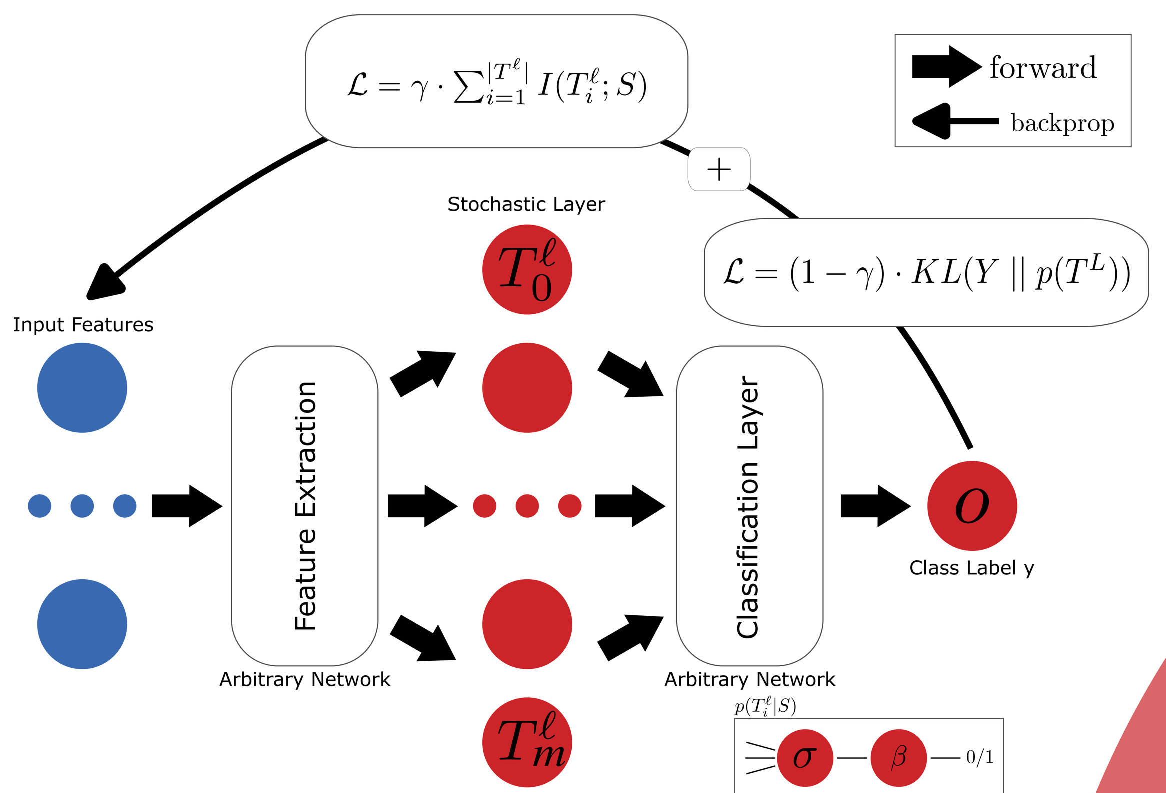
Adaptive Machine Learning on FPGAs: Bridging Simulated and Real-World Data in High-Energy Physics

Mattia Cerrato[1] & Marius Köppel[2]
[1]Institute of Computer Science Johannes Gutenberg-Universität Mainz
[2]Institute for Particle Physics and Astrophysics ETH Zürich

Abstract

- Use **binary stochastically activations** to treat neurons as random variables [1]



Introduction

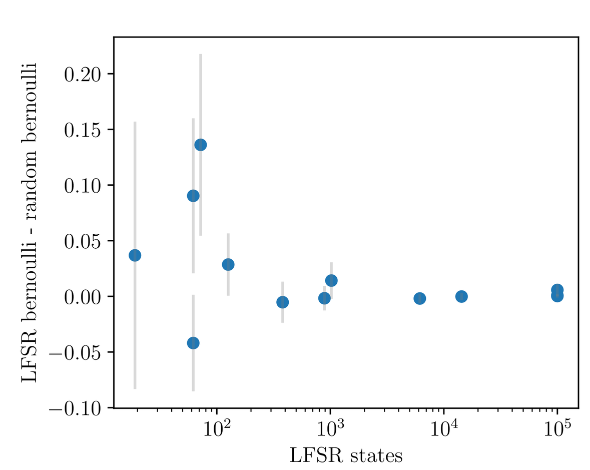
- Be invariant to **differences** in **simulation** and **data**
- **Adversarial classifier** need **2x** neurons to perform domain transformation task
- Computing of mutual information in **deterministic** neural networks is **hard** or even **impossible**

Method

- **Stochastic quantization** neurons by **Bernoulli**

$$H(T_i^l) = -(1-\theta_i^l) \cdot \log_2(1-\theta_i^l) - \theta_i^l \cdot \log_2(\theta_i^l)$$

- On **FPGAs** use linear-feedback-shift-register for Bernoulli distribution



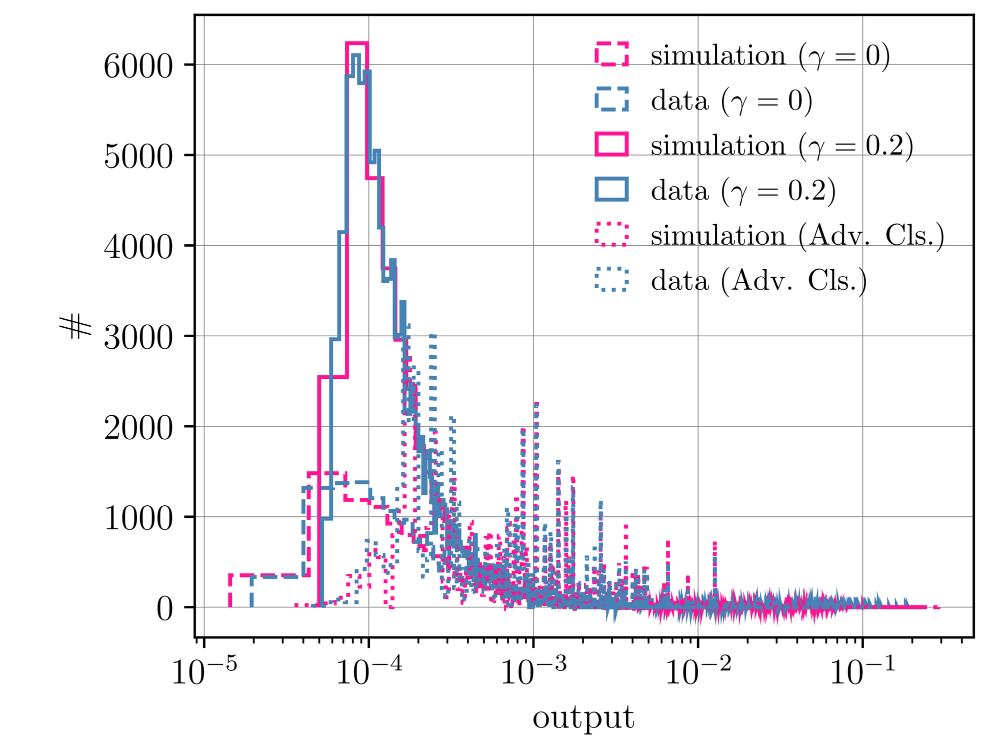
$$H(T_i^l) \qquad H(S)$$

Direct minimization of mutual information in a full precision network using binary stochastically-activated layers for obtaining invariant representations

$$\sum_{i=1}^{|T_l|} I(T_i^l; S) \geq I(T_l; S)$$

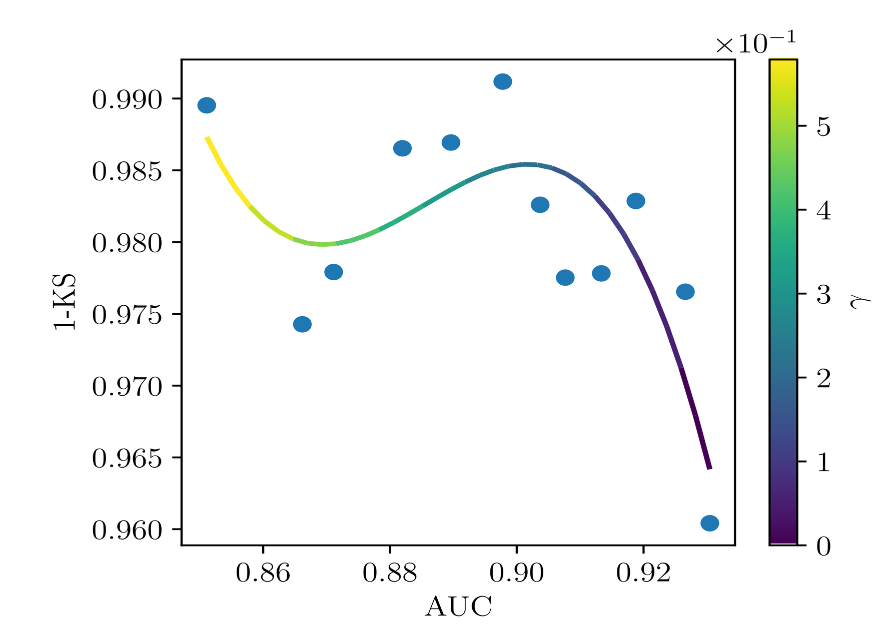
Experiments

- We analyze the method on **LHCb data** [2]



Results

- Method can **outperform adversarial classifier** & **normal NN**
- Method obtains **stable** invariant/accuracy **tradeoffs**



Outlook

- Use method to be **pile-up** invariant
- Revisit **information bottleneck** theory

References
[1] https://arxiv.org/abs/2208.02656
[2] https://www.kaggle.com/competitions/flavours-of-physics

ETH zürich JG|U