Machine learning based full jet p_{T} reconstruction

Hannah Bossi in collaboration with Laura Havener, Raymond Ehlers, and Constantin Loizides

ORNL Physics Division Seminar, August 27th 2020 (Remote)

Funded by U.S. DEPARTMENT OF ENERGY

Recap of ML background estimator

Use machine learning (ML) to create a mapping to correct the jet for the background!

Jet Properties

(Including constituent properties)

$\xrightarrow{\textbf{ML}}$

Corrected Jet p_T

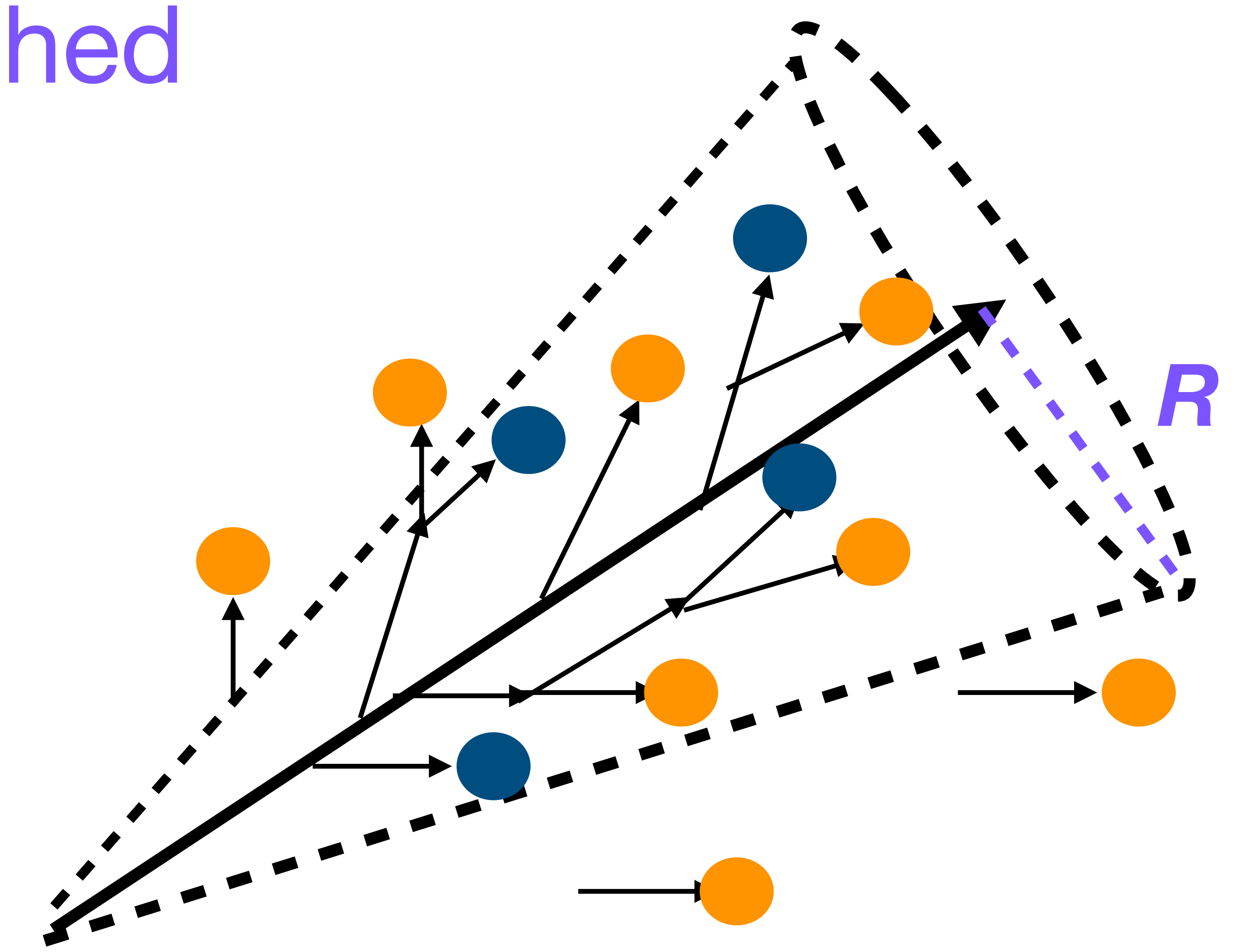
$\longrightarrow$

Unfold for fluctuations and detector effects

Reduces residual fluctuations, allowing the measurement to be pushed to lower p_T and larger R with reduced systematic uncertainties.

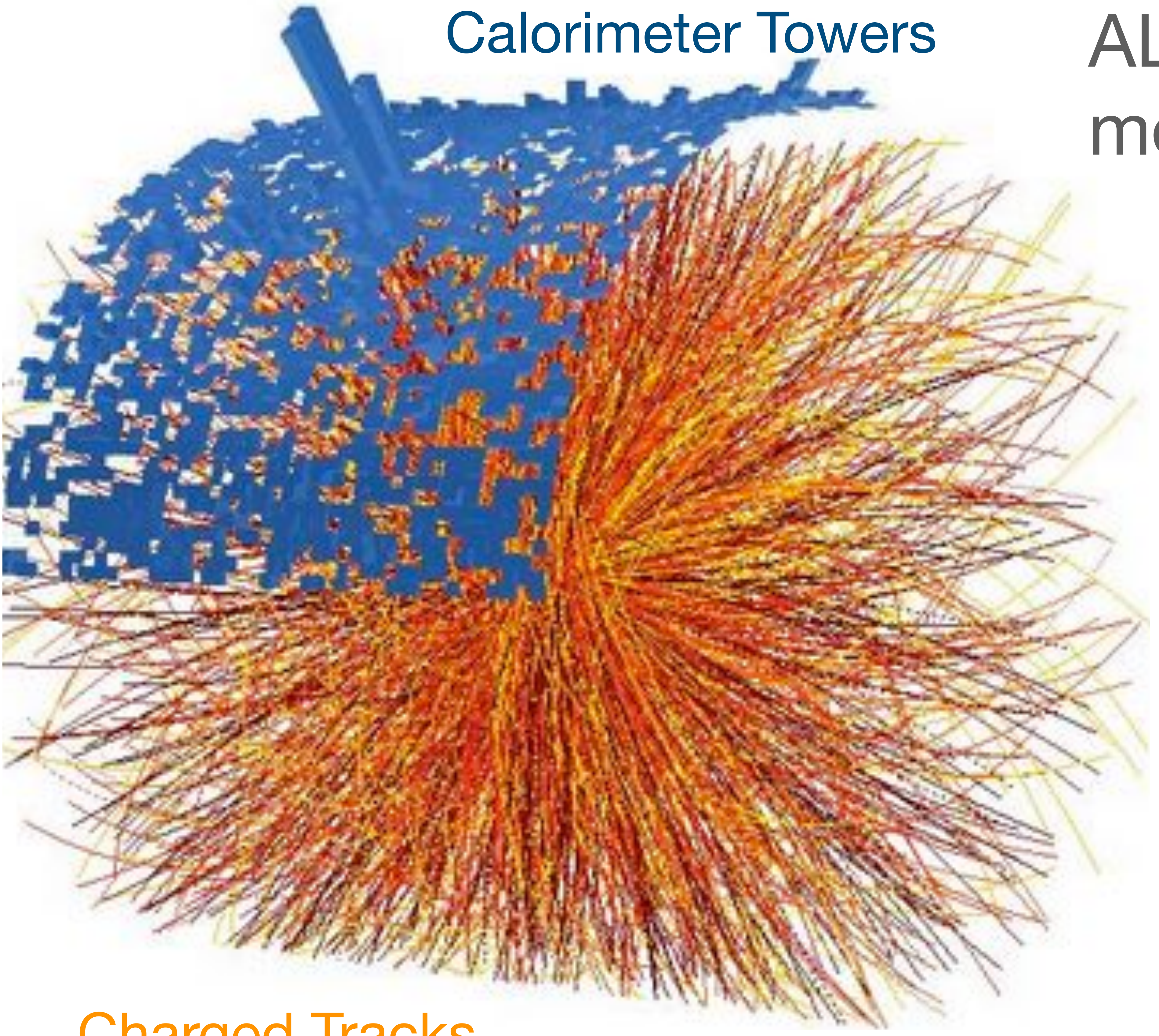
Including constituent information in training introduces a fragmentation bias which needs to be investigated.

This talk: focusing on the extension of this method to full jets and the bias studies!

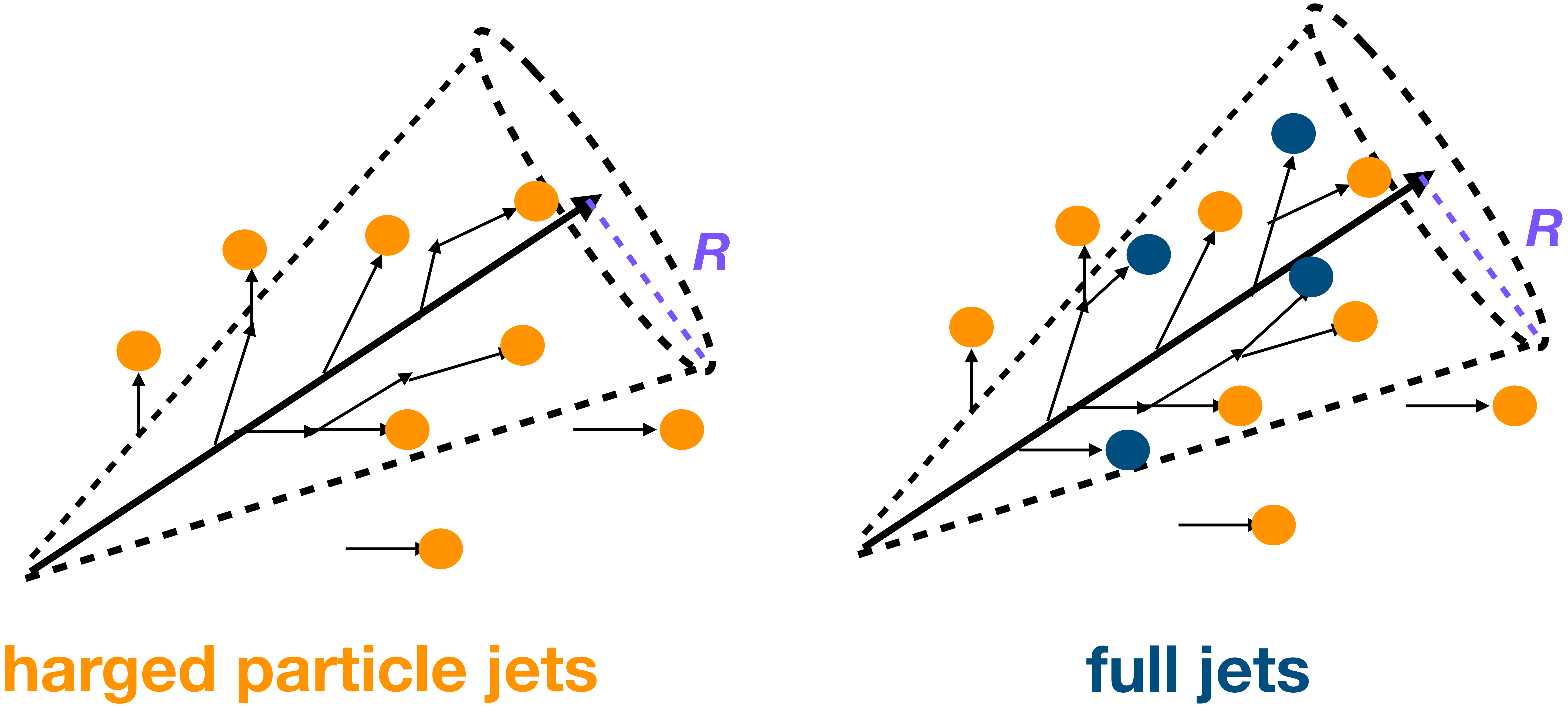


R.Haake, C. Loizides Phys. Rev. C 99, 064904 (2019)

Charged vs. full jets

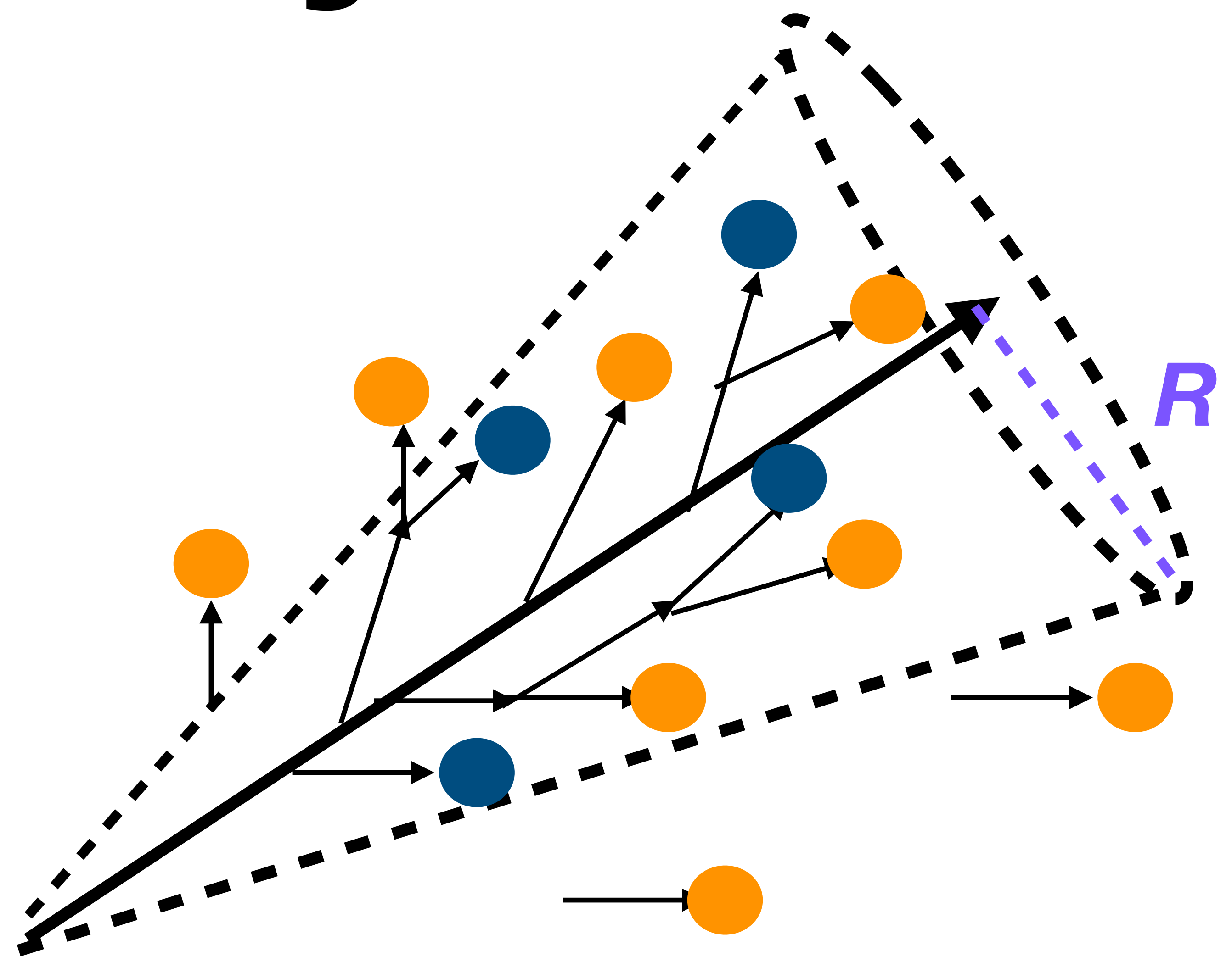


ALICE uses the electromagnetic calorimeter (EMCal) to measure neutral particles



ALICE uses the time protection chamber (TPC) and inner tracking system (ITS) to measure charged particles

Why do we want full jets?

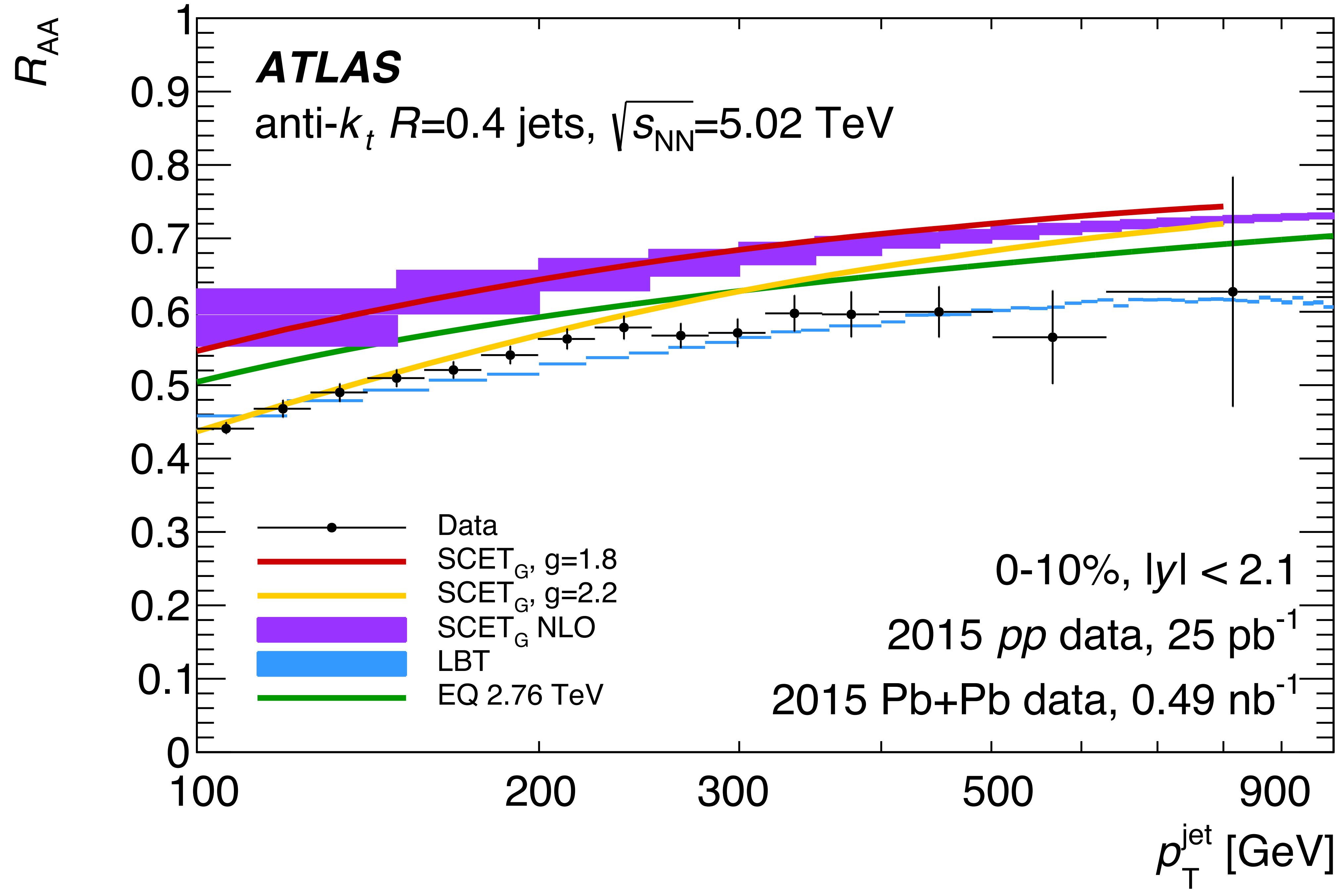


full jets

Full jets pose an additional experimental challenge by utilizing two detector systems. Why do we care?

Main Reason: Full jets allow for direct comparison to theory and other experiments!

Phys. Lett. B 790 (2019) 108



The traditional definition of a jet includes both charged and neutral particles, full jets are consistent with this definition.

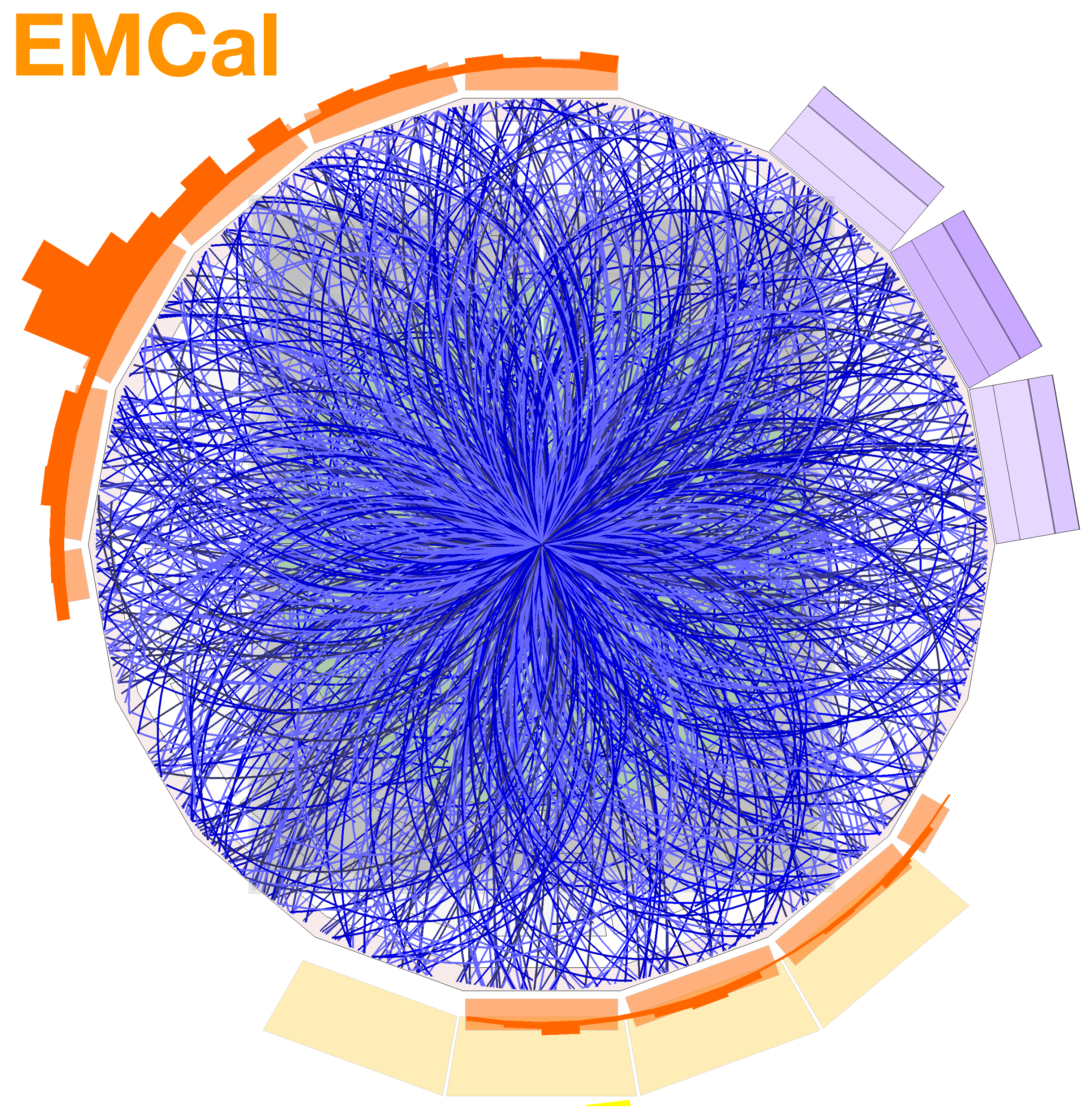
Area based method with full jets

Like for charged jets, we usually do a ρ-based correction for the background.

Our statistics for ρ limited from EMCal coverage.

Solution: Scale ρ_{charged} in a way that accounts for the ratio of possible constituents in full versus charged jets.

$$s(C) = \frac{(\sum p_{\mathrm{T,track}}^{\mathrm{calo}} + p_{\mathrm{T,cluster}}^{\mathrm{calo}})/A_{\mathrm{calo}}}{\sum p_{\mathrm{T,track}}^{\mathrm{TPC}}/A_{\mathrm{TPC}}}$$



Final ρ is then written by $\rho_{\text{full}}(c) = \rho_{\text{charged}} * s(C)$

Would we be able to improve this correction (similar to charged particle jets) with ML?

Training/testing process

Training (PYTHIA fragmentation)

Train on “hybrid event” created by embedding PYTHIA jets into Pb-Pb Background

Shallow neural network implemented in *scikit-learn*.

3 layers [100,100,50] nodes

Testing

Apply ML estimator to hybrid events not used in training.

For charged jets we used a thermal toy background. Difficult to implement for full jets, **we will use real Pb—Pb data instead**.

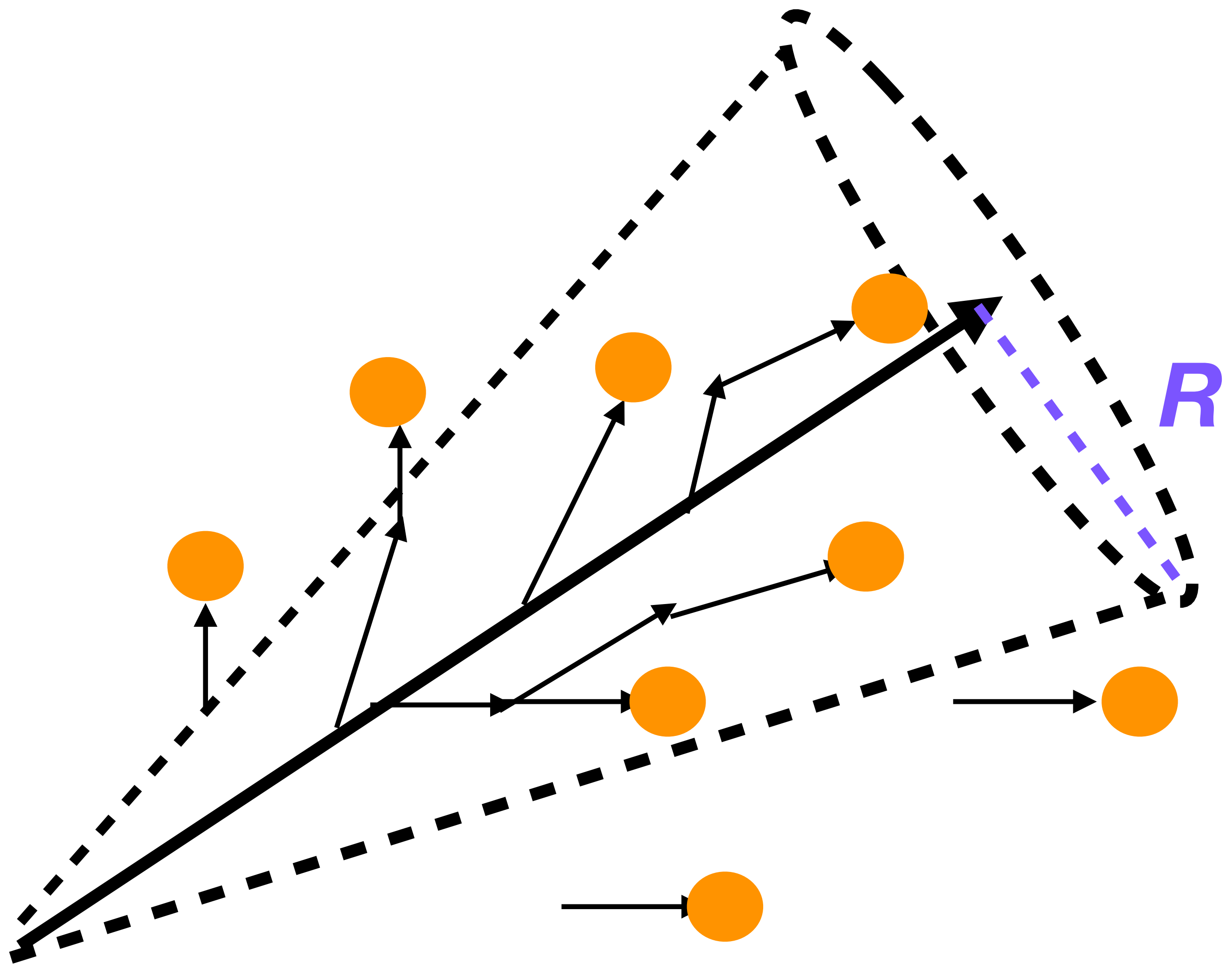
Background is realistic!

How do the features change for full jets?

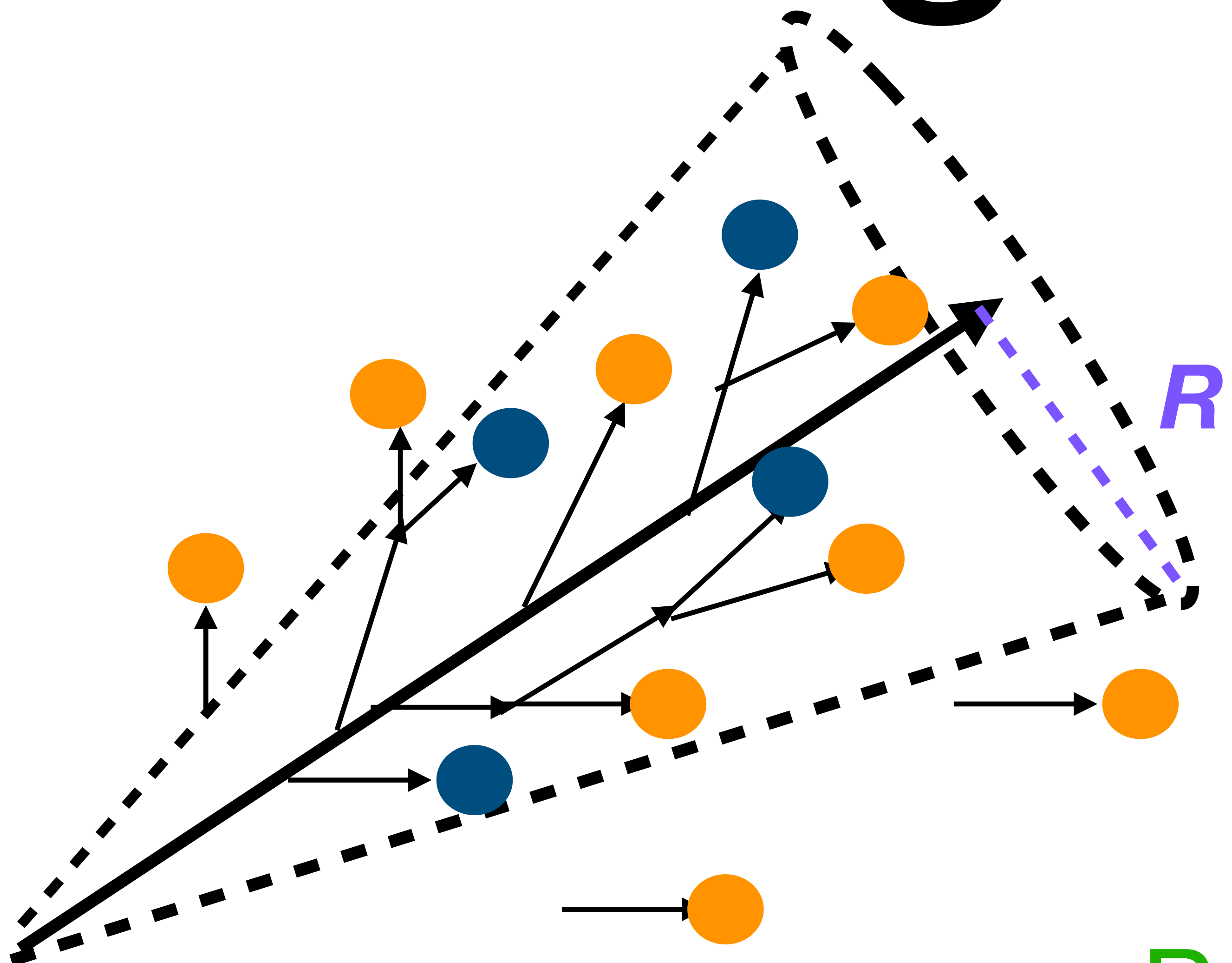
Do we get back the signal we put in?

Regression target is now **PYTHIA jet p_{T} we embedded.**

Features for training - full jets



charged particle jets



full jets

Full jets have more constituents than charged particles jets → **include more constituents in training.**

Process to choose features is the same →iteratively remove unimportant or highly correlated features.

Jet Angularity

Jet p_{T} (area-based corrected)

Number of Constituents within Jet

p_{T} of 12 Leading Constituents

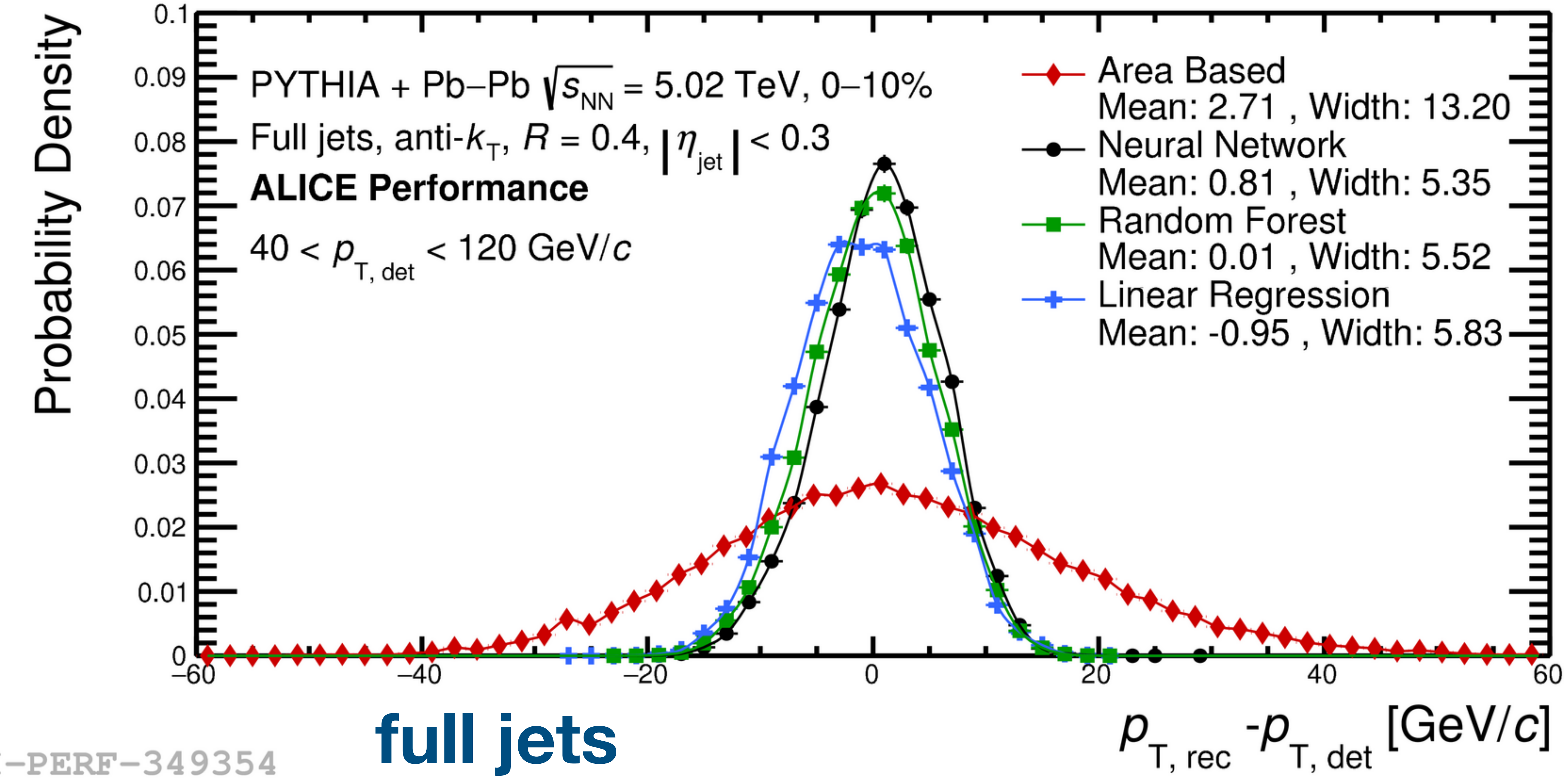
Now, constituents are either charged or neutral.

Evaluating the performance

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

$p_{\mathrm{T,rec}} \rightarrow p_{\mathrm{T}}$ predicted by ML

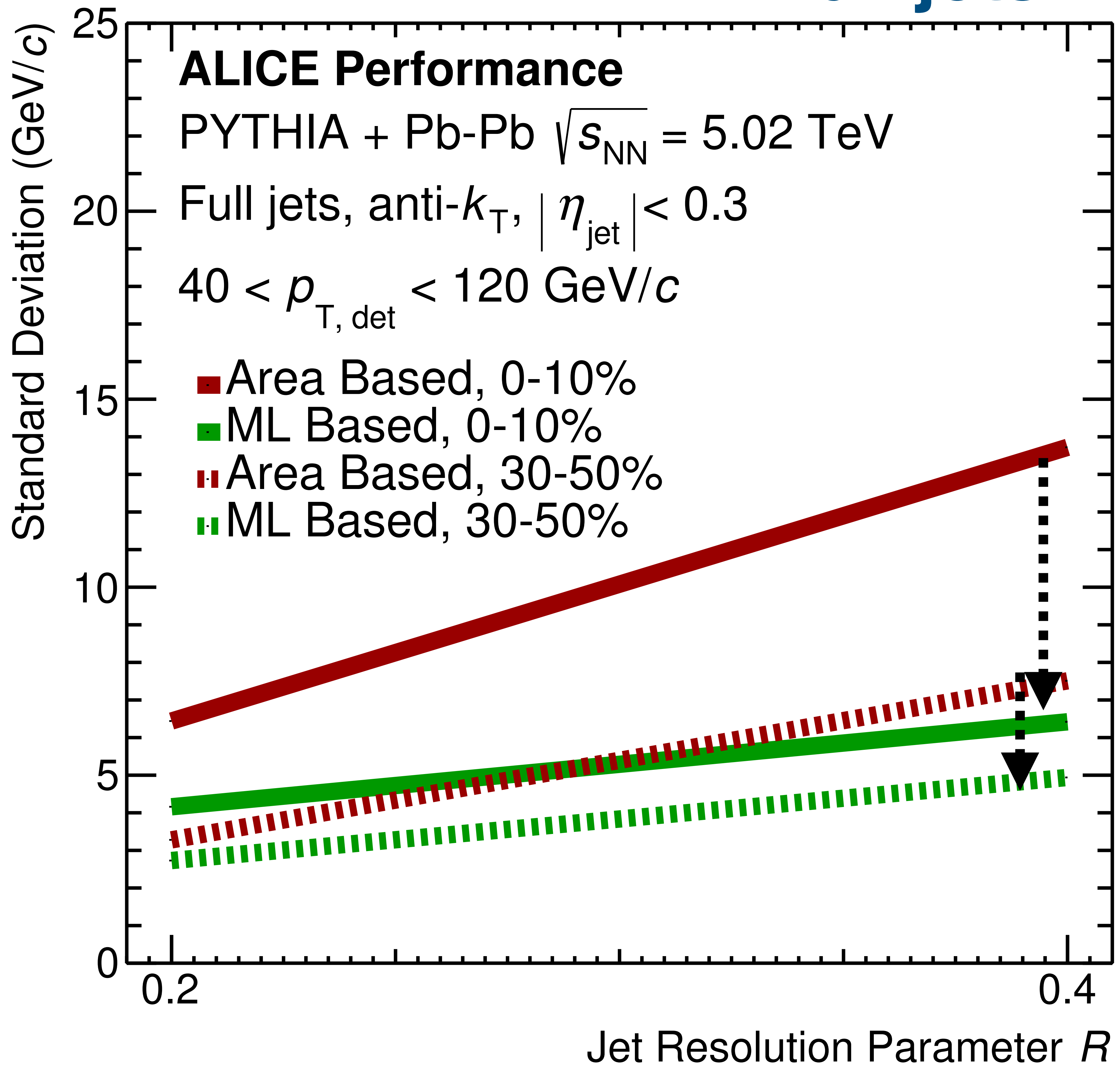
$p_{\mathrm{T,true}} \rightarrow$ PYTHIA detector level jet p_{T}



ALI-PERF-349354

full jets

full jets

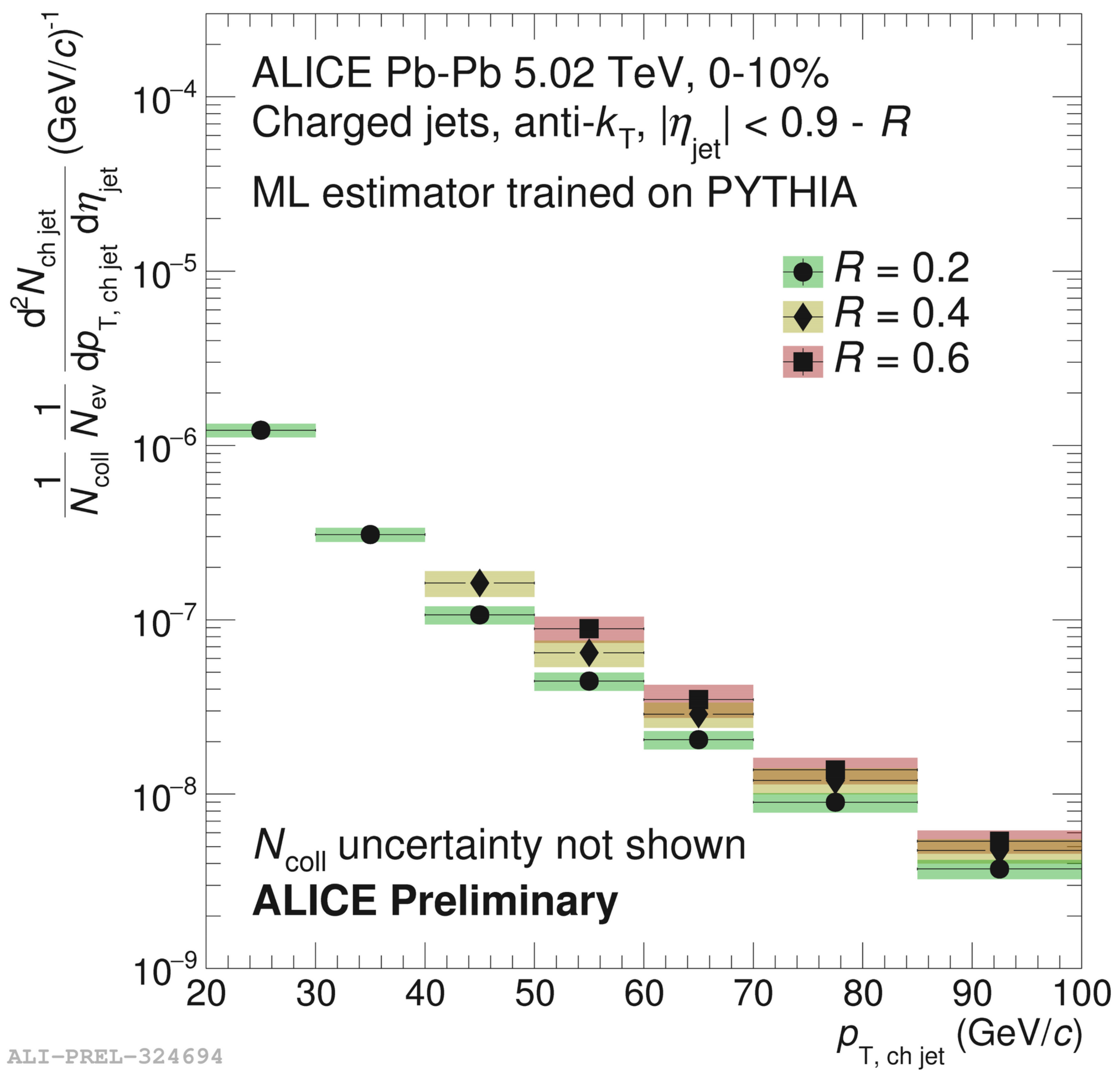


ALI-PERF-339976

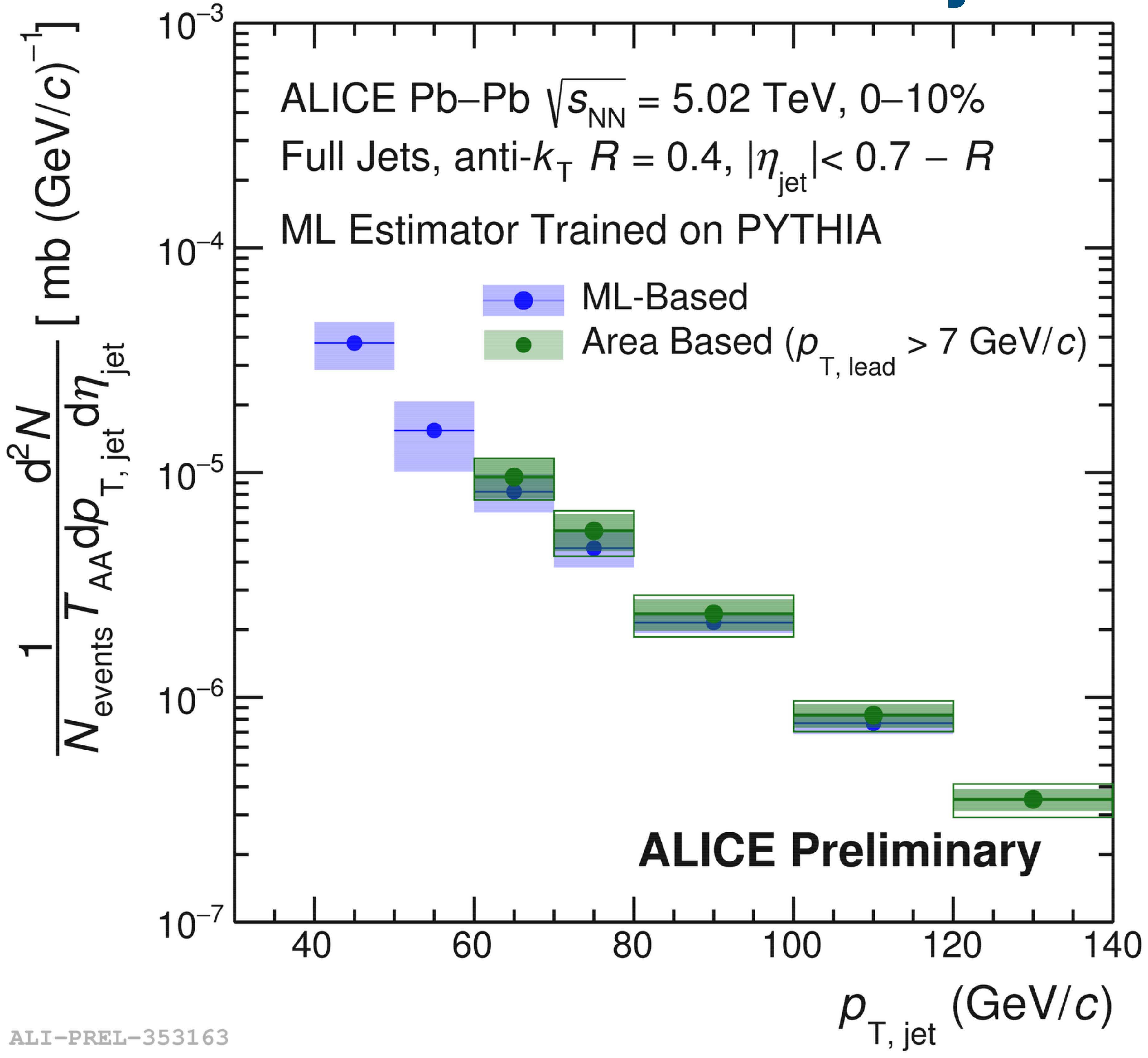
Narrow δp_{T} → Reduced residual fluctuations significantly reduced!

Results - inclusive jet spectra

charged particle jets



full jets



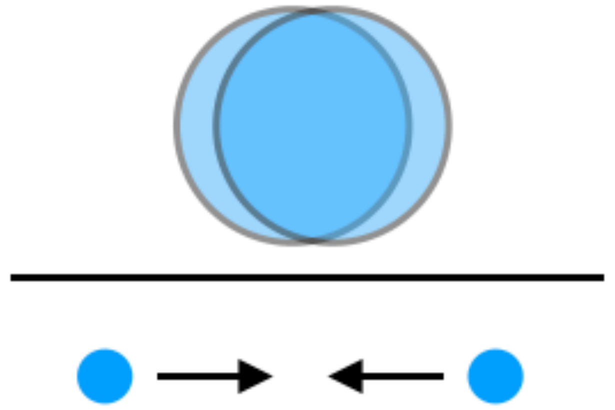
Unfolding systematics dominate at lower p_T.

Tracking efficiency systematics dominate at high p_T.

	Lower p_T Cutoff (GeV/c)	
R	Charged Particle Jets	Full Jets
0.2	20	40
0.3	50	60
0.4	40	40
0.6	50	N/A

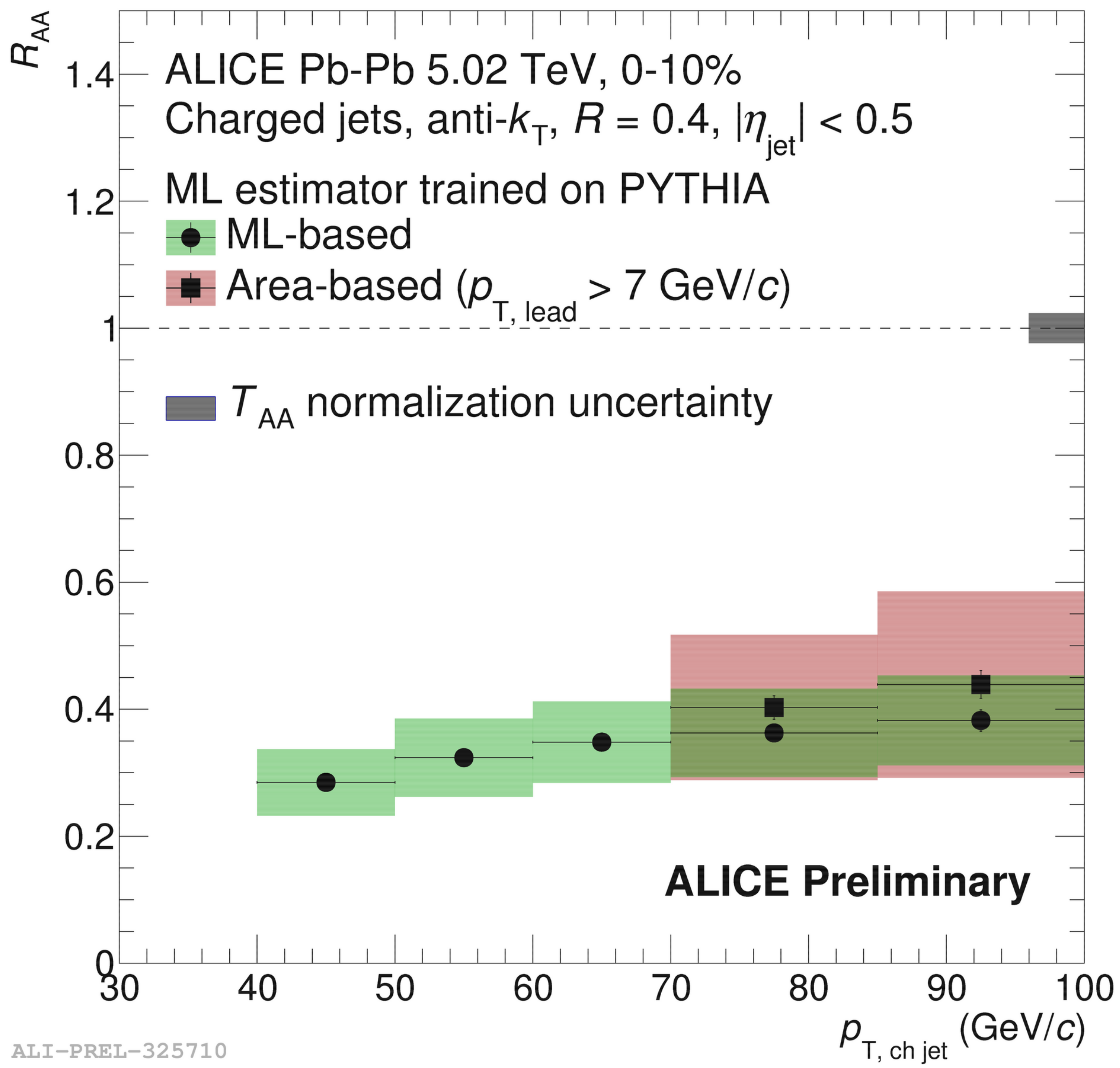
Able to extend measurements to lower p_T and larger *R*!

Results - jet R_{AA}

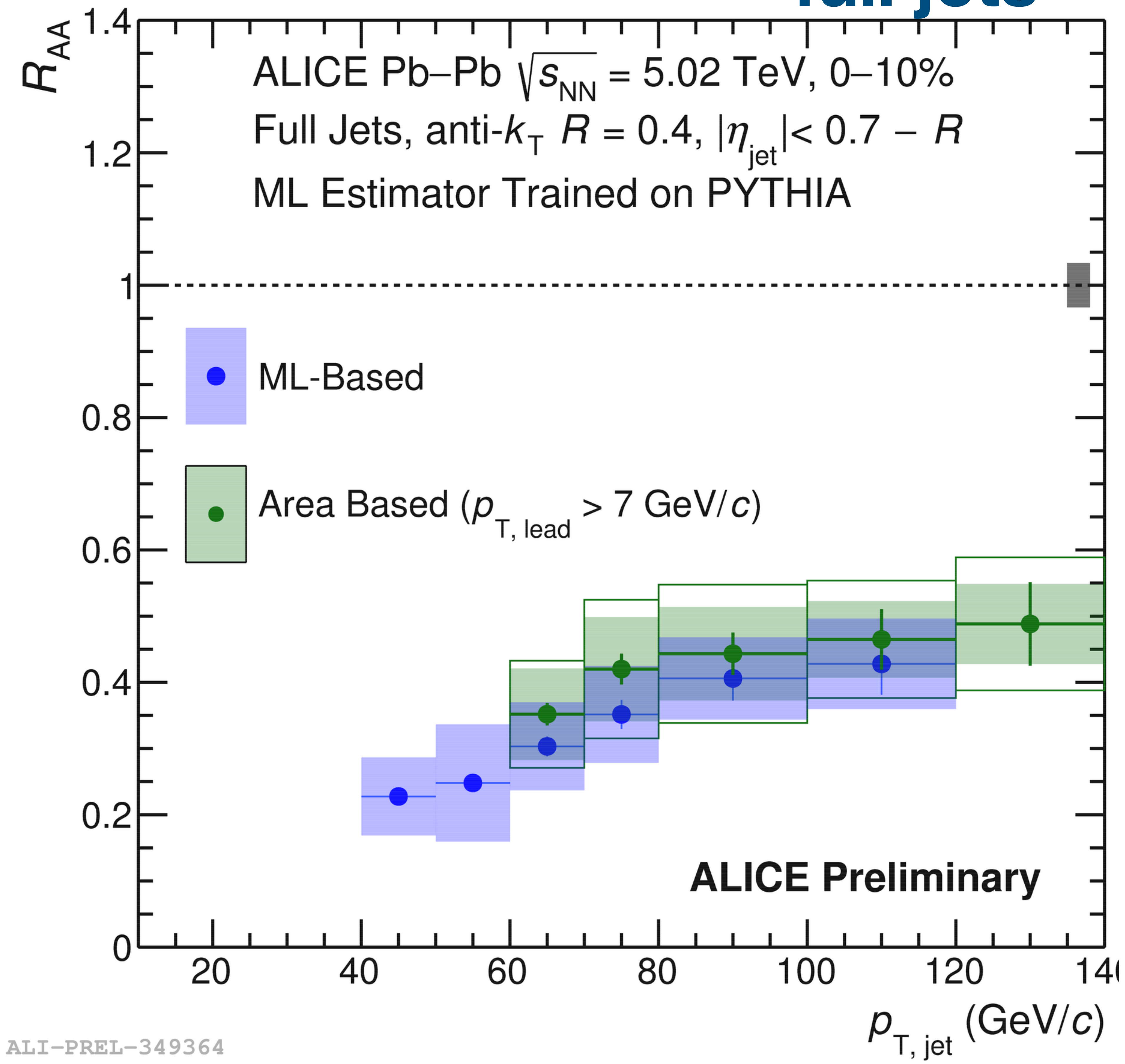


$$R_{\mathrm{AA}} = \frac{\left.\frac{1}{N_{\mathrm{event}}}\frac{d^2N^{\mathrm{PbPb}}_{\mathrm{jet}}}{dp_{\mathrm{T}}dy}\right|_{\mathrm{cent}}}{\langle T_{\mathrm{AA}}\rangle \frac{d^2\sigma^{\mathrm{pp}}_{\mathrm{jet}}}{dp_{\mathrm{T}}dy}}$$

charged particle jets



full jets

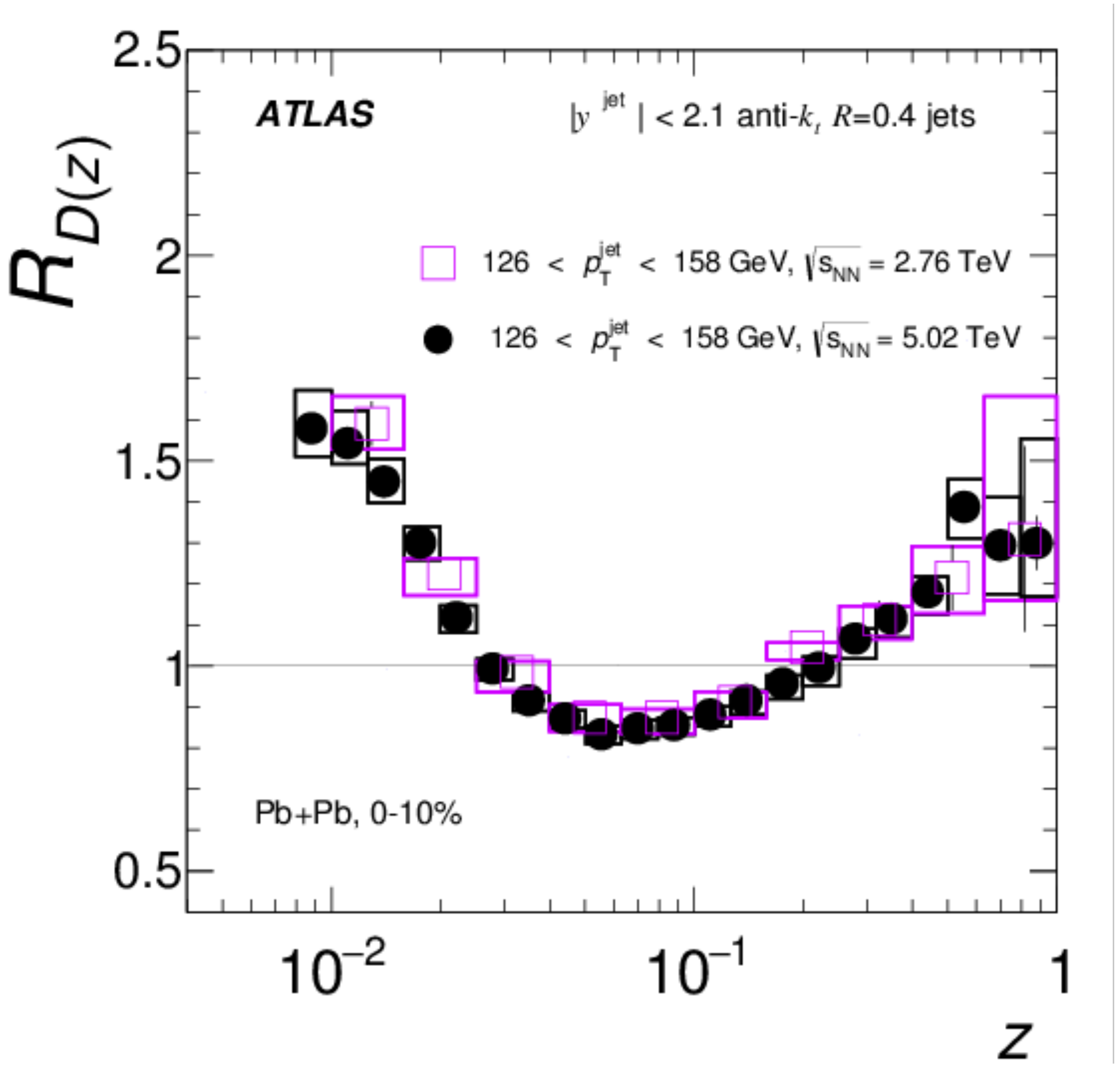


See significant jet suppression down to 40 GeV/c!

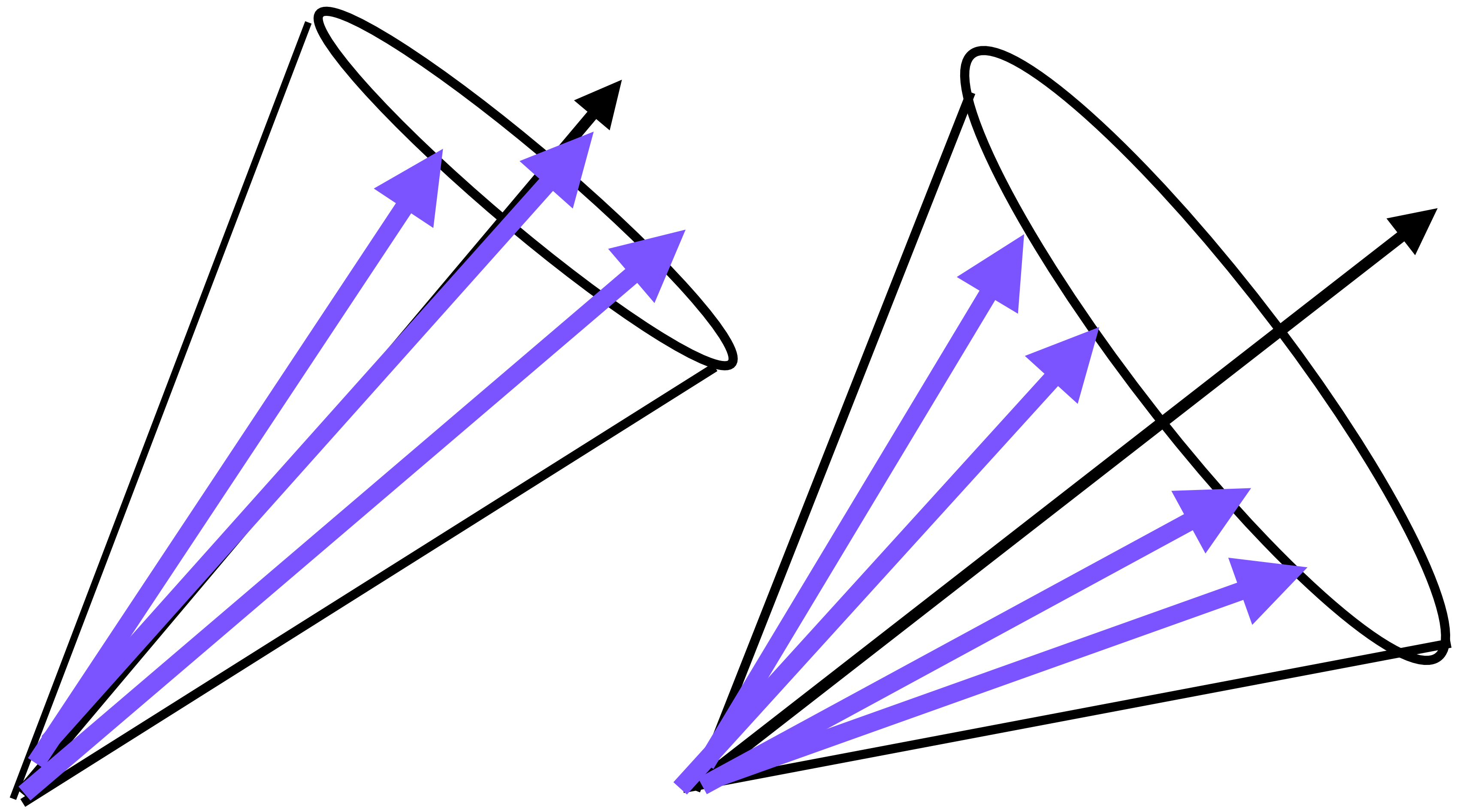
Systematic uncertainties are reduced.

Fragmentation bias

Learning on constituents introduces a *fragmentation* bias.



Phys. Rev. C 98, 024908 (2018)



We learn on a PYTHIA fragmentation.

We know that fragmentation in heavy-ion collisions is modified by the presence of the medium.

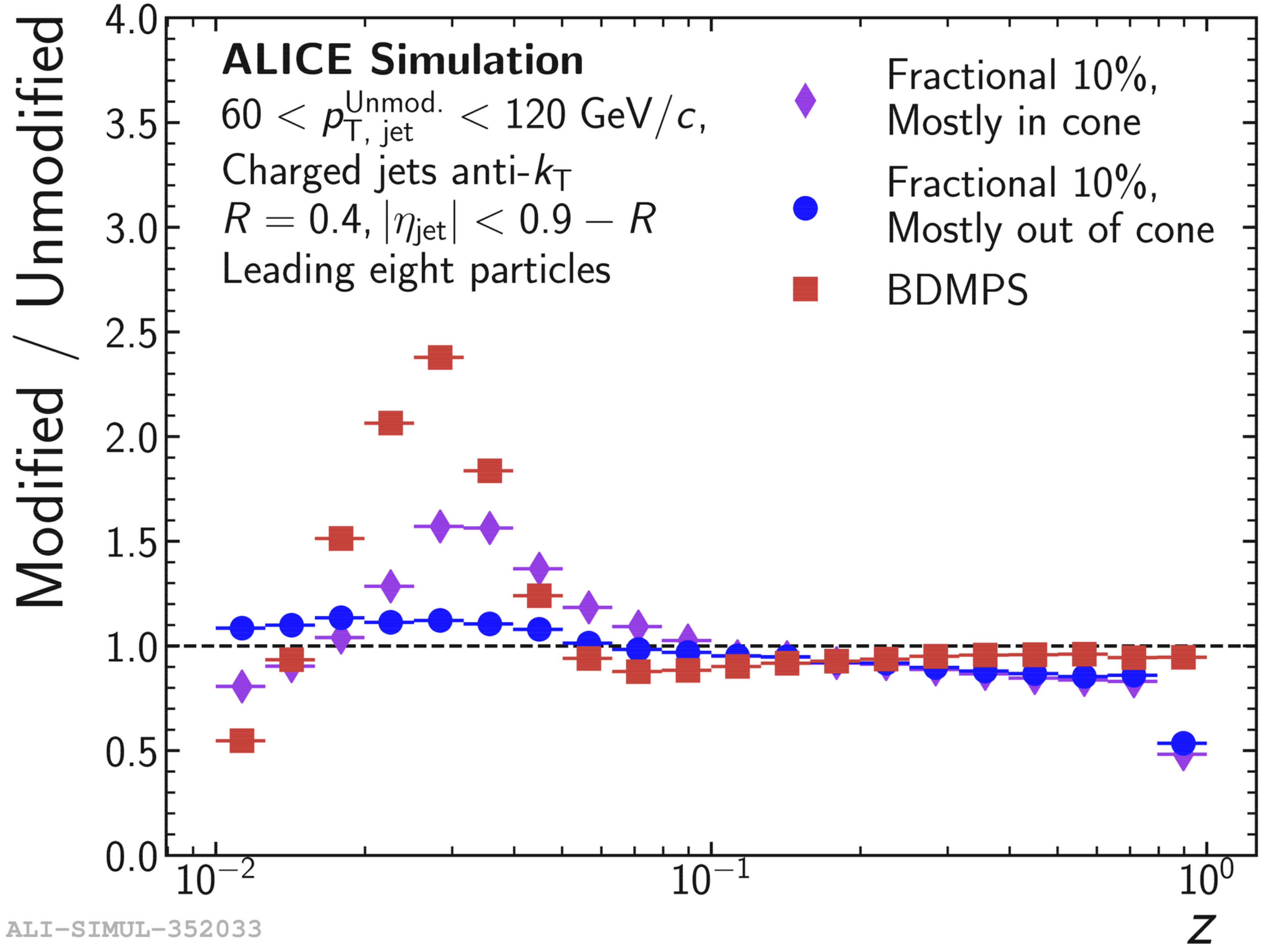
We want to investigate how this impacts the final result we get with ML!

Toy model studies

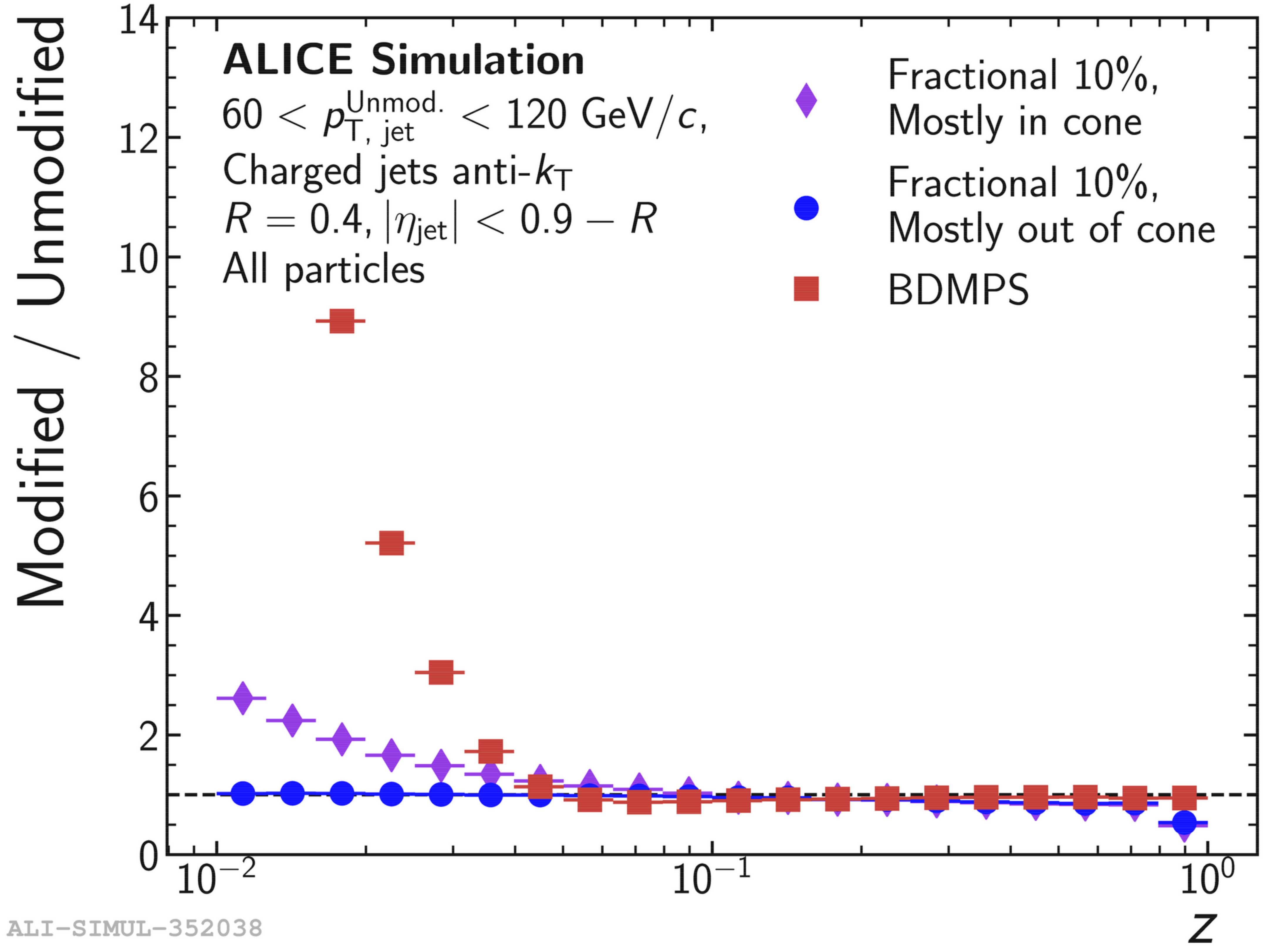
We study a toy model with three different ways to alter constituents of the jet, changing the fragmentation.

Modification to the fragmentation function

Leading 8 particles



Inclusive particles



Toy model modifications indeed modify the fragmentation, some modifications are more extreme than others.

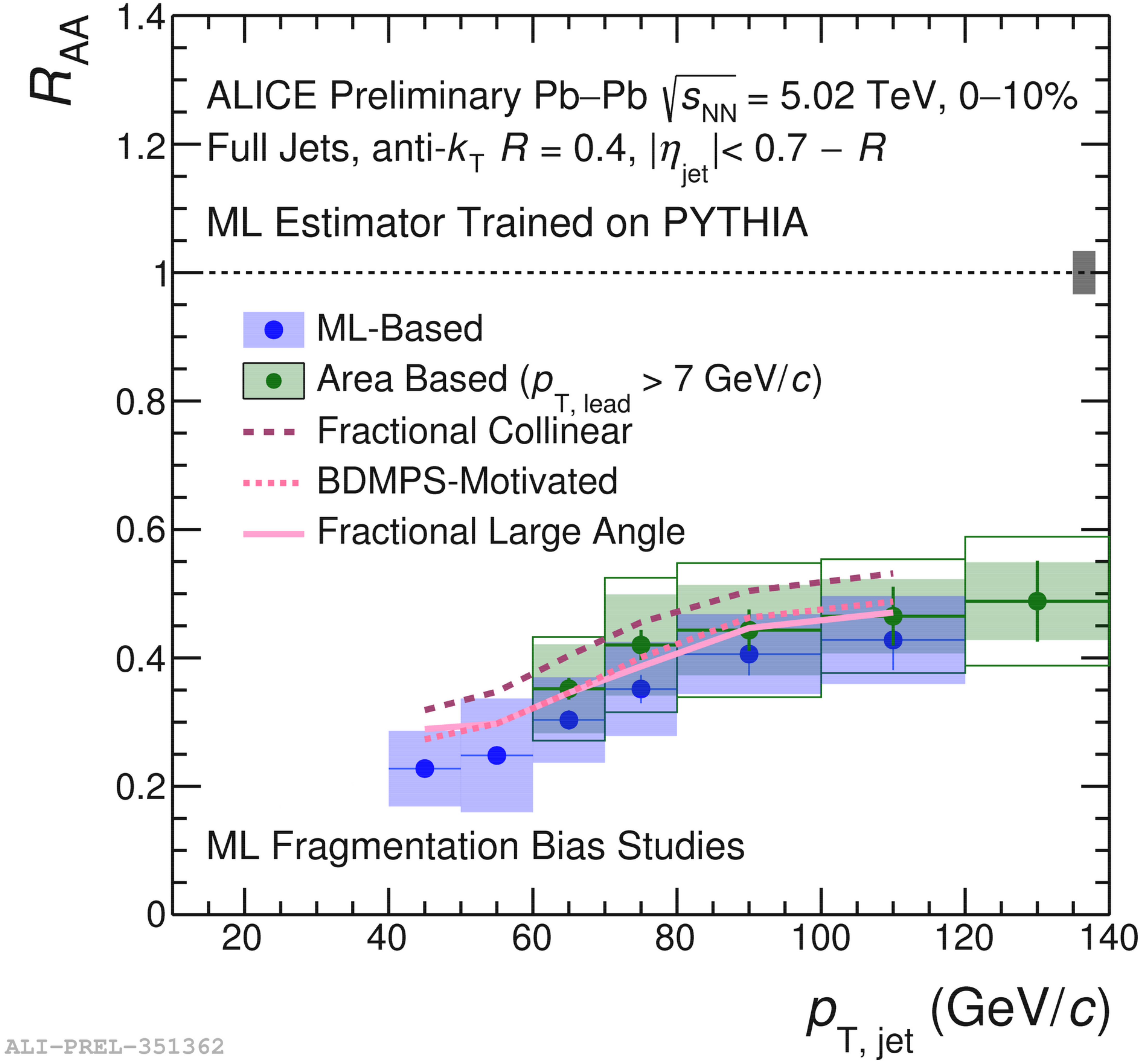
8 leading particles are what we chose to train on.

Illustration of potential bias

Train on the modified toy model and apply to data; measure bias.

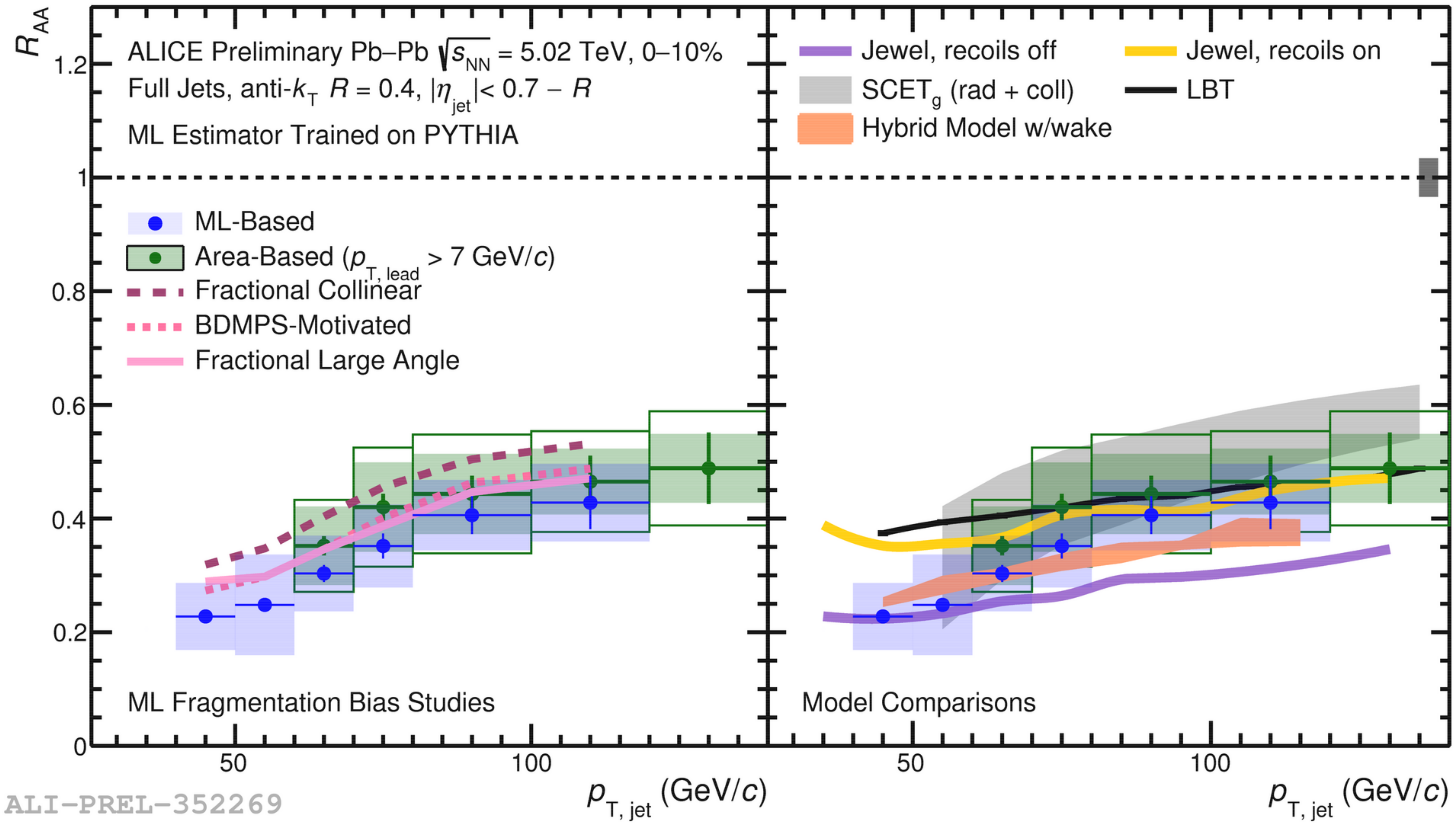
Method is relatively robust to the explored biases!

Lower p_T is a largely unexplored region. Machine learning provides us with an opportunity to study this.



Comparing to models

Keeping previous studies in mind, let's compare to models!



Aiming to constrain models at low p_T with new measurement technique!

JEWEL: Scattering and radiative energy loss, with/without recoiling medium.
JHEP 1707 (2017) 141

SCETg: Interactions with medium mediated by Glauber gluon exchange.
JHEP 07 (2019) 148

Hybrid Model: medium response via wake. AdS/CFT non-pert. regime.
Phys. Rev. Lett. 124, 052301

LBT: hydrodynamic medium, jet-medium interactions, recoils.
Phys. Rev. C 99 (2019) 054911

Summary and conclusion

We present a novel machine learning based background correction, which allows for the extension of inclusive jet measurements to lower p_{T} and larger *R* than previously possible in ALICE.

See significant jet suppression down to low p_{T} for large resolution parameters.

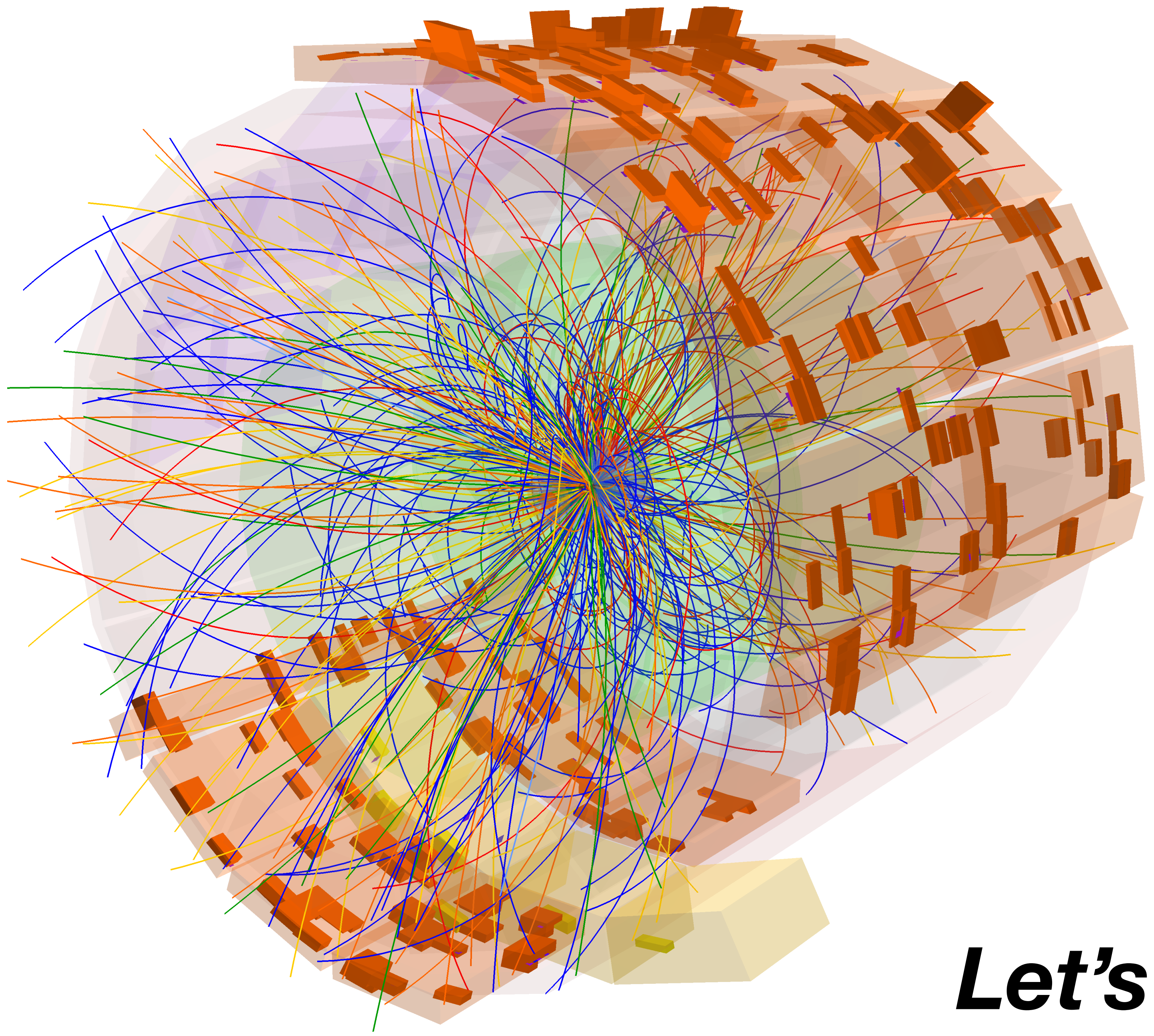
We study the fragmentation bias introduced by training the neural network on the constituents from PYTHIA using a toy model with three different modifications, estimating the effect of these modifications on the R_{AA}.

What's next?

Where do we go from here?

Our toy models are only simple tests of embedding a quenched signal, how do we get closer to the true case?

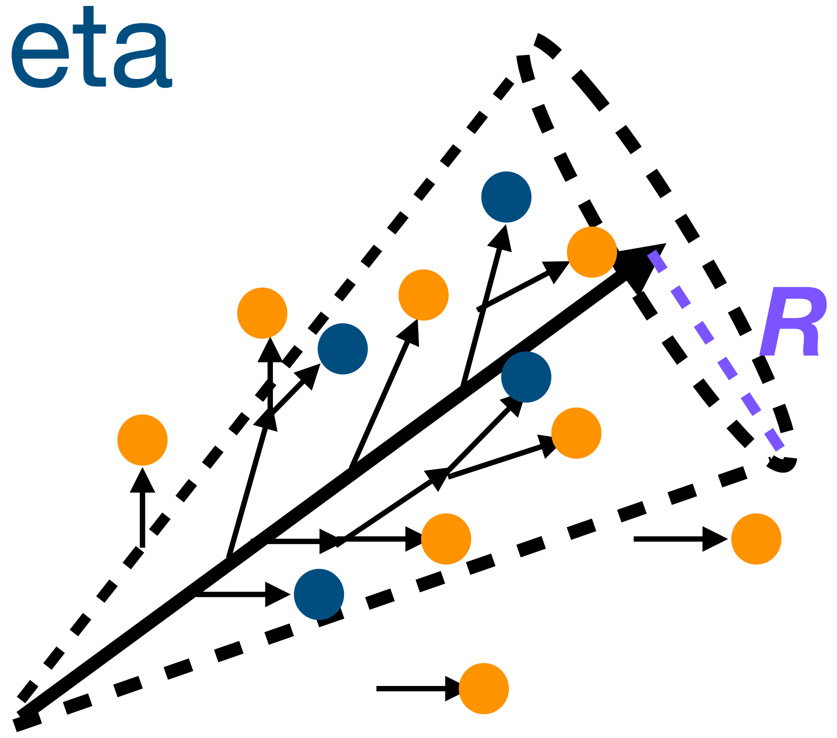
→ Train on a quenched MC: JEWEL, JETSCAPE, etc.



How far out can we go in R with ALICE?

charged particle jets: Limited to R = 0.9 max from eta acceptance of TPC.

full jets: Limited to R = 0.7 max from eta acceptance of EMCAL.



Let's see how far we can go! Stay tuned!

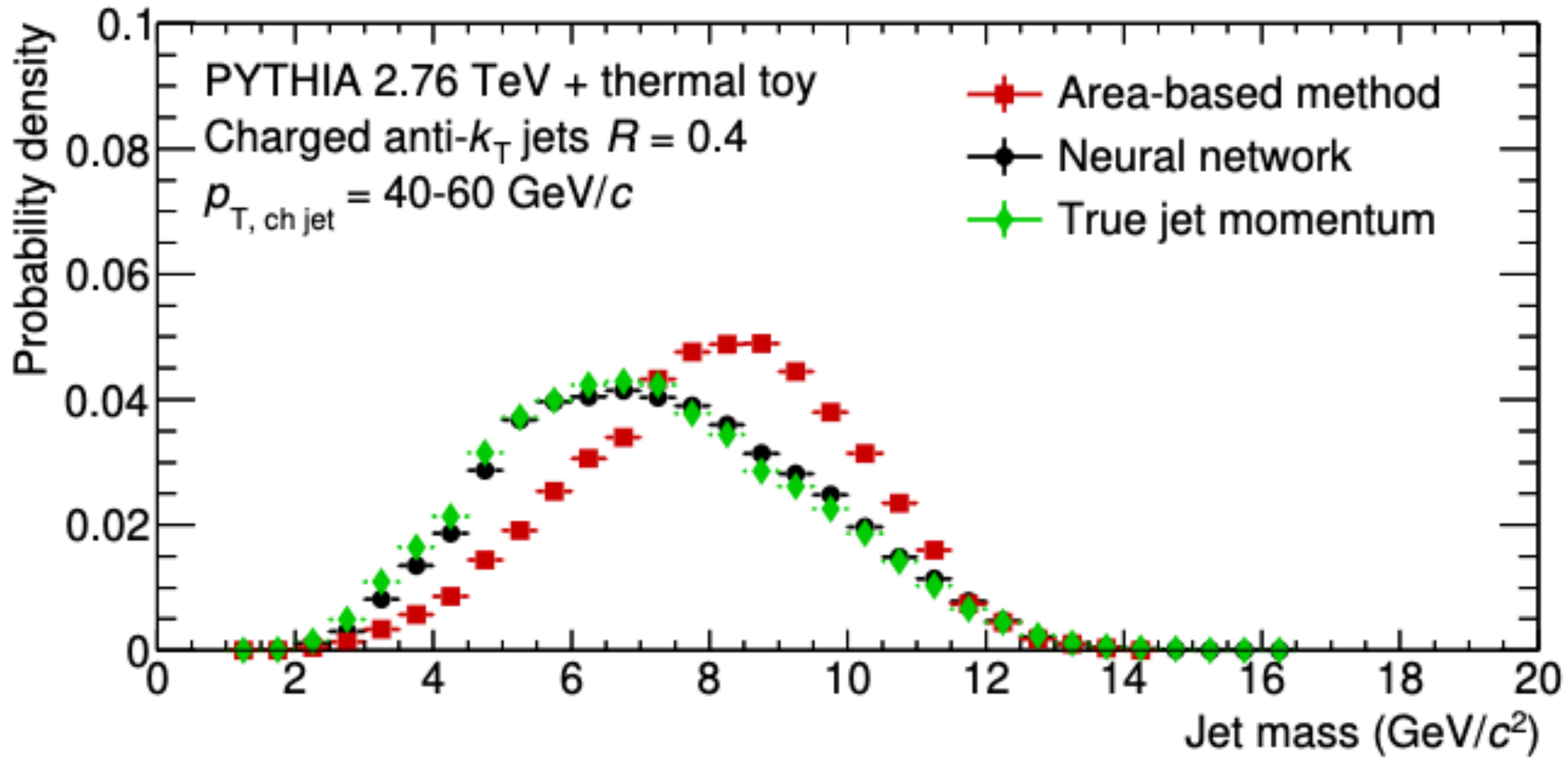
Backup

What variables could we use ML for?

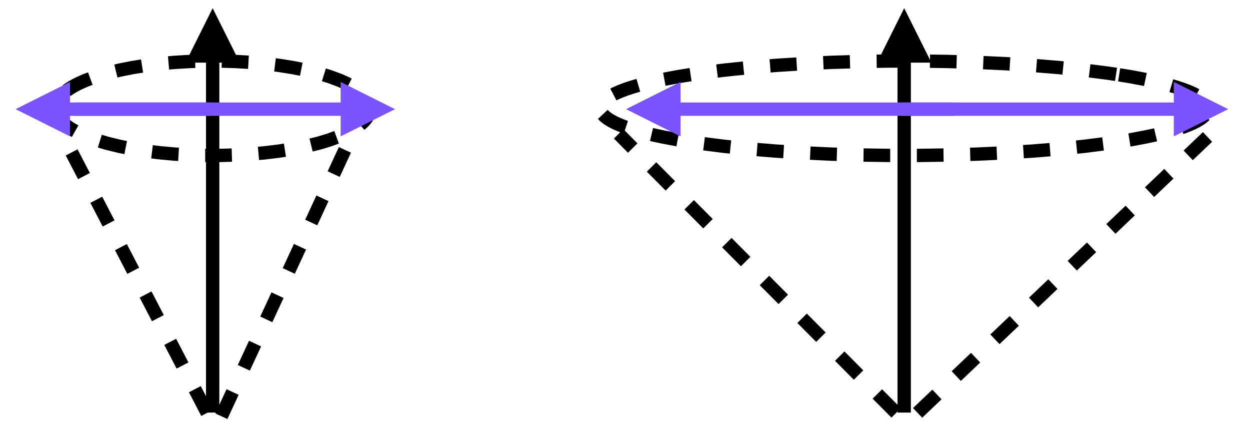
Jet mass is a good candidate for ML → binned in p_{T}!

→ better determination of p_{T}, better determination of jet mass

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

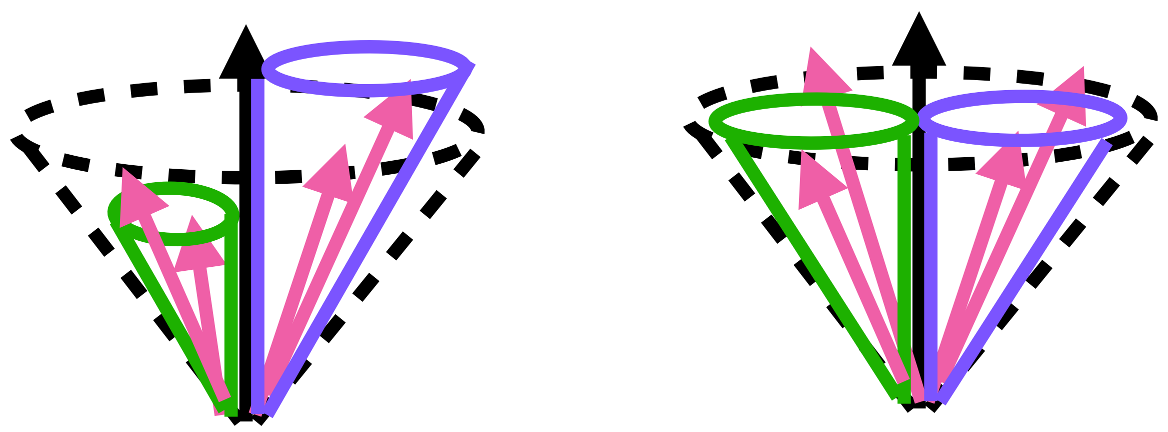


Jet Mass



Already see good performance!

Jet Splittings



Next frontier: Could we use ML for substructure??

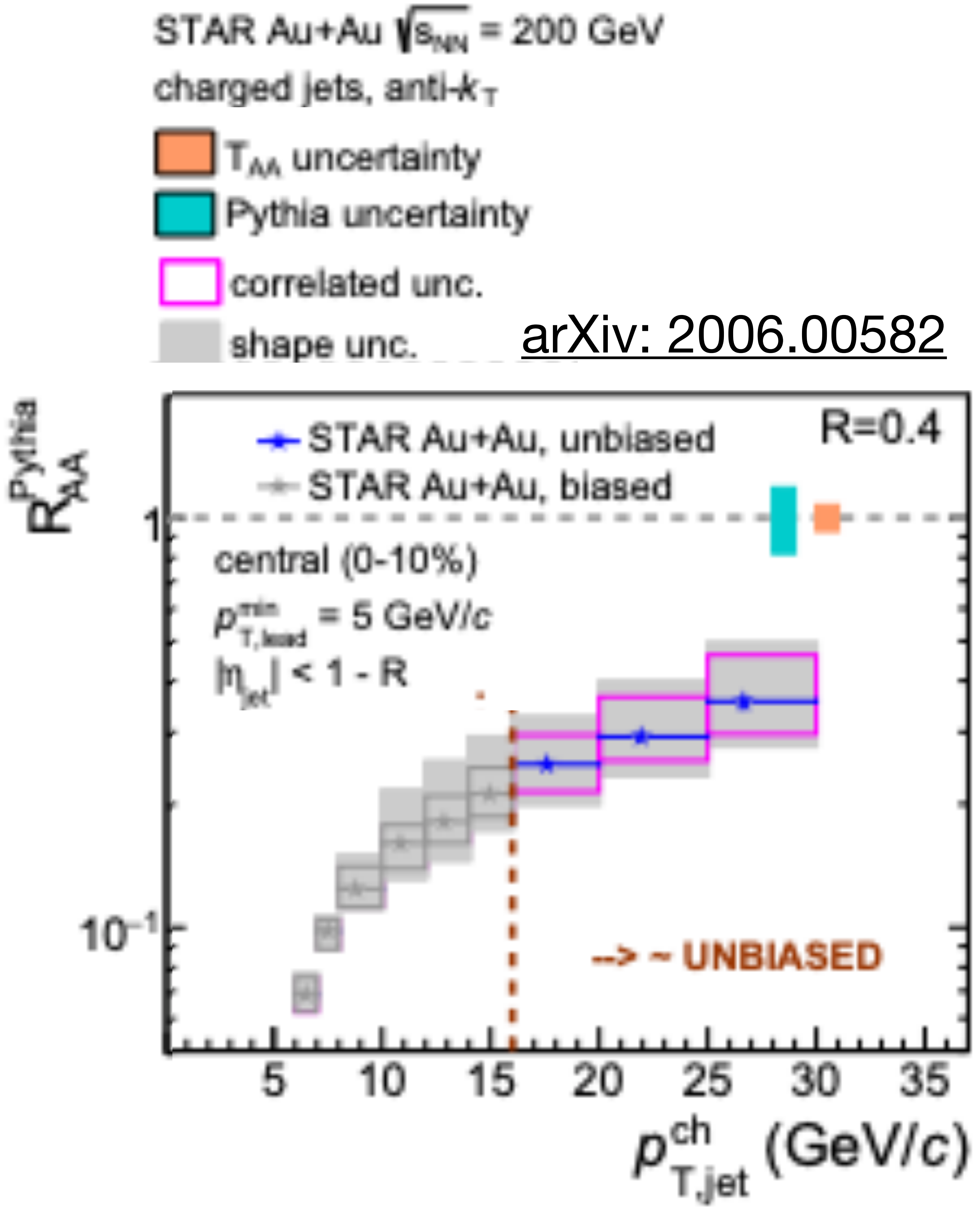
Measurements of inclusive jet R_{AA}

Many measurements of the inclusive jet R_{AA} for R = 0.4 jets in central 0-10% collisions.

See suppression across many different scales!

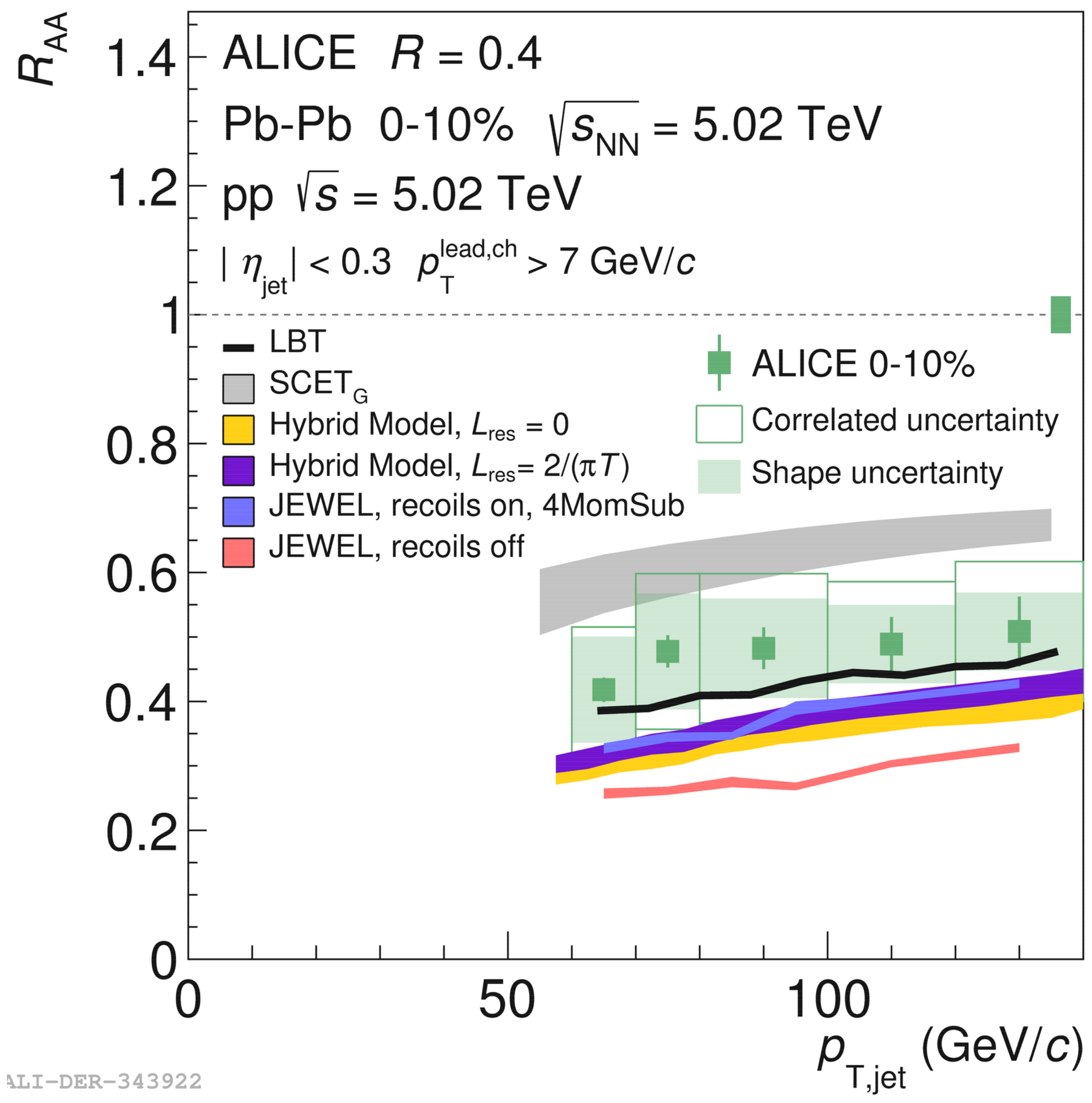
STAR

arXiv: 2006.00582



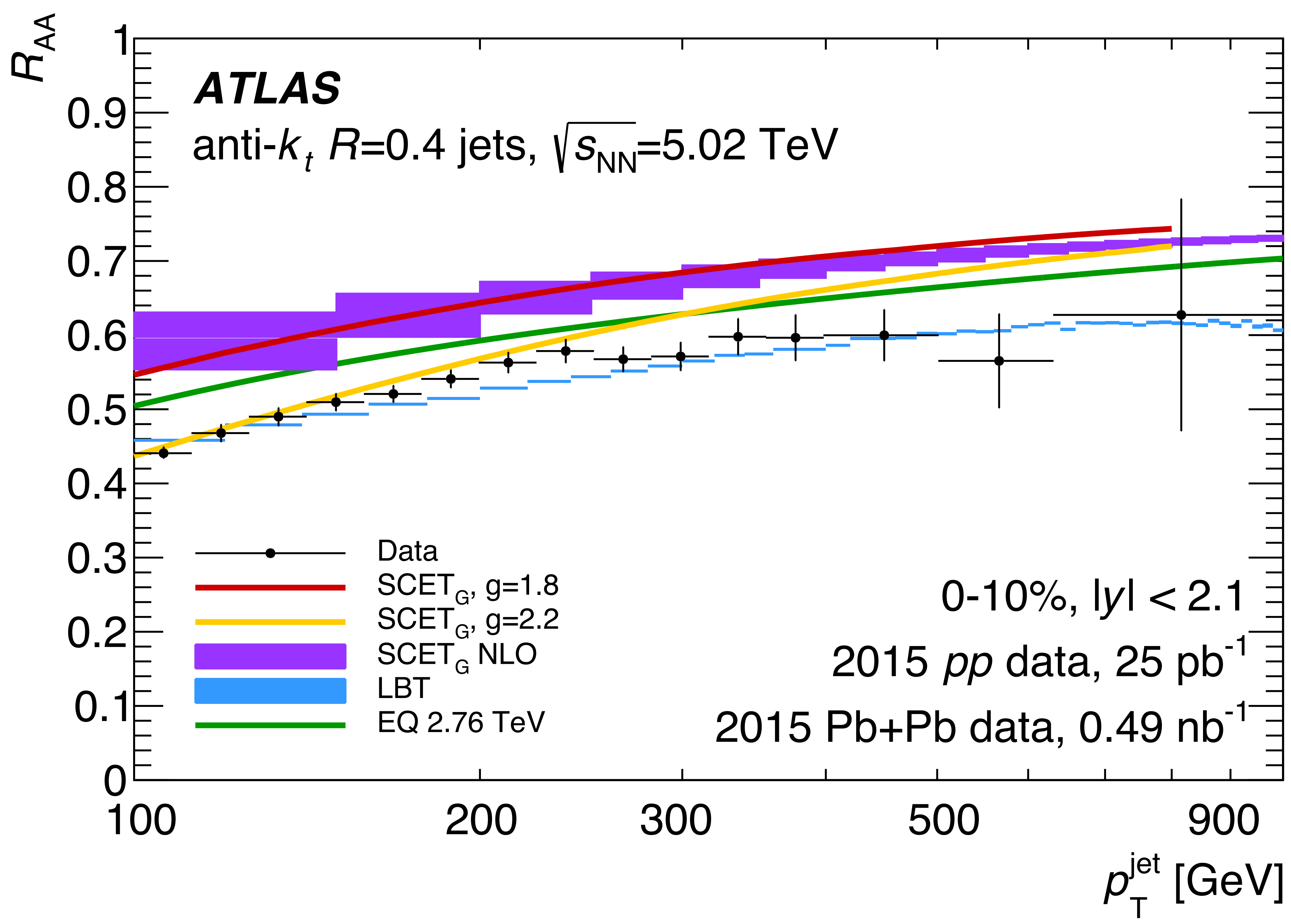
ALICE

Phys. Rev. C 101 034911 (2020)



ATLAS EXPERIMENT

Phys. Lett. B 790 (2019) 108

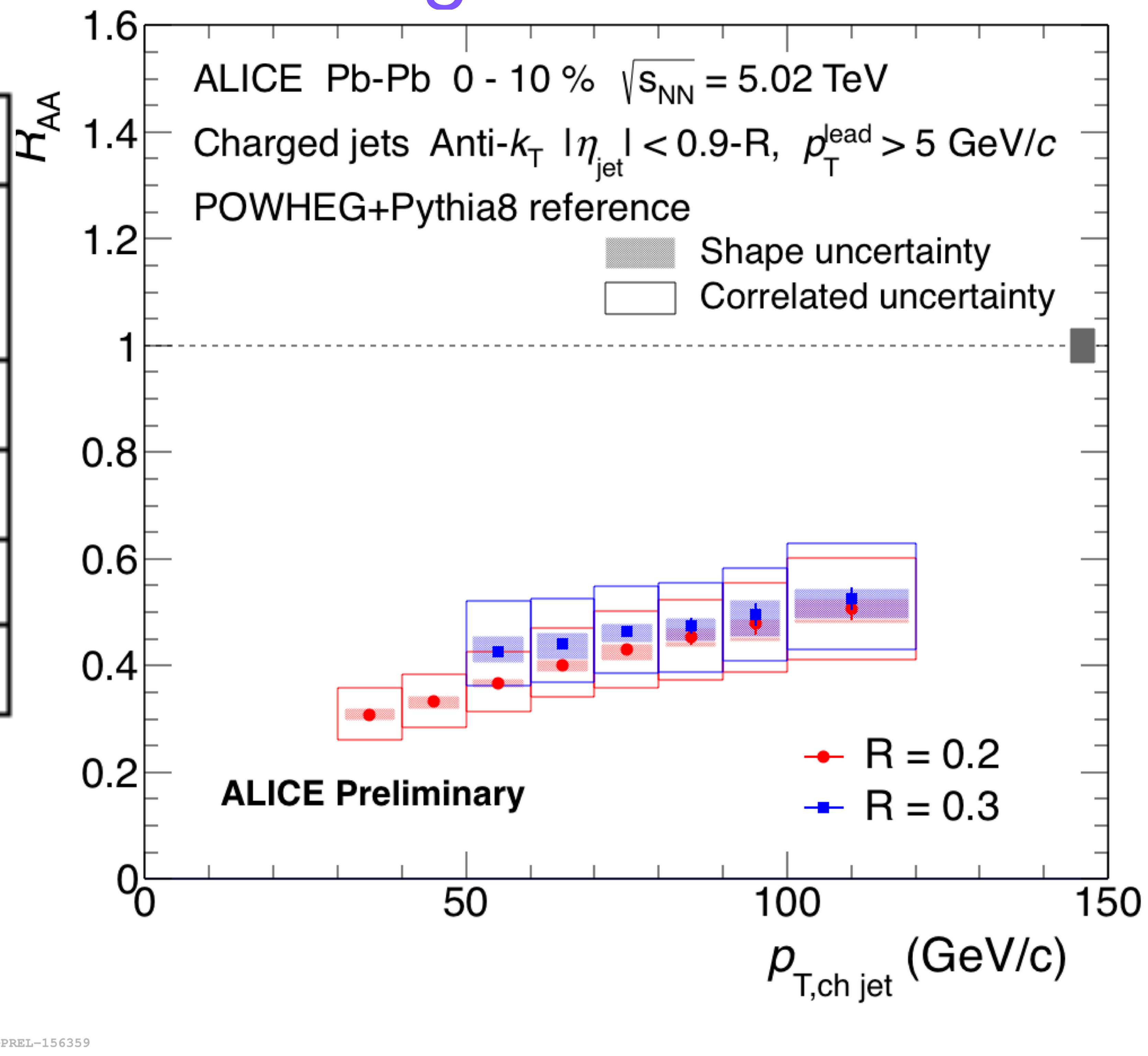


Where are we now in ALICE? *(area based method)*

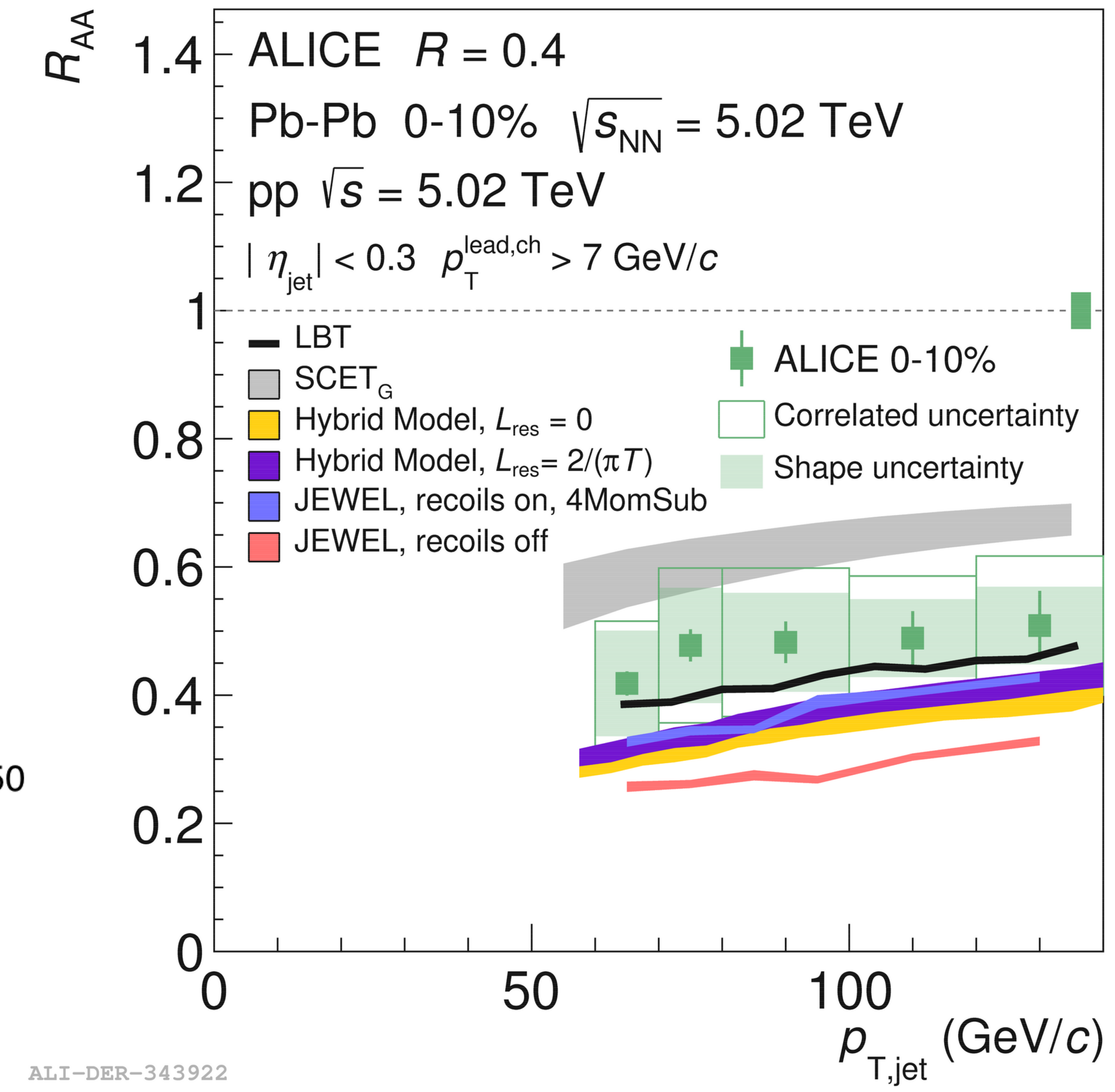
Inclusive Jet Measurement Summary

	Lower p_T Cutoff (GeV/c)	
R	Charged Particle Jets	Full Jets
0.2	30	40
0.3	50	60
0.4	N/A	60
0.6	N/A	N/A

Charged Particle Jets



Full Jets



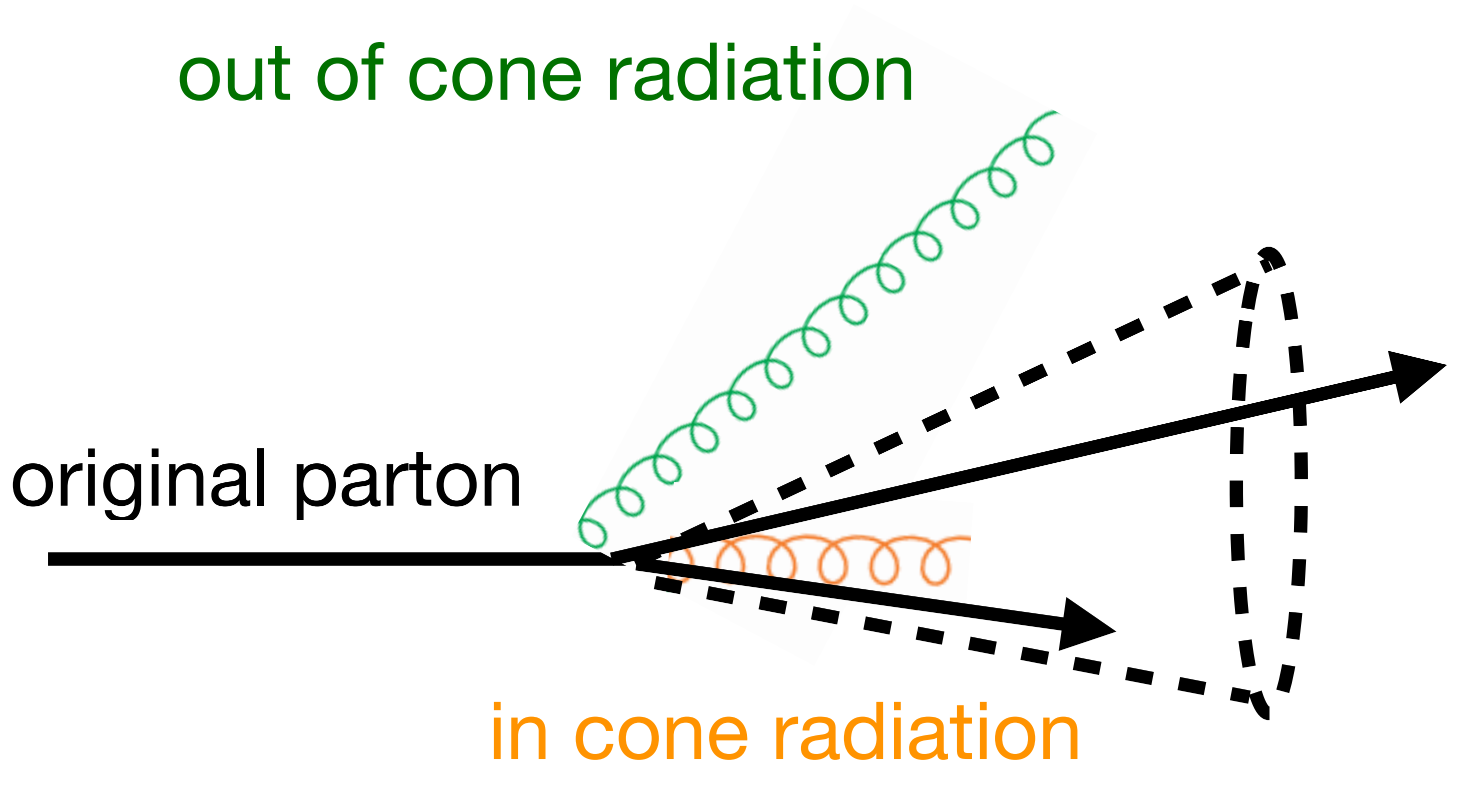
We see a suppression!

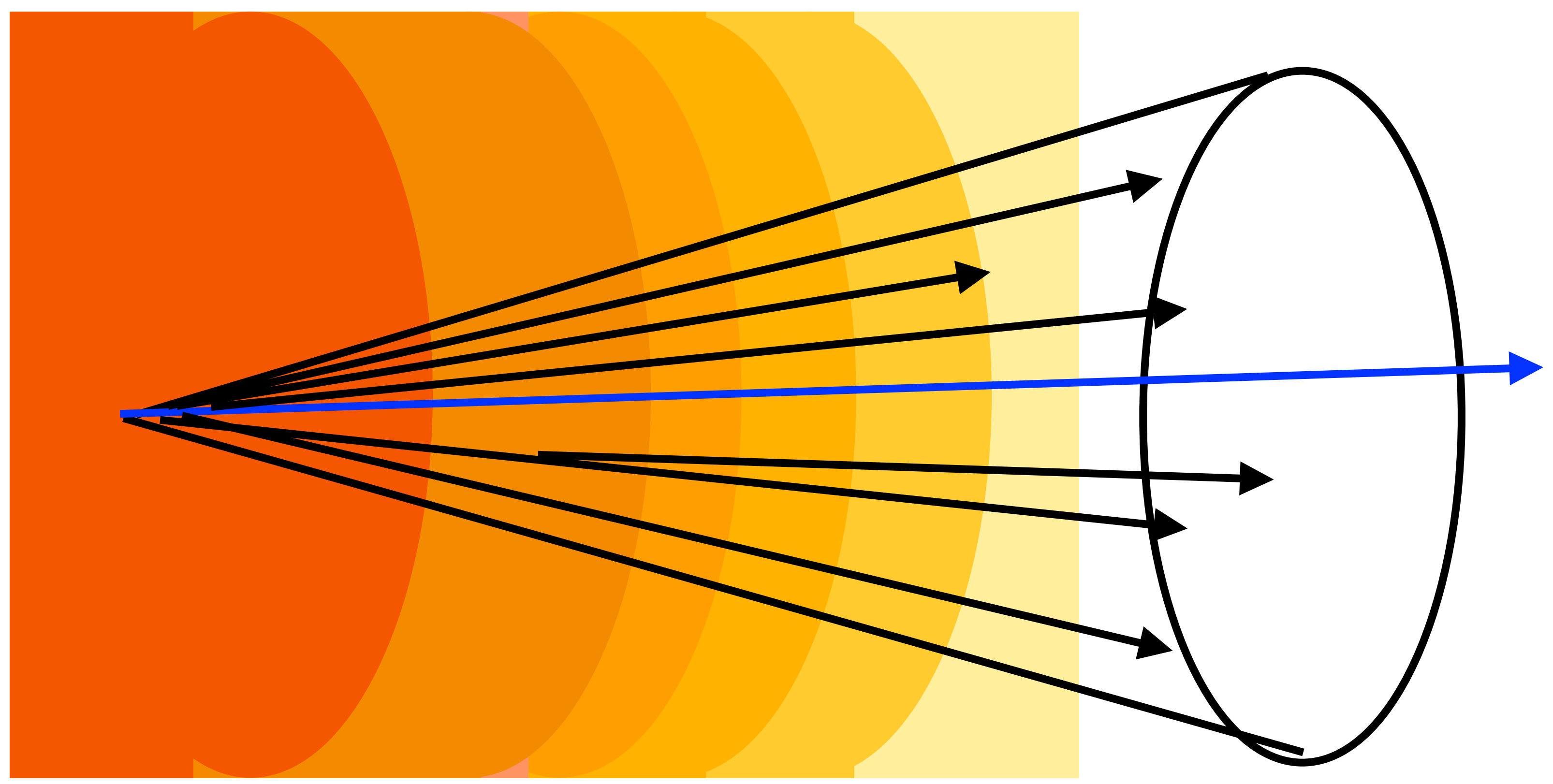
ALICE benefits from precise tracking at low p_T.

Prevented from going lower by large fake jet contribution at these low jet p_Ts!

Pushing to low p_T and large *R*

Many differential measurements of nuclear modification separate out energy loss effects.





Momentum broadening causes energy to be lost outside of the jet cone → R_{AA} ↘

Recover energy deposited in the medium → R_{AA} ↗

Recoiling medium adds energy to jet cone → R_{AA} ↗

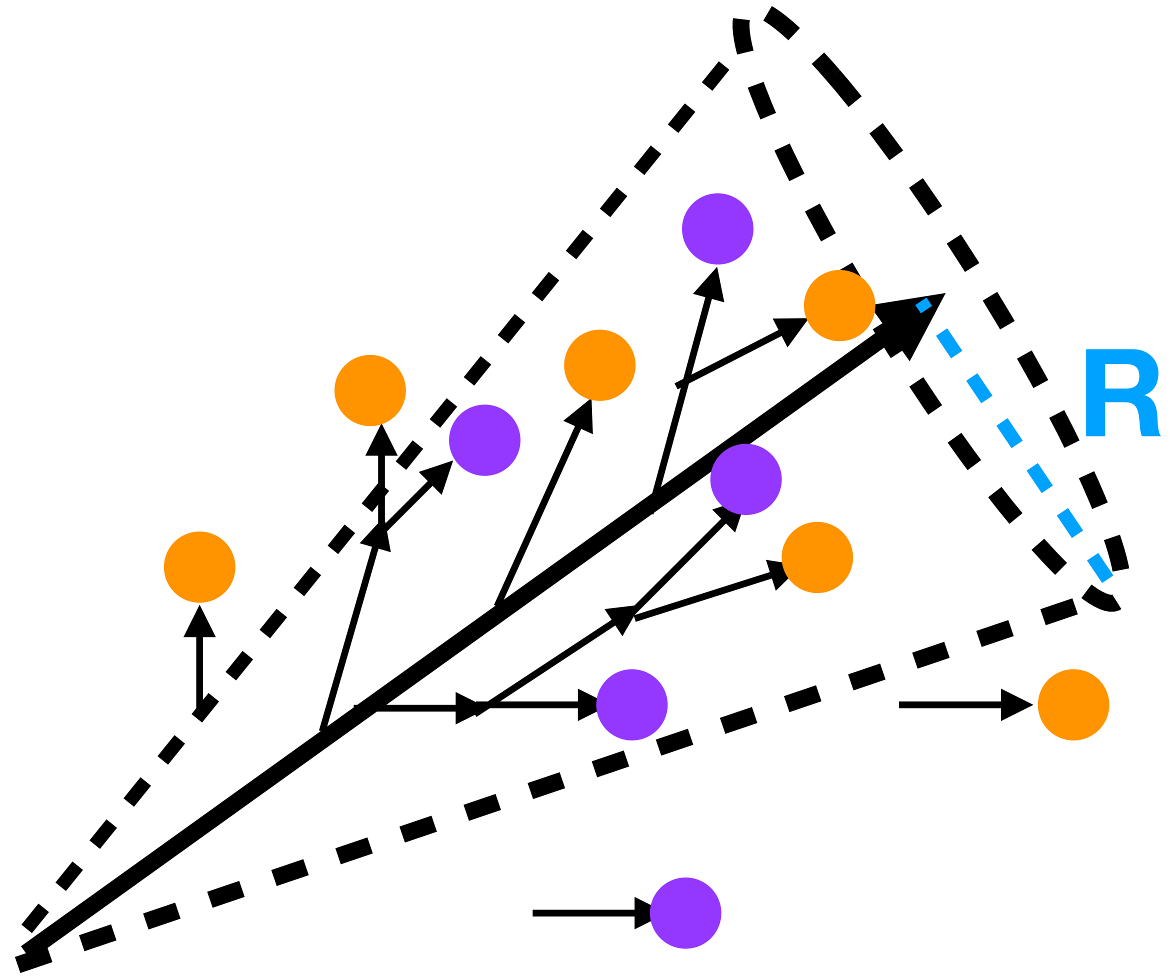
Wider jets have more complex structure, which could experience more quenching → R_{AA} ↘

Different jets with different structure experience these effects differently

→measure dependence of R_{AA} on p_T ***and*** *R!*

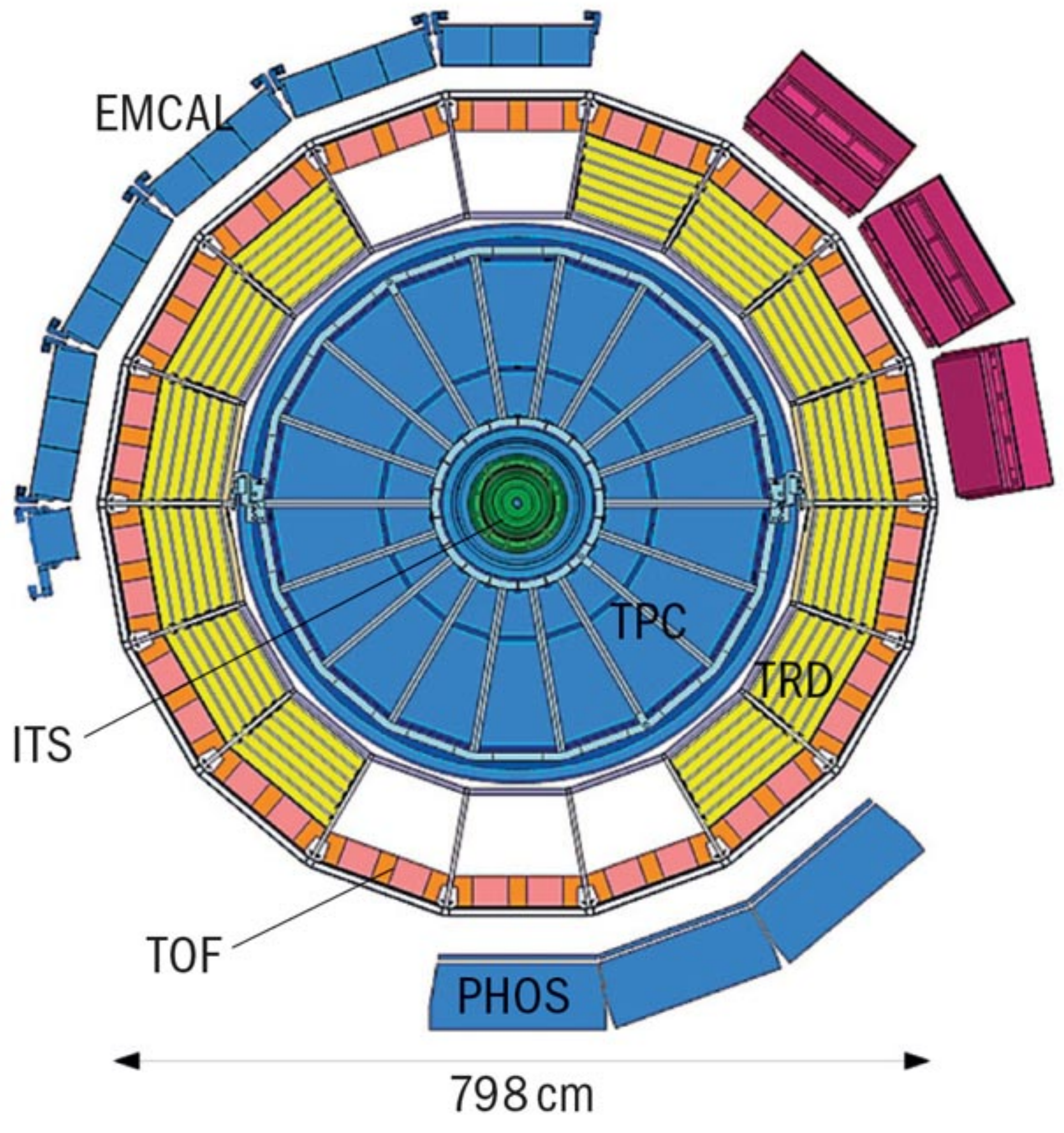
Remember: Low p_T and large *R* are difficult regions to study with inclusive jet probes.

Analysis details



Inclusive Pb—Pb jet sample at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV $L \sim 250\ \mu b^{-1}$ with the ALICE detector in 2015.

anti-k_{T} jets with various resolution parameters *R* and centralities



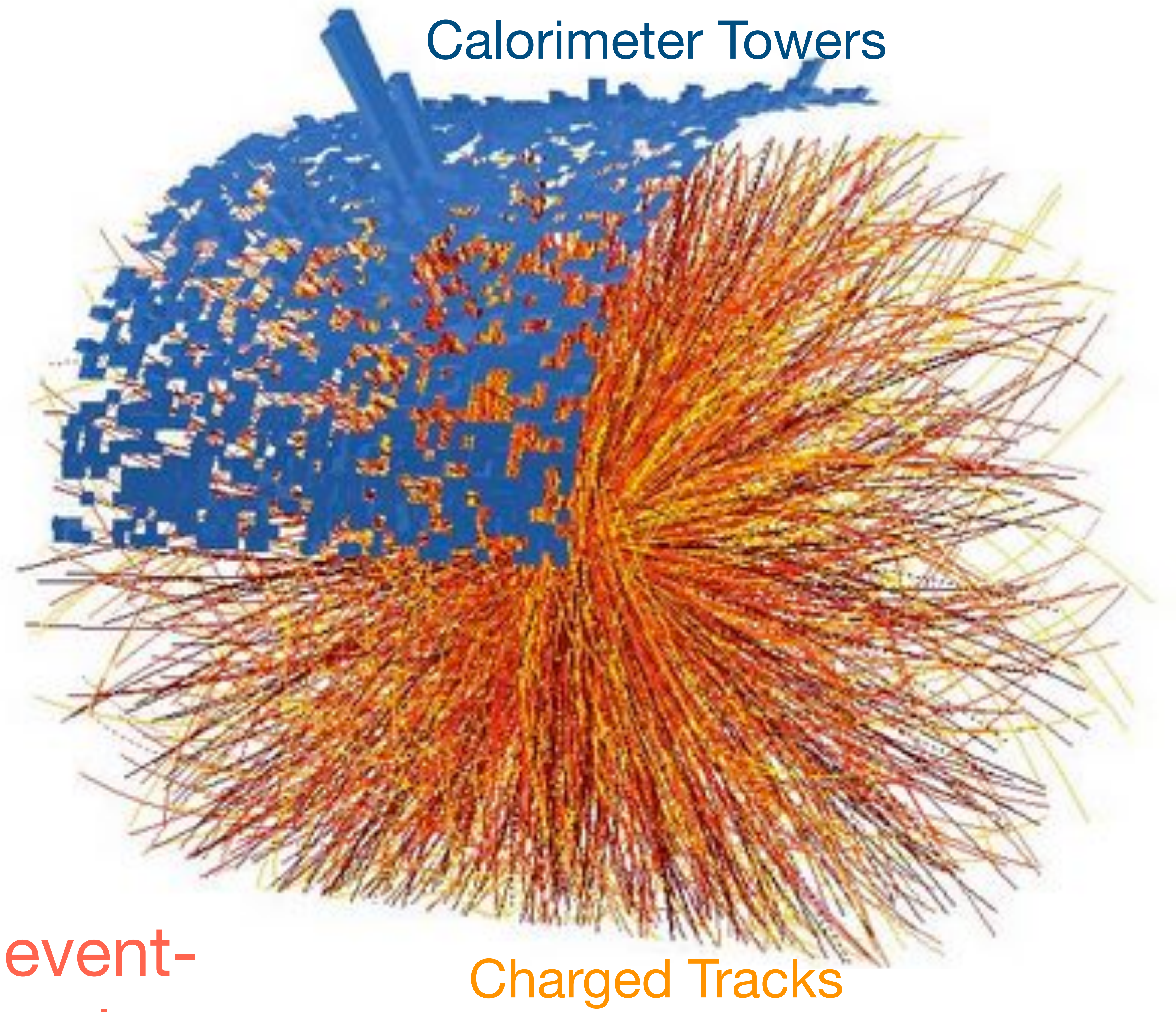
Charged particle jets → contain the charged component of the jet
→ measured with tracking detectors

Full jets → contain charged and neutral components of the jet
→ measured with electromagnetic calorimeter
→ limited to fiducial phi acceptance

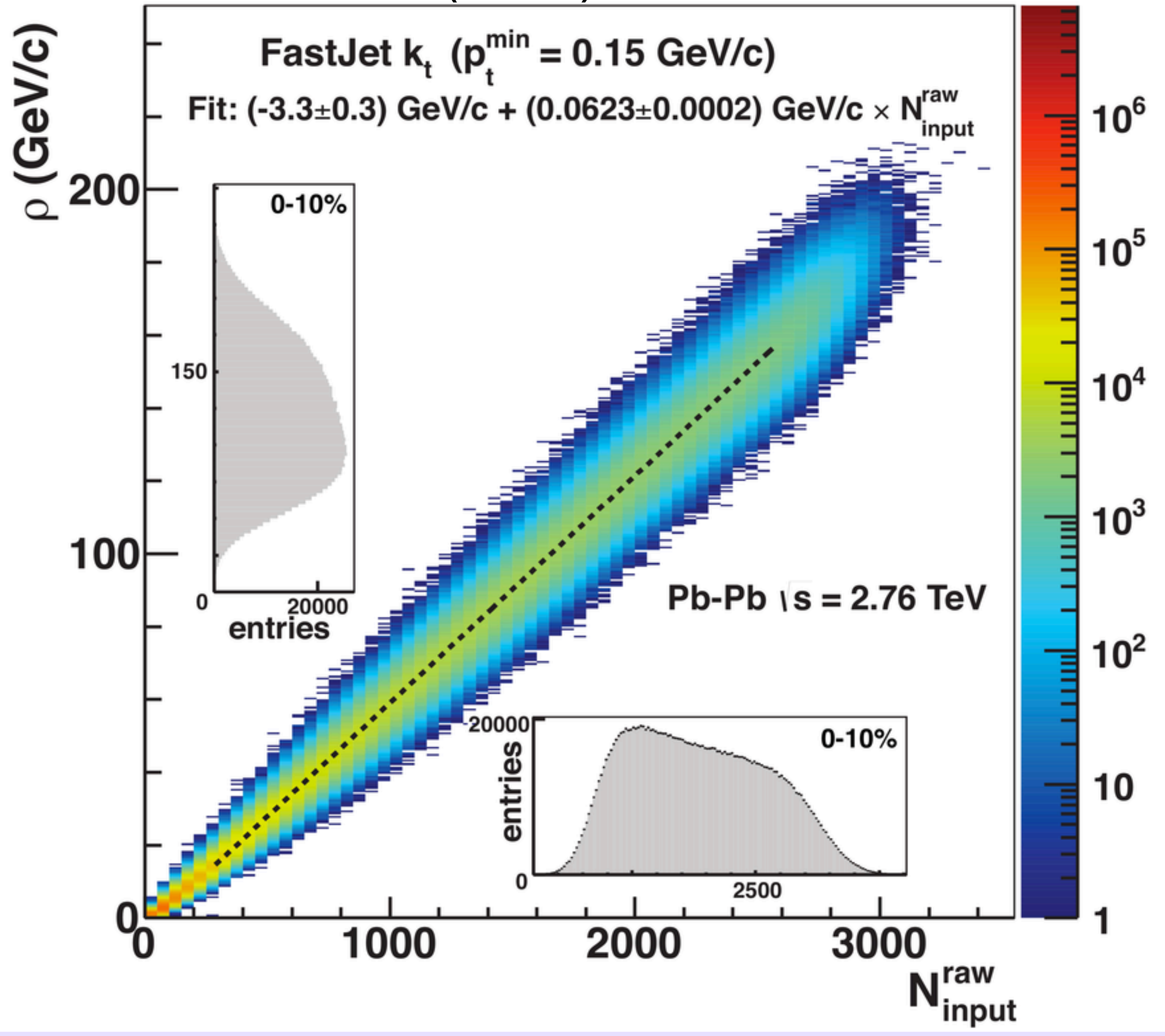
Reconstructing jet p_T

Reconstruction of inclusive jet p_T in heavy-ion collisions is made difficult by the large fluctuating background from the underlying event.

Fluctuations can be on the order of jet itself!



JHEP 1203 (2012) 053



Area based method:

Pedestal subtraction of event-averaged momentum density.

1. Estimate and subtract the pedestal

2. Leading track bias to remove fake contributions

3. Correct for residual fluctuations via unfolding

$$p_{T,rec} = p_{T,raw} - \rho A$$

Pushing to low p_T and large R

Low p_T and large R are less studied regions with inclusive jet probes.

Modification varies at different scales.

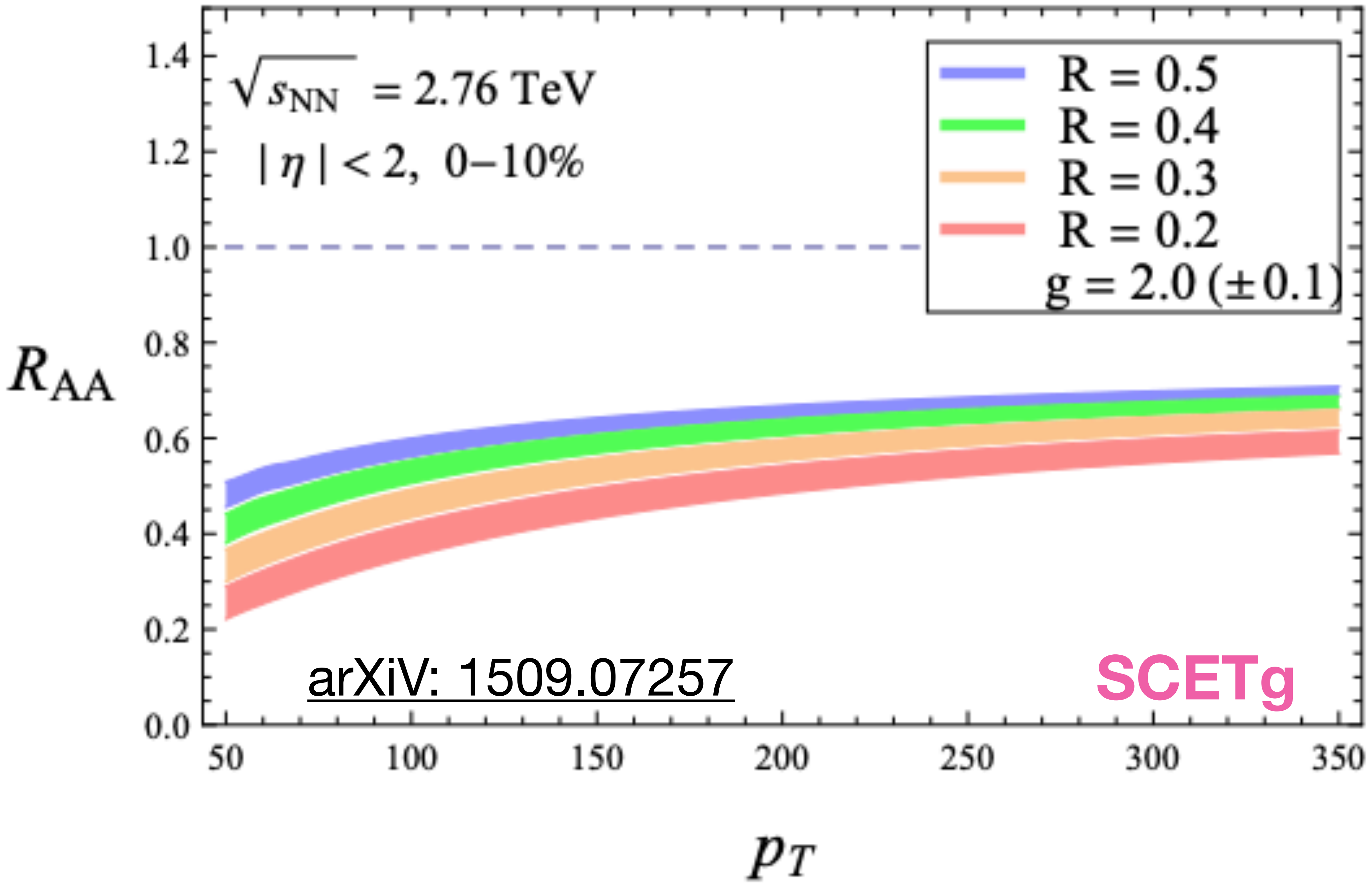
Allow for comparisons with jets at RHIC.

Are large R jets modified?

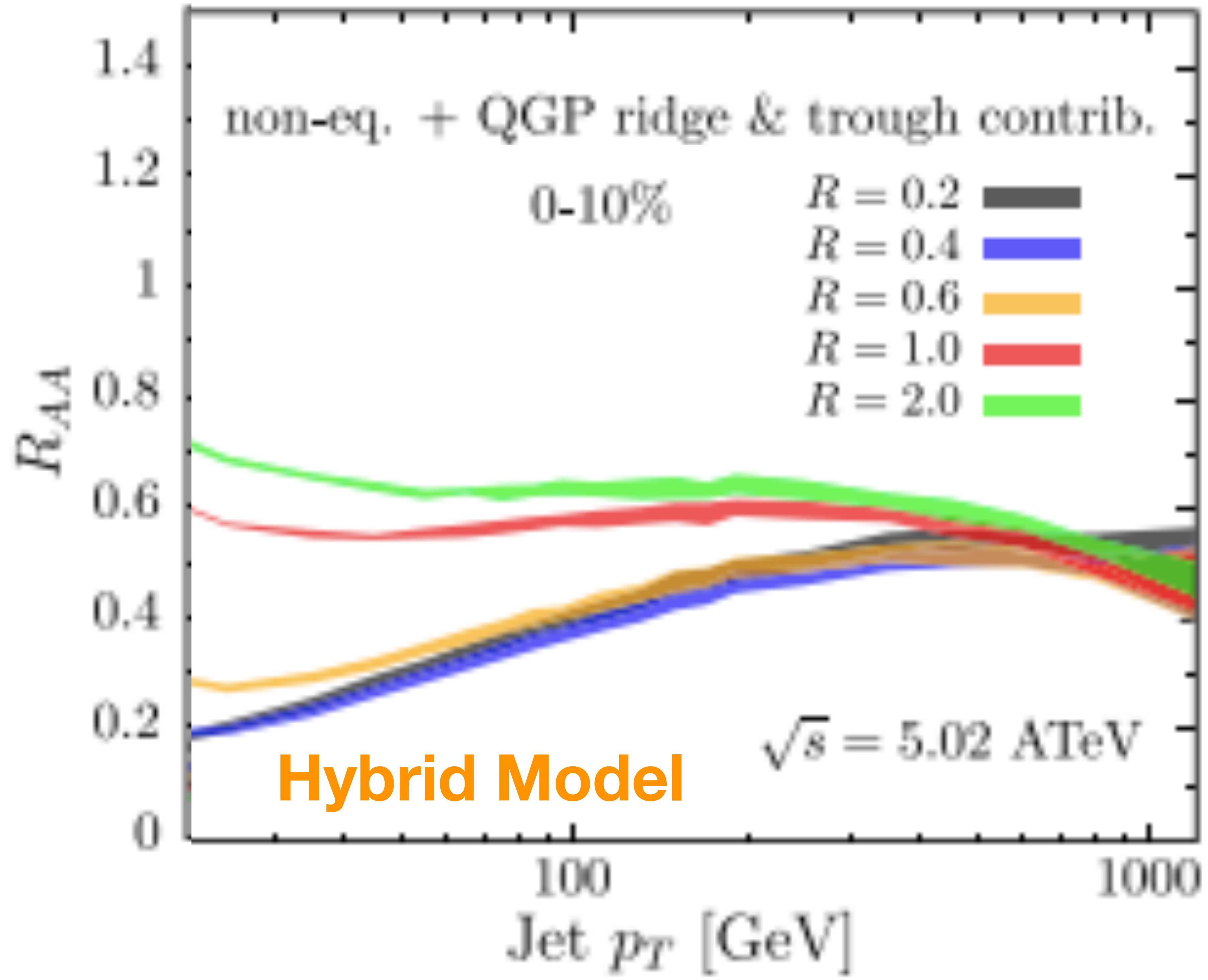
For more on the modification of jet substructure, check out talks by James Mulligan and Raymond Ehlers!

Jet substructure measurements in Pb-Pb collisions at 5.02 TeV with ALICE

Exploring large-R jets and substructure in Pb-Pb collisions at 5.02 TeV with ALICE



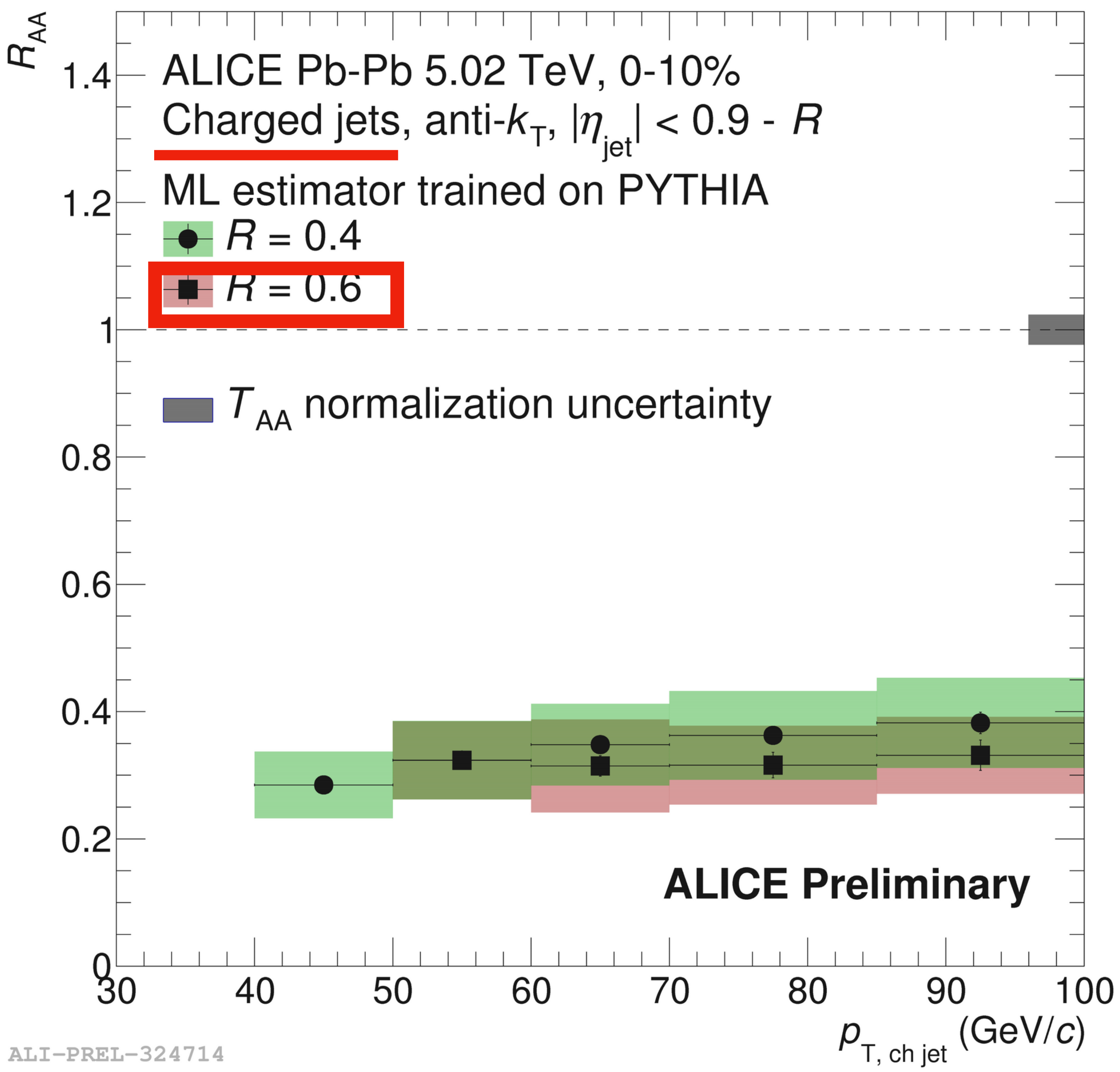
Phys. Rev. Lett. 124, 052301



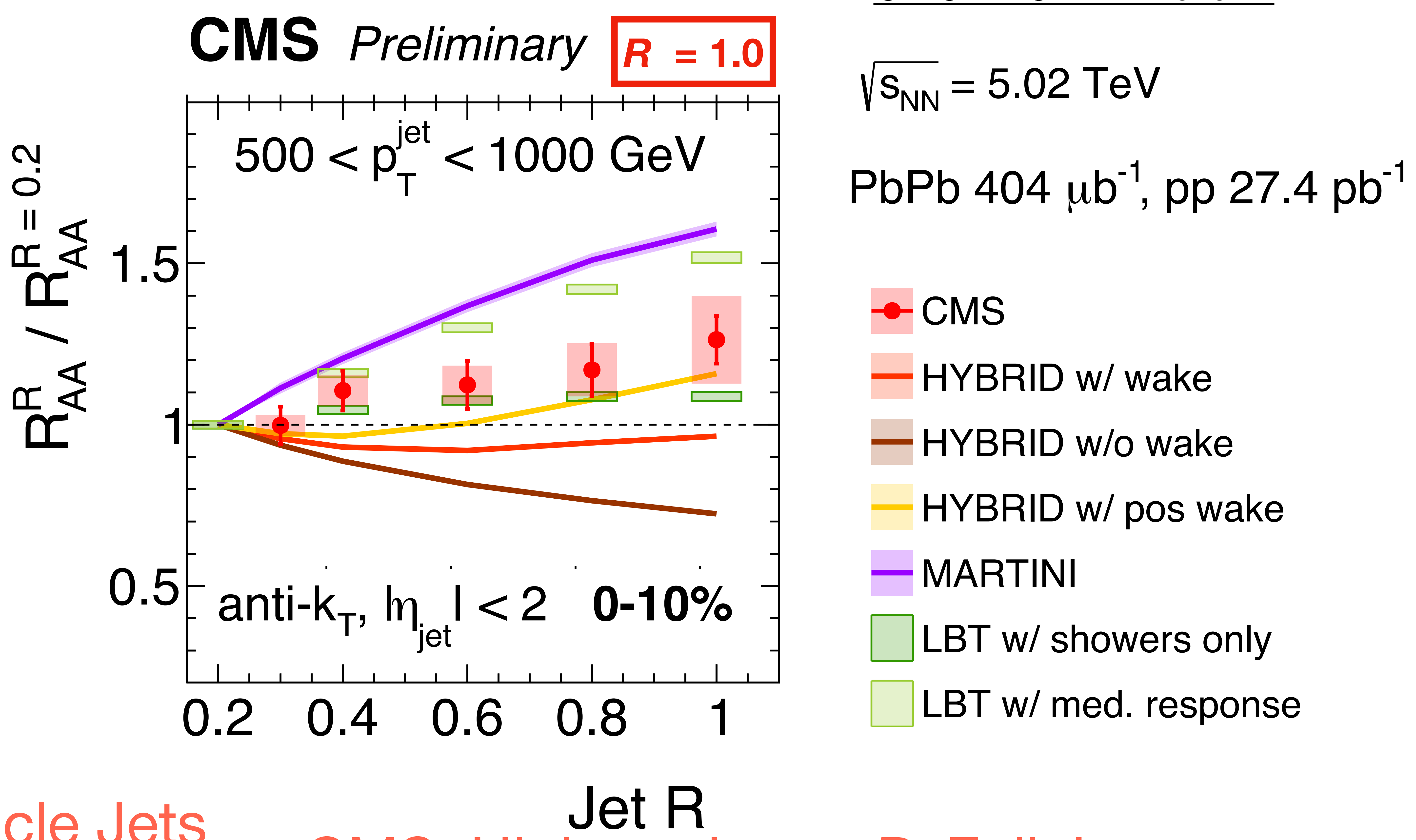
Pushing to low p_{T} and large *R*

Low p_{T} and large *R* are less studied regions with inclusive jet probes, lots of recent experimental efforts to extend measurements to these regions!

ML-Based Correction



CMS-PAS-HIN-18-014

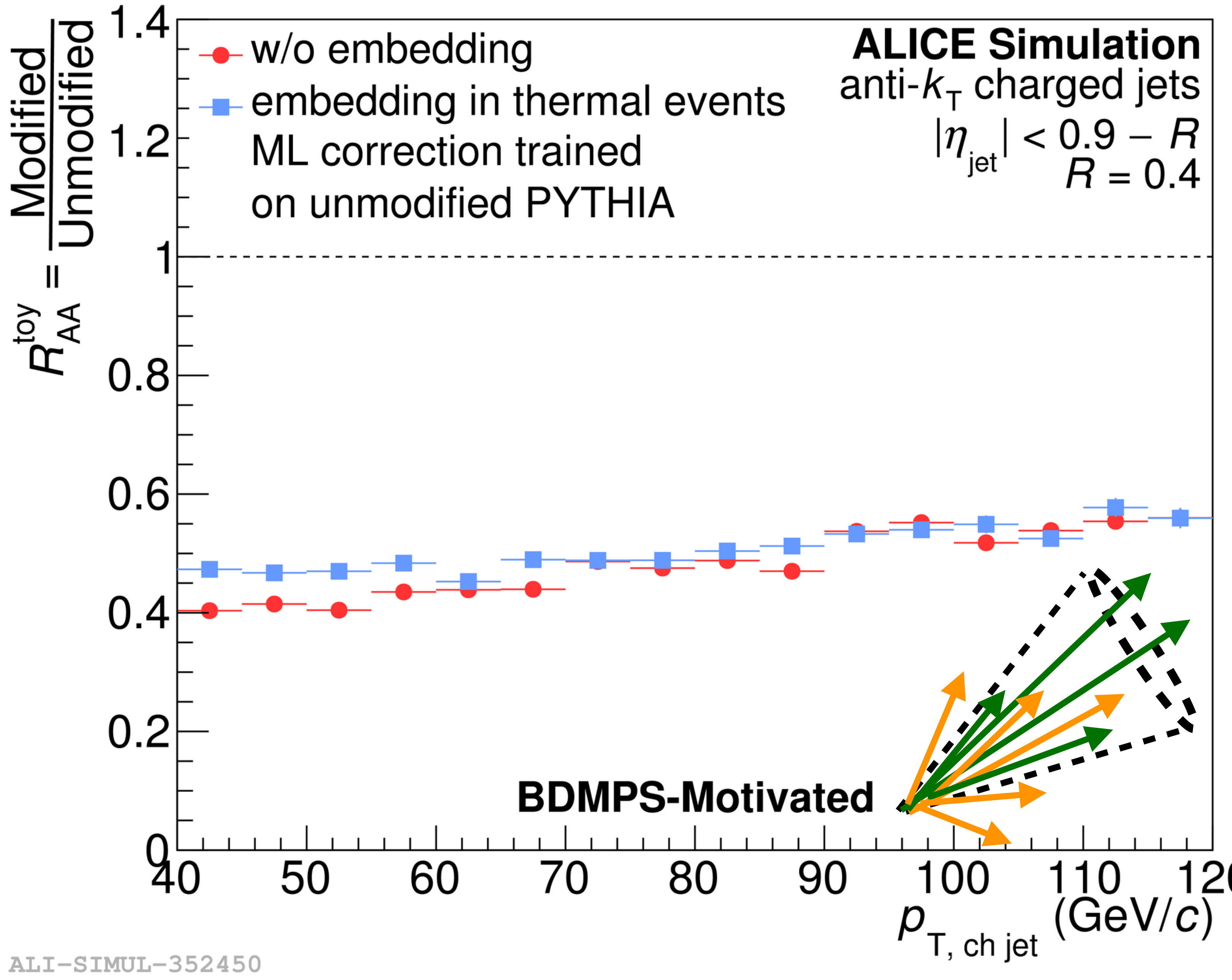


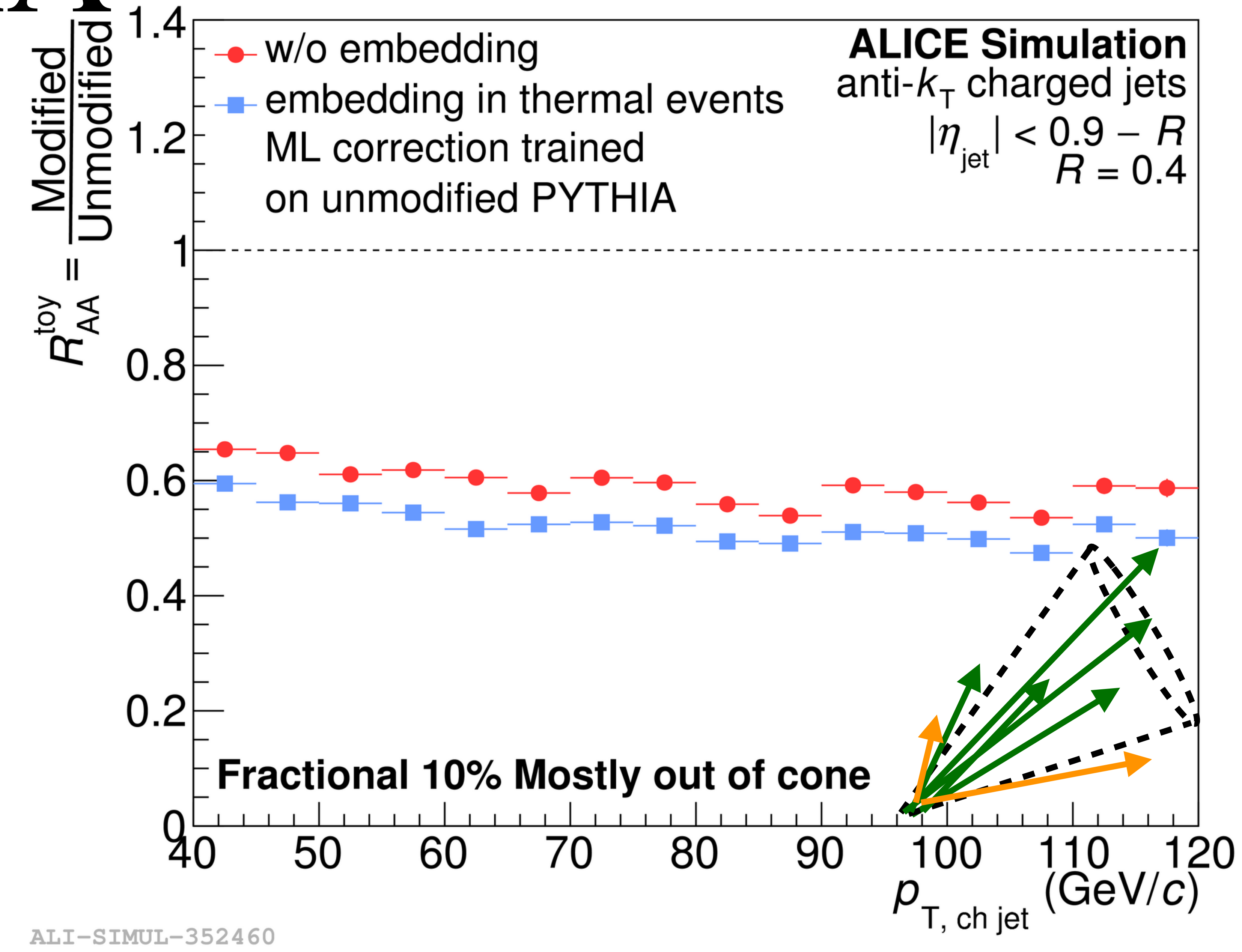
ALICE: Low p_{T}, Large *R*, Charged Particle Jets

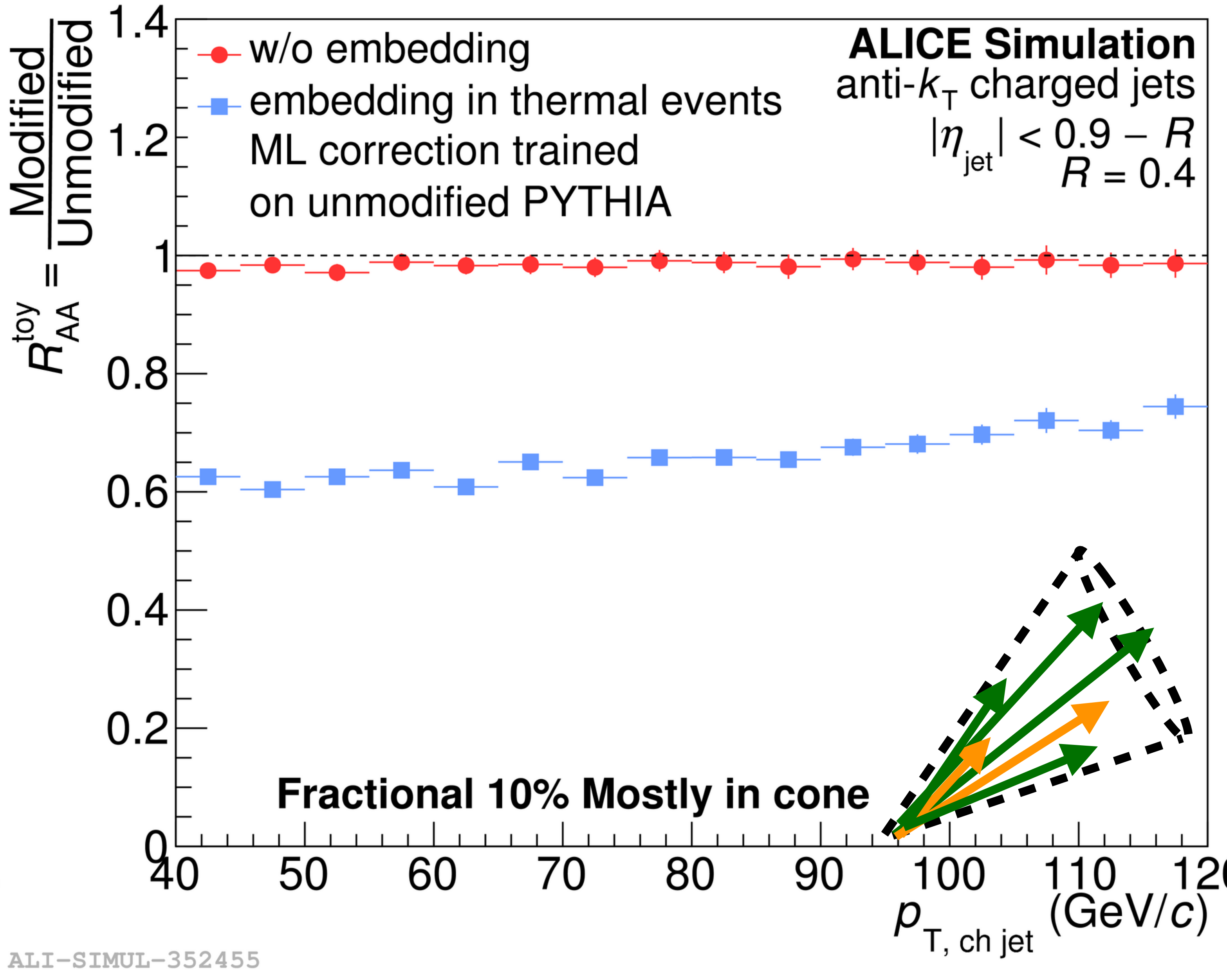
→Extend to full jets!

CMS: High p_{T}, Large *R*, Full Jets

Looking at $R_{\mathrm{AA}}^{\mathrm{toy}}$







1. Modify PYTHIA jets

2. Apply ML trained on unmodified PYTHIA

3. Look at $R_{\mathrm{AA}}^{\mathrm{toy}} = \frac{\textbf{Modified}}{\textbf{Unmodified}}$

Here, we focus on the difference between PYTHIA and Embedded (ML).

Largest difference for the mostly in cone case.

Will we see similar biases in the final result if we trained on modified toy?

Technical Details of the ML

Regression task where the regression target is the detector level jet p_{T}.

Here we are prioritizing a simple model!

Training sample 10%, testing sample 90%.

Implemented in *scikit-learn*. Default parameters used unless otherwise specified.

Shallow Neural Network

Shallow, 3 layers with [100, 100, 50] nodes

ADAM optimizer, stochastic gradient descent algorithm.

Nodes/neurons activated by a ReLU activation function.

Linear Regression

Normalization set to true by default.

Random Forest

Ensemble of 30 decision trees.

Maximum number of features set to 15.

Features for training

Ask ourselves two questions

How important is the feature to the model? → Feature Scores

Higher the feature score, more often variable is used in training.

How correlated is the feature with other features?

Feature	Score	Feature	Score
Jet p_T (no corr.)	**0.1355**	$p^1_{T,\mathrm{const}}$	0.0012
Jet mass	0.0007	$p^2_{T,\mathrm{const}}$	**0.0039**
Jet Area	0.0005	$p^3_{T,\mathrm{const}}$	0.0015
Jet p_T (area based corr.)	**0.7876**	$p^4_{T,\mathrm{const}}$	0.0011
LeSub	0.0004	$p^5_{T,\mathrm{const}}$	0.0009
Radial moment	0.0005	$p^6_{T,\mathrm{const}}$	0.0009
Momentum dispersion	0.0007	$p^7_{T,\mathrm{const}}$	0.0008
Number of constituents	0.0008	$p^8_{T,\mathrm{const}}$	0.0007
Mean of constituent p_Ts	**0.0585**	$p^9_{T,\mathrm{const}}$	0.0006
Median of Constituent p_Ts	0.0023	$p^{10}_{T,\mathrm{const}}$	0.0007

Iteratively remove unimportant or highly correlated features!

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

Features for training

Ask ourselves two questions

How important is the feature to the model? → Feature Scores

Higher the feature score, more often variable is used in training.

How correlated is the feature with other features?

Final List: Prioritizing a simple model!

Jet p_{T} (area-based corrected)

Number of Constituents within Jet

Jet Angularity

p_{T} of 12 Leading Constituents

BDMPS Toy Model Modification

$$P(\theta_g, \omega) = \alpha\omega\theta_q^3\sqrt{\frac{2\omega}{\hat{q}}}Le^{\frac{-\theta_q^2\omega^2}{\sqrt{2\omega\hat{q}}}}$$

JHEP 0109 (2001) 033

Modify the constituents of the jet by sampling the BDMPS gluon emission spectrum in the emission angle and energy.

For this study we use values of $\hat{q}$ = 2 and L = 7 fm and p_{loss} = 1.0 .

Motivation behind this is to emit from a probability distribution dictated by quenching theory.