Chinese Physics C

ACCEPTED MANUSCRIPT • OPEN ACCESS

Probing Singlet Vector-Like Top Quarks in the Hadronic tZ Channel at the HL-LHC using Machine and Deep Learning Architectures

To cite this article before publication: Haroon Sagheer *et al* 2026 *Chinese Phys. C* in press https://doi.org/10.1088/1674-1137/aec2cd

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **Copyright © The Author(s)**.



As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 216.73.216.56 on 19/11/2026 at 08:54

Chinese Physics C

Probing Singlet Vector-Like Top Quarks in the Hadronic tZ Channel at the HL-LHC using Machine and Deep Learning Architectures

Journal:	*Chinese Physics C*
Manuscript ID	CPC-2026-0641.R3
Manuscript Type:	Paper
Date Submitted by the Author:	11-Nov-2026
Complete List of Authors:	Sagheer, Haroon; Riphah International University, Department of Physics Tayyab Javaid, Muhammad Ahmed, Ijaz; Physics Hassan, Ali Ahmad, Farzana
Keywords:	VLQ-T, HL-LHC, 14 TeV, Hadronic Decay, ML
Subject:	TeV physics/new physics < Particle theory:

SCHOLARONE MANUSCRIPTS

Accepted Manuscript

Probing Singlet Vector-Like Top Quarks in the Hadronic tZ Channel at the HL-LHC using Machine and Deep Learning Architectures

Haroon Sagheer,[1,*] M. Tayyab Javaid,[2,†] Ijaz Ahmed,[2,‡] Ali Hassan,[3,§] and Farzana Ahmad[4,¶]

[1]*Riphah International University, Islamabad, Pakistan*

[2]*Federal Urdu University of Arts, Science and Technology, Islamabad, Pakistan*

[3]*Department of Physics and Astronomy "Ettore Majorana", University of Catania, Piazza Università, Italy*

[4]*SETI, & College of Engineering, Konkuk University, Seoul 05029, South Korea*

(Dated: November 10, 2026)

Abstract

We study single production of a singlet vector-like top partner T at the 14 TeV HL-LHC in the channel $pp \to Tj$, with $T \to tZ$, $t \to bjj$ and $Z \to \nu\bar{\nu}$, for $m_T = 1.8$–3.1 TeV at the benchmark $g^* = 0.4$, $R_L = 0.5$. After a hadronic preselection and a χ^2-based reconstruction of the hadronic top quark, signal and background are separated with an XGBoost classifier and a graph neural network (GNN). Both classifiers receive identical event information and are trained on common process-stratified partitions with three random seeds at every mass. Their test AUCs lie in the range 0.984–0.993 and differ by at most 0.0023, so the two architectures perform equivalently once the input information is matched. The background model comprises eighteen Standard Model processes, including $Z(\nu\bar{\nu})$+jets, lost-lepton top-quark, W+jets and $t\bar{t}V$ production, and the high-score background tail is evaluated with simulated samples that are statistically independent of the training. Taking the total background at its 95% confidence-level upper bound, the expected discovery significance at 3000 fb^{-1} is 3.16–4.23 at $m_T = 1.8$ TeV, and three-standard-deviation evidence and 95% confidence-level exclusion are reached up to $m_T \simeq 1.82$ TeV and $m_T \simeq 2.13$ TeV, respectively, with both classifiers. The small signal-to-background ratio in the selected region makes a data-driven constraint of the background normalization essential for realizing this reach. We also give expected limits on $\sigma(pp \to Tj)\,\mathcal{B}(T \to tZ)$, which lie between 41 and 64 fb for $m_T = 1.8$–3.1 TeV and correspond to $g^* \simeq 0.28$ at 1.8 TeV for $R_L = 0.5$.

PACS numbers: 12.60.Fr, 12.60.-i, 14.65.Jk, 14.80.-j

Keywords: Vector-like quark, HL-LHC, XGBoost, graph neural network, collider phenomenology, sensitivity

*Electronic address: haroonsagheer663@gmail.com

†Electronic address: Ch.tayyab119933@gmail.com

‡Electronic address: ijaz.ahmed@fuuast.edu.pk

§Electronic address: ali.hassan@phd.unict.it; Electronic address: ali.hassan@lns.infn.it

¶Electronic address: farzana@konkuk.ac.kr

Accepted Manuscript For Review Only

I. INTRODUCTION

The Standard Model (SM) describes existing collider measurements with remarkable precision, yet it leaves open basic questions about electroweak naturalness, flavor, and the structure of physics at the TeV scale. A broad class of extensions predicts vector-like quarks (VLQs), heavy color-triplet fermions whose left- and right-handed components have the same gauge quantum numbers [3, 15, 18]. Such states arise in composite Higgs models, warped extra dimensions, and little Higgs theories, where gauge-invariant mass terms and mixing with SM quarks can help address Higgs-sector radiative corrections [6, 37, 41].

For a singlet top partner T, the dominant decay channels are $T \to bW$, $T \to tZ$, and $T \to th$, with asymptotic branching ratios dictated by Goldstone equivalence at high mass [3, 17]. At the LHC, pair production is QCD-driven and approximately independent of electroweak mixing, but rapidly loses phase space at multi-TeV masses. Single production is model dependent and directly sensitive to the effective mixing coupling. Dedicated phenomenological studies have investigated the tZ, $Z \to \nu\bar{\nu}$ channel at the LHC and future hadron colliders [32, 34].

ATLAS and CMS searches for singly produced T quarks in $t\bar{Z}$ and tH final states already push singlet-T limits into the TeV range [2, 23–25, 44, 45]. Particularly relevant to the present topology, CMS used 137 fb^{-1} of 13 TeV data in the jets-plus-missing-momentum final state and set observed narrow-width limits on $\sigma\mathcal{B}(T \to tZ)$ ranging from 602 to 15 fb over $m_T = 0.6$–1.8 TeV [24]. ATLAS subsequently used an XGBoost-based boosted-top classifier in the same broad top-plus-missing-momentum signature and excluded $m_T < 1.8$ TeV for $\kappa_T = 0.5$ and $\mathcal{B}(T \to tZ) = 25\%$ [1]. A later ATLAS combination of the Ht and Zt channels excluded masses below 2.1 TeV for the same singlet benchmark [2]. These experimental results use profile-likelihood inference and data-constrained backgrounds; their κ_T convention is not automatically identical to the g^* convention used here. The HL-LHC target luminosity of ~ 3000 fb^{-1} can extend the reach, but only if complex hadronic backgrounds are controlled [5, 14].

Single production is the dominant production mode at the masses of interest. Table I compares the LO cross sections of single production at $R_L = 0.5$ with those of QCD pair production, computed with the same model implementation. At the benchmark coupling $g^* = 0.4$, single production exceeds pair production by a factor of 1 028–49 235 for $m_T = 1.8$–3.1 TeV, and even at $g^* = 0.1$ by at least a factor of 64; next-to-leading-order corrections to pair production do not change this conclusion.

We therefore focus on the single-production topology $pp \to Tj$ with $T \to tZ$, $t \to bW \to bjj$,

Accepted Manuscript
For Review Only

TABLE I: LO cross sections in fb at $\sqrt{s} = 14$ TeV for single production, $pp \to Tj + \bar{T}j$, at $R_L = 0.5$ and two values of g^*, and for QCD pair production, $pp \to T\bar{T}$. Single production scales as $(g^*)^2$.

m_T (TeV)	$\sigma(pp \to Tj)$, $g^* = 0.1$	$\sigma(pp \to Tj)$, $g^* = 0.4$	$\sigma(pp \to T\bar{T})$
1.8	35.4	566	0.550
2.0	25.3	404	0.210
2.2	18.0	288	0.0817
2.5	11.2	179	0.0206
2.7	8.11	130	8.29×10^{-3}
2.9	5.96	95.3	3.40×10^{-3}
3.1	4.34	69.5	1.41×10^{-3}

and $Z \to \nu\bar{\nu}$. The hadronic top decay provides a large branching fraction, while the invisible Z yields a missing-momentum handle. The irreducible and reducible backgrounds of this final state are $t\bar{t}$, tZj, $ZZjj$ and $WZjj$ production, as in the cut-based study of Ref. [34]. Because the lepton veto is never fully efficient, processes with a leptonically decaying W boson also enter through lost leptons. We therefore extend the background model by dileptonic $t\bar{t}$, tW, t-channel single-top, $t\bar{t}W$, $t\bar{t}Z$, W+jets (light and heavy flavour) and WZ+heavy-flavour production, and we include $Z(\nu\bar{\nu})$+jets production with light and heavy-flavour jets, which, together with $t\bar{t}$ and W+jets, forms the main background of searches in this final state [24]. The background model then comprises eighteen processes. Heavy-T kinematics induce correlated structures in jet momenta, angular separations, and event-level variables, motivating multivariate methods beyond a cut-based approach.

Simulation samples were generated with MadGraph5_aMC@NLO and processed with Pythia8 and Delphes [4, 29, 46]. After the hadronic preselection ($N_j \geq 3$, $N_b \geq 1$, $N_\ell = 0$) and kinematic cuts, signal and background are separated with XGBoost [22] and with a GNN in which jets are nodes connected by relational edges [33, 38–40, 43, 47]. To isolate the effect of the architecture, both classifiers receive exactly the same event information: XGBoost reads it as a flat vector and the GNN as a graph. Both are trained on identical event partitions with identical preprocessing, at seven masses and with three random seeds each.

Compared with the cut-based study of Ref. [34], the present analysis adds tree-based and graph-based classifiers within the same 5FS detector-level topology. XGBoost itself is not a novel ingredient for this signature because ATLAS has already used it in a search for a boosted top quark

Accepted Manuscript
For Review Only

plus missing transverse momentum [1]. Our aims are therefore (i) a controlled comparison of the two architectures under matched information, (ii) a complete accounting of signal and background efficiencies and yields at fixed signal-efficiency working points, and (iii) an estimate of the HL-LHC sensitivity in which the sparsely populated high-score background tail is evaluated with simulated samples that were not used in training.

The model is parameterized by (m_T, g^*, R_L). The signal samples are generated at $g^* = 0.4$ and $R_L = 0.5$, and the sensitivity at this benchmark is obtained directly from them. The expected limits on other couplings in Sec. V G use the exact quadratic dependence of on-shell single production on g^*, together with a correction for the finite T width at the benchmark.

The paper is organized as follows. Sec. II summarizes the singlet-T model and parameter definitions. Sec. III details event generation, detector simulation, and selection cuts. Sec. IV presents the multivariate analysis, the background estimate and the statistical treatment. Sec. V presents the classifier performance, the selected yields and the expected sensitivity. Sec. VII concludes with a summary and outlook.

II. THE SINGLET VECTOR-LIKE TOP-QUARK MODEL

We summarize the singlet-T framework used in the analysis and fix our notation. The parameterization is that of the model-independent framework of Ref. [15], whose UFO implementation is used to generate the signal samples; see also Refs. [3, 34].

A. Field Content, Couplings, and Parameterization

We consider a vector-like top partner T with electric charge $+2/3$, transforming as an $SU(2)_L$ singlet. In a gauge-invariant realization, the singlet $T = (T_L, T_R)$ has a vector-like mass term and Yukawa couplings to the SM quark doublets q_L^i,

$$\mathcal{L} \supset -\lambda_i \, \bar{q}_L^i \, \tilde{H} \, T_R - M \, \bar{T}_L T_R + \text{h.c.}, \tag{1}$$

with $\tilde{H} = i\sigma_2 H^*$. After electroweak symmetry breaking, these terms mix T with the up-type quarks. For $\lambda_i v \ll M$, the left-handed mixing angles are $s_L^i \simeq \lambda_i v/(\sqrt{2}M)$, the mass of the heavy eigenstate is $m_T \simeq M$, and the right-handed mixing angles are suppressed by m_{u_i}/m_T [3, 15]. To leading order in s_L^i and neglecting the masses of the light quarks, the couplings of the heavy mass

Accepted Manuscript For Review Only

eigenstate are

$$\mathcal{L}_T = \sum_{i=1,3} s_L^i \left[\frac{g}{\sqrt{2}} \bar{T}_L \gamma^\mu W_\mu^+ d_L^i + \frac{g}{2c_W} \bar{T}_L \gamma^\mu Z_\mu u_L^i - \frac{g\, m_T}{2m_W} \bar{T}_R h\, u_L^i \right] - s_L^3 \frac{g\, m_t}{2m_W} \bar{T}_L h\, t_R + \text{h.c.}, \quad (2)$$

where g is the $SU(2)_L$ gauge coupling, $c_W \equiv \cos\theta_W$, h is the physical Higgs field, $u^i = (u, c, t)$ and $d^i = (d, s, b)$. We assume that T mixes with the first and third generations only (Sec. II C). The mixing angles are parameterized by an overall coupling g^* and a generation-sharing parameter R_L,

$$s_L^i = g^* \sqrt{\frac{\zeta_i}{2}}, \qquad \zeta_1 \equiv \frac{R_L}{1+R_L}, \qquad \zeta_3 \equiv \frac{1}{1+R_L}, \qquad \zeta_1 + \zeta_3 = 1, \quad (3)$$

so that $g^{*2} = 2\left[(s_L^u)^2 + (s_L^t)^2\right]$ and $R_L = (s_L^u/s_L^t)^2$. Thus $R_L = 0$ corresponds to pure third-generation mixing, whereas $R_L \to \infty$ gives pure first-generation mixing. Equations (2) and (3) are the couplings of the UFO implementation of Ref. [15] with $\kappa_T \equiv g^*$, the singlet branching-ratio parameters $\xi_W = 1/2$, $\xi_Z = \xi_H = 1/4$, and $\zeta_u = \zeta_1$, $\zeta_t = \zeta_3$; this implementation is used to generate the signal samples (Sec. III). The kinematic normalization factors of that implementation, which approach unity at large m_T, are omitted in Eq. (2).

B. Partial Widths, Branching Ratios, and Limiting Behavior

The dominant third-generation decay modes of a singlet T are

$$T \to bW, \qquad T \to tZ, \qquad T \to tH, \quad (4)$$

with $T \to dW, uZ, uh$ opening when $R_L \neq 0$. In the limit $m_T \gg m_t, m_W, m_Z, m_h$, Eq. (2) gives [15, 17]

$$\Gamma(T \to W d_i) \simeq \zeta_i \frac{(gg^*)^2}{128\pi} \frac{m_T^3}{m_W^2}, \quad (5)$$

$$\Gamma(T \to Z u_i) \simeq \zeta_i \frac{(gg^*)^2}{256\pi} \frac{m_T^3}{m_Z^2 c_W^2}, \quad (6)$$

$$\Gamma(T \to h u_i) \simeq \zeta_i \frac{(gg^*)^2}{256\pi} \frac{m_T^3}{m_W^2}, \quad (7)$$

where $i = 1, 3$. Since $m_Z c_W = m_W$ at tree level, the asymptotic ratio for each generation is

$$\Gamma(W d_i) : \Gamma(Z u_i) : \Gamma(h u_i) = 2 : 1 : 1. \quad (8)$$

Summing over both generations gives

$$\frac{\Gamma_T}{m_T} \simeq \frac{(gg^*)^2}{64\pi} \frac{m_T^2}{m_W^2}, \quad (9)$$

Accepted Manuscript For Review Only

provided that no nonstandard decay modes are open. For $m_T = 2$ TeV and $g^* = 0.1$, the model implementation gives $\Gamma_T = 26.3$ GeV, in agreement with Eq. (9) to better than 1%. At the benchmark $g^* = 0.4$, Eq. (9) gives $\Gamma_T/m_T \approx 0.21$ at $m_T = 2$ TeV and 0.51 at 3.1 TeV, so the T is a broad resonance at the higher masses. The signal samples are generated with these widths (Sec. III).

The branching ratio for the signal channel is

$$\mathrm{Br}(T \to tZ) = \frac{\Gamma(T \to Zt)}{\Gamma_{\mathrm{tot}}}. \tag{10}$$

In the heavy-mass limit the overall factor $(gg^*)^2$ cancels, and Eqs. (3) and (7) imply

$$\mathrm{Br}(T \to tZ) \longrightarrow \frac{\zeta_3}{4} = \frac{1}{4(1+R_L)}. \tag{11}$$

Finite-mass phase space produces the weak residual m_T dependence visible in Fig. 1.

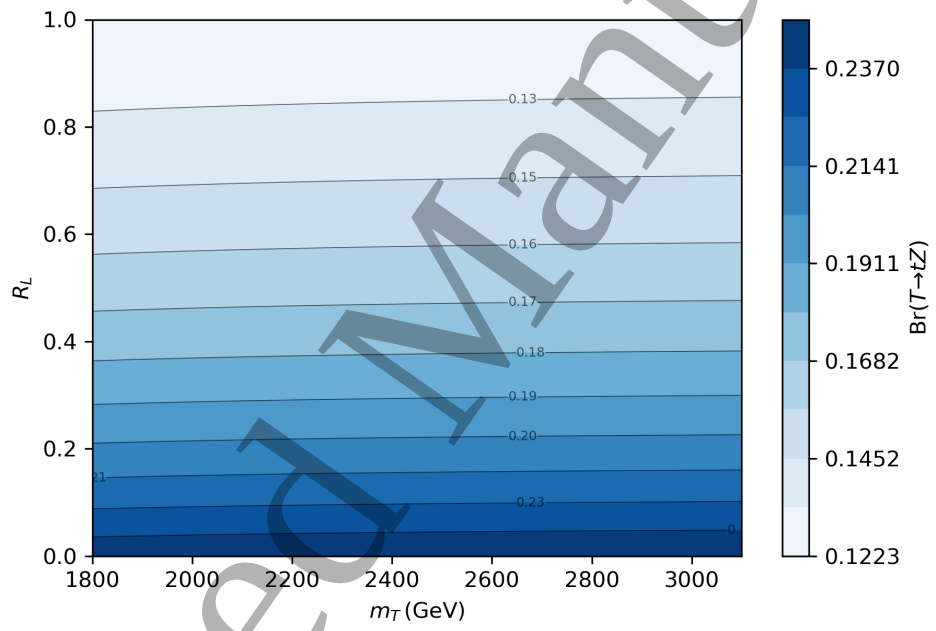


FIG. 1: Branching ratio $\mathrm{Br}(T \to tZ)$ in the (R_L, m_T) plane.

For $R_L = 0$, couplings are purely to the third generation and $\mathrm{Br}(T \to tZ)$ is maximal within the singlet framework. As R_L increases, the coupling is shared with the first generation, reducing $\mathrm{Br}(T \to tZ)$ while opening light-quark modes.

C. Flavour structure and indirect constraints

Accepted Manuscript

Mixing of T with more than one SM generation induces flavour-changing couplings of the Z boson at tree level. In the mass basis, the Z coupling to up-type quarks acquires the non-diagonal term

$$\mathcal{L} \supset -\frac{g}{2c_W}\, s_L^i s_L^j\, \bar{u}_L^i \gamma^\mu u_L^j Z_\mu \qquad (i \neq j), \tag{12}$$

as a consequence of the non-unitarity of the 3×3 up-quark mixing matrix. Mixing with both the first and second generations would therefore generate tree-level u–c transitions and contribute to D^0–$\bar{D}^0$ mixing, which is strongly constrained; following Ref. [34], we set the mixing with the second generation to zero. The remaining constraints on Eq. (3) are the following.

- *CKM unitarity.* Mixing with the up quark reduces the first-row sum to $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1 - (s_L^u)^2$. The measured value, 0.9983 ± 0.0007 [36], implies $(s_L^u)^2 \lesssim 3 \times 10^{-3}$ at 95% confidence level, i.e. $g^* \lesssim 0.13$ for $R_L = 0.5$.

- *Top-quark FCNC.* For $i = 1$, $j = 3$, Eq. (12) induces $t \to uZ$ with

$$\mathcal{B}(t \to uZ) \simeq \frac{(s_L^u s_L^t)^2}{2}\, \frac{(1 - x_Z)^2(1 + 2x_Z)}{(1 - x_W)^2(1 + 2x_W)} \simeq 0.46\, (s_L^u s_L^t)^2, \tag{13}$$

  with $x_V = m_V^2/m_t^2$. The ATLAS limit $\mathcal{B}(t \to uZ) < 6.2 \times 10^{-5}$ for a left-handed coupling [10] requires $s_L^u s_L^t < 0.012$, i.e. $g^* < 0.22$ for $R_L = 0.5$.

- *Electroweak precision data.* Mixing with the top quark contributes to the oblique parameters, dominantly to T; fits require $s_L^t \lesssim 0.16$ for $m_T \approx 1$ TeV, with stronger bounds at larger m_T because the contribution does not decouple at fixed mixing [11, 12]. This corresponds to $g^* \lesssim 0.28$ for $R_L = 0.5$ and $g^* \lesssim 0.23$ for $R_L = 0$ at 1 TeV.

At the benchmark $g^* = 0.4$, $R_L = 0.5$ of Ref. [34], which we adopt for the collider analysis, the mixing angles are $s_L^u = 0.16$ and $s_L^t = 0.23$, and all three bounds are exceeded. In the minimal model, in which T is the only new state, the benchmark therefore serves as a reference point for the signal kinematics rather than as a viable parameter point, and the allowed region corresponds to $g^* \lesssim 0.13$ for $R_L = 0.5$. The indirect bounds rely on this minimal assumption, whereas a direct search tests the production rate itself. In Sec. V G we therefore also express the expected sensitivity as an upper limit on $\sigma(pp \to Tj)\,\mathcal{B}(T \to tZ)$ and on g^*, which can be compared with any parameter point.

Accepted Manuscript
For Review Only

D. Benchmark Choices

The signal samples are generated at the benchmark $g^* = 0.4$ and $R_L = 0.5$ for the mass points $m_T = 1.8$, 2.0, 2.2, 2.5, 2.7, 2.9, and 3.1 TeV.

We adopt the five-flavor scheme (5FS), in which the b quark is treated as a parton in the proton. This is standard for single production of heavy top partners and enhances the relevance of the Tj topology at large masses. Details of event generation, PDF choice, and matching are given in Sec. III.

At fixed R_L, single production of an on-shell T scales quadratically with g^*:

$$\sigma(pp \to Tj)\,\mathrm{Br}(T \to tZ) \propto (g^*)^2\, F(m_T, R_L). \tag{14}$$

For a broad resonance, the rate of the full decay chain deviates from this scaling; the corresponding correction is evaluated in Sec. V G. Changing R_L requires the full production factor $F(m_T, R_L)$, including the different light- and heavy-quark parton luminosities; rescaling only $\mathrm{Br}(T \to tZ)$ is not valid.

III. EVENT GENERATION AND SELECTION

A. Signal and principal background samples

Signal and background samples were generated at LO with MadGraph5_aMC@NLO v3.5.11 [4]. The signal process is

$$pp \to Tj, \qquad T \to tZ, \qquad t \to bW \to bjj, \qquad Z \to \nu\bar{\nu}, \tag{15}$$

including the charge-conjugate mode $pp \to \bar{T}j$. Signal samples were generated at $g^* = 0.4$ and $R_L = 0.5$ for $m_T = 1.8$, 2.0, 2.2, 2.5, 2.7, 2.9 and 3.1 TeV. The principal backgrounds are $t\bar{t}$, tZj, $ZZjj$, and $WZjj$, including charge-conjugate channels. The analysis uses the five-flavor scheme (5FS), in which the b quark is a proton parton.

Parton-level generation cuts are imposed to stabilize the matrix-element integration and reflect basic detector acceptance:

$$\begin{aligned} &p_T(j),\, p_T(b),\, p_T(\ell) > 25 \text{ GeV}, \\ &|\eta(j)| < 5.0, \quad |\eta(b)| < 5.0, \quad |\eta(\ell)| < 2.5, \\ &\Delta R(x, y) > 0.4 \quad (x, y \in \{j, b, \ell\} \text{ where applicable}). \end{aligned} \tag{16}$$

Accepted Manuscript
For Review Only

We use the LHAPDF6 set 331100 (`NNPDF40_nnlo_as_01180`) [9, 16]. Electroweak input parameters are $m_t = 172.52$ GeV, $m_h = 125.11$ GeV, $m_Z = 91.1876$ GeV, $m_W = 80.379$ GeV, $\alpha_s(m_Z) = 0.118$, $1/\alpha(m_Z) = 132.185$, and $G_F = 1.1663787 \times 10^{-5}$ GeV^{-2} [36]. Parton showering and hadronization were performed with PYTHIA8 [46]. No additional multijet merging was applied to the signal; the recoil jet is generated at matrix-element level in $pp \to Tj$.

The samples are normalized to their LO cross sections multiplied by K factors of 1.616 for $t\bar{t}$ [28], 1.10 for tZj, 1.15 for $ZZjj$, 1.14 for $WZjj$ and 0.95 for the signal [19, 31]. All yields refer to an integrated luminosity of 3000 fb^{-1}.

B. Detector simulation and event selection

Detector effects were simulated with DELPHES [29] using its CMS card, with the object definitions adapted to this analysis. The card is listed in full in the material available from the authors; its relevant contents are as follows. Charged particles are reconstructed in the tracker for $|\eta| < 2.5$, and the hadronic calorimeter covers $|\eta| < 5$ with energy resolution $\sigma_E/E = 150\%/\sqrt{E/\text{GeV}} \oplus 5\%$ for $|\eta| < 3$ and $270\%/\sqrt{E/\text{GeV}} \oplus 13\%$ in the forward region. Jets are clustered from particle-flow candidates with the anti-k_T algorithm and $R = 0.4$ [20, 21] and retained for $p_T > 25$ GeV before the jet-energy-scale correction. Electrons (muons) with $p_T > 25$ GeV and $|\eta| < 2.5$ are reconstructed with efficiency 0.95 for $|\eta| < 1.5$ and 0.85 (0.95) for $1.5 < |\eta| < 2.5$, and must satisfy a relative isolation $I < 0.12$ (0.25) within $\Delta R = 0.4$. The missing transverse momentum is the negative vector sum of all particle-flow candidates. The b-tagging efficiency and the misidentification rates are the p_T-dependent parameterizations of the CMS card, listed in Table II. Pileup is not simulated.

The parameterization corresponds to the performance of the Run 2 CMS detector. For comparison, Table II also lists the DELPHES parameterization for HL-LHC studies, which assumes the upgraded tracking of the Phase-2 detectors: a b-tagging efficiency of 0.75 up to 500 GeV and misidentification rates of 0.10 for c jets and 0.01 for light jets. Relative to it, the card used here has a lower b-tagging efficiency at high p_T and higher misidentification rates. The HL-LHC parameterization would therefore increase the signal and the backgrounds with genuine b jets by about 10% and reduce the W+jets and Z+jets backgrounds, in which the b-tagged jet is mostly a misidentified light or charm jet, so the present choice is conservative for this analysis. At the HL-LHC, up to 200 simultaneous interactions per bunch crossing degrade the missing-momentum resolution; the upgraded detectors are designed to mitigate these effects, which are not included in the present study.

Accepted Manuscript
For Review Only

TABLE II: b-tagging efficiency ϵ_b and misidentification rates ϵ_c, ϵ_{light} for $|\eta| < 2.5$ as functions of the jet p_T, for the CMS card used in this analysis and for the DELPHES HL-LHC card.

	CMS card (this analysis)			HL-LHC card		
p_T (GeV)	ϵ_b	ϵ_c	ϵ_{light}	ϵ_b	ϵ_c	ϵ_{light}
50	0.64	0.17	0.012	0.75	0.10	0.010
100	0.71	0.21	0.014	0.75	0.10	0.010
200	0.72	0.20	0.018	0.75	0.10	0.010
500	0.56	0.15	0.029	0.68	0.09	0.009
1000	0.33	0.11	0.048	0.60	0.08	0.008

After detector simulation, events must satisfy a hadronic preselection:

$$N_j \geq 3, \qquad N_b \geq 1, \qquad N_\ell = 0. \tag{17}$$

Jets are ordered in p_T; j_1, j_2, j_3 denote the three leading jets irrespective of flavour, b_1 the leading b-tagged jet, and light jets are the jets without a b tag. The hadronic top quark is reconstructed by testing every b-tagged jet together with every pair of light jets and keeping the combination that minimizes

$$\chi^2_t = \left(\frac{m_{jj} - m_W}{\sigma_W}\right)^2 + \left(\frac{m_{bjj} - m_t}{\sigma_t}\right)^2, \tag{18}$$

with $m_W = 80.379$ GeV, $m_t = 172.52$ GeV, $\sigma_W = 15$ GeV and $\sigma_t = 25$ GeV. The selected b jet, light-jet pair and bjj system define m_{jj}, m_{bjj} and the top-candidate four-momentum used below. The kinematic selection is

$$\begin{aligned}
&p_T(b_1) > 60 \text{ GeV}, \qquad p_T(j_1) > 120 \text{ GeV},\\
&\Delta R(j_1, j_2) < 3.0, \qquad \Delta R(j_1, j_3) < 3.5,\\
&E_T^{\text{miss}} > 160 \text{ GeV}, \qquad N(\text{light jets}) \geq 2,\\
&40 \text{ GeV} < m_{jj} < 140 \text{ GeV}, \qquad \chi^2_t < 60,\\
&80 \text{ GeV} < m_{bjj} < 320 \text{ GeV}.
\end{aligned} \tag{19}$$

These requirements target the reconstructed hadronic top and W candidates while maintaining sensitivity to heavy-T kinematics. Table III gives the cutflow for $m_T = 1.8$ and 2.0 TeV and the four principal backgrounds.

Accepted Manuscript
For Review Only

TABLE III: Cutflow for the signal benchmarks (T1800, T2000) and the four principal backgrounds at 14 TeV. Entries are effective cross sections in fb, including the K factors of Sec. III A, quoted to three significant figures.

Cuts	Signal (fb)	Backgrounds (fb)			
	T1800 (T2000)	$t\bar{t}$	tZj	$ZZjj$	$WZjj$
σ (before cuts)	9.83 (6.68)	3.40×10^5	44.0	217	548
Trigger	5.09 (3.25)	1.40×10^5	23.1	50.2	69.8
$p_T[b_1] > 60\,\text{GeV}$	4.49 (2.85)	9.78×10^4	14.1	27.9	39.8
$p_T[j_1] > 120\,\text{GeV}$	4.44 (2.82)	3.94×10^4	6.09	13.4	23.8
$\Delta R[j_1, j_2] < 3.0$	2.74 (1.71)	2.30×10^4	2.24	8.07	14.0
$\Delta R[j_1, j_3] < 3.5$	1.60 (0.982)	2.07×10^4	1.82	7.29	12.4
$E_T^{\text{miss}} > 160\,\text{GeV}$	1.59 (0.978)	3.36×10^3	0.750	3.14	5.77
N(light jets) ≥ 2	1.55 (0.951)	3.06×10^3	0.675	2.67	5.50
$40 < m_{jj} < 140\,\text{GeV}$	0.890 (0.542)	2.42×10^3	0.401	1.58	3.25
$\chi_t^2 < 60$	0.639 (0.371)	2.33×10^3	0.386	1.36	2.69
$80 < m_{bjj} < 320\,\text{GeV}$	0.585 (0.336)	2.28×10^3	0.379	1.29	2.53

Table III shows the cutflow for two representative benchmarks, $m_T = 1800$ and 2000 GeV, together with the principal backgrounds. The pattern follows the expected physics of the hadronic tZ topology. The E_T^{miss} requirement targets the invisible $Z \to \nu\bar{\nu}$ decay, while the hard recoil and collimated jets from a boosted top motivate the $p_T(j_1)$ and ΔR selections. The m_{jj} window and χ_t^2 enforce consistency with a hadronic $W \to jj$ and $t \to bjj$ reconstruction. Quantitatively (using the summed background columns), the trigger already removes $\sim 59\%$ of the background while keeping ~ 49–52% of the signal. The $p_T(j_1)$ cut yields a further $\sim 60\%$ background reduction with $\sim 99\%$ signal retention, and the E_T^{miss} cut is the single most powerful step, suppressing backgrounds by $\sim 84\%$ while retaining $\gtrsim 99.5\%$ of the signal. The two ΔR cuts reduce the background by $\sim 42\%$ and $\sim 10\%$ per step with ~ 58–62% signal retention, and the m_{jj} window removes another $\sim 21\%$ of background at $\sim 57\%$ signal efficiency. The final top-consistency cuts (χ_t^2 and the m_{bjj} window) are milder ($\sim 3\%$ and $\sim 2\%$ background reduction) but improve the kinematic fidelity of the

Accepted Manuscript
For Review Only

reconstructed top.[1]

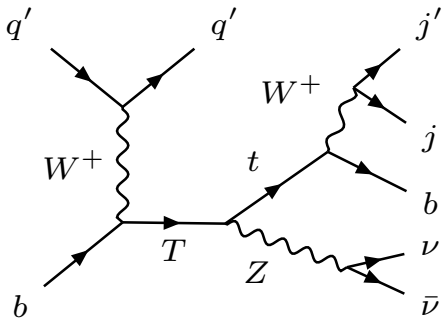


FIG. 2: Representative LO t-channel diagram for single production $pp \to Tj$ in the 5FS with $T \to tZ$, $t \to bW \to bjj$, and $Z \to \nu\bar{\nu}$.

C. Additional background processes

Events with a leptonically decaying W boson pass the lepton veto when the charged lepton is outside the acceptance, not isolated, or below threshold. To include these lost-lepton contributions we use eleven further simulated processes: dileptonic $t\bar{t}$; tW with one or two leptonic W decays; t-channel single-top production; $t\bar{t}W$; $t\bar{t}Z$ with $Z \to \nu\bar{\nu}$; W+jets with light, $b\bar{b}$ and $c\bar{c}$ jets; and $WZ(\nu\bar{\nu})$ with $b\bar{b}$ or $c\bar{c}$ jets. In every sample at least one W boson decays to $e\nu$ or $\mu\nu$. Independent samples of semileptonic $t\bar{t}$, of tZj with a leptonic top decay, and of $WZ(\nu\bar{\nu})jj$ were produced in the same way and are used in Sec. IV D.

These samples were generated with MadGraph5_aMC@NLO v3.6.2 at LO with the NNPDF2.3 LO PDF set (LHAPDF 247000) [13], the jet and b-quark generation cuts of Eq. (16), and Pythia8 showering. The charged leptons from W-boson decays are not restricted at generation level, so lepton losses at all momenta and pseudorapidities are included. Detector simulation used the standard Delphes CMS card, which differs from the analysis card only in the object thresholds and lepton isolation cone. The analysis-card jet and lepton thresholds are imposed exactly at the reconstruction level: jets are retained if their p_T before the jet-energy-scale correction exceeds 25 GeV, and only electrons and muons with $p_T > 25$ GeV and $|\eta| < 2.5$ enter the lepton veto. The selection of Sec. III B, including the χ^2_t assignment of Eq. (18), is then applied

[1] All quoted reductions and efficiencies are step-wise (relative to the preceding cut) and are consistent between the $m_T = 1800$ and 2000 GeV benchmarks in Table III.

Accepted Manuscript
For Review Only

unchanged. The K factors are 1.616 for $t\bar{t}$, 1.15 for tW and W+jets, 1.08 for single top, 1.25 for $t\bar{t}V$ and 1.14 for WZ.

The samples differ from the principal ones in the PDF set, in the treatment of τ leptons and in the lepton isolation cone, all of which affect the fraction of events that pass the lepton veto. These differences are absorbed into a transfer factor measured directly on processes present in both sets. For $p = t\bar{t}$, tZj and $WZjj$, R_p is the ratio of the selected cross section of the principal sample (last row of Table III) to that of the corresponding independent sample. We find $R_{t\bar{t}} = 1.65$, $R_{tZj} = 3.07$ and $R_{WZjj} = 1.84$. The additional top-quark processes are scaled by $R_{t\bar{t}}$ and the W-boson processes by R_{WZjj}. In the conservative estimate of Sec. IV D all eleven are scaled by the largest value, $R_{\max} = 3.07$. Table IV lists the samples. Multijet production with instrumental missing momentum is not simulated; in the CMS analysis of this final state its contribution is smaller than those of $t\bar{t}$, W+jets and Z+jets by more than an order of magnitude [24].

$Z(\nu\bar{\nu})$+jets. Production of a Z boson decaying to neutrinos in association with jets differs from $W(\ell\nu)$+jets production only by the invisible boson decay, and its jet kinematics are obtained from the $W(\ell\nu)$+jets samples above. In events with exactly one reconstructed electron or muon, the lepton is treated as invisible by adding its transverse momentum to $\vec{p}_T^{\,\mathrm{miss}}$, and the selection of Sec. III B is applied to the modified events. The normalization is fixed separately for light, $c\bar{c}$ and $b\bar{b}$ jets by the ratio of the LO cross sections of $Z(\nu\bar{\nu})$+jets and $W(\ell\nu)$+jets, computed with MadGraph5_aMC@NLO with identical process definitions, generation cuts and PDF set: $\sigma_Z/\sigma_W = 0.30$, 0.93 and 1.19, respectively. Events in which the lepton is not reconstructed are restored by dividing by the fraction of W events in the selected topology whose lepton is reconstructed, $A_{\mathrm{reco}} = 0.66$–0.70. The same setup reproduces the cross sections of the $W(\ell\nu)$+jets samples to within 2.0%. A K factor of 1.15 is used, as for W+jets.

Hadronic τ decays in W+jets. The additional samples contain W-boson decays to electrons and muons. Decays to $\tau\nu$ followed by a hadronic τ decay always pass the lepton veto and add a jet. They are obtained from the same $W(\ell\nu)$+jets templates: the reconstructed lepton is replaced by a τ lepton with the same momentum, of which a fraction x is carried by the visible hadrons and forms a jet, while the rest is added to $\vec{p}_T^{\,\mathrm{miss}}$. The fraction x is sampled from the collinear decay spectra, which are flat in $[m^2_{\mathrm{had}}/m^2_\tau, 1]$, for the π, ρ and remaining hadronic modes with their measured branching fractions [36]; ten decay configurations are drawn per event. The normalization uses lepton universality, $\mathcal{B}(\tau \to \mathrm{hadrons}) = 0.648$ [36], and the same reconstruction fraction A_{reco} as above. For the top-quark processes, the corresponding contribution is included through the transfer factor $R_{t\bar{t}}$, which is measured with respect to the principal $t\bar{t}$ sample.

Accepted Manuscript
For Review Only

After the full selection, the additional processes, including $Z(\nu\bar{\nu})$+jets, amount to $\sigma_{\rm sel} = 1\,708$ fb, compared with 2 287 fb for the principal backgrounds, i.e. 75% before the classifier requirement.

TABLE IV: Additional and independent background samples. $N_{\rm gen}$ is the number of generated events, σK the normalized cross section, $N_{\rm sel}$ and $\epsilon_{\rm sel}$ the number and fraction of events passing the full selection of Sec. III B, and $\sigma_{\rm sel} = \sigma K \epsilon_{\rm sel}$. For the additional simulated processes $\sigma_{\rm sel}$ includes the nominal transfer factor R; for $Z(\nu\bar{\nu})$+jets, $N_{\rm sel}$ counts the selected emulated events, σK is the $Z(\nu\bar{\nu})$+jets cross section, and $\sigma_{\rm sel}$ includes the lepton-reconstruction correction $1/A_{\rm reco}$.

Process	$N_{\rm gen}$	σK (fb)	$N_{\rm sel}$	$\epsilon_{\rm sel}$ (%)	$\sigma_{\rm sel}$ (fb)
Additional processes ($\sigma_{\rm sel}$ includes the factor R)					
$t\bar{t}$ (dileptonic)	500 000	44 980	684	0.137	101.3
tW (one lepton)	300 000	10 526	503	0.168	29.05
tW (dileptonic)	300 000	3 511	107	0.036	2.062
tq (t-channel)	1 000 000	44 517	321	0.032	23.52
$t\bar{t}W$	200 000	138.5	3 252	1.626	3.708
$t\bar{t}Z(\nu\bar{\nu})$	200 000	50.22	8 724	4.362	3.606
W+light jets	1 000 000	2.499×10^6	166	0.017	762.0
$Wb\bar{b}j$	750 000	8 747	1 617	0.216	34.64
$Wc\bar{c}j$	750 000	9 557	846	0.113	19.80
$WZ(\nu\bar{\nu})b\bar{b}$	300 000	1.963	2 284	0.761	2.745×10^{-2}
$WZ(\nu\bar{\nu})c\bar{c}$	300 000	2.216	1 330	0.443	1.804×10^{-2}
Independent samples of processes in the original set					
$t\bar{t}$ (one lepton)	1 000 000	2.699×10^5	5 138	0.514	1 387
$tZ(\nu\bar{\nu})j$	200 000	40.42	611	0.305	0.123
$WZ(\nu\bar{\nu})jj$	500 000	581.9	1 182	0.236	1.376
$Z(\nu\bar{\nu})$+jets, emulated from the $W(\ell\nu)$+jets samples					
$Z(\nu\bar{\nu})$+light jets	1 000 000	7.509×10^5	321	0.032	369.0
$Z(\nu\bar{\nu})b\bar{b}j$	750 000	10 280	3 827	0.510	75.29
$Z(\nu\bar{\nu})c\bar{c}j$	750 000	8 640	1 893	0.252	32.41
Hadronic τ decays in W+jets, emulated from the $W(\ell\nu)$+jets samples					
$W(\tau_h\nu)$+light jets	1 000 000	8.097×10^5	1 897	0.019	232.6
$W(\tau_h\nu)b\bar{b}j$	750 000	2 834	23 272	0.310	12.49
$W(\tau_h\nu)c\bar{c}j$	750 000	3 096	10 370	0.138	6.214

Accepted Manuscript
For Review Only

IV. MULTIVARIATE ANALYSIS

The hadronic tZ topology with missing transverse momentum exhibits correlated kinematic and angular structure that can benefit from multivariate classification. Machine-learning methods are now standard tools in collider analyses [8, 33, 38–40, 43].

Both classifiers map the event information $\mathbf{x}$ of Sec. IV A to a score $s(\mathbf{x}) \in [0, 1]$ and are trained by minimizing the same weighted binary cross-entropy,

$$\mathcal{L}_{\mathrm{BCE}} = -\frac{1}{N}\sum_{i=1}^{N} w_i \left[y_i \ln s(\mathbf{x}_i) + (1 - y_i) \ln\big(1 - s(\mathbf{x}_i)\big) \right], \tag{20}$$

where $y_i = 1$ (0) for signal (background) and w_i are the class-balanced training weights of Sec. IV A. The minimizer of Eq. (20) is a monotonic function of the signal-to-background likelihood ratio, so both methods approximate the same optimal discriminant and differ only in the family of functions used to represent it and in its regularization.

XGBoost represents the score as $s(\mathbf{x}) = \sigma\big(\sum_{k=1}^{K} f_k(\mathbf{x})\big)$, with $\sigma(z) = 1/(1+e^{-z})$ and regression trees f_k built sequentially on the 61-component input vector. Each tree is chosen to minimize Eq. (20) plus the complexity penalty [22]

$$\Omega(f_k) = \gamma T_k + \frac{\lambda}{2}\|w_k\|^2, \tag{21}$$

where T_k is the number of leaves of tree k, w_k are its leaf values, and γ, λ control the model complexity.

The GNN represents each event as a graph whose nodes are the four jets of Table V and whose edges connect all ordered pairs of nodes [43, 47]. Node features are first embedded in a hidden space; in each of two message-passing layers, every node then receives a message from each of its three neighbours, computed from the two node embeddings and the edge feature, and updates its embedding from the average message,

$$h_v^{(l+1)} = \phi^{(l)}\left(h_v^{(l)},\ \frac{1}{3}\sum_{u \neq v} \psi^{(l)}\left(h_v^{(l)}, h_u^{(l)}, e_{vu} \right) \right), \tag{22}$$

where $\psi^{(l)}$ and $\phi^{(l)}$ are single-layer neural networks. The embeddings after the last layer are averaged over the nodes, which makes the representation independent of the node ordering, concatenated with the global quantities, and passed through a hidden layer to a sigmoid output. The parameters are obtained by minimizing Eq. (20) with stochastic gradient descent, regularized by weight decay and dropout. The explicit layer definitions are given in Sec. C.

Accepted Manuscript
For Review Only

A. Common inputs and event partitions

Both classifiers receive the same information. The 13 global quantities are the transverse mass of the top candidate and $\vec{p}_T^{\,\rm miss}$,

$$m_T(bjj, E_T^{\rm miss}) = \sqrt{\left(E_T^{bjj} + E_T^{\rm miss}\right)^2 - \left|\vec{p}_T^{\,bjj} + \vec{p}_T^{\,\rm miss}\right|^2}, \qquad E_T^{bjj} = \sqrt{m_{bjj}^2 + \left|\vec{p}_T^{\,bjj}\right|^2}, \tag{23}$$

m_{bjj}, $E_T^{\rm miss}$, H_T, S_T, H_T/S_T, $\Delta\phi(\vec{p}_T^{\,bjj}, \vec{p}_T^{\,\rm miss})$, $\Delta R(j_1, j_2)$, $\Delta R(j_1, j_3)$, $\Delta R(j_2, j_3)$, $\Delta R(b, q_1)$, $\Delta R(b, q_2)$ and $p_T(j_1)$. Here $H_T = p_T(j_1) + p_T(j_2) + p_T(j_3)$, $S_T = H_T + E_T^{\rm miss}$, b is the b jet selected by Eq. (18), and $q_{1,2}$ are the two leading light jets.

The graph has four nodes: j_1, j_2, j_3 and the selected b jet, which may coincide with one of the leading jets. Each node carries $\log(1 + p_T/\text{GeV})$, η, $\sin\phi$, $\cos\phi$, $\log[1 + \max(m, 0)/\text{GeV}]$ and a four-component one-hot label of its role. Every ordered pair of nodes is joined by an edge carrying ΔR computed from the node coordinates. XGBoost receives the node, edge and global quantities as one 61-component vector; the GNN receives the same quantities as node, edge and global tensors. Event identifiers, process labels and weights are not used as inputs. Table V summarizes the inputs.

TABLE V: Input information, identical for XGBoost and the GNN.

Category	Quantities	Components
Global	$m_T(bjj, E_T^{\rm miss})$, m_{bjj}, $E_T^{\rm miss}$, H_T, S_T, H_T/S_T, $\Delta\phi(\vec{p}_T^{\,bjj}, \vec{p}_T^{\,\rm miss})$, $\Delta R(j_1, j_2)$, $\Delta R(j_1, j_3)$, $\Delta R(j_2, j_3)$, $\Delta R(b, q_1)$, $\Delta R(b, q_2)$, $p_T(j_1)$	13
Nodes (j_1, j_2, j_3, b)	$\log(1 + p_T)$, η, $\sin\phi$, $\cos\phi$, $\log(1 + m)$, role label	4×9
Edges	ΔR between all ordered node pairs	12

The signal sample of each mass and every background process are separately divided into 60% training, 20% validation and 20% test events. The background events and their partition assignments are the same at all masses, so no event used for training at one mass is used for validation or testing at another. Each classifier is trained at a single mass. At $m_T = 2$ TeV the training, validation and test partitions contain 12 999, 4 334 and 4 336 events. Continuous inputs are standardized with moments computed on the training partition only. During training the signal weights are scaled so that signal and background have equal total weight; all efficiencies, ROC curves and AUC values use the physical event weights.

B. Classifiers and training

Accepted Manuscript
For Review Only

The hyperparameters are listed in Table VI; no hyperparameter search is performed. XGBoost training stops after 20 rounds without improvement of the validation log loss, and the best iteration is used. The GNN is trained with the Adam optimizer in mini-batches of 512 events, and the checkpoint with the lowest validation loss is kept (patience of ten epochs). Every configuration is trained with the seeds 42, 137 and 2026 on the fixed partitions; seed 42 is used for the figures. The analysis uses XGBoost 3.1.2 and PyTorch 2.9.1. Reloading the 42 saved models reproduces the stored predictions to better than 2×10^{-7}.

TABLE VI: Classifier hyperparameters.

Model	Hyperparameter	Value
XGBoost	maximum trees, depth, learning rate	350, 3, 0.05
XGBoost	row and column subsampling	0.7, 0.7
XGBoost	γ, minimum child weight	0.2, 2
XGBoost	λ, α	1, 0
XGBoost	early stopping	20 rounds
GNN	message-passing layers, node embedding, head	2, 32, 64
GNN	activation, dropout	ReLU, 0.2
GNN	optimizer, learning rate, weight decay	Adam, 10^{-3}, 10^{-4}
GNN	maximum epochs, batch size, patience	60, 512, 10

C. Working points and yields

Score thresholds s_0 are chosen on the validation partition to give signal efficiencies of 50%, 70% and 80% and are then applied unchanged to the test partition; events with $s \geq s_0$ are selected. Yields at 3000 fb^{-1} are

$$S(s_0) = \sum_{i \in \mathrm{sig}} w_i \, \Theta(s_i - s_0), \qquad B(s_0) = \sum_{i \in \mathrm{bkg}} w_i \, \Theta(s_i - s_0), \tag{24}$$

where the sums run over test events and each weight is scaled by the inverse of the test fraction of its process, so that $S(-\infty)$ and $B(-\infty)$ equal the selected yields of Table III. The efficiencies ϵ_S and ϵ_B are defined relative to the same selected yields.

D. Background estimate in the high-score region

At the working points the test partitions of the principal backgrounds retain only a few events, each representing up to $\sim 10^4$ events at 3000 fb^{-1}. The classifier efficiency of the dominant

Accepted Manuscript For Review Only

backgrounds is therefore measured on the independent samples of Sec. III C, which were never used in training and contain several times more selected events with correspondingly smaller weights. For $p = t\bar{t}$, tZj and $WZjj$,

$$B_p(s_0) = \sigma_{\mathrm{sel},p}\, L\, \frac{n_p(s \geq s_0)}{N_{\mathrm{sel},p}}, \tag{25}$$

where $\sigma_{\mathrm{sel},p}$ is the selected cross section of the principal sample, $L = 3000$ fb^{-1}, and n_p of the $N_{\mathrm{sel},p}$ selected independent events pass the threshold. $ZZjj$ uses its own test events. The additional processes contribute $R\,\sigma K L\, n/N_{\mathrm{gen}}$. For W+light jets, whose large cross section is represented by comparatively few selected events, the efficiency is factorized: the fraction passing a loose threshold (98% signal efficiency) is counted on the W+light-jet sample, and the conditional efficiency from the loose threshold to s_0 is taken from the $W{+}b\bar{b}$ and $W{+}c\bar{c}$ samples, which have the same W-boson kinematics. $Z(\nu\bar{\nu})$+jets and the hadronic-τ contribution are treated in the same way, using the emulated events of Sec. III C: the heavy-flavour components are counted directly, and the light-flavour component is factorized with the emulated heavy-flavour events.

These ingredients are validated in Sec. B. The fraction of $t\bar{t}$, tZj and $WZjj$ events passing a given threshold agrees between the principal test sample and the independent sample within the statistical precision of the test sample: the ratios averaged over all masses, seeds and classifiers are 1.01 ($t\bar{t}$), 0.84 (tZj) and 1.22 ($WZjj$). The factorization predicts the directly counted W+light-jet efficiency at 90% and 95% signal efficiency to within a factor of 1.63 and 1.28, respectively.

We quote a central estimate B and a conservative estimate B_{95}. For B_{95}, the transferred components are increased by 25%, which covers the largest closure deviation, all additional simulated processes are scaled by $R_{\max}$, and the factorized W+light-jet component, the $Z(\nu\bar{\nu})$+jets estimate and the hadronic-τ estimate are doubled. Each accepted count n is then assigned a Gamma$(n+1)$ distribution, corresponding to a Poisson likelihood with a flat prior; the weighted components are summed in pseudo-experiments, and B_{95} is the 95th percentile of the total. A process with no accepted events therefore contributes at the level of its sensitivity, rather than zero.

E. Changes with respect to the original analysis

The analysis differs from that of the original submission in the following respects. Each change addresses a specific limitation of the original procedure.

1. *Classifier inputs.* Originally, XGBoost used ten engineered event-level observables, whereas

Accepted Manuscript
For Review Only

the GNN used per-jet node features, pairwise edges and global observables. Both classifiers now receive identical information (Sec. IV A), so that differences in performance can be attributed to the architectures. As a result, the larger AUC of the original GNN (0.983–0.990, compared with 0.976–0.983 for XGBoost) is no longer observed (Sec. V C).

2. *Training and partitions.* The original GNN was trained on the events of all masses together. Since the background events are common to all masses, background events used for validation at one mass were used for training at another. Each classifier is now trained at a single mass, with background partitions that are identical at all masses (Sec. IV A).

3. *Normalization of the yields.* In the original GNN analysis, the significances were computed from event weights summed over the test partition without scaling it to the full sample, so that they did not correspond to 3000 fb^{-1}. All yields are now scaled to the full sample (Sec. IV C).

4. *Working points.* The common threshold $s_0 = 0.99$ corresponds to different signal efficiencies for the two classifiers. Thresholds are now chosen on validation events at fixed signal efficiency (Sec. IV C).

5. *Background model.* The four principal backgrounds are complemented by the processes of Sec. III C, and the high-score background tail is evaluated with independent samples (Sec. IV D).

6. *Interpretation.* The original (g^*, m_T) contours were inconsistent with the significances computed at the generated benchmark, showing five-standard-deviation crossings at couplings below $g^* = 0.4$, where the benchmark itself remained below five standard deviations. They are replaced by results computed directly at the generated benchmark and by limits on $\sigma\mathcal{B}$ and g^* (Secs. V E and V G).

F. Statistical treatment

For a counting experiment with expected signal S and background B, the median discovery significance and the median exclusion significance for the background-only hypothesis are [26, 27]

$$Z_{\text{dis}} = \sqrt{2\left[(S+B)\ln\left(1+\frac{S}{B}\right) - S\right]}, \qquad Z_{\text{exc}} = \sqrt{2\left[S - B\ln\left(1+\frac{S}{B}\right)\right]}. \tag{26}$$

Accepted Manuscript For Review Only

We use $Z_{\rm dis} \geq 5$ for discovery, $Z_{\rm dis} \geq 3$ for evidence and $Z_{\rm exc} \geq 1.645$ for exclusion at 95% confidence level. With a relative background-normalization uncertainty δ_B, $\sigma_B = \delta_B B$,

$$Z_{\rm dis}(\sigma_B) = \sqrt{2\left[(S+B)\ln\left(\frac{(S+B)(B+\sigma_B^2)}{B^2+(S+B)\sigma_B^2}\right) - \frac{B^2}{\sigma_B^2}\ln\left(1+\frac{\sigma_B^2 S}{B(B+\sigma_B^2)}\right)\right]}. \quad (27)$$

The sensitivity is quoted at the 50% signal-efficiency working point, which gives the largest mass reach of the three working points for both classifiers; its threshold is fixed on validation events for each mass and seed. Results are given for the least sensitive of the three training seeds. The mass reach is obtained by log-linear interpolation of Z between the simulated masses. A complete experimental result would require a profile-likelihood fit with nuisance parameters and a CL_s construction [42].

V. RESULTS AND DISCUSSION

Figure 3 shows the global inputs after the selection of Sec. III B. Compared with the background, the signal has much harder $E_T^{\rm miss}$, $p_T(j_1)$, S_T and $m_T(bjj, E_T^{\rm miss})$ distributions and a lower H_T/S_T.

A. Input correlations

The Pearson correlation coefficients of the 13 global inputs, computed with physical event weights after the selection of Sec. III B, are shown separately for the 2 TeV signal and for the background in Figs. 10 and 11 of Sec. A, together with the matrices of all 45 non-constant inputs, including node and edge quantities. The strongest correlations follow from the definitions: $S_T = H_T + E_T^{\rm miss}$, and H_T, S_T and $p_T(j_1)$ are correlated with coefficients of 0.84–0.96 in both samples. Other correlations differ markedly between signal and background. In the signal the recoiling heavy system gives $\rho(E_T^{\rm miss}, H_T) = 0.79$ and $\rho(E_T^{\rm miss}, S_T) = 0.93$, compared with 0.39 and 0.68 in the background, and m_{bjj} is correlated with $m_T(bjj, E_T^{\rm miss})$ (0.50), whereas in the background m_{bjj} is nearly uncorrelated with all other inputs ($|\rho| \leq 0.22$). Such differences in the correlation structure are part of the information used by the classifiers. Correlated inputs are retained, since neither boosted trees nor the GNN require uncorrelated features; the deterministic relation among H_T, $E_T^{\rm miss}$ and S_T adds no information but does not bias either classifier.

B. Training diagnostics

Accepted Manuscript
For Review Only

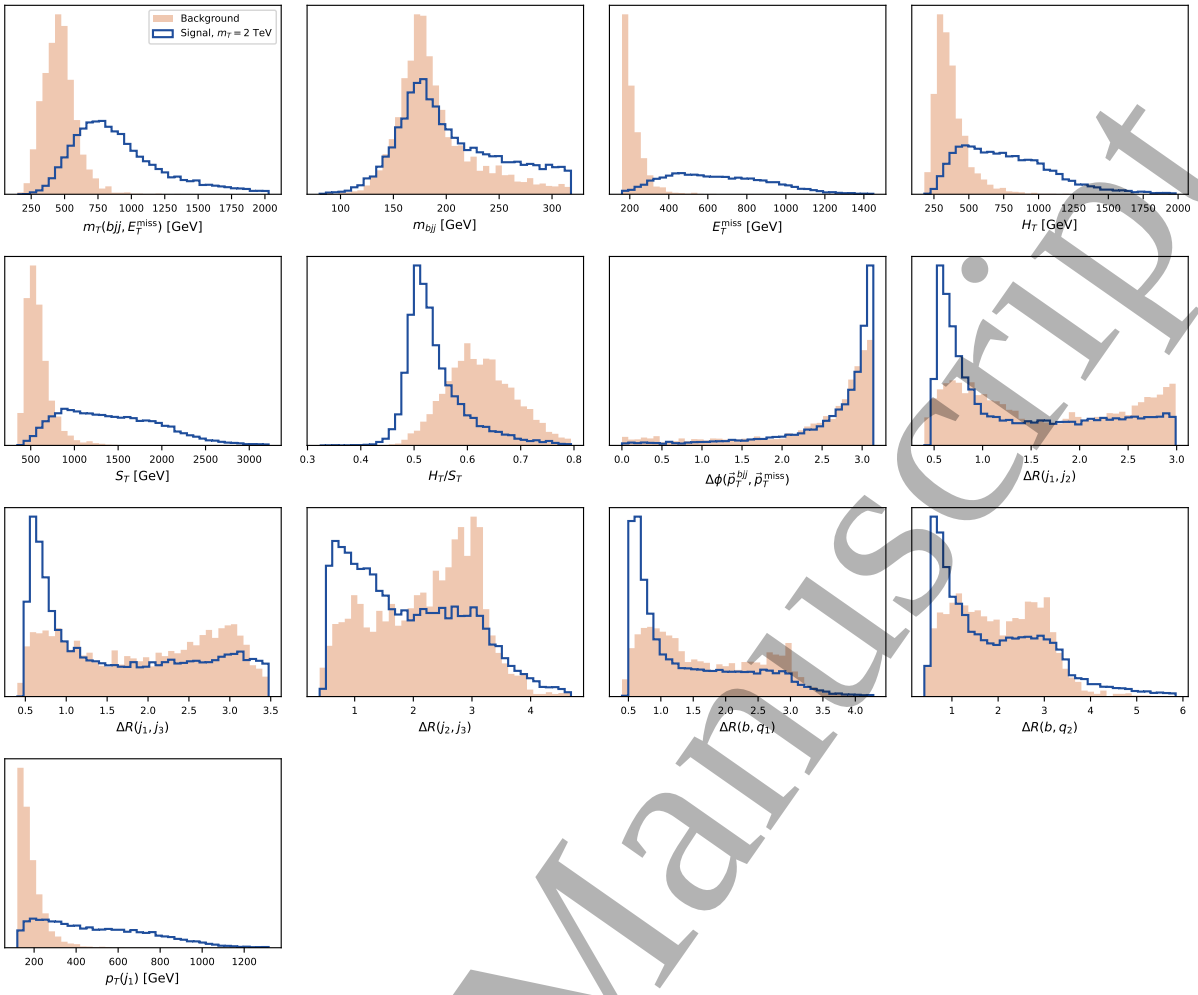


FIG. 3: Distributions of the 13 global inputs for the $m_T = 2$ TeV signal and for the principal backgrounds after the selection of Sec. III B, normalized to unit area with physical event weights.

Figure 4 shows the training and validation loss and AUC as functions of the boosting round and the epoch at $m_T = 2$ TeV. The selected XGBoost round and GNN epoch lie where the validation loss has reached its minimum, and the training–validation difference remains small. Figure 5 compares the score distributions of the training and test partitions separately for signal and background, with a logarithmic ordinate that resolves the high-score region. Training and test distributions agree within statistical fluctuations. The Kolmogorov–Smirnov statistics [35] for the three seeds are given in Table VII; 1 of the 12 p-values lies below 0.05 ($p = 0.018$), consistent with the number of comparisons. The train and test AUCs in Table VIII differ by less than 0.01.

Accepted Manuscript Only

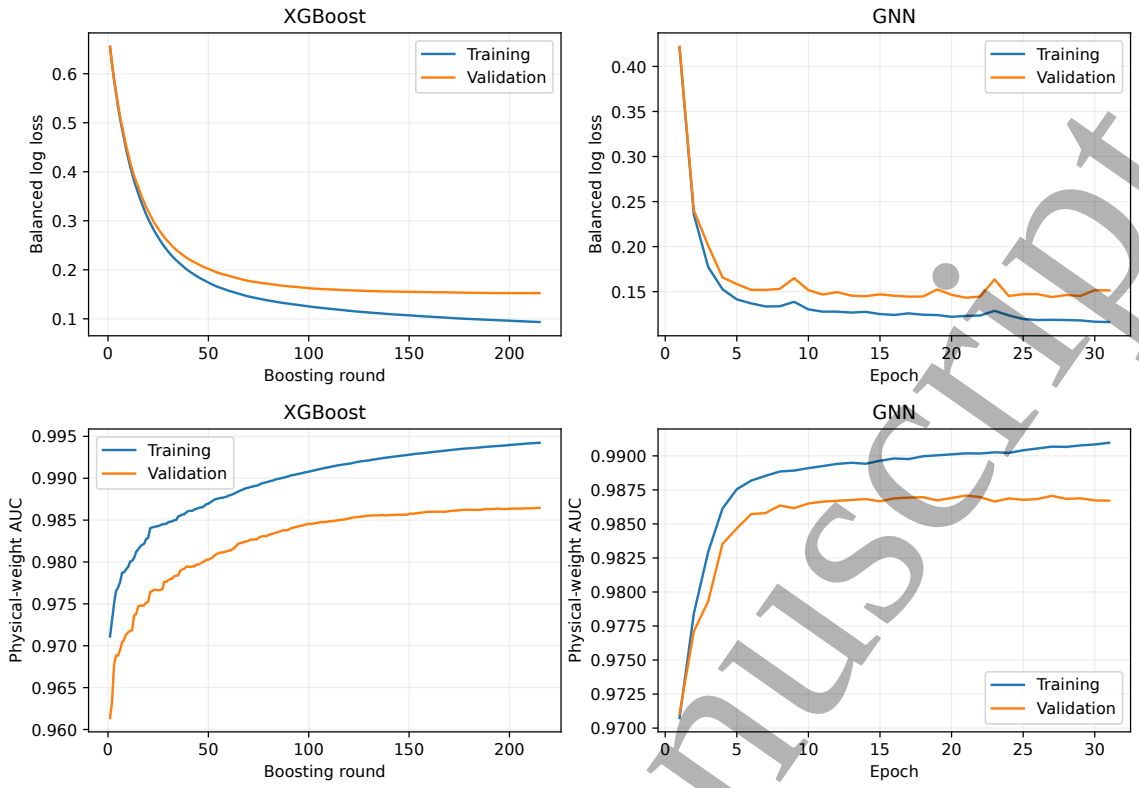


FIG. 4: Training and validation loss (top) and physical-weight AUC (bottom) for XGBoost and the GNN at $m_T = 2$ TeV (seed 42).

Accepted Manuscript
eview Only

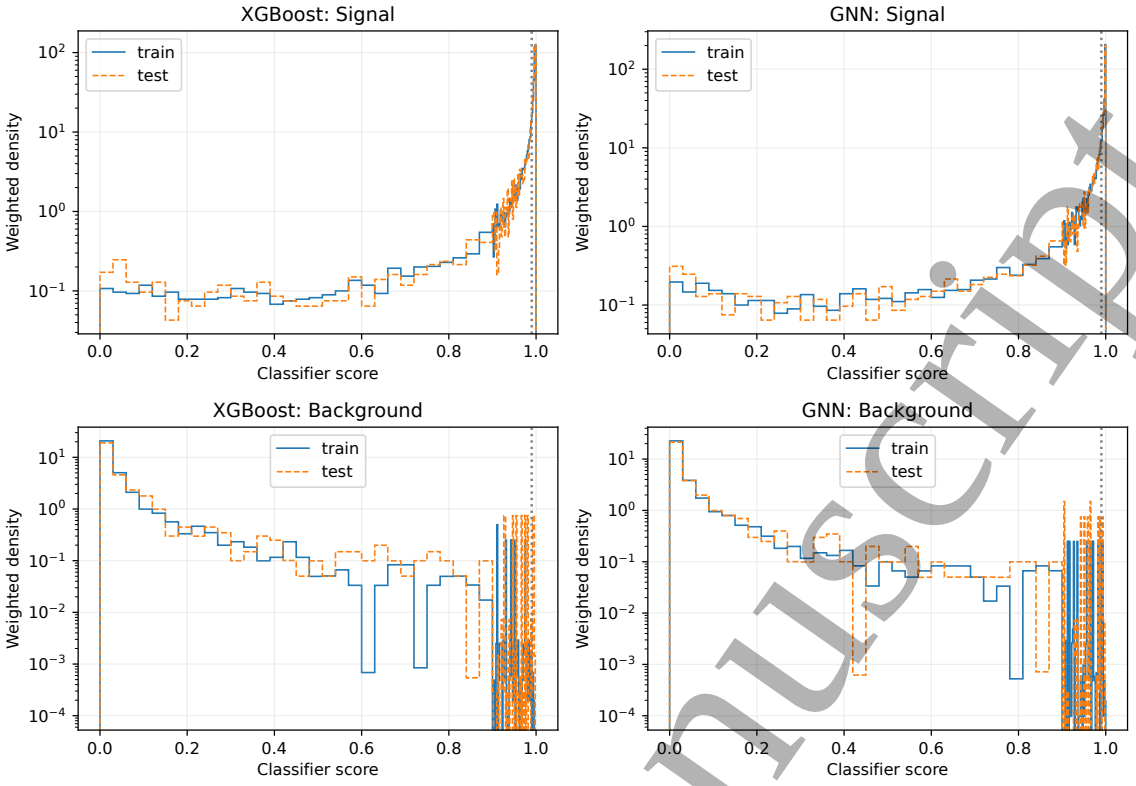


FIG. 5: Classifier-score distributions of the training and test partitions for signal (top) and background (bottom) at $m_T = 2$ TeV (seed 42), each normalized to unit area with physical weights. The dotted line marks $s = 0.99$.

TABLE VII: Kolmogorov–Smirnov comparison of training and test score distributions at $m_T = 2$ TeV, computed on unweighted events (9 333 and 3 111 signal, 3 666 and 1 225 background events).

		Signal		Background	
Seed	Model	D	p	D	p
42	XGBoost	0.0132	0.807	0.0438	0.058
42	GNN	0.0158	0.603	0.0333	0.254
137	XGBoost	0.0131	0.815	0.0504	0.018
137	GNN	0.0134	0.792	0.0396	0.109
2026	XGBoost	0.0168	0.518	0.0399	0.104
2026	GNN	0.0164	0.552	0.0362	0.175

C. Classifier performance

Accepted Manuscript
eview only

Table VIII lists the AUCs at $m_T = 2$ TeV for all seeds, and Table IX the test-AUC range over seeds at every mass. With identical inputs, both classifiers reach test AUCs between 0.984 and 0.993. XGBoost is higher in all 21 paired runs, by 0.0002–0.0023, a difference comparable to the spread between seeds and without effect on the sensitivity below. In the originally submitted analysis the GNN used per-jet information that XGBoost did not receive, and its AUCs (0.983–0.990) exceeded those of XGBoost (0.976–0.983). With matched information this difference disappears, so it reflected the input content rather than the architecture. Figure 6 shows the ROC curves, and Table X the background rejection at fixed signal efficiency.

TABLE VIII: Physical-weight AUC at $m_T = 2$ TeV. The selected step is the boosting round or epoch of the retained model.

Seed	Model	Train AUC	Validation AUC	Test AUC	Selected step
42	XGBoost	0.99386	0.98637	0.98623	195
42	GNN	0.99018	0.98708	0.98449	21
137	XGBoost	0.99485	0.98674	0.98655	235
137	GNN	0.99018	0.98690	0.98428	22
2026	XGBoost	0.99359	0.98639	0.98562	172
2026	GNN	0.99035	0.98694	0.98472	25

TABLE IX: Range of the test AUC over the seeds 42, 137 and 2026 at each mass.

m_T (TeV)	XGBoost	GNN
1.8	0.98735–0.98771	0.98535–0.98603
2.0	0.98562–0.98655	0.98428–0.98472
2.2	0.99209–0.99231	0.99166–0.99173
2.5	0.98936–0.98992	0.98804–0.98838
2.7	0.99069–0.99088	0.98998–0.99045
2.9	0.99330–0.99337	0.99222–0.99263
3.1	0.98948–0.98957	0.98859–0.98887

Accepted Manuscript
For Review Only

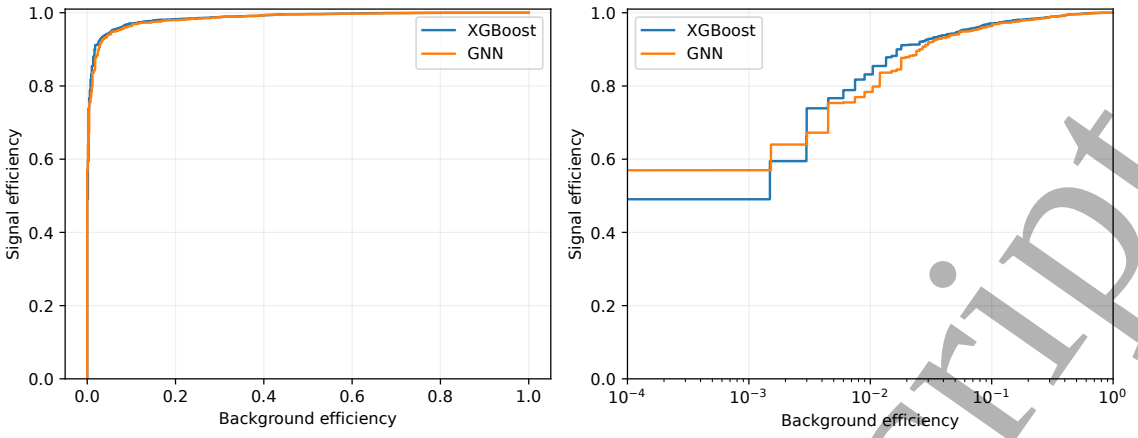


FIG. 6: Test ROC curves at $m_T = 2$ TeV (seed 42) with linear (left) and logarithmic (right) background-efficiency axes.

TABLE X: Background rejection $1-\epsilon_B$ of the principal backgrounds on the test partition at fixed signal efficiency, $m_T = 2$ TeV (seed 42).

	XGBoost		GNN	
Target ϵ_S (%)	ϵ_S (%)	$1-\epsilon_B$ (%)	ϵ_S (%)	$1-\epsilon_B$ (%)
50	50.016	99.8510	50.016	99.9971
70	70.010	99.7006	70.010	99.5502
80	80.006	99.2510	80.006	98.8048

D. Efficiencies and yields at the working points

Table XI gives the thresholds, efficiencies and yields at $m_T = 2$ TeV obtained from the test partitions of the principal samples, for the validation-selected working points and for a common threshold $s_0 = 0.99$. The common threshold corresponds to different operating points for the two classifiers: it retains 66.1% of the signal for XGBoost and 60.9% for the GNN. Comparisons are therefore made at equal signal efficiency. The column n_B gives the number of accepted test background events and σ_B^{MC} the corresponding statistical uncertainty; with only a few accepted events, these yields are dominated by simulation statistics, which is addressed by the estimate of Sec. IV D.

Accepted Manuscript
For Review Only

TABLE XI: Thresholds, efficiencies and yields at 3000 fb^{-1} on the test partitions of the principal samples, $m_T = 2$ TeV (seed 42). The 50%, 70% and 80% thresholds are selected on the validation partition.

Model	Point	s_0	ϵ_S (%)	ϵ_B (%)	S	B	n_B	σ_B^{MC}
XGBoost	50%	0.9957	49.69	0.1490	500.9	10225	2	10205
XGBoost	70%	0.9856	71.46	0.2994	720.2	20536	6	14432
XGBoost	80%	0.9608	80.68	0.7491	813.2	51387	17	22819
XGBoost	$s_0 = 0.99$	0.9900	66.09	0.2979	666.1	20434	4	14432
GNN	50%	0.9958	48.76	0.0014	491.5	98	1	98
GNN	70%	0.9774	70.75	0.4497	713.1	30851	11	17676
GNN	80%	0.9380	80.65	1.1952	812.9	81990	21	28864
GNN	$s_0 = 0.99$	0.9900	60.88	0.1519	613.6	10422	4	10206

E. Background composition and expected sensitivity

Figure 7 shows the background in the high-score region for the full background model, together with the principal-sample test events and the signal. Table XII lists the accepted events and yields of every process at the 50% working point for $m_T = 2$ TeV. In this region $Z(\nu\bar{\nu})$+jets is the largest background, contributing between 36% and 76% of the central estimate across masses, seeds and classifiers, followed by W+jets and semileptonic $t\bar{t}$; hadronic τ decays in W+jets contribute 2–19%, and together the processes beyond the four principal ones contribute 63–99%. These three processes are also the main backgrounds in the CMS analysis of this final state [24]. At the 50% working point each of these components rests on only a few accepted simulated events, so the central estimate fluctuates between seeds and classifiers; the conservative estimate B_{95} accounts for this statistical precision.

Accepted Manuscript
For Review Only

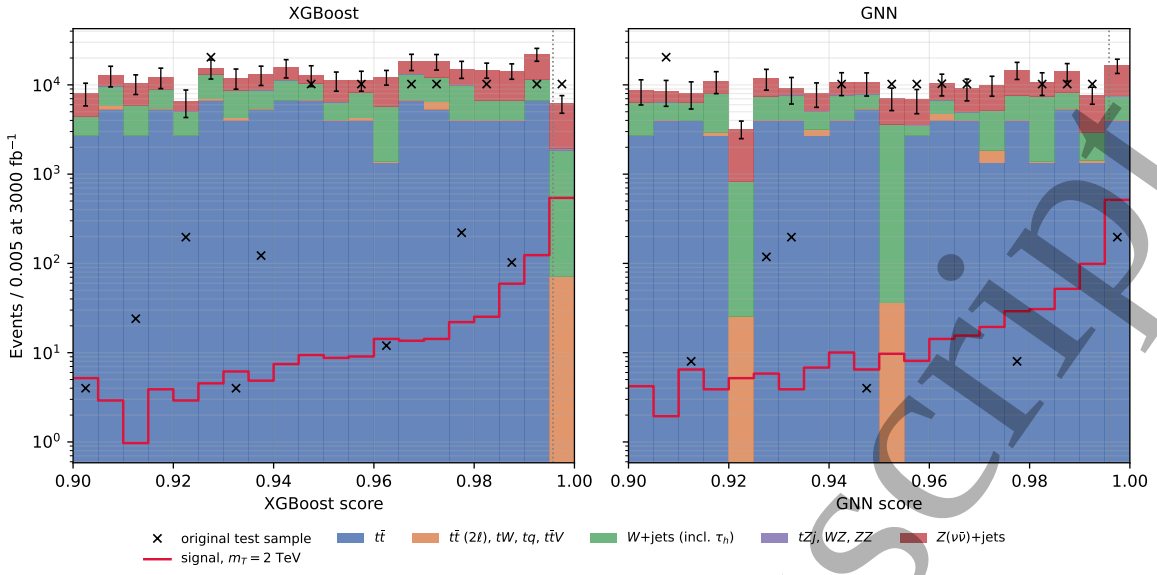


FIG. 7: Background score distribution above $s = 0.9$ at $m_T = 2$ TeV (seed 42) for the full background model, stacked by process group, at 3000 fb^{-1}. Error bars give the statistical uncertainty of the simulated samples. Crosses show the test events of the principal samples and the line the signal. The dotted line marks the 50% working point.

Accepted Manuscript
For Review Only

TABLE XII: Accepted simulated events n and central yields at 3000 fb^{-1} for each background process at the 50% working point, $m_T = 2$ TeV (seed 42). The first three processes use the independent samples normalized as in Eq. (25); for the light-jet components of W+jets, $Z(\nu\bar{\nu})$+jets and $W(\tau_h\nu)$+jets, n counts events above the working point and the yield uses the factorized efficiency. The $Z(\nu\bar{\nu})$+jets and $W(\tau_h\nu)$+jets entries use the emulated events of Sec. III C; for $W(\tau_h\nu)$+jets, n counts decay configurations, of which ten are drawn per event.

	XGBoost		GNN	
Process	n	Yield	n	Yield
$t\bar{t}$ (one lepton)	0	0.0	2	2665.4
$tZ(\nu\bar{\nu})j$	0	0.0	2	3.7
$WZ(\nu\bar{\nu})jj$	11	70.5	17	109.0
$ZZjj$	1	20.0	0	0.0
$t\bar{t}$ (dileptonic)	0	0.0	0	0.0
tW (one lepton)	0	0.0	0	0.0
tW (dileptonic)	0	0.0	0	0.0
tq (t-channel)	0	0.0	0	0.0
$t\bar{t}W$	2	6.8	5	17.1
$t\bar{t}Z(\nu\bar{\nu})$	36	44.6	36	44.6
W+light jets	0	1036.0	0	1288.7
$Wb\bar{b}j$	1	64.3	1	64.3
$Wc\bar{c}j$	0	0.0	0	0.0
$WZ(\nu\bar{\nu})b\bar{b}$	9	0.3	7	0.3
$WZ(\nu\bar{\nu})c\bar{c}$	8	0.3	12	0.5
$Z(\nu\bar{\nu})$+light jets	0	2726.4	1	6281.6
$Z(\nu\bar{\nu})b\bar{b}j$	7	413.2	17	1003.4
$Z(\nu\bar{\nu})c\bar{c}j$	5	256.8	9	462.3
$W(\tau_h\nu)$+light jets	0	385.9	0	1407.9
$W(\tau_h\nu)b\bar{b}j$	8	12.9	38	61.2
$W(\tau_h\nu)c\bar{c}j$	8	14.4	15	27.0
Total B		5052		13437
B_{95}		30332		49537
Signal S		501		491

Table XIII and Fig. 8 give the expected sensitivity at every mass. Using the conservative background B_{95}, the discovery significance at $m_T = 1.8$ TeV is 3.16 for the GNN and 4.23 for

Accepted Manuscript
For Review Only

XGBoost; three-standard-deviation evidence is reached up to $m_T \simeq 1.82$ TeV and 95% confidence-level exclusion up to $m_T \simeq 2.13$ TeV with both classifiers. The evidence and exclusion reaches of XGBoost and the GNN differ by 0.12 and 0.04 TeV; these differences, like that at 1.8 TeV, reflect the few accepted simulated background events rather than the classifiers. The central estimates are higher, with discovery up to $m_T \simeq 1.87$ TeV and exclusion up to $m_T \simeq 2.46$ TeV for both classifiers, but because the accepted $t\bar{t}$ count fluctuates between zero and a few events, the central values are not monotonic in m_T; we therefore base our conclusions on B_{95}. The exclusion reach is comparable to the current 13 TeV exclusions of 1.8–2.1 TeV for related singlet benchmarks [1, 2], whose coupling conventions differ from ours.

TABLE XIII: Expected sensitivity at the 50% working point and 3000 fb^{-1}, for the least sensitive of the three seeds. B is the central and B_{95} the conservative background estimate of Sec. IV D.

m_T (TeV)	Model	ϵ_S (%)	S	B	B_{95}	$Z_{\mathrm{dis}}(B)$	$Z_{\mathrm{dis}}(B_{95})$	$Z_{\mathrm{exc}}(B_{95})$
1.8	GNN	51.3	899	25549	80890	5.59	3.16	3.15
1.8	XGBoost	51.3	900	11668	44923	8.22	4.23	4.22
2.0	GNN	48.8	491	13437	49537	4.21	2.20	2.20
2.0	XGBoost	49.9	503	7703	35393	5.66	2.66	2.66
2.2	GNN	56.8	344	15345	59500	2.77	1.41	1.41
2.2	XGBoost	54.7	331	11903	48911	3.02	1.49	1.49
2.5	GNN	48.0	140	8119	39913	1.55	0.70	0.70
2.5	XGBoost	48.8	143	4943	30951	2.02	0.81	0.81
2.7	GNN	48.9	89	5932	37878	1.16	0.46	0.46
2.7	XGBoost	49.1	90	3725	29544	1.46	0.52	0.52
2.9	GNN	47.4	56	7340	39015	0.66	0.29	0.29
2.9	XGBoost	46.9	56	5157	38644	0.78	0.28	0.28
3.1	GNN	49.2	39	4908	33602	0.55	0.21	0.21
3.1	XGBoost	49.6	39	4662	29527	0.57	0.23	0.23

Accepted Manuscript
For Review Only

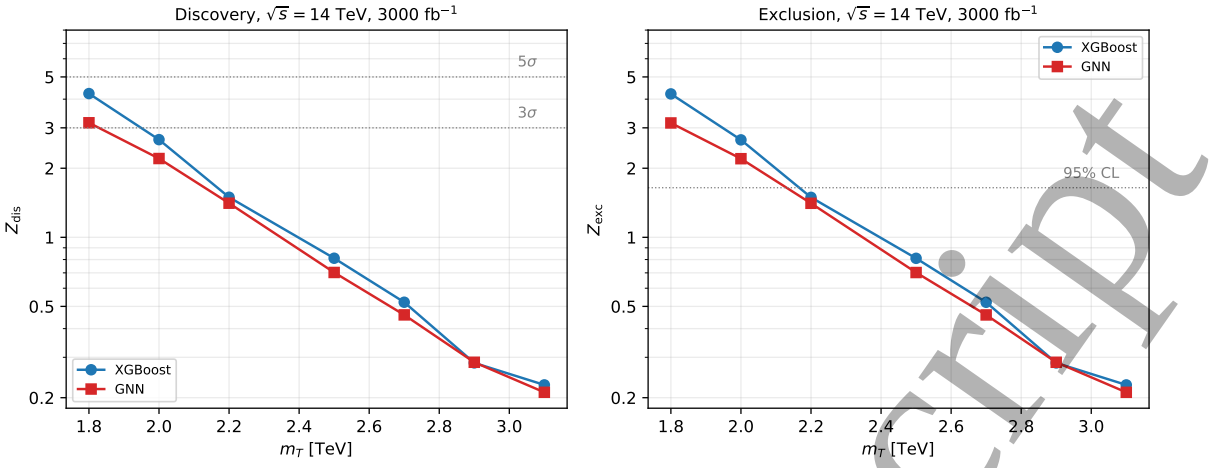


FIG. 8: Expected discovery (left) and exclusion (right) significance at 3000 fb^{-1} with the conservative background estimate B_{95}, at the 50% working point and for the least sensitive seed.

F. Dependence on the background normalization

The selected region has a small signal-to-background ratio, so the sensitivity depends on how well the background normalization is known. Table XIV shows Z_{dis} from Eq. (27) for relative uncertainties δ_B up to 10% with the conservative background, and Table XV up to 20% with the central background. The signal-to-background ratio is a few per cent with the central estimate and about one per cent with the conservative one. A background-normalization uncertainty of $\delta_B = 1\%$ therefore reduces the conservative discovery significance by more than half, and with the central estimate $\delta_B = 5\%$ brings it below two at every mass. The sensitivity in this channel is thus limited by the knowledge of the $Z(\nu\bar{\nu})$+jets, W+jets and $t\bar{t}$ normalizations rather than by the signal yield. At the HL-LHC these backgrounds would be normalized in control regions, for example $Z(\ell\ell)$+jets and $W(\ell\nu)$+jets events for $Z(\nu\bar{\nu})$+jets and W+jets, and events failing the classifier or top-reconstruction requirements for $t\bar{t}$, as in Ref. [24]; with 3000 fb^{-1} the statistical precision of such control regions can reach the per-cent level. These results quantify the background precision that such a measurement must achieve.

Accepted Manuscript
For Review Only

TABLE XIV: Discovery significance with the conservative background B_{95} for relative background-normalization uncertainties δ_B.

m_T (TeV)	Model	S/B_{95}	$Z_{\rm dis}(B_{95})$ for $\delta_B =$				
			0	1%	2%	5%	10%
1.8	GNN	0.011	3.16	1.05	0.55	0.22	0.11
1.8	XGBoost	0.020	4.23	1.80	0.97	0.40	0.20
2.0	GNN	0.010	2.20	0.90	0.48	0.20	0.10
2.0	XGBoost	0.014	2.66	1.25	0.68	0.28	0.14
2.2	GNN	0.006	1.41	0.53	0.28	0.11	0.06
2.2	XGBoost	0.007	1.49	0.62	0.33	0.13	0.07

TABLE XV: Discovery significance with the central background B for relative background-normalization uncertainties δ_B.

m_T (TeV)	Model	S/B	$Z_{\rm dis}(B)$ for $\delta_B =$			
			0	5%	10%	20%
1.8	GNN	0.035	5.59	0.69	0.35	0.17
1.8	XGBoost	0.077	8.22	1.48	0.75	0.38
2.0	GNN	0.037	4.21	0.71	0.36	0.18
2.0	XGBoost	0.065	5.66	1.25	0.63	0.32
2.2	GNN	0.022	2.77	0.44	0.22	0.11
2.2	XGBoost	0.028	3.02	0.54	0.27	0.14

G. Expected limits on the cross section and the coupling

The benchmark lies above the indirect bounds of Sec. II C. We therefore also express the sensitivity in a form that applies to other couplings. For each mass and classifier, the expected 95% confidence-level upper limit on the signal strength μ relative to the benchmark follows from $Z_{\rm exc}(\mu S, B_{95}) = 1.645$, using the conservative background and the least sensitive seed. At $g^* = 0.4$ the T is a broad resonance (Sec. II B), and the cross section of the full decay chain is smaller than the quadratic extrapolation from small couplings. We account for this with the ratio $c(m_T)$ of the decay-chain cross sections at $g^* = 0.4$ and at $g^* = 0.1$, the latter multiplied by $(0.4/0.1)^2$, computed with the same model implementation; it ranges from 0.68 to 0.92. The expected limit

Accepted Manuscript
For Review Only

for a narrow T is then

$$[\sigma(pp\to Tj)\,\mathcal{B}(T\to tZ)]_{95} = \mu_{95}\,c(m_T)\;[\sigma(pp\to Tj)\,\mathcal{B}(T\to tZ)]_{g^*=0.4}\,, \qquad g^*_{95} = 0.4\sqrt{\mu_{95}\,c(m_T)}, \tag{28}$$

where the benchmark cross section is the LO single-production cross section of Table I, multiplied by the K factor of the signal samples and by $\mathcal{B}(T\to tZ) = 1/6$ at $R_L = 0.5$; g^*_{95} follows because single production of an on-shell T scales exactly as $(g^*)^2$. The selection efficiency is taken from the benchmark samples. Table XVI and Fig. 9 give the results. For $R_L = 0.5$, the expected limit on $\sigma(pp\to Tj)\,\mathcal{B}(T\to tZ)$ lies between 41 and 64 fb for $m_T = 1.8$–3.1 TeV for the less sensitive classifier, corresponding to $g^*_{95} = 0.28$ at 1.8 TeV and 0.32 at 2 TeV. These values lie above the bounds $g^* \lesssim 0.13$ from CKM unitarity and $g^* < 0.22$ from $t\to uZ$ decays in the minimal model (Sec. II C). In that model, the HL-LHC sensitivity in this channel is therefore not sufficient to probe the parameter space allowed by indirect constraints at $R_L = 0.5$, and the limits on $\sigma\mathcal{B}$ are relevant for extended models in which these constraints are modified.

Using Eq. (9), the coupling limits in Table XVI imply $\Gamma_T/m_T \simeq 6\%$–8% at $m_T = 1.8$ TeV and 11%–13% at 2.0 TeV; these benchmarks are more compatible with the approximate narrow-width treatment and may serve as approximate sensitivity estimates. The corresponding ratios increase to approximately 24%–25% at 2.2 TeV and 53%–62% at 2.5 TeV, reaching approximately 2.5–2.7 at 3.1 TeV when the same formula is extrapolated. The higher-mass coupling limits should therefore be regarded as formal extrapolations beyond the narrow-width regime, rather than quantitatively reliable exclusions of the stated model.

Accepted Manuscript
For Review Only

TABLE XVI: Expected 95% confidence-level upper limits on $\sigma(pp \to Tj)\,\mathcal{B}(T \to tZ)$ and on g^* at $R_L = 0.5$ and 3000 fb^{-1}, for a narrow T, obtained with the conservative background and the least sensitive seed. $c(m_T)$ is the width correction of Eq. (28), and the third column gives the benchmark cross section at $g^* = 0.4$. The coupling limits at $m_T \geq 2.2$ TeV are formal extrapolations beyond the narrow-width regime and should not be interpreted as quantitatively reliable exclusions of the stated model.

			$\sigma\mathcal{B}$ limit (fb)		g^* limit	
m_T (TeV)	$c(m_T)$	$\sigma\mathcal{B}$, $g^* = 0.4$ (fb)	XGBoost	GNN	XGBoost	GNN
1.8	0.916	89.6	31.9	42.8	0.24	0.28
2.0	0.875	64.0	34.6	41.8	0.29	0.32
2.2	0.848	45.7	42.7	45.3	0.39	0.40
2.5	0.795	28.3	45.8	52.8	0.51	0.55
2.7	0.757	20.5	49.2	55.9	0.62	0.66
2.9	0.722	15.1	63.3	62.9	0.82	0.82
3.1	0.681	11.0	54.4	58.6	0.89	0.92

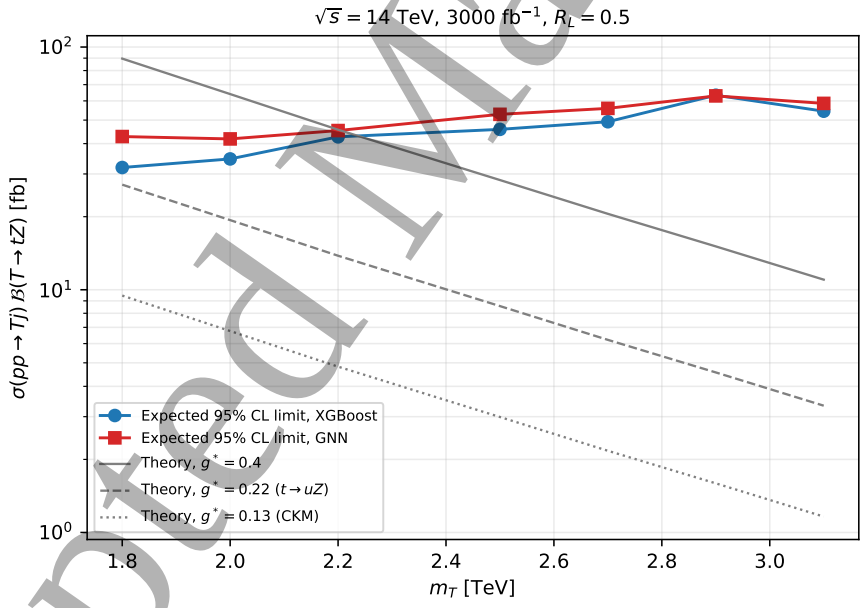


FIG. 9: Expected 95% confidence-level upper limits on $\sigma(pp \to Tj)\,\mathcal{B}(T \to tZ)$ at 3000 fb^{-1} for XGBoost and the GNN, compared with the LO prediction at $R_L = 0.5$ for $g^* = 0.4$ and for the largest couplings allowed by $t \to uZ$ ($g^* = 0.22$) and by CKM unitarity ($g^* = 0.13$).

For Review Only
Accepted Manuscript

VI. COMPARISON WITH EXISTING STUDIES

The cut-based analysis of Ref. [34] projects, at 14 TeV and 3000 fb^{-1} and for $R_L = 0.5$, two- and five-standard-deviation reaches of 2.52 and 2.22 TeV for couplings up to $g^* = 0.5$. Our results apply to $g^* = 0.4$, include $Z(\nu\bar{\nu})$+jets and lost-lepton backgrounds that are absent from that study, and use a conservative treatment of the background tail; a direct comparison would in addition require a common normalization of the coupling, which differs between the Lagrangian displayed in that reference and the model implementation used here. The CMS search in the same final state [24] excludes $\sigma(pp \to Tbq)\,\mathcal{B}(T \to tZ) > 15$ fb at $m_T = 1.8$ TeV with 137 fb^{-1} at 13 TeV, for a narrow T produced in association with a b quark. Our expected limits at 3000 fb^{-1}, 41–64 fb with the conservative background and 22–28 fb with the central background estimate, are not stronger. Besides the different production mode, this reflects the analysis strategy. CMS combines merged, partially merged and resolved top-quark candidates, classifies events by the presence of a forward jet, and fits the transverse-mass distribution with backgrounds constrained in control regions. The present study uses a resolved top-quark reconstruction, a single counting region and a conservative estimate of the simulated background tail. An HL-LHC analysis of this channel would benefit from combining the classifiers studied here with these techniques. Table XVII summarizes relevant published searches. They constrain model-dependent cross sections or κ_T benchmarks at 13 TeV with data-constrained backgrounds; a numerical conversion to g^* would require matching the coupling, width and branching-fraction conventions.

TABLE XVII: Selected published searches for singly produced vector-like T quarks. Quoted values are observed 95% confidence-level results.

Study	Data set	Relevant feature	Published result
CMS [24]	13 TeV, 137 fb^{-1}	Jets plus missing momentum, $T \to tZ$	Narrow-width $\sigma\mathcal{B}$ limits of 602–15 fb for $m_T = 0.6$–1.8 TeV; a 5%-width singlet benchmark is excluded below 0.98 TeV.
ATLAS [23]	13 TeV, 139 fb^{-1}	Combined Ht and Zt final states	Singlet couplings $\kappa_T > 0.53$ are excluded for all tested masses below 2.3 TeV; $\kappa_T = 0.35$ is excluded at 1.6 TeV.
ATLAS [1]	13 TeV, 139 fb^{-1}	Boosted top plus missing momentum; XGBoost top tagger	$m_T < 1.8$ TeV excluded for $\kappa_T = 0.5$ and $\mathcal{B}(T \to tZ) = 25\%$.
CMS [25]	13 TeV, 138 fb^{-1}	All-hadronic tH/tZ final state	Cross-section limits of 1260–68 fb over $m_T = 0.6$–1.2 TeV.
ATLAS [2]	13 TeV, 139 fb^{-1}	Combination of single-T Ht and Zt searches	Singlet benchmark with $\kappa_T = 0.5$ excluded below 2.1 TeV.

Accepted Manuscript
For Review Only

VII. CONCLUSION

We have studied single production of a singlet top partner in the hadronic tZ channel with $Z \to \nu\bar{\nu}$ at the 14 TeV HL-LHC. Signal and background are separated with XGBoost and a GNN that receive identical event information and are trained on common, process-stratified partitions at seven masses with three seeds each. Both classifiers reach test AUCs of 0.984–0.993 and differ by at most 0.0023; once the input information is matched, the tree-based and graph-based architectures perform equivalently for this signature.

The background model includes eighteen processes, and the high-score background tail is evaluated with simulated samples independent of the training, validated by closure tests. $Z(\nu\bar{\nu})$+jets, W+jets and $t\bar{t}$ form the background in the signal region, with $Z(\nu\bar{\nu})$+jets the largest. With the total background at its 95% upper bound, the expected discovery significance at 3000 fb^{-1} for $g^* = 0.4$ and $R_L = 0.5$ is 3.16–4.23 at $m_T = 1.8$ TeV, and evidence and 95% confidence-level exclusion are reached up to $m_T \simeq 1.82$ TeV and $\simeq 2.13$ TeV, comparable to the current 13 TeV limits for related singlet benchmarks. Because the signal-to-background ratio in the selected region is a few per cent, realizing this reach requires the background normalizations to be constrained in data control regions at the per-cent level; Tables XIV and XV quantify the required precision.

In the minimal model with a single vector-like quark, the benchmark exceeds the bounds from CKM unitarity, $t \to uZ$ decays and electroweak precision data. We therefore also provide expected limits on $\sigma(pp \to Tj)\,\mathcal{B}(T \to tZ)$, between 41 and 64 fb for $m_T = 1.8$–3.1 TeV and $R_L = 0.5$. They correspond to $g^* \simeq 0.28$ at 1.8 TeV, above the region allowed by CKM unitarity and $t \to uZ$ decays in the minimal model, and apply directly to extended models in which the indirect constraints are modified.

VIII. STATEMENTS AND DECLARATIONS

A. Acknowledgments

The authors thank Mr. Kamran Ahmad (AM Physics, National Centre for Physics, Pakistan) for valuable discussions and time.

Accepted Manuscript
For Review Only

B. Funding

This research did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

C. Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

D. Data and code availability

The Delphes card, the event-selection code, the event-level input features, the partition assignments, the trained models with their preprocessing constants, the classifier scores and the scripts used for the background estimate and the statistical analysis are available from the authors on reasonable request.

Appendix A: Correlations of the input variables

Figures 10 and 11 show the Pearson correlation coefficients of the 13 global inputs for the 2 TeV signal and for the principal backgrounds, and Figs. 12 and 13 those of all 45 non-constant inputs: the 13 global quantities, the five continuous features of each of the four nodes and the 12 edge distances. The constant role labels are omitted. The axes use $\ell_{p_T} \equiv \log(1 + p_T/\mathrm{GeV})$ and $\ell_m \equiv \log[1 + \max(m, 0)/\mathrm{GeV}]$. Physical event weights are used, after the selection of Sec. III B.

Accepted Manuscript
For Review Only

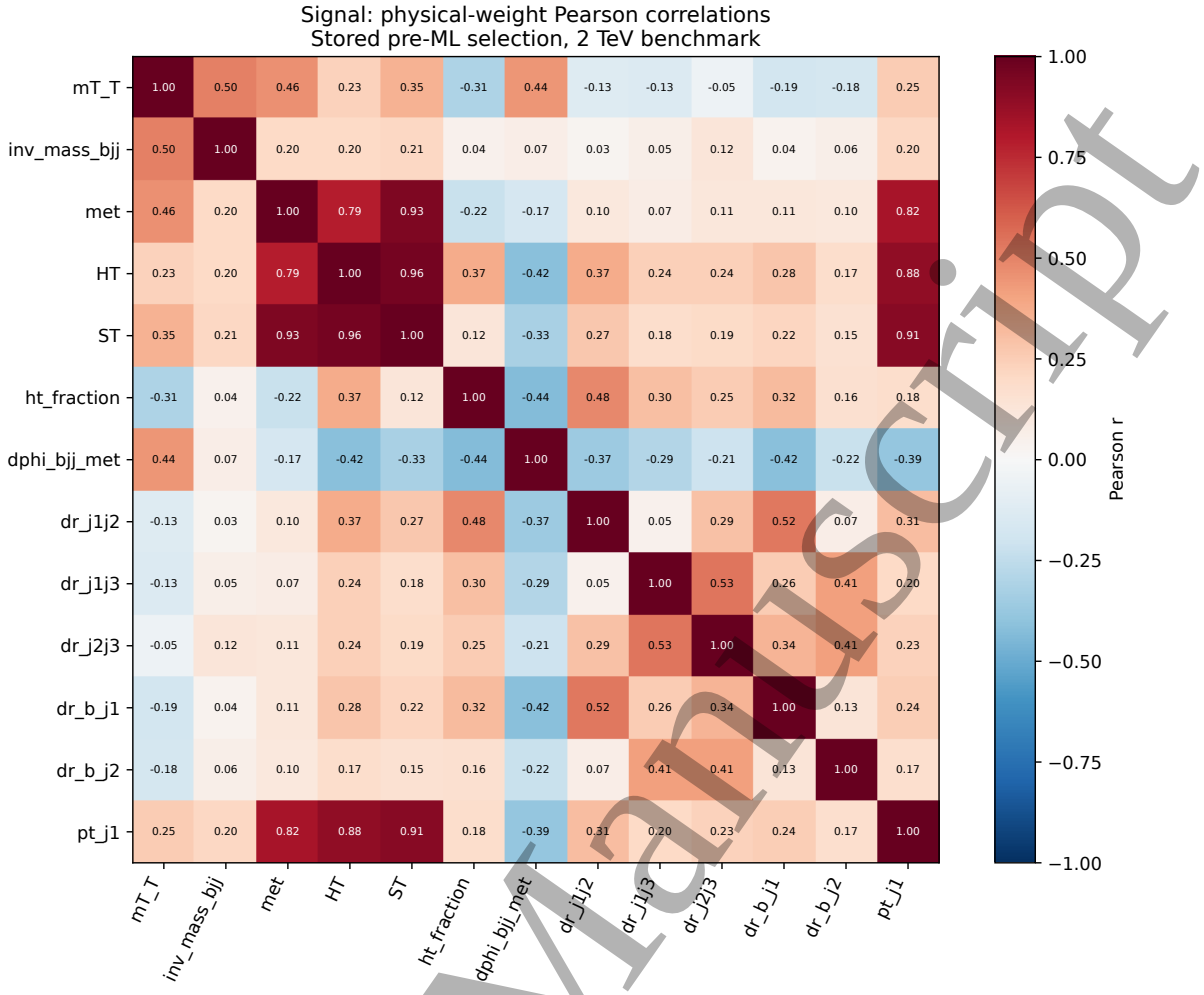


FIG. 10: Pearson correlation coefficients of the global inputs for the $m_T = 2$ TeV signal after the selection of Sec. III B.

Accepted Manuscript
For Review Only

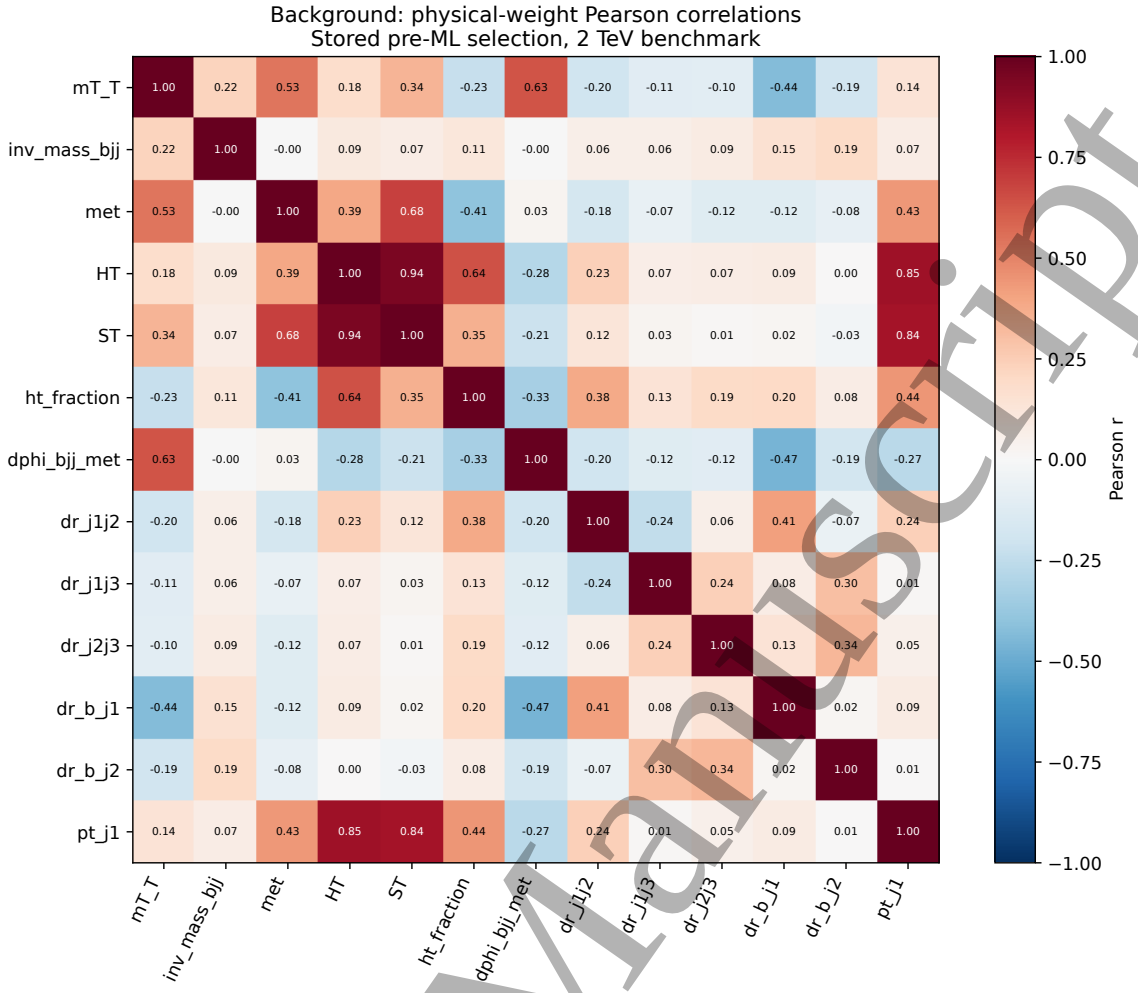


FIG. 11: Pearson correlation coefficients of the global inputs for the background after the selection of Sec. III B, using physical event weights.

Accepted Manuscript
For Review Only

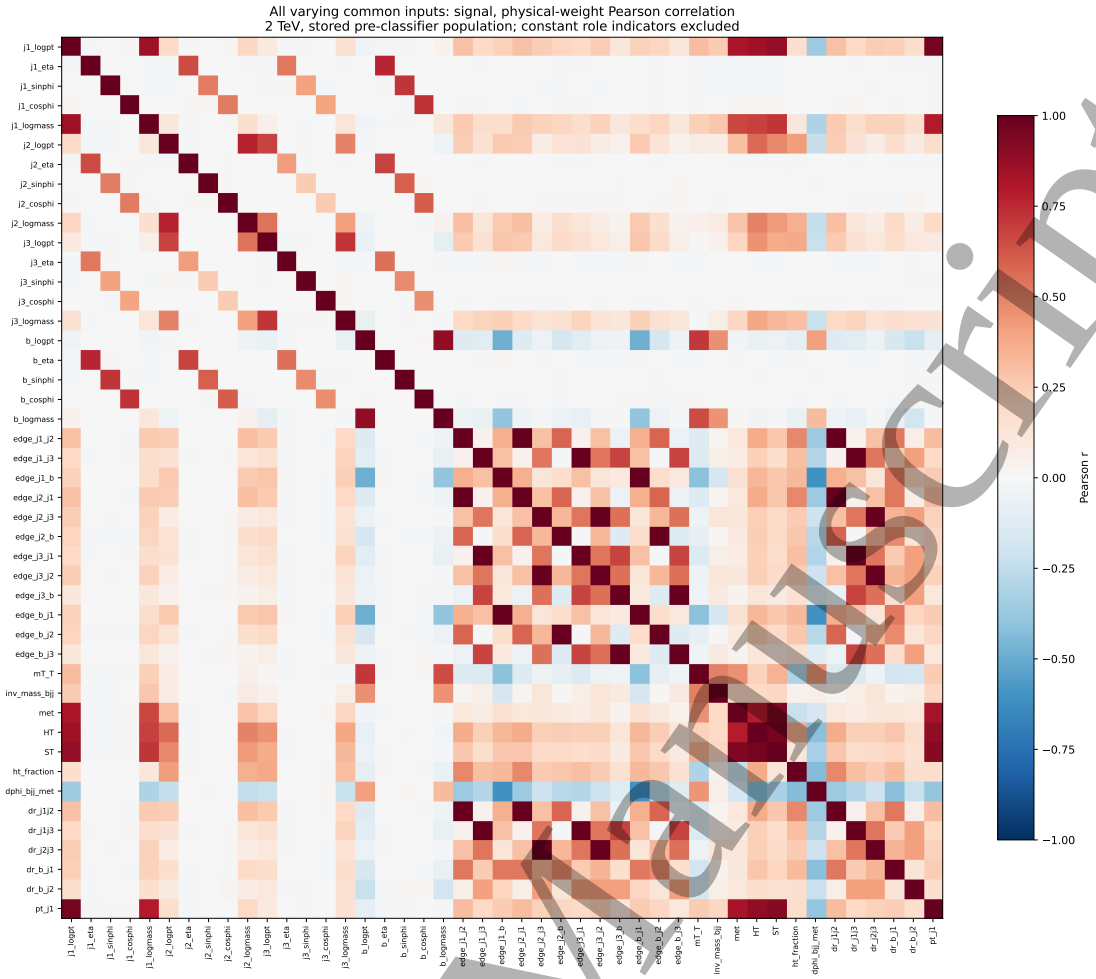


FIG. 12: Correlation coefficients of all non-constant inputs for the $m_T = 2$ TeV signal.

Accepted Manuscript Only

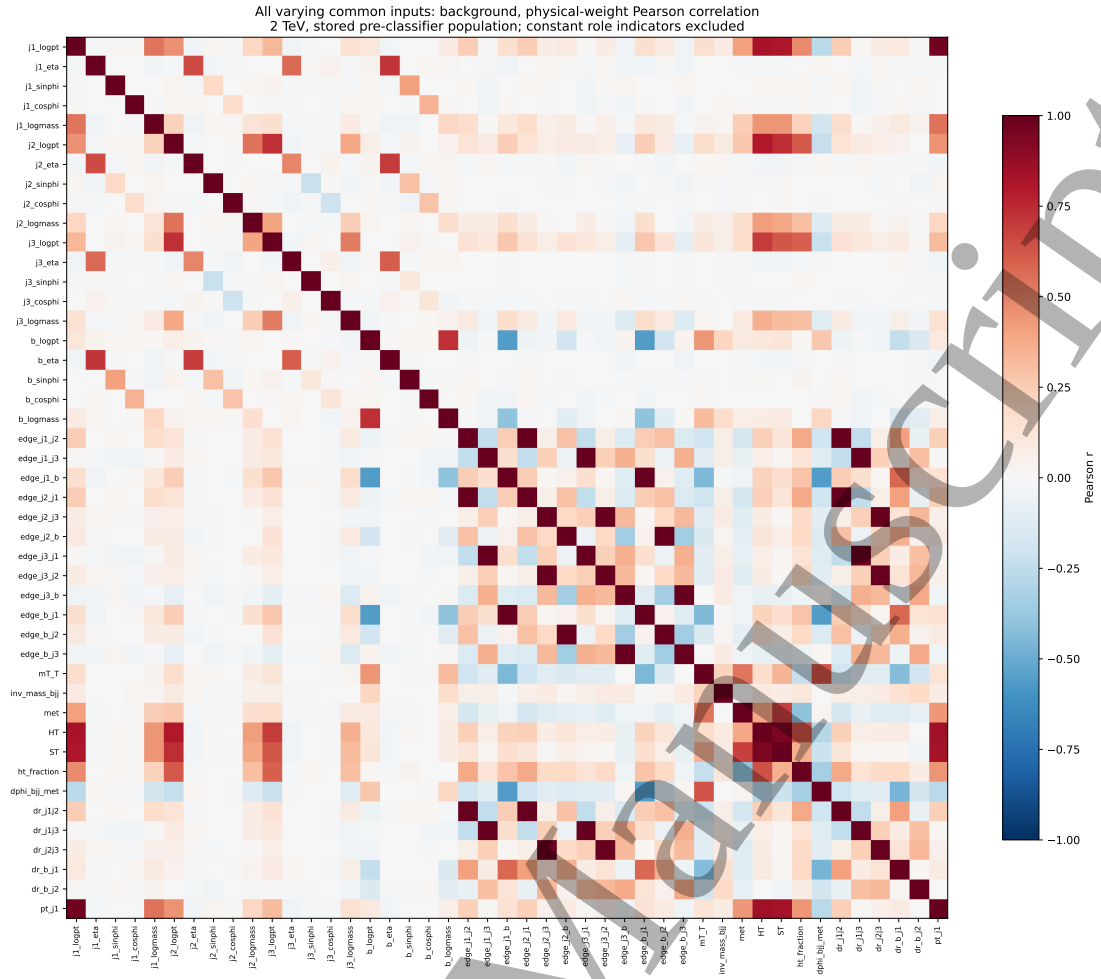


FIG. 13: Correlation coefficients of all non-constant inputs for the principal backgrounds.

Accepted Manuscript Only

Appendix B: Closure tests of the background estimate

For each of the seven masses, three seeds and two classifiers, score thresholds are placed at signal efficiencies of 90%, 95% and 98%, where the test partitions of the principal samples still contain enough background events. The fraction of selected events above the threshold is compared between the test partition of the principal sample, ϵ_{orig}, and the independent sample, ϵ_{ind}. Table XVIII gives the ratio averaged over the 126 configurations, its range, and the largest deviation in units of the combined statistical uncertainty. The ratios are compatible with unity for $t\bar{t}$, which dominates the background; the average ratios of 0.84 for tZj and 1.22 for $WZjj$ are covered by the 25% margin of the conservative estimate. For the W+light-jet factorization, the number of W+light-jet events passing the 90% and 95% thresholds, summed over all configurations, is 1.63 and 1.28 times the factorized prediction, within the factor of two applied in the conservative estimate.

TABLE XVIII: Ratio of classifier efficiencies between independent and principal samples of the same process. The range and the largest pull refer to the individual mass, seed and classifier configurations.

Process	Target ϵ_S (%)	$\langle\epsilon_{\mathrm{ind}}/\epsilon_{\mathrm{orig}}\rangle$	range	max |pull|
$t\bar{t}$ (one lepton)	90	0.86	0.56–1.15	1.4
$t\bar{t}$ (one lepton)	95	1.01	0.77–1.25	1.4
$t\bar{t}$ (one lepton)	98	1.03	0.91–1.22	1.7
$tZ(\nu\bar{\nu})j$	90	0.76	0.55–0.98	1.7
$tZ(\nu\bar{\nu})j$	95	0.84	0.63–1.03	1.9
$tZ(\nu\bar{\nu})j$	98	0.86	0.67–1.01	2.3
$WZ(\nu\bar{\nu})jj$	90	1.29	0.84–1.87	1.8
$WZ(\nu\bar{\nu})jj$	95	1.26	1.05–1.48	1.8
$WZ(\nu\bar{\nu})jj$	98	1.18	1.07–1.45	1.8

Appendix C: Graph neural network architecture

Each event is represented by a graph with four nodes, $v \in \{j_1, j_2, j_3, b\}$, and twelve directed edges connecting all ordered pairs of distinct nodes. The node feature vector $\mathbf{x}_v \in \mathbb{R}^9$ contains the five standardized kinematic quantities of Table V and a one-hot label of the node role; the edge feature e_{vu} is the standardized ΔR between nodes v and u; and $\mathbf{u} \in \mathbb{R}^{13}$ contains the standardized global quantities. With $\mathrm{ReLU}(z) = \max(z, 0)$, square brackets denoting concatenation, and D

Accepted Manuscript
For Review Only

denoting dropout with probability 0.2 during training, the network is

$$h_v^{(0)} = \mathrm{ReLU}\big(W_0\,\mathbf{x}_v + b_0\big), \tag{C1}$$

$$m_{v\leftarrow u}^{(l)} = \mathrm{ReLU}\big(W_m^{(l)}\,[h_v^{(l)};\, h_u^{(l)};\, e_{vu}] + b_m^{(l)}\big), \tag{C2}$$

$$h_v^{(l+1)} = D\Big(\mathrm{ReLU}\big(W_h^{(l)}\,[h_v^{(l)};\, \tfrac{1}{3}\textstyle\sum_{u\neq v} m_{v\leftarrow u}^{(l)}] + b_h^{(l)}\big)\Big), \qquad l = 0, 1, \tag{C3}$$

$$h_G = \frac{1}{4}\sum_v h_v^{(2)}, \tag{C4}$$

$$s = \sigma\Big(w_2^{\mathrm{T}}\, D\big(\mathrm{ReLU}(W_1\,[h_G;\, \mathbf{u}] + b_1)\big) + b_2\Big), \tag{C5}$$

with $h_v \in \mathbb{R}^{32}$, $W_0 \in \mathbb{R}^{32\times 9}$, $W_m^{(l)} \in \mathbb{R}^{32\times 65}$, $W_h^{(l)} \in \mathbb{R}^{32\times 64}$, $W_1 \in \mathbb{R}^{64\times 45}$ and $w_2 \in \mathbb{R}^{64}$. The message passing is equivariant, and the readout invariant, under a common relabelling of the nodes; the role labels retain the information about the p_T ordering of the jets, which is also available to XGBoost. The network has 11 713 trainable parameters and is trained by minimizing Eq. (20) with the Adam optimizer (learning rate 10^{-3}, weight decay 10^{-4}).

[1] G. Aad et al, JHEP, **05**: 263 (2024)

[2] G. Aad et al, Phys. Rev. D, **111**: 012012 (2025)

[3] J. A. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer et al, Phys. Rev. D, **88**: 094010 (2013)

[4] J. Alwall, R. Frederix, S. Frixione et al, JHEP, **07**: 079 (2014)

[5] G. Apollinari, I. Bejar Alonso, O. Bruning et al, Tech. Rep. CERN-2015-005, (2015)

[6] N. Arkani-Hamed, A. G. Cohen, E. Katz et al, JHEP, **07**: 034 (2002)

[7] A. Atre, G. Azuelos, M. Carena et al, JHEP, **08**: 080 (2011)

[8] P. Baldi, P. Sadowski, and D. Whiteson, Nat. Commun., **5**: 4308 (2014)

[9] R. D. Ball et al, Eur. Phys. J. C, **82**: 428 (2022)

[10] G. Aad et al. (ATLAS Collaboration), Phys. Rev. D, **108**: 032019 (2023)

[11] S. Dawson and E. Furlan, Phys. Rev. D, **86**: 015021 (2012)

[12] C.-Y. Chen, S. Dawson, and I. M. Lewis, Phys. Rev. D, **90**: 035016 (2014)

[13] R. D. Ball, V. Bertone, S. Carrazza et al, Nucl. Phys. B, **867**: 244-289 (2013)

[14] A. Banerjee, E. Bergeaas Kuutmann, V. Ellajosyula et al, SciPost Phys. Core, **7** (4): 079 (2024)

[15] M. Buchkremer, G. Cacciapaglia, A. Deandrea et al, Nucl. Phys. B, **876**: 376-417 (2013)

[16] A. Buckley, J. Ferrando, S. Lloyd et al, Eur. Phys. J. C, **75**: 132 (2015)

[17] G. Cacciapaglia, A. Deandrea, D. Harada et al, JHEP, **11**: 159 (2010)

[18] G. Cacciapaglia, A. Deandrea, L. Panizzi et al, JHEP, **03**: 070 (2012)

[19] G. Cacciapaglia, A. Carvalho, A. Deandrea et al, Phys. Lett. B, **793**: 206-211 (2019)

[20] M. Cacciari, G. P. Salam, and G. Soyez, JHEP, **04**: 063 (2008)

Accepted Manuscript For Review Only

[21] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C, **72**: 1896 (2012)
[22] T. Chen and C. Guestrin, Proc. KDD, (2016)
[23] G. Aad, B. Abbott, K. Abeling et al., (ATLAS Collaboration), JHEP, **08**: 153 (2023).
[24] A. Tumasyan, W. Adam, J. W. Andrejkovic et al., (CMS Collaboration), JHEP, **05**: 093 (2022)
[25] A. Hayrapetyan, A. Tumasyan W. Adam et al., (CMS Collaboration), Phys. Rev. D, **110**: 072012 (2024)
[26] G. Cowan, K. Cranmer, E. Gross et al, Eur. Phys. J. C, **71**: 1554 (2011)
[27] K. Cranmer, Proc. 2011 Eur. School High Energy Phys., 267-308 (2015)
[28] M. Czakon, P. Fiedler, and A. Mitov, Phys. Rev. Lett., **110**: 252004 (2013)
[29] J. de Favereau, C. Delaere, P. Demin et al, JHEP, **02**: 057 (2014)
[30] A. Deandrea, T. Flacke, B. Fuks et al, JHEP, **08**: 107 (2021)
[31] B. Fuks and H.-S. Shao, Eur. Phys. J. C, **77** (2): 135 (2017)
[32] L. Han, S. Wang, L. Shang et al, Chin. Phys. C, **47**: 043108 (2023)
[33] P. T. Komiske, E. M. Metodiev, and J. Thaler, JHEP, **01**: 121 (2019)
[34] H. Li, J. Chao, and G. Zhang, Nucl. Phys. B, **994**: 116310 (2023)
[35] F. J. Massey, Jr., J. Am. Stat. Assoc., **46** (253): 68-78 (1951)
[36] S. Navas et al, Phys. Rev. D, **110**: 030001 (2024)
[37] G. Panico and A. Wulzer, Lect. Notes Phys., **913** (2016)
[38] H. Qu and L. Gouskos, Phys. Rev. D, **101**: 056019 (2020)
[39] H. Qu, C. Li, and S. Qian, Proc. Mach. Learn. Res., **162**: 18281-18292 (2022)
[40] A. Radovic, M. Williams, D. Rousseau et al, Nature, **560**: 41-48 (2018)
[41] L. Randall and R. Sundrum, Phys. Rev. Lett., **83**: 3370-3373 (1999)
[42] A. L. Read, J. Phys. G, **28**: 2693-2704 (2002)
[43] J. Shlomi, P. Battaglia, and J.-R. Vlimant, Mach. Learn. Sci. Technol., **2**: 021001 (2021)
[44] A. M. Sirunyan et al, JHEP, **05**: 029 (2017)
[45] A. M. Sirunyan et al, Phys. Lett. B, **781**: 574-600 (2018)
[46] T. Sjostrand, S. Ask, J. R. Christiansen et al, Comput. Phys. Commun., **191**: 159-177 (2015)
[47] M. Zaheer, S. Kottur, S. Ravanbakhsh et al, Adv. Neural Inf. Process. Syst., **30**: 3391-3401 (2017)

Accepted Manuscript
For Review Only