

Machine learning-based b-jet tagging in pp collisions at $\sqrt{s} = 13$ TeV

Hadi Hassan^{1,*}, Neelkamal Mallick^{2,†} and D. J. Kim^{2,3,‡}

¹*University of Tsukuba, Tennodai 1-1-1, Tsukuba, Ibaraki 305-0006, Japan*

²*Department of Physics, University of Jyväskylä, P.O. Box 35, FI-40014, Jyväskylä, Finland*

³*Helsinki Institute of Physics, P.O. Box 64, FI-00014, Helsinki, Finland*



(Received 29 May 2025; accepted 29 June 2026; published 4 August 2026)

Studying heavy-flavor jets in pp collisions is important since it can test perturbative QCD calculations and be used as a reference for heavy-ion collisions. Jets in this analysis are reconstructed from charged particles using the anti- k_T algorithm with a resolution parameter $R = 0.4$ at midrapidity $|\eta| < 0.5$. Beauty jets are tagged using a machine learning model that uses a convolutional neural network trained on information extracted from the jet, tracks, and secondary vertices. Results show that this model is superior in b-jet-tagging compared to other traditional tagging methods.

DOI: [10.1103/rw87-lyw8](https://doi.org/10.1103/rw87-lyw8)

I. INTRODUCTION

Heavy-flavor jets are important for the understanding of quantum chromodynamics (QCD) [1]. They can provide tests for the perturbative QCD calculations, and have a large mass that acts as a hard scale which suppresses soft-gluon radiations at small angles [2,3]. They are produced very early in the collision, so they carry information on the hard collision and information about the high-energy-density medium produced in heavy-ion collisions [4,5]. Identification of jets formed from hadronization of b-quarks, so-called “b-jets,” is important for measurements such as cross section [6–8], fragmentation, jet substructure [9,10], and for understanding the quark-gluon plasma in heavy-ion collisions [11,12].

Various tagging algorithms have been developed to achieve the heavy-flavor jet identification. The problem with traditional b-jet tagging algorithms, such as impact parameter (IP) or secondary vertex (SV) methods [7,13–15] that exploit the long lifetime of the b-hadron, is that they rely on cuts imposed on a single variable. Machine learning (ML) methods can combine several parameters from different reconstruction levels, such as track-level observables (e.g., track impact parameters), jet-level properties (e.g., jet mass), or secondary vertex features (e.g., decay length), and extract complex multidimensional correlations and patterns

among the event characteristics. Previous studies from ATLAS and CMS experiments show that the ML models are superior compared to the traditional methods in terms of performance [16–19]. In this study, to simulate b-jet events in pp collisions, the PYTHIA8 Monte Carlo (MC) event generator is used [20,21], mainly since it is widely used and carefully validated by the LHC experiments [22,23]. Experimental measurements, such as those performed by ALICE, are affected by detector effects such as finite tracking, pointing resolutions, acceptance, and efficiency. Although the analysis methods can be optimized for experimental conditions such as the ALICE environment, we wish to keep the discussion at the methodological level. Hence, we do not perform a complete MC simulation including the ALICE detector [24] using Geant4 [25], but instead smear the transverse momentum (p_T) and distance to the closest approach (DCA) distributions to mimic realistic experimental conditions.

This paper is organized as follows: Sec. II describes the event generation, spectra smearing, jet reconstruction, feature extraction, and machine learning methodology. Section III presents the performance of the ML algorithm for the tagging of beauty jets and compares it with traditional methods such as the impact parameter and the secondary vertex methods. Finally, Sec. IV summarizes the findings of this analysis.

II. METHODOLOGY

A. Event generation and jet reconstruction

Proton-proton (pp) collisions at c.m. energy $\sqrt{s} = 13$ TeV are generated using the PYTHIA8 [20,21] MC event generator with the Monash 2013 tune [26] and all the hard QCD processes. For a more realistic description of the ALICE detector [24], events with primary vertices within

*Contact author: hadi.hassan@cern.ch

†Contact author: neelkamal.mallick@cern.ch

‡Contact author: dong.jo.kim@jyu.fi

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). 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.

$|z| < 10$ cm in the beam direction, and final state charged particles (after parton shower, hadronization and decays) with transverse momentum $p_T > 0.15$ GeV/ c , within the ALICE tracker acceptance $|\eta| < 0.9$, are used in this analysis.

Then a transverse momentum smearing is applied on these tracks based on the parameters from the ALICE inner tracking system 2 (ITS2) momentum resolution [27] using the following equation:

$$\sigma(p_T)/p_T = \sqrt{a^2 + b^2/p_T^2 + c^2 \cdot p_T^2}. \quad (1)$$

Here, $a = 0.005$, $b = 8 \times 10^{-4}$ GeV/ c , and $c = 0.001$ (GeV/ c) $^{-1}$ are constants that represents the momentum resolution of the ALICE ITS2 detector. In addition to the p_T smearing, a smearing of the DCA of the particles with respect to the primary vertex has been applied according to the DCA resolution of the ITS2 detector [27,28],

$$\sigma(\text{DCA}_{xy/z})(\mu\text{m}) = \sqrt{a^2 + b^2/p_T^2}. \quad (2)$$

Here, $a = 7$ and 6 μm , and $b = 23$ and 25 $\mu\text{m} \cdot \text{GeV}/c$ are the values of the constants for the DCA resolution in the transverse plane and z direction, respectively.

The selected particles with p_T and DCA smearing are passed to the jet clustering algorithm. Jets are reconstructed using the anti- k_T jet clustering algorithm [29] from the FastJet package [30] with a resolution parameter $R = 0.4$. This ensures that most of the initial partonic energy is contained inside the clustered jet. The energies of the tracks are combined using the energy-recombination scheme (E -scheme), which includes the mass of the particles in the jet four-momentum calculation with the best strategy. It should be noted here that the real masses of the particles are used for the jet reconstruction. Further, a reconstructed jet is classified as a beauty jet (b-jet) if it contains a beauty quark (from the PYTHIA particle record), a charm jet (c-jet) if it contains a charm quark, and a light-flavor jet (lf-jet) if neither of the heavy quarks are present. Jets with beauty quarks coming from the gluon splitting are classified as b-jets.

B. Feature extraction

For the machine learning algorithm, several features are extracted from the reconstructed jet, tracks, and secondary vertices. For jets, the extracted features are the jet transverse momentum (p_T), pseudorapidity (η), azimuthal angle (ϕ), number of track constituents, number of reconstructed secondary vertices (explained later), and the jet mass. For the constituents of the jet, the extracted features are particle p_T , particle η , the dot product of the momenta of the track and jet ($\vec{p}_{\text{jet}} \cdot \vec{p}_{\text{part}}$), the dot product of the track and the jet divided by the jet momentum ($\vec{p}_{\text{jet}} \cdot \vec{p}_{\text{part}}/|\vec{p}_{\text{jet}}|$), and the

distance between the track and the jet axis in the $\eta - \phi$ plane calculated as $\Delta R = \sqrt{(\eta_{\text{jet}} - \eta_{\text{track}})^2 - (\phi_{\text{jet}} - \phi_{\text{track}})^2}$. In addition, other features such as the signed DCA in the transverse (xy) plane, the signed DCA in z direction, and the distance between the track and the closest reconstructed secondary vertex are calculated with the same distance equation mentioned before. The DCA sign is calculated as the sign of the dot product of the jet axis and the DCA vector ($\text{sign}(\vec{\text{DCA}}_{\text{track}} \cdot \vec{p}_{\text{jet}})$). Tracks coming from secondary vertices have positive DCA signs due to their long flight distance, while primary tracks coming from the primary vertex could have positive or negative DCA signs. Some of these features extracted from the jet and tracks are very discriminative between the different jet flavors, especially the signed DCA of the track to the primary vertex, as seen in Fig. 1.

Prior to particle selection, a maximum DCA (in both xy and z) of 1 cm is applied to reject tracks originating from strange particle decays. Since the DCA is the most discriminative observable, the tracks in each jet are sorted

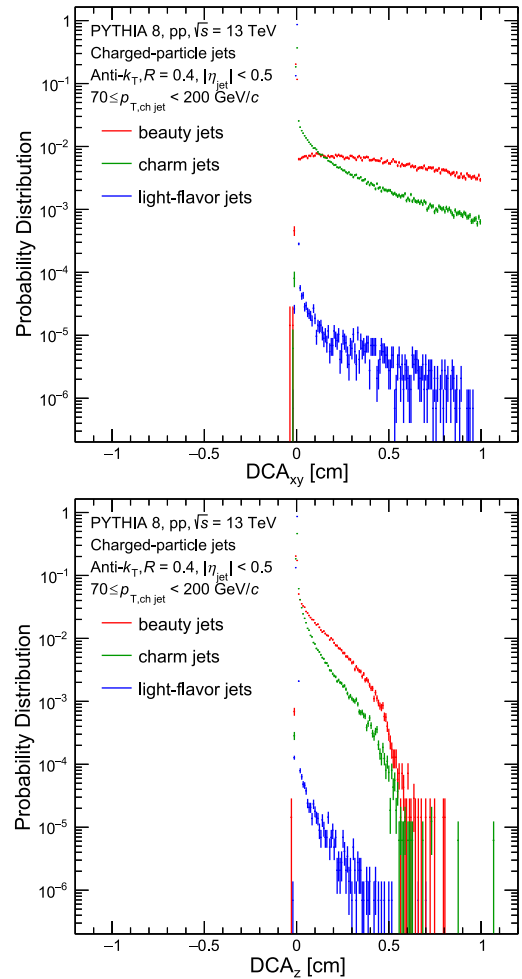


FIG. 1. Transverse (top) and longitudinal (bottom) DCA distributions for different jet flavors.

in descending order of their signed DCAs, and the top ten tracks are selected as input to the ML model.

Secondary vertices within jets are reconstructed from their constituent tracks. First, a kinematic cut is applied by selecting tracks with $p_T > 0.3$ GeV/ c , which helps to exclude tracks significantly affected by detector resolution. Triplets of the selected tracks are then combined to identify potential secondary vertices under the assumption that they may originate from a common decay point. For each triplet, a vertex fitting algorithm is employed to determine the optimal vertex position by minimizing the distance between the tracks and the vertex—quantified by the χ^2 fit. The resulting candidate vertices are subjected to additional quality criteria to ensure that only well-reconstructed and physically meaningful secondary vertices are retained. It should be noted here that no smearing on the secondary vertices is applied since the DCA and momentum of the tracks, which are used for the secondary vertex reconstruction, are already smeared.

The features from the reconstructed secondary vertices inside the jet are derived from the tracks associated to vertex. These features include: SV p_T , η , the distance between the secondary vertex and the jet calculated using the ΔR equation mentioned above, the SV mass, the energy fraction carried by the secondary vertex in the jet, the distance of closest approach in the bending plane of the SV trajectory with respect to the primary vertex, the cosine pointing angle (CPA) which is the cosine of the angle between the SV momentum and the vector pointing from the primary vertex (PV) to SV, χ^2 of the SV reconstruction fit, and the dispersion $D = \sqrt{(d_1^2 + d_2^2 + d_3^2)/N}$ which is the square root of squared summed DCAs of the daughter tracks with respect to the secondary vertex normalized by

the number of tracks. The decay length which is the flight distance between the primary and secondary vertices in the transverse plane and in the z direction is also included in the list of input features.

Figure 2 (right) shows the decay length distribution for the different jet flavors, and Fig. 2 (left) shows the cosine pointing angle distribution for the reconstructed secondary vertex. As can be seen in the figure, the decay length is very discriminative for the different jet flavors. This behavior is expected because beauty and charm hadrons have short lifetimes, so they decay away from the primary vertex, producing a measurable decay length. Light-flavor hadrons are produced from the primary vertex so they do not form a secondary vertex, resulting in a very small decay length. Since the decay length is the most discriminative feature for secondary vertices, all SVs inside the jets are sorted in descending order based on their decay length. The top ten SVs are then selected as input for the ML model.

The number of input tracks and SVs per jet is limited to 10. This choice is motivated by a dedicated study in which the number of input objects was varied between 3 and 20. The discrimination performance improves with increasing multiplicity up to approximately eight objects, after which it saturates, with negligible gain beyond this point. Therefore, ten inputs are used as a compromise between performance and computational cost.

Tables I–III in the Appendix summarize all the features from jet, tracks, and secondary vertices and their definitions that are used for training the ML model.

C. Machine learning model

Convolutional neural networks (CNNs) [31] are neural networks used specifically for image classification by

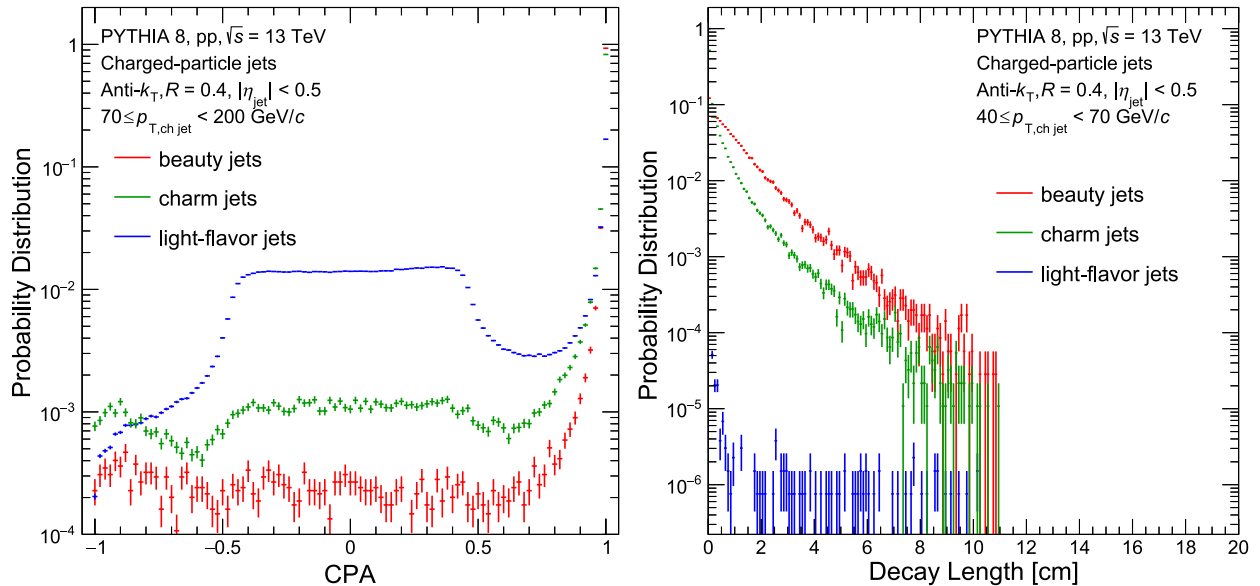


FIG. 2. Left: cosine pointing angle distribution of the secondary vertices for the different jet flavors. Right: secondary vertex decay length.

identifying patterns and features in data. They consist of three types of layers: a convolution layer that applies kernel filters that extract patterns, a pooling layer that reduces the number of inputs, and fully connected layers that will make the final decision and connect the CNN to the output layer. In this analysis, a machine learning algorithm comprising three branches has been used. Two CNN branches correspond to the ten tracks and ten SVs connected at the end to a recurrent neural network (RNN) layer with long short-term memory (LSTM) units [32], and a branch that corresponds to the global jet parameters.

The convolution branches consist of four convolution layers made of 128, 64, 64, and 32 filters with kernel sizes 5, 3, 3, and 2, respectively. Two max-pooling layers of size 2 are applied after the second and third convolutions. A stride of 1, same padding, and ReLU activation function are used. A dropout of 0.1 is applied after each layer. At the end of the convolution process, the output is reshaped and then connected to an LSTM layer with 64 RNN units, batch normalization, and a dropout of 0.1 at the end. The third branch, which corresponds to the jet features, will be directly concatenated with the two other branches. Then

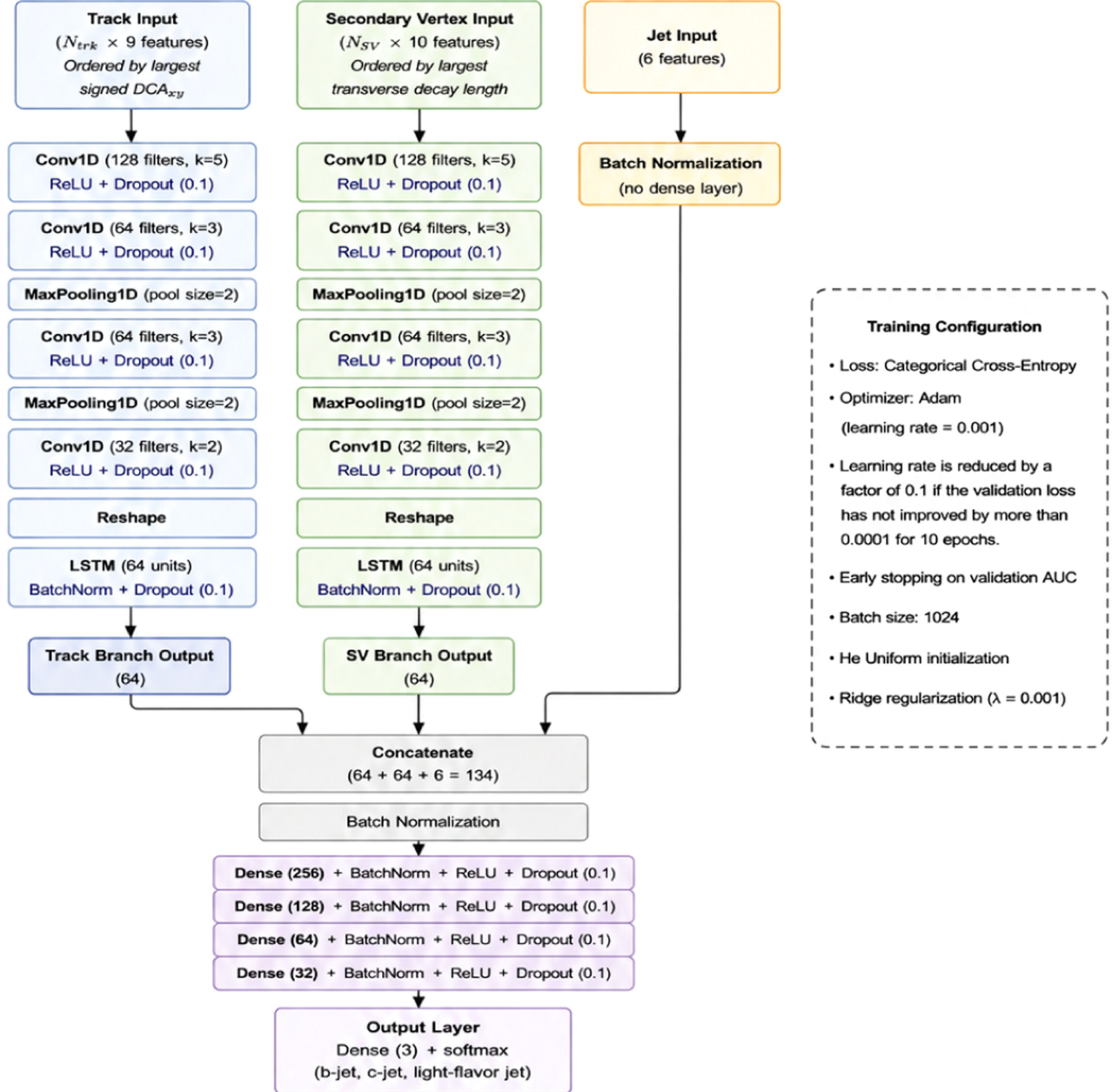


FIG. 3. Machine learning model architecture combining local track and secondary vertex features processed through 1D CNNs with global jet-level features.

the flattened layer will be connected to four fully connected layers with 256, 128, 64, and 32 neurons. A dropout of 0.1 is again applied after each layer. The output is then connected to a softmax layer with three classes corresponding to the different jet flavors. Batch normalization is applied after each layer. The loss function used is the categorical cross entropy with Adam optimizer, He Uniform kernel initializer, ridge regularization with regularization parameter of 0.001, and a learning rate of 0.001. The learning rate is reduced by a factor of 0.1 if the validation loss has not improved by more than 0.0001 for ten epochs. The model architecture is summarized in a diagram as shown in Fig. 3. The model is a multiclass classifier that provides three output probabilities corresponding to beauty, charm, and light-flavor jets. These probabilities are combined into a single one-dimensional b-jet discriminator defined as D_b [33], as follows:

$$D_b = \log\left(\frac{P_b}{f_c P_c + (1 - f_c) P_{lf}}\right), \quad (3)$$

where P_b , P_c , and P_{lf} are the probabilities given by the model for a jet to be classified as a b-, c-, or lf-jet, respectively, and f_c is the parameter corresponding to the charm-jet fraction in the background. In this paper, the parameter f_c was chosen to be 0.08, as expected from the inclusive jet composition in PYTHIA8.

In order to train this model, 160×10^6 pp collisions at $\sqrt{s} = 13$ TeV are generated using PYTHIA8. From all of these events, around 2×10^6 b-jets were obtained and used as input for the TensorFlow [34] ML model. The sample is constructed from three classes corresponding to b-, c-, and light-flavor jets with equal fractions. This corresponds to a prior probability of 1/3 for each class. The sample is randomly split into 70% used for training, 15% for validation, and 15% for testing. Each subsample contains approximately equal fractions of b-, c-, and light-flavor jets.

After compiling the neural network with all its CNN and dense layers, it is trained on the training sample using a batch size of 1000. A maximum of 300 epochs is allowed, with early stopping applied based on the validation accuracy. Training is stopped if the validation accuracy has not improved by more than 10^{-4} for ten consecutive epochs. In practice, convergence is reached much earlier, and the optimal model is selected after 53 epochs, thereby preventing overfitting. The performance is then evaluated on the testing sample. Finally, the trained network is converted into ONNX [35] format, which is loaded and applied to the reconstructed jets during event generation.

III. RESULTS

Figure 4 (top) shows the receiver operating characteristic (ROC) curve of the ML algorithm under study. The ROC curve is a graphical representation for the model

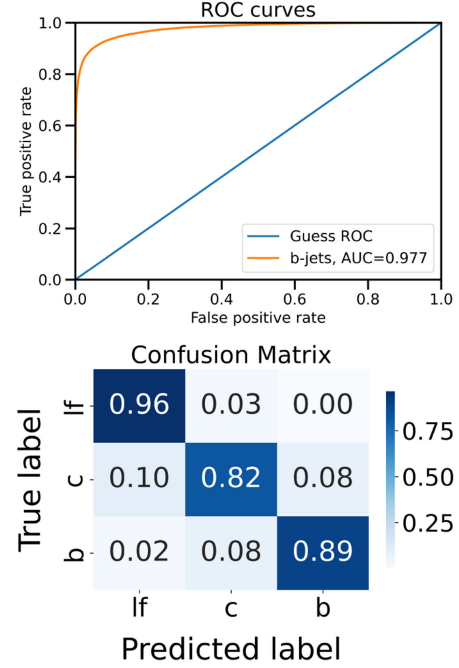


FIG. 4. ROC curves and confusion matrix.

performance. It shows the true-positive rate (tagged b-jets) versus the false positive rate (mistagged c-/lf-jets) at various thresholds. As can be seen in the figure, there is a small decrease in the true-positive rate compared to a sharp decrease in false-positive rate. This figure also shows the area under the curve (AUC), which is the area under the ROC curve; if the AUC is 1.0, it means perfect separation between signal and background, while 0.5 means random guessing. The AUC shown in Fig. 4 (top) is around 0.977, which shows that the model is able to efficiently separate the different jet flavors. The whole training and evaluation procedures were repeated five times using different random seeds for splitting the data between train/validation/test in order to test the model stability. By comparing the ROC curves obtained from these different trainings, the results were found to be consistent with an average AUC of 0.9754 ± 0.0012 .

Figure 4 (bottom) shows the confusion matrix, which is a table that summarizes the performance of the model by comparing the true labels of jets to the predicted ones from the ML model. It displays the fraction of correctly classified b-, c-, and lf-jets along the diagonal entries, and the fraction of misidentified jets elsewhere. This misidentification can happen either due to the presence of secondaries in the lf-jets or due to the presence of a displaced secondary vertex in the case of charm jets. As shown in the figure, the fraction of misidentification is very small compared to the fraction of correctly identified jets, which again reflects the high purity of the tagged jets sample. The b-jets are tagged with an accuracy of almost 90%, which is a decent score based on the complexity of the problem.

Figure 5 (top) shows the D_b distribution of the different jet flavors as defined in Eq. (3). As one can observe, the

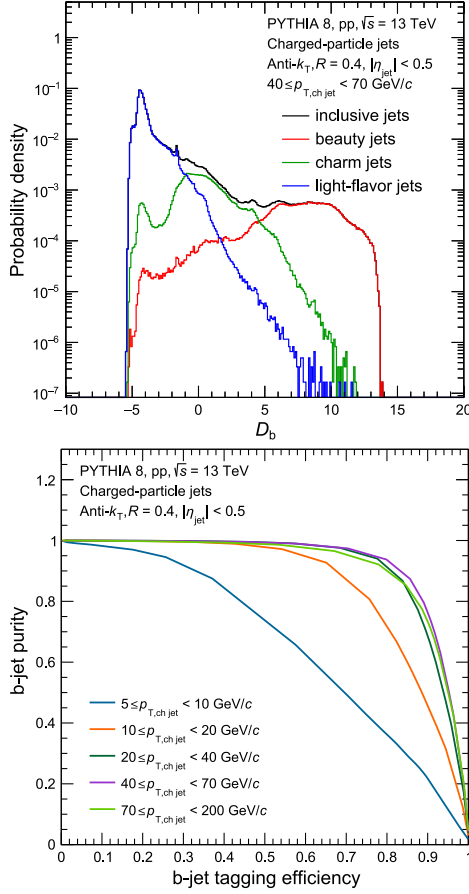


FIG. 5. Top: D_b distribution for the different jet flavors. Bottom: b-jet purity as a function of the tagging efficiency from the ML model.

distribution decreases sharply for the light-flavor jets. For charm jets, the distribution increases slightly to a peak at $D_b = 0$ before experiencing a sharp decrease. For beauty jets, however, the distribution gradually increases and reaches a peak around $D_b = 6$ before decreasing at higher values. This demonstrates the model's separation ability between b-jets and c-/lf-jets.

Figure 5 (bottom) shows the purity versus efficiency of the b-tagging model for different jet p_T bins. This figure is crucial since it illustrates the performance of the model in different kinematic regions. It is made by changing the D_b threshold from -5 to 13 and selecting all the jets that pass this threshold. As shown in the figure, low score threshold values will give high efficiencies but low purity, since this allows for more background jets to be tagged. Conversely, the high score threshold values lead to lower efficiencies but, in return, give a very high purity since it suppresses a large fraction of the background. High efficiency might be needed in measurements that require a maximum signal yield, while purity is crucial for analyses where the background contamination must be minimized, such as when measuring b-jets in Pb-Pb collisions at LHC energies.

When comparing the model performance in the different p_T regions, one can see that the performance of the ML model is reliable (it reaches 90% purity at 80% efficiency) for high p_T jets, which is due to the larger DCA of the tracks and the larger decay length of the secondary vertices compared to charm and light-flavor jets. The performance at the lowest p_T bins is not as good as the high p_T ones, which is due to the low particle multiplicity inside the jet compared to the high p_T jets. This is also due to the shorter decay lengths which might make the jet resemble charm or

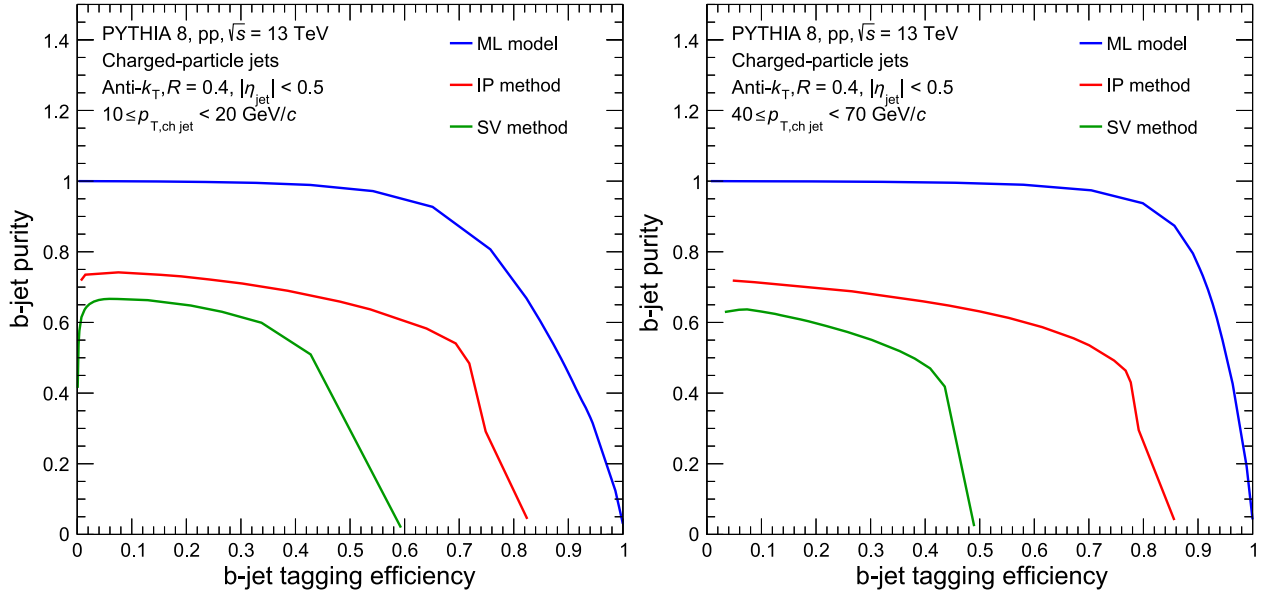


FIG. 6. ML model performance (purity versus efficiency) compared with the performance of traditional methods (IP and SV methods) in two different p_T ranges.

light-flavor jet, imposing higher complexity for the jet flavor tagging at low- p_T .

Finally, to validate the use of machine learning-based models for jet flavor tagging, Fig. 6 shows the performance of the ML model compared to traditional cut-based methods, such as the secondary vertex and track counting methods. These approaches rely on the track with the largest DCA in the jet and are implemented following the procedure described in Ref. [7]. As can be seen, the ML model used in this work achieves a much higher purity for the same efficiencies. For example, at 50% efficiency, the ML model achieves 30% higher purity compared to the IP method. This means that the ML model is superior since other methods rely on single observables such as the SV decay length or DCA, while the ML model, working on a multivariate analysis framework, fuses all the possible observables from the jet, tracks, and secondary vertices. This demonstrates the advantage of using machine learning models in jet flavor tagging.

IV. CONCLUSION

Events in pp collisions at $\sqrt{s} = 13$ TeV are generated using PYTHIA8 (Monash) model, where charged particles with $p_T > 0.15$ GeV/ c and $|\eta| < 0.9$ are selected. The p_T and DCA of the selected particles are smeared according to the ALICE ITS2 detector transverse momentum and DCA resolutions, respectively. Then, jets are reconstructed from these charged particles using the anti- k_T algorithm with

$R = 0.4$. Then a machine learning model that uses CNNs is trained and tested on these jets. The result shows a better performance in terms of purity and tagging efficiency of the ML model compared to the traditional methods.

The current study is performed only on events generated using PYTHIA8 simulation. This model is planned to be used on the ALICE experimental data in the future, which requires a good agreement between data and MC for the input variables and discriminator output (D_b). Possible discrepancies between data and MC could affect the performance of the model and may require further investigation in the future.

ACKNOWLEDGMENTS

We would like to thank the Jyväskylä ALICE group for fruitful discussions. We acknowledge CSC–IT Center for Science in Espoo, Finland, for allocating computational resources. N.M. and D.J.K. are supported by the Academy of Finland through the Center of Excellence in Quark Matter (Grant No. 346328), and H. H. through Grant No. 346327. H. H. also acknowledges support from JSPS KAKENHI Grant No. JP24H00003.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX: THE FEATURES USED FOR TRAINING THE ML MODEL

TABLE I. List of jet-level features used as input for the ML model.

Feature	Description
Jet p_T	Transverse momentum of the jet
Jet η	Pseudorapidity of the jet
Jet φ	Azimuthal angle of the jet
Number of track constituents	Number of tracks associated with the jet
Number of secondary vertices	Number of reconstructed SVs inside the jet
Jet mass	Invariant mass of the jet

TABLE II. List of jet constituent (track-level) features used as input for the ML model.

Feature	Description
Particle p_T	Transverse momentum of the particle
Particle η	Pseudorapidity of the particle
$\vec{p}_{\text{jet}} \cdot \vec{p}_{\text{part}}$	Dot product of jet and particle momenta
$\vec{p}_{\text{jet}} \cdot \vec{p}_{\text{part}} / \vec{p}_{\text{jet}} $	Normalized dot product by jet momentum
ΔR (jet, track)	Distance between jet axis and track in η - φ plane
Signed DCA (xy)	Signed distance of closest approach in transverse plane
Signed DCA (z)	Signed distance of closest approach in z direction
Distance to closest SV	Distance between track and closest SV

TABLE III. List of SV features used as input for the ML model.

Feature	Description
SV p_T	Transverse momentum of the SV
SV η	Pseudorapidity of the SV
ΔR (SV, jet)	Distance between SV and jet using ΔR
SV mass	Invariant mass of the SV
Energy fraction	Fraction of jet energy carried by the SV
DCA (bending plane)	Closest approach of SV trajectory to the primary vertex
CPA	Cosine of angle between SV momentum and the pointing vector $PV \rightarrow SV$
χ^2 of SV fit	χ^2 from the SV reconstruction fit
Dispersion D	$\sqrt{(d_1^2 + d_2^2 + d_3^2)/N}$; summed squared DCAs of daughter tracks divided by N
Decay length (transverse, z)	Flight distance between PV and SV in transverse plane and z

- [1] H. D. Politzer, Asymptotic freedom: An approach to strong interactions, *Phys. Rep.* **14**, 129 (1974).
- [2] Y. L. Dokshitzer, V. A. Khoze, and S. I. Troian, On specific QCD properties of heavy quark fragmentation (“dead cone”), *J. Phys. G* **17**, 1602 (1991).
- [3] S. Acharya *et al.* (ALICE Collaboration), Direct observation of the dead-cone effect in quantum chromodynamics, *Nature (London)* **605**, 440 (2022); **607**, E22 (2022).
- [4] B.-W. Zhang, E. Wang, and X.-N. Wang, Heavy quark energy loss in nuclear medium, *Phys. Rev. Lett.* **93**, 072301 (2004).
- [5] Y. L. Dokshitzer and D. E. Kharzeev, Heavy quark colorimetry of QCD matter, *Phys. Lett. B* **519**, 199 (2001).
- [6] S. Chatrchyan *et al.* (CMS Collaboration), Inclusive b -jet production in pp collisions at $\sqrt{s} = 7$ TeV, *J. High Energy Phys.* **04** (2012) 084,
- [7] S. Acharya *et al.* (ALICE Collaboration), Measurement of inclusive charged-particle b -jet production in pp and p -Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, *J. High Energy Phys.* **01** (2022) 178,
- [8] A. Tumasyan *et al.* (CMS Collaboration), Measurement of the production cross section for $Z + b$ jets in proton-proton collisions at $\sqrt{s} = 13$ TeV, *Phys. Rev. D* **105**, 092014 (2022).
- [9] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of b jet shapes in proton-proton collisions at $\sqrt{s} = 5.02$ TeV, *J. High Energy Phys.* **05** (2021) 054.
- [10] A. Tumasyan *et al.* (CMS Collaboration), Search for medium effects using jets from bottom quarks in PbPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, *Phys. Lett. B* **844**, 137849 (2023).
- [11] G. Aad *et al.* (ATLAS Collaboration), Measurement of the nuclear modification factor of b -jets in 5.02 TeV Pb + Pb collisions with the ATLAS detector, *Eur. Phys. J. C* **83**, 438 (2023).
- [12] S. Chatrchyan *et al.* (CMS Collaboration), Evidence of b -jet quenching in PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, *Phys. Rev. Lett.* **113**, 132301 (2014); **115**, 029903(E) (2015).
- [13] CMS Collaboration, Identification of b -quark jets with the CMS experiment, *J. Instrum.* **8**, P04013 (2013).
- [14] ATLAS Collaboration, Secondary vertex finding for jet flavour identification with the ATLAS detector, Technical Report, CERN, Geneva, 2017, <https://cds.cern.ch/record/2270366>.
- [15] ATLAS Collaboration, Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run, Technical Report, CERN, Geneva, 2017, <https://cds.cern.ch/record/2273281>.
- [16] G. Aad *et al.* (ATLAS Collaboration), Fast b -tagging at the high-level trigger of the ATLAS experiment in LHC Run 3, *J. Instrum.* **18**, P11006 (2023).
- [17] G. Aad *et al.* (ATLAS Collaboration), ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset, *Eur. Phys. J. C* **83**, 681 (2023).
- [18] A. Cagnotta, F. Carnevali, and A. De Iorio, Machine learning applications for jet tagging in the CMS experiment, *Appl. Sci.* **12**, 10574 (2022).
- [19] E. Bols, J. Kieseler, M. Verzetti, M. Stoye, and A. Stakia, Jet flavour classification using DeepJet, *J. Instrum.* **15**, P12012 (2020).
- [20] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA 8.2, *Comput. Phys. Commun.* **191**, 159 (2015).
- [21] C. Bierlich *et al.*, A comprehensive guide to the physics and usage of PYTHIA 8.3, *SciPost Phys. Codebases* **2022**, 8 (2022).
- [22] S. Acharya *et al.* (ALICE Collaboration), Pseudorapidity distributions of charged particles as a function of mid- and forward rapidity multiplicities in pp collisions at $\sqrt{s} = 5.02$, 7 and 13 TeV, *Eur. Phys. J. C* **81**, 630 (2021).
- [23] A. M. Sirunyan *et al.* (CMS Collaboration), Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements, *Eur. Phys. J. C* **80**, 4 (2020).
- [24] K. Aamodt *et al.* (ALICE Collaboration), The ALICE experiment at the CERN LHC, *J. Instrum.* **3**, S08002 (2008).

- [25] S. Agostinelli *et al.*, Geant4—A simulation toolkit, *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).
- [26] P. Skands, S. Carrazza, and J. Rojo, Tuning PYTHIA 8.1: The Monash 2013 tune, *Eur. Phys. J. C* **74**, 3024 (2014).
- [27] B. Abelev *et al.* (ALICE Collaboration), Technical design report for the upgrade of the ALICE Inner Tracking System, Technical Report, 2014, <https://cds.cern.ch/record/1625842>.
- [28] A. S. Triolo (On behalf of the ALICE Collaboration), Calibration and performance of the upgraded Alice Inner Tracking System, *J. Instrum.* **20**, C01030 (2025).
- [29] M. Cacciari, G. P. Salam, and G. Soyez, The anti- k_t jet clustering algorithm, *J. High Energy Phys.* **04** (2008) 063.
- [30] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, *Eur. Phys. J. C* **72**, 1896 (2012).
- [31] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE* **86**, 2278 (1998).
- [32] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Comput.* **9**, 1735 (1997).
- [33] ATLAS Collaboration, Graph neural network jet flavour tagging with the ATLAS detector, Technical Report, CERN, Geneva, 2022, <https://cds.cern.ch/record/2811135>. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2022-027>.
- [34] TensorFlow Community, TensorFlow, <https://www.tensorflow.org/> (2025).
- [35] ONNX Community, ONNX: Open Neural Network Exchange, <https://onnx.ai> (2025).