Universität Zürich UZH · Imperial College London · University of BRISTOL · UNIVERSITY OF CAMBRIDGE

Guillermo Hijano, Davide Lancierini, Alex Marshall, Andrea Mauri, Patrick Owen, Mitesh Patel, Konstantinos Petridis, Shah Rukh Qasim, Nicola Serra, William Sutcliffe, Hanae Tilquin

Motivation

- Physics analyses can involve a large amount of **background candidates** (for instance, heavy flavor physics)
- **No systematic method** to automatically determine the relevant backgrounds
  - Process relies on physicist's intuition
  - Risk of human error
- Our tool would serve to:
  - automatize and **accelerate the workflow**
 - **mitigate human errors**

Approach
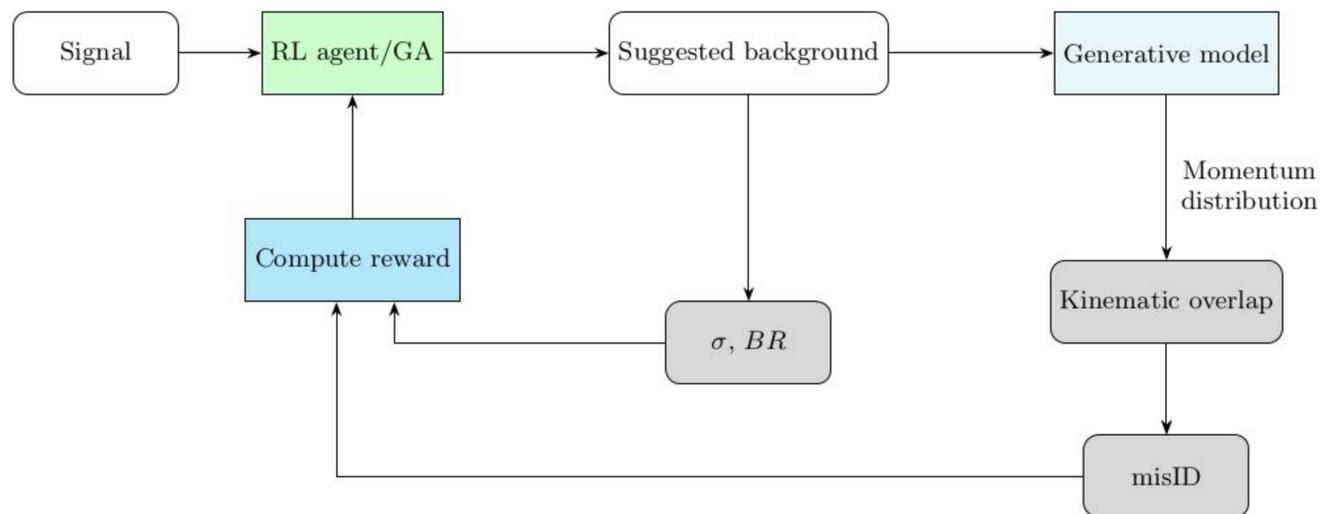
$$Reward = \frac{\sigma(pp \to M^B)}{\sigma(pp \to M^S)} \frac{BR_b}{BR_s} \cdot \mathcal{M} \cdot \mathcal{K}$$

Reinforcement Learning for background determination in particle physics

Background finder

- Approach combines **Genetic Algorithms (GAs)**, **Reinforcement Learning (RL)** and **transformers**
  - GAs perform an **efficient exploration**
    - Identify **successful trajectories** in the environment
  - RL agent uses these solutions as high quality training data, and **learns** to predict relevant backgrounds for new signals
    - AlphaZero algorithm to handle large trajectory space and purely terminal rewards
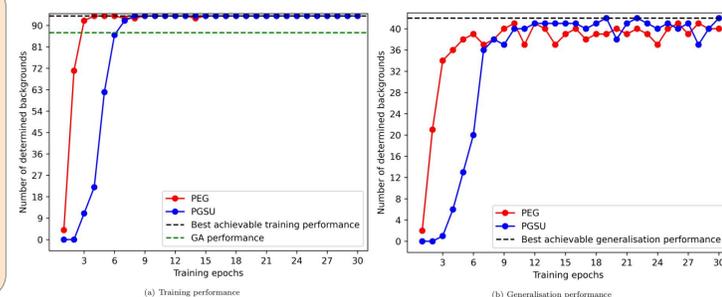 - Transformer architecture for agent to handle **token sequences** representing decays
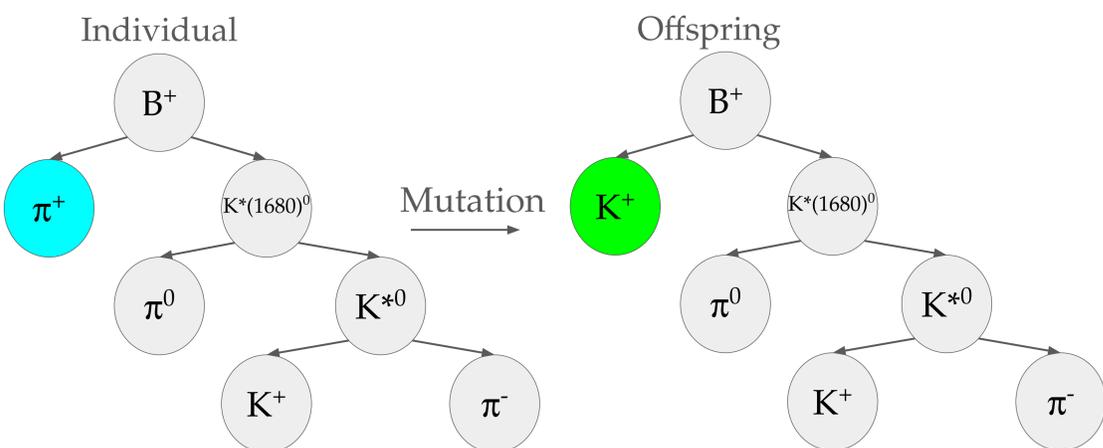
- **State** is represented by a sequence of tokens
  - Initial state encodes just the signal
- **Agent** predicts next token in each step, iteratively constructing the background
- We can deal with **intermediate resonances**, **misidentifications** and **partially reconstructed backgrounds** with a few tokens
- Simplified example:
  - Signal: $B^0 \to K^+ \pi^0 \pi^-$ K- and π- are misidentified
  - Background: $B^+ \to K^{*+}(K^+ \pi^0)(\pi^+)K^-$
    π+ is not detected

Results

- Trained the agent with **17 training signals** (94 backgrounds)
- Checked its **generalization** ability with **8 new signals** (42 backgrounds). Chosen to be similar to the training signals.
- **Performance**:
  - GA performance (using a population of 6,000 individuals and an evolution of 40 generations):
    - Found 87/94 training backgrounds
  - RL agent performance (measured building 100,000 sequences per signal):
    - 94/94 training backgrounds learnt by agent
    - 42/42 generalization backgrounds found by agent

Performance before fine-tuning:

Performance after fine-tuning: