Baku State University: Journal of Physics & Space Sciences, **2025**, v2 (4), pp. 23-31

journal homepage: http://bsuj.bsu.edu.az/en

The CP structure of the top quark-Higgs Yukawa coupling is probed through an analysis of ttH/tH events

Nazim A. Huseynov*, Anastasiya D. Tropina

CERN, Geneva, Switzerland

Received 01-Sep-2025; Accepted 21-Oct-2025

DOİ: https://doi.org/10.30546/209501.101.2025.2.04.0108

Abstract

Investigation of the interaction between the Higgs boson and the top quark is of high priority interest. The large mass of the top quark of approximately 173.3 GeV requires that it couples strongly to the Higgs boson in the SM. Consequently, the top quark is suspected to play a special role in the electroweak symmetry breaking. The Yukawa coupling of the top quark can be measured directly through the inclusive cross-section of the associated production of the Higgs boson with a top-quark pair ($t\bar{t}H$). The measurement of the Higgs boson production cross-section in association with a single top quark (tH) and the kinematic properties of tH and $t\bar{t}H$ events can also provide information about the CP nature of the coupling. In the SM the Higgs boson is a scalar and its interactions are CP-even. The pure pseudoscalar hypothesis with CP-odd interactions with weak vector bosons and fermions has been excluded. However, it remains experimentally allowed that the Higgs boson is a CP-mixed state, which arises in extended Higgs sectors, and would provide a new source of CP-violation beyond the SM.

Keywords: Higgs boson; top quark; Yukawa coupling; CP-violation;
PACS Numbers: 14.65.Ha; 14.80.Cp; 14.80.-j;

1. Introduction

Following the experimental confirmation of the Higgs boson in 2012 [1,2], the primary research objective has transitioned to the precise determination of its properties. A main task is the quantitative assessment of its diverse production

*Corresponding author – Tel.: +41 22 76 67376
e-mail: nazim.huseynov@cern.ch; ORCID ID: 0000-0003-2201-5572

cross sections and decay branching fractions, which are subsequently evaluated against the corresponding Standard Model (SM) theoretical expectations.

The Higgs boson couples to fermions through Yukawa interactions, with the coupling to the top quark (y_t) being the largest. This near-unity magnitude renders y_t critically influential in several key theoretical domains. It dominates radiative corrections to the Higgs potential, thereby affecting the potential's geometry and the stability of the electroweak vacuum. Furthermore, y_t significantly modifies the renormalization group evolution of SM parameters, which alters the energy scale associated with potential new physics phenomena.

Consequently, the top-quark Yukawa coupling serves as a sensitive probe for physics beyond the Standard Model. In particular, precise investigations of its charge-parity (CP) structure may offer crucial insights into the mechanisms responsible for the observed baryon asymmetry of the universe. Within the SM, the Higgs boson is a scalar particle with purely CP-even couplings. While experimental results have excluded a pure pseudoscalar state with CP-odd interactions with fermions and weak bosons, the possibility of a CP-mixed state remains experimentally allowed. Such a state, predicted in models with extended Higgs sectors, would introduce a new source of CP violation beyond the SM.

Among the SM production modes, the associated production of a Higgs boson with a single top quark (tH) constitutes the fifth most probable channel. Its predicted cross section, approximately 74 fb [3], is suppressed by an order of magnitude relative to the fourth-ranked $t\bar{t}H$ associated production mode. This significant suppression is theoretically attributed to a destructive quantum interference between the two leading-order Feynman diagrams (Figure 1).

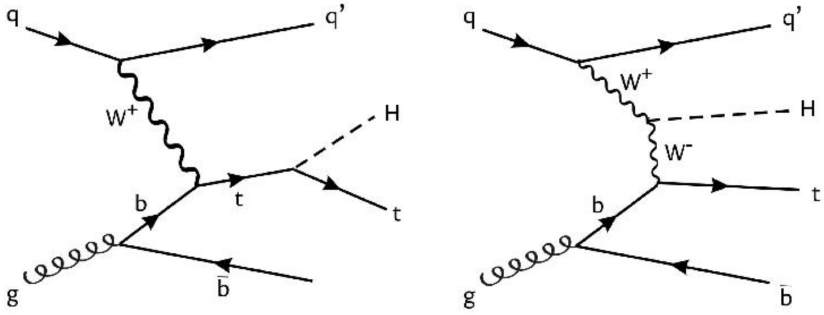


Fig. 1. The leading Feinman diagrams for the $\boldsymbol{pp \to tH}$ production.

While its magnitude is consistent with SM predictions — as established through $t\bar{t}H$ measurements — the sign of the coupling remains experimentally unconstrained. The associated production of a Higgs boson with a single top quark (tH) provides a unique probe this sign. The simplest minimal extension to the SM is the

Inverted Top Coupling (ITC) model, in which $y_t^{ITC} = -y_t^{SM}$.In this scenario, the destructive interference that suppresses the SM tH cross-section becomes constructive. This results in a predicted cross-section of approximately 848 fb [3], over an order of magnitude larger than the SM expectation. Consequently, searches for tH production constitute a highly sensitive probe for physics beyond the SM.

2. Analysis strategy

The analysis uses a dataset corresponding to an integrated luminosity of $140\ fb^{-1}$ collected by the ATLAS detector at a center-of-mass energy of $\sqrt{s} = 13\ TeV$ during the LHC Run 2 (2015–2018). tH production channel is studied separately and in the simultaneous measurement of the $t\bar{t}H$ and tH signal strengths.

At the LHC, tH proceeds via three electroweak production modes: $tHbq$ (t-channel), tWH, and the s-channel. Combined $t\bar{t}H$ and tH analysis targets the $tHbq$ process, as it possesses the largest cross section and exhibits a more distinct experimental signature from the dominant $t\bar{t}H$ background compared to the tWH mode. In separate tH analysis tWH mode is also accounted for. The s-channel, characterized by a negligible production rate, is not considered in both analyses.

2.1. tH analysis

The search is organized into three orthogonal analysis channels, defined by the multiplicity of light-flavor leptons (electrons or muons) in the final state:

- One-lepton ($1l$) channel
- Two same-sign leptons ($2lSS$) channel
- Three-lepton ($3l$) channel

Events containing hadronically decaying tau leptons are vetoed across all channels. Higgs boson decays are primarily targeted via the $H \rightarrow b\bar{b}$ mode in the 1l channel and the $H \rightarrow WW^*$ mode in the multilepton ($2lSS$ and $3l$) channels. The multilepton channels also include sub-leading contributions from $H \rightarrow \tau\tau$ and $H \rightarrow ZZ^*$ decays, incorporated through their light-lepton final states.

Due to their distinct background compositions, the channels are analyzed independently and combined afterwards. Each channel is subdivided into multiple regions: a Signal Region (SR) optimized for signal sensitivity, and several Control Regions (CRs) used to constrain the dominant backgrounds. Specifically: the 1l analysis employs one SR and one CR dedicated to constraining $t\bar{t}$ background. The 2ℓSS and 3ℓ analyses each utilize one SR, three CRs to constrain backgrounds from non-prompt and misidentified leptons, and one CR targeting $t\bar{t}W$ production.

The analysis employs multivariate techniques, specifically Boosted Decision Trees (BDTs), across all channels to optimize the discrimination between the tH signal and background processes. These classifiers are utilized both to define the analysis regions and to enhance the final sensitivity in the statistical inference. The results derived from the BDT-based approach were validated by a complementary analysis employing sequential cuts on key kinematic variables (the cut-and-count, or C&C, analysis). The C&C analysis yielded pre-fit signal significances of 0.224σ for the SM tH prediction and 3.63σ for the ITC hypothesis, improving upon significances of 0.156σ and 2.45σ obtained in the initial preselection regions, respectively. Given its superior performance, the BDT-based analysis was selected to derive the final results.

The signal strength parameter, defined as $\mu_{tH} = \sigma_{measured}/\sigma_{predicted}$, is extracted via a profile likelihood fit. This fit is performed simultaneously across all analysis regions and channels, allowing for the simultaneous estimation of μ_{tH} and the data-driven constraint of the normalization for dominant background processes. The systematic uncertainties are incorporated in the likelihood function through nuisance parameters The results are derived first from channel-specific fits and subsequently from a combined fit of all channels.

2.2. Combined tH and $t\bar{t}H$ analysis

The analysis is performed by statistically combining six independent channels. The channels are defined based on the multiplicities of hadronically decaying tau leptons (τ_{had}) and light leptons. The channel composition is as follows: two channels with zero τ_{had} ($2lSS0\tau_{had}$ and $3l0\tau_{had}$), one single-τ_{had} channel ($2lSS1\tau_{had}$), two channels with two-τ_{had} channels ($1l2\tau_{had}$ and $2l2\tau_{had}$), and one channel with no explicit τ_{had} requirement ($4l$). The signatures target $t\bar{t}H$ production where at least one top quark decays leptonically, and the Higgs boson decays via $H \to WW^*$, $H \to \tau\tau$ or $H \to ZZ^*$, with final states containing combinations of light leptons and τ_{had}. This section is focused on $2lSS0\tau_{had}$, $3l0\tau_{had}$ and $1l2\tau_{had}$ channels where the event selection criteria were also optimized to retain significant fraction of tH events.

Multivariate analysis (MVA) techniques, based on either Boosted Decision Trees (BDTs) or deep neural networks (DNNs), are implemented to enhance the discrimination between signal and background events. The distributions of input variables used for MVA training were validated against data, with no significant mismodelling observed in the simulation. The classifier output scores are employed to define orthogonal event categories. Categories exhibiting high signal ($t\bar{t}H$ or tH) purity are designated as signal regions (SRs), while those dominated by specific background processes are utilized as control regions (CRs) to constrain background normalizations and shapes.

The differential cross-section measurement is performed within the Simplified Template Cross-Section (STXS). For $t\bar{t}H$ category this formalism defines six bins based on the generator-level Higgs boson transverse momentum (p_t^H), labeled STXS

1-6, corresponding to p_t^H intervals of 0-60 GeV, 60-120 GeV, 120-200 GeV, 200-300 GeV, 300-450 GeV, and >450 GeV. To assign events to these STXS bins, reconstructed p_t^H estimators are employed. In the $2lSS0\tau_{had}$ and $3l0\tau_{had}$ channels, a Graph Neural Network (GNN) is trained to estimate p_t^H. For the $1l2\tau_{had}$ channel a Boosted Decision Tree (BDT) is utilized for the same purpose. The outputs of these algorithms are used to categorize events into the corresponding STXS regions for the differential measurement.

A simultaneous profile-likelihood fit to all signal and control region bins is performed to simultaneously measure the $t\bar{t}H$ and tH cross sections differentially in the STXS bins and to constrain the normalizations of the major background processes. Systematic uncertainties are incorporated into the likelihood model via constrained nuisance parameters.

2.3. Search for CP-violation

A search for CP violation in the top–Higgs interaction is conducted using the previously established region definitions. This search employs a dedicated fit framework parameterized to accommodate a potential mixture of CP-odd and CP-even Higgs boson states. Such a mixture would modify both the cross sections and the kinematic properties of $t\bar{t}H$, $tHbq$, and tWH production processes.

Indirect, model-dependent constraints on the CP structure of the top-quark Yukawa coupling can be derived from complementary measurements, including searches for the electron electric dipole moment, the gluon-fusion Higgs production cross section, and the $H \to \gamma\gamma$ decay rate. To test this physics scenario, the Standard Model top–Higgs interaction Lagrangian is extended with an additional CP-odd component, expressed as [4]:

$$\mathcal{L}_{t\bar{t}H} = -\kappa_t' y_t \phi \overline{\psi_t}(\cos\alpha + i\gamma_5 \sin\alpha)\psi_t, \qquad (2)$$

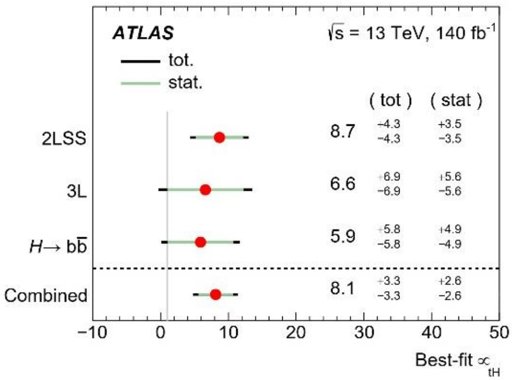
where y_t is the SM Yukawa coupling, modified by a coupling modifier κ_t', α is the CP-mixing angle, ϕ is the Higgs field, ψ_t and $\overline{\psi_t}$ are top-quark spinor fields and γ_5 is a Dirac matrix. In the SM case $\kappa_t' = 1$ and $\alpha = 0°$. Pure CP-odd hypotheses correspond to $\kappa_t' = 1$ and $\alpha = 90°$.

The parameters α and κ_t' are determined via a binned profile-likelihood fit, where the likelihood function depends on the expected signal yields for $t\bar{t}H$, $tHqb$, and tWH processes, which are themselves functions of α and κ_t'. By evaluating the test statistic on a grid of these parameters, two-dimensional exclusion contours in the (κ_t', α) plane are derived.

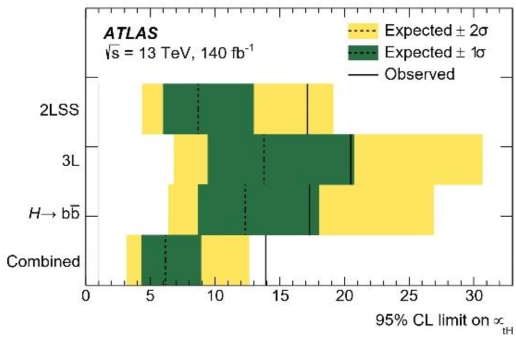
This analysis utilizes only the $2\ell SS0\tau_{\text{had}}$, $3\ell 0\tau_{\text{had}}$, $1\ell 2\tau_{\text{had}}$, and $2\ell 2\tau_{\text{had}}$ channels. For the $2\ell SS0\tau_{\text{had}}$ and $3\ell 0\tau_{\text{had}}$ channels, the same discriminants used in the STXS measurement are employed for the CP analysis. In the $1\ell 2\tau_{\text{had}}$ and $2\ell 2\tau_{\text{had}}$ channels, dedicated BDTs are trained to distinguish between CP-even and CP-odd signal hypotheses. Separate BDT classifiers are constructed for $t\bar{t}H$ and $tHqb$ processes.

3. Results and conclusions

A primary result of the separate tH analysis is the measurement of the tH signal strength, μ_{tH}. Assuming the SM hypothesis, was measured $\mu_{tH}^{\text{SM}} = 8.1 \pm 2.6\,(\text{stat.}) \pm 2.0\,(\text{syst.})$ (Figure 2a), corresponding to a cross section of 720 ± 270 fb. The observed excess has a significance of 2.8σ over the background-only hypothesis. At the 95% confidence level, tH production cross sections exceeding 13.9 times the SM prediction were excluded (Figure 2b).



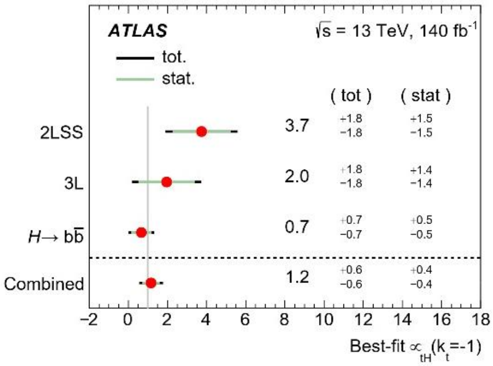
(**a**)



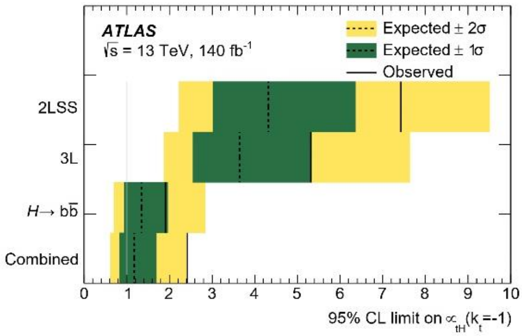
(**b**)

Fig. 2. Results under SM hypothesis: (a) The best-fit values for the Standard Model signal strength, $\boldsymbol{\mu_{tH}^{SM}}$. (b) The observed 95% confidence level (CL) upper limit. The observed limits (solid black lines) are shown alongside the expected limits (dotted black lines) derived under the background-only hypothesis [5].

For the ITC hypothesis, was obtained a signal strength of $\mu_{tH}^{\text{ITC}} = 1.2 \pm 0.4\,(\text{stat.}) \pm 0.5\,(\text{syst.})$ (Figure 3a), equivalent to a cross section of 1020 ± 540 fb. At the 95% confidence level, cross sections larger than 2.4 times the ITC prediction are excluded (Figure 3b). While the data exhibit a mild preference for the ITC interpretation, both the SM and ITC signal hypotheses remain compatible with the observed excess within the measured uncertainties.



(**a**)



(**b**)

Fig. 3. Results under ITC hypothesis: (a) The best-fit values for the Standard Model signal strength, $\boldsymbol{\mu_{tH}^{SM}}$. (b) The observed 95% confidence level (CL) upper limit. The observed limits (solid black lines) are shown alongside the expected limits (dotted black lines) derived under the background-only hypothesis [5].

For combined $t\bar{t}H$ and tH analysis a simultaneous fit to all channels, with the $t\bar{t}H$ and $tHqb$ signal strengths ($\mu_{t\bar{t}H}$ and μ_{tHqb}) as free parameters and the tWH contribution fixed to its SM prediction, yields the results shown in Figure 4. The best-fit values are $\mu_{t\bar{t}H} = 0.59^{+0.22}_{-0.20}$ and $\mu_{tHqb} = 7.2^{+4.6}_{-4.0}$ (including uncertainties on the SM predictions). The measured $tHqb$ signal strength is compatible with the dedicated tH result.

3.1. CP interpretation

The resulting exclusion contours in the (κ_t', α) parameter plane (Equation 1) are presented in Figure 5. The measurements are consistent with the SM prediction, which corresponds to a purely CP-even state ($\alpha = 0°$). The observed deviation from a pure CP-odd hypothesis is excluded at the 1.8σ level. Owing to the relative flatness of the likelihood function near its minimum, the precise best-fit point is subject to minor numerical variations. At the 68% confidence level, the observed lower limit on the absolute value of the CP-mixing angle is $|\alpha| > 62°$.

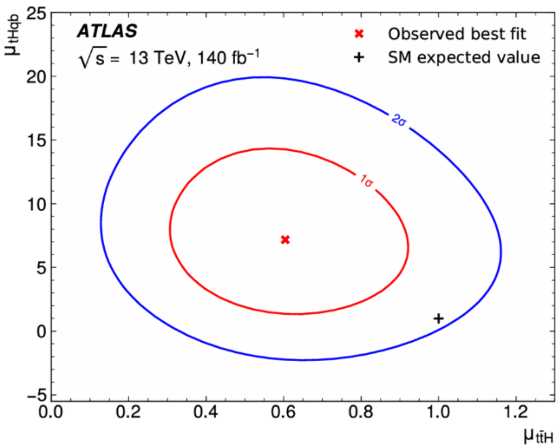
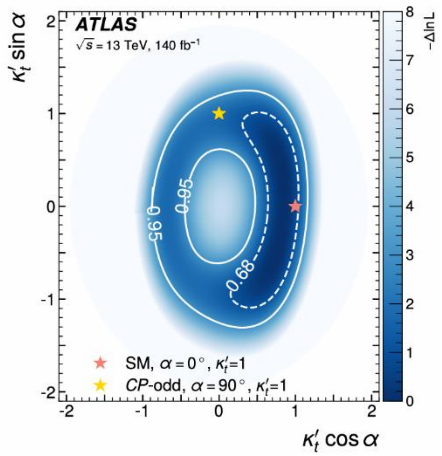


Fig. 4. Observed exclusion contours at the 1σ and 2σ confidence levels in the ($\boldsymbol{\mu_{t\bar{t}H}, \mu_{tHbq}}$) parameter plane. The best-fit value is indicated by a cross. The contours incorporate both experimental uncertainties and theoretical uncertainties on the SM predictions [6].



Fig, 5. Observed exclusion contours in the plane defined by the parameters $\boldsymbol{\kappa'_t \cos\alpha}$ and $\boldsymbol{\kappa'_t \sin\alpha}$. The regions enclosed by the dashed and solid curves correspond to the best-fit result at the 68% and 95% confidence levels, respectively [6].

References

[1] Aad G. et al. [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.* **2012**, *716*, pp. 1–29, 10.1016/j.physletb.2012.08.020

[2] Chatrchyan S. et al. [CMS Collaboration]. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B.* **2012**, *716*, pp. 30–61, 10.1016/j.physletb.2012.08.021

[3] de Florian D. et al. [LHC Higgs Cross Section Working Group Collaboration]. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. *CERN Yellow*

Rep. Monogr. **2017**, 2, pp. 1–869

[4] Demartin F.; Maltoni F.; Mawatari K.; Page B.; Zaro M. Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction. *Eur. Phys. J. C*. **2014**, 74, p. 3065, 10.1140/epjc/s10052-014-3065-2

[5] Aad G. et al. [ATLAS Collaboration]. Search for the production of a Higgs boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. JHEP. 2025, 10, p. 093, 10.1007/JHEP10(2025)093

[6] Aad G. et al. [ATLAS Collaboration]. Measurement of the Higgs boson production in association with top quarks in multilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. JHEP (accepted)