Nuclear Science and Techniques (2025) 36:218
https://doi.org/10.1007/s41365-025-01789-y

DATA ARTICLE

Experimental dataset from BESIII detector at Beijing electron–positron collider

Ming-Hua Liao[1] · Jian-Shu Liu[1] · Xin-Nan Wang[2] · Sheng-Sen Sun[2,3] · Zheng-Yun You[1]

Received: 1 April 2025 / Revised: 8 May 2025 / Accepted: 19 June 2025 / Published online: 12 September 2025

Abstract
In the BESIII detector at Beijing electron–positron collider, billions of events from e^+e^- collisions were recorded. These events passing through the trigger system were saved in raw data format files. They play an important role in the study of physics in τ-charm energy region. Here, we published an e^+e^- collision dataset containing both Monte Carlo simulation samples and real data collected by the BESIII detector. The data pass through the detector trigger system, file format conversion, and physics information extraction and finally save the physics information and detector response in text format files. This dataset is publicly available and is intended to provide interested scientists and those outside of the BESIII collaboration with event information from BESIII, which can be used to understand physics research in e^+e^- collisions, developing visualization projects for physics education, public outreach, and science advocacy.

Keywords Electron–positron collision · BESIII · Data sharing · Education · Visualization

1 Introduction

The Beijing Electron–Positron Collider (BEPC) [1] has been in operation since 1989 and was upgraded to BEPCII in 2008. The BEPCII was designed to operate in τ-charm energy region. It is a double-ring multi-beam electron–positron collider with a circumference of 237.5 m, beam energy ranging from 1.0 to 2.45 GeV, and luminosity of 1×10^{33} cm^{-2}s^{-1}. BEPCII has 93 bunches arranged at 8 ns or 2.4 m intervals in each ring. The electron and positron beams collided at a horizontal crossing angle of ± 11 mrad at the intersection.

Beijing Spectrometer III (BESIII) [2–4], a spectrometer running at BEPCII, can accurately record the information of the final state particles from e^+e^- collisions. This information was recorded using a two-level trigger system [5]. Level-1 (L1) is hardware-based, and level-3 (L3) is software-based, known as the event filter, which operates on an online computer farm. The L1 trigger aims to efficiently identify high-quality physics events and reduce cosmic rays and beam-related background noise. Data stored in the front-end buffers are retrieved once an L1 accept signal is received, approximately 6.4 ms after a collision. Subsequently, the data were sent to an online computer farm for event processing. The event stream undergoes filtering to further suppress background events through the L3 trigger before saving to the permanent storage device in the raw file format.

The BESIII Offline Software System (BOSS) [6, 7], which was developed using C++ language and object-oriented techniques, uses the event data service of Gaudi as a data manager to read raw data information. Reconstruction algorithms [8–10] can retrieve information from raw data and reconstruct physical information, which will be stored for the subsequent physics analysis.

This work was supported by the National Key Research and Development Program of China (No. 2023YFA1606000), National Natural Science Foundation of China (Nos. 12175321, 11975021, and U1932101), and National College Students Science and Technology Innovation Project of Sun Yat-sen University.

✉ Zheng-Yun You
youzhy5@mail.sysu.edu.cn

[1] School of Physics, Sun Yat-Sen University, Guangzhou 510275, China

[2] Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

[3] University of Chinese Academy of Sciences, Beijing 100049, China

Open datasets of experiments, particularly from large scientific facilities, play an important role in developing new techniques, physics education, public outreach, and science advocacy [11–13]. For example, experiments at the Large Hadron Collider (LHC) have released data for unsupervised new physics detection [13]. Open datasets are made available in a format that is easy to observe and access directly, allowing users to selectively use specific data to achieve their goals, such as extracting the momentum and initial position of charged particle tracks.

In this study, we prepared datasets based on the BESIII experiment. The main purpose is to provide a small fraction of massive data events collected by BESIII. The dataset can be used to understand the physics of e^+e^- collisions and to develop event visualization techniques by displaying physics events collected by the BESIII detector. Visualization can help physicists deepen their understanding of some physics processes and has been applied in optimizing the corresponding physics analyses [14–21]. Additionally, the dataset can effectively support physicists in public outreach and scientific advocacy. The dataset contains some typical physics processes in e^+e^- collisions in the τ-charm energy region, with approximately 100 events for every process. The data are stored in text format, including the detector hit responses and the reconstructed particle information for each event, making it easy for the public to access directly.

This remainder of this study is structured as follows: Sect. 2 introduces the BESIII detector and data conversion for the open dataset. Section 3 describes the physical processes and related information in the dataset. Section 4 presents the technical validation, usage notes, and code availability.

2 Experimental design, materials and methods

The structure of the BESIII detector is described as follows: The cylindrical core consists of a helium-based multilayer drift chamber (MDC), a plastic scintillator time-of-flight system (TOF), and a CsI(Tl) electromagnetic calorimeter (EMC), which are all enclosed in a superconducting solenoidal magnet (SSM) providing a 1.0 T magnetic field. The solenoid is supported by an octagonal flux-return yoke with a resistive plate chamber muon identifier module (MUC) interleaved with steel. The acceptance of charged particles and photons was 93% over 4π solid angle. The charged particle momentum resolution at 1 GeV/c was 0.5%, and the dE/dx resolution was 6% for electrons from the Bhabha scattering. The EMC measured photon energies with a resolution of 2.5% (5%) at 1 GeV in the barrel (end-cap) region. The time resolution of the TOF barrel region was 68 ps, whereas that of the end cap was 110 ps. A visualization of the BESIII detector is shown in Fig. 1. A more detailed description of the BESIII detector is available in Ref. [2].

Specifications Table	
Subject	Particles Physics
Specific subject area	Electron–Positron collider experiment.
Data format	Raw&Analyzed
Type of data	Text format file
How data were acquired	Measurements by the BESIII detector and processed with BESIII offline software.
Parameters for data collection	Electron–positron collisions at different energy points.
Description of data collection	Data were collected by the BESIII detector and reconstructed for physics analysis.
Data collection	The data were collected from electron–positron collisions at BEPCII using the BESIII detector.
Data source location	Institution: Institute of High Energy Physics, Chinese Academy of Sciences
	Country: China
Data accessibility	Repository name: Science Data Bank
	Data identification number: https://cstr.cn/31253.11.sciencedb.21486
	Direct URL to data: https://doi.org/10.57760/sciencedb.21486
Related research article	Z. J. Li, et al. Front. Phys. **19**, 64201 (2024). doi:10.1007/s11467-024-1422-7

The open dataset was provided in text format, and the data conversion flow is shown in Fig. 2. First, it originates from the raw files of real e^+e^- collisions or ROOT [22] raw (RTRAW) files generated from BESIII simulated physics processes. These events pass through the trigger system and are saved on a disk [5]. Second, the raw information can

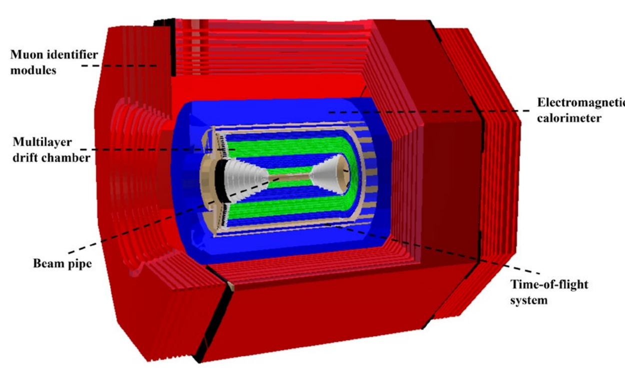

Fig. 1 (Color online) The BESIII detector

be converted into physical information through reconstruction and saved in reconstruction (REC) format files containing both raw and reconstructed information. These REC files can be viewed effectively using the visualization software BesVis [23] provided by BESIII. Third, the event information in the REC files was used to extract physical information with the analysis package, such as the momentum of charged particles, energy deposit in the EMC, and particle penetration length in the MUC. This information was saved in ROOT files. Next, some classic and interesting events can be displayed using BesVis, with their event and run number recorded, allowing the selection of good events to be saved in a new filtered ROOT file. Finally, the event information in the filtered ROOT file was converted into text format for the open dataset.

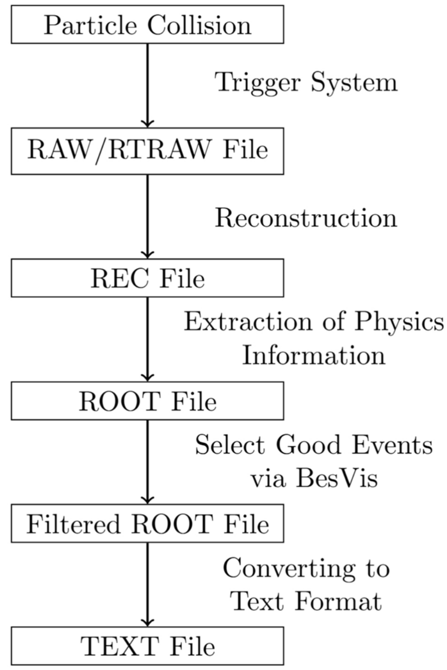

Fig. 2 The data conversion flow of BESIII event files

3 Data records

The dataset can be accessed via the repository listed in the specification table. Table 1 lists the physical processes, selected number of events, and data types included in the dataset. The dataset consisted of two parts: one based on Monte Carlo (MC) simulation samples using the BESIII offline software and the other based on real data collected by the BESIII detector. The events in the dataset were saved in a text format with hit and track information from the detector response. The content of the dataset is introduced in the following sections.

Because some classical physics processes are submerged in huge background events in real data, it is not easy to select them directly. However, the behavior of MC events is close to real data from event displays and can help us determine whether rare physics events are signals. The MC samples include three major decay modes with large branching ratios for the ground state of vector charmonium J/ψ, such as the electromagnetic decay $J/\psi \to e^+e^-/\mu^+\mu^-$, and the hadronic decay $J/\psi \to \pi^+\pi^-\pi^0$ [24], in which $J/\psi \to e^+e^-$ is the decay channel where the charm quark was discovered in 1974 [25, 26]. Additionally, the decay modes of charmed meson D, such as $D^0 \to K^-\pi^+, \bar{D}^0 \to K^+\pi^-, D^+ \to K^-\pi^+\pi^+$, and $D^- \to K^+\pi^-\pi^-$ [27, 28], are provided, as they are the main single-tag reconstruction channels in D meson studies widely used in charm physics analysis. Correspondingly, the main decay channel of the lightest charmed baryon Λ_c^+, $\Lambda_c^+ \to pK^-\pi^+$ [29], is provided. The dataset also contains the semi-leptonic process $\Lambda_c^+ \to ne^+\nu_e$, and the $\bar{\Lambda}_c^-$ is reconstructed using some main decay channels [30]. The neutrons in the final states have a clear signature in the EMC with a distinct hit distribution. This physics process has also been applied in neutron identification studies at BESIII based on deep learning techniques [30].

Table 1 The physics processes, number of events, and data types in the dataset

S.N	Physics processes	Events	Types
1	$J/\psi \to e^+e^-$	128	MC
2	$J/\psi \to \mu^+\mu^-$	126	MC
3	$J/\psi \to \pi^+\pi^-\pi^0$	155	MC
4	$D^0\bar{D}^0,\ D^0 \to K^-\pi^+, \bar{D}^0 \to K^+\pi^-$	122	MC
5	$D^+D^-,\ D^+ \to K^-\pi^+\pi^+,\ D^- \to K^+\pi^-\pi^-$	140	MC
6	$\Lambda_c^+\bar{\Lambda}_c^-,\ \Lambda_c^+ \to pK^-\pi^+,\ \bar{\Lambda}_c^- \to \bar{p}K^+\pi^-$	113	MC
7	$\Lambda_c^+ \to ne^+\nu_e$	104	MC
8	$e^+e^- \to \pi^+\pi^-J/\psi$ ($Z_c(3900)$)	10	Data
9	$e^+e^- \to e^+e^-$	141	Data
10	$e^+e^- \to \mu^+\mu^-$	121	Data
11	$e^+e^- \to \gamma\gamma$	150	Data
12	3097 MeV	150	Data
13	3686 MeV	150	Data
14	3773 MeV	144	Data
15	Cosmic ray	10	Data

The BESIII detector collected a large amount of real data in the τ-charm energy region, with the highest integrated luminosity at energy points of 3097 MeV, 3686 MeV, and 3773 MeV, which correspond to the production thresholds of J/ψ, $\psi(3686)$, and $D\bar{D}$, respectively. The dataset includes real data at these energy points, with approximately 150 events at each energy point, as well as cosmic ray backgrounds and quantum electrodynamics (QED) processes, such as $e^+e^- \rightarrow e^+e^-/\mu^+\mu^-/\gamma\gamma$. Additionally, a flagship discovery from real data, the four-quark state $Z_c(3900)$ decay events [31, 32], was also provided in the dataset. For some events in the detector, the particle trajectories appear to resemble the shape of ψ, which are widely used in outreach and journal cover images [33].

The filenames in the dataset follow a style that includes the data type, physics process, and recorded information, such as "Data_Zc3900_track_hit.txt". Each text file contains detailed reconstructed track and hit information for every event and type setting sequentially as follows: run number, event number, MDC track, MDC hit, TOF track, TOF hit, EMC track, EMC hit, MUC track, and MUC hit. Tracks and hits contain a variety of digital information, such as momentum, energy, and detector identifiers. The definitions of digits in text files are stored in the specification document in the dataset named "Parameters_definition.txt". Users can refer to the documents for detailed data information.

4 Recommended repositories to store and find data

With approval from the BESIII collaboration, the dataset has been made public via the Science Data Bank with the link: https://doi.org/10.57760/sciencedb.21486. Users can obtain additional information from the BESIII official website (http://bes3.ihep.ac.cn/).

5 Technical validation

The dataset records a few important physical processes, as well as major background events. They have some remarkable features in terms of visualization. Herein, we present various visualization techniques to demonstrate the validity of the dataset.

The physical processes in the dataset can be visualized using BesVis, as shown in Fig. 3. It has an X-Y cross sectional view of the BESIII detector, where the tracks and hits can be intuitively displayed.

In addition to the two-dimensional (2D) display, a three-dimensional display was also provided using OpenGL, as shown in Fig. 4. This demonstrates a $\Lambda_c^+ \rightarrow ne^+\nu_e$ event, which contains four charged tracks and a clear cluster of EMC hits from neutron energy deposits.

Moreover, new visualization techniques have been developed to display these physical processes. For example, unity-based event displays have been used in BESIII [14, 34], JUNO [17], and other experiments [21]. It has the features of local running, multi-platform deployment, and better visualization effects. Figure 5 shows the 2D display of a $J/\psi \rightarrow \mu^+\mu^-$ event, where the muons penetrate the MUC, and a $Z_c(3900)$ event, which shows the approximate ψ-shaped tracks.

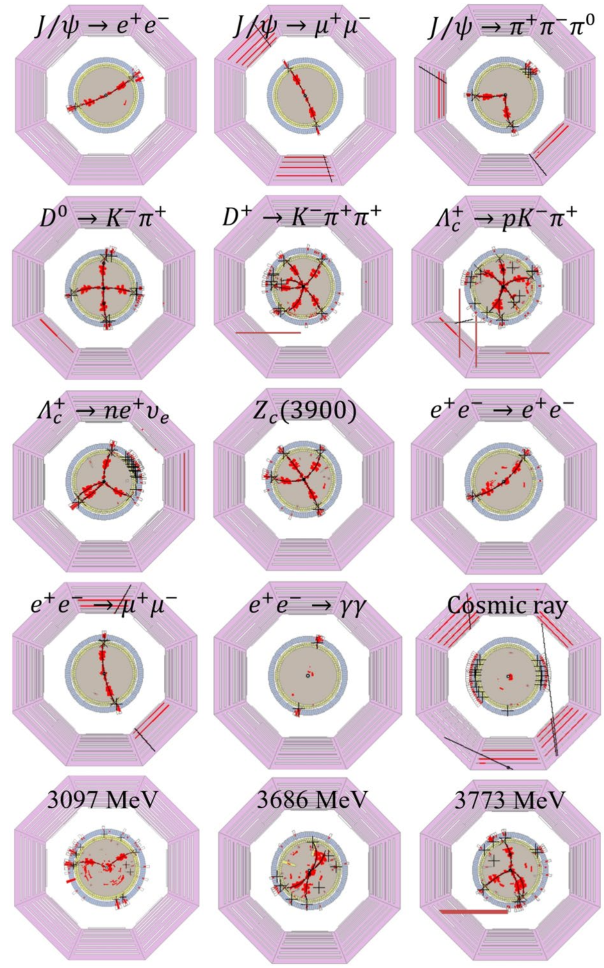

Fig. 3 (Color online) Various physics processes displayed with BesVis in X-Y view. From the inside out, the subdetectors are MDC, TOF, EMC, and MUC

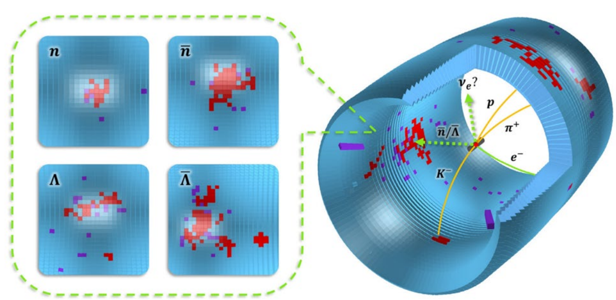

Fig. 4 (Color online) Visualization of a $\Lambda_c^+ \rightarrow ne^+\nu_e$ event using OpenGL [30]. The blue blocks represent the EMC crystals, and the red blocks indicate the hits of particles with energy deposited in the EMC

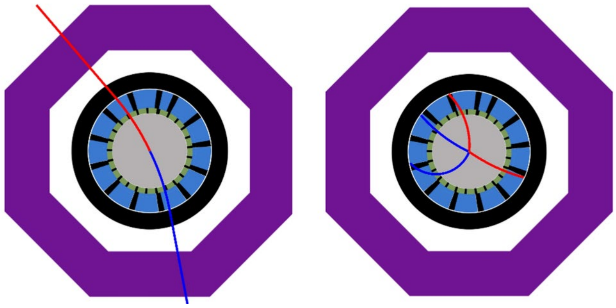

Fig. 5 (Color online) Unity-based visualization of a $J/\psi \to \mu^+\mu^-$ event and a $Z_c(3900)$ event. The red and blue curves represent positively and negatively charged particle tracks, respectively. From the inside out, the subdetectors are MDC, TOF, EMC, SSM, and MUC

6 Usage notes

We propose three potential applications with this dataset.

1) The real data contain not only J/ψ, $\psi(3686)$, and $D\bar{D}$ production events, but also continuum backgrounds including QED processes, cosmic ray backgrounds, and other events. The dataset can be used to study the characteristics of different processes.
2) The dataset has full detector response information at the hit level and track level. Both MC simulation data and real data are provided. It can be used to study the detector response in simulation and track reconstruction algorithms for development of data processing techniques in offline software.
3) Visualization plays a crucial role in every aspect of high energy physics experiments. The dataset describes full event information in e^+e^- collisions. Combined with the BESIII detector geometry, it can be used to develop event display software with various new visualization techniques, not only for physics researches, but also for physics education and public outreach.

Author Contributions S.S. Sun and Z.Y. You conceived the experiment, M.H. Liao and X.N. Wang conducted the experiment, and M.H. Liao and J.S. Liu analyzed the results. All authors have reviewed the manuscript.

Code availability The dataset was publicly available and could be accessed without restrictions. The authors encourage the use of this dataset for visualization research, education, and public outreach. The code is hosted in a public repository to ensure transparency and facilitate community contributions.

Declarations

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

References

1. M.H. Ye, Z.P. Zheng, BEPC, the Beijing electron positron collider. Int. J. Mod. Phys. A **2**, 1707–1725 (1987). https://doi.org/10.1142/S0217751X87000880
2. M. Ablikim et al., [BESIII], Design and construction of the BESIII detector. Nucl. Instrum. Meth. A **614**, 345–399 (2010). https://doi.org/10.1016/j.nima.2009.12.050
3. D.M. Asner, T. Barnes, J.M. Bian et al., Physics at BES-III. Int. J. Mod. Phys. A **24**, S1-794 (2009). https://doi.org/10.48550/arXiv.0809.1869
4. M. Ablikim, M.N. Achasov, P. Adlarson et al., [BESIII], Future physics programme of BESIII. Chin. Phys. C **44**, 040001 (2020). https://doi.org/10.1088/1674-1137/44/4/040001
5. Z.A. Liu, W. Gong, Y. Guo et al., Trigger system of BESIII. 2007 15th IEEE-NPSS Real-Time Conference, Batavia, IL, USA, pp. 1-4, (2007). https://doi.org/10.1109/RTC.2007.4382859
6. J. Zou, W. Li, Q. Ma et al., Offline data processing system of the BESIII experiment. Eur. Phys. J. C **84**, 937 (2024). https://doi.org/10.1140/epjc/s10052-024-13241-3
7. https://lhcb-comp.web.cern.ch/Frameworks/Gaudi/
8. Y.T. Liang, K. Liu, Z.Y. You et al., Track reconstruction in the BESIII muon counter. Chin. Phys. C **33**, 666–672 (2009). https://doi.org/10.1088/1674-1137/33/8/013
9. L.K. Jia, Z.P. Mao, W.D. Li et al., Study of low momentum track reconstruction for the BESIII main drift chamber. Chin. Phys. C **34**, 1866–1873 (2010). https://doi.org/10.1088/1674-1137/34/12/014
10. Y. Zhang, J. Zhang, Y. Yuan et al., Track reconstruction using the hough transform at BESIII. PoS ICHEP2018 888 (2019). https://doi.org/10.22323/1.340.0888
11. G. Guerrieri [ATLAS], Open data at ATLAS: Bringing TeV collisions to the World. [arXiv:2502.21133 [hep-ex]]. https://doi.org/10.48550/arXiv.2502.21133
12. Y. Elangovan, B. Satyanarayana, R. Shinde et al., Design and development of portable RPC-based cosmic muon tracker. [arXiv:2503.01764 [hep-ex]]. https://doi.org/10.48550/arXiv.2503.01764
13. E. Govorkova, E. Puljak, T. Aarrestad et al., LHC physics dataset for unsupervised new physics detection at 40 MHz. Sci. Data **9**, 118 (2022). https://doi.org/10.1038/s41597-022-01187-8
14. K.X. Huang, Z.J. Li, Z. Qian et al., Method for detector description transformation to Unity and application in BESIII. Nucl. Sci. Tech. **33**, 142 (2022). https://doi.org/10.1007/s41365-022-01133-8
15. Z.J. Li, M.K. Yuan, Y.X. Song et al., Visualization for physics analysis improvement and applications in BESIII. Front. Phys. **19**, 64201 (2024). https://doi.org/10.1007/s11467-024-1422-7
16. Z. You, K. Li, Y. Zhang et al., A ROOT based event display software for JUNO. JINST **13**, T02002 (2018). https://doi.org/10.1088/1748-0221/13/02/T02002
17. J. Zhu, Z. You, Y. Zhang et al., A method of detector and event visualization with Unity in JUNO. JINST **14**, T01007 (2019). https://doi.org/10.1088/1748-0221/14/01/T01007
18. M.H. Liao, K.X. Huang, Y.M. Zhang et al., A ROOT-based detector geometry and event visualization system for JUNO-TAO. Nucl. Sci. Tech. **36**, 39 (2025). https://doi.org/10.1007/s41365-024-01604-0
19. R.M. Bianchi et al., [ATLAS], Virtual reality and game engines for interactive data visualization and event displays in HEP, an example from the ATLAS experiment. EPJ Web Conf. **214**, 02013 (2019). https://doi.org/10.1051/epjconf/201921402013
20. A.M. Tadel et al., [CMS], EVE-7 and FireworksWeb: the next generation event visualization tools for ROOT and CMS. EPJ Web Conf. **245**, 08027 (2020). https://doi.org/10.1051/epjconf/202024508027

21. Z.Y. Yuan, Z. Yuan, T.Z. Song et al., Method for detector description conversion from DD4hep to Filmbox. Nucl. Sci. Tech. **35**, 146 (2024). https://doi.org/10.1007/s41365-024-01506-1
22. R. Brun, F. Rademakers, ROOT: an object oriented data analysis framework. Nucl. Instrum. Meth. A **389**, 81–86 (1997). https://doi.org/10.1016/S0168-9002(97)00048-X
23. S. Huang, Z. You, Update of the BESIII event display system. J. Phys. Conf. Ser. **1085**, 042027 (2018). https://doi.org/10.1088/1742-6596/1085/4/042027
24. S. Navas et al., [Particle Data Group], Review of particle physics. Phys. Rev. D **110**, 030001 (2024). https://doi.org/10.1103/PhysRevD.110.030001
25. J.J. Aubert et al., [E598], Experimental observation of a heavy particle J. Phys. Rev. Lett. **33**, 1404–1406 (1974). https://doi.org/10.1103/PhysRevLett.33.1404
26. C. Bacci, R.B. Celio, M. Berna-Rodini et al., Preliminary result of Frascati (ADONE) on the nature of a new 3.1-GeV particle produced in e+ e- annihilation. Phys. Rev. Lett. **33**, 1408 (1974). [erratum: Phys. Rev. Lett. **33**, 1649 (1974)]. https://doi.org/10.1103/PhysRevLett.33.1408.
27. M. Ablikim et al., [BESIII], Measurements of absolute branching fractions for D mesons decays into two pseudoscalar mesons. Phys. Rev. D **97**, 072004 (2018). https://doi.org/10.1103/PhysRevD.97.072004
28. G. Bonvicini et al., [CLEO], Updated measurements of absolute D^+ and D^0 hadronic branching fractions and $\sigma(e^+e^- \to D\bar{D})$ at $E_{\rm cm}$ = 3774 MeV. Phys. Rev. D **89**, 072002 (2014). https://doi.org/10.1103/PhysRevD.89.072002
29. M. Ablikim et al., [BESIII], Measurements of absolute hadronic branching fractions of Λ_c^+ baryon. Phys. Rev. Lett. **116**, 052001 (2016). https://doi.org/10.1103/PhysRevLett.116.052001
30. M. Ablikim et al., [BESIII], Observation of a rare beta decay of the charmed baryon with a graph neural network. Nature Commun. **16**, 681 (2025). https://doi.org/10.1038/s41467-024-55042-y
31. M. Ablikim et al., [BESIII], Observation of a charged charmoniumlike structure in $e^+e^- \to \pi^+\pi^- J/\psi$ at $\sqrt{s}$ =4.26 GeV. Phys. Rev. Lett. **110**, 252001 (2013). https://doi.org/10.1103/PhysRevLett.110.252001
32. Z.Q. Liu et al., [Belle], Study of $e^+e^- \to \pi^+\pi^- J/\psi$ and observation of a charged charmoniumlike state at Belle. Phys. Rev. Lett. **110**, 252002 (2013). https://doi.org/10.1103/PhysRevLett.110.252002
33. G. Huang et al., [BESIII], Probing the internal structure of baryons. Natl. Sci. Rev. **8**, nwab187 (2021). https://doi.org/10.1093/nsr/nwab187
34. J. Li, Z. You, Event visualization with unity in BESIII experiment. PoS ICHEP2024, 1047 (2025). https://doi.org/10.22323/1.476.1047