GENERATIVE MODELS FOR PARTICLE SHOWER SIMULATION IN FUNDAMENTAL PHYSICS

Anonymous authors
Paper under double-blind review

ABSTRACT

Simulations in fundamental physics connect underlying first-principle theories and empirical descriptions of detectors with experimentally observed data. They are indispensable, e.g., in particle physics, where they underpin statistical inference tasks and experiment design. However, these Monte-Carlo-based simulations are computationally costly. Growing data volumes produced by upcoming experiments exacerbate this problem and need to be matched by a commensurate growth in simulated statistics for precision measurements. Generative machine learning models offer a potential way to leverage the effectiveness of classical simulation via amplification. This contribution presents state-of-the-art performance for simulating the interaction of two different types of elementary particles with different characteristics — photons and pions — with high granularity detectors of 27k and 8k read-out channels, respectively.

1 INTRODUCTION

Fundamental physics probes the physical laws of Nature at length scales of 10^{-18} meters by investigating the interactions of high-energy particles at collider experiments. There, beams of particles (e.g., protons or electrons) close to the speed of light are brought into collision. These interactions produce sprays of secondary particles that are eventually measured in complex and highly-granular detectors. Investigating the statistical distributions of these decay products, for example, allowed the discovery of the Higgs boson particle in 2012 (CMS Collaboration, 2012; ATLAS Collaboration, 2012).

Detailed simulations of the occurring processes are essential to carry out this research. Monte Carlo methods are traditionally used to describe the interaction probabilities due to the quantum-mechanical, statistical theories and empirical models of the particle interactions with detectors. These simulations enable precise measurements of fundamental constants, the construction of hypothesis tests for discovering new particles and interaction laws, and the design and further improvement of detectors.

To be statistically viable, the number of simulated collision events should approximate the number of recorded events, leading to the required simulation of billions of examples; further growth is needed due to improvements in the collider and future experiments. Given a simulation time of seconds-to-minutes per event, this constitutes a significant expenditure of resources and a substantial bottleneck for this avenue of fundamental research.

A promising approach lies in using generative machine learning models to amplify the available statistics(Butter et al., 2020; Paganini et al., 2018a). Methods using Generative Adversarial Network (GANs) (Goodfellow et al., 2014), Variational Autoencoders (VAE) (Kingma & Welling, 2014), and autoregressive flows Huang et al. (2018) have been investigated for different aspects of this challenge such as event generation (Di Sipio et al., 2019; Otten et al., 2019; Butter et al., 2019; Alanazi et al., 2020), parton showering (Bothmann & Debbio, 2019; de Oliveira et al., 2017; Monk, 2018) and detector simulation (Paganini et al., 2018b; Erdmann et al., 2018; 2019; ATLAS Collaboration, 2018; 2019; Ghosh, 2019; Belayneh et al., 2019; Buhmann et al., 2021).

This contribution shows progress on a particular simulation challenge: particle showers in highly-granular detectors. We investigate two types of particles with different decay characteristics and present state-of-the-art methods for their generation, including a detailed statistical analysis of the obtained differential distributions of relevant features. The remainder of this work is structured as follows: Sec. 2 introduces the physics challenges and training datasets; Sec. 3 discusses the used generative architectures; Sec. 4 presents the obtained results; and Sec. 5 summarises our findings and shows an outlook on future work.

2 DATA SETS

We simulate the resulting chains of particle emissions, so-called *showers*, obtained by shooting two different species of primary particles — photons and pions — into a detector block. A realistic simulation will need to be conditioned on the particle type, but we treat photon and pion showers as two separate problems for development purposes. We selected these two species as they offer distinct difficulties in their simulation: photon showers are more compact and comparatively easier to simulate, while due to a mixture of electromagnetic and nuclear interactions present in their showers, pions exhibit a significantly larger topological variety. These interactions are illustrated in Fig.1.

Simulating the planned International Linear Detector (ILD) (Abramowicz et al., 2020) prototype was chosen as a target as this detector design, due to high granularity and high expected particle rates, offers a particularly compelling simulation challenge. Different sub-detectors are especially sensitive to different particle types. For photons this is the *electromagnetic calorimeter* (ECal) while for pions the *Analogue Hadron Calorimeter* (AHCal) is particularly relevant. We, therefore, simulate photon interactions with the ECal and pion interactions with the AHCal subdetectors, respectively. Modern calorimeter detectors are constructed in a sandwich structure consisting of dense passive material layers that initiate showers and active sensor materials measuring the interaction. Details on both data sets are provided in Table 1. For both particle types, the primary particle energies are sampled uniformly between 10 and 100 GeV, to provide a realistic variety. The initial simulation is carried out using the GEANT4 (Agostinelli et al., 2003) software package. Artifacts from projecting the physical geometry to a fixed grid are corrected such that every pixel corresponds to precisely one sensor.

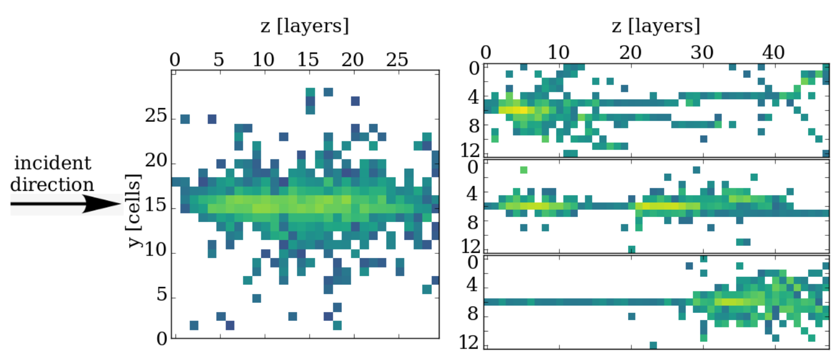

Figure 1: Projections of one photon shower (left) and three pion showers (right), indicating the large diversity among pion shower shapes. Note also the different shape of the $30 \times 30 \times 30$ photon and $48 \times 13 \times 13$ pion data.

3 GENERATIVE MODELS

We train and evaluate three generative models in this study: GANs, GANs optimizing a Wasserstein loss, and Bounded Information Bottleneck Autoencoders (BIB-AEs). All are implemented in PYTORCH (Paszke et al., 2019) using convolutional and transpose-convolutional operations. Additional details of the network architectures are provided in Ref. Buhmann et al. (2021).

GANs (Goodfellow et al., 2014) were first proposed for simulating showers in particle physics in Ref. Paganini et al. (2018a). To obtain realistic showers, we parametrize the generator and discrim-

Table 1: Parameter summary of the two data sets used.

	Photon / ECal	Pion / AHCal
Material (Active/Passive)	Silicon / Tungsten	Polystyrene / Steel
Sensor size	5×5 mm^2	$30 \times 30 \times 3$ mm^3
Data dimensions (pixels)	$30 \times 30 \times 30$	$13 \times 13 \times 48$
Training sample size	950k showers	500k showers

inator on the primary particle energy E (Baldi et al., 2016). We use minibatch discrimination (Salimans et al., 2016) following Karras et al. (2017) to reduce mode collapse, increase the variety of generated showers, and obtain better stability in the training process for the GAN model.

Using the Wasserstein distance (Cédric, 2009) — e.g. approximated following the Kantorovich-Rubinstein duality via a critic (Gulrajani et al., 2017) network — as loss can improve the stability of generative model training. Following Erdmann et al. (2019), we condition the generator and critic network on the primary particle energy and use an additional constrained network to improve fidelity. For simulating pion showers, the WGAN model training is improved via Latent Optimization (Wu et al., 2019; 2020).

The BIB-AE (Voloshynovskiy et al., 2019) unifies features of GANs and autoencoders. A pair of Wasserstein-GAN-like critics evaluate the quality of reconstructed images: one judges the realism of the reconstructed outputs, and one compares the output and input image, ensuring that they look

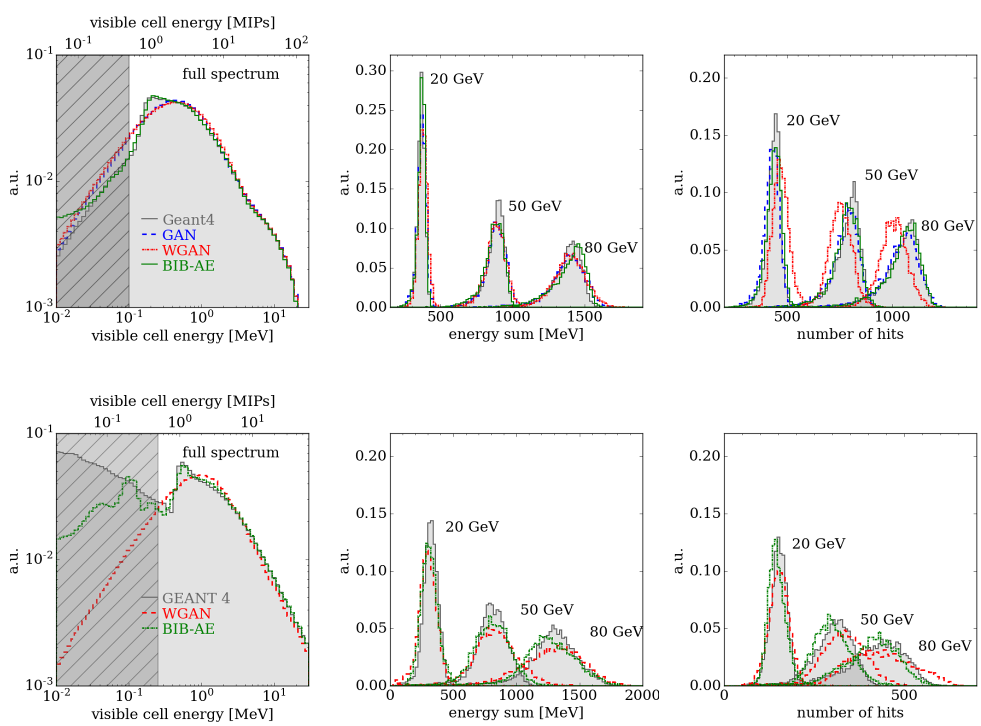

Figure 2: Differential distributions comparing physics quantities between GEANT4 (ground truth) and the different generative models. The top (bottom) row shows photon (pion) showers. The energy per-cell is measured in MeV for the bottom axis and in multiples of the expected energy deposit of a minimum ionizing particle (MIP) for the top axis.

alike. Latent regularisation is improved by an additional critic and an MMD term (Gretton et al., 2008). An additional Post-Processor network, trained in a second step, is used to improve per-pixel energies. BIB-AE training is further improved by minibatch-discrimination during training.

4 RESULTS

For simulations in particle physics, a visual comparison of individual images is less revealing than a statistical analysis of the obtained distributions. In Fig. 2, we show examples of three relevant observables: the energy contained in a single sensor (visible cell energy, left), the total energy sum over all pixels in a shower (center), and the number of non-zero pixels (right). The top row shows photon showers, while pions are presented in the bottom row. In all cases, we compare the ground truth (GEANT4) to the learned models.

Due to electronic noise, only single-cell energies above half the energy deposited by a minimally ionizing particle (MIP) are physically relevant. This energy is different for ECal and AHCal detectors and provided as top scale. The greyed-out region in the left panels represents this threshold. All models describe high energies above 5 MIP well. The BIB-AE with Post Processor network also correctly learns the structure around 1 MIP. The total energy is similarly well described by all models and in general is better modeled for pion showers. The WGAN model has difficulties modeling the number of non-zero pixels, while GAN and BIB-AE better describe this distribution for photons and the BIB-AE for pions.

The main goal of training generative model approximations is to speed-up the sampling process. Table 2 shows the time to simulate single showers using GEANT4 and the different generative models. For assessing practical benefits, the time needed to produce the training data, the time to train generative models, and the possible amount of statistical amplification will need to be taken into account, but the observed speed-ups of three-orders of magnitude are encouraging for future work.

Table 2: Computational performance of WGAN and BIB-AE generators on a single core of an Intel® Xeon® CPU E5-2640 v4 (CPU) and NVIDIA® V100 with 32 GB of memory (GPU) compared to GEANT4. For the generative models, the best performing batch size is shown and given by the mean and standard deviation obtained for sets of 5000 showers.

Hardware	Simulator	Photons time/shower [ms]	Photons speed-up	Pions time/shower [ms]	Pions speed-up
CPU	GEANT4	4082 ± 170	$\times 1$	2684 ± 125	$\times 1$
	WGAN	61.44 ± 0.03	$\times 66$	195.67 ± 0.56	$\times 14$
	BIB-AE	95.98 ± 0.08	$\times 43$	36.05 ± 0.82	$\times 74$
GPU	WGAN	3.93 ± 0.03	$\times 1039$	2.695 ± 0.004	$\times 996$
	BIB-AE	1.60 ± 0.03	$\times 2551$	1.101 ± 0.004	$\times 2438$

5 CONCLUSIONS AND OUTLOOK

Fundamental physics is facing growing difficulties from the computing resources expanded on slow Monte-Carlo-based simulations. While these simulations encode valuable physical knowledge and are challenging to replace, better use of generated statistics is possible by using generative models. In this contribution, we show the performance of GAN, Wasserstein-GAN, and the BIB-AE architecture on this task. We observe accurate modeling of physically relevant quantities over several orders of magnitude and a speed-up over three orders of magnitude. For practical applications, the fidelity will need to be further improved — down to percent level in a larger number of distributions — and the generator needs to be more flexible by conditioning on particle type, incident position, and incident angle in addition to the energy.

References

Halina Abramowicz et al. International Large Detector: Interim Design Report. 3 2020. 2003.01116.

S. Agostinelli et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3): 250 – 303, 2003.

Yasir Alanazi, N. Sato, Tianbo Liu, W. Melnitchouk, Michelle P. Kuchera, Evan Pritchard, Michael Robertson, Ryan Strauss, Luisa Velasco, and Yaohang Li. Simulation of electron-proton scattering events by a Feature-Augmented and Transformed Generative Adversarial Network (FAT-GAN). 1 2020. 2001.11103.

ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, Sep 2012. 1207.7214.

ATLAS Collaboration. Deep generative models for fast shower simulation in ATLAS. Technical Report ATL-SOFT-PUB-2018-001, CERN, Geneva, Jul 2018.

ATLAS Collaboration. VAE for photon shower simulation in ATLAS. Technical Report ATL-SOFT-SIM-2019-007, CERN, October 2019.

Pierre Baldi, Kyle Cranmer, Taylor Faucett, Peter Sadowski, and Daniel Whiteson. Parameterized Machine Learning for High-Energy Physics. *Eur. Phys. J. C*, 76(5):235, 2016. 1601.07913.

Dawit Belayneh et al. Calorimetry with Deep Learning: Particle Simulation and Reconstruction for Collider Physics. 2019. 1912.06794.

Enrico Bothmann and Luigi Debbio. Reweighting a parton shower using a neural network: the final-state case. *JHEP*, 01:033, 2019. 1808.07802.

Erik Buhmann, Sascha Diefenbacher, Engin Eren, Frank Gaede, Gregor Kasieczka, Anatolii Korol, and Katja Krüger. Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed. 2021. 2005.05334.

Anja Butter, Tilman Plehn, and Ramon Winterhalder. How to GAN LHC Events. *SciPost Phys.*, 7: 075, 2019. 1907.03764.

Anja Butter, Sascha Diefenbacher, Gregor Kasieczka, Benjamin Nachman, and Tilman Plehn. GAN-plifying Event Samples. 8 2020. 2008.06545.

Villani Cédric. *Optimal Transport: Old and New*. Springer, Berlin, 2009.

CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, Sep 2012. 1207.7235.

Luke de Oliveira, Michela Paganini, and Benjamin Nachman. Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis. *Comput. Softw. Big Sci.*, 1(1):4, 2017. 1701.05927.

Riccardo Di Sipio, Michele Faucci Giannelli, Sana Ketabchi Haghighat, and Serena Palazzo. DijetGAN: A Generative-Adversarial Network Approach for the Simulation of QCD Dijet Events at the LHC. 2019. 1903.02433.

Martin Erdmann, Lukas Geiger, Jonas Glombitza, and David Schmidt. Generating and refining particle detector simulations using the Wasserstein distance in adversarial networks. *Comput. Softw. Big Sci.*, 2(1):4, 2018. 1802.03325.

Martin Erdmann, Jonas Glombitza, and Thorben Quast. Precise simulation of electromagnetic calorimeter showers using a Wasserstein Generative Adversarial Network. *Comput. Softw. Big Sci.*, 3(1):4, 2019. 1807.01954.

Aishik Ghosh. Deep generative models for fast shower simulation in ATLAS. Technical Report ATL-SOFT-PROC-2019-007, CERN, Geneva, Jun 2019.

Ian J. Goodfellow et al. Generative Adversarial Nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, pp. 2672–2680, Cambridge, MA, USA, 2014.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander J. Smola. A Kernel Method for the Two-Sample Problem. *CoRR*, 2008. `0805.2368`.

Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. Improved Training of Wasserstein GANs. In *Advances in Neural Information Processing Systems 30*, pp. 5767–5777, 2017.

Chin-Wei Huang, David Krueger, Alexandre Lacoste, and Aaron C. Courville. Neural Autoregressive Flows. *CoRR*, 2018. `1804.00779`.

Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive Growing of GANs for Improved Quality, Stability, and Variation. 2017. `1710.10196`.

Diederik P Kingma and Max Welling. Auto-Encoding Variational Bayes. 2014. `1312.6114`.

J. W. Monk. Deep Learning as a Parton Shower. *JHEP*, 12:021, 2018. `1807.03685`.

Sydney Otten, Sascha Caron, Wieske de Swart, Melissa van Beekveld, Luc Hendriks, Caspar van Leeuwen, Damian Podareanu, Roberto Ruiz de Austri, and Rob Verheyen. Event Generation and Statistical Sampling for Physics with Deep Generative Models and a Density Information Buffer. 2019. `1901.00875`.

Michela Paganini, Luke de Oliveira, and Benjamin Nachman. Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters. *Phys. Rev. Lett.*, 120(4):042003, 2018a. `1705.02355`.

Michela Paganini, Luke de Oliveira, and Benjamin Nachman. CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks. *Phys. Rev.*, D97(1):014021, 2018b. `1712.10321`.

Adam Paszke et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems 32*, pp. 8024–8035, 2019. Advances in Neural Information Processing Systems 32 pp. 8024–8035.

Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved Techniques for Training GANs. 2016. `1909.10578`.

Slava Voloshynovskiy, Mouad Kondah, Shideh Rezaeifar, Olga Taran, Taras Holotyak, and Danilo Jimenez Rezende. Information bottleneck through variational glasses. 2019. `1912.00830`.

Yan Wu, Mihaela Rosca, and Timothy Lillicrap. Deep Compressed Sensing. 2019. `1905.06723`.

Yan Wu, Jeff Donahue, David Balduzzi, Karen Simonyan, and Timothy Lillicrap. LOGAN: Latent Optimisation for Generative Adversarial Networks. 2020. `1912.00953`.