# The information content of quenched jets

James Mulligan in collaboration with Yue Shi Lai, Mateusz Płoskoń, Felix Ringer

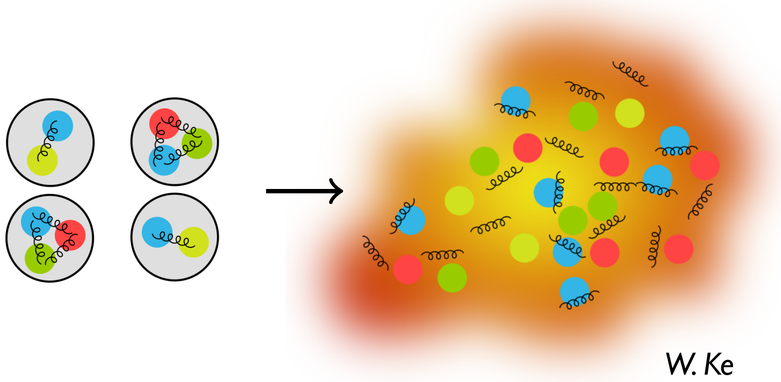
Lawrence Berkeley National Laboratory

ML4Jets
University of Heidelberg (virtual)
July 6 2021

# Jet quenching in the quark-gluon plasma

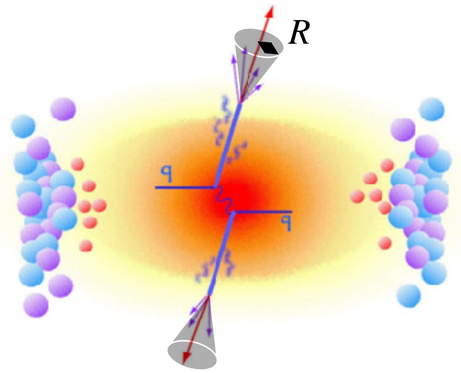
**In heavy-ion collisions at LHC/RHIC, we study the high-temperature, deconfined state of QCD known as the quark-gluon plasma (QGP)**

- What are the relevant degrees of freedom of the QGP?
- How does a strongly-coupled system arise from QFT?




**Jets can be used as a probe of the QGP**

- As jets traverse the QGP, their fragmentation is modified due to interactions with the medium
- These modifications can tell us about the microscopic structure of the QGP

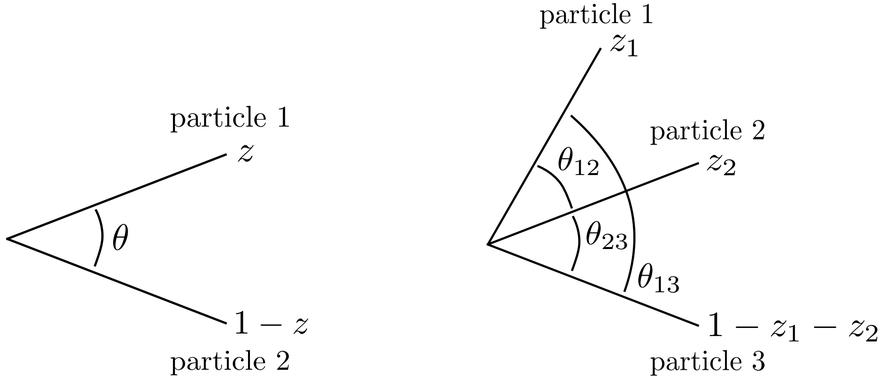



→ **We seek to understand how jets in heavy-ion collisions are different than jets in proton-proton collisions**

# The information content of jets in vacuum

*Datta, Larkoski JHEP 06 (2017) 073*

A jet with $K$ particles can be fully specified by $3K - 4$ observables

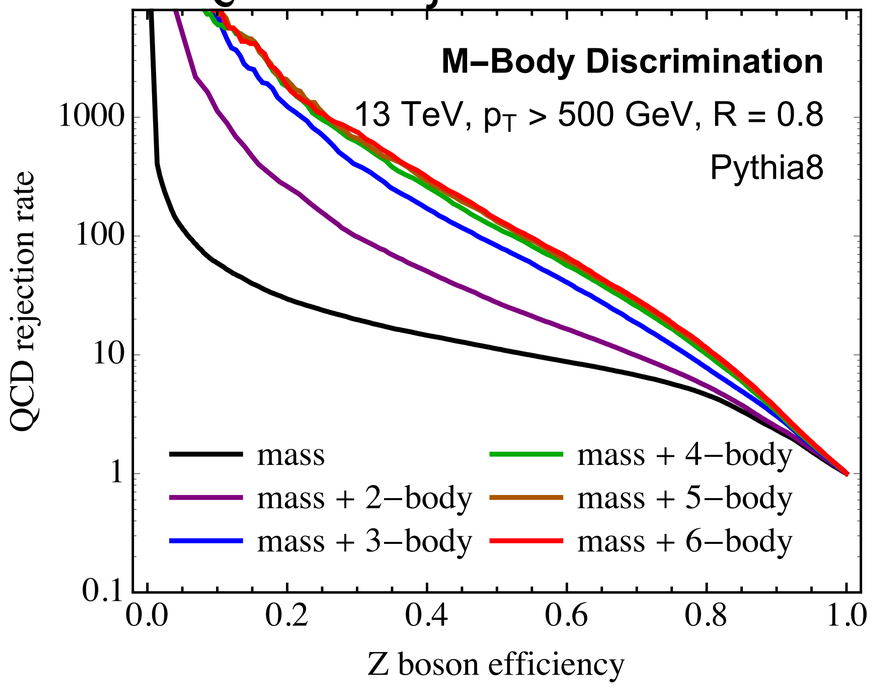



e.g. *N*-subjettiness basis:

$$\left\{\tau_1^{(0.5)}, \tau_1^{(1)}, \tau_1^{(2)}, \tau_2^{(0.5)}, \tau_2^{(1)}, \tau_2^{(2)}, \ldots, \tau_{M-2}^{(0.5)}, \tau_{M-2}^{(1)}, \tau_{M-2}^{(2)}, \tau_{M-1}^{(1)}, \tau_{M-1}^{(2)}\right\}$$

By constructing a complete set of IRC-safe observables, one can study at what point the information content saturates

*QCD vs. Z jet discrimination*



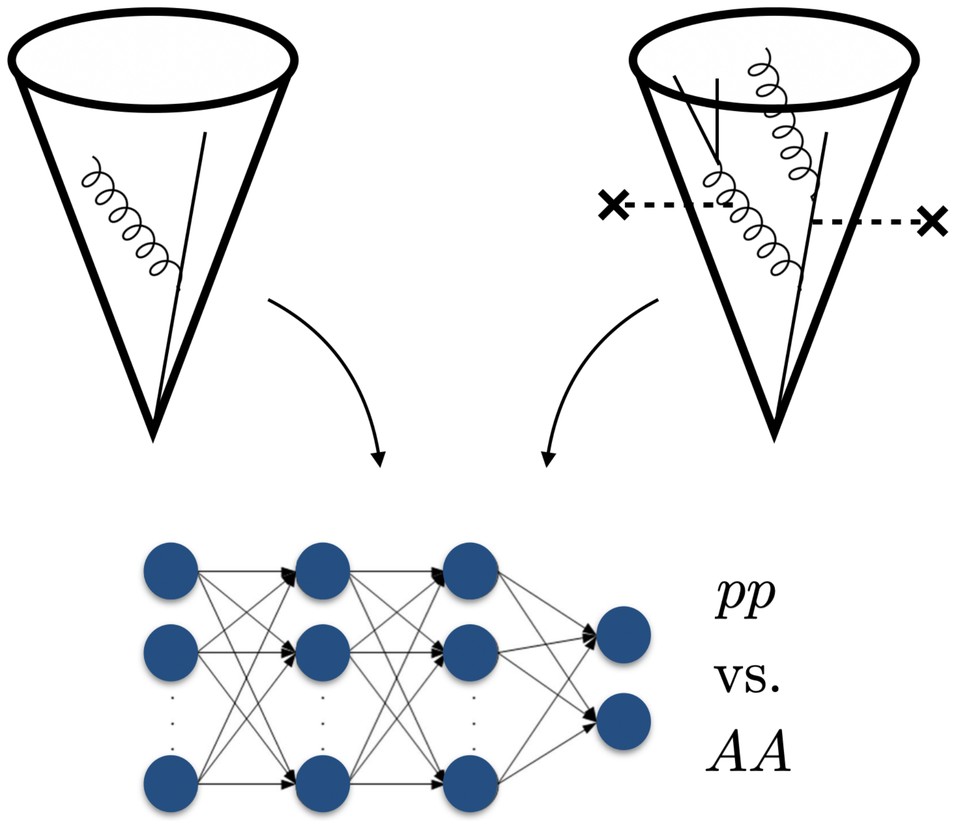
# The information content of quenched jets

In this talk we extend this idea to jet quenching

→ Binary classification problem

Determine the minimal set of observables to optimally discriminate pp vs. AA jets

- Quantify *K*-body discriminating power
- Find observables that capture the most discriminating aspects of jet modification




See also:

*Chien, Elayavalli 1803.03589*
*Lai 1810.00835*
*Du, Pablos, Tywoniuk JHEP 03 (2021) 206*
*Apolinário et al. 2106.08869*

# Two approaches

## *N*-subjettiness basis with Dense Neural Network (DNN)

Input layer: Complete set of jet substructure observables

$N$-subjettiness: *Thaler, Tilburg JHEP 03 (2011) 015*

$$\tau_N^{(\beta)} = \frac{1}{p_T} \sum_{i \in \text{Jet}} p_{Ti} \min\left\{R_{1i}^{\beta}, R_{2i}^{\beta}, \ldots, R_{Ni}^{\beta}\right\}$$

$K$-body phase space: *Datta, Larkoski JHEP 06 (2017) 073*

$$\left\{\tau_1^{(0.5)}, \tau_1^{(1)}, \tau_1^{(2)}, \tau_2^{(0.5)}, \tau_2^{(1)}, \tau_2^{(2)}, \ldots, \tau_{M-2}^{(0.5)}, \tau_{M-2}^{(1)}, \tau_{M-2}^{(2)}, \tau_{M-1}^{(1)}, \tau_{M-1}^{(2)}\right\}$$

DNN: $3K - 4$ inputs, 3 layers, tensorflow/keras

Only includes IRC-safe information

## Particle Flow Network (PFN)

*Komiske, Metodiev, Thaler JHEP 01 (2019) 121*

Deep sets *Zaheer et al. 1703.06114*
*Wagstaff et al. 1901.09006*
*Bloem-Reddy, Teh JMLR 21 90 (2020)*

Permutation-invariant neural network

$$f(p_1, \ldots, p_M) = F\left(\sum_{i=1}^{M} \Phi(p_i)\right)$$

latent space $d = 256$

Classifier

DNNs

Includes IRC-unsafe information

# Setup

Simulate jets using PYTHIA (pp) and JEWEL (AA)

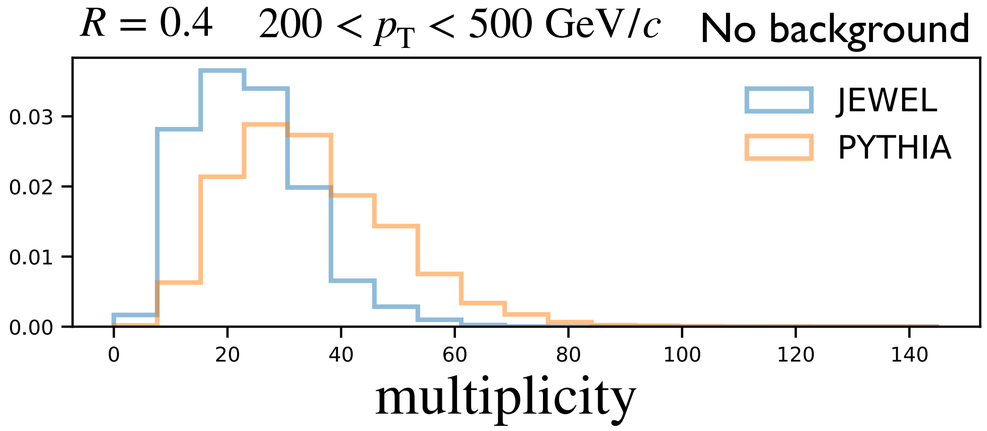
$$\sqrt{s_{\mathrm{NN}}} = 5.02 \text{ TeV}, \quad R = 0.4, \quad |\eta| < 2$$

$$200 < p_{\mathrm{T,jet}} < 500 \text{ GeV}/c$$

PhaseSpace : bias2SelectionRef = 20
PhaseSpace : bias2SelectionPow = 5.7

Note: JEWEL without recoils

No mass information used



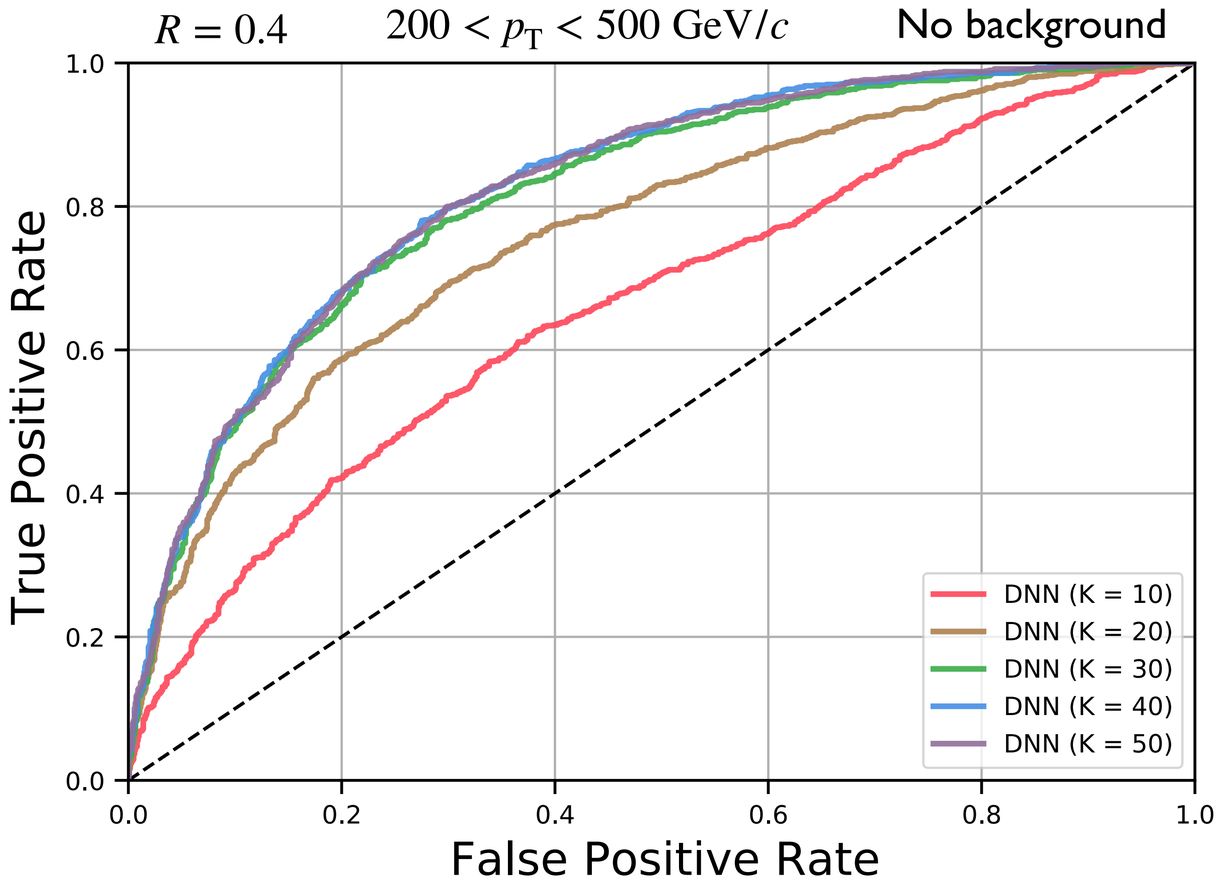

Consider (i) hard event only, (ii) hard event + heavy-ion background

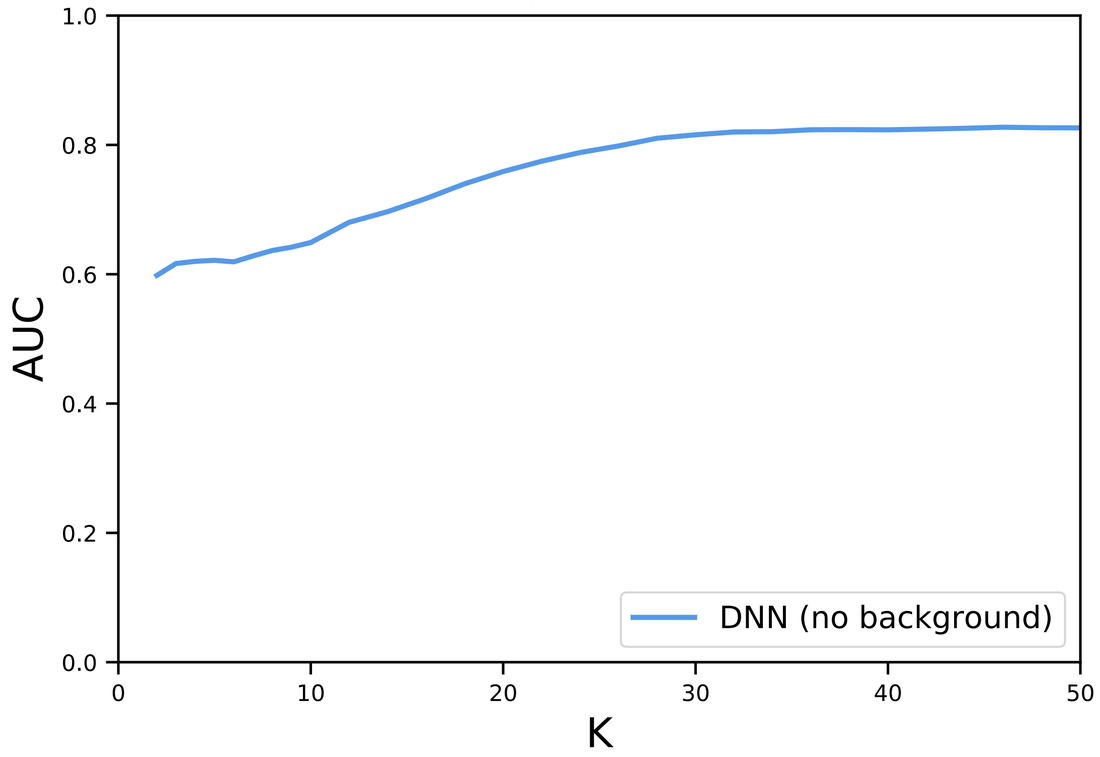
Background:

- Thermal background tuned to $0 - 10\,\%$ central $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV
  $dN/d\eta \approx 2500, \quad \langle p_{\mathrm{T}} \rangle \approx 0.8$ GeV/$c$
- Event-by-event constituent subtraction: $R_{\mathrm{max}} = 0.25$
- Jets matched to hard jet $p_{\mathrm{T}}$ scale

Results shown here are preliminary with modest statistics ~20k jets per $p_{\mathrm{T}}$ bin

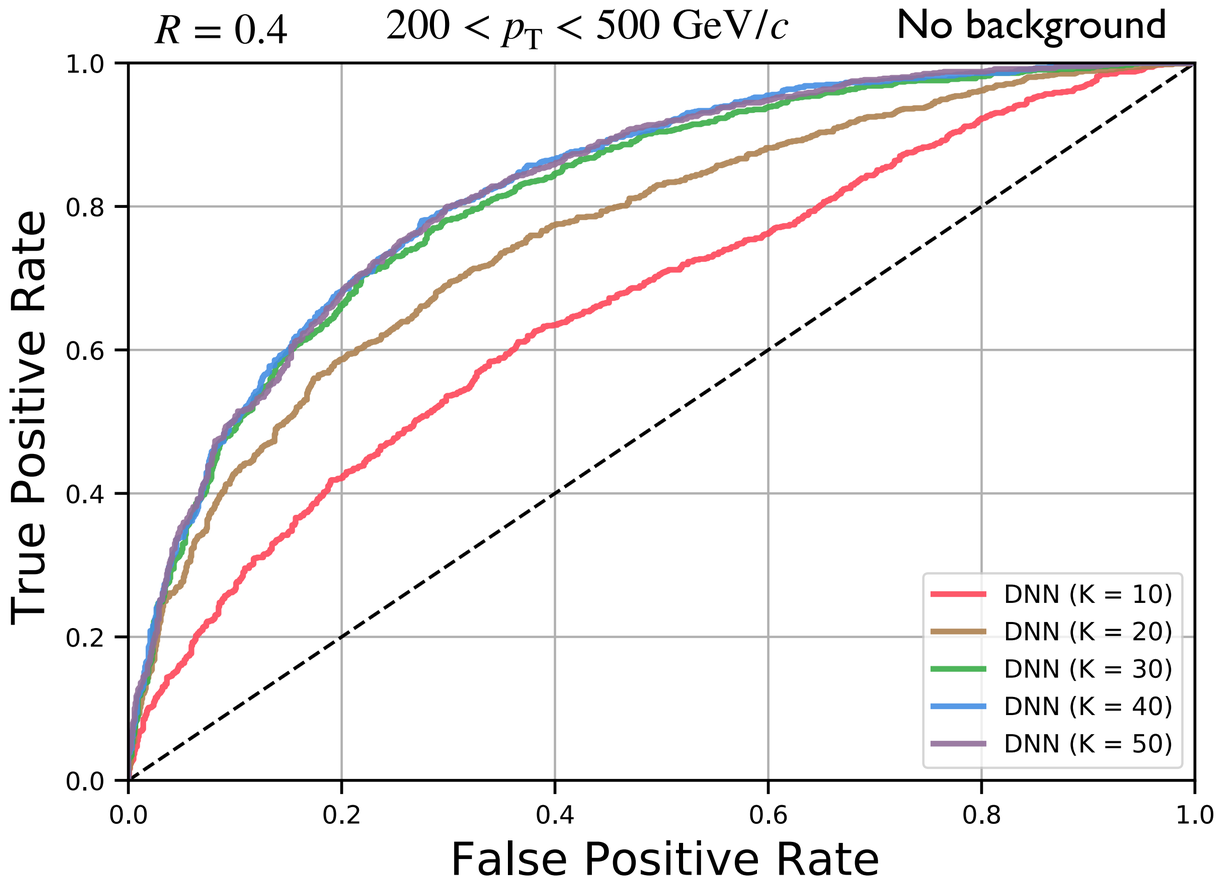
# Results — pp vs. AA



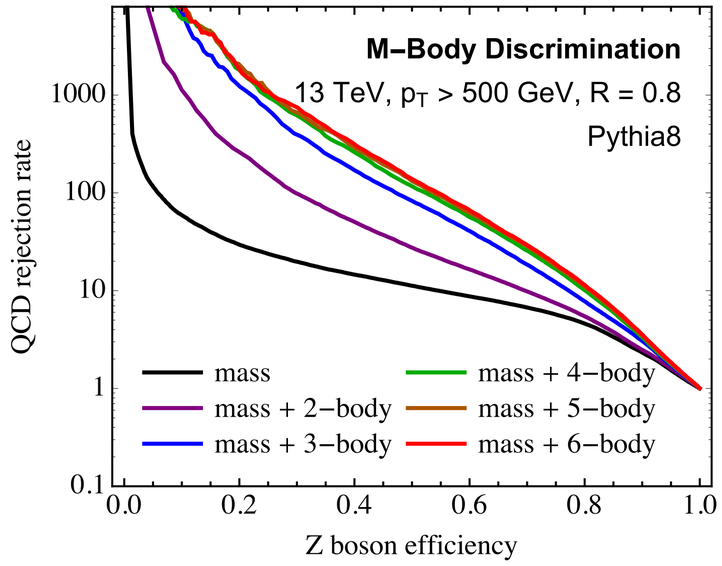




Significant information in quenched jets up to $K \approx 30$

# Results — pp vs. AA



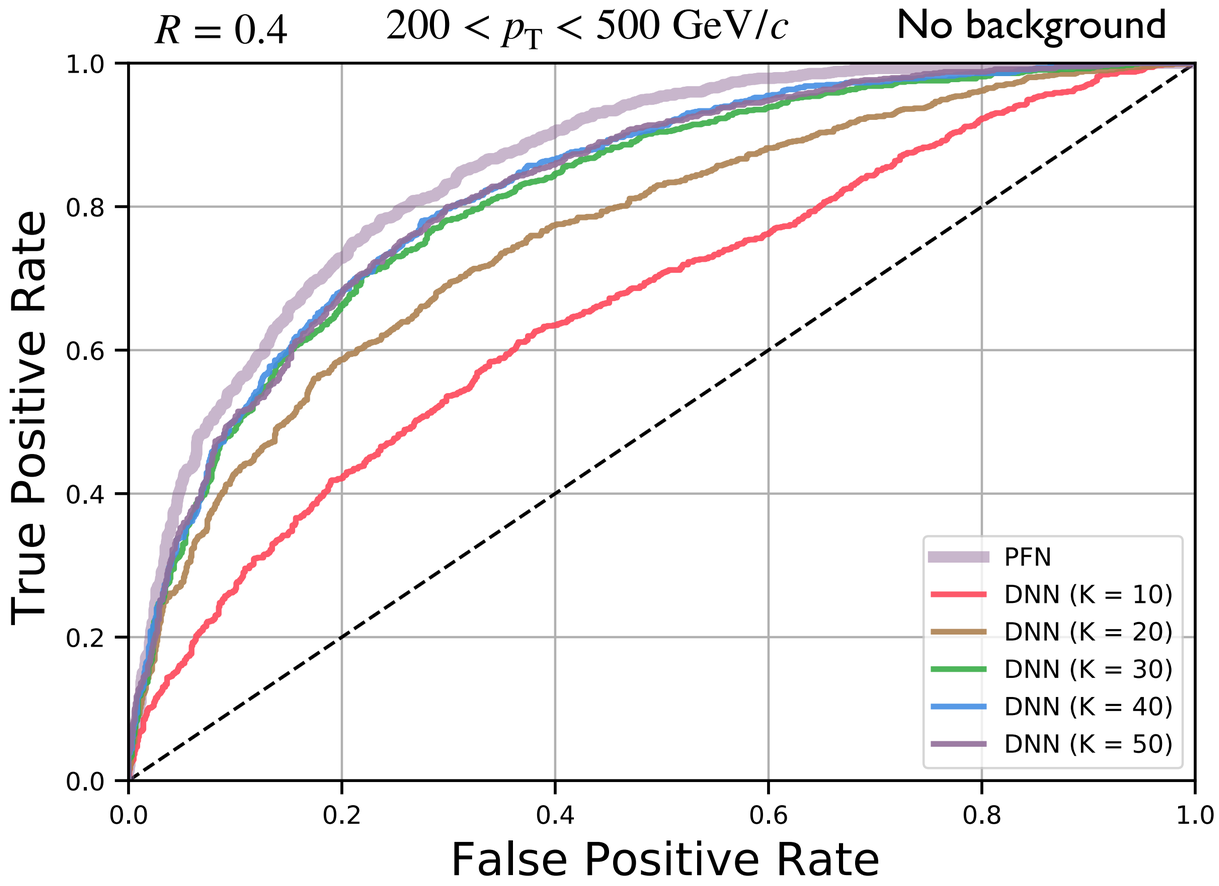

*Datta, Larkoski JHEP 06 (2017) 073*




→ Unlike QCD vs. Z jets (which saturate at $K = 4$), vacuum vs. quenched jets contain discriminating power in soft physics (high $K$-body phase space)

**Significant information in quenched jets up to $K \approx 30$**

# Results — pp vs. AA




Deep set data representation (PFN) performs slightly better than *N*-subjettiness basis (DNN)

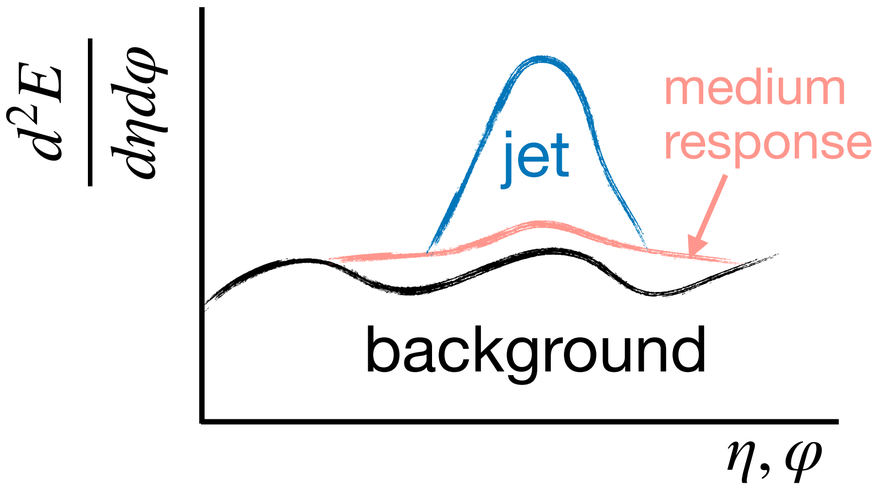
The difference can be due to:

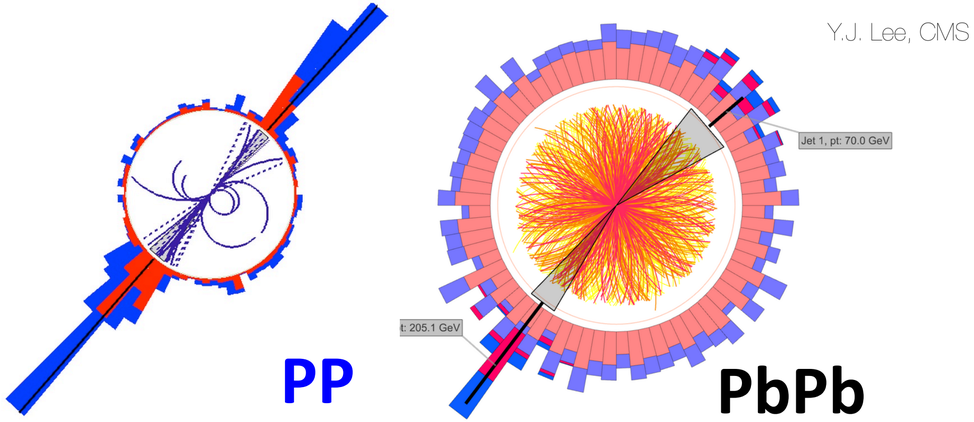
- IRC-unsafe information in PFN
- Different data representations / training / hyperparameter performance

Significant information in quenched jets up to $K \approx 30$

# Heavy-ion background

The soft collisions in a heavy-ion event produce a large, fluctuating underlying event

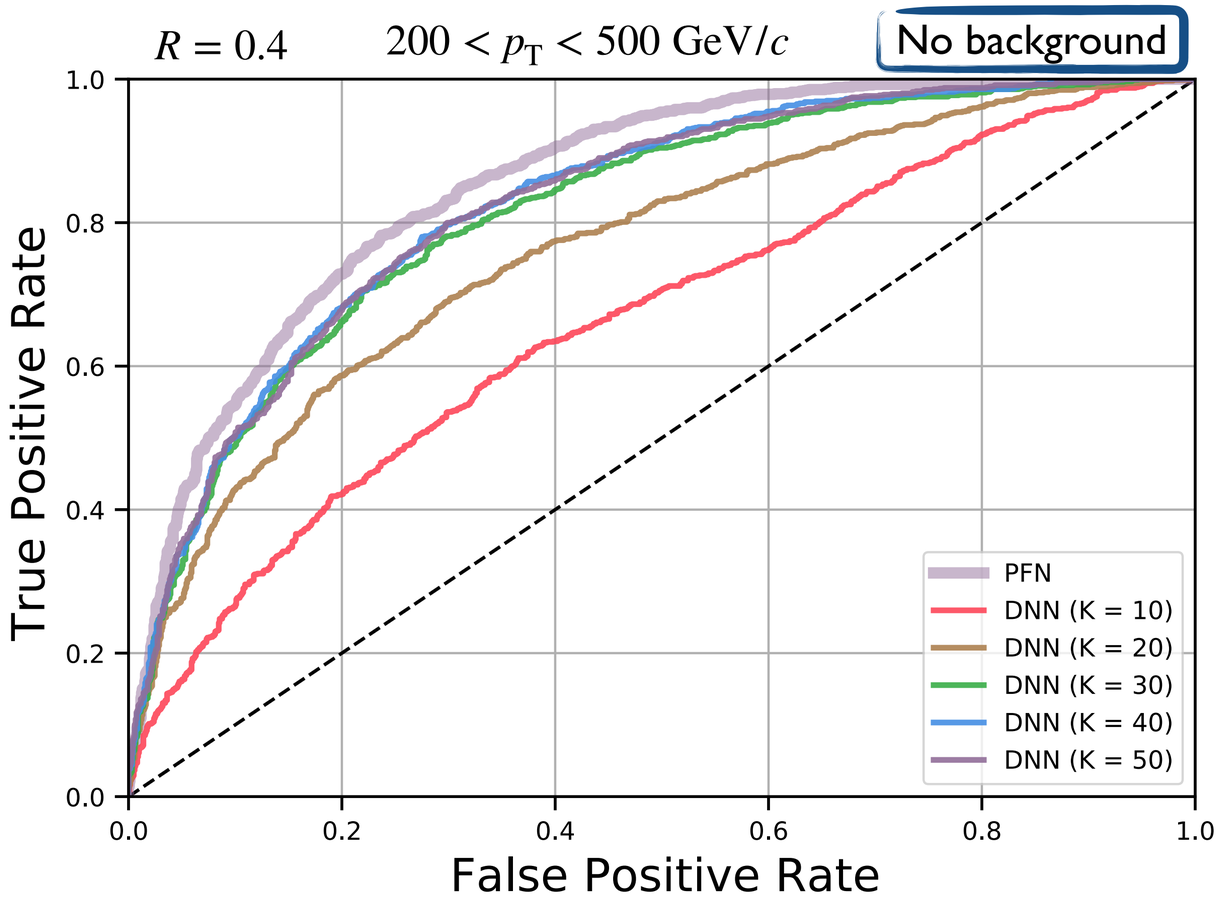


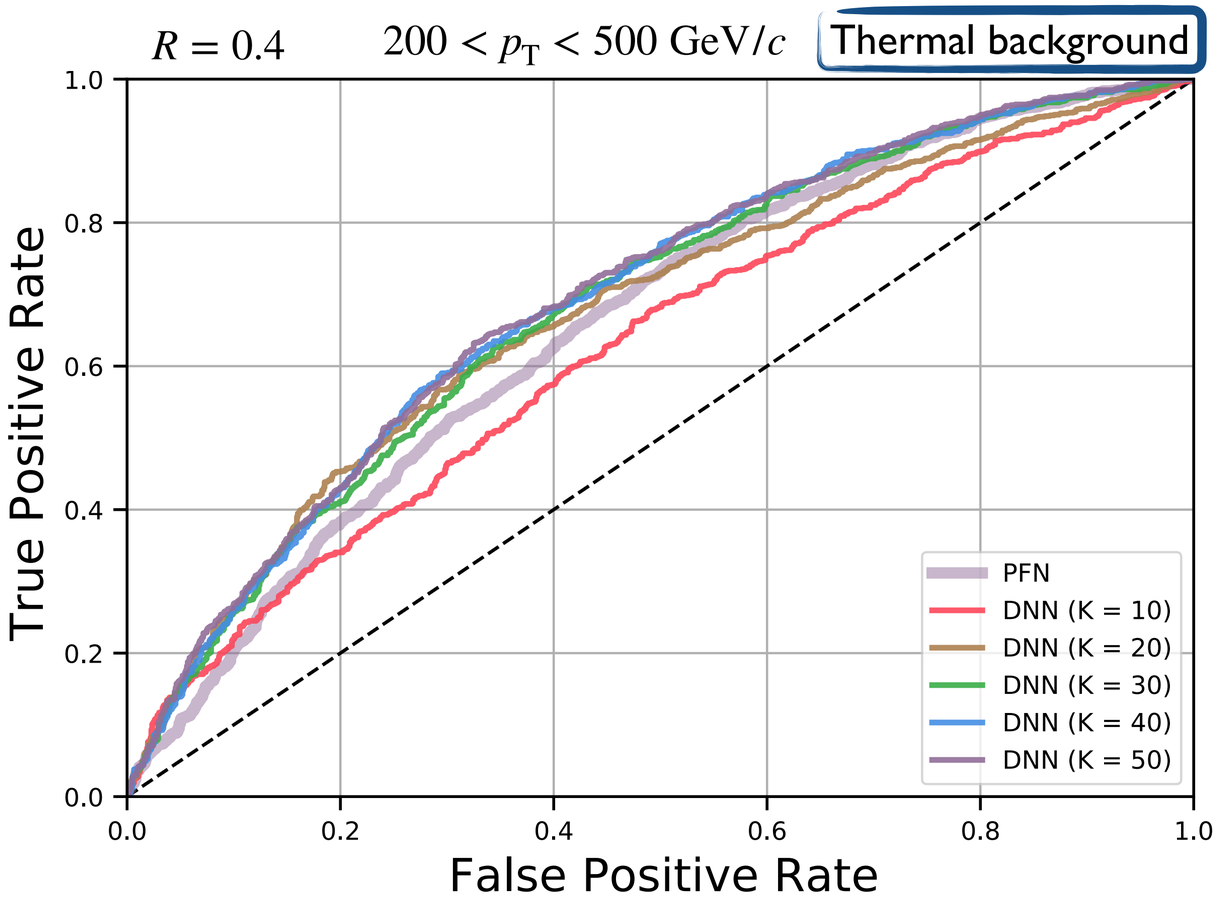




This is a major experimental and theoretical hurdle

**To what extent does the background destroy discriminating power?**

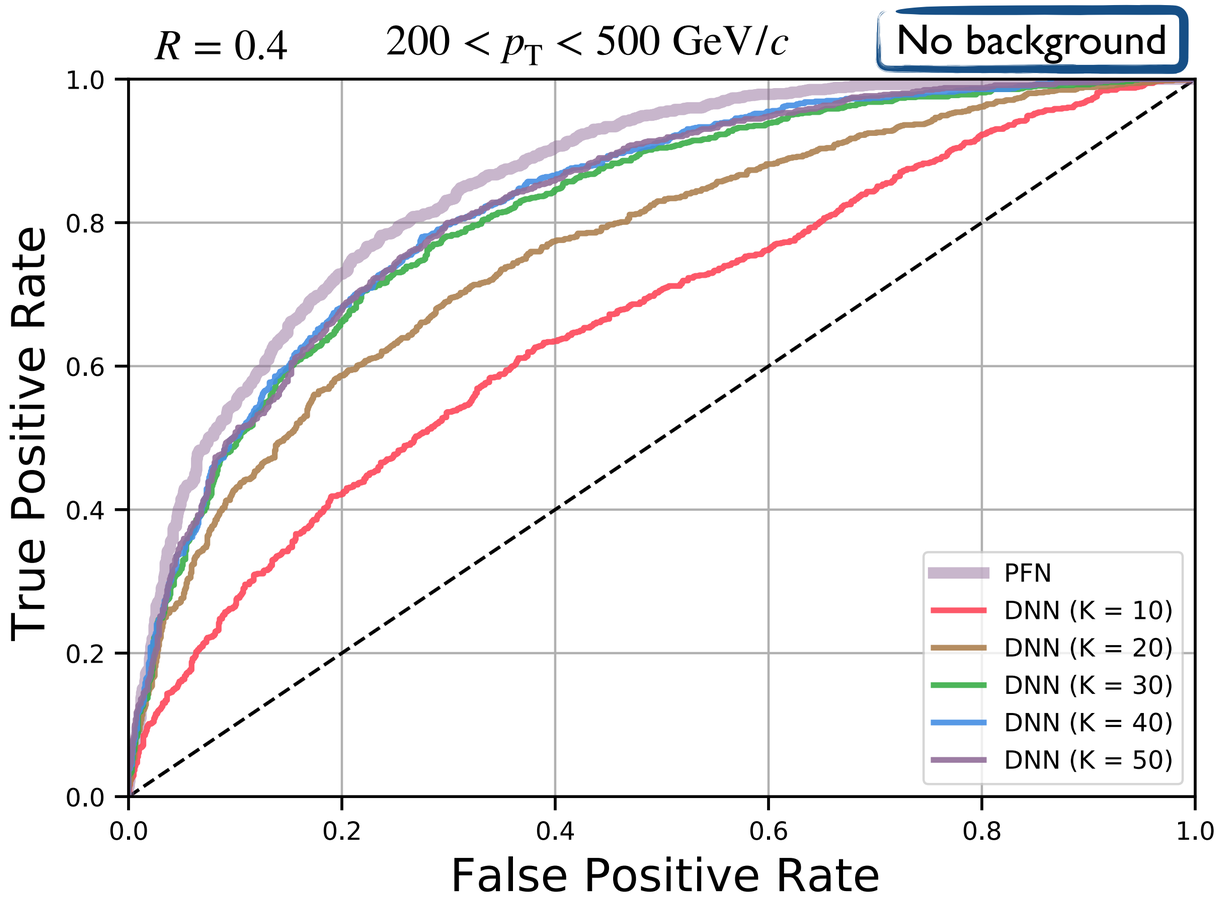
# Results — w/o vs. w/ background

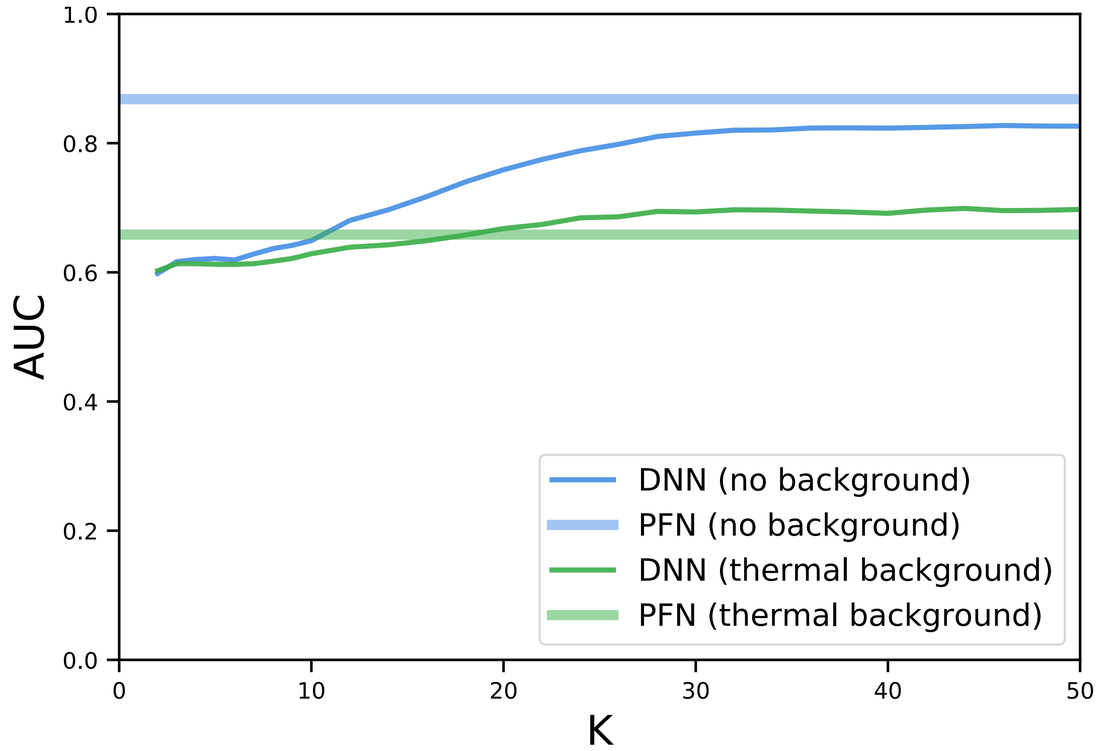



Discriminating power is highly reduced by the fluctuating underlying event

# Results — w/o vs. w/ background

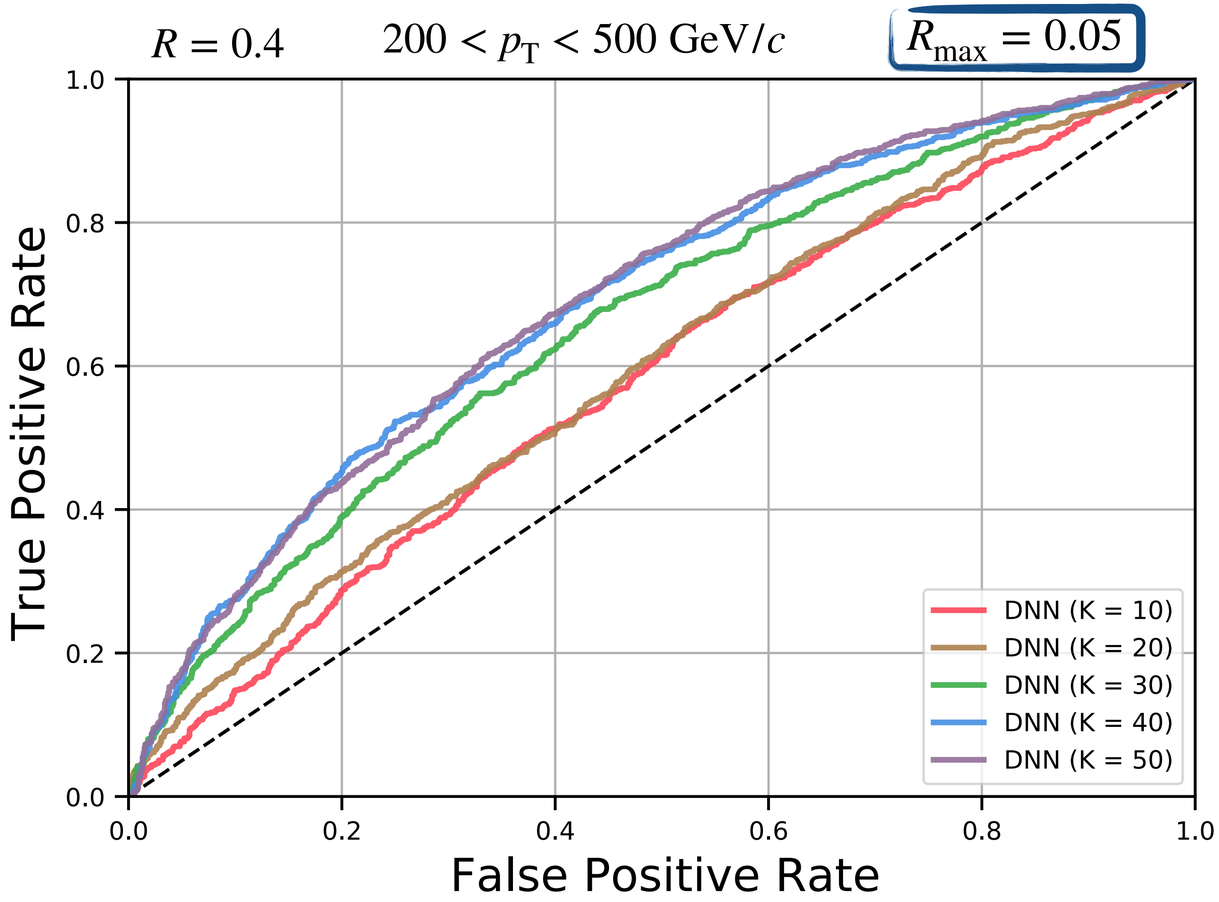


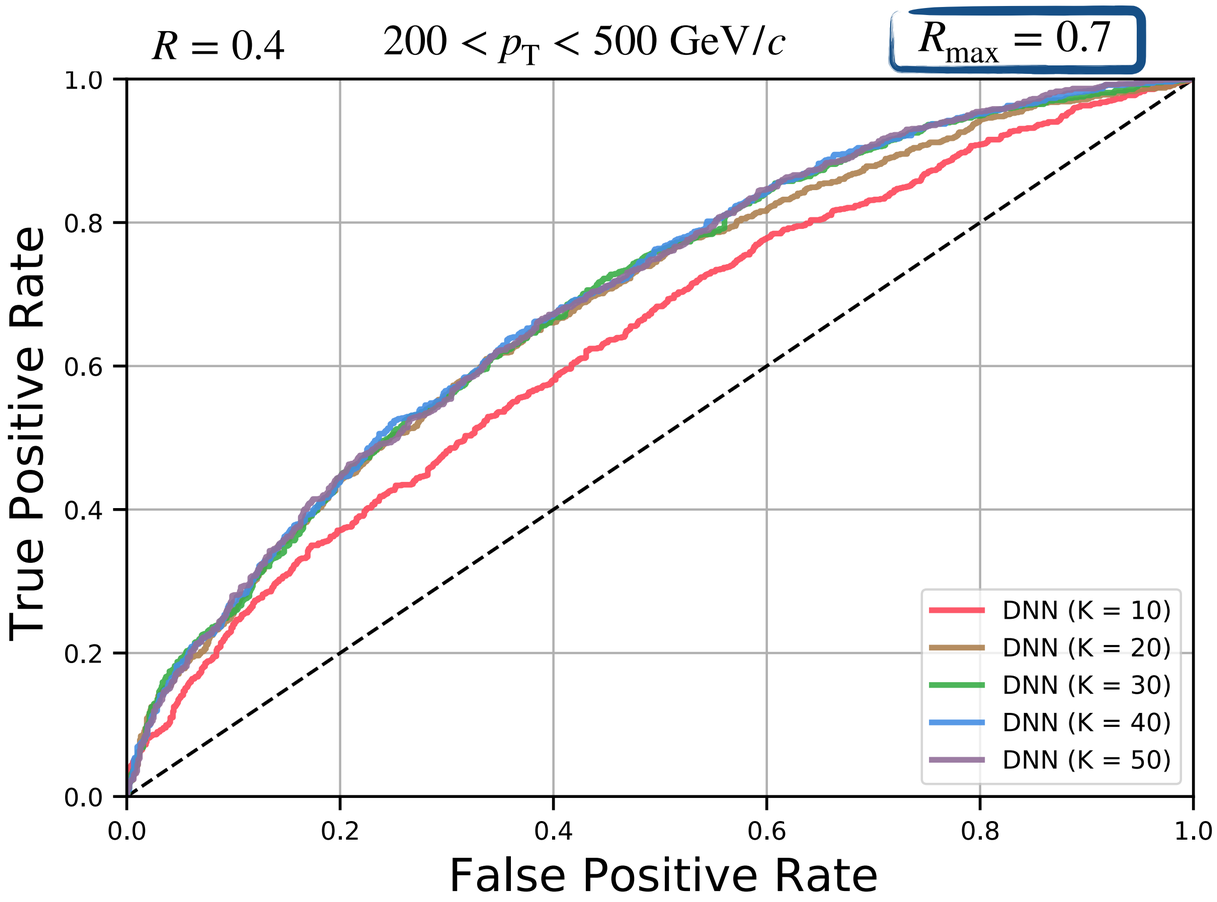




PFN performs slightly worse w/background — to be investigated

Discriminating power is highly reduced by the fluctuating underlying event

# Results — $R_{\max}$ dependence




The details of the background subtraction affect the distribution of information — but total discriminating power is fairly robust

# Automated design of observables

*Lai 1810.00835*
*Datta, Larkoski JHEP 03 (2018) 086*
*Datta, Larkoski, Nachman PRD 100, 095016 (2019)*

Now that we have demonstrated an ML classifier, we can find observable(s) that can approximate the classifier

→ Theoretical interpretability

Approximate the $3K - 4$ N-subjettiness observables with e.g. product observables

Product observable: Sudakov safe

$$O = \prod_{N<K,\ \beta\in\{0.5,1,2\}} \left(\tau_N^{\beta}\right)^{c_{N\beta}}$$

# Automated design of observables

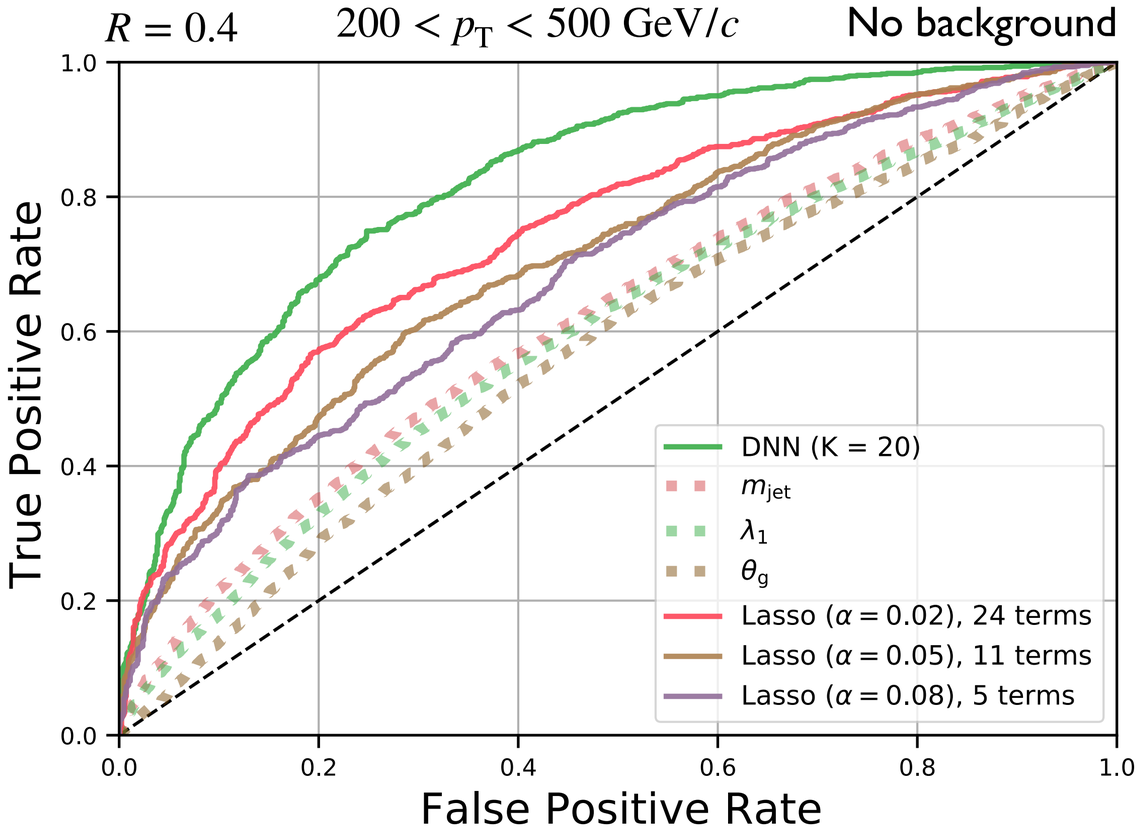
**Lasso regression**

$$O = \prod_{N<K,\ \beta\in\{0.5,1,2\}} \left(\tau_N^{\beta}\right)^{c_{N\beta}}$$

Stronger regularization drives $c_{N\beta}$ to zero

$\alpha = 0.02$ → 24 terms
$\alpha = 0.05$ → 11 terms
$\alpha = 0.08$ → 5 terms

e.g. $(\tau_1^2)^{1.437}(\tau_5^2)^{0.068}(\tau_6^2)^{1.712} \times \ldots$




Balancing the tradeoff of discriminating power and complexity, we can design optimal observables for distinguishing pp and AA jets

# Summary

**We trained binary classifiers to distinguish jets in heavy-ion collisions from those in proton-proton collisions**

- Using IRC-safe *N*-subjettiness basis: much more **discriminating information from soft physics** is needed for pp vs. AA jets than for QCD vs. Z-jets
  - → Motivates design of analytically calculable observables that are sensitive to soft physics
- Construct optimally discriminating observables using Lasso regression

**These methods can be applied directly to experimental data — labels are known**

- Can also apply to full events: LHC, EIC
- Is there an optimal basis, where the information content saturates more quickly?