Simulation-Assisted Decorrelation for Resonant Anomaly Detection

Luc Le Pottert, Kees Benkendorff, and Benjamin Nachman
University of Michigan, Lawrence Berkeley National Laboratory, and Reed College

ABSTRACT: Machine learning approaches to anomaly detection have recently been shown to significantly extend the search program at the Large Hadron Collider and elsewhere. One of the prototypical examples for these methods is the search for resonant new physics, where a bump hunt can be performed in an invariant mass spectrum; however, methods which focus this example rely entirely on data are susceptible to sculpting artificial bumps from the dependence of the machine learning classifier on the invariant mass. We explore two solutions to this challenge by minimally incorporating simulation into the learning.

SOLUTIONS

- The most robust solutions have false positive mitigation built into the anomaly detection techniques themselves.
- We present two solutions with this characteristic:
  1. Eliminate resonant feature differences by comparing SR to itself, as in the SALAD algorithm.
  2. Penalize the classifier for learning the resonant features, as in the SA-CWOLA algorithm.

DATA AND SELECTIONS

- 2020 LHC Olympics dataset used for prototyping
- LHC-like detector sim, using Delphes-3.4.1 and the CMS detector card (incl. both Pythia & Herwig simulations)
- Signal is $W+W$, with $m_W=3.5$ TeV, $m_H=500$ GeV, and $m_\phi=100$ GeV.
- Signal is resonant in the invariant mass distribution $m_{jj}$ for the selected SR and SB for this example are indicated by the grey lines in Figure 5.

Problem with Resonant AD

Correlations between training & resonant features cause problems for certain ML AD methods, i.e. an increase in false positives and decreased signal sensitivity.

ILLUSTRATED PROBLEM: CWOLA

The CWOLA classifier is sometimes able to infer the resonant feature set $m$ from the discriminating feature set $x$ and correctly tag the SR/SB. This effectively runs performance.

- CWOLA trained on dijet data, with some strong correlation between the jet masses and the invariant dijet mass.
- This allows the classifier to tag the entire signal region correctly as “signal-like,” ignoring the signal.
- This results in distribution sculpting, as in Figure 4.

Note that the ATLAS result in [5] avoided this with explicit decorrelation.

Conclusions & Paper

The algorithms presented here—SA-CWOLA & SALAD—show both robustness to correlation and good signal sensitivity, indicating them to be promising new analysis techniques.

See our paper for more info: arxiv.org/abs/2009.02205

Figure 1: Signal/Background specificity

A graphical representation of searches for new physics in terms of the background and signal model dependence, in terms of (a) achieving signal sensitivity and (b) background specificity. Image from Ref. [3].

Figure 2: CWOLA saminics

A visualization of the CWOLA training data (left) and an example of feature localization (right). All images from [5].

Figure 3: Model performances

ROC curves (left), relative significance improvement curves (center) for a dijet injected signal. At right, the p-value SR curves for the case of no injected signal.

REFERENCES


Figure 4: Resonant Feature Distributions

Invariant dijet mass $m_{jj}$ distributions for the various cuts on NN data (top/nr. of cuts), for CWOLA (blue) and SA-CWOLA (orange).

Figure 5: LHC Resonance

Invariant dijet mass spectrum for the LHC Olympics dataset, with SR/SB bands, and Pythia “Data / Herwig / Simulation.”

Figure 6: SALAD Reweighting

Feature distributions for simulation (“Sim”) and reweighted simulation (“Sim + DCR”) for an arbitrary 519 discriminating feature, the 1-subtlety ratio.