Novel tools and challenges for jet physics in heavy-ion collisions

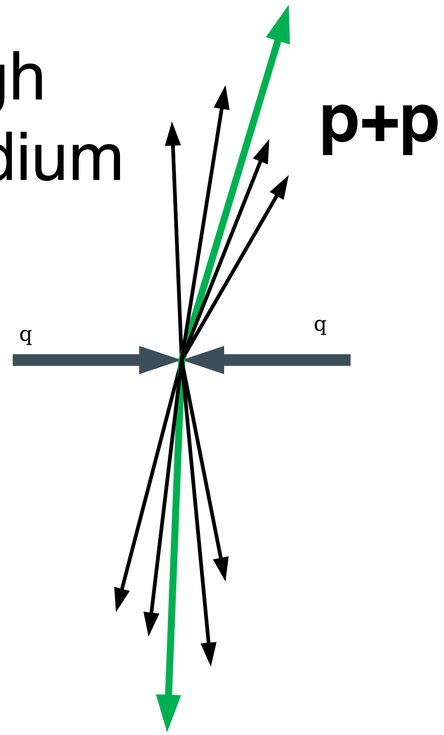
Laura Havener, Yale University
LHCP 2020, Remote
Tuesday, May 26th, 2020

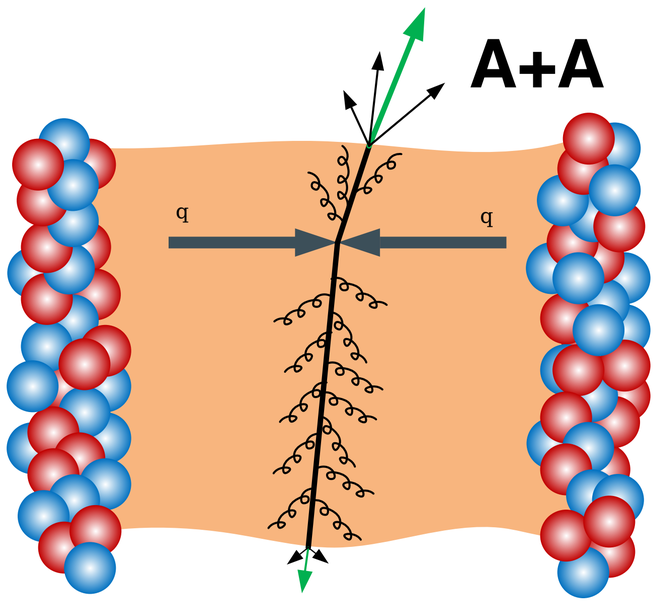
Wright Laboratory

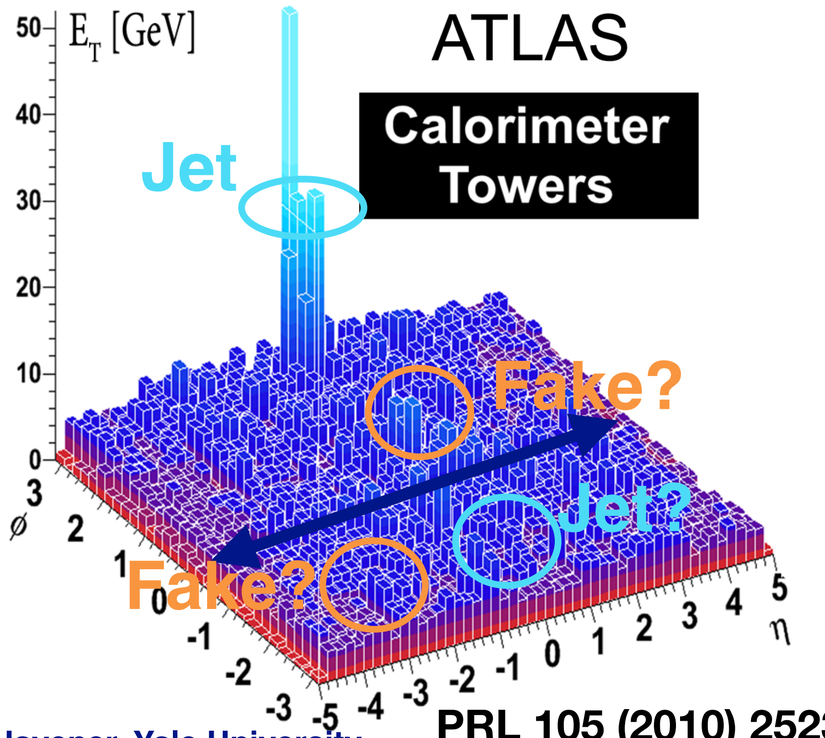
Yale

Jets in heavy-ion collisions

- Jet quenching: partons lose energy through strong interactions with the hot dense medium
  - ➡ Jet energy loss and modification of internal substructure
  - ▸ Study over large range in jet p_T and radii







PRL 105 (2010) 252303

- Large uncorrelated background due to underlying event (UE) fluctuations can be of the order of the jet itself
  - ➡ Be careful with fake jets from upward UE fluctuations (prohibits unfolding)

Measuring jets in HI collisions

- Challenges for measuring jets in heavy-ion collisions at the LHC:

➡ Large underlying background in heavy-ion collisions that is difficult to remove

- Restricts measurements in kinematic regions where we are still interested in the physics

Solution: improved background subtraction including machine-learning techniques

➡ Always searching for new variables that are sensitive to the physics we are interested in

- Especially ones that are can be compared to first principle calculations to help constrain models

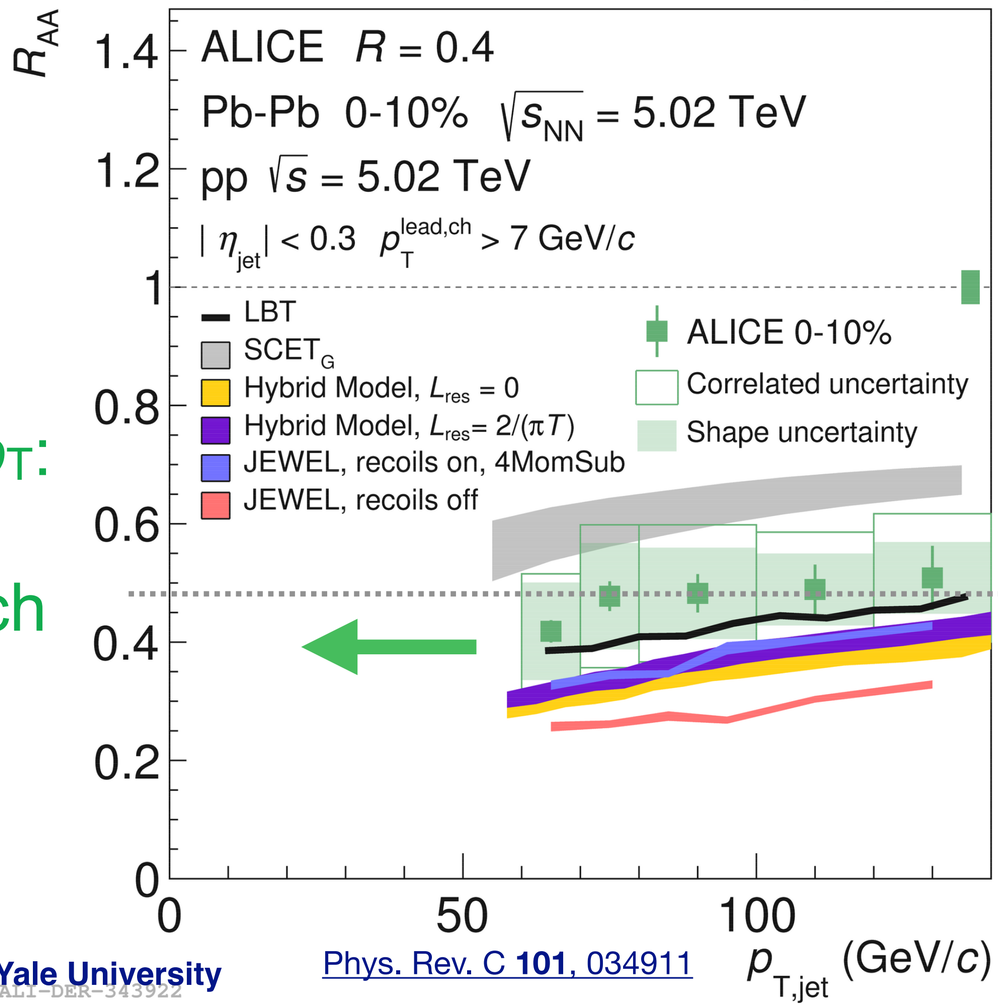
Solution: jet substructure tools and improved background subtraction for jet substructure

Inclusive jet suppression: LHC

- Inclusive jet suppression over a large jet p_T range

$$R_{AA} = \frac{\frac{1}{N_{event}} \frac{d^2 N_{jet}^{PbPb}}{dp_T dy}\Big|_{cent}}{\frac{< N_{coll} >_{cent}}{\sigma_{inel}} \times \frac{d^2\sigma_{jet}^{pp}}{dp_T dy}} \quad \frac{\text{\# Pb-Pb}}{\text{\# pp}}$$

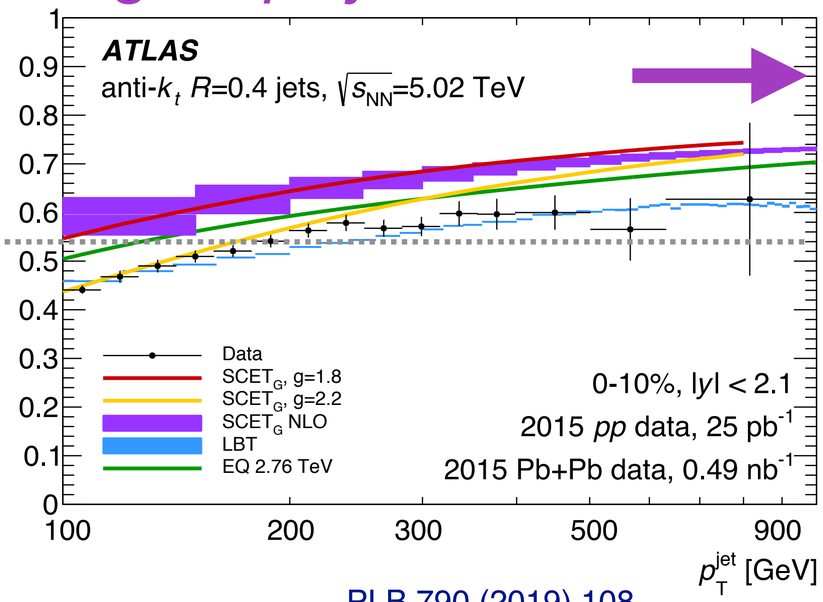
ALICE focuses on lower p_T jets



Lower p_T: ML approach

Phys. Rev. C **101**, 034911

ATLAS /CMS focus on higher p_T jets out to a TeV



Larger R: higher p_T

PLB 790 (2019) 108

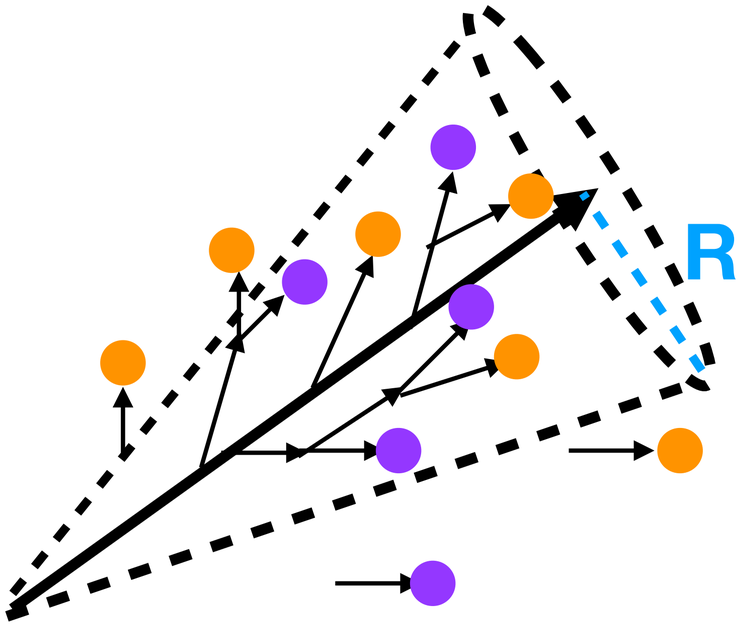
Machine learning approach

- Standard ALICE procedure misses residual fluctuations

Phys. Rev. C 99, 064904 (2019)

Includes a leading track bias

- ML technique learns a data-driven mapping to correct the jet p_T by exploiting the difference between the signal jets and the background



Introduces a potential fragmentation bias

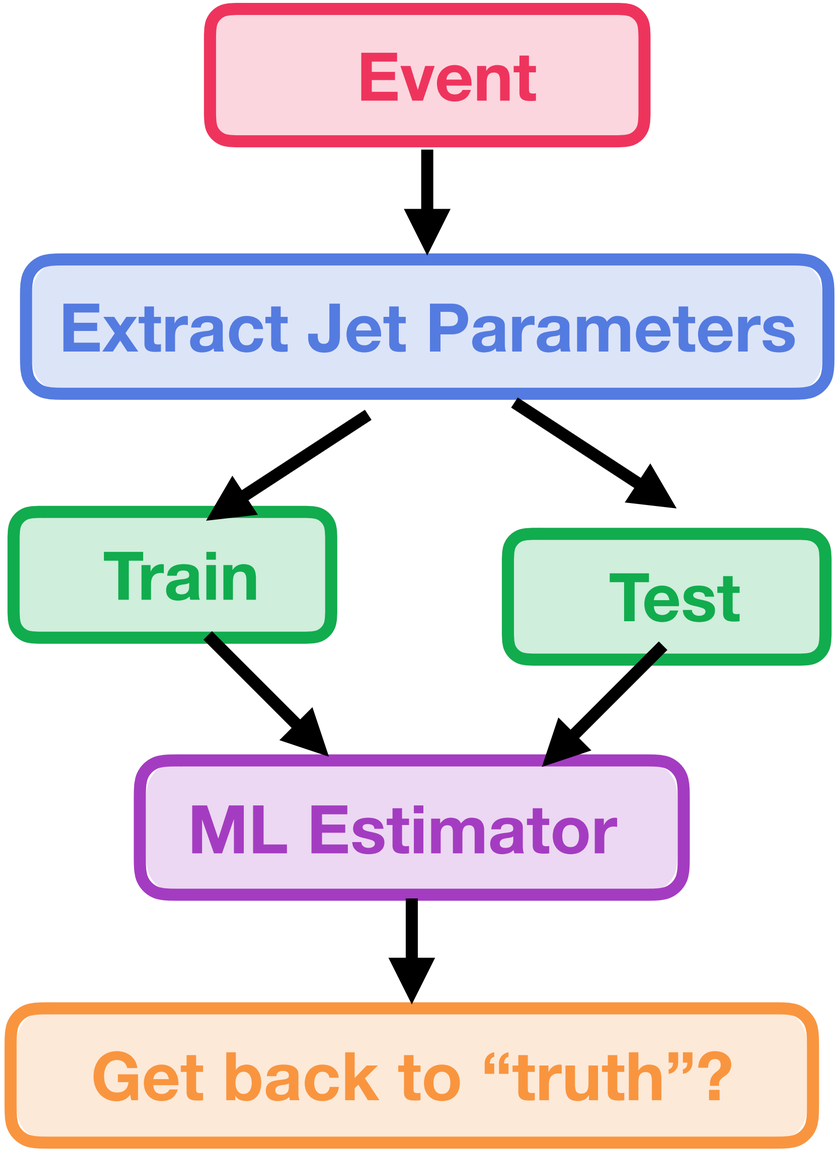
Jet properties including standard corrected p_T and jet constituents

ML

Correct jet p_T

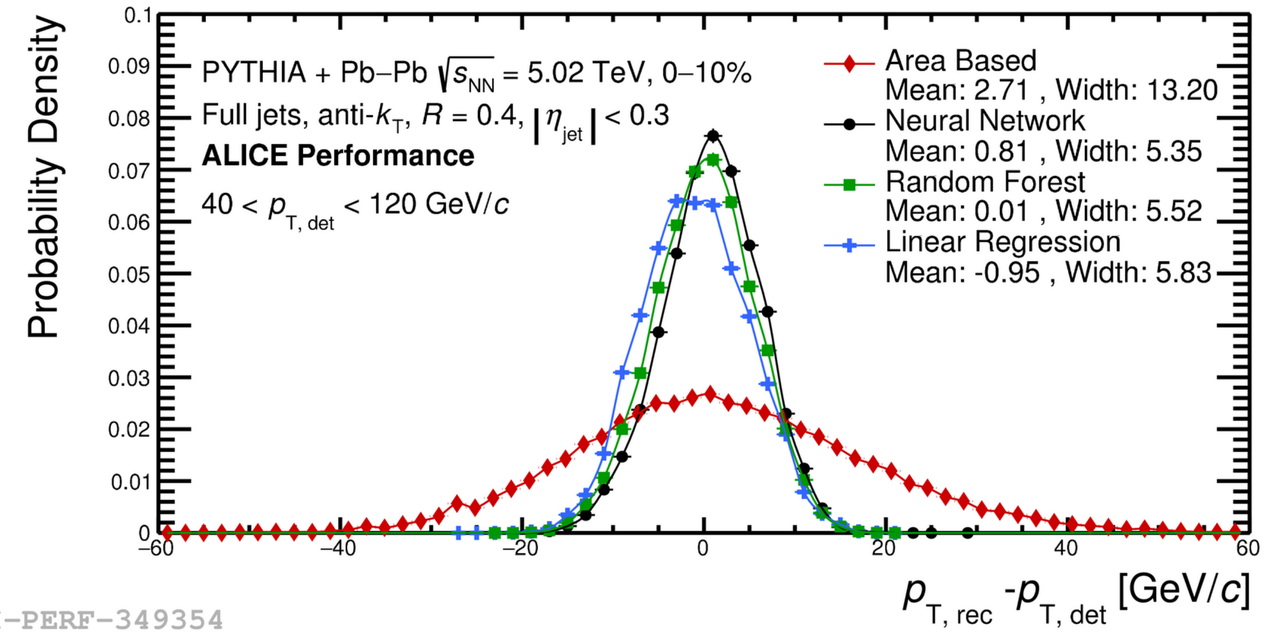
- Applied to charged or full jets (contain charged tracks and neutral clusters, measured in the TPC and EMCal, respectively)

ML approach: method

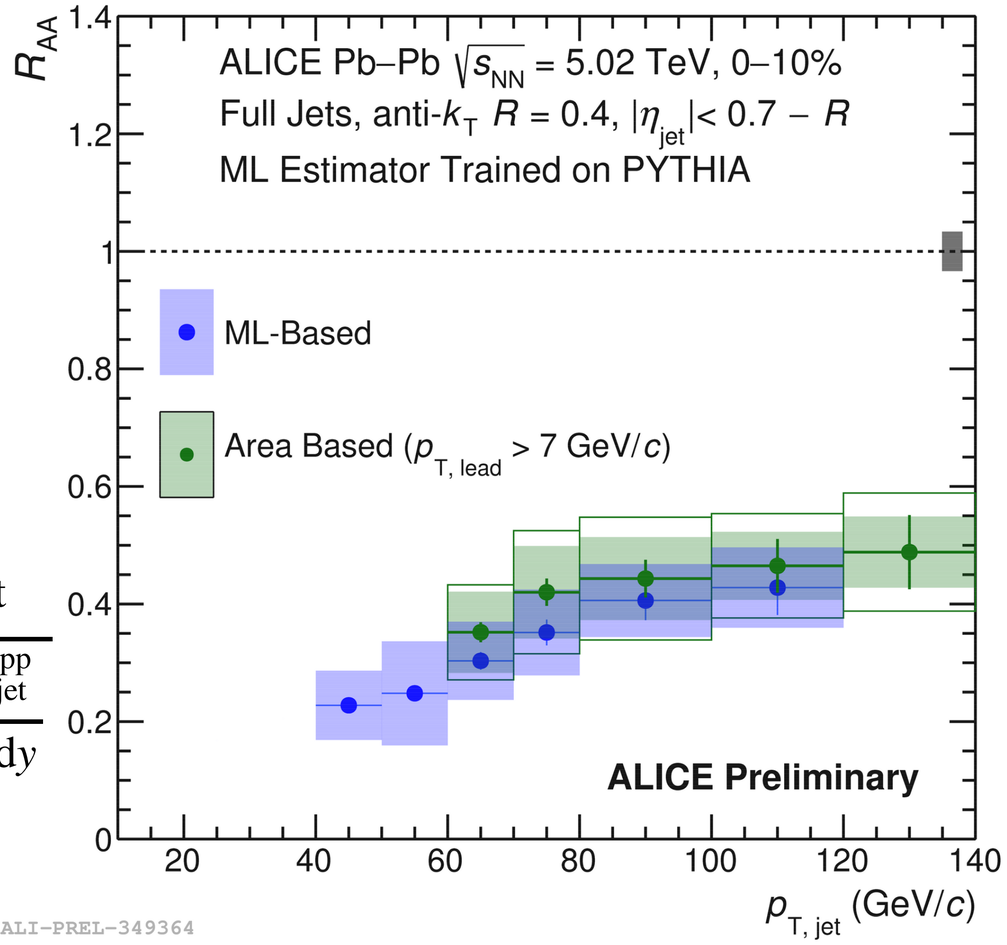


- Embedding pp PYTHIA events into real Pb-Pb data
- Optimized for method performance and how important/correlated they are
  - Area-based corrected
  - 12 leading constituents
  - jet angularity
  - number of constituents
  - 10% train
  - 90% test
  - shallow neural network (100, 100, 50)
  - random forest
  - linear regression
- Regression task to predict the corrected jet p_T
  - “truth” = detector level PYTHIA jet p_T

ML approach: low p_T inclusive jets



- ML method improves resolution over area-based -> unfold to lower p_T



- The R_{AA} in 0-10% ALICE data with the ML-based and area-based method with a leading track bias are consistent!

$$R_{AA} = \frac{\frac{1}{N_{event}} \frac{d^2 N_{jet}^{PbPb}}{dp_T dy}\Big|_{cent}}{\frac{< N_{coll} >_{cent}}{\sigma_{inel}} \times \frac{d^2\sigma_{jet}^{pp}}{dp_T dy}}$$

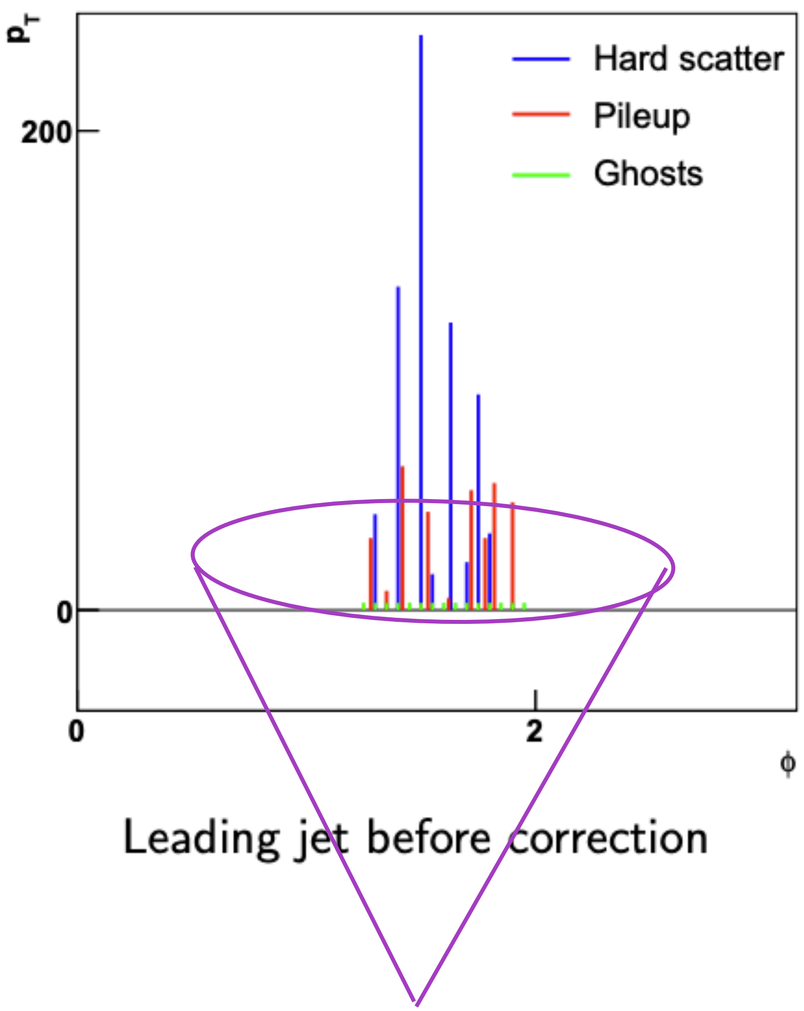
p_T reach now down to 40 GeV/c, reduced uncertainties

Inclusive jets at large R

CMS-PAS-HIN-18-014

- Larger statistics datasets -> differential measurement at high jet p_T
- Jet-by-jet constituent subtraction used for pile-up removal in pp with inclusion of flow modulation improves resolution
  - Estimate background energy density in each jet

JHEP 1908 (2019) 175

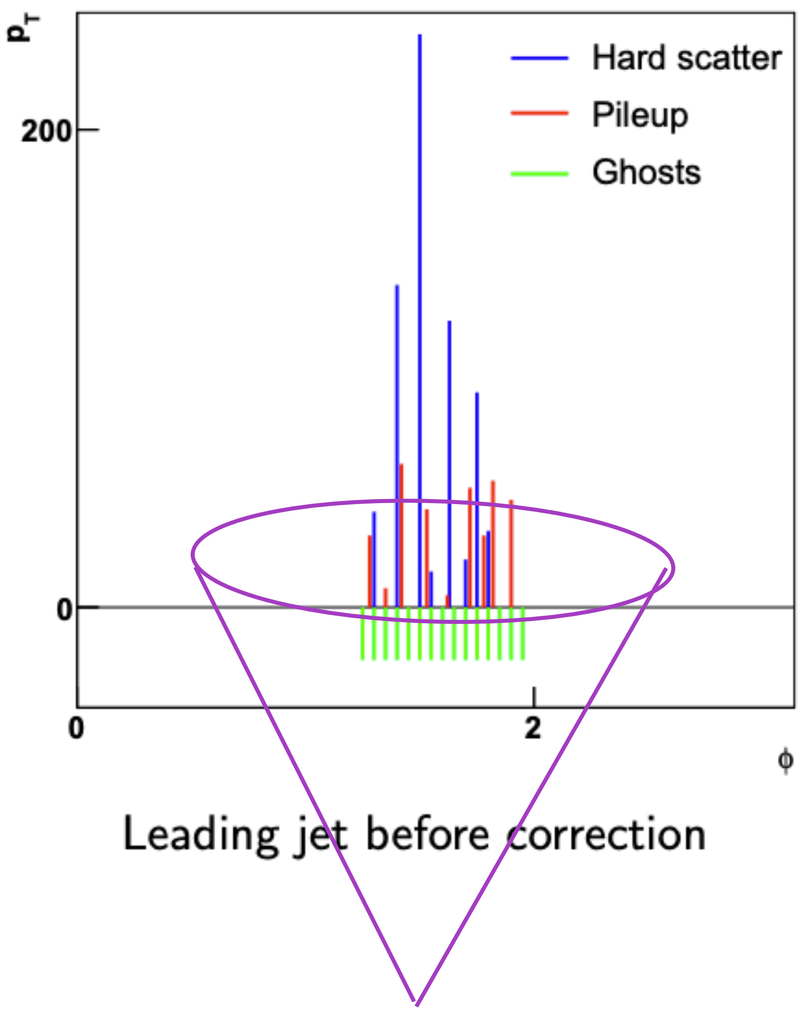


Inclusive jets at large R

CMS-PAS-HIN-18-014

- Larger statistics datasets -> differential measurement at high jet p_T
- Jet-by-jet constituent subtraction used for pile-up removal in pp with inclusion of flow modulation improves resolution
  - Estimate background energy density in each jet
  - Add infinitesimal ghosts having negative the background contribution

JHEP 1908 (2019) 175

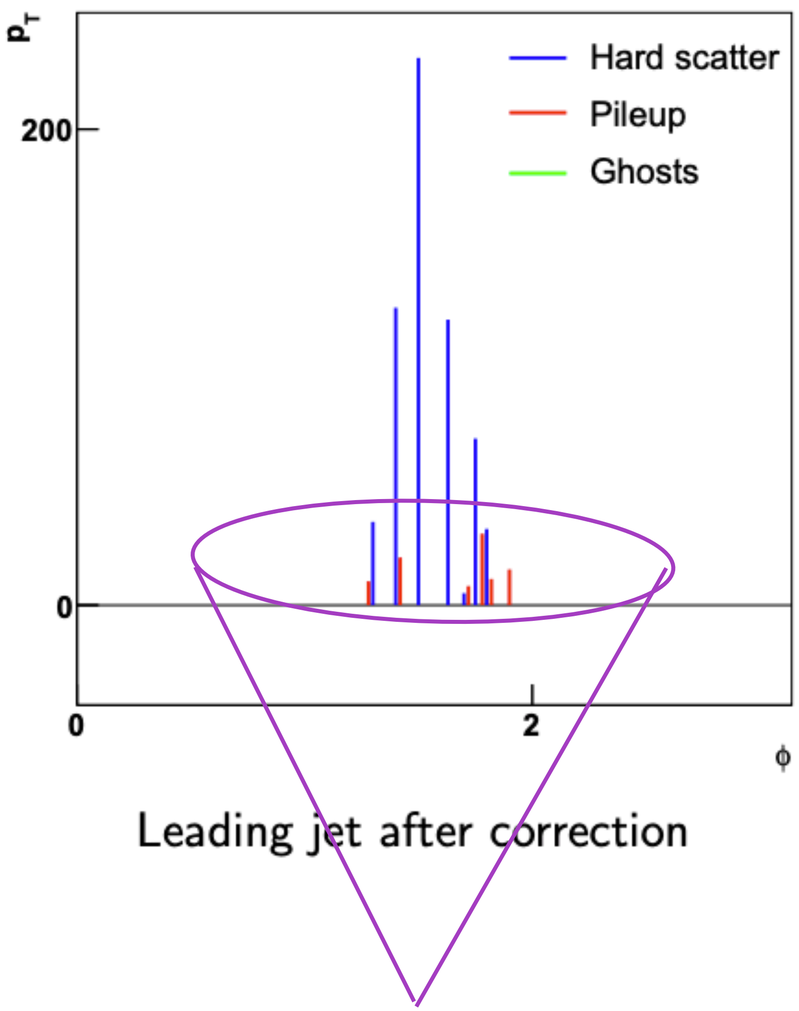


Inclusive jets at large R

CMS-PAS-HIN-18-014

- Larger statistics datasets -> differential measurement at high jet p_T
- Jet-by-jet constituent subtraction used for pile-up removal in pp with inclusion of flow modulation improves resolution
  - Estimate background energy density in each jet
  - Add infinitesimal ghosts having negative the background contribution
  - Combine ghosts and particles based on how close they are and remove ones with < 0 momentum

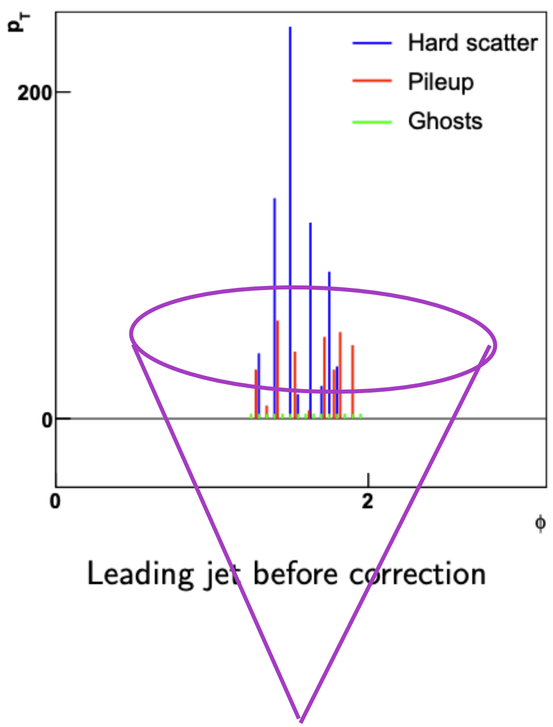
JHEP 1908 (2019) 175

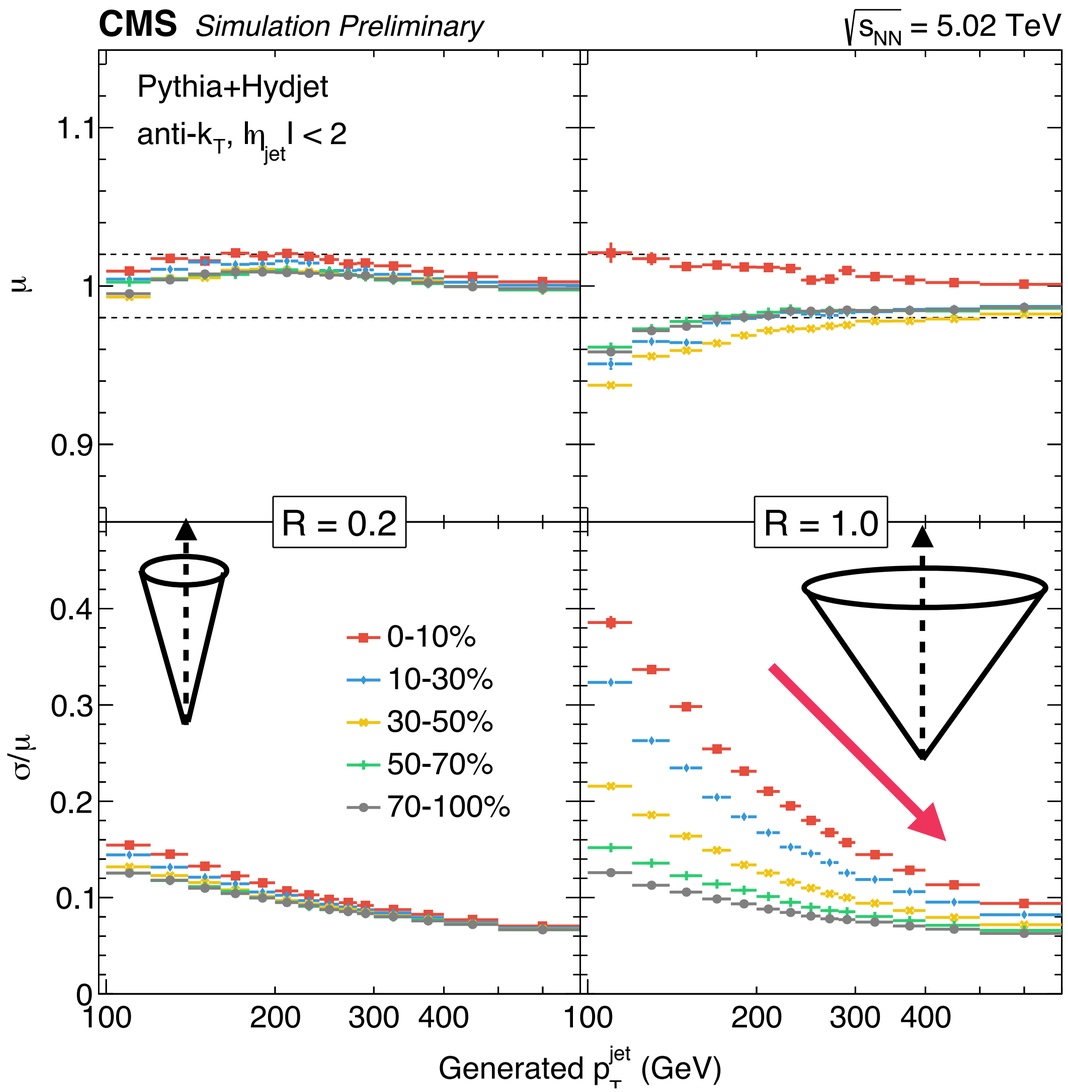


Inclusive jets at large R

CMS-PAS-HIN-18-014

- Larger statistics datasets -> differential measurement at high jet p_T
- Jet-by-jet constituent subtraction used for pile-up removal in pp with inclusion of flow modulation improves resolution
  - Estimate background density in each jet and remove using ghosts





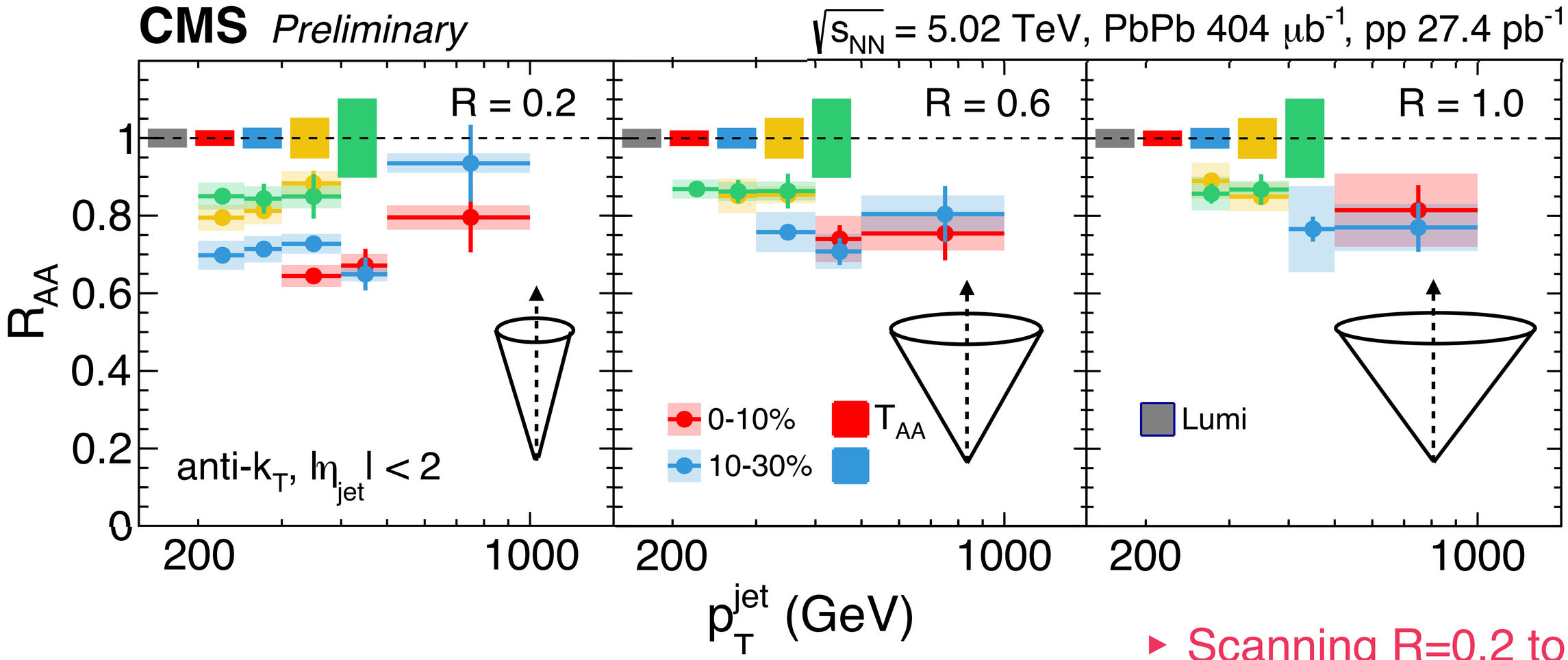
- Higher jet p_T -> improved resolution allowing for measurement at large R

R_{AA} at large R

$$R_{AA} = \frac{\frac{1}{N_{event}} \frac{d^2 N_{jet}^{PbPb}}{dp_T dy}\Big|_{cent}}{\frac{< N_{coll} >_{cent}}{\sigma_{inel}} \times \frac{d^2 \sigma_{jet}^{pp}}{dp_T dy}}$$

CMS-PAS-HIN-18-014

- R_{AA} at larger R and high p_T now possible!

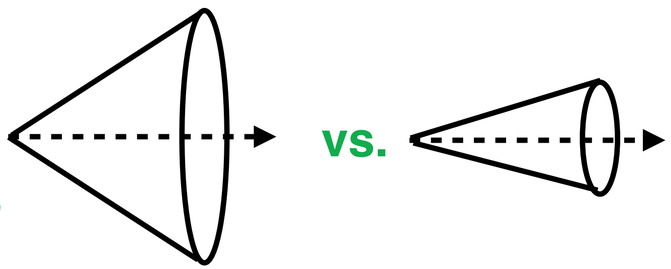


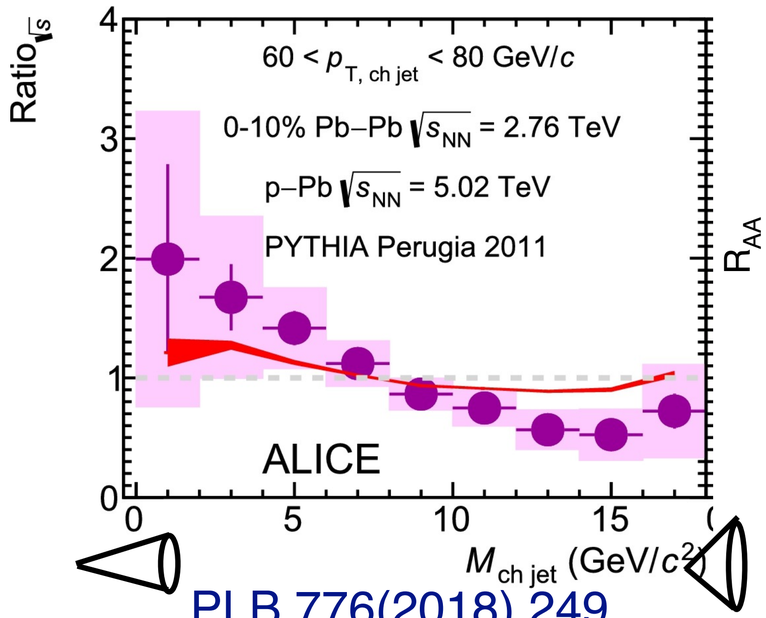
- Scanning R=0.2 to 1.0!

Jet substructure

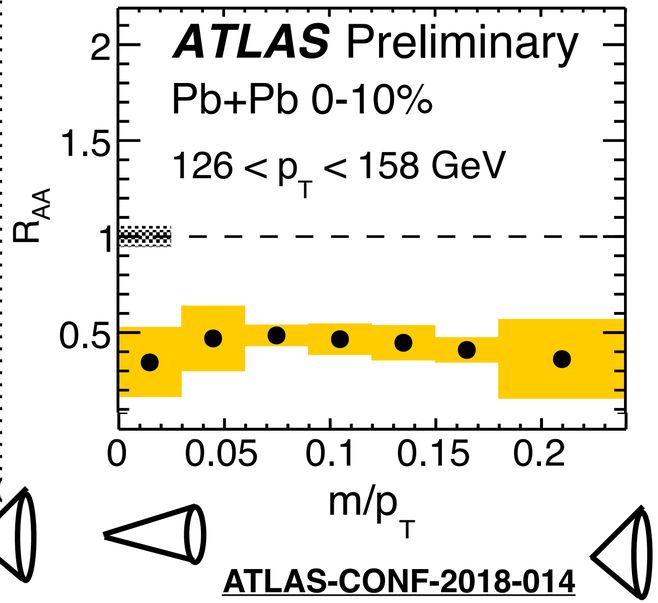
- Many jet substructure measurements in HIs
- Ex: jet mass

How wide are the jets?



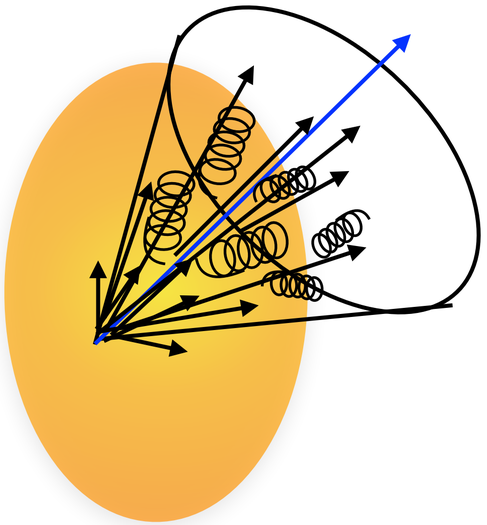


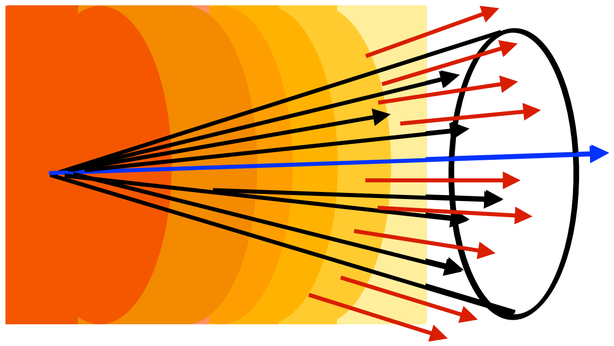
PLB 776(2018) 249



ATLAS-CONF-2018-014

- Insensitive to medium effects?
  - Possible cancellation of effects:
    - Softening and broadening of constituents
      - Inside cone->larger mass
      - Outside cone->smaller mass
    - Medium recoil->larger mass

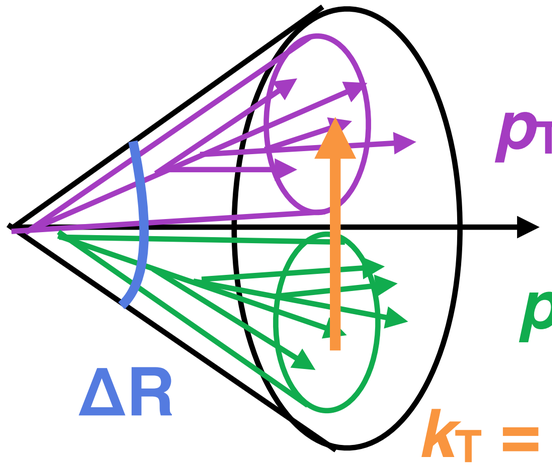




Jet splitting tools

- Methods from pp select on harder splittings inside the jet to remove the soft background contribution:

$$z_g = \frac{\min(p_{Ti}, p_{Tj})}{p_{Ti} + p_{Tj}}$$



$p_{T1} = (1-z)p_T$

$p_{T2} = zp_T$

$k_T = p_{T2}\sin(\Delta R)$

- **Jet grooming:** Soft Drop
  - Recluster jets with Cambridge/Aachen (C/A)* to enforce angular ordering and enforce SD condition

    *JHEP 9708:001,1997

$$z_g > z_{cut}\theta^{\beta} \qquad \theta = \frac{\Delta R}{R}$$

  - Default condition: **$z_{cut} = 0.1$**, **$\beta = 0$**

- **Jet trimming:**
  - Recluster smaller jets already found at R=0.2 into larger R jets at R=1.0 using the k_T algorithm and take the first splitting

    ATLAS-CONF-2019-054

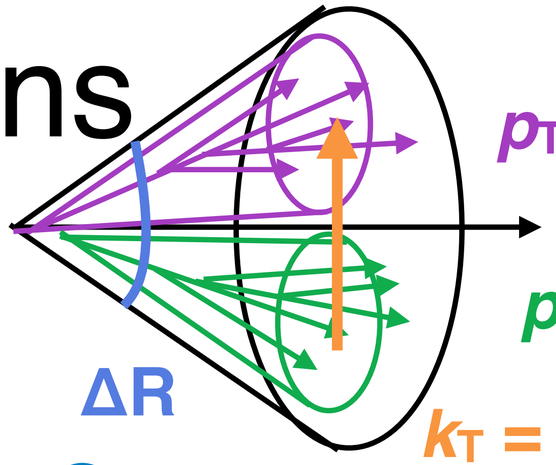
  - Measures the k_T splitting scale $\sqrt{d_{12}}$

$$\sqrt{d_{12}} = p_{T2}\Delta R_{12}$$

Jet splitting tools: HI collisions

- Methods from pp select on harder splittings inside the jet to remove the soft background contribution:

$$z_g = \frac{\min(p_{Ti}, p_{Tj})}{p_{Ti} + p_{Tj}}$$



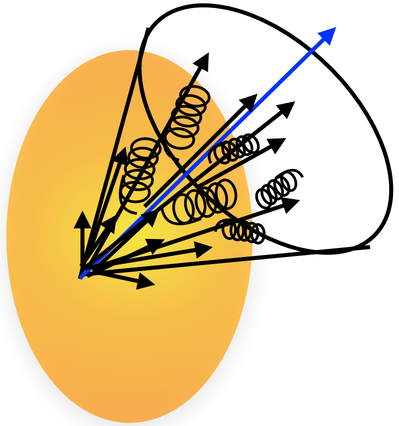
$p_{T1} = (1-z)p_T$

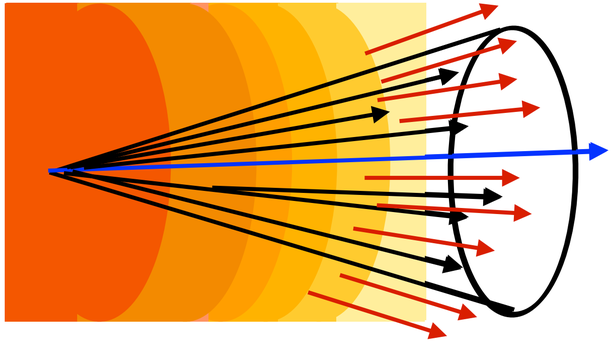
$p_{T2} = zp_T$

$k_T = p_{T2}\sin(\Delta R)$

- **Jet grooming:** Soft Drop
- **Jet trimming:**

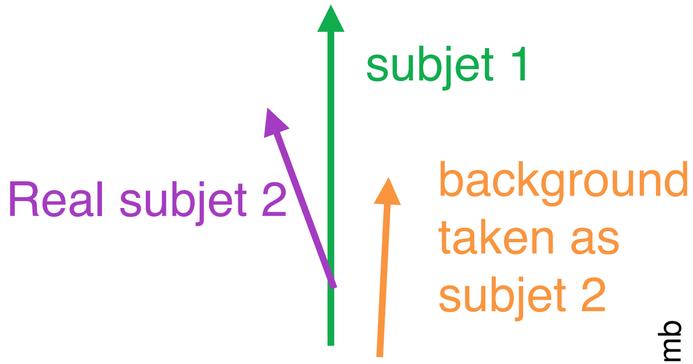
- Applying to HIs:
  - Helps remove soft background
  - Removes soft signal from softening of jet constituents and medium response!
    - Separate medium effects that might cancel in other variables
    - Removes some of the physics signal



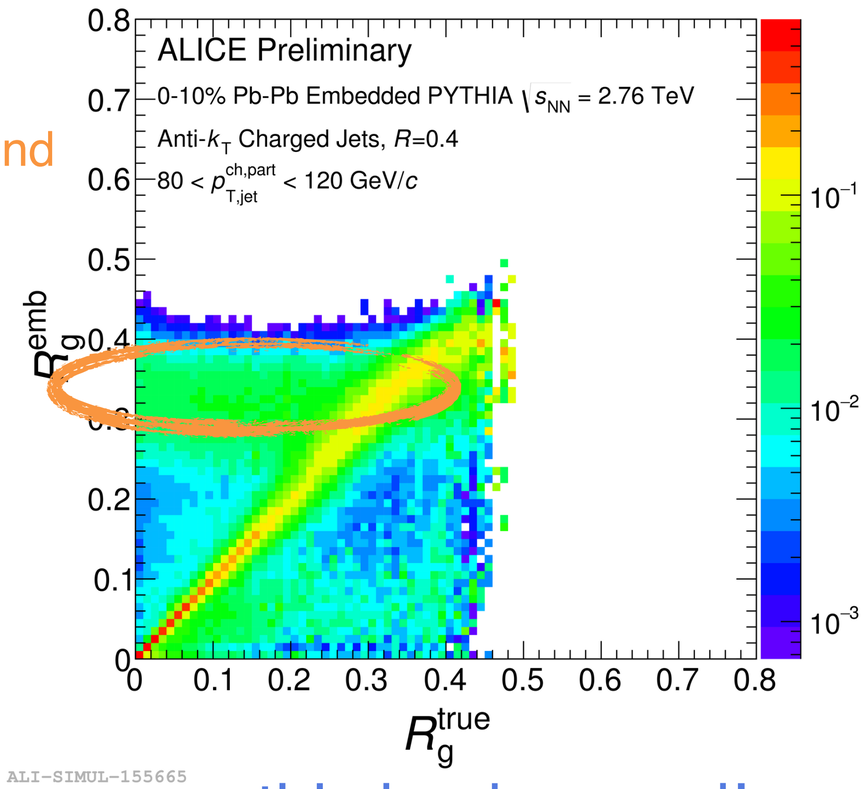


Background treatment

- Uncorrelated background leads to subjets picked up as incorrect splittings



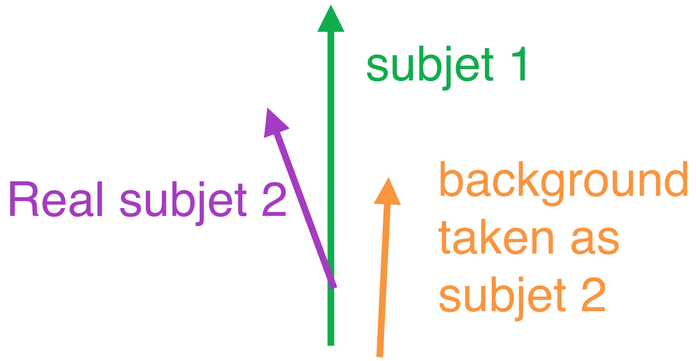
- Response built from embedding Pythia into real minimum bias Pb-Pb data to mimic background
  - *Non-diagonal response prohibits unfolding*



- Unfolding desired to compare to models: need to suppress this background!

Background treatment

- Uncorrelated background leads to subjets picked up as incorrect splittings



- Unfolding desired to compare to models: need to suppress this background!

- Solutions:

1. Smaller jet radii
2. Semi-central collisions

ALICE

R=0.2 jets between 60-80 GeV

R=0.4 jets between 60-80 GeV

More symmetric splittings:

SD $z_{cut} = 0.2$

SD $z_{cut} = 0.4$

First unfolded jet splittings at lower p_T

3. Higher jet p_T

R=1.0 jets between 200-500 GeV

ATLAS EXPERIMENT

p_T (R=0.2) > 35 GeV

First jet splittings at large R

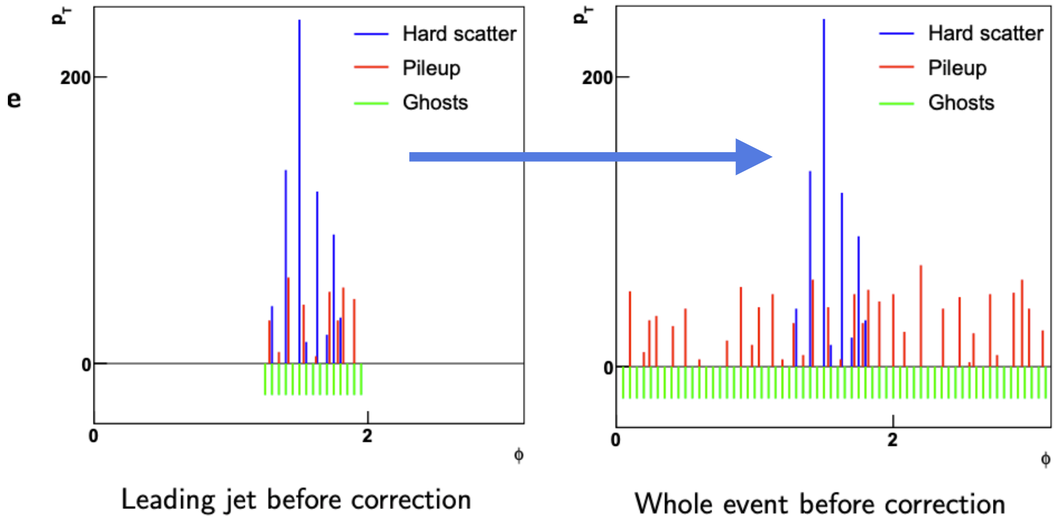
Both constrain models in different phase spaces!

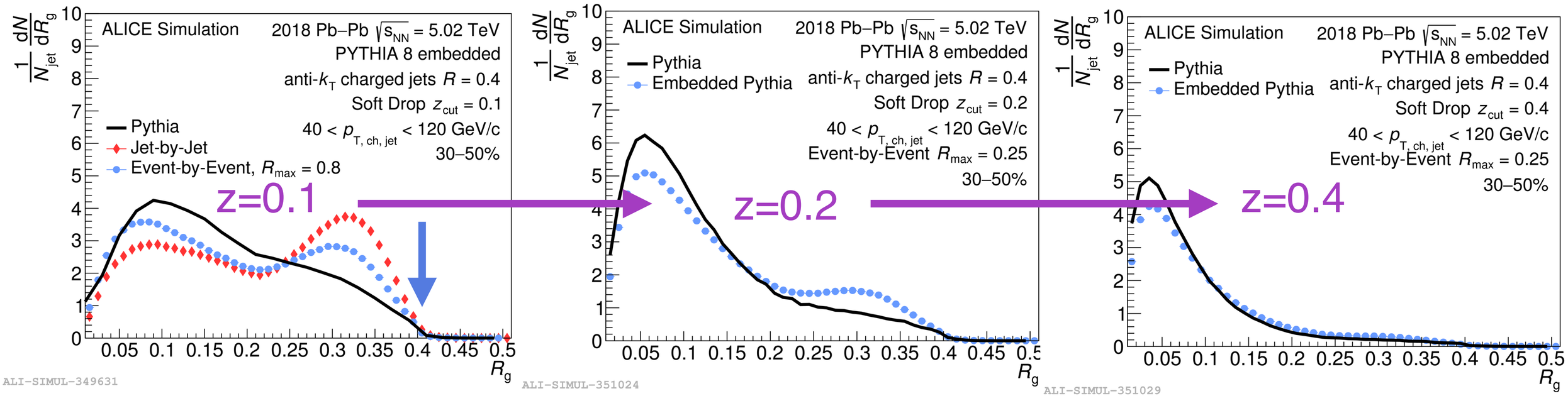
Jet splittings: SD grooming

$60 < p_T^{ch} < 80$ GeV

JHEP 1908 (2019) 175

- ALICE uses an improved background subtraction:
  - Clear secondary background peak
  - Event-by-event constituent subtraction instead of jet-by-jet reduces the peak
  - Tighter SD condition reduces remaining peak





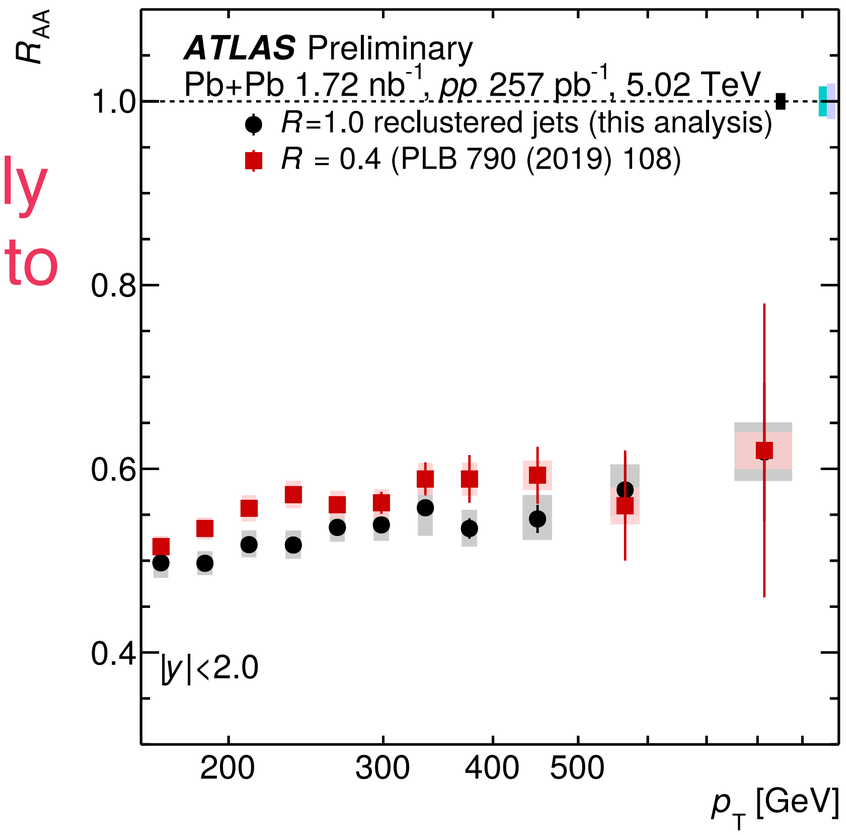
Jet splittings: large R

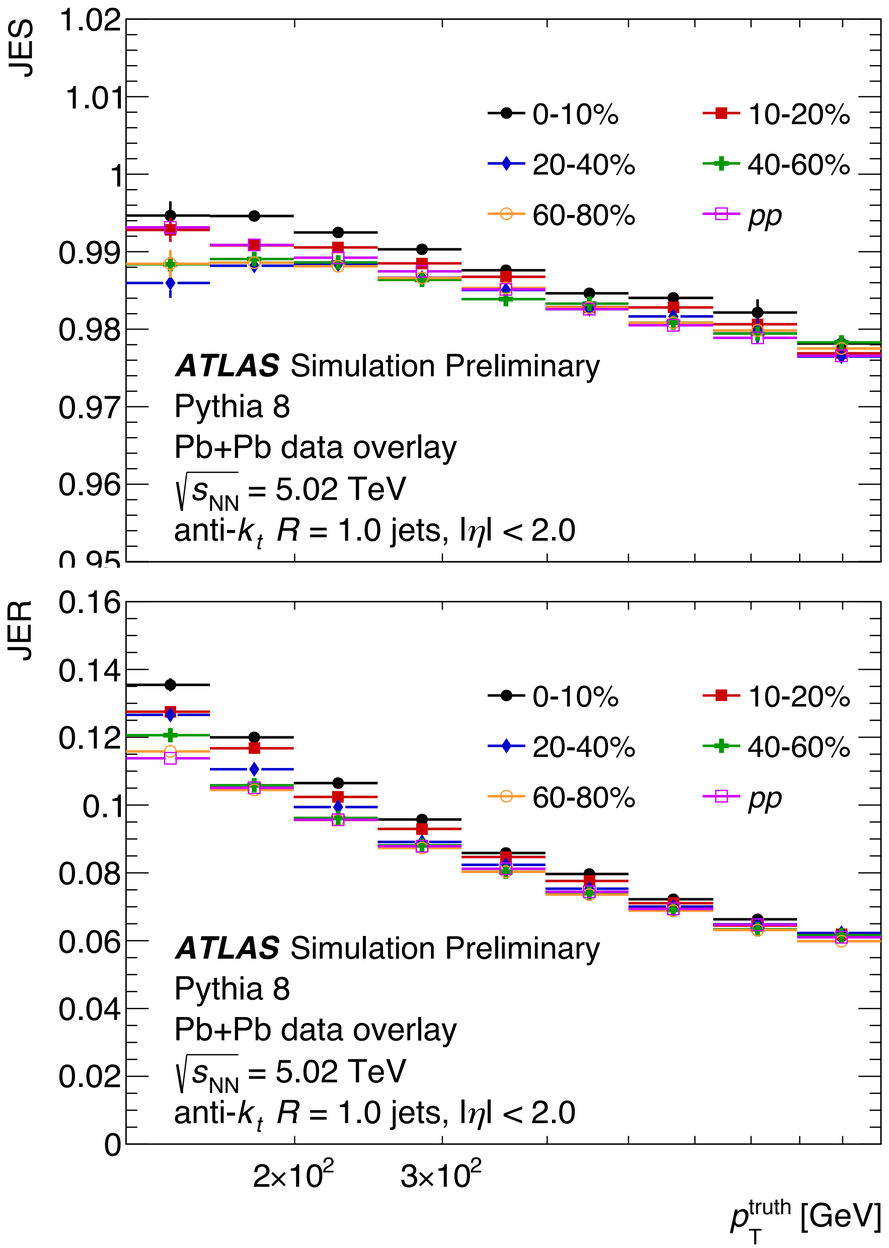
$200 < p_T < 500$ GeV

ATLAS-CONF-2019-054 R= 1.0

- Combining R=0.2 jets into R=1.0 jets allows ATLAS to take advantage of precise background subtraction for R=0.2 jets
  - JER is ~10% at 200 GeV for 0-10%

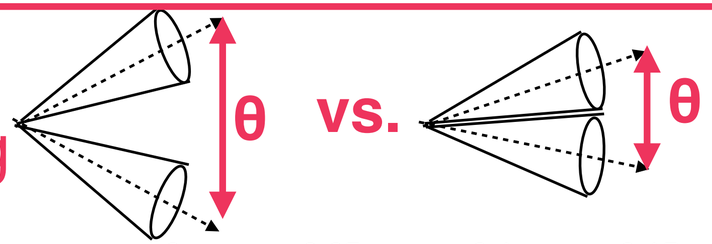
- Reclustered R=1.0 jets have qualitatively similar suppression to standard R=0.4 jets
  - Energy radiated between subjets removed



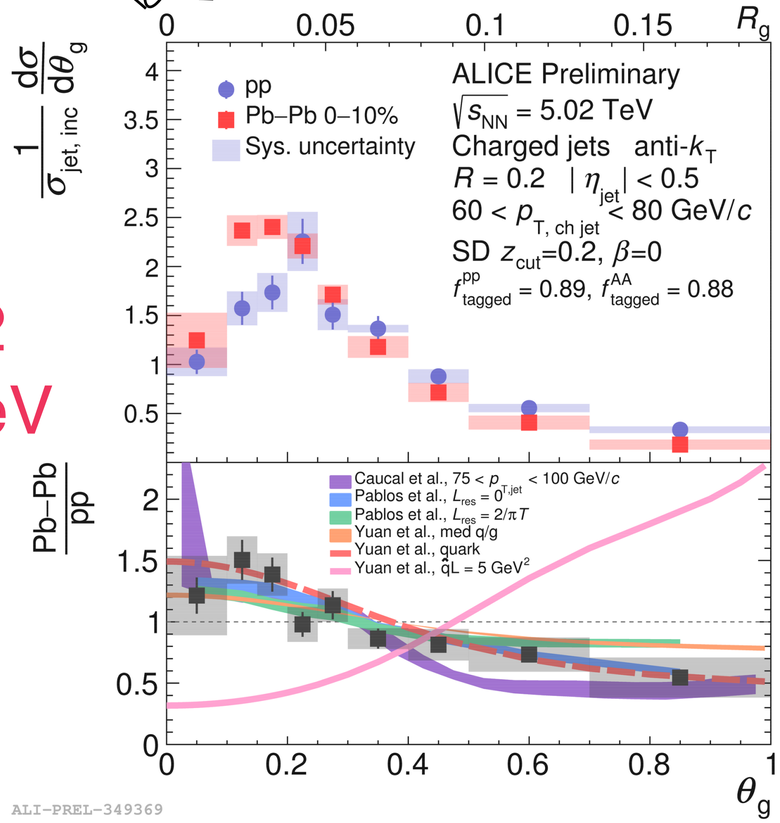


Jet splittings: results

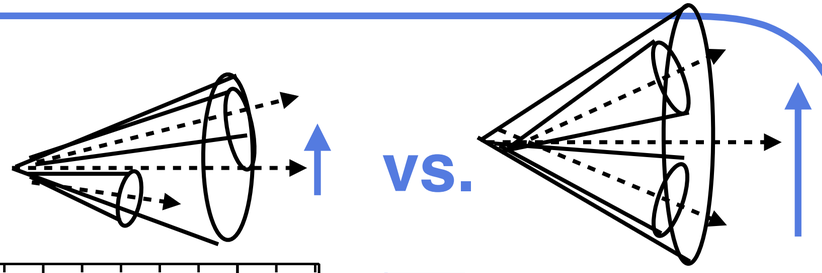
R_g: grooming angle



0-10%
R=0.2
SD z=0.2
60-80 GeV

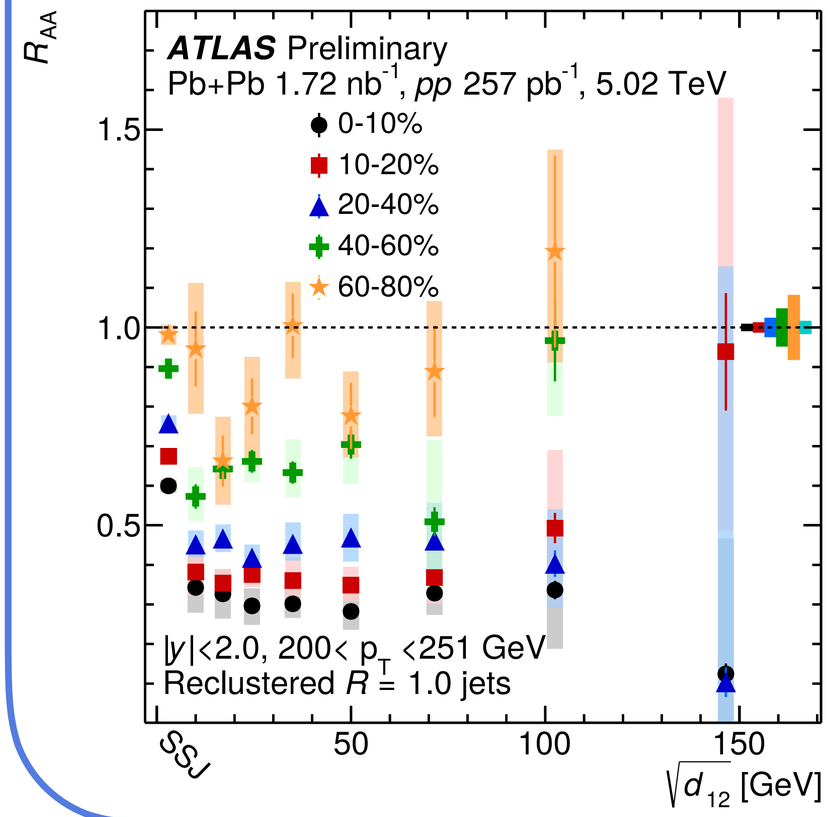


k_T splitting scale



$\sqrt{d_{12}} = p_{T2}\Delta R_{12}$

0-10%
R=1.0
200-500 GeV



Both demonstrate a modification in Pb-Pb vs. pp

Fully unfolded in 2D so precise comparison can be made to models!

Conclusion

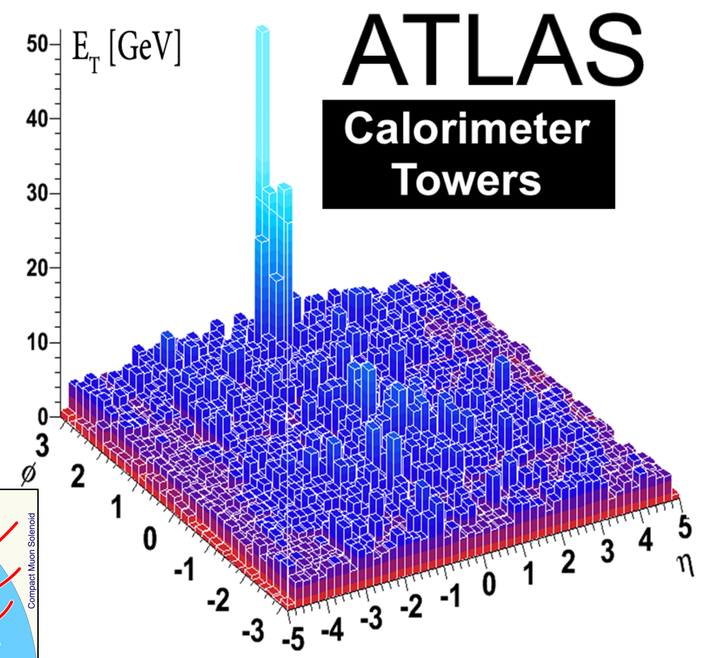
- Significant progress in background subtraction techniques:
  - Opportunity to measure inclusive jets at lower p_T and larger R
- New developments in jet substructure tools:
  - Selecting on hard jet splittings seems to be sensitive to jet quenching
  - Improved background removal techniques allow results to be unfolded at both lower p_T and larger R
- Improving jet tools allow for direct comparison of unfolded results to theoretical calculations to help constrain models
- New tools constantly being developed
  - looking forward to future measurements!

Zoom room for additional questions over coffee! *link here*

Backup

Heavy-ion background subtraction

- Large uncorrelated background due to the underlying event (UE) that contributes energy inside the jet cone
  - Fluctuations can be of the order of the jet itself
- Have to be careful with fake jets due to upward UE fluctuations



PRL 105 (2010) 252303

ALICE estimates ρ from tracks with the two highest jets in the event removed and subtracts jet-by-jet

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}} - \rho \mathrm{A}$$

ATLAS estimates ρ from calorimeter towers with jets removed, taking into flow modulation and subtracting tower-by-tower with an iteration step

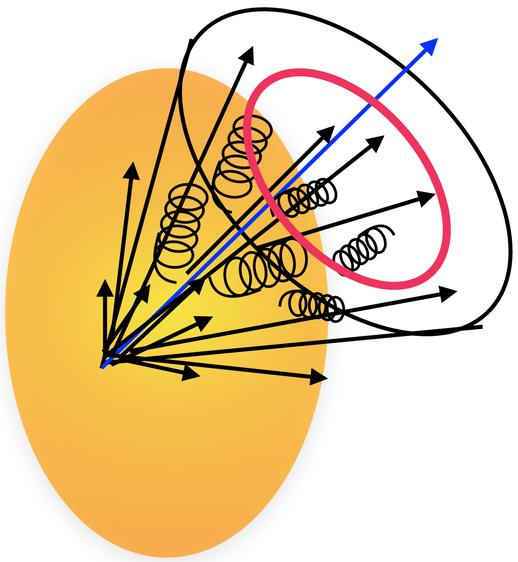
$$E_{Tj}^{sub} = E_{Tj} - A_j \rho_i(\eta_j)(1 + 2v_{ni}\cos 2(\phi_j - \psi_n))$$

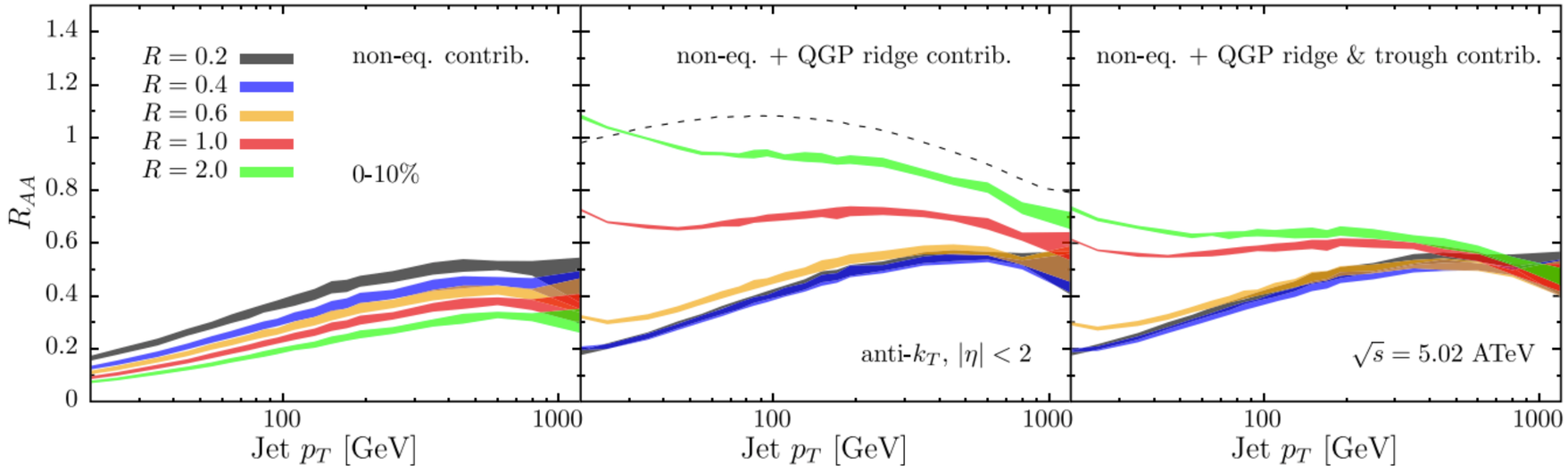
CMS estimates ρ with flow modulation and uses the constituent subtraction method to perform the subtraction

JHEP 1908 (2019) 175

Large R/low p_T motivation

- Larger radii
  - ➡ Possible recovery of the jet energy because of out-of-cone radiation
  - ➡ Possible difference in modification for larger jets
- Lower jet p_T
  - ➡ Probes different scale and modification expected to be different
  - ➡ Connection to RHIC
  - ➡ Difference between jet radii could be larger at lower p_T



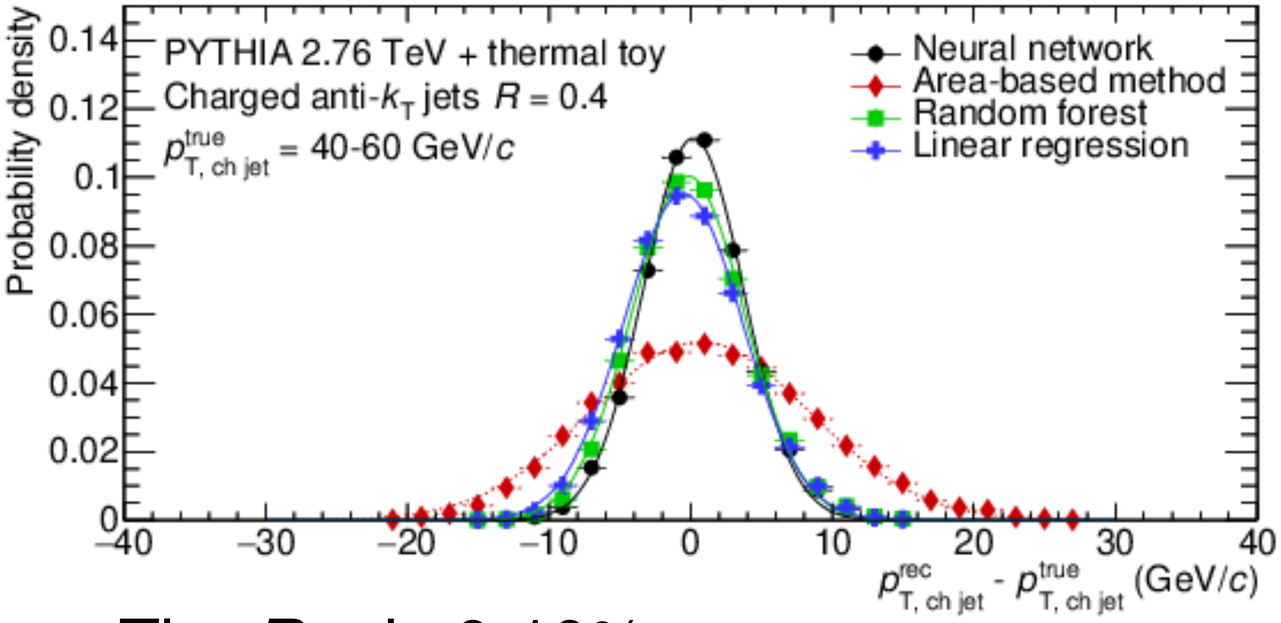


Hybrid model

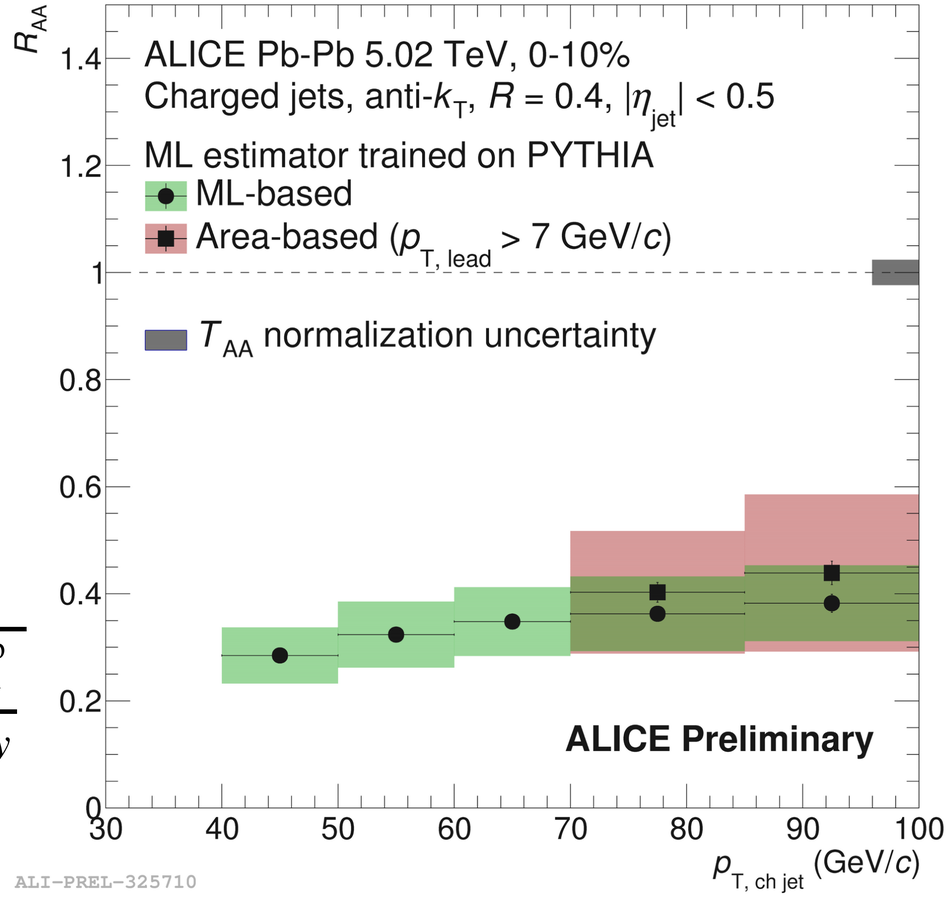
Phys.Rev.Lett. 124 (2020)

ML approach: charged jets

Phys. Rev. C 99, 064904 (2019)



- ML method improves resolution over area-based -> unfold to lower p_T



- The R_{AA} in 0-10% ALICE data with the ML-based and area-based method with a leading track bias are consistent!

$$R_{AA} = \frac{\frac{1}{N_{event}} \frac{d^2 N^{PbPb}_{jet}}{dp_T dy}\Big|_{cent}}{\frac{< N_{coll} >_{cent}}{\sigma_{inel}} \times \frac{d^2\sigma^{pp}_{jet}}{dp_T dy}}$$

p_T reach now down to 40 GeV/c, reduced uncertainties

Laura Havener, Yale University

ML configurations

- Regression task that is prioritizing a simple model
- Implemented in scikit-learn with defaults unless otherwise specified

1. Shallow neural network
   - Shallow, three-layer network with [100, 100, 50] nodes
   - ADAM optimizer, stochastic gradient descent algorithm
   - Nodes/neutrons activated by a ReLU activation function
2. Linear regression
   - Normalization set to the default
3. Random Forest
   - Ensemble of 30 decision trees
   - Maximum number of features set to 15

ML: features for training

- In order to determine the features for training, ask two questions:

1. How important is the feature in this model? -> feature score
   - Higher score, more often it is used in training

2. How correlated is the feature with other features?

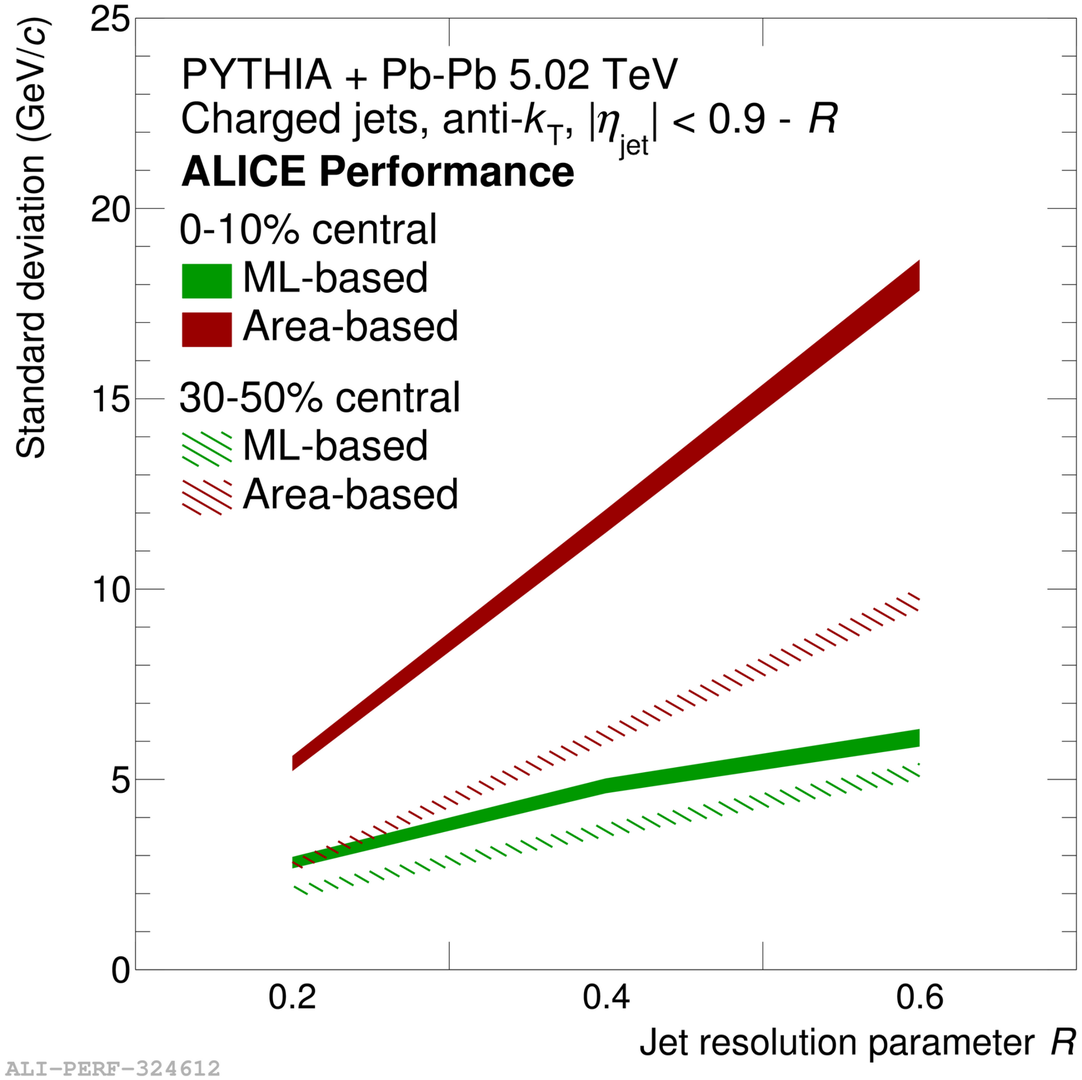
Charged Particle Jets

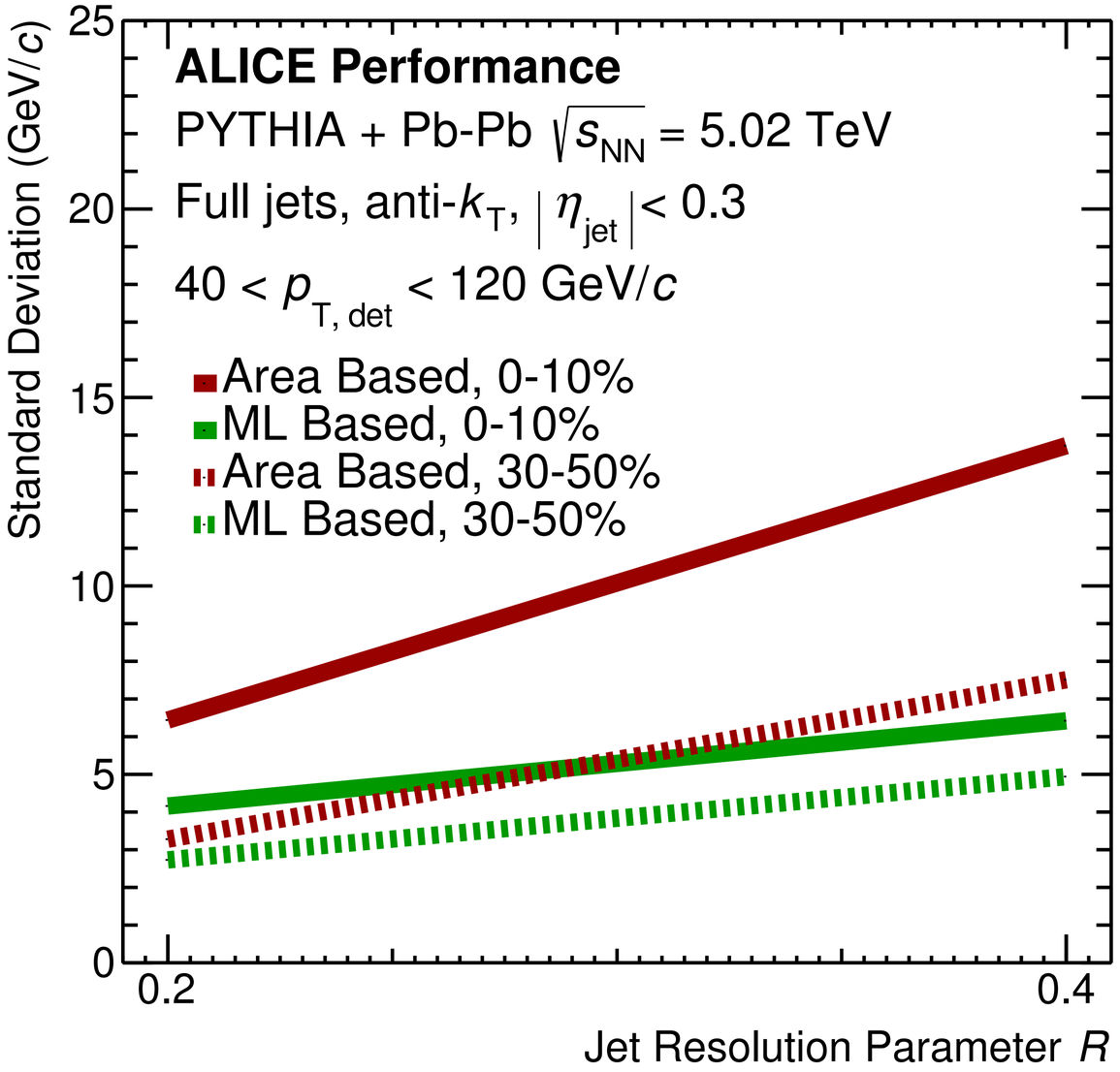
Feature	Score	Feature	Score
Jet p_T (no corr.)	**0.1355**	$p^1_{T,const}$	0.0012
Jet mass	0.0007	$p^2_{T,const}$	**0.0039**
Jet Area	0.0005	$p^3_{T,const}$	0.0015
Jet p_T (area based corr.)	**0.7876**	$p^4_{T,const}$	0.0011
LeSub	0.0004	$p^5_{T,const}$	0.0009
Radial moment	0.0005	$p^6_{T,const}$	0.0009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

Phys. Rev. C 99, 064904 (2019)

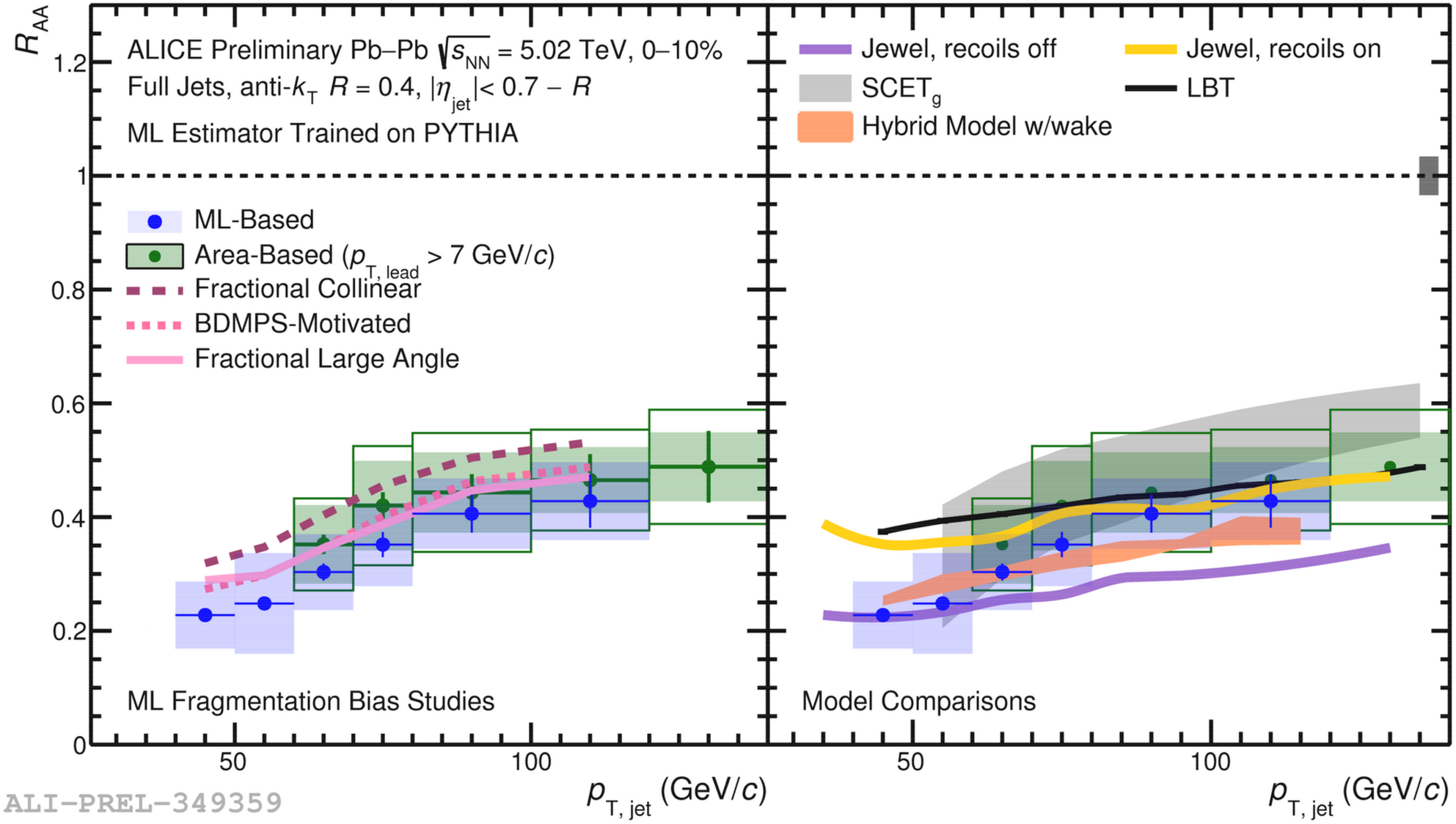
- Iteratively remove unimportant and/or highly correlated feature!

ML approach: performance





ML approach: fragmentation bias



Laura Havener, Yale University

ML approach: R_{AA}

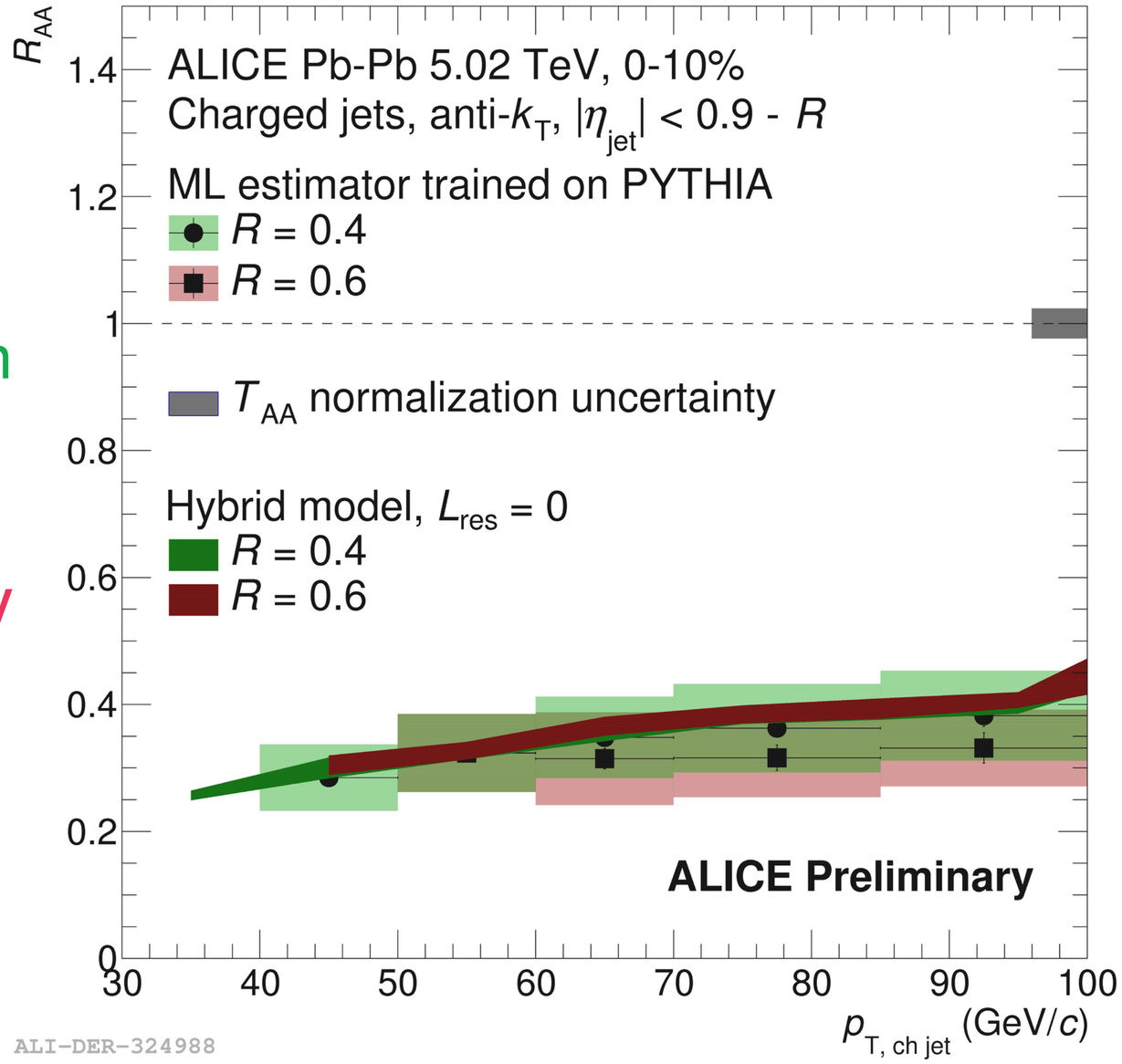
- Evaluate the R_{AA} in 0-10% ALICE data using the ML-based estimator and compare to the area-based method with a leading track bias (consistent!)

➡ See the p_T reach is much lower (down to 40 GeV) and the systematics are reduced!

➡ R=0.6 is also possible down to 60 GeV

- Comparison to hybrid model!

$$R_{AA} = \frac{\frac{1}{N_{event}} \frac{d^2 N_{jet}^{PbPb}}{dp_T dy}\Big|_{cent}}{\langle T_{AA} \rangle_{cent} \times \frac{d^2 \sigma_{jet}^{pp}}{dp_T dy}}$$



Constituent subtraction (CS)

- Estimate background density in each jet or event

$$\rho = \mathrm{med}\left(\frac{p_{\mathrm{T,jet}}^{\mathrm{raw,i}}}{A_{\mathrm{jet}}^{\mathrm{i}}}\right)$$

$$\rho_m = \mathrm{med}\left(\frac{m_i}{A_{\mathrm{jet}}^{\mathrm{i}}}\right)$$

- Add infinitesimally small ghosts to the event
- Set the p_T for each ghost to negative values

$$p_{\mathrm{T,g}} = A_{\mathrm{g}}\rho$$

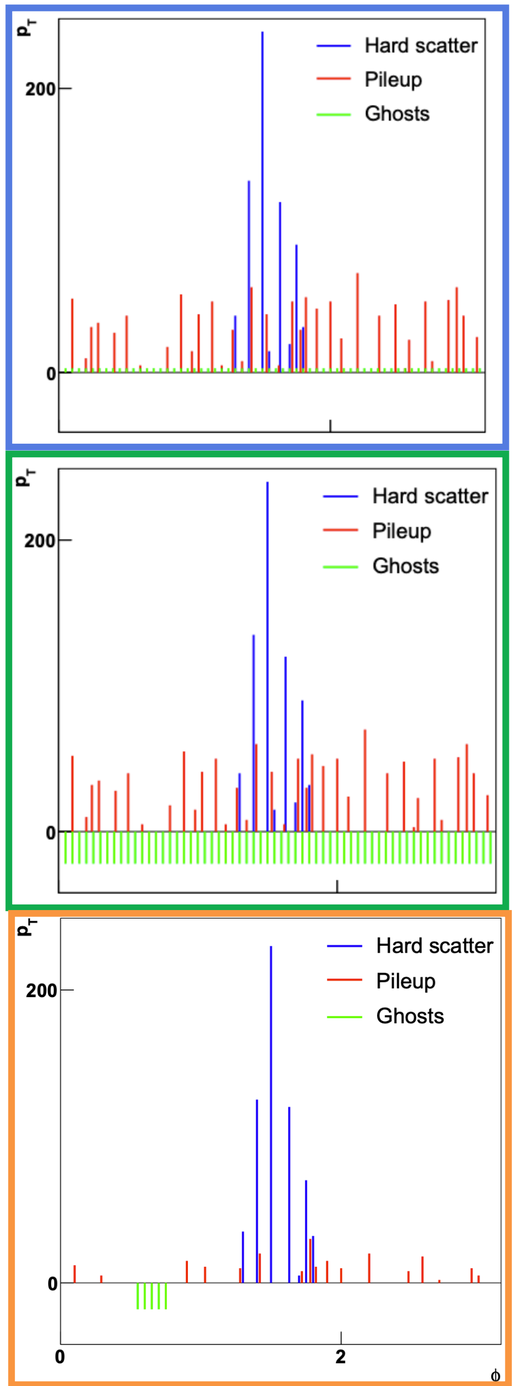
$$m_{\mathrm{g}} = A_{\mathrm{g}}\rho_m$$

- Calculate distance between each particle and ghost for each pair and sort in ascending order
- Iteratively change the momentum and mass of each ghost/particle until no more pairs remain

if ($p_T > p_T^g$) $p_T = p_T - p_T^g$ $\quad$ $p_T^g = p_T^g - p_T$

$p_T^g = 0$ $\quad$ $p_T = 0$

- Discard particles with 0 momentum

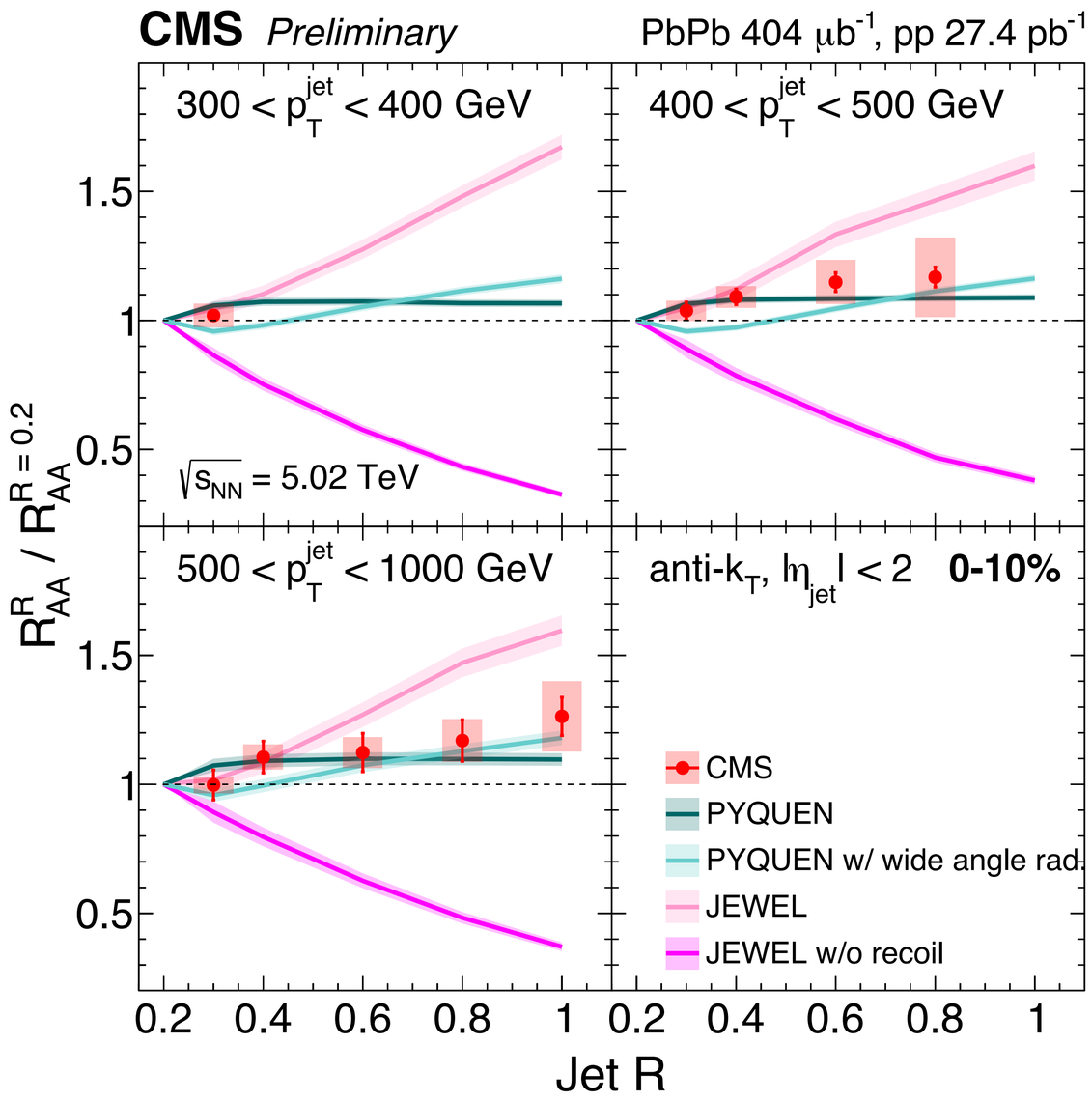


JHEP 1908 (2019) 175

R_{AA} at large R: model comparisons

CMS-PAS-HIN-18-014

- Compare R_{AA} at larger R to R_{AA} at R=0.2 as a function of R
  - Can be used to constrain the vast number of jet quenching models!
    - Event generators

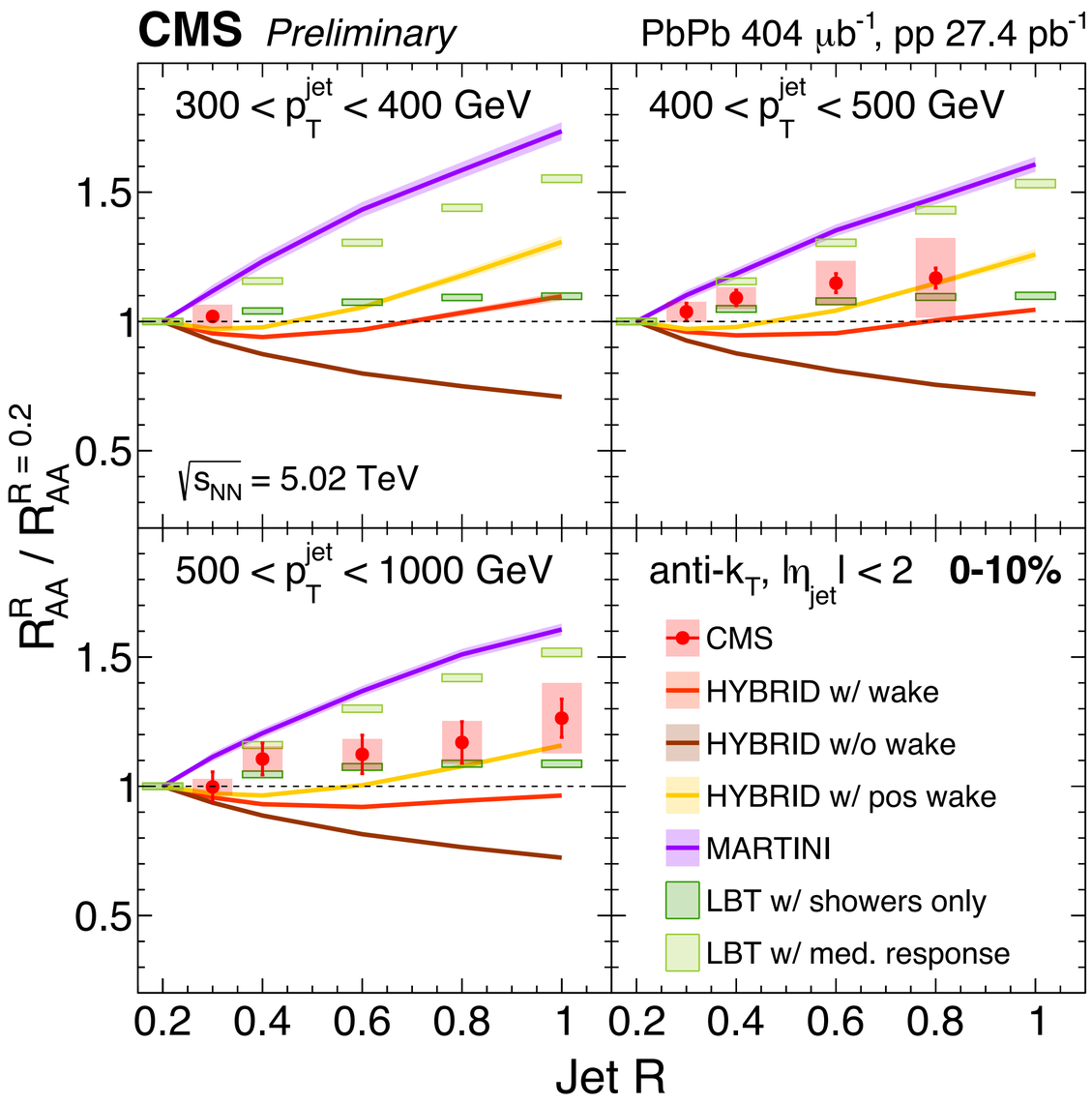


R_{AA} at large R: model comparisons

CMS-PAS-HIN-18-014

- Compare Raa at larger R to Raa at R=0.2 as a function of R

- Can be used to constrain the vast number of jet quenching models!
  - Event generators
  - Monte carlos



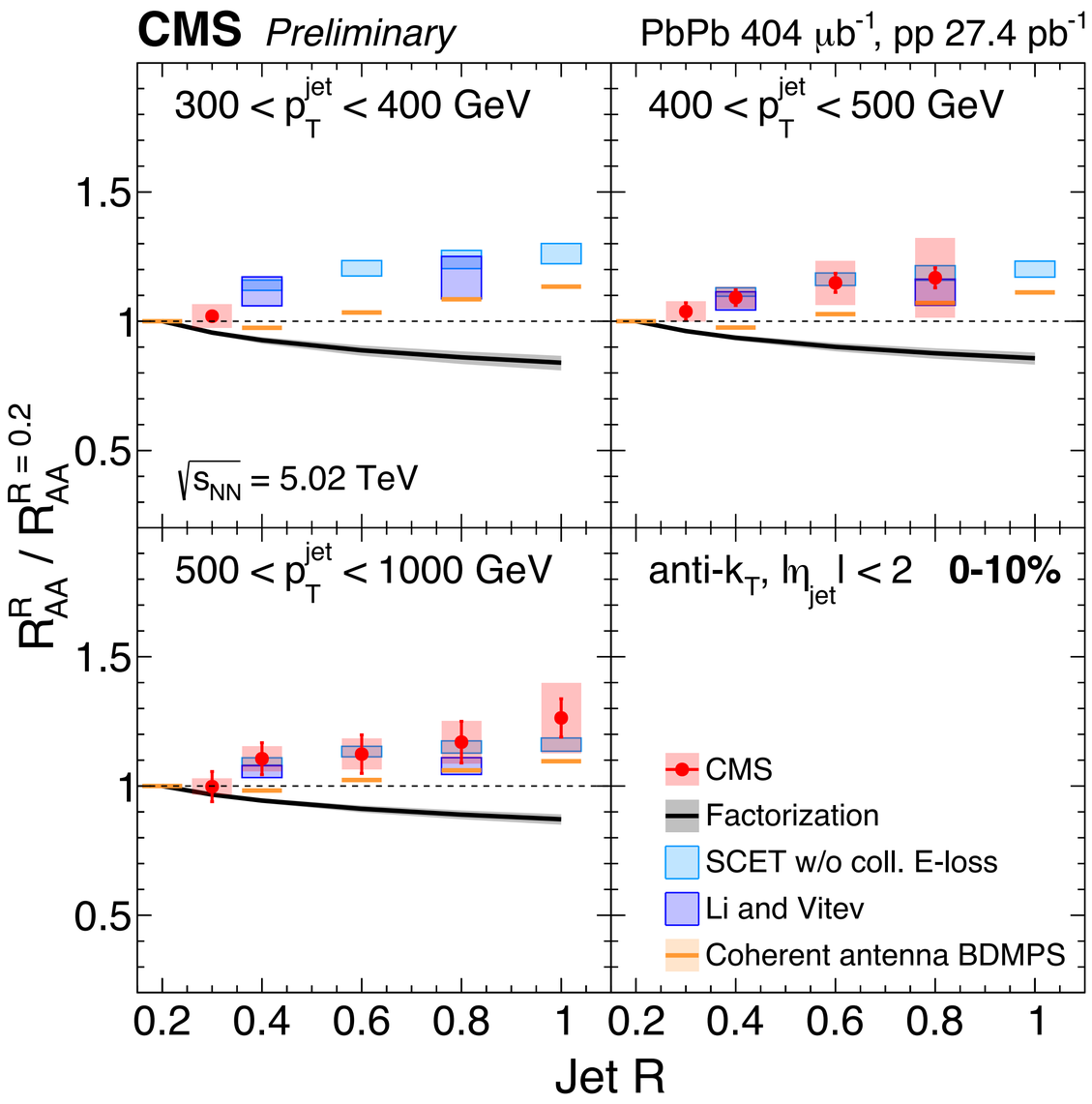
R_{AA} at large R: model comparisons

CMS-PAS-HIN-18-014

- Compare Raa at larger R to Raa at R=0.2 as a function of R

- Can be used to constrain the vast number of jet quenching models!
  - Event generators
  - Monte carlos
  - Calculations

- See a large spread in the models, allowing us to learn something about the various underlying mechanisms at play in different models



Exploring the Lund Plane: in vacuum

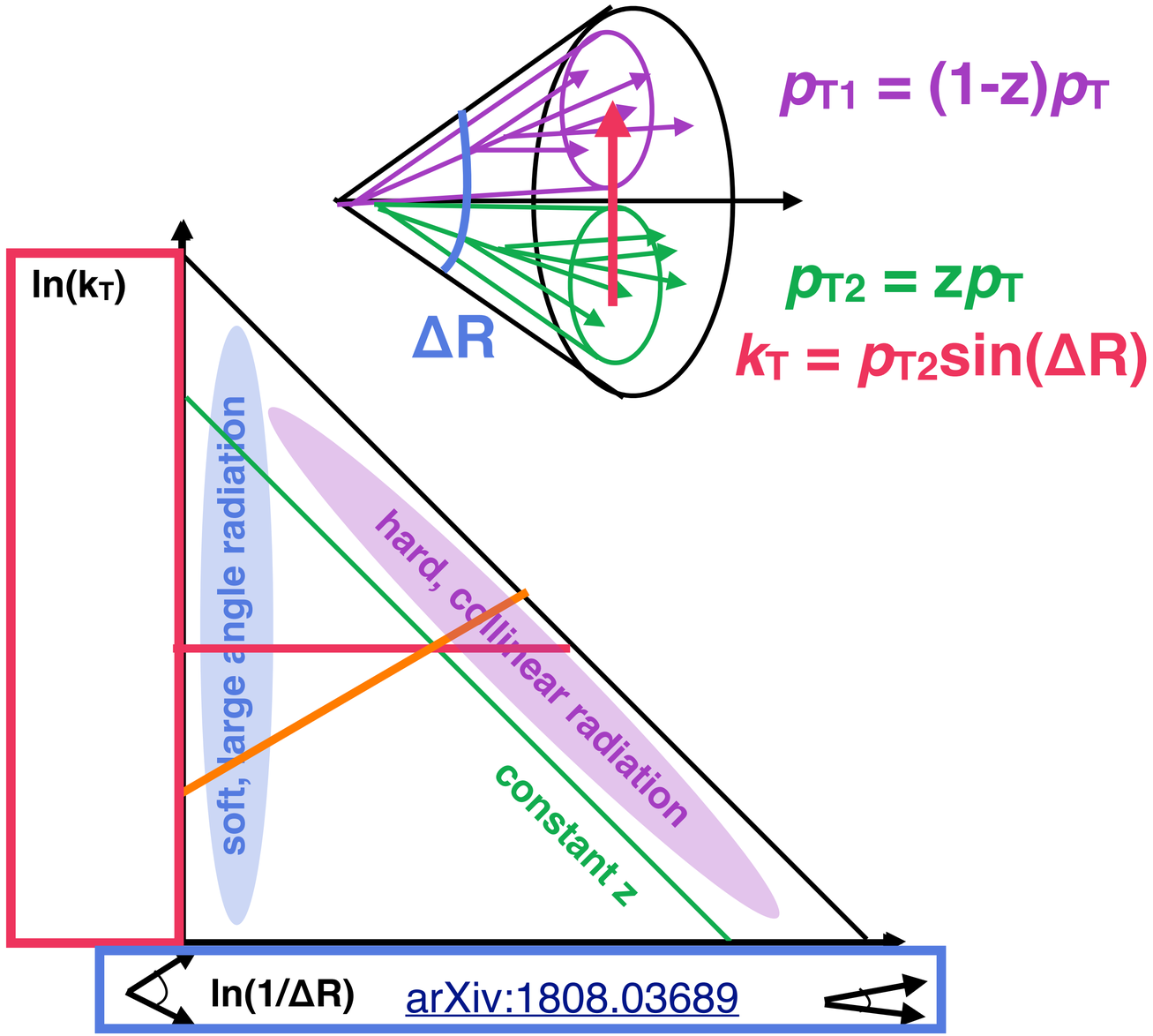
- Lund Diagram*: phase space of jet splitting

*Z. Phys. C43 (1989)
JHEP 12 (2018)

- ΔR: opening angle between subjets
- k_T: relative transverse momentum of subjets
- $\log(k_T) > 0$ separates perturbative from non-perturbative regime
- Formation time: how long until the splitting occurred

$$t_f = \frac{1}{(1-z)k_T \Delta R}$$

Y. L. Dokshitzer, et.al.



arXiv:1808.03689

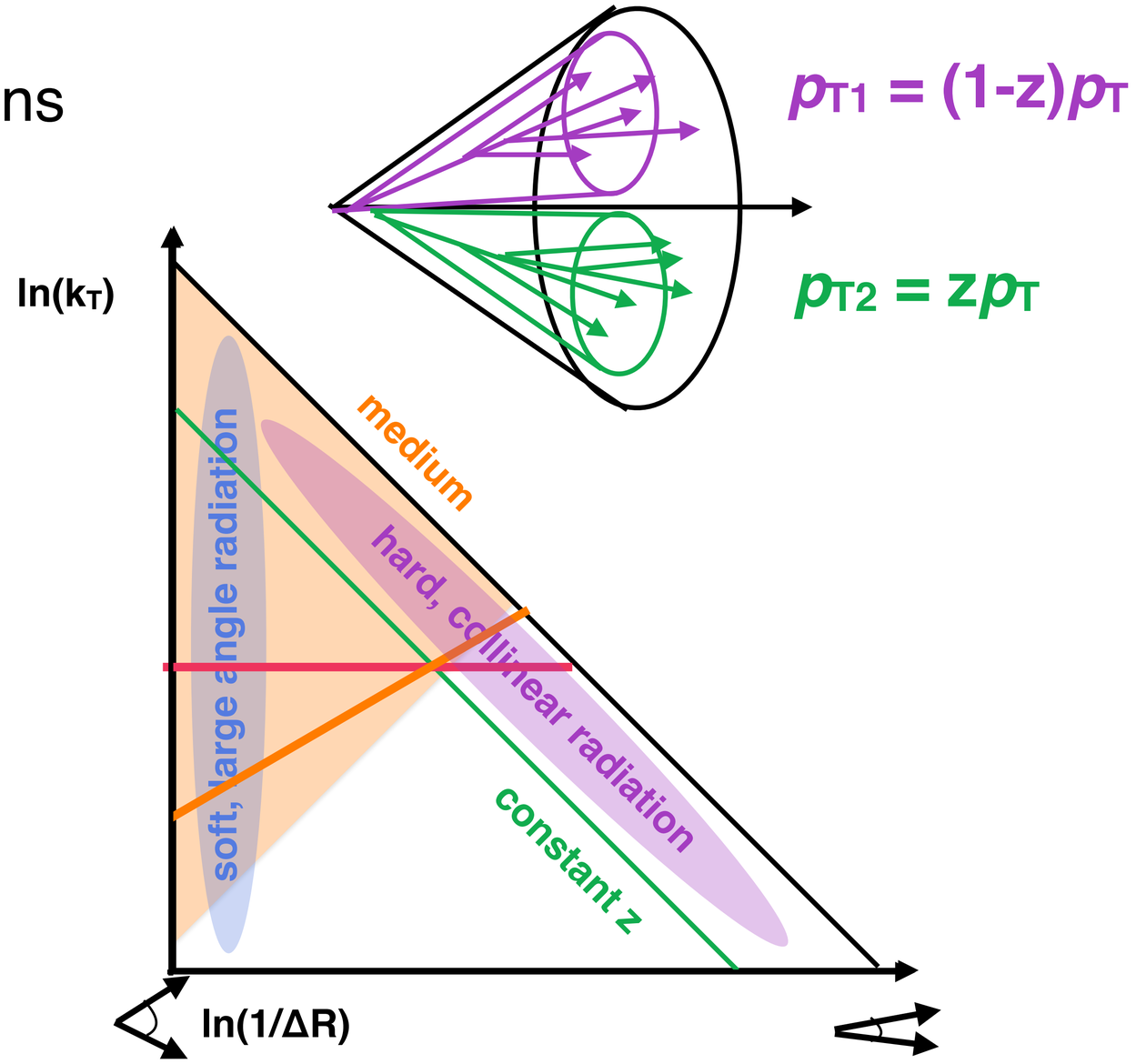
Exploring the Lund Plane: in medium

- Jet splittings in heavy-ion (HI) collisions

➡ in/out of medium splittings
  - Earlier/wider splittings experience more medium

➡ Vacuum splittings vs. non-perturbative in-medium splittings

➡ Coherence vs. decoherence
  - Split jets should be more quenched



Laura Havener, Yale University

Soft drop grooming

- Reconstruct anti-k_T R=0.4 charged jets between 80-120 GeV/c with jet-by-jet constituent background subtraction* (in HI collisions)

*JHEP 06 (2014) 092

- Recluster jets with Cambridge/Aachen (C/A)* to enforce angular ordering and fill *primary* Lund diagram with splitting information

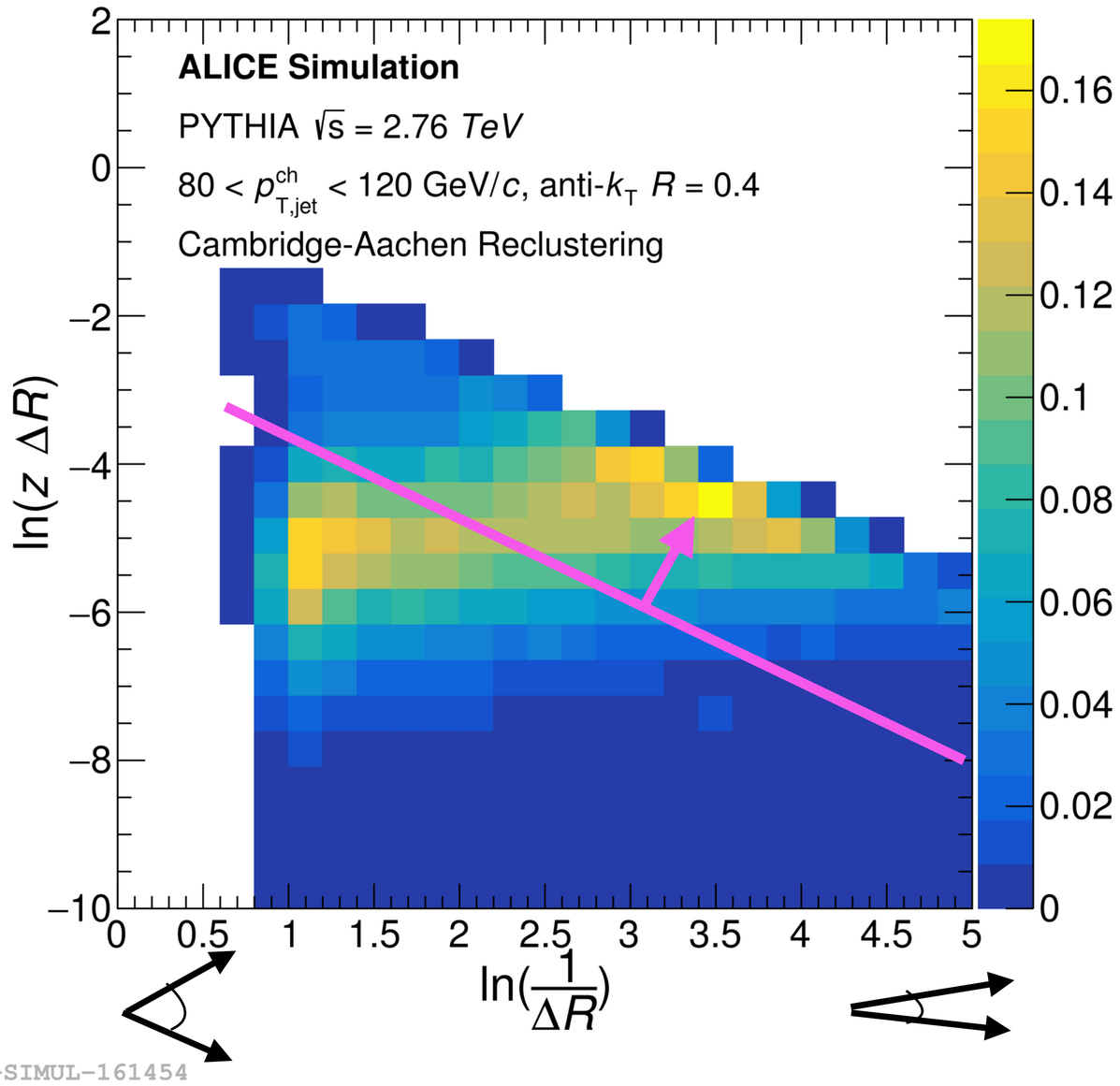
*JHEP 9708:001,1997

- Soft drop grooming to access hard splitting

$$z_g = \frac{\min(p_{Ti}, p_{Tj})}{p_{Ti} + p_{Tj}}$$

$$z_g > z_{cut}\,\theta^{\beta} \qquad \theta = \frac{\Delta R}{R}$$

- Default condition: $z_{cut} = 0.1$, $\beta = 0$



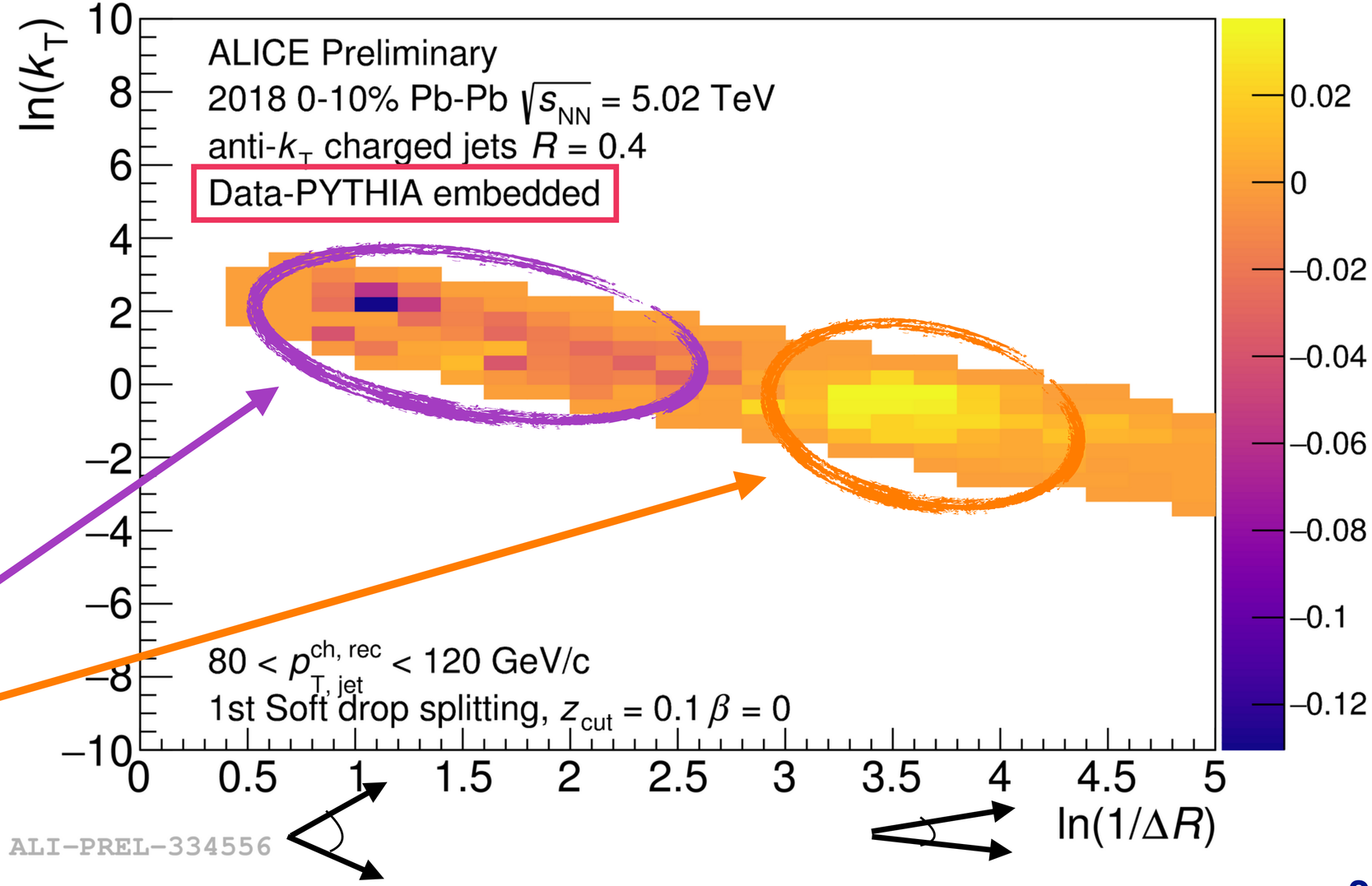
Exploring the Lund Plane in Data

New 2018 0-10% Pb-Pb collision data at 5.02 TeV

- Compare to PYTHIA8 embedded into real 0-10% Pb-Pb collisions

- Subtract the embedded MC from the data in order to remove the effects from the large HI background

- Suppression at large ΔR and enhancement at small ΔR

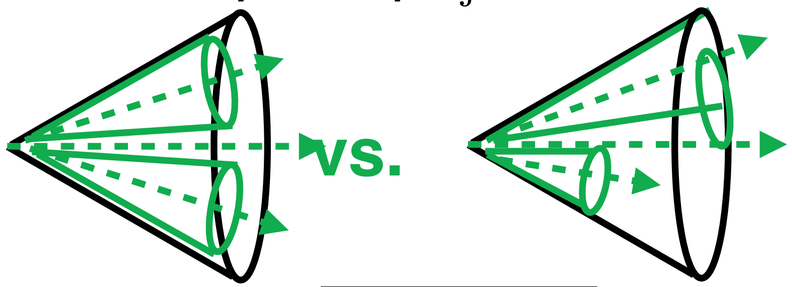


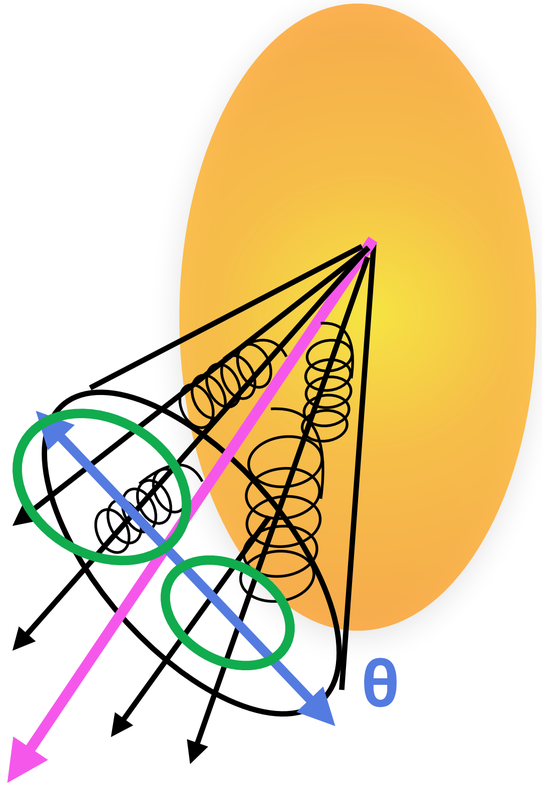
Groomed variables

- Soft drop grooming variables probe jet splitting

➡ z_g: shared momentum fraction between two hardest subjets in parton shower

$$z_g = \frac{\min(p_{Ti}, p_{Tj})}{p_{Ti} + p_{Tj}}$$

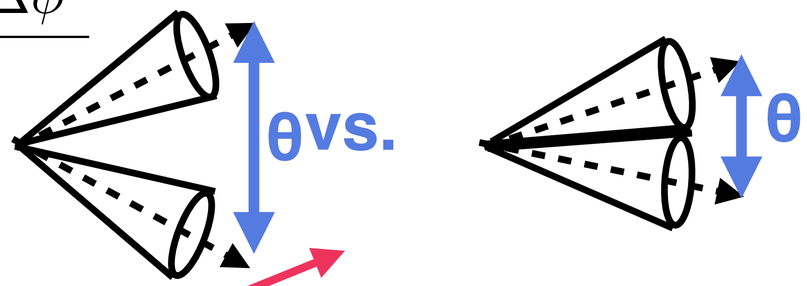




How symmetric is the jet splitting?

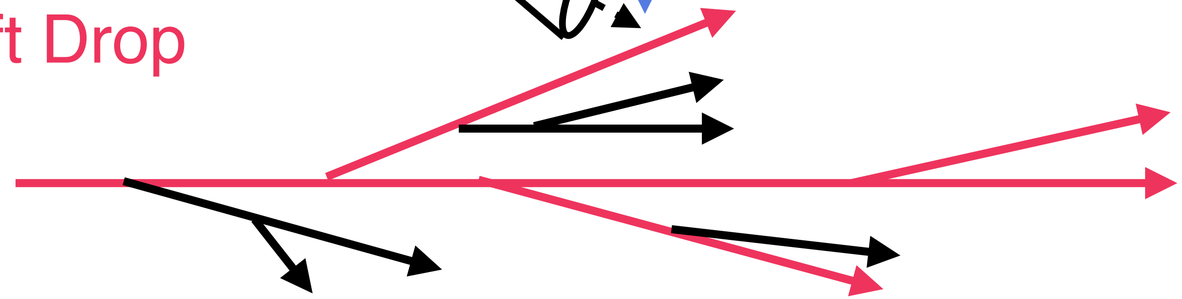
➡ θ_g: distance between subjets

$$\theta_g = \frac{R_g}{R} = \frac{\sqrt{\Delta\eta^2 + \Delta\phi^2}}{R}$$



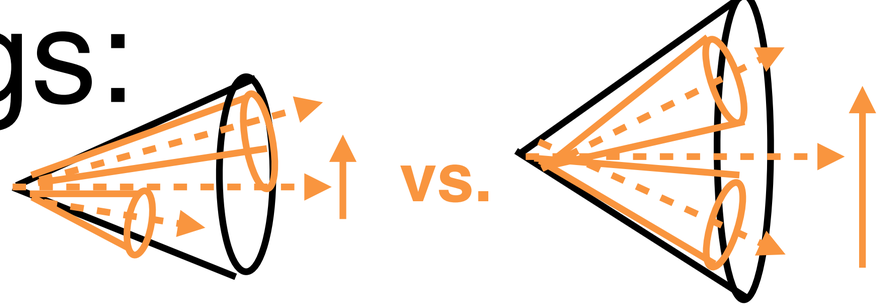
How far apart are the subjets?

➡ n_{SD}: number of splittings passing Soft Drop



Number of subjets within a jet?

Jet splittings: large R



$200 < p_T < 500$ GeV

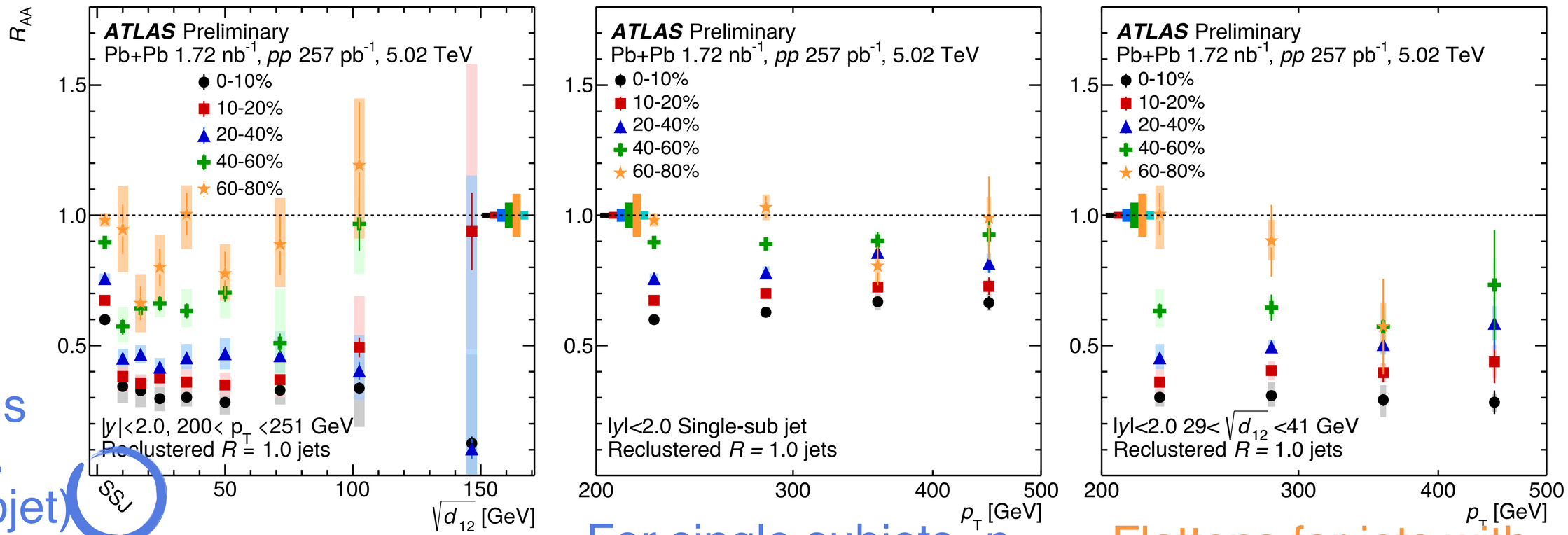
R= 1.0

$\sqrt{d_{12}} = p_{T2}\Delta R_{12}$

- Fully unfolded in 2D such that it can be directly compared to models

$\sqrt{d_{12}} = 0$

SSJ: no splittings is found (i.e. single subjet)



See more suppression when a splitting is found

For single subjets, p_T dependence similar to inclusive case

Flattens for jets with a splitting

ATLAS-CONF-2019-054

Consistent with ALICE picture of split jets being suppressed