



Contribution ID: 182

Type: Poster + Flashtalk

## ExpertSim: Fast Particle Detector Simulation Using Mixture of Experts for Generative Models

High-energy physics experiments at the Large Hadron Collider (LHC) at CERN rely on simulations to model particle interactions and understand experimental data. These simulations, crucial for reconstructing collision events, are traditionally performed using Monte Carlo-based methods, which are highly computationally demanding. With hundreds of thousands of CPU cores dedicated to these tasks annually, the need for more efficient, high-fidelity alternatives is pressing.

Recently, generative deep learning models have emerged as a promising solution, offering faster synthetic data generation. However, applying standard generative architectures to calorimeter simulations remains challenging due to the complex and multi-modal nature of detector responses. Different particles produce diverse energy deposition patterns, making it difficult for a single model to capture the full variability without sacrificing accuracy.

In our work, we focus on simulating the Zero Degree Calorimeter (ZDC) in the ALICE experiment at CERN, which plays a critical role in measuring the energy of non-interacting nucleons in heavy-ion collisions. The ZDC responses fall into multiple distinct categories, corresponding to different types of particle interactions and energy depositions. Attempting to model all of these variations with a single generative network leads to trade-offs between fidelity and efficiency, as a single model lacks the capacity to fully represent the complex, structured nature of the data distribution.

To overcome these limitations, we propose ExpertSim, a novel Mixture-of-Generative-Experts (MoE) architecture designed specifically for ZDC simulation. Instead of relying on a single generative model, ExpertSim employs multiple specialized experts, each trained to simulate a specific subset of the data distribution. A router network dynamically assigns incoming particle events to the most appropriate expert based on their physical properties, ensuring that each expert is specialized in a particular response pattern. This division of tasks improves the overall fidelity of the simulation while maintaining computational efficiency.

A key component of our approach is the Expert Differentiation Loss, which ensures that each expert specializes in a distinct subset of the data rather than redundantly modeling overlapping distributions. This loss function penalizes similarity between experts by encouraging diversity in their generated outputs. By explicitly maximizing the difference between experts' mean energy intensities, Expert Differentiation Loss forces the model to partition the data more effectively, leading to clearer specialization and higher-quality generated responses.

In addition to the expert-based architecture, ExpertSim incorporates diversity regularization, intensity constraints, and an auxiliary regressor, which enhance the accuracy of the generated responses. The diversity regularization mitigates GAN mode collapse by encouraging variation in generated samples, while the intensity regularization ensures that the total deposited energy aligns with real detector signals. Furthermore, the auxiliary regressor aids in learning spatial correlations in energy deposition, improving the geometric consistency of the simulations.

Through extensive experiments, we demonstrate that ExpertSim outperforms existing generative models on the task of simulating the response of the ZDC. Our model achieves a 15% reduction in Wasserstein distance between the distribution of real and generated data compared to prior methods while preserving the significant computational speedup that generative models offer over Monte Carlo-based approaches.

### AI keywords

generative models, fast simulation, mixture of experts, generative adversarial networks

**Primary authors:** Mr BĘDKOWSKI, Patryk (Warsaw University of Technology); DUBIŃSKI, Jan (Warsaw University of Technology; IDEAS NCBR); Mr SZATKOWSKI, Filip (Warsaw University of Technology; IDEAS NCBR); Prof. TRZCIŃSKI, Tomasz (Warsaw University of Technology; Tooploox); Prof. ROKITA, Przemysław (Warsaw University of Technology); Dr DEJA, Kamil (Warsaw University of Technology; IDEAS NCBR)

**Presenter:** Mr BĘDKOWSKI, Patryk (Warsaw University of Technology)

**Track Classification:** Simulations & Generative Models