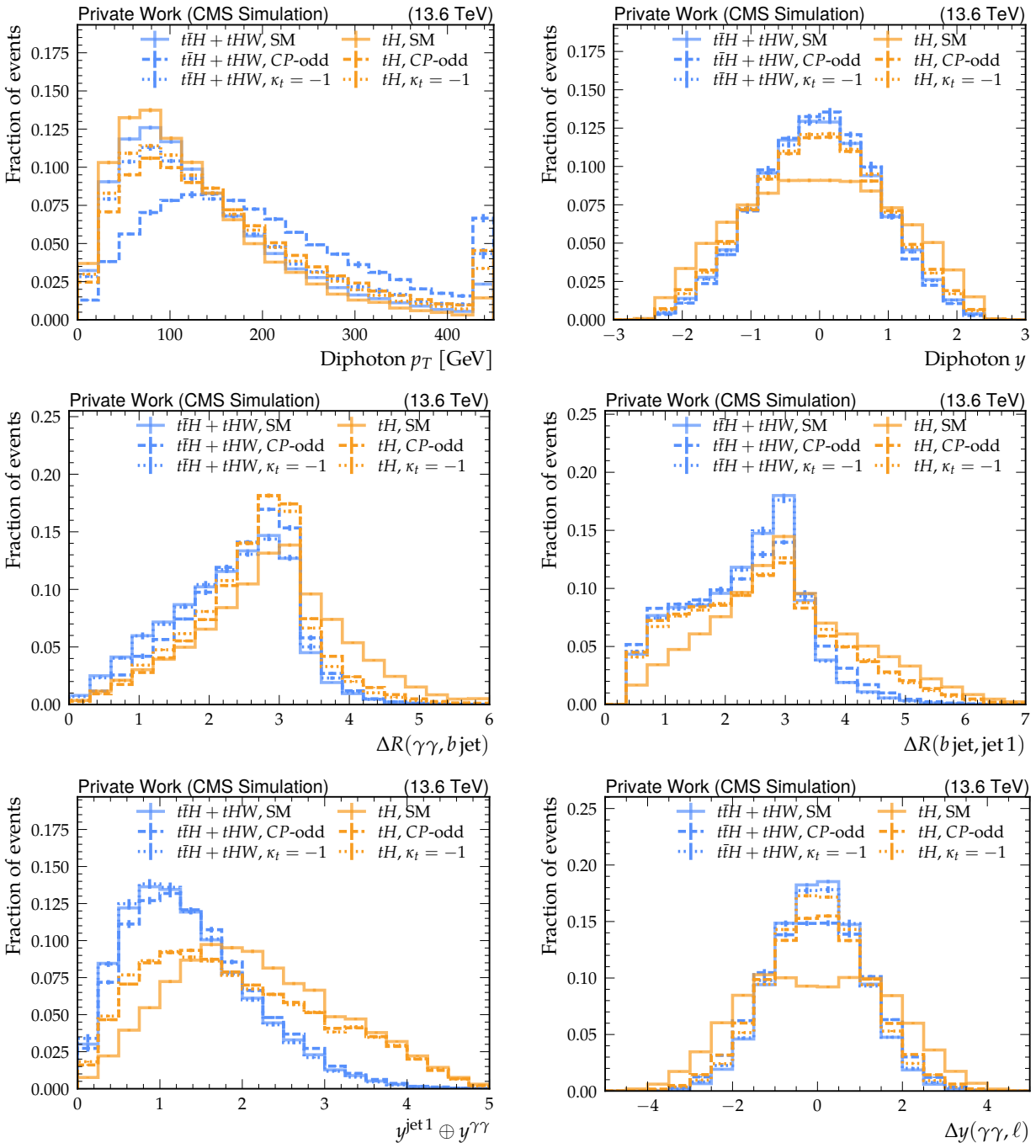


Figure 6.3: Kinematic observables with stronger dependence on the considered coupling scenarios. Events passing the hadronic selection are shown as an example, except for the distribution of the rapidity difference between the reconstructed diphoton system and the lepton, which is only defined for events passing the leptonic selection. The coupling dependence is indicated by comparing the SM (solid), CP-odd (dashed), and inverted-sign (dotted) coupling scenarios. Statistical modelling uncertainties are marked and possible overflow is included in the rightmost bin.

6.2. Background modelling for classifier training and categorisation optimisation

The event-level background model used for the optimisation studies is specified in the following. Its role is to provide a sufficiently realistic description of the continuum background for classifier training, validation, and categorisation optimisation.

6.2.1. Common modelling strategy

The dominant background processes in analyses in the $H \to \gamma\gamma$ decay channel are diphoton, γ+jet, and multijet production. Their relative contributions in an analysed phase space depend on the criteria placed on photon shower shape and isolation, which are designed to suppress non-prompt photon candidates. The modelling of these processes is particularly challenging in the phase space relevant for this analysis, which is characterised by the presence of two reconstructed photon candidates and a b-jet candidate. The available γ+jet and multijet simulation samples are produced only at LO in QCD. The diphoton sample also has LO accuracy for each included parton multiplicity, but contains matrix elements with up to three additional jets as well as the loop-induced $gg \to \gamma\gamma$ contribution, as described in Sec. 5.2.3. For the diphoton process, the LO cross section is known to underestimate the true rate substantially. It was shown that NLO and even NNLO corrections are crucial to describe the diphoton fiducial cross section [200–202], while perturbative convergence is found at N3LO [203]. Similarly, the predicted yield of the available γ+jet and multijet processes in the analysed diphoton phase space cannot be assumed to be fully correct, since both the cross sections and, in particular, the $j \to \gamma$ misidentification rate in simulation can induce discrepancies with the data. These main background processes are therefore normalised in a data-driven approach.
For background components with only one or no genuine photons, a severe limitation of the MC modelling is the very small efficiency for populating the selected phase space with two reconstructed photons that pass the preselection shower-shape and isolation requirements and the minimum photon ID score requirement of -0.7. This is particularly relevant for multijet, but also for $t\bar{t}$ production, where the signal-like phase space can be reached only if two jets are misidentified as photons. Since the cross sections of these processes are nevertheless much larger than the signal cross section, relevant contributions in the selected phase space are expected. To increase the effective statistical power of such simulation samples, the event processing is performed without applying the usual photon-quality preselection criteria with the HiggsDNA framework. In practice, this means that reconstructed photon candidates are retained without the standard requirements on the shower-shape and isolation observables and, most importantly, without the cut on the multivariate photon ID score. Only a loose baseline photon definition of $H/E < 0.15$ present in the NanoAOD input files remains in addition to the kinematic acceptance criteria based on p_{T} and η. This procedure substantially enlarges the available event samples, but it also removes the correlation between the hadronic misidentification rate and the object kinematics induced by the photon preselection.
In the later analysis, the photon shower-shape and isolation observables are no longer used directly; only the ID score and the information on the presence of a pixel seed are retained. The pixel-seed flag determined during the upstream reconstruction is retained unchanged. The photon ID scores, which accumulate strongly around -1 for $j \to \gamma$ misidentifications, need to be mapped back into the analysis phase space. To restore a realistic BDT response, the ID scores of hadronic photon candidates are reassigned from templates derived from $j \to \gamma$ misidentifi-

cations in the γ+jet simulation. Possible correlations between the retained pixel-seed flag and the reassigned photon ID score are neglected. While this procedure is used in both channels of the analysis, its details and impact are discussed in the hadronic subsection below, where they are most illustrative. The correction chain developed here is inspired by an approach applied in previous measurements with the full Run-2 dataset [26]. There, data events in which one photon candidate fails the minimum photon-ID criterion were used to replace the γ+jet and multijet simulation. As this sample is from data, it does not suffer from systematic biases in simulation. A possible drawback is overlap arising between this data-driven sample and simulation samples, which could induce relevant discrepancies. The approach developed here is expected to provide a more robust description when the data sample is comparatively small but the simulated datasets are large and the selection efficiency is the main bottleneck.

Overlap between simulation samples

An additional technical issue is the possible overlap between inclusive top-quark samples and photon-enriched samples. Although the inclusive $t\bar{t}$ and single-top quark samples do not contain photons at matrix-element level, isolated prompt photons can still be produced by QED radiation in the PS. These events can overlap with samples in which one or two photons are generated explicitly in the hard process. For the top-quark pair and single-top quark production processes, this overlap is treated using generator-level criteria. The processes $t\bar{t}$, $t\bar{t}\gamma$, and $t\bar{t}\gamma\gamma$, and analogously the corresponding single-top quark samples, are simulated with `MadGraph5_aMC@NLO`. In the photon-enriched samples, photons are included at matrix-element level only in the production process, not in the subsequently simulated top-quark decays. These generated photons are required to satisfy a dynamic smooth-cone isolation criterion as introduced in Ref. [204], which provides an infrared-safe definition of isolated prompt photons. The samples are then made exclusive according to the number of isolated prompt photons associated with the production process: inclusive $t\bar{t}$ and single-top quark events are vetoed if such a photon is present, while $t\bar{t}\gamma$ and $t\gamma$ events are vetoed if a second such photon is present. Prompt photons originating from the top-quark decay chain are retained, since they are not part of the overlapping phase space between the inclusive and photon-enriched production samples.
For the diphoton, γ+jet, and multijet samples, a corresponding generator-level veto is not applied, because these processes are produced with different generators and the associated isolation settings are not sufficiently harmonised to define a consistent prescription. Here, the event partition is instead based on the origin of the reconstructed photons. Events in the multijet sample are retained only if both reconstructed photon candidates are of $j \to \gamma$ origin, while for the γ+jet sample exactly one such misidentification is required.

6.2.2. Hadronic channel

The hadronic channel poses the main challenge for the optimisation studies because the background level is much higher than in the leptonic channel. A good background description is therefore a crucial ingredient for the analysis. Figure 6.4 illustrates the limitations of a baseline description obtained from the simulation samples without the specific corrections. The falling diphoton-mass spectrum is reproduced qualitatively, but visible deficits remain over the full diphoton mass range. The overall yield in data exceeds the simulation by a factor of approximately two. As a second example, the subleading photon ID score is shown, which shows a particular relative deficit in the low-score region. This indicates that the fractions of prompt and

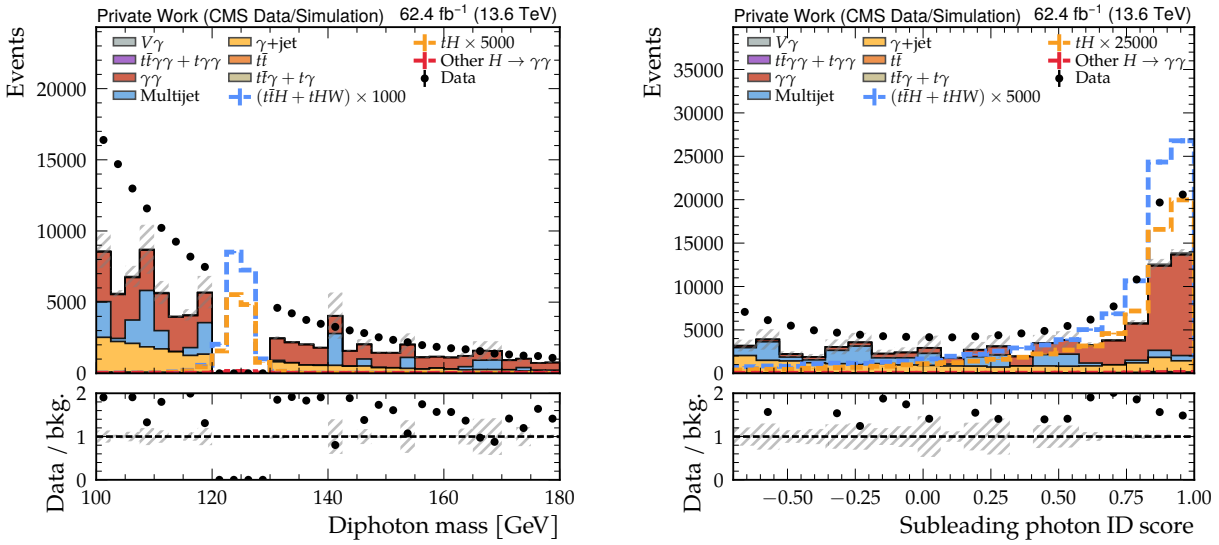
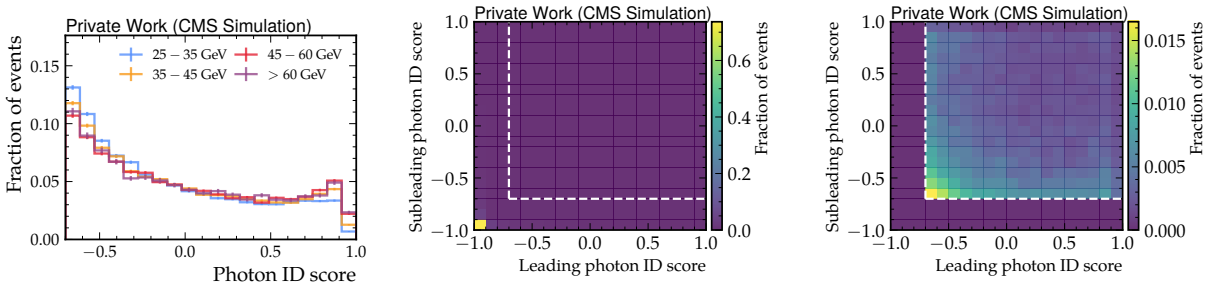


Figure 6.4: Representative distributions in the hadronic channel obtained from the simulated background samples before the dedicated modelling corrections are applied. The data and continuum background exclude the 120–130 GeV signal window. The $H \to \gamma\gamma$ processes are overlaid with the signal scaled up. Statistical uncertainties due to the limited number of simulated events are indicated.

non-prompt photons are not modelled correctly. The impact of the limited event counts of the $j \to \gamma$-dominated samples after the hadronic selection is particularly evident here and driven by the multijet sample. With the standard photon preselection applied, only 45 simulated multijet events remain, whereas the predicted weighted yield from this poorly modelled sample corresponds to about 15% of all data events and is thus non-negligible. With the relaxed photon requirements, approximately 16 500 multijet events pass the hadronic selection. In the same way, $t\bar{t}$ events with two $j \to \gamma$ misidentifications are treated and scaled to data as discussed below; these events are referred to in the following as $t\bar{t}(jj \to \gamma\gamma)$ events for brevity. This is done because without this contribution, a sizeable deficit of continuum events in the $t\bar{t}H$-like phase space of classifier outputs is found later on. Here, the originally predicted yield is small, as it amounts to less than 1% of the data. However, the simulated event count is increased from approximately 1 000 to 680 000. These large enhancements in the statistical power are crucial for the training of the DNN-based classifiers which operate in a high-dimensional feature space.

The templates used for the reassignment of photon ID scores are extracted separately for $j \to \gamma$ candidates with superclusters in the EB and EE. Four bins in p_T are used with boundaries of 25–35 GeV, 35–45 GeV, 45–60 GeV, and above 60 GeV. The resulting templates are illustrated in Fig. 6.5, together with the original and the morphed two-dimensional ID-score spectrum obtained from the multijet sample without photon-quality preselection. As expected for $j \to \gamma$ misidentifications, the ID score distributions have a sharp peak towards -1, which is the lower bound of the underlying BDT output domain. Only a small tail extends above -0.7. As shown by the γ+jet events, this tail is rather flat towards high scores and becomes more pronounced for large reconstructed photon p_T, where neutral-meson decays are least distinguishable from single prompt photons based on shower-shape criteria. Because of the sharp peak of the original distribution, most events populate the bin with the lowest possible scores in the two-dimensional



(a) $j \to \gamma$ ID score (EB) **(b)** Multijet before morphing **(c)** Multijet after morphing

Figure 6.5: Photon ID score distributions (left) obtained for $j \to \gamma$ misidentifications from the γ+jet MC sample. Reconstructed photons with their supercluster in the EB are shown as an example, as used to assign photon ID scores to events processed with relaxed preselection. The effect of this morphing procedure on the multijet event sample is shown in the middle and right panels.

score plane. After the morphing, by contrast, the events are spread over the full two-dimensional phase space of classifier outputs above -0.7, reflecting the wider tail of the distributions from which the values are drawn.

Using samples produced with relaxed photon requirements improves the statistical precision, but it also distorts the kinematic properties of the selected sample. To mimic the kinematic dependence of the photon-preselection efficiency, a four-dimensional reweighting is derived for the samples with the relaxed preselection as a function of the leading- and subleading-photon p_{T} and $|\eta|$, by comparing the deficit in the prediction with the dominant multijet yield. The weights are defined as

$$w_{jj\to\gamma\gamma}(\mathbf{x}) = \frac{N_{\mathrm{data}}(\mathbf{x}) - N_{\mathrm{other}}(\mathbf{x})}{N_{\mathrm{multijet}}(\mathbf{x})}, \qquad \mathbf{x} = (p_T^{\gamma_1}, p_T^{\gamma_2}, |\eta_{\gamma_1}|, |\eta_{\gamma_2}|). \tag{6.2}$$

Here, N_{data} denotes the yield observed in the data sidebands and N_{other} is the sum of all other background contributions in a given four-dimensional bin.

In a second step, the dominant diphoton, γ+jet, and multijet components are renormalised with inclusive scale factors. A weighted least-squares fit to the leading- and subleading-photon ID score distributions, which are sensitive to the mixture of prompt and non-prompt photons, is performed to obtain the normalisation factors. They are derived separately for the different data-taking periods to incorporate possible changes in the preselection efficiency between these processes. Such an effect is expected because of the higher pileup level in the 2023 data taking and the different noise thresholds in the calorimeters, which slightly affect the photon shower shapes and isolation variables. The resulting scaling factors for the diphoton process are 1.48 and 1.67 for 2022 and 2023, respectively, while the corresponding factors for the γ+jet events are 1.60 and 1.69. The multijet samples are rescaled by 1.59 and 2.12, although the initial normalisation is arbitrary due to the kinematic reweighting and the omitted photon preselection.

In a last step, the normalisation of the $t\bar{t}(jj \to \gamma\gamma)$ component is determined. The overall contribution of these events in data is assumed to be sufficiently small such that their normalisation can be determined after the main background components are fixed. The corresponding rate is constrained with a weighted least-squares fit to a distribution that is more sensitive to the presence of $t\bar{t}$ events, for which the b-tagging discriminant of the selected b jet is chosen. The

resulting fitted normalisation factor is 1.43. As for the multijet component, this value is defined relative to the arbitrary initial normalisation of the enhanced sample and therefore has no direct physical interpretation.
Example kinematic distributions of the mass–sideband data and the background model with the full correction chain applied are shown in Fig. 6.6. The correction chain brings the overall normalisation of the continuum model close to the sideband data and visibly improves the agreement for the diphoton-mass spectrum and the photon ID scores. The remaining shape agreement is also adequate for the kinematic variables most relevant for classifier training, such as the diphoton transverse momentum, $\Delta R(\gamma\gamma)$, the jet multiplicities, and the b-tagging discriminant. Quantitatively, the agreement between sideband data and the continuum model is at the 10%-level in the bulk of the distributions over most kinematic observables. Residual differences remain in the high-p_{T} tails of some distributions, for example for the diphoton system or the b jet, where this model predicts higher rates than observed in data. A possible improvement could be achieved by performing this correction chain in more granular regions of phase space, for example in several bins of the diphoton p_{T}. This is not pursued here because of the limited number of events in data and simulation, but it will likely become useful once more data-taking periods and the associated simulation are included in the analysis.
In the following, the discriminative observables between the signals and the background model are analysed. Multiple variables play a crucial role. For example, the backgrounds tend to have much lower jet multiplicity than the signals. While the central jet multiplicity peaks at two for the background, the highest signal densities are found for three central jets for tH and five central jets for $t\bar{t}H$, where the background density is reduced by factors of approximately two and 25, respectively. Another strongly discriminating variable is the b jet b-tagging score. Most background events tend towards lower values, which shows that the selected b-jet candidates mostly originate from light jets or c jets. For the signal, by contrast, jets actually initiated by b quarks are present such that the distribution peaks sharply towards one. Processes with non-prompt photons can be identified with the photon ID score, since events with one or two misidentified photon candidates tend to have at least one lower score. Events from the continuum background also differ from the signal in kinematic photon properties. The continuum background and the other $H \to \gamma\gamma$ processes often have a separation between the photons around $\Delta R(\gamma\gamma) = \pi$, while the typically more boosted signal tends to produce photons with smaller separation and higher p_{T}. These kinematic differences between signal and background are therefore important inputs to the neural-network classifiers developed in this thesis.

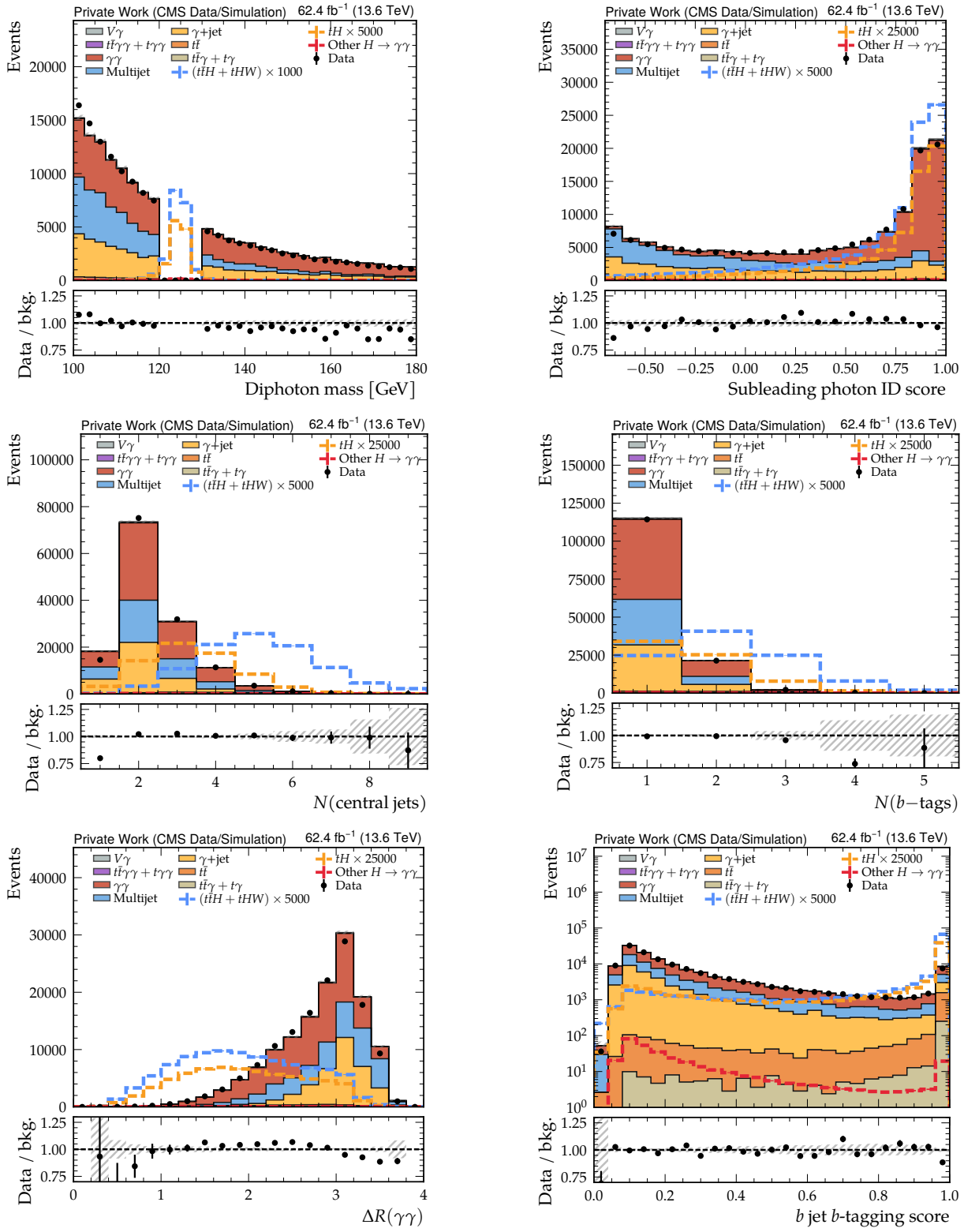


Figure 6.6: Representative photon and diphoton observables in the hadronic channel after applying the dedicated background corrections used for the optimisation studies. The data and continuum background exclude the 120–130 GeV signal window. Statistical uncertainties due to the limited number of simulated events are indicated, and overflow is included in the rightmost bin.

6.2.3. Leptonic channel

In the leptonic channel, the number of events passing the selection is substantially lower in both simulation and data. While 162 125 data events pass the hadronic selection, only 898 events are present in the leptonic channel. In simulation, several processes yield vanishing or negligible event numbers, such that the corresponding event-level model is not used for classifier training or categorisation optimisation in order to avoid unstable behaviour driven by very large statistical uncertainties. In particular, only 26 γ+jet and zero multijet events pass the selection. This is expected, because these processes require a misidentified lepton in addition to misidentified photon candidates and therefore have tiny selection efficiencies. For the γ+jet process, however, the predicted yield is not negligible and corresponds to approximately 7% of the data. Because of the small number of events passing the selection, this process is not included in the final simulated background model in order to reduce the statistical uncertainty. For simulated diphoton events, the efficiency is large enough that around 4 000 events pass the selection. Because of the lower statistical power in the leptonic channel, the diphoton normalisation obtained in the hadronic channel is applied; a possible bias in the lepton-misidentification rate is therefore not corrected separately here.
The background composition also differs qualitatively from that in the hadronic channel. An important background contribution in this channel is DY production in association with a photon, with an electron-positron pair in the final state; $Z/\gamma^* + \gamma \to e^+e^-\gamma$. These events can enter the selected phase space if one $e \to \gamma$ misidentification is present, which is a non-negligible effect given the large DY cross section and the fact that the $e \to \gamma$ misidentification rate is much higher than the $j \to \gamma$ rate. That these events are likely under-predicted is indicated by the kinematic distributions before corrections are applied, shown in Fig. 6.7. The distributions of the invariant masses $m(\ell, \gamma_1)$ and $m(\ell, \gamma_2)$ exhibit the characteristic structure around the Z-boson mass expected from $Z \to e^+e^-$ events in which one electron is reconstructed as a photon, where it is visible that the data exceed the prediction in this region by a factor of approximately two. At the same time, the fraction of photon candidates with a reconstructed pixel seed is higher in data than in the uncorrected simulation, which also points to an $e \to \gamma$ misidentification rate that is too low in the simulation. Furthermore, the fraction of events including top-quark pair production is substantial in this channel, because here leptons and b jets are abundant and therefore, the selection efficiency is comparably high.

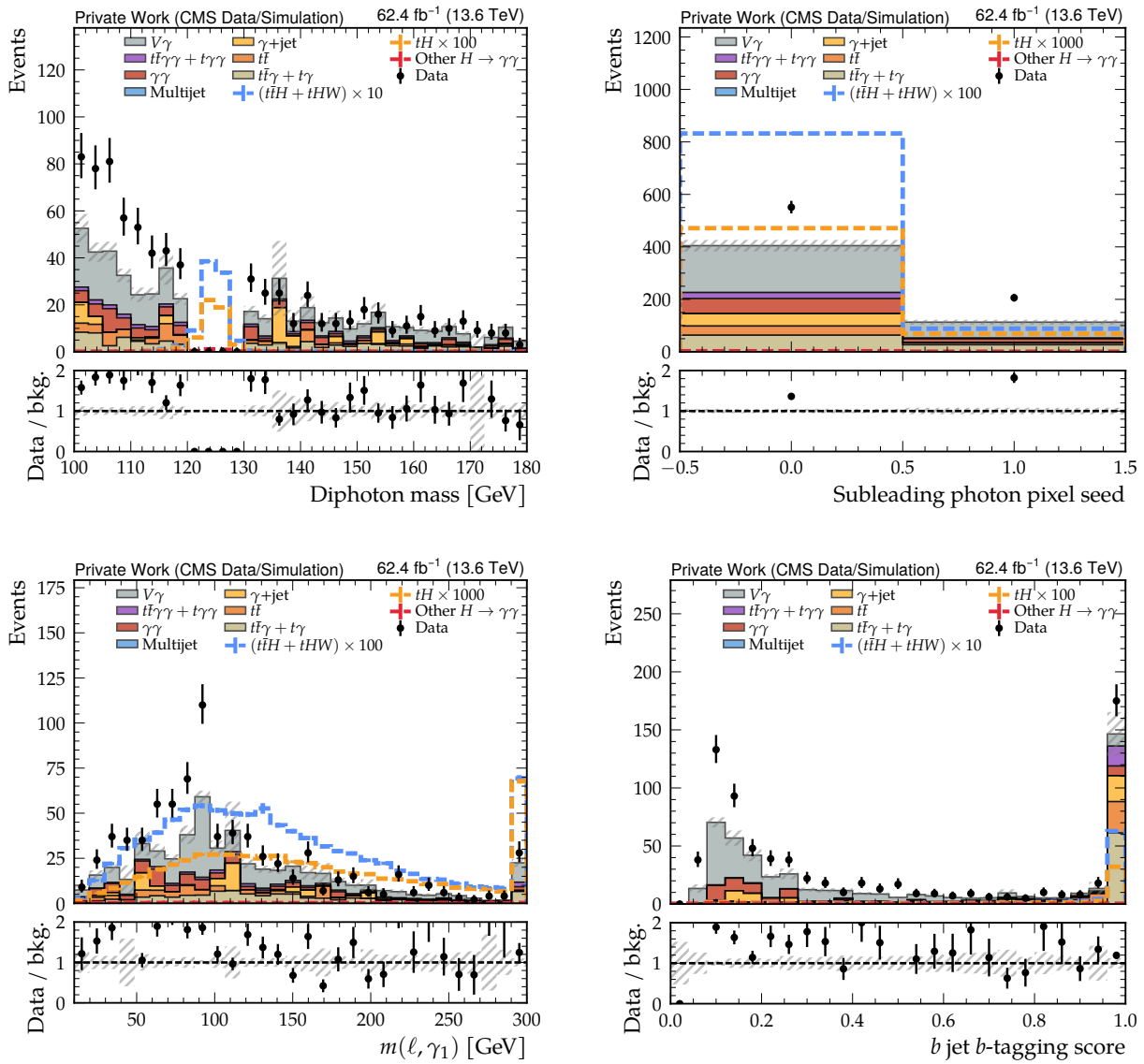


Figure 6.7: Example observables in the leptonic channel showing the simulated background model before the specific corrections developed in this thesis are applied. The data and continuum background exclude the 120–130 GeV signal window. Statistical uncertainties due to the limited number of simulated events are indicated, and overflow is included in the rightmost bin.

The correction strategy in the leptonic channel focuses on improving the modelling of the DY+γ and $t\bar{t}(jj \to \gamma\gamma)$ contributions. The former is regarded as the leading source of events with $e \to \gamma$ misidentifications and is normalised with a data-driven scale factor. A fully process-independent treatment of $e \to \gamma$ misidentification would require classification of events based on the origin of the reconstructed photon candidates across all background components. In the present work, however, the background model is constructed at the level of individual simulated processes, and the corresponding correction is therefore applied to the dominant DY+γ contribution only. For the $t\bar{t}(jj \to \gamma\gamma)$ modelling, the statistical uncertainty is reduced with the same enhancement strategy as in the hadronic channel by relaxing the photon-quality requirements. These two contributions are then normalised simultaneously using the b jet b-tagging discriminant, which is sensitive to their relative yield. The normalisation obtained for the DY+γ process is 2.14. The corresponding fitted normalisation factor for the $t\bar{t}(jj \to \gamma\gamma)$ component is 0.29. As in the hadronic channel, this value is defined relative to the arbitrary initial normalisation of the enhanced sample and therefore has no direct physical interpretation. Figure 6.8 shows example observables in the leptonic channel after applying all corrections. A much better description of variables sensitive to the fraction of DY+γ events is observed than in the plain simulation. The description of the most important variables for the discrimination between signal and background, such as the jet multiplicity and the b-tagging discriminant, is satisfactory within the statistical precision of the available data and simulation. For most kinematic observables, the agreement between data and the continuum model is at the 25%-level in the bulk of the distributions. Some discrepancies nevertheless remain; for example, the photon ID scores indicate that the fraction of $j \to \gamma$ misidentifications is still under-predicted in this phase space. This again suggests that processes such as γ+jet or multijet production are not entirely negligible here. The omission of the small γ+jet event sample is nevertheless preferable despite the possible residual contribution, because no statistically stable event-level model can be constructed from only 26 events and the remaining impact is judged to be subleading for the optimisation. Overall, the modelling with these corrections is also assessed to be sufficient in the leptonic channel for classifier training and categorisation optimisation.

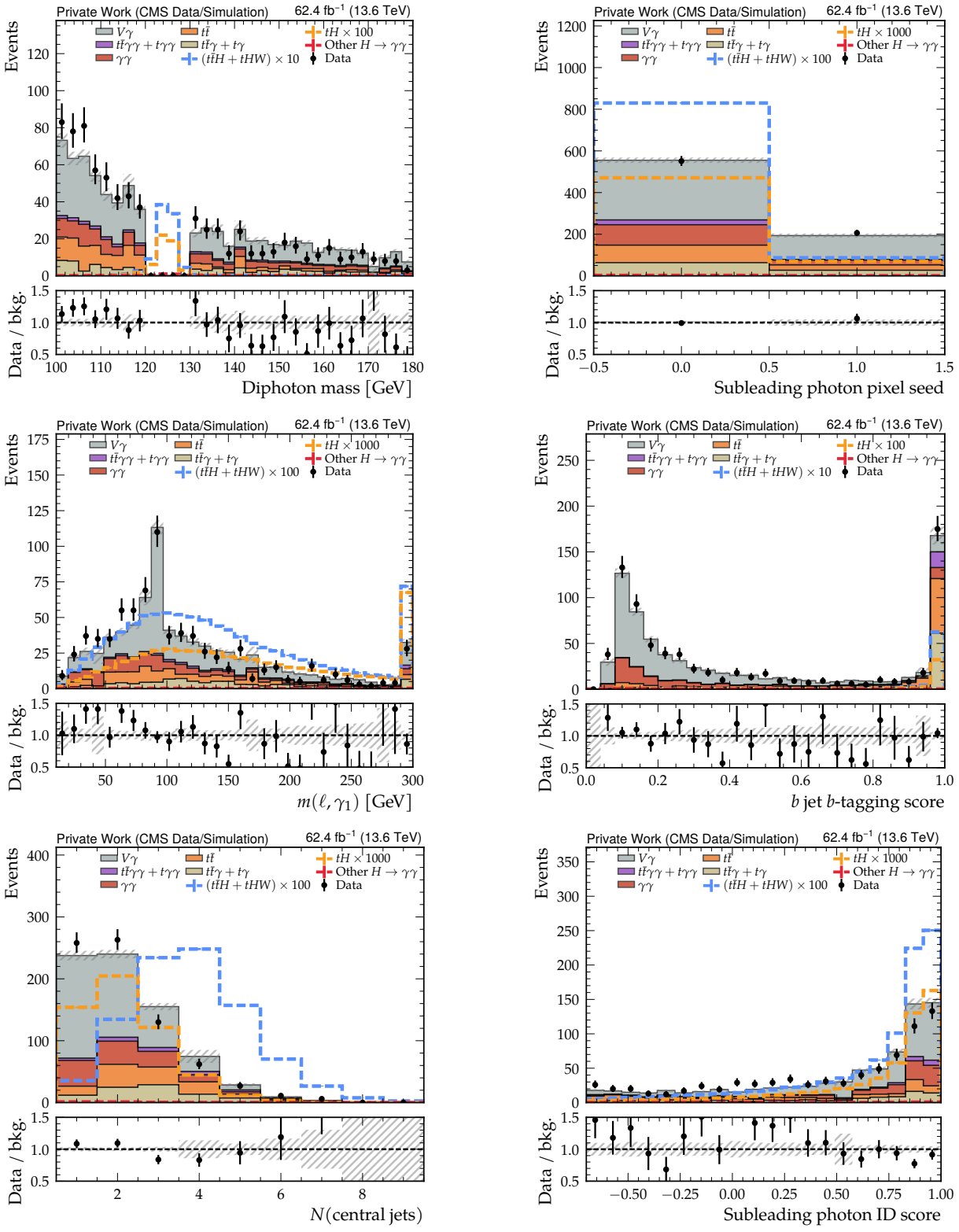


Figure 6.8: Example observables in the leptonic channel with the specific corrections developed in this thesis applied. The data and continuum background exclude the 120–130 GeV signal window. Statistical uncertainties due to the limited number of simulated events are indicated, and overflow is included in the rightmost bin. Overall, a reasonable description of the kinematic properties of the data is achieved.

7. Classification of events

The selected event sample of the $t\bar{t}H$ and tH cross-section measurement remains strongly background-dominated after the reconstruction and selection steps described in Chap. 4. At the same time, the two signal processes populate overlapping but distinct regions of phase space. A multivariate treatment is therefore useful to exploit the correlated kinematic information of the reconstructed photons, leptons, jets, and missing transverse momentum more efficiently than a purely cut-based strategy. For this purpose, DNNs are developed to separate $t\bar{t}H + tHW$ from tH events and to distinguish each signal from the combined background processes. Their output scores define the discriminant space used for the categorisation optimisation in Chap. 8. Because the measured cross sections should depend as little as possible on assumptions about the top-quark Yukawa coupling, the classifier-output distributions are required to remain stable under variations of the strength and CP structure of the interaction.
Before presenting the classifiers used in the analysis, this chapter first recalls the machine-learning concepts most relevant for the following discussion. It then briefly summarises a published study on Kolmogorov–Arnold networks (KANs) [205], which investigated an alternative neural-network architecture for the binary separation of $t\bar{t}H$ and tH events and represented the first application of KANs in high-energy physics. The main conclusion relevant for this thesis is already stated here. KANs can match the performance of conventional multilayer perceptrons (MLPs) in this task, but they do not provide a sufficiently strong practical advantage to replace the MLP classifiers used in the analysis.

7.1. Brief overview of machine-learning concepts

This section gives only a short orientation on the concepts most relevant for this chapter; for a comprehensive physics-oriented overview, the reader is referred to Ref. [206]. Following the classical definition of Mitchell [207], a program is said to learn

> *from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.*

In the supervised machine-learning setup used here, the experience E is represented by a training dataset of simulated events with known process labels. The task T is the event classification according to these process labels, and the performance measure P can be understood as the classification quality on statistically independent events. In practice, this amounts to learning a function that maps reconstructed observables to a discriminative score. For an MLP, this function is built from layers of nodes with trainable weights and biases. Non-linear activation functions are applied between layers, allowing the network to model correlations that a linear classifier could not describe. The weights and biases are adjusted during training with gradient-descent-based optimisation, for which the required gradients are computed by error backpropagation. This requires a differentiable objective function. Since target measures such as classification accuracy are not always convenient differentiable objectives, they are replaced by a differentiable

surrogate loss. For binary classification tasks in this chapter, the classification loss is the binary cross-entropy (BCE). For target labels y_i and predicted probabilities $\hat{y}_i$ of N events, it is defined as

$$\mathrm{BCE}(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right]. \tag{7.1}$$

During training, the loss obtained on the training sample is minimised by adapting the trainable parameters, while the precise parameter update is determined by the chosen optimiser. Before observables are used as network inputs, they are usually transformed to comparable numerical scales. Throughout this chapter, studentisation refers to subtracting the mean value of each feature computed on the training sample and dividing by the corresponding standard deviation. The same transformation, with the mean and standard deviation fixed from the training sample, is then applied to the validation and test samples.
The trained classifier should generalise to statistically independent events rather than learn statistical noise or sample-specific details of the training dataset. The difference between the performance on the training sample and on statistically independent validation or test samples is the generalisation gap. A large generalisation gap can indicate overtraining, where the classifier has adapted to sample-specific fluctuations and therefore performs worse on independent events. To monitor this behaviour, the available events are separated into training, validation, and test samples. The training sample is used for the optimisation, the validation sample for hyperparameter choices and performance monitoring during the training, and the independent test sample for final performance and overtraining checks. This separation avoids tuning directly on the sample used to quote the final classifier performance. A standard countermeasure against overtraining is regularisation. For the L_2 regularisation used in this thesis, the optimisation objective is augmented by a penalty proportional to $\lambda_{\mathrm{L2}} \|\theta\|_2^2$, where θ denotes the trainable model parameters and λ_{L2} controls the regularisation strength.

7.2. Study on Kolmogorov–Arnold networks

This section summarises the study published in Ref. [205]. The text below is summarised from that publication in a condensed form, with emphasis on the aspects that motivate the classifier choice made in this thesis.

7.2.1. Motivation

Conventional MLPs are a standard tool for event classification in collider physics and are motivated by the universal approximation theorem [208]. Recently, KANs have been proposed as an alternative architecture motivated by the Kolmogorov–Arnold representation theorem [209]. In the formulation that motivates the architecture, a continuous multivariate function can be written as

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{2n+1} \phi_i \left(\sum_{j=1}^{n} \phi_{ij}(x_j) \right), \tag{7.2}$$

i.e. as a composition of sums and univariate functions [209]. This motivates a neural-network architecture replacing the fixed activation functions on the nodes in MLPs by learnable univariate functions on the edges [210], followed by simple summation on nodes. In practice, the KAN implementation used in the study parameterises these edge functions by spline-based activation functions. This makes KANs attractive in principle for two reasons: they can provide a compact

model description and, at least for small architectures, the learned univariate transformations can be inspected directly or even represented with analytical functions.

Ref. [210] demonstrates striking improvements over MLPs for several low-dimensional function-approximation benchmark problems. The architecture also appears highly parameter efficient in that setting compared to MLPs. It is therefore an important question for the high-energy physics community whether these advantages persist for realistic collider datasets with substantially larger input dimensionality and more complex correlation structures. The goal of the study published in Ref. [205] is to test this in a realistic collider event-classification problem. For this purpose, KANs are benchmarked against conventional MLPs in the binary separation of $t\bar{t}H$ and tH production in the $H \to \gamma\gamma$ final state. The task is closely related to the measurement presented in this thesis, since the separation of these two processes is an important ingredient of the later categorisation.

7.2.2. Simulated dataset and event selection

The datasets used in Ref. [205] were generated specifically for that study and are therefore independent of the samples of the thesis analysis, which are produced within the CMS Collaboration. Events for $t\bar{t}H$ and tH production at $\sqrt{s} = 14\,\text{TeV}$ are generated at LO accuracy in perturbative QCD with `MadGraph5_aMC@NLO` [183] using the `NNPDF23_lo_as_0130_qed` PDFs [211]. The tH sample contains the dominant t-channel process together with the charge-conjugate contribution and is simulated in the four-flavour scheme following Ref. [30], while $t\bar{t}H$ is simulated in the five-flavour scheme. The events are interfaced to `Pythia` 8.3 for the $H \to \gamma\gamma$ decay, PS, and hadronisation [84], and to `Delphes` 3.5 with a CMS detector card for the fast detector simulation [212]. Jets are reconstructed with the anti-k_T algorithm with radius parameter $R = 0.5$ [141, 142]. Only final states with at least one semi-leptonically decaying top quark are generated, including decays via τ leptons.

A simple analysis preselection is applied to define a realistic event-classification problem, similar to the selection for the leptonic channel of the signal region in this analysis as described in Sec. 4.3. Events are required to contain exactly two photons with transverse momenta above 35 and 25 GeV, respectively, with $|\eta| < 2.5$, and with a diphoton invariant mass in the range $100 < m_{\gamma\gamma} < 180\,\text{GeV}$. In addition, the ratios of the leading and subleading photon transverse momenta to the diphoton mass are required to exceed $1/3$ and $1/4$, respectively. Furthermore, at least one electron or muon with $p_\text{T} > 10\,\text{GeV}$ and $|\eta| < 2.4$ is required. The selection also requires at least one b-tagged jet with $p_\text{T} > 25\,\text{GeV}$ and $|\eta| < 2.5$, and at least one additional jet with $p_\text{T} > 25\,\text{GeV}$ and $|\eta| < 4.7$. After this selection, 100 000 events per process are used for training, 33 000 for validation, and an independent test sample of 100 000 events per process is reserved for a precise final performance comparison.

The classifiers are trained on 22 reconstructed input observables. These include photon, lepton, and jet kinematics, missing transverse momentum, jet and lepton multiplicities, and high-level observables tailored to the topologies. Since representative feature distributions are already discussed in the previous chapter, they are not repeated here. More relevant for the KAN–MLP comparison is the non-trivial correlation structure of the inputs, shown in Figure 7.1. The correlations differ significantly between $t\bar{t}H$ and tH events, illustrating that the separation task is genuinely multivariate and cannot be reduced to a few almost independent one-dimensional inputs.

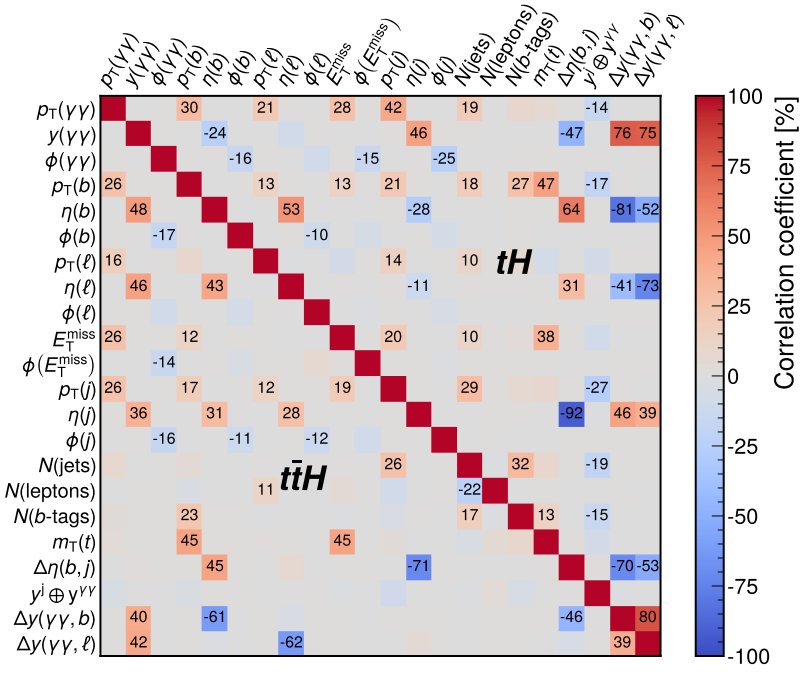


Figure 7.1: Matrix of Pearson correlation coefficients for the 22 input features used in the KAN study. The lower triangle corresponds to $t\bar{t}H$ events and the upper triangle to tH events. This figure is taken from Ref. [205].

7.2.3. Training strategy and network configurations

For the KANs, the `pykan` implementation provided together with Ref. [210] is used, based on `PyTorch` [213]. Several architectures are compared, ranging from a single-layer KAN with node structure 22–1 to multi-layer models such as 22–3–1, 22–45–1, 22–10–5–2–1 and 22–45–10–5–2–1. The default spline settings of the package are retained for this comparison, corresponding to cubic splines and a grid parameter of $G = 5$. All models are trained with binary cross-entropy and early stopping. In addition to the Adam optimiser [214], the L-BFGS optimiser [215] advocated in the original KAN paper is tested, but no relevant performance improvement is observed in this application.

A dedicated preprocessing is required for the KANs because the spline activations depend explicitly on the chosen input domain. Typical collider observables contain long tails and outliers, which can stretch the spline range far beyond the densely populated region of phase space. To mitigate this effect, transverse momenta and the transverse top-quark mass are first mapped with $\tilde{x} = \ln(1 + x)$ and all input features are then scaled to the interval $[0, 1]$. For the MLP baselines, the studentisation procedure introduced above is sufficient. This difference is already instructive from a practical point of view: while KANs are trainable in a realistic high-energy physics setup, they require more careful handling of preprocessing and domain updates than standard MLPs. In particular, the domains of the hidden layers of KANs need to be updated iteratively during training to match the output range of the layers upstream. This additional procedure leads to noticeable instabilities and is not required for MLPs.

For the comparison to KANs, a set of shallow and deep MLPs is trained on the same dataset, from a small single-hidden-layer network with eight nodes up to a five-hidden-layer model

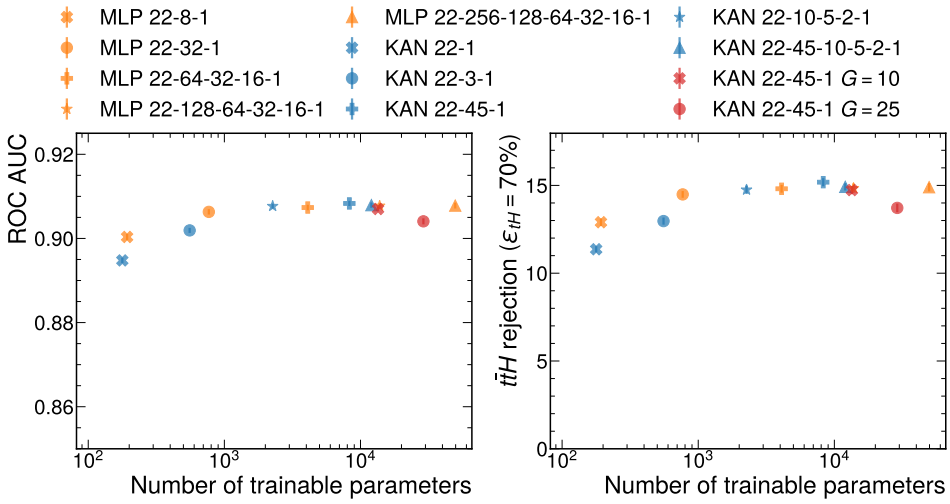


Figure 7.2: Performance comparison between KANs and MLPs in the $t\bar{t}H$–tH classification task. Comparison of the ROC AUC and of the $t\bar{t}H$ rejection at fixed tH efficiency as a function of the number of trainable parameters. This figure is taken from Ref. [205].

with node structure 22–256–128–64–32–16–1, implemented in `TensorFlow` 2.17 [216, 217]. The hidden layers use rectified linear unit (ReLU) activation functions and the output nodes use sigmoid activations. This setup spans models with roughly two hundred to about sixty thousand trainable parameters and allows a direct comparison of performance and parameter efficiency.

7.2.4. Results and interpretation

The main result of Ref. [205] is that KANs with at least one hidden layer achieve a classification performance very similar to that of well-tuned MLPs. The single-layer KAN is visibly weaker, which is consistent with its limited expressiveness. Once a hidden layer is added, however, the gap closes quickly. The best KAN considered in the study, the 22–45–1 architecture, reaches a receiver operating characteristic (ROC) area under the curve (AUC) of 0.908 on the independent test set, essentially identical to the best MLP. Even a small MLP with only one hidden layer yields a ROC AUC of 0.906, such that the remaining numerical differences are minor in the context of a full physics analysis.

This behaviour is also reflected in the parameter-efficiency comparison shown in Figure 7.2. For very small models, the MLPs are more parameter efficient: the smallest MLP already achieves strong background rejection with substantially fewer trainable parameters than the smallest competitive KAN. For intermediate and large model sizes, the two architectures reach similar performance. The best KAN is numerically slightly better than the best MLP for the specific classification task considered. However, this comparison is based only on classification metrics evaluated on simulated test events. Such a small difference is unlikely to translate into a relevant improvement in the final cross-section measurement after event categorisation and the likelihood fit, whose precision is largely limited by data statistics and is additionally affected by systematic uncertainties. The study therefore does not reveal a compelling performance argument for replacing MLPs by KANs in the thesis analysis.

One distinctive observation of Ref. [205] concerns the single-layer KAN. This architecture differs from the two-step composition suggested by Eq. (7.2), and such single-layer architectures are not

explored in Ref. [210]. With only one layer, each input variable is transformed through a single learned univariate spline, and the network output is given by the sum of these one-dimensional transformations followed by the final sigmoid. In this case, the study shows that the learned spline shapes closely resemble the one-dimensional log-likelihood ratios of the respective features, as illustrated in Figure 7.3. This is precisely the behaviour expected for a binary classifier with additive univariate structure, and it provides a rare example in which the learned representation of high-dimensional data can be interpreted directly in physics terms.
The comparison shown in Figure 7.3 also makes clear why this result is special to the one-layer network by comparing to splines learned by a 22–3–1 KAN. Once a hidden layer is added, multiple splines act on the same input and are combined non-trivially in subsequent layers. The learned functions then no longer resemble simple one-dimensional likelihood ratios and the transparent interpretation is largely lost. Already the competitive 22–45–1 KAN contains more than one thousand learned activation functions and is difficult to inspect meaningfully by eye. In the regime where KANs become competitive with MLPs in terms of classification performance, their practical interpretability is therefore not guaranteed. For the purpose of the present thesis, this weakens the main argument in favour of KANs.
The study in Ref. [205] constitutes the first application of KANs in high-energy physics and shows that this architecture is viable for collider event classification. Nevertheless, the subsequent analysis in this thesis relies on conventional MLPs rather than KANs. This choice is driven by practical considerations: MLPs achieve essentially the same performance for the $t\bar{t}H$–tH classification, which serves here as a proxy for a typical task in high-energy physics, and they are simpler to train and maintain. KANs remain interesting as interpretable small models and as a direction for future studies, but they do not offer a decisive advantage for the analysis strategy pursued here.

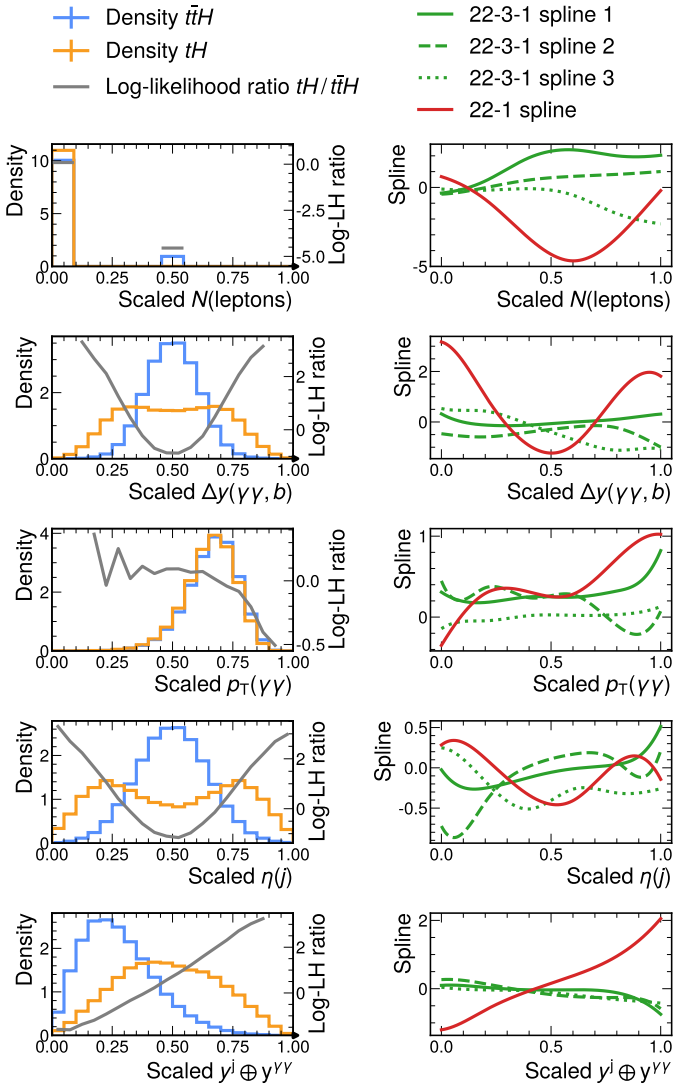


Figure 7.3: Comparison of learned KAN splines with one-dimensional likelihood-ratio information for selected input features. The left panels show the feature distributions for tH and $t\bar{t}H$ together with the corresponding one-dimensional log-likelihood ratios. The right panels compare these ratios to the learned spline of the single-layer KAN and to the first-layer splines of the 22–3–1 KAN. This figure is taken from Ref. [205].

7.3. Datasets and preprocessing for the cross-section measurement

Following the analysis strategy motivated in Chap. 4, the classifier training does not rely solely on the SM signal hypothesis. To this end, the training dataset is expanded by signal samples corresponding to the *CP*-odd and the inverted-sign coupling scenarios. These scenarios are intentionally chosen to be rather extreme: although they are already in tension with the current Run-2 constraints, they provide a conservative basis for enforcing a classifier response that remains stable under smaller deviations from the SM expectation. This differs from previous cross-section measurements of top-quark associated Higgs boson production, which typically employed classifiers and event categories based on the SM signal model [28, 35, 39, 40, 159, 218–220]. Dedicated studies of the *CP* structure of the top-quark Yukawa coupling introduced additional *CP*-sensitive discriminants, but not for the purpose of defining the classifiers used for a cross-section measurement [26, 27].

The numbers of simulated events per process available after the analysis preselection are summarised in Table 7.1. As expected from the different branching fractions and acceptance effects, the hadronic-channel selection retains substantially larger event samples than the leptonic selection. Processes with event counts smaller than 500 after the preselection are not included in the training. All datasets are split reproducibly into training, validation, and test subsets using fractions of 60%, 20%, and 20%, respectively.

A broad set of reconstructed observables is considered as potential input to the classifiers. These include kinematic properties of the diphoton system, individual photons, the leading lepton, the selected b-tagged jet, and up to five further jets. For all these reconstructed objects, the kinematic quantities p_{T}, η, and ϕ are considered; for jets, the reconstructed mass is included in addition. The missing transverse momentum and its azimuthal angle are also used. In addition, event-level summary quantities such as jet, b-tag, and lepton multiplicities, invariant masses between pairs of jets, pairs of leptons and photons, and angular separations ΔR of relevant object combinations are used. The initial feature set is intentionally chosen generously and is then reduced using the feature-importance studies described below. In this way, the final networks retain the full classification performance to within the percent level while avoiding an unnecessarily large input space.

Particular care is taken to avoid classifier inputs that could induce artificial structures in the diphoton invariant-mass distribution. The later statistical inference models the continuum background from the shape of the $m_{\gamma\gamma}$ distribution in data, separately in each analysis category. The mass spectra in different categories may therefore have different smoothly falling shapes, but the classifiers should not reject continuum background by directly exploiting the reconstructed diphoton mass. Otherwise, the categorisation could create signal-like features in the mass spectrum and thereby bias the final measurement. For this purpose, the transverse momenta of the diphoton system and of the individual photons are expressed relative to the diphoton mass as $p_{\mathrm{T}}^{\gamma\gamma}/m_{\gamma\gamma}$ and $p_{\mathrm{T}}^{\gamma_i}/m_{\gamma\gamma}$. This substantially reduces the possibility that the network learns residual diphoton-mass information from the inputs.

Before the variables are passed to the MLPs, the studentisation procedure introduced in Sec. 7.1 is applied. In practice, it improves the conditioning of the optimisation problem and leads to visibly more stable training and faster convergence of the networks.

Process	N_{lep}	N_{had}	Used (lep)	Used (had)
$t\bar{t}H$ (SM)	1 334 000	2 892 000	yes	yes
$t\bar{t}H$ (CP-odd)	766 000	1 713 000	yes	yes
tHq (SM)	692 000	2 072 000	yes	yes
tHq (CP-odd)	767 000	2 354 000	yes	yes
tHq ($\kappa_t = -1$)	176 000	476 000	yes	yes
tHW (SM & reweights)	1 021 000	2 078 000	yes	yes
ggH	370	285 000	no	yes
VBF	304	189 000	no	yes
VH	33 000	191 000	yes	yes
$b\bar{b}H$	461	59 000	no	yes
$\gamma\gamma$	3 977	1 752 000	yes	yes
γ+jet	24	14 000	no	yes
Multijet	18	16 000	no	yes
$t\bar{t}\gamma\gamma + tq\gamma\gamma$	51 000	160 000	yes	yes
$t\bar{t}\gamma + tq\gamma$	4 629	20 000	yes	yes
$t\bar{t}$	165 000	720 000	yes	yes
tq	2	31	no	no
$V\gamma$	16 000	31 000	yes	yes

Table 7.1: Event counts per process and channel after the analysis preselection, summed over training, validation, and test subsamples. Counts above 10^4 are rounded to the nearest thousand, while smaller counts are quoted exactly. The last two columns indicate whether the corresponding process is used in the signal–background trainings in the leptonic and hadronic channels.

Weighting of events during training

For all trainings, event weights are chosen such that the classifier learns the relevant physics problem rather than being influenced by the available MC sample composition. In the signal–background classifiers, a special treatment is required for the continuum background. Restricting the training to a narrow Higgs boson mass window would lead to an unnecessarily small number of continuum events. Instead, the non-resonant background is used over the full selected range $100\,\text{GeV} < m_{\gamma\gamma} < 180\,\text{GeV}$. It is then down-weighted such that its total effective yield corresponds to that expected in the central part of the signal window, 123–127 GeV, which contains the majority of the signal events. This rescaling is obtained from an exponential fit to the continuum spectrum and corresponds to a weight factor of approximately 0.06 for continuum events. In this way, the available statistical power is retained while the training remains representative of the phase space relevant for the final measurement.

Another technical complication arises from negative event weights in the `MadGraph5_aMC@NLO` samples used for training. If such weights are used directly in the binary cross-entropy loss, the loss is no longer bounded from below and the optimisation can exploit this in an unphysical way. To avoid this instability in the training loss, the absolute value of each event weight is taken and the resulting weights are renormalised per simulated process such that the total expected yield remains unchanged. This procedure was found to stabilise the optimisation without altering the overall class composition. It is only a prescription for classifier training; the physics normalisation of the simulated samples in the later statistical analysis is not changed by this treatment.

Within the signal definition, tHW is treated as part of the $t\bar{t}H$-like class. The relative contribution of $t\bar{t}H$ and tHW is set according to the product of production cross section, acceptance, and selection efficiency for the respective coupling hypothesis. For trainings of classifiers aimed to be agnostic to the coupling state, the three coupling scenarios are then normalised such that the

total signal weight is identical for each coupling scenario within a given process class. Finally, the two classes of the binary problem, namely either $t\bar{t}H+tHW$ versus tH or signal versus background, are normalised to equal total weight. This prevents the optimisation from developing a bias towards the more abundant class.

7.4. Neural-network setup for coupling-robust classification

The KAN study in Section 7.2 already indicated that the MLP performance saturates for comparatively modest model sizes, and no relevant gain was observed when significantly larger MLPs were tested on that dataset. All classifiers used for the measurement are implemented as MLPs with four hidden layers of 128, 64, 32, and 16 nodes, respectively, followed by one output node normalised with the sigmoid function. This architecture was found to be sufficient in capacity for all classification tasks considered here, as deeper and wider models did not yield a significant performance improvement and only tend to overtrain more quickly. Using a single common architecture throughout the study also facilitates a consistent comparison between tasks and channels, and lowers the dimensionality of the remaining hyperparameter space to be explored.

The hidden layers use ReLU activation functions, while the network parameters are optimised with the Adam optimiser using a learning rate of 10^{-4}. A comparatively large batch size of 1 024 is chosen, because each batch must contain a sufficient number of events from the different coupling hypotheses for the correlation metrics to be estimated reliably. Early stopping with a patience of 25 epochs is employed, and the best model parameters according to a stopping criterion are restored after training. Overtraining and a significant generalisation error are mitigated by the L_2 regularisation introduced above and are subsequently checked explicitly, such that the training samples can later be incorporated in the statistical analysis without introducing a relevant bias. The regularisation strength is tuned separately for the different channels.

The classifier response should be as insensitive as possible to the underlying Yukawa-coupling hypothesis. This problem is closely related to the broader machine-learning task of learning predictive functions that are robust against nuisance parameters or protected attributes. Several approaches to this problem have been proposed in the literature. The adversarial method from Ref. [221] enforces approximate independence by introducing a secondary network that attempts to recover the nuisance information from the classifier output. The adversary is optimised to predict the nuisance label as well as possible, while the classifier is optimised to keep the task-relevant information and to make the nuisance prediction difficult. This competition is implemented by adding the adversarial loss to the classifier objective with the opposite sign. If the procedure is successful, the adversary can no longer infer the nuisance information from the classifier output. Penalty-based methods instead augment the classifier loss by differentiable measures of dependence, with distance correlation (DisCo) providing one particularly convenient choice in high-energy physics applications [222]. DisCo estimates general, including non-linear, statistical dependence from samples. Beyond these two approaches, invertible transformations based on conditional normalising flows [223] and more recently optimal-transport constructions [224] have also been explored. In the present thesis, the adversarial and DisCo methods are studied and compared. A more extensive comparison of both approaches for the $t\bar{t}H$–tH classification problem was carried out in the Bachelor's thesis in Ref. [225] during the development of this analysis; the central result relevant here is reproduced below.

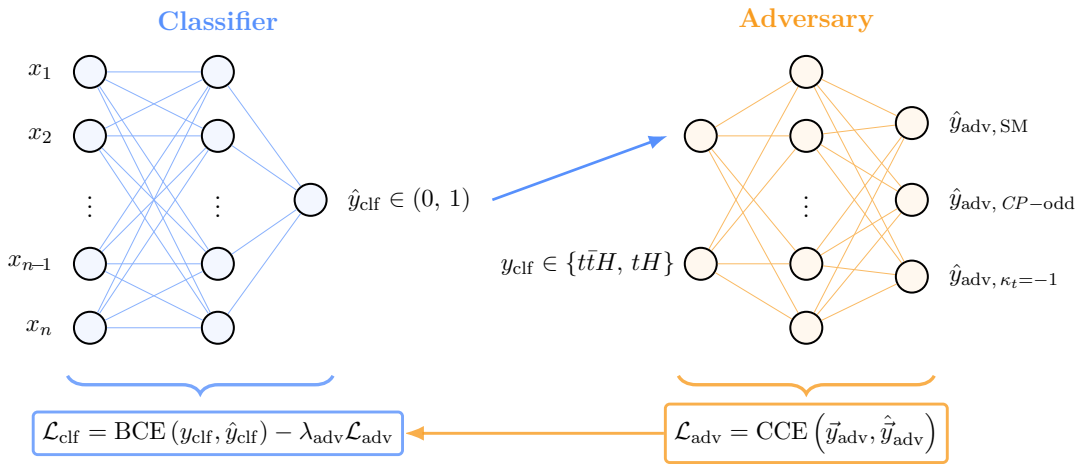


Figure 7.4: Schematic illustration of the adversarial training prescription used for the optimisation of event classifiers with the corresponding loss functions. Event-level observables $\vec{x}$ are processed by a binary classifier trained to separate two analysis classes, corresponding to either $t\bar{t}H{+}tHW$ versus tH, $t\bar{t}H{+}tHW$ versus background, or tH versus background. The classifier output is passed to the adversary together with an auxiliary signal-process label, such that the adversary attempts to infer the underlying Yukawa-coupling hypothesis of the signal events.

In the implementation used in this thesis, the adversary predicts three classes for signal events, referring to the coupling scenarios used. The corresponding training prescription is illustrated schematically in Figure 7.4. For the $t\bar{t}H$–tH classifier, all events belong to one of the two signal classes and therefore carry such a coupling label. For the signal–background classifiers, background events do not have a top-quark Yukawa-coupling label; the adversarial loss is therefore evaluated on signal events only, while the binary classifier loss is evaluated on both signal and background events. A signal-process label indicating whether an event is part of the $t\bar{t}H{+}tHW$ or the tH class is provided as an auxiliary input to the adversary. Using the binary class label y_{clf} and classifier prediction $\hat{y}_{\text{clf}}$, the classifier objective in the adversarial setup is

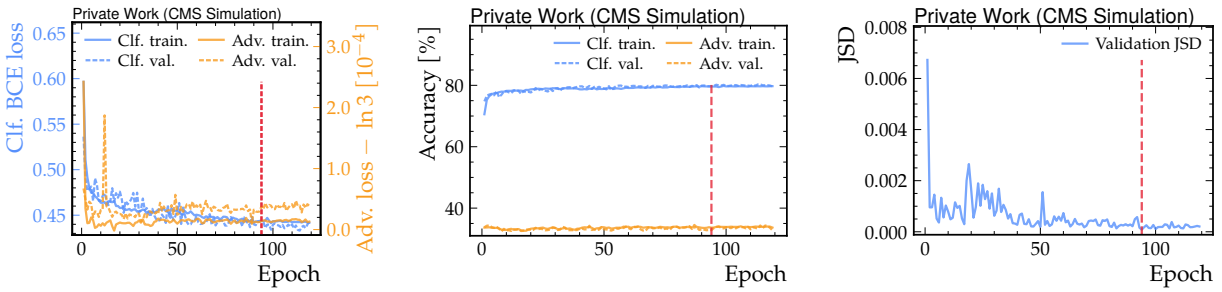
$$\mathcal{L}_{\text{clf}} = \text{BCE}(y_{\text{clf}}, \hat{y}_{\text{clf}}) - \lambda_{\text{adv}}\,\text{CCE}(\vec{y}_{\text{adv}}, \hat{\vec{y}}_{\text{adv}}) + \lambda_{\text{L2}}\,\|\theta_{\text{clf}}\|_2^2, \tag{7.3}$$

where the categorical cross-entropy (CCE) loss of the adversary's class labels and predictions is given by $\text{CCE}(\vec{y}_{\text{adv}}, \hat{\vec{y}}_{\text{adv}})$. The parameter λ_{adv} controls the trade-off between discrimination power and coupling robustness achieved by the decorrelation.

The adversary is also implemented as an MLP with ReLU activations in the hidden layers and a softmax output over the coupling labels. Ref. [225] documents that using deep adversaries with multiple hidden layers leads to a more stable training and a stronger decorrelation than using a single-layer adversary. For all adversaries, the same architecture with six hidden layers of 512, 256, 128, 64, 32 and 16 nodes is used in the training of the final classifiers. The optimisation is performed with alternating updates, using two gradient steps for the adversary followed by one gradient step for the classifier on separately iterated sets of mini-batches. As a reference for each task, an additional classifier is trained on the SM samples only using the BCE loss.

7.5. $t\bar{t}H$–tH classification and comparison of decorrelation methods

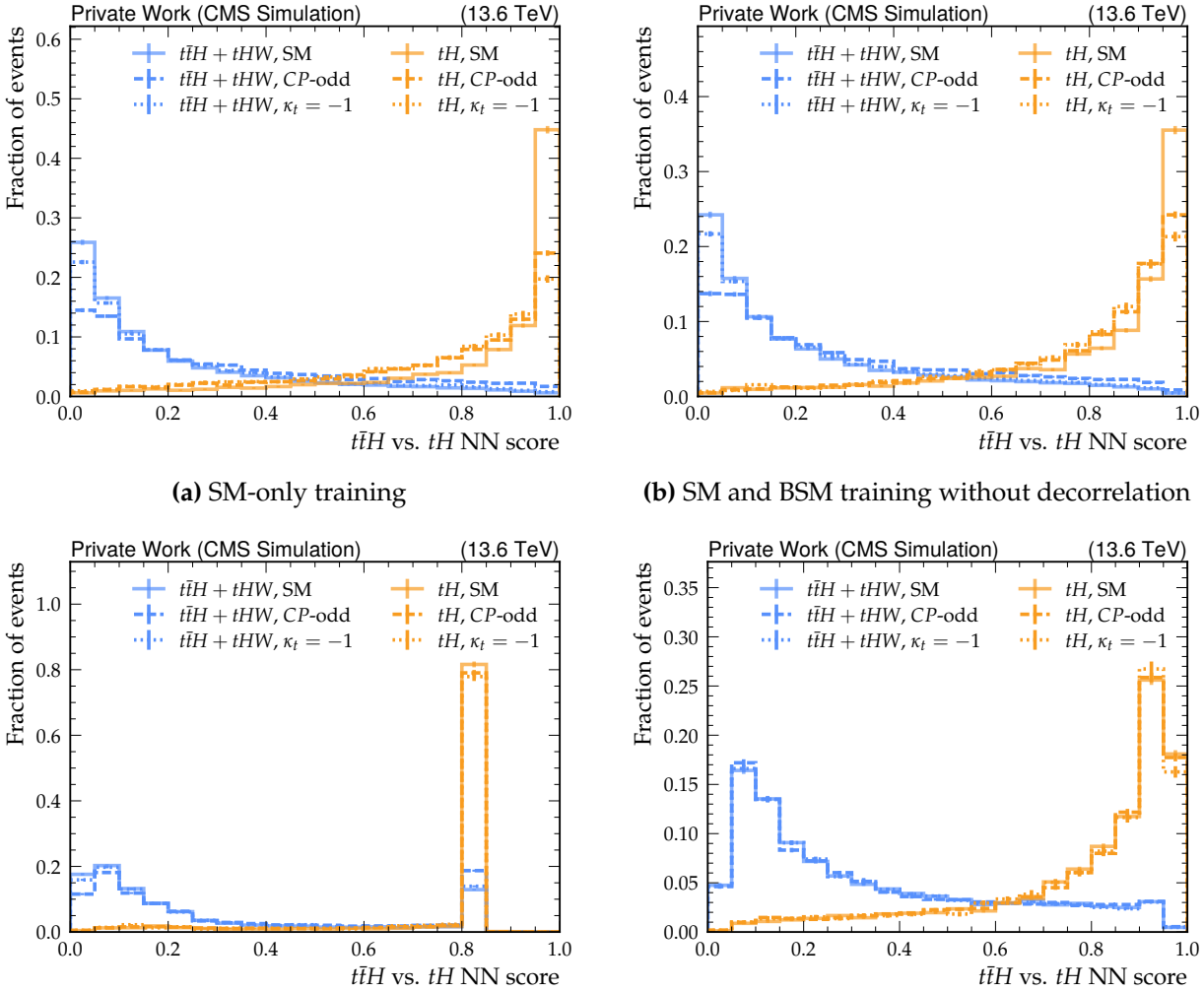
The separation of $t\bar{t}H{+}tHW$ from tH events is used as the central reference task for developing and comparing the decorrelation strategy. In this task, both classes correspond to genuine Higgs boson production with the decay into two prompt photons, such that individual photon-ID and quality observables are not useful for the classification. Accordingly, the individual photon observables are omitted and only the kinematic properties of the reconstructed diphoton system are included. However, the diphoton mass is excluded, which is reserved for the final statistical inference. The dominant discriminating information instead arises from the different event topologies of the two production modes. Building on the kinematic discussion in Chap. 6, tH production is characterised by the t-channel topology with more forward activity, while $t\bar{t}H$ events tend to exhibit a larger number of central jets. For this reason, separate counts of central and forward jets are provided to the classifier.
The training evolution of the hadronic adversarial $t\bar{t}H{+}tHW$ versus tH classifier is shown in Figure 7.5. The classifier binary cross-entropy decreases smoothly, while the adversarial loss remains close to $\ln 3$, which is the value expected for random guessing in a balanced three-class problem. For readability, only the small deviation from this limiting value is displayed. At the same time, the classifier accuracy improves substantially, whereas the adversary accuracy remains close to $1/3$. On its own, this could suggest that the adversary stops learning after only a few epochs. However, the validation Jensen–Shannon divergence (JSD) between the SM and CP-odd classifier-output distributions continues to decrease up to approximately the 80th epoch, showing that the residual coupling dependence of the classifier output is still being reduced. The final stopping point is therefore determined from the geometric mean of the validation classifier loss and the validation JSD, rather than from the loss alone. This criterion was found to provide a useful compromise between discrimination power and coupling robustness. The inverted-sign scenario is not included in the JSD calculation here, because the number of training examples in this scenario is limited and the JSD is therefore not estimated reliably for this scenario.
To compare adversarial decorrelation with the DisCo approach, several DNNs were trained with both methods while scanning the strength of the decorrelation term over multiple orders of magnitude. The purpose of this scan is twofold. First, it determines the range in which the optimisation remains numerically stable. Second, it identifies the region in which an improved decorrelation is still achieved without an excessive degradation of the classification performance. In practice, unstable or strongly distorted trainings are observed for $\lambda_{\text{DisCo}} \gtrsim 1$ and for $\lambda_{\text{adv}} \gtrsim 10^4$.



(a) Losses **(b)** Accuracies **(c)** JSD

Figure 7.5: Evolution of the training and validation metrics for the hadronic adversarial $t\bar{t}H{+}tHW$ versus tH classifier. The left panel shows the classifier and adversarial losses, the middle panel the corresponding classifier and adversarial accuracies, and the right panel the validation JSD obtained from comparing the SM and CP-odd output distributions. The vertical red line marks the epoch selected for the final model from the geometric mean of the classifier loss and the JSD obtained from the validation dataset.

The resulting output distributions are compared in Figure 7.6. A network trained only on the SM samples provides the best separation for exactly that hypothesis, but at the price of a pronounced model dependence. In the most tH-like region, for example, the efficiency above a classifier output of 0.95 is approximately 0.45 for SM tH events, but only about 0.25 and 0.20 for the CP-odd and $\kappa_t = -1$ scenarios, respectively. Likewise, the efficiency of $t\bar{t}H$ events shows a pronounced relative increase by more than a factor of three for the CP-odd hypothesis in that region. A similar effect is observed in the $t\bar{t}H$-like region of the output score. Training on all three coupling scenarios without an explicit decorrelation term improves the response for the BSM samples only mildly; the output shapes remain visibly different. The DisCo-based training reduces these differences to some extent, especially in the tH distributions, but still leaves clear distortions in the $t\bar{t}H$ class between the coupling scenarios. In addition, the output range becomes visibly compressed. This behaviour is consistent with the broader difficulty of balancing classification performance and DisCo regularisation. Unstable or otherwise undesirable solutions have been reported when the DisCo-penalty strength becomes too strong [222, 224, 226]. The adversarially decorrelated network instead aligns the output distributions of the three coupling hypotheses much more closely over the full score range. Its performance on the SM hypothesis is somewhat reduced compared with the SM-only training, but it remains much more stable when the coupling hypothesis is changed.



(a) SM-only training

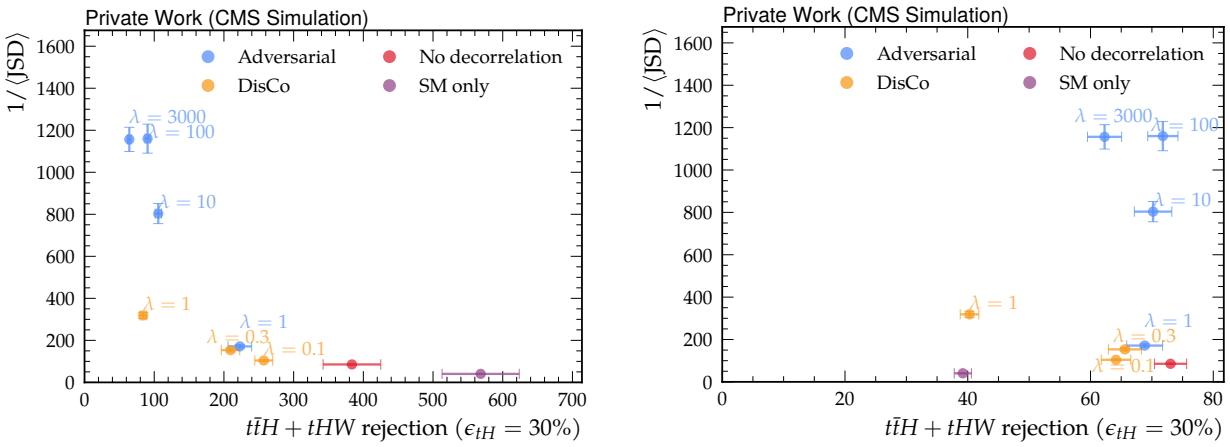
(b) SM and BSM training without decorrelation

(c) DisCo decorrelation

(d) Adversarial decorrelation

Figure 7.6: Classifier-output distributions for the hadronic $t\bar{t}H$–tH task under different training prescriptions. The panels compare a training using SM samples only, a training with all coupling hypotheses without explicit decorrelation, a DisCo-regularised training, and the adversarially decorrelated network used in the analysis.

This qualitative behaviour is quantified in Figure 7.7, which summarises the trade-off between classification performance and residual coupling dependence. As a performance metric, the rejection of the $t\bar{t}H{+}tHW$ class at fixed tH efficiency of 30% is shown. The degree of decorrelation is measured by the inverse of the averaged JSD obtained from comparing the SM output distribution with the corresponding BSM distributions. For the SM-only reference network, the rejection in the SM scenario reaches approximately 560, but this value collapses to roughly 40 when evaluated under the CP-odd hypothesis. The network trained on all coupling hypotheses without explicit decorrelation still exhibits a pronounced mismatch between SM and BSM output distributions, with the corresponding rejection values reaching approximately 380 and 75, respectively. The DisCo-regularised models improve the JSD only moderately and do not achieve a regime in which the performance on the three coupling hypotheses becomes consistent. A strong decorrelation of $\lambda_{\text{DisCo}} = 1$ leads to a substantial reduction of the performance on the SM scenario. The BSM performance is only mildly improved in this case, such that the SM and CP-odd rejections are approximately 80 and 40, respectively. The adversarial models instead populate a substantially more favourable region of the plane. At a given rejection, they achieve markedly higher decorrelation levels as quantified by $1/\text{JSD}$, and for sufficiently strong decorrelation the performance obtained on the SM sample is closely reproduced for the BSM scenarios. For the strongest tested decorrelation strength where stable training was still possible, $\lambda_{\text{adv}} = 3\,000$, the rejection obtained from the SM sample is approximately 65 and closely matches the rejection of the CP-odd case. The tails of the respective score distributions are also visibly aligned between the coupling scenarios, which is a relevant property for the present measurement and constitutes the main reason for adopting adversarial decorrelation for all classifiers used in the measurement. The same qualitative conclusion was also found in the leptonic channel and in the signal–background classifications, consistent with the more detailed comparison provided in Ref. [225].



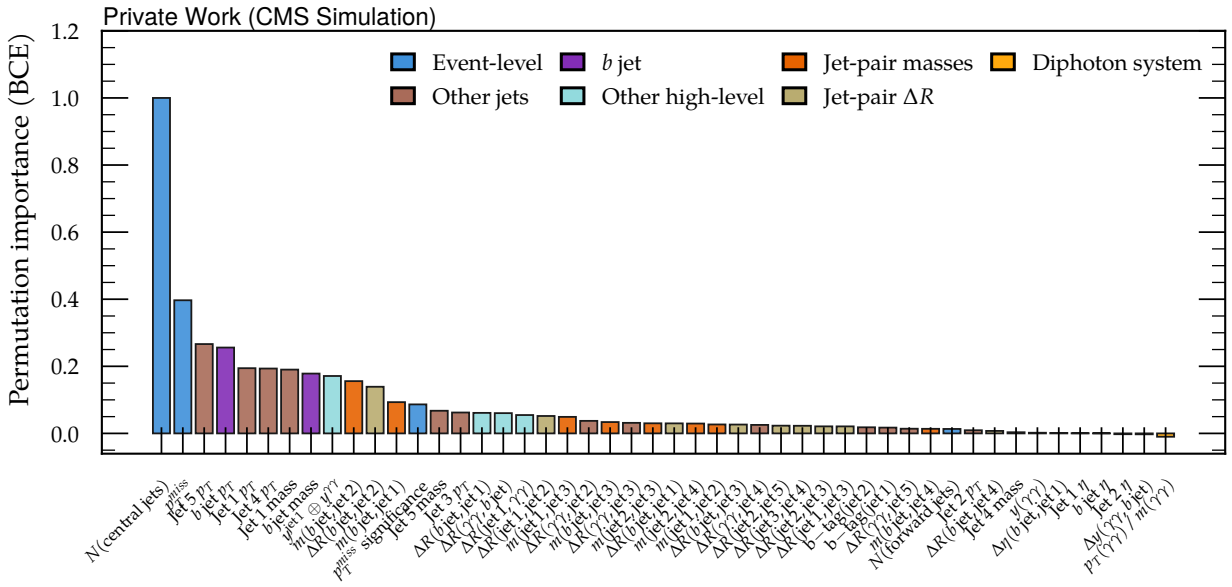
(a) Performance on the SM hypothesis

(b) Performance on the CP-odd hypothesis

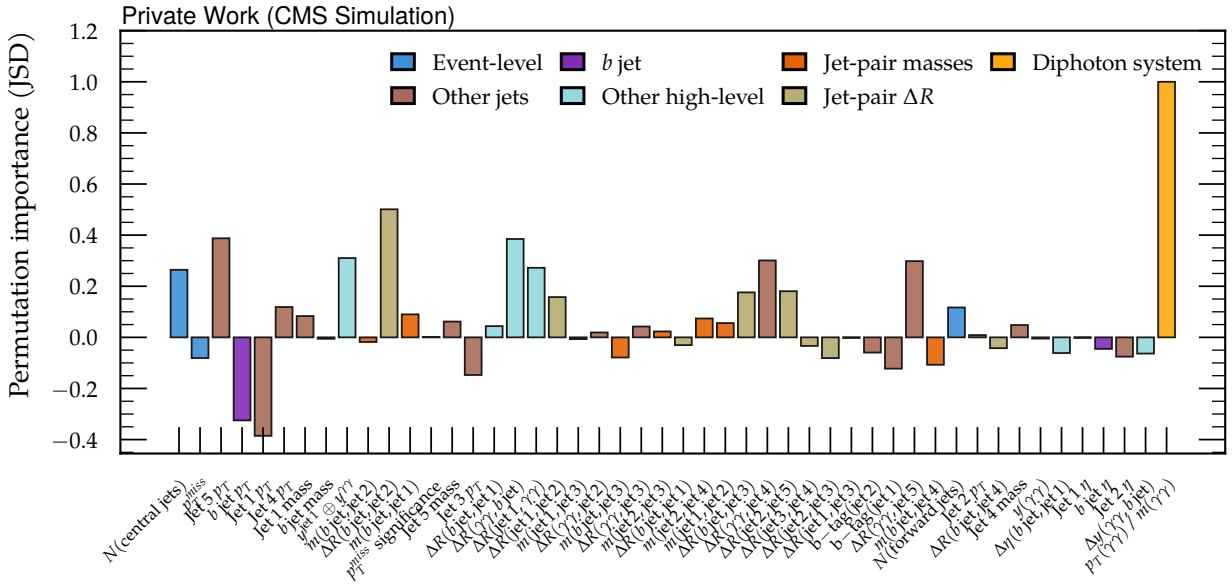
Figure 7.7: Comparison of classification performance and residual coupling dependence obtained from different $t\bar{t}H{+}tHW$–tH training setups. The horizontal axis shows the rejection of the $t\bar{t}H{+}tHW$ class at fixed tH efficiency of 30%, the vertical axis shows the inverse of the averaged JSD between the output distributions. The decorrelation penalty strengths λ are indicated. The error bars are obtained from bootstrap resampling and indicate the statistical uncertainty.

Feature engineering and input reduction

A reduction of the initially large input feature set is desirable for several reasons. Smaller input spaces lead to faster evaluation, reduce the risk of learning unstable fluctuations in weakly constrained variables, and make the classifier behaviour easier to interpret. In addition, the feature-importance analysis provides a useful partial view into the otherwise largely opaque behaviour of the DNNs.
The reduction procedure starts by training a classifier on the full feature set, and the relevance of each feature is evaluated using permutation feature importance [227]. For a given variable, the corresponding values are permuted across the events, thereby removing its event-by-event association with the class label and the other input variables while preserving its overall distribution. The trained network is then re-evaluated on this modified sample and the change in a chosen validation metric is recorded. In the present case, this is done both for the BCE and for the average pairwise JSD between the SM samples and both additional coupling hypotheses, denoted $\langle \mathrm{JSD} \rangle$. The former probes the classification performance, while the latter probes the residual coupling dependence. Variables whose permutation changes neither of these objectives appreciably are discarded.
Figure 7.8 shows the resulting feature-importance values of the hadronic $t\bar{t}H$–tH classifier after removing the least useful features. The most important variables with respect to the cross-entropy are the number of central jets, the transverse momenta of jets, and the missing transverse momentum. The importance of the kinematic properties of the jets is partly correlated with the overall jet multiplicity. This is consistent with the expected topology difference between the two processes.
An interesting observation in this study is that some variables exhibit negative permutation importance with respect to the classification metric, and the respective features may vary between different trained models. This does not imply that these variables are irrelevant. Instead, the same variables typically show a large positive importance with respect to the JSD, indicating that they carry precisely the coupling-sensitive information that the decorrelation is intended to suppress. Hence, this observation reflects the fact that the network is trained under a simultaneous discrimination and decorrelation objective. Variables with a strong sensitivity to the coupling hypothesis can improve the class separation, but may worsen the robustness of the classifier output. When such a variable is permuted, the classifier may therefore become even slightly better at separating the classes while at the same time becoming less stable across coupling scenarios. This behaviour is observed in particular for variables such as $p_{\mathrm{T}}^{\gamma\gamma} / m_{\gamma\gamma}$ and rapidity differences involving the diphoton system. These are the most coupling-sensitive observables, as discussed in Sec. 6.1.



(a) Impact on the binary cross-entropy



(b) Impact on the coupling dependence measured by the JSD

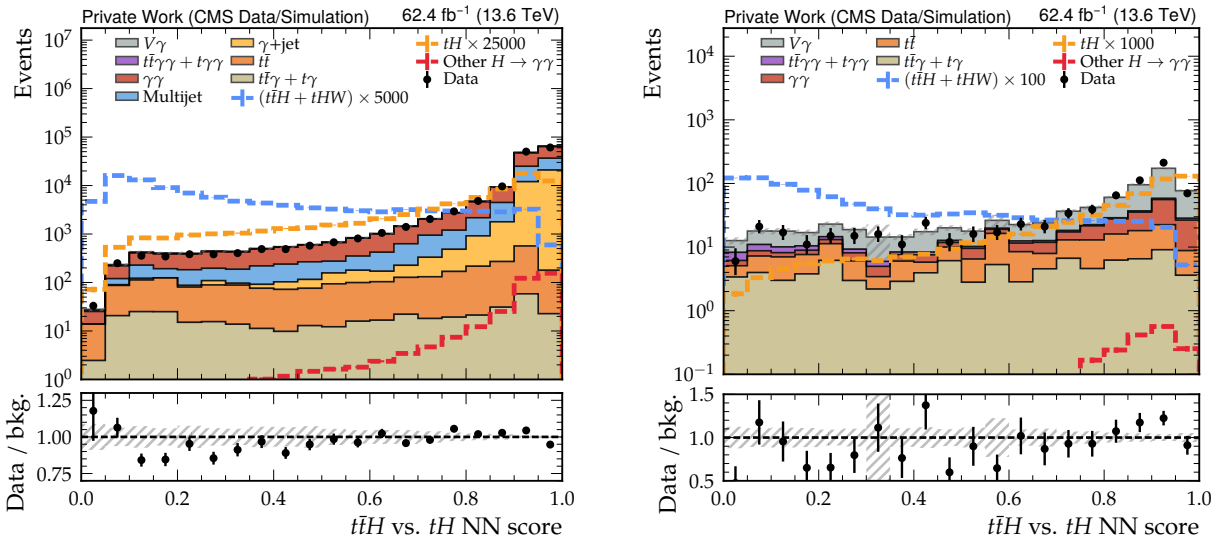
Figure 7.8: Permutation feature importance for the hadronic $t\bar{t}H$–tH classifier. The upper panel shows the change in the validation BCE when a given feature is permuted, while the lower panel shows the corresponding change in the validation JSD between the SM and CP-odd output distributions.

To summarise the impact of adversarial decorrelation beyond the hadronic example discussed above, the corresponding ROC AUC values for both channels are collected in Tab. 7.2. In the hadronic channel, the SM-only classifier achieves the highest ROC AUC on the SM hypothesis, but its performance degrades strongly when evaluated on the CP-odd and inverted-sign hypotheses. After adversarial decorrelation, the absolute ROC AUC on the SM sample is reduced, but the values become much more closely aligned across the three coupling scenarios. The same qualitative pattern is observed in the leptonic channel.

Channel	Training setup	SM	CP-odd	$\kappa_t = -1$
Hadronic	SM-only	0.938	0.857	0.891
Hadronic	Adversarially decorrelated	0.886	0.879	0.882
Leptonic	SM-only	0.953	0.892	0.917
Leptonic	Adversarially decorrelated	0.914	0.904	0.907

Table 7.2: ROC AUC for the $t\bar{t}H$–tH classifiers in the hadronic and leptonic channels, evaluated for the three coupling hypotheses. For both channels, the adversarial training reduces the dependence of the classifier performance on the assumed coupling hypothesis, at the price of a moderate reduction in the ROC AUC obtained on the SM hypothesis.

Beyond the performance metrics discussed above, the classifier-score distributions used in the analysis must be compared between simulation and data in the selected phase space used later for categorisation. The corresponding distributions, including data from the diphoton-mass sidebands, are shown in Fig. 7.9. In both channels, the data and simulated background distributions show sufficient agreement for the later categorisation procedure. At the same time, the majority of background events populate the tH-like region of the classifier output. This is most pronounced in the hadronic channel, where the background density increases over three orders of magnitude from the $t\bar{t}H$-like to the tH-like region. This underlines the central challenge of the analysis: the tH-enriched hadronic phase space is precisely where the continuum and resonant backgrounds are most abundant, making an efficient signal–background suppression essential for the final sensitivity.



(a) Hadronic channel

(b) Leptonic channel

Figure 7.9: Output distributions of the $t\bar{t}H$–tH classifiers used in the analysis in the hadronic and leptonic channels. Data and continuum backgrounds are shown from the diphoton-mass sidebands excluding the 120–130 GeV signal region, while the $H \to \gamma\gamma$ processes are overlaid from simulation. Uncertainties from limited MC event numbers are indicated.

7.6. Signal–background classification

Having established the DNN architecture and decorrelation strategy on the challenging $t\bar{t}H$–tH problem, the same methodology is applied to the signal–background classification tasks. For each channel, two additional binary classifiers are trained: one for tH versus background and one for $t\bar{t}H$+tHW versus background. These classifiers are designed to suppress the large continuum and resonant backgrounds while retaining good efficiency for the corresponding signal topology.

In these tasks, the individual photon observables become important inputs. Unlike the $t\bar{t}H$–tH separation, the signal–background problem profits directly from variables sensitive to non-prompt photons and to diphoton systems that do not stem from a heavy resonance decay. For this reason, photon ID scores and pixel-seed information are included in the input space in addition to the event-topology variables. The former help discriminate against $j \to \gamma$ misidentifications, while the latter help identify the $e \to \gamma$ contributions. The angular separation ΔR of the two reconstructed photons and the Collins–Soper angle [228] are also used.

A further important aspect is that the adversary is trained simultaneously on both signal processes, while the classifier BCE loss is defined only with respect to the signal class relevant for the corresponding binary task. Concretely, for the tH–background classifier, $t\bar{t}H$+tHW events are masked from the BCE loss but retained for the adversarial loss; conversely, tH events are masked from the BCE loss of the $t\bar{t}H$+tHW–background classifier. This is necessary because decorrelating only with respect to, for example, the tH process was found to leave a strong residual dependence on the $t\bar{t}H$ coupling states in the output distribution, and vice versa. Since the later categorisation relies on the combined discriminant space of all three classifiers, the more global adversarial prescription was found to lead to a more stable overall signal efficiency across coupling assumptions.

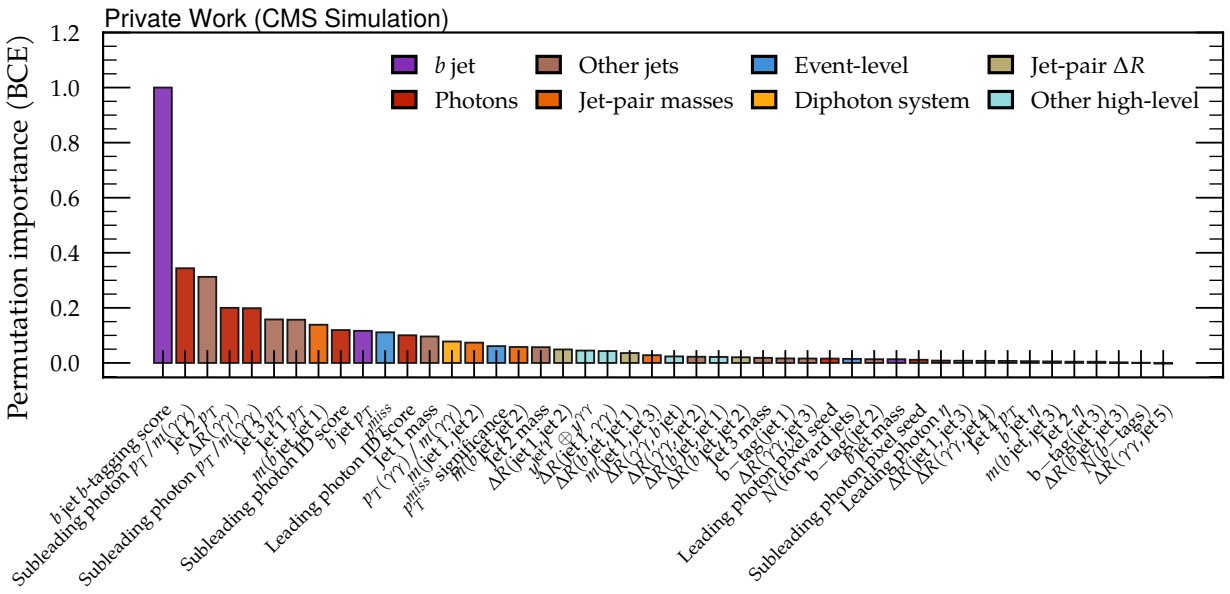


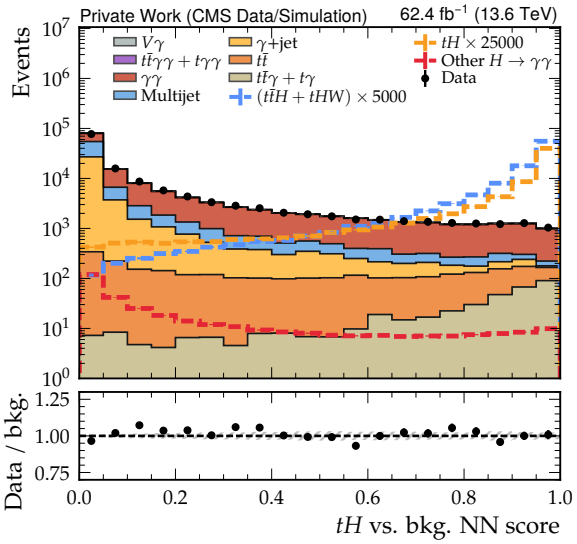
Figure 7.10: Permutation feature importance for the representative hadronic tH–background classifier evaluated with respect to the BCE loss.

This section focuses on the hadronic tH–background classifier as the representative and most challenging example. This choice is motivated both by its large background contamination and by the comparatively limited prior exploration of this phase space in $H \to \gamma\gamma$ measurements. Even without explicit decorrelation, the signal–background classifiers are less strongly model dependent than the $t\bar{t}H$–tH classifiers. Nevertheless, significant differences between the coupling hypotheses remain, mainly because individual photon and diphoton transverse momenta are both signal-discriminating and coupling sensitive. Here, the CP-odd and inverted-sign scenarios tend to have a harder photon spectrum than the SM scenario, which leads to an improved separability of these samples from all backgrounds. Adversarial decorrelation reduces these residual differences further and is therefore retained also for the signal–background classifiers.

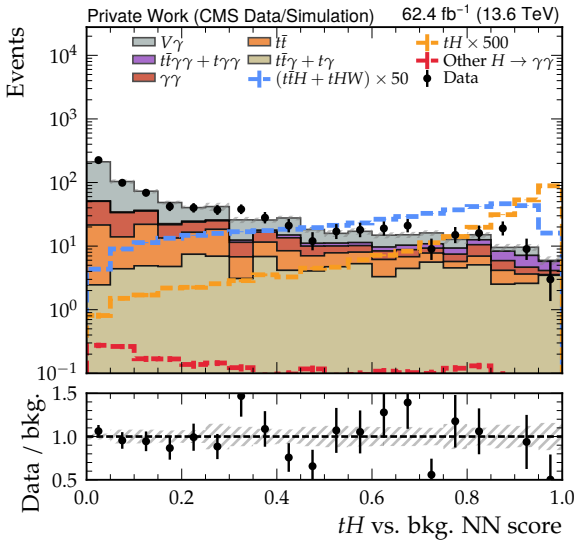
The permutation feature importance with respect to the classifier BCE loss for the hadronic tH versus background classifier is shown in Figure 7.10 as an example. The most important variable is the b-tagging score of the selected b jet, reflecting the large contamination from backgrounds without a genuine b-flavoured jet. As expected, the photon transverse momenta, the angular separation of the two photons, their ID scores, and the transverse momenta of the leading jets are among the most relevant features. In the leptonic channel, the dominant features differ notably, as the background composition is different. There, the invariant masses $m(\gamma_1, e)$ and $m(\gamma_2, e)$ are particularly important because they identify $e \to \gamma$ misidentifications from DY+γ events. Pixel-seed information and jet multiplicities also rank highly in that case.

The hadronic tH–background classifier reaches ROC AUC values of 0.971, 0.973, and 0.971 for the SM, CP-odd, and $\kappa_t = -1$ hypotheses, respectively. It therefore provides a strong suppression of the combined background while maintaining a stable response across the coupling hypotheses. The output distribution of this classifier, shown in Figure 7.11a, exhibits a decrease of the background density by roughly two orders of magnitude over the displayed range. The strongest rejection is achieved against multijet background, for which the ROC AUC exceeds 0.99. By contrast, the backgrounds least distinguishable from the signal are $t\gamma\gamma$- and $t\bar{t}\gamma\gamma$-like processes, which have final states very similar to the signal and therefore yield ROC AUC val-

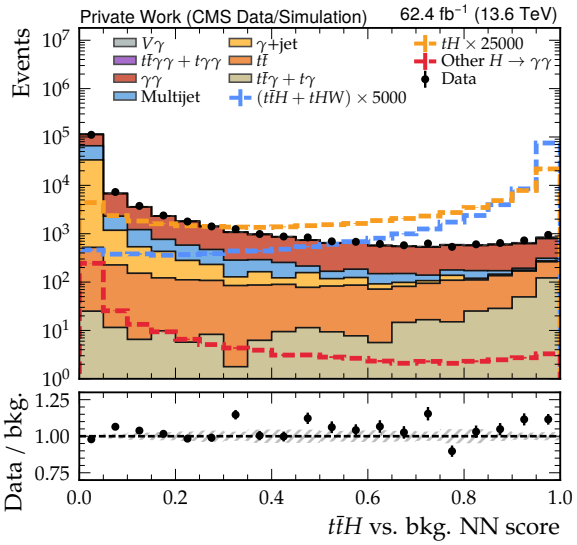
ues only around 0.6. As a consequence, the signal-like region is enriched in these irreducible backgrounds as well as in $t\bar{t}\gamma$, $t\gamma$ and, to a lesser extent, $t\bar{t}$ events. Owing to the much larger production rate of diphoton events, however, the dominant continuum contribution in the most tH-like hadronic phase space remains diphoton production in association with heavy-flavour jets. Among the resonant backgrounds, the small tails of the ggH and VBF production modes enter the highest-score region as relevant backgrounds.
The output distributions obtained for all four signal–background classifiers are shown in Figure 7.11. All four networks achieve a clear separation of the corresponding signal process from the combined background. At the same time, the agreement between mass–sideband data and continuum–background simulation is reasonably good, supporting the use of these discriminants together with the simulation-based background model in the categorisation optimisation.



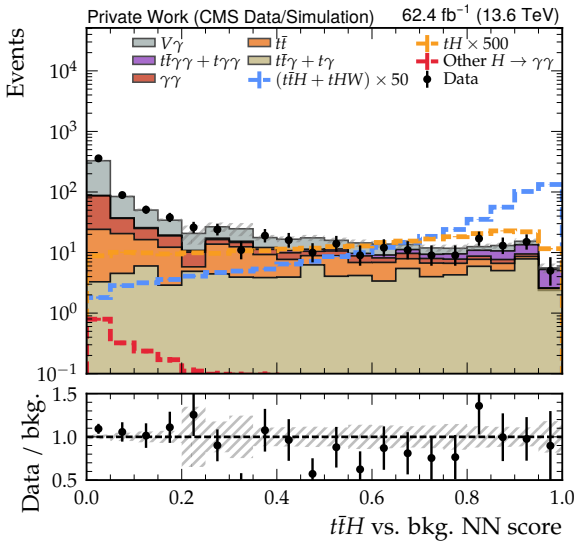
(a) Hadronic channel



(b) Leptonic channel



(c) Hadronic channel



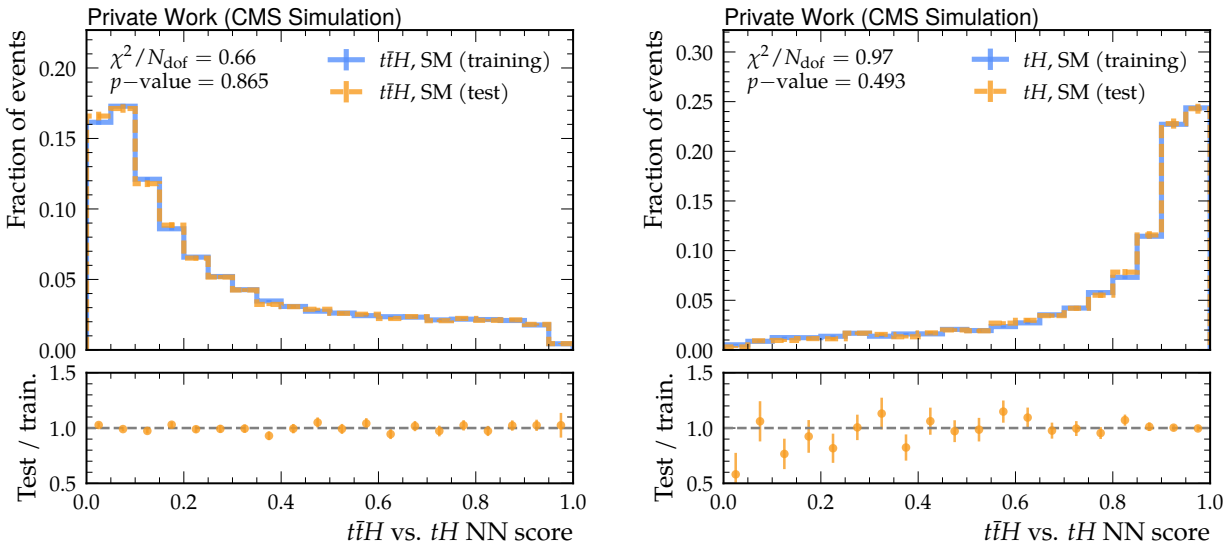
(d) Leptonic channel

Figure 7.11: Output distributions of the four signal–background classifiers used in the measurement. Data and continuum backgrounds are taken from the diphoton-mass sidebands excluding the 120–130 GeV signal region, while the simulated $H \to \gamma\gamma$ processes are overlaid. Uncertainties from limited MC event numbers are indicated.

7.7. Validation and robustness checks

In addition to the coupling-robust training prescription, several dedicated validation studies are performed to ensure that the classifiers can be used reliably in the statistical analysis. The first check concerns overtraining and possible generalisation differences between the training and test samples. For each classifier, the output distributions obtained on the training and test samples are compared for all simulated $H \to \gamma\gamma$ processes relevant to the measurement. Such differences would be relevant because the training samples are later also used in the statistical model in order to reduce the statistical uncertainty of the simulated signal and resonant background templates. The regularisation strength is adjusted such that no significant trend remains in the ratio of training to test set distributions. The agreement is quantified through a χ^2 probability, and only classifiers with probabilities above 0.05 for all relevant samples are retained. In addition, unweighted events are used to evaluate a Kolmogorov–Smirnov (KS) test as a complementary check of the overall agreement between the training and test distributions, and p-values above 0.05 are required. While the χ^2 test is more sensitive to local deviations in individual bins, the KS test probes the global consistency of the cumulative distributions and is therefore useful for identifying broader trends in the classifier response. In practice, these criteria require regularisation strengths λ_{L2} of order 10^{-3} for the hadronic classifiers and the leptonic signal discriminator, where the training statistics are large. Somewhat stronger regularisation of order 10^{-2} is used for the leptonic signal–background networks, which would otherwise be particularly prone to overtraining. The output distributions in Figure 7.11 are obtained with these final regularisation settings. These larger regularisation strengths also improve the numerical stability of the optimisation.

Figure 7.12 shows representative examples of the comparison between training and test set distributions for the $t\bar{t}H$ and tH samples obtained from the leptonic $t\bar{t}H$–tH classifier. The corresponding χ^2 probabilities are 0.87 and 0.49, respectively; the associated KS p-values are likewise above the acceptance threshold. This level of agreement is representative of the final models retained for the analysis. It indicates that the networks learn stable properties of the phase space rather than statistical fluctuations specific to the training subset, and that no sizeable generalisation difference is visible in the classifier-output distributions. No significant bias is therefore expected from using the training data also in the statistical analysis in order to reduce the statistical uncertainty on the signal and resonant background models.



(a) $t\bar{t}H$ sample **(b)** tH sample

Figure 7.12: Representative comparison of classifier-output distributions between the training and test samples for the leptonic $t\bar{t}H$–tH classifier. The SM samples are shown as an example.

A second and complementary validation is performed by analysing the network scores in the $Z \to e^+e^-$ control region defined in Sec. 4.4. The corresponding distributions are displayed in Figure 7.13. For the $t\bar{t}H$–tH classifier, the expected separation between $t\bar{t}\ell\ell$ and tZ topologies is clearly visible, in close analogy with the separation between the two signal classes in the diphoton phase space. A similar pattern is observed for the two signal–background classifiers, where the control-region analogues of the respective signal processes populate the more signal-like output regions. Compared with the signal region, the signal–background classifiers fall less steeply towards high output values in the control region, such that even the signal-like parts of the phase space are dominated by DY events. This is expected, since the predominant background in the control region is from $Z \to e^+e^-$ decays with additional jets from the DY sample, which already contain the photon-like signature from electrons reconstructed as photons. The classifiers, however, exploit photon-ID variables that are designed to discriminate against photon candidates of hadronic origin, while for electrons, the photon ID scores match the photon distribution closely. Overall, the data follow the simulation within the uncertainties in the majority of the displayed bins. The most pronounced discrepancy occurs in the $t\bar{t}H$-like region of the $t\bar{t}H$–tH classifier. This region is mainly populated by events with large jet multiplicities, for which the modelling of the dominant DY background is expected to be least reliable. Only up to two jets are described at matrix-element level, while additional radiation is generated by the PS. No indication is found that this behaviour points to a problem in the classifier itself or in the modelling of the signal events, where the underlying jet modelling is expected to be more accurate. Instead, the generally good agreement in the $Z \to e^+e^-$ region supports the conclusion that the classifier outputs are sufficiently well modelled for the intended use in the measurement. On the basis of this control-region study, no additional systematic uncertainty is assigned solely to classifier-output modelling. Residual effects in the signal and resonant-background category populations are covered by the systematic object-level and signal modelling uncertainties discussed later, while the continuum background is constrained from data in the final diphoton-

mass fit in each category. Any remaining mismodelling not already covered by these effects is judged to be negligible compared with the total uncertainty of the measurement.

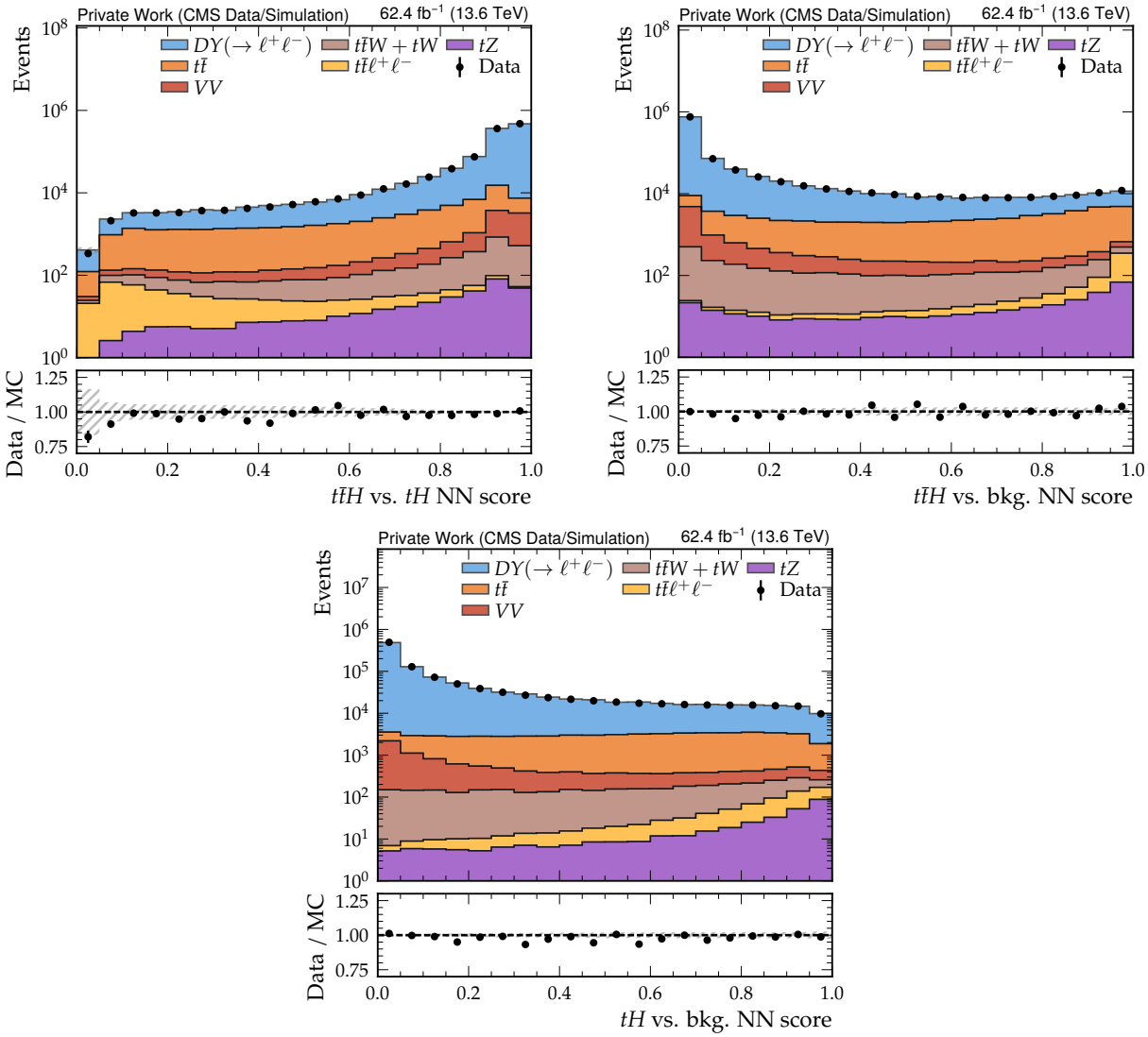


Figure 7.13: Output distributions of the hadronic classifiers in the $Z \to e^+e^-$ control region. The panels show the $t\bar{t}H$–tH, $t\bar{t}H$+tHW versus background, and tH versus background classifier scores. Uncertainties from limited MC event numbers are indicated.

8. Differentiable optimisation of the event categorisation

Events selected for the $t\bar{t}H$ and tH cross-section measurement in the $H \to \gamma\gamma$ channel are background-dominated, while the signals are both rare and kinematically different. A central analysis step is therefore an event categorisation that enables a sensitive simultaneous extraction of both cross sections. The foundation for this step is set with the neural network classifiers described in Chapter 7, which provide three discriminants to separate the signals from each other and each signal from the background. The choice of category boundaries therefore directly affects the expected precision: if events with different expected signal-to-background ratios or different expected $t\bar{t}H$-to-tH signal ratios are grouped together, part of the separation power of the continuous classifier outputs is lost. This chapter describes the categorisation optimisation framework used in this thesis, presenting a novel approach to learn optimal bin boundaries in multi-dimensional discriminant spaces using differentiable optimisation. This work has been published in Ref. [229], where the method is applied to a toy dataset and compared with a Bayesian optimisation [230–232] approach in detail. Parts of this chapter are reproduced or adapted from that publication. The text below follows the motivation and methodology from that publication, while the toy studies are shown in a condensed form. The application to the $t\bar{t}H$ and tH analysis is discussed in the latter part of this chapter, presenting the first application of the differentiable categorisation optimisation from Ref. [229] to collider data.

As in the measurement presented in this thesis, many other studies of collider data depend on a categorisation scheme on which the statistical inference is based. This includes simple cases where events are binned according to one-dimensional discriminants for likelihood fits, more complex cases with multidimensional discriminants, and more generally, any kind of statistical analysis performed on categorised data. Therefore, finding an optimal solution to the categorisation problem is of interest to a broad range of high-energy physics analyses. However, bin boundaries are often chosen manually or by simple rules such as equidistant or quantile-based binning. For multi-class classifier outputs, a common approach is to assign each event to the class with the largest score ("argmax") and then bin the one-dimensional projections in each class separately. These choices are convenient but not guaranteed to be optimal for the chosen figure of merit.

The number of published works directly addressing this problem is limited, motivating the detailed study presented in Ref. [229] and used for this measurement. The most important related works are summarised here. In the adaptive "Bayesian Blocks" histogramming approach [233], variable-width bins are defined by maximising a Poisson likelihood for piecewise-constant rates within these bins, penalised for adding more boundaries. This approach, however, targets a minimal loss of density-estimation quality from the binning procedure in one-dimensional distributions rather than an analysis figure of merit such as the significance of a signal. The ThickBrick method [234] optimises categorisation thresholds for a significance metric through a clustering-like procedure. Inference-aware optimisation approaches of machine-learning models, such as INFERNO [235] and related methods [236, 237], replace the non-differentiable histogramming

step with differentiable approximations while optimising a likelihood-derived loss. These methods keep bin edges fixed a priori and optimise the classifier outputs to perform the categorisation directly.
In many cases in high-energy physics analyses, however, classifiers are trained without such end-to-end optimisation pipelines. In particular, the study from Ref. [238] constructs categories directly in multi-dimensional score spaces using an unsupervised k-means clustering of the simulated events, hence possibly exploiting correlations but not targeting a specific sensitivity metric. The approach presented here instead focuses on learning the bin boundaries, while directly optimising a chosen figure of merit in a discriminant space of arbitrary dimension. This is achieved by parameterising the bin boundaries such that the chosen significance metric is differentiable with respect to their positions and shapes, while keeping the approach sufficiently simple to integrate into typical analysis workflows. The optimisation approach, labelled Gradient-based cATegorisation Optimiser (GATO), is implemented in the open-source Python package `gato-hep`, available on PyPI [239]. The implementation is based on TensorFlow 2.19 [216] and uses the RMSprop [240] optimiser in the experiments shown here.

8.1. Categorisation approach

The categorisation in a discriminant space of dimension d is parameterised using a Gaussian mixture model (GMM) with K components. The number of components equals the number of analysis categories. For an event with discriminant value $x \in \mathbb{R}^d$, the likelihood under component k is given by the multivariate normal distribution,

$$\ell_k(x) = \mathcal{N}\big(x \mid \boldsymbol{\mu}_k, \Sigma_k\big), \tag{8.1}$$

with component mean $\boldsymbol{\mu}_k$ and covariance matrix Σ_k. Including a mixture weight π_k, the logarithmic score is defined as

$$s_k(x) = \log \ell_k(x) + \log \pi_k. \tag{8.2}$$

In the hard-assignment limit as applied in the inference, each event is assigned to the bin corresponding to the largest score $s_k(x)$, and bin boundaries are defined by the equal-score surfaces between components. This defines flexible, non-rectangular category shapes via equal-score surfaces, despite each component being Gaussian.
Positive definiteness is ensured by parameterising each covariance via a Cholesky factorisation

$$\Sigma_k = L_k L_k^T, \tag{8.3}$$

with a lower-triangular matrix L_k. Starting from unconstrained parameters L_k^{raw}, we define an off-diagonal and a diagonal part,

$$O_k = \kappa \cdot \text{lower}(L_k^{\text{raw}}), \qquad D_k = \exp\big(\text{diag}(L_k^{\text{raw}})\big), \tag{8.4}$$

and construct L_k as the sum of these parts, $L_k = O_k + D_k$. Here, $\text{lower}(\cdot)$ extracts the lower-triangular part of a matrix, and $\text{diag}(\cdot)$ extracts the diagonal elements as a vector. The exponential ensures strictly positive diagonal elements of the Cholesky factor, and the factor κ is a hyperparameter introduced to damp the off-diagonal evolution for stability. In all studies presented here, we set $\kappa = 0.1$. Mixture weights are constrained to be positive and to sum to unity via a softmax normalisation of unconstrained parameters.

Two parameterisations for the component means are supported in `gato-hep`, with the goal to keep the means within the discriminant phase space. For discriminants that live on a simplex (e.g. multi-class softmax outputs), one can omit a coordinate and apply a softmax to unconstrained trainable parameters so that the means reside on the simplex. This approach is used in the toy studies in Ref. [229] which are summarised in Sec. 8.2. In the analysis of this thesis, where the three-dimensional discriminant space is restricted to the unit hypercube $[0,1]^3$, the means are mapped to this domain via a sigmoid transformation instead. The one-dimensional sigmoid-boundary model is optimised analogously, using differentiable bin weights defined by sigmoids as smoothed step functions, which is the standard approach in differentiable programming.

8.1.1. Differentiable optimisation

To enable gradient-descent based optimisation of the GMM parameters, the hard assignment of an event to one bin is replaced by a differentiable approximation using a temperature-scaled softmax over the scores,

$$\gamma_k(x) = \mathrm{softmax}_k\left(\frac{s_k(x)}{T}\right). \tag{8.5}$$

These normalised weights sum to one over all categories and can hence be interpreted as fractional bin memberships. Expected bin yields are then obtained by summing the original event weights multiplied by $\gamma_k(x)$ for category k. For high temperatures T, event memberships are distributed broadly across bins, while in the low-temperature limit $T \to 0$, the hard-assignment behaviour is recovered. A cosine-annealing schedule is used for the temperature and learning rate to encourage exploration at early stages of the optimisation and to converge towards harder assignments.

Sigmoid-based model

In addition to the GMM-based categorisation, a simpler one-dimensional binning model is used as a comparison that parameterises the bin edges directly and replaces the hard category assignment by a smooth, differentiable approximation based on sigmoid functions. For a discriminant value $x \in [x_0, x_1]$ and K bins, $K-1$ ordered cut positions are introduced,

$$x_0 < c_1 < c_2 < \cdots < c_{K-1} < x_1,$$

which are constructed such that they remain strictly ordered (i.e. boundaries cannot cross) by parameterising them as positive cumulative increments from softmax-activated trainable parameters.

To obtain differentiable bin memberships, each cut at position c_j is mapped to a smooth step via a sigmoid,

$$s_j(x) = \sigma\big(k\,(x - c_j)\big) = \frac{1}{1 + \exp\big(-k\,(x - c_j)\big)},$$

where $k > 0$ controls the steepness of the transition. Analogous to the temperature parameter used in the GMM-based soft assignment, larger values of k yield a sharper (more step-like) binning, while smaller values of k increase smoothness and stabilise gradient-based optimisation.

From these cumulative steps, non-negative bin weights that sum to unity are constructed:

$$w_1(x) = 1 - s_1(x), \qquad w_m(x) = s_{m-1}(x) - s_m(x) \;\; (m = 2, \ldots, K-1), \qquad w_K(x) = s_{K-1}(x).$$

In the limit $k \to \infty$, this recovers a hard assignment to the bins defined by the cuts c_j.

In multiple dimensions, this construction can be applied per coordinate to obtain rectangular bins. For an event with discriminant vector $x = (x^{(1)}, \ldots, x^{(d)})$, the superscript a labels the coordinate in the d-dimensional space. If m_a denotes the bin index along coordinate a, and $w^{(a)}_{m_a}(x^{(a)})$ is the corresponding one-dimensional bin weight for that coordinate, the weight of the multi-dimensional rectangular bin indexed by $(m_1, \ldots, m_d)$ is given by the product

$$w_{m_1,\ldots,m_d}(x) = \prod_{a=1}^{d} w^{(a)}_{m_a}(x^{(a)}),$$

which corresponds to the outer product of the per-coordinate bin weights. This yields a differentiable approximation to axis-aligned cuts while retaining the one-dimensional interpretability of the learned boundaries.

8.1.2. Optimisation objective and penalties

The optimisation target can be any continuous, differentiable function of the expected signal and background yields in each bin, denoted by S_k and B_k for bin k. In this analysis, the Asimov signal significance [241] is used, calculated per bin as

$$Z_k = \sqrt{2\left[(S_k + B_k)\ln\left(1 + \frac{S_k}{B_k}\right) - S_k\right]}, \tag{8.6}$$

and combined in quadrature over bins. This quantity approximates the median expected discovery significance under the signal-plus-background hypothesis and is a suitable proxy for sensitivity optimisation in a cross-section measurement. Throughout this chapter, significance values computed in this way are quoted as equivalent Gaussian standard deviations.

To avoid bin configurations that are inappropriate in realistic analyses, optional penalty terms are included. First, a low-yield penalty suppresses bins with background yields below a threshold $B_{\min}$,

$$P_{\text{low}}(B) = \sum_{k=1}^{K} \left[\max(0, B_{\min} - B_k)\right]^2. \tag{8.7}$$

A second term is introduced to discourage bins with large relative background uncertainties. For a given bin k, the expected background yield is given by the sum of the weights of simulated events assigned to that bin,

$$B_k = \sum_i w_i, \tag{8.8}$$

where w_i denotes the weight of event i assigned to that bin. Assuming independent Poissonian fluctuations of the underlying event counts, the variance of a weighted sum is given by the sum of squared weights,

$$\text{Var}(B_k) = \sum_i w_i^2. \tag{8.9}$$

This yields the relative statistical uncertainty on the background prediction in bin k,

$$\sigma_{\mathrm{rel},k} = \frac{\sqrt{\sum_i w_i^2}}{B_k}. \tag{8.10}$$

Accordingly, the penalty with respect to a threshold r is defined as

$$P_{\mathrm{unc}}(B) = \sum_{k=1}^{K} \left[\max(0,\, \sigma_{\mathrm{rel},k} - r)\right]^2. \tag{8.11}$$

A related strategy has recently been explored in the fully differentiable analysis pipeline of Ref. [242]: bins with insufficient MC statistics are smoothly down-weighted or masked during training to stabilise the optimisation.
The penalties used here are weighted by tuneable hyperparameters and added to the main objective to encode analysis constraints such as minimal expected background yields and stable uncertainty estimates.

8.2. Toy examples

First, the categorisation optimisation method is demonstrated on toy datasets as presented in Ref. [229]. The set of toy datasets is constructed from multivariate Gaussian distributions in a three-dimensional feature space, with one or two signals and five background processes, which differ in their means. A one-dimensional discriminant is defined by an analytic likelihood ratio of the signal over the combined background for a binary classification example. For the three-class case, one likelihood ratio is constructed for each of the two signals and the combined background, and the resulting three values are normalised with a softmax transformation. Two scenarios are considered for the three-class case, which primarily differ in the separability of the two signal classes. In the first setup, the ROC AUC for separating the two signals is about 0.8, whereas it exceeds 0.98 in the second setup, representing a nearly perfect separation.
Fig. 8.1 confirms that the optimised binning approaches in one dimension achieve a higher significance than fixed equidistant binning for a given number of bins. All optimised strategies, i.e., the Bayesian approach and the two GATO approaches based on the GMM and the simpler sigmoid-based approach, converge towards similar significances. The solutions found by the two GATO approaches typically agree closely, indicating that the optimisation landscape is well-behaved and that the solutions are close to a local maximum of the significance.
In multi-dimensional cases, the GMM-based optimisation can exploit the full structure of the discriminant space. Here, the loss is the negative geometric mean of the two signal significances in the three-class cases, encouraging balanced sensitivity to both signals. Figure 8.2 illustrates this behaviour for two representative three-class scenarios. The score densities in the top row show that the main difference lies in the overlap of the two signal classes: in the first scenario, the 68% contours overlap significantly, whereas they are clearly separated in the second scenario. The bin boundaries learned in a setup with five bins, shown in the middle row, visualise how the GMM components adapt to the discriminant structure. The bottom row compares the achieved significance of the GMM-based approach with two baselines. In the first, the binning within the argmax projections onto each signal score is optimised separately with the Bayesian approach. The second uses the k-means clustering algorithm from Ref. [238]. When the signal separation is imperfect, the GMM optimisation yields a clear improvement over argmax-

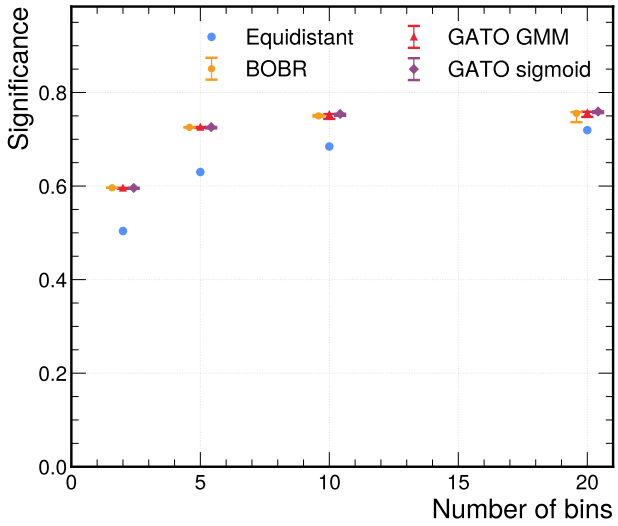
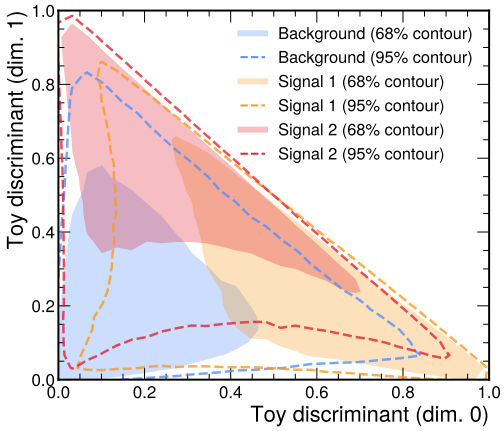
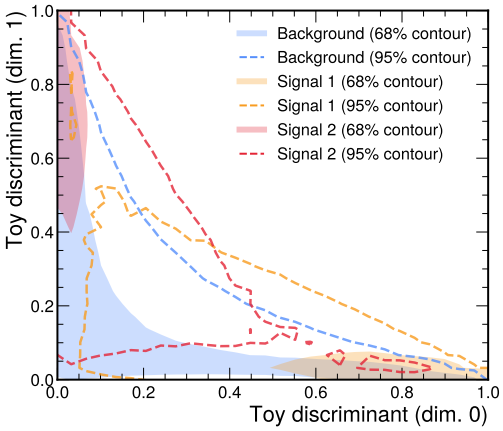


Figure 8.1: Comparison of the significance obtained in the one-dimensional setup. Error bars are obtained from 20 independent optimisation runs, marking the 16th to 84th percentiles. BOBR denotes the Bayesian optimisation approach. Markers are shifted horizontally around the tested bin counts to enhance the visibility. This figure was published in Ref. [229].

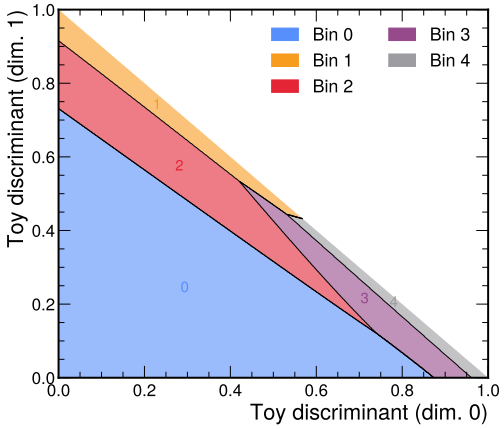
based projections and clustering baselines, while for strong separation, the argmax-projected and subsequently one-dimensionally optimised approach becomes competitive. Nevertheless, the GATO approach achieves competitive or better significances in both scenarios, demonstrating the flexibility and effectiveness of the method in multi-dimensional discriminant spaces.



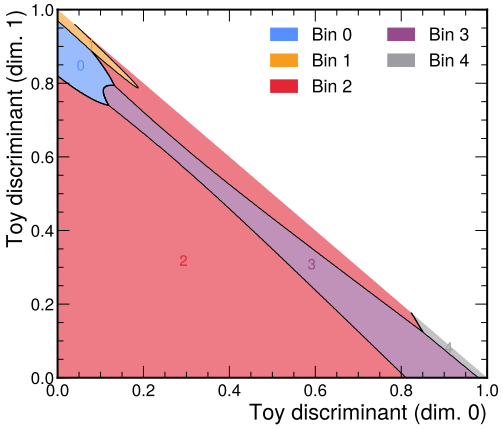
(a) Toy score densities: scenario 1



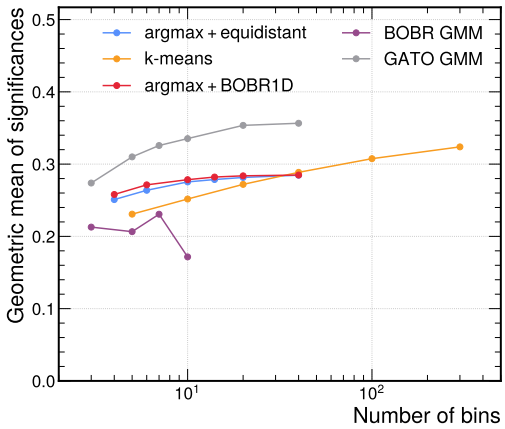
(b) Toy score densities: scenario 2



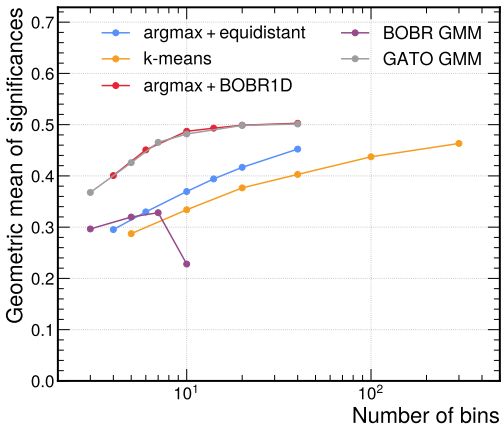
(c) Optimised GMM boundaries: scenario 1



(d) Optimised GMM boundaries: scenario 2



(e) Scenario with overlapping signals



(f) Scenario with well-separated signals

Figure 8.2: Three-class toy discriminant space with density contours (top row), the resulting five-bin GMM boundaries (middle row), and comparison of the geometric mean of the two signal significances (bottom row) for the two scenarios. Only the two signal scores are shown in the contour plots, with the background score given implicitly by the softmax constraint. This figure is adapted from Ref. [229].

8.3. Application to the $t\bar{t}H$ and tH cross-section measurement

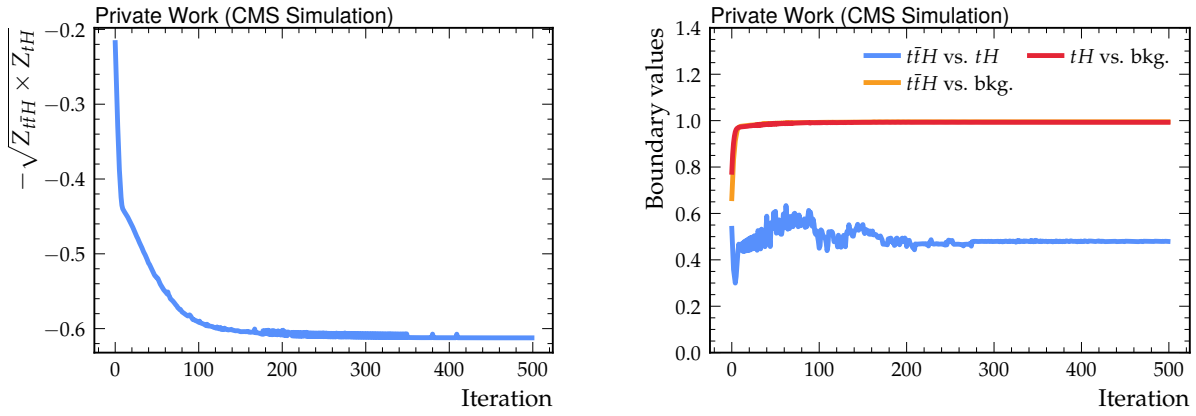
The categorisation is performed separately in the hadronic and leptonic channels. The negative geometric mean of the expected significances for the $t\bar{t}H + tHW$ and tH signals,

$$L = -\sqrt{Z_{t\bar{t}H} \times Z_{tH}},$$

is used as the primary loss function to be minimised. In the application to this analysis, the optimisation studies described below use signal samples generated with the SM top-quark Yukawa coupling; samples with alternative CP assumptions are introduced only in Sec. 8.4, where they are used to construct a dedicated robustness penalty. Only simulation samples are used for the optimisation to avoid potential biases of the measurement from statistical fluctuations in the observed sideband data. The expected signal and background yields are computed in a narrow diphoton mass window of $125 \pm 2\,\text{GeV}$, which contains the majority of the signal events. In this window, however, the number of simulated continuum background events is limited, in particular in the signal-like phase space, where the background is strongly suppressed and a precise determination of the bin boundaries is important. To improve the statistical precision of the continuum-background estimate, sideband events outside this window are included. The continuum yield in the signal window is then obtained by reweighting the total continuum contribution to match the smaller signal window yield. For this purpose, an exponential function is fitted to the simulated continuum diphoton mass spectrum. The ratio of the fitted yield in the signal window to the total weighted yield over the full range is used as an event-wise scaling factor when estimating the continuum contribution to B_k. Because the shape of the diphoton mass distribution changes across the discriminant space, this reweighting is performed separately for each category and repeated every 25 iterations of the GATO training. Using the entire mass range is not expected to bias the optimisation, since the analysis chain is designed to be diphoton-mass agnostic and the final statistical analysis determines the continuum background from fits of smoothly falling functions to observed data.

Each run is performed for 500 epochs using full-batch training, because the statistical uncertainty would be too large in mini-batches. The learning rate is scheduled with a cosine profile from 0.1 to 0.001, and the temperature is annealed from 0.5 to 0.01. At the final temperature, the expected significance computed using the differentiable fractional memberships differs by less than 1 % from that obtained after assigning each event to the category with the largest score. The low-background penalty is weighted by $\lambda = 0.001$ using the threshold $B_{\text{min}} = 10$ expected background events per category. This threshold is motivated by bias studies described in Ref. [243]. Even when using the discrete profiling method [244] to mitigate biases from the choice of analytic background functions, residual biases from the choice of the fit function may remain in categories with very low event counts. These were quantified using pseudo-datasets by comparing significances obtained with fixed and floating background functions. A minimum of ten expected events per category was found to be the lowest value that still yields acceptable residual biases (at the 20 % level), and is therefore adopted here as a pragmatic constraint to ensure robust statistical inference.

The two categorisation approaches supported in `gato-hep` are compared in this analysis. The first defines categories with rectangular boundaries based on the sigmoid model as a baseline. Here, events are classified as either $t\bar{t}H$- or tH-like using a single cut on the corresponding classifier output, and each region is further split into multiple categories by cutting on the relevant background-discriminant score. The model learns all cut positions simultaneously to maximise



(a) Loss

(b) Boundary positions

Figure 8.3: Optimisation history for the sigmoid-based categorisation with four categories in the hadronic channel, showing the loss (left) and the category boundaries (right).

the expected sensitivity. In the simplest case, this leads to four categories: two $t\bar{t}H$ and two tH categories, corresponding to high- and low-purity regions defined by cuts on the signal–background scores.

The evolution of metrics during the optimisation, using the hadronic channel as an example, is shown in Fig. 8.3 for this setup. The loss decreases and converges, while the boundaries move from their random initialisation towards values that maximise the expected significance. The largest changes of the boundary positions appear within the first epochs, where the loss also decreases most rapidly. The threshold on the $t\bar{t}H$–tH split evolves less smoothly because the loss function is comparatively flat in this region, and convergence is reached after about 300 epochs. In this example, the final boundaries are obtained to be approximately 0.479 for the $t\bar{t}H$–tH split and 0.993 and 0.995 for the tH–background and $t\bar{t}H$–background splits, respectively.

For this particular setup, a validation scan is performed by fixing two boundary values and scanning the third. Here, the background yields are determined from exponential fits over the entire diphoton mass range for the continuum background, while the signal and resonant background yields are taken by simply summing the event weights in the signal window. Fig. 8.4 shows the scan results. The figure of merit is flat over a wide range of cut positions for the $t\bar{t}H$–tH split, while being particularly sensitive to the position of the signal–background-discriminant thresholds. As expected from the classifier design, moving the boundary on the $t\bar{t}H$–background discriminant has the strongest impact on the significance of the $t\bar{t}H$ signal, while the tH–background boundary mainly affects the tH significance. No pronounced secondary local maxima are observed, indicating a well-behaved optimisation landscape. This test suggests that the optimisation converges to a local maximum of the significance in this simplest setup.

The second setup uses the GMM in the three-dimensional discriminant space. In the GMM approach, component means are mapped to the $[0, 1]^3$ hypercube via a sigmoid transformation. The GMM components define non-rectangular categories and are hence less constrained than the sigmoid-based approach; however, optimality is harder to validate in this case. The approach adopted is therefore to compare the expected significances obtained with both schemes and to select the configuration yielding the best sensitivity, while keeping the number of categories as small as feasible for the final analysis.

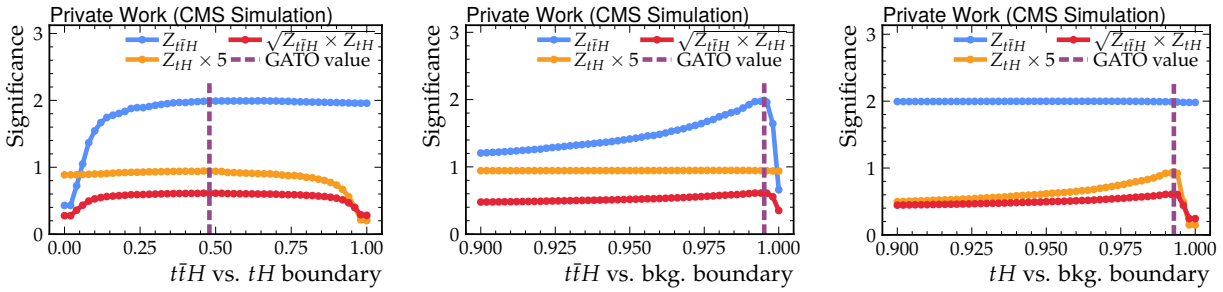
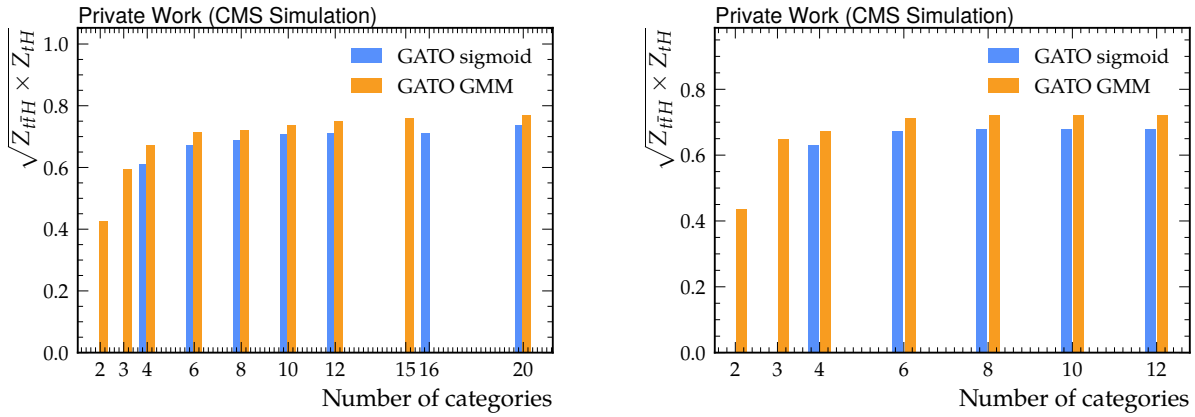


Figure 8.4: Validation scans for the sigmoid-based categorisation with one trainable cut point for each of the three neural network discriminants, while keeping the other two fixed to the optimised values. The learned cut points are indicated by the dashed vertical lines, which coincide with the maxima of the geometric mean of the signal significances.

For both setups, the number of categories is scanned from low to high, and the point where the expected significance saturates is chosen. The lowest number of categories is four for the sigmoid-based approach, as described above, if the same number of boundaries is placed on the two signal–background discriminants. In the GMM approach, the minimum number of components is two and all components are freely movable within the phase-space restrictions. Fig. 8.5 compares both approaches as a function of the number of categories in the hadronic and leptonic channels. The general trend shows that the expected significance increases with the number of categories up to a saturation point, indicating successful exploitation of the discriminant space. For a given number of categories, the GMM approach outperforms the sigmoid-based approach in both channels, as expected from its higher flexibility. In the hadronic channel, the significances continue to increase up to the largest tested category count of 20, with only a small improvement of around 1 % when increasing from 15 to 20 components. Therefore, higher category numbers are not tested. In the leptonic channel, where substantially fewer events are selected, the saturation occurs earlier at around eight categories. In both channels, the improvement of the GMM approach over the rectangular cuts is around 5–10 % for a given number of categories. Analysing the individual signal sensitivities underlying the geometric means in Fig. 8.5, the leptonic channel is found to be more sensitive to the $t\bar{t}H$ signal, with an expected significance of 2.6 compared to 2.3 in the hadronic channel. Conversely, the tH sensitivity is higher in the hadronic channel, reaching 0.25 compared to 0.19 in the leptonic channel.

In the following, categories are named according to the signal with the larger selection efficiency per category. If a category is more efficient for $t\bar{t}H+tHW$ than for tH, it is labelled "$t\bar{t}H$ had. i" or "$t\bar{t}H$ lep. i", otherwise "tH had. i" or "tH lep. i". The index i increases with decreasing category significance, i.e., the highest-significance category is labelled with $i = 1$. Categories that are not sensitive to either signal are grouped into a single background control category, "bkg. had." or "bkg. lep.", in order to simplify the subsequent statistical analysis. This includes all categories that do not improve the total significance for either signal by at least 1 %.

Fig. 8.6 visualises the learned GMM components for a hadronic three-component model in the three-dimensional discriminant space as a simple example. The ellipsoids show iso-density surfaces of each Gaussian component around its fitted mean and highlight how the model adapts to the event density. One component expands widely towards low signal–background scores and has a large volume, while the other components cluster in the high signal–background discriminant region and have a smaller volume, indicating that they both enrich a small number of signal-like events. The tH-sensitive and $t\bar{t}H$-sensitive components are mostly separated in



(a) Hadronic channel

(b) Leptonic channel

Figure 8.5: Comparison of the expected geometric mean of the signal significances obtained from sigmoid and GMM approaches as a function of the number of categories. The left subfigure shows the hadronic channel, and the right subfigure shows the leptonic channel.

the $t\bar{t}H$–tH discriminant. The components do not align with the coordinate axes, demonstrating the advantage of the GMM approach in exploiting correlations between the discriminants over argmax-based projections. It is important to note that the displayed ellipsoids are not the category boundaries themselves, which are instead given by equal-score surfaces between components. However, the surfaces shown illustrate the component shapes from which these boundaries are derived.

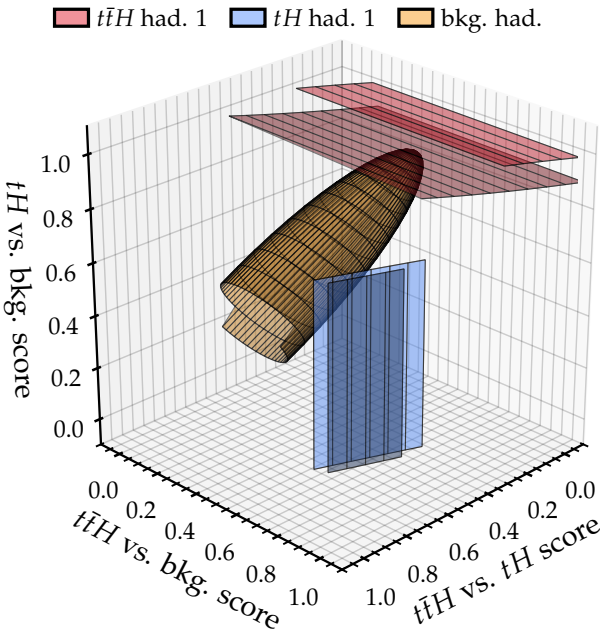


Figure 8.6: Visualisation of a trained example GATO GMM with three components in the three-dimensional DNN output space for the hadronic channel. The ellipsoids show iso-density surfaces of the Gaussian components, chosen such that they enclose a fixed probability fraction of the corresponding multivariate Gaussian distribution. They visualise the Gaussian components corresponding to the selected categories and are truncated at the displayed phase-space boundaries. Components can appear open because only their means are constrained to lie within the data range, while their covariance matrices are not explicitly restricted to remain inside it.

8.4. Categorisation penalty for robustness against CP assumptions

Even though decorrelated classifiers are used with their output distributions agreeing at the 10% level between the different coupling scenarios considered (see Chap. 7), the signal efficiencies in the most sensitive categories can still differ substantially. In particular, in the most tH-like categories, the $t\bar{t}H$ yield is small, so even moderate absolute coupling-dependent changes in the $t\bar{t}H$ tail can correspond to large relative efficiency differences among coupling hypotheses. It is observed that $t\bar{t}H$ events produced with a purely CP-odd top-quark Yukawa coupling tend to populate the tH-like categories more than the corresponding events with a CP-even Yukawa coupling, and the resulting relative efficiency shifts can even exceed the expected tH yields in the same categories. In the likelihood fits described in Chap. 9, this effect is observed to bias the extracted tH signal strength when the assumed CP state differs from the model used to generate toy datasets, as discussed there in detail. Conversely, the remaining CP-dependent migration of the tH tail into $t\bar{t}H$-like categories does not lead to significant biases because the lower total tH yield makes this migration less impactful. Effects of similar magnitude are observed for both the GMM-based and the rectangular cut-based approaches.

To stabilise the analysis against this effect, an additional penalty term is introduced that enforces similar per-category efficiencies for $t\bar{t}H$ events produced with CP-even and CP-odd top-quark Yukawa couplings, particularly in the most tH-sensitive categories. With S_k^{even} and S_k^{odd} denot-

ing the expected $t\bar{t}H{+}tHW$ signal yields in category k for the CP-even and CP-odd hypotheses, the per-category efficiencies are given by

$$\varepsilon_k^{\text{even}} = \frac{S_k^{\text{even}}}{\sum_j S_j^{\text{even}}}, \qquad \varepsilon_k^{\text{odd}} = \frac{S_k^{\text{odd}}}{\sum_j S_j^{\text{odd}}}. \tag{8.12}$$

The CP-decorrelation penalty is defined by the squared relative difference,

$$P_{CP} = \sum_k w_k \left(\frac{\varepsilon_k^{\text{even}} - \varepsilon_k^{\text{odd}}}{\varepsilon_k^{\text{even}} + \varepsilon_k^{\text{odd}} + \epsilon} \right)^2, \tag{8.13}$$

with a small ϵ for numerical stability. The weights w_k are chosen to be proportional to a cross-section-agnostic Asimov significance of the tH signal per category, obtained by normalising the tH signal to unit total yield for both coupling states. In this way, the weights depend on where the signal is in the phase space, not on the absolute cross section, which would favour coupling states with larger cross sections. This focuses the penalty on the most tH-sensitive categories where the potential bias is most pronounced, while leaving other categories largely unconstrained. This penalty is multiplied by a tuneable hyperparameter and added to the total loss, together with the significance objective and the low-yield penalty.
The effect of CP-decorrelation on efficiencies in optimised categories is illustrated in Fig. 8.7, where acceptance effects are included. In the most tH-like category, the ratio of CP-odd to CP-even efficiency times acceptance is about 5 (3) in the hadronic (leptonic) channel without CP-decorrelated training. With CP-decorrelated classifiers but no CP penalty in the GATO optimisation, the ratio drops to about 3.2 (1.7). With a CP-penalty strength of 30 and CP-decorrelated neural networks, the ratio reduces further to about 1.45 (1.40), only mildly above the $\sim$ 20 % higher acceptance for CP-odd events. However, this reduction of the model dependence of the categorisation comes at the cost of a reduction in expected tH sensitivity, by around 20 % in the hadronic and 5 % in the leptonic channel. An analogous figure for the tH efficiencies is shown in Appendix A. Importantly, repeating the category-count scan confirms that the preference for the GMM-based categorisation remains unchanged; the conclusions from the comparison in Fig. 8.5 without the CP penalty therefore still hold qualitatively. After including the CP-decorrelation penalty, models with 15 and 10 components are used in the hadronic and leptonic channels, respectively. With the naming and merging scheme described above, sixteen categories are used in the statistical analysis in total. This includes four $t\bar{t}H$-like and four tH-like categories in the hadronic channel, four $t\bar{t}H$-like and two tH-like categories in the leptonic channel, and one background-enriched category in each channel.

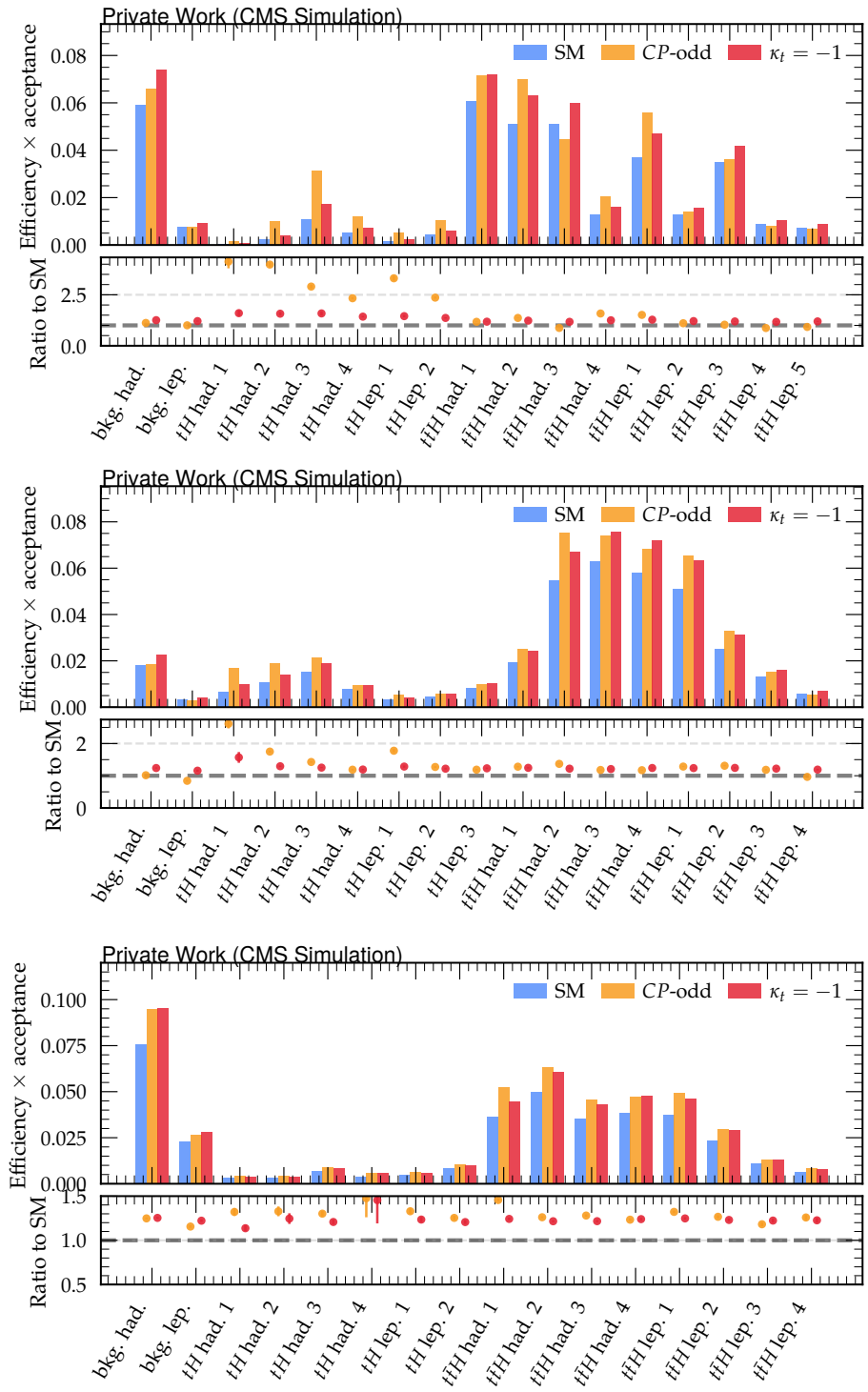


Figure 8.7: Per-category $t\bar{t}H$+tHW efficiencies (including acceptance) in the hadronic channel. From top to bottom: baseline setup with classifiers trained only on SM samples; CP-decorrelated classifiers without a CP penalty in the GATO optimisation; CP-decorrelated classifiers with a CP penalty of strength 30.

As a complementary cross-check, observed and simulated events in the $Z \to e^+e^-$ validation region are categorised with the final hadronic GMM model, i.e., the model trained with the CP penalty introduced above. The resulting category yields are compared in Fig. 8.8. The observed data follow the overall category structure predicted by the simulation well, with agreement at the level of about 10 % in most categories and about 20 % in the most signal-like ones. These differences do not exceed the uncertainty scales later assigned to the signal and resonant background processes in the statistical analysis. At the same time, the statistical uncertainty of the simulated prediction is of comparable size in the most signal-like categories, so this comparison is interpreted as a qualitative cross-check rather than as a precision closure test. The category composition is also consistent with the intended behaviour of the categorisation: $t\bar{t}Z$ events are enriched in the $t\bar{t}H$-like categories, while tZ contributes more prominently to the tH-enriched ones. No significant indication of mismodelling is observed.

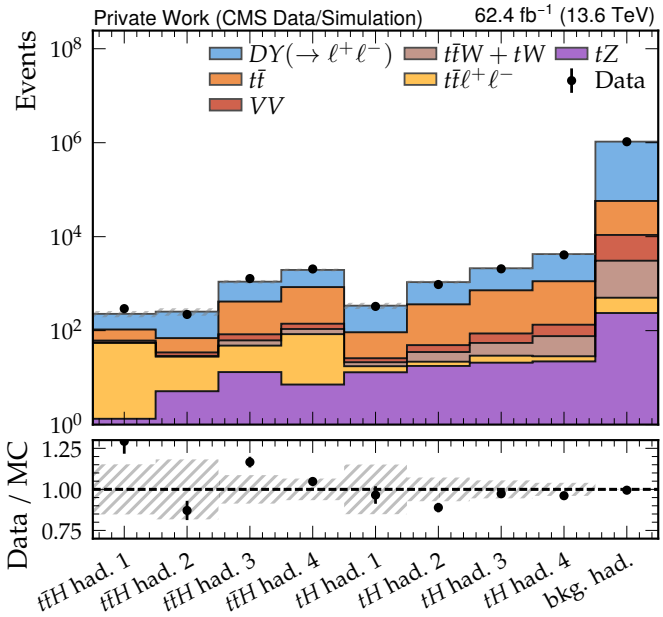


Figure 8.8: Modelling of the hadronic category composition in the $Z \to e^+e^-$ validation region. Observed event yields are compared with the simulated process composition after the final category assignment. Uncertainties due to the limited size of the simulation samples are shown.

9. Statistical analysis: methodology and results

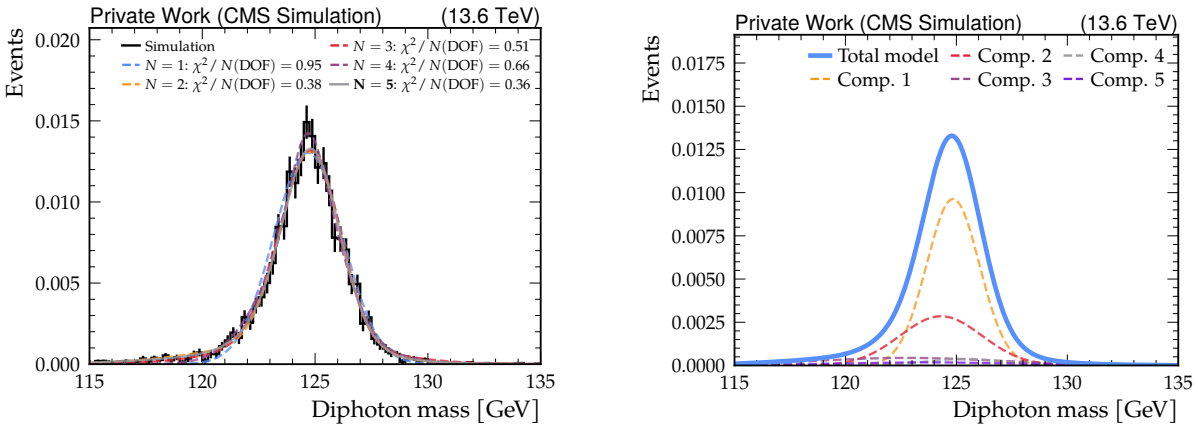
This chapter describes the statistical procedure used to extract the parameters of interest (POIs) from the observed data using parametric models for the signal and background contributions. The statistical analysis uses the categorisation defined in Chap. 8 and is performed on the diphoton mass spectrum $m_{\gamma\gamma}$. The $H \to \gamma\gamma$ processes are modelled from MC simulation, with the signals tH and $t\bar{t}H + tHW$ expressed as functions of the POIs, while the resonant backgrounds are described by their SM predictions. Nuisance parameters (NPs) incorporate effects of systematic uncertainties. In contrast, the continuum background is modelled directly from the observed $m_{\gamma\gamma}$ spectrum in data, assuming a smooth functional shape that does not qualitatively change from the broad sidebands to the narrow signal window. Signal and background are combined into a parametric model, and a binned profile likelihood fit is performed to determine the values of the POIs. The bin width is set to 0.25 GeV, which is much smaller than the diphoton mass resolution. Even in the best categories used in this analysis, the effective resolution is around 1.5 GeV, so the information loss due to the mass binning is negligible. The fits are performed with the `Flashgg FinalFit` package [245], built on the CMS COMBINE framework [246]. The parametric models are implemented with ROOFIT [247].
The POIs are defined as rate parameters for the fiducial cross sections relative to the nominal SM predictions,

$$r_{tH} = \frac{\sigma_{\text{fid}}(tH)}{\sigma_{\text{fid}}^{\text{SM}}(tH)}, \quad r_{t\bar{t}H+tHW} = \frac{\sigma_{\text{fid}}(t\bar{t}H + tHW)}{\sigma_{\text{fid}}^{\text{SM}}(t\bar{t}H + tHW)}. \tag{9.1}$$

For expected results, the statistical model is evaluated on an Asimov dataset corresponding to the median expectation under the assumed SM hypothesis [241]. This dataset is constructed from the nominal signal-plus-background model and therefore contains the expected bin contents without statistical fluctuations with respect to that model. Statistical uncertainties nevertheless enter the expected inference through the likelihood model. For the continuum background entering this expected model, the analytic functions are determined from sideband fits excluding the 120–130 GeV signal window, whereas the final observed results are obtained from fits to data including the signal and background models over the full mass range.

9.1. Signal and resonant background model

The shape of the diphoton mass peak mostly depends on the energy resolution and the vertex reconstruction efficiency. As the photon energy resolution as well as the vertex finding performance depend on the kinematic properties of the photons and overall event properties, a single signal model does not describe all effects adequately. Instead, the signal models are built separately per production process, for events either passing or failing the fiducial phase space criteria, and per category to account for the kinematic effects, as well as per data-taking period to capture changes in detector and beam conditions. The same parametric modelling strategy



(a) Model complexity scan

(b) Signal model composition

Figure 9.1: Illustration of the signal fitting procedure, showing the $t\bar{t}H$ process in the *$t\bar{t}H$ had. 1* category with the *2022postEE* dataset. The dependence of χ^2/N_{DOF} on the number of Gaussian components is shown in the left sub-figure, and the individual fitted components of the $N = 5$ model which minimises the metric are shown on the right.

is also applied to the resonant Higgs boson background from production modes other than tH and $t\bar{t}H{+}tHW$, while its normalisation is fixed to the corresponding SM prediction up to the nuisance-parameter variations described below.

Each signal shape is parameterised by a sum of up to five Gaussian components,

$$f(\hat{m}_{\gamma\gamma}) = \sum_{i=1}^{N\leq 5} w_i \, \mathcal{N}(\hat{m}_{\gamma\gamma} \,|\, \mu_i, \sigma_i), \tag{9.2}$$

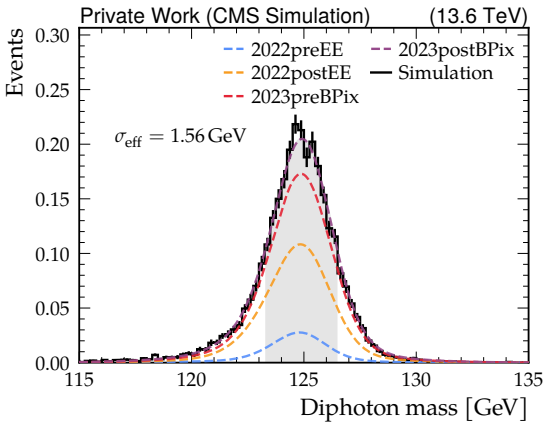
where $\hat{m}_{\gamma\gamma} \equiv m_{\gamma\gamma}/\mathrm{GeV}$ is dimensionless and w_i labels the relative contribution of component i to the model, such that $\sum_i w_i = 1$. The optimal number of components is selected by fitting models with $N = 1, \ldots, 5$ components and evaluating the reduced χ^2, χ^2/N_{DOF}, where N_{DOF} is the number of bins minus the number of free fit parameters. The configuration with the smallest reduced χ^2 is chosen. This procedure is illustrated for the $t\bar{t}H$ process with the category *$t\bar{t}H$ had. 1* and data-taking period *2022postEE* as an example in Fig. 9.1. Overall, a good modelling of the shape of the signal peaks is observed. The signal models are built from simulation with the Higgs boson mass set to $m_H = 125\,\mathrm{GeV}$ and are then shifted to the reference value $m_H = 125.08\,\mathrm{GeV}$, corresponding to the result from the most precise CMS measurement [97]. No interpolation of the acceptance or signal shapes beyond this small offset is performed, as the corresponding effects are assumed to be negligible. In categories with insufficient MC statistics for a given process, the corresponding shape is replaced by the $t\bar{t}H$ shape in the same category, using the larger sample to obtain a stable parametrisation, whereas the normalisation prediction is kept. This approximation assumes that the residual process-dependent differences in the diphoton mass shape are negligible compared to the statistical precision of the measurement.

Systematic uncertainties affecting signal shapes or yields are incorporated either by reweighting simulated events or by varying reconstructed quantities that induce migrations between categories or into and out of the selected event sample. They enter the fit as NPs, summarised by

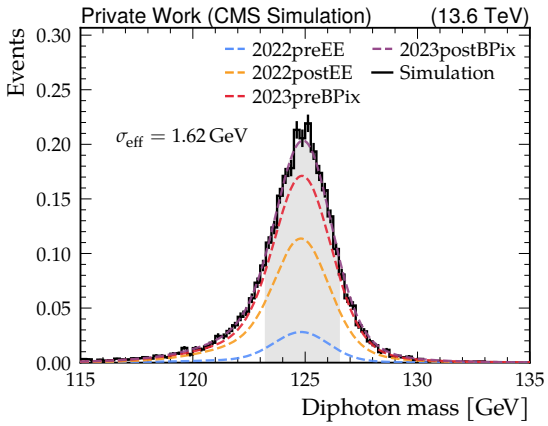
the vector $\vec{\theta}$. The signal normalisation for process p in category c and data-taking period y with integrated luminosity $\mathcal{L}_{\text{int},y}$ can then be expressed as

$$S_{c,p,y} = r_p \, \sigma_p(m_H, \vec{\theta}) \, A_{c,p}(\vec{\theta}) \, \epsilon_{c,p,y}(\vec{\theta}) \, \mathcal{L}_{\text{int},y}, \tag{9.3}$$

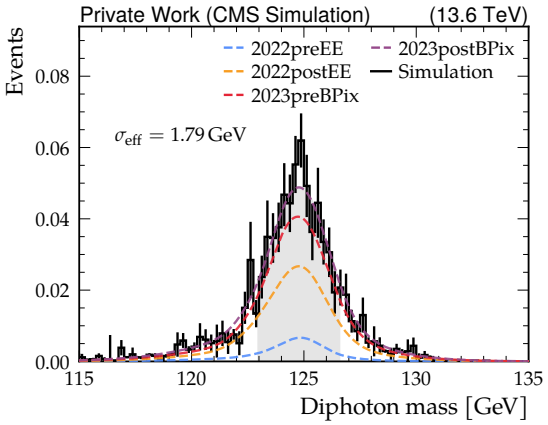
where r_p is the rate parameter relative to the nominal SM prediction, σ_p the production cross section (including its dependence on m_H), and $A_{c,p}$ and $\epsilon_{c,p,y}$ the acceptance and efficiency. The resulting parametric models, when combining all eras and production modes within example categories, are shown in Fig. 9.2. The categories which are most pure in $t\bar{t}H$, most pure in tH, and most enriched in backgrounds are shown for both the hadronic and leptonic channels. The diphoton mass resolution varies significantly between the most signal-like and background-like categories, as photon quality criteria strongly contribute to the neural network predictions as shown in Chap. 7. The smallest values are found in the most $t\bar{t}H$-dominated regions with approximately 1.6 GeV in both channels, while in the background-like categories, the effective resolution is around 2.2 GeV.



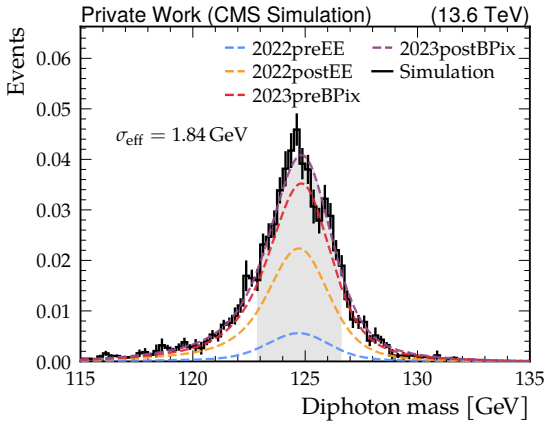
(a) $t\bar{t}H$ had. 1



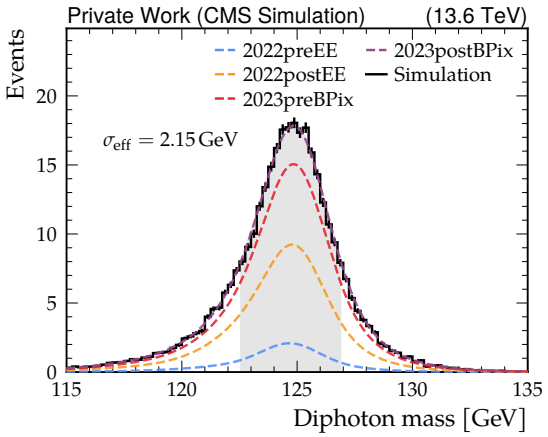
(b) $t\bar{t}H$ lep. 1



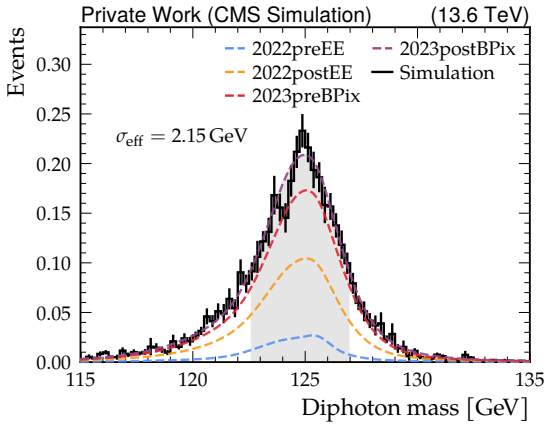
(c) tH had. 1



(d) tH lep. 1



(e) Bkg. had.



(f) Bkg. lep.

Figure 9.2: Sums of the signal and resonant background models of selected example categories. The dashed lines mark the contributions from the individual data-taking periods which are modelled individually, stacked on top of each other. The black overlaid histograms show the simulation before the parametric modelling, where the total statistical uncertainty is shown. Its magnitude is driven by the resonant background contributions. The width of the total signal model, σ_{eff}, is indicated.

Systematic uncertainties

Systematic uncertainties enter the statistical model through NPs and affect the expected signal and resonant-background yields and, where relevant, the parametric description of the diphoton-mass shapes. For the continuum background, the uncertainty associated with the choice of the analytic background function is treated separately through the discrete profiling method described in Sec. 9.2. The considered sources are grouped into theoretical and experimental uncertainties below.

Theoretical uncertainties The primary result is a fiducial cross section measurement of the considered signal modes $t\bar{t}H+tHW$ and tH. Theory uncertainties that only affect the prediction of the fiducial cross sections (e.g. the inclusive cross section and the fiducial acceptance) for these production modes are therefore reported as uncertainties on the theoretical prediction and are not included as NPs in the fit. For the signal processes, only the category-migration effects induced by such theoretical variations are included as NPs, in order to avoid double counting uncertainties on the fiducial cross sections themselves. For the resonant Higgs boson background processes, uncertainties on the corresponding total cross sections are included as separate NPs, using the values from Ref. [85], while the category-migration effects are included as separate NPs in the same way as for the signal processes. The theoretical uncertainties are treated as fully correlated across data-taking periods.

- *Renormalisation and factorisation scales:* variations of the renormalisation and factorisation scales μ_R and μ_F are used to estimate the impact of missing higher-order terms in the perturbative expansion. A three-component scheme is employed: varying μ_R by factors of 0.5 and 2.0 with μ_F fixed to its nominal value, varying μ_F by the same factors with μ_R fixed, and varying μ_R and μ_F simultaneously up and down by a factor of two. These variations are treated as independent between production processes.

- *PDF and α_s:* the uncertainty from the proton PDFs is evaluated using the NNPDF3.1 NNLO global fit with $\alpha_s(M_Z) = 0.118$ (LHAPDF set 325300) [78], following the recommendations of Ref. [248]. For acceptance and category-migration effects, the corresponding 100 PDF-weight variations are applied. The uncertainty associated with the strong coupling constant $\alpha_s(M_Z)$ is evaluated using the dedicated variations provided in the same set. The PDF- and α_s-related NPs are treated as correlated across production processes.

- *PS modelling:* the strength of initial-state radiation (ISR) and final-state radiation (FSR) is varied by changing the corresponding parameters of the PS. The resulting acceptance and category-migration effects are treated as correlated across production processes.

- *Additional heavy-flavour contributions in resonant backgrounds:* conservative normalisation uncertainties are assigned to the ggH, VBF, and VH production modes in association with additional heavy-flavour jets, partly motivated by Ref. [249]. The corresponding event weights are varied by $\pm 100\%$ and separate NPs are implemented for each process and also for bottom and charm jets. This accounts for the fact that the efficiency of these processes in the signal-sensitive phase space strongly depends on the modelling of additional heavy-flavour contributions, which are not well constrained experimentally and are possibly subject to large theoretical uncertainties.

Experimental uncertainties Experimental effects are evaluated either via event reweighting when only efficiencies are affected, or by propagating object-level variations through the full event selection and categorisation when kinematic quantities are modified. In the latter case, separate collections of selected events are produced for the corresponding variations in HiggsDNA. Diphoton mass shape uncertainties are only taken into account for uncertainties affecting the measurement of kinematic photon properties. These are the energy scale and resolution, the modelling of the material budget in front of the ECAL, and the modelling of light-collection non-uniformity as explained in the following. For these sources, alternative signal parameterisations are derived and the changes in the fitted signal-shape parameters define the corresponding NP effects.

- *Photon and electron energy scale corrections:* uncertainties are evaluated by varying the input datasets and selections used to derive the corrections with the method of Ref. [140] (ID criteria and mass requirements) and are propagated by applying the variations to the simulation samples. As a preliminary version of these corrections is used in this analysis, the photon energy-scale uncertainties are doubled and treated separately for barrel and endcap superclusters. This inflation is intended to cover additional uncertainties associated with transferring a calibration derived from electrons reconstructed as photons to genuine photons.

- *ECAL longitudinal response non-uniformity:* an additional uncertainty in the photon energy scale arises from the imperfect modelling of the longitudinal non-uniformity of the ECAL crystal response (front non-uniformity (FNUF)) [250]. Since the photon energy scale corrections are derived from electrons in $Z \to e^+e^-$ events, residual mismodelling of the depth dependence of light collection can induce a small bias when transferring the calibration from electrons to photons. This effect is propagated as a correlated photon energy-scale NP affecting the signal mass shape.

- *Material upstream of the ECAL*: The modelling of the tracker material budget influences the predicted conversion probability of photons and thus the mixture of converted and unconverted candidates entering the analysis. Since these categories exhibit different energy response and resolution, imperfect material modelling induces a residual uncertainty in the diphoton mass scale and resolution. This uncertainty was assessed by varying the upstream material description in simulation in previous measurements and the impact on photon p_{T} is propagated to the signal mass-shape parameterisation.

- *Pileup modelling:* the expected pileup distribution is varied by changing the total inelastic proton-proton interaction cross section used to derive the expected pileup distribution, where a value of (69.2 ± 3.2) mb is used conventionally in the CMS Collaboration. The event weights are derived by reweighting the simulated true interaction multiplicity distribution to match the nominal and varied pileup distributions, and the resulting variations in the expected yields are implemented as NPs correlated between data-taking periods.

- *Integrated luminosity:* the luminosity uncertainty is implemented as a global normalisation uncertainty affecting all simulated processes. A reduced Run-3 correlation scheme is used, implemented through two NPs: one component affecting the 2022 data with 1.38% [251] and the 2023 data with 0.17%, and a second component affecting only the 2023 data with 1.27%. This corresponds to a partially correlated treatment of the luminosity uncertainty between the 2022 and 2023 data-taking periods.

- *Diphoton trigger and photon preselection efficiencies:* uncertainties are evaluated with tag-and-probe studies in $Z \to e^+e^-$ events by varying the signal and background models used in the efficiency extraction.

- *Electron-veto efficiency:* the efficiency of the electron-veto requirement is measured using $Z \to \mu^+\mu^-\gamma$ events. The corresponding uncertainties are dominated by the finite sample sizes and are treated as uncorrelated between data-taking periods.

- *Lepton efficiencies and calibrations:* electron reconstruction and ID efficiencies and their uncertainties are derived from tag-and-probe studies in $Z \to e^+e^-$ events. Muon ID and isolation efficiencies and their uncertainties, as well as momentum scale and resolution uncertainties, are derived from $Z \to \mu^+\mu^-$ events.

- *Missing transverse momentum:* The uncertainty in the unclustered component of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is evaluated from the momentum resolution of the PF candidates contributing to that component, with separate treatments for the different PF candidate types. It is dominated by neutral hadrons and PF candidates from the forward hadron calorimeter [156]. It is therefore distinct from the dedicated uncertainties assigned to jets, photons, electrons, and muons, for which $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is recalculated and included in the corresponding NPs.

- *Jet energy scale and resolution:* jet energy scale and resolution uncertainties are taken from Ref. [146]. An implementation is used which employs one NP summarising all sources of the total jet energy scale uncertainties and one NP for the jet energy resolution. These are treated as uncorrelated with each other and uncorrelated between data-taking periods.

- *b-tagging:* uncertainties on b-tagging discriminator-shape scale factors are evaluated in a $t\bar{t}$-enriched control sample. The corresponding NPs are split into components associated with the jet energy scale dependence of the scale-factor measurement, heavy-/light-flavour contamination, statistical components (including year-dependent components), and additional components related to the charm-jet fraction.

9.2. Continuum background model

In contrast to the signal model, the continuum background model is determined inclusively over all eras to ensure sufficient event counts for the parametric fits. For each category, analytic functions are fitted to the observed $m_{\gamma\gamma}$ spectrum in data over a broad window of 100–180 GeV. To avoid biases from the choice of a specific analytic background shape, the discrete profiling method [244] is used, following the common practice in measurements performed in the $H \to \gamma\gamma$ decay channel at the CMS experiment. For each category, multiple function families and orders are tested and the choice of specific functions is treated as a discrete NP that is profiled in the likelihood fits. The candidate families considered are exponentials, power laws, Bernstein polynomials, and Laurent series. Their implemented functional forms are summarised in Tab. 9.1.
The functions entering the envelope for the discrete profiling are selected as follows: For each family and category, a sequence of probability density functions (pdfs) with increasing functional complexity is constructed. The exact mapping between the order and the number of free shape parameters depends on the function family. For the selection of functions entering the envelope, fits are performed in the sidebands, excluding the 120–130 GeV signal window.

Table 9.1: Candidate analytic function families for the background model in the form implemented in the fit framework. Here, $x \equiv (\hat{m}_{\gamma\gamma} - \hat{m}_{\min})/(\hat{m}_{\max} - \hat{m}_{\min}) \in [0,1]$ denotes the mass variable mapped to the fit range. For the exponential and power-law families, the implemented order N is odd and corresponds to $n = (N+1)/2$ basis components. An overall normalisation factor is omitted for each family.

Family	Implemented form $f(\hat{m}_{\gamma\gamma})$
Exponential family (odd order $N = 2n-1$)	$f_N(\hat{m}_{\gamma\gamma}) = \sum_{i=1}^{n} c_i \exp(\lambda_i \hat{m}_{\gamma\gamma})$, with $\lambda_i < 0$, $c_i > 0$, $\sum_{i=1}^{n} c_i = 1$
Power-law family (odd order $N = 2n-1$)	$f_N(\hat{m}_{\gamma\gamma}) = \sum_{i=1}^{n} c_i \, \hat{m}_{\gamma\gamma}^{\alpha_i}$, with $c_i > 0$, $\sum_{i=1}^{n} c_i = 1$
Bernstein family	$f_N(x) \propto \sum_{i=0}^{N} b_i \binom{N}{i} x^i (1-x)^{N-i}$, with $b_0 = 1$, $b_i > 0$ for $i = 1, \ldots, N$
Laurent family	$f_N(\hat{m}_{\gamma\gamma}) = c_0 \, \hat{m}_{\gamma\gamma}^{-4} + \sum_{i=1}^{n_\ell} c_i^{(-)} \, \hat{m}_{\gamma\gamma}^{-4-i} + \sum_{j=1}^{n_h} c_j^{(+)} \, \hat{m}_{\gamma\gamma}^{-4+j}$, with $n_\ell = \lceil N/2 \rceil$, $n_h = N - n_\ell$, and all coefficients positive

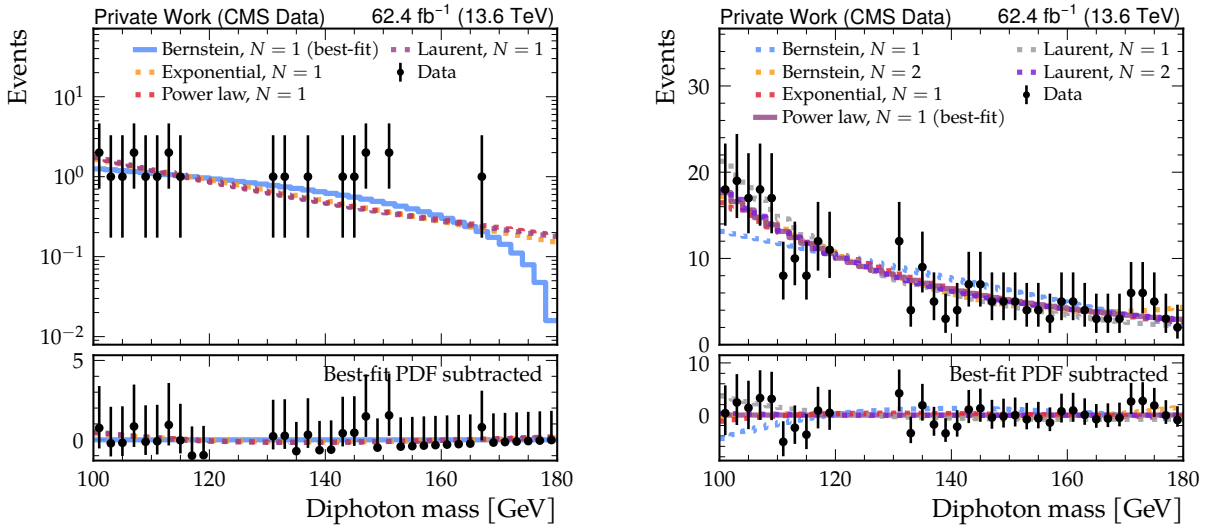
The minimum order is determined by a likelihood-ratio test. For two candidate functions within the same family with increasing complexity, the test statistic is given by the difference in the negative log-likelihoods of the fits to data in the sidebands,

$$q = -2 \left(\ln \mathcal{L}_{\text{simple}} - \ln \mathcal{L}_{\text{complex}} \right) . \tag{9.4}$$

In the asymptotic approximation, q is evaluated using a χ^2 distribution with a number of degrees of freedom equal to the difference in the number of free parameters Δn_{par} between the two models. The order is increased only while the additional complexity yields a statistically significant improvement, $p = \text{Prob}(q, \Delta n_{\text{par}}) < 0.05$. The scan stops once $p \geq 0.05$ or a family-dependent practical upper order is reached. For the envelope construction, the scan is repeated with a looser threshold $p < 0.1$. In the current implementation, Bernstein polynomials are tested up to order 5 and Laurent series up to order 6, while the exponential and power-law families are implemented only for odd orders and are therefore scanned as $N = 1, 3, 5$.

For each eligible order, a sideband goodness-of-fit p-value is computed, either asymptotically assuming a χ^2 distribution or using toys when the average number of events per bin is below five, which is the case in most categories in this analysis. Candidate envelope members are retained only if their sideband goodness-of-fit satisfies GOF > 0.01, while the order selected by the tighter likelihood-ratio criterion is always kept for the corresponding family in order to avoid an unphysically narrow envelope. The order-selection test itself still uses the asymptotic Wilks χ^2 approximation in low-statistics categories and is therefore heuristic. In practice, this procedure retains at least one representative from each candidate family.

Technically, the envelope is implemented as a collection of indexed probability density functions (pdfs) in a `RooMultiPdf` format, and the final fit profiles over the corresponding discrete



(a) $t\bar{t}H$ had. 1

(b) $t\bar{t}H$ had. 4

Figure 9.3: Selected background fit functions entering the discrete-profiling envelope in two representative hadronic categories. The order N of the functions is displayed and the best-fit function is highlighted. The data in the signal window from 120–130 GeV are not shown. The lower panel shows the data points with the best fit function subtracted.

pdf index. Within the envelope, one function per category is labelled as the best-fit pdf by minimising the penalised score $-2 \ln \mathcal{L} + N_B$, where N_B is the number of background parameters. This best-fit choice is used for visualisation and to initialise the background normalisation, while all envelope members remain part of the multi-pdf construction and are profiled as discrete alternatives in the final fit, with the same penalty applied to discourage overly complex shapes.

Two representative examples of the selected background-fit functions entering the discrete-profiling envelope are shown in Fig. 9.3. The categories *$t\bar{t}H$ had. 1* and *$t\bar{t}H$ had. 4* are chosen to illustrate the behaviour at different background levels. The corresponding plots for the remaining categories are collected in Appendix C.

9.3. Likelihood framework and inference

The inference of the POIs is based on a likelihood function $\mathcal{L}(\vec{r}, \vec{\theta})$, which models the probability of observing the data given a set of POIs $\vec{r}$ and NPs $\vec{\theta}$. The data entering the fit are the binned diphoton mass spectra in the individual analysis categories. The fit therefore corresponds to a binned extended maximum-likelihood fit to the $m_{\gamma\gamma}$ spectrum, where the narrow Higgs boson signal is extracted on top of a smooth continuum background. The observed event counts in the different categories and diphoton mass bins are modelled as conditionally independent Poisson counting experiments given the POIs and nuisance parameters. The likelihood therefore factorises over categories c and diphoton mass bins b, with correlations from systematic uncertainties introduced through the shared nuisance parameters, and can be schematically expressed as

$$\mathcal{L}(\vec{r}, \vec{\theta}) = \prod_c \prod_b \mathrm{Pois}\left(n_{c,b} \,\middle|\, \nu_{c,b}(\vec{r}, \vec{\theta})\right) \times \prod_j \pi_j(\theta_j). \tag{9.5}$$

Here, $n_{c,b}$ denotes the number of observed events in bin b of category c, $\nu_{c,b}$ the corresponding expected event yield, and $\pi_j(\theta_j)$ the constraint term associated with NP θ_j. The Poisson factor $\text{Pois}(n_{c,b} \mid \nu_{c,b})$ describes the probability of observing $n_{c,b}$ events for a given expectation value $\nu_{c,b}$.
The expected event yield in each bin can be decomposed as

$$\nu_{c,b}(\vec{r},\vec{\theta}) = \sum_{p\in\{tH,\, t\bar{t}H+tHW\}} \left[S^{\text{in}}_{c,b,p}(r_p,\vec{\theta}) + S^{\text{OOA}}_{c,b,p}(\vec{\theta})\right] + B_{\text{res.},c,b}(\vec{\theta}) + B_{\text{cont.},c,b}(\vec{\theta}), \tag{9.6}$$

where $S^{\text{in}}_{c,b,p}$ denotes the in-fiducial signal contribution for process p, $S^{\text{OOA}}_{c,b,p}$ the corresponding out of acceptance (OOA) contribution, $B_{\text{res.},c,b}$ the resonant Higgs boson background from production modes other than tH and $t\bar{t}H+tHW$, and $B_{\text{cont.},c,b}$ the continuum background contribution. For the two measured signal processes, only the contributions that pass the fiducial phase space selection scale with the corresponding POIs, while the OOA contributions are fixed to their SM expectation, or to the corresponding alternative coupling hypothesis where explicitly stated. The expected bin yields are obtained by integrating the corresponding parametric signal and background probability density functions over each mass bin.
The terms $\pi_j(\theta_j)$ are auxiliary terms in the likelihood and encode the effect of experimental and theoretical systematic uncertainties. For the continuous NPs considered in this analysis, these auxiliary constraint terms are Gaussian in the NPs themselves. Uncertainties that affect only the overall event yields per category and process are implemented as multiplicative variations of the expected rates, which ensures a positive prediction and corresponds to a log-normal treatment of the yield uncertainty. By contrast, the NPs that affect the signal-shape model enter the parametrisation of the diphoton mass distribution directly and are constrained around their nominal values by Gaussian terms.
Statistical inference is performed using likelihood-ratio test statistics. Unless stated otherwise, the definitions and notation in this subsection follow Ref. [241].

Confidence intervals

For simple hypotheses, likelihood-ratio tests are optimal in the sense of the Neyman–Pearson lemma [252]. In the presence of composite hypotheses and NPs, the profile-likelihood ratio remains a standard choice and admits well-controlled asymptotic behaviour, which is exploited throughout this analysis following Ref. [241]. For given POIs $\vec{r}$, the profile-likelihood ratio test statistic is defined as

$$t_{\vec{r}} = -2\ln\frac{\mathcal{L}\left(\vec{r},\,\hat{\hat{\theta}}_{\vec{r}}\right)}{\mathcal{L}\left(\hat{\vec{r}},\,\hat{\theta}\right)} \equiv -2\ln\lambda(\vec{r}) \equiv -2\Delta\ln\mathcal{L}(\vec{r}), \tag{9.7}$$

where $\left(\hat{\vec{r}},\hat{\theta}\right)$ denotes the unconditional maximum-likelihood estimators and $\hat{\hat{\theta}}_{\vec{r}}$ denotes the conditional maximum-likelihood estimator of the NPs for fixed $\vec{r}$. Profiling over $\vec{\theta}$ propagates the impact of systematic uncertainties into the inferred POIs and typically broadens the likelihood as a function of $\vec{r}$.
In one-dimensional scans, one POI is varied on a grid while the other POI and all NPs are profiled at each scan point. Concretely, when scanning r_{tH} ($r_{t\bar{t}H+tHW}$), the parameter $r_{t\bar{t}H+tHW}$ (r_{tH}) and all NPs are re-optimised at each grid point. Under the asymptotic approximation, t_r follows

a χ^2 distribution with one degree of freedom (Wilks' theorem) [253], such that $-2\Delta \ln \mathcal{L} = 1.0$ and 3.84 define the boundaries of the 68 % and 95 % confidence intervals, respectively. In two-dimensional scans in $(r_{tH}, r_{t\bar{t}H+tHW})$, the corresponding asymptotic distribution has two degrees of freedom, and the 68 % and 95 % confidence regions are obtained from $-2\Delta \ln \mathcal{L} = 2.30$ and 5.99, respectively.

For the fiducial cross-section measurement, the POIs are not constrained to be non-negative in these likelihood scans. This avoids a bias of the fitted central values in regions of limited sensitivity, where downward fluctuations of the data would otherwise accumulate at the physical boundary at zero, while upward fluctuations remain unrestricted. Allowing negative values therefore leads to better-behaved profile-likelihood scans and a more regular confidence-interval construction. Such negative values are not interpreted as physical negative cross sections, but only as unconstrained maximum-likelihood estimators. The scan domain is nevertheless restricted to points for which all expected event yields remain non-negative, since the Poisson likelihood is not defined otherwise. By contrast, for the significance and upper-limit constructions discussed below, the physical constraint of non-negative signal rates is imposed explicitly.

Signal significance

The signal significance quantifies the degree to which the observed data are incompatible with the background-only hypothesis. For this purpose, a one-sided test statistic is used that enforces the physical constraint of a non-negative rate parameter, $r \geq 0$. Following Ref. [241], a modified profile-likelihood ratio is defined as

$$\tilde{\lambda}(r) = \begin{cases} \dfrac{\mathcal{L}\left(r, \hat{\hat{\theta}}(r)\right)}{\mathcal{L}\left(\hat{r}, \hat{\theta}\right)}, & \hat{r} \geq 0 \\ \dfrac{\mathcal{L}\left(r, \hat{\hat{\theta}}(r)\right)}{\mathcal{L}\left(0, \hat{\hat{\theta}}(0)\right)}, & \hat{r} < 0 \end{cases}, \tag{9.8}$$

and the corresponding test statistic is $\tilde{t}_r = -2 \ln \tilde{\lambda}(r)$. The test statistic for the background-only hypothesis ($r = 0$) is then

$$q_0 = \tilde{t}_0 = \begin{cases} -2 \ln \lambda(0), & \hat{r} \geq 0 \\ 0, & \hat{r} < 0 \end{cases}, \tag{9.9}$$

which ensures that downward fluctuations of the fitted rate parameter do not generate an apparent excess. The observed value $q_{0,\text{obs}}$ is converted to a p-value,

$$p_0 = \int_{q_{0,\text{obs}}}^{\infty} f(q_0 \mid H_0)\, dq_0, \tag{9.10}$$

where $f(q_0 \mid H_0)$ denotes the probability density of q_0 under the background-only hypothesis. The p-value is reported as an equivalent one-sided Gaussian significance,

$$Z = \Phi^{-1}(1 - p_0), \tag{9.11}$$

with Φ^{-1} the inverse cumulative distribution function of a standard normal distribution.

Upper limits

Upper limits are constructed using a one-sided test statistic derived from the modified profile-likelihood ratio in Eq. (9.8), again under the constraint $r \geq 0$,

$$\tilde{q}_r = \begin{cases} -2\ln\tilde{\lambda}(r), & r \geq \hat{r} \\ 0, & r < \hat{r} \end{cases}, \tag{9.12}$$

which sets the test statistic to zero when the tested rate parameter is below the fitted one, such that no incompatibility is assigned in that regime. In principle, an upper limit could be obtained by determining the value of r for which the corresponding p-value falls below 0.05. However, in the presence of limited experimental sensitivity, downward fluctuations of the data may lead to overly stringent exclusions. To mitigate this behaviour, the modified frequentist CL_s prescription [254, 255] is used. The quantities p_{s+b} and $1 - p_b$ denote the upper-tail probabilities of $\tilde{q}_r$ under the signal-plus-background and background-only hypotheses, respectively. The modified CL is defined as

$$\mathrm{CL}_s = \frac{p_{s+b}}{1 - p_b}. \tag{9.13}$$

The upper limit at 95 % CL is defined as the rate-parameter value for which $\mathrm{CL}_s = 0.05$. This construction corresponds to the standard CL_s upper limit implemented in the Combine framework and follows the conventions widely adopted in searches for the Higgs boson by the ATLAS and CMS Collaborations [246, 256].
For the test statistics defined above, the probability densities under the relevant hypotheses are evaluated using the asymptotic approximations of Ref. [241]. This avoids the computational cost of toy MC ensembles while providing accurate results in the large-sample limit.

Goodness of fit

To quantify the global compatibility of the fitted statistical model with the observed data, a goodness-of-fit test is performed. A test statistic comparing the observed binned diphoton mass spectra with the corresponding post-fit expectation is evaluated across all categories. Its distribution under the fitted model is obtained from pseudo-data generated from the post-fit model, and the resulting p-value is defined as the probability to obtain a discrepancy at least as large as that observed in data.

9.4. Robustness against assumptions on the top-quark Yukawa coupling

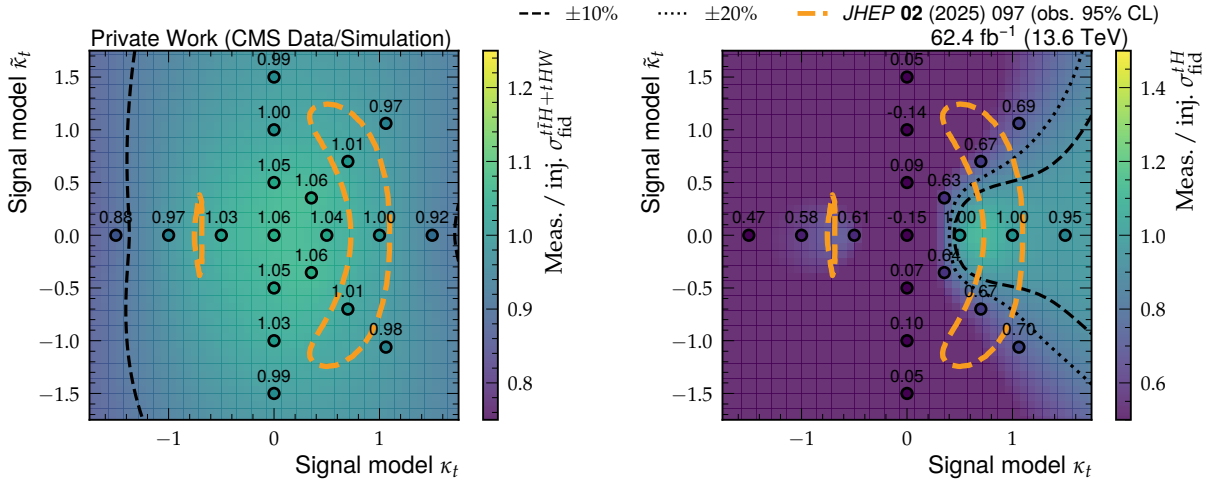
The extraction of the fiducial tH and $t\bar{t}H$ cross sections relies on signal templates that depend on the assumed structure of the top-quark Yukawa interaction, parameterised by the coupling modifiers $(\kappa_t, \tilde{\kappa}_t)$. To assess the stability of the measurement with respect to these assumptions, dedicated robustness studies are performed with Asimov datasets in the asymptotic likelihood framework.
In a first study, the Asimov dataset is generated according to the SM expectation, while the signal templates used in the fit are varied across the $(\kappa_t, \tilde{\kappa}_t)$ plane. The ratio of the fitted to the injected fiducial cross section is used to quantify potential biases for $t\bar{t}H{+}tHW$ and tH. The coupling plane is sampled on a polar grid consisting of 19 points, defined by radii 0.5, 1.0, and

1.5, and by azimuthal angles 0, $\pi/4$, $\pi/2$, π, $3\pi/2$, and $7\pi/4$, as well as the point with vanishing coupling $(0, 0)$. The discrete results obtained at these points are interpolated using radial basis functions to obtain a smooth representation of the bias across the plane. Technically, the underlying signal models are constructed as follows. For $t\bar{t}H$, detector-level templates corresponding to general coupling hypotheses are obtained by superposing the simulation samples with the SM and purely CP-odd top-quark Yukawa couplings according to the parameterisation from Eq. (5.1). For tH, dedicated simulation samples are available for coupling points on the unit circle and for vanishing Yukawa coupling. For intermediate points, the expected yield is modelled with a second-order polynomial of the coupling modifiers, separately for each category. The SM sample is reweighted to match this yield, hence the fraction of in-fiducial over OOA events matches the SM prediction for the intermediate configurations, which is validated to introduce only negligible sub-percent level biases. A similar procedure is applied for tHW, where event-level reweighting factors available in the LO `MadGraph` simulation are used to determine the polynomial expression.

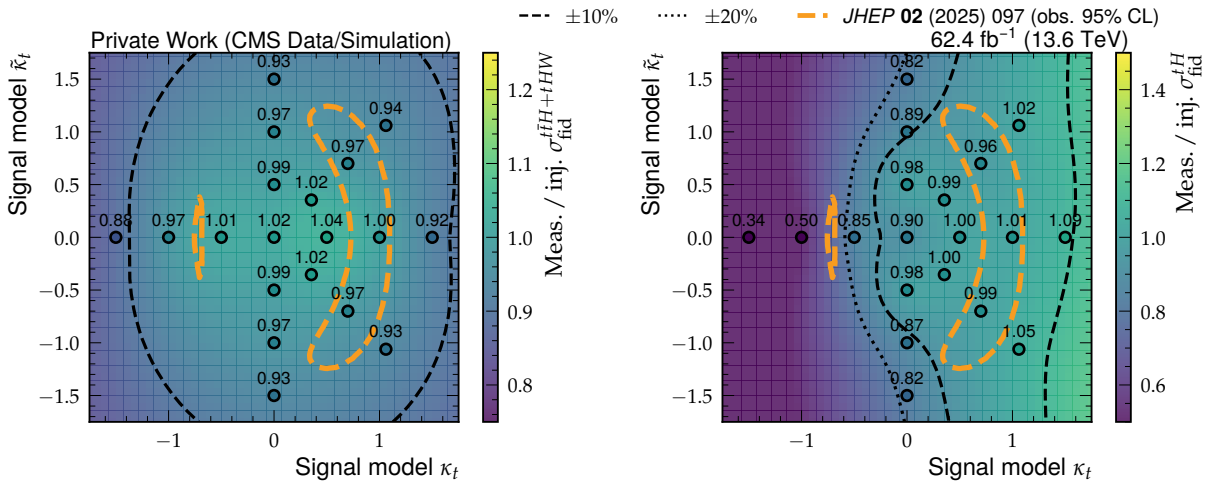
The resulting biases in the best-fit POIs as a function of the coupling modifiers are shown in Fig. 9.4. Two setups are compared: a categorisation based on neural networks trained exclusively on SM samples without explicit CP decorrelation, and the networks trained with adversarial decorrelation together with the categorisation optimisation including a CP-sensitive penalty term, as introduced in Chap. 7 and Chap. 8.

In the SM-trained categorisation, the extracted $t\bar{t}H$+tHW cross section remains stable at the level of a few percent across the coupling plane. In contrast, the tH cross section exhibits pronounced biases. For example, in a CP-admixed scenario with $(\kappa_t, \tilde{\kappa}_t) = (1/\sqrt{2}, \pm 1/\sqrt{2})$, the fitted tH cross section is reduced by approximately one third with respect to the injected value. In the purely CP-odd scenario $(0, \pm 1)$, the fitted tH cross section becomes compatible with zero or slightly negative. These effects originate from changes in the category composition across the coupling plane, in particular from a migration of $t\bar{t}H$ events into tH-sensitive categories. When such signal models are fitted to an SM Asimov dataset, which predicts fewer events in these regions, the fit achieves compatibility between the model and the Asimov dataset by reducing the extracted tH normalisation.

The decorrelated categorisation shows a markedly improved robustness. The $t\bar{t}H$+tHW cross section remains stable at a level comparable to the SM-trained setup, while the biases in the tH cross section are substantially reduced. Residual deviations arise from two effects. First, there remains a mild dependence of the per-category acceptance and efficiency on the coupling hypothesis, although significantly weaker than in the non-decorrelated case. A second source of biases is induced by assumptions on the OOA events, which are treated as background-like contributions. These contributions reach 5–10% of the expected yields in individual detector-level categories and exhibit shapes similar to the respective signal. In scenarios with strongly enhanced total cross sections, such as the inverted-sign case, this contribution is similarly scaled up as the fiducial contribution, leading to a downward bias when fitting only the in-fiducial component against an SM Asimov dataset. For CP-admixed scenarios, the resulting biases remain at the level of a few percent in the decorrelated setup. For the CP-odd and inverted-sign assumptions, the biases increase to approximately 10% and 50%, respectively. Importantly, for the half-plane with positive κ_t, which is of primary phenomenological interest as it contains the majority of the parameter space compatible with a combination of previous measurements using the CMS Run-2 data [257], the biases remain below 10% everywhere in the decorrelated setup for both processes, while the corresponding low-bias region is considerably smaller in the SM-trained categorisation, particularly for tH.



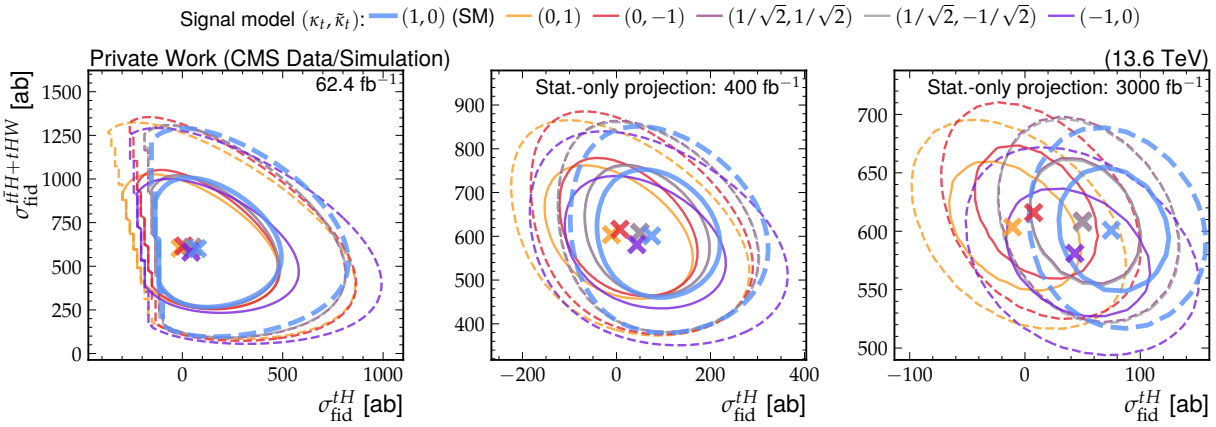
(a) SM-trained categorisation (comparison setup)



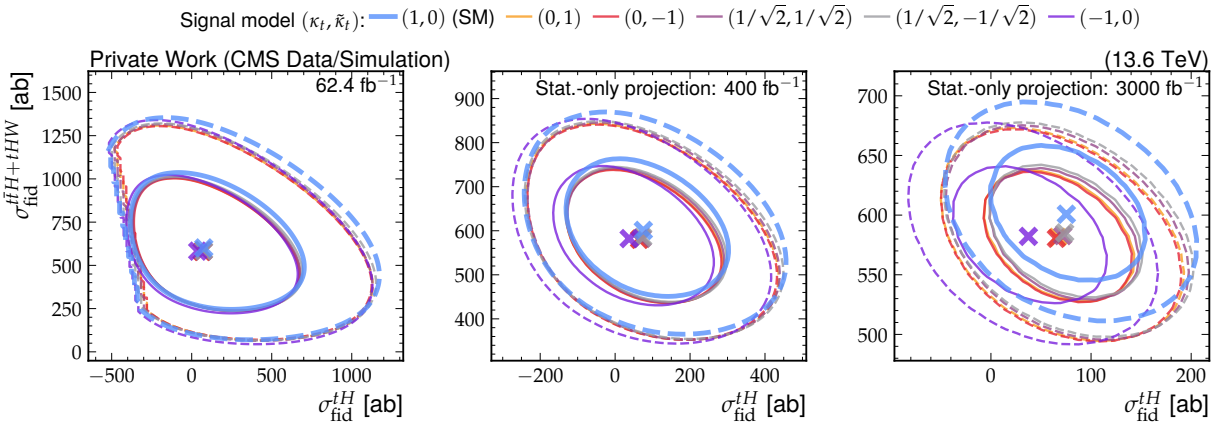
(b) Decorrelated categorisation (nominal setup)

Figure 9.4: Robustness study with fixed SM Asimov dataset: the ratio of the measured to the injected fiducial cross section is shown for different signal models used across the $(\kappa_t, \tilde{\kappa}_t)$ plane. The top row shows the results for the SM-trained categorisation, while the bottom row shows the results for the decorrelated categorisation. The left and right panels show the results for $t\bar{t}H{+}tHW$ and tH cross sections, respectively. The coloured heatmap shows an interpolation obtained with radial basis functions based on the discrete results at the sampled coupling points, which are indicated by the overlaid numbers. The black contours correspond to interpolated biases of 10% and 20% for the respective process. For comparison, the parameter space compatible with a combination of previous measurements with the CMS experiment at 95% CL is overlaid with the orange dashed contours [257].

While the discussion above focuses on biases of the extracted central values, the phenomenological relevance of signal-model assumptions depends strongly on the experimental precision of the measurement. This is illustrated in Fig. 9.5, which shows the expected two-dimensional confidence regions in the fiducial cross-section plane obtained from fitting the SM Asimov dataset with signal models of different coupling assumptions and including only statistical uncertainties. On the far negative-r_{tH} side, the scans are truncated once the expected yield in at least one category would become negative. This typically occurs first in the low-background *tH lep. 1* region, so the outer contours in that direction terminate at the non-negativity boundary of the statistical model rather than being closed by the likelihood threshold.
At the current integrated luminosity of 62.4 fb^{-1}, the dominant effect of varying the signal model is a change in the extent and orientation of the confidence regions rather than a statistically significant shift of the best-fit point. In the categorisation based on SM-trained neural networks without decorrelation, the outer edges of the 68 % and 95 % confidence regions vary by approximately 20% across coupling assumptions. In contrast, in the decorrelated categorisation, the corresponding variation is reduced to about 5%. This demonstrates that, at the present level of experimental precision, the decorrelated setup yields confidence intervals that are substantially more stable with respect to assumptions on the top-quark Yukawa coupling. The same conclusion is obtained for the median expected CL_s upper limit on the fiducial tH cross section in the same fits with only statistical uncertainties included. In the decorrelated categorisation, replacing the SM signal model by the purely CP-odd scenario changes the expected limit by -3.1%, while the inverted-sign scenario changes it by only $+0.6\%$. In the SM-trained comparison setup, the corresponding expected limits increase by about 18% and 24%, respectively, confirming that the decorrelation also stabilises the search sensitivity used for the quoted tH upper limit.
For larger datasets, the interpretation changes qualitatively. As the statistical uncertainty decreases, variations in the size of the confidence regions become subdominant, and the focus shifts such that the bias of the extracted central values relative to the expected uncertainty becomes significant. This regime is illustrated by the projections to 400 fb^{-1}, representative of the full Run-3 programme, and to 3000 fb^{-1}, corresponding to the High-Luminosity LHC (HL-LHC) expectation. In the SM-trained categorisation, shifts of the best-fit point reach the level of approximately 1–2 standard deviations for some coupling scenarios at 3000 fb^{-1}. In contrast, the decorrelated categorisation remains stable with tensions below about 0.5 standard deviations. In this high-precision regime, the reduced bias of the central values and the more reliable coupling interpretation become the primary motivation for the decorrelated approach.
Analogous studies based on CP-odd and inverted-sign Asimov datasets are presented in Appendix B and lead to the same qualitative conclusion.



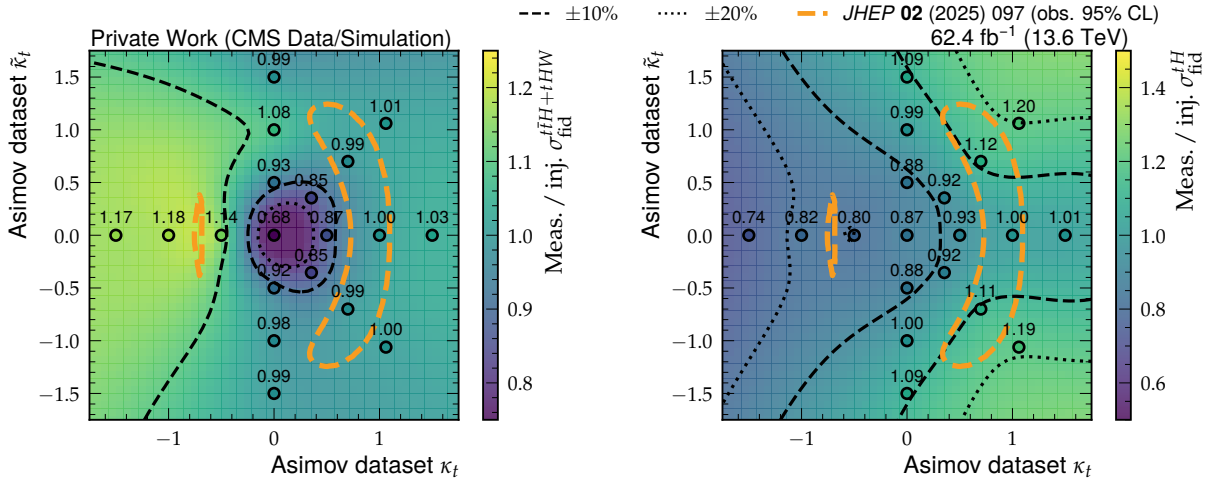
(a) No decorrelation (comparison setup)



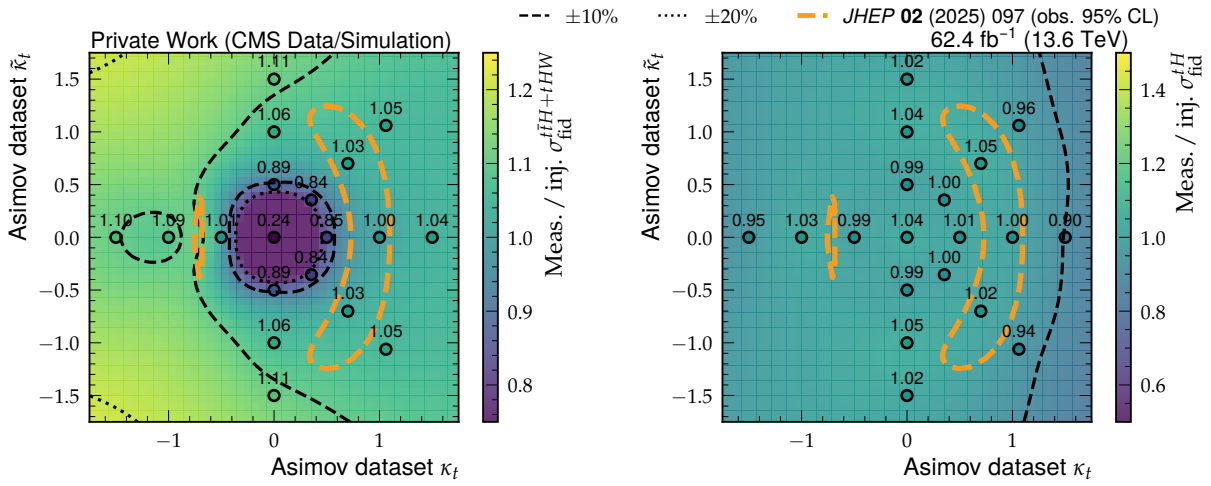
(b) Decorrelated categorisation (nominal setup)

Figure 9.5: Expected two-dimensional confidence regions in the fiducial cross-section plane obtained from an SM Asimov dataset and signal models of different coupling scenarios. The solid and dashed contours represent the 68 % and 95 % confidence regions, respectively, while the cross markers indicate the best-fit points. The scans are shown for integrated luminosities of 62.4 fb^{-1} (this analysis), 400 fb^{-1}, and 3000 fb^{-1}, corresponding to expectations for the full Run-3 and the HL-LHC data taking. The upper row shows the results for the categorisation based on neural networks trained on SM signal samples without explicit *CP* decorrelation, while the lower row shows the corresponding results for the nominal decorrelated categorisation. Only statistical uncertainties are included in the fits.

A complementary robustness test is performed by fixing the signal model to the SM configuration, as done in conventional measurements, and generating Asimov datasets for different coupling hypotheses across the $(\kappa_t, \tilde{\kappa}_t)$ plane. The resulting best-fit cross sections are shown in Fig. 9.6 for both setups. In this configuration, the biases observed for $t\bar{t}H{+}tHW$ remain small across the plane, while the deviations for tH are generally smaller than in the previous study, where the signal model was varied in the fits. For the SM-trained categorisation, the biases appear smaller than in the above study, which can be attributed to migrations of $t\bar{t}H$ into the tH-sensitive categories and vice versa as was discussed in Chap. 8, which lead to a partial compensation of the bias when fitting an SM signal model against an Asimov dataset generated with a different coupling hypothesis. The biases in the decorrelated setup are smaller than in the above study, because the SM point is close to the minimum of the tH cross section, such that large enhancements with the associated OOA effect are absent in the signal model used. Within the region of coupling space not excluded by Run 2 measurements at 95 % CL, the biases reach up to approximately 20% in the SM-trained categorisation and remain below 5% in the decorrelated setup.
Overall, the decorrelated categorisation exhibits a significantly improved robustness against assumptions on the top-quark Yukawa coupling, at the cost of a reduced sensitivity to the tH signal of approximately 25% in the SM scenario. In CP-odd and inverted-sign scenarios, the sensitivity is comparable to that of the SM-trained setup while showing no relevant bias with respect to the signal model choice. The sensitivity to $t\bar{t}H$ production is largely unaffected by the decorrelation procedure across the coupling plane. In view of the substantially reduced model dependence, the decorrelated categorisation is adopted for the extraction of the fiducial cross sections from the observed data in the following.



(a) SM-trained categorisation (comparison setup)



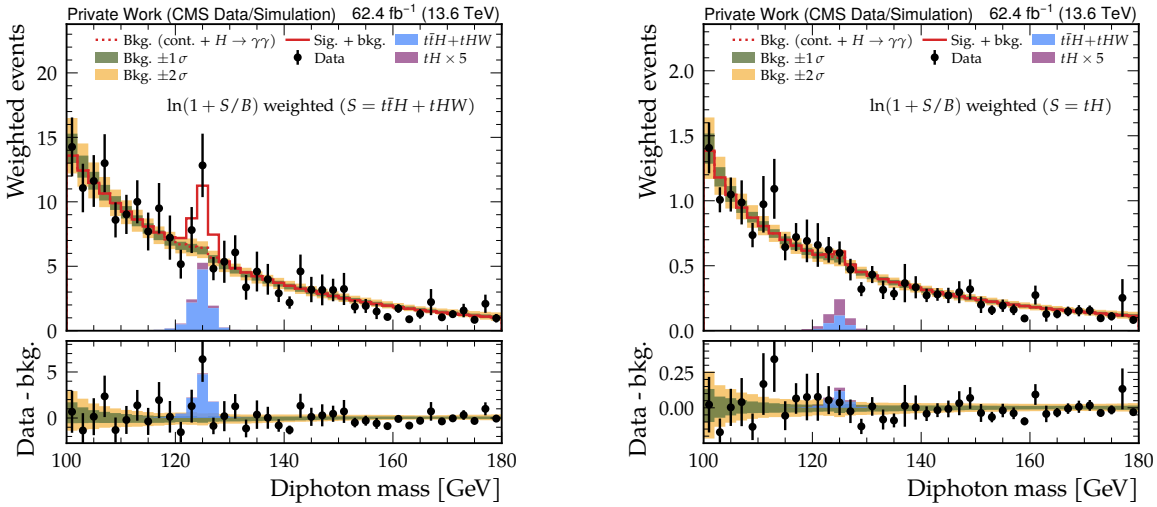
(b) Decorrelated categorisation (nominal setup)

Figure 9.6: Best-fit fiducial cross sections obtained when the signal model used in the fit is fixed to the SM hypothesis, while the Asimov dataset is generated for different points in the $(\kappa_t, \tilde{\kappa}_t)$ plane. The upper panel shows the results for the categorisation based on neural networks trained on SM signal samples without explicit CP decorrelation, while the lower panel shows the corresponding results for the nominal decorrelated categorisation. The left and right panels show the fitted $t\bar{t}H{+}tHW$ and tH cross sections, respectively. The coloured heatmap shows an interpolation obtained from the discrete sampled coupling points indicated by the overlaid numbers. For comparison, the parameter space compatible with a combination of previous measurements with the CMS experiment at 95% CL is overlaid with the orange dashed contours [257].

9.5. Results from observed data

The results from applying the above methodology to fit the combined signal and background model to the observed data are documented in the following. In all profile likelihood fits performed, the SM signal model and the decorrelated categorisation are used. All 16 categories are included. The contributions from other Higgs boson production modes are fixed to their SM predictions, with NPs included to account for the corresponding cross-section and modelling uncertainties.

A compact visualisation of the post-fit agreement between data and the model is provided in Fig. 9.7 by combining categories with $\ln(1 + S/B)$ weights. Here, S and B denote the expected signal and background yields in a narrow signal window in each category, where S is taken to be either the $t\bar{t}H{+}tHW$ or the tH contribution, depending on the weighting shown. The observed data are in good agreement with the post-fit model as summarised with these two weighted spectra, which are designed to highlight the presence of the signal contributions in the combined fit. The full set of unweighted, per-category spectra of the observed data is provided in Appendix D.



(a) Weighted for $t\bar{t}H + tHW$

(b) Weighted for tH

Figure 9.7: Weighted observed diphoton mass spectra (top panel) and background-subtracted residuals (bottom panel), where categories are combined using $\ln(1 + S/B)$ weights and S and B denote the expected signal and background yields in a narrow signal window per category. The red dotted line denotes the post-fit background model, and the red solid line shows the total post-fit model of the signals and all backgrounds. The filled bands around the background model indicate the one- and two-standard-deviation post-fit uncertainty bands on the background prediction, shown in green and yellow, respectively. They are obtained by drawing 500 toys per category from the best-fit background model; the envelope of the fitted background models is shown. The overlaid signal shapes for $t\bar{t}H{+}tHW$ (tH) show the post-fit shapes, but are normalised to $r_p = 1$ in the lower residual panels, while the tH signal contribution is scaled to $r_{tH} = 5$ in the upper right panel to improve visibility.

The primary result of the measurement is the two-dimensional profile-likelihood scan in the fiducial cross-section plane shown in Fig. 9.8, which provides the best-fit point and the 68 %

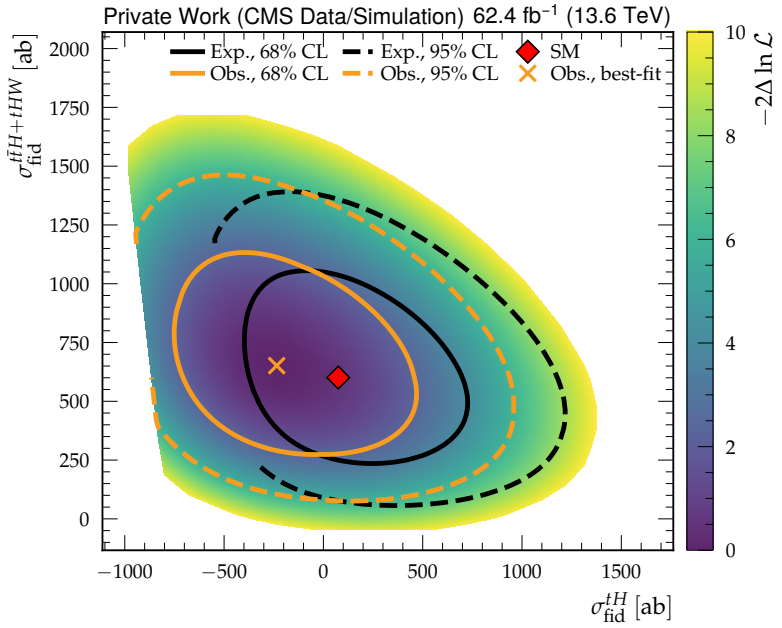


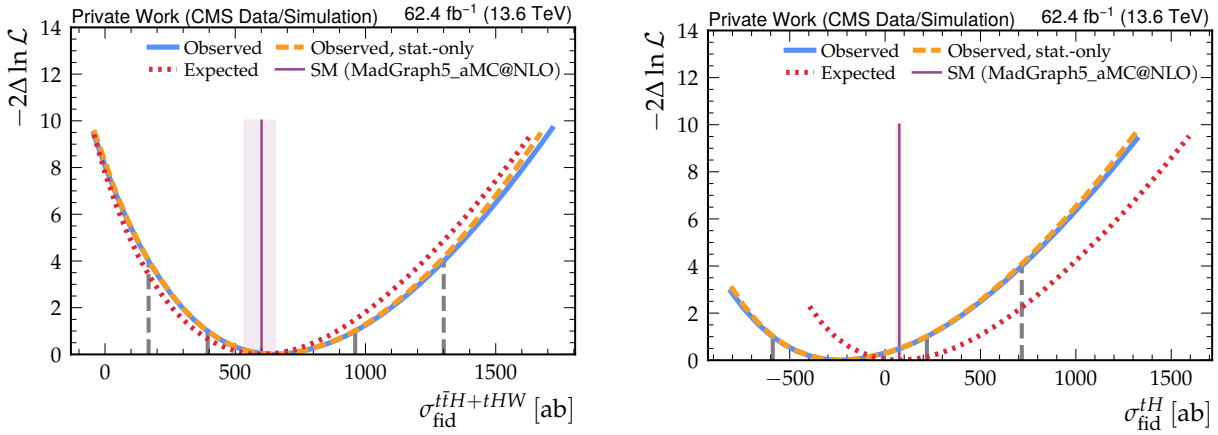
Figure 9.8: Two-dimensional profile-likelihood scan in the fiducial cross-section plane using the decorrelated categorisation and the full systematic uncertainty model. The coloured background shows the observed profile test statistic evaluated at each point as continuous interpolation over the scanned grid, while the black solid (dashed) contours indicate the observed 68% (95%) confidence regions. The expected 68% and 95% contours are included in orange. The red diamond marks the SM point, and the black cross denotes the observed best-fit value.

and 95 % confidence contours while retaining the correlation between tH and $t\bar{t}H$. It is obtained from a scan of 10000 grid points, distributed equidistantly over the displayed parameter plane, at which the profile-likelihood test statistic is evaluated. The confidence contours shown in the figure are obtained by interpolating the resulting likelihood surface. As in the Asimov studies above, the scan is truncated on the far negative-r_{tH} side where the expected yield in at least one category would become negative, which affects only the outermost part of that half-plane and not the proximity of the best-fit point. The observed best-fit point lies within the expected 68% confidence region. Quantified in terms of the p-value of the SM point in the two-dimensional likelihood plane, the compatibility is $p = 0.78$ and hence there is no evidence for a significant deviation from the SM expectation.

The observed (expected) individual fiducial cross sections are

$$\sigma_{\text{fid}}^{t\bar{t}H+tHW} = 650^{+310}_{-260}(\text{stat.})^{+50}_{-30}(\text{syst.}) \left(600^{+290}_{-250}(\text{stat.})^{+40}_{-20}(\text{syst.})\right) \text{ ab} = 650^{+310}_{-260} \left(600^{+290}_{-250}\right) \text{ ab},$$
$$\sigma_{\text{fid}}^{tH} = -240^{+450}_{-350}(\text{stat.})^{+60}_{-60}(\text{syst.}) \left(74^{+400}_{-330}(\text{stat.})^{+70}_{-40}(\text{syst.})\right) \text{ ab} = -240^{+450}_{-350} \left(74^{+410}_{-340}\right) \text{ ab}.$$

The quoted uncertainties are obtained from the one-dimensional profile-likelihood scans shown in Fig. 9.9, where the statistical uncertainty is the predominant component. These scans are produced by evaluating 100 points on a one-dimensional grid for each POI, profiling all NPs and the other POI. The negative central value obtained for tH does not imply a physical negative cross section, but reflects the unconstrained maximum-likelihood estimator used for the fiducial cross-section measurement, as discussed in Sec. 9.3. For comparison, the same figure also



(a) $\sigma_{\text{fid}}^{t\bar{t}H+tHW}$ scan

(b) σ_{fid}^{tH} scan

Figure 9.9: One-dimensional profile-likelihood scans, expressed in terms of the cross sections in the fiducial phase space. The dashed curves indicate scans with only the statistical uncertainties included, and the solid curves show the result with the full statistical model. The red dotted curve shows the expected result obtained with the SM Asimov dataset. The solid and dashed grey vertical lines indicate the 68% and 95% CL intervals, respectively. For comparison, the SM predictions and their uncertainties as obtained with the `MadGraph5_aMC@NLO` generator are shown by the vertical line and the shaded band, respectively.

shows the corresponding SM predictions. In the fiducial phase space considered, they are obtained as $\sigma_{\text{fid,SM}} = \sigma_{\text{tot,SM}} \cdot A_{\text{fid,SM}} \cdot \mathcal{B}(H \to \gamma\gamma)_{\text{SM}}$, using the simulation samples generated with `MadGraph5_aMC@NLO` to determine the fiducial acceptance $A_{\text{fid,SM}}$ and the total cross-section predictions from Ref. [85]. This yields

$$\sigma_{\text{fid,SM}}^{t\bar{t}H+tHW} = 602^{+37}_{-56}\,(\text{scales})^{+28}_{-26}\,(\text{PDFs}\,\&\,\alpha_s) \pm 17\,(\mathcal{B})\,\text{ab} = 602^{+50}_{-64}\,\text{ab},$$
$$\sigma_{\text{fid,SM}}^{tH} = 74.2^{+5.5}_{-11.6}\,(\text{scales})^{+3.2}_{-3.1}\,(\text{PDFs}\,\&\,\alpha_s) \pm 2.1\,(\mathcal{B})\,\text{ab} = 74.2^{+6.5}_{-12.2}\,\text{ab}.$$

The quoted theory uncertainties are obtained by combining acceptance variations derived from the varied event weights with the corresponding uncertainties on the inclusive production cross sections taken from Ref. [85]. The considered uncertainties comprise renormalisation- and factorisation-scale variations, PDF and α_s uncertainties, as well as the uncertainty on $\mathcal{B}(H \to \gamma\gamma)$ from Ref. [52]. Expressed relative to the nominal SM predictions, the measured values can be written as

$$r_{\text{t}\bar{\text{t}}\text{H+tHW}} = 1.08^{+0.51}_{-0.43}\left(1.00^{+0.48}_{-0.42}\right), \quad r_{\text{tH}} = -3.1^{+6.0}_{-4.7}\left(1.0^{+5.5}_{-4.5}\right).$$

The overall compatibility of the observed data with the fitted statistical model is quantified with a goodness-of-fit test. For this purpose, a test statistic comparing the observed data with the best-fit model is evaluated over all categories and mass bins and compared with the corresponding distribution obtained from pseudo-data generated from the fitted model. The resulting p-value is 0.72, indicating no evidence for a significant discrepancy between the data and the model. The fitted POIs are moderately anti-correlated, with correlation coefficient $\rho(r_{tH}, r_{t\bar{t}H+tHW}) = -0.31$. This reflects a mild trade-off between the two signal components in the simultaneous fit, while the categorisation still provides substantial separation between them.

Significance and upper limit

Following the prescription from Sec. 9.3, the significance of the $t\bar{t}H{+}tHW$ signal and an upper limit on the tH cross section are determined. The observed $t\bar{t}H + tHW$ signal significance corresponds to 2.8 standard deviations, in agreement with the expected significance of 2.8 standard deviations. No significant signal is observed for tH production and an upper limit is reported using the CL_s technique. The observed 95% CL upper limit on the fiducial tH cross section is 755 ab, whereas the expected limit is 830 ab. These values correspond to 10.2 and 11.2 times the nominal SM prediction, respectively. The result is visualised in Fig. 9.10, which shows the observed and expected CL_s upper limits together with the 68% and 95% CL expected bands. The observed limit is within the 68% CL expected band under the SM expectation, which is obtained using the asymptotic approximation for the distributions of the profile-likelihood test statistic entering the CL_s construction.

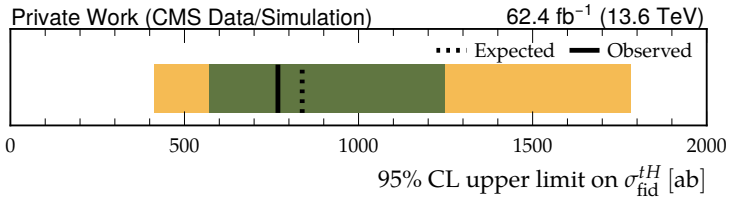
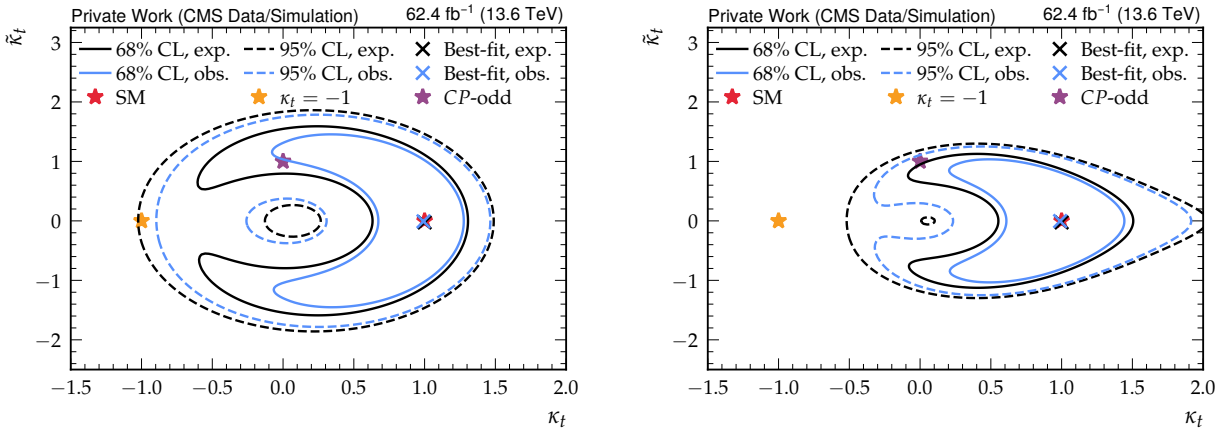


Figure 9.10: Observed and expected 95 % CL upper limits on the fiducial tH cross section obtained with the CL_s prescription and the asymptotic approximation. The expected limit corresponds to the median expectation for the SM signal-plus-background model, while the green and yellow bands indicate the central 68% and 95% intervals of the expected limit distribution in the asymptotic approximation.

Constraints on top-quark Yukawa coupling parameters

The likelihood in the fiducial cross-section plane is re-expressed in terms of the Yukawa-coupling parameters $(\kappa_t, \tilde{\kappa}_t)$ by mapping each coupling point to the corresponding predicted pair of fiducial cross sections and evaluating the profiled likelihood ratio from the two-dimensional likelihood scan shown in Fig. 9.8. Two variants of this reinterpretation are considered. Both use the total cross sections from Eqs. (5.1)–(5.3) together with the fiducial acceptance shown in Fig. 6.1. In the first interpretation, $\mathcal{B}(H \to \gamma\gamma)$ is fixed to its SM value in order to remain consistent with the conventions used in previous published measurements and to isolate the tree-level impact of the coupling modifiers on the production cross sections. In the second interpretation, the coupling dependence of $\mathcal{B}(H \to \gamma\gamma)$ is included according to Eq. (2.26), which treats the scaling with $(\kappa_t, \tilde{\kappa}_t)$ consistently between production and decay. The resulting constraints on the coupling parameters are shown in Fig. 9.11.
The observed best-fit point is very close to the SM prediction in both variants. The observed confidence regions are tighter than expected in the region with negative κ_t or non-zero $\tilde{\kappa}_t$, where the tH cross section would be strongly enhanced. This is consistent with the slight depletion observed in the most tH-sensitive categories. When the coupling dependence of $\mathcal{B}(H \to \gamma\gamma)$ is included, the constraints in the region with negative κ_t become significantly tighter, because Eq. (2.26) predicts an enhanced diphoton branching ratio there. As a result, the observed data exclude the largest part of that half-plane. At the same time, the constraints in the region with κ_t exceeding the value of 1 are much weaker in that interpretation, because the effect of enhanced



(a) Fixed $\mathcal{B}(H \to \gamma\gamma)$ **(b)** Coupling-dependent $\mathcal{B}(H \to \gamma\gamma)$

Figure 9.11: Constraints on the top-quark Yukawa coupling-parameter plane $(\kappa_t, \tilde{\kappa}_t)$ obtained by reinterpreting the likelihood scan in the fiducial cross-section plane. The left panel keeps $\mathcal{B}(H \to \gamma\gamma)$ fixed at its SM value, whereas the right panel varies it according to Eq. 2.26. The expected and observed confidence regions are shown together with the best-fit points. As reference, the SM prediction of $(\kappa_t, \tilde{\kappa}_t) = (1, 0)$, the purely CP-odd scenario $(\kappa_t, \tilde{\kappa}_t) = (0, 1)$, and the inverted-sign case $(\kappa_t, \tilde{\kappa}_t) = (-1, 0)$ are shown with coloured markers.

production cross sections and decreased $\mathcal{B}(H \to \gamma\gamma)$ cancel each other. The purely CP-odd scenario $(\kappa_t, \tilde{\kappa}_t) = (0, 1)$ lies close to the observed 68 % CL contour in the fixed-$\mathcal{B}(H \to \gamma\gamma)$ interpretation and close to the observed 95 % CL contour once the branching-ratio scaling is included. For the inverted-sign scenario $(\kappa_t, \tilde{\kappa}_t) = (-1, 0)$, the exclusion is above 95 % CL for fixed $\mathcal{B}(H \to \gamma\gamma)$ and reaches a p-value below 10^{-7} with branching-ratio scaling, corresponding to more than five Gaussian standard deviations.

Impacts of nuisance parameters

The impact of the NPs introduced in Sec. 9.3 on the two POIs is analysed in the following. For each NP, the fit is repeated after fixing that parameter to values one standard deviation above and below its post-fit value, while all remaining NPs and the POIs are profiled under this condition. The resulting shift in the corresponding POI defines the impact attributed to that NP. In addition, the post-fit central values and uncertainties of the NPs themselves are obtained to assess which uncertainty sources are most strongly constrained or shifted by the fit. The results are summarised in Fig. 9.12, which ranks the dominant NPs according to their impacts for each of the two POIs. The impact pattern is qualitatively different for the two measured production modes. For $r_{t\bar{t}H+tHW}$, the leading contributions arise from b-tagging uncertainties, in particular the light-flavour contamination and the charm-quark related component, as well as from theory uncertainties affecting the $t\bar{t}H$ signal modelling. There, the simultaneous variation of the factorisation and renormalisation scale shows the largest impact, while PS variations also contribute to a similar extent. The corresponding shifts are modest, with the largest individual impact amounting to around 3% in $r_{t\bar{t}H+tHW}$. The observed and expected impact patterns for $r_{t\bar{t}H+tHW}$ are very similar, both in sign and in magnitude for the leading NPs. This indicates that the observed fit remains close to the Asimov expectation in the $t\bar{t}H$-sensitive direction and

supports the conclusion that the $t\bar{t}H$ measurement is statistically dominated, with systematic uncertainties providing only a moderate additional broadening of the confidence interval.

For r_{tH}, the dominant NPs are different. The largest impacts originate from the conservative heavy-flavour associated uncertainties assigned to the ggH, VBF, and VH production modes, followed by photon-energy-scale and material-related uncertainties and, to a smaller extent, theory uncertainties in the $t\bar{t}H$ signal model. This reflects the smaller purity of tH-sensitive categories in that signal mode and the weaker sensitivity of the fit to r_{tH}: because the likelihood is broad in the tH direction, relatively small changes in the modelling of other Higgs boson production modes or in the photon-energy response can induce sizeable shifts in the profiled best-fit value of r_{tH}. The largest post-fit impacts on r_{tH} are correspondingly much larger than for $r_{t\bar{t}H+tHW}$, reaching values of order 0.6 for the leading NPs. In contrast to the $t\bar{t}H$ case, the observed and expected impact patterns for r_{tH} differ notably: for several dominant NPs, the sign of the impact changes and large relative changes in the magnitude between the observed and Asimov fits are observed. This behaviour is also consistent with the location of the fitted rate parameters: the observed best-fit value of $r_{t\bar{t}H+tHW}$ remains close to the SM expectation, whereas the observed best-fit value of r_{tH} is shifted farther away. Since the impact is a local property of the profiled likelihood around the respective fit point, such a shift of the fit point can change the local response to an NP variation and the resulting impact. The impact ranking is nevertheless physically sensible and consistent with the corresponding one-dimensional scans, which show that the tH measurement remains strongly limited by statistical uncertainties and substantially less constrained than the $t\bar{t}H$ measurement.

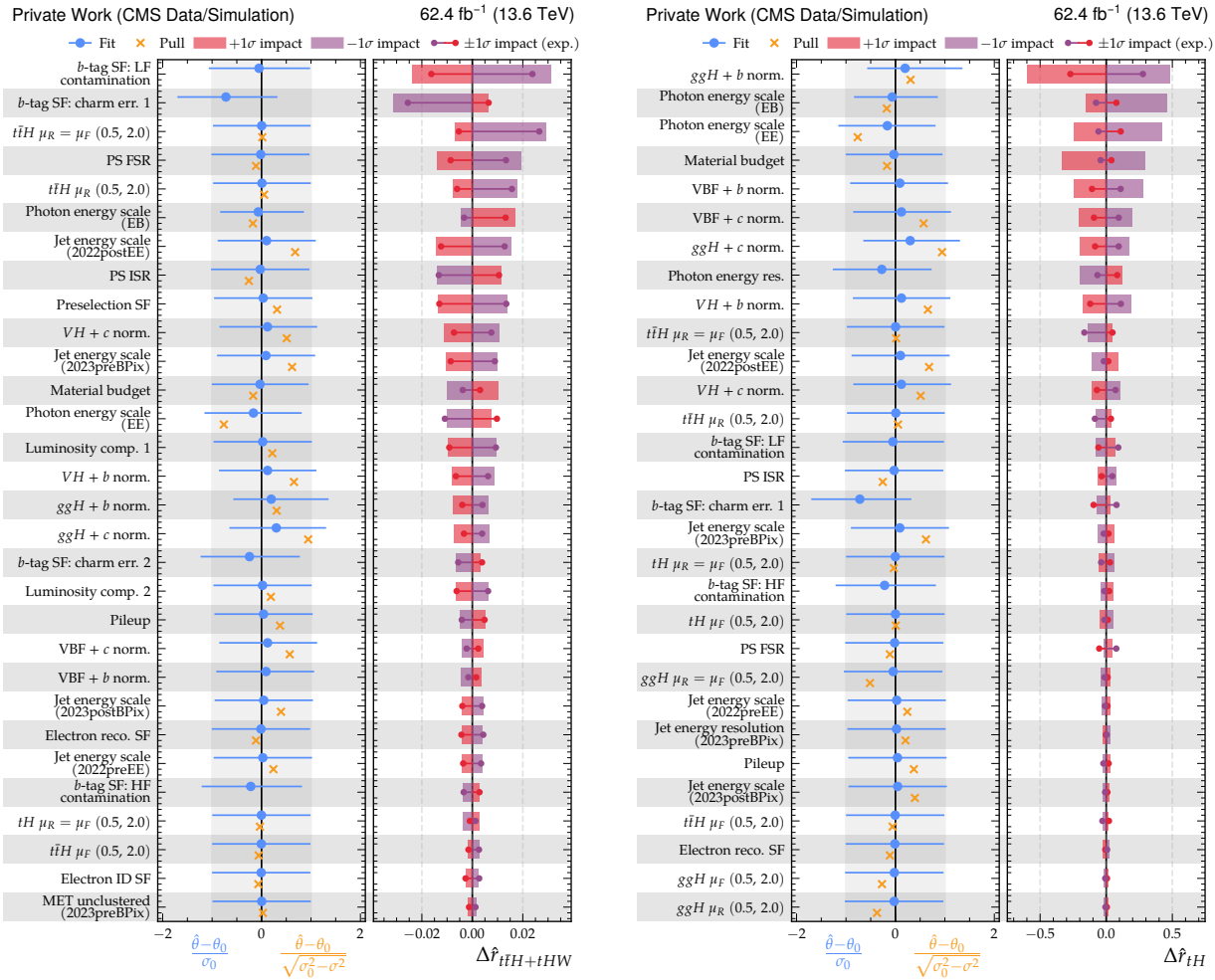
The post-fit NP values are also inspected to verify that the variations preferred by the fit correspond to physically understandable directions in parameter space. The largest shift is found for the b-tagging NP *b-tag SF: charm err. 1*, which is fitted to a value about 0.7 standard deviations below its nominal value. A negative variation of this nuisance mainly increases the expected yields in the $t\bar{t}H$ hadronic categories while decreasing them in the leptonic categories, matching the pattern seen in the observed data. This behaviour is plausible because the uncertainty changes the shape of the b-tagging discriminant for charm jets after removing the inclusive normalisation effect, which naturally induces residual anticorrelations between events of different charm-jet multiplicity. The hadronic categories have a higher average charm-jet multiplicity than the leptonic categories, since they explicitly target hadronic W-boson decays including $W \to cs$. All other NPs are fitted to values within 0.25 standard deviations of their nominal values. Slight constraints are observed for the photon energy scale NPs, where the corresponding uncertainty was conservatively assigned, hence such constraints are expected.

For each NP, the fitted value is shown together with the pull quantity

$$g = \frac{\hat{\theta} - \theta_0}{\sqrt{\sigma_0^2 - \sigma^2}}, \tag{9.14}$$

where θ_0 and σ_0 denote the nominal value and prefit uncertainty, and $\hat{\theta}$ and σ the corresponding post-fit quantities [258]. In this definition, the shift of the NP away from its nominal value is normalised to the information gain provided by the fit, as encoded in the reduction of the variance from σ_0^2 to σ^2. The pull is therefore useful as a diagnostic of whether an NP is both shifted and constrained by the data. For a few weakly constrained NPs, however, the relevant post-fit uncertainty is slightly larger than the corresponding prefit uncertainty, so that the denominator becomes ill-defined. In these cases, the fit still returns a well-defined post-fit central value and

uncertainty, and the corresponding impacts remain meaningful, while only the pull quantity itself is not defined.



(a) Impact ranking for $r_{t\bar{t}H+tHW}$

(b) Impact ranking for r_{tH}

Figure 9.12: Ranking of the dominant post-fit nuisance-parameter effects on $r_{t\bar{t}H+tHW}$ and r_{tH} in the nominal unblinded fit. In each panel, the left-hand side shows the fitted nuisance-parameter values relative to their nominal values: blue points with horizontal bars indicate the post-fit central values and their uncertainties, while orange crosses denote the corresponding pull quantity wherever it is well defined. On the right-hand side, the impact of each NP on the corresponding POI is shown as horizontal bars, with the green and red bars indicating the downward and upward POI shifts induced by varying the nuisance within its post-fit uncertainty. The NPs are ordered from top to bottom by decreasing impact on the POI.

10. Conclusions

This thesis presented a measurement of Higgs boson production cross section in association with a top-quark pair or with a top quark and a W boson, $t\bar{t}H + tHW$, together with a search for the t-channel production with a single top quark, tH, in the diphoton decay channel of the Higgs boson. The analysis used proton–proton collision data recorded with the CMS detector in 2022 and 2023 at $\sqrt{s} = 13.6\,\text{TeV}$, corresponding to an integrated luminosity of $62.4\,\text{fb}^{-1}$. Events with two photon candidates, at least one b-tagged jet, and additional jets or leptons were selected and split into hadronic and leptonic channels. By formulating the measurement in a fiducial phase space defined at the particle level, it remained closely aligned with the experimentally accessible region, and no extrapolation to the total phase space was required. This is particularly relevant for the interpretation in terms of the top-quark Yukawa interaction, where the analysis was designed to reduce the dependence of the result on assumptions about the signal kinematics.

A central methodological development of this work is a machine-learning-based classification and categorisation strategy that is sensitive to the rare signals while remaining robust under variations of the top-quark Yukawa coupling. Deep neural-network classifiers were trained on signal samples spanning different coupling hypotheses, parameterised by the modifiers κ_t and $\tilde{\kappa}_t$, and an adversarial decorrelation procedure was used to suppress the direct impact of these assumptions on the classifier outputs. On top of this, a differentiable optimisation method was developed to determine category boundaries directly in the three-dimensional discriminant space. Its Gaussian-mixture-based parameterisation proved flexible enough to exploit the available separation power efficiently and constitutes a generally useful tool for event categorisation beyond this analysis. The analysis strategy substantially reduced the model dependence of the extracted cross sections, in particular for tH, such that the remaining variations of the expected confidence regions across the $(\kappa_t, \tilde{\kappa}_t)$ plane are small compared with the measurement uncertainty, while preserving a competitive overall sensitivity. Several further improvements were required to achieve the final performance. A particular methodological advancement was the inclusion of hadronic tH categories, which substantially increased the sensitivity compared with the purely leptonic approaches used in Refs. [39, 40]. This was enabled by an effective continuum-background description that allowed the training of machine-learning classifiers.

The measurement yielded an observed and expected significance of 2.8 standard deviations for the combined $t\bar{t}H + tHW$ signal. The corresponding fiducial cross section was measured to be $650^{+310}_{-260}\,\text{ab}$, in good agreement with the standard model (SM) prediction of $602^{+50}_{-64}\,\text{ab}$. No significant tH signal was observed, and an upper limit of 755 ab was set at 95 % CL, corresponding to 10.2 times the nominal SM prediction and close to the expected limit of 830 ab. The measurement is strongly dominated by the statistical uncertainty. Because the diphoton decay is part of the fiducial phase-space definition, the measured cross sections are largely independent of assumptions on $\mathcal{B}(H \to \gamma\gamma)$, and the corresponding branching-ratio dependence enters mainly through the coupling reinterpretation. In the measurement, only the top-quark Yukawa interaction was varied in the signal processes, while the remaining Higgs boson couplings, in particular the HWW coupling relevant for tH production, were kept at their SM values. Other Higgs boson production modes were normalised to their SM predictions, but large uncertainties were

assigned to their components with additional heavy-flavour jets because these contributions are difficult to predict in the selected phase space. As a consequence, the statistical model is also expected to have sufficient flexibility to absorb variations in these background contributions without inducing significant model-dependent biases in the extracted signal cross sections.
In a first reinterpretation of the measured likelihood from the cross-section plane to the κ_t–$\tilde{\kappa}_t$ plane, $\mathcal{B}(H \to \gamma\gamma)$ is fixed to its SM value in order to isolate the tree-level coupling dependence of the production cross sections. Here, the inverted-sign scenario $(\kappa_t, \tilde{\kappa}_t) = (-1, 0)$ is excluded at 95 % CL, while the purely CP-odd scenario $(0, 1)$ lies close to the observed 68 % CL contour. When the coupling dependence of $\mathcal{B}(H \to \gamma\gamma)$ is included consistently, the constraints on these coupling scenarios become more stringent: the $(\kappa_t, \tilde{\kappa}_t) = (-1, 0)$ case is excluded with a p-value below 10^{-7}, while the $(\kappa_t, \tilde{\kappa}_t) = (0, 1)$ case lies close to the 95 % CL contour. These constraints are complementary to those obtained in Run-2 measurements that relied on dedicated CP-sensitive observables [26, 27], because the sensitivity here arises from a coupling-agnostic measurement of the $t\bar{t}H + tHW$ and tH cross sections.
The deliberate reduction of model dependence mildly reduces the sensitivity relative to an analysis optimised purely for the SM kinematic hypothesis, but yields a result that is more reliable for reinterpretations away from the SM assumption. A direct luminosity rescaling must be interpreted cautiously, because the current categorisation is significantly constrained by the minimum event yield required for background estimation and therefore scales conservatively with luminosity. Nevertheless, a naive scaling to the full Run-2 luminosity gives an expected $t\bar{t}H + tHW$ significance of about 4.2 standard deviations, close to the 4.7 standard deviations expected in the dedicated CMS Run-2 $t\bar{t}H$ diphoton measurement [26]. For tH, the same exercise gives an expected limit of about 7.5 times the SM prediction, comparable to the expected limits of 6.8 from ATLAS [40] and 8 from CMS [39]. These results, however, also included the tHW contribution, which is enhanced in the leptonic channel; consequently, the quoted limits cannot be interpreted as constraints on the t-channel tH process alone. The resulting expected sensitivity to tH achieved in this thesis is particularly interesting because several Run-2 measurements showed mild excesses in tH-sensitive regions [34–36], whereas the present analysis finds no such indication in the statistically independent dataset analysed. At the time of finalising this thesis, an ATLAS analysis of 164 fb^{-1} of Run-3 data using related hadronic tH categories set an observed upper limit of 6.2 times the SM prediction on the combined t-channel tH and tHW production rate in the diphoton final state; the combined Run-2 and Run-3 interpretation excluded a purely CP-odd top-quark Yukawa coupling at 5.8 standard deviations [259].
The most important next step is the extension of this measurement to the substantially larger Run-3 dataset that has been recorded after 2023 and is not yet included here. Further gains should come from combinations with other Higgs boson decay channels as in Ref. [28] and from augmenting the present strategy with dedicated CP-sensitive discriminants, focusing on both $t\bar{t}H$ and tH production. The methodological developments of this thesis are not restricted to this analysis. The differentiable categorisation framework can be applied to a broad class of optimisation problems, in particular those with many signal classes or high-dimensional discriminant spaces, and the effective MC-based continuum-background description can be useful in other $H \to \gamma\gamma$ analyses where data control samples are statistically limited. Overall, this thesis demonstrates that the fiducial analysis strategy designed for reduced model dependence can achieve competitive sensitivity to $t\bar{t}H$ and tH production in the $H \to \gamma\gamma$ decay channel already with the first Run-3 dataset, while providing a robust basis for future coupling interpretations.

A. Effect of CP assumptions on tH efficiencies

Figure A.1 shows the per-category tH efficiencies for the same three CP-decorrelation configurations discussed in Sec. 8.4. The figure mirrors the $t\bar{t}H{+}tHW$ case, illustrating how the CP-decorrelation strategy affects the tH efficiencies across the optimised categories.

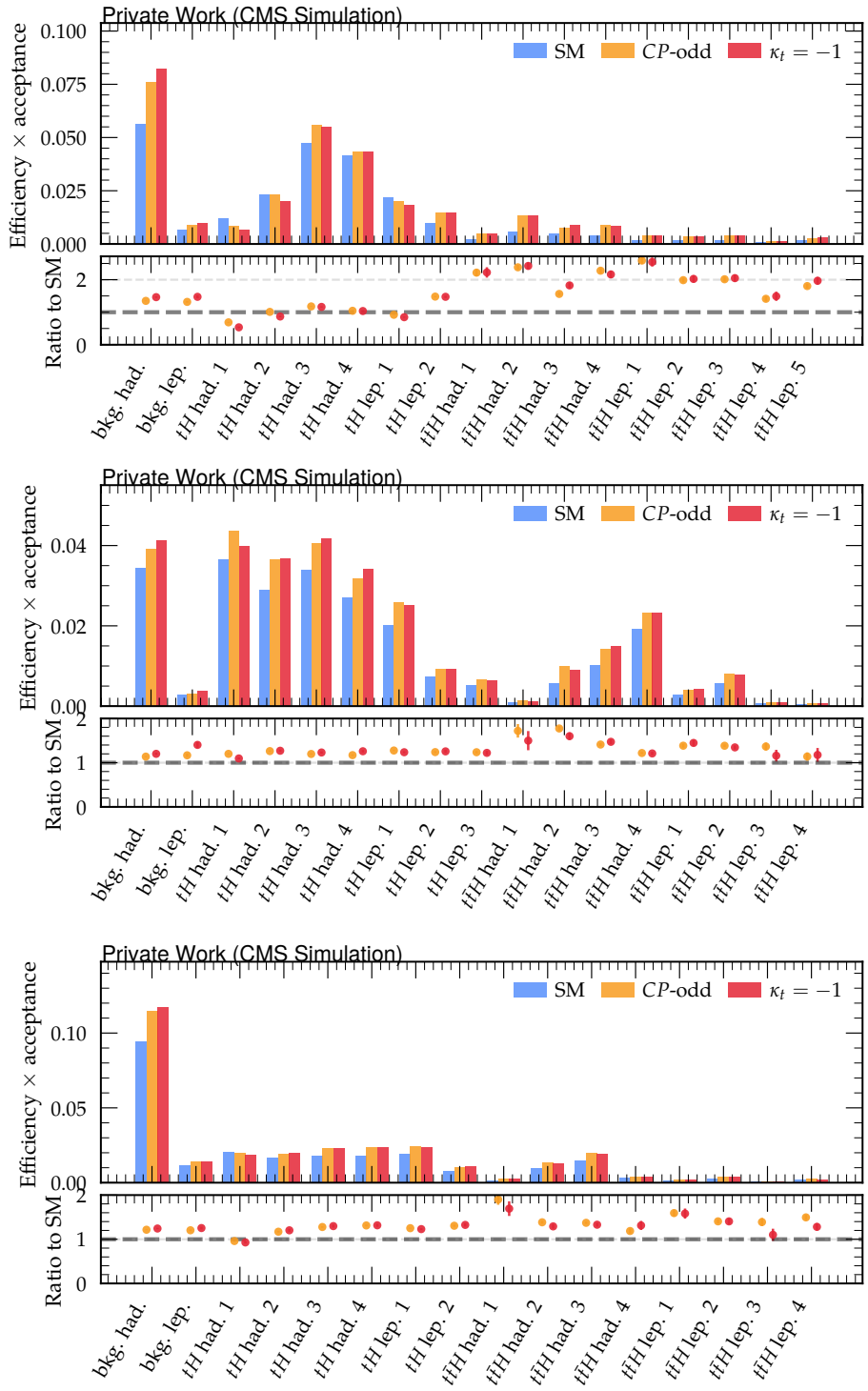
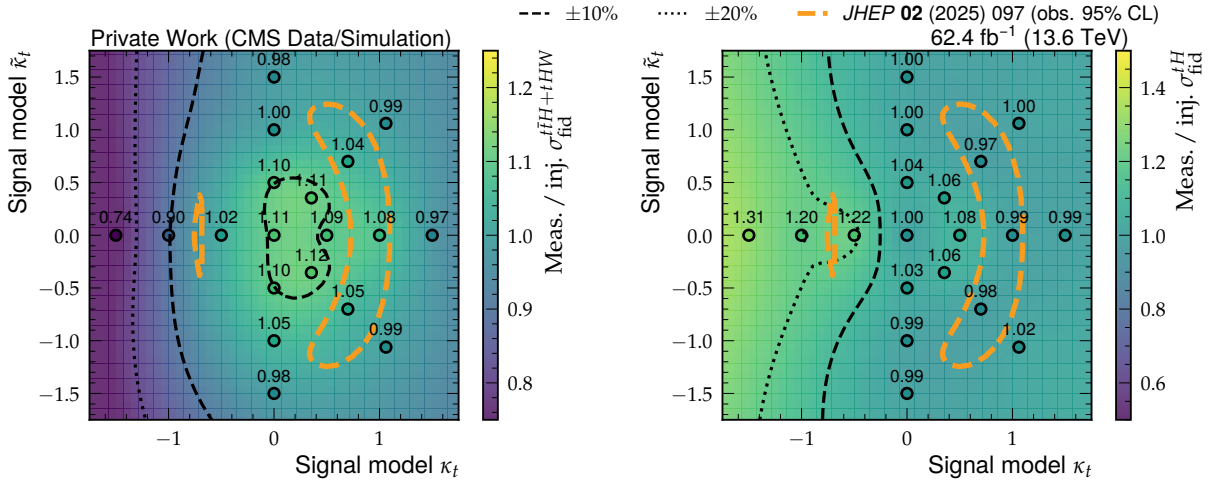


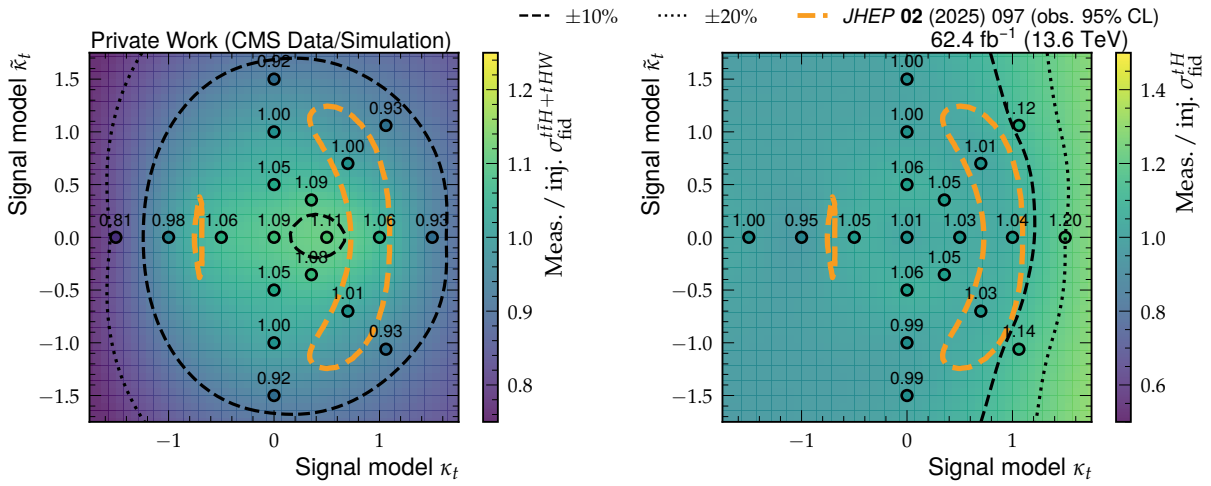
Figure A.1: Per-category tH efficiencies (including acceptance) in the hadronic channel. From top to bottom: comparison setup based on classifiers purely trained with SM samples; CP-decorrelated classifiers without a CP penalty in the GATO optimisation; CP-decorrelated classifiers with a CP penalty of 30. The penalty reduces the relative CP-even/odd efficiency differences, most visibly in tH-like categories that otherwise receive only a small $t\bar{t}H$ tail.

B. Robustness studies with alternative Asimov datasets

This appendix presents robustness studies based on Asimov datasets generated under two alternative coupling assumptions: a purely CP-odd top-quark Yukawa coupling and an inverted-sign configuration with $\kappa_t = -1$. For each hypothesis, the interpolated ratio of measured to injected fiducial cross sections across the $(\kappa_t, \tilde{\kappa}_t)$ plane is shown together with the corresponding 68% and 95% CL contours, comparing the SM-trained categorisation to the decorrelated setup. In both scenarios, the decorrelated categorisation yields visibly smoother bias maps and tighter, more stable contours than the SM-trained setup. The residual deviations remain largest in regions where the coupling choice enhances the tH contribution, but the decorrelated configuration consistently reduces the migration-induced bias compared to the non-decorrelated case. The inverted-sign case exhibits larger distortions overall, yet the qualitative improvement from decorrelation persists.

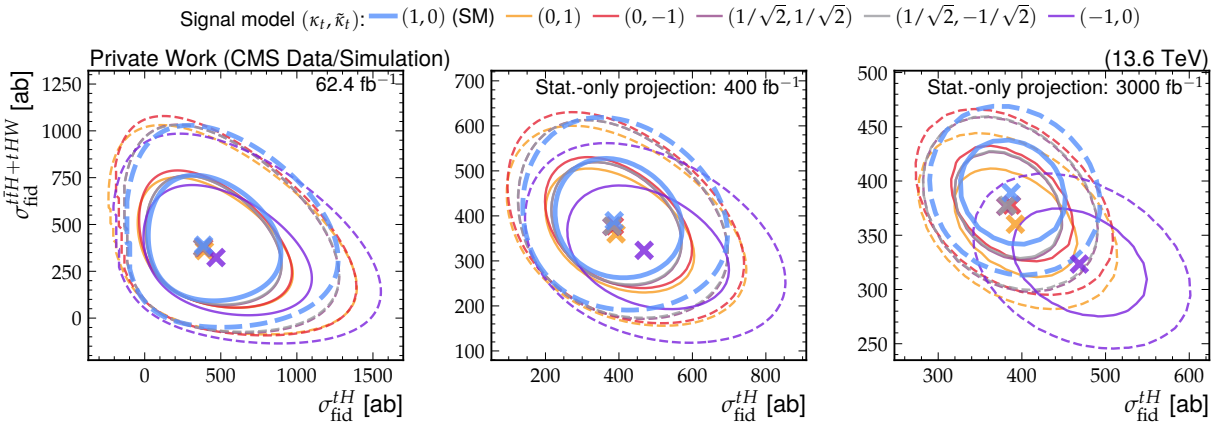


(a) SM-trained categorisation

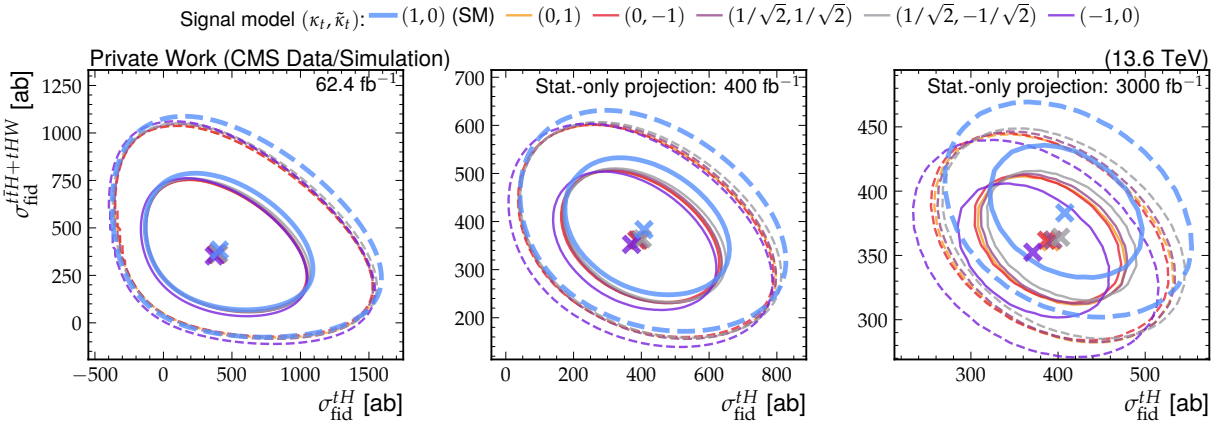


(b) Decorrelated categorisation

Figure B.1: Robustness study with fixed CP-odd Asimov dataset: the ratio of the measured to the injected fiducial cross section is shown for different signal models used across the $(\kappa_t, \tilde{\kappa}_t)$ plane. The top row shows the results for the SM-trained categorisation, while the bottom row shows the results for the decorrelated categorisation. The left and right parts show the results for $t\bar{t}H{+}tHW$ and tH cross sections, respectively. The colored heatmap shows an interpolation obtained with radial basis functions based on the discrete results at the sampled coupling points, which are indicated by the imprinted numbers. The black contours represent the parameters at which the interpolated bias reaches 10% and 20% for the respective process. For comparison, the parameter space compatible with a combination of previous measurements with the CMS experiment at 95% CL is overlaid with the orange dashed contours [257].

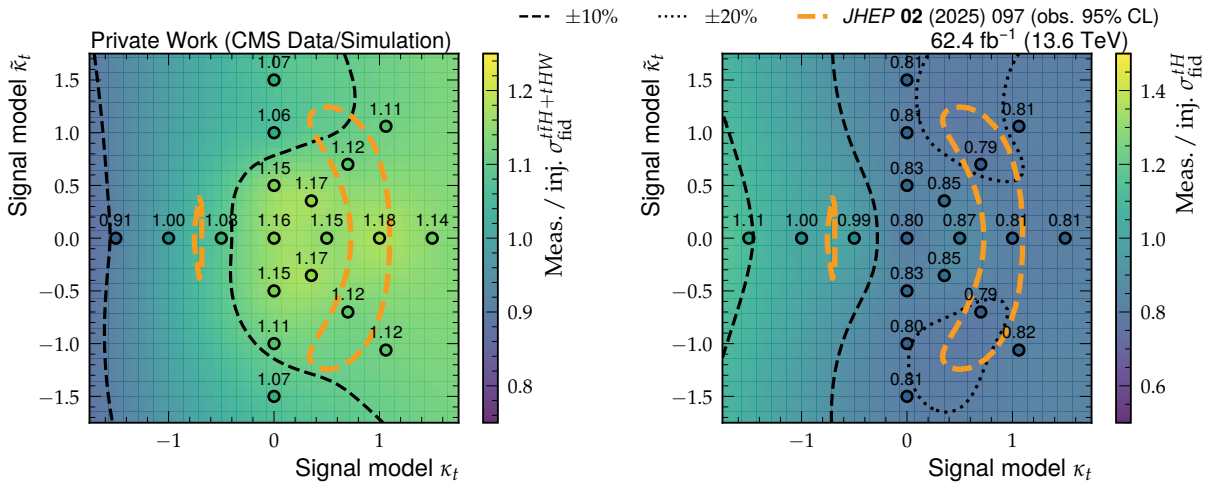


(a) *CP*-odd Asimov: SM-trained categorisation

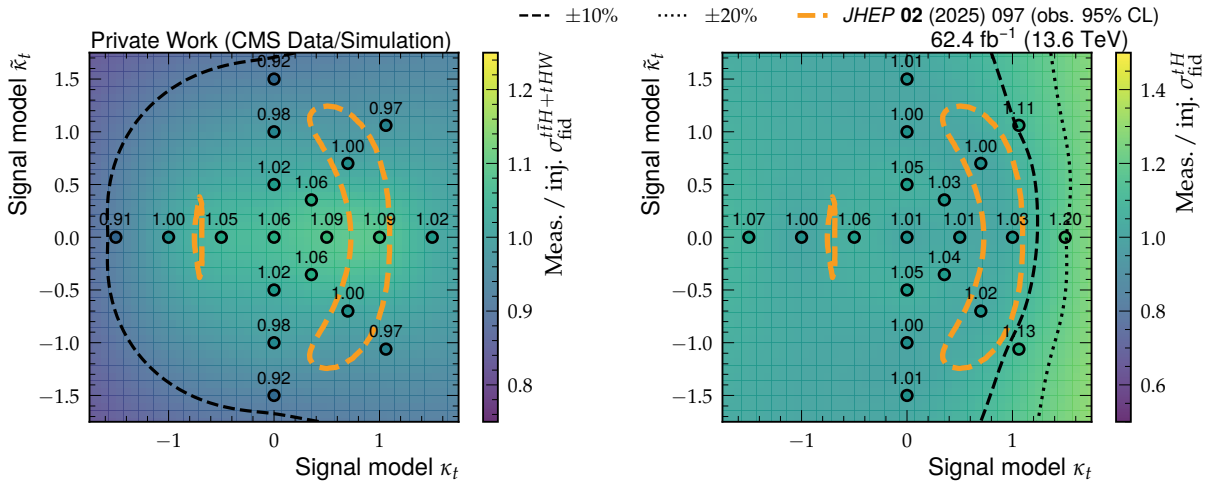


(b) *CP*-odd Asimov: decorrelated categorisation

Figure B.2: Expected two-dimensional confidence regions in the fiducial cross-section plane obtained from a *CP*-odd Asimov dataset and signal models of different coupling scenarios. The solid and dashed contours represent the 68 % and 95 % confidence regions, respectively, while the cross markers indicate the best-fit points. The scans are shown for integrated luminosities of 62.4 fb^{-1} (this analysis), 400 fb^{-1}, and 3000 fb^{-1}, corresponding to expectations for the full Run-3 and the HL-LHC data taking. The upper row shows the results for the categorisation based on neural networks trained on SM signal samples without explicit *CP* decorrelation, while the lower row shows the corresponding results for the nominal decorrelated categorisation. Only statistical uncertainties are included in the fits.

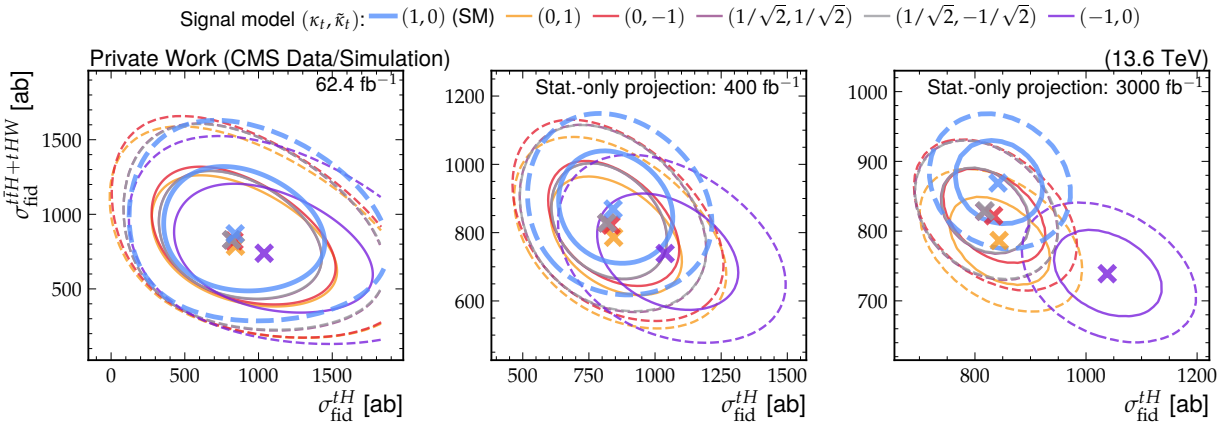


(a) SM-trained categorisation

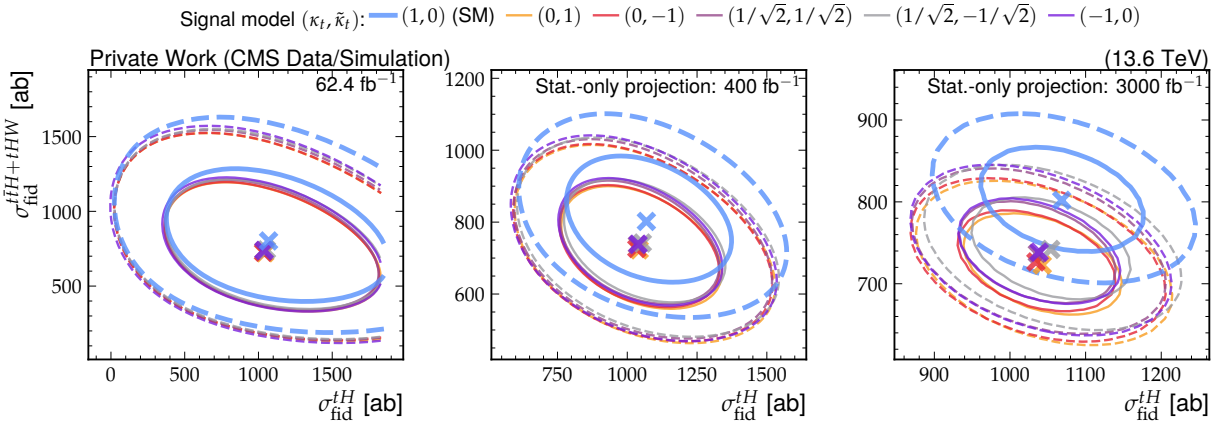


(b) Decorrelated categorisation

Figure B.3: Robustness study with fixed Asimov dataset with the inverted-sign scenario $\kappa_t = -1$: the ratio of the measured to the injected fiducial cross section is shown for different signal models used across the $(\kappa_t, \tilde{\kappa}_t)$ plane. The top row shows the results for the SM-trained categorisation, while the bottom row shows the results for the decorrelated categorisation. The left and right parts show the results for $t\bar{t}H{+}tHW$ and tH cross sections, respectively. The colored heatmap shows an interpolation obtained with radial basis functions based on the discrete results at the sampled coupling points, which are indicated by the imprinted numbers. The black contours represent the parameters at which the interpolated bias reaches 10% and 20% for the respective process. For comparison, the parameter space compatible with a combination of previous measurements with the CMS experiment at 95% CL is overlaid with the orange dashed contours [257].



(a) $\kappa_t = -1$ Asimov: SM-trained categorisation

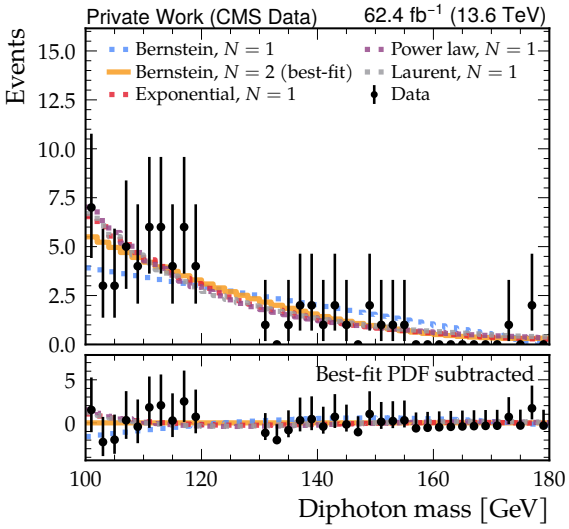


(b) $\kappa_t = -1$ Asimov: decorrelated categorisation

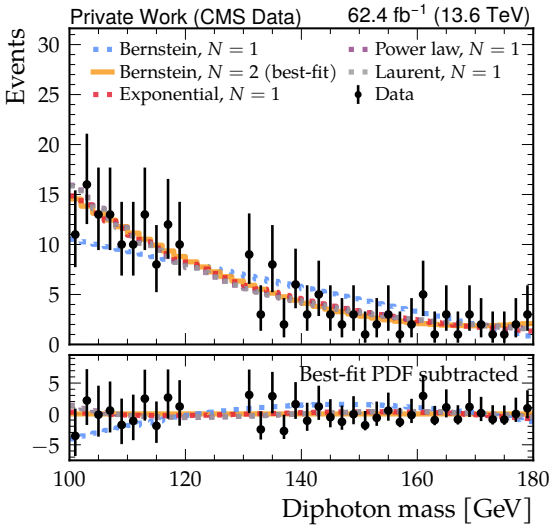
Figure B.4: Expected two-dimensional confidence regions in the fiducial cross-section plane obtained from an Asimov dataset with the inverted-sign scenario $\kappa_t = -1$ and signal models of different coupling scenarios. The solid and dashed contours represent the 68 % and 95 % confidence regions, respectively, while the cross markers indicate the best-fit points. The scans are shown for integrated luminosities of 62.4 fb^{-1} (this analysis), 400 fb^{-1}, and 3000 fb^{-1}, corresponding to expectations for the full Run-3 and the HL-LHC data taking. The upper row shows the results for the categorisation based on neural networks trained on SM signal samples without explicit *CP* decorrelation, while the lower row shows the corresponding results for the nominal decorrelated categorisation. Only statistical uncertainties are included in the fits.

C. Continuum-background modelling

This appendix collects the selected background-fit functions entering the discrete-profiling envelope for the categories not shown explicitly in the main text. The order N of the functions is displayed in each plot and the best-fit function is highlighted. The data in the signal window from 120–130 GeV are not shown.

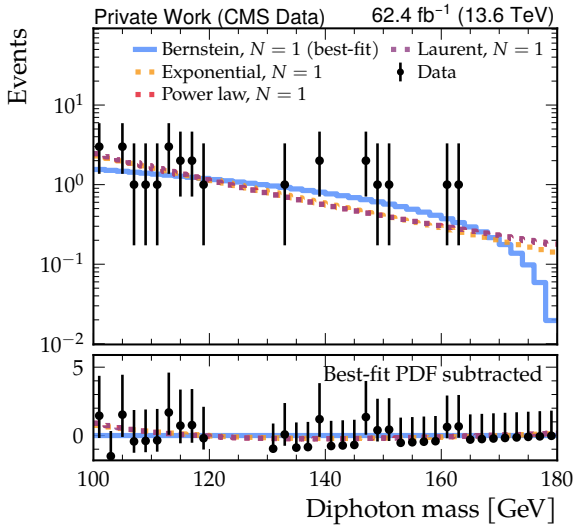


(a) $t\bar{t}H$ had. 2

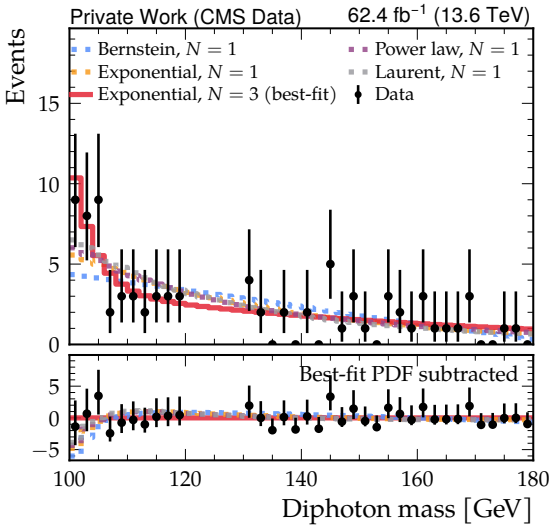


(b) $t\bar{t}H$ had. 3

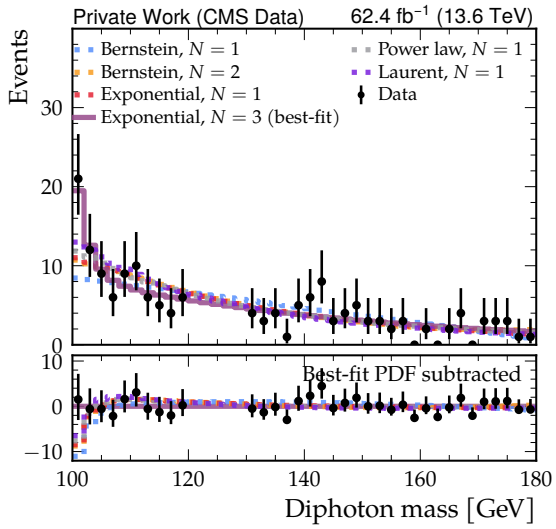
Figure C.1: Selected background fit functions entering the discrete-profiling envelope for the remaining hadronic $t\bar{t}H$ categories. The lower panel in each plot shows the data points with the best-fit function subtracted.



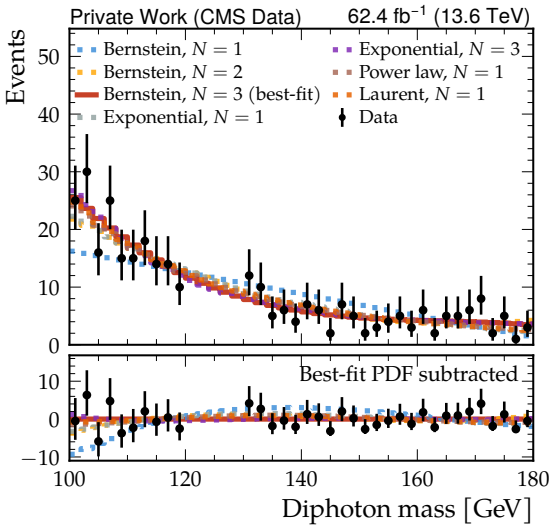
(a) tH had. 1



(b) tH had. 2

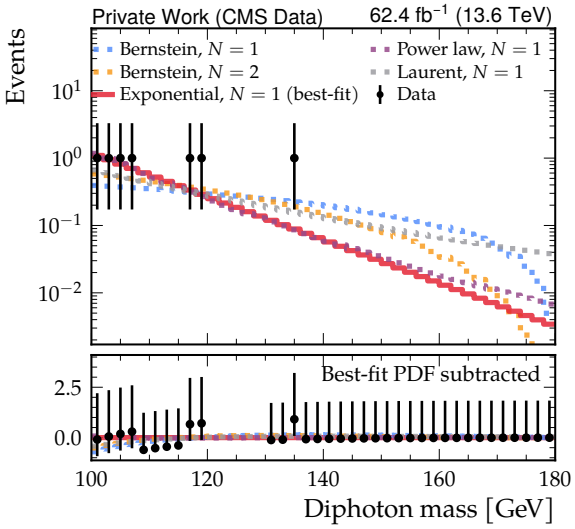


(c) tH had. 3

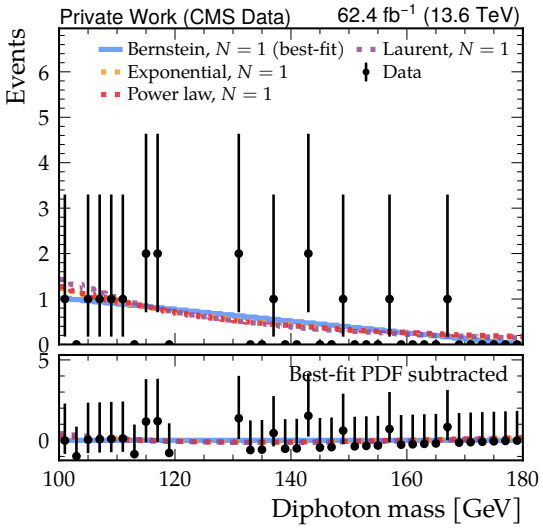


(d) tH had. 4

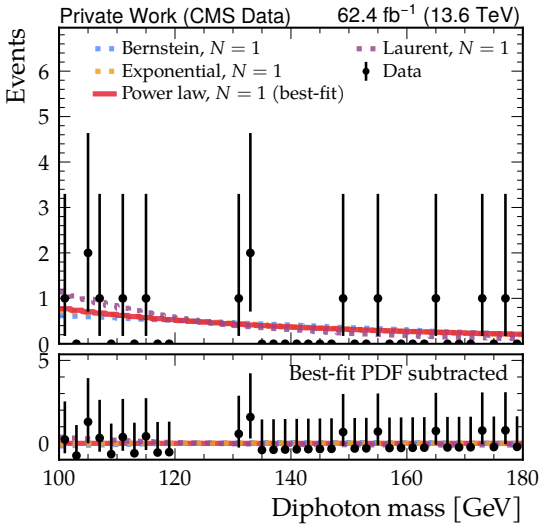
Figure C.2: Selected background fit functions entering the discrete-profiling envelope for the remaining hadronic tH categories. The lower panel in each plot shows the data points with the best-fit function subtracted.



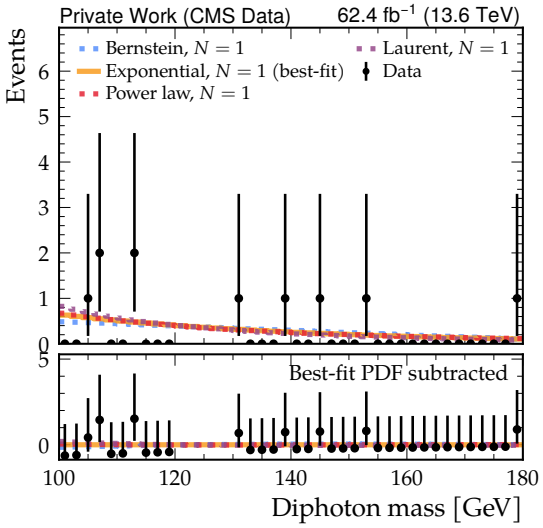
(a) $t\bar{t}H$ lep. 1



(b) $t\bar{t}H$ lep. 2

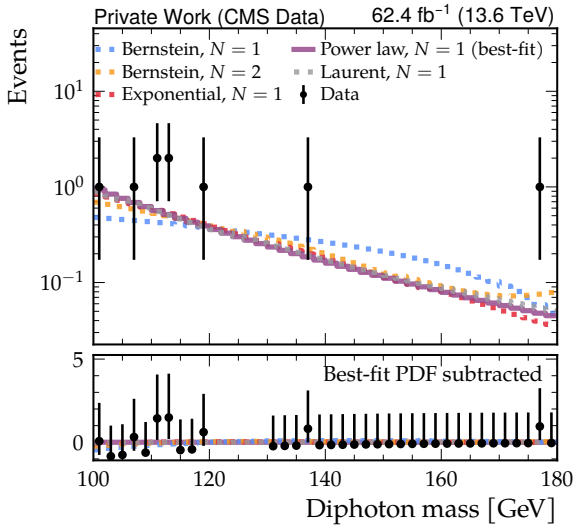


(c) $t\bar{t}H$ lep. 3

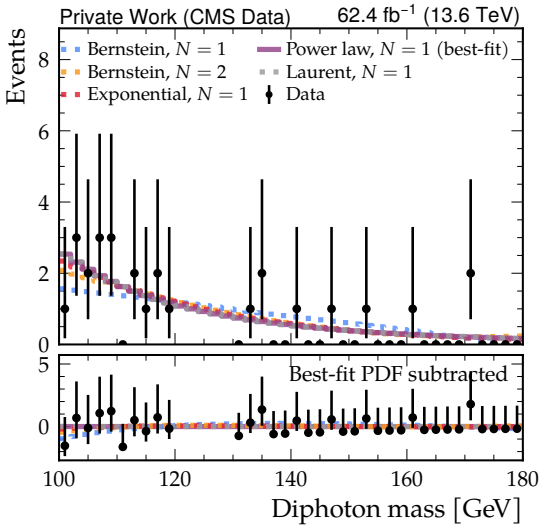


(d) $t\bar{t}H$ lep. 4

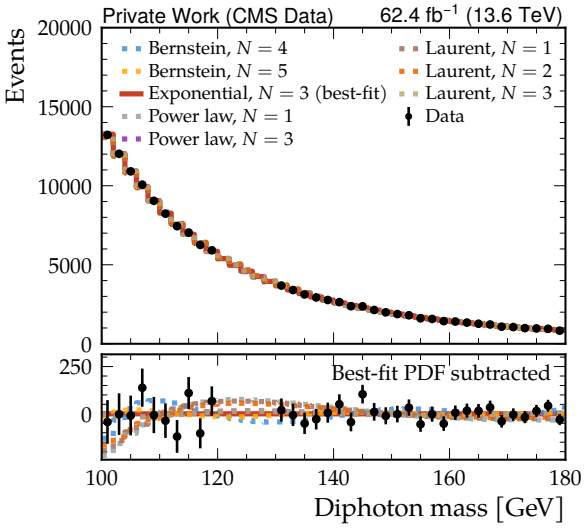
Figure C.3: Selected background fit functions entering the discrete-profiling envelope for the leptonic $t\bar{t}H$ categories. The lower panel in each plot shows the data points with the best-fit function subtracted.



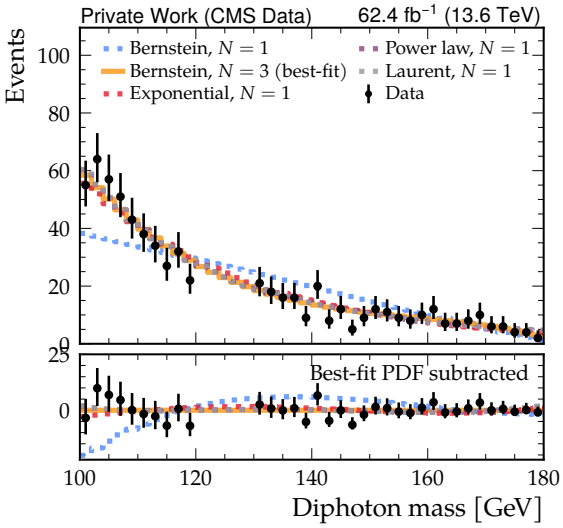
(a) tH lep. 1



(b) tH lep. 2



(c) Bkg. had.

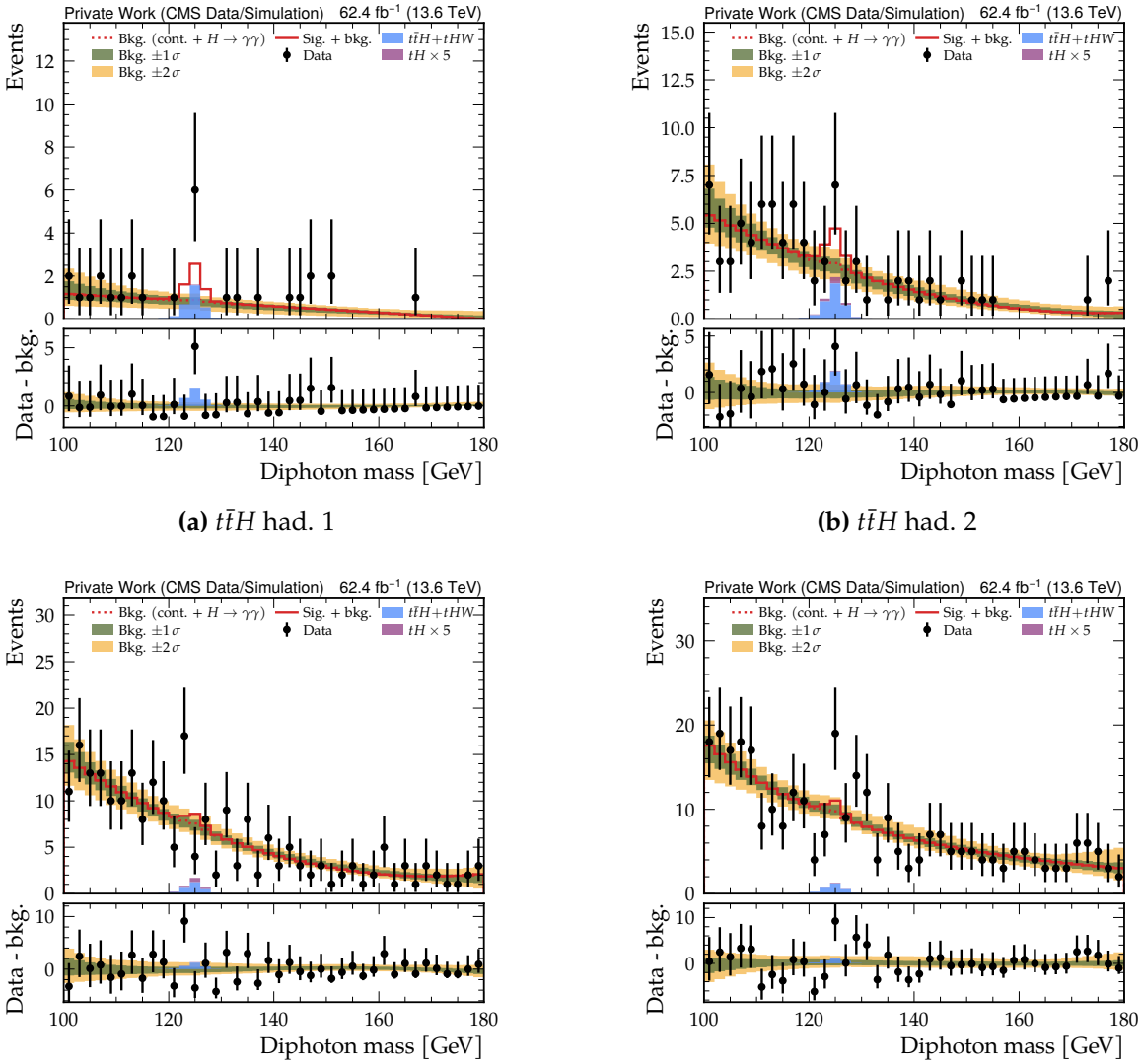


(d) Bkg. lep.

Figure C.4: Selected background fit functions entering the discrete-profiling envelope for the remaining leptonic tH categories and the two background-enriched categories. The lower panel in each plot shows the data points with the best-fit function subtracted.

D. Observed per-category diphoton mass spectra

This appendix shows the unweighted diphoton mass spectra for all analysis categories used in the weighted $S+B$ visualisation. The figures show the spectra used in the statistical model and are grouped by category type following the naming convention from Sec. 8.3. Figures D.1–D.4 show the $t\bar{t}H$ hadronic categories, the $t\bar{t}H$ leptonic categories, the tH hadronic categories, and the tH leptonic plus background categories, respectively.



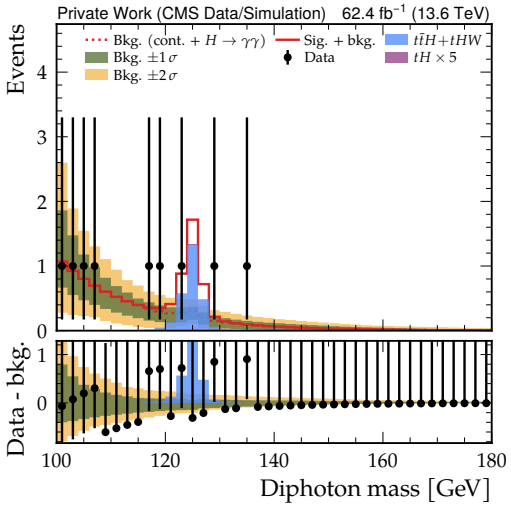
(a) $t\bar{t}H$ had. 1

(b) $t\bar{t}H$ had. 2

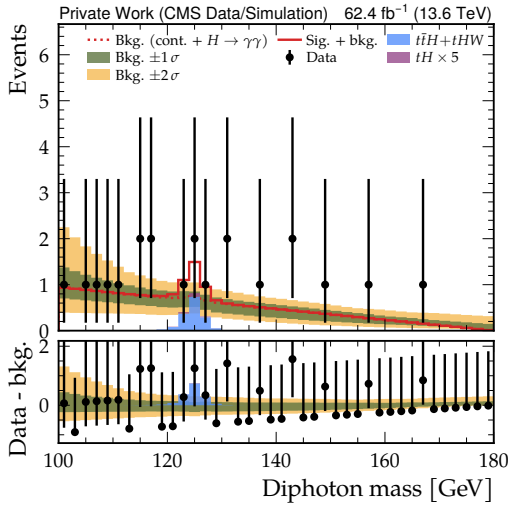
(c) $t\bar{t}H$ had. 3

(d) $t\bar{t}H$ had. 4

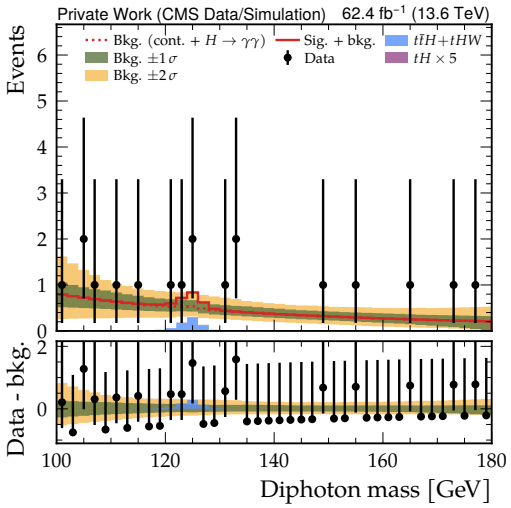
Figure D.1: Observed diphoton mass spectra (top) and background-subtracted residuals (bottom). The red dashed line denotes the post-fit background model, and the red solid line shows the total post-fit model of the signals and all backgrounds. The filled bands around the background model indicate the $\pm1\sigma$ (green) and $\pm2\sigma$ (yellow) post-fit uncertainty on the background prediction. The overlaid signal shapes for $t\bar{t}H$+tHW and tH are shown with their pre-fit SM-expectation in the lower panel, the tH signal is enhanced by a factor 5 in the upper panel to improve visibility.



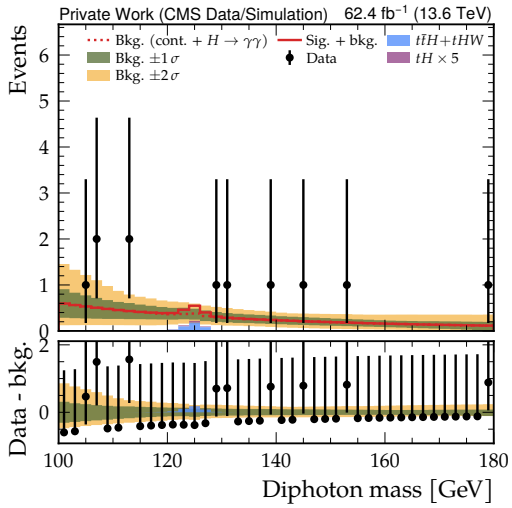
(a) $t\bar{t}H$ lep. 1



(b) $t\bar{t}H$ lep. 2

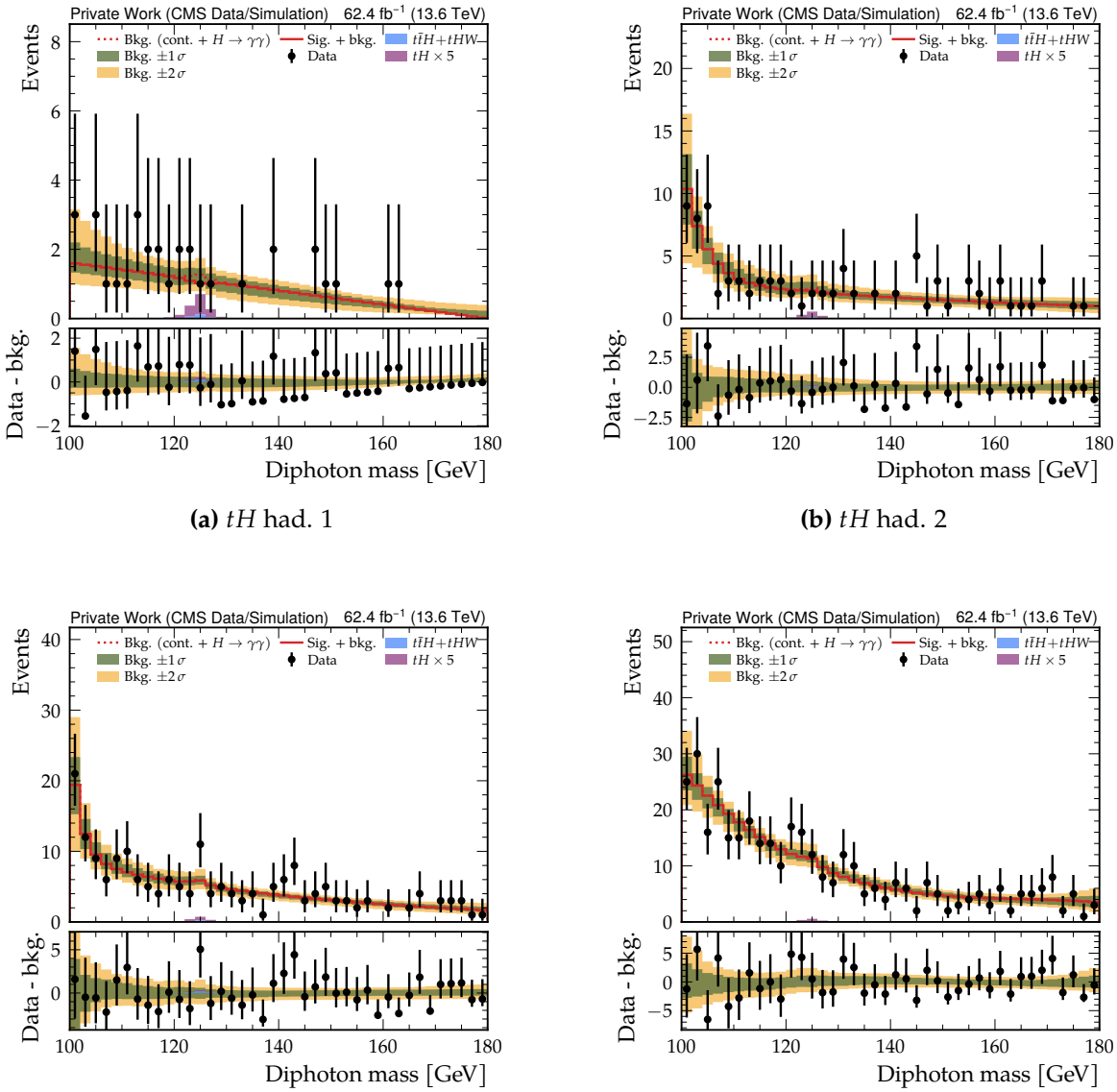


(c) $t\bar{t}H$ lep. 3



(d) $t\bar{t}H$ lep. 4

Figure D.2: Observed diphoton mass spectra (top) and background-subtracted residuals (bottom). The red dashed line denotes the post-fit background model, and the red solid line shows the total post-fit model of the signals and all backgrounds. The filled bands around the background model indicate the $\pm 1\sigma$ (green) and $\pm 2\sigma$ (yellow) post-fit uncertainty on the background prediction. The overlaid signal shapes for $t\bar{t}H$+tHW and tH are shown with their pre-fit SM-expectation in the lower panel, the tH signal is enhanced by a factor 5 in the upper panel to improve visibility.



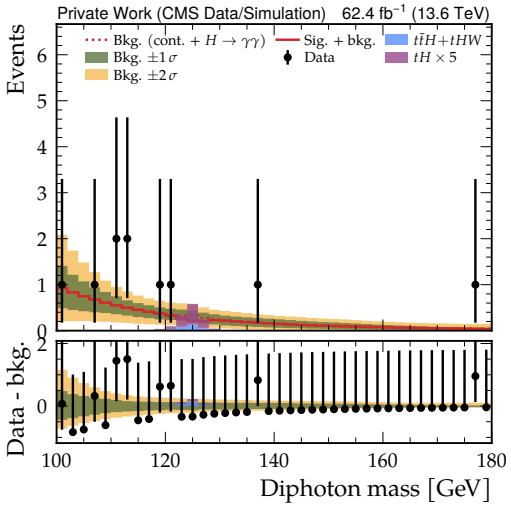
(a) tH had. 1

(b) tH had. 2

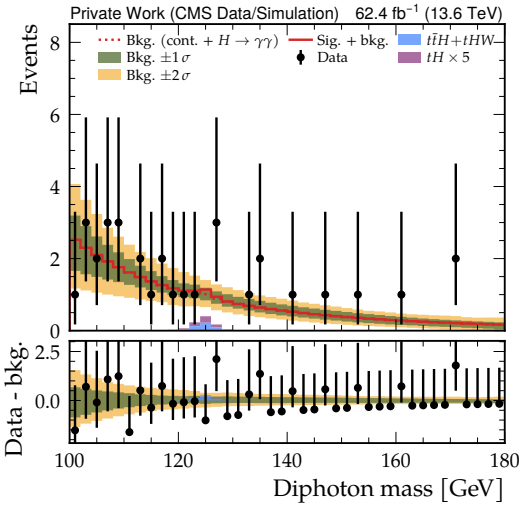
(c) tH had. 3

(d) tH had. 4

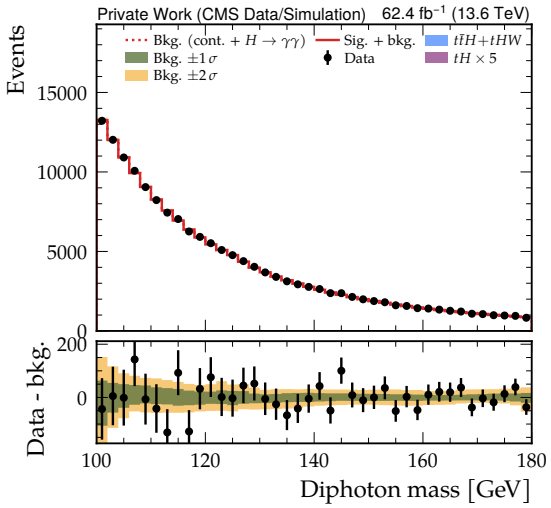
Figure D.3: Observed diphoton mass spectra (top) and background-subtracted residuals (bottom). The red dashed line denotes the post-fit background model, and the red solid line shows the total post-fit model of the signals and all backgrounds. The filled bands around the background model indicate the $\pm1\sigma$ (green) and $\pm2\sigma$ (yellow) post-fit uncertainty on the background prediction. The overlaid signal shapes for $t\bar{t}H$+tHW and tH are shown with their pre-fit SM-expectation in the lower panel, the tH signal is enhanced by a factor 5 in the upper panel to improve visibility.



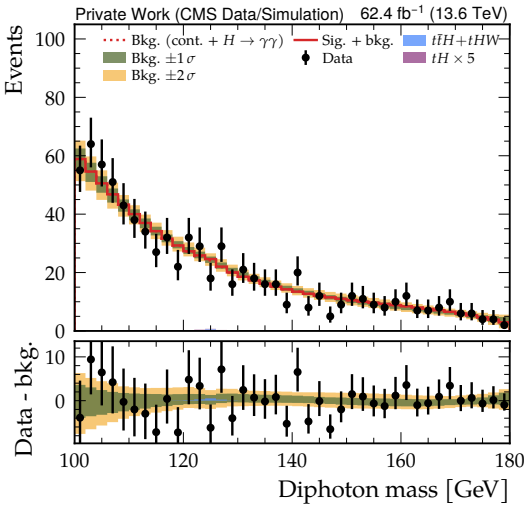
(a) tH lep. 1



(b) tH lep. 2



(c) Bkg. had.



(d) Bkg. lep.

Figure D.4: Observed diphoton mass spectra (top) and background-subtracted residuals (bottom). The red dashed line denotes the post-fit background model, and the red solid line shows the total post-fit model of the signals and all backgrounds. The filled bands around the background model indicate the $\pm1\sigma$ (green) and $\pm2\sigma$ (yellow) post-fit uncertainty on the background prediction. The overlaid signal shapes for $t\bar{t}H$+tHW and tH are shown with their pre-fit SM-expectation in the lower panel, the tH signal is enhanced by a factor 5 in the upper panel to improve visibility.

E. Overview of personal scientific work and publications

This appendix summarises the author's scientific work and publications connected to the doctoral project. Its purpose is to document the collaborative context of this work and to distinguish the author's contributions from common infrastructure from the CMS Collaboration and from contributions made by collaborators, as referenced in the declaration of authorship. In addition to the main analysis of the top-quark associated Higgs boson production presented in this thesis, the work includes complementary research projects closely connected to its scientific scope. These projects concern the analysis of $H \to \gamma\gamma$ events in the CMS experiment, the development of machine-learning methods for collider-data analysis and precision photon reconstruction, and therefore provide methodological and phenomenological context for this thesis. The author's contributions to the studies listed below were carried out under the supervision of Prof. Dr. Johannes Erdmann.

Contributions to the $t\bar{t}H$–tH analysis

The $t\bar{t}H$–tH analysis presented in this thesis was carried out within the CMS Collaboration and builds on the centrally provided reconstruction, calibration, simulation, and analysis infrastructure. It also benefited from regular discussions within the CMS Collaboration, in the working group of Prof. Dr. Johannes Erdmann, and within the Research Training Group "Physics of the Heaviest Particles at the LHC" at the RWTH Aachen University. Within this collaborative environment, the analysis-specific work presented in this thesis was carried out primarily by the author. This includes the development of the analysis strategy and code, the production of most of the signal samples, the development of relevant parts of the CMS $H \to \gamma\gamma$ analysis framework HiggsDNA, the processing of the recorded and simulated datasets, the optimisation of the event selection in the hadronic and the leptonic analysis channel, the construction of the effective continuum-background model, all classifier trainings presented in this thesis, the development of the categorisation-optimisation technique, the statistical analysis and the inclusion of systematic uncertainties, and the study and reduction of the signal-model dependence of the results. All figures and scientific results related to this analysis and presented in this thesis were produced by the author.

First Run-3 measurement of $H \to \gamma\gamma$ cross sections at the CMS experiment

The first measurement of inclusive and differential fiducial Higgs boson production cross sections in the $H \to \gamma\gamma$ decay channel at $\sqrt{s} = 13.6\,\text{TeV}$ was performed within the CMS Collaboration using the first Run-3 data collected in 2022 and published as:

> CMS Collaboration, "Measurements of inclusive and differential Higgs boson production cross sections at $\sqrt{s} = 13.6\,\text{TeV}$ in the $H \to \gamma\gamma$ decay channel", *JHEP* **09** (2025) 070, *doi*:10.1007/JHEP09(2025)070, *arXiv*:2504.17755.

The author's contributions to this measurement include the validation of the photon reconstruction performance for the full 2022 data-taking period, significant work on the HiggsDNA framework, the validation of photon-related corrections and systematic uncertainties, and the optimisation of the categorisation used in the measurement.

Differentiable multi-dimensional categorisation

The differentiable optimisation technique for event categorisation developed for this thesis is described in the following more general methodological study, which has since been accepted for publication in *Eur. Phys. J. C*:

> J. Erdmann, N. K. Kasaraguppe, and F. Mausolf, "Learning to bin: differentiable and Bayesian optimization for multi-dimensional discriminants in high-energy physics", 2026, *arXiv*:2601.07756, accepted for publication in *Eur. Phys. J. C*.

The study applies the approach to toy examples and compares it to a Bayesian optimisation strategy developed by co-authors, while Chapter 8 discusses both the method and its application to the $t\bar{t}H$–tH analysis. The author's contributions include the development of the differentiable optimisation method, the toy-data generation framework, the conceptual design of the toy studies, and major contributions to the writing of the manuscript. The toy-study figures from this study that are included in Chapter 8 were produced by N. K. Kasaraguppe, using code and methodological input contributed by the author and the other co-authors.

Kolmogorov–Arnold networks for event classification

The first application of Kolmogorov–Arnold networks in high-energy physics was published as:

> J. Erdmann, F. Mausolf, and J. L. Späh, "KAN We Improve on HEP Classification Tasks? Kolmogorov–Arnold Networks Applied to an LHC Physics Example", *Comput. Softw. Big Sci.* **9** (2025) 9, *doi*:10.1007/s41781-025-00138-3, *arXiv*:2408.02743.

Their performance, parameter efficiency, and interpretability were studied using the $t\bar{t}H$–tH classification as a representative LHC event-classification task. The study is summarised in Chapter 7. The author's contributions include the preparation of the simulated datasets, with technical support and code contributions from J. L. Späh for sample production, all network trainings and comparison studies shown in the publication, the production of the corresponding plots, major contributions to the writing and review process, and the presentation of the work at the international conference ML4Jets 2024 in Paris.

Study on super resolution for photons

A study on single-image super-resolution techniques for electromagnetic calorimeter data investigated the use of generative adversarial networks to enhance the spatial granularity of photon shower images. Using a simplified calorimeter setup inspired by the CMS electromagnetic calorimeter, it was shown that super-resolved images improve shower-shape reconstruction and position resolution, with a resulting gain in photon-identification performance in the case of small training samples. The work was published as:

> J. Erdmann, A. van der Graaf, F. Mausolf, and O. Nackenhorst, "SR-GAN for SR-gamma: super resolution of photon calorimeter images at collider experiments", *Eur. Phys. J. C* **83** (2023) 1001, *doi*:10.1140/epjc/s10052-023-12178-3, *arXiv*:2308.09025.

The author's contributions include the optimisation and evaluation of the trained networks, major contributions to writing the paper, and the presentation of the work at the international conference ML4Jets 2023 in Hamburg.

Fast calorimeter simulation parameterised on the material budget

A study on a novel fast simulation approach that is parametric in the density and position of material placed in front of an electromagnetic calorimeter was published as:

> J. Erdmann, J. Kann, F. Mausolf, and P. Wissmann, "ParaFlow: Fast calorimeter simulations parameterized in upstream material configurations", *Eur. Phys. J. C* **85** (2025) 857, *doi*:10.1140/epjc/s10052-025-14604-0, *arXiv*:2503.21461.

The author's contributions include the technical supervision of the bachelor's student performing the main parts of the study and contributions to the writing of the paper.

Other work within the CMS Collaboration

Further contributions within the CMS Collaboration include the development of a geometric summary view for the CMS Online Monitoring System for DT muon detectors, enabling fast visual assessment of the data quality during data taking and commissioning, and shifts for monitoring the DT function during data taking. In addition, the author contributed to the ongoing development of a low-level deep-learning approach to photon identification by providing the code framework to prepare the corresponding training datasets and by providing technical supervision for two master's students and a bachelor's student who carried out the initial training and evaluation studies. The author also served as contact person between the CMS electron and photon group and the machine-learning group, improving the information exchange and coordination between both groups.

Bibliography

[1] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] CMS Collaboration, "Observation of a New Boson with Mass Near 125 GeV in pp Collisions at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[4] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[5] CMS Collaboration, "The CMS Experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[6] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.

[7] S. L. Glashow, "Partial-symmetries of weak interactions", *Nucl. Phys.* **22** (1961) 579, doi:https://doi.org/10.1016/0029-5582(61)90469-2.

[8] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (1967) 1264, doi:10.1103/PhysRevLett.19.1264.

[9] A. Salam, "Weak and Electromagnetic Interactions", *Conf. Proc. C* **680519** (1968) 367, doi:10.1142/9789812795915_0034.

[10] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132, doi:https://doi.org/10.1016/0031-9163(64)91136-9.

[11] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", *Phys. Rev. Lett.* **13** (1964) 321, doi:10.1103/PhysRevLett.13.321.

[12] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 508, doi:10.1103/PhysRevLett.13.508.

[13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles", *Phys. Rev. Lett.* **13** (1964) 585, doi:10.1103/PhysRevLett.13.585.

[14] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons", *Phys. Rev.* **145** (1966) 1156, doi:10.1103/PhysRev.145.1156.

[15] T. W. B. Kibble, "Symmetry breaking in non-abelian gauge theories", *Phys. Rev.* **155** (1967) 1554, doi:10.1103/PhysRev.155.1554.

[16] S. Alekhin, A. Djouadi, and S. Moch, "The top quark and Higgs boson masses and the stability of the electroweak vacuum", *Phys. Lett. B* **716** (2012) 214, doi:10.1016/j.physletb.2012.08.024, arXiv:1207.0980.

[17] G. Degrassi et al., "Higgs mass and vacuum stability in the Standard Model at NNLO", *JHEP* **08** (2012) 098, doi:10.1007/JHEP08(2012)098, arXiv:1205.6497.

[18] B. A. Dobrescu and C. T. Hill, "Electroweak symmetry breaking via top condensation seesaw", *Phys. Rev. Lett.* **81** (1998) 2634, doi:10.1103/PhysRevLett.81.2634, arXiv:hep-ph/9712319.

[19] R. S. Chivukula, B. A. Dobrescu, H. Georgi, and C. T. Hill, "Top Quark Seesaw Theory of Electroweak Symmetry Breaking", *Phys. Rev. D* **59** (1999) 075003, doi:10.1103/PhysRevD.59.075003, arXiv:hep-ph/9809470.

[20] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pene, "Standard model CP violation and baryon asymmetry", *Mod. Phys. Lett. A* **9** (1994) 795, doi:10.1142/S0217732394000629, arXiv:hep-ph/9312215.

[21] P. Huet and E. Sather, "Electroweak baryogenesis and standard model CP violation", *Phys. Rev. D* **51** (1995) 379, doi:10.1103/PhysRevD.51.379, arXiv:hep-ph/9404302.

[22] CMS Collaboration, "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV", *Phys. Rev. D* **92** (2015) 012004, doi:10.1103/PhysRevD.92.012004, arXiv:1411.3441.

[23] ATLAS Collaboration, "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector", *Eur. Phys. J. C* **75** (2015) 476, doi:10.1140/epjc/s10052-015-3685-1, arXiv:1506.05669. [Erratum: Eur. Phys. J. C **76** (2016) 152].

[24] CMS Collaboration, "Analysis of the CP structure of the Yukawa coupling between the Higgs boson and τ leptons in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **06** (2022) 012, doi:10.1007/JHEP06(2022)012, arXiv:2110.04836.

[25] ATLAS Collaboration, "Measurement of the CP properties of Higgs boson interactions with τ-leptons with the ATLAS detector", *Eur. Phys. J. C* **83** (2023) 563, doi:10.1140/epjc/s10052-023-11583-y, arXiv:2212.05833.

[26] CMS Collaboration, "Measurements of $t\bar{t}H$ production and the CP structure of the Yukawa interaction between the Higgs boson and top quark in the diphoton decay channel", *Phys. Rev. Lett.* **125** (2020) 061801, doi:10.1103/PhysRevLett.125.061801, arXiv:2003.10866.

[27] ATLAS Collaboration, "CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and tH Processes Using $H \to \gamma\gamma$ with the ATLAS Detector", *Phys. Rev. Lett.* **125** (2020) 061802, doi:10.1103/PhysRevLett.125.061802, arXiv:2004.04545.

[28] CMS Collaboration, "Measurement of the associated production of a top-antitop-quark pair and a Higgs boson, or of a single top quark and a Higgs boson, in the $H \to b\bar{b}$ decay channel in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **02** (2025) 097, doi:10.1007/JHEP02(2025)097, arXiv:2407.10896.

[29] F. Demartin et al., "Higgs characterisation at NLO in QCD: *CP* properties of the top-quark Yukawa interaction", *Eur. Phys. J. C* **74** (2014) 3065, doi:10.1140/epjc/s10052-014-3065-2, arXiv:1407.5089.

[30] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, "Higgs production in association with a single top quark at the LHC", *Eur. Phys. J. C* **75** (2015) 267, doi:10.1140/epjc/s10052-015-3475-9, arXiv:1504.00611.

[31] F. Demartin et al., "tWH associated production at the LHC", *Eur. Phys. J. C* **77** (2017) 34, doi:10.1140/epjc/s10052-017-4601-7, arXiv:1607.05862.

[32] ATLAS Collaboration, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector", *Phys. Lett. B* **784** (2018) 173, doi:10.1016/j.physletb.2018.07.035, arXiv:1806.00425.

[33] CMS Collaboration, "Observation of $t\bar{t}H$ production", *Phys. Rev. Lett.* **120** (2018) 231801, doi:10.1103/PhysRevLett.120.231801, arXiv:1804.02610.

[34] CMS Collaboration, "Combined measurements and interpretations of Higgs boson production and decay in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", 2026, arXiv:2602.18611.

[35] ATLAS Collaboration, "Search for the production of a Higgs boson in association with a single top quark in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *JHEP* **10** (2025) 093, doi:10.1007/JHEP10(2025)093, arXiv:2508.14695.

[36] ATLAS Collaboration, "Measurement of the Higgs boson production in association with top quarks in multilepton final states in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", 2025, arXiv:2510.23755.

[37] CMS Collaboration, "A portrait of the Higgs boson by the CMS experiment ten years after the discovery.", *Nature* **607** (2022) 60, doi:10.1038/s41586-022-04892-x, arXiv:2207.00043. [Erratum: Nature **623** (2023) E4].

[38] ATLAS Collaboration, "A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery", *Nature* **607** (2022) 52, doi:10.1038/s41586-022-04893-w, arXiv:2207.00092. [Erratum: Nature **612** (2022) E24].

[39] CMS Collaboration, "Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **07** (2021) 027, doi:10.1007/JHEP07(2021)027, arXiv:2103.06956.

[40] ATLAS Collaboration, "Measurement of the properties of the Higgs boson production at $\sqrt{s} = 13\,\text{TeV}$ in the $H \to \gamma\gamma$ channel using $139\,\text{fb}^{-1}$ of pp collision data with the ATLAS experiment", *JHEP* **07** (2023) 088, doi:10.1007/JHEP07(2023)088, arXiv:2207.00348.

[41] CMS Collaboration, "Measurement of the Higgs boson inclusive and differential fiducial production cross sections in the diphoton decay channel with pp collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **07** (2023) 091, doi:10.1007/JHEP07(2023)091, arXiv:2208.12279.

[42] ATLAS Collaboration, "Measurements of the Higgs boson inclusive and differential fiducial cross-sections in the diphoton decay channel with pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **08** (2022) 027, doi:10.1007/JHEP08(2022)027, arXiv:2202.00487.

[43] CMS Collaboration, "A measurement of the Higgs boson mass in the diphoton decay channel with the CMS detector", CMS-PAS-HIG-24-007 (2026), https://cds.cern.ch/record/2957193.

[44] ATLAS Collaboration, "Measurement of the Higgs boson mass with $H \to \gamma\gamma$ decays in 140 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector", *Phys. Lett. B* **847** (2023) 138315, doi:10.1016/j.physletb.2023.138315, arXiv:2308.07216.

[45] CMS Collaboration, "Measurements of inclusive and differential Higgs boson production cross sections at $\sqrt{s} = 13.6$ TeV in the $H \to \gamma\gamma$ decay channel", *JHEP* **09** (2025) 070, doi:10.1007/JHEP09(2025)070, arXiv:2504.17755.

[46] ATLAS Collaboration, "Measurement of the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ cross-sections in pp collisions at $\sqrt{s} = 13.6$ TeV with the ATLAS detector", *Eur. Phys. J. C* **84** (2024) 78, doi:10.1140/epjc/s10052-023-12130-5, arXiv:2306.11379.

[47] ATLAS Collaboration, "Search for anomalies in vector-boson fusion production of the Higgs boson in $H(\to \gamma\gamma)jj$ events using 164 fb^{-1} of pp collision data collected at $\sqrt{s} = 13.6$ TeV with the ATLAS detector", 2026, arXiv:2603.20087.

[48] H. Bahl et al., "Indirect CP probes of the Higgs-top-quark interaction: current LHC constraints and future opportunities", *JHEP* **11** (2020) 127, doi:10.1007/JHEP11(2020)127, arXiv:2007.08542.

[49] CMS Collaboration, "Model-independent measurement of the Higgs boson associated production with two jets and decaying to a pair of W bosons in proton–proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2026) 073, doi:10.1007/JHEP03(2026)073, arXiv:2509.07958.

[50] J. Huston, K. Rabbertz, and G. Zanderighi, "Quantum Chromodynamics", *Phys. Rev. D* **110** (2024) 030001, doi:10.1103/PhysRevD.110.030001. Chap. 9 in *Review of Particle Physics*.

[51] J. de Blas, S. Dittmaier, and R. Kogler, "Electroweak Model and Constraints on New Physics", *Phys. Rev. D* **110** (2024) 030001, doi:10.1103/PhysRevD.110.030001. Chap. 10 in *Review of Particle Physics*.

[52] M. Cepeda, L. Reina, and P. Savard, "Status of Higgs boson physics", *Phys. Rev. D* **110** (2024) 030001, doi:10.1103/PhysRevD.110.030001. Chap. 11 in *Review of Particle Physics*.

[53] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories", *Phys. Rev. Lett.* **30** (1973) 1343, doi:10.1103/PhysRevLett.30.1343.

[54] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys. Rev. Lett.* **30** (1973) 1346, doi:10.1103/PhysRevLett.30.1346.

[55] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys. Rev. Lett.* **10** (1963) 531, doi:10.1103/PhysRevLett.10.531.

[56] M. Kobayashi and T. Maskawa, "*CP* Violation in the Renormalizable Theory of Weak Interaction", *Prog. Theor. Phys.* **49** (1973) 652, doi:10.1143/PTP.49.652.

[57] Super-Kamiokande Collaboration, "Evidence for oscillation of atmospheric neutrinos", *Phys. Rev. Lett.* **81** (1998) 1562, doi:10.1103/PhysRevLett.81.1562, arXiv:hep-ex/9807003.

[58] SNO Collaboration, "Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory", *Phys. Rev. Lett.* **89** (2002) 011301, doi:10.1103/PhysRevLett.89.011301, arXiv:nucl-ex/0204008.

[59] D0 Collaboration, "Observation of the top quark", *Phys. Rev. Lett.* **74** (1995) 2632, doi:10.1103/PhysRevLett.74.2632, arXiv:hep-ex/9503003.

[60] CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions", *Phys. Rev. Lett.* **74** (1995) 2626, doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002.

[61] C. Burgard, "Standard model of physics", 2016, https://texample.net/tikz/examples/model-physics/, accessed 2026-05-18.

[62] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln", *Helv. Phys. Acta* **6** (1933) 110, doi:10.1007/s10714-008-0707-4.

[63] V. C. Rubin, J. Ford, W. K., and N. Thonnard, "Rotational Properties of 21 Sc Galaxies with a Large Range of Luminosities and Radii, from NGC 4605 ($R = 4$ kpc) to UGC 2885 ($R = 122$ kpc)", *Astrophys. J.* **238** (1980) 471, doi:10.1086/158003.

[64] A. D. Sakharov, "Violation of *CP* invariance, *C* asymmetry, and baryon asymmetry of the universe", *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32, doi:10.1070/PU1991v034n05ABEH002497.

[65] L. Susskind, "Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory", *Phys. Rev. D* **20** (1979) 2619, doi:10.1103/PhysRevD.20.2619.

[66] K. Agashe, R. Contino, and A. Pomarol, "The Minimal composite Higgs model", *Nucl. Phys. B* **719** (2005) 165, doi:10.1016/j.nuclphysb.2005.04.035, arXiv:hep-ph/0412089.

[67] G. C. Branco et al., "Theory and phenomenology of two-Higgs-doublet models", *Phys. Rept.* **516** (2012) 1, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.

[68] D. Fontes, J. C. Romão, R. Santos, and J. P. Silva, "Large pseudoscalar Yukawa couplings in the complex 2HDM", *JHEP* **06** (2015) 060, doi:10.1007/JHEP06(2015)060, arXiv:1502.01720.

[69] F. Boudjema, R. M. Godbole, D. Guadagnoli, and K. A. Mohan, "Lab-frame observables for probing the top-Higgs interaction", *Phys. Rev. D* **92** (2015) 015019, doi:10.1103/PhysRevD.92.015019, arXiv:1501.03157.

[70] P. Artoisenet et al., "A framework for Higgs characterisation", *JHEP* **11** (2013) 043, doi:10.1007/JHEP11(2013)043, arXiv:1306.6464.

[71] I. Brivio and M. Trott, "The Standard Model as an Effective Field Theory", *Phys. Rept.* **793** (2019) 1, doi:10.1016/j.physrep.2018.11.002, arXiv:1706.08945.

[72] D. de Florian et al., "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", *CERN Yellow Rep. Monogr.* **2** (2017) 1, doi:10.23731/CYRM-2017-002, arXiv:1610.07922.

[73] C. Degrande et al., "Single-top associated production with a Z or H boson at the LHC: the SMEFT interpretation", *JHEP* **10** (2018) 005, doi:10.1007/JHEP10(2018)005, arXiv:1804.07773.

[74] CMS Collaboration, "Combined effective field theory interpretation of Higgs boson, electroweak vector boson, top quark, and multi-jet measurements", *Eur. Phys. J. C* **86** (2026) 331, doi:10.1140/epjc/s10052-025-14997-y, arXiv:2504.02958.

[75] A. V. Gritsan, R. Röntsch, M. Schulze, and M. Xiao, "Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques", *Phys. Rev. D* **94** (2016) 055023, doi:10.1103/PhysRevD.94.055023, arXiv:1606.03107.

[76] E. Aschenauer, R. Thorne, and R. Yoshida, "Structure functions", *Phys. Rev. D* **110** (2024) 030001, doi:10.1103/PhysRevD.110.030001. Chap. 18 in *Review of Particle Physics*.

[77] S. Forte and G. Watt, "Progress in the Determination of the Partonic Structure of the Proton", *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 291, doi:10.1146/annurev-nucl-102212-170607, arXiv:1301.6754.

[78] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[79] V. Gribov and L. Lipatov, "Deep inelastic ep scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438.

[80] L. Lipatov, "The parton model and perturbation theory", *Sov. J. Nucl. Phys.* **20** (1975) 94.

[81] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language", *Nucl. Phys. B* **126** (1977) 298, doi:10.1016/0550-3213(77)90384-4.

[82] Y. L. Dokshitzer, "Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics", *Sov. Phys. JETP* **46** (1977) 641.

[83] J. C. Collins, D. E. Soper, and G. F. Sterman, "Factorization of Hard Processes in QCD", *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1, doi:10.1142/9789814503266_0001, arXiv:hep-ph/0409313.

[84] C. Bierlich et al., "A comprehensive guide to the physics and usage of PYTHIA 8.3", *SciPost Phys. Codeb.* (2022) 8, doi:10.21468/SciPostPhysCodeb.8, arXiv:2203.11601.

[85] A. Karlberg et al., "Ad interim recommendations for the Higgs boson production cross sections at $\sqrt{s} = 13.6\,\text{TeV}$", 2024, arXiv:2402.09955.

[86] R. Balsach et al., "State-of-the-art cross sections for $t\bar{t}H$: NNLO predictions matched with NNLL resummation and EW corrections", *SciPost Phys. Comm. Rep. 10* (2025) doi:10.21468/SciPostPhysCommRep.10, arXiv:2503.15043.

[87] C. Biello, C. Savoini, C. Signorile-Signorile, and M. Wiesemann, "Next-to-next-to-leading order event generation for $t\bar{t}H$ production with approximate two-loop amplitude", 2026, arXiv:2603.06143.

[88] J. Hermann, D. Stremmer, and M. Worek, "CP structure of the top-quark Yukawa interaction: NLO QCD corrections and off-shell effects", *JHEP* **09** (2022) 138, doi:10.1007/JHEP09(2022)138, arXiv:2205.09983.

[89] G. Bevilacqua, M. Reinartz, and M. Worek, "SMEFT everywhere: a NLO study of $pp \to t\bar{t}H$ with decaying tops", 2026, arXiv:2605.11747.

[90] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, "Determining the Structure of Higgs Couplings at the LHC", *Phys. Rev. Lett.* **88** (2002) 051801, doi:10.1103/PhysRevLett.88.051801, arXiv:hep-ph/0105325.

[91] D. Pagani, I. Tsinikos, and E. Vryonidou, "NLO QCD+EW predictions for tHj and tZj production at the LHC", *JHEP* **08** (2020) 082, doi:10.1007/JHEP08(2020)082, arXiv:2006.10086.

[92] Particle Data Group, "Review of Particle Physics", *Phys. Rev. D* **110** (2024) 030001, doi:10.1103/PhysRevD.110.030001.

[93] ATLAS Collaboration, "Combined measurements of Higgs boson production and decay at $\sqrt{s} = 13$ TeV using up to 140 fb^{-1} of data collected by the ATLAS Experiment", ATLAS-CONF-2025-006 (2025), https://cds.cern.ch/record/2937634.

[94] CMS Collaboration, "Evidence for Higgs boson decay to a pair of muons", *JHEP* **01** (2021) 148, doi:10.1007/JHEP01(2021)148, arXiv:2009.04363.

[95] ATLAS Collaboration, "Evidence for the Dimuon Decay of the Higgs Boson in pp Collisions with the ATLAS Detector", *Phys. Rev. Lett.* **135** (2025) 231802, doi:10.1103/gzdh-p159, arXiv:2507.03595.

[96] ATLAS, CMS Collaboration, "Evidence for the Higgs Boson Decay to a Z Boson and a Photon at the LHC", *Phys. Rev. Lett.* **132** (2024) 021803, doi:10.1103/PhysRevLett.132.021803, arXiv:2309.03501.

[97] CMS Collaboration, "Measurement of the Higgs boson mass and width using the four-lepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **111** (2025) 092014, doi:10.1103/PhysRevD.111.092014, arXiv:2409.13663.

[98] CMS Collaboration, "Constraints on the Higgs boson total decay width using signal-background interference in the diphoton final state with proton–proton collisions at $\sqrt{s} = 13$ TeV", CMS-PAS-HIG-25-004 (2025), https://cds.cern.ch/record/2941115.

[99] ATLAS Collaboration, "Evidence of off-shell Higgs boson production from ZZ leptonic decay channels and constraints on its total width with the ATLAS detector", *Phys. Lett. B* **846** (2023) 138223, doi:10.1016/j.physletb.2023.138223, arXiv:2304.01532.

[100] CMS Collaboration, "Analysis of the CP structure of the Yukawa coupling between the Higgs boson and tau leptons in proton-proton collisions at $\sqrt{s} = 13.6\,\text{TeV}$", 2026, arXiv:2606.03510.

[101] CMS Collaboration, "Constraints on anomalous Higgs boson couplings to vector bosons and fermions in its production and decay using the four-lepton final state", *Phys. Rev. D* **104** (2021) 052004, doi:10.1103/PhysRevD.104.052004, arXiv:2104.12152.

[102] CMS Collaboration, "Constraints on anomalous Higgs boson couplings from its production and decay using the WW channel in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **84** (2024) 779, doi:10.1140/epjc/s10052-024-12925-0, arXiv:2403.00657.

[103] ATLAS Collaboration, "Combination of measurements of CP properties of Higgs boson interactions with vector bosons using proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", 2026, arXiv:2603.20117.

[104] ATLAS Collaboration, "Determination of the relative sign of the Higgs boson couplings to W and Z bosons using WH production via vector-boson fusion with the ATLAS detector", *Phys. Rev. Lett.* **133** (2024) 141801, doi:10.1103/PhysRevLett.133.141801, arXiv:2402.00426.

[105] CMS Collaboration, "Study of WH production through vector boson scattering and extraction of the relative sign of the W and Z couplings to the Higgs boson in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Lett. B* **860** (2025) 139202, doi:10.1016/j.physletb.2024.139202, arXiv:2405.16566.

[106] N. Berger et al., "Simplified Template Cross Sections - Stage 1.1", 2019, arXiv:1906.02754.

[107] J. Brod, U. Haisch, and J. Zupan, "Constraints on CP-violating Higgs couplings to the third generation", *JHEP* **11** (2013) 180, doi:10.1007/JHEP11(2013)180, arXiv:1310.1385.

[108] CMS Collaboration, "Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the lepton+jets final state in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Rev. D* **100** (2019) 072007, doi:10.1103/PhysRevD.100.072007, arXiv:1907.01590.

[109] CMS Collaboration, "Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the dilepton final state in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Rev. D* **102** (2020) 092013, doi:10.1103/PhysRevD.102.092013, arXiv:2009.07123.

[110] ATLAS Collaboration, "Measurement of the top-quark Yukawa coupling from $t\bar{t}$ production in the lepton+jets final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **01** (2026) 117, doi:10.1007/JHEP01(2026)117, arXiv:2509.16039.

[111] CMS Collaboration, "Search for production of four top quarks in final states with same-sign or multiple leptons in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **80** (2020) 75, doi:10.1140/epjc/s10052-019-7593-7, arXiv:1908.06463.

[112] ATLAS Collaboration, "Observation of four-top-quark production in the multilepton final state with the ATLAS detector", *Eur. Phys. J. C* **83** (2023) 496, doi:10.1140/epjc/s10052-023-11573-0, arXiv:2303.15061.

[113] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[114] LHCb Collaboration, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[115] CMS Collaboration, "Public CMS Luminosity Information", https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults, 2026, Accessed: 2026-06-01.

[116] R. Steerenberg, "Accelerator Report: Full house in the LHC", 2023, https://home.cern/news/news/accelerators/accelerator-report-full-house-lhc, accessed 2026-05-11.

[117] CMS Collaboration, "Development of the CMS detector for the CERN LHC Run 3", *JINST* **19** (2024) P05064, doi:10.1088/1748-0221/19/05/P05064, arXiv:2309.05466.

[118] CMS Collaboration, "Description and Performance of Track and Primary-Vertex Reconstruction with the CMS Tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[119] CMS Tracker Group Collaboration, "The CMS Phase-1 Pixel Detector Upgrade", *JINST* **16** (2021) P02027, doi:10.1088/1748-0221/16/02/P02027, arXiv:2012.14304.

[120] CMS Collaboration, "Performance of the CMS phase-1 pixel detector with Run 3 data", CMS-DP-2022-047 (2022), https://cds.cern.ch/record/2839741.

[121] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton–Proton Collisions at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[122] CMS Collaboration, "Performance of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **19** (2024) P09004, doi:10.1088/1748-0221/19/09/P09004, arXiv:2403.15518.

[123] CMS ECAL Collaboration, "Laser monitoring for the CMS ECAL", *Nucl. Instrum. Meth. A* **623** (2010) 267, doi:10.1016/j.nima.2010.02.217.

[124] CMS Collaboration, "The ultimate CMS ECAL calibration and performance for the legacy reprocessing of LHC Run 2 data", *PoS* **ICHEP2022** (2023) 643, doi:10.22323/1.414.0643.

[125] CMS Collaboration, "CMS electromagnetic calorimeter calibration and alignment", *Nucl. Instrum. Meth. A* **936** (2019) 121, doi:10.1016/j.nima.2018.10.027.

[126] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[127] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[128] CMS Collaboration, "CMSSW", GitHub repository, 2026, https://github.com/cms-sw/cmssw, accessed 2026-05-04.

[129] D. Contardo et al., "Technical Proposal for the Phase-II Upgrade of the CMS Detector", CMS-TDR-15-02 (2015), https://cds.cern.ch/record/2020886.

[130] CMS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **11** (2018) 185, doi:10.1007/JHEP11(2018)185, arXiv:1804.02716.

[131] CMS Collaboration, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton–Proton Collisions at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015) P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.

[132] TMVA Collaboration, "TMVA - Toolkit for Multivariate Data Analysis", 2007, arXiv:physics/0703039.

[133] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021) P05014, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

[134] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System", in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, p. 785, 2016, doi:10.1145/2939672.2939785.

[135] CMS Collaboration, "Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8\,\text{TeV}$", *Eur. Phys. J. C* **76** (2016) 13, doi:10.1140/epjc/s10052-015-3853-3, arXiv:1508.07819.

[136] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7\,\text{TeV}$", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[137] C. C. Daumann et al., "One Flow to Correct Them all: Improving Simulations in High-Energy Physics with a Single Normalising Flow and a Switch", *Comput. Softw. Big Sci.* **8** (2024) 15, doi:10.1007/s41781-024-00125-0, arXiv:2403.18582.

[138] G. Papamakarios, T. Pavlakou, and I. Murray, "Masked autoregressive flow for density estimation", 2018, arXiv:1705.07057.

[139] G. Papamakarioset al., "Normalizing flows for probabilistic modeling and inference", *J. Mach. Learn. Res.* **22** (2021) 1, arXiv:1912.02762.

[140] F. Couderc, P. Gaigne, and M. Ö. Sahin, "Lepton energy scale and resolution corrections based on the minimization of an analytical likelihood: IJazZ2.0", 2026, arXiv:2602.17300.

[141] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[142] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[143] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup Per Particle Identification", *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[144] CMS Collaboration, "Jet Energy Scale and Resolution Measurements Using Run 3 Data Collected by CMS in 2022 and 2023 at 13.6 TeV", CMS-DP-2024-039 (2024), https://cds.cern.ch/record/2902862.

[145] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/P09018, arXiv:2003.00503.

[146] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[147] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[148] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[149] CMS Collaboration, "Performance summary of AK4 jet b tagging with data from 2022 proton–proton collisions at 13.6 TeV with the CMS detector", CMS-DP-2024-025 (2024), https://cds.cern.ch/record/2898464.

[150] H. Qu, C. Li, and S. Qian, "Particle Transformer for Jet Tagging", 2022, arXiv:2202.03772.

[151] A. Vaswani et al., "Attention is all you need", *Adv. Neural Inf. Process. Syst.* **30** (2017) 5998, arXiv:1706.03762.

[152] A. Stein et al., "Improving Robustness of Jet Tagging Algorithms with Adversarial Training", *Comput. Softw. Big Sci.* **6** (2022) 15, doi:10.1007/s41781-022-00087-1, arXiv:2203.13890.

[153] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton–proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[154] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth. A* **262** (1987) 444, doi:10.1016/0168-9002(87)90887-4.

[155] CMS Collaboration, "Muon momentum calibration with proton–proton collisions at $\sqrt{s} = 13.6$ TeV", CMS-DP-2024-065 (2024), https://cds.cern.ch/record/2904701.

[156] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton–proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[157] CMS Collaboration, "Simultaneous Probe of the Charm and Bottom Quark Yukawa Couplings Using $t\bar{t}H$ Events", *Phys. Rev. Lett.* **136** (2026) 011801, doi:10.1103/9nwb-splk, arXiv:2509.22535.

[158] ATLAS Collaboration, "Probing the CP nature of the top-Higgs Yukawa coupling in $t\bar{t}H$ and tH events with $H \to b\bar{b}$ decays using the ATLAS detector at the LHC", *Phys. Lett. B* **849** (2024) 138469, doi:10.1016/j.physletb.2024.138469, arXiv:2303.05974.

[159] CMS Collaboration, "Search for associated production of a Higgs boson and a single top quark in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Rev. D* **99** (2019) 092005, doi:10.1103/PhysRevD.99.092005, arXiv:1811.09696.

[160] G. P. Salam and E. Slade, "Cuts for two-body decays at colliders", *JHEP* **11** (2021) 220, doi:10.1007/JHEP11(2021)220, arXiv:2106.08329.

[161] CMS Collaboration, "HiggsDNA", GitLab repository, 2026, https://gitlab.cern.ch/cms-analysis/general/HiggsDNA, accessed 2026-05-04.

[162] A. Rizzi, G. Petrucciani, and M. Peruzzi, "A further reduction in CMS event data for analysis: the NANOAOD format", *EPJ Web Conf.* **214** (2019) 06021, doi:10.1051/epjconf/201921406021.

[163] M. Peruzzi, G. Petrucciani, A. Rizzi, and for the CMS Collaboration, "The NanoAOD event data format in CMS", *J. Phys. Conf. Ser.* **1525** (2020) 012038, doi:10.1088/1742-6596/1525/1/012038.

[164] N. C. Smith and L. A. Gray, "Evaluating awkward arrays, uproot, and coffea as a query platform for High Energy Physics data", *J. Phys. Conf. Ser.* **2438** (2023) 012033, doi:10.1088/1742-6596/2438/1/012033.

[165] J. Pivarski, J. Nandi, D. Lange, and P. Elmer, "Columnar data processing for HEP analysis", *EPJ Web Conf.* **214** (2019) 06026, doi:10.1051/epjconf/201921406026.

[166] J. Pivarski, D. Lange, and P. Elmer, "Nested data structures in array frameworks", *J. Phys. Conf. Ser.* **1525** (2020) 012053, doi:10.1088/1742-6596/1525/1/012053.

[167] J. Pivarski, P. Elmer, and D. Lange, "Awkward Arrays in Python, C++, and Numba", *EPJ Web Conf.* **245** (2020) 05023, doi:10.1051/epjconf/202024505023.

[168] CMS Collaboration, "flashgg", GitHub repository, 2026, https://github.com/cms-analysis/flashgg, accessed 2026-05-04.

[169] R. Brun and F. Rademakers, "ROOT — An object oriented data analysis framework", *Nucl. Instrum. Meth. A* **389** (1997) 81, doi:10.1016/s0168-9002(97)00048-x.

[170] E. M. Fajardo et al., "How much higher can HTCondor fly?", *J. Phys. Conf. Ser.* **664** (2015) 062014, doi:10.1088/1742-6596/664/6/062014.

[171] M. A. Jette and T. Wickberg, "Architecture of the Slurm Workload Manager", in *Job Scheduling Strategies for Parallel Processing*, p. 3, 2023, doi:10.1007/978-3-031-43943-8_1.

[172] Y. Babuji et al., "Parsl: Pervasive parallel programming in python", in *Proc. 28th ACM Int. Symp. High-Performance Parallel and Distributed Computing (HPDC)*, p. 25, 2019, doi:10.1145/3307681.3325400.

[173] CMS Collaboration, "Search for HH production decaying into two b quarks and two photons in pp collisions at 13.6 TeV with a partial CMS Run 3 dataset", CMS-PAS-HIG-25-007 (2026), https://cds.cern.ch/record/2945062.

[174] CMS Collaboration, "Search for a cH signal in the associated production of at least one charm quark with a Higgs boson in the diphoton decay channel in pp collisions at $\sqrt{s} = 13$ TeV", *JHEP* **11** (2025) 060, doi:10.1007/JHEP11(2025)060, arXiv:2503.08797.

[175] CMS Collaboration, "Luminosity measurement in proton–proton collisions at 13.6 TeV in 2022 at CMS", CMS-PAS-LUM-22-001 (2024), https://cds.cern.ch/record/2890833.

[176] CMS Collaboration, "Measurement of the offline integrated luminosity for the CMS proton–proton collision dataset recorded in 2023", CMS-DP-2024-068 (2024), https://cds.cern.ch/record/2904808.

[177] S. van der Meer, "Calibration of the effective beam height in the ISR", CERN-ISR-PO-68-31 (1968), https://cds.cern.ch/record/296752.

[178] P. Grafström and W. Kozanecki, "Luminosity determination at proton colliders", *Prog. Part. Nucl. Phys.* **81** (2015) 97, doi:10.1016/j.ppnp.2014.11.002.

[179] CMS Collaboration, "Electron and photon reconstruction and identification performance at CMS in 2022 and 2023", CMS-DP-2024-052 (2024), https://cds.cern.ch/record/2904365.

[180] GEANT4 Collaboration, "GEANT4 - A Simulation Toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[181] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[182] V. Hirschi and O. Mattelaer, "Automated event generation for loop-induced processes", *JHEP* **10** (2015) 146, doi:10.1007/JHEP10(2015)146, arXiv:1507.00020.

[183] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[184] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[185] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.

[186] J. S. Gainer et al., "Exploring Theory Space with Monte Carlo Reweighting", *JHEP* **10** (2014) 078, doi:10.1007/JHEP10(2014)078, arXiv:1404.7129.

[187] O. Mattelaer, "On the maximal use of Monte Carlo samples: re-weighting events at NLO accuracy", *Eur. Phys. J. C* **76** (2016) 674, doi:10.1140/epjc/s10052-016-4533-7, arXiv:1607.00763.

[188] W. Beenakker, R. Hopker, M. Spira, and P. M. Zerwas, "Squark and gluino production at hadron colliders", *Nucl. Phys. B* **492** (1997) 51, doi:10.1016/S0550-3213(97)80027-2, arXiv:hep-ph/9610490.

[189] S. Dawson and L. Reina, "QCD corrections to associated Higgs boson production", *Phys. Rev. D* **57** (1998) 5851, doi:10.1103/PhysRevD.57.5851, arXiv:hep-ph/9712400.

[190] U. Haisch, P. Pani, and G. Polesello, "Determining the *CP* nature of spin-0 mediators in associated production of dark matter and $t\bar{t}$ pairs", *JHEP* **02** (2017) 131, doi:10.1007/JHEP02(2017)131, arXiv:1611.09841.

[191] S. Dittmaier et al., "Higgs radiation off quarks in supersymmetric theories at e^+e^- colliders", *Phys. Lett. B* **478** (2000) 247, doi:10.1016/S0370-2693(00)00278-1, arXiv:hep-ph/0002035.

[192] K. Hamilton, P. Nason, and G. Zanderighi, "MINLO: Multi-Scale Improved NLO", *JHEP* **10** (2012) 155, doi:10.1007/JHEP10(2012)155, arXiv:1206.3572.

[193] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, "NNLOPS simulation of Higgs boson production", *JHEP* **10** (2013) 222, doi:10.1007/JHEP10(2013)222, arXiv:1309.0017.

[194] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[195] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[196] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[197] Sherpa Collaboration, "Event Generation with Sherpa 2.2", *SciPost Phys.* **7** (2019) 034, doi:10.21468/SciPostPhys.7.3.034, arXiv:1905.09127.

[198] ATLAS Collaboration, "Observation of $t\bar{t}\gamma\gamma$ production at $\sqrt{s}$ =13 TeV with the ATLAS detector", *Phys. Lett. B* (2026) 140195, doi:10.1016/j.physletb.2026.140195, arXiv:2506.05018.

[199] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi, "Disentangling Higgs-Top Couplings in Associated Production", *JHEP* **04** (2014) 004, doi:10.1007/JHEP04(2014)004, arXiv:1312.5736.

[200] J. M. Campbell, R. K. Ellis, Y. Li, and C. Williams, "Predictions for diphoton production at the LHC through NNLO in QCD", *JHEP* **07** (2016) 148, doi:10.1007/JHEP07(2016)148, arXiv:1603.02663.

[201] S. Catani et al., "Diphoton production at the LHC: a QCD study up to NNLO", *JHEP* **04** (2018) 142, doi:10.1007/JHEP04(2018)142, arXiv:1802.02095.

[202] ATLAS Collaboration, "Measurement of the production cross section of pairs of isolated photons in pp collisions at 13 TeV with the ATLAS detector", *JHEP* **11** (2021) 169, doi:10.1007/JHEP11(2021)169, arXiv:2107.09330.

[203] M. Czakon, F. Eschment, T. Generet, and R. Poncelet, "Next-to-next-to-next-to-leading order QCD corrections to photon-pair production", 2026, arXiv:2604.12613.

[204] S. Frixione, "Isolated photons in perturbative QCD", *Phys. Lett. B* **429** (1998) 369, doi:10.1016/S0370-2693(98)00454-7, arXiv:hep-ph/9801442.

[205] J. Erdmann, F. Mausolf, and J. L. Späh, "KAN We Improve on HEP Classification Tasks? Kolmogorov–Arnold Networks Applied to an LHC Physics Example", *Comput. Softw. Big Sci.* **9** (2025) 9, doi:10.1007/s41781-025-00138-3, arXiv:2408.02743.

[206] M. Erdmann, J. Glombitza, G. Kasieczka, and U. Klemradt, "Deep Learning for Physics Research", World Scientific, 2021, doi:10.1142/12294.

[207] T. M. Mitchell, "Machine Learning", McGraw Hill, 1997, p. 2, ISBN 978-0-07-042807-2.

[208] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks* **2** (1989) 359, doi:10.1016/0893-6080(89)90020-8.

[209] A. N. Kolmogorov, "On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition", *Dokl. Akad. Nauk SSSR* **114** (1957) 953.

[210] Z. Liu et al., "KAN: Kolmogorov–arnold networks", 2024, arXiv:2404.19756.

[211] R. D. Ball et al., "Parton distributions with LHC data", *Nucl. Phys. B* **867** (2013) 244, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.

[212] DELPHES 3 Collaboration, "DELPHES 3, A modular framework for fast simulation of a generic collider experiment", *JHEP* **02** (2014) 057, doi:10.1007/JHEP02(2014)057, arXiv:1307.6346.

[213] A. Paszke et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library", in *Proc. NeurIPS 2019*, 2019, arXiv:1912.01703.

[214] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", in *Proc. ICLR 2015*, 2015, arXiv:1412.6980.

[215] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization", *Math. Program.* **45** (1989) 503, doi:10.1007/BF01589116.

[216] M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems", 2015, arXiv:1603.04467.

[217] TensorFlow Developers, "TensorFlow v2.17.0", 2024, doi:10.5281/zenodo.12726004.

[218] CMS Collaboration, "Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **81** (2021) 378, doi:10.1140/epjc/s10052-021-09014-x, arXiv:2011.03652.

[219] ATLAS Collaboration, "Search for the associated production of a Higgs boson and a single top quark in the $H \to \tau\tau$ decay mode and combined measurement of tH production using pp collisions at 13 TeV with the ATLAS detector", ATLAS-CONF-2026-002 (2026), https://cds.cern.ch/record/2957351.

[220] CMS Collaboration, "Search for the associated production of a Higgs boson with a single top quark in proton–proton collisions at $\sqrt{s} = 8$ TeV", *JHEP* **06** (2016) 177, doi:10.1007/JHEP06(2016)177, arXiv:1511.03610.

[221] G. Louppe, M. Kagan, and K. Cranmer, "Learning to Pivot with Adversarial Networks", *Adv. Neural Inf. Process. Syst.* **30** (2017) 981, arXiv:1611.01046.

[222] G. Kasieczka and D. Shih, "DisCo Fever: Robust Jet Classifiers through Distance Correlation", *Phys. Rev. Lett.* **125** (2020) 122001, doi:10.1103/PhysRevLett.125.122001, arXiv:2001.05310.

[223] S. Klein et al., "Decorrelation with conditional normalizing flows", *SciPost Phys.* **15** (2023) 124, doi:10.21468/SciPostPhys.15.3.124, arXiv:2211.02486.

[224] M. Algren, J. A. Raine, and T. Golling, "Decorrelation using optimal transport", *Mach. Learn. Sci. Tech.* **5** (2024) 035051, doi:10.1088/2632-2153/ad5d22, arXiv:2307.05187.

[225] J. Hermanns, "Study of decorrelation methods for the classification of $t\bar{t}H$ and tHq events in the $H \to \gamma\gamma$ decay channel at the CMS experiment", Bachelor's Thesis, RWTH Aachen University, 2025.

[226] CMS Collaboration, "Machine learning method for enforcing variable independence in background estimation with LHC data: ABCDisCoTEC", 2025, arXiv:2506.08826.

[227] L. Breiman, "Random forests", *Mach. Learn.* **45** (2001) 5, doi:10.1023/A:1010933404324.

[228] J. C. Collins and D. E. Soper, "Angular Distribution of Dileptons in High-Energy Hadron Collisions", *Phys. Rev. D* **16** (1977) 2219, doi:10.1103/PhysRevD.16.2219.

[229] J. Erdmann, N. K. Kasaraguppe, and F. Mausolf, "Learning to bin: differentiable and Bayesian optimization for multi-dimensional discriminants in high-energy physics", 2026, arXiv:2601.07756.

[230] H. J. Kushner, "A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise", *J. Basic Eng.* **86** (1964) 97, doi:10.1115/1.3653121.

[231] J. Močkus, "On Bayesian Methods for Seeking the Extremum", in *Proc. Optimization Techniques*, p. 400, 1974, doi:10.1007/3-540-07165-2_55.

[232] A. G. Zhilinskas, "Single-step Bayesian search method for an extremum of functions of a single variable", *Cybernetics* **11** (1975) 160, doi:10.1007/BF01069961.

[233] B. Pollack, S. Bhattacharya, and M. Schmitt, "Bayesian Block Histogramming for High Energy Physics", 2017, arXiv:1708.00810.

[234] K. T. Matchev and P. Shyamsundar, "ThickBrick: optimal event selection and categorization in high energy physics. Part 1: Signal discovery", *JHEP* **03** (2021) 291, doi:10.1007/JHEP03(2021)291, arXiv:1911.12299.

[235] P. De Castro and T. Dorigo, "INFERNO: Inference-Aware Neural Optimisation", *Comput. Phys. Commun.* **244** (2019) 170, doi:10.1016/j.cpc.2019.06.007, arXiv:1806.04743.

[236] S. Wunsch, S. Jörger, R. Wolf, and G. Quast, "Optimal Statistical Inference in the Presence of Systematic Uncertainties Using Neural Network Optimization Based on Binned Poisson Likelihoods with Nuisance Parameters", *Comput. Softw. Big Sci.* **5** (2021) 4, doi:10.1007/s41781-020-00049-5, arXiv:2003.07186.

[237] N. Simpson and L. Heinrich, "neos: End-to-End-Optimised Summary Statistics for High Energy Physics", *J. Phys. Conf. Ser.* **2438** (2023) 012105, doi:10.1088/1742-6596/2438/1/012105, arXiv:2203.05570.

[238] S. Diekmann, N. Eich, and M. Erdmann, "Binning high-dimensional classifier output for HEP analyses through a clustering algorithm", *EPJ Web Conf.* **295** (2024) 06005, doi:10.1051/epjconf/202429506005.

[239] F. Mausolf, "gato-hep: Gradient-based cATegorization Optimizer for High-Energy Physics", PyPI package, 2026, https://pypi.org/project/gato-hep/, accessed 2026-05-17.

[240] G. Hinton, N. Srivastava, and K. Swersky, "Lecture 6c: RMSProp – Neural Networks for Machine Learning", 2012, Online lecture, University of Toronto / Coursera, https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.

[241] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727.

[242] IceCube Collaboration, "Using End-to-End Optimized Summary Statistics to Improve IceCube's Diffuse Galactic Fits", *PoS* **ICRC** (2025) 1132, doi:10.22323/1.501.1132, arXiv:2507.06666.

[243] CMS Collaboration, "Search for the nonresonant and resonant production of a Higgs boson in association with an additional scalar boson in the $\gamma\gamma\tau\tau$ final state in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **05** (2026) 186, doi:10.1007/JHEP05(2026)186, arXiv:2506.23012.

[244] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, "Handling uncertainties in background shapes: the discrete profiling method", *JINST* **10** (2015) P04015, doi:10.1088/1748-0221/10/04/P04015, arXiv:1408.6865.

[245] CMS Collaboration, "flashggFinalFit", GitHub repository, 2026, https://github.com/cms-analysis/flashggFinalFit, accessed 2025-05-13.

[246] CMS Collaboration, "The CMS Statistical Analysis and Combination Tool: Combine", *Comput. Softw. Big Sci.* **8** (2024) 19, doi:10.1007/s41781-024-00121-4, arXiv:2404.06614.

[247] W. Verkerke and D. Kirkby, "The RooFit Toolkit for Data Modeling", in *Proc. CHEP03*, 2003, arXiv:physics/0306116.

[248] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[249] S. Manzoni et al., "Taming a leading theoretical uncertainty in HH measurements via accurate simulations for $b\bar{b}H$ production", *JHEP* **09** (2023) 179, doi:10.1007/JHEP09(2023)179, arXiv:2307.09992.

[250] CMS Collaboration, "A measurement of the Higgs boson mass in the diphoton decay channel", *Phys. Lett. B* **805** (2020) 135425, doi:10.1016/j.physletb.2020.135425, arXiv:2002.06398.

[251] CMS Collaboration, "Luminosity measurement in proton–proton collisions at 13.6 TeV in 2022 at CMS", CMS-PAS-LUM-22-001 (2024), https://cds.cern.ch/record/2890833.

[252] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses", *Philos. Trans. Roy. Soc. Lond. A* **231** (1933) 289, doi:10.1098/rsta.1933.0009.

[253] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses", *Ann. Math. Statist.* **9** (1938) 60, doi:10.1214/aoms/1177732360.

[254] A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[255] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[256] ATLAS Collaboration, CMS Collaboration, and LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", ATL-PHYS-PUB-2011-011, CMS NOTE-2011-005 (2011), https://cds.cern.ch/record/1379837.

[257] CMS Collaboration, "Measurement of the $t\bar{t}H$ and tH production rates in the $H \to b\bar{b}$ decay channel using proton–proton collision data at $\sqrt{s} = 13$ TeV", *JHEP* **02** (2025) 097, doi:10.1007/JHEP02(2025)097, arXiv:2407.10896.

[258] L. Demortier and L. Lyons, "Everything you always wanted to know about pulls", CDF/ANAL/PUBLIC/5776 (2002), http://physics.rockefeller.edu/luc/technical_reports/cdf5776_pulls.pdf.

[259] ATLAS Collaboration, "Probing the Higgs-top Yukawa interaction in the $t\bar{t}H$ and tH processes using $H \to \gamma\gamma$ with the ATLAS detector", 2026, arXiv:2606.04855.

Acknowledgements

My deepest gratitude goes to my primary supervisor, Prof. Johannes Erdmann. Thank you for your continuous support, guidance, and encouragement throughout my doctoral studies. Your expertise and insights have been invaluable in shaping this work, while leaving sufficient freedom for investigating those aspects of the research that particularly interested me. Your mentorship has shaped me far beyond the scientific content of this thesis, and I deeply appreciate the supportive and personal environment you have provided. I would also like to thank my secondary supervisor, Prof. Malgorzata Worek, for her valuable feedback and support in many discussions over the years, and for the particularly welcoming atmosphere in these situations. Moreover, I would like to thank you for agreeing to serve as a reviewer for this thesis.

I gratefully acknowledge the Studienstiftung des deutschen Volkes for supporting my doctoral studies through a doctoral scholarship. This support gave me the freedom and stability to pursue my doctorate with the necessary focus and independence.

My sincere thanks also go to all current and former members of Prof. Erdmann's working group for the inspiring and enjoyable working environment. In particular, I would like to thank Jan Lukas, Caio, and Nitish for the great collaboration in the many projects we worked on together over the years. I am equally grateful to all members of the group for the many enjoyable moments we shared, both in and outside of work. My particular thanks also go to all bachelor's and master's students I had the pleasure to supervise in their technical projects, from whose research I learned a lot.

I would like to thank the members of the CMS Collaboration for their excellent work and dedication, which provided the foundation for this research. In particular, I am thankful to the members of the CMS $H \to \gamma\gamma$ group for all contributions to the common tools and infrastructure and for the insightful feedback during the development of this analysis. This includes especially the members of this analysis team from the CEA Saclay and the NISER groups; thank you for the countless discussions on the analysis. Furthermore, I would like to thank Felipe Silva for the fruitful collaboration on the DT OMS development. Your help was crucial for the success of that project and it was always a great pleasure to work with you.

I also want to thank all members of the RTG "Physics of the heaviest particles at the LHC". The retreats, the PhenoChat, colloquia, and social events created many opportunities to discuss this work from different perspectives and to stay connected to a broader research community. I acknowledge the RTG for financial support during the final phase of my doctorate. Furthermore, I would like to thank the members of the III. Physikalisches Institut A at RWTH Aachen University for their support throughout my doctoral studies, in particular for the high-performance computing infrastructure that was essential for this work and for the administrative assistance.

Finally, my deepest thanks go to my girlfriend Natalie, my family, and the friends who accompanied me beyond the academic part of this journey. Thank you for your endless support, patience, encouragement, and understanding throughout these years. You helped me keep perspective during difficult phases and made the time of this doctorate much richer beyond work. Without you, this journey would not have been possible.