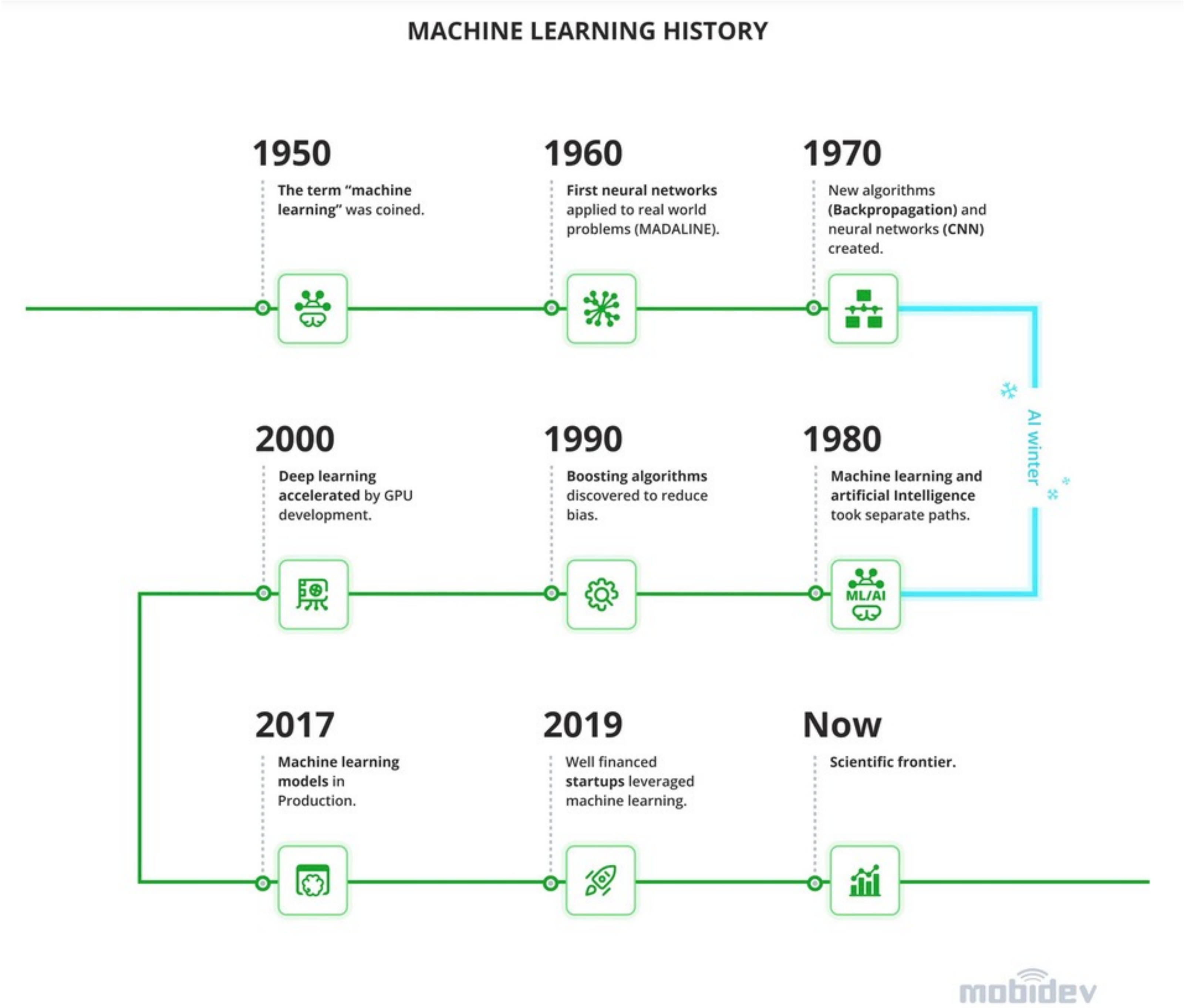
Machine learning approaches to jet quenching

Google engineer put on leave after saying AI chatbot has become sentient

Blake Lemoine says system has perception of, and ability to express thoughts and feelings equivalent to a human child

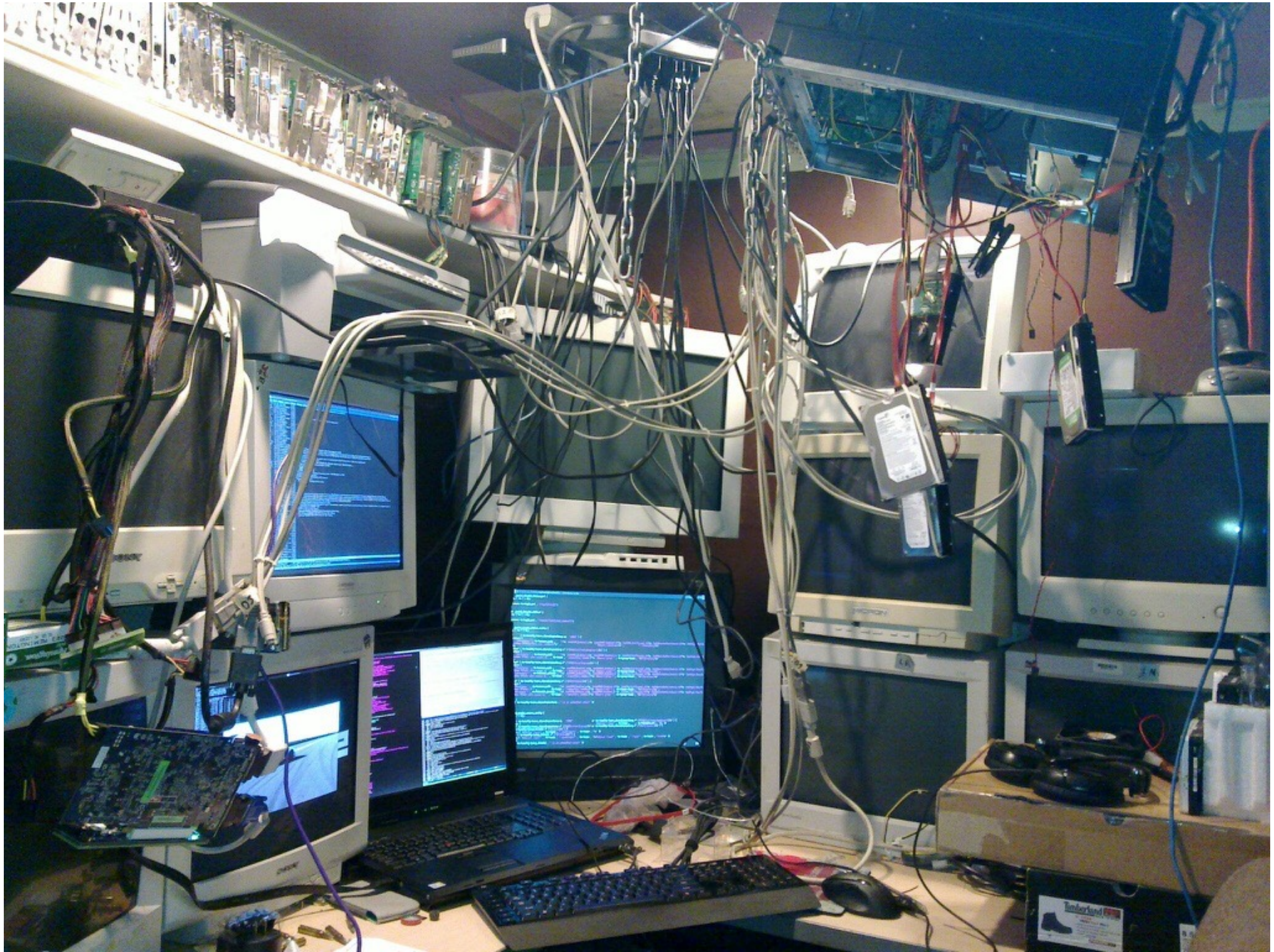
Search
The Guardian
US edition
News website of the year

Revelation has put new scrutiny on the capacity of, and secrecy surrounding, the world of artificial intelligence (AI). Photograph: Boris Roessler/EPA


MACHINE LEARNING HISTORY
1950
The term "machine learning" was coined.
1960
First neural networks applied to real world problems (MADALINE).
1970
New algorithms (Backpropagation) and neural networks (CNN) created.
AI winter
1980
Machine learning and artificial Intelligence took separate paths.
ML/AI
1990
Boosting algorithms discovered to reduce bias.
2000
Deep learning accelerated by GPU development.
2017
Machine learning models in Production.
2019
Well financed startups leveraged machine learning.
Now
Scientific frontier.
mobidev

It’s just the beginning…

But we should rapidly move beyond proof of concepts and generic ML tech – towards physics aware / white box / interpretable AI and data-ready … it requires a conceptual change…



ML for jet quenching: What’s the question(s)?

- Tell me if *this* jet is quenched and I take it from there…
- What to measure to expose jet quenching?
  - ⇔ maximize the information on quenching
- Frequentist vs. statistical inference
  - Simulation based, variational, Bayesian inference
  - Constrain physical model parameters with data (aka fit)
  - Question: how to discover missing features of the model/theory?
- Mitigate background effects?
- Understand uncertainties after ML application – some pitfalls:
  - **ML performance as good as the data set itself**
  - **ML performance as good as the ML model itself**

ML for jet quenching: important problems

- Model dependence – "how to" model independent learning
- Data (preparation) – data level, biases, inefficiencies
- Black box vs. white box ⇔ explain ability
- Applicability – can't ignore* the UE in heavy-ion collisions
- Uncertainty quantification, model biases – applying MC tuned model to data?

A few of using ML
- Infer physics (probability dist. of params)
- Design new observables – optimize exp. sensitivity
- Select/tag-and-study
- Discover features / anomaly detection

Suggested methodology:
- *Learn about the ML model using Monte Carlo*
- *Learn about physics model using data (not MC)*

*) ignoring means pure MC studies of quenching w/o prescription how to do work with data that contain the background effects – see https://arxiv.org/abs/2006.01812 for an example

Nota bene…

- HEP driven ML effort much larger - we are learning (sic!) how to benefit…
- https://arxiv.org/abs/2102.02770
- https://github.com/iml-wg/HEPML-LivingReview

Classification

Parameterized classifiers

Jet images, event images, sequences, trees, graphs, sets (point clouds), physics-inspired basis, W/Z tagging, quarks/gluons, top quark tagging, strange jets, b-tagging, flavor physics, BSM, PID, neutrino detectors, direct DM detectors, cosmology/astro/c.rays, tracking, heavy-ions/NP, hyperparameters, weak/semi supervision, unsupervised, reinforcement learning, quantum ML, feature ranking, attention, regularization, optimal transport, software, hardware/firmware, deployment

Regression

Pileup, Calibration, Recasting, Matrix Elements, Parameter Estimation, Parton Distribution Functions (and related), Lattice Gauge Theory, Function Approximation, Symbolic Regression

Decorrelation methods

Generative models / density estimation

GANs, Autoencoders, Normalizing flows, Physics-inspired, Mixture Models, Phase space generation, Gaussian processes

Anomaly detection

Simulation-based (`likelihood-free') Inference

Parameter estimation, Unfolding, Domain adaptation, BSM, Differentiable Simulation

Uncertainty Quantification

Interpretability, Estimation, Mitigation, Uncertainty- and inference-aware learning

Experimental Results

Performance studies, Searches and measurements were ML reconstruction is a core component, Final analysis discriminate for searches, Measurements using deep learning directly (not through object reconstruction)

Some ML applications – related to jet quenching problems…

Jet pT corrections

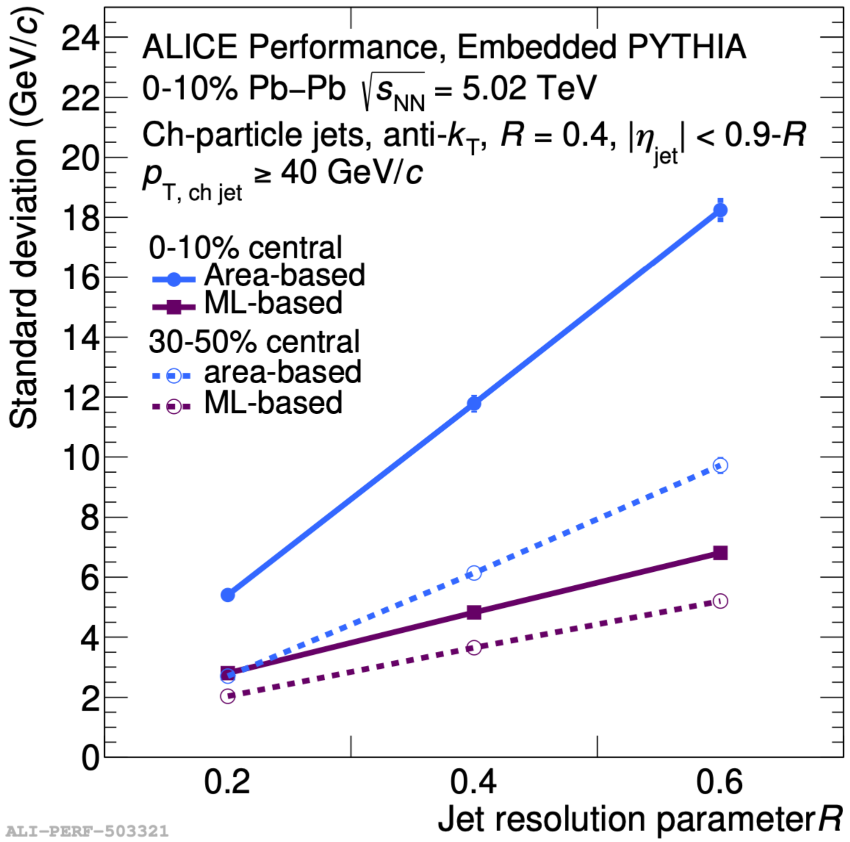
Jet pT corrections

https://arxiv.org/abs/1810.06324 + ALICE
https://arxiv.org/abs/1909.01639

- Modest goal but high gain: improve momentum resolution – enable low-pT large-R jet measurements
- Actual application of ML in AA data
- Regression, Random Forest (decision trees), NN
- Predict jet pT based on structure features – analysis of sensitivity done
- **Model dependence: mitigation with FF variations + uncertainty quantification**

https://indi.to/3mH5h

$$\delta p_{\mathrm{T}} = p_{\mathrm{T,rec}} - p_{\mathrm{T,true}}$$

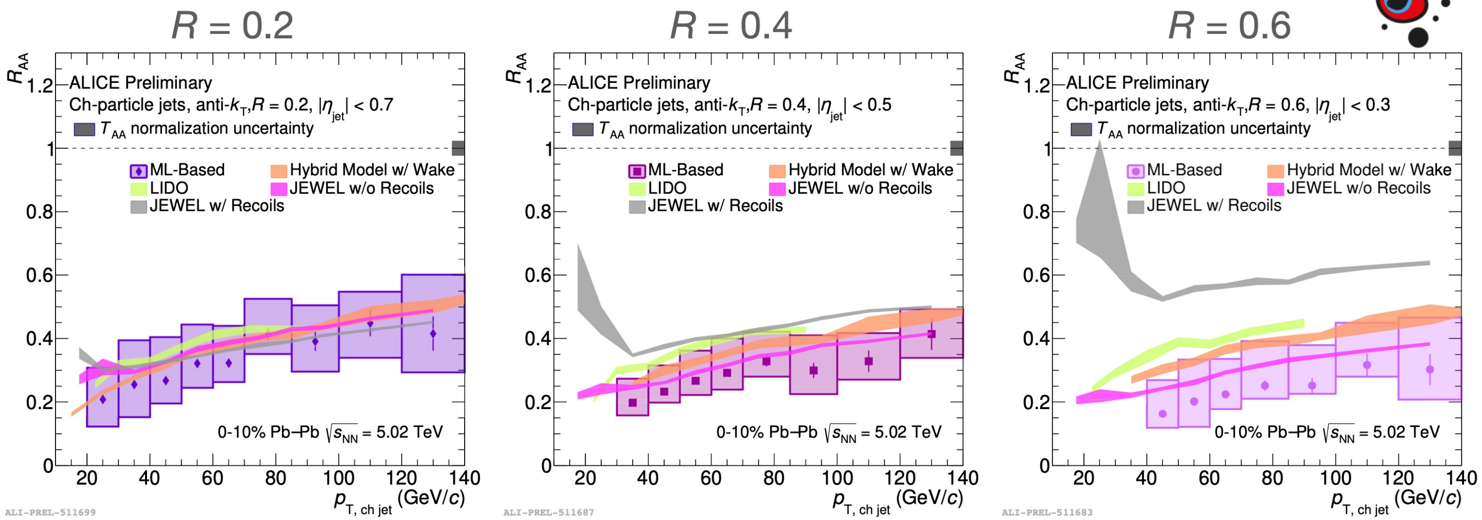


Nuclear modification factors (0-10%)

New for QM

ALICE

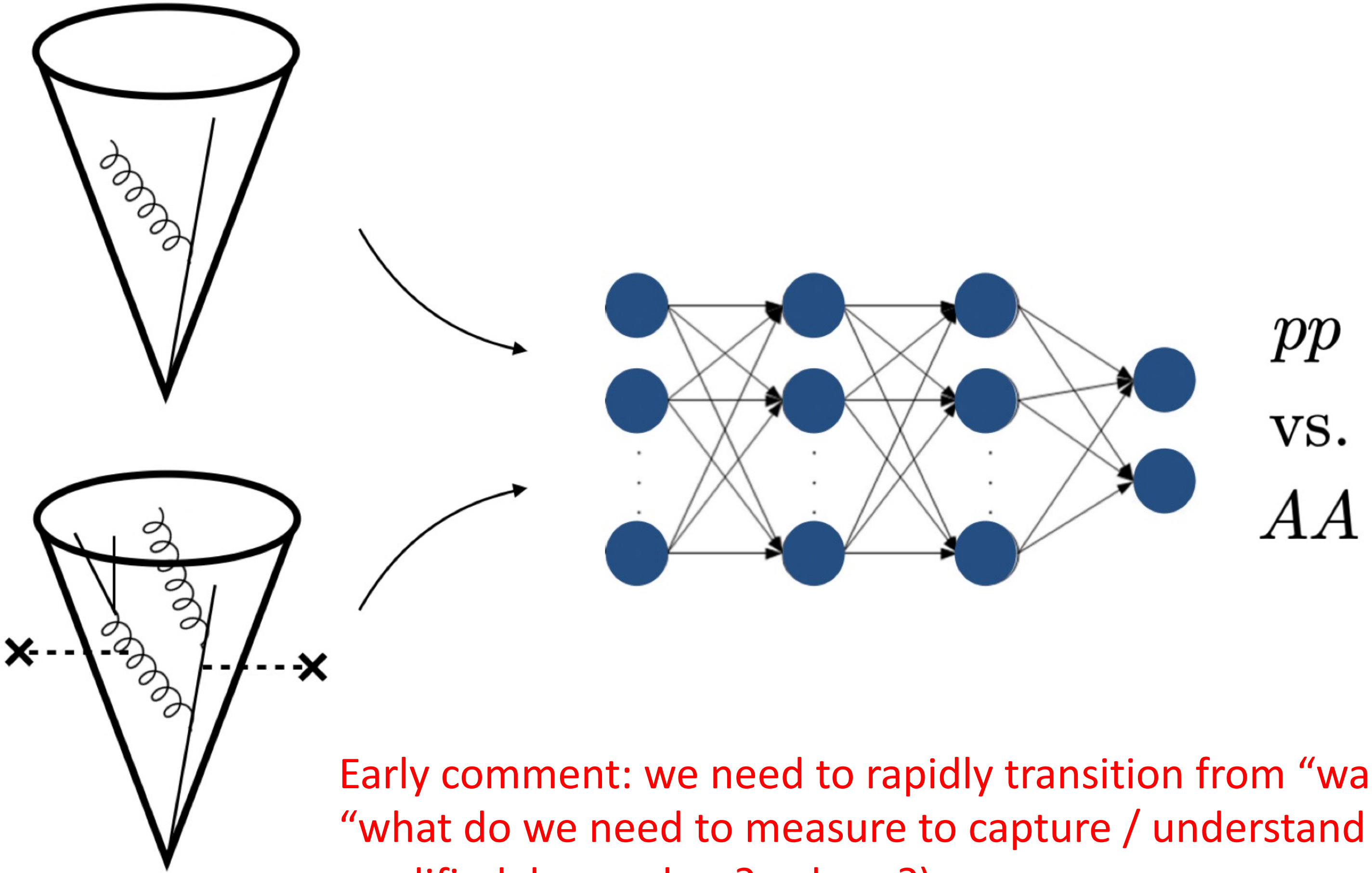
LIDO: JHEP 05 (2021) 041 JEWEL: JHEP 1707 (2017) 141 Hybrid Model: Phys. Rev. Lett. 124, 052301 (2020)



→ Measuring down to lower p_{T} and larger R than ever before in heavy-ions at the LHC!

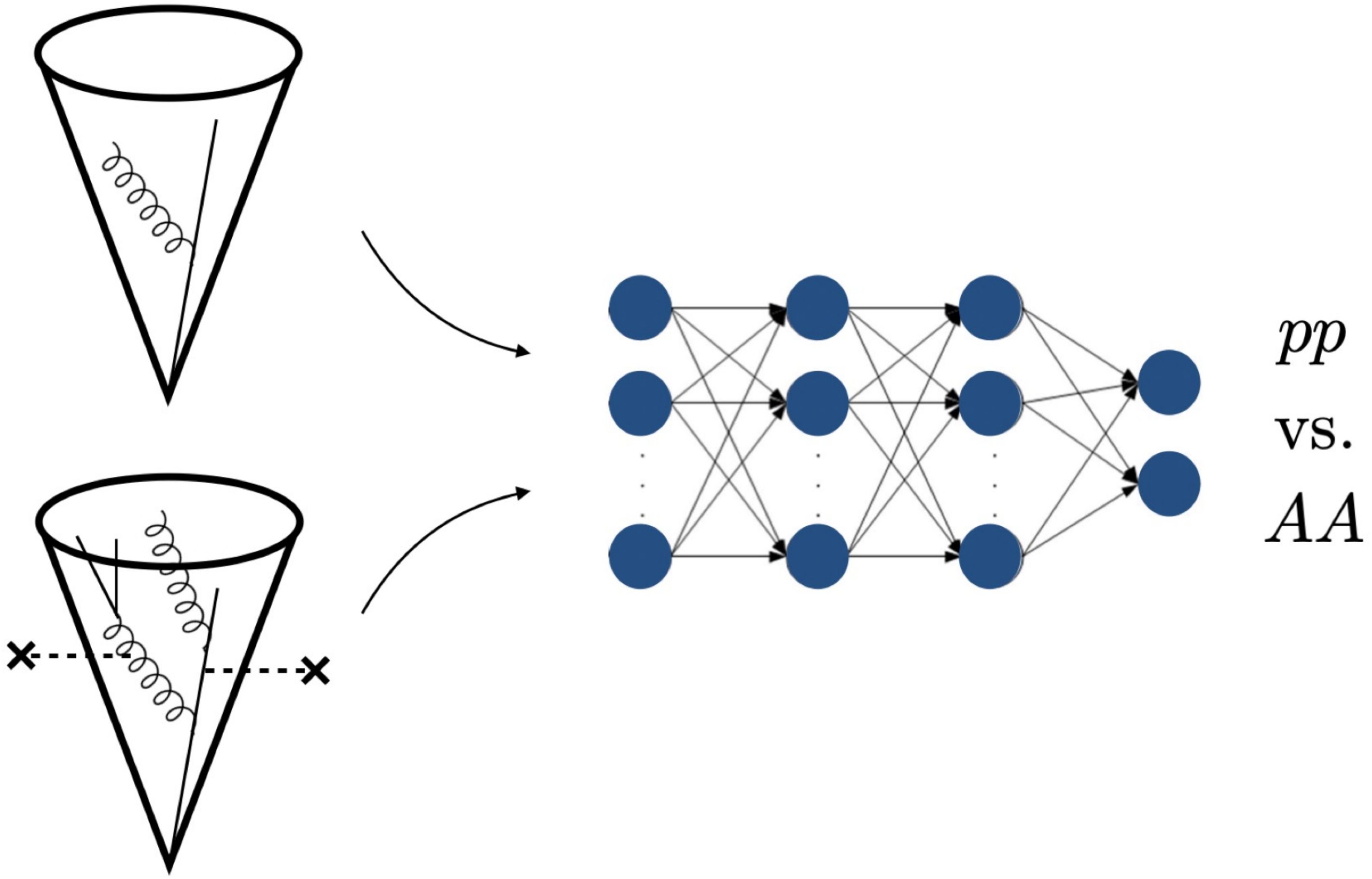
→ Models generally agree with data, but can they describe the R-dependence?

Classification: quenched or not quenched?



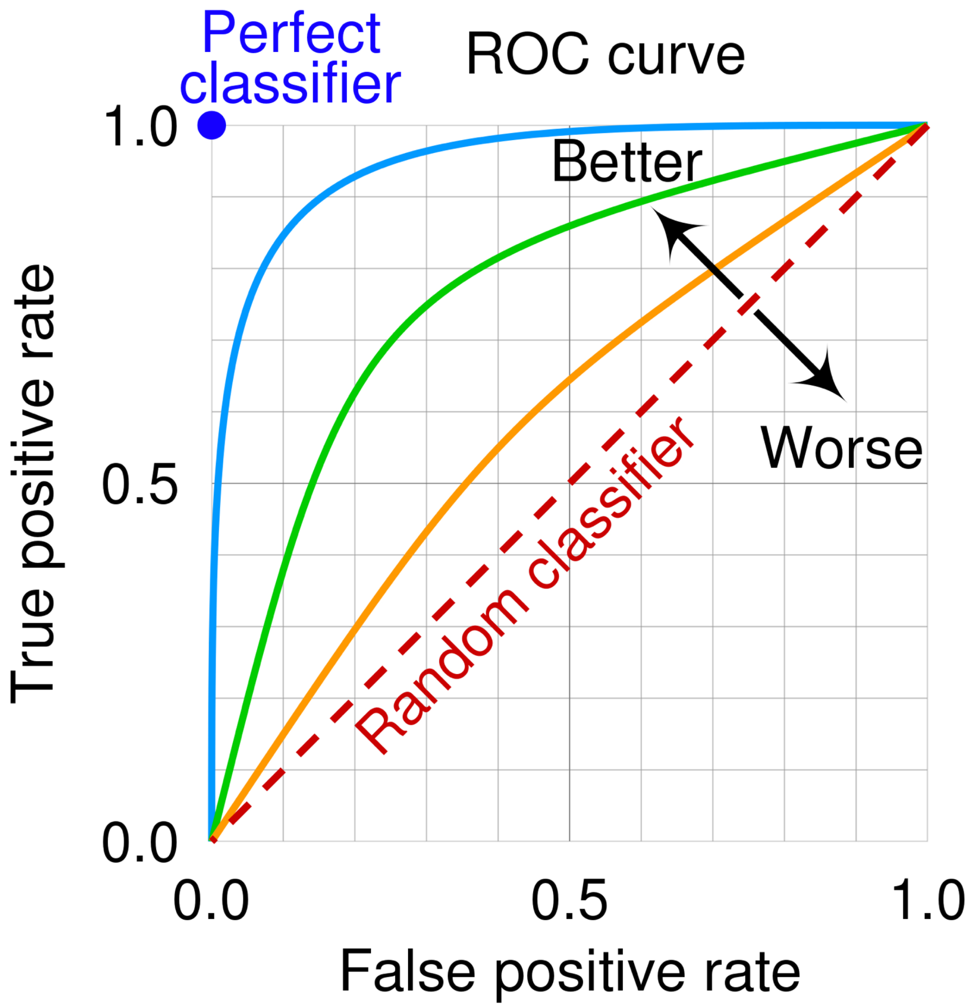
Early comment: we need to rapidly transition from “was the jet quenched or not?” to “what do we need to measure to capture / understand jet quenching?” (what modified, how, when?, where?)

Classification: quenched or not quenched?



Quick reminder…

receiver operating characteristic curve



diagnostic ability of a [binary classifier] system as its discrimination threshold is varied

Yang-Ting Chien, Raghav Kunnawalkam Elayavalli

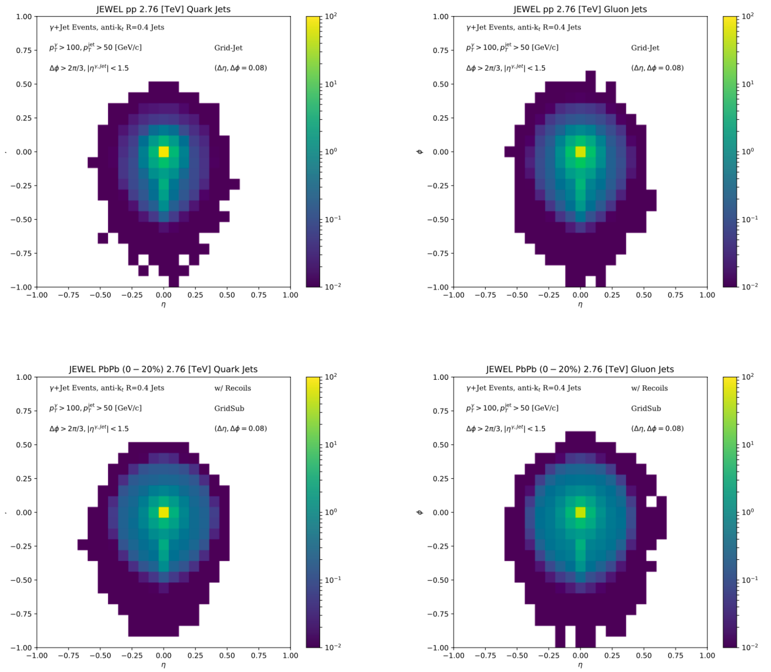
Discrimination: quark vs. gluon jets…

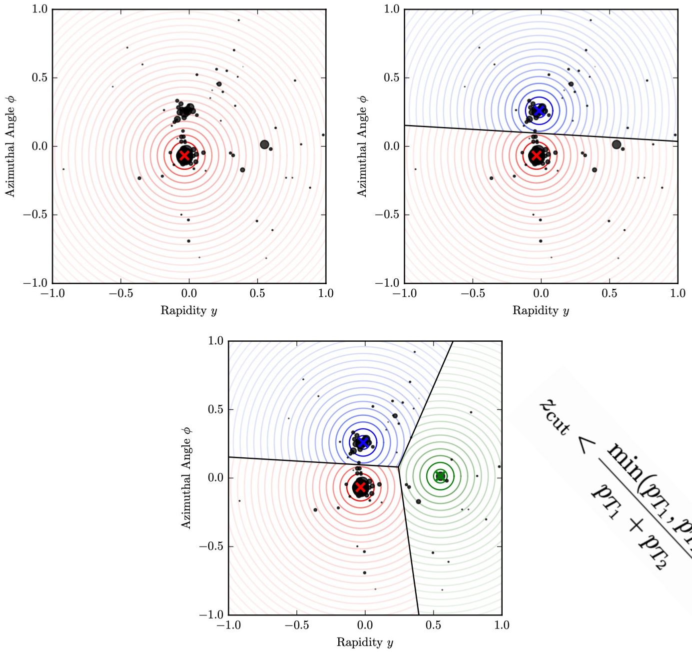
https://arxiv.org/abs/1803.03589

JEWEL
Background: ON/OFF
pT > 50 GeV/c
POP study

- Probing heavy ion collisions using quark and gluon jet substructure
- Multi-layer perceptron (MLP) with Jet mass, two radial moments including the girth, the $p_{T,D}$, and the pixel multiplicity
- Deep convolutional neural network (DCNN) on discretized images of quark jets and gluon jets (η,ϕ)
- Telescoping deconstruction framework exploiting subjet kinematics – pT, mass (use MLP)
- **"We find that the quark gluon discrimination performance worsens in heavy ion jets due to significant soft event activity affecting the soft jet substructure."**

https://arxiv.org/abs/1310.7584
- HEP related





$$z_{\mathrm{cut}} < \frac{\min(p_{T_1}, p_{T_2})}{p_{T_1} + p_{T_2}} \equiv z_{T_3}$$

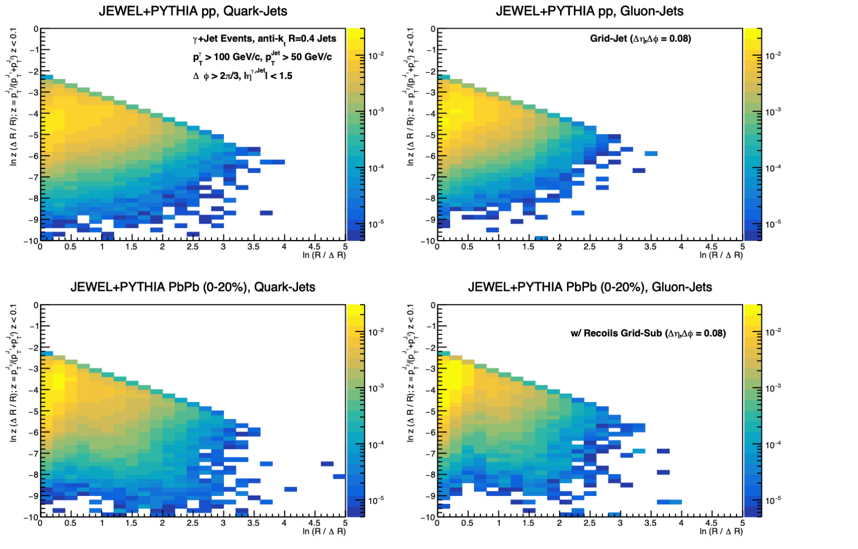


Figure 11: Lund diagrams for branches removed in the soft-drop procedure with $z < z_{\mathrm{cut}} = 0.1$ for quark (left) and gluon (right) jets in pp (top) and central PbPb collisions in JEWEL. See text for the definition of the axes.

Quark vs. gluon jets…

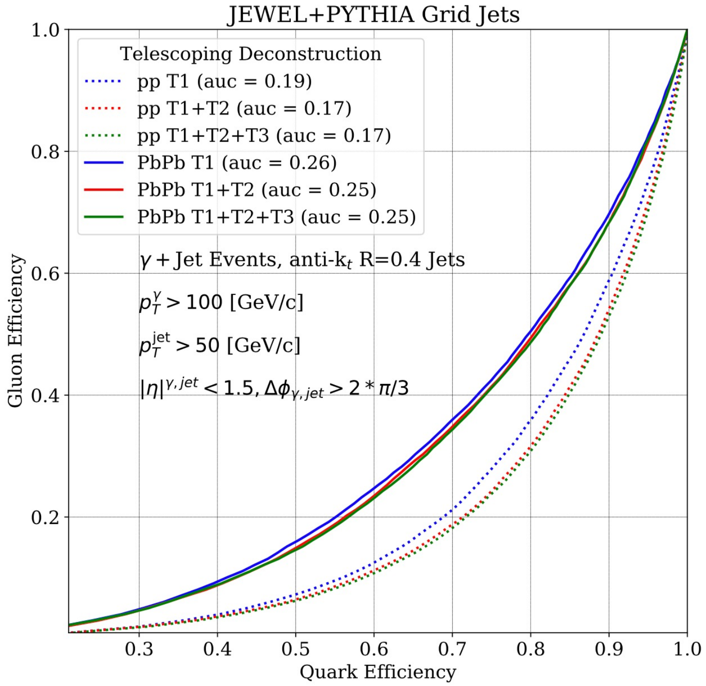
Yang-Ting Chien, Raghav Kunnawalkam Elayavalli

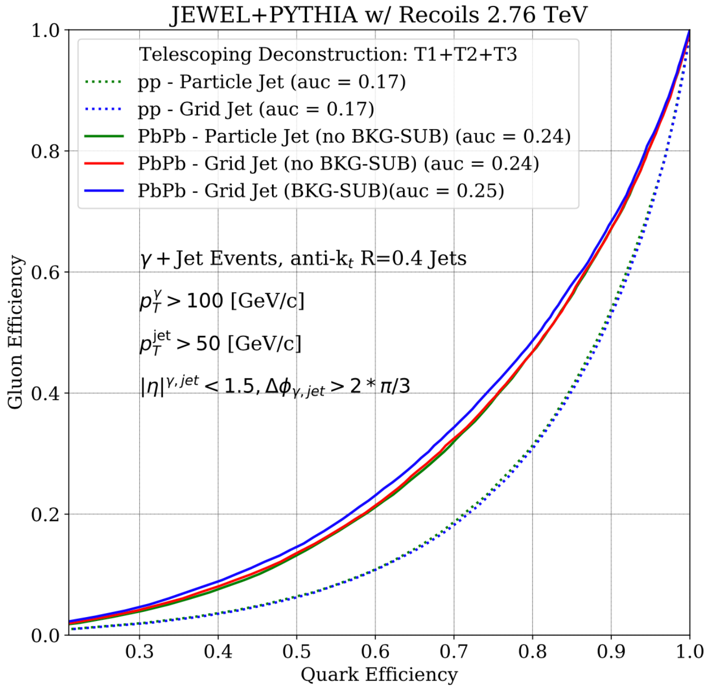
https://arxiv.org/abs/1803.03589

JEWEL
Background: ON/OFF
pT > 50 GeV/c
POP study

- Probing heavy ion collisions using quark and gluon jet substructure
- Multi-layer perceptron (MLP) with Jet mass, two radial moments including the girth, the $p_{T,D}$, and the pixel multiplicity
- Deep convolutional neural network (DCNN) on discretized images of quark jets and gluon jets (η,ϕ)
- Telescoping deconstruction framework exploiting subjet kinematics – pT, mass (use MLP)
- **"We find that the quark gluon discrimination performance worsens in heavy ion jets due to significant soft event activity affecting the soft jet substructure.“**

https://arxiv.org/abs/1310.7584
- HEP related





"Through multiple methods and observables, we **consistently find the dominant feature of the Jewel jet quenching model to be the increase of soft particle multiplicity due to medium recoils throughout the jet region.** This is closely related to the loss of information in subleading subjets, which is a characteristic feature of Jewel."

"The CNN architecture has not been tuned exhaustively, therefore its ROC curves serve to give a general sense of performance. "

Yang-Ting Chien, Raghav Kunnawalkam Elayavalli

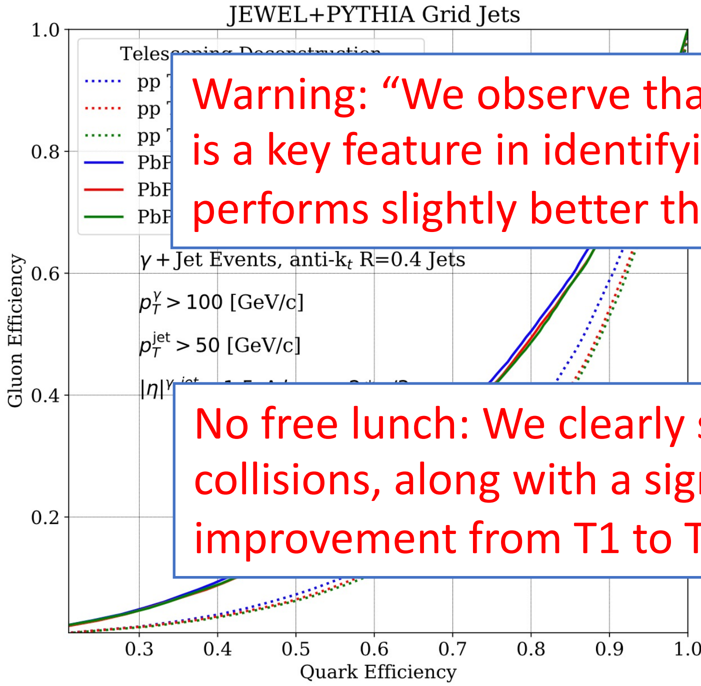
https://arxiv.org/abs/1803.03589

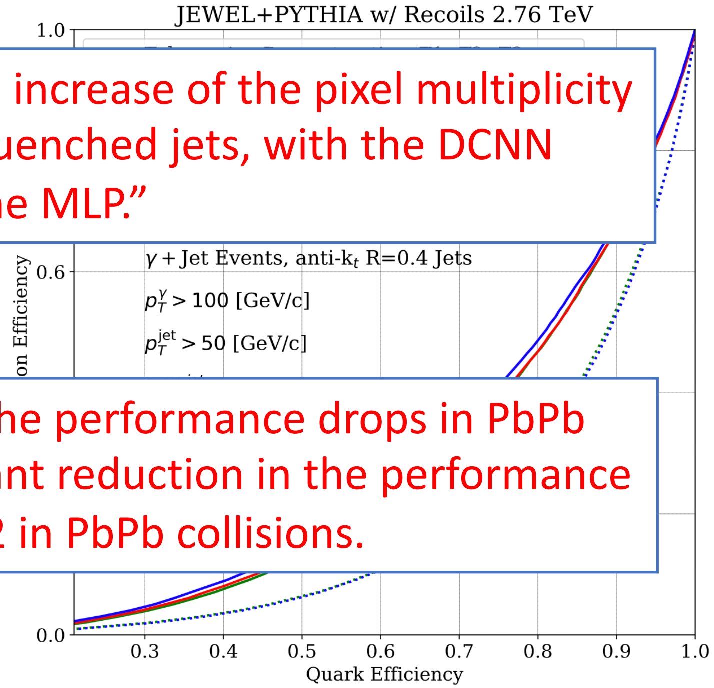
Quark vs. gluon jets…

JEWEL
Background: ON/OFF
pT > 50 GeV/c
POP study

- Probing heavy ion collisions using quark and gluon jet substructure
- Multi-layer perceptron (MLP) with Jet mass, two radial moments including the girth, the $p_{T,D}$, and the pixel multiplicity
- Deep convolutional neural network (DCNN) on discretized images of quark jets and gluon jets (η,ϕ)
- Telescoping deconstruction framework exploiting subjet kinematics – pT, mass (use MLP)
- **"We find that the quark gluon discrimination performance worsens in heavy ion jets due to significant soft event activity affecting the soft jet substructure."**

https://arxiv.org/abs/1310.7584
- HEP related





Warning: "We observe that the increase of the pixel multiplicity is a key feature in identifying quenched jets, with the DCNN performs slightly better than the MLP."

No free lunch: We clearly see the performance drops in PbPb collisions, along with a significant reduction in the performance improvement from T1 to T1+T2 in PbPb collisions.

"Through multiple methods and observables, we **consistently find the dominant feature of the Jewel jet quenching model to be the increase of soft particle multiplicity due to medium recoils throughout the jet region.** This is closely related to the loss of information in subleading subjets, which is a characteristic feature of Jewel."

"The CNN architecture has not been tuned exhaustively, therefore its ROC curves serve to give a general sense of performance. "

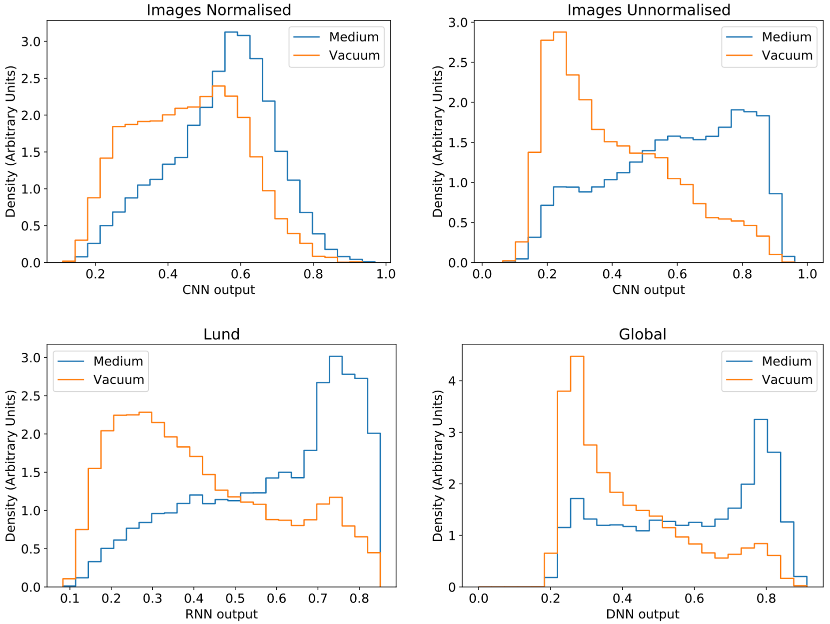
Deep Learning for the Classification of Quenched Jets

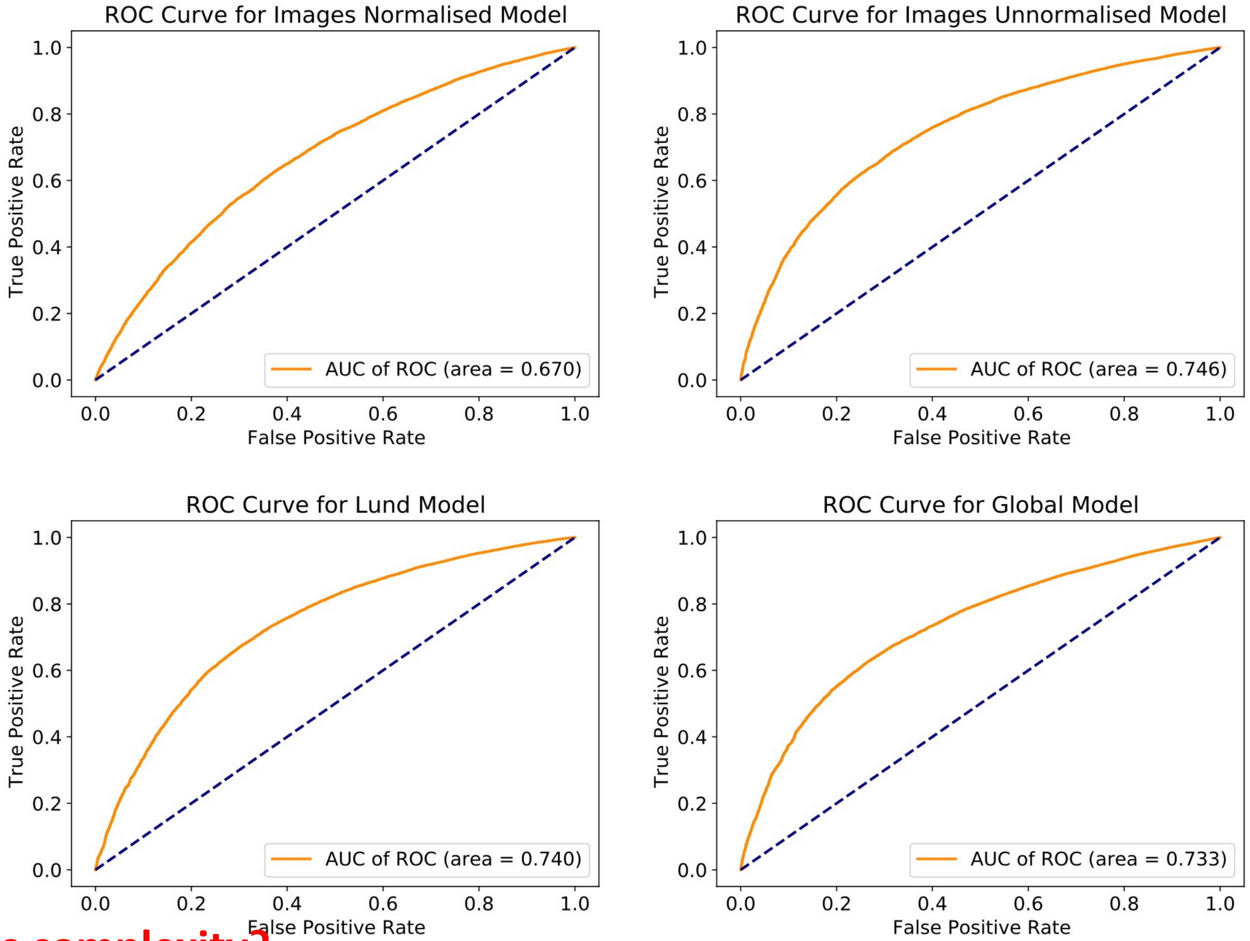
https://arxiv.org/abs/2106.08869

L. Apolinário, N. F. Castro, M. Crispim Romão, J. G. Milhano, R. Pedro, F. C. R. Peres

JEWEL
Background: OFF
High-pT
POP study

- Images: Convolutional Neural Network (CNN) for the jet (η,ϕ) images. In addition, we further considered the case that the image channels were *normalised* or left *unnormalised*
- Lund: Recurrent Neural Network (RNN) for the sequence of the C/A re-clustered sequence of the primary Lund plane coordinates
- Global: Dense Neural Network (DNN) for the tabular data of the global jet transverse momentum and the number of constituents, (pT,jet,nconst)





Comments:

- **Do not throw away data – secondary Lund plane? – adds complexity?**
- **Studies w/o background – unrealistic performance?**
- **Low-pT jets more interesting than high-pT jets?**

Deep Learning for the Classification of Quenched Jets

https://arxiv.org/abs/2106.08869

L. Apolinário, N. F. Castro, M. Crispim Romão, J. G. Milhano, R. Pedro, F. C. R. Peres

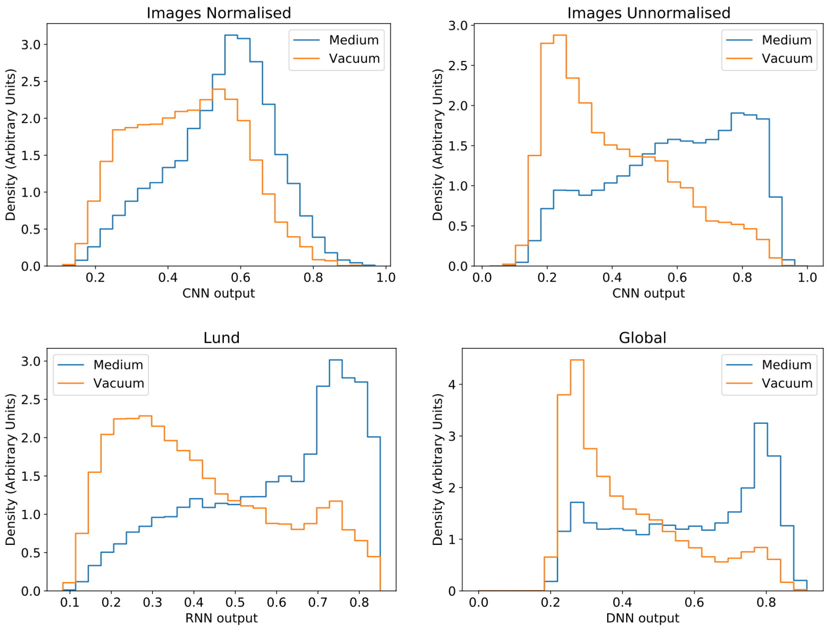
JEWEL

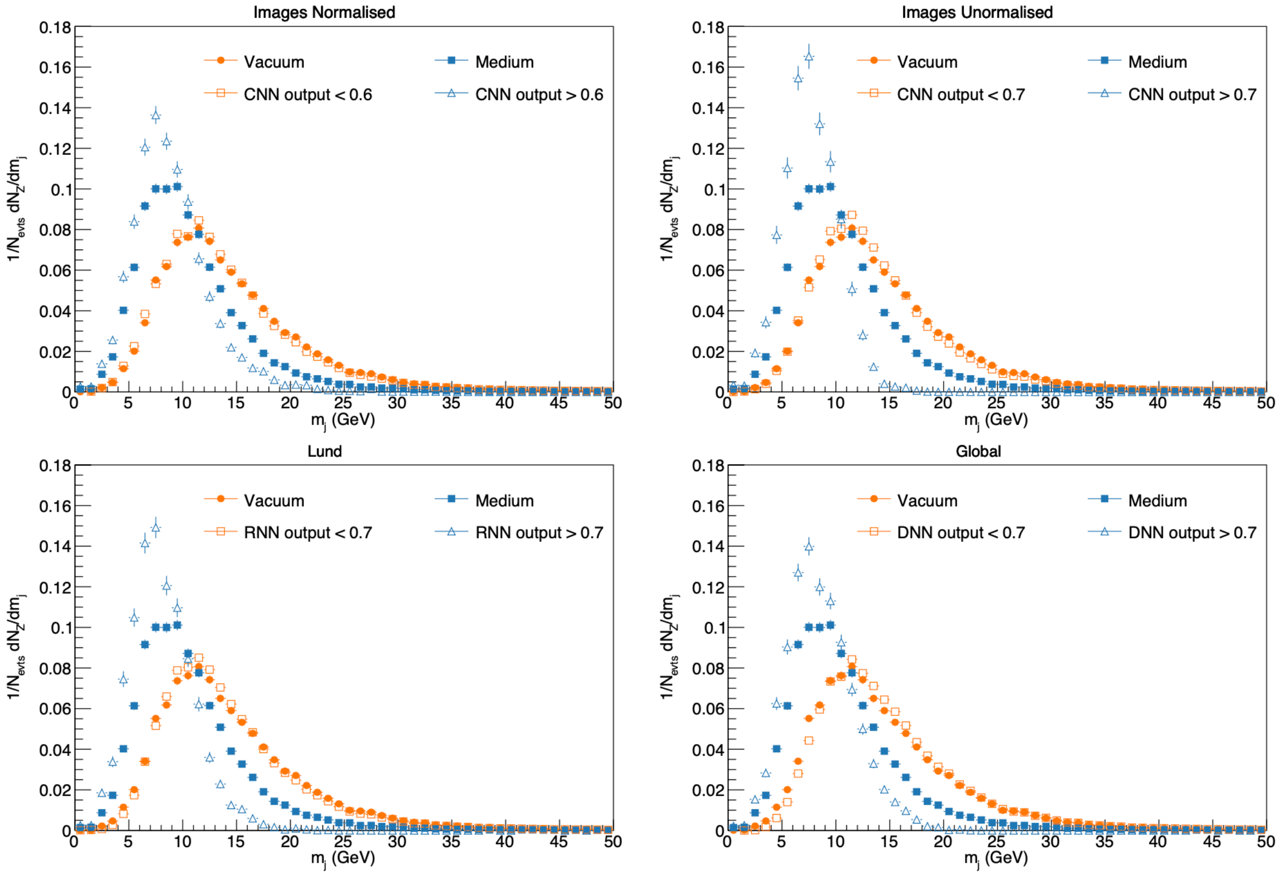
Background: OFF

High-pT

POP study

- Images: Convolutional Neural Network (CNN) for the jet (η,ϕ) images. In addi- tion, we further considered the case that the image channels were *normalised* or left *unnormalised*
- Lund: Recurrent Neural Network (RNN) for the sequence of the C/A re-clustered sequence of the primary Lund plane coordinates
- Global: Dense Neural Network (DNN) for the tabular data of the global jet transverse momentum and the number of constituents, (pT,jet,nconst)





Comments:

- **Do not throw away data – secondary Lund plane? – adds complexity?**
- **Studies w/o background – unrealistic performance?**
- **Low-pT jets more interesting than high-pT jets?**

Quenched jet tagging

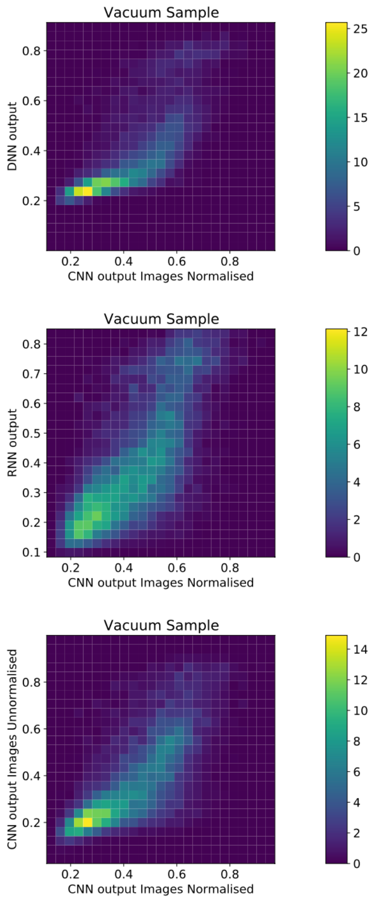
https://arxiv.org/abs/2106.08869

L. Apolinário, N. F. Castro, M. Crispim Romão, J. G. Milhano, R. Pedro, F. C. R. Peres

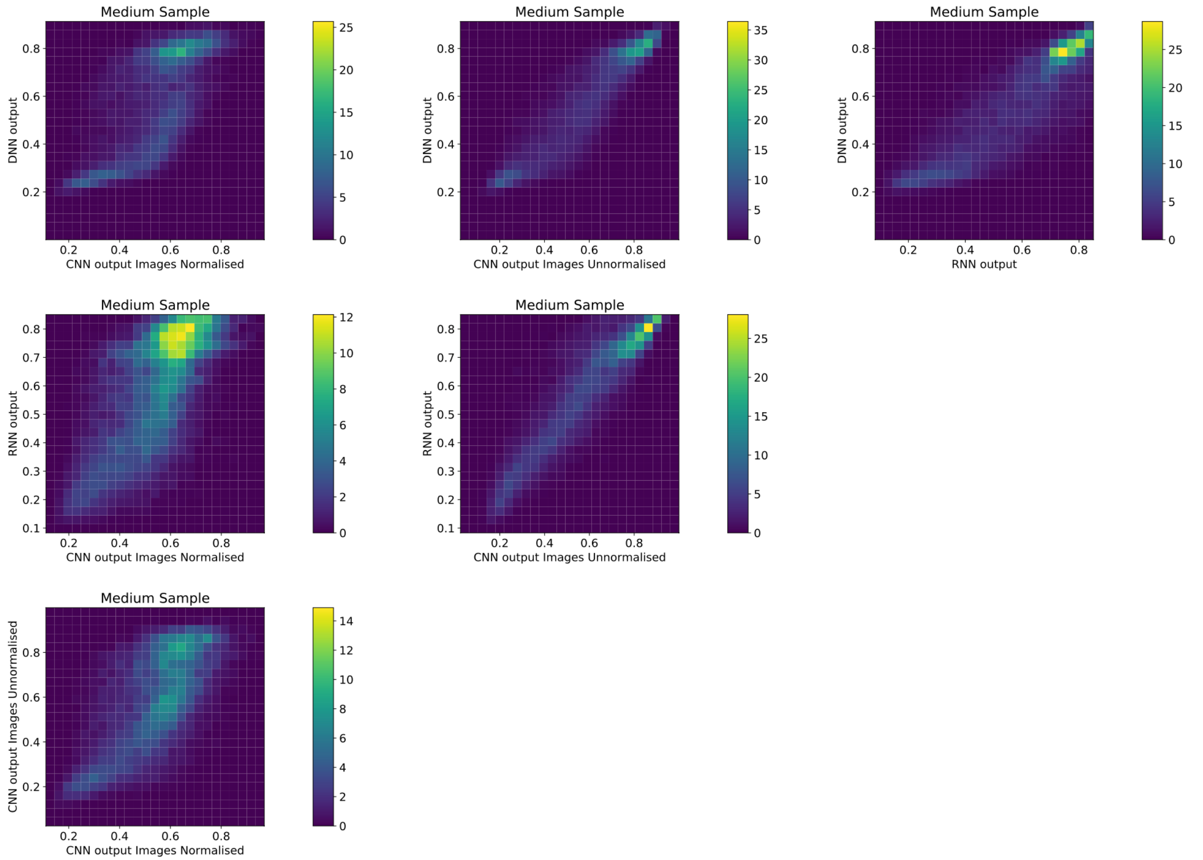
JEWEL
Background: OFF
High-pT
POP study

Interesting appendix! A Correlation between Deep Neural Networks

VACUUM



MEDIUM



=> What does it mean!? => INTERPRETABILITY!

Deep learning jet modifications in heavy-ion collisions

https://arxiv.org/abs/2012.07797

Yi-Lun Du, Daniel Pablos, Konrad Tywoniuk

- convolutional neural network (CNN) to diagnose modifications from **jet images** where the training and validation is performed using the **hybrid strong/weak coupling model**
- the **angular distribution of soft particles** in the jet cone and their relative contribution to the total jet energy **contain significant discriminating power**, which can be exploited to **tailor observables** that provide a good estimate of the energy loss ratio => study a set of jet observables
    - Mass, jet shape ρ(r), fragmentation distribution D
    - Groomed (SD) zg, Rg, nSD, Mg
- potential of deep learning techniques in the analysis of the **geometrical aspects** of jet quenching such as the in-medium **traversed length** or the position of the hard scattering in the transverse plane, opening up new possibilities for **tomographic studies**

Study:

$$\chi_{jh} \equiv \frac{E^h_f}{E^h_i}$$

medium / vacuum

Defined matching procedure for vacuum-medium

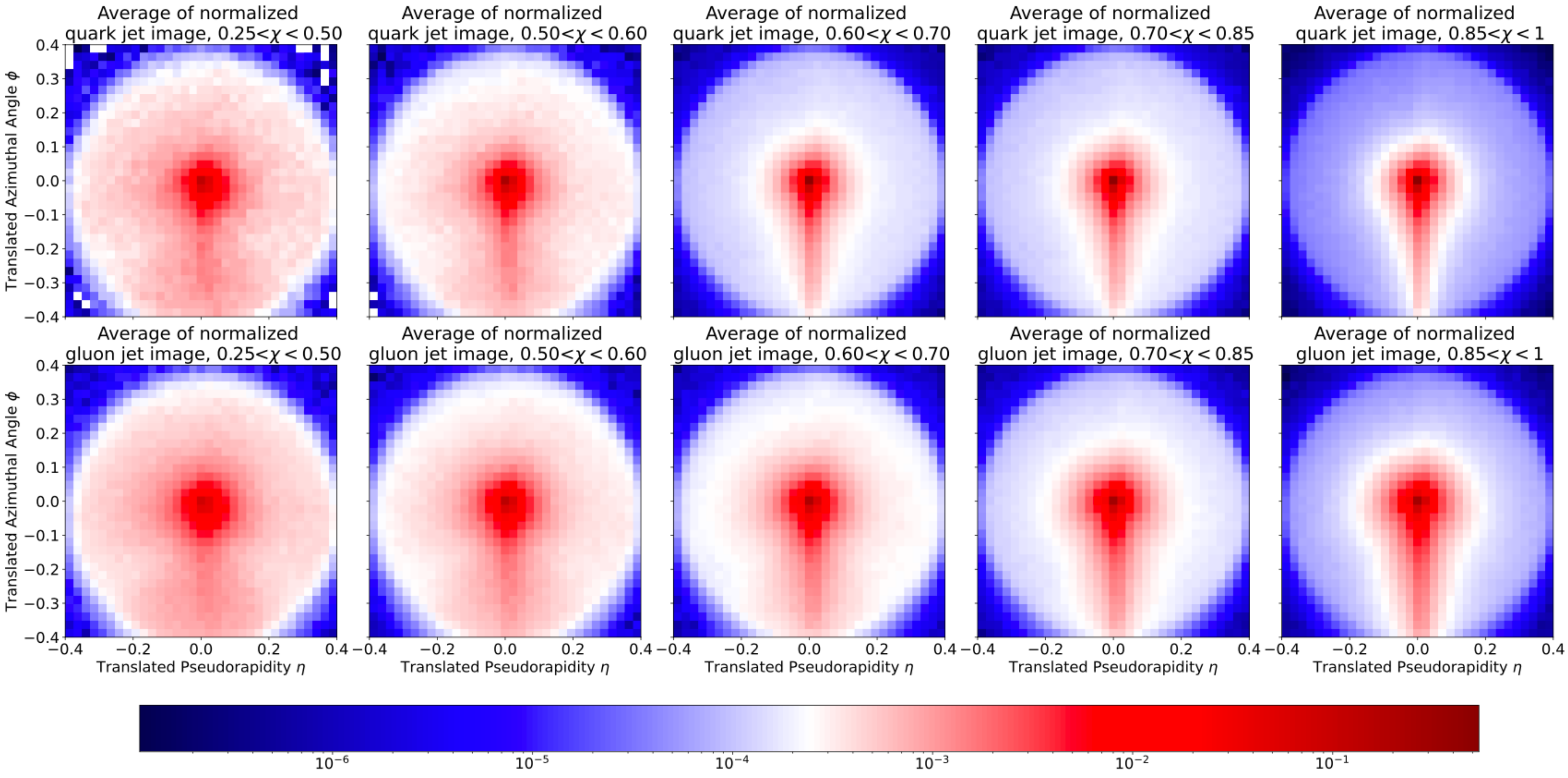
$$L = \frac{\sum_{i\in\mathrm{jet}} p_{T,i} L_i}{\sum_{i\in\mathrm{jet}} p_{T,i}}$$

i – enumerates partons
Jet traversed L := pT weighted sum of individual partons in a jet

Classification of q/g jets in hot QCD medium with deep learning

https://arxiv.org/abs/2012.07797

Yi-Lun Du, Daniel Pablos, Konrad Tywoniuk



Classification of q/g jets in hot QCD medium with deep learning

https://arxiv.org/abs/2112.00681
Yi-Lun Du, Daniel Pablos, Konrad Tywoniuk

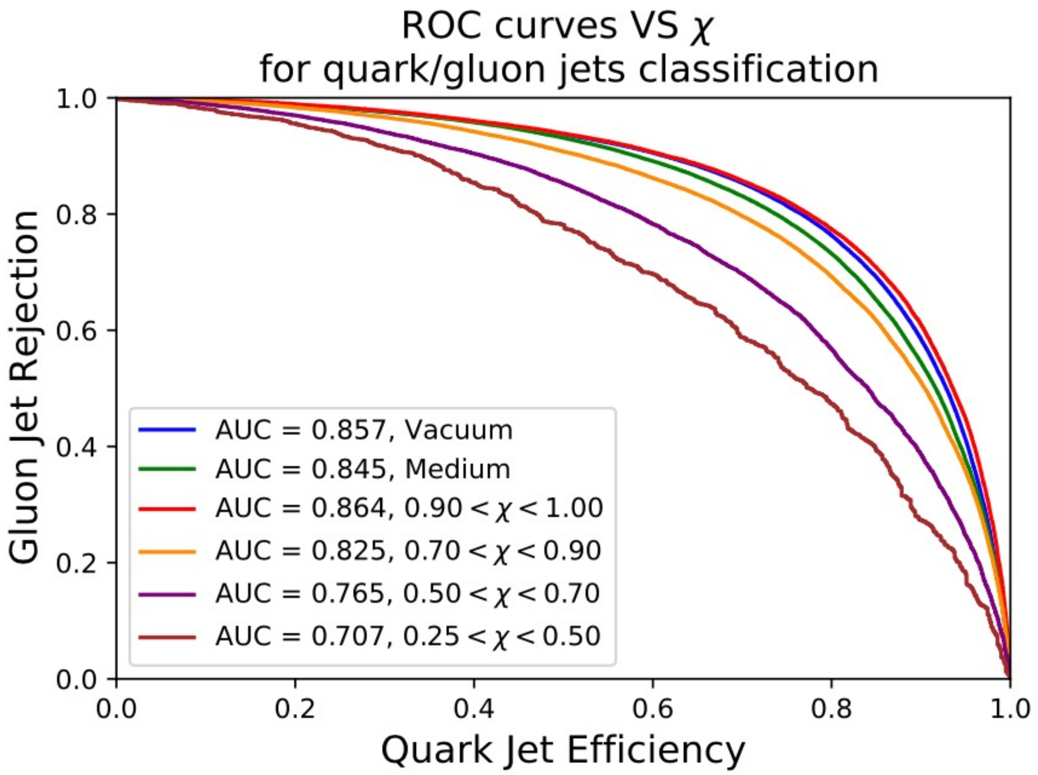


Figure 4: ROC curves of quark efficiency versus gluon rejection for jets in pp collisions and jets in PbPb collisions for inclusive and sliced in χ samples.

Input (size)	Accuracy
Jet shape (8)	72.2%
JFF (10)	73.0%
Jet features (7)	73.6%
JFF, jet shape (18)	74.9%
JFF, jet shape, features (25)	75.8%
Jet image (33×33)	75.9%

Table 1: Classification performance with different inputs. Jet features include: jet p_T, z_g, n_{SD}, R_g, M, M_g, Multiplicity.

"It has been found that the greater the energy loss is, the more difficult it is to classify the jets."

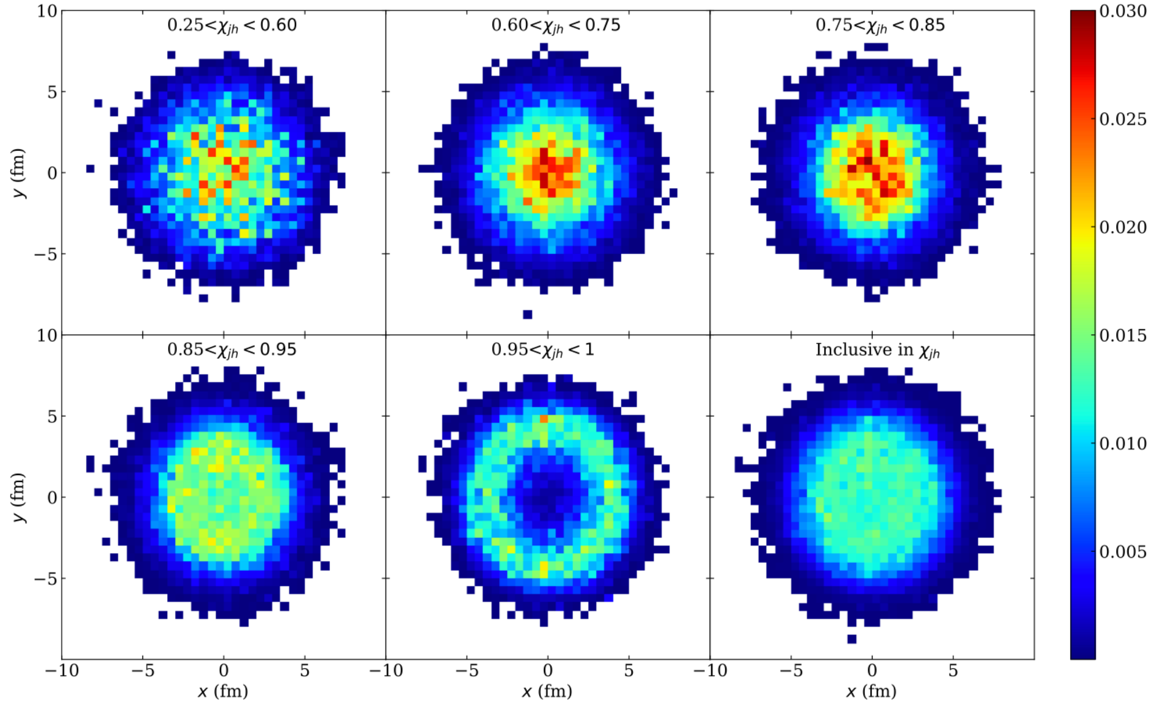
Classification of q/g jets in hot QCD medium with deep learning

https://arxiv.org/abs/2012.07797

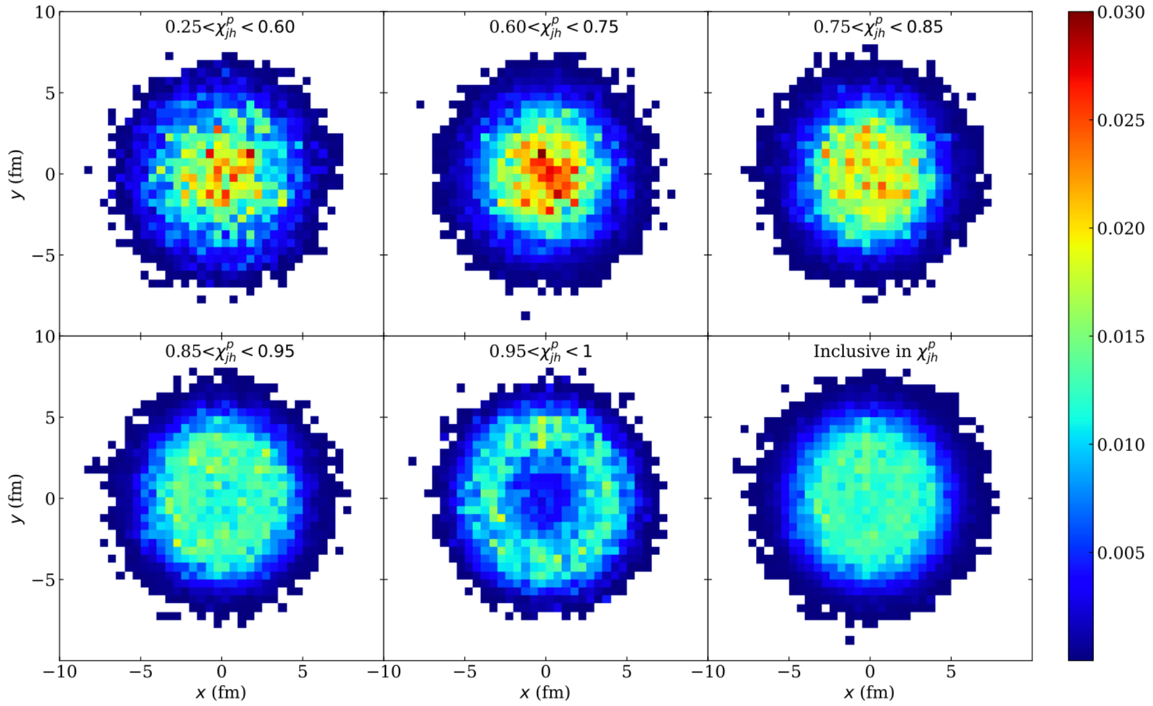
Yi-Lun Du, Daniel Pablos, Konrad Tywoniuk

$$\chi_{jh} \equiv \frac{E_f^h}{E_i^h}$$

Scattering points for true value χjh



Scattering points for **predicted** value χjh

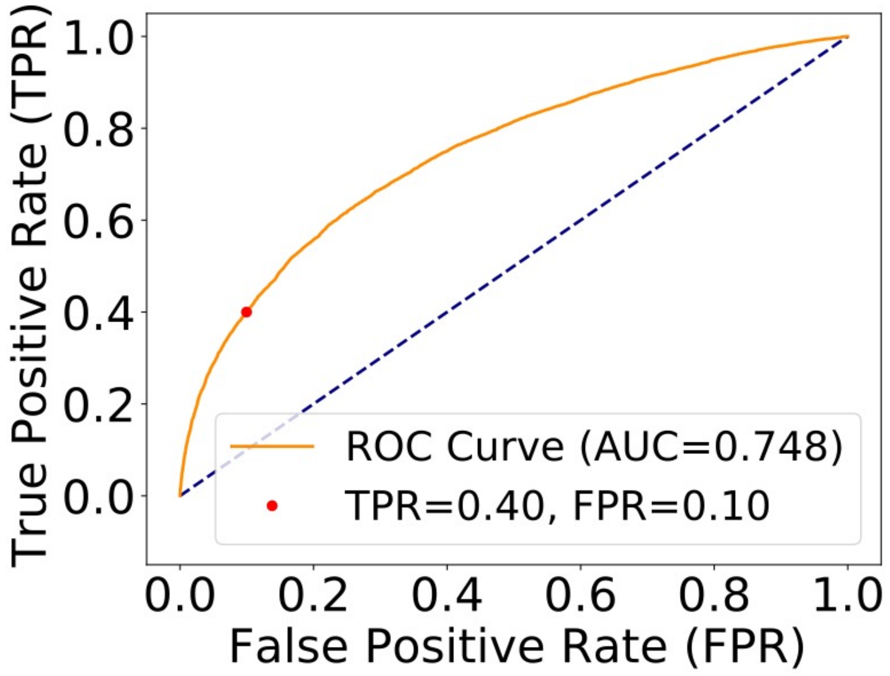


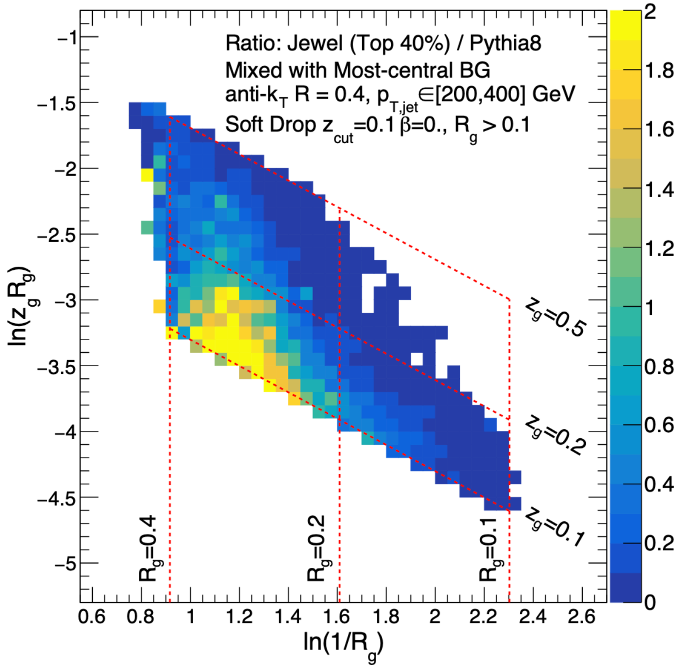
Identification of quenched jets…

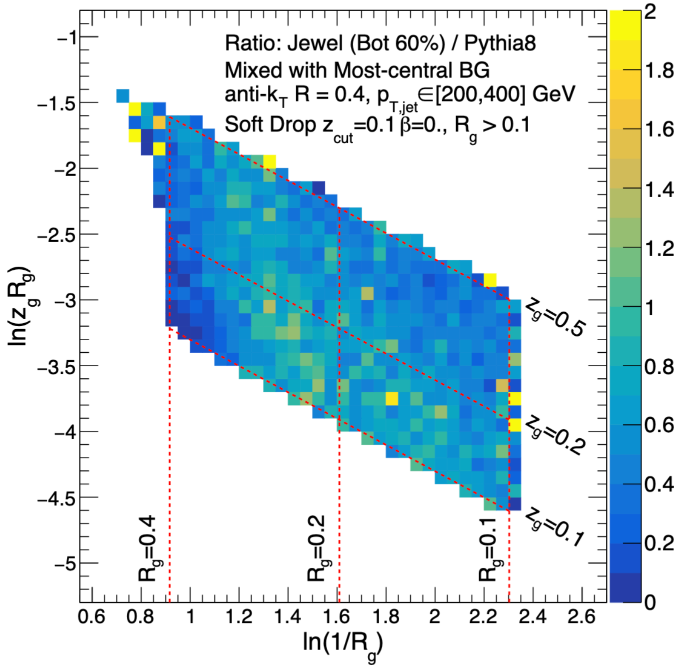
https://arxiv.org/abs/2206.01628
Lihan Liu, Marta Verweij, Julia Velkovska

JEWEL, PYTHIA8
Background: ON
Constit. Subtr.: ON
pT > 200 GeV/c
Supervised ML
POP study

- Deconstruct the clustering sequence (C/A)
- Study the **sequence** of splitting's (feature set: z, θ, k_T …)
- Long short-term memory network => **classify jets: quenched or not**
- Concern: path to application in data, residual background



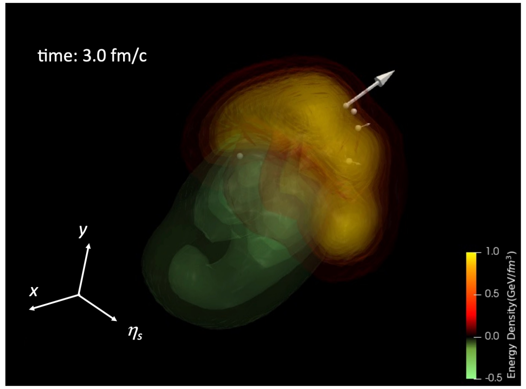




https://arxiv.org/abs/2206.02393

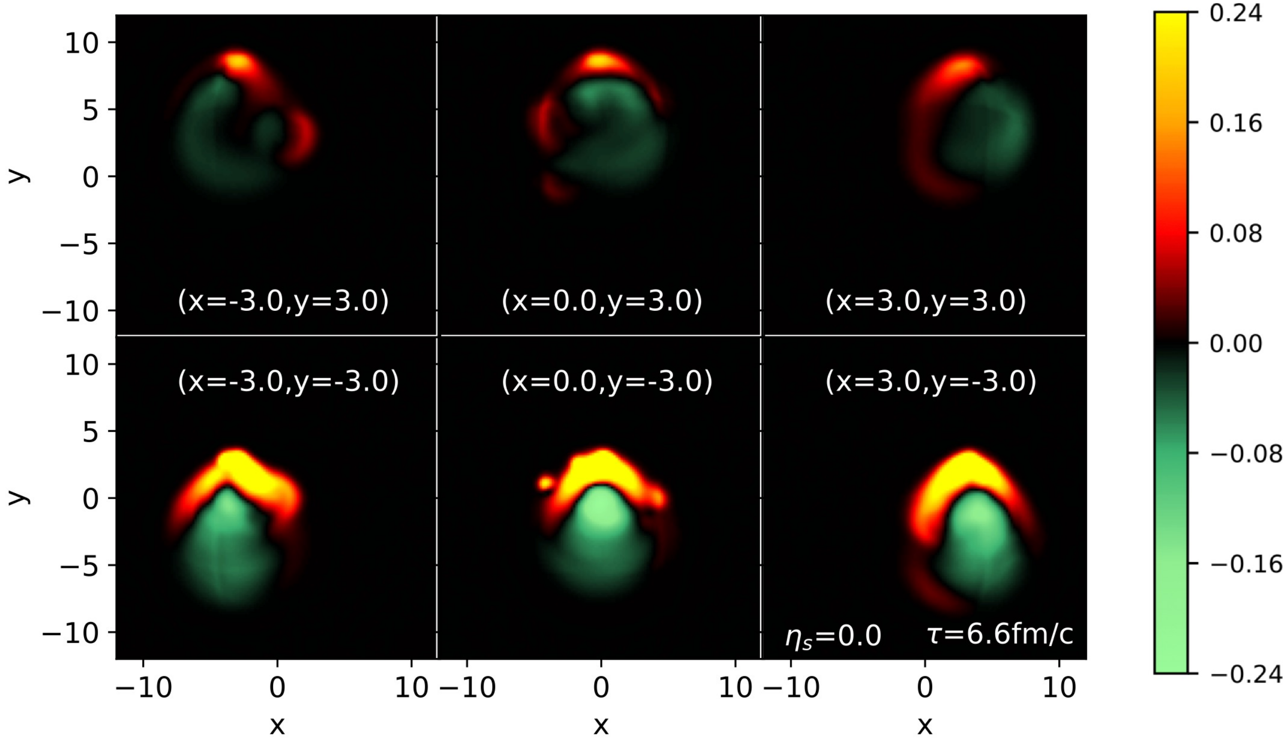
Jets and selected geometry

Zhong Yang, Yayun He, Wei Chen, Wei-Yao Ke, Long-Gang Pang, Xin-Nian Wang

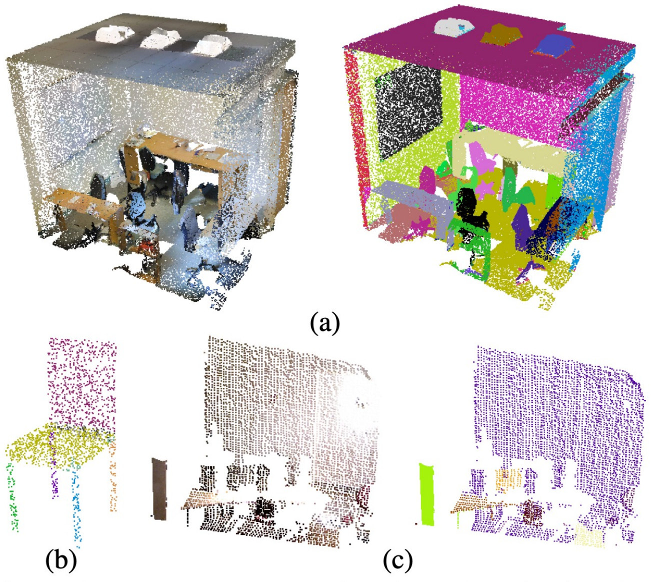


- Deep learning assisted jet tomography for the study in QGP
  - Is it possible to determine the initial jet production positions using the momenta of final state particles?
  - Will the signal of Mach cones and the diffusion wake be amplified if initial jet production positions are constrained to specific regions with long jet propagation lengths or fixed propagation direction relative to the radial flow?
  - How reliable is the new deep learning assisted method?

https://arxiv.org/abs/1711.08588



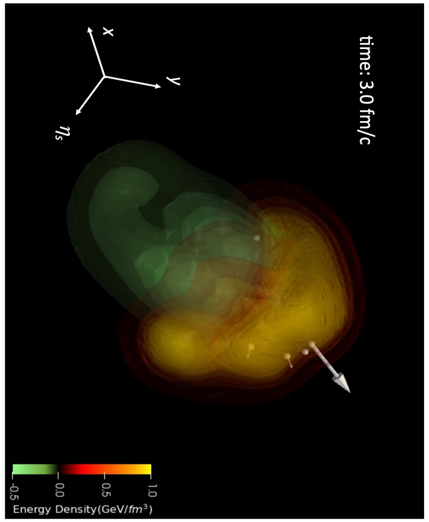
CoLBT-hydro, LIDO (train)
Background: ON
Constit. Subtr.: ON
pT > 100 GeV/c
Supervised ML
POP study



https://arxiv.org/abs/2206.02393

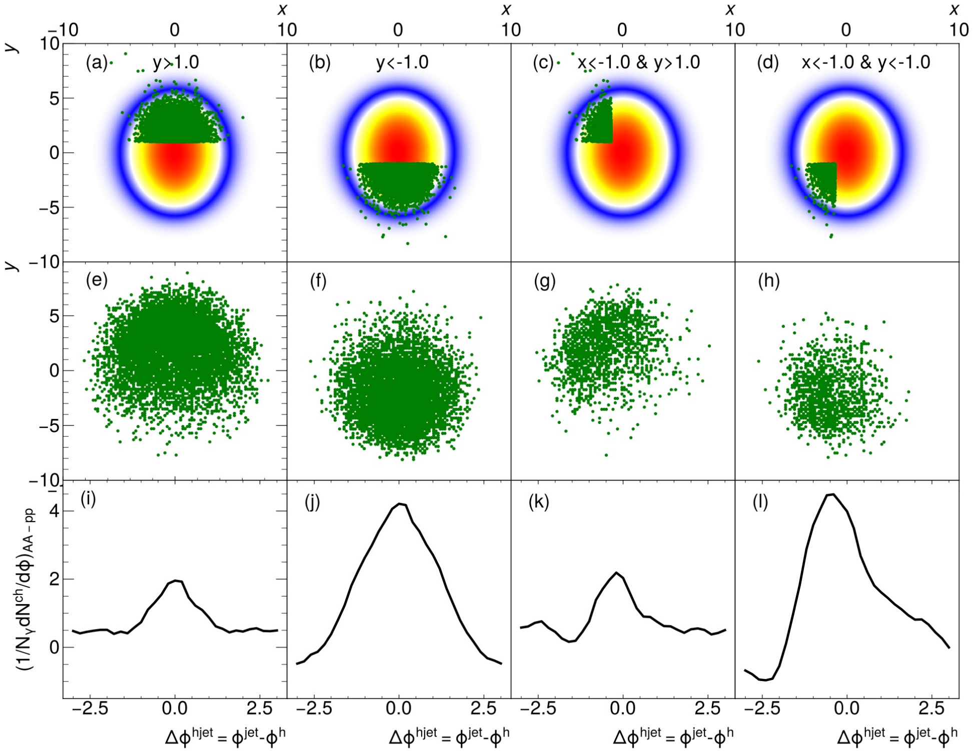
Jets and selected geometry

Zhong Yang, Yayun He, Wei Chen, Wei-Yao Ke, Long-Gang Pang, Xin-Nian Wang

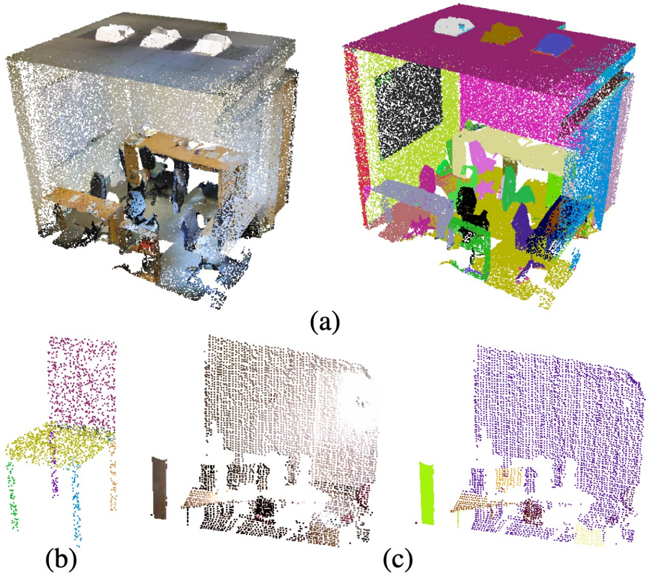
- Deep learning assisted jet tomography for the study of  in QGP

  - Is it possible to determine the initial jet production positions using the momenta of final state particles?
  - Will the signal of Mach cones and the diffusion wake be amplified if initial jet production positions are constrained to specific regions with long jet propagation lengths or fixed propagation direction relative to the radial flow?
  - How reliable is the new deep learning assisted method?

https://arxiv.org/abs/1711.08588



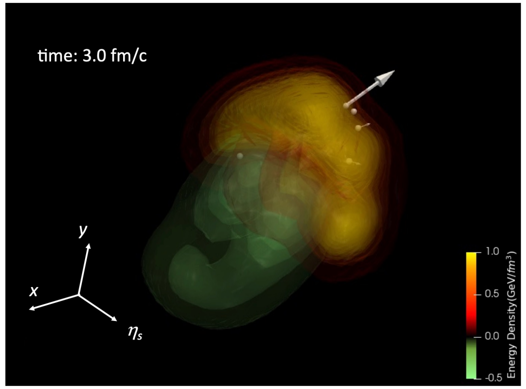
CoLBT-hydro, LIDO (train)
Background: ON
Constit. Subtr.: ON
p_T > 100 GeV/c
Supervised ML
POP study



https://arxiv.org/abs/2206.02393

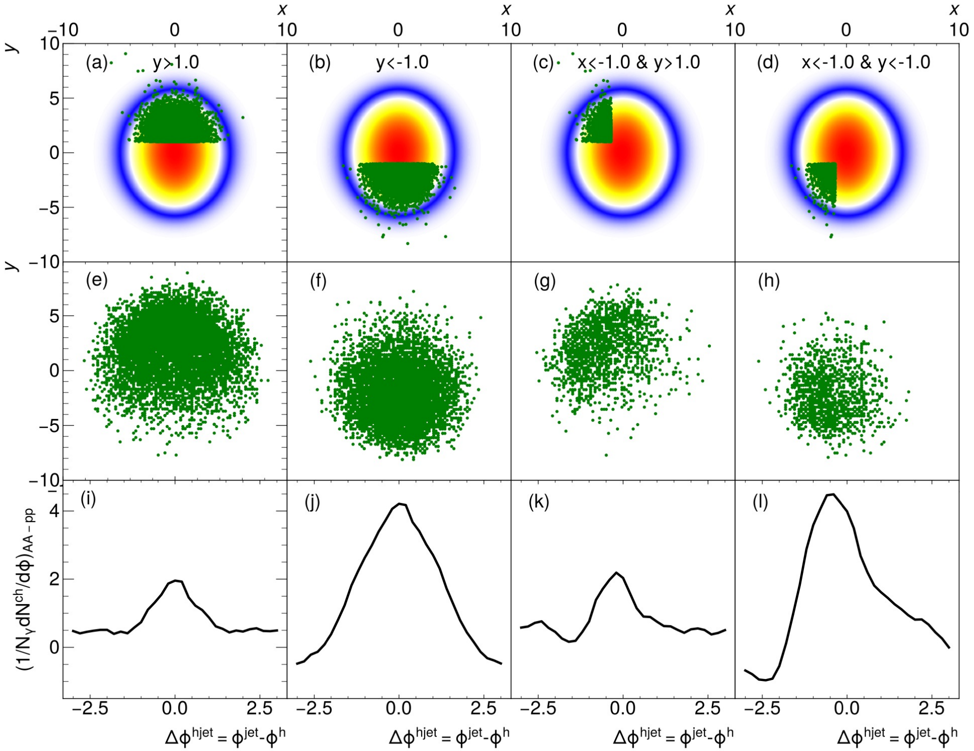
Jets and selected geometry

Zhong Yang, Yayun He, Wei Chen, Wei-Yao Ke, Long-Gang Pang, Xin-Nian Wang

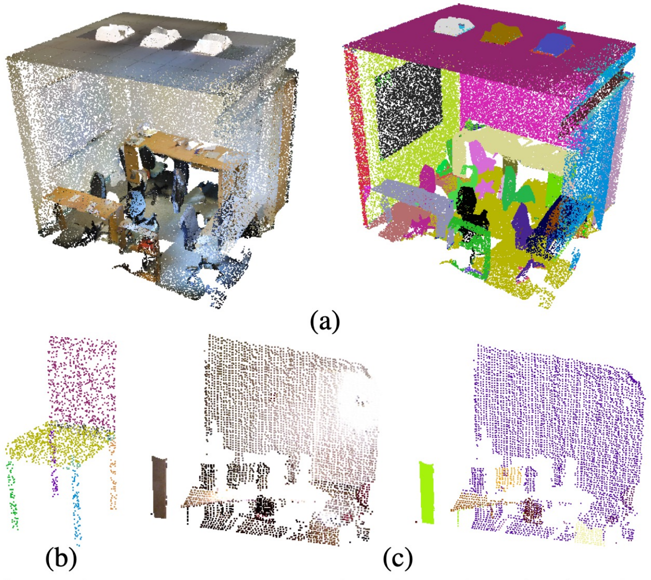


- Deep learning assisted jet tomography for the study in QGP
  - Is it possible to deter n positions using the momenta of final state particles?
  - => with large fluctuations
  - Will the signal of Mach cones and the diffusion wake be amplified if initial jet production positions are constrained to specific regions with long jet propagation lengths or fixed propagation direction relative to the radial flow?
  - How reliable is the new deep learning assisted method?

https://arxiv.org/abs/1711.08588



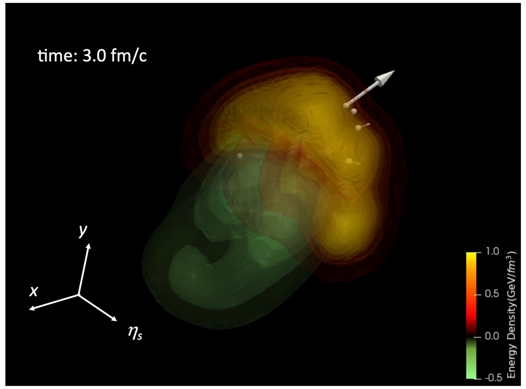
CoLBT-hydro, LIDO (train)
Background: ON
Constit. Subtr.: ON
p_T > 100 GeV/c
Supervised ML
POP study



https://arxiv.org/abs/2206.02393

Jets and selected geometry

Zhong Yang, Yayun He, Wei Chen, Wei-Yao Ke, Long-Gang Pang, Xin-Nian Wang



- Deep learning assisted jet tomography for the study in QGP
  - Is it possible to deter n positions using the momenta of final state particles?

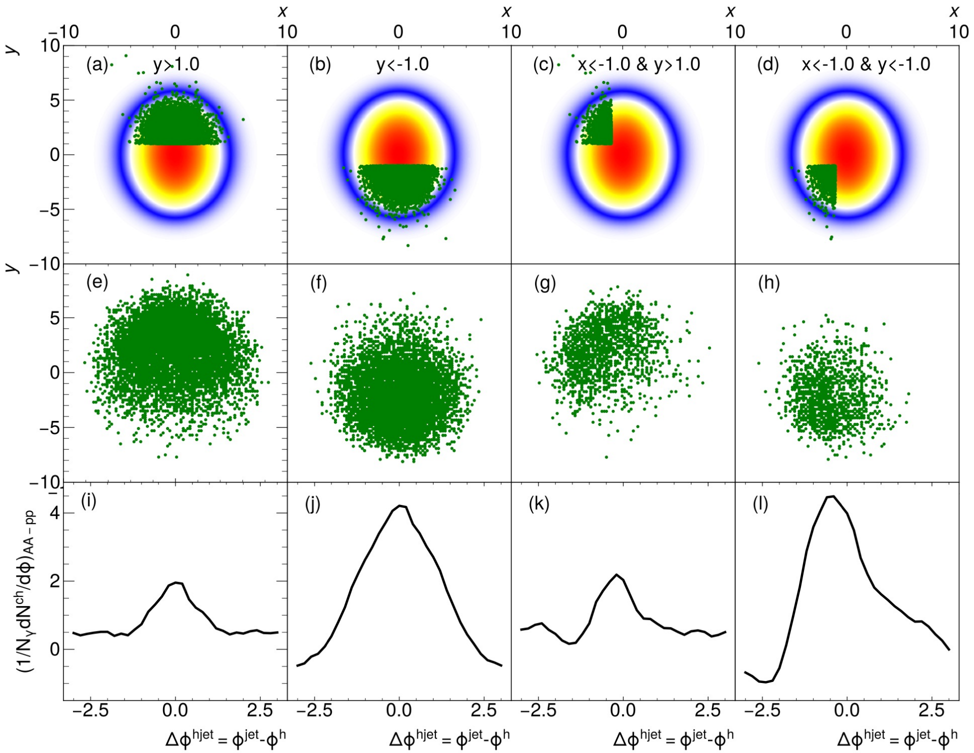
=> with large fluctuations

  - W initial jet pr t pr dial flow?

=> somewhat yes, but quantitative details important; scrutiny in the experiment?

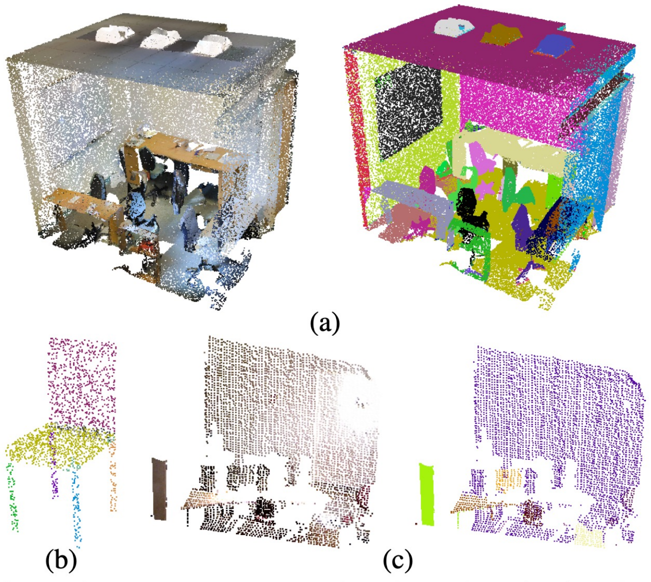
  - How reliable is the new deep learning assisted method?

https://arxiv.org/abs/1711.08588



CoLBT-hydro, LIDO (train)
Background: ON
Constit. Subtr.: ON
p_T > 100 GeV/c
Supervised ML
POP study

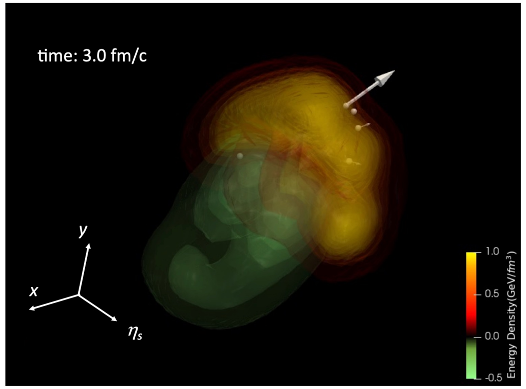
Note the asymmetric distributions



https://arxiv.org/abs/2206.02393

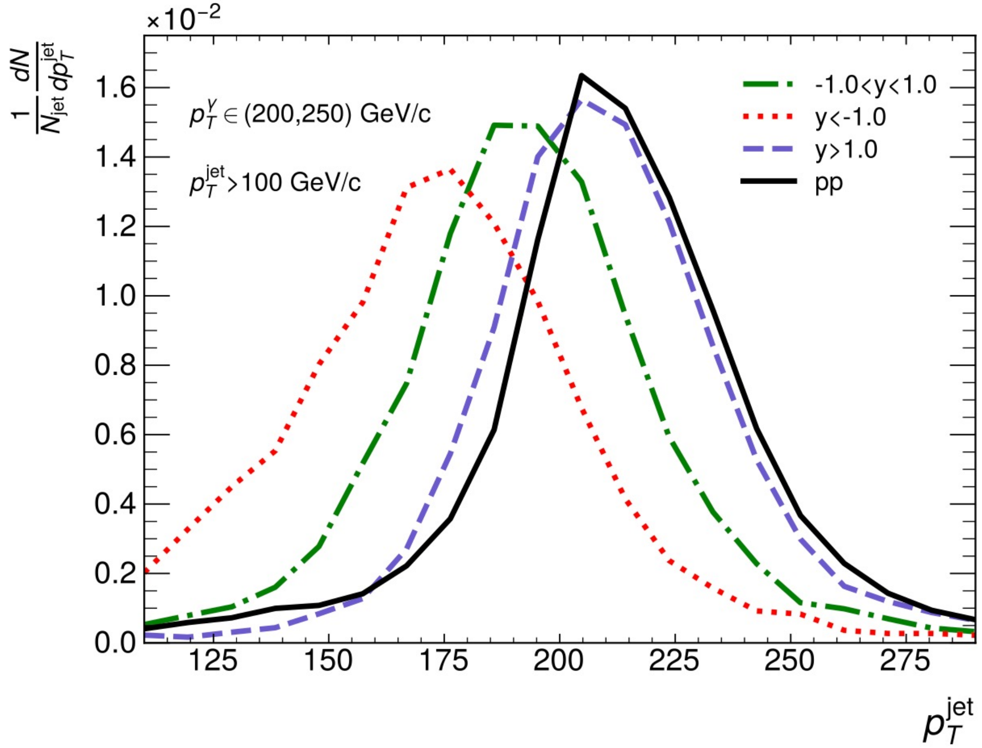
Jets and selected geometry

Zhong Yang, Yayun He, Wei Chen, Wei-Yao Ke, Long-Gang Pang, Xin-Nian Wang

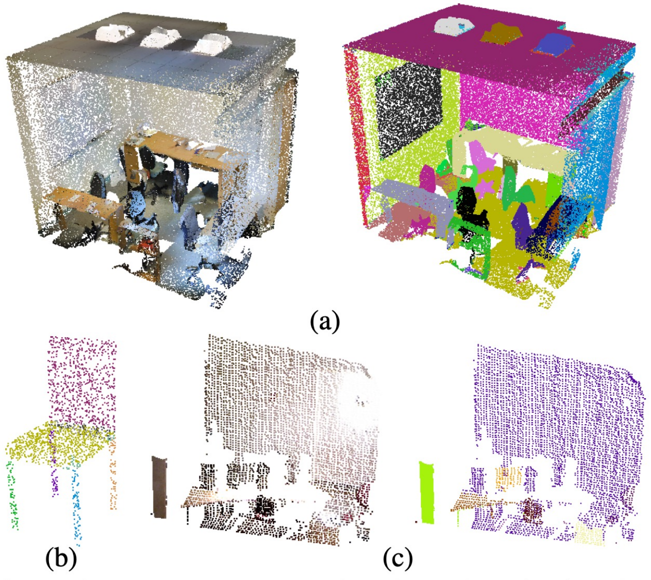


- Deep learning assisted jet tomography for the study in QGP
  - Is it possible to deter n positions using the momenta of final state particles?
  - => with large fluctuations
  - W initial jet pr t pr dial flow?
  - => somewhat yes, but quantitative details important; scrutiny in the experiment?
  - How reliable is
  - observation level only – quantitative level?

https://arxiv.org/abs/1711.08588



CoLBT-hydro, LIDO (train)
Background: ON
Constit. Subtr.: ON
pT > 100 GeV/c
Supervised ML
POP study



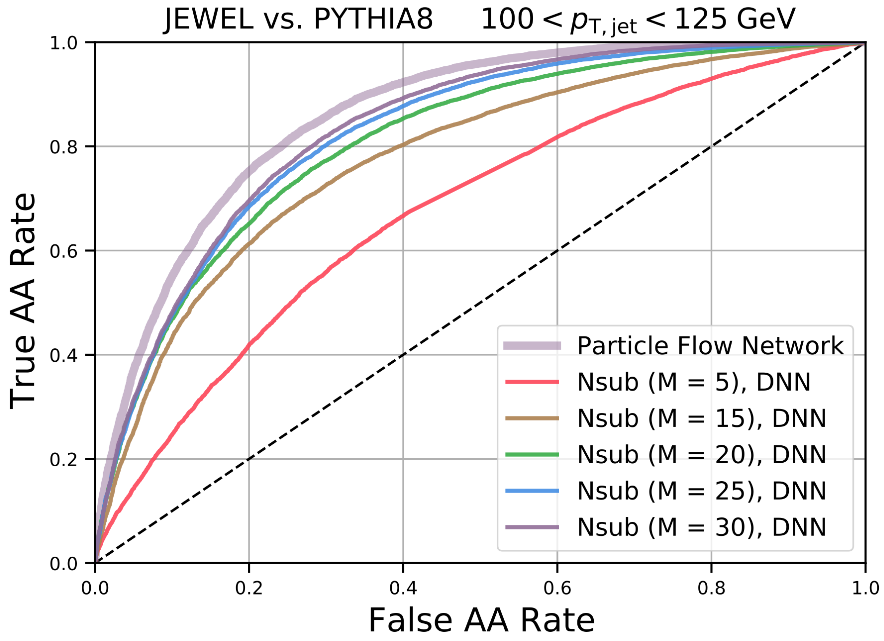
Information in quenched jets

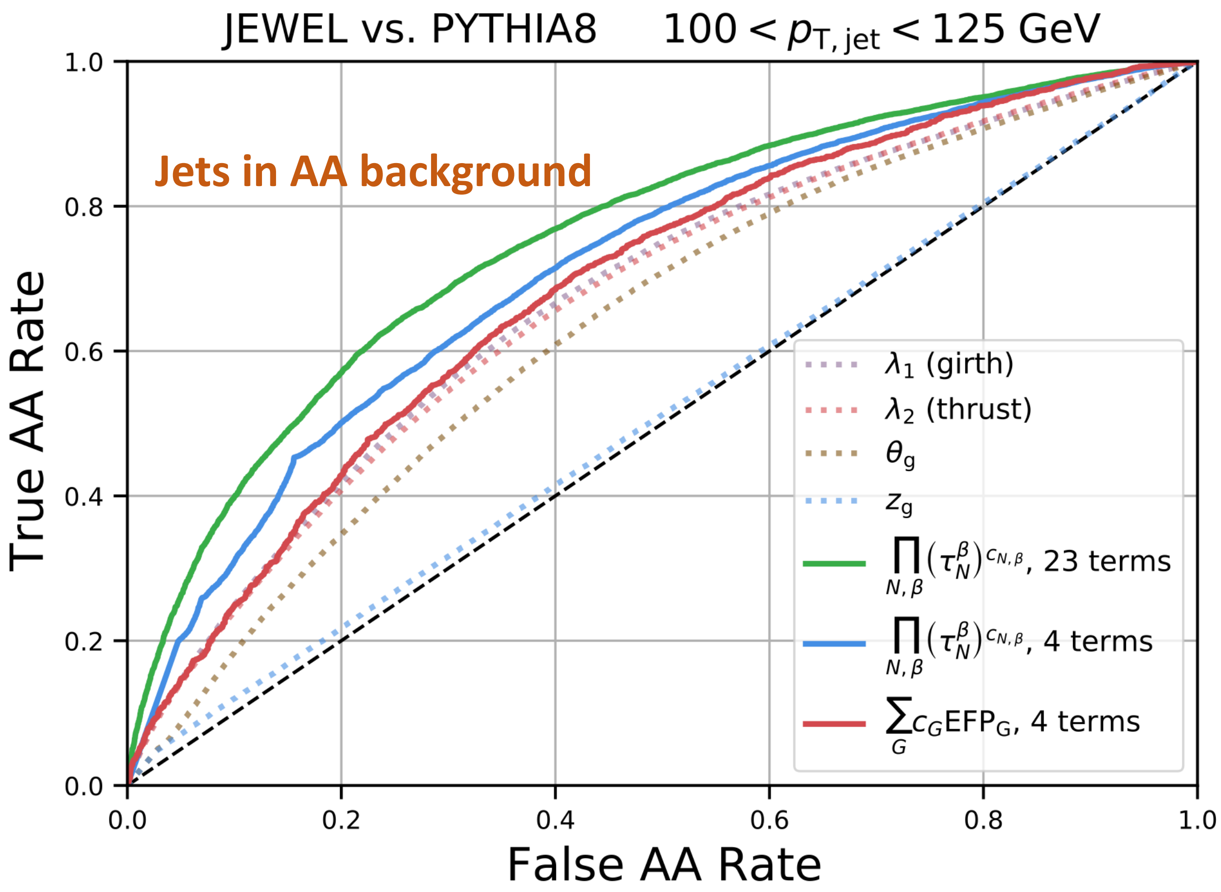
https://arxiv.org/abs/2111.14589
Yue Shi Lai, James Mulligan, MP, Felix Ringer

- Binary classification to design observables
  - Quantify information: IRC safe – hard vs. soft
  - **Define new, optimized discriminating observable**
  - Study information loss in AA
- Sensitivity to quenching in soft emissions and IRC-unsafe physics inside the jet
- **training labels exactly known(!) => use experimental data without reliance on modeling**

N-subjetiness - minimal basis of the M-body phase space of emissions inside the jet

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





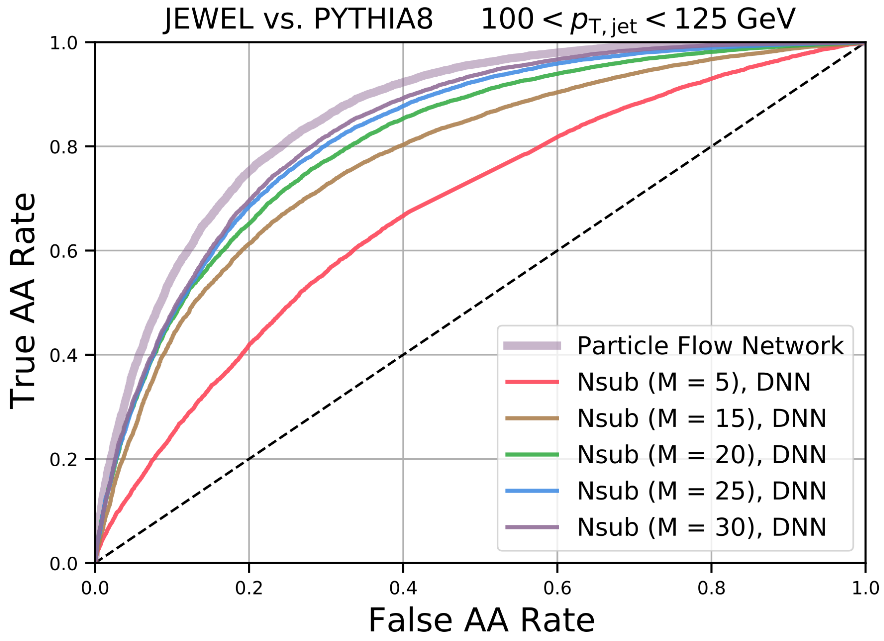
https://arxiv.org/abs/2111.14589
Yue Shi Lai, James Mulligan, MP, Felix Ringer

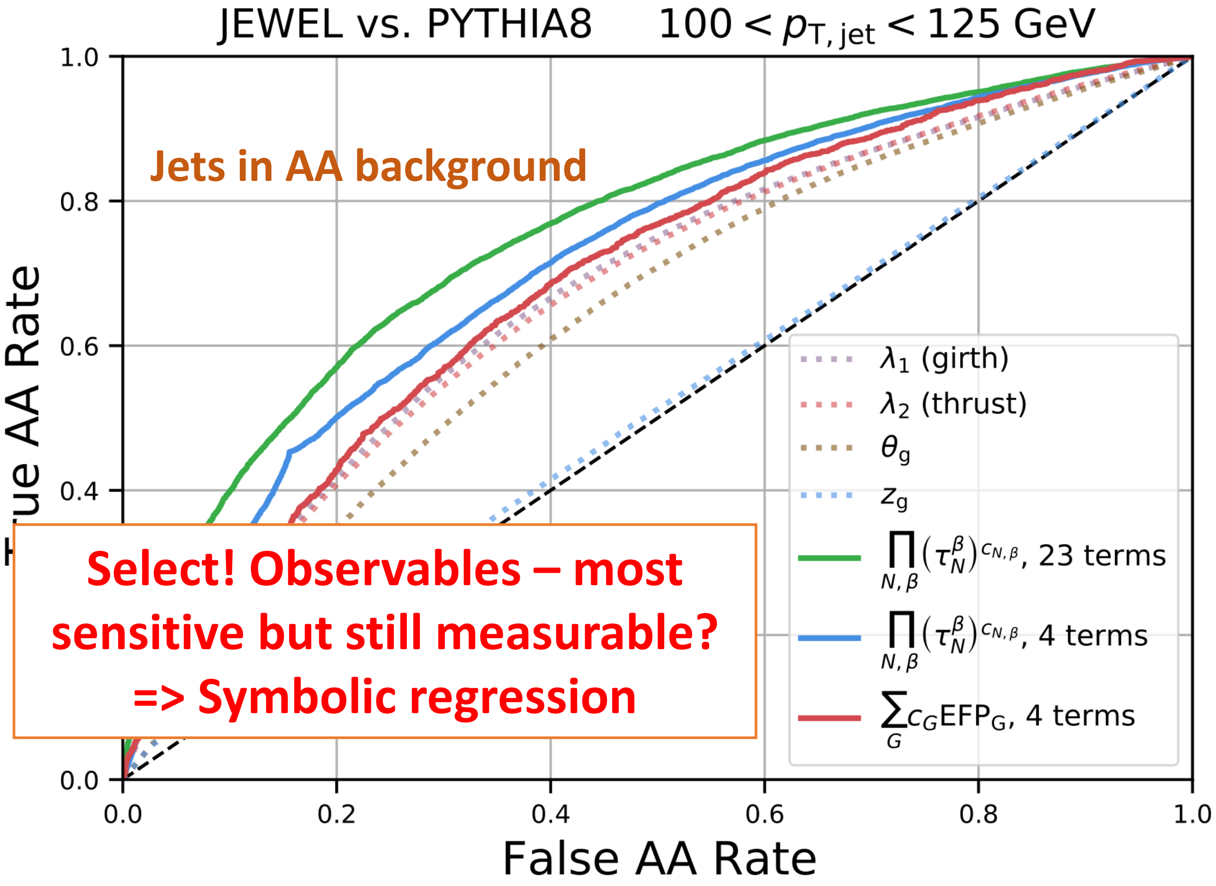
Information in quenched jets

- Binary classification to design observables
  - Quantify information: IRC safe – hard vs. soft
  - **Define new, optimized discriminating observable**
  - Study information loss in AA
- Sensitivity to quenching in soft emissions and IRC-unsafe physics inside the jet
- **training labels exactly known(!) => use experimental data without reliance on modeling**

N-subjetiness - minimal basis of the M-body phase space of emissions inside the jet

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



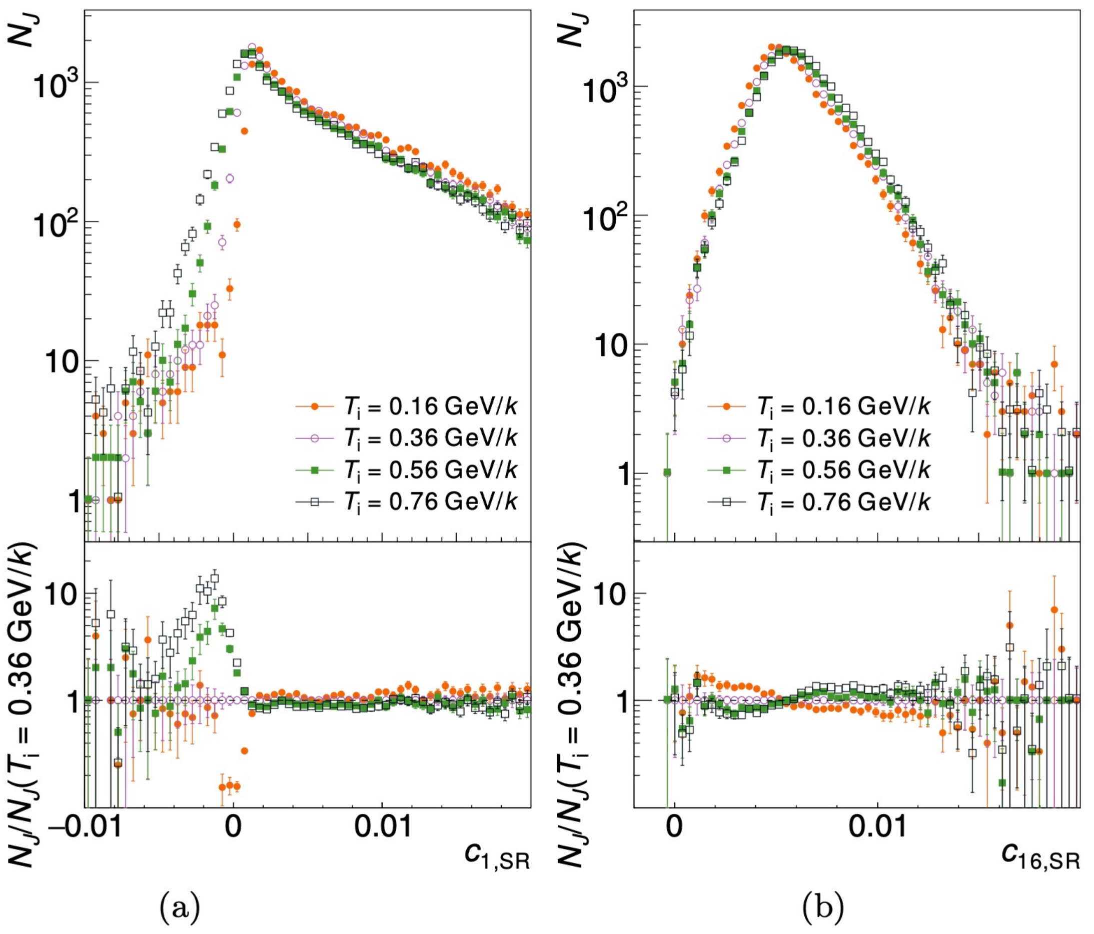


Select! Observables – most sensitive but still measurable?
=> Symbolic regression

Inferring medium properties?

https://arxiv.org/abs/2012.06582
Yue Shi Lai

- "Automated Discovery of Jet Substructure Analyses"
- Use architecture a la cNN to enable feature extraction
- Perform symbolic regression to constrain algebraic form of *the* observable sensitive to features in data…
- **example: sensitivity to initial temperature parameter in JEWEL**



The distribution of the symbolic regression approximated neuron (a) $c_{1,SR}$ and (b) $c_{16,SR}$ for various T_i in Jewel for $100 < p_{T,J} < 300$ GeV/c, with the ratio relative to $T_i = 0.36$ GeV

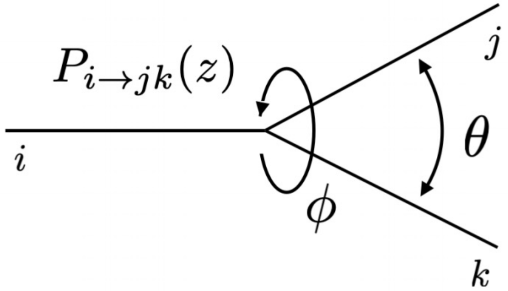
Machine learning physics

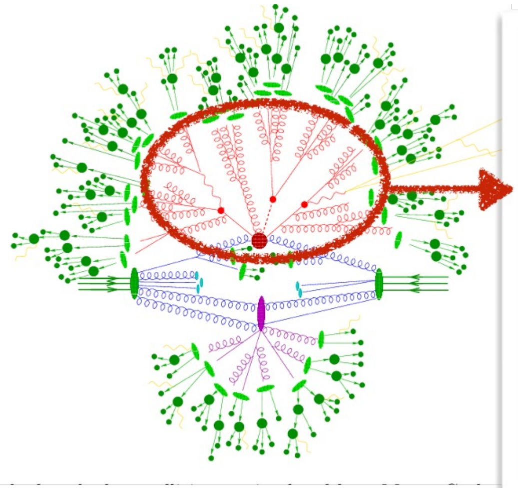
https://arxiv.org/abs/2012.06582

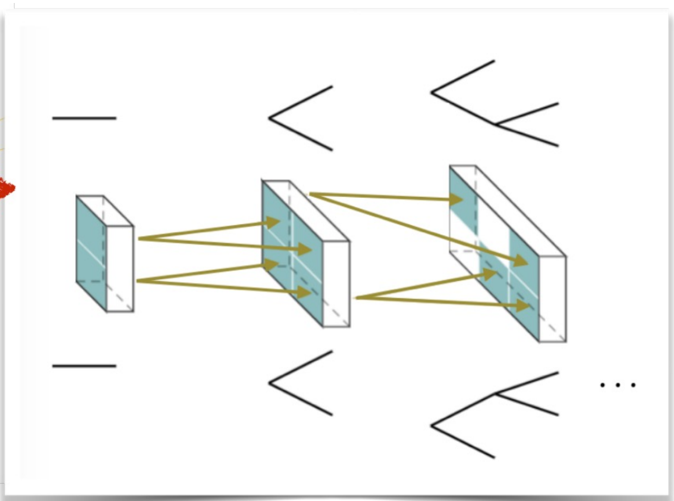
Yue Shi Lai, Duff Neill, MP, Felix Ringer

- Not jet quenching example but illustrative for a broader context => white box AI
  - Modus operandi: Design explainable physics-aware AI/ML
  - Ultimate goal: Infer the underlying physics directly from data (e.g. RHIC, LHC,…, future EIC)
  - At this point: Proof of concept using a parton shower

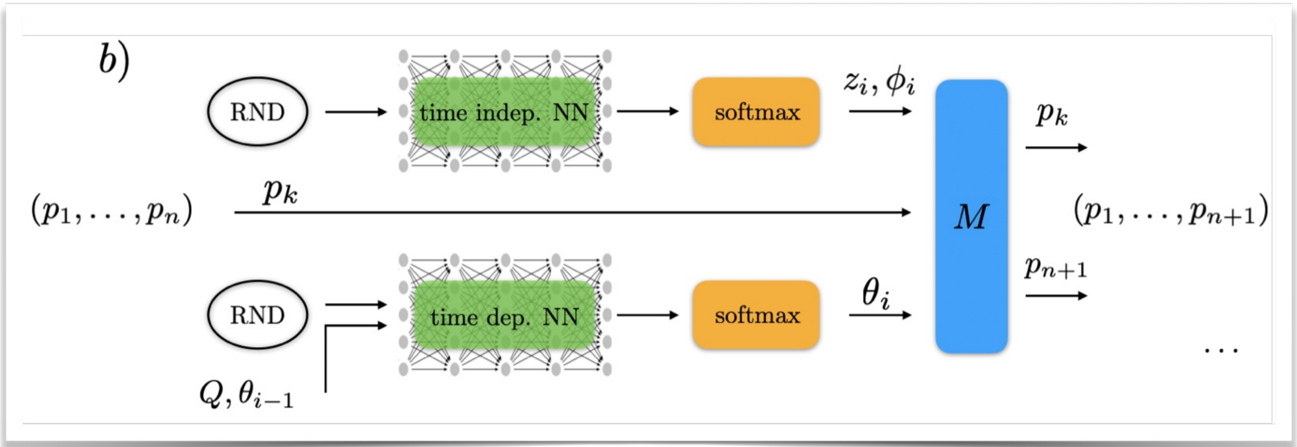
A GAN reconstructs the parton shower (encoding kinematics of splittings)







ith splitting process (n → n+1 partons)



Parton shower: multiple partons and their splittings

Single parton splitting encoded

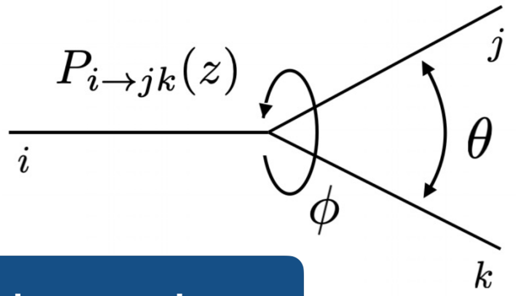
Machine learning physics

https://arxiv.org/abs/2012.06582

Yue Shi Lai, Duff Neill, MP, Felix Ringer

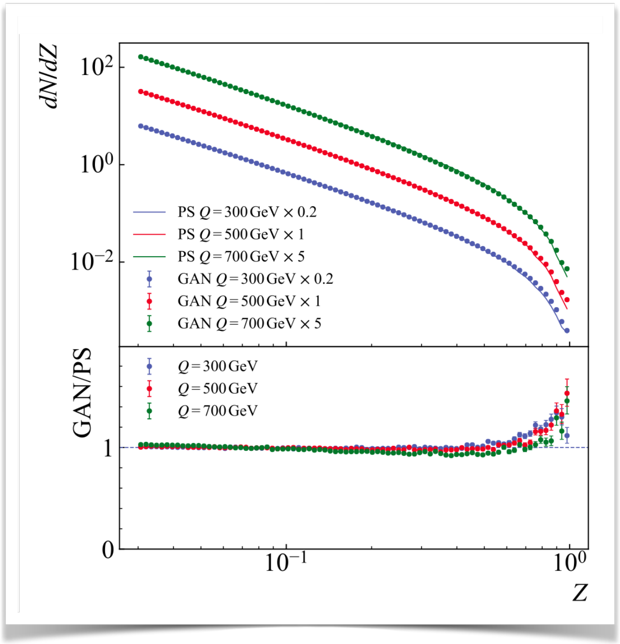
- Not jet quenching example but illustrative for a broader context => white box AI
  - Modus operandi: Design explainable physics-aware AI/ML
  - Ultimate goal: **Infer the underlying physics directly from data** (e.g. RHIC, LHC,…, future EIC)
  - At this point: Proof of concept using a parton shower

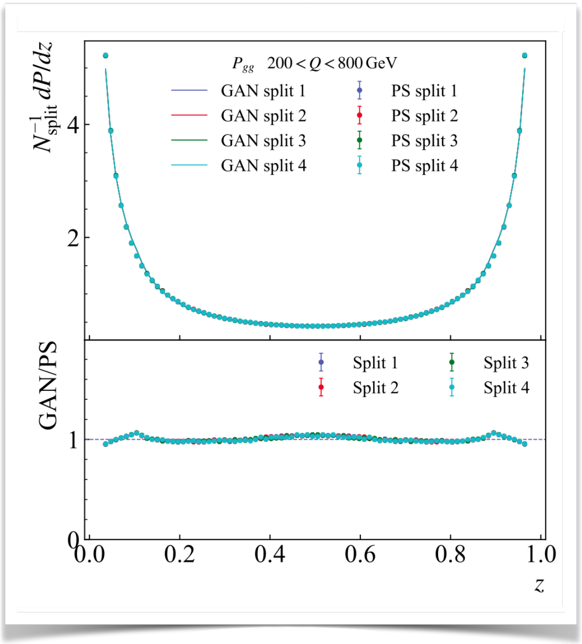
A GAN reconstructs the parton shower (encoding kinematics of splittings)



Input: generated (PS) distribution of final state (FS) partons - GAN reconstructs the FS precisely …

… but internally the NN learns the physical distribution for a single splitting - the DGLAP splitting function





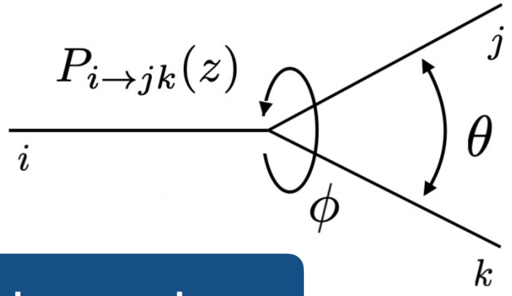
Machine learning physics

https://arxiv.org/abs/2012.06582

Yue Shi Lai, Duff Neill, MP, Felix Ringer

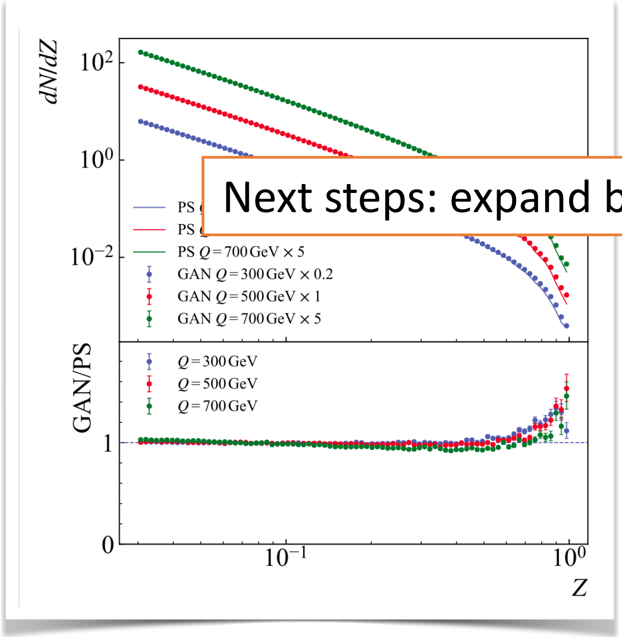
- Not jet quenching example but illustrative for a broader context => white box AI
  - Modus operandi: Design explainable physics-aware AI/ML
  - Ultimate goal: **Infer the underlying physics directly from data** (e.g. RHIC, LHC,…, future EIC)
  - At this point: Proof of concept using a parton shower

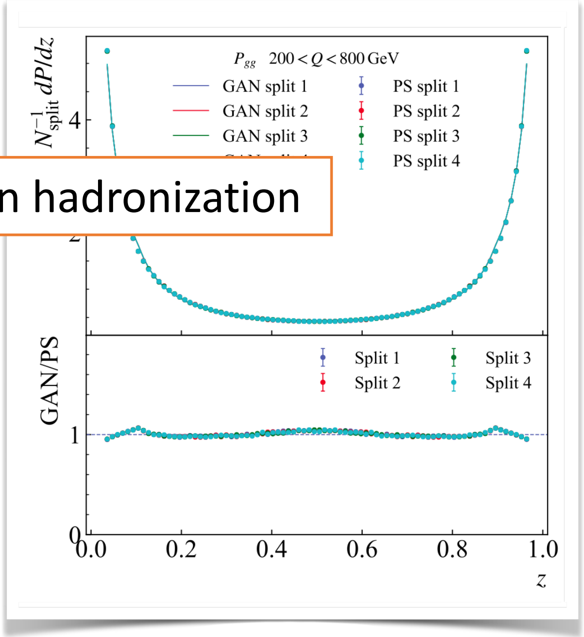
A GAN reconstructs the parton shower (encoding kinematics of splittings)



Input: generated (PS) distribution of final state (FS) partons - GAN reconstructs the FS precisely …

… but internally the NN learns the physical distribution for a single splitting - the DGLAP splitting function





Next steps: expand beyond 1->2; built-in hadronization

Summary

- Physics model dependent statements:
  - With some precision quark vs. gluon jet discrimination in simulation
  - Jet quenching information in soft components of the jet
  - It is difficult to select on geometry with control (fluctuations) but qualitatively selection on points of scattering, path length possible (in models)
- Physics model independent statements – data ready:
  - There is no free lunch
  - Model dependence is a problem that needs to be eliminated/controlled
  - Quantification of uncertainties is a problem that needs addressing

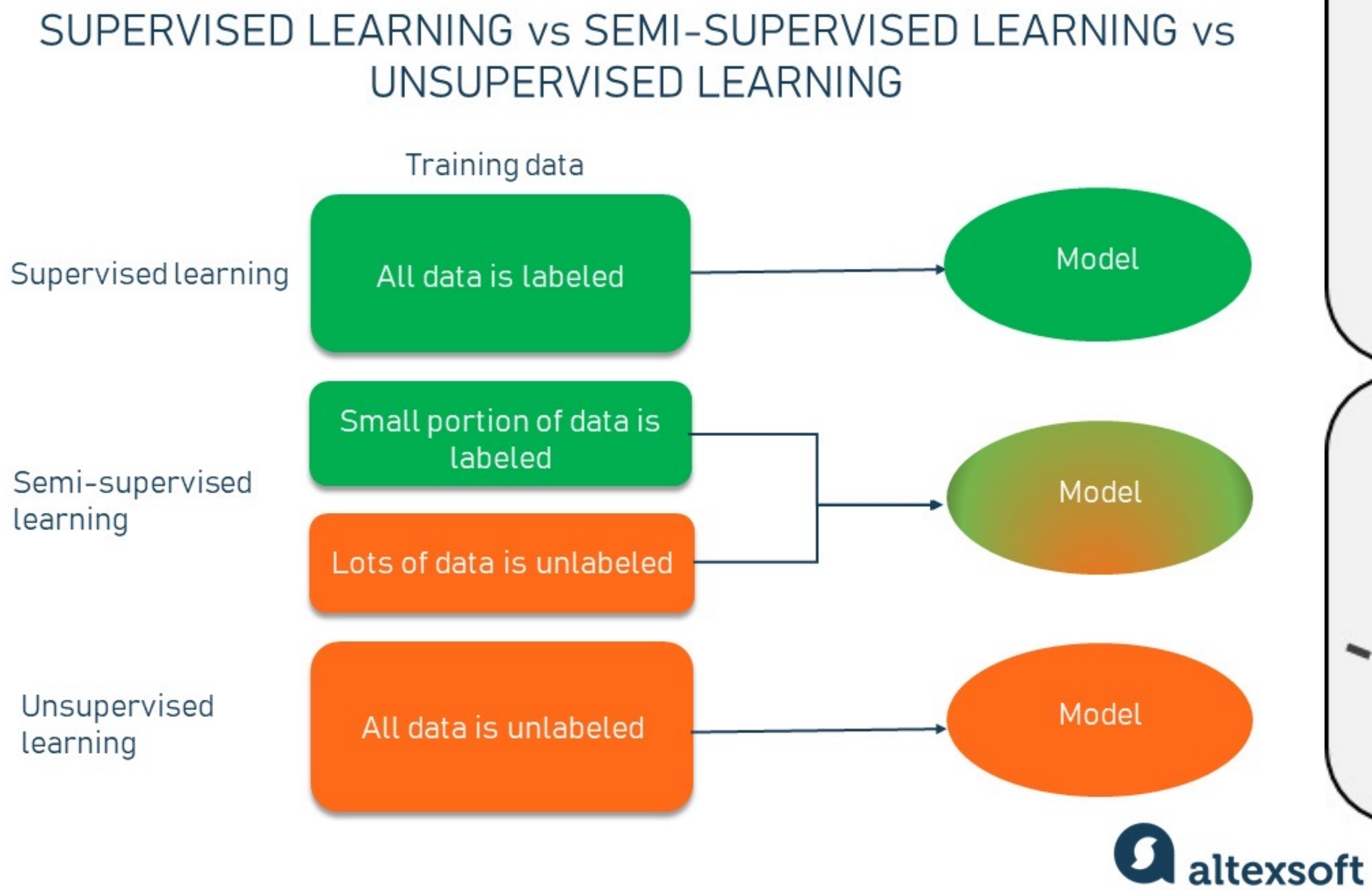
Future (wanted) developments – aka Outlook

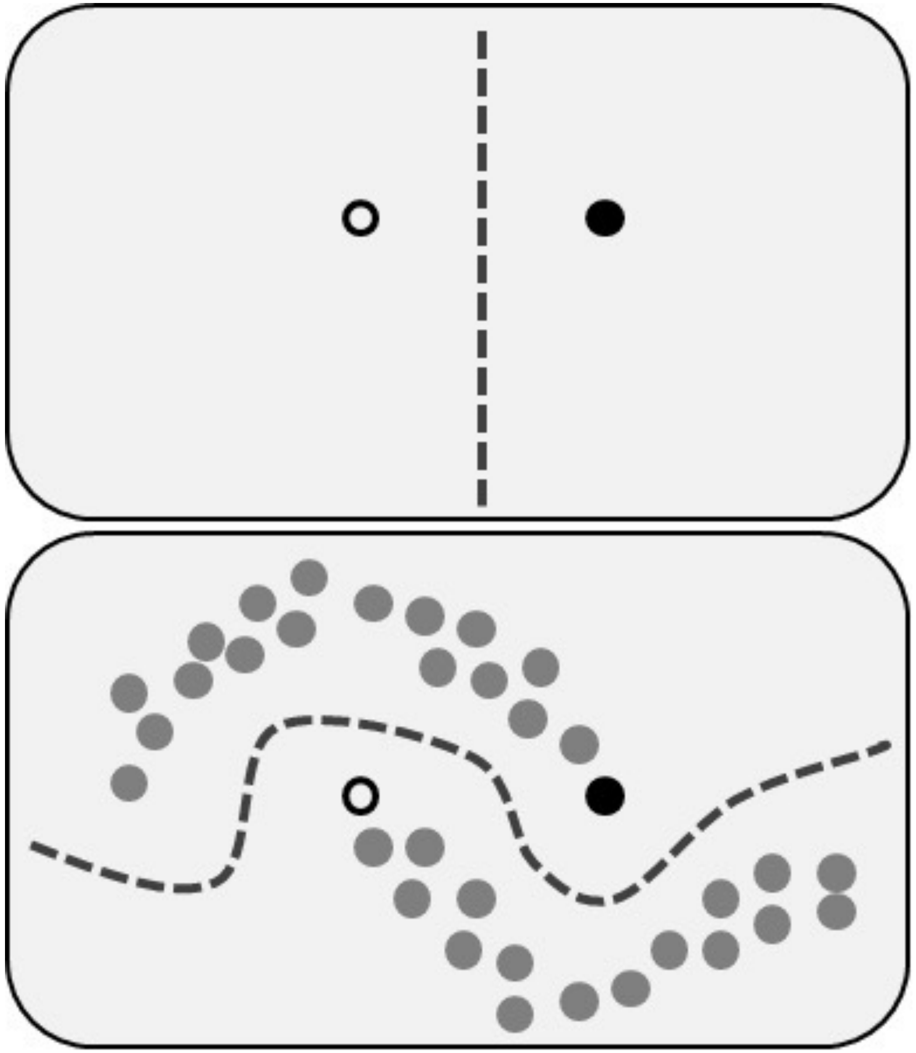
- Move away from image based deep learning to sequence based data
  - Avoid information loss, enable large latent space, no padding
  - deep sets, graphNNs
- A tip: benchmark on known / public data sets (the same data sets)
- **Physics aware models ⬄ white box (as opposed to black box) models**
  - **Interpretability! => experiment theory feedback?**
- (Un)supervised learning on data? Or semi-supervised = data+MC
  - **Jet classification directly in AA data** – connection to j.q. modelling? => inference of physics
- Uncertainty inference / quantification
- 'New' directions:
  - More GAN based approaches?
  - Decorrelation methods (!) – look for smallest set of maximally discriminating observables

bayesfast

https://www.bayesfast.org

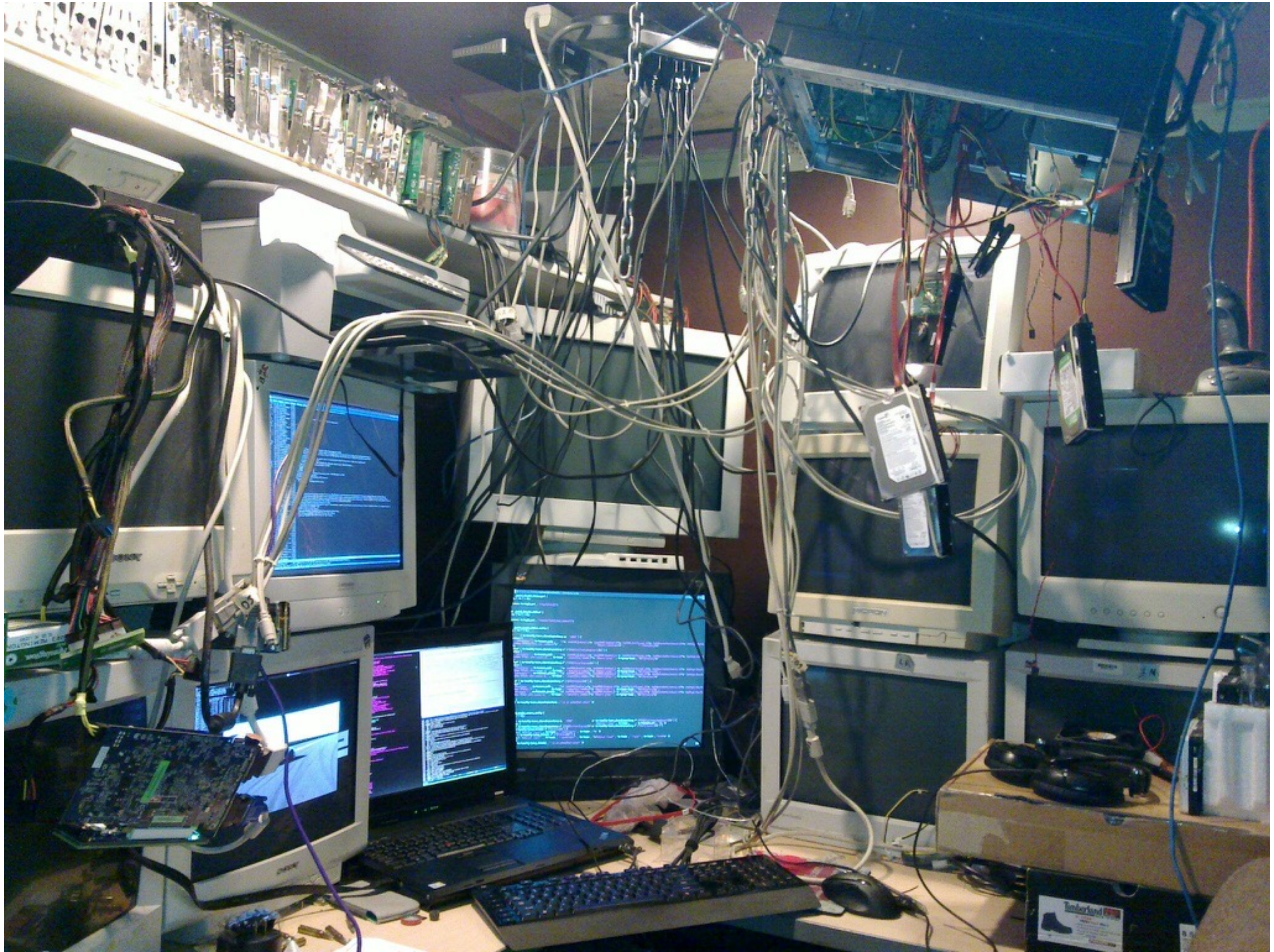
Intriguing opportunity: semi-supervised, unsupervised? <=> not all (none) labels given





It’s just the beginning…

But we should rapidly move beyond proof of concepts and generic ML tech – towards physics aware / white box / interpretable AI and data-ready … it requires a conceptual change…



Thanks!

Optimal ML “output”

- What features of jets change medium-vacuum?
  - Observables
  - Constituents (‘medium recoil’ not distinguishable from jet particles by def.)
- Connection the theory?
- Assistive: select quenched jets and study those with a traditional “microscope” (observables)
  - What means quenched jets in model-agnostic scenario?

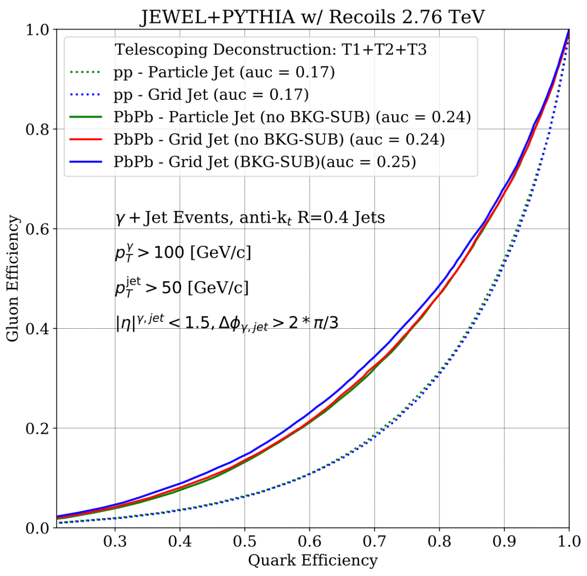
Discrimination: quark vs. gluon jets…

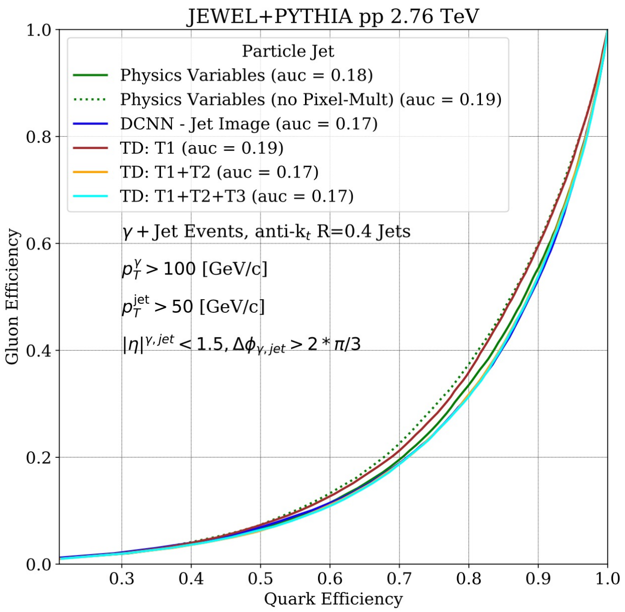
https://arxiv.org/abs/1803.03589

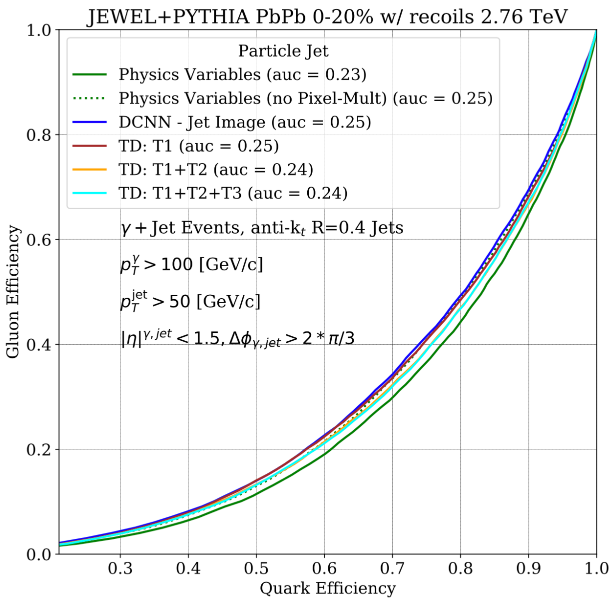
JEWEL
Background: ON/OFF
pT > 50 GeV/c
POP study

- Probing heavy ion collisions using quark and gluon jet substructure
- Jet mass, two radial moments including the girth, the $p_{T,D}$, and the pixel multiplicity => multi-layer perceptron (MLP);
- Deep convolutional neural network (DCNN) on discretized images of quark jets and gluon jets (η,ϕ)
- Telescoping deconstruction framework exploiting subjet kinematics – pT, mass (use MLP)
- **"We find that the quark gluon discrimination performance worsens in heavy ion jets due to significant soft event activity affecting the soft jet substructure."**

https://arxiv.org/abs/1310.7584
- HEP related







"Through multiple methods and observables, we **consistently find the dominant feature of the Jewel jet quenching model to be the increase of soft particle multiplicity due to medium recoils throughout the jet region**. This is closely related to the loss of information in subleading subjets, which is a characteristic feature of Jewel."

"The CNN architecture has not been tuned exhaustively, therefore its ROC curves serve to give a general sense of performance. "

Discrimination: quark vs. gluon jets…

https://arxiv.org/abs/1803.03589

JEWEL
Background: ON/OFF
pT > 50 GeV/c
POP study

- Probing heavy ion collisions using quark and gluon jet substructure
- Jet mass, two radial moments including the girth, the $p_{T,D}$, and the pixel multiplicity => multi-layer perceptron (MLP);
- Deep convolutional neural network on discretized images of quark jets and gluon jets (η,ϕ)
- Telescoping deconstruction framework exploiting subjet kinematics – pT, mass (use MLP)
- **"We find that the quark gluon discrimination performance worsens in heavy ion jets due to significant soft event activity affecting the soft jet substructure."**

https://arxiv.org/abs/1310.7584
- HEP related

The higher the better

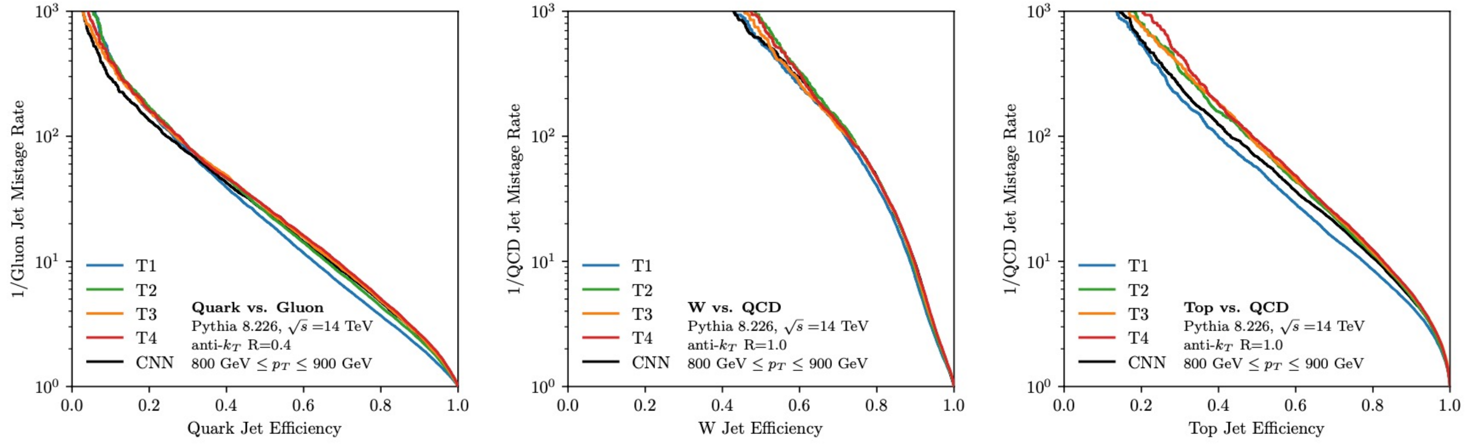
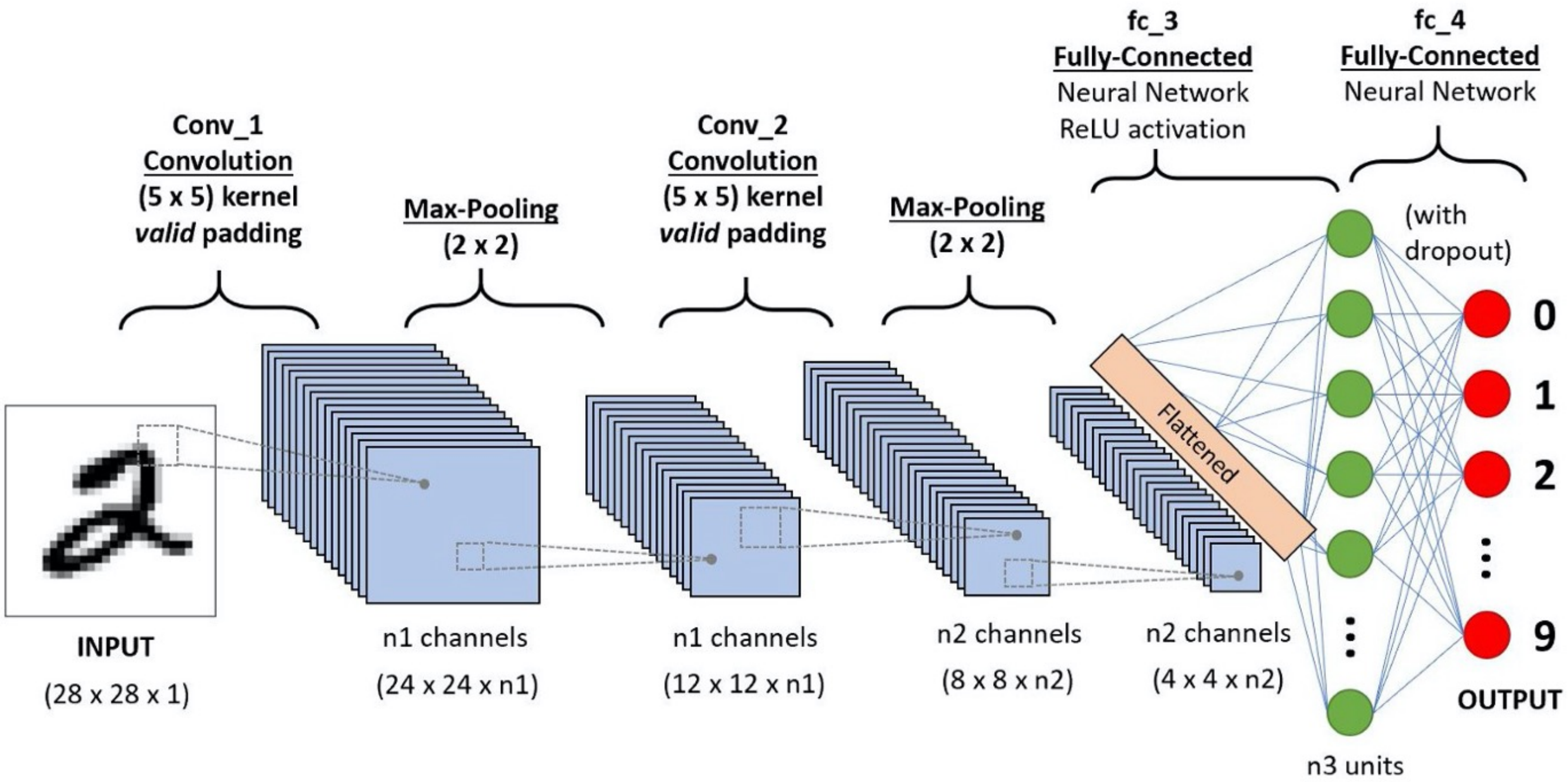


Figure 19: ROC curves for the DNNs trained on cumulant telescoping deconstruction observables up to T1 through T4 orders and the jet image method using CNNs for quark/gluon discrimination (left panel), boosted W (middle panel) and top (right panel) tagging. The TN performance approximately saturates at T2 (quark/gluon), T1 (W tagging), and T2 (top tagging) orders.

"Through multiple methods and observables, we **consistently find the dominant feature of the Jewel jet quenching model to be the increase of soft particle multiplicity due to medium recoils throughout the jet region.** This is closely related to the loss of information in subleading subjets, which is a characteristic feature of Jewel."

"The CNN architecture has not been tuned exhaustively, therefore its ROC curves serve to give a general sense of performance. "

DCNN example



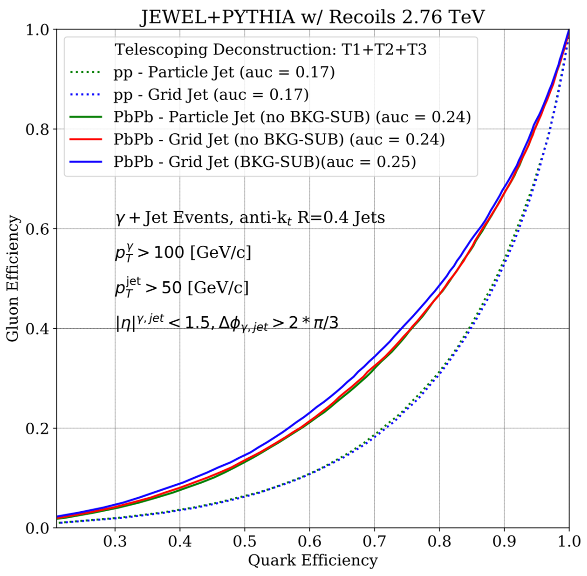
Discrimination: quark vs. gluon jets…

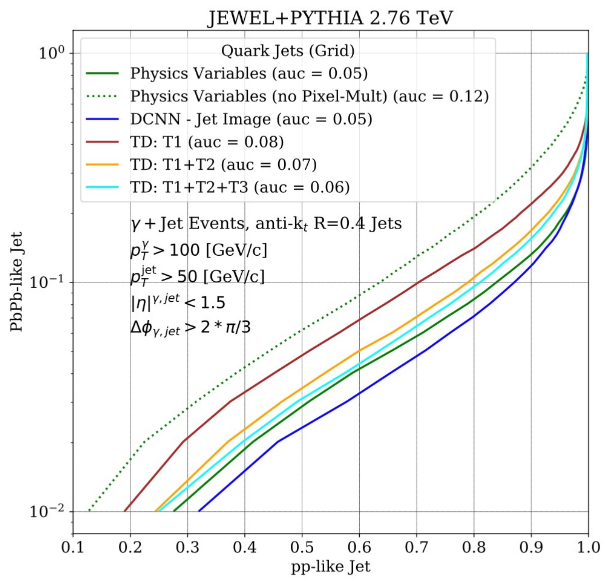
https://arxiv.org/abs/1803.03589

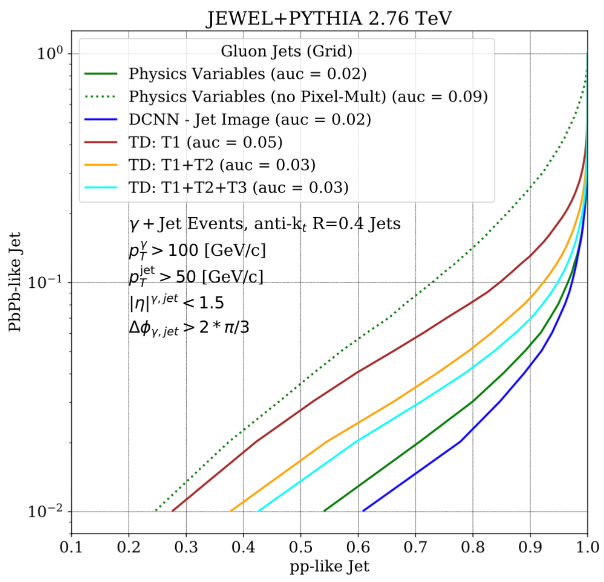
JEWEL
Background: ON/OFF
pT > 50 GeV/c
POP study

- Probing heavy ion collisions using quark and gluon jet substructure
- Jet mass, two radial moments including the girth, the $p_{T,D}$, and the pixel multiplicity => multi-layer perceptron (MLP);
- Deep convolutional neural network (DCNN) on discretized images of quark jets and gluon jets (η,ϕ)
- Telescoping deconstruction framework exploiting subjet kinematics – pT, mass (use MLP)
- **"We find that the quark gluon discrimination performance worsens in heavy ion jets due to significant soft event activity affecting the soft jet substructure."**

https://arxiv.org/abs/1310.7584
- HEP related







"Through multiple methods and observables, we **consistently find the dominant feature of the Jewel jet quenching model to be the increase of soft particle multiplicity due to medium recoils throughout the jet region**. This is closely related to the loss of information in subleading subjets, which is a characteristic feature of Jewel."

"The CNN architecture has not been tuned exhaustively, therefore its ROC curves serve to give a general sense of performance. "