Published by IOP Publishing for Sissa Medialab

Received: *March 10, 2023*
Accepted: *May 8, 2023*
Published: *June 5, 2023*

Generative adversarial networks for scintillation signal simulation in EXO-200

S. Li,[a,*] **I. Ostrovskiy,**[b,*] **Z. Li,**[c] **L. Yang,**[c] **S. Al Kharusi,**[d] **G. Anton,**[e] **P.S. Barbeau,**[g] **I. Badhrees,**[f,1] **D. Beck,**[a] **V. Belov,**[h] **T. Bhatta,**[i,2] **M. Breidenbach,**[j] **T. Brunner,**[d,9] **G.F. Cao,**[l] **W.R. Cen,**[l,3] **C. Chambers,**[d,4] **B. Cleveland,**[m,5] **M. Coon,**[a] **A. Craycraft,**[n] **T. Daniels,**[o] **L. Darroch,**[d] **S.J. Daugherty,**[p,6] **J. Davis,**[j] **S. Delaquis,**[j,7] **A. Der Mesrobian-Kabakian,**[m,8] **R. DeVoe,**[q] **J. Dilling,**[k,9] **A. Dolgolenko,**[h] **M.J. Dolinski,**[r] **J. Echevers,**[a,10] **W. Fairbank Jr.,**[n] **D. Fairbank,**[n] **J. Farine,**[m] **S. Feyzbakhsh,**[s] **P. Fierlinger,**[t] **Y.S. Fu,**[l] **D. Fudenberg,**[q,11] **P. Gautam,**[r,12] **R. Gornea,**[f,9] **G. Gratta,**[q] **C. Hall,**[u] **E.V. Hansen,**[r,8] **J. Hoessl,**[e] **P. Hufschmidt,**[e] **M. Hughes,**[b] **A. Iverson,**[n] **A. Jamil,**[v,13] **C. Jessiman,**[f] **M.J. Jewell,**[q,14] **A. Johnson,**[j] **A. Karelin,**[h] **L.J. Kaufman,**[j,15] **T. Koffas,**[f] **R. Krücken,**[k,16] **A. Kuchenkov,**[h] **K.S. Kumar,**[s] **Y. Lan,**[k] **A. Larson,**[i] **B.G. Lenardo,**[q,17] **D.S. Leonard,**[w] **G.S. Li,**[l] **C. Licciardi,**[m] **Y.H. Lin,**[r,18] **R. MacLellan,**[i,19] **T. McElroy,**[d] **T. Michel,**[e] **B. Mong,**[j] **D.C. Moore,**[v] **K. Murray,**[d] **O. Njoya,**[x] **O. Nusair,**[b] **A. Odian,**[j] **A. Perna,**[m] **A. Piepke,**[b] **A. Pocar,**[s] **F. Retière,**[k] **A.L. Robinson,**[m] **P.C. Rowson,**[j] **J. Runge,**[g] **S. Schmidt,**[e] **D. Sinclair,**[f,9] **K. Skarpaas,**[j] **A.K. Soma,**[r] **V. Stekhanov,**[h] **M. Tarka,**[s,20] **S. Thibado,**[s] **J. Todd,**[n] **T. Tolba,**[l,21] **T.I. Totev,**[d]

[1]Permanent address: King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia
[2]Present address: Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506, U.S.A.
[3]Present address: Witmem Technology Co., Ltd., No.56 Beisihuan West Road, Beijing, China
[4]Now at: TRIUMF, Vancouver, British Columbia V6T 2A3, Canada
[5]Also at SNOLAB, Sudbury, ON, Canada
[6]Present address: Carleton University, Ottawa, Ontario K1S 5B6, Canada
[7]Deceased
[8]Present address: Commissariat à l'Energie Atomique et aux énergies alternatives, France
[9]Now at: Oak Ridge National Laboratory, Oak Ridge, TN, U.S.A.
[10]Present address: Department of Physics at the University of California, Berkeley, California 94720, U.S.A..
[11]Present address: Qventus, 295 Bernardo Ave, Suite 200, Mountain View, California 94043, U.S.A.
[12]Present address: Department of Physics, University of Virginia, Charlottesville, VA 22904
[13]Present address: Department of Physics, Princeton University, Princeton, New Jersey, U.S.A.
[14]Present address: Wright Laboratory, Department of Physics, Yale University, New Haven, Connecticut 06511, U.S.A.
[15]Also at Physics Department and CEEM, Indiana University, Bloomington, IN, U.S.A.
[16]Present address: Lawrence Berkeley National Laboratory, Berkeley, CA, U.S.A.
[17]Present Address: SLAC National Accelerator Laboratory, Menlo Park, CA, U.S.A.
[18]Present address: SNOLAB, Sudbury, ON, Canada
[19]Present address: Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506, U.S.A.
[20]Present address: SCIPP, University of California, Santa Cruz, CA, U.S.A.
[21]Present address: Institute for Experimental Physics, Hamburg University, 22761 Hamburg, Germany

2023 JINST 18 P06005

https://doi.org/10.1088/1748-0221/18/06/P06005

R. Tsang,[b] B. Veenstra,[f] V. Veeraraghavan,[b,22] P. Vogel,[y] J.-L. Vuilleumier,[z] M. Wagenpfeil,[e] J. Watkins,[f] M. Weber,[q,23] L.J. Wen,[l] U. Wichoski,[m] G. Wrede,[e] S.X. Wu,[q,24] Q. Xia,[v,16] D.R. Yahne,[n] Y.-R. Yen,[r,25] O.Ya. Zeldovich[h] and T. Ziegler[e]

[a] *Physics Department, University of Illinois, Urbana-Champaign, Illinois 61801, U.S.A.*

[b] *Department of Physics and Astronomy, University of Alabama, Tuscaloosa, Alabama 35487, U.S.A.*

[c] *Physics Department, University of California, San Diego, La Jolla, CA 92093, U.S.A.*

[d] *Physics Department, McGill University, Montreal, Quebec H3A 2T8, Canada*

[e] *Erlangen Centre for Astroparticle Physics (ECAP), Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen 91058, Germany*

[f] *Physics Department, Carleton University, Ottawa, Ontario K1S 5B6, Canada*

[g] *Department of Physics, Duke University, and Triangle Universities Nuclear Laboratory (TUNL), Durham, North Carolina 27708, U.S.A.*

[h] *Institute for Theoretical and Experimental Physics named by A.I. Alikhanov of National Research Centre "Kurchatov Institute," Moscow 117218, Russia*[26]

[i] *Department of Physics, University of South Dakota, Vermillion, South Dakota 57069, U.S.A.*

[j] *SLAC National Accelerator Laboratory, Menlo Park, California 94025, U.S.A.*

[k] *TRIUMF, Vancouver, British Columbia V6T 2A3, Canada*

[l] *Institute of High Energy Physics, Beijing 100049, China*

[m] *Department of Physics, Laurentian University, Sudbury, Ontario P3E 2C6, Canada*

[n] *Physics Department, Colorado State University, Fort Collins, Colorado 80523, U.S.A.*

[o] *Department of Physics and Physical Oceanography, University of North Carolina at Wilmington, Wilmington, NC 28403, U.S.A.*

[p] *Physics Department and CEEM, Indiana University, Bloomington, Indiana 47405, U.S.A.*

[q] *Physics Department, Stanford University, Stanford, California 94305, U.S.A.*

[r] *Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, U.S.A.*

[s] *Amherst Center for Fundamental Interactions and Physics Department, University of Massachusetts, Amherst, MA 01003, U.S.A.*

[t] *Technische Universität München, Physikdepartment and Excellence Cluster Universe, Garching 80805, Germany*

[u] *Physics Department, University of Maryland, College Park, Maryland 20742, U.S.A.*

[v] *Wright Laboratory, Department of Physics, Yale University, New Haven, Connecticut 06511, U.S.A.*

[w] *IBS Center for Underground Physics, Daejeon 34126, Korea*

[x] *Department of Physics and Astronomy, Stony Brook University, SUNY, Stony Brook, New York 11794, U.S.A.*

[y] *Kellogg Lab, Caltech, Pasadena, California 91125, U.S.A.*

[z] *LHEP, Albert Einstein Center, University of Bern, Bern, Switzerland*

E-mail: sli135uiuc@126.com, iostrovskiy@ua.edu

[22] Present Address: Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, U.S.A.

[23] Present address: Descartes Labs, 100 North Guadalupe, Santa Fe, New Mexico 87501, U.S.A.

[24] Present Address: Canon Medical Research US Inc., Vernon Hills, IL, U.S.A.

[25] Present address: Los Angeles, CA 90025, U.S.A.

[26] Now a division of National Research Center "Kurchatov Institute," Moscow 123182, Russia

* Corresponding author.

2023 JINST 18 P06005

Abstract: Generative Adversarial Networks trained on samples of simulated or actual events have been proposed as a way of generating large simulated datasets at a reduced computational cost. In this work, a novel approach to perform the simulation of photodetector signals from the time projection chamber of the EXO-200 experiment is demonstrated. The method is based on a Wasserstein Generative Adversarial Network — a deep learning technique allowing for implicit non-parametric estimation of the population distribution for a given set of objects. Our network is trained on real calibration data using raw scintillation waveforms as input. We find that it is able to produce high-quality simulated waveforms an order of magnitude faster than the traditional simulation approach and, importantly, generalize from the training sample and discern salient high-level features of the data. In particular, the network correctly deduces position dependency of scintillation light response in the detector and correctly recognizes dead photodetector channels. The network output is then integrated into the EXO-200 analysis framework to show that the standard EXO-200 reconstruction routine processes the simulated waveforms to produce energy distributions comparable to that of real waveforms. Finally, the remaining discrepancies and potential ways to improve the approach further are highlighted.

Keywords: Analysis and statistical methods; Double-beta decay detectors; Simulation methods and programs; Time projection chambers

2023 JINST 18 P06005

Contents

2023 JINST 18 P06005

1 Introduction

Computer simulations play a crucial role in many aspects of experimental nuclear and particle physics, including detector design optimization, data analysis, and new physics searches. A typical simulation uses a Monte Carlo approach that starts with a generation of a primary particle that is then propagated through the detailed detector geometry taking into account the stochastic nature of relevant physics processes and detector responses. For simulation of the response to scintillation light produced by ionizing radiation, tens of thousands of photons need to be generated for each O(MeV) energy deposition in liquid xenon. Since the trajectory of each photon is tracked throughout its propagation by simulation packages, such as Geant4 [1], the process is computationally expensive and time consuming. Recent development has significantly sped up the scintillation light simulation through the use of software packages that utilize graphical processing units (GPUs), like Chroma [2–5], but even these highly parallel simulations still require large computational resources. Another issue is that optical properties of materials and detector geometry are often not precisely known, which contributes to differences between the simulation and experimental detector response.

Recent developments in machine learning based generative models offer alternative approaches to physics event generation, including Generative Adversarial Networks (GANs) [6], Variational Autoencoders (VAEs) [7], and Normalizing Flows (NFs) [8]. These techniques allow for non-parametric learning of the data distribution, with sampling being as fast as a single forward pass through the neural network. GAN was invented for computer image generation. It consists of two competing networks, a generator and a discriminator. While the generator does its best to mimic the training images, the discriminator aims to separate the real images from the generated image [9]. Successful training of the

GAN network can lead to the generation of new images indistinguishable from the training images. An obvious caveat is that training images must be available, either from the specific detector to be studied or from essentially similar detectors. Several groups have demonstrated GANs as a tool for fast simulation of Cherenkov detectors [10], muon production through the interaction of a proton beam with dense targets [11], liquid argon time projection chambers [12], and high-granularity calorimeters [13].

In this work, we apply the GAN technique to the simulation of scintillation light in the EXO-200 experiment [14]. EXO-200 is a 175-kg liquid xenon (LXe) detector built to search for the neutrinoless double beta decay of ^{136}Xe. The scintillation light is collected by two planes of avalanche photo-diodes (APDs). While some progress was made on developing a detailed optical simulation of the EXO-200 detector, the discrepancies between data and simulation, likely caused by poorly known optical properties, and computational costs of photon tracking through a complex detector geometry led to EXO-200 using a simplified, parametric optical simulation of the overall light yield per one array of APDs. In this work, we demonstrate that one can train a GAN network directly with calibration data from EXO-200, bypassing the needs for detailed knowledge of optical properties and detector geometry. Once well-trained, the generator is able to produce accurate response of individual APDs at given positions and energies, with better fidelity and faster speed compared with conventional MC simulation. Section 2 provides a brief overview of the EXO-200 detector, its simulation and event reconstruction. Section 3 describes the GAN models developed for generating raw signals deposited on APD channels, including section 3.4 that explains our approach to simulating APD signals with GANs. Results of the simulations are presented in section 4. The last section summarizes the findings and discusses the limitations and future directions of the work.

2 The EXO-200 experiment

The EXO-200 detector is a cylindrical time projection chamber containing 175 kg of LXe enriched to 80.6% in ^{136}Xe. The chamber is divided into two equal drift volumes by a photo-etched phosphor bronze cathode plane in the middle, as shown in figure 1. At each end of the chamber there are two instrumented wire planes crossed at 60° to measure charge. When ionizing radiation interacts in liquid xenon, the produced electrons are drifted towards the wire planes by a main drift field applied between the wires and the cathode plane. The V wires are located closer to the cathode and measure the induction of the passing electrons, which are collected onto the U wires that are located 9 mm behind. The U- and V-wire planes together have 91.8% optical transparency. The APDs are located on a platter behind the wire planes looking at the drift volume. The APDs combine high quantum efficiency for the scintillation light with ultra-low levels of radioactivity [15]. The wire signals provide two-dimensional position information of the event. The position in the drift direction can be obtained using the known drift speed and the time difference between the light signal collected by the APDs and charge signals collected by the U wires. Given the drift speed and the size of the detector, the maximum drift time is on the order of 120 µs. The charge and light signals in LXe are anti-correlated [16, 17]. Consequently, the two signals are combined for energy measurements to achieve optimal energy resolution. Both the wire and APD signals are read out by charge-sensitive preamplifiers outside the lead shielding. A detailed description of the detector can be found in [14].

When a detector trigger condition is met, the data acquisition (DAQ) system records digitized raw waveforms of 152 charge channels and 74 light channels. Each waveform consists of 2048 samples taken

2023 JINST 18 P06005

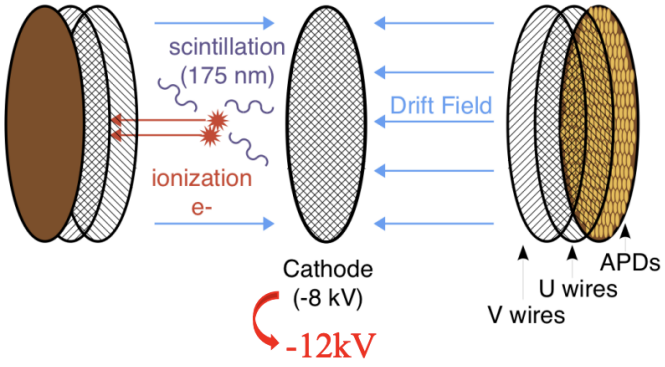


Figure 1. Schematic of the inner EXO-200 TPC. The cathode voltage was set to −8 kV (−12 kV) for Phase I(II) operations. The ionization electrons are drifted to the U wire plane that is grounded and serves as an anode. The scintillation light of ∼ 175 nm [18] is detected by APDs behind the grids.

at 1 MS/s with 1024 samples occurring before the trigger and 1024 samples after the trigger. The raw waveforms are processed by an algorithm that reconstructs the energy depositions inside the detector. Briefly, an initial signal-finding stage identifies channels containing a signal above a given noise threshold. After the signal-finding stage is the parameter estimation stage, where each of the identified signals is analyzed to extract parameters relevant for the analysis. Then, the set of signals and signal parameters for each event is combined to determine event topology and event energy. The reconstructed signals are calibrated by radioactive sources emitting γ rays with known energy. The charge and light signals are calibrated separately, then combined to form a rotated energy scale with improved energy resolution.

To understand the detector response to energy deposits in the detector volume, a Geant4-based Monte Carlo simulation is employed. The simulation can accurately simulate the charge depositions. It does not produce an accurate response for individual APD channels because of the uncertainty of optical properties and difficulties with implementing the anti-correlation between charge and light responses. Instead, it avoids the computational cost of tracking individual photons by using a parameterized light response function to simulate the light yield on each plane of the APDs based on the position of the energy deposit in the detector. The light response is then evenly distributed among all APD channels of the given plane with randomized noise added to each waveform. This light simulation is used only to approximately simulate the light reconstruction threshold and is not used to determine the Monte Carlo energy. Following other successful applications of Chroma to optical simulations [2, 3], a Chroma-based photon tracking was later developed in EXO-200. This approach matched with the data better but still required tuning of material optical properties and has not yet been published. The difficulties with the scintillation light simulation motivate us to apply the GAN approach to improve the simulation quality and speed.

3 GAN Description

3.1 APD waveform image

Since the GAN framework was developed for image generation, we convert the 74 channels of APD digitized waveforms into images for ease of integration into the software framework. We remove the baseline of the waveforms by subtracting the average of the first 300 µs. We do not scale the

2023 JINST 18 P06005

waveforms, as it did not improve training. A typical waveform's values range between −100 to 500 ADC units. To reduce the number of dimensions of the target space, we select 350 µs samples of the waveform around the signal. Since the rest of the samples only contain noise, this achieves a substantial reduction in complexity without an appreciable reduction in accuracy. The location of the signal depends on the trigger type — at the center of the waveform for events triggered by APDs and off-center but within the maximal drift time for events triggered by wires. The resulting image of an event consists of 74 channels, each having 350 entries, as shown in figure 2.

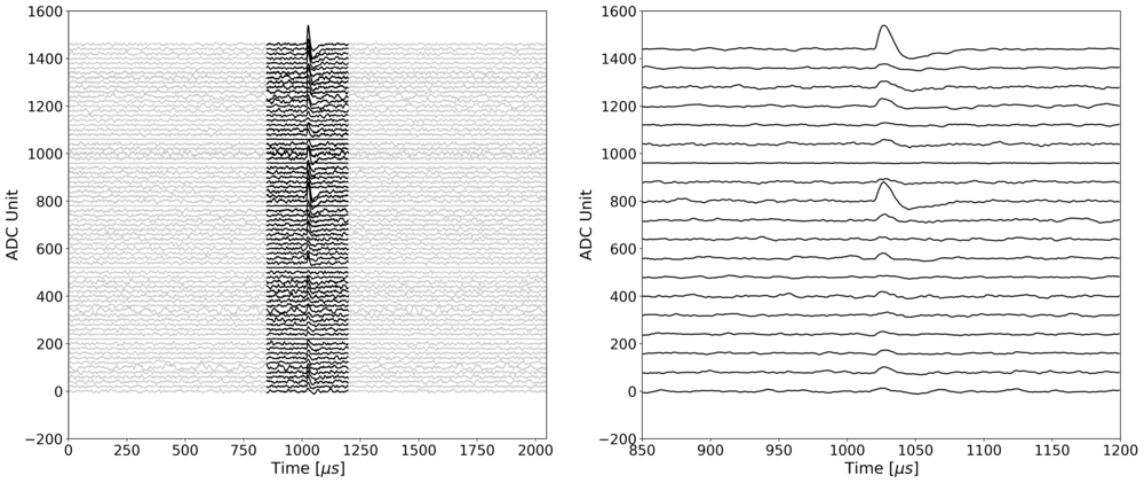


Figure 2. (Left) Example waveforms of the APDs showing pulses from a cluster at certain position and energy. For clarity, the signals on each APD are offset vertically in order of channel number and separated by a constant of 20 units in the y direction. Due to computational reasons, raw waveforms of 2048 µs (gray) are truncated to 350 µs (black) before being fed to the neural network. (Right) Same as the left panel but zoomed-in around the peak positions and showing every fourth APD channel for clarity.

3.2 Network structure

To simulate the raw waveforms of APDs, we use the Keras framework [19] to build a network based on Wasserstein GAN [6] with label conditioning. The network consists of a discriminator and a generator. The generator takes in an array of uniform random values, specified position and energy of the scintillation cluster as inputs, then generates artificial waveforms mimicking the real APD responses. The discriminator measures the similarity between the generated and real samples, providing feedback to the generator to improve its performance. Training of such an adversarial framework allows the generator to produce data-like samples $x_g = G(z)$ out of noise z. The original GAN [9] architecture has shown impressive results but turned out to be unstable and hard to monitor during the training process. The Wassertein GAN is, with its improvement, a network that allows for a stabilized training procedure by delivering adequate gradients to the generator, which provide a meaningful loss metric not susceptible to training collapse. The Wassertein GAN is enforced by the Lipschitz constraint on the discriminator to calculate the 1-Wasserstein (or, simply, Wasserstein) distance. A differentiable function f is 1-Lipschitz if and only if it has gradients with norm at most 1 everywhere, $\forall x : \ |f'(x)| \leq 1$. A later study [20] proves that points interpolated between the real and generated data should have a gradient norm of 1 for f. Hence, instead of applying weight

2023 JINST 18 P06005

clipping which manually forces hyperparameters of network in the Lipschitz constraint, Wassertein GAN gradient penalty (WGAN-GP) penalizes the model if the gradient norm moves away from its target norm value of 1, as shown in eq. 3.1 from [20].

$$L = \underbrace{\mathop{\mathbb{E}}_{\tilde{x}\sim P_g}[D(\tilde{x})] - \mathop{\mathbb{E}}_{x\sim P_r}[D(x)]}_{\text{Wasserstein distance}} + \underbrace{\lambda \mathop{\mathbb{E}}_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla D(\hat{x})\|_2 - 1)^2]}_{\text{gradient penalty}} \tag{3.1}$$

where D is the discriminator, $\hat{x} = \epsilon x + (1-\epsilon)\tilde{x}, \epsilon \in U(0,1)$, λ is the gradient penalty's weighting coefficient, and $\|\|_2$ denotes the Euclidean norm. The gradient penalty term, $(\|\nabla D(\hat{x})\|_2 - 1)^2$, encourages the norm of the gradient to go towards 1. The point x used to calculate the gradient norm is any point sampled between the GAN-generated distribution, P_g, and real data distribution, P_r. A gradient penalty is a soft version of the Lipschitz constraint that removes the undesirable behaviour of gradient explosion/vanishing when the weight clipping parameter is not carefully tuned in the earlier Wassertein GAN design.

Wasserstein Distance is also known as Earth mover's distance, as it defines the cost for moving a distribution onto a target distribution using optimal transport. The original GAN is trained to minimize the Jensen-Shannon Divergence (JD) [9]. Comparing with JD, the Wasserstein Distance has the following advantages:

- Wasserstein Distance is a continuous and almost differentiable function which is easier to optimize.
- As the discriminator gets better, JD locally saturates and thus the gradient becomes zero and vanishes.
- Wasserstein Distance is a meaningful function as its converges to 0 while two distributions are getting closer together and diverges when they are moving apart.
- Wasserstein Distance is more stable than JD, and the model is hard to collapse when using Wasserstein Distance as the objective function.

Lastly, to generate samples with specific characteristics, in this case, the position and energy of the scintillation, the generator is provided with the labels through label conditioning by an Auxilliary Classifier, as described in [21]. Using this Generator model, the training process becomes more stable and can now be used to generate images of a specific type using the class label.

The schematic diagram in figure 3 shows the overall structure of the GAN model. The detailed layer structure of the generator and the discriminator network can be found in appendix A. The generator dependency can be written as $G = G(N, E, P)$, where N is random noise and E and P are physics labels corresponding to scintillation energy E and the deposit coordinates $P = (P_x, P_y, P_z)$. Because we use calibration data for training the network, we do not know the precise energy and position of the scintillation cluster, and need to use the EXO-200 reconstruction algorithm to provide labels for the events. There are two choices for the scintillation energy label, one is the reconstructed scintillation energy and the other is a combined, also called rotated, reconstructed energy that uses both charge and light information. Although the rotated energy has better energy resolution of ~1.2% at the Q value of the ^{136}Xe double-beta decay [22], it removes the anti-correlation between the charge

2023 JINST 18 P06005

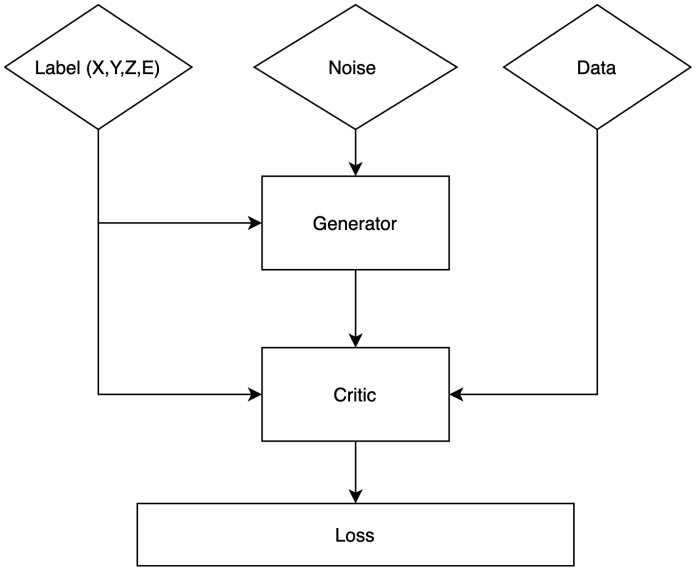


Figure 3. The schematic training diagram of generator and discriminator for the GAN network. Labels, (X,Y,Z,E), are also inputs of discriminator to allow training of generator with specific characteristics.

and light responses and so is a biased estimate of an event's true scintillation energy. The reconstructed scintillation energy, despite its larger energy resolution of $\sim 5.1\%$ at the same energy, is thus a better estimate of the true scintillation energy and is chosen as the energy label for the data in this work.

The physics labels are tiled to the convolutional layers as shown in figure 10 in appendix A. Each value of the input dimensions, x_i, is copied to a new convolutional layer filled with that value, so four additional layers are concatenated to the image input, changing the dimension from $(74, 350, 1)$ to $(74, 350, 5)$. Furthermore, we also concatenate the label vector to the dense representation layers at the end of the convolutional part of the discriminator network. This design is inspired by the concept of residual learning [23] which simplifies learning when the output is expected to be close to input.

3.3 Training and test datasets

The training dataset is obtained from ^{228}Th calibration runs of the EXO-200 experiment. The data are separated into two phases: phase I, which took place from 2011 to 2014, and phase II (2016 to 2018). The noise of the APD channels is different in the two phases because the frontend electronics were upgraded at the start of phase II. In this work, we focus on the simulation of the phase II data. The training set is created by randomly picking events from the calibration runs. The events are selected to only have one scintillation and one charge cluster, with fully reconstructed energy and position. The training objects are created by combining the waveforms of all individual APD channels to form an image. The pixel values of the image correspond to the waveforms' amplitudes at a particular time. The reconstructed energy and position serve as event labels. To reduce bias, the energy distribution of the training sets is flattened. The same method was applied to flatten the spatial distribution of events. However, since the calibration source positions are limited, and certain regions of the detector have very few events due to a limited amount of nearby source deployments, we cannot completely flatten the training dataset in the spatial dimension, as shown in figure 4. The final training set contains 215,789 events from phase II. Similarly, a test dataset is obtained in the

2023 JINST 18 P06005

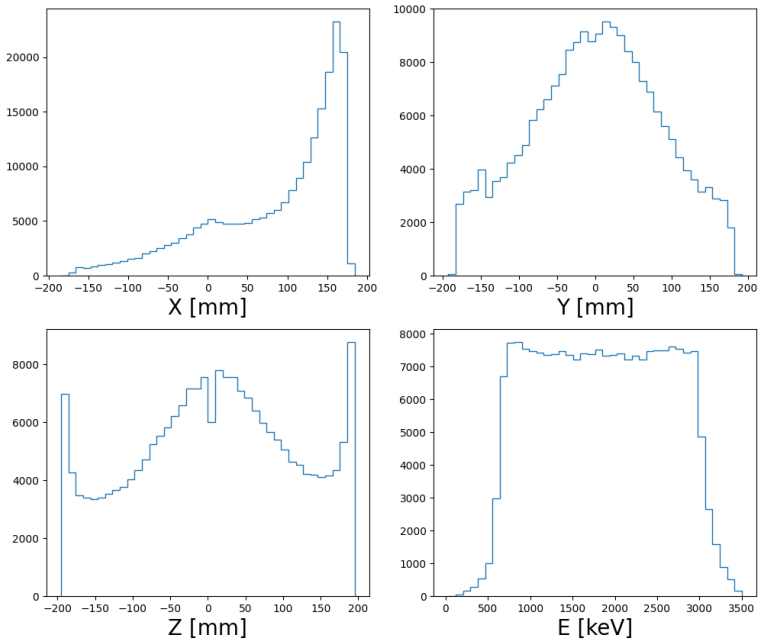


Figure 4. The spatial and energy distribution of training dataset.

same way after event selection for the training dataset. The ratio of the total number of events of test and training datasets is close to 7:3.

3.4 Network training process

Because a GAN contains two separate networks, its training algorithm must deal with two different kinds of training. Typically, the algorithm proceeds in alternating training of the discriminator and the generator. During training of the discriminator, the real data and the generated data serve as inputs. Then the training switches to the generator. The generator performance improves with the discriminator providing the feedback. The cycle continues until the Wasserstein distance becomes close to zero, indicating that the generator produces objects that the discriminator cannot effectively distinguish from the real ones. The convergence of the GAN network can be difficult to achieve and identify, as it requires a delicate balance between the discriminator and the generator networks.

In this work, we train both generator and discriminator using the Adam optimizer [24] with the learning rate starting at 10^{-4} at the beginning of the training process and with the optimizer parameters $\beta_1 = 0.5$ and $\beta_2 = 0.9$. The loss function coefficient number $\lambda = 10$. We train the discriminator for five iterations, before updating the generator for one iteration. In each iteration step, the network is trained with a 20 events batch. One epoch counts as a loop over the entire phase II training set. We train the networks on a computing cluster in successive batch jobs running for 10 epochs each, while decreasing training rate gradually from 10^{-4} to 10^{-8} for the discriminator and 5 times faster for the generator.

2023 JINST 18 P06005

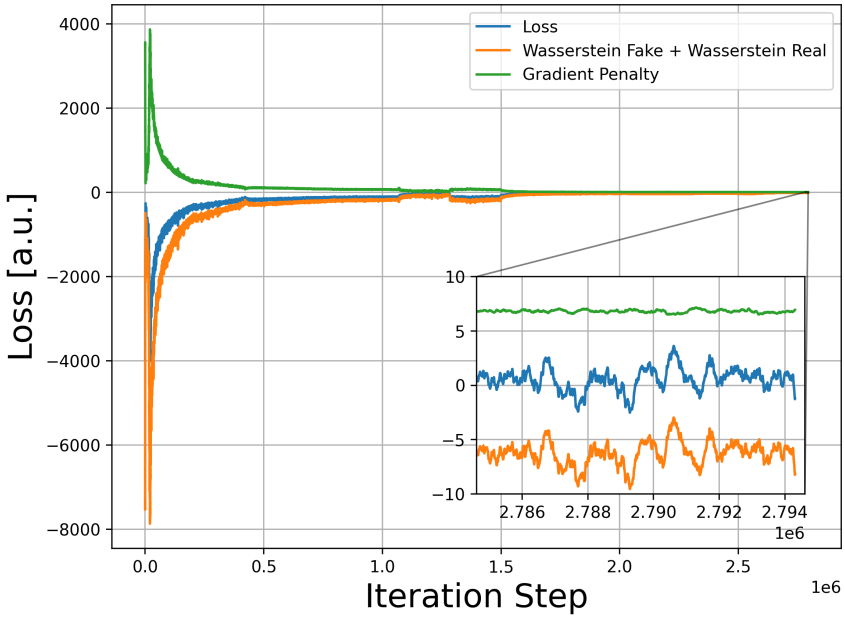


Figure 5. Loss curves of the moving average of 300 iterations for total discriminator loss during the training (blue), the Wasserstein distance (orange) and gradient penalty (green). The Wasserstein term contains "wasserstein real" as $-D(x)$, the outputs of real waveforms and "wasserstein fake" as $D(G(z))$, the outputs from GAN-generated waveforms.

The gradient penalty and Wasserstein distance drop quickly during the initial training, but the rate of decrease slows down and gradually approaches zero, as shown in figure 5. The training results were evaluated during the training using the metrics described in the next section. We stop the training if absolute values of the gradient penalty and Wasserstein distance are less than 10, and the loss is close to zero, which typically occurs after 2.8 million batches, or about 260 epochs. The training process takes about 500 GPU-hours on the Nvidia GeForce 1080Ti GPU with 11 GB RAM.

4 Results and validation

For evaluating the model, we pair up the events from the test dataset with the ones generated with the GAN for the same energy and position parameters. The generated events not only need to "look similar" but also need to have similar features such as the energy and position information after reconstruction. Thus we introduce a set of metrics including the raw waveforms, the features extracted from the pulses, such as amplitudes and spatial dependency, and reconstructed physical information, such as scintillation and rotated energy. The validation tests for the training results are conducted in following levels:

- Raw waveform level: directly compare raw waveforms
- Signal level: compare signal features and its spatial dependence
- Reconstruction level: compare information from the EXO-200 reconstruction, e.g., scintillation energy spectrum.

2023 JINST 18 P06005

Because the length of the GAN-generated waveforms is 350 µs, shorter than the full EXO-200 event size of 2048 µs, they are stitched back to the full waveform with the section centered around the event trigger replaced by the generated waveforms. We visually compare the generated and real events in different energy ranges, low (< 1 MeV), middle (1–2 MeV), and high (> 2 MeV), and find that the generated waveforms reproduce the main features of the real events. The signals in each of the waveforms are aligned in time as one would expect from scintillation events. The signal shapes, channel noise, and amplitude distribution in the generated image is also similar to the real events. Figure 6 shows a sample event with energy > 2 MeV.

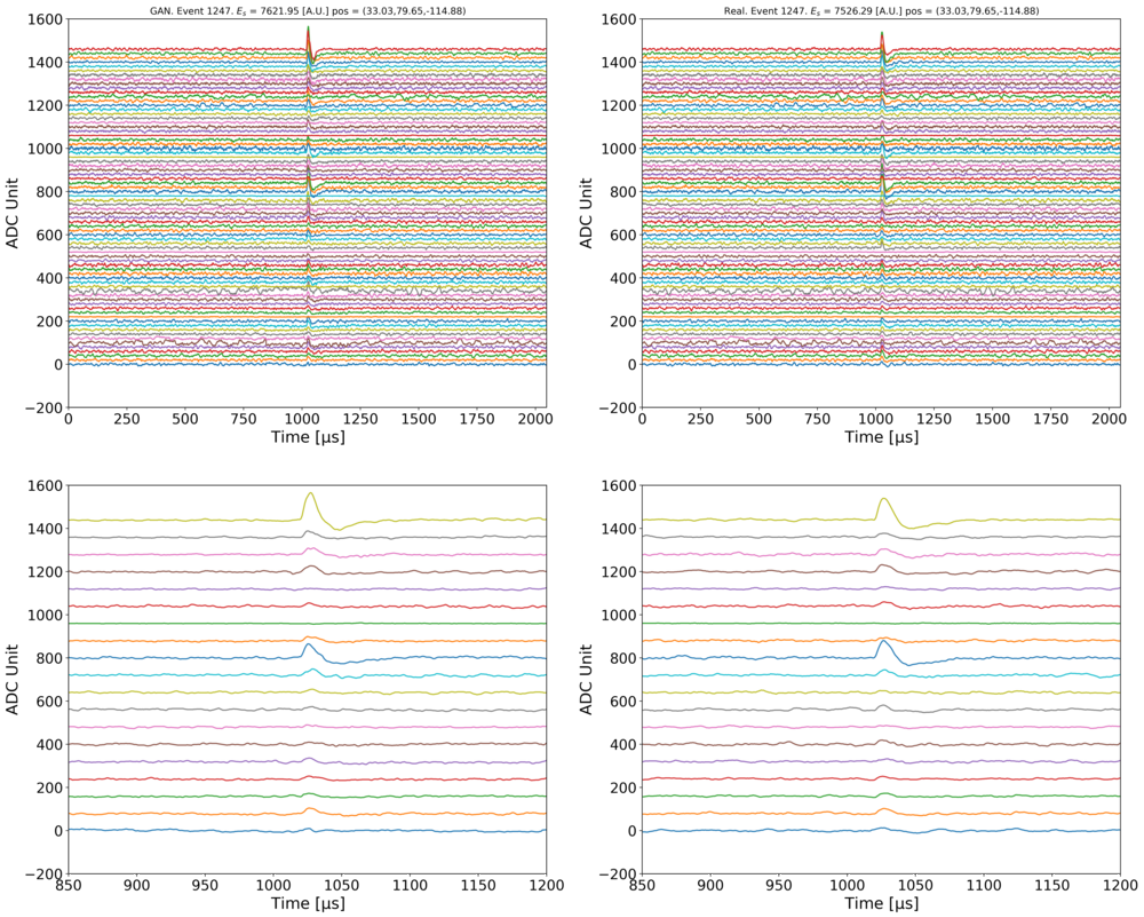
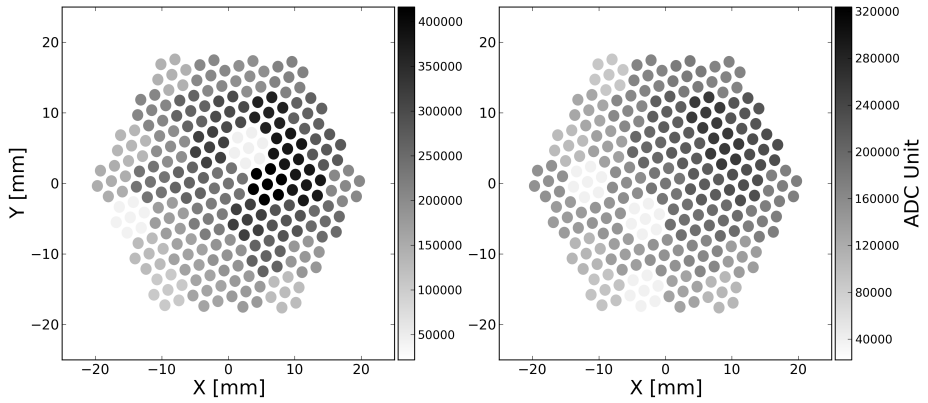


Figure 6. Example of the generated (top left) and real (top right) APD signal waveforms. Vertical and horizontal axes correspond to the amplitude + channel number × A.U. and time bins, respectively. The waveforms are obtained for the same values of the conditional variables and look qualitatively very similar. The two bottom panels are the same as the top ones but zoomed-in around the peak positions and showing every fourth APD channel for clarity.

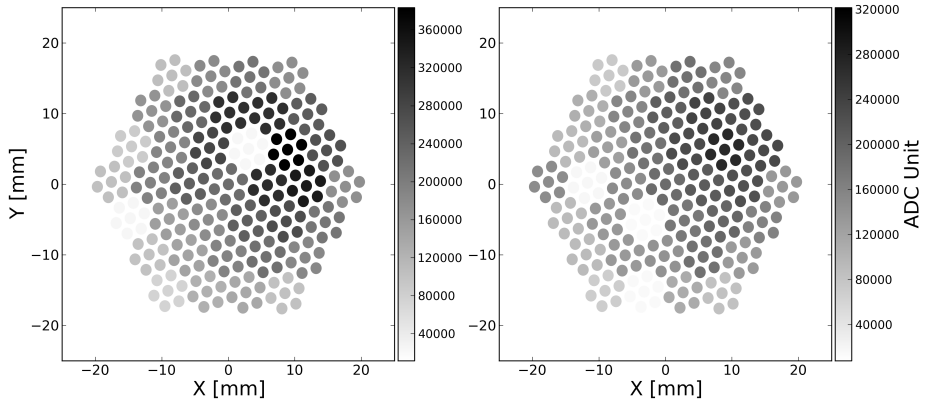
Another feature that is easy to spot is that several waveforms do not have signals. These correspond to five channels disconnected due to excessive leakage current of the APDs. We can examine the disconnected channels further by looking at the average signal amplitude of all events in one typical calibration run. Figure 7 shows the pattern of APDs in two planes and its corresponding

2023 JINST 18 P06005

signal amplitude for both GAN and real waveforms. The APDs are arranged in two hexagonal planes and grouped in gangs of 6 or 7. The amplitudes are represented in grey-scale in the plot with darker color representing larger amplitudes. As shown in figure 7(a) and 7(b), the model finds the disconnected APD channels and reproduces an average amplitude pattern similar to the real calibration data. The average amplitude of the disconnected APDs in the generated waveforms does not completely go to zero, as some GAN-generated events have small residual pulses in these channels. The average amplitudes are around 5 ADC units slightly larger than the noise RMS value of 2 ADC units, but negligible compared to amplitudes of connected channels.



(a) APD channels' luminosity of GAN-generated waveforms.



(b) APD channels' luminosity of real waveforms.

Figure 7. Summed amplitude of all events in a run for each of the APD channels. GAN-generated (top) and real (bottom) data are shown. Two APD planes are shown separately. Darker color means higher amplitude. White areas show the dead channels. We can identify the disconnected channels in GAN data (a) that matches with calibration data in (b).

To make a more precise evaluation, we apply the EXO-200 reconstruction algorithm to both GAN-generated waveforms and the real data. Since the reconstruction requires wire signals to determine the position of the events, we copied the wire signals of the real data to the generated APD waveforms. Since the wire signals are identical, the reconstruction returns the same position and charge energy information for both the GAN and real data. The only difference comes from the reconstructed scintillation energy. To extract the scintillation energy, the reconstruction algorithm

2023 JINST 18 P06005

adds all waveforms from plane 1 and plane 2 to produce two sum waveforms, then fits the sum waveforms to extract the summed amplitudes. In figure 8, we plot the spatial dependence of the sum waveform amplitude of each APD plane for both real data and data generated with the energy and position as labels. Plane 1 is located at z = 200 mm while plane 2 is located at $z = -200$ mm, thus the plane 1 amplitude becomes higher when a cluster is closer to the plane 1 while the plane 2 amplitude becomes lower. There is no large R or ϕ dependence due to high reflectivity of the PTFE reflector surrounding the TPC. The plot shows that GAN accurately reproduces the plane 1 and plane 2 average amplitudes and that standard deviation matches well with the real data in all three cylindrical spatial dimensions.

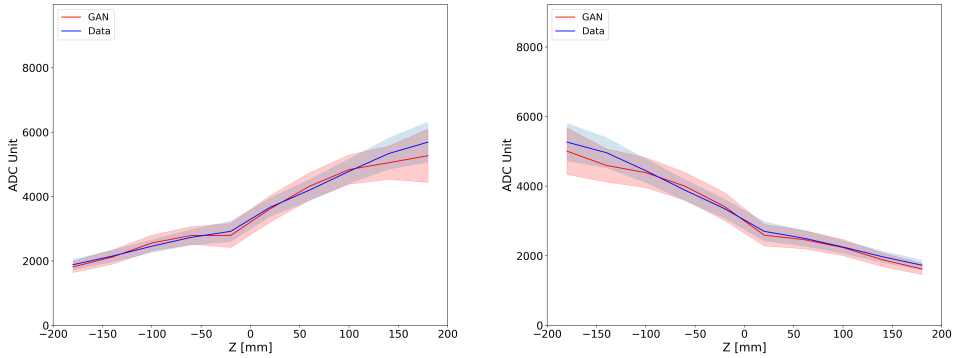
After signal fitting, the extracted amplitudes go through energy calibration procedure to convert them into scintillation energy measurements. The scintillation spectra of the generated and real data for one ^{228}Th calibration run are shown in figure 9(a). The agreement is fairly good across the entire energy range, but the GAN generated spectrum has slightly worse energy resolution. The energy resolution (σ/E) of the 2.614 MeV ^{228}Th peak is 7.6% for real data, and 9.1% for GAN-generated data. The degradation of the resolution is likely a consequence of the uncertainty of the energy labels used to train the network, rather than the inherent issue with the GAN network.

As mentioned in section 2, using a linear combination of the charge and light signals can improve the energy resolution of the detector. In a 2-D scatter plot of scintillation energy and charge energy, this corresponds to a projection to a rotated axis, as shown in figure 9(b). The GAN-generated data reproduces the anti-correlation between the charge and light signals, confirming that the good agreement between true and generated scintillation energies is achieved on the event-by-event basis. The optimal rotation angle, which in EXO-200 was treated as a free parameter chosen to optimize rotated resolution, is slightly different from the real data. If we project the GAN and real datasets onto their optimal axes, the rotated spectra match well with each other, as shown in figure 9(c). The energy resolution of the ^{228}Th peak is 2.24% for the real data and 2.47% for the GAN generated data.

The energy resolution can be further improved by correcting position dependency of the light response. To that end, the detector is divided into cylindrical voxels with 13 sections in Z and R directions and 8 sections in ϕ direction. The ^{228}Th calibration data in each voxel is fitted to extract the position of the 2.614 MeV peak. The ratio between the voxel peak position and the overall peak position is the correction factor for the voxel. We call this the light map correction [25]. The light map records the correction factor in the reconstruction space. If we apply the light map to both the real and GAN generated data, the energy resolution improves to $\sim$ 1.3% for the real data, and $\sim$2.0% for the GAN data, as shown in figure 9(d). The fact that the light map correction improves the GAN resolution demonstrates that GAN network is able to reproduce, at least to a good degree, the position dependence of the scintillation signal. Like in the previous two tests, the worse resolution for the GAN images is likely also impacted by the imperfect energy labels used in training. If so, then using a light source with known intensity adjustable to span the relevant energy range and deployed at relevant positions inside the detector [26–30] to train the network would lead to better results.

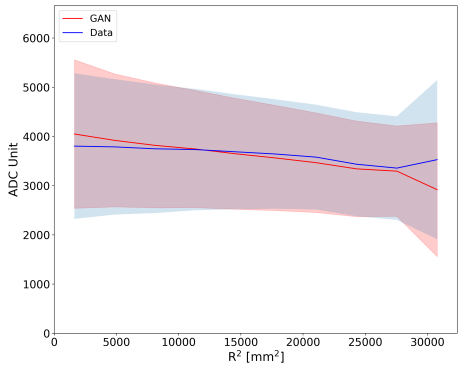
To compare the simulation speed, we generate scintillation waveforms for several thousands ^{228}Th events inside the detector using the EXO-200 Monte Carlo framework and GAN network. The EXO-200 framework first utilizes Geant4 to simulate energy depositions of the ^{228}Th gammas, then uses a fast parametric optical simulation that produces expected number of photons detected by the APD planes, and finally generates the APD waveforms using transfer functions that depend on the

2023 JINST 18 P06005

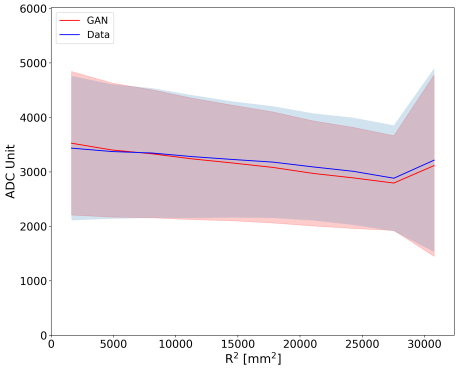


(a) Plane 1 Amplitude vs. Z

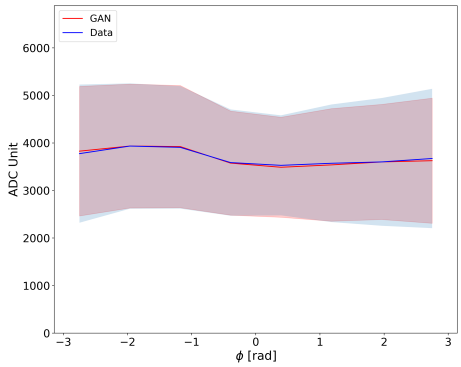
(b) Plane 2 Amplitude vs. Z



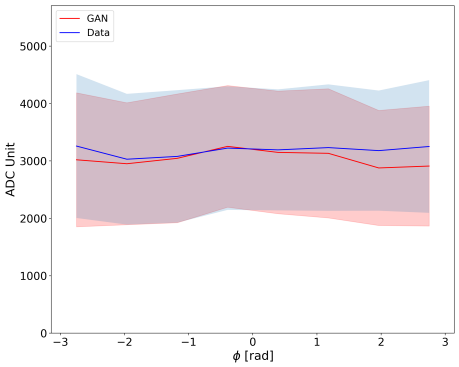
(c) Plane 1 Amplitude vs. R^2



(d) Plane 2 Amplitude vs. R^2



(e) Plane 1 Amplitude vs. ϕ

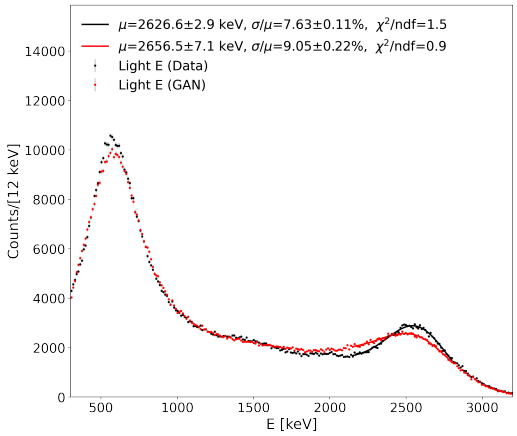


(f) Plane 2 Amplitude vs. ϕ

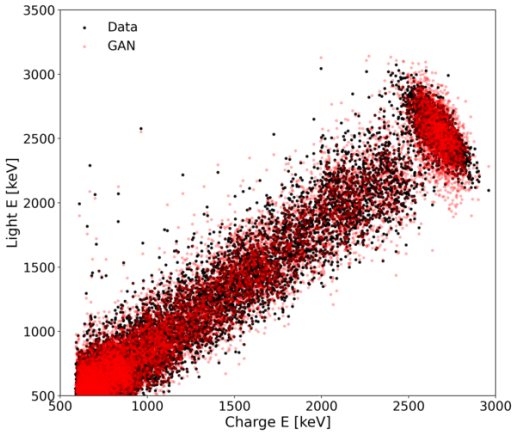
Figure 8. Profiles of the validation metric distributions as the function of the spatial components. For each metric i, we show its average value μ_i (middle thick line), as well as the one standard deviation band of the metric distribution $\mu_i \pm \sigma_i$ (top and bottom band edge). GAN-generated distributions (red) show a strong overlap with real data (blue).

parameters of the APD electronics. By default, an analogous "digitization" process is performed for the wires, but we switch the wire digitization off for a more fair comparison. We observe an average simulation rate of ~ 4.2 events/s, with the Geant4 and digitization steps taking roughly the same time. The test was run on a machine with an Intel Xeon Gold 6226R CPU [31] and 12 GB RAM. In contrast, GAN directly simulates APD waveforms given ^{228}Th event energy and position. We observe an average simulation rate of ~ 69 events/s when ran on Nvidia GeForce 1080Ti GPU [32]

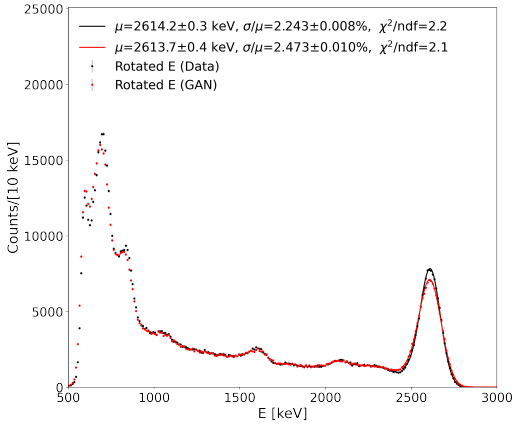
2023 JINST 18 P06005



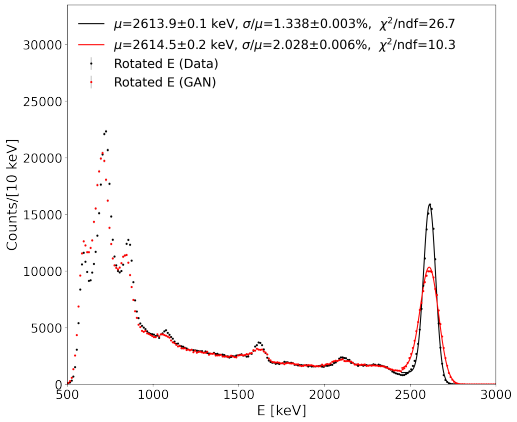
(a) The ^{228}Th light energy spectra of real (black) and GAN-generated (red) events without position correction. Fitted resolutions are also given.



(b) Scatter plot of reconstructed light and charge energies from GAN (red) and real (black) waveforms. Both GAN and real data show anti-correlation between the light and charge signals.



(c) Histograms show the projection of the scatter points from figure 9(b) onto the corresponding rotated axis, $\theta_r = 0.102$ and $\theta_g = 0.082$. The red and blue lines show the fitting results of the rotated E peak.



(d) Histograms show the same projection of events after the light map calibration onto corresponding rotated axis, $\theta_r = 0.224$ and $\theta_g = 0.143$. The red and blue lines show the fitting results of the rotated E peak.

Figure 9. Reconstructed light energy distributions for the GAN-generated and real data.

with 11 GB RAM. While not a precise back-to-back comparison, the test shows that the new approach can simulate light waveforms at a rate that is roughly an order of magnitude faster than with the traditional approach.

5 Conclusions

In this work, we applied Wasserstein GAN technique to generate raw APD signal waveforms for the EXO-200 experiment. Detector calibration data was used as the training samples bypassing the need for computationally intensive scintillation simulation and inaccuracies of the MC simulation models. Using reconstructed event position and energy as labels, the network was successfully

2023 JINST 18 P06005

trained to generate raw waveforms that mimic main features of the real waveforms, including signal arriving time, pulse shape, and channel noise. At the signal level, the network learned the missing channels and spatial dependency of the pulse amplitudes. Furthermore, the GAN could reproduce the energy spectrum of the calibration data after reconstruction. Combined with corresponding charge waveforms, the anti-correlation between the scintillation and charge channels was reproduced, confirming that the good agreement between true and generated scintillation energies is done event-by-event. The energy resolution of the scintillation and rotated spectra are slightly worse for the GAN-generated data than for the real data.

In this study, events that contain single spatially-distinct charge deposition (so-called single-site events) for both training and validation were used. Although there are also multi-site events in the dataset, the only difference between them as far as this work is concerned is the spatial distribution of the scintillation light. Since the GAN was able to reproduce spatial distribution of the signal for single-site events, it is reasonable to expect a similar performance on multi-site events, which are effectively a combination of several independent sites. We concentrated on the single-site events for simplicity in this novel work, leaving the multi-site events as a potential future direction. It should also be noted that since the approach described here requires training on real data, it can not be used to optimize the design of future detectors as effectively as the traditional MC frameworks.

We have demonstrated that using the GAN network and calibration data can be a powerful approach to generate simulation data faster and more accurately than with traditional Monte Carlo simulation. One drawback of the approach is that the labels of the training data are derived from reconstructed scintillation energy, which in case of EXO-200 carries a $\sim$ 5% uncertainty. Experiments that can utilize a set of training events with better truth labels would see better results. For example, some experiments may be able to use an LED or laser source with known, variable intensity, and which can illuminate all relevant parts of the detector. An interesting alternative possibility, which we also consider an avenue for future work, is training the network on both charge and light waveforms simultaneously. If the network is able to deduce the anti-correlation between the two signals, then one could potentially harness the advantage of the improved accuracy of the rotated energy labels.

In summary, GAN is a promising tool to accelerate and improve simulation for particle physics. Using calibration data directly as inputs can simplify the simulation and avoid inaccuracies of the simulation model.

Acknowledgments

This work is supported by a Department of Energy Grant No. DE-SC0019261. EXO-200 is supported by DOE and NSF in the United States, NSERC in Canada, IBS in Korea, DFG in Germany, and CAS and ISTCP in China. EXO-200 data analysis and simulation uses resources of the National Energy Research Scientific Computing Center (NERSC). We gratefully acknowledge the support of Nvidia Corporation with the donation of three Titan Xp GPUs used for the optical simulations.

A Network architecture

The neural network architectures for the discriminator and generator are shown in 10.

2023 JINST 18 P06005

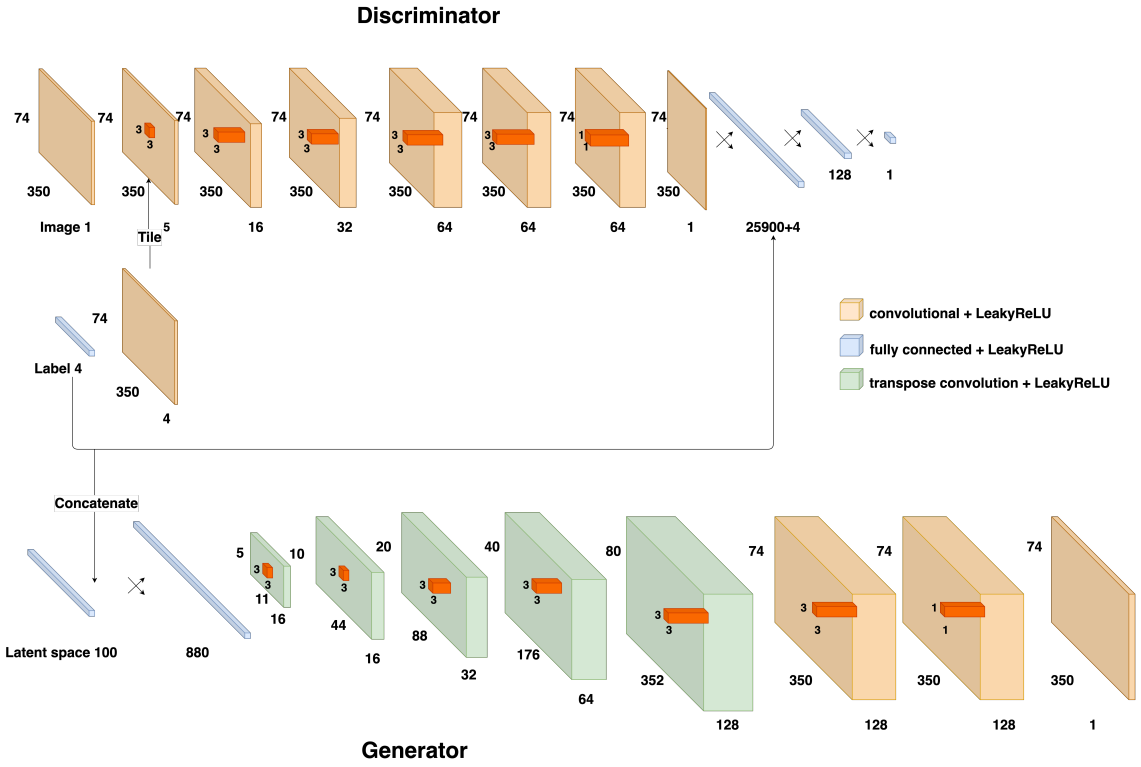


Figure 10. Architecture of the GAN with corresponding kernel size, number of feature maps. The discriminator (generator) has 1,755,333 (1,273,953) trainable parameters.

References

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

[2] S. Seibert and A. LaTorre, *Fast Optical Monte Carlo Simulation With Surface-Based Geometries Using* Chroma, 2011, https://www.tlatorre.com/chroma/_downloads/chroma.pdf.

[3] T. Kaptanoglu, M. Luo, B. Land, A. Bacon and J. Klein, *Spectral Photon Sorting For Large-Scale Cherenkov and Scintillation Detectors*, *Phys. Rev. D* **101** (2020) 072002 [arXiv:1912.10333].

[4] DARWIN collaboration, *GPU-based optical simulation of the DARWIN detector*, 2022 *JINST* **17** P07018 [arXiv:2203.14354].

[5] nEXO collaboration, *nEXO: neutrinoless double beta decay search beyond 10^{28} year half-life sensitivity*, *J. Phys. G* **49** (2022) 015104 [arXiv:2106.16243].

[6] M. Arjovsky, S. Chintala and L. Bottou, *Wasserstein GAN*, [arXiv:1701.07875].

[7] D.P. Kingma and M. Welling, *Auto-Encoding Variational Bayes*, arXiv:1312.6114.

[8] I. Kobyzev, S.J. Prince and M.A. Brubaker, *Normalizing flows: An introduction and review of current methods*, *IEEE Trans. Pattern Anal. Machine Intell.* **43** (2021) 3964.

[9] M. Mirza and S. Osindero, *Conditional Generative Adversarial Nets*, in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Montreal, Canada, 8–13 December 2014, vol. 2, p. 2672–2680 arXiv:1411.1784.

2023 JINST 18 P06005

[10] LHCb collaboration, *Fast Data-Driven Simulation of Cherenkov Detectors Using Generative Adversarial Networks*, *J. Phys. Conf. Ser.* **1525** (2020) 012097 [arXiv:1905.11825].

[11] SHiP collaboration, *Fast simulation of muons produced at the SHiP experiment using Generative Adversarial Networks*, 2019 *JINST* **14** P11028 [arXiv:1909.04451].

[12] P. Lutkus, T. Wongjirad and S. Aeron, *Towards Designing and Exploiting Generative Networks for Neutrino Physics Experiments using Liquid Argon Time Projection Chambers*, in *Proceedings of the 9th International Conference on Learning Representations*, 3–7 May 2022 [arXiv:2204.02496].

[13] G.R. Khattak, S. Vallecorsa, F. Carminati and G.M. Khan, *Fast simulation of a high granularity calorimeter by generative adversarial networks*, *Eur. Phys. J. C* **82** (2022) 386 [arXiv:2109.07388].

[14] M. Auger et al., *The EXO-200 detector, part I: Detector design and construction*, 2012 *JINST* **7** P05010 [arXiv:1202.2192].

[15] R. Neilson et al., *Characterization of large area APDs for the EXO-200 detector*, *Nucl. Instrum. Meth. A* **608** (2009) 68 [arXiv:0906.2499].

[16] E. Aprile, K.L. Giboni, P. Majewski, K. Ni and M. Yamashita, *Observation of Anti-correlation between Scintillation and Ionization for MeV Gamma-Rays in Liquid Xenon*, *Phys. Rev. B* **76** (2007) 014115 [arXiv:0704.1118].

[17] EXO-200 collaboration, *Correlated fluctuations between luminescence and ionization in liquid xenon*, *Phys. Rev. B* **68** (2003) 054201 [hep-ex/0303008].

[18] K. Fujii, Y. Endo, Y. Torigoe, S. Nakamura, T. Haruyama, K. Kasami et al., *High-accuracy measurement of the emission spectrum of liquid xenon in the vacuum ultraviolet region*, *Nucl. Instrum. Meth. A* **795** (2015) 293.

[19] F. Chollet et al., *Keras*, https://github.com/keras-team/keras.

[20] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. Courville, *Improved Training of Wasserstein GANs*, in *Advances in Neural Information Processing Systems*, vol. 30, 2017 [arXiv:1704.00028].

[21] A. Odena, C. Olah and J. Shlens, *Conditional Image Synthesis With Auxiliary Classifier GANs*, in *Proceedings of the 34th International Conference on Machine Learning*, Sydney, NSW, Australia, 6–11 August 2017, pp. 2642–2651, [arXiv:1610.09585].

[22] EXO-200 collaboration, *Search for Neutrinoless Double-β Decay with the Complete EXO-200 Dataset*, *Phys. Rev. Lett.* **123** (2019) 161802 [arXiv:1906.02723].

[23] K. He, X. Zhang, S. Ren and J. Sun, *Deep Residual Learning for Image Recognition*, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, U.S.A., 27–30 June 2016, pp. 770–778 arXiv:1512.03385.

[24] D.P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2014 [arXiv:1412.6980].

[25] EXO-200 collaboration, *Improved measurement of the $2\nu\beta\beta$ half-life of ^{136}Xe with the EXO-200 detector*, *Phys. Rev. C* **89** (2014) 015502 [arXiv:1306.6106].

[26] SNO collaboration, *The Sudbury neutrino observatory*, *Nucl. Instrum. Meth. A* **449** (2000) 172 [nucl-ex/9910016].

[27] J. Liu et al., *Automated calibration system for a high-precision measurement of neutrino mixing angle θ_{13} with the Daya Bay antineutrino detectors*, *Nucl. Instrum. Meth. A* **750** (2014) 19 [arXiv:1305.2248].

[28] Double Chooz collaboration, *Double Chooz Calibration*, *Nucl. Phys. B Proc. Suppl.* **229-232** (2012) 431.

2023 JINST 18 P06005

[29] Borexino collaboration, *Borexino calibrations: Hardware, Methods, and Results*, 2012 *JINST* **7** P10018 [arXiv:1207.4816].

[30] Y. Zhang, J. Liu, M. Xiao, F. Zhang and T. Zhang, *Laser Calibration System in JUNO*, 2019 *JINST* **14** P01009 [arXiv:1811.00354].

[31] https://www.intel.com/content/www/us/en/products/sku/199347/intel-xeon-gold-6226r-processor-22m-cache-2-90-ghz/specifications.html.

[32] https://www.nvidia.com/en-us/geforce/products/10series/geforce-gtx-1080-ti/.

2023 JINST 18 P06005