Hunting for top partner with a new signature at the LHC

Daohan Wang,[1,2,3] Lei Wu,[4,*] and Mengchao Zhang[1,†]

[1]*Department of Physics and Siyuan Laboratory, Jinan University, Guangzhou 510632, P.R. China*
[2]*CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, P. R. China*
[3]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, P. R. China*
[4]*Department of Physics and Institute of Theoretical Physics, Nanjing Normal University, Nanjing 210023, P. R. China*

(Received 22 November 2020; accepted 11 May 2021; published 14 June 2021)

A vector-like top partner plays a central role in many new physics models that attempt to address the hierarchy problem. The top partner is conventionally assumed to decay to a quark and a Standard Model (SM) boson. However, it is also possible that the top partner decays to a non-SM scalar and a top quark. Such an exotic decay channel can be the main decay channel of top partners, and thus provides new windows to search for top partners at the LHC. In this paper, with the classical machine learning method of boosted decision trees, we perform a model-independent study of the discovery potential of this new signature at the LHC. In order to suppress the main QCD background, we consider the subdominant but clean decay channel $a \to \gamma\gamma$. For completeness, the single production process and pair production process of top partners are taken into account. We find that, for both single and pair production, the future high-luminosity LHC can exclude a top partner mass up to the TeV scale through the channel $T \to ta$ $(a \to \gamma\gamma)$, even if $\mathrm{BR}(a \to \gamma\gamma)$ is as small as $\mathcal{O}(0.1\%)$. Besides, our result shows that single production can overmatch pair production at the 14 TeV LHC, provided that the top partner is heavier than 800–900 GeV.

DOI: 10.1103/PhysRevD.103.115017

I. INTRODUCTION

The observation of the Higgs boson has completed the Standard Model (SM) [1,2]. However, the radiative stability of the Higgs boson mass is widely considered as a major theoretical motivation for new physics beyond the SM. A popular method to cure this problem is to introduce a softly-broken supersymmetry (SUSY), and then quadratic divergences can be canceled exactly by the superpartners. Due to the large top Yukawa coupling, the spin-0 top squark plays a central role in SUSY searches. A spin-$1/2$ vector-like top partner (denoted as T) can also arise in new physics models that attempt to stabilize the Higgs mass, like the composite Higgs model with partial compositeness [3–10]. Through the mixing with the top quark, the spin-$1/2$ vector-like top partner decays to bW^+, tZ, and th [11–13]. Current direct searches, which are designed for these conventional decay channels, have excluded the mass of the top partner up to about 1 TeV [14–23].

However, it is possible for the vector-like top partner to decay exotically [24–39]. For example, if the ultraviolet (UV) completion of a composite Higgs model is constructed by introducing new fermions that are charged under a new strong gauge interaction, then generally there are other light pseudo-Nambu-Goldstone bosons (pNGBs) in addition to the SM Higgs doublet [40–43]. In such a UV construction, a light CP-odd pseudoscalar (noted as a), which is associated with a nonanomalous axial $U(1)$ global symmetry, always arises. This can lead to a new decay channel of the vector-like top quark, $T \to ta$, if this is allowed by kinematics.[1] If a is heavier than 350 GeV, it mainly decays to $t\bar{t}$ and results in six top quark final states [28]. If a is lighter than 350 GeV, its dominant decay channel can be $a \to b\bar{b}$ or $a \to gg$. In the former case, due to multiple b jets in the final state, current data can exclude m_T up to about 1 TeV [33]. In the latter case, the huge QCD background greatly reduces the sensitivity of current LHC searches, and an m_T around 400–550 GeV can still survive under current direct search bounds [33].

*Co-corresponding author.
leiwu@njnu.edu.cn
†Corresponding author.
mczhang@jnu.edu.cn

[1] a or other pNGBs can also be directly probed by searching for diboson or fermion pair signals [44,45].

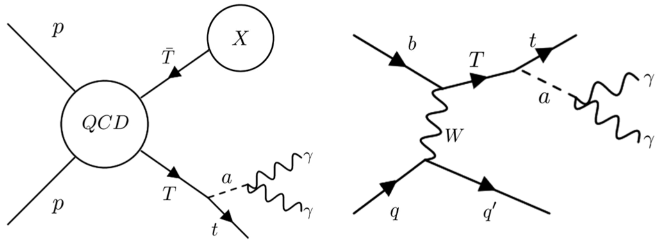

FIG. 1. Feynman diagrams: the pair production process $pp \to T\bar{T}$ (left) and the single production process $pp \to Tj$ (right) at the LHC. The bottom quark is considered as a parton in the proton.

To overcome the difficulties in the case of $\mathrm{BR}(a \to gg) \approx 1$, in this paper we consider the subdominant but much cleaner decay channel of the pNGB a, $a \to \gamma\gamma$. Different from previous works [30], we will adopt the classical machine learning method of boosted decision trees (BDT) [46] to improve the search sensitivity,[2] and focus on the single production process of T, $pp \to Tj$ [cf. Fig. 1 (right)]. The pair production process of T, $pp \to \bar{T}T$, will also be considered as a comparison with the single production process [cf. Fig. 1(left)].[3] Due to the large QCD coupling, $pp \to \bar{T}T$ is the conventional production process in top partner searches. In contrast with the pair production process, $pp \to Tj$ is induced by electroweak coupling [50,51] which is much weaker than QCD coupling, and thus its cross section is generally considered to be negligible. But, single production of the top partner has a larger phase space and can be enhanced by the collinear effect from the light quark emitting a W boson in the high-energy region [52,53]. These features may make the single production process a sensitive probe of the top partner at the LHC, especially when the top partner is heavy.[4] In addition, our analysis is performed in a model-independent way, and can be easily applied to a concrete model.

The rest of this paper is organized as follows. In Sec. II we present the model framework, which is the SM extended by a vector-like top partner T and a light pNGB a. In Sec. III we perform a detailed Monte Carlo simulation of our signal process and main background process. Model-independent exclusion limits for the single and pair production processes and a search sensitivity comparison are given. In Sec. IV we use two benchmark models to show how to obtain the exclusion limits for concrete models by using our model-independent results. Finally, we conclude our work in Sec. V.

[2]Other machine learning methods have been used in top partner searches, e.g., Ref. [47].

[3]In composite Higgs models, the top partner could have more a exotic production process, e.g., Refs. [48,49].

[4]For a recent study of singly produced vector-like quarks, see Ref. [54].

II. MODEL FRAMEWORK

We consider the SM extended by a vector-like top partner T (with electric charge $+2/3$) and a light pseudoscalar a. This simplified scenario can be embedded in many new physics models [30–32,42,55–58]. To be specific, the relevant Lagrangian of the vector-like top partner can be expressed as [59]

$$\begin{aligned}\mathcal{L}_T = \bar{T}(i\not{D} - m_T)T + \bigg(\kappa^T_{W,L}\frac{g}{\sqrt{2}}\bar{T}\not{W}^+P_L b \\ + \kappa^T_{Z,L}\frac{g}{2c_W}\bar{T}\not{Z}P_L t - \kappa^T_{h,L}\frac{m_T}{v}\bar{T}hP_L t \\ + i\kappa^T_{a,L}\frac{m_T}{v}\bar{T}aP_L t + L \leftrightarrow R + \mathrm{H.c.}\bigg),\end{aligned} \tag{1}$$

where m_T is the mass of the top partner, and the couplings κ^T_i describe the effective interactions between the top partner and other particles. Since we are interested in the single top partner production process $pp \to T(\to ta)j$, which comes from the electroweak coupling, we only keep $\kappa^T_{W,L}$ and $\kappa^T_{a,L}$,[5] and neglect other κ^T_i in this work for simplicity. The appearance of κ^T_W will inevitably lead to the conventional decay channel $T \to bW^+$. Current direct searches [17] already exclude the mass of T up to 1.1 TeV, provided $\mathrm{BR}(T \to bW^+)$ is 0.5. However, κ^T_W is determined by the $SU(2)$ gauge coupling and the mixing angle between the top partner and the elementary top [26]. But κ^T_a is induced by the coupling between the top partner and the pNGB a, which can come from the strong interaction sector. So κ^T_a can be much larger than κ^T_W, and makes $T \to ta$ the main decay channel. $\mathrm{BR}(T \to bW^+)$ can be small enough to escape current direct search bounds. Concrete models and a discussion will be given in Sec. IV.

On the other hand, the Lagrangian of the pseudoscalar a and its interactions with SM particles can be expressed as

$$\begin{aligned}\mathcal{L}_a = \frac{1}{2}(\partial_\mu a)(\partial^\mu a) - \frac{1}{2}m_a^2 a^2 + \sum_f \frac{iC^a_f m_f}{f_a} a\bar{f}\gamma^5 f \\ + \frac{g_s^2 K^a_g}{16\pi^2 f_a} aG^a_{\mu\nu}\tilde{G}^{a\mu\nu} + \frac{e^2 K^a_\gamma}{16\pi^2 f_a} aA_{\mu\nu}\tilde{A}^{\mu\nu} \\ + \frac{g^2 c_W^2 K^a_Z}{16\pi^2 f_a} aZ_{\mu\nu}\tilde{Z}^{\mu\nu} + \frac{egc_W K^a_{Z\gamma}}{8\pi^2 f_a} aA_{\mu\nu}\tilde{Z}^{\mu\nu} \\ + \frac{g^2 K^a_W}{16\pi^2 f_a} aW_{\mu\nu}\tilde{W}^{\mu\nu},\end{aligned} \tag{2}$$

[5]If we inversely select a purely right-handed coupling, or a mixture of these two cases, the b jet coming from the $T \to ta \to bWa$ decay chain will be slightly less energetic (see, e.g., Ref. [60]). But the b jet p_T is not a sensitive parameter in our following signal-background discrimination, and thus our final exclusion bounds will hardly change if the coupling is not purely left-handed. For conciseness, $\kappa^T_{W,L}$ and $\kappa^T_{a,L}$ will be denoted as κ^T_W and κ^T_a in the rest of this paper.

where the gauge boson field strength and its conjugate are denoted as $V_{\mu\nu} \equiv \partial_\mu V_\nu - \partial_\nu V_\mu$ and $\tilde{V}_{\mu\nu} \equiv \epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}$. $G^a_{\mu\nu}$, $A_{\mu\nu}$, $Z_{\mu\nu}$, and $W_{\mu\nu}$ are the field strengths of the gluon, photon, Z boson, and W boson, respectively. m_a and m_f are the masses of the pseudoscalar a and SM fermions. C^a_f and K^a_V are dimensionless coupling coefficients of $a\bar{f}f$ and $aV\tilde{V}$ terms. g_s, e, g, and c_W are the strong coupling constant, electric charge, $SU(2)$ coupling constant, and the cosine of the Weinberg angle, respectively.

If C^a_f are comparable with K^a_V, then $\mathrm{BR}(a \to \bar{b}b)$ will be a dominant decay channel of a when $m_a < 350$ GeV. In this case, multiple bottom quark jets in the final state help to suppress the QCD background, and $m_T \lesssim 1$ TeV can be excluded by current data [33]. However, if C^a_f is negligible, $a \to gg$ will be the dominant decay channel and current direct searches become insensitive, especially when $m_a \lesssim 100$ GeV [33]. This is because it is difficult for us to identify the jet pair coming from a decay in the huge QCD background. To enhance the search sensitivity, we can consider the minor decay channel $a \to \gamma\gamma$. Due to the hierarchy between g_s^2 and e^2, $\mathrm{BR}(a \to \gamma\gamma)$ is generally much smaller than $\mathrm{BR}(a \to gg)$ in most models. But the photon pair signature from this decay channel is very clean, and it helps to greatly suppress the QCD background. In the rest of this paper, we will assume that C^a_f is negligible and focus on the photon pair signature in the collider analysis.

III. COLLIDER SIMULATION AND ANALYSIS

In Fig. 2, we present the cross sections of single and pair production of the top partner at the 14 TeV LHC. We normalize the leading-order results, which are calculated using MadGraph5 [61], to the next-to-leading-order QCD predictions by using K-factors 0.95 [62] and 1.3 [63] for the single and pair production process, respectively. It can be seen that, as m_T increases, the cross section of T pair production decreases faster than the cross section of single T production. In the parameter region where m_T is greater than 1.2 TeV (1.7 TeV), the single production can have a cross section larger than the pair production at the 14 TeV LHC, when the value of k^T_W is 0.1 (0.05).

In our basic setting, we only maintain the top partner's couplings κ^T_W and κ^T_a, and thus T can only decay to ta or bW^+. As we explained in Sec. II, we will only consider the decay channel $T \to ta$ and neglect $T \to bW^+$. In order to suppress the QCD background, for both the single and pair production processes, we require a photon pair to appear in the final state. Thus, for single production the only process we need to consider is $pp \to Tj \to taj$ followed by $a \to \gamma\gamma$. For pair production, due to a generally very small $\mathrm{BR}(a \to \gamma\gamma)$, the signal process we consider is $pp \to T\bar{T} \to t\bar{t}aa$, followed by one a decaying to $\gamma\gamma$ and another a decaying to gg. After such a choice, we

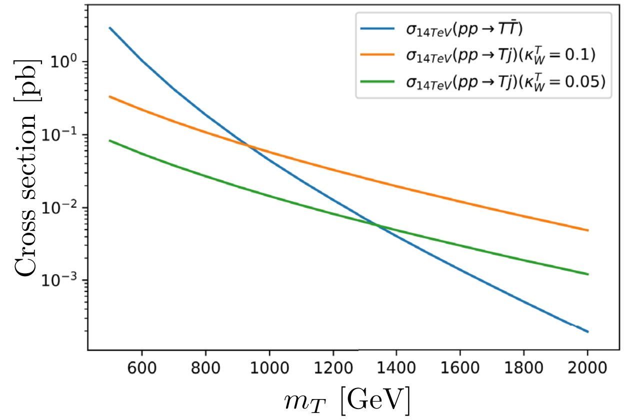

FIG. 2. Cross sections of single T and T pair production at the 14 TeV LHC. The conjugate process $pp \to \bar{T}j$ is included in single production as well. The coupling k^T_W is set to 0.1 and 0.05 as examples.

can treat the production cross sections $\mathrm{BR}(T \to ta)$, and $\mathrm{BR}(a \to \gamma\gamma)$ as undetermined parameters, and focus on the kinematic variables' distribution.[6] Thus, for a model-independent analysis, only m_T and m_a are relevant.

For Monte Carlo simulation, we implement our effective Lagrangian in a UFO model file [64] using FeynRules [65], and generate the parton-level signal and background events with MadGraph5 [61]. Parton shower and hadronization are performed with PYTHIA8 [66]. The detector effect is simulated using Delphes [67]. We assume the b-tagging efficiency to be 70% and the rate of mistagging a light-quark jet or gluon jet as a b jet to be 1%. A jet might be mistagged as a photon in a hadron collider environment. We use the jet-faking-photon rate given in Ref. [68] to estimate this effect.

A. Single production of the top partner

The full signal process we consider is $pp \to Tj \to t(\to bl^+\nu_l)a(\to \gamma\gamma)j$. The SM backgrounds are from the resonant processes $pp \to t\bar{t}h$, $pp \to Whjj$, and $pp \to thj$, where the SM Higgs decays to a photon pair. Besides, there are also the nonresonant backgrounds $pp \to t\bar{t}\gamma\gamma$, $pp \to tj\gamma\gamma$, and $pp \to Wjj\gamma\gamma$, where two photons come from the radiation of charged particles. Each background process, e.g., $pp \to t\bar{t}\gamma\gamma$, can be faked by $pp \to t\bar{t}j\gamma$ ($pp \to t\bar{t}jj$) with one (two) hard jets mistagged as photons. We find that, due to a low jet-faking-photon rate (generally smaller than 0.05% [68]), the cross section of a process with a faked photon is at least 1 order of magnitude smaller than the cross section of a process without a faked photon. So, we neglect the jet-faking-photon effect in the rest of this paper, and only consider

[6] $\mathrm{BR}(a \to gg)$ is almost equal to 1 after we assume that C^a_f is negligible in the Lagrangian (2).

TABLE I. Cut-flow table of our basic selection criteria for the single production process.

Basic Selection	BP1 (fb)	BP2 (fb)	$t\bar{t}h(h\to\gamma\gamma)$ (fb)	$Whjj(h\to\gamma\gamma)$ (fb)	$t\bar{t}\gamma\gamma$ (fb)	$tj\gamma\gamma$ (fb)	$Wjj\gamma\gamma$ (fb)	$thj(h\to\gamma\gamma)$ (fb)
$1b$ jet with $p_{\mathrm{T}} > 20$ GeV and $\lvert\eta\rvert < 2.5$	1.039	1.033	0.417	0.0744	5.30	10.72	32.85	0.0535
1 lepton with $p_{\mathrm{T}} > 20$ GeV and $\lvert\eta\rvert < 2.5$	0.238	0.237	0.135	0.015	1.67	2.08	4.60	0.011
2γ with $p_{\mathrm{T}} > 20$ GeV and $\lvert\eta\rvert < 2.5$	0.075	0.068	0.085	0.0089	0.60	0.73	1.64	0.006
1 jet with $p_{\mathrm{T}} > m_T/8$ and $\lvert\eta\rvert < 2.5$	0.026	0.021	0.0073	0.0011	0.081	0.19	0.279	0.0014

$pp\to t\bar{t}h$, $pp \to Whjj$, $pp \to thj$, $pp \to t\bar{t}\gamma\gamma$, $pp \to tj\gamma\gamma$, and $pp \to Wjj\gamma\gamma$ as our background processes.[7]

We use basic selection criteria to select the events used in our analysis:

(1) Exactly one b jet with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$.
(2) Exactly one isolated lepton (electron or muon) with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$.
(3) Exactly two isolated photons with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$.
(4) Exactly one jet with $p_{\mathrm{T}} > m_T/8$ and $|\eta| < 2.5$. This jet cannot be tagged as a b jet.

The first three selection criteria simply require the particles in the final state of a signal process to exist. The final-state parton recoiled with the top partner is generally quite hard, and thus we require a normal jet (i.e., light-quark jet) to have a large p_{T}, and this p_{T} cut also depends on the value of m_T. Furthermore, we require only one normal jet with $p_{\mathrm{T}} > m_T/8$. This requirement is used to exclude multijet events from background processes and the pair production process. Events from the pair production process can also pass the above basic selection criteria with a low probability. In the Appendix we show that this contamination from the pair production process can be ignored if k_W^T is not too small. Finally, we need to emphasize that the cone size used in photon isolation is $\Delta R = 0.2$, and thus signal events with two final-state photons that are too close to each other will be discarded.

To design a cut flow that can be used to enhance search significance, we need to study the kinematics of the final states for both signal and background processes. To illustrate the distribution of kinematics variables, we consider two benchmark points (BPs):

$$\mathrm{BP\ 1}: m_T = 800\ \mathrm{GeV}, \qquad m_a = 50\ \mathrm{GeV}, \qquad (3)$$

$$\mathrm{BP\ 2}: m_T = 800\ \mathrm{GeV}, \qquad m_a = 200\ \mathrm{GeV}. \qquad (4)$$

We also fix κ_W^T, $\mathrm{BR}(T\to ta)$, and $\mathrm{BR}(a\to\gamma\gamma)$ to 0.1, 100%, and 1%, respectively,[8] We use this setting as an example to show the cross section of signal and background after the basic selection. The cut-flow table of our basic selection for these two BPs and the main backgrounds is given in Table I.

After the basic selection, the type and number of final-state particles are the same for signal and background events, while the cross section of the background is still about 2 orders of magnitude larger than the cross section of the signal process. To further suppress the background and enhance search sensitivity, we need to study the kinematics for both signal and background. Here we study the following variables that we think can be used in signal and background discrimination.

(1) Pseudorapidity η of the hardest jet. The hardest jet in the signal process is evolved from the parton recoiled with the heavy top partner T, while the hardest jet in the background process comes from parton radiation. Thus, the hardest jet in the signal process tends to be more central than the hardest jet in the background process.
(2) We define a new variable which is called the "reconstructed top partner mass" and is denoted as $\tilde{m}_T$:

$$\tilde{m}_T = \sqrt{(p_{\mathrm{visible}} + p_{\mathrm{invisible}})^2}. \qquad (5)$$

Here, the visible 4-momentum p_{visible} is the sum of the 4-momenta of the b jet, lepton, and two photons. The invisible 4-momentum $p_{\mathrm{invisible}}$ is defined as $(\not{E}_{\mathrm{T}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, 0)$. The difference between p_ν and $p_{\mathrm{invisible}}$ is the longitudinal momentum of the neutrino. Due to the rather long decay chain of T, the momentum carried by the neutrino is not so large compared with m_T. Thus, missing the longitudinal momentum of the neutrino will not significantly change the reconstruction of the top partner mass, and we can expect $\tilde{m}_T \sim m_T$ for the signal process.

[7]Pure QCD processes like $pp \to t\bar{t}jj$ have a cross section that is much larger than that of $pp \to t\bar{t}\gamma\gamma$. So, turning off the negligible QCD processes like $pp \to t\bar{t}jj$ helps us to save lots of simulation time.

[8]For a concrete model, $\mathrm{BR}(T \to ta)$ surely cannot be 100%. In this section we only perform a model-independent analysis and focus on the variable distribution.

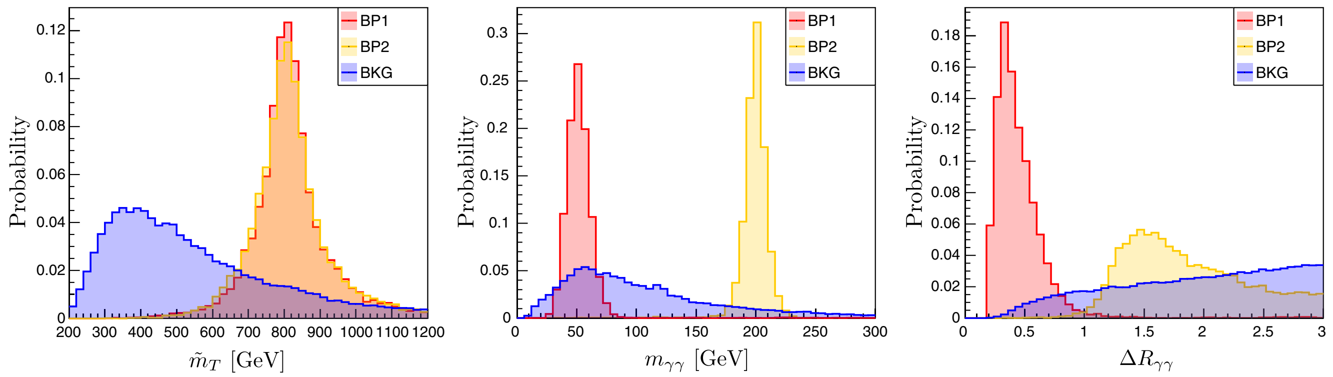

FIG. 3. Normalized distributions of the variables that are very different for signal and background processes. Here, BKG (background) means the combination of all six background processes. The proportion of each background process is proportional to their cross section after the basic selection.

(3) The distance between two photons $\Delta R_{\gamma\gamma}$, which is defined as

$$\Delta R_{\gamma\gamma} = \sqrt{(\Delta\eta_{\gamma\gamma})^2 + (\Delta\phi_{\gamma\gamma})^2}, \tag{6}$$

where $\Delta\eta_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ are the pseudorapidity and azimuthal angle difference between the two photons. Because a is highly boosted when T is much heavier than a, $\Delta R_{\gamma\gamma}$ tends to be small in signal events. $\Delta R_{\gamma\gamma}$ in background processes would be quite random because the photons mainly come from charged particle radiation. The Higgs in the background process is generally not too boosted, and thus the two photons from the decay of the Higgs tend to go back to back.

(4) Invariant mass of the photon pair $m_{\gamma\gamma}$:

$$m_{\gamma\gamma} = \sqrt{(p_{\gamma 1} + p_{\gamma 2})^2}, \tag{7}$$

$m_{\gamma\gamma}$ should be around m_a if Γ_a is not too large. Because the decay constant f_a in the Lagrangian (2) is generally larger than a TeV [45], $\Gamma_a < 1$ GeV can be satisfied in almost all of the parameter space and we can observe a spiky $m_{\gamma\gamma}$ distribution in the signal process.

(5) The transverse momentum of visible final-state particles also provides useful information. The final-state visible particles in the signal process come from the decay of a heavy T or the particle recoiled with it, while the final-state visible particles in the background process are from the radiation or the decay of $W/h/t$. So, we can expect that the p_{T} of the final-state visible particles in the signal process are larger than the p_{T} of the final-state visible particles in the background process. Here we consider the following variables for discrimination:

p_{T} of the hardest jet, p_{T} of the second hardest jet,
p_{T} of b jet, p_{T} of lepton,
p_{T} of the first photon, p_{T} of the second photon.

In Fig. 3 we show the normalized distributions of three variables that are quite different for our BPs and the SM background. As we expected, for the signal process the mass of the top partner T and pseudoscalar a can be reconstructed very well, and the photon pair is more collinear. Variables we do not show here also have different distributions for signal and background, but they are subdominant in our signal and background discrimination.

In order to fully utilize all the information to distinguish signal and background, we use a BDT [46], which is implemented in TMVA-Toolkit [69], to do a multiple-variable analysis. All of the variables after the basic selection are used as input:

$$\left\{ \begin{array}{c} \text{First jet } \eta, \tilde{m}_T, \Delta R_{\gamma\gamma}, m_{\gamma\gamma}, \text{ first jet } p_{\mathrm{T}}, \\ \text{Second jet } p_{\mathrm{T}}, \text{ first } \gamma p_{\mathrm{T}}, \text{ second } \gamma p_{\mathrm{T}}, l p_{\mathrm{T}}, b \text{ jet } p_{\mathrm{T}} \end{array} \right\}. \tag{8}$$

In the BDT setting, we use 200 decision trees, choose the minimum in the leaf node as 2.5%, and set the maximum depth as 3. Half of the events are chosen as test events, and the Kolmogorov-Smirnov test is required to be larger than 0.01 to avoid overtraining. BDT maps the multiple variables to a BDT response. A signal-like event tends to get a large response, while a background-like event tends to get a small response. Thus, the cut can be easily performed by requiring the BDT response to be greater than a certain value.[9] We denote the amount of signal and background events after the

[9]As present in Fig. 4, the variable distributions for BP1 and BP2 are different, especially $\Delta R_{\gamma\gamma}$ and $m_{\gamma\gamma}$, so we do not expect to be able to use one single BDT to distinguish all model points. Instead, for each m_T and m_a, we need to train a corresponding BDT for discrimination.

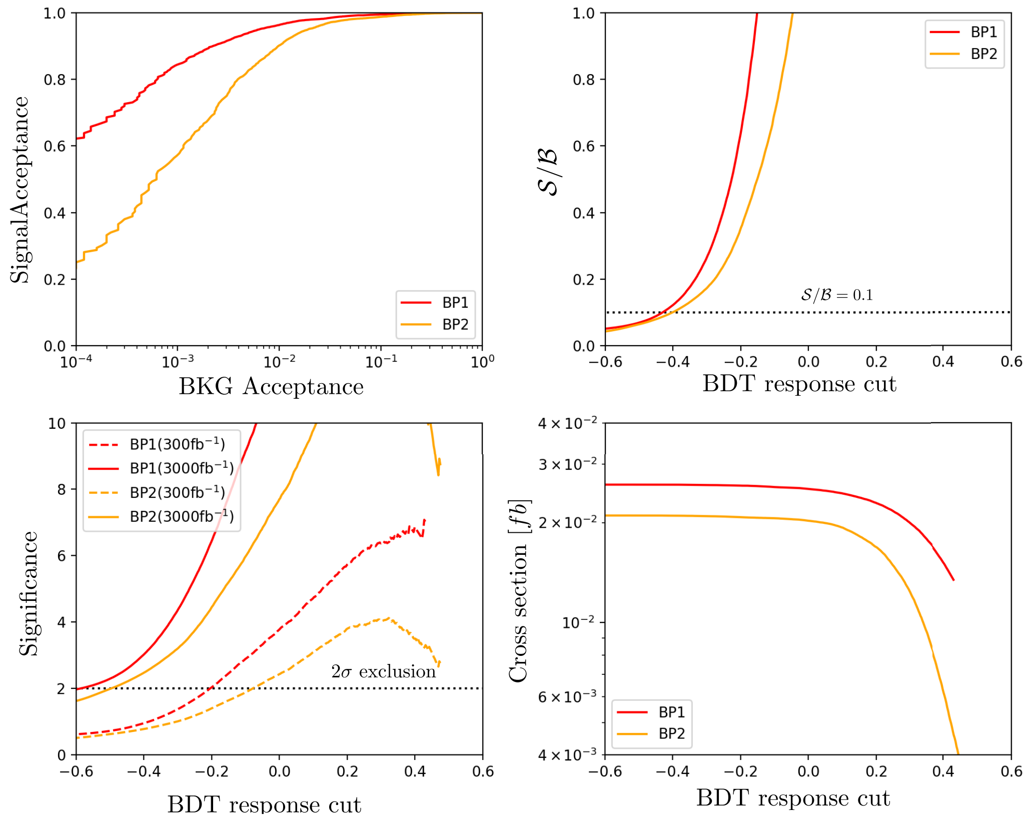

FIG. 4. Upper left: ROC curve for signal and background discrimination. Upper right: $\mathcal{S}/\mathcal{B}$ as functions of the BDT response cut. Bottom left: significance as a function of the BDT response cut. Unlike $\mathcal{S}/\mathcal{B}$, the significance also depends on the integrated luminosity. Bottom right: cross section of the signal process as a function of the BDT response cut.

BDT response cut as $\mathcal{S}$ and $\mathcal{B}$, respectively. The statistical significance is evaluated using the Poisson formula

$$\text{Significance} = \sqrt{2\left[(\mathcal{S}+\mathcal{B})\ln\left(1+\frac{\mathcal{S}}{\mathcal{B}}\right)-\mathcal{S}\right]}. \quad (9)$$

If Significance > 2 and the experimental result is consistent with the SM expectation, then this model is excluded at the 2σ confidence level. However, sometimes our BDT cut can remove the background very effectively and make the search almost background free ($\mathcal{B}\sim 0$). In that case we cannot use Eq. (9) to estimate the statistical significance, but rather should use $\mathcal{S}=3$ to determine the 2σ exclusion limits.[10] So we require $\mathcal{S}$ to be always greater than 3 to prevent overestimating the Significance. Considering the systematics at a hadron collider, we also require $\mathcal{S}/\mathcal{B}$ to be larger than 0.1. In Fig. 4 (upper left) we present receiver operating characteristic (ROC) curves obtained using the BDT response cut. It clearly shows that we can lower the background by 3 or 4 orders of magnitude without hurting the signal too much. In Fig. 4 (upper right, bottom left, and bottom right) we present $\mathcal{S}/\mathcal{B}$, the Significance, and the cross section of the signal process as functions of the BDT response cut. It can be seen that we can change the BDT response cut to rapidly increase the Significance and $\mathcal{S}/\mathcal{B}$, and thus Significance > 2, $\mathcal{S}/\mathcal{B} > 0.1$, and $\mathcal{S}\geq 3$ can be satisfied at the same time. So, both BP1 and BP2 can be excluded by the single production process at the 14 TeV LHC, regardless of whether the luminosity is 300 or 3000 fb^{-1}.

Inversely, if we treat the production cross section and BRs as free parameters, then the distributions of those variables can be used to determine the 2σ exclusion upper limit on $\sigma(pp\to Tj)\times \text{BR}(T\to at)\times \text{BR}(a\to\gamma\gamma)$, as a function of (m_T, m_a). For certain (m_T, m_a), the upper limit of $\sigma(pp\to Tj)\times \text{BR}(T\to at)\times \text{BR}(a\to\gamma\gamma)$ is its minimum value that satisfies Significance > 2, $\mathcal{S}/\mathcal{B} > 0.1$, and $\mathcal{S}\geq 3$, under the optimal BDT response cut. Repeating this process for each (m_T, m_a), we obtain Fig. 5, which shows the upper limits of $\sigma(pp\to Tj)\times \text{BR}(T\to at)\times$ $\text{BR}(a\to\gamma\gamma)$ in the $m_a - m_T$ plane.[11] The upper limit with a

[10]This is also called the "rule of three" in statistics.

[11]In Fig. 5, the upper limit on m_a is set to 350 GeV, which is roughly twice the mass of the top quark. Due to the large value of the top Yukawa coupling, $a\to t\bar{t}$ should be the main decay channel of the pseudoscalar a, provided $m_a > 350$ GeV. This scenario has been studied in a previous work [28], so in this work we only focus on the region with $m_a < 350$ GeV.

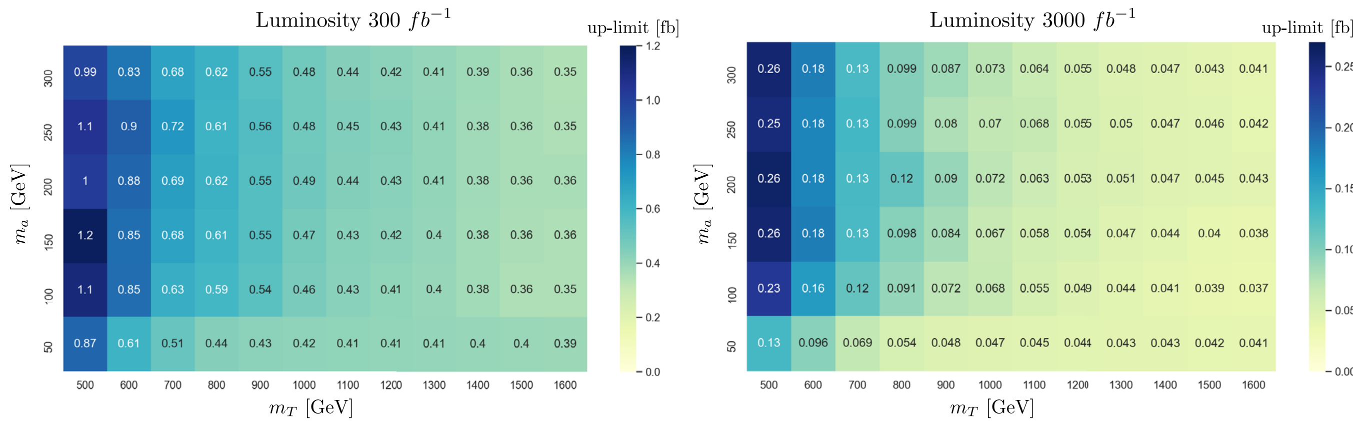

FIG. 5. Model-independent 2σ exclusion limit on $\sigma(pp \to Tj) \times \mathrm{BR}(T \to at) \times \mathrm{BR}(a \to \gamma\gamma)$ as a function of (m_T, m_a), with an integrated luminosity of 300 fb^{-1} (left) and 3000 fb^{-1} (right).

luminosity of 3000 fb^{-1} is certainly much smaller than the upper limit with a luminosity of 300 fb^{-1}, because the Significance is approximately proportional to the square root of the integrated luminosity. Besides, the upper limit on $\sigma(pp \to Tj) \times \mathrm{BR}(T \to at) \times \mathrm{BR}(a \to \gamma\gamma)$ decreases as m_T increases. This is simply because the decay products of T become more and more energetic as m_T increases, and thus it will be easier to distinguish the signal from the SM background.

The upper limits we obtained here can be applied to a concrete model by simply calculating $\sigma(pp \to Tj)$ and branching ratios. We will show this in Sec. IV.

B. Pair production of the top partner

For the pair production process, due to a small $\mathrm{BR}(a \to \gamma\gamma)$, we only require one photon pair in the final state. The full process is $pp \to T\bar{T}aa \to t\bar{t}\gamma\gamma gg$, followed by t decaying leptonically. Similar to what we did for the single production process, we also use "basic selection + BDT" to study the pair production process. The basic selection criteria used for pair production are as follows:

(1) Exactly two b jets with $p_\mathrm{T} > 20$ GeV and $|\eta| < 2.5$.
(2) Exactly two leptons (electron or muon) with $p_\mathrm{T} > 20$ GeV and $|\eta| < 2.5$.
(3) Exactly two photons with $p_\mathrm{T} > 20$ GeV and $|\eta| < 2.5$.

Here we simply require the final-state particles that can be used to suppress the QCD background to exist.

For illustration, we use the same benchmark points as we used in the last subsection:

$$\mathrm{BP}\ 1: m_T = 800\ \mathrm{GeV}, \qquad m_a = 50\ \mathrm{GeV}, \tag{10}$$

$$\mathrm{BP}\ 2: m_T = 800\ \mathrm{GeV}, \qquad m_a = 200\ \mathrm{GeV}. \tag{11}$$

We also fix $\mathrm{BR}(T \to ta)$ and $\mathrm{BR}(a \to \gamma\gamma)$ to 100% and 1% to estimate the search sensitivity of these two BPs. The value of κ_W^T is not important here, because we assume $\mathrm{BR}(T \to bW)$ to be negligible. The cut-flow table of our basic selection for these two BPs and the main backgrounds is given in Table II. Because we require two b jets and two leptons in the final state, compared with the result presented in Table I, the cross sections for background (BKG) processes after the basic selection are much smaller in Table II.

In our pair process, there are two branches of particles coming from the decay of $T\bar{T}$, so it is difficult for us to distinguish which final-state particles originate from the same mother particle. Thus, it is not easy to reconstruct the mass of the top partner. Instead, we replace the reconstructed top partner mass $\tilde{m}_T$ by the scalar p_T sum of all final-state objects and missing energy, H_T:

$$H_T = \sum_i p_{\mathrm{T,i}} + \not{E}_\mathrm{T}. \tag{12}$$

Here the index "i" denotes two selected b jets, two selected leptons, two selected photons, and the first and second hardest jets. H_T roughly reflects the energy scale of the hard process, and helps to distinguish signal and background. Furthermore, we replace the η of the hardest jet

TABLE II. Cut-flow table of our basic selection criteria for the pair production process.

Basic Selection	BP1 (fb)	BP2 (fb)	$t\bar{t}h(h\to\gamma\gamma)$ (fb)	$Whjj(h\to\gamma\gamma)$ (fb)	$t\bar{t}\gamma\gamma$ (fb)	$tj\gamma\gamma$ (fb)	$Wjj\gamma\gamma$ (fb)	$thj(h\to\gamma\gamma)$ (fb)
$2b$ jet with $p_\mathrm{T} > 20$ GeV and $\lvert\eta\rvert < 2.5$	0.656	0.881	0.179	0.00327	2.39	1.38	1.46	0.00698
2 lepton with $p_\mathrm{T} > 20$ GeV and $\lvert\eta\rvert < 2.5$	0.090	0.105	0.0133	0.00009	0.172	0.02107	0.0132	0.00030
2γ with $p_\mathrm{T} > 20$ GeV and $\lvert\eta\rvert < 2.5$	0.0295	0.0328	0.00821	0.00005	0.0615	0.00596	0.00602	0.00020

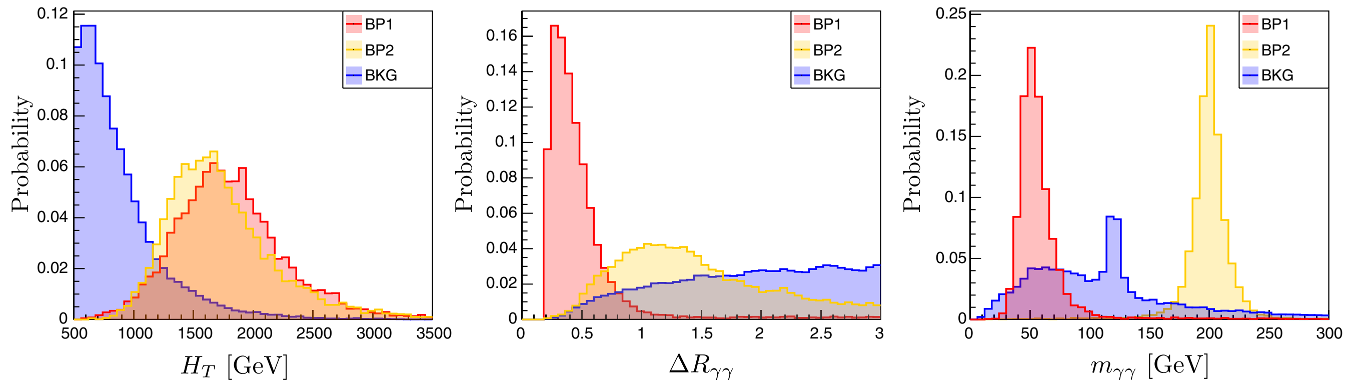

FIG. 6. Normalized distributions of the variables that are very different for signal and background processes. Here BKG (background) means the combination of all six background processes. The proportion of each background process is proportional to their cross section after the basic selection.

with the missing energy $\not{E}_\mathrm{T}$, and because of the two heavy resonances T, generally speaking, $\not{E}_\mathrm{T}$ in the signal process is larger than $\not{E}_\mathrm{T}$ in the SM background process.

For the pair production process, we use the following variables as input for the BDT to distinguish signal and background:

$$\left\{ \begin{matrix} \not{E}_\mathrm{T}, H_T, \Delta R_{\gamma\gamma}, m_{\gamma\gamma}, \text{ first } b \text{ jet } p_\mathrm{T}, \text{ second } b \text{ jet } p_\mathrm{T}, \\ \text{first } \gamma\ p_\mathrm{T}, \text{ second } \gamma\ p_\mathrm{T}, \text{ first lepton } p_\mathrm{T}, \text{ second lepton } p_\mathrm{T} \end{matrix} \right\}.$$

In Fig. 6 we show the normalized distributions of three variables that are very different for our benchmark signal settings and the SM backgrounds. As we expected, the distribution of H_T picks around $2m_T$ for signal process. The BDT setting we used here is the same as the one we used in the last subsection. As we did before, in Fig. 7 we present ROC curves, $\mathcal{S}/\mathcal{B}$, Significance, and the cross section of the signal process as functions of the BDT response cut. To compare the search sensitivity of single production and pair production, in Fig. 7 (upper left) we also show the ROC curves obtained from the single production channel. It can be seen that the area under the ROC curves from the single production process is larger than the area under the ROC curves from the pair production process, while the $\mathcal{S}/\mathcal{B}$ and Significance curves indicate that our BPs can be excluded (or detected) with a much larger Significance through the pair production process.[12] This is simply because the BKG cross section becomes much smaller, as we have shown in Table II.

Similar to what we did for the single production process, we also present the 2σ exclusion upper limit on $\sigma(pp \to \bar{T}T) \times [\mathrm{BR}(T \to at)]^2 \times \mathrm{BR}(a \to \gamma\gamma)$ as a function of (m_T, m_a). Figure 8 is our result. A comparison of the exclusion ability of single and pair production cannot be directly obtained from Figs. 5 and 8, because these two processes have different cross sections. In the next subsection we use some plots which are easier to understand to make a comparison.

C. Single production vs pair production

Another issue we try to study in this work is the comparison of the search sensitivity between single production and the conventional pair production. As illustrated in Fig. 2, $\sigma(pp \to Tj)$ will exceed $\sigma(pp \to T\bar{T})$ when m_T is large enough. But for a realistic analysis, we also need to consider the effect of background. It is obvious from Figs. 5 and 8 that the exclusion upper limits of single production are much smaller than the exclusion upper limits of pair production. But the production cross sections of these two processes are also different. In order to clearly compare the search sensitivity, it is better to show the minimal integrated luminosity at the 14 TeV LHC that is needed for a parameter point to be excluded at the 2σ level. For this purpose, we fix $\mathrm{BR}(T \to ta) = 1$ and $\kappa_W^T = 0.1$, and then choose several specific values for m_a and $\mathrm{BR}(a \to \gamma\gamma)$. Then, the 2σ exclusion integrated luminosity can be given as a function of m_T. Figure 9 is the result. It clearly shows that pair production is more sensitive when $m_T \lesssim 750$ GeV (850 GeV), and single production is more sensitive when $m_T \gtrsim 750$ GeV (850 GeV), when m_a is 50 GeV (300 GeV).

[12]However, the Significance curve is calculated using Eq. (9), and it does not necessarily mean that the pair production process has a better exclusion ability when $m_T = 800$ GeV. When we decrease $\mathrm{BR}(a \to \gamma\gamma)$ to a lower order, we need to reduce $\mathcal{B}$ to close to zero. In this case, the 2σ exclusion limits should be obtained from $\mathcal{S} = 3$.

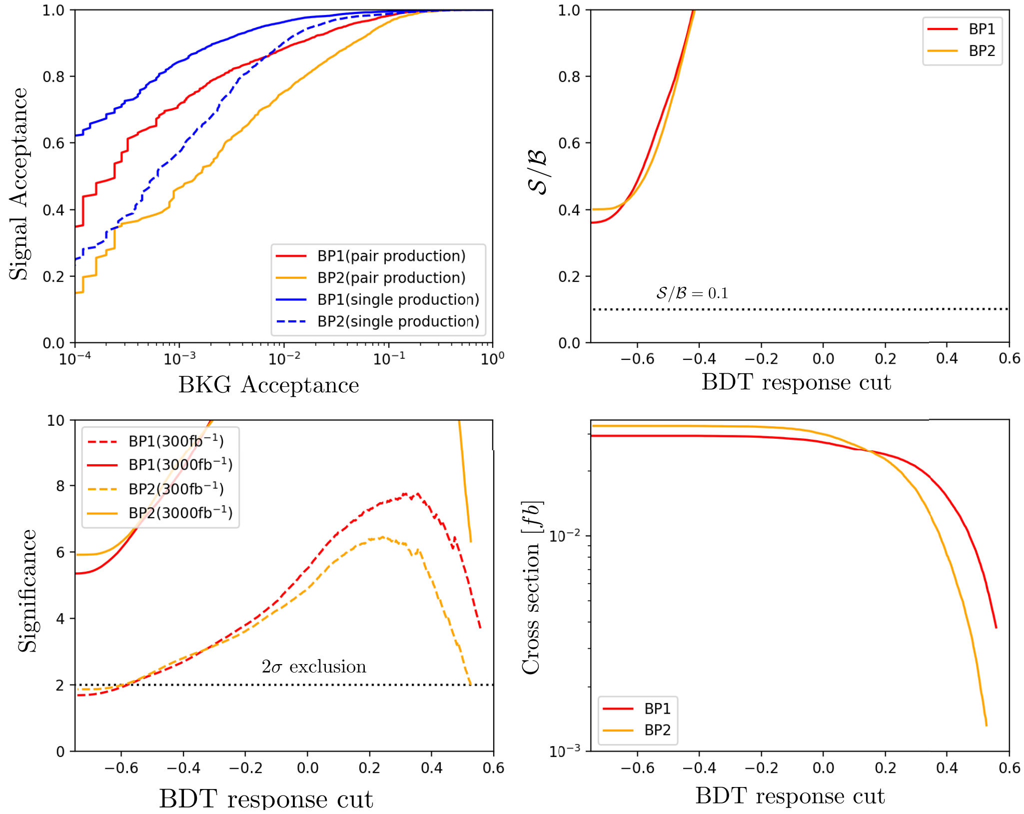

FIG. 7. Upper left: ROC curve for signal and background discrimination. Upper right: $\mathcal{S}/\mathcal{B}$ as functions of the BDT response cut. Bottom left: Significance as a function of the BDT response cut. Bottom right: cross section of the signal process a as function of the BDT response cut.

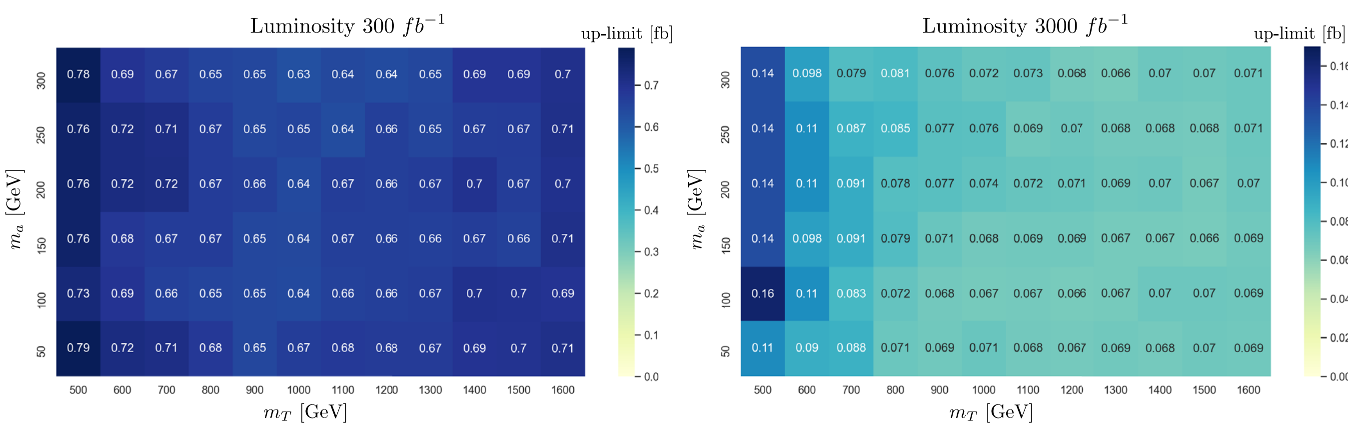

FIG. 8. Model-independent 2σ exclusion limit on $\sigma(pp \to \bar{T}T) \times [\mathrm{BR}(T \to at)]^2 \times \mathrm{BR}(a \to \gamma\gamma)$ as a function of (m_T, m_a), with an integrated luminosity of 300 (left) and 3000 fb^{-1} (right).

A more realistic comparison can be performed by fixing the 14 TeV LHC integrated luminosity to 3000 fb^{-1} and treating $\mathrm{BR}(a \to \gamma\gamma)$ as a free parameter. Then, the pair production and single production channels can exclude two different regions in the m_T vs $\mathrm{BR}(a \to \gamma\gamma)$ plane at the 2σ level. In Fig. 10 we present the excluded region, and it clearly shows that the single production channel excludes more parameter space when $m_T \gtrsim 800$ GeV (900 GeV) and $m_a = 50$ GeV (300 GeV).

IV. INTERPRETATION OF CONCRETE MODEL

In this section we apply the model-independent exclusion limits we obtained in the last section to concrete

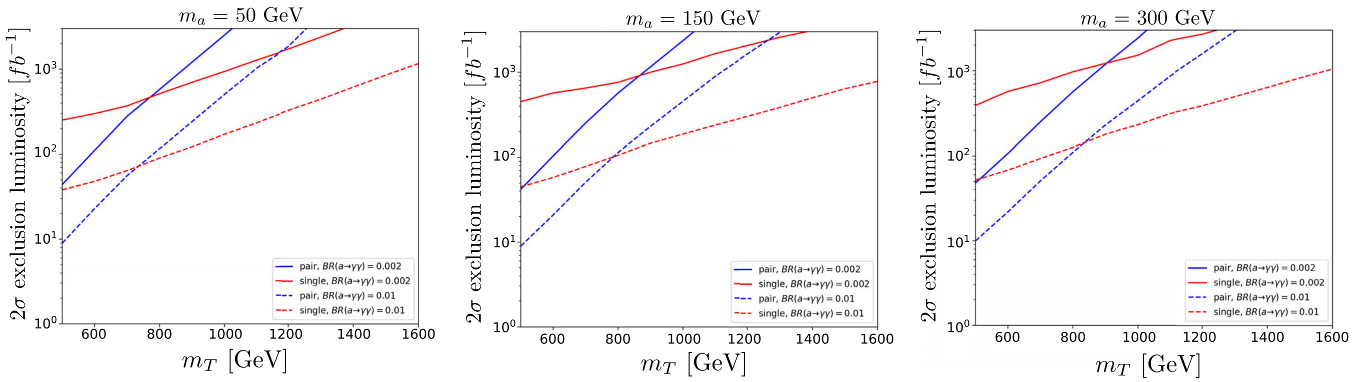

FIG. 9. 2σ exclusion integrated luminosity as a function of m_T for single and pair production at the 14 TeV LHC. Several values for m_a and $\mathrm{BR}(a \to \gamma\gamma)$ are chosen. $\mathrm{BR}(T \to ta)$ and κ_W^T are fixed to 1 and 0.1, respectively.

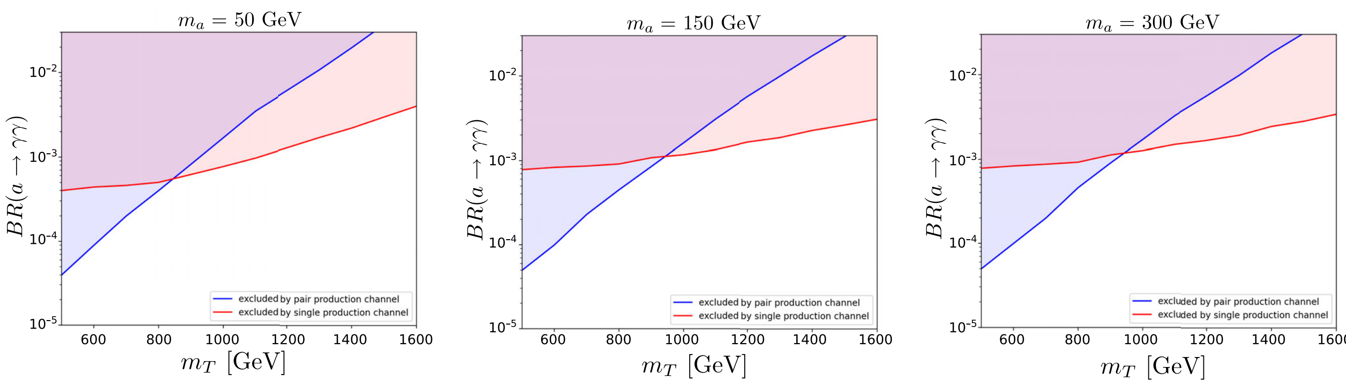

FIG. 10. 2σ exclusion limit in the m_T vs $\mathrm{BR}(a \to \gamma\gamma)$ plane for single and pair production. Here we consider the 14 TeV LHC with an integrated luminosity of 3000 fb^{-1}. $\mathrm{BR}(T \to ta)$ and κ_W^T are fixed to 1 and 0.1, respectively.

models. As explained in Sec. II, in this work we treat $T \to Zt$, $T \to ht$, and $a \to \bar{f}f$ as negligible, and for simplicity we further close all of the right-hand couplings. Thus, our Lagrangian can be simplified as

$$\mathcal{L}_T = \bar{T}(i\not{D} - m_T)T + \left(\kappa_W^T \frac{g}{\sqrt{2}} \bar{T} \not{W}^+ P_L b + i\kappa_a^T \frac{m_T}{v} \bar{T} a P_L t + \mathrm{H.c.}\right), \quad (13)$$

$$\mathcal{L}_a = \frac{1}{2}(\partial_\mu a)(\partial^\mu a) - \frac{1}{2} m_a^2 a^2 + \frac{g_s^2 K_g^a}{16\pi^2 f_a} a G_{\mu\nu}^a \tilde{G}^{a\mu\nu} + \frac{e^2 K_\gamma^a}{16\pi^2 f_a} a A_{\mu\nu} \tilde{A}^{\mu\nu} + \frac{g^2 c_W^2 K_Z^a}{16\pi^2 f_a} a Z_{\mu\nu} \tilde{Z}^{\mu\nu} + \frac{e g c_W K_{Z\gamma}^a}{8\pi^2 f_a} a A_{\mu\nu} \tilde{Z}^{\mu\nu}. \quad (14)$$

$\sigma(pp \to Tj)$ is proportional to $(\kappa_W^T)^2$, and the decay width of T can be expressed as

$$\Gamma(T \to Wb) = (\kappa_W^T)^2 \frac{m_T^3 g^2}{64\pi m_W^2} \Gamma_W(m_T, m_W, m_b), \quad (15)$$

$$\Gamma(T \to ta) = (\kappa_a^T)^2 \frac{m_T^3 g^2}{64\pi m_W^2} \Gamma_a(m_T, m_a, m_t). \quad (16)$$

These kinematic functions are

$$\Gamma_W(m_T, m_W, m_b) = \lambda^{\frac{1}{2}}\left(1, \frac{m_b^2}{m_T^2}, \frac{m_W^2}{m_T^2}\right)\left[\left(1 - \frac{m_b^2}{m_T^2}\right)^2 + \frac{m_W^2}{m_T^2} - 2\frac{m_W^4}{m_T^4} + \frac{m_W^2 m_b^2}{m_T^4}\right],$$

$$\Gamma_a(m_T, m_a, m_t) = \frac{1}{2}\lambda^{\frac{1}{2}}\left(1, \frac{m_t^2}{m_T^2}, \frac{m_a^2}{m_T^2}\right)\left[1 + \frac{m_t^2}{m_T^2} - \frac{m_a^2}{m_T^2}\right], \quad (17)$$

with the phase-space function $\lambda(a, b, c)$ defined as

$$\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2bc. \quad (18)$$

For the pseudoscalar a, we actually only need to know the ratio between $\Gamma(a \to \gamma\gamma)$ and $\Gamma(a \to gg)$. This is because the strong coupling g_s is much larger than the

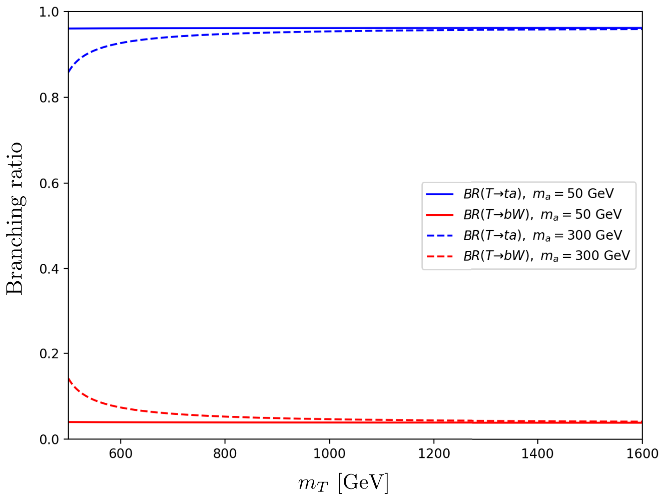

FIG. 11. $\mathrm{BR}(T \to ta)$ and $\mathrm{BR}(T \to bW)$ in Models 1 and 2 as functions of m_T, for different values of m_a.

electroweak coupling g, and thus the total decay width of a is approximately determined by $\Gamma(a \to gg)$. Then, $\mathrm{BR}(a \to \gamma\gamma)$ can be estimated by

$$\mathrm{BR}(a \to \gamma\gamma) \approx \frac{\Gamma(a \to \gamma\gamma)}{\Gamma(a \to gg)} = \frac{\alpha_{\mathrm{EM}}^2 (K_\gamma^a)^2}{8\alpha_S^2 (K_g^a)^2} \tag{19}$$

Here we treat (m_T, m_a) as undetermined parameters. K_Z^a, $K_{Z\gamma}^a$, K_W^a, and f_a are not relevant to our collider analysis. Limits from diboson resonance searches can be easily avoided by assuming that f_a has a large value. So, except for m_T and m_a, we only have four input parameters for a simplified model:

$$\kappa_W^T, \qquad \kappa_a^T, \qquad K_\gamma^a, \qquad K_g^a. \tag{20}$$

For illustration, we consider two different model settings:

Model 1 :

$$\kappa_W^T = 0.1, \quad \kappa_a^T = 0.5, \quad K_\gamma^a = 1.0, \quad K_g^a = 0.5; \tag{21}$$

Model 2 :

$$\kappa_W^T = 0.1, \quad \kappa_a^T = 0.5, \quad K_\gamma^a = 1.2, \quad K_g^a = 0.8. \tag{22}$$

The difference between these two models is the value of $\mathrm{BR}(a \to \gamma\gamma)$, which is 0.22% and 0.12% for Model 1 and Model 2, respectively. $\mathrm{BR}(T \to ta)$ and $\mathrm{BR}(T \to bW)$ are the same for the two models, but they depend on m_a and m_T. In Fig. 11 we present $\mathrm{BR}(T \to ta)$ and $\mathrm{BR}(T \to bW)$ as functions of m_T, with m_a fixed to 50 and 300 GeV. It can be seen that in most of the parameter space, $\mathrm{BR}(T \to bW)$ is smaller than 5%, and thus the current direct search bound can be escaped [17].

After calculating $\mathrm{BR}(T \to at)$ and $\mathrm{BR}(a \to \gamma\gamma)$ from these couplings and spectra, we can compare $\sigma(pp \to Tj) \times \mathrm{BR}(T \to at) \times \mathrm{BR}(a \to \gamma\gamma)$ and $\sigma(pp \to T\bar{T}) \times [\mathrm{BR}(T \to at)]^2 \times \mathrm{BR}(a \to \gamma\gamma)$ for each (m_T, m_a) with the upper limit presented in Figs. 5 and 8. Then, the 2σ exclusion limit in the m_T-m_a plane can be obtained. Figure 12 shows the exclusion limits in the m_T-m_a plane for both models. It shows that if $\mathrm{BR}(T \to ta) \approx 1$ and $\mathrm{BR}(a \to \gamma\gamma)$ is around $\mathcal{O}(0.1\%)$, we can exclude m_T up to the TeV scale at the future high-luminosity LHC. The detection ability of single production is more sensitive to the value of $\mathrm{BR}(a \to \gamma\gamma)$. The sensitivity of the single production channel is greatly enhanced when $\mathrm{BR}(a \to \gamma\gamma)$ increases from 0.12% to 0.22%. Our results also show that the single production channel becomes more sensitive when the pNGB a becomes lighter. On the contrary, the pair production search channel is more robust against the values of $\mathrm{BR}(a \to \gamma\gamma)$ and m_a.

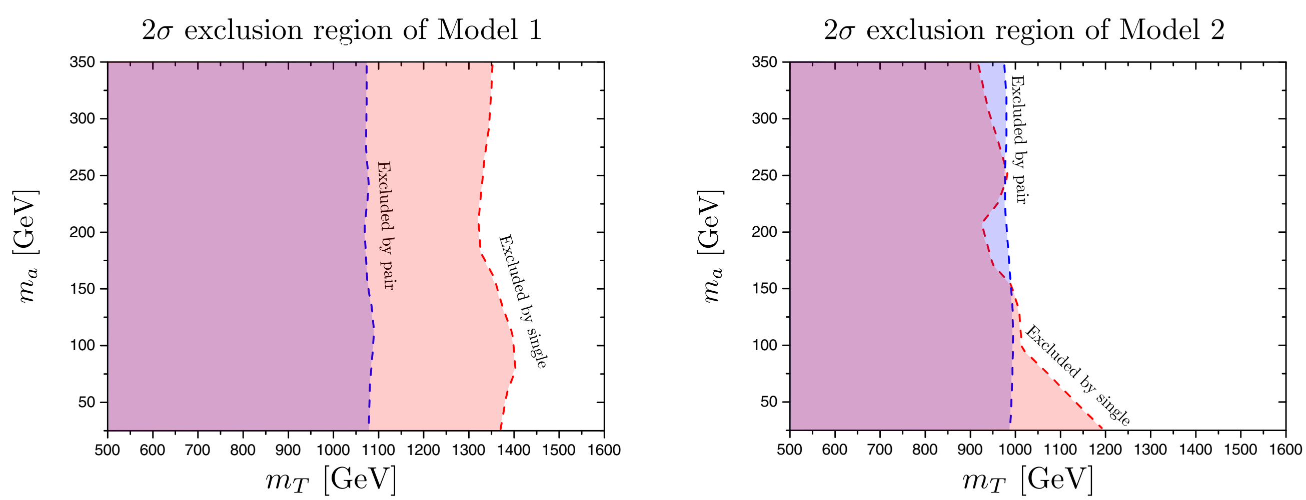

FIG. 12. Left: 2σ exclusion region in the m_T-m_a plane for Model 1. Left: 2σ exclusion region in the m_T-m_a plane for Model 2. Here we only show the exclusion limits for 3000 fb^{-1} at the 14 TeV LHC.

TABLE III. Acceptance for single and pair production processes after basic selection for the single production process. m_a is fixed to 150 GeV in this table.

m_T (GeV)	500	600	700	800	900	1000	1100	1200	1300	1400	1500	1600
Acceptance for single production process	1.39%	1.59%	1.83%	1.94%	2.09%	2.21%	2.31%	2.37%	2.61%	2.71%	2.88%	2.99%
Acceptance for pair production process	0.029%	0.024%	0.052%	0.068%	0.10%	0.14%	0.25%	0.13%	0.10%	0.21%	0.19%	0.15%

V. CONCLUSIONS

The vector-like top partner and the pseudo-Nambu-Goldstone boson are two key features of composite Higgs models. In this paper, we studied the observability of a new signature of the top partner decaying to a pNGB, $T \to ta$, through the production processes $pp \to Tj$ and $pp \to T\bar{T}$ at the LHC. We found that the clean decay channel of pNGB, $a \to \gamma\gamma$, helps to suppress the huge QCD background, and even the branching ratio of $a \to \gamma\gamma$ is as small as $\mathcal{O}(0.1\%)$. Model-independent exclusion limits for single and pair production were presented, and can be easily applied to concrete model. We also compared the direct search sensitivity of the single and pair production channels. We found that the single production process $pp \to Tj$ can be more sensitive than the conventional pair production process $pp \to \bar{T}T$ when m_T is larger than 800–900 GeV.

ACKNOWLEDGMENTS

Part of this work was done while the authors were visiting Jinan University. This work was supported by the National Natural Science Foundation of China (NNSFC) under Grants No. 11947118 and 11705093.

APPENDIX: CONTAMINATION OF PAIR PRODUCTION PROCESS TO SINGLE PRODUCTION PROCESS

One of the main purposes of this work is to compare the detection ability of the single production and pair production processes. Thus, we hope that the basic selection criteria used for single and pair production can separate the events from these two processes.

The basic selection criteria for the pair production process can exclude the events from the single production process, because two b jets and two leptons in the final state cannot be faked by the single production process. But the basic selection criteria for the single production process cannot fully exclude the events from the pair production process if one W boson in the pair production process $T\bar{T} \to t\bar{t}aa \to WWbbaa$ decays hadronically and one b jet is mistagged as a normal jet. So we need to know the acceptance for both the single and pair production processes after we apply the following basic selection criteria:

(1) Exactly one b jet with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$.
(2) Exactly one isolated lepton (electron or muon) with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$.
(3) Exactly two isolated photons with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$.
(4) Exactly one jet with $p_{\mathrm{T}} > m_T/8$ and $|\eta| < 2.5$. This jet can not be tagged as b jet.

In Table III we present the acceptance for both the single and pair production processes as m_T varies from 500 to 1600 GeV, and m_a is fixed to 150 GeV. We find that the acceptance does not change much if we change the value of m_a. Table III and the production cross section plot in Fig. 2 indicate that the contamination of the pair production process to the single production process can be ignored when $m_T \gtrsim 600$ GeV and $\kappa_W^T \gtrsim 0.1$.

[1] G. Aad *et al.* (ATLAS Collaboration), Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. B **716**, 1 (2012).
[2] S. Chatrchyan *et al.* (CMS Collaboration), Observation of a new boson at a mass of 125 GeV with the CMS Experiment at the LHC, Phys. Lett. B **716**, 30 (2012).
[3] D. B. Kaplan and H. Georgi, $SU(2) \times U(1)$ breaking by vacuum misalignment, Phys. Lett. **136B**, 183 (1984).
[4] D. B. Kaplan, Flavor at SSC energies: A new mechanism for dynamically generated fermion masses, Nucl. Phys. **B365**, 259 (1991).
[5] M. Redi and A. Tesi, Implications of a light Higgs in composite models, J. High Energy Phys. 10 (2012) 166.
[6] R. Contino, L. Da Rold, and A. Pomarol, Light custodians in natural composite Higgs models, Phys. Rev. D **75**, 055014 (2007).

[7] O. Matsedonskyi, G. Panico, and A. Wulzer, Light top partners for a light composite Higgs, J. High Energy Phys. 01 (2013) 164.
[8] D. Marzocca, M. Serone, and J. Shu, General composite Higgs models, J. High Energy Phys. 08 (2012) 013.
[9] S. Blasi and F. Goertz, Softened Symmetry Breaking in Composite Higgs Models, Phys. Rev. Lett. **123**, 221801 (2019).
[10] S. Blasi, C. Csaki, and F. Goertz, A natural composite Higgs via universal boundary conditions, SciPost Phys. **10**, 121 (2021).
[11] K. Agashe, R. Contino, and A. Pomarol, The minimal composite Higgs model, Nucl. Phys. **B719**, 165 (2005).
[12] A. D. Simone, O. Matsedonskyi, R. Rattazzi, and A. Wulzer, A first top partner Hunter's guide, J. High Energy Phys. 04 (2013) 004.
[13] G. Cacciapaglia, A. Deandrea, D. Harada, and Y. Okada, Bounds and decays of new heavy vector-like top partners, J. High Energy Phys. 11 (**2010**) 159.
[14] M. Aaboud *et al.* (ATLAS Collaboration), Search for pair production of up-type vector-like quarks and for four-top-quark events in final states with multiple b-jets with the ATLAS detector, J. High Energy Phys. 07 (2018) 089.
[15] M. Aaboud *et al.* (ATLAS Collaboration), Search for pair- and single-production of vector-like quarks in final states with at least one Z boson decaying into a pair of electrons or muons in pp collision data collected with the ATLAS detector at $\sqrt{s} = 13$ TeV, Phys. Rev. D **98**, 112010 (2018).
[16] M. Aaboud *et al.* (ATLAS Collaboration), Search for pair production of vector-like top quarks in events with one lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector, J. High Energy Phys. 08 (2017) 052.
[17] M. Aaboud *et al.* (ATLAS Collaboration), Search for pair production of heavy vector-like quarks decaying to high-p_T W bosons and b quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 10 (**2017**) 141.
[18] M. Aaboud *et al.* (ATLAS Collaboration), Search for pair production of heavy vector-like quarks decaying into hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Rev. D **98**, 092005 (2018).
[19] M. Aaboud *et al.* (ATLAS Collaboration), Search for pair production of heavy vector-like quarks decaying into high-p_T W bosons and top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, J. High Energy Phys. 08 (2018) 048.
[20] M. Aaboud *et al.* (ATLAS Collaboration), Combination of the Searches for Pair-Produced Vector-Like Partners of the Third-Generation Quarks at $\sqrt{s} = 13$ TeV with the ATLAS Detector, Phys. Rev. Lett. **121**, 211801 (2018).
[21] A. M. Sirunyan *et al.* (CMS Collaboration), Search for vector-like T and B quark pairs in final states with leptons at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 08 (2018) 177.
[22] A. M. Sirunyan *et al.* (CMS Collaboration), Search for pair production of vector-like quarks in the $bW\bar{b}W$ channel from proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B **779**, 82 (2018).
[23] A. M. Sirunyan *et al.* (CMS Collaboration), Search for pair production of vectorlike quarks in the fully hadronic final state, Phys. Rev. D **100**, 072001 (2019).
[24] A. Anandakrishnan, J. H. Collins, M. Farina, E. Kuflik, and M. Perelstein, Odd top partners at the LHC, Phys. Rev. D **93**, 075009 (2016).
[25] S. Kraml, U. Laa, L. Panizzi, and H. Prager, Scalar versus fermionic top partner interpretations of $t\bar{t} + E_T^{\mathrm{miss}}$ searches at the LHC, J. High Energy Phys. 11 (2016) 107.
[26] N. Bizot, G. Cacciapaglia, and T. Flacke, Common exotic decays of top partners, J. High Energy Phys. 06 (2018) 065.
[27] J. Serra, Beyond the minimal top partner decay, J. High Energy Phys. 09 (2015) 176.
[28] H. Han, L. Huang, T. Ma, J. Shu, T. M. P. Tait, and Y. Wu, Six top messages of new physics at the LHC, J. High Energy Phys. 10 (2019) 008.
[29] H. Alhazmi, J. H. Kim, K. Kong, and I. M. Lewis, Shedding light on top partner at the LHC, J. High Energy Phys. 01 (2019) 139.
[30] R. Benbrik, E. B. Kuutmann, D. Buarque Franzosi, V. Ellajosyula, R. Enberg, G. Ferretti, M. Isacson, Y. B. Liu, T. Mandal, T. Mathisen *et al.*, Signatures of vector-like top partners decaying into new neutral scalar or pseudoscalar bosons, J. High Energy Phys. 05 (2020) 028.
[31] J. A. Aguilar-Saavedra, D. E. Lpez-Fogliani, and C. Muoz, Novel signatures for vector-like quarks, J. High Energy Phys. 06 (2017) 095.
[32] M. Chala, Direct bounds on heavy toplike quarks with standard and exotic decays, Phys. Rev. D **96**, 015028 (2017).
[33] G. Cacciapaglia, T. Flacke, M. Park, and M. Zhang, Exotic decays of top partners: Mind the search gap, Phys. Lett. B **798**, 135015 (2019).
[34] J. H. Kim and I. M. Lewis, Loop induced single top partner production and decay at the LHC, J. High Energy Phys. 05 (2018) 095.
[35] K. P. Xie, G. Cacciapaglia, and T. Flacke, Exotic decays of top partners with charge 5/3: Bounds and opportunities, J. High Energy Phys. 10 (2019) 134.
[36] J. C. Criado and M. Perez-Victoria, Vector-like quarks with non-renormalizable interactions, J. High Energy Phys. 01 (2020) 057.
[37] J. A. Aguilar-Saavedra, J. Alonso-Gonzlez, L. Merlo, and J. M. No, Exotic vectorlike quark phenomenology in the minimal linear σ model, Phys. Rev. D **101**, 035015 (2020).
[38] M. Ramos, Composite dark matter phenomenology in the presence of lighter degrees of freedom, J. High Energy Phys. 07 (2020) 128.
[39] J. H. Kim, S. D. Lane, H. S. Lee, I. M. Lewis, and M. Sullivan, Searching for dark photons with maverick top partners, Phys. Rev. D **101**, 035041 (2020).
[40] J. Barnard, T. Gherghetta, and T. S. Ray, UV descriptions of composite Higgs models without elementary scalars, J. High Energy Phys. 02 (2014) 002.
[41] G. Ferretti and D. Karateev, Fermionic UV completions of composite Higgs models, J. High Energy Phys. 03 (2014) 077.
[42] G. Ferretti, Gauge theories of partial compositeness: Scenarios for Run-II of the LHC, J. High Energy Phys. 06 (2016) 107.

[43] T. DeGrand, M. Golterman, E. T. Neil, and Y. Shamir, One-loop chiral perturbation theory with two fermion representations, Phys. Rev. D **94**, 025020 (2016).
[44] A. Belyaev, G. Cacciapaglia, H. Cai, G. Ferretti, T. Flacke, A. Parolini, and H. Serodio, Di-boson signatures as standard candles for partial compositeness, J. High Energy Phys. 01 (2017) 094; Erratum, J. High Energy Phys. 12 (2017) 088.
[45] G. Cacciapaglia, G. Ferretti, T. Flacke, and H. Serdio, Light scalars in composite Higgs models, Front. Phys. **7**, 22 (2019).
[46] B. P. Roe, H. J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, Boosted decision trees, an alternative to artificial neural networks, Nucl. Instrum. Methods Phys. Res., Sect. A **543**, 577 (2005).
[47] M. R. Crispim, N. F. Castro, R. Pedro, and T. Vale, Transferability of deep learning models in searches for new physics at colliders, Phys. Rev. D **101**, 035042 (2020).
[48] J. P. Araque, N. F. Castro, and J. Santiago, Interpretation of vector-like quark searches: Heavy gluons in composite Higgs models, J. High Energy Phys. 11 (2015) 120.
[49] S. Dasgupta, S. K. Rai, and T. S. Ray, Impact of a colored vector resonance on the collider constraints for top-like top partner, Phys. Rev. D **102**, 115014 (2020).
[50] A. Atre, G. Azuelos, M. Carena, T. Han, E. Ozcan, J. Santiago, and G. Unel, Model-independent searches for new quarks at the LHC, J. High Energy Phys. 08 (2011) 080.
[51] A. Atre, M. Chala, and J. Santiago, Searches for new vector like quarks: Higgs channels, J. High Energy Phys. 05 (2013) 099.
[52] S. S. D. Willenbrock and D. A. Dicus, Production of heavy quarks from W gluon fusion, Phys. Rev. D **34**, 155 (1986).
[53] L. Wu, Enhancing thj production from top-Higgs FCNC couplings, J. High Energy Phys. 02 (2015) 061.
[54] A. Roy, N. Nikiforou, N. Castro, and T. Andeen, Novel interpretation strategy for searches of singly produced vector-like quarks at the LHC, Phys. Rev. D **101**, 115027 (2020).
[55] R. Contino and A. Pomarol, Holography for fermions, J. High Energy Phys. 11 (**2004**) 058.
[56] G. Cacciapaglia, G. Marandella, and J. Terning, Dimensions of supersymmetric operators from AdS/CFT, J. High Energy Phys. 06 (2009) 027.
[57] M. Schmaltz, Physics beyond the Standard Model (theory): Introducing the little Higgs, Nucl. Phys. B, Proc. Suppl. **117**, 40-49 (2003).
[58] R. Dermek, E. Lunghi, and S. Shin, Hunting for vectorlike quarks, J. High Energy Phys. 04 (2019) 019.
[59] M. Buchkremer, G. Cacciapaglia, A. Deandrea, and L. Panizzi, Model independent framework for searches of top partners, Nucl. Phys. **B876**, 376 (2013).
[60] R. Godbole, M. Guchait, C. K. Khosa, J. Lahiri, S. Sharma, and A. H. Vijay, Boosted top quark polarization, Phys. Rev. D **100**, 056010 (2019).
[61] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, MadGraph 5: Going beyond, J. High Energy Phys. 06 (2011) 128.
[62] G. Cacciapaglia, A. Carvalho, A. Deandrea, T. Flacke, B. Fuks, D. Majumder, L. Panizzi, and H. S. Shao, Next-to-leading-order predictions for single vector-like quark production at the LHC, Phys. Lett. B **793**, 206 (2019).
[63] B. Fuks and H. S. Shao, QCD next-to-leading-order predictions matched to parton showers for vector-like quark models, Eur. Phys. J. C **77**, 135 (2017).
[64] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, UFO—The Universal FeynRules Output, Comput. Phys. Commun. **183**, 1201 (2012).
[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] T. Sjstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA 8.2, Comput. Phys. Commun. **191**, 159 (2015).
[67] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), DELPHES 3, A modular framework for fast simulation of a generic collider experiment, J. High Energy Phys. 02 (2014) 057.
[68] ATLAS Collaboration, Expected performance for an upgraded ATLAS detector at high-luminosity LHC, Report No. ATL-PHYS-PUB-2016-026.
[69] A. Hocker *et al.*, TMVA—Toolkit for multivariate data analysis, *Proc. Sci.*, ACAT2007 (2007) 040 [arXiv:physics/0703039].