# Excursion: Active Learning and Gaussian processes to estimate level sets of black box functions

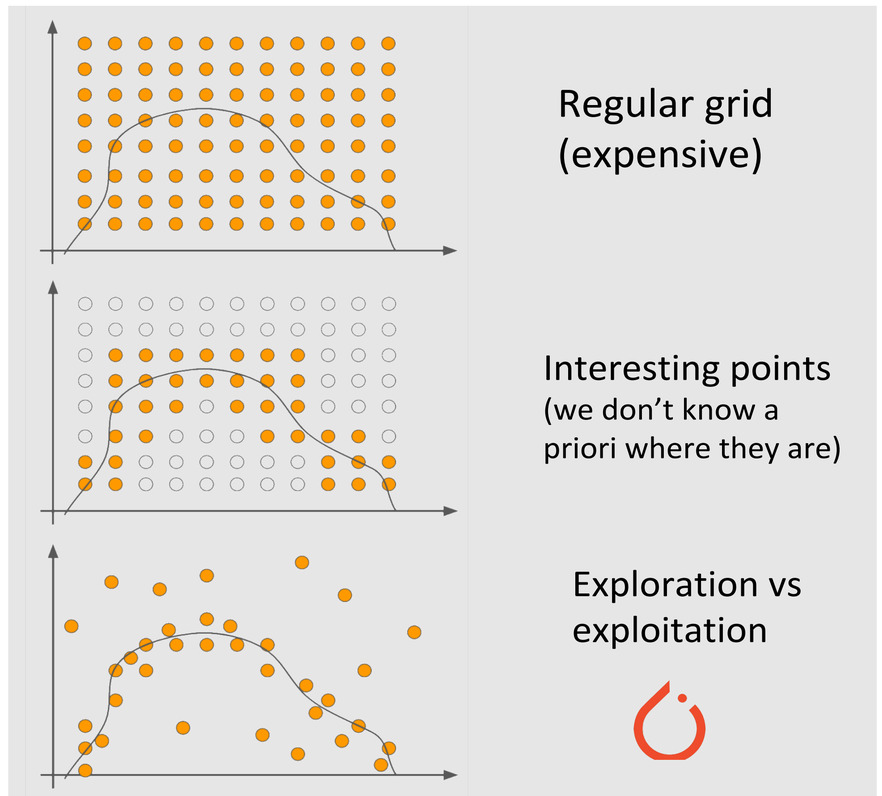
Kyle Cranmer [1] **Irina Espejo** [1] Lukas Heinrich [2] Giles Louppe [3]

Institutions: [1] New York University [2] CERN [3] University of Liege

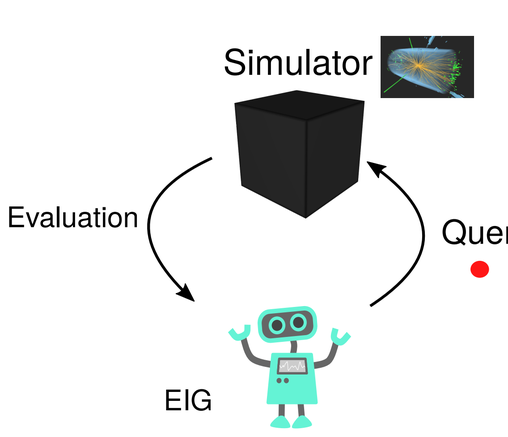
The SCAILFIN Project

## The goal and why

- Goal is to find *level sets of black-box functions* that are expensive to evaluate. Examples: test statistics from complex simulations.
- Evaluate the black box function at *interesting points only* instead of evaluating over a regular grid. We use a *Gaussian process* to: interpolate between samples and model uncertainty in the knowledge of the black box function.
  The *acquisition function* controls the exploration vs exploitation tradeoff. Select one that *minimizes global uncertainty* of the location of the excursion set.
- Future: efforts will focus on *scalability* wrt the dimensionality of the function domain space. Example, likelihood ratio as function of mass, charge, spin,...

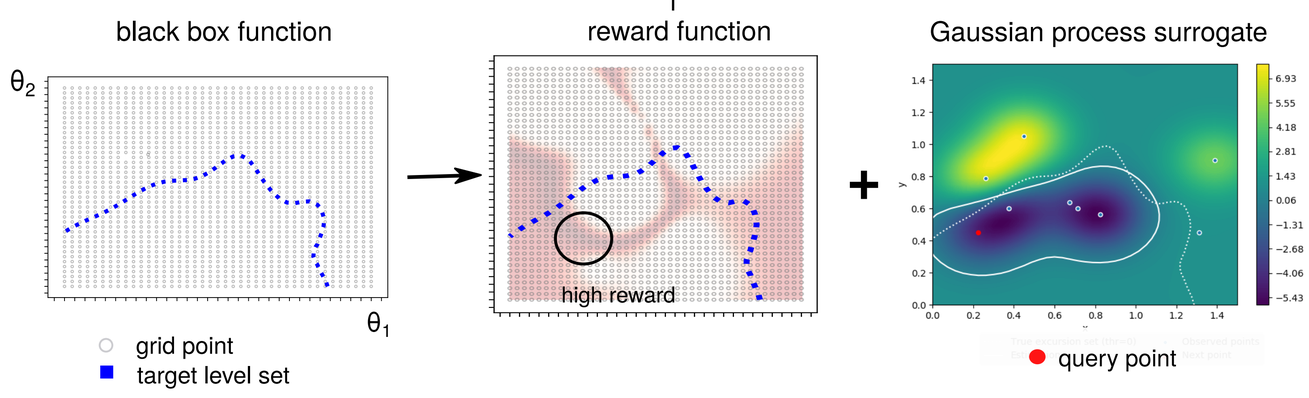



## Method

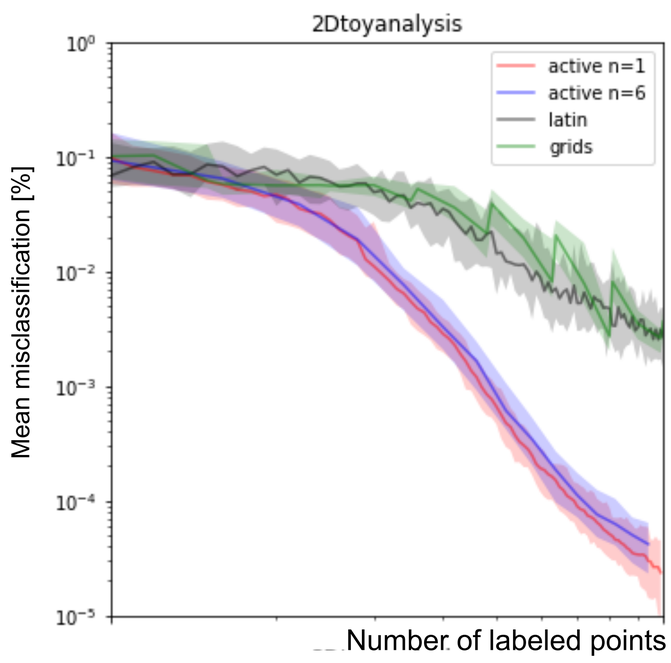



$$u_{MES}(x) = -\sum_{c_j} p_j(x) \log p_j(x) \quad \text{with } p_j(x) = \int_{t_j}^{t_{j+1}} \exp\left(-\frac{(y-\mu(x))^2}{2k(x,x')^2}\right) dy$$

Gaussian process: $Y(x) \sim \mathcal{N}(\mu(x), k(x,x'))$ thresholds: $t_j$



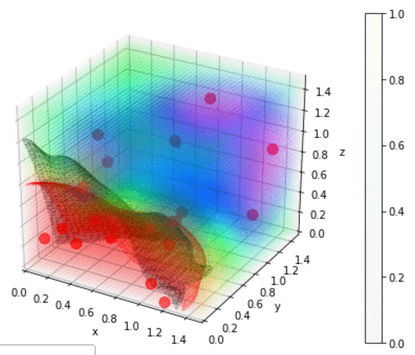

## Results

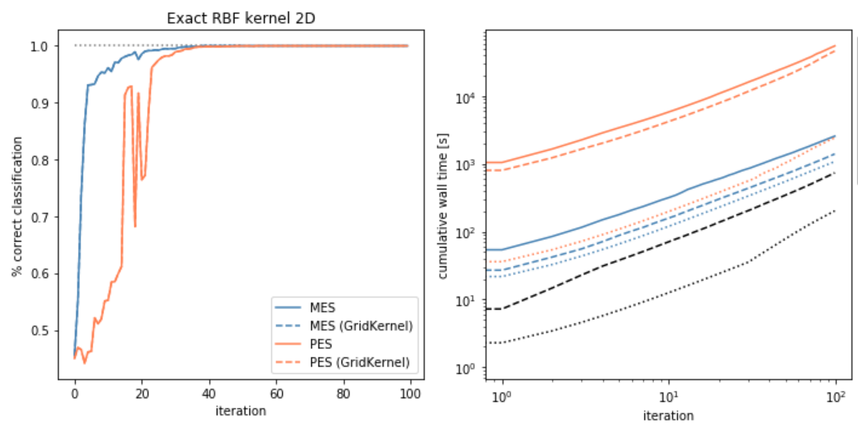



*Is it all worth it? Yes! [3]*

diana-hep/excursion

irinaespejo/excursion









[3] L. Heinrich, G. Louppe, K. Cranmer, *Excursion Set Estimation using Sequential Entropy Reduction for Efficient Searches for New Physics at the LHC*, ACAT 2019

**National Science Foundation (DGE-1922658)**
**National Science Foundation (ACI-1450310)**
**National Science Foundation (OAC-841471)**

This project is supported by National Science Foundation under Cooperative Agreement OAC-1836650. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.