Contribution ID: **109** Type: **not specified**

DeGeSim: Conditional Denoising Diffusion Probabilistic Models as Multi-Dimensional Density Mappers for Continuous and Discrete State Spaces

Friday 10 November 2023 09:30 (15 minutes)

As the performance of the Large Hadron Collider (LHC) continues to improve in terms of energy reach and instantaneously luminosity, ATLAS faces an increasingly challenging environment. High energy proton-proton (pp) interactions, known as *hard scatters*, are produced in contrast to low energy inelastic proton-proton collisions referred to as *pile-up*. From the perspective of data analyses, hard scatter events are processes of interest that probe the quantum scale, whilst pile-up is conceptually no different from noise and so is often removed from reconstructed objects (e.g. jets) during a measurement. As the High Luminosity LHC (HL-LHC) era approaches, current simulations of pile-up are inadequate at addressing the environment presented by an estimated 200 pile-up interactions per bunch crossing. This poses a significant problem for precision measurements at the HL-LHC.

In order to address this, *Deep Generative models for fast and precise physics Simulations* (DeGeSim) endeavours to utilise deep generative image synthesis techniques to emulate calorimeter images of soft quantum chromodynamic (QCD) pile-up data collected by ATLAS at the LHC. The project ultimately uses Denoising Diffusion Probabilistic Models (DDPMs) to synthesize calorimeter images based on instances of real (observed) pile-up data collected by the ATLAS detector. However, instead of seeding the generation from gaussian noise, MC simulated images of pile-up are used. This is achieved by harnessing the intrinsic markov chain process of diffusion models to map MC images to data images, allowing for semantic based image alteration. The intention is to replace MC generated calorimeter images with *data informed edited* versions of the image within the ATLAS simulation chain, thereby yielding images that better resemble data.

The work that will be presented is a sub-component of the aforementioned model, which addresses a key problem in probability density mapping techniques, such as density ratio estimation, of disjoint probability density functions in which the state spaces lack support. Specifically, we demonstrate that a conditional denoising diffusion probabilistic model (DDPM) augmented with self-conditioning can be used to map between otherwise disjoint pdfs. This is achieved by utilising the conditional behaviour of DDPMs to solve a pseudo-inverse problem of generating a pdf with parameter set $\vec{\theta}'$ from an initial data point obtained by sampling the sample space of a different disjoint pdf with parameter set $\vec{\theta}$. This is achieved via the use of density ratio estimators and classifier free guidance. In addition, the proposed model utilises analog bit representations of discrete state spaces to solve the instabilities introduced when dealing with datasets that occupy both continuous and discrete state spaces, as is common in High Energy Particle physics problems.

Author: JIGGINS, Stephen (Deutsches Elektronen-Synchrotron (DE))

Co-author: KATZY, Judith (Deutsches Elektronen-Synchrotron (DE))

Presenters: KATZY, Judith (Deutsches Elektronen-Synchrotron (DE)); JIGGINS, Stephen (Deutsches Elektronen-Synchrotron (DE))

Session Classification: Results, Observables & Techniques