**THE EUROPEAN
PHYSICAL JOURNAL C**

On the sensitivity reach of LQ production with preferential couplings to third generation fermions at the LHC

A. Flórez[1,a] , **J. Jones-Pérez**[2,b] , **A. Gurrola**[3,c], **C. Rodriguez**[1,d] , **J. Peñuela-Parra**[1,e]

[1] Departamento de Física, Universidad de Los Andes, Cra. 1 # 18a-12, Bogotá, Colombia
[2] Sección Física, Departamento de Ciencias, Pontificia Universidad Católica del Perú, Apartado 1761, Lima, Peru
[3] Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA

Abstract Leptoquarks (LQs) are hypothetical particles that appear in various extensions of the Standard Model (SM), that can explain observed differences between SM theory predictions and experimental results. The production of these particles has been widely studied at various experiments, most recently at the Large Hadron Collider (LHC), and stringent bounds have been placed on their masses and couplings, assuming the simplest beyond-SM (BSM) hypotheses. However, the limits are significantly weaker for LQ models with family non-universal couplings containing enhanced couplings to third-generation fermions. We present a new study on the production of a LQ at the LHC, with preferential couplings to third-generation fermions, considering proton-proton collisions at $\sqrt{s} = 13$ TeV and $\sqrt{s} = 13.6$ TeV. Such a hypothesis is well motivated theoretically and it can explain the recent anomalies in the precision measurements of B-meson decay rates, specifically the $R_{D^{(*)}}$ ratios. Under a simplified model where the LQ masses and couplings are free parameters, we focus on cases where the LQ decays to a τ lepton and a b quark, and study how the results are affected by different assumptions about chiral currents and interference effects with other BSM processes with the same final states, such as diagrams with a heavy vector boson, Z′. The analysis is performed using machine learning techniques, resulting in an increased discovery reach at the LHC, allowing us to probe new physics phase space which addresses the B-meson anomalies, for LQ masses up to 5.00 TeV, for the high luminosity LHC scenario.

[a] e-mail: ca.florez@uniandes.edu.co (corresponding author)

[b] e-mail: jones.j@pucp.edu.pe

[c] e-mail: alfredo.gurrola@vanderbilt.edu

[d] e-mail: c.rodriguez45@uniandes.edu.co

[e] e-mail: j.penuela@uniandes.edu.co

1 Introduction

After more than ten years collecting data, the LHC has confirmed that the Standard Model (SM) is indeed the correct theory describing particle physics for energies below the TeV scale. Nevertheless, there exist reasons to expect the SM to be a low-energy effective realization of a more complete theory. On the theoretical side, we do not know if gravity should be quantized, or if the gauge interactions should be unified, and if so, we do not know how to solve the associated hierarchy problems on the Higgs mass. Moreover, we have no explanation for fermion family replication, nor for the lack of CP violation in the strong sector. This expectation for physics beyond the SM (BSM) is reinforced experimentally, where the observation of neutrino masses, dark matter, and the baryon asymmetry in the Universe, cannot be explained by the SM.

Leptoquarks (LQs) are hypothetical bosons carrying both baryon and lepton number, thus interacting jointly with a lepton and a quark. They are a common ingredient in SM extensions where quarks and leptons share the same multiplet. Typical examples of these can be found in the Pati-Salam [1] and SU(5) GUT [2] models. In addition, they can also be found in theories with strong interactions, such as compositeness [3]. Due to their exotic coupling which allows quark-lepton transitions, they have a diverse phenomenology, which naturally leads to several constraints. An important one comes from proton decay, which forces the LQ mass to values close to the Planck scale, unless baryon and lepton numbers are not violated. Furthermore, in models where the latter are conserved, the LQ can still be subject to a wide variety of bounds [4–9]. Examples of these come from meson mixing, electric and magnetic dipole moments, atomic parity violation tests, rare decays, and direct searches. Nevertheless, the significance of each bound is a model dependent question.

In the last years, an increased interest in low scale LQs has emerged due to the anomalies in the precision measurements of the B-meson decay rates. As it is well known, these corresponded mainly to deviations in the $R_{K^{(*)}}$ [10–13] and $R_{D^{(*)}}$ [14–25] ratios, which measure the violation of lepton flavour universality (LFU). What followed was a very intense theoretical development, aiming to explain the anomalies by TeV scale LQ exchange at tree level [26–41]. Before the end of 2022, it was generally agreed that, within proposed single LQ solutions, the only candidate capable of addressing all B-meson anomalies simultaneously and surviving all other constraints was a vector LQ (U_1), transforming as (**3**, **1**, 2/3), and coupling mainly to third-generation fermions via b τ and t ν_τ vertices [36,39]. In spite of a recent re-analysis of $R_{K^{(*)}}$ data showing this ratio to be compatible with the SM prediction [42–45], the solution to the $R_{D^{(*)}}$ anomaly is still an open question and remains a valid motivation for the study of scenarios where new particles have preferential couplings to third-generation fermions. Thus, it is still of interest to continue exploring the possibility of observing the U_1 LQ at the LHC [41].

As expected, the theoretical community has extensively participated in probing LQ models by scrutinizing search strategies, recasting LHC results, and predicting the reach in the parameter space via different searches involving third-generation fermions (see for instance [46–55]). In addition, several 13 TeV searches for LQs decaying into t/b and τ/ν final states have been performed by the CMS [56–64] and ATLAS [65–71] collaborations.

Of the searches above, we find [62] particularly interesting. Here, the CMS collaboration explores signals corresponding to t ν b τ and t ν τ final states, with 137 fb^{-1} of proton-proton (p p) collision data. The former is motivated by LQ pair production, with one LQ decaying into t ν and the other into b τ, while the latter arises from a single LQ produced in association with a τ, with a subsequent LQ decay into t ν (see Fig. 1 for the corresponding diagrams). From the combination of both production channels, the search excludes U_1 masses under $1.3 - 1.7$ TeV, with this range depending on the U_1 coupling to gluons and on its coupling g_U in the b$_L$ τ_L vertex.

What makes this search particularly attractive is that, for the first time, an LHC collaboration directly places (mass dependent) bounds on g_U. This is important, since having information on this parameter is crucial in order to understand if the U_1 is really responsible for the $R_{D^{(*)}}$ anomaly. The inclusion of the single-LQ production mode is important, since its cross-section is directly proportional to g_U^2. However, as can be seen in Fig. 6 of [62], the current constraints are dominated by pair production, with single-LQ production playing a subleading role. While this is expected [49], it still leads us to ponder the possibility of improving the sensitivity of LHC searches to single-LQ production, and thus
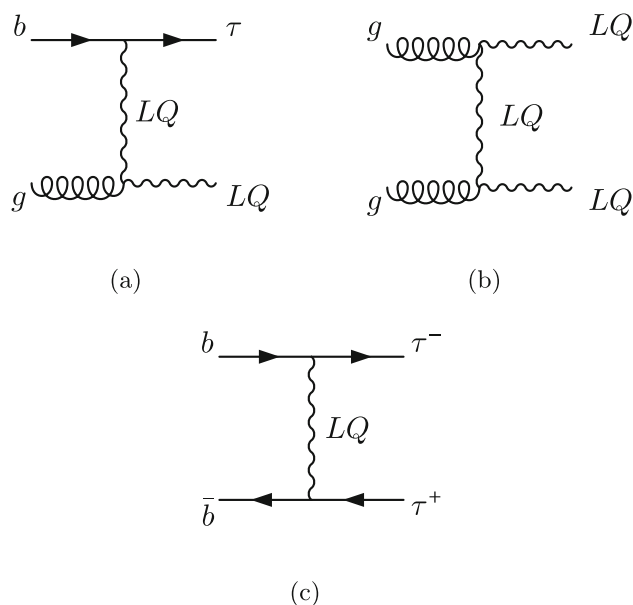

Fig. 1 Representative Feynman diagrams of single (**a**), pair (**b**), and non-resonant (**c**) leptoquark production in proton-proton collision experiments. In single and pair production, the diagrams shown involve t-channel LQ exchange, dominant for lower LQ mass. However, for larger mass there exist s-channel diagrams featuring a virtual bottom quark and gluon, respectively

on achieving better constraints on g_U. Other complementary and similar searches to [62] were carried out by both ATLAS [70] and CMS [64].

It is also well known, though, that searches for an excess in the high-p_T tails of τ lepton distributions can strongly probe g_U, up to very large LQ masses. Indeed, as shown in [41,72], the new physics effective operators contributing to $R_{D^{(*)}}$ also contribute to an enhancement in the p p $\to$ $\tau\tau$ production rates. This has motivated a large number of recasts [39,41,49,51,73–77], as well as a CMS search explicitly providing constraints in terms of U_1 [63]. Nevertheless, it is important to note that for these p p $\to$ $\tau\tau$ processes, the LQ participates non-resonantly, so contributions to the p p $\to$ $\tau\tau$ rates and kinematic distributions from non-LQ BSM diagrams containing possible virtual particles, such as a heavy neutral vector boson Z$'$, could spoil a straightforward interpretation of any possible excess [51]. Thus, it is also necessary to understand how the presence of other virtual particles can affect the sensitivity of an analysis probing g_U.

In this work we study the projected LQ sensitivity at the LHC, considering already available p p data as well as the expected amount of data to be acquired during the High-Luminosity LHC (HL-LHC) runs. We explore a proposed analysis strategy which utilizes a combination of single-, double-, and non-resonant-LQ production, targeting final states with varying τ-lepton and b-jet multiplicities. The

studies are performed considering various benchmark scenarios for different LQ masses and couplings, also taking into account distinct chiralities for the third-generation fermions in the LQ vertex. We also assess the impact of a companion Z', which is typical of gauge models, in non-resonant LQ probes, and find that interference effects can have a significant effect on the discovery reach. We consider this effect to be of high interest, given that non-resonant LQ production can have the largest cross-section, and thus could be an important channel in terms of discovery potential.

An important aspect of this work is that the analysis strategy is developed using a machine learning (ML) algorithm based on Boosted Decision Trees (BDT)[78]. The output of the event classifier is used to perform a profile-binned likelihood test to extract the overall signal significance for each model considered in the analysis. The advantage of using BDTs and other ML algorithms has been demonstrated in several experimental and phenomenological studies [50,79–84]. In our studies, we find that the BDT algorithm gives sizeable improvement in signal significance.

This paper is organized as follows. In Sect. 2 we present our simplified model and review the model parameters which are relevant for solving the B-meson anomalies. Section 3 describes the details associated with the analysis strategy and the simulation of signal and background samples. Section 4 contains the results of the study, including the projected sensitivity for different benchmark scenarios considered. Finally, in Sect. 5 we discuss the implication of our results and prospects for future studies.

2 A simplified model for the U_1 leptoquark

Extending the SM with a massive U_1 vector LQ is not straightforward, as one has to ensure the renormalizability of the model. Most of the theoretical community has focused on extensions of the Pati-Salam (PS) models which avoid proton decay, such as the scenario found in [85]. Other examples include PS models with vector-like fermions [86–88], the so-called 4321 models [89–91], the twin PS^2 model [92,93], the three-site PS^3 model [94–96], as well as composite PS models [97–99].

In what follows, we shall restrict ourselves to a simplified non-renormalizable lagrangian, understood to be embedded into a more complete model. The SM is thus extended by adding the following terms featuring the U_1 LQ:

$$\mathcal{L}_{U_1} = -\frac{1}{2} U_{\mu\nu}^\dagger U^{\mu\nu} + M_U^2\, U_{1\mu}^\dagger U_1^\mu$$
$$-ig_s\, U_{1\mu}^\dagger\, T^a\, U_{1\nu}\, G^{a\mu\nu} - i\frac{2}{3} g'\, U_{1\mu}^\dagger U_{1\nu} B^{\mu\nu}$$
$$+\frac{g_U}{\sqrt{2}}[U_{1\mu}(\bar{Q}_3\,\gamma^\mu L_3 + \beta_L^{s\tau}\,\bar{Q}_2\,\gamma^\mu L_3$$
$$+\beta_R\,\bar{b}_R\,\gamma^\mu \tau_R) + \text{h.c.}] \tag{1}$$

where $U_{\mu\nu} \equiv \mathcal{D}_\mu U_{1\nu} - \mathcal{D}_\nu U_{1\mu}$, and $\mathcal{D}_\mu \equiv \partial_\mu + ig_s T^a G_\mu^a + i\frac{2}{3} g' B_\mu$. As evidenced by the second line above, we assume that the LQ has a gauge origin.[1]

The third and fourth lines in in Eq. (1) shows the LQ interactions with SM fermions, with coupling g_U, which we have chosen as preferring the third generation.[2] These are particularly relevant for the LQ decay probabilities, as well as for the single-LQ production cross-section. The $\beta_L^{s\tau}$ parameter, which is the LQ $\to s\tau$ coupling in the β_L matrix (see footnote), is chosen to be equal to 0.2, following the fit done in [75], in order to simultaneously solve the $R_{D^{(*)}}$ anomaly and satisfy the p p $\to \tau^+\tau^-$ constraints. Although $\beta_L^{s\tau}$ technically alters the single-LQ production cross-section and LQ branching fractions, we have confirmed that a value of $\beta_L^{s\tau} = 0.2$ results in negligible impact on our collider results, and thus is ignored in our subsequent studies.

The LQ right-handed coupling is modulated with respect to the left-handed one by the β_R parameter. The choice of β_R is important phenomenologically, as it affects the LQ branching ratios,[3] as well as the single-LQ production cross-section. To illustrate the former, Fig. 2 (top) shows the LQ $\to$ bτ and LQ $\to$ tν branching ratios as functions of the LQ mass, for two values of β_R. For large LQ masses, we confirm that with $\beta_R = 0$ then BR(LQ $\to$ bτ) $\approx$ BR(LQ $\to$ tν) $\approx \frac{1}{2}$. However, for $\beta_R = -1$, as was chosen in [38], the additional coupling adds a new term to the total amplitude, leading to BR(LQ $\to$ bτ) $\approx \frac{2}{3}$. The increase in this branching ratio can thus weaken bounds from LQ searches targeting decays into tν final states, which motivates exploring the sensitivity in bτ final states exclusively. Note that although a BR(LQ $\to$ bτ) $\approx$ 1 scenario is possible by having the LQ couple exclusively to right-handed currents (i.e, $g_U \to 0$, but $g_U\beta_R \neq 0$), it does not solve the observed anomalies in the $R_{D^{(*)}}$ ratios. Therefore, although some LHC searches assume BR(LQ $\to$ bτ) = 1, we stress that in our studies we assume values of the model parameters and branching ratios that solve the $R_{D^{(*)}}$ ratios.

To further understand the role of β_R at colliders, Fig. 2 (bottom) shows the cross-section for single-LQ (sLQ), double-LQ (dLQ), and non-resonant (non-res) production, as a function of mass and for a fixed coupling $g_U = 1.8$, assuming p p collisions at $\sqrt{s} = 13$ TeV. We note that this benchmark scenario with $g_U = 1.8$ results in a LQ $\to$ bτ decay width that is <5% of the LQ mass, for mass values

---

[1] The couplings in the second line of Eq. (1) can be found in the literature as $g_s \to g_s(1 - \kappa_U)$ and $g' \to g'(1 - \tilde{\kappa}_U)$, in order to take into account the possibility of an underlying strong interaction.

[2] Before the demise of the $R_{K^{(*)}}$ anomaly [42–45], a 3 × 3 β_L matrix would be used instead, with values fitted to solve all B meson anomalies.

[3] Having $\beta_L^{s\tau}$ different from zero also opens new decay channels. These, however, are either suppressed by $\beta_L^{s\tau}$ and powers of λ_{CKM}. In any case, this effect would decrease BR(LQ $\to$ bτ) and BR(LQ $\to$ tν) by less than 3%.
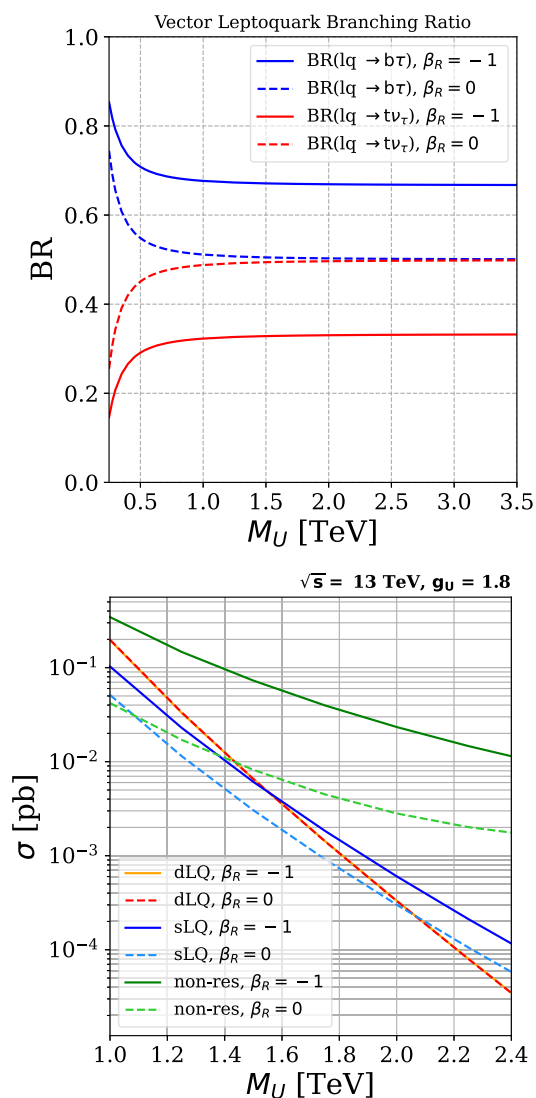
Fig. 2 Top: The LQ → bτ and LQ → tν branching ratios for $\beta_R = 0$ (solid lines) and $\beta_R = -1$ (dashed lines). Bottom: Signal cross-section as a function of the LQ mass, for $\sqrt{s} = 13$ TeV, with $g_U = 1.8$. We show single, pair, and non-resonant production, for $\beta_R = -1, 0$ in solid and dashed lines, respectively

from 250 GeV to 2.5 TeV. In the figure, we observe that, since dLQ production is mainly mediated by events from quantum chromodynamic processes, the choice of β_R does not affect the cross-section. However, for sLQ production, a non-zero value for β_R increases the cross-section by about a factor of 2 and by almost one order of magnitude in the case of non-res production. These results shown in Fig. 2 are easily understood by considering the diagrams shown in Fig. 1. The LQ mass value where the sLQ production cross-section exceeds the dLQ cross-section depends on the choice of g_U.

We also note that to solve the $R_{D^{(*)}}$ anomaly, the authors of [75] point out that the wilson coefficient $C_U \equiv g_U^2 \, v_{SM}^2/(4 \, M_U^2)$ is constrained to a specific range of values,

and this range depends on the value of the β_R parameter. Therefore, the allowed values of the coupling g_U depend on M_U and β_R, and thus our studies are performed in this multi-dimensional phase space.

As noted in Sect. 1, we study the role of a Z′ boson in $pp \to \tau\tau$ production. The presence of a Z′ boson in LQ models has been justified in various papers, for example, in [51]. The argument is that minimal extensions of the SM which include a massive gauge U_1 LQ, uses the gauge group $SU(4) \times SU(3)' \times SU(2)_L \times U(1)_{T_R^3}$. Such an extension implies the presence of an additional massive boson, Z′, and a color-octet vector, G′, arising from the spontaneous symmetry breaking into the SM.[4] The Z′ in particular can play an important role in the projected LQ discovery reach, as it can participate in $pp \to \tau\tau$ production by s-channel exchange, both resonantly and as a virtual mediator. To study the effect of a Z′ on the $pp \to \tau\tau$ production cross-sections and kinematics, we extend our benchmark Lagrangian in Eq. (1) with further non-renormalizable terms involving the Z′. Accordingly, we assume the Z′ only couples to third-generation fermions. Our simplified model is thus extended by:

$$\mathcal{L}_{Z'} = -\frac{1}{4} Z'_{\mu\nu} Z'^{\mu\nu} + \frac{1}{2} M_{Z'}^2 Z'_\mu Z'^\mu$$
$$+ \frac{g_{Z'}}{2\sqrt{6}} Z'^\mu (\zeta_q \bar{Q}_3 \gamma_\mu Q_3 + \zeta_t \bar{t}_R \gamma_\mu t_R$$
$$+ \zeta_b \bar{b}_R \gamma_\mu b_R - 3\zeta_\ell \bar{L}_3 \gamma_\mu L_3 - 3\zeta_\tau \bar{\tau}_R \gamma_\mu \tau_R) \tag{2}$$

where the constants $M_{Z'}, g_{Z'}, \zeta_q, \zeta_t, \zeta_b, \zeta_\ell, \zeta_\tau$, are model dependent.

We study two extreme cases for the Z′ mass, following [100], namely $M_{Z'} = \sqrt{\frac{1}{2}} M_U < M_U$ and $M_{Z'} = \sqrt{\frac{3}{2}} M_U > M_U$. We also assume the LQ and Z′ are uniquely coupled to left-handed currents, i.e. $\zeta_q = \zeta_\ell = 1$ and $\zeta_t = \zeta_b = \zeta_\tau = 0$. With these definitions, Fig. 3 shows the effect of the Z′ on the $\tau\tau$ production cross-section, considering $g_U = 1$, $\beta_R = 0$, and different $g_{Z'}$ couplings. On the top panel, the cross-sections corresponding to the cases where $M_{Z'} = \sqrt{\frac{1}{2}} M_U$ are shown. As expected, the $\tau\tau$ production cross-section for the inclusive case (i.e., $g_{Z'} \neq 0$) is larger than that for the LQ-only non-res process ($g_{Z'} = 0$, depicted in blue). This effect increases with $g_{Z'}$ and, within the evaluated values, can exceed the LQ-only cross-section by up to two orders of magnitude. In contrast, a more intricate behaviour can be seen in the bottom panel of Fig. 3, which corresponds to $M_{Z'} = \sqrt{\frac{3}{2}} M_U$. Here, for low values of M_U, a similar

---

[4] Naively, the LQs are associated to the breaking of $SU(4) \to SU(3)_{[4]} \times U(1)_{B-L}$, the G′ arises from $SU(3)_{[4]} \times SU(3)' \to SU(3)_c$, and the Z′ comes from the breaking of $U(1)_{B-L} \times U(1)_{T_R^3} \to U(1)_Y$. Notice that the specific pattern of breaking, and the relations between the masses and couplings, are connected to the specific scalar potential used.
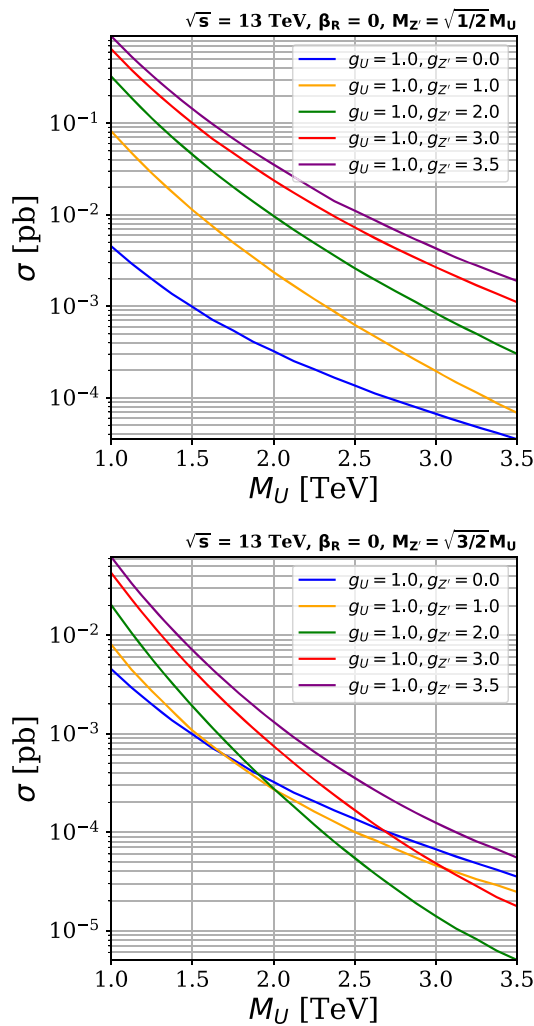
Fig. 3 $\tau\tau$ cross-section as a function of the LQ mass for different values of g_U and $g_{Z'}$. The estimates are performed at $\sqrt{s} = 13$ TeV, $\beta_R = 0$, $M_{Z'} = \sqrt{1/2}M_U$ (top), and $M_{Z'} = \sqrt{3/2}M_U$ (bottom)
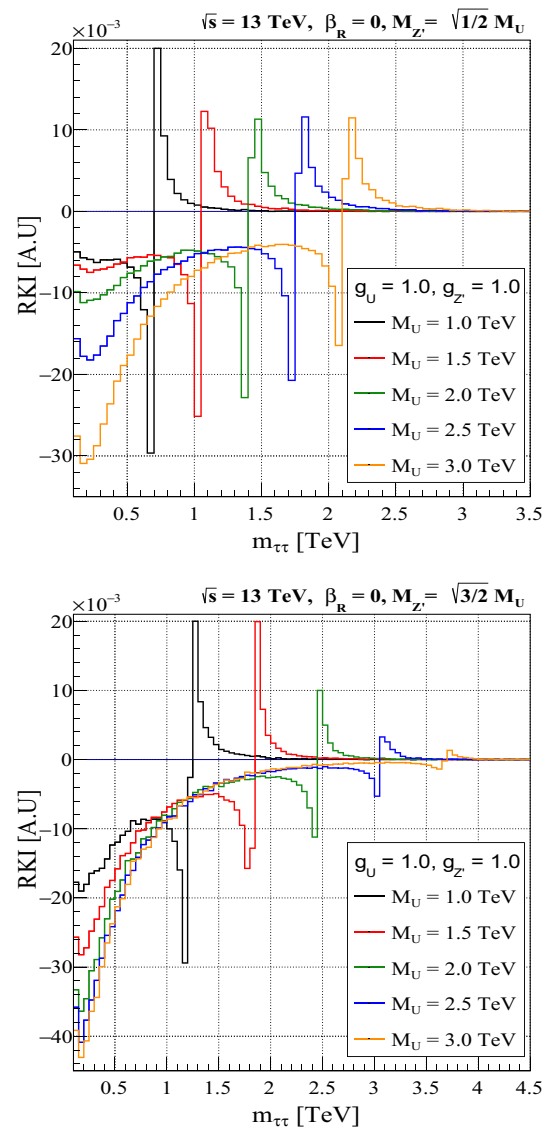

Fig. 4 The relative kinematic interference (RKI), as a function of the reconstructed mass of two taus, for different LQ masses. The studies are performed assuming $\sqrt{s} = 13$ TeV, $\beta_R = 0$, $g_U = 1.0$, $g_{Z'} = 1.0$, $M_{Z'} = \sqrt{1/2}M_U$ (top), and $M_{Z'} = \sqrt{3/2}M_U$ (bottom)

increase in the cross-section is observed. However, for higher values of M_U, the inclusive $pp \rightarrow \tau\tau$ cross-section is smaller than the LQ-only $\tau\tau$ cross-section. This behaviour suggests the presence of a dominant destructive interference at high masses, leaving its imprint on the results.

In order to further illustrate the effect, Fig. 4 shows the relative kinematic interference (RKI) as a function of the reconstructed invariant mass $m_{\tau\tau}$, for $g_{Z'} = 1$ and varying values of M_U. The RKI parameter is defined as

$$\text{RKI}(m_{\tau\tau}) = \frac{1}{\sigma_{\text{LQ}+Z'}} \left[\frac{d\sigma_{\text{LQ}+Z'}}{dm_{\tau\tau}} - \left(\frac{d\sigma_{\text{LQ}}}{dm_{\tau\tau}} + \frac{d\sigma_{Z'}}{dm_{\tau\tau}} \right) \right],$$

(3)

where σ_X is the production cross-section arising due to contributions from X particles. For example, $\sigma_{\text{LQ}+Z'}$ represents the inclusive cross-section where both virtual LQ and s-channel Z' exchange contribute. For both cases, we can

observe the presence of deep valleys in the RKI curves when $m_{\tau\tau} \rightarrow 0$, indicating destructive interference between the LQ and the Z' contributions. This interference generates a suppression of the differential cross-section for lower values of $m_{\tau\tau}$ and, therefore, in the integrated cross-section.

The observed interference effects are consistent with detailed studies on resonant and non-res $pp \rightarrow t\bar{t}$ production, performed in reference [101].

3 LQ search strategy and simulation

Our proposed analysis strategy utilizes single-LQ (i.e. $pp \rightarrow \tau$ LQ), double-LQ (i.e. $pp \rightarrow$ LQ LQ), and non-resonant LQ

production (i.e. $pp \rightarrow \tau\tau$) as shown in Fig. 1. At leading order in α_s, since we focus on $U_1 \rightarrow b\tau$ decays, the sLQ process results in the $b\tau\tau$ mode, the dLQ process results in the $bb\tau\tau$ mode, and the non-res process results in the $\tau\tau$ mode. Therefore, in all cases we obtain two τ leptons, with either 0, 1, or 2 b jets. The τ leptons decay to hadrons (τ_h) or semi-leptonically to electrons or muons (τ_ℓ, $\ell = e$ or μ). To this end, we study six final states: $\tau_h\tau_{h/\ell}$, $b\tau_h\tau_{h/\ell}$, and $bb\tau_h\tau_{h/\ell}$, which can be naively associated to non-res, sLQ and dLQ production, respectively. Nevertheless, experimentally it is possible for b jets to not be properly identified or reconstructed, leading, for instance, to a fraction of dLQ signal events falling into the $b\tau_h\tau_{h/\ell}$ and $\tau_h\tau_{h/\ell}$ categories. Similarly, soft jets can fake b jets, such that non-res processes can contribute to the $b\tau_h\tau_{h/\ell}$ and $bb\tau_h\tau_{h/\ell}$ final states. This kind of signal loss and mixing is taken into account in our analysis.[5]

The contributions of signal and background events are estimated using Monte Carlo (MC) simulations. We implemented the U_1 model from [51], adjusted to describe the lagrangian in Eqs. (1) and (2), using FeynRules (v2.3.43) [102,103]. The branching ratios and cross-sections have been calculated using MadGraph5_aMC (v3.1.0) [104,105], the latter at leading order in α_s. The corresponding samples are generated considering pp collisions at $\sqrt{s} = 13\,\text{TeV}$ and $\sqrt{s} = 13.6\,\text{TeV}$. All samples are generated using the NNPDF3.0 NLO [106] set for parton distribution functions (PDFs) and using the full amplitude square SDE strategy for the phase-space optimization due to strong interference effects with the Z′ boson. Parton level events are then interfaced with the PYTHIA (v8.2.44) [107] package to include parton fragmentation and hadronization processes, while DELPHES (v3.4.2) [108] is used to simulate detector effects, using the input card for the CMS detector geometric configurations, and for the performance of particle reconstruction and identification.

At parton level, jets and leptons are required to have a minimum transverse momentum (p_T) of 20 GeV, while b jets are required to have a minimum p_T of 30 GeV. Additionally, we constrain the pseudorapidity (η) to $|\eta| < 2.5$ for b jets and leptons, and $|\eta| < 5.0$ for jets. The production cross-sections shown in the bottom panel of Figs. 2 and 3 are obtained with the aforementioned selection criteria.

Table 1 shows the preliminary event selection criteria for each channel at analysis level. The channels are divided based on the multiplicity of b jets, $N(b)$, number of light leptons, $N(\ell)$, number of hadronic tau leptons, $N(\tau_h)$, and kinematic

---

[5] Note that further signal mixing can also occur at the event generation level by including terms at larger order in α_s. For example, in the non-res diagram in Fig. 1, one of the initial b could come from a $g \rightarrow b\bar{b}$ splitting, leading to non resonant production of $b\tau_h\tau_{h/\ell}$. Simulating and studying the role of such NLO contributions is outside the scope of this work.

Table 1 Preliminary event selection criteria used to filter events before feeding them to the BDT algorithm. A $\Delta R(p_i, p_j) > 0.3$ requirement is imposed between all pairs of reconstructed particle candidates p_i, p_j

Variable	Threshold					
	$\tau_h\tau_h$	$b\tau_h\tau_h$	$bb\tau_h\tau_h$	$\tau_h\tau_\ell$	$b\tau_h\tau_\ell$	$bb\tau_h\tau_\ell$
$N(b)$	$= 0$	$= 1$	≥ 2	$= 0$	$= 1$	≥ 2
$p_T(b)$	–	$\geq 30\,\text{GeV}$		–	$\geq 30\,\text{GeV}$	
$|\eta(b)|$	–	≤ 2.4		–	≤ 2.4	
$N(\ell)$	$= 0$			$= 1$		
$p_T(e)$	–			$\geq 35\,\text{GeV}$		
$p_T(\mu)$	–			$\geq 30\,\text{GeV}$		
$|\eta(\ell)|$	–			≤ 2.4		
$N(\tau_h)$	$= 2$			$= 1$		
$p_T(\tau_h)$	$\geq 50\,\text{GeV}$					
$|\eta(\tau_h)|$	≤ 2.3					
$\Delta R(p_i, p_j)$	≥ 0.3					

criteria based on η, p_T and spatial separation of particles in the detector volume ($\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$). The minimum p_T thresholds for leptons are chosen following references [62,63,67], based on experimental constrains associated to trigger performance. Following reference [109], we use a flat identification efficiency for b jets of 70% across the entire p_T spectrum with misidentification rate of 1%. These values correspond with the "medium working point" of the CMS algorithm to identify b jets, known as DeepCSV. We also explored the "Loose" ("Tight") working point using an efficiency of 85% (45%) and mis-identification rate of 10% (0.1%). The "medium working point" was selected as it gives the best signal significance for the analysis.

For the performance of τ_h identification in DELPHES, we consider the latest technique described in [110], which is based on a deep neural network (i.e. DeepTau) that combines variables related to isolation and τ-lepton lifetime as input to identify different τ_h decay modes. Following [110], we consider three possible DeepTau "working points": (i) the "Medium" working point of the algorithm, which gives a 70% τ_h-tagging efficiency and 0.5% light-quark and gluon jet mis-identification rate; (ii) the "Tight" working point, which gives a 60% τ_h-tagging efficiency and 0.2% light-quark and gluon jet mis-identification rate; and (iii) the "VTight" working point, which gives a 50% τ_h-tagging efficiency and 0.1% light-quark and gluon jet mis-identification rate. Similar to the choice of b-tagging working point, the choice of τ_h-tagging working point 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 [111–113].

Table 2 The number of simulated events for the signal and background samples

Sample	$t\bar{t}$	single t	VV	V+jets	signals
$N_{events} \times 10^{-6}$	24.31	11.50	32.35	39.45	0.60

After applying the preliminary selection criteria, the primary sources of background are production of top quark pairs ($t\bar{t}$), and single-top quark processes (single t), followed by production of vector bosons with associated jets from initial or final state radiation (V+jets), and pair production of vector bosons (VV). The number of simulated MC events used for each sample is shown in Table 2.

We use two different sets of signal samples. The first set includes various $\{M_U, g_U\}$ scenarios, for two different values of $\beta_R \in \{0, -1\}$. We generate signal samples for M_U values between 250 and 5000 GeV, in steps of 250 GeV. The considered g_U coupling values are between 0.25 and 3.5, in steps of 0.25. Although the signal cross-sections depend on both M_U and g_U, the efficiencies of our selections only depend on M_U (for all practical purposes) since the decay widths are relatively small compared to the mass of M_U ($\frac{\Gamma_U}{M_U} < 5\%$), and thus more sensitive to experimental resolution. In total there are 280 $\{M_U, g_U, \beta_R\}$ scenarios simulated for this first set of signal samples, and for each of these scenarios two subsets of samples are generated, which are used separately for the training and testing of the machine learning algorithm. The second set of signal samples is used to evaluate interference effects between LQs and the Z′ bosons in non-res production. Using benchmark values $g_U = 1.8$ and $\beta_R = 0$, we consider various $\{M_U, g_{Z'}\}$ scenarios for two different Z′ mass hypotheses, $(M_{Z'}/M_U)^2 \in \{\frac{1}{2}, \frac{3}{2}\}$. The M_U values vary between 500 and 5000 GeV, in steps of 250 GeV. The $g_{Z'}$ coupling values are between 0.25 and 3.5, in steps of 0.25. Therefore, in total there are 280 $\{M_U, g_{Z'}, (M_{Z'}/M_U)^2\}$ scenarios simulated for this second set of signal samples, and for each of these scenarios a total of 6.0×10^5 MC events are generated.

As noted previously, the simulated signal and background events are initially filtered using selections which are motivated by experimental constraints, such as the geometric constraints of the CMS detector, the typical kinematic thresholds for reconstruction of particle objects, and the available triggers. The remaining events after the preliminary event selection criteria are used to train and execute a BDT algorithm for each signal point in the $\{M_U, g_U\}$ space, in order to maximize the probability to detect signal amongst background events. The BDT algorithm is implemented using the `scikit-learn` [114] and `xgboost` (XGB) [115] python libraries. We use the the `XGBClassifier` class from the `xgboost` library, a 10-fold cross validation using

the `scikit-learn` method (`GridCV`[6]) for a grid in a hyperparameter space with 75, 125, 250, and 500 estimators, maximum depth in 3, 5, 7, 9, as well as learning rates of 0.01, 0.1, 1, and 10. For the cost function, we utilize the default mean square error (`MSE`). Additionally, we use the tree method based on the approximate greedy algorithm (histogram-optimized), referred to as `hist`, with a uniform sample method. These choices allow us to maximize the detection capability of the BDT algorithm by carefully tuning the hyperparameters, selecting an appropriate cost function, and utilizing an optimized tree construction method.

For each of the six analysis channels and $\{M_U, g_U\}$ signal point, the binary XGB classifier was trained (tested) with 20% (80%) of the simulated events, for each signal and background MC sample. Over forty kinematic and topological variables were studied as input for the XGB. These included the momenta of b jets and $\tau_{h,\ell}$ candidates; both invariant and transverse masses of pairs of τ objects and of b τ combinations; angular differences between b jets, between τ objects, and between the $\tau_{h,\ell}$ and b jets; and additional variables derived from the missing momentum in the events. After studying correlations between variables and their impact on the performance of the BDT, we found that only eight variables were necessary and responsible for the majority of the sensitivity of the analysis. The variable that provides the best signal to background separation is the scalar sum of the p_T of the final state objects (τ_h, $\tau_{h/\ell}$, and b jets) and the missing transverse momentum, referred to as S_T^{MET}:

$$S_T^{MET} = |\vec{p}_T^{\ miss}| + \sum_{\tau_h, \tau_{h/\ell}, b} |\vec{p}_T| \qquad (4)$$

The S_T^{MET} variable has been successfully used in LQ searches at the LHC, since it probes the mass scale of resonant particles involved in the production processes. Other relevant variables include the magnitude of the vectorial difference in p_T between the two lepton candidates ($|\Delta \vec{p}_T|_{\tau_h \tau_{h/\ell}}$), the $\Delta R_{\tau_h \tau_{h/\ell}}$ separation between them, the reconstructed dilepton mass $m_{\tau_h \tau_{h/\ell}}$, and the product of their electric charges ($Q_{\tau_h} \times Q_{\tau_{h/\ell}}$). We also use the $|\Delta \vec{p}_T|$ between the τ_h candidate and $\vec{p}_T^{\ miss}$, and (if applicable) the $|\Delta \vec{p}_T|$ between the τ_h candidate and the leading b jet. For the final states including two τ_h candidates, the one with the highest p_T is used.

Figure 5 shows some relevant topological distributions, including S_T^{MET} on the top, for the b $\tau_h \tau_\ell$ category. In the Figure we include all signal production modes to this channel, with each component weighted with respect to their total contribution to the combined signal. The combined signal distribution is normalised to unity. We also show all background processes contributing to this channel, each of them

---

[6] GridCV is a method that allows to find the best combination of hyperparameter values for the model, as this choice is crucial to achieve an optimal performance.
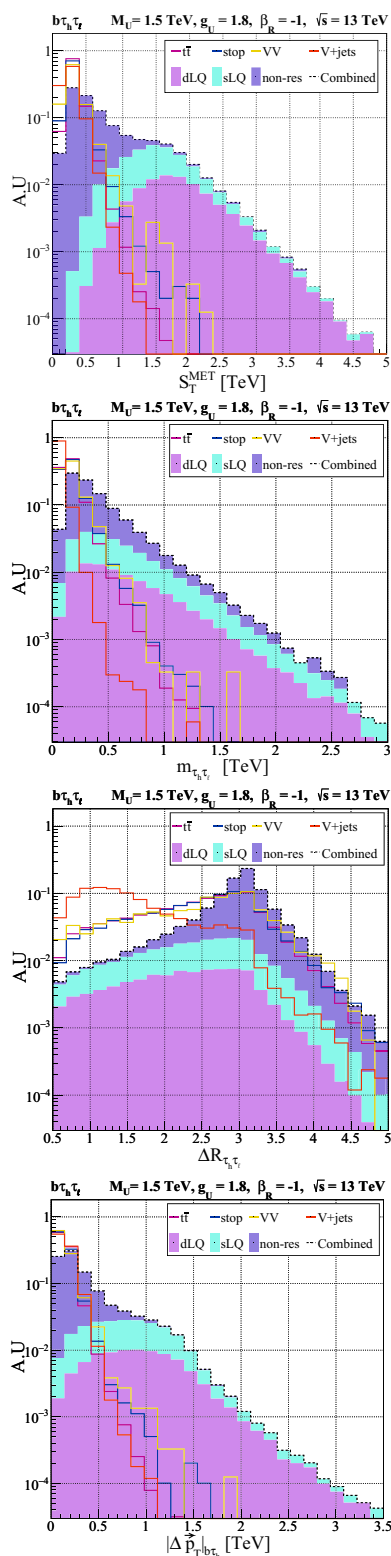
Fig. 5 S_T^{MET}, $m_{\tau_h \tau_\ell}$, $\Delta R_{\tau_h \tau_\ell}$, $|\Delta \vec{p}_T|_{b\tau_h}$ signal and background distributions for the $b\tau_h \tau_\ell$ channel. The signal distributions are generated for a benchmark sample with LQ mass of 1.5 TeV maximally coupled to right-handed currents. The combined distribution (shown as a stacked histogram) is the sum of the distributions, correctly weighted according to their respective cross-sections, assuming a coupling $g_U = 1.8$

individually normalised to unity. We find that the combined signal is dominated by sLQ production for large values of S_T^{MET}, while non-res production dominates for small S_T^{MET}. Interestingly, the backgrounds also sit at low S_T^{MET} values, since S_T^{MET} is driven by the mass scale of the SM particles being produced, in this case top quarks and Z/W bosons. This suggest that the sLQ and dLQ signals can indeed be separated from the SM background. As expected, the S_T^{MET} sLQ and dLQ signal distributions have a mean near M_U, representative of resonant production, and a broad width as expected for large mass M_U hypotheses when information about the z-components of the momenta of objects is not utilised in the S_T^{MET} calculation.

Figure 5 (second from the top) shows the reconstructed mass of the ditau system, for the $b\tau_h \tau_\ell$ search channel. Since the two τ candidates in signal events arise from different production vertices (e.g., each τ candidate in dLQ production comes from a different LQ decay chain), the ditau mass distribution for signal scales as $m_{\tau_h \tau_\ell} \sim p_T(\tau_h) + p_T(\tau_\ell)$, and thus has a tail which depends on M_U and sits above the expected SM spectrum. On the other hand, the SM $m_{\tau_h \tau_\ell}$ distributions sit near $m_{Z/W}$ since the τ candidates in SM events arise from Z/W decays.

Figure 5 (third from the top) shows the $\Delta R_{\tau_h \tau_\ell}$ distribution for the $b\tau_h \tau_\ell$ channel. In the case of the $pp \rightarrow \tau\tau$ non-res signal distribution, the two τ leptons must be back-to-back to preserve conservation of momentum. Therefore, the visible τ candidates, τ_h and τ_ℓ, give rise to a $\Delta R_{\tau_h \tau_\ell}$ distribution that peaks near π radians. In the case of sLQ production, although the LQ and associated τ candidate must be back-to-back, the second τ candidate arising directly from the decay of the LQ does not necessarily move along the direction of the LQ (since the LQ also decays to a b quark). As a result, the $\Delta R_{\tau_h \tau_\ell}$ distribution for the sLQ signal process is smeared out, is broader, and has a mean below π radians. On the other hand, the τ_h candidate in $t\bar{t}$ events is often a jet being misidentified as a genuine τ_h. When this occurs, the fake τ_h candidate can arise from the same top quark decay chain as the τ_ℓ candidate, thus giving rise to small $\Delta R_{\tau_h \tau_\ell}$ values. This difference in the signal and background distributions provides a nice way for the ML algorithm to help decipher signal and background processes.

As noted above, the $|\Delta \vec{p}_T|$ distribution between the visible τ candidates and the b-quark jets is an important variable to help the BDT distinguish between signal and background processes. The discriminating power can be seen in Fig. 5 (bottom), which shows the $|\Delta \vec{p}_T|$ between the τ_h and b-jet candidate of the $b\tau_h \tau_\ell$ channel. In the case of dLQ production, the b quarks and τ leptons from the LQ $\rightarrow$ $b\tau$ decay acquire transverse momentum of $p_T \sim M_U/2$. However, when the τ lepton decays hadronically (i.e. $\tau \rightarrow \tau_h \nu$), a large fraction of the momentum is lost to the neutrino. Therefore, the $|\Delta \vec{p}_T|_{b\tau_h}$ distribution for the dLQ (and sLQ)
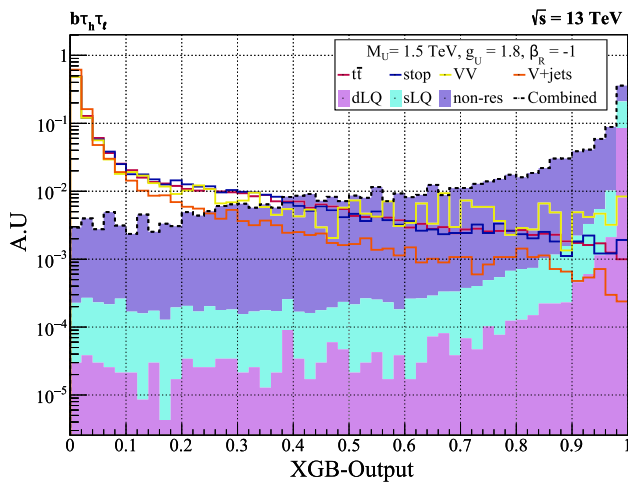
Fig. 6 Postfit XGB-output normalised distribution in the $b\,\tau_h\tau_\ell$ channel, for LQ mass of 1.5 TeV, constant coupling $g_U = 1.8$, and maximally coupled to right-handed currents

process peaks below $M_U/2$. On the other hand, for a background process such as V+jets, the b jet arises due to initial state radiation, and thus must balance the momentum of the associated vector boson (i.e. $p_T(b) \sim p_T(V) \sim m_V$). Since the visible τ candidate is tyically produced from the V boson decay chain, its momentum (on average) is approximately $p_T(\tau_h) \sim p_T(V)/4 \sim m_V/4$. Therefore, to first order, the $|\Delta \vec{p}_T|$ distribution for the V+jets background is expected to peak below the m_V mass.

Lets us turn to the results of the $b\tau_h\tau_\ell$ BDT classifier, which is shown in Fig. 6 for the different signal production modes and backgrounds. Similar to Fig. 5, the distribution for each individual signal production mode is weighted with respect to their total contribution to the combined signal. The background distributions and combined signal distribution are normalized to an area under the curve of unity. Figure 6 shows the XGB distributions for a signal benchmark point with $M_U = 1.5$ TeV, $g_U = 1.8$, and $\beta_R = -1$. The XGB output is a value between 0 and 1, which quantifies the likelihood that an event is either signal-like (XGB output near 1) or background-like (XGB output near 0). We see that the presence of the sLQ and dLQ production modes is observed as an enhancement near a XGB output of unity, while the backgrounds dominate over the low end of the XGB output spectrum, especially near zero. In fact, over eighty percent of the sLQ and dLQ distributions reside in the last two bins, XGB output greater than 0.96, while more than sixty percent of the backgrounds fall in the first two bins, XGB output less than 0.04. It is also interesting to note that in comparison to the sLQ and dLQ distributions in Fig. 6, non-res is broader and not as narrowly peaked near XGB output of 1, which is expected due to the differences in kinematics described

above. Overall, if we focus on the last bin in this distribution, we find approximately 0.2% of the background, in contrast to 22% of the non-res, 78% of the sLQ, and 91% of the dLQ signal distributions. These numbers highlight the effectiveness of the XGB output in reducing the background in the region where the signal is expected.

The output signal and background distributions of the XGB classifier, normalised to their cross section times preselection efficiency times luminosity, are used to perform a profile binned likelihood statistical test in order to determine the expected signal significance. The estimation is performed using the RooFit [116] package, following the same methodology as in Refs. [117–132]. The value of the significance (Z_{sig}) is measured using the probability to obtain the same outcome from the test statistic in the background-only hypothesis, with respect to the signal plus background hypothesis. This allows for the determination of the local p-value and thus the calculation of the signal significance, which 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 as nuisance parameters, considering log-priors for normalization and Gaussian priors for shape uncertainties. Our consideration of systematic uncertainties includes both experimental and theoretical effects, focusing on the dominant sources of uncertainty. Following [133], we consider a 3% systematic uncertainty on the measurement of the integrated luminosity at the LHC. A 5% uncertainty arises due to the choice of the parton distribution function used for the MC production, following the PDF4LHC prescription [134]. The chosen PDF set only has an effect on the overall expected signal and background yields, but the effect on the shape of the XGB output distribution is negligible. Reference [110] reports a systematic uncertainty of 2–5%, depending on the p_T and η of the τ_h candidate. Therefore, we utilize a conservative 5% uncertainty per τ_h candidate, independent of p_T and η, which is correlated between signal and background processes with genuine τ_h candidates, and correlated across XGB bins for each process. We assumed a 5% τ_h energy scale uncertainty, independent of p_T and η, following the CMS measurements described in [110]. Finally, we assume a conservative 3% uncertainty per b-jet candidate, following reference [135], and an additional 10% uncertainty in all the background predictions to account for possible mismodeling by the simulated samples. The uncertainties on the background estimates are typically derived from collision data using dedicated control samples that have negligible signal contamination and are enriched with events from the specific targeted background. The systematic uncertainties on the background estimates are treated as uncorrelated between background processes.
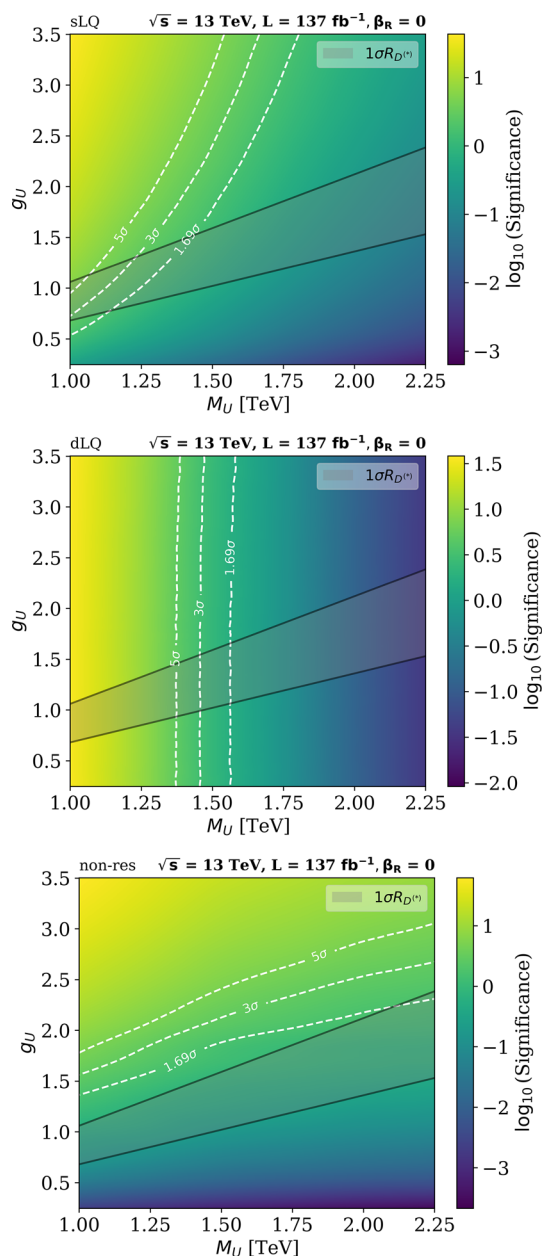
Fig. 7 Signal significance for different coupling scenarios and LQ masses, without right-handed currents, using the combination of all search channels. The results pertaining to sLQ, dLQ and non-res production are displayed respectively from the top. These results are for $\sqrt{s} = 13\,\text{TeV}$ and 137 fb^{-1}

4 Results

The expected signal significance for sLQ, dLQ and non-res production, and their combination, is presented in Fig. 7. Here, the significance is shown as a heat map in a two dimensional plane of g_U versus M_U, considering exclusive couplings to left-handed currents, i.e. BR(LQ $\rightarrow$ b τ) $= \frac{1}{2}$. The dashed lines show the contours of constant signal significance. The 1.69σ contour represents exclusion at 95%

confidence level, and the 3-5σ contours represent potential discovery. The grey band defines the set of $\{M_U, g_U\}$ values that can explain the B-meson anomalies, $C_U \sim 0.01$ for this scenario. The estimates are performed under the conditions for the second run, RUN-II, of the LHC ($\sqrt{s} = 13\,\text{TeV}$ and $L = 137\,\text{fb}^{-1}$). We find that the dLQ interpretation plot (Fig. 7 second from the top) does not depend on g_U, which is expected due to dLQ production arising exclusively from interactions with gluons. For this reason, the dLQ production process provides the best mode for discovery when g_U is small. On the other hand, the non-res channel is more sensitive to changes in the coupling parameter g_U, as its production cross-section depends on g_U^4. Therefore, the non-res production process provides the best mode for discovery when g_U is large. These results confirm the expectations from previous analyses (see for instance [49]), in the sense that the dLQ and non-res processes complement each other nicely at low and high g_U scenarios. The sLQ channel combines features from both the dLQ and non-res channels, in principle making it an interesting option to explore different scenarios and gain a better understanding of LQ properties, but the evolution of the signal significance in the full phase space is more complicated as it involves resonant LQ production with a cross-section that depends non-trivially on M_U, g_U, and the LQ coupling to gluons. However, Fig. 7 shows that the sLQ production process can provide complementary and competitive sensitivity to the non-res and dLQ processes, in certain parts of the phase space.

The top panel of Fig. 8 presents the sensitivity of all signal production processes combined, and compares our expected exclusion region with the latest one from the ATLAS Collaboration [70]. The comparison suggests that our proposed analysis strategy provides better sensitivity than current methods being carried out at ATLAS, especially at large values of g_U. In particular, we find that with the pp data already available from RUN-II, our expected exclusion curves begin to probe solutions to the B-anomalies for LQ masses up to 2.25 TeV.

Figure 8 shows the expected signal significance considering BR(LQ $\rightarrow$ b τ) $= 1$, in order to compare our analysis with the corresponding results from the CMS [64] and ATLAS [71] Collaborations. Let us emphasize again that BR(LQ $\rightarrow$ b τ) depends on β_R, as illustrated on the top panel of Fig. 2. Thus, although the BR(LQ $\rightarrow$ b τ) $= 1$ scenario is a possible physical case, it does not solve the observed anomalies in the $R_{D^{(*)}}$ ratios, as it corresponds to the case where LQs couple exclusively to right-handed currents.

With this in mind, the scenario studied by CMS in [64] considers couplings only to left-handed currents, setting artificially the condition BR(LQ $\rightarrow$ b τ) $= 1$. In order to compare, we scale the efficiency×acceptance of our selection criteria for $\beta_R = 0$, by a factor of 2.0 for sLQ and 4.0 for dLQ. According to Fig. 8, the ML approach that we have followed
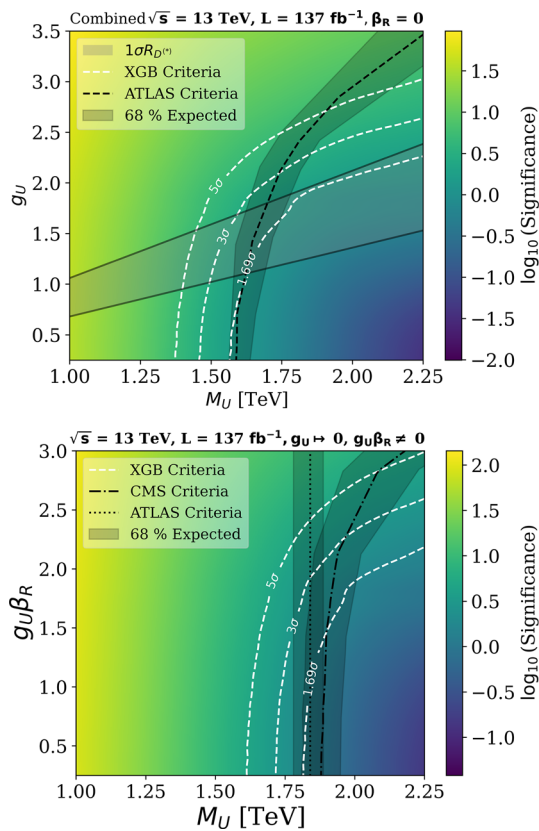
Fig. 8 The top (bottom) panel shows signal significance comparison with ATLAS [70] (CMS and ATLAS [64,71]) background only hypothesis, for the combination of all channels, with uniquely coupling to left-handed (right-handed) currents. The estimates are performed at $\sqrt{s} = 13$ TeV and 137 fb^{-1}
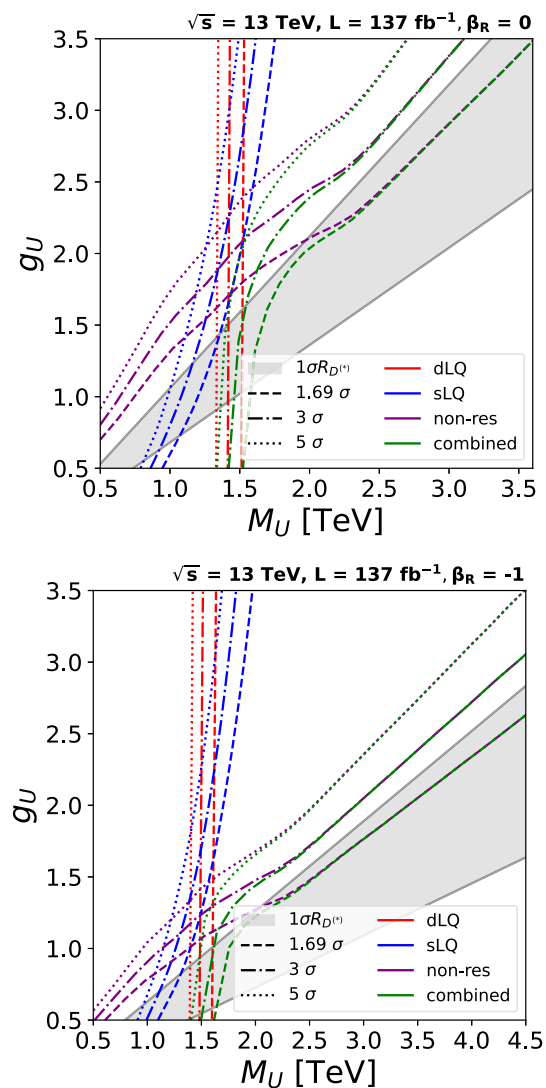

Fig. 9 Signal significance for different coupling scenarios and LQ masses for all channels. This plot summarizes our results with $\beta_R = 0$ (without right-handed currents) and $\beta_R = -1$ (maximally coupled to right-handed currents). The estimates are performed at $\sqrt{s} = 13$ TeV and 137 fb^{-1}

again suggests an optimisation of the signal and background separation, having the potential of improving the regions of exclusion (1.69 σ) with respect to that of CMS. In the bottom panel of the figure we have also included a similar exclusion by ATLAS [71]. However, since ATLAS only considers dLQ production in the analysis, the results are not entirely comparable, so are included only as a reference.

We now turn to the role of β_R, and our capacity of probing the regions solving the B-meson anomalies. Figure 9 shows the maximum significant contours, under LHC RUN-II conditions, for the different LQ production mechanisms and their combination, considering scenarios with only left-handed currents ($\beta_R = 0$, top) and with maximal right-handed currents ($\beta_R = -1$, bottom). We find a noticeable improvement in signal significance in all channels when taking $\beta_R = -1$, as is expected from the increase in BR(LQ $\rightarrow$ b τ) branching ratio and production cross-sections (see Fig. 2). However, the region solving the B-meson anomalies also changes, preferring lower values of g_U, such that in both cases we find ourselves just starting to probe this region at large M_U.

The combined significance contours for the different BR scenarios that have been considered is presented in Fig. 10.

These contours illustrate the regions of exclusion for the three cases of interest, namely exclusive left-handed currents (BR(LQ $\rightarrow$ b τ) = $\frac{1}{2}$, $\beta_R = 0$), maximal left and right couplings (BR(LQ $\rightarrow$ b τ) = $\frac{2}{3}$, $\beta_R = -1$), and exclusive right-handed currents (BR(LQ $\rightarrow$ b τ) = 1, $g_U \rightarrow 0$, $g_U \beta_R = 1$). For small g_U, we find that the exclusive right-handed scenario is most sensitive, while the exclusive left-handed case is the worst. The reason for this is that this region is excluded principally by dLQ production, such that having the largest branching ratio is crucial in order to have a large number of events. For larger couplings, both exclusive scenarios end up having similar exclusion regions, with the $\beta_R = -1$ case being significantly more sensitive. The reason in this case is that the exclusion is dominated by
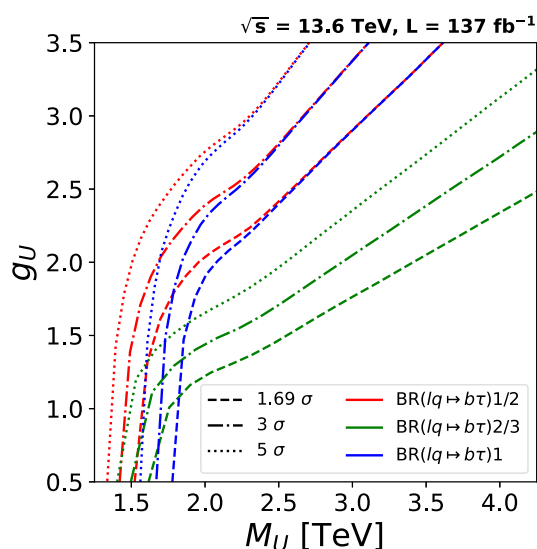
Fig. 10 Signal significance for different coupling scenarios and LQ masses, considering the case without coupling to right-handed currents $BR(LQ \to b\tau) = \frac{1}{2}$, the case maximally coupled to right- and left-handed currents $BR(LQ \to b\tau) = \frac{2}{3}$, and the case uniquely coupled to right-handed currents $BR(LQ \to b\tau) = 1$. The estimates are performed at $\sqrt{s} = 13\,\text{TeV}$ and $137\,\text{fb}^{-1}$

non-res, which has a much larger production cross-section if both currents are turned on.

In order to finalise our analysis of the LQ-only model, we show in Fig. 11 the expected combined significance in the relatively near future. For this, considering $\sqrt{s} = 13.6\,\text{TeV}$, we show contours for the sensitivity corresponding to integrated luminosities of $137\,\text{fb}^{-1}$, $300\,\text{fb}^{-1}$, and $3000\,\text{fb}^{-1}$, for scenarios with only left-handed currents (top) and with maximal coupling to right-handed currents (bottom). Note that for $\beta_R = 0$ ($\beta_R = -1$), couplings g_U close to 3.18 (1.85) and $M_U = 5.0\,\text{TeV}$ can be excluded with 1.69σ significance for the high luminosity LHC era, allowing us to probe the practically the entirety of the B-meson anomaly favored region. Note that the background yields for the high luminosity LHC might be larger due to pileup effects. Nevertheless, as it was mentioned in Sect. 3, we have included a conservative 10% systematic uncertainty associated with possible fluctuations on the background estimations. Although effects from larger pileup might be significant, they can be mitigated by improvements in the algorithms for particle reconstruction and identification, and also on the data-analysis techniques.

As commented on the Introduction, non-res production can be significantly affected by the presence of a companion Z′, which provides additional s-channel diagrams that add to the total cross-section and can interfere destructively with the LQ t-channel process (see Figs. 3 and 4). From our previous results, we see that non-res always is of high importance in determining the exclusion region, particularly at large M_U
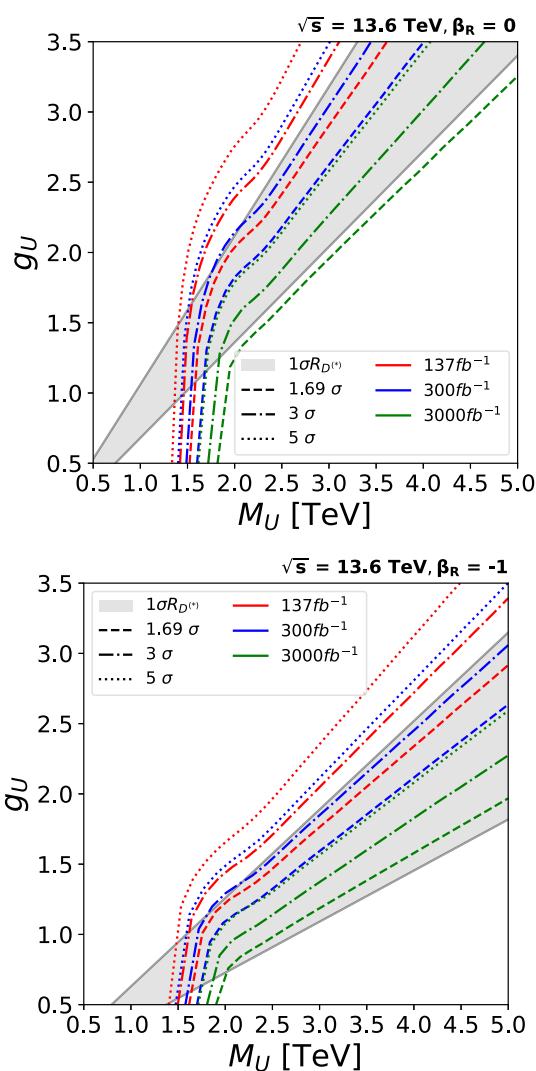

Fig. 11 Projected signal significance for different coupling scenarios and LQ masses maximally coupled to right-handed currents. The estimates are performed at $\sqrt{s} = 13.6\,\text{TeV}$, $137\,\text{fb}^{-1}$, $300\,\text{fb}^{-1}$ and $3000\,\text{fb}^{-1}$

and g_U, meaning it is crucial to understand how this role is affected in front of a Z′ with similar mass.

The change in sensitivity on the non-res signal significance due this interference effect with the Z′ boson is shown in Fig. 12. We consider two opposite cases for the Z′ mass: $M_{Z'}^2 = M_U^2/2$ (top) and $M_{Z'}^2 = 3\,M_U^2/2$ (bottom). Our results are shown on the $g_{Z'}$ - M_U plane, for a fixed $g_U = 1.8$ and $\beta_R = 0$. For the $M_{Z'}^2 = M_U^2/2$ scenario, there is an overall increase in the total cross-section, with a larger $g_{Z'}$ implying a larger sensitivity. This means that our ability to probe smaller values of g_U could be enhanced, as a given observation would be reproduced with both a specific g_U and vanishing $g_{Z'}$, or a smaller g_U with large $g_{Z'}$. Thus, for a large enough $g_{Z'}$, it could be possible to enhance non-res to the point that the entire region favoured by B-anomalies could
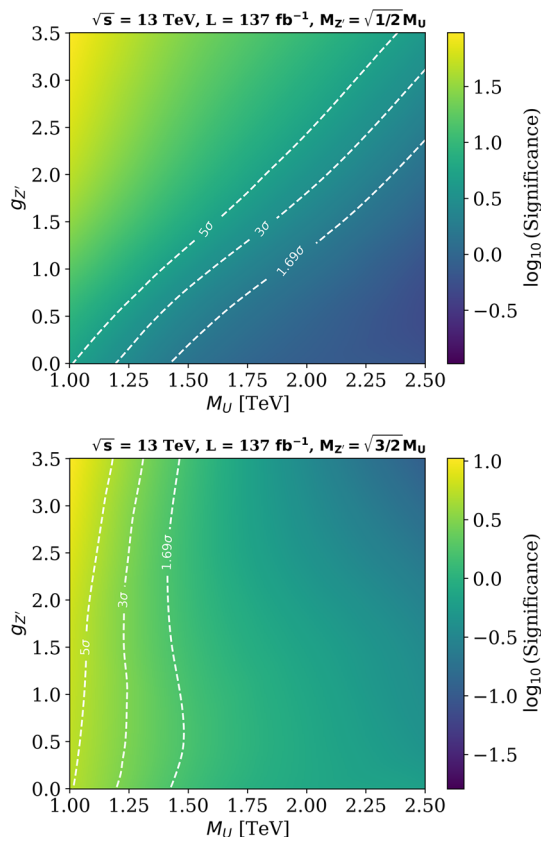
Fig. 12 Change on the non-res signal significance for different Z' coupling scenarios and LQ masses. The estimates are performed at $\sqrt{s} = 13.0$ TeV, $\beta_R = 0$, $g_U = 1.8$, $M_{Z'} = \sqrt{1/2}M_U$ (top), and $M_{Z'} = \sqrt{3/2}M_U$ (bottom)

be ruled out. In contrast, for $M_{Z'}^2 = 3\,M_U^2/2$ the cross-section is strongly affected by the large destructive interference, such that a larger $g_{Z'}$ does not necessarily imply an increase in sensitivity. In fact, as can be seen in the bottom panel, for large M_U the significance is reduced as $g_{Z'}$ increases, leading to the opposite conclusion than above, namely, that a large $g_{Z'}$ could reduce the effectiveness of non-res.

The impact of the above can be seen in Fig. 13, which shows our previous sensitivity curves on the $M_U - g_U$ plane, but this time with a Z' contribution to non-res. We use the same values of $M_{Z'}$ as before, but fix $g_{Z'} = 3.5$. For smaller $M_{Z'}$ (top), the non-res contribution is enhanced so much, that both sLQ and dLQ play no role whatsoever in determining the exclusion region. We find that, for small g_U, the sensitivity is dominated by Z' production such that, since M_U is related to $M_{Z'}$, LQ masses up to ~ 3 TeV are excluded. This bound is slightly relaxed for larger values of g_U, which is attributed to destructive interference effects due to an increased LQ contribution.

The bottom panel of Fig. 13 shows that case where $M_{Z'}$ is larger than M_U. As expected from our previous discussion, the behaviour and impact of non-res is modified. For small
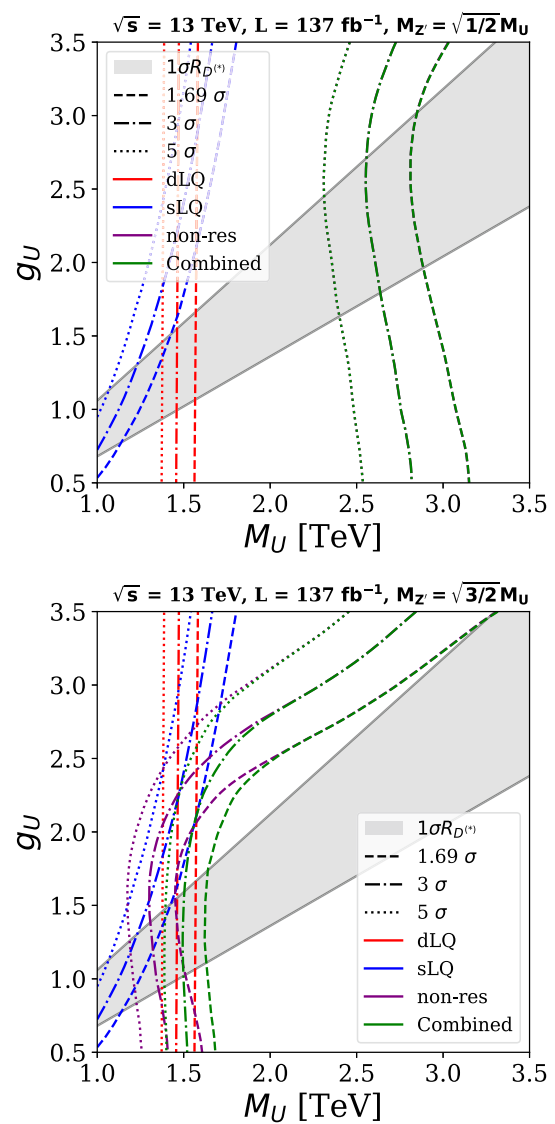

Fig. 13 Signal significance for different coupling scenarios and LQ masses, for all channels, with an additional Z' contribution to non-res production. We set $\beta_R = 0$ and $g_{Z'} = 3.5$, taking $M_{Z'}^2$ equal to $M_U^2/2$ ($3M_U^2/2$) on the top (bottom) panel

g_U, we again have the pure Z' production dominating the non-res cross-section, leading to a null sensitivity on g_U, similar to what happens in dLQ. In contrast, for very large g_U, we find that the pure LQ non-res production is the one that dominates, and we recover sensitivity regions with a slope similar to those shown in Figs. 7, 8, 9, 10 and 11, shifted towards larger values of g_U. For intermediate values of this coupling, the destructive interference have an important effect again, twisting the exclusion region slightly towards the left. Still, even in this case, we find that sLQ plays a marginal role in defining the combined exclusion region, and that the final result again depends primarily on dLQ and non-res production.

5 Discussion and conclusions

Experimental searches for LQs with preferential couplings to third generation fermions are currently of great interest due to their potential to explain observed tensions in the $R_{(D)}$ and $R_{(D*)}$ decay ratios of B mesons with respect to the SM predictions. Although the LHC has a broad physics program on searches for LQs, it is very important to consider the impact of each search within wide range of different theoretical assumptions within a specific model. In addition, in order to improve the sensitivity to detect possible signs of physics beyond the SM, it is also important to strongly consider new computational techniques based on machine learning (ML). Therefore, we have studied the production of U_1 LQs with preferential couplings to third generation fermions, considering different couplings, masses and chiral currents. These studies have been performed considering p p collisions at $\sqrt{s} = 13$ TeV and 13.6 TeV and different luminosity scenarios, including projections for the high luminosity LHC. A ML algorithm based on boosted decision trees is used to maximize the signal significance. The signal to background discrimination output of the algorithm is taken as input to perform a profile binned-likelihood test statistic to extract the expected signal significance.

The expected signal significance for sLQ, dLQ and non-res production, and their combination, is presented as contours on a two dimensional plane of g_U versus M_U. We present results for the case of exclusive couplings to left-handed, mixed, and exclusive right-handed currents. For the first two, the region of the phase space that could explain the B meson anomalies is also presented. We confirm the findings of previous works that the largest production cross-section and best overall significance comes from the combination of dLQ and non-res production channels. We also find that the sensitivity to probe the parameter space of the model is highly dependent on the chirality of the couplings. Nevertheless, the region solving the B-meson anomalies also changes with each choice, such that in all evaluated cases we find ourselves just starting to probe this region at large M_U.

Our studies compare our exclusion regions with respect to the latest reported results from the ATLAS and CMS Collaborations. The comparison suggests that our ML approach has a better sensitivity than the standard cut-based analyses, especially at large values of g_U. In addition, our projections for the HL-LHC cover the whole region solving the B-anomalies, for masses up to 5.00 TeV.

Finally, we consider the effects of a companion Z$'$ boson on non-res production. We find that such a contribution can have a considerable impact on the LQ sensitivity regions, depending on the specific masses and couplings. In spite of this, we still consider non-res production as an essential channel for probing LQs in the future.

References

1. J.C. Pati, A. Salam, Lepton number as the fourth color. Phys. Rev. D **10**, 275–289 (1974). (**Erratum: Phys. Rev. D 11, 703-703 (1975)**)
2. H. Georgi, S.L. Glashow, Unity of all elementary particle forces. Phys. Rev. Lett. **32**, 438–441 (1974)
3. B. Schrempp, F. Schrempp, Light leptoquarks. Phys. Lett. B **153**, 101–107 (1985)
4. M. Leurer, A comprehensive study of leptoquark bounds. Phys. Rev. D **49**, 333–342 (1994)
5. S. Davidson, D.C. Bailey, B.A. Campbell, Model independent constraints on leptoquarks from rare processes. Z. Phys. C **61**, 613–644 (1994)
6. M. Leurer, Bounds on vector leptoquarks. Phys. Rev. D **50**, 536–541 (1994)
7. J.L. Hewett, T.G. Rizzo, Much ado about leptoquarks: a comprehensive analysis. Phys. Rev. D **56**, 5709–5724 (1997)
8. F.S. Queiroz, K. Sinha, A. Strumia, Leptoquarks, dark matter, and anomalous LHC events. Phys. Rev. D **91**(3), 035006 (2015)
9. I. Doršner, S. Fajfer, A. Greljo, J.F. Kamenik, N. Košnik, Physics of leptoquarks in precision experiments and at particle colliders. Phys. Rep. **641**, 1–68 (2016)
10. LHCb Collaboration, Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays. Phys. Rev. Lett. **113**(2014)
11. LHCb Collaboration, Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays. JHEP **08**, 055 (2017)
12. LHCb Collaboration, Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays. Phys. Rev. Lett. **122**(19), 191801 (2019)
13. LHCb Collaboration, Test of lepton universality in beauty-quark decays. Nat. Phys. **18**(3), 277–282 (2022)

14. B.B. Collaboration, Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays. Phys. Rev. Lett. **109**, 101802 (2012)

15. B.B. 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**(7), 072012 (2013)

16. Belle Collaboration. Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method (2019)

17. B. 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)

18. B. 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)

19. B. 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)

20. B. 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)

21. 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**(11), 111803 (2015). [Erratum: Phys.Rev.Lett. 115, 159901 (2015)]

22. 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)]

23. 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)

24. LHCb Collaboration, Test of lepton flavor universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays. Phys. Rev. D **97**(7), 072013 (2018)

25. LHCb Collaboration. Measurement of the ratios of branching fractions $\mathcal{R}(D^*)$ and $\mathcal{R}(D^0)$ (2023)

26. G. Hiller, M. Schmaltz, R_K and future $b \to s\ell\ell$ physics beyond the standard model opportunities. Phys. Rev. D **90**, 054014 (2014)

27. B. Gripaios, M. Nardecchia, S.A. Renner, Composite leptoquarks and anomalies in B-meson decays. JHEP **05**, 006 (2015)

28. R. Alonso, B. Grinstein, J.M. Camalich, Lepton universality violation and lepton flavor conservation in B-meson decays. JHEP **10**, 184 (2015)

29. L. Calibbi, A. Crivellin, T. Ota, Effective field theory approach to $b \to s\ell\ell^{(\prime)}$, $B \to K^{(*)}\nu\bar{\nu}$ and $B \to D^{(*)}\tau\nu$ with third generation couplings. Phys. Rev. Lett. **115**, 181801 (2015)

30. S. Fajfer, N. Košnik, Vector leptoquark resolution of R_K and $R_{D^{(*)}}$ puzzles. Phys. Lett. B **755**, 270–274 (2016)

31. M. Bauer, M. Neubert, Minimal leptoquark explanation for the $R_{D^{(*)}}$, R_K, and $(g-2)_\mu$ anomalies. Phys. Rev. Lett. **116**(14), 141802 (2016)

32. D. Bečirević, N. Košnik, O. Sumensari, R. Zukanovich-Funchal, Palatable leptoquark scenarios for lepton flavor violation in exclusive $b \to s\ell_1\ell_2$ modes. JHEP **11**, 035 (2016)

33. A. Crivellin, D. Muller, T. Ota, Simultaneous explanation of $R(D^*)$ and $b \to \mu^+\mu^-$:the last scalar leptoquark standing. JHEP **09**, 040 (2017)

34. G. D'Amico, M. Nardecchia, P. Panci, F. Sannino, A. Strumia, R. Torre, A. Urbano, Flavour anomalies after the R_{K^*} measurement. JHEP **09**, 010 (2017)

35. G. Hiller, I. Nisandzic, R_K and R_{K^*} beyond the standard model. Phys. Rev. D **96**(3), 035003 (2017)

36. D. Buttazzo, A. Greljo, G. Isidori, D. Marzocca, B-physics anomalies: a guide to combined explanations. JHEP **11**, 044 (2017)

37. D. Bečirević, I. Doršner, S. Fajfer, N. Košnik, D.A. Faroughy, O. Sumensari, Scalar leptoquarks from grand unified theories

38. to accommodate the B-physics anomalies. Phys. Rev. D **98**(5), 055003 (2018)

38. C. Cornella, J. Fuentes-Martín, G. Isidori, Revisiting the vector leptoquark explanation of the B-physics anomalies. JHEP **07**, 168 (2019)

39. A. Angelescu, D. Bečirević, D.A. Faroughy, F. Jaffredo, O. Sumensari, Single leptoquark solutions to the B-physics anomalies. Phys. Rev. D **104**(5), 055017 (2021)

40. G. Belanger et al., Leptoquark manoeuvres in the dark: a simultaneous solution of the dark matter problem and the $R_{D^{(*)}}$ anomalies. JHEP **02**, 042 (2022)

41. J. Aebischer, G. Isidori, M. Pesut, B.A. Stefanek, F. Wilsch, Confronting the vector leptoquark hypothesis with new low- and high-energy data. **83**(2), 153 (2023)

42. LHCb Collaboration. Test of lepton universality in $b \to \ell^+\ell^-$ decays (2022)

43. LHCb Collaboration. Measurement of lepton universality parameters in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays (2022)

44. A. Greljo, J. Salko, A. Smolković, P. Stangl, Rare b decays meet high-mass Drell-Yan. JHEP **5**, 2023 (2023)

45. M. Ciuchini, M. Fedele, E. Franco, A. Paul, L. Silvestrini, M. Valli, Constraints on lepton universality violation from rare b decays. Phys. Rev. D **107**, 055036 (2023)

46. B. Diaz, M. Schmaltz, Y. Zhong, The leptoquark Hunter's guide: Pair production. JHEP **10**, 097 (2017)

47. I. Doršner, A. Greljo, Leptoquark toolbox for precision collider studies. JHEP **05**, 126 (2018)

48. N. Vignaroli, Seeking leptoquarks in the $t\bar{t}$ plus missing energy channel at the high-luminosity lhc. Phys. Rev. D **99**, 035021 (2019)

49. M. Schmaltz, Y. Zhong, The leptoquark Hunter's guide: large coupling. JHEP **01**, 132 (2019)

50. A. Biswas, D. Kumar-Ghosh, N. Ghosh, A. Shaw, A.K. Swain, Collider signature of U_1 leptoquark and constraints from $b \to c$ observables. J. Phys. G **47**(4), 045005 (2020)

51. M.J. Baker, J. Fuentes-Martín, G. Isidori, M. König, High-p_T signatures in vector–leptoquark models. Eur. Phys. J. C **79**(4), 334 (2019)

52. U. Haisch, G. Polesello, Resonant third-generation leptoquark signatures at the Large Hadron Collider. JHEP **05**, 057 (2021)

53. A. Bhaskar, T. Mandal, S. Mitra, M. Sharma, Improving third-generation leptoquark searches with combined signals and boosted top quarks. Phys. Rev. D **104**(7), 075037 (2021)

54. J. Bernigaud, M. Blanke, I.M. Varzielas, J. Talbert, J. Zurita, LHC signatures of τ-flavoured vector leptoquarks. JHEP **08**, 127 (2022)

55. R. Leonardi, O. Panella, F. Romeo, A. Gurrola, H. Sun, S. Xue. Phenomenology at the LHC of composite particles from strongly interacting Standard Model fermions via four-fermion operators of NJL type. EPJC **80**(309) (2020)

56. C.M.S. Collaboration, Search for heavy neutrinos or third-generation leptoquarks in final states with two hadronically decaying τ leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. JHEP **03**, 077 (2017)

57. C.M.S. Collaboration, Search for third-generation scalar leptoquarks and heavy right-handed neutrinos in final states with two tau leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. JHEP **07**, 121 (2017)

58. C.M.S. Collaboration, Search for third-generation scalar leptoquarks decaying to a top quark and a τ lepton at $\sqrt{s} = 13$ TeV. Eur. Phys. J. C **78**, 707 (2018)

59. C.M.S. Collaboration, Constraints on models of scalar and vector leptoquarks decaying to a quark and a neutrino at $\sqrt{s} = 13$ TeV. Phys. Rev. D **98**(3), 032005 (2018)

60. C.M.S. Collaboration, Search for a singly produced third-generation scalar leptoquark decaying to a τ lepton and a bottom

quark in proton-proton collisions at $\sqrt{s} = 13$ TeV. JHEP **07**, 115 (2018)

61. C.M.S. Collaboration, Search for heavy neutrinos and third-generation leptoquarks in hadronic states of two τ leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. JHEP **03**, 170 (2019)

62. C.M.S. 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)

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)

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 (2022)

65. ATLAS Collaboration, Searches for third-generation scalar leptoquarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. JHEP **06**, 144 (2019)

66. ATLAS Collaboration, Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector. Eur. Phys. J. C **80**(8), 737 (2020)

67. ATLAS Collaboration, Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a τ-lepton in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. JHEP **06**, 179 (2021)

68. ATLAS Collaboration, Search for new phenomena in final states with b-jets and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. JHEP **05**, 093 (2021)

69. 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)

70. ATLAS Collaboration. Search for leptoquarks decaying into the $b\tau$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector (2023)

71. ATLAS Collaboration. Search for pair production of third-generation leptoquarks decaying into a bottom quark and a τ-lepton with the atlas detector (2023)

72. D.A. Faroughy, A. Greljo, J.F. Kamenik, Confronting lepton flavor universality violation in B decays with high-p_T tau lepton searches at LHC. Phys. Lett. B **764**, 126–134 (2017)

73. A. Angelescu, D. Bečirević, D.A. Faroughy, O. Sumensari, Closing the window on single leptoquark solutions to the B-physics anomalies. JHEP **10**, 183 (2018)

74. A. Bhaskar, D. Das, T. Mandal, S. Mitra, C. Neeraj, Precise limits on the charge-2/3 U1 vector leptoquark. Phys. Rev. D **104**(3), 035016 (2021)

75. C. Cornella, D.A. Faroughy, J. Fuentes-Martín, G. Isidori, M. Neubert, Reading the footprints of the B-meson flavor anomalies. JHEP **08**, 050 (2021)

76. L. Allwicher, D.A. Faroughy, F. Jaffredo, O. Sumensari, F. Wilsch, Drell-Yan tails beyond the standard model. JHEP **3**, 2023 (2023)

77. U. Haisch, L. Schnell, S. Schulte, Drell-Yan production in third-generation gauge vector leptoquark models at NLO+PS in QCD. JHEP **2**, 2023 (2023)

78. J. H. Friedman. Greedy function approximation: A gradient boosting machine. Ann. Stat. **29**(5), 1189–1232 (2001). Publisher: Institute of Mathematical Statistics

79. X. Ai, S.C. Hsu, K. Li, C.T. Lu, Probing highly collimated photon-jets with deep learning. JPCS **2438**(1), 012114 (2023)

80. ATLAS Collaboration, 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)

81. S. Chigusa, S. Li, Y. Nakai, W. Zhang, Y. Zhang, J. Zheng, Deeply learned preselection of Higgs dijet decays at future lepton colliders. Phys. Lett. B **833**, 137301 (2022)

82. Y.L. Chung, S.C. Hsu, B. Nachman, Disentangling boosted Higgs boson production modes with machine learning. JINST **16**, P07002 (2021)

83. J. Feng, M. Li, Q.S. Yan, Y.P. Zeng, H.H. Zhang, Y. Zhang, Z. Zhao, Improving heavy Dirac neutrino prospects at future hadron colliders using machine learning. JHEP **9**, 2022 (2022)

84. D. Barbosa, F. Díaz, L. Quintero, A. Flórez, A. Sanchez, A. Gurrola, E. Sheridan, and F. Romeo. Probing a Z′ with non-universal fermion couplings through top quark fusion, decays to bottom quarks, and machine learning techniques. EPJC **83**(413) (2023)

85. N. Assad, B. Fornal, B. Grinstein, Baryon Number and Lepton Universality Violation in Leptoquark and Diquark Models. Phys. Lett. B **777**, 324–331 (2018)

86. L. Calibbi, A. Crivellin, T. Li, Model of vector leptoquarks in view of the B-physics anomalies. Phys. Rev. D **98**(11), 115002 (2018)

87. M. Blanke, A. Crivellin, B Meson anomalies in a Pati-Salam model within the Randall-Sundrum background. Phys. Rev. Lett. **121**(1), 011801 (2018)

88. S. Iguro, J. Kawamura, S. Okawa, Y. Omura, TeV-scale vector leptoquark from Pati-Salam unification with vectorlike families. Phys. Rev. D **104**(7), 075008 (2021)

89. L. Di Luzio, A. Greljo, M. Nardecchia, Gauge leptoquark as the origin of B-physics anomalies. Phys. Rev. D **96**(11), 115011 (2017)

90. A. Greljo, B.A. Stefanek, Third family quark–lepton unification at the TeV scale. Phys. Lett. B **782**, 131–138 (2018)

91. L. Di Luzio, J. Fuentes-Martín, A. Greljo, M. Nardecchia, S. Renner, Maximal flavour violation: a Cabibbo mechanism for leptoquarks. JHEP **11**, 081 (2018)

92. S.F. King, Twin Pati-Salam theory of flavour with a TeV scale vector leptoquark. JHEP **11**, 161 (2021)

93. M. Fernández-Navarro, S.F. King, B-anomalies in a twin Pati-Salam theory of flavour including the 2022 LHCb $R_{K^{(*)}}$ analysis. JHEP **2**, 2023 (2023)

94. M. Bordone, C. Cornella, J. Fuentes-Martín, G. Isidori, A three-site gauge model for flavor hierarchies and flavor anomalies. Phys. Lett. B **779**, 317–323 (2018)

95. M. Bordone, C. Cornella, J. Fuentes-Martín, G. Isidori, Low-energy signatures of the PS^3 model: from B-physics anomalies to LFV. JHEP **10**, 148 (2018)

96. J. Fuentes-Martín, G. Isidori, J.M. Lizana, N. Selimovic, B.A. Stefanek, Flavor hierarchies, flavor anomalies, and Higgs mass from a warped extra dimension. Phys. Lett. B **834**, 137382 (2022)

97. B. Gripaios, Composite leptoquarks at the LHC. JHEP **02**, 045 (2010)

98. R. Barbieri, C.W. Murphy, F. Senia, B-decay anomalies in a composite leptoquark model. Eur. Phys. J. C **77**(1), 8 (2017)

99. R. Barbieri, A. Tesi, B-decay anomalies in Pati-Salam SU(4). Eur. Phys. J. C **78**(3), 193 (2018)

100. J. Fuentes-Martín, G. Isidori, M. König, N. Selimović, Vector leptoquarks beyond tree level. III. Vectorlike fermions and flavor-changing transitions. Phys. Rev. D **102**, 115015 (2020)

101. A. Djouadi, J. Ellis, A. Popov, J. Quevillon, Interference effects in $t\bar{t}$ production at the LHC as a window on new physics. JHEP **03**, 119 (2019)

102. N.D. Christensen, C. Duhr, FeynRules - Feynman rules made easy. Comput. Phys. Commun. **180**, 1614–1641 (2009)

103. A. Alloul, N.D. Christensen, C. Degrande, C. Duhr, B. Fuks, FeynRules 2.0 - A complete toolbox for tree-level phenomenology. Comput. Phys. Commun. **185**, 2250–2300 (2014)

104. J. Alwall, C. Duhr, B. Fuks, O. Mattelaer, D.G. Öztürk, C.H. Shen, Computing decay rates for new physics theories with FeynRules

and MadGraph 5_aMC@NLO. Comput. Phys. Commun. **197**, 312–323 (2015)

105. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.S. Shao, T. Stelzer, P. Torrielli, M. Zaro, 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)
106. NNPDF Collaboration. Parton distributions for the LHC Run II. JHEP **04**, 040 (2015)
107. T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C.O. Rasmussen, P.Z. Skands, An introduction to PYTHIA 8.2. Comput. Phys. Commun. **191**, 159–177 (2015)
108. DELPHES 3 Collaboration. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. JHEP **02**, 057 (2014)
109. C.M.S. Collaboration, Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. JINST **13**(04), P05011 (2018)
110. C.M.S. Collaboration, Identification of hadronic tau lepton decays using a deep neural network. JINST **17**(07), P07023 (2022)
111. CMS Collaboration. Study of the Discovery Reach in Searches for Supersymmetry at CMS with 3000/fb. CMS Physics Analysis Summary CMS-PAS-FTR-13-014, CERN, 2013. Accessed on January 07, (2023)
112. A.M. 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)
113. A.M. Sirunyan et al., Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. JINST **16**(05), P05014 (2021)
114. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
115. T. Chen and C. Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. Association for Computing Machinery
116. The ROOT team. Root - an object oriented data analysis framework. *Nucl. Inst. & Meth. in Phys. Res. A*, 389:81–86, 1997. Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996. See also "ROOT"[software], Release v6.18/02, and https://root.cern/manual/roofit/
117. D. Barbosa, F. Díaz, L. Quintero, A. Flórez, M. Sanchez, A. Gurrola, E. Sheridan, F. Romeo, Probing a Z′ with non-universal fermion couplings through top quark fusion, decays to bottom quarks, and machine learning techniques. Eur. Phys. J. C **83**(5), 413 (2023)
118. A. Flórez, A. Gurrola, W. Johns, P. Sheldon, E. Sheridan, K. Sinha, B. Soubasis, Probing axionlike particles with $\gamma\gamma$ final states from vector boson fusion processes at the LHC. Phys. Rev. D **103**(9), 095001 (2021)
119. A. Flórez, A. Gurrola, W. Johns, J. Maruri, P. Sheldon, K. Sinha, S.R. Starko, Anapole dark matter via vector boson fusion processes at the LHC. Phys. Rev. D **100**(1), 016017 (2019)
120. A. Flórez, Y. Guo, A. Gurrola, W. Johns, O. Ray, P. Sheldon, S. Starko, 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)

121. A. Flórez, K. Gui, A. Gurrola, C. Patiño, D. Restrepo, Expanding the reach of heavy neutrino searches at the LHC. Phys. Lett. B **778**, 94–100 (2018)
122. A. Flórez, A. Gurrola, W. Johns, Y. Do Oh, P. Shendon, D. Teague, T. Weiler, Searching for new heavy neutral gauge bosons using vector boson fusion processes at the LHC. Phys. Lett. B **767**, 126–132 (2017)
123. A. Flórez, L. Bravo, A. Gurrola, C. Ávila, M. Segura, P. Sheldon, W. Johns, 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)
124. B. Dutta, S. Ghosh, A. Gurrola, D. Julson, T. Kamon, J. Kumar. Probing an MeV-Scale Scalar Boson in Association with a TeV-Scale Top-Quark Partner at the LHC. JHEP **03**(164) (2023)
125. R. Arnowitt, B. Dutta, A. Gurrola, T. Kamon, A. Krislock, D. Toback, Determining the Dark Matter Relic Density in the Minimal Supergravity Stau-Neutralino Coannihilation Region at the Large Hadron Collider. Phys. Rev. Lett. **100**, 231802 (2008)
126. B. Dutta, A. Gurrola, T. Kamon, A. Krislock, A.B. Lahanas, N.E. Mavromatos, D.V. Nanopoulos, Supersymmetry Signals of Supercritical String Cosmology at the Large Hadron Collider. Phys. Rev. D **79**, 055002 (2009)
127. C. Avila, A. Flórez, A. Gurrola, D. Julson, S. Starko, Connecting Particle Physics and Cosmology: Measuring the Dark Matter Relic Density in Compressed Supersymmetry at the LHC. Phys. Dark Univ. **27**, 100430 (2020)
128. B. Dutta, A. Gurrola, W. Johns, T. Kamon, P. Sheldon, K. Sinha, Vector Boson Fusion Processes as a Probe of Supersymmetric Electroweak Sectors at the LHC. Phys. Rev. D **87**, 035029 (2013)
129. A. Delannoy, B. Dutta, A. Gurrola, W. Johns, T. Kamon, E. Luiggi, A. Melo, P. Sheldon, K. Sinha, K. Wang, S. Wu, Probing Dark Matter at the LHC using Vector Boson Fusion Processes. Phys. Rev. Lett. **111**, 061801 (2013)
130. B. Dutta, T. Ghosh, A. Gurrola, W. Johns, T. Kamon, P. Sheldon, K. Sinha, S. Wu, Probing Compressed Sleptons at the LHC using Vector Boson Fusion Processes. Phys. Rev. D **91**, 055025 (2015)
131. B. Dutta, W. Flanagan, A. Gurrola, W. Johns, T. Kamon, P. Sheldon, K. Sinha, K. Wang, S. Wu, Probing Compressed Top Squarks at the LHC at 14 TeV. Phys. Rev. D **90**, 095022 (2014)
132. B. Dutta, W. Flanagan, A. Gurrola, W. Johns, T. Kamon, P. Sheldon, K. Sinha, K. Wang, S. Wu, Probing Compressed Bottom Squarks with Boosted Jets and Shape Analysis. Phys. Rev. D **92**, 095009 (2015)
133. Paul Lujan The Pixel. Luminosity, Telescope: A detector for luminosity measurement at CMS using silicon pixel sensors. Technical Report CMS-DN-21-008, et al., *Accepted by Eur* (Phys, J. C., 2022)
134. J. Butterworth, S. Carrazza, A. Cooper-Sarkar, A. De Roeck, J. Feltesse, S. Forte, J. Gao, S. Glazov, J. Huston, Z. Kassabov, R. McNulty, A. Morsch, P. Nadolsky, V. Radescu, J. Rojo, R. Thorne, PDF4LHC recommendations for LHC Run II. J. Phys. G **43**, 023001 (2016)
135. A.M. Sirunyan et al., Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. JINST **13**, P05011 (2018)