ParaFlow: fast calorimeter simulations parameterized in upstream material configurations

Johannes Erdmann, Jonas Kann, Florian Mausolf, Peter Wissmann

RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany

Abstract

We study whether machine-learning models for fast calorimeter simulations can learn meaningful representations of calorimeter signatures that account for variations in the full particle detector's configuration. This may open new opportunities in high-energy physics measurements, for example in the assessment of systematic uncertainties that are related to the detector geometry, in the inference of properties of the detector configuration, or in the automated design of experiments. As a concrete example, we parameterize normalizing-flow-based simulations in configurations of the material upstream of a toy calorimeter. We call this model ParaFlow, which is trained to interpolate between different material budgets and positions, as simulated with Geant4. We study ParaFlow's performance in terms of photon shower shapes that are directly influenced by the properties of the upstream material, in which photons can convert to an electron-positron pair. In general, we find that ParaFlow is able to reproduce the dependence of the shower shapes on the material properties at the few-percent level with larger differences only in the tails of the distributions.

1 Introduction

The simulation of the detector response with GEANT4 [1] is an essential ingredient in almost all measurements in current high-energy physics, such as at the Large Hadron Collider (LHC). However, the computational cost that is associated with the production of simulated samples is large [2] and driven by the simulation of the shower development in the calorimeters. Alternatives that require less resources ("fast simulations") are hence needed. Traditionally, fast simulations use parameterizations of the calorimeter response [3–7]. Generative models based on deep learning, however, have been proposed as a promising approach for fast calorimeter simulations [8]. They have gained much attention (for recent reviews see Refs. [9, 10]) and first models are now in use at LHC experiments [5, 11].

These generative models are surrogates that are trained to reproduce predictions based on the "full simulation" with GEANT4 as closely as possible. Various generative machine learning approaches have been shown to be capable of simulating calorimeter showers of high quality, including models based on variational auto-encoders [12–21], generative adversarial networks [8, 22–36], flow-based models [37–46], diffusion models [47–56], and autoregressive models [57–61], following the taxonomy of Ref. [9]. A detailed comparison of the performance in terms of generation quality and timing benchmarks for a large range of models can be found in the HEP community challenge on calorimeter simulations [62].

Most of these works include models that have been trained to reproduce signatures in a specific detector. So far, only a few studies have gone beyond. In Refs. [58, 59, 63], a "geometry-aware" approach was proposed, which allows for varying sizes of the calorimeter readout cells. Alternatively, in Refs. [41, 42, 49, 50], "geometry-independent" models were developed that generate point clouds in a given detector material and are hence independent of the exact readout geometry. In Ref. [64], a "detector-agnostic" model was studied that is intended to be easily adaptable to other calorimeters. This was inspired by Ref. [65], which explored meta-learning to interpolate between calorimeters with different materials. In addition, an approach called "geometry adaption" was pursued in Ref. [52], which relies on a latent mapping of the calorimeter geometry. All of these works have focussed on relaxing the assumptions on the calorimeter geometry. While some of these works addressed aspects related to the readout segmentation for a given distribution of detector material [49, 50, 58, 59, 63], others aimed to learn the translation between different distributions of material in the calorimeters for a small set of scenarios [64, 65]. Although calorimeter signatures are directly influenced by the amount of material ("material budget") in front of the calorimeter

("upstream material"), the dependence of fast calorimeter simulations on other parts of the full particle detector has not gained much attention.

We ask whether generative models for calorimeter simulations can learn a meaningful representation of variable configurations for the full particle detector, i.e., in a single generative model. Such a model would have important implications for high-energy physics experiments: it could allow for a more efficient and more fine-grained evaluation of systematic uncertainties that are related to the detector configuration and which can be important [66, 67]; it could be used to estimate parameters of the detector configuration directly from data; and it could serve the automated design of particle physics experiments [68–76], where surrogate models can be used for gradient-based end-to-end optimizations.

We study this for a CMS-like toy calorimeter with an iron block with variable thickness and position in front, which serves as a proxy for the upstream material, i.e., the tracking detector, support material etc. The basic concept of the study is independent of the chosen generative model. As a state-of-the-art example, we use an architecture that is closely inspired by the CaloFlow model [37], which is based on normalizing flows (NFs) [77, 78]. We study photon signatures, as they are directly affected by the amount and position of the upstream material. The more upstream material is present, the more conversions to electron-positron pairs are expected to occur, resulting in wider showers in the calorimeter on average. In addition, the further away the material is positioned from the calorimeter, the longer the flight distance of electrons and positrons in the magnetic field, which again results in wider showers.

We choose a simple two-parameter model for the upstream material with one parameter for the thickness of the iron block and another parameter for its distance to the calorimeter. We train a single NF, which we call ParaFlow, on photons simulated with Geant to learn the dependence of the photon's calorimeter signature as a function of these parameters. We then compare ParaFlow's predictions to the expectation from Geant4 simulations to assess the quality of the surrogate model. We finally conclude whether ParaFlow is able to learn a meaningful representation of the calorimeter signature as a function of the detector configuration parameters.

2 Simulated samples

We simulate a toy calorimeter that is inspired by the electromagnetic barrel calorimeter of the CMS detector [79]. The simulation is based on the framework of Ref. [8] with GEANT4 11.2.2 and builds on the CMS-like toy calorimeter from Ref. [36]. We simulate PbWO₄ scintillating crystals with a length of 230 mm and a front face of $22 \times 22 \,\mathrm{mm}^2$. The front of the calorimeter is placed at a distance of 1.29 m from a GEANT4 particle gun that simulates photons. These photons have a uniformly distributed energy between 20 GeV and 100 GeV, where the direction of the photons is perpendicular to the front face of the detector. This energy range is chosen to represent typical photon energies analysed at the LHC, for example, in measurements targetting $H \to \gamma \gamma$ decays. To distribute the energy deposition in the detector such that not all photons hit the toy calorimeter centrally, the position of the source is smeared. For that, a Gaussian distribution with a width of 44 mm is chosen. This corresponds to the width of two crystals. A magnetic field with a field strength of $B=4\,\mathrm{T}$ is added in x-direction, as shown in Fig. 1. The magnetic field widens the photon showers in one direction, as in typical particle detectors, and its strength is chosen to be close to the field strength of the CMS solenoid. To mimic the effect of material upstream of the calorimeter on the photon conversion rate and on the shower formation, an iron plate with a front surface of 1×1 m² is added between the source and the calorimeter. Its distance to the calorimeter, b, and its thickness, d are parameters of the simulation setup. They were varied uniformly in the ranges $50 \text{ cm} \le b \le 90 \text{ cm}$ and $0.5 X_0 \le d \le 1.5 X_0$, where $X_0 = 1.757 \text{ cm}$ is the radiation length of iron. The central values of the intervals are inspired by the CMS detector [80]. The interval ranges are chosen such that the variations have a significant impact on the shower properties within the statistics of the simulated sample. We did not simulate noise in the calorimeter, nor other aspects of digitization. Approximately 1,000,000 photon showers were simulated in total, of which 60% were used for training, 10% for validation, and 30% for testing purposes.

3 Machine-learning architecture

Normalizing flows map the input data to a base space via a learnable bijective transformation. Generated samples in the input space are obtained by sampling from the known distribution in the base space, typically a multidimensional Gaussian distribution. As in Ref. [37], we use masked autoregressive flows (MAFs) [81] with monotonic rational quadratic spline (RQS) transformations [82].

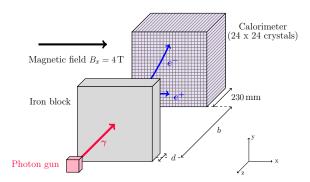


Figure 1: Sketch of the toy detector. The particle gun shoots photons perpendicular to the PbWO₄ calorimeter. The photons may undergo conversion in the iron block, which represents the material upstream of the calorimeter, producing an e^+e^- pair. The whole volume contains a homogenous magnetic field. The distance b between the centre of the iron block and the calorimeter front face, as well as the thickness d of the iron block are variable.

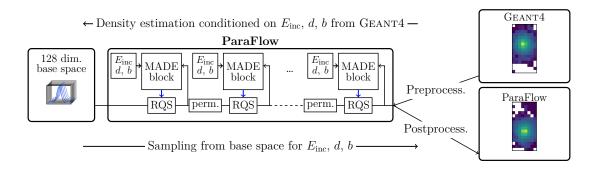


Figure 2: Illustration of the NF architecture. The figure design is inspired by Ref. [37].

The parameters of the RQS transformations are calculated by MADE blocks [83]. In Ref. [37], two NFs were used, where the first predicts the layer-wise energy deposition in a multi-layer calorimeter, and the second generates the distribution of the energy within a layer. As our calorimeter consists of a single layer, we use a single NF.

The architecture is illustrated in Fig. 2. The energy depositions of GEANT4-simulated photons in the calorimeter crystals ("calorimeter images") are mapped to a 128-dimensional base space, conditioned on the incoming photon energy $E_{\rm inc}$ and the material variables d and b. The calorimeter images are flattened before being passed to the NF, appended by the conditions. Between the MADE blocks, the variables are randomly permuted.

The main hyperparameters are the number of MADE blocks, the structure of the multi-layer perceptrons (MLPs) within the MADE blocks, and the number of bins for the RQS transformations. We build the NF with ten MADE blocks and ten spline bins. These parameters were optimised in a coarse grid search based on the area under the ROC curve of a ResNet [84] classifier that was trained to discriminate generated from Geant-simulated photons. The MADE blocks consist of MLPs with two hidden layers and 196 nodes each, activated with ReLU. Wider MLPs were not found to improve the performance. The resulting model has approximately 8 million parameters.

The calorimeter images are preprocessed in the following way: The 24×24 simulated energy depositions in the calorimeter crystals are cropped to the 8×16 window that contains the maximum energy sum. The window's asymmetry accounts for the impact of the magnetic field on the shower formation. Even for showers generated with the thickest iron plate and largest distance from the calorimeter, i.e., where the widest showers are expected, on average less than 1% of the energy is deposited outside of the window. As in Ref. [37], we add a small noise, which we choose to be uniformly distributed between 0 and 20 keV. The energy deposits in the individual crystals are then divided by the incoming photon energy $E_{\rm inc}$ and processed by a logit transformation, as in Ref. [37]. The conditions c_i are logarithmically transformed, such that they are mapped to the domain [-1,1]:

$$c_i \rightarrow 2 \cdot \frac{\log(c_i - (c_{i,\min} - 1))}{\log(c_{i,\max} - (c_{i,\min} - 1))} - 1 \in [-1, 1], \text{ for } c_i \in [c_{i,\min}, c_{i,\max}]$$
 (1)

The NFs are implemented in PyTorch [85] using the Zuko package [86]. The training is performed with a batch-size of 1000 using the Adam optimizer [87] with an initial learning rate of 10^{-3} and a cosine annealing learning-rate scheduler with warm restarts [88]. The restarting learning rate is reduced by a factor of 0.5 on plateaus of the validation loss with a patience of 10 epochs. The training is stopped if the validation loss did not improve over 15 epochs.

4 Results

4.1 Qualitative comparisons

Figure 3 shows example generated calorimeter images for varying values of material budget (d) and position (b) upstream of the calorimeter, for an energy of $E_{\rm inc}=30\,{\rm GeV}$. These are compared with randomly selected Geant4 images with very similar values of d and b. As expected, the generated and the Geant4-simulated photons show a trend towards more energy clusters in the image when the material budget increases, as the conversion probability rises. The clusters also tend to be separated more strongly when the distance of the iron block and the calorimeter increases, because of the longer distance that electrons and positrons from such conversions travel in the magnetic field.

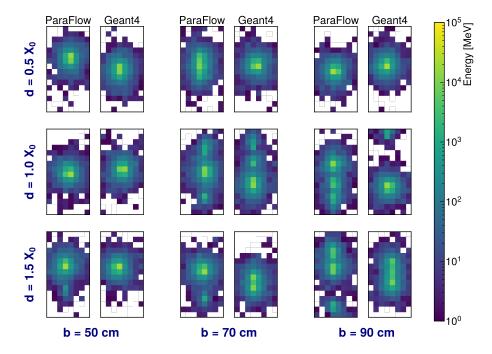


Figure 3: Generated calorimeter images of 30 GeV photons for varying material conditions d and b, compared with randomly selected Geant4 images with very similar values of d and b.

Figure 4 shows the average of all GEANT4 calorimeter images in the test dataset (300,000 photons), the average of 1,000,000 calorimeter images generated with ParaFlow with uniform conditions, as well as their relative difference. Among the central 4×4 crystals, the highest relative deviation amounts to 2%. For the crystals with larger distances from the centre, the deviations mostly remain below 15%. Since the average energies in these crystals is of the order of $10\,\mathrm{MeV}$, they are expected to only contribute very little to reconstructed shower properties, such as the shower shape observables discussed below.

4.2 Distributions of shower shape observables

We compare the predictions obtained from the ParaFlow samples with the corresponding Geant4 results using shower shape observables. We quantify the agreement using the separation power of histograms, also used in Ref. [62], and given by

$$S(h_1, h_2) = \frac{1}{2} \sum_{i} \frac{(h_{1,i} - h_{2,i})^2}{h_{1,i} + h_{2,i}},$$
(2)

ParaFlow Average images

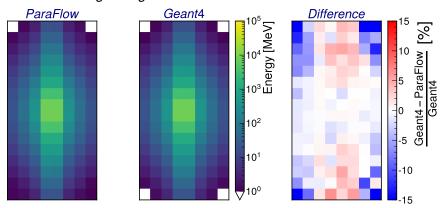


Figure 4: Averages of Geant4-simulated and ParaFlow-generated calorimeter images, taken across all energies and detector conditions. On the right, their relative difference is shown.

where $h_{n,i}$ is the count of the i^{th} bin of the n^{th} (n = 1, 2) normalised histogram. A separation power of S = 0 indicates perfectly matching histograms, and S = 1 quantifies vanishing agreement.

To quantify the overall performance of ParaFlow in describing the GEANT4 samples, the energy distributions in the brightest and second brightest crystals are shown in Fig. 5. These observables are largely insensitive to variations in the material upstream of the calorimeter and allow us to assess whether our model shows the good description that is expected for a CaloFlow-based model. In contrast to Ref. [37], however, we did not normalise the energies in Fig. 5 by the total deposited energy in our one-layer calorimeter, as we expect ParaFlow to also learn the overall energy. For the energy distribution of the brightest crystal, we see a good agreement between the ParaFlow samples and the GEANT4 data. The deviations remain below 5% over almost the entire range. Similarly, we find deviations below 5% in the bulk of the distribution of the second-brightest crystal. Deviations occur in the high-energy tail that may be due to an insufficient representation of shower images with a high-energy second-brightest crystal in the training data. Overall, the histograms show a good separation power of $\mathcal{O}(10^{-4})$, which indeed indicates good performance of our model.

To assess whether ParaFlow is able to accurately capture the dependence of the detector simulation on the material upstream of the calorimeter, we divide the ParaFlow test set and the GEANT4-generated dataset into equal-sized subsets, either divided by d or b (i.e., inclusive in the other variable). As in Fig. 4, the ParaFlow test set is comprised of 1,000,000 ParaFlow-generated images with uniform conditions, and the GEANT4 test set contains 300,000 simulated photons. The subsets are defined by the boundaries $[0.5, 0.75, 1.0, 1.25, 1.5] X_0$ for d, and [50, 60, 70, 80, 90] cm for b. We use the following shower shape observables to assess ParaFlow's performance in describing the dependence on the material. These observables are particularly sensitive to the location and probability of conversions in the material upstream of the calorimeter:

• The width of the shower is defined as

$$w = \frac{\sum_{i} r_i E_i}{\sum_{i} E_i} \,, \tag{3}$$

with E_i the energy deposited in the i^{th} crystal and r_i its Euclidean distance to the barycentre of the shower in units of crystal widths.

• The variable R_9 is defined as the sum of the energies in the 3×3 crystals surrounding the crystal with the highest energy deposition, divided by the total energy, $E_{\text{tot}} = \sum_i E_i$:

$$R_9 = \frac{E_{3\times3}}{E_{\text{tot}}} \tag{4}$$

This quantity is commonly used in the CMS experiment to distinguish converted from unconverted photons [89, 90]. Lower values are expected for converted photons, in particular for early conversions, i.e., large values of b.

• We use the DBSCAN clustering algorithm [91] to define the number of energy clusters in the calorimeter image. This algorithm relies on three parameters: ε , the maximum distance

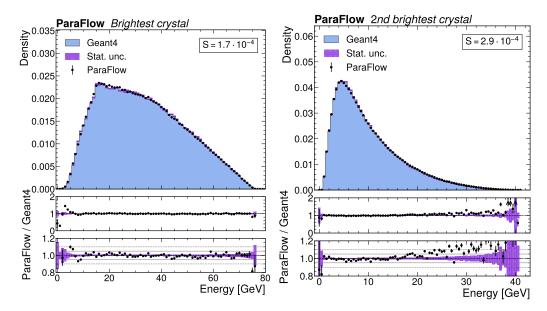


Figure 5: Distributions of the energy in the brightest (left) and second-brightest (right) crystal in the calorimeter in the full dataset, i.e., integrated over the two parameters that describe the material upstream of the calorimeter, as well as over photon energies. The plots below show the ratio of ParaFlow over Geant4 on two different y-axis scales. Both observables are largely insensitive to the variations of the upstream material and are used to assess the overall performance of ParaFlow.

between points to be considered neighbours, minPts, the minimum number of points needed to form a cluster, and the energy threshold E_{\min} to filter out low-energy crystals. Crystals with $E > E_{\min}$ and which meet the minPts requirement within the radius ε form clusters. We choose $\varepsilon = 1.5$ crystal widths to include only direct neighbours, minPts= 1, and $E_{\min} = 400 \, \text{MeV}$, to stay above typical experimental noise levels.

Figures 6-8 show the performance of ParaFlow in describing the material dependence of these sensitive observables:

- Fig. 6 shows the distributions of the shower width. It is evident that the shower width depends on the characteristics of the material upstream of the calorimeter. An increase in material thickness d leads to larger shower widths, as expected for a larger conversion rate. Larger shower widths are also expected when conversions occur earlier, i.e., with a larger distance b from the calorimeter. These trends are clearly visible in the Geant simulation, and well reproduced in the shower images generated with ParaFlow.
- Fig. 7 shows the distribution of R_9 . The distributions show the expected trends, i.e., lower average values and stronger tails towards low values of R_9 for a larger material budget (larger values of d) and a larger distance between the conversion vertices and the calorimeter (b). Again, ParaFlow reproduces these trends and closely follows the distributions from GEANT4.
- Fig. 8 shows the distributions of the number of energy clusters. We observe the expected dependence: with increasing conversion rates and for earlier conversions, the number of detected clusters tends to grow. ParaFlow reproduces these dependencies properly. In particular, the share of one-cluster images (unconverted photons) and two-cluster images (converted photons) is correctly reproduced at the percent level. For larger numbers of clusters, however, we find an underestimation by ParaFlow in comparison to the Geant4-simulated calorimeter images. This is possibly due to the under-representation of shower images with very large cluster counts (≥ 4) in the training set.

In general, the calorimeter images generated with ParaFlow agree with the predictions from GEANT4 within better than 5% in the bulk of the distributions and with separation powers of $\mathcal{O}(10^{-4})$. The agreement is at the same level as observed for the inclusive distribution of material-insensitive observables (Fig. 5), with larger differences observed in the tails of some distributions. However, no systematic trends are observed that would indicate that ParaFlow's performance depends on the configuration of the upstream material that it is parameterized in.

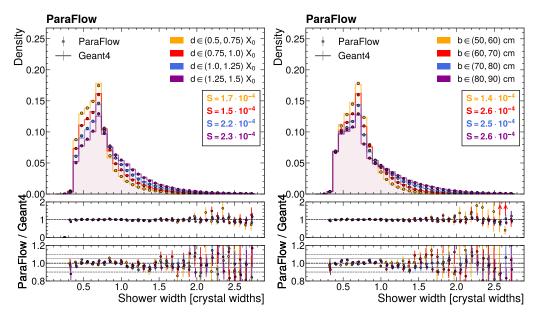


Figure 6: Distributions of the shower width, w, subdivided according to the values of the material parameters: distributions for varying material budget d (left) and for varying material distance b (right). The plots below show the ratio of ParaFlow over Geant4 on two different y-axis scales.

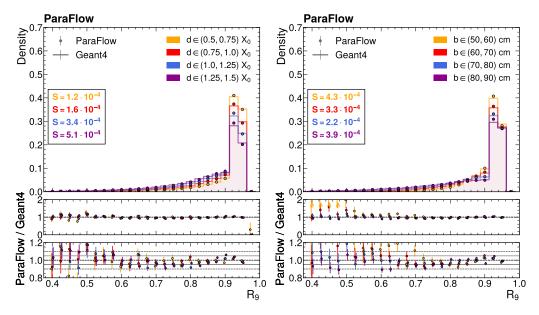


Figure 7: Distributions of R_9 , subdivided according to the values of the material parameters: distributions for varying material budget d (left) and for varying material distance b (right). The plots below show the ratio of ParaFlow over Geant4 on two different y-axis scales.

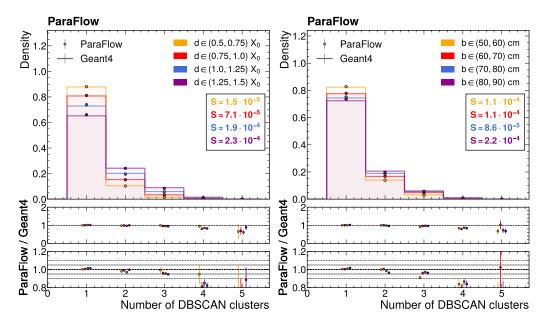


Figure 8: Distributions of the number of energy clusters, as found by the DBSCAN algorithm, subdivided according to the values of the material parameters: distributions for varying material budget d (left) and for varying material distance b (right). The plots below show the ratio of ParaFlow over GEANT4 on two different y-axis scales.

Figure 9 shows the mean values of these three observables (shower width, R_9 , number of clusters), now as a function of both parameters of the upstream material, d and b. We observe the expected trends of the shower shape observables, i.e., wider showers, lower values of R_9 and more clusters, the more material is present upstream of the calorimeter and the further away it is from the calorimeter. ParaFlow reproduces these trends very well. The relative differences in the mean values are below 6% for all of these shower shapes. For the R_9 variable, the deviations are even smaller, which we attribute to the fact that the distribution of R_9 is generally narrower (cf. Fig. 7). Figure 9 hence provides further evidence that ParaFlow is able to capture the dependence of the calorimeter images on the amount and position of the upstream material.

ParaFlow

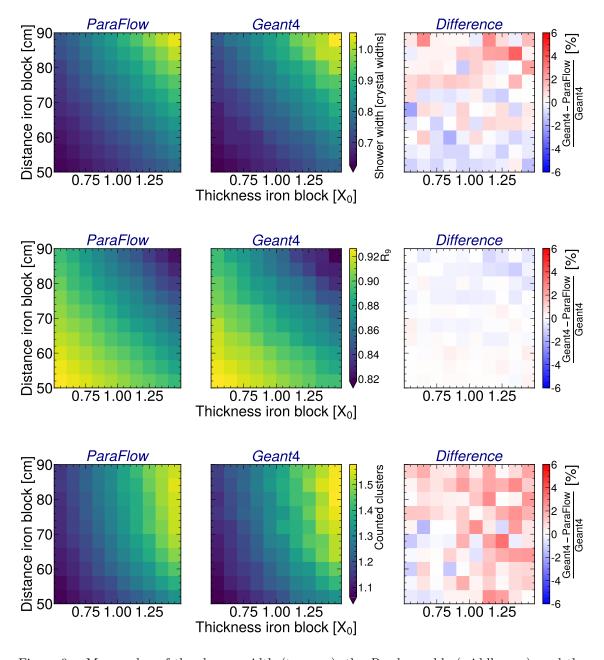


Figure 9: Mean value of the shower width (top row), the R_9 observable (middle row), and the number of clusters (bottom row) as a function the upstream material budget (thickness of the iron block, d, in units of X_0) and of the distance between the upstream material from the calorimeter (distance of the iron block, b, in units of cm), shown for ParaFlow (left column) and for Geant4 (middle column). Their relative difference (right column) is also shown.

5 Conclusions

We have proposed ParaFlow, a fast calorimeter simulation that is parameterized as a function of the particle detector's configuration. While the concept is independent of the specific generative model, we studied it based on CaloFlow [37], which uses normalizing flows as the generative model. We studied ParaFlow for a CMS-like toy calorimeter and for the example case of varying configurations of the upstream material. As a simplified model for this material, we used an iron block that is characterized by its thickness and its distance from the calorimeter. We observed the expected changes in photon calorimeter signatures in GEANT4 simulations, which we then used to train ParaFlow. We evaluated the performance of ParaFlow by comparing its predictions for photon shower shapes to the expectation from GEANT4 simulations, as a function of the material configurations. For the comparison, we chose a set of shower shape observables that is particularly sensitive to the properties of photon conversions to electron-positron pairs in the upstream material.

In general, we observe that ParaFlow describes the GEANT4 samples well, with an agreement of better than 5% in the bulk of the distributions for variations of the amount of upstream material, as well as for the variations of its distance from the calorimeter. Larger deviations are only observed in the tails of the distributions, which are not strongly represented in the training dataset. Overall, we observe a similar separation power of $\mathcal{O}(10^{-4})$ for these shower shape observables. We investigated the dependence of the means of the shower shape observables when both, material budget and distance, were varied simultaneously. We find again very good agreement of the ParaFlow predictions with the expectations from the GEANT4 simulations. We conclude that ParaFlow is indeed able to a learn a representation of calorimeter signatures as a function of properties of the upstream material in a single model.

This work provides evidence that fast calorimeter simulations can be used to learn a representation of variable configurations of the full particle detector. This is not possible in GEANT4, as it requires a separate simulation for each single configuration. In future studies, ParaFlow should be extended to parameterizations of more complex variations in the detector configurations. This may open the door to important applications: it would allow for an efficient and more fine-grained evaluation of systematic uncertainties that are related to uncertainties in the detector configurations, such as uncertainties in the overall material budget or material distribution; it may also allow to infer parameters of detector configurations directly from data and hence improve the understanding of the detector in particle physics experiments; it may also be used as a differential surrogate model in gradient-based optimizations of detector designs.

Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (DFG) under grants 400140256 - GRK 2497 (The physics of the heaviest particles at the LHC, JE and FM) and 686709 - ER 866/1-1 (Heisenberg Programme, JE), and by the Studienstiftung des deutschen Volkes (FM).

References

- [1] S. Agostinelli et al., Geant4 a simulation toolkit, Nucl. Instrum. Meth. A 506 (2003) 250.
- [2] ATLAS collaboration, ATLAS HL-LHC Computing Conceptual Design Report, 2020, CERN-LHCC-2020-015, LHCC-G-178, https://cds.cern.ch/record/2729668.
- [3] ATLAS collaboration, M. Beckingham, M. Duehrssen, E. Schmidt, M. Shapiro, M. Venturi et al., The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim, 2010, ATL-PHYS-PUB-2010-013 https://cds.cern.ch/record/1300517.
- [4] ATLAS collaboration, Performance of the Fast ATLAS Tracking Simulation (FATRAS) and the ATLAS Fast Calorimeter Simulation (FastCaloSim) with single particles, 2014, ATL-SOFT-PUB-2014-001, https://cds.cern.ch/record/1669341.
- [5] ATLAS collaboration, G. Aad et al., AtlFast3: The Next Generation of Fast Simulation in ATLAS, Comput. Softw. Big Sci. 6 (2022) 7, [2109.02551].
- [6] CMS collaboration, S. Abdullin, P. Azzi, F. Beaudette, P. Janot and A. Perrotta, The Fast Simulation of the CMS Detector at LHC, J. Phys. Conf. Ser. 331 (2011) 032049.
- [7] CMS collaboration, M. Hildreth, V. N. Ivanchenko and D. J. Lange, *Upgrades for the CMS simulation*, J. Phys. Conf. Ser. 898 (2017) 042040.

- [8] M. Paganini, L. de Oliveira and B. Nachman, CaloGAN: Simulating 3D high energy particle showers in multilayer electromagnetic calorimeters with generative adversarial networks, Phys. Rev. D 97 (2018) 014021, [1712.10321].
- [9] B. Hashemi and C. Krause, Deep generative models for detector signature simulation: A taxonomic review, Rev. Phys. 12 (2024) 100092, [2312.09597].
- [10] F. Y. Ahmad, V. Venkataswamy and G. Fox, A Comprehensive Evaluation of Generative Models in Calorimeter Shower Simulation, 2406.12898.
- [11] M. Barbetti, Lamarr: LHCb ultra-fast simulation based on machine learning models deployed within Gauss, Proceedings of the 21st International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT 2022), 2303.11428.
- [12] A. Ghosh on behalf of the ATLAS Collaboration, Deep generative models for fast shower simulation in ATLAS, J. Phys. Conf. Ser. 1525 (2020) 012077.
- [13] K. Deja, J. Dubiński, P. Nowak, S. Wenzel and T. Trzciński, End-to-end Sinkhorn Autoencoder with Noise Generator, 2006.06704.
- [14] E. Buhmann, S. Diefenbacher, E. Eren, F. Gaede, G. Kasieczka, A. Korol et al., Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed, Comput. Softw. Big Sci. 5 (2021) 13, [2005.05334].
- [15] E. Buhmann, S. Diefenbacher, E. Eren, F. Gaede, G. Kasieczka, A. Korol et al., Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network, EPJ Web Conf. 251 (2021) 03003, [2102.12491].
- [16] E. Buhmann, S. Diefenbacher, D. Hundhausen, G. Kasieczka, W. Korcari, E. Eren et al., Hadrons, better, faster, stronger, Mach. Learn. Sci. Tech. 3 (2022) 025014, [2112.09709].
- [17] S. Diefenbacher, E. Eren, F. Gaede, G. Kasieczka, A. Korol, K. Krüger et al., New angles on fast calorimeter shower simulation, Mach. Learn. Sci. Tech. 4 (2023) 035044, [2303.18150].
- [18] A. Hariri, D. Dyachkova and S. Gleyzer, Graph Generative Models for Fast Detector Simulations in High Energy Physics, 2104.01725.
- [19] A. Abhishek, E. Drechsler, W. Fedorko and B. Stelzer, CaloDVAE: Discrete Variational Autoencoders for Fast Calorimeter Shower Simulation, Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS 2021) Workshop on Machine Learning and the Physical Sciences, 2210.07430.
- [20] J. C. Cresswell, B. L. Ross, G. Loaiza-Ganem, H. Reyes-Gonzalez, M. Letizia and A. L. Caterini, CaloMan: Fast generation of calorimeter showers with density estimation on learned manifolds, Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS 2022) Workshop on Machine Learning and the Physical Sciences, 2211.15380.
- [21] Q. Liu, C. Shimmin, X. Liu, E. Shlizerman, S. Li and S.-C. Hsu, Calo-VQ: Vector-Quantized Two-Stage Generative Model in Calorimeter Simulation, 2405.06605.
- [22] L. de Oliveira, M. Paganini and B. Nachman, Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis, Comput. Softw. Big Sci. 1 (2017) 4, [1701.05927].
- [23] L. de Oliveira, M. Paganini and B. Nachman, Controlling Physical Attributes in GAN-Accelerated Simulation of Electromagnetic Calorimeters, J. Phys. Conf. Ser. 1085 (2018) 042017, [1711.08813].
- [24] M. Paganini, L. de Oliveira and B. Nachman, Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters, Phys. Rev. Lett. 120 (2018) 042003, [1705.02355].
- [25] 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, [2109.07388].

- [26] G. R. Khattak, S. Vallecorsa and F. Carminati, Three Dimensional Energy Parametrized Generative Adversarial Networks for Electromagnetic Shower Simulation, Proceedings of the 25th IEEE International Conference on Image Processing (ICIP 2018) (2018) 3913.
- [27] S. Vallecorsa, F. Carminati and G. Khattak, 3D convolutional GAN for fast simulation, EPJ Web Conf. 214 (2019) 02010.
- [28] D. Belayneh et al., Calorimetry with deep learning: particle simulation and reconstruction for collider physics, Eur. Phys. J. C 80 (2020) 688, [1912.06794].
- [29] P. Musella and F. Pandolfi, Fast and Accurate Simulation of Particle Detectors Using Generative Adversarial Networks, Comput. Softw. Big Sci. 2 (2018) 8, [1805.00850].
- [30] V. Chekalina, E. Orlova, F. Ratnikov, D. Ulyanov, A. Ustyuzhanin and E. Zakharov, Generative Models for Fast Calorimeter Simulation: the LHCb case, EPJ Web Conf. 214 (2019) 02034, [1812.01319].
- [31] S. Diefenbacher, E. Eren, G. Kasieczka, A. Korol, B. Nachman and D. Shih, DCTRGAN: improving the precision of generative models with reweighting, JINST 15 (2020) P11004, [2009.03796].
- [32] K. Jaruskova and S. Vallecorsa, Ensemble Models for Calorimeter Simulations, J. Phys. Conf. Ser. 2438 (2023) 012080.
- [33] M. Faucci Giannelli and R. Zhang, CaloShowerGAN, a generative adversarial network model for fast calorimeter shower simulation, Eur. Phys. J. Plus 139 (2024) 597, [2309.06515].
- [34] M. Erdmann, J. Glombitza and T. Quast, Precise Simulation of Electromagnetic Calorimeter Showers Using a Wasserstein Generative Adversarial Network, Comput. Softw. Big Sci. 3 (2019) 4, [1807.01954].
- [35] F. Carminati, A. Gheata, G. Khattak, P. Mendez Lorenzo, S. Sharan and S. Vallecorsa, Three dimensional Generative Adversarial Networks for fast simulation, J. Phys. Conf. Ser. 1085 (2018) 032016.
- [36] J. Erdmann, A. van der Graaf, F. Mausolf and O. Nackenhorst, SR-GAN for SR-gamma: super resolution of photon calorimeter images at collider experiments, Eur. Phys. J. C 83 (2023) 1001, [2308.09025].
- [37] C. Krause and D. Shih, Fast and accurate simulations of calorimeter showers with normalizing flows, Phys. Rev. D 107 (2023) 113003, [2106.05285].
- [38] C. Krause and D. Shih, Accelerating accurate simulations of calorimeter showers with normalizing flows and probability density distillation, Phys. Rev. D 107 (2023) 113004, [2110.11377].
- [39] C. Krause, I. Pang and D. Shih, CaloFlow for CaloChallenge dataset 1, SciPost Phys. 16 (2024) 126, [2210.14245].
- [40] I. Pang, D. Shih and J. A. Raine, Calorimeter shower superresolution, Phys. Rev. D 109 (2024) 092009, [2308.11700].
- [41] S. Schnake, D. Krüker and K. Borras, Generating Calorimeter Showers as Point Clouds, Paper at Workshop Machiner Learning and the Physical Sciences at the 37th conference on Neural Information Processing Systems (NeurIPS 2022), https://ml4physicalsciences.github.io/2022/files/NeurIPS_ML4PS_2022_77.pdf.
- [42] S. Schnake, D. Krücker and K. Borras, CaloPointFlow II Generating Calorimeter Showers as Point Clouds, 2403.15782.
- [43] T. Buss, F. Gaede, G. Kasieczka, C. Krause and D. Shih, Convolutional L2LFlows: generating accurate showers in highly granular calorimeters using convolutional normalizing flows, JINST 19 (2024) P09003, [2405.20407].
- [44] E. Dreyer, E. Gross, D. Kobylianskii, V. Mikuni, B. Nachman and N. Soybelman, Automated Approach to Accurate, Precise, and Fast Detector Simulation and Reconstruction, Phys. Rev. Lett. 133 (2024) 211902, [2406.01620].

- [45] F. Ernst, L. Favaro, C. Krause, T. Plehn and D. Shih, Normalizing flows for high-dimensional detector simulations, SciPost Phys. 18 (2025) 081, [2312.09290].
- [46] L. Favaro, A. Ore, S. P. Schweitzer and T. Plehn, CaloDREAM Detector response emulation via attentive flow matching, SciPost Phys. 18 (2025) 088, [2405.09629].
- [47] V. Mikuni and B. Nachman, Score-based generative models for calorimeter shower simulation, Phys. Rev. D 106 (2022) 092009, [2206.11898].
- [48] V. Mikuni and B. Nachman, CaloScore v2: single-shot calorimeter shower simulation with diffusion models, JINST 19 (2024) P02001, [2308.03847].
- [49] E. Buhmann, S. Diefenbacher, E. Eren, F. Gaede, G. Kasieczka, A. Korol et al., CaloClouds: fast geometry-independent highly-granular calorimeter simulation, JINST 18 (2023) P11025, [2305.04847].
- [50] E. Buhmann, F. Gaede, G. Kasieczka, A. Korol, W. Korcari, K. Krüger et al., CaloClouds II: ultra-fast geometry-independent highly-granular calorimeter simulation, JINST 19 (2024) P04020, [2309.05704].
- [51] S. Diefenbacher, V. Mikuni and B. Nachman, Refining Fast Calorimeter Simulations with a Schrödinger Bridge, 2308.12339.
- [52] O. Amram and K. Pedro, Denoising diffusion models with geometry adaptation for high fidelity calorimeter simulation, Phys. Rev. D 108 (2023) 072014, [2308.03876].
- [53] F. T. Acosta, V. Mikuni, B. Nachman, M. Arratia, B. Karki, R. Milton et al., Comparison of point cloud and image-based models for calorimeter fast simulation, JINST 19 (2024) P05003, [2307.04780].
- [54] D. Kobylianskii, N. Soybelman, E. Dreyer and E. Gross, Graph-based diffusion model for fast shower generation in calorimeters with irregular geometry, Phys. Rev. D 110 (2024) 072003, [2402.11575].
- [55] D. Kobylianskii, N. Soybelman, N. Kakati, E. Dreyer, B. Nachman and E. Gross, Advancing set-conditional set generation: Diffusion models for fast simulation of reconstructed particles, Phys. Rev. D 110 (2024) 092013, [2405.10106].
- [56] T. Madula and V. M. Mikuni, CaloLatent: Score-based Generative Modelling in the Latent Space for Calorimeter Shower Generation, Paper at Workshop Machiner Learning and the Physical Sciences at the 38th conference on Neural Information Processing Systems (NeurIPS 2023), https://ml4physicalsciences.github.io/2023/files/NeurIPS_ML4PS_2023_19.pdf.
- [57] Y. Lu, J. Collado, D. Whiteson and P. Baldi, Sparse autoregressive models for scalable generation of sparse images in particle physics, Phys. Rev. D 103 (2021) 036012, [2009.14017].
- [58] J. Liu, A. Ghosh, D. Smith, P. Baldi and D. Whiteson, Geometry-aware Autoregressive Models for Calorimeter Shower Simulations, Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS 2022) Workshop on Machine Learning and the Physical Sciences, 2212.08233.
- [59] J. Liu, A. Ghosh, D. Smith, P. Baldi and D. Whiteson, Generalizing to new geometries with Geometry-Aware Autoregressive Models (GAAMs) for fast calorimeter simulation, JINST 18 (2023) P11003, [2305.11531].
- [60] S. Diefenbacher, E. Eren, F. Gaede, G. Kasieczka, C. Krause, I. Shekhzadeh et al., L2LFlows: generating high-fidelity 3D calorimeter images, JINST 18 (2023) P10017, [2302.11594].
- [61] M. R. Buckley, C. Krause, I. Pang and D. Shih, Inductive simulation of calorimeter showers with normalizing flows, Phys. Rev. D 109 (2024) 033006, [2305.11934].
- [62] C. Krause, M. Faucci Giannelli, G. Kasieczka, B. Nachman, D. Salamani, D. Shih et al., CaloChallenge 2022: A Community Challenge for Fast Calorimeter Simulation, 2410.21611.
- [63] D. Smith, A. Ghosh, J. Liu, P. Baldi and D. Whiteson, Fast multi-geometry calorimeter simulation with conditional self-attention variational autoencoders, 2411.05996.

- [64] P. Raikwar, R. Cardoso, N. Chernyavskaya, K. Jaruskova, W. Pokorski, D. Salamani et al., Transformers for Generalized Fast Shower Simulation, EPJ Web Conf. 295 (2024) 09039.
- [65] D. Salamani, A. Zaborowska and W. Pokorski, MetaHEP: Meta learning for fast shower simulation of high energy physics experiments, Phys. Lett. B 844 (2023) 138079.
- [66] ATLAS collaboration, G. Aad et al., Measurement of the Higgs boson mass with $H \to \gamma \gamma$ decays in 140 fb⁻¹ of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector, Phys. Lett. B 847 (2023) 138315, [2308.07216].
- [67] CMS collaboration, A. M. Sirunyan et al., A measurement of the Higgs boson mass in the diphoton decay channel, Phys. Lett. B 805 (2020) 135425, [2002.06398].
- [68] MODE collaboration, A. G. Baydin et al., Toward Machine Learning Optimization of Experimental Design, Nuclear Physics News 31 (2021) 25.
- [69] MODE collaboration, T. Dorigo et al., Toward the end-to-end optimization of particle physics instruments with differentiable programming, Reviews in Physics 10 (2023) 100085, [2203.13818].
- [70] S. Shirobokov, V. Belavin, M. Kagan, A. Ustyuzhanin and A. G. Baydin, Black-Box Optimization with Local Generative Surrogates, Proceedings of Advances in Neural Information Processing Systems 34 (NeurIPS 2020), 2002.04632.
- [71] C. Neubüser, J. Kieseler and P. Lujan, Optimising longitudinal and lateral calorimeter granularity for software compensation in hadronic showers using deep neural networks, Eur. Phys. J. C 82 (2022) 92, [2101.08150].
- [72] M. Aehle et al., Progress in End-to-End Optimization of Detectors for Fundamental Physics with Differentiable Programming, 2310.05673.
- [73] M. Aehle, M. Novák, V. Vassilev, N. R. Gauger, L. Heinrich, M. Kagan et al., Optimization Using Pathwise Algorithmic Derivatives of Electromagnetic Shower Simulations, 2405.07944.
- [74] G. C. Strong et al., TomOpt: differential optimisation for task- and constraint-aware design of particle detectors in the context of muon tomography, Mach. Learn. Sci. Tech. 5 (2024) 035002, [2309.14027].
- [75] MODE collaboration, K. Schmidt et al., End-to-End Detector Optimization with Diffusion models: A Case Study in Sampling Calorimeters, 2502.02152.
- [76] K. A. Woźniak, S. Mulligan, J. Kieseler, M. Klute, F. Fleuret and T. Golling, End-to-End Optimal Detector Design with Mutual Information Surrogates, 2503.14342.
- [77] E. G. Tabak and C. V. Turner, A Family of Nonparametric Density Estimation Algorithms, Commun. Pure Appl. Math. 66 (2013) 145.
- [78] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed and B. Lakshminarayanan, Normalizing Flows for Probabilistic Modeling and Inference, J. Mach. Learn. Res. 22 (2021) 1, [1912.02762].
- [79] CMS collaboration, S. Chatrchyan et al., The CMS Experiment at the CERN LHC, JINST 3 (2008) S08004.
- [80] CMS collaboration, S. Chatrchyan et al., Description and performance of track and primary-vertex reconstruction with the CMS tracker, JINST 9 (2014) P10009, [1405.6569].
- [81] G. Papamakarios, T. Pavlakou and I. Murray, Masked Autoregressive Flow for Density Estimation, Proceedings of Advances in Neural Information Processing Systems 30 (NIPS 2017), 1705.07057.
- [82] C. Durkan, A. Bekasov, I. Murray and G. Papamakarios, Neural Spline Flows, Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2019), 1906.04032.

- [83] M. Germain, K. Gregor, I. Murray and H. Larochelle, MADE: Masked Autoencoder for Distribution Estimation, Proceedings of the 32nd International Conference on Machine Learning (ICML 2015), PMLR 37 (2015) 881, 1502.03509.
- [84] K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition, Proceedings of the 2016 Conference on Computer Vision and Pattern Recognition (CVPR 2016), 1512.03385.
- [85] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan et al., PyTorch: An Imperative Style, High-Performance Deep Learning Library, Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2019), 1912.01703.
- [86] F. Rozet et al., Zuko: Normalizing flows in PyTorch, https://pypi.org/project/zuko.
- [87] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015), 1412.6980.
- [88] I. Loshchilov and F. Hutter, SGDR: Stochastic Gradient Descent with Warm Restarts, Proceedings of the 5th International Conference on Learning Representations (ICLR 2017), 1608.03983.
- [89] CMS collaboration, V. Khachatryan et al., Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV, JINST 10 (2015) P08010, [1502.02702].
- [90] CMS collaboration, A. M. Sirunyan et al., Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC, JINST 16 (2021) P05014, [2012.06888].
- [91] M. Ester, H. Kriegel, J. Sander and X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96) 226, https://cdn.aaai.org/KDD/1996/KDD96-037.pdf.