Simulation-Assisted Decorrelation for Resonant Anomaly Detection

Luc Le Pottier[1,2], Kees Benkendorfer[2,3], & Benjamin Nachman[2]

[1]University of Michigan, [2]Lawrence Berkeley National Laboratory, [3]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 follow this example that 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.

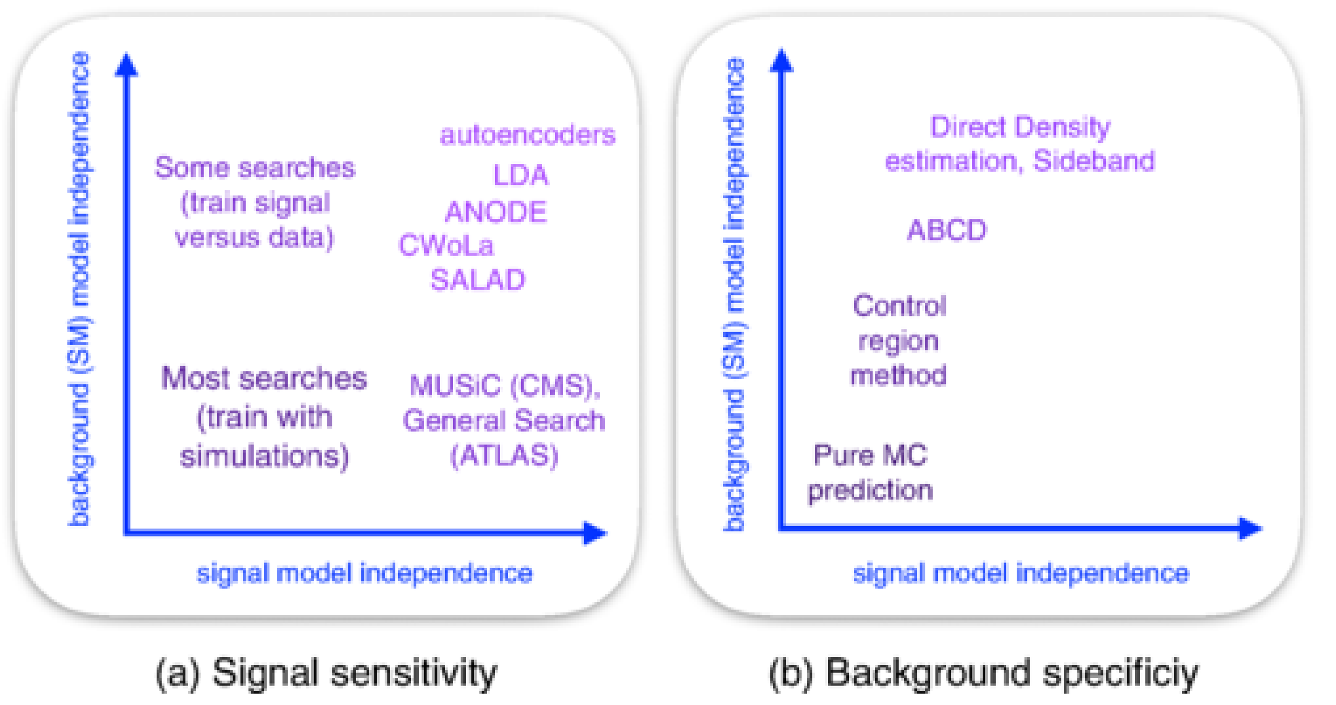

Figure 1: Signal/Background specificity

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

Motivation

While effective, the Standard Model of particle physics (SM) is still incomplete.

- No particles beyond the standard model (BSM) yet found
- LHC searches for BSM physics are generally **model-dependent**; that is, given a model, search for evidence of that model.
- There remains a vast space of signal models with no dedicated search.

Model-independent search methods have thus become more common for anomaly detection (AD)

- **Machine learning** (ML) has played an integral part
  - e.g. CWoLa [1,4,5], SALAD [2], and ANODE [3].
- Beginning to see applications in data
  - ATLAS analysis using CWoLa in Ref. [4] is the first ML-based anomaly hunt from an LHC general experiment

Resonant Anomaly Detection

Resonant signals are signals localized by a set of m **resonant features,** with some additional **discriminatory features** x

Using m, one can split data into a **signal region (SR)** where the signal is localized, and a **sideband (SB)** where it is not

Resonant anomaly detection is the process of searching for this type of signal, knowing only the SR and SB cut information

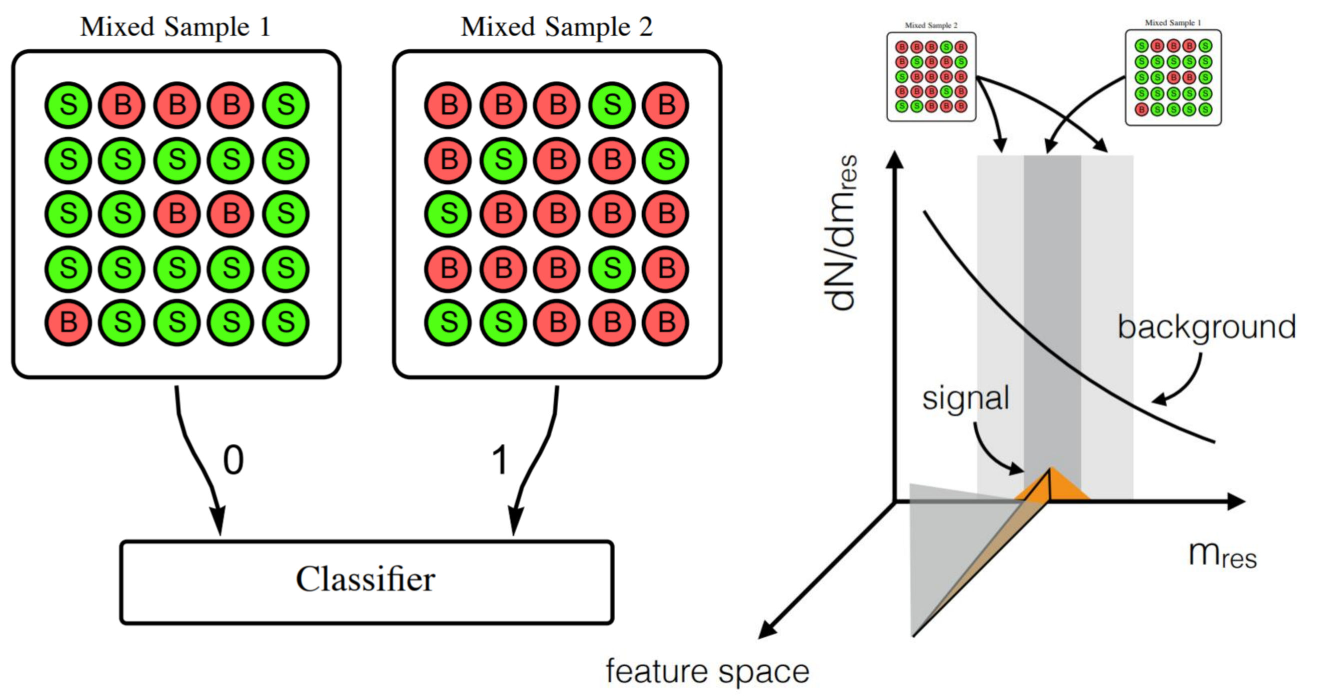

Figure 2: CWoLa schematics

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

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 $W' \to XY$, with $m_{W'} = 3.5$TeV, $m_X = 500$GeV, and $m_Y = 100$GeV
- Signal is resonant in the invariant mass distribution m_{JJ}; the selected SR and SB for this example are indicated by the grey lines in Figure 5

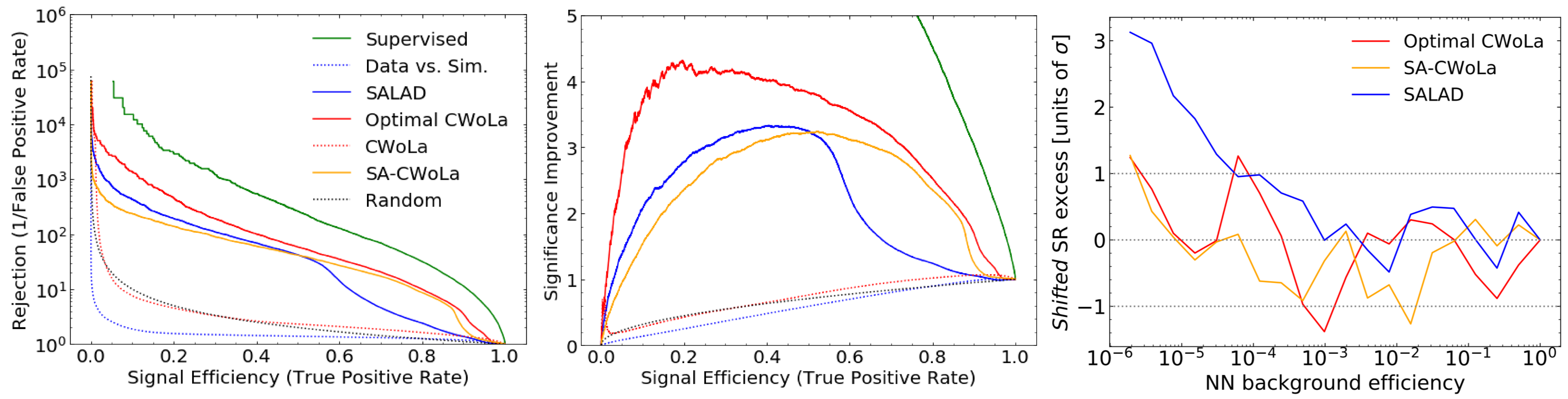

Figure 3: Model performances

ROC Curves (left), relative significance improvement curves (center) for a 2-sigma injected signal. At right, the post-fit SR excess for the case of no injected signal

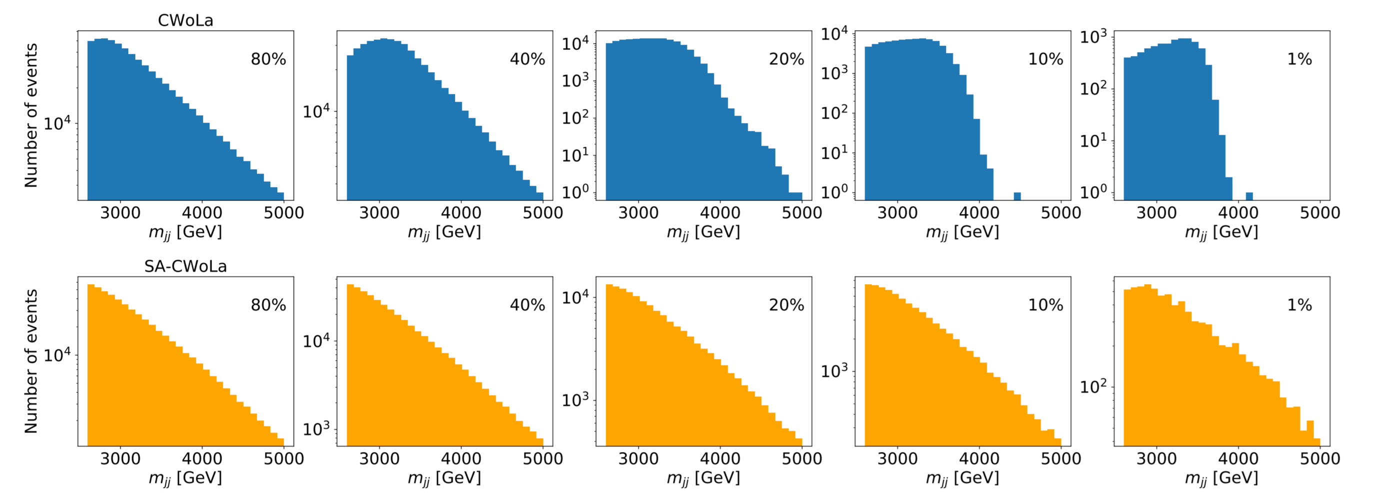

Figure 4: Resonant Feature Distributions

Invariant dijet mass m_{JJ} distributions for the various cuts on NN data (top right numbers), for CWoLa (blue) and SA-CWoLa (orange)

CWoLa

1. Split data into SR/SB and train a supervised classifier *f(x)* to distinguish between them (see figure 2).
2. For similar SR/SB, such a tagger will learn instead to tag signal. Use *f(x)* emphasize signal significance for rare signals.

SALAD

Idea: eliminate SR/SB differences by training on the SR

1. Train a classifier *f(x,m)* to distinguish data and simulation in the **SB**, using both feature sets x & m. Reweight the simulation using *f*, with weights given by

$$w(x|m) = \frac{f(x,m)}{1 - f(x,m)}$$

2. Train a classifier *g(x)* to distinguish data and reweighted simulation in the **SR**, with only discriminatory features x

SA-CWoLa

Idea: penalize the CWoLa classifier for distinguishing SR and SB in a simulated dataset.

In this way, the loss function is minimized **only by signal detection**, if the simulation is good. The loss function is given by

$$\mathcal{L}_{\text{SA-CWola}}[f] = \overbrace{-\sum_{i\in \text{SR,data}} \log(f(x_i)) - \sum_{i\in \text{SB,data}} \log(1-f(x_i))}^{\text{normal CWoLa}}$$
$$\underbrace{+\lambda\left(\sum_{i\in \text{SR,sim.}} \log(f(x_i)) + \sum_{i\in \text{SB,sim.}} \log(1-f(x_i))\right)}_{\text{penalty term}}$$

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

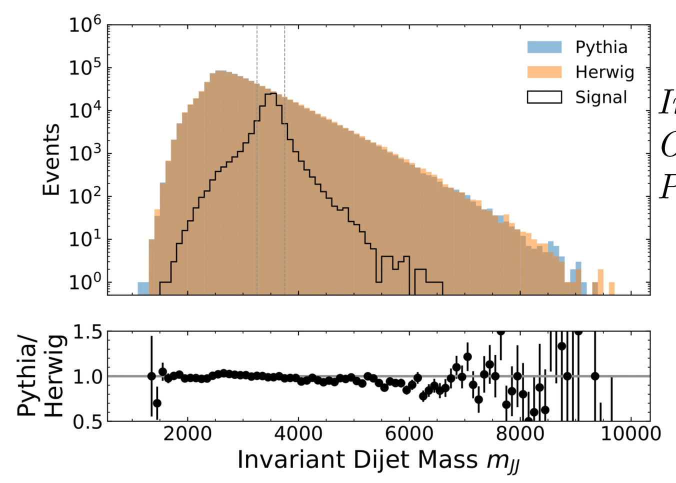

Figure 5: LHCO Resonance

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

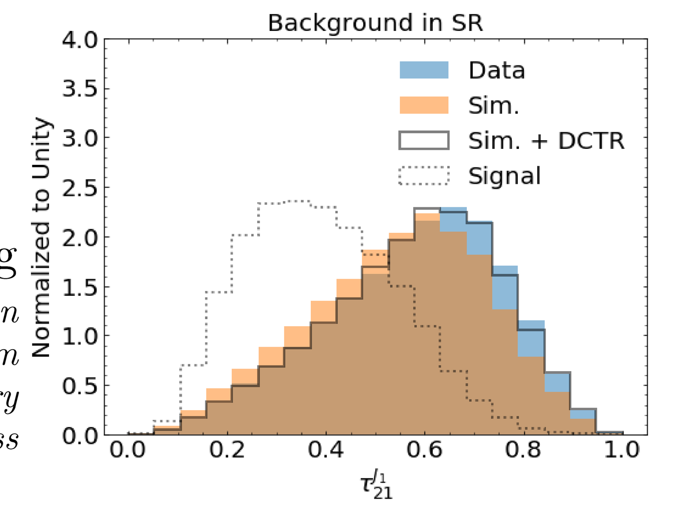

Figure 6: SALAD Reweighting

Feature distributions for simulation ("Sim") and reweighted simulation ("Sim + DCTR") for an arbitrary discriminating feature, the n-subjettiness ratio

This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program.

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.

References

1. J.H. Collins, K. Howe, and B. Nachman. "Anomaly Detection for Resonant New Physics with Machine Learning," *Physical Review Letters*, vol. 121, p. 241803, Dec. 2018.
2. A. Andreassen, B. Nachman, and D. Shih, "Simulation Assisted Likelihood-free Anomaly Detection," Physical Review D **101**, 095004 (2020)
3. B. Nachman and D. Shih. Anomaly Detection with Density Estimation. Phys. Rev. D, 101(7):075042, April 2020.
4. ATLAS Collaboration. Dijet resonance search with weak supervision using $\sqrt{s}$ = 13 TeV pp collisions in the ATLAS detector. [hep-ex], May 2020.
5. E. Metodiev, B. Nachman & J. Thaler. Classification without labels: Learning from mixed samples in high energy physics. JHEP, 2017(10):174, 10/2017.