CMS Physics Analysis Summary

Differential cross section measurement of t$\bar{\text{t}}$H production in proton-proton collisions at $\sqrt{s} = 13$ TeV in CMS

The CMS Collaboration

Abstract

The production of a Higgs (H) boson in association with two top quarks (t$\bar{\text{t}}$H) in final states containing multiple electrons, muons, or hadronically decaying tau leptons is measured using proton-proton collisions recorded at a center-of-mass energy of 13 TeV with the CMS detector. The analyzed data correspond to an integrated luminosity of 138 fb^{-1}. The analysis aims at events that contain H $\rightarrow$ WW or H $\rightarrow \tau\tau$ decays and the top quarks decay into final states with leptons or hadrons. The signal sensitivity is maximized by partitioning the selected events depending on the lepton multiplicity into three exclusive event categories: 2ℓ "same sign" + 0 hadronic tau leptons, 2ℓ "same sign" + 1 hadronic tau lepton, and 3ℓ "same sign" + 0 hadronic tau leptons, where ℓ denotes charged light leptons (e, μ). Differential production rates are measured as a function of the H boson transverse momentum and of the mass of the t$\bar{\text{t}}$H system and found to be compatible with predictions from the standard model of particle physics. This result is the first differential measurement of t$\bar{\text{t}}$H production to date by the CMS Collaboration.

1 Introduction

The discovery of the Higgs (H) boson by the ATLAS and CMS Collaborations in 2012 [1, 2] marked a significant milestone in particle physics. The properties of this boson are found, so far, to be compatible with the expectations from the standard model (SM) and have been studied in a series of precision measurements [3]. Direct results from the LHC Run 2 (corresponding to data collected between 2015 and 2018) indicate that couplings of the H boson, particularly to the W and Z bosons, align with SM predictions [4]. The Yukawa couplings (y_f) of the H boson to fermions are proportional to the fermion mass (m_f) and expressed as $y_f = \sqrt{2}m_f/v$, where $v \approx 246$ GeV is the vacuum expectation value of the H field.

The top quark, with a mass $m_t = 173.34 \pm 0.76$ GeV [5], is the heaviest known fermion: its Yukawa coupling y_t is expected to be of the order of one. The determination of y_t is particularly important for understanding the mechanism of electroweak symmetry breaking [6–8], as deviations from the SM prediction would indicate the presence of new physics beyond the SM.

The CMS Collaboration measured the ratio $\kappa_t = y_t/y_t^{SM}$ to be $-0.9 < \kappa_t < -0.7$ or $0.7 < \kappa_t < 1.1$, at 95% confidence level (CL) [9], by investigating the associated production of the H boson with a top quark pair and subsequent decay to leptonic final states. The charge-parity (*CP*) structure of the Yukawa interaction between the H boson and the top quark has also been determined by the CMS Collaboration in a recent publication reporting two-dimensional confidence regions for the *CP*-odd and *CP*-even coupling modifiers, as well as constraining the presence of a fractional *CP*-odd contribution [10, 11].

The analysis described in this note targets the production of a H boson in association with two top quarks ($t\bar{t}H$) in a multilepton final state: we follow closely the strategy adopted in Ref. [9], focussing on its subset of most sensitive final states to explore the differential behaviour of the $t\bar{t}H$ system. The differential cross section of $t\bar{t}H$ production is measured as a function of the transverse momentum p_T of the H boson p_T, and the visible mass $m_{t\bar{t}H}$ of the $t\bar{t}H$ system. Leading order (LO) Feynman diagrams for $t\bar{t}H$ and tH production are shown in Figs. 1 and 2.

Figure 1: Feynman diagrams at LO for $t\bar{t}H$ production.

The measurements are based on data recorded by the CMS experiment in proton-proton (pp) collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV collected during the LHC Run 2, corresponding to an integrated luminosity of 138 fb^{-1}. Final states with electrons, muons, and hadronically decaying tau leptons (τ_h) are considered. Three exclusive final state signatures are considered: $2\ell ss + 0\tau_h$, $2\ell ss + 1\tau_h$, and $3\ell + 0\tau_h$, where ℓ denotes charged light leptons (e, μ), and "ss" indicates indicate lepton pairs with the same electric charge ("same-sign").

Such categories correspond to those that provided the best precision for the inclusive $t\bar{t}H$ production cross section measurement in Ref. [9]. Machine learning techniques, such as Artificial
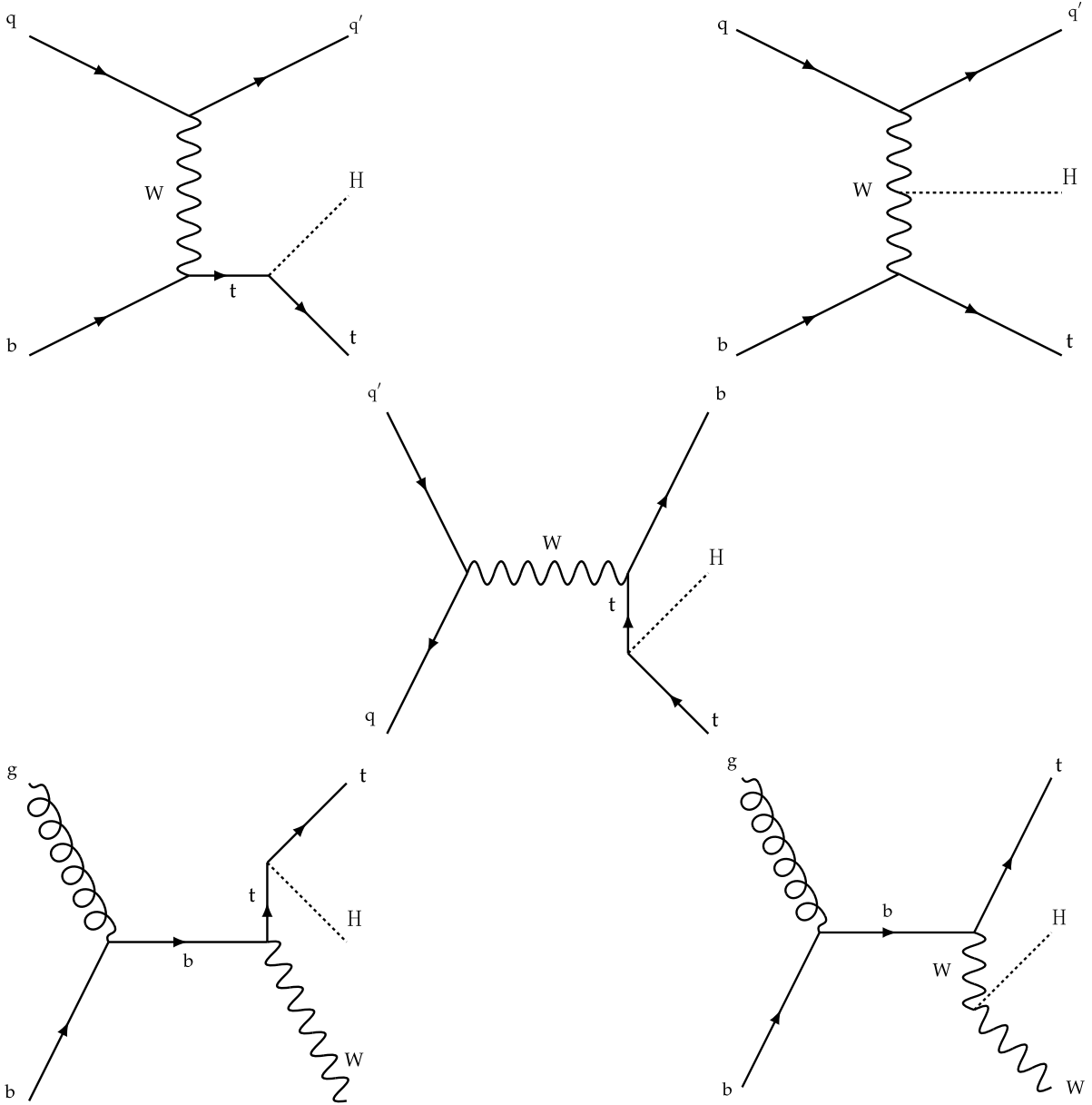
Figure 2: Feynman diagrams at LO for tH production via the *t*-channel (tHq in upper left and upper right) and *s*-channel (middle) processes, and for associated production of a H boson with a single top quark and a W boson (tHW in lower left and lower right). The tHq and tHW production processes are shown for the five-flavour scheme described in Section 3.

Neural Networks (ANN), are employed for improved separation between the $t\bar{t}H$ signal and the backgrounds. All differential observables are reconstructed from final-state objects (visible decay products and transverse energy): the $t\bar{t}H$ mass is reconstructed as the sum of the four-momenta of final-state objects, whereas the H boson p_T is regressed via ANNs that use several features of the final state objects as inputs. This note is structured as follows: Section 2 provides a brief introduction to the CMS detector, Section 3 describes the data sets and simulated event samples used, Section 4 explains the event selection strategy, and Section 5 details the background estimation. In Sections 6 and 7 the reconstruction of the observables used and the list of the systematic uncertainties is presented. Finally, in Section 8, the results of the differential cross section measurement are shown.

2   The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. A silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections, are positioned within the solenoid volume. The silicon tracker measures charged particles within the pseudorapidity range $|\eta| < 2.5$. The ECAL is a fine-grained hermetic calorimeter with quasi-projective geometry, and is segmented into the barrel region of $|\eta| < 1.48$ and in two endcaps that extend up to $|\eta| < 3.0$. The hadron calorimeter barrel and endcaps similarly cover the region $|\eta| < 3.0$. Forward calorimeters extend the coverage up to $|\eta| < 5.0$. Muons are measured and identified in the range $|\eta| < 2.4$ by gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A two-level trigger system [12] is used to reduce the rate of recorded events to a level suitable for data acquisition and storage. The first level of the CMS trigger system, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events with a fixed latency of $4\,\mu$s. The high-level trigger processor farm further decreases the event rate from around 100 kHz to about 1 kHz. Details of the CMS detector and its performance, together with a definition of the coordinate system and the kinematic variables used in the analysis, are reported in Ref. [13].

3   Data sets and simulations

This analysis uses pp collision data recorded at $\sqrt{s} = 13$ TeV at the CERN LHC during 2016–2018. Data-taking periods during which the CMS detector was not fully operational are excluded from the analysis. The total integrated luminosity of the analyzed data set amounts to $138\,\text{fb}^{-1}$, out of which 36.3 [14], 41.5 [15], and $59.8\,\text{fb}^{-1}$ [16] have been recorded in 2016, 2017, and 2018, respectively.

The backgrounds arising from nonprompt or misidentified leptons (collectively denoted as *nonprompt*) and from lepton charge mismeasurement are estimated from data, regardless of the processes that originated them, using the methods described in Section 5. All the other contributions from background processes are determined from simulation.

The signal samples for $t\bar{t}H$, tHW (tHq) are generated at next-to-leading-order (NLO) accuracy in perturbative quantum chromodynamics (QCD) via POWHEG v2.0 [17–19]. The POWHEG generator is also used to simulate the the majority of the background processes. These include $t\bar{t}$, single top quark, and tW production, diboson production ($W^{\pm}W^{\mp}$, WZ, and ZZ), SM H boson production via gluon (ggH) and vector boson (qqH) fusion , and the production of SM H bosons in association with W and Z bosons (WH, ZH) and with top quark pairs ($t\bar{t}WH$, $t\bar{t}ZH$).

The four-flavour (4 FS) and five-flavour (5 FS) schemes [20, 21] are used to simulate the tHq and tHW processes, respectively. In the 5 FS, bottom quarks are considered as sea quarks of the proton and may appear in the initial state of pp scattering processes, as opposed to the 4 FS, where only up, down, strange, and charm quarks are considered as valence or sea quarks of the proton, while bottom quarks are produced by gluon splitting at the matrix element level and therefore only appear in the final state [20]. The event samples for tHW (tHq) samples of events are generated using NLO 5 FS calculations in both cases. The contribution from s-channel tH production is negligible and is not considered in this analysis.

Background arising from Drell–Yan (DY), $W\gamma$, $Z\gamma$, and $t\bar{t}$ production in association with W and

Z bosons (t$\bar{\text{t}}$W, t$\bar{\text{t}}$Z), from triboson (WWW, WWZ, WZZ, ZZZ, and WZγ) production, as well as from the production of four top quarks (t$\bar{\text{t}}$t$\bar{\text{t}}$) are generated with MADGRAPH5_aMC@NLO at NLO accurancy in QCD. The modeling of the t$\bar{\text{t}}$W process includes additional electroweak corrections simulated using MADGRAPH5_aMC@NLO proportional to $\alpha_S\alpha_{EW}^3$ [22–24], where α_S (α_{EW}) is the strong (electroweak) coupling constant. The t$\bar{\text{t}}\gamma$, t$\bar{\text{t}}\gamma^*$, tZ, t$\bar{\text{t}}$WW, and W+jets backgrounds are simulated at LO using the program MADGRAPH5_aMC@NLO v2.6.5 [25–28]. For simulated samples generated via MADGRAPH5_aMC@NLO at NLO (LO) in QCD, the matching between matrix element and parton shower is performed with the FxFx [28] (MLM [26]) scheme. The top quark mass m_{t} assumed in all samples is 172.5 GeV.

All simulations use the NNPDF3.1 [29] sets of parton distribution functions. Parton showering, hadronization, the underlying event description, and the decays of tau leptons, including polarization effects, are modelled using PYTHIA [30] v8.240 with the CP5 tune [31].

The presence of simultaneous pp collisions in the same or nearby bunch crossings, referred to as pileup (PU), is modelled by superimposing inelastic pp interactions, simulated using PYTHIA, to all Monte Carlo (MC) events. Simulated events are weighed so that the PU distribution of simulated samples matches the one observed in the data. All MC events are passed through a detailed simulation of the CMS apparatus, based on GEANT4 [32, 33], and are processed using the same version of the CMS event reconstruction software used for the data.

Simulated events are corrected to account for residual differences between data and simulation arising in: trigger efficiencies; reconstruction and identification efficiencies for electrons, muons, and τ_{h}; the energy scale of τ_{h} and jets; the efficiency in identifying jets originating from the hadronization of bottom quarks and the corresponding misidentification rates for light-quark and gluon jets; and the resolution in missing transverse momentum ($p_{\text{T}}^{\text{miss}}$). Object corrections are detailed in Section 4. The respective corrections are typically at the level of a few percent [34–38] and are measured using a variety of SM processes, such as DY ($Z/\gamma^* \to$ ee, $Z/\gamma^* \to \mu\mu$, and $Z/\gamma^* \to \tau\tau$), t$\bar{\text{t}}$, and γ+jets production.

4 Event reconstruction and selection

The event reconstruction and selection aims at selecting t$\bar{\text{t}}$H events where the H boson decays into a pair of W bosons or tau leptons. The W bosons can subsequently decay either hadronically or into electrons or muons, while the tau can decay to electrons, muons, or hadrons. Decays of the H boson to Z bosons are not explicitly explored with dedicated event categories but are included in the analysis. Therefore, the analysis targets events with two leptons with the same electric charge, with or without the addition of one hadronically decaying tau lepton, or three leptons with at least one pair of leptons with opposite charge.

4.1 Object reconstruction

The CMS particle-flow (PF) algorithm [39] provides a comprehensive event description by combining information from all subdetectors to reconstruct and identify individual particles, which are classified into five categories: electrons, muons, photons, charged hadrons, and neutral hadrons.

Electrons are reconstructed using information from the tracker and ECAL [40] and must satisfy $p_{\text{T}} > 7$ GeV and $|\eta| < 2.5$. Their identification relies on a multivariate (MVA) algorithm that combines observables such as energy-momentum consistency, cluster compactness, and bremsstrahlung effects. Additional selection criteria reject electron candidates from photon conversions and ensure charge consistency in specific channels [40].

Muons are reconstructed by linking tracks in the silicon tracker to hits in the muon detectors embedded in the steel flux-return yoke [41]. To ensure high-quality identification, muons must satisfy $p_T > 5\,\text{GeV}$ and $|\eta| < 2.4$, with stringent requirements on the spatial matching of tracks and the relative uncertainty in curvature measurements for certain channels.

Electrons and muons passing these criteria are classified as "loose leptons". To distinguish prompt leptons (from W/Z bosons or leptonic τ decays) from nonprompt leptons (from b hadron decays), MVA discriminants are employed. These take as input charged and neutral particles within a cone around the lepton direction, as well as observables related to the lepton and reconstructed jets. Key features include the ratio of the lepton p_T to the jet p_T and the perpendicular momentum component relative to the jet direction, alongside traditional observables such as relative isolation [42, 43] and impact parameters relative to the primary vertex. Leptons passing the MVA discriminants are referred to as "tight leptons".

The leptonic decay products of tau leptons are selected as described previously, while hadronic decays are reconstructed and identified using the "hadrons-plus-strips" algorithm [37], which reconstructs hadronic tau lepton decays such as $\tau^- \to h^- \nu_\tau, \tau^- \to h^- \pi^0 \nu_\tau, \tau^- \to h^- \pi^0 \pi^0 \nu_\tau$, and similar modes, including their charge conjugates. Photons from neutral pion decays that convert in the silicon tracker are reconstructed using rectangular strips in the ECAL, which account for energy deposit broadening caused by magnetic bending and bremsstrahlung. The identification of τ_h candidates employs the "DeepTau" algorithm [44], a convolutional ANN [45] that uses both high- and low-level observables. High-level observables include the p_T, η, ϕ, and mass of the τ_h candidate, its decay mode, isolation characteristics, and its production-decay displacement. Low-level information consists of particle activity in $\eta \times \phi$ grids centered on the τ_h direction, with an inner grid (0.2×0.2, cells of 0.02×0.02) and an overlapping outer grid (0.5×0.5, cells of 0.05×0.05). The τ_h candidates must have $p_T > 20\,\text{GeV}$ and $|\eta| < 2.3$, and pass ANN-based selections tailored to different analysis channels. The selection corresponding to the very-loose working point (WP) of the tau identification criteria is applied. The very-loose working point (WP) of the tau identification criteria is applied. This working point achieves a τ_h vs. jets identification efficiency of around 90% and a τ_h vs. lepton identification efficiency exceeding 95%.

Jets are reconstructed using the anti-k_T algorithm [46] with a distance parameter of 0.4 and PF-reconstructed particles as inputs. Corrections for PU effects [47] and calibrations based on p_T and η [35] are applied. Jets must satisfy $p_T > 25\,\text{GeV}$ and $|\eta| < 5.0$, pass noise rejection criteria [48], and avoid overlap with leptons or τ_h within $\Delta R < 0.4$. For $2.7 < |\eta| < 3.0$, a tighter $p_T > 60\,\text{GeV}$ threshold is used to mitigate calorimeter noise. Jets are categorized as central ($|\eta| < 2.4$) or forward ($2.4 < |\eta| < 5.0$), with high-p_T forward jets serving as a signature for tH production.

Jets within $|\eta| < 2.4$ from b quarks are identified as b jets using the DEEPJET algorithm [49], which employs ANN-based observables related to the long lifetime, particle multiplicity, and mass of b jets, including properties of secondary vertices and jet constituents. Selection efficiencies of b jets are 84% (loose WP) and 70% (tight WP), with respective mistag rates of 11% and 1.1% for light-quark/gluon jets, and 50% and 15% for jets originating from a c quark.

The missing transverse momentum vector, $\vec{p}_T^{\text{miss}}$, is computed as the negative vector p_T sum of all PF-reconstructed particles. Its magnitude, p_T^{miss}, is combined with the magnitude of the p_T sum of electrons, muons, τ_h, and jets, denoted as H_T^{miss}, to define the linear discriminant $L_D := 0.6 p_T^{\text{miss}} + 0.4 H_T^{\text{miss}}$ [50]. This discriminant reduces PU effects and background contributions by leveraging the higher resolution of p_T^{miss} and the robustness of H_T^{miss}.

4.2 Event selection

The experimental signature consists of multiple electrons, muons, and τ_{h}, $p_{\mathrm{T}}^{\mathrm{miss}}$ from neutrinos, one (for tH) or two (for $\mathrm{t\bar{t}H}$) b jets, and additional light-quark jets. The analysis defines three non-overlapping channels based on the event signature: $2\ell\mathrm{ss} + 0\tau_{\mathrm{h}}$, $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$, and $3\ell + 1\tau_{\mathrm{h}}$.

Events are selected based on single-, double-, and triple-lepton triggers and lepton+τ_{h} triggers. Offline selections ensure that reconstructed objects match the trigger-level requirements. Trigger thresholds vary across data-taking periods: e.g., single-electron (single-muon) triggers vary in the range 25–35 (22–27) GeV.

The charge of the leptons and τ_{h}'s is required to match the expected signature of the $\mathrm{t\bar{t}H}$ and tH signal processes. In the $3\ell + 0\tau_{\mathrm{h}}$ and $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$ channels, the charge sum of the leptons and τ_{h} is required to be either $+1$ or -1. For reference, the charge of the τ lepton is determined by summing the charges of its decay products [37].

The $2\ell\mathrm{ss} + 0\tau_{\mathrm{h}}$ channel targets events where one lepton originates from the H boson decay and the other from a top quark decay. Requiring same-sign (ss) leptons reduces the signal yield by about half but significantly improves the signal-to-background ratio by removing a large background contribution from $\mathrm{t\bar{t}}$+jets production with dileptonic decays of the top quarks. The better signal-to-background ratio with ssleptons compared to opposite-sign (OS) lepton pairs motivates the separate analysis of events with two leptons and one τ_{h} in the channel $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$.

The selection criteria for b jets are designed to maximize the efficiency for the $\mathrm{t\bar{t}H}$ and tH signal. Two b jets are required: one b jet may fail the p_{T} or η acceptance or b-tagging criteria, provided the other b jet passes the tight b-tagging requirements. This approach is motivated by the fact that major background processes, such as single top and top pair production with W, Z, photons, and jets, exhibit a similar b jet multiplicity to the $\mathrm{t\bar{t}H}$ and tH signals.

Background contributions from $\mathrm{t\bar{t}Z}$, tZ, WZ, and DY production are reduced by vetoing events that contain same-flavor opposite-signed (SFOS) lepton pairs, where the leptons pass loose selection criteria and have an invariant mass, $m_{\ell\ell}$, within 10 GeV of the Z boson mass ($m_{\mathrm{Z}} = 91.19$ GeV) [51]. This selection criterion is referred to as the "Z boson veto". In the $2\ell\mathrm{ss} + 0\tau_{\mathrm{h}}$ and $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$ channels, the veto is also applied to ss electron pairs, due to the significantly higher probability of mismeasuring the charge of electrons compared to muons.

Further suppression of background contributions from DY production is achieved by imposing a requirement on the linear discriminant, $L_{\mathrm{D}} > 30$ GeV, which is adjusted based on specific conditions to either enhance the selection efficiency for $\mathrm{t\bar{t}H}$ and tH signal events or to reject more background events.

In the $2\ell\mathrm{ss} + 0\tau_{\mathrm{h}}$ and $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$ channels, the L_{D} requirement is applied only to events where both reconstructed leptons are electrons. This condition suppresses background contributions from DY processes that stem from mismeasuring the electron charge. In the $3\ell + 0\tau_{\mathrm{h}}$ channel, the distribution of the number of jets for the DY background decreases sharply as the number of jets increases. Therefore, events with a large number of jets contribute very little to the DY background. As a result, the L_{D} requirement is applied only to events with three or fewer jets.

Events with leptons passing loose selection criteria and an invariant mass below 12 GeV are vetoed to remove low-mass DY production, quarkonium decays, and cascade decays of heavy-flavor hadrons that are poorly modeled in the simulation.

Additionally, in the $3\ell + 0\tau_{\mathrm{h}}$, events containing four leptons with an invariant mass below

140 GeV are vetoed to avoid overlap with the analysis of signals decaying via H $\rightarrow$ ZZ $\rightarrow$ $\ell^+\ell^-\ell^+\ell^-$, as described in Ref. [52].

Events selected in any of the categories are required to pass a filter algorithm that removes events that are subject to different types of spurious detector signals. These algorithms are referred to as "p_T^{miss} filters" and perform additional event cleaning according to beamhalo effects, detector noise, etc. [53].

Table 1 provides a detailed description of the $2\ell ss + 0\tau_h$, $2\ell ss + 1\tau_h$, and $3\ell + 1\tau_h$ category selections.

Table 1: Event selections applied in the $2\ell ss + 0\tau_h$, $2\ell ss + 1\tau_h$, $3\ell + 0\tau_h$, and $3\ell + 1\tau_h$ channels. The p_T thresholds applied to the lepton of highest, second-highest, and third-highest p_T are separated by slashes. The symbol "—" indicates that no requirement is applied.

Selection step	$2\ell ss + 0\tau_h$	$2\ell ss + 1\tau_h$	$3\ell + 0\tau_h$
Targeted $t\bar{t}H$ decay	$t \rightarrow b\ell\nu, t \rightarrow bqq'$ with $H \rightarrow WW \rightarrow \ell\nu qq'$	$t \rightarrow b\ell\nu, t \rightarrow bqq'$ with $H \rightarrow \tau\tau \rightarrow \ell\nu\nu\tau_h\nu$	$t \rightarrow b\ell\nu, t \rightarrow bqq'$ with $H \rightarrow WW \rightarrow \ell\nu qq'$
Targeted tH decays	$t \rightarrow b\ell\nu,$ $H \rightarrow WW \rightarrow \ell\nu qq'$	$t \rightarrow b\ell\nu,$ $H \rightarrow \tau\tau \rightarrow \ell\tau_h + \nu's$	$t \rightarrow b\ell\nu,$ $H \rightarrow \tau\tau \rightarrow \ell\tau_h + \nu's$
Trigger	Single- and double-lepton triggers		
Lepton η	$|\eta| < 2.5$ (e) or 2.4 (μ)		
τ_h p_T	—	$p_T > 20$ GeV	—
τ_h η	—	$|\eta| < 2.3$	—
τ_h identification	—	very-loose	—
Charge requirements	2 ss leptons	2 ss leptons	$\sum_{\ell,\tau_h} q = 0$
	and charge quality requirements	and charge quality requirements $\sum_{\ell,\tau_h} q = \pm 1$	
Multiplicity of central jets	$\geq$3 jets	$\geq$3 jets	$\geq$3 jets
b-tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets		
Missing transverse momentum	$L_D > 30$ GeV†		
Dilepton invariant mass	$|m_{\ell\ell} - m_Z| > 10$ GeV‡ and $m_{\ell\ell} > 12$ GeV		

4.3 Event classification

To enhance the separation between signal and background and consequently improve the sensitivity of the analysis, a Deep Neural Network (DNN) classifier is trained for each signal region (SR).

The observables used as input for the DNN are selected to maximize the discriminating power and sensitivity of the analysis. These optimizations are performed separately for each of the three analysis channels. The key observables include the transverse momentum of the reconstructed leptons, τ_h, and jets. These are categorized based on the number of electrons, muons, forward jets, central jets, and jets that satisfy either loose or tight b tagging criteria. Additional observables include the three-momentum of leptons, τ_h, and jets, L_D, and angular separations between leptons, τ_h, and jets. The average ΔR between jet pairs, the sum of charges for various combinations of leptons and τ_h, and specific observables related to top quark and H boson decay modes are also included. The input variables include observables related to the reconstruction of specific top quark and H boson decay modes. These comprise the transverse mass of a lepton, defined as $m_T = \sqrt{2p_T^\ell p_T^{miss}(1 - \cos\Delta\phi)}$, where $\Delta\phi$ represents the angle in the transverse plane between the lepton momentum and the $\vec{p}_T^{miss}$ vector; the invariant masses of various combinations of leptons and τ_h; and the invariant mass of the jet pair with the highest and second-highest b-tagging discriminants. These observables are further complemented

by the outputs of MVA algorithms, described in Ref. [50], that reconstruct hadronic top quark decays and identify jets originating from $H \to WW \to \ell^+ \nu_\ell q\overline{q}'$ decays.

The classification involves three main classes: $t\overline{t}H$-like events, tH-like events, and background-like events. An additional category for $t\overline{t}W$-like events is used in the $2\ell ss + 0\tau_h$ category. However, the classification of $t\overline{t}H$ events can be improved by exploiting the correlation between visible final state objects and the H boson p_T. To this extent, the $t\overline{t}H$-like category is partitioned into two subcategories: one with $t\overline{t}H$-like events having a generator-level H boson p_T below 300 GeV and another with $t\overline{t}H$-like events having a generator-level H boson p_T above 300 GeV. This has found to provide an improved separation power with respect to the other background categories high H boson p_T phase-space, that is crucial for the transverse momentum regression described in Section 6.1 and for the differential measurement that is performed. When the DNN classifier is evaluated on data and simulated events, its output scores can be interpreted as the probability of the event belonging to one of the event classes. Events are therefore divided into exclusive categories and assigned to the category that corresponds to the node that gives the highest value for the event.

5 Background estimation

The SRs defined above receive contributions from the $t\overline{t}H$ signal as well as from several background processes, which are estimated using different methods following the methodology employed in previous publications [9, 54, 55], where these methods are described in more detail.

The contributions of background processes are categorized into irreducible and reducible components. A background is classified as irreducible if all reconstructed electrons and muons are genuine prompt leptons, and all reconstructed τ_h candidates originate from genuine hadronic tau lepton decays. In the $2\ell ss + 0\tau_h$ and $2\ell ss + 1\tau_h$ channels, an additional requirement is imposed that the measured charge of reconstructed electrons and muons matches their true charge. Irreducible background contributions are modelled using simulated events that satisfy these criteria to prevent double counting.

All other background contributions are treated as reducible and are primarily determined from data. Three sources of reducible background contributions are identified: misidentified leptons and τ_h (referred to as "misidentified leptons"), asymmetric photon conversions into electrons (referred to as "conversions"), and lepton charge mismeasurement.

The misidentified lepton and τ_h background arises from events where at least one reconstructed electron or muon is attributed to the misidentification of a nonprompt lepton or hadron, or where at least one reconstructed τ_h results from the misidentification of a quark or gluon jet. The dominant contribution to this background originates from $t\overline{t}$+jets production, due to its substantial cross section.

The conversion background consists of events in which one or more reconstructed electrons arise from the conversion of a photon. This background is typically dominated by $t\overline{t}\gamma$ events, where one electron or positron from the photon conversion retains most of the energy of the converted photon, while the other electron or positron is low in energy and fails to be reconstructed. Such cases are referred to as asymmetric conversions.

The charge mismeasurement background is specific to the $2\ell ss + 0\tau_h$ and $2\ell ss + 1\tau_h$ channels and involves events where the charge of a reconstructed lepton is mismeasured. The primary source of this background is $t\overline{t}$+jets production, where both top quarks decay semi-leptonically.

In the $2\ell\mathrm{ss} + 1\tau_\mathrm{h}$ channel, this process additionally involves the misidentification of a quark or gluon jet as a τ_h. Electron charge mismeasurement typically arises from the emission of a hard bremsstrahlung photon, followed by an asymmetric conversion of the photon. The reconstructed electron or positron is typically the particle that retains most of the energy of the converted photon, leading to an equal probability of the reconstructed electron having the same or opposite charge compared to the particle that emitted the bremsstrahlung photon [40]. The probability of mismeasuring the charge of muons is negligible in this analysis.

To ensure mutual exclusivity among the three types of reducible backgrounds, events are prioritized as follows: the misidentified lepton background is given precedence over the charge mismeasurement and conversion backgrounds, and the charge mismeasurement background is given precedence over the conversions background when an event qualifies for more than one type. The misidentified leptons and charge mismeasurement backgrounds are estimated from data, while the conversions background is modeled in simulation. The methods for estimating the misidentified leptons and charge mismeasurement backgrounds are detailed in Sections 5.1 and 5.2, respectively.

To prevent double-counting between data- and simulation-driven background estimates, the reconstructed electrons, muons, and τ_h in simulated events are matched to their generator-level counterparts and simulated signal and background events that qualify as misidentified leptons or charge mismeasurement backgrounds in the SR are vetoed.

Regarding irreducible backgrounds, contributions other than those from $\mathrm{t\bar{t}W(W)}$, $\mathrm{t\bar{t}Z}$, $\mathrm{t\bar{t}}$+jets, DY, and diboson backgrounds, as well as SM H boson production via ggH, qqH, WH, ZH, $\mathrm{t\bar{t}WH}$, and $\mathrm{t\bar{t}ZH}$, are collectively referred to as "rare" backgrounds. These rare backgrounds contribute minimally to the three analysis channels and include processes such as tW and tZ production, same-sign W boson pair production, triboson production, and $\mathrm{t\bar{t}t\bar{t}}$ production. The contribution of tZq is estimated from simulation, while the the normalizations of the other irreducible backgrounds are determined in-situ from control regions (CRs) in data, as described in Section 8.

5.1 Estimation of the "misidentified leptons" background

The contribution from nonprompt leptons is estimated following the misidentification probability (MP) method [55]. This estimate is performed by selecting events that pass all the requirements that define the SRs, but at least one of the leptons fails some of the selection requirements specified in Section 4. This region, labeled as application region, is enriched in events with nonprompt leptons, which tend to fail the lepton selection criteria. The nonprompt contribution to our SRs is estimated by weighting events in the application region by a weight w, defined as

$$w = (-1)^{n+1} \prod_{i=1}^{n} \frac{f(p_\mathrm{T}, \eta)_i}{1 - f(p_\mathrm{T}, \eta)_i}, \tag{1}$$

where n is the number of leptons failing the selection and $f(p_\mathrm{T}, \eta)_i$ is the MP, defined as the probability of a nonprompt lepton to pass the SR selection criteria. The MP is measured in a region enriched in multijet events containing a nonprompt lepton, that are collected by low-p_T lepton triggers.

5.2 Estimation of the "charge mismeasurement" background

The charge mismeasurement background, relevant for events containing one or two reconstructed electrons in the $2\ell\mathrm{ss} + 0\tau_\mathrm{h}$ and $2\ell\mathrm{ss} + 1\tau_\mathrm{h}$ channels, is estimated using a method analogous to the MP method. Events are selected that satisfy all SR selection criteria, except that

both leptons are required to be of opposite sign rather than same sign, and are assigned appropriately defined weights. In the $2\ell\mathrm{ss} + 0\tau_\mathrm{h}$ channel, the weight corresponds to the sum of the probabilities for the charge of either lepton to be mismeasured. In the $2\ell\mathrm{ss} + 1\tau_\mathrm{h}$ channel, only the lepton sharing the same charge as the τ_h is considered, since only events where the charge of this lepton is mismeasured satisfy the condition $\sum_{\ell,\tau_\mathrm{h}} q = \pm 1$ applied in the SR of this channel.

The probability for mismeasurement of the electron charge, referred to as the electron charge misidentification rate, is determined using $Z/\gamma^* \to$ ee events. These events are selected by requiring an electron pair with an invariant mass m_ee in the range $60 < m_\mathrm{ee} < 120\,\mathrm{GeV}$, with no requirement on the charge of the pair. Contributions to the selected sample from processes other than DY production of electron pairs are determined via a maximum likelihood fit to the m_ee distribution. Denoting the number of $Z/\gamma^* \to$ ee events with reconstructed ss and OS electron pairs as N_ss and N_OS, respectively, the electron charge misidentification rate is calculated as the ratio $N_\mathrm{ss}/(N_\mathrm{OS} + N_\mathrm{ss})$. This ratio is measured as a function of the electron p_T and η, ranging from 5.1×10^{-5} for electrons with low p_T in the ECAL barrel to 1.6×10^{-3} for electrons with high p_T in the ECAL endcap.

5.3 Control regions for irreducible backgrounds

The accuracy of the simulation-based modeling of the primary irreducible backgrounds, originating from $\mathrm{t\bar{t}W(W)}$, $\mathrm{t\bar{t}Z}$, WZ, and ZZ production, is validated using three CRs. The first one, referred to as the 3ℓ-CR, is based on the SR of the $3\ell + 0\tau_\mathrm{h}$ channel and targets the $\mathrm{t\bar{t}Z}$ and WZ backgrounds. The selection criteria for the 3ℓ-CR differ from the SR by removing the Z boson veto, requiring at least one SFOS lepton pair with an invariant mass $m_{\ell\ell}$ satisfying $|m_{\ell\ell} - m_\mathrm{Z}| < 10\,\mathrm{GeV}$, relaxing the jet multiplicity requirement to at least one jet, and removing the requirement on b-tagged jets. Events are categorized by lepton flavour (eee, eeμ, e$\mu\mu$, $\mu\mu\mu$), jet multiplicity, and b-tagged jet multiplicity.

The second CR, termed the 4ℓ-CR, is based on the SR of the $4\ell + 0\tau_\mathrm{h}$ channel and validates the ZZ background. Modifications include removing the Z boson veto, requiring at least one SFOS lepton pair with $|m_{\ell\ell} - m_\mathrm{Z}| < 10\,\mathrm{GeV}$, and relaxing the requirements on jet and b-tagged jet multiplicities. Events are binned by the number of SFOS lepton pairs with $|m_{\ell\ell} - m_\mathrm{Z}| < 10\,\mathrm{GeV}$ and by the number of b-tagged jets to distinguish ZZ from $\mathrm{t\bar{t}Z}$ contributions.

The third CR targets the $\mathrm{t\bar{t}W(W)}$ background and is identical to the SR of the $2\ell\mathrm{ss} + 0\tau_\mathrm{h}$ channel, except that the DNN output node with the highest activation must correspond to the $\mathrm{t\bar{t}W}$ background. Contributions from misidentified leptons and charge mismeasurement are estimated using methods described in Sections 5.1 and 5.2.

Uncertainties in the determination of the backgrounds normalizations in the CRs include both statistical and systematic sources, combined in quadrature. Systematic uncertainties are consistent with those applied in the SR, which are detailed in Section 7.

6 Differential cross section extraction

The signal cross section is extracted from data via a binned maximum likelihood fit [56] to the distributions of selected observables in each SR defined for the $2\ell\mathrm{ss} + 0\tau_\mathrm{h}$, $2\ell\mathrm{ss} + 1\tau_\mathrm{h}$, and $3\ell + 0\tau_\mathrm{h}$ categories as well as in the so-called 3ℓ and 4ℓ CRs described in Section 5.3. Each SR is further separated into subcategories based on the different output nodes provided by the DNN multiclassifiers described in Section 4.3.

The 3ℓ and 4ℓ CRs are used to estimate the ZZ and WZ contributions from data, while the $t\bar{t}W$ and $t\bar{t}Z$ contributions are estimated using the corresponding classes of the DNN output and the tails of the 3ℓ and 4ℓ CRs. In the $2\ell ss + 0\tau_h$ category, events are further classified depending on the flavour of the leptons ($ee/e\mu/\mu\mu$), except in the $t\bar{t}H$ high- and low-p_T nodes, where no flavour categorization is applied. In the $3\ell + 0\tau_h$ category, the tH and $t\bar{t}H$ nodes are separated according to the number of b-tagged jets.

6.1 Reconstruction of $m_{t\bar{t}H}$ and regression of p_T(H)

The $m_{t\bar{t}H}$ is reconstructed by computing the invariant mass of the visible objects assigned to the $t\bar{t}H$ system. The four momentum of the $t\bar{t}H$ system is computed by summing together the four momenta of the leptons and the ones of the up to leading six jets when sorted from highest to lowest p_T. A DNN regression of $m_{t\bar{t}H}$ was also studied. However, since the performance of the DNN was similar to that provided by summing the four vector of the visible decay products, the simpler strategy has been chosen.

A DNN is used to estimate the H boson p_T in each final state category. In each category, 70% of the simulated signal events constitute the training set, while 30% constitute the test set. These DNNs take as input the kinematic and high-level variables detailed in Table 2. The list of variables used in each event category is optimized to provide the best performance. The most important features are the kinematic ones, in particular the H Jet Tagger and Resolved Hadronic Top Tagger algorithms described in Ref. [9].

Each DNN was trained via gradient descent by minimizing a loss function defined targeting the generator-level H boson p_T distribution in $t\bar{t}H$ signal events. The loss function is based on the mean square loss but has an additional term to penalize large differences of the variance between the true and the regressed value. This penalty term is needed as all three DNNs tended to regress to a wider H boson p_T distribution than the truth data.

6.2 Signal region binning strategy

As already described in the previous section, events in each SR are separated based on the output values provided by the DNN multiclassifier. An event is assigned to a given class ($t\bar{t}H$, tH, $t\bar{t}W$, etc) based on its highest classifier output score in each final state category.

In the $2\ell ss + 0\tau_h$ category, background events get further split by their lepton flavor combination (ee, $e\mu$, $\mu\mu$). In the $3\ell + 0\tau_h$ category, background events get split according to their number of b-tagged jets using the medium identification WP. Finally, in the $2\ell ss + 1\tau_h$ category, no additional separation based on lepton flavor or number of b jets is made because of the low amount of events present in the category with the current luminosity.

In each group of background classes, quantiles of the amount of non-class-like events are used to divide the classifier scores into bins. The quantile of the background within each bin remains approximately constant. This approach increases the purity of class-like events in subsequent bins of the same group, while the purity of non-class-like events stays approximately constant.

The H boson p_T measurement is interpreted in terms of a custom binning ($[0, 120, 200, \infty]$) optimized for sensitivity. Similarly, the $m_{t\bar{t}H}$ measurement is performed with a custom binning of $[0, 750, 1000, \infty]$.

Depending on whether the measurement is performed in terms of the H boson p_T or $m_{t\bar{t}H}$, the $t\bar{t}H$-like classes are split differently. For the H boson p_T measurement, $t\bar{t}H$-like events are binned according to the regressed H boson p_T and split into low- and high-p_T^H classes. Events

Table 2: Input variables for the H boson p_T DNN-based regression. A check mark ($\checkmark$) indicates the variable is used in a given channel, whereas a long dash (—) indicates the variable is not used in that channel. The sum of the variables listed in the table corresponds to the sum of their four vectors. The list of variables has been optimized to ensure the best H boson p_T regression per each final state. The most important variables for the classifier are the sum of the first five jets followed by the lepton variables.

Variable description	$2\ell ss + 0\tau_h$	$2\ell ss + 1\tau_h$	$3\ell + 0\tau_h$
Jet variables*			
$j_1(p_T, \eta, \phi)$	—	$\checkmark$	—
$j_2(p_T, \eta, \phi)$	—	$\checkmark$	—
Score of DeepJet Discriminator $(j_{1,2})$	—	$\checkmark$	—
Number of jets in event $(j_{1,2})$	—	$\checkmark$	—
Lepton variables*			
$l_1(p_T, \eta, \phi)$	$\checkmark$	$\checkmark$	$\checkmark$
$l_1(m_T)$	—	$\checkmark$	—
$l_2(p_T, \eta, \phi)$	$\checkmark$	$\checkmark$	$\checkmark$
$l_2(m_T)$	—	$\checkmark$	—
$l_3(p_T, \eta, \phi)$	—	—	$\checkmark$
$\tau_h(p_T, \eta, \phi)$	—	$\checkmark$	—
kinematic features			
p_T^{miss}	$\checkmark$	$\checkmark$	$\checkmark$
$\Phi(p_T^{miss})$	$\checkmark$	—	$\checkmark$
High level variables			
$t_{had}(p_T, \eta, \phi)^\dagger$	$\checkmark$	—	$\checkmark$
$t_{had}(\text{BDT score})^\dagger$	$\checkmark$	—	$\checkmark$
Higgs jet tagger†	—	$\checkmark$	—
Jet is from from hadronic top flag† $(j_{1,2})$	—	$\checkmark$	—
(p_T, η, ϕ) of vectorial sum of variables of first five jets $\left(\sum_{n=1}^{5} j_1(p_T, \eta, \phi)\right)$	$\checkmark$	—	$\checkmark$
(p_T, η, ϕ) of vectorial sum of variables of remaining jets $(\sum_{n>5} j_1(p_T, \eta, \phi))$	$\checkmark$	—	$\checkmark$
(p_T, η, ϕ) from the vectorial sum of all jet and fakeable lepton variables $(\sum_n j_n + \sum_n l_n(p_T, \eta, \phi))$	$\checkmark$	—	$\checkmark$
Avg. ΔR distance among all jets	—	$\checkmark$	—
$m_{t\bar{t}H}$	—	$\checkmark$	—
Total number of variables	21	29	21

* Order determined by p_T
$\dagger$ Variables come from the "resolved hadronic top tagger" algorithm described in Ref. [9]

in the low-p_T^H class use the binning as a function of the regressed H boson p_T $[0, 120, 200, \infty]$. Conversely, for events in the high-p_T^H class, the lower p_T^H bins are sparsely populated. Therefore, the binning $[0, 200, 300, \infty]$ is used.

For the $m_{t\bar{t}H}$ measurement, all $t\bar{t}H$-like events are grouped together and binned according to a binning scheme, $[0, 750, 1000, \infty]$. Similar to the background classes, the binning oth the $t\bar{t}H$-like category is done in such a way that the fraction of background events remain approximately constant in each bin. This approach ensures that the background content stays consistent across bins, while the number of signal events increases with each iteration, thus maximising the sensitivity to the signal.

A comprehensive list of the number of bins for each channel is shown in Table 3 for the H boson p_T measurement and in Table 4 for the $m_{t\bar{t}H}$ measurement.

The signal extraction is accomplished through a profile likelihood ratio fit to the distributions of the observables described in Section 6.2. Several SRs for $2\ell ss + 0\tau_h$, $2\ell ss + 1\tau_h$, and $3\ell + 0\tau_h$ categories, along with CRs for the 3ℓ and 4ℓ categories, are utilized.

Table 3: Binning for the H boson p_T measurement for each channel.

Class	$2\ell\text{ss} + 0\tau_h$	$2\ell\text{ss} + 1\tau_h$	$3\ell + 0\tau_h$
low p_T	8	4	4
high p_T	4	4	4
tH	8	4	4
bkg	9	7	7
$t\bar{t}W$	13	—	—
Total	42	21	21

Table 4: Binning for the $m_{t\bar{t}H}$ measurement for each channel.

Class	$2\ell\text{ss} + 0\tau_h$	$2\ell\text{ss} + 1\tau_h$	$3\ell + 0\tau_h$
$t\bar{t}H$	18	12	12
tH	8	4	4
bkg	9	5	5
$t\bar{t}W$	13	—	—
Total	42	21	21

7 Systematic uncertainties

Several sources of uncertainties affect the predicted event yields for both signal and background processes. These arise from theory-related and experimental effects, as well as from the methods used to estimate data-driven backgrounds, such as nonprompt leptons and charge misidentification. These uncertainties, commonly referred to as systematic uncertainties, are modeled as nuisance parameters in the profile likelihood ratio fits used for signal extraction. These nuisance parameters are allowed to modify the event yield, accounting for the migration of events among regions and different bins in the distributions fitted in each region. Two categories of nuisance parameters are considered: those that purely affect the yield in a category (rate uncertainties) are assigned a log-normal probability density function, whereas those that also affect the shape of the distributions (shape uncertainties) are modelled via a polynomial interpolation with a Gaussian constraint and are also allowed to modify the event yields in a category [57].

The efficiencies of triggers based on one, two, or three leptons (electrons or muons) are measured as a function of the lepton multiplicity, with uncertainties ranging from 1% to 2%, using $t\bar{t}$+jets and diboson events recorded with p_T^{miss}-based triggers [55].

Efficiencies for electrons and muons to satisfy offline reconstruction and identification criteria are determined as functions of lepton p_T and η using the "tag-and-probe" method [34] in $Z/\gamma^* \to$ ee and $Z/\gamma^* \to \mu\mu$ events.

The τ_h identification efficiency and energy scale are measured in $Z/\gamma^* \to \tau\tau$ events [37] with uncertainties of 5% and 1.2%, respectively.

Jet energy scale uncertainties, ranging from a few percent depending on jet p_T and η, are derived using the p_T-balance method applied to dijet, and multijet events [35]. These uncertainties are propagated to signal and background expectations by varying jet energies in simulation, recalculating kinematic observables, and reapplying event selection. Uncertainties in jet

energy resolution are also evaluated but have smaller impacts.

The b-tagging efficiency is measured with an uncertainty of a few percent in $t\bar{t}$+jets and multijet events as a function of jet p_T and η. Heavy-flavor enrichment in multijet events is achieved by requiring a muon in the event. Mistag rates for light-quark and gluon jets are measured in multijet events, with uncertainties of 5–10% for the loose and 20–30% for the tight b-tagging criteria, depending on p_T and η [36].

The integrated luminosities for the 2016, 2017, and 2018 data-taking periods are known with uncertainties of 2.3–2.5% [58–60]. The total Run 2 (2016–2018) integrated luminosity is known with an improved uncertainty of 1.8%, reflecting the uncorrelated time evolution of certain systematic effects.

The uncertainties in the PU are evaluated by varying the number of inelastic pp interactions in simulated events by 4.6% [61]. The resulting impact on the $t\bar{t}$H and tH signal and background contributions modeled using simulation is less than 1%.

Theory-related uncertainties affecting event yields and the distributions of the DNN classifier outputs, used for signal extraction, are evaluated for the $t\bar{t}$H and tH signals as well as the main irreducible backgrounds ($t\bar{t}$W, and $t\bar{t}$Z production). Production cross section uncertainties are between –9.9 and +6.8 for $t\bar{t}$H, –7.3 and +5.1 for tH, –12.2 and +13.5 for $t\bar{t}$W, –10.2 and +11.7 for $t\bar{t}$Z, respectively [62]. These uncertainties arise from missing higher-order corrections, PDF variations, and uncertainties in the strong coupling constant α_S. The $t\bar{t}$W and $t\bar{t}$Z absolute rate constrained in situ with the following post-fit uncertainties as reported in Section 8.

Cross section uncertainties are important for quoting production rates relative to their SM expectations and for setting limits on the H boson coupling to the top quark. The impact of missing higher-order corrections on discriminating observables is assessed by varying renormalization and factorization scales up and down by a factor of two, following Refs. [63–65], avoiding opposite-direction variations. PDF uncertainties are evaluated according to Ref. [66].

The extrapolation of WZ and ZZ background rates from the 3ℓ- and 4ℓ CRs to the SR depends on the heavy-flavour content of these backgrounds. According to MC simulation, most b-jets in WZ and ZZ events result from misidentified light-quark or gluon jets, rather than from charm or bottom quarks. A 40% uncertainty is assigned to the modeling of heavy-flavour content, based on differences in jet multiplicity distributions between data and simulation in the 3ℓ CR. An additional systematic uncertainty covers the misidentification of light-quark or gluon jets as b jets.

The uncertainty in the nonprompt background is taken from Ref. [55], amounting to 30% in each channel.

Systematic uncertainties from the same source are treated as fully correlated across all the analysis channels. Theoretical uncertainties are fully correlated across all data-taking periods, while experimental uncertainties are treated as uncorrelated between the 2016, 2017, and 2018 datasets. This treatment is justified as the auxiliary measurements used to validate and, if necessary, correct the MC simulation are primarily statistical in nature and independent across the three data-taking periods, reflecting changes in detector conditions over time.

The systematic uncertainities in the cross section differential measurement of the $t\bar{t}$H signal rates are summarized in Table 5. The largest contributions arise from the uncertainties in the estimation of misidentified lepton and charge mismeasurement backgrounds, the statistical uncertainty of the observed data, and the theoretical uncertainties affecting yields and discriminating observables for the $t\bar{t}$H signal.

Table 5: Summary of the main systematic uncertainty sources, their type, and the correlations across the three data-taking years.

Uncertainty source	Type	Correlation 2016–2018
Trigger efficiency	Norm. / Shape	Uncorrelated
Identification and isolation efficiency for e and μ	Shape	Uncorrelated
Identification efficiency for τ_h	Shape	Uncorrelated
Energy scale of e, μ, and τ_h	Shape	Uncorrelated
Jet energy scale	Shape	Uncorr. / Corr.
b-tag efficiency and mistag rate	Shape	Uncorr. (stat.) / Corr. (exp.)
E_T^{miss} resolution and response	Shape	Correlated
Theoretical uncertainty	Norm. / Shape	Correlated
Fake background rate	Shape	Uncorrelated
Charge mismeasurement background rate	Norm.	Correlated
Electroweak and rare background rates	Norm.	Correlated
Luminosity	Norm.	Uncorr. / Corr.
Level-1 ECAL prefiring	Shape	Uncorrelated

8   Results

The signal extraction is performed using a profile likelihood ratio fit applied to the distributions of the DNN classifier described in Section 6.2. The analysis utilizes SRs corresponding to the $2\ell ss + 0\tau_h$, $2\ell ss + 1\tau_h$, and $3\ell + 0\tau_h$ categories, along with CRs defined for the 3ℓ and 4ℓ categories. Systematic uncertainties are treated as nuisance parameters in the profile likelihood ratio fits, as detailed in Section 7.

The most significant backgrounds in this analysis are the $t\bar{t}W$ and $t\bar{t}Z$ processes. These background contributions are estimated directly from data using the CRs specified in Section 5.3. Factors for $t\bar{t}W$ and $t\bar{t}Z$ are determined to be 1.53 pm 0.25 and 1.04 pm 0.25, respectively. The distribution of the number of jets is shown in the 3ℓ and 4ℓ CRs. The 3ℓ-CR is dominated by the WZ background, while the 4ℓ-CR is primarily influenced by the ZZ background.

This note presents two cross section measurements: one based on a custom binning scheme for the H boson p_T and the other based on a custom binning scheme for the $m_{t\bar{t}H}$ variable. Both measurements are found to be consistent with SM predictions, with p-values of 0.02 and 0.52, respectively.

Postfit distributions of the DNN classifier for the $2\ell ss + 0\tau_h$, $2\ell ss + 1\tau_h$, and $3\ell + 0\tau_h$ categories, as well as the 3ℓ and 4ℓ CRs, are shown in Fig. 4. All nodes of the DNN classifier are represented in these plots. As expected, the $t\bar{t}W$ and $t\bar{t}Z$ processes are the most prominent backgrounds in the $t\bar{t}H$ node.

The results of the cross section measurement and the signal strength (defined as σ/σ_{SM}), derived using the custom binning for the H boson p_T, are presented in Fig. 5 and Table 6. It is observed that the differential measurement is primarily limited by statistical uncertainties, which remain the dominant source of uncertainty. The main contributions to systematic uncertainties arise from the estimation of misidentified lepton and charge mismeasurement backgrounds, as well as theoretical uncertainties.

The same measurement is performed as a function of the $m_{t\bar{t}H}$ binning $[0, 750, 1000, \infty]$. The postfit distributions are presented in Fig. 6, while the cross section measurement and the signal strength as a function of the $m_{t\bar{t}H}$ binning are shown in Fig. 5 and Table 6. The differential measurement is dominated by statistical uncertainties. The primary sources of systematic un-
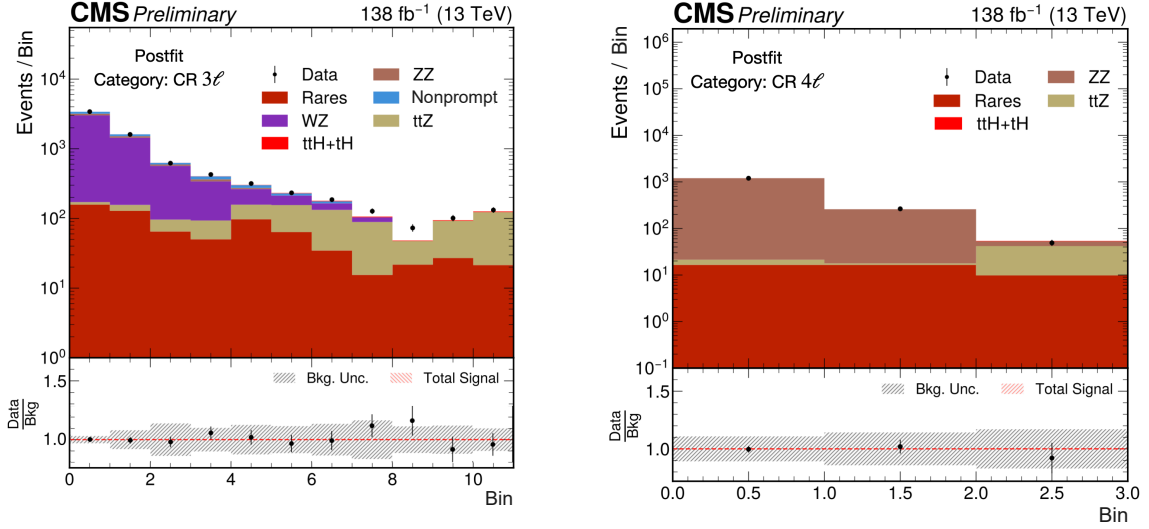
Figure 3: The distribution of the number of jets in the 3ℓ and 4ℓ CRs. In the 3ℓ CR, bins one to four correspond to events without b jets, bins five to eight correspond to events with exactly one b jet, and bins nine to twelve correspond to events with more than one b jet. In the 4ℓ CR, the first bin corresponds to events without jets, the second bin corresponds to events with more than zero jets and exactly one b jet, and the third bin corresponds to events with more than one jet and more than one b jet. The 3ℓ CR is dominated by WZ background, while the 4ℓ CR is dominated by ZZ.

Table 6: Measured signal strengths and corresponding uncertainties (68% CL) in different H boson p_T bins and $m_{t\bar{t}H}$ bins.

H boson p_T bin	Signal strength
$[0, 120)$ GeV	-0.78 ± 0.64 (stat) ± 0.48 (syst)
$[120, 200)$ GeV	4.36 ± 1.42 (stat) ± 0.80 (syst)
$[200, \infty)$ GeV	-1.32 ± 0.86 (stat) ± 0.46 (syst)
$m_{t\bar{t}H}$ bin	Signal strength
$[0, 750)$ GeV	0.71 ± 0.58 (stat) ± 0.36 (syst)
$[750, 1000)$ GeV	0.90 ± 1.23 (stat) ± 0.63 (syst)
$[1000, \infty)$ GeV	0.92 ± 1.02 (stat) ± 0.55 (syst)

certainties originate from the estimation of misidentified lepton backgrounds, charge mismeasurement backgrounds, and theoretical predictions.

9 Summary

The production of a Higgs boson in association with two top quarks ($t\bar{t}H$) is measured in final states containing multiple electrons, muons, or tau leptons decaying to hadrons and a neutrino, using proton-proton collisions recorded at a center-of-mass energy of 13 TeV by the CMS experiment. The analyzed data correspond to an integrated luminosity of 138 fb^{-1}. The analysis is optimised for events that contain $H \rightarrow WW$ or $H \rightarrow \tau\tau$ decays where each of the top quarks decays either semileptonically or exclusively to jets. The sensitivity to the signal process is maximized by including three signatures in the analysis, depending on the lepton multiplicity. The separation among the $t\bar{t}H$ signal and the background processes is enhanced through
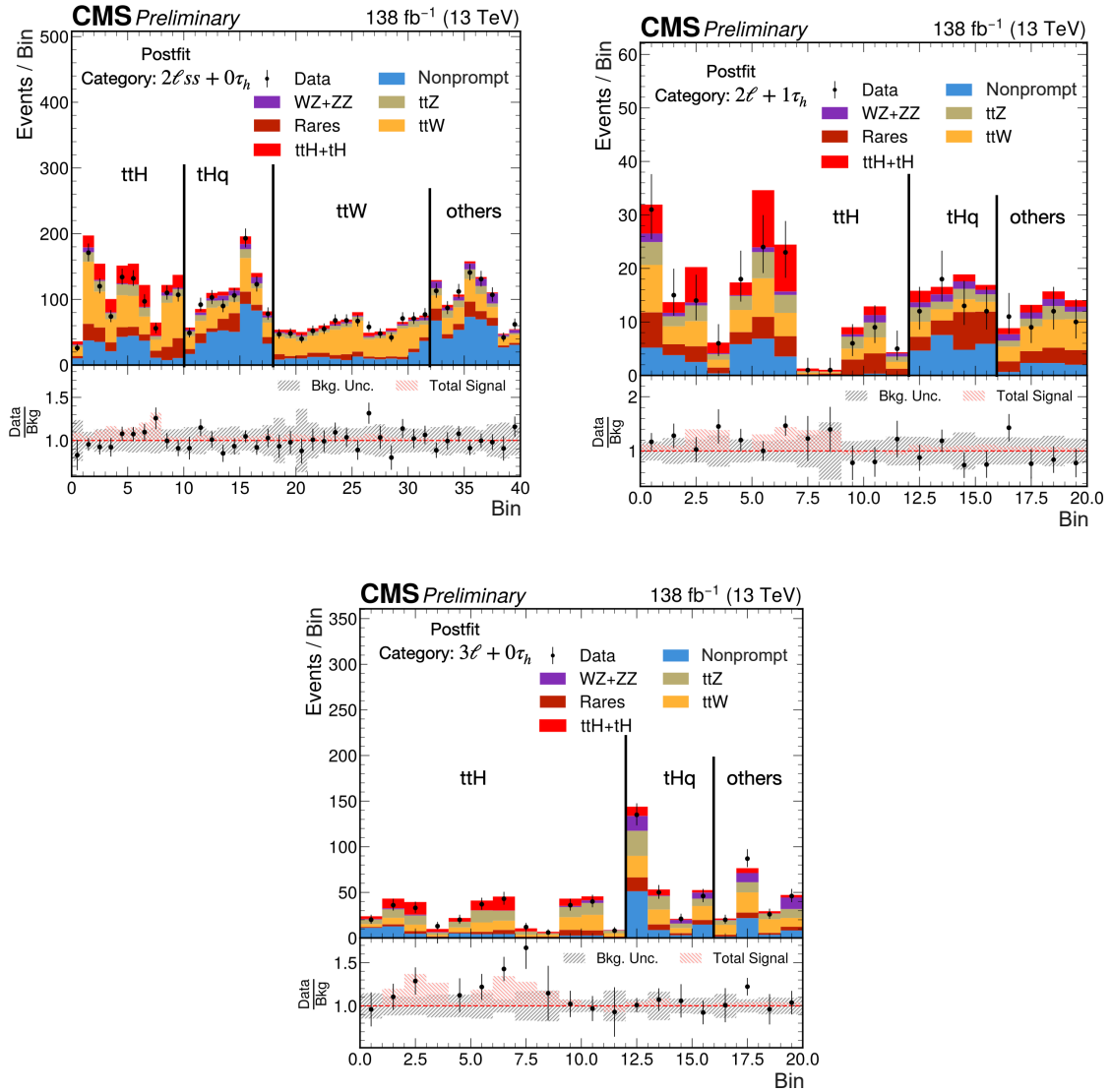
Figure 4: Postfit distributions of the DNN discriminant for $2\ell ss + 0\tau_{\mathrm{h}}$ (upper left), $2\ell ss + 1\tau_{\mathrm{h}}$ (upper right) and $3\ell + 0\tau_{\mathrm{h}}$ (bottom) categories. All nodes of the DNN classifier are shown in the plots, $t\bar{t}W$ and $t\bar{t}Z$ are the most abundant backgrounds in the $t\bar{t}H$ node.
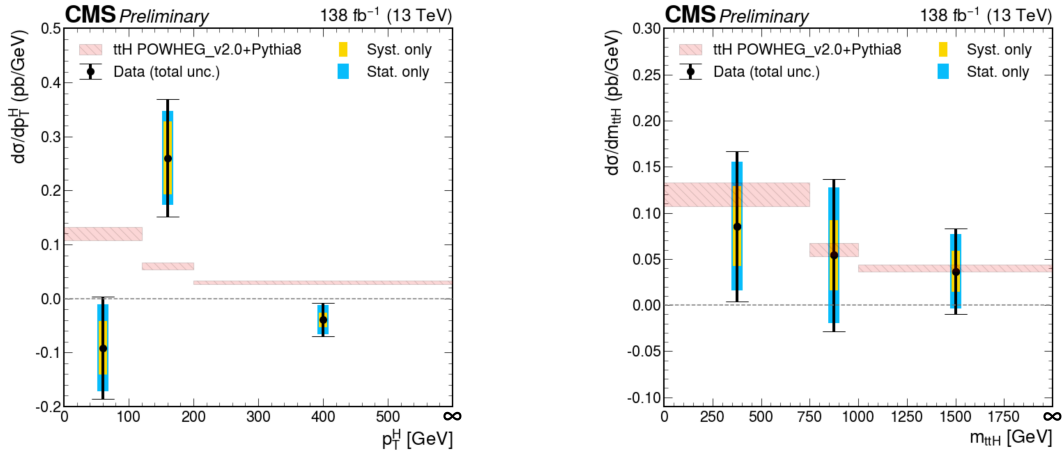
Figure 5: Measured differential cross section and uncertainties as a function of the p_T (left) and $m_{t\bar{t}H}$ (right) relative to the SM. Yellow band represents the impact of the systematic uncertainties, while the azure band represents the impact of the systematic uncertainties, while the azure band represents the impacts of the statistical uncertainties.

machine-learning techniques. Differential production rates are measured as a function of the Higgs boson transverse momentum and of the mass of the $t\bar{t}H$ system and found to be compatible with predictions from the standard model. This result is the first differential measurement of $t\bar{t}H$ production to date by the CMS Collaboration.

References

[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, "A portrait of the Higgs boson by the CMS experiment ten years after the discovery", *Nature* **607** (July, 2022) 6068, `doi:10.1038/s41586-022-04892-x`.

[4] CMS Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J.* **75** (2015), no. 5, 212, `doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662`.

[5] ATLAS, CDF, CMS and D0 Collaborations, "First combination of Tevatron and LHC measurements of the top quark mass", 2014. `https://arxiv.org/abs/1403.4427`.

[6] 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`.

[7] R. S. Chivukula, B. A. Dobrescu, H. Georgi, and C. T. Hill, "Top Quark Seesaw Theory of Electroweak Symmetry Breaking", *Phys. Rev.* **59** (1999) 075003, `doi:10.1103/PhysRevD.59.075003, arXiv:hep-ph/9809470`.
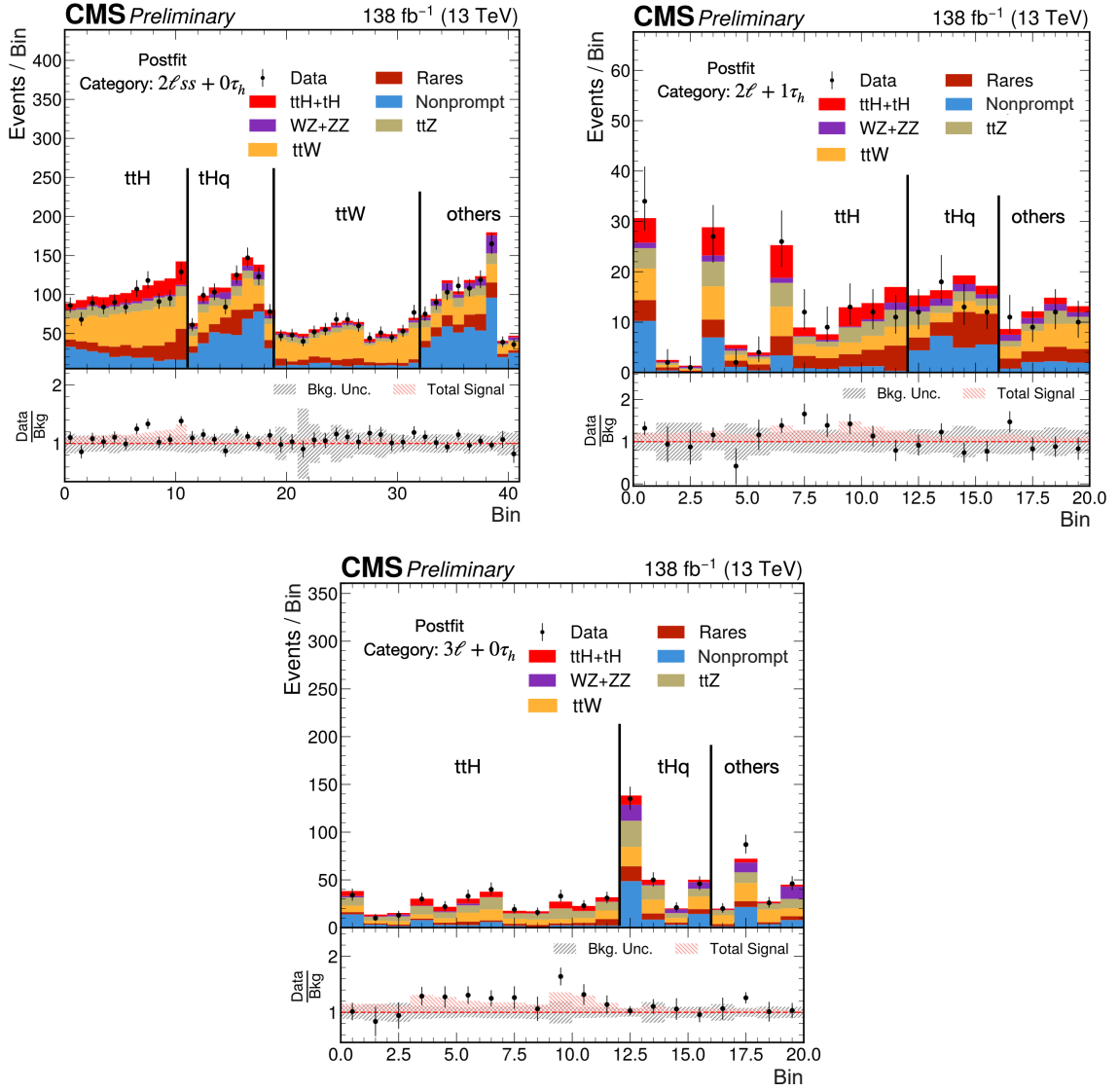
Figure 6: Postfit distributions of the DNN discriminant for $2\ell ss + 0\tau_h$ (upper left), $2\ell ss + 1\tau_h$ (upper right) and $3\ell + 0\tau_h$ (bottom). All nodes of the DNN classifier are shown in the plots, $t\bar{t}W$ and $t\bar{t}Z$ are the most abundant backgrounds in the $t\bar{t}H$ node.

[8] D. Delepine, J. M. Gerard, and R. Gonzalez Felipe, "Is the standard Higgs scalar elementary?", *Phys. Lett.* **372** (1996) 271, `doi:10.1016/0370-2693(96)00048-2`, `arXiv:hep-ph/9512339`.

[9] CMS Collaboration, "Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **81** (2021) 378, `doi:10.1140/epjc/s10052-021-09014-x`, `arXiv:2011.03652`.

[10] 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", 2024. `https://arxiv.org/abs/2407.10896`.

[11] CMS Collaboration, "Search for CP violation in ttH and tH production in multilepton channels in proton-proton collisions at 13 TeV", *Journal of High Energy Physics* **2023** (July, 2023) `doi:10.1007/jhep07(2023)092`.

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

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

[14] CMS Collaboration, "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS", *Eur. Phys. J. C* **81** (2021) 800, `doi:10.1140/epjc/s10052-021-09538-2`, `arXiv:2104.01927`.

[15] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[16] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.

[17] 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`.

[18] 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`.

[19] 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`.

[20] F. Maltoni, G. Ridolfi, and M. Ubiali, "b-initiated processes at the LHC: a reappraisal", *JHEP* **07** (2012) 022, `doi:10.1007/JHEP07(2012)022`, `arXiv:1203.6393`. [Erratum: `doi:10.1007/JHEP04(2013)095`].

[21] 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`.

[22] R. Frederix and I. Tsinikos, "Subleading EW corrections and spin-correlation effects in $t\bar{t}W$ multi-lepton signatures", *Eur. Phys. J. C* **80** (2020) 803, `doi:10.1140/epjc/s10052-020-8388-6`, `arXiv:2004.09552`.

[23] J. A. Dror, M. Farina, E. Salvioni, and J. Serra, "Strong tW Scattering at the LHC", *JHEP* **01** (2016) 071, `doi:10.1007/JHEP01(2016)071`, `arXiv:1511.03674`.

[24] L. Buonocore et al., "Precise Predictions for the Associated Production of a W Boson with a Top-Antitop Quark Pair at the LHC", *Phys. Rev. Lett.* **131** (2023), no. 23, 231901, `doi:10.1103/PhysRevLett.131.231901`, `arXiv:2306.16311`.

[25] 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`.

[26] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, `doi:10.1140/epjc/s10052-007-0490-5`, `arXiv:0706.2569`.

[27] 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`.

[28] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, `doi:10.1007/JHEP12(2012)061`, `arXiv:1209.6215`.

[29] NNPDF Collaboration, "Parton distributions for the LHC Run 2", *JHEP* **04** (2015) 040, `doi:10.1007/JHEP04(2015)040`, `arXiv:1410.8849`.

[30] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, `doi:10.1016/j.cpc.2015.01.024`, `arXiv:1410.3012`.

[31] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, `doi:10.1140/epjc/s10052-019-7499-4`, `arXiv:1903.12179`.

[32] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, `doi:10.1016/S0168-9002(03)01368-8`.

[33] J. Allison et al., "Recent developments in GEANT4", *Nucl. Instrum. Meth. A* **835** (2016) 186, `doi:10.1016/j.nima.2016.06.125`.

[34] CMS Collaboration, "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2011) 080, `doi:10.1007/JHEP01(2011)080`, `arXiv:1012.2466`.

[35] 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`.

[36] 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`.

[37] CMS Collaboration, "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **13** (2018) P10005, `doi:10.1088/1748-0221/13/10/P10005`, `arXiv:1809.02816`.

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

[39] 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`.

[40] 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`.

[41] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **13** (2018), no. 06, P06015, `doi:10.1088/1748-0221/13/06/P06015`, `arXiv:1804.04528`.

[42] K. Rehermann and B. Tweedie, "Efficient identification of boosted semileptonic top quarks at the LHC", *JHEP* **03** (2011) 059, `doi:10.1007/JHEP03(2011)059`, `arXiv:1007.2221`.

[43] CMS Collaboration, "Search for new physics in same-sign dilepton events in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **76** (2016) 439, `doi:10.1140/epjc/s10052-016-4261-z`, `arXiv:1605.03171`.

[44] CMS Collaboration, "Performance of the DeepTau algorithm for the discrimination of taus against jets, electrons, and muons", CMS Detector Performance Summary CMS-DP-2019-033, 2019.

[45] Y. Lecun, "Generalization and network design strategies", Technical Report of the University of Toronto, CRG-TR-89-4.

[46] 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`.

[47] M. Cacciari, G. P. Salam, and G. Soyez, "The catchment area of jets", *JHEP* **04** (2008) 005, `doi:10.1088/1126-6708/2008/04/005`, `arXiv:0802.1188`.

[48] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2016.

[49] CMS Collaboration, "CMS Phase 1 heavy flavour identification performance and developments", CMS Detector Performance Summary CMS-DP-2017-013, 2017.

[50] CMS Collaboration, "Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **08** (2018) 066, `doi:10.1007/JHEP08(2018)066`, `arXiv:1803.05485`.

[51] Particle Data Group, P. A. Zyla et al., "Review of particle physics", *Prog. Theor. Exp. Phys.* **2020** (2020) 083C01, `doi:10.1093/ptep/ptaa104`.

[52] CMS Collaboration, "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV", *JHEP* **11** (2017) 047, `doi:10.1007/JHEP11(2017)047`, `arXiv:1706.09936`.

[53] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ tev using the cms detector", *Journal of Instrumentation* **14** (July, 2019) P07004, `doi:10.1088/1748-0221/14/07/p07004`.

[54] CMS Collaboration, "Search for physics beyond the standard model in top quark production with additional leptons in the context of effective field theory", *Journal of High Energy Physics* **2023** (December, 2023) `doi:10.1007/jhep12(2023)068`.

[55] CMS Collaboration, "Search for CP violation in $t\bar{t}$H and tH production in multilepton channels in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **07** (2023) 092, `doi:10.1007/JHEP07(2023)092`, `arXiv:2208.02686`.

[56] T. Adye, "Unfolding algorithms and applications", *Journal of Physics: Conference Series* **368** (2011) 012029, `doi:10.1088/1742-6596/368/1/012029`.

[57] CMS Collaboration, "The CMS statistical analysis and combination tool: Combine", *Computing and Software for Big Science* **8** (November, 2024) `doi:10.1007/s41781-024-00121-4`.

[58] CMS Collaboration, "CMS luminosity measurements for the 2016 data-taking period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

[59] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2017.

[60] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2018.

[61] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2018) 161, `doi:10.1007/JHEP07(2018)161`, `arXiv:1802.02613`.

[62] D. de Florian et al., "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", CERN Report CERN-2017-002-M, 2016. `doi:10.23731/CYRM-2017-002`, `arXiv:1610.07922`.

[63] M. Cacciari et al., "The $t\bar{t}$ cross-section at 1.8 TeV and 1.96 TeV: a study of the systematics due to parton densities and scale dependence", *JHEP* **04** (2004) 068, `doi:10.1088/1126-6708/2004/04/068`, `arXiv:hep-ph/0303085`.

[64] S. Catani, D. de Florian, M. Grazzini, and P. Nason, "Soft gluon resummation for Higgs boson production at hadron colliders", *JHEP* **07** (2003) 028, `doi:10.1088/1126-6708/2003/07/028`, `arXiv:hep-ph/0306211`.

[65] R. Frederix et al., "Four-lepton production at hadron colliders: MADGRAPH5_aMC@NLO predictions with theoretical uncertainties", *JHEP* **02** (2012) 099, `doi:10.1007/JHEP02(2012)099`, `arXiv:1110.4738`.

[66] J. Butterworth et al., "PDF4LHC recommendations for LHC Run 2", *J. Phys. G* **43** (2016) 023001, `doi:10.1088/0954-3899/43/2/023001`, `arXiv:1510.03865`.