Nuclear Science and Techniques (2023) 34:158
https://doi.org/10.1007/s41365-023-01320-1

Applying the Kalman filter particle method to strange and open charm hadron reconstruction in the STAR experiment

Xin-Yue Ju[1,2] · Yue-Hang Leung[2,3] · Sooraj Radhakrishnann[2,4] · Petr Chaloupka[5] · Xin Dong[2] · Yury Fisyak[6] · Pavol Federic[7] · Ivan Kisel[8,9,10] · Hong-Wei Ke[6] · Michal Kocan[5] · Spyridon Margetis[4] · Ai-Hong Tang[6] · Iouri Vassiliev[10] · Yi-Fei Zhang[1] · Xiang-Lei Zhu[11] · Maksym Zyzak[10]

Received: 24 March 2023 / Revised: 5 July 2023 / Accepted: 11 July 2023 / Published online: 30 October 2023

Abstract
We applied KF Particle, a Kalman Filter package for secondary vertex finding and fitting, to strange and open charm hadron reconstruction in heavy-ion collisions in the STAR experiment. Compared to the conventional helix swimming method used in STAR, the KF Particle method considerably improved the reconstructed Λ, Ω, and D^0 significance. In addition, the Monte Carlo simulation with STAR detector responses could adequately reproduce the topological variable distributions reconstructed in real data using the KF Particle method, thereby retaining substantial control of the reconstruction efficiency uncertainties for strange and open charm hadron measurements in heavy-ion collisions.

Keywords Heavy-ion collisions · Secondary vertex finding · Kalman filter

1 Introduction

In high-energy particle and nuclear physics experiments, strange and heavy flavor hadrons aid in studying the electroweak and strong interactions in the Standard Model [1–3]. These particles are predominantly short-lived, and their ground-state particles such as K_S^0, Λ, D^0, and Λ_c^+ exhibit proper lifetimes ($c\tau$) varying from tens of micrometers to several centimeters [4]. The experimental reconstruction of the decay positions and their separation from the collision vertices is imperative for achieving precise measurements [5–7]. This becomes extremely critical in high-energy heavy-ion experiments at the RHIC and LHC, where the collision vertex produces thousands of particles. The process of secondary vertex reconstruction can significantly reduce the combinatorial background in these collisions while achieving a finite reconstruction efficiency, especially for low-momentum particles [5–7]. Therefore, the balance between the combinatorial background and the reconstruction efficiency must be considered for the particle of interest to achieve the best experimental measurement precision.

This work was supported by the National Natural Science Foundation of China (Nos. 11890712 and 12061141008) and the National Key R &D Program of China (Nos. 2018YFE0104700 and 2018YFE0205200). This work was supported in part by the Offices of NP and HEP within the U.S. DOE Office of Science; Yue-Hang Leung was partially supported by the GSI-Heidelberg cooperation contract.

✉ Yi-Fei Zhang
ephy@ustc.edu.cn

1 University of Science and Technology of China, Hefei 230026, China

2 Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

3 University of Heidelberg, 69120 Heidelberg, Germany

4 Kent State University, Kent, OH 44242, USA

5 Czech Technical University in Prague, Prague, Czech Republic

6 Brookhaven National Laboratory, Upton, NY 11973, USA

7 Nuclear Physics Institute of the Czech Academy of Sciences, Prague, Czech Republic

8 Goethe-Universität Frankfurt, Frankfurt am Main, Germany

9 Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany

10 GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, Germany

11 Tsinghua University, Beijing 100084, China

The STAR detector at RHIC serves as a general purpose detector dedicated to heavy-ion experiments [8]. The primary tracking subsystem, the time projection chamber (TPC) [9], provides a pointing resolution of ~ 1 mm to the collision vertex for charged tracks, which enables topological separation of strange hadron weak decay positions from the primary collision point. A high-resolution silicon detector, the Heavy Flavor Tracker (HFT), was operated from 2014 to 2016, which improved the charged track pointing resolution to more than ~ 50 μm for 750 MeV/c charged kaon tracks [10]. This enables the topological reconstruction of various open-charm hadron decays in heavy-ion collisions [5, 11–15] and significantly improves the precision of the measurements without necessitating the detection of the decay vertex [16]. Furthermore, the vertex resolution is sufficient to separate the open charm and open beauty hadron decays, which facilitates the measurement of beauty decay electrons to reveal mass-dependent parton energy loss in the hot-dense medium [17–19].

Conventionally, secondary vertex reconstruction in STAR has been conducted by determining the distance between the closest approach (DCA) points of two charged track helices, referred to as the helix swimming method (HS). Earlier, the decay position was regarded as the middle of the two DCA points, and this method has demonstrated adequate performance in reconstructing strange and open-channel hadrons in heavy-ion collisions [5, 6]. The key topological variables employed in this method is schematically represented in Fig. 1: DCA of daughter particles to the primary vertex (DCA_{v1}, DCA_{v2}), DCA between two daughter particles (DCA_{12}), decay length from the decay vertex position to the primary vertex (d), θ denotes the angle between the particle momentum vector of interest and the decay length vector, and/or the DCA between the interested particle helix and the primary vertex (b). The calculations were performed based on the mathematical helix model for the daughter tracks. The experimentally estimated uncertainties were excluded from the reconstruction method.

Recently, within STAR, an experimentally estimated error matrix on the track helix-fitted parameters was rendered in the offline analysis software infrastructure. Simultaneously, the KF Particle package, a Kalman Filter method used for secondary vertex finding and fitting utilizing the estimated track helix error matrices, was deployed for STAR offline analysis. Overall, this study aims to improve the secondary particle reconstruction with constraints provided by additional knowledge on the error matrices of various topological variables.

This paper reports the results of applying the KF Particle method to the reconstruction of strange (Λ,Ω^-) and open-charm (D^0) hadrons in heavy-ion collisions in the STAR experiments. A toolkit for multivariate analysis (TMVA) package deployed in ROOT [20] was used to optimize the topological selection cuts for the best signal significance in both the helix swimming and KF Particle methods. The remainder of this paper is organized as follows. Section 2 describes the mechanism followed by the KF Particle method to manage the secondary particle reconstruction and fitting. The application of the KF Particle method to the STAR data is discussed in Sect. 3. The optimized signal performance of the helix-swimming method and the KF Particle method are comparatively analyzed as well. The topological variable distributions from the KF Particle method obtained through the real data are comparatively analyzed with those derived from Monte Carlo (MC) simulations. Finally, the present findings are summarized in Sect. 4.

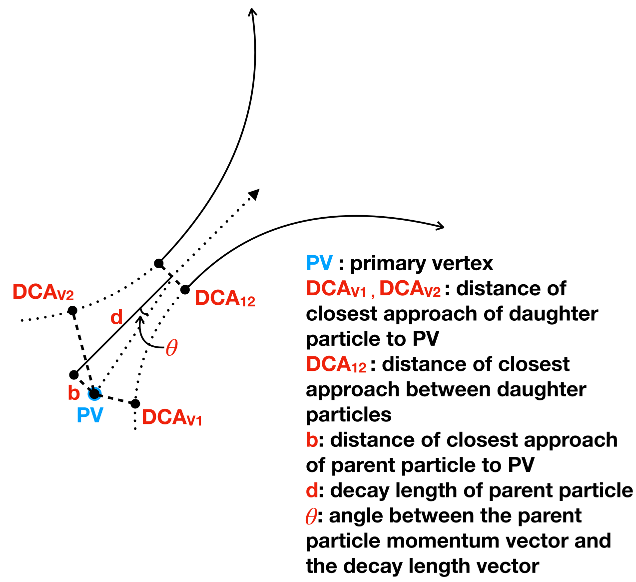

Fig. 1 Sketch of key topological variables used by the helix swimming method

2 KF particle method

The Kalman Filter (KF) [21] represents a recursive method for analyzing linear discrete dynamic systems described by a vector of parameters called the state vector $\boldsymbol{r}$ according to a series of measurements observed over time. It estimates the unknown vector parameters with high accuracy and is widely used in tracking and data prediction tasks.

In particle experiments, the Kalman filter can be employed to solve various tasks, such as track finding, particle reconstruction, and event vertex reconstruction [22]. In particular, the KF particle package utilizes the Kalman filter for the reconstruction of short-lived particles and vertex finding has been developed and is currently applied to STAR data analysis.

In the KF Particle framework, each particle is described by a state vector with eight parameters [23] $\boldsymbol{r} = (x, y, z, p_x, p_y, p_z, E, s)$, where (x, y, z), (p_x, p_y, p_z), and E are the position, and $s = l/p$, where l denotes the length of the trajectory in the laboratory coordinate system and p refers the total momentum of particle. This natural particle parametrization renders the algorithm independent of the detector system geometry. The reconstructed state vector and its covariance matrix ($\boldsymbol{C}$) contain all the necessary information regarding the particle, which enables the calculation of physical quantities such as momentum, energy, and lifetime with their accuracy and the χ^2 values during the reconstruction, i.e., to estimate the quality of reconstruction.

To simplify the calculation, the momentum and energy of the mother particle were calculated from the sum of all the daughter particles, and only the vertex position was fitted. After transporting the daughter particle to the current estimation of the decay vector $(r_k, \boldsymbol{C}_k)$, the state vector of the daughter particle can be deemed as a measurement $(m_k, \boldsymbol{V}_k)$ of the state vector of the mother particle. Using the residual ζ_k between r_k and m_k and the Kalman gain matrix $\boldsymbol{K}_k$ evaluated from $\boldsymbol{C}_k$ and $\boldsymbol{V}_k$, the estimation of the mother particle vector can be updated as $(r_{k+1}, \boldsymbol{C}_{k+1})$ according to Eq. (1).

$$\zeta_k = r_k - m_k,\ r_{k+1} = r_k + \boldsymbol{K}_k\zeta_k,\ \boldsymbol{C}_{k+1} = \boldsymbol{C}_k - \boldsymbol{K}_k\boldsymbol{C}_k^{'} \tag{1}$$

The χ^2 criterion for this estimation is obtained simultaneously. A basic filtering algorithm was developed by conducting this process on all daughter tracks. A complete description of the algorithm and its mathematical justification is detailed in Refs. [23, 24]. Herein, we briefly outline the scheme for short-lived particle reconstruction displayed in Fig. 2.

1. Sort the final state particles into primary and secondary according to its χ^2 to collision vertex.
2. Selection of an initial secondary decay point, often as the DCA point to the collision vertex from the first daughter track. Set the mother particle initial parameters $(\boldsymbol{r_0}, \boldsymbol{C_0})$, $\boldsymbol{C_0}$ is often set as an infinite diagonal matrix.
3. Extrapolation of the k-th daughter particle to the point of the closest approach with the current estimation of the decay point and update its parameters.
4. Correction of the decay vertex according to k-th daughter particle and adding the 4-momentum of the daughter particle to the 4-momentum of the mother particle.
5. Iteration of over all n daughter particles and calculation of an optimum estimation of the decay vector and its covariance matrix $(\boldsymbol{r}_n^i, \boldsymbol{C}_n^i)$ and the χ^2 probabilities.
6. If the production vertex of the mother particle (typically, the primary vertex) is known, the mother particles are transported to it. Thereafter, the position of the production vertex is filtered and the χ^2 probabilities of the origination are calculated from the production vertex.
7. Set $\boldsymbol{r}_n^i$ and $\boldsymbol{C}_n^i$ as the initial parameters of the mother particle and repeat steps 3–6 N times.
8. Finalize the precision of the mother particle parameters $(\boldsymbol{r_n}, \boldsymbol{C_n})$.

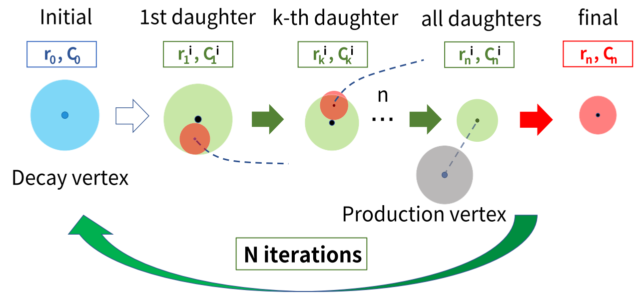

Fig. 2 (Color online) Schematic of short-lived particle reconstruction with the KF Particle package. The main steps are as follows: initialize the parameters of the mother particle; extrapolate a daughter particle to the DCA point with a current estimation of the mother particle; correct the mother particle with the parameters of its daughter particle; after correcting overall daughter particles, the optimum estimation of the mother particle is obtained; the parameters of current mother particle are input to the initial step and iterated several instances until the results converge

Compared with the traditional helix swimming method, the KF Particle method offers several crucial advantages.

- Usage of the daughter particle track parameters covariance matrices adds information on the detector performance and the track reconstruction quality, improving the mother particle reconstruction accuracy and efficiency.
- Statistical criteria (χ^2 based cuts) were calculated and used for background rejection, for instance, using χ^2 between the daughter track parameters and the collision point parameters instead of DCA to better discriminate primary and secondary particles.
- The natural and simple interface enables the reconstruction of the complicated decay chains [24].
- Usage of parallel programming provides high computational speed for the above-mentioned rather complicated calculations.

3 Application to data

We applied the KF Particle method to the reconstruction of strange (Λ, Ω^-) and open-charm (D^0) hadrons using the data collected from the STAR experiment. Recent experimental datasets of Au+Au collisions at $\sqrt{s_{\mathrm{NN}}}$ = 27 GeV (for Λ and Ω^-) and 200 GeV (for D^0), containing the error matrix information of the track parameters, were applied in this analysis.

3.1 Λ reconstruction

Λ particles were reconstructed using the decay channel $\Lambda \rightarrow p + \pi^-$, which offers a branching ratio of 69.2% [4]. Λ particles decayed with an appropriate decay length of $c\tau \approx 79$ mm after they were produced in Au+Au collisions. Protons and pions were identified based on the ionization energy loss in the TPC gas. In practice, charged tracks with $|n\sigma_X| < 3$ for any particle of interest X were selected, where $n\sigma_X$ is defined as follows.

$$n\sigma_X = \frac{1}{\sigma_X} \log \frac{\langle dE/dx \rangle_{\text{measured}}}{\langle dE/dx \rangle_X^{\text{Bischel}}}, \tag{2}$$

where $\langle dE/dx \rangle_{\text{measured}}$ denotes the average energy loss per unit length measured by the TPC of the STAR detector, $\langle dE/dx \rangle_X^{\text{Bischel}}$ represents the expected energy loss $\langle dE/dx \rangle$ for a certain particle species X (in this case, protons or pions), and σ_{particle} denotes the $\langle dE/dx \rangle$ resolution measured by the TPC (typically ≈ 8% [9]). For each proton or pion track, we required a minimum of 15 hits in the TPC to ensure adequate track quality.

Using the data collected in the STAR experiment from Au+Au collisions at $\sqrt{s_{NN}} = 27$ GeV, Λ particles were reconstructed using the KF Particle method, and various kinematic and topological variables such as mass, p_T, and decay length were calculated. As depicted in Fig. 3, clear Λ mass peaks were observed in the invariant mass $m_{p\pi^-}$ distributions in the p_T range of 0.4–0.6 GeV/c from collision events with 0–5% (left panel) and 30–40% (right panel) centrality.

To ensure that the KF Particle method can be reliably used for extracting the physical yields, we applied the KF Particle method to a Monte Carlo simulated sample generated using an embedding technique detailed as follows: Simulated Λ particles with flat p_T and rapid distributions were propagated through a GEANT3 [25] simulation of the STAR TPC. The Λ particles decayed inside the simulated detector and the electronic signals originating from the decay particles were mixed with those from a given event from the real data. The number of simulated Λ particles was 5% of the measured charged-particle multiplicity of the event in which the simulated particles were embedded, and the simulated Λ particles all originated from the primary vertex of that event. The combined electronic signals were subsequently processed using the STAR tracking software, which is used for real data processing as well. Thereafter, the KF Particle package was deployed on the resultant tracks for Λ reconstruction.

We compared the performance of KF particles on real data and MC simulation samples. The topological variables listed below (Table 1) were used to select the Λ candidates during the KF Particle reconstruction.

Statistical criteria were used instead of geometric quantities correspondingly ($DCA_{\text{prim},\pi} \rightarrow \chi^2_{\text{prim},\pi}$, $DCA_{\text{prim},p} \rightarrow \chi^2_{\text{prim},p}$, $DCA_{p-\pi} \rightarrow \chi^2_{p-\pi}$, b and $\theta \rightarrow \chi^2_{\text{topo},\Lambda}$, $d_\Lambda \rightarrow d_\Lambda/\sigma d_\Lambda$). Comparisons of these variables between the data and MC simulation for Λ candidates with $p_T = 0.4 - 1.2$ GeV/c and a centrality between 0 and 10% are depicted in Fig. 4. In general, the distributions of these topological variables from the data are appropriately described by the MC simulations for all centralities and p_T.

To achieve the optimal significance of the Λ signal, the Toolkit for Multivariate data A analysis is used. TMVA is a family of supervised learning algorithms that can be used to differentiate between signals and backgrounds. For further details, please refer to Refs. [20]. Signal and background samples were prepared as inputs for training. The signal samples were obtained from a GEANT3 simulation as described above. For the background sample, we selected sidebands ($3\sigma < |m_{p\pi} - m_{\Lambda,\text{PDG}}| < 6\sigma$) $p\pi$ pairs in the real data around the Λ mass peak, where σ is the width of the Λ mass peak and $m_{p\pi}$ and $m_{\Lambda,\text{PDG}}$ are the masses of the $p\pi$ pair and the Λ baryon from the PDG, respectively. These signal

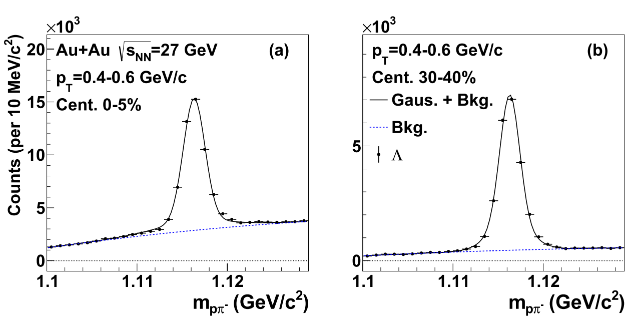

Fig. 3 $p\pi^-$ invariant mass distributions of $p_T = 0.4 - 0.6$ GeV/c in Au+Au collisions at $\sqrt{s_{NN}} = 27$ GeV with centrality 0–5% (left) and 30–40% (right). Black data points depict all $p\pi^-$ pair distributions and are fitted with the Gaussian function in addition to the combinatorial background distributions depicted in the blue lines, which were estimated via side-band fitting

Table 1 Topological variables for Λ reconstruction

Variable	Description
$\chi^2_{\text{prim},\pi}$	χ^2 deviation of π track to the primary vertex
$\chi^2_{\text{prim},p}$	χ^2 deviation of p track to the primary vertex
$\chi^2_{\text{topo},\Lambda}$	χ^2 of primary vertex to the reconstructed Λ
$\chi^2_{p-\pi}$	χ^2 of daughter particle (p-π) fit
d_Λ	decay length of Λ
$d_\Lambda/\sigma d_\Lambda$	decay length normalized by its uncertainty

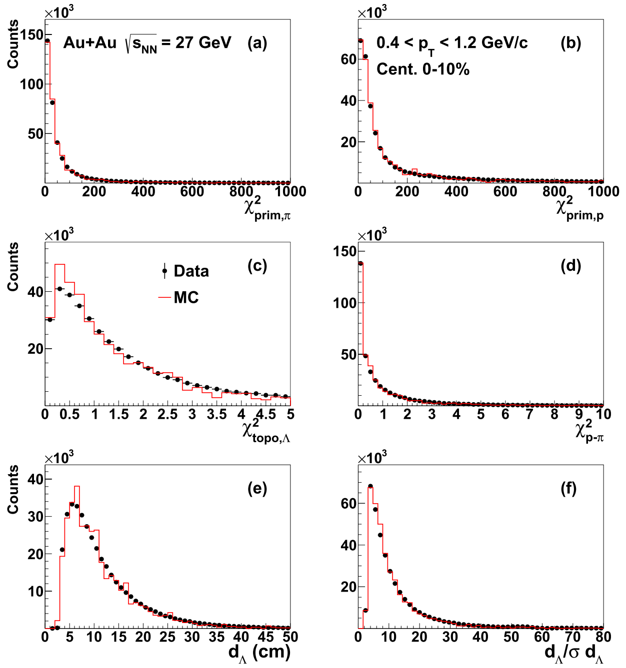

Fig. 4 Key topological variable distributions: **a** $\chi^2_{\text{prim},\pi}$, **b** $\chi^2_{\text{prim},p}$, **c** $\chi^2_{\text{topo},\Lambda}$, **d** $\chi^2_{p-\pi}$, **e** d_Λ, **f** $d_\Lambda/\sigma d_\Lambda$ used in KF Particle method for Λ reconstruction. Data (black points) and MC simulations (red curves) are compared

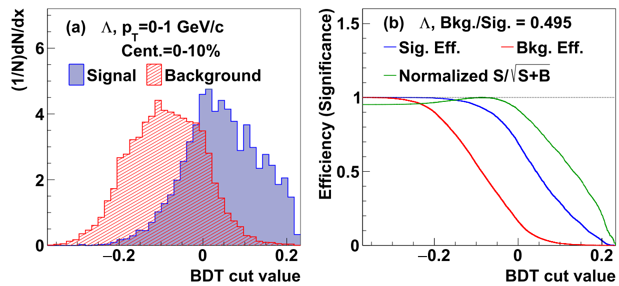

Fig. 5 (Color online) (left) BDT response value distributions for signal (blue) and background (red) Λ candidates in the p_T range 0–1 GeV/c for 0–10% centrality. (right) Efficiency for signal (blue line) and background (red line) Λ candidates in the p_T range 0–1 GeV/c for 0–10% centrality as a function of the cut value placed on the BDT response value. Significance (green line) achieves its maximum value when the cut value is −0.09

and background samples are further divided into different p_T and centrality classes. We used the Boosted Decision Tree method for training. Decision-tree learning uses a set of input features and splits the input data recursively based on these features. In our case, the input features were the topological variables listed in Table 1 and the input data are the signal and background samples depending on these variables. BDT combines multiple decision trees to strengthen the differentiation power; a detailed discussion can be found in Ref. [26]. Training considers the correlations between different topological variables and collapses them into a single value, referred to as the BDT response value.

The BDT response value distributions from the signal and background samples for Λ candidates with $p_T = 0-1\,\text{GeV}/c$ and centrality 0–10% are shown in the left panel of Fig. 5. We observe that the BDT response values for the signal and background are significantly different from each other, and thus serve as a good measure for differentiating between the signal and background. To select a BDT response cut value to optimize the significance $S/\sqrt{S+B}$, where S stands for signal counts and B stands for background counts, we used the TMVA package to calculate the signal and background efficiency as a function of the BDT response cut value, $\varepsilon_S(\text{BDT cut})$ and $\varepsilon_B(\text{BDT cut})$, using the signal and background samples, respectively. The signal and background efficiencies for Λ candidates in the p_T range 0–1 GeV/c and a centrality 0–10% are shown in the right panel of Fig. 5 for blue and red lines. The estimated significance as a function of the BDT cut-off value can then be calculated using Eq. (3):

$$Sig.(\text{BDT cut}) = \frac{S_0\varepsilon_S}{\sqrt{S_0\varepsilon_S + B_0\varepsilon_B}}, \tag{3}$$

where S_0 and B_0 are the number of signal and background counts in the dataset before the BDT cut is applied. S_0 is obtained from the estimated Λ counts without performing any cut on the topological variables, and B_0 is obtained from the number of sideband $p\pi^-$ pairs without the BDT cut. The calculated significance as a function of the cut value applied to the BDT response value for Λ candidates in the p_T range 0–1 GeV/c and 0–10% centrality is shown in the right panel of Fig. 5 as a green line. We find that a cut value of −0.09 maximizes the significance of this particular p_T and centrality bin, and we choose this cut value for Λ reconstruction. This procedure is then repeated for each p_T and the centrality bin. Generally, as the signal-to-background ratio decreases, a stricter BDT selection cut is necessary to optimize the significance.

We extracted the number of signals and background counts for each p_T and the centrality bin using the tuned BDT cuts obtained, as explained above. We then used the standard helix swimming method used in previous STAR analyses [6], tuned the topological cuts in the HS method by the BDT, extracted the corresponding number of signals and background counts using the same procedure, and compared the significance obtained using these two methods. For a fair comparison, the track quality and particle identification cuts were identical. The ratios of significance as functions of p_T for the three centrality selections are shown in Fig. 6. The

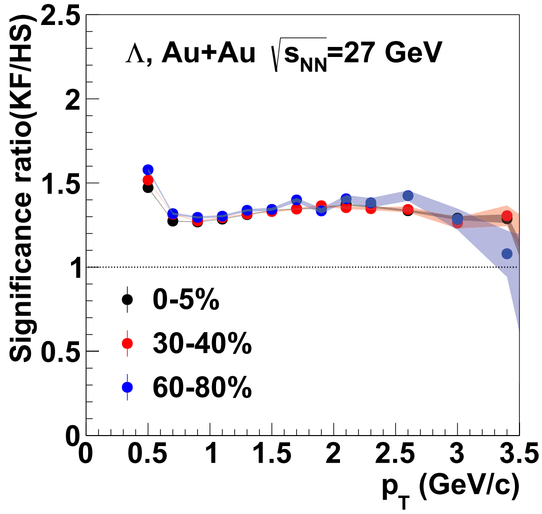

Fig. 6 (Color online) Ratio of significance for Λ particles using the KF Particle method in conjunction with BDT training over those using the helix swimming (HS) method in conjunction with BDT training as a function of p_T of the Λ particles for centrality selection 0–5% (black), 30–40% (red) and 60–80% (blue). The shaded bands indicate the statistical uncertainties

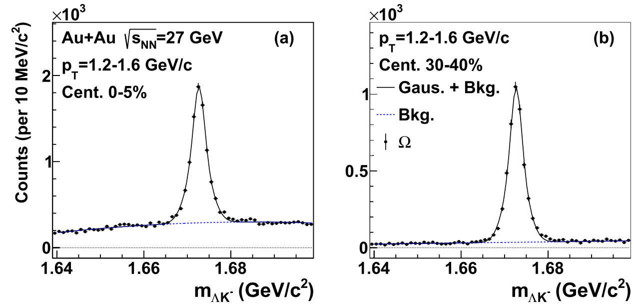

Fig. 7 ΛK^- invariant mass distributions of $p_T = 1.2 - 1.6$ GeV/c in Au+Au collisions at $\sqrt{s_{NN}} = 27$ GeV with centrality 0–5% (left), and 30–40% (right). Black data points depict all ΛK^- pair distributions and are fitted with the Gaussian function plus the combinatorial background distributions shown in the blue lines, which are estimated via side-band fitting

increase in significance is approximately independent of the centrality, $\approx$ 30% in the p_T range 1–3 GeV/c, and increases at low p_T to $\approx$ 50%. This demonstrates that the KF Particle method is more significant for Λ signal extraction in Au+Au collisions at $\sqrt{s_{NN}} = 27$ GeV in the STAR experiments.

3.2 Ω reconstruction

Next, we turn to Ω baryon. Ω baryons were reconstructed using the decay channel $\Omega \rightarrow \Lambda + K^- \rightarrow p + \pi^- + K^-$. Ω particles decayed with a proper decay length of $c\tau \approx 25$ mm [4], and the Λ daughters decayed again soon thereafter. The final daughter tracks were detected using STAR TPC. Similarly, for each proton, kaon or pion track, a minimum of 15 hits were required to ensure good track quality. We reconstruct the Λ baryons with the KF Particle method first and then treat it as a daughter track to reconstruct the Ω production vertex. In Fig. 7, clear Ω mass peaks were observed in the ΛK^- invariant mass distributions using the KF Particle method.

Because the decay topology for Ω baryons is more complicated than that for Λ baryons, more topological variables can be used for training to facilitate the differentiation between the signal and background. The topological variables are listed in Table 2 were used in the selection of Ω baryon candidates during KF Particle reconstruction.

Similar to the Λ baryon study, we generated an MC sample of the reconstructed Ω baryons using a GEANT3 simulation of the STAR TPC. The data-MC comparison of key topological variables is shown in Fig. 8.

We find reasonable agreement between the data and MC simulations, which suggests proper estimation and usage of the covariance matrix of the Λ daughters and gives us confidence that the KF Particle method may be reliably used for the extraction of Ω baryon yields. We then generated a signal and background sample using the same method as in Λ analysis to supply inputs for TMVA training using the BDT method. The BDT response value distribution for Ω candidates with $p_T = 1 - 4$ GeV is shown in the left panel of Fig. 9. The signal efficiency, background efficiency, and significance are shown in the right panel of Fig. 9. As in the case of Λ analysis, we selected the BDT response cut-off value that optimizes significance.

This process is repeated for each p_T and the centrality bin. The significance of using the optimized BDT response cuts for each p_T and centrality bin was extracted. We then performed signal extraction using the default helix swimming method, with candidate selection cuts chosen to be

Table 2 Topological variables for Ω reconstruction

Variable	Description
$\chi^2_{\mathrm{prim},\pi}$	χ^2 deviation of π track to the primary vertex
$\chi^2_{\mathrm{prim},p}$	χ^2 deviation of p track to the primary vertex
$\chi^2_{\mathrm{prim},K}$	χ^2 deviation of K track to the primary vertex
$\chi^2_{\mathrm{topo},\Lambda}$	χ^2 of primary vertex to the reconstructed Λ
$\chi^2_{p-\pi}$	χ^2 of daughter particle (p-π) fit
$\chi^2_{\mathrm{topo},\Omega}$	χ^2 of primary vertex to the reconstructed Ω
$\chi^2_{\Lambda-K}$	χ^2 of daughter particle (Λ-K) fit
d_Λ	decay length of Λ
$d_\Lambda/\sigma_{d_\Lambda}$	Λ decay length normalized by its uncertainty
d_Ω	decay length of Ω
$d_\Omega/\sigma d_\Omega$	Ω decay length normalized by its uncertainty

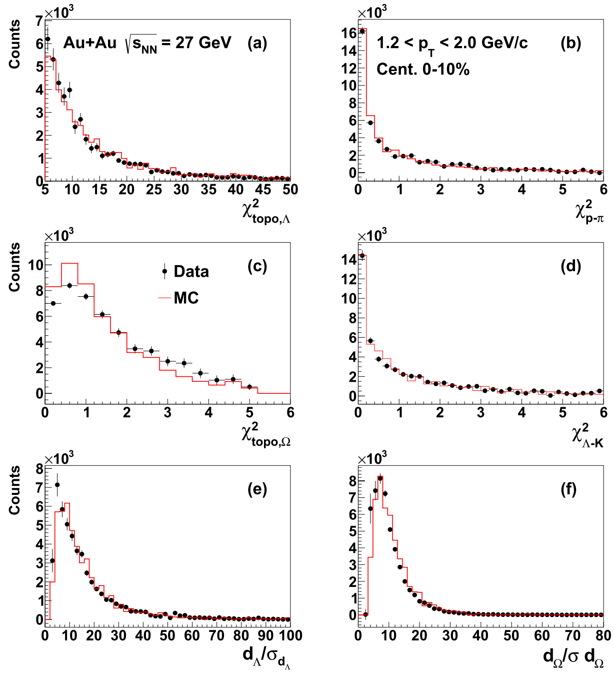

Fig. 8 Key topological variable distributions: **a** $\chi^2_{\mathrm{topo},\Lambda}$, **b** $\chi^2_{p-\pi}$, **c** $\chi^2_{\mathrm{topo},\Omega}$, **d** $\chi^2_{\Lambda-K}$, **e** $d_\Lambda/\sigma_{d_\Lambda}$, **f** $d_\Omega/\sigma d_\Omega$ used in KF Particle method for Ω reconstruction. Data (black points) and MC simulations (red curves) are compared

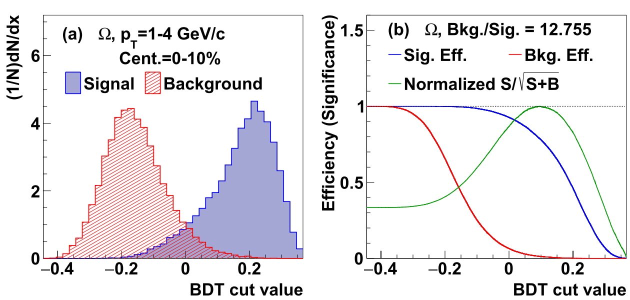

Fig. 9 (Color online) (left) BDT response value distributions for signal (blue) and background (red) Ω candidates in the p_{T} range 1–4 GeV/c for 0–10% centrality. (right) Efficiency for signal (blue line) and background (red line) Ω candidates in the p_{T} range 1–4 GeV/c for 0–10% centrality as a function of the cut value placed on the BDT response value. The significance (green line) achieves its maximum value when the cut value is 0.09

the same as in the previous Ω analyses at the same collision energy [6, 27], also tuned by the BDT method. The signal and background counts were extracted using the default helix swimming method, and the ratios of the significances were calculated using these two methods, as shown in Fig. 10. We observe an ≈ 50% increase in significance in the p_{T} range of 1–4 GeV/c. This increase is higher than that for Λ, likely owing to the more complex

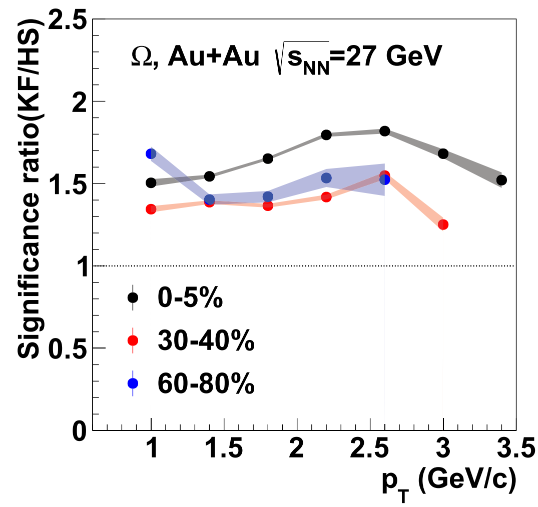

Fig. 10 (Color online) Ratio of significance for Ω particles using the KF Particle method in conjunction with BDT training over those using the helix swimming (HS) method in conjunction with BDT training as a function of p_{T} of the Ω particles for centrality selection 0–5% (black), 30–40% (red) and 60–80% (blue). The shaded bands indicate the statistical uncertainties

decay topology with two decay vertices reconstructed by KF particles and a larger background. Further studies using KF particles are underway to extend the Ω measurement to low p_{T} beyond 1 GeV/c; however, this is beyond the scope of this study.

3.3 D^0 reconstruction

D^0 (and $\bar{D}^0$) particles were reconstructed via the decay channel $D^0 \rightarrow K^-\pi^+$ and its charge conjugation with an appropriate decay length of $c\tau \approx 123\ \mu$m [4]. Because this decay length is less than the spatial resolution of the TPC detector, information from the microvertex detector HFT is used to distinguish the D^0 decay vertex from the primary collision vertex. For each kaon or pion daughter track, a minimum of 15 hits in the TPC and a match with the HFT detector with at least three hits were required to ensure good track quality. For kaon and pion particle identification, in addition to the requirement that $|n\sigma_\pi| < 3$ and $|n\sigma_{\mathrm{K}}| < 2$, we also utilized information from the time-of-flight (TOF) detector [5] when available to help identify the particles at high p_{T} by requiring the measured inverse velocity ($1/\beta$) to be within three standard deviations The topological variables are listed in Table 3 were used to select the D^0 meson candidates in KF Particle reconstruction. $p_{\mathrm{T},\pi}$ and $p_{T,K}$ cuts are added to reject combinatorial background at low p_{T}.

Similar to the Λ and Ω baryon studies, we generated an MC sample of reconstructed D^0 mesons using a GEANT3 simulation of the STAR TPC, HFT, and TOF and

Table 3 Topological variables for D^0 reconstruction

Variable	Description
$\chi^2_{\mathrm{prim},\pi}$	χ^2 deviation of π track to the primary vertex
$p_{T,\pi}$	transverse momentum of π track
$\chi^2_{\mathrm{prim},K}$	χ^2 deviation of K track to the primary vertex
$p_{T,K}$	transverse momentum of K track
$\chi^2_{\mathrm{topo},D^0}$	χ^2 of primary vertex to the reconstructed D^0
$\chi^2_{K,\pi}$	χ^2 of daughter particle (K-π) fit
$L_{D^0}/\sigma_{L_{D^0}}$	D^0 decay length normalized by its uncertainty

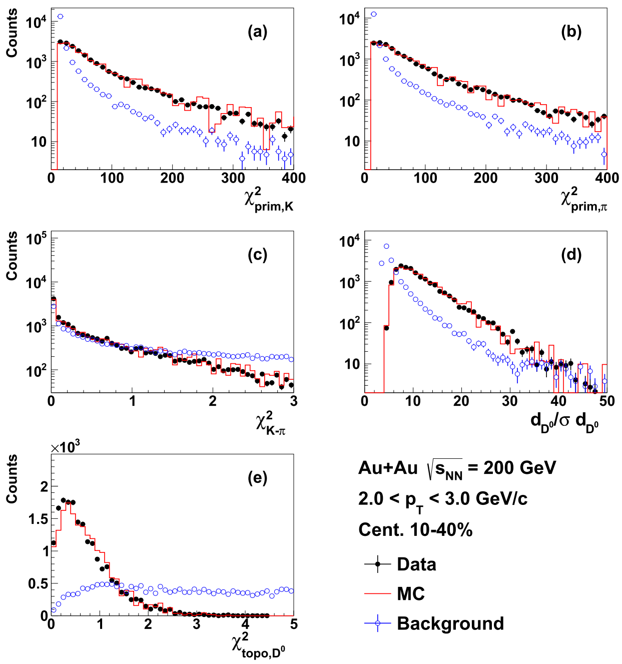

Fig. 11 (Color online) Key topological variables distributions: **a** $\chi^2_{\mathrm{prim},\pi}$, **b** $\chi^2_{\mathrm{prim},K}$, **c** $\chi^2_{K,\pi}$, **d** $d_{D^0}/\sigma_{d_{D^0}}$, **e** $\chi^2_{\mathrm{topo},D^0}$ used in KF Particle method for D^0 reconstruction. Data (black points), MC simulations (red curves), and background (blue circles) are compared

processed it through full detector tracking, as was done in the real data reconstruction with the previously mentioned embedding technique. The HFT simulator was tuned to reproduce the single-track efficiency and DCA pointing resolution observed in real data. However, the consistency in the topological variable distributions between the data and MC for D^0 signals is yet to be demonstrated. Figure 11 shows a comparison of several key topological variables used in the KF Particle method for D^0 reconstruction between the data (black data points) and MC (red histograms). We found good agreement between the data and MC simulations for these variables, which means that this multiple-detector-combined MC simulation can generate

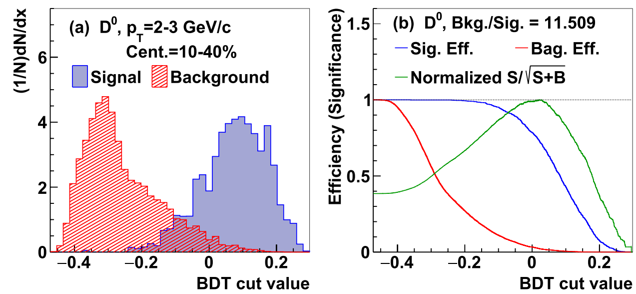

Fig. 12 (Color online) (left) BDT response value distributions for signal (blue) and background (red) D^0 candidates in the p_T range 2–3 GeV/c for 10–40% centrality. (right) Efficiency for signal (blue line) and background (red line) D^0 candidates in the p_T range 2–3 GeV/c for 10–40% centrality as a function of the cut value placed on the BDT response value. The significance (green line) achieves its maximum value when the cut value is 0.05

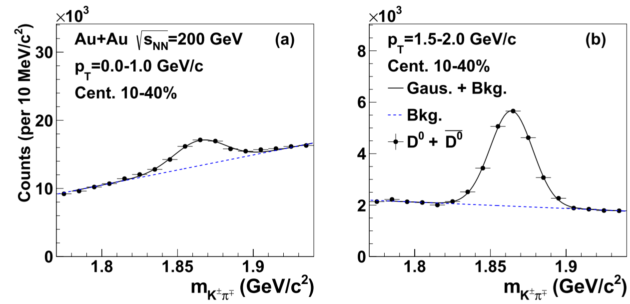

Fig. 13 $K^\pm\pi^\mp$ invariant mass distributions using the KF Particle method in 10–40% Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV in the region of $p_T = 0-1$ GeV/c (left) and $p_T = 1.5-2.0$ GeV/c (right). Black data points depict all $K^\pm\pi^\mp$ pair distributions and are fitted with the Gaussian function plus the combinatorial background distributions shown in the blue lines, which are estimated via side-band fitting

D^0 signals reasonably. The background distributions are shown in Fig. 11 (blue circles)). They are estimated from real data using the sideband method, in which background candidates are selected by requiring an invariant mass of $K\pi$ pairs within $3\sigma < |M_{\mathrm{inv}} - M_{D^0}| < 6\sigma$ (σ is the Gaussian width of the D^0 signal). The signal and background behaviors can be distinguished well, particularly from $d_{D^0}/\sigma_{d_{D^0}}$ and $\chi^2_{\mathrm{topo},D^0}$ distributions.

Thereafter, we used the signal sample generated from the MC simulation and the background sample from the sideband candidates in the data to conduct TMVA training with the BDT method to determine the topological selection working point for the best signal significance. Figure 12 left panel exhibits the BDT response value distributions for the D^0 signal and background in the region of $p_T = 2-3$ GeV/c with a centrality 10–40%; the efficiencies of the signal and background are also shown in the

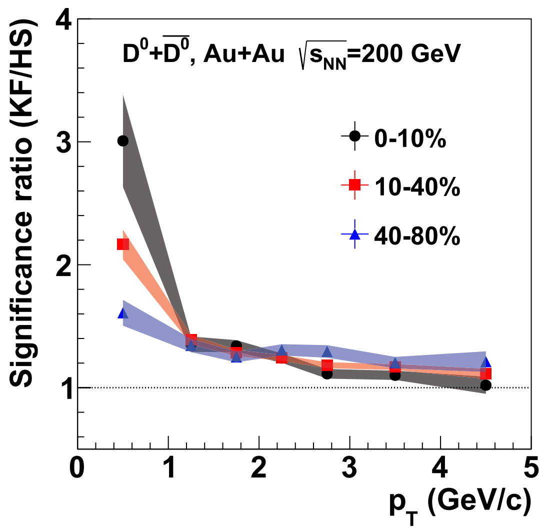

Fig. 14 (Color online) Ratio of significance for D^0 particles using the KF Particle method in conjunction with BDT training over those using the helix swimming(HS) method in conjunction with BDT training as a function of p_T of the D^0 particles for centrality selection 0–10% (black circles), 10–40% (red squares), and 40–80% (blue triangles). Shaded bands indicate the statistical uncertainties

right panel. The significance was normalized to the maximum value. We determine the BDT response cut value to optimize the significance of D^0 for each p_T and centrality class.

Thereafter, we applied the optimized BDT selection cuts to real data analysis. Figure 13 displays the D^0-invariant mass distributions derived using the KF Particle method for 10–40% Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV in the regions $p_T = 0 - 1$ GeV/c (left) and $p_T = 1.5 - 2.0$ GeV/c (right), respectively. Black lines depict the function fits of the data with a Gaussian function for the D^0 signal and linear background. Compared with the Λ and Ω reconstructions, more combinatorial backgrounds are included because the D^0 decay points are closer to the collision points and blend with the particles from them, especially at $p_T = 0 - 1$ GeV/c.

Thereafter, the signal significance was calculated from the invariant mass distributions of D^0 candidates. The signal counts were obtained from a Gaussian function fit of the D^0 peak, whereas the background counts were determined based on a linear background function fit within a mass window of $|M_{inv} - M_{D^0}| < 3\sigma$, where σ denotes the width of D^0 peak. We compared the significance of D^0 from the KF Particle method to the helix swimming (HS) method used in a previous analysis [5] which was in conjunction with BDT training; the ratio of the D^0 significance between these two methods is illustrated in Fig. 14. The shaded bands indicate the statistical uncertainties in this calculation. This comparison demonstrates that the KF Particle method improves the reconstructed D^0 signal significance, especially for low p_T and more central collisions. In 0–10% of the central Au+Au collisions at $p_T < 1$ GeV/c, the improvement can signify as a factor of $\approx$3, potentially because of the enormous amount of combinatorial background (hundreds of signals) in that particular range, and the cuts based on statistical criteria work appropriately in discriminating the particles originating from a secondary vertex.

4 Summary

In summary, we applied the KF Particle method to reconstruct Λ, Ω^- hyperons, and D^0 mesons in the STAR experiment. The KF Particle method, which utilizes covariant matrices of tracking parameters, improves the reconstructed Λ (Ω) significance by approximately 30% (50%) compared with the traditional helix swimming method in $\sqrt{s_{NN}} = 27$ GeV Au+Au collisions. The improvement in D^0 significance by applying the KF Particle method has a p_T dependence at $\sqrt{s_{NN}} = 200$ GeV Au+Au collisions, with the largest improvement as significant as a factor of $\approx$3 in $p_T < 1$ GeV/c and 0–10% central collisions. The present findings demonstrated that the Monte Carlo simulation can reproduce the topological variable distributions used in the KF Particle method, thereby establishing KF particles as a robust method for analyzing the strange and open charm hadrons in the STAR experiment. As the KF Particle method is independent of the detector geometry, it can be applied to other experiments, especially in analyses with a small signal-to-background ratio [28–31].

Acknowledgements The authors thank the STAR Collaboration, RHIC Operations Group, RCF at BNL, and NERSC Center at LBNL for their support.

Author Contributions All authors contributed to the study conception and design. Material preparation, and analysis were performed by Xin-Yue Ju and Yue-Hang Leung. Data collection are performed by the RHIC-STAR collaboration and the Monte Carlo simulation thanks to Sooraj Radhakrishnann and Xiang-Lei Zhu. The first draft of the manuscript was written by Xin-Yue Ju,Yue-Hang Leung, Xin Dong and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability The data that support the findings of this study are openly available in Science Data Bank at https://www.doi.org/10.57760/sciencedb.j00186.00250 and https://cstr.cn/31253.11.sciencedb.j00186.00250.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

1. P. Koch, B. Muller, J. Rafelski, Strangeness in relativistic heavy ion collisions. Phys. Rep. **142**, 167–262 (1986). https://doi.org/10.1016/0370-1573(86)90096-7
2. S. Frixione, M.L. Mangano, P. Nason et al., Heavy quark production. Adv. Ser. Direct. High Energy Phys. **15**, 609–706 (1998). https://doi.org/10.1142/9789812812667_0009
3. X. Dong, Y.-J. Lee, R. Rapp, Open heavy-flavor production in heavy-ion collisions. Ann. Rev. Nucl. Part. Sci. **69**, 417–445 (2019). https://doi.org/10.1146/annurev-nucl-101918-023806
4. P.A. Zyla, R.M. Barnett, J. Beringer et al., Review of particle physics. Prog. Theor. Exp. Phys. **2020**, 083C01 (2020). https://doi.org/10.1093/ptep/ptaa104
5. J. Adam, L. Adamczyk, J.R. Adams et al., Centrality and transverse momentum dependence of D^0-meson production at mid-rapidity in Au+Au collisions at $\sqrt{s_{\rm NN}}$ = 200 GeV. Phys. Rev. C **99**(3), 034908 (2019). https://doi.org/10.1103/PhysRevC.99.034908
6. J. Adam, L. Adamczyk, J.R. Adams et al., Strange hadron production in Au+Au collisions at $\sqrt{s_{\rm NN}}$ =7.7, 11.5, 19.6, 27, and 39 GeV. Phys. Rev. C **102**(3), 034909 (2020). https://doi.org/10.1103/PhysRevC.102.034909
7. M.S. Abdallah, B.E. Aboona, J. Adam et al., Measurements of H^3_Λ and H^4_Λ lifetimes and yields in Au+Au collisions in the high-baryon density region. Phys. Rev. Lett. **128**(20), 202301 (2022). https://doi.org/10.1103/PhysRevLett.128.202301
8. K. Ackermann, N. Adams, C. Adler et al., STAR detector overview. Nucl. Instrum. Meth. A **499**, 624–632 (2003). https://doi.org/10.1016/S0168-9002(02)01960-5
9. 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
10. G. Contin, L. Greiner, J. Schambach et al., The STAR MAPS-based PiXeL detector. Nucl. Instrum. Meth. A **907**, 60–80 (2018). https://doi.org/10.1016/j.nima.2018.03.003
11. L. Adamczyk, J.K. Adkins, G. Agakishiev et al., Measurement of D^0 azimuthal anisotropy at midrapidity in Au+Au collisions at $\sqrt{s_{NN}}$=200 GeV. Phys. Rev. Lett. **118**(21), 212301 (2017). https://doi.org/10.1103/PhysRevLett.118.212301
12. J. Adam, L. Adamczyk, J.R. Adams et al., First measurement of Λ_c baryon production in Au+Au collisions at $\sqrt{s_{\rm NN}}$ = 200 GeV. Phys. Rev. Lett. **124**(17), 172301 (2020). https://doi.org/10.1103/PhysRevLett.124.172301
13. J. Adam, L. Adamczyk, J.R. Adams et al., Observation of $D_s^\pm/D^0$ enhancement in the Au+Au collisions at $\sqrt{s_{_{NN}}}$ = 200 GeV. Phys. Rev. Lett. **127**, 092301 (2021). https://doi.org/10.1103/PhysRevLett.127.092301
14. 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
15. X. Luo, S. Shi, N. Xu et al., A study of the properties of the QCD phase diagram in high-energy nuclear collisions. Particles **3**(2), 278–307 (2020). https://doi.org/10.3390/particles3020022
16. L. Adamczyk, J.K. Adkins, G. Agakishiev et al., Observation of D^0 meson nuclear modifications in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. Phys. Rev. Lett. **113**(14), 142301 (2014). https://doi.org/10.1103/PhysRevLett.113.142301
17. M.S. Abdallah, B.E. Aboona, J. Adam et al., Evidence of mass ordering of charm and bottom quark energy loss in Au+Au collisions at RHIC. Eur. Phys. J. C **82**(12), 1150 (2022). https://doi.org/10.1140/epjc/s10052-022-11003-7
18. F. Si, X. Chen, L. Zhou et al., Charm and beauty isolation from heavy flavor decay electrons in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV at RHIC. Phys. Lett. B **805**, 135465 (2020). https://doi.org/10.1016/j.physletb.2020.135465
19. D. Li, F. Si, Y. Zhao et al., Charm and beauty isolation from heavy flavor decay electrons in p+p and Pb+Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV at LHC. Phys. Lett. B **832**, 137249 (2022). https://doi.org/10.1016/j.physletb.2022.137249
20. H. Voss, A. Höcker, J. Stelzer et al., TMVA, the Toolkit for Multivariate Data Analysis with ROOT. PoS **ACAT**, 040 (2009). https://doi.org/10.22323/1.050.0040
21. R.E. Kalman, A new approach to linear filtering and prediction problems. J. Basic Eng. **82**(1), 35–45 (1960). https://doi.org/10.1115/1.3662552
22. R. Frühwirth, et al., Data Analysis Techniques for High-Energy Physics, 2nd Ed. Cambridge (2000)
23. S. Gorbunov, On-line reconstruction algorithms for the CBM and ALICE experiments. PhD. Thesis (2013). https://nbn-resolving.org/urn:nbn:de:hebis:30:3-295385
24. M. Zyzak, Online selection of short-lived particles on many-core computer architectures in the CBM experiment at FAIR. PhD. Thesis (2016). https://nbn-resolving.org/urn:nbn:de:hebis:30:3-414288
25. R. Brun, F. Bruyant, M. Maire et al., GEANT 3: user's guide Geant 3.10, Geant 3.11; rev. version. CERN, Geneva (1987). https://cds.cern.ch/record/1119728
26. H. Drucker, C. Cortes, Boosting decision trees. In Proceedings of the 8th International Conference on Neural Information Processing Systems Vol. 8, pp. 479-485 (1995)
27. L. Adamczyk, J.K. Adkins, G. Agakishiev et al., Probing parton dynamics of QCD matter with Ω and ϕ production. Phys. Rev. C **93**(2), 021903 (2016). https://doi.org/10.1103/PhysRevC.93.021903
28. R. Ralf, Bottomonium suppression in heavy-ion collisions and the in-medium strong force. Nucl. Sci. Tech. **34**, 63 (2023). https://doi.org/10.1007/s41365-023-01213-3
29. Y. Ma, Hypernuclei as a laboratory to test hyperon-nuucleon interactions. Nucl. Sci. Tech. **34**(6), 97 (2023). https://doi.org/10.1007/s41365-023-01248-6
30. N. Li, Z. Sun, X. Liu et al., Perfect DD^* molecular prediction matching the T_{cc} observation at LHCb. Chin. Phys. Lett. **38**(9), 092001 (2021). https://doi.org/10.1088/0256-307X/38/9/092001
31. R. Aaij, Search for the doubly charmed baryon Ω^+_{cc}. Sci. China Phys. Mech. **64**(10), 101062 (2021). https://doi.org/10.1007/s11433-021-1742-7