Zero Degree Calorimeter Fast Simulation with Normalizing Flows

Emilia Majerz[1,*] and Witold Dzwinel[1]

on behalf of the ALICE Collaboration

[1]AGH University of Krakow, *majerz@agh.edu.pl

ZERO DEGREE CALORIMETER FAST SIMULATION

The Zero Degree Calorimeter (ZDC) of the ALICE experiment at CERN is traditionally simulated using a **computationally expensive** Monte Carlo approach. With their recent advances, Generative Neural Networks offer a promising alternative to perform this task. Our work focuses on the application of **Normalizing Flows** (NFs) in the simulation of the ZDC responses, with a particular focus on **transfer learning** and **fine-tuning** techniques.

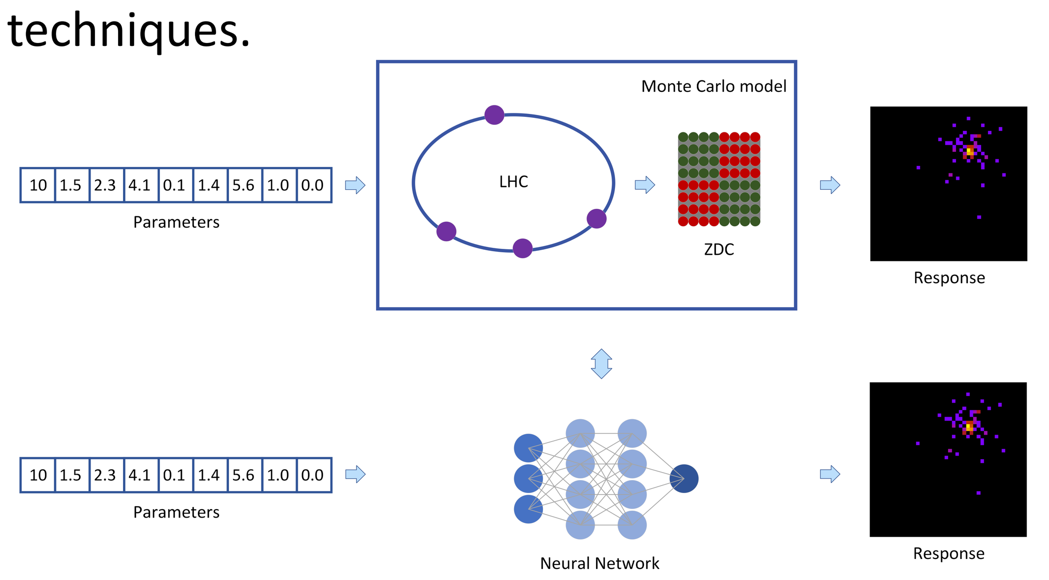

Fig. 1. Experiment overview.

DATASET CHALLENGES

Dataset structure

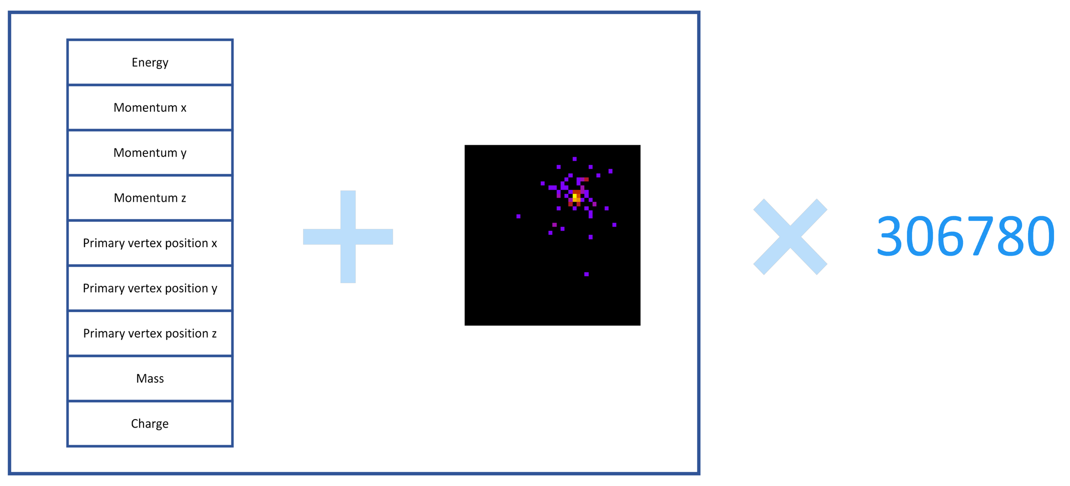

Fig. 2. Dataset structure - 306780 pairs: primary particle features + detector response.

Data imbalance

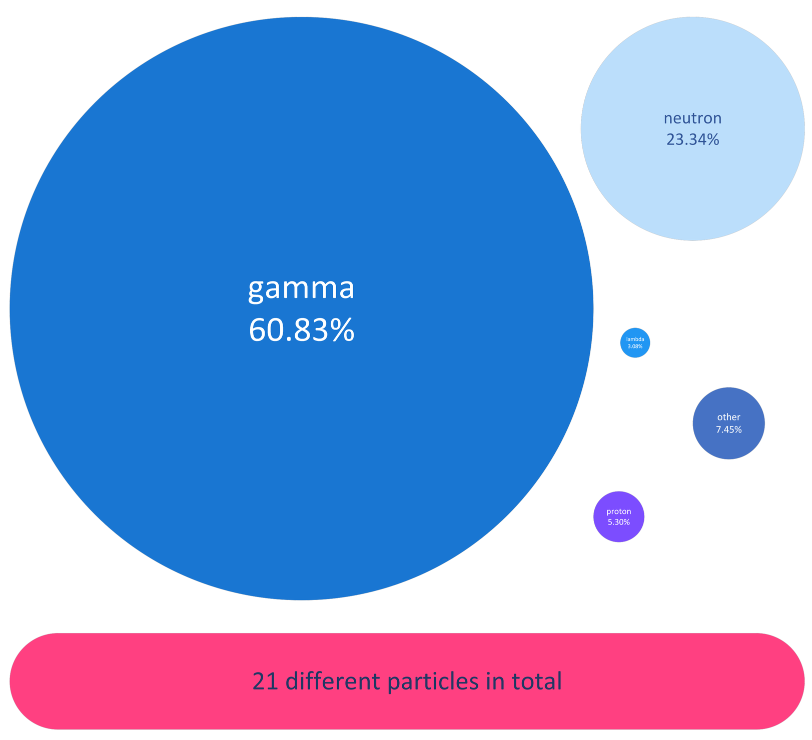

Fig. 3. Dataset structure. Over 90% of the dataset is covered by four types of particles, with 21 different particles in total.

Detector responses diversity

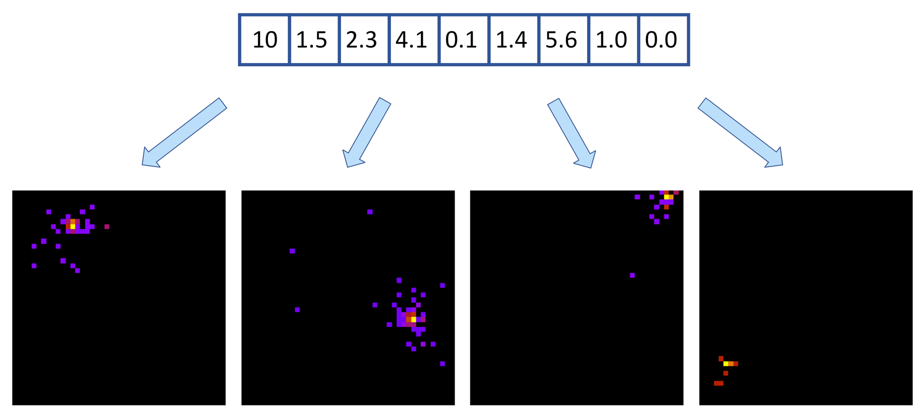

Fig. 4. The same input triggers different detector responses.

MODEL SELECTION

TWO-STAGE MODELLING

1. Bayesian Neural Network (BNN) for predicting the **total number of photons** in the detector response.
2. NF for **shower shape** modelling.

WASSERSTEIN METRIC FOR ASSESSING MODEL QUALITY

Applying checkerboard masks to detector responses and **comparing distributions** of such generated numbers of photons between original and generated images.

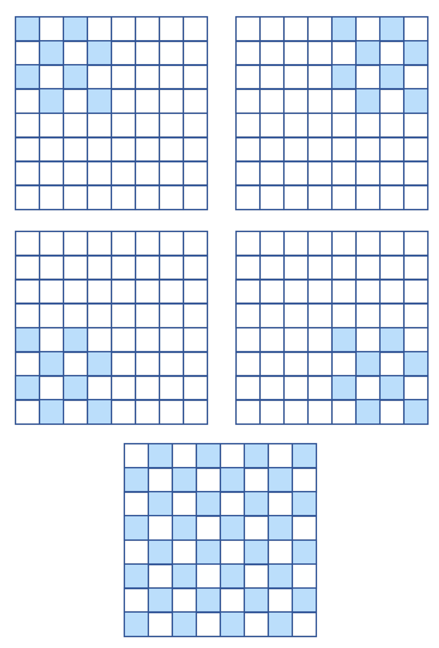

Fig. 5. Checkerboard masks applied to detector responses.

TRANSFER LEARNING + FINE-TUNING

The model trained using the whole dataset is then **fine-tuned** for specific particles using the *gradual unfreezing* technique. We propose two approaches:
1. Starting the unfreezing **from layers close to the base Gaussian distribution** (TL_*nepochs*_GD).
2. Starting the unfreezing **from layers close to the data distribution** (TL_*nepochs*_DG).

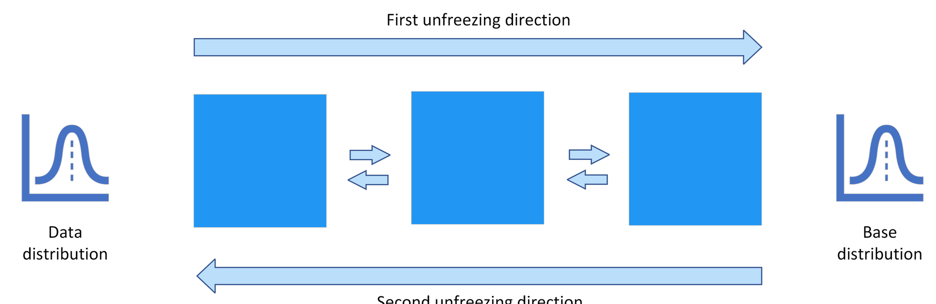

Fig. 6. Fine-tuning schema for NFs.

RESULTS IMPROVEMENT WITH TRANSFER LEARNING AND FINE-TUNING

The baseline NF model (referred to as *Baseline*_100) achieved a Wasserstein score of 2.34 ± 0.02 and was later fine-tuned for Gamma, Neutron, Lambda, K(S)0, and Sigma+ particles, separately. The model was trained for 100 epochs and then fine-tuned for another 100 (*TL_100_GD* and *TL_100_DG*). The performance was also compared with a baseline model trained for 200 epochs (*Baseline*_200) and a model trained for the specified particle from scratch (*Individual*), also for 200 epochs.

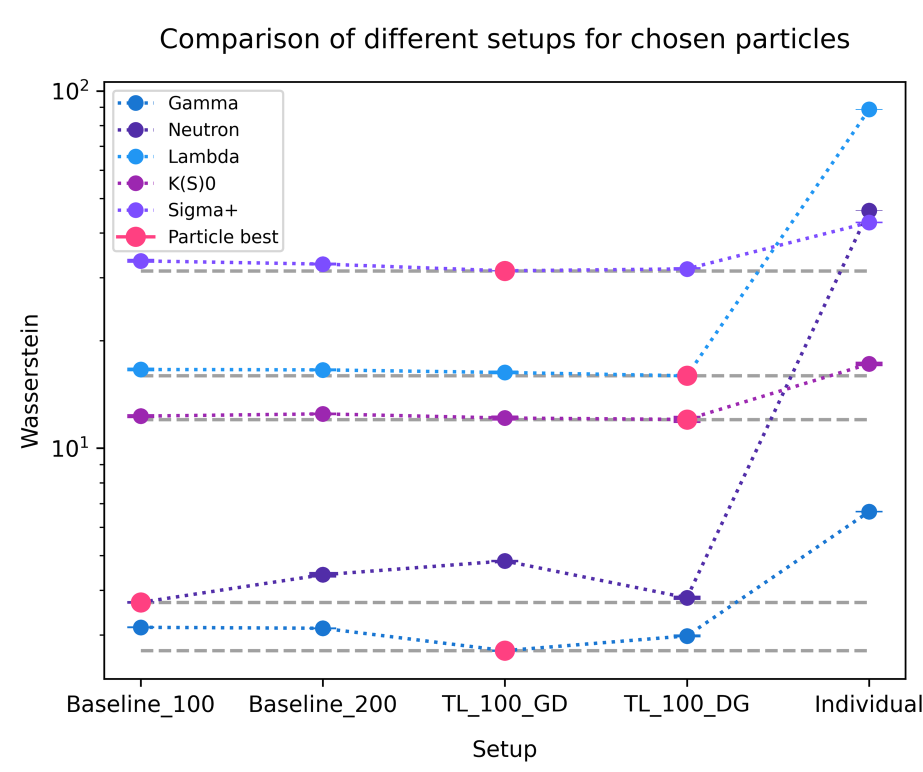

Fig. 7. Comparison of model performances for different particles.

The **transfer learning + fine-tuning setup outperformed** other models in 4 out of 5 cases. For neutrons, the baseline model was already fit so well to the data, that no improvements were possible.

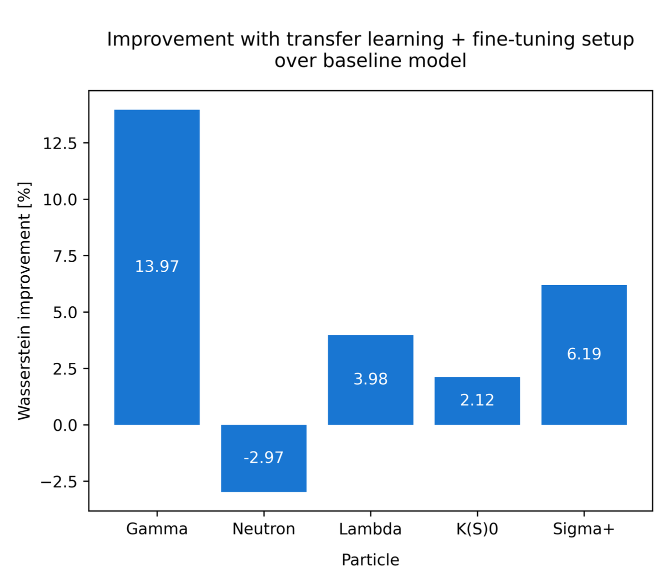

Fig. 8. The best transfer learning + fine-tuning setup improvements for different particles.

UNDERSTANDING THE MODEL

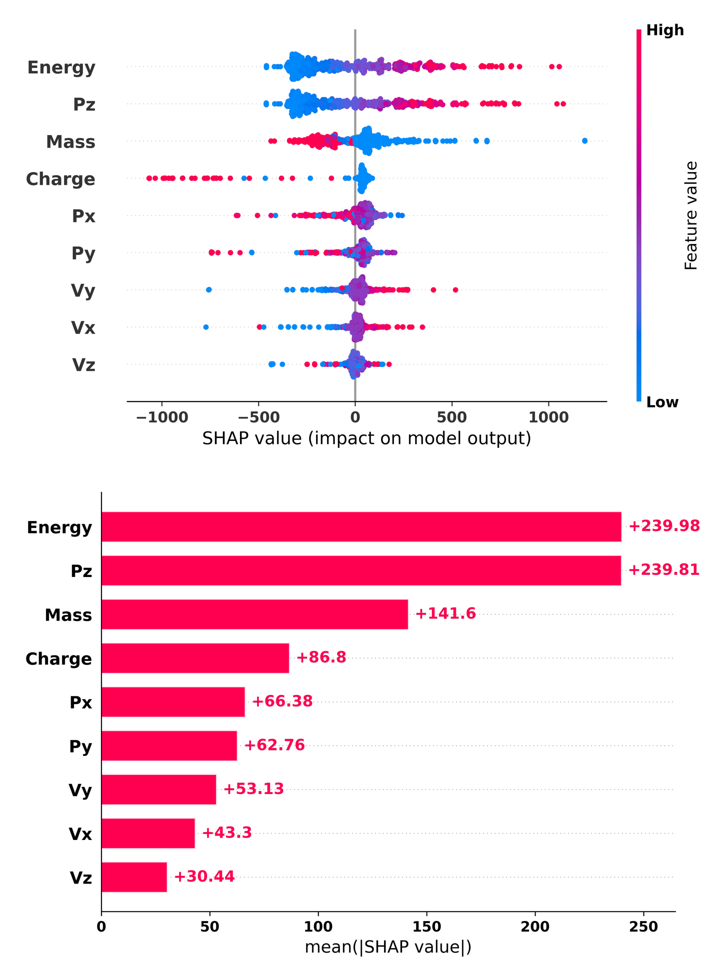

Fig. 9. Shap values (top) and feature importances with their impact on model output (bottom) calculated for the BNN model for predicting **the numbers of photons** in the detector output. As physically expected, **Energy** and **momentum in the *z* direction** (Pz) are the most important features, and their bigger values correspond to bigger numbers of photons.

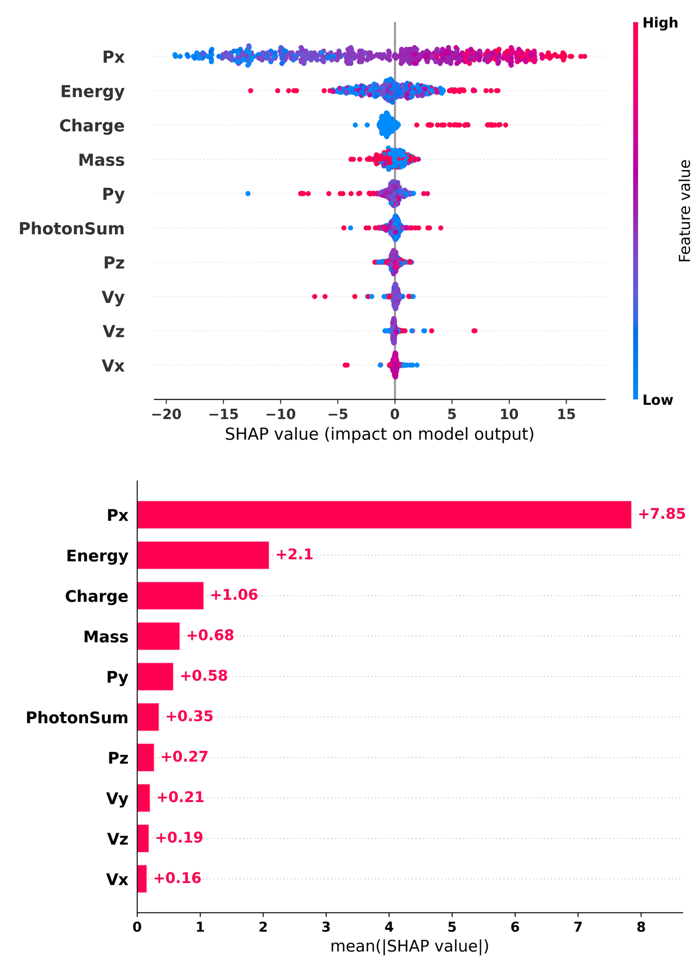

Fig. 10. Shap values (top) and feature importances with their impact on model output (bottom) calculated for a BNN surrogate model for predicting the coordinates of the **center of the shower** in the detector output trained on the outputs of the baseline NF model. The figure shows the analysis for the ***x* coordinate**. As physically expected, **momentum in the *x* direction** (Px), **Energy**, and **Charge** are the most important features.

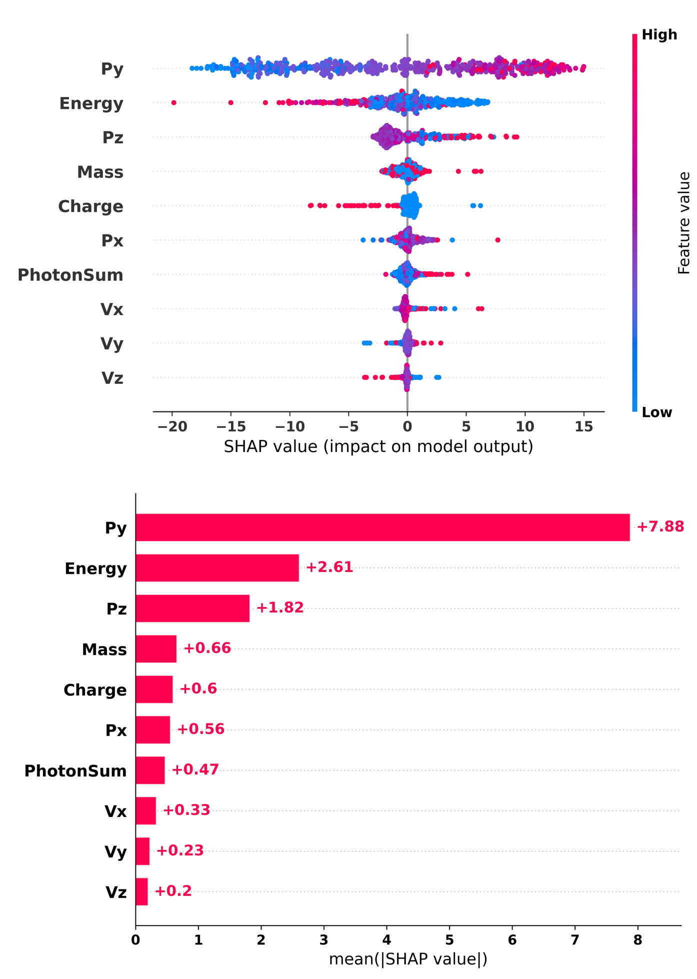

Fig. 11. Shap values (top) and feature importances with their impact on model output (bottom) calculated for a BNN surrogate model for predicting the coordinates of the **center of the shower** in the detector output trained on the outputs of the baseline NF model. The figure shows the analysis for the ***y* coordinate**. As physically expected, **momentum in the *y* direction** (Py) and **Energy** are the most important features.

CONCLUSIONS

1. NFs offer a **strong-performing alternative** to the Monte Carlo approach on the task of the ZDC shower simulation.
2. NFs can **benefit from the transfer learning + fine-tuning schema**, and the layer unfreezing can be executed either starting from the layers close to the data or to the base ditributions.
3. The **reasoning** behind NFs trained on the task of the ZDC shower simulation **follows physical intuition**.

This work is co-financed and in part supported by the Ministry of Science and Higher Education (Agreement Nr 2023/WK/07) by the program entitled "PMW" and by the Ministry funds assigned to AGH University in Krakow. We gratefully acknowledge Polish high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational grants no. PLG/2023/016410 and PLG/2024/017264.