The Information Content of Jet Quenching and Machine Learning Assisted Observable Design

Yue Shi Lai (Lawrence Berkeley National Lab)
James Mulligan (Lawrence Berkeley National Lab/U.C. Berkeley)
Mateusz Płoskoń (Lawrence Berkeley National Lab)
Felix Ringer (Stony Brook University/C. N. Yang Institute for Theoretical Physics)

arXiv:2111.14589 [hep-ph], submitted to JHEP

Introduction

Several guiding principles have been used to identify suitable observables in order to extract properties of the quark–gluon plasma (QGP):

- Infrared-collinear (IRC) safety
- Observables that can be measured well in the heavy-ion collision background
- Observables chosen based on their sensitivity to specific medium effects

We explore a new guiding principle: the systematic quantification of the relative information content of quenched vs. vacuum jets.

We examine jet quenching as a classification problem — distinguishing jets in proton–proton collisions from jets in heavy-ion collisions, as illustrated in Fig. 1. We utilize machine learning to quantify the features and patterns that distinguish these two classes of jets.

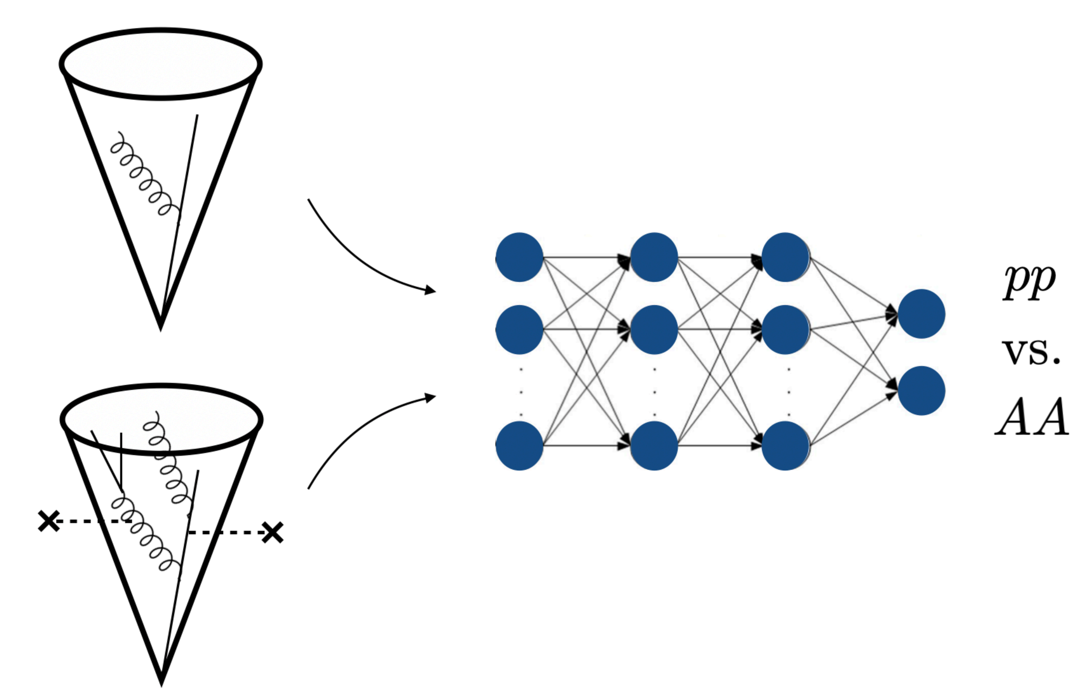

Fig. 1. Schematic illustration of jets in *pp* (top) and heavy-ion *AA* (bottom) collisions. By training a binary classifier, the machine learns the relevant information that distinguishes jets in *pp* and *AA* collisions.

Rather than training the classifier purely as a black box, we focus on formulating the classifier in a way that is theoretically interpretable. For the first time we can dissect the machine-learned information which allows for a degree of interpretability and provides guidance for constructing observables to study the QGP.

Monte Carlo Jet Sample

We compare PYTHIA8 [1] tune Monash2013 [2] and JEWEL [3] 2.2.0 (T_i = 590 MeV/k_B, τ_i = 0.4 $\hbar$/GeV) at $\sqrt{(s_{NN})}$ = 5.02 TeV. Recoil particles in JEWEL are not included for simplicity due to poorly understood physics of medium response. PYTHIA8 multi-parton interactions (MPI) is disabled in order to match JEWEL.

Jets are reconstructed using R = 0.4 anti-k_T algorithm within $|\eta_{\rm jet}| < 2$. Particle are unidentified and assumed to have the pion mass.

IRC Safe vs. Unsafe Information Content

We compare the performance using the IRC-unsafe particle flow network (PFN) [4]

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

with the IRC-safe (EFN) [4]

$$\tilde{f}(p_1, \ldots, p_M) = F\left(\sum_{i=1}^{M} z_i \Phi(\hat{p}_i)\right)$$

Φ and F are parametrized in term of deep neural networks (DNNs) using the EnergyFlow package [4] with Keras/TensorFlow [5]. Φ has two hidden layers with a latent space dimension of 256. F has three layers with 100 hidden neurons.

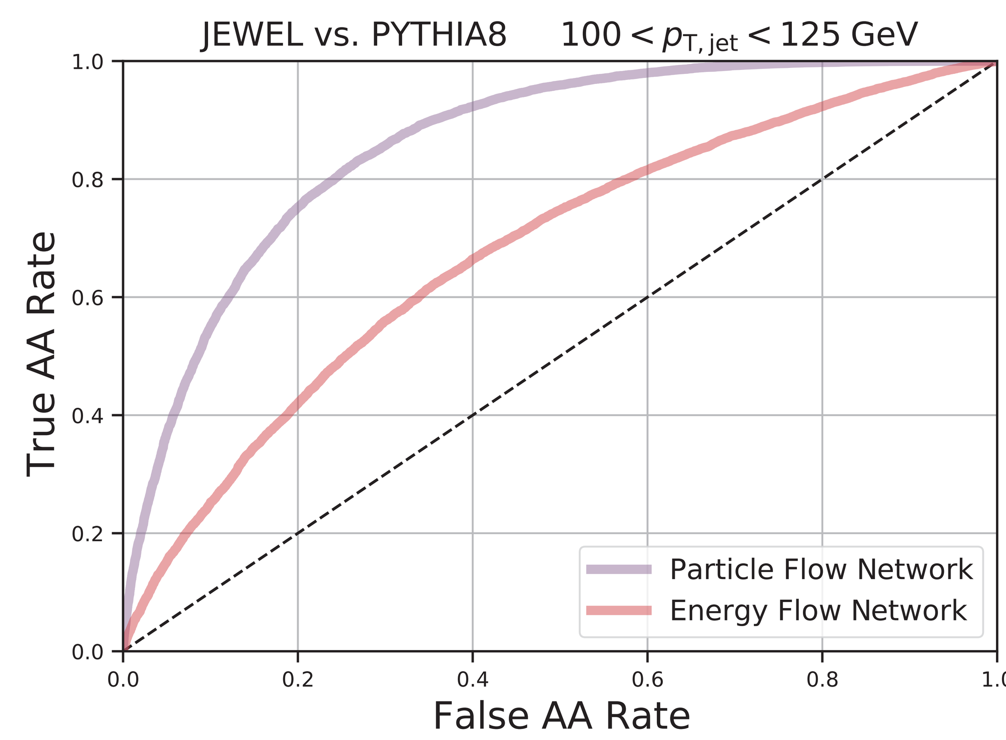

Fig. 2. Classification performance of *pp* vs. *AA* jets in terms of receiver operating characteristic (ROC) [6] curves using IRC-unsafe PFNs (area under the curve, AUC = 0.860) and IRC-safe EFNs (AUC = 0.675).

Findings suggest it will be valuable to measure jet substructure observables in heavy-ion collisions that go beyond IRC-safety.

Hard vs. Soft Information Content

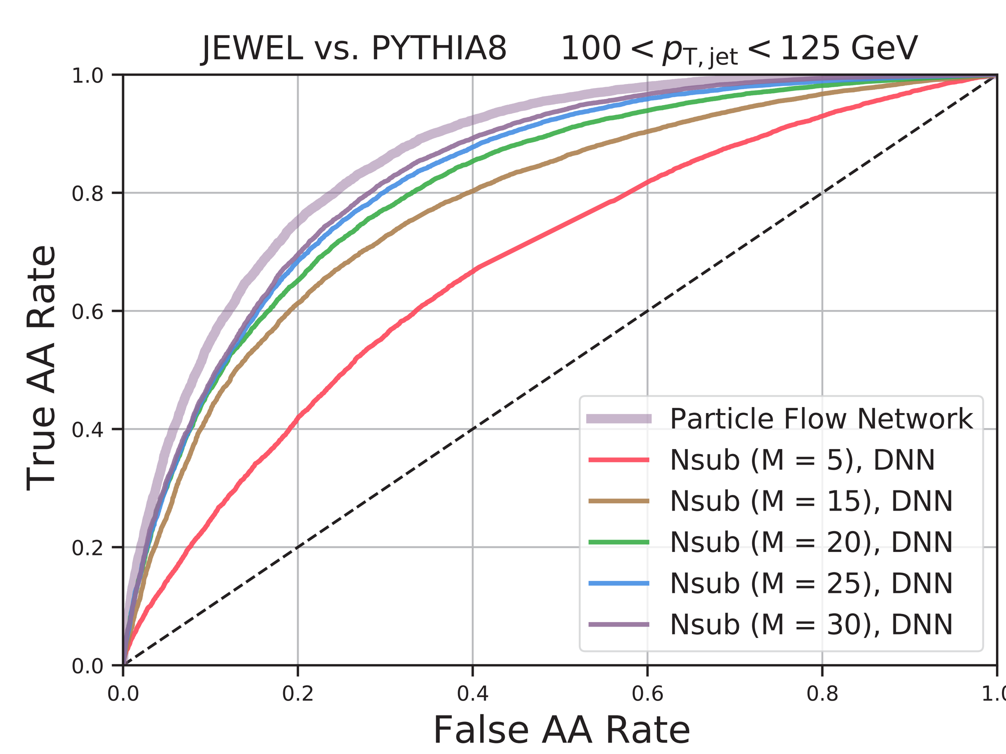

Fig. 3. ROC curves for jets in *pp* vs. *AA* collisions using the *N*-subjettiness basis. For comparison we also show the result obtained using the classifier based on PFNs.

Findings suggest that it will be necessary to measure new soft-sensitive jet substructure observables in heavy-ion collisions to fully make use of the available information recorded by the experimental collaborations. This information can be accessed by *N*-subjettiness observables for large values of *N*.

Observable Design

Lasso (least absolute shrinkage and selection operator) [7] classifier are constructed for a product of *N*-subjettiness variables, which is generally Sudakov safe [8]. The regularization parameter λ provides a handle to balance the performance of the classifier with the simplicity of the resulting observable. For several values of λ, following observables are found without background for at most M = 15 partices.

$$\lambda = 0.5: \quad \mathcal{O}^{\rm ML}_{N-{\rm sub}} = \tau^{(1)}_{14},$$

$$\lambda = 0.1: \quad \mathcal{O}^{\rm ML}_{N-{\rm sub}} = \left(\tau^{(1)}_{10}\right)^{0.071} \left(\tau^{(1)}_{11}\right)^{0.157} \left(\tau^{(1)}_{14}\right)^{0.649} \tau^{(2)}_{14},$$

$$\lambda = 0.01: \quad \mathcal{O}^{\rm ML}_{N-{\rm sub}} = \left(\tau^{(0.5)}_{2}\right)^{0.608} \left(\tau^{(2)}_{4}\right)^{-0.186} \times \ldots \times \tau^{(2)}_{14} \quad (23\ \text{terms}).$$

Bibliography

[1] T. Sjöstrand *et al.*, *Comput. Phys. Comm.* **178**, 852 (2008).
[2] P. Skands *et al.*, *Eur. Phys. J. C* **74**, 3024 (2014).
[3] K. C. Zapp *et al.*, *JHEP* **03**, 080 (2013); K. C. Zapp, *Eur. Phys. J. C* **74**, 2762 (2014).
[4] P. T. Komiske *et al.*, *JHEP* **01**, 121 (2019).
[5] F. Chollet, https://github.com/fchollet/keras; M. Abadi *et al.*, 12th USENIX Symposium OSDI (2016).
[6] W. Peterson *et al.*, *Trans. IRE Prof. Group Inform. Theory* **4**, 171 (1954)
[7] F. Santosa *et al.*, *SIAM J. Sci. Stat. Comput.* **7**, 1307 (1986).
[8] K. Datta *et al.*, *JHEP* **03**, 086 (2018); K. Datta *et al.*, *Phys. Rev. D* **100**, 095016 (2019)
[9] P. T. Komiske *et al.*, *JHEP* **04**, 013 (2018)
[10] J. Mulligan, M. Płoskoń, *Phys. Rev. C* **102**, 044913 (2020).

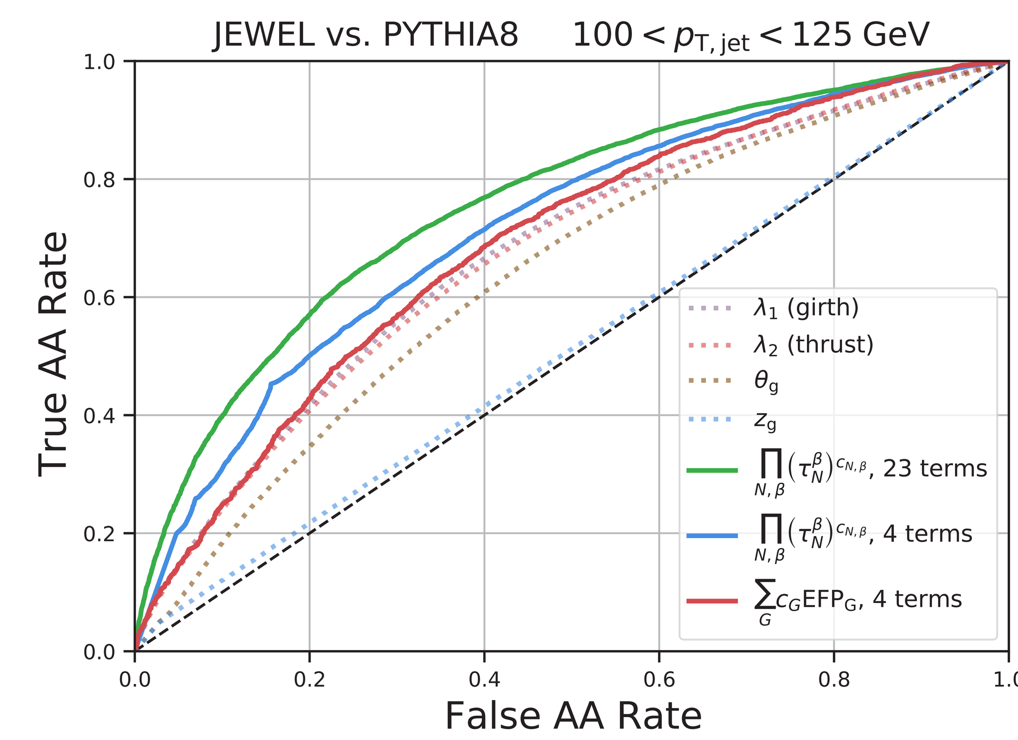

Fig. 4. ROC curves for the Lasso regression using the N-subjettiness basis and energy-flow polynomials (EFPs) [9]. For comparison we also show the result for typical observables in heavy-ion collisions.

The Lasso regression generally prefers large values of *N*. For sufficiently large values of λ, the Lasso regression always picks only one observable which turns out to be one of the *N*-subjettiness observables with the largest allowed value of *N*.

Information Loss from Underlying Event

The particle multiplicities, fluctuations, and transverse momentum observed in 0–10% central Pb–Pb data is approximated based on a thermal model consisting of *N* particles drawn from a Gaussian with $\langle dN/d\eta \rangle \approx 2500$ and p_T sampled event-by-event from a Gamma distribution, $f_\Gamma(p_T;\, \alpha, \beta) \propto p_T^{\alpha-1}\exp(-p_T/\beta)$ with $\alpha = 2$, $\beta = 0.4$ (or $\langle p_T \rangle \approx 0.8$ GeV) [10].

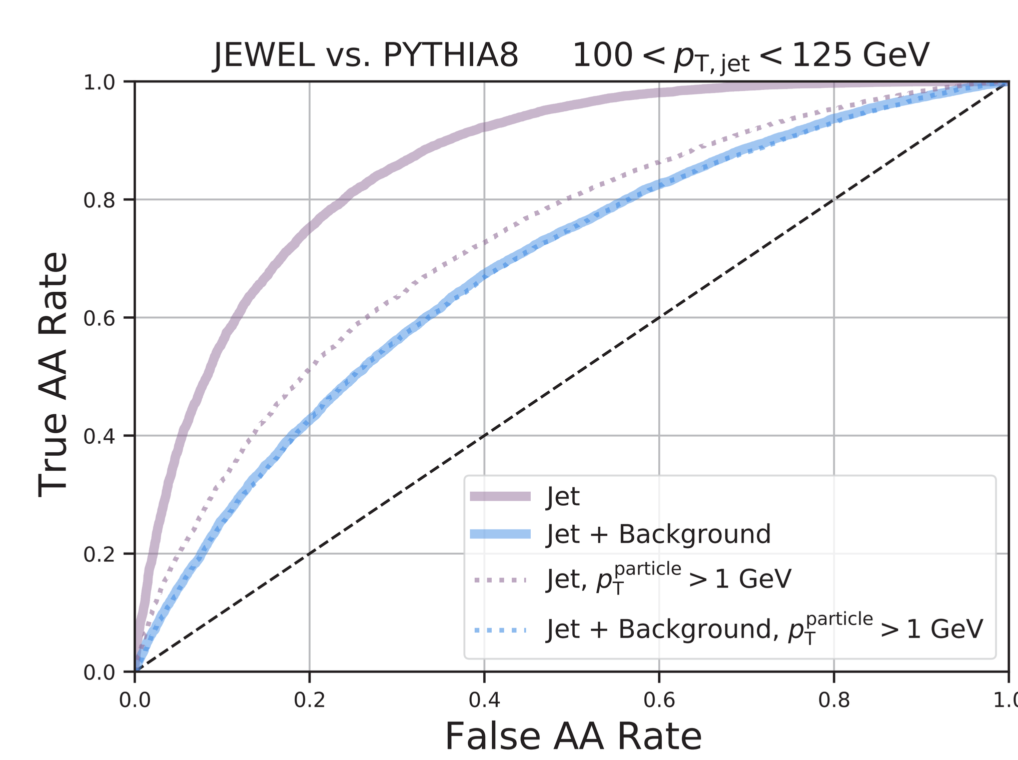

Fig. 5. ROC curves for PFNs trained with PYTHIA8/JEWEL jets without and with a thermal background (with event-wide constituent subtraction), and compared to the performance only considering jet constituents with $p_T > 1$ GeV.

There is a dramatic decrease in the classification power due to the presence of the underlying event. In the presence of background, sufficiently soft discrimination is no longer useful — and the discrimination is dominated by hard physics. Soft information is crucial to maximize discrimination between quenched and unquenched jets, yet the fluctuating underlying event fundamentally prevents much of this information from being accessed.

Conclusions

There is a significant ability to distinguish jets in the absence of the heavy-ion underlying event. Substantial amount of this information resides in IRC-unsafe physics. Using complete sets of IRC safe observables, the performance saturates only when a large number of observables are included, demonstrating that a substantial amount of information is contained in soft emissions inside the jet. The ability to distinguish jets in *AA* from jets in *pp* collisions appears to substantially decrease in the presence of the underlying event. We designed new observables using Lasso regression that maximize the discrimination power between jets in heavy-ion collisions and jets in proton-proton collisions and are generally analytically tractable in perturbative QCD.