Computing and ML for SHiP

Oliver Lantwin *on behalf of the SHiP Collaboration*

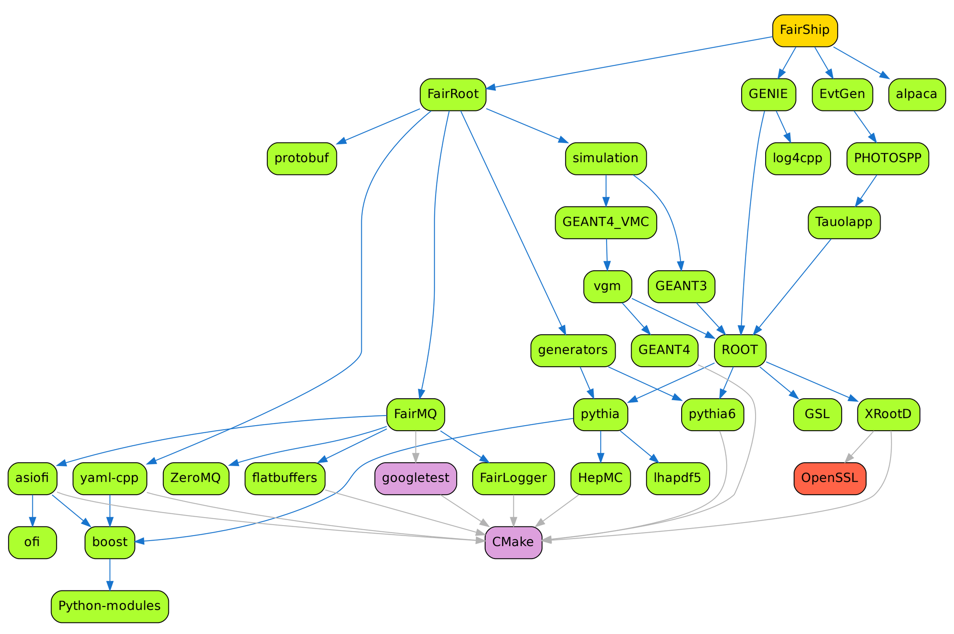
oliver.lantwin@cern.ch

The SHiP experiment

- The Search for Hidden Particles (SHiP) experiment has been approved in March 2024, with commissioning expected in 2032.
- Designed to be the world-leading experiment for the search for feebly interacting particles (FIPs), such as dark photons (DPs), dark scalars (DSs), Heavy Neutral Leptons (HNLs) and axion-like particles (ALPs), and other beyond Standard Model particles, using the High-Intensity ECN3 facility at CERN's SPS.
- **High intensity** and **zero background** result in a unique and interesting mix of computing constraints and requirements

The SHiP software framework

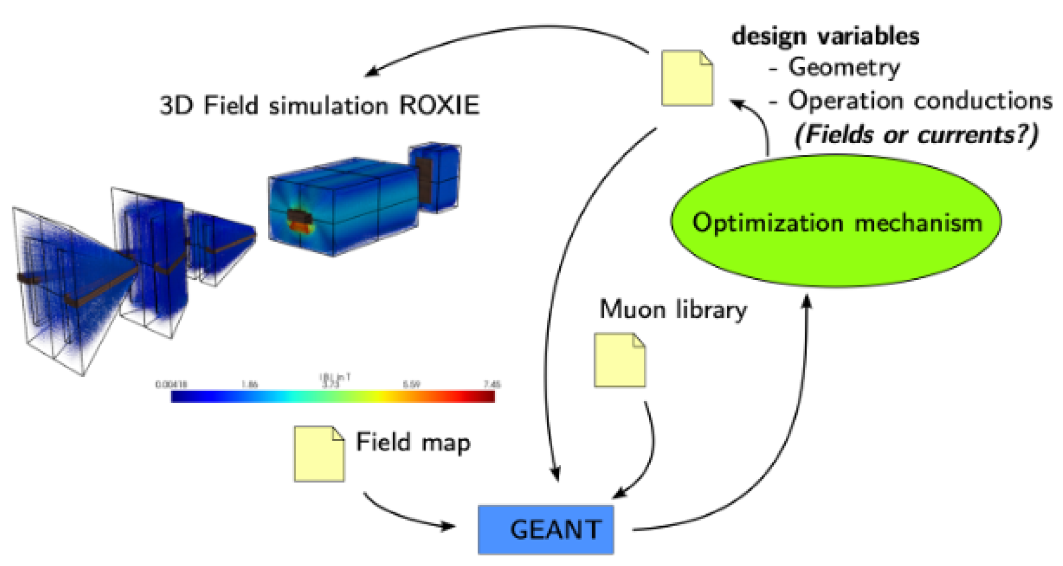
- The SHiP software framework, `FairShip`, is **based on the `FairRoot` framework**, making use of `PyROOT` to present a pythonic interface to users.
- Dependencies managed using ALICE's `alibuild` with prebuilt packages distributed via CVMFS



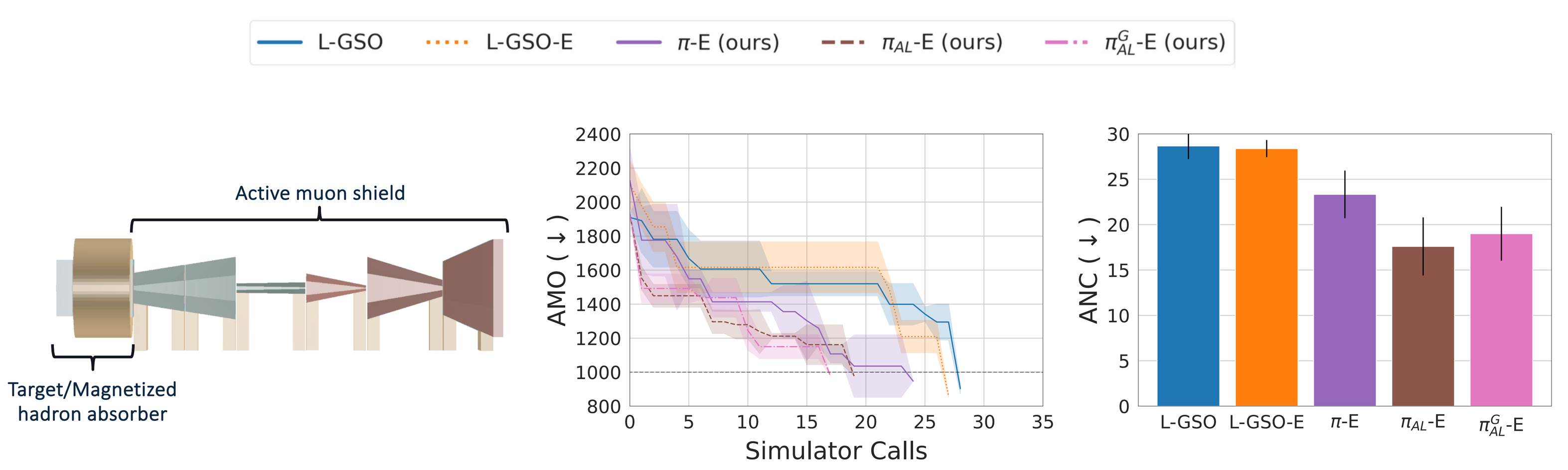
- Framework already used in the real in world dedicated experiments (**SHiP muon flux** and **SHiP charm-cross section** measurements), as the **basis of the SND@LHC software framework** (taking data at LHC since 2022) and for its planned AdvSND upgrade

Optimisation of the muon shield

- The **muon shield is critical to reaching SHiP's physics goals**
- Optimisation of muon shield in the past performed using approximate fields and Bayesian optimisation
  - 42 parameters
  - **Noisy, black-box optimisation**
  - Full simulation using `Geant4` for each configuration
  - 100 configurations tested in parallel
- Moving Bayesian optimisation to GPUs and redesign of optimisation workflow now allow us to test $\mathcal{O}100\,000$ muon shield configurations per day
- **On-the-fly calculation of field maps** becomes plausible by integrating with CERN's ROXIE package



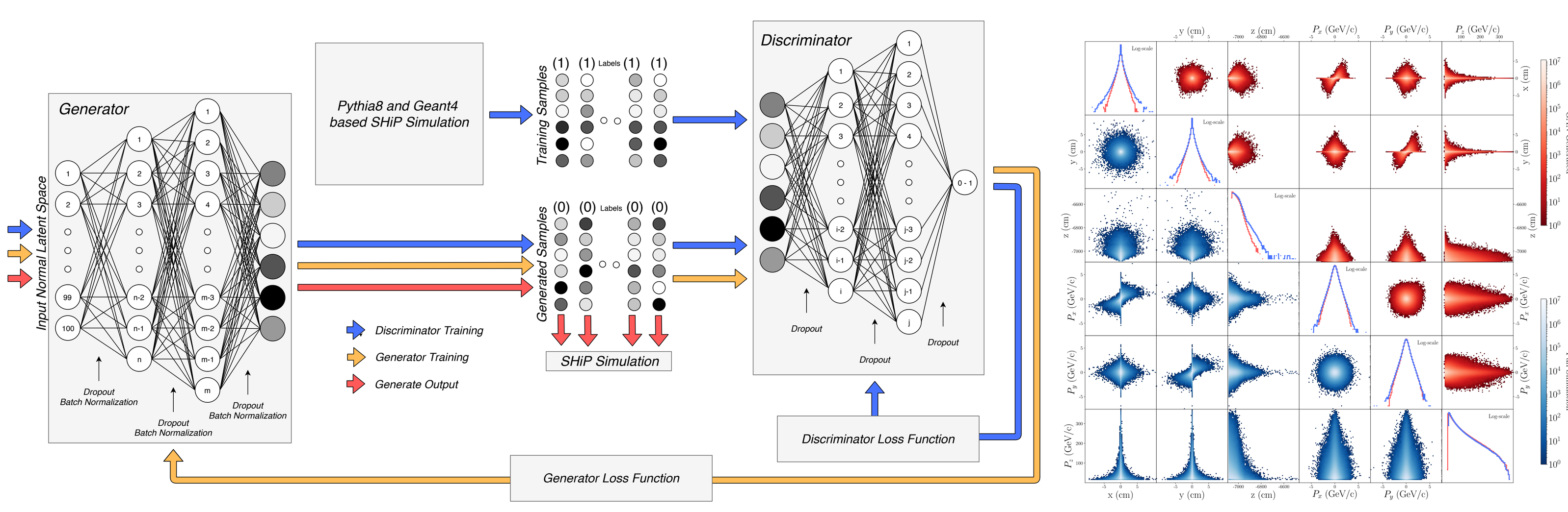
- **Local Generative Surrogates** (L-GSO) allow us to simulate fewer configurations with full simulation, drastically speeding up optimisation
- **Reinforcement learning** shows promise to further improve on Bayesian Optimisation and L-GSO



Comparison of L-GSO and reinforcement learning ($\pi - E$) approaches, from

Rejecting known knowns: Background simulation

- Full simulation available for only a **small fraction of a second**, but need to be confident about **background across 15 years**
- Bias simulations and force specific interactions for neutrino and muon DIS allow us to reach expected statistics >15 years
- Use **Generative Adversarial Networks** (GANs) to improve muon statistics



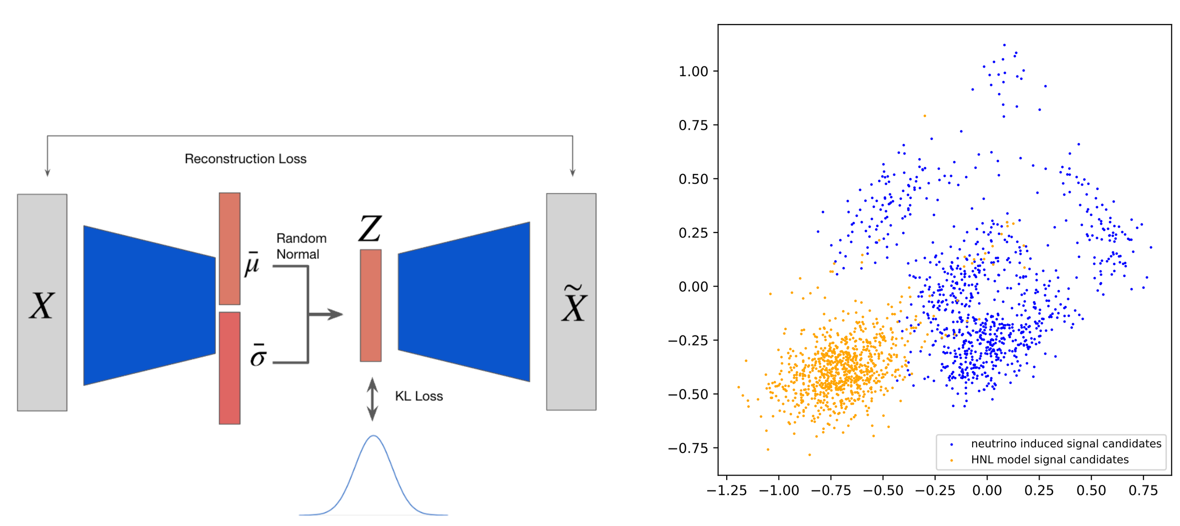
- Work ongoing to merge events and simulate non-uniform spill time structure
- We are developing a **background-tagging algorithm using a Graph Neural Network** (GNN) to tag parts of events as background while preserving high signal efficiency

Looking for known unknowns: `SensCalc` & `EventCalc`

- **Direct interface with `EventCalc` generator** to allow studying a variety of FIP channels without having to implement each separately
- `EventCalc` is the evolution of `SensCalc`, an experiment-independent **semi-analytical** sensitivity calculation tool, validated against many experiments in-house sensitivity calculations

Looking for unknown unknowns: Anomaly detection

- In order to be sensitive to **unexpected signatures**, we need to develop techniques to distinguish them from the background with minimal assumptions
- After **successful proof-of-concept**, we are investigating using Variational Autoencoders (VAEs) to develop a selection for new physics signals using only our knowledge of the detector backgrounds, **complementary to targeted selections**



Left: Schematic of the VAE architecture; Right: Separation of signal (HNL) and background (neutrino) in the latent space of the VAE.

Towards data taking

2025 **Major technology decisions informed by simulation** need to be taken
  - Electronic or emulsion-based Scattering and neutrino detector (SND)
  - Partially super-conducting or fully warm muon shield

2027 TDRs (including online and offline computing) to be submitted

2032 **Commissioning with beam** and then data taking for 15 years