Lawrence Berkeley National Laboratory
Recent Work

Title
Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters.

Permalink

Journal

ISSN

Authors
Paganini, Michela
de Oliveira, Luke
Nachman, Benjamin

Publication Date

DOI

Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters

Michela Paganini,[1,2,*] Luke de Oliveira,[2,†] and Benjamin Nachman[2,‡]

[1]*Yale University, New Haven, Connecticut 06520, USA*
[2]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

Physicists at the Large Hadron Collider (LHC) rely on detailed simulations of particle collisions to build expectations of what experimental data may look like under different theoretical modeling assumptions. Petabytes of simulated data are needed to develop analysis techniques, though they are expensive to generate using existing algorithms and computing resources. The modeling of detectors and the precise description of particle cascades as they interact with the material in the calorimeter are the most computationally demanding steps in the simulation pipeline. We therefore introduce a deep neural network-based generative model to enable high-fidelity, fast, electromagnetic calorimeter simulation. There are still challenges for achieving precision across the entire phase space, but our current solution can reproduce a variety of particle shower properties while achieving speedup factors of up to $100\,000\times$. This opens the door to a new era of fast simulation that could save significant computing time and disk space, while extending the reach of physics searches and precision measurements at the LHC and beyond.

DOI: 10.1103/PhysRevLett.120.042003

Introduction.—High-precision modeling of the interactions of particles with media is important across many physical sciences, enabling and accelerating new findings. Similar to complex weather or cosmological modeling, the detailed simulation of subatomic particle collisions and interactions, as captured by detectors at the LHC, is a computationally demanding task which annually requires billions of CPU hours, constituting more than half of the LHC experiments' computing resources [1–3].

The Nobel-prize-winning Higgs boson discovery [4,5] would not have been possible without extensive simulation. Before its experimental observation, its fundamental properties, such as its mass, were unknown, but synthetic particle collisions could be generated to simulate the outcome of various measurements under different model assumptions.

Today, as several questions remain unanswered about the nature of known particles (such as neutrinos) and hypothetical ones (such as the supersymmetric partners of the standard model particles), modern nuclear and particle physics research continues to strongly depend on detailed simulations for developing analysis techniques, interpreting results, and designing new experiments.

Cutting-edge software libraries such as GEANT4 [6] provide the backbone to construct complex detector geometries and accurately model physical processes and interactions happening at distance scales as small as 10^{-20} m.

The shortcoming of this method is its computational footprint. The high-precision description of electromagnetic and nuclear processes that govern the evolution of particle showers in calorimeters can require minutes per event on modern computing platforms [7,8], making this the most computationally expensive step in the simulation pipeline. Because of the high simulation cost, significant resources are also invested in storing generated data sets, which can occupy petabytes of disk space.

This bottleneck becomes apparent at the scale at which events need to be simulated to enable physics analyses at the high luminosity phase of the LHC. The ATLAS and CMS experiments are expected to observe about 10^8 Higgs boson events [9], buried in $\sim 10^{17}$ background events [10,11]. Hundreds of billions of simulated collisions will be required to reduce the Monte Carlo uncertainty and measure some of the Higgs boson's as yet unprobed properties.

Approximate calorimeter simulation techniques exist [12–15], but they provide compromises that lie on different, yet similarly suboptimal, parts of the accuracy-speedup trade-off curve.

Full detector simulations are too slow to meet the growing analysis demands; current fast simulations are not precise enough to serve the entire physics program. We therefore introduce a deep learning model, named CaloGAN, for a high-fidelity fast simulation of particle showers in electromagnetic calorimeters. Its goal is to be both quick and precise, by

significantly reducing the accuracy cost incurred with increased speedup. A fast simulation technique of this kind also addresses the issue of data storage and transfer, as the gained generation simplicity and speedup make real-time, on-demand simulation a possibility.

Similar techniques have been tested in cosmology [16,17], condensed matter physics [18], and oncology [19]. However, the sparsity, high dynamic range, and highly location-dependent features present in this application make it uniquely challenging. In addition to enabling physics analysis at the LHC, an approach similar to CaloGAN's may be useful for other applications in particle and nuclear physics, nuclear medicine, and space science that require a detailed modeling of particle interactions with matter.

Method.—To alleviate the computational burden of simulating electromagnetic showers, we introduce a method based on generative adversarial networks (GANs) [20] in order to directly simulate component readouts in electromagnetic calorimeters. GANs are an increasingly popular approach to learning a generative model using deep neural networks, and they have shown great promise in generating clear samples from natural images [21].

Though the GAN formulation, by design, does not admit an explicit probability density or explicit likelihood, we gain the ability to sample from the learned generative model in a efficient manner. The GAN training uses a minimax game-theoretic framework and admits a function g as an artifact that maps a d-dimensional latent vector, $z \sim p_z(z) \in \mathbb{R}^d$ to a point in the space of realistic samples. We would like the implicit density learned by g to be close to the distribution f that governs the simulated data distribution. Since g is a neural network, a forward pass to generate new samples is highly efficient on modern computing platforms [22].

Previous work [23] investigated GAN-based methods for jet images [24], which are similar to one-layer calorimeters with square pixels (except that jet generators such as PYTHIA [25] are much faster than GEANT4). This Letter addresses the complexity introduced by modeling a realistic sampling detector with heterogeneous longitudinal and transverse segmentation. We exploit the location specificity of the calorimeter and utilize weight locality at the model level. We also follow the guidelines outlined in Ref. [23] in order to deal with both high dynamic range and high sparsity levels. Our neural-network architecture per calorimeter layer is a function of the readout grid dimensionality, and it is augmented with an attentional component [26] that provides a mechanism to carry information from layer to layer [27]. This allows CaloGAN to model the physical sequential dependence among the calorimeter layers.

To ensure the realism of the CaloGAN setup, we impose an additional constraint to encourage the generator to produce a given energy shower. That is, the learned, implicit probability density function f needs to converge to the hypothetical data generating function g for any initial nominal energy E_0, i.e., $f(x|E = E_0) \rightarrow g(x|E = E_0)$ for all cases where $E_0 \in [E_{min}, E_{max}]$.

To encourage this to be well modeled, a physics-specific loss component is introduced to penalize absolute deviation between the nominal energy E_0 and the reconstructed energy $\hat{E}$. A noteworthy subtlety is that this penalization scheme, coupled with minibatch discrimination [28], invites the network to learn the distribution of $|E_0 - \hat{E}|$, a desirable characteristic for a readily applicable practical system to augment fast simulation. Such a formulation also encourages conservation of energy through the generation process. The simulation only includes models of energy deposition, not digitization (a nonlinear effect that can violate reconstructed energy conservation). The energy per layer includes the contribution from inactive material (see below). Therefore, aside from leakage beyond the calorimeter (relevant mostly for charged pions), energy must be conserved and provides a useful constraint on the generation.

Experimental results.—From a series of simulated showers, CaloGAN is tasked with learning the simulated data distributions of γ, e^+, and π^+ generated by GEANT4 with uniform energy spectrum [1,100] GeV, and incident perpendicular to the center of a three-layer, heterogeneously segmented, liquid argon (LAr) calorimeter cube of side length 480 mm. The training data set [29] is represented in image format by three figures of dimensions 3×96, 12×12, and 12×6, each representing the shower energy depositions per pixel in each calorimeter layer. The energy per layer includes the active and inactive contributions. For calorimeter calibrations [30], e.g., it is important to have the inactive component; in the future, one could add separate layers for the inactive component or add a second step for dividing the energy per layer into two components. The flexible CaloGAN architecture allows for a straightforward extension to related detector geometries that have more sampling layers or different cell sizes per layer [31].

Our analysis establishes that it is possible to generate three-dimensional electromagnetic showers in a multilayer sampling LAr calorimeter with uneven spatial segmentation while attempting to preserve spatiotemporal relationships among layers.

For performance evaluation, we choose application-driven methods focused on sample quality. A first qualitative assessment is accompanied by a quantitative evaluation based on physics-driven similarity metrics. The choice reflects the domain specific procedure for Monte Carlo–data comparisons. However, it is also important to examine high-dimensional behavior because CaloGAN is not anchored by parametrized models the way traditional fast simulators are. While the adversarial classifier provides some high-dimensional validation, we also use particle classification performance. Visualization and validation are still key challenges for multidimensional generators parametrized by a neural network.

Qualitative evaluation.—The average calorimeter deposition per voxel (Fig. 1) suggests that the learned generative models of γ, e^+, and π^+ showers capture aspects of the underlying physical processes. For photon showers, for

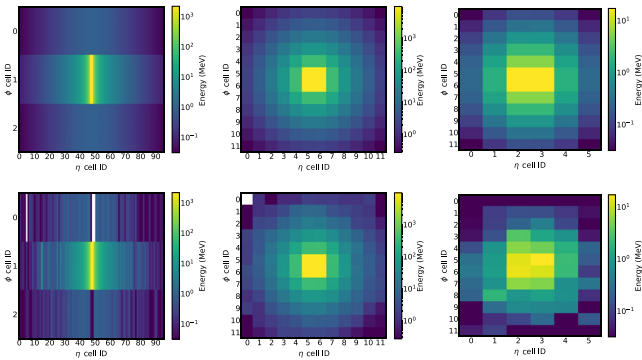FIG. 1. Average γ GEANT4 shower (top row), and average γ CaloGAN shower (bottom row), with progressive calorimeter depth (from left to right).

instance, the mean per-layer cell variations only show ∼4% and ∼1% discrepancy in the first two layers, respectively where most energy is deposited for e/γ. This level of agreement is promising, but it is important to analyze more than the mean energy pattern to fully study the strengths and weaknesses of the proposed approach.

The CaloGAN-generated samples are checked for adequate diversity and lack of direct memorization of the GEANT4 samples used for training. The nearest (by Euclidean distance) GEANT4 image is found for each of a random selection of CaloGAN images in order to verify the desired characteristics (Fig. 2). The samples show strong inter- and intraclass diversity and no evidence of memorization since the closest images do not look exactly the same.
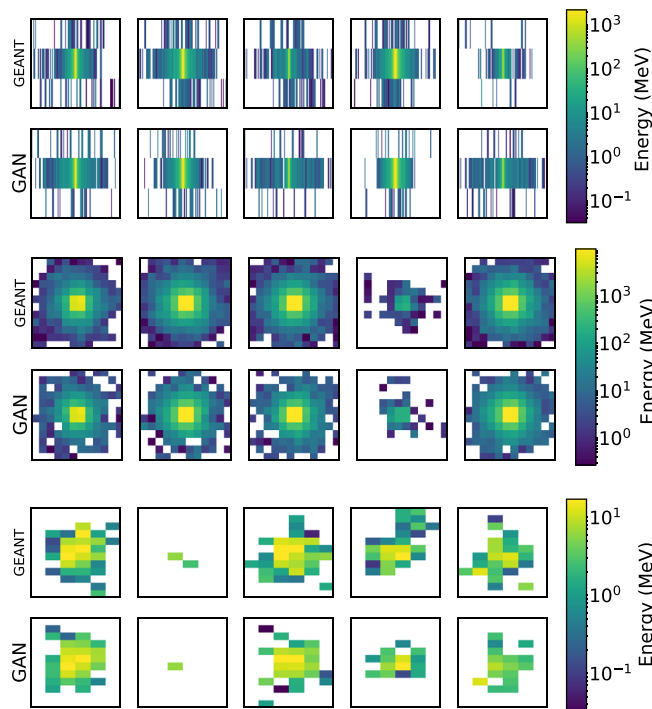

FIG. 2. Five randomly selected γ showers per calorimeter layer from GEANT4 (top rows) and their five nearest neighbors (by Euclidean distance) from a set of CaloGAN candidates.

Shower shape description.—Geometrically and physically motivated shower shape variables [32] are used as further validation and introspection into the capabilities of CaloGAN to adequately model and capture nonlinear functional representations of the simulated data distribution (Fig. 3). In fact, it is desirable for CaloGAN to recover the target distribution of these 1D statistics.

The network is not shown any shower shape variables (only pixel values) at training time—therefore, it is encouraging to note that CaloGAN recovers the simulated data distribution for a variety of shower shapes across the three particle types. However, certain features of some distributions are not well described. This is a challenge for the future and will likely require improvements to the architecture and training procedure. Longer trainings of higher capacity architectures have shown promise in rectifying some of these issues.

Examining 1D statistics does not probe correlations between shower shapes or higher dimensional aspects of the probability distribution. One way to examine the full shower phase space is to study classification performance, as described in the next section.

Classification as a performance proxy.—When training a six-layer, fully connected classification model on the 504-dimensional pixel space of the concatenated representation of shower energy depositions across all calorimeter layers, no major classification degradation is observed for out-of-domain learning when trained on the full simulation, i.e., when the network is trained on GEANT4 samples but evaluated on CaloGAN samples. Specifically, although the classification accuracy always reaches 99% when evaluating performance on CaloGAN showers—which points to an overdifferentiation among particle types in the CaloGAN data set—in both $e^+ - \gamma$ and $e^+ - \pi^+$ discrimination tasks, the evaluation of the network trained on GEANT4 images results in no accuracy decrease in the former task (∼70%), and only a 2% decrease in the latter (∼97% versus ≳99% accuracy), when compared to the classifier tested on CaloGAN samples. The stability of the accuracy metric implies that CaloGAN succeeds at representing at least as much variation among showers initiated by different particles as is necessary to classify them using the same features in GEANT4. Training on CaloGAN and testing on GEANT4 does show significant degradation, indicating that the GAN is inventing new class-dependent features or underrepresenting class-independent features. While percent-level variations may be important for some applications, using classification as a generator diagnostic is an important tool for exposing the modeling of interclass shower variations.

Computational performance.—Directly generating deposited energy per calorimeter cell rather than particle dynamics renders the model's time-complexity invariant to nominal energy, whereas GEANT4 shower simulation runtime increases significantly with higher energy. Therefore, CaloGAN affords sizable simulation-time speedups compared to GEANT4. All benchmarks are performed on Intel Xeon® 2.6 GHz processors for CPU time and a single
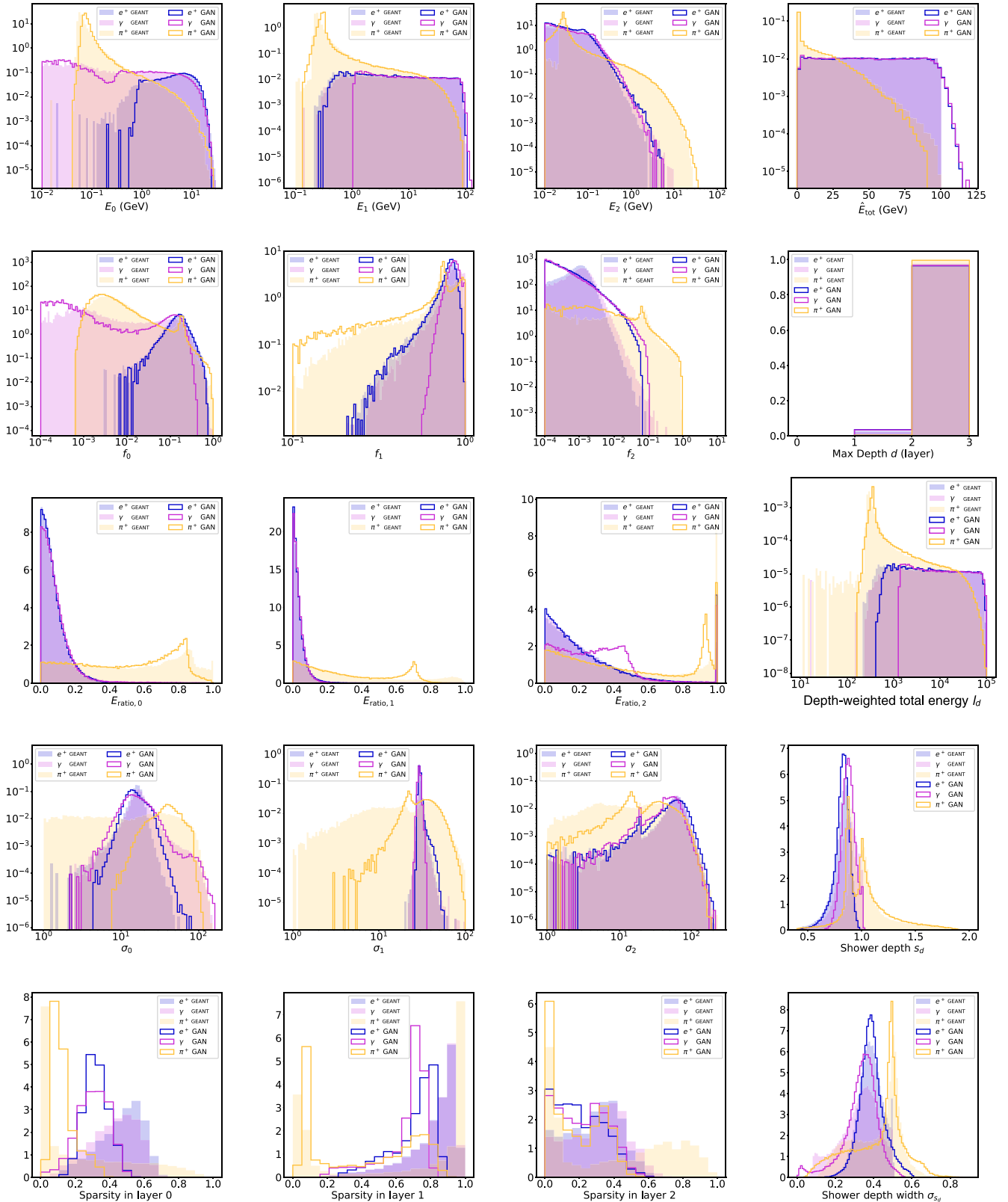
FIG. 3. Comparison of shower shape variables and other variables of interest, such as the sparsity level per layer, for the GEANT4 and CaloGAN data sets for e^+, γ, and π^+. See Ref. [33] for detailed definition.

NVIDIA® K80 for GPU time. When simulating a single e^+ in a uniform energy range between 1 and 100 GeV, CaloGAN is $O(10^2)$ times faster than GEANT4 on both CPU and GPU.

However, when batching is utilized, the CaloGAN throughput significantly improves—when batching of size 1024 is allowed (not unrealistic given the embarrassingly parallel

nature of electromagnetic showering), the per-e^+ generation time is $O(10^3)$ times faster on CPU and $O(10^5)$ times faster on GPU.

Outlook and future work.—This Letter demonstrates that the generative adversarial network technology represents a powerful new tool for efficient simulation. Our ability to infuse physics domain knowledge into the neural network documents the flexibility and extensibility of the method for field-specific applications and explicit mismodeling mitigation.

Prior to this work, the prospect of a GAN-based calorimeter simulation had generated considerable excitement within the high energy physics community. The availability and performance of CaloGAN has attracted further interest as a concrete and publicly available demonstration of the power and drawbacks of a GAN-based calorimeter simulation. In addition to the applicability within individual experiments, variations of CaloGAN are also being studied as a generic tool for future GEANT software versions. While CaloGAN is currently structured as a fast simulation tool, in the future it could also be trained on test-beam data to replace or augment a full simulation tool.

Future work will focus on incorporating the most recent cutting-edge innovations from the GAN literature to stabilize the training procedure and improve convergence to optimal solutions [34–37]. While our primary effort will be to improve and maintain this technique for event simulation at the LHC, this neural-network approach retains generalization power to other fields in which computationally expensive simulation inhibits result productivity.

---

[*]michela.paganini@yale.edu
[†]lukedeoliveira@lbl.gov
[‡]bnachman@cern.ch
[1] J. Flynn, CERN Report No. CERN-RRB-2015-014, 2015.
[2] E. Karavakis, J. Andreeva, S. Campana, S. Gayazov, S. Jezequel, P. Saiz, L. Sargsyan, J. Schovancova, and I. Ueda, J. Phys. Conf. Ser. **513**, 062024 (2014).
[3] C. Bozzi, CERN Reports No. LHCb-PUB-2015-004 and No. CERN-LHCb-PUB-2015-004, 2015.
[4] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[5] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).
[6] GEANT4 Collaboration, Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).
[7] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **70**, 823 (2010).
[8] R. Rahmat, R. Kroeger, and A. Giammanco, J. Phys. Conf. Ser. **396**, 062016 (2012).
[9] D. de Florian *et al.* (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector* (CERN Yellow Reports, Geneva, 2017).
[10] M. Aaboud *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **117**, 182002 (2016).
[11] CERN Report No. CMS-PAS-FSQ-15-005, 2016.
[12] G. Grindhammer and S. Peters, in *Proceedings of the International Conference on Monte Carlo Simulation in High-Energy and Nuclear Physics (MC 93), Tallahassee, Florida, 1993*, edited by P. Dragovitsch, S. Linn, and M. Burbank (World Scientific, Singapore, 1994), p. 140.
[13] M. Beckingham *et al.* (ATLAS Collaboration), CERN Report No. ATL-PHYS-PUB-2010-013, 2010.
[14] G. Grindhammer, M. Rudowicz, and S. Peters, in *Proceedings of the SSC Workshop on Calorimetry for the Superconducting Super Collider, Tuscaloosa, 1989* (1990) [Nucl. Instrum. Methods Phys. Res., Sect. A **290**, 469 (1990)].
[15] E. Barberio, J. Boudreau, B. Butler, S. L. Cheung, A. Dell'Acqua, A. D. Simone, E. Ehrenfeld, M. V. Gallas, A. Glazov, Z. Marshall, J. Mueller, R. Plaakyte, A. Rimoldi, P. Savard, V. Tsulaia, A. Waugh, and C. C. Young, J. Phys. Conf. Ser. **160**, 012082 (2009).
[16] S. Ravanbakhsh, F. Lanusse, R. Mandelbaum, J. Schneider, and B. Poczos, arXiv:1609.05796.
[17] K. Schawinski, C. Zhang, H. Zhang, L. Fowler, and G. K. Santhanam, Mon. Not. R. Astron. Soc. **467**, L110 (2017).
[18] L. Mosser, O. Dubrule, and M. J. Blunt, Phys. Rev. E **96**, 043309 (2017).
[19] A. Kadurin, A. Aliper, A. Kazennov, P. Mamoshina, Q. Vanhaelen, K. Khrabrov, and A. Zhavoronkov, Oncotarget **8**, 10883 (2016).
[20] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial networks, in *Advances in Neural Information Processing Systems 27* (Curran Associates, Inc., New York, 2014), pp. 2672–2680.
[21] A. Radford, L. Metz, and S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, in *Proceedings of the International Conference on Learning Representations (ICLR), 2015* [arXiv:1511.06434].
[22] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, and E. Shelhamer, arXiv:1410.0759.
[23] L. de Oliveira, M. Paganini, and B. Nachman, Comput. Software Big Sci. **1**, 4 (2017).
[24] J. Cogan, M. Kagan, E. Strauss, and A. Schwarztman, J. High Energy Phys. 02 (2015) 118.
[25] T. Sjostrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.
[26] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, in *Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 2015*, Proceedings of Machine Learning Research Vol. 37, edited by F. Bach and D. Blei (PMLR, Lille, France, 2015), p. 2048.
[27] H. Zhang, T. Xu, H. Li, S. Zhang, X. Huang, X. Wang, and D. Metaxas, StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks, in *The IEEE International Conference on Computer Vision*

(ICCV) (IEEE Computer Society, Washington, 2017), pp. 5908–5916.

[28] T. Salimans, I. J. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, Improved techniques for training GANs, in *Advances in Neural Information Processing Systems 29*, edited by D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc., New York, 2016), pp. 2234–2242.

[29] B. Nachman, L. de Oliveira, and M. Paganini, Electromagnetic calorimeter shower images, DOI: 10.17632/pvn3xc3wy5.1 (2017).

[30] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **77,** 490 (2017).

[31] B. Nachman, L. de Oliveira, and M. Paganini, hep-lbdl/CaloGAN: CaloGAN generation, training, and analysis code, DOI: 10.5281/zenodo.584155 (2017).

[32] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C **38,** 090001 (2014).

[33] M. Paganini, L. de Oliveira, and B. Nachman, arXiv:1712.10321 [Phys. Rev. D (to be published)].

[34] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, Improved training of Wasserstein GANs, in *Advances in Neural Information Processing Systems 30*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc., New York, 2017), pp. 5769–5779.

[35] S. Nowozin, B. Cseke, and R. Tomioka, f-GAN: Training generative neural samplers using variational divergence minimization, in *Advances in Neural Information Processing Systems 29*, edited by D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc., New York, 2016), pp. 271–279.

[36] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, G. Klambauer, and S. Hochreiter, GANs trained by a two time-scale update rule converge to a local Nash equilibrium, in *Advances in Neural Information Processing Systems 30*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc., New York, 2017), pp. 6629–6640.

[37] D. Berthelot, T. Schumm, and L. Metz, arXiv:1703.10717.