ICHEP 2022 BOLOGNA
ICHEP 2022 XLI International Conference on High Energy Physics Bologna (Italy)
6 13 07 2022

Contribution ID: **1364** Type: **Parallel Talk**

Non-Parametric Data-Driven Background Modelling using Conditional Probabilities

Saturday, 9 July 2022 18:00 (15 minutes)

Background modelling is one of the main challenges of particle physics analyses at hadron colliders. Commonly employed strategies are the use of simulations based on Monte Carlo event generators or the use of parametric methods. However, sufficiently accurate simulations are not always available or may be computationally costly to produce in high statistics, leading to uncertainties that can limit the sensitivity of searches. On the other hand, parametric methods rely on the use of a functional form with free parameters to fit the observed data, which may bias the extraction of a potential signal.

A novel approach for non-parametric data-driven background modelling is presented, which addresses these issues for a broad class of searches and measurements [1]. This approach relies on a relaxed version of the event selection to estimate conditional probability density functions. Two different methods are provided for its implementation. The first is based on ancestral sampling and uses the data from the relaxed selection to obtain a graph of probability density functions of the relevant variables, accounting for the most significant correlations. A background model is generated by sampling events from this graph, before the full event selection is applied. This provides a robust implementation for cut-and-count based analyses. The strategy is further expanded in the second implementation, in which a generative adversarial network is trained to estimate the joint probability density function of the variables used in the analysis, conditioned on the variable used to blind the signal region. This training proceeds in the sidebands, and the conditional probability density function is interpolated into the signal region to estimate the background. The application of each implementation is presented and their performance is discussed.

[1] https://arxiv.org/abs/2112.00650

In-person participation

Yes

Primary authors: CHISHOLM, Andrew Stephen (University of Birmingham); NEEP, Thomas (University of Birmingham); NIKOLOPOULOS, Konstantinos (University of Birmingham); OWEN, Rhys (UKRI-STFC Rutherford Appleton Laboratory); REYNOLDS, Elliot (Berkeley National Laboratory); SILVA, Julia Manuela (University of Birmingham)

Presenter: SILVA, Julia Manuela (University of Birmingham)

Session Classification: Computing and Data handling

Track Classification: Computing and Data handling