HL-LHC sensitivity to an ultraheavy S_{uu} diquark in the $u\chi$ channel

Matei S. Filip,[1,2] Calin Alexa,[1,*] Daniel C. Costache,[1,2] Anca M. Dinu,[1,2] Ioan M. Dinu,[1] Ioana Duminica,[1,2] and Gabriel C. Majeri[2]

[1]*Particle Physics Department, IFIN-HH, 077125 Măgurele, Romania*
[2]*Faculty of Physics, University of Bucharest, 077125 Măgurele, Romania*

(Received 22 January 2026; accepted 13 July 2026; published 29 September 2026)

We study the HL-LHC sensitivity to an ultraheavy diquark S_{uu} produced in up-quark fusion and decaying through a vectorlike quark (VLQ) χ as $S_{uu} \to u\chi$, with $\chi \to Wb, Zt, h^0 t$. For fully hadronic decays of the W, Z and top quark, these modes give rise to multijet final states with different jet multiplicities. Within the same model framework used previously for the $S_{uu} \to \chi\chi$ six-jet channel, we consider S_{uu} masses in the multi-TeV range and VLQ masses of the order of a few TeV, and simulate proton–proton collisions at $\sqrt{s} = 13.6$ TeV with integrated luminosities up to the HL-LHC target. The analysis employs a machine learning-based discriminant adapted from the six-jet study to the $S_{uu} \to u\chi$ topology, and is used to derive the corresponding discovery reaches and exclusion limits. We find that the $u\chi$ channel provides a complementary signature for ultraheavy diquark searches, especially in regions where the branching ratio $B(S_{uu} \to u\chi)$ is sizable, thereby broadening the sensitivity program for S_{uu} production at the HL-LHC.

DOI: 10.1103/8vcn-frj7

I. INTRODUCTION

The search for TeV-scale resonances remains an important objective of the ATLAS and CMS programs at the High-Luminosity LHC (HL-LHC). Multijet final states provide sensitivity to a broad class of beyond Standard Model (BSM) scenarios, including color-sextet diquark scalars and vectorlike quarks (VLQs) [1–4]. Recent CMS searches for ultraheavy resonances in four-jet final states have reported sensitivity to masses up to about 9 TeV [1], highlighting the potential of the HL-LHC to extend the reach for new colored resonances. Hints of excesses in dijet and di-dijet searches at the LHC have also been discussed in Refs. [5,6].

One such BSM framework [7–9] contains a color sextet scalar diquark S_{uu} with Yukawa couplings y_{uu} to right-handed up quarks, $y_{u\chi}$ to a Standard Model (SM) up quark and a vectorlike quark χ, and $y_{\chi\chi}$ to same-sign VLQ pairs. The S_{uu} field is a complex scalar with electric charge $+4/3$, while χ is a color-triplet fermion with charge $+2/3$. Depending on the mass hierarchy and coupling strengths, S_{uu} can decay into two up-type quarks, a single VLQ and an up quark, or a pair of VLQs. These channels lead to distinctive hadron-collider signatures, characterized by high jet multiplicity and same-sign content.

In our previous studies, we analyzed the six-jet final state $S_{uu} \to \chi\chi \to (Wb)(Wb) \to (jjb)(jjb)$, developing a machine learning based signal selection [10] and a statistical interpretation of the discovery and exclusion prospects at the HL-LHC [11].

The present study complements those analyses by exploring the single-VLQ channel $S_{uu} \to u\chi$, where the VLQ decays as $\chi \to Wb, Zt, h^0 t$, considering fully hadronic final states for each decay mode. This extends the sensitivity study of S_{uu} production to a new set of multijet topologies and provides the basis for future combined analyses. We assess the HL-LHC sensitivity using a machine learning based signal selection and a statistical framework for discovery and exclusion projections.

The paper is organized as follows. Section II summarizes the BSM theoretical framework and the Monte Carlo generation of signal and background samples, while the machine learning method for signal–background discrimination is discussed in Sec. III A. Section III B presents the selected signal and background yields and discusses the impact of model parameters on the random forest (RF) based analysis. The statistical framework and the corresponding discovery and exclusion results are given in Sec. IV. Finally, Sec. V summarizes the main findings and outlines future prospects.

II. THEORETICAL FRAMEWORK AND EVENT GENERATION

This section summarizes the theoretical framework and event generation setup used in the analysis. We first

*Contact author: calin.alexa@cern.ch

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

introduce the S_{uu} scalar-diquark model and the relevant interaction terms for the $S_{uu} \to u\chi$ decay channel. We then describe the Monte Carlo generation of signal and background samples and the detector-level reconstruction.

Theoretical framework and decay channels The theoretical model underlying this study is the scalar diquark framework introduced in Refs. [7–9]. In addition to the Standard Model fields, the model contains a color-sextet scalar S_{uu}, with $(6, 1, 4/3)$ under $SU(3)_c \times SU(2)_L \times U(1)_Y$, and a vectorlike quark χ, with $(3, 1, 2/3)$. The relevant interaction terms are

$$\frac{y_{uu}}{2} S_{uu} \bar{u}_R u_R^c + \frac{y_{\chi\chi}}{2} S_{uu} \bar{\chi}_R \chi_R^c + y_{u\chi} S_{uu} \bar{u}_R \chi_R^c + \text{H.c.} \quad (1)$$

Here y_{uu}, $y_{u\chi}$, and $y_{\chi\chi}$ denote the Yukawa couplings of S_{uu} to two right-handed up quarks, to an up quark and a vectorlike quark, and to two same-sign vectorlike quarks, respectively. These interactions give rise to the decay modes $S_{uu} \to uu$, $S_{uu} \to u\chi$ and $S_{uu} \to \chi\chi$, depending on the coupling strengths and the available phase space. In this work, we assume that these are the only relevant interactions of the S_{uu} diquark.

The production and decay strengths are controlled by the Yukawa couplings y_{uu}, $y_{\chi\chi}$, and $y_{u\chi}$, together with the masses M_S and m_χ. Thus, the model contains five free parameters. In this study, we reduce this number by fixing the ratio $y_{\chi\chi}/y_{uu} = 1.5$, following the benchmark choice of Ref. [7]. This ratio affects both the total width of the S_{uu} resonance and the relative importance of the uu, $u\chi$ and $\chi\chi$ decay modes. Larger values of $y_{\chi\chi}/y_{uu}$ lead to a broader resonance, while smaller values make the phenomenology more closely aligned with conventional dijet searches.

The present work is devoted to $S_{uu} \to u\chi$, followed by fully hadronic χ decay modes $\chi \to Wb, \chi \to Zt$ and $\chi \to h^0 t$, as shown in Fig. 1. Semileptonic modes are expected to benefit from reduced QCD multijet background due to the presence of isolated leptons and missing transverse momentum, but this comes at the cost of smaller branching fractions and different reconstruction and systematic requirements.

The tree-level partial decay width of the diquark scalar into an up quark and a vectorlike quark is [7]

$$\Gamma(S_{uu} \to u\chi) = \frac{M_S}{16\pi} |y_{u\chi}|^2 \left(1 - \frac{m_\chi^2}{M_S^2}\right)^2 \quad (2)$$

where M_S is the S_{uu} mass and m_χ is the VLQ mass. This expression is obtained at leading-order for a two-body scalar-fermion-fermion decay, neglecting the up-quark mass. For $m_\chi = 2$ TeV, the branching ratio $B(S_{uu} \to u\chi)$ slightly decreases with diquark mass, from 14.2% at $M_S = 7$ TeV to 13.7% at $M_S = 10$ TeV.

Event generation and detector simulation The signal samples were generated with MadGraph5_aMC@NLO v3.3.2

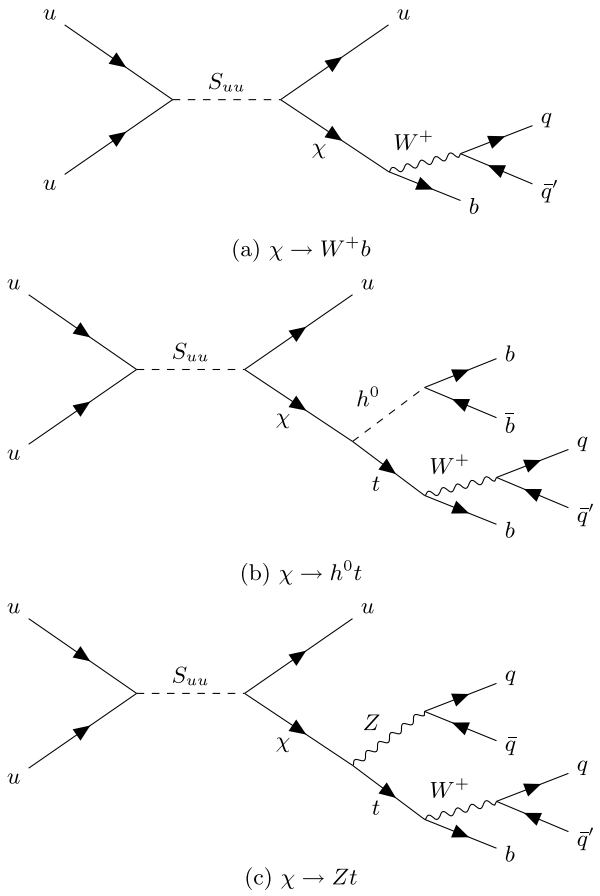


FIG. 1. The Feynman diagrams of the $S_{uu} \to u\chi$ decay modes in fully hadronic final states.

[12] and subsequently passed to Pythia8.310 [13] for parton showering, hadronization, initial- and final-state radiation (ISR/FSR) and multiparton interactions (MPI). The configuration used the NNPDF23LO parton distribution functions (PDF) [14] and the default Monash 2013 tune [15]. The Standard Model background samples, taken to be the same as in Ref. [10], were generated directly with Pythia8.310 using the same shower, hadronization, and tune settings. All signal and background events were processed through DELPHES 3.5.0 [16] using the default ATLAS detector card. Jets were reconstructed with the anti-k_t algorithm implemented in FastJet 3.3.4 [17]. Matching and merging settings were not applied in the present phenomenological setup. Additional jet activity is modeled by the parton shower, which provides an approximate description of soft and collinear radiation. This approximation may affect the modeling of hard extra jets in high-mass multijet final states, and a treatment including matrix-element merging is left for future work.

For consistency with Ref. [10], the same simulation chain was used for background samples. The considered Standard Model backgrounds include fully hadronic $2 \to 2$ QCD, W + jets, Higgs, diboson, and $t\bar{t}$ processes, generated in high-invariant-mass phase space relevant for the S_{uu}

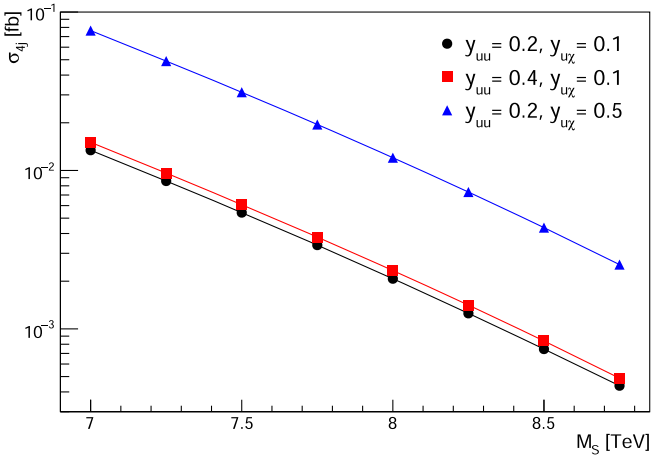


FIG. 2. $pp \to S_{uu} \to u\chi \to u(Wb) \to u(jjb)$ cross section as a function of M_S for three Yukawa coupling scenarios.

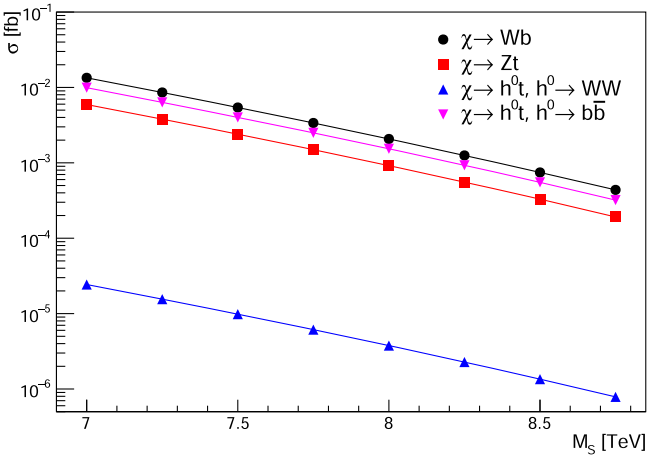


FIG. 3. Fully hadronic cross sections for different χ decay modes in $S_{uu} \to u\chi$ at $\sqrt{s} = 13.6$ TeV, for $y_{uu} = 0.2$ and $y_{u\chi} = 0.1$.

search. An invariant mass threshold was imposed at generation level using the `mHatMin` option, denoted $\hat{m}_{\min}$, in Pythia8.310 [13], in order to efficiently populate the diquark-mass region considered in this analysis. The present work is restricted to fully hadronic final states. Semileptonic modes provide orthogonal search channels defined by mutually exclusive lepton-multiplicity requirements and are therefore left for future studies. This separation allows independent background strategies, dominated by QCD multijets production in the fully hadronic case and by $W+$jets and semileptonic $t\bar{t}$ production in leptonic channels and facilitates a future combined statistical interpretation.

We consider three benchmark coupling scenarios. In the first two, following our previous studies, the product $y_{uu}B(\chi \to Wb)$ is set to 0.1 and 0.2, with $y_{u\chi} = 0.1$. The third benchmark assumes an enhanced $u\chi$ coupling $y_{u\chi} = 0.5$ to assess the sensitivity of the $S_{uu} \to u\chi$ channel in a high-rate scenario. Figure 2 shows the dependence of the four-jet cross section on y_{uu} and $y_{u\chi}$. For fixed $y_{u\chi}$, increasing y_{uu} changes the cross section by about 11%, whereas increasing $y_{u\chi}$ to 0.5 enhances it by a factor of about 6.

TABLE I. $pp \to S_{uu} \to u\chi \to u(Wb) \to u(jjb)$ process cross section vs M_S for $m_\chi = 1.5$, 2 TeV at $\sqrt{s} = 13.6$ TeV, in the conservative $y_{uu} = 0.2$, $y_{u\chi} = 0.1$ scenario.

	$M_S = 7$ TeV	$M_S = 7.5$ TeV	$M_S = 8$ TeV	$M_S = 8.5$ TeV
		$m_\chi = 2$ TeV		
σ_{4j} [fb]	1.34E-02	5.42E-03	2.07E-03	7.46E-04
		$m_\chi = 1.5$ TeV		
σ_{4j} [fb]	1.31E-02	5.30E-03	2.04E-03	7.34E-04

Searches by the ATLAS [4] and CMS [18] Collaborations set lower bounds on same-sign vectorlike quark masses of about 1.36 TeV and 1.4 TeV, respectively, in pair-production processes. As shown in Table I, reducing m_χ from 2 TeV to 1.5 TeV changes the cross section only slightly, by approximately 2.5%.

While the CMS analyses [1,2] target the scenario of $B(\chi \to jj) = 100\%$, we consider three vectorlike quark decay modes with the branching ratios: $B(\chi \to Wb) = 50\%$, $B(\chi \to h^0 t) = 25\%$ and $B(\chi \to Zt) = 25\%$. The corresponding fully hadronic cross sections are shown in Fig. 3, assuming $t \to Wb$ and hadronic decays of the W and Z bosons. In the following, we neglect the $\chi \to h^0 t$, $h^0 \to WW$ contributions, whose cross sections are negligible compared with the other channels.

III. CLASSIFIER SETUP AND YIELD DEPENDENCE ON MODEL PARAMETERS

This section describes the machine learning framework used for signal–background classification and examines how the resulting selected signal and background yields depend on the principal model parameters. We first summarize the classifier configuration, training strategy, and event selection procedure, and then present the variation of postselection event yields across the considered parameter space, including the dependence on the heavy diquark and vectorlike quark masses as well as the relevant coupling assumptions. Together, these elements establish the basis for the subsequent statistical interpretation.

A. Classifier setup

The MC-generated samples are processed with a Machine Learning classifier designed to enhance separation from Standard Model backgrounds. We use the random forest framework developed in Ref. [10], with the classifier

output being used as the main discriminant in the statistical analysis. Because the different χ decay modes lead to different jet multiplicities, the preprocessing is adapted to the signal hypothesis. For the $\chi \to Wb$ channel, the four highest-energy jets are used, while for the $\chi \to h^0 t$ and $\chi \to Zt$ channels the selection is extended to the six highest-energy jets when constructing the corresponding input observables.

The classifier uses 75 features characterizing the event topology. These include the kinematics (p_T, η, ϕ) of the leading jets, all dijet and trijet invariant masses and the corresponding angular separation (ΔR). The feature set also includes global observables such as b-tag and jet multiplicities, the maximum p_T sum of dijet pairs, and the number of dijet combinations compatible with the W-boson mass window. The full list of the 75 input observables used in the random forest training is provided in the Supplemental Material [19].

During the development of this analysis, additional channel-specific observables based on reconstructed Higgs and Z boson invariant masses were tested. However, these variables did not lead to a measurable improvement relative to the baseline set of 75 input observables, since the RF classifier already captures the relevant correlations through the dijet mass combinations included in the feature set. Similarly, the variable associated with the W boson mass window has a limited feature importance, below 3% across the considered topologies. It is nevertheless retained, since hadronically decaying W bosons appear in all signal decay chains studied here.

This behavior is consistent with the feature-importance analysis. The most relevant discriminants across the studied topologies are the transverse momenta of the three leading jets, specific combinatorial dijet invariant masses, and the invariant mass of the dijet pair with the smallest angular separation, $(m_{2j}^{\Delta R_{\min}})$. Together, these leading variables account for more than 40% of the total feature importance, indicating that the classifier is mainly driven by the hard-process kinematics. For the $\chi \to Wb$ channel, the trijet invariant mass also appears among the highest-ranked observables.

By contrast, the transverse momenta of the fourth to sixth leading jets contribute significantly less to the model's decision-making process. This is consistent with the boosted kinematics of the signal, where a large fraction of the event energy is carried by a few hard jets or collimated jet systems rather than being distributed uniformly over higher jet multiplicities. The angular variables η and ϕ have among the lowest feature importances, as they provide less direct separation between the signal and background samples.

The training and validation procedure uses an 80/20 split, with 80% of the simulated events used for training and 20% reserved for testing. To assess the stability of the classifier output and estimate the associated uncertainties, a 5-fold cross-validation is performed. Each event is assigned a continuous RF score $P \in [0, 1]$, interpreted as the probability for the event to be signal-like. The final selection is defined by a discriminator threshold D, such that events with $P > D$ are retained as signal candidates. The selected event counts are converted into physical yields, S_{ev} and B_{ev}, by weighting the simulated samples with the corresponding production cross sections, integrated luminosity, and selection efficiencies.

An optimization of the RF hyperparameters is also performed by evaluating three distinct candidate configurations [11]. Model 0 uses 100 estimators with unrestricted tree depth, employing minimal requirements of 2 samples per split and 1 sample per leaf. Model 1 increases the ensemble size to 250 estimators while introducing stricter regularization through a maximum depth of 10, alongside higher thresholds of 50 samples for splitting and 20 samples per leaf. Model 2 provides an intermediate configuration with 400 estimators, a maximum depth of 12, 20 samples per split, and 5 samples per leaf. A fixed random seed of 42 is used in all cases to ensure reproducibility.

For each signal hypothesis, the optimal configuration is selected by maximizing the scoring function

$$\langle Z_{\text{score}} \rangle = \frac{1}{3} \cdot \sum_{D} \frac{S_{\text{ev}}(D)}{\sqrt{B_{\text{ev}}(D)}} \cdot \frac{1}{1 + \sigma_{B_{\text{ev}}(D)}/B_{\text{ev}}(D)} \tag{3}$$

with $D \in \{0.8, 0.9, 0.925\}$. The first factor represents the signal significance at a given discriminator threshold, while the second penalizes large variations of the background yield across the cross-validation folds. This criterion favors configurations with strong signal-background discrimination and stable background estimates.

B. Dependence of selected yields on the main model parameters

In this subsection, we present the selected signal and background yields obtained after applying the random forest discriminator. The results are studied as a function of the main model and analysis choices, including the vectorlike quark mass, the Yukawa couplings y_{uu} and $y_{u\chi}$, and the χ decay mode. This allows us to assess the stability of the signal selection across the relevant $S_{uu} \to u\chi$ benchmark scenarios.

Dependence on m_χ. As discussed before, lowering the vectorlike quark mass from 2 to 1.5 TeV has only a small effect on the four-jet cross section, with $B(S_{uu} \to u\chi)$ decreasing slightly from 13.9% to 13.7%. The corresponding signal (S_{ev}) and background (B_{ev}) event yields for both mass points, for discriminator thresholds $D \in [0.8, 0.95]$, are shown in Table II. In line with the small change in cross section, S_{ev} is slightly lower for the lighter χ, while the background event yields vary only within statistical uncertainties.

TABLE II. S_{ev} and B_{ev} for $S_{uu} \to u\chi \to u(Wb)$ with $m_\chi = 1.5$, 2 TeV, $M_S = 8$ TeV and $\hat{m}_{min} = 7.5$ TeV.

	$D = 0.80$	$D = 0.90$	$D = 0.925$	$D = 0.95$
		$m_\chi = 2$ TeV		
S_{ev}	6.20	6.02	5.77	5.07
B_{ev}	12.12 ± 0.70	2.53 ± 0.44	1.54 ± 0.25	0.72 ± 0.18
		$m_\chi = 1.5$ TeV		
S_{ev}	6.09	5.82	5.49	4.74
B_{ev}	11.96 ± 0.71	2.66 ± 0.52	1.60 ± 0.26	0.86 ± 0.26

TABLE III. S_{ev} and B_{ev} for $S_{uu} \to u\chi \to u(Wb)$ in the fully hadronic channel, with $m_\chi = 2$ TeV, $M_S = 8$ TeV and $\hat{m}_{min} = 7.5$ TeV, for 3 Yukawa couplings scenarios.

	$D = 0.80$	$D = 0.90$	$D = 0.925$	$D = 0.95$
S_{ev} (i)	6.20	6.02	5.77	5.07
S_{ev} (ii)	6.98	6.78	6.49	5.72
S_{ev} (iii)	36.0	34.9	33.5	29.5
B_{ev}	11.96 ± 0.71	2.66 ± 0.52	1.60 ± 0.26	0.86 ± 0.26

Dependence on Yukawa couplings. We consider three coupling scenarios for the fully hadronic $u\chi$ final state: (i) a conservative choice $y_{uu} = 0.2$, $y_{u\chi} = 0.1$; (ii) an enhanced diquark coupling $y_{uu} = 0.4$ with $y_{u\chi} = 0.1$; (iii) a scenario in which $u\chi$ channel is favored, with $y_{u\chi} = 0.5$, as motivated in Refs. [7–9]. The corresponding signal event yields S_{ev} are summarized in Table III, with the largest S_{ev} obtained for $y_{u\chi} = 0.5$ and the smallest S_{ev} in the conservative scenario. Since the three scenarios differ only in the Yukawa coupling values, while the masses, decay channel, and event selection are kept fixed, the background yields are unchanged.

Dependence on the χ decay mode. For a comprehensive study of fully hadronic final states in the $S_{uu} \to u\chi$ channel, in addition to the $\chi \to Wb$ mode, we also include the decays $\chi \to Zt$ and $\chi \to h^0 t$.

As shown in Table IV, the $\chi \to Wb$ fully hadronic mode yields the highest signal-to-background ratio, driven by its larger branching ratio fraction and production cross section relative to the other channels.

Nevertheless, despite the different multiplicities, the random forest achieves similar background rejection efficiencies for all three channels, within the cross-validation uncertainties.

Taken together, these studies indicate that the search sensitivity depends more strongly on the Yukawa couplings and the χ decay mode than on the precise value of m_χ in the range studied.

TABLE IV. Sampled signal (S_{ev}) and background (B_{ev}) event counts obtained with the random forest model for 3χ decay channels, $y_{uu}B(\chi \to Wb) = 0.1$, $m_\chi = 2$ TeV, $M_S = 8$ TeV and $\hat{m}_{min} = 7.5$ TeV.

	$D = 0.80$	$D = 0.90$	$D = 0.925$	$D = 0.95$
		$S_{uu} \to u\chi \to u(Wb)$		
S_{ev}	6.20	6.02	5.77	5.07
B_{ev}	11.96 ± 0.71	2.66 ± 0.52	1.60 ± 0.26	0.86 ± 0.26
		$S_{uu} \to u\chi \to u(h^0 t), h^0 \to b\bar{b}$		
S_{ev}	4.60	4.47	4.27	3.75
B_{ev}	14.1 ± 0.97	2.28 ± 0.36	1.33 ± 0.21	0.54 ± 0.25
		$S_{uu} \to u\chi \to u(Zt)$		
S_{ev}	2.74	2.62	2.47	2.14
B_{ev}	11.2 ± 0.74	2.31 ± 0.36	1.44 ± 0.29	0.61 ± 0.25

These results show that the highest selected signal yields are obtained for the fully hadronic $\chi \to Wb$ mode and for model parameter choices with enhanced $y_{u\chi}$.

IV. STATISTICAL INTERPRETATION

To assess the observation potential of resonant S_{uu} production in the fully hadronic $u\chi$ channel, we follow the statistical procedure of Ref. [11], implemented with ROOFIT [20] and ROOSTATS [21]. Local p-values are computed using a toy-based profiled-likelihood approach, while 95% CL upper limits on the model dependent signal strength multiplier μ are obtained with the asymptotic CL_s method. The resulting limits, denoted by μ^{95}, are evaluated for each scenario discussed in Sec. III B.

Uncertainty sources and estimates. The uncertainties associated with the PDFs and with the renormalization and factorization scale variations are independently estimated using MadGraph5_aMC@NLO v3.3.2 [12]. Scale uncertainties are evaluated with the standard seven point prescription.

For the signal process, the scale uncertainties vary across the S_{uu} mass range from $^{+19.1\%}_{-15.4\%}$ to $^{+26.3\%}_{-20.6\%}$. The corresponding PDF uncertainties increase with M_S over the interval $M_S \in [7, 9.5]$ TeV, ranging from $\pm 9.61\%$ to $\pm 21.5\%$.

For the background, uncertainty estimates are assigned for the three dominant components after the ML selection: $h^0 + jj$, $W + j$, and $q\bar{q} \to gg$. Since these processes account for more than 90% of the total selected background, they are taken to dominate the background uncertainty budget. The scale uncertainties depend on the process, with the smallest variations obtained for $W + j$, ranging from $^{+31.7\%}_{-20.7\%}$ to $^{+33.2\%}_{-20.8\%}$, and the largest for $h^0 + jj$, ranging from $^{+36.7\%}_{-21.7\%}$ to $^{+43.0\%}_{-22.2\%}$. The PDF uncertainties span a wider range for the backgrounds, from $\pm 27.8\%$

for $W + j$ at $\hat{m}_{\min} = 6.5$ TeV to $\pm 83.5\%$ for $q\bar{q} \to gg$ at $\hat{m}_{\min} = 9$ TeV.

The scenarios with PDF uncertainties reduced by factors of two and four are considered only as illustrative future projections of possible HL-LHC improvements and are not used as the baseline in the sensitivity estimates.

The uncertainties associated with the HL-LHC integrated luminosity are taken to be 1%, following [22].

Although pile-up was not included in the event simulation, we account for its potential impact through an additional systematic uncertainty. This uncertainty was estimated by considering scenarios with 25, 50, and 100 interactions per bunch crossing, yielding variations of up to $\pm 6.1\%$.

The impact of the anti-k_t jet radius parameter was evaluated by varying the nominal value $R = 0.6$ to $R = 0.2$ and $R = 1.0$. The corresponding jet reconstruction uncertainty is estimated to be $\pm 8.0\%$.

The JES and JER uncertainties were estimated with a seven-point variation, analogous to the theoretical scale-variation procedure, yielding a combined effect of $\pm 5.6\%$.

The uncertainties derived from the variation across the ML k-fold cross-validation samples are also included in the statistical analysis.

In the likelihood model, all of the aforementioned uncertainty sources are implemented as normalization nuisance parameters affecting the expected signal and background yields. PDF and scale uncertainties are treated as independent sources. For the background model, nuisance parameters are assigned to the dominant post-selection components mentioned above. The uncertainty sources are treated independently between the signal and background processes, with correlations appearing in the case of JES/JER, pile-up, luminosity and jet clustering cone radius.

Expected exclusion limits. The analysis is performed for a fixed ML discriminator threshold $D = 0.9$. Unlike in the $S_{uu} \to \chi\chi$ channel, the expected limits in the $S_{uu} \to u\chi$ channel show only a weak dependence on y_{uu}, $y_{u\chi}$, and the VLQ mass. The exclusion limits are clustered around $M_S^{\lim} \simeq 8.2$ TeV, with variations of at most 0.1 TeV, as illustrated in Figs. 4 and 5. These results indicate that varying m_χ and the Yukawa couplings within the considered benchmark scenarios does not significantly change the search sensitivity in the $u\chi$ channel.

For the $y_{u\chi} = 0.5$ benchmark, the larger $S_{uu} \to u\chi$ rate significantly improves the mass reach. As shown in Fig. 6, the 95% CL limits extend by about 0.7 TeV, up to $M_S^{\lim} \simeq 8.9$ TeV.

The additional decay channels $\chi \to Zt$ and $\chi \to h^0 t$ are presented in Fig. 7. The scans are performed for $D = 0.9$ and $y_{uu} = 0.2$. For these channels, the mass reach is reduced relative to the $\chi \to Wb$ mode, with intersections at $M_S^{\lim} \simeq 7.6$ TeV for $\chi \to Zt$ and $M_S^{\lim} \simeq 8.0$ TeV for $\chi \to h^0 t$. The $M_S = 8.75$ TeV point is not shown for the

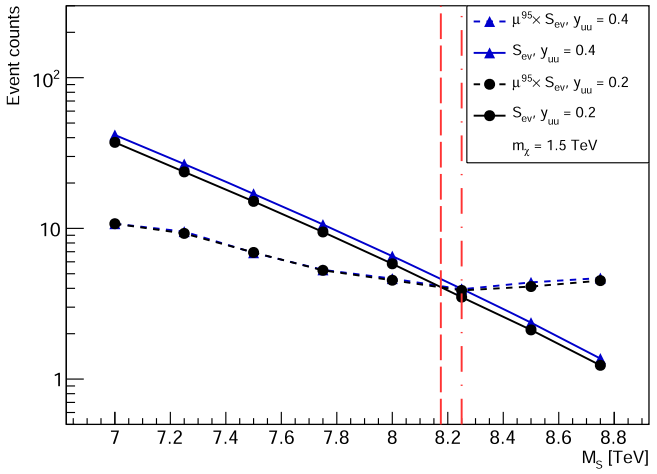


FIG. 4. 95% CL upper limits on $\mu^{95} \times S_{\text{ev}}$ for $\chi \to Wb$, comparing $y_{uu} = 0.2$ and 0.4 at $m_\chi = 1.5$ TeV.

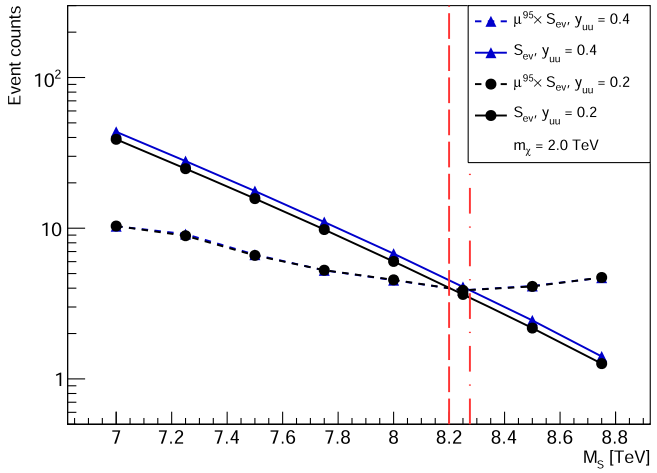


FIG. 5. 95% CL upper limits on $\mu^{95} \times S_{\text{ev}}$ for $\chi \to Wb$, comparing $y_{uu} = 0.2$ and 0.4 at $m_\chi = 2$ TeV.

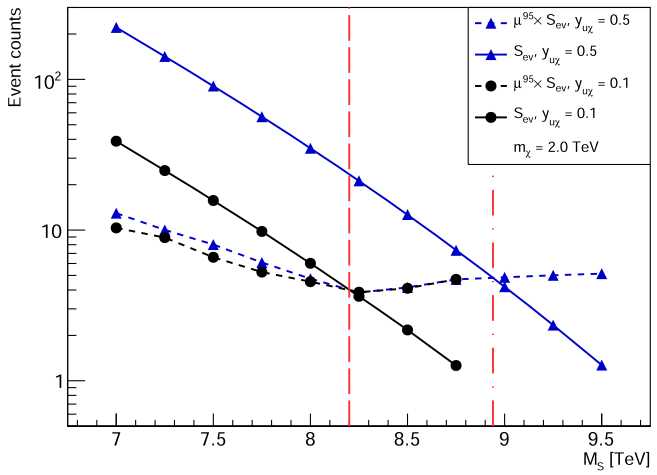


FIG. 6. 95% CL upper limits on $\mu^{95} \times S_{\text{ev}}$ for $\chi \to Wb$, comparing $y_{u\chi} = 0.1$ and 0.5 at fixed $y_{uu} = 0.2$.

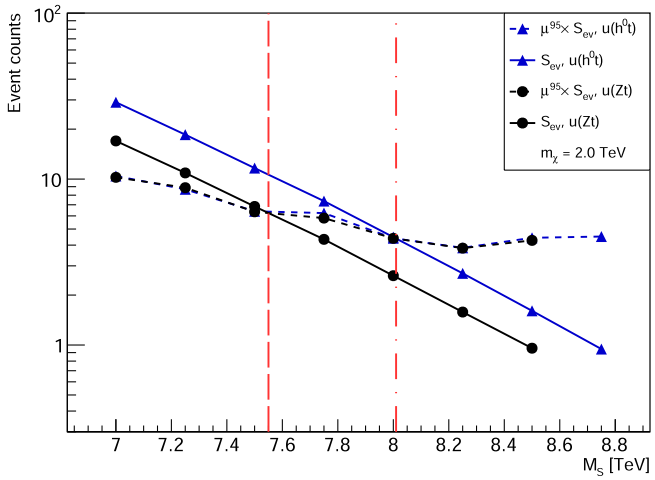


FIG. 7. 95% CL upper limits on μS_{ev} comparing $\chi \to h^0 t$ and $\chi \to Zt$ at $y_{uu} = 0.2$ and $m_\chi = 2$ TeV.

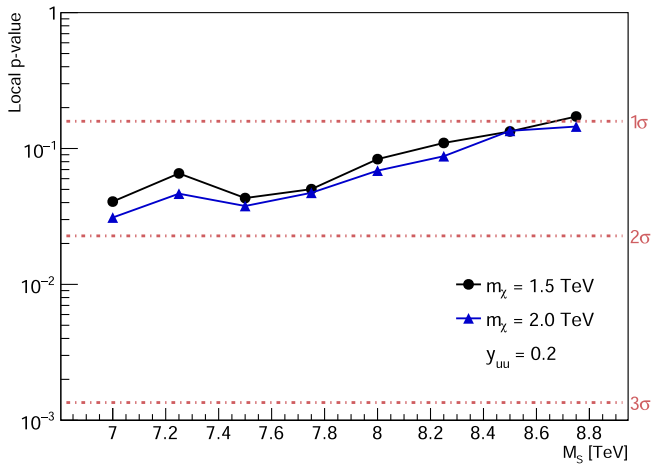


FIG. 8. Local p-value and $y_{uu} = 0.2$, comparing $m_\chi = 1.5$ and 2 TeV.

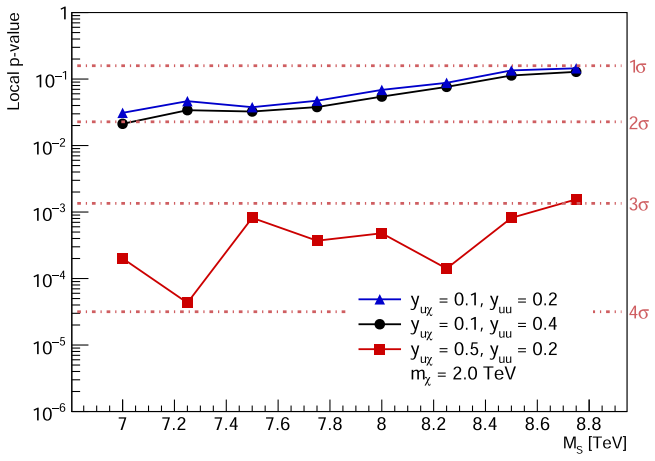


FIG. 9. Local p-value for $\chi \to Wb$ and $m_\chi = 2$ TeV, comparing three different Yukawa coupling combinations.

$u(Zt)$ channel because the expected event yield is well below one event, where the adopted limit-setting procedure is no longer reliable.

Local significances. We also compute the local p-values for a fixed ML discriminator threshold $D = 0.9$. Figure 8 compares the two VLQ mass hypotheses, $m_\chi = 1.5$ TeV and $m_\chi = 2$ TeV. Unlike in our previous study [11], the p-values are slightly smaller for $m_\chi = 2$ TeV, corresponding to a modestly higher discovery sensitivity for the heavier VLQ benchmark. At lower S_{uu} masses, p_0 reaches the $\sim 2\sigma$ level, with very similar results for the two m_χ choices.

The dependence on the Yukawa couplings is shown in Fig. 9. For the enhanced coupling benchmark $y_{u\chi} = 0.5$, the expected local significance reaches the $\sim 3 - 4\sigma$ level. The scenarios with $y_{u\chi} = 0.1$ give lower discovery sensitivity, with the $y_{uu} = 0.4$ case performing slightly better than the conservative benchmark.

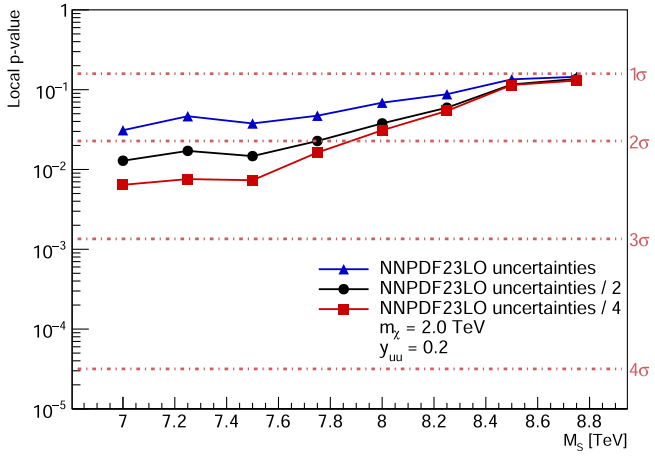


FIG. 10. Local p-value for $\chi \to Wb$, $y_{uu} = 0.2$ and $m_\chi = 2$ TeV, comparing different PDF uncertainty scenarios.

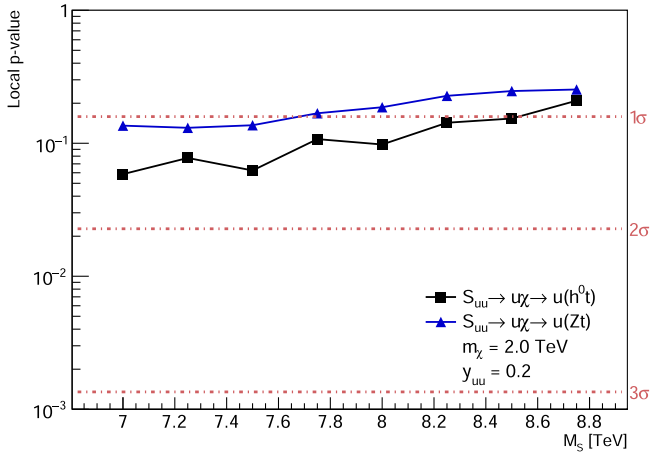


FIG. 11. Local p-value for $\chi \to Wb$, $y_{uu} = 0.2$ and $m_\chi = 2$ TeV, comparing $\chi \to h^0 t$ and $\chi \to Zt$ channels.

In Fig. 10, we study two scenarios in which the current PDF uncertainties are reduced by factors of two and four. These choices are motivated by the expected improvement of PDF constraints at the HL-LHC [23]. A visible improvement is obtained in the conservative benchmark, $y_{uu} = 0.2$, $y_{u\chi} = 0.1$.

For the enhanced coupling benchmark $y_{u\chi} = 0.5$, the reduced PDF uncertainty scenarios lead to extremely small p-values. In this regime, a toy-based evaluation becomes computationally impractical, since an unfeasibly large number of pseudoexperiments would be required to populate the relevant tail of the test statistic distribution. As an estimate, we therefore use the asymptotic formula of Ref. [24], obtaining significances as large as $Z = 5.9\sigma$, corresponding to $p_0 = 1.6 \times 10^{-9}$, for $M_S = 7.25$ TeV.

For the additional decay channels, the sensitivity is considerably weaker than for $\chi \to Wb$, with local significances reaching only about 1-1.5σ, as shown in Fig. 11. This reduction is mainly due to the smaller cross sections of these processes.

Overall, the limits in the fully hadronic $u\chi$ channel show only a weak dependence on m_χ and y_{uu}, but improve significantly for larger $y_{u\chi}$. The $\chi \to Wb$ mode provides the strongest sensitivity, while the $\chi \to Zt$ and $\chi \to h^0 t$ channels give weaker limits and lower local significances. The results are mainly affected by the signal rate and by the treatment of large PDF uncertainties at high masses, emphasizing the importance of improved PDF constraints for HL-LHC projections.

V. CONCLUSIONS

With the High Luminosity LHC upgrade, the more than one order of magnitude increase in integrated luminosity will substantially enhance the reach for high-mass resonances. In this context, diquark scalars and vectorlike quarks have attracted renewed interest, especially following the CMS and ATLAS reports of high invariant mass multijet events [4,18].

In this work, we investigated the $S_{uu} \to u\chi$ decay channel, whose cross section is smaller than that of the $S_{uu} \to \chi\chi$ mode studied previously in Refs. [10,11]. We used the same simulation and analysis framework to generate signal and background samples and to perform signal–background discrimination. Fully hadronic final-state cross sections and signal-selection efficiencies were compared for two vectorlike quark masses, three benchmark choices of Yukawa couplings, and different χ decay modes.

Lowering the vectorlike quark mass from 2 to 1.5 TeV has only a small effect: the selected signal yield decreases slightly with m_χ, in agreement with the theoretical expectations of Ref. [7]. Among the three Yukawa benchmarks, the enhanced $y_{u\chi}$ scenario provides the best sensitivity in the $u\chi$ channel. For all considered χ decay modes, the random forest classifier achieves good discrimination performance, supporting the motivation for a future combined analysis including multiple multijet final state channels.

In the fully hadronic $S_{uu} \to u\chi \to Wb$ channel, the 95% CL exclusion limits are around $M_S^{\lim} \simeq 8.2$ TeV for $m_\chi = 1.5$, 2 TeV and $y_{uu}B(\chi \to Wb) = 0.1$, 0.2, with a discriminator threshold $D = 0.9$. The mass reach is reduced for the other χ decay modes, yielding $M_S^{\lim} \simeq 7.6$ TeV for $\chi \to Zt$, and $M_S^{\lim} \simeq 8.0$ TeV when $\chi \to h^0 t$. In the benchmark with enhanced $y_{u\chi}$, the 95% CL limit improves to $M_S^{\lim} \simeq 8.9$ TeV.

Compared with the previously studied $S_{uu} \to \chi\chi$ channel, the $u\chi$ channel gives slightly weaker limits for the common benchmark points. For example, for $y_{uu} = 0.2$ and $m_\chi = 2$ TeV, the previous $\chi\chi$ analysis yielded $M_S^{\lim} \simeq 8.3$ TeV, compared with $M_S^{\lim} \simeq 8.2$ TeV in the present $u\chi$ study. The discovery sensitivity is also somewhat lower in the $u\chi$ channel, mainly because of the smaller production cross section and correspondingly lower signal yields.

Although we adopt a different theoretical benchmark, the exclusion limits obtained in this work are broadly consistent with the CMS search for resonantly produced dijet pairs via broad mediators [1], which reports four-jet invariant-mass limits of about 8 TeV for narrow resonances and 9 TeV for broader resonances. Since this is a phenomenological study, our statistical results should be understood as estimates of the model sensitivity to the diquark and vectorlike quark masses, rather than as definitive experimental exclusions.

The local significances range from about 1σ to 4σ, depending on the decay mode, Yukawa couplings, M_S, and m_χ. These results show that the $S_{uu} \to u\chi$ channel provides a complementary signature for ultraheavy diquark searches at the HL-LHC.

A natural next step is to include this channel in a combined diquark search program, together with the previously studied $\chi\chi$ and uu final states, within a unified likelihood framework. Since the $u\chi$, $\chi\chi$, and uu channels probe complementary branching-ratio and coupling regimes, such a combined likelihood is expected to improve the global coverage of the S_{uu} parameter space, although the gain in mass reach is likely to be moderate for the benchmark points studied here. Such a combination, incorporating realistic detector effects and systematic uncertainties, would provide more robust constraints on the S_{uu} mass and its Yukawa couplings, while improving the overall discovery potential through the use of multiple signatures. On the theory side, extending the study to include next-to-leading order (NLO) QCD corrections, a more detailed treatment of PDF uncertainties at large x, and alternative flavor structures for the S_{uu} couplings would further clarify the model dependence of the results. These

developments would help establish ultraheavy diquark scalars and vectorlike quarks as benchmark targets for the HL-LHC multijet physics program.

ACKNOWLEDGMENTS

The work of M.-S. F., C. A., D.-C. C., A.-M. D, I.-M. D., and I. D. was supported by IFIN-HH under Contract No. PN-23210104 with the Romanian Ministry of Education and Research. The work of G. M. was supported by the project "Romanian Hub for Artificial Intelligence–HRIA", Smart Growth, Digitization and Financial Instruments Program, 2021–2027, MySMIS no. 351416.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

[1] A. Hayrapetyan *et al.* (CMS Collaboration), Search for resonant production of pairs of dijet resonances through broad mediators in proton-proton collisions at $\sqrt{s}=13$ TeV, arXiv:2507.17884.
[2] A. Tumasyan *et al.* (CMS Collaboration), Search for resonant and nonresonant production of pairs of dijet resonances in proton-proton collisions at $\sqrt{s}=13$ TeV, J. High Energy Phys. 07 (2023) 161.
[3] G. Aad *et al.* (ATLAS Collaboration), Pursuit of paired dijet resonances in the Run 2 dataset with ATLAS, Phys. Rev. D **108**, 112005 (2023).
[4] G. Aad *et al.* (ATLAS Collaboration), Search for pair-production of vector-like quarks in lepton + jets final states containing at least one b-tagged jet using the Run 2 data from the ATLAS experiment, Phys. Lett. B **854**, 138743 (2024).
[5] A. Crivellin, C. A. Manzari, B. Mellado, and S.-E. Dahbi, Consistency and interpretation of the LHC dijet excesses, Phys. Rev. D **107**, 054045 (2023).
[6] P. Bittar, S. Roy, and C. E. M. Wagner, *r*-parity violation and 8 tev four-jet events at the LHC, Phys. Rev. D **112**, 115030 (2025).
[7] B. A. Dobrescu, R. M. Harris, and J. Isaacson, Ultraheavy resonances at the LHC: Beyond the QCD background, arXiv:1810.09429.
[8] B. A. Dobrescu, LHC probes of the 10 TeV scale, arXiv:1912.13155.
[9] B. A. Dobrescu, TeV-scale particles and LHC events with dijet pairs, J. High Energy Phys. 06 (2025) 113.
[10] I. Duminica, C. Alexa, I. M. Dinu, B. A. Dobrescu, and M.-S. Filip, Ultraheavy diquark decaying into vectorlike quarks at the LHC, Phys. Rev. D **111**, 115025 (2025).
[11] D. C. Costache, C. Alexa, A. M. Dinu, I. M. Dinu, I. Duminica, M. S. Filip, and G. C. Majeri, Observability of an ultraheavy diquark decaying into vectorlike quarks at the LHC, Phys. Rev. D **113**, 115056 (2026).
[12] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, J. High Energy Phys. 07 (2014) 079.
[13] C. Bierlich *et al.*, A comprehensive guide to the physics and usage of Pythia8.3, SciPost Phys. Codebases **2022**, 8 (2022).
[14] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions for the LHC Run II, J. High Energy Phys. 04 (2015) 040.
[15] P. Skands, S. Carrazza, and J. Rojo, Tuning Pythia8.1: The monash 2013 tune, Eur. Phys. J. C **74**, 3024 (2014).
[16] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), DELPHES 3, A modular framework for fast simulation of a generic collider experiment, J. High Energy Phys. 02 (2014) 057.
[17] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, Eur. Phys. J. C **72**, 1896 (2012).
[18] A. Tumasyan *et al.* (CMS Collaboration), Search for pair production of vector-like quarks in leptonic final states in proton-proton collisions at $\sqrt{s}=13$ TeV, J. High Energy Phys. 07 (2023) 020.
[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/8vcn-frj7 for the full list of input features used in the analysis.
[20] W. Verkerke and D. Kirkby, The RooFit toolkit for data modeling, in *Statistical Problems in Particle Physics, Astrophysics and Cosmology* (Imperial College Press, London), pp. 186–189, 10.1142/9781860948985_0039.
[21] L. Moneta, K. Belasco, K. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, The RooStats project, Proc. Sci. ACAT2010 (**2010**) 057. [arXiv:1009.1003].
[22] B. Giacobbe and on behalf of ATLAS LUCID Group, Lucid-3: The upgrade of the atlas luminosity detector for high-luminosity lhc, J. Instrum. **19**, C03053 (2024).
[23] R. A. Khalek, S. Bailey, J. Gao, L. Harland–Lang, and J. Rojo, Towards ultimate parton distributions at the high-luminosity LHC, Eur. Phys. J. C **78**, 982 (2018),
[24] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71**, 1554 (2011).