Project is supported in part by DOE grant Grant No. DE-FG05-92ER40712

Financial support from The Gordon and Betty Moore Foundation and the American Physical Society to present this work at the GHP 2023 workshop

* *Paper available on: https://arxiv.org/abs/2206.01628*

Author: Lihan Liu, Julia Velkovska, Marta Verweij, Yilun Wu

Identifying Quenching Effect in Heavy-ion Collisions with Machine Learning *

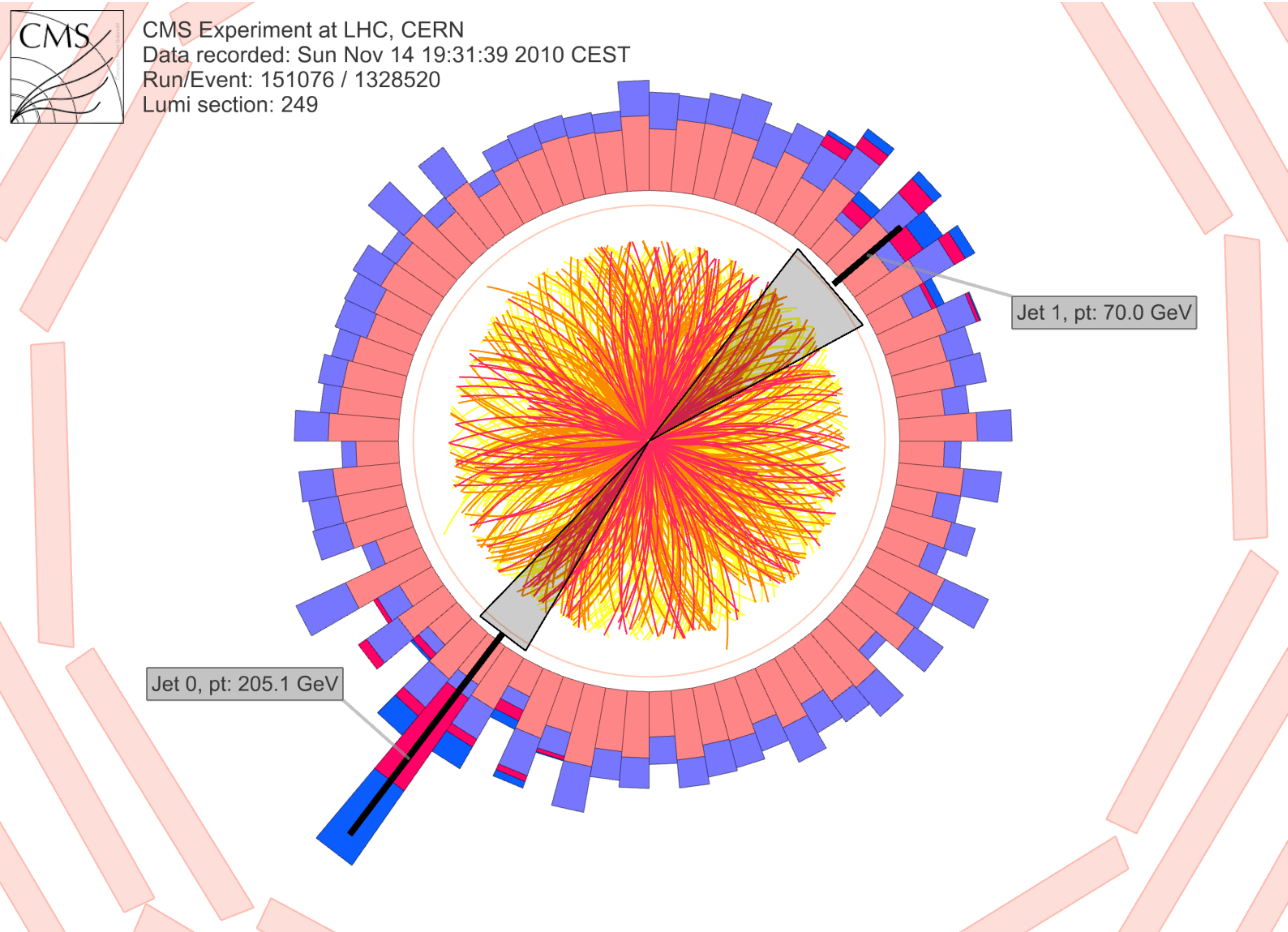
GHP 2023 WORKSHOP

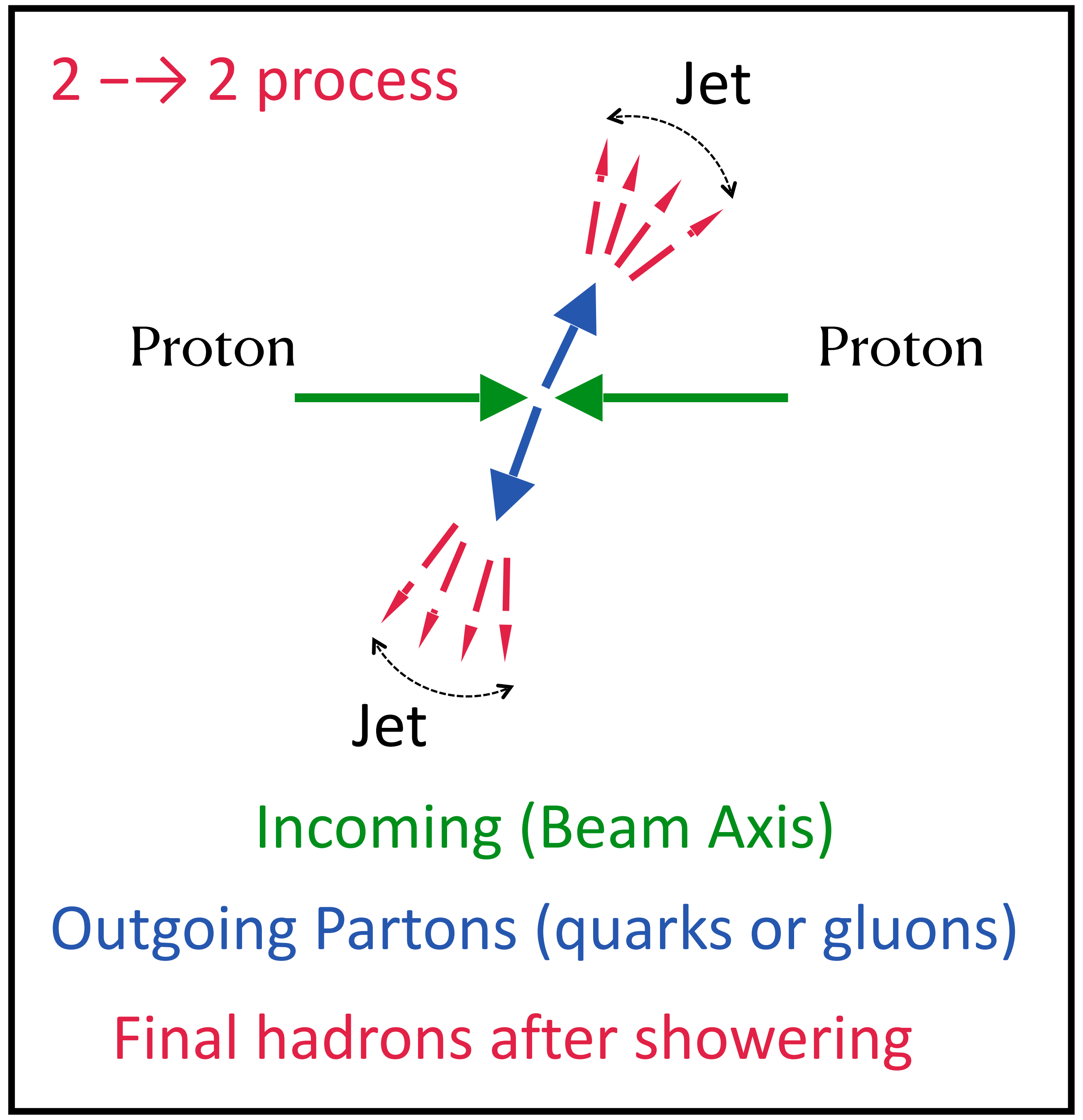
Yilun Wu

Vanderbilt University

Background: Jet Quenching Phenomenon

Heavy ion collision

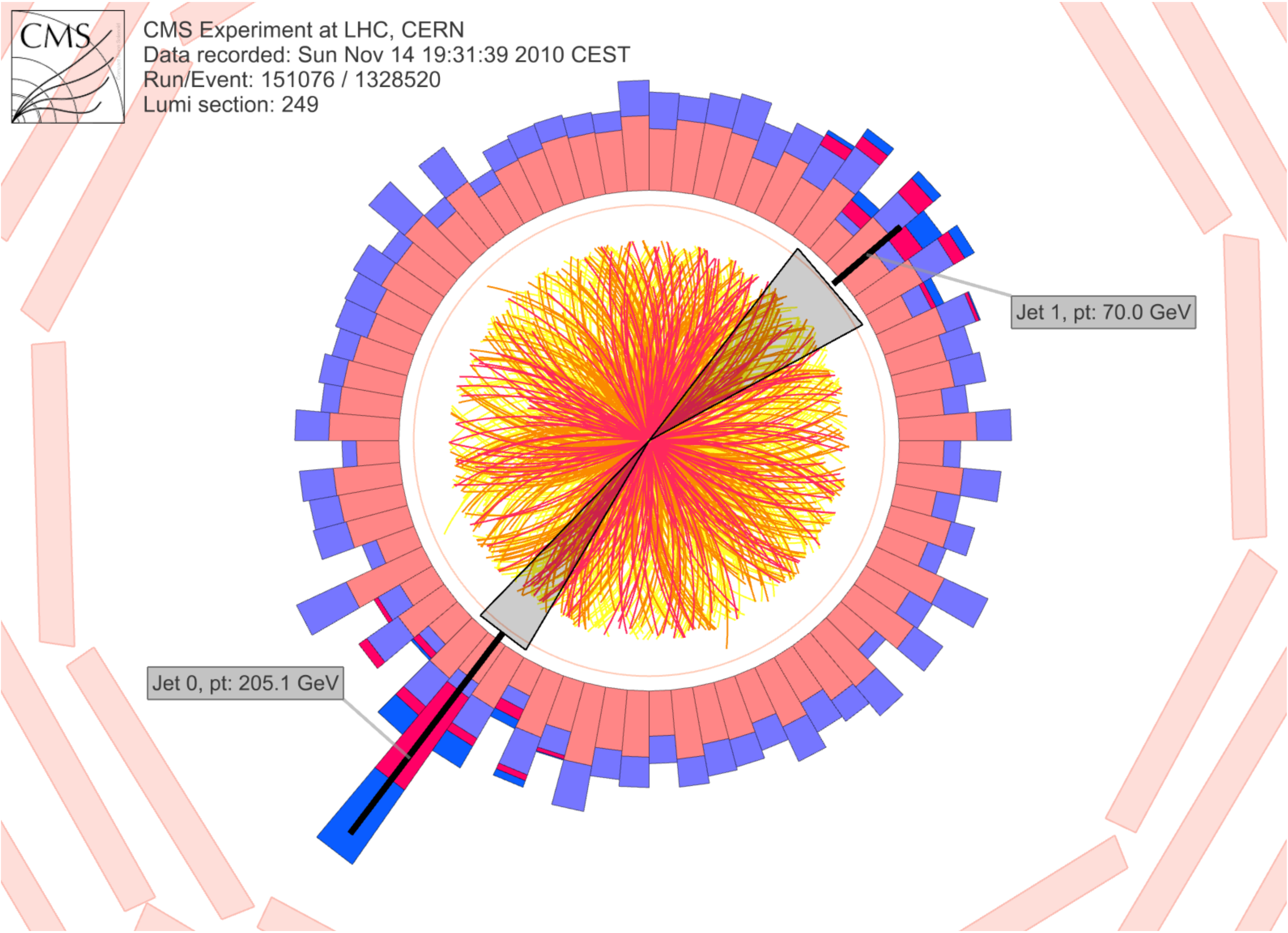




▲ Jets traveling in opposite direction with equal initial transverse momentum pT

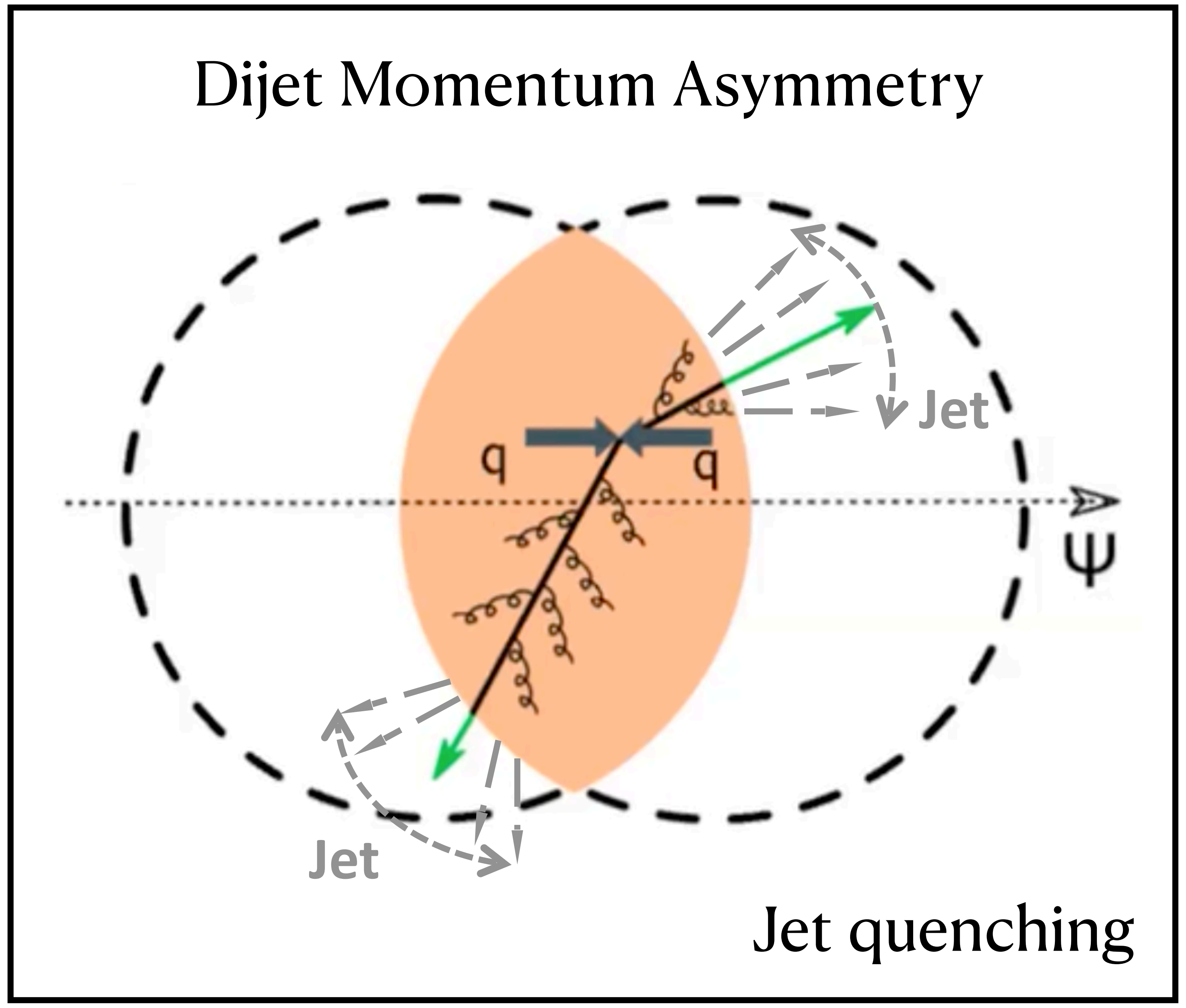
◀ Final state pT is not equal

Background: Jet Quenching Phenomenon

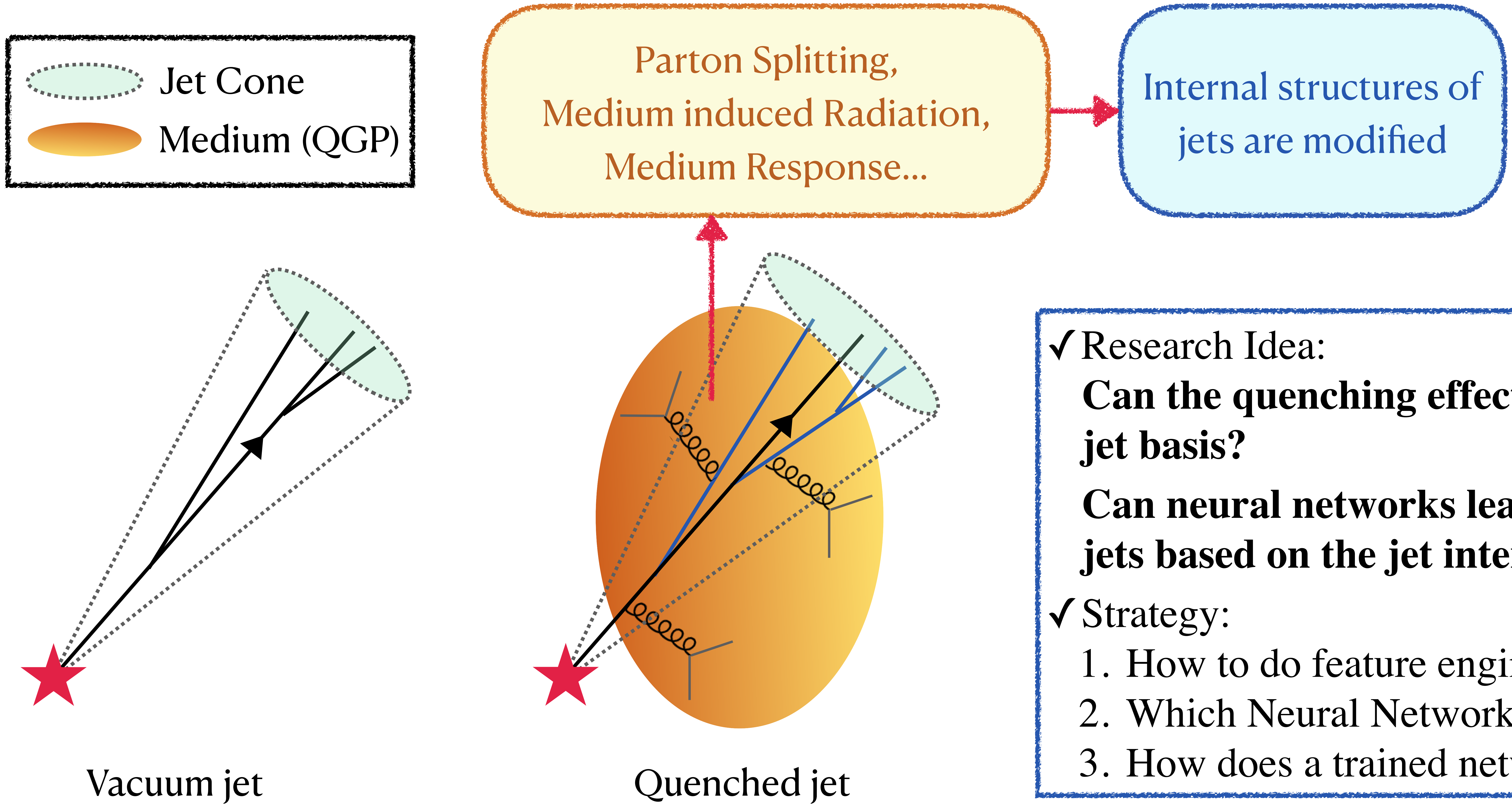


QGP signature: Jet quenching phenomenon

- Jets interact with the QGP medium and lose energy.
- Back-to-back jets traverse different path length of the QGP medium.
- **Jets are quenched(modified) in different levels when traversing the QGP.**



Motivation



Jet Substructures

Plenary talk "Jet Substructure and its utility in small and large systems"
—-Raghav Kunnawalkam Elayavalli

✓ Research Idea:

Can the quenching effect be studied on a jet-by-jet basis?

Can neural networks learn to identify quenched jets based on the jet internal structures?

✓ Strategy:

1. How to do feature engineering?
2. Which Neural Network? How to do training?
3. How does a trained network behave?

How to do feature engineering?

It is necessary to introduce the thermal background effect to the feature engineering.

JEWEL simulation for dijet events:

Non-quenched jets (vacuum class)

Quenched jets (medium class)

Embedding the simulated event with a thermal background:

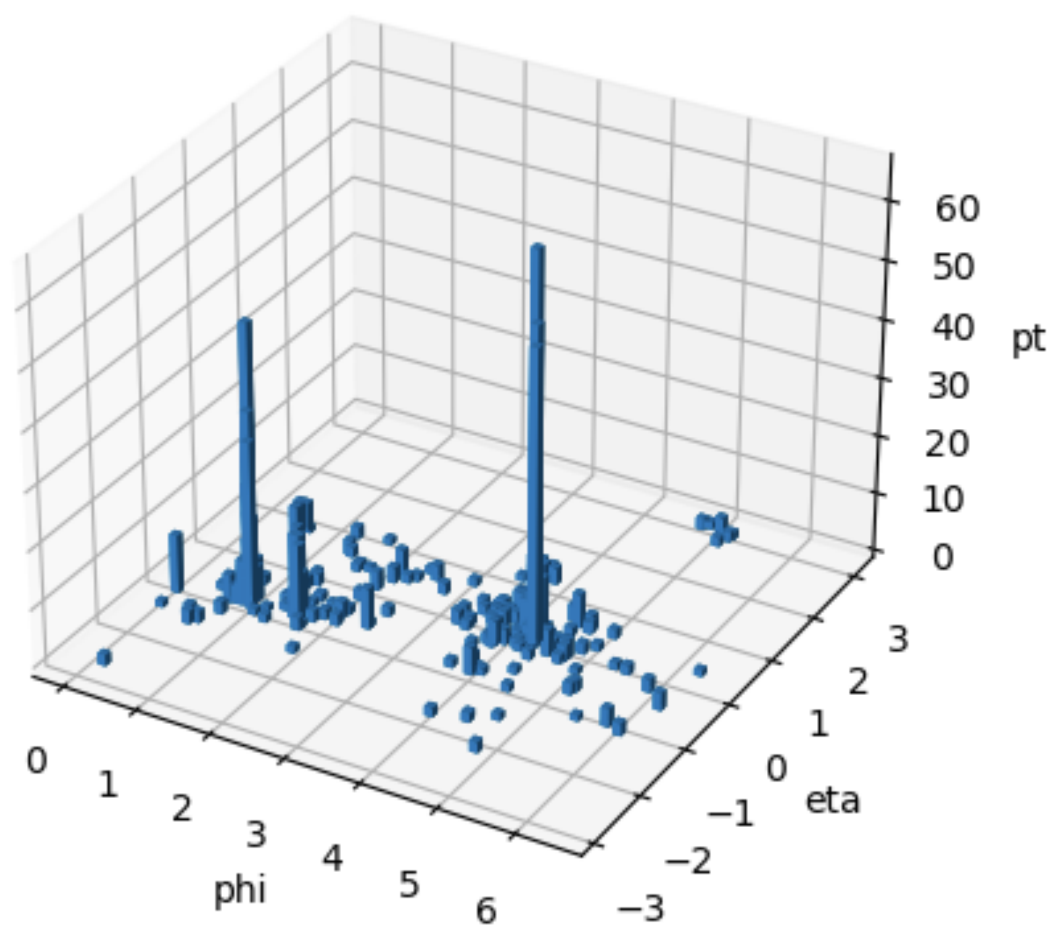
Also presents in the experiment

Background subtraction algorithm:

Event-wide Constituent Subtraction

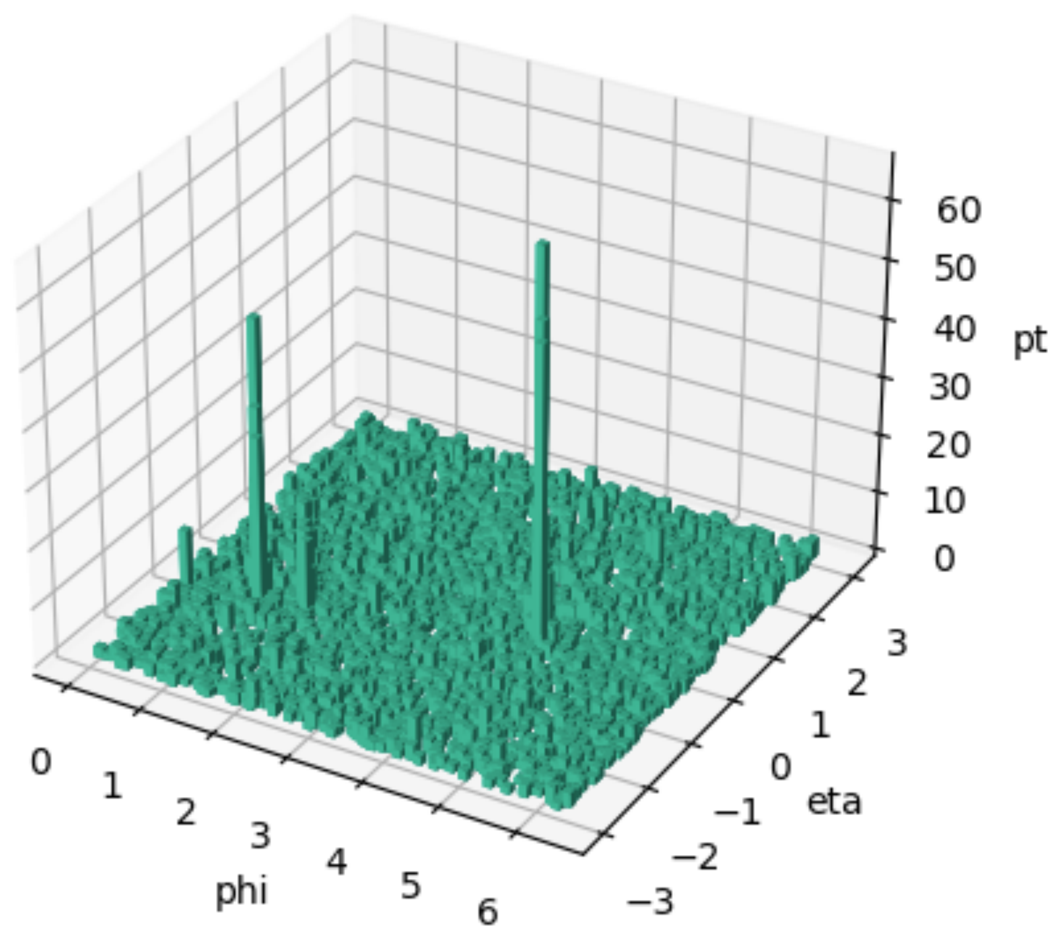
We use the jets reconstructed from the bkg-subtracted events for next step

0-10% Centrality

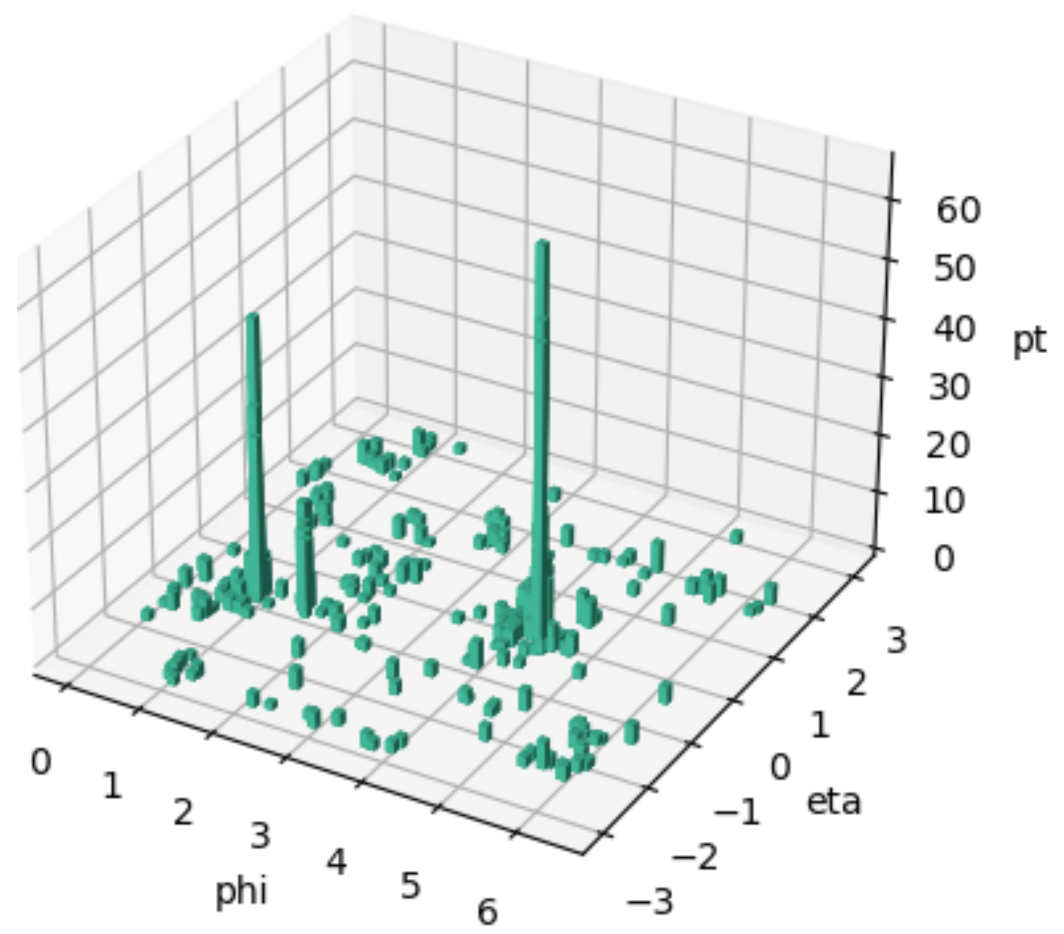


dijet hard event

+ Most central background



mixed event



bkg-sub event

How to do feature engineering?

Jet observable that represents the <u>internal structure</u> of a jet:

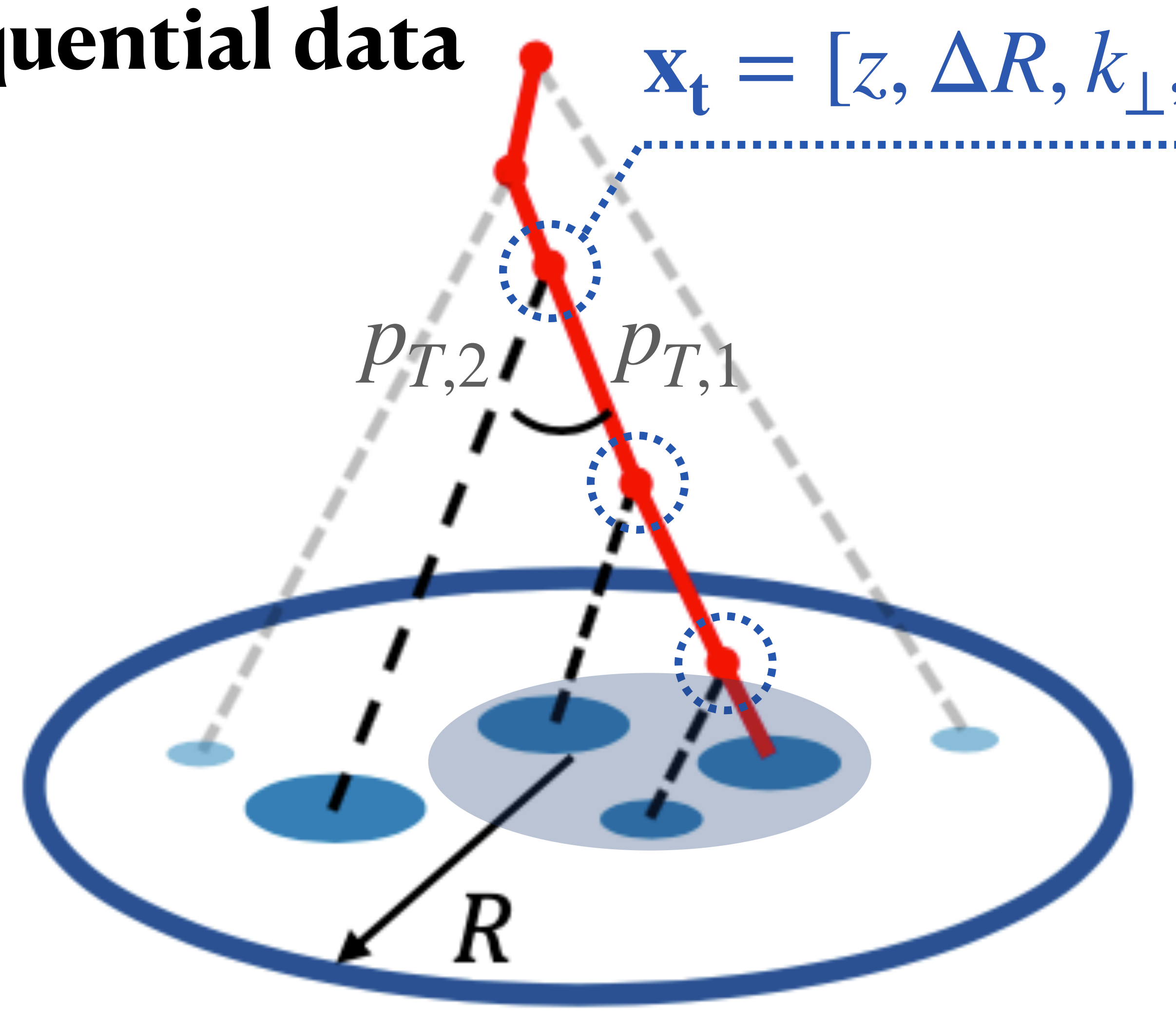
- **Jet substructure**

Input

Long Short-Term Memory Neural Network

- learning from sequential data
- Improved RNN (Recurrent Neural Network)

Sequential data $\mathbf{x_t} = [z, \Delta R, k_\perp, m, \ldots]$



Jet substructures

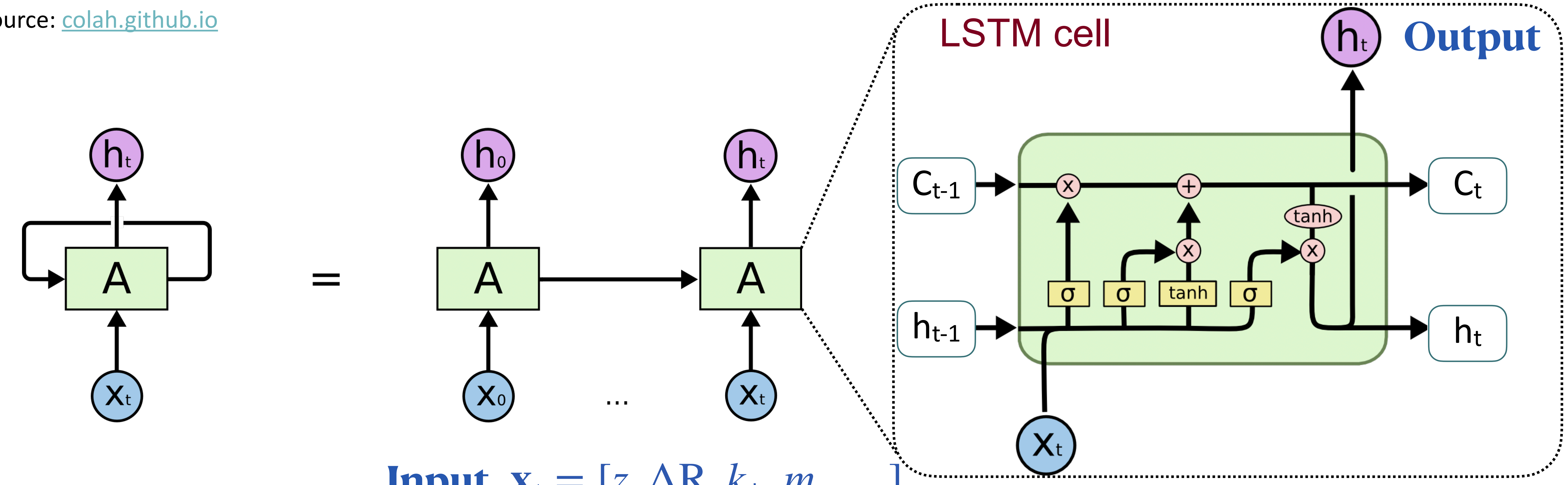
$$z = \frac{\min(p_{T,1},\ p_{T,2})}{p_{T,1} + p_{T,2}}$$ Shared momentum ratio

$$\Delta R = \sqrt{(\varphi_1 - \varphi_2)^2 + (\eta_1 - \eta_2)^2}$$ Angular separation

$$k_\perp = p_{T,2} * \Delta \mathrm{R}$$ Perpendicular momentum

$$m = inv\_mass(j_1, j_2)$$ Invariant mass

Image source: colah.github.io

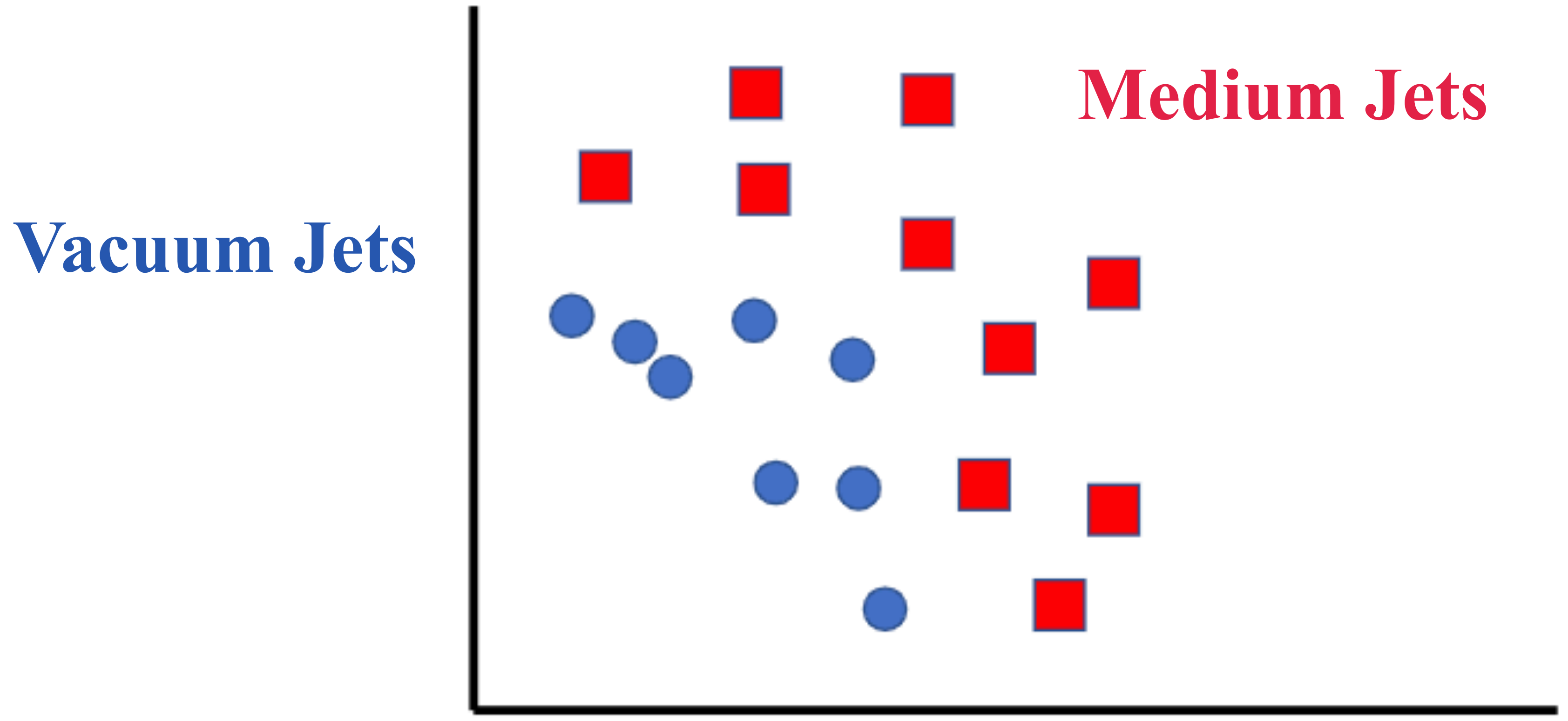


Input $\mathbf{x_t} = [z, \Delta \mathrm{R}, k_\perp, m, \ldots]$

How to do training?

Binary classification problem

Classification finds a function separating classes in a high dim space.



✓**Start from the binary classification problem**

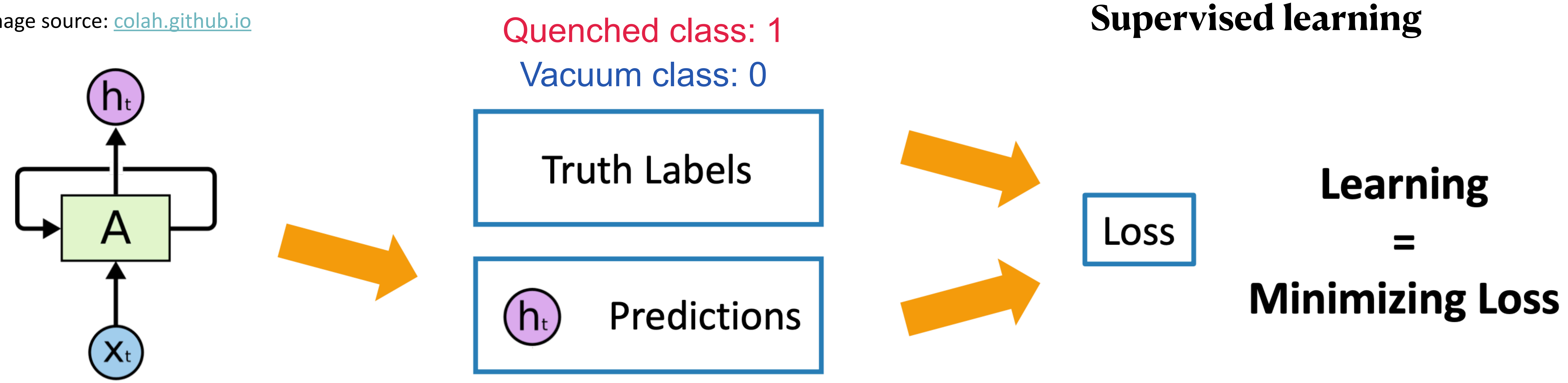
- Separating medium (quenched) jets from vacuum jets
- Four dimension space spanned from the jet substructures: $\mathbf{x_t} = [z, \Delta R, k_\perp, m, \ldots]$

✓**Supervised learning**

- Calculating the loss between truth labels and predictions
- Keeping training until the loss is minimized

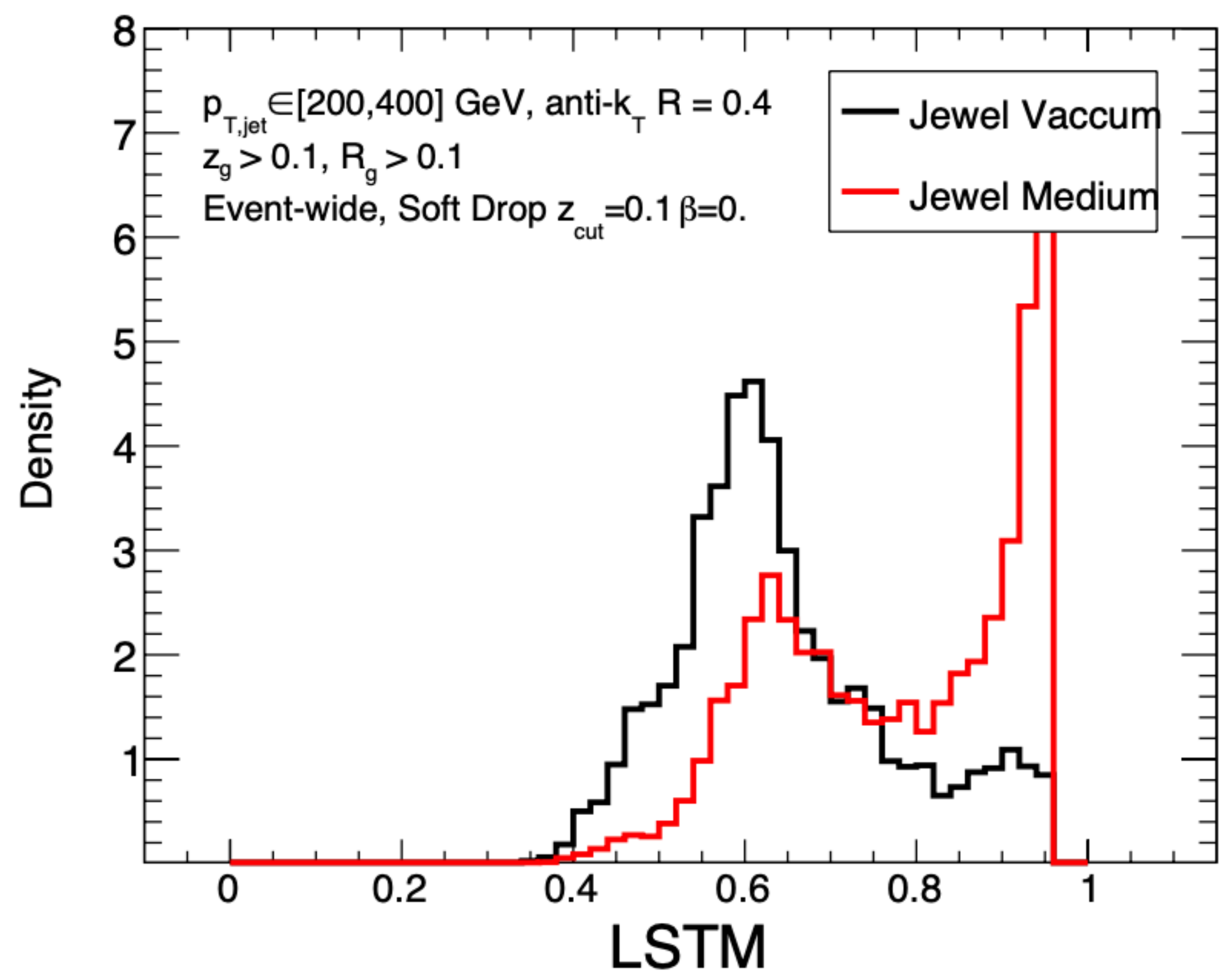
Supervised learning

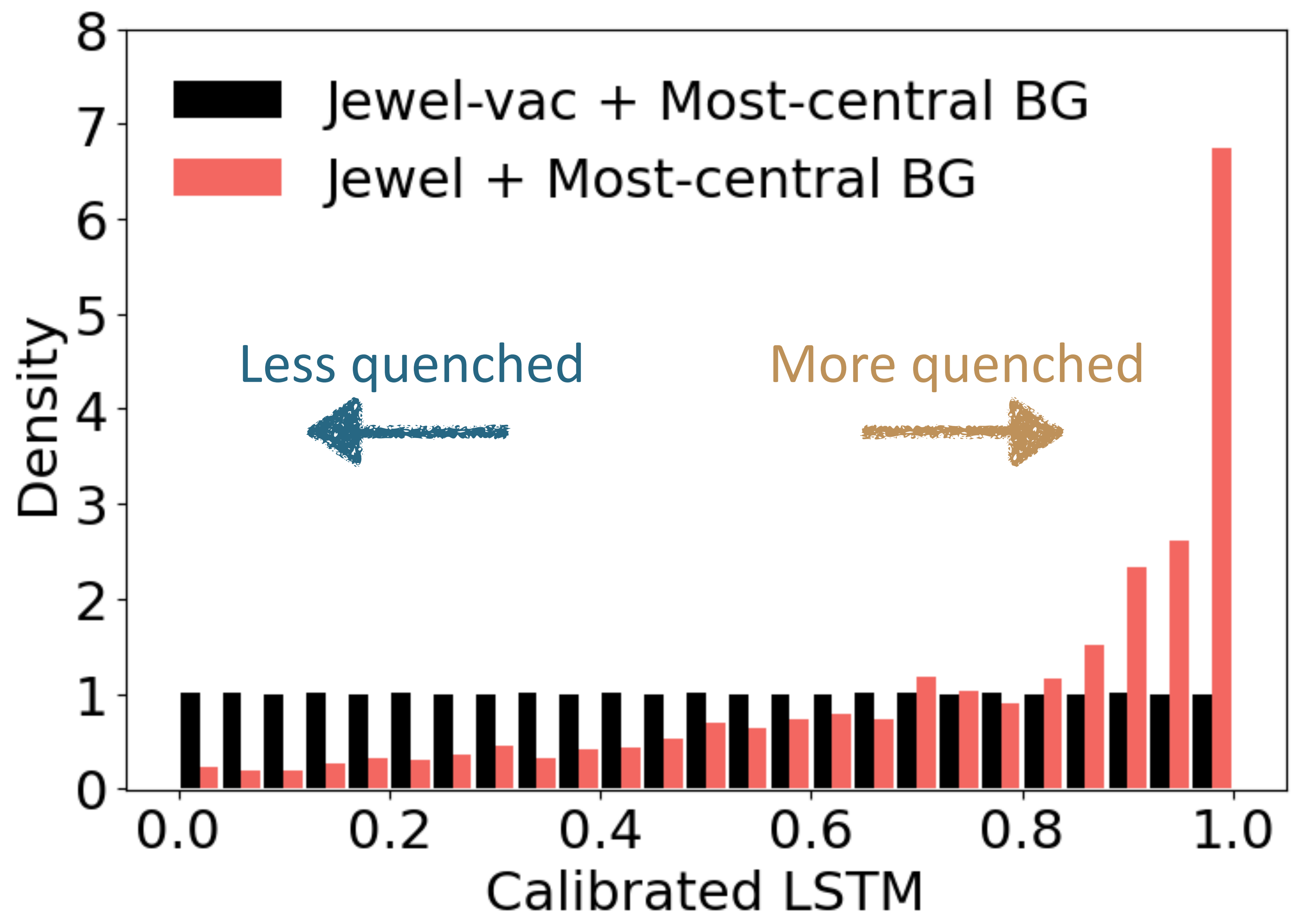
Image source: colah.github.io



How does a trained network behave?

Binary class labeling: Jewel(PbPb) jets: 1; Jewel-vac(pp) jets: 0
Histogram: Distributions of discriminators (predictions from a trained neural network).





Medium jets is separated from vacuum jets. But there are similarities between the two classes.

Calibration process: using the whole vacuum jets as reference, the quenching amount of each medium jet is determined—"Quenchness"

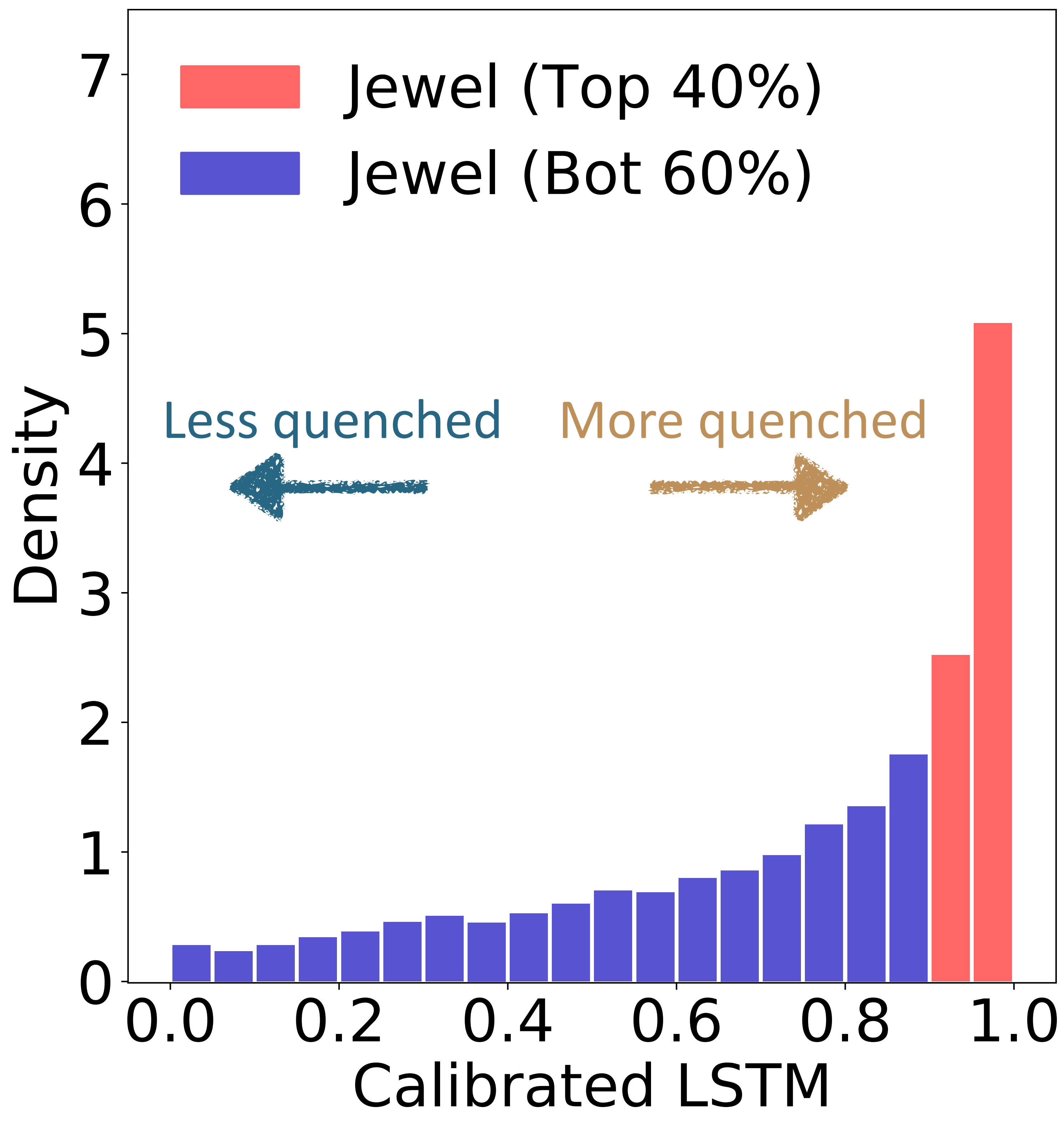
How does a trained network behave?

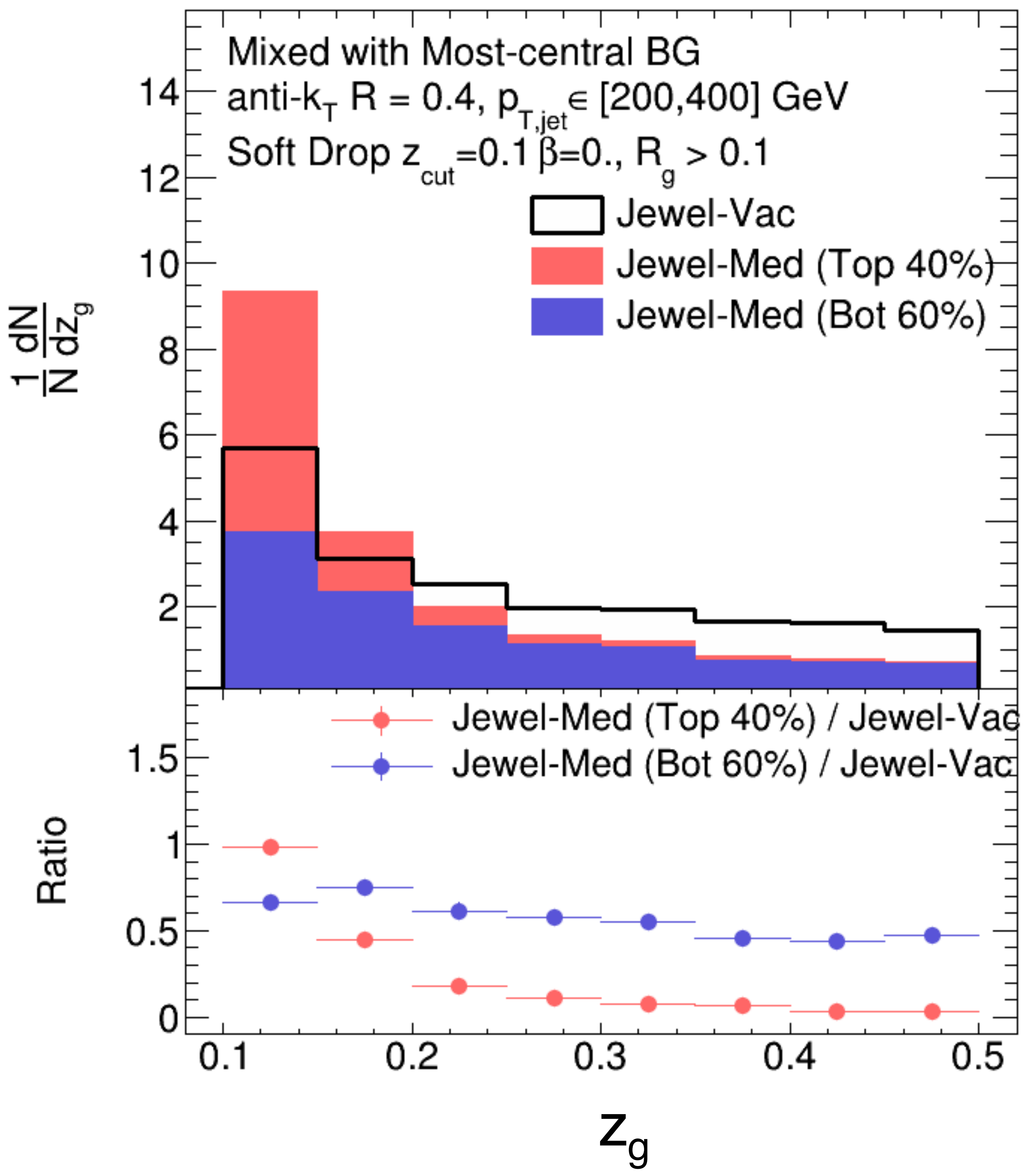
Q: How do we know that the LSTM indeed predicts the "quenchness" on a jet-by-jet basis?

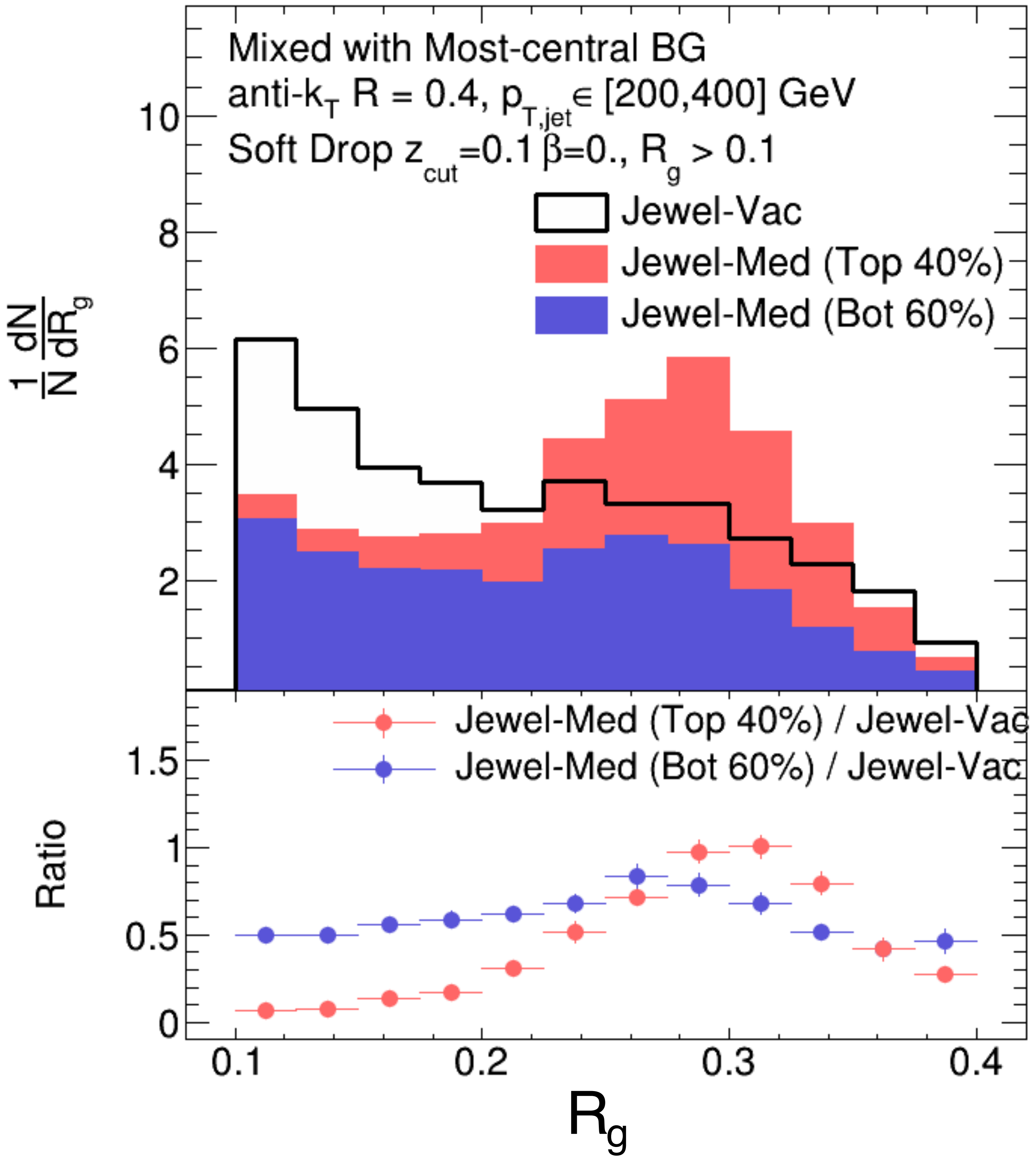
A: Using the LSTM outputs ("quenchness" predictions) to analyze jet observables

- ✓ Jet substructures
- ✓ Lund plane
- ✓ Jet fragmentation function
- ✓ Jet shape

Jet Quenchness Identification Results: — Jet Substructures

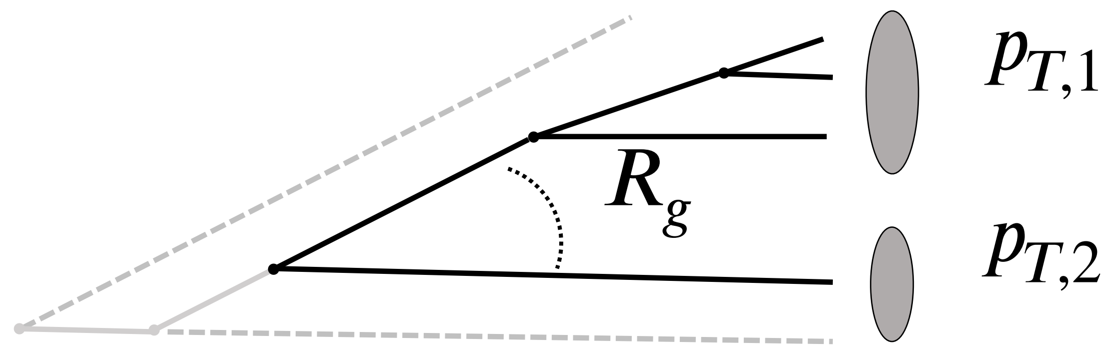




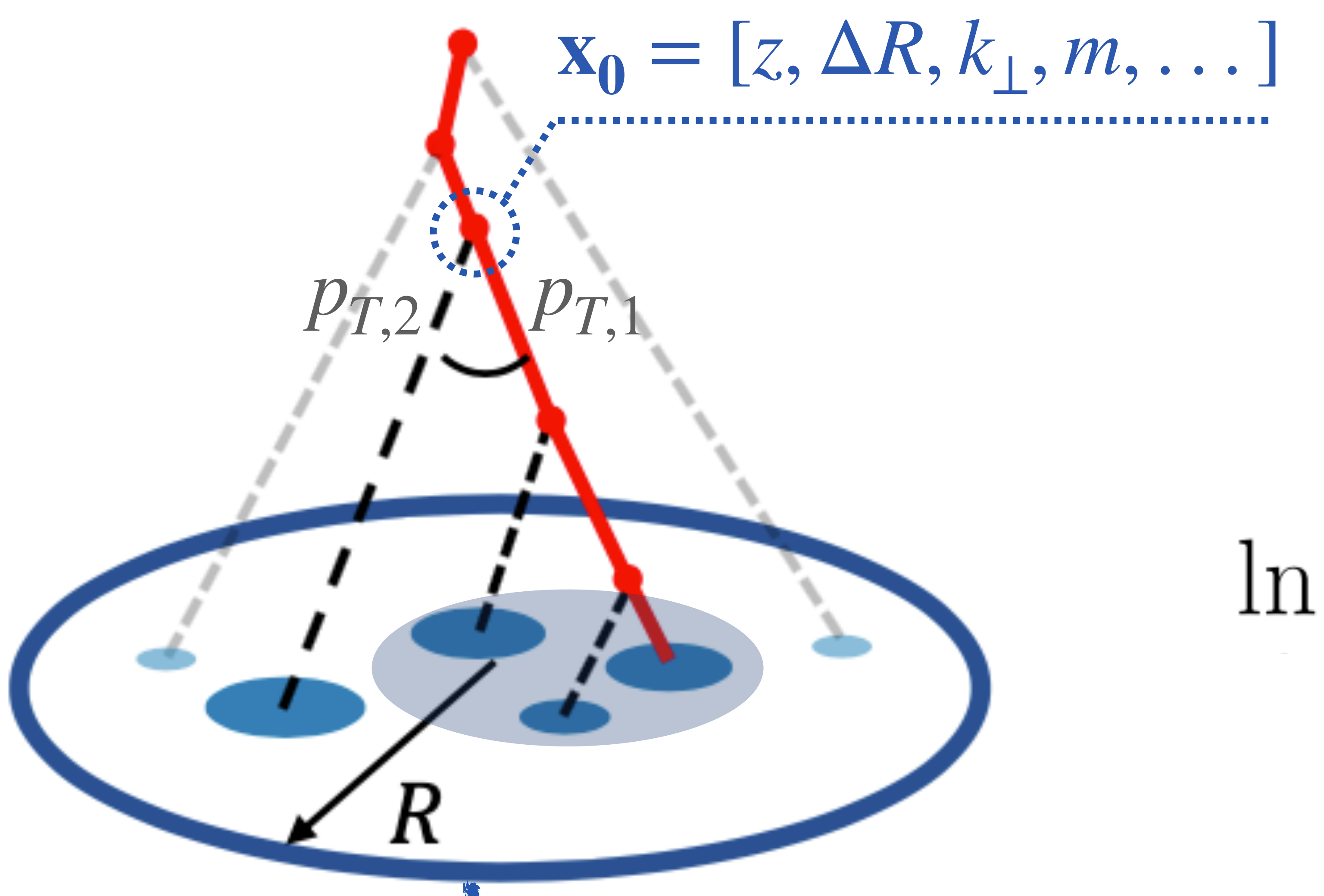


Quenchness: The LSTM output for each medium jet. If the value is closer to 1, then the jet is more quenched. And vice versa.

$$z_g = \frac{\min(p_{T,1},\ p_{T,2})}{p_{T,1} + p_{T,2}}$$



What is Lund Jet Plane?



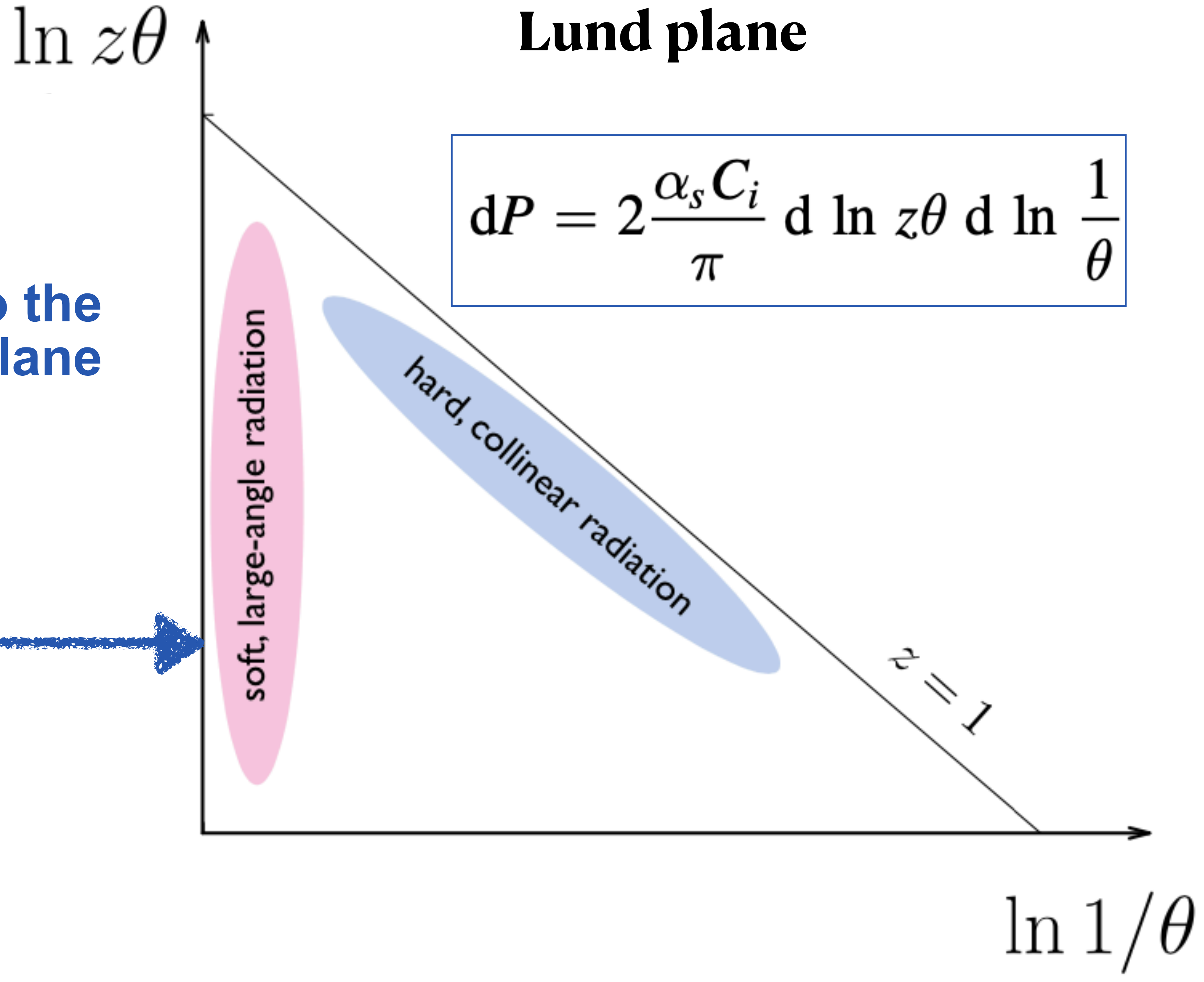
Lund jet plane characterizes the relative energy and angle of the first jet splitting (or *emission*). It contains all radiation pattern at the first splitting point.

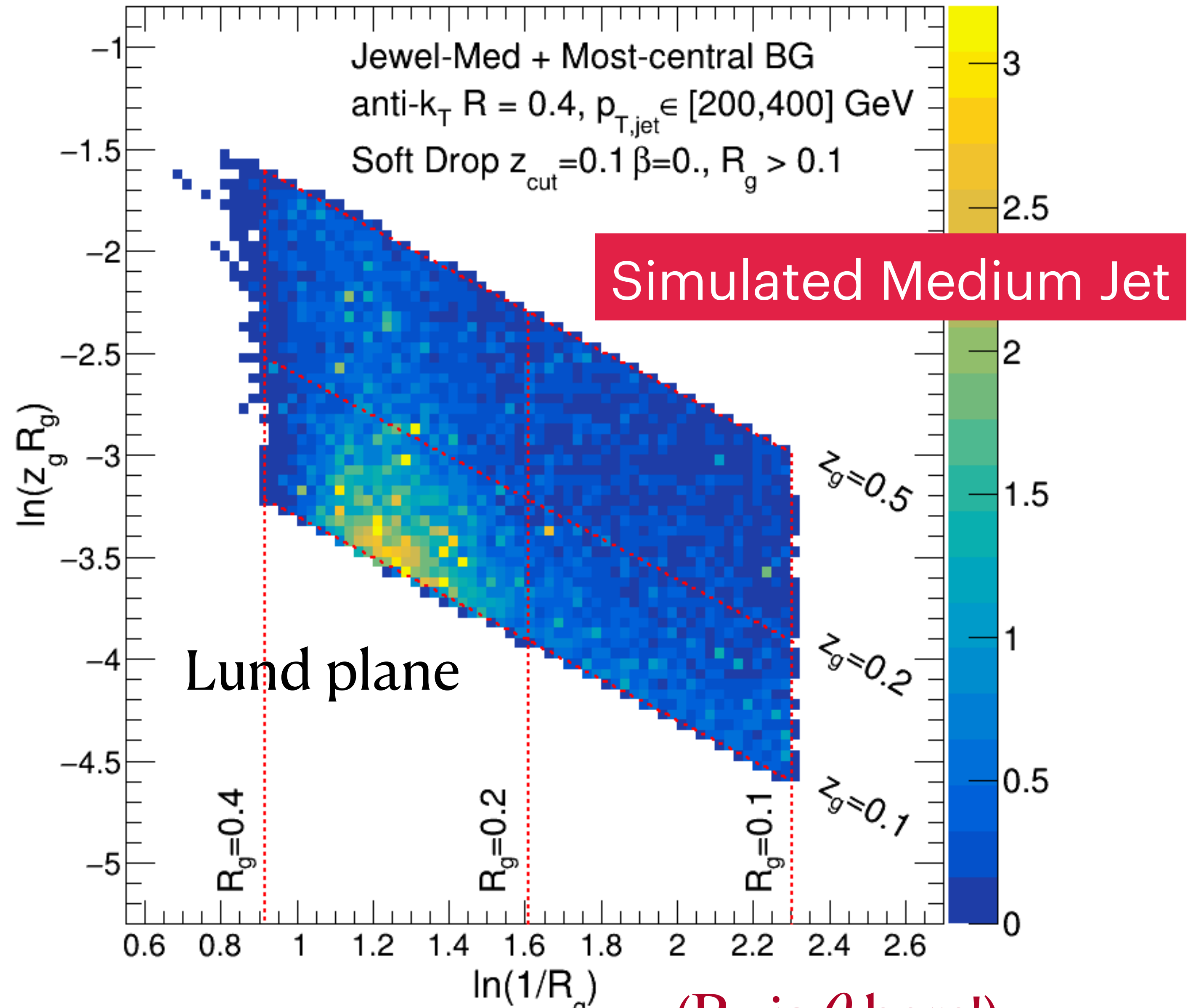
Map the jet to the Lund plane

z: **energy sharing fraction**

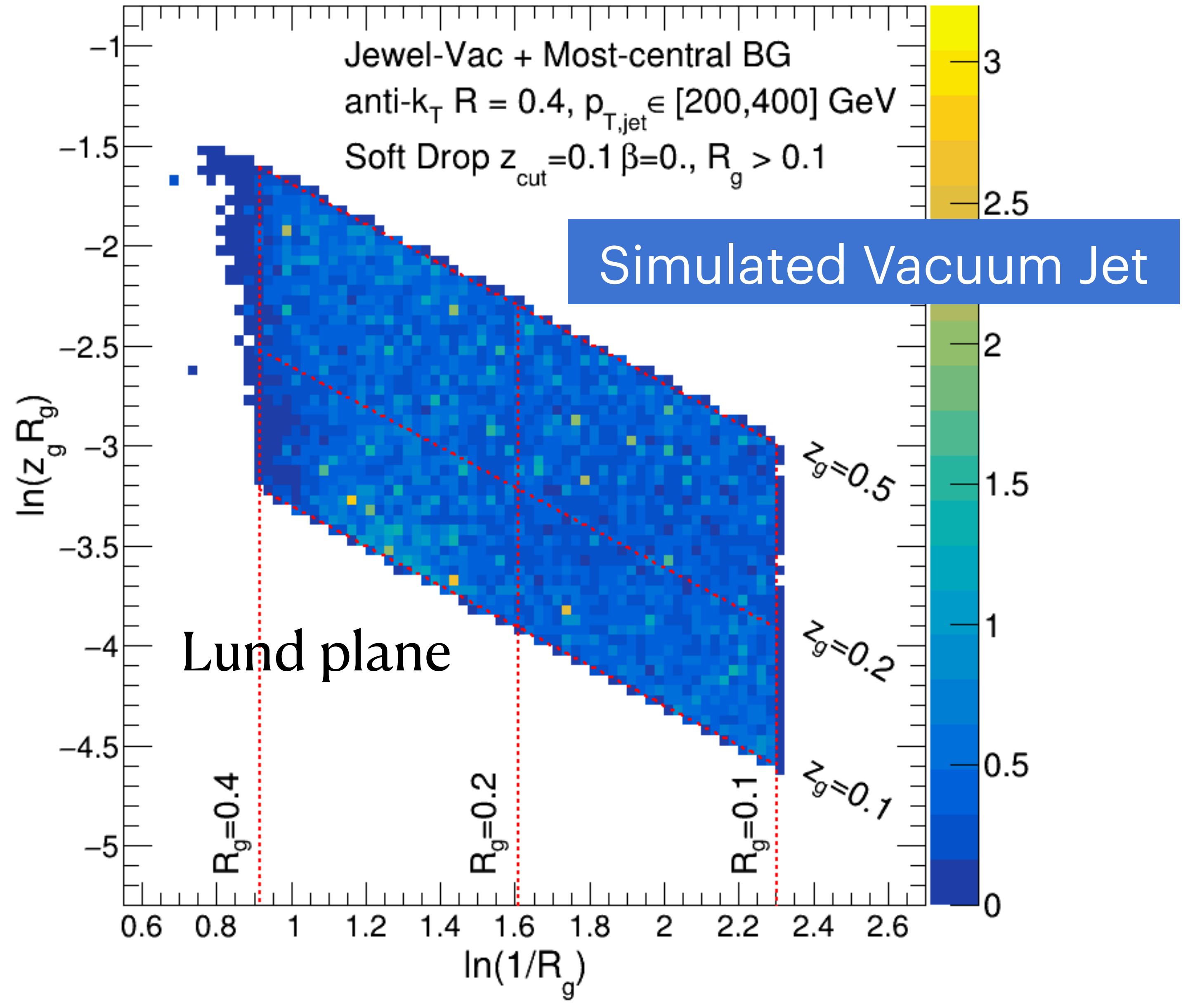
$$z = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}}$$

θ: **splitting angle**





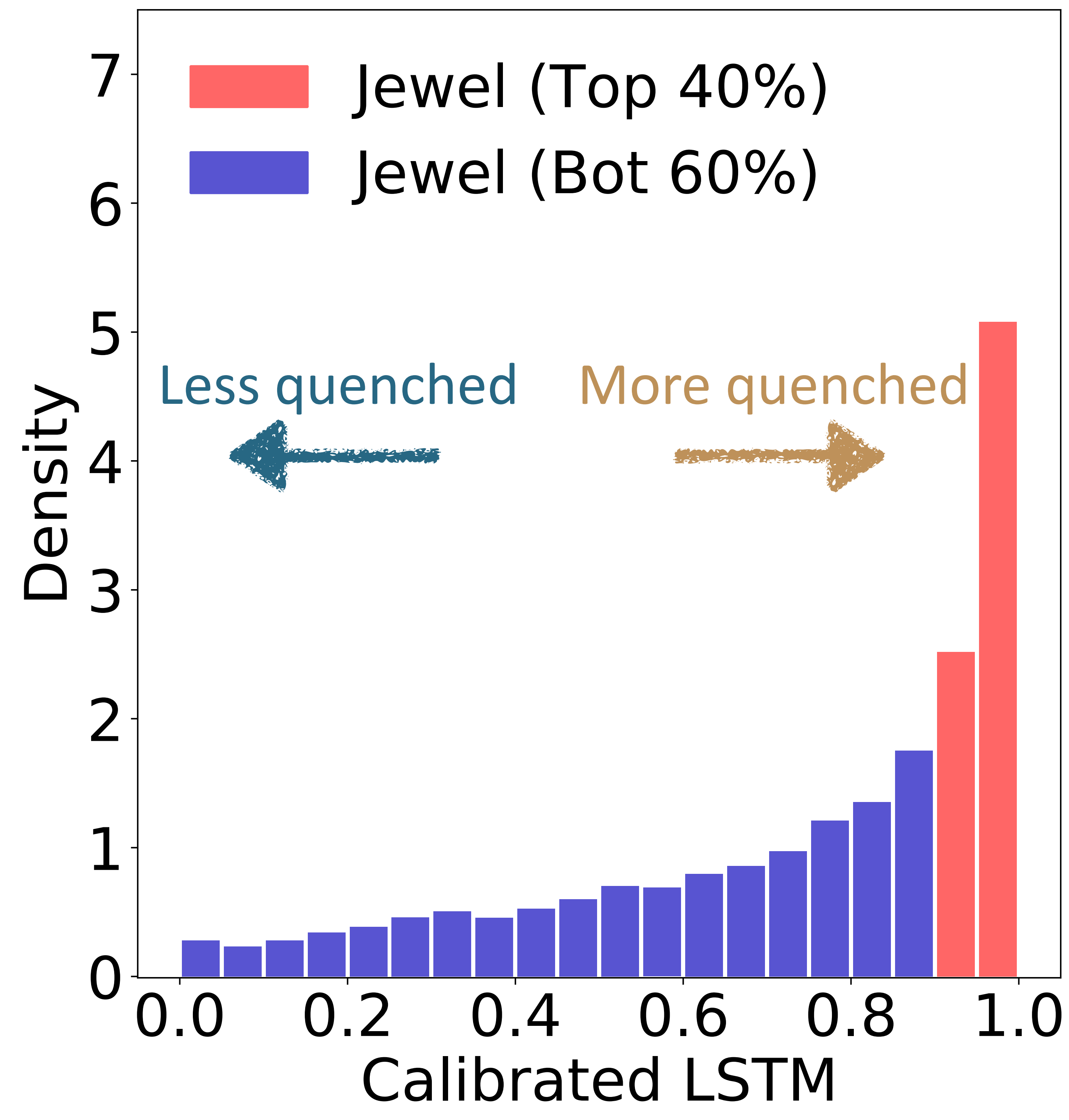
(R_g is θ here!)

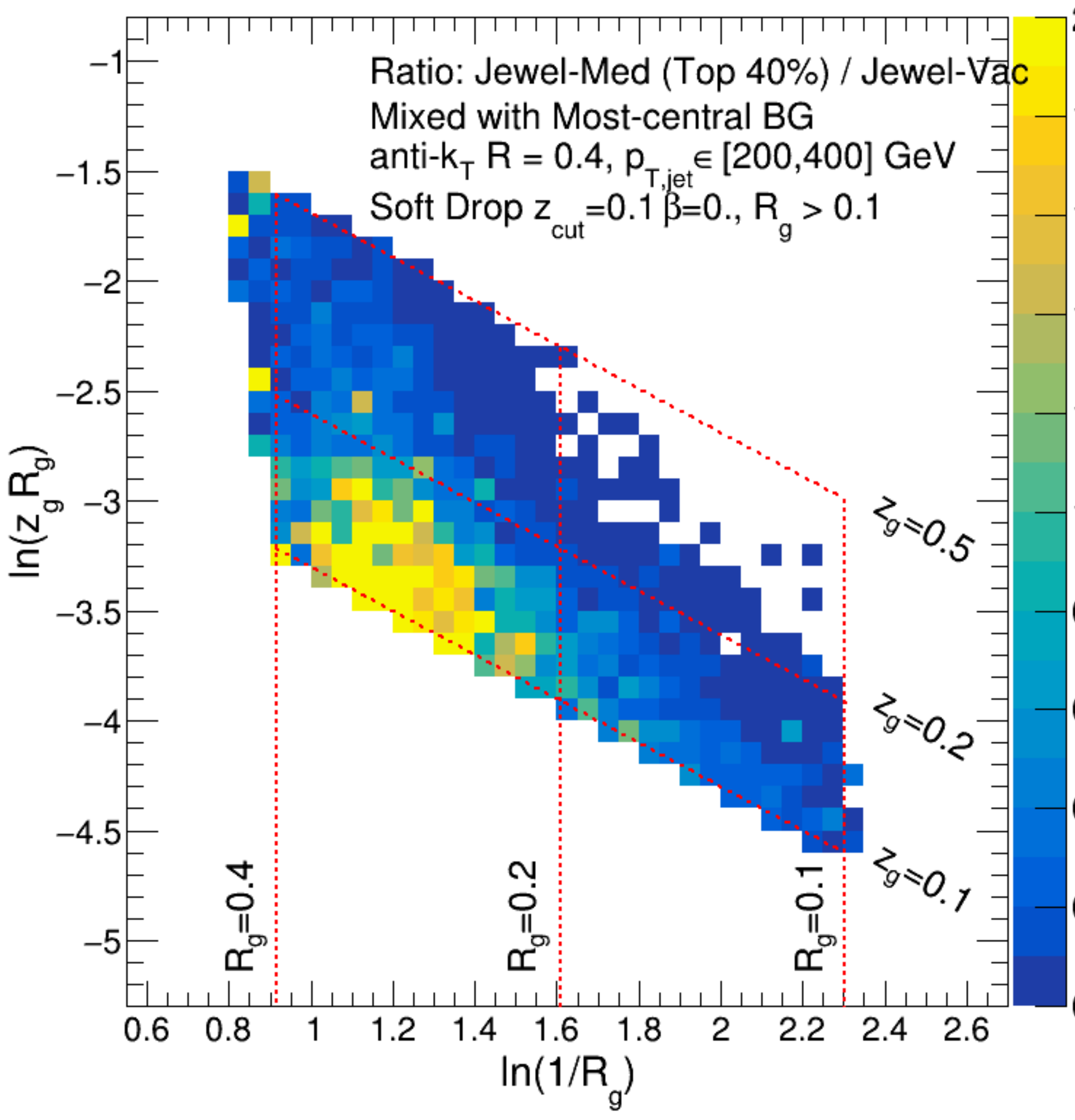


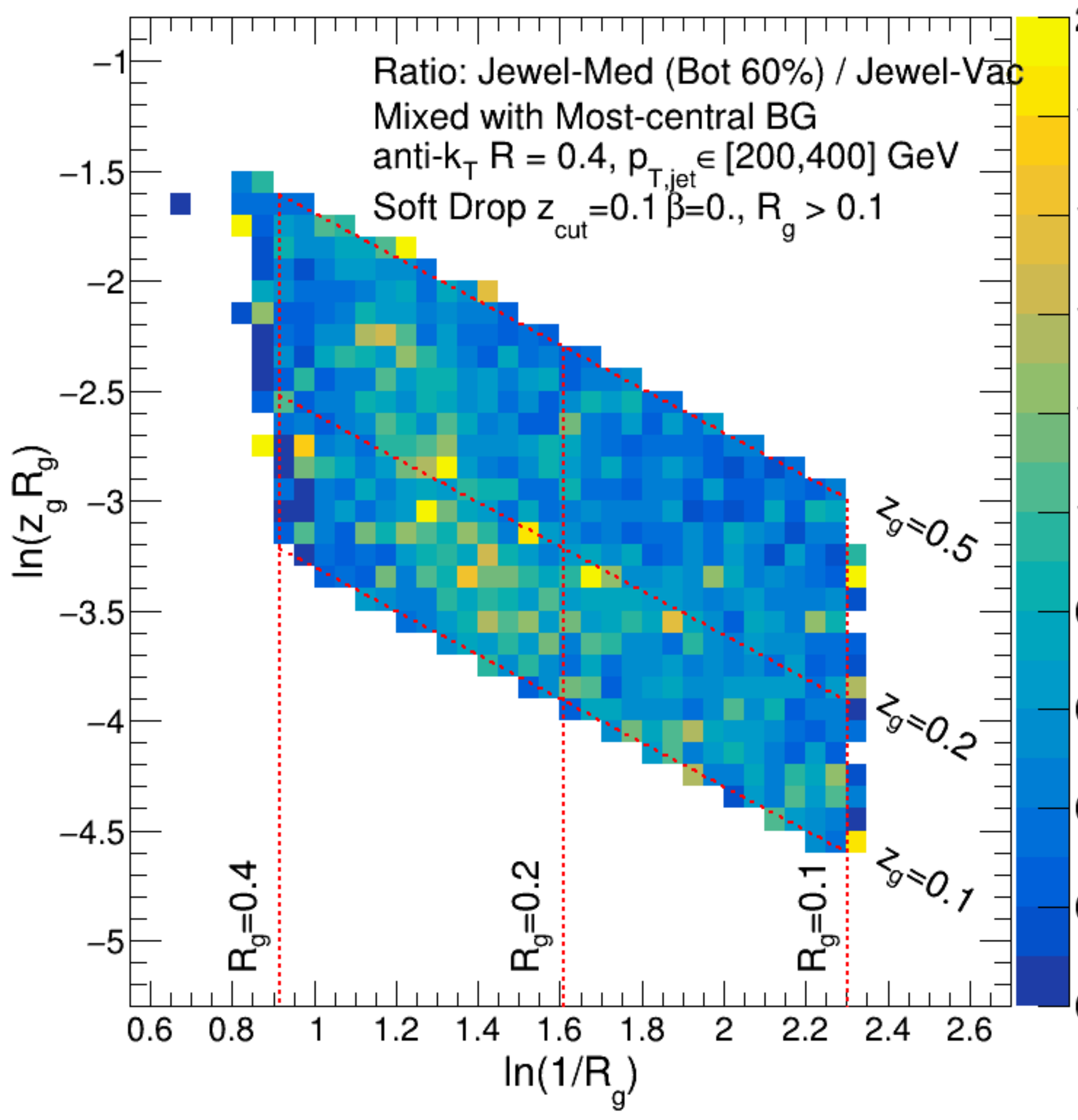
Figures are from: https://arxiv.org/abs/2206.01628

Jet Quenchness Identification Results: — Lund Plane

Top 40% “quenchness" jet & Bottom 60% “quenchness” jet predicted by LSTM



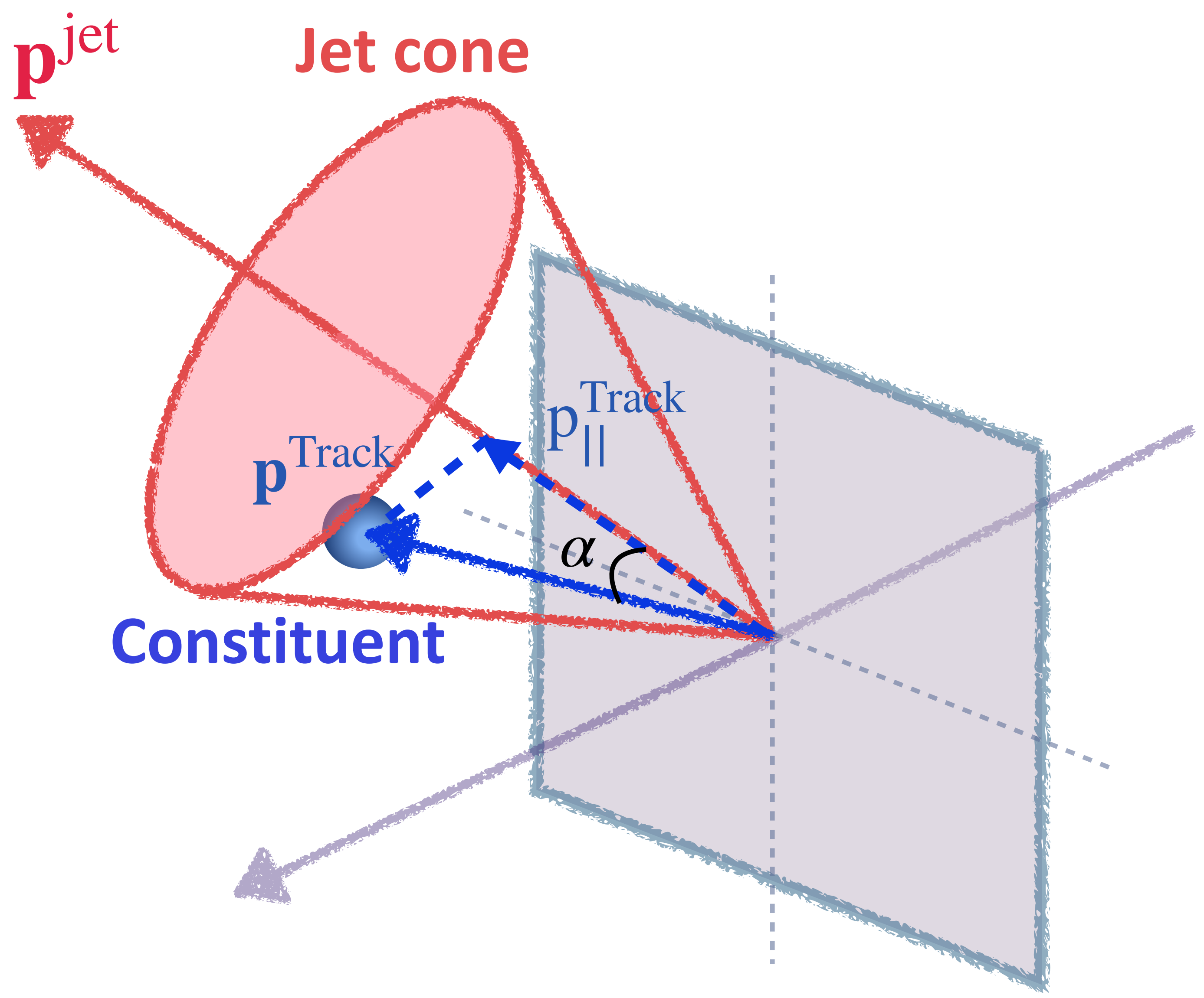




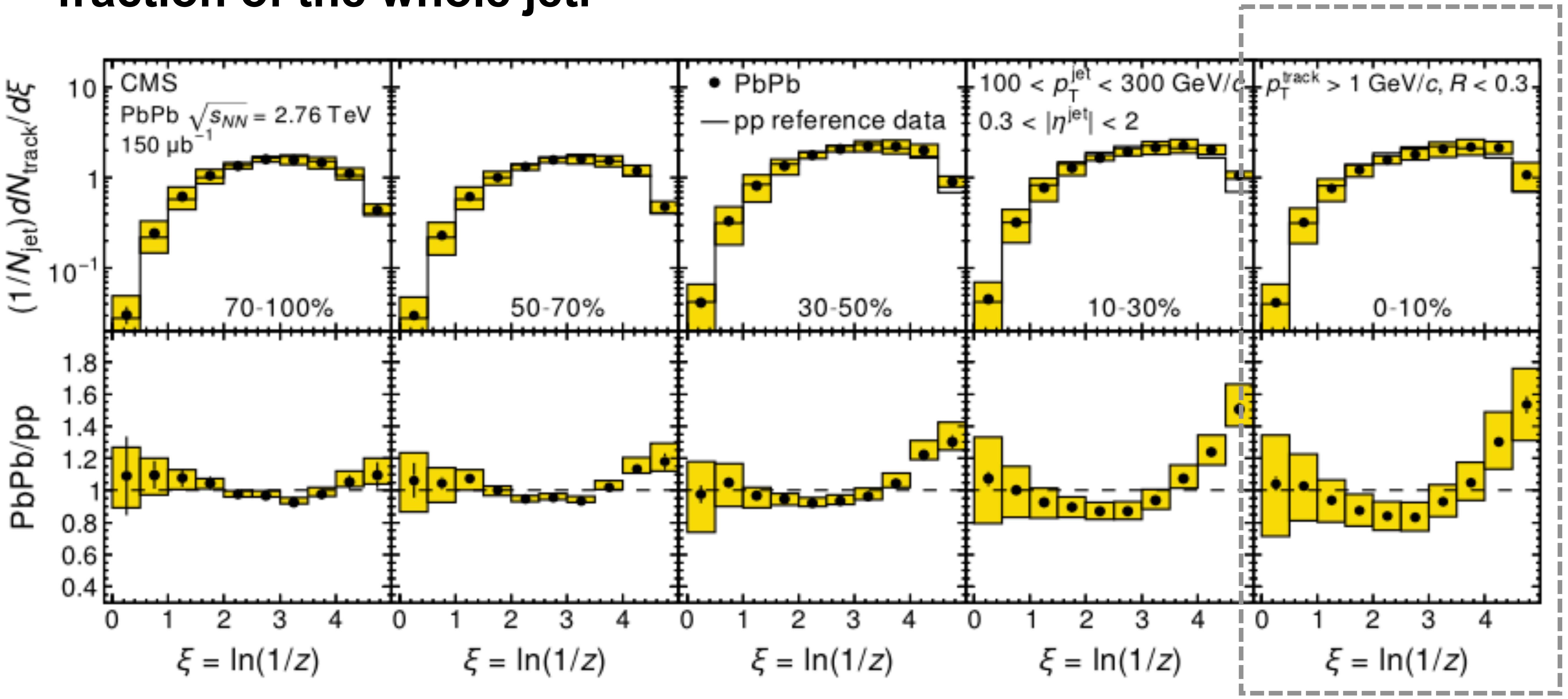
Quenchness: The LSTM output for each medium jet.
If the value is closer to 1, then the jet is more quenched. And vice versa.

What is Jet Fragmentation Function (JFF) ξ?

- For quantifying the partition of the jet energy into its constituent particles: $z = \frac{p_{||}^{\text{track}}}{p^{\text{jet}}}$, $\xi = \ln(\frac{1}{z})$



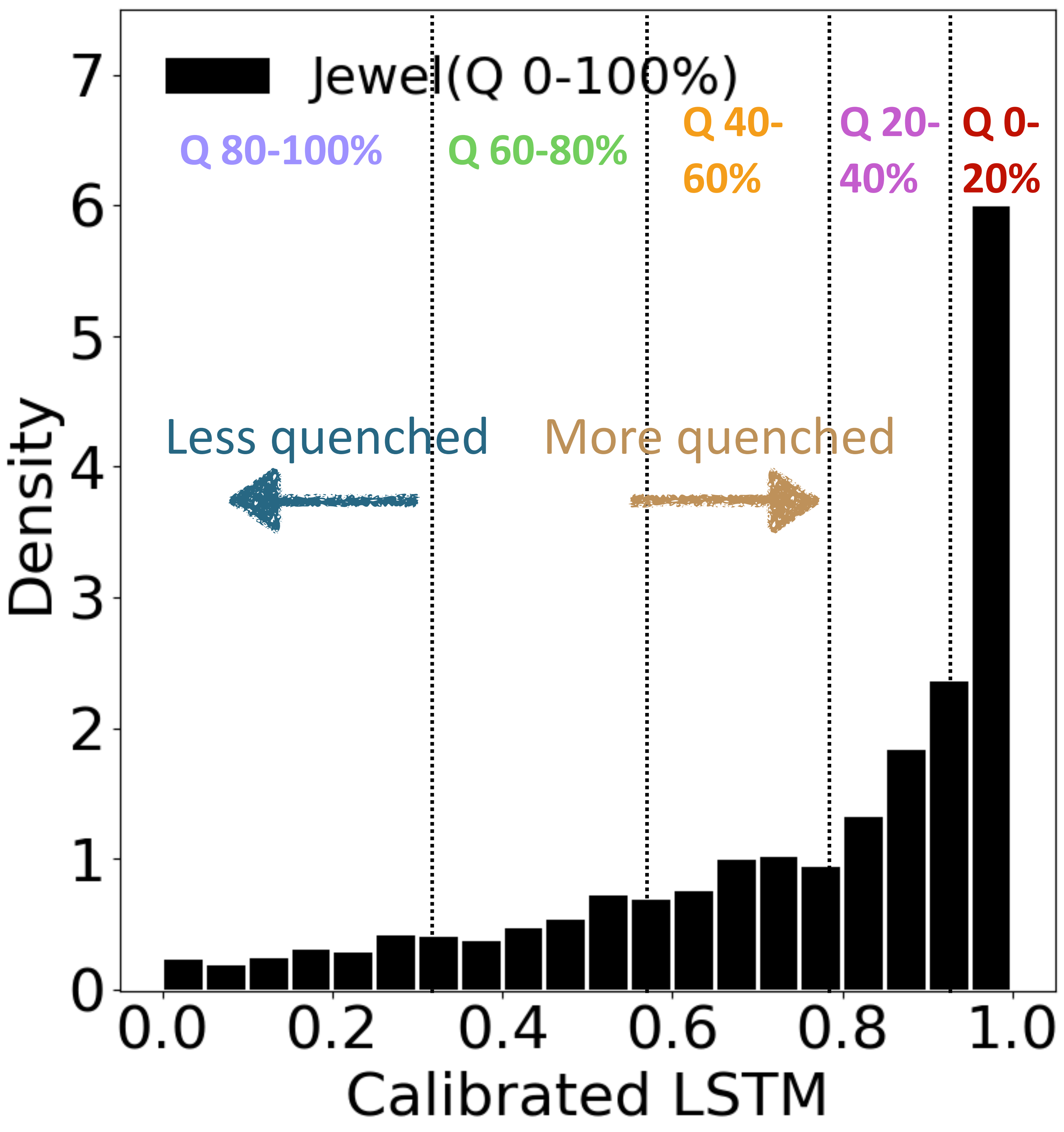
- Low ξ values correspond to high p_{T} particles within the jet cone, vice versa. It provides information about **the probability of finding one hadron inside jet cone containing certain a longitudinal energy fraction of the whole jet.**



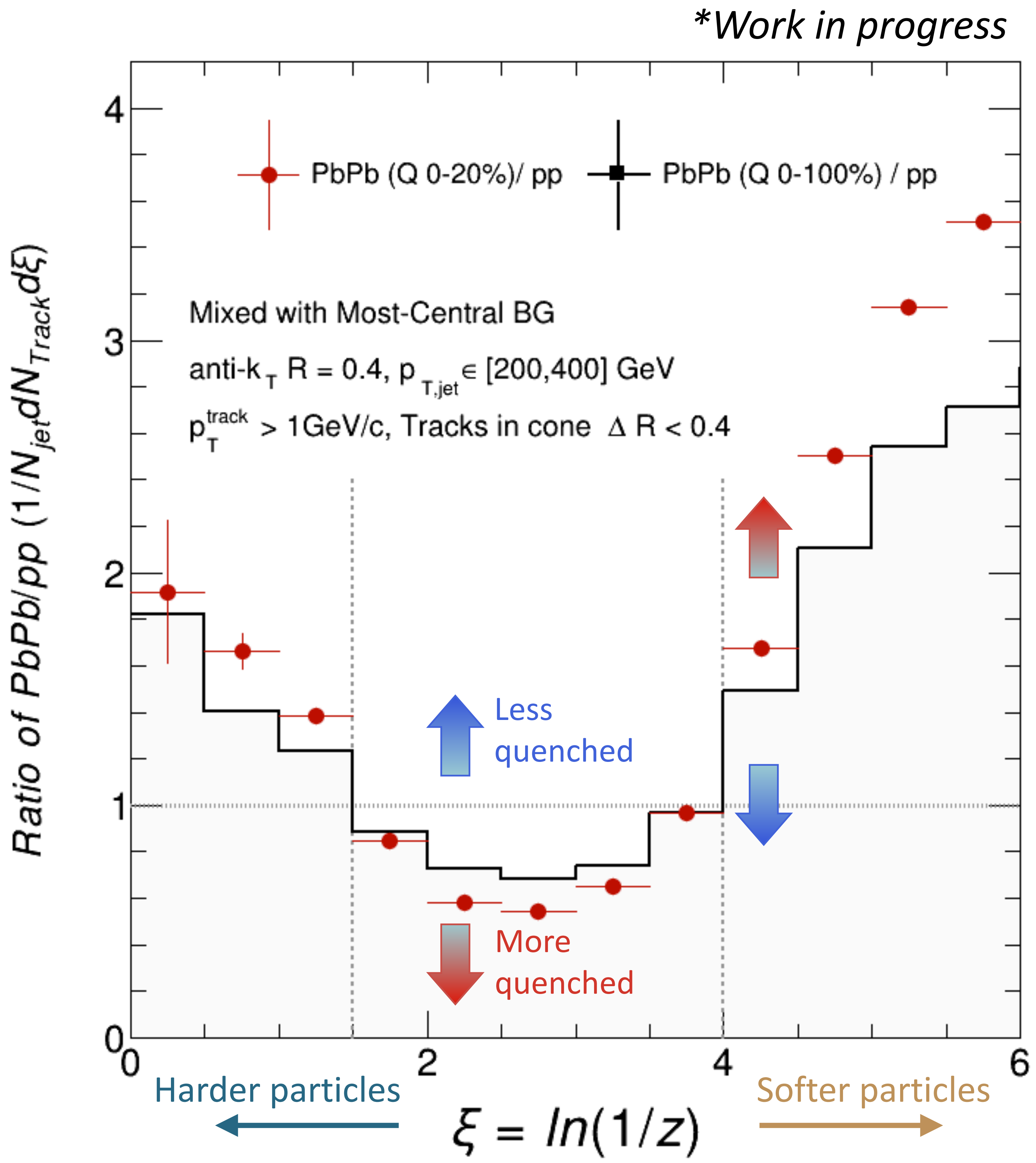
CMS Collaboration, Phys. Rev. C **90**, 024908

- For the most central collisions, a significant enhancement on high ξ values (p_T^{track}< 3GeV) is observed, with depletion in the intermediate region as compensation. **The result shows an enhancement of soft particle contribution to the jet energy and a suppression of high p_T particles in central PbPb collisions, compared to pp collisions.**

Jet Quenchness Identification Results: JFF ratio



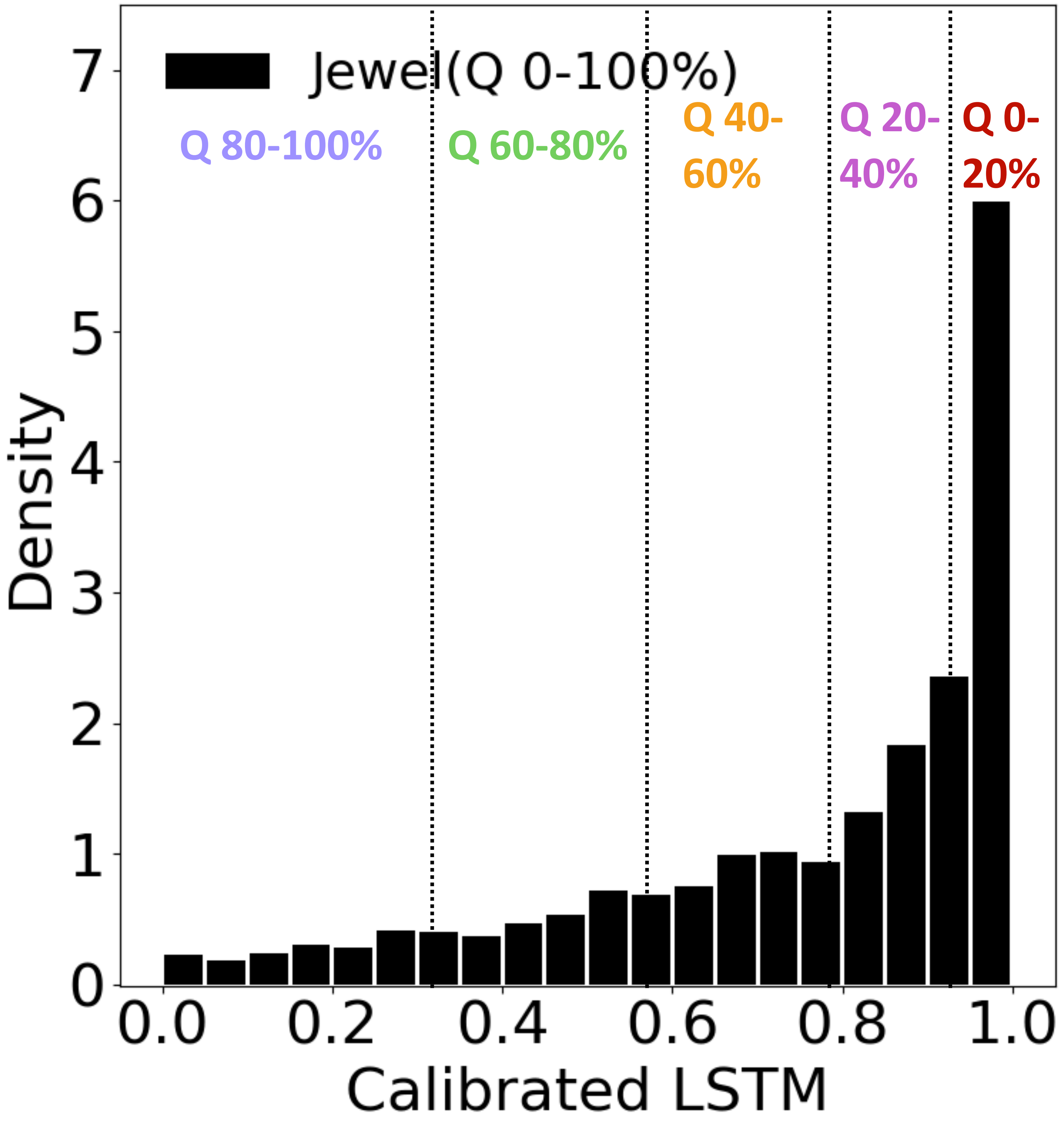
Five "quenchness" classes of JEWEL jets predicted by the LSTM



The JFF ratios from five "quenchness" classes of JEWEL jets divided by the Jewel-vac jets

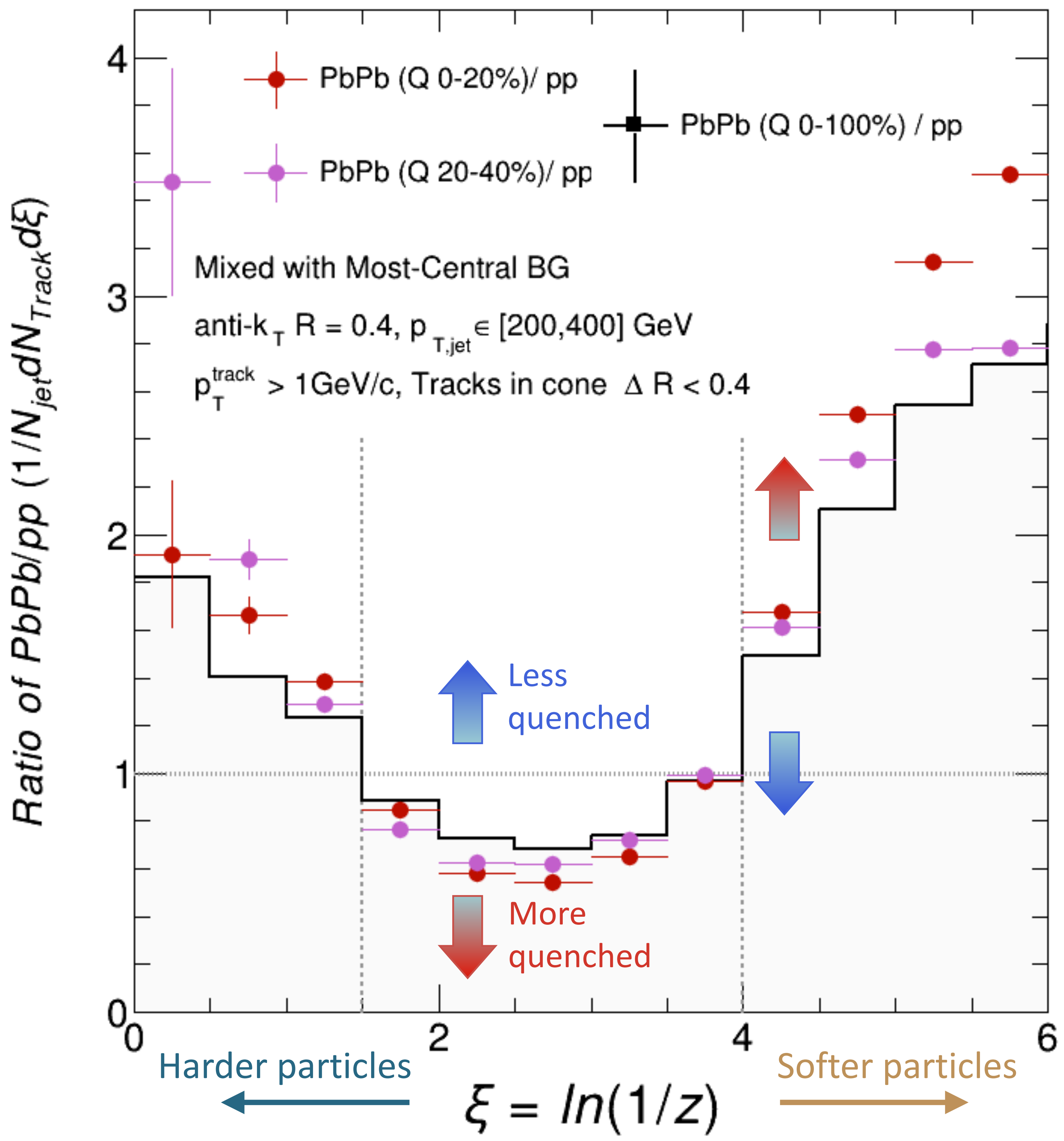
- For **0-20% quenchness** jets, the large ξ is **most** enhanced with a depletion of intermediate ξ.

Jet Quenchness Identification Results: JFF ratio



Five "quenchness" classes of JEWEL jets predicted by the LSTM

*Work in progress

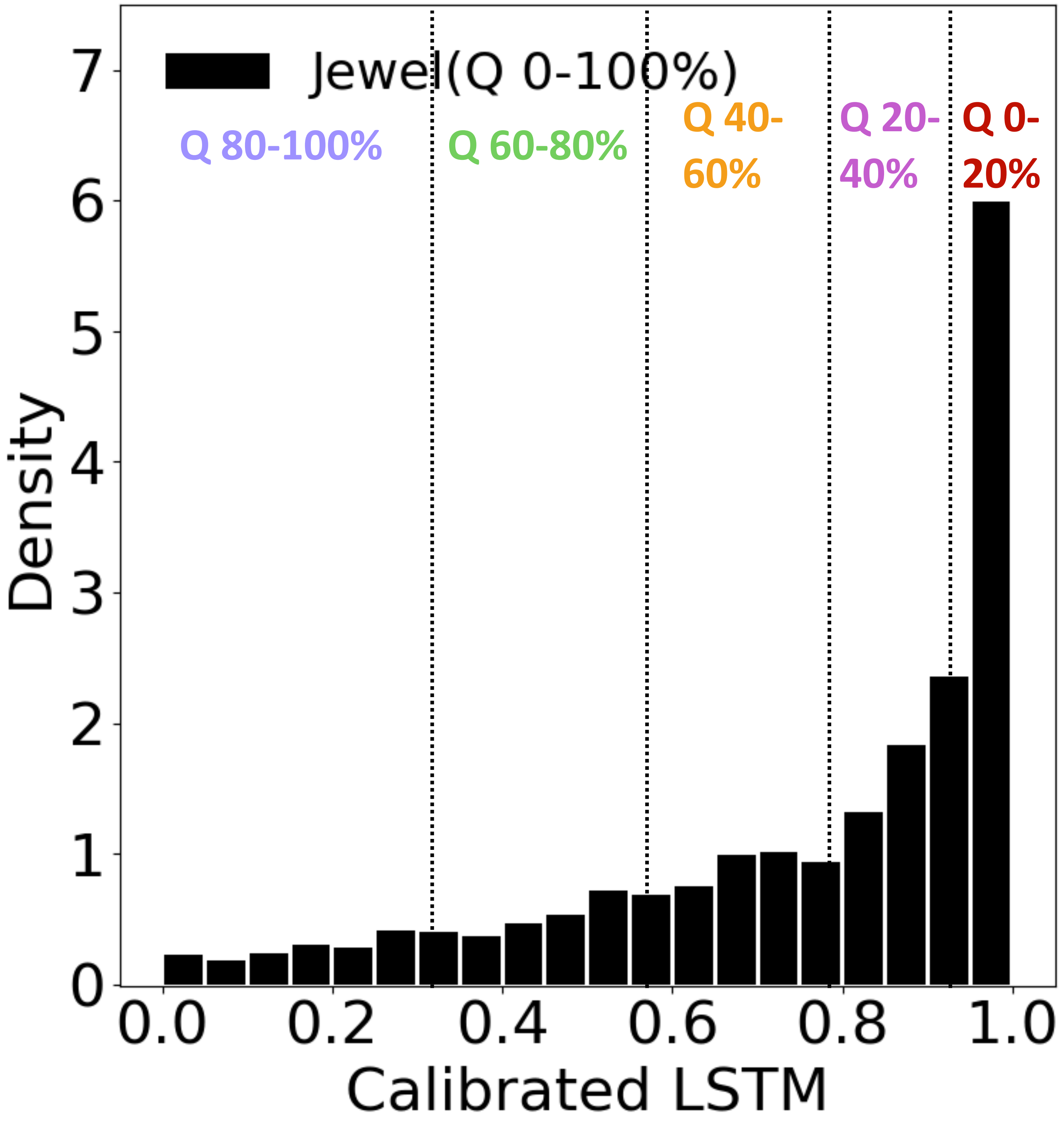


The JFF ratios from five "quenchness" classes of JEWEL jets divided by the Jewel-vac jets

- For **0-20% quenchness** jets, the large ξ is **most** enhanced with a depletion of intermediate ξ.
- For **20-40% quenchness** jets, the **most** enhanced region is of small ξ;

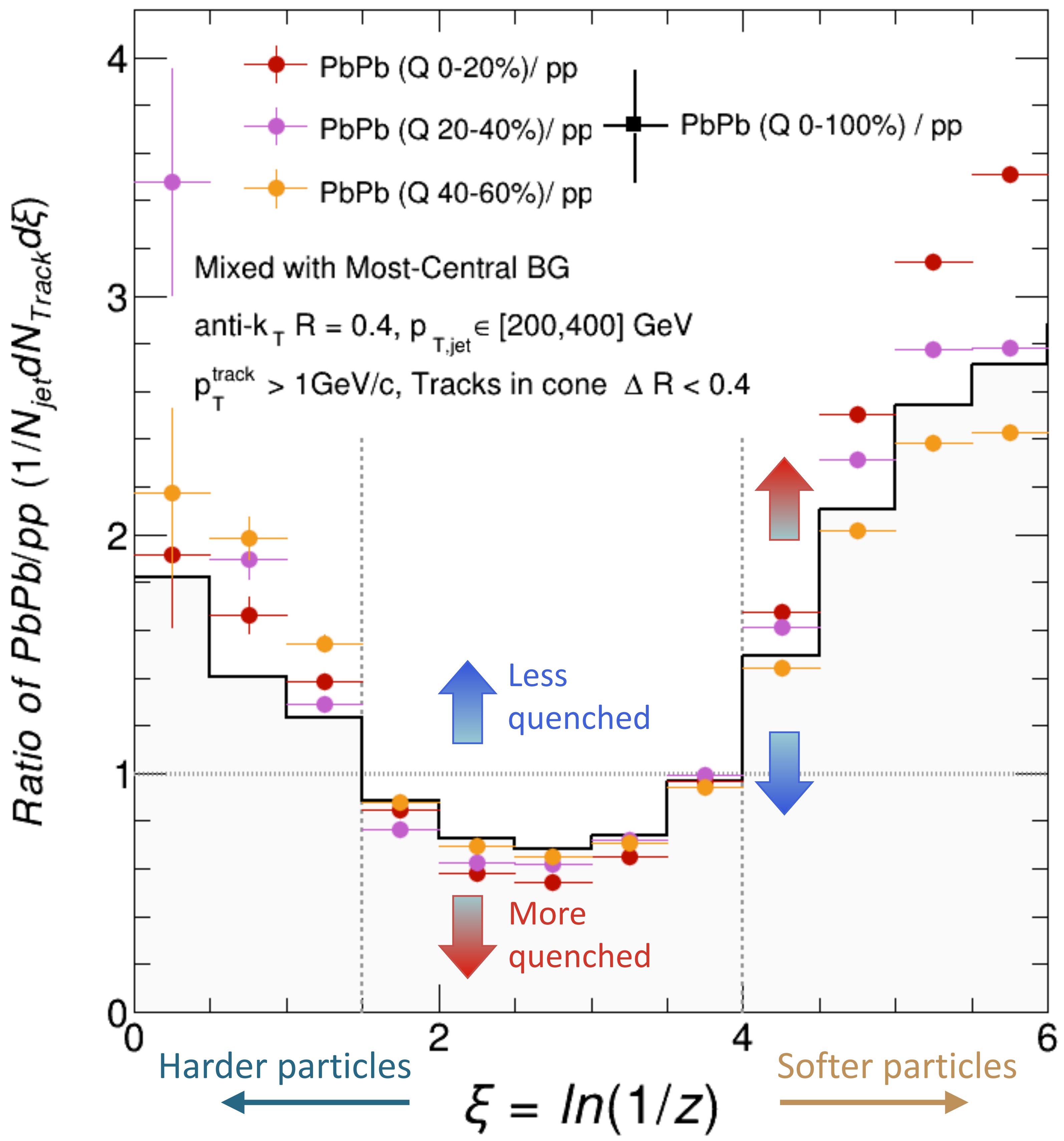
A bias towards jets that are less fragmented than the average quenched jets

Jet Quenchness Identification Results: JFF ratio



Five "quenchness" classes of JEWEL jets predicted by the LSTM

*Work in progress

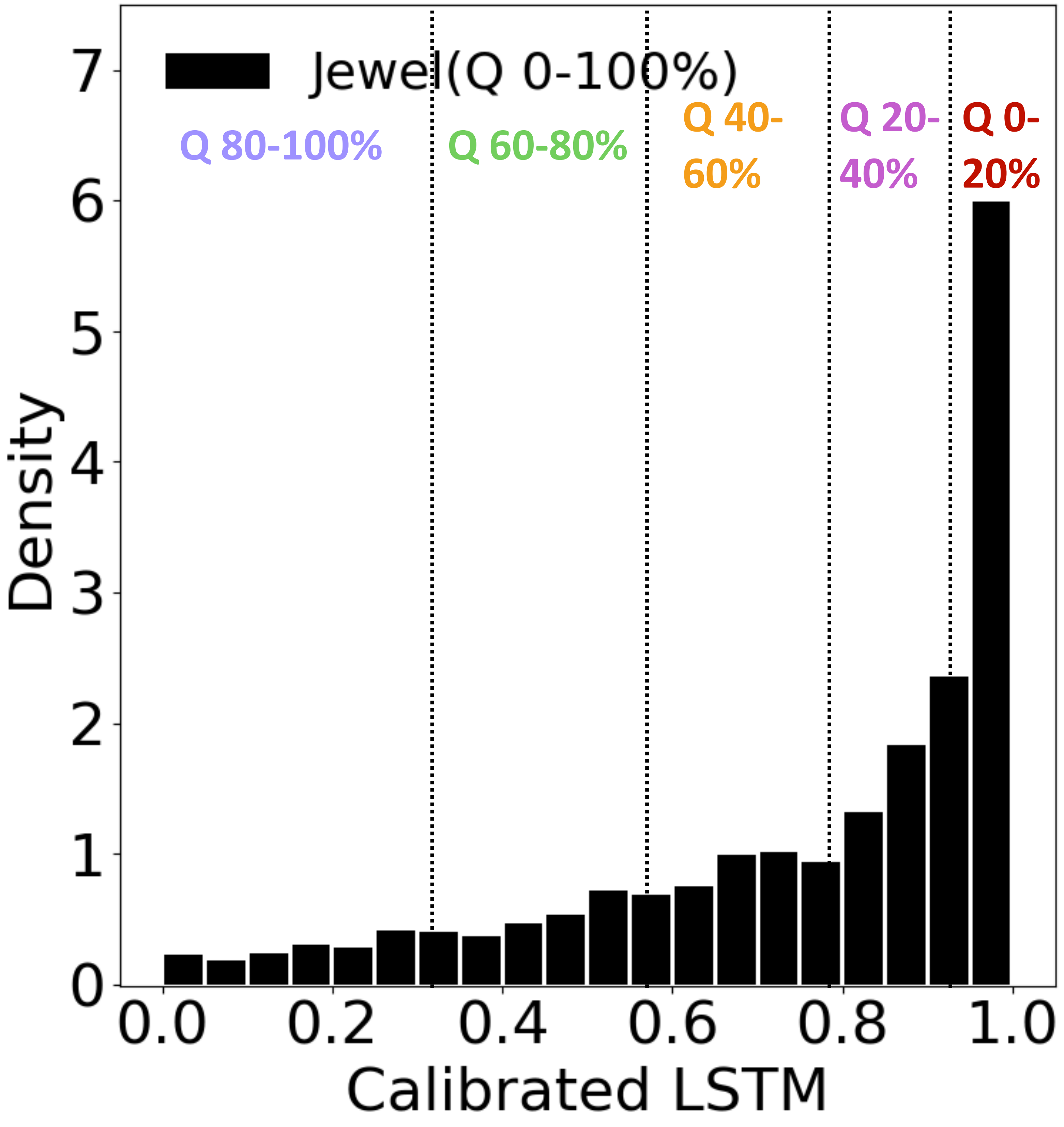


The JFF ratios from five "quenchness" classes of JEWEL jets divided by the Jewel-vac jets

- For **0-20% quenchness** jets, the large ξ is **most** enhanced with a depletion of intermediate ξ.
- For **20-40% quenchness** jets, the **most** enhanced region is of small ξ; Same for **40-60% quenchness** jets.

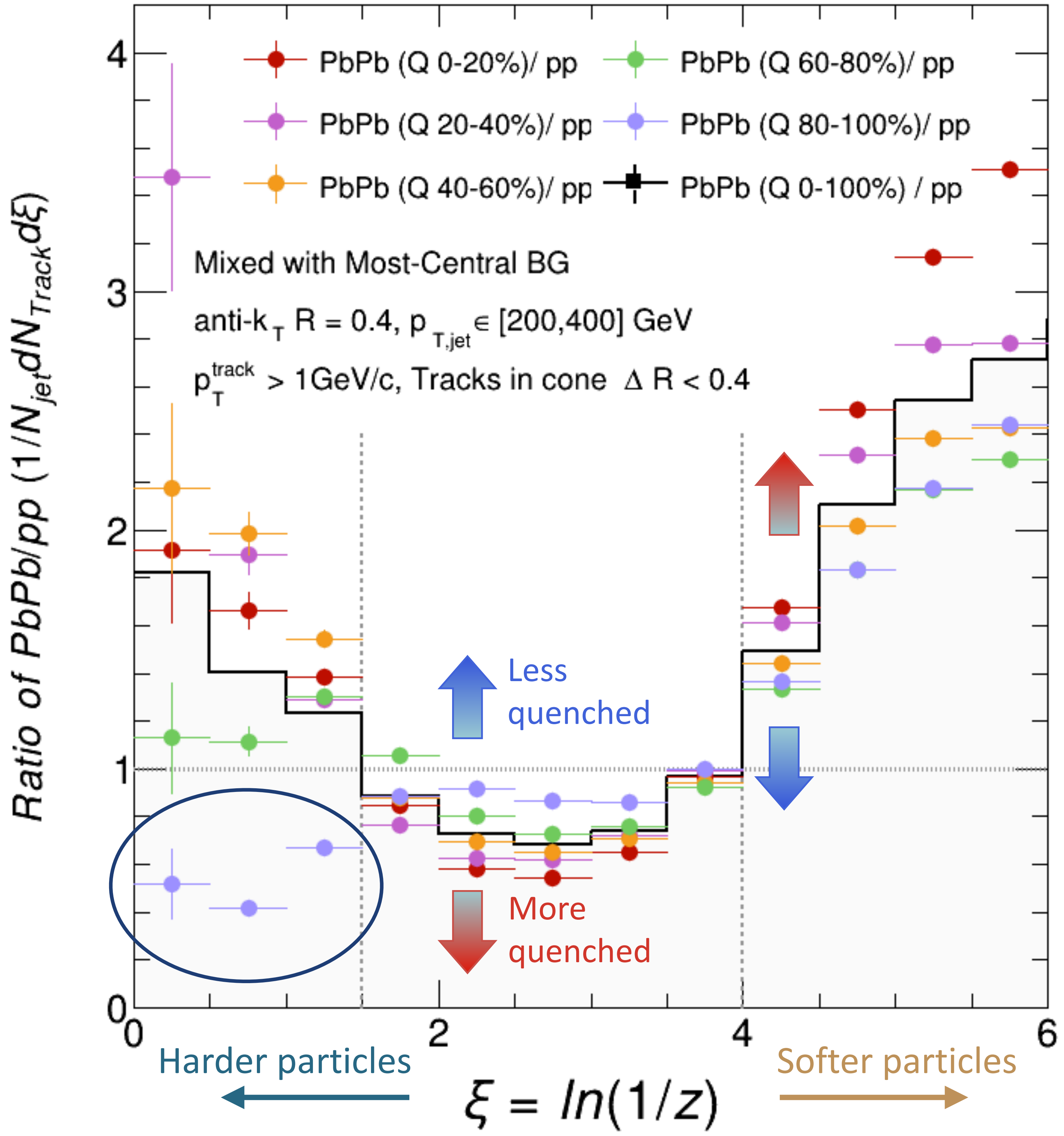
A bias towards jets that are less fragmented than the average quenched jets

Jet Quenchness Identification Results: JFF ratio



Five “quenchness” classes of JEWEL jets predicted by the LSTM

*Work in progress



The JFF ratios from five “quenchness” classes of JEWEL jets divided by the Jewel-vac jets

- For **0-20% quenchness** jets, the large ξ is **most** enhanced with a depletion of intermediate ξ.
- For **20-40% quenchness** jets, the **most** enhanced region is of small ξ; Same for **40-60% quenchness** jets.
- For **80-100% quenchness** jets, the ratio between the them and vacuum still deviates from unity.

"*they tends to be narrower and less fragmented than the average jet population in vacuum*" — *J. High Energ. Phys. 2021, 206 (2021)*

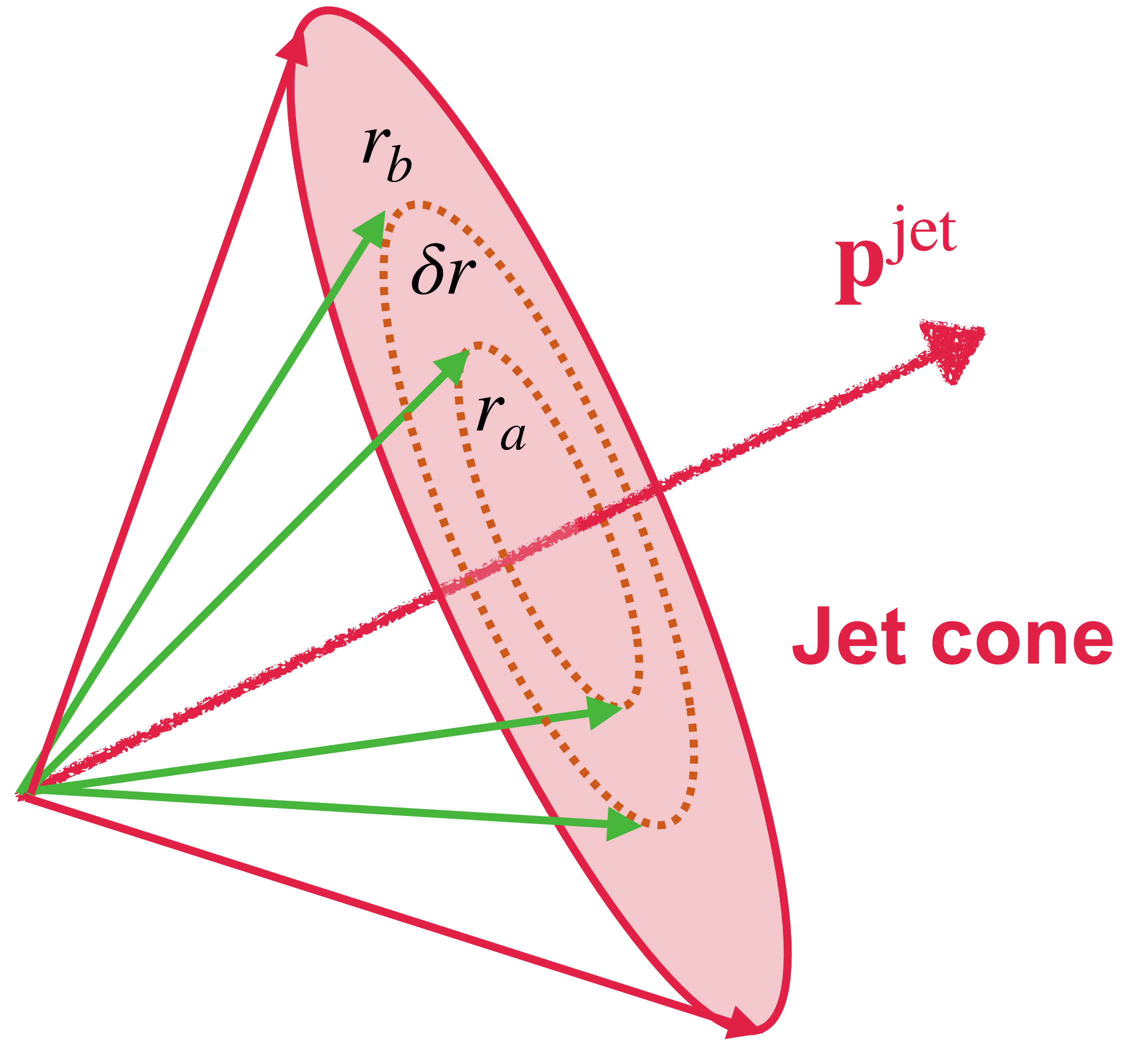
They behave like biased vacuum jets (with small LSTM values) in the small ξ region.

What is Jet Shape Function $\rho(r)$?

- The jet shape function,

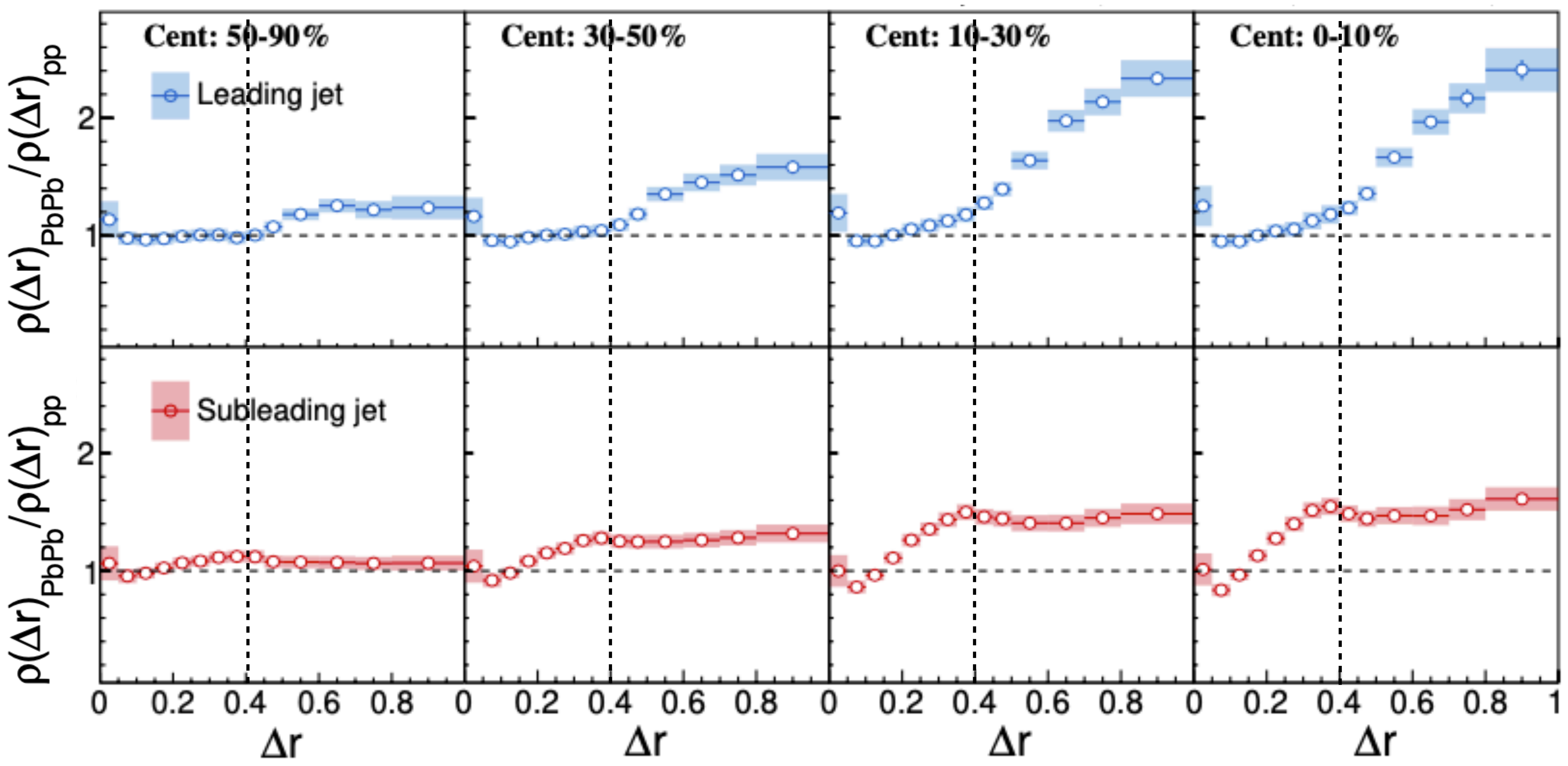
$$\rho(r) = \frac{1}{\delta r} \frac{1}{N_{\text{jet}}} \sum_{\text{jets}} \frac{\Sigma_{\text{tracks} \in [r_a, r_b)} p_T^{\text{track}}}{p_T^{\text{jet}}},$$

provides information about the radial distribution of the momentum carried by the jet constituents (fragments).

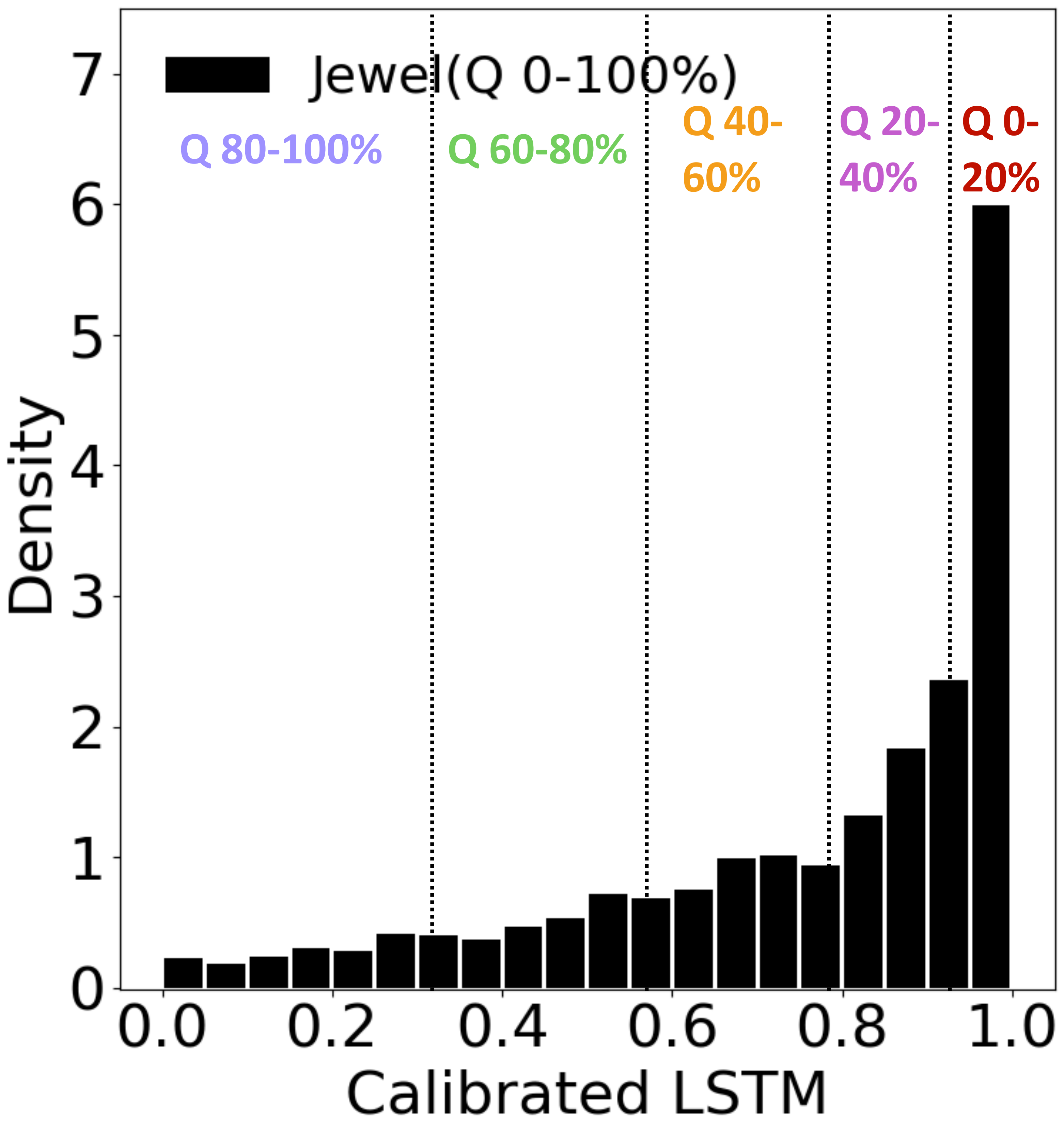


- The jet shape ratios between PbPb and pp show a redistribution of jet energy to softer particles extending to large angles away from the jet axis.
- The energy lost due to parton propagation in QGP is observed to be recovered by soft hadrons at large angles with respect to the jet axis.

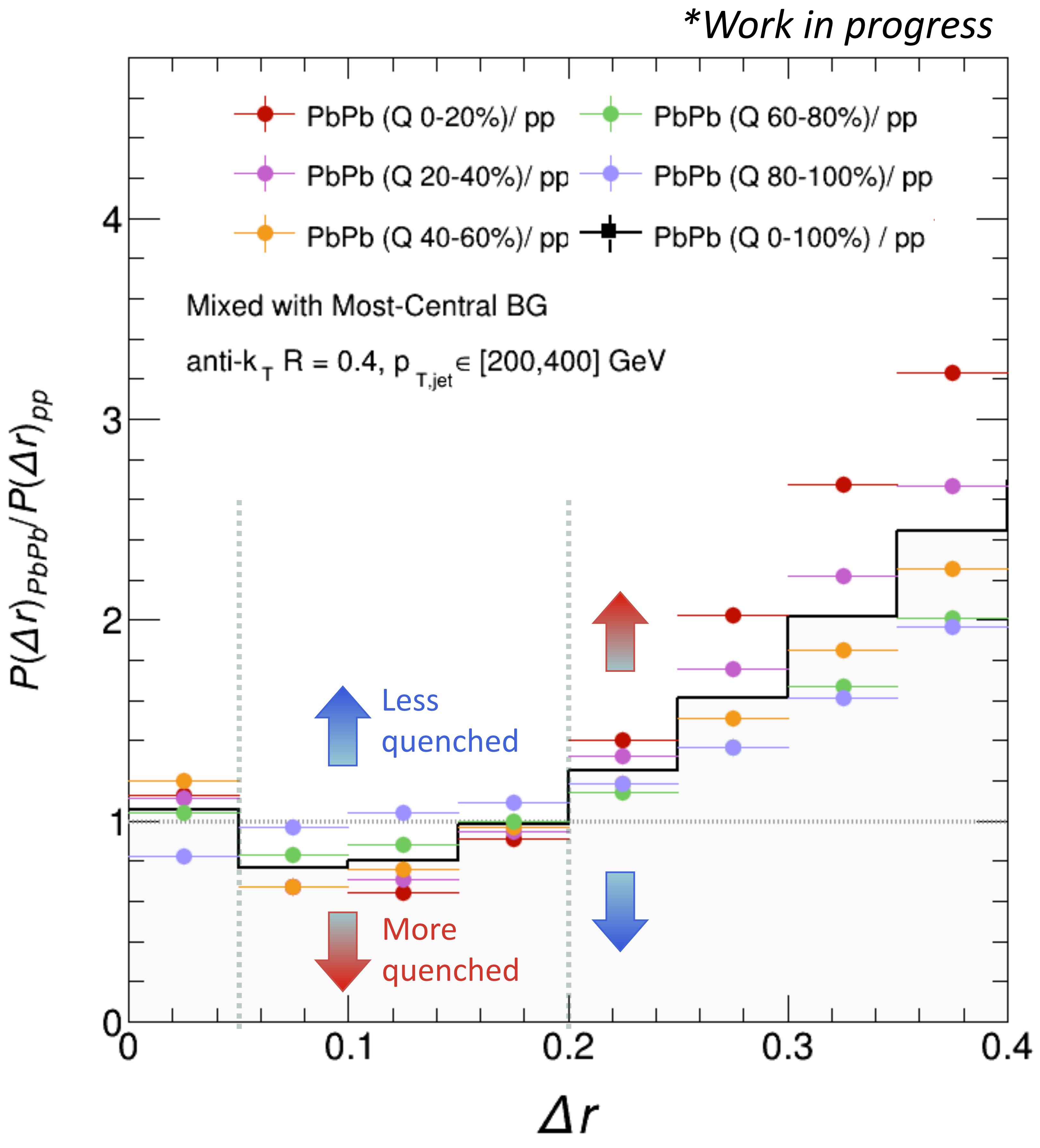
CMS Collaboration, J. High Energ. Phys. 2021, 116 (2021)



Jet Quenchness Identification Results: Jet Shape(JS) ratio



Five "quenchness" classes of JEWEL jets predicted by the LSTM



The JS ratios from five "quenchness" classes of JEWEL jets divided by the Jewel-vac jets

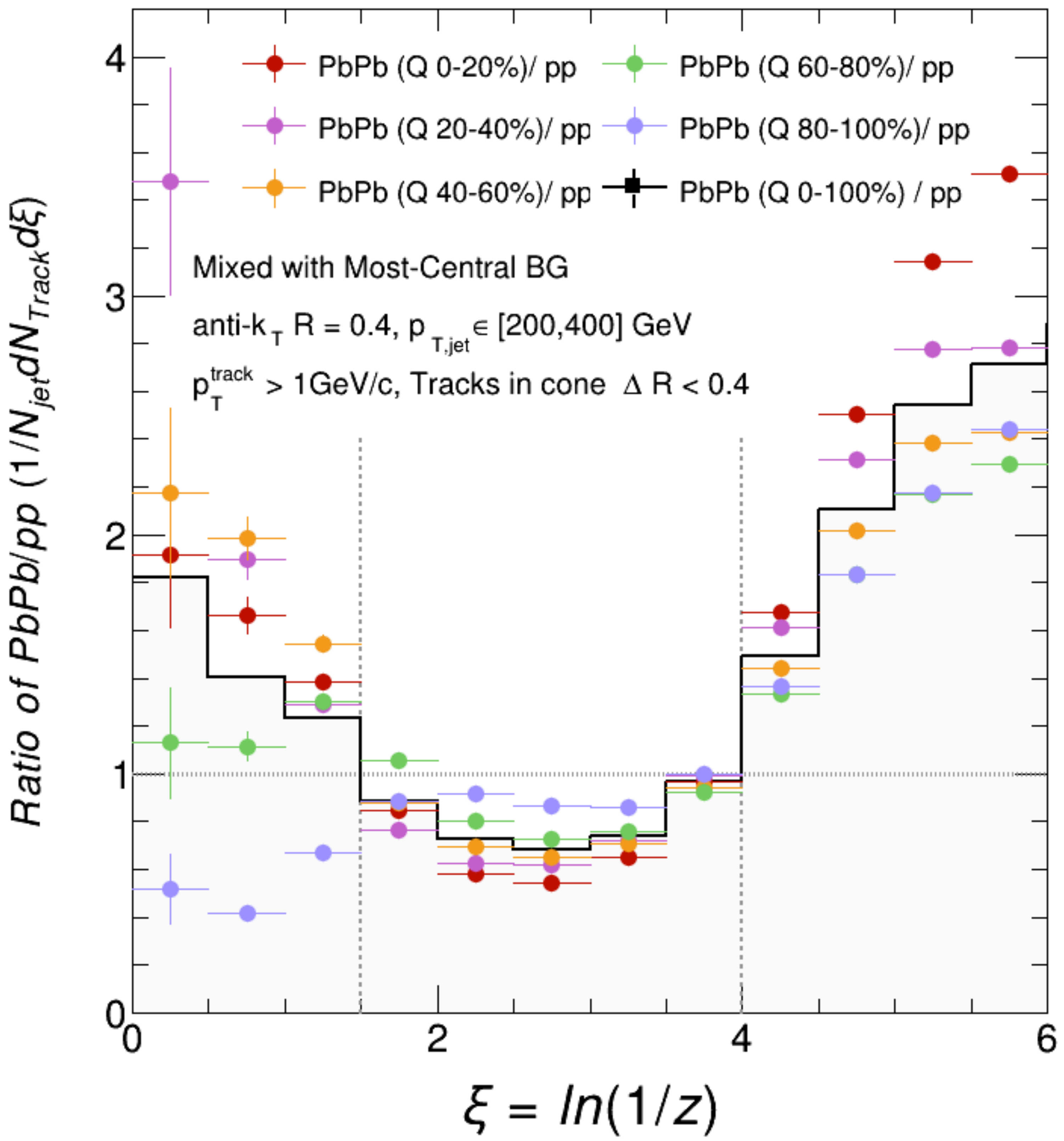
The JS ratios from five quenchness classes divided by the vacuum jets also show various jet quenching modes, consistent with the JFF ratio results.

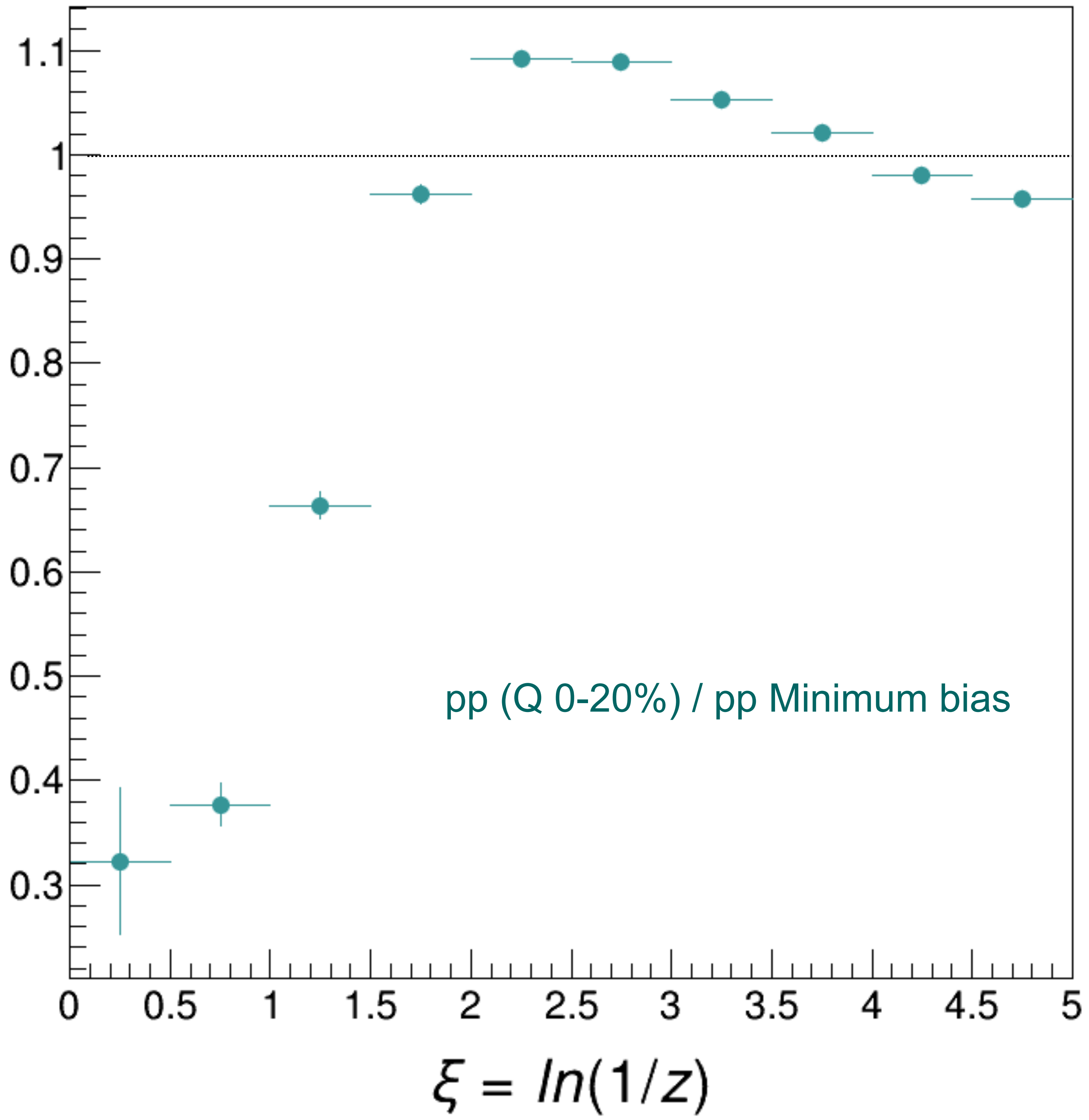
Conclusions and Next Step

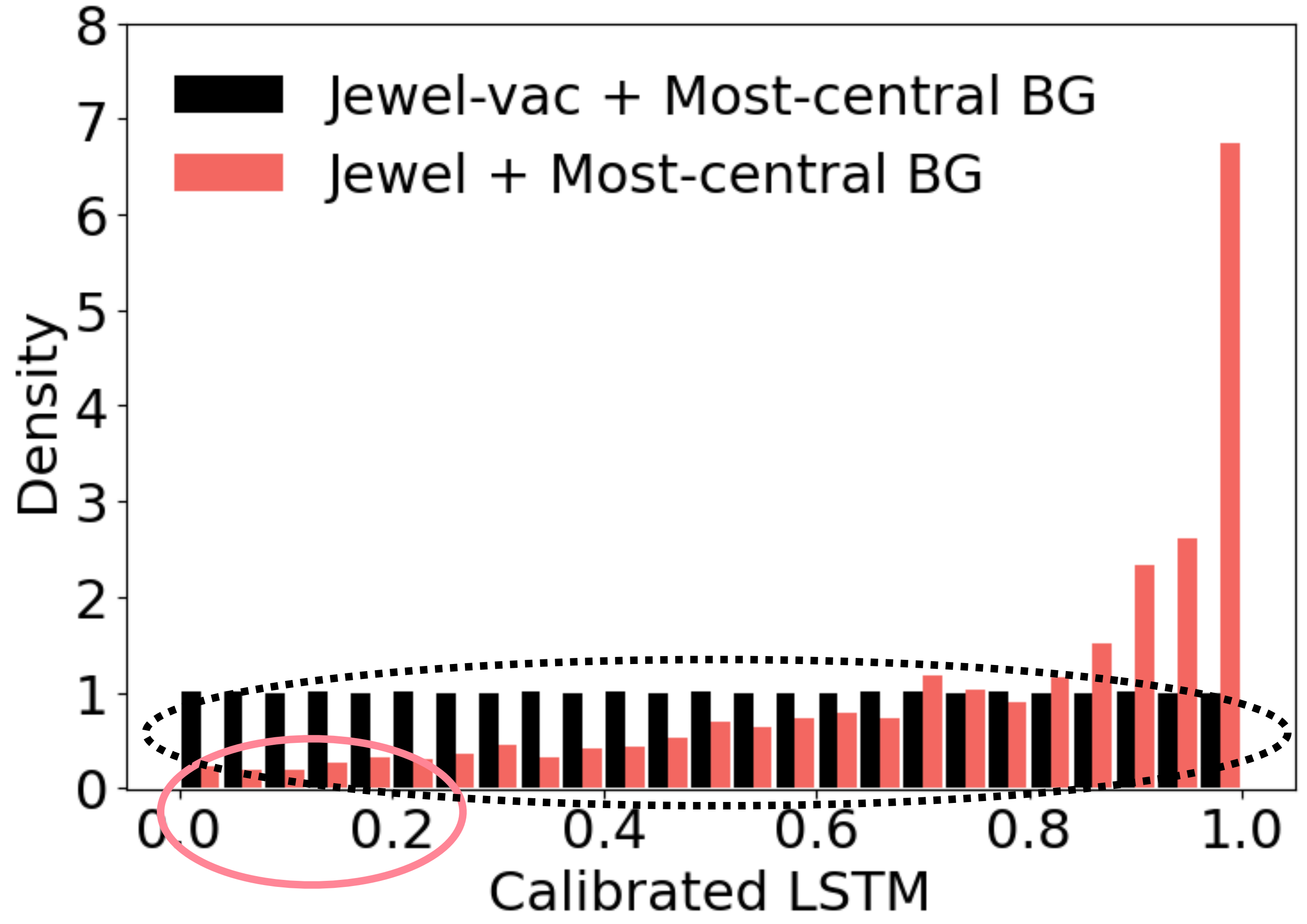
- ❖ A well trained neural network is capable to identify different quenching levels, on a jet-by-jet base.
  - ✓ From the jet substructure perspective, there is similarity between vacuum jets and quenched jets.
  - ✓ If we are able to prove the correlation between the “jet quenchness” predicted by the LSTM output and the energy loss, it may be a better way to study the medium-induced modification than the averaging jet observables.
- ❖ We are working on training the neural network(NN) with experimental data, in order to make sure the NN doesn’t learn from the algorithm differences between event generators.

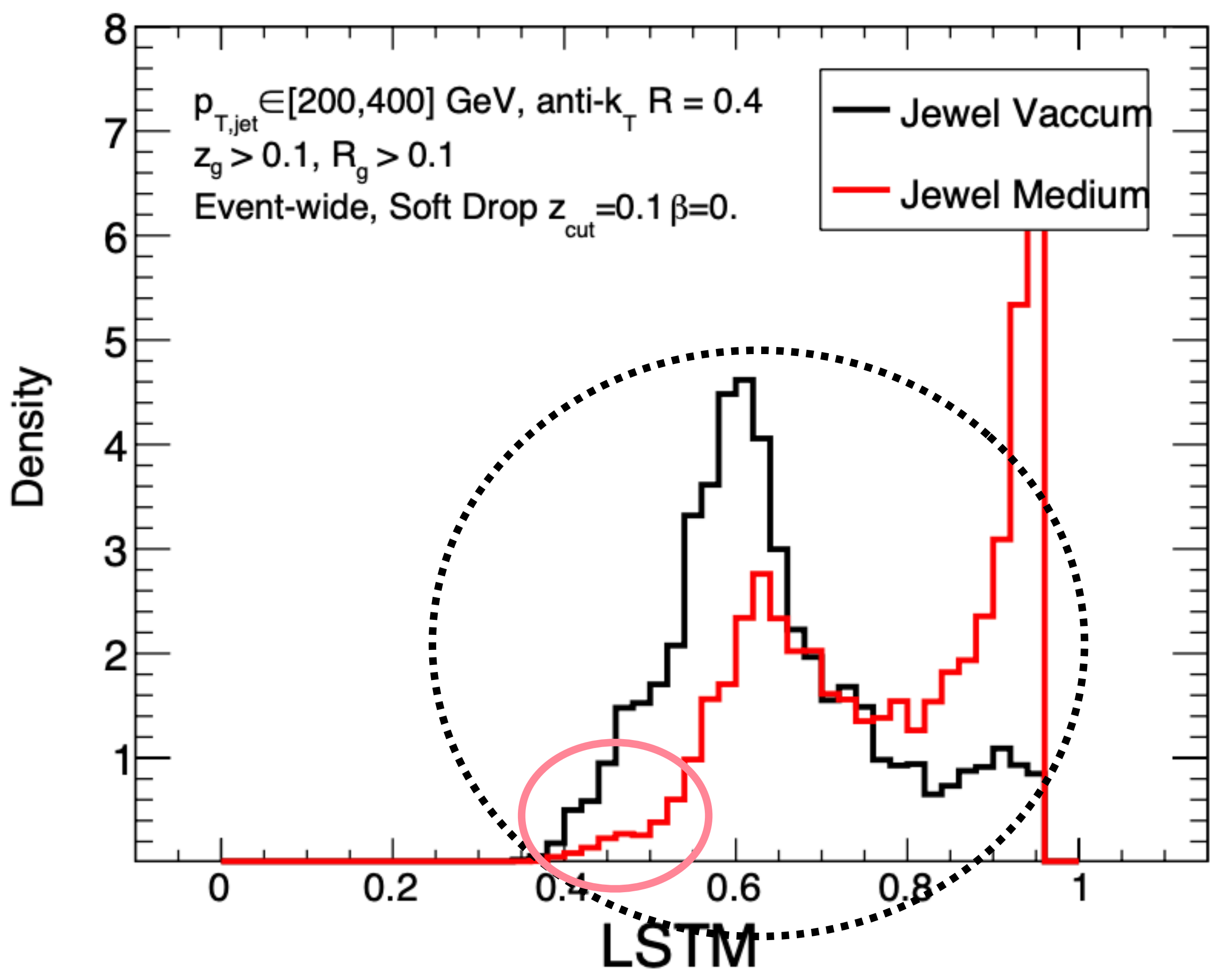
Backups

Biased Vacuum Jets



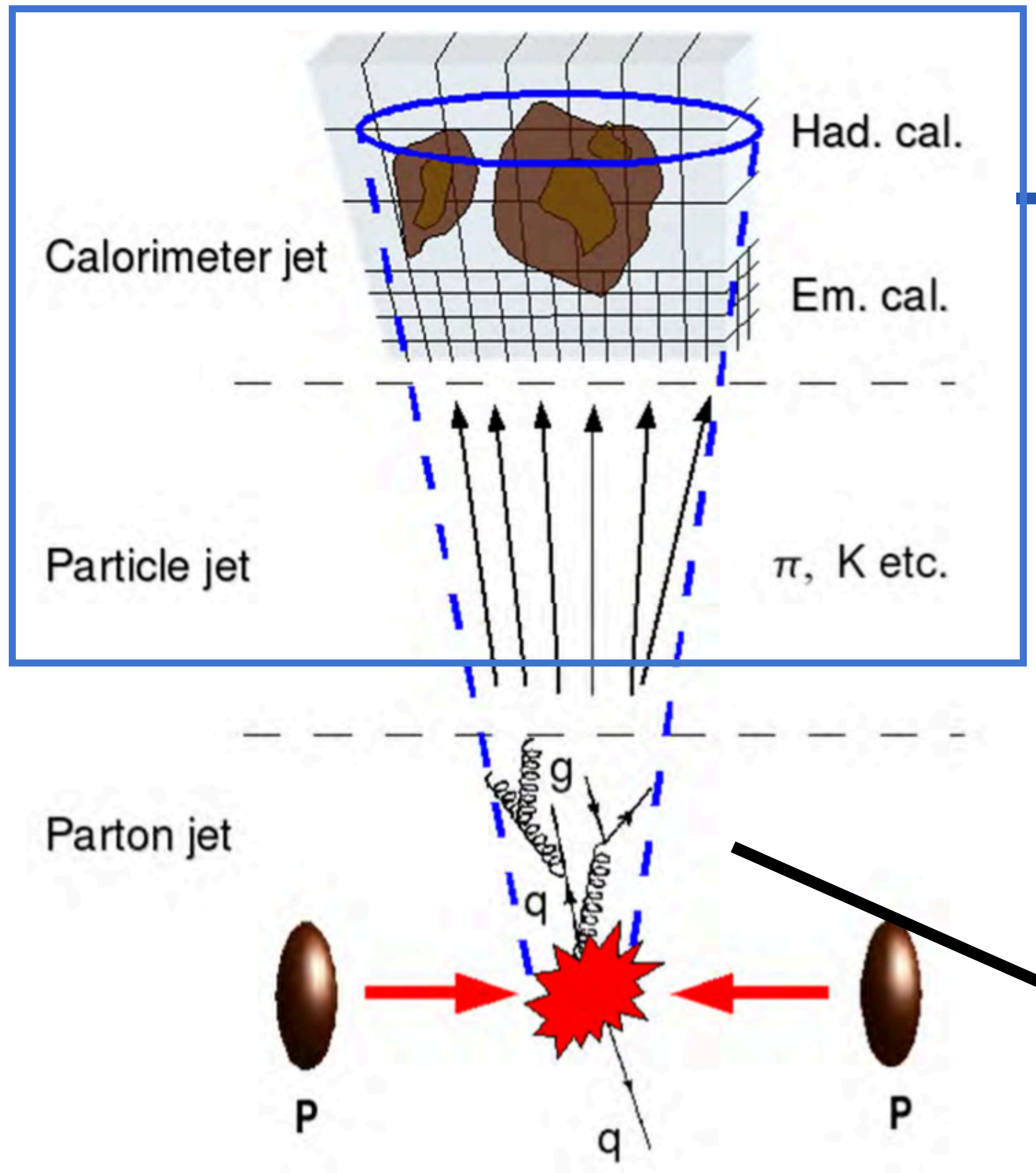




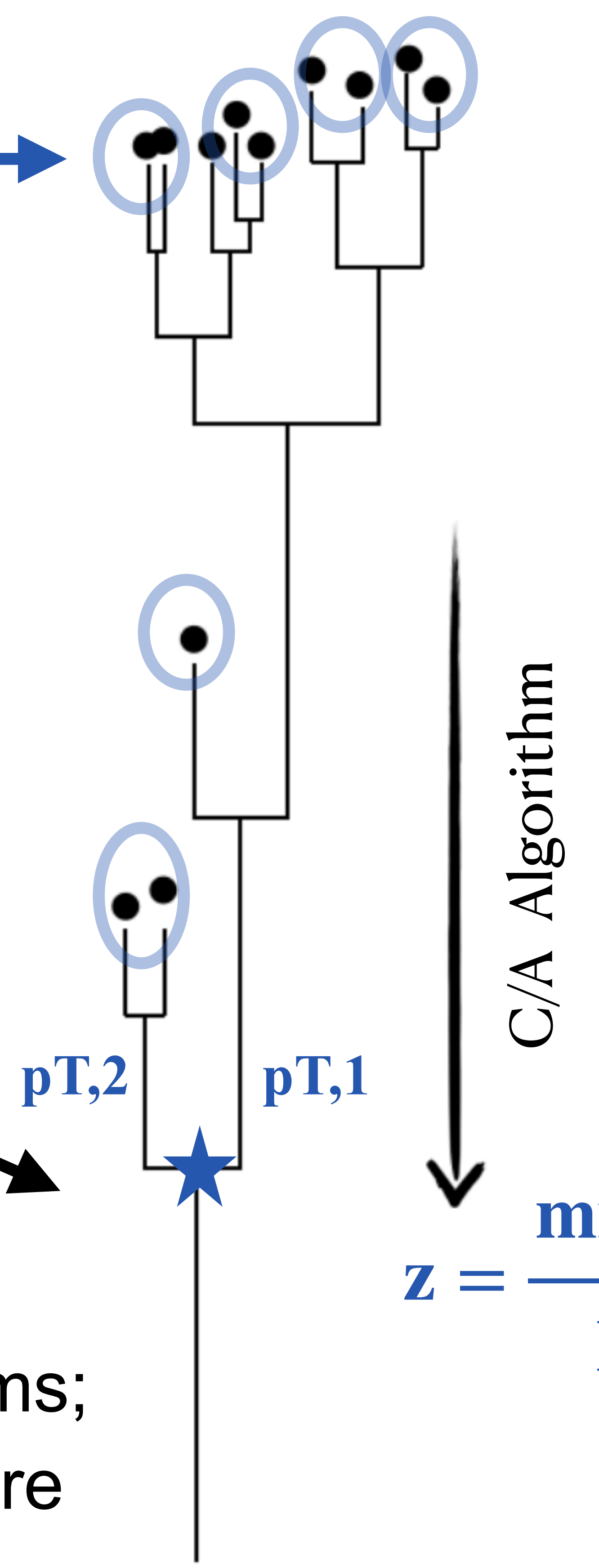


Measurable Observables: Jet Substructures

Jets in detectors



Jet in a binary tree structure



$$z = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}}$$

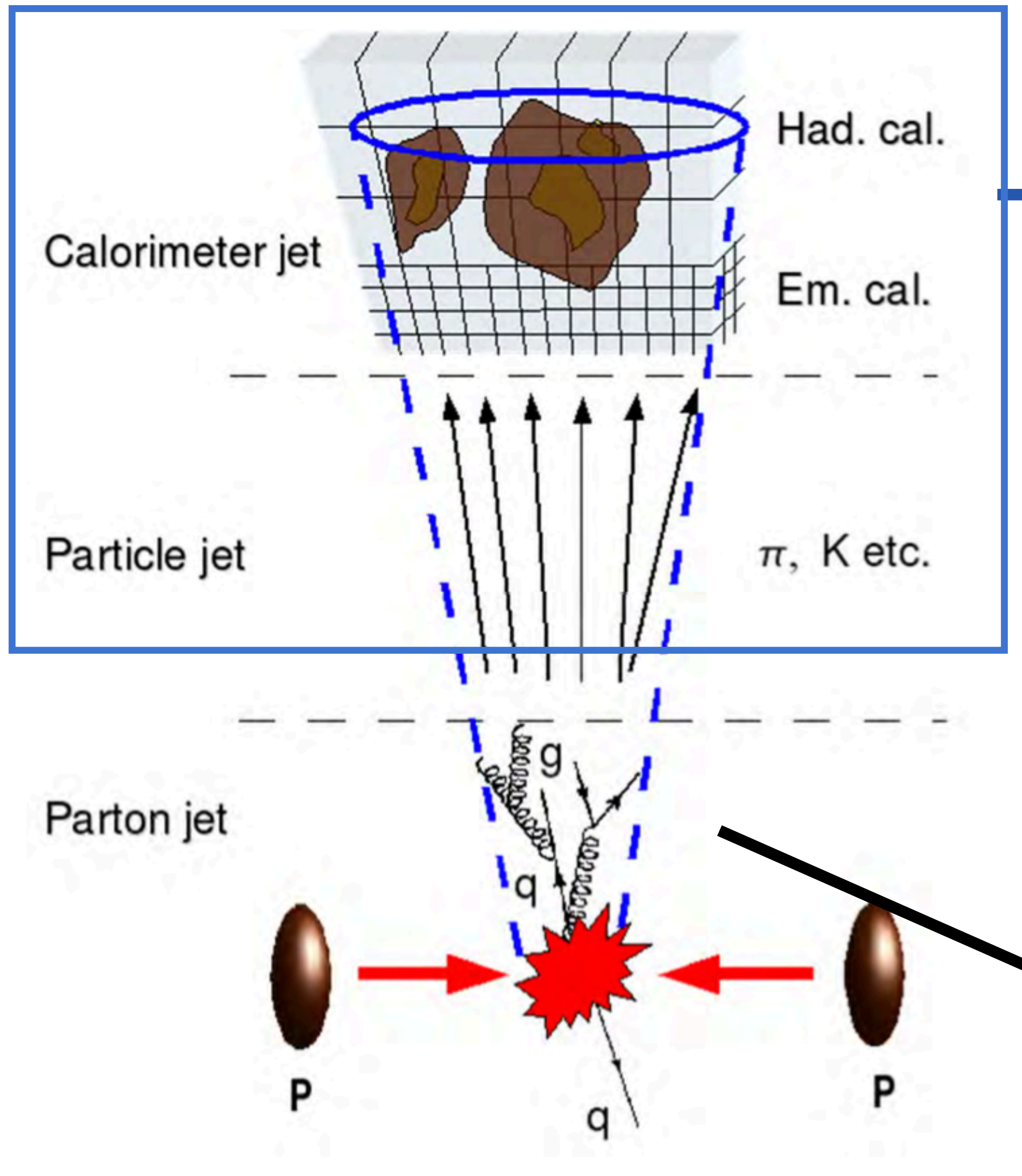
1st step: each time cluster the two closest items; eventually get the binary tree structure

2nd step: use the soft drop to discard the softer splitting of the two branches

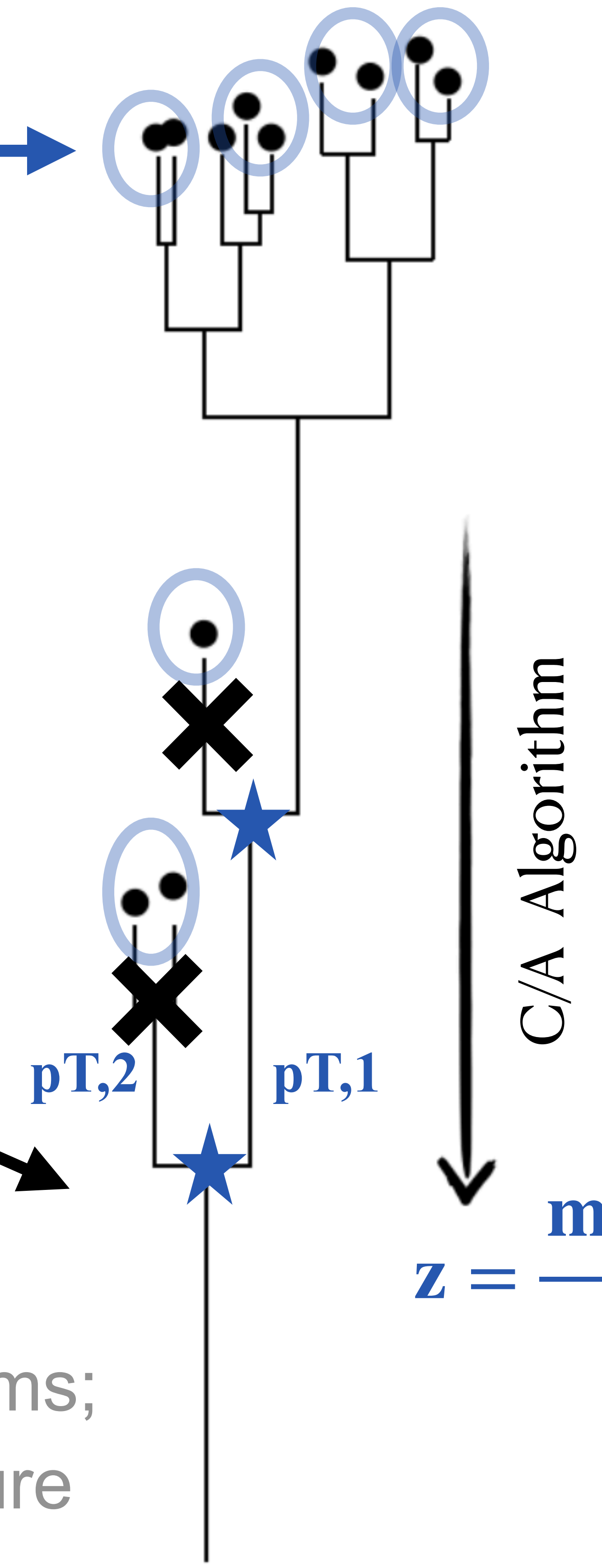
Jets are the collimated bunches of hadrons measured in our detectors

Measurable Observables: Jet Substructures

Jets in detectors



Jet in a binary tree structure



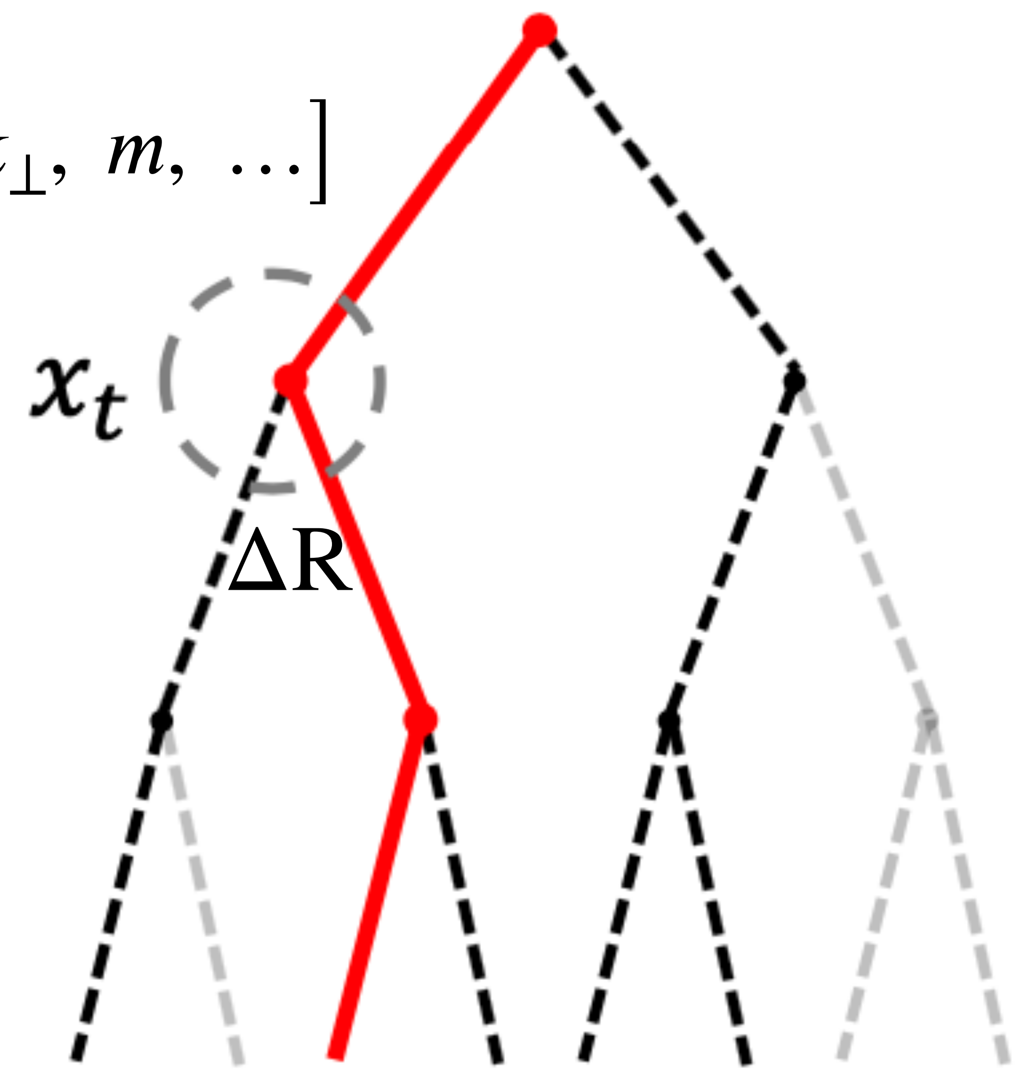
$$z = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}}$$

1st step: each time cluster the two closest items; eventually get the binary tree structure

2nd step: use the soft drop to discard the softer splitting of the two branches

Hardest branch of the jet

$$x_t = [z,\ \Delta R, k_\perp,\ m,\ \ldots]$$

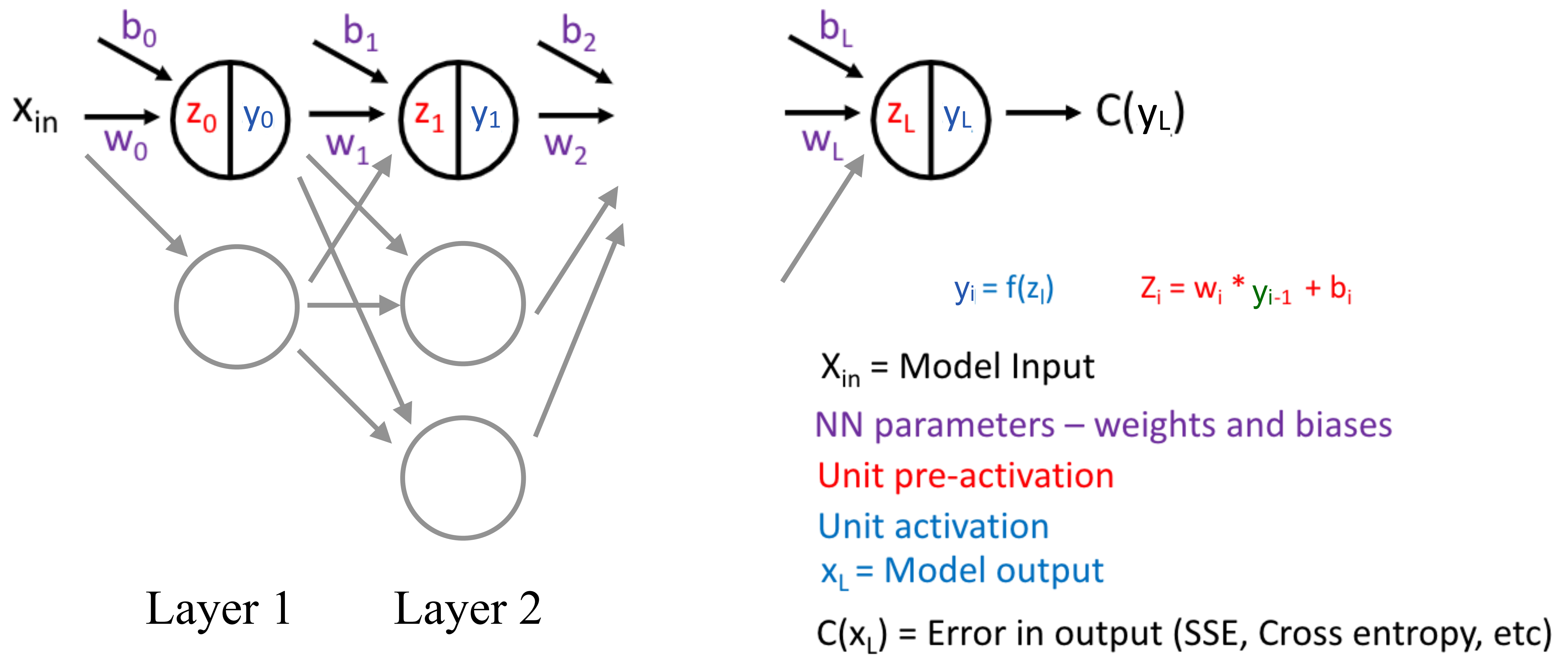


Jet substructure variables are defined at the splitting points of the jet. They are sensitive to jet-induced medium response. Thus, they are good tools to study the jet energy loss in medium

Neural Network and Feature Engineering

```
space = hp.choice('hyper_parameters',[
    {
        'size_batch': hp.quniform('size_batch', 2000, 10000, 1000),
        'num_epochs': hp.quniform('num_epochs', 30, 50, 5),
        'num_layers': hp.quniform('num_layers', 2, 4, 1),
        'Hidden_size 0': hp.quniform('hidden_size0', 8, 20, 2),
        'hidden_size1': hp.quniform('hidden_size1', 4, 8, 2),
        'learning_rate': hp.uniform('learning_rate', 0.01, 0.05),
        'decay_factor': hp.uniform('decay_factor', 0.9, 0.99),
        'loss_func' : hp.choice('loss_func', ['mse']),
    }
])
```

Hyper parameter space

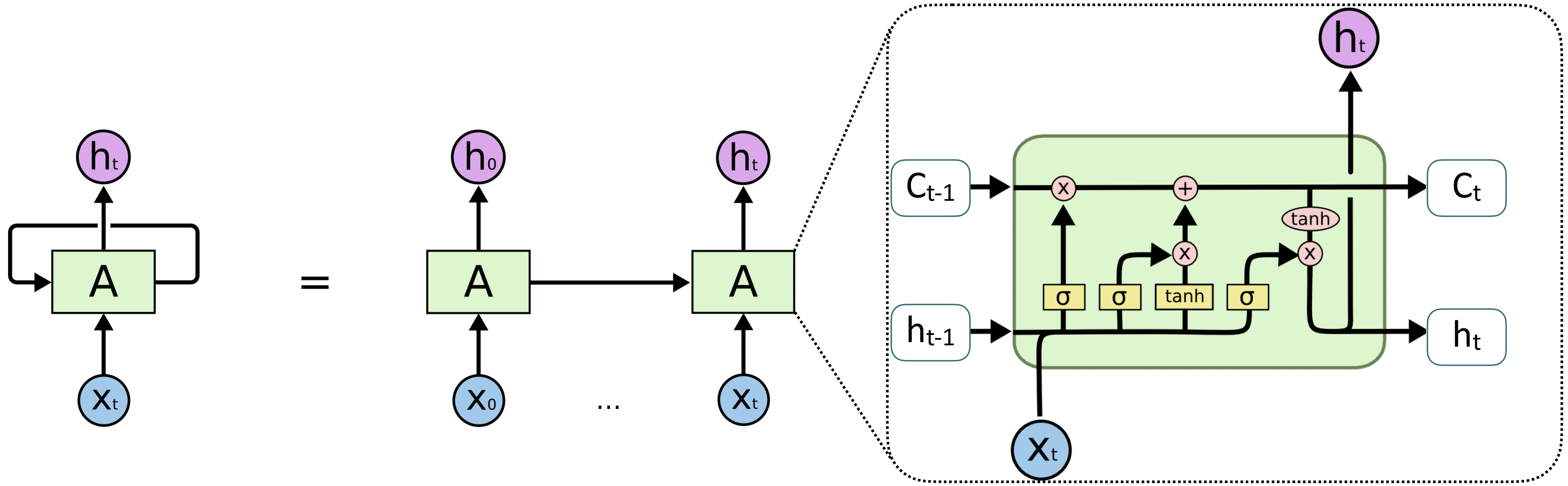


Fully Connected layers

Stacked LSTM layers + 2 full-connect layers. Output of the last step from the top LSTM layer is directed to two full-connect layers.

Both the input and output dimensions of the first full-connect layer are the hyper-parameters defining the architecture of the neural network.

* *https://arxiv.org/abs/2206.01628*



Training+Validation

Select jets from dataset to form batches:
Non-quenched jets from Jewel-vacuum
Quenched jets (Medium jets) from Jewel

Mean square error (MSE) batch loss

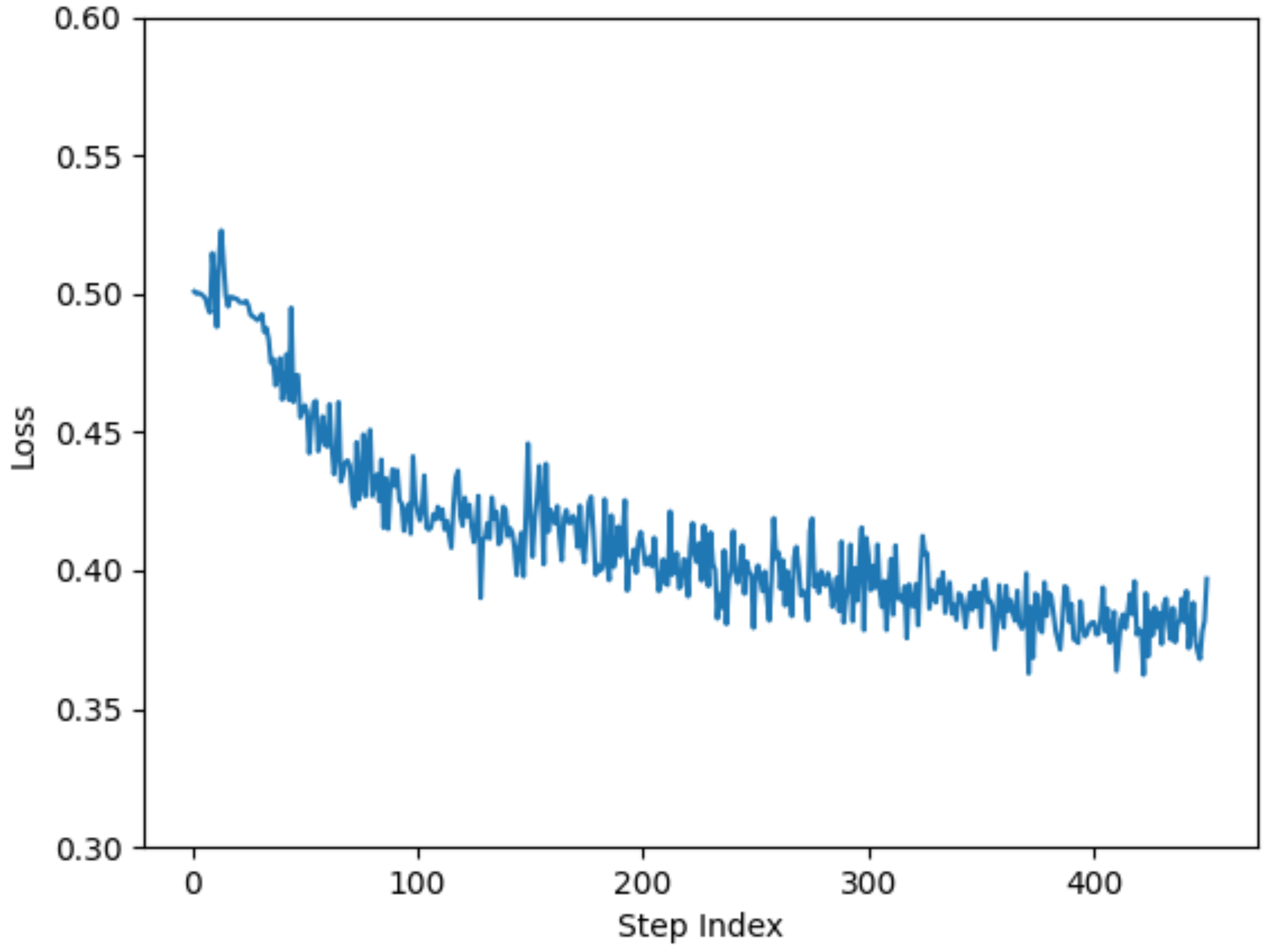
$$L = \frac{\sum_{batch} \omega_i * (x_i - y_i)^2}{\sum_{batch} \omega_i}$$

ω_i: event weight
x_i: predictive label
y_i: truth label

($\omega_i = 1$ for real experimental samples)

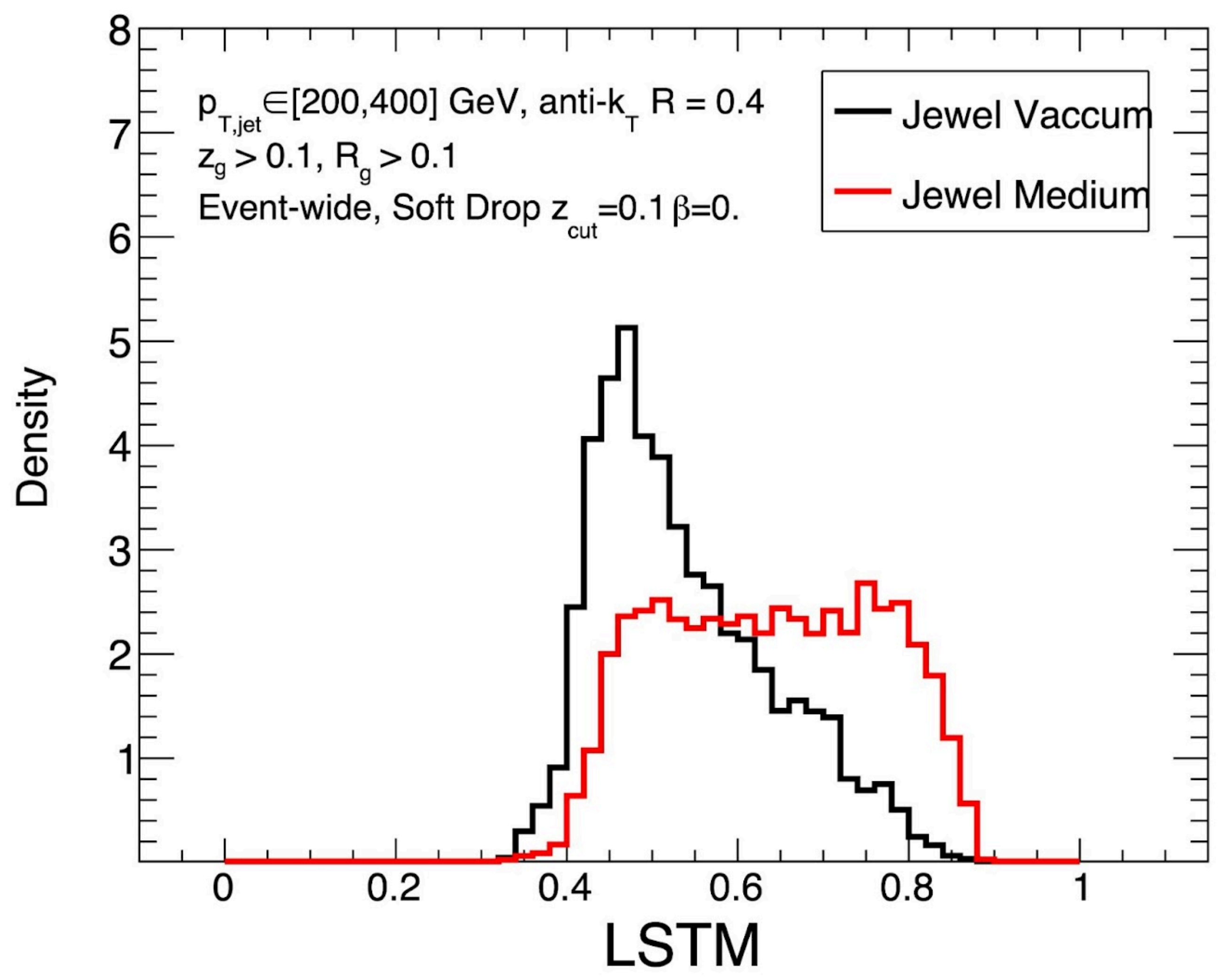
Input dataset: 200k events 200k events

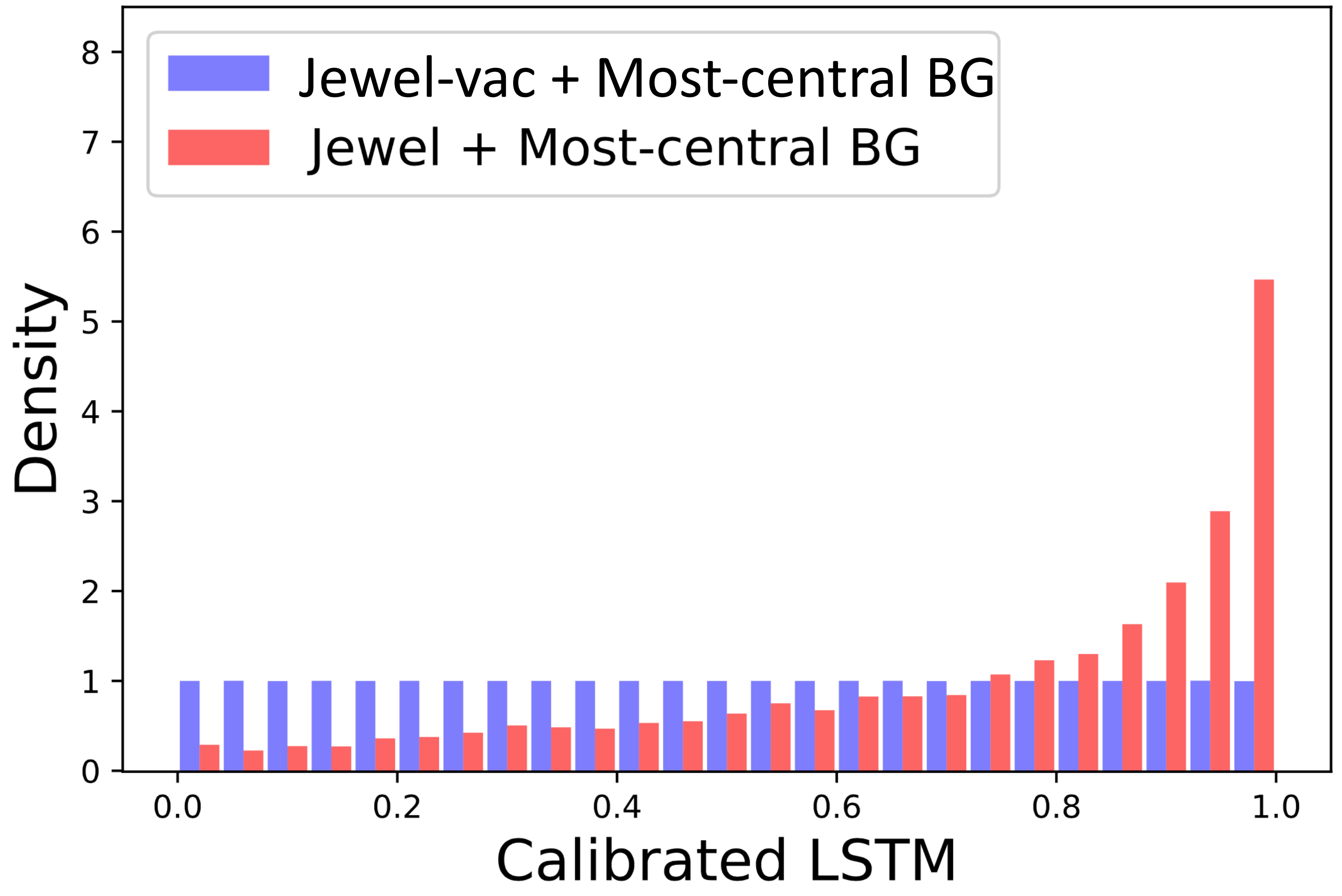
No. of Jets	Training Set (w/wo cuts)	Validation Set (w/wo cuts)
Non-quenched jets	**42535**/310332	**42272**/310276
Medium jets	**52954**/298675	**52967**/298876



Example of batch loss decreasing in the training

Calibration





One complete training loops epochs first. And each epoch loops over all the jet batches, but the order of jets in each batch is shuffled