Anomaly Detection in $Y \rightarrow XH$ & FEB2 Upgrades of the ATLAS Detector

Gabriel Matos

Columbia University

August 6, 2021
1 Introduction
   - The Standard Model
   - Detecting New Physics: LHC and ATLAS

2 Anomaly Detection in $Y \rightarrow XH$
   - $Y \rightarrow XH$
   - Anomaly Detection
   - Performance of Anomaly Score in $Y \rightarrow XH$

3 FEB2 Upgrades of the ATLAS Detector
   - Hi-Lumi LHC
   - Slice Testboard
Introduction

- The Standard Model
- Detecting New Physics: LHC and ATLAS

Anomaly Detection in $Y \rightarrow XH$

- $Y \rightarrow XH$
- Anomaly Detection
- Performance of Anomaly Score in $Y \rightarrow XH$

FEB2 Upgrades of the ATLAS Detector

- Hi-Lumi LHC
- Slice Testboard
The Standard Model

- Quantum field theory describing all known fundamental particles and their interaction
The Standard Model

- Quantum field theory describing all known fundamental particles and their interaction

- Fermions
  1. Quarks
     - $+\frac{2}{3}e$, $-\frac{1}{3}e$; color charge
     - Ex: top, bottom quarks
  2. Leptons
     - $\pm e$, neutral
     - Ex: electron, neutrinos
     - Come in three mass generations
The Standard Model

- Quantum field theory describing all known fundamental particles and their interaction

- Bosons
  1. Gauge bosons
     - Photon ($\gamma$)
     - Gluon ($g$)
     - $W^\pm$, $Z^0$
  2. Scalar boson
     - Higgs ($H$)
       - Gauge bosons mediate forces
       - Higgs gives particles mass
The Standard Model: Not All That There Is!

Decades of experiments → robust credibility to SM
- Discovery of quarks, $W^\pm$, $Z^0$ bosons, the Higgs boson...

Yet still many unanswered questions...
1. Gravity?
2. Dark matter, dark energy?
3. Matter-antimatter asymmetry?
4. Free parameters?

Motivates searches for “Beyond the Standard Model” (BSM) physics!
Producing New Physics

How to go about producing new physics...?
Producing New Physics

How to go about producing new physics...?

Build a huge collider!
The Large Hadron Collider (LHC) is a proton-proton circular particle collider in Geneva, Switzerland, built by CERN

- The largest and most powerful particle collider in the world
  - Collision energy scales of $\sqrt{s} = E_{cm} = 13\, \text{TeV}$
  - Straddles French-Swiss border in a 27 km underground tunnel
Detecting New Physics

How to go about detecting new physics...?

Build a huge detector!
The ATLAS Detector

- A Toroidal LHC ApparatuS (ATLAS) → largest detector at the LHC
- Products of proton-proton collisions are mapped by the different layers of the detector

- Going radially outwards...
  1. The inner detector (ID)
  2. EM and hadronic calorimeters
  3. Muon spectrometers
1 Introduction
   - The Standard Model
   - Detecting New Physics: LHC and ATLAS

2 Anomaly Detection in $Y \rightarrow XH$
   - $Y \rightarrow XH$
   - Anomaly Detection
   - Performance of Anomaly Score in $Y \rightarrow XH$

3 FEB2 Upgrades of the ATLAS Detector
   - Hi-Lumi LHC
   - Slice Testboard
Searching for New Physics: Model-Generic Searches

- Imperfections to the SM $\rightarrow$ search for BSM physics!
- Theoretical predictions imply new physics at energy scales reachable by the LHC
  - Key extensions of SM predict heavy diboson decay channels
  - Higgs couples to mass...

Motivates search for a generic new particle $Y$ at the TeV scale, decaying to another new particle $X$ ($\rightarrow q\bar{q}$) and a SM Higgs ($\rightarrow b\bar{b}$): $Y \rightarrow XH$

- Search in 2D space of unknown $Y$ and $X$ masses
- $X$ and $Y$ are generic $\rightarrow$ model-generic search!
Mass of SM Higgs and of the X are assumed to be in $O(100 \text{ GeV})$ scale → produced with high momentum
- Decay products are highly Lorentz boosted and collimated
- Because of quark confinement: $\star$ → quarks → hadronization → jets!

High momentum in $Y \rightarrow XH$ causes individual jets to be merged
- Sensitivity is regained by reconstructing decay product as large-radius jet instead of two individual jets
**X Tagging & Anomaly Detection**

- How to distinguish background from interesting events? **Tagging!**
- How are X jets tagged?
  - Past analysis constrained X to decay to two quarks → two-prong substructure
  - Used a $D_2$ cut to quantify two-pronginess and do the X tagging
- Why claim the X decays only to two quarks? Limiting sensitivity to new potential BSM physics
- Remediate...

**Anomaly Detection: the identification of outliers from known data**

- Implies no signal hypothesis → only need to learn features of background/data
- Allows for sensitivity to any signature that is “anomalous”
How to Detect Anomalies?

- How is anomaly detection done? **Neural Networks!**
- Design a neural network that...
  1. Takes a dataset as input ($x$)
  2. Encode data points into their most salient features (latent space)
  3. Reconstructs (decodes) data points as output ($y$)

- The difference between the output $y$ and input $x$ is called the “loss”
- Train neural network over data → learns general features of data
  - **Good** reconstruction of things that resemble average data (small loss)
  - **Bad** reconstruction of things that don’t resemble data (large loss)
- Loss can be used as discriminant for anomalous data!
Can apply anomaly detection to physics analysis with Variational Recurrent Neural Network (VRNN)

Use VRNN’s loss function to build an **Anomaly Score** (AS)
- Train directly over data
- High anomaly score $\rightarrow$ interesting anomalous substructure

Study performance over $Y \rightarrow XH$ signal samples + special samples (w/ different substructure hypotheses) for model independence:

1. Find optimal AS cut that maximizes sensitivity $S/\sqrt{B}$
2. Cut on AS and look for bump in reconstructed jet masses $m_{JJ}$
3. Evaluate performance by computing percent increase in sensitivity
Cut Optimization: Anomaly Score Distributions

- AS distributions for $Y \rightarrow XH$ and “special” signals
- Scan over different AS cuts, plot sensitivity $S/\sqrt{B}$ metric, pick average max for all 6 distributions $\rightarrow$ AS $> 0.4$
Results: Background-Only Distros & Mass Sculpting

Selection on Anomaly Score does not produce significant bump (or false signal) when no signal is present
- No significant sculpting of $m_{JJ}$ shape
- Little sculpting on most $m_X$ spectrum except low masses
Results: $Y \rightarrow XH$ Signal-Injected $m_{JJ}$

- Plot background with signal injected before and after AS cut → look for bump in distribution + increase in $S/\sqrt{B}$
- Selection on AS enhances significance on bump in $m_{JJ}$ by $\sim 16\%$

![Graphs showing data distribution before and after signal injection](image)

VcXH (2-prong)
Results: “Special” Signal-Injected $m_{JJ}$

- Selection on AS enhances significance on bump in $m_{JJ}$ by $\sim 10\%$
- AS lets us be sensitive to both 2-prong AND 3-prong

![Graphs showing data before and after selection on AS with CWoLa 3-prong results](image)
That’s not all...

- **Anomaly Score** lets us be sensitive to substructure hypotheses we weren’t before!
  - Part of bigger study that shows AS boosts sensitivity for two other substructure hypotheses by $\sim 5\text{-}15\%$
  - Across all boosted $Y \rightarrow XH$ signal by $\sim 5\text{-}30\%$
  - Lets us be *truly* model-independent search and sensitive to many types of new physics!

- **Some next steps...**
  - Implementation in $Y \rightarrow XH$ requires creating new analysis region for background estimation
  - Perform tests with different trainings
    - Gauge performance varying epoch, training selection, etc.
Introduction
- The Standard Model
- Detecting New Physics: LHC and ATLAS

Anomaly Detection in $Y \rightarrow XH$
- $Y \rightarrow XH$
- Anomaly Detection
- Performance of Anomaly Score in $Y \rightarrow XH$

FEB2 Upgrades of the ATLAS Detector
- Hi-Lumi LHC
- Slice Testboard
Electromagnetic (EM) Calorimeter

- The ATLAS calorimeter → layers of lead immersed in liquid argon (LAr)
- Particles passing through the calorimeter are captured by the metal layers and converted to showers of lower energy particles
  - Showers ionise the LAr → current → measure with readout electronics
  - The calorimeter cells are sampled at 40 MHz by the readout system
Hi-Lumi LHC

- Upgrades to the LHC → High Luminosity (HL) LHC
  - Increase instantaneous luminosity
  - Pile up: number of simultaneous proton collisions per bunch crossing
- Need to develop new LAr calorimeter readout to accommodate higher luminosity requirements
Front End Boards Upgrade

- Readout Front-End Boards (FEBs) → read out + digitize LAr calorimeter signals
- FEB2 Upgrade → replaces current FEB boards for HL-LHC

FEB2

- 128 calorimeter channels; 1542 boards
- Dataflow
  1. Input calorimeter signal
  2. Preamplifier/shaper (LAUROC chips)
  3. Digitization → 16-bit words @ 640 Mbps (COLUTA chips)
  4. Repackaging (lpGBT chips)
  5. Off detector
R&D → Slice Testboard (sliceboard)

- "Slice" of final FEB2
  - 32 of 128 channels implemented
  - Same chip density + layout as final FEB2

- Full board control + readout implemented through GUI
  - Can configure and read back all LAUROC, COLUTA, and lpGBT chips
Slice Testboard

- R&D → Slice Testboard (sliceboard)
- “Slice” of final FEB2
  - 32 of 128 channels implemented
  - Same chip density + layout as final FEB2
- Full board control + readout implemented through GUI
  - Can configure and read back all LAUROC, COLUTA, and lpGBT chips
The graphical user interface (GUI) allows for data taking and configures and calibrates all chips in the sliceboard.
The graphical user interface (GUI) allows for data taking and configures and calibrates all chips in the sliceboard.
Clock Configurations

▶ Each COLUTA chip has 8 channels that map to lpGBT registers
▶ Timing settings between COLUTA and lpGBT chips need to match for data to come out correct → determine clock configurations
▶ Three parameters:
  1. INV/Delay640: adjusts when COLUTA outputs data in 640 MHz ticks (applies to all channels)
  2. xPhaseSelect: determines when to write data to lpGBT (can change @ channel level)
  3. lpGBT phase: determines when lpGBT reads data (can change @ register level)
▶ Determine working clock parameters → clock scan
Clock Scan

How to scan clock parameters?

- Each COLUTA chip/channel combination has a serial pattern that we know
  - Can vary INV/Delay640 and xPhaseSelect parameters
  - Read serial data
  - Determine if it’s phase shifted or invalid (wrong pattern)
  - Find lpGBT phase so that pattern is correct
- Scan through all possible clock parameters → find all possible lpGBT phases

![Clock Scan Result Table]

*Gabriel Matos*  
Anomaly Detection in $Y \rightarrow XH \& FEB2$  
August 6, 2021
Finding Best Clock Parameters

- Delay parameters are $O(\text{ps}) \rightarrow$ may be subject to fluctuations
- Want to pick lpGBT phase that is in “island” of same phases
  - If fluctuation does occur, won’t read gibberish
- Algorithm that determines score for each clock configuration...

- Clock scan rewrite configuration files for the board, implements new parameters
- Phase determined can be validated by a serializer data “checker”
A Lot More Happening...

▶ A lot of nextsteps!
  • 32-bit mode clock scan...
  • SAR/MDAC Calibration...
  • Energy resolution studies...
  • COLUTAV4 development...

▶ Plan to have next version of test board with full 128 channels by next summer!

▶ A lot happening in other fronts too: most recent LAr week talks
Acknowledgements

A special thanks to...

- Everyone at Nevis and to John, Georgia, and Amy for organizing the REU program
- John, Julia, Alan, and Elena for getting me up to speed on a lot of the content in this talk
- The others REUs I’ve gotten to know over the past two months
- The NSF for funding this research

Thank you!
BACKUP
The masses of the $X$ and $Y$ are unknown $\rightarrow$ 2D grid of simulated signal points.
Autoencoders

- **Autoencoder (AE):** a model that is able to encode an input to a lower-dimensional latent space and pick out its most salient features, and also decode from that latent space, checking for reconstruction errors.

- **Variational Autoencoder (VAE):** encodes to a probability distribution in the latent space and allows for Bayesian inference by sampling from this latent space.

![VAE Cell Diagram]
Variational Recurrent Neural Network (VRNN): recurrent neural network that updates a variational autoencoder at each time step and accommodates for variable-length inputs.

\[ \mathcal{L}(t) = |y(t) - x(t)|^2 + \lambda D_{KL}(z || z_t) \]

Mean Squared Error + Kullback-Leibler Divergence

- Train over data using large-R jet constituent 4-vectors as inputs to the NN
Anomaly Score Definition

- Anomaly Score is defined based on the KL-Divergence term of the VRNN loss function
  \[
  \text{Anomaly Score} = \rho = 1 - e^{-D_{KL}}
  \]

- The AS as is gives signal events a score near zero → transform such that signal gives higher anomaly score
  \[
  \rho' = 1 - \left( \frac{\rho}{2\bar{\rho}} \right)
  \]
Results evaluated at epoch 57 (out of 100) of training

1. 3 sample $Y \rightarrow XH$ simulation signals were chosen: (2000, 160), (3400, 500), and (5000, 200)
2. 3 “special” signals corresponding to other substructure: CWoLa bb, CWoLa 3-prong, Dark Jets

- Signals = normalized to 36.1 fb$^{-1}$
- Data = taken from HSB, normalized to 36.1 fb$^{-1}$
- Anomaly Score treated as sole analysis selection (no Higgs tagging/information)
Training Results: ROCs

VcXH (2000, 160)

VcXH (3400, 500)

VcXH (5000, 200)

CWoLa bb

CWoLa 3-prong

Dark Jets
Full Results: $Y \rightarrow XH$ Signal-Injected $m_{JJ}$

Before:

![Histograms showing event distributions before](before_histograms)

VcXH (2000, 160)

VcXH (3400, 500)

VcXH (5000, 200)

After:

![Histograms showing event distributions after](after_histograms)

VcXH (2000, 160)

VcXH (3400, 500)

VcXH (5000, 200)
Full Results: “Special” Signal-Injected mJJ

Before:

After:
Computing percent increase in sensitivity for every signal point in signal grid after AS selection

AS cut gives \(\sim 5\text{-}30\%\) increase in sensitivity to boosted \(Y \rightarrow XH\) signal points

Sensitivity suffers in resolved regime (less prominent correlations between constituents in large-R jets)