Artificial Intelligence for the LHCb Simulation

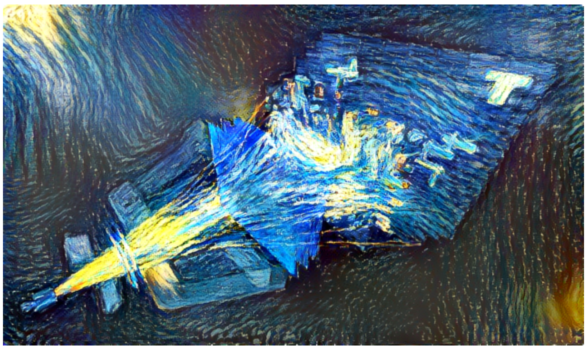
Advisor:	Lucio Anderlini	< lucio.anderlini1@fi.infn.it >
Co-Advisors:	Michele Veltri	< michele.veltri@uniurb.it >
	Matteo Barbetti	< matteo.barbetti@fi.infn.it >

In order to claim or disprove New Physics effects in the collected data, the LHCb experiment compares the acquired datasets with Simulated data. Simulations employ Monte Carlo techniques to predict the outcome of the proton-proton collisions producing heavy hadrons, their decay mechanisms and the interaction of the most long-lived particles with the detector material. The energy deposits of these particles in the detector are then converted into electronic signals and processed with the very same algorithms as used to reconstruct acquired data.

Such a process is extremely computing intensive, requiring a large distributed computing infrastructure on which the Collaboration invests several million euros every year. A large fraction of the computing resources is spent to simulate the detector response to hadrons and muons produced in various heavy hadron decays. Different decaying hadrons and different decay modes may result into significantly different distributions of the kinematic and geometrical variables describing the particles traversing the detector, but once those are set, the detector response to their traversal will be almost completely independent of their parent particle. Hence, reusing the simulation of the interaction of particles with the detector for multiple decay modes and decaying hadrons may allow to save a huge fraction of the computing resources needed to explore the Physics of proton-proton collisions.

Reusing the simulation, however, is not a simple task. The number of geometrical and kinematical variables describing the particles produced in a heavy hadron decay is too large to build a library of reconstructed hadrons.

At the Florence unit of INFN, we are working on a different approach: an Artificial Intelligence observes the fully simulated data to build a statistical model able to predict the outcome of the simulation for a given set of kinematic and geometrical variables. The trained agent is then able to replace the simulation procedure with a computational costs reduced by at least two orders of magnitude.



Graphical representation of LHCb event passed through a Style Transfer algorithm that returns it as the Van Gogh's "Starry Night" (stolen from the Andrey Ustyuzhanin's talk).

The Ph.D. candidates will take part in this challenging research involving Physics, High Throughput HEP computing techniques and cutting-edge Smart Computing technologies, including hardware accelerators and Cloud technologies. Depending on the ambitions and attitude of the candidate, the activity can span from the increase of the Technology Readiness Level of the AI-based Simulations to the development of advanced algorithms to model effects due to particle-to-particle correlations, such as hadronic jets or shared calorimetric clusters.

Further reading

- V. Chekalina, *et al.*, *Generative Models for Fast Calorimeter Simulation: the LHCb case*, EPJ Web Conf. **214** (2019) 02034
- M. Rama and G. Vitali, *Calorimeter fast simulation based on hit libraries LHCb Gauss framework*, EPJ Web Conf. **214** (2019) 02040
- A. Maevskiy *et al.*, *Fast Data-Driven Simulation of Cherenkov Detectors Using Generative Adversarial Networks*, `arXiv:1905.11825`
- LHCb Collaboration, *Performance of the Lamarr Prototype: the ultra-fast simulation option integrated in the LHCb simulation framework*, LHCB-FIGURE-2019-017
- M. Barbetti, *Techniques for parametric simulation with deep neural networks and implementation for the LHCb experiment at CERN and its future upgrades*, Master's thesis, University of Florence, 2020

- L. Anderlini, *Machine Learning for the LHCb Simulation*, arXiv:2110.07925
- F. Sergeev *et al.*, *Fast simulation of the LHCb electromagnetic calorimeter response using VAEs and GANs*, J. Phys.: Conf. Ser. **1740** (2021) 012028
- F. Ratnikov and A. Rogachev, *Fast simulation of the electromagnetic calorimeter response using Self-Attention Generative Adversarial Networks*, EPJ Web Conf. **251** (2021) 03043
- G. Graziani, *et al.*, *A Neural-Network-defined Gaussian Mixture Model for particle identification applied to the LHCb fixed-target programme*, JINST **17** (2022) P02018
- L. Anderlini, *et al.*, *Towards Reliable Neural Generative Modeling of Detectors*, arXiv:2204.09947
- LHCb Collaboration, *Machine-Learnt parametrizations for the Ultra-Fast Simulation of the LHCb detector*, LHCB-FIGURE-2022-004
- LHCb Collaboration, *Validation of the Lamarr framework with $\Lambda_b^0 \to \Lambda_c^+ \mu^- X$ decays*, LHCB-FIGURE-2022-014

About the group

Lucio Anderlini is Researcher at INFN - Firenze. He has been coordinating the Machine Larning group of the Institute since 2019. Currently, he is coordinator of the *Detailed & Fast Simulation* working group of the LHCb and lead developer of the *Lamarr* framework, designed to inject machine-learning models in the distributed-computing architecture of the LHCb experiment.

Michele Veltri is Professor of Physics at the University of Urbino. Prof. Veltri is a leading expert on Geant4, a simulation framework used to predict the radiation-matter interactions in High Energy Physics and coordinates the activities related to Geant4 in the LHCb Simulation Project.

Matteo Barbetti is Ph.D Student in Smart Computing with several years of experience on Generative Models and in particular on Generative Adversarial Networks. His Master Thesis on the application of Generative Models to the LHCb Simulation was recently awarded with the Giulia Vita Finzi prize by the National INFN Computing & Network committee. Matteo develops and maintains TFGans, a repository of implementations of generative models in TensorFlow designed to be applicable to problems in High Energy Physics.