Anomaly Awareness for New Physics Searches

Charanjit Kaur Khosa
University of Sussex

Anomaly Detection Workshop, 17th July 2020

In collaboration with Veronica Sanz
Based on arxiv[hep-ph]: 2007.XXXX (appearing next week)

UNIVERSITY OF SUSSEX

newton international fellowships

Anomaly Awareness (AA)

We will present a new algorithm for anomaly detection called Anomaly Awareness (AA)

The algorithm learns about normal events (SM) while is made aware of the possibility of anomalies (BSM)

in a way that it becomes sensitive to ANY kind of BSM anomalies

In this talk we will show how AA works in a well-known topology for new physics searches
FAT JETS
and test it against an array of BSM scenarios: EFT Higgs, Resonances -> leading jet with 2, 3 or 4 subsets

New potential for New Physics Searches

(Deep)Neural Networks, CNNs, (V)Autoencoders, Clustering,..

Classification, Jet tagging, Anomalous Jet, Anomalous Events, Limit setting, Resonance,..

Assisting BSM detection

This talk:

- **BSM**: Boosted Regime
- **Input data**: Jet Images
- **Model**: CNN

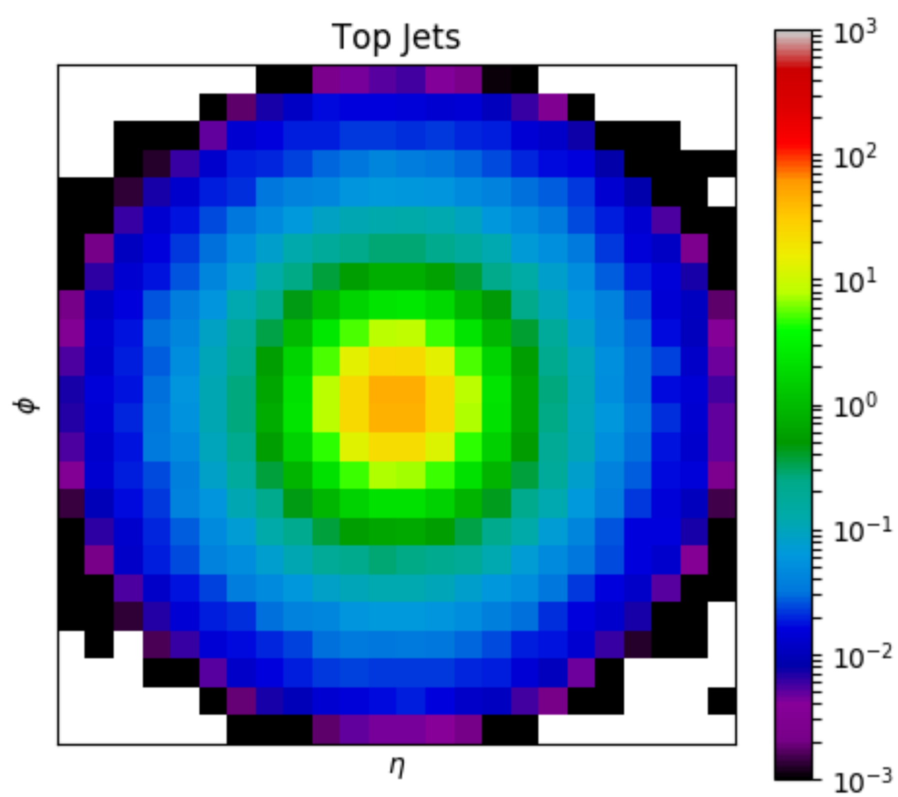
Top and QCD Jets

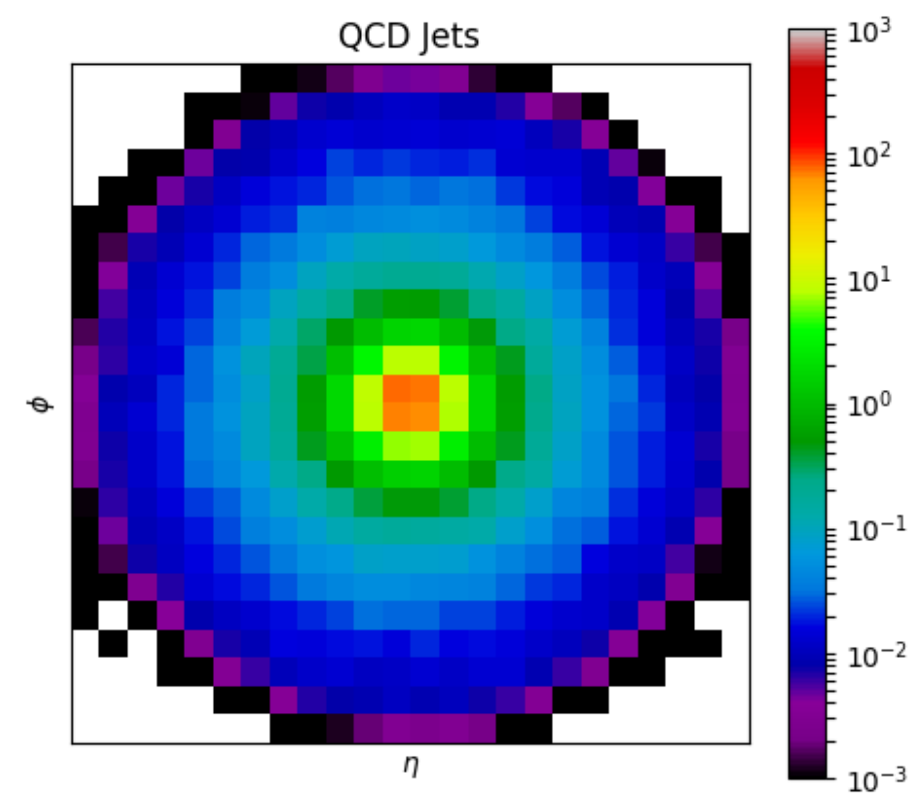
SM $t\bar{t}$ and QCD diet production, $\sqrt{s} = 13$ TeV

Madgraph + pythia

Leading jet with pt > 750 GeV, R=1 Anti-kt jet

$\Delta\eta = 0.087, \Delta\phi = 0.087$





Leading jet (Averaged over 50K events)

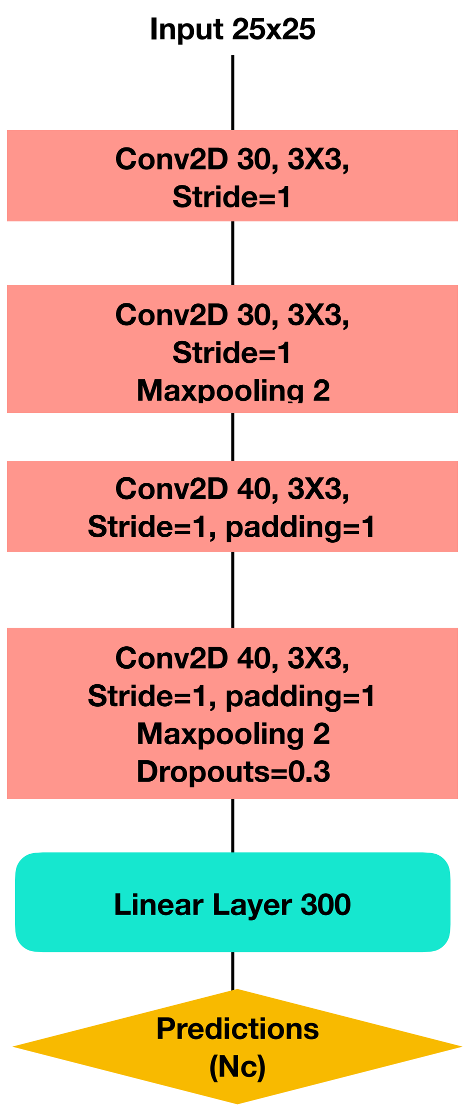
CNN for Top vs QCD Classification

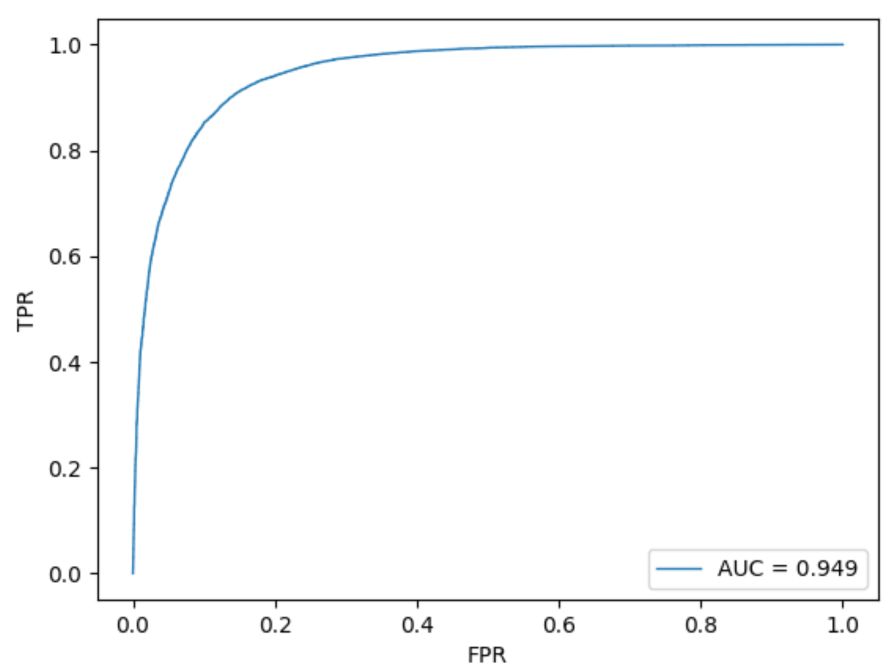
(2C Baseline Classification)

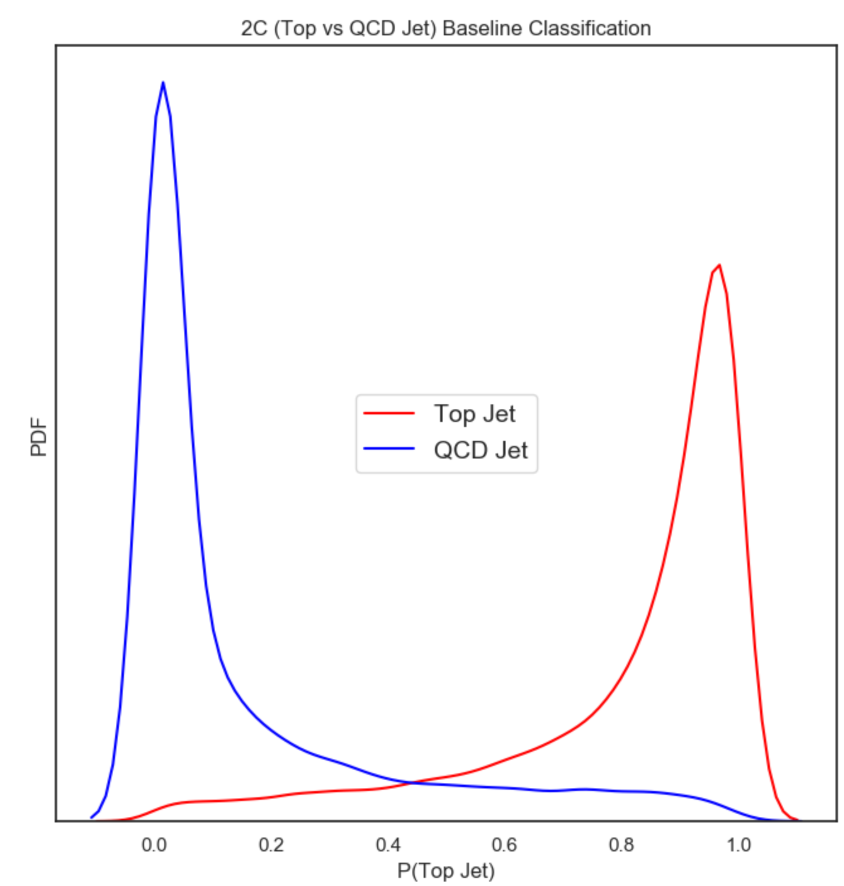
Batch Size=100
Epochs=100

100K Images (balanced data)
Training:Test data= 70:30%

Cross-Entropy Loss function







Related work
G.Kasieczka, T.Plehn, M.Russell and T.Schell, JHEP 05 (2017), 006
S.Macaluso and D.Shih, JHEP 10(2018), 121

BSM Benchmarks

SM

$$W jet : pp \to W^+W^-, W \to jj$$

HEFT $C_{HW} = 0.3 \;\; ([D^\mu H^\dagger T_{2k} D^\nu H] W^k_{\mu\nu})$

$$\mathrm{EFT} : pp \to HZ, H \to b\bar{b}, Z \to l^+l^-$$

Resonance (Graviton) decay, mY=3 TeV

$$R_2 : pp \to Y \to ZZ, Z \to jj$$

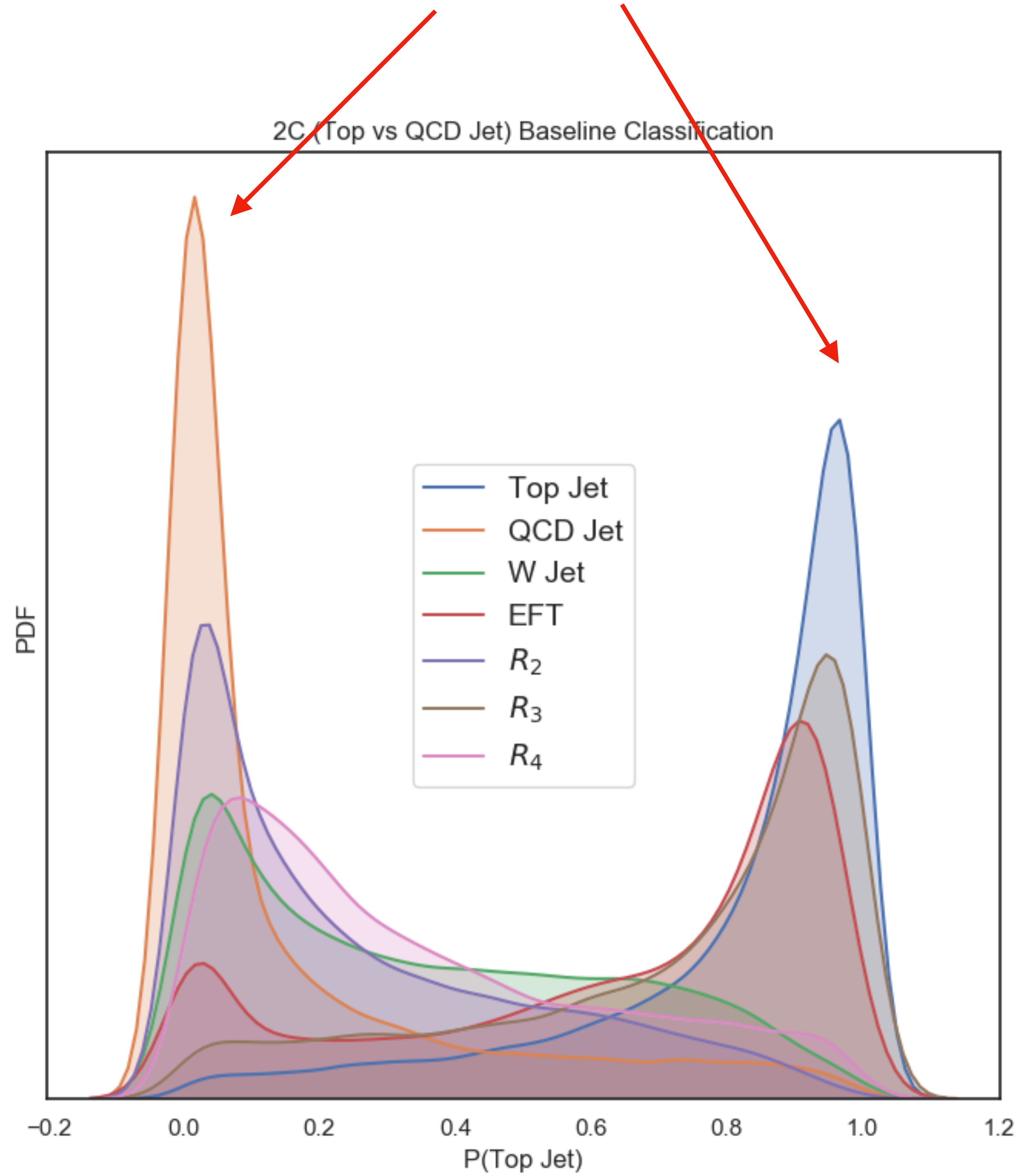
$$R_3 : pp \to Y \to t\bar{t}, t \to bW, W \to jj$$

$$R_4 : pp \to Y \to HH, H \to W^+W^-, W \to jj$$

Leading jet, same cuts as baseline jets

2C Baseline Classification

Training on



Softmax Probability

Anomaly Awareness

Algorithm 1 Anomaly Awareness (AA).

Prior Run

Initialize test:train splitting of *Normal* dataset (N)
Initialize hyper parameters
Initialize Model (CNN architecture)
for Training over the epochs **do**
  Cross entropy loss
  Update model parameters.
end for
Get accuracy for D_{test} and D_{train}
This run sets the hyper-parameters for the AA run

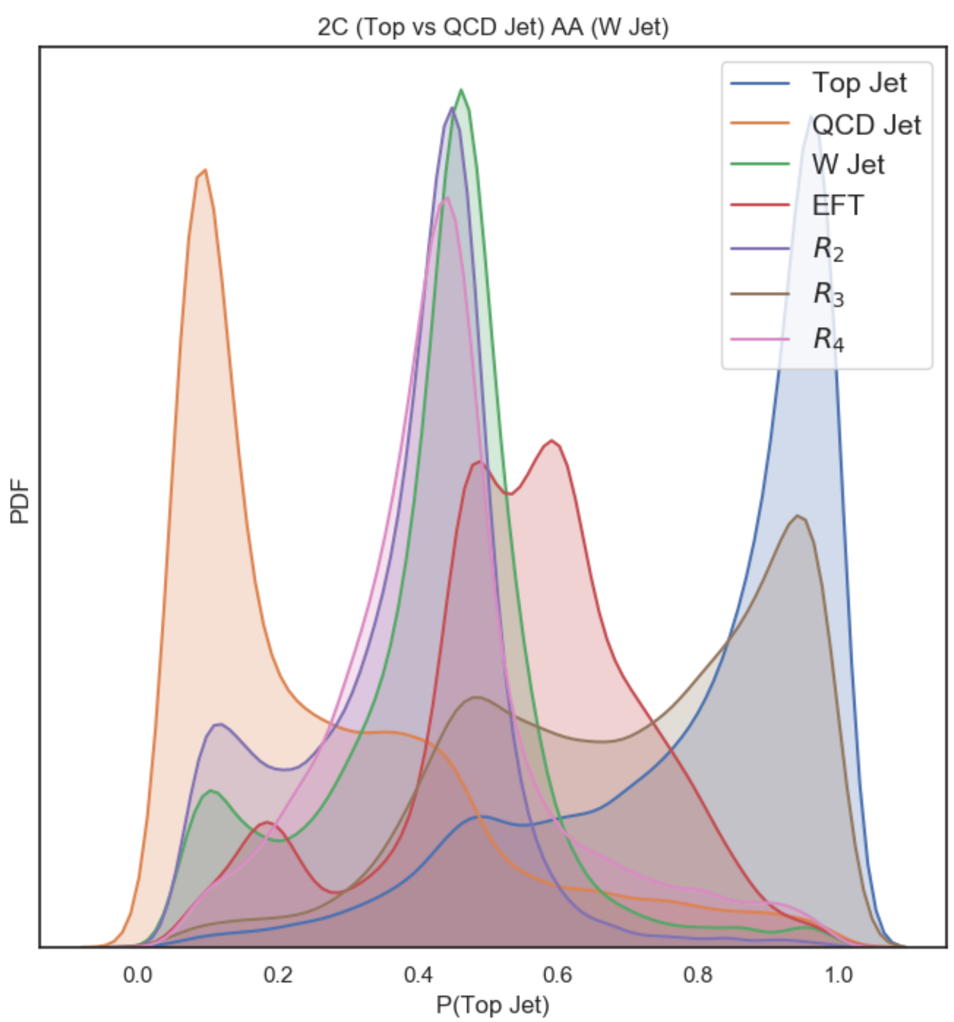
Anomaly Detection Run

Load the *Anomaly* (An) dataset
Initialize amount of data w.r.t. the *Normal* dataset
Initialize λ_{AA}
for Training over the epochs **do**
  l_1 = Cross entropy loss (*Normal* dataset)
  l_2 = Cross entropy loss (*Anomaly* dataset with Uniform Distribution)
  Loss = $l_1 + \lambda_{AA} l_2$
end for
Get softmax probabilites for all the data sets,
p_i, $i = N$, An
Select datapoints in a range $[p_{An}^{min}, p_{An}^{max}]$
optimized to select anomaly over normal events

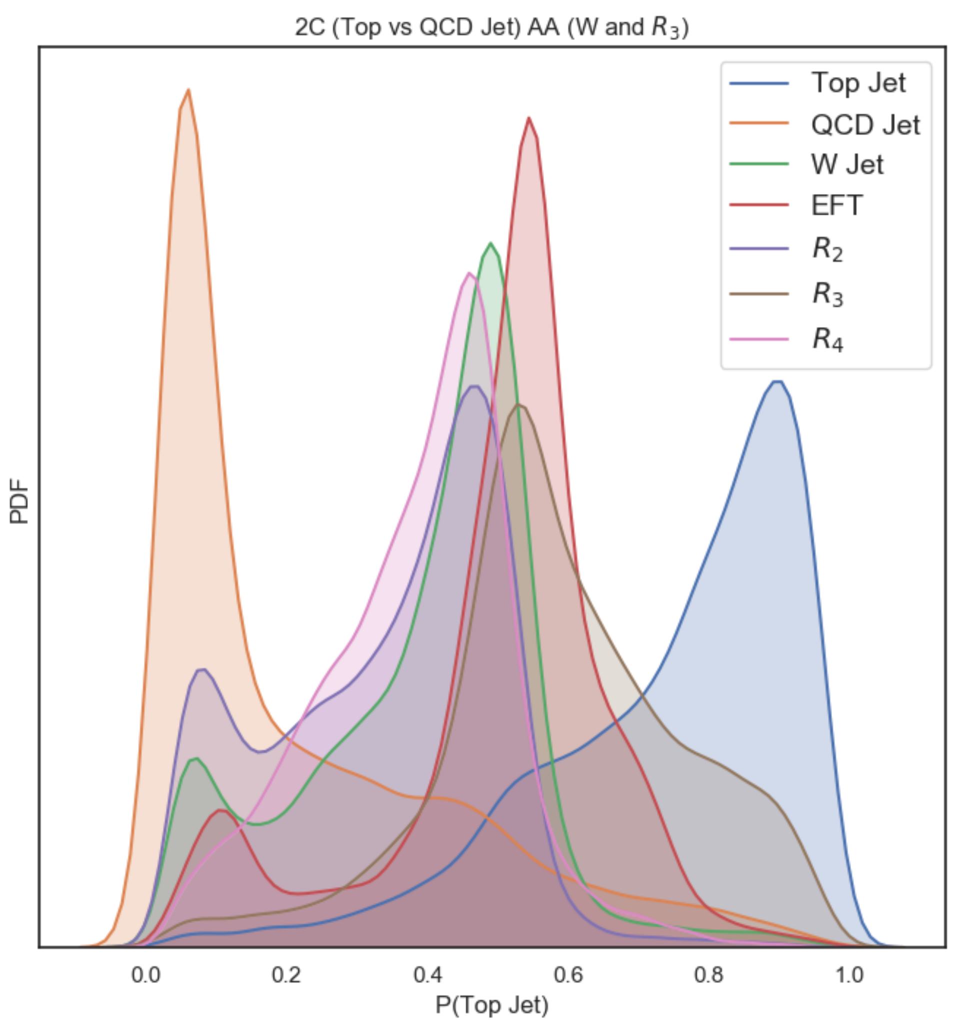
Anomaly Awareness for 2C Data

We see the effect of adding awareness to the classification task

One type of anomaly data set



Two types of anomaly data set



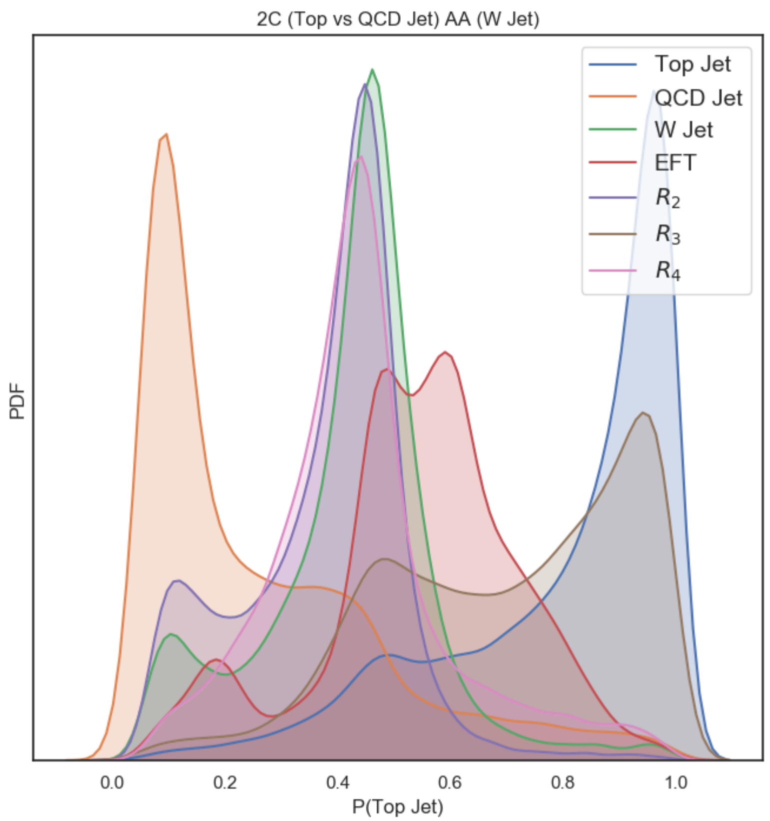
As we add more types of BSM examples, ALL BSMs gather in the centre

Uniform Distribution to the AA term for all the BSM events

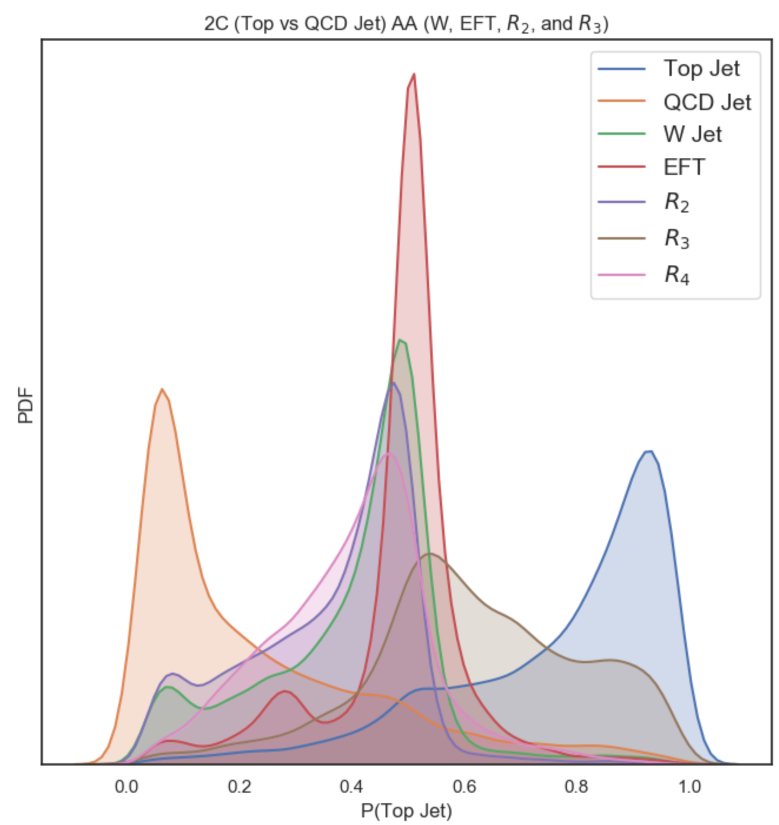
Robust Anomaly Detector

Awareness of a variety of anomalies

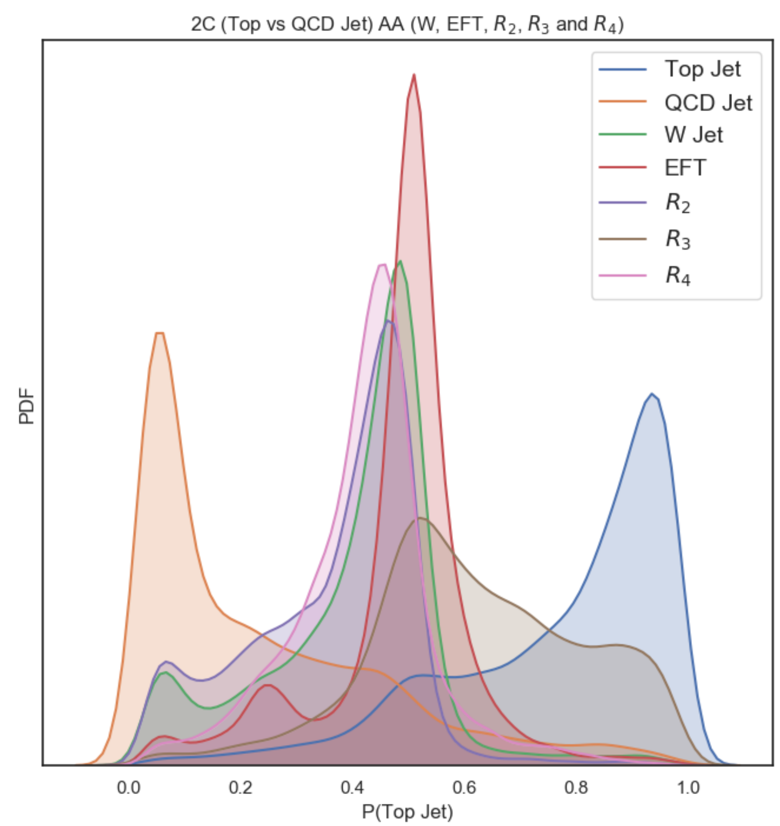
AA term with 1 type of anomaly



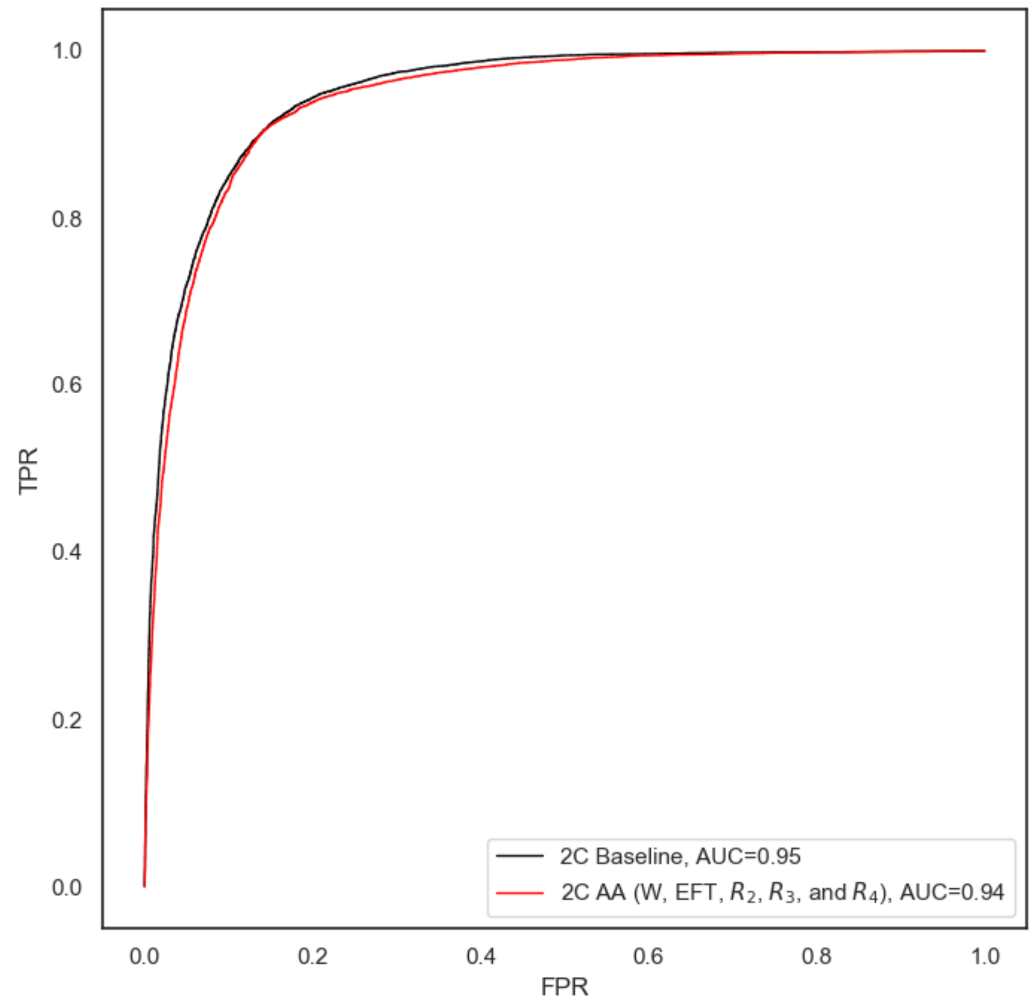
AA term with 4 types of anomalies



AA term with 5 types of anomalies

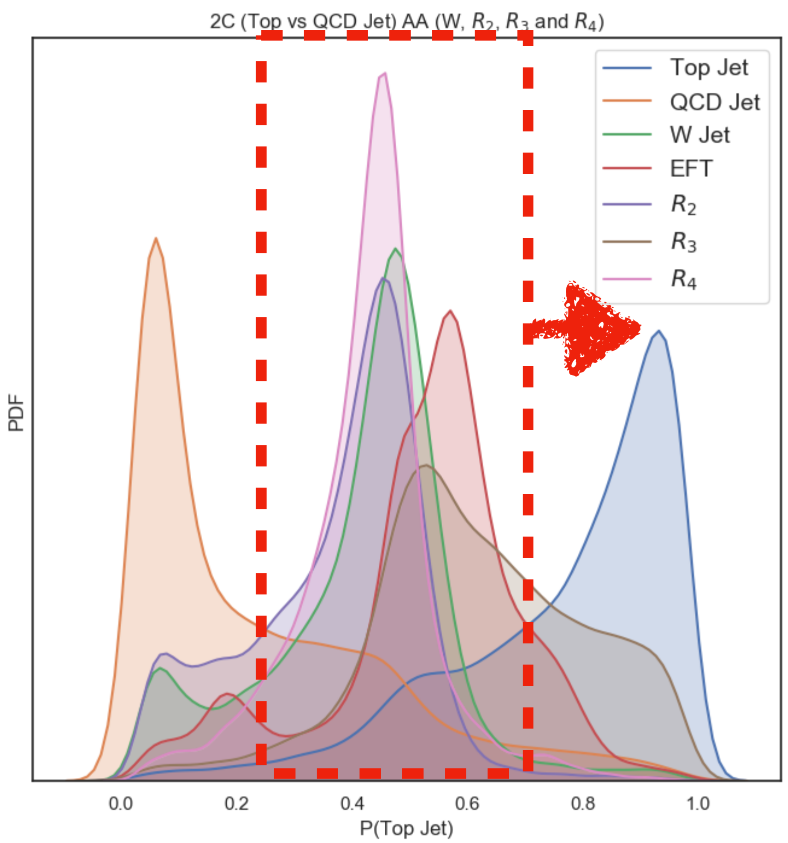


Baseline vs AA Comparison



The addition of the AA term does not degrade the baseline classification but adds the ability to use its output for anomaly detection

Signal Cross-section Reach



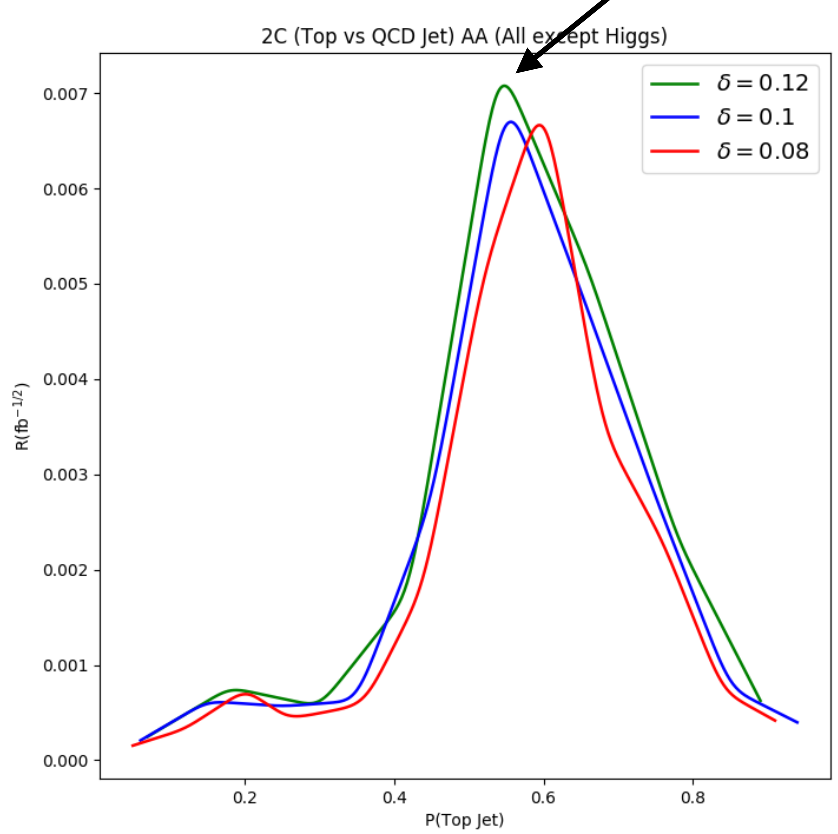
We scan on windows of the classifier output
Cutting a small window around 0.5
anomaly detection is enhanced
We use S/Sqrt[B] as an example of quantity to maximise (S=BSM, B=SM)

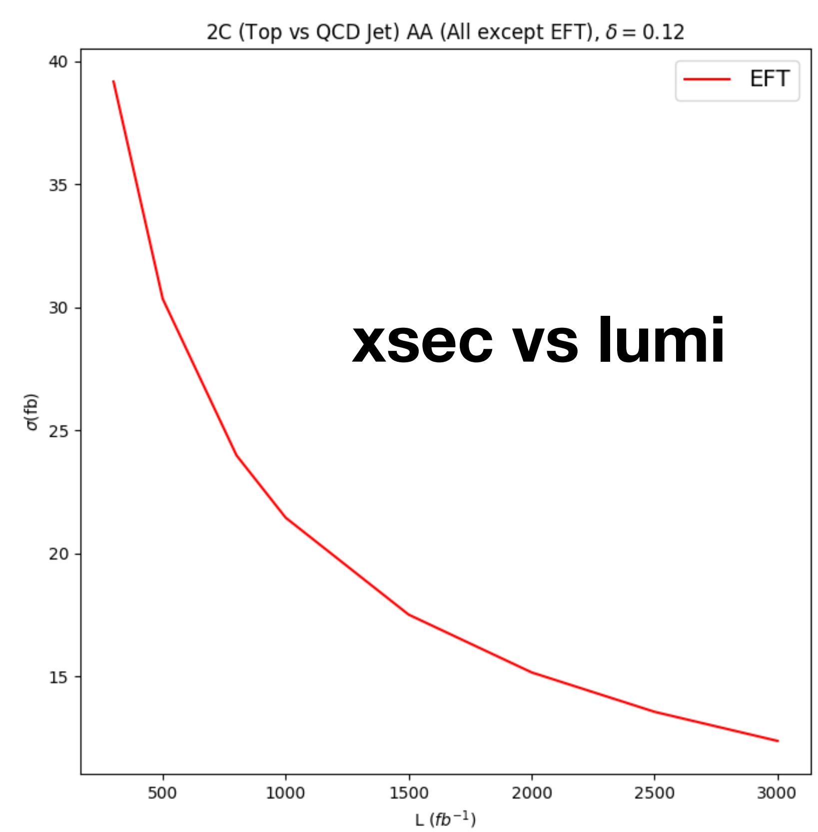
$$R = \frac{\epsilon_S}{\sqrt{\sigma_B \epsilon_B}}$$

$$\frac{S}{\sqrt{B}} = 5$$

optimal cut

translation into a generic anomaly xsec



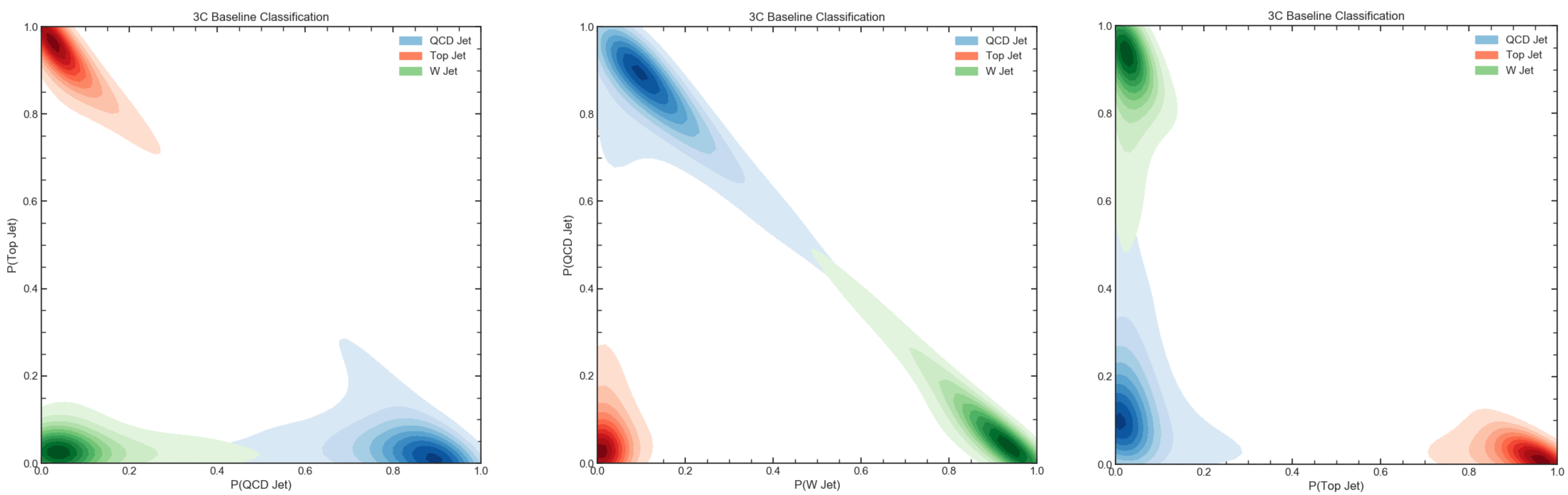


Three-classes (baseline) Classification

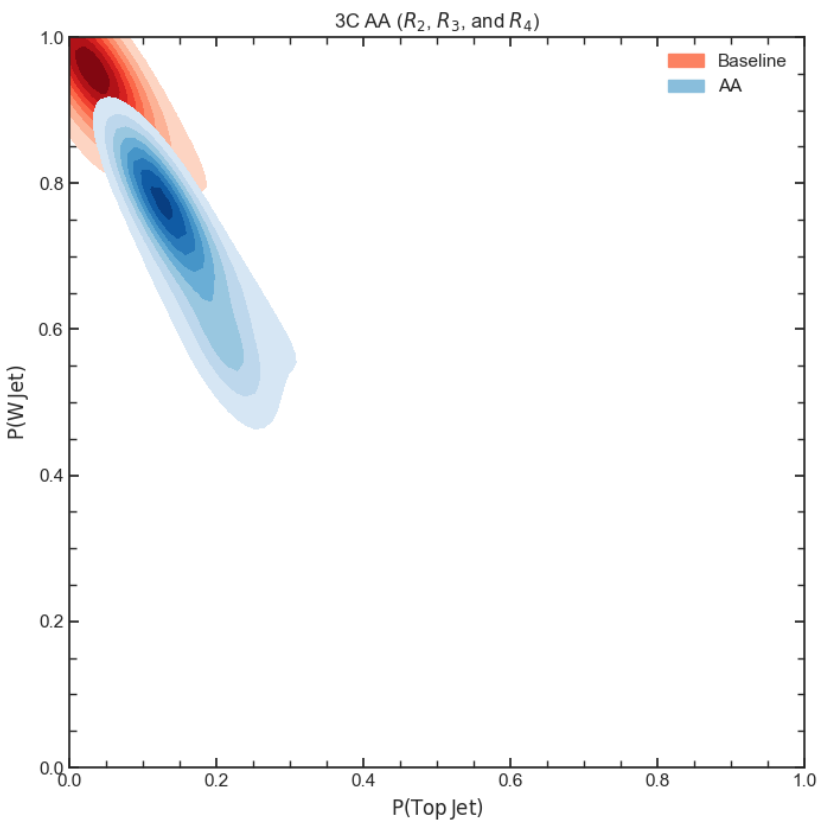
This procedure can be generalized beyond binary classification

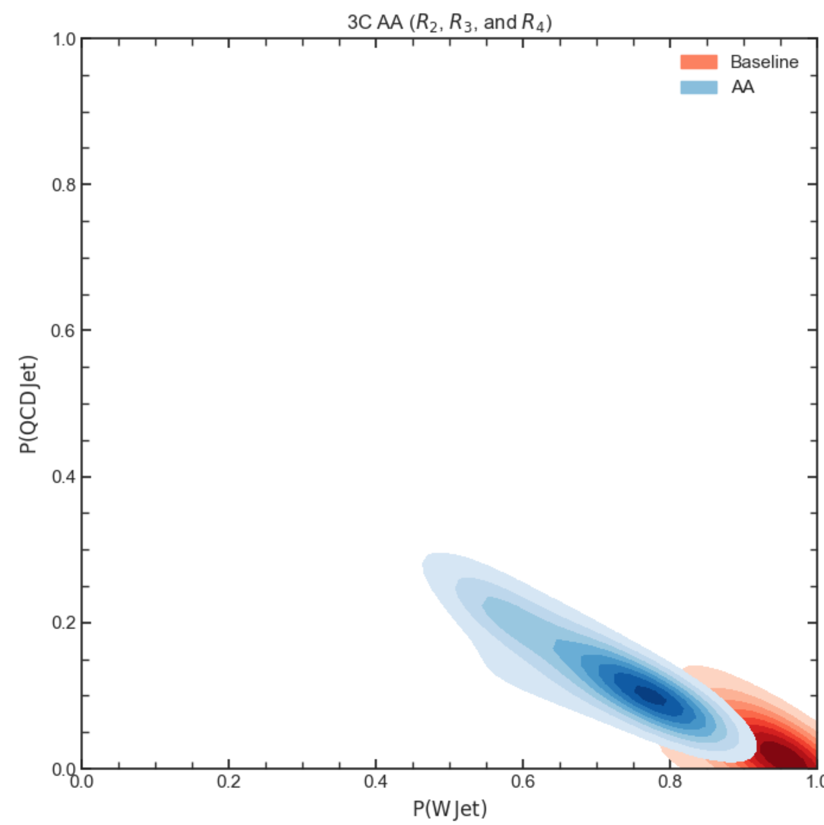
Top Jet, QCD jet, W-jet

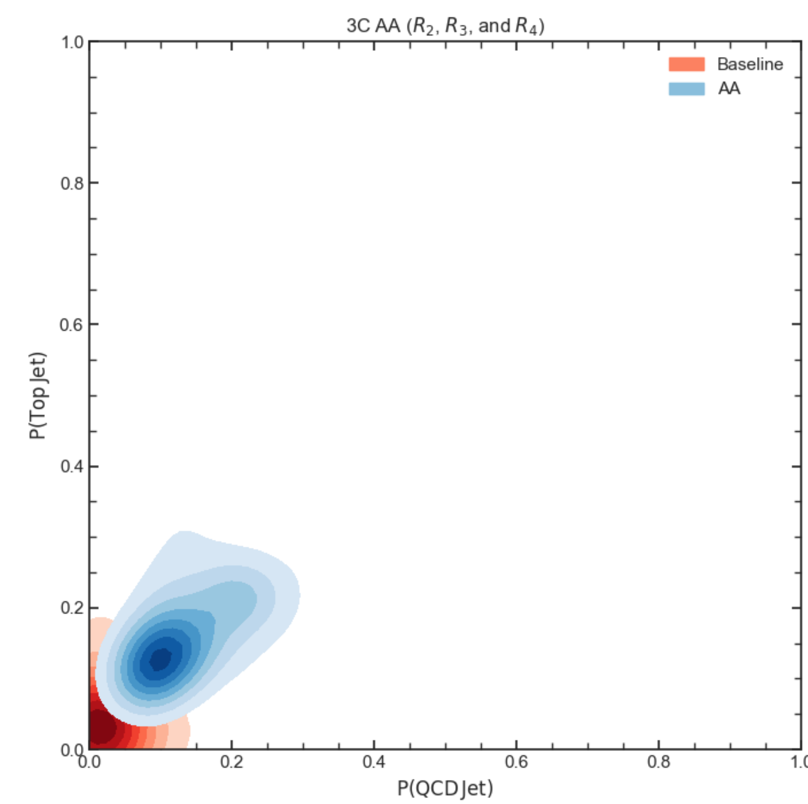
150 K images (balanced data set), training:test=70:30%



Anomaly Awareness for 3 Classes







Unseen Data Set: EFT

Summary and Outlook

We present a new algorithm for anomaly detection
It is based on the procedure of classifying 'normal' (SM) events
and make aware during that classification of the presence of anomalies (BSM)

We find that the procedure is effective on BSM anomalies not seen before and becomes robust as we make the algorithm aware of a varied-enough set of anomalies

- We demonstrate the potential of anomaly awareness method for the boosted regime using jet images for the event representation and CNN classification model

Next Steps:

- Using different models
- Use it for a large set of kinematic variables to capture variety of new physics
- Comparison with other anomaly detection methods
- Using it for LHCO data set