Decorrelated Jet Substructure Tagging using Adversarial Neural Networks

Chase Shimmin
Yale University

IML Workshop
March 22, 2017

Details can be found in:

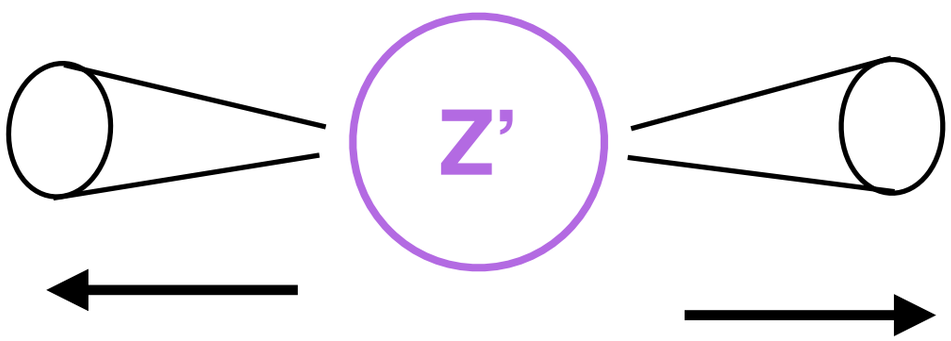
"Decorrelated Jet Substructure Tagging using Adversarial Neural Networks"

C. Shimmin, P. Sadowski, P. Baldi, E. Weik, D. Whiteson, E. Goul, A. Søgaard

https://arxiv.org/abs/1703.03507

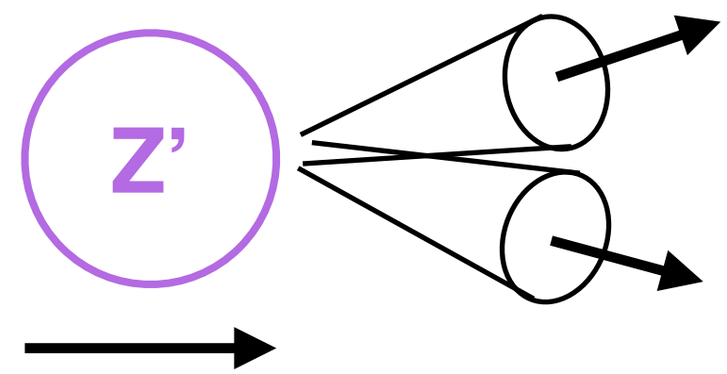
Boosted Objects

at rest



$p_T^{jet} \sim m_{Z'}/2$

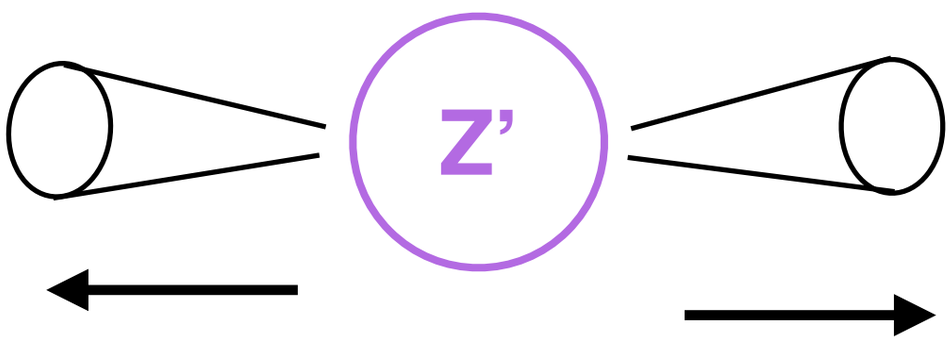
boosted



$p_T^{Z'}/m_{Z'} > \sim 1$

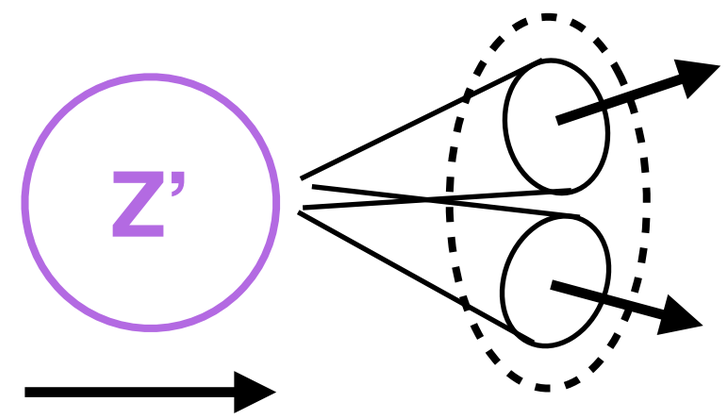
Boosted Objects

at rest



$p_T^{jet} \sim m_{Z'}/2$

boosted

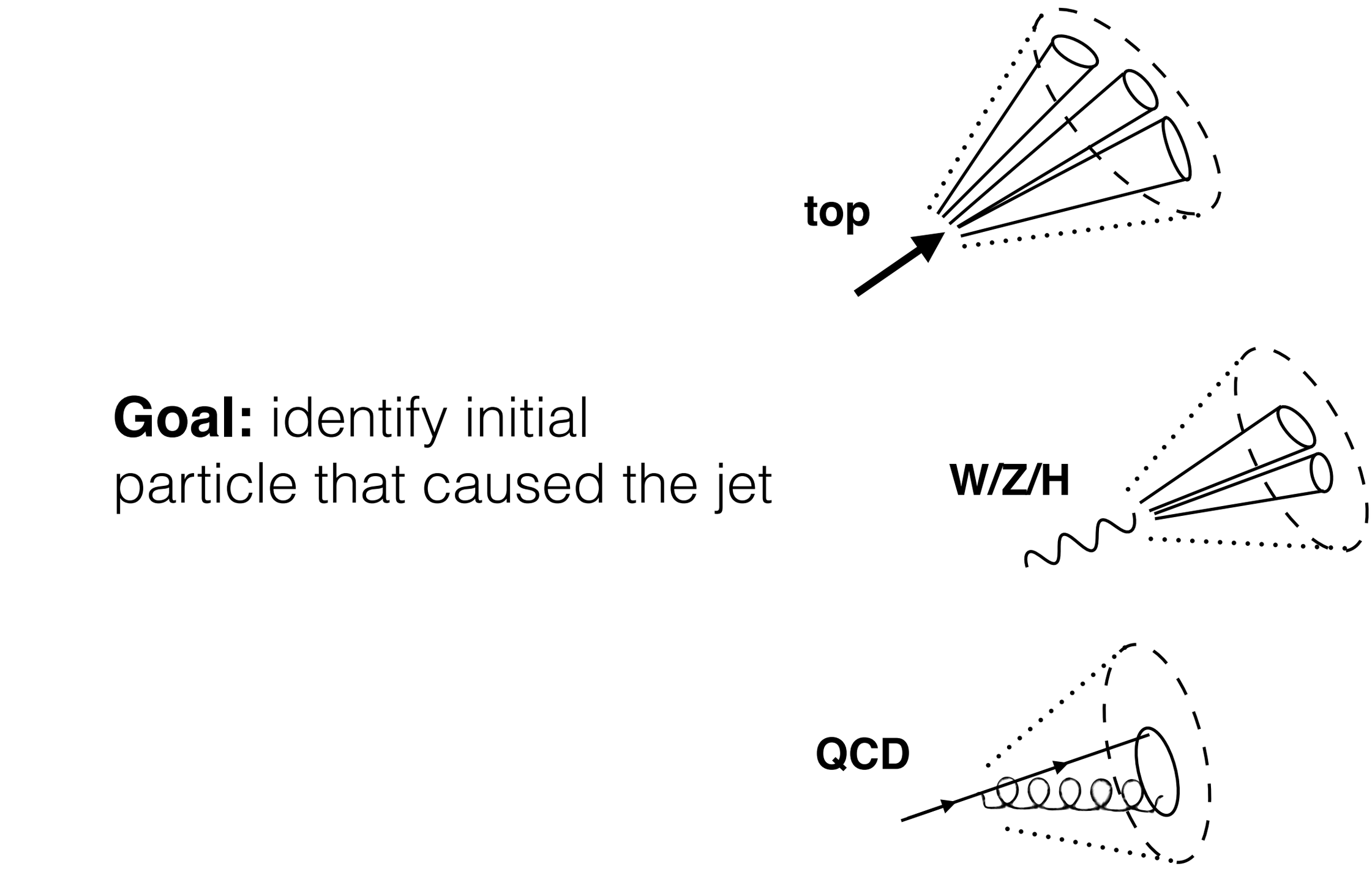


$p_T^{Z'}/m_{Z'} > \sim 1$

Large radius jet

(boosted) Jet Tagging

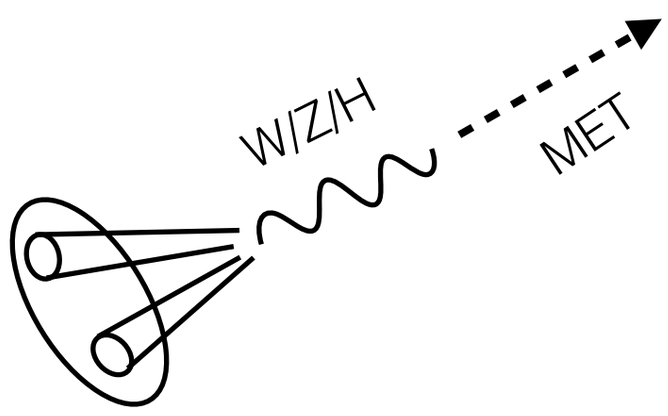
Goal: identify initial particle that caused the jet



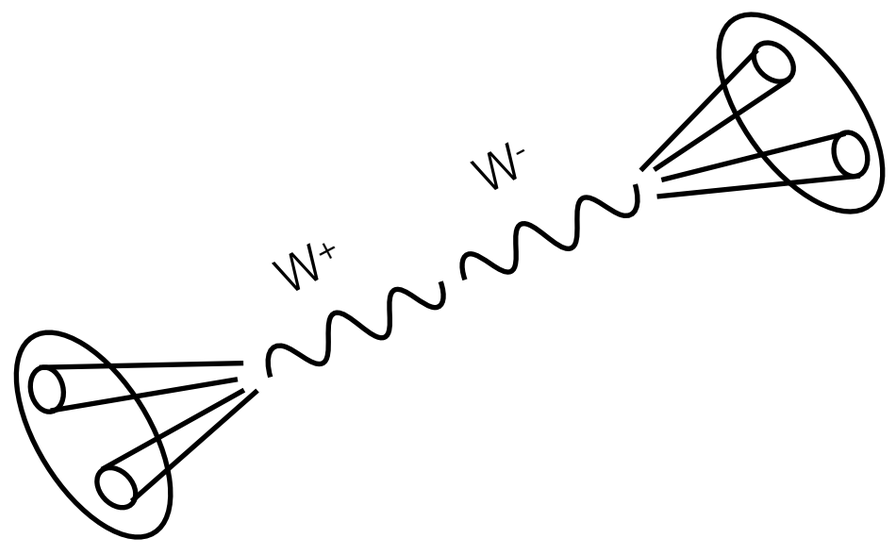
Analysis Applications

Generally want to enhance signal w/
known objects over QCD background:

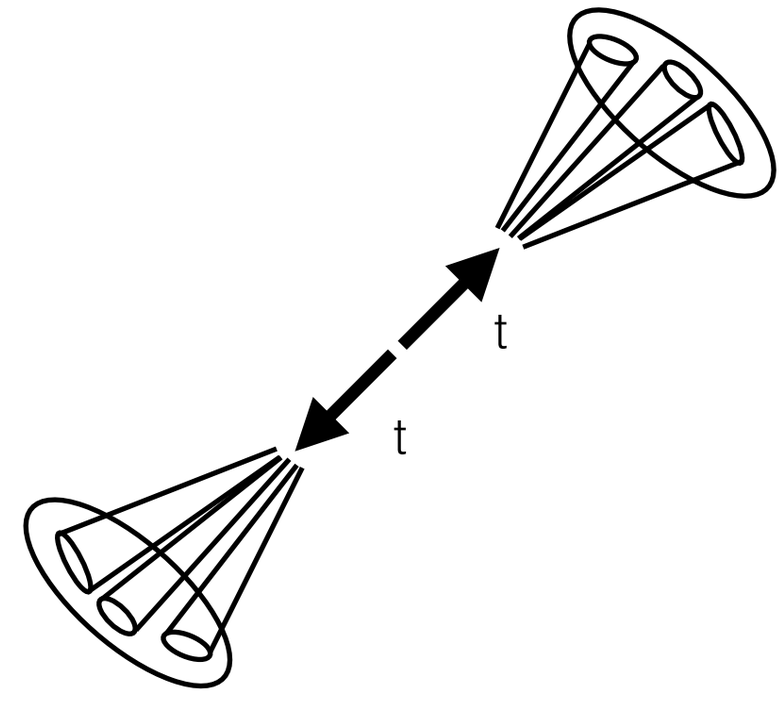
"Mono-X"



VV resonance



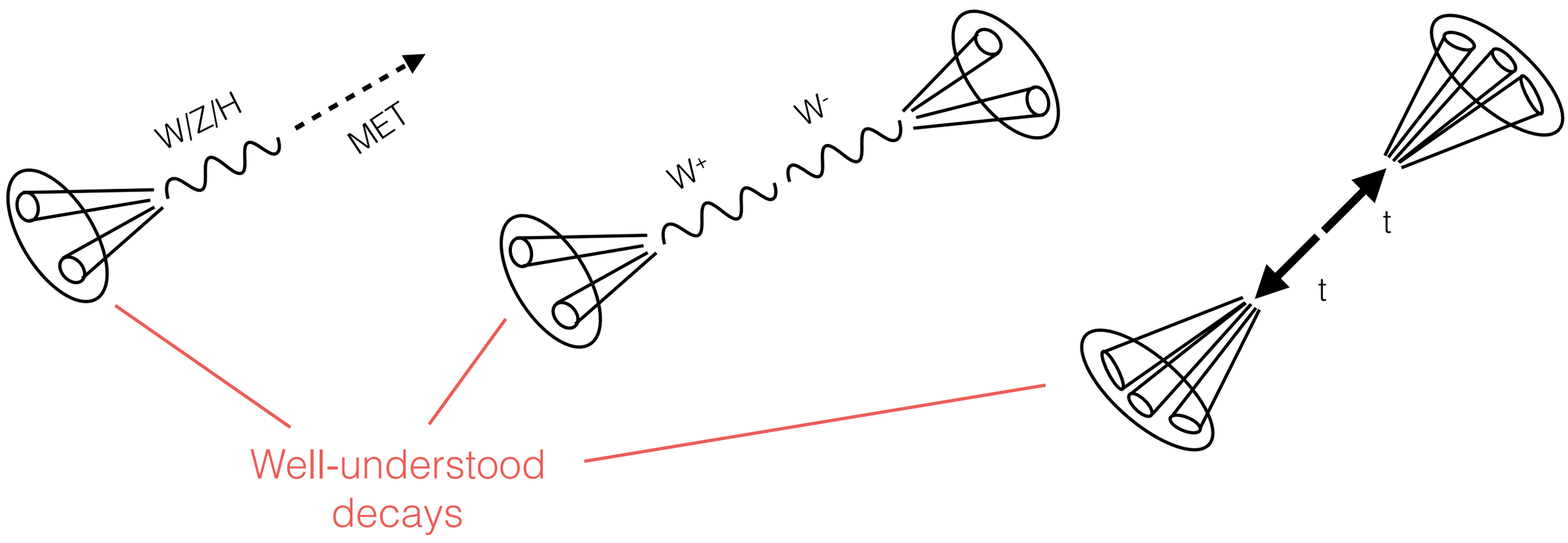
(heavy) Z'→tt



Analysis Applications

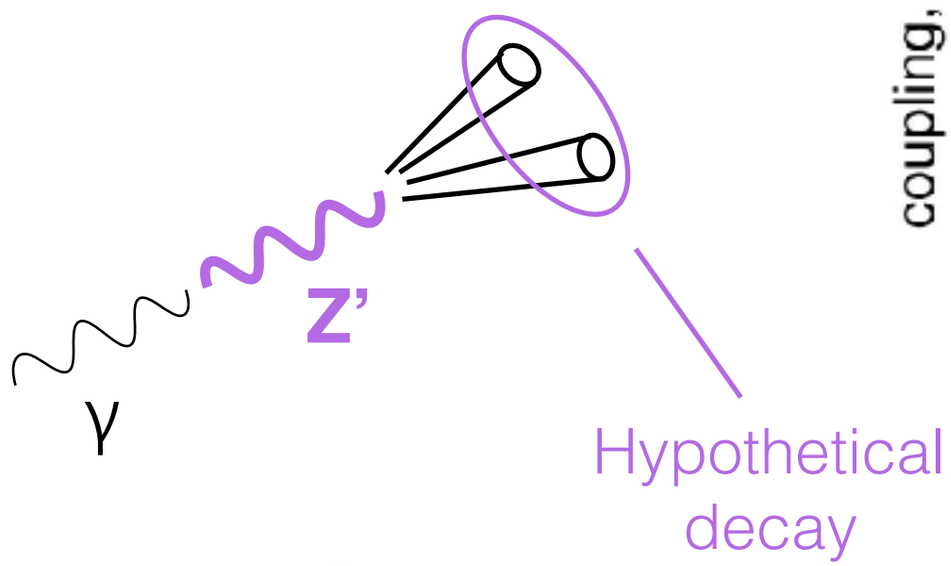
Generally want to enhance signal w/
known objects over QCD background:

“Mono-X” **VV resonance** **(heavy) Z’→tt**



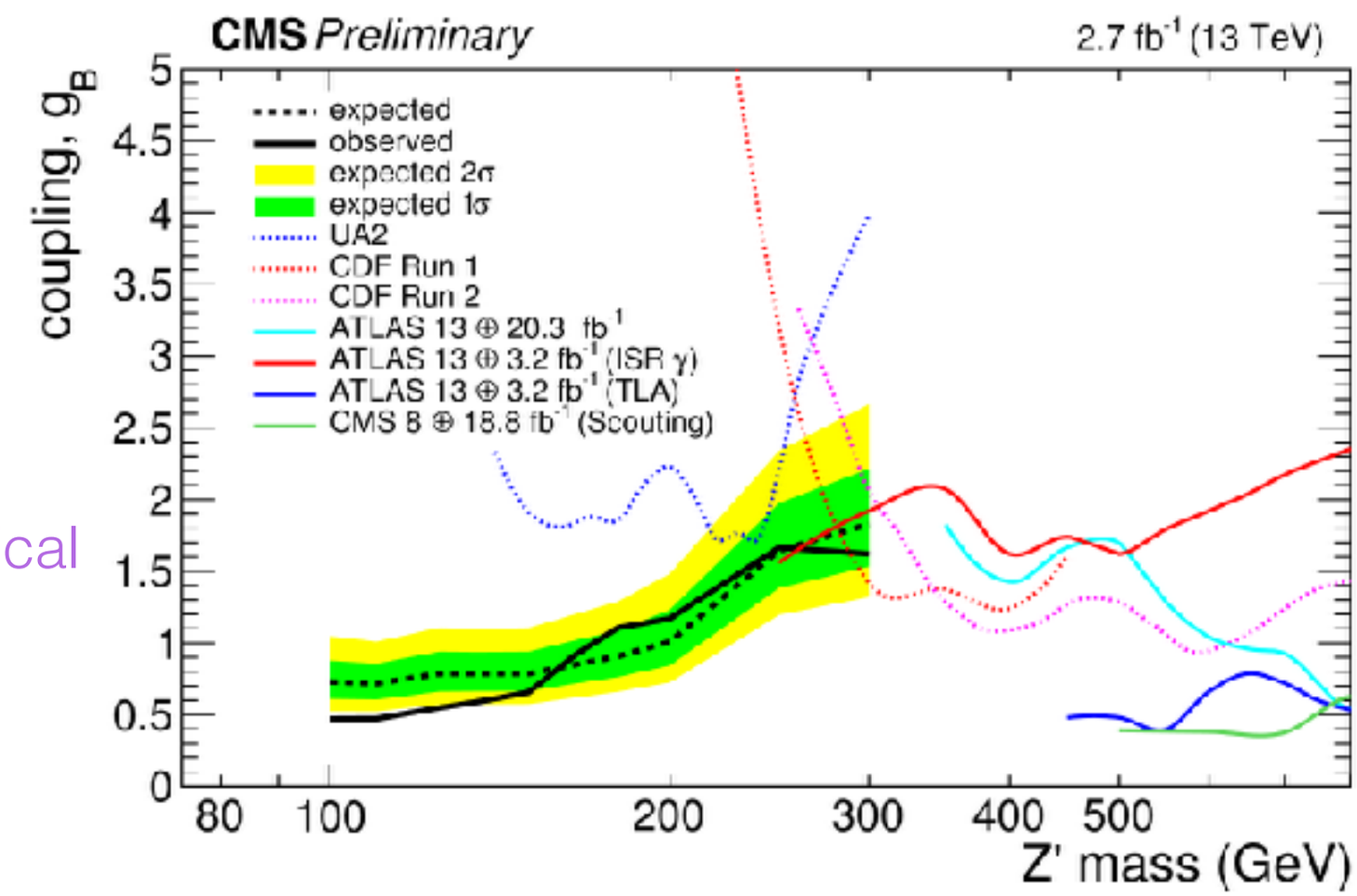
Analysis Applications

Low-mass leptophobic resonance



$p_T^{\gamma} \sim 150$ GeV

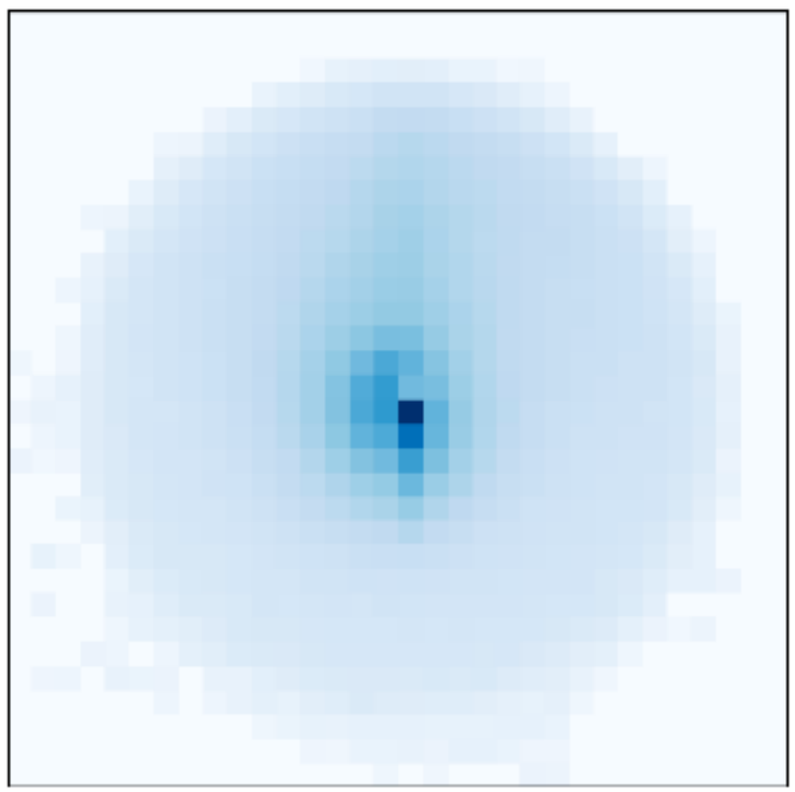
$m_{Z'} < \sim 200$ GeV



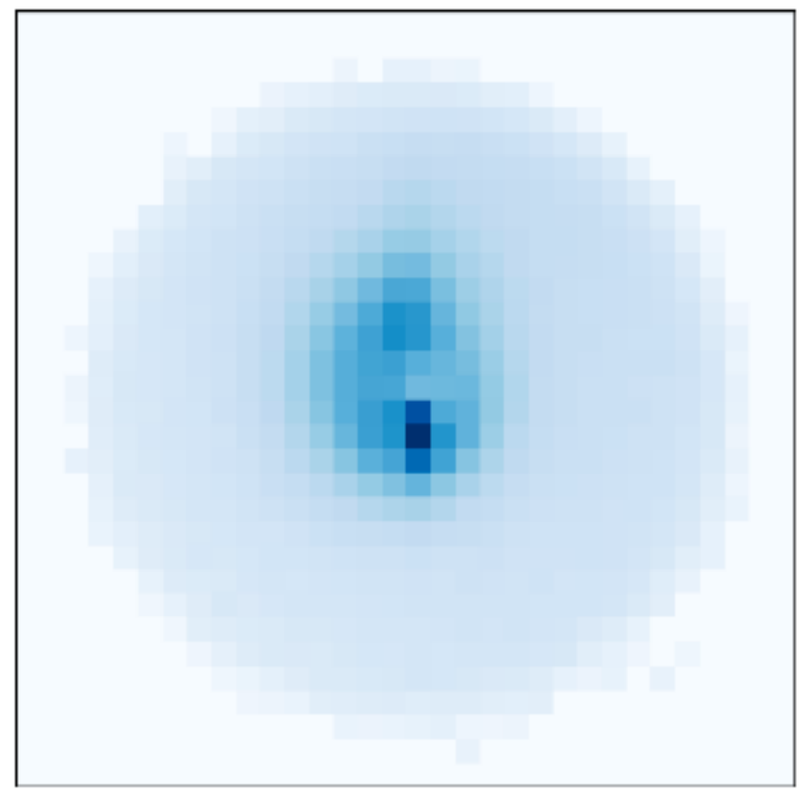
Jet Substructure

In addition to resonance, boosted jets have distinctive structure:

QCD jet



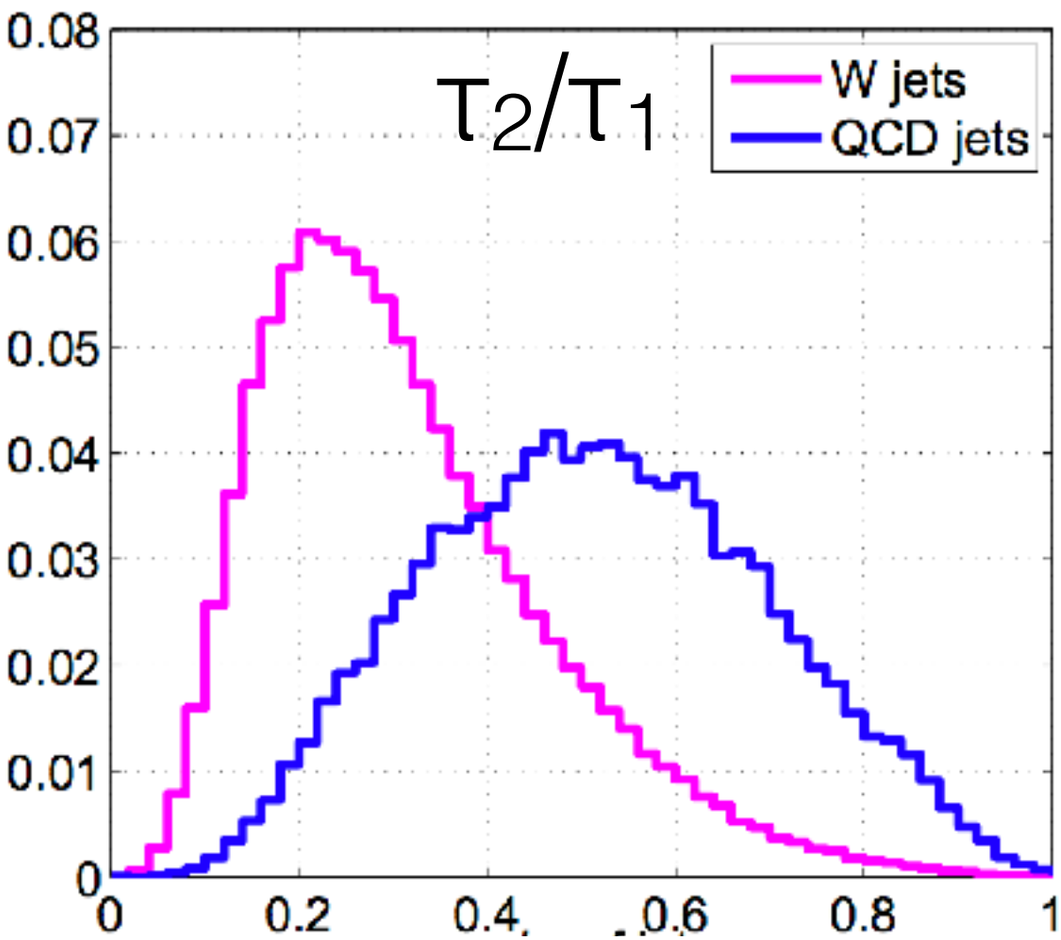
W jet

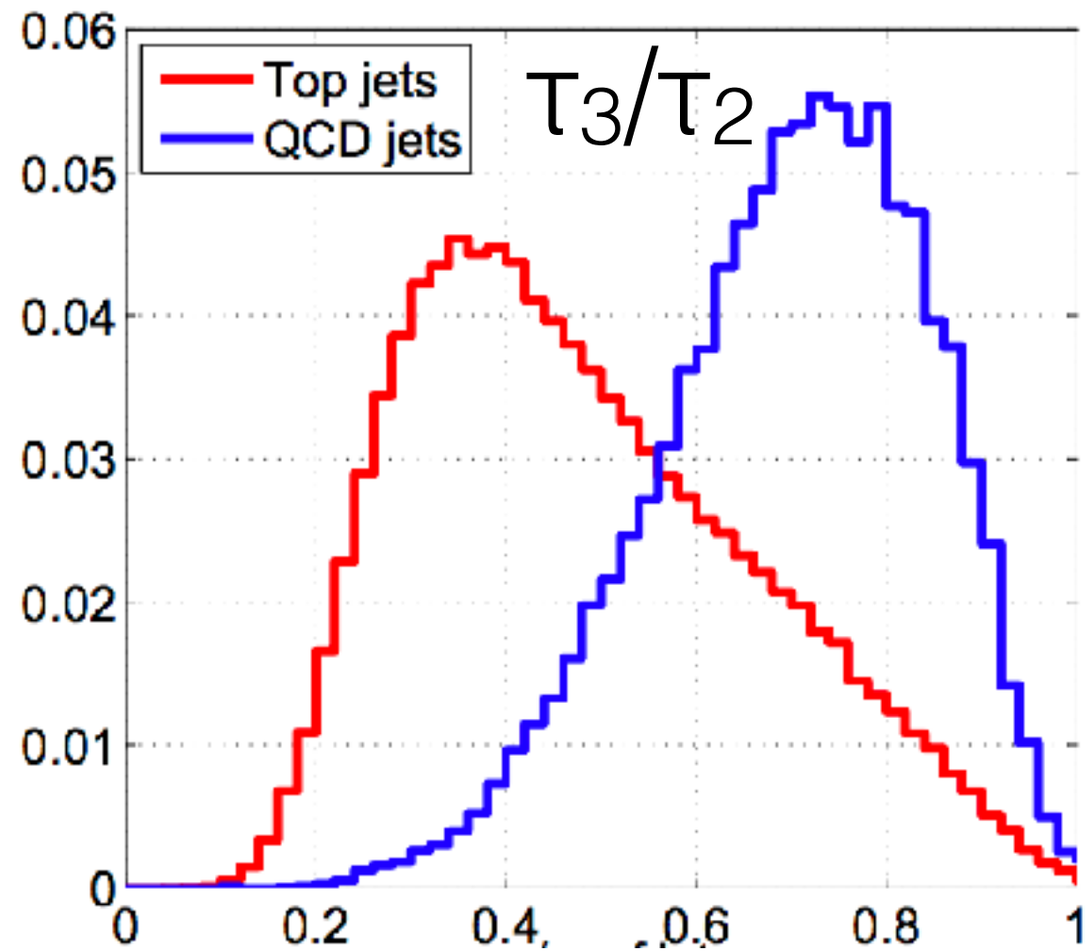


arXiv:1603.09349

Substructure Variables

- Many theoretically motivated tools to quantify jet substructure, e.g. N-subjettiness, ECF…

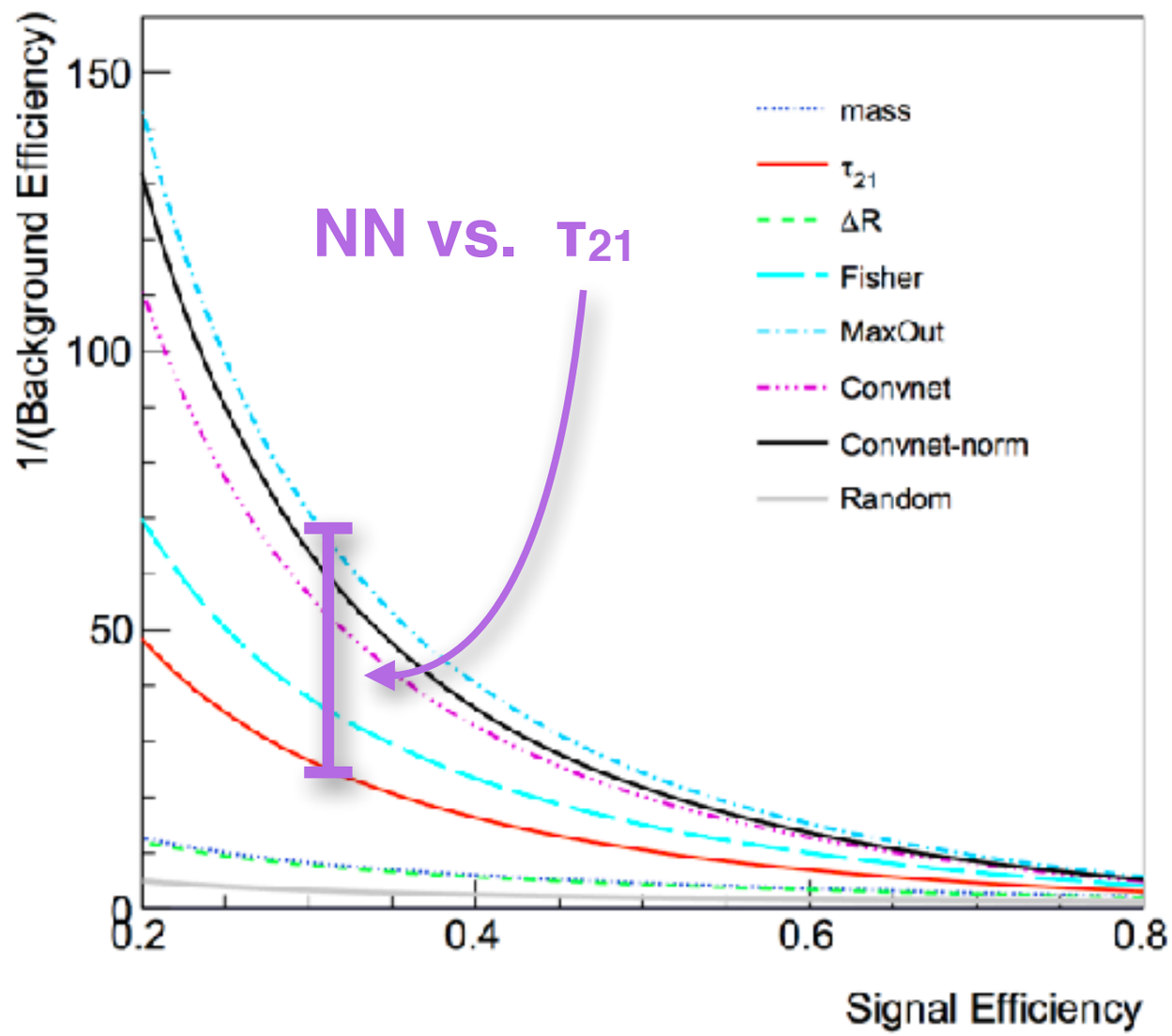




arXiv:1011.2268

Multivariate Taggers

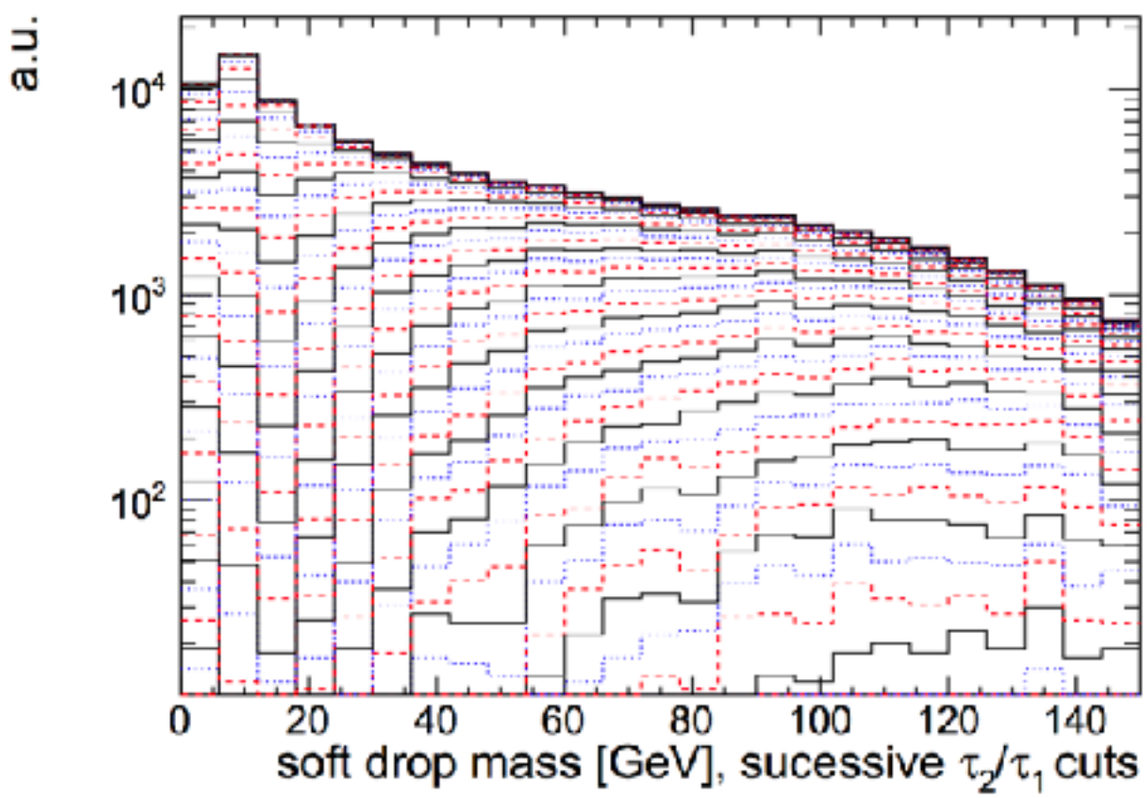
Multivariate taggers (BDT, NN) in general can do even better!



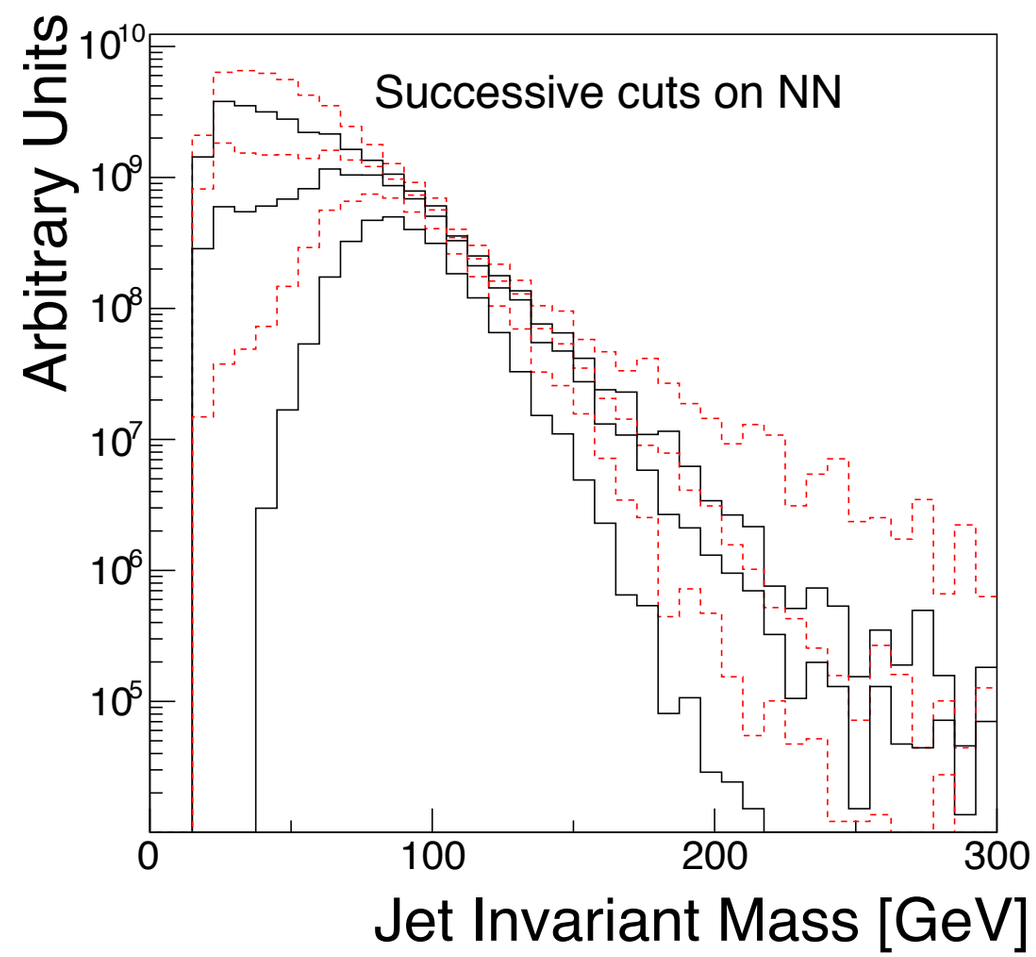
arXiv:1511.05190

Mass Correlation

But… cutting on taggers **distorts mass spectrum**



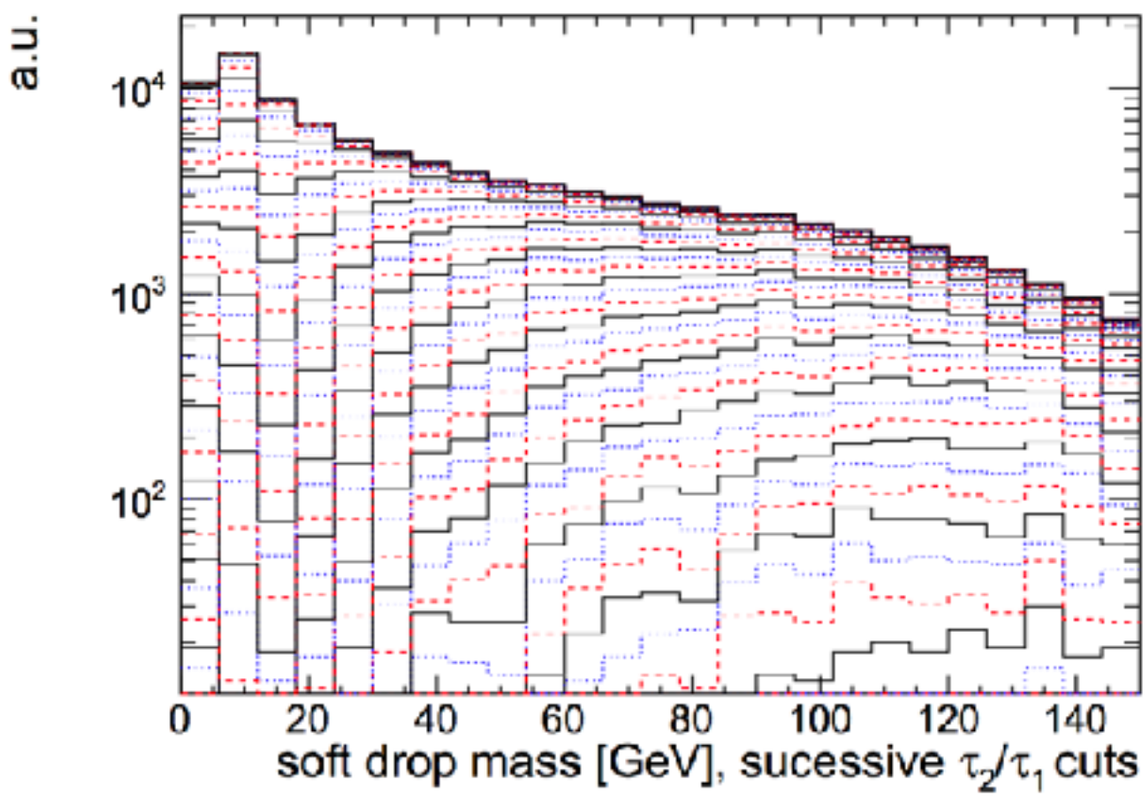
arXiv:1603.00027



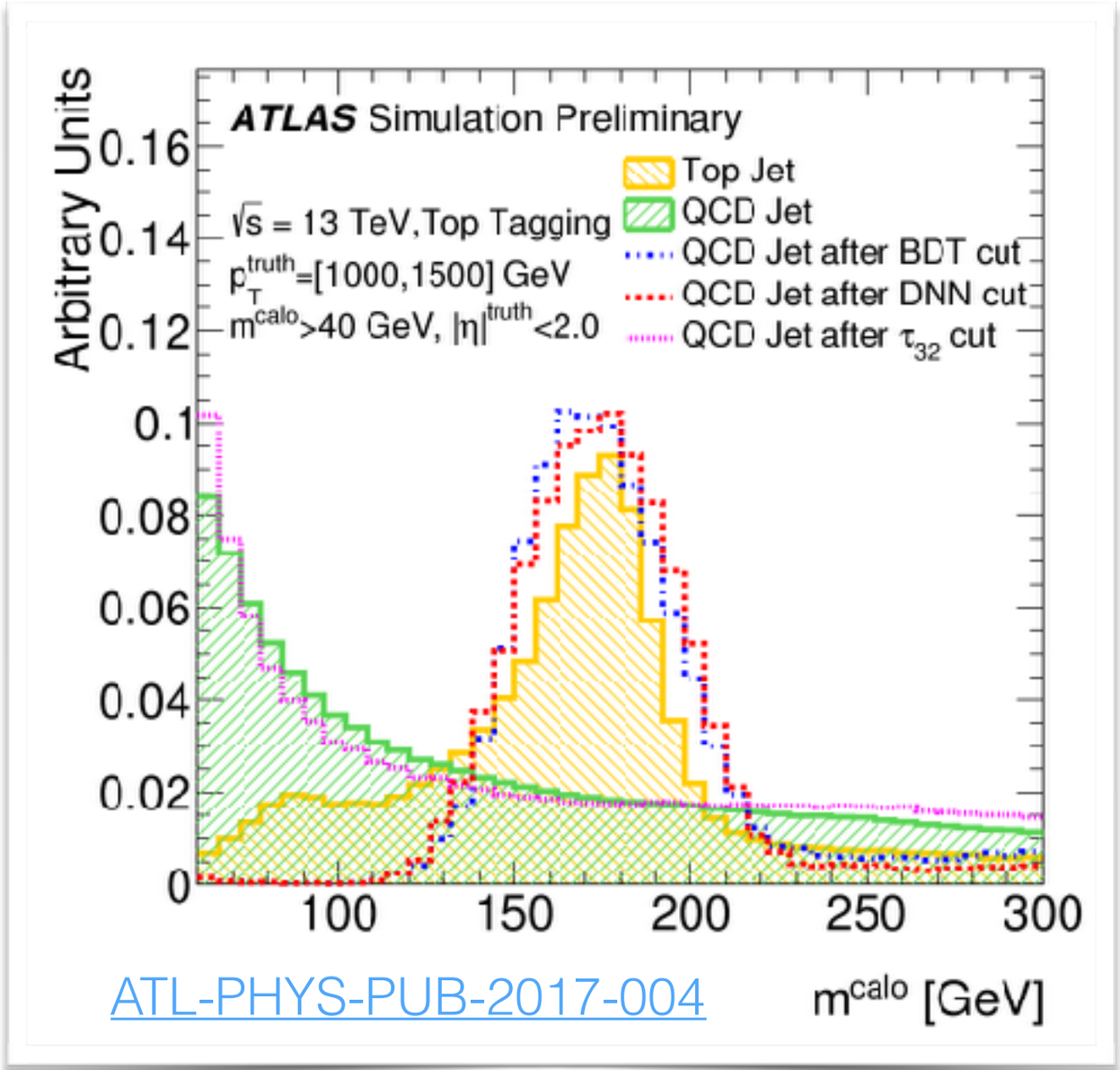
arXiv:1703.03507

Mass Correlation

But: cutting on taggers **distorts mass spectrum**



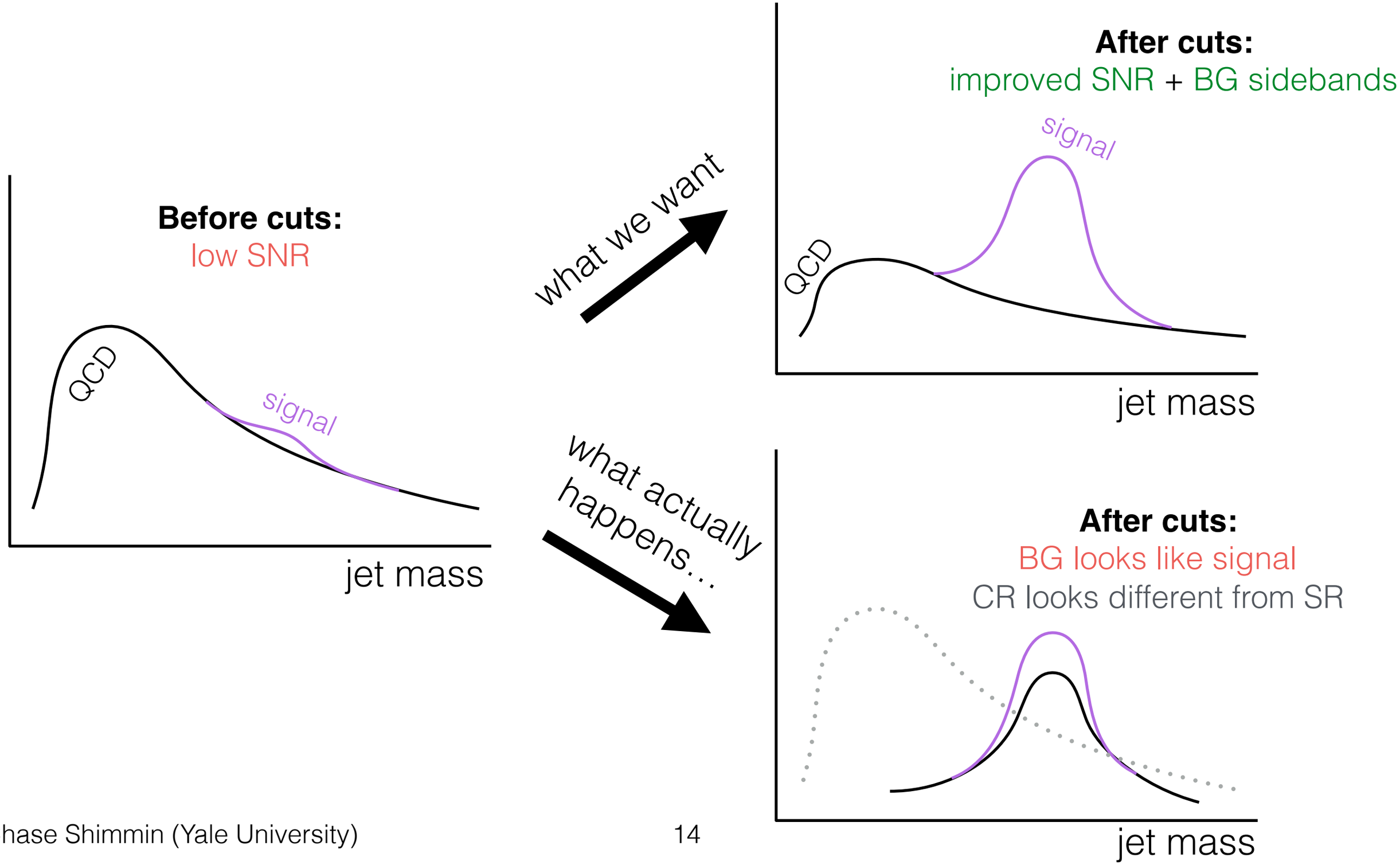
arXiv:1603.00027



ATL-PHYS-PUB-2017-004

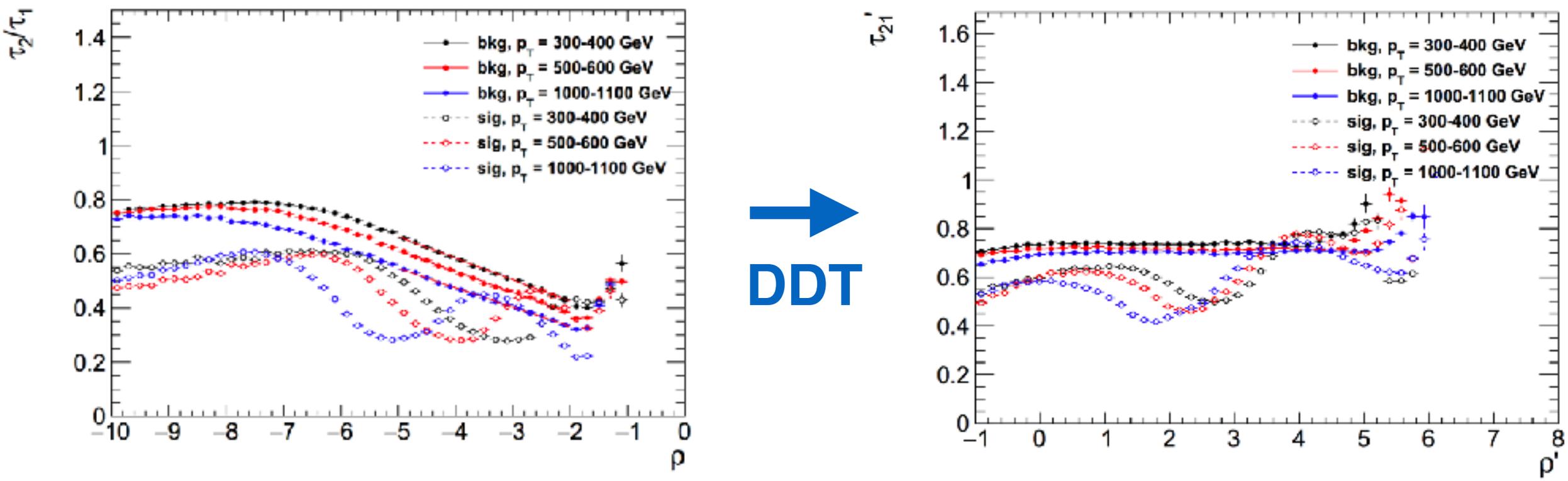
Mass Correlation

Correlation with the observable of interest is bad!



De-Correlation

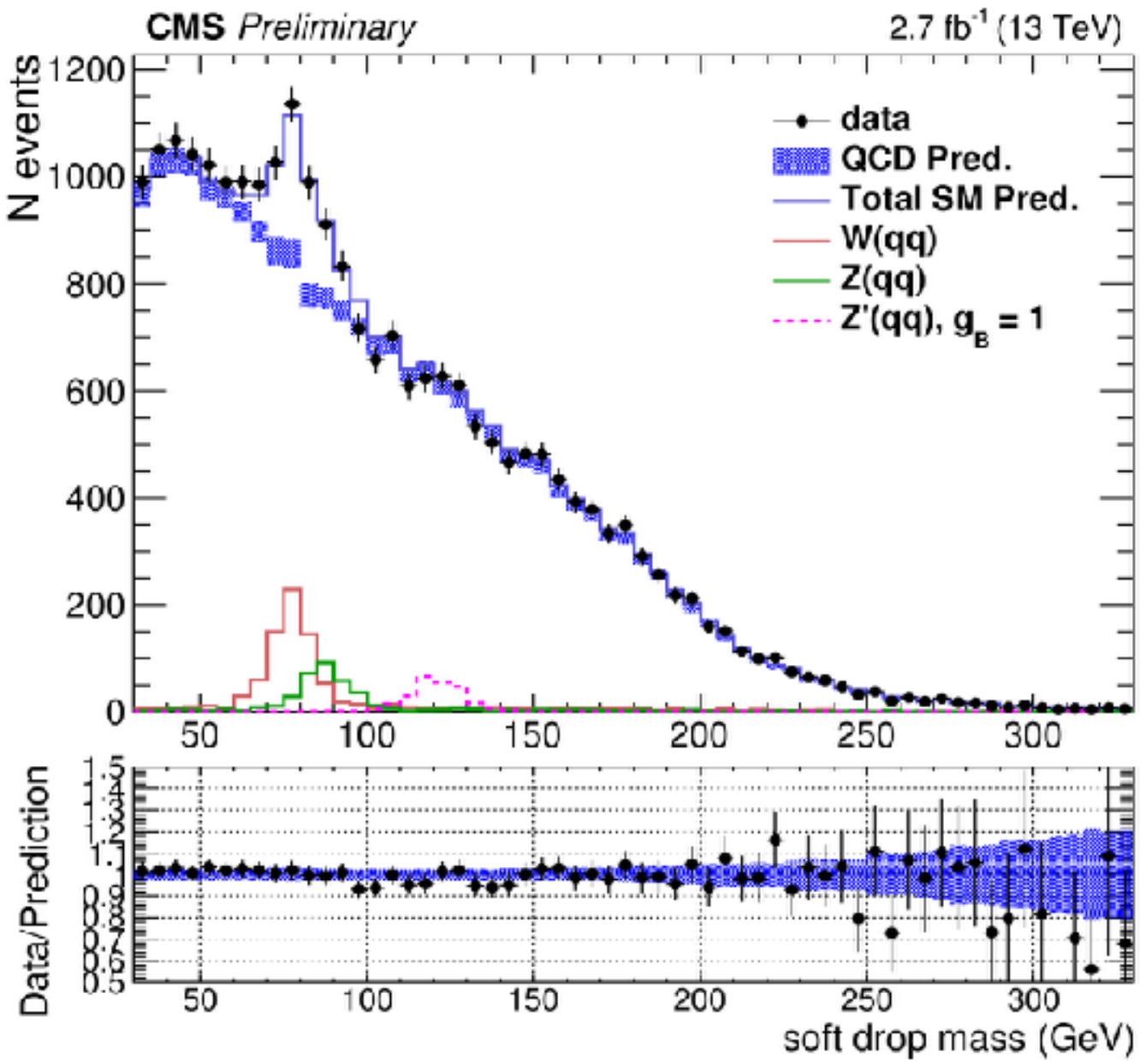
- "DDT" paper proposes explicit transformation to decorrelate τ_{21} variable



arXiv:1603.00027

De-Correlation

- DDT method was used by CMS in low-mass Z' search



CMS-PAS-EXO-16-030

Generalization

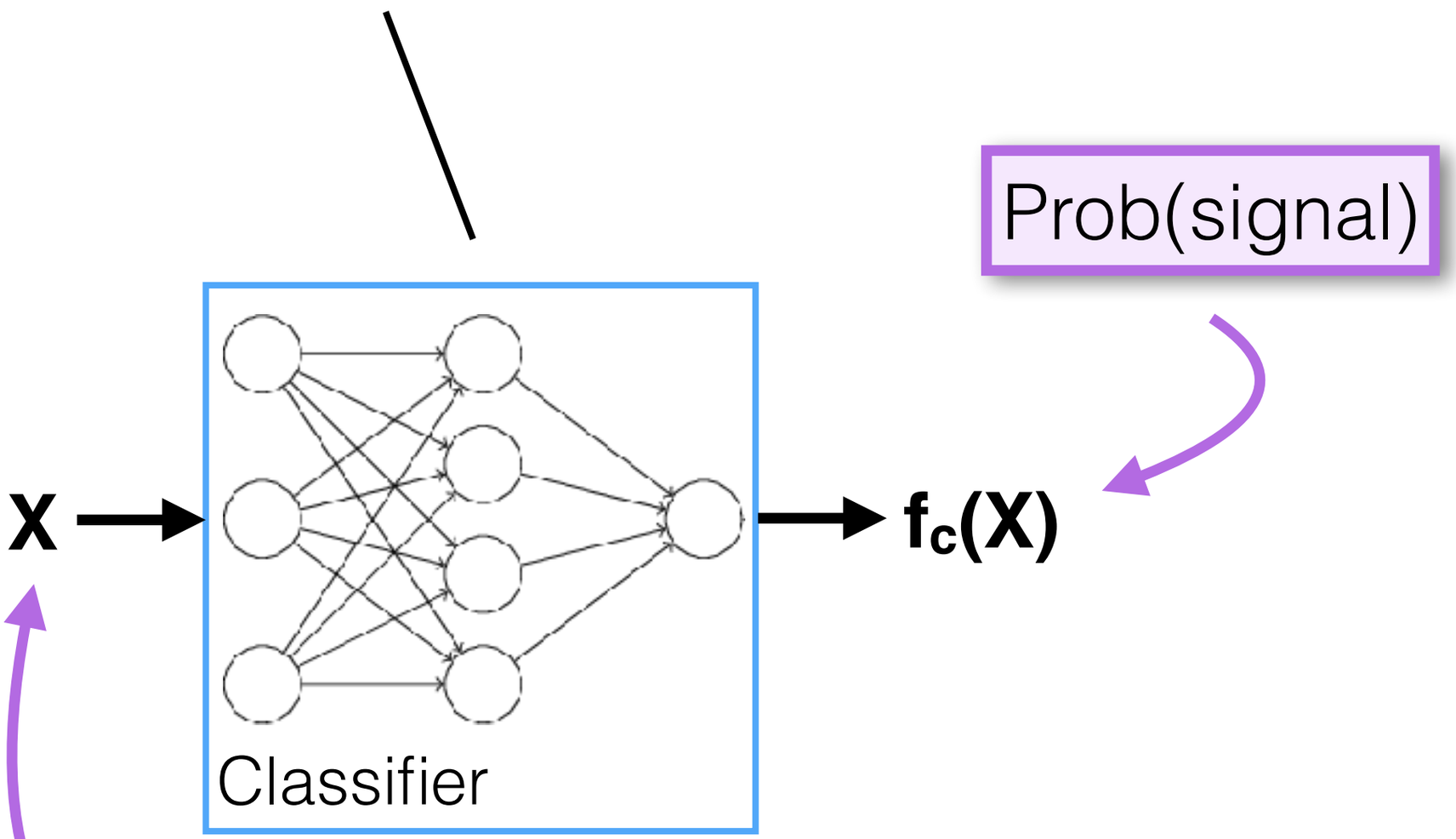
- We would like to **generalize** this decorrelation approach for arbitrary classifiers
- Some proposed approaches:
  - multivariate DDT via PCA arXiv:1603.00027
  - uGBoost: add loss to enforce “flatness” arXiv:1410.4140
  - ★ Adversarial “pivot” / domain adaptation: arXiv:1611.01046

We investigate this approach

Adversarial Decorrelation

Basic idea:

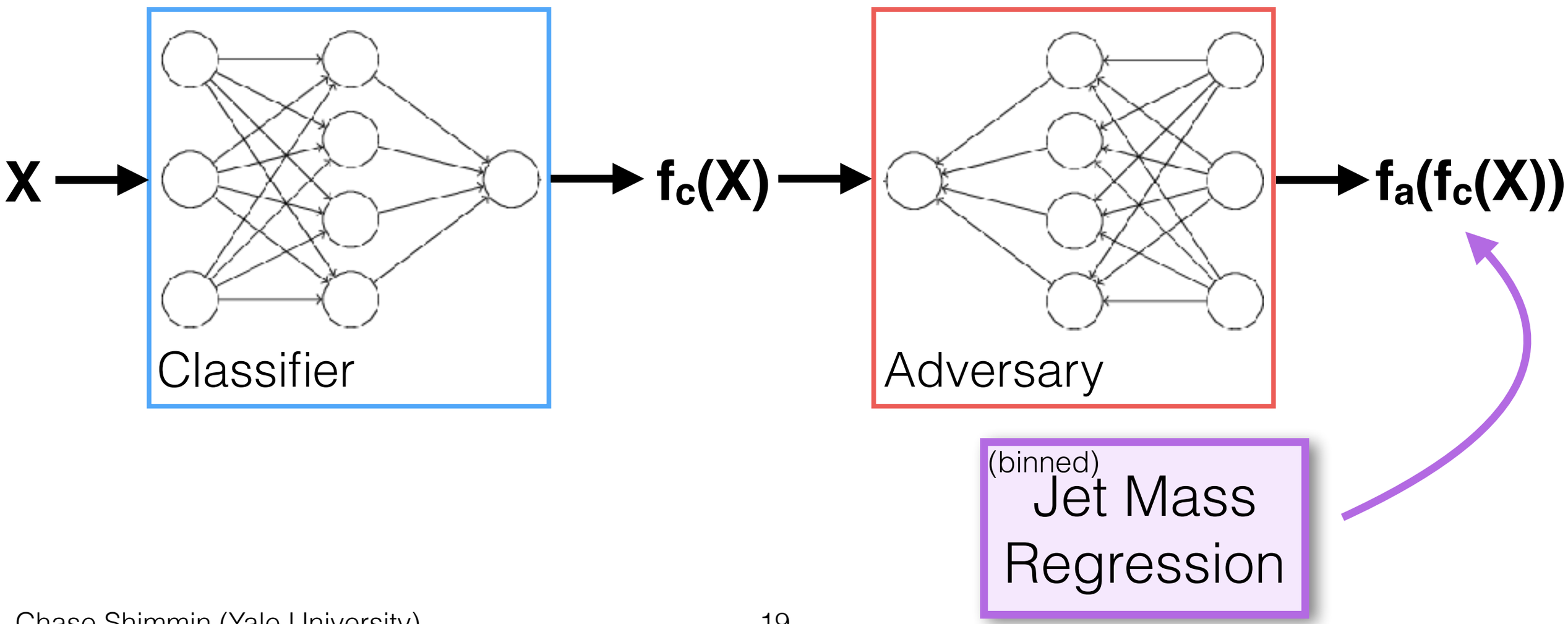
Classifier is trained to identify signal jets



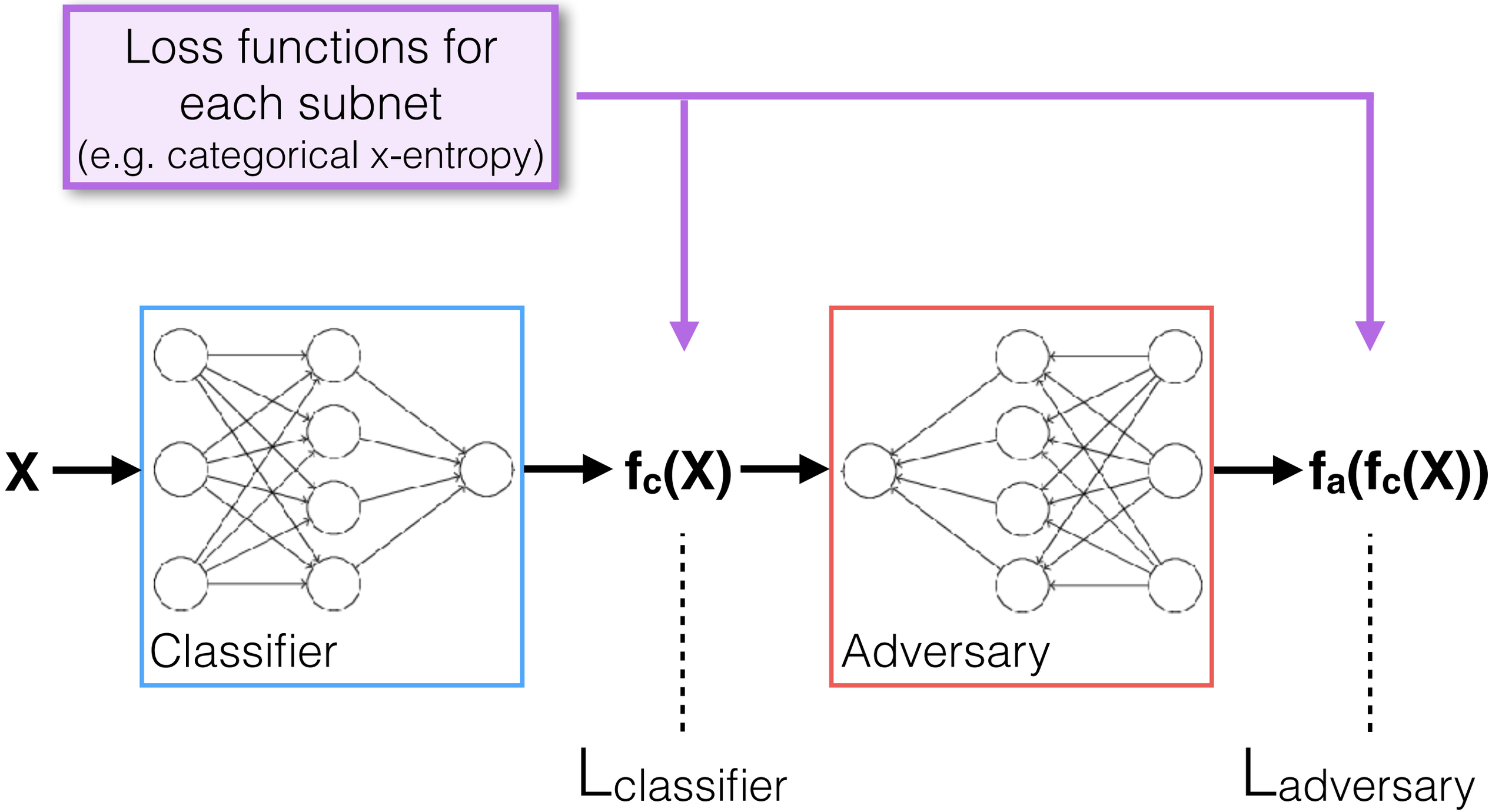
Jet features
(p_T, m, τ_{21}, …)

Adversarial Decorrelation

Adversary is trained to predict jet mass



Adversarial Decorrelation



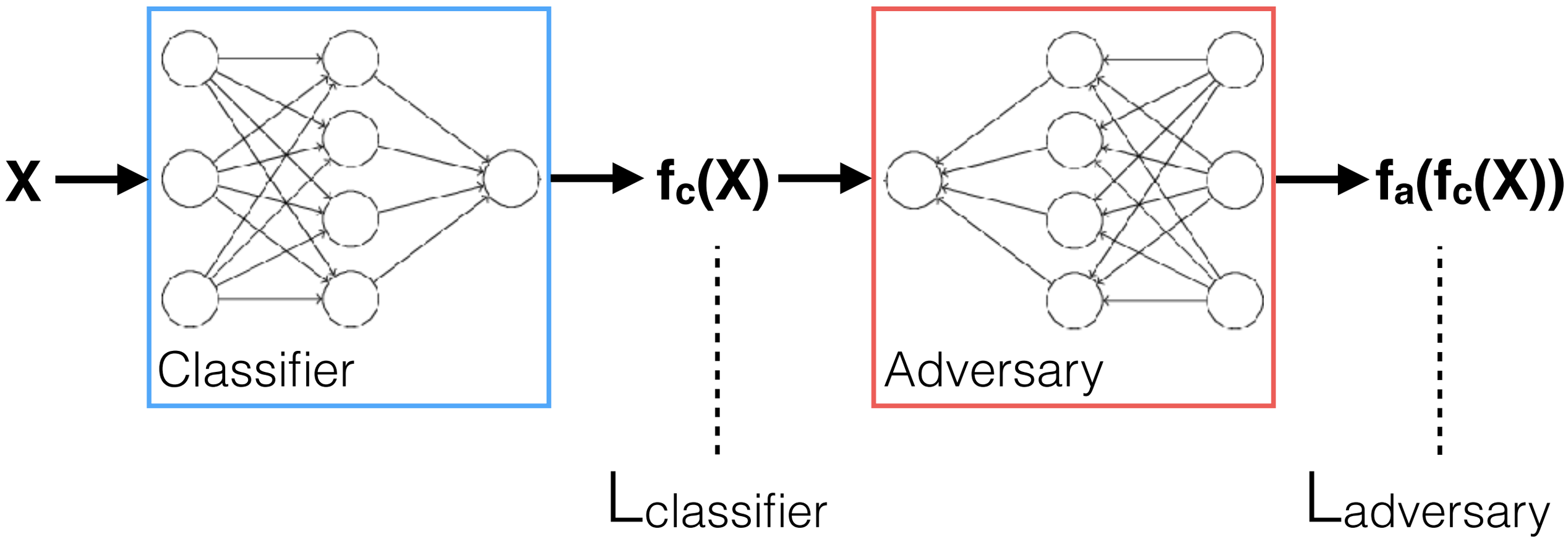
Adversarial Decorrelation

Simultaneously minimize:

$L_{adversary}$

and

$L_{tagger} = L_{classifier} - \lambda L_{adversary}$



Adversarial Decorrelation

Simultaneously minimize:

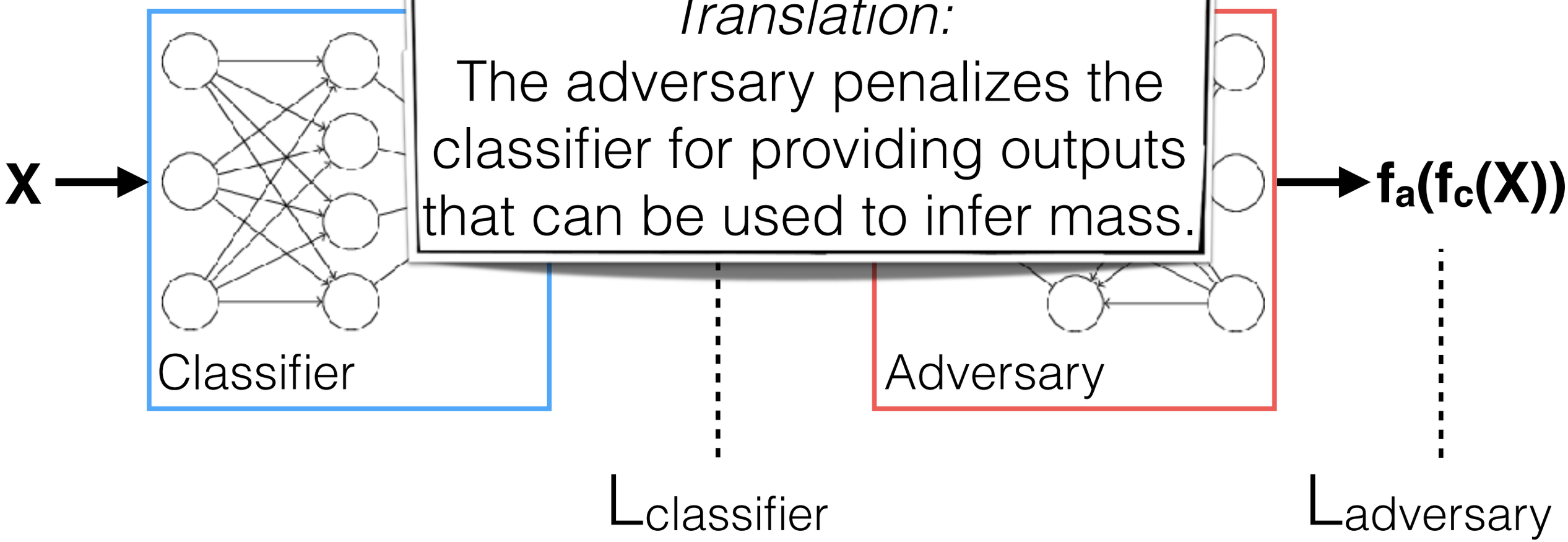
$L_{adversary}$

and

$L_{tagger} = L_{classifier} - \lambda L_{adversary}$

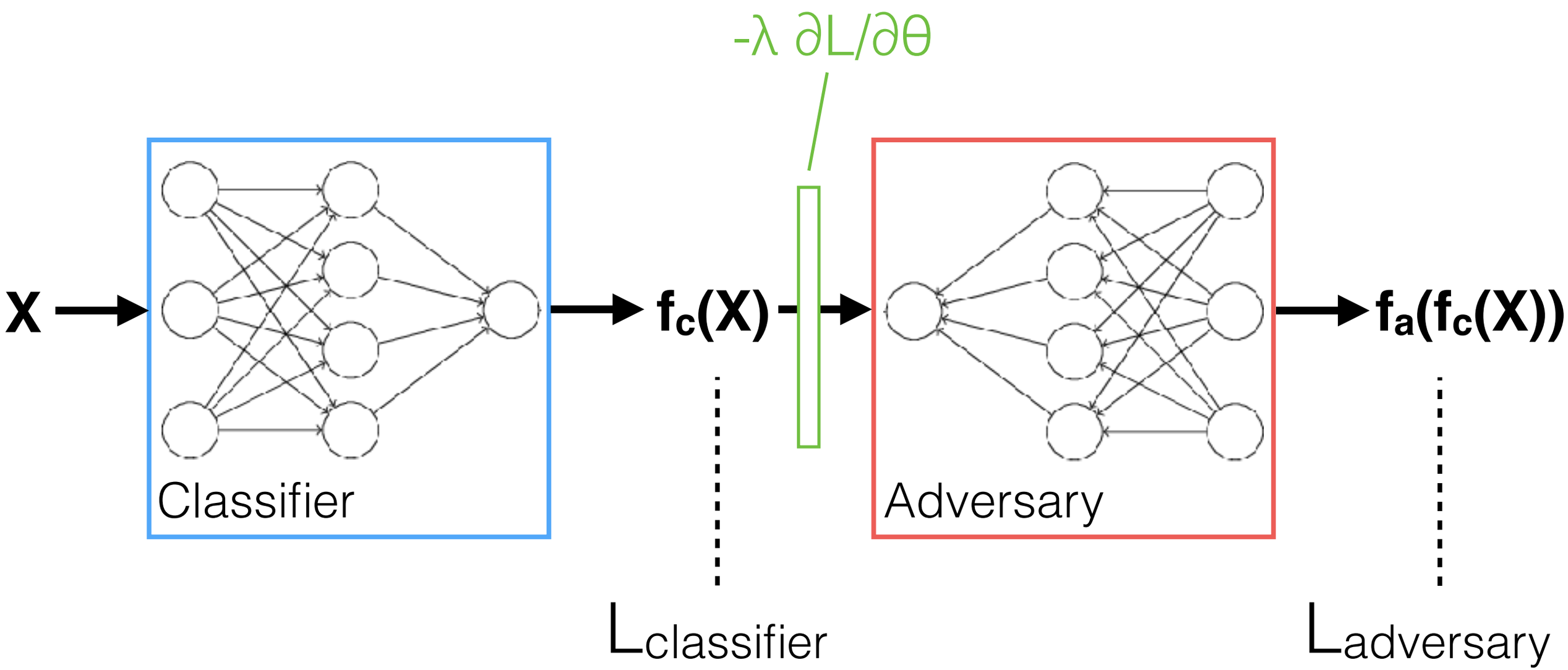
Translation:

The adversary penalizes the classifier for providing outputs that can be used to infer mass.



Training

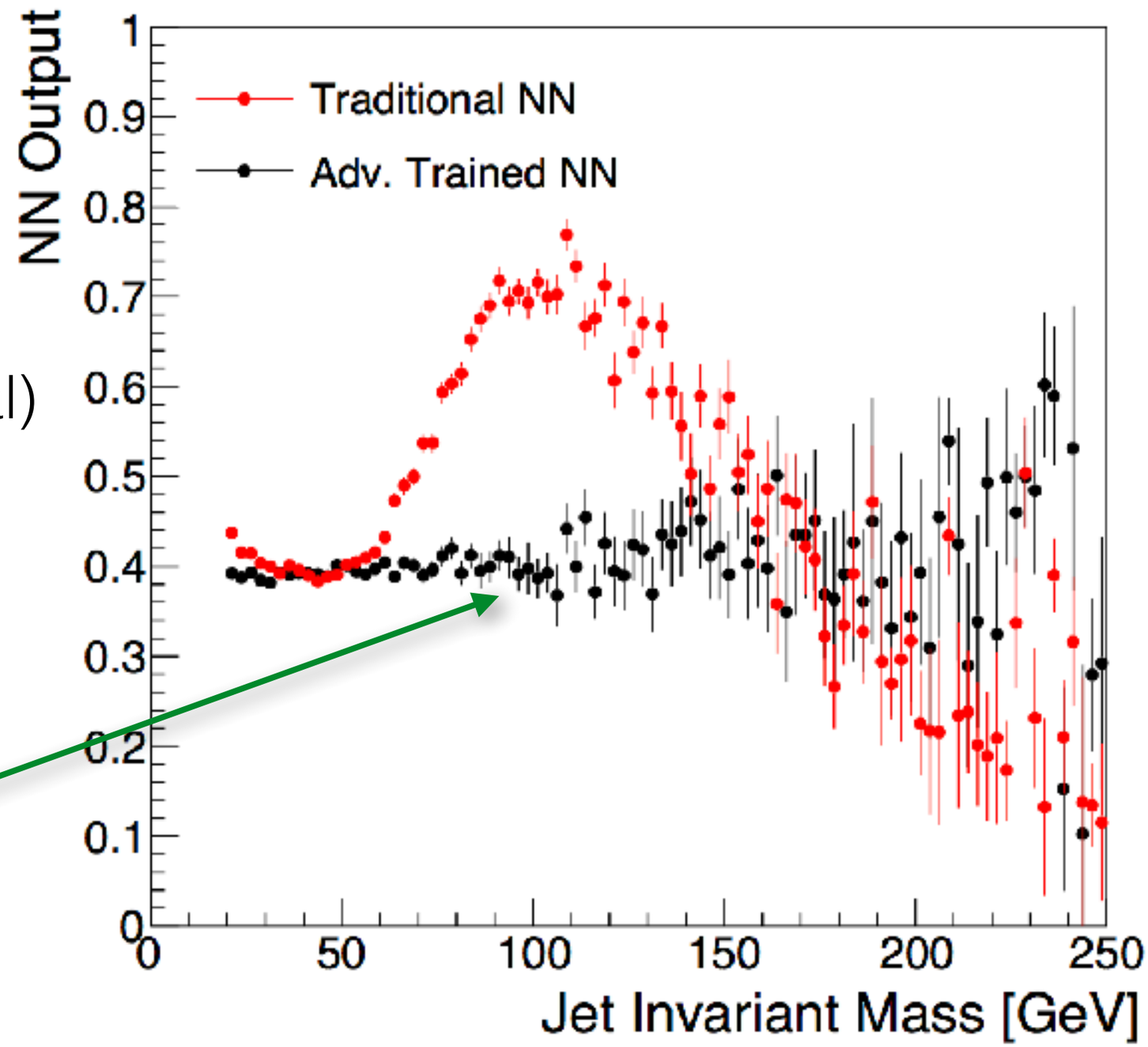
- Simultaneous optimization achieved with **gradient scaling layer**
- Signal events are given zero weight in adversary loss



Results

Training on ~200k MC events:

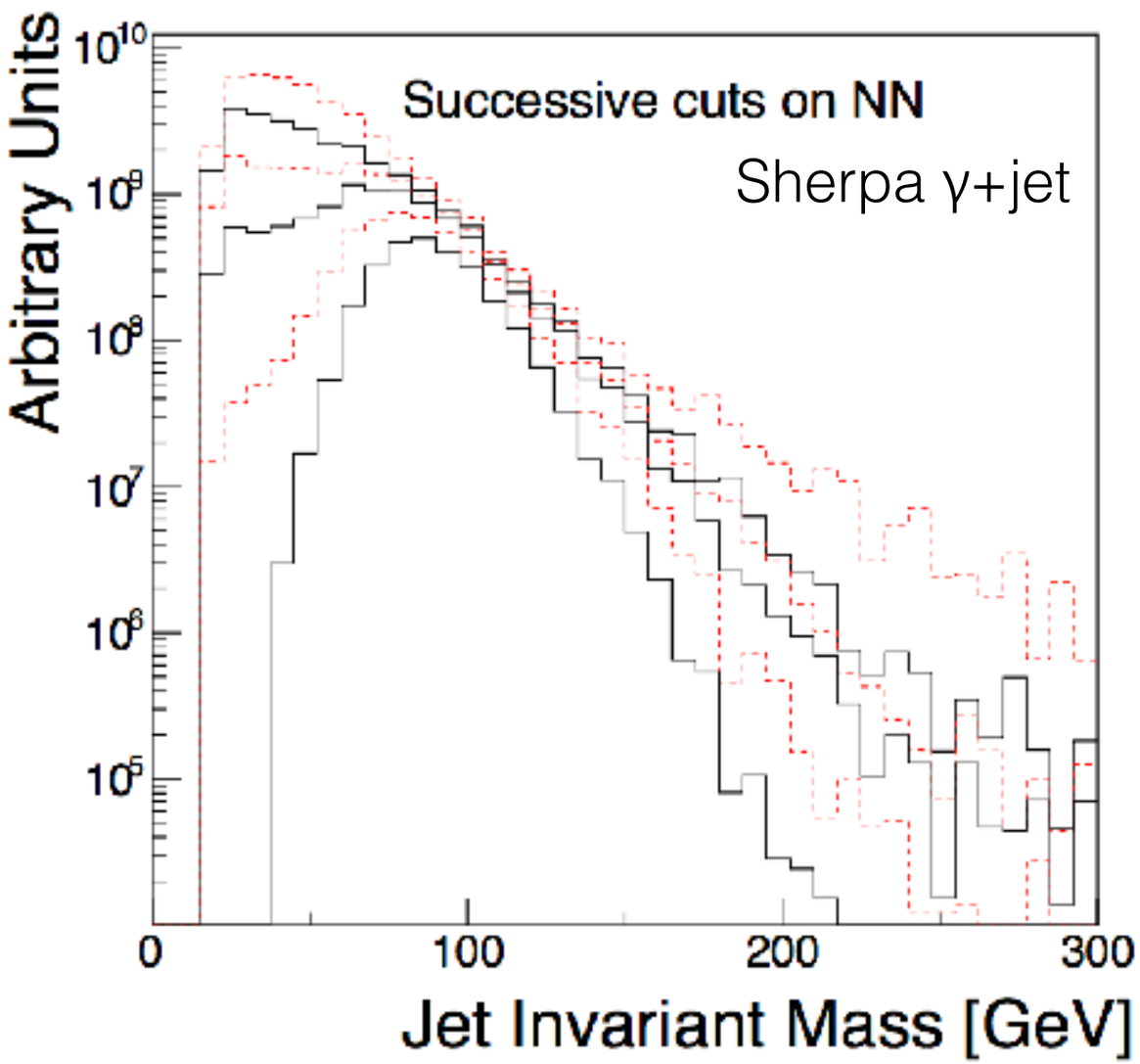
Sherpa γ+jet (BG)
MG5 γ+Z' (Signal)
Pythia + Delphes (Both)

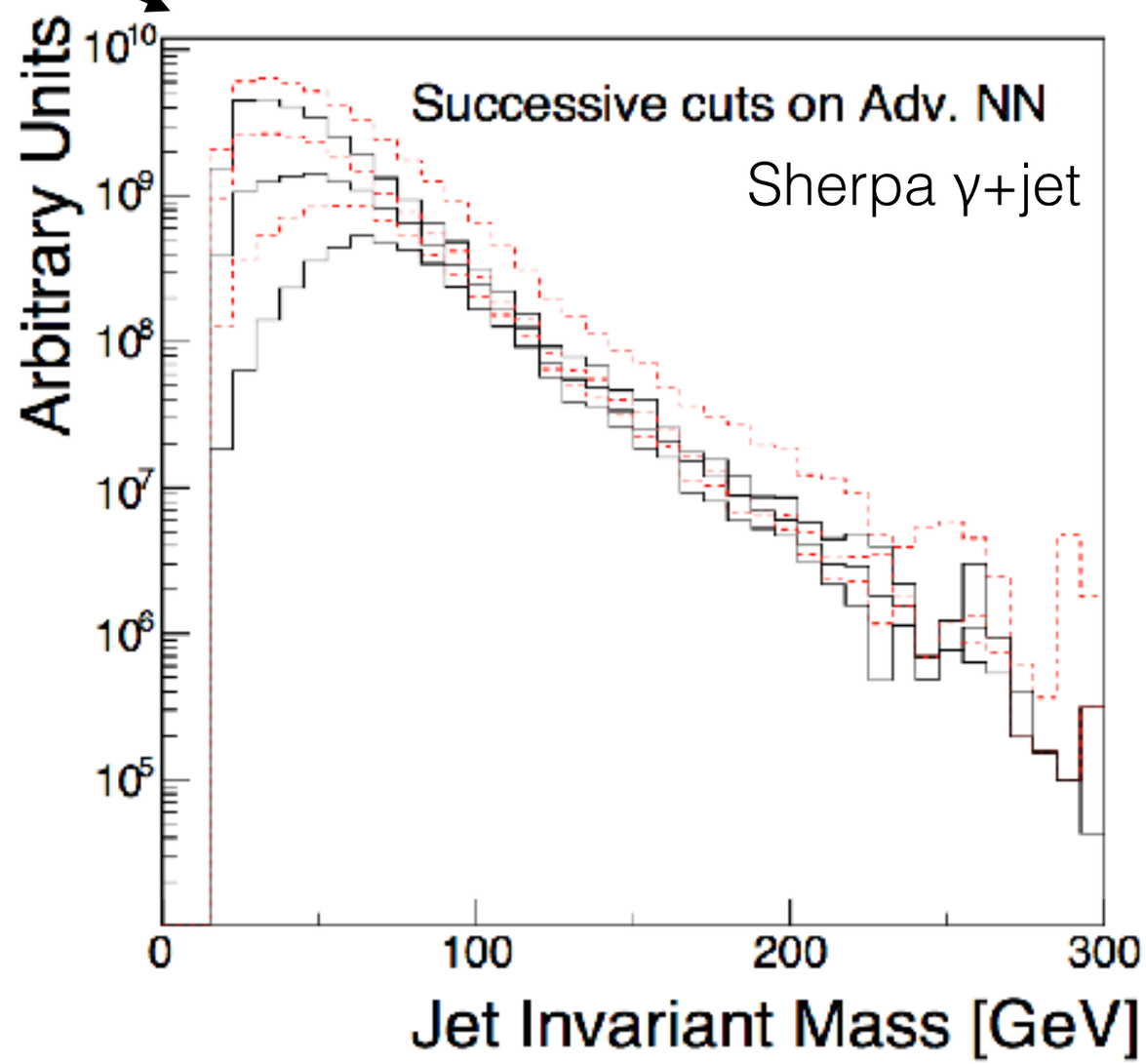


✓ Tagger profile much flatter

Results

✓ BG distortion considerably reduced

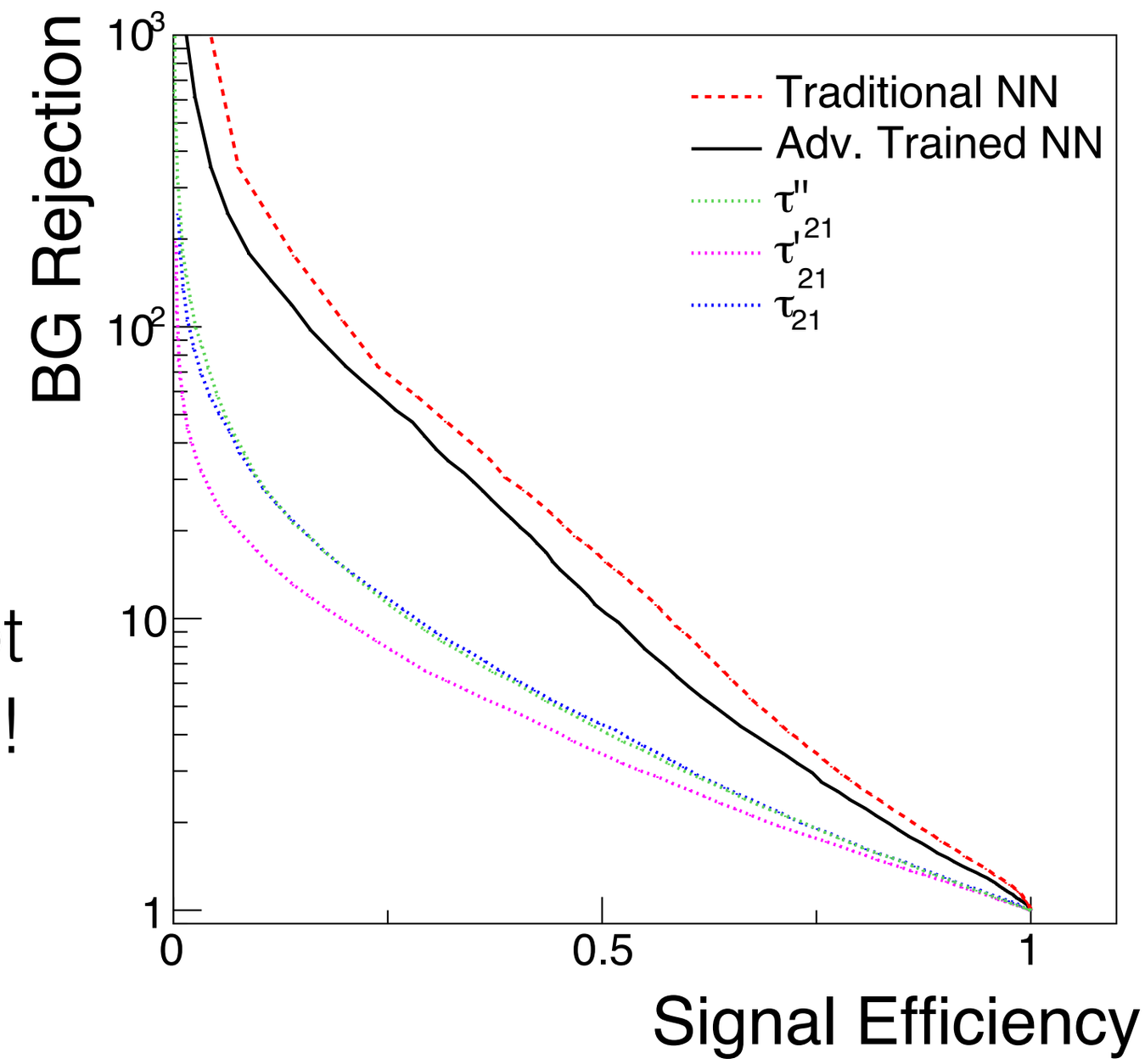




ROC Performance

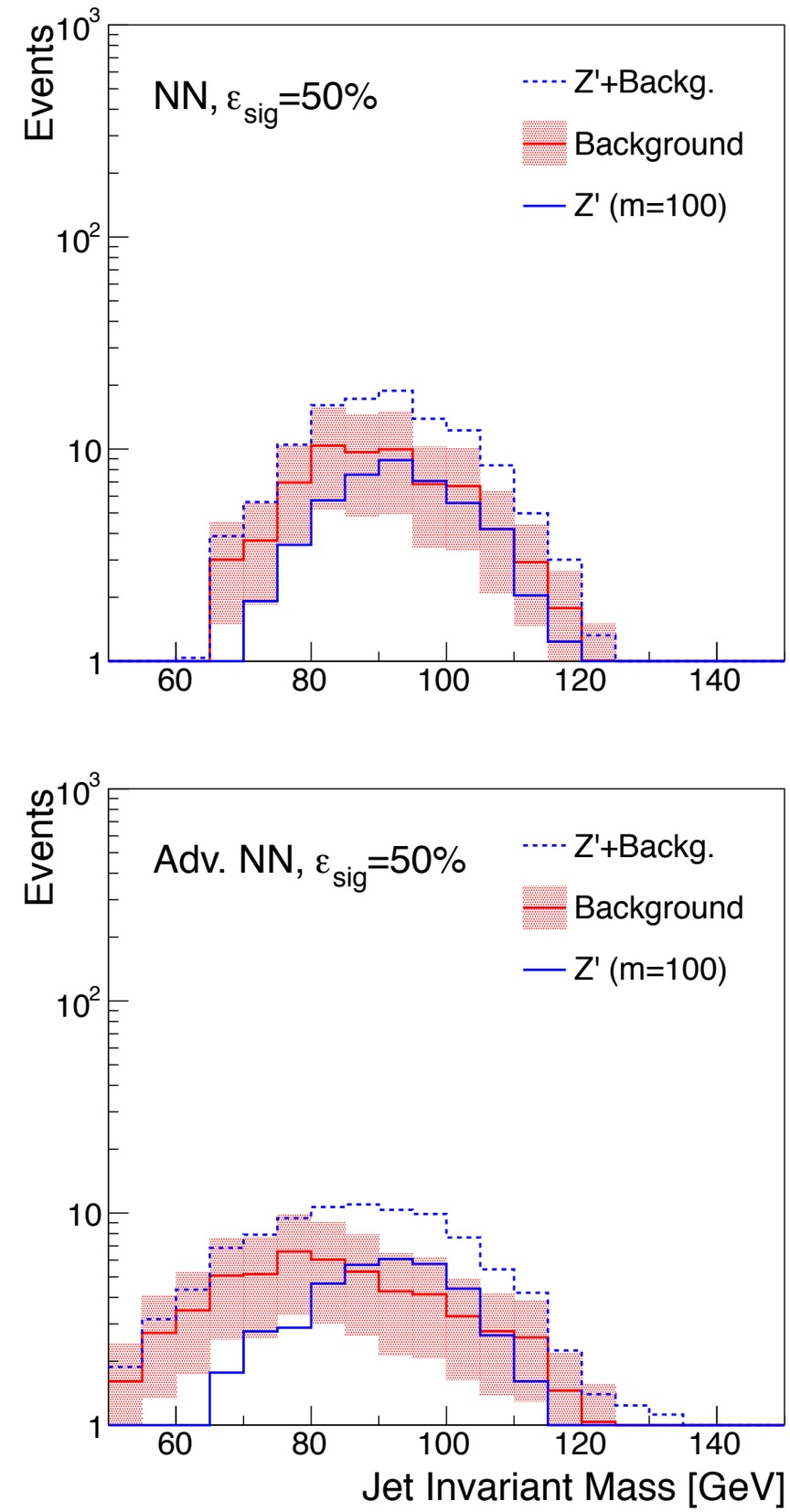
Adversarial method:
slightly lower AUC

… however this is not our figure of merit!



Statistical Significance

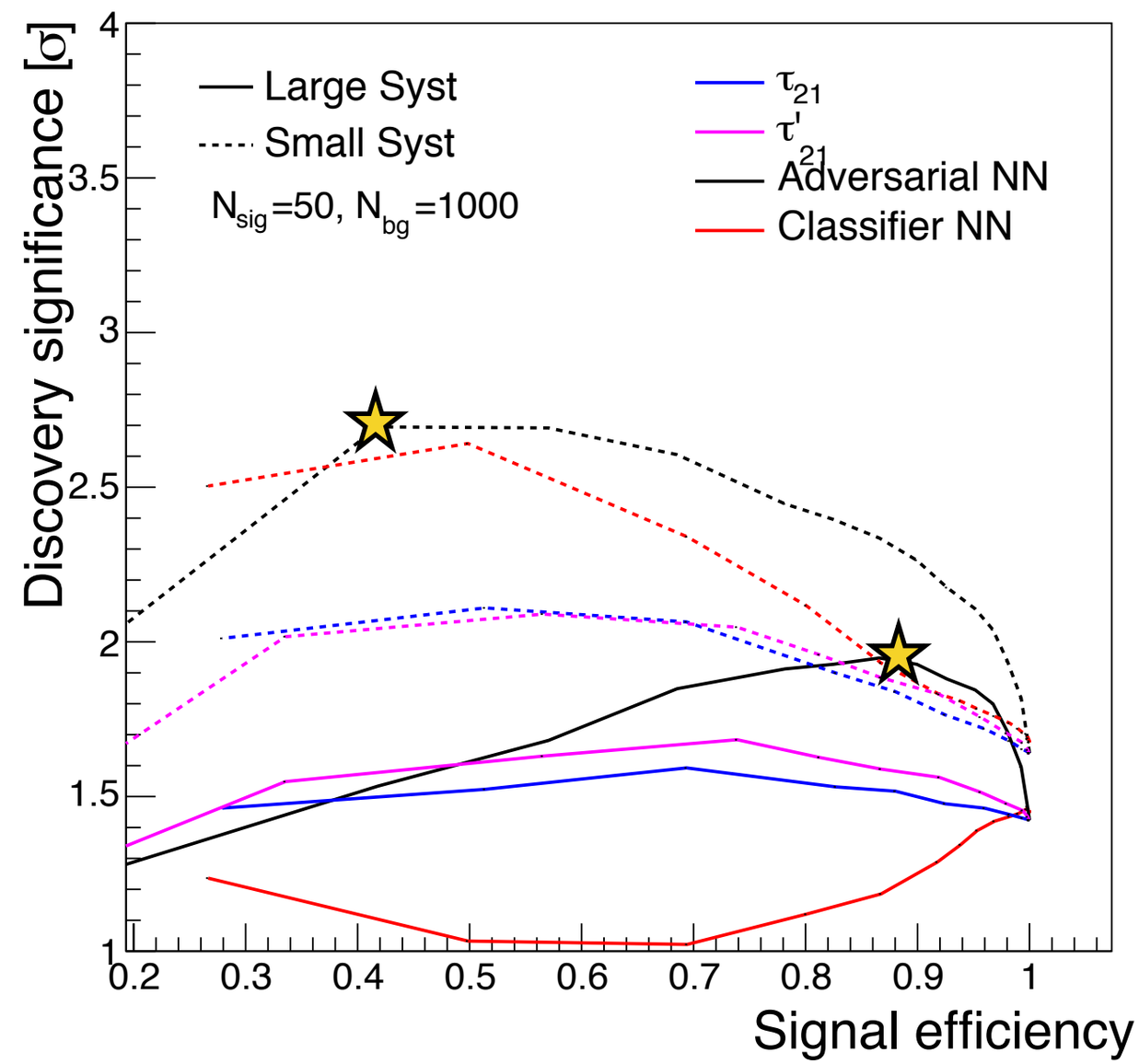
- Toy statistical model:
  - MC template fit
  - BG normalization uncertainty



Statistical Significance

- Toy statistical model:
  - MC template fit
  - BG normalization uncertainty

✓ Adversarial method attains highest discovery significance

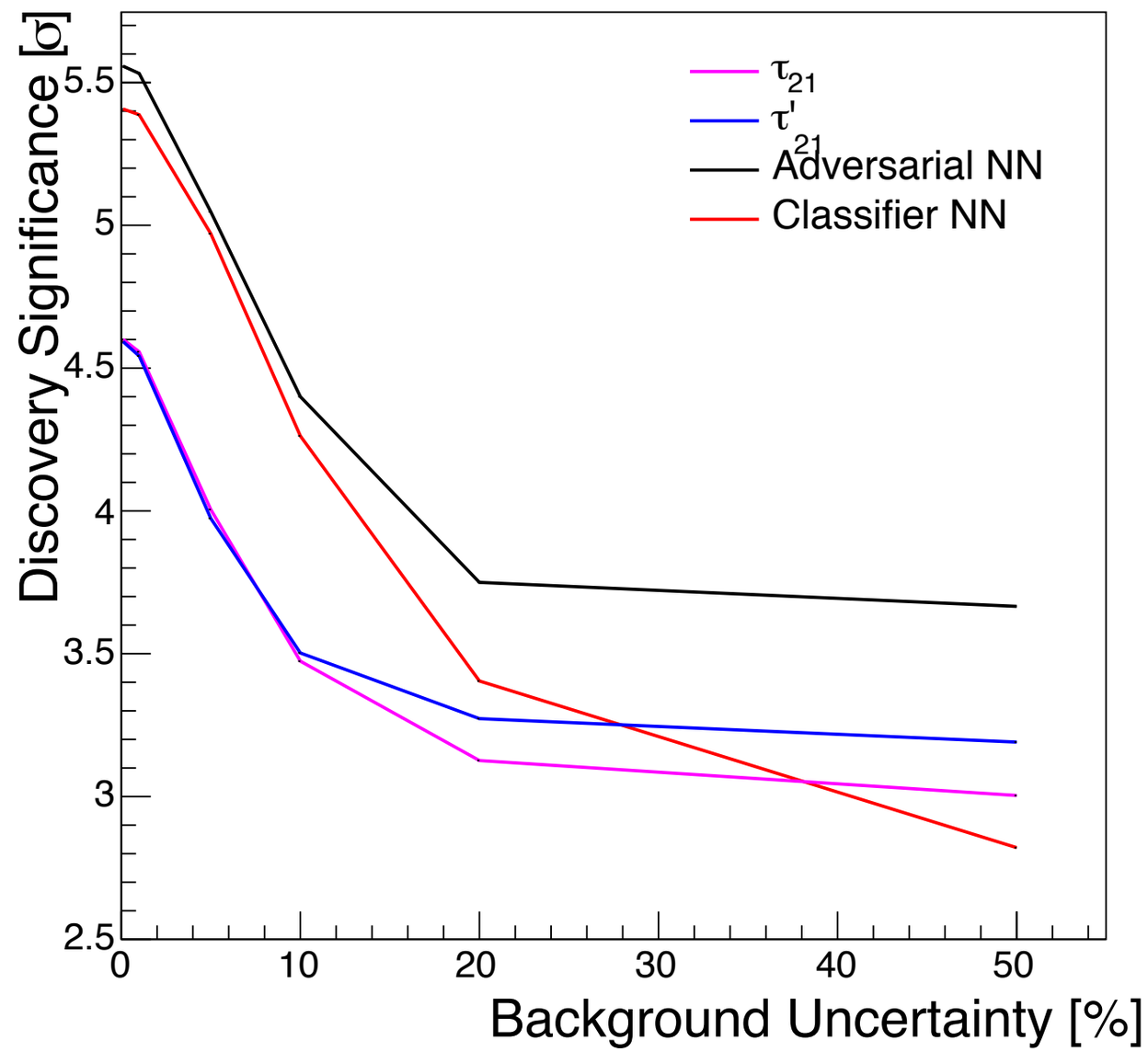


Statistical Significance

- Toy statistical model:
  - MC template fit
  - BG normalization uncertainty

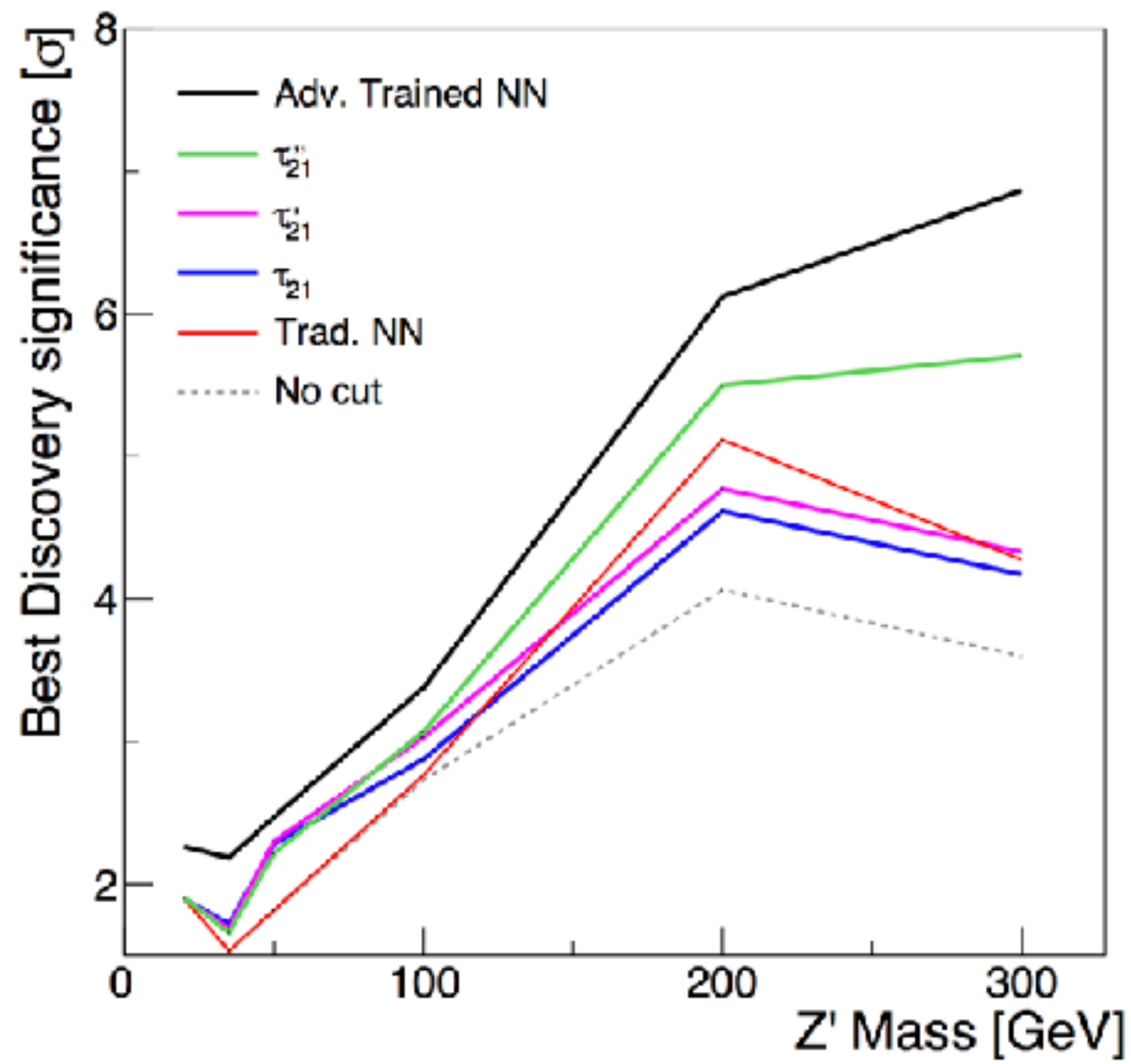
✓ Adversarial method attains highest discovery significance

- Larger systematics
  ⇒ stronger improvement



Parameter Scans

➡ Architecture can be extended to include parametric dependence on hypothesis mass, $M_{Z'}$

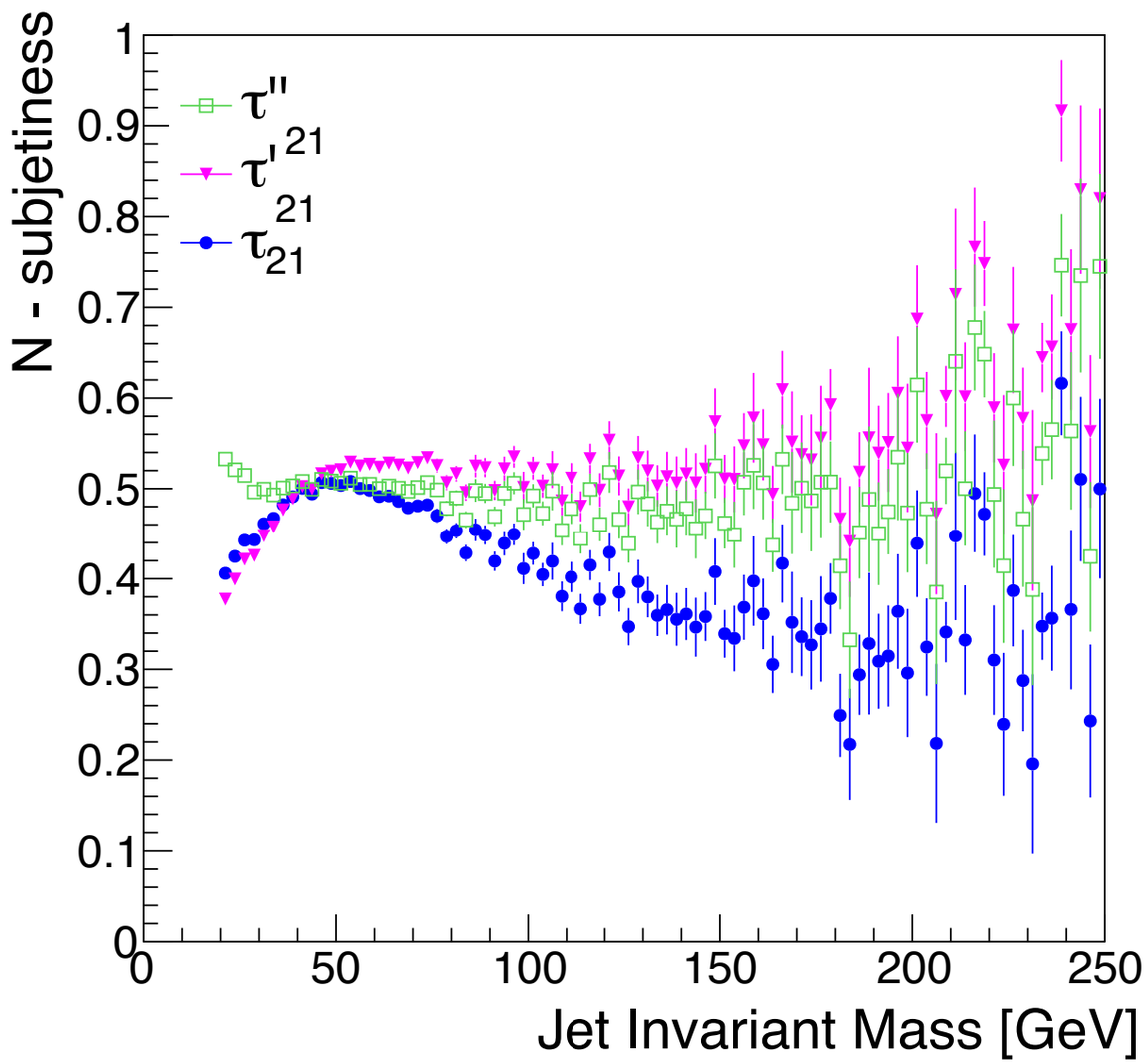


Summary / Conclusion

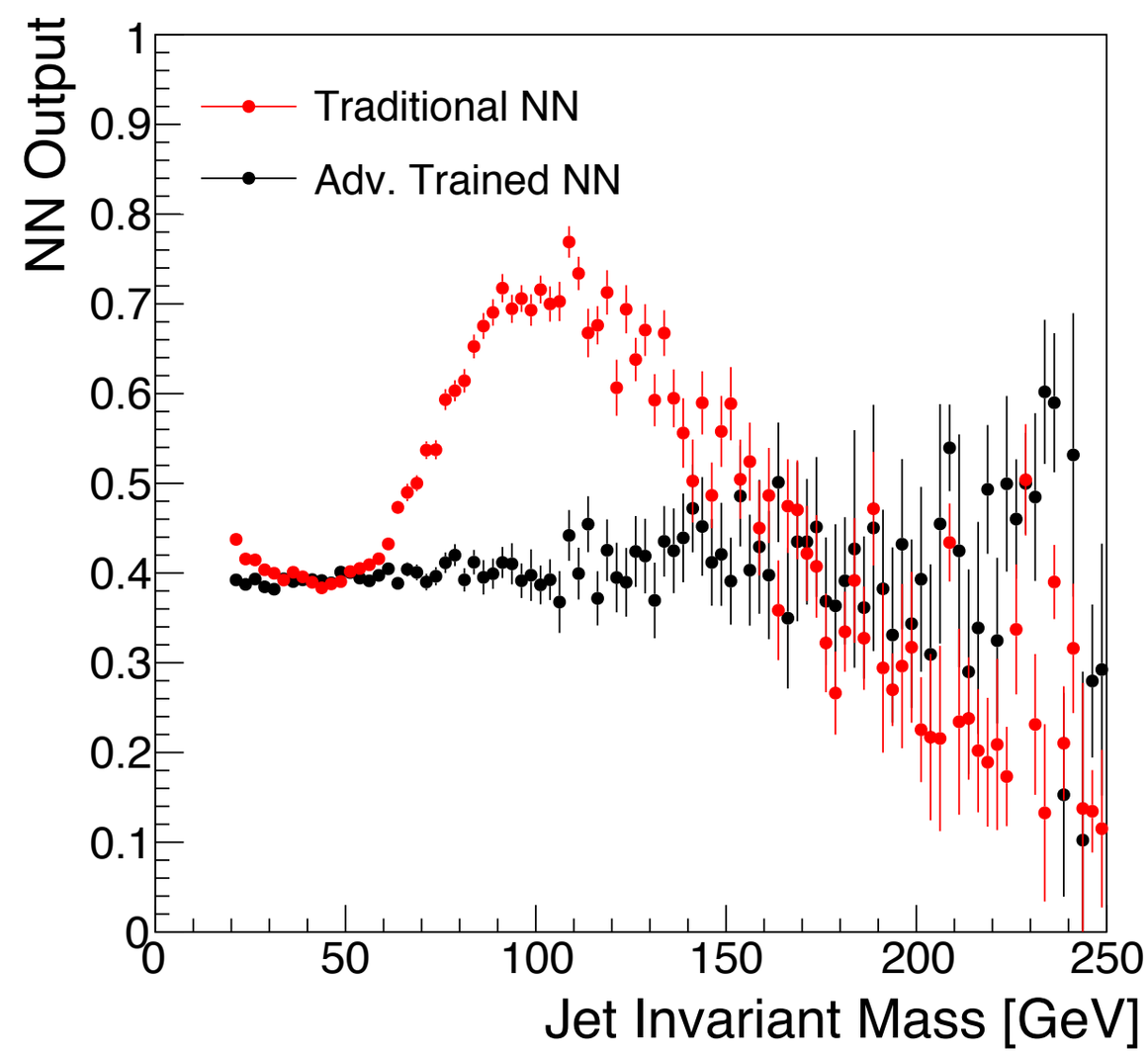
- Multivariate taggers are powerful tools for many signals
- However, correlation with analysis observables results in **reduced sensitivity** in the presence of **BG modeling systematics**
- Adversarial techniques can enforce decorrelation for arbitrarily complex classifiers
- The resulting classifier may outperform both theoretically-motivated variables as well as conventional neural networks
- The method is generic and should work for different object taggers and/or analysis observables

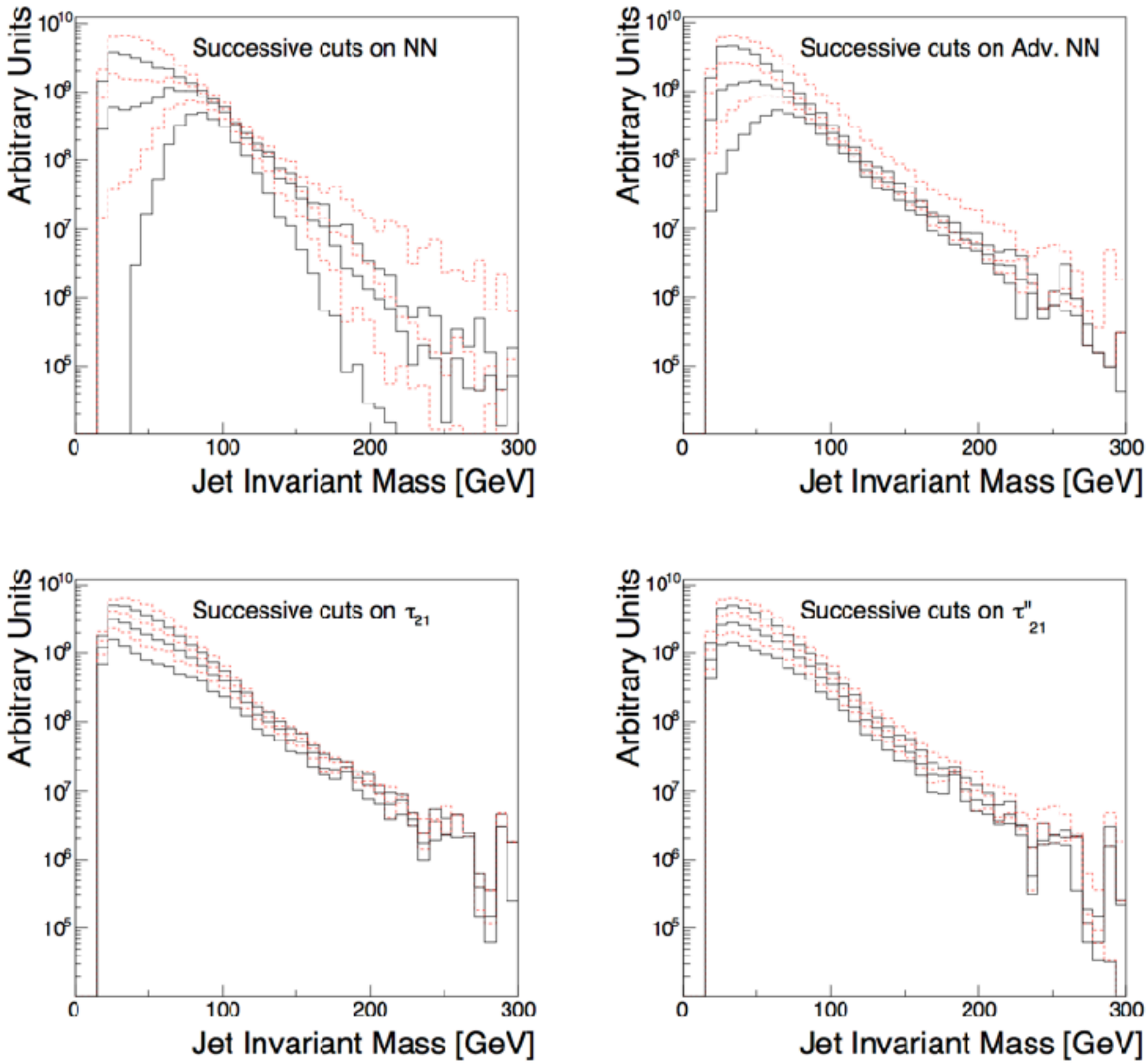
End

N-subjettiness profiles

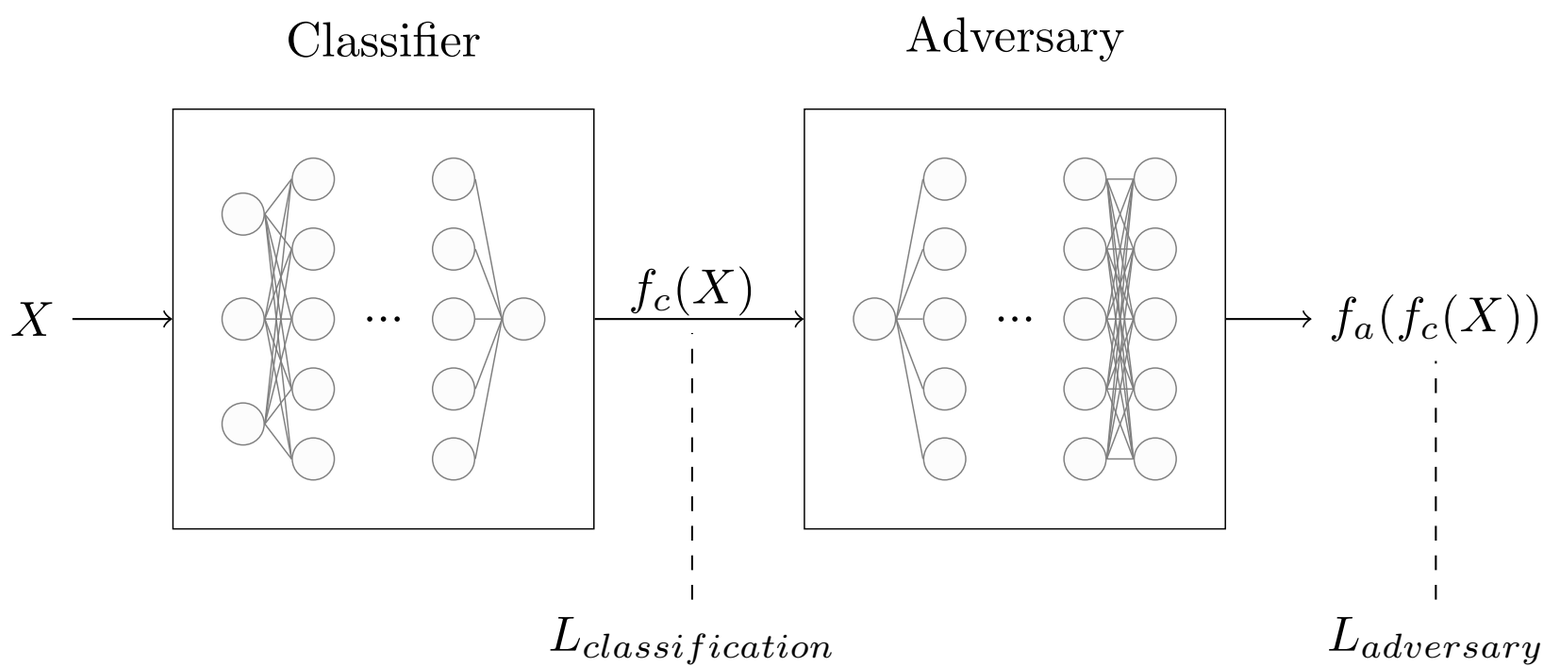


NN profiles

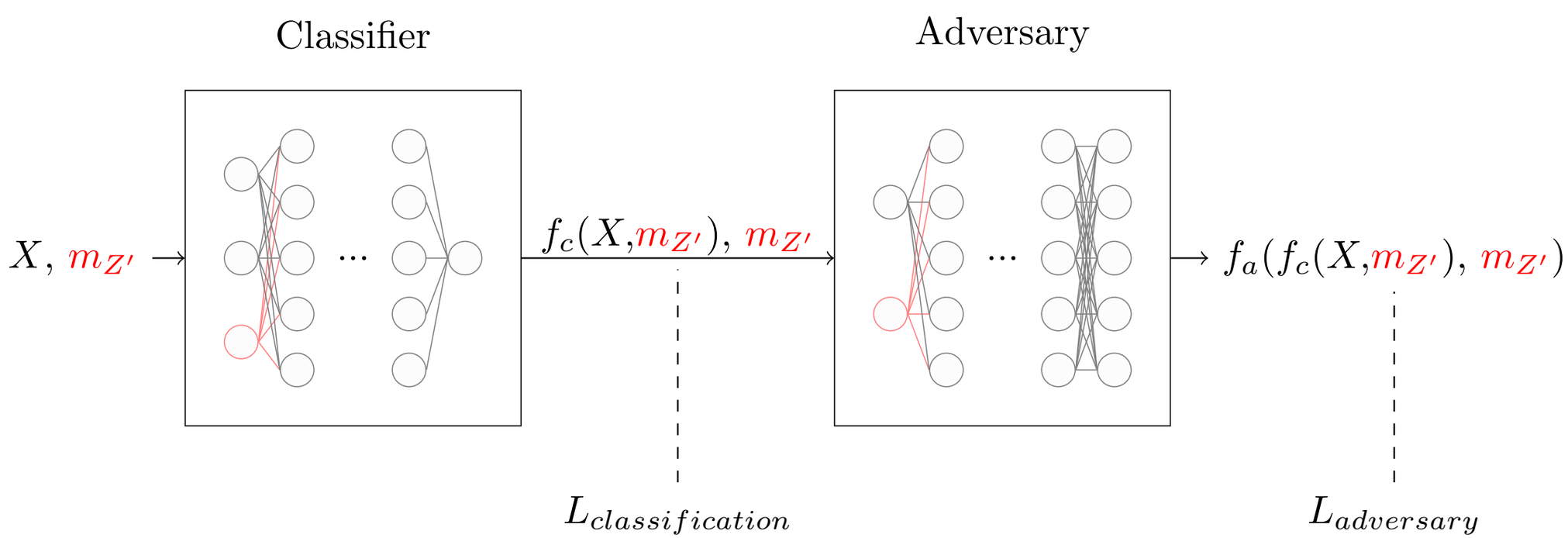



Successive cuts on NN
Successive cuts on Adv. NN
Successive cuts on τ_{21}
Successive cuts on τ''_{21}
Arbitrary Units
Jet Invariant Mass [GeV]
10^{10}
10^{9}
10^{8}
10^{7}
10^{6}
10^{5}
0
100
200
300

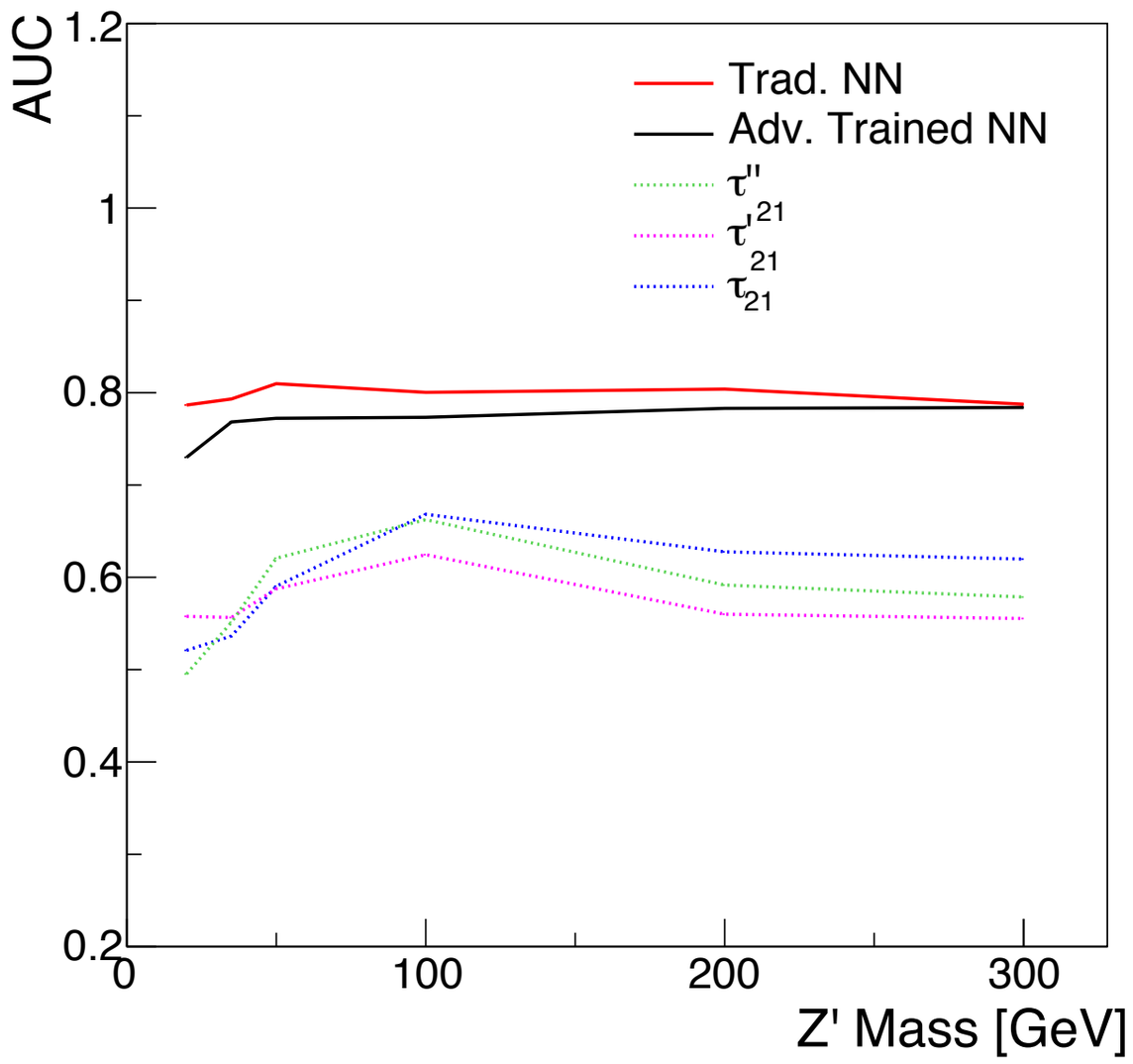
Adv. NN

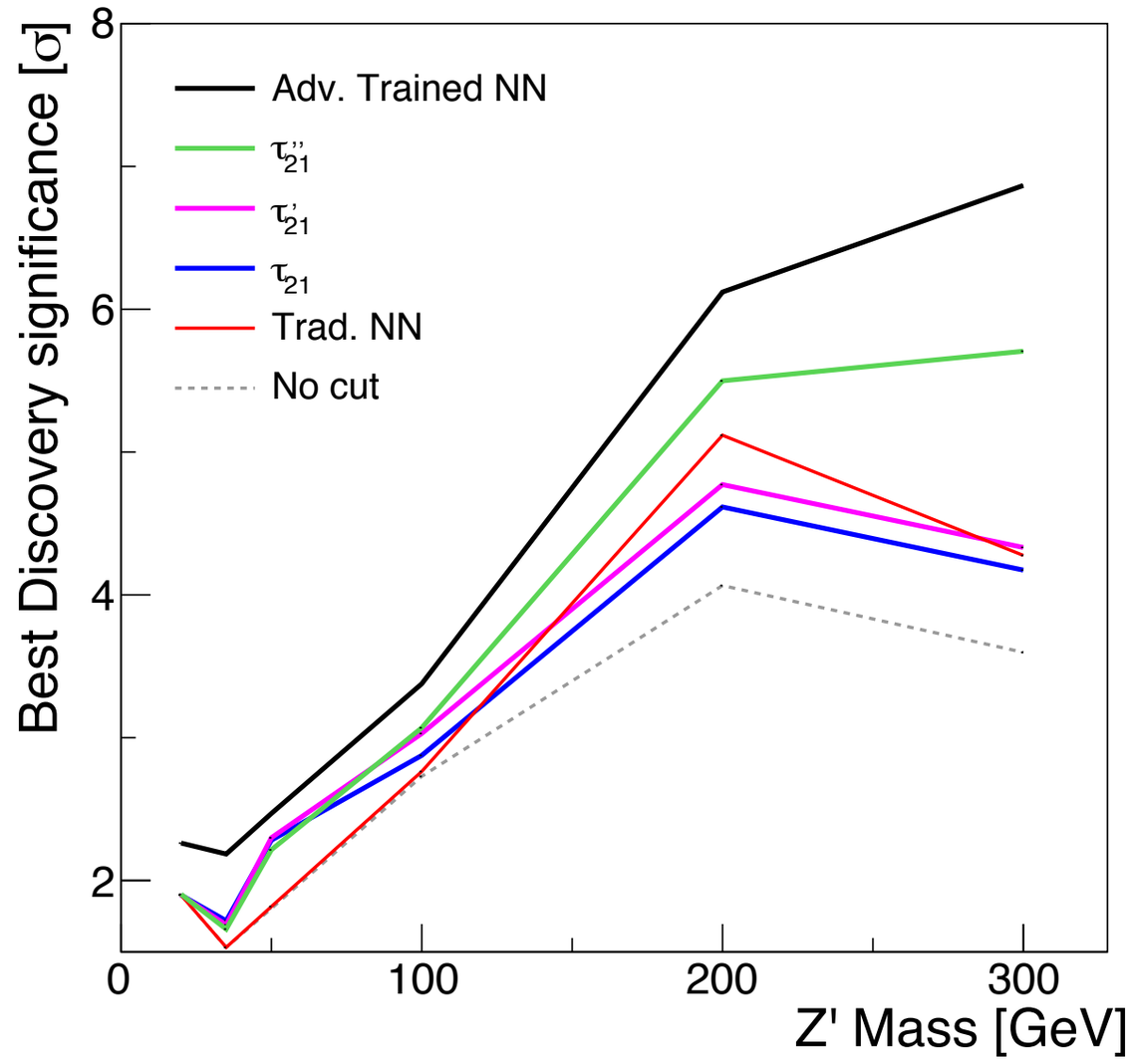


Parametric Adv. NN



AUC and significance





pT dependence