

Exploring top-quark signatures of heavy flavor-violating scalars at the LHC

Alexandre Alves^{1,2,*}, Eduardo da Silva Almeida^{3,†}, Alex G. Dias², and Diego S. V. Gonçalves^{2,‡}

¹*Departamento de Física, Universidade Federal de São Paulo, UNIFESP, 09972-270, Diadema-SP, Brazil*

²*Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, UFABC, 09210-580, Santo André-SP, Brazil*

³*Departamento de Física do Estado Sólido, Universidade Federal da Bahia, UFBA, 40170-115, Salvador-BA, Brazil*



(Received 6 August 2025; revised 25 September 2025; accepted 29 September 2025; published 21 October 2025)

In this work, we study flavor-violating scalars (flavons) in a range of large masses that have not been explored previously. We model the interactions with an effective field theory formulation where the flavon is heavier than the top quark. In addition, we assume that the flavon only couples to fermions of the standard model in a flavor-changing way, making it easier to escape experimental constraints. As the flavon couples strongly to top quarks, same-sign and opposite-sign top-quark pair signals can be explored in the search for those particles. We investigate the possibility of probing flavons through their interactions with a top quark plus up and charm quarks for effective couplings of order 10^{-2} TeV and masses in the ~ 200 – 1400 GeV range in the 14 TeV LHC in scenarios with up to 3000/fb.

DOI: [10.1103/8kcm-wfpt](https://doi.org/10.1103/8kcm-wfpt)

I. INTRODUCTION

Hypothetical scalar particles with flavor-violating interactions are part of the spectra of several models proposed to address problems unresolved within the standard model (SM). Among the models predicting such particles, there is an effective field theory (EFT) known as the Froggatt-Nielsen mechanism [1], motivated by its explanation of the flavor structure of the Yukawa couplings required to generate the fermion mass hierarchy and the Cabibbo–Kobayashi–Maskawa (CKM) mixing matrix of the SM. In this mechanism, the Yukawa couplings $Y_{ij} = y_{ij}(\langle\phi\rangle/M)^{n_f}$ are effectively generated through nonrenormalizable operators that are invariant under a global $U(1)_F$ symmetry. This symmetry is broken by the vacuum expectation value (vev) $\langle\phi\rangle$ of a singlet field ϕ , and the operators are suppressed by a high-energy scale $M > \langle\phi\rangle$, i.e., $y_{ij}(\frac{\phi}{M})^{n_f}\bar{F}_{iL}\Phi f'_{jR}$. The couplings y_{ij} are of order 1, and $\bar{F}_{iL}\Phi f'_{jR}$ are the SM Yukawa operators carrying charge $-n_f < 0$ of $U(1)_F$ involving the chiral fermion multiplets, F_{iL} and f'_{jR} , along with the Higgs doublet $\Phi = H$ (or

$i\sigma_2 H^*$ depending on the fermion f'_{jR}). The particle excitation of the pseudo-Nambu-Goldstone boson field associated with the spontaneous breaking of the $U(1)_F$ symmetry is the flavon, which exhibits flavor-violating interactions. Axions and axionlike particles [2,3] can also have flavor-violating interactions when their associated symmetry behaves similarly to a flavor symmetry [4–8]. As is typical of these particles, their flavor-violating interactions with matter are proportional to the mass of the fermions involved but suppressed by the energy scale of new physics. Especially for axionlike particles and flavons,¹ if the suppression scale is too high, their production becomes negligible, rendering them effectively unobservable [10].

In this work, we investigate signals of heavy flavons at the 14 TeV LHC under the hypothesis that their main interactions violate flavor (the maximally violating flavons), with perturbative couplings proportional to the quark masses. We explore the effective top-flavon couplings, which are not yet severely constrained, unlike the flavor-violating couplings involving the down-type quarks. We also assume that flavor-conserving couplings of the flavons to the fermions are negligibly small compared to the considered flavor-violating ones. Because the gauge bosons need to couple to same-flavor fermions, the flavon–gauge boson couplings cannot be generated at the loop level.

¹In typical axion models, the suppression scale is too high [9]. Nonetheless, axion flavor interactions can still be constrained through meson decays [5].

*Contact author: aalves@unifesp.br

†Contact author: almeidae@ufba.br

‡Contact author: diego.vieira@ufabc.edu.br

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

Our analysis focuses on flavon masses higher than 200 GeV, which decay predominantly into a top quark plus an up or charm quark. We identify a peculiar phenomenology of these new particles where final states contain both same-sign and opposite-sign top-quark pairs, which impacts the signal-to-background separation.

Two experimental studies have searched for the types of signals we propose in this work. A CMS study [11] searches for same-sign top signals in the channel $qg \rightarrow t + a(\bar{q}t)$, where a can be either a scalar or pseudoscalar from a two-Higgs doublet model (2HDM). The 2HDM was studied in an alignment limit where all additional Higgs quartic couplings are of order 1 and no Z_2 symmetry is imposed [12]. In this case, flavor-changing neutral Higgs couplings involving the bosons H and A are allowed. Such a model may provide a viable explanation for electroweak baryogenesis. In addition, they assume a specific framework where the scalar and pseudoscalar have nearly degenerate masses to preserve some interference effects while maintaining reasonably high cross sections. Same-sign top signals from flavor-violating scalars in 2HDMs have also been searched for by ATLAS [13]. In the case of this ATLAS study, they consider final states with up to four top quarks, which can only be significantly produced when diagonal couplings are present. However, the scenario where diagonal top couplings are turned off is investigated, and in that case, limits on production cross section times' branching ratios can impact the EFT formulation that we consider. We thus take the exclusion limits derived in Ref. [13] into account. ATLAS Collaboration is also searching for flavor-changing neutral current (FCNC) couplings between the top quark and the Higgs boson in multilepton final states [14].

Contrary to the same-sign top-quark signal, which leads to two hard same-sign leptons and has low backgrounds, the opposite-sign option is plagued by the $t\bar{t}$ events and other backgrounds. In this case, separating the signal events is challenging. We thus employed a parametrized neural network (pNN) [15–17] for the signal background classification to enhance the sensitivity to our EFT parameters. The pNN option is adequate when we want the neural network to learn to separate signals from backgrounds as a continuous function of the flavon mass parameter.

Our main result is that flavons in the mass interval of ~ 200 – 1400 GeV can be discovered or excluded at the 14 TeV High-Luminosity LHC (HL-LHC) for effective couplings up to order 10^{-2} TeV $^{-1}$ after 3000/fb, considerably extending the search prospects for these particles, with predominance of the same-sign signal channel. The findings of the present analysis represent a complementary study to previous works [18,19] on flavons lighter than the top quark.

In Sec. II, we present the EFT formulation of the interactions. In Sec. III, the signal and background simulations are discussed. A cut-based analysis strategy with

results is addressed in Sec. IV, while the multivariate analysis using pNN is presented in Sec. V. The conclusions can be found in Sec. VI.

II. EFFECTIVE LAGRANGIAN

We assume the flavon field originates from the spontaneous symmetry breaking of a global $U(1)_F$ flavor symmetry, as in the Froggatt-Nielsen scheme [1]. The main interactions of the flavon with fermions arise from effective operators of the form $y_{ij}(\frac{\phi}{M})^{n_f} \bar{F}_{iL} \Phi f'_{jR}$, where $\phi(x) = (\rho(x) + \sqrt{2}\langle\phi\rangle)e^{ia(x)/\Lambda}/\sqrt{2}$ is a scalar SM singlet containing the flavon field a , and n_f are integers fixed by the flavor symmetry. Here, F_{iL} and f'_{jR} denote the SM fermion multiplets, and $\Phi = H$ (or $i\sigma_2 H^*$ depending on f'_{jR}) is the Higgs doublet. The scale Λ may depend on the vevs of other scalar fields charged under $U(1)_F$.² With the vevs of H and ϕ , the flavon interactions with SM fermions are described by the effective Lagrangian (e.g., [4])

$$\mathcal{L} \supset -i \frac{a}{\Lambda} \sum_{f,i,j} \bar{f}_i [(m_j - m_i)v_{ij} + (m_j + m_i)a_{ij}\gamma_5] f_j, \quad (1)$$

where m_i is the mass of the mass eigenstate fermion field f_i , while a_{ij} and v_{ij} are the axial and vectorial coefficients, respectively. In Appendix A, we outline how such a Lagrangian could be derived from a UV Froggatt-Nielsen (FN) model.

Depending on the coefficients a_{ij} and v_{ij} , and given the mass proportionality in Eq. (1), the most intense interactions of the flavon are expected to involve the charged fermions of the third generation. Since a is a pseudoscalar field, the flavor-violating couplings v_{ij} imply parity violation. However, the CP symmetry is preserved if v_{ij} and a_{ij} are real, which is the case we will study.

It is also assumed that the flavon acquires a mass through an explicit breaking mechanism of the $U(1)_F$ symmetry, either via an anomalous feature, as in the case of the axion, or through a potential term like $V = (\lambda\phi^n/M^{n-4} + \text{H.c.})$. For our purposes, the exact origin of this mechanism is not relevant, so we assume a flavon mass m_a .

Our analysis will be based on the following hypotheses regarding the couplings: the flavon diagonal couplings with fermions and vector bosons, as well as the nondiagonal couplings with the down-type quarks, are all negligible, the dominant couplings with quarks are the nondiagonal ones involving the top quark plus charm or up quark, and the couplings with leptons play a subdominant role, as the flavon couples to the fermion mass. We note that subdominant diagonal couplings to fermions are sufficient to

²This is the flavon decay constant and may depend on the $U(1)_F$ charge c_i of the scalar fields φ_i acquiring vevs, as $\Lambda = \sqrt{2(\sum_i c_i^2 \langle\varphi_i\rangle^2)}$.

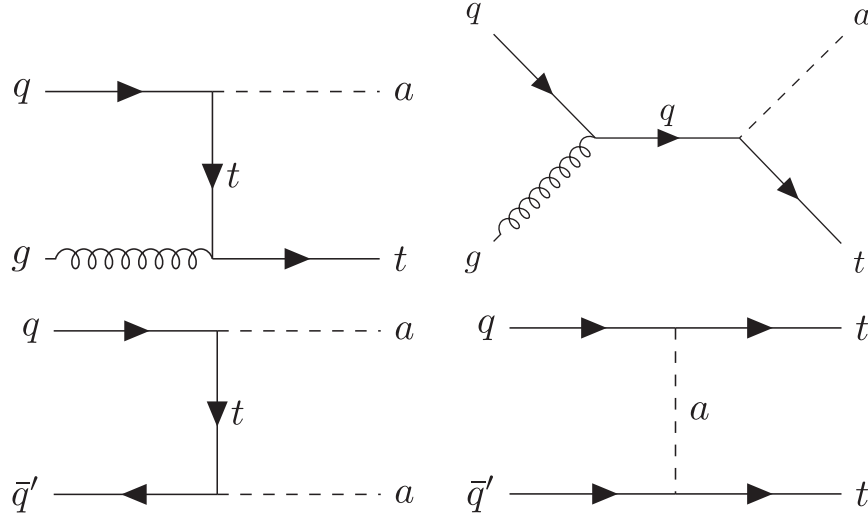


FIG. 1. Feynman diagrams for top-pair signals. The diagrams in the top row illustrate contributions from single flavon production. The lower left panel shows double flavon production, and the lower right diagram depicts the virtual flavon contribution in the t -channel. Here, $q = u, c$.

ensure that the corresponding couplings to neutral gauge bosons are also subdominant, since the latter arise at the loop level and are induced by the diagonal fermion couplings with the flavon. One example of a working model that has the assumed texture under the FN mechanism hypothesis is shown in Appendix B.³ This way, the set of parameters we consider is the following:

$$\begin{cases} m_a, \Lambda & \text{flavon mass and new physics scale,} \\ v_{tc}, a_{tc}, v_{tu}, a_{tu} & \text{nondiagonal top-quark couplings.} \end{cases} \quad (2)$$

These couplings avoid dangerous tree-level FCNCs and strong experimental constraints on flavor-diagonal interactions.

To further restrict the parameter space, we will also assume universal flavon couplings to quarks, that is, $v_{tc} = v_{tq} \equiv v_{tq}$ and $a_{tc} = a_{tq} \equiv a_{tq}$. Moreover, the cross sections will depend only on the ratio between $c_{tq} \equiv \sqrt{v_{tq}^2 + a_{tq}^2}$ and Λ , thus our searches can be performed on the c_{tq}/Λ versus the flavon mass, m_a . This parameter space reduction is derived directly from the assumption $m_a > m_t \gg m_q$, allowing us to effectively neglect the light-quark masses in the expressions for the decay rates $\Gamma(a \rightarrow tq)$ and, consequently, the production cross sections, as outlined in Appendix C.

Our analysis is restricted to flavons with masses in the 200–1400 GeV range, which preferentially decay into a top quark plus up or charm quarks. Even though the effective

flavon couplings to down-type quarks are assumed to be negligibly small, flavons can decay via $a \rightarrow bq$ with $q = \bar{d}, \bar{s}$ at one loop. However, this decay is suppressed compared to the top-quark decay channel since it arises from next-to-next-to-leading-order corrections [18].

III. SIGNAL AND BACKGROUND SIMULATIONS

We simulated signals and backgrounds in leading order at the 14 TeV LHC with MadGraph5 [20,21] using the UFO [22] files generated by FeynRules [23] from a private implementation of our EFT model. The parton shower and detector simulations were performed with default settings of Pythia 8.3 [24] and Delphes3 [25], respectively.

The signal we consider involves the production of two top quarks with the same or opposite charges and up to two non- b jets after the decay of flavons into a top quark plus a c or u quark. The W bosons from top quarks decay leptonically into electrons and muons plus neutrinos. We assume that flavons with masses higher than $m_t + m_c$ decay exclusively and equally into the dominating flavor-violating channels, $a \rightarrow t\bar{u}, t\bar{c}, \bar{t}u, \bar{t}c$. The signal receives the following contributions:

- Flavon single production, as shown in the two panels on the first row of Fig. 1. Again, we have opposite-sign (OS) and same-sign (SS) top-quark pairs, but with a non- b jet from the flavon decay. The flavon total width is small for all masses we consider in this work, so the narrow width approximation is valid to compute the cross sections which, as a consequence of $\text{BR}(a \rightarrow tq) \sim 1$ and $\text{BR}(t \rightarrow Wb) \sim 1$, factorize, in this case, as $(c_{tq}/\Lambda)^2$ (see Appendix C).
- Double flavon production, as shown in the lower left panel of Fig. 1. Now, we have OS and SS top-quark

³We are currently working on explicit realizations [choice of particle charges under $U(1)_F$ and other parameters] of the FN mechanism that fulfill those requirements. The results will be published elsewhere.

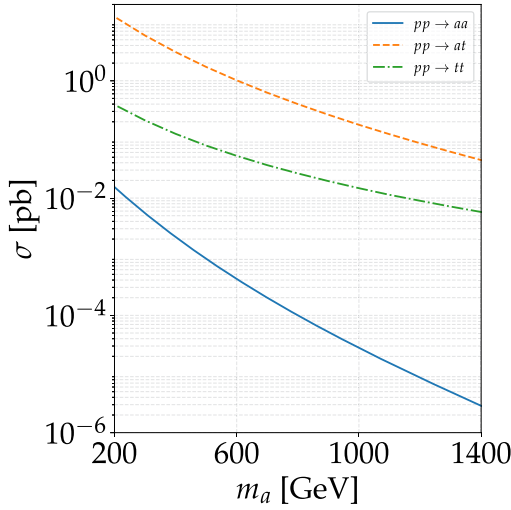


FIG. 2. Production cross sections for all three production modes considered in this work: pair production ($pp \rightarrow aa$), single production ($pp \rightarrow at$), and t -channel contribution ($pp \rightarrow tt$). In the last case, all combinations of top-quark charges are included.

pairs, along with two non- b jets from the flavon decay. The number of signal events from this contribution scales as $(c_{tq}/\Lambda)^4$ (see Appendix C).

- (c) Flavon in the t -channel, as shown in the lower right panel of Fig. 1. In this case, two top quarks of any charge combination are produced, tt , $\bar{t}\bar{t}$, or $t\bar{t}$. After decaying, the top quarks produce two b jets and SS or OS leptons ($\ell = e, \mu$). The number of signal events

from this contribution scales as $(c_{tq}/\Lambda)^4$ (see Appendix C). Because the final states of the three contributions differ, there are no interference effects at the leading order.

Summarizing, we propose to look for heavy flavons in the channels

$$pp \rightarrow tt, \bar{t}\bar{t} + \text{jets} \rightarrow bb(\bar{b}\bar{b}) + \ell^\pm \ell^\pm + E_T + \text{jets}, \quad (3)$$

$$pp \rightarrow t\bar{t} + \text{jets} \rightarrow b\bar{b} + \ell^\pm \ell^\mp + E_T + \text{jets}. \quad (4)$$

The production cross sections, in pb, for all three channels are shown in Fig. 2. In the mass range of interest, the single production dominates, followed by the t -channel contribution. Double flavon production is the most suppressed, primarily due to the flavon mass.

The dominant background for OS lepton signals is $t\bar{t}$ production, with subdominant contributions from next-to-next-to-leading-order Drell-Yan plus two b jets ($\ell\ell bb$) and single top tW production. SS lepton signals receive small reducible backgrounds from lepton and jet misidentification related to these sources. The irreducible background for SS signals is vector boson fusion (VBF), $qq' \rightarrow W^\pm W^\pm + bb$, which is CKM-suppressed when the W boson couples to bottom quarks. When it couples charm quarks to the initial-state up and strange quarks, there is no CKM suppression. However, the charm quark would be misidentified as a bottom quark. Moreover, the bottom jets would be produced in the forward regions of the detector, where b -tagging is less efficient, while the signals tend to

TABLE I. Efficiencies after the sequential cuts from Eq. (5), for SS leptons in the final state shown for selected signal masses and relevant backgrounds. The cuts are applied cumulatively in the following order: (I) at least one b jet, (II) missing transverse energy $\cancel{p}_T > 20$ GeV, (III) exactly two SS leptons, and (IV) dilepton invariant mass $M_{\ell\ell} > 100$ GeV. After all cuts, the only significant background source is $W^\pm W^\pm jj$, which is the only irreducible SS background.

Cut	200 GeV	700 GeV	1200 GeV	$t\bar{t}$	tW	$\ell^+ \ell^- bb$	$W^\pm W^\pm jj$
(I)	0.81	0.85	0.84	0.80	0.65	0.75	0.11
(II)	0.77	0.83	0.83	0.77	0.62	0.29	0.11
(III)	0.094	0.096	0.082	$\sim 10^{-4}$	$\sim 10^{-4}$	$\sim 10^{-4}$	0.03
(IV)	0.039	0.051	0.050	$\sim 10^{-5}$	$\sim 10^{-5}$	$\sim 10^{-5}$	0.02

TABLE II. Efficiencies after the sequential cuts from Eq. (5), for OS leptons in the final state shown for selected signal masses and relevant backgrounds. The cuts are applied cumulatively in the following order: (I) at least one b jet, (II) missing transverse energy $\cancel{p}_T > 20$ GeV, (III) exactly two opposite-sign leptons, and (IV) dilepton invariant mass $M_{\ell\ell} > 100$ GeV.

Cut	200 GeV	700 GeV	1200 GeV	$t\bar{t}$	tW	$\ell^+ \ell^- bb$
(I)	0.81	0.85	0.84	0.80	0.64	0.75
(II)	0.77	0.83	0.83	0.77	0.62	0.29
(III)	0.092	0.087	0.073	0.28	0.26	0.14
(IV)	0.036	0.047	0.042	0.12	0.12	0.13

be produced in the central regions. Yet, the VBF background contributes comparably to the reducible sources in the SS background event sample. We show the efficiency of the SS and OS cuts in Tables I and II, respectively.

IV. CUT ANALYSIS

In order to clean the signals from backgrounds, we apply kinematic cuts to both SS and OS signals and backgrounds. Particles that do not satisfy the following selection criteria are excluded⁴:

- (I) at least 1 b jet with $p_{T_{j,b}} > 20$ GeV, $|\eta_{j,b}| < 2.5$,
- (II) $\cancel{p}_T > 20$ GeV,
- (III) exactly 2 leptons with $p_{T_\ell} > 10$ GeV, $|\eta_\ell| < 2.5$,
- (IV) $M_{\ell\ell} > 100$ GeV. (5)

The b -tag efficiency is $\sim 50\%$ for $p_{T_b} = 20$ GeV, raising to $\sim 70\%$ at $p_{T_b} = 100$ GeV, as parametrized by Delphes3. On the other hand, requiring two leptons reduces the signal efficiency, especially for heavy masses, since it becomes more difficult to find isolated leptons when the flavon decay products become increasingly collimated, see the lower left plot of Fig. 4. The cut on the invariant mass of two leptons, $M_{\ell\ell}$, helps to tame the large Drell-Yan background. In the high-luminosity regime, the transverse momentum threshold for observation might be raised for trigger requirements in the presence of higher pile-up. We expect, however, that a higher p_T should impact the background efficiency more than the signal.

In Table III, we display the cross section times' acceptance after imposing the cuts of Eq. (5), taking hadronization and detector effects into account. The background rates are still much larger than the signals. Selecting only SS lepton events thus eliminates most backgrounds. However, since SS and OS events are easily separable, in the next section, we will investigate whether OS events can be separated from backgrounds utilizing a neural network classifier. For now, we will select only the much cleaner SS events. As shown in Table III, the expected number of backgrounds is approximately 3100, compared to at least 400 signal events, after at least 3000/fb.

Considering only same-sign leptons, the results for the 95% confidence level (CL) and 5σ discovery regions in the coupling versus mass plane are shown in Fig. 3. The region $m_a > \Lambda$, highlighted in magenta, refers to the validity limit of the EFT. The orange-shaded region is the one excluded at the 95% CL by the ATLAS study [13].

⁴We require exactly two opposite-sign leptons to enforce a clean dilepton topology and suppress contamination from events with extra leptons as $t\bar{t}$ + jets and dibosons.

TABLE III. The cross sections after the cuts of Eq. (5) for both SS and OS events are shown for backgrounds and for signal benchmarks with flavon masses of 200, 600, 1000, and 1400 GeV. The signal cross sections were computed assuming $\Lambda = 1$ TeV, and the resulting event yields are presented for an integrated luminosity of 3000/fb.

Type	SS cross section	SS event yield	OS cross section	OS event yield
$t\bar{t}$	0.69 fb	2100	3.1 pb	9.3×10^6
tW	0.06 fb	189	0.34 pb	1.0×10^6
$\ell^+\ell^-bb$	4.2 ab	12	0.02 pb	7.2×10^4
W^+W^-jj	3.3 ab	10	17.1 fb	5.1×10^4
$W^\pm W^\pm jj$	0.26 fb	780	0.05 ab	0.15
200 GeV	22.75 fb	6.8×10^4	22.31 fb	6.7×10^4
600 GeV	2.87 fb	8611	2.91 fb	8755
1000 GeV	0.57 fb	1700	0.57 fb	1690
1400 GeV	0.14 fb	425	0.12 fb	374

The statistical significances were computed using the Asimov formula [26], which accounts for both statistical and systematic uncertainties in the background. In the background-dominated regime ($B \gg S$) and in the absence of systematics ($\epsilon_{\text{syst}} = 0$), the significance increases with luminosity as $\sqrt{L}$. When a nonzero systematic uncertainty is included, the denominator acquires an additional $(\epsilon_{\text{syst}}B)^2$ term, and once this dominates over the statistical fluctuations $\sqrt{B}$, the significance asymptotically approaches $Z_A \sim S/(\epsilon_{\text{syst}}B)$. This saturation limit is independent of luminosity, implying that beyond a certain threshold, improvements in L no longer enhance the sensitivity. Meaningful gains from luminosity increases restrict systematic errors in the background yields to scenarios up to 5%.

Both neutral and charged currents can turn $q\bar{q}'$ bound states into $q'\bar{q}$, causing a meson-antimeson oscillation. In particular, a $D(\bar{u}c)$ meson can oscillate into $\bar{D}(\bar{c}u)$, leading to a measurable mass shift [27]. In our flavon model, this phenomenon is caused by a box-type diagram where the flavon couples u or c quarks to top quarks. We adapted the results from [28,29] to our case in the scenario where the scalar and axial couplings are equal, $v_{tq} = a_{tq}$, to estimate the impact of those constraints on the model, see Appendix D. We depict, in Fig. 3, the points in the parameter space that can explain the $D - \bar{D}$ data with its 68% CL band from the observable value of the dimensionless mass difference of the D -meson system, $x_D = (4.34_{-1.39}^{+1.26}) \times 10^{-3}$ [27]. Currently, the SM prediction is not completely clear due to nonperturbative QCD contributions [28], so we can consider two scenarios: exclusion, where if the data confirm the SM, the region above the upper 1σ blue line in Fig. 3 will be excluded, and discovery, where if the SM turns out to be in disagreement with the data, then the blue band represents points where our model can explain the $D - \bar{D}$ mixing.

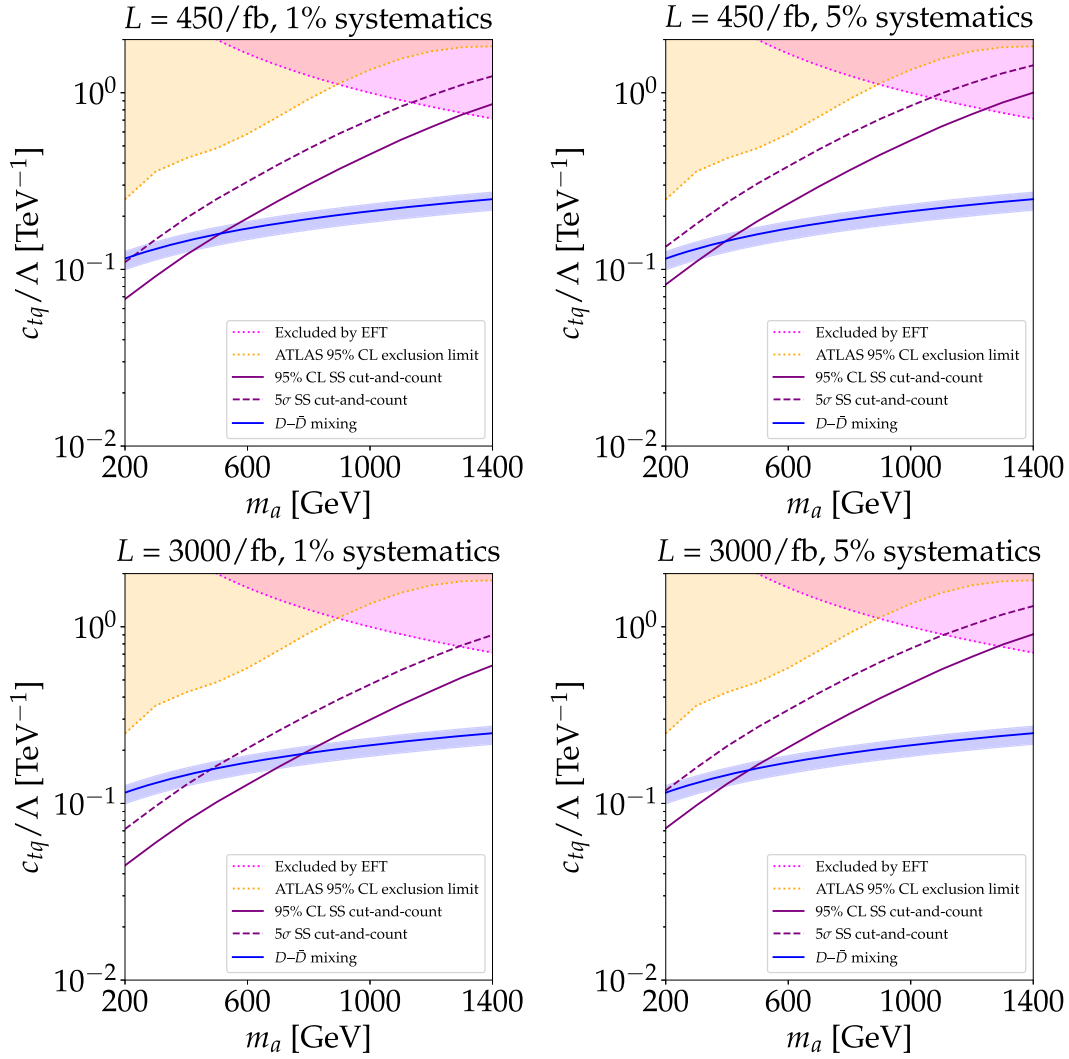


FIG. 3. The 95% CL and 5σ limits for the cut-based approach are shown for integrated luminosities of 450/fb and 3000/fb at the 14 TeV LHC. The region where $m_a > c_{tq}/\Lambda$, where the EFT is not valid, is shown in magenta. The orange band represents the ATLAS constraint for a flavon with no diagonal couplings [13]. The cut-based significance appears insensitive to luminosity because the dominant background ($t\bar{t}$) is already in the systematics-dominated regime. In this case, the significance $Z \sim S/(\epsilon_{\text{sys}}B)$ saturates and no longer grows with L , explaining why increasing the integrated luminosity does not change the extracted c_{tq}/Λ for a fixed target significance.

The upper panels of Fig. 3 show that the D -meson constraints already exclude the region that the LHC could discover flavons with 450/fb. There is some room for exclusion at the 95% CL, though. With 3000/fb, some complementarity between D -meson data and LHC data is possible. Moreover, the 14 TeV HL-LHC can probe models with coupling $0.04 \lesssim c_{tq}/\Lambda \lesssim 0.2 \text{ TeV}^{-1}$ and masses up to $\sim 700 \text{ TeV}$.

V. MULTIVARIATE ANALYSIS

We trained neural networks to separate signals and backgrounds to maximize the signal significance for a given mass hypothesis for both OS and SS events, despite

the challenging channel being the OS one. The coupling only rescales the remaining number of events after the neural network separation, but the mass impacts the distribution of the kinematic variables of the signals. In principle, a neural network trained for a given flavon mass might perform poorly in identifying signals of other masses, which would force us to train many algorithms to account for the change in the flavon mass parameter. It was demonstrated in [15,16] that a neural network parametrized by a small, fixed grid of mass parameters can learn how to interpolate between the masses of the grid used to train the algorithm. In our case, we trained our pNN with signal events of the mass grid: 200, 500, 800, 1100, and 1400 GeV. We then used the model to classify events of

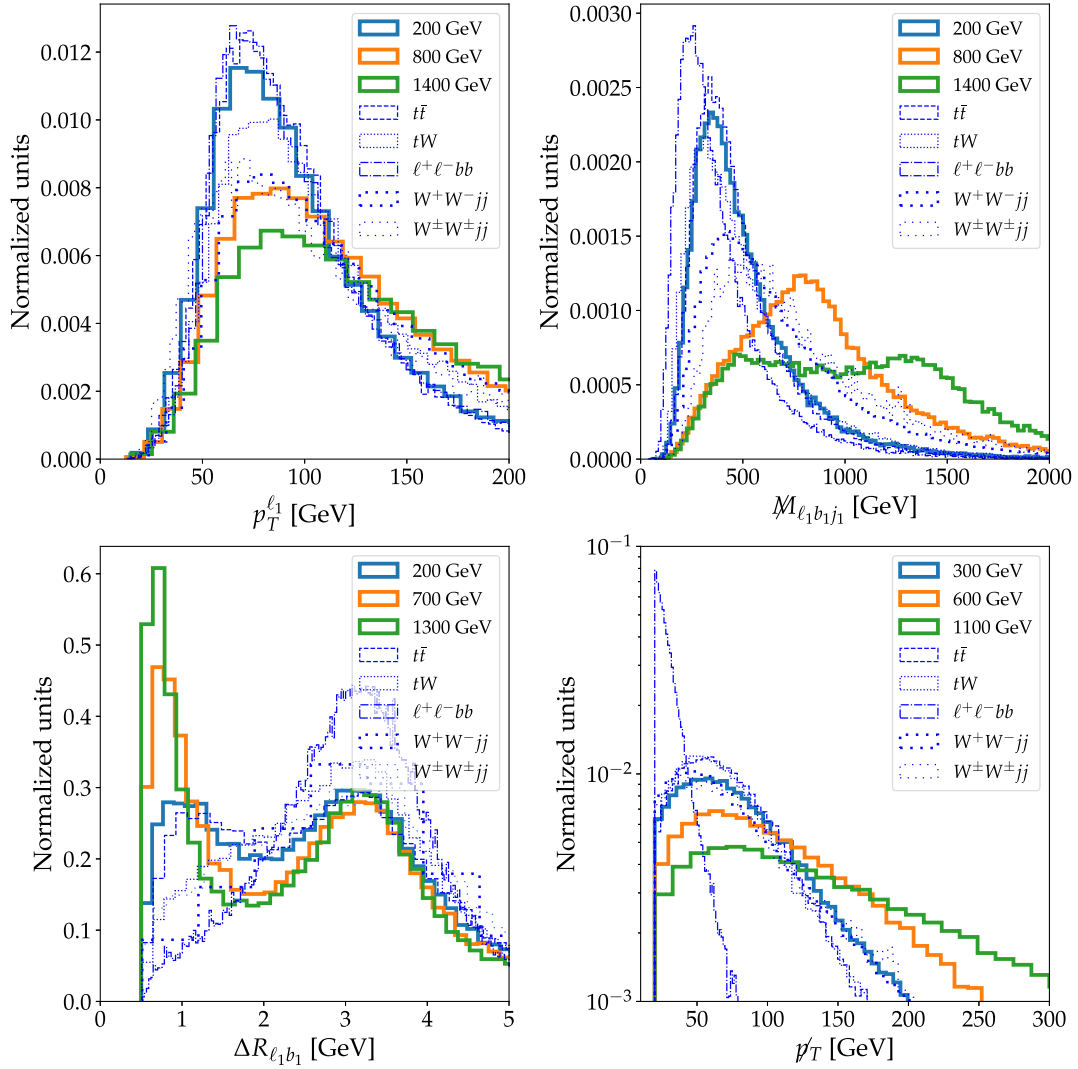


FIG. 4. Examples of kinematic distributions used as high-level variables in the pNN training. For smaller values of m_a , all distributions closely resemble those of the $t\bar{t}$ and tW backgrounds. As m_a increases, the signal distributions become more distinct, which partly explains why the pNN achieves higher accuracy in identifying signals with larger m_a .

these and eight other different masses. We verified that the pNN learns to distinguish the intermediate masses correctly by training dedicated neural networks for a few intermediate masses.

The variables used to represent the events may facilitate the performance of the neural networks to separate the signal and background events. Many high-level features can be constructed with the four-momenta of the leptons, bottom-, charm-, and light-quark jets. We tested many configurations and found a useful representation of our events in terms of the transverse momenta, rapidity, and energy in the lab frame of the jets and leptons, the missing transverse momentum of the events, invariant mass combinations, differences in the azimuthal angle, and the difference in the $\eta - \phi$ plane: $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Some of these variable distributions are shown in Fig. 4. As expected, the lepton p_T and the missing transverse

momentum spectra become harder as the flavon mass increases. Additionally, for heavy flavons, the decay products tend to become collimated. This effect is visible in the $\Delta R_{\ell_1 b_1}$ distribution, where the heavier flavons are peaked toward smaller distribution values. The $M_{\ell_1 b_1 j_1}$ is a proxy for the total mass of the system, which cannot be reconstructed without the neutrinos. This distribution receives contributions from the three production mechanisms. The subleading one, $pp \rightarrow t\bar{t}$, populates the low-mass bins, while the resonant ones, the high-mass bins, are mainly the dominant process $pp \rightarrow at$. The onset of the latter contribution can be seen in the features of the solid lines for 200, 800, and 1400 GeV flavons, where a resonance starts to rise.

To train the neural network, we divided the available data into two parts: one for training and validation, and a separate portion reserved for the final test.

TABLE IV. Hyperparameters and neural network architecture used to train the pNN. The hyperparameters were optimized with Hyperopt [30].

Hidden layers	Neurons	Activation	Drop-out	Learning rate	Batch size
3	80	ELU	0.265	0.00156	80

Hyperparameter optimization was performed with the Hyperopt package [30], using a stratified 10-fold cross-validation scheme. For each trial set of hyperparameters (learning rate, batch size, number of hidden layers, activation functions, dropout rate, and number of neurons per

layer), the network was trained 10 times, each time holding out a different fold for validation and using the remaining nine for training. The average accuracy across the 10 folds was used to rank the trials, and the best-performing hyperparameters are reported in Table IV.

Once the hyperparameters were fixed, the final training was also carried out within the 10-fold cross-validation framework, using the Adam optimizer [31] to minimize the binary cross-entropy loss function. To reduce the risk of overfitting, training was stopped if the validation loss did not improve for 25 consecutive epochs, with a maximum limit of 200 epochs. The 10 resulting models were later combined to provide central predictions and their associated variability.

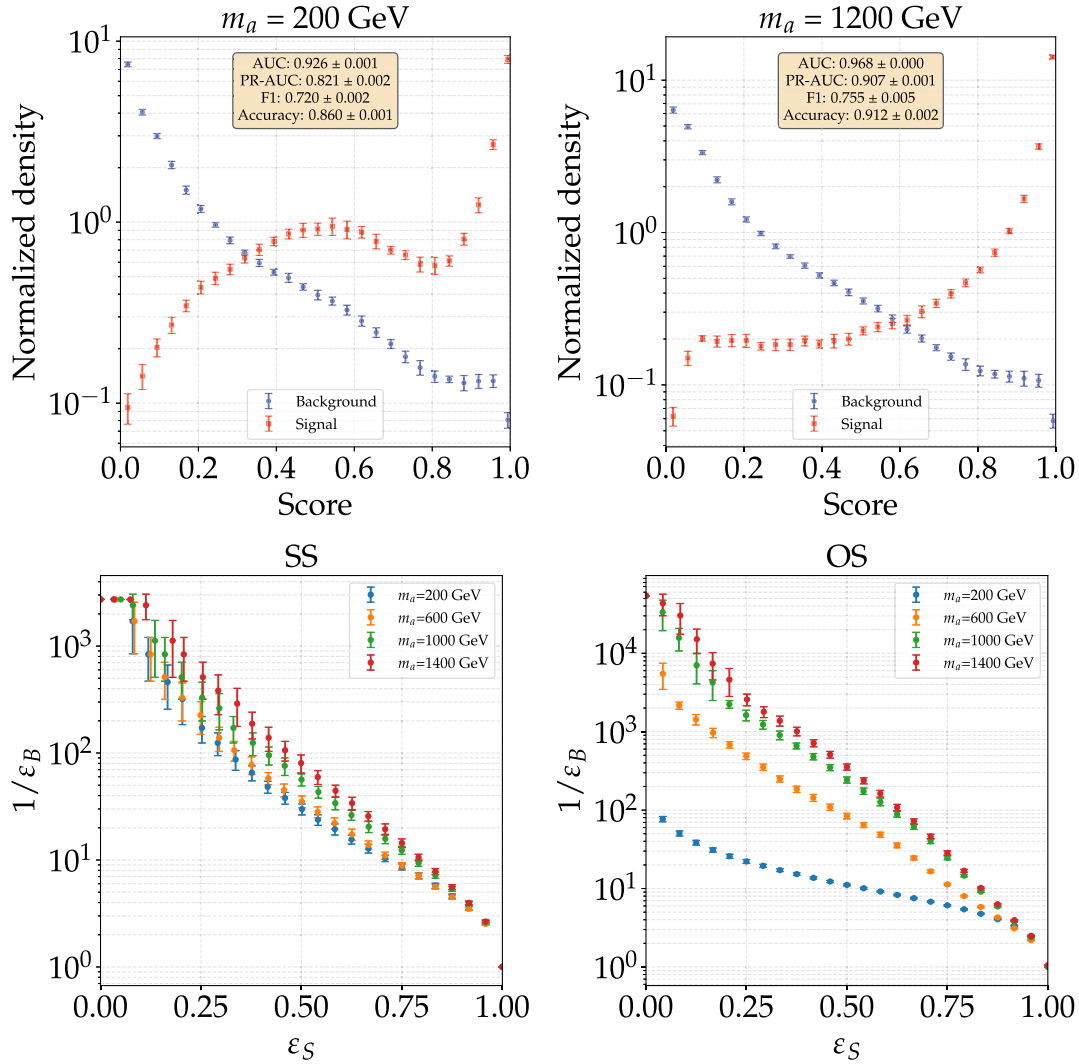


FIG. 5. Upper panels: output scores of signal (red) and background (blue) from the pNN classifier. Higher flavon masses are more easily separated, as their kinematic distributions become increasingly distinct from those of the background. Lower panels: background rejection, $1/\epsilon_B$, as a function of signal efficiency, ϵ_S , obtained by applying cuts on the pNN output scores for the SS and OS cases. At fixed signal efficiency, the rejection improves with increasing flavon mass. The SS channel provides better rejection at low masses, while OS dominates at higher masses. Vertical error bars indicate the standard deviation of the sample mean across the 10 cross-validation folds, reflecting the variability of the metrics.

To ensure a reliable classification, the signal-to-background ratio was kept approximately balanced in the Monte Carlo training samples, following the strategy of Ref. [16]. The training dataset consisted of 270,000 events, evenly distributed among signals from the mass grid and the five background sources. Each of the $t\bar{t}$, tW , W^+W^-jj , and $\ell^+\ell^-bb$ backgrounds was represented by 31,750 events, while $W^\pm W^\pm jj$ contributed 8000 events. This smaller number reflects the reduced efficiency of producing $W^\pm W^\pm jj$ events that satisfy Eq. (5), as also seen in the event yields of Table III. We verified that the pNN successfully distinguished all classes under these conditions.

After training and cross-validation, the models were applied to the independent test set, which contains signal events with masses from 200 to 1400 GeV in 100 GeV steps, including masses not present in the training grid. As illustrated in Fig. 5, the pNN maintains nearly the same performance on these interpolated masses as on the trained points, with the reported metrics reflecting the averaged predictions across the 10-fold models.

The upper panels of Fig. 5 show the distribution of the signal (red) and background (blue) events as functions of the classifier output score. Lighter flavons are more difficult to discriminate from the background, as could have been anticipated from the distributions in Fig. 4. In the lower panels of Fig. 5, we depict the algorithm performance for each test mass in terms of the signal efficiency, ε_S , and background rejection, $1/\varepsilon_B$, for events that survive a given threshold score cut for SS and OS events. The neural network presents a good discerning power in both cases.

Still, it is best in the OS scenario for smaller masses, where it keeps, for example, only $\sim 10\%$ (0.1%) of background events for a signal efficiency of 50% in the best(worst) mass cases. The accuracy also increases from light to heavy flavon masses, as can be seen from the upper panels of Fig. 5.

After training the pNN, the number of surviving signal and background events was determined by applying cuts on the test set scores. The cut on the pNN output was optimized to maximize the SS and OS statistical significances, subject to the requirement of a minimum number of signal events: two events for a 95% CL exclusion and 25 events for a 5σ discovery, following the Poisson likelihood treatment in the low-background regime [26].

Unlike the traditional cut-and-count method, the pNN actively identifies regions of phase space with the highest expected significance. By selecting score thresholds where the signal-to-background ratio is largest, it effectively suppresses the background contribution and alleviates the saturation of sensitivity in the systematics-dominated regime, thereby preserving sensitivity to increasing luminosity.

In Fig. 6, we present the limits obtained from the multivariate analysis in comparison with those from the cut-based approach. The use of neural networks improves the coupling limits by about 20% for a 200 GeV flavon and up to 50% for a 1400 GeV flavon, based on the average over the 10-fold cross-validation. To exploit both channels, we combine the OS and SS signal significances in quadrature, although the SS contribution dominates. Since the OS and SS channels share common sources of systematic

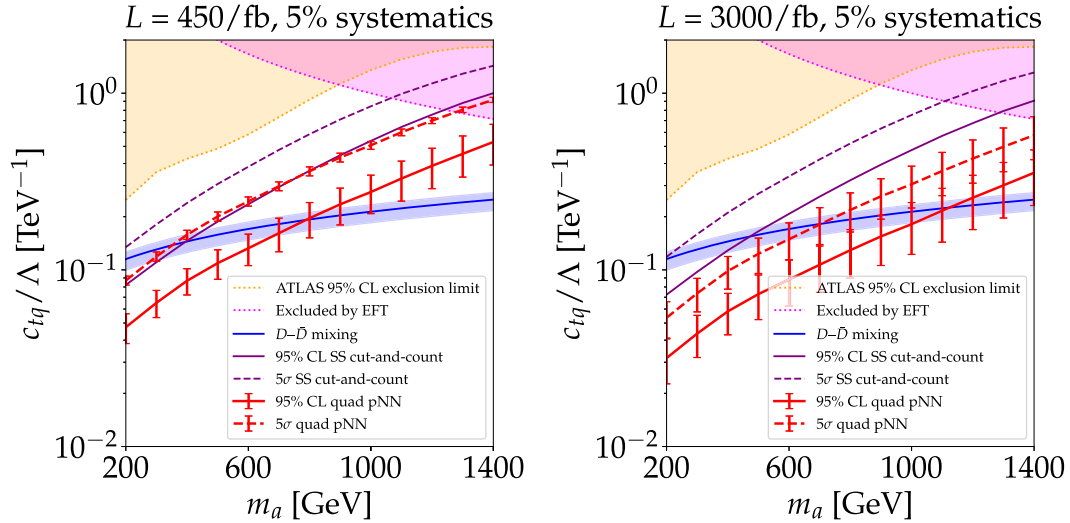


FIG. 6. Exclusion (95% CL) and discovery limits at the 14 TeV LHC with 450/fb (left) and 14 TeV LHC with 3000/fb (right). For the pNN limits, the error bars correspond to the standard deviation of the sample mean across the 10 folds of the cross-validation. The blue line corresponds to the current experimental limit [27], with the shaded blue band representing the 1σ uncertainty. The magenta region, where $m_a > c_{tq}/\Lambda$, lies beyond the validity of the EFT description adopted in this work. The orange band shows the ATLAS constraint for flavons without diagonal couplings [13]. Our model can accommodate the D -meson constraint for flavon masses up to 700 GeV after 450/fb and up to 1.1 TeV after 3000/fb.

uncertainty (e.g., luminosity and b -tagging), the effect of correlations is expected to be small. Nevertheless, a full profile likelihood combination would provide the most accurate treatment, even if the results are not expected to differ significantly from those shown here. With 450/fb we see, in the left panel of Fig. 6, that masses up to ~ 700 GeV can be probed with couplings in the interval 0.05 to 0.1/TeV, considering the D -meson constraint. After accumulating 3000/fb, however, the flavon mass that can be probed extends to ~ 1100 GeV in the coupling interval of ~ 0.03 to 0.2/TeV, as we can see in the right panel of Fig. 6. If the D -meson data confirm new physics, we see that a flavon of 1 TeV and coupling of ~ 0.2 /TeV can be discovered at the HL-LHC.

VI. CONCLUSIONS

Flavor-violating processes at colliders are an interesting avenue for new physics searches and have been intensively pursued by LHC collaborations. Contrary to light quarks and leptons, flavor-violating interactions involving top quarks and up or charm quarks are much less constrained by searches for FCNCs. Models like the Froggatt-Nielsen mechanism, which aim to explain the Yukawa hierarchy, predict the existence of an axionlike particle that can mediate such flavor-violating interactions, referred to as the flavon. We assume these scalars predominantly change the flavor of the fermions they interact with, thereby avoiding constraints from flavor-violation searches such as $D - \bar{D}$ oscillations, as well as from direct searches at colliders.

This work explored a distinctive phenomenology of flavons at the 14 TeV LHC within an effective field theory framework. Pair and single flavon production, along with virtual t -channel contributions, give rise to pairs of OS and SS top quarks. The decay of SS top quarks into SS leptons leads to an almost background-free signal, albeit at the cost of halving the number of signal events. A cut-and-count strategy can probe flavon couplings of the order of 10^{-2} TeV $^{-1}$. Rather than discarding all OS top-quark events, which are subject to higher background, we trained a neural network to enhance the signal relative to $t\bar{t}$ and other backgrounds. We employed parametrized neural networks to obtain a smooth and convenient classifier as a function of the flavon mass. After classification, the samples were separated according to the lepton-pair charge, and the SS and OS signal significances were combined in quadrature, thereby making full use of the flavon production channels.

Our analysis shows that training the pNN also with the SS signals provides the best performance, improving the reach for couplings by up to 50% compared to the cut-based analysis. Although the pNN performs very well in cleaning the OS channel, especially for heavy flavons, the OS contribution is subdominant when the statistical significances are combined in quadrature. The use of more

advanced algorithms in the future may further improve the results. Previous experimental studies used multivariate analysis to separate signals and backgrounds in the multi-top channel [13] and top + flavon [11], and we confirm its usefulness and convenience in increasing the significance of the signal over the full range of flavon mass considered.

The 450/fb scenario can exclude 200(700) GeV flavons for couplings as weak as $c_{tq}/\Lambda \sim 5.5 \times 10^{-2} (2 \times 10^{-1})$ TeV $^{-1}$ at the 14 TeV HL-LHC. It can also discover 200(350) GeV flavons with couplings $c_{tq}/\Lambda \gtrsim 9 \times 10^{-2} (1.2 \times 10^{-1})$ TeV $^{-1}$. With 3000/fb, the 14 TeV LHC can exclude 200(1100) GeV flavons for couplings as weak as $c_{tq}/\Lambda \sim 3.5 \times 10^{-2} (2 \times 10^{-1})$ TeV $^{-1}$. The discovery prospects are also promising, as 200 (700) GeV flavons can be discovered for couplings $c_{tq}/\Lambda \gtrsim 5 \times 10^{-2} (1.5 \times 10^{-1})$ TeV $^{-1}$. Our findings extend previous searches for light flavons up to 10 GeV [19], and heavy flavons with masses up to ~ 200 GeV, where decays into tau leptons are favored [18]. We also estimate the impact of the $D - \bar{D}$ oscillation data on the model's parameters. We found that, in the case where the data confirm the SM predictions, the HL-LHC will be able to probe flavons with masses up to ~ 1 TeV.

ACKNOWLEDGMENTS

We thank Andreas Crivellin for the helpful and enlightening discussions. We thank our dear friend Alex Gomes Dias for all these years being a true scientist with an endless effort to refine each work he was involved in, for all his ideas, and for his personality, which was a magnet for great people. You will always be missed. This study was financed in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), via the Grants No. 307317/2021-8 (A.A.) and No. 308280/2023-7 (A.G.D.), and in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Finance Code 001 (D.S.V.G.). A.A. also acknowledges support from the FAPESP (No. 2021/01089-1) Grant. E.d.S.A. partially thanks FAPESP for its support in the initial steps of this work (Grant No. 2022/07185-5).

DATA AVAILABILITY

The data that support the findings of this article are openly available [32].

APPENDIX A: EFFECTIVE LAGRANGIAN

In FN models, SM Yukawa operators arise from higher-dimensional operators

$$\mathcal{L} \supset y_{ij} \left(\frac{\phi}{M} \right)^{n_{ij}} \bar{F}_{iL} \Phi f'_{jR} + \text{H.c.}, \quad (\text{A1})$$

where the flavon multiplet is written nonlinearly as

$$\phi(x) = \frac{1}{\sqrt{2}}(\rho(x) + \sqrt{2}\langle\phi\rangle)e^{ia(x)/\Lambda}. \quad (\text{A2})$$

After electroweak symmetry breaking, $\Phi \rightarrow (0, (v+h)/\sqrt{2})^T$, and defining $\varepsilon \equiv \langle\phi\rangle/M$, one gets

$$(Y_f)_{ij} = y_{ij}\varepsilon^{n_{ij}}, \quad (M_f)_{ij} = \frac{v}{\sqrt{2}}(Y_f)_{ij}. \quad (\text{A3})$$

Because ϕ appears with a phase, Eq. (A1) includes a factor $e^{in_{ij}a/\Lambda}$. Expanding to first order in a/Λ gives, in the gauge/flavor basis,

$$\mathcal{L} \supset \frac{ia}{\Lambda} \bar{f}_{Li}(M_f)_{ij}n_{ij}f_{Rj} + \text{H.c.} \quad (\text{A4})$$

Now, we rotate to mass eigenstates $f_{L,R} \rightarrow U_{L,R}^f f_{L,R}$ with $U_L^{f\dagger} M_f U_R^f = \text{diag}(m_1, \dots)$, and define the charge misalignment matrices

$$X_L \equiv U_L^{f\dagger} n U_L^f, \quad X_R \equiv U_R^{f\dagger} n U_R^f. \quad (\text{A5})$$

Using Eq. (A4) and standard identities for Dirac bilinears, it is convenient to trade the interaction for a derivative coupling to flavor currents (equivalent at leading order in derivatives),

$$\mathcal{L} \supset -\frac{\partial_\mu a}{\Lambda} \bar{f}_i \gamma^\mu (v_{ij} - a_{ij} \gamma^5) f_j, \quad (\text{A6})$$

where v and a are real, Hermitian matrices (for CP -conserving benchmarks) that encode the misalignment between $U(1)_F$ charges and fermion mass eigenstates (they are functions of $X_{L,R}$).

Integrating by parts and using the fermion equations of motion between on-shell external states,

$$\partial_\mu (\bar{f}_i \gamma^\mu f_j) = i(m_i - m_j) \bar{f}_i f_j, \quad (\text{A7})$$

$$\partial_\mu (\bar{f}_i \gamma^\mu \gamma^5 f_j) = i(m_i + m_j) \bar{f}_i \gamma^5 f_j, \quad (\text{A8})$$

one obtains the nonderivative interaction

$$\mathcal{L} \supset -\frac{ia}{\Lambda} \bar{f}_i [(m_j - m_i)v_{ij} + (m_j + m_i)a_{ij}\gamma^5] f_j, \quad (\text{A9})$$

which is exactly Eq. (1) in the manuscript. The diagonal entries are proportional to fermion masses and can be suppressed by suitable charge assignments, while off-diagonal entries are unsuppressed when charge and mass bases are misaligned. With real v_{ij} and a_{ij} , the interaction is CP -conserving, as used throughout our analysis.

APPENDIX B: EXAMPLE OF A FLAVOR-VIOLATING-DOMINATED FN MODEL

It is possible to realize models within the FN framework that exhibit the generic features adopted as working assumptions in this study, namely, scenarios with dominant couplings to flavor-violating fermion pairs. For illustrative purposes, we present here one such model.

In the FN mechanism, unlike in the SM, the Yukawa matrices in the unbroken phase are taken to be generic complex matrices with entries of order unity. A convenient parametrization of a 3×3 Yukawa matrix is

$$M = \begin{pmatrix} g_{11}\varepsilon^{n_{11}} & g_{12}\varepsilon^{n_{12}} & g_{13}\varepsilon^{n_{13}} \\ g_{21}\varepsilon^{n_{21}} & g_{22}\varepsilon^{n_{22}} & g_{23}\varepsilon^{n_{23}} \\ g_{31}\varepsilon^{n_{31}} & g_{32}\varepsilon^{n_{32}} & g_{33}\varepsilon^{n_{33}} \end{pmatrix} = \begin{pmatrix} |g_{11}|e^{i\theta_{11}}\varepsilon^{n_{11}} & e^{i\theta_{12}}\varepsilon^{n_{12}} & e^{i\theta_{13}}\varepsilon^{n_{13}} \\ e^{i\theta_{21}}\varepsilon^{n_{21}} & |g_{22}|e^{i\theta_{22}}\varepsilon^{n_{22}} & e^{i\theta_{23}}\varepsilon^{n_{23}} \\ e^{i\theta_{31}}\varepsilon^{n_{31}} & e^{i\theta_{32}}\varepsilon^{n_{32}} & |g_{33}|e^{i\theta_{33}}\varepsilon^{n_{33}} \end{pmatrix}, \quad (\text{B1})$$

where the coefficients g_{ij} are assumed to be complex numbers of order 1, n_{ij} denote charge-dependent exponents, and ε is the expansion parameter associated with the FN symmetry breaking.

The set of charges and the value of ε determine the predictions for fermion mass hierarchies, CKM matrix elements, and flavon decay branching ratios into charged lepton pairs. Since the phases θ_{ij} and relative magnitudes of g_{ij} are not fixed *a priori*, different realizations of the $\mathcal{O}(1)$ parameters lead to different numerical predictions. To capture the statistical behavior of these variations, we sample the parameter space and report the mean and standard deviation of the resulting distributions. In practice, we generate $N_{\text{samples}} = 25,000$ random realizations. The model under consideration is specified by the FN charges given in Eq. (B2),

$$\begin{pmatrix} q_{Q_1} & q_{Q_2} & q_{Q_3} \\ q_u & q_c & q_t \\ q_d & q_s & q_b \\ q_{L_1} & q_{L_2} & q_{L_3} \\ q_e & q_\mu & q_\tau \end{pmatrix} = \begin{pmatrix} 2 & 4 & 4 \\ 0 & -4 & 2 \\ -1 & -2 & -1 \\ -4 & -3 & -3 \\ 0 & 2 & 4 \end{pmatrix}. \quad (\text{B2})$$

For this benchmark model we fix the FN expansion parameter to

$$\varepsilon = 0.175. \quad (\text{B3})$$

The diagonal entries of the Yukawa matrices are sampled with magnitudes $|g_{ii}| \in [0.5, 5]$, while the phases θ_{ij} are drawn uniformly from the interval $[0.01, 2\pi]$. The resulting branching fractions of the flavon are displayed in Fig. 7, together with the standard deviation of the sample mean. Despite the variation induced by the random $\mathcal{O}(1)$ parameters, the dominant decay mode remains $a \rightarrow c\bar{t}$.

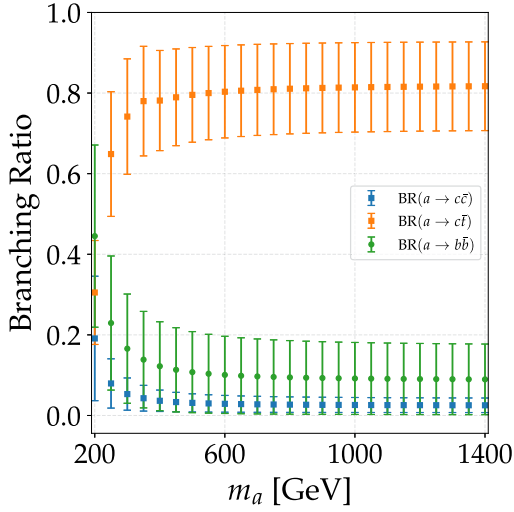


FIG. 7. Branching ratios of the flavon in the FN benchmark model. Error bars indicate the standard deviation of the sample mean obtained from 25,000 random realizations.

Flavor-conserving final states are also present with non-negligible rates, but they are suppressed by approximately 1 order of magnitude compared to the leading flavor-violating channel. This feature renders the benchmark a representative example of a flavor-violating FN realization.

Table VI summarizes the predictions for the charged-fermion masses obtained from the sampling procedure, compared against their experimental values. Although the central values do not perfectly reproduce the observed spectrum, the correct mass hierarchies are captured within order-of-magnitude accuracy, consistent with the FN expectation.

Similarly, Table V presents the CKM matrix elements. The sampled means do not exactly reproduce the experimental central values but consistently capture the observed hierarchical structure and approximate magnitudes of the entries.

TABLE V. CKM matrix elements in the FN benchmark model. The first column shows the experimental values, while the second and third columns give, respectively, the mean and the standard deviation of the sample mean obtained from 25,000 random realizations.

Particle	Value [GeV]	Mean [GeV]	std [GeV]
m_u	0.00216	0.000386	0.0002
m_c	1.27	1.35	0.8
m_t	172.7	175.5	0.8
m_d	0.00467	0.0127	0.006
m_s	0.0934	0.0792	0.03
m_b	4.18	2.81	1
m_e	0.000511	0.000992	0.0007
m_μ	0.1057	0.06313	0.03
m_τ	1.776	1.417	0.07

TABLE VI. Charged-fermion masses in the FN benchmark model. The first column shows the experimental values, while the second and third columns give, respectively, the mean and the standard deviation of the sample mean obtained from 25,000 random realizations.

Parameter	Value	Mean	std
$(V_{\text{CKM}})_{11}$	0.974	0.779	0.2
$(V_{\text{CKM}})_{12}$	0.225	0.541	0.2
$(V_{\text{CKM}})_{13}$	0.00369	0.0206	0.01
$(V_{\text{CKM}})_{21}$	0.225	0.54	0.2
$(V_{\text{CKM}})_{22}$	0.973	0.775	0.2
$(V_{\text{CKM}})_{23}$	0.0418	0.0905	0.04
$(V_{\text{CKM}})_{31}$	0.00857	0.0478	0.03
$(V_{\text{CKM}})_{32}$	0.0411	0.0765	0.04
$(V_{\text{CKM}})_{33}$	0.999	0.995	0.004

APPENDIX C: THE EMERGENCE OF THE EFFECTIVE c_{tq} COUPLINGS

Let us exemplify the factorization of the couplings with the decay rate $a \rightarrow tq$, which is given by

$$\Gamma(a \rightarrow tq) = \frac{N_c [(m_a^2 - (m_t - m_q)^2)(m_t + m_q)^2 a_{tq}^2 + (m_a^2 - (m_t + m_q)^2)(m_t - m_q)^2 v_{tq}^2]}{8\pi m_a^3 \Lambda^2} \times \sqrt{(m_a^2 - (m_t - m_q)^2)(m_a^2 + (m_t + m_q)^2)},$$

where $N_c = 3$. The assumptions we take are $m_a > m_t \gg m_q$, allowing us to effectively take the limit $m_q \rightarrow 0$, yielding

$$\Gamma(a \rightarrow tq) \approx \frac{N_c m_a m_t^2}{8\pi} \left(1 - \frac{m_t^2}{m_a^2}\right)^2 \frac{(a_{tq}^2 + v_{tq}^2)}{\Lambda^2},$$

so, we define $c_{tq} = \sqrt{a_{tq}^2 + v_{tq}^2}$.

Now, the narrow-width approximation can be used to calculate the cross sections we are investigating. Assuming $\text{BR}(t \rightarrow Wb) \sim 1$, the $pp \rightarrow at$ signal scales as

$$\sigma(pp \rightarrow at \rightarrow \ell^+ \ell^+ bbj) \propto \sigma(pp \rightarrow at) \text{Br}(a \rightarrow tq) \sim \left(\frac{c_{tq}}{\Lambda}\right)^2 \sigma_0(pp \rightarrow at),$$

where σ_0 is the baseline production cross section with $c_{tq} = 1$.

By similar calculations, we find that the signal $pp \rightarrow aa$ scales as

$$\sigma(pp \rightarrow aa \rightarrow \ell^+ \ell^+ bbjj\nu_\ell \nu_\ell) \propto \sigma(pp \rightarrow aa) \text{Br}(a \rightarrow tq)^2 \sim \left(\frac{c_{tq}}{\Lambda}\right)^4 \sigma_0(pp \rightarrow aa),$$

while $pp \rightarrow tt$ scales as

$$\sigma(pp \rightarrow tt \rightarrow \ell^+ \ell^+ bb) \propto \left(\frac{c_{tq}}{\Lambda}\right)^4 \sigma_0(pp \rightarrow tt).$$

APPENDIX D: $D-\bar{D}$ MIXING

The dimensionless mass difference of the D -meson system in models with flavor-violating couplings between the scalar sector and fermions was calculated in Ref. [28], where the Cheng-Sher ansatz [33] was used to set the texture of the flavor-violating couplings in the mass basis. Since we do not assume this ansatz, we quote their result in terms of general flavor-violating couplings $\xi_{h^0 f_i f_j}$,

$$x_D \approx \frac{\xi_{h^0 ut}^2 \xi_{h^0 ct}^2}{3\pi^2 m_a^2 \Gamma_D} f_D^2 M_D B_D. \quad (\text{D1})$$

where Γ_D and M_D are the averaged total decay width and mass of the D -meson system, f_D is the decay constant of D , B_D is the matrix element associated with the $\bar{u}c \leftrightarrow u\bar{c}$ transition, and the approximate symbol reflects the order-1 contributions of the EFT matching and the box loop

function. We thus identify the equivalent expression for the constants $\xi_{h^0 f_i f_j}$ in our model,

$$\xi_{h^0 f_i f_j} = \frac{1}{\Lambda} [v_{ij}(m_j - m_i) + a_{ij}(m_i + m_j)], \quad (\text{D2})$$

which leads to the following approximation:

$$x_D \approx \frac{(v_{ct} + a_{ct})^2 (v_{ut} + a_{ut})^2 m_t^4}{3\pi^2 \Lambda^4 m_a^2 \Gamma_D} f_D^2 M_D B_D. \quad (\text{D3})$$

Since we have adopted $v_{tq} = a_{tq} = 1$ throughout the work, we apply the same assumption when evaluating this bound. Plugging in the relevant constants, we express the cutoff scale explaining the mass difference as a function of the flavon mass,

$$\frac{c_{tq}}{\Lambda} \approx \frac{1}{\sqrt{2}} \left(\frac{3\pi^2 m_a^2 \Gamma_D x_D}{f_D^2 M_D B_D} \right)^{1/4}, \quad (\text{D4})$$

where the factor $2^{-1/2}$ ensures both the exclusion limits and the mixing constraint have compatible y-axes in the special case where $v_{ij} = a_{ij}$.

-
- [1] C. D. Froggatt and H. B. Nielsen, *Nucl. Phys.* **B147**, 277 (1979).
 - [2] L. Di Luzio, M. Giannotti, E. Nardi, and L. Visinelli, *Phys. Rep.* **870**, 1 (2020).
 - [3] J. Jaeckel and A. Ringwald, *Annu. Rev. Nucl. Part. Sci.* **60**, 405 (2010).
 - [4] Y. Ema, K. Hamaguchi, T. Moroi, and K. Nakayama, *J. High Energy Phys.* **01** (2017) 096.
 - [5] F. Björkeröth, E. J. Chun, and S. F. King, *J. High Energy Phys.* **08** (2018) 117.
 - [6] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, *J. High Energy Phys.* **09** (2022) 056.
 - [7] Y. Jho, S. Knapen, and D. Redigolo, *J. High Energy Phys.* **10** (2022) 029.
 - [8] A. G. Dias, J. Leite, and D. S. V. Gonçalves, *Phys. Rev. D* **104**, 075014 (2021).
 - [9] L. Di Luzio, F. Mescia, and E. Nardi, *Phys. Rev. D* **96**, 075003 (2017).
 - [10] M. Bauer, M. Neubert, and A. Thamm, *J. High Energy Phys.* **12** (2017) 044.
 - [11] A. Hayrapetyan *et al.* (CMS Collaboration), *Phys. Lett. B* **850**, 138478 (2024).
 - [12] W.-S. Hou and M. Kikuchi, *Europhys. Lett.* **123**, 11001 (2018).
 - [13] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **12** (2023) 081.
 - [14] S. Wahdan (ATLAS Collaboration), *arXiv:2411.14026*.
 - [15] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, *Eur. Phys. J. C* **76**, 235 (2016).

- [16] L. Anzalone, T. Diotalevi, and D. Bonacorsi, *Mach. Learn. Sci. Technol.* **3**, 035017 (2022).
- [17] W. Esmail, A. Hammad, A. Jueid, and S. Moretti, *arXiv*: 2405.16499.
- [18] A. Alves, A. G. Dias, E. da Silva Almeida, and D. S. V. Gonçalves, *Phys. Rev. D* **109**, 055013 (2024).
- [19] C. Cornella, P. Paradisi, and O. Sumensari, *J. High Energy Phys.* **01** (2020) 158.
- [20] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *J. High Energy Phys.* **07** (2014) 079.
- [21] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, *J. High Energy Phys.* **07** (2018) 185; **11** (2021) 85.
- [22] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, *Comput. Phys. Commun.* **183**, 1201 (2012).
- [23] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *Comput. Phys. Commun.* **185**, 2250 (2014).
- [24] C. Bierlich *et al.*, *SciPost Phys. Codebases* **2022**, 8 (2022).
- [25] 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.
- [26] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* **71**, 1554 (2011).
- [27] A. Schwartz and M. Gersaback, Global fit for D^0 - $\bar{D}^0$ mixing (2023), https://hflav-eos.web.cern.ch/hflav-eos/charm/CKM23/results_mix_cpv.html.
- [28] E. Golowich, J. Hewett, S. Pakvasa, and A. A. Petrov, *Phys. Rev. D* **76**, 095009 (2007).
- [29] A. Crivellin, A. Kokulu, and C. Greub, *Phys. Rev. D* **87**, 094031 (2013).
- [30] J. Bergstra, D. Yamins, and D. D. Cox, in *Proceedings of the 30th International Conference on Machine Learning* (PMLR, Georgia, USA, 2013), <http://proceedings.mlr.press/v28/bergstra13.pdf>.
- [31] D. P. Kingma and J. Ba, *arXiv*:1412.6980.
- [32] <https://doi.org/10.5281/zenodo.16745102>
- [33] T. P. Cheng and M. Sher, *Phys. Rev. D* **35**, 3484 (1987).