Nuclear Science and Techniques (2025) 36:224
https://doi.org/10.1007/s41365-025-01799-w

DATA ARTICLE

Experimental dataset from CDEX's high-purity germanium detectors in China Jinping Underground Laboratory

Li-Tao Yang[1] · Zhen-Yu Zhang[1] · Hao Ma[1] · Qian Yue[1] · Zhi Zeng[1]

Received: 7 January 2025 / Revised: 19 March 2025 / Accepted: 10 April 2025 / Published online: 20 September 2025

Abstract
Founded in 2009, the China Dark Matter Experiment (CDEX) collaboration was dedicated to the detection of dark matter (DM) and neutrinoless double beta decay using high-purity germanium (HPGe) detectors in the China Jinping Underground Laboratory. HPGe detectors are characterized by a high energy resolution, low analysis threshold, and low radioactive background, making them an ideal platform for the direct detection of DM. Over the years, CDEX has accumulated a massive amount of experimental data, based on which various results on DM detection and neutrinoless double beta decay have been presented. Because the dataset was collected in a low-background environment, apart from the analysis of DM-related physical channels, it has great potential as an indicator in other rare physical events searches. Furthermore, by providing raw pulse shapes, the dataset can serve as a tool for effectively understanding the internal mechanisms of HPGe detectors.

Keywords Low-background experiment · Pulse shapes · Raw data · HPGe detectors · CDEX

1 Introduction

The China Dark Matter Experiment (CDEX) collaboration was established in 2009. Using high-purity germanium (HPGe) detectors, CDEX aims to detect Dark Matter (DM) and neutrinoless double beta decay ($0\nu\beta\beta$). The experiment was conducted at the China Jinping Underground Laboratory (CJPL), which is currently the world's deepest underground laboratory, and a cosmic-ray muon flux of 61.7 $\mathrm{yr}_{-1}\mathrm{m}^{-2}$ is achieved [1, 2]. With a low cosmic-ray flux, the overall radioactive background can be maintained at a very low level for experiments conducted inside the CJPL, which helps increase detector sensitivity, and makes the detection of rare physical events possible.

Among the different DM detection techniques, CDEX has adopted the high-purity germanium (HPGe) detector technique. HPGe detectors are characterized by their low radioactive background, low energy threshold (160 eVee, where "eVee" represents electron equivalent energy derived from a charge calibration), and excellent energy resolution. These advantages make HPGe detectors an ideal platform for DM and $0\nu\beta\beta$ detection. To date, CDEX has achieved many competitive results on several physical channels.

Two initial CDEX stages, CDEX-1 and CDEX-10, were already completed. CDEX is now moving forward to its next stages, CDEX-50 and CDEX-300, to achieve better sensitivity for DM and $0\nu\beta\beta$ detection by upscaling the mass of the involved germanium detector system. Herein, the dataset accumulated from CDEX-1 and CDEX-10 [3] is presented for possible reuse by interested parties or individuals for studies on HPGe detectors or rare physical events detection.

The CDEX-1 experiment using a 1-kg-scale single-element p-type point-contact germanium (pPCGe) detector proceeded in two stages, CDEX-1A and CDEX-1B. In both stages, the germanium detector was cooled by a cold finger connected to a 30-L dewar filled with liquid nitrogen (LN_2). Then, toward a future ton-scale DM search experiment, the second-generation CDEX experiment (CDEX-10) was conducted with a total detector-crystal mass of approximately 10 kg. The CDEX-10 experiment comprised three encapsulated detector strings labeled as C10-A, C10-B, and C10-C,

This work was supported by the National Key Research and Development Program of China (Nos. 2023YFA1607100 and 2022YFA1605000) and the National Natural Science Foundation of China (No. 12322511).

✉ Zhi Zeng
zengzhi@tsinghua.edu.cn

1 Department of Engineering Physics, Tsinghua University, Beijing 100084, China

each comprising three point-contact germanium detectors encapsulated in a copper vacuum tube and directly immersed in LN_2. In this manner, the previous high-Z shielding of the germanium crystals was replaced by low-Z LN_2, thereby providing an integrated shielding and cooling system that effectively suppresses the radioactive background.

The final spectra used for the physical analysis have already been published in previous CDEX studies. The final spectrum was obtained after the data processing pipeline developed by the CDEX collaboration, which comprised (1) removal of abnormal data to ensure quality, (2) information extraction from raw pulse shapes collected by the data acquisition (DAQ) system, (3) removal of abnormal events, (4) energy calibration, (5) physical/noise discrimination, and (6) efficiency corrections and determination of uncertainties. However, the raw data before data analysis have never been disclosed.

After the analysis procedure, the final energy spectrum only provides information on the radioactive background measured by the CDEX experimental system, whereas the abundant information embedded in the raw data collected by the experimental system is eliminated. By disclosing the original data, CDEX collaboration hopes that interested parties or individuals can benefit from three potential aspects: (1) analysis of other rare physical events, (2) study of the radioactive environment of CJPL, and (3) study of the HPGe detectors' internal mechanisms.

2 Experimental Design, Materials and Methods

The experimental setup for the CDEX-10 is outlined in this section. The CDEX-10 system resides in a polyethylene room with 1-m-thick walls at the CJPL. The entire system comprises an LN_2 tank made of stainless steel and integrated with copper shielding, a 10-kg point-contact germanium detector array, and a DAQ system. CDEX-10 relies on passive shielding to reduce the radioactive background. From outside to inside, the passive shielding system comprises 2400 m of rock overburden, a 1-m-thick polyethylene layer, and 20-cm-thick high-purity oxygen-free high-conductivity copper immersed in the LN_2 surrounding the detector array. The HPGe detectors were directly immersed in LN_2 and maintained at a stable operating temperature near the LN_2 temperature [4].

The data originated from the signals generated by the detectors and were further collected by the DAQ system. The signals from the pPCGe detectors were fed into a pulsed reset preamplifier and the output signals were divided into five identical signals. Each of the three shaping amplifiers and two timing amplifiers in the DAQ system receive one of these outputs for further processing and digitization. The two high-gain shaping amplifiers S_{p6} and S_{p12} cover a 0–12 keVee energy range for light-DM analysis with shaping times of 6 μs and 12 μs, respectively, and Gaussian filtering. The involved high-gain timing amplifier (Tp) measures the signal' rise times in the 0–12 keVee energy range to distinguish between bulk and surface events. The low-gain shaping amplifier and low-gain timing amplifier covered the high-energy range to collect the background of the experimental system. The system is triggered by an S_{p6} signal loaded onto a leading-edge discriminator. Moreover, random trigger (RT) signals with a 0.05 Hz frequency were generated by a signal generator and fed into the DAQ system. Using RT signals, the zero-point energy resolution can be determined, the dead time of the DAQ can be estimated, and the efficiencies of energy-independent cuts within the data analysis procedure can be determined.

The output signals of the system described above were digitized via 14-bit 100-MHz flash analog-to-digital converters (FADCs). Once triggered by an S_{p6} or RT signal, the FADCs record the raw output pulses from those amplifiers in the time window of 120 μs. Notably, the pulsed reset preamplifier generates discharge signals (inhibit signal), which can induce significant noises and false triggers in the subsequent time period. To avoid negative effects induced by inhibit signals, each inhibit signal is coupled with a 29.5-ms veto signal provided by a timer.

The CDEX-10 DAQ system is shown in Fig. 1. The digitized outputs of the DAQ system are the dataset to be shared herein. These raw data are stored in the form of binary files. Apart from raw pulse shapes of HPGe detectors, root files with information extracted from these pulse shapes will also be presented for easier usage.

3 Data Records

The data can be accessed via the repository listed in the specifications table.

Specifications table

Subject	Particle physics
Specific subject area	Ultra-low-background experiment
Data format	Raw&Analyzed
Type of data	Binary file and ROOT file
How data were acquired	Measurements were taken using HPGe detectors in CJPL.
Parameters for data collection	Refer to Materials and Methods section.
Description of data collection	Data were collected by CDEX-10 for direct detection of DM.

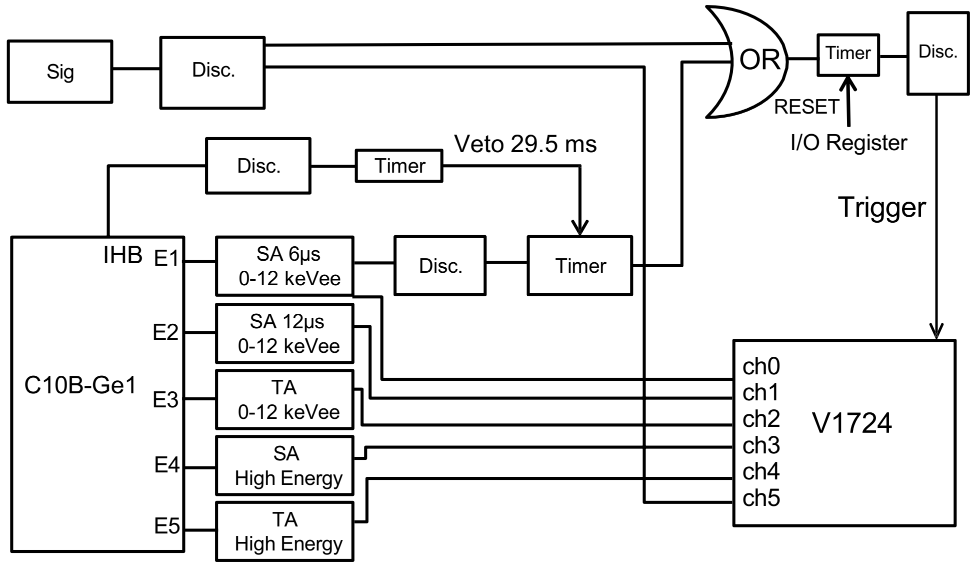

Fig. 1 Schematic of the DAQ system of the C10B-Ge1 detector. Sig: signal generator. Disc: discriminator. SA: shaping amplifier. TA: timing amplifier

Specifications table

Subject	Particle physics
Data collection	The data were collected from CDEX-10 using the HPGe detectors with or without calibration
	sources nearby.
Data source location	Institution: Department of Engineering Physics, Tsinghua University
	Country: China
Data accessibility	Repository name: ScienceDataBank
	Data identification number: https://cstr.cn/31253.11.sciencedb.17724
	Direct URL to data: https://doi.org/10.57760/sciencedb.17724
Related research article	H. Jiang et al. Sci. China-Phys. Mech. Astron. 62, 31012 (2019) doi: 10.1007/s11433-018-8001-3

The dataset includes two types of files: binary files (.bin) and ROOT (.root) files. Binary files contain original data pulses collected from different channels of the FADC, and ROOT files contain information extracted from the pulse shapes in the binary files. Entries corresponding to the same signal were connected by the file name and index of each entry within the data file.

Binary files cannot be read directly using traditional data processing tools. The binary files are written in a predefined structure to save the outputs from the involved FADC. Therefore, to process the saved pulse shapes, the binary files must be decoded according to the predefined structure. For easy of usage, the decoding script used in CDEX-10 was also supplemented in our repository. The relevant binary file structures are explicitly marked in the comments within the script. Users can refer to the script for more detailed information. Moreover, the script is easy for users to use to transform binary data into readable pulse shapes. Each binary file contains 10,000 events with pulse shapes from the five channels described previously.

Pulse shapes contain all the information about events in the HPGe detectors. However, one pulse shape contains 12,000 data points, rendering data processing extremely data-consuming. In most cases, the data analysis procedure is not performed directly on the original pulse shapes; instead, it is performed on the parameterized information of the raw pulse shapes. CDEX collaboration uses characteristics extracted from pulse shapes, including maximum values, pulse rise time, pulse FWHM, etc. The extracted information is stored in the form of ROOT files. Each ROOT file contained information on 10,000 events. An information extraction script was also added to the data repository. Users can refer to the scripts' comments for more detailed information.

4 Recommended repositories to store and find data

The dataset was made public via the Science Data Bank, and the link to the repository can be found in the Specifications Table. The dataset can also be found on the official CDEX website (https://cdex.ep.tsinghua.edu.cn/). Users can contact the CDEX for detailed information.

5 Technical Validation

The dataset recorded potential rare physical events and radioactive background events near the experimental setup. If the dataset is valid, peaks from the environmental radionuclides should be identified. Herein, we demonstrate the calibration result of the dataset in both low- and high-energy ranges as well as the energy spectrum derived from the dataset in Figs. 2 and 3 to demonstrate the validity of the dataset.

For energy calibration in the 0–12 keVee energy range (Fig. 2a), the optimal integrated area of the pulse from S_{p12} was selected to define the energy for its excellent energy

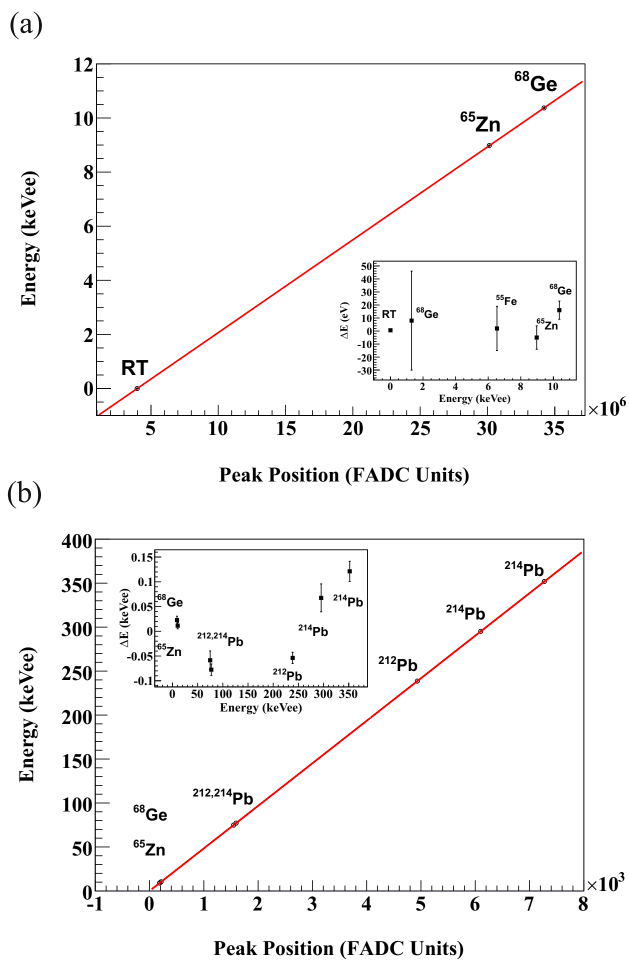

Fig. 2 (Color online) **a** Energy calibration in the low-energy region using the 10.37 keV peak of ^{68}Ge and the 8.98 keV peak of ^{65}Zn, and the zero energy defined by the RT events. The inset shows the energy deviations of these calibrated peaks, the 1.30 keV peak of ^{68}Ge, and the 6.54 keV peak of ^{55}Fe. The energy deviation from linearity is smaller than 16 eVee. **b** Energy calibration in the high-energy region. As shown in the inset, the energy deviations are below 0.15 keVee in the 0–350 keVee energy range

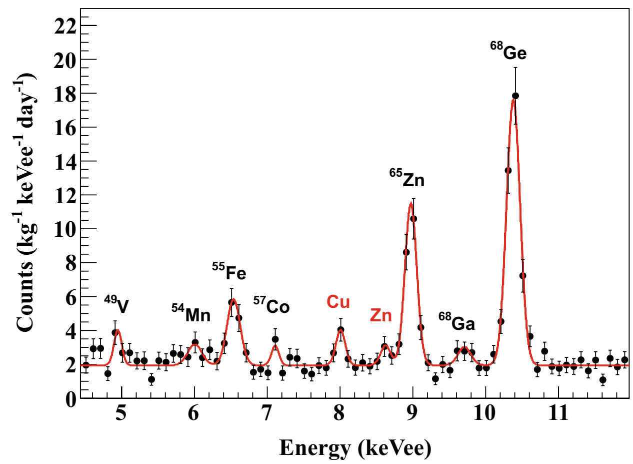

Fig. 3 (Color online) Energy spectrum in the 4.5–12.0 keVee energy range from the experimental data. Peaks from different radionuclides are identified

linearity in the low energy region. Energy calibration was done with the internal characteristic K-shell X-rays (KX) originating from the electron capture of the cosmogenic radioisotopes: 10.37 keV of ^{68}Ge and 8.98 keV of ^{65}Zn, and the zero energy defined by the RT events. The inset figure displays the energy difference (Δ) between the calibrated energy and the real energy of these three peaks, together with the other two peaks observed in the measured background spectrum: a 1.30 keV peak from ^{68}Ge and a 6.54 keV peak from ^{55}Fe. The energy deviation from the linear fit is less than 16 eVee.

In the high-energy region (Fig. 2b), the spectrum containing several K-shell X-rays and γ-rays from ^{65}Zn (8.98 keV), ^{68}Ge (10.37 keV), and ^{212,214}Pb (74.8 keV, 77.1 keV, 238.6 keV, 295.2 keV, 351.9 keV) was calibrated using the fitted height of the Tp pulse. The energy deviation from linearity was below 0.15 keVee in the energy range 0–350 keVee.

As shown in Fig. 3, seven KX peaks are attributed to the internal cosmogenic isotopes ^{68}Ge (10.37 keV), ^{68}Ga (9.66 keV), ^{65}Zn (8.98 keV), ^{57}Co (7.11 keV), ^{55}Fe (6.54 keV), ^{54}Mn (5.99 keV), ^{49}V (4.97 keV) in the Ge crystal. The remaining two X-rays were inferred to originate from the Cu (8.0 keV) and Zn (8.6 keV) materials close to the crystal, which are excited by high-energy γ-rays. This inference was obtained through an in-depth analysis of their distribution characteristics within the detector, based on their pulse shape characteristics and understanding of the detector structure [4].

6 Usage notes

We intuitively propose three potential applications.

1. The dataset is collected in the radiopure environment in CJPL; therefore, it can potentially be used for other rare event detection.
2. Understanding the radioactive environment is crucial for research on CJPL. The dataset can be used to identify components within the radioactive background of the CJPL.
3. The internal mechanisms of HPGe detectors are complicated, and understanding these mechanisms can help improve the detector performance. The mechanism involved within the detector can be inferred by studying the pulse shapes collected by HPGe detectors.

Author contributions Q. Yue and Z. Zeng conceived the experiment, L.T. Yang and H. Ma conducted the experiment, and L.T. Yang and Z.Y. Zhang analyzed the results. All authors have reviewed the manuscript.

Code availability The code developed for this dataset is publicly available and can be accessed without restriction. The authors encourage the use and reuse of this code for research and educational purposes. 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. J.P. Cheng, K.J. Kang, J.M. Li et al., The China Jinping underground laboratory and its early science. Annu. Rev. Nucl. Part. Sci. **67**, 231 (2017). https://doi.org/10.1146/annurev-nucl-102115-044842
2. Y.C. Wu, X.Q. Hao, Q. Yue et al., Measurement of cosmic ray flux in the China JinPing underground laboratory. Chin. Phys. C **37**, 086001 (2013). https://doi.org/10.1088/1674-1137/37/8/086001
3. L.T. Yang, Q. Yue, Experimental Dataset from CDEX's High Purity Germanium Detectors in China Jinping Underground Laboratory[DS/OL]. V1. Science Data Bank, 2024[2024-12-02]. https://doi.org/10.57760/sciencedb.17724
4. H. Jiang, L.T. Yang, Q. Yue et al., Performances of a prototype point-contact germanium detector immersed in liquid nitrogen for light dark matter search. Sci. China-Phys. Mech. Astron. **62**, 31012 (2019). https://doi.org/10.1007/s11433-018-8001-3