# Applications of machine learning in jet quenching

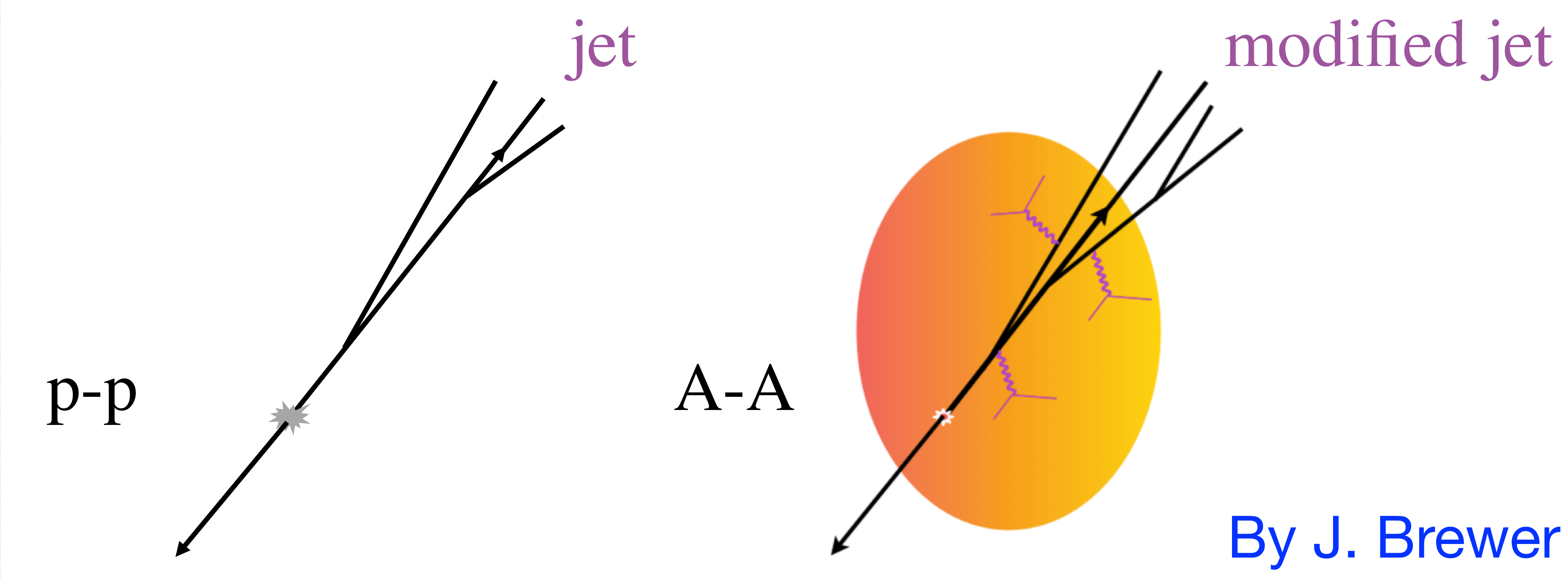
Based on Ran Li（李然）, YLD, Shanshan Cao, PRC 113, 024912 (2026), PLB 870 (2025) 139940
& YLD, Daniel Pablos and Konrad Tywoniuk, JHEP03(2021)206, PRL 128, 012301 (2022)

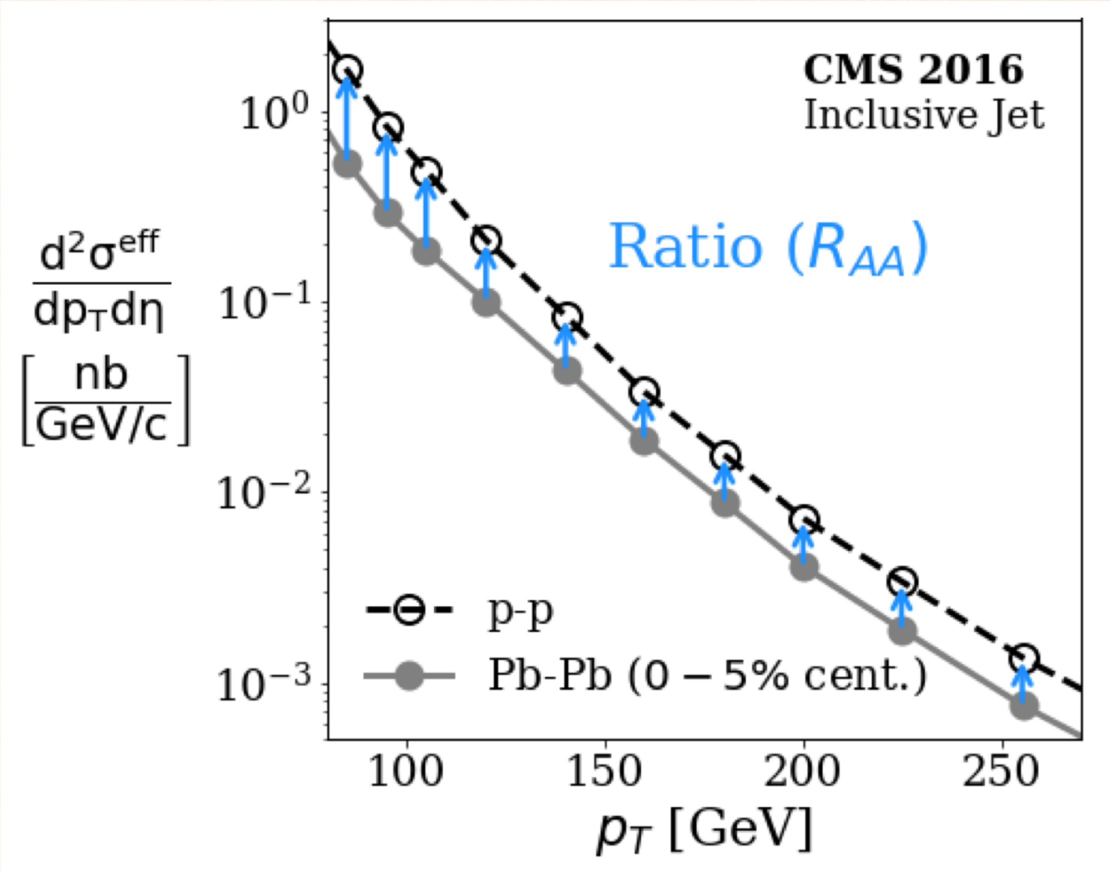
Yi-Lun Du（杜轶伦）

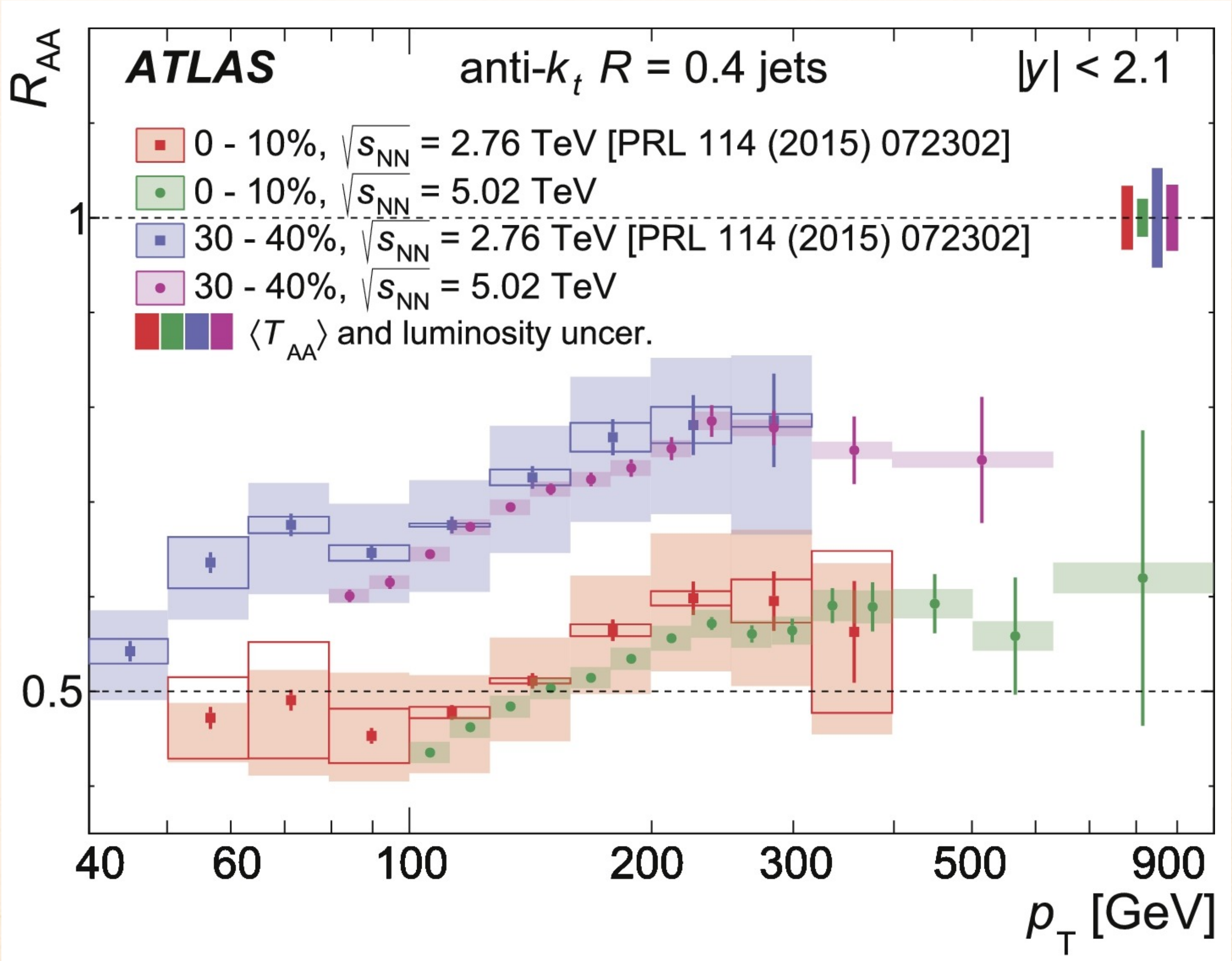
Nanjing Normal University

C3NT Workshop: Jet-Soft Correlations
March 22-April 4 2026, Wuhan

# QGP and Jet Modifications

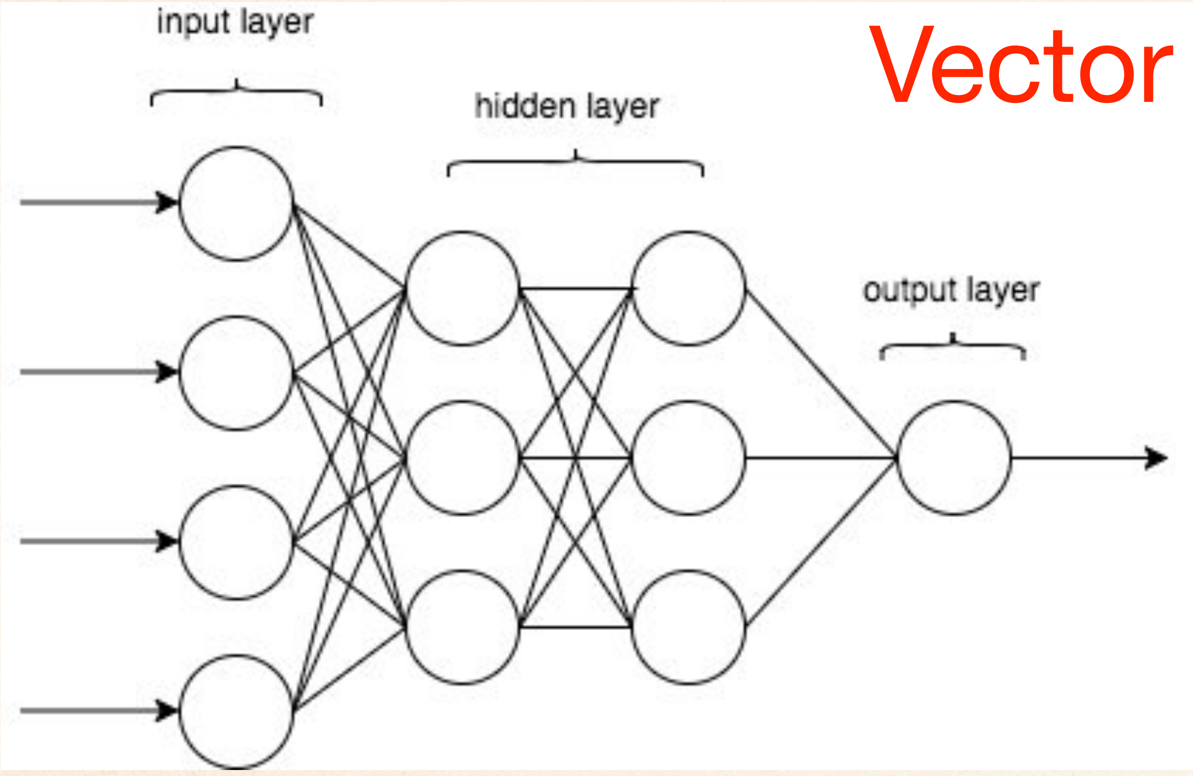


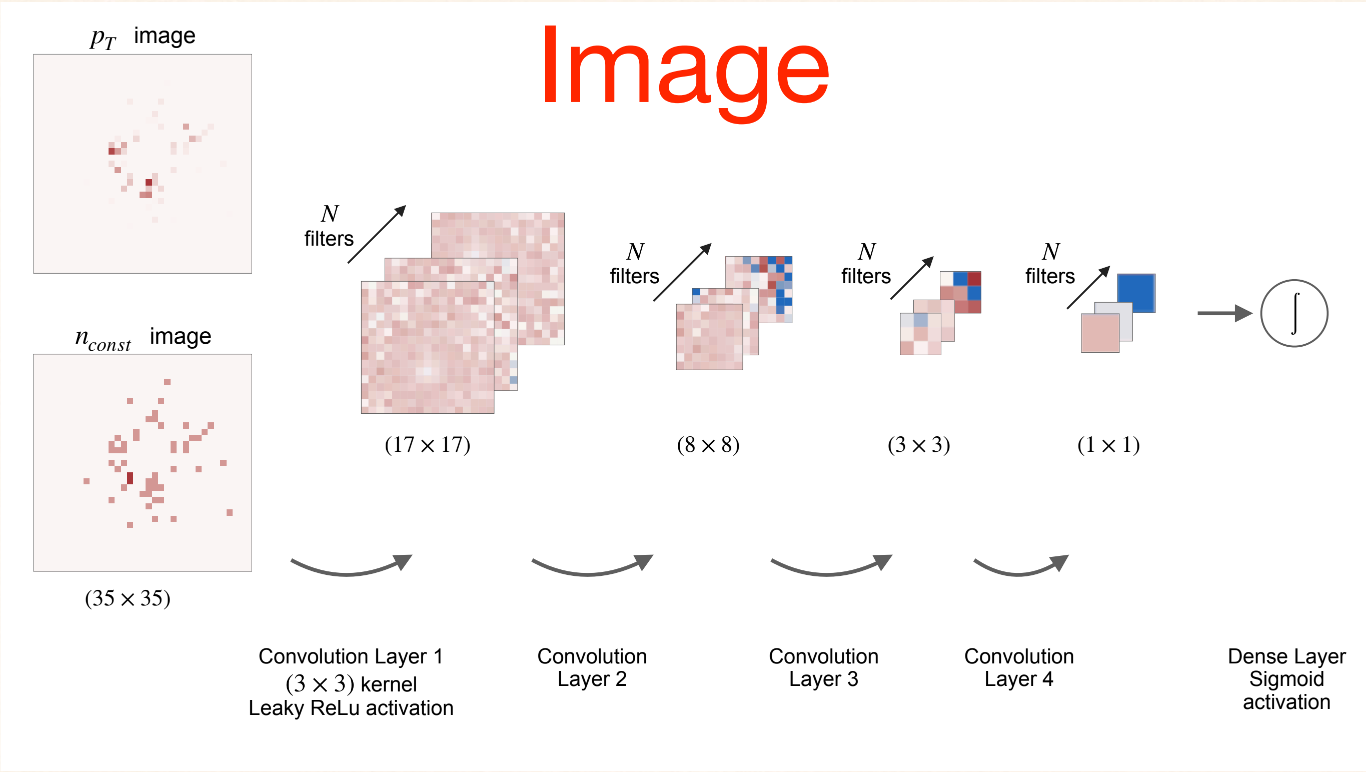



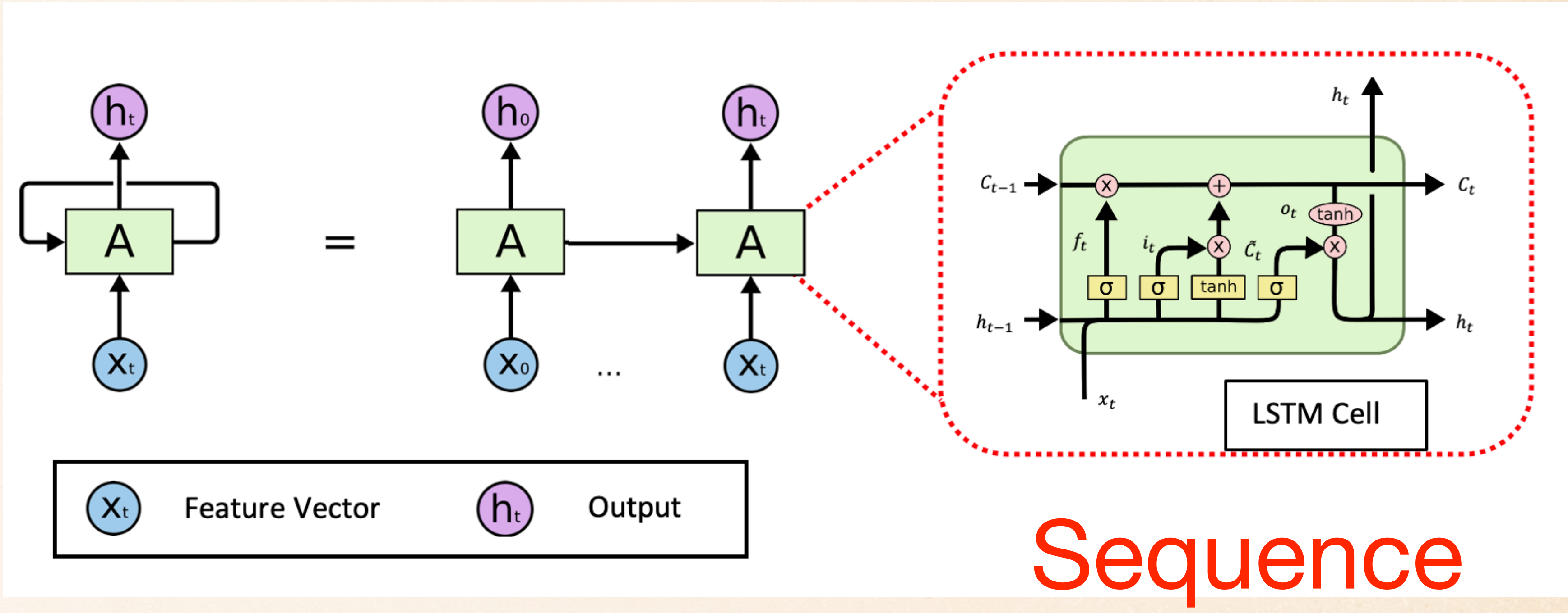



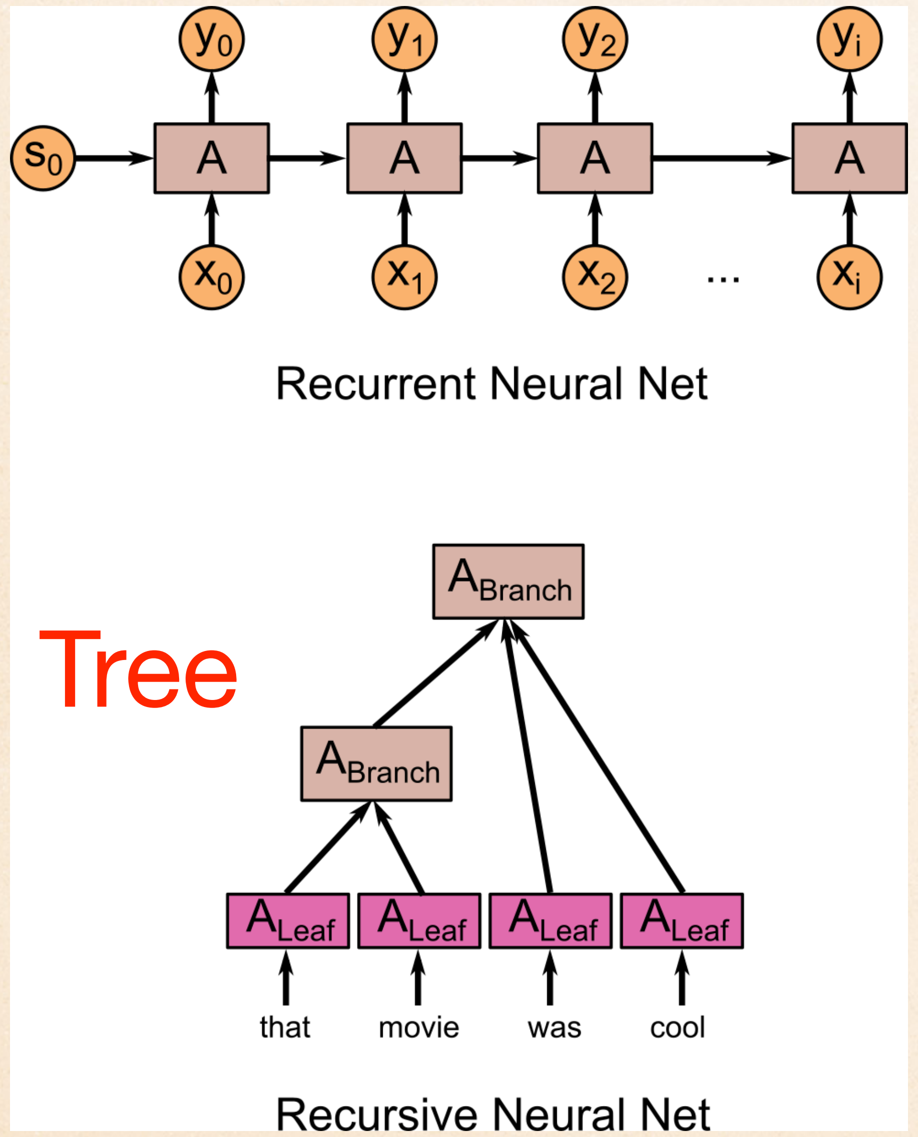

- Jets are unique hard probes to QGP
  - QGP: AA/AB collisions at $\sqrt{s}$ with different centralities
  - Jets (dijets, $\gamma$/Z-jets ): initiator's flavor, energy, position, direction and jet substructures developed at early stages
  - Jet-medium interactions
- Towards more precise probes & jet tomography with ML

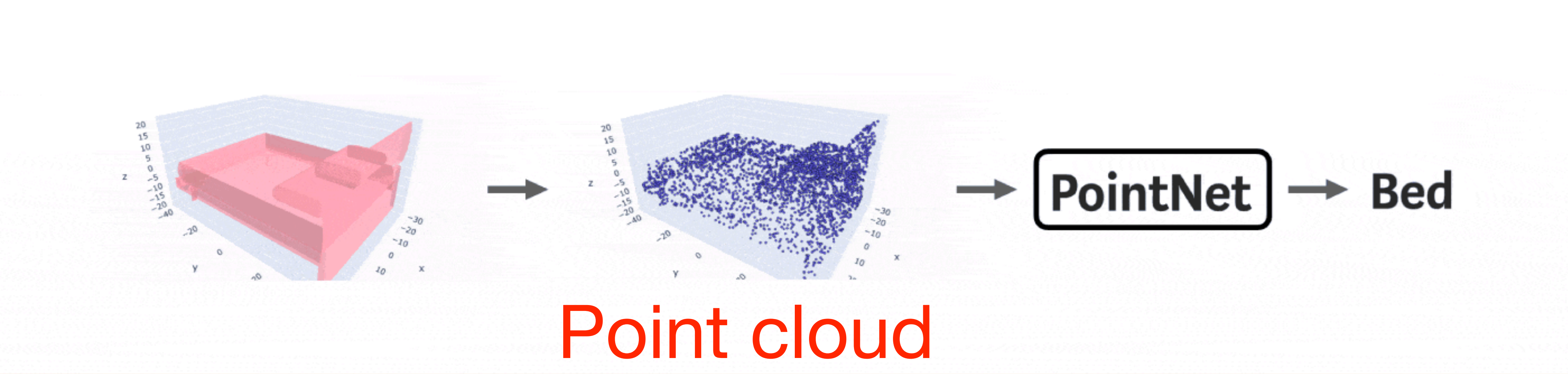
# Machine Learning Tools for Jets

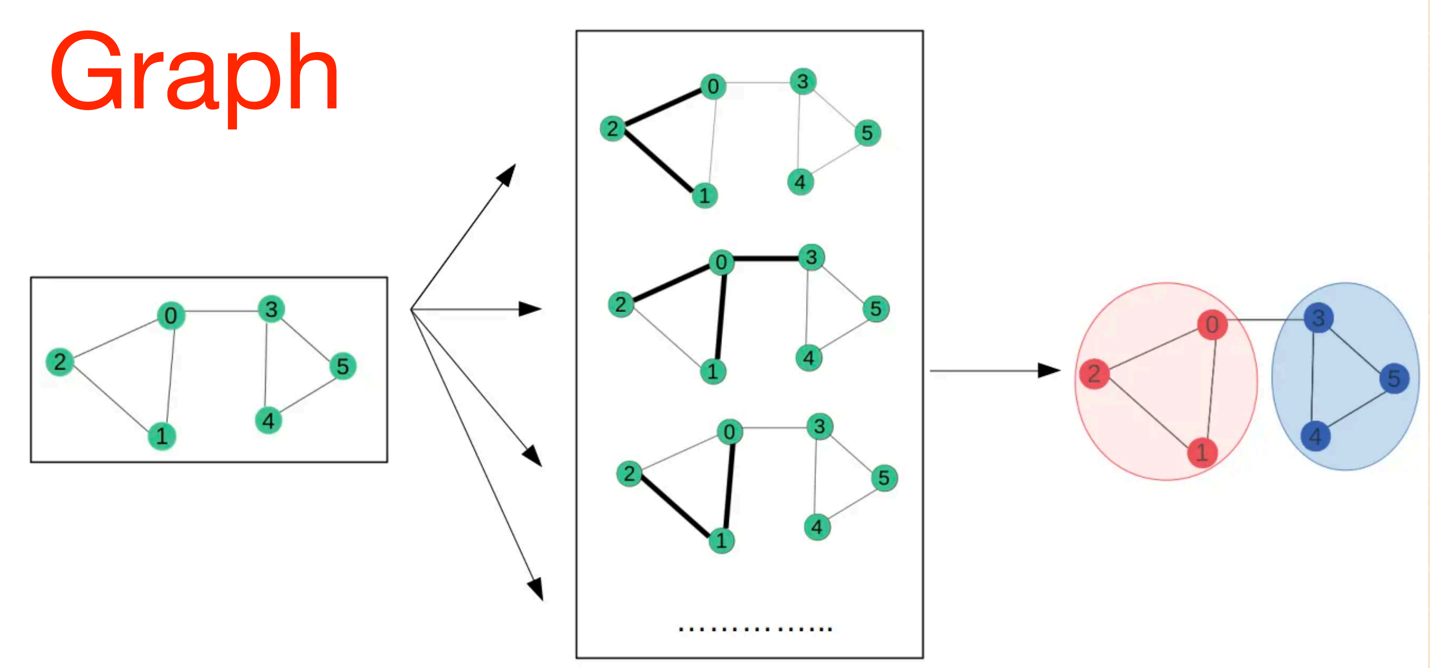



- Many tools suit a lot for jets
- A tremendous amount of applications in HEP arXiv:2102.02770

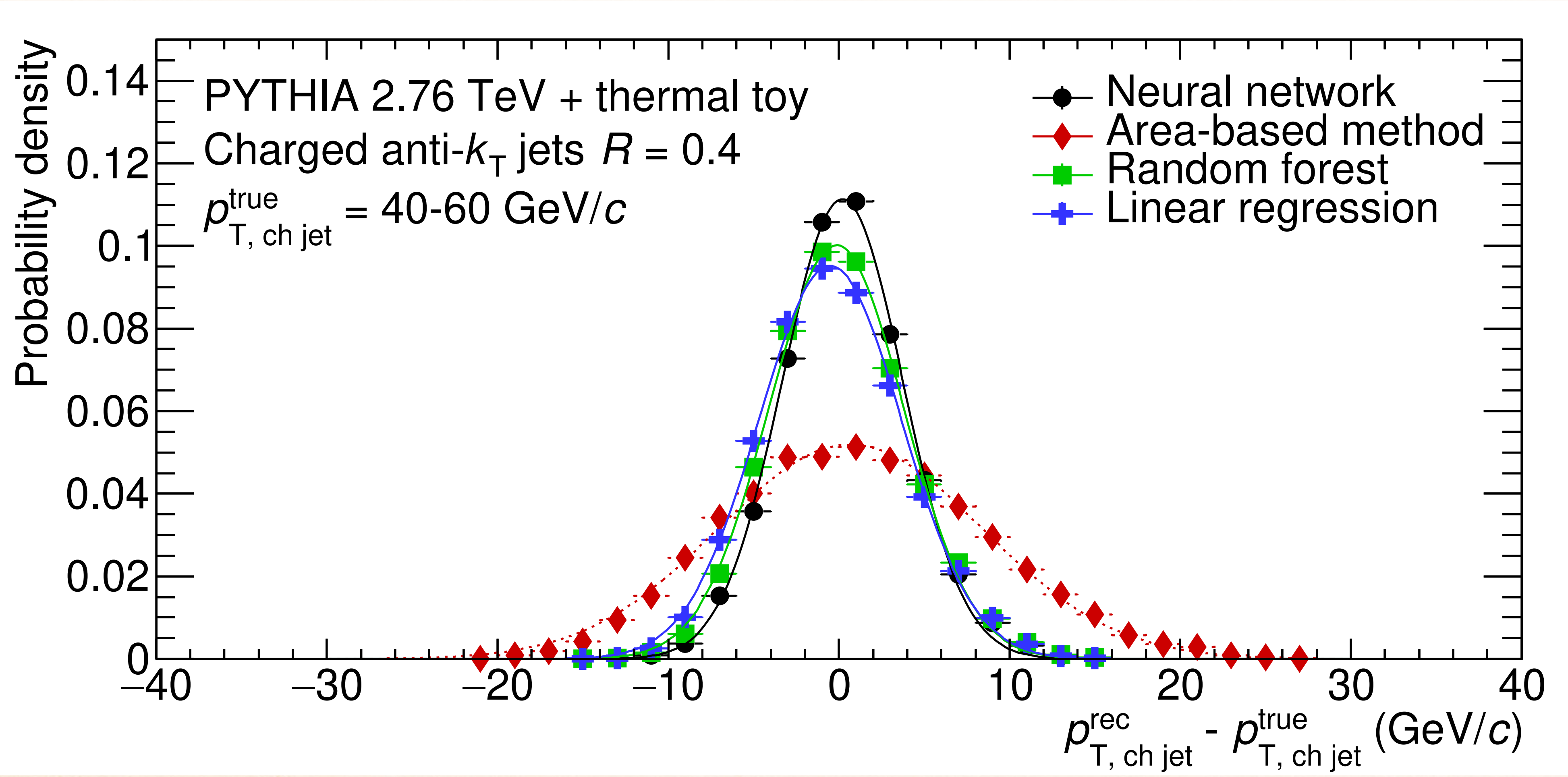
# Outline

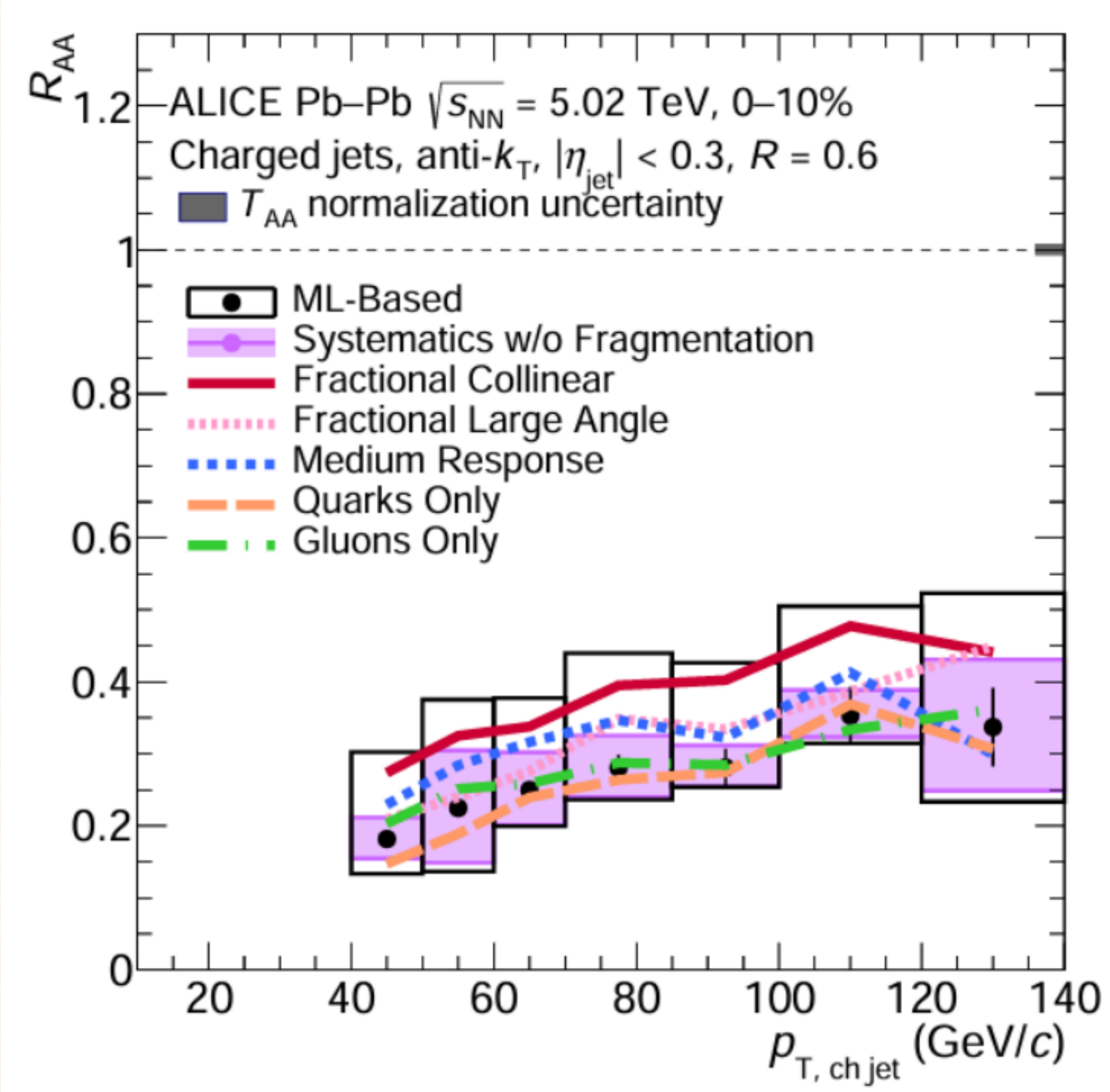
- Reconstruction of Jet Momentum
- Prediction of Jet Energy Loss

# Jet Momentum Reconstruction in ALICE

R. Haake and C. Loizides, Phys. Rev. C 99, 064904 (2019)

ALICE: PLB 849 (2024): 138412



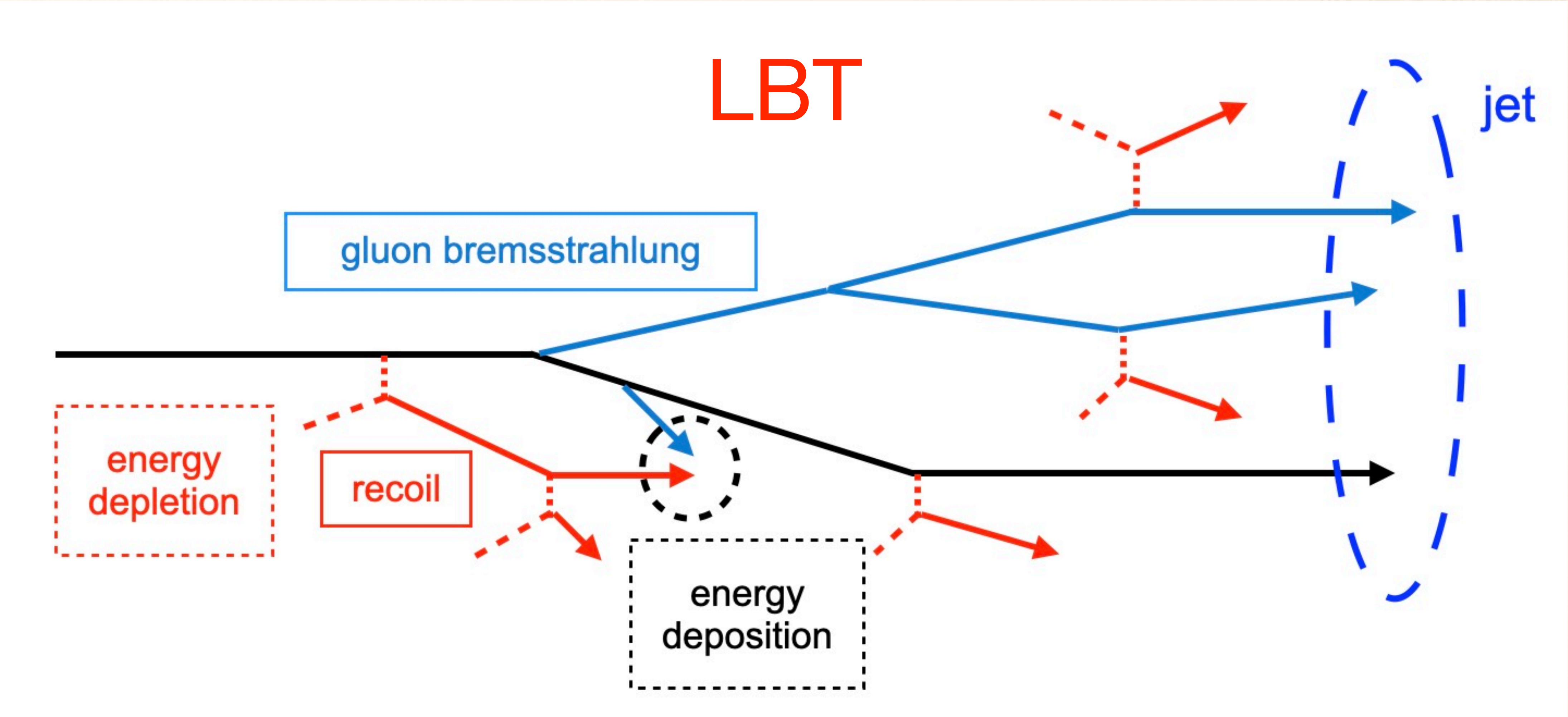




- ML-based methods outperform the traditional Area-based method in reconstructing jet momentum in the presence of the QGP background
- There ML models are trained using PYTHIA jets without quenching effects

- Successful application in ALICE measurement
- ML models are trained separately using PYTHIA jets + toy models of different quenching effects to avoid the training bias and quantify the systematic errors
- Why not train different jet fragmentation patterns simultaneously?

# PYTHIA & LBT Models

- Vacuum jets: PYTHIA8 model
- Jet interaction with QGP: Linear Boltzmann Transport (LBT) model
- Pb-Pb collisions in 0-10% centrality at $\sqrt{S}$=5.02TeV
- QGP background: a toy thermal model
- Reconstructed jets with anti-$k_T$, R=0.4




- Target $p_T$ —sum of the (PYTHIA/LBT) jet particle $p_T$ within jet cone, in the presence of background particles

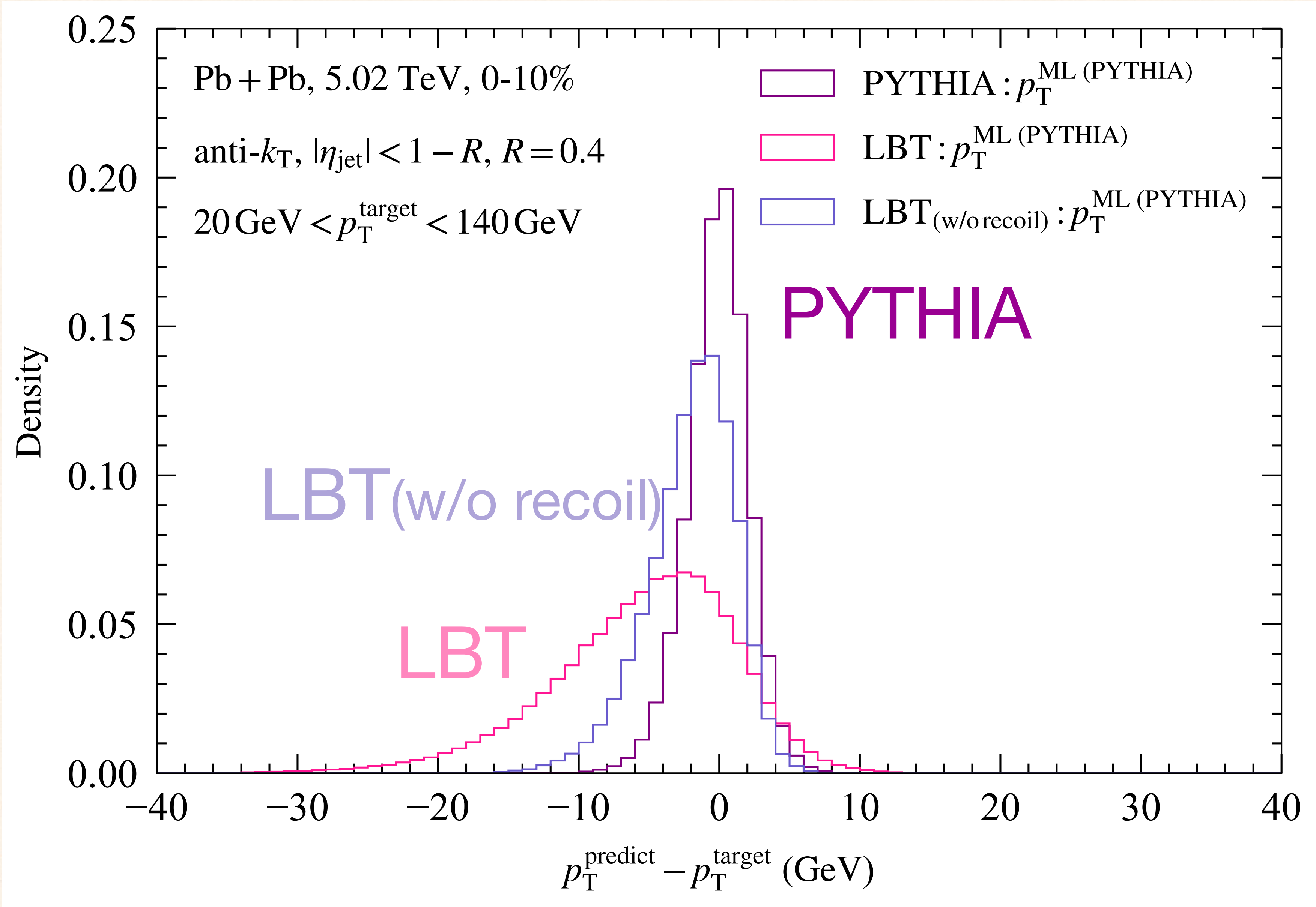
Background

| $\pi^+ + \pi^-$ | |
|---|---|
| $\mathrm{d}N/\mathrm{d}y$ | $< p_T >$ |
| 1699.80 | 0.5682 GeV |

$p_T$ follows a Boltzmann distribution

# Effects of realistic medium modifications on ML performance

Ran Li, YLD, Shanshan Cao, PLB 870 (2025) 139940



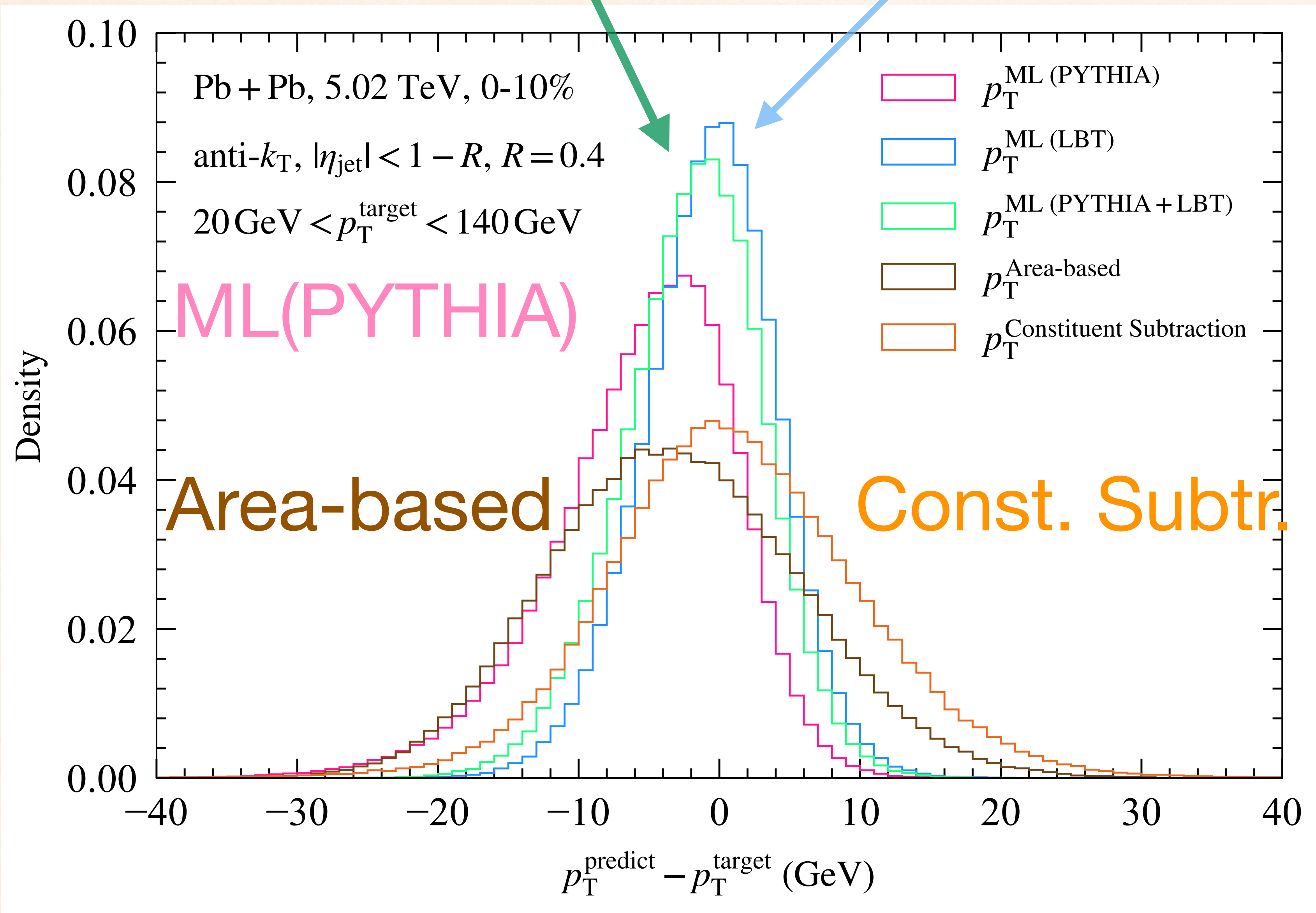

Test PYTHIA-trained ML on different datasets

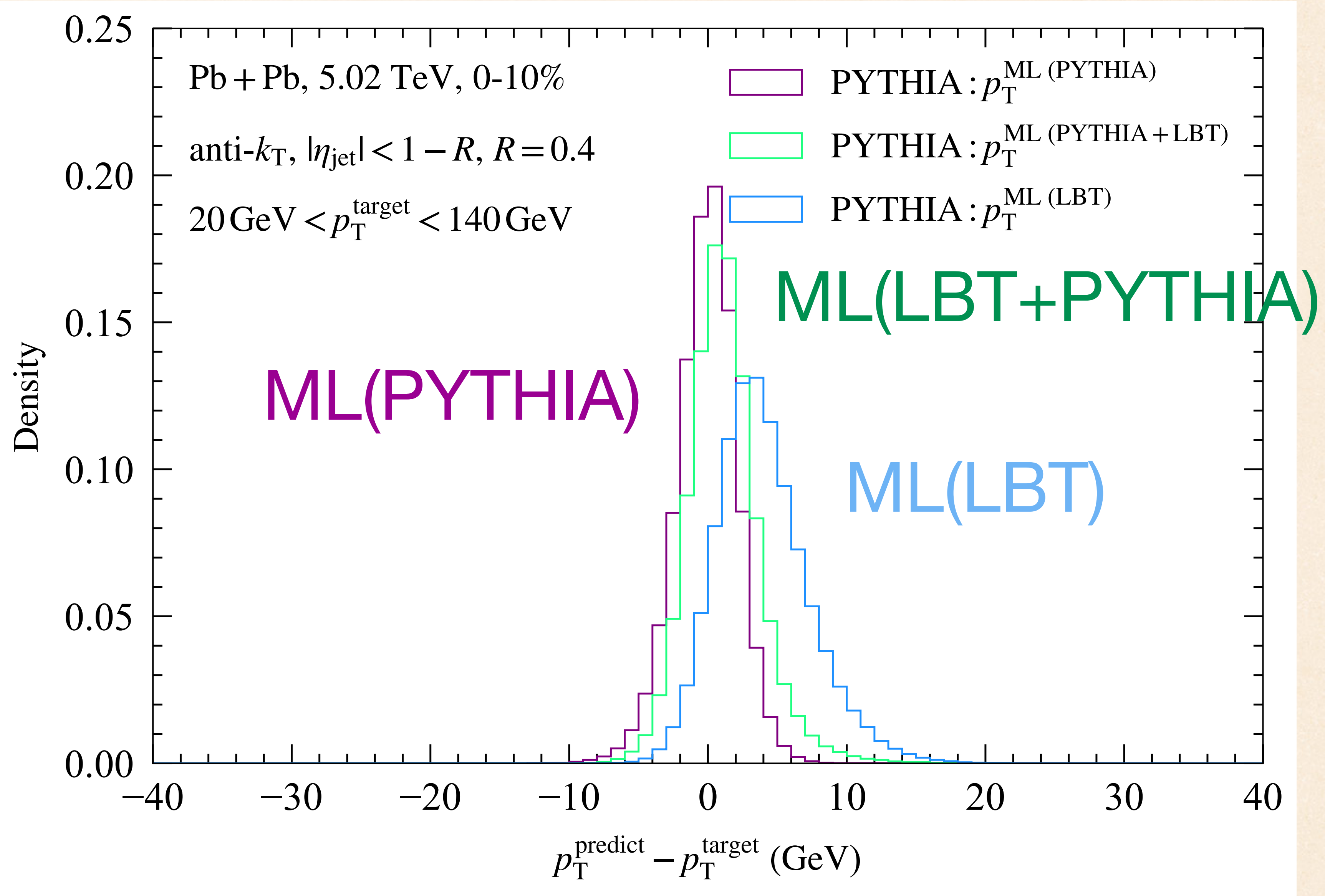
- PYTHIA-trained ML has a prediction bias when applied on LBT jets.
- This bias is reduced when the ML model is applied on LBT jets without recoil particles: Recoil particles may be considered as the background.
- Conclusion: ML model trained using vacuum jets is incapable of recognizing medium response particles in quenched jets.

# Performance of different ML models

Ran Li, YLD, Shanshan Cao, PLB 870 (2025) 139940

ML(LBT+PYTHIA) ML(LBT)




Testing on LBT data

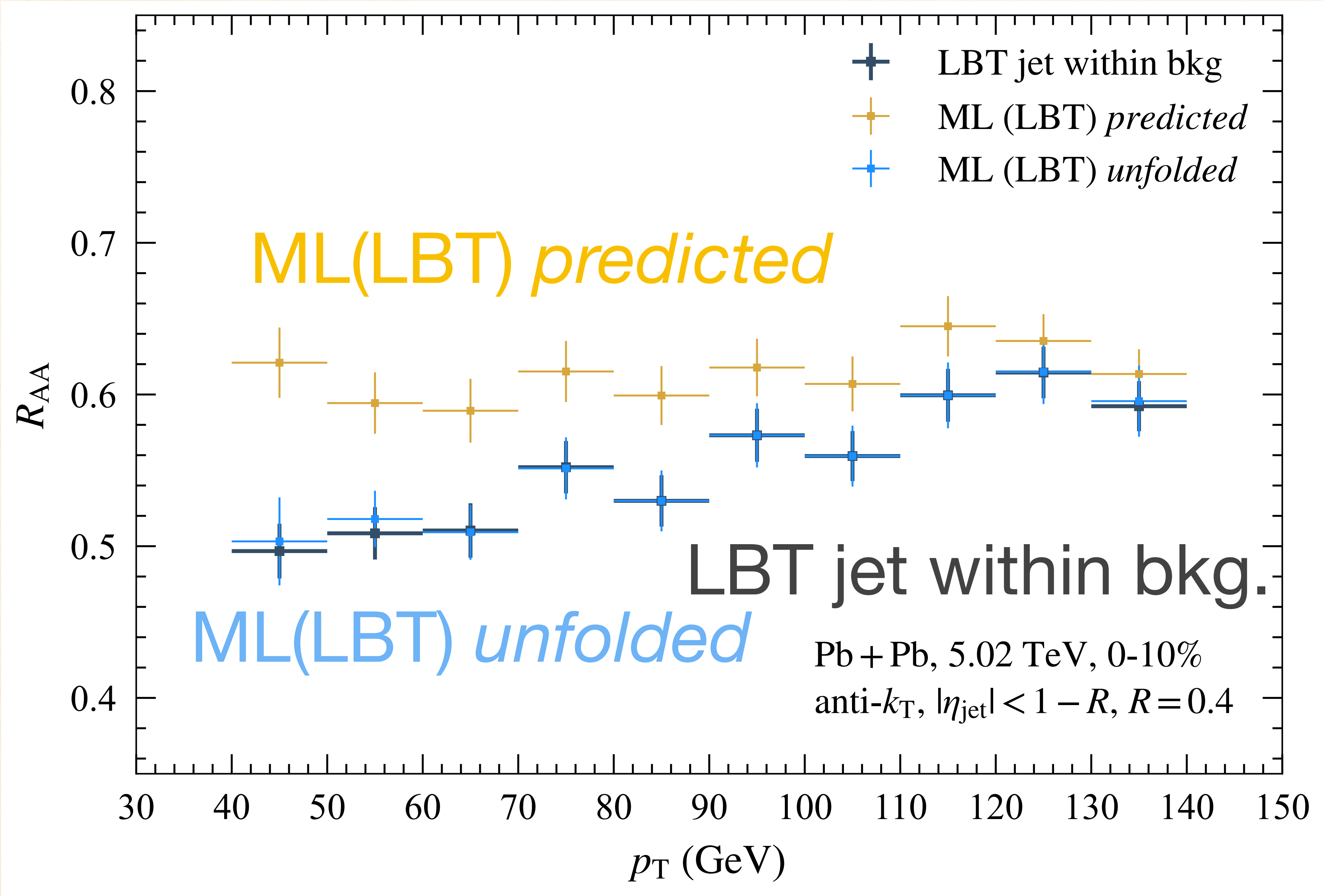
Testing on PYTHIA data

- The bias is reduced when the ML model is trained by LBT jets directly or by LBT+PYTHIA jets.
- LBT+PYTHIA-trained ML can make accurate predictions on PYTHIA and LBT data simultaneously, demonstrating strong robustness.
- ML methods are more accurate than conventional methods (Area-based & Constituent Subtraction ).

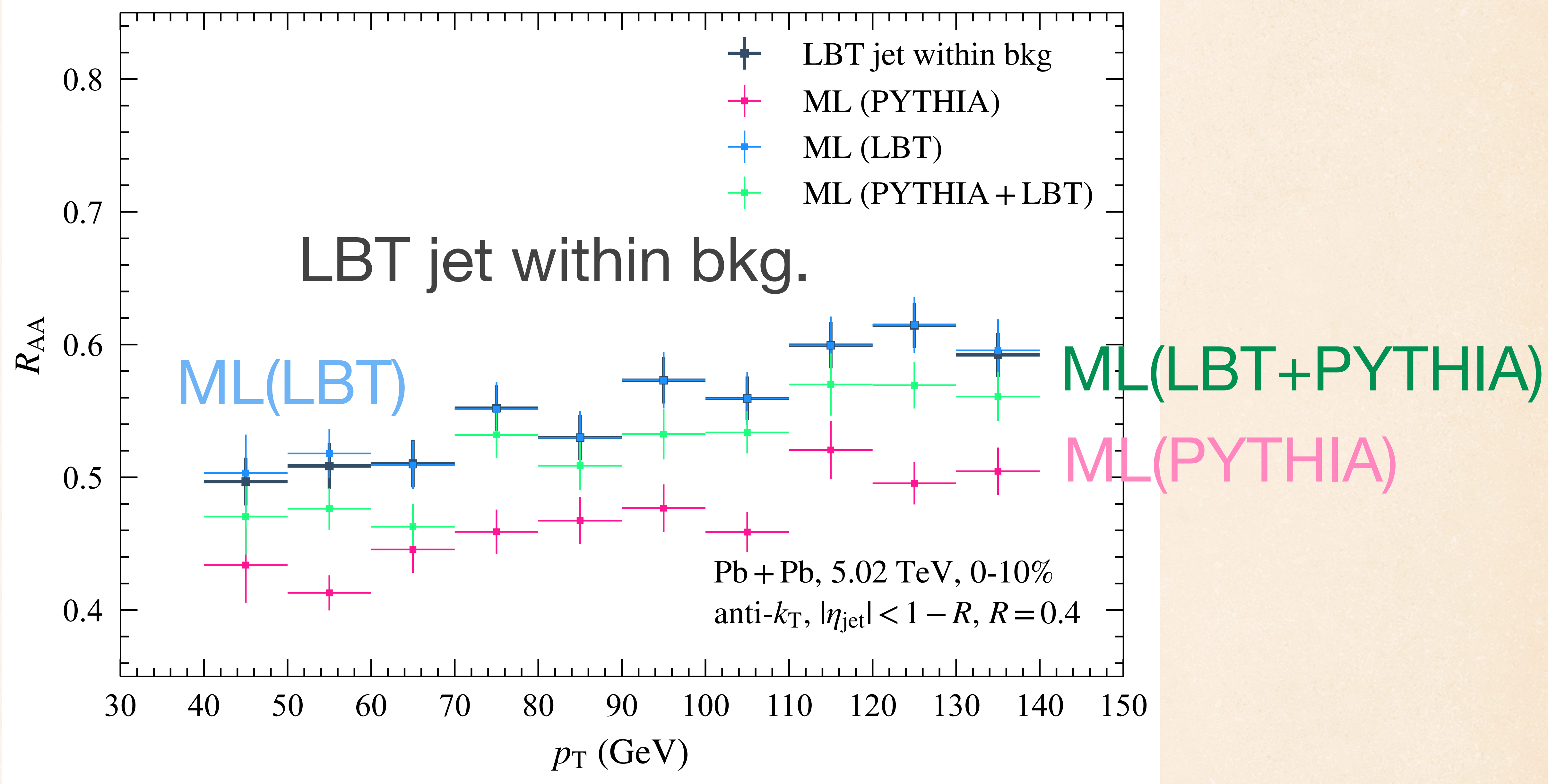
# Jet Nuclear Modification Factor $R_{AA}$

Ran Li, YLD, Shanshan Cao, PLB 870 (2025) 139940

Effect of unfolding

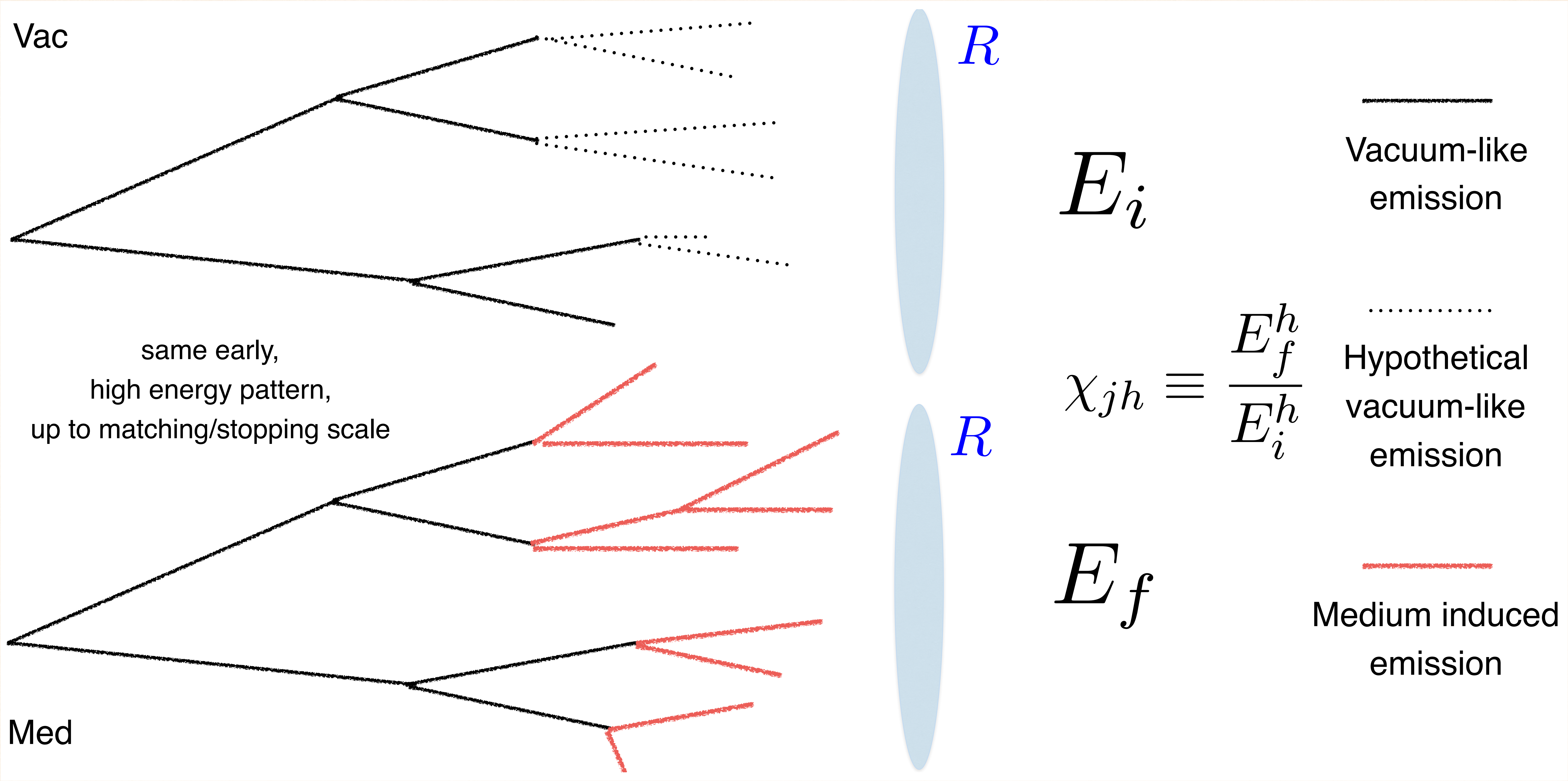



Different ML models




- Direct application of ML(LBT) over-predicts the jet $R_{AA}$ due to convolution of residual distribution of $p_T^{\text{target}}$ & steeply falling $p_T$ spectra of jets, which can be effectively corrected by unfolding procedure.
- $R_{AA}$ from ML(LBT) & ML(LBT+PYTHIA) are closer to that of the target baseline "LBT jet within bkg." than $R_{AA}$ from ML(PYTHIA).
- Conclusion: it is necessary to include realistic quenching effects in training ML models.

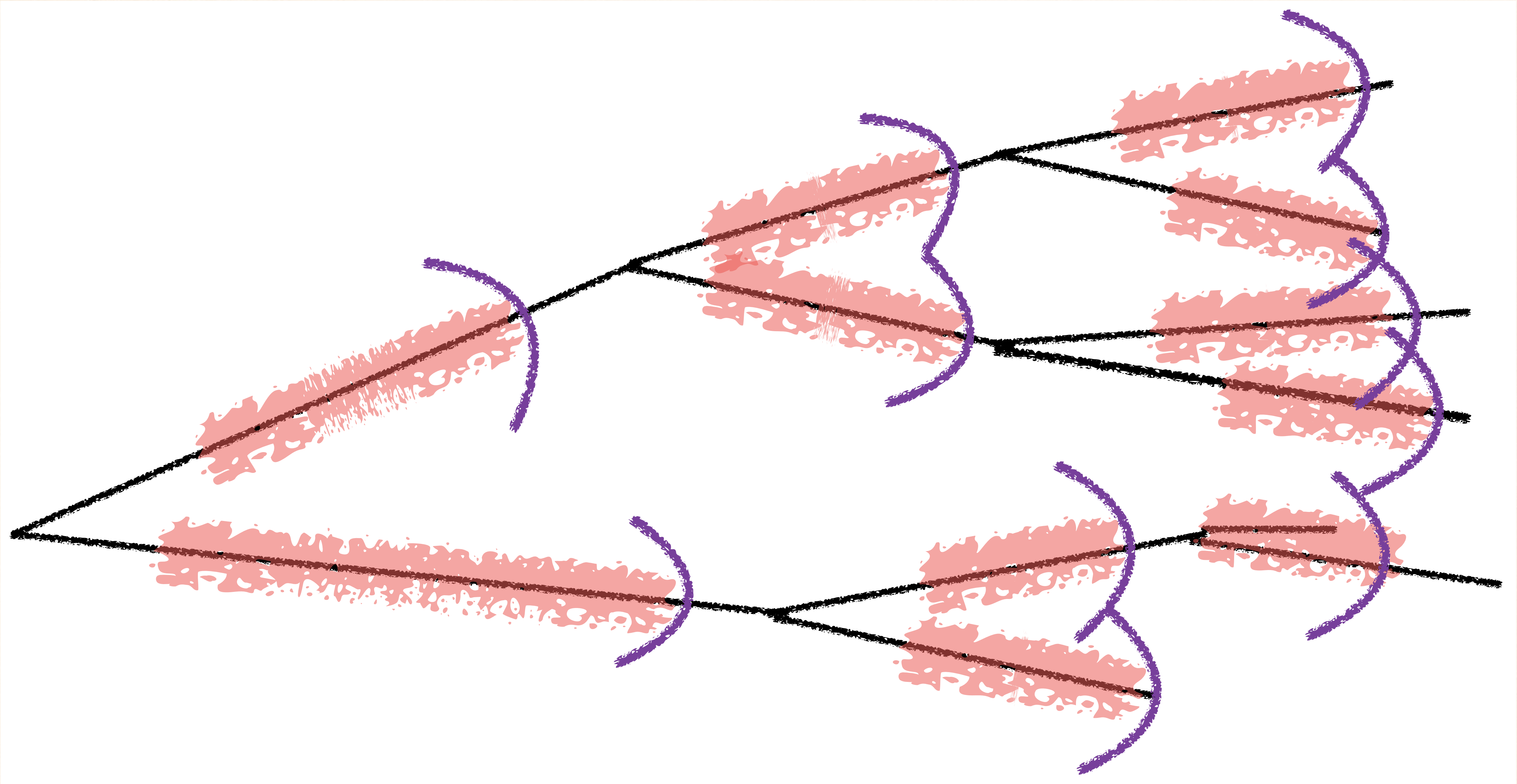
# From event-averaged observables to jet-by-jet energy loss



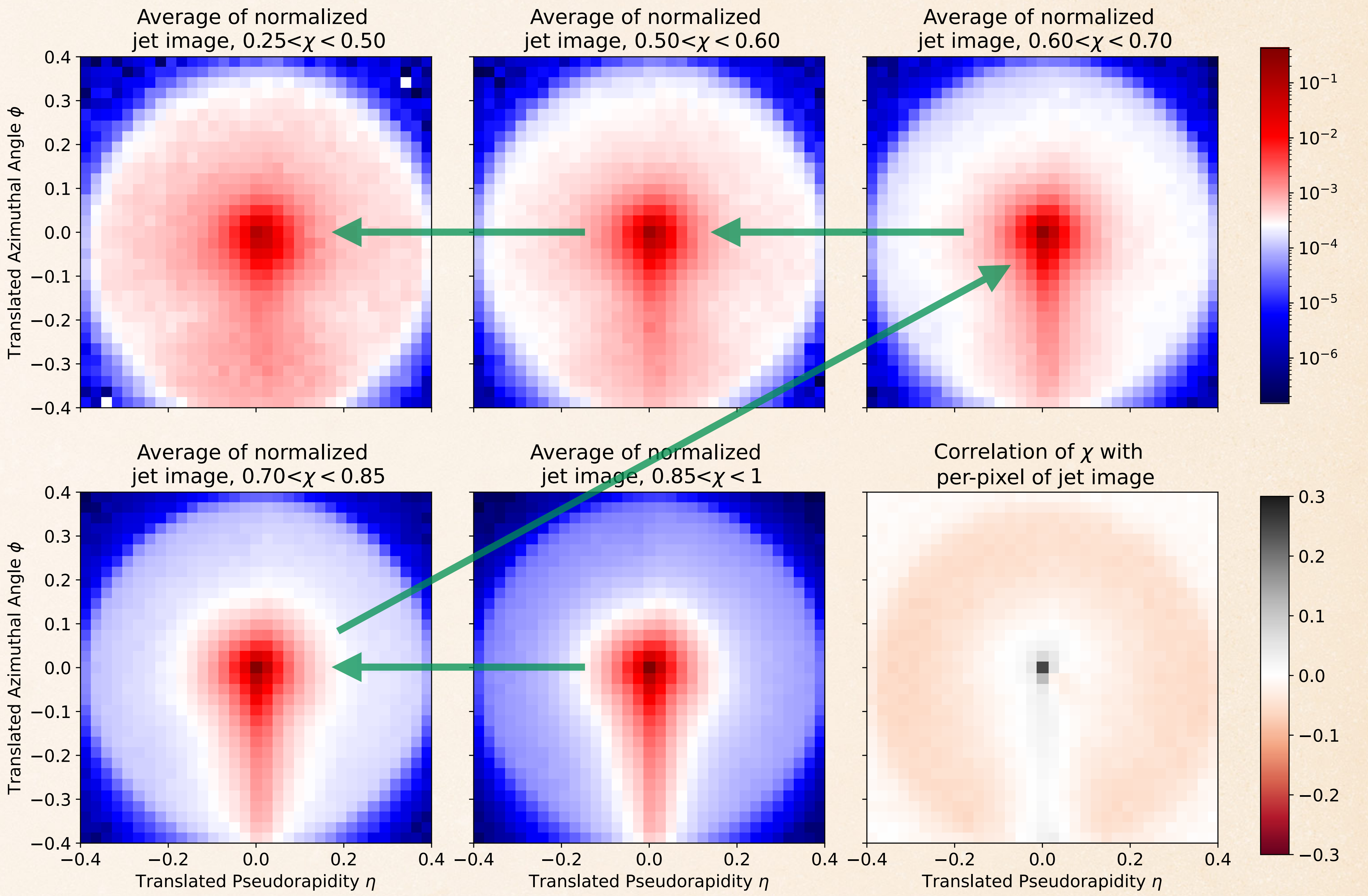

Generalizable jet energy loss ratio

# Prior efforts: Strong/weak hybrid model

YLD, D. Pablos and K. Tywoniuk, JHEP03(2021)206



Strongest quenching




Weakest quenching

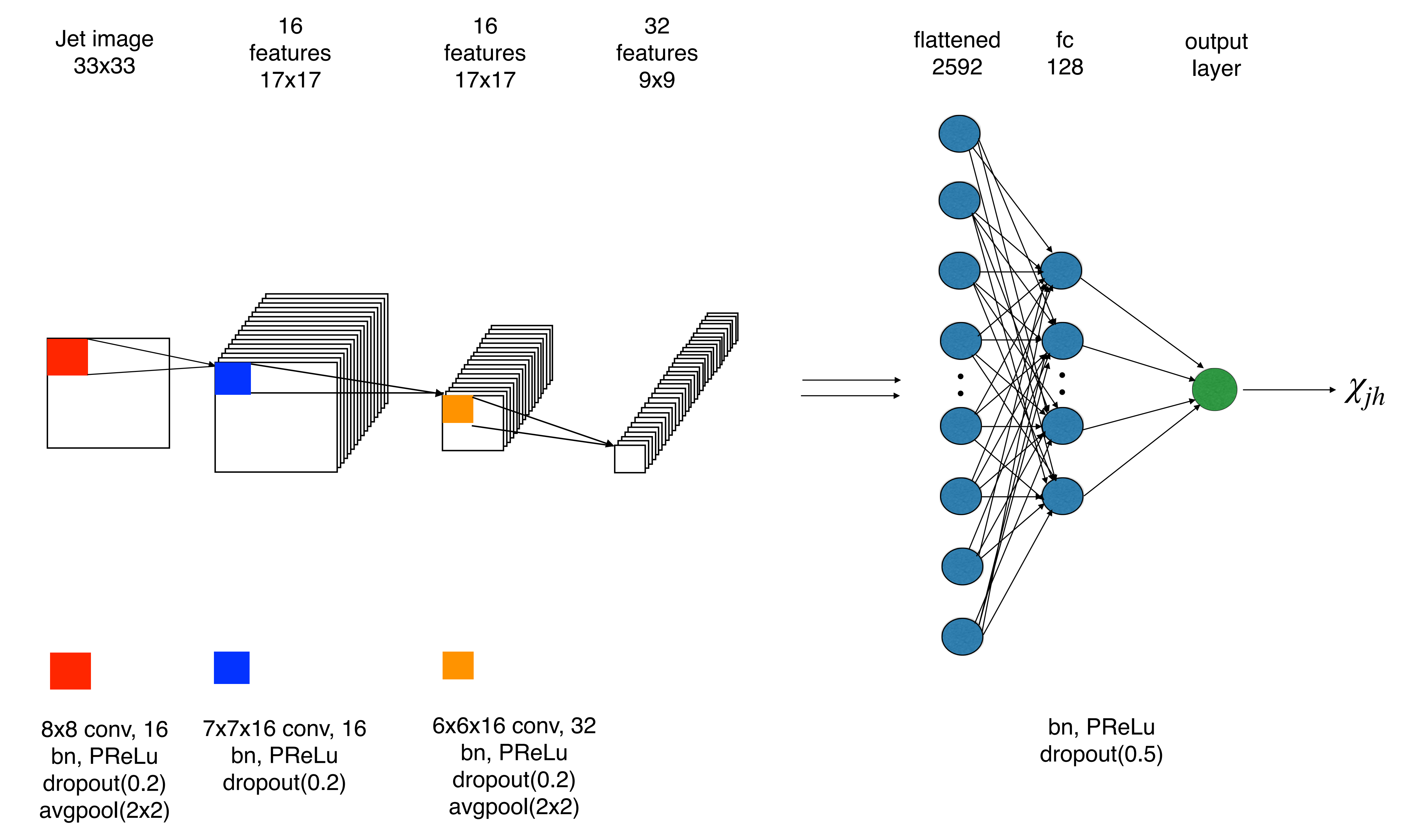
Pearson Coefficients

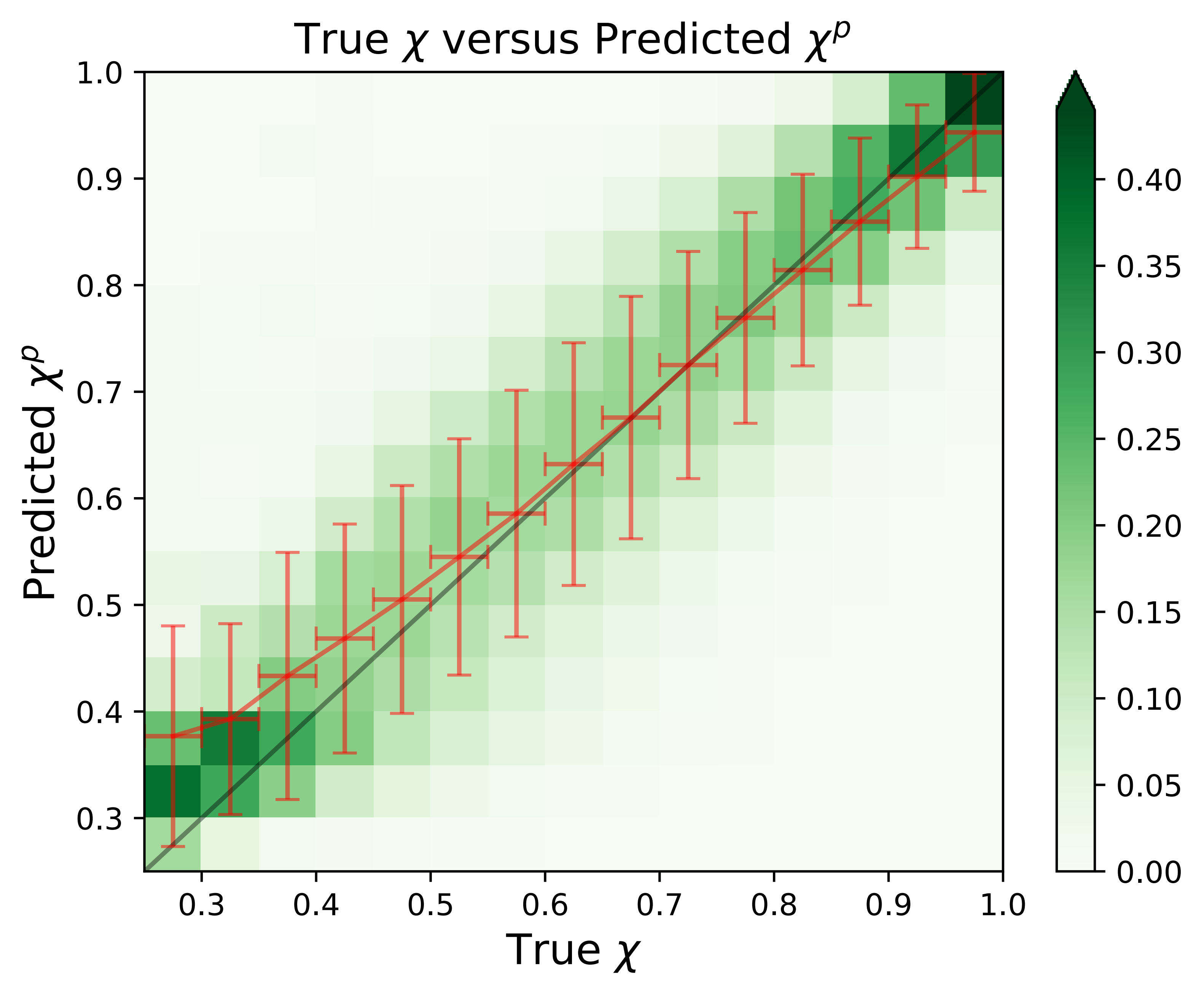
- PYTHIA8 down to hadronization scale
- Strongly coupled energy loss at every stage
- Hadrons from the hydro. wake (medium response)

Jet-soft correlation: quenching transfer the energy of soft particles to large angles

# CNN Prediction of Jet Energy Loss

YLD, D. Pablos and K. Tywoniuk, JHEP03(2021)206




- Regression analysis of jet energy loss between the twin jets with CNN.
- Well predicted for a wide range of $\chi$.
- Infer out-of-the-cone jet energy loss from in-cone jet constituents.

Performance: Jet image
=jet shape+FF+substructures

Interpretability!

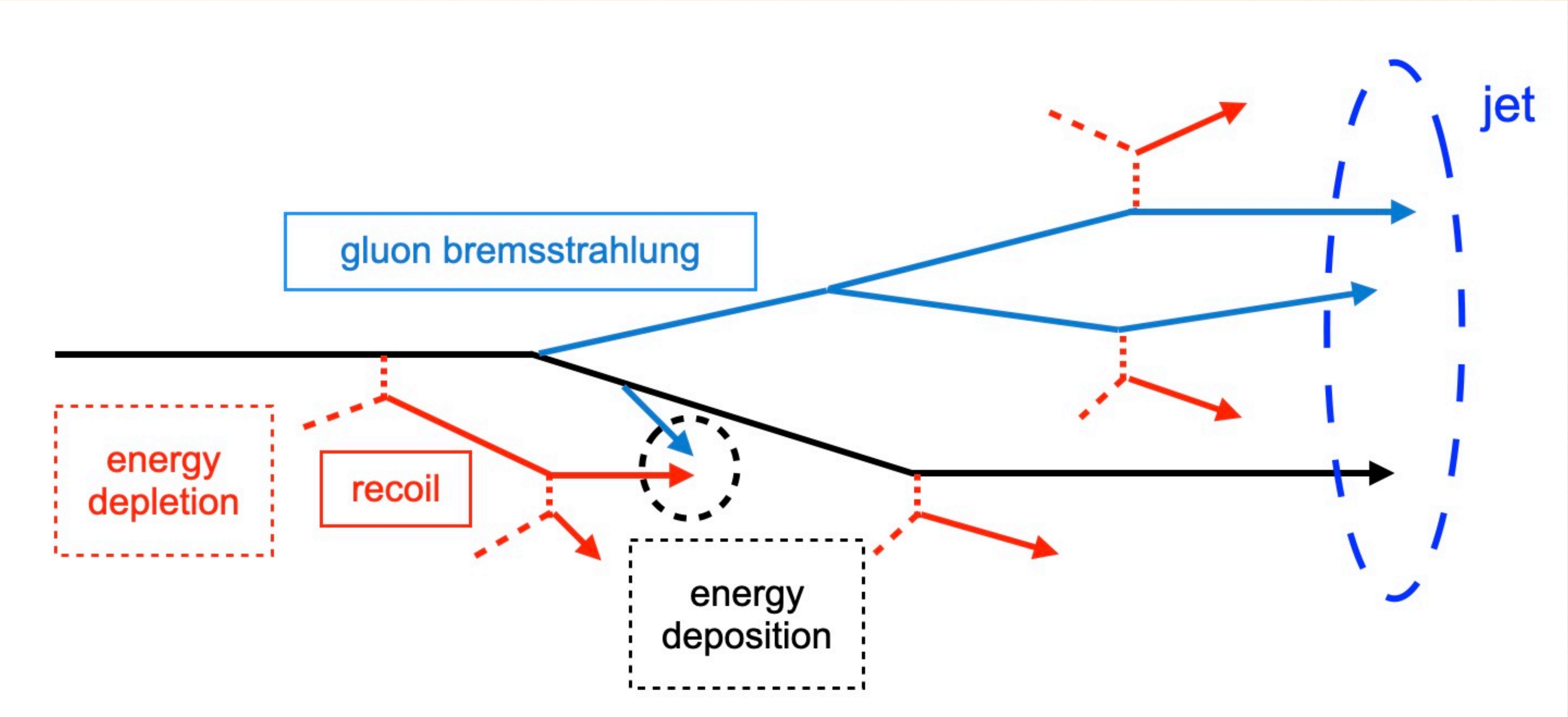
# Challenges from the QGP background

- Vacuum jets: PYTHIA8 model
- Jet interaction with QGP: Linear Boltzmann Transport (LBT) model
- Pb-Pb collisions in 0-10% centrality at √S=5.02TeV
- QGP background: a toy thermal model

Background

| $\pi^+ + \pi^-$ | |
|---|---|
| $\mathrm{d}N/\mathrm{d}y$ | $< p_T >$ |
| 1699.80 | 0.5682 GeV |

$p_T$ follows a Boltzmann distribution



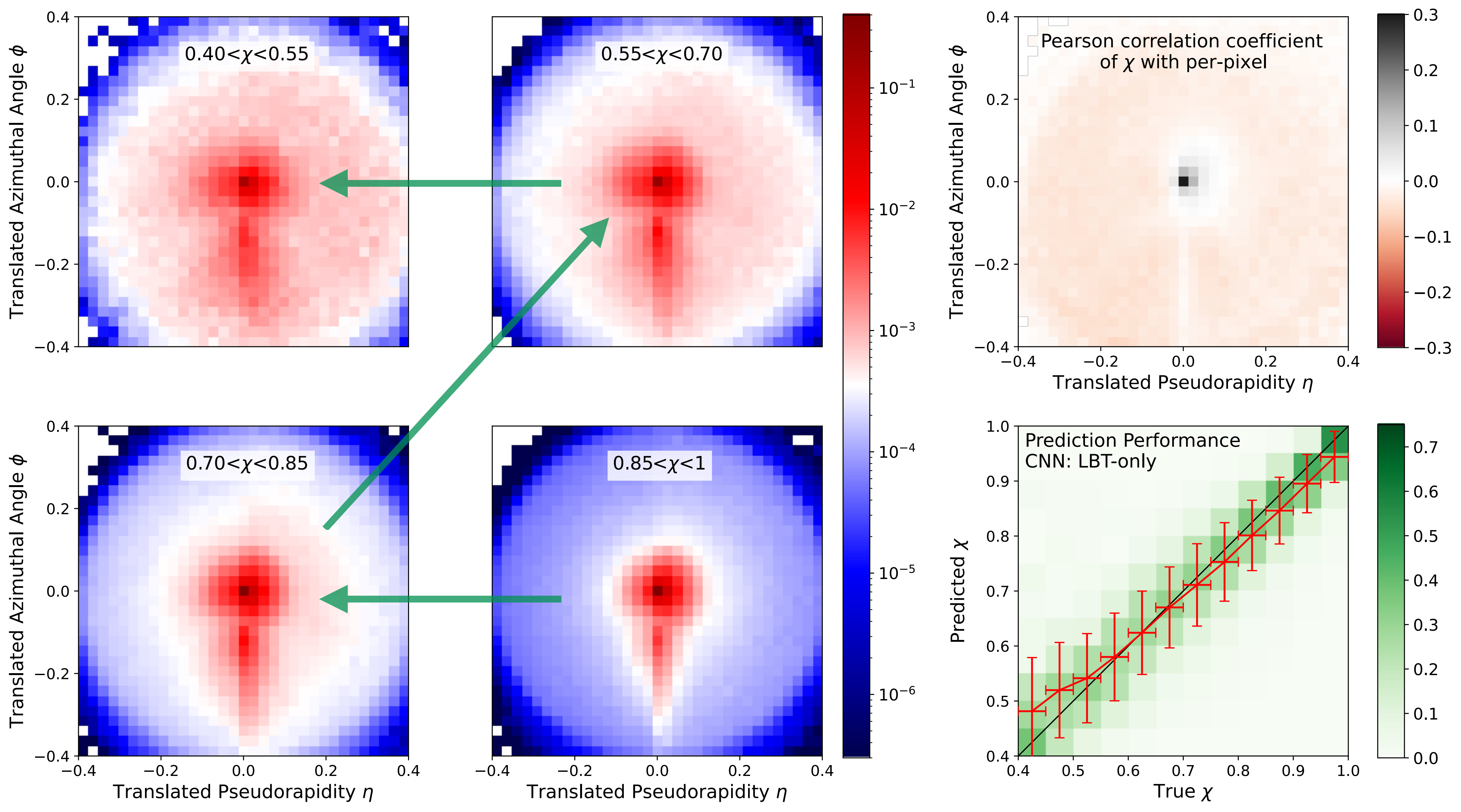

LBT model

- Jet $p_T$ —sum of the LBT jet particle $p_T$ within jet cone, in the presence of background particles

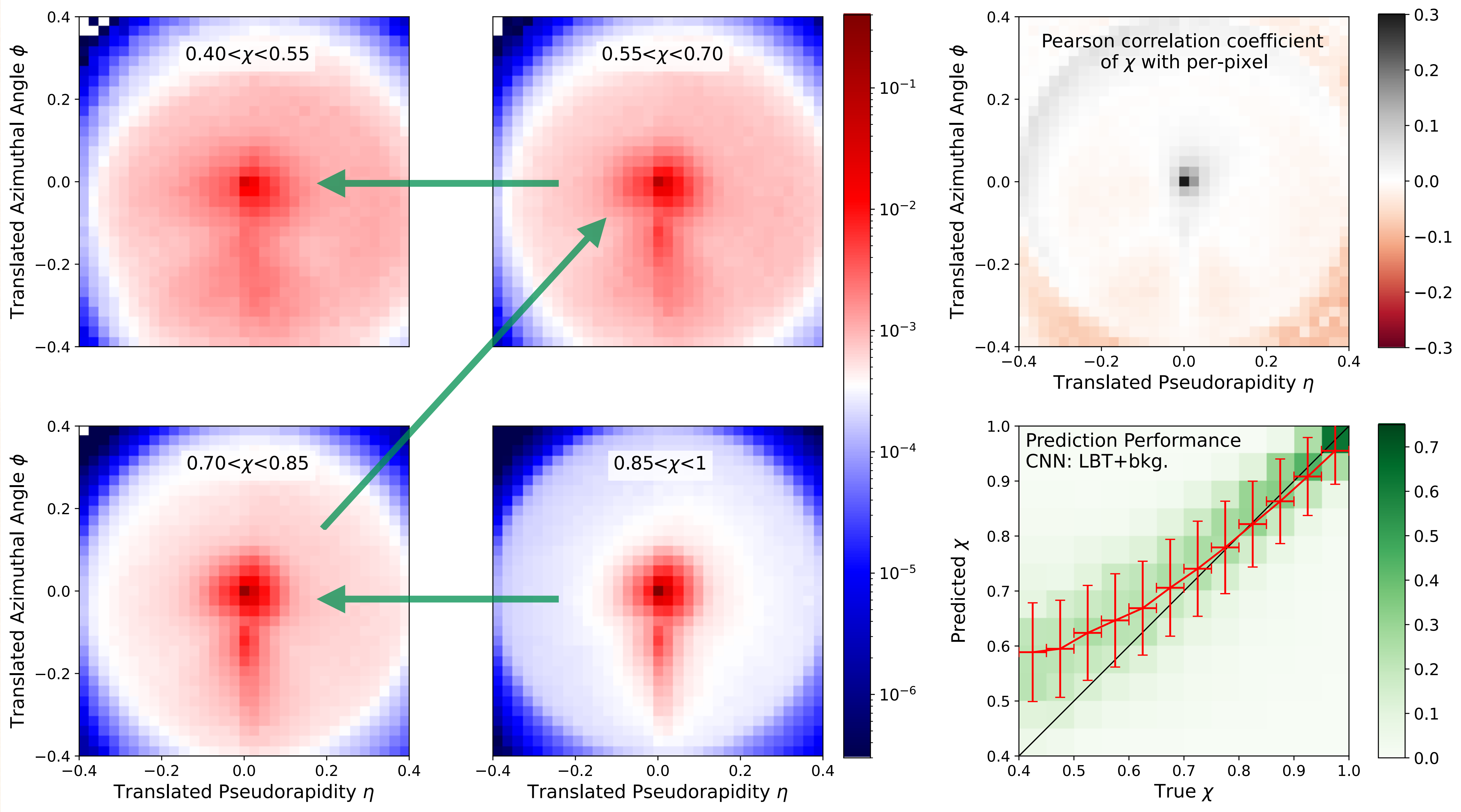
# CNN prediction performance on LBT jets

Ran Li, YLD, Shanshan Cao, PRC 113, 024912 (2026)

LBT jets without QGP background



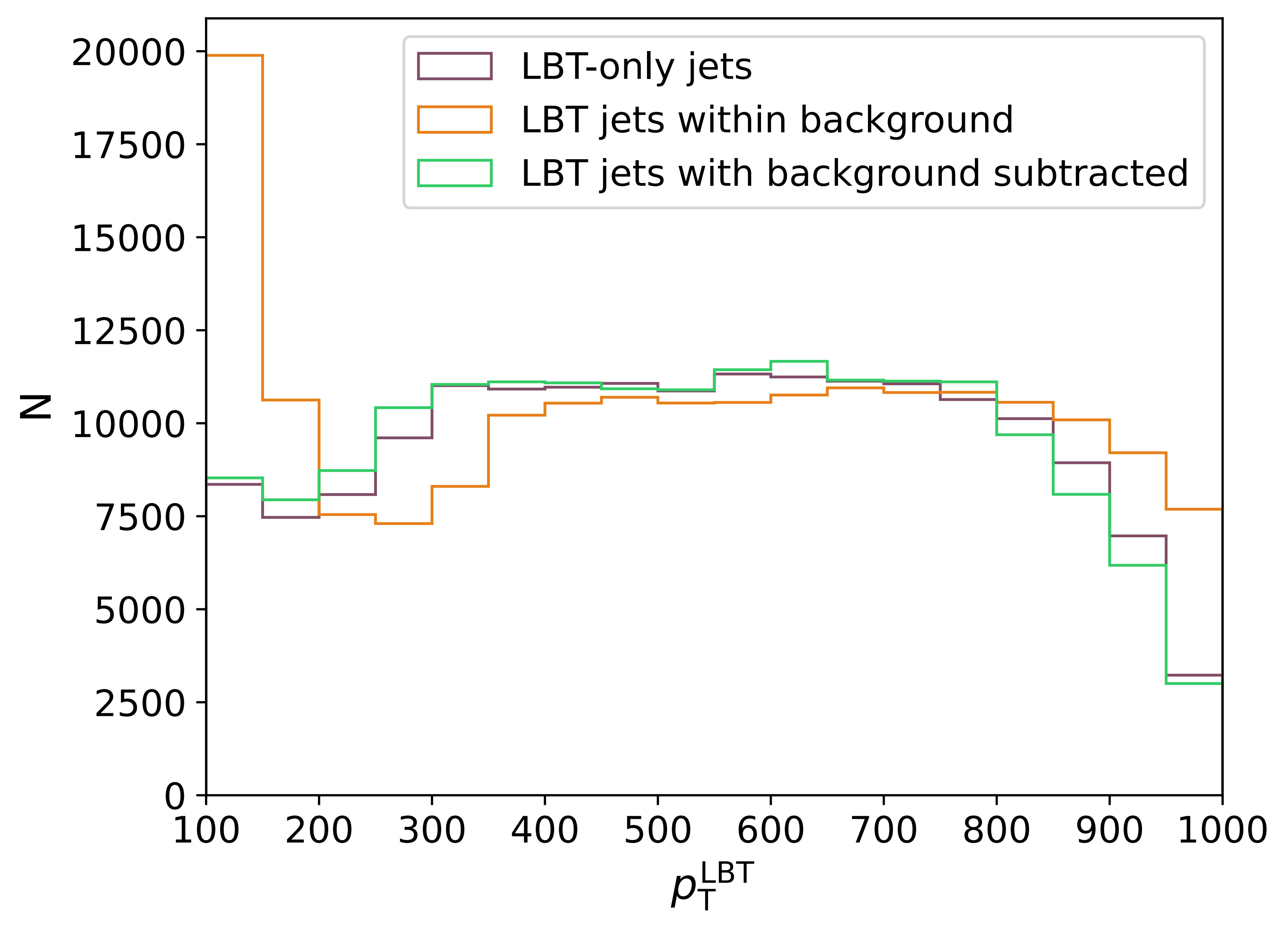

LBT jets within QGP background

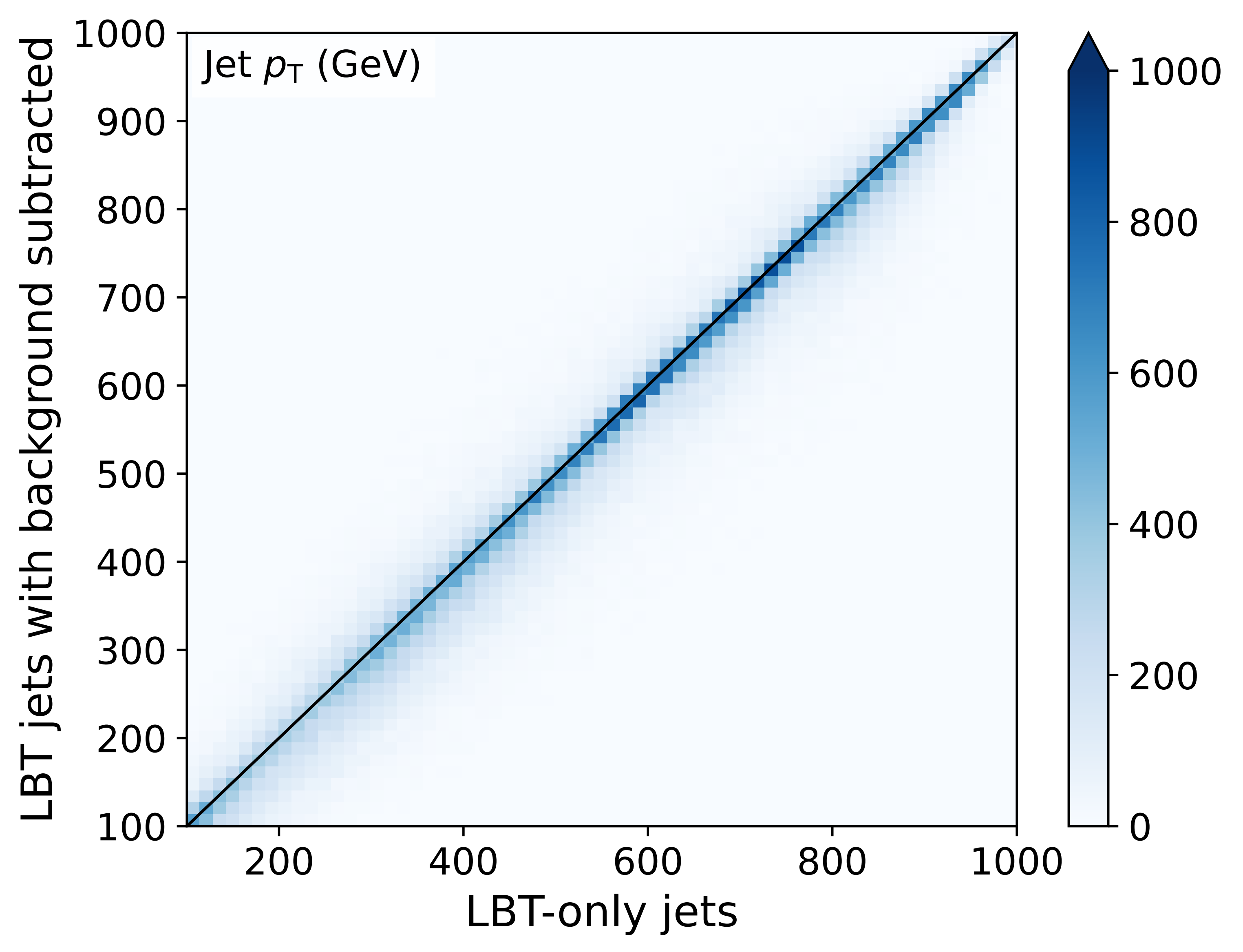


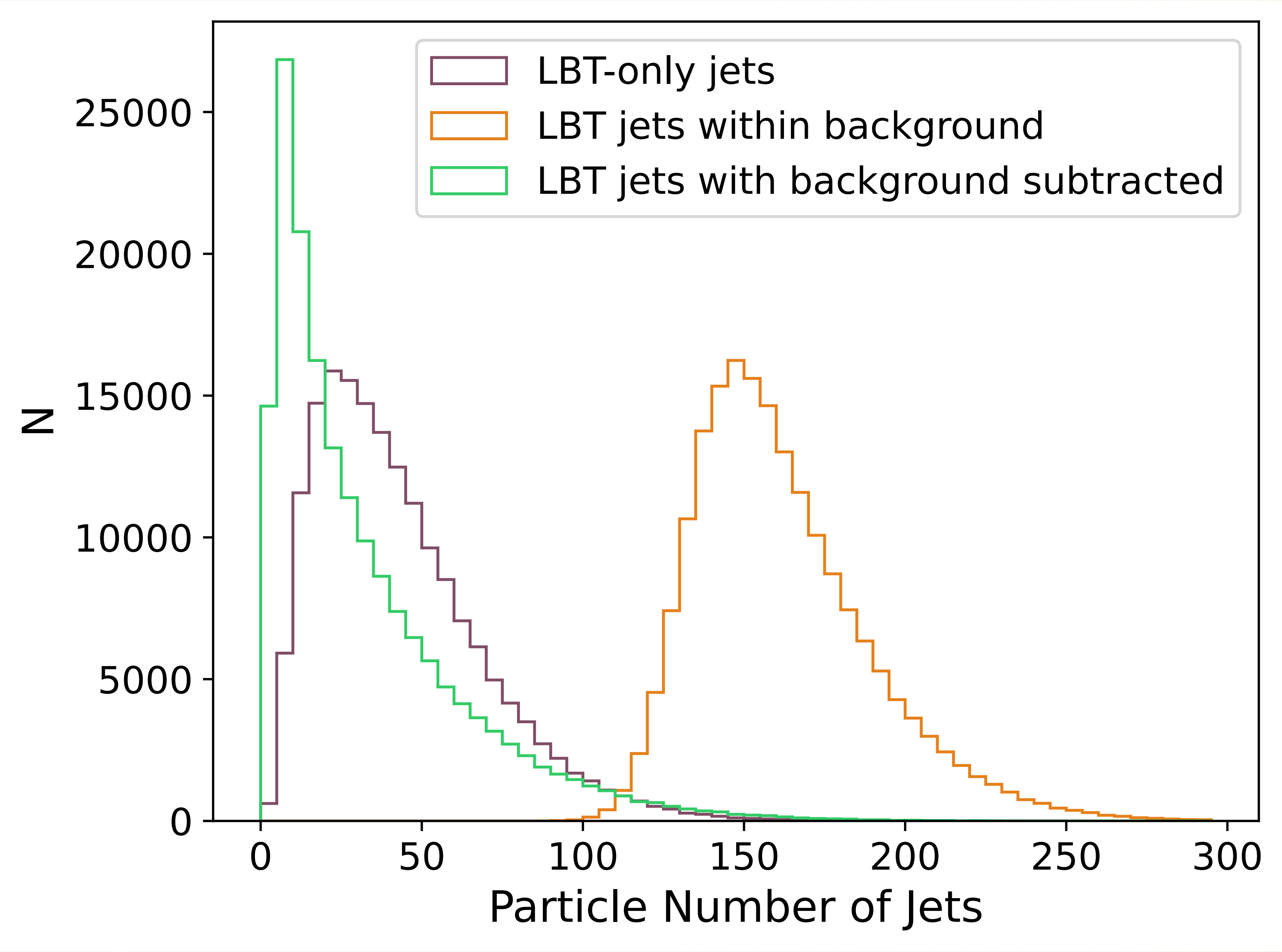

- Retrained CNN prediction performance on LBT jets is comparable with that on jets from Hybrid model
- QGP background weakens the correlations between jet images and energy loss ratio, leading to degraded performance

# Background Removal of Constituent Subtraction

Ran Li, YLD, Shanshan Cao, PRC 113, 024912 (2026)



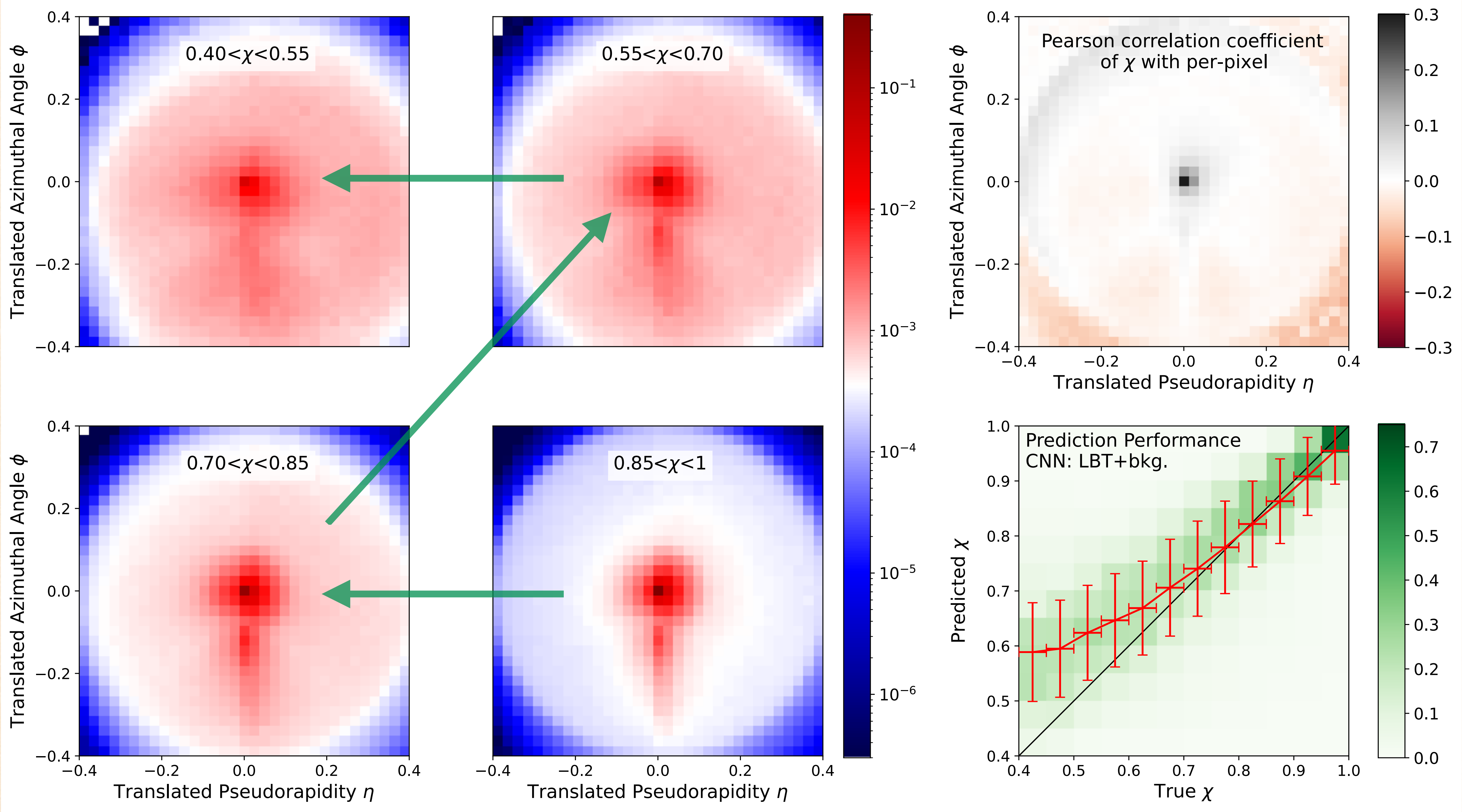




- Jet $p_T$ can be well reproduced.
- Particle number of jets is slightly over-subtracted.

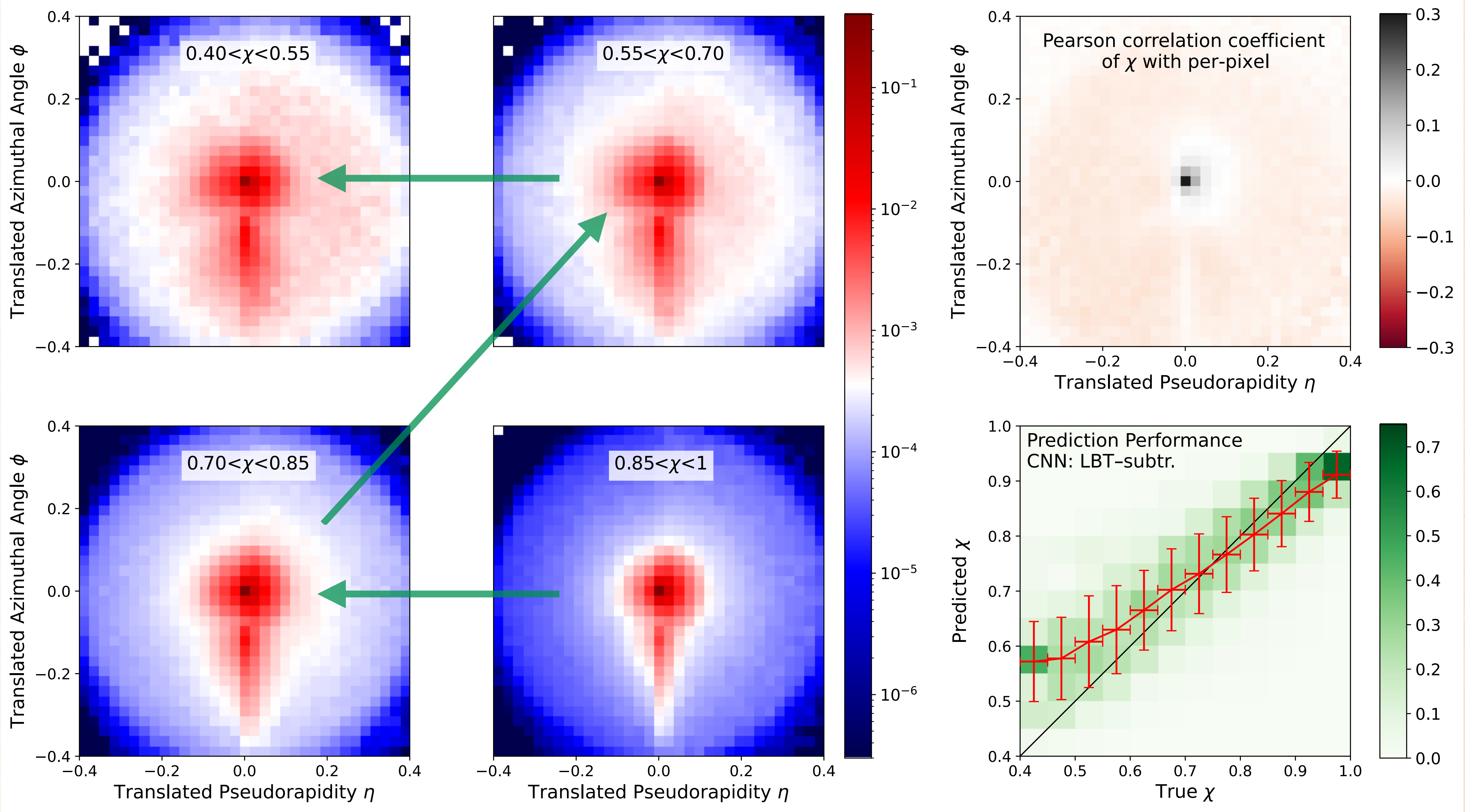
# CNN prediction performance on LBT jets

Ran Li, YLD, Shanshan Cao, PRC 113, 024912 (2026)

LBT jets within QGP background




LBT jets with QGP bkg. subtracted

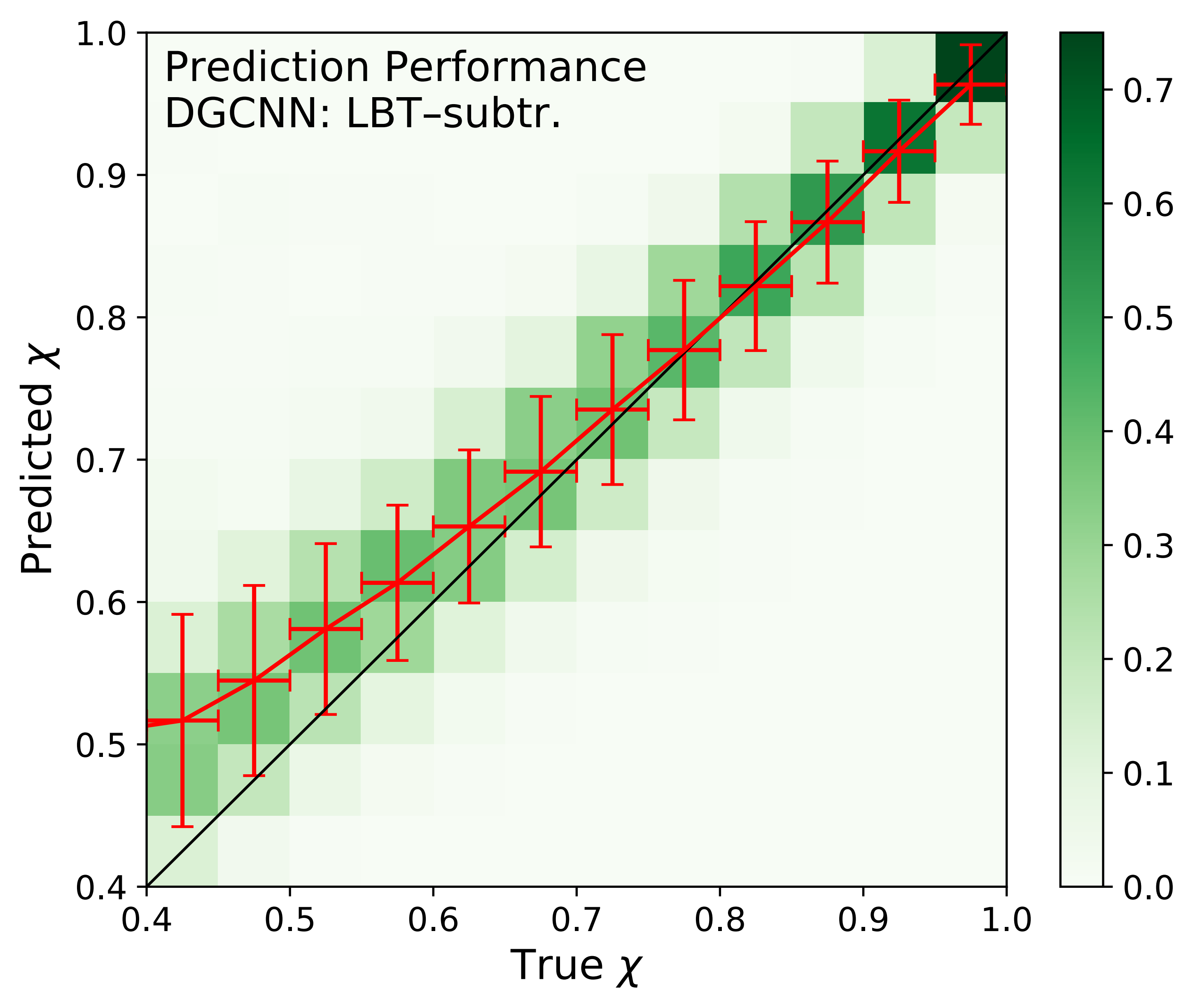



- The correlations between jet images and energy loss ratio is largely recovered via subtraction of background particles (Constituent Subtraction method), as well as the CNN prediction performance
- How to further improve?

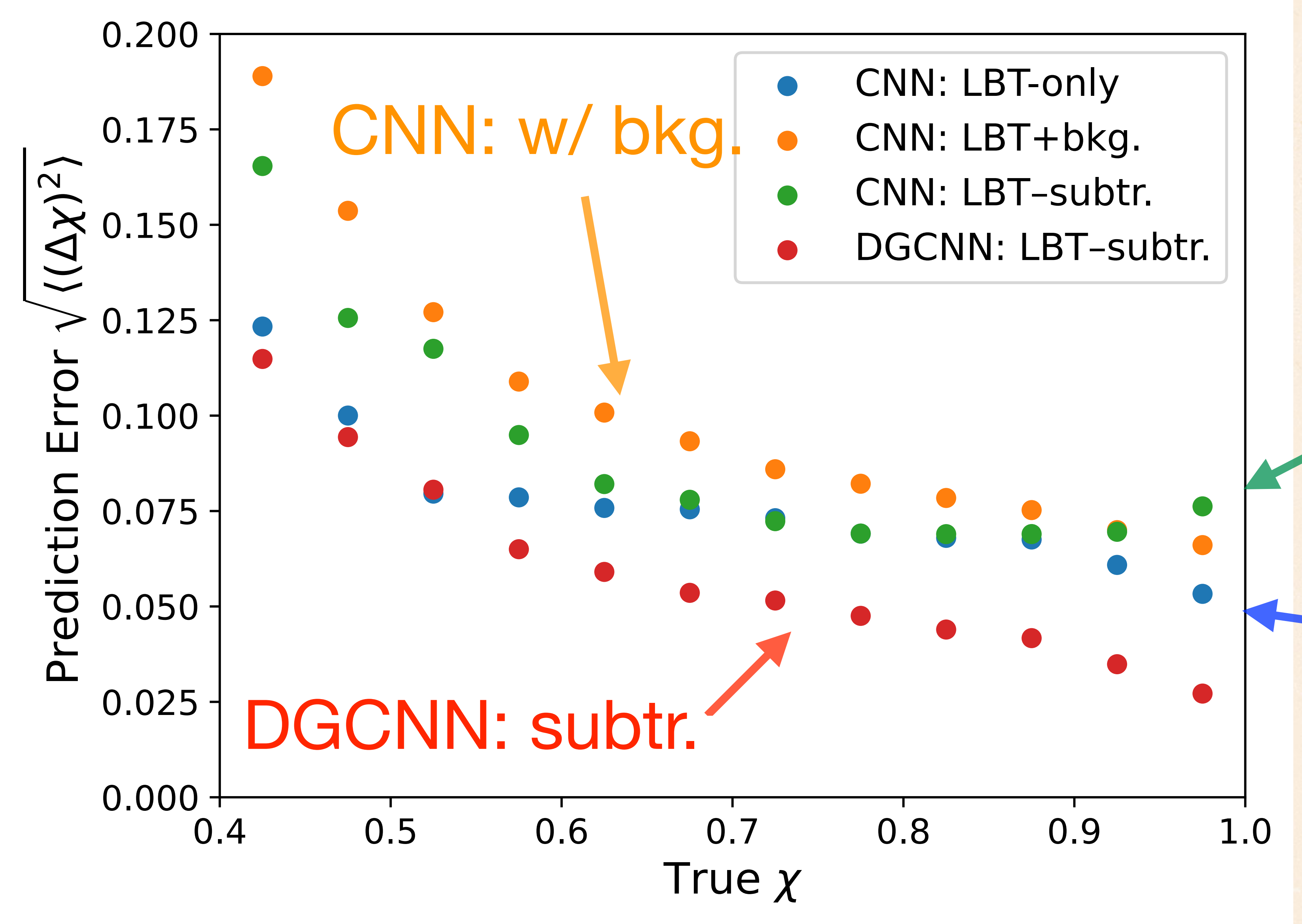
# DGCNN on point cloud of LBT jets after bkg. subtr.

Ran Li, YLD, Shanshan Cao, PRC 113, 024912 (2026)

DGCNN Performance




Prediction Errors




CNN: subtr.

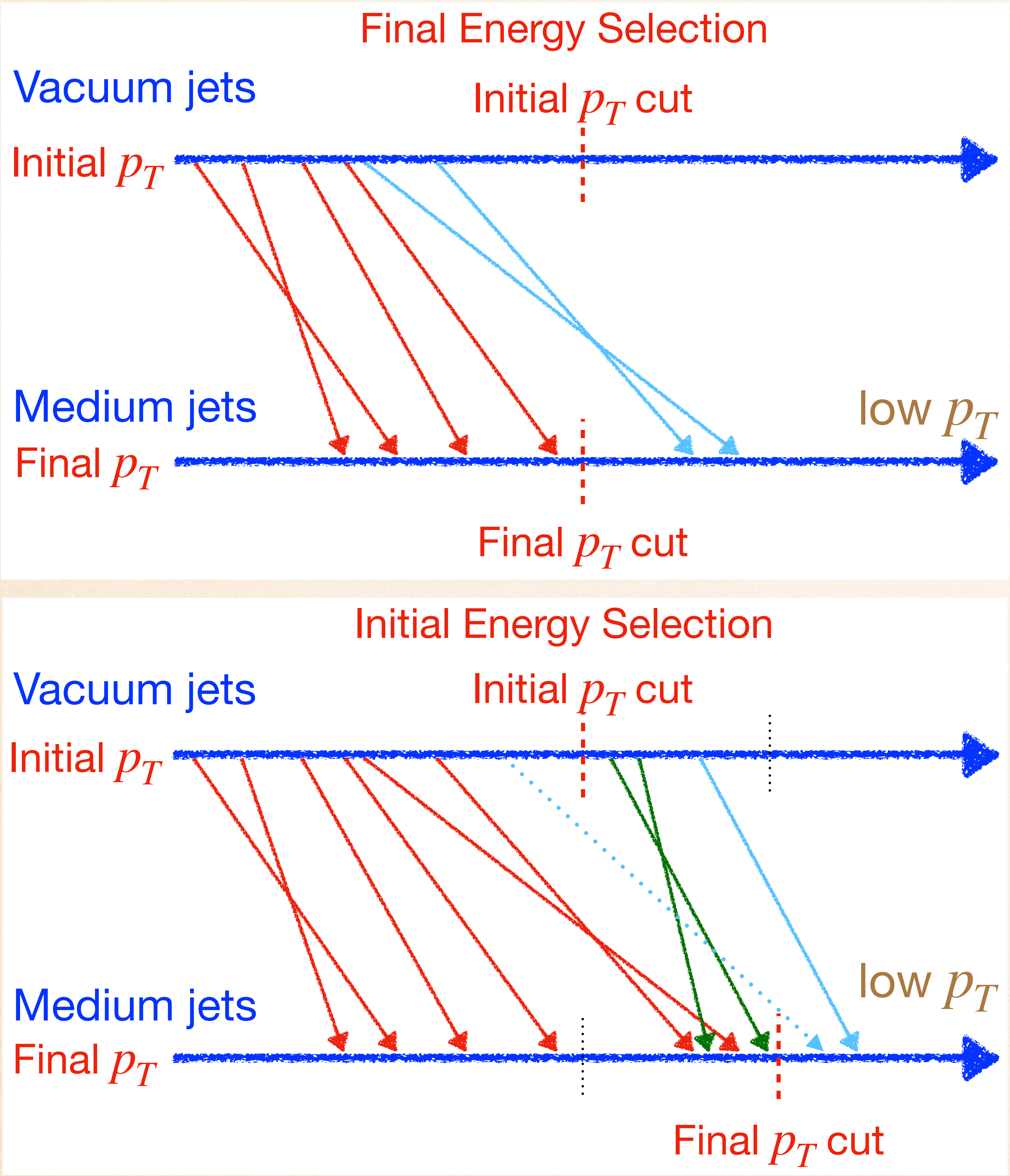
CNN w/o bkg.

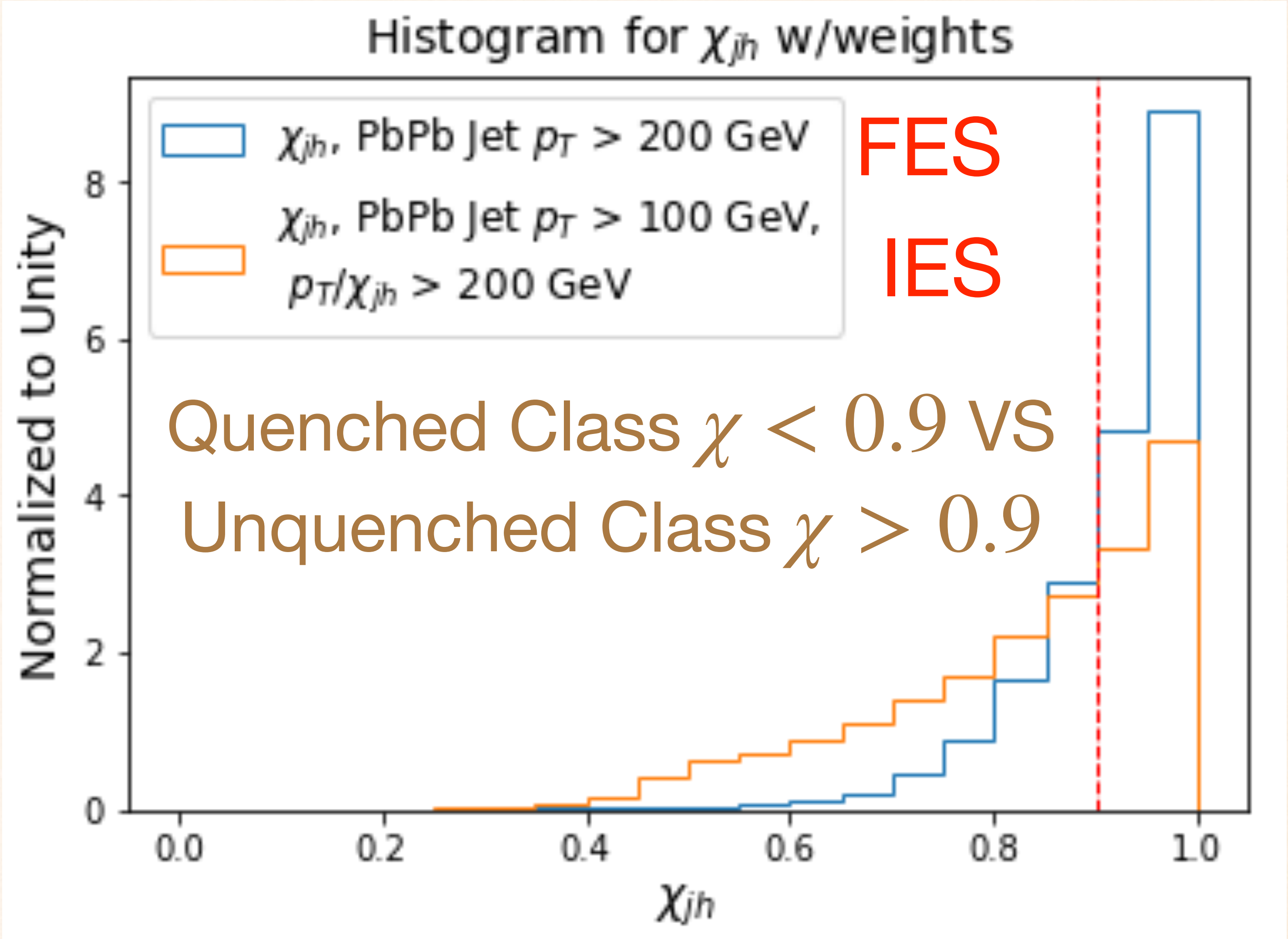
More quenched ⟷ Less quenched

- DGCNN takes point cloud representation of jets as input, keeping the full jet information
- DGCNN achieves the best performance in all scenarios

# Applications: Initial VS Final Energy Selection

YLD, D. Pablos and K. Tywoniuk, JHEP03(2021)206



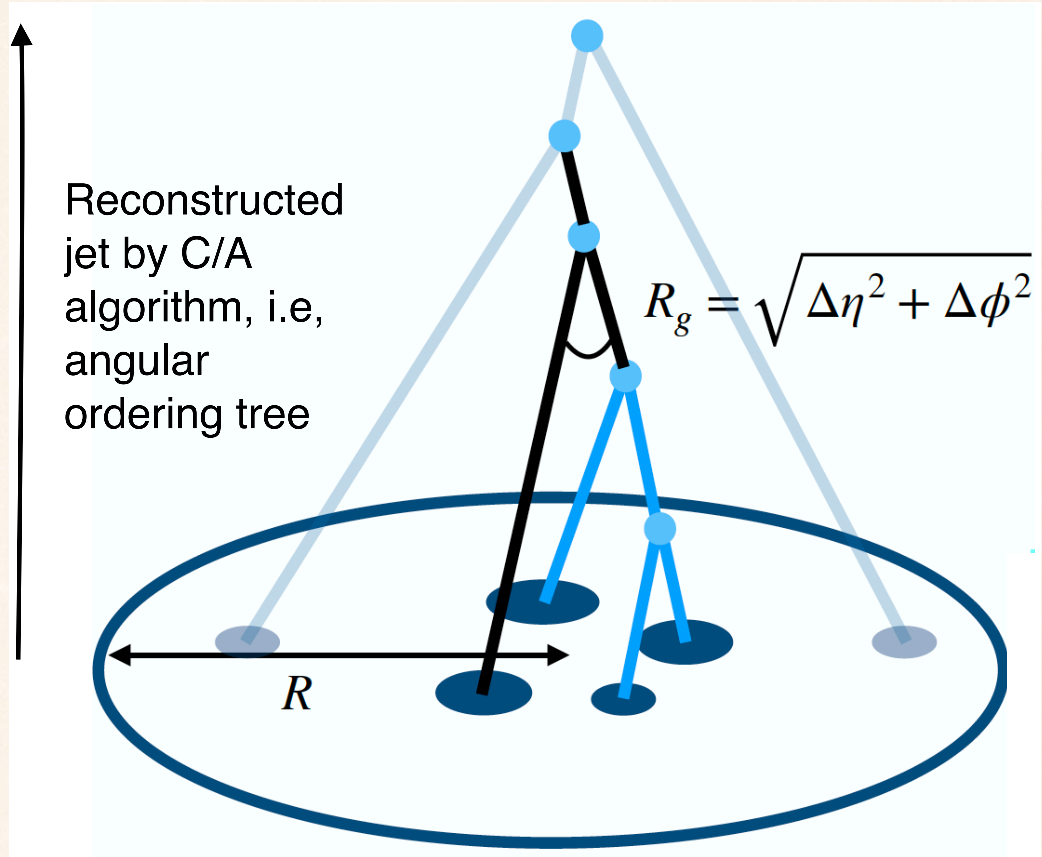

- Final Energy Selection (FES): impose $p_T$ cut on final energy $p_T$>200 GeV → Steeply falling energy loss $\chi$ distribution, dominated by little quenched samples!
- Initial Energy Selection (IES): impose $p_T$ cut on initial energy $p_T/\chi$>200 GeV & final energy $p_T$>100 GeV → More support of fairly quenched jets. More distinguishable!

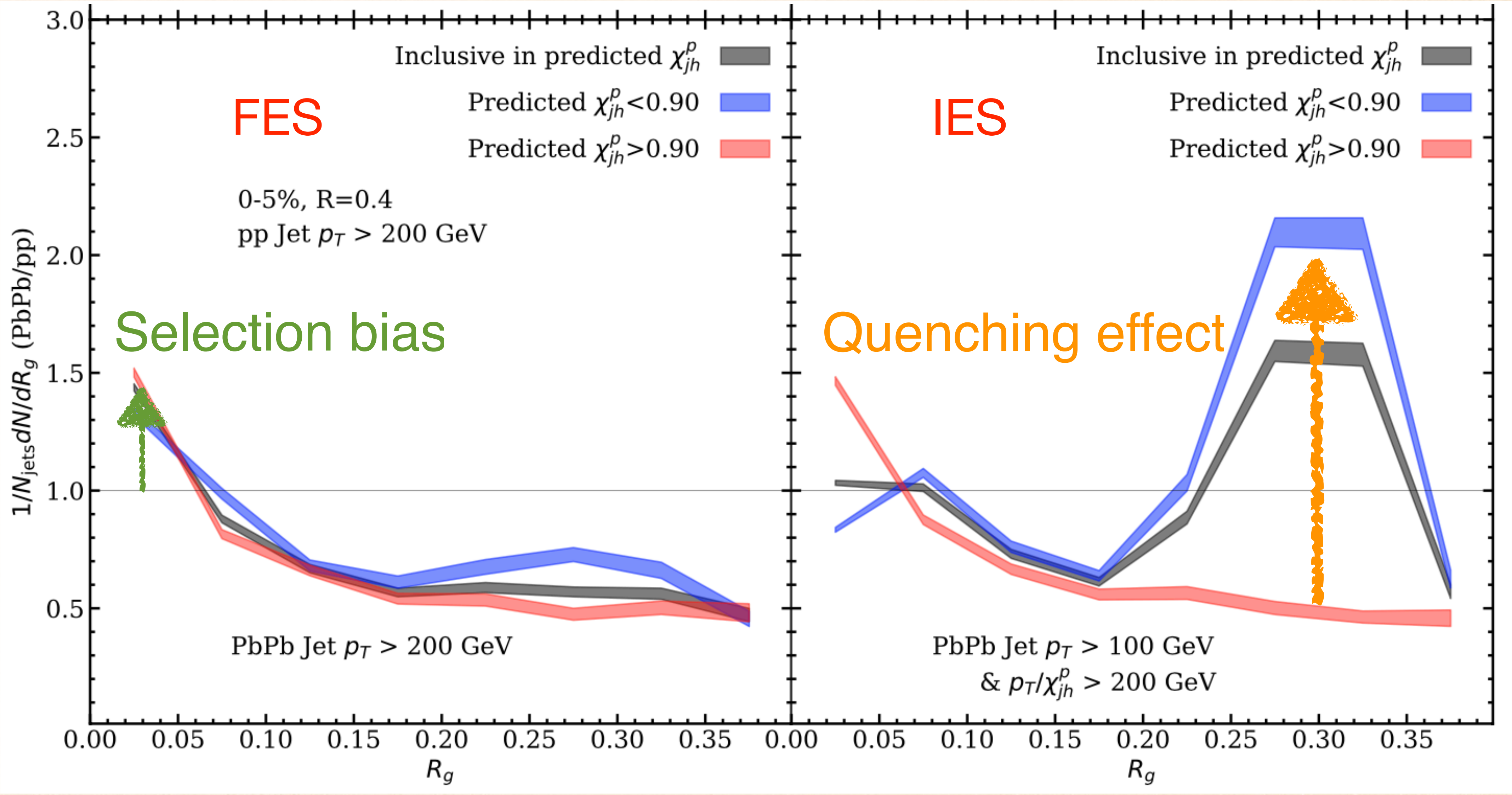
# Reveal and remove selection bias: Apple-to-apple comparison

YLD, D. Pablos and K. Tywoniuk, JHEP03(2021)206

$R_g$ ratio between PbPb jets and all pp jets

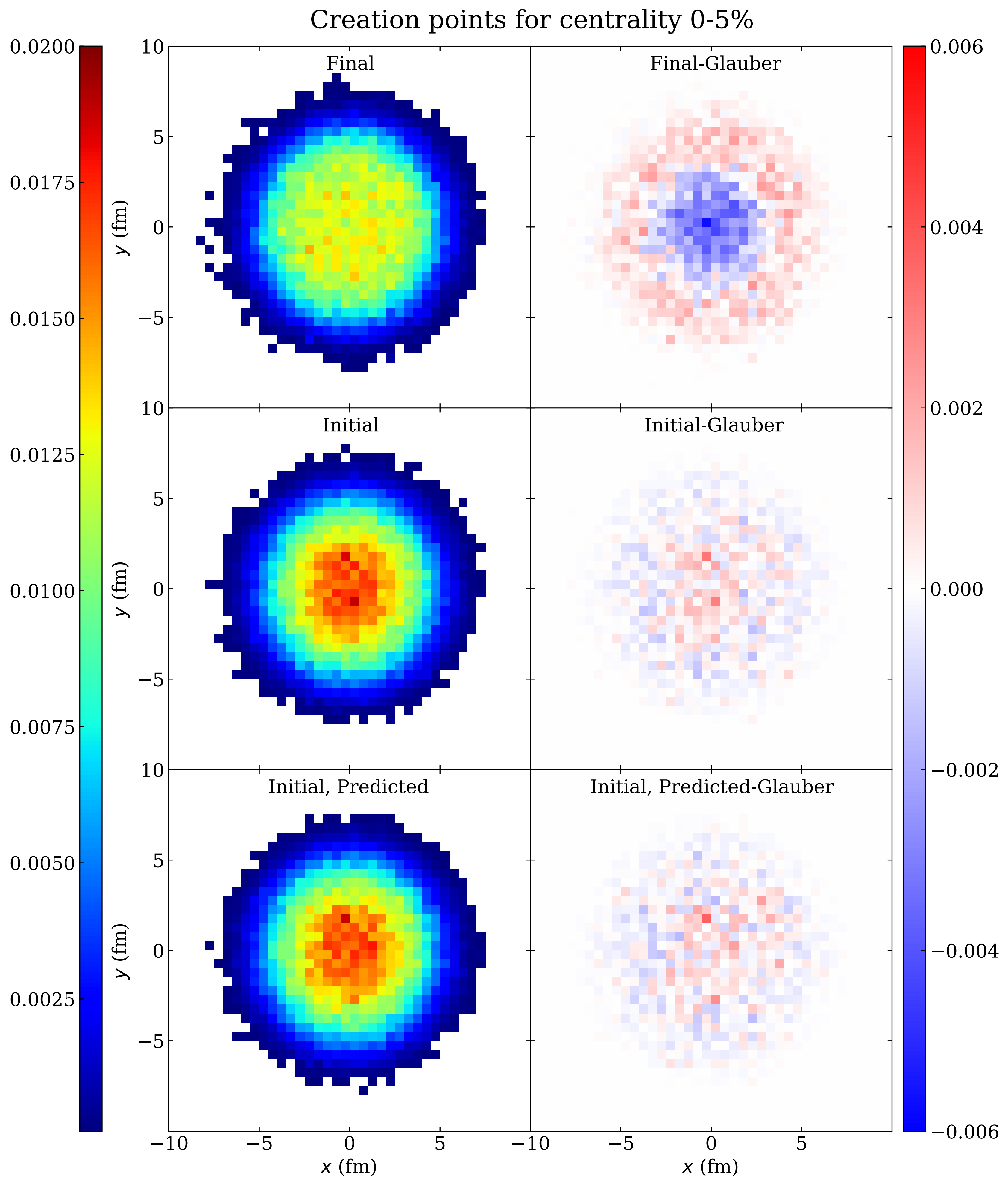
- FES: Selection bias towards jets with smaller $R_g$, originated by final energy $p_T$ cut.
- IES:
  - Unquenched class: still biased due to $\chi > 0.9$ cut: to belong to this class, a jet had better to be with smaller $R_g$, compared with all pp jets.
  - Quenched class presents features related to energy loss, compared with unquenched class: jet quenching leads to enhancement of large $R_g$- creation of a new, semi-hard branch at large angle from medium response.

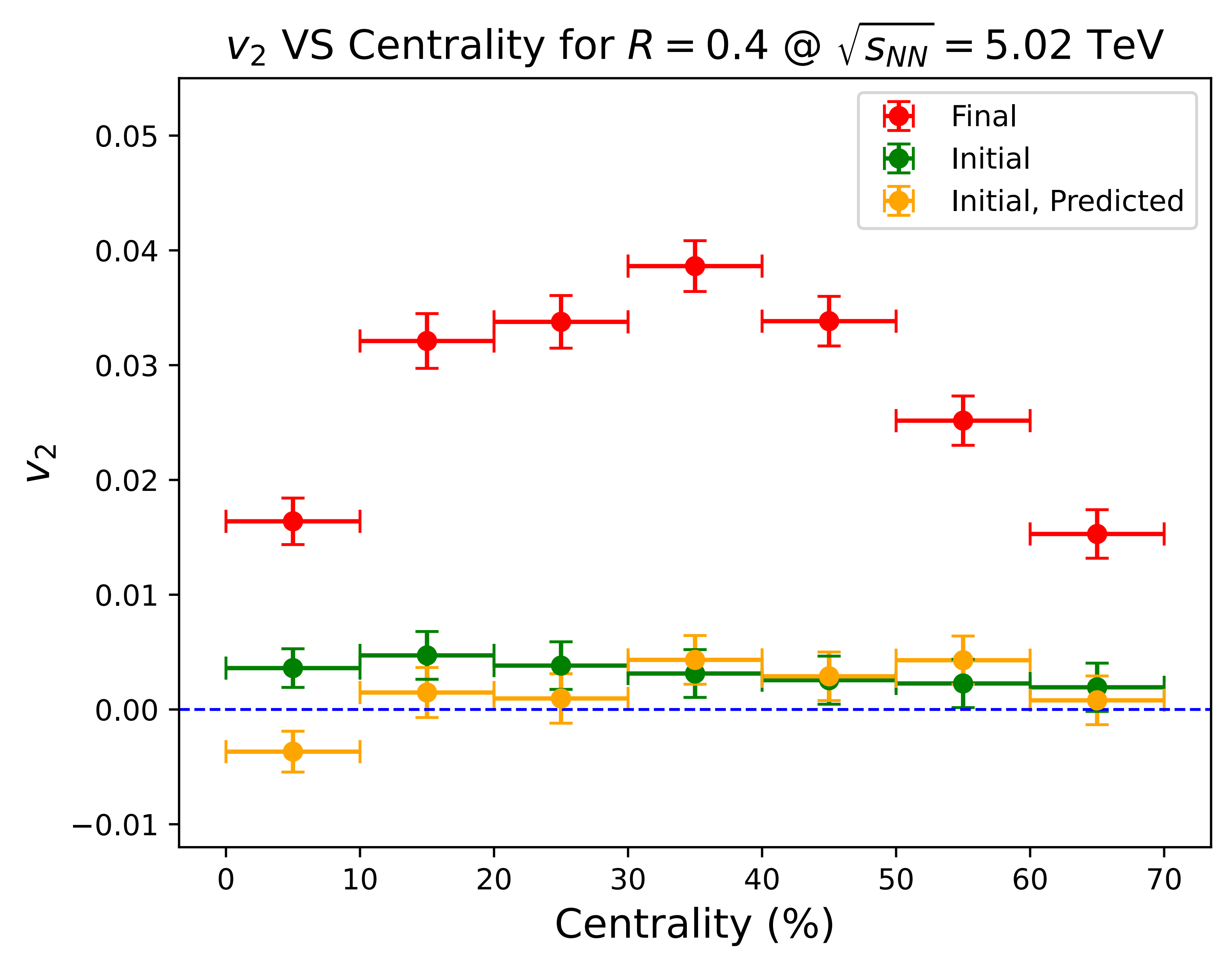


# Genuine Configuration Profile

YLD, D. Pablos and K. Tywoniuk, Phys. Rev. Lett. 128, 012301 (2022)



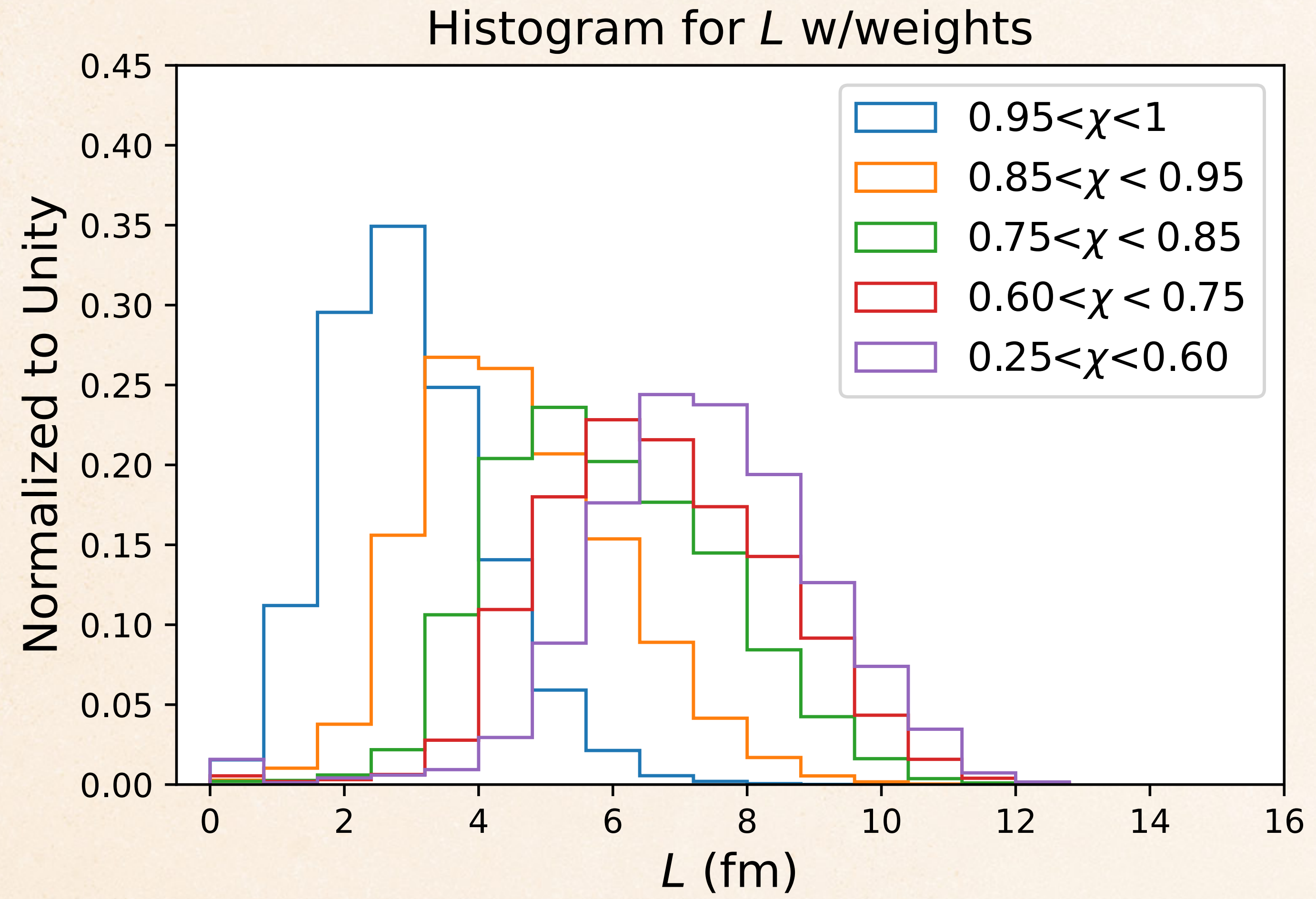



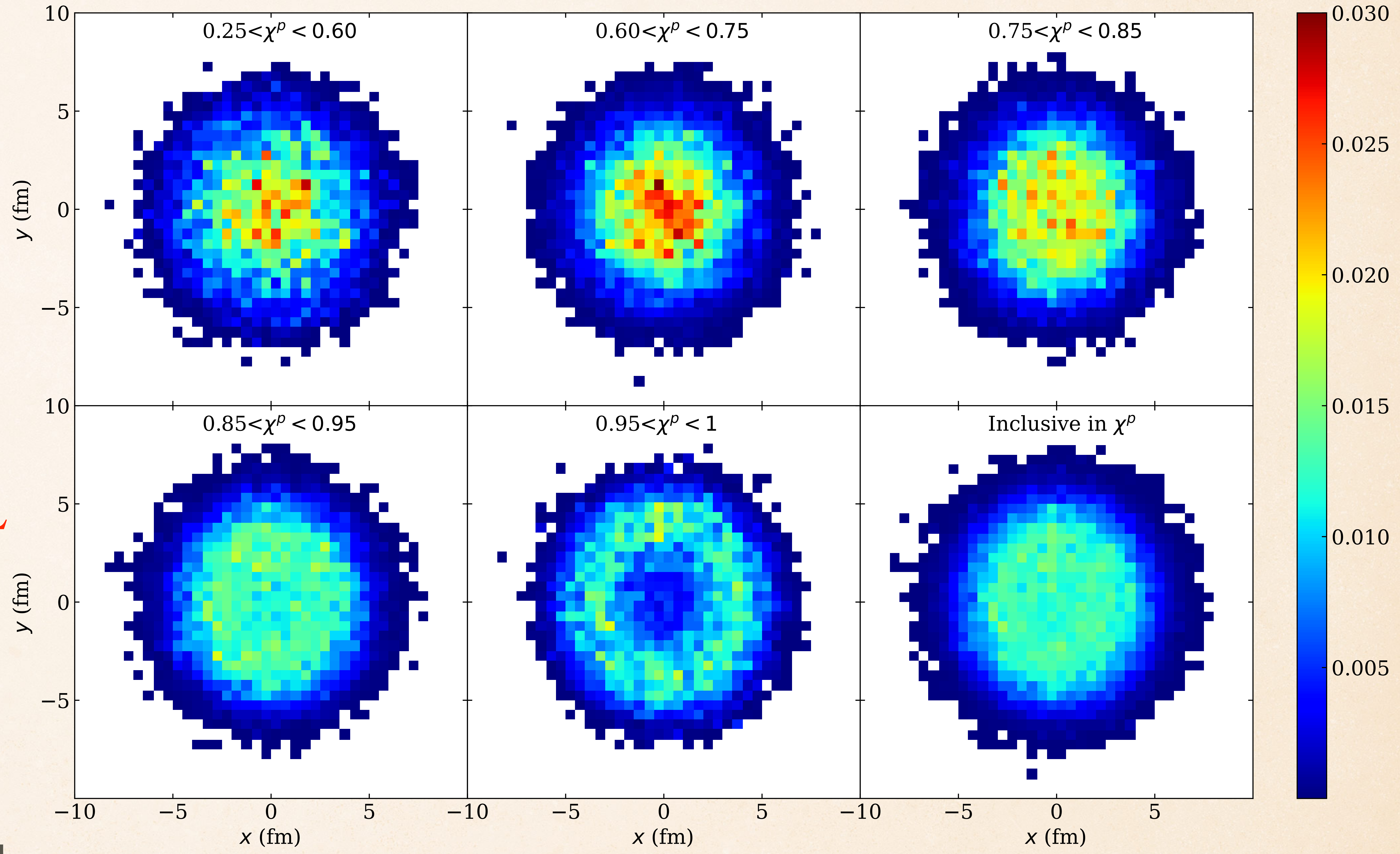

- Initial Energy Selection (IES) "removes" final state interactions (selection bias), since we record "all" jets

- IES provides access to the genuine jet creation point (path length) distribution and possible initial-state jet anisotropy

# Towards jet tomography

YLD, D. Pablos and K. Tywoniuk, JHEP03(2021)206



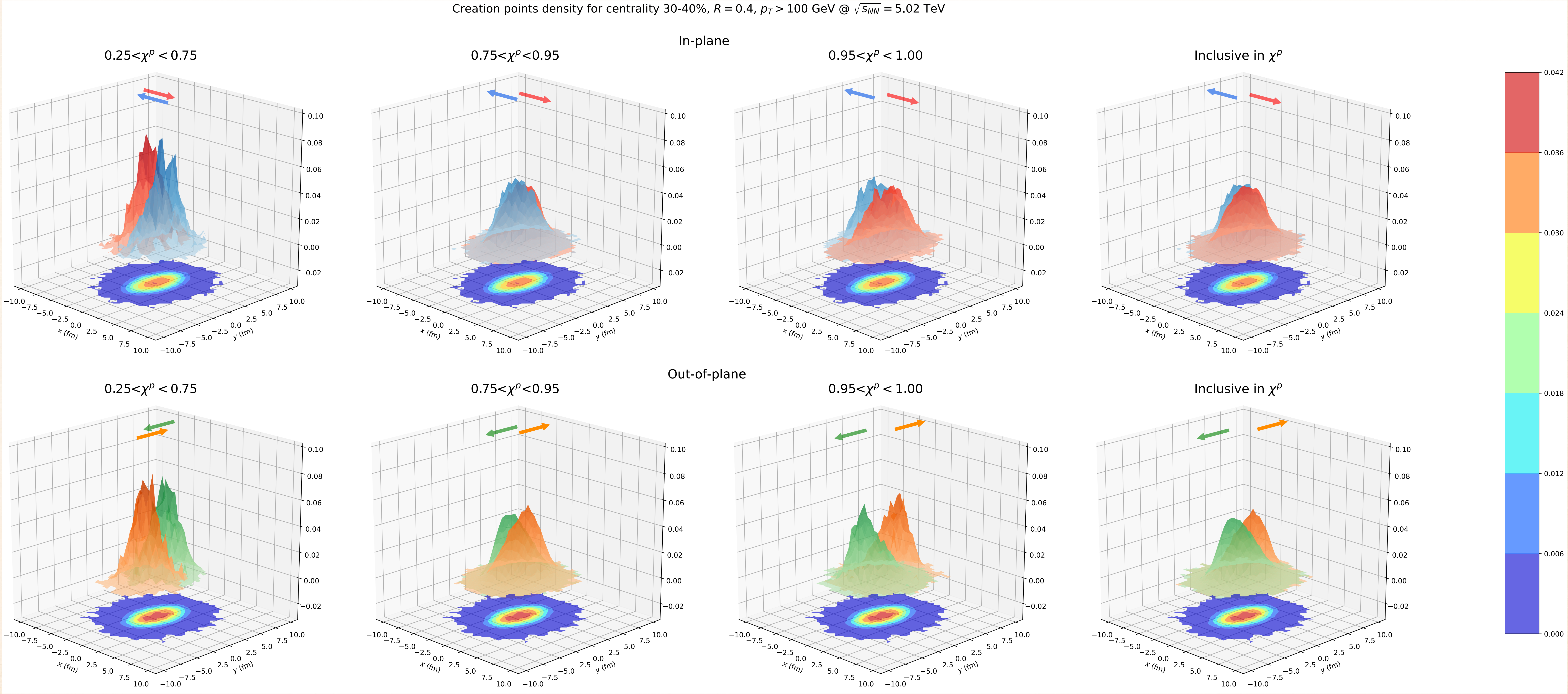

- Strong correlation between jet traversed length $L$ and energy loss $\chi$
- Selecting jets with different $\chi$ will naturally select jets that traversed different $L$
- Great potential to make tomographic application!

# Towards jet tomography

YLD, D. Pablos and K. Tywoniuk, Phys. Rev. Lett. 128, 012301 (2022)

In-plane jets
($v_2 > 0$) going left ($p_x < 0$)
and right ($p_x > 0$)

Out-of-plane jets
($v_2 < 0$) going upper ($p_y > 0$)
and down ($p_y < 0$)




- To get very quenched, jets have to travel longer in medium. So $v_2$ & $p_{x,y}$ are helpful for jet tomography.

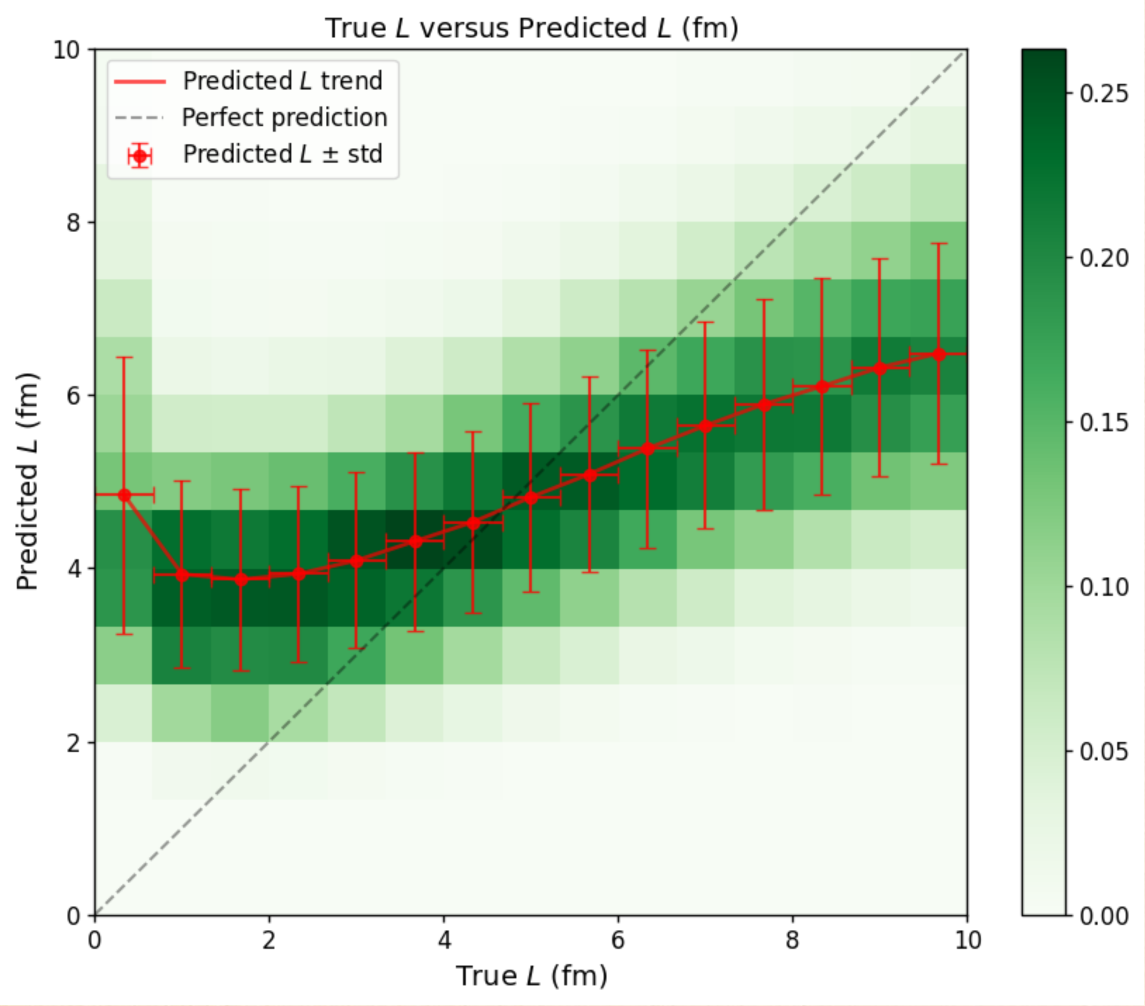
# Summary & Outlook

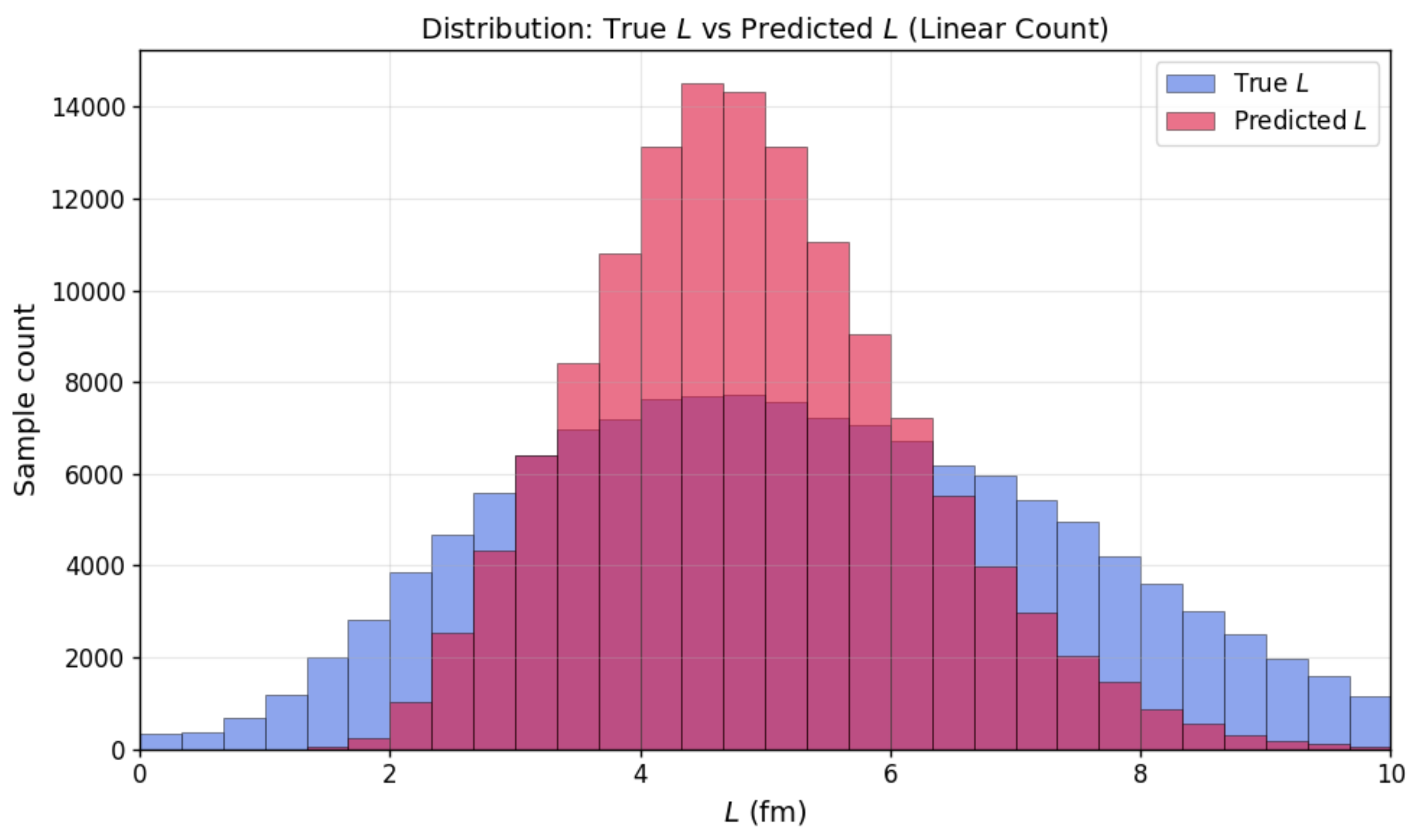
- To better reconstruct jet momentum with machine leaning, training ML models with various realistic quenched jets simultaneously are recommended to obtain strong robustness and generalizability.
- DGCNN, *taking full jet information*, can obtain good prediction performance on jet energy loss in the presence of QGP background.
- With energy loss ratio: Mitigate selection bias and reveal medium effects on various jet observables.

- Develop novel ML models, e.g., transformer, to deal with jet & background particles directly to improve the prediction accuracy and avoid the procedure of background removal.
- Reduce model dependence: Incorporate quenched jets from various MC models.
- Include the realistic anisotropic QGP background for better generalizability in experiments.

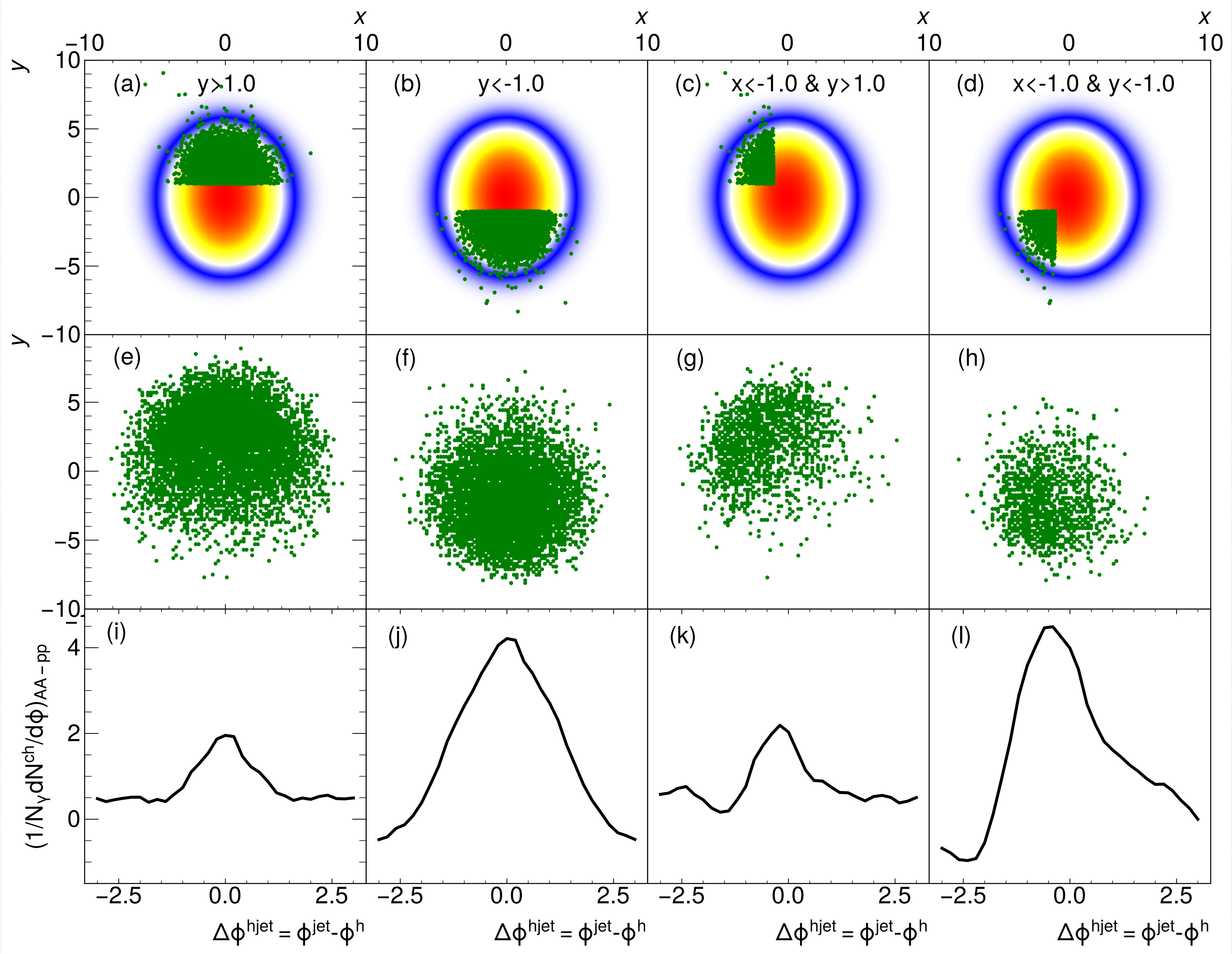
# How to deal with imperfect regression?

Z. Yang, Y. He, W. Chen, *et al.*, Eur. Phys. J. C 83, 652 (2023)




Jet in-medium traversed length $L$




Network selection

Actual distribution

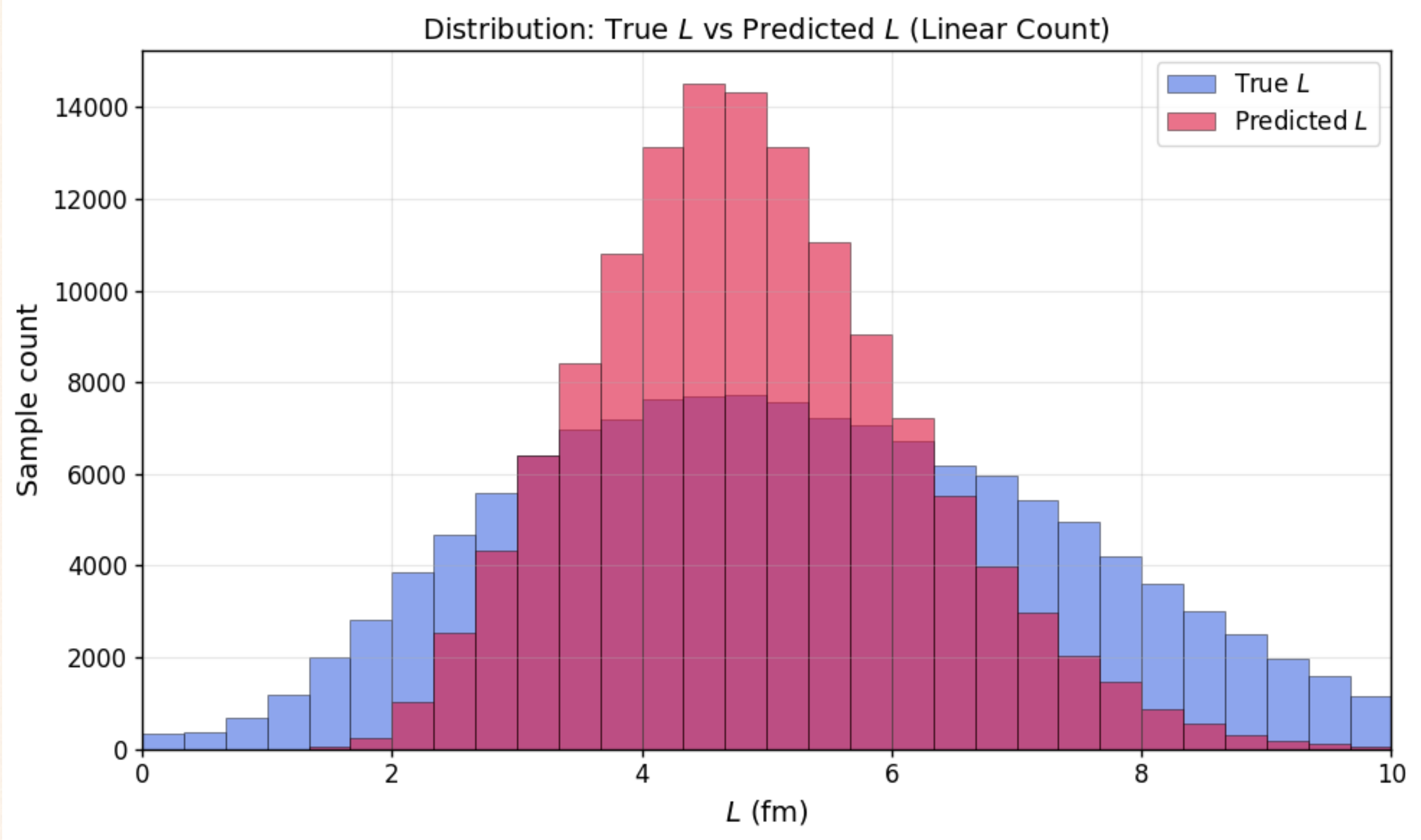
Jet Production Positions

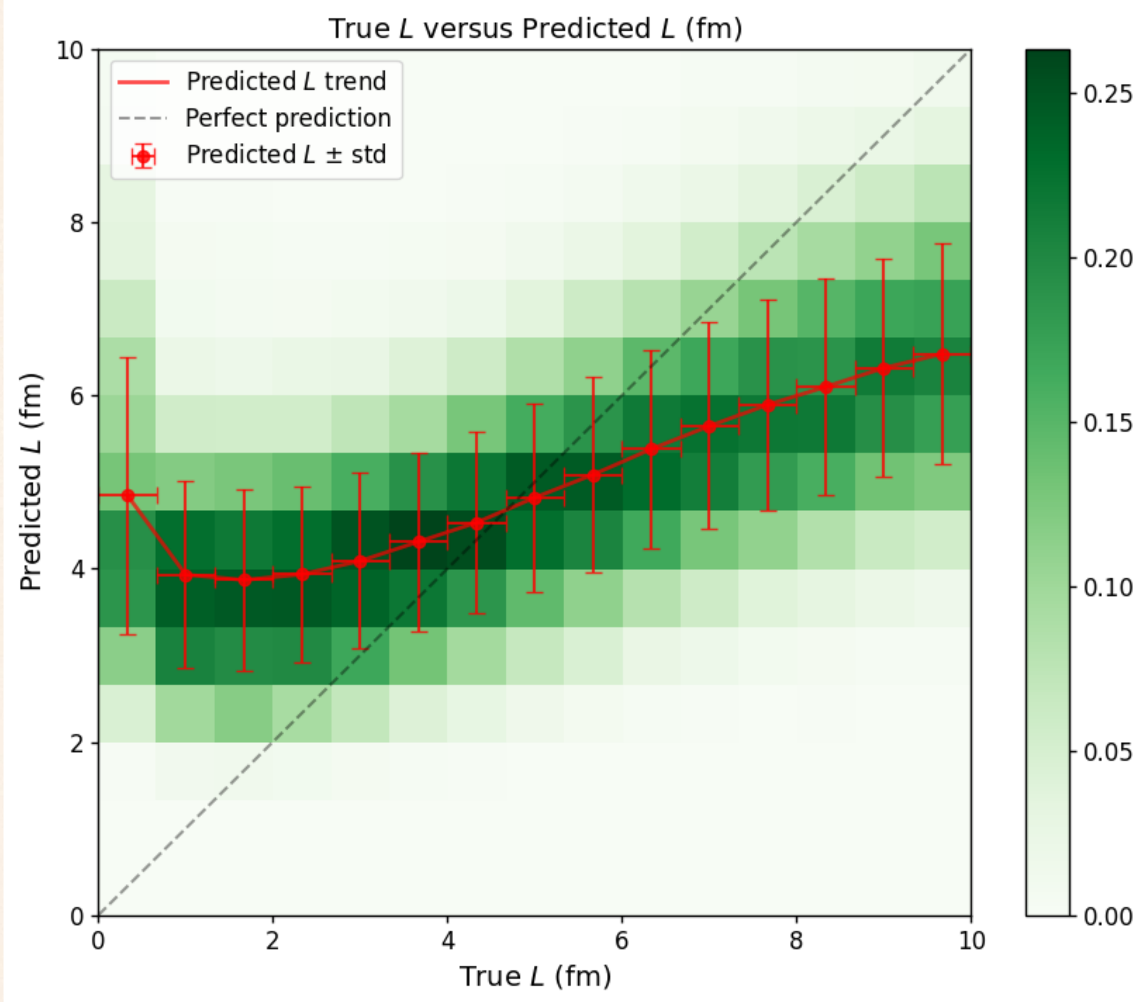
# New training strategy: also required to follow the distribution

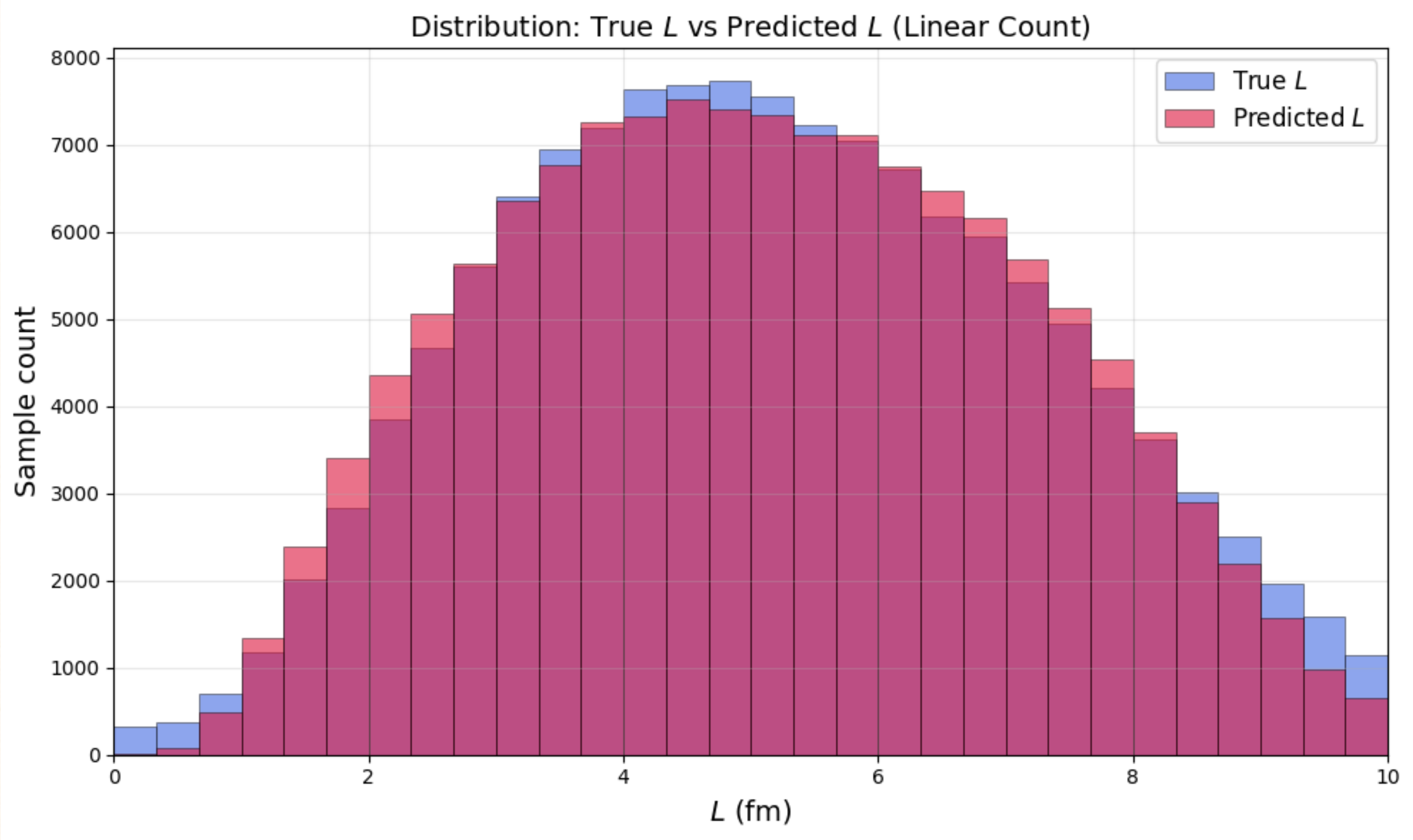
Jet in-medium traversed length $L$

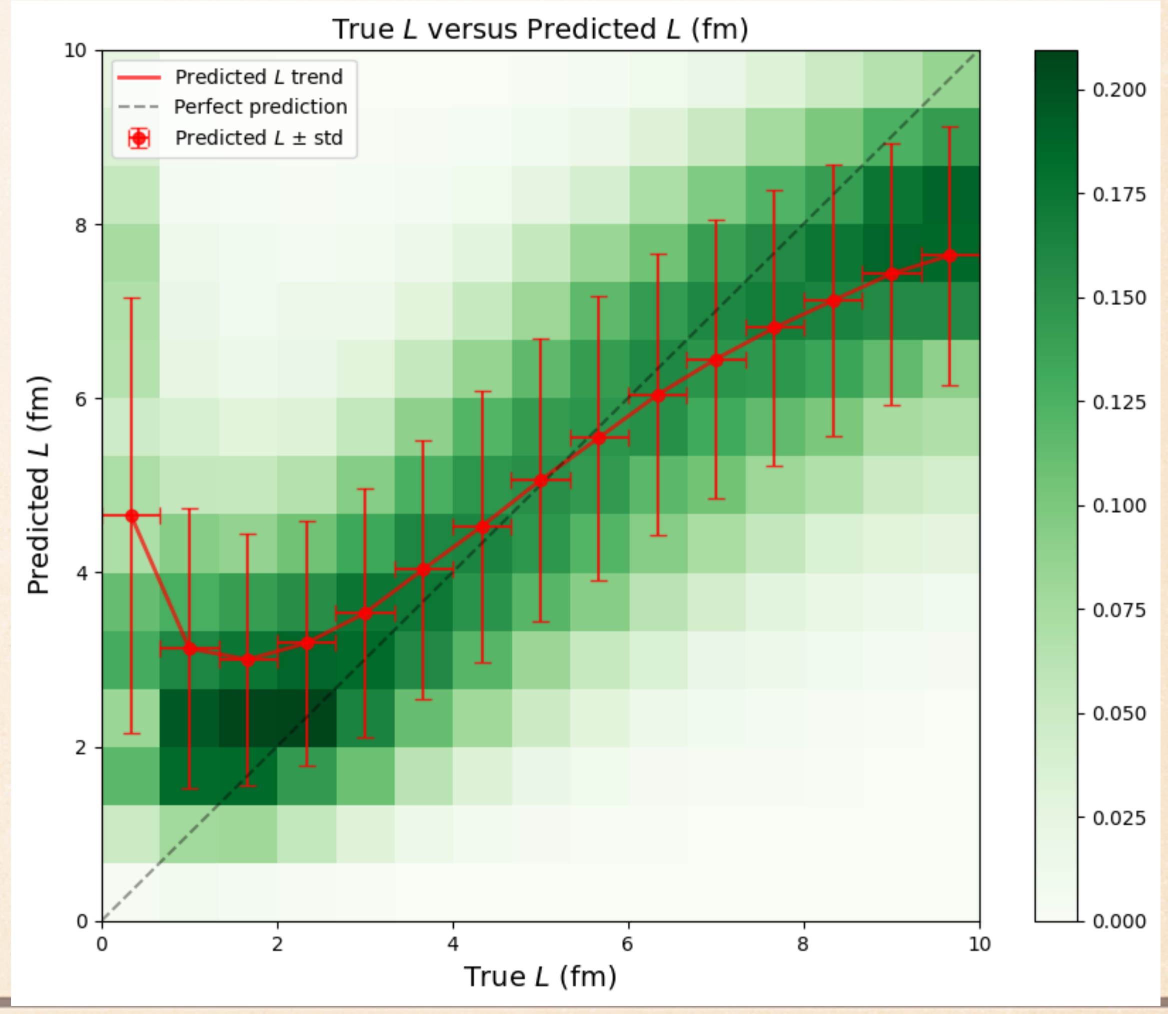
$$\mathscr{L} = \underbrace{\frac{1}{N}\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2}_{\text{MSE loss}}$$

$$\mathscr{L} = \underbrace{\frac{1}{N}\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2}_{\text{MSE loss}} + \lambda\ \underbrace{D\left(P(\hat{y}), P(y)\right)}_{\text{dist. distance}}$$



# Dynamical Graph Convolutional Neural Network

Huilin Qu, Loukas Gouskos, Phys. Rev. D 101, 056019