Simulation study of particle identification using cluster counting technique for the BESIII drift chamber

Shuiting Xin, Guang Zhao, Linghui Wu, Mingyi Dong, Gang Li, Shengsen Sun and Xinchou Lou

Institute of High Energy Physics,
19B, Yuquan Road, Shijing District, Beijing 100049, China
University of Chinese Academy of Sciences (CAS),
19A Yuquan Road, Shijing District, Beijing 100049, China
Physics Department, University of Texas at Dallas,
Richardson TX, U.S.A.

E-mail: zhaog@ihep.ac.cn

ABSTRACT: The particle identification of charged hadrons, especially for the separation of $K$ and $\pi$, is crucial for the flavour physics study. Ionization measurement with the cluster counting technique, which has a much less fluctuation than traditional $dE/dx$ measurement, is expected to provide better particle identification for the BESIII experiment. Simulation studies, including a Garfield++ based waveform analysis and a performance study of $K/\pi$ identification in the BESIII offline software system have been performed. The results show that $K/\pi$ separation power and PID efficiency would be improved appreciably in the momentum range above 1.2 GeV/c using cluster counting technique even with a conservative resolution assumption.

KEYWORDS: Particle identification methods; Gaseous detectors; Wire chambers (MWPC, Thin-gap chambers, drift chambers, drift tubes, proportional chambers etc)

ArXiv ePrint: 2210.01845

*Corresponding author.
1 Introduction

The Beijing Spectrometer (BESIII) [1] detector at the Beijing Electron Positron Collider (BEPCII) is a general purpose detector designed for studies of hadron physics and \( \tau \)-charm physics. It has collected data of around 40 fb\(^{-1}\) at different center-of-mass energies from 2 to 5 GeV since 2009. To consider further physics opportunities and extend the physics potential of BESIII, an upgrade of BEPCII was initiated [2]. For this upgrade, high luminosity and increased beam energy were proposed and discussed. Hence, excellent performance of particle identification (PID) is essential for both high precision measurements and new physics searches in the future program.

The current PID of BESIII is based on the momentum and ionization energy loss (\( dE/dx \)) measurements by the multilayer drift chamber (MDC) and the time of flight measurement by the time-of-flight (TOF) counters. The \( dE/dx \) resolution is about 6\% for minimum-ionizing particles (MIP) \( \pi \) [3]. The time resolution of TOF is 68 ps in the barrel region and 60 ps in the end cap region [4, 5]. Given the need for a highly accurate charged particle identification, the cluster counting (\( dN/dx \)) technique [6], an alternative method of ionization measurement, can be a choice for the BESIII drift chamber.

Cluster counting has been demonstrated in experiments for several decades [6]. When a charged particle passes through a drift chamber, a sequence of clusters of one or more ion-electron pairs emerge along the track. The Landau distribution with an infinite long tail describes the fluctuations of energy loss. On the other hand, the primary ionization itself allows a preferable understanding of the ionization behavior of gas, because the Poisson nature of the number of clusters provides a smaller uncertainty than a Landau distribution. Therefore, counting the primary clusters can conceivably reach an intrinsic resolution of the ionization process, and offer an improved particle separation capability. In recent years, the development of electronics technology facilitates realization of the cluster counting method. The technique was reviewed in ref. [8]. It was validated with single-cell prototype drift chambers at TRIUMF using test beam data [9]. For some future \( e^+e^- \) experiments, cluster counting has been proposed for the IDEA drift chamber [10] and is also being studied in the CEPC drift chamber [11].
In this paper, we report on a Monte Carlo (MC) study of applying the cluster counting technique for PID in the BESIII drift chamber. To estimate the $dN/dx$ resolution, a Garfield++ based simulation is carried out. A further study of $K/\pi^+$ identification capability is practiced in the BESIII offline software system (BOSS). This paper proceeds as follows: section 2 introduces the simulation study using the Garfield++ software [12]. Section 3 presents the MC study in BOSS [13]. Section 4 gives a brief conclusion.

2 Garfield based simulation

A two-step simulation work is performed in this study. In order to probe the potential of cluster counting technique, a theoretical estimation of ionization measurement is firstly obtained from Garfield++ simulation on ionization process. The study of $dE/dx$ and $dN/dx$ behaviors is illustrated in section 2.1. For the second step, the resolution degradation of $dN/dx$ is considered in the simulation of waveform processing. The waveform analysis is described in section 2.2.

2.1 Ionization simulation for $dE/dx$ and $dN/dx$

The ionization simulation is to model the energy loss and ionization clusters by charged particles in an MDC. The MDC measures momentum and energy loss along the charged particle trajectory ($dE/dx$ method) to derive the mass of charged particle. To achieve the cluster counting technique, the peaks associated with the ionization clusters need to be identified in the induced signal waveform with a peak finding algorithm. A typical simulated signal waveform with identified peaks is shown in figure 1.

![Figure 1. A typical waveform of induced signal in a time window of 500 nanosecond (ns). The red triangles are identified peaks.](image)

The simulation of $dE/dx$ and $dN/dx$ for MDC is carried out with the Garfield++ program and interfaced to Heed. The program Heed, an implementation of the photon-absorption ionization (PAI) model simulates the ionization produced along the track of charged particles [15].

---

1 Throughout this paper, the terms "$K^-$", "$\pi^-$" usually indicate $K^-$, $\pi^-$. 

---
In the simulation, the configuration of detector geometry and material of MDC are implemented in the Garfield++ program. The drift chamber consists of 43 cylindrical layers, the cell size is around 12 mm × 12 mm for the inner eight layers and 16.2 mm × 16.2 mm for the outer layers. The chamber is filled with a helium based mixture He-C$_3$H$_8$ 60:40 and operates in a 1T magnetic field [1]. Figure 2 shows the $dN/dx$ and $dE/dx$ distributions of 2 GeV/c π in one cell, respectively. As expected, the energy loss distribution, by recording the true energy deposition, is peaked with a long tail, as shown in figure 2(a), where a fit is performed using a Landau convolving a Gaussian function. The δ electrons received enough energy from the incident particle ionize other atoms and result in such events in the tail of energy loss distribution. While the $dN/dx$ distribution is a symmetric Gaussian in 2(b), independent of that extra ionization process.

![Figure 2. $dE/dx$ and $dN/dx$ distribution of 2 GeV/c π in one cell at truth ionization level, where the fit function for $dE/dx$ is a Landau convolving a Gaussian.](image)

For a track that passes through the chamber, to reduce the impact of the long Landau tail, $dE/dx$ is obtained using the truncated mean method, where each cell keeps hits within 5% to 75% range of charge integral. The $dE/dx$ of a track considering all MDC cells is displayed in figure 3(a). The $dN/dx$ is averaged over hits from all cells of a track, as shown in figure 3(b). The $dE/dx$ resolution defined as $\sigma_{dE/dx}$ is about 5.6% at this momentum. While $dN/dx$ resolution defined as $\sigma_{dN/dx}$ reaches a value below 3%, which is a factor of 2 better than $dE/dx$.

Figure 4 shows the evolution of those two quantities along the momentum of π and K. $dN/dx$ apparently has narrow bands of K and π than $dE/dx$, giving a clear trend on separation. Apart from that, the intersection point of the π and K curves around 1.1 GeV/c infers that two particle species are indistinguishable at that momentum. As a consequence, neither energy loss measurement nor cluster counting is effective at all in the region overlapped within 1σ band. In this case, the identification of the two particle species has to be achieved with the TOF.
Figure 3. $dE/dx$ (left) and $dN/dx$ (right) distribution of 2 GeV/c $\pi$ (blue) and $K$ (red) for one track through 43 layers at truth ionization level. The track length is around 66 cm.

Figure 4. $dE/dx$ and $dN/dx$ versus momentum of $\pi$ and $K$. The track length is 66 cm. The error bands are drawn with $1\sigma$ and $2\sigma$ uncertainty from a Gaussian fit.

2.2 $dN/dx$ simulation based on waveform processing

The critical content of cluster counting technique implementation is the waveform processing. It is composed of the waveform digitization of signal produced by individual ionization and waveform analysis with suitable algorithm applied to detect peaks associated with clusters.

As described in section 2.1, the drift and avalanche of electron-ion pairs are simulated for events in one cell. Collecting all the charges from the electrodes, the induced current is shaped as a sequence of pulses. The induced signal shown in figure 1 is simulated with the Garfield++ program. For the simulation used by the current study, electronic noise and shaping time are not considered. The sampling rate is 2 Gsamples/s. The single electron diffusion is around a few ns.
Figure 5. Performance of cluster counting in waveform analysis. (a) The found clusters against truth clusters; (b) found clusters and truth clusters distribution for single cell events, with resolution of 16.1%, 19.6%, respectively; (c) ratio of found clusters over truth clusters. The average ratio is 0.91.

A peak finding algorithm named TSpectrum, provided by ROOT [14], is adapted for identification of peaks. Figure 5(a) shows the number of found clusters against the truth clusters from the TSpectrum algorithm for a statistic of one thousand $\pi$ events. Figure 5(b) shows the distributions of truth and found clusters from the waveform analysis. Compared with the truth distribution, the average ratio of found clusters over truth clusters is about 91%, in other words, the mean value decreases by 10%, and the sigma value is degraded from 16% to 19.6% for a single cell. The ratio of found clusters over truth clusters by event is presented in 5(c). The $dE/dx$ of an event can be reconstructed from the charge integral of the waveform. Assuming each cell has the same measured $dN/dx$ and $dE/dx$ distributions, it can be shown that the reconstructed $dN/dx$ and $dE/dx$ resolution of waveform analysis model is 2.93% and 5.92%, respectively, for 77.4 cm (43 layers with 1.8 cm cell size) track length. The same truncation range, 5% to 75%, introduced in section 2.1 is applied to the $dE/dx$ calculation in the extrapolation.

The main reason for the inefficiency of counting is high cluster density. For the gas mixture of He-C$_3$H$_8$ 60:40, the cluster density of the primary ionization is 26 per cm for MIP $\pi$ and 29 per cm for 2 GeV/c $\pi$. As the pileup region shown in figure 1, some clusters are very close in time on the waveform due to the high cluster density and the wide longitudinal diffusion, which makes peak finding difficult, and leads to a worse resolution.
There are more factors that cause the degradation of the resolution in the experiment: 1) The sampling rate of the electronics has substantial impacts on the digitization process so as on the peak finding efficiency. 2) Noise makes miscounting from the actual peaks. 3) Secondary ionization causes overestimation of the primary counts. In order to estimate degradation considering all those effects, 30% and 60% degradations of $dN/dx$ resolution are proposed.

A PID performance study is carried out with the $dN/dx$ considering the resolution degradation. To quantify the PID performance, the separation power $S_{K,\pi}^M$ between $K$ and $\pi$ of measurement $M$ is introduced, where $M$ refers to either $dE/dx$ or $dN/dx$.

$$S_{K,\pi}^M = \frac{|M_K - M_\pi|}{\sqrt{\sigma(M_K)^2 + \sigma(M_\pi)^2}}$$

(2.1)

In the equation, $M_K$ and $M_\pi$ indicate the average value of the distribution of measurement $M$, $\sigma(M_K)$ and $\sigma(M_\pi)$ are the corresponding standard deviation of the measurement $M$ for $K$ and $\pi$, respectively. The separation power as a function of momentum is drawn in figure 6 for $dE/dx$ and $dN/dx$. The $K/\pi$ separation with cluster counting surpasses energy loss measurement for the whole range of momentum, except for momenta of particle from 0.9 GeV/c to 1.2 GeV/c, which could be recovered by the TOF detector. It is found that the resolution of $dN/dx$ is the crucial point in evaluating PID performance. A gain of 250% is seen from the separation using the truth $dN/dx$ information. In the most degradation scenario, $dN/dx$ still delivers an advantage over $dE/dx$ in the high momenta region, which offers an increase of about 170%.

Figure 6. Separation power of $\pi/K$ as a function of momentum for truth $dN/dx$, 30% degradation of $\sigma_{dN/dx}$, 30% degradation of $\sigma_{dN/dx}$ and reconstructed $dE/dx$ based on charge integration of waveform.
3 Monte-Carlo study in the BESIII offline software system

In this section, a MC study is performed in the BESIII offline software system (BOSS). The BOSS utilizes full offline data processing including detector simulation, reconstruction and calibration of sub-detectors and analysis toolkit. A $dN/dx$ model based on parameterization method is implemented in the analysis toolkit of BOSS. The $dN/dx$ estimation is modeled from the previous simulation at track level. The predicted cluster counts are expressed as a function of $\beta\gamma$ of the charged track. $dN/dx$ is calculated for each track with two particle assumptions of $\pi, K$ according to their $\beta\gamma$ for the next step of statistic calculation. $dE/dx$ is obtained with the BOSS algorithm after hit and track level correction.

In order to apply the PID analysis, probability variables are calculated for each particle hypothesis. Depending on the combination of inputs, $M$ could be $dN/dx$, $dE/dx$ or TOF. The procedure for the calculation of particle probability is following:

1. For each particle hypothesis, we calculate a statistic variable $\chi^M_{\text{par}} (\text{par} = \pi, K)$ for measurement $M$, with expected mean value and deviation: $\chi^M_{\text{par}} = \frac{M_{\text{par}} - M_{\text{par,exp}}}{\sigma_{M_{\text{par}}}}$

2. Find the sum of square of $\chi^M_{\text{par}}$, it can be used to combine information: $\chi^2_{\text{par}} = \sum^n_M (\chi^M_{\text{par}})^2$, where $n$ is number of degrees of freedom (ndof), denotes the number of measurements.

3. Derive the probability of each particle hypothesis for a certain $\chi^2_{\text{par}}$, $P_{\text{par}} (\chi^2_{\text{par}}|n)$, which denotes the upper tail probability of Chi-squared distribution of degree $n$.

The particle type is determined by the hypothesis with the highest probability. For instance, a charged track is identified as $K$ if $P_K > P_\pi$, and vice versa. Consequently, the PID efficiency of a particle $\text{par}$ ($K$ or $\pi$) can be defined,

$$\epsilon_{\text{par}} = \frac{N_{\text{par} \rightarrow \text{par}}}{N_{\text{par}}}, \quad (3.1)$$

where $N_{\text{par}}$ is the total number of generated $\text{par}$, $N_{\text{par} \rightarrow \text{par}}$ is the number of correctly identified $\text{par}$.

Single particle events of $K$ and $\pi$ are simulated for PID efficiency study. The Monte Carlo samples are produced with GEANT4 [16], which include the geometry description of the BESIII detector and the detector response. Only events having exactly one good quality track are selected, events with more than one reconstructed tracks are removed. Charged tracks detected in MDC are required to be within a polar angle ($\theta$) range of $|\cos \theta| < 0.93$, where $\theta$ is defined with respect to the $z$-axis in MDC.

The PID efficiencies of $K$ and $\pi$ are displayed in figure 7, there are four different cases considered: $dE/dx$ only, $dN/dx$ only, $dE/dx + \text{TOF}$ and $dN/dx + \text{TOF}$. The combinations with TOF are performed using the Chi-squared based approach in the previous paragraph with ndof = 2. Comparing the results with and without TOF, the TOF information helps ameliorate PID efficiency from 50% to around 90% at momentum around 1.1 GeV/$c$, but contributes little in high momentum range. The combined $dN/dx$ and TOF information provides the best PID efficiency for $K$ and $\pi$, denoted by the $dN/dx$ curves in figure 7(b), 7(d). Compared to the efficiency of $dN/dx$ and $dE/dx$ for $K$ without TOF in figure 7(a), the one with cluster counting of a 30% degradation is 20% superior to efficiency of energy loss method. With a 60% degradation of $dN/dx$, a gain over 10% is still possible even in a conservative estimation which considers a peak finding efficiency in noisy background.
Another common figure of merit to evaluate the performance of a binary classifier is the receiver operating characteristic (ROC). Larger area under ROC curve indicates a better performance of a classifier. One can also extract information of signal efficiency for a given background rejection level from the ROC curve. Figure 8 plots the $K$ selection efficiency versus $\pi$ rejection efficiency obtained by cutting on $\pi$ probability, for $dE/dx$ and $dN/dx$. The possibilities without TOF and combined with TOF are depicted on the left and right, respectively. The ROC curves of PID methods show that combined $dN/dx$ and TOF information provides the best PID performance. We provide benchmarks with $\pi$ rejection efficiency at 70\%, 90\% and 99\%, shown as orange dash lines. Consequently, with a 90\% of $\pi$ rejection in without TOF scenario, an over 80\% $K$ selection efficiency is expected using the truth $dN/dx$ information, increased by at least 20\% with respect to the $dE/dx$ method.

4 Conclusion

This proof-of-concept study demonstrates the profit of cluster counting technique for the BESIII Helium-based drift chamber. Specifically, we utilize Garfield++ to generate waveforms of the drift chamber signal, different assumptions on $dN/dx$ resolution are applied to evaluate PID performance. The separation power between $\pi$ and $K$ with cluster counting foresees a promising improvement
Figure 8. ROC curves made by $K$ selection efficiency versus $K$ selection efficiency. Left: $dN/dx$ and $dE/dx$ without TOF; right: $dN/dx$ and $dE/dx$ with TOF. Orange dash line indicates $\pi$ rejection efficiencies at 70%, 90% and 99% respectively.

over traditional energy loss measurement in the range of momentum above 1.2 GeV/c, gains of a factor of 2.5 and 1.7 are expected in the theoretical case and in the conservative case within a 60% degradation of $dN/dx$ resolution, respectively. Furthermore, a $dN/dx$ model is parameterized in the BOSS analysis toolkit. The PID efficiency combining the drift chamber and the time-of-flight counters information is evaluated using Monte Carlo simulation. The results show that the PID efficiency of $\pi$ and $K$ can be enhanced by over 20% in the theoretical case. To achieve an excellent PID capability delivered by cluster counting technique, high performance of front-end readout electronics with low noise is necessary to resolve signal pulses from different ionization clusters. Aside from the simulation study, prototype tests are expected to validate the feasibility of cluster counting technique.

Acknowledgments

This research was funded by National Key R&D Program of China under Contracts Nos. 2020YFA0406304, National Natural Science Foundation of China (NSFC) under Contract Nos. 12275296, U1832204, 11521505 and National 1000 Talents Program of China.

References

[1] BESIII collaboration, A. Buzulutskov, ed., The construction of the BESIII experiment, Nucl. Instrum. Meth. A 598 (2009) 7.
[2] BESIII collaboration, Future Physics Programme of BESIII, Chin. Phys. C 44 (2020) 040001 [arXiv:1912.05983].
[3] X.-X. Cao et al., Studies of $dE/dx$ measurements with the BESIII, Chin. Phys. C 34 (2010) 1852.
[4] S. Sun, Time calibration for barrel TOF system of BESIII, J. Phys. Conf. Ser. 396 (2012) 022051.
[5] Y.-X. Guo, S.-S. Sun, F.-F. An, R.-X. Yang, M. Zhou, Z. Wu et al., The study of time calibration for upgraded end cap TOF of BESIII, *Radiat. Detect. Technol. Meth.* 1 (2017) 15.

[6] A.H. Walenta, The Time Expansion Chamber and Single Ionization Cluster Measurement, *IEEE Trans. Nucl. Sci.* 26 (1979) 73.

[7] F. Lapique and F. Piuz, Simulation of the measurement by primary cluster counting of the energy lost by a relativistic ionizing particle in argon, *Nucl. Instrum. Meth.* 175 (1980) 297.

[8] J. Vavra, PID techniques: Alternatives to RICH Methods, *Nucl. Instrum. Meth. A* 639 (2011) 193.

[9] J.-F. Caron et al., Improved Particle Identification Using Cluster Counting in a Full-Length Drift Chamber Prototype, *Nucl. Instrum. Meth. A* 735 (2014) 169 [arXiv:1307.8101].

[10] FCC collaboration, *FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2*, *Eur. Phys. J. ST* 228 (2019) 261.

[11] CEPC Study Group, *CEPC Conceptual Design Report: Volume 2 — Physics & Detector*, IHEP-CEPC-DR-2018-02 (2018) [arXiv:1811.10545].

[12] Garfield++, https://garfieldpp.web.cern.ch/garfieldpp.

[13] Z.Y. Deng et al., Object-Oriented BESIII detector simulation system, *Chin. Phys. C* 30 (2006) 371.

[14] R. Brun and F. Rademakers, ROOT: An object oriented data analysis framework, *Nucl. Instrum. Meth. A* 389 (1997) 81 [Proceedings of the AIHENP’96 Workshop, Lausanne, Switzerland, 2–6 September 1996].

[15] D. Pfeiffer et al., Interfacing Geant4, Garfield++ and Degrad for the Simulation of Gaseous Detectors, *Nucl. Instrum. Meth. A* 935 (2019) 121 [arXiv:1806.05880].

[16] GEANT4 collaboration, GEANT4 — a simulation toolkit, *Nucl. Instrum. Meth. A* 506 (2003) 250.