Search for heavy scalar or pseudoscalar states in tt̄ events at CMS

Laurids Jeppe, on behalf of the CMS collaboration

Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany

laurids.jeppe@cern.ch

The 17th International Workshop on
Top Quark Physics (TOP2024)
Saint-Malo, France, 22-27 September 2024
doi:[10.21468/SciPostPhysProc.18](https://doi.org/10.21468/SciPostPhysProc.18)

Abstract

A search for scalar or pseudoscalar states decaying to a top quark-antiquark pair (tt̄), using $138\,\text{fb}^{-1}$ of pp collision data taken at $\sqrt{s} = 13\,\text{TeV}$ using the CMS detector, is presented. Events with one or two leptons are analyzed using the invariant tt̄ mass ($m_{\text{tt̄}}$) as well as angular and spin correlation observables. An excess in the data is observed for low values of $m_{\text{tt̄}}$, preferring a pseudoscalar over a scalar hypothesis. It is interpreted in terms of a generic model of (pseudo)scalar boson production, as well as a simplified model of a tt̄ bound state (η_{t}), yielding good agreement with the data. Moreover, limits on the couplings of additional (pseudo)scalar bosons to top quarks are set.

doi:[10.21468/SciPostPhysProc.18.020](https://doi.org/10.21468/SciPostPhysProc.18.020)

1 Introduction

The Higgs sector of the Standard Model (SM) is often considered a promising avenue to search for Beyond the Standard Model (BSM) physics, and many models including extended Higgs sectors, such as the Two-Higgs Doublet Model (2HDM) or supersymmetric models have been proposed. The new particles predicted in these models often exhibit Yukawa-like couplings to fermions, leading to large couplings to the top quark due to its large mass. In particular, for new electrically neutral states with masses above $2m_{\text{t}}$, the decay to a top quark-antiquark pair (tt̄) is often dominant in large areas of parameter space, motivating searches for such states in tt̄ events at the LHC. In this work, heavy spin-0 bosons with scalar or pseudoscalar couplings to the top quark are considered.

At the same time, tt̄ bound states are expected to form in the SM according to several calculations (see e.g. Refs. [1–3]), manifesting as a broad peak in the invariant tt̄ mass spectrum slightly below the tt̄ threshold. At the LHC, they are expected to be dominated by a pseudoscalar component, leading to a similar signature as an additional pseudoscalar boson, which suggests a search for these effects using the same methodology.

This work presents such a search [4], performed with the CMS experiment [5] in proton-proton collisions at the LHC with the full Run 2 dataset, corresponding to an integrated luminosity of $138\,\text{fb}^{-1}$. A similar search was presented by ATLAS in Ref. [6]. The search presented here updates on a previous work, which considered only $35.9\,\text{fb}^{-1}$ of data [7]. Two analysis channels are considered, targeting the dilepton ($\ell\ell$) and lepton+jets (ℓj) decay channels of $t\bar{t}$, and the invariant $t\bar{t}$ mass as well as angular and spin correlation observables are employed to isolate the signals from the SM background.

2 Signal modeling

For the interpretation of a generic pseudoscalar or scalar boson (denoted A and H, respectively), the signal is assumed to be produced in gluon fusion through a top quark loop, with only a Yukawa-like coupling to the top quark considered. As a result, the free parameters of the model are the masses, widths and coupling strength modifiers ($g_{\text{A}t\bar{t}}$ resp. $g_{\text{H}t\bar{t}}$). Because the resulting final state is the same as in $t\bar{t}$ production in the SM, interference with the SM is expected, leading to a peak-dip structure in the $m_{t\bar{t}}$ spectrum.

For $t\bar{t}$ bound states, on the other hand, no full calculation that can be directly compared to data is available at the time of writing. Calculations of the expected $m_{t\bar{t}}$ spectrum can be performed in the framework of non-relativistic QCD (NRQCD) [1], which predicts an attractive potential and a resulting peak for the color-singlet component of $t\bar{t}$, and a repulsive potential, leading to a suppression, for the color-octet component. In this work, a simplified model (proposed in Ref. [8]) for this peak is used, denoted η_{t}. It consists of a generic spin-0, color-singlet, pseudoscalar resonance coupling directly to gluons and top quarks, with the mass and width of the resonance extracted from a fit to an NRQCD prediction, which yields $m(\eta_{\text{t}}) = 343\,\text{GeV}$ and $\Gamma(\eta_{\text{t}}) = 7\,\text{GeV}$. To not influence the $t\bar{t}$ continuum, the model is restricted to invariant masses of $|m_{t\bar{t}} - m(\eta_{\text{t}})| < 6\,\text{GeV}$. The prediction from this model is then added to the continuum $t\bar{t}$ prediction as given by perturbative QCD (pQCD). While this model is not expected to reproduce the details of the predicted lineshape, which is anyway not well known in the first place, it is deemed sufficient for this analysis due to the coarse experimental $m_{t\bar{t}}$ resolution of around 15% close to the $t\bar{t}$ threshold rendering the details irrelevant.

3 Analysis setup

In the ℓj channels, events with exactly one lepton (e or μ) and 3 or more jets, of which at least 2 are b-tagged, are selected and sorted into four categories based on the lepton flavor as well as the number of jets (3 or $\geq$4). The NeutrinoSolver algorithm [9] is used to reconstruct the $t\bar{t}$ system, and an energy correction factor is applied for events with exactly three jets to account for jets that were merged or lost [10]. Two-dimensional templates are constructed based on $m_{t\bar{t}}$ and $|\cos\theta^{\star}|$, where $\theta^{\star}$ is the scattering angle of the leptonically decaying top quark with respect to the beam axis. This variable has discriminating power because SM $t\bar{t}$ production peaks at small scattering angles, while the A, H and η_{t} signals are isotropic.

In the $\ell\ell$ channels, events with exactly two leptons and at least 2 jets, with at least one b-tagged, are selected. They are split into categories by lepton flavor, and in the same-flavor channels additional cuts are applied to reject Z+jets events. Again, the four-momenta of the $t\bar{t}$ system is reconstructed, employing an analytical approach which assumes that the two neutrinos in the $t\bar{t}$ decay are the sole source of missing transverse momentum ($p_{\text{T}}^{\text{miss}}$) and that the top quarks and W bosons are on-shell. The finite detector resolution is taken into account by repeating the reconstruction 100 times per events with inputs (lepton and jet four-momenta as well as $p_{\text{T}}^{\text{miss}}$) randomly smeared and taking a weighted average over all real solutions.

Three-dimensional templates are constructed from $m_{t\bar{t}}$ as well as two spin correlation observables c_{hel} and c_{han}. They are defined in the helicity basis for the top and antitop spins [11], corresponding to boosting the lepton four-momenta into the rest frames of their parent top quarks. The observable c_{hel} is the scalar product of the lepton directions of flight in the helicity basis, while c_{han} is defined as a similar scalar product containing a negative sign in the top quark direction of flight. The distributions of both of these observables (before phase space cuts) are straight lines, whose slopes depend on entries in the $t\bar{t}$ spin density matrix and thus directly probe the spin state of the $t\bar{t}$ system. For pure $^1S_0$ states, as produced by the pseudoscalars A and η_t, the slope of c_{hel} is maximally positive, while for pure $^3P_0$ states (produced by the scalar H), the slope of c_{han} is maximally negative [12]. As a result, the observables provide good discrimination for the different signal hypotheses.

The dominant background in all channels consists of SM $t\bar{t}$ production, which is generated at NLO in QCD using Powheg v2, interfaced to Pythia 8, and reweighted to higher order predictions at NNLO in QCD and NLO in electroweak processes using a two-dimensional bins of $m_{t\bar{t}}$ and the top scattering angle $\cos\theta^\star$. Further backgrounds are tW and t-channel single-top production (estimated from MC), Z+jets production ($\ell\ell$ only, estimated from MC with data-driven normalization) and QCD multijet as well as W+jets production (ℓj only, estimated from a data sideband).

4 Results

Across all channels, an excess of data compared to the continuum pQCD prediction is observed at low values of $m_{t\bar{t}}$, translating directly to an excess of the observed over the expected limits for the coupling strengths $g_{At\bar{t}}$ and $g_{Ht\bar{t}}$. The excess is found to be stronger in the pseudoscalar case A.
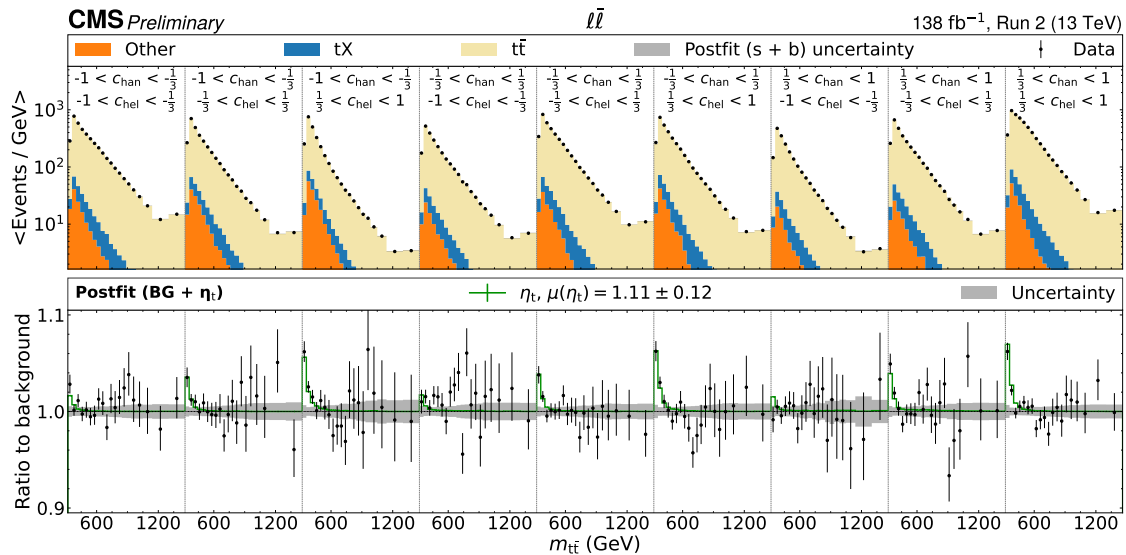
Figure 1: Postfit three-dimensional distribution of $m_{t\bar{t}}$, c_{hel} and c_{han} in the $\ell\ell$ channel in a signal+background fit with η_t as the signal. The perturbative QCD SM predictions are shown in the upper panel as the colored stack and the observed data as the black dots. The lower panel shows the ratio of data to the pQCD prediction, the postfit uncertainty (grey band), and the η_t signal normalized to the postfit yield (green line). This figure is taken from the auxiliary material of Ref. [4].
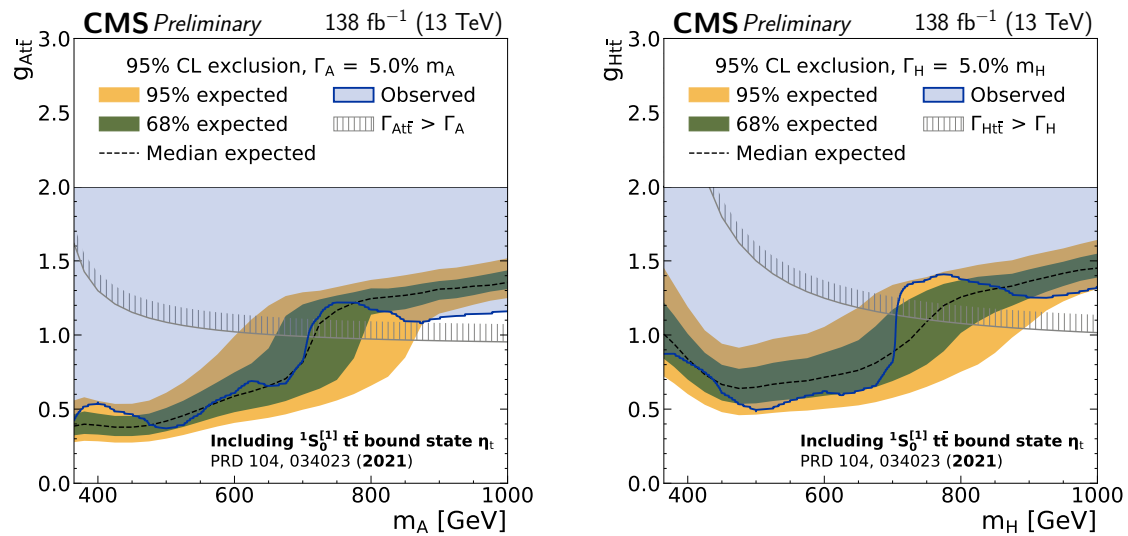
Figure 2: Exclusion limits at 95% confidence level on the coupling strength modifier of the pseudoscalar A (left) and scalar H (right) as a function of their respective mass for an example width of 5%. The observed (expected) limit is shown in blue (black), and the 68% (95%) intervals are shown as the green (yellow) band. This figure is taken from Ref. [4].

Based on these observations, the excess is interpreted as a possible $t\bar{t}$ bound state using the η_t model as described in Sec. 2 by extracting the η_t cross section, corresponding to the difference of the observed data yield to the pQCD prediction, in a signal+background fit.

The resulting cross section is $\sigma(\eta_t) = 7.1 \pm 0.8\,\text{pb}$. By comparing to the prediction given in Ref. [8] of $\sigma(\eta_t)^{\text{pred}} = 6.43\,\text{pb}$, obtained by fitting a NRQCD prediction, good agreement is found. The parameterized model seems to describe the data well, as shown as an example for the $\ell\ell$ channel in form of a postfit distribution in Fig. 1. It can in particular be seen that both the signal simulation and the data exhibit an increase in the slope of c_{hel} in the low $m_{t\bar{t}}$ bins compared to the pQCD prediction, which is consistent with a pseudoscalar contribution. Nonetheless, it should be cautioned that the η_t model considered here is not a complete description of a $t\bar{t}$ bound state, but a simple parameterization, missing for example (but not limited to) contributions from soft gluons changing color-octet into color-singlet states.

The uncertainty on the η_t cross section is dominated by systematic effects from the pQCD background modeling, including the electroweak corrections to $t\bar{t}$, parton shower uncertainties, missing higher orders in the matrix elements, PDF uncertainties and the uncertainty in the top quark mass. For a full overview of the systematic uncertainties, see Ref. [4].

Following the observation that the data is well described by the η_t model, limits on additional (BSM) spin-0 states A and H are set by including the η_t contribution in the background prediction, with its normalization freely floating in the fit. The excess is now no longer present, and limits on the couplings $g_{\text{At}\bar{t}}$ and $g_{\text{Ht}\bar{t}}$ are set in the mass range of $365 - 1000$ GeV for different A/H widths. An example is shown in Fig. 2 for a relative width of 5%.

5 Conclusion

A search for heavy scalar or pseudoscalar spin-0 states in $t\bar{t}$ events, performed with the CMS detector with 138 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV, is presented [4]. The dilepton and lepton+jets decay channels of $t\bar{t}$ are analyzed using the invariant $t\bar{t}$ mass as well

as angular and spin correlation observables. An excess, localized in low bins of $m_{t\bar{t}}$, is observed, and found to be consistent with a pseudoscalar state. It is interpreted in terms of a generic model for scalar or pseudoscalar boson production (A or H), as well as a parameterized model of a $t\bar{t}$ bound state (η_t). The cross section of η_t, understood as the difference to the pQCD background prediction, is extracted as $\sigma(\eta_t) = 7.1 \pm 0.8$ pb. Good agreement with the data is found, and stringent exclusion limits are set on the couplings of A and H by adding η_t to the background prediction.

Acknowledgments

Funding information L.J. acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2121 "Quantum Universe" – 390833306. This work has been partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 491245950.

References

[1] Y. Kiyo, J. H. Kühn, S. Moch, M. Steinhauser and P. Uwer, *Top-quark pair production near threshold at LHC*, Eur. Phys. J. C **60**, 375 (2009), doi:10.1140/epjc/s10052-009-0892-7 [preprint doi:10.48550/arXiv.0812.0919].

[2] Y. Sumino and H. Yokoya, *Bound-state effects on kinematical distributions of top quarks at hadron colliders*, J. High Energy Phys. **09**, 034 (2010), doi:10.1007/JHEP09(2010)034 [preprint doi:10.48550/arXiv.1007.0075].

[3] W.-L. Ju, G. Wang, X. Wang, X. Xu, Y. Xu and L. L. Yang, *Top quark pair production near threshold: Single/double distributions and mass determination*, J. High Energy Phys. **06**, 158 (2020), doi:10.1007/JHEP06(2020)158 [preprint doi:10.48550/arXiv.2004.03088].

[4] CMS collaboration, *Search for heavy pseudoscalar and scalar bosons decaying to a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Tech. Rep. (CERN-EP-2025-124), CERN, Geneva, Switzerland (2025), https://cds.cern.ch/record/2937738 [preprint doi:10.48550/arXiv.2507.05119].

[5] CMS collaboration: S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, J. Instrum. **3**, S08004 (2008), doi:10.1088/1748-0221/3/08/S08004.

[6] ATLAS collaboration: G. Aad et al., *Search for heavy neutral Higgs bosons decaying into a top quark pair in 140 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, J. High Energy Phys. **08**, 013 (2024), doi:10.1007/JHEP08(2024)013 [preprint doi:10.48550/arXiv.2404.18986].

[7] CMS collaboration: A. M. Sirunyan et al., *Search for heavy Higgs bosons decaying to a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV*, J. High Energy Phys. **04**, 171 (2020), doi:10.1007/JHEP04(2020)171 [preprint doi:10.48550/arXiv.1908.01115].

[8] B. Fuks, K. Hagiwara, K. Ma and Y.-J. Zheng, *Signatures of toponium formation in LHC run 2 data*, Phys. Rev. D **104**, 034023 (2021), doi:10.1103/PhysRevD.104.034023 [preprint doi:10.48550/arXiv.2102.11281].

[9] B. A. Betchart, R. Demina and A. Harel, *Analytic solutions for neutrino momenta in decay of top quarks*, Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. **736**, 169 (2014), doi:10.1016/j.nima.2013.10.039 [preprint doi:10.48550/arXiv.1305.1878].

[10] R. Demina, A. Harel and D. Orbaker, *Reconstructing $t\bar{t}$ events with one lost jet*, Nucl. Instrum. Methods Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. **788**, 128 (2015), doi:10.1016/j.nima.2015.03.069 [preprint doi:10.48550/arXiv.1310.3263].

[11] 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*, J. High Energy Phys. **12**, 001 (2015), doi:10.1007/JHEP12(2015)026 [preprint doi:10.48550/arXiv.1508.05271].

[12] F. Maltoni, C. Severi, S. Tentori and E. Vryonidou, *Quantum detection of new physics in top-quark pair production at the LHC*, J. High Energy Phys. **03**, 099 (2024), doi:10.1007/JHEP03(2024)099 [preprint doi:10.48550/arXiv.2401.08751].