



Phenomenology at the large hadron collider with deep learning: the case of vector-like quarks decaying to light jets

Felipe F. Freitas^{1,a}, João Gonçalves^{1,b} , António P. Morais^{1,c}, Roman Pasechnik^{2,d}

¹ Departamento de Física da Universidade de Aveiro and Centre for Research and Development in Mathematics and Applications (CIDMA), Campus de Santiago, 3810-183 Aveiro, Portugal

² Department of Astronomy and Theoretical Physics, Lund University, Sölvegatan 14A, 223-62 Lund, Sweden

Received: 5 May 2022 / Accepted: 9 September 2022 / Published online: 18 September 2022
© The Author(s) 2022

Abstract In this work, we continue our exploration of TeV-scale vector-like fermion signatures inspired by a Grand Unification scenario based on the trinification gauge group. A particular focus is given to pair-production topologies of vector-like quarks (VLQs) at the LHC, in a multi-jet plus a charged lepton and a missing energy signature. We employ Deep Learning methods and techniques based in evolutive algorithms that optimize hyper-parameters in the neural network construction, whose objective is to maximise the Asimov estimate for distinct VLQ masses. In this article, we consider the implications of an innovative approach by simultaneously combining detector images (also known as jet images) and tabular data containing kinematic information from the final states. With this technique we are able to exclude VLQs, that are specific for the considered model, up to a mass of 800 GeV in both the high-luminosity the Run-III phases of the LHC programme.

Contents

1 Introduction	1
2 VLQ sector of the SHUT model	2
3 VLQ search at the LHC	4
3.1 LHC analysis: hybrid net approach	5
3.2 Numerical results	6
4 Conclusions	10
A Low-scale LS-SHUT model	11
B VLQ Feynman rules	12

^a e-mail: felipefreitas@ua.pt

^b e-mail: jpedropino@ua.pt (corresponding author)

^c e-mail: aapmorais@ua.pt

^d e-mail: Roman.Pasechnik@thep.lu.se

C Angular and kinematic distributions at the ATLAS detector	12
References	15

1 Introduction

The Standard Model (SM) of particle physics has, so far, been recognized as the most successful model in theoretical physics, proving its validity over a wide range of subatomic phenomena and energy scales [1–7]. However, despite its immense success, there is an experimental evidence that indicates certain missing features for which the SM can not provide a satisfactory explanation. Among these are the lack of a dynamical mechanism behind neutrino mass generation and the absence of a plausible dark matter (DM) candidate, as well as the emergence of slowly but steadily growing anomalies such as those of the magnetic moment of the muon [8, 9], the R_{D,D^*} [10, 11] and the R_{K,K^*} flavour observables [12, 13]. Regarding the muon $g - 2$ and B-physics anomalies, one should stress that the corresponding deviations for the SM predictions do not pass the 5σ threshold for discovery, and thus are only considered as an indication for possible New Physics (NP) effects.

To address these shortcomings one typically extends the SM framework by introducing new beyond-the-SM (BSM) fields and new interactions. There are three potential avenues for one to follow in this case. One possibility relies in a phenomenology driven approach, and either simply add new BSM fields that address a given problem [14, 15], or work under the effective field theory (EFT) formalism attempting to establish a bound on the NP energy scale by confronting the model predictions with the data [16, 17]. One can also adopt a more fundamental avenue and build a high-energy scale framework from which the SM gauge group,

its Yukawa sector and particle content are emergent at low energy-scale. In the latter approach, one typically resorts to the so-called Grand Unified Theories (GUTs), potentially providing a first-principles understanding of the origins of the specific SM structure. For such a class of models, one attempts to encapsulate the electroweak (EW) symmetry group, $SU(3)_C \times SU(2)_L \times U(1)_Y$ into a larger gauge group, with the SM and NP fields being embedded into its representations. This strategy may result in various types of unification in the sectors of matter fields and forces such as gauge and/or Yukawa couplings unification, Higgs-lepton unification, quark-lepton unification etc, in some cases even yielding relatively light NP fields in the low-energy SM-like EFT spectrum. The latter may be observable at current or near-future collider experiments, offering a potential of probing the GUT-scale physics through either direct observation of those new states or studies of their impact on precision SM observables, such as the muon $g - 2$ and flavour changing neutral current (FCNC) processes. In this context, some of the authors have previously constructed a novel class of GUT models based on the trinification gauge group supplemented with a horizontal family symmetry in both supersymmetric (SUSY) and non-SUSY versions [18–23]. In particular, it has been demonstrated that successful unification of the gauge couplings can be achieved in a SUSY formulation in consistency with proton decay constraints [21,22]. For this work, we shall focus on the latter scenario, which shares many common features with other GUT implementations and thus represents a good playground for collider searches.

The SHUT (Supersymmetric Higgs-Unified Trinification) model [21,22] predicts various types of NP at low energies. The extended fermion sectors include heavy left-handed and sterile right-handed neutrinos as well as vector-like fermions (VLFs), such as EW-singlet vector-like quarks (VLQs) and EW-doublet vector-like leptons (VLLs). The latter can potentially be at the reach of Run-III and the high-luminosity (HL) phases of the Large Hadron Collider (LHC) [24]. An extended scalar sector features a minimum of three Higgs doublets at low scales, required for consistency with the measured Cabibbo–Kobayashi–Maskawa (CKM) quark mixing matrix.

In this article we continue the phenomenological exploration of SHUT inspired models, focusing on the collider signatures of a down-type VLQs, expected to emerge at a TeV energy scale. Three generations of VLQs are predicted within such a GUT framework with the first two generations being light enough to be potentially accessible at the LHC, with masses of $\mathcal{O}(1)$ TeV, whereas the 3rd generation VLQ is expected to hold a mass of $\mathcal{O}(100)$ TeV i.e. it is too heavy to be probed in collider measurements in the near future. As a particular promising discovery channel, we consider multi-jet signatures with one charged lepton and missing energy, probing the couplings of the VLQ with light quarks. For this,

we apply Deep Learning (DL) techniques to help separating signal from background classes. To this end, we rely on two types of data in the training phase, namely, we use the abstract jet images [25] and the numerical data containing kinematic information from the final states. To train the Neural Network (NN) on the data, we construct a specific Hybrid Net, which is designed to make use of both data formats. Optimization of the NN hyper-parameters is performed by using an evolution algorithm, which optimises the Asimov metric. For a thorough description of our methods and techniques we refer the reader to our previous work [24].

The paper is organized as follows. In Sect. 2 we provide a brief introduction into the SHUT model, providing particular details relevant for our phenomenological analysis. In Sect. 3 we present the methodology implemented in the search for VLQs at the LHC, including a full description of the chosen topology as well as a detailed description of the DL algorithms and datasets used for solving the considered classification problem. This section is finalised with the presentation and subsequent discussion of the numerical results. Finally, concluding remarks are given in Sect. 4. Additional information is presented in appendix, including the VLQ Feynman rules B as well as kinematic and angular distributions C.

2 VLQ sector of the SHUT model

In this section, we begin by briefly introducing the SHUT model, with the focus on the main properties that arise from the GUT description, without delving deep into the mathematical inner workings of this framework. A more detailed description of the considered GUT framework can be found in earlier works e.g. in Refs. [18–22], while a comprehensive discussion of the low-energy limit, including the physical Lagrangian that is used in this work, can be found in Ref. [24]. In what follows we provide the theoretical background that was taken as a physics case to our study. For completeness of information a more extensive discussion of the model can be found in Appendix A.

The SHUT model is a particularly promising flavoured GUT (F-GUT) scenario whose group-theoretical structure is inspired by the E_6 exceptional gauge group extended with a local $SU(2)_F \times U(1)_F$ family symmetry, where all matter content, including the Higgs sector, is unified in fundamental $(\mathbf{27}, \mathbf{2})$ and $(\mathbf{27}, \mathbf{1})$ representations. It then proposes a first-principles' explanation for some of the well-known problems within the SM, such as a common origin of all fundamental gauge forces (with exception of gravity) as well as the observed flavour hierarchies in the SM fermionic sectors. In practical terms, the SHUT model starts off with the trinification gauge symmetry, supplemented with the $SU(2)_F \times U(1)_F$ family symmetry, whose representations

Table 1 Trinification fields representations under $[SU(3)]^3 \times SU(2)_F \times U(1)_F$ for each of the fundamental chiral superfields shown in Eq. (2.1)

	$SU(3)_L$	$SU(3)_R$	$SU(3)_C$	$SU(2)_F$	$U(1)_F$
L^i	3	$\bar{3}$	1	2	1
L^3	3	$\bar{3}$	1	1	-2
Q_L^i	$\bar{3}$	1	3	2	1
Q_L^3	$\bar{3}$	1	3	1	-2
Q_R^i	1	3	$\bar{3}$	2	1
Q_R^3	1	3	$\bar{3}$	1	-2

Bold values indicate charges of non-abelian $SU(3)/SU(2)$

have a common origin from the referred (27,2) and (27,1) multiplets of the $E_6 \times SU(2)_F \times U(1)_F$ gauge group.

Due to SUSY, the Higgs sector is unified with the charged lepton and neutrinos sectors, thus, sharing common flavor properties with fermions at high energies. The superpotential of the SHUT model reads [22]

$$W_3 = \varepsilon_{IJ} (\mathcal{Y}_1 L^I Q_L^3 Q_R^J - \mathcal{Y}_2 L^I Q_L^J Q_R^3 + \mathcal{Y}_2 L^3 Q_L^I Q_R^J), \tag{2.1}$$

where ε_{IJ} is the two-dimensional Levi-Civita symbol, \mathcal{Y}_1 and \mathcal{Y}_2 are the two unified Yukawa couplings, while subscripts L/R denote $SU(3)_{L/R}$ superfields. The indices $I, J = 1, 2$ represent $SU(2)_F$ doublets, while flavour singlets are denoted with the “3” in superscript. These massless superfields form bi-triplets of the trinification symmetry ($SU(3)_R \times SU(3)_L \times SU(3)_C$) and transform in accordance with the representations shown in Table 1. Expanding these fields in terms of $SU(3)$ indices, we write the superfields in component form as [22]

$$\begin{aligned} (L^{i,3})_r^l &= \begin{pmatrix} \mathcal{N}_R & \mathcal{E}_L & e_L \\ \mathcal{E}_R & \mathcal{N}_L & \nu_L \\ e_R & \nu_R & \phi \end{pmatrix}^{i,3}, \\ (Q_L^{i,3})_l^x &= (u_L^x \quad d_L^x \quad D_L^x), \\ (Q_R^{i,3})_x^r &= (u_{R_x} \quad d_{R_x} \quad D_{R_x})^T \quad i,3, \end{aligned} \tag{2.2}$$

where l, r and x are $SU(3)_L, SU(3)_R$ and $SU(3)_C$ indices, respectively. In here, we note that the left-handed and right-handed components of the SM-like leptons are embedded in the L superfield (e_R and e_L). Right-handed neutrinos also reside in this representation (which are embedded inside \mathcal{N}_R, ν_R and ϕ), which helps address the neutrino mass problem via a radiative seesaw mechanism [20,22]. The SM quarks are then found in the $d_{L,R}/u_{L,R}$ components of $Q_{L,R}$, whereas the new exotic down-type VLQs reside in the $D_{L,R}$ components.

At low energies, the SM gauge group is fully realized, with additional exotic physics, including an extended scalar

sector and vector-like fermions. Full unification of the field content results in a reduced freedom in the Yukawa sector, with the presence of only two \mathcal{Y}_1 and \mathcal{Y}_2 couplings at high scales, which provide the leading contributions to the generation of quark masses at tree level and the necessary means to radiatively induce strong hierarchies as observed in Nature. Indeed, as soon as the scalar components of L develop vacuum expectation values (VEVs) that break the SHUT symmetry down to that of the SM, the down-type VLQs, second- and third-generation SM-like quarks gain tree-level masses, while the charged leptons and first-generation quarks develop their masses at quantum level via loop corrections. This fact offers the necessary ingredients to potentially explain the observed splitting in the fermion spectra implying the relative lightness of the charged leptons, neutrinos and the quarks of first generation.

In Ref. [22], various benchmark scenarios assuming three light Higgs doublets yielding a consistent CKM mixing matrix, were explored in detail. The further presence of new VLQs results in an extended 3×6 quark mixing matrix, defined as

$$V_{CKM} = U_L^u \cdot P \cdot (U_L^d)^\dagger = \begin{pmatrix} V_{CKM}^{SM} & V_{CKM}^{VLQs} \end{pmatrix}, \tag{2.3}$$

where P is a projection operator, defined as $P = (\mathbb{1}_{3 \times 3} | 0_{3 \times 3})$. Here, a 3×3 matrix V_{CKM}^{SM} represents the CKM-like quark mixing of the SM while a 3×3 matrix V_{CKM}^{VLQs} contains a mixing between the down-type VLQs with up-type chiral quarks of the SM. Two of them are relatively light, with a certain hierarchy between them, and are expected to emerge at a TeV energy scale. A third heavy VLQ appears beyond the reach of current or near-future colliders.

In the present work we explore the LHC discovery potential of the lightest VLQ denoted as D , ignoring its heavier counterparts for simplicity. Their mass reads as

$$m_D \sim \omega \mathcal{Y}_2 \quad m_S \sim p \mathcal{Y}_2 \quad m_B \sim p \mathcal{Y}_1, \tag{2.4}$$

with $0.01 \sim \mathcal{Y}_2 \ll \mathcal{Y}_1 \sim 1$ fixed by the charm and top quark masses respectively, and with $\omega \sim 100$ TeV being an intermediate Left-Right symmetry breaking scale and $p \gg \omega$, associated to the breaking of the family symmetry. This results in the hierarchy $m_B \gg m_S \gg m_D$ with the latter expected to emerge not far above the TeV scale. As a particular benchmark scenario inspired by the phenomenologically relevant ranges of the SHUT model parameters found in Ref. [22], we adopt the following values of the vector-like D -quark mixing elements in the extended CKM matrix:

$$V_{uD} \simeq 5.1 \times 10^{-6}, \quad V_{cD} \simeq 2.6 \times 10^{-5}, \quad V_{tD} \simeq 0.016. \tag{2.5}$$

In our numerical analysis below, we consider only light quark jets in the final states of VLQ decays $D \rightarrow W^+ \text{jets}$, so only V_{uD} and V_{cD} will be relevant in what follows. These

will be fixed as above, whereas a numerical scan will be performed over the VLQ mass m_D within a phenomenologically acceptable range (see below).

3 VLQ search at the LHC

The generation of events in proton-proton collisions at the LHC follows the standard methods employed by the particle physics community. In this work, we use a specially designed chain of computational tools developed in the previous work of Ref. [24], and here we only briefly elaborate on the corresponding machinery. First, the Lagrangian density is constructed utilizing the SARAH package [26], which can be used to derive interaction vertices between the fields, their spectra and mixing. The latter are then outputted into UFO [27] files that can be used in Monte-Carlo event generators. In particular, here we employ MadGraph [28] for hard-scattering events and Pythia8 [29] for hadronization/showering processes. Fast-simulation of the detector is implemented using Delphes [30] (with the ATLAS default card), while the kinematic data extraction is done in ROOT [31]. We generate 250k events, for both VLQ production signal and backgrounds, at a centre-of-mass energy of $\sqrt{s} = 14$ TeV and all at leading-order (LO) precision. We also employ the parton-distribution function NNPDF2.3 [32], which fixes the strong-coupling constant, α_s , and its evolution.

First, it is instructive to analyse the current constraints from direct searches of down-type VLQs. From the experimental side, among the well-known NP states a VLQ is one of the most searched for at the LHC. Since it is a coloured particle, its production rate is expected to be relatively high. In this regard, typical lower bounds on the VLQ mass range from 1.4 TeV up to 2.0 TeV (see the latest exclusion bounds, as of July 2021, in the summary plots of Ref. [33]). However, one must carefully analyse the assumptions made in these analyses. First, the vast majority of current searches focus on dominant couplings to the third generation of chiral quarks (see e.g. Refs. [34–39]). Such an assumption is not the most general, as there is no reason for the VLQ to not couple to lighter quarks. Indeed, there have also been the searches assuming dominant light-quark couplings – see, for instance, Refs. [40,41]. In Ref. [40], in particular, a search for pair production of heavy VLQs has been performed, with their subsequent decays into light jets, that can go via either neutral ($D \rightarrow Zq$) or charged ($D \rightarrow Wq$) currents. This work excludes the VLQs with masses below 690 GeV at 95% confidence level. Such a constraint is quite relevant since the same topology will be used in our analysis below. In the second work [41], the same VLQ decay modes into light jets have been considered, but in single VLQ production channels. For the neutral $D \rightarrow Zq$ decay channel, the VLQ mass has been constrained to be below 760 GeV, whereas a tighter

lower bound of 900 GeV is obtained in the charged $D \rightarrow Wq$ decay channel. These constraints, however, are not relevant for the particular benchmark point that we have shown in the previous section. For example, we have calculated the production cross-section in MadGraph for single production of the VLQ (using the benchmark point shown in Eq. (2.5)), for which we obtained $\sigma = 1.678 \times 10^{-4}$ pb for a mass of 800 GeV, which is consistent with the cross-section limits in Ref. [41]. For the sake of consistency, in this work we consider the VLQ mass to be greater than 800 GeV. In fact, it was verified that for the $D \rightarrow Wt$ channel this particular benchmark point yields a cross-section of 0.067 pb compatible with the upper 95% C.L. limit in [42], where $\sigma \approx 0.07$ pb and where it is assumed that the sum of the decay branching fractions to Zb , Hb and Wt is one, as opposed to our scenario where $\text{Br}(D \rightarrow Wt) + \text{Br}(D \rightarrow Wq) \approx 1$.

As mentioned before, the chosen topology of the signal is equivalent to that of Ref. [40], with VLQ pair-production undergoing via gluon fusion. An example for the LO Feynman diagram, with the contribution from the gluon triple vertex, is shown in Fig. 1. All tree-level graphs that contribute to the matrix-element are automatically included at the level of event generation in MadGraph. The signal is characterized by one charged lepton (either an electron or a muon) and missing energy, both arising due to W^- boson decay. Additionally, four jets from first and second generation quarks are present, with two of these arising from the decay of one W boson and the other two originating from the decays of the heavy VLQ states. Due to the presence of multiple jets, the matching and merging to the original partons is important. For this, we rely on the automatic matching procedure of MadGraph, which employs the kT-MLM scheme [43]. For tagging the jets, we consider the following procedure. First, for each event, we begin by picking all jets and order them by its p_T value. From this list, we subsequently impose the kinematic constraints, namely $|\eta(j)| < 2.5$ and $p_T(j) > 25$ GeV. From the surviving list we check whether the jets are tagged as coming from a b quark. If the jet is not tagged as b , then it is accepted, else, it is discarded. If there are less than 4 jets that obey these constraints then the event is discarded. We do not consider any tau tagging information when selecting valid jet candidates.

The main irreducible backgrounds for this topology include:

- Pair-production of top quarks, $t\bar{t}$, with one of the W bosons decaying fully leptonically, whereas the other one decays into two light jets;
- Diboson production plus two light jets. In this case, one of the vector bosons decays into two light jets whereas the other one (W boson) decays into a charged lepton and a neutrino;

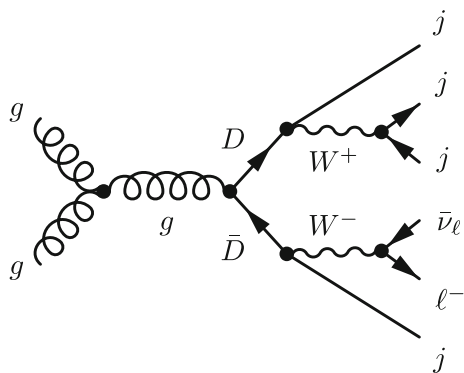


Fig. 1 One of the LO Feynman diagrams for the pair-production of a D -type VLQ. Here, g corresponds to gluons from initial colliding protons. The D -quark decays into a light jet $j = u, c, \bar{u}, \bar{c}$ and a W boson, with one of the W 's further decaying via a leptonic channel, with one charged lepton ($\ell^- = e^-, \mu^-$) and the corresponding neutrino ($\bar{\nu}_e, \bar{\nu}_\mu$), and the other decaying via a hadronic channel into light jets. In the hadronic W decay, both the up- and down-type SM quarks are considered

- $V + \text{jets}$, with V either a W or a Z^0 boson. While for the former one considers the decay into a charged lepton and its corresponding neutrino, for the latter case we consider the fully leptonic decay channel (muon or electron modes).

Since the Monte-Carlo simulations are done at LO, we must reweight our events based on higher-order corrections available in literature. In particular, we consider approximate $N^3\text{LO}$ corrections for $t\bar{t}$ [44], NLO corrections for $W + \text{jets}$ [45] and $WW + \text{jets}$ [46], and NNLO corrections to $Z^0 + \text{jets}$ [47].

While the NN models will do most of the heavy lifting, we will also consider simple selection criteria to help discriminate the signal events from the background. In particular, we consider that

$$p_T(\ell^\pm) > 25 \text{ GeV},$$

$$\text{MET} > 20 \text{ GeV},$$

$$|\eta(\ell^\pm)| \leq 2.5,$$

with $\eta = -\ln(\tan(\theta/2))$, θ is the angle with respect to the beam axis, and MET is defined as a missing transverse energy. With this, we extract all relevant kinematic and angular information about the final states' configurations, such as the transverse momentum, pseudo-rapidity and mass distributions of pairs of produced particles. Low-level observables include the transverse momenta, energy, pseudo-rapidity and the azimuthal angle of the four jets (j_1, j_2, j_3 and j_4) with the charged leptons, as well as MET. The jets are ordered with respect to their total transverse momenta, that is, j_1 is the leading jet with the highest p_T and j_4 is that with the lowest p_T . Additionally, observables reconstructed from the combination of various objects

are also considered. Namely, one uses the mass distributions $M(j_1, j_2), M(j_1, j_3), M(j_1, j_4), M(j_2, j_3), M(j_2, j_4), M(j_3, j_4), M(\text{MET}, \ell^-, j_1), M(j_2, j_3, j_4), M(\text{MET}, \ell^-, j_2), M(j_1, j_3, j_4)$, the cosine of the polar angle $\cos \Delta\theta(\ell^-, \text{MET}), \cos \Delta\theta(\ell^-, j_1), \cos \Delta\theta(\ell^-, j_2), \cos \Delta\theta(\ell^-, j_3), \cos \Delta\theta(\ell^-, j_4), \cos \Delta\theta(j_1, j_2), \cos \Delta\theta(j_1, j_3), \cos \Delta\theta(j_1, j_4), \cos \Delta\theta(j_2, j_3), \cos \Delta\theta(j_2, j_4), \cos \Delta\theta(j_3, j_4)$, the azimuthal angle distributions $\Delta\Phi(j_1, \ell^-), \Delta\Phi(j_2, \ell^-), \Delta\Phi(j_3, \ell^-), \Delta\Phi(j_4, \ell^-), \Delta\Phi(j_1, j_2), \Delta\Phi(j_1, j_3), \Delta\Phi(j_1, j_4), \Delta\Phi(j_2, j_3), \Delta\Phi(j_2, j_4), \Delta\Phi(j_3, j_4)$ and the ΔR distributions $\Delta R(j_1, \ell^-), \Delta R(j_2, \ell^-), \Delta R(j_3, \ell^-), \Delta R(j_4, \ell^-), \Delta R(j_1, j_2), \Delta R(j_1, j_3), \Delta R(j_1, j_4), \Delta R(j_2, j_3), \Delta R(j_2, j_4), \Delta R(j_3, j_4)$.¹

3.1 LHC analysis: hybrid net approach

In this section, we present the DL methodology used in our analysis, which involves a combination of both image and tabular datasets. The network, which we name as Hybrid Net in what follows, will take into account both data types in the classification of background and signal classes. The image datasets will be trained in a convolutional model, whereas kinematic data are trained through a linear sequential model, after which we combine the predictions of both.

Jet images, first introduced in Refs. [48,49], are constructed by mapping particle hits in the detector into a grid in the (η, ϕ) plane, where the intensity of the pixel is directly proportional to the energy deposited by the particles. These methods, however, can be quite inefficient since in a typical jet-image there can exist plenty of blank pixels, making it computationally inefficient when passing through convolutional layers in the training phase. While various approaches have been proposed in the literature (such as particle clouds [50] or through the use of graph networks [51,52]), in this work we consider abstract jet-images as first discussed by one of the authors and collaborators in Ref. [25]. In this approach, besides associating each pixel with the energy of a particle, we also map characteristics of the particle into various geometrical shapes. For example, if it is a charged particle in a detector, we map it to a circle. In this scenario, the radius of the polygons is directly proportional to the particle energy. With this, we then provide additional features for the neural model to learn. Examples of these types of images can be seen in Fig. 2.

In order to construct such images, we impose the following conditions:

¹ ΔR is defined as the Euclidean distance in the (η, ϕ) plane, i.e. $\Delta R(i, j) = \sqrt{\Delta\Phi(i, j)^2 + \Delta\eta^2(i, j)}$ with $\Delta\Phi(i, j) = \phi_j - \phi_i$ and $\Delta\eta(i, j) = \eta_j - \eta_i$. $\Delta\theta(i, j)$ is identically defined, with $\Delta\theta(i, j) = \theta_j - \theta_i$. Mass distributions are reconstructed from the observables inside parenthesis, e.g. $M(j_1, j_2) \equiv M(j_1^\mu + j_2^\mu)$, where j^μ are four-momentum vectors.

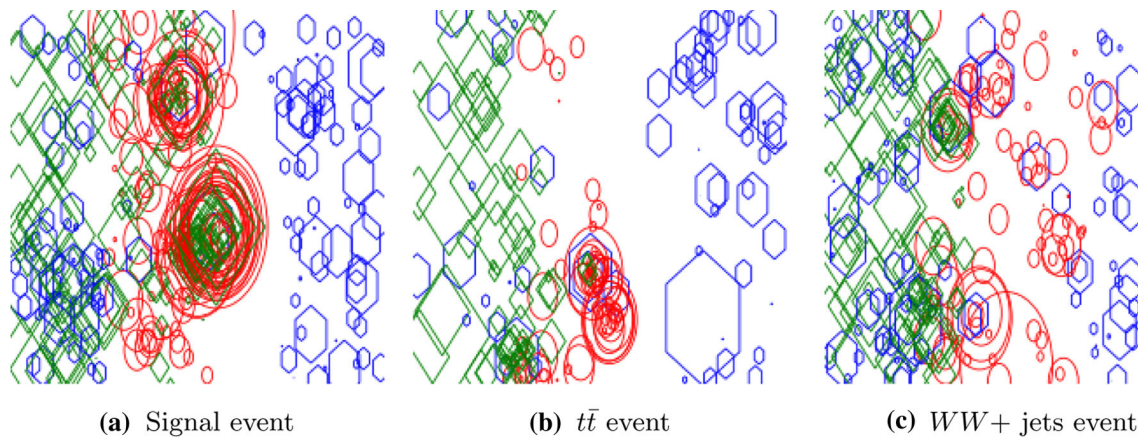


Fig. 2 Example of an abstract jet image in the (η, ϕ) plane. Green polygons are representative of photons, red circles are charged particles while blue hexagons indicate neutral hadrons. The radius of the polygons is proportional to the energy left by the particle in the calorimeter

tower. In this image we have an event for a proton-proton collision at $\sqrt{s} = 14$ TeV, with subsequent VLQ double production (to the right) and the corresponding background events ($t\bar{t}$ in the centre and WW +jets in the right)

- Charged particles hitting the detector are mapped into blue circles, whose radius is given as $r = \log p_T$, where p_T is the transverse momentum of the particle;
- Neutral hadrons are mapped into hexagons, with the radius given as $r = \log(E_T)$, where E_T is the transverse energy;
- Photons are mapped into squares, with the radius given by the same formula as the neutral hadrons.

- Bottleneck architecture²: True or False;
- Initialisers in the sequential model: Normal, He normal and He uniform;
- Optimisers in the sequential model: Adam, SGD, Adamax and nadam;
- Vision models: ResNet-34.

The resulting images are then stored in the .png format, with the pixel size of 224×224 , requiring that both hadrons and photons are restricted to be in the kinematic ranges of $|\eta| \leq 5.0$ and $|\phi| \leq \pi$.

We use both the image and tabular information and feed it into a NN. The images pass through a convolutional network, whereas the corresponding kinematic information is passed through a linear model. We then combine the results by summing together the predictions of both networks leading to the final result. The hyper-parameter optimization designed to maximize the statistical significance is implemented via evolutionary algorithms as described in previous works [24,53,54]. Training of the NNs based on the statistical significance follows the analysis first performed by A. Elwood and D. Krücker [55]. In this work, only linear networks are part of the optimization procedure, whereas the convolutional network for the images is fixed to a ResNet-34 network [56]. Furthermore, we consider the following list of hyper-parameters:

- Number of neurons of the sequential model: 256, 512, 1024 and 2048;
- Number of layers of the sequential model: 1, 2, 3, 4 and 5;

The evolutionary algorithm builds sets of Hybrid Net models. Both the linear and convolution parts of the algorithms are constructed using PyTorch [57]. In the convolution part, we also utilized FastAI [58], which employs PyTorch as backend. Given this methodology, let us now turn to a discussion of numerical results.

3.2 Numerical results

We first begin by presenting the results for a particular (fixed) VLQ mass. Focusing on the lightest D -VLQ with a mass of 800 GeV and applying the selection criteria described above, the obtained LO cross-sections for the signal, and corresponding higher-order rates for the backgrounds, are showcased in Table 2. Additionally, the total expected number of events at both run-III and the HL phase of the LHC are also indicated. The variation of the post-selection cross-section with the VLQ mass is shown in Fig. 3. In this figure, we also plot dashed lines to indicate the allowed sensitivity expected for future runs of the LHC. In particular, we show that for the benchmark values of the VLQ-light quark mixing elements indicated in Eq. (2.5), the run-III phase of the LHC can potentially probe D -type VLQs up to 1.9 TeV in this channel, while the HL phase has a sufficient sensitivity

² We define a bottleneck architecture as a neural model with an ever decreasing number of neurons at each layer.

to exclude VLQs up to 2.37 TeV. Note that this approximate limits are set for the best case scenario, that is, in the limit of experimental errors going to zero and therefore can be seen as guideline to the mass scale where the VLQ can be potentially probed. Indeed, in the scenario where only one event is produced, the systematics of the backgrounds completely overshadow the signal.

In order to evaluate the performance of the network, we show in Fig. 4 the receiver operating characteristic (ROC) plots for three distinct situations: for classification with only kinematic data (top panel), for classification with only jet images (middle panel) and classification where both are taken into account (bottom panel). These plots can serve as a measurement of how well the network performs on the data. In particular, these results are shown for the validation dataset, i.e. for the dataset that the network was not trained on. Within the labeling of the curves, we also indicate the area under the curve (AUC), which is related to the accuracy. As one can readily observe, the combination of both the jet images and the kinematic data offers the greater discriminating power with the highest accuracy, clearly validating the improved performance of the developed hybrid network. Indeed, we have found that the tabular data is one that performs the worst, with the jet image classification giving the best results. Such a good accuracy indicates an efficient separation between the signal and background classes. Further confirmation of this fact can be given by the predicted confidence scores for the samples plotted on the left-hand-side of Fig. 4. Here, we can clearly see that the model successfully separates the signal classes (mainly populating the convolutional NN scores of around one) from the background (which mainly populates a score of around zero). Of course, such separation is not perfect, with some contamination still present, in a sense that some background and signal overlap in some regions of the convolutional NN parameter domain. All subsequent analysis on the data will be done based on the best results, i.e. on the hybrid network scores with both kinematic and image data.

The histograms with the kinematic and angular observables that are utilized by the NN are plotted in appendix, see Figs. 7 and 8. One can observe that the angular distributions for the signal tends to be overshadowed by the backgrounds, with the signal following the same structure as the background, with the only exception being the pseudo-rapidity distributions. Note that the signal has a dominant peak at $\eta \sim 0$, whereas the backgrounds spread over the entirety of the allowed range $|\eta| \leq 2.5$. The kinematics, on the other hand, offers the most significant discriminating power, with distributions being peaked in kinematic regions where the SM backgrounds are less dominant. This is true for all transverse momentum, energy and MET distributions. Similarly, mass distributions also peak in higher invariant mass domains compared to those of the SM backgrounds.

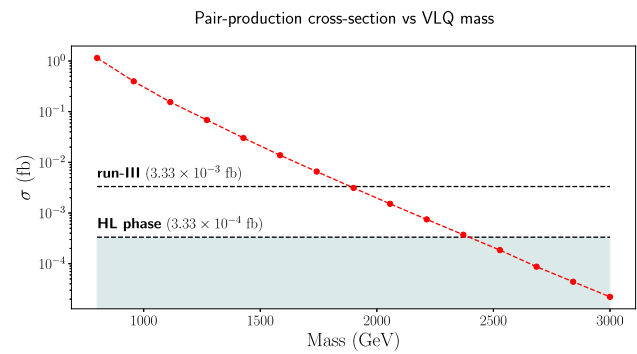


Fig. 3 The production cross-section (in femtobarn) as a function of the VLQ mass (in GeV). Two horizontal dashed lines are shown, one for the run-III target luminosity and the other for the HL phase of the LHC, which represent the minimal cross section for which a single event can be produced, i.e. $N_{\text{exp}} = \sigma \mathcal{L} = 1$ for the dashed lines. The area filled in gray represents the region where neither the HL-LHC or run-III LHC have the required sensitivity to detect the D -VLQ for the chosen benchmark (2.5). The y-axis is shown in logarithmic scale

Each individual event corresponds to a particular value of the physical observables transported over into the tabular data. Each event will have an associated jet image. A combination of both is used in the classification procedure, and then the statistical significance is computed using the validation dataset. We present the results using the Asimov estimate, which is defined as [59]

$$\mathcal{Z}_A = \left[2 \left((s + b) \ln \left(\frac{(s + b)(b + \sigma_b^2)}{b^2 + (s + b)\sigma_b^2} \right) - \frac{b^2}{\sigma_b^2} \ln \left(1 + \frac{\sigma_b^2 s}{b(b + \sigma_b^2)} \right) \right) \right]^{1/2}, \tag{3.1}$$

where s is the number of signal events, b is the number of background events and σ_b is the uncertainty of the background. For this metric, we consider three scenarios with either $\sigma_b = 15\%$, $\sigma_b = 25\%$ or the much more conservative scenario of $\sigma_b = 50\%$ systematic uncertainty. The first two are the most realistic.

Considering the HL phase of the LHC, which is expected to reach an integrated luminosity of $\mathcal{L} = 3000 \text{ fb}^{-1}$, we can determine the median significance for the considered metrics as a function of the NN score, as shown in Fig. 5. The maximal significances for each metric are then determined to be

- $\mathcal{Z}_A(\sigma_b = 15\%) = 16.89\sigma$,
- $\mathcal{Z}_A(\sigma_b = 25\%) = 10.33\sigma$,
- $\mathcal{Z}_A(\sigma_b = 50\%) = 5.21\sigma$,

i.e. we obtain significances greater than 5σ for all systematic uncertainties, even in the most conservative scenario of 50%. Such a strong signature of VLQs at the HL run of the LHC can be expected, since VLQs are coloured particles, and their production is enhanced at hadron colliders.

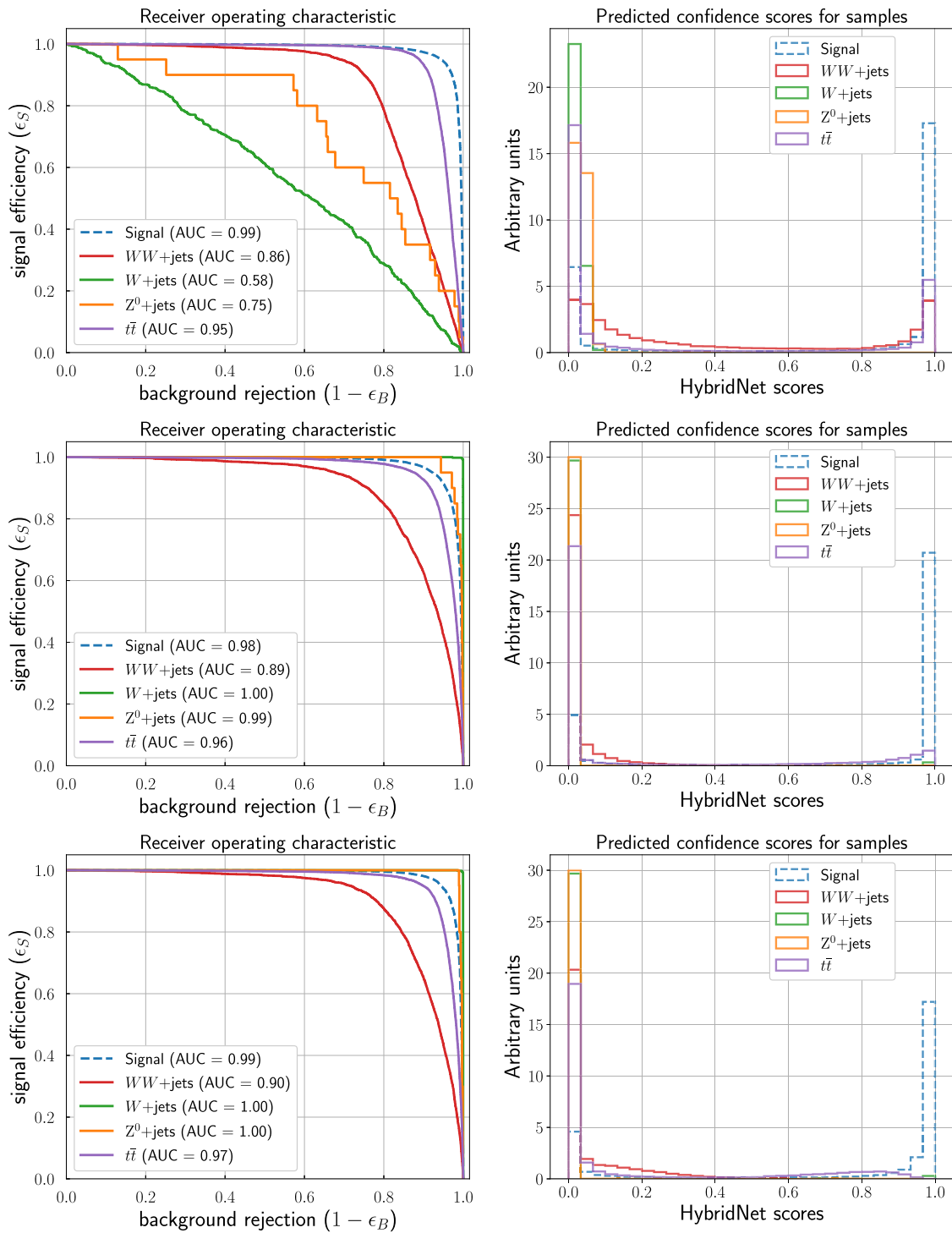


Fig. 4 Performance results for the Hybrid network. On each panel, we plot the ROC on the left-hand side and, on the right, – the predicted confidence score for our samples. In the top panel, the results are shown for classification with only kinematic data, in the middle panel we have

the results only for classification with the jet images, and at the bottom, the results when both data types were part of the training. The signal is indicated with a dashed blue line, while the $WW+jets$ background – in red, the $W+jets$ – in green, Z^0+jets – in orange and $t\bar{t}$ – in purple

Table 2 Predicted total cross section (in fb) for both the signal and each respective background, before and after the cut selection. In the last two columns we indicate the overall number of expected events for the analysis, calculated as $N = \sigma \mathcal{L}$, for run-III and the HL phase of the LHC

	σ (before cuts, in fb)	σ (after cuts, in fb)	Events at run-III	Events at HL-LHC
VLQ 800 GeV	1.402	1.063	318	3189
WW + jets	1.24×10^5	7.72×10^3	2.316×10^6	2.316×10^7
W + jets	7.95×10^8	3.46×10^5	1.038×10^8	1.038×10^9
Z^0 + jets	6.33×10^7	5.70×10^3	1.71×10^6	1.71×10^7
$t\bar{t}$	9.89×10^5	1.83×10^5	5.49×10^7	5.49×10^8

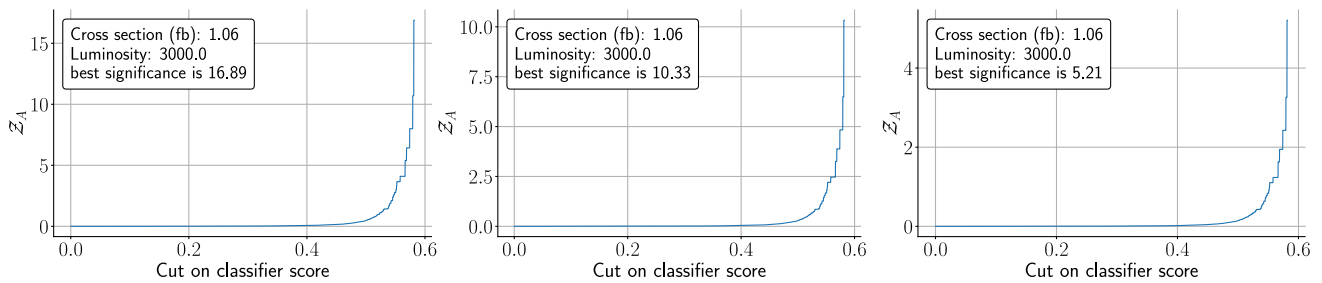
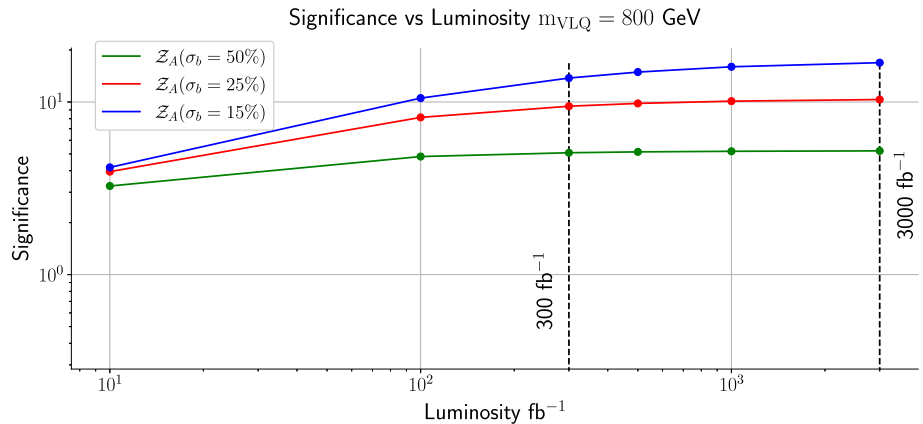


Fig. 5 The statistical significance for a hypothetical discovery of VLQs as a function of the NN score for different values of the systematic uncertainties considered in this work. The VLQ mass is 800 GeV, and the collider luminosity is $\mathcal{L} = 3000$ fb. The significance is computed fol-

lowing the implementation of an evolution algorithm whose method is described in [24]. From left to right, we have $\sigma_b = 15\%$, $\sigma_b = 25\%$ and $\sigma_b = 50\%$

Fig. 6 The statistical significance as a function of the systematic uncertainties for pair-production of 800 GeV D -VLQs. The green curve corresponds to $\sigma_b = 50\%$, the red curve – to $\sigma_b = 25\%$ and the blue line – to $\sigma_b = 15\%$. There are two vertical dashed lines marking the run-III luminosity (300 fb^{-1}) and the HL-LHC luminosity (3000 fb^{-1}). Both the y-axis and x-axis are shown in logarithmic scale



It is also relevant to analyse the impact for other luminosities, in particular, for the run-III of the LHC, which is set to deliver $\mathcal{L} = 300 \text{ fb}^{-1}$ of data. In Fig. 6 we plot the statistical significance as a function of the systematic errors. Specifically, the maximal significances obtained for the run-III luminosity read as follows:

- $\mathcal{Z}_A(\sigma_b = 15\%) = 13.76\sigma$,
- $\mathcal{Z}_A(\sigma_b = 25\%) = 9.45\sigma$,
- $\mathcal{Z}_A(\sigma_b = 50\%) = 5.08\sigma$,

i.e. the significance is above 5σ for all considered systematics, indicating that the VLQ production can also be probed at run-III.

In order to complement these results further, we implement the same machinery in a numerical scan over the VLQ mass, up to the detectability limit of the HL phase, which corresponds up to the VLQ mass of around 2.2 TeV. For these additional mass points, we have used the same network as the one for the 800 GeV benchmark, with the results summarised in Table 3. As the main result, a 5σ significance is obtained for VLQ masses of up to 800 GeV. In particular, for a VLQ of 1.2 TeV, the highest significance was of 1.64σ , at the HL phase of the LHC. Indeed, for points with a larger mass, a significance above the discovery threshold requires a greater control of the background systematics. For example, if we consider a systematic uncertainty at the percent level at the high-luminosity phase of the LHC, for masses of order 1.2 TeV one obtains a remarkable enhance-

Table 3 The statistical significance for different values of the systematics considered in this work, for various masses of the VLQ and different collider luminosities. Scenarios which pass the 5σ requirement for

potential exclusion (or hypothetical discovery) are displayed in bold. Points marked with “–” indicate that the cross-section is not enough to produce one event at the corresponding luminosity

Benlmarks	800 GeV	1200 GeV	1700 GeV	2200 GeV
$\mathcal{Z}_A(\sigma_b = 15\%) 300 \text{ fb}^{-1} \mathcal{Z}_A(\sigma_b = 25\%) \mathcal{Z}_A(\sigma_b = 50\%)$	13.76σ	1.23 σ	0.06 σ	–
	9.45σ	0.89 σ	0.04 σ	–
	5.08σ	0.49 σ	0.02 σ	–
$\mathcal{Z}_A(\sigma_b = 15\%) 1000 \text{ fb}^{-1} \mathcal{Z}_A(\sigma_b = 25\%) \mathcal{Z}_A(\sigma_b = 50\%)$	15.99σ	1.51 σ	0.07 σ	–
	10.11σ	0.98 σ	0.05 σ	–
	5.18σ	0.51 σ	0.02 σ	–
$\mathcal{Z}_A(\sigma_b = 15\%) 3000 \text{ fb}^{-1} \mathcal{Z}_A(\sigma_b = 25\%) \mathcal{Z}_A(\sigma_b = 50\%)$	16.89σ	1.64 σ	0.08 σ	0.014 σ
	10.33σ	1.01 σ	0.05 σ	0.0071 σ
	5.21σ	0.51 σ	0.02 σ	0.0036 σ

ment in the significance up to $\mathcal{Z}_A = 17.82\sigma$, while for a 1.7 TeV VLQ we could expect up to $\mathcal{Z}_A = 3.68\sigma$. These results demonstrate that for larger masses systematic uncertainties become increasingly relevant and, in the best interest of the LHC physics programme, must be well understood and studied by our experimentalist colleagues. In principle, it is expected that an improvement in the systematic uncertainties is achieved over time opening up the possibility to probe larger VLQ masses above 800 GeV in the context of the scenario discussed in this article. Note that this picture can be potentially improved through the combination of various production channels, namely single-production and double production channels, such as the ones involving decays into top quarks, which may also be relevant for the considered model.

4 Conclusions

In this paper, we continue the DL phenomenological analysis, as first started in our previous work [24], of exotic vector-like fermions that naturally arise in the low-energy limit of a F-GUT model based on the trinification gauge group. In particular, we have focused on possible collider signatures of the weak-isospin-singlet D -type VLQ at the LHC, via pair-production topologies, with subsequent VLQ decay into lighter chiral quarks. The latter topology results in a single charged lepton, light jets and a neutrino in the final state particularly convenient for experimental analysis at the LHC. In order to accomplish this task, we have performed Monte-Carlo simulations and utilized sophisticated DL methods for efficient separation of SM backgrounds from the VLQ signal. This enables us to compute the statistical significance for a hypothetical discovery of the considered VLQ. In the context of the VLQ detection prospects at the LHC, we have utilised an evolutionary algorithm that optimises the construction of a NN based on the maximisation of the statistical significance.

For this purpose, we have exploited both the kinematic tabular data as well as the detector jet images as inputs in the classification procedure through the use of a Hybrid Net. We have found that the combination of both data types improves the accuracy with respect to their individual contribution.

With this technique, we have found that the down-type VLQs with a mass of up to 800 GeV can be excluded both at the run-III and the HL phase of the LHC in the decay channels to light quarks. In particular, we have determined that at $\mathcal{L} = 3000 \text{ fb}^{-1}$ one rules out VLQ at this mass scale with statistical significance of $\mathcal{Z}_A = 16.89\sigma$, assuming systematic uncertainties of $\sigma_b = 15\%$. For this same point, we have found $\mathcal{Z}_A = 13.76\sigma$ for run-III luminosity. Additionally, we have found that the statistical significance of heavier VLQs decaying to light quarks remains below the discovery threshold and an exclusion bound can not be promptly determined for the considered channel. Instead, it was shown that a better control over systematic uncertainties is required in order to open up the possibility to probe larger masses in the light quarks channel.

Acknowledgements The work developed in this article is supported by the projects PTDC/FIS-PAR/31000/2017 and CERN/FIS-PAR/0014/2019. J.G., F.F.F., and A.P.M. are supported by the Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT – Fundação para a Ciência e a Tecnologia), references UIDB/04106/2020 and UIDP/04106/2020. A.P.M. is also supported by national funds (OE), through FCT, I.P., in the scope of the framework contract foreseen in the numbers 4, 5 and 6 of the article 23, of the Decree-Law 57/2016, of August 29, changed by Law 57/2017, of July 19. J.G. is also directly funded by FCT through a doctoral program grant with the reference 2021.04527.BD. R.P. is supported in part by the Swedish Research Council grant, contract number 2016-05996, as well as by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 668679).

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: Due to the big memory size, the data was not deposited. However, it can be requested from the corresponding author. Analysis codes to generate jet

images/kinematic data and the plots of this work is available at <https://github.com/Mrazi09/VLQ-Deep-Learningproject>.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Funded by SCOAP³. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

A Low-scale LS-SHUT model

Here, we consider a possible low-energy realisation of the Yukawa sector of the LS-SHUT model. The full physical Lagrangian of the model was previously shown in Ref. [24]. Here, we are going to specify the relevant terms that impact the (chiral and vector-like) quark sector. The charges of the SM and the VLQs w.r.t. the SM gauge symmetry are shown in Table 4. Note that, in the low-energy limit of the LS-SHUT, the third generation of the VLQs becomes heavy, and can therefore be integrated out below the scale of Left-Right symmetry breaking [22]. Hence, in Table 4, only two generations of VLQs are considered.

The SU(2)_L doublets are defined as

$$Q_L^i = \begin{bmatrix} u_L \\ d_L \end{bmatrix}^i \quad L^i = \begin{bmatrix} \nu_{eL} \\ e_L \end{bmatrix}^i \quad \phi^i = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix}^i, \quad (A.1)$$

Table 4 Quantum numbers of the phenomenologically relevant light SM fermion, light Higgs and VLQ sectors of the LS-SHUT model. Here, Q_L and L are the SM quark and lepton left-handed doublets. d_R , u_R and e_R are the right-handed down, up and lepton SM singlets, respectively. The new physics is encoded in the last two rows, with $D_{L,R}$ corresponding to the VLQs and ϕ to the three generations of Higgs doublets containing the SM-like Higgs boson

Field	SU(3) _C	SU(2) _L	U(1) _Y	# of generations
Q_L	3	2	1/3	3
L	1	2	−1	3
d_R	3	1	−2/3	3
u_R	3	1	4/3	3
e_R	1	1	−2	3
$D_{L,R}$	3	1	−2/3	2
ϕ	1	2	1	3

Bold values indicate charges of non-abelian SU(3)/SU(2)

where Q_L^i with the flavour index $i = 1, 2, 3$ corresponding to three generations of the left-handed SM-like quark SU(2)_L-doublets that arise from the Q_L superfield, whereas the SM-like lepton doublets L^i , come from the L superfield, whose interactions are determined by the superpotential (2.1). Taking into account the representations shown in Table 4, the most general renormalisable Yukawa Lagrangian of the SM-like fermions is given by

$$\begin{aligned} \mathcal{L}_Y = & (Y^a)_{iJ} (\bar{Q}_L)^i (D_R)^J \phi_a + (\Gamma^a)_{ij} (\bar{Q}_L)^i (d_R)^j \phi_a \\ & + (\Delta^a)_{ij} (\bar{Q}_L)^i (u_R)^j \tilde{\phi}_a + \\ & + (\Pi^a)_{ij} (\bar{L})^i (e_R)^j \phi_a + \text{h.c.} \end{aligned} \quad (A.2)$$

Here, we consider only two generations of the VLQs, with $J = 1, 2$, such that Y is a 3×2 Yukawa matrix, whereas

Γ , Δ and Π are the Yukawa 3×3 textures. Note that in the main text we further consider the also valid scenario where only one generation of VLQs is light enough to be present in the low-scale particle spectrum. Due to the vector-like nature of the BSM quarks additional bilinear terms are allowed by the gauge symmetry and read

$$\mathcal{L}_{\text{bil}} = (M_D)_{IJ} (\bar{D}_L)^I (D_R)^J + (\Xi)_{IJ} (\bar{D}_L)^I (d_R)^J. \quad (A.3)$$

Following spontaneous EW symmetry breaking, the mass Lagrangian of the quark fields takes the form

$$\begin{aligned} \mathcal{L}_{q,\text{SB}} = & \frac{v_a}{\sqrt{2}} (Y^a)_{iJ} (\bar{d}_L)^i (\bar{D}_R)^J + \frac{v_a}{\sqrt{2}} (\Gamma^a)_{ij} (\bar{d}_L)^i (\bar{d}_R)^j \\ & + \frac{v_a}{\sqrt{2}} (\Delta^a)_{ij} (\bar{u}_L)^i (\bar{u}_R)^j \\ & + (M_D)_{IJ} (\bar{D}_L)^I (\bar{D}_R)^J + (\Xi)_{IJ} (\bar{D}_L)^I (\bar{d}_R)^J, \end{aligned} \quad (A.4)$$

with v_a corresponding to the Higgs doublet VEVs.

For the up-quark sector, the mass matrix is given as

$$M_u = \frac{v_a}{\sqrt{2}} \begin{bmatrix} \Delta_{11}^a & \Delta_{12}^a & \Delta_{13}^a \\ \Delta_{21}^a & \Delta_{22}^a & \Delta_{23}^a \\ \Delta_{31}^a & \Delta_{32}^a & \Delta_{33}^a \end{bmatrix}, \quad (A.5)$$

whose eigenvalues give rise to the masses of the up, charm and top quarks in the SM. On the other hand, the down-type mass matrix takes the form

$$M_d = \begin{bmatrix} \frac{v_a}{\sqrt{2}} \Gamma_{11}^a & \frac{v_a}{\sqrt{2}} \Gamma_{12}^a & \frac{v_a}{\sqrt{2}} \Gamma_{13}^a & \frac{v_a}{\sqrt{2}} Y_{11}^a & \frac{v_a}{\sqrt{2}} Y_{12}^a \\ \frac{v_a}{\sqrt{2}} \Gamma_{21}^a & \frac{v_a}{\sqrt{2}} \Gamma_{22}^a & \frac{v_a}{\sqrt{2}} \Gamma_{23}^a & \frac{v_a}{\sqrt{2}} Y_{21}^a & \frac{v_a}{\sqrt{2}} Y_{22}^a \\ \frac{v_a}{\sqrt{2}} \Gamma_{31}^a & \frac{v_a}{\sqrt{2}} \Gamma_{32}^a & \frac{v_a}{\sqrt{2}} \Gamma_{33}^a & \frac{v_a}{\sqrt{2}} Y_{31}^a & \frac{v_a}{\sqrt{2}} Y_{32}^a \\ \Xi_{11} & \Xi_{21} & \Xi_{31} & (M_D)_{11} & (M_D)_{12} \\ \Xi_{12} & \Xi_{22} & \Xi_{32} & (M_D)_{21} & (M_D)_{22} \end{bmatrix}. \quad (A.6)$$

In the considered model, NP effects contribute only to the down-quark sector at tree-level through a mixing with down-type VLQs. Where the lightest VLQ is denoted as D , whereas the heaviest VLQ is denoted as S .

B VLQ Feynman rules

In this appendix, we present the relevant Feynman rules, in the mass basis, that contribute to collider analysis in our work. Here, we define g as the $SU(2)_L$ gauge coupling, g_s as the $SU(3)_C$ gauge coupling, δ_{ab} as the Kronecker delta in colour space, δ_{ij} as the Kronecker delta in generation space and λ as the well-known Gell–Mann matrices. The projector operators are defined in the standard way as $P_L = (1 - \gamma_5)/2$ and $P_R = (1 + \gamma_5)/2$. Generation indices run as $i = d, s, b, D, \dots$ and $j = u, c, t$, where dots denote heavier down-type VLQs ignored in the current analysis. Besides, V_{CKM} is the extended CKM matrix defined in (2.3).

$$= \frac{ig}{\sqrt{2}} \delta_{ab} (V_{CKM})_{ij} \gamma^\mu P_L$$

(B. 1)

$$= -\frac{ig_s}{2} \delta_{ij} (\lambda^c)_{ab} \gamma^\mu P_L - \frac{ig_s}{2} \delta_{ij} (\lambda^c)_{ab} \gamma^\mu P_R$$

(B. 2)

C Angular and kinematic distributions at the ATLAS detector

See Figs. 7 and 8.

VLQ dimensionfull variables. M = 800 GeV

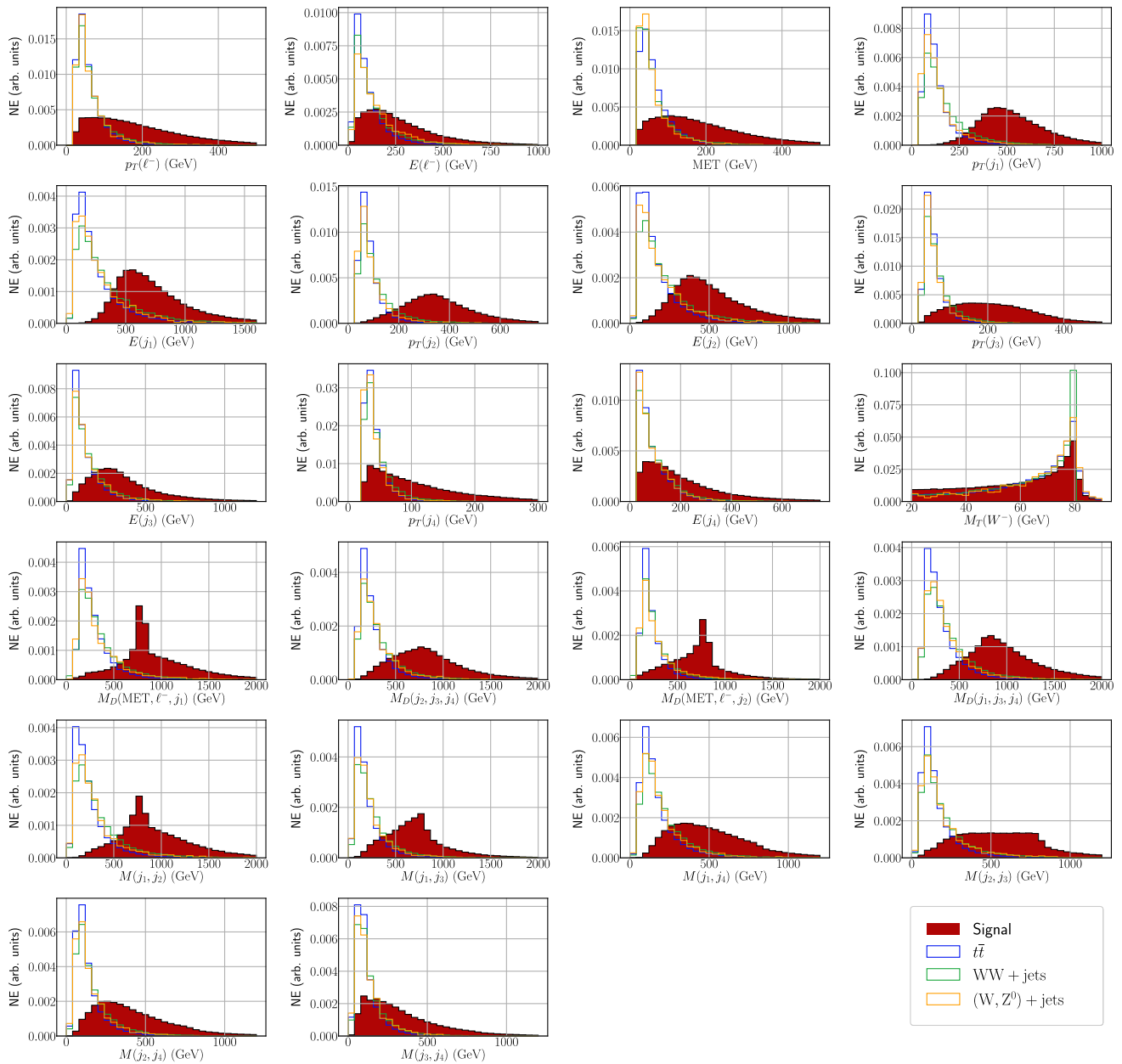


Fig. 7 The dimensionfull kinematic distributions (in units of GeV) for pair-production of 800 GeV VLQs. From top to bottom and left to right, we have $p_T(\ell^-)$, $E(\ell^-)$, MET, $p_T(j_1)$, $E(j_1)$, $p_T(j_2)$, $E(j_2)$, $p_T(j_3)$, $E(j_3)$, $p_T(j_4)$, $E(j_4)$, $M_T(W^-)$, $M(W_{jj}^+)$, $M_D(\text{MET}, \ell^-, j_1)$,

$M_D(j_2, j_3, j_4)$, $M_D(\text{MET}, \ell^-, j_2)$, $M_D(j_1, j_3, j_4)$, $M(j_1, j_2)$, $M(j_1, j_3)$, $M(j_1, j_4)$, $M(j_2, j_3)$, $M(j_2, j_4)$ and $M(j_3, j_4)$. The y-axis represents normalised events, and we select 30 bins for both the signal and the background

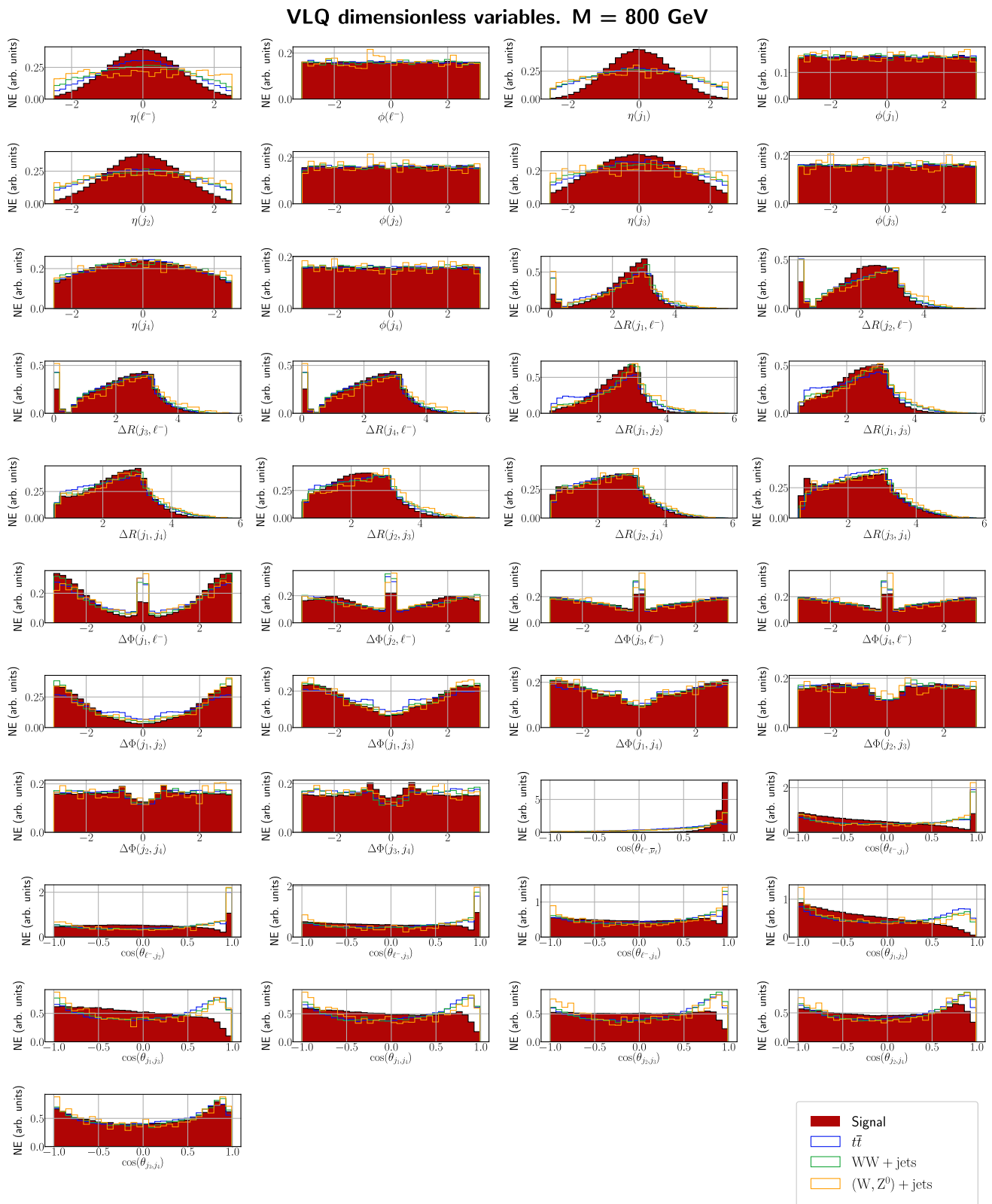


Fig. 8 The dimensionless angular distributions for pair-production of 800 GeV VLQs. From top to bottom and left to right, we have $\eta(\ell^-)$, $\phi(\ell^-)$, $\eta(j_1)$, $\phi(j_1)$, $\eta(j_2)$, $\phi(j_2)$, $\eta(j_3)$, $\phi(j_3)$, $\eta(j_4)$, $\phi(j_4)$, $\Delta R(j_1, \ell^-)$, $\Delta R(j_2, \ell^-)$, $\Delta R(j_3, \ell^-)$, $\Delta R(j_4, \ell^-)$, $\Delta R(j_1, j_2)$, $\Delta R(j_1, j_3)$, $\Delta R(j_1, j_4)$, $\Delta R(j_2, j_3)$, $\Delta R(j_2, j_4)$, $\Delta R(j_3, j_4)$, $\Delta\Phi(j_1, \ell^-)$, $\Delta\Phi(j_2, \ell^-)$, $\Delta\Phi(j_3, \ell^-)$, $\Delta\Phi(j_4, \ell^-)$, $\Delta\Phi(j_1, j_2)$,

$\Delta\Phi(j_1, j_3)$, $\Delta\Phi(j_1, j_4)$, $\Delta\Phi(j_2, j_3)$, $\Delta\Phi(j_2, j_4)$, $\Delta\Phi(j_3, j_4)$, $\cos(\theta_{\ell^-, \bar{\nu}_\ell})$, $\cos(\theta_{\ell^-, j_1})$, $\cos(\theta_{\ell^-, j_2})$, $\cos(\theta_{\ell^-, j_3})$, $\cos(\theta_{\ell^-, j_4})$, $\cos(\theta_{j_1, j_2})$, $\cos(\theta_{j_1, j_3})$, $\cos(\theta_{j_1, j_4})$, $\cos(\theta_{j_2, j_3})$, $\cos(\theta_{j_2, j_4})$ and $\cos(\theta_{j_3, j_4})$. The y-axis represents normalised events and we select 30 bins for both the signal and the background

References

1. ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B* **716**, 1 (2012). [arXiv:1207.7214](#)
2. D0 Collaboration, Observation of the top quark. *Phys. Rev. Lett.* **74**, 2632 (1995). [arXiv:hep-ex/9503003](#)
3. CDF Collaboration, Observation of top quark production in $\bar{p}p$ collisions. *Phys. Rev. Lett.* **74**, 2626 (1995). [arXiv:hep-ex/9503002](#)
4. Gargamelle Neutrino Collaboration, Observation of neutrino like interactions without muon or electron in the Gargamelle neutrino experiment. *Phys. Lett. B* **46**, 138 (1973)
5. UA1 Collaboration, Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Phys. Lett. B* **122**, 103 (1983)
6. UA1 Collaboration, Experimental observation of lepton pairs of invariant mass around 95-GeV/c² at the CERN SPS collider. *Phys. Lett. B* **126**, 398 (1983)
7. UA2 collaboration, Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ collider. *Phys. Lett. B* **129**, 130 (1983)
8. G.W. Bennett, B. Bousquet, H.N. Brown, G. Bunce, R.M. Carey, P. Cushman et al., Final report of the e821 muon anomalous magnetic moment measurement at BNL. *Phys. Rev. D* **73** (2006)
9. Muon $g - 2$ Collaboration, Measurement of the positive muon anomalous magnetic moment to 0.46 ppm. *Phys. Rev. Lett.* **126**, 141801 (2021)
10. LHCb Collaboration, Measurement of the ratio of the $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ -lepton decays. *Phys. Rev. Lett.* **120**, 171802 (2018). [arXiv:1708.08856](#)
11. Belle Collaboration, Measurement of the τ lepton polarization and $r(D^*)$ in the decay $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$. *Phys. Rev. Lett.* **118**, 211801 (2017)
12. BELLE Collaboration, Test of lepton flavor universality and search for lepton flavor violation in $B \rightarrow K\ell\ell$ decays. *JHEP* **03**, 105 (2021). [arXiv:1908.01848](#)
13. LHCb Collaboration, Test of lepton universality in beauty-quark decays. [arXiv:2103.11769](#)
14. G. Arcadi, L. Calibbi, M. Fedele, F. Mescia, Muon $g - 2$ and b anomalies from dark matter. *Phys. Rev. Lett.* **127**, 061802 (2021)
15. I. Doršner, S. Fajfer, O. Sumensari, Muon $g - 2$ and scalar leptoquark mixing. *JHEP* **06**, 089 (2020). [arXiv:1910.03877](#)
16. B. Capdevila, A. Crivellin, S. Descotes-Genon, J. Matias, J. Virto, Patterns of new physics in $b \rightarrow s\ell^+\ell^-$ transitions in the light of recent data. *JHEP* **01**, 093 (2018). [arXiv:1704.05340](#)
17. D. Buttazzo, A. Greljo, G. Isidori, D. Marzocca, B-physics anomalies: a guide to combined explanations. *JHEP* **11**, 044 (2017). [arXiv:1706.07808](#)
18. J.E. Camargo-Molina, A.P. Morais, R. Pasechnik, J. Wessén, On a radiative origin of the standard model from trinification. *JHEP* **09**, 129 (2016). [arXiv:1606.03492](#)
19. J.E. Camargo-Molina, A.P. Morais, A. Ordell, R. Pasechnik, M.O. Sampaio, J. Wessén, Reviving trinification models through an E6-extended supersymmetric GUT. *Phys. Rev. D* **95**, 075031 (2017). [arXiv:1610.03642](#)
20. J.E. Camargo-Molina, A.P. Morais, A. Ordell, R. Pasechnik, J. Wessén, Scale hierarchies, symmetry breaking and particle spectra in SU(3)-family extended SUSY trinification. *Phys. Rev. D* **99**, 035041 (2019). [arXiv:1711.05199](#)
21. A.P. Morais, R. Pasechnik, W. Porod, Grand unified origin of gauge interactions and families replication in the standard model. *Universe* **7**, 461 (2021). [arXiv:2001.04804](#)
22. A.P. Morais, R. Pasechnik, W. Porod, Prospects for new physics from gauge left-right-colour-family grand unification hypothesis. *Eur. Phys. J. C* **80**, 1162 (2020). [arXiv:2001.06383](#)
23. A.E. Cárcamo Hernández, D.T. Huong, S. Kovalenko, A.P. Morais, R. Pasechnik, I. Schmidt, How low-scale trification sheds light in the flavor hierarchies, neutrino puzzle, dark matter, and leptogenesis. *Phys. Rev. D* **102**, 095003 (2020). [arXiv:2004.11450](#)
24. F.F. Freitas, J. Gonçalves, A.P. Morais, R. Pasechnik, Phenomenology of vector-like leptons with deep learning at the large hadron collider. *JHEP* **01**, 076 (2021). [arXiv:2010.01307](#)
25. A. Alves, F.F. Freitas, Towards recognizing the light facet of the Higgs Boson. *Mach. Learn. Sci. Technol.* **1**, 045025 (2020). [arXiv:1912.12532](#)
26. F. Staub, SARAH 4: a tool for (not only SUSY) model builders. *Comput. Phys. Commun.* **185**, 1773 (2014). [arXiv:1309.7223](#)
27. C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, T. Reiter, UFO—the universal Feyn rules output. *Comput. Phys. Commun.* **183**, 1201 (2012). [arXiv:1108.2040](#)
28. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer 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). [arXiv:1405.0301](#)
29. T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., An introduction to PYTHIA 8.2. *Comput. Phys. Commun.* **191**, 159 (2015). [arXiv:1410.3012](#)
30. DELPHES 3 Collaboration, DELPHES 3, a modular framework for fast simulation of a generic collider experiment. *JHEP* **02**, 057 (2014). [arXiv:1307.6346](#)
31. R. Brun, F. Rademakers, ROOT: an object oriented data analysis framework. *Nucl. Instrum. Meth. A* **389**, 81 (1997)
32. NNPDF Collaboration, Parton distributions with QED corrections. *Nucl. Phys. B* **877**, 290 (2013). [arXiv:1308.0598](#)
33. A. Collaboration, ATLAS heavy particle searches—95% CL upper exclusion limits. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-033/fig_01.png (2021)
34. ATLAS Collaboration, Search for large missing transverse momentum in association with one top-quark in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP* **05**, 041 (2019). [arXiv:1812.09743](#)
35. ATLAS Collaboration, Search for new phenomena in events with same-charge leptons and b -jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP* **12**, 039 (2018). [arXiv:1807.11883](#)
36. ATLAS Collaboration, Search for pair production of heavy vector-like quarks decaying into hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D* **98**, 092005 (2018). [arXiv:1808.01771](#)
37. ATLAS Collaboration, Search for pair production of heavy vector-like quarks decaying to high- p_T W bosons and b quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP* **10**, 141 (2017). [arXiv:1707.03347](#)
38. ATLAS Collaboration, Search for single production of vector-like quarks decaying into Wb in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Eur. Phys. J. C* **76**, 442 (2016). [arXiv:1602.05606](#)
39. ATLAS Collaboration, Search for production of vector-like quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP* **08**, 105 (2015). [arXiv:1505.04306](#)
40. ATLAS Collaboration, Search for pair production of a new heavy quark that decays into a W boson and a light quark in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev. D* **92**, 112007 (2015). [arXiv:1509.04261](#)
41. ATLAS Collaboration, Search for heavy vector-like quarks coupling to light quarks in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Lett. B* **712**, 22 (2012). [arXiv:1112.5755](#)
42. ATLAS Collaboration, Combination of the searches for pair-produced vector-like partners of the third-generation quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. Lett.* **121**, 211801 (2018). [arXiv:1808.02343](#)

43. M.L. Mangano, M. Moretti, F. Piccinini, M. Treccani, Matching matrix elements and shower evolution for top-quark production in hadronic collisions. *JHEP* **01**, 013 (2007). [arXiv:hep-ph/0611129](#)
44. C. Muselli, M. Bonvini, S. Forte, S. Marzani, G. Ridolfi, Top quark pair production beyond NNLO. *JHEP* **08**, 076 (2015). [arXiv:1505.02006](#)
45. G. Balossini, G. Montagna, C.M. Carloni Calame, M. Moretti, O. Nicrosini, F. Piccinini et al., Combination of electroweak and QCD corrections to single W production at the Fermilab Tevatron and the CERN LHC. *JHEP* **01**, 013 (2010). [arXiv:0907.0276](#)
46. J.M. Campbell, R.K. Ellis, C. Williams, Vector boson pair production at the LHC. *JHEP* **07**, 018 (2011). [arXiv:1105.0020](#)
47. S. Catani, L. Cieri, G. Ferrera, D. de Florian, M. Grazzini, Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO. *Phys. Rev. Lett.* **103**, 082001 (2009). [arXiv:0903.2120](#)
48. J. Cogan, M. Kagan, E. Strauss, A. Schwartzman, Jet-images: computer vision inspired techniques for jet tagging. *JHEP* **02**, 118 (2015). [arXiv:1407.5675](#)
49. L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, A. Schwartzman, Jet-images — deep learning edition. *JHEP* **07**, 069 (2016). [arXiv:1511.05190](#)
50. H. Qu, L. Gouskos, ParticleNet: jet tagging via particle clouds. *Phys. Rev. D* **101**, 056019 (2020). [arXiv:1902.08570](#)
51. X. Ju et al., Graph neural networks for particle reconstruction in high energy physics detectors. In *33rd Annual Conference on Neural Information Processing Systems*, vol. 3. [arXiv:2003.11603](#) (2020)
52. S. Farrell et al., Novel deep learning methods for track reconstruction. In: *4th International Workshop Connecting the Dots 2018*, vol. 10. [arXiv:1810.06111](#) (2018)
53. A.P. Morais, A. Onofre, F.F. Freitas, J. Gonçalves, R. Pasechnik, R. Santos, Deep learning searches for vector-like leptons at the LHC and electron/muon colliders. [arXiv:2108.03926](#)
54. C. Bonilla, A.E.C. Hernández, J. Gonçalves, F.F. Freitas, A.P. Morais, R. Pasechnik, Collider signatures of vector-like fermions from a flavor symmetric 2HDM. [arXiv:2107.14165](#)
55. A. Elwood, D. Krücker, Direct optimisation of the discovery significance when training neural networks to search for new physics in particle colliders. [arXiv:1806.00322](#)
56. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. [arXiv:1512.03385](#)
57. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan et al., Pytorch: An imperative style, high-performance deep learning library. [arXiv:1912.01703](#)
58. J. Howard, S. Gugger, fastai: a layered API for deep learning. [arXiv:2002.04688](#)
59. G. Cowan, K. Cranmer, E. Gross, O. Vitells, Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C* **71**, 1554 (2011). [arXiv:1007.1727](#)