Fast Calorimeter Simulation with Wasserstein Generative Adversarial Networks

Anonymous Author(s)
Affiliation
Address
email

Abstract

This work employs WGANs to generate calorimetric data that will be on high demand in the event of the CMS Phase II Upgrade. The generated samples resemble the real ones from physics-based simulations; the particle showers are especially well-reproduced longitudinally along the calorimeter. The performance of the model is evaluated using the Wasserstein distance between the real and generated distribution of energy projected in each dimension of the calorimeter.

1 Motivation

In High Energy Physics (HEP) collider experiments, detectors register the outcome of particle collisions. In the reconstruction of collision events, the particle traces left in the detector are analyzed, for example, to identify the particle type and estimate its energy. The highly granular calorimeter (HGCAL) will be one of the greatest novelties of the Compact Muon Solenoid (CMS) Phase II upgrade. After the upgrade, there will be four times more particle interactions in the collider, introducing a new demand for large and accurate simulation samples; such simulations then become even more computationally expensive and will hardly be able to supply all the demand given the resource constraints of the experiment.

Deep generative models offer a possibility to speed up the data generation compared to simulations. This work uses Wasserstein Generative Adversarial Networks (WGANs) [1] as an alternative to supply the high demand for fast simulation for the HGCAL.

GANs have been applied to calorimeter simulation problems in HEP, such as in the CaloGAN [4] architecture for ATLAS' electromagnetic calorimeter; in the Linear Collider Detector's three-dimensional calorimeter [2]; and WGANs have been applied to calorimeter data from CERN's Super Proton Synchrotron test beam [3]. Thus being a promising direction to explore.

2 Data and model

This work considers a three-dimensional calorimeter of hexagonal-shaped cells, receiving particles from varied incoming angles; this angle variation originates realistic events and diverse samples. The dataset consists of Monte Carlo simulations of electron showers generated with the Geant4 software, containing 150,000 events with energies ranging from 0 to 500 GeV.

The model architecture adapts the Deep Convolutional GAN to fit the dimensions of the data, using two-dimensional convolutions. The calorimetric data can be seen as a 3D image, and the sensors resemble voxels, convolutional layers are suitable to model spatial correlations.

The data generation is conditioned on the energy of the particle. For such, the normalized energy value is sampled from a normal distribution that represents the energy distribution of the real data,

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

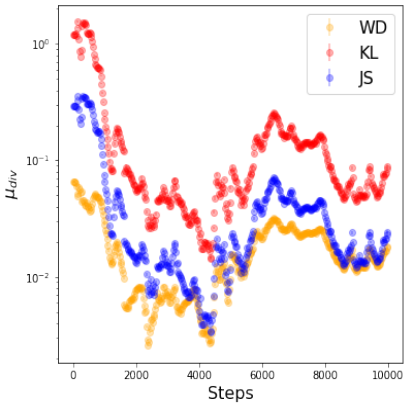

Figure 1: Evaluation metrics during training.

and this value is multiplied to the latent space vector of the generator. An additional task is added to the critical network: regression on the energy of the particle. The total loss function is weighted on this regression task, in addition to the WGAN loss.

3 Results

The generated samples obtained are similar to the real ones, but can overestimate the total energy of the particles. The energy deposition in each axis of the calorimeter reveals that the shower development along the calorimeter is well-reproduced accross the longitudinal layers. However, even though the energy deposition pattern on the cross section of the calorimeter is centralized, it still requires further improvement.

To evaluate the model performance, the Wasserstein distance (WD) between the distribution of real and generated samples is computed for the projection of the energy distribution in each axis of the calorimeter. The evaluation metric is the average of the Wasserstein distance for each axis. The Kullback-Leibler (KL) divergence and Jensen-Shannon (JS) divergence are also computed for the same purposes, showing that the performance of the model was consistent regardless of the metric, as in Figure 1.

This approach achieved a reasonable computational speedup: 1000 samples can be generated in less than a second using 10 CPU cores Skylake. The model trained for approximately 40 minutes in a V100 GPU. Nevertheless, the inference time of this model cannot yet be compared with the running time of physics-based simulations due to the different complexity of the events.

References

[1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein GAN. *arXiv:1701.07875 [cs, stat]*, January 2017. arXiv: 1701.07875.

[2] Federico Carminati, Gulrukh Khattak, Maurizio Pierini, Amir Farbin, Benjamin Hooberman, Wei Wei, Matt Zhang, Vitória Barin Pacela, Sofia Vallecorsa, Maria Spiropulu, and Jean-Roch Vlimant. Calorimetry with Deep Learning: Particle Classication, Energy Regression, and Simulation for High-Energy Physics. page 6, December 2017.

[3] Martin Erdmann, Jonas Glombitza, and Thorben Quast. Precise simulation of electromagnetic calorimeter showers using a Wasserstein Generative Adversarial Network. *arXiv:1807.01954 [physics]*, July 2018. arXiv: 1807.01954.

[4] Michela Paganini, Luke de Oliveira, and Benjamin Nachman. Accelerating Science with Generative Adversarial Networks: An Application to 3d Particle Showers in Multi-Layer Calorimeters. *Physical Review Letters*, 120(4), January 2018. arXiv: 1705.02355.