BSM reach of four-top production at the LHC

Anisha,[1,*] Oliver Atkinson,[1,†] Akanksha Bhardwaj,[1,‡] Christoph Englert,[1,§]
Wrishik Naskar,[1,∥] and Panagiotis Stylianou[2,¶]

[1]*School of Physics & Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom*
[2]*Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*

(Received 2 March 2023; accepted 18 May 2023; published 1 August 2023)

Many scenarios of beyond the Standard Model (BSM) physics give rise to new top-philic interactions that can be probed at proton machines such as the Large Hadron Collider through a variety of production and decay modes. On the one hand, this will enable a detailed determination of the BSM model's parameters when a discovery is made and additional sensitivity in nondominant production modes can be achieved. On the other hand, the naive narrow width approximation in dominant production modes such as gluon fusion might be inadequate for some BSM parameter regions due to interference effects, effectively making less dominant production modes more relevant in such instances. In this work, we consider both these questions in the context of four top quark final states at the LHC. First, we show that the SM potential can be enhanced through the application of targeted graph neural network techniques that exploit data correlations beyond cut-and-count approaches. Second, we show that destructive interference effects that can degrade BSM sensitivity of top-philic states from gluon fusion are largely avoided by turning to four top final states. This achieves considerable exclusion potential, e.g., the two Higgs doublet model. This further motivates four top final states as sensitive tools for BSM discovery in the near future of the LHC.

DOI: 10.1103/PhysRevD.108.035001

I. INTRODUCTION

The absence of evidence for new physics beyond the Standard Model (BSM) in searches at the high-energy frontier of the Large Hadron Collider (LHC) is on the one hand puzzling, as the fundamental origin of the electroweak scale and its stability remains a mystery. On the other hand, the BSM benchmarking programme leading up to the LHC data was, perhaps, overly optimistic. The absence of a BSM signal has highlighted less abundant collider processes as important tools for BSM discrimination, drawing from model-independent and theoretically motivated techniques as well as improved approaches to better exploiting model-(un)specific correlations in searches for new interactions. The resurgence of machine learning techniques applied to the LHC realm is an example of the latter; effective field theory (EFT) in its linear or non-linear formulation is an example of the former.

Particularly interesting processes along these lines are four top quark final states, with the ATLAS and CMS collaborations recently reporting evidence of SM-like production [1–6]. Four top final states are relevant processes when considering EFT deformations of expected SM correlations [7–9], lending considerable sensitivity to large-scale EFT fits, but they also enable a targeted snapshot of electroweak Higgs properties [10]. Recently in [11], it was demonstrated that the large kinematical information that can be exploited in such a busy final state leads to improved sensitivity when final state correlations are tensioned against the SM backgrounds using machine learning approaches. Reference [12] discussed Bayesian techniques for four tops.

For top-philic resonance extensions, four top final states are naively less motivated as gluon fusion provides a large production cross section and top pair resonance searches are experimentally less challenging than the analysis of four top states. Given the generic vulnerability of $pp \to t\bar{t}$ final states to quantum interference with QCD backgrounds [13,14], this might well be a fallacy: when BSM signals are rendered small through interference in the dominant production and decay channels, only those with limited production cross sections remain. Higgs pair production as a relatively clean tool in such circumstances can be statistically limited and four top final states might well provide a more robust alternative route for discovery.

*anisha@glasgow.ac.uk
†o.atkinson.1@research.gla.ac.uk
‡akanksha.bhardwaj@glasgow.ac.uk
§christoph.englert@glasgow.ac.uk
∥w.naskar.1@research.gla.ac.uk
¶panagiotis.stylianou@desy.de

The aim of this work is to clarify both these avenues: building on sensitivity enhancements in discriminating SM four top production, we quantitatively analyse the sensitivity improvement when turning to motivated extensions of the SM. A further emphasis is given to the relevance of interference effects in four top final states.

This paper is organized as follows. In Sec. II, we provide details on event simulation and inclusive fiducial selection for the different (lepton-dominated) partonic collider processes that form the basis of this study: same-sign dilepton in association with b quark and light jets

$$pp \to t\bar{t}t\bar{t} \to \ell^\pm \ell^\pm + \text{jets} + b\,\text{quarks} \quad (\text{2SSDL}), \tag{1.1a}$$

and three lepton production in association with b and light flavor jets

$$pp \to t\bar{t}t\bar{t} \to \ell^\pm \ell^\pm \ell^\mp + \text{jets} + b\,\text{quarks} \quad (\text{3L}). \tag{1.1b}$$

In the following, $\ell = e, \mu$ and we include leptonic decays of τ leptons (see Refs. [1,3] for recent explorations by the ATLAS and CMS experiments). Subsequently, we detail in Sec. II our analysis implementation in terms of a graph neural network (GNN) that provides a particularly motivated approach to exploiting the various particle-level correlations expected in signal and background distributions. Section III is devoted to results and we first discuss the sensitivity yield for SM four top production (Sec. III A), which provides the baseline for the subsequent BSM sensitivity discussion in Secs. III B and III C. Our BSM discussion is split into addressing the relevance of GNN techniques when facing resonant and nonresonant BSM extensions to gauge the potential of the approach for different phenomenological situations. For resonant BSM extensions, we also touch on the relevance of signal-background interference, which is known to be large in resonant ditop final states [15–21], but turns out to be negligible for the dominant four top production modes. A summary and conclusions are provided in Sec. IV.

II. ANALYSIS FRAMEWORK

A. Event simulation and fiducial selection

To prepare our datasets we need to preselect the events appropriate for 2SSDL and 3L final state topologies. Events are generated from simulating proton-proton collisions at $\sqrt{s} = 13$ TeV using MADGRAPH_AMC@NLO [22] with leading order precision. These events are subsequently showered and hadronized using PYTHIA 8.3 [23]. We then reconstruct the final states particles using MADANALYSIS [24] that interfaces FASTJET [25,26].

To identify our inclusive search region we employ the following baseline selection criteria. Light leptons (electrons and muons) are defined from a threshold of $p_T > 10$ GeV within the detector coverage of the electromagnetic calorimeter $|\eta| < 2.5$, where p_T and η denote the transverse momentum and pseudorapidity, respectively. Selected b-tagged jets are identified from a threshold $p_T > 20$ GeV within the tracker $|\eta| < 2.5$. Light flavor jets that fall into the hadronic calorimeter coverage, $|\eta| < 4.5$, are fed to the analysis if they satisfy $p_T > 20$ GeV. Furthermore, we require a significant amount of missing transverse energy of at least 20 GeV. To isolate the mismeasurement of missing transverse energy ($\vec{\not{E}}_T$) from jets we select the events with azimuthal angle difference $|\phi(j) - \phi(\overrightarrow{\not{E}_T})| > 0.2$.

Our analysis for the two signal topologies 2SSDL and 3L closely follows the ATLAS and CMS studies of Ref. [1–6] (see also [27–29] for precision SM analyses informing experimental studies). For the 2SSDL final state, we require at least 2 leptons, at least two b jets and at least one additional jet according to the above criteria. To select the events for the 3L search, we require at least three leptons, at least two b jets without restriction on the light flavor jets. Throughout, we employ a 70% flat b-tagging efficiency, which corresponds to a pessimistic b tagging working point [30] (see also [31]). In turn this removes large contributions from misidentified b jets [31] for the considered p_T range of b jets, hence we neglect effects from misidentified b jets in this study.

Having defined our fiducial regions, we turn to the dominant SM backgrounds contributing to these final states:

(1) The $t\bar{t}W^\pm$ channel with leptonic decays of the tops and W is the most dominant background for the 2SSDL case; however its contributions are subdominant for the 3L channel.
(2) Conversely, background contributions from $t\bar{t}Z$ are significant for the 3L case, while contamination of the 2SSDL signal region from this channel is reduced (yet substantial) compared to others.
(3) $W^+W^-W^\pm + 2b$ jets, with subsequent decays $W \to \ell\nu_\ell$, is the second large background for 2SSDL and remains a considerable background for the 3L selection.
(4) $W^+W^-Z + 2b$ jets with leptonic decays of vector bosons contributes as a background for both the channels, and yields the largest contamination in the 3L case.
(5) $ZW^\pm + 2b$ jets is a subdominant background to both channels when both Z and W bosons decay leptonically. However, its contribution to 2SSDL is negligible and therefore dropped from Table I.

The cross sections after decay and baseline selection are shown in Tables I and II contributing to 2SSDL and 3L channels, respectively.

We prepare datasets for the background class of our neural network setup by combining individual background samples according to their cross section to obtain a realistic composition expected at a fixed luminosity. Representative kinematic distributions are shown in Fig. 1, comparing signal

TABLE I. Cross sections for the different SM processes having the most significant contributions to the 2SSDL background, in accordance with our baseline cuts. The subscript on each particle describes the particle's decay channel ($\ell^\pm$ describes leptonic decay and h refers to a hadronic decay channel). We include the uncertainties from the Monte-Carlo simulation.

Processes	Cross Section (fb)
$pp \to t_{\ell^+}\bar{t}_h W^+_{\ell^+} + t_h \bar{t}_{l_-} W^-_{\ell^-}$	57.67 ± 0.06
$pp \to t_{\ell^+}\bar{t}_h Z_{\ell^+\ell^-} + t_h\bar{t}_{\ell^-} Z_{\ell^+\ell^-} + t_{\ell^+}\bar{t}_{\ell^-} Z_{\ell^+\ell^-}$	10.65 ± 0.01
$pp \to (W^+_{\ell^+} W^-_h W^+_{\ell^+} + W^+_h W^-_{\ell^-} W^-_{\ell^-}) b\bar{b}$	43.29 ± 0.05
$pp \to (W^+_{\ell^+} W^-_h Z_{\ell^+\ell^-} + W^+_h W^-_{\ell^-} Z_{\ell^+\ell^-} + W^+_{\ell^+} W^-_{\ell^-} Z_{\ell^+\ell^-}) b\bar{b}$	12.65 ± 0.02

TABLE II. Summary of background cross sections and Monte-Carlo uncertainties for the different SM processes contributing to the 3L selection, given our baseline cuts. The subscript on each particle denotes the decay channel as in Table I.

Processes	Cross Section (fb)
$pp \to t_{\ell^+}\bar{t}_{\ell^-} W^\pm_{\ell^\pm}$	3.421 ± 0.004
$pp \to t_{\ell^+}\bar{t}_h Z_{\ell^+\ell^-} + t_h\bar{t}_{\ell^-} Z_{\ell^+\ell^-} + t_{\ell^+}\bar{t}_{\ell^-} Z_{\ell^+\ell^-}$	10.65 ± 0.01
$pp \to Z_{\ell^+\ell^-} W^\pm_{\ell^\pm} b\bar{b}$	3.296 ± 0.003
$pp \to W^+_{\ell^+} W^-_{\ell^-} W^\pm_{\ell^\pm} b\bar{b}$	3.614 ± 0.004
$pp \to (W^+_{\ell^+} W^-_h Z_{\ell^+\ell^-} + W^+_h W^-_{\ell^-} Z_{\ell^+\ell^-} + W^+_{\ell^+} W^-_{\ell^-} Z_{\ell^+\ell^-}) b\bar{b}$	12.65 ± 0.02

and background at hadron level. It is evident that H_T and M_{4t}[1] are the most discriminating observables and hardness-related observables can be expected as the dominant drivers of signal vs. background discrimination. The variables H_T and M_{4t} are significantly correlated since both use energy-momentum information of the process, as shown in Fig. 2. The neural network can exploit such correlations (also extending to other observables, see Fig. 3) in a nonrectangular and optimized way, thus enabling efficient signal-background discrimination. We can therefore expect machine-learned observables to carry significant discriminating power in isolating (B)SM four top production.

B. Graph neural network architecture and data representation

Given the complex final state that can be expected from four top topologies, in addition to the process-specific correlations of final state objects, GNNs are ideal candidate architectures to discriminate the signal characteristics from the expected background. Applications of GNNs to particle physics data have a short yet successful history. Their versatility in efficiently discriminating BSM data from the SM expectation beyond traditional observables in a robust way has been highlighted in series of papers, ranging from designing anomaly detection methods [32–34], over jet tagging [31,35] to constraining EFT operator deformations in the BSM top sector fits [36].

Taking inspiration from this evolving success story, we employ a GNN to discriminate the four top signal from the relevant SM background in this work. To implement the graph structure, we use the DEEP GRAPH LIBRARY [37] with the PYTORCH [38] backend and choose an Edge Convolution (EdgeConv) network to classify signal and background (final state particle events are interfaced employing PYLHE [39]).

The GNN can be divided into (i) message passing followed by (ii) node readout. Edge convolution is known to be particularly suited for extracting internode information (edge) from the given low-level node features (i.e., particle-level properties, see below). The message passing function for the edge convolution is defined as

$$\vec{x}_i^{(l+1)} = \frac{1}{|\mathcal{N}(i)|} \sum_{j\in\mathcal{N}(i)} \mathrm{RELU}\Big(\Theta\cdot\Big(\vec{x}_j^{(l)} - \vec{x}_i^{(l)}\Big) + \Phi\cdot\big(\vec{x}_i^{(l)}\big)\Big). \tag{2.1}$$

Here, $\vec{x}_i^{(l)}$ represents the node features of node i in the lth message passing layer, with $l = 0$ denoting the input node features of the graph. The neighborhood set $\mathcal{N}(i)$ consists of all nodes in the graph connected to node i. The linear layers Θ and Φ take the input vectors $(\vec{x}_j^{(l)} - \vec{x}_i^{(l)})$ and $\vec{x}_i^{(l)}$, respectively, and map them onto alternate dimension spaces where one applies an activation function on their vector sum. The dimensionality of the "hidden" feature space is chosen such that performance is optimized, while avoiding overfitting.

For the four top vs. background discrimination task, we use a fully connected bidirectional graph as input to the GNN. Each particle is represented by a node, which is associated with a node feature vector $[p_T, \eta, \phi, E, m, \mathrm{PID}]$ (representing transverse momentum, pseudorapidity, azimuthal angle, energy, mass and particle identification number, respectively). For our analyses, after a fixed number of message passing steps Eq. (2.1), we employ a mean graph readout operation to the node features of the final message passing layer to extract a vector capturing the graph properties of each event. The graph representation obtained this way is then fed into a linear layer with a RELU activation function. The final linear layer maps the result to a two dimensional vector, normalized by the SoftMax function, which corresponds to the probability of an event arising from a signal or a background process.

We use the ADAM optimizer [40] to minimize the cross-entropy loss function with an initial learning rate of 0.001. The network is trained separately with data corresponding

[1]At hadron level we identify M_{4t} as the total invariant mass of the visible final states and H_T as the scalar p_T sum of all visible final states of the system.

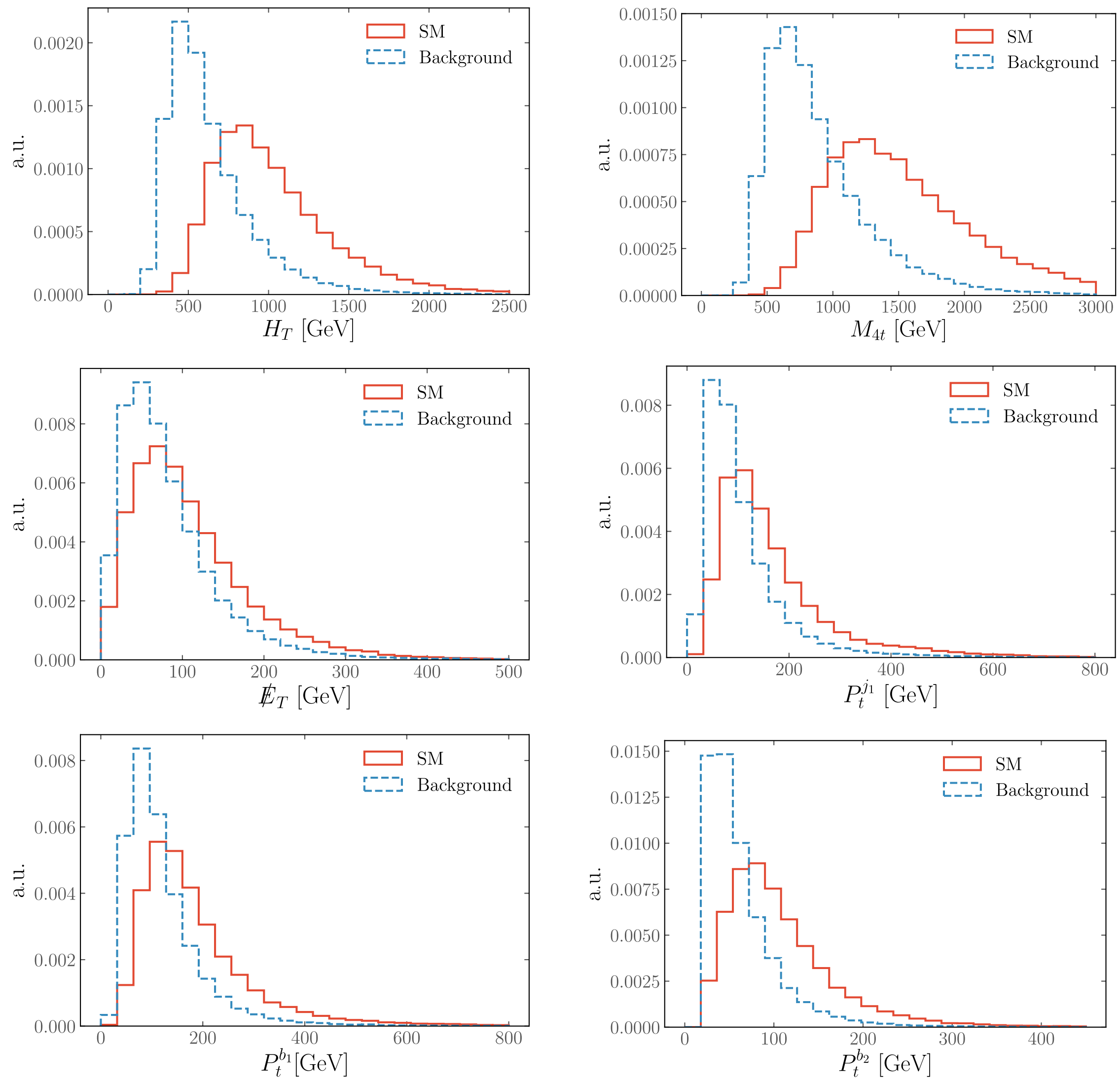

FIG. 1. Normalized distributions of the kinematic variables related to 2SSDL channel for SM four top signal and SM backgrounds contributing to the process at the LHC running at 13 TeV.

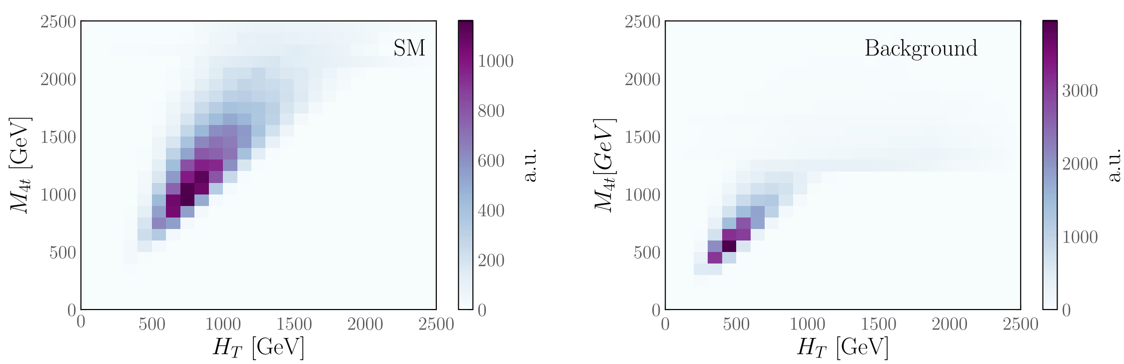

FIG. 2. Normalized two-dimensional distributions of the kinematic variables H_T and M_{4t} for 2SSDL channel for SM four tops signal and SM backgrounds contributing to the process at the LHC running at 13 TeV.

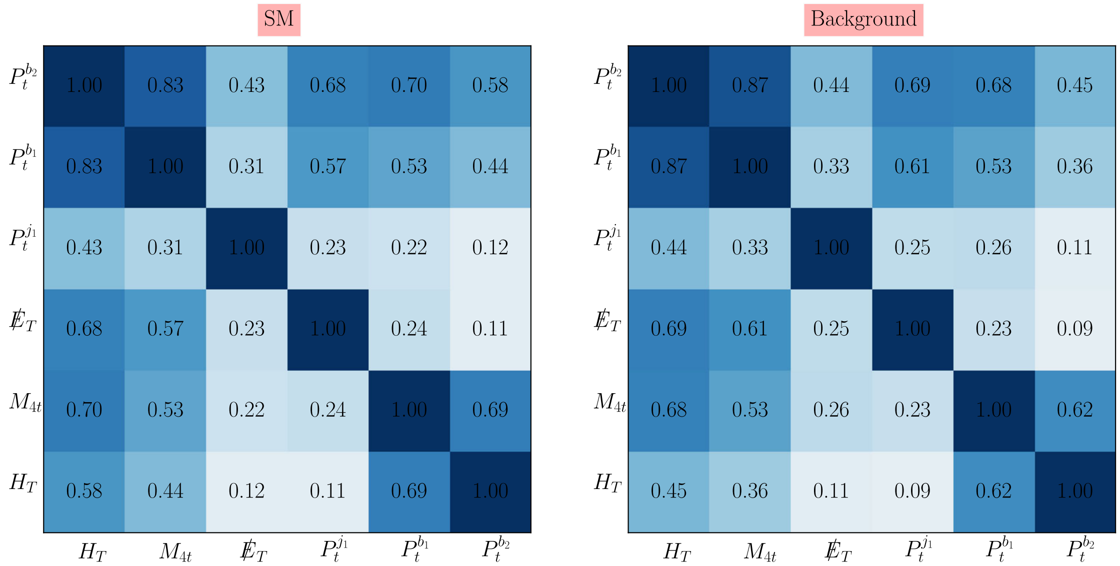

FIG. 3. Correlation plots for 2SSDL channel for SM four tops signal and SM backgrounds for different kinematic variables extending Fig. 2.

to both final states. In each case, the data was split into a 75%–25% training-test ratio. The 75% training data was further split into a 75%–25% training-validation ratio. The learning rate decays with a factor of 0.1 if the loss function has not decreased for three consecutive epochs. We train the models for 100 epochs in minibatches of 100 graphs and an early stopping condition on the loss for ten epochs. By using various combinations of EdgeConv and hidden linear layers, we identify two convolutional layers together with two linear layers and 40 nodes each as a suitable set of hyperparameters, as shown in Fig. 4. We verify that the validation and the training accuracies to check for

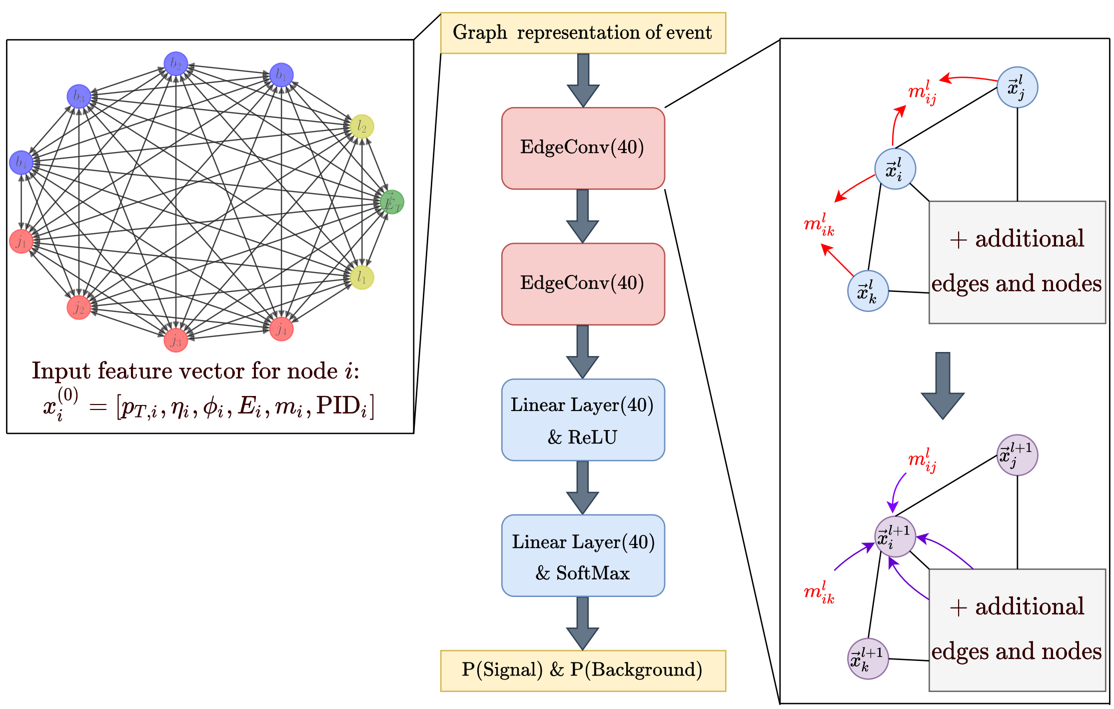

FIG. 4. The architecture of the utilized GNN that maps the events embedded into graphs to the probability that they were sourced by signal or background processes. On the left, an example of an input graph is shown with different colors indicating the different final states that are included in our analyses. On the right the EdgeConv operation is indicated where during the first step (red arrows) the messages $m_{ij}^{(l)} = \mathrm{RELU}(\Theta \cdot (\vec{x}_j^{(l)} - \vec{x}_i^{(l)}) + \Phi \cdot (\vec{x}_i^{(l)}))$ are calculated according to Eq. (2.1) and subsequently in the second step they are used to obtain the updated node features by taking the mean (purple arrows).

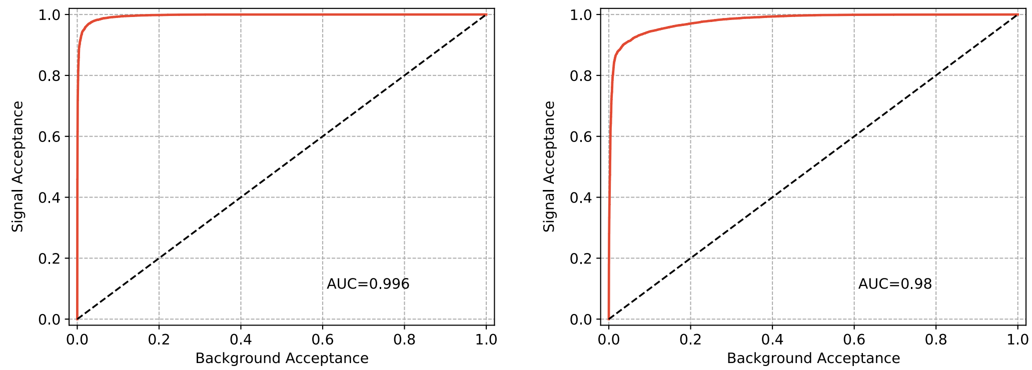

FIG. 5. ROC curves for discriminating the SM four top quark signal in the (a) 2SSDL and (b) 3L final states.

overtraining (we find that deeper networks lead to overtraining by investigating accuracy and loss of the network).

Some comments regarding similar analysis strategies for the four top final state as given above are in order: a variety of ML algorithms have been explored in Ref. [11] to assess their efficiency for the four top process, including the graph-based PARTICLENET [41] utilizing the EdgeConv operation.[2] The authors of Ref. [11] demonstrate that PARTICLENET does indeed provide superior performance in comparison with the alternative algorithms that are considered there. We do find a qualitatively similar behavior.

III. RESULTS

A. Four tops at the LHC: Isolating the SM signal

The Standard Model four top final state has been explored by the ATLAS and CMS collaborations with considerable effort. ATLAS currently observes four tops with a sensitivity of 1.9σ over the background-only hypothesis with a production cross section of 26 fb [3]. In comparison, CMS observes 2.6σ at a cross section of 12.6 fb [1].

To comment on the GNN performance for SM four top production, we first turn to the sensitivity estimates given by the backgrounds described in Sec. II. We first train our network for the 2SSDL samples for SM four top ($t\bar{t}t\bar{t}$) as the signal class and the corresponding backgrounds (listed in Table I) as the background class. A similar procedure is followed for the 3L final state. We demonstrate the network performance using the receiver operating characteristic (ROC) curves of the network which are shown in Fig. 5 for the two different final states under consideration. The areas under the curves (AUCs) of 0.994 for 2SSDL and 0.968 for 3L, imply a very high classifier performance, as expected by the correlations of Figs. 1 and 2 (as well as the findings of Ref. [11]).

[2]PARTICLENET uses the k-nearest-neighbors algorithm for the embedding of particles to graphs according to their separation $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, while we rely on fully-connected graphs.

With the GNN setup optimized, we can cast the sensitivity encoded in the ROC curve into exclusion limits by considering the significance $\sigma = N_s/\sqrt{N_s + N_b}$, where N_s is the number of signal events and N_b is the number of background events at a particular luminosity. The number of signal event N_s is given by $\sigma^s \times \varepsilon_{s(\mathrm{GNN})} \times L$ and number of background event is given is given by N_B is given by $\sigma^b \times \varepsilon_{b(\mathrm{GNN})} \times L$, where $\sigma^{s/b}$ is the cross section of signal and background after the baseline selection, L is the integrated luminosity and $\varepsilon_{s/b(\mathrm{GNN})}$ are the efficiencies obtained at the best working point from the ROC curve such that $\sigma = N_s/\sqrt{N_s + N_b}$ is maximized. We also explore the impact of systematic uncertainties by considering 20% and 50% overall systematic uncertainties in the analysis.

For the 2SSDL final state, the four top production cross section can be measured as 4.4 fb with a systematic uncertainty of 50%, and 4.1 fb with a systematic uncertainty of 20%, at an integrated luminosity of 139 fb^{-1} with a significance of 2.7σ over the background-only hypothesis using the GNN. Although we cannot claim comparability with the realistic experimental environments, these findings indicate that the GNN selection could indeed provide an avenue to further hone the sensitivity in ATLAS and CMS by integrating these techniques into their analysis flow. For the 3L final state, the four top production cross section can be measured as 18.6 fb with a systematic uncertainty of 50%, and 17.3 fb with a systematic uncertainty of 20%, at an integrated luminosity of 139 fb^{-1} with a significance of 2 standard deviations over the background-only hypothesis using GNN, implying a similar potential improvement as for 3L. The performance of the GNN in either channel reveals excellent discriminating power for the busy multitop states at the LHC with prospects of improving cross section measurements significantly.

B. Nonresonant new interactions: Modified Higgs boson interactions

Having demonstrated that the GNN indeed provides a suitable approach for extracting the four top SM signal, we can now turn to the relevance of the GNN for departures of the four top final states from the SM expectation. New physics modifications of four top final states from nonresonant interactions have been considered in a range of phenomenological studies [8,9]. To highlight the relevance of the above approach, we consider the so-called $\hat{H}$ parameter as a motivated EFT-related example; Ref. [10] particularly emphasized the relevance of four top final states for associated searches (bounds have since been provided by the CMS experiment in [1]). Considering this particular nonresonant modification is a well-motivated test bed to motivate a more comprehensive analysis of four top interactions from the perspective of effective field theory [9,36].

The $\hat{H}$ parameter is the analog of the $\hat{W}$, $\hat{Y}$ operators [42] of the gauge sector and can be understood as an oblique correction taking the form

$$\mathcal{L}_{\hat{H}} = \frac{\hat{H}}{m_H^2} |D_\mu D^\mu \Phi|^2, \tag{3.1}$$

where D_μ is the covariant derivative action on the Higgs doublet Φ.[3] The presence of this interaction modifies the propagation of the physical Higgs boson (suppressing the Higgs width which is not relevant as the Higgs boson is probed off-shell in four top production)

$$-i\Delta(p^2, m_H^2) = \frac{1}{p^2 - m_H^2} - \frac{\hat{H}}{m_H^2} \tag{3.2}$$

with associated coupling modifications of the Higgs boson's couplings to massive vectors $V = W^\pm, Z$ and heavy fermions (here the top quark)

$$\frac{g^{\hat{H}}_{VVH(p^2)}}{g^{\rm SM}_{VVH}} = 1 - \hat{H}\left(1 - \frac{p^2}{m_H^2}\right), \qquad \frac{g^{\hat{H}}_{t\bar{t}H}}{g^{\rm SM}_{t\bar{t}H}} = 1 - \hat{H}, \tag{3.3}$$

for canonically normalized fields. The correlation of HVV interactions and H propagator resulting in a cancellation is a consequence of gauge symmetry [10]

$$g^{\hat{H}}_{VVH}\Delta(p^2, m_H^2) = g^{\rm SM}_{VVH}\Delta^{\rm SM}_{\hat{H}=0}(p^2, m_H^2) + \mathcal{O}(\hat{H}^2), \tag{3.4}$$

which highlights multi-top final states as particularly suitable process to put constraints on the interaction of Eq. (3.1) beyond Higgs coupling measurements.

To include the effects of $\hat{H}$ in our study, we modify the HELAS routines [47,48] of our $pp \to t\bar{t}t\bar{t}$ implementation to

[3]It is worthwhile mentioning that the modification of the propagator probes a genuine direction of Higgs effective field theory (HEFT), that can be radiatively sourced in the so-called κ framework [43], see e.g. [44,45]. The associated coupling modifications can then be understood as a probe of the linearity of the Higgs boson's interactions in the SM and its dimension-6 extension [46].

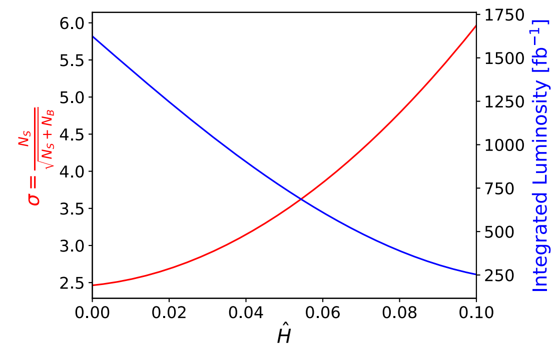

FIG. 6. Significance as a function of $\hat{H}$ at 13 TeV LHC with a 1000 fb^{-1} integrated luminosity. We also present the required luminosity as a 3σ confidence level for different $\hat{H}$ values.

reflect the modifications of propagators, vertices and their cancellation to linear order in the amplitude. As a benchmark for nonresonant new interactions, we change $\hat{H} = 0.1$ as a new physics modification (this value is motivated from the HL-LHC measurement sensitivity provided in Ref. [10]). We select a SM-rich sample by further applying a cut on H_T greater than 900 GeV to enable a coupling measurement that is more geared to SM modifications, see Fig. 1. We then train a GNN network to discriminate between SM interaction with $\hat{H} = 0$ from the nonresonant new interaction with $\hat{H} = 0.1$. The kinematical differences are not dramatic (which is highlighted by the $\sim$10% sensitivity in the first place), and therefore further motivate the inclusion of as much correlated information as possible achievable through GNN applications.

We find an AUC of 60%, which is testimony to the difficulty of extracting electroweak properties from QCD-busy final states. The observed sensitivity can be used to set limits for in the parameter range $\hat{H} \in [0.0, 0.10]$ shown in Fig. 6. The blue curve shows the required luminosity to obtain the 3σ bound on that particular $\hat{H}$ value, whereas the red curve shows the significance obtained at a fixed luminosity $L = 1000$ fb^{-1}. Our GNN analysis again gives rise to an expected sensitivity improvement beyond the estimates of cut-and-count analyses [1,10]. While the realistic analysis in the LHC environment is likely more limited in sensitivity, the addition of the GNN techniques described above could significantly improve the sensitivity to $\hat{H}$ considerably, which also indicates an improved sensitivity to generic EFT deformations.

C. Resonant new interactions: Interference effects and 2HDM reach

Four top final states in the context of resonant extensions of the SM have been studied in a range of analyses [49–57], typically relying on traditional collider observables alongside cut-and-count strategies. While new resonant structures lend themselves to such approaches, we can also

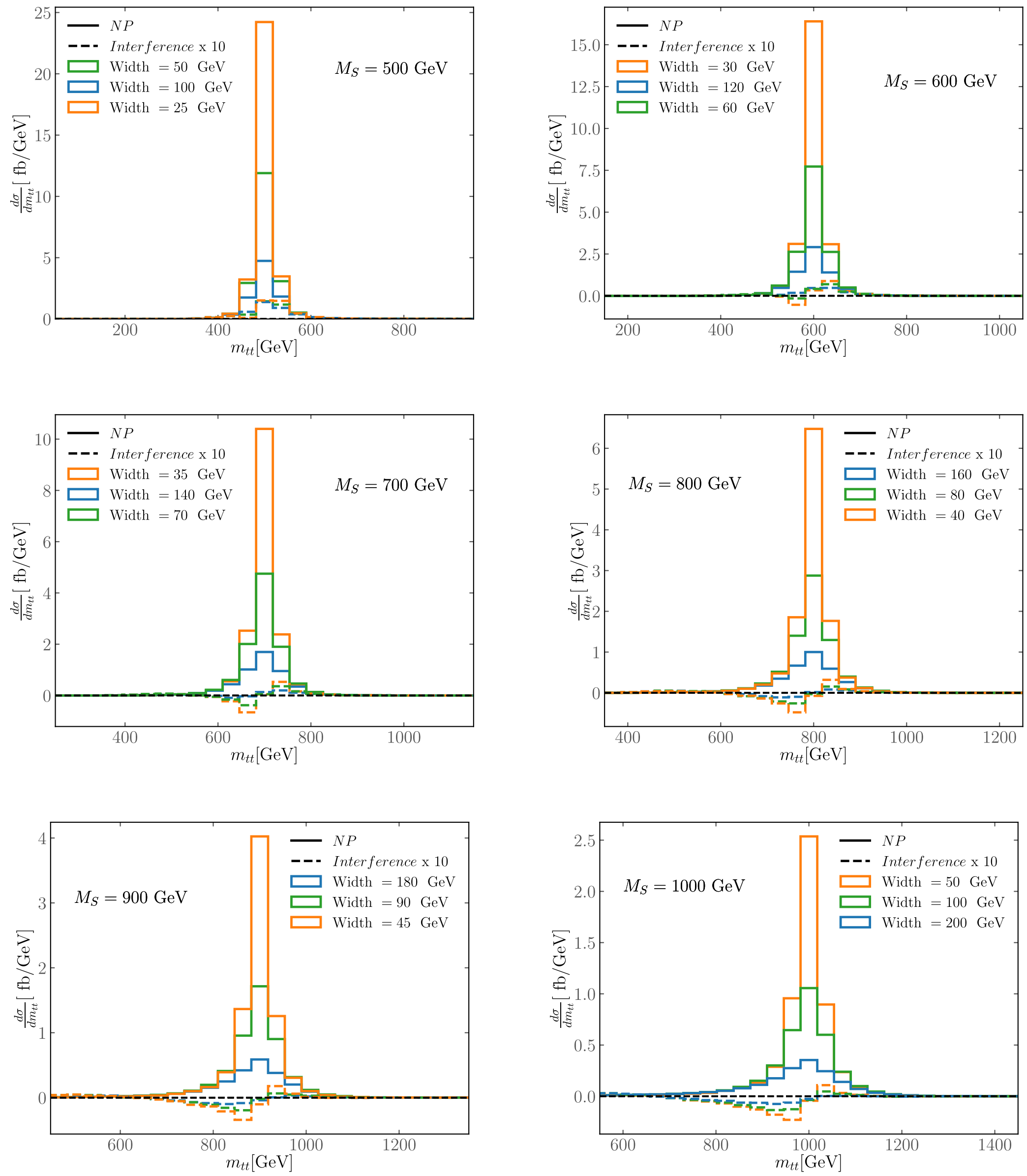

FIG. 7. The invariant mass distribution of m_{tt} system closest to the mass of the new scalar S as detailed in the text.

expect a significant sensitivity enhancement when turning to ML approaches (see also [58–63]).

Before we comment on the GNN performance, however, we will perform a qualitative analysis of the interference effects that are known to be large in direct gluon fusion production $gg \to t\bar{t}$ of extra scalar resonances [13,14]. Such interference effects can severely limit the sensitivity estimates and are not straightforward to include in ML-based selections due to negative weights. To this end we introduce a simplified scalar resonance S with Higgs-like couplings to the top quark[4]

$$\mathcal{L}_{\rm simp} = \frac{1}{2}(\partial S)^2 - \frac{M_S^2}{2}S^2 - \frac{m_t}{v}[\xi_S \bar{t}_L t_R S + \text{H.c.}], \quad (3.5)$$

keeping the width of S as a free parameter in a scan to qualitatively assess S resonance-distortion (ξ_S plays the role of the Higgs coupling modifier for $S = H$). To study the interference effects in the four top final states, we compare the resonance structure stemming for the new-physics-only (resonance) interactions

$$d\sigma^{\rm new} \sim |\mathcal{M}_{\rm res}|^2 \text{dLIPS} \quad (3.6)$$

with interference-only contributions

$$d\sigma^{\rm inf} \sim 2\text{Re}(\mathcal{M}_{\rm bkg}\mathcal{M}^*_{\rm res})\text{dLIPS}, \quad (3.7)$$

[4]We model the phenomenological Lagrangian with FEYNRULES [64,65] and export the relevant Feynman Rules in the UFO [66] format. We use the same toolchain as before with the addition of MADSPIN [67,68].

TABLE III. Significances for different masses M_S for a coupling choice $\xi_S = 1$ and $\Gamma_S/M_S = 0.1$, Eq. (3.5). The luminosity is taken to be 3000 fb^{-1} for 13 TeV LHC collisions.

M_S (GeV)	significance for 2SSDL	significance for 3L
500	10.6σ	1.4σ
600	9.9σ	1.3σ
700	8.4σ	1.1σ
800	6.9σ	0.9σ
900	5.5σ	0.7σ
1000	4.3σ	0.5σ

where “bkg” refers to any nonresonant amplitude contribution (e.g., the continuum QCD background). A good qualitative understanding of the signal distribution distortion can be obtained by isolating the peaked $m_{t\bar{t}}$ distribution: we construct all possible combinations of $m_{t\bar{t}}$ and choose the invariant mass which is closest to the candidate mass M_S. The resulting distributions are shown in Fig. 7 for a range of mass $M_S = [500, 600, 700, 800, 1000]$ GeV and width choices $\Gamma_S/M_S = 5\%, 10\%, 20\%$ for $\xi_S = 1$.

From the plots shown in Fig. 7, we can infer that the interference effects that distort the mass peak are comparably small. Closer to the decoupling limit $\xi_S \to 0$ we will encounter more sensitivity-limiting distortion, however this will happen at quickly vanishing BSM resonance cross sections so that a (ML-assisted) bump hunt will not provide any sensitivity when considering the backgrounds. In the case of imaginary phases of ξ_S, i.e., CP-odd coupling structures, the interference effects will be qualitatively similar see Refs. [15] (they can also vanish in processes when interference with CP-even backgrounds removes such effects [60]).

Having identified interference effects as largely irrelevant, we return to the GNN analysis of the BSM resonance structures. We train our network for the same two final states, this time with the added new interactions (i.e., the newly added CP-even scalar with $\xi_S = 1$) for the different benchmark points listed in Table III. The ROC curves for one such benchmark point ($M_S = 600$ GeV) for both the final states and their corresponding AUCs are shown in Fig. 8. The corresponding significance for each benchmark point for the different decay channels is also listed in Table III for $\xi_S = 1$ and $\Gamma_S/M_S = 0.1$. We interpret these limits as bounds on the coupling modifier ξ_S as shown in Fig. 9.

Although we can indeed put constraints $\xi_S < 1$ (predominantly through the 2SSDL channel), the resulting bounds are relatively weak. For instance in singlet mixing scenarios, the parameter range that this analysis is sensitive to is already highly constrained from Higgs signal strength measurements. Turning to phenomenologically richer scenarios we further examine the 2HDM of type II, which features new neutral CP odd and CP even states. In this model, the Higgs sector consists of two distinct complex doublets, both with a nonzero vacuum expectation value (vev) (reviews can be found in Refs. [69,70]). One of these doublets couples to the up-type quarks and the other to down-type quarks and leptons in the type II model, thereby leading to fermion mass terms. After electroweak symmetry breaking, one is left with 5 Higgs bosons; 4 of which are CP even; h, H and $H^\pm$, and a CP odd particle, A. The lighter of the two neutral CP even states is identified as the observed Higgs boson. The phenomenology of this model then depends on the masses of the new Higgs bosons, the ratio of the two vevs, defined as $\tan\beta = v_2/v_1$ and $\cos(\beta-\alpha)$, where α is the mixing angle between the two neutral CP even states. The simplified Lagrangian of Eq. (3.5) can be mapped onto the 2HDM type II Yukawa sector

$$\mathcal{L}_{\text{2HDM}} \supset -\frac{m_t}{v}(\xi_h \bar{t}th + \xi_H \bar{t}tH - i\xi_A \bar{t}\gamma^5 tA), \tag{3.8}$$

where we have suppress charged Higgs contributions as well as nontop interactions. For the 2HDM type the coupling modifiers related to the top quark are

$$\xi_h = \sin(\beta-\alpha) + \cos(\beta-\alpha)\cot\beta, \tag{3.9a}$$

$$\xi_H = \cos(\beta-\alpha) - \sin(\beta-\alpha)\cot\beta, \tag{3.9b}$$

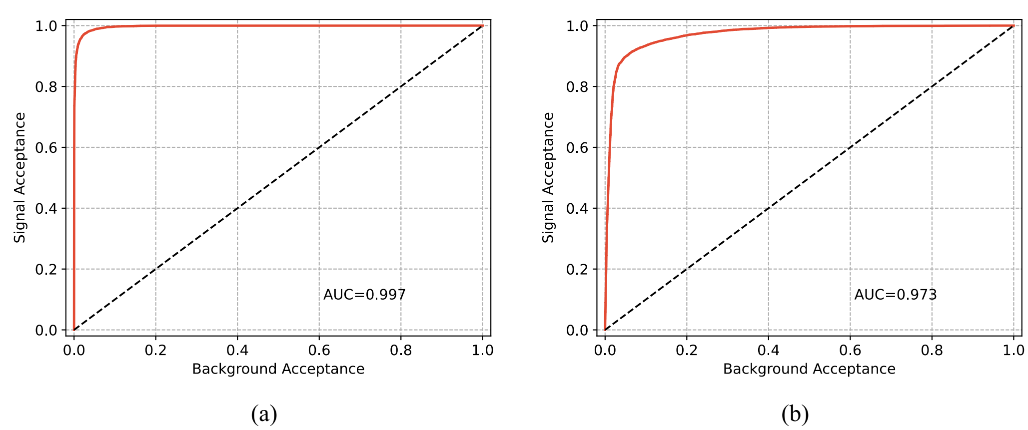

FIG. 8. ROC curves for training a four top quark signal for a BSM CP-even scalar with $M_S = 600$ GeV for (a) 2SSDL and (b) 3L final states.

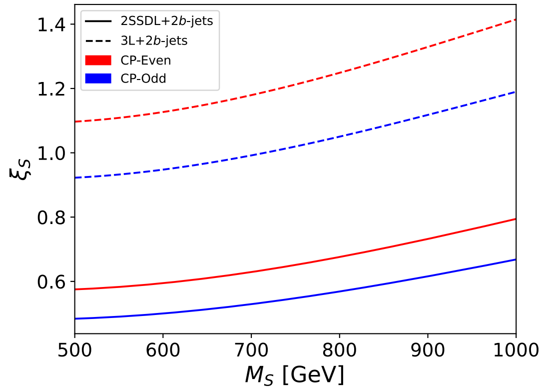

FIG. 9. 2σ confidence level limits on scalar mass with respect to CP-even and CP-odd couplings for 3000 fb^{-1} for 13 TeV LHC collisions.

$$\xi_A = \cot\beta. \tag{3.9c}$$

The 2HDM is a particularly well motivated BSM model, as it is, in principle, capable of resolving a number of tensions between current experimental values and SM predictions, as well as opening avenues to satisfy the Sakharov criteria for required baryogenesis [71]. As such, this model is well studied in the literature [72–80]. Here we use this model to explore the validity of the benchmark points examined above. In order to satisfy the theoretical considerations the masses of the new particles are set to be entirely degenerate. Similarly, to fall in line with the signal strength data of the SM Higgs we set $\cos(\beta-\alpha)=0$; the alignment limit which recovers exactly the SM phenomenology of h.

To examine the mass range of the above benchmark points in the face of the collider searches for additional Higgs bosons we make use of the 2HDECAY [81–86] and HIGGSBOUNDS [87–93] packages. Additionally, we extrapolate the LHC search data to 3000 fb^{-1} at 13 TeV. Performing a scan over 20000 random points, the results of this analysis are shown in Fig. 10, with points in blue allowed by current and extrapolated data, while orange points are allowed only by the current data. The improvement in sensitivity is lead by searches for the decays $H^+ \to t\bar{b}$ and $H^0 \to \tau^+\tau^-$ [94,95]. This shows that a large swathe of the parameter space that we take for benchmark points in the above analysis is possible within the framework of the 2HDM of type II.

Given that four top final states are far less abundant than other processes, they can be expected to exhibit a reduced sensitivity to new states (this is already visible from Fig. 9). Nonetheless, repeating the above analysis also including the CP odd state, we find that complementary sensitivity can be achieved for the 2HDM scenario with the GNN-assisted resonance search detailed above: By overlaying the constraints found from 2SSDL final states (shaded area) we see that there is a nontrivial parameter range of the 2HDM type II's alignment limit that can be accessed four top final states at 2σ confidence and 3000 fb^{-1}. This example contains only the more sensitive channel of the ones discussed in this work and additional final states are likely to further improve this result.

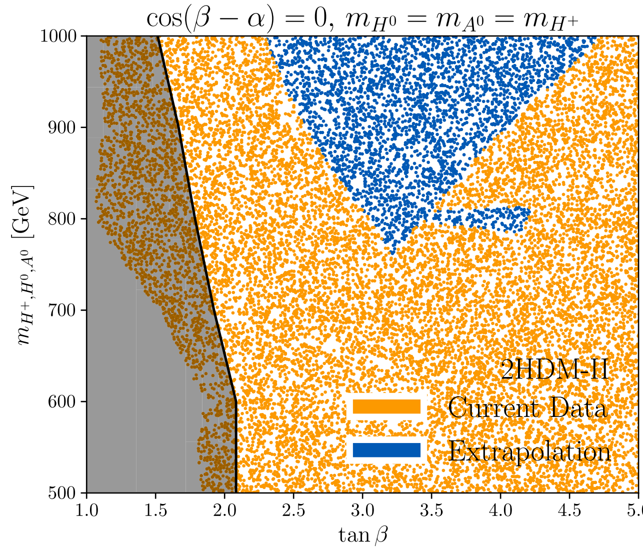

FIG. 10. Scan of the parameter space of the type II 2HDM in the face of current collider data, and LHC search data extrapolated to 3000 fb^{-1} at 13 TeV, with a 2σ confidence level exclusion contour from 2SSDL final states. Points in blue are allowed by current and expected future datasets, orange points by current data only. The shaded area gives the sensitivity of the 2SSDL resonance search detailed in the text at 3000 fb^{-1}. Interference effects from $gg \to t\bar{t}$ sensitivity estimates are not included.

IV. SUMMARY AND CONCLUSIONS

The growing sensitivity to four top final states at the Large Hadron Collider [1–6] bears great potential for the discovery of new physics beyond the Standard Model. While the complexity of the final states requires a dedicated search strategy compared to lower multiplicity processes, the plethora of kinematic information that can be accessed in these processes to highlight departures from the SM can be formidably exploited by means of adapted search strategies, e.g. building on machine learning approaches. These do not only provide support in establishing sensitivity to SM four top production as shown in Sec. III A (see also [11]), but they become particularly relevant when four top final states will be scrutinized from a BSM perspective in the future. We have demonstrated that significant sensitivity to well-motivated representative nonresonant physics can be achieved (Sec. III B), but also that sensitivity to resonances is gained by turning to graph neural networks (Sec. III C). The GNN exploits the hierarchical graph structure of particle physics events that is particularly accentuated when dealing with new resonances. In this case, the four top final states can also fill a potential sensitivity gap in s-channel gluon fusion production of top-philic states $gg \to t\bar{t}$ that results from destructive signal-background interference, distorting or removing the resonance structure. The dominant partonic channels of four top final states show no such behavior and therefore remain robust processes even when interference removes sensitivity in top pair gluon fusion production. In particular, this can add complementary sensitivity to BSM searches for (but not limited to) two Higgs doublet models as we have shown in a representative parameter scan for the 2HDM type II.

ACKNOWLEDGMENTS

A. and this work is funded by the Leverhulme Trust under Grant No. RPG-2021-031. O. A. is supported by the UK Science and Technology Facilities Council (STFC) under Grant No. ST/V506692/1. A. B. and C. E. are supported by the STFC under Grant No. ST/T000945/1. C. E. acknowledges further support by the IPPP Associateship Scheme. W. N. is funded by a University of Glasgow College of Science and Engineering Scholarship. P. S. is supported by the Deutsche Forschungsgemeinschaft under Germany's Excellence strategy EXC2121 "Quantum Universe"—390833306. This work has been partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—No. 491245950.

[1] A. M. Sirunyan *et al.* (CMS Collaboration), Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV, Eur. Phys. J. C **80**, 75 (2020).
[2] G. Aad *et al.* (ATLAS Collaboration), Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Eur. Phys. J. C **80**, 1085 (2020).
[3] G. Aad *et al.* (ATLAS Collaboration), Measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 11 (2021) 118.
[4] A. M. Sirunyan *et al.* (CMS Collaboration), Evidence for the simultaneous production of four top quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV, Technical Report No. CMS-TOP-21-005, CERN-EP-2023-014, 2022.
[5] A. M. Sirunyan *et al.* (CMS Collaboration), Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 11 (2019) 082.
[6] M. Aaboud *et al.* (ATLAS Collaboration), Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Rev. D **99**, 052009 (2019).
[7] C. Zhang, Constraining $qqtt$ operators from four-top production: A case for enhanced EFT sensitivity, Chin. Phys. C **42**, 023104 (2018).
[8] G. Banelli, E. Salvioni, J. Serra, T. Theil, and A. Weiler, The present and future of four top operators, J. High Energy Phys. 02 (2021) 043.
[9] R. Aoude, H. El Faham, F. Maltoni, and E. Vryonidou, Complete SMEFT predictions for four top quark production at hadron colliders, J. High Energy Phys. 10 (2022) 163.
[10] C. Englert, G. F. Giudice, A. Greljo, and M. McCullough, The $\hat{H}$-parameter: An oblique Higgs view, J. High Energy Phys. 09 (2019) 041.
[11] L. Builtjes, S. Caron, P. Moskvitina, C. Nellist, R. R. de Austri, R. Verheyen *et al.*, Climbing four tops with graph networks, transformers and pairwise features, arXiv:2211.05143.
[12] E. Alvarez, B. M. Dillon, D. A. Faroughy, J. F. Kamenik, F. Lamagna, and M. Szewc, Bayesian probabilistic modeling for four-top production at the LHC, Phys. Rev. D **105**, 092001 (2022).
[13] K. J. F. Gaemers and F. Hoogeveen, Higgs production and decay into heavy flavors with the gluon fusion mechanism, Phys. Lett. **146B**, 347 (1984).
[14] D. Dicus, A. Stange, and S. Willenbrock, Higgs decay to top quarks at hadron colliders, Phys. Lett. B **333**, 126 (1994).
[15] S. Jung, J. Song, and Y. W. Yoon, Dip or nothingness of a Higgs resonance from the interference with a complex phase, Phys. Rev. D **92**, 055009 (2015).
[16] R. Frederix and F. Maltoni, Top pair invariant mass distribution: A window on new physics, J. High Energy Phys. 01 (2009) 047.
[17] M. Carena and Z. Liu, Challenges and opportunities for heavy scalar searches in the $t\bar{t}$ channel at the LHC, J. High Energy Phys. 11 (2016) 159.
[18] B. Hespel, F. Maltoni, and E. Vryonidou, Signal background interference effects in heavy scalar production and decay to a top-anti-top pair, J. High Energy Phys. 10 (2016) 016.
[19] A. Djouadi, J. Ellis, A. Popov, and J. Quevillon, Interference effects in $t\bar{t}$ production at the LHC as a window on new physics, J. High Energy Phys. 03 (2019) 119.
[20] M. Aaboud *et al.* (ATLAS Collaboration), Search for Heavy Higgs Bosons A/H Decaying to a Top Quark Pair in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector, Phys. Rev. Lett. **119**, 191803 (2017).
[21] A. M. Sirunyan *et al.* (CMS Collaboration), Search for heavy Higgs bosons decaying to a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 04 (2020) 171.
[22] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, J. High Energy Phys. 07 (2014) 079.

[23] C. Bierlich *et al.*, A comprehensive guide to the physics and usage of PYTHIA8.3, SciPost Phys. 8 (2022).

[24] E. Conte, B. Fuks, and G. Serret, MadAnalysis 5, a user-friendly framework for collider phenomenology, Comput. Phys. Commun. **184**, 222 (2013).

[25] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, Eur. Phys. J. C **72**, 1896 (2012).

[26] M. Cacciari and G. P. Salam, Dispelling the N^3 myth for the k_t jet-finder, Phys. Lett. B **641**, 57 (2006).

[27] G. Bevilacqua and M. Worek, Constraining BSM physics at the LHC: Four top final states with NLO accuracy in perturbative QCD, J. High Energy Phys. 07 (2012) 111.

[28] F. Maltoni, D. Pagani, and I. Tsinikos, Associated production of a top-quark pair with vector bosons at NLO in QCD: Impact on $t\bar{t}$H searches at the LHC, J. High Energy Phys. 02 (2016) 113.

[29] R. Frederix, D. Pagani, and M. Zaro, Large NLO corrections in $t\bar{t}W^{\pm}$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions, J. High Energy Phys. 02 (2018) 031.

[30] G. Aad *et al.* (ATLAS Collaboration), Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector, Report No. ATLAS-CONF-2012-043, 2012.

[31] E. Bols, J. Kieseler, M. Verzetti, M. Stoye, and A. Stakia, Jet flavour classification using DeepJet, J. Instrum. **15**, P12012 (2020).

[32] A. Blance and M. Spannowsky, Unsupervised event classification with graphs on classical and photonic quantum computers, J. High Energy Phys. 08 (2020) 170.

[33] O. Atkinson, A. Bhardwaj, C. Englert, P. Konar, V. S. Ngairangbam, and M. Spannowsky, IRC-safe graph autoencoder for unsupervised anomaly detection, Front. Artif. Intell. **5**, 943135 (2022).

[34] O. Atkinson, A. Bhardwaj, C. Englert, V. S. Ngairangbam, and M. Spannowsky, Anomaly detection with convolutional graph neural networks, J. High Energy Phys. 08 (2021) 080.

[35] F. A. Dreyer and H. Qu, Jet tagging in the Lund plane with graph networks, J. High Energy Phys. 03 (2021) 052.

[36] O. Atkinson, A. Bhardwaj, S. Brown, C. Englert, D. J. Miller, and P. Stylianou, Improved constraints on effective top quark interactions using edge convolution networks, J. High Energy Phys. 04 (2022) 137.

[37] M. Wang, D. Zheng, Z. Ye, Q. Gan, M. Li, X. Song *et al.*, Deep graph library: A graph-centric, highly-performant package for graph neural networks, arXiv:1909.01315.

[38] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan *et al.*, Pytorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems 32* (Curran Associates, Inc., 2019), pp. 8024–8035.

[39] L. Heinrich, M. Feickert, and E. Rodrigues, pylhe: v0.5.0, 10.5281/zenodo.1217031 (2023).

[40] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.

[41] H. Qu and L. Gouskos, ParticleNet: Jet tagging via particle clouds, Phys. Rev. D **101**, 056019 (2020).

[42] R. Barbieri, A. Pomarol, R. Rattazzi, and A. Strumia, Electroweak symmetry breaking after LEP-1 and LEP-2, Nucl. Phys. **B703**, 127 (2004).

[43] S. Dittmaier *et al.* (LHC Higgs Cross Section Working Group Collaboration), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables* (CERN, Geneva, 2011).

[44] M. J. Herrero and R. A. Morales, One-loop renormalization of vector boson scattering with the electroweak chiral Lagrangian in covariant gauges, Phys. Rev. D **104**, 075013 (2021).

[45] Anisha, O. Atkinson, A. Bhardwaj, C. Englert, and P. Stylianou, Quartic Gauge-Higgs couplings: Constraints and future directions, J. High Energy Phys. 10 (2022) 172.

[46] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, Dimension-six terms in the standard model Lagrangian, J. High Energy Phys. 10 (2010) 085.

[47] H. Murayama, I. Watanabe, and K. Hagiwara, HELAS: HELicity amplitude subroutines for Feynman diagram evaluations, Report No. KEK-91-11, 1992.

[48] P. de Aquino, W. Link, F. Maltoni, O. Mattelaer, and T. Stelzer, ALOHA: Automatic libraries of helicity amplitudes for Feynman diagram computations, Comput. Phys. Commun. **183**, 2254 (2012).

[49] G. Aad *et al.* (ATLAS Collaboration), Search for $t\bar{t}H/A \to t\bar{t}t\bar{t}$ production in the multilepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Report No. CERN-EP-2022-170, 2022.

[50] F. Blekman, F. Déliot, V. Dutta, and E. Usai, Four-top quark physics at the LHC, Universe **8**, 638 (2022).

[51] E. Alvarez, A. Juste, and R. M. S. Seoane, Four-top as probe of light top-philic new physics, J. High Energy Phys. 12 (2019) 080.

[52] E. Alvarez, D. A. Faroughy, J. F. Kamenik, R. Morales, and A. Szynkman, Four tops for LHC, Nucl. Phys. **B915**, 19 (2017).

[53] S. Kanemura, H. Yokoya, and Y.-J. Zheng, Searches for additional Higgs bosons in multi-top-quarks events at the LHC and the international linear collider, Nucl. Phys. **B898**, 286 (2015).

[54] L. Darmé, B. Fuks, and F. Maltoni, Top-philic heavy resonances in four-top final states and their EFT interpretation, J. High Energy Phys. 09 (2021) 143.

[55] L. Darmé, B. Fuks, and M. Goodsell, Cornering sgluons with four-top-quark events, Phys. Lett. B **784**, 223 (2018).

[56] P. S. Bhupal Dev and A. Pilaftsis, Maximally symmetric two Higgs doublet model with natural standard model alignment, J. High Energy Phys. 12 (2014) 024.

[57] Q.-H. Cao, J.-N. Fu, Y. Liu, X.-H. Wang, and R. Zhang, Probing top-philic new physics via four-top-quark production, Chin. Phys. C **45**, 093107 (2021).

[58] A. Butter, G. Kasieczka, T. Plehn, and M. Russell, Deep-learned top tagging with a Lorentz layer, SciPost Phys. **5**, 028 (2018).

[59] J. H. Collins, K. Howe, and B. Nachman, Extending the search for new resonances with machine learning, Phys. Rev. D **99**, 014038 (2019).

[60] O. Atkinson, C. Englert, and P. Stylianou, On interference effects in top-philic decay chains, Phys. Lett. B **821**, 136618 (2021).

[61] F. F. Freitas, J. a. Gonçalves, A. P. Morais, and R. Pasechnik, Phenomenology of vector-like leptons with deep learning at the large hadron collider, J. High Energy Phys. 01 (2021) 076.

[62] J. Hollingsworth and D. Whiteson, Resonance searches with machine learned likelihood ratios, arXiv:2002.04699.
[63] O. Kitouni, B. Nachman, C. Weisser, and M. Williams, Enhancing searches for resonances with machine learning and moment decomposition, J. High Energy Phys. 04 (2021) 070.
[64] N. D. Christensen and C. Duhr, FeynRules—Feynman rules made easy, Comput. Phys. Commun. **180**, 1614 (2009).
[65] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, FeynRules 2.0—A complete toolbox for tree-level phenomenology, Comput. Phys. Commun. **185**, 2250 (2014).
[66] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, UFO—The universal FeynRules output, Comput. Phys. Commun. **183**, 1201 (2012).
[67] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations, J. High Energy Phys. 04 (2007) 081.
[68] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations, J. High Energy Phys. 03 (2013) 015.
[69] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, Theory and phenomenology of two-Higgs-doublet models, Phys. Rep. **516**, 1 (2012).
[70] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, *The Higgs Hunter's Guide* (2000), Vol. 80.
[71] A. D. Sakharov, Violation of *CP* Invariance, C asymmetry, and baryon asymmetry of the universe, Pis'ma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).
[72] P. Basler, M. Krause, M. Muhlleitner, J. Wittbrodt, and A. Wlotzka, Strong first order electroweak phase transition in the *CP*-conserving 2HDM revisited, J. High Energy Phys. 02 (2017) 121.
[73] F. Kling, J. M. No, and S. Su, Anatomy of exotic Higgs decays in 2HDM, J. High Energy Phys. 09 (2016) 093.
[74] P. Arnan, D. Bečirević, F. Mescia, and O. Sumensari, Two Higgs doublet models and $b \to s$ exclusive decays, Eur. Phys. J. C **77**, 796 (2017).
[75] A. Crivellin, D. Müller, and C. Wiegand, $b \to s\ell^+\ell^-$ transitions in two-Higgs-doublet models, J. High Energy Phys. 06 (2019) 119.
[76] X.-F. Han and H.-X. Wang, Revisiting wrong sign Yukawa coupling of type II two-Higgs-doublet model in light of recent LHC data, Chin. Phys. C **44**, 073101 (2020).
[77] O. Atkinson, M. Black, A. Lenz, A. Rusov, and J. Wynne, Cornering the two Higgs doublet model type II, J. High Energy Phys. 04 (2022) 172.
[78] O. Atkinson, M. Black, C. Englert, A. Lenz, A. Rusov, and J. Wynne, The flavourful present and future of 2HDMs at the collider energy frontier, J. High Energy Phys. 11 (2022) 139.
[79] T. Biekötter, S. Heinemeyer, J. M. No, M. O. Olea-Romacho, and G. Weiglein, The trap in the early Universe: Impact on the interplay between gravitational waves and LHC physics in the 2HDM, J. Cosmol. Astropart. Phys. 03 (2023) 031.
[80] S. Dawson, D. Fontes, S. Homiller, and M. Sullivan, Role of dimension-eight operators in an EFT for the 2HDM, Phys. Rev. D **106**, 055012 (2022).
[81] M. Krause, M. Mühlleitner, and M. Spira, 2HDECAY—A program for the calculation of electroweak one-loop corrections to Higgs decays in the two-Higgs-doublet model including state-of-the-art QCD corrections, Comput. Phys. Commun. **246**, 106852 (2020).
[82] A. Djouadi, J. Kalinowski, and M. Spira, HDECAY: A program for Higgs boson decays in the standard model and its supersymmetric extension, Comput. Phys. Commun. **108**, 56 (1998).
[83] A. Djouadi, J. Kalinowski, M. Muehlleitner, and M. Spira, HDECAY: Twenty$_{++}$ year after, Comput. Phys. Commun. **238**, 214 (2019).
[84] M. Krause, M. Muhlleitner, R. Santos, and H. Ziesche, Higgs-to-Higgs boson decays in a 2HDM at next-to-leading order, Phys. Rev. D **95**, 075019 (2017).
[85] A. Denner, S. Dittmaier, and J.-N. Lang, Renormalization of mixing angles, J. High Energy Phys. 11 (2018) 104.
[86] T. Hahn and M. Perez-Victoria, Automatized one loop calculations in four-dimensions and D-dimensions, Comput. Phys. Commun. **118**, 153 (1999).
[87] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein, and K. E. Williams, HiggsBounds: Confronting arbitrary Higgs sectors with exclusion bounds from LEP and the Tevatron, Comput. Phys. Commun. **181**, 138 (2010).
[88] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein, and K. E. Williams, HiggsBounds 2.0.0: Confronting neutral and charged Higgs sector predictions with exclusion bounds from LEP and the Tevatron, Comput. Phys. Commun. **182**, 2605 (2011).
[89] P. Bechtle, O. Brein, S. Heinemeyer, O. Stal, T. Stefaniak, G. Weiglein *et al.*, Recent developments in HiggsBounds and a preview of HiggsSignals, Proc. Sci., CHARGED2012 (**2012**) 024 [arXiv:1301.2345].
[90] P. Bechtle, O. Brein, S. Heinemeyer, O. Stål, T. Stefaniak, G. Weiglein *et al.*, HiggsBounds – 4: Improved tests of extended Higgs sectors against exclusion bounds from LEP, the tevatron and the LHC, Eur. Phys. J. C **74**, 2693 (2014).
[91] P. Bechtle, S. Heinemeyer, O. Stal, T. Stefaniak, and G. Weiglein, Applying exclusion likelihoods from LHC searches to extended Higgs sectors, Eur. Phys. J. C **75**, 421 (2015).
[92] P. Bechtle, D. Dercks, S. Heinemeyer, T. Klingl, T. Stefaniak, G. Weiglein, and J. Wittbrodt, HiggsBounds-5: Testing Higgs sectors in the LHC 13 TeV era, Eur. Phys. J. C **80**, 1211 (2020).
[93] H. Bahl, V. M. Lozano, T. Stefaniak, and J. Wittbrodt, Testing exotic scalars with HiggsBounds, Eur. Phys. J. C **82**, 584 (2022).
[94] A. M. Sirunyan *et al.* (CMS Collaboration), Search for charged Higgs bosons decaying into a top and a bottom quark in the all-jet final state of pp collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 07 (2020) 126.
[95] G. Aad *et al.* (ATLAS Collaboration), Search for Heavy Higgs Bosons Decaying into Two Tau Leptons with the ATLAS Detector Using pp Collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **125**, 051801 (2020).