Sparse 3D Images: Point Cloud or Image methods?

Fernando Torales Acosta  
ftoralesacosta@lbl.gov  
Physics Division, Lawrence Berkeley National Laboratory  
Berkeley, CA 94720, USA

Vinicius Mikuni  
National Energy Research Scientific Computing Center  
Berkeley Lab, Berkeley, CA 94720, USA

Benjamin Nachman  
Physics Division, Lawrence Berkeley National Laboratory  
Berkeley Institute for Data Science, University of California  
Berkeley, CA 94720, USA

Miguel Arratia  
Department of Physics and Astronomy, University of California  
Riverside, CA 92521, USA  
Thomas Jefferson National Accelerator Facility, Newport News  
Virginia 23606, USA

Bishnu Karki  
Department of Physics and Astronomy, University of California  
Riverside, CA 92521, USA

Ryan Milton  
Department of Physics and Astronomy, University of California  
Riverside, CA 92521, USA

Piyush Karande  
Computational Engineering Division, Lawrence Livermore National Laboratory, Livermore CA 94550

Aaron Angerami  
Nuclear and Chemical Science Division, Lawrence Livermore National Laboratory, Livermore, CA 94550

Abstract

Score based generative models are a new class of generative models that have been shown to accurately generate high dimensional datasets. Recent advances in generative models have used images with 3D voxels to represent and model complex detector data. Point clouds, however, are likely a more natural representation for many of these data sets, particularly in calorimeters with high granularity that produce very sparse images. Point clouds preserve all of the information of the orig-
we study this problem in the context of a specific example: Calorimeters. Since most cells in a
θ
where the time-dependent network output with trainable parameters
α
and
σ
with time-dependent parameters
α
The data generation from the trained diffusion models is implemented using the DDIM sampler
t
velocity of the perturbed data at time
x
∼
R
2
Deep Learning Models

Diffusion models are a class of generative neural networks that allow for stable training paired with
high flexibility in the model design. Data is slowly perturbed over time using a time parameter
t
∈
R
that determines the perturbation level. The task of the neural network is to approximate the

\begin{align}
\mathcal{L}_\theta &= \mathbb{E}_{t, \epsilon} \| \mathbf{v}_t - \hat{\mathbf{v}}_{t, \theta} \|^2, \\
\mathbf{v}_t &\equiv \alpha_t \epsilon - \sigma_t \mathbf{x}, \text{ with } \epsilon \sim \mathcal{N}(0, 1). \\
\mathbf{x}_t &= \alpha_s \hat{x}_t(x_t) + \sigma_s \frac{x_t - \alpha_s \hat{x}_t(x_t)}{\sigma_t}.
\end{align}

For a fair comparison, all diffusion models are trained using the same score-matching strategy and
fixed number of 512 time steps during sampling.
The fast point cloud diffusion model (FPCD) follows [5], where a permutation equivariant estimation of the score function is obtained by the combination of a DeepSets [13] architecture with attention layers [14]. During the point cloud simulation, two models are also defined: one that learns the number of non-empty cells, conditioned on the initial energy of the incoming particle, and one model that learns the score function of the normalized point cloud, also conditioned on the momentum of the particle to be simulated and the number of hits to be generated.

The model trained on the image dataset (CALO SCORE) is adapted from [15] with a few modifications. Compared to the original implementation, the calorimeter simulation task is now broken down in two diffusion models: one that learns only the energy deposits in each layer of the calorimeter, conditioned only on the initial energy of the particle to be simulated, and one model that learns to generate normalized voxels per layer, conditioned on the energy deposition in each layer and the initial energy of the particle to be simulated. Additionally, the original U-NET [16] model is combined with attention layers.

3 Detector and Data Descriptions

The DD4HEP framework [17] is used to run GEANT simulations of a high-granularity iron-scintillator calorimeter based on previous designs for use in colliders [18, 8]). The sampling structure comprises of 55 $10 \times 10 \times 0.3$ cm scintillator tiles sandwiched between 2.0 cm thick steel absorber plates. The calorimeter is 1.2 meters long, with it’s front set at $z=3.8$ m. 1.7 million events of single $\pi^+$ particles are generated with momentum $1 < P < 125$ GeV/c in rings within the detector (see Figure 1).

Dataset 1 and Dataset 2 used in training share the same parent GEANT simulation, such that the fast point-cloud diffusion model and the image model are trained on different representations of the same set of calorimeter showers.

Dataset 1 is stored as zero-suppressed point-cloud representation. The GEANT data is stored in files containing two datasets, clusters and cells. The cluster dataset contains the $P_{\text{Gen}}$ of the incident pion, as well as the number of hits in the calorimeter. The cell data is comprised of a constant number of 200 cells per event. Each cell contains energy, $x$, $y$, and $z$ coordinate values. Empty cells, or cells with deposited energy below the threshold are 0-masked and ignored during training.

Dataset 2 is created by converting the point cloud dataset into an image format. Images at the full granularity of the detector would be unrealistic for an real-world detector. For example, the detector in [8], contains only 7 readout channels along the z-dimension, no where near 55. Additionally, images at full resolution would result in an unmanageably large datasets (see Table 1), and would represent the largest calorimeter image training ever done. The calorimeter cells were therefore clustered into groups of 5 along each axis of the detector to create voxels, where $5 \times 5 \times 5$ cells = 1 voxel. Energy in each of the cells making up the voxel were summed and assigned to the final voxel’s total energy. The final image format consists of $11 \times 11$ voxels. A hit in the voxelized dataset is defined as any voxel with energy deposition above threshold.

For the final comparison, generated samples from the point cloud model are voxelized using the same method for Dataset 2. All comparisons are in this image format, at the same resolution of $11 \times 11 \times 11$ voxels per image.

4 Results

A variety of distributions are used to evaluate the quality of the generated images. For all comparisons, the Earth mover’s distance (EMD) [19], also known as the 1-Wasserstein distance [20], between generated distributions and GEANT distributions is calculated. The EMD score a distance-like measure of the dissimilarity between two distributions. All EMD scores in Figures 2 are calculated on the final voxelized distributions.

Table 1 shows the model size, size of each dataset, and time to generate 100k calorimeter showers. The point cloud model is smaller by a factor of 4 compared to the image based model, and samples events 3 times faster. Lastly, the point cloud dataset requires over 100 times less disk space than the image format at full granularity. The AUC is obtained from a classifier trained to distinguish the samples of both models only in the voxelized image format. Both models have very good AUC,
<table>
<thead>
<tr>
<th>Model</th>
<th># Parameters</th>
<th>Disk Size (Full)</th>
<th>Sample Time</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image</td>
<td>2,572,161</td>
<td>1016MB (62GB)</td>
<td>8036.19s</td>
<td>0.673</td>
</tr>
<tr>
<td>Point Cloud</td>
<td>620,678</td>
<td>509 MB</td>
<td>2631.41s</td>
<td>0.726</td>
</tr>
</tbody>
</table>

Table 1: Comparison of model size, size of data representation on disk, generation time, and AUC of the same classifier trained to distinguish between the model and the original GEANT showers. All comparisons are done for 100k calorimeter showers. All results in the image row were obtained with $11 \times 11 \times 11$ voxel images. Disk size of the image dataset at full granularity is shown in parenthesis.

Figure 1: The 2-dimensional distribution of the mean deposited energy in 5th voxelized layer of the calorimeter. The first column is the original Geant simulation. The second column is the fast point-cloud based diffusion model (FPCD), and the 3rd column is the image-based model (CALOscore).

Figure 2(a) compares the total energy deposited in the calorimeter. Both models are in good agreement with GEANT at small deposited energies, deviating no more than 10%. At the highest deposited

Figure 2: Comparison of point-cloud model (orange), image based model (grey-blue) and GEANT (black). Sum of all voxel energies is shown in (a), the total number of voxel hits is shown in (b), and the mean energy per $z$-layer is shown in (c). The dashed red lines in the bottom panels represent the 10% deviation interval of the generated samples from the original GEANT simulation. The earth mover’s distance (EMD) between each distribution and the GEANT distribution is also shown.
energies, however, both diffusion models begin to fall away from GEANT, with the point-cloud model generating less energy, and the image based model generating slightly more energy than GEANT. Events in this region are rare, however, and statistical fluctuations begin to dominate even GEANT. Figure 2(b) comparing total number of calorimeter hits. At low number of hits, both models show good agreement with GEANT, with deviations slightly above 10%. At 15 or more hits, both models begin to deviate well past 10%, with the point cloud model oversampling the number of hits, and the image based model undersampling the number of hits.

Figure 2(c) shows the average deposited energy along $z$. The mean deposited energy in $z$-coordinates in panel (c) show both models in very good agreement with the original GEANT predictions. The $z$ distribution shows the point cloud samples are systematically lower than the original GEANT distributions. This indicates the point cloud model would benefit from learning the energy per layer directly, as is done in the image model described Sec. 2.

5 Conclusion and Outlook

This work makes the first direct comparison between two score based generative models using either images or point clouds as representations of the same calorimeter training data. Both models perform well for most distributions, with very similar AUCs, but the image-based diffusion model invariably has a lower EMD in each comparison to GEANT. Overall, the performance of the point-cloud diffusion model is very close to the image model. This is despite the point cloud model being disadvantaged: the second model in it’s architecture is not conditioned on the energy per layer, unlike the image based model. At the same time, the point cloud model offers several advantages over the image model: Vastly smaller dataset size - about 100x smaller at full resolution, individual cell hits do not need to be summed in a voxelization procedure, resulting in information loss, and 3x faster generation times.

This work establishes a benchmark for future research on generative models, offering valuable insights into the challenges of modeling hadronic showers in highly granular calorimeters using image-based techniques, while also exploring the potential of point-cloud methods. The current advantages of point clouds, in combination with improvements to close the remaining performance gap described earlier, will likely make point cloud based models a clear choice for highly granular calorimeters. This work should serve as a reference for studies utilizing future calorimeters based on the CALICE design, including those intended for use in CMS at the LHC and ePIC at the EIC.

References


