



Reducing systematic bias in machine learning applications to J/ψ signal extraction in high-energy nuclear physics

Yan Wang¹ · Rong-Rong Ma² · Kai-Feng Shen¹ · Ze-Bo Tang¹ · Wang-Mei Zha¹

Received: 29 September 2025 / Revised: 5 December 2025 / Accepted: 4 January 2026 / Published online: 20 May 2026

© The Author(s), under exclusive licence to China Science Publishing & Media Ltd. (Science Press), Shanghai Institute of Applied Physics, the Chinese Academy of Sciences, Chinese Nuclear Society 2026

Abstract

Machine learning techniques are increasingly used in high-energy nuclear physics because they can exploit multivariate correlations more efficiently than conventional cut-based analyses. A central challenge is the construction of training samples that faithfully reproduce the detector response observed in data. Signal samples are usually derived from detector simulations; therefore, mismatches between simulation and data can degrade classifier performance and introduce systematic biases. This work presents two practical correction procedures, namely cumulative distribution function (CDF) mapping and a shift-and-scale transformation, to align simulated signal features with those measured in data. Their performance is demonstrated with J/ψ yield measurements in $\sqrt{s_{NN}} = 200$ GeV Ru+Ru and Zr+Zr collisions recorded by STAR. A set of self-consistency tests shows that these procedures substantially suppress the systematic bias associated with data-simulation discrepancies in machine-learning-based signal extraction.

Keywords Machine learning · J/ψ reconstruction · Cumulative distribution function mapping · Shift-and-scale

1 Introduction

Machine learning has become an indispensable tool across a broad range of scientific and technological disciplines, including image recognition, natural-language processing, recommendation systems, and speech analysis. Its main advantage is the ability to identify complex structures in high-dimensional datasets and to provide powerful tools for classification, regression, and clustering [1–6]. In nuclear physics, machine learning has been used to extract or predict key observables such as charge radii, nuclear masses, and binding energies [7–11]. In relativistic heavy-ion collisions,

it is especially valuable for classification tasks, for example separating rare physics signals from large combinatorial backgrounds [12–17]. Representative algorithms include boosted decision trees (BDTs) [18, 19], deep neural networks [20–22], and ensemble methods [23], all of which can exploit multidimensional correlations more effectively than conventional straight-cut analyses. Independent of the specific algorithm, a reliable classifier requires representative signal and background samples. Whereas background samples can often be obtained directly from data, signal samples are usually produced with Monte Carlo simulations. Because detector simulations do not perfectly reproduce the data, training on mismatched signal samples can degrade classifier performance and bias the extracted physics observables, as illustrated in Appendix A.

To address this challenge, we propose two correction methods: cumulative distribution function (CDF) mapping and shift-and-scale. CDF mapping systematically transforms a simulated feature distribution into the corresponding distribution measured in data without requiring an explicit parametric description of the shape, which makes it flexible and broadly applicable. The shift-and-scale procedure is more restrictive but simpler to implement, and it is well suited to cases in which the simulated and measured distributions

This work was supported by the National Key R&D Program of China (No. 2024YFA1610803) and the National Natural Science Foundation of China (Nos. 12361141827 and 12175223).

✉ Ze-Bo Tang
zbtang@ustc.edu.cn

✉ Wang-Mei Zha
first@ustc.edu.cn

¹ Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China

² Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA

Table 1 Electron preselection cuts and the optimal BDT cut criterion from machine learning

Track p_T	Detectors used	Electron preselection cuts	BDT criterion
$p_T \leq 1.0$ GeV/c	TPC, TOF	$ 1/\beta - 1 < 0.035$; for $p > 0.8$ GeV/c: $-1 < n\sigma_e < 2.25$; for $p \leq 0.8$ GeV/c: $(3 \times p - 3.4) < n\sigma_e < 2.25$;	BDT > 0.72
$p_T > 1.0$ GeV/c	TPC, TOF and BEMC no matching	$ 1/\beta - 1 < 0.035$; $-1 < n\sigma_e < 2.25$	BDT > 0.72
	TPC, BEMC and TOF no matching	$-1.25 < n\sigma_e < 2.25$; $0.3 < E_0/p < 1.8$	
	TPC, TOF and BEMC	$ 1/\beta - 1 < 0.035$; $-1.75 < n\sigma_e < 2.25$; $0.3 < E_0/p < 1.8$	

already have similar shapes and differ mainly in location and width.

J/ψ is a classic probe of the medium created in relativistic heavy-ion collisions and is widely used to investigate the properties of the Quark–Gluon Plasma [24, 25]. Accurate yield measurements are therefore of continuing interest [26–34]. In this work, a BDT implemented with the eXtreme Gradient Boosting (XGBoost) framework [35] is used to separate the J/ψ signal from background in Ru+Ru and Zr+Zr collisions at $\sqrt{s_{NN}} = 200$ GeV recorded by STAR at RHIC. This case study demonstrates both the implementation of the proposed correction procedures and their effectiveness in improving the robustness of machine-learning-based signal extraction. Section 2 describes the J/ψ reconstruction, Sects. 3 and 4 present the construction of the training samples and the model training procedure, Sect. 5 discusses the self-consistency tests and results, and Sect. 6 summarizes the main conclusions.

2 J/ψ Reconstruction

In 200 GeV Ru+Ru and Zr+Zr collisions, J/ψ mesons are reconstructed through the dielectron decay channel [27, 34, 36], with electron¹ identification based on the Time-of-Flight (TOF) [37], the Time Projection Chamber (TPC) [38, 39], and the Barrel Electromagnetic Calorimeter (BEMC) [40] detectors. Loose preselection requirements on the particle-identification (PID) observables are first imposed to suppress the dominant background while retaining sufficient signal statistics for the subsequent machine learning analysis.

The TOF detector measures the arrival time of a charged particle. Together with the collision start time and the momentum measured in the TPC, one obtains $1/\beta \equiv c/v$, where c is the speed of light and v is the particle velocity. Figure 1 shows the distribution of $1/\beta$ as a function of momentum p for charged particles. Because particles with different masses have different velocities at the same

momentum, distinct PID bands appear in this representation. A preselection cut of $|1/\beta - 1| < 0.035$ is applied, shown as dashed lines in Fig. 1 and listed in Table 1.

The TPC measures the specific ionization energy loss, dE/dx , which provides additional separation between electrons and hadrons. PID based on dE/dx is commonly expressed through the normalized deviation of the measured value from the expectation for a given particle species, denoted by $n\sigma_e$, $n\sigma_\pi$, and so forth. For electrons, the $n\sigma_e$ distribution is expected to be approximately Gaussian with mean 0 and width 1 after proper calibration. Figure 2 shows the $n\sigma_e$ distribution as a function of p after imposing a tighter TOF requirement, $|1/\beta - 1| < 0.025$, to improve the electron purity for visualization. The dashed curves indicate the momentum-dependent preselection criteria used before machine learning training.

At high momentum, the BEMC provides additional PID through the ratio of the deposited electromagnetic energy to the track momentum. In practice, electron identification commonly uses the largest tower energy within the BEMC cluster, E_0 , rather than the total cluster energy E_{tot} , because E_0 has better resolution and is less sensitive to the high detector occupancy in heavy-ion collisions. Figure 3 compares the E_0/p distributions for protons and electrons in the

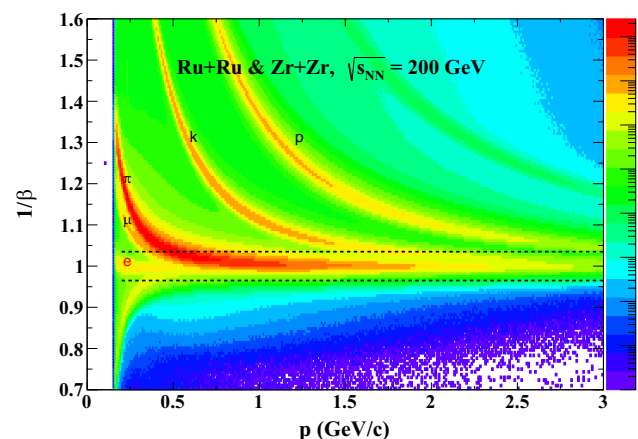


Fig. 1 (Color online) Distribution of $1/\beta$ versus momentum for charged particles, where the dashed horizontal lines indicate the $1/\beta$ cuts for electron preselection

¹ Here, “electron” denotes both electrons and positrons unless otherwise specified.

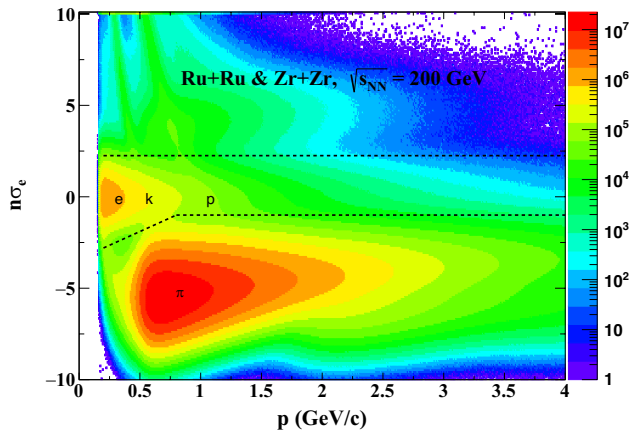


Fig. 2 (Color online) Distribution of $n\sigma_e$ versus momentum for charged particles after applying the $|1/\beta - 1| < 0.025$ cut. The dashed curves correspond to the $n\sigma_e$ cuts used for electron preselection

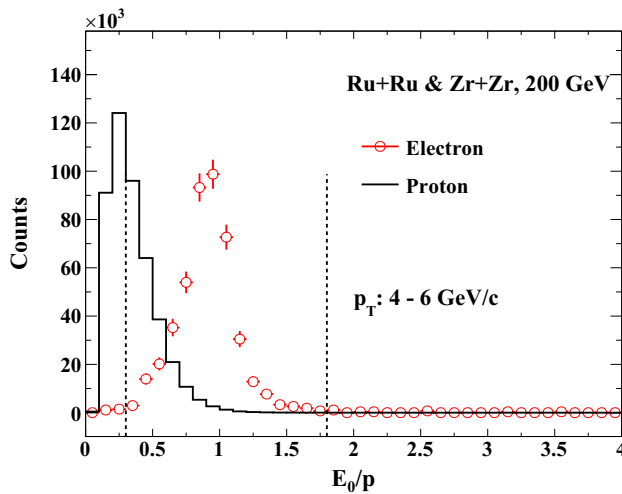


Fig. 3 (Color online) Distributions of E_0/p for electrons (circles) and protons (solid line) within $4 < p_T < 6$ GeV/c. The vertical dashed lines correspond to the E_0/p cuts used for electron preselection

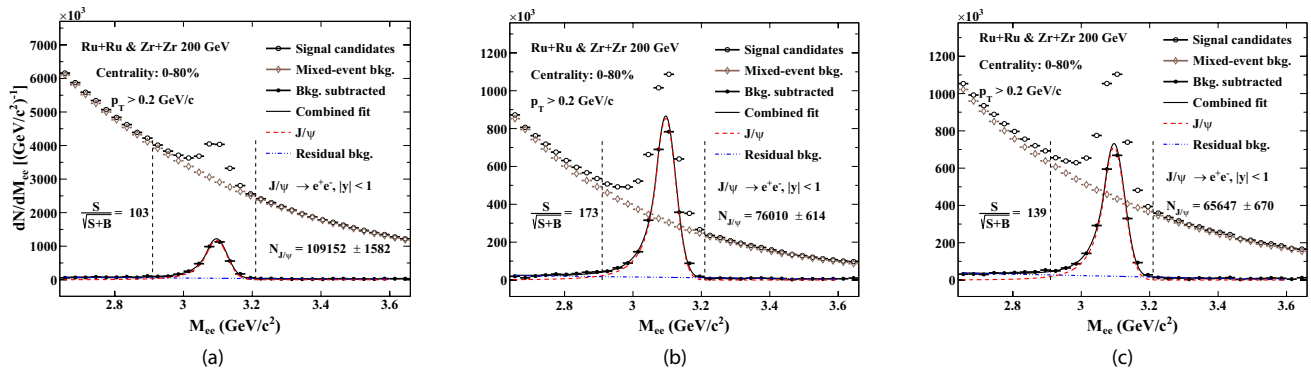


Fig. 4 (Color online) Invariant mass distributions of J/ψ candidates reconstructed via the dielectron channel, within $|y| < 1$ and $p_T > 0.2$ GeV/c, for Ru+Ru and Zr+Zr collisions. Different panels correspond

range $4 < p_T < 6$ GeV/c. The electron sample, often referred to as photonic electrons, is obtained from photon conversions and neutral-meson Dalitz decays [41]. As expected, the electron distribution peaks at a larger E_0/p value than the proton distribution. The dashed vertical lines indicate the preselection requirement on E_0/p .

The use of the TOF, TPC, and BEMC preselection criteria depends on the track kinematics and on whether matching information is available in the TOF and BEMC, as summarized in Table 1. Electron and positron candidates that satisfy these requirements are paired to form J/ψ candidates, and the resulting invariant-mass distribution is shown in Fig. 4a. The combinatorial background is estimated with the mixed-event technique [42]. After subtracting this component, the remaining invariant-mass distribution is fitted with two terms: a Crystal Ball function for the J/ψ signal and a linear term for the residual background arising mainly from semileptonic heavy-flavor decays and Drell–Yan production. The raw J/ψ yield is obtained by counting the electron–positron pairs in the mass window $[2.91, 3.21]$ GeV/ c^2 and subtracting the fitted residual-background contribution in the same interval. Approximately, 1.1×10^5 J/ψ candidates are reconstructed, with a signal significance of 103. The following sections show how machine learning further improves this significance.

3 Construction of training samples

For J/ψ reconstruction, the classification targets are electron–positron pairs that form J/ψ candidates. To determine whether a candidate is signal-like or background-like, a set of discriminating features must be defined. In principle, any observable that improves the separation without introducing physics bias can be used. In this work, the selected features are $n\sigma_e$, $n\sigma_\pi$, E_0/p , and the distance of closest approach (DCA, defined as the minimum distance between

to different electron identification cuts. **a** With preselection cuts listed in Table 1; **b** With preselection and BDT cuts listed in Table 1; **c** With straight cuts listed in Table 2

a charged-particle trajectory and the collision vertex) for both daughter tracks, labeled as “tr1” and “tr2”.

The background training sample is selected from electron–positron pairs in the invariant-mass sidebands, namely (2.65, 2.85), (3.25, 3.5), and (3.8, 4.5) GeV/ c^2 , well separated from the J/ψ peak. The interval (3.5, 3.8) GeV/ c^2 is excluded because it contains a small contribution from $\psi(2S)$ mesons. This sideband method assumes that the background in the sidebands is representative of the background beneath the J/ψ peak. Therefore, the features used for classification should not depend strongly on the pair invariant mass, which constitutes an important feature-selection criterion. Figure 5 shows the correlations between the selected features and the invariant mass for the background sample; no strong mass dependence is observed. As expected, $n\sigma_e$ and $n\sigma_\pi$ are strongly correlated because both are derived from the same dE/dx measurement.

The signal training sample is generated with Monte Carlo simulation. Electrons and positrons from J/ψ decays are propagated through a GEANT3-based [43] model of the STAR detector, and the simulated hits are embedded into real Ru+Ru and Zr+Zr events before reconstruction with the same software chain used for data. Inevitably, the simulation does not perfectly reproduce the data because detector conditions evolve with time and the heavy-ion environment

has high occupancy. The quality of the simulation can be assessed by comparing the $n\sigma_e$, $n\sigma_\pi$, E_0/p , and DCA distributions for electrons from simulated J/ψ decays with those for photonic electrons in data, as shown in Fig. 6. The DCA distributions agree reasonably well, whereas the $n\sigma_e$ and $n\sigma_\pi$ distributions exhibit clear offsets and the E_0/p distribution shows a noticeable shape mismatch. If left uncorrected, these discrepancies can degrade the machine learning performance and bias the extracted J/ψ yield.

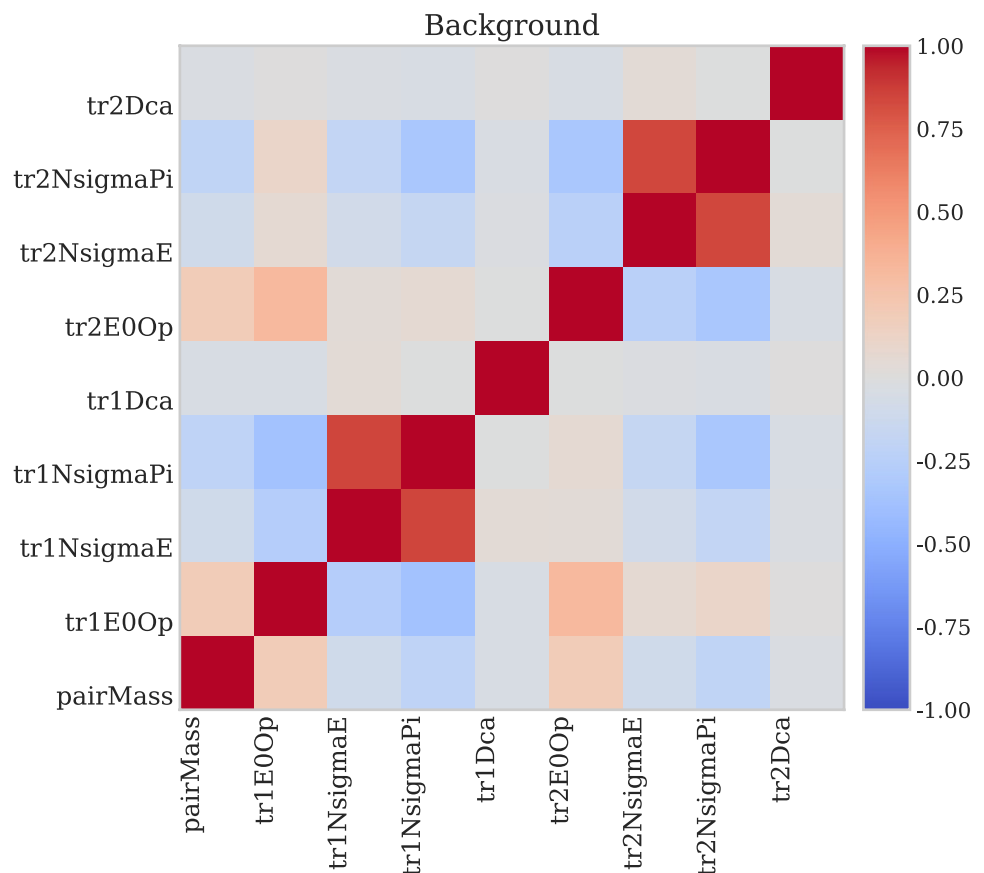
Because the simulated and measured $n\sigma_e$ and $n\sigma_\pi$ distributions have similar shapes, they can be corrected with a shift-and-scale transformation:

$$n\sigma_{i,\text{cor}} = (n\sigma_i - \mu_{\text{simu}}) \cdot (\sigma_{\text{data}}/\sigma_{\text{simu}}) + \mu_{\text{data}}. \quad (1)$$

Here, μ_{data} (σ_{data}) and μ_{simu} (σ_{simu}) are the means (standard deviations) of the corresponding $n\sigma_i$ distributions in data and simulation, respectively, with $i = e$ or π . After this transformation, the simulated $n\sigma_e$ and $n\sigma_\pi$ distributions agree well with those observed in data, as shown in Fig. 6a and b.

For the E_0/p distribution, the mismatch between simulation and data is not limited to a simple offset or width difference, so shift-and-scale is insufficient. We therefore introduce a cumulative distribution function (CDF) mapping method. Let $f_{\text{simu}}(x)$ and $f_{\text{data}}(x)$, with $x \in [x_{\min}, x_{\max}]$, denote

Fig. 5 (Color online) Correlations between all selected features and the reconstructed electron–positron pair invariant mass for the background training sample. Each cell represents the correlation coefficient between a pair of variables, with positive correlations shown in red and negative correlations in blue



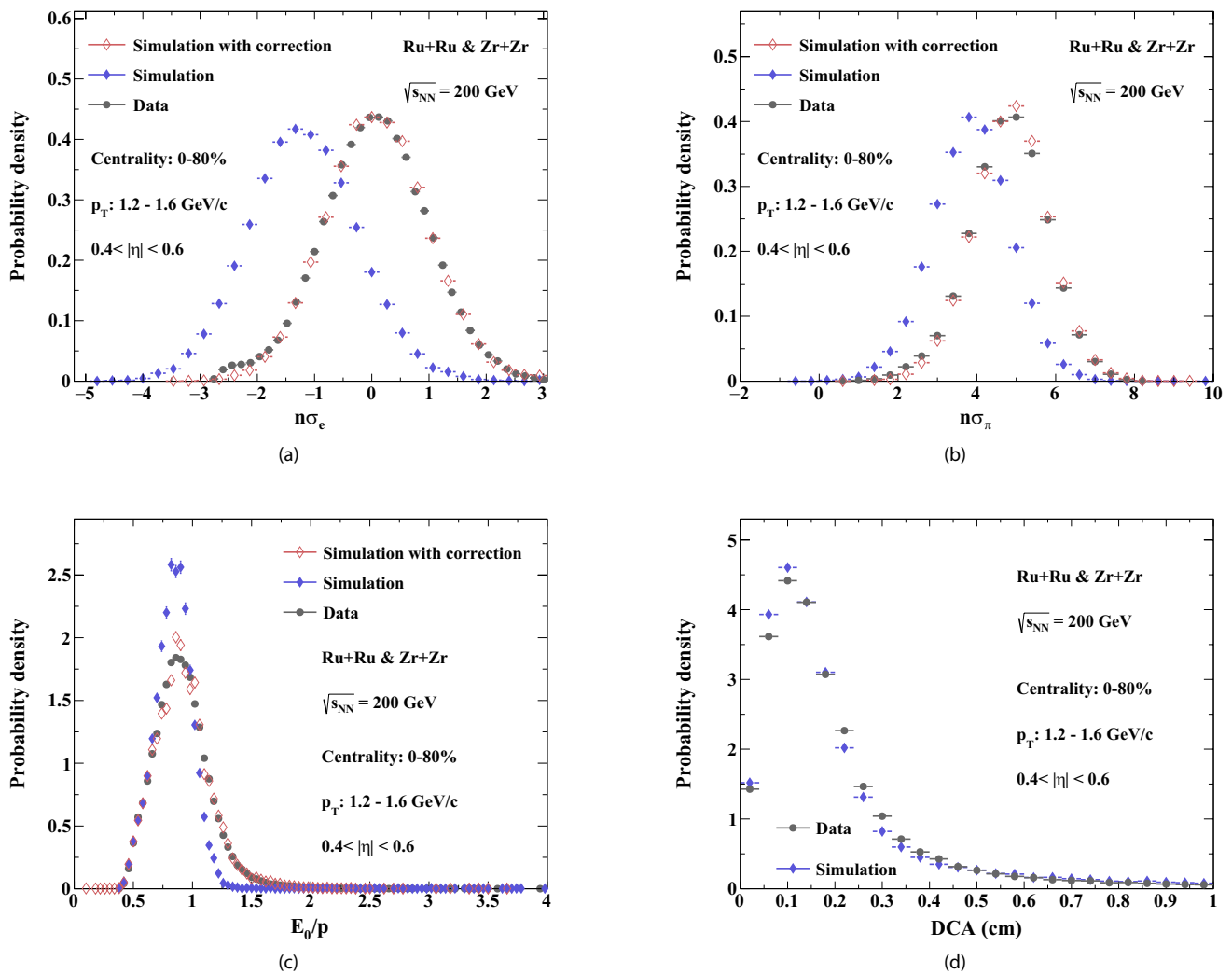


Fig. 6 (Color online) Comparisons of feature distributions for electrons within $1.2 < p_T < 1.6$ GeV/c and $0.4 < |\eta| < 0.6$ from simulations, both with (open diamonds) and without (filled diamonds) cor-

rections, to those from real data (filled circles). **a** $n\sigma_e$ distributions; **b** $n\sigma_\pi$ distributions; **c** E_0/p distributions; **d** DCA distributions

the probability density functions of a feature in simulation and data, respectively. The corresponding CDF is

$$C(x) = \int_{x_{\min}}^x f(x') dx'.$$

For an event in simulation with feature value A , its cumulative probability is

$$C_{\text{simu}}(A) = \int_{x_{\min}}^A f_{\text{simu}}(x') dx'.$$

The corrected value B is then obtained from the data distribution through

$$C_{\text{data}}(B) = C_{\text{simu}}(A) \quad \Rightarrow \quad B = C_{\text{data}}^{-1}(C_{\text{simu}}(A)).$$

This procedure defines an event-by-event mapping from the original value A to the corrected value B , such that the transformed simulation reproduces the data distribution. After CDF mapping, the corrected simulated E_0/p distribution (open diamonds) is consistent with the distribution measured in data (solid circles), as shown in Fig. 6c. In practice, the PDFs are represented by discrete histograms, so the method is implemented numerically by constructing discrete CDFs and searching for the closest cumulative probability. When the simulation and data differ strongly in shape, this discrete matching can produce bin-to-bin fluctuations at the level of one histogram bin, leading to localized excesses in some bins and deficits in neighboring bins. This effect explains the statistical irregularities visible in Fig. 6c. Because the classifier is trained on event-by-event continuous feature values rather than on the displayed histogram counts themselves,

this binning artifact does not compromise the model training. It can be further reduced by decreasing the histogram bin width.

Compared with shift-and-scale, CDF mapping is more general but also more involved to implement. Both procedures act deterministically on each event and therefore retain the event-level correspondence among observables, which is essential for multivariate classifiers to benefit from joint feature information. For the same reason, the signal sample cannot be constructed by independently sampling one-dimensional feature distributions from data, because that procedure would destroy the physically relevant inter-feature dependence structure.

4 Model training and application

A BDT classifier implemented with XGBoost is trained to further improve the J/ψ signal significance. The signal and background samples are each divided into two equal subsets, one for training and the other for testing. After training, the model is applied to both subsets. For every electron–positron pair, the classifier returns an output between 0 and 1, with values close to 1 corresponding to signal-like candidates and values close to 0 corresponding to background-like candidates. The output distributions are shown in Fig. 7. The red band and red points denote the signal distributions for the training and testing samples, respectively, whereas the blue band and blue points show the corresponding background distributions. The strong separation between signal and background indicates good discriminating power, and the agreement between the training and testing samples shows that overtraining is not evident.

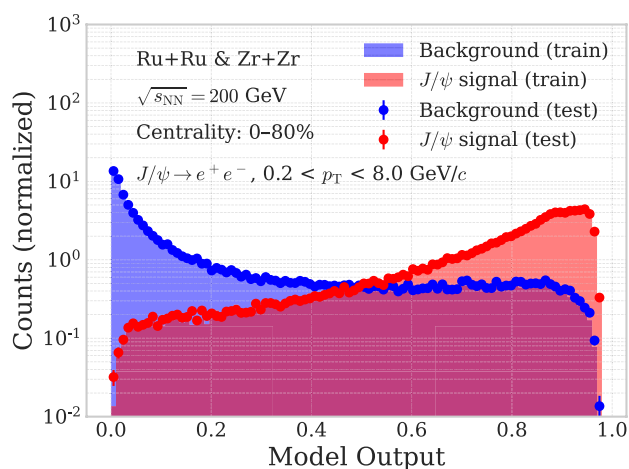


Fig. 7 (Color online) Model output distributions for both the training and testing samples. The red (blue) histogram represents the training signal (background) sample, while the red (blue) points are for the testing signal (background) sample

In the final analysis, a threshold on the BDT output is chosen so that pairs with scores above the threshold are retained as signal candidates and those below it are rejected as background. For a given threshold, the signal efficiency is defined as the fraction of signal events that survive the cut, and the background efficiency is defined analogously. Plotting the signal efficiency against the background efficiency while scanning the threshold from 0 to 1 yields the receiver operating characteristic (ROC) curve, where the horizontal and vertical axes correspond to the false-positive rate (FPR) and true-positive rate (TPR), respectively. The ROC curves for the training and testing samples are shown in Fig. 8. The area under the curve (AUC) quantifies the overall classification performance; values closer to 1 indicate better separation. The AUC obtained here is 0.91, demonstrating strong discriminating power. The nearly identical ROC curves and AUC values for the training and testing samples provide additional evidence that the model is not overfit.

5 Model validation via self-consistency checks

Beyond establishing that the model is not overtrained, it is necessary to verify that the BDT efficiencies for both signal and background are correctly reproduced in real data. This validation is carried out through two self-consistency tests.

The first test examines the J/ψ yield corrected for the BDT efficiency as a function of the BDT threshold. Because the measured yield should not depend on the particular analysis threshold, the corrected yield is expected to remain constant within uncertainties. The J/ψ efficiency for a given BDT threshold is derived from the signal output distribution in Fig. 7, and the resulting efficiency curves are shown in

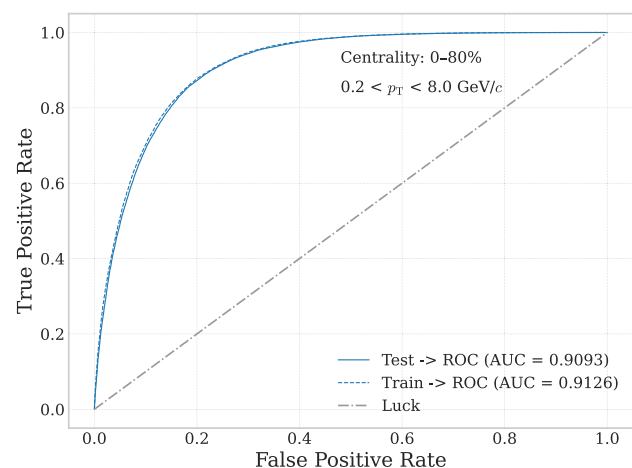


Fig. 8 ROC curves for the training (dashed) and testing (solid) samples. The reported AUC values quantify the overall performance, while the diagonal line represents a random classifier for comparison

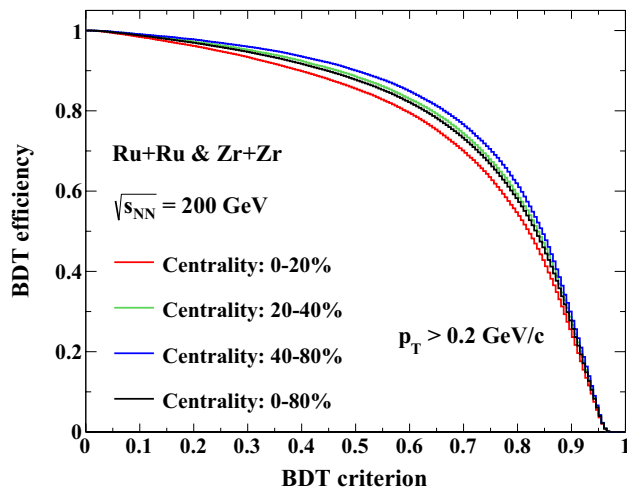


Fig. 9 (Color online) BDT cut efficiencies for J/ψ with $p_T > 0.2$ GeV/c in different centrality classes of 200 GeV Ru+Ru and Zr+Zr collisions

Fig. 9 for several centrality intervals in Ru+Ru and Zr+Zr collisions. Here, centrality characterizes the geometric overlap of the two nuclei: 0–20% corresponds to the most central collisions, whereas 40–80% corresponds to more peripheral collisions. After applying the preselection criteria in Table 1 and varying the BDT threshold, the raw J/ψ counts are extracted from data with the procedure described in Sect. 2. Figure 10 then shows the efficiency-corrected J/ψ counts as a function of the BDT threshold for four centrality classes. Although the signal efficiency changes by roughly an order of magnitude as the threshold increases from 0 to 0.9, the corrected counts remain stable, with variations below 2%. This stability strongly suggests that the signal-efficiency estimate from the machine learning model accurately represents the efficiency in data, thereby validating the feature-correction procedure described in Sect. 3.

A more stringent validation compares the J/ψ significance as a function of the BDT threshold between the machine learning expectation and the measurement in data, because this test depends on both the signal and background efficiencies. The expected J/ψ significance is calculated as

$$\begin{aligned} S(x) &= S_0 \cdot \text{eff}_s(x)/\text{eff}_s(x_0), \\ B(x) &= B_0 \cdot \text{eff}_b(x)/\text{eff}_b(x_0), \\ \text{Significance}_{\text{expected}}(x) &= S(x)/\sqrt{S(x) + B(x)}. \end{aligned} \quad (2)$$

In Eq. 2, x is the BDT threshold, and $\text{eff}_s(x)$ and $\text{eff}_b(x)$ are the corresponding signal and background efficiencies. The quantities $S(x)$ and $B(x)$ denote the signal and background counts after the threshold is applied, while S_0 and B_0 are the counts at a reference threshold x_0 . The choice of x_0 is arbitrary and is used only to normalize the prediction to the data at one point. The resulting expected significance is shown as

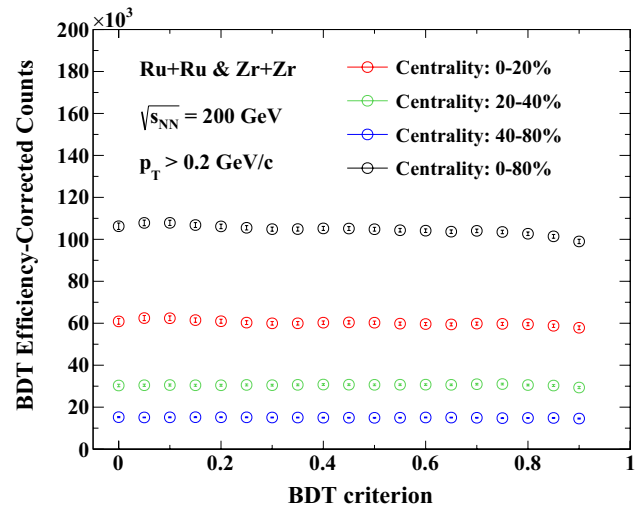


Fig. 10 (Color online) J/ψ counts, corrected by BDT cut efficiencies, as a function of BDT criterion in different centrality classes of 200 GeV Ru+Ru and Zr+Zr collisions

the gray band in Fig. 11. It first increases and then decreases as the threshold is tightened from 0 to 0.9, reflecting the competition between improved background rejection and reduced signal efficiency. The prediction from the trained model is compared with the significance extracted directly from data, shown as open circles. For clarity, the two curves are aligned at a threshold of 0.5, namely $x_0 = 0.5$ in Eq. 2. Their consistency indicates that the training samples reproduce the behavior of both signal and background in data with sufficient accuracy for this analysis.

The expected-significance curve from the trained model also provides a principled way to determine the optimal BDT threshold. Compared with scanning the data directly, an optimization based on the training samples is less sensitive to statistical fluctuations in the measured spectrum and therefore less prone to introducing threshold-selection bias, which is particularly important for low-statistics signals. As shown in Fig. 11, the J/ψ significance reaches its maximum at a BDT threshold of 0.72, which is therefore chosen as the working point. Figure 4b shows the raw signal extraction obtained with this optimal threshold. The corresponding significance is 173, representing a 68% improvement over the result with only the preselection criteria (Fig. 4a) and a 24% improvement over the result obtained with straight PID cuts (Fig. 4c and Table 2).

The same self-consistency tests are also useful for diagnosing problematic input features. If adding a given feature destroys the agreement between the model expectation and the data, that feature should be corrected or removed from the training. As an example, Fig. 12 in the Appendix shows the result obtained without CDF mapping for E_0/p . In that case, the efficiency-corrected signal counts deviate

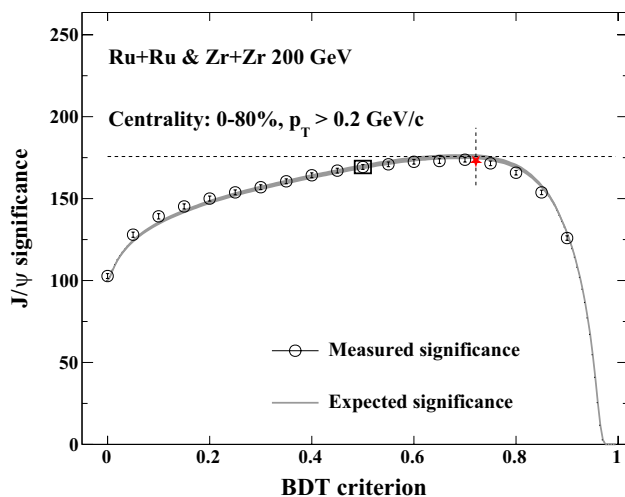


Fig. 11 (Color online) J/ψ significance as a function of BDT criterion. The black open circles represent the J/ψ significance directly obtained from real data while the black band shows the expected significance estimated from the machine learning model. The two distributions are aligned at BDT value of 0.5 as indicated by the black open square. The vertical dashed line at 0.72 marks the point where the expected significance reaches its maximum value of 176 as indicated by the horizontal dashed line. The red star represents the J/ψ significance of 173 obtained from real data at the same BDT criterion

progressively from the expected flat behavior as the BDT threshold increases, reaching 29% at a threshold of 0.9. Likewise, the expected J/ψ significance becomes inconsistent with the significance measured in data, as shown in Fig. 13. These discrepancies provide direct evidence that the feature-correction procedures introduced in Sect. 3 are necessary for unbiased signal extraction.

6 Summary

This work develops two practical procedures, CDF mapping and shift-and-scale, to align simulated feature distributions with those measured in data while preserving the event-level multivariate information required for machine learning classification. These methods address a common source of systematic bias in experimental analyses, namely the mismatch between simulated and real detector responses. Using J/ψ yield extraction in $\sqrt{s_{NN}} = 200$ GeV Ru+Ru and Zr+Zr collisions as a case study, we show that the corrected training samples lead to stable efficiency-corrected yields and to a predicted significance consistent with the significance measured in data. The self-consistency tests therefore validate both the necessity and the effectiveness of the

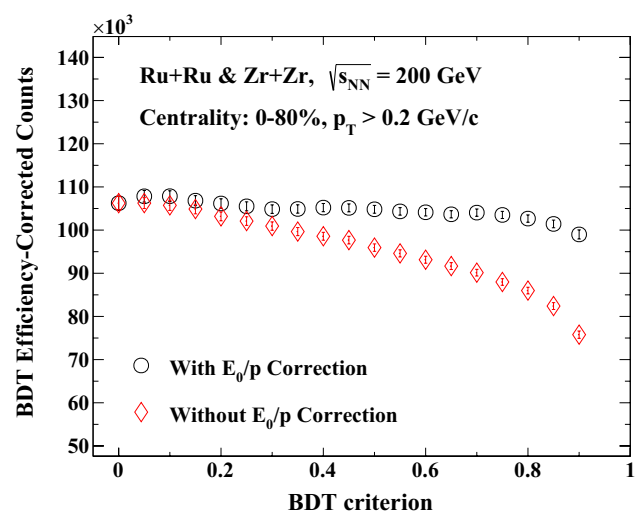


Fig. 12 (Color online) J/ψ counts, corrected for the BDT cut efficiencies, as a function of BDT criterion, for scenarios that the signal sample is incorporated with (circles) or without (diamonds) the E_0/p correction

proposed corrections. The same strategy should be broadly useful in machine learning applications throughout high-energy nuclear and particle physics whenever simulated signal samples do not fully reproduce the data.

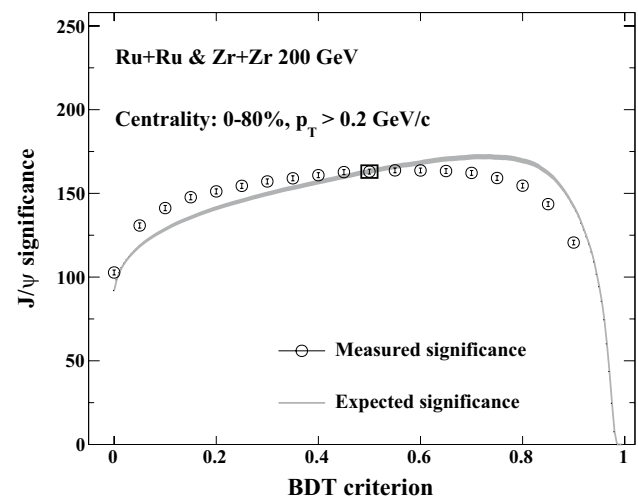


Fig. 13 (Color online) Expected J/ψ significance as a function of the BDT criterion from the machine learning model trained using the signal sample without the E_0/p correction. It is compared to the measured significance in real data (open circles). The two distributions are aligned at the BDT criterion of 0.5, indicated by the open square, to facilitate the comparison

Table 2 Straight cuts used for electron identification

Track p_T	Detectors used	Electron identification cuts
$p_T \leq 1.0 \text{ GeV}/c$	TPC, TOF	$ 1/\beta - 1 < 0.025$; for $p > 0.8 \text{ GeV}/c$: $-0.75 < n\sigma_e < 2$; for $p \leq 0.8 \text{ GeV}/c$: $(3 \times p - 3.15) < n\sigma_e < 2$
$p_T > 1.0 \text{ GeV}/c$	TPC, TOF and BEMC no matching	$ 1/\beta - 1 < 0.025$; $-0.75 < n\sigma_e < 2$
	TPC, BEMC and TOF no matching	$-1 < n\sigma_e < 2$; $0.5 < E_0/p < 1.5$
	TPC, TOF and BEMC	$ 1/\beta - 1 < 0.025$; $-1.5 < n\sigma_e < 2$; $0.5 < E_0/p < 1.5$

Appendix A: Results without the E_0/p correction

Figure 12 compares the BDT-efficiency-corrected J/ψ yield as a function of the BDT threshold for two cases: with and without aligning the E_0/p distribution of the training signal sample to that observed in data.

Figure 13 compares the expected J/ψ significance from a model trained without CDF mapping for E_0/p with the significance measured directly in data.

Appendix B: Straight electron identification cuts

The table below lists the conventional straight PID cuts used as the reference selection for comparison with the machine learning-based approach.

Acknowledgements The authors thank the STAR Collaboration and the RHIC Operations Group for providing the data used in this study. They also acknowledge the RCF facility at Brookhaven National Laboratory for the computational resources required for this analysis. Helpful discussions with Peng-Zhong Lu and Zhen-Jun Xiong are gratefully acknowledged.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Yan Wang, Rong-Rong Ma, Kai-Feng Shen, Ze-Bo Tang, and Wang-Mei Zha. The first draft of the manuscript was written by Yan Wang, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Generative AI use During the preparation of this work, the authors used ChatGPT-4.0 for language editing and grammar checking only. The tool was used to improve the readability and clarity of the text. After using this service/tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article. No AI tool was used to generate scientific data, perform analysis, or formulate conclusions.

References

1. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015). <https://doi.org/10.1038/nature14539>
2. L. Pang, K. Zhou, X. Wang, Deep learning for nuclear physics. *Nucl. Phys. Rev.* **37**(3), 720–726 (2020). <https://doi.org/10.11804/NuclPhysRev.37.2019CNPC41>
3. F.P. Li, L.G. Pang, G.Y. Qin, QCD equation of state at finite μ_B using a deep learning-assisted quasi-parton model. *Phys. Lett. B* **868**, 139692 (2025). <https://doi.org/10.1016/j.physletb.2025.139692>
4. W. He, Q. Li, Y. Ma et al., Machine learning in nuclear physics at low and intermediate energies. *Sci. China Phys. Mech. Astron.* **66**(8), 282001 (2023). <https://doi.org/10.1007/s11433-023-2116-0>
5. Z. Li, Y. Wang, Q. Li et al., Machine-learning predictions for the nuclear charge radius: Bayesian method versus decision-tree-based algorithm. *Phys. Rev. C* **112**(1), 014312 (2025). <https://doi.org/10.1103/vj25-zwd3>
6. Z.M. Niu, J.Y. Fang, Y.F. Niu, Comparative study of radial basis function and Bayesian neural network approaches in nuclear mass predictions. *Phys. Rev. C* **100**(5), 054311 (2019). <https://doi.org/10.1103/PhysRevC.100.054311>
7. B. Zhou, Y.Y. Cao, J.Y. Guo, Predictions of nuclear charge radii based on the convolutional neural network. *Nucl. Sci. Tech.* **34**(10), 152 (2023). <https://doi.org/10.1007/s41365-023-01308-x>
8. Q.F. Li, Y.J. Wang, Z.P. Gao et al., Machine learning the nuclear mass. *Nucl. Sci. Tech.* **32**(10), 109 (2021). <https://doi.org/10.1007/s41365-021-00956-1>
9. L. Tang, Z.H. Zhang, Nuclear charge radius predictions by kernel ridge regression with odd-even effects. *Nucl. Sci. Tech.* **35**(2), 19 (2024). <https://doi.org/10.1007/s41365-024-01379-4>
10. Z.Z. Ren, Z.Y. Yuan, D. Bai et al., Reliable calculations of nuclear binding energies by the gaussian process of machine learning. *Nucl. Sci. Tech.* **35**(6), 105 (2024). <https://doi.org/10.1007/s41365-024-01463-9>
11. G.Q. Zhang, X.G. Cao, G.Y. Cheng et al., The study of intelligent algorithm in particle identification of heavy-ion collisions at low and intermediate energies. *Nucl. Sci. Tech.* **35**(2), 33 (2024). <https://doi.org/10.1007/s41365-024-01388-3>
12. K. Albertsson, P. Alton, D. Anderson et al., Machine learning in high energy physics community white paper. *J. Phys. Conf. Ser.* **1085**(2), 022008 (2018). <https://doi.org/10.1088/1742-6596/1085/2/022008>
13. A.M. Sirunyan, A. Tumasyan, W. Adam et al., Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques. *J. Instrum.* **15**(06), P06005 (2020). <https://doi.org/10.1088/1748-0221/15/06/P06005>
14. S. Acharya, D. Adamová, A. Agarwal et al., Measurement of beauty production via non-prompt charm hadrons in p–Pb

- collisions at $\sqrt{s_{NN}} = 5.02$ TeV. *J. High Energy. Phys.* **2024**, 148 (2024). [https://doi.org/10.1007/JHEP11\(2024\)148](https://doi.org/10.1007/JHEP11(2024)148)
15. S. Acharya, D. Adamová, A. Adler et al., First measurement of Ω_c^0 production in pp collisions at $\sqrt{s}=13$ TeV. *Phys. Lett. B* **846**, 137625 (2023). <https://doi.org/10.1016/j.physletb.2022.137625>
 16. S. Acharya, D. Adamová, A. Adler et al., First measurement of Λ_c^+ production down to $p_T = 0$ in pp and p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. *Phys. Rev. C* **107**(6), 064901 (2023). <https://doi.org/10.1103/PhysRevC.107.064901>
 17. S. Acharya, D. Adamová, G. Aglieri Rinella et al., Study of flavor dependence of the baryon-to-meson ratio in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D* **108**(11), 112003 (2023). <https://doi.org/10.1103/PhysRevD.108.112003>
 18. Y. Coadou, Boosted decision trees. in *Artificial Intelligence for High Energy Physics*, ed. by P. Calafiura, D. Rousseau, K. Terao (World Scientific, Singapore, 2022), pp. 9–58. https://doi.org/10.1142/9789811234033_0002
 19. B.P. Roe, H.J. Yang, J. Zhu et al., Boosted decision trees, an alternative to artificial neural networks. *Nucl. Instrum. Meth. A* **543**(2–3), 577–584 (2005). <https://doi.org/10.1016/j.nima.2004.12.018>
 20. J. Schmidhuber, Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015). <https://doi.org/10.1016/j.neunet.2014.09.003>
 21. C.F. Madrazo, I.H. Cacha, L.L. Iglesias et al., Application of a convolutional neural network for image classification for the analysis of collisions in high energy physics. *EPJ Web Conf.* **214**, 06017 (2019). <https://doi.org/10.1051/epjconf/201921406017>
 22. P. Baldi, K. Bauer, C. Eng et al., Jet substructure classification in high-energy physics with deep neural networks. *Phys. Rev. D* **93**(9), 094034 (2016). <https://doi.org/10.1103/PhysRevD.93.094034>
 23. I. Bentley, J. Tedder, M. Gebran et al., High precision binding energies from physics-informed machine learning. *Phys. Rev. C* **111**(3), 034305 (2025). <https://doi.org/10.1103/PhysRevC.111.034305>
 24. T. Matsui, H. Satz, J/ψ suppression by Quark-Gluon plasma formation. *Phys. Lett. B* **178**, 416–422 (1986). [https://doi.org/10.1016/0370-2693\(86\)91404-8](https://doi.org/10.1016/0370-2693(86)91404-8)
 25. M. Gyulassy, The QGP discovered at RHIC. In: W. Greiner, M.G. Itkis, J. Reinhardt, M.C. Güçlü (eds) *Structure and dynamics of elementary matter*. NATO Science Series (Springer, Dordrecht, 2004), vol 166, pp. 159–182. https://doi.org/10.1007/978-1-4020-2705-5_14
 26. L. Adamczyk et al., (STAR), J/ψ production at low p_T in Au + Au and Cu + Cu collisions at $\sqrt{s_{NN}} = 200$ GeV with the STAR detector. *Phys. Rev. C* **90**(2), 024906 (2014). <https://doi.org/10.1103/PhysRevC.90.024906>
 27. L. Adamczyk, G. Agakishiev, M.M. Aggarwal et al., J/ψ production at high transverse momenta in p+p and Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV. *Phys. Lett. B* **722**, 55–62 (2013). <https://doi.org/10.1016/j.physletb.2013.04.010>
 28. L. Adamczyk, J.K. Adkins, G. Agakishiev et al., Energy dependence of J/ψ production in Au+Au collisions at $\sqrt{s_{NN}} = 39, 62.4$ and 200 GeV. *Phys. Lett. B* **771**, 13–20 (2017). <https://doi.org/10.1016/j.physletb.2017.04.078>
 29. A. Adare, S. Afanasiev, C. Aidala et al., J/ψ Production vs. centrality, transverse momentum, and rapidity in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV. *Phys. Rev. Lett.* **98**, 232301 (2007). <https://doi.org/10.1103/PhysRevLett.98.232301>
 30. B.B. Abelev, J. Adam, D. Adamová et al., Centrality, rapidity and transverse momentum dependence of J/ψ suppression in Pb-Pb collisions at $\sqrt{s_{NN}}=2.76$ TeV. *Phys. Lett. B* **734**, 314–327 (2014). <https://doi.org/10.1016/j.physletb.2014.05.064>
 31. J. Adam, D. Adamová, M.M. Aggarwal et al., Inclusive, prompt and non-prompt J/ψ production at mid-rapidity in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *J. High Energy. Phys.* **2015**, 051 (2015). [https://doi.org/10.1007/JHEP07\(2015\)051](https://doi.org/10.1007/JHEP07(2015)051)
 32. J. Adam, L. Adamczyk, J.R. Adams et al., Measurement of inclusive J/ψ suppression in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV through the dimuon channel at STAR. *Phys. Lett. B* **797**, 134917 (2019). <https://doi.org/10.1016/j.physletb.2019.134917>
 33. Y. Wang et al., (STAR Collaboration), J/ψ production in Ru + Ru and Zr + Zr collisions at $\sqrt{s_{NN}} = 200$ GeV with the STAR experiment. *PoS HardProbes2023*, 110 (2024). <https://doi.org/10.22323/1.438.0110>
 34. Z. Tang, W. Zha, Y. Zhang, An experimental review of open heavy flavor and quarkonium production at RHIC. *Nucl. Sci. Tech.* **31**(8), 81 (2020). <https://doi.org/10.1007/s41365-020-00785-8>
 35. T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System (2016). <https://doi.org/10.1145/2939672.2939785>
 36. J. Chen, X. Dong, X.H. He et al., Properties of the QCD matter: review of selected results from the relativistic heavy ion collider beam energy scan (RHIC BES) program. *Nucl. Sci. Tech.* **35**(12), 214 (2024). <https://doi.org/10.1007/s41365-024-01591-2>
 37. W.J. Llope et al., (STAR Collaboration), The large-area time-of-flight (TOF) upgrade for the STAR detector. *AIP Conf. Proc.* **1099**(1), 778–781 (2009). <https://doi.org/10.1063/1.3120153>
 38. H. Wieman, E. Anderssen, A. Banerjee et al., STAR TPC at RHIC. *IEEE Trans. Nucl. Sci.* **44**, 671–678 (1997). <https://doi.org/10.1109/23.603731>
 39. M. Anderson, J. Berkovitz, W. Betts et al., The star time projection chamber: a unique tool for studying high multiplicity events at RHIC. *Nucl. Instrum. Meth. A* **499**, 659–678 (2003). [https://doi.org/10.1016/S0168-9002\(02\)01964-2](https://doi.org/10.1016/S0168-9002(02)01964-2)
 40. M. Beddo, E. Bielick, T. Fornek et al., The STAR barrel electromagnetic calorimeter. *Nucl. Instrum. Meth. A* **499**, 725–739 (2003). [https://doi.org/10.1016/S0168-9002\(02\)01970-8](https://doi.org/10.1016/S0168-9002(02)01970-8)
 41. L. Adamczyk, J.K. Adkins, G. Agakishiev et al., Measurements of dielectron production in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV from the STAR experiment. *Phys. Rev. C* **92**(2), 024912 (2015). <https://doi.org/10.1103/PhysRevC.92.024912>
 42. A. Adare, S. Afanasiev, C. Aidala et al., Detailed measurement of the e^+e^- pair continuum in p+p and Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV and implications for direct photon production. *Phys. Rev. C* **81**, 034911 (2010). <https://doi.org/10.1103/PhysRevC.81.034911>
 43. B.I. Abelev, M.M. Aggarwal, Z. Ahammed et al., Systematic measurements of identified particle spectra in pp, d+Au and Au+Au collisions from STAR. *Phys. Rev. C* **79**, 034909 (2009). <https://doi.org/10.1103/PhysRevC.79.034909>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.