Eur. Phys. J. C (2023) 83:413
https://doi.org/10.1140/epjc/s10052-023-11506-x

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Probing a Z′ with non-universal fermion couplings through top quark fusion, decays to bottom quarks, and machine learning techniques

Diego Barbosa[2], Felipe Díaz[2], Liliana Quintero[2], Andrés Flórez[2], Manuel Sanchez[2], Alfredo Gurrola[1,a], Elijah Sheridan[1], Francesco Romeo[1]

[1] Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA
[2] Physics Department, Universidad de los Andes, Bogotá, Colombia

Received: 27 January 2023 / Accepted: 12 April 2023

Abstract The production of heavy neutral mass resonances, Z′, has been widely studied theoretically and experimentally. Although the nature, mass, couplings, and associated quantum numbers of this hypothetical particle are yet to be determined, current LHC experimental results have set strong constraints assuming the simplest beyond Standard Model (SM) hypotheses. We present a new feasibility study on the production of a Z′ boson at the LHC, with family non-universal couplings, considering proton–proton collisions at $\sqrt{s} = 13$ and 14 TeV. Such a hypothesis is well motivated theoretically and it can explain observed differences between SM predictions and experimental results, as well as being a useful tool to further probe recent results in searches for new physics considering non-universal fermion couplings. We work under two simplified phenomenological frameworks where the Z′ masses and couplings to the SM particles are free parameters, and consider final states of the Z′ decaying to a pair of b quarks. The analysis is performed using machine learning techniques to maximize the sensitivity. Despite being a well motivated physics case in its own merit, such scenarios have not been fully considered in ongoing searches at the LHC. We note the proposed search methodology can be a key mode for discovery over a large mass range, including low masses, traditionally considered difficult due to experimental constrains. In addition, the proposed search is complementary to existing strategies.

1 Introduction

The standard model (SM) of particle physics is a successful theory to explain a plethora of experimental observations involving weak, electromagnetic, and strong interactions over the last few decades. However, as experiments probe new questions and increasing energies, observations indicate the SM is incomplete and might be a low-energy remnant of a more complete theory. There are a multitude of theoretical models proposed to overcome the SM limitations. Although the initial motivations and resulting implications of these models can vary, a common characteristic is the manifestation of new particles that can be probed in proton–proton (pp) collisions at the CERN's Large Hadron Collider (LHC).

Numerous ideas have been proposed to probe physics beyond the SM, motivating a large volume of searches at the LHC. Nonetheless, extensive searches have found no firm indication of new phenomena, largely constraining theories and setting exclusion limits up to multi-TeV on the masses of new particles predicted by those theories [1–8]. Possible explanations for the lack of evidence point to either new particles being too massive or having too low a production rate in existing colliders, or new physics having different features compared to what is traditionally assumed in many beyond SM theories and searches, thus remaining concealed in processes not yet investigated. In particular, many searches conducted so far at the LHC rely heavily on the assumption that these hypothesized new particles have similar couplings to all generations of fermions, including couplings to the partons inside the proton, thus favoring LHC production modes through light quarks. Therefore, if new phenomena are within the reach of the LHC, both in energy and production rate, they might manifest with different features compared to what is assumed in searches at high energy colliders, thus requiring new efforts and experimental quests.

In this paper, we consider a different scenario in which new particles have non-universal fermion couplings, favoring

[a] e-mail: Alfredo.Gurrola@Vanderbilt.Edu (corresponding author)

Published online: 16 May 2023

higher-generation fermions, which we refer to as *anogenophilic* particles. In particular, we consider a new neutral vector gauge boson, Z', with only couplings to third generation fermions, referred to as *tritogenophilic*. This physics case is also interesting theoretically and because of recent results in precision measurements, offering a new physics phase space not yet fully explored at the LHC.

An anogenophilic resonance is predicted in several theories that extend the SM and in different contexts such as in minimal $U(1)_X$ extensions [9], top-assisted technicolor models [10], Randall–Sundrum models with Kaluza–Klein excitations of the graviton [11–13], two-Higgs doublet models that address the naturalness of the electroweak symmetry breaking scale [14–17], left-right extensions of the SM [18], models with a color-sextet or color-octet [19–24], with composite particles [25–38], or with dark matter mediators [39–45].

Besides well-motivated theoretical extensions of the SM, it is interesting to speculate how new physics with non-equal fermion couplings may reconcile with some experimental results from precision measurements that, although not yet confirmed, appear to indicate that such a hypothesis is viable and worth exploring at the energies accessible at the LHC. In this perspective, recent results seem to confirm long-standing tensions with the SM in the measurement of the muon anomalous magnetic moment [46] and measurements related to the branching ratios R_D, $R_{D^{(*)}}$ [47–55], all showing significant deviations from the SM expectation [46,56]. The scenario of a tritogenophilic new particle is also interesting in light of recent results of the measured cross sections for $pp \to t\bar{t} + b\bar{b}$ ($t\bar{t}t\bar{t}$) production from the ATLAS [57,58] [59] and CMS [60,61] collaborations, which are found to be higher than the expectations from the SM.

Lastly, direct searches for additional Higgs bosons have recently reported some tensions with the SM expectations at about 95 GeV in the diphoton [62] final state, and 100 GeV in the ditau [63] final state, which are consistent with each other within the resolution of the reconstructed invariant mass of the $\tau\tau$ system, and that urge new experimental quests for particles coupling to higher generation fermions (such as top and bottom quarks investigated in this work). Interestingly, the analysis in [63] also reports a deviation at the TeV scale in a search for leptoquarks coupling to higher generation fermions, which also appears in a separate dedicated search at CMS [64].

The mass, quantum numbers, and couplings of new hypothetical mediators can be open parameters to be determined experimentally, making the new physics phase space broadly defined. Thus, initial ATLAS/CMS searches for these new type of particles were conducted considering models with democratic couplings to all fermion families, and focused on Drell–Yan production mechanisms with light quarks (e.g., $q\bar{q} \to Z'$), and final states with muons and electrons with high signal acceptance and a narrow "bump" in the reconstructed invariant mass spectrum of lepton pairs sitting above a smooth and steeply falling background distribution [65,66]. However, from the phenomenological point of view, when couplings to light quarks are suppressed in pp colliders, relative to higher-generation fermions, new production mechanisms become dominant to generate and discover beyond SM resonances. They are produced in association with other SM particles and give origin to rare and peculiar signatures. The phenomenology of purely top-philic Z' [67–69] scenarios, as well as models with a Z' that couples to top quarks and tau/muon [70,71] leptons, have already been studied in the literature. Furthermore, a CMS search has been performed for a neutral resonance coupling to top quarks and decaying to muons or electrons [72].

In this paper, we perform a previously unexamined feasibility study on the production of a more general tritogenophilic Z' produced through the fusion of a $t\bar{t}$ pair ($t\bar{t}Z'$) and decaying to a pair of b quarks ($Z' \to b\bar{b}$), as in Fig. 1. We consider the final state where one of the two remaining tops from the fusion process, referred to as *spectator* top quarks, decays to bW and the W boson subsequenly decays to an electron or muon plus its neutrino. Such a choice balances the lower $W \to \ell\nu$ branching fraction compared to W boson decays into two quarks, with a cleaner final state. This has the double advantage of mitigating the large background from full-hadronic SM quantum chromodynamics (QCD) processes, and of overcoming the otherwise overwhelming events rate that is outside the typical trigger bandwidth at the LHC, rendering the search sensitive to a wide range of

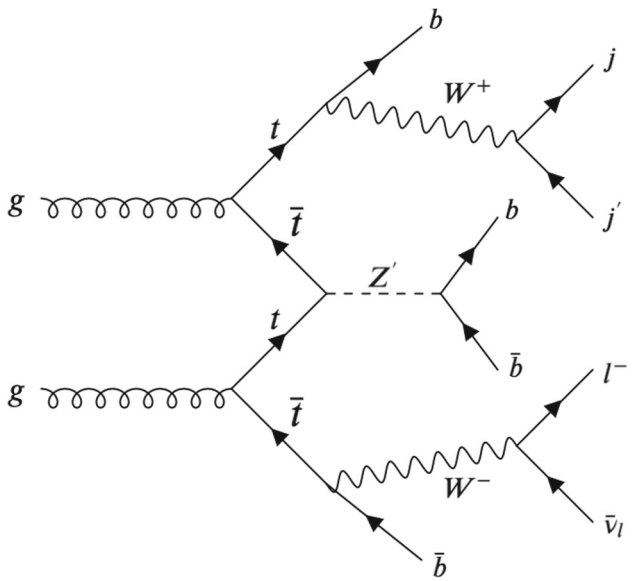

Fig. 1 Representative Feynman diagram for the production of a Z' boson through the fusion of a top quark pair, where the Z' decays to a pair of bottom quarks and the two spectator top quarks decay semi-leptonically

Z' masses. For Z' masses below the $2m_t$ kinematic production threshold, where $Z' \to t\bar{t}$ decays are not permitted, the Z' decay to $b\bar{b}$ is the dominant discovery mode. Furthermore, the analysis strategy proposed in this paper provides enhanced sensitivity compared to other approaches already used in searches at the LHC [1,4,73–75]. Above $2m_t$, the reduced jet multiplicity of the $Z' \to b\bar{b}$ final state, in comparison to $Z' \to t\bar{t}$, favors the experimental reconstruction of the Z' mass. In this work, machine learning techniques are used to maximize the experimental sensitivity.

2 Samples and simulation

Signal and background samples are generated with MadGraph5_aMC (v2.6.3.2) [76] considering pp beams colliding with a center-of-mass energy of $\sqrt{s} = 13$ TeV and $\sqrt{s} = 14$ TeV. All samples are generated using the NNPDF3.0 NLO [77] set for parton distribution functions (PDFs). Parton level events are then interfaced with the PYTHIA (v8.2.05) [78] package to include parton fragmentation and hadronization processes, while DELPHES (v3.4.1) [79] is used to simulate detector effects, using the CMS detector geometric configurations and parameters, for performance of particle reconstruction and identification. At parton level, jets are required to have a minimum transverse momentum (p_T) of 20 GeV and pseudorapidity (η) $|\eta| < 5.0$. The cross sections in this paper are obtained with the aforementioned parton-level selections. The MLM algorithm [80] is used for jet matching and jet merging. The xqcut and qcut variables of the MLM algorithm, related with the minimal distance between partons and the energy spread of the clustered jets, are set to 30 and 45, respectively, as a result of an optimization process requiring the continuity of the differential jet rate as a function of jet multiplicity.

The signal samples are generated considering the production of a Z' and two associated top quarks ($pp \to Z't\bar{t}$), inclusive in α and α_s. For our benchmark signal scenario, we consider the simplified model in Ref. [81] where the Z' masses and couplings to the SM particles are free parameters, and defined as variations of the SM Z boson couplings (i.e., variations of the so-called Sequential Standard Model, SeqSM). The Z' coupling to the first and second generation SM quarks is defined as $g_{Z'q\bar{q}} = g_q \times g_{Zq\bar{q}}$, where $g_{Zq\bar{q}}$ is the SM Z boson coupling to first and second generation quarks and g_q is a "modifier" for the coupling. Similarly, the Z' coupling to the third generation SM quarks is defined as $g_{Z',b/t,\bar{b}/\bar{t}} \times g_{Z,b/t,\bar{b}/\bar{t}}$, where $g_{Z',b/t,\bar{b}/\bar{t}}$ is the modifier to the SeqSM coupling. We refer to this model as "simplified phenomenological model 1" (SPM1). In all cases considered, the modifiers for the Z' couplings to $t\bar{t}$ and $b\bar{b}$ are equal to each other, and thus for simplicity we henceforth refer to those modifiers as $g_{Z't\bar{t}}$. Therefore, a scenario with $g_{Z't\bar{t}} = 1$ has similar Z' couplings to top/bottom quarks as the SeqSM. Signal samples were created for $m(Z')$ ranging from 250 to 2000 GeV. Table 1 lists the production cross sections for different Z' masses, considering pp collisions at $\sqrt{s} = 13$ TeV and 14 TeV, and for two representative g_q coupling scenarios with $g_{Z't\bar{t}} = 1$. The $g_q = 0$ case is a proxy for the tritogenophilic scenarios, where the couplings of the Z' to light quarks are suppressed. The $g_q = 1$ case allows for non-negligible couplings to light quarks, and thus other $t\bar{t}Z'$ production processes can contribute, such as initial state radiation of a Z' from a light quark.

Table 1 Signal cross sections, calculated with MadGraph, for different Z' masses and couplings to first and second generation quarks. The values in this table are calculated with $g_{Z't\bar{t}} = 1$

Z' mass (GeV)	13 TeV		14 TeV	
	$\sigma_{g_q=0}$(fb)	$\sigma_{g_q=1}$(fb)	$\sigma_{g_q=0}$(fb)	$\sigma_{g_q=1}$(fb)
250	51.34	72.87	64.32	90.21
300	32.61	47.73	41.22	58.94
325	26.40	39.47	33.48	48.33
350	21.58	32.84	27.14	40.53
375	17.71	27.48	22.50	33.77
400	14.58	23.14	18.97	28.62
500	7.379	12.53	9.557	15.67
750	1.700	3.546	2.315	4.516
1000	0.502	1.285	0.703	1.681
2000	0.011	0.066	0.017	0.093

In addition to our primary signal benchmark model described above, we also consider a tritogenophilic scenario where the Z' is a color singlet vector particle whose effective couplings are not suppressed by factors of the electroweak mixing angles (as in the SeqSM) and whose relevant interactions to top/bottom quarks are given by the following renormalizable Lagrangian: $\mathcal{L}_{\text{int}} = \bar{t}\gamma_\mu(c_L P_L + c_R P_R)tZ'^\mu = c_{\text{eff}}\bar{t}\gamma_\mu(\cos\theta P_L + \sin\theta P_R)tZ'^\mu$, where $P_{R/L} = (1 \pm \gamma_5)/2$ are the projection operators, $c_{\text{eff}} = \sqrt{c_L^2 + c_R^2}$ is the Z' coupling to top/bottom quarks, and $\tan\theta = c_R/c_L$ is the tangent of the chirality angle. We consider the case where the Z' couplings to top and bottom quarks are equal to each other, and thus for simplicity we henceforth refer to those couplings as c_t. This type of simplified model, which we refer to as "simplified phenomenological model 2" (SPM2), has been studied in Refs. [67–69], and it has been shown that $t\bar{t}Z'$ production is independent of θ. We have checked that this is indeed the case. Thus, we only consider $\theta = \pi/2$. Although the signal kinematic distributions for this particular model are similar to those of SPM1, the $t\bar{t}Z'$ production cross sections for SPM2 are larger than those of SPM1, when $c_t = g_{Z't\bar{t}}$, since the SPM2 Lagrangian does not contain suppression terms from the electroweak mixing angles. Our primary motivation in

Table 2 Cross sections calculated with MadGraph for the dominant background processes

Process	13 TeV (fb)	14 TeV (fb)
$t\bar{t}h$	393.5	476.3
$t\bar{t}X$	13,600	16,620
$t\bar{t}t\bar{t}$	8.973	11.80

using SPM2 is to compare the projected discovery reach of the proposed analysis strategy in this paper, with other strategies, such as those in Ref. [68], which considers the $Z' \to t\bar{t}$ decay mode.

For each signal model (SPM1 and SPM2 with $q_q = 0$ and 1), we generate signal samples for Z' mass values between 250 GeV and 2000 GeV, in steps of 25 GeV between 250 and 500 GeV, and steps of 250 GeV between 500 and 2000 GeV. The considered Z'-$t\bar{t}$ coupling values are between 1 and 4, in steps of 0.5. In total there are 476 $\{g_q, g_{Z't\bar{t}}, m(Z'), \text{SPM model}\}$ signal scenarios simulated, and for each of these scenarios two sets of samples are generated, each with one million simulated events, which are used separately for the training and testing of the machine learning algorithm.

Several sources of background are considered for our studies, including production of top quark pairs ($t\bar{t}$), Z/W bosons with associated jets (V+jets), QCD multijet, associated production of a Higgs (h) or a Z/γ^* boson from $t\bar{t}$ fusion processes (denoted $t\bar{t}h$ and $t\bar{t}X$), and associated production of four t quarks ($t\bar{t}t\bar{t}$). Since our signal topology targets final states with four bottom quarks ($Z' \to b\bar{b}$ and $t\bar{t} \to bWbW$), the $t\bar{t}$, V+jets, and QCD multijet backgrounds do not meaningfully contribute to our studies ($\ll 1\%$ of the total background). The $t\bar{t}h$, $t\bar{t}X$, and $t\bar{t}t\bar{t}$ processes are the dominant sources of background events. The $tt\bar{h}$ and $tt\bar{X}$ processes become important backgrounds when h and Z/γ^* decay to a pair of bottom quarks. Table 2 shows the production cross sections for the dominant backgrounds, at $\sqrt{s} = 13$ TeV and 14 TeV.

The total event rates are determined using $N = \sigma \times L \times \epsilon$, where N represents the total yield of events, L the integrated luminosity considered (for this study, 150 fb^{-1}, 300 fb^{-1}, and 3000 fb^{-1}), and ϵ represents any efficiencies which might reduce the total event yield (e.g., particle identification efficiencies). The $L = 150$ fb^{-1} scenario represents an estimate for the amount of data already collected by the ATLAS and CMS experiments, while the other luminosity scenarios are the expectations for the next decade of pp data taking at the LHC. All production cross sections are computed at tree level. Since the k-factors associated with higher-order corrections to QCD production cross sections are typically greater than one, our estimates of the sensitivity are conservative.

Following Ref. [82], we consider three possible "working points" for the identification of the b-jet candidates in DELPHES: (i) the "Loose" working point of the DeepCSV algorithm, which gives a 85% b-tagging efficiency and 10% light quark mis-identification rate; (ii) the "Medium" working point of the DeepCSV algorithm, which gives a 70% b-tagging efficiency and 1% light quark mis-identification rate; and (iii) the "Tight" working point of the DeepCSV algorithm, which gives a 45% b-tagging efficiency and 0.1% light quark mis-identification rate. The choice of b-tagging working points is determined through an optimization process which maximizes discovery reach. The "Medium" working point was ultimately shown to provide the best sensitivity and therefore chosen for this study. For muons (electrons), the assumed identification efficiency is 95% (85%), with a 0.3% (0.6%) mis-identification rate [83–85].

3 Data analysis using the gradient boost algorithm

The analysis of signal and background events is performed using a machine learning event classifier, namely a gradient boosted decision trees (BDTs) [86]. Machine learning offers advantages over traditional event classification methods. In particular, machine learning models consider all kinematic variables in tandem, efficiently traversing the high-dimensional space of event kinematics, thereby enabling them to enact complicated selection criteria which incorporates that high-dimensional space in its entirety.

This method iteratively trains decision trees to learn the residuals between predictions and expected values yielded by the tree trained just before it, thereby greedily minimizing error at each iteration. BDTs have been employed to great effect previously in classification problems arising in collider physics (e.g., [87–93]). We note that although neural networks have also been successfully used for similar tasks, e.g., in Refs. [94,95], the complex nature of the studies in this work (particle objects considered, experimental constraints in a high luminosity LHC, etc.) motivate the use of a BDT because of its usefulness, efficiency, and simplicity in understanding the machine learning output and underlying nature of the samples being analyzed.

Simulated signal and background events are initially filtered, before being passed to the BDT algorithm, requiring at least four well reconstructed and identified b-jet candidates, at least two jets not tagged as b jets, and exactly one identified light lepton (ℓ), that could be either an electron (e) or a muon (μ). The filtering selections are motivated by experimental constraints, such as the geometric constraints of the CMS/ATLAS detectors, the typical kinematic thresholds for reconstruction of particle objects, and the available lepton triggers which also drive the minimal kinematic thresholds. Selected jets must have $p_T > 30$ GeV

Table 3 Preliminary event selection criteria used to filter events that are passed to the gradient boosting algorithm. A $\Delta R(p_i, p_j) > 0.3$ requirement is applied to all the particle candidate pairs p_i, p_j

Variable	Threshold
$p_T(j)$	> 30 GeV
$\lvert\eta(j)\rvert$	< 5.0
$\lvert\eta(\text{b jets})\rvert$	< 2.5
$p_T(\text{b jets})$	> 30 GeV
$N(\text{b jets})$	> 3
$N(\ell)$	$= 1$
$\lvert\eta(\ell)\rvert$	< 2.5
$p_T(\ell)$	> 25 GeV
$\Delta R(p_i, p_j)$	> 0.3

and $|\eta(j)| < 5.0$, while b-jet candidates with $p_T > 30$ GeV and $|\eta(b)| < 2.5$ are chosen. The ℓ object must pass a $p_T > 25$ GeV threshold and be within a $|\eta(\ell)| < 2.5$. Overlapping objects in $\eta - \phi$ space are removed using a minimum ΔR among all particle candidates (p_i) above 0.3, where $\Delta R(p_i, p_j) = \sqrt{(\Delta\phi(p_i, p_j))^2 + (\Delta\eta(p_i, p_j))^2}$. These filtering criteria will be henceforth referred to as pre-selections. The efficiency of the pre-selections depends on $m(Z')$, but is typically about 10%. Table 3 summarizes these pre-selections for the analysis.

Events passing this pre-selection are used as input for the BDT algorithm, which classifies them as signal or background, using a probability factor. We implement the BDT algorithm using the canonical `scikit-learn` [96] and `xgboost` [97] libraries. In particular, we employed the `XGBClassifier` class in the latter library with 250 iterations, a max depth of 7, a learning rate of 0.1, and default parameters otherwise, although we note that model performance was found to be largely independent of hyperparameters.

Figures 2, 3, 4 and 5, show relevant kinematic distributions for two SPM1 signal points and dominant backgrounds, normalized to the area under the curve (unity). The distributions correspond to the b-jet candidate with the highest p_T (b_1), the second b-jet candidate with the highest p_T (b_2), the ΔR separation between the b_1 and b_2 candidates, and the reconstructed mass between the b_1 and b_2, $m(b_1, b_2)$, respectively. These distributions are among the variables identified by the BDT algorithm with the highest signal to background discrimination power.

As can be seen from Figs. 2 and 3, for $m(Z')$ values beyond the electroweak scale, the relatively large leading and subleading b-jet p_T is a key feature attributed to the heavy Z' with respect to the mass of the bottom quarks, thus resulting in an average $p_T(b_{1,2})$ of approximately $m(Z')/2$. This kinematic feature provides a nice handle to discriminate high $m(Z')$ sig-

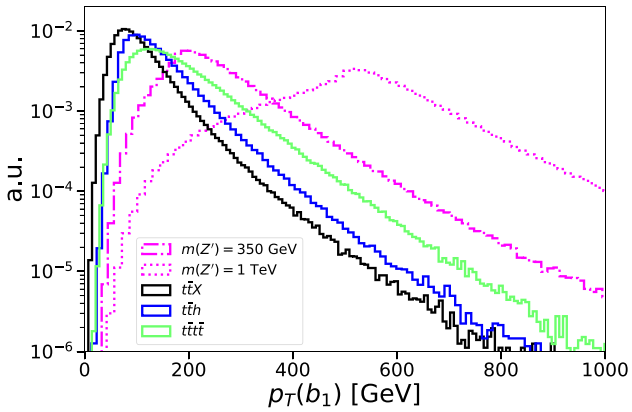

Fig. 2 Transverse momentum distributions for the b quark jet with the highest transverse momentum, for two signal points with masses of 350 GeV and 1000 GeV and dominant backgrounds

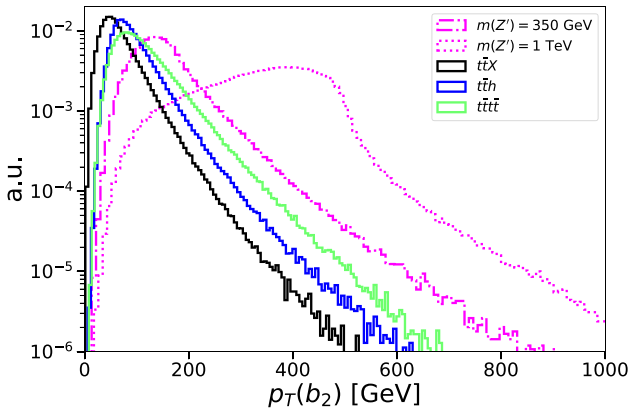

Fig. 3 Transverse momentum distributions for the b quark jet with the second highest transverse momentum, for two signal points with masses of 350 GeV and 1000 GeV and dominant backgrounds

nal events amongst the large SM backgrounds, which have lower average $p_T(b_{1,2})$ constrained by the top quark and/or higgs masses. The ΔR separation between b_1 and b_2 is determined by the amount of momentum transfer to the resonant particles in each process (Z', h, or t), which in turn depends on the masses of those particles. Therefore, Fig. 4 shows greater discrimination between background and signal processes as $m(Z')$ becomes larger. Finally, as noted previously, an advantage of the $Z' \to b\bar{b}$ final state in comparison to $Z' \to t\bar{t}$ is the experimental reconstruction of the Z' mass, which is observed as a peak in the $m(b_1, b_2)$ signal distributions in Fig. 5 near the true $m(Z')$ value. On the other hand, the background $m(b_1, b_2)$ distributions show a peak near $m(h) = 125$ GeV for the $t\bar{t}h$ background, or a broad distribution for the other backgrounds, indicative of the combination of two b jets from different decay vertices. We note that the $Z' \to b\bar{b}$ decay width depends on $g^2_{Z't\bar{t}} \times \frac{m_b^2}{m(Z')^2}$ and is thus suppressed by the relatively small bottom quark mass with respect to the $g_{Z't\bar{t}}$ and $m(Z')$ values considered in these studies. Therefore,

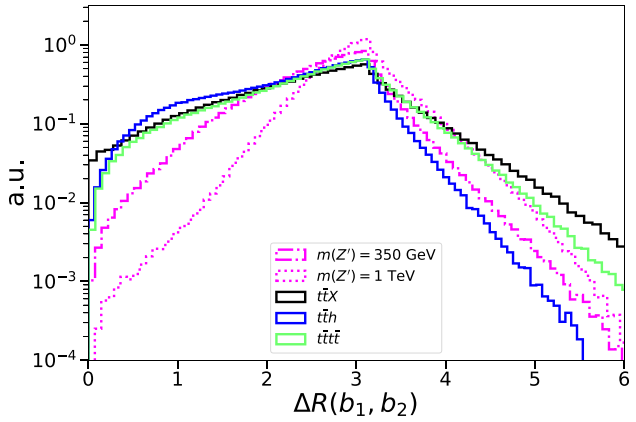

Fig. 4 Distributions for the ΔR angular separation between the the highest (b_1) and second highest (b_2) transverse momentum b quark pair, for two signal points with masses of 350 GeV and 1000 GeV and dominant backgrounds

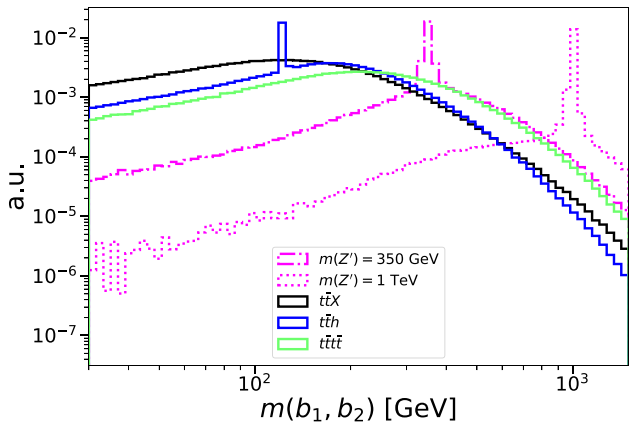

Fig. 5 Invariant mass distributions for the highest (b_1) and second highest (b_2) transverse momentum b quark pair, for two signal points with masses of 350 GeV and 1000 GeV and dominant backgrounds

the width of the $m(b_1, b_2)$ signal distributions is driven by the experimental resolution in the reconstruction of the b-jet momenta, as well as the probability that the two leading b jets are the correct pair from the Z′ decay.

In addition to these aforementioned variables in Figs. 2, 2, 4, 6 and 5, a variety of other kinematic variables were included as inputs to the BDT algorithm. In particular, 47 such variables were used in total, and these included the momenta of b and light quark jets (not tagged as b jets); invariant masses of pairs of b jets and of the two leading light jets; angular differences between b jets, between light quark jets, and between the lepton and b jets; and transverse masses derived from the lepton-p_{T}^{miss} pair and lepton-p_{T}^{miss}-b triplets. The variables $m(b_i, b_j)$ for $i, j \neq 1$ provide some additional discrimination between signal and background when the leading b-jets are not a Z′ decay candidate. The transverse mass variables are designed to be sensitive to a leptonic decay of the W boson and t quark (i.e., m_{jj} and

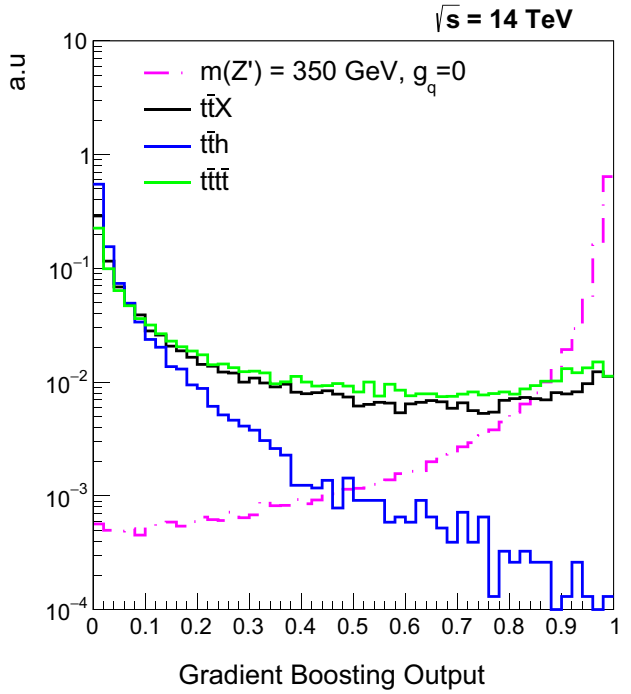

Fig. 6 Output of the gradient boosting algorithm for a Z′ signal with mass of 350 GeV and $g_q = 0$ coupling, and the dominant backgrounds. The distributions are normalized to unity

$m_{\mathrm{T}}(\ell, p_{\mathrm{T}}^{miss})$ should be near m_{W}, and $m_{\mathrm{T}}(\ell, \mathrm{b}, p_{\mathrm{T}}^{miss})$ near m_{t}), as this is an important feature in our signal (Fig. 1). A trained BDT can return the discriminating power of each of its inputs: we found that the plotted kinematic variables (i.e., $p_{\mathrm{T}}(b_1)$, $p_{\mathrm{T}}(b_2)$, $\Delta R(b_1, b_2)$, and $m(b_1, b_2)$) were among the most productive variables from this standpoint, producing about 60-75% of signal significance (depending on Z′ mass), but the inclusion of all 47 variables does provide a non-trivial enhancement.

Figure 6 shows the distributions for the output of the BDT algorithm for a SPM1 signal benchmark point with $m(Z') = 350$ GeV and $\{g_q, g_{Z't\bar{t}}\} = \{0, 1\}$, and the dominant backgrounds. Figure 7 shows the BDT output for $m(Z') = 500$ GeV and $\{g_q, g_{Z't\bar{t}}\} = \{1, 1\}$. The output of the BDT algorithm is a value between 0 and 1, which quantifies the likelihood that an event is either signal-like (BDT output near 1) or background-like (BDT output near 0). The distributions in Figs. 6 and 7 are normalized to an area under the curve of unity. Figure 8 shows the true positive rate (TPR), defined as the probability with which signal events are selected using the BDT output, as a function of the false positive rate (FPR), defined as the probability with which background events are selected. For example, for $m(Z') = 500$ GeV, when signal events are selected at 50% probability, the background is selected at 2×10^{-4} probability. Table 4 shows the expected total event rates for each process, for a particular choice of bin ranges of the BDT output, assuming an integrated lumi-

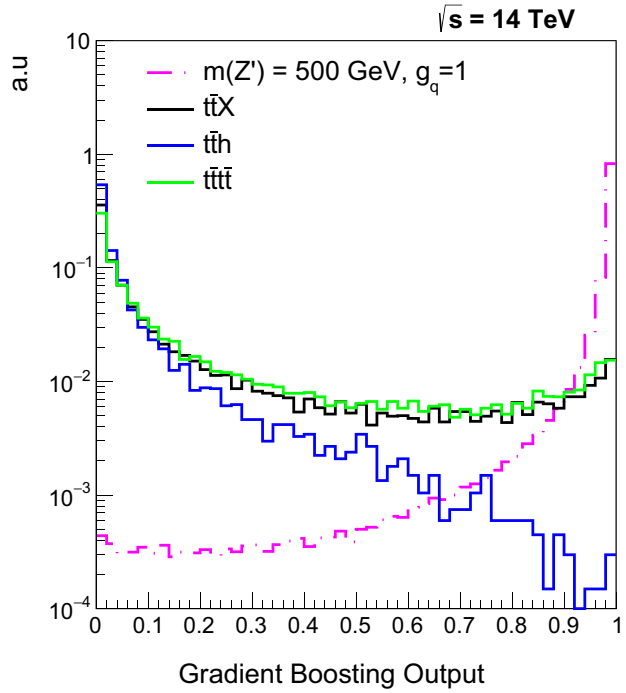

Fig. 7 Output of the gradient boosting algorithm for a Z′ signal with mass of 500 GeV and $g_q = 1$ coupling, and for the most relevant backgrounds. The distributions are normalized to unity

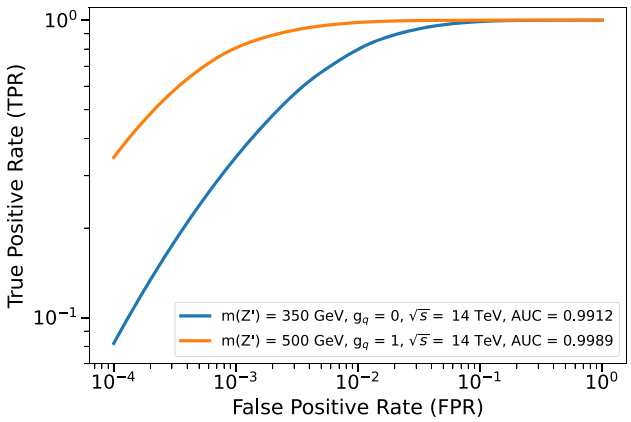

Fig. 8 True positive rate versus false positive rate of the BDT algorithm, for two different signal benchmark scenarios

Table 4 Event yields for the main backgrounds and the signal point for $m(Z') = 1.0$ TeV, for some of the bin entries for the output of the gradient boosting algorithm. The events correspond to 14 TeV, $g_q = 0$, and 3000 fb^{-1} luminosity scenario

$t\bar{t}\,Z$	$t\bar{t}\,h$	$t\bar{t}t\bar{t}$	$m(Z') = 1.0$ TeV
Events for bin entries 1–10			
3.686e+06	7.582e+04	7719	0.110
1.088e+05	682.5	229.6	0.060
4.969e+04	230.0	105.4	0.062
2.753e+04	118.7	59.53	0.045
1.853e+04	83.45	42.68	0.043
1.258e+04	49.14	27.44	0.029
1.06e+04	39.87	24.52	0.037
7984	25.96	16.45	0.031
6774	13.91	12.11	0.035
5468	12.98	11.70	0.025
Events for bin entries 41–50			
870.9	0	1.412	0.023
822.6	1.855	1.211	0.019
919.3	0.927	1.211	0.013
435.5	1.855	1.110	0.014
290.3	1.855	1.513	0.019
629.0	1.855	1.211	0.019
725.8	0.927	0.908	0.020
338.7	1.855	0.505	0.020
580.6	1.855	1.009	0.014
387.1	0.0	0.706	0.030
Events for bin entries 91–100			
387.1	0.0	1.11	0.182
629	0.927	1.412	0.240
580.6	0.0	0.8071	0.247
725.8	1.855	1.009	0.335
774.2	1.855	1.11	0.419
387.1	2.782	1.513	0.589
919.3	1.855	1.614	0.886
629	0.9273	2.724	1.605
1403	0.927	3.027	3.754
2952	3.709	3.935	214.0

nosity of 3000 fb^{-1}. The values in Table 4 are determined using $N = \sigma \times \mathrm{L} \times \epsilon$, where ϵ contains the efficiency of the pre-selection criteria times the probability that a given process will have a BDT output in a particular bin range. The bins are counted from 1 to 100, going from left to right, such that bin 1 is the leftmost bin near BDT output of 0, and bin 100 is the rightmost bin near a BDT output of 1. The backgrounds dominate over the SPM1 benchmark signal yields in a large part of the BDT output spectrum, especially near zero, where the background yields are about six orders of magnitude larger. The presence of signal will be observed as an enhancement in the yields near a BDT output of unity.

4 Results

Using the BDT distributions normalized to cross section times pre-selection efficiency times luminosity, we calculate the expected experimental signal significance of the proposed search methodology, for different signal models, LHC operation conditions, and integrated luminosity scenarios. As

noted earlier, we consider three values for the total integrated luminosity at the LHC: (i) 150 fb^{-1}, which is approximately the amount of pp data already collected by the ATLAS and CMS experiments; (ii) 300 fb^{-1}, expected in the next few years; and (iii) 3000 fb^{-1}, expected by the end of the High Luminosity LHC era. The significance is calculated using the expected bin-by-bin yields of the BDT output distribution in a profile likelihood fit, using the ROOTFit [98] package developed by CERN. Similar to Refs. [81,99–103], the signal significance Z_{sig} is determined using the probability of obtaining the same test statistic with the background-only hypothesis and the signal plus background hypothesis, defined as the local p-value. The value of Z_{sig} corresponds to the point where the integral of a Gaussian distribution between Z_{sig} and ∞ results in a value equal to the local p-value.

Systematic uncertainties are incorporated into the significance calculation as nuissance parameters, using a log-normal prior for normalization and a Gaussian prior for shape related uncertainties. The systematic uncertainties are based on both experimental and theoretical constraints. A 3% systematic uncertainty is used to account for experimental errors on the estimation of the integrated luminosity collected by experiments. This is a reasonable and conservative choice based on Ref. [104]. A systematic uncertainty is included due to the choice of PDF, with respect to the default set used to produce the simulated signal and background samples. The PDF uncertainties were calculated following the PDF4LHC prescription [98], and results in up to 5% systematic uncertainty, depending on the process. The effect of the chosen PDF set on the shape of the BDT output distribution is negligible. Other theoretical uncertainties were considered, such as the absence of higher-order contributions to the signal cross sections, which can alter the pre-selection efficiency and shapes of kinematic distributions which are fed into the BDT algorithm. This uncertainty is calculated by varying the renormalization and factorization scales by a factor of two with respect to the nominal value, and by considering the full change in the bin-by-bin yields of the BDT output distribution. They are found to be at most 3% in a given bin. For experimental uncertainties related to the reconstruction and identification of bottom quarks, Ref. [105] reports a systematic uncertainty of 1–5%, depending on p_{T} and η of the b-jet candidate. However, we assume a conservative 5% uncertainty per b-jet candidate, independent of p_{T} and η, which is correlated between signal and background processes with genuine bottom quarks, and correlated across BDT bins for each process. The electron and muon reconstruction, identification, and isolation requirements have an uncertainty of 2%, while a conservative 3% systematic uncertainty is set on the variation of the electron and muon energy/momentum scale and resolution [106,107]. We assumed 2–5% jet energy scale uncertainties, depending on η and p_{T}, resulting in shape-based uncertainties on the BDT output distribution that range

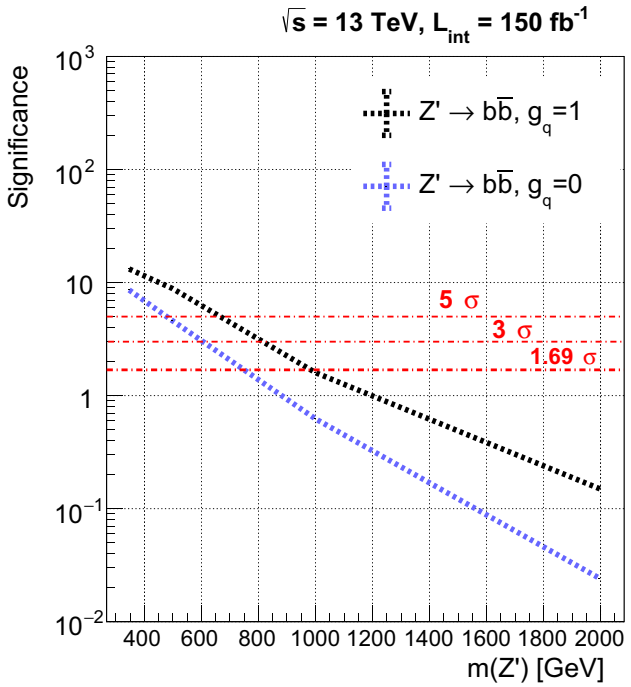

Fig. 9 Expected signal significance as function of reconstructed mass, at $\sqrt{s} = 13$ TeV and 150fb^{-1} luminosity, for the $g_q = 0$, 1 and $g_{Z'\mathrm{t\bar{t}}} = 1$ benchmark coupling scenarios. The 1.69σ reference point for exclusion, and the 3σ and 5σ points for discovery sensitivity are shown as red-dashed lines

from 1 to 4%, depending on the BDT bin. Finally, we consider a 10% systematic uncertainty associated with possible errors on the background predictions, which are uncorrelated between background processes.

Figure 9 shows the SPM1 signal significance as function of Z′ mass, for the $\{g_q, g_{Z'\mathrm{t\bar{t}}}\} = \{0, 1\}$ and $\{g_q, g_{Z'\mathrm{t\bar{t}}}\} = \{1, 1\}$ coupling scenarios, assuming $\sqrt{s} = 13$ TeV and 150 fb^{-1}. A signal significance of 1.69σ is our threshold to define expected exclusion at 90% confidence level, while 3σ (5σ) significance defines evidence (discovery) of new physics. For the $\{g_q, g_{Z'\mathrm{t\bar{t}}}\} = \{1, 1\}$ scenario, the analysis shows potential to exclude masses below 1.0 TeV, and achieve greater than 3σ (5σ) signal sensitivity for Z′ masses below 800 (675) GeV. For the SPM1 scenario with $\{g_q, g_{Z'\mathrm{t\bar{t}}}\} = \{0, 1\}$, the expected exclusion range is $m(\mathrm{Z}') < 780$ GeV, and the 3σ (5σ) reach is $m(\mathrm{Z}') < 600$ (500) GeV. Figure 10 shows the results for the same scenarios, but considering pp collisions at $\sqrt{s} = 14$ TeV and integrated luminosities of 300 fb^{-1} and 3000 fb^{-1}. For the $\{g_q, g_{Z'\mathrm{t\bar{t}}}\} = \{1, 1\}$ scenario and assuming an integrated luminosity of 3000 fb^{-1}, the expected exclusion bound goes up to $m(\mathrm{Z}') < 1.7$ TeV, while the 3σ reach improves to $m(\mathrm{Z}') < 1.45$ TeV. We note that the $\{g_q, g_{Z'\mathrm{t\bar{t}}}\} = \{1, 1\}$ scenario is a useful benchmark to compare the sensitivity to existing family-universal Z′ searches at CMS/ATLAS. The projected sensitivity obtained in this work for the $\{g_q, g_{Z'\mathrm{t\bar{t}}}\} = \{1, 1\}$ case is superseded by cur-

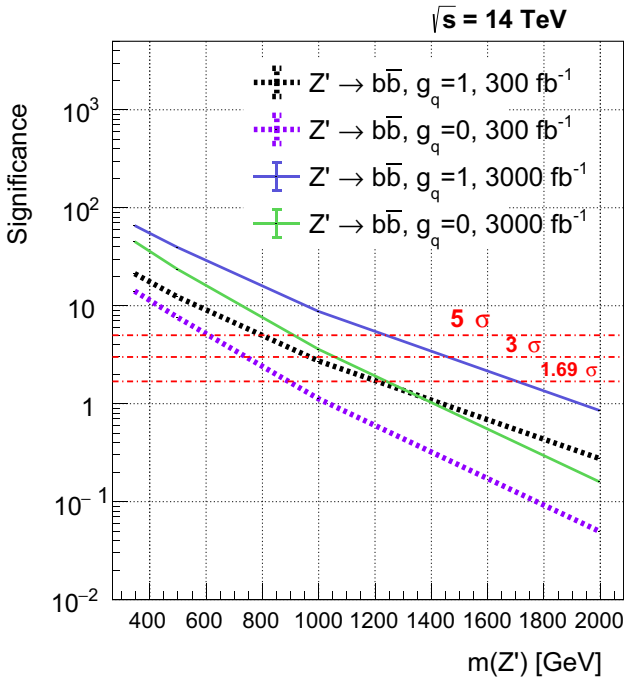

Fig. 10 Expected signal significance as function of reconstructed mass, at $\sqrt{s} = 14$ TeV and 300 fb^{-1} (3000fb^{-1}) luminosity, for the $g_q = 0, 1$ and $g_{Z't\bar{t}} = 1$ benchmark benchmark coupling scenarios. The 1.69σ reference point for exclusion, and the 3σ and 5σ points for discovery sensitivity are shown as red-dashed lines

rent and traditional searches at the LHC, such as searches targeting Drell–Yan production of $Z' \to jj$. However, even if a new gauge boson with family-universal couplings is first discovered via other analysis strategies, the proposed $pp \to t\bar{t}Z'(\to b\bar{b})$ search strategy remains important to measure the Z' couplings to third-generation fermions. On the other hand, the proposed search strategy in this work can provide the best mode for discovery in the case of a Z' coupling dominantly to the third generation.

We also estimate the expected signal significance for different SPM1 coupling scenarios of the Z' boson to t/b quarks. Figure 11 shows the signal significance for different $g_{Z't\bar{t}}$ and $m(Z')$ scenarios, with suppressed couplings to first and second generation quarks ($g_q = 0$), assuming $\sqrt{s} = 13$ TeV and 150 fb^{-1}. Figure 12 shows the corresponding results for the same $\{g_{Z't\bar{t}}, m(Z')\}$ combinations, but using $g_q = 1$. The results for $\sqrt{s} = 14$ TeV, assuming 300 fb^{-1} and 3000 fb^{-1}, are presented in Figs. 13, 14, 15 and 16 for different $\{g_q, g_{Z't\bar{t}}, m(Z')\}$ combinations.

Table 5 shows the SPM2 signal significance as function of $m(Z')$ and integrated luminosity, for the $\{c_t, \theta\} = \{1, \pi/2\}$ scenario, assuming $\sqrt{s} = 14$ TeV. The expected SPM2 exclusion range is $m(Z') < 1.5$ TeV at L = 300 fb^{-1}, while the 5σ discovery reach is $m(Z') < 1.5$ TeV for the 3000 fb^{-1} expected by the end of the high luminosity LHC era.

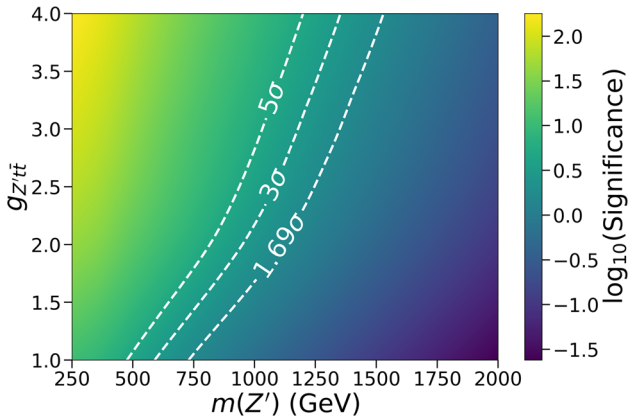

Fig. 11 Projected signal significance for the $g_q = 0$ benchmark model for different g_{tt} coupling scenarios and Z' masses. The estimates are performed at $\sqrt{s} = 13$ TeV and 150 fb^{-1}

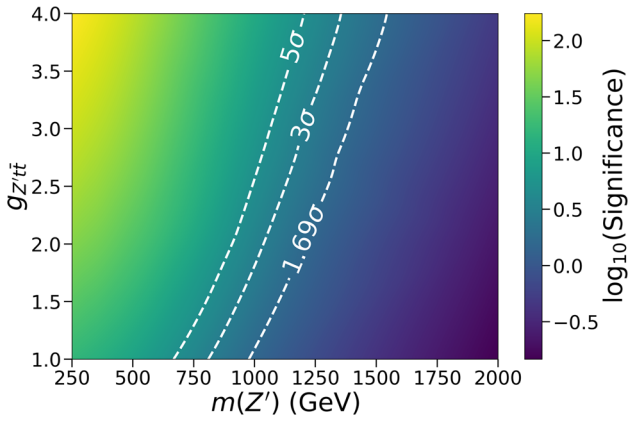

Fig. 12 Projected signal significance for the $g_q = 1$ benchmark model for different g_{tt} coupling scenarios and Z' masses. The estimates are performed at $\sqrt{s} = 13$ TeV and 150 fb^{-1}

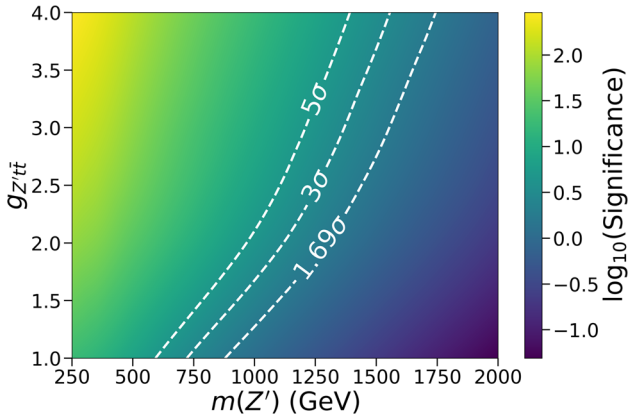

Fig. 13 Projected signal significance for the $g_q = 0$ benchmark model for different g_{tt} coupling scenarios and Z' masses. The estimates are performed at $\sqrt{s} = 14$ TeV and 300 fb^{-1}

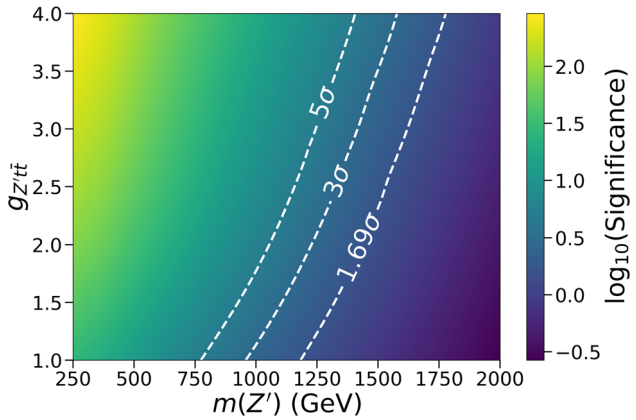

Fig. 14 Projected signal significance for the $g_q = 1$ benchmark model for different g_{tt} coupling scenarios and Z′ masses. The estimates are performed at $\sqrt{s} = 14$ TeV and 300 fb^{-1}

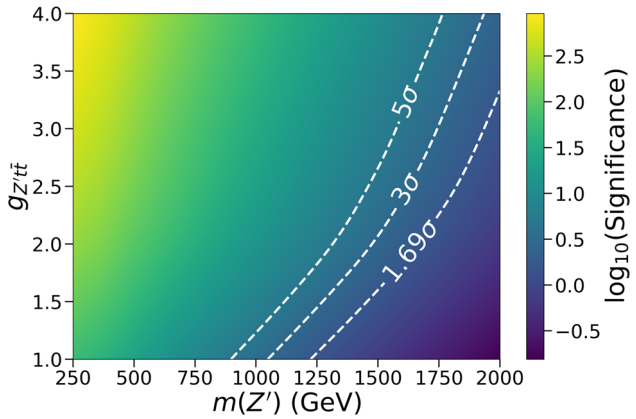

Fig. 15 Projected signal significance for the $g_q = 0$ benchmark model for different g_{tt} coupling scenarios and Z′ masses. The estimates are performed at $\sqrt{s} = 14$ TeV and 3000 fb^{-1}

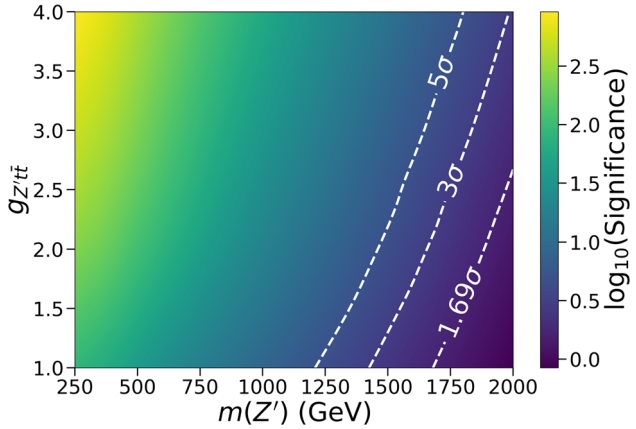

Fig. 16 Projected signal significance for the $g_q = 1$ benchmark model for different g_{tt} coupling scenarios and Z′ masses. The estimates are performed at $\sqrt{s} = 14$ TeV and 3000 fb^{-1}

Table 5 Projected signal significance for our second simplified model, considering the $c_t = 1$ coupling scenario with varying Z′ masses. The calculations are performed at $\sqrt{s} = 14$ TeV and assuming both 300 fb^{-1} and 3000 fb^{-1}

m(Z′)	300 fb $^{-1}$	3000 fb^{-1}
250 GeV	59.7	188.8
350 GeV	45.1	142.8
500 GeV	23.8	75.4
1000 GeV	3.31	10.48
1500 GeV	1.68	5.41
2000 GeV	0.135	0.427

5 Discussion

As the LHC continues to run with pp collisions at the highest energy, and with the slow increase in luminosity expected of the high-luminosity program of the accelerator, it is an important matter to ponder why certain searches for new physics have not provided strong evidence for discovery, and consider unexplored possibilities. In this work, we examine the phenomonology of a Z′ boson favoring higher-generation fermions (anogenophilic), in particular coupling to third generation fermions (tritogenophilic). This scenario is well motivated and arises in many theories that extend the SM [10–37,39–45,108]. It also seems to appear as a possible, although not yet confirmed, pattern in precision measurements of the B-physics sector [47–55,109–117] and the measurement of the muon anomalous magnetic moment [46]. An anogenophilic Z′ has already been investigated phenomenologically or experimentally for the case in which the new boson is produced in association with two top quarks and decays to two top quarks (top-philic [67–69]), tau/muon leptons [70,71], or muon/electron leptons [72]. Here we have presented a feasibility study for the Z′ decay into two b quarks. The study has been performed under the context of pp collisions at the LHC, at $\sqrt{s} = 13$ TeV and $\sqrt{s} = 14$ TeV, using a BDT algorithm to optimize the signal to background separation and maximize exclusion or discovery potential. Various coupling scenarios for the Z′ have been considered, including suppressed couplings to light flavour quarks ($g_q = 0$), enhanced couplings to third generation fermions, and preferential couplings to top and bottom quarks ($g_{Z't\bar{t}}$). Under the SPM1 $g_q = 1$ ($g_q = 0$) scenario, at $\sqrt{s} = 13$ TeV and integrated luminosity of 150 fb^{-1}, Z′ masses up to 1.0 TeV (780 GeV) can be excluded at 95% confidence level, while 5σ discovery potential exists for masses below 675 GeV (500 GeV). For the high luminosity era of the LHC with $\sqrt{s} = 14$ TeV and integrated luminosity of 3000 fb^{-1}, Z′ masses up to 1.70 TeV (1.25 TeV) can be excluded for the SPM1 $g_q = 1$ ($g_q = 0$) scenario, while the 5σ discovery reach is m(Z′) < 1.25 TeV (900 GeV). For the SPM2

benchmark scenario with $c_t = 1$ and $\theta = \pi/2$, the discovery (exclusion) reach is 1.5 (1.7) TeV at $\sqrt{s} = 14$ TeV and integrated luminosity of 3000 fb^{-1}. As noted previously, the projected sensitivity using the SPM2 scenario serves as a good comparison with other search strategies. For example, the authors of Ref. [68] examined the high luminosity LHC sensitivity to these anogenophilic scenarios using the $pp \to t\bar{t}Z' \to t\bar{t}t\bar{t}$ final state with boosted top tagging algorithms, and reported a projected 2σ reach of approximately $m(Z') < 1.5$ TeV for the same coupling scenario of $c_t = 1$, assuming an integrated luminosity of 3000 fb^{-1}. That result is to be compared with the stronger projected significance of $> 5.41\sigma$ for $m(Z') < 1.5$ TeV in Table 5, using the strategy presented in this paper. Additionally, Ref. [68] reports that a $> 5\sigma$ discovery reach is attainable for $m(Z') = 1.5$ TeV if $c_t > 1.65$. For comparison, Table 5 already shows a significance of 5.41σ for $m(Z') = 1.5$ TeV with a smaller coupling of $c_t = 1$. We also point out that these comparisons are conservative since the studies outlined in Ref. [68] assume a 100% branching ratio of $Z' \to t\bar{t}$, which would not be the case if Z' couples to both top and bottom quarks.

The main result of this paper is that probing heavy neutral gauge bosons produced in association with spectator top quarks, and decaying to a pair of bottom quarks, can be a key search methodology. It represents the most important anogenophilic/tritogenophilic mode for discovery at $m(Z') < 2m_t$ where the $Z' \to t\bar{t}$ decay is kinematically forbidden, and remains competitive with the $Z' \to t\bar{t}$ decay mode at TeV scale masses, benefiting from the possibility to reconstruct the Z' mass from the two highest-p_T b jets and resulting in events with reduced jet multiplicity. Furthermore, even if a Z' boson is discovered in other search channels when $m(Z')$ is large, a $t\bar{t}Z' \to t\bar{t}b\bar{b}$ search remains a key part of the search program at the LHC in order to establish the couplings of the Z' to all fermions. In particular, whereas a $t\bar{t}Z' \to t\bar{t}t\bar{t}$ search can measure the Z' mass and coupling to top quarks, the proposed $t\bar{t}Z' \to t\bar{t}b\bar{b}$ search can additionally measure the Z' coupling to bottom quarks.

The proposed data analysis represents a competitive alternative to complement searches already being conducted at the LHC. Those searches are based on the analysis of the mass distribution of two b-quark jets, in the resolved or boosted regime, using events whose triggers require high-p_T jets [1,4,75], b-quark jets [74], or a photon [73]. In the analysis strategy considered here instead, we can rely on the presence of an electron or muon lepton originating from the decay of a spectator top, which allows an unbiased selection of b-quark jets originating from the Z', or on the possibility to define a trigger using both a light lepton and jets, in order to select particles with lower energy, and thus probe lower values of $m(Z')$.

Because of the above reasons, we deem that the proposed analysis strategy should be considered in future Z' searches at the LHC, by both the ATLAS and the CMS Collaborations.

Acknowledgements We thank the constant and enduring financial support received for this project from the faculty of science at Universidad de Los Andes (Bogotá, Colombia), the Colombian Science Ministry MinCiencias, with the Grant program 70141, Contract number 164-2021, the Physics and Astronomy department at Vanderbilt University and the US National Science Foundation. This work is supported in part by NSF Award PHY-1945366 and a Vanderbilt Seeding Success Grant.

Data availability statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.]

Funded by SCOAP^3. SCOAP^3 supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

1. ATLAS Collaboration, Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. JHEP **03**, 145 (2020)
2. ATLAS Collaboration, Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector. Phys. Rev. D **104**(11), 112005 (2021)
3. ATLAS Collaboration, Search for a right-handed gauge boson decaying into a high-momentum heavy neutrino and a charged lepton in pp collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV. Phys. Lett. B **798**, 134942 (2019)
4. CMS Collaboration, Search for narrow resonances in the b-tagged dijet mass spectrum in proton–proton collisions at $\sqrt{s} = 13$ TeV. (2022). arXiv:2205.01835
5. CMS Collaboration, Search for resonant and nonresonant new phenomena in high-mass dilepton final states at $\sqrt{s} = 13$ TeV. JHEP **07**, 208 (2021)
6. CMS Collaboration, Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at $\sqrt{s} = 13$ TeV. JHEP **05**, 033 (2020)
7. CMS Collaboration, Search for singly and pair-produced leptoquarks coupling to third-generation fermions in proton-proton collisions at s=13 TeV. Phys. Lett. B **819**, 136446 (2021)
8. CMS Collaboration, Probing heavy Majorana neutrinos and the Weinberg operator through vector boson fusion processes in proton–proton collisions at $\sqrt{s} = 13$ TeV. (2022). arXiv:2206.08956

9. A. Das et al., Probing the minimal U(1)X model at future electron-positron colliders via fermion pair-production channels. Phys. Rev. D **105**(11), 115030 (2022)
10. C.T. Hill, Topcolor assisted technicolor. Phys. Lett. B **345**, 483–489 (1995)
11. L. Randall, R. Sundrum, A large mass hierarchy from a small extra dimension. Phys. Rev. Lett. **83**, 3370–3373 (1999)
12. L. Randall, R. Sundrum, An alternative to compactification. Phys. Rev. Lett. **83**, 4690–4693 (1999)
13. H. Davoudiasl et al., Experimental probes of localized gravity: on and off the wall. Phys. Rev. D **63**, 075004 (2001)
14. G.C. Branco et al., Theory and phenomenology of two-Higgs-doublet models. Phys. Rep. **516**, 1–102 (2012)
15. S. Gori et al., Closing the wedge: search strategies for extended Higgs sectors with heavy flavor final states. Phys. Rev. D **93**(7), 075038 (2016)
16. P.S. Bhupal Dev, A. Pilaftsis, Maximally symmetric two Higgs doublet model with natural standard model alignment. JHEP **12**, 024 (2014) [Erratum: JHEP 11, 147 (2015)]
17. J.F. Gunion et al., *The Higgs Hunter's Guide*, vol. 80. Fontiers in Physics (2000)
18. R.N. Mohapatra, J.C. Pati, "Natural" left-right symmetry. Phys. Rev. D **11**, 2558 (1975)
19. B.A. Dobrescu et al., Massive color-octet bosons and pairs of resonances at hadron colliders. Phys. Lett. B **670**, 119–123 (2008)
20. C.-R. Chen et al., Color sextet scalars at the CERN large hadron collider. Phys. Rev. D **79**, 054002 (2009)
21. Y. Bai, B.A. Dobrescu, Heavy octets and tevatron signals with three or Four b jets. JHEP **07**, 100 (2011)
22. E.L. Berger et al., Color sextet scalars in early LHC experiments. Phys. Rev. Lett. **105**, 181802 (2010)
23. H. Zhang et al., Color sextet vector bosons and same-sign top quark pairs at the LHC. Phys. Lett. B **696**, 68–73 (2011)
24. M. Gerbush et al., Color-octet scalars at the CERN LHC. Phys. Rev. D **77**, 095003 (2008)
25. N. Arkani-Hamed et al., The littlest Higgs. JHEP **07**, 034 (2002)
26. A. Pomarol, J. Serra, Top quark compositeness: feasibility and implications. Phys. Rev. D **78**, 074026 (2008)
27. O. Matsedonskyi et al., Light top partners for a light composite Higgs. JHEP **01**, 164 (2013)
28. B. Gripaios et al., Search strategies for top partners in composite Higgs models. JHEP **08**, 171 (2014)
29. D. Liu, R. Mahbubani, Probing top-antitop resonances with $t\bar{t}$ scattering at LHC14. JHEP **04**, 116 (2016)
30. G. Cacciapaglia et al., Composite scalars at the LHC: the Higgs, the sextet and the octet. JHEP **11**, 201 (2015)
31. R. Barbieri et al., The Higgs boson from an extended symmetry. Phys. Rev. D **76**, 115008 (2007)
32. G. Panico, A. Wulzer, The discrete composite Higgs model. JHEP **09**, 135 (2011)
33. S. De Curtis et al., The 4D composite Higgs. JHEP **04**, 042 (2012)
34. D. Marzocca et al., General composite Higgs models. JHEP **08**, 013 (2012)
35. B. Bellazzini et al., Composite Higgs sketch. JHEP **11**, 003 (2012)
36. G. Panico et al., On the tuning and the mass of the composite Higgs. JHEP **03**, 051 (2013)
37. B. Bellazzini et al., Composite Higgses. Eur. Phys. J. C **74**(5), 2766 (2014)
38. R. Leonardi et al., Phenomenology at the LHC of composite particles from strongly interacting Standard Model fermions via four-fermion operators of NJL type. Eur. Phys. J. C **80**, 309 (2020)
39. U. Haisch, E. Re, Simplified dark matter top-quark interactions at the LHC. JHEP **06**, 078 (2015)
40. M.R. Buckley et al., Scalar simplified models for dark matter. Phys. Rev. D **91**, 015017 (2015)
41. P. Cox et al., Novel collider and dark matter phenomenology of a top-philic Z′. JHEP **06**, 110 (2016)
42. Y. Zhang, Top quark mediated dark matter. Phys. Lett. B **720**, 137–141 (2013)
43. S. Baek et al., Top-philic scalar dark matter with a vector-like fermionic top partner. JHEP **10**, 117 (2016)
44. C. Arina et al., A comprehensive approach to dark matter studies: exploration of simplified top-philic models. JHEP **11**, 111 (2016)
45. J. D'Hondt et al., Signatures of top flavour-changing dark matter. JHEP **03**, 060 (2016)
46. Muon g-2 Collaboration, Measurement of the positive muon anomalous magnetic moment to 0.46 ppm. Phys. Rev. Lett. **126**(14), 141801 (2021)
47. BaBar Collaboration, Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays. Phys. Rev. Lett. **109**, 101802 (2012)
48. BaBar Collaboration, Measurement of an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays and implications for charged Higgs bosons. Phys. Rev. D **88**, 072012 (2013)
49. Belle Collaboration, Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle. Phys. Rev. D **92**, 072014 (2015)
50. Belle Collaboration, Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method. Phys. Rev. D **94**, 072007 (2016)
51. Belle Collaboration, Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$. Phys. Rev. Lett. **118**, 211801 (2017)
52. Belle Collaboration, Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ with one-prong hadronic τ decays at Belle. Phys. Rev. D **97**, 012004 (2018)
53. Belle Collaboration, Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method. Phys. Rev. Lett. **124**, 161803 (2020)
54. LHCb Collaboration, Measurement of the ratio of branching fractions $mathcal B(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/mathcal B(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$. Phys. Rev. Lett. **115**, 111803 (2015) [Erratum: Phys.Rev.Lett. 115, 159901 (2015)]
55. LHCb Collaboration, Measurement of the ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays. Phys. Rev. Lett. **120**, 171802 (2018)
56. HFAG Collaboration, Average of $R_D^{\tau/\ell'}$ and $R_{D^{(*)}}^{\tau/\ell'}$ measurements. https://hflav-eos.web.cern.ch/hflav-eos/semi/spring19/html/RDsDsstar/RDRDs.html. Accessed 7 Jan 2023
57. ATLAS Collaboration, Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. JHEP **04**, 046 (2019)
58. 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**(11), 1085 (2020)
59. 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. JHEP **11**, 118 (2021)
60. CMS Collaboration, Measurement of the cross section for $t\bar{t}$ production with additional jets and b jets in pp collisions at $\sqrt{s}$ = 13 TeV. JHEP **07**, 125 (2020)
61. CMS Collaboration, Measurement of the $t\bar{t}b\bar{b}$ production cross section in the all-jet final state in pp collisions at $\sqrt{s}$ = 13 TeV. Phys. Lett. B **803**, 135285 (2020)
62. A.M. Sirunyan et al., Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s}$ = 8 and 13 TeV. Phys. Lett. B **793**, 320–347 (2019)
63. CMS Collaboration, Searches for additional Higgs bosons and for vector leptoquarks in $\tau\tau$ final states in proton–proton collisions at $\sqrt{s}$ = 13 TeV. (2022). arXiv:2208.02717

64. CMS Collaboration, The search for a third-generation leptoquark coupling to a τ lepton and a b quark through single, pair and nonresonant production at $\sqrt{s} = 13$ TeV. Geneva, cMS-EXO-19-016 (2022). http://cds.cern.ch/record/2815309?ln=en. Accessed 7 Jan 2023
65. G. Aad et al., Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector. Phys. Lett. B **796**, 68–87 (2019)
66. CMS Collaboration, Search for a narrow resonance in high-mass dilepton final states in proton-proton collisions using 140 fb^{-1} of data at $\sqrt{s} = 13$ TeV. CERN, Geneva, Technical report (2019). https://cds.cern.ch/record/2684757. Accessed 7 Jan 2023
67. N. Greiner et al., Model-independent production of a top-philic resonance at the LHC. JHEP **04**, 029 (2015)
68. J.H. Kim et al., Probing TeV scale top-philic resonances with boosted top-tagging at the high luminosity LHC. Phys. Rev. D **94**(3), 035023 (2016)
69. P.J. Fox et al., Top-philic Z' forces at the LHC. JHEP **03**, 074 (2018)
70. M. Abdullah et al., A heavy neutral gauge boson near the Z boson mass pole via third generation fermions at the LHC. Phys. Lett. B **803**, 135326 (2020)
71. J.F. Kamenik et al., Lepton flavor universality violation without new sources of quark flavor violation. Phys. Rev. D **97**(3), 035002 (2018)
72. CMS Collaboration, Search for physics beyond the standard model in multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. JHEP **03**, 051 (2020)
73. ATLAS Collaboration, Search for low-mass resonances decaying into two jets and produced in association with a photon using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Phys. Lett. B **795**, 56–75 (2019)
74. CMS Collaboration, Search for narrow resonances in the b-tagged dijet mass spectrum in proton-proton collisions at $\sqrt{s} = 8$ TeV. Phys. Rev. Lett. **120**(20), 201801 (2018)
75. CMS Collaboration, Search for low-mass resonances decaying into bottom quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV. Phys. Rev. D **99**(1), 012005 (2019)
76. J. Alwall et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. JHEP **07**, 079 (2014)
77. R.D. Ball et al., Parton distributions for the LHC Run II. JHEP **04**, 040 (2015)
78. T. Sjöstrand et al., An introduction to PYTHIA 8.2. Comput. Phys. Commun. **191**, 159–177 (2015)
79. J. de Favereau et al., DELPHES 3, a modular framework for fast simulation of a generic collider experiment. JHEP **02**, 057 (2014)
80. J. Alwall et al., Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. Eur. Phys. J. C **53**, 473–500 (2008)
81. A. Florez et al., Searching for new heavy neutral gauge bosons using vector boson fusion processes at the LHC. Phys. Lett. B **767**, 126–132 (2017)
82. A. Sirunyan et al., Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. JINST **13**(04), P05011 (2018)
83. CMS Collaboration, Study of the discovery reach in searches for supersymmetry at CMS with 3000/fb. Geneva, cMS-PAS-FTR-13-014 (2013). https://cds.cern.ch/record/1607141. Accessed 7 Jan 2023
84. A. Sirunyan et al., Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at $\sqrt{s} = 13$ TeV. JINST **15**(02), P02027 (2020)
85. A. Sirunyan et al., Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. JINST **16**(05), P05014 (2021)
86. J.H. Friedman, Greedy function approximation: a gradient boosting machine. Ann. Stat. **29**(5), 1189–1232 (2001)
87. X. Ai et al., Probing highly collimated photon-jets with deep learning. (2022). arXiv:2203.16703
88. M. Aaboud et al., Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. Phys. Rev. D **97**(7), 072016 (2018)
89. S. Chigusa et al., Deeply learned preselection of Higgs dijet decays at future lepton colliders. Phys. Lett. B **833**, 137301 (2022)
90. Y.-L. Chung et al., Disentangling boosted Higgs boson production modes with machine learning. JINST **16**, P07002 (2021)
91. J. Feng et al., Improving heavy Dirac neutrino prospects at future hadron colliders using machine learning. **9**, 141 (2022). https://doi.org/10.1088/1748-0221/16/07/P07002
92. V. Khachatryan et al., Search for the associated production of the Higgs boson with a top-quark pair. JHEP **09**, 087 (2014) [Erratum: JHEP 10, 106 (2014)]
93. R. Santos et al., Machine learning techniques in searches for $t\bar{t}h$ in the $h \to b\bar{b}$ decay channel. JINST **12**(04), P04014 (2017)
94. G. Kasieczka et al., The machine learning landscape of top taggers. SciPost Phys. **7**, 014 (2019)
95. X. Vidal et al., How to use machine learning to improve the discrimination between signal and background at particle colliders. Appl. Sci. **11**(22), 11076 (2021)
96. F. Pedregosa et al., Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**(null), 2825–2830 (2011)
97. T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD '16. New York, NY, USA* (Association for Computing Machinery, 2016), pp. 785–794
98. J. Butterworth et al., PDF4LHC recommendations for LHC Run II. J. Phys. G **43**, 023001 (2016)
99. A. Flórez et al., Probing axionlike particles with $\gamma\gamma$ final states from vector boson fusion processes at the LHC. Phys. Rev. D **103**(9), 095001 (2021)
100. A. Flórez et al., Anapole dark matter via vector boson fusion processes at the LHC. Phys. Rev. D **100**(1), 016017 (2019)
101. A. Flórez et al., Probing heavy spin-2 bosons with $\gamma\gamma$ final states from vector boson fusion processes at the LHC. Phys. Rev. D **99**(3), 035034 (2019)
102. A. Flórez et al., Expanding the reach of heavy neutrino searches at the LHC. Phys. Lett. B **778**, 94–100 (2018)
103. A. Flórez et al., Probing the stau-neutralino coannihilation region at the LHC with a soft tau lepton and a jet from initial state radiation. Phys. Rev. D **94**(7), 073007 (2016)
104. P. Lujan et al., CMS BRIL collaboration, The Pixel Luminosity Telescope: a detector for luminosity measurement at CMS using silicon pixel sensors. Eur. Phys. J. C. (2022). arXiv:2206.08870
105. A. Sirunyan et al., Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. JINST **13**, P05011 (2018)
106. CMS Collaboration, Search for heavy resonances and quantum black holes in eμ, eτ, and $\mu\tau$ final states in proton–proton collisions at $\sqrt{s}$ = 13 TeV. (2022). arXiv:2205.06709
107. ATLAS Collaboration, Search for the charged-lepton-flavor-violating decay $Z \to e\mu$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. (2022). arXiv:2204.10783
108. S. Baek, Dark matter contribution to $b \to s\mu^+\mu^-$ anomaly in local $U(1)_{L_\mu - L_\tau}$ model. Phys. Lett. B **781**, 376–382 (2018)
109. LHCb Collaboration, Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$. Phys. Rev. Lett. **120**, 121801 (2018)
110. Belle Collaboration, Lepton-flavor-dependent angular analysis of $B \to K^*\ell^+\ell^-$. Phys. Rev. Lett. **118**, 111801 (2017)

111. LHCb Collaboration, Measurement of form-factor-independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$. Phys. Rev. Lett. **111**, 191801 (2013)
112. LHCb Collaboration, Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays. JHEP **06**, 133 (2014)
113. LHCb Collaboration, Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays. Phys. Rev. Lett. **113**, 151601 (2014)
114. LHCb Collaboration, Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity. JHEP **02**, 104 (2016)
115. LHCb Collaboration, Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$. JHEP **09**, 179 (2015)
116. LHCb Collaboration, Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays. JHEP **08**, 055 (2017)
117. LHCb Collaboration, Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays. Phys. Rev. Lett. **122**, 191801 (2019)