

Machine learning approaches to top quark flavor-changing four-fermion interactions in trilepton signals at the LHC

Meisam Ghasemi Bostanabad[✉] and Mojtaba Mohammadi Najafabadi[✉]

*School of Particles and Accelerators, Institute for Research in Fundamental Sciences (IPM),
P.O. Box 19395-5531, Tehran, Iran*



(Received 13 March 2025; accepted 20 May 2025; published 11 June 2025)

We explore the top quark flavor-changing four-fermion interactions ($tuee$ and $tcee$) with scalar, vector, and tensor structures using machine learning models to analyze trilepton processes at the LHC. The study is performed using $t\bar{t}$ and tW processes, where a top quark decays into $u/c + e^+ + e^-$. The analysis incorporates both reducible and irreducible backgrounds while accounting for realistic detector effects. The dominant backgrounds for these trilepton signatures arise from $t\bar{t}$ production, single top quark production in association with V , and VV production (where $V = W, Z$). These backgrounds are significantly reduced using machine-learning-based classification models, which optimize event selection and improve signal sensitivity. For an integrated luminosity of 3000 fb^{-1} at the LHC, we find that the expected 95% confidence level (CL) limits on the scale of four-fermion flavor-changing neutral current interactions reach $\Lambda \leq 5.5 \text{ TeV}$ for $tuee$ and $\Lambda \leq 5.7 \text{ TeV}$ for $tcee$ in the $t\bar{t}$ channel, and $\Lambda \leq 1.9 \text{ TeV}$ ($tuee$) and $\Lambda \leq 2.0 \text{ TeV}$ ($tcee$) in the tW channel. We also provide an interpretation of our EFT analysis in the context of a specific Z' model, illustrating how the derived constraints translate into bounds on the parameter space of a heavy neutral gauge boson mediating flavor-changing interactions.

DOI: [10.1103/mk8x-nrpn](https://doi.org/10.1103/mk8x-nrpn)

I. INTRODUCTION

With its large mass and strong Higgs boson coupling, the top quark is key to probing flavor physics beyond the Standard Model (SM), particularly in searches for new interactions violating flavor conservation. Its substantial mass makes it particularly sensitive to various forms of new physics (NP), especially in relation to flavor-changing interactions and CP violation. In many scenarios of new physics, flavor-changing neutral current (FCNC) effects involving the top quark are enhanced due to the top quark's mass relative to the scale of the new physics. As a result, studying top quark interactions provides a powerful probe into potential new flavor dynamics beyond the SM. In the SM, FCNCs are forbidden at tree level due to the Glashow-Iliopoulos-Maiani (GIM) mechanism [1], which ensures their suppression. Consequently, FCNC processes can only occur at higher orders through loop-level diagrams and are highly suppressed, leading to branching fractions far below current experimental sensitivities. In particular, transitions like $t \rightarrow u$ and $t \rightarrow c$, either in top decays or in production processes, are extraordinarily rare under SM predictions,

with branching ratios far below current experimental sensitivities.

In the SM, the branching ratios for FCNC top quark decays, $t \rightarrow u/c + \ell^+ + \ell^-$, are exceedingly small, on the order of 10^{-14} to 10^{-13} [2,3]. In the SM, the loop amplitudes for these decays are dominated by contributions from down-type quarks, with the bottom quark (b) playing the most significant role. Consequently, the scale of these amplitudes is set by the bottom quark mass m_b . For instance, the partial widths for $t \rightarrow cZ$ decays can be expressed as [2,3]

$$\Gamma(t \rightarrow Z + c) \sim |V_{tb}^* V_{cb}|^2 \alpha_{\text{em}} \times g^4 \times m_t (m_b/m_W)^4 \times \mathcal{F},$$

where α_{em} represents the fine structure constant, g is the weak coupling constant, and V_{tb}, V_{cb} are CKM matrix elements. $\mathcal{F} \sim (1 - m_Z^2/m_t^2)^2$ is a phase-space and polarization factor obtained by neglecting the charm quark mass m_c . The factor $(m_b/m_W)^4$ arises from the GIM mechanism, which introduces an additional suppression beyond naive expectations based on dimensional analysis, power counting, and CKM matrix elements. This fourth power of the mass ratio is responsible for the extreme suppression of these decays, as the GIM mechanism cancels out leading-order contributions and leaves only terms proportional to the mass differences of the down-type quarks in the loop. This suppression of FCNC effects within the SM implies that even the smallest deviation from these expectations,

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) 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.

either in direct measurements or in indirect observables, may provide compelling evidence for new physics beyond the SM. Top quark FCNC processes are thus of great interest, as they offer a clean signal for exploring potential NP effects that may not be observable in lighter quark sectors. Theoretical and experimental efforts have been dedicated to the study of top quark FCNCs, both within model-independent effective field theory (EFT) frameworks and in the context of specific extensions to the SM, such as models involving new heavy gauge bosons, extra dimensions, or supersymmetry [2–26].

In this study, we investigate the FCNC processes described by the effective Lagrangian which will be described in the next section. Our analysis probes two key channels: top pair production, where one of the top quarks decays into a light up-type quark (u or c) and two same-flavor, opposite-sign charged electrons; and single top quark production in association with a W boson, where the top quark undergoes a flavor-changing decay as described by the effective Lagrangian. In the top quark pair production case, the semileptonic decay of the other top quark is considered, and for the tW production process, the leptonic decay of the W is probed. Therefore, the final state is characterized by a trilepton signature accompanied by missing transverse momentum and additional jets. This focus on trilepton final states is strategic, as these states provide a promising environment for observing FCNC processes with minimal SM background interference and reduced sensitivity to systematic uncertainties. The analysis is further refined by employing advanced machine learning techniques, enhancing signal-to-background discrimination through optimized classification algorithms. These methods significantly boost sensitivity, enabling a clearer distinction of rare FCNC-driven top quark processes from dominant SM backgrounds, thus advancing the precision measurement of these elusive interactions.

The structure of the paper is as follows: Section II is dedicated to describing the theoretical framework and the flavor-changing four-fermion interactions of the top quark, along with their experimental constraints. Section III provides a detailed description of the signal and background process definitions and their generation. In Sec. IV, we present the machine learning algorithms employed to achieve optimal discrimination between signal and background processes. The statistical methods used to derive exclusion limits, along with the final results, are discussed in Sec. V. In Sec. VI, an interpretation of the EFT analysis in the context of a specific flavor-changing Z' model is presented. Lastly, a summary of findings and conclusions is given in Sec. VII.

II. TOP QUARK FLAVOR-CHANGING FOUR-FERMION INTER-ACTIONS AND EXPERIMENTAL CONSTRAINTS

To study flavor-changing four-fermion interactions involving the top quark, a model-independent approach

is adopted. Such approaches typically focus on parametrizing new physics effects in terms of higher-dimensional operators, such as four-fermion interactions, that modify the SM Lagrangian. These operators are suppressed by the new physics scale Λ , but the large mass of the top quark significantly enhances their potential observability. In this study, we adopt the parametrization from Refs. [27–30] for the effective Lagrangian describing flavor-changing $t\bar{u}\ell^+\ell^-$ contact interactions, given by

$$\mathcal{L}_{t\bar{u}\ell\ell} = \frac{1}{\Lambda_\ell^2} \sum_{i,j=L,R} [V_{ij}^\ell (\bar{\ell}\gamma_\mu P_i \ell) (\bar{t}\gamma^\mu P_j u) + S_{ij}^\ell (\bar{\ell} P_i \ell) (\bar{t} P_j u) + T_{ij}^\ell (\bar{\ell}\sigma_{\mu\nu} P_i \ell) (\bar{t}\sigma^{\mu\nu} P_j u)], \quad (1)$$

where u represents up and charm quarks; ℓ represents the electron, muon, and τ ; and the different couplings are given by V_{ij}^ℓ (vector-like), S_{ij}^ℓ (scalar-like), and T_{ij}^ℓ (tensor-like). In this formulation, the Wilson coefficients V_{ij}^ℓ , S_{ij}^ℓ , T_{ij}^ℓ , and the new physics scale Λ_ℓ are assumed to be different for different lepton flavors. This assumption is well motivated by both theoretical and experimental considerations. Many BSM scenarios predict nonuniversal couplings across different lepton generations, potentially leading to lepton flavor universality violation, as suggested by anomalies observed in B physics, which refers to the study of processes involving B mesons, particularly those containing a bottom (beauty) quark. Furthermore, experimental constraints on FCNCs involving different leptons vary significantly. While electron-related FCNCs are tightly constrained by high-precision, low-energy experiments, those involving muons and τ 's are subject to comparatively looser bounds, allowing for different new physics scales across lepton flavors [31–33]. The decay rate for the three-body process $t \rightarrow u + \ell^+ + \ell^-$ is given by

$$\Gamma(t \rightarrow u/c + \ell^+ + \ell^-) = \frac{m_t^5}{6144\pi^3\Lambda_\ell^4} \sum_{i,j=L,R} [|V_{ij}^\ell|^2 \mathcal{I}_V + |S_{ij}^\ell|^2 \mathcal{I}_S + |T_{ij}^\ell|^2 \mathcal{I}_T], \quad (2)$$

where $\mathcal{I}_V = 4$, $\mathcal{I}_S = 1$, and $\mathcal{I}_T = 48$. The interference terms vanish in this approximation, as they are proportional to m_ℓ or m_u . The tensor coupling T_{ij}^ℓ dominates with respect to the vector- and scalar-type couplings, which is mainly due to its momentum-dependent terms and Lorentz structure, which enhance its contribution when integrating over the phase space.

In previous studies, the flavor-changing four-fermion interactions involving top quarks, such as $t\bar{u}\ell\ell$ and $t\bar{c}\ell\ell$ described by the Lagrangian presented in Eq. (1), have been extensively explored as potential signatures of NP beyond the SM [29,30,34–40]. These interactions are hypothesized to arise from high-energy processes involving heavy

mediators, such as new scalar, vector, or tensor fields, which induce higher-dimensional operators in effective field theories. These FCNC interactions introduced in Eq. (1) open up novel production channels at the LHC, such as single top production accompanied by a dilepton pair ($pp \rightarrow t\ell\ell$), or with an additional light jet ($pp \rightarrow t\ell\ell + j$). These processes can manifest in distinct experimental signatures, notably in the form of dilepton + b -jet events ($pp \rightarrow \ell\ell + b\text{-jet} + X$) and trilepton final states ($pp \rightarrow \ell\ell\ell' + X$), where two of the leptons are opposite-sign same-flavor (OSSF) pairs [30]. An early study explored the use of neural networks to search for top-quark FCNCs mediated by scalar interactions, demonstrating the potential of machine learning techniques, although it focused on different final states, as detailed in Ref. [35].

LEP2 analyses set limits on the $t\ell\ell\ell$ and $t\ell\ell\ell'$ four-fermion interactions, constraining the new physics scale Λ_e to approximately 1 TeV [34,36]. However, the LHC has the potential to significantly improve these limits, providing a more sensitive probe of such interactions. For example, analyses of the dimuon channel ($pp \rightarrow \mu\mu + b\text{-jet} + X$) with existing LHC data corresponding to 140 fb^{-1} have provided that 95% CL bounds on the NP scale reach $\Lambda_\mu \leq 5 \text{ TeV}$ for tensor-like interactions and $\Lambda_\mu \leq 3.2 \text{ TeV}$ for vector-like interactions for the $t\mu\mu\mu$ interaction [30].

The precise measurement performed by the MOLLER experiment [41] places stringent constraints on new physics, particularly on four-fermion contact interactions that contribute to electron-electron scattering. From this measurement, the following limits have been derived [24]: $\Lambda_e/\sqrt{c_{tuee}} > 3.0 \text{ TeV}$ and $\Lambda_e/\sqrt{c_{tcee}} > 1.7 \text{ TeV}$ for the $tuee$ and $tcee$ interactions, respectively. Consequently, these constraints impose strong upper limits on the branching ratios of rare top decays, specifically $\text{Br}(t \rightarrow c + e^+ + e^-) \lesssim 3 \times 10^{-7}$ and $\text{Br}(t \rightarrow u + e^+ + e^-) \lesssim 3 \times 10^{-8}$.

Figure 1 presents the branching fraction contours for $t \rightarrow u/c + \ell^+ + \ell^-$, expressed in units of the SM branching fraction, in the plane of vector and tensor coupling strengths. The contour levels are set to 50, 500, and 5000 times the SM value for illustration, where the SM prediction is taken to be approximately 10^{-13} . As observed, the current experimental limits remain significantly above the SM expectation. The excluded regions for the vector and tensor couplings, derived using 140 fb^{-1} of LHC data, are shown by dotted and dashed lines with a yellow-shaded region [30].

Looking ahead to the High-Luminosity LHC (HL-LHC), with a projected dataset of 3000 fb^{-1} , the discovery potential is even more promising. The anticipated 95% CL bounds on the scale of the $t\mu\mu\mu$ tensor (vector) interactions are expected to reach $\Lambda_\mu \leq 7.1 \text{ TeV}$ (4.7 TeV), while for the $t\ell\mu\mu$ interactions, the bounds are projected to be $\Lambda_\mu \leq 2.4 \text{ TeV}$ (1.5 TeV) for tensor (vector) couplings [30]. These enhanced sensitivities demonstrate the LHC's capability to probe the high-energy regime of new flavor-changing physics and place stringent limits on the possible existence of these four-fermion operators.

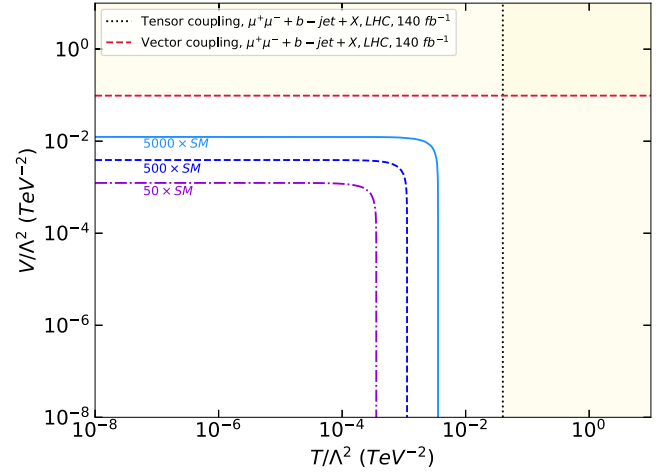


FIG. 1. The branching fraction contours for $t \rightarrow u + \ell^+ + \ell^-$, expressed in units of the SM branching fraction, are shown in the plane of vector and tensor coupling strengths. The contour levels are set to 50, 500, and 5000 times the SM value, which is taken to be 2.8×10^{-13} , for illustration. The excluded regions for tensor and vector couplings, derived using 140 fb^{-1} of LHC data, are indicated by dotted and dashed lines with a yellow-shaded area [30].

In a recent study by the CMS experiment, searches for charged lepton flavor violation (CLFV) in top quark production and decay were reported, focusing on potential four-fermion interactions involving $tue\mu$ and $tce\mu$ [42]. These interactions, which are forbidden in the SM, serve as a sensitive probe for new physics beyond the SM, particularly in scenarios involving FCNCs. The analysis utilized data from proton-proton collisions at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 138 fb^{-1} , collected by the CMS detector at the LHC. Events were selected based on the presence of an oppositely charged electron-muon pair and at least one b -tagged jet, which is a characteristic signature of top quark decays. The analysis employed an EFT framework to parametrize the lepton-flavor-violating interactions of the top quark. Both the production and decay modes of the top quark mediated by these effective operators were included in the search, enabling a comprehensive investigation of CLFV effects in the top sector. For the $tue\mu$ coupling, the 95% CL lower limits on $\Lambda/\sqrt{c}$ are found to be 6.7, 4.8, and 3.3 TeV for the tensor, vector, and scalar couplings, respectively, where c represents the corresponding Wilson coefficient. Similarly, for the $tce\mu$ coupling, the 95% CL lower limits on $\Lambda/\sqrt{c}$ are measured to be 3.4, 2.4, and 1.6 TeV for the tensor, vector, and scalar couplings, respectively.

III. SIMULATED SAMPLES AND SELECTION CRITERIA FOR FCNC EVENTS

In this analysis, two signal processes, $t\bar{t}$ and tW , are examined to probe the flavor-changing couplings of the top

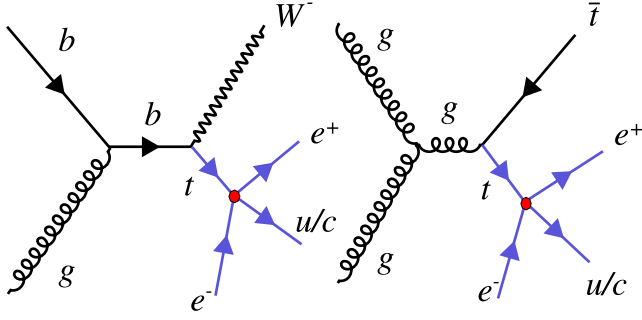


FIG. 2. Representative Feynman diagrams for the $pp \rightarrow tW$ (left) and $pp \rightarrow t\bar{t}$ (right) signal events. The red circle signifies the FCNC interaction of the top quark, which is the focus of this search.

quark through the decays $t \rightarrow u + e^+ + e^-$ and $t \rightarrow c + e^+ + e^-$. The study of FCNC couplings to up and charm quarks is performed independently for each channel. In the $t\bar{t}$ process, one top quark undergoes an FCNC decay into a light up or charm quark accompanied by two opposite-sign (OS) electrons, while the other top quark decays conventionally into a W boson and a b quark. For both the $t\bar{t}$ and tW signals, the W boson decays leptonically, contributing an additional electron in the final state. Consequently, signal events are characterized by the presence of three electrons (two OS electrons from the FCNC decay and one from the W decay), along with at least a light-flavor jet (from the up or charm quark) and exactly one b jet in the case of $t\bar{t}$ production. The Feynman diagrams for these signal processes are shown in Fig. 2, illustrating the expected final states for each channel.

For both processes, the expected number of signal events, denoted as $\mathcal{N}_{S(tW)}$ and $\mathcal{N}_{S(t\bar{t})}$, is given by

$$\mathcal{N}_{S(tW,\bar{t})} = \sigma_{(tW,\bar{t})} \times \mathcal{B}(t \rightarrow u/c + e^+ + e^-) \times \epsilon_{(tW,\bar{t})} \times \mathcal{L}, \quad (3)$$

where $\sigma_{(tW,\bar{t})}$ represents the production cross section of the respective process; $\mathcal{B}(t \rightarrow u/c + e^+ + e^-)$ denotes the branching fraction of the FCNC decay; $\epsilon_{(tW,\bar{t})}$ accounts for the overall selection efficiency specific to the tW and $t\bar{t}$ channels; and $\mathcal{L}$ is the integrated luminosity of the dataset, taken as 3000 fb^{-1} . The efficiency factor $\epsilon_{(tW,\bar{t})}$ incorporates multiple contributions, including the impact of event selection criteria, detector acceptance, and reconstruction performance, which are inherently dependent on the interaction type. Given that the reconstruction efficiency of leptons varies with their flavor, $\epsilon_{(tW,\bar{t})}$ is lepton-flavor dependent. Electrons generally experience lower detection efficiency compared to muons due to their higher probability of undergoing bremsstrahlung and energy loss in the detector material, leading to broader energy deposits and more stringent isolation requirements. Additionally, electron

identification is more sensitive to detector resolution effects, particularly at high pseudorapidity, which further reduces efficiency relative to muons. In contrast, muons, which traverse the entire detector and leave distinct signatures in the muon spectrometer, are less affected by such effects, resulting in a comparatively higher reconstruction efficiency. Consequently, the sensitivity in the electron channel is expected to be weaker compared to the muon channel when analyzed under similar conditions.

The primary background processes contributing to both signal channels include top-antitop ($t\bar{t}$) production, diboson processes such as ZZ and $W^\pm Z$, associated top-quark production with vector bosons ($t\bar{t}W^\pm$, $t\bar{t}Z$, tZq), and four-top-quark ($4t$) production. The dominant SM background for the $t\bar{t}$ signal arises primarily from $t\bar{t}$ production, where electron misidentification can occur due to jet misidentification or photon conversion, leading to an irreducible background contribution. In the case of the tW production channel, the most significant backgrounds originate from diboson production (VV , where $V = W, Z$) and single top quark production in association with a Z boson (tZq), with subsequent leptonic decays of the vector bosons. The analysis is conducted at the LHC with a center-of-mass energy of 13 TeV, assuming an integrated luminosity of 3000 fb^{-1} , corresponding to the dataset anticipated for the High-Luminosity LHC phase.

The effective Lagrangian described in Eq. (1) was implemented into the FeynRules framework [43], generating a Universal Feynman Output (UFO) model [44]. This model was then integrated into MadGraph5_aMC@NLO [45], enabling precise numerical cross-section calculations and event generation. Both signal and background events were simulated with MadGraph5_aMC@NLO, followed by parton showering, hadronization, and underlying event simulations performed with Pythia v8 [46].

All samples were subsequently processed using Delphes 3 [47], employing the high-luminosity CMS detector configuration to simulate realistic detector effects and to reconstruct final-state objects such as electrons, jets, and missing transverse energy. Electron reconstruction relied on isolation criteria within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.3$, with minimum transverse momentum (p_T) set to 20 GeV and pseudorapidity restricted to $|\eta| < 3$. Jets were reconstructed using the anti- k_t clustering algorithm [48] with a radius parameter of $R = 0.4$, implemented through the FastJet package [49,50]. Jet candidates were selected with a minimum transverse momentum requirement of $p_T > 30 \text{ GeV}$ and a pseudorapidity constraint of $|\eta| < 3$. In the analysis of the $t\bar{t}$ and tW processes, the number of jets was required to be at least two and one, respectively. For the $t\bar{t}$ signal process, exactly one b -tagged jet was required.

To reconstruct the top quark mass from the FCNC vertex, three algorithms were developed, as the results illustrated in Fig. 3. In the first method (green histogram), the top quark mass is reconstructed using the two opposite-sign (OS)

electrons with the smallest separation, along with the leading non- b -tagged jet. In the second method (red histogram), all available electrons and the leading non- b -tagged jet are used to minimize the difference $|m_{q\ell\ell} - m_{\text{top}}|$. The third and optimal algorithm (blue histogram) utilizes all electrons and all non- b -tagged jets to minimize $|m_{q\ell\ell} - m_{\text{top}}|$, producing the most accurate reconstruction of the top quark mass. All methods yield maximum values around the expected top quark mass, with the blue algorithm demonstrating superior performance in achieving this objective.

Figure 4 shows the distribution of the reconstructed top quark mass from the FCNC vertex for both $t\bar{t}$ (left) and tW (right) signal processes. The distributions include the main

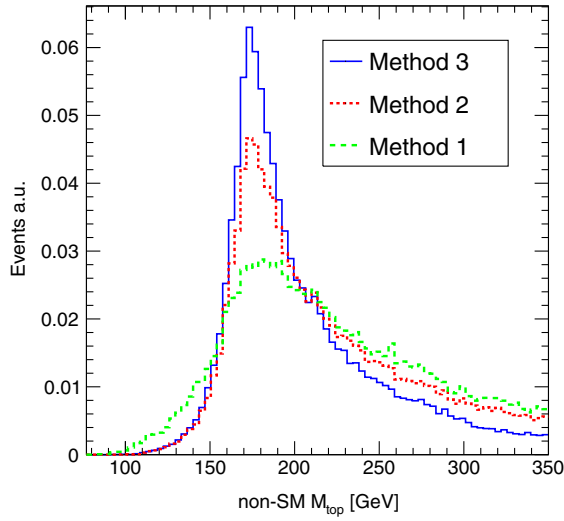


FIG. 3. Top quark mass distribution reconstructed from the FCNC vertex using three algorithms. The green, red, and blue histograms correspond to methods utilizing different combinations of opposite-sign electrons and non- b -tagged jets to optimize the reconstruction accuracy. The blue histogram, representing the most comprehensive algorithm, provides the closest approximation to the true top quark mass.

background contributions and reflect the application of the selection criteria described in the previous sections.

In order to discriminate the signal events from the SM backgrounds, several variables were defined as the input for the machine learning models listed below:

- (1) p_T , η , and ϕ of the leading light flavor jet.
- (2) p_T and η of the b -tagged jet.
- (3) p_T , η , and ϕ of the leading electron.
- (4) Cosine of the angle between the first and second opposite-sign electrons $[\cos(\Delta\phi(e_1, e_2))]$.
- (5) Difference between the first and second opposite-sign electrons in η $[\Delta\eta(e_1, e_2)]$.
- (6) Magnitude of the missing transverse momentum (p_T^{miss}).
- (7) The invariant mass of the non-SM top quark—i.e., m_{eeu} , indicated as $M_{\text{top}}^{\text{non-SM}}$.

Figure 5 illustrates the roles of input feature heat maps for $t \rightarrow ue^-e^+$ in this analysis. The final row (“isSignal”) shows whether each event is classified as signal or background. This label is determined during the event-generation process, where signal and background samples are simulated separately. The most important feature, with the highest correlation coefficients relative to the “isSignal” row, is the nonstandard model top quark mass, with a coefficient of -0.23 . The negative correlation coefficient for the non-Standard Model top quark mass indicates that as the value of the non-Standard Model top quark mass increases, the probability of the data being background becomes higher.

At the end of this section, we discuss the validity of the EFT framework, which requires that contributions from new physics remain within the bounds imposed by tree-level unitarity. In particular, new-physics-induced cross sections must not exceed these limits to ensure the consistency of the EFT expansion. The four-fermion operators considered in this study can arise from different ultraviolet (UV) completions, such as a heavy neutral boson mediating interactions between quarks and leptons, or a scalar or vector particle coupling to quark-lepton pairs. In the case of

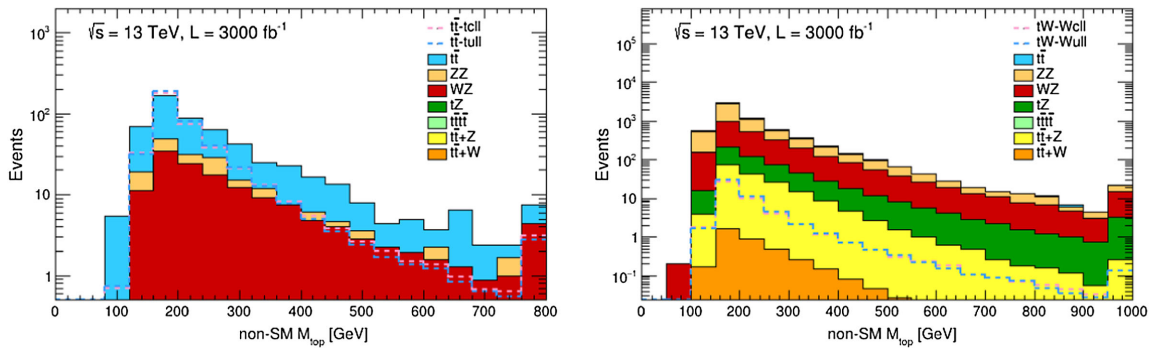


FIG. 4. Distribution of top quark mass from a nonstandard vertex for $t\bar{t}$ (left) and tW (right) signal scenarios with $S_{RR} = V_{RR} = T_{RR} = 1$ and $\Lambda = 1$ TeV, overlaid with the SM backgrounds. The last bin (overflow bin) collects all values exceeding the upper range of the histogram.



FIG. 5. Heat map of input feature correlations for the $t \rightarrow ue^-e^+$ decay channel. Each cell represents the correlation coefficient between two variables, highlighting relationships between features. The last row, labeled “isSignal,” indicates the correlation of each feature with the classification of events as either signal or background.

an intermediate heavy boson, the EFT description remains valid as long as the cutoff scale Λ_e satisfies $m_{e^+e^-}^{\text{max}} < \Lambda_e$, where $m_{e^+e^-}^{\text{max}}$ represents the maximal invariant mass of the dielectron [involved in the $tu(c)ee$ vertex] in the process. This condition is applied to ensure that the characteristic energy of the interaction remains below the new physics scale, thereby preserving the validity of the EFT approximation and maintaining the reliability of the effective expansion.

IV. MACHINE LEARNING MODELS

The traditional cut-based selection method is a relatively crude approach, potentially resulting in the loss of significant signal due to the absolute thresholds imposed. Machine learning techniques are keys in the search for flavor-changing processes involving the top quark. Methods like deep learning and boosted decision trees effectively distinguish between signal events and background processes [51,52], improving the sensitivity of experimental searches. For the implementation of the neural network (NN), we used Keras [53], an open-source library that acts as an interface for the TensorFlow package. The Scikit-learn [54] library was employed for various other classification algorithms, including boosted decision trees, random forests, support vector machines, and more.

A. Classification neural network

A neural network consists of interconnected nodes arranged in layers. Each node processes information and transfers it to the next layer. The links among nodes are determined by a weight, the values of which are adjusted throughout the training process. The neural network modifies its weighted connections based on a loss function (such as mean squared error, which measures the difference between actual and desired outputs) in an effort to minimize the loss value. Each iteration where the algorithm processes the entire dataset is termed an epoch, while the extent to which weights are adjusted during each epoch is denoted as the learning rate. If the learning rate is excessively high, the model may rapidly converge to a suboptimal solution, whereas an overly low learning rate can impede progress, leading to stagnation. We employed a set of significant kinematic properties, such as non-SM top quark mass, jet and electron transverse momentum, missing transverse momentum, etc., as input features. The neural network model was trained using labeled data, where signal instances were assigned a score of 1 and background instances a score of 0. The dataset was partitioned into training and testing sets, with an 80–20% split, where the test set was utilized to evaluate performance.

Neural network hyperparameter optimization is a critical step in designing efficient and effective deep learning models. These hyperparameters, such as learning rate, batch size, number of layers, neurons per layer, optimizers, and activation functions influence the performance of the network. Techniques for hyperparameter optimization aim to find the optimal combination of hyperparameters that minimize the model's loss function and improve its predictive accuracy. In this search, we used a random search method to explore the hyperparameter space and identify the best configuration. To mitigate overfitting in the neural network, two strategies are employed. First, early stopping, which is an effective approach where the training process is halted when the model's performance on a validation set begins to degrade, preventing it from further memorizing noise in the training data. Additionally, dropout can be implemented during training to randomly deactivate a fraction of neurons, encouraging the network to learn more robust features and reducing reliance on specific neurons.

The neural network used in this analysis consists of an input layer followed by three hidden layers, each containing 128 neurons. The scaled exponential linear unit (SELU) activation function is employed in all hidden layers to promote self-normalizing behavior. To mitigate overfitting, dropout regularization is applied after each hidden layer with a dropout rate of 10%. The output layer uses a sigmoid activation function to produce a binary classification score. The network is trained using the Adam optimizer with the learning rate provided by the Keras library (0.001), and the loss function is binary cross-entropy, defined as [55]

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \quad (4)$$

where y_i and $\hat{y}_i$ denote the true and predicted labels, respectively. Prior to training, all input features are standardized using a StandardScaler transformation to zero mean and unit variance.

The effectiveness of a binary neural network, consisting of only two outputs, can be evaluated using the receiver operating characteristic (ROC) curve. This curve illustrates the relationship between the true positive rate (signals which are labeled as signals) and the false positive rate (backgrounds which are labeled as signals) across different threshold settings, as depicted in Fig. 6. A greater area under the ROC curve generally indicates higher performance of the neural network. The neural network equipped with fine-tuned hyperparameters and dropout layers (shown as green in Fig. 6) achieves the highest performance in detecting signal events.

A comparison of the performance metrics for the baseline and optimized DNN models is presented in Table I, highlighting the improvements in both area under the receiver operating characteristic curve (AUC) and accuracy

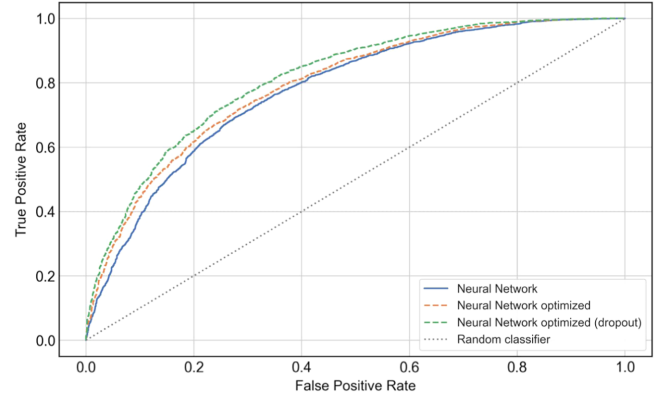


FIG. 6. ROC curve showing the performance of the three neural network classification models, with the x axis representing the false positive rate (FPR) and the y axis representing the true positive rate (TPR).

achieved through dropout regularization and optimizer tuning.

To define signal regions, we used approximate median significance (AMS):

$$\text{AMS} = \sqrt{2 \left((TPR + FPR + b_r) \ln \left(1 + \frac{TPR}{FPR + b_r} \right) - TPR \right)}, \quad (5)$$

where TPR and FPR are the true and false positive rates, and b_r is some number chosen to reduce the variance of the AMS such that the selection region is not too small. The significance of an observed signal is then evaluated based on the value of the AMS. We have plotted AMS for several values of b_r to see how it affects selection for the threshold value and found that $b_r = 0.001$ maximizes the AMS plots. A higher AMS value indicates a greater deviation of the observed signal from the expected background, suggesting a more significant observation; one should then select the value of the threshold that maximizes the AMS. The impact of the threshold on the AMS values for all neural network models for the $t\bar{t}$ signal is shown in Fig. 7. As observed, the optimized neural network with dropout layers exhibits a higher AMS value, particularly around the 70%–80% threshold.

TABLE I. Performance comparison of different neural network architectures in terms of AUC and accuracy (mean $\pm$ standard deviation) across multiple runs.

Model	AUC	Accuracy
Baseline DNN	0.7805 ± 0.0069	0.7310 ± 0.0040
DNN with dropout	0.8102 ± 0.0064	0.7537 ± 0.0019
DNN with optimizer tuning	0.7958 ± 0.0025	0.7427 ± 0.0009

B. Other classification models

In addition to the neural network model, various other classification machine learning algorithms have been explored in this study. Among these are the random forest [56], decision tree [57], support vector machine (SVM) [58], k -nearest neighbors (KNN) [59], and simple logistic regression. Each of these models offers unique advantages and characteristics that make it suitable for different types of classification tasks. Random forest and decision tree models, for example, are known for their ability to handle nonlinear relationships and interactions between features, making them effective in capturing complex patterns in the data. Support vector machines excel at separating classes with a clear margin of separation, while k -nearest neighbors

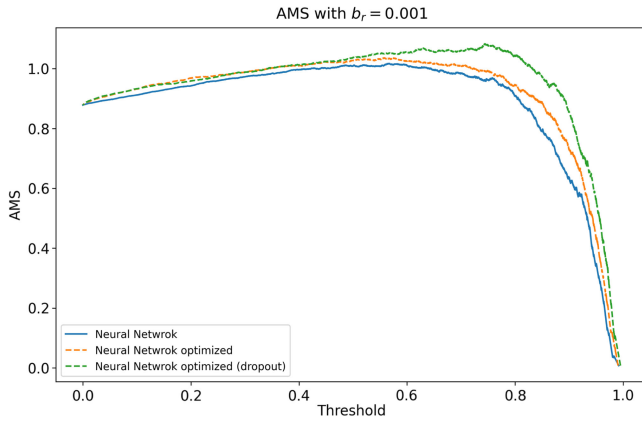


FIG. 7. AMS curve for a simple neural network (blue), optimized neural network (orange), and optimized neural network with dropout layers (green) for the $t\bar{t}$ signal. The threshold with the highest AMS value is the best option for defining the signal region.

models rely on similarity metrics to classify data points. By exploring a range of classification algorithms, we aim to identify the most suitable model for accurately detecting and characterizing top quark flavor-changing events in simulated data. Figure 8 illustrates the accuracy of various classification machine learning models across a spectrum of known classifiers. Compared to the 62% and 63% accuracy achieved by the gradient boosting and random forest models, the optimized neural network with dropout layers achieves a higher accuracy of 76%.

V. RESULTS

In this sensitivity analysis, our objective is to evaluate the sensitivity of the experimental results to the scale of new physics, represented by the parameter Λ . We begin by calculating the p -value for each hypothesis, both in the presence and absence of signal events, using the BinomialExpP function within the RooFit framework [60]. This approach allows us to quantify the probability of observing the expected event distribution under each hypothesis, based on simulated data. To further interpret the results, we perform a CL test, a statistical method widely adopted in particle physics analyses, to establish the 95% confidence level (CL) exclusion thresholds for the parameter Λ . These exclusion thresholds define the range of Λ values excluded at a 95% CL, providing constraints on the scale of new physics under investigation.

In Tables II and III, we present the expected number of signal and total background events per 3000 fb^{-1} of integrated luminosity for two distinct signals, $t\bar{t}$ and tW , where the top quark decays to a light-flavor jet (u or c) and an electron-positron pair. Results are listed for the optimized neural network model with dropout layers (the most accurate model), along with two threshold settings that vary

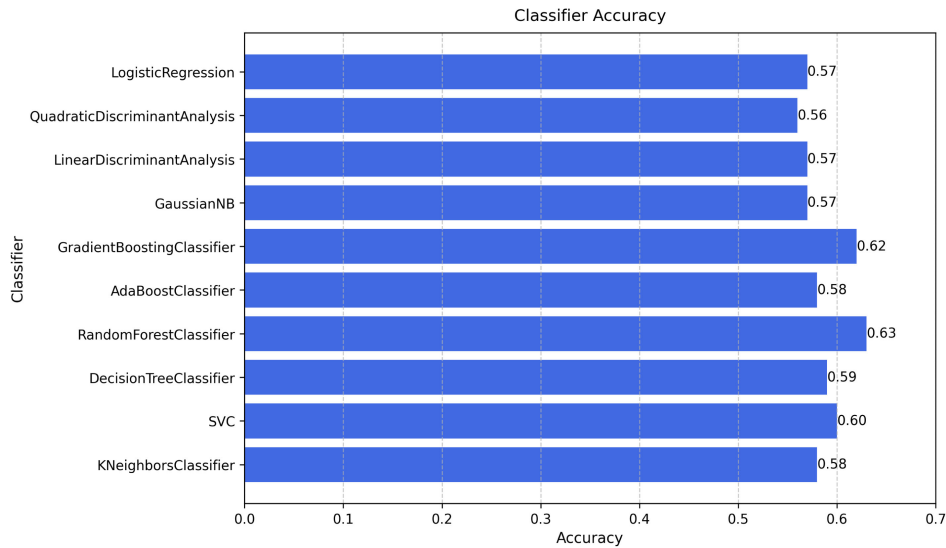


FIG. 8. Accuracy for classified models such as random forest, support vector machine, and gradient boosting, etc. The hyperparameters for random forest and decision tree classifiers are optimized to achieve the maximum accuracy.

TABLE II. Number of signal ($\mathcal{N}_S$) and background ($\mathcal{N}_B$) events, and signal-to-background ratio per 3000 fb⁻¹ of integrated luminosity, for the process $pp \rightarrow t\bar{t} \rightarrow ue^-e^+\bar{t}$ (top) and $pp \rightarrow t\bar{t} \rightarrow ce^-e^+\bar{t}$ (bottom).

Process	NN threshold $\geq 75\%$			NN threshold $\geq 80\%$		
	$\mathcal{N}_S$	$\mathcal{N}_B$	$\mathcal{N}_S/\mathcal{N}_B$	$\mathcal{N}_S$	$\mathcal{N}_B$	$\mathcal{N}_S/\mathcal{N}_B$
$pp \rightarrow t\bar{t} \rightarrow ue^-e^+\bar{t}$	538.1	345.6	1.6	421.0	245.1	2.2
$pp \rightarrow t\bar{t} \rightarrow ce^-e^+\bar{t}$	570.8	240.7	2.4	492.8	173.9	3.3

 TABLE III. Number of signal ($\mathcal{N}_S$) and background ($\mathcal{N}_B$) events, and signal-to-background ratio per 3000 fb⁻¹ of integrated luminosity, for the process $pp \rightarrow tW \rightarrow ue^-e^+W$ (top) and $pp \rightarrow tW \rightarrow ce^-e^+W$ (bottom).

Process	NN threshold $\geq 65\%$			NN threshold $\geq 70\%$		
	$\mathcal{N}_S$	$\mathcal{N}_B$	$\mathcal{N}_S/\mathcal{N}_B$	$\mathcal{N}_S$	$\mathcal{N}_B$	$\mathcal{N}_S/\mathcal{N}_B$
$pp \rightarrow tW \rightarrow ue^-e^+W$	47.3	4087.4	0.013	38.7	2955.1	0.019
$pp \rightarrow tW \rightarrow ce^-e^+W$	38.1	3353.7	0.014	30.1	2261.9	0.020

depending on the signal scenario. As observed, increasing the cut on the neural network threshold leads to a higher signal-to-background ratio ($\mathcal{N}_S$), enhancing the separation between signal and background. This improvement allows for greater sensitivity in distinguishing potential signal events from background noise, contributing to more robust constraints on the parameter Λ .

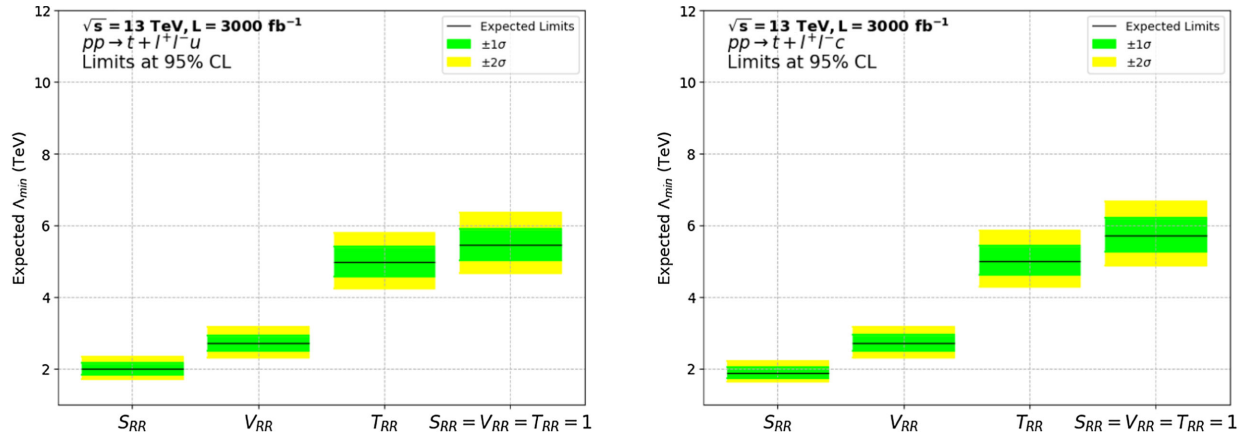
The expected 95% CL exclusion limits on the new physics scale Λ for the $t\bar{t}$ process are shown in Fig. 9, for three coupling scenarios: scalar-like ($S_{RR} = 1$), vector-like

($V_{RR} = 1$), tensor-like ($T_{RR} = 1$), and a mixed scenario ($S_{RR} = V_{RR} = T_{RR} = 1$). These results are obtained using the optimized neural network model with dropout layers, and they are presented alongside $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands. Among these scenarios, the scalar-like configuration demonstrates a lower exclusion limit, due to its reduced production cross section. The sensitivity of the $t\bar{t}$ channel to tensor (vector) coupling extends to a lower limit of approximately $\Lambda_e > 5(2.8)$ TeV for both $tuee$ and $tcee$ at an integrated luminosity of 3000 fb⁻¹. The sensitivity achieved in the analysis of a single top plus dilepton to tensor (vector) coupling was found to be $\Lambda_\mu > 7.1(4.7)$ TeV and $\Lambda_\mu > 2.4(1.5)$ TeV for the corresponding $t\mu\mu$ and $t\bar{c}\mu\mu$ using an integrated luminosity of 3000 fb⁻¹ [30].

As shown in Fig. 10, for both $t\bar{t}$ channels, increasing the NN threshold from 0.70 to 0.80 yields a modest but consistent improvement in the expected limits on the NP scale. While the central values increase slightly across the thresholds, the associated $\pm 1\sigma$ and $\pm 2\sigma$ bands remain stable, indicating improved signal-to-background separation without a significant increase in statistical uncertainty. This trend aligns with expectations: higher NN thresholds enhance purity, reducing background contamination and improving sensitivity.

Figure 11 shows the corresponding 95% CL exclusion limits on Λ_e for the tW process across three distinct neural network threshold settings, all utilizing the optimized network with dropout layers. The results indicate consistent performance across thresholds, with lower exclusion limits reaching up to 2 TeV. These findings underscore the robustness of the analysis across both processes, demonstrating competitive sensitivity in probing the scale of new physics in FCNC interactions.

To benchmark the performance of our machine learning strategy, we performed a traditional cut-and-count analysis using high-level observables and a manually defined signal


 FIG. 9. Expected 95% CL lower limit on Λ_e of the $tuee$ (left) and $tcee$ (right) for four signal scenarios: $S_{RR} = 1$, $V_{RR} = 1$, $T_{RR} = 1$, and $S_{RR} = V_{RR} = T_{RR} = 1$, with a total integrated luminosity of 3000 fb⁻¹.

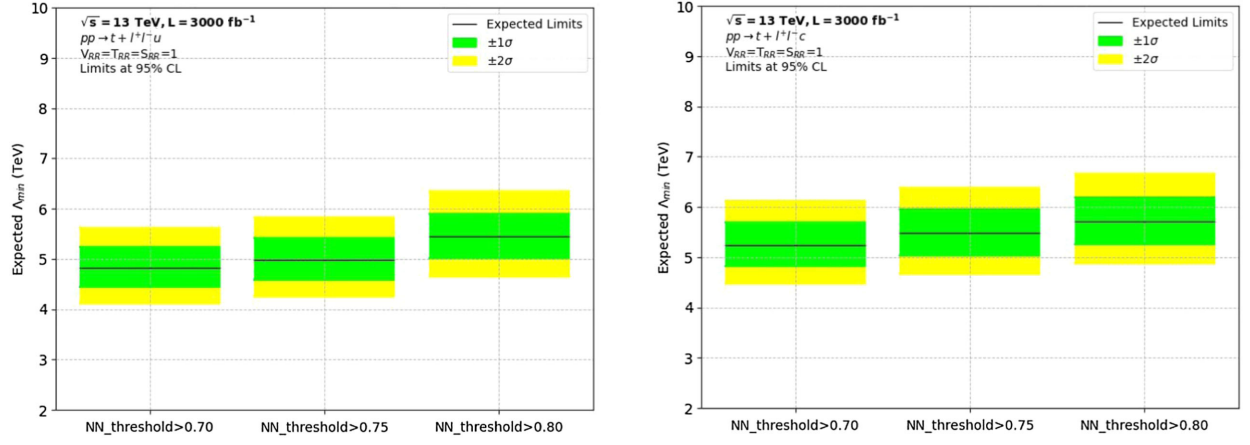


FIG. 10. Expected 95% CL lower limit on Λ_e of the *tuee* (left) and *tcee* (right) of the $t\bar{t}$ process for three neural network thresholds, with a total integrated luminosity of 3000 fb^{-1} .

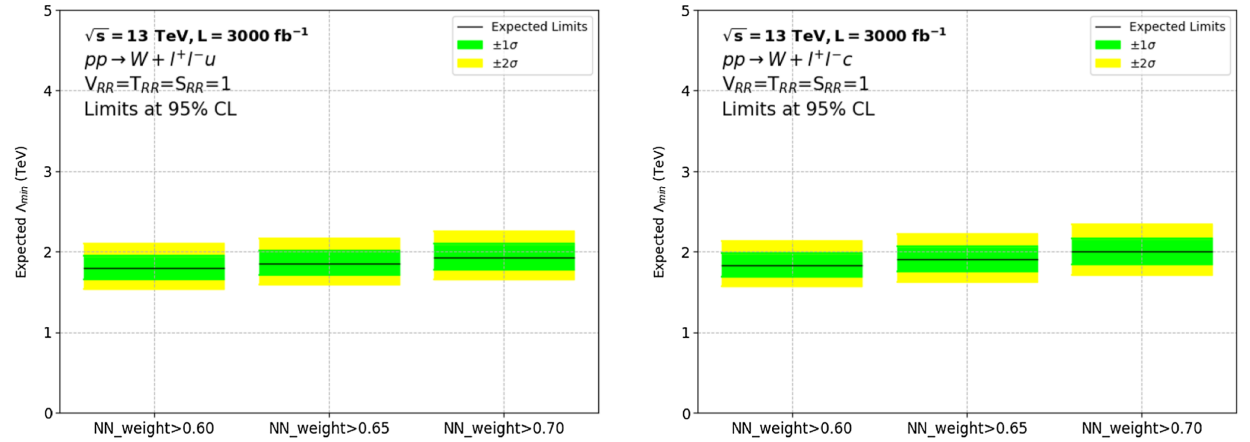


FIG. 11. Expected 95% CL lower limit on Λ_e of the *tuee* (left) and *tcee* (right) of the tW process for three neural network thresholds, with a total integrated luminosity of 3000 fb^{-1} .

region ($150 \text{ GeV} < M_{\text{top}}^{\text{non-SM}} < 200 \text{ GeV}$ and $N_{\text{jet}} \geq 3$). The resulting signal yields were 98.27 for the *tcee* operator and 105.30 for *tuee*, with 98.53 background events (Table IV). This yields an NP scale sensitivity of approximately 3.6 TeV at 95% CL for both channels. While this demonstrates that even simple selection criteria can yield meaningful results, the deep neural network approach improves background suppression and signal discrimination, ultimately leading to stronger limits on the NP scale, surpassing the cut-and-count baseline.

TABLE IV. Cut-and-count results showing the number of signal ($\mathcal{N}_S$) and background ($\mathcal{N}_B$) events, and the signal-to-background ratio for up and charm signals.

Process	$\mathcal{N}_S$	$\mathcal{N}_B$	$\mathcal{N}_S/\mathcal{N}_B$
$pp \rightarrow t\bar{t} \rightarrow ue^-e^+\bar{t}$	105.3	98.5	1.07
$pp \rightarrow t\bar{t} \rightarrow ce^-e^+\bar{t}$	98.3	98.5	1.00

VI. INTERPRETATION OF RESULTS IN THE CONTEXT OF A Z' MODEL

In this section, as an example, we interpret the results of our EFT analysis in the context of a specific Z' model. The Z' boson is a new gauge boson associated with an additional $U(1)'$ gauge symmetry beyond the SM. The Lagrangian describing the interactions of the Z' boson with the SM fermions is given by

$$\mathcal{L} \supset g_{tu} Z'_\mu \bar{t} \gamma^\mu P_L u + g_\ell Z'_\mu \bar{\ell} \gamma^\mu P_L \ell + \text{H.c.}, \quad (6)$$

where Z'_μ is the Z' gauge field, and g_{tu} and g_ℓ are the couplings of the Z' boson to the top quark (t), up/charm quark (u), and leptons (ℓ), respectively. This model introduces FCNCs at tree level, and the Z' boson can mediate processes such as $t \rightarrow u Z'$ and $Z' \rightarrow \ell^+ \ell^-$. By matching the EFT to this Z' model, we can translate the bounds on the EFT Wilson coefficients into constraints on

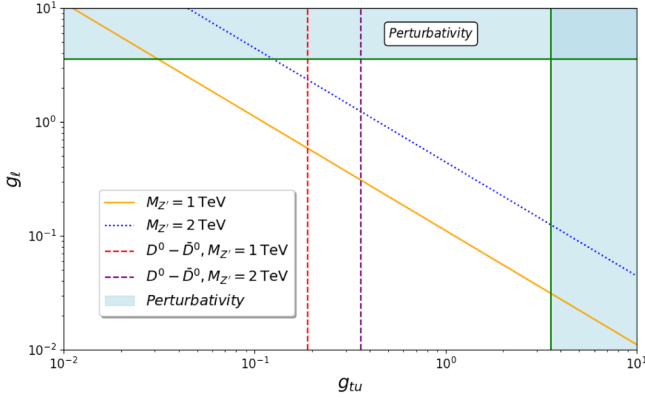


FIG. 12. Parameter space constraints derived from the interpretation of rare top decays at the HL-LHC for Z' boson masses of $m_{Z'} = 1$ and 2 TeV. The light blue shaded region represents the exclusion derived from perturbativity conditions. The dashed lines are the limits obtained from $D^0 - \bar{D}^0$ mixing, assuming $g_{tc} = 0.01$ for $m_{Z'} = 1$ and 2 TeV.

the Z' couplings and mass. This interpretation allows us to explore the implications of our EFT results for new physics scenarios involving a Z' boson with flavor-violating interactions. In particular, the matching condition is given by $\frac{V}{\Lambda^2} \propto g_{tu}g_\ell/M_{Z'}^2$.

This matching condition arises from integrating out the heavy Z' boson and expressing its effects in terms of higher-dimensional operators in the EFT. The Wilson coefficient V encapsulates the effects of the Z' boson at energies below its mass scale, allowing us to connect the EFT framework to the specific UV completion provided by the Z' model. The matching condition plays a crucial role in translating the bounds on the Wilson coefficient, obtained from experimental analysis and EFT analyses, into constraints on the fundamental parameters of the Z' model: the couplings g_{tu} and g_ℓ , and the Z' mass $m_{Z'}$. Figure 12 shows the parameter space region that can be bounded by the interpretation of the present analysis of rare top decays in the HL-LHC with $m_{Z'} = 1, 2$ TeV.

In addition to experimental constraints, the Z' model must satisfy theoretical requirements, such as perturbativity, to ensure its consistency as a quantum field theory. The decay $t \rightarrow u + e^+ + e^-$ is mediated by the Z' boson, and its amplitude depends on the product of the couplings g_{tu} and g_ℓ . To ensure perturbativity, the following conditions must hold: $|g_{tu}| \leq \sqrt{4\pi}$ and $|g_\ell| \leq \sqrt{4\pi}$. These conditions ensure that the perturbative expansion of the decay amplitude remains valid, preventing higher-order terms from dominating the leading-order results and preserving the consistency of the theory. The constraints derived from these conditions are illustrated in Fig. 12, where the excluded region is highlighted in light blue shading.

We also present limits on the flavor-changing part of the Z' couplings using $D^0 - \bar{D}^0$ mixing, considering its effects on the up-type quarks, and constraining the couplings and

mass of Z' . Within the SM, $D^0 - \bar{D}^0$ mixing is a manifestation of FCNCs that occurs because the flavor eigenstates differ from the physical mass eigenstates of the $D^0 - \bar{D}^0$ system. Both short-range quark-level transitions and long-range processes contribute to $D^0 - \bar{D}^0$ oscillations. The short-range contributions proceed through loops where virtual particles are mediated [61–63]. This makes the study of $D^0 - \bar{D}^0$ mixing particularly interesting, as new physics models with additional FCNC degrees of freedom can be examined through it. The parameters of the Z' FCNC interactions between two up-type quarks can be significantly constrained using the measurement of $D^0 - \bar{D}^0$ mixing, which is affected at both the tree level and loop level. Assuming the mediator has up-charm (g_{uc}), up-top (g_{tu}), and charm-top (g_{tc}) couplings, it can have a significant contribution to neutral D -meson mixing. If the mediator is Z' , its contribution to the mass difference between the two mass eigenstates is given by [64]

$$\Delta M_D = \frac{f_D^2 M_D^2 B_D}{48 m_{Z'}^2} \left[g_{uc}^2 + g_{tu}^2 g_{tc}^2 \frac{x}{8\pi^2} (32 f_{Z'}(x) - 5 g_{Z'}(x)) \right],$$

$$\text{where } f_{Z'}(x) = \frac{1}{2} \frac{1}{(1-x)^3} [1 - x^2 + 2x \log x],$$

$$g_{Z'}(x) = \frac{2}{(1-x)^3} [2(1-x) + (1+x) \log x]. \quad (7)$$

The first term in the square bracket represents the tree-level contribution, while the second term is due to the box contribution. Here, $x \equiv m_{Z'}^2/m_t^2$, $M_D \sim 1.9$ GeV is the mass of the D meson, $f_D \sim 0.223$ GeV is its decay constant, and $B_D \sim 1$ is the bag model parameter. In this work, for simplicity, we set $g_{uc} = 0$ and only concentrate on the FCNCs of top and up/charm quarks via Z' . The current bound on D -meson mixing is $\Delta M_D \lesssim 10^{-15}$ [65,66], which allows us to derive constraints on g_{tu} , as presented in Fig. 12 for $g_{tc} = 0.01$ and $m_{Z'} = 1, 2$ TeV. The bounds from D -meson mixing on g_{tu} and g_{tc} are similar.

VII. SUMMARY

In this study, we have investigated the effects of FCNC couplings in the top quark sector through the processes $pp \rightarrow t\bar{t}$ and $pp \rightarrow tW$, effectively probed using a three-electron selection and a b -jet tagging strategy. SM backgrounds were carefully analyzed, demonstrating that an optimal selection criterion based on the reconstructed top quark mass from the FCNC vertex provides excellent separation between NP signals and SM backgrounds. Several machine learning models were trained and optimized to enhance signal-background discrimination, with the neural network incorporating dropout layers yielding

the highest performance in classification metrics. Furthermore, we calculated exclusion limits on the scale of the $tuee$ and $tcee$ FCNC interactions, with expected 95% CL bounds of $\Lambda_e \leq 5.5$ TeV for a $tuee$ interaction and $\Lambda_e \leq 5.7$ TeV for a $tcee$ interaction, assuming $S_{RR} = V_{RR} = T_{RR} = 1$. For the only tensor interaction, the expected 95% CL bounds of $\Lambda_e \lesssim 5$ TeV for both $tuee$ and $tcee$ interactions have been obtained.

These results can be compared to the recent study [30] that explores FCNC interactions in $t\mu^+\mu^-$ and $tc\mu^+\mu^-$ processes, based on 3000 fb^{-1} of HL-LHC data, which reports 95% CL bounds of $\Lambda_\mu \leq 7.1$ TeV and $\Lambda_\mu \leq 2.4$ TeV for the tensor coupling via these trilepton signals. Despite the lepton flavor difference, our analysis achieves comparable sensitivity, with the expected limit for the tce^+e^- channel reaching a value of around 2.5 TeV higher.

This study enhances FCNC sensitivity at the LHC by applying machine learning techniques to three-electron final states. The exclusion limits derived here are

competitive with previous analyses, highlighting the power of advanced classification models in enhancing new physics searches. In addition, we have interpreted our EFT results obtained from the analysis in the context of a specific Z' model, demonstrating how the derived constraints map onto the parameter space of a heavy neutral gauge boson that mediates flavor-changing interactions.

ACKNOWLEDGMENTS

We are grateful to Gh. Haghighat for his insightful comments on the manuscript. We would like to express our sincere gratitude to IPM for providing the servers and computational resources necessary for this research.

DATA AVAILABILITY

The data that support the findings of this article are openly available [67].

-
- [1] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Phys. Rev. D* **2**, 1285 (1970).
 - [2] G. Eilam, J. L. Hewett, and A. Soni, *Phys. Rev. D* **44**, 1473 (1991); **59**, 039901(E) (1999).
 - [3] S. Bejar, J. Guasch, and J. Sola, [arXiv:hep-ph/0101294](#).
 - [4] J. J. Liu, C. S. Li, L. L. Yang, and L. G. Jin, *Phys. Lett. B* **599**, 92 (2004).
 - [5] J. Cao, Z. Heng, L. Wu, and J. M. Yang, *Phys. Rev. D* **79**, 054003 (2009).
 - [6] S. Bejar, J. Guasch, D. Lopez-Val, and J. Sola, *Phys. Lett. B* **668**, 364 (2008).
 - [7] J. L. Diaz-Cruz, M. A. Perez, G. Tavares-Velasco, and J. J. Toscano, *Phys. Rev. D* **60**, 115014 (1999).
 - [8] J. A. Aguilar-Saavedra, *Acta Phys. Pol. B* **35**, 2695 (2004).
 - [9] G. Durieux, F. Maltoni, and C. Zhang, *Phys. Rev. D* **91**, 074017 (2015).
 - [10] S. Khatibi and M. Mohammadi Najafabadi, *Phys. Rev. D* **89**, 054011 (2014).
 - [11] J. A. Aguilar-Saavedra, *Nucl. Phys.* **B812**, 181 (2009).
 - [12] H. Khanpour, S. Khatibi, M. Khatiri Yanehsari, and M. Mohammadi Najafabadi, *Phys. Lett. B* **775**, 25 (2017).
 - [13] S. Khatibi and H. Khanpour, *Nucl. Phys.* **B967**, 115432 (2021).
 - [14] A. Hayrapetyan *et al.* (CMS Collaboration), *Phys. Rev. D* **109**, 072004 (2024).
 - [15] K. Y. Oyulmaz, A. Senol, H. Denizli, A. Yilmaz, I. Turk Cakir, and O. Cakir, *Eur. Phys. J. C* **79**, 83 (2019).
 - [16] J. Ebadi, F. Elahi, M. Khatiri, and M. Mohammadi Najafabadi, *Phys. Rev. D* **98**, 075012 (2018).
 - [17] E. Alici and M. Köksal, *Mod. Phys. Lett. A* **34**, 1950298 (2019).
 - [18] G. Aad *et al.* (ATLAS Collaboration), *Phys. Lett. B* **842**, 137379 (2023); **847**, 138286(E) (2024).
 - [19] Y. Afik, S. Bar-Shalom, K. Pal, A. Soni, and J. Wudka, *Nucl. Phys.* **B980**, 115849 (2022).
 - [20] P. J. Fox, Z. Ligeti, M. Papucci, G. Perez, and M. D. Schwartz, *Phys. Rev. D* **78**, 054008 (2008).
 - [21] A. Falkowski, M. González-Alonso, and K. Mimouni, *J. High Energy Phys.* **08** (2017) 123.
 - [22] S. Bhattacharya, S. Jahedi, S. Nandi, and A. Sarkar, *J. High Energy Phys.* **07** (2024) 061.
 - [23] S. Sun, Q. S. Yan, X. Zhao, and Z. Zhao, *Phys. Rev. D* **108**, 075016 (2023).
 - [24] W. Altmannshofer, S. Gori, B. V. Lehmann, and J. Zuo, *Phys. Rev. D* **107**, 095025 (2023).
 - [25] L. Shi and C. Zhang, *Chin. Phys. C* **43**, 113104 (2019).
 - [26] W. Liu and H. Sun, *Phys. Rev. D* **100**, 015011 (2019).
 - [27] B. Grzadkowski, Z. Hioki, and M. Szafranski, *Phys. Rev. D* **58**, 035002 (1998).
 - [28] B. Grzadkowski, *Acta Phys. Pol. B* **27**, 921 (1996).
 - [29] S. Bar-Shalom and J. Wudka, *Phys. Rev. D* **60**, 094016 (1999).
 - [30] Y. Afik, S. Bar-Shalom, A. Soni, and J. Wudka, *Phys. Rev. D* **103**, 075031 (2021).
 - [31] D. Guadagnoli and P. Koppenburg, *Annu. Rev. Nucl. Part. Sci.* **73**, 1 (2023).
 - [32] F. U. Bernlochner, M. F. Sevilla, D. J. Robinson, and G. Wormser, *Rev. Mod. Phys.* **94**, 015003 (2022).
 - [33] D. Bryman, V. Cirigliano, A. Crivellin, and G. Inguglia, *Annu. Rev. Nucl. Part. Sci.* **72**, 69 (2022).
 - [34] J. Abdallah *et al.* (DELPHI Collaboration), *Eur. Phys. J. C* **71**, 1555 (2011).
 - [35] J. Aguilar Saavedra and G. C. Branco, *Phys. Lett. B* **945**, 347 (2000).
 - [36] P. Achard *et al.* (L3 Collaboration), *Phys. Lett. B* **549**, 290 (2002).

- [37] S. Davidson, M. L. Mangano, S. Perries, and V. Sordini, *Eur. Phys. J. C* **75**, 450 (2015).
- [38] M. Chala, J. Santiago, and M. Spannowsky, *J. High Energy Phys.* **04** (2019) 014.
- [39] J. A. Aguilar-Saavedra, *Nucl. Phys.* **B843**, 638 (2011); **B851**, 443(E) (2011).
- [40] N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang, *J. High Energy Phys.* **04** (2019) 100.
- [41] J. Benesch *et al.* (MOLLER Collaboration), arXiv:1411.4088.
- [42] A. Hayrapetyan *et al.* (CMS Collaboration), *Phys. Rev. D* **111**, 012009 (2025).
- [43] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *Comput. Phys. Commun.* **185**, 2250 (2014).
- [44] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, *Comput. Phys. Commun.* **183**, 1201 (2012).
- [45] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *J. High Energy Phys.* **06** (2011) 128.
- [46] T. Sjostrand, S. Mrenna, and P. Z. Skands, *J. High Energy Phys.* **05** (2006) 026.
- [47] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (Delphes 3 Collaboration), *J. High Energy Phys.* **02** (2014) 057.
- [48] M. Cacciari, G. P. Salam, and G. Soyez, *J. High Energy Phys.* **04** (2008) 063.
- [49] M. Cacciari, G. P. Salam, and G. Soyez, *Eur. Phys. J. C* **72**, 1896 (2012).
- [50] M. Cacciari and G. P. Salam, *Phys. Lett. B* **641**, 57 (2006).
- [51] G. Aad *et al.*, *Eur. Phys. J. C* **83**, 561 (2023).
- [52] G. Aad *et al.*, *Phys. Lett. B* **856**, 138938 (2024).
- [53] F. Chollet *et al.*, Keras, <https://keras.io> (2015).
- [54] F. Pedregosa *et al.*, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [55] Ian Goodfellow and Yoshua Bengio, *Deep Learning* (MIT Press, Cambridge, England, 2016).
- [56] L. Breiman, *J. Mach. Learn. Res.* **45**, 5 (2001).
- [57] A. M. Ahmed, A. Rizaner, A. Hakan Ulusoy, and Y. Deng, *PLoS One* **13**, e0194168 (2018).
- [58] S. Xie and R. T. Ogden, *Biostatistics* **25**, kxae007 (2024).
- [59] A. Basheer *et al.*, arXiv:2003.09187.
- [60] W. Verkerke and D. Kirkby, *eConf C* **0303241**, MOLT007 (2003).
- [61] S. Bianco, F. L. Fabbri, D. Benson, and I. Bigi, *Riv. Nuovo Cimento* **26N7**, 1 (2003).
- [62] G. Burdman and I. Shipsey, *Annu. Rev. Nucl. Part. Sci.* **53**, 431 (2003).
- [63] M. Artuso, B. Meadows, and A. A. Petrov, *Annu. Rev. Nucl. Part. Sci.* **58**, 249 (2008).
- [64] A. Arhrib, K. Cheung, C. W. Chiang, and T. C. Yuan, *Phys. Rev. D* **73**, 075015 (2006).
- [65] Y. S. Amhis *et al.* (HFLAV Collaboration), *Phys. Rev. D* **107**, 052008 (2023).
- [66] G. Yeghiyan, arXiv:1209.1854.
- [67] M. G. Bostanabad and M. M. Najafabadi, Zenodo (2024), 10.5281/zenodo.15012286.