Contribution ID: **63**
Type: **Oral**

CaloDiT: Diffusion with transformers for fast shower simulation

Wednesday 13 March 2024 17:10 (20 minutes)

Recently, transformers have proven to be a generalised architecture for various data modalities, i.e., ranging from text (BERT, GPT3), time series (PatchTST) to images (ViT) and even a combination of them (Dall-E 2, OpenAI Whisper). Additionally, when given enough data, transformers can learn better representations than other deep learning models thanks to the absence of inductive bias, better modelling of long-range dependencies, and interpolation and extrapolation capabilities. On the other hand, diffusion models are the state-of-the-art approach for image generation, which still use conventional U-net models for generation, mostly consisting of convolution layers making little use of the advantages of transformers. While these models show good generation performance it lacks the generalisation capabilities obtained from the transformer model. Standard diffusion models with an Unet architecture have already proven to be able to generate calorimeter showers, while transformer-based models, like those based on a VQ-VAE architecture, also show promising results. A combination of a diffusion model with a transformer architecture should bridge the quality of the generation sample obtained from diffusion with the generalisation capabilities of the transformer architecture. In this paper, we propose CaloDiT, to model our problem as a diffusion process with transformer blocks. Furthermore, we show the ability of the model to generalise to different calorimeter geometries, bringing us closer to a foundation model for calorimeter shower generation.

Significance

This contribution presents a novel transformer-based machine learning model for general fast shower simulation, fitting perfectly the focus theme of ACAT 2024 - the foundation models.

References

Experiment context, if any

Primary authors: DA COSTA CARDOSO, Renato Paulo (CERN); RAIKWAR, Piyush (CERN); ZABOROWSKA, Anna (CERN); SALAMANI, Dalila (CERN); JARUSKOVA, Kristina (CERN); Dr VALLECORSA, Sofia (CERN); YEO, Kyongmin (IBM Research); EKAMBARAM, Vijay (IBM Research); NGUYEN, Nam (IBM Research); KALAGNANAM, Jayant (IBM Research); SRIVATSA, Mudhakar (IBM Research)

Presenter: RAIKWAR, Piyush (CERN)

Session Classification: Track 1: Computing Technology for Physics Research