Prepared for submission to JINST

ML-Driven Optimization of BEGe Detector Event Selection in the VIP Experiment

S.H. Yip,[1,2,*] **S. Manti,**[2,*] **M. Bazzi,**[2] **N. Bortolotti,**[3,4] **M. Bragadireanu,**[5,2] **M. Cargnelli,**[6,2] **A. Clozza,**[2] **L. De Paolis,**[2] **R. Del Grande,**[7,2] **C. Guaraldo,**[2,†] **M. Iliescu,**[2] **M. Laubestein,**[8] **J. Marton,**[6,2] **F. Napolitano,**[9,10] **F. Nola,**[2] **K. Piscicchia,**[3,2] **A. Porcelli,**[2,3,11,12] **A. Scordo,**[2] **F. Sgaramella,**[2] **D. Sirghi,**[2,3,5] **F. Sirghi,**[2,5] **J. Zmeskal**[6,†] **and C. Curceanu**[2]

[1] *University of California, Berkeley, CA 94720, USA*

[2] *Laboratori Nazionali di Frascati INFN, Frascati, Italy*

[3] *Centro Ricerche Enrico Fermi, Museo Storico della Fisica e Centro Studi e Ricerche "Enrico Fermi", Roma, Italy*

[4] *Physics Department, "Sapienza" University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy*

[5] *Horia Hulubei National Institute of Physics and Nuclear Engineering (IFIN-HH), Măgurele, Romania*

[6] *Stefan-Meyer-Institut für Subatomare Physik, Vienna, Austria*

[7] *Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, 115 19, Prague, Czech Republic*

[8] *Laboratori Nazionali del Gran Sasso, Istituto Nazionale di Fisica Nucleare, Via G. Acitelli 22, 67100, Assergi, Italy*

[9] *Via A. Pascoli 06123, Perugia (PG), Italy, Dipartimento di Fisica e Geologia, Università degli studi di Perugia*

[10] *INFN Sezione di Perugia, Via A. Pascoli, 06123 Perugia, Italia*

[11] *Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, ul. prof. Stanislawa Lojasiewicza 11, 30-348, Kraków, Poland*

[12] *Centro de Investigacíon, Tecnología, Educacíon y Vinculacíon Astronómica, Universidad de Antofagasta, Avenida Angamos 601, 1240000, Antofagasta, Chile*

[*] *Corresponding Authors*

[†] *Deceased*

E-mail: simone.manti@lnf.infn.it, jason17@berkeley.edu

Abstract: Experiments probing foundational questions in quantum mechanics, such as potential violations of the Pauli Exclusion Principle and spontaneous wave-function collapse, demand detectors with exceptional sensitivity in a low-background environment.

Keywords: Only keywords from JINST's keywords list please

ArXiv ePrint: 1234.56789

Contents

1 Introduction

Experimental tests of the foundations of quantum mechanics require extreme sensitivity, as they target phenomena that, if observed, would occur at very reduced rates. Such rare-event searches provide a unique window into possible deviations from the standard quantum framework, complementing more conventional precision tests and opening opportunities to probe new physics at the interface between microscopic and macroscopic scales [1, 2].

Among the most fundamental principles of quantum theory is the Pauli Exclusion Principle (PEP) [3], which underlies the structure of atoms, the stability of matter [4], and even the behavior of astrophysical objects like neutron stars [5]. Its validity is directly linked to the Spin–Statistics Theorem [6], which dictates that fermions (bosons) obey antisymmetric (symmetric) wave functions under exchange of particles. Any violation of PEP would have profound implications, pointing to physics beyond the established quantum framework. From an experimental point of view, such violations could manifest as examples of anomalous atomic transitions, shifted by a few tenths of a keV from the standard spectral lines [7]. Establishing increasingly stringent limits on such effects is therefore a central goal of dedicated rare-event searches.

Another direction involves models of spontaneous wave-function collapse, which attempt to resolve the measurement problem by introducing additional physical mechanisms that suppress quantum superpositions at macroscopic scales. Prominent examples include the Continuous Spontaneous Localization (CSL) model [8] and the Diósi–Penrose (DP) model [9, 10]. Both predict the spontaneous emission of radiation, in particular X-rays in the keV range, which can serve as experimental signatures of the particular collapse model especially in the low-energy regime [11].

Detecting such rare and subtle signatures requires not only excellent energy resolution but also extremely low background. Underground laboratories play a crucial role in this context. The Gran Sasso National Laboratory (LNGS) provides a uniquely low-background environment for the

study of rare events. Located under approximately 1400 m of rock-equivalent to 3800 m w e water-equivalent (m w.e.), the overburden reduces the cosmic muon flux to only about 3.41×10^{-4} m^{-2} s^{-1} [12]. This substantial attenuation of the muon flux enables experiments at LNGS to achieve significantly reduced background levels. This makes LNGS one of the world's premier environments for rare-event searches, hosting experiments ranging from neutrino physics to dark matter and double beta decay [13–15].

Within this environment, the VIP (Violation of the Pauli Exclusion Principle) collaboration has progressively tightened the upper limits on PEP-violating events. The original VIP experiment set an upper limit of $\beta^2/2 < 4.7 \times 10^{-29}$ [16], the most stringent constraint at the time. Its successor, VIP-2, pushed this further, establishing limits of $\beta^2/2 \leq 6.8 \times 10^{-42}$ (Bayesian, 90% CL) and $\beta^2/2 \leq 7.1 \times 10^{-42}$ (Frequentist CL$_s$, 90% CL) [17]. The ongoing VIP-3 effort is extending the search across the periodic table, with high-Z targets such as zirconium, silver, palladium, and tin [18].

A central element of these experiments is the detector technology. Broad Energy Germanium (BEGe) detectors combine the features of coaxial and low-energy HPGe designs, covering a wide dynamic range from a few keV up to several MeV. They are particularly suited for rare-event searches because of their excellent resolution and efficiency at low energies. However, achieving thresholds of a few keV is challenging due to microphonic noise and electronic instabilities, which can obscure faint signals precisely in the energy range of interest for PEP-violating transitions and collapse-induced radiation [11].

In this work, we report recent advances in the operation of a BEGe detector within the VIP program at LNGS. Two key improvements were implemented. First, the detector setup was mechanically and acoustically isolated using a soundproof box and pneumatic suspension, suppressing microphonic contributions. Second, we developed and validated a machine–learning–based event selection strategy, employing a denoising autoencoder and a convolutional neural network to enhance discrimination between signal and background. These methods were applied to data acquired in 2021, demonstrating improved effective energy threshold and a higher signal-to-background ratio.

The remainder of this paper is organized as follows. Section 2 describes the experimental setup, including the BEGe detector and its isolation system. Section 3 introduces the machine-learning framework for event selection, with emphasis on the denoising and classification stages. Section 4 presents the performance validation with the 2021 dataset and discusses the improvements in resolution, threshold, and background rejection. Section 5 concludes with a summary and an outlook for future developments in the VIP program and related rare-event searches.

2 Experimental Setup

The experimental setup for the BEGe detector is illustrated in Figure 1. The detector is housed inside a soundproof enclosure, which also functions as a Faraday cage, providing both acoustic and electromagnetic isolation. The setup features a cylindrical BEGe detector (Canberra model BE3830) with a diameter of 71 mm, a thickness of 30 mm, and an active area of 3800 mm^2. Cooling to approximately 80 K is achieved using liquid nitrogen, supplied by a Canberra 7915-30-CD-ULB

cryostat placed within a dewar that is refilled weekly. The dewar and detector are mounted on pneumatic suspensions, which effectively isolate the system from mechanical vibrations.

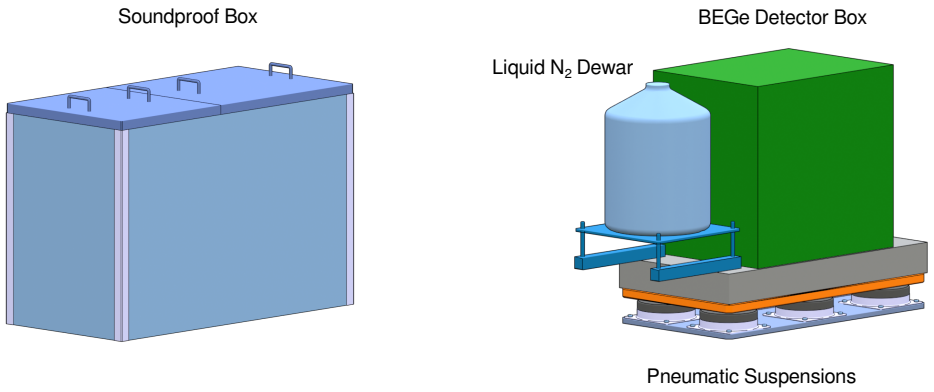


Figure 1: The BEGe setup installed at LNGS. On the left, the soundproof enclosure provides both acoustic and electromagnetic isolation, functioning as a Faraday cage. On the right (not to scale), the setup includes the cryostat (light blue), the germanium detector (green), the support structure (blue), and the vibration isolation platform (gray/orange).

3 BEGe Event Selection Strategy

~~One of the primary objectives of the VIP experiment is to characterize and reduce the background in the search for new physics phenomena, such as Non-Commutative Quantum Gravity (NCQG) models, which predict violations of the Pauli Exclusion Principle (PEP) as a consequence of the Spin-Statistics Theorem (SST) [**PauliSpinStatistics1940**], and dynamical collapse models, including Continuous Spontaneous Localization (CSL) [**ghirardi1986unified**, **pearle1976reduction**, **pearle1979toward**, **pearle1989combining**, **ghirardi1990markov**] and Diósi–Penrose (DP) models [**diosi2014gravitation**, **diosi1987universal**, **diosi1989models**, **penrose1996gravity**, **penrose2014gravitization**]. These latter models address the quantum measurement problem by predicting spontaneous emission of faint X-ray radiation [11]. The BEGe detector is dedicated to acquiring and analyzing background events to optimize signal selection and improve sensitivity for the upcoming effort to set limits on the collapse model parameters.~~

~~A clear understanding and reduction of the background is crucial, as the rare signals of interest, such as PEP-violating transitions in Ge, which are shifted by ~0.3 keV from the standard K_α line at ~9.9 keV, can only be reliably identified if the background is well characterized. Furthermore, a well-characterized background is important for distinguishing the subtle signatures predicted by collapse models, as these models may produce unique spontaneous radiation patterns that could be obscured by noise. Achieving a low energy threshold (down to ~6 keV) is essential, but in the sub-25 keV range, electronic noise becomes significant in the experimental setup described in Section 2 [11].~~ To address this, a denoising autoencoder (DAE) is used to reconstruct the underlying

events in the presence of substantial background noise, allowing more accurate event analysis and potentially lowering the effective energy threshold [11].

3.1 Denoising with Denoising Autoencoder (DAE)

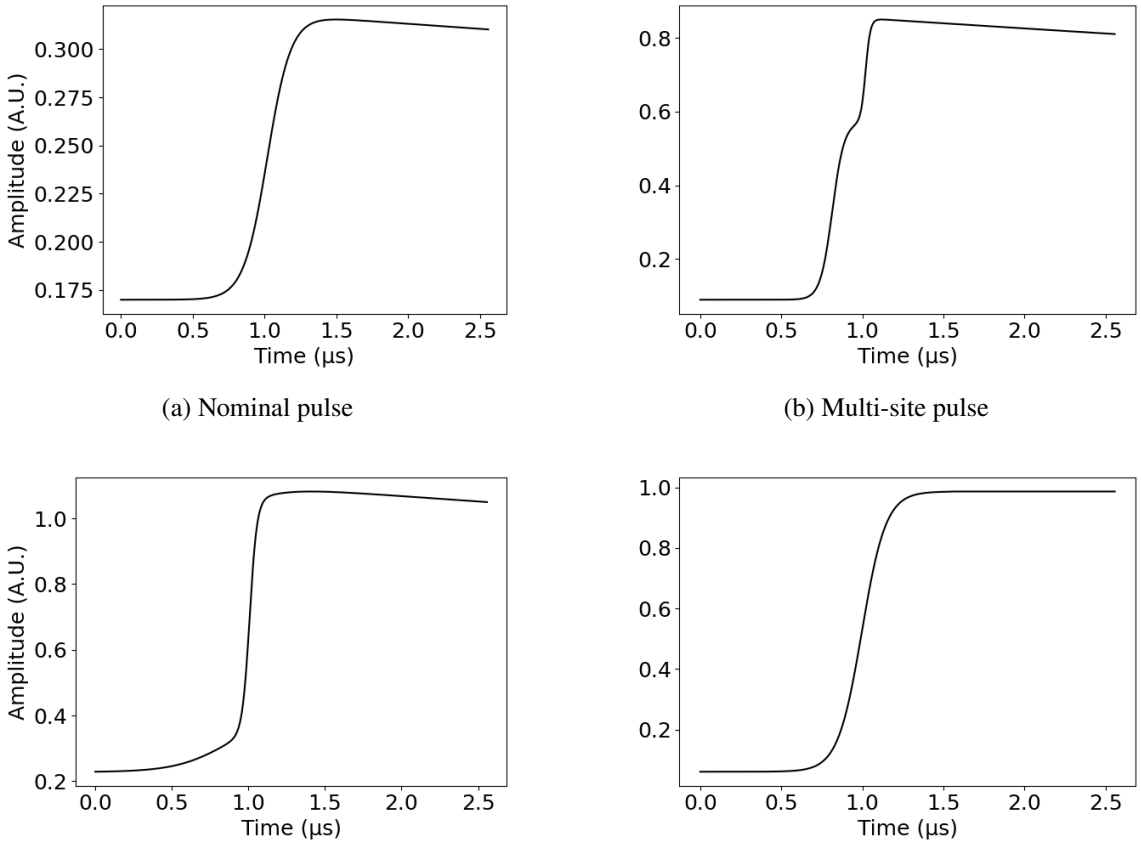
Traditional signal denoising and smoothing methods, such as exponential smoothing and Savitzky-Golay filtering, have been widely used to mitigate noise in time series and pulse data [**brown1956exponential**, **savitzky1964smoothing**]. These approaches typically rely on carefully chosen parameters (e.g., smoothing factors or window sizes), which must be finely tuned to achieve optimal performance. However, in the presence of significant or nonstationary noise, such parameter selection becomes increasingly challenging and may not yield robust results across varying datasets. A further discussion of denoising methods is given in ref. [**Anderson_2022**]. In contrast, data-driven methods such as denoising autoencoders (DAEs) can learn complex noise distributions directly from data, offering improved adaptability and robustness without the need for manual parameter adjustment.

All events are preprocessed prior to training through baseline subtraction and normalization by amplitude. For each pulse, the baseline is first estimated as the mean of the initial 100 samples and subtracted from the entire waveform to remove DC offsets. The resulting signal is then processed with a trapezoidal filter, and the amplitude is extracted as the maximum absolute value of the filtered output beyond the filter response region. The baseline-subtracted pulse is then normalized by this amplitude, resulting in a unit-amplitude, baseline-corrected signal. This standardized preprocessing pipeline is applied uniformly to both synthetic and experimental datasets to ensure robust and consistent inputs throughout model training and evaluation. It is worth noting that without this baseline subtraction and normalization, the DAE fails to accurately reconstruct pulse shapes.

In order to achieve accurate reconstruction of signal pulses in the presence of variational background noise, a data augmentation strategy is employed based on the simulation of idealized detector pulses. Synthetic pulses are generated using analytic mathematical models that capture the features of the signal events in the BEGe detector, as described in Section 2. Four classes of pulses are included in the dataset to reflect the event topologies: nominal, multi-site, and slow-rise pulses.

Nominal pulses (Figure 2a) correspond to single-site energy depositions occurring within the bulk of the detector, where the electric field is uniform and charge carriers are efficiently collected at the electrodes. These events are characterized by a prompt, steep rising edge and an exponential decay, and represent the expected signal. **Multi-site pulses** (Figure 2b) are generated to mimic the response to events in which energy is deposited at multiple, spatially separated locations within the detector volume. Such events can arise from, for example, Compton scattering or coincident background interactions. The resulting pulse shapes typically exhibit multiple rising steps or a more complex rise profile, reflecting the superposition of charge collection from each interaction site. **Slow-rise pulses** (Figure 2c) are included to model interactions occurring near the detector surfaces, where the electric field strength is reduced. In these regions, charge carriers drift more slowly to the electrodes, resulting in a notably slower rise time compared to bulk events. As observed in our data, these surface events are more likely to originate from background processes rather than the rare signal of interest. The inclusion of slow-rise pulses in the training set allows the autoencoder (AE) to learn features associated with surface backgrounds, which can be subsequently suppressed through pulse shape discrimination, thereby improving the overall signal-to-noise ratio. **Flat-top**

(saturated) pulses (Figure 2d) are also incorporated in the dataset [1]. These pulses are characterized by a rapid rise followed by a plateau, which can occur when the detector response saturates due to high energy depositions or limitations in the readout electronics. Including flat-top pulses in the training set is crucial to ensure that the AE can recognize and appropriately process saturated signals, which may otherwise be misidentified or improperly reconstructed.



(a) Nominal pulse

(b) Multi-site pulse

(c) Slow-rise pulse

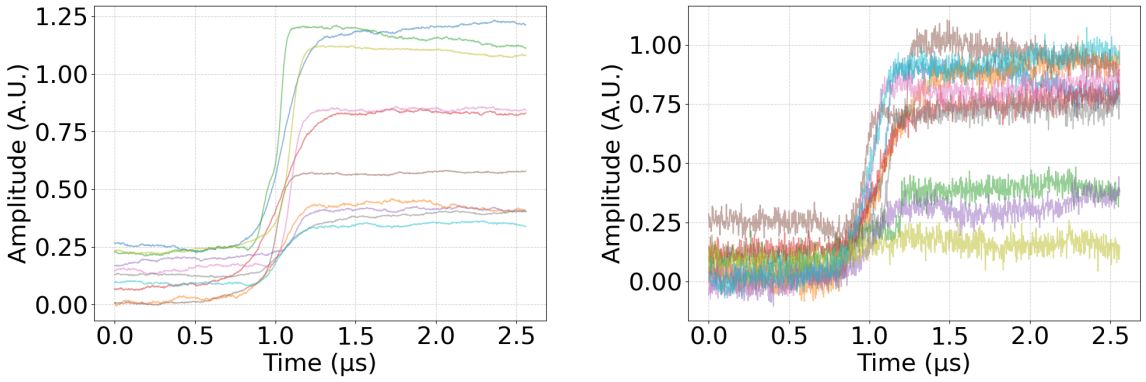
(d) Flat-top (saturated) pulse

Figure 2: Examples of the four pulse types used in the augmented dataset before preprocessing. Different amplitudes are generated to ensure the same preprocessing technique applied to detector data works properly.

Each simulated pulse is superimposed with realistic noise. Both white Gaussian noise and colored (correlated) noise are included, with the latter generated by integrating a white noise process to reproduce the low-frequency fluctuations typical of the experimental setup. Two distinct noise levels are applied (see Figure 3a and 3b), with the final dataset comprising 5000 simulated events, evenly divided between the two noise conditions. This simulation strategy ensures that the AE is exposed to a wide range of noise environments during training, thereby enhancing its ability to robustly reconstruct clean pulses from noise in real data.

The DAE is trained using the preprocessed augmented pulse dataset, where the input $\mathbf{X}$ consists

[1] Saturated pulses are added primarily for the new data acquisition in 2025, as discussed in Section 4.



(a) (b)

Figure 3: Example of 10 randomly selected augmented pulses shown with two different noise levels before preprocessing. **(a)** Pulses with lower noise amplitude. **(b)** Pulses with higher noise amplitude.

of noisy pulses and the target $\mathbf{Y}$ corresponds to the associated clean signals. The model is designed to learn a mapping $f_\theta : \mathbf{X} \rightarrow \mathbf{Y}$ that reconstructs the underlying clean pulse from its noisy counterpart.

The AE architecture is based on a 1D convolutional neural network (CNN) with an encoding and decoding structure. The encoder comprises three convolutional layers with progressively decreasing filter sizes, each followed by the rectified linear unit (ReLU) activation function, and max-pooling operations, reducing the temporal dimension and capturing hierarchical features of the pulse shape [**glorot2011deep**]. The decoder mirrors this structure, using upsampling layers and additional convolutional layers with ReLU activation to restore the original pulse length, followed by a final convolutional layer with linear activation to produce the reconstructed signal. The model is trained end-to-end using the Adam optimizer with a learning rate of 10^{-4} [**kingma2017adammethodstochasticoptimization**]. After training, the same preprocessing pipeline is applied to experimental data, which are then fed into the trained AE to reconstruct the normalized, denoised pulses.

3.2 ML-Based Event Selection

The reconstructed pulses produced by the DAE serve as inputs for the subsequent binary CNN classifier. Each event is represented by two input channels: (1) the normalized pulse, which is directly taken from the output of the DAE and requires no additional normalization, and (2) the normalized first derivative of the pulse, scaled to the range [0, 1]. By including both the pulse and its first derivative as separate input channels, the CNN can capture complex inter-relationships between the overall pulse shape and the subtle changes such as multi-site events. Signal events for training are defined as those that satisfy all of the following conditions:

1. **Feature constraints:** The rise time, full-width-at-half-maximum (FWHM) time, and L1 norm each fall within one standard deviation (1σ) of their respective total distributions. *Rise time* is the time interval for the pulse to increase from 10% to 90% of its maximum amplitude after baseline subtraction. *FWHM time* is the width of the pulse at half its maximum value,

measured on both sides of the peak in the normalized first derivative space. *L1 norm* is the sum of absolute differences between the observed pulse and a reference pulse after baseline subtraction.

2. **Peak count:** The event contains exactly one peak in the normalized first derivative, as determined using the `find_peaks` function from SciPy [**2020SciPy-NMeth**].

3. **Labeling:** The event is carefully reviewed and assigned a *signal* label.

Background events are defined as those assigned a *background* label. The total number of signal events for training is set to be two times the total number of background events. 70% of the selected dataset is used for training, and the remaining events are split evenly between testing and validation dataset (15% each).

The CNN model is composed of sequential 1D convolutional layers with increasing filter sizes and decreasing kernel widths, each followed by batch normalization and max pooling to extract and condense features from the input. A fully connected dense layer interprets the extracted features, with a dropout layer included to mitigate overfitting. The final output layer uses a sigmoid activation function to produce a probability score for binary classification. The model is trained using the Adam optimizer with a learning rate of 10^{-5} and binary cross-entropy loss [**kingma2017adammethodstochasticoptimization**], and evaluated using accuracy, precision, and recall metrics. Early stopping of 10 epochs is employed to prevent overfitting and ensure optimal model weights.

After training, the CNN assigns a probability to each event, representing the likelihood of it being a signal. The optimal classification threshold is obtained from the receiver operating characteristic (ROC) curve as the point that maximizes TPR − FPR, corresponding to the highest separation from a random classifier [**youden1950index**].

The trained model is then applied to all reconstructed events, assigning a probability score to each. Events are classified as signal or background based on whether their assigned probability exceeds the optimal threshold determined from the ROC.

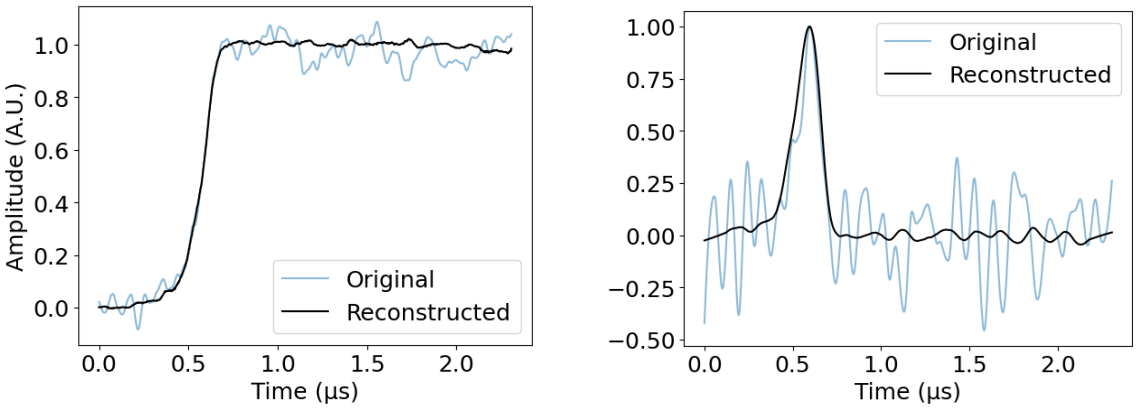
4 Results & Discussion

The performance of the DAE was evaluated through both quantitative and qualitative means. The training and validation loss as a function of epochs decrease rapidly in the initial epochs and reach a plateau after approximately 100 epochs, stabilizing at a value around 5×10^{-4}. With the correct normalization, this shows that the DAE is able to reconstruct the underlying clean signals with high accuracy. The close agreement between the training and validation curves suggests that the model generalizes well and does not suffer from overfitting.

The effectiveness of the DAE in recovering signal pulses from noisy backgrounds is further illustrated in Figure 4. In both panels, the original waveforms and the DAE-reconstructed waveforms are shown for an event corresponding to roughly 15 keV. In panel (a), it can be seen that the DAE effectively suppresses background noise, accurately recovering the underlying signal. In a more challenging case, shown in Figure 4b, the first derivative of the original waveform is heavily

contaminated by noise, but the DAE is able to extract and reconstruct the essential features of the pulse, showing its robustness even in the presence of significant background noise.

This denoising step offers two key advantages for subsequent signal/background classification. First, by providing cleaner inputs, the DAE reduces the complexity required for the downstream CNN classifier. This simplifies the architecture and training process, lowering the computational requirements and improving classification performance [2]. Second, the DAE is particularly effective at preserving low-energy events, which are otherwise more susceptible to noise and could be lost or misclassified. As shown in Figure 4b, such events are challenging to classify directly from the raw waveform but become much more tractable after denoising.



(a) Original

(b) First derivative

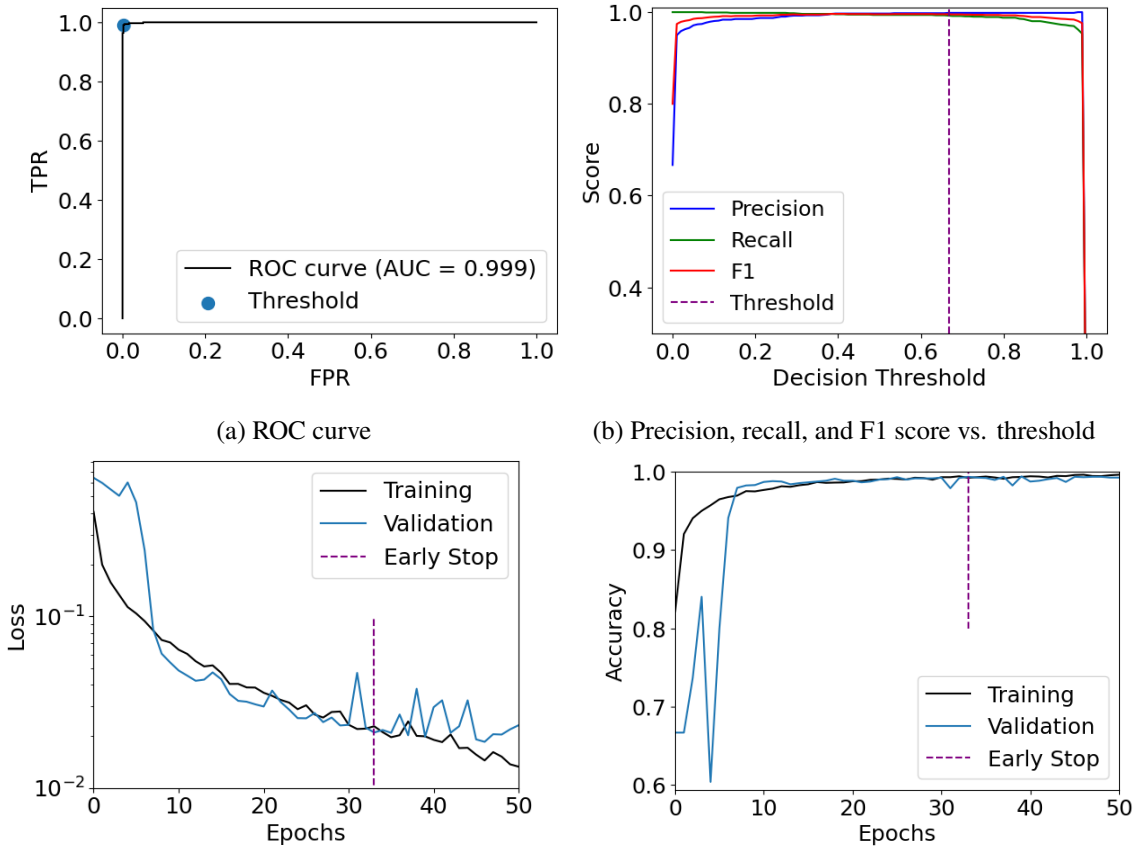
Figure 4: Example of a reconstructed pulse overlaid on top of the raw pulse. All pulses shown are preprocessed. This pulse correspond to ~15keV.

After denoising, a CNN was trained to classify events as signal or background. The performance of the training is illustrated in Figure 5. The ROC curve shown in Figure 5a demonstrates a near perfect separation between signal and background, with an area under the curve (AUC) of 0.999. The optimal decision threshold is found to be ~0.76, at which the classifier achieves maximal discrimination power.

The training and validation loss curves (Figure 5c) show that both losses converge quickly as the training progresses. Different early stopping steps are tested, and the results show consistency that the validation loss plateaus after about 33 epochs while the training loss continues to decrease, confirming the current 33 epochs as the optimal early stopping. In addition to loss, the training and validation accuracy curves in Figure 5c are monitored throughout training. Both accuracies increase sharply and stabilize at values very close to 1.0, with the validation accuracy closely following the training accuracy after the initial epochs, indicating a stable generalization and confirming the absence of overfitting. All metrics (precision, recall, and F1 score) remain close to unity near the optimal threshold, finalizing the robustness of the CNN classifier.

Once the training is complete, signal events are selected as those with a classification probability above the chosen threshold. An energy spectrum, fitted using a binned maximum likelihood method,

[2] An analysis of applying ML methods directly on raw signal showed a consistent rejection of low-energy signals due to significant background noise, regardless of the complexity of the architecture.



(a) ROC curve

(b) Precision, recall, and F1 score vs. threshold

(c) Training and validation loss

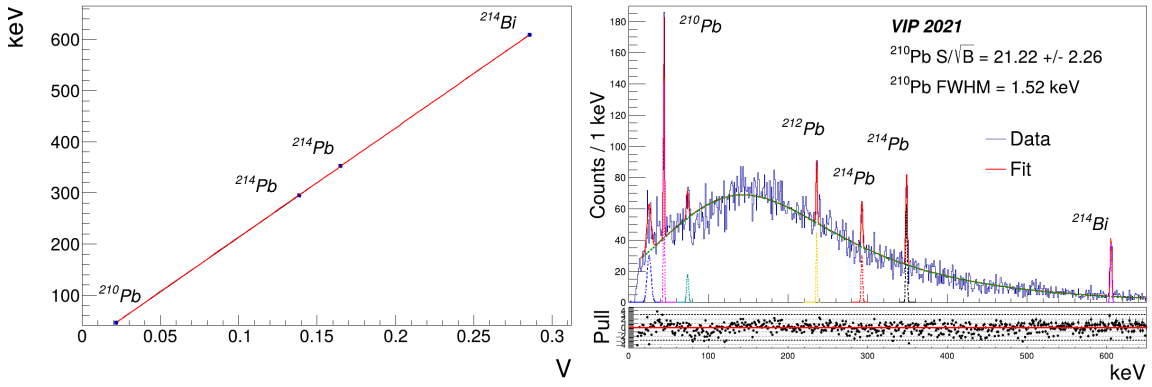
(d) Training and validation accuracy

Figure 5: Performance of the CNN model.

is then constructed both before (Figure 6b) and after (Figure 7) cleaning, using amplitudes extracted with a trapezoidal filter. Using the DAE-CNN workflow, the resolution (FWHM) at the ^{210}Pb peak and the energy threshold are determined to be approximately 1.47 keV and 11 keV, respectively. The calibration is performed using four peaks present in the spectrum, originating from the decay chains of ^{238}U: ^{210}Pb at 46 keV, ^{214}Pb at 295 keV and 352 keV, and ^{352}Bi at 609 keV, as shown in Figure 6a. Each spectrum is fitted with the following function:

$$\begin{aligned} f(x) = {} & e^{a_1+b_1x}\,\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{x-\mu_1}{\sigma_1}\right)\right] \\ & + e^{a_2+b_2x}\,\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{x-\mu_2}{\sigma_2}\right)\right] \\ & + \sum_i A_i\,\exp\left(-\frac{(x-\mu_{g,i})^2}{2\sigma_{g,i}^2}\right). \end{aligned} \tag{4.1}$$

The background is modeled as the sum of two error functions, where the first describes the exponential increase below 100 keV, and the second describes the exponential decrease above 100 keV. Each transition line is modeled with a Gaussian distribution.



(a) (b)

Figure 6: **(a)** Calibration curve used to relate amplitudes to energies. **(b)** Energy spectrum without any data cleaning.

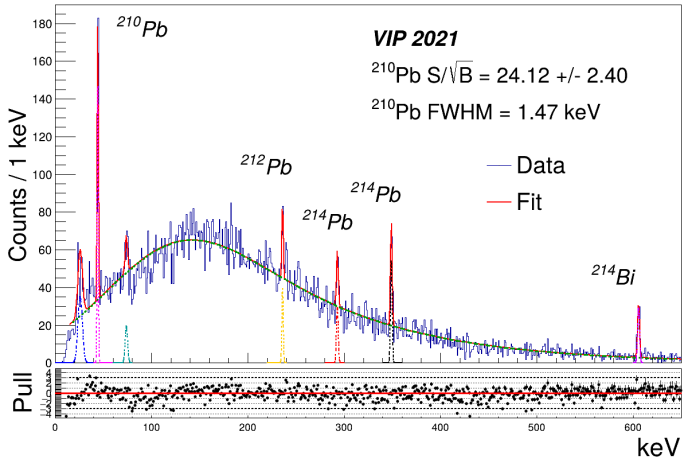


Figure 7: Energy spectrum after applying machine learning.

The application of the DAE in the preprocessing stage has demonstrated a significant improvement in the reconstruction of low-energy events in the BEGe detector. The quick convergence of loss and epochs is facilitated by the DAE preprocessing, which provides cleaner waveforms and thus accelerates the learning process. The high accuracy, along with the strong AUC, precision, recall, and F1 scores, indicates that the combined DAE-CNN workflow provides highly effective event selection, as illustrated in Figures 5. Without the DAE, events in the sub-15 keV range are dominated by electronic noise, making reliable analysis and event classification challenging. In contrast, the DAE effectively suppresses background noise and reconstructs the underlying signal, lowering the effective energy threshold to 11 keV. This directly addresses the experimental goal mentioned in Section 3, where lowering the energy threshold is essential to enhance the sensitivity to rare processes. Furthermore, the use of the DAE reduces the computational complexity of the subsequent CNN classification, as cleaner input waveforms allow for a simpler architecture without compromising classification performance. In addition, the DAE-preprocessing procedure allows for alternative analysis approaches, including traditional non-ML methods and feature-based ML

techniques based on the extracted parameters from the denoised events. The ML-based approach yields a 14% improvement in the signal-to-background ratio compared to the baseline analysis, which is particularly impactful in the low-energy regime where background contributions are most significant.

Beyond immediate performance gains, this procedure can be further utilized as a scalable and transferable solution for low-background rare-event searches. The DAE-CNN workflow can be directly applied to ongoing datasets acquired with the BEGe detector in 2025 or similar experimental configurations, allowing the optimization of event selection strategies without requiring substantial retraining or hardware modifications. In addition, a Python package `vipbege` is provided alongside this work, allowing for the extraction of the features discussed here, as well as the generation of simulated waveforms both with and without noise, allowing its use for analysis in similar experimental configurations.

5 Conclusion

Data availability statement

Acknowledgments

We thank H. Schneider, L. Stohwasser, and D. Stückler from Stefan-Meyer-Institut for their fundamental contribution in designing and building the VIP-2 setup. We thank the Gran Sasso Underground Laboratory of INFN, INFN-LNGS, and its Director, Ezio Previtali, the LNGS staff, and the low radioactivity laboratory for the experimental activities dedicated to high-sensitivity tests of the Pauli exclusion principle.

[1] Stephen L Adler et al. "Collapse Models with Non-White Noises". In: *Journal of Physics A: Mathematical and Theoretical* 40.50 (Nov. 2007), p. 15083. ISSN: 1751-8121. DOI: 10.1088/1751-8113/40/50/012.

[2] Angelo Bassi et al. "Models of Wave-Function Collapse, Underlying Theories, and Experimental Tests". In: *Reviews of Modern Physics* 85.2 (Apr. 2013), pp. 471–527. DOI: 10.1103/RevModPhys.85.471.

[3] W. Pauli. "The Connection Between Spin and Statistics". In: *Physical Review* 58.8 (Oct. 1940), pp. 716–722. DOI: 10.1103/PhysRev.58.716.

[4] Freeman J. Dyson et al. "Stability of Matter. I". In: *Journal of Mathematical Physics* 8.3 (Mar. 1967), pp. 423–434. ISSN: 0022-2488. DOI: 10.1063/1.1705209.

[5] Norman K. Glendenning. *Special and General Relativity: With Applications to White Dwarfs, Neutron Stars and Black Holes*. Springer Science & Business Media, Apr. 2010. ISBN: 978-0-387-47109-9.

[6] Raymond Frederick Streater et al. *PCT, Spin and Statistics, and All That*. Princeton University Press, 2000. ISBN: 978-0-691-07062-9.

[7] Erik Ramberg et al. "Experimental Limit on a Small Violation of the Pauli Principle". In: *Physics Letters B* 238.2 (Apr. 1990), pp. 438–441. ISSN: 0370-2693. DOI: 10.1016/0370-2693(90)91762-Z.

[8] G. C. Ghirardi. “Unified Dynamics for Microscopic and Macroscopic Systems”. In: *Physical Review D* 34.2 (1986), pp. 470–491. DOI: 10.1103/PhysRevD.34.470.

[9] L. Diósi. “Gravitation and Quantum-Mechanical Localization of Macro-Objects”. In: *Physics Letters A* 105.4 (Oct. 1984), pp. 199–202. ISSN: 0375-9601. DOI: 10.1016/0375-9601(84)90397-9.

[10] Roger Penrose. “On Gravity’s Role in Quantum State Reduction”. In: *General Relativity and Gravitation* 28.5 (May 1996), pp. 581–600. ISSN: 1572-9532. DOI: 10.1007/BF02105068.

[11] Kristian Piscicchia et al. “Optimization of a BEGe Detector Setup for Testing Quantum Foundations in the Underground LNGS Laboratory”. In: *Condensed Matter* 9.2 (June 2024), p. 22. ISSN: 2410-3896. DOI: 10.3390/condmat9020022.

[12] G. Bellini et al. “Cosmic-Muon Flux and Annual Modulation in Borexino at 3800 m Water-Equivalent Depth”. In: *Journal of Cosmology and Astroparticle Physics* 2012.05 (May 2012), p. 015. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2012/05/015.

[13] XENON Collaboration et al. “First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment”. In: *Physical Review Letters* 131.4 (July 2023), p. 041003. DOI: 10.1103/PhysRevLett.131.041003.

[14] Majorana Collaboration et al. “Search for Neutrinoless Double-$\ensuremath{\beta}$ Decay in $^{76}\mathrm{Ge}$ with the Majorana Demonstrator”. In: *Physical Review Letters* 120.13 (Mar. 2018), p. 132502. DOI: 10.1103/PhysRevLett.120.132502.

[15] GERDA Collaboration et al. “Final Results of GERDA on the Search for Neutrinoless Double-$\ensuremath{\beta}$ Decay”. In: *Physical Review Letters* 125.25 (Dec. 2020), p. 252502. DOI: 10.1103/PhysRevLett.125.252502.

[16] S. Bartalucci et al. “New Experimental Limit on the Pauli Exclusion Principle Violation by Electrons”. In: *Physics Letters B* 641.1 (Sept. 2006), pp. 18–22. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2006.07.054.

[17] Fabrizio Napolitano et al. “Testing the Pauli Exclusion Principle with the VIP-2 Experiment”. In: *Symmetry* 14.5 (May 2022), p. 893. ISSN: 2073-8994. DOI: 10.3390/sym14050893.

[18] Simone Manti et al. “Testing the Pauli Exclusion Principle across the Periodic Table with the VIP-3 Experiment”. In: *Entropy* 26.9 (2024). ISSN: 1099-4300. DOI: 10.3390/e26090752.