CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2024/09/28

Search for heavy pseudoscalar and scalar bosons decaying to top quark pairs in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

The CMS Collaboration

Abstract

A search for heavy pseudoscalar or scalar bosons decaying to a top quark pair ($\mathrm{t\bar{t}}$) in final states with one or two charged leptons is presented, using $138\,\text{fb}^{-1}$ of proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ recorded by the CMS experiment at the CERN LHC. The invariant mass of the reconstructed $\mathrm{t\bar{t}}$ system and variables sensitive to its spin state are used to discriminate against the standard model $\mathrm{t\bar{t}}$ background. An excess of the data above the background prediction, as modeled using perturbative quantum chromodynamics (QCD) only, is observed. The excess is located close to the $\mathrm{t\bar{t}}$ production threshold and it significantly favors the pseudoscalar signal hypothesis over the scalar hypothesis. It is compatible with the production of a ${}^1S_0^{[1]}$ $\mathrm{t\bar{t}}$ bound state (η_t), as predicted by a simplified model of nonrelativistic QCD, with a cross section of 7.1 pb and an uncertainty of 11%. The excess has a significance of above five standard deviations. Including the η_t contribution in the background modeling, exclusion limits at 95% confidence level are set on the coupling of further pseudoscalar or scalar bosons to top quarks in a mass range of $365-1000\,\text{GeV}$ and relative widths of $0.5-25\%$.

1 Introduction

The observation of a Higgs boson with mass of ≈125 GeV by the ATLAS and CMS Collaborations in 2012 [1–3] confirmed the existence of an elementary spin-0 state, a crucial ingredient of the standard model (SM) of particle physics. While only one such state is required in the SM, many beyond-the-SM (BSM) extensions predict additional spin-0 states, such as the two-Higgs-doublet models (2HDMs) [4], models predicting a new electroweak singlet [5], and models with a combination of singlet and doublet fields [6]. These additional bosons may also provide a portal to dark matter by acting as a mediator between SM and dark matter particles [7, 8]. The new states introduced in these BSM extensions usually include pseudoscalar (CP-odd) neutral bosons, scalar (CP-even) neutral bosons, and charged bosons. We use the symbol A to denote pseudoscalar neutral states, H for scalar neutral states not identified as the one with a mass of 125 GeV, and Φ as a common symbol to refer either to A or H bosons.

Due to their large mass, top quarks play a key role in searches for new physics. Provided that additional Φ bosons couple to fermions via a Yukawa interaction with coupling strength proportional to the fermion mass, Φ bosons with mass larger than twice the top quark mass m_t may have the decay to a top quark pair ($t\bar{t}$) as the dominating channel. This is true especially for A bosons, whose decays to weak vector bosons are suppressed by CP symmetry, as well as for H bosons in 2HDMs in the vicinity of the alignment limit [9].

In this paper, we consider a Yukawa-like coupling between Φ bosons and top quarks. The corresponding terms in the Lagrangian for the two CP eigenstates are:

$$\mathcal{L}_{\text{Yukawa,A}} = i g_{At\bar{t}} \frac{m_t}{v} \bar{t} \gamma_5 t A, \qquad \mathcal{L}_{\text{Yukawa,H}} = -g_{Ht\bar{t}} \frac{m_t}{v} \bar{t} t H, \tag{1}$$

where the strength of the couplings is controlled by real-valued coupling modifiers $g_{\Phi t\bar{t}} \geq 0$, and v is the vacuum expectation value of the SM Higgs field. We probe Φ boson masses in the range 365–1000 GeV and total widths of 0.5–25% relative to the mass.

The production of Φ bosons is dominated by the gluon fusion process with only top quarks in the loop, which together with the Φ bosons decaying into $t\bar{t}$ interferes with the SM $t\bar{t}$ production. Feynman diagrams for both processes are shown in Fig. 1. The resonant component results in a Breit–Wigner peak in the $t\bar{t}$ invariant mass ($m_{t\bar{t}}$) distribution, while the interference component may be either destructive or constructive, with shape and magnitude of the $m_{t\bar{t}}$ distribution depending on the phase space region, the specific signal model, and the types of particles that appear in the loop of the production diagram [10, 11]. The sum of the components may result in a peak-dip structure in the $m_{t\bar{t}}$ distribution [12–14].

Decays of the A and H bosons produce $t\bar{t}$ systems in the $^1S_0$ and $^3P_0$ states, respectively [14], while the SM $gg \to t\bar{t}$ production results in a mixture of states that changes with the partonic center-of-mass energy. Furthermore, top quarks often decay before hadronization, and spin

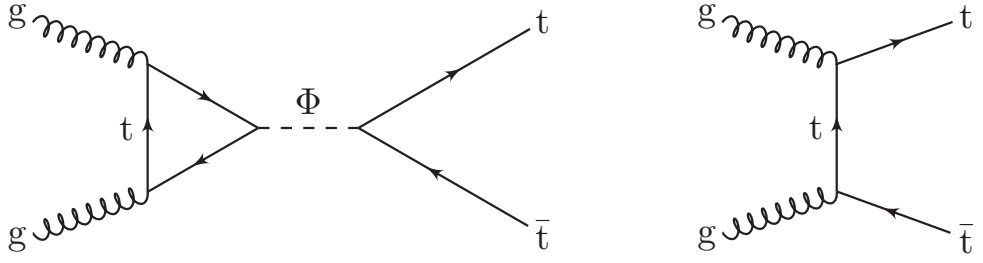

Figure 1: The Feynman diagram for the signal process (left) and an example diagram for the SM production of top quark pairs (right).

information is preserved in angular distributions of its decay products [15, 16]. Therefore, we exploit angular observables to discriminate between the different $\mathrm{t\bar{t}}$ spin states present in signal and background production.

In the SM, $\mathrm{t\bar{t}}$ production is described in quantum chromodynamics (QCD). State-of-the-art cross section predictions rely on perturbative QCD (pQCD) and include electroweak (EW) corrections. Additional enhancements of $\mathrm{t\bar{t}}$ production near its threshold are predicted in nonrelativistic QCD, in part due to the production of toponium bound states [17–20]. We use a simplified model of the production of the color-singlet pseudoscalar bound state ${}^1S_0^{[1]}$, referred to as η_t [20].

This paper describes a search for pseudoscalar and scalar bosons decaying to $\mathrm{t\bar{t}}$ in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ using the CMS detector at the CERN LHC. The analyzed data set corresponds to an integrated luminosity of $138\,\mathrm{fb}^{-1}$, collected in 2016–2018 [21–23]. Events in the single-lepton (ℓj) channel are selected with exactly one electron or muon that is associated with at least three jets, and in the dilepton ($\ell\bar{\ell}$) channel with exactly two electrons or muons and at least two jets. The top quark four-momenta are estimated with dedicated kinematic reconstruction algorithms and the resulting $m_{\mathrm{t\bar{t}}}$ together with additional observables sensitive to the spin state of the $\mathrm{t\bar{t}}$ system are used to search for Φ bosons.

The data is compared with the pQCD background prediction alone , as well as with predictions that include Φ boson and/or η_t production. For Φ boson production, constraints on the coupling modifier between the Φ and the top quarks $g_{\Phi\mathrm{t\bar{t}}}$ are derived as a function of the Φ boson mass and width in three signal configurations—one A boson, one H boson, or both one A and one H boson—and two background scenarios—excluding and including η_t production as part of the background. In addition, we test whether η_t production alone can sufficiently describe the data by extracting its cross section $\sigma(\eta_\mathrm{t})$ in a scenario with no Φ boson contributions.

This search updates a similar analysis performed by the CMS experiment using $35.9\,\mathrm{fb}^{-1}$ of data collected in 2016, where a moderate signal-like deviation compatible with A boson production with a mass of 400 GeV was found [24], without inclusion of any contribution from $\mathrm{t\bar{t}}$ bound states as background. Similar searches have also been performed by the ATLAS experiment using $20.3\,\mathrm{fb}^{-1}$ of 8 TeV data [25] and $140\,\mathrm{fb}^{-1}$ of 13 TeV data [26].

2 The CMS detector and event reconstruction

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity (η) coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. More detailed descriptions of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Refs. [27, 28].

Events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of $4\,\mu$s [29]. The second level, known as the high-level trigger, consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to

around 1 kHz before data storage [30].

The primary vertex (PV) is taken to be the vertex corresponding to the hardest scattering in the event, evaluated using tracking information alone, as described in Section 9.4.1 of Ref. [31]. The particle-flow (PF) algorithm [32] aims to reconstruct and identify each individual particle in an event, with an optimized combination of information from the various elements of the CMS detector. The reconstructed particles are referred to as PF candidates in the following. The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from a combination of the electron momentum at the PV as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

For each event, hadronic jets are clustered from the PF candidates using the infrared and collinear safe anti-k_T algorithm [33, 34] with a distance parameter of 0.4. Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5–10% of the true momentum over the entire transverse momentum (p_T) spectrum and detector acceptance. Additional proton-proton interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, charged particles identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions. Jet energy corrections are derived from simulation to bring the measured response of jets to that of particle level jets on average. In situ measurements of the momentum balance in dijet, photon+jet, Z+jet, and multijet events are used to account for any residual differences in the jet energy scale between data and simulation [35]. The jet energy resolution amounts typically to 15–20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV [35]. Additional selection criteria are applied to each jet to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures. To be considered in the data analysis, jets are required to satisfy $|\eta| < 2.4$, to have $p_T > 30$ (20) GeV in the ℓj ($\ell\bar{\ell}$) channel, and to be separated by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.4$ from any selected lepton, where $\Delta\eta$ and $\Delta\phi$ are the η and azimuthal angle differences between lepton and jet, respectively.

Jets originating from b quarks are identified with the DEEPJET algorithm [36–38]. The employed working point has a selection efficiency for b quark jets of about 78%, and a misidentification rate of 15 (2)% for c quark jets (light-quark and gluon jets evaluated together and referred to as light jets in the following), as evaluated in simulated $t\bar{t}$ samples. Differences between data and simulation in the b tagging efficiency and misidentification rate are accounted for by scale factors that depend on the jet p_T and η.

Electrons are measured in the pseudorapidity range $|\eta| < 2.5$ as energy deposits in the ECAL matched to a track. The momentum resolution for electrons with $p_T \approx 45$ GeV from $Z \to ee$ decays ranges from 1.6 to 5%. It is generally better in the barrel region than in the endcaps, and also depends on the bremsstrahlung energy emitted by the electron as it traverses the material in front of the ECAL [39, 40]. Only electrons with $|\eta| < 2.4$ and $p_T > 20$ GeV are considered in the data analysis. In the $\ell\bar{\ell}$ channel, well-identified electron candidates are selected using identification criteria based on boosted decision trees with a working point targeting a 90%

efficiency, with a misidentification rate of 1% and 3% in the barrel and endcap regions respectively. In the ℓj channel, well-identified electrons are selected using the "tight" working point of cut-based criteria, with an additional requirement of being consistent with originating from the PV. The efficiency of the "tight" working point of cut-based criteria is about 70%, with a misidentification rate of 1% and 2% in the barrel and endcap regions respectively. Furthermore, the "veto" working point of the same cut-based identification criteria is used to define a sample of loosely identified electrons used to veto events in the ℓj channel. All of these criteria are described in detail in Ref. [39].

Muons are measured in the pseudorapidity range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. Matching muons to tracks measured in the silicon tracker results in a relative p_{T} resolution, for muons with p_{T} up to 100 GeV, of 1% in the barrel and 3% in the endcaps. The p_{T} resolution in the barrel is better than 7% for muons with p_{T} up to 1 TeV. Only muons with $p_{\mathrm{T}} > 20$ GeV are considered in the data analysis. For use in the main event selection, well-identified muon candidates are required to pass the "tight" working point of the cut-based identification criteria described in Ref. [41]. The selection efficiency of well-identified muons, together with the isolation requirements described below, is 75 to 85%. The misidentification rate for well-identified muons is 0.1 to 0.3%, and the probability to incorrectly label muons within jets as isolated is 5 to 15%. Loosely identified muons are those passing the "loose" working point of the cut-based identification criteria, and are used in the ℓj channel to veto events.

Lepton candidates are required to be isolated from other activity in the event, to enrich the sample of leptons resulting from the prompt, direct decays of massive W and Z bosons. The relative isolation I_{rel} is calculated as the p_{T} sum of charged-hadron, neutral-hadron, and photon PF candidates inside a cone of $\Delta R = 0.4$ around the lepton, divided by the lepton p_{T}. An estimated contribution from pileup is subtracted in this calculation [39, 41]. Well-identified muons are required to have $I_{\mathrm{rel}} < 0.15\,(0.25)$ in the ℓj ($\ell\bar{\ell}$) channel, while loosely identified muons are required to have $I_{\mathrm{rel}} < 0.25$. For electrons, the isolation is already included in the identification criteria defined above [39]. Simulation-to-data scale factors that depend on the lepton p_{T} and η are used to correct for small differences in lepton trigger, identification, and I_{rel} efficiency.

The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [42]. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is modified to account for corrections to the energy scale of the reconstructed jets in the event.

3 Data and simulated event samples

The analyzed data have been recorded in 2016–2018 using triggers that require the presence of a single isolated electron or muon, or the presence of two such leptons including all possible flavor combinations. Four independent data-taking eras are considered: 2016pre (19.5 fb^{-1}), 2016post (16.8 fb^{-1}), 2017 (41.5 fb^{-1}), and 2018 (59.8 fb^{-1}). The 2016 data set is split into two eras because of a modification of the APV readout chip settings that affects the efficiency of the track hit reconstruction during the 2016 data-taking period [43], where the identifiers "pre" and "post" refer to the periods before and after this modification. The 2016pre, 2016post, and 2017 eras are also affected by an inefficiency caused by the gradual shift in the timing of the inputs to the ECAL L1 trigger in the regions $|\eta| > 2.0$ [29]. Correction factors are computed from data and applied to the acceptance evaluated by simulation to account for this effect.

In order to compare the collected data to theoretical predictions, Monte Carlo (MC) samples are produced with events simulating the signal and SM background processes. Different event generators are employed for the calculation of matrix elements (MEs). In all cases, the generators employ the NNPDF3.1 parton distribution functions (PDFs) [44] and are interfaced with PYTHIA 8.240 [45] for fragmentation and hadronization using the CP5 underlying event tune [46, 47]. The nominal value of m_t is set to 172.5 GeV in all samples involving top quarks, as well as in the computation of theoretical corrections that are applied to them. The simulated events are processed through the CMS detector simulation based on the GEANT4 program [48]. Separate MC samples are generated corresponding to the data-taking conditions of each of the four eras. Pileup events are generated with PYTHIA and overlaid in all samples. The simulated events are weighted to reproduce the distribution of the number of pileup interactions observed in data, assuming a minimum bias cross section of 69.2 mb. On average, there are 23 collisions per bunch crossing in 2016, 33 in 2017, and 32 in 2018 [49].

The $\Phi \to t\bar{t}$ signal process is simulated at leading-order (LO) accuracy in pQCD using a custom model in the MADGRAPH5_aMC@NLO 2.6.5 event generator [50] that implements the full kinematics of the top quark loop of the gluon fusion production, including m_t effects, via an effective coupling between the Φ bosons and gluons [51]. Event samples are produced for different Φ mass and relative total width values, such that a good coverage throughout the region of phase space probed in this search is obtained. They are reweighted to target signal hypotheses by the event-by-event ratios of the squared MEs of the target signal hypothesis and the one used in the original event simulation. The target signal hypotheses in this search are Φ bosons with masses of 365, 380, and 400–1000 GeV (in steps of 25 GeV), and relative total widths of 0.5–3 (in steps of 0.5), 4–8 (in steps of 1), 10, 13, 15, 18, 21, and 25%. We use the notation "A$(400, 5\%)$" to refer to Φ bosons of a particular CP eigenstate, mass (in GeV), and relative total width. The factorization and renormalization scales, μ_F and μ_R, are set on an event-by-event basis to $m_{t\bar{t}}/2$, following the choice in Ref. [52]. The top quarks from the Φ boson decay are decayed in MADGRAPH5_aMC@NLO, preserving their spin correlations.

Separate samples are generated for events corresponding to resonant Φ boson production, and for events corresponding to interference terms in the ME calculation between Φ boson and SM $t\bar{t}$ background production. Events in the interference samples can receive negative weights, reflecting the sign of the corresponding part of the squared ME in the presence of a destructive interference. Since the Φ boson is produced via gluon fusion with a top quark loop, the $\Phi t\bar{t}$ coupling appears twice in the ME. As a result, events originating from the resonance ME terms correspond to a cross section proportional to $g^4_{\Phi t\bar{t}}$, while those from interference correspond to a cross section proportional to $g^2_{\Phi t\bar{t}}$.

We calculate cross sections for resonant Φ boson production at next-to-next-to-LO (NNLO) accuracy with the SUSHI 1.7.0 program [53, 54] in the context of Type-II 2HDM models, where the 2HDMC program [55] is used to calculate the remaining model parameters for a given signal hypothesis. The coupling modifiers of the Φ bosons to bottom and charm quarks are set to zero. The ratio of the NNLO cross section to the LO cross section calculated with MADGRAPH5_aMC@NLO is used as K factor to normalize the resonant part of the signal samples, with typical values around 2.

For the interference component of the signal samples, we apply K factors corresponding to the geometric mean of those applied to the resonant signal and the SM $t\bar{t}$ process [52]. Here, the SM $t\bar{t}$ production K factor is calculated as the ratio between the $t\bar{t}$ cross section at NNLO in pQCD with next-to-next-to-leading logarithmic (NNLL) soft-gluon resummation, as described below, and at LO in pQCD with leading logarithmic resummation. The nominal value of the

K factor is 1.49, and is within 1.42 and 1.55 for different m_t values and scale choices used in the computation. For the H signal, we have compared the resonance and interference K factors with a recent explicit NLO calculation in the scope of a one-Higgs-singlet extension of the SM in Ref. [56]. We find good agreement for the resonance component and significant differences for the interference component. However, we have checked that this discrepancy does not significantly alter the conclusions of this work.

The η_t process is simulated at LO accuracy in pQCD, utilizing a custom model obtained from Ref. [20], implemented into the MADGRAPH5_aMC@NLO 2.6.5 event generator. The model is similar to the one used for the $\Phi \rightarrow t\bar{t}$ signal generation, although its effective gluon-pseudoscalar coupling is implemented as an effective contact interaction instead of via the top quark loop. Samples of resonant $\eta_t \rightarrow WbWb$ events are produced, to allow contributions from off-shell top quarks. The η_t mass and width are set to 343 and 7 GeV, respectively. Other simulation parameters are set following the recommendations of the model authors. Only events with $|m_{WbWb} - 343\,\text{GeV}| < 6\,\text{GeV}$ at the generator level are kept for further analysis. The version of the model used here does not include the nonrelativistic Hamiltonian reweighting mentioned in Ref. [20]. This is expected to have a negligible effect on this analysis, which is performed on reconstructed distributions, considering that the reweighting has a very small effect on parton-level distributions [57]. The η_t signal is normalized to a nominal cross section of 6.43 pb in accordance with Ref. [20].

The main background contribution originates from the SM $t\bar{t}$ production process, and is simulated at next-to-LO (NLO) order in pQCD using the POWHEG v2 generator [58–61]. The μ_F and μ_R scales are set to $\sqrt{m_t^2 + p_{T,t}^2}$, where m_t and $p_{T,t}$ are the mass and p_T of the top quarks in the underlying Born-level configuration. The sample is normalized to the predicted $t\bar{t}$ production cross section of $833.9^{+20.5}_{-30.0}$ pb, as calculated with the TOP++2.0 program at NNLO in pQCD, and including soft-gluon resummation at NNLL order [62]. The uncertainty is derived from the independent variations of μ_F and μ_R. To improve the theoretical description of the SM $t\bar{t}$ production process, the sample is reweighted to account for NNLO QCD and NLO EW corrections. The NNLO QCD prediction is calculated using the MATRIX program [63], using a nominal scale choice of $0.5(\sqrt{m_t^2 + p_{T,t}^2} + \sqrt{m_t^2 + p_{T,\bar{t}}^2})$. The NLO EW prediction is calculated using the HATHOR program [64–67], using the same nominal scale choice as the NNLO QCD prediction. Both predictions are computed using the same PDF set as the SM $t\bar{t}$ sample. The weights are applied double-differentially at the generator level as a function of $m_{t\bar{t}}$ and the cosine of the angle between the direction of the top quark in the zero-momentum frame (ZMF) of the $t\bar{t}$ system and the direction of the $t\bar{t}$ system in the laboratory frame, $\cos\theta_t^*$. The latter is sensitive to the spin state of the $t\bar{t}$ system and correlated to the observables employed in this search.

Other SM background events originate from single top quark production, single vector boson production (Drell–Yan Z/γ^* and W), diboson production (WW, WZ, and ZZ), $t\bar{t}$ production in association with a vector boson (referred to as $t\bar{t}V$), and events composed uniquely of jets produced through the strong interaction, referred to as QCD multijet processes. The single top quark production processes, via the t, tW, and s channels, are generated at NLO using POWHEG v2, POWHEG, and MADGRAPH5_aMC@NLO, respectively [68, 69]. The samples are normalized using the NLO cross section predictions for the t and s channels [64, 70], and approximate NNLO prediction for the tW channel [71]. The Z/γ^* process is generated with the POWHEG event generator [59, 60] with a multi-scale-improved NNLO accuracy in QCD [72, 73], matched with PYTHIA 8 for initial-state radiation and the PHOTOS package [74, 75] for final-state radiation. The W event samples are generated using MADGRAPH5_aMC@NLO at LO with up

to four additional partons, and the MLM matching scheme [76] is used to combine the different parton multiplicities. The single vector boson production cross sections are calculated at NNLO [77, 78], however, in the $\ell\bar{\ell}$ channel, the normalization of the Z/γ^* contribution is directly determined from a control region in data. Events simulating the diboson processes are generated using PYTHIA and normalized to the respective NNLO (WW) [79] or NLO (WZ and ZZ) [80] cross sections. For the WW process, we checked that instead explicitly simulating non-resonant WWb$\bar{\text{b}}$ production, which leads to the same final state as t$\bar{\text{t}}$ production, does not change the results of this work. The t$\bar{\text{t}}$V events are generated at NLO with MADGRAPH5_aMC@NLO, applying MC@NLO [81] merging, and are normalized using NLO cross section predictions. Finally, the QCD multijet events are simulated with PYTHIA 8.

4 Data analysis in the ℓj channel

In the ℓj channel, events are selected for further analysis that contain exactly one well-identified lepton with $p_T > 30\,\text{GeV}$, as defined in Section 2. For data recorded during 2018 and most of 2017 except for an early period, a higher threshold of $p_T > 34\,\text{GeV}$ is applied if the lepton is an electron, in order to account for higher trigger-level thresholds. Events containing additional loosely identified leptons (as defined in Section 2) with $p_T > 20\,\text{GeV}$ are rejected. Events are required to contain at least three jets with $p_T > 30\,\text{GeV}$, of which at least two are required to be b tagged.

4.1 Kinematic reconstruction

Each selected event is reconstructed under the assumption of a t$\bar{\text{t}}$ pair with one leptonically and one hadronically decaying top quark. The first step is to determine the neutrino four-momentum based on the measured p_T^{miss}, and the second step is to assign jets to the four final-state quarks. Different procedures are followed for events with at least four or exactly three jets, as described below.

The neutrino four-momentum p^ν is reconstructed with the algorithm described in Ref. [82], separately using each b jet in the event as candidate for the b jet originating from the leptonically decaying W boson. Mass constraints of the W boson and top quark are formulated, and the p^ν that satisfies these constraints and minimizes the distance $D_\nu = |p_T^{\text{miss}} - p_T^\nu|$ is used as unique solution for this b jet candidate [82]. If no solution is found for any b jet, the event is rejected.

For events with four or more jets, a likelihood function is constructed using the product of the probability density of the minimal D_ν and the two-dimensional probability density of the invariant masses of the hadronically decaying top quark and W boson. The probability densities are evaluated from simulated events in which all jets are correctly identified. All possible assignments of jets to the four final-state quarks are evaluated, provided that only b-tagged jets are assigned as b and $\bar{\text{b}}$ quark candidates. The best jet assignment is taken to be the one that maximizes this likelihood.

For events with exactly three jets, the techniques described in Ref. [83] are applied. A likelihood function is constructed using the product of the probability density of the minimal D_ν and the probability density of the invariant mass of the two jets assigned as originating from the hadronically decaying top quark. As with the case of four or more jets, the best assignment is taken to be the one that maximizes this likelihood. There are two typical cases of t$\bar{\text{t}}$ events that only have three jets. The first and more common case is when one or more quarks from the t$\bar{\text{t}}$ decay lie outside of the detector acceptance, which we refer to as lost-jet events. The

second case typically occurs in the high-momentum regime, where the ΔR between the top quark decay products are lower, leading to multiple quarks being clustered into one jet. These events are referred to as partially merged events. Once the best jet assignment is identified, a correction is applied to the four-momentum of the hadronically decaying top quark as a function of its reconstructed mass. The correction factor is larger for lost-jet events and is close to one for partially merged events, since a significant energy loss is expected only in the former case.

In events where all visible $\mathrm{t\bar{t}}$ decay products are available, the correct combination is reconstructed in 74% of all cases for four-jet reconstruction and 83% for three-jet reconstruction. With respect to all $\mathrm{t\bar{t}}$ events, these correspond to rates of 37 and 61%, respectively.

The signal is extracted using 2D templates built using the $m_{\mathrm{t\bar{t}}}$ and $|\cos\theta^*_{\mathrm{t}_\ell}|$ variables. The angle $\theta^*_{\mathrm{t}_\ell}$ is defined between the reconstructed leptonically decaying top quark in the ZMF and the direction of the $\mathrm{t\bar{t}}$ system in the laboratory frame, analogously to θ^*_{t} introduced in Section 3. The spin-0 nature of the signals leads to the top quarks being emitted isotropically in the $\mathrm{t\bar{t}}$ ZMF, resulting in a flat $\cos\theta^*_{\mathrm{t}_\ell}$ distribution at the generator level in the absence of kinematic selections. The SM distribution on the other hand peaks toward high values of $|\cos\theta^*_{\mathrm{t}_\ell}|$, due to the contribution from other spin states. As a result, the $|\cos\theta^*_{\mathrm{t}_\ell}|$ distribution will be enriched with signal events at low values.

4.2 Background estimation

The SM background in the ℓj channel is estimated from MC simulation for $\mathrm{t\bar{t}}$ and single top quark production, while QCD multijet production and EW processes (mostly W boson and small contributions from Z/γ^*, diboson, and $\mathrm{t\bar{t}V}$ production) are estimated using control samples in data without b-tagged jets, with the expected single top quark and $\mathrm{t\bar{t}}$ contributions subtracted out. The ratio of simulated background events in the signal and control regions is applied as normalization factor to the obtained background distributions. Since the QCD multijet background is difficult to simulate for the $\mathrm{t\bar{t}}$ selection, a normalization uncertainty of 50% is estimated based on the statistical uncertainty in the simulation and the differences of the observed and expected event yields in the control region.

To estimate the effect of changing the b-tagging requirements on the kinematic distributions, this procedure is repeated for three different selections of the highest allowed b-tagging discriminant value in the event. The shape differences between the central selection and the selections with a higher and lower allowed value of the highest b-tagging discriminant are taken into account as uncertainties in the background estimation. This procedure has been validated in simulation, and the kinematic distributions obtained from the control region are compatible with those in the signal region. As an additional uncertainty, we take into account a variation of the subtracted single top quark and $\mathrm{t\bar{t}}$ contributions, in which their expected contributions are scaled by the ratio of observed and expected events in the control regions.

5 Data analysis in the $\ell\bar{\ell}$ channel

In the $\ell\bar{\ell}$ channel, events are selected that contain exactly two oppositely charged well-identified leptons, one with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and the other $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. Events are rejected if they contain additional well-identified electrons or muons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. Furthermore, the invariant mass $m_{\ell\bar{\ell}}$ of the dilepton pair is required to be larger than 20 GeV, to suppress events from low mass dilepton resonances, and for same-flavor pairs to be outside of the Z boson

mass window, i.e., $76 < m_{\ell\bar{\ell}} < 106\,\text{GeV}$. To further suppress Z/γ^* background contributions, events in the ee and $\mu\mu$ channels are required to have $p_\text{T}^\text{miss} > 40\,\text{GeV}$. In all lepton flavor final states, among the selected jets with $p_\text{T} > 20\,\text{GeV}$, at least one is required to be b tagged and at least two are required to have $p_\text{T} > 30\,\text{GeV}$.

5.1 Kinematic reconstruction

Each selected event is reconstructed under the assumption that the final state consists of a top quark pair that decays into two leptonically decaying W bosons. A kinematic reconstruction algorithm [84] is applied to reconstruct the $\text{t}\bar{\text{t}}$ system, which consists of two steps: First, out of all jets in an event, two are identified as the b and $\bar{\text{b}}$ quark candidates. Second, these two candidates, together with the two leptons as well as p_T^miss, are used to determine the t and $\bar{\text{t}}$ quark four-momenta by applying mass constraints on the W bosons and top quarks, taking into account experimental resolutions via smearings.

First, to find the best assignment of jets to the b and $\bar{\text{b}}$ quarks, candidate pairs of jets are constructed depending of the number of b-tagged jets in the event. For events with two or more b-tagged jets, only those are considered as candidates, while for events with exactly one b-tagged jet, this jet is paired with all other jets in the event. The invariant masses of the visible top quark decay products $m_{\ell^+\text{b}}$ and $m_{\ell^-\bar{\text{b}}}$ are calculated for each $\text{b}\bar{\text{b}}$ candidate pair, and a likelihood is constructed as the product of the truth-level probability densities of the two invariant masses, evaluated from simulated events. The candidate pair that maximizes this likelihood is chosen for the next step of the reconstruction.

Second, a system of equations for the top quark four-momenta is constructed from energy and momentum conservation as well as additional constraints: the top quark and the W bosons are assumed to be on-shell, and the two neutrinos from the W boson decays are assumed to be the sole source of p_T^miss. These equations are solved for the neutrino momenta using an analytic method [85], and the top quark four-momenta calculated as the vectorial sum of the decay products. In case there are multiple real solutions to the equations, the one with the lowest reconstructed value of $m_{\text{t}\bar{\text{t}}}$ is used.

In many cases, this procedure on its own does not give real solutions for the $\text{t}\bar{\text{t}}$ system since it does not take into account the detector resolution. To remedy this, the system of equations is solved 100 times per event with random smearings applied to the energies and directions of the $\text{b}\bar{\text{b}}$ candidates and the leptons. These smearings are sampled, respectively, from distributions of the relative energy difference and angular distance between reconstructed and truth-level objects, as evaluated in simulated events. For all samplings that result in a real solution to the system of equations, weighted averages of the t and $\bar{\text{t}}$ quark four-momenta are computed over all samplings, with the weight given by the same likelihood based on $m_{\ell^+\text{b}}$ and $m_{\ell^-\bar{\text{b}}}$ as used for the $\text{b}\bar{\text{b}}$ quark candidate assignment. These averages are then considered as the final result of the reconstruction.

The performance of the $\text{t}\bar{\text{t}}$ reconstruction algorithm is studied using simulated SM $\text{t}\bar{\text{t}}$ events in the $\ell\bar{\ell}$ final state. The algorithm produces a solution for 90% of the events. In 78% of these events at least one b quark jet is correctly assigned, while in 61% both jets are correctly assigned. The precision of the achieved $m_{\text{t}\bar{\text{t}}}$ solution is assessed by its resolution, which is the standard deviation of its relative error to the generator-level $m_{\text{t}\bar{\text{t}}}$. The resolution is in the range of 18% (achieved at low generator-level $m_{\text{t}\bar{\text{t}}}$ values near the threshold region) to around 30% at high generator-level $m_{\text{t}\bar{\text{t}}}$ values above 1000 GeV. The average $m_{\text{t}\bar{\text{t}}}$ resolution is 23%.

The search is performed by building 3D templates using $m_{\text{t}\bar{\text{t}}}$ and two observables c_hel and c_han

that probe the spin correlations of the $t\bar{t}$ system. Spin correlation variables have been discussed in detail in the literature [15, 86–88], and we follow the coordinate system and sign convention of Ref. [86]. The observable c_{hel} (referred to as $\cos\varphi$ in Refs. [15, 86] and $-\cos\theta_{ab}$ in Ref. [87]) is defined as the scalar product $c_{\text{hel}} = \hat{\ell}_t^+ \cdot \hat{\ell}_{\bar{t}}^-$, where $\hat{\ell}_t^+$ and $\hat{\ell}_{\bar{t}}^-$ are the unit vectors of the momenta of the two leptons in their respective helicity frames, obtained by first boosting the leptons into the $t\bar{t}$ ZMF and then further boosting them into the rest frames of their parent top (anti)quarks. The observable c_{han} (identified with $-\cos\theta'_{ab}$ in Ref. [87]) is obtained by flipping the sign of the component parallel to the top quark direction (the $\hat{k}$ direction in Ref. [86]) for either $\hat{\ell}_t^+$ or $\hat{\ell}_{\bar{t}}^-$, and then calculating a similar scalar product. Both c_{hel} and c_{han} are sensitive to the spin and CP state of the $t\bar{t}$ system.

At the generator level and with no requirements on acceptance, the distributions of c_{hel} and c_{han}, integrated over the phase space of all other variables, follow a straight line, as shown in Fig. 2 for SM $t\bar{t}$, resonant Φ boson, and η_t production. For c_{hel}, the slope is maximally positive for a pseudoscalar resonance, mildly positive for the SM $t\bar{t}$ production, and mildly negative for a scalar resonance. On the other hand, for c_{han}, the slope is mildly positive for a pseudoscalar resonance, approximately flat for the SM $t\bar{t}$ production, and maximally negative for a scalar resonance. These features allow discriminating both between the signal and background processes and between the A and H states.

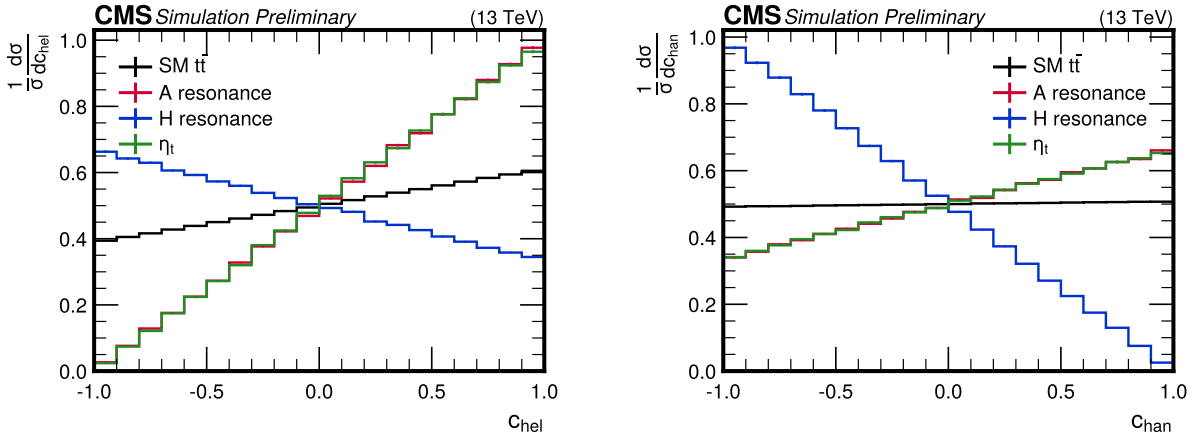

Figure 2: Normalized differential cross sections in the spin correlation observables c_{hel} (left) and c_{han} (right) at the parton level in the $\ell\bar{\ell}$ channel, with no requirements on acceptance, for SM $t\bar{t}$ (black), resonant A (red), resonant H (blue), η_t (green) production.

5.2 Background estimation

All background processes in the $\ell\bar{\ell}$ channel, namely $t\bar{t}$, single top quark, Z/γ^*, diboson, and $t\bar{t}V$ production, are estimated from simulated events. In the case of Z/γ^* production, the total yield of the simulation is corrected using data inside the Z boson mass window, which is removed in the main event selection, following a modified version of the procedure described in Ref. [89]. The same selection criteria except for the $m_{\ell\bar{\ell}}$ requirements are applied to the data inside the Z boson mass window. We assume that there, the Z/γ^* contribution is negligible in the $e\mu$ channel compared to the ee and $\mu\mu$ channels, and that other backgrounds contribute equally to the three channels up to a combinatoric factor. Thus, we can estimate the Z/γ^* contribution in data inside the Z boson mass window by subtracting the yield in the $e\mu$ channel from the yield in the ee and $\mu\mu$ channels, correcting for lepton reconstruction effects.

Next, to estimate the ratio of the Z/γ^* contribution inside and outside the Z boson mass win-

dow, denoted as $R_{\text{in/out}}$, we define a second sideband containing events with no b-tagged jets. The ratio in this region, $R^{\text{0b}}_{\text{in/out}}$, can be measured directly by comparing the Z/γ^* yields in data inside and outside the Z boson mass window. We then assume the ratio of ratios $R^{\geq\text{1b}}_{\text{in/out}}/R^{\text{0b}}_{\text{in/out}}$ in the regions with one or more b tags and zero b tags, respectively, to be well-described by simulation, which is a looser assumption compared to that in Ref. [89]. From this, we can infer $R^{\geq\text{1b}}_{\text{in/out}}$, and thus the total Z/γ^* yield outside the Z boson mass window, for events with one or more b tags, as used in the main selection.

The yield is separately estimated for the ee and $\mu\mu$ channels, and used to normalize the simulated Z/γ^* contribution. We quote the results in the form of ratios to the yields predicted by simulation in Table 1. For the eμ channel, where the Z/γ^* contribution is small, the geometric mean of the ratios to simulation is used.

Table 1: Derived scale factors for the Z/γ^* event yield in the different lepton flavor final states, and their statistical uncertainties.

	2016pre	2016post	2017	2018
ee	0.961 ± 0.010	0.965 ± 0.008	0.866 ± 0.006	0.883 ± 0.005
eμ	0.961 ± 0.007	0.968 ± 0.005	0.883 ± 0.004	0.892 ± 0.003
$\mu\mu$	0.961 ± 0.009	0.971 ± 0.006	0.901 ± 0.005	0.902 ± 0.004

6 Systematic uncertainties

Various sources of uncertainty affect the distributions of the observables used in this analysis, and are implemented as nuisance parameters in the binned maximum-likelihood fit described in Section 7. For each considered experimental and theoretical systematic effect, variations of the predicted signal and background distributions are evaluated. Uncertainties that affect only the normalization of a process are modeled using log-normal constraints as described in Section 4.2 of Ref. [90]. Gaussian constraints are imposed for all other uncertainties, which are referred to as shape uncertainties and can include a log-normal constrained variation of the overall normalization, by modifying the product of the event acceptance and the cross sections of the relevant processes. Unless stated otherwise, all uncertainties are evaluated on signal as well as background processes and treated as fully correlated among the processes, lepton channels, and analysis eras. The uncertainties are summarized in Table 2, and described in detail in the following.

The uncertainty in the jet p_{T} scale [35] is evaluated by varying the corresponding corrections within their uncertainties, resulting in a total of 17 nuisance parameters that correspond to the absolute and relative jet energy scales, calibration uncertainties in specific detector regions, p_{T} balance in dijet or Z/γ^* events used in the jet energy calibration, and flavor-dependent jet response split into one source for b quark jets and another for all other. Of these, 12 nuisance parameters affect individual analysis eras. The uncertainty in the jet p_{T} resolution measured in calibration data is propagated to the scale correction and smearing of the jet p_{T} resolution in simulation. An uncertainty in the unclustered component of $p_{\text{T}}^{\text{miss}}$ is computed by shifting the energies of PF candidates not clustered into jets with $p_{\text{T}} > 15\,\text{GeV}$ according to the energy resolution for each type of PF candidate [42].

Uncertainties in the b tagging efficiency scale factors applied to simulated events are evaluated by varying them within their respective uncertainties [36], independently for heavy-flavor (b and c quarks) and other (light quarks and gluon) jets. We assign 20 nuisance parameters for the

Table 2: The systematic uncertainties considered in the analysis, indicating the number of corresponding nuisance parameters (if not one) in the statistical model, the type (affecting only normalization or also the shape of the search templates), and the affected processes and analysis channels they are applicable to.

Uncertainty (# of parameters)	Type	Process	Channel
Jet p_T scale (17)	shape	all	all
Jet p_T resolution (4)	shape	all	all
Unclustered p_T^{miss} (4)	shape	all	all
b tagging heavy-flavor jets (20)	shape	all	all
b tagging light-flavor jets (5)	shape	all	all
Single-electron trigger	shape	all	ej
Single-muon trigger (5)	shape	all	μj
Dilepton triggers (12)	shape	all	ee, eμ, $\mu\mu$
Electron identification (2)	shape	all	ej, ee, eμ
Muon identification (10)	shape	all	μj, eμ, $\mu\mu$
ECAL L1 trigger inefficiency (3)	shape	all	all
Pileup	shape	all	all
Integrated luminosity (7)	norm.	all	all
Top quark Yukawa coupling	shape	SM $t\bar{t}$	all
EW correction scheme	shape	SM $t\bar{t}$	all
m_t	shape	SM $t\bar{t}$, Φ, η_t	all
ME μ_R (5)	shape	SM $t\bar{t}$, Φ, single top, Z/γ^*	all
ME μ_F (6)	shape	SM $t\bar{t}$, Φ, η_t, single top, Z/γ^*	all
PS ISR (6)	shape	SM $t\bar{t}$, Φ, η_t, single top, Z/γ^*	all
PS FSR (6)	shape	SM $t\bar{t}$, Φ, η_t, single top, Z/γ^*	all
Color reconnection (2)	shape	SM $t\bar{t}$	all
h_{damp}	shape	SM $t\bar{t}$	all
PDF (2)	shape	SM $t\bar{t}$	all
Single top quark normalization	norm.	Single top	all
EW+QCD normalization	norm.	Data-driven EW+QCD	ℓj
EW+QCD shape (20)	shape	Data-driven EW+QCD	ℓj
$t\bar{t}V$ normalization	norm.	$t\bar{t}V$	$\ell\bar{\ell}$
Z/γ^* normalization	norm.	Z/γ^*	$\ell\bar{\ell}$
Diboson normalization	norm.	Diboson	$\ell\bar{\ell}$
MC statistical (3920)	shape	all	all

heavy-flavor jet scale factors that correspond to the parton shower (PS) modeling, the presence of leptons within the jet, the jet p_T scale, the number of pileup interactions, and differences between different scale factor estimation methods. Of these, 4 nuisance parameters affect individual analysis eras. For the light-flavor jet scale factors, 5 nuisance parameters are assigned, of which 4 affect individual analysis eras.

Uncertainties in the trigger, electron identification, and muon identification scale factors are considered [39, 41], including also effects from the isolation requirement and the track reconstruction at the trigger level. For the single-muon trigger and muon identification scale factors,

each uncertainty component is further split into statistical components that are uncorrelated across analysis eras and a correlated systematic component. The effects of the inefficiency caused by the gradual shift in the timing of the inputs of the ECAL L1 trigger [29] are considered by assigning one nuisance parameter each to the 2016pre, 2016post, and 2017 analysis eras.

The effective inelastic proton-proton cross section used for pileup reweighting in the simulation is varied by 4.6% from its nominal value. Additionally, the uncertainty in the integrated luminosity amounts to 1.6% [21–23] and affects the normalization of all simulated processes, and is split into 7 nuisance parameters with different correlation assumptions between the analysis eras.

The prediction of the SM $\mathrm{t\bar{t}}$ production is affected by various sources of theoretical uncertainty. The computation of the NLO EW correction, discussed in Section 3, depends on the value of the SM top quark Yukawa coupling through interference with diagrams containing virtual SM Higgs bosons. An uncertainty in the coupling is considered by varying its value by $1.00^{+0.11}_{-0.12}$, where the range is given by the experimental measurement reported in Ref. [91]. Furthermore, the uncertainty in the application scheme of the NLO EW corrections when combined with NNLO QCD corrections is considered by taking the difference between the multiplicative and additive approaches, as recommended in Ref. [67]. The uncertainty in m_{t} is considered by shifting its value in simulation by $\pm 3\,\mathrm{GeV}$, and the induced variations are then rescaled by a factor of $1/3$ to emulate a more realistic top quark mass uncertainty of 1 GeV [92]. The effect of the choice of μ_{R} and μ_{F} in the ME calculation is evaluated by varying these scales independently by a factor of 2 and $1/2$. The effects of the m_{t}, μ_{R}, and μ_{F} variations on the acceptance and shape of the search templates are considered at NLO accuracy, while the effects on the overall SM $\mathrm{t\bar{t}}$ normalization is considered at NNLO+NNLL accuracy [62, 93]. Decoupling the theoretical nuisance parameters based on their effects—one each for the acceptance and shape, and one additional parameter for the overall SM $\mathrm{t\bar{t}}$ normalization—does not alter the conclusions of this analysis.

The scales used to evaluate α_{S} in the PS simulation of initial- and final-state radiation (ISR and FSR) are also varied independently by a factor of 2 in each direction. The effect of the uncertainties in the underlying event tune is estimated by varying the parameters of the CP5 underlying event tune [47]. Two uncertainties are assigned for the color reconnection model, with one based on the "QCD-inspired" model [94], and the other by switching on the early resonance decay option in PYTHIA 8.240 [95].

The uncertainty in the matching scale between the ME and the PS is evaluated by varying the POWHEG parameter h_{damp}, which controls the suppression of radiation of additional high-p_{T} jets. The nominal value of h_{damp} in the simulation and its variations are $1.58^{+0.66}_{-0.59}\,m_{\mathrm{t}}$ [96]. The uncertainty arising from the choice of the PDF set is evaluated by reweighting the simulated $\mathrm{t\bar{t}}$ events using 100 replicas of the NNPDF3.1 set. A principal component analysis is performed on the variations from the PDF replicas to construct base variations in the space of the predicted event yields in each bin of the search templates, from which the one with the largest eigenvalue is used as the PDF uncertainty. The second largest eigenvalue is found to be almost two orders of magnitude smaller than the largest one, thus the base variations corresponding to it and smaller eigenvalues are not considered. The uncertainty in the α_{S} parameter used in the PDF set forms a second independent PDF variation uncertainty.

The μ_{R} and μ_{F} scale uncertainties in the Φ signal simulation are treated independently for the resonance and interference components. Compared to the alternative of varying the scales for

the two components simultaneously, we found this to be the more conservative option. The effect on the acceptance and shapes of the search templates is considered at LO accuracy, while the effect on signal cross section is considered at NNLO accuracy. The scales used in the PS simulation of ISR and FSR are also varied independently by a factor of 2 in each direction and are treated independently for the resonant and interference components.

The uncertainty in m_t for the signal is considered by varying its value in simulation by $\pm 1\,\text{GeV}$. Its effect on acceptance, shape, and cross section is considered in the same way as μ_R and μ_F variations. Given that this is a variation on the same physical parameter, it is treated as fully correlated across all signal and background processes. Other theoretical uncertainties in the signal, such as the PDF, are neglected as they are small compared to those already considered.

The η_t signal simulation considers μ_F, ISR, FSR, and m_t uncertainties, affecting only acceptance and shape. They are handled identically to the corresponding variations in the Φ signal simulation, except for the absence of variations on the overall normalization, which is always taken to be freely floating in this analysis. Since the used model describes effective η_t production via a contact interaction, without the emission of extra partons at the LO ME level, the model encodes no dependence on α_S. Therefore, μ_R variations have no effect on the η_t prediction.

The μ_R, μ_F, ISR, and FSR scale uncertainties are also independently considered for the Z/γ^* and single top quark production processes. For these processes, the μ_R and μ_F uncertainties affect only acceptance and shape, not normalization.

The expected yields for most of the non-$t\bar{t}$ background processes are derived using theoretical predictions for the cross sections at NLO or higher accuracy. The uncertainties assumed in the normalization of these processes are conservative and always exceed those of the corresponding theoretical computations. For single top quark production, we assign an uncertainty of 15%, based on relevant cross section measurements [97–99]. In the single-lepton channels, the normalization uncertainty of the EW+QCD background estimate evaluated from control samples in data is taken to be 50%, and shape uncertainties as described in Section 4 are considered as well. The uncertainties corresponding to the change in shape induced by varying the b tagging requirements are considered separately for the single-lepton channels, but correlated across analysis eras. Statistical uncertainties in the $t\bar{t}$ and single top quark subtraction are considered separately for each channel and era. In the $\ell\bar{\ell}$ channels, the uncertainty in the $t\bar{t}V$ production is taken to be 30% [100, 101]. The uncertainty of the Z/γ^* production is taken to be 5% [102]. To account for the fact that this search probes a restricted region of the phase space of the corresponding processes, we assign a normalization uncertainty of 30% for diboson production, which has little impact on the overall sensitivity due to the small contribution of these processes.

The nominal background prediction is affected by the limited size of the simulated MC event samples. This statistical uncertainty is evaluated using the "light" Barlow–Beeston method [103], by introducing one additional nuisance parameter for every bin of the search distributions. These parameters are uncorrelated across all channels and analysis eras.

Several systematic variations, most notably those constructed from dedicated MC samples, are affected by statistical fluctuations. We suppress these fluctuations with the smoothing procedure described in Ref. [24].

In general, the relative importance of different systematic uncertainties depends greatly on the signal hypothesis, especially the mass of the scalar bosons. Close to the $t\bar{t}$ production threshold, the variations in the value of the top quark Yukawa coupling and m_t become important, while for larger m_Φ the PDF, μ_R, and μ_F variations in the SM $t\bar{t}$ background become dominant.

Among the experimental uncertainty sources, those due to jet p_{T} scale and heavy-flavor jet tagging are important. In addition, MC statistical uncertainties, when grouped together, often outweigh every other individual uncertainty.

7 Statistical analysis

To evaluate the consistency of the observed data with the presence of a signal, we perform a statistical analysis using the search distributions described in Sections 4–5. The ℓj and $\ell\bar{\ell}$ final states do not overlap as they correspond to orthogonal lepton selection criteria.

The statistical model is defined by the likelihood function

$$\begin{aligned} L(\boldsymbol{p}_\Phi, \mu(\eta_{\mathrm{t}}), \boldsymbol{\nu}) &= \left(\prod_i \frac{\lambda_i(\boldsymbol{p}_\Phi, \mu(\eta_{\mathrm{t}}), \boldsymbol{\nu})^{n_i}}{n_i!} \mathrm{e}^{-\lambda_i(\boldsymbol{p}_\Phi, \mu(\eta_{\mathrm{t}}), \boldsymbol{\nu})} \right) G(\boldsymbol{\nu}), \\ \lambda_i(\boldsymbol{p}_\Phi, \mu(\eta_{\mathrm{t}}), \boldsymbol{\nu}) &= S_i^{\Phi}(\boldsymbol{p}_\Phi, \boldsymbol{\nu}) + S_i^{\eta_{\mathrm{t}}}(\mu(\eta_{\mathrm{t}}), \boldsymbol{\nu}) + B_i(\boldsymbol{\nu}), \end{aligned} \tag{2}$$

with B_i denoting the combined background yield in a given bin i, S_i^{Φ} the Φ signal yield dependent on signal model parameters $\boldsymbol{p}_\Phi$, $S_i^{\eta_{\mathrm{t}}}$ the yield of the η_{t} contribution dependent on the signal strength $\mu(\eta_{\mathrm{t}})$, $\boldsymbol{\nu}$ the vector of nuisance parameters (on which the signal and background yields generally depend), and n_i the observed yield. The external constraints on the nuisance parameters are taken into account in this likelihood via a product of corresponding probability density functions, $G(\boldsymbol{\nu})$.

The Φ signal yield S_i^{Φ} is given by

$$S_i^{\Phi}(\boldsymbol{p}_\Phi, \boldsymbol{\nu}) = \sum_{\Phi=\mathrm{A,H}} \left(g_{\Phi \mathrm{t}\bar{\mathrm{t}}}^4 \, s_{R,i}^{\Phi}(m_\Phi, \Gamma_\Phi, \boldsymbol{\nu}) + g_{\Phi \mathrm{t}\bar{\mathrm{t}}}^2 \, s_{I,i}^{\Phi}(m_\Phi, \Gamma_\Phi, \boldsymbol{\nu}) \right), \tag{3}$$

where $s_{R,i}^{\Phi}$ and $s_{I,i}^{\Phi}$ are the yields for the resonant and interference part, respectively. The vector $\boldsymbol{p}_\Phi$ represents the signal model parameters and comprises the Φ boson mass m_Φ, width Γ_Φ, and $g_{\Phi \mathrm{t}\bar{\mathrm{t}}}$. Equation (3) is kept generic by including contributions from both A and H. Since there is no interference between them, the corresponding signal distributions are trivially added together.

The yield of the η_{t} contribution $S_i^{\eta_{\mathrm{t}}}$ is given by

$$S_i^{\eta_{\mathrm{t}}}(\mu(\eta_{\mathrm{t}}), \boldsymbol{\nu}) = \mu(\eta_{\mathrm{t}}) \, s_i^{\eta_{\mathrm{t}}}(\boldsymbol{\nu}), \tag{4}$$

where $s_i^{\eta_{\mathrm{t}}}$ are the predicted η_{t} signal yields and $\mu(\eta_{\mathrm{t}})$ is the signal strength modifier, which is allowed to float freely when η_{t} is part of either the signal or background models. When η_{t} is considered as a signal, $g_{\Phi \mathrm{t}\bar{\mathrm{t}}}$ is always set to zero. When η_{t} is considered as a background, i.e., when Φ is the signal, the interference between the A boson and η_{t} is not taken into account. Because of the freely floating η_{t} signal strength and the mass splitting between η_{t} and the lightest A considered in this analysis, this is expected to have negligible impact.

The background-only model is constructed by setting $g_{\Phi \mathrm{t}\bar{\mathrm{t}}} = 0$ and/or $\mu(\eta_{\mathrm{t}}) = 0$, as appropriate. The compatibility between data and a given hypothesis is determined by performing profiled likelihood ratio or fully frequentist scans over the parameters of the signal models in different scenarios, as described below.

7.1 Methodology for single Φ interpretation

In the single Φ interpretation, constraints on the coupling strength modifier $g_{\Phi t\bar{t}}$ are derived as a function of m_Φ for fixed relative widths, separately for A and H. This is done while setting the coupling modifier for the other CP state in Eq. (2) to zero, thus excluding it from the statistical model. The scan is performed for the mass and relative width values listed in Section 3. Coupling strength values up to 3 are probed to guarantee that the amplitudes preserve perturbative unitarity for all calculations, in accordance with the lower bound $\tan\beta = 1/g_{At\bar{t}} \gtrsim 0.3$ given in Ref. [4] in the context of 2HDMs.

A variant of the LHC profile likelihood ratio test statistic $\tilde{q}_\mu$ equivalent to those described in Refs. [104, 105] is utilized:

$$\tilde{q}_\mu(\boldsymbol{p}_\Phi) = -2\ln\frac{L(\mu, \boldsymbol{p}_\Phi, \hat{\boldsymbol{\nu}}_{\mu,\boldsymbol{p}_\Phi})}{L(\hat{\mu}, \boldsymbol{p}_\Phi, \hat{\boldsymbol{\nu}}_{\hat{\mu},\boldsymbol{p}_\Phi})}, \quad 0 \leq \hat{\mu} \leq \mu. \tag{5}$$

Because the Φ signal scales nonlinearly with the coupling modifiers $g_{\Phi t\bar{t}}$, we introduce an auxiliary overall signal strength modifier μ in terms of which the test statistic is expressed, in the same way as in Ref. [24]. This facilitates testing different Φ signal hypotheses in a computationally efficient way. The auxiliary parameter scales the overall Φ signal yield in Eq. (3), keeping the other parameters in $\boldsymbol{p}_\Phi$ fixed . The likelihood in the numerator is maximized with respect to the nuisance parameters, and $\hat{\boldsymbol{\nu}}_{\mu,\boldsymbol{p}_\Phi}$ denotes the vector of their values at the maximum for a given $\boldsymbol{p}_\Phi$. Depending on the scenario considered, the η_t signal strength is kept freely floating and treated as part of the nuisance parameters, or it is fixed to $\mu(\eta_t) = 0$ in both numerator and denominator. A similar notation is used in the denominator, where the likelihood is maximized with respect to both μ and $\boldsymbol{\nu}$, under the additional constraint $0 \leq \hat{\mu} \leq \mu$. The requirement $\hat{\mu} \geq 0$ excludes cases in which the shape of the overall BSM contribution gets flipped, resulting in a qualitatively different effect from what is targeted in this search. The condition $\hat{\mu} \leq \mu$ prevents the exclusion of a signal hypothesis if the data are more compatible with a model that predicts the BSM contribution of a similar shape but a larger overall size.

For each signal hypothesis, we perform a test according to the CL_s criterion [106, 107]. An asymptotic approximation [104] is employed to efficiently construct the distributions of the adopted test statistic. We exclude a configuration $\boldsymbol{p}_\Phi$ at 95% confidence level (CL) if the CL_s value computed for $\mu = 1$, which reproduces the nominal signal expectation, is smaller than 0.05.

7.2 Methodology for A+H interpretation

In the A+H interpretation, we consider the more general case where two Φ states exist at the same time. We confine ourselves to the case with exactly one A and exactly one H, i.e., the case considered in 2HDMs [4]. Constraints in the $g_{\Phi t\bar{t}}$ plane are set using the following test statistic:

$$\tilde{q}_{\boldsymbol{p}_\Phi} = -2\ln\frac{L(\boldsymbol{p}_\Phi, \hat{\boldsymbol{\nu}}_{\boldsymbol{p}_\Phi})}{L(\hat{\boldsymbol{p}}_\Phi, \hat{\boldsymbol{\nu}}_{\hat{\boldsymbol{p}}_\Phi})}, \tag{6}$$

expressed directly in terms of $g_{\Phi t\bar{t}}$. In contrast to the single A/H interpretation, the asymptotic approximation is not exploited, rendering the auxiliary parameter μ unnecessary. Furthermore, $\mu(\eta_t)$ is always considered as part of the nuisance parameters $\boldsymbol{\nu}$.

For each $g_{\Phi t\bar{t}}$ configuration under consideration, its compatibility with the data is evaluated with the Feldman–Cousins prescription [108]. An iterative procedure is applied to reduce the

number of points for which the test statistic needs to be evaluated. An initially sparse grid of $g_{\Phi t\bar{t}}$ configurations are evaluated and refined around the region of the exclusion contour boundary at a given CL. The procedure is repeated until the distance of two neighboring $g_{\Phi t\bar{t}}$ configurations in the plane is below granularity threshold.

8 Results

Various models are tested and presented as results. Section 8.1 shows the fits for a single Φ boson with the background prediced only by the pQCD modeling, showing an excess around the $t\bar{t}$ production threshold. Upon confirming the pseudoscalar nature of the excess (see Appendix A), a fit using the simplified color-singlet $t\bar{t}$ bound state model η_t is performed in Section 8.2, showing a good description of the data. Consequently, the results for additional Φ contributions with η_t included in the background are presented in Section 8.3, and the data are found to be in good agreement with this background prediction without the need for a Φ boson contribution. Finally, Section 8.4 shows fully frequentist exclusion contours for the simultaneous presence of A and H for a few examples of A/H masses and widths.

8.1 Single Φ interpretation with the pQCD-only background model

When including only the pQCD predictions included in the background model, we observe a deviation from the background expectation. This discrepancy, occurring mainly at low $m_{t\bar{t}}$, is shown with prefit normalizations in the upper two panels of Figs. 3–5 for the three channels considered.

We test whether this observed prefit discrepancy can be explained within the single Φ interpretation. The 95% CL exclusion limits, derived as described in Section 7.1, are shown in Figs. 6–7. The expected constraints on $g_{\Phi t\bar{t}}$ evolve in accordance with the signal cross section, as A/H mass and width values increase. The relatively sharper decline in sensitivity for A/H with masses between 700–900 GeV and larger widths is due to the cancellation between cross sections of the resonance and interference signal components.

The obtained constraints on $g_{\Phi t\bar{t}}$ improve with respect to the previous results presented in Ref. [24], which exploited a smaller data set with a simpler analysis strategy. In the ℓj channel, the addition of the three jets category increased the statistical power of the analysis. In the $\ell\bar{\ell}$ channel, the addition of c_{han} as a search variable improved the search sensitivity, particularly for H. These improvements significantly extended the range of excluded $g_{\Phi t\bar{t}}$ values from the previous result, for most of the CP, mass, and width values considered in this analysis.

A significant deviation with respect to the background expectation is observed, for both A and H interpretations at low m_Φ values. The data is better described by the A signal hypothesis, where in general narrower widths are preferred. We investigate the structure of the deviation in greater detail in Appendix A. The observed local significance of the best fit point for the $A(365, 2\%)$ signal hypothesis over the background-only hypothesis is beyond five standard deviations. The ratio of the observed to the postfit predicted distributions is shown in the third panel of Figs. 3–5.

8.2 Cross section measurement for the η_t signal model

Instead of the single A/H interpretation, we also evaluate whether the observed prefit discrepancy can be explained within the SM by a color-singlet $t\bar{t}$ bound state, as provided by our η_t signal model specified in Section 3. The maximum likelihood estimate of the η_t produc-

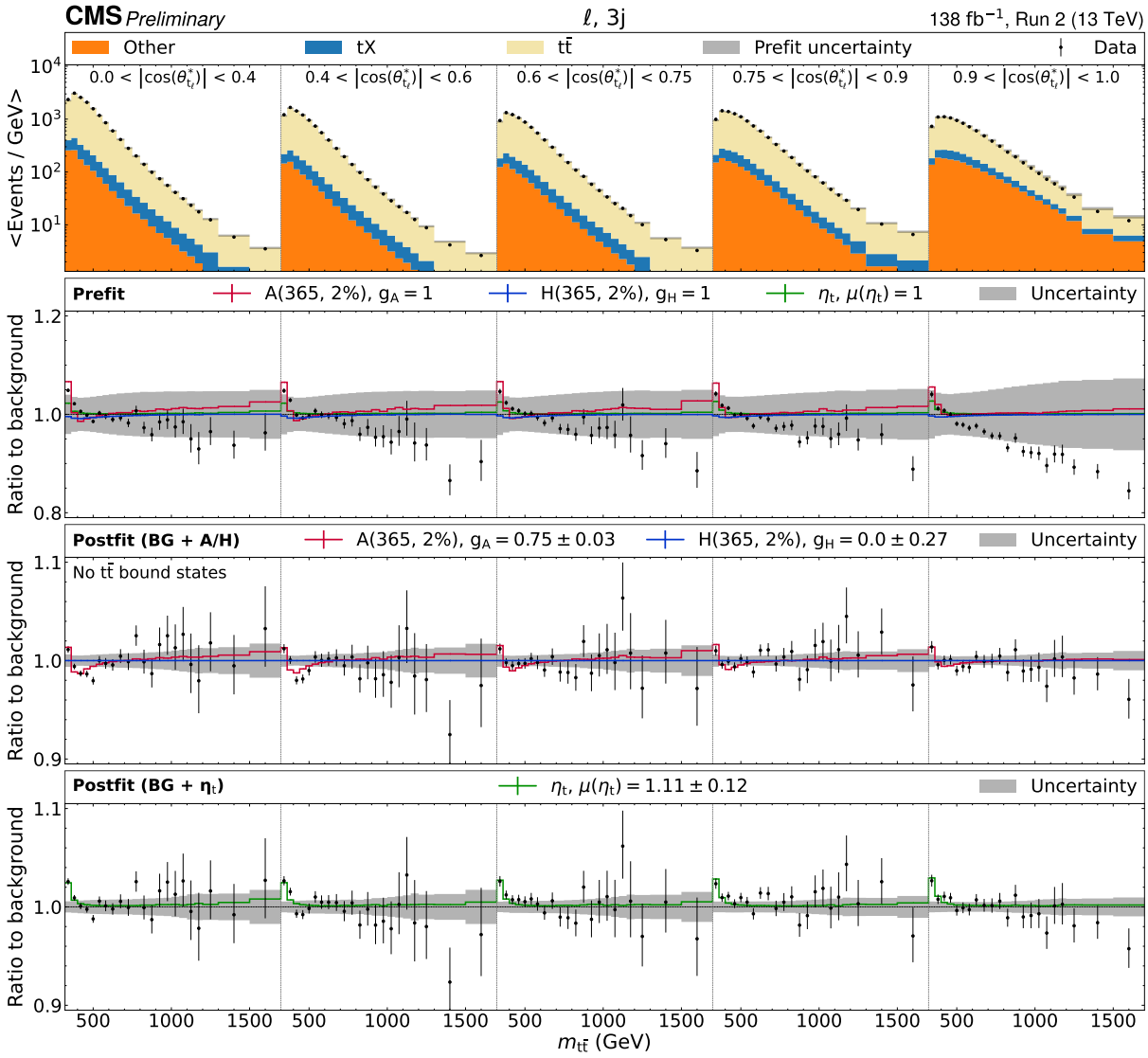

Figure 3: Observed and expected $m_{\mathrm{t\bar{t}}}$ distribution in $|\cos\theta^{*}_{\mathrm{t}_\ell}|$ bins, shown for the ℓ, 3j channel summed over lepton flavors and analysis eras. In the first panel, the data (points with statistical error bars) and predicted pQCD-only background (colored histograms) are compared before the fit to the data, and the corresponding prefit uncertainty is shown with a gray band. In the second panel, the ratio of the data to the sum of the pQCD-only background is shown, and three signal hypotheses ($\mathrm{A}(365, 2\%)$, $\mathrm{H}(365, 2\%)$, and η_t) are overlaid for illustration. In the third and fourth panels, the ratio is shown with the best fit normalization applied for two different interpretations: either fitting only the A/H signal with no η_t considered in the background (third panel), or fitting only η_t (fourth panel). In both cases, the gray band shows the postfit uncertainty, and the respective signals are overlaid with their best fit model parameters.

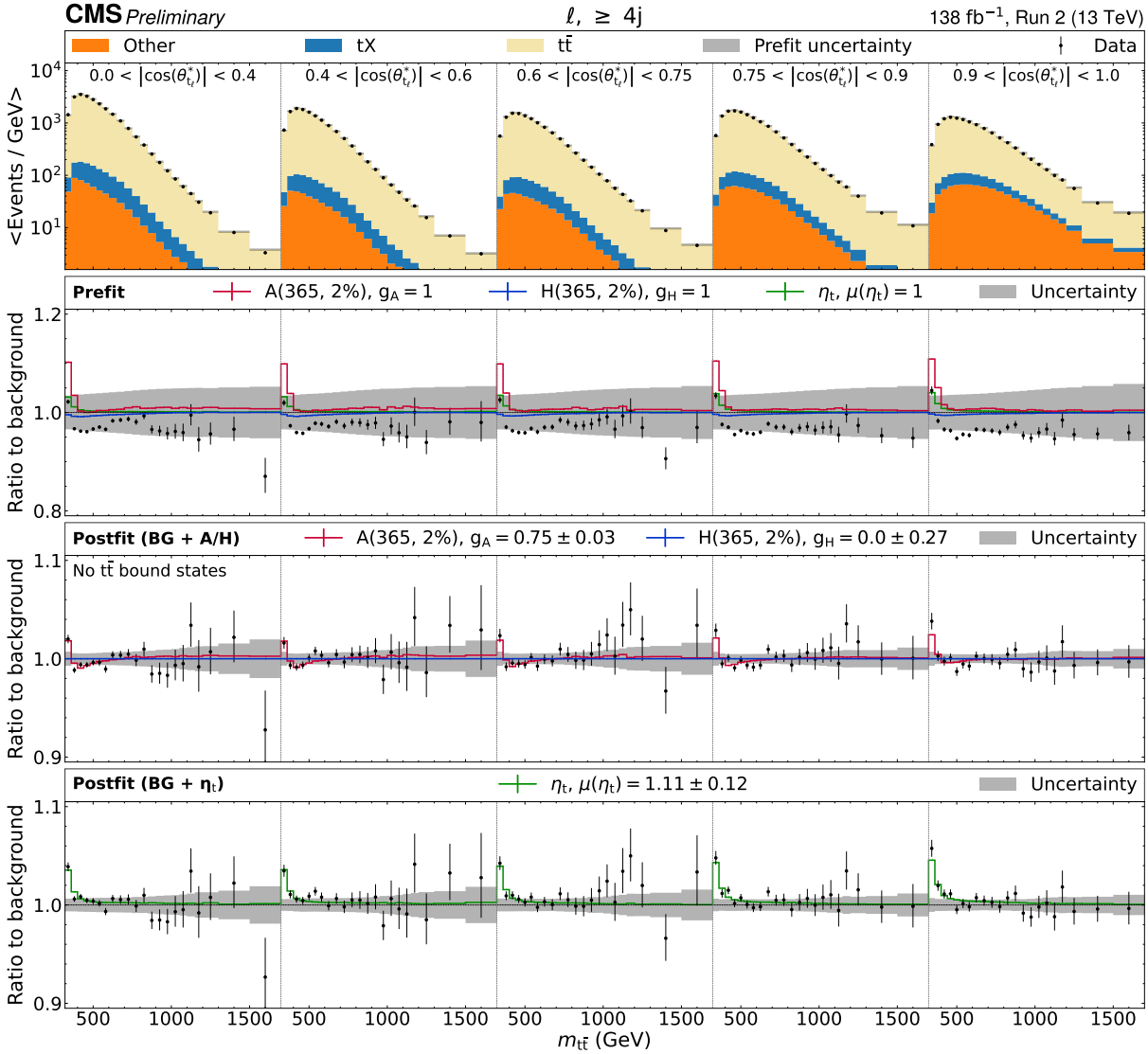

Figure 4: Observed and expected $m_{t\bar{t}}$ distribution in $|\cos\theta^*_{t_\ell}|$ bins, shown for the ℓ, $\geq$ 4j channel summed over lepton flavors and analysis eras. Notations as in Fig. 3.

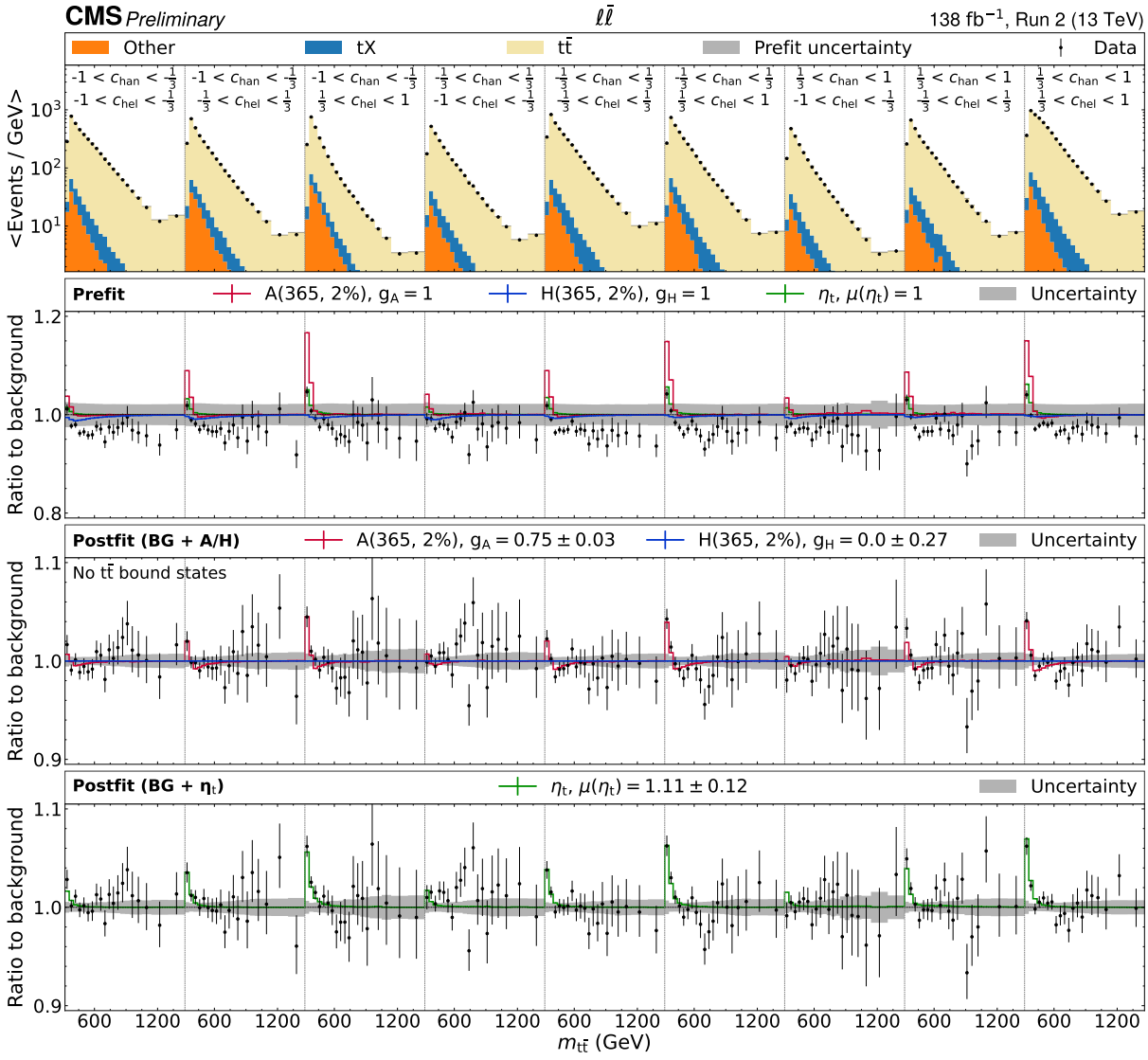

Figure 5: Observed and expected $m_{t\bar{t}}$ distribution in c_{hel} and c_{han} bins, shown for the $\ell\bar{\ell}$ channel summed over lepton flavors and analysis eras. Notations as in Fig. 3.

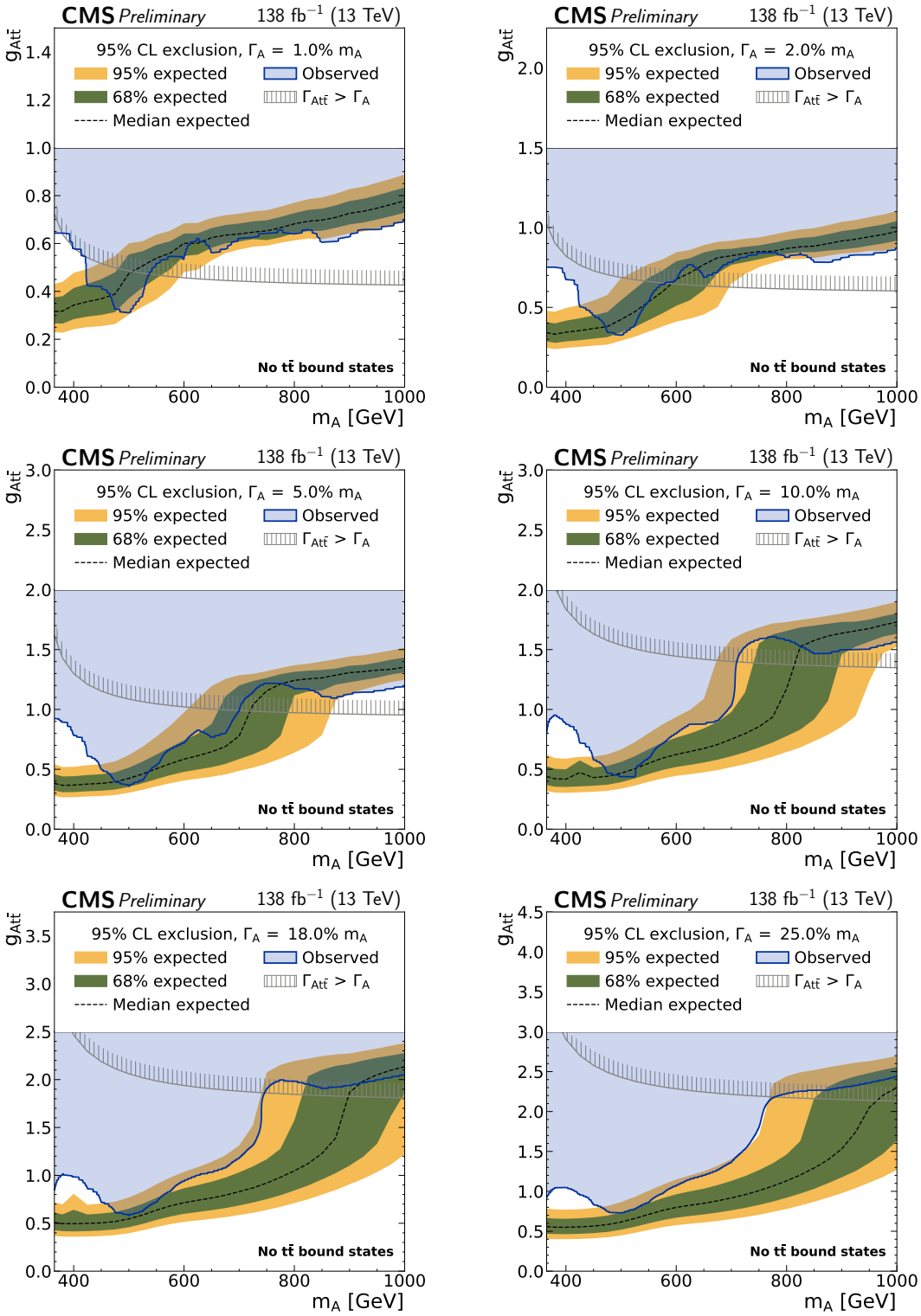

Figure 6: Model-independent constraints on $g_{At\bar{t}}$ as a function of the A mass for relative widths of 1, 2, 5, 10, 18, and 25%. The observed constraints are indicated by the blue shaded area. The inner green band and the outer yellow band indicate the regions containing 68 and 95%, respectively, of the distribution of constraints expected under the background-only hypothesis. The unphysical region of phase space in which the partial width $\Gamma_{A\to t\bar{t}}$ becomes larger than the A total width is indicated by the hatched line.

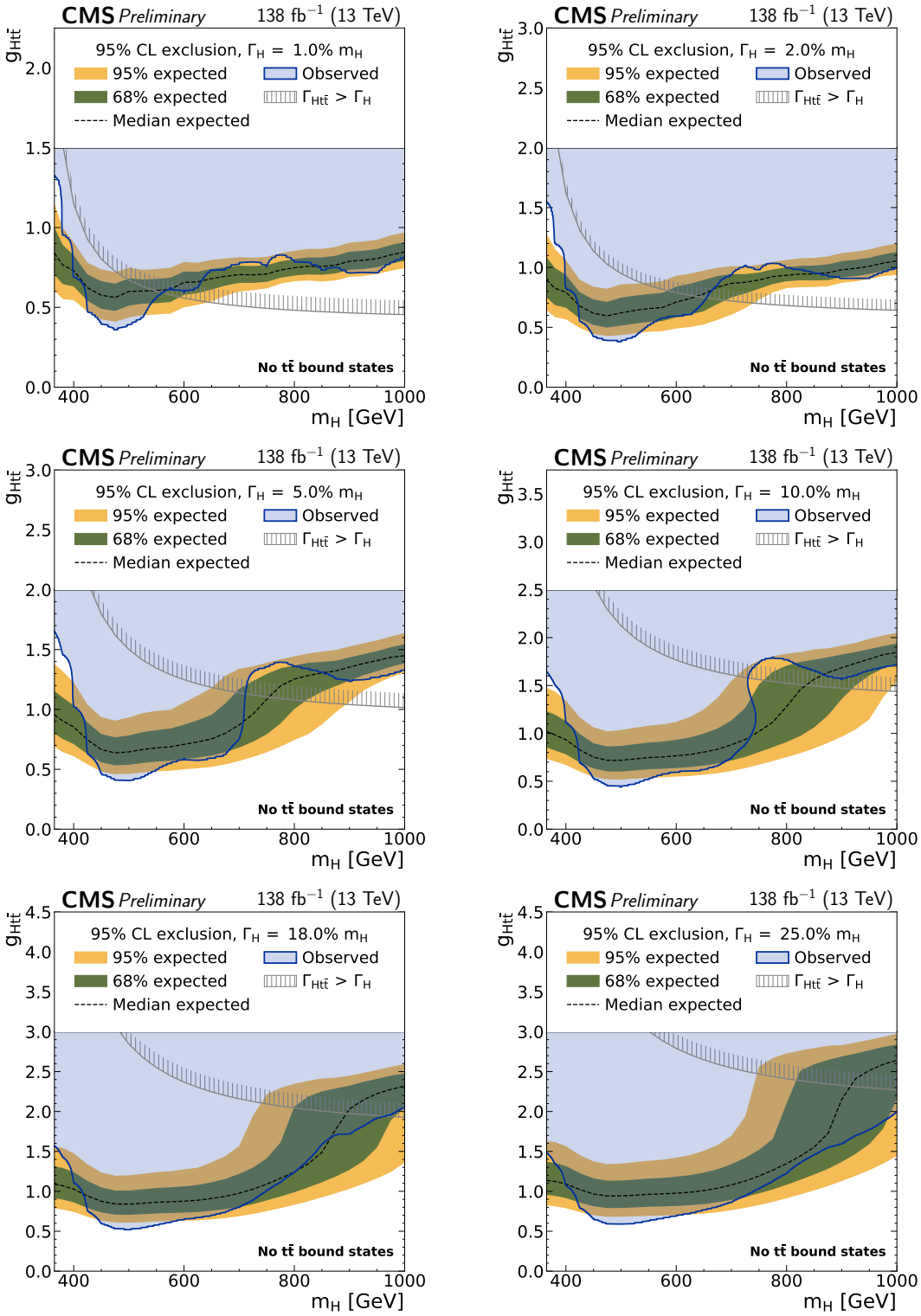

Figure 7: Model-independent constraints on $g_{Ht\bar{t}}$ as a function of the H mass, for relative widths of 1, 2, 5, 10, 18, and 25%. The observed constraints are indicated by the blue shaded area. The inner green band and the outer yellow band indicate the regions containing 68 and 95%, respectively, of the distribution of constraints expected under the background-only hypothesis. The unphysical region of phase space in which the partial width $\Gamma_{H\to t\bar{t}}$ becomes larger than the H total width is indicated by the hatched line.

tion cross section, combining the ℓj and $\ell\bar{\ell}$ final states, is extracted to be $\sigma(\eta_t) = 7.1\,\text{pb}$. The uncertainty of the extracted cross section is about 11%, assuming the adequacy of the background model and their associated uncertainties (see Secs. 3 and 6)—in particular its use of resonant $t\bar{t}$ production at NLO accuracy in pQCD reweighted double-differentially to NNLO accuracy in pQCD and NLO accuracy in EW. The robustness of this assumption is examined in Appendix B. The extracted cross section is in agreement with the theory prediction of 6.43 pb given in Ref. [20], which was obtained by fitting the results of a nonrelativistic QCD calculation from Ref. [57] with the pQCD contribution subtracted.

The pulls and impacts (as defined in Ref. [90]) of the nuisance parameters with the highest impact on the η_t cross section are shown in Fig. 8. The dominant contributions arise from modeling uncertainties, particularly those affecting the $m_{t\bar{t}}$ threshold region. The correlation matrix between the η_t signal strength and the nuisance parameters is shown in Fig. 9. There is some anticorrelation between the η_t signal strength and the top quark Yukawa coupling, as also shown by the corresponding nuisance parameter impact. This is because both effects induce an enhancement in the $t\bar{t}$ threshold region. However, the degeneracy is mild as the shape effects of the top quark Yukawa coupling do not vary strongly as a function of the angular observables employed in this analysis. For η_t, in contrast, the pseudoscalar nature results in a

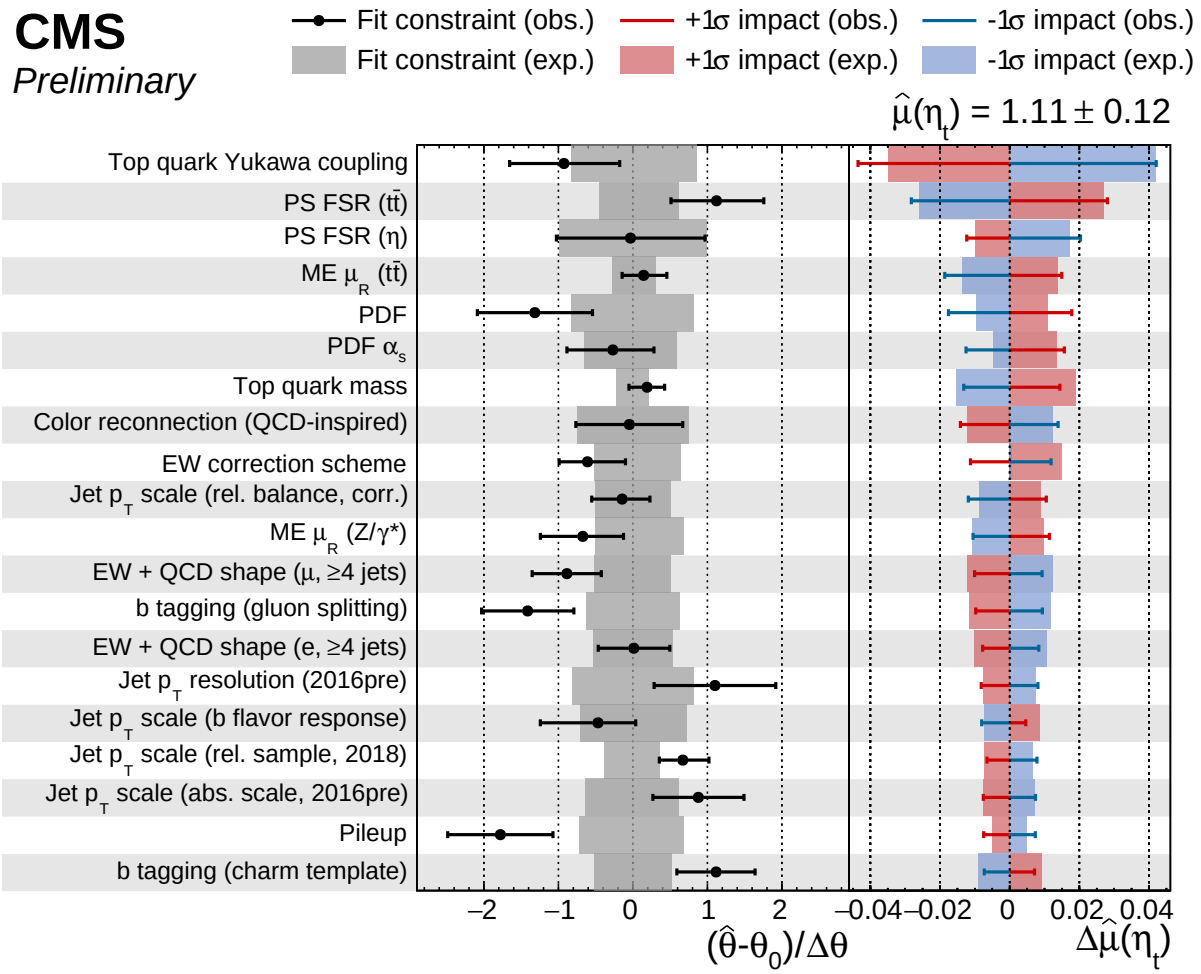

Figure 8: Pulls and impacts of the nuisance parameters with the largest impact in the η_t interpretation. A detailed description of the nuisance parameters relating to the jet p_T scale is provided in Ref. [35].

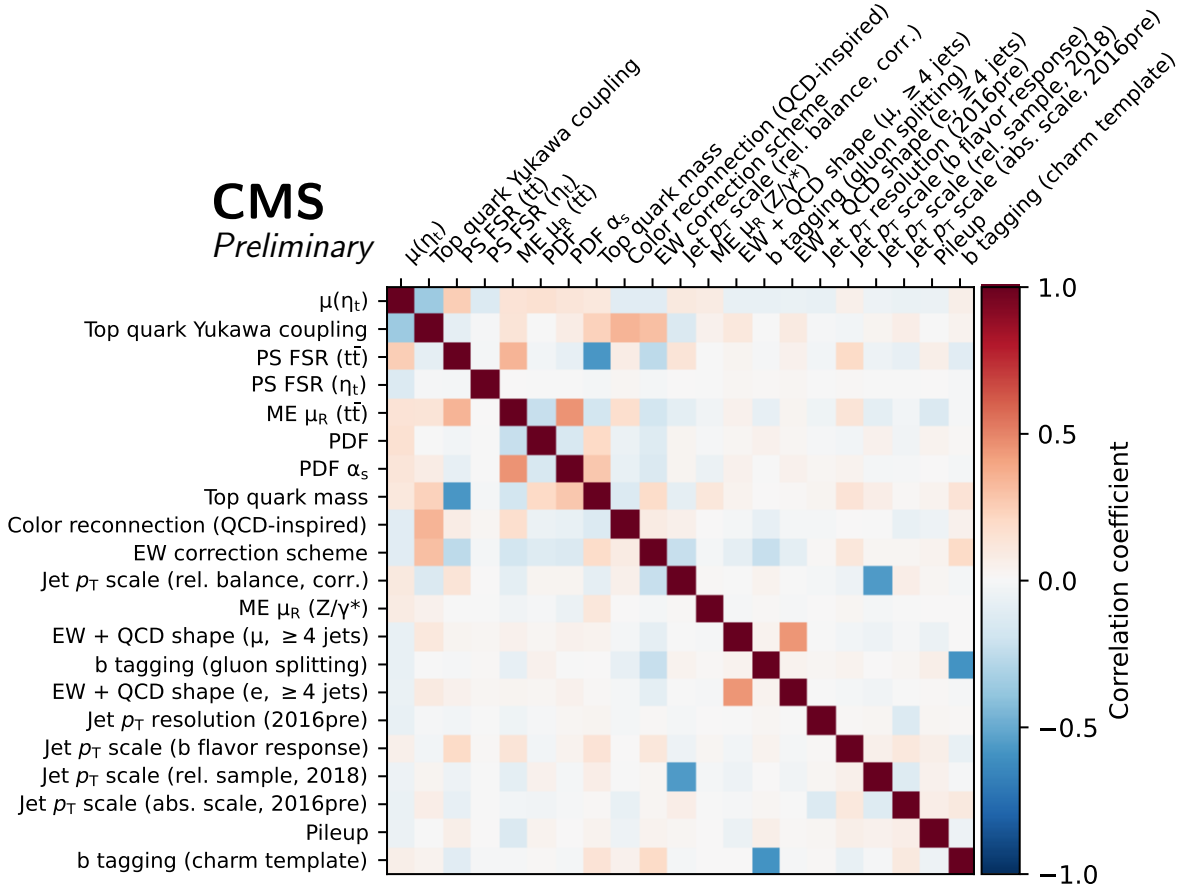

Figure 9: Correlation matrix between the η_t signal strength and nuisance parameters with the largest impact in the η_t interpretation. A detailed description of the nuisance parameters relating to the jet p_T scale is provided in Ref. [35].

strong angular dependence of the predicted signal contribution, in particular on c_{hel}.

The ratio of observed to the postfit predicted distributions is shown in the fourth panels of Figs. 3–5. By adding the η_t signal contribution to the pQCD-only SM background model, good agreement with the observed data is achieved.

8.3 Single Φ interpretation with the pQCD and η_t background model

After adding the η_t contribution to the pQCD-only SM background model with a freely floating η_t normalization nuisance parameter, as discussed in Section 7, we repeat the single Φ interpretation. The obtained 95% CL exclusion limits are shown in Figs. 10–11. Compared to Figs. 6–7, where the pQCD-only background model is assumed, the expected and observed constraints are in good agreement, indicating that the data is well described by the background model without the need for a BSM contribution as provided by a single Φ boson.

8.4 The A+H interpretation with the pQCD and η_t background model

Although the results in Section 8.3 suggest that there are no additional Φ bosons present in the $t\bar{t}$ spectrum beyond the deviation at the $t\bar{t}$ threshold, a more precise investigation is necessary to assess the exclusion in the case multiple Φ bosons are present, as their contributions may partically cancel out as a result of the peak-dip structure of the signal. We explore this sce-

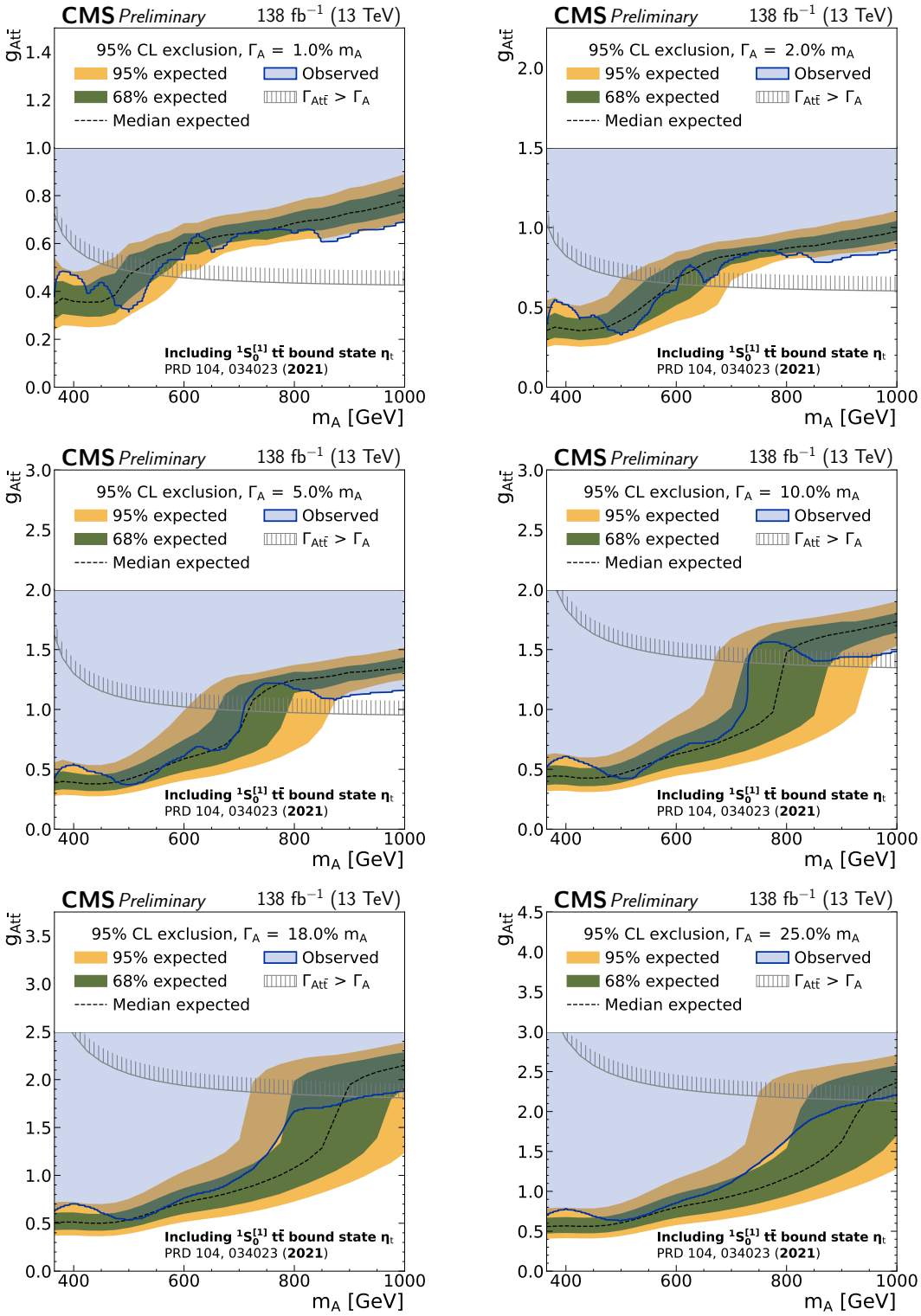

Figure 10: Model-independent constraints on $g_{At\bar{t}}$ as a function of the A mass for relative widths of 1, 2, 5, 10, 18, and 25%. The observed constraints are indicated by the blue shaded area. The inner green band and the outer yellow band indicate the regions containing 68 and 95%, respectively, of the distribution of constraints expected under the background-only hypothesis. The unphysical region of phase space in which the partial width $\Gamma_{A\to t\bar{t}}$ becomes larger than the A total width is indicated by the hatched line.

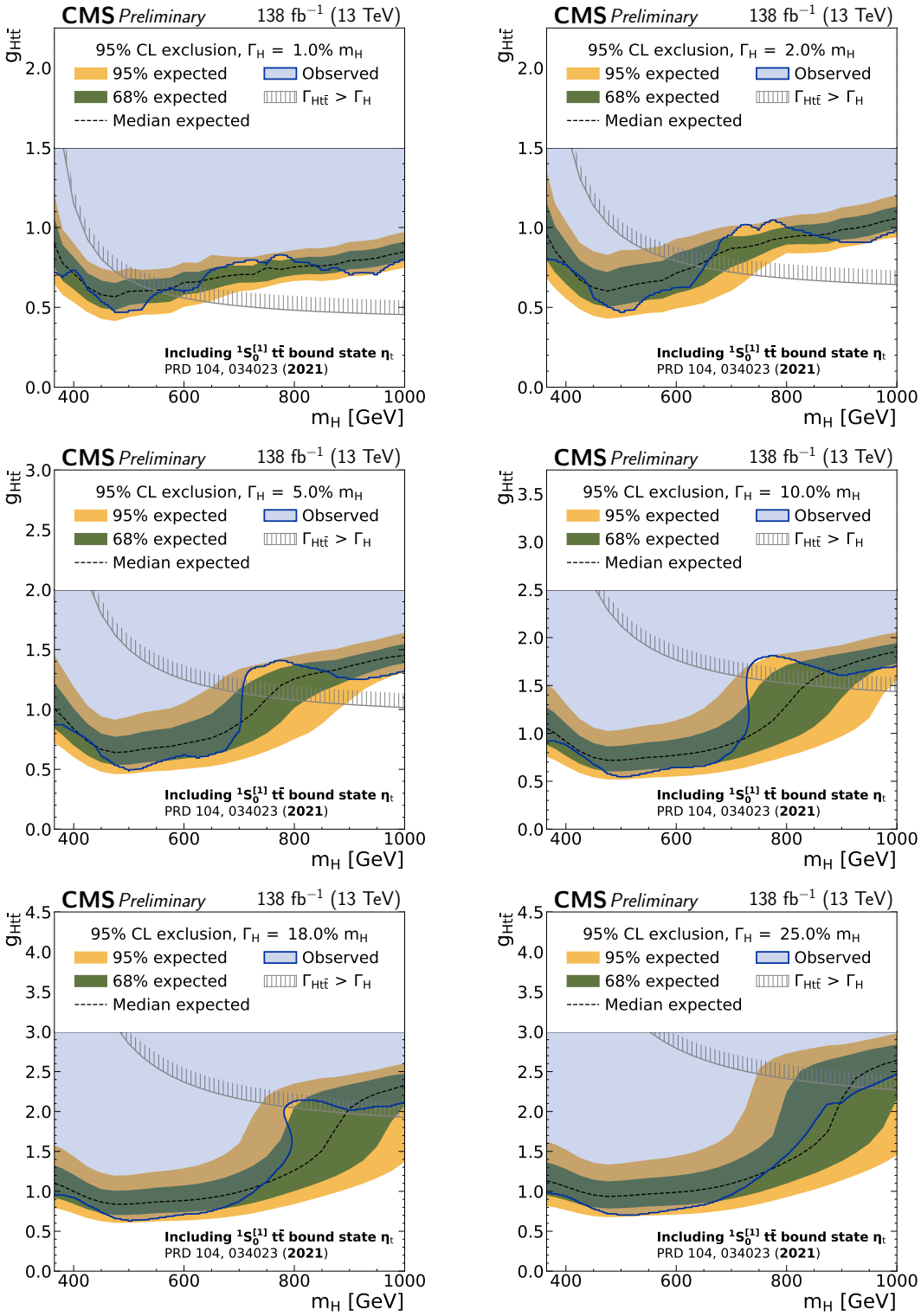

Figure 11: Model-independent constraints on $g_{Ht\bar{t}}$ as a function of the H mass, for relative widths of 1, 2, 5, 10, 18, and 25%. The observed constraints are indicated by the blue shaded area. The inner green band and the outer yellow band indicate the regions containing 68 and 95%, respectively, of the distribution of constraints expected under the background-only hypothesis. The unphysical region of phase space in which the partial width $\Gamma_{H\to t\bar{t}}$ becomes larger than the H total width is indicated by the hatched line.

nario by performing the simultaneous A+H interpretation, as introduced in Section 7.2. Like in Section 8.3, in this study the η_t contribution considered as part of the background model. The results are shown in Fig. 12 for four example cases, consisting of combinations of A and H bosons with masses and widths of 365 GeV, 2% and 1000 GeV, 5%. The observed exclusion contours are compatible with zero A+H contribution in all the example cases shown.

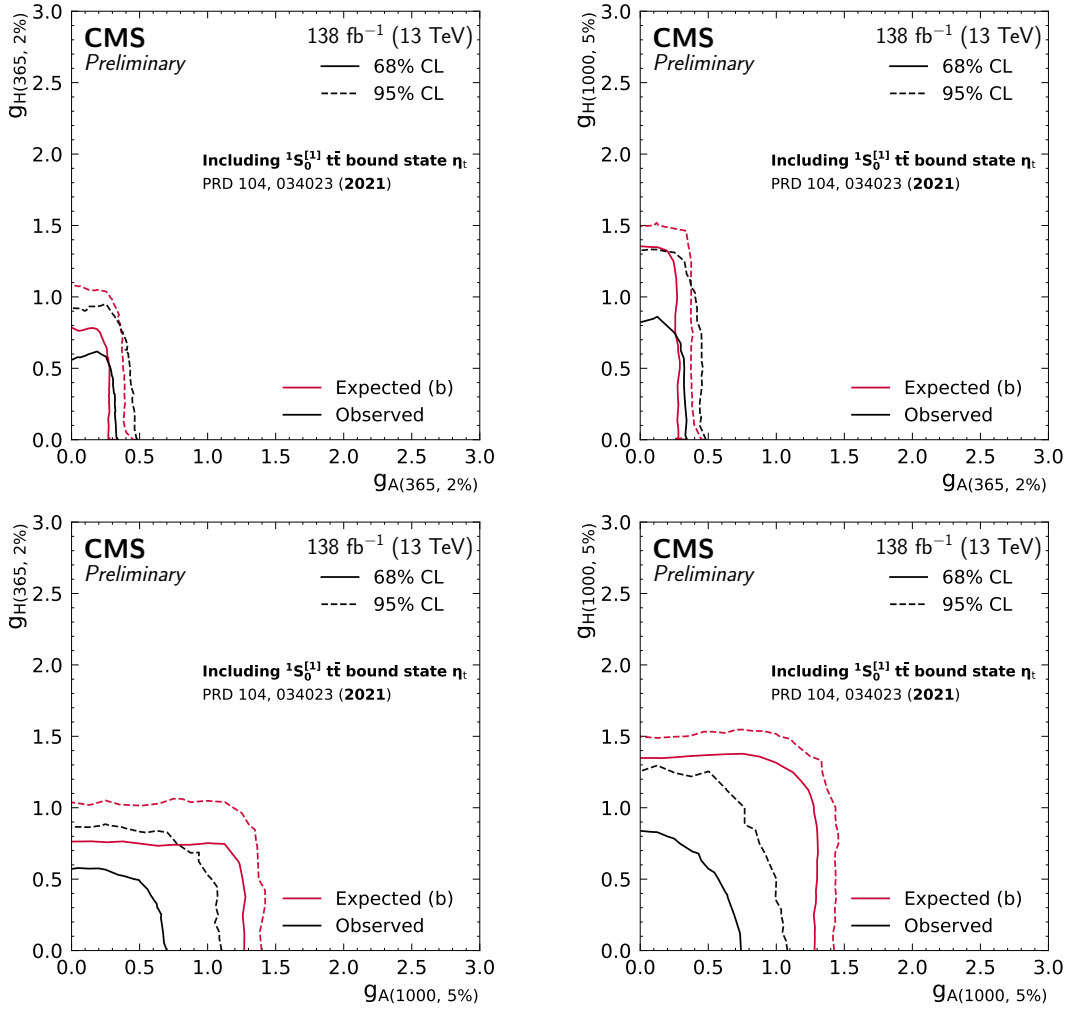

Figure 12: Frequentist 2D exclusion contours for $g_{At\bar{t}}$ and $g_{Ht\bar{t}}$ in the A+H interpretation for four different signal hypotheses: $A(365, 2\%) + H(365, 2\%)$ (upper left), $A(365, 2\%) + H(1000, 5\%)$ (upper right), $A(1000, 5\%) + H(365, 2\%)$ (lower left), and $A(1000, 5\%) + H(1000, 5\%)$ (lower right). The expected and observed contours, evaluated with the Feldman–Cousins prescription [108], are shown in black and red, respectively, with the solid and dashed lines corresponding to exclusions at 68 and 95% CL, and the respective best-fit points for $g_{At\bar{t}}$ and $g_{Ht\bar{t}}$ are shown as the colored crosses. In all cases, the η_t contribution is considered as part of the background.

9 Summary

A search for the production of heavy pseudoscalar or scalar bosons produced in proton-proton collisions at $\sqrt{s} = 13$ TeV and decaying to a top quark pair ($t\bar{t}$) in the final states with one or two charged leptons is presented, using data corresponding to an integrated luminosity of 138 fb^{-1} recorded with the CMS detector at the LHC. The invariant mass of the reconstructed $t\bar{t}$ system

and angular variables sensitive to its spin are used to discriminate the signal from the standard model $\mathrm{t\bar{t}}$ background. Both resonant production of the new boson and interference terms with the perturbative QCD (pQCD) $\mathrm{t\bar{t}}$ background are included in the signal model.

A deviation from the background prediction, modeled only using pQCD, is observed. It is located close to the $\mathrm{t\bar{t}}$ production threshold, similar to the moderate deviation observed in a previous CMS search based on a data sample corresponding to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$ [24]. This deviation significantly favors the pseudoscalar signal hypothesis over the scalar hypothesis. It is compatible with the production of a ${}^1S_0^{[1]}$ $\mathrm{t\bar{t}}$ bound state η_t, as predicted by a simplified model of nonrelativistic QCD. The cross section of this contribution is found to be $\sigma(\eta_\mathrm{t}) = 7.1\,\mathrm{pb}$, with an uncertainty of 11%. The excess has a significance of above five standard deviations. Further investigations by both the experimental and theoretical communities are necessary to elucidate the nature of this excess.

Including η_t production with an unconstrained normalization in the background prediction leads to a good description of the observed data, with no hint for further new pseudoscalar or scalar boson production. Exclusion limits at 95% confidence level are set on the coupling strength between top quarks and new bosons, covering masses of 365–1000 GeV and relative widths of 0.5–25%. Stringent constraints are found for the three cases of a new pseudoscalar boson, a new scalar boson, and the simultaneous presence of one new pseudoscalar and one new scalar boson, excluding coupling values as low as 0.4 (0.6) in the pseudoscalar (scalar) case.

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, "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] 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.

[5] K. Huitu et al., "Probing pseudo-Goldstone dark matter at the LHC", *Phys. Rev. D* **100** (2019) 015009, doi:10.1103/PhysRevD.100.015009, arXiv:1812.05952.

[6] M. Mühlleitner, M. O. P. Sampaio, R. Santos, and J. Wittbrodt, "Phenomenological comparison of models with extended Higgs sectors", *JHEP* **08** (2017) 132, doi:10.1007/JHEP08(2017)132, arXiv:1703.07750.

[7] J. Abdallah et al., "Simplified models for dark matter searches at the LHC", *Phys. Dark Univ.* **9–10** (2015) 8, doi:10.1016/j.dark.2015.08.001, arXiv:1506.03116.

[8] C. Arina et al., "A comprehensive approach to dark matter studies: exploration of simplified top-philic models", *JHEP* **11** (2016) 111, doi:10.1007/JHEP11(2016)111, arXiv:1605.09242.

[9] N. Craig, J. Galloway, and S. Thomas, "Searching for signs of the second Higgs doublet", 2013. arXiv:1305.2424.

[10] M. Carena and Z. Liu, "Challenges and opportunities for heavy scalar searches in the $t\bar{t}$ channel at the LHC", *JHEP* **11** (2016) 159, doi:10.1007/JHEP11(2016)159, arXiv:1608.07282.

[11] A. Djouadi, J. Ellis, A. Popov, and J. Quevillon, "Interference effects in $t\bar{t}$ production at the LHC as a window on new physics", *JHEP* **03** (2019) 119, doi:10.1007/JHEP03(2019)119, arXiv:1901.03417.

[12] K. J. F. Gaemers and F. Hoogeveen, "Higgs production and decay into heavy flavours with the gluon fusion mechanism", *Phys. Lett. B* **146** (1984) 347, doi:10.1016/0370-2693(84)91711-8.

[13] D. Dicus, A. Stange, and S. Willenbrock, "Higgs decay to top quarks at hadron colliders", *Phys. Lett. B* **333** (1994) 126, doi:10.1016/0370-2693(94)91017-0, arXiv:hep-ph/9404359.

[14] W. Bernreuther, M. Flesch, and P. Haberl, "Signatures of Higgs bosons in the top quark decay channel at hadron colliders", *Phys. Rev. D* **58** (1998) 114031, doi:10.1103/PhysRevD.58.114031, arXiv:hep-ph/9709284.

[15] W. Bernreuther, A. Brandenburg, Z. G. Si, and P. Uwer, "Top quark pair production and decay at hadron colliders", *Nucl. Phys. B* **690** (2004) 81, doi:10.1016/j.nuclphysb.2004.04.019, arXiv:hep-ph/0403035.

[16] G. Mahlon and S. J. Parke, "Spin correlation effects in top quark pair production at the LHC", *Phys. Rev. D* **81** (2010) 074024, doi:10.1103/PhysRevD.81.074024, arXiv:1001.3422.

[17] A. H. Hoang et al., "Top-antitop pair production close to threshold: Synopsis of recent NNLO results", in *Proc. 4th Workshop of the 2nd ECFA/DESY Study on Physics and Detectors for a Linear Electron-Positron Collider: Oxford, UK, March 20–23, 1999*. 2000. arXiv:hep-ph/0001286. [Eur. Phys. J. direct 2 (2000) 1]. doi:10.1007/s1010500c0003.

[18] Y. Kiyo et al., "Top-quark pair production near threshold at LHC", *Eur. Phys. J. C* **60** (2009) 375, doi:10.1140/epjc/s10052-009-0892-7, arXiv:0812.0919.

[19] W.-L. Ju et al., "Top quark pair production near threshold: single/double distributions and mass determination", *JHEP* **06** (2020) 158, doi:10.1007/JHEP06(2020)158, arXiv:2004.03088.

[20] B. Fuks, K. Hagiwara, K. Ma, and Y.-J. Zheng, "Signatures of toponium formation in LHC run 2 data", *Phys. Rev. D* **104** (2021) 034023, doi:10.1103/PhysRevD.104.034023, arXiv:2102.11281.

[21] 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.

[22] 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.

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

[24] CMS Collaboration, "Search for heavy Higgs bosons decaying to a top quark pair in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **04** (2020) 171, doi:10.1007/JHEP04(2020)171, arXiv:1908.01115.

[25] ATLAS Collaboration, "Search for heavy Higgs bosons A/H decaying to a top quark pair in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector", *Phys. Rev. Lett.* **119** (2017) 191803, doi:10.1103/PhysRevLett.119.191803, arXiv:1707.06025.

[26] ATLAS Collaboration, "Search for heavy neutral Higgs bosons decaying into a top quark pair in $140\,\text{fb}^{-1}$ of proton-proton collision data at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *JHEP* **08** (2024) 013, doi:10.1007/JHEP08(2024)013, arXiv:2404.18986.

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

[28] 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.

[29] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

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

[31] CMS Collaboration, "Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid", CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015. doi:10.17181/CERN.VU8I.D59J.

[32] 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.

[33] 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.

[34] 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.

[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] E. Bols et al., "Jet flavour classification using DeepJet", *JINST* **15** (2020) P12012, doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.

[38] CMS Collaboration, "Performance summary of AK4 jet b tagging with data from proton-proton collisions at 13 TeV with the CMS detector", CMS Detector Performance Note CMS-DP-2023-005, 2023.

[39] 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.

[40] CMS Collaboration, "ECAL 2016 refined calibration and Run 2 summary plots", CMS Detector Performance Note CMS-DP-2020-021, 2020.

[41] 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.

[42] 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.

[43] CMS Collaboration, "Simulation of the silicon strip tracker pre-amplifier in early 2016 data", CMS Detector Performance Note CMS-DP-2020-045, 2020.

[44] NNPDF Collaboration, "Parton distributions for the LHC run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[45] 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.

[46] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", *Eur. Phys. J. C* **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[47] 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.

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

[49] 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.

[50] 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.

[51] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas, "Higgs boson production at the LHC", *Nucl. Phys. B* **453** (1995) 17, doi:10.1016/0550-3213(95)00379-7, arXiv:hep-ph/9504378.

[52] B. Hespel, F. Maltoni, and E. Vryonidou, "Signal background interference effects in heavy scalar production and decay to a top-anti-top pair", *JHEP* **10** (2016) 016, doi:10.1007/JHEP10(2016)016, arXiv:1606.04149.

[53] R. V. Harlander, S. Liebler, and H. Mantler, "SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the standard model and the MSSM", *Comput. Phys. Commun.* **184** (2013) 1605, doi:10.1016/j.cpc.2013.02.006, arXiv:1212.3249.

[54] R. V. Harlander, S. Liebler, and H. Mantler, "SusHi bento: Beyond NNLO and the heavy-top limit", *Comput. Phys. Commun.* **212** (2017) 239, doi:10.1016/j.cpc.2016.10.015, arXiv:1605.03190.

[55] D. Eriksson, J. Rathsman, and O. Stål, "2HDMC—two-Higgs-doublet model calculator", *Comput. Phys. Commun.* **181** (2010) 189, doi:10.1016/j.cpc.2009.09.011, arXiv:0902.0851. [Erratum: doi:10.1016/j.cpc.2009.12.026].

[56] A. Banfi et al., "Higgs interference effects in top-quark pair production in the 1HSM", 2023. arXiv:2309.16759. Submitted to *JHEP*.

[57] Y. Sumino and H. Yokoya, "Bound-state effects on kinematical distributions of top quarks at hadron colliders", *JHEP* **09** (2010) 034, doi:10.1007/JHEP09(2010)034, arXiv:1007.0075. [Erratum: doi:10.1007/JHEP06(2016)037].

[58] 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.

[59] 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.

[60] 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.

[61] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, "Top-pair production and decay at NLO matched with parton showers", *JHEP* **04** (2015) 114, doi:10.1007/JHEP04(2015)114, arXiv:1412.1828.

[62] M. Czakon and A. Mitov, "TOP++: a program for the calculation of the top-pair cross-section at hadron colliders", *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[63] M. Grazzini, S. Kallweit, and M. Wiesemann, "Fully differential NNLO computations with MATRIX", *Eur. Phys. J. C* **78** (2018) 537, doi:10.1140/epjc/s10052-018-5771-7, arXiv:1711.06631.

[64] M. Aliev et al., "HATHOR: Hadronic top and heavy quarks cross section calculator", *Comput. Phys. Commun.* **182** (2011) 1034, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.

[65] J. H. Kühn, A. Scharf, and P. Uwer, "Electroweak corrections to top-quark pair production in quark-antiquark annihilation", *Eur. Phys. J. C* **45** (2006) 139, doi:10.1140/epjc/s2005-02423-6, arXiv:hep-ph/0508092.

[66] J. H. Kühn, A. Scharf, and P. Uwer, "Electroweak effects in top-quark pair production at hadron colliders", *Eur. Phys. J. C* **51** (2007) 37, doi:10.1140/epjc/s10052-007-0275-x, arXiv:hep-ph/0610335.

[67] J. H. Kühn, A. Scharf, and P. Uwer, "Weak interactions in top-quark pair production at hadron colliders: An update", *Phys. Rev. D* **91** (2015) 014020, doi:10.1103/PhysRevD.91.014020, arXiv:1305.5773.

[68] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[69] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].

[70] P. Kant et al., "HATHOR for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions", *Comput. Phys. Commun.* **191** (2015) 74, doi:10.1016/j.cpc.2015.02.001, arXiv:1406.4403.

[71] N. Kidonakis, "Top quark production", in *Proc. Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): Dubna, Russia, July 15–28, 2013*. 2013. arXiv:1311.0283. [DESY-PROC-2013-03]. doi:10.3204/DESY-PROC-2013-03/Kidonakis.

[72] P. F. Monni et al., "MiNNLO$_{\rm PS}$: a new method to match NNLO QCD to parton showers", *JHEP* **05** (2020) 143, doi:10.1007/JHEP05(2020)143, arXiv:1908.06987. [Erratum: doi:10.1007/JHEP02(2022)031].

[73] P. F. Monni, E. Re, and M. Wiesemann, "MiNNLO$_{\rm PS}$: optimizing $2 \to 1$ hadronic processes", *Eur. Phys. J. C* **80** (2020) 1075, doi:10.1140/epjc/s10052-020-08658-5, arXiv:2006.04133.

[74] E. Barberio and Z. Wąs, "PHOTOS—a universal Monte Carlo for QED radiative corrections: version 2.0", *Comput. Phys. Commun.* **79** (1994) 291, doi:10.1016/0010-4655(94)90074-4.

[75] P. Golonka and Z. Wąs, "PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays", *Eur. Phys. J. C* **45** (2006) 97, doi:10.1140/epjc/s2005-02396-4, arXiv:hep-ph/0506026.

[76] 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.

[77] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_S^2)$", *Phys. Rev. D* **74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[78] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in FEWZ", *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.

[79] T. Gehrmann et al., "W^+W^- production at hadron colliders in next to next to leading order QCD", *Phys. Rev. Lett.* **113** (2014) 212001, doi:10.1103/PhysRevLett.113.212001, arXiv:1408.5243.

[80] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", in *Proc. 10th DESY Workshop on Elementary Particle Theory: Loops and Legs in Quantum Field Theory (LL2010): Wörlitz, Germany, April 25–30, 2010*. 2010. arXiv:1007.3492. [Nucl. Phys. B Proc. Suppl. 205–206 (2010) 10]. doi:10.1016/j.nuclphysbps.2010.08.011.

[81] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **06** (2002) 029, doi:10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.

[82] B. A. Betchart, R. Demina, and A. Harel, "Analytic solutions for neutrino momenta in decay of top quarks", *Nucl. Instrum. Meth. A* **736** (2014) 169, doi:10.1016/j.nima.2013.10.039, arXiv:1305.1878.

[83] R. Demina, A. Harel, and D. Orbaker, "Reconstructing $t\bar{t}$ events with one lost jet", *Nucl. Instrum. Meth. A* **788** (2015) 128, doi:10.1016/j.nima.2015.03.069, arXiv:1310.3263.

[84] CMS Collaboration, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s}$ = 8 TeV", *Eur. Phys. J. C* **75** (2015) 542, doi:10.1140/epjc/s10052-015-3709-x, arXiv:1505.04480.

[85] L. Sonnenschein, "Analytical solution of $t\bar{t}$ dilepton equations", *Phys. Rev. D* **73** (2006) 054015, doi:10.1103/PhysRevD.73.054015, arXiv:hep-ph/0603011. [Erratum: doi:10.1103/PhysRevD.78.079902].

[86] W. Bernreuther, D. Heisler, and Z.-G. Si, "A set of top quark spin correlation and polarization observables for the LHC: Standard model predictions and new physics contributions", *JHEP* **12** (2015) 026, doi:10.1007/JHEP12(2015)026, arXiv:1508.05271.

[87] J. A. Aguilar-Saavedra and J. A. Casas, "Improved tests of entanglement and Bell inequalities with LHC tops", *Eur. Phys. J. C* **82** (2022) 666, doi:10.1140/epjc/s10052-022-10630-4, arXiv:2205.00542.

[88] F. Maltoni, C. Severi, S. Tentori, and E. Vryonidou, "Quantum detection of new physics in top-quark pair production at the LHC", *JHEP* **03** (2024) 099, doi:10.1007/JHEP03(2024)099, arXiv:2401.08751.

[89] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section and the top quark mass in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **07** (2011) 049, doi:10.1007/JHEP07(2011)049, arXiv:1105.5661.

[90] CMS Collaboration, "The CMS statistical analysis and combination tool: COMBINE", 2024. arXiv:2404.06614. Accepted by *Comput. Softw. Big Sci.*

[91] CMS Collaboration, "Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **79** (2019) 421, doi:10.1140/epjc/s10052-019-6909-y, arXiv:1809.10733.

[92] CMS Collaboration, "Measurement of the top quark mass using proton-proton data at $\sqrt{s} = 7$ and 8 TeV", *Phys. Rev. D* **93** (2016) 072004, doi:10.1103/PhysRevD.93.072004, arXiv:1509.04044.

[93] M. Botje et al., "The PDF4LHC working group interim recommendations", 2011. arXiv:1101.0538.

[94] J. R. Christiansen and P. Z. Skands, "String formation beyond leading colour", *JHEP* **08** (2015) 003, doi:10.1007/JHEP08(2015)003, arXiv:1505.01681.

[95] CMS Collaboration, "CMS PYTHIA 8 colour reconnection tunes based on underlying-event data", *Eur. Phys. J. C* **83** (2023) 587, doi:10.1140/epjc/s10052-023-11630-8, arXiv:2205.02905.

[96] CMS Collaboration, "Investigations of the impact of the parton shower tuning in PYTHIA 8 in the modelling of $\mathrm{t\bar{t}}$ at $\sqrt{s} = 8$ and 13 TeV", CMS Physics Analysis Summary CMS-PAS-TOP-16-021, 2016.

[97] ATLAS Collaboration, "Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **04** (2017) 086, doi:10.1007/JHEP04(2017)086, arXiv:1609.03920.

[98] CMS Collaboration, "Measurement of the single top quark and antiquark production cross sections in the t channel and their ratio in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **800** (2020) 135042, doi:10.1016/j.physletb.2019.135042, arXiv:1812.10514.

[99] CMS Collaboration, "Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **10** (2018) 117, doi:10.1007/JHEP10(2018)117, arXiv:1805.07399.

[100] CMS Collaboration, "Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **08** (2018) 011, doi:10.1007/JHEP08(2018)011, arXiv:1711.02547.

[101] ATLAS Collaboration, "Measurement of the $\mathrm{t\bar{t}Z}$ and $\mathrm{t\bar{t}W}$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. D* **99** (2019) 072009, doi:10.1103/PhysRevD.99.072009, arXiv:1901.03584.

[102] ATLAS Collaboration, "Measurements of top-quark pair to Z-boson cross-section ratios at $\sqrt{s} = 13, 8, 7$ TeV with the ATLAS detector", *JHEP* **02** (2017) 117, doi:10.1007/JHEP02(2017)117, arXiv:1612.03636.

[103] R. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993) 219, doi:10.1016/0010-4655(93)90005-W.

[104] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[105] ATLAS and CMS Collaborations, and LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.

[106] 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.

[107] A. L. Read, "Presentation of search results: The $\mathrm{CL_s}$ technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[108] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals", *Phys. Rev. D* **57** (1998) 3873, doi:10.1103/PhysRevD.57.3873, arXiv:physics/9711021.

[109] T. Ježo et al., "An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects", *Eur. Phys. J. C* **76** (2016) 691, doi:10.1140/epjc/s10052-016-4538-2, arXiv:1607.04538.

[110] T. Ježo, J. M. Lindert, and S. Pozzorini, "Resonance-aware NLOPS matching for off-shell $t\bar{t} + tW$ production with semileptonic decays", *JHEP* **10** (2023) 008, doi:10.1007/JHEP10(2023)008, arXiv:2307.15653.

[111] T. Ježo and P. Nason, "On the treatment of resonances in next-to-leading order calculations matched to a parton shower", *JHEP* **12** (2015) 065, doi:10.1007/JHEP12(2015)065, arXiv:1509.09071.

A Characterization of the deviation in the $t\bar{t}$ threshold region

As discussed in Section 8.1, the significant deviation from the pQCD-only background expectation is found at low $m_{t\bar{t}}$ values, and can be described by a model with a single Φ boson with low m_Φ, where the best description is obtained with a pseudoscalar. The A and H signal hypothesis differ significantly in the predicted signal rates for m_Φ close to the $t\bar{t}$ threshold. To evaluate if the quoted preference for the A boson is just a consequence of the size of the deviation or indeed points towards specific spin and CP properties, we present a dedicated study in this appendix.

A modified A+H interpretation is performed, assuming the pQCD-only background model and using the A/H(365, 2%) signal configuration. Only the resonant component of the signal model is used, and both the A and H contributions are independently normalized to an arbitrary nominal cross section of 10 pb. The normalization of the templates corresponding to systematic variations are such that their relative deviation with respect to the nominal templates is preserved. The result of this interpretation is shown in Fig. A.1.

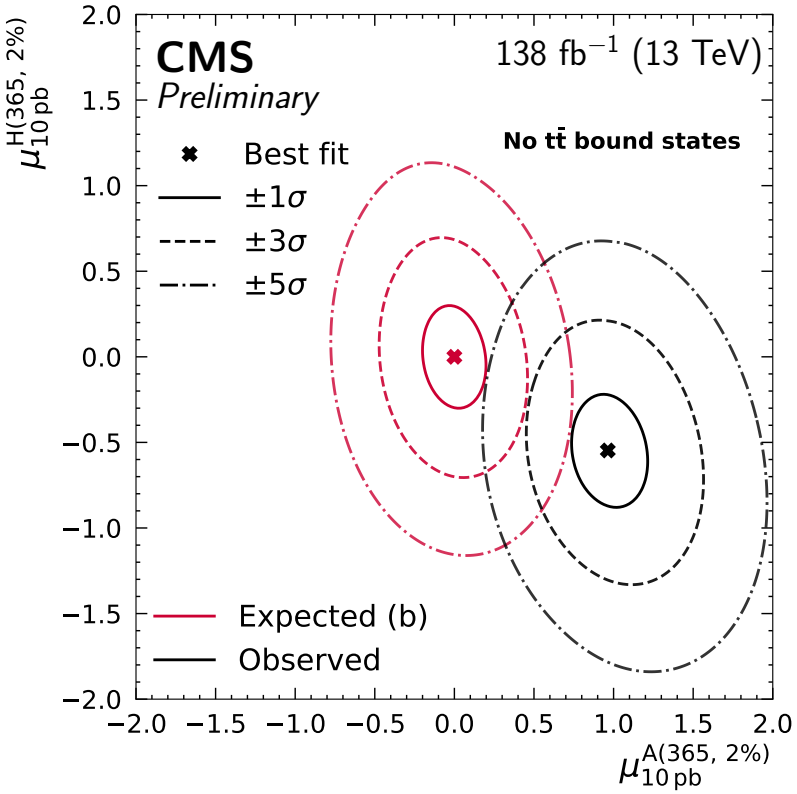

Figure A.1: Local significance contours for the pair of A/H(365, 2%), considering only the resonant signal components. Different line styles are used to indicate the regions compatible with the data at progressive confidence levels.

The best fit values of the arbitrary signal strengths $\mu^{A/H(365,2\%)}_{10\,pb}$ lie far from the origin of zero signal contribution, with a local significance beyond five standard deviations. Comparing the A and H directions, we see that although the observed data is compatible with zero H(365, 2%) contribution within three standard deviations, the case of $\mu^{A(365,2\%)}_{10\,pb} = 0$ is not compatible within five standard deviations at any $\mu^{H(365,2\%)}_{10\,pb}$ value. Based on this, we conclude that the nature of the deviation is significantly more pseudoscalar-like than scalar-like.

B Investigation of alternate predictions for the SM $t\bar{t}$ background

To check the robustness of the cross section extracted using the η_t model in Section 8.2, the fit is repeated with an alternate prediction for the SM $t\bar{t}$ and tW background, generated using the `b_bbar_4l` subprocess in POWHEG vRES [109–111]. This generator calculates the full matrix element for the process $pp \to b\bar{b}\ell\bar{\ell}\nu\bar{\nu}$ at NLO in QCD including off-shell top quark and finite top quark width effects, which might be relevant in the $t\bar{t}$ threshold region where the top quarks are off-shell. Because `b_bbar_4l` is at the time of writing only available for $t \to b\ell\bar{\nu}$ decays, this test is only performed in the $\ell\bar{\ell}$ channels. The same NNLO QCD and NLO EW corrections as for POWHEG v2, as described in Section 3, are applied to the $t\bar{t}$ part of `b_bbar_4l`, defined by the resonance histories of the events [110]. Only the nominal prediction is exchanged, while the relative effect of systematic uncertainties is estimated from POWHEG v2 as before. The results on the η_t cross section can be found in Table B.1, compared to the $\ell\bar{\ell}$-only result for the default setup. It is found to be about 2 standard deviations lower than the nominal result, with no significant change in the absolute uncertainty.

Table B.1: Results on the η_t cross section, using only the $\ell\bar{\ell}$ channels, for the `b_bbar_4l` background prediction and for the default setup. The quoted uncertainty for `b_bbar_4l` assumes the same uncertainty as for the nominal result.

Prediction for SM $t\bar{t}$ and tW	Extracted η_t cross section	Uncertainty
`b_bbar_4l` (POWHEG vRES)	5.9 pb	18%
Default (POWHEG v2)	7.5 pb	13%

C Comparison of the η_t and single Φ boson interpretations

The compatibility of the η_t, single A boson, and single H boson interpretations with the observed excess is further quantified by providing the difference in two times the negative log-likelihood, $-2\ln L$, between the best-fit point for a given interpretation and the background-only hypothesis. This corresponds directly to the test statistics used for the different interpretations, as defined in Equation (5). The results are provided in Table C.1. The η_t interpretation yields the largest difference to the background-only hypotheses, signifying the best statistical compatibility with the data. It should be noted that part of this difference might result from the choice of mass range for Φ bosons with a lower bound of $m_\Phi = 365\,\text{GeV}$, which is higher than the mass of 343 GeV assumed for η_t.

Table C.1: Best-fit values of the signal strengths or coupling modifiers as well as differences in $-2\ln L$ between the best-fit point and the background-only hypothesis for the η_t, single A boson, and single H boson interpretations.

Interpretation	Best-fit point	Difference in $-2\ln L$
η_t	$\mu(\eta_t) = 1.11$	-86.2
Single A boson	$m_A = 365\,\text{GeV}$, $\Gamma_A/m_A = 2\%$, $g_{At\bar{t}} = 0.78$	-72.6
Single H boson	$m_H = 365\,\text{GeV}$, $\Gamma_H/m_H = 2\%$, $g_{Ht\bar{t}} = 1.45$	-10.4