CP Properties of Higgs Boson Interactions with Top Quarks in the \((\text{tt})\overline{\text{t}}\overline{\text{H}}\) and \(t\overline{\text{t}}H\) Processes Using \(H \rightarrow \gamma\gamma\) with the ATLAS Detector

Aad, G.; Abbott, B.; Abbott, DC; Abud, AA; Abeling, K.; Abhayasinghe, DK; Abidi, S.H.; AbouZeid, Ossama Sherif Alexander; Abraham, NL; Abramowicz, H.; Abreu, H.; Abulaiti, Y.; Alonso Diaz, Alejandro; hqz214, hqz214; Camplani, Alessandra; Dam, Mogens; Galster, Gorm Aske Gram Krohn; Hansen, Jørn Dines; Hansen, Peter Henrik; Hansen, Jørgen Beck; Ignazzi, Rosanna; Monk, James William; Petersen, Troels Christian; Wiglesworth, Graig; Stark, Simon Holm; Xella, Stefania; ATLAS Collaboration

Published in:
Physical Review Letters

DOI:
10.1103/PhysRevLett.125.061802

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and $tH$ Processes
Using $H \to \gamma\gamma$ with the ATLAS Detector

G. Aad et al.
(ATLAS Collaboration)

(Received 10 April 2020; accepted 15 June 2020; published 5 August 2020)

A study of the charge conjugation and parity (CP) properties of the interaction between the Higgs boson and top quarks is presented. Higgs bosons are identified via the diphoton decay channel ($H \to \gamma\gamma$), and their production in association with a top quark pair ($t\bar{t}H$) or single top quark ($tH$) is studied. The analysis uses 139 fb$^{-1}$ of proton–proton collision data recorded at a center-of-mass energy of $\sqrt{s} = 13$ TeV with the ATLAS detector at the Large Hadron Collider. Assuming a CP-even coupling, the $t\bar{t}H$ process is observed with a significance of 5.2 standard deviations. The measured cross section times $H \to \gamma\gamma$ branching ratio is $1.64^{+0.38}_{-0.36}(\text{stat})^{+0.17}_{-0.14}(\text{sys})$ fb, and the measured rate for $t\bar{t}H$ is $1.43^{+0.37}_{-0.31}(\text{stat})^{+0.21}_{-0.15}(\text{sys})$ times the Standard Model expectation. The $tH$ production process is not observed and an upper limit on its rate of 12 times the Standard Model expectation is set. A CP-mixing angle greater (less) than 43 (−43)° is excluded at 95% confidence level.

DOI: 10.1103/PhysRevLett.125.061802

The observation of Higgs boson production in association with top quarks at the LHC [1,2] provides an opportunity to probe the charge conjugation and parity (CP) properties of the Yukawa coupling of the Higgs boson to the top quark. The Standard Model (SM) of particle physics predicts the Higgs boson to be a scalar particle ($J^{CP} = 0^{++}$) with a prescribed coupling to the top quark. However, the presence of a $J^{CP} = 0^{+-}$ pseudoscalar admixture, which introduces a second coupling to the top quark, has not yet been excluded. Any measured CP-odd contribution would be a sign of physics beyond the SM and could account for the explanation of the observed baryon asymmetry of the universe. This Letter presents a search for CP violation in this coupling and measurements of the production rate of the Higgs boson, via its decay into two photons, in association with top quarks. Recently, the CMS Collaboration performed a similar study [3].

Studies of CP properties of the Higgs boson interactions with gauge bosons have been performed by the ATLAS and CMS experiments [4–9]; the results show no deviations from the SM predictions. However, these measurements probe the bosonic couplings in which CP-odd contributions enter only via higher-order operators that are suppressed by powers of $1/\Lambda^2$ [10], where $\Lambda$ is the scale of the new physics in an effective field theory (EFT). In the case of the Yukawa couplings, the CP-odd contributions are not suppressed by powers of $1/\Lambda^2$.

The CP properties of the top Yukawa coupling can be probed directly using Higgs boson production in association with top quarks: $t\bar{t}H$ and $tH$ processes. The couplings impact the production rates [11–14] and some kinematic distributions. The $tH$ rate is particularly sensitive to deviations from SM couplings due to destructive interference in the SM between diagrams where the Higgs boson radiates from a top quark and from a $W$ boson. The presence of CP-mixing in the top Yukawa coupling also modifies the gluon–gluon fusion (ggF) production rate and the $H \to \gamma\gamma$ decay rate.

This analysis is performed using 139 fb$^{-1}$ of $\sqrt{s} = 13$ TeV proton–proton ($pp$) collision data recorded from 2015 to 2018 with the ATLAS detector. The ATLAS detector [15–17] is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and near 4\pi coverage in solid angle [18]. The trigger system consists of a hardware-based first-level trigger and a software-based high-level trigger [19]. Events used in this analysis were triggered by requiring two photons with a loose identification requirement [20] in the 2015–2016 data-taking period and transverse energies of at least 25 GeV and 35 GeV for the subleading and leading photons, respectively. Due to the greater instantaneous luminosity, the photon trigger identification requirement was tightened in the 2017–2018 data-taking period. The average trigger efficiency is over 98% for events passing the full diphoton event selection for this analysis.

The EFT definition used in this Letter is provided by the Higgs characterization model [21], which is implemented in the MadGraph5_AMC@NLO generator [22]. Within

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3.
this model, the term in the effective Lagrangian that describes the top Yukawa coupling is

\[ \mathcal{L} = -\frac{m_t}{v} \left[ \bar{\psi} \gamma_i \frac{\kappa}{c} \cos(\alpha) + i \sin(\alpha) \gamma_5 \right] \psi H \]

where \( m_t \) is the top quark mass, \( v \) is the Higgs vacuum expectation value, \( \kappa_i \) (\( i > 0 \)) is the top Yukawa coupling parameter, and \( \alpha \) is the CP-mixing angle. The SM corresponds to a CP-even coupling with \( \alpha = 0 \) and \( \kappa_i = 1 \), while a CP-odd coupling is realized when \( \alpha = 90^\circ \).

Simulated \( t\bar{t}H \) and \( tH \) samples were generated using MadGraph5_AMC@NLO 2.6.2 at next-to-leading order in QCD for different \( \alpha \) and \( \kappa_i \) for \( tH \) values, with the NNPDF30NLO [23] parton distribution function (PDF) set used for the matrix element (ME) evaluation, and interfaced to PYTHIA 8 [24] using the NNPDF23LO [25] PDF set for parton showering (PS). The A14 parameter set [26], tuned to data, was used for both PS and underlying event (UE).

From these samples, the yields for \( t\bar{t}H \) and \( tH \) are parameterized as functions of \( \alpha \) and \( \kappa_i \), which are used in the statistical interpretations. Samples for other Higgs boson production processes, \( ggF \) [27], vector-boson fusion (VBF) [28], and vector-boson associated production (\( VH \)) [29,30] were produced with POWHEG-BOX v2 generator [31] using the PDF4LHC15 PDF set [32] for ME, with the A14 PDF set [33] and PYTHIA 8 for PS using the CTEQ6L1 [34] PDF set. Samples generated with Herwig 7 [35] are used for systematic uncertainty studies that involve modeling of the PS, hadronization and UE. The simulated Higgs boson samples are normalized to the SM cross sections (Refs. [36–54]) times the SM branching ratio (BR) to diphotons (Refs. [36,55–58]) with a Higgs boson mass of 125.09 GeV [59], and specifically for \( t\bar{t}H \), the SM predicted cross section times the \( H \to \gamma\gamma \) BR is \( \sigma_{t\bar{t}H} \times B_{\gamma\gamma} = 1.15^{+0.09}_{-0.12} \) fb.

Although this analysis relies on a data-driven approach for background estimations, a simulated background sample for the \( t\bar{t}\gamma\gamma \) process was generated to optimize the event selection and develop the background model. This sample was generated using the MadGraph5_AMC@NLO generator, with the NNPDF23LO PDF set and showered with PYTHIA 8.

All generated Higgs boson events were passed through a full simulation of the ATLAS detector response [60] using GEANT 4 [61]. The \( t\bar{t}\gamma\gamma \) events were processed with a fast simulation in which the full simulation of the calorimeter is replaced with a parameterization of the calorimeter response [62]. The effects of multiple \( pp \) interactions in the same or neighboring bunch crossings are included using events generated with PYTHIA 8. Events are weighted such that the distribution of the average number of interactions per bunch crossing matches that observed in data, which is typically around 30 to 40.

Events are selected by requiring two isolated photon candidates with transverse momenta \( p_T \) greater than 35 GeV and 25 GeV. Both photons must satisfy the tight identification requirement [20]. The identification is constructed from a cut-based selection using the electromagnetic shower shape variables. The leading (subleading) photon must have \( p_T / m_{\gamma\gamma} > 0.35 \) (0.25), and the diphoton invariant mass \( m_{\gamma\gamma} \) is required to be in the range \( m_{\gamma\gamma} \in [105,160] \) GeV. Jets are reconstructed using the anti-\( k_t \) algorithm [63] with a radius parameter of \( R = 0.4 \). Events are required to have at least one jet with \( p_T > 25 \) GeV containing a \( b \)-hadron (\( b \)-jet), identified using a \( b \)-tagging algorithm with an efficiency of 77% and a mistagging rate of 0.9% for light-flavor jets [64].

Selected events are sorted into two \( t\bar{t}H \)-enriched regions. The \( \text{“Lep”} \) region, targeting top quark decays in which at least one of the resulting \( W \) bosons decays leptonically, requires events to have at least one isolated lepton (muon or electron) candidate with \( p_T > 15 \) GeV passing medium identification requirements (Refs. [20,65]). The \( \text{“Had”} \) region targets hadronic top quark decays (as well as top quark decays to both hadronically decaying \( \tau \) leptons and unreconstructed leptons) and requires events to have at least two additional jets with \( p_T > 25 \) GeV and no selected lepton.

A boosted decision tree (BDT) used for the top quark reconstruction, denoted by “Top Reco BDT,” is trained with the \( t\bar{t}H \) sample by using the xgboost package [66] to extract the three-jet (triplet) combination best matching the hadronic decay products of a top quark. This BDT uses \( p_T, \eta, \phi \), and the energy \( E \) of \( W \) boson and \( b \) jet (where the \( W \) boson candidate is formed by a pair of jets). Furthermore, this BDT uses the angular distance \( \Delta R_{Wb} \) between the \( W \) boson and \( b \) jet, \( \Delta R_{jj} \) between the two jets composing the \( W \) boson candidate, and \( b \)-tagging information about all three jets in the triplet and the invariant mass of the triplet. For events in the Had region, the triplet with the highest Top Reco BDT score is taken as the primary top quark candidate \( t_1 \). In the Lep region, for events containing only one lepton, a \( W \) boson candidate is first constructed from the lepton and missing transverse momentum \( E_T^{\text{miss}} \). Then \( t_1 \) is reconstructed from this leptonic \( W \) boson candidate and the jet giving the highest Top Reco BDT score. No top quark candidate is reconstructed for events containing more than one lepton. After \( t_1 \) is selected, if there are at least three additional jets, a second top quark candidate \( t_2 \) is reconstructed by selecting the triplet with the highest BDT score from the remaining jets; if there is only one or two additional jets, then \( t_2 \) is taken as the sum of the remaining jets; otherwise no \( t_2 \) is reconstructed.

To improve the analysis sensitivity, selected events are categorized using partitions of a two-dimensional BDT space. Two independent BDTS are trained using the xgboost algorithm: “Background Rejection BDT” and “CP BDT,” and each of them is trained separately in the Had and Lep regions. The Background Rejection BDT is trained to separate \( t\bar{t}H \)-like events from background that
are mainly nonresonant diphoton production processes, including $\gamma\gamma$ + jets and $t\bar{t}\gamma\gamma$. A detailed discussion of this methodology is given in Ref. [1]. The CP BDT is trained to separate CP-even from CP-odd couplings using $t\bar{t}H$ and $tH$ processes. The CP BDT uses $p_T$ and $\eta$ of the diphoton system, $p_T$ and $\eta$ of $t_1$ and $t_2$, their azimuthal angles calculated relative to the diphoton system $\phi_{t_1t_2}$, $\phi_{t\bar{t}t\bar{t}}$ as well as their Top Rec BDT scores. It also uses differences in pseudorapidity and azimuthal angle $\Delta\eta_{t_1t_2}$ and $\Delta\phi_{t_1t_2}$ between the two top quark candidates, the invariant mass of the diphoton and primary top quark system $m_{t\bar{t}t\bar{t}}$, the invariant mass of the two top quark candidates $m_{t_1t_2}$, the scalar $p_T$ sum of jets $H_T$, the $E_T^{\text{miss}}$ divided by $\sqrt{H_T}$, the number of jets and $b$-tagged jets, and the minimum and second smallest angular differences $\Delta R_{tj}$ between a photon and a jet.

Figure 1 shows the BDT discriminant distributions in the data as well as those expected from CP-even and CP-odd Higgs boson signals in the Had region. The discriminating power can be seen by comparing the CP-even, CP-odd, and data shapes. Events with low values of the Background Rejection BDT response are removed, and the remaining events are categorized. The number of categories and the boundary locations are chosen to optimize the $t\bar{t}H$ significance and the discriminating power between the CP-even and CP-odd cases. There are 20 categories in total: 12 in the Had region and 8 in the Lep region.

The results are impacted by three distinct types of uncertainties: the statistical uncertainty associated with the data, theoretical modeling systematic uncertainties, and experimental systematic uncertainties. The first dominates. Theoretical uncertainties for $t\bar{t}H$ and $tH$ rates in the various categories are assessed. The following effects are considered: the value of the strong coupling constant; alternative generator for the PS, hadronization, and UE; and PDF uncertainty. In the three (two) most CP-even sensitive Had (Lep) categories, each of these uncertainties is less than 10%. The background from ggF is less than 0.25 events in each of the most sensitive categories; conservative uncertainties, including a 100% theoretical uncertainty in the modeling of the radiation of additional heavy-flavor jets, are assigned to it in the Had region. The same heavy-flavor uncertainty is also assigned to the VBF and VH processes.

Experimental uncertainties arise from identification and isolation criteria for photons, electrons, and muons and from their energy scale and resolution [20,65]. Jets have uncertainties from $b$ tagging [64] and vertex identification [67] in addition to the energy scale and resolution [68]. Uncertainties in the measurement of $E_T^{\text{miss}}$ [69], which is used in the leptonic categories, are also included. These experimental effects impact the expected event yield in each category and can cause events to migrate between the categories. The overall uncertainty is less than 20% in each category. In addition, uncertainties in the luminosity [70] obtained using the LUCID-2 detector [71] and trigger efficiency [19] are responsible for uncertainties in the overall event yield of 1.7% and 0.4%, respectively.

A simultaneous maximum-likelihood fit is performed to the $m_{\gamma\gamma}$ spectra in all the categories. Signal and background shapes are modeled by analytic functions using the strategy discussed in Ref. [6]. The chosen background function is

![Graph](image)

FIG. 1. Left: two-dimensional BDT distribution in the selected data events ($m_{\gamma\gamma} \in [105, 160]$ GeV) from the Had region showing the Background Rejection BDT and CP BDT. The inner (outer) contours capture 25% (50%) of the $t\bar{t}H$ and $tH$ signal events for CP-even (blue) and CP-odd (red) hypotheses. Right: projections onto the background rejection and CP BDT axes. Contributions from CP-even (blue) and CP-odd (red) $t\bar{t}H/tH$ processes and the data (black) are shown and normalized to unit area. The error bars on data are statistical.
based on the simulated $t\bar{t}\gamma\gamma$ events following the procedure in Ref. [1], which imposes stringent conditions on potential biases in the extracted signal yield to avoid losses in sensitivity. The parameters of the background model and background normalization in each category are left free in the fit. The profile likelihood ratio is used as the test statistic, and the asymptotic approximation [72] is used for statistical interpretations. Yields from $t\bar{t}H$ and $tH$ are extracted after subtracting the very small contribution from other Higgs boson production modes using their SM expected values. Figure 2 shows the distributions of the reconstructed masses for the diphoton system and primary top quark. The events are weighted by $\ln(1 + S/B)$ with $S$ and $B$ being the fitted signal and background yields in the smallest $m_{\gamma\gamma}$ interval containing 90% of the signal in each category. The $p$ value associated with the compatibility between the observed spectra and the fit model using the goodness-of-fit test method described in Ref. [73] is 35%. Assuming a $CP$-even coupling, the $\sigma_{t\bar{t}H} \times B_{\gamma\gamma}$ is derived by constraining all the non-$t\bar{t}H$ Higgs boson processes to their SM predictions and measured to be $1.64^{+0.38}_{-0.36} (\text{stat})^{+0.17}_{-0.14} (\text{sys})$ fb. The measured rate for $t\bar{t}H$ is $1.43^{+0.33}_{-0.31} (\text{stat})^{+0.21}_{-0.15} (\text{sys})$ times the SM expectation. The background-only hypothesis is rejected with an observed (expected) significance of 5.2$\sigma$ (4.4$\sigma$). The rate for $tH$ is derived by constraining all the non-$t\bar{t}H/tH$ Higgs boson processes to their SM prediction without prior constraint on the rate of $t\bar{t}H$. Using the CLs method [74], this yields a 95% confidence level (CL) upper limit of 12 times the SM prediction, the same as expected assuming the presence of SM $tH$ signal. This is stricter than the previous best limit of 25 times the SM prediction on $tH$ from the CMS analysis performed using 35.9 fb$^{-1}$ of data at $\sqrt{s} = 13$ TeV [75] with the $t\bar{t}H$ process constrained to the SM prediction.

Extraction of values for the top Yukawa coupling requires additional information. In particular, the BR of $H \rightarrow \gamma\gamma$ is needed to recover the total Higgs boson production rate, and the Higgs boson coupling to gluons is needed to account for the small ggF background. The corresponding Higgs boson coupling modifiers $\kappa_\gamma$ and $\kappa_g$ are measured in the Run 2 Higgs boson combination [76]. This combination includes the first 80 fb$^{-1}$ of data used in this paper, and $t\bar{t}H$ and $tH$ analyses from other decay channels. The combination analysis is repeated without the $t\bar{t}t$ and $tH$ inputs, and this result is used to constrain $\kappa_\gamma$ and $\kappa_g$. The impact on $\kappa_\gamma$ and $\kappa_g$ of removing input $t\bar{t}H$ and $tH$ analyses from the combination is small. The correlation of the systematic uncertainties between the Higgs boson coupling combination and this analysis is neglected. The correlation has a small impact on $\alpha$, and a similar effect on $\kappa$, as on signal strength reported in Ref. [76]. This analysis is insensitive to the potential modifications of ggF kinematics due to $CP$ mixing, which is therefore neglected. The results of the fit for $\kappa_\gamma \cos(\alpha)$ and $\kappa_\gamma \sin(\alpha)$ are shown as contours in Fig. 3. A limit on $\alpha$ is set without prior constraint on $\kappa_\gamma$ in the fit: $|\alpha| > 43^\circ$ is excluded at 95% CL. The expected exclusion is $|\alpha| > 63^\circ$ under the $CP$-even hypothesis. A value of $\alpha = 90 (180)^\circ$ is excluded at 3.9$\sigma$ (2.5$\sigma$). A comparable study from the CMS experiment excluded $\alpha = 90^\circ$ at 3.2$\sigma$ [3]. If $\kappa_\gamma$ and $\kappa_g$ are parameterized using $\alpha$ and $\kappa_\gamma$, the observed (expected) exclusion is $|\alpha| > 43 (56)^\circ$ without prior constraint on $\kappa_\gamma$ in the fit. The impact of the systematic uncertainties is negligible.
FIG. 3. Two-dimensional likelihood contours for $\kappa, \cos(\alpha)$ and $\kappa, \sin(\alpha)$ with $ggF$ and $H \rightarrow \gamma\gamma$ constrained by the Higgs boson coupling combination.

In summary, the production rate of the Higgs boson in association with top quarks is measured, and the $CP$ property of the top Yukawa coupling is studied. The no-$t\bar{t}H$ hypothesis is rejected with a significance of $5.2\sigma$, and the measured $\sigma_{t\bar{t}H} \times B_{\gamma\gamma} = 1.64^{+0.36}_{-0.31} (\text{stat})^{+0.17}_{-0.15} (\text{sys})$ fb. The measured rate for $t\bar{t}H$ is $1.43^{+0.33}_{-0.21} (\text{stat})^{+0.17}_{-0.15} (\text{sys})$ times the SM expectation. The $t\bar{t}H$ process is not observed, and an upper limit of 12 times the SM expectation is set on its rate at 95% CL. All measurements are consistent with the SM expectations, and the possibility of $CP$-odd couplings between the Higgs boson and top quark is severely constrained. A pure $CP$-odd coupling is excluded at 3.9$\sigma$, and $|\alpha| > 43^\circ$ is excluded at 95% CL.

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; STFC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC, and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST, and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR, and VSC CR, Czech Republic; DFG and DZIF, Germany; INFN, Italy; INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK), and BNL (USA), the Tier-2 facilities worldwide, and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [77].
K. Mimasu, V. Sanz, and C. Williams, Higher order QCD

P. Nason and C. Oleari, NLO Higgs boson production via

F. Boudjema, D. Guadagnoli, R. M. Godbole, and K. A.

K. Hamilton, P. Nason, E. Re, and G. Zanderighi, NNLOPS

ATLAS Collaboration, ATLAS Pythia8 tunes to 7 TeV data,


ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the center of the detector and the z axis along the beam pipe. The x axis points from the interaction point to the center of the LHC ring, and the y axis points upwards. Cylindrical coordinates ($r$, $\phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the z axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.


G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, $HH^*/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO, J. High Energy Phys. 10 (2013) 083.


[64] ATLAS Collaboration, ATLAS \( b \)-jet identification performance and efficiency measurement with \( t\bar{t} \) events in \( pp \) collisions at \( \sqrt{s} = 13 \) TeV, Eur. Phys. J. C 79, 970 (2019).


[76] ATLAS Collaboration, Combined measurements of Higgs boson production and decay using up to 80 fb\(^{-1}\) of proton-proton collision data at \( \sqrt{s} = 13 \) TeV collected with the ATLAS experiment, Phys. Rev. D 101, 012002 (2020).

Institute for Theoretical and Experimental Physics named by A.I. Alikhanov of National Research Centre "Kurchatov Institute", Moscow, Russia