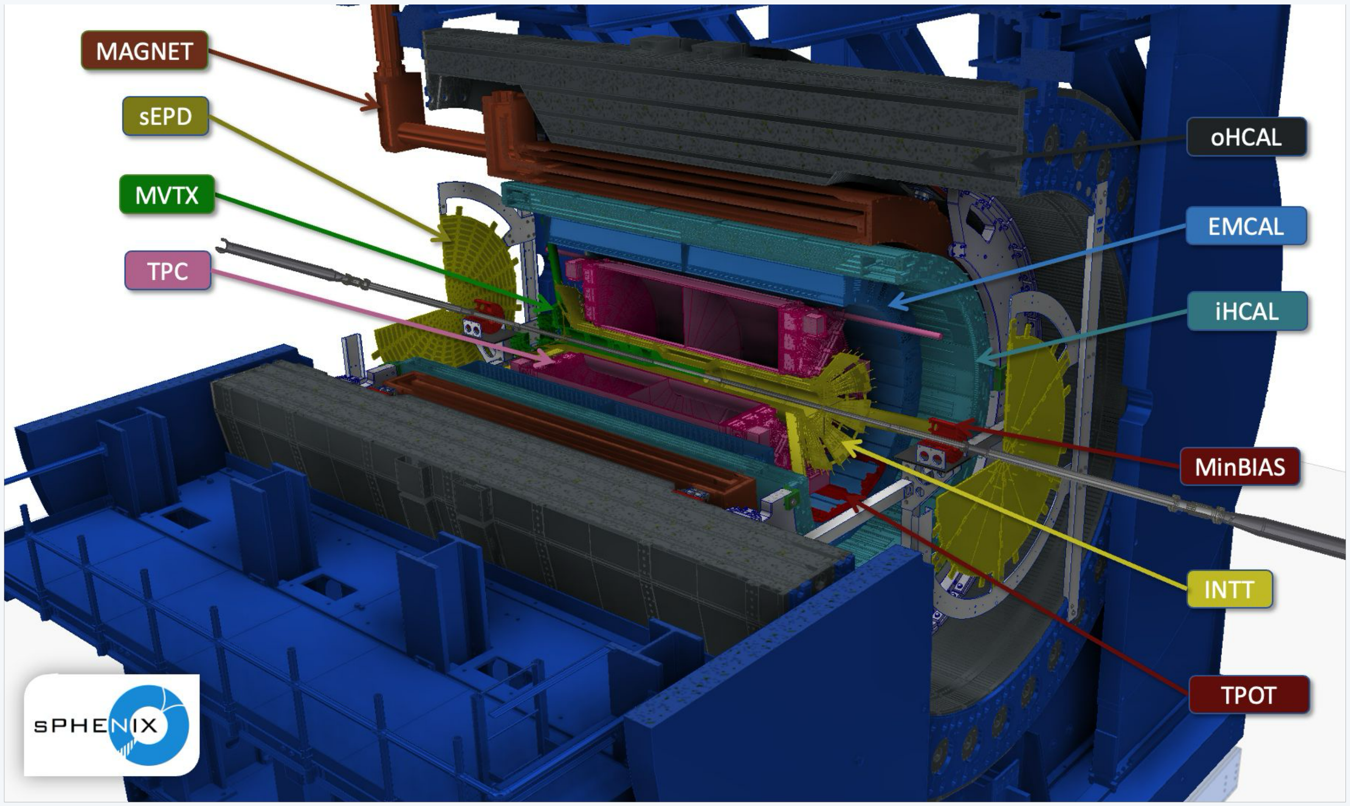
Jet Substructure in p+p Collisions with the sPHENIX Detector

Jennifer L. James - Vanderbilt University

Hard Probes Satellite Workshop on Jets and Heavy Flavor Physics
June 27–28, 2026 | Atlanta, GA

VANDERBILT UNIVERSITY
U.S. DEPARTMENT of ENERGY
sPHENIX

The sPHENIX Detector



Superconducting Magnet:
(1.4 T solenoid) formerly BaBar

Tracking system:
MAPS - based Vertex Tracker (MVTX)
Intermediate Silicon Tracker (INTT)
Time Projection Chamber (TPC)
TPC Outer Tracker (TPOT)

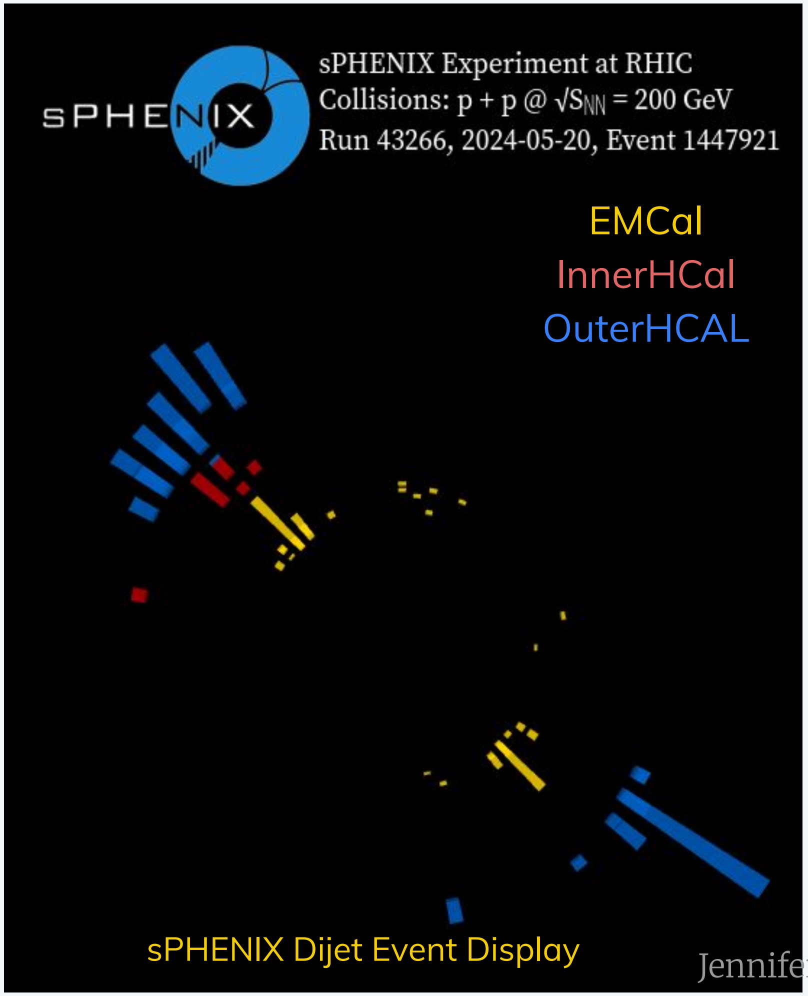
Calorimetry:
Electromagnetic Calorimeter (EMCal)
inner Hadronic Calorimeter (iHCal)
outer Hadronic Calorimeter (oHCal)

High rate:
Up to 15 kHz Trigger, >10 GB/s DAQ
Large acceptance:
Full φ coverage; for $|z|<10$ cm: $|\eta|\leq 1.1$

EMCal: $16\%/\sqrt{E} \oplus 5\%$, granularity 0.025×0.025

iHCal + oHCal: $100\%/\sqrt{E}$, granularity 0.1×0.1

Calorimeter Jet Reconstruction



Tower inputs

- EMCal retowered to match HCal 0.1 ×0.1 geometry
- Tower 4-momentum from spatial center + energy
- Vertex-corrected η, ϕ per tower

Standard Jet Finding at sPHENIX

- Anti-k_T algorithm via FastJet
- R = 0.2–0.8 jets, $|\eta_{jet}| \leq 1.1 - R$
- Jet energy scale corrections applied
- |z_vtx| ≤ 30 cm or 60 cm

Why Jet Substructure at RHIC?

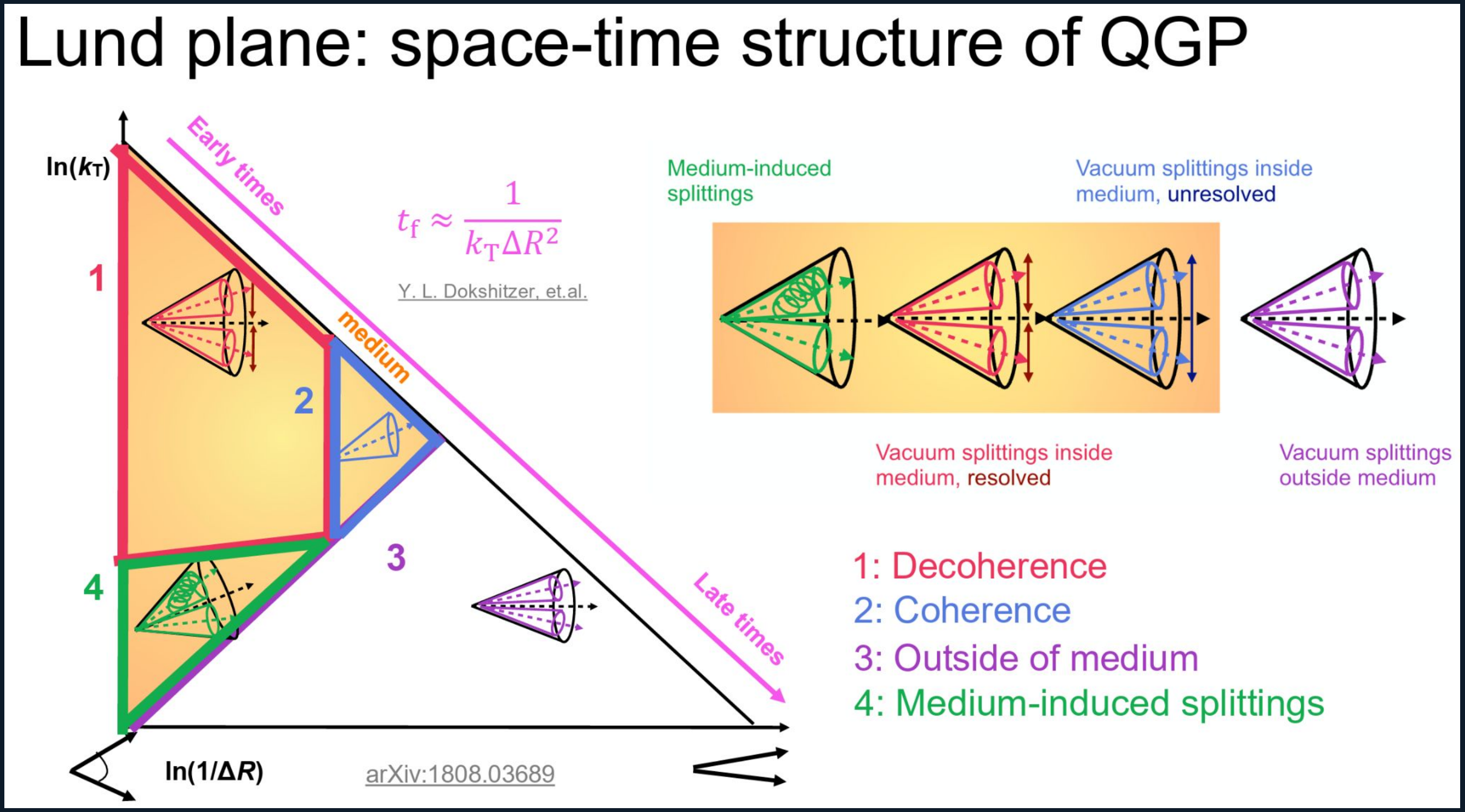
QGP & Medium Modification

- Jets traverse the QGP and lose energy through medium interactions
- Substructure encodes how this energy loss depends on angular scale and splitting kinematics
- Explore coherence vs. decoherence regime determined by QGP coherence length
  - $\lambda_T \sim (\hat{q}\, \theta_c^{\,2})^{-1/3}$

Why p+p first?

- p+p establishes the vacuum baseline for all substructure observables
- Validates MC simulations (PYTHIA & HERWIG) and unfolding procedures
- Essential reference for Au+Au medium modifications
- sPHENIX 2024 p+p dataset:
  - 16.6 pb^{-1} at $\sqrt{s}$ = 200 GeV

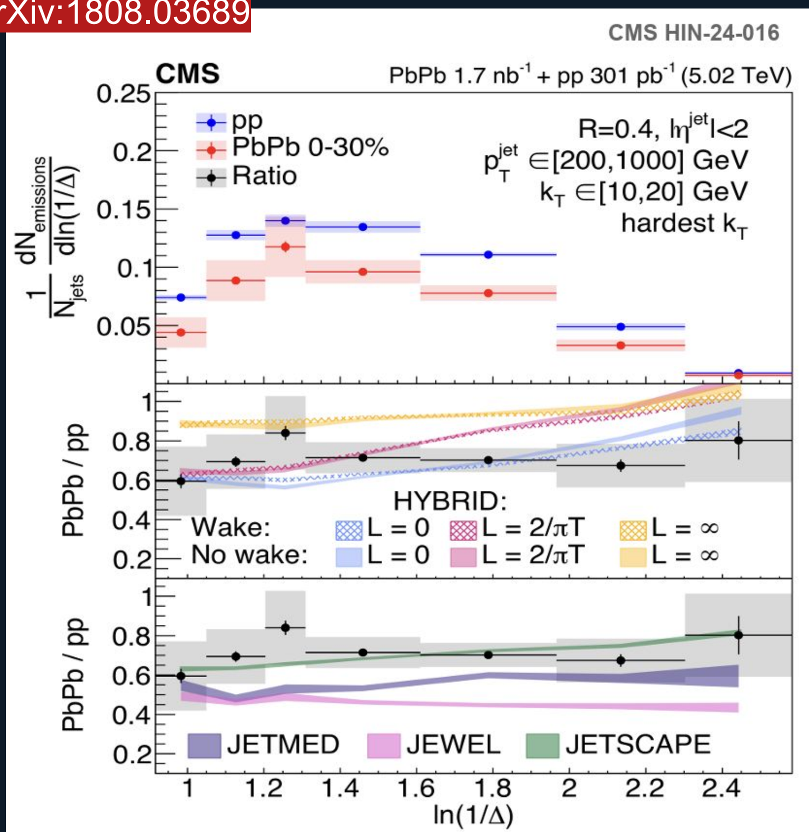
Why Jet Substructure at RHIC?

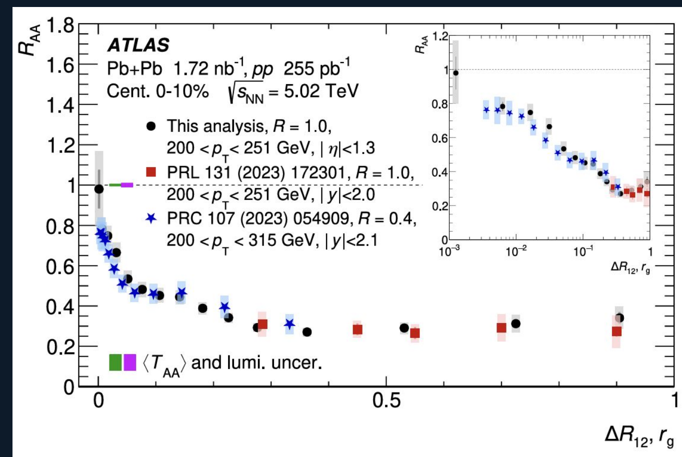


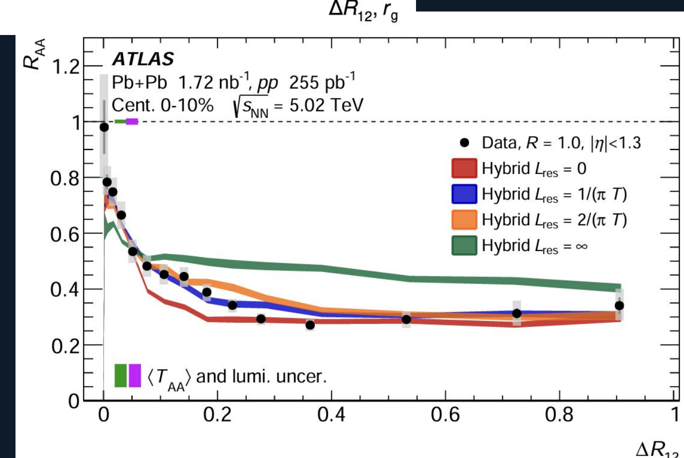
Laura Havener, Yale U. @ HP26

Why Jet Substructure at RHIC?

arXiv:1808.03689







arXiv:2504.04805v2

arXiv:2509.08881

Anne Sickles
HP 26, Tuesday

Soft Drop Grooming

Identifying the hard splitting core of a jet

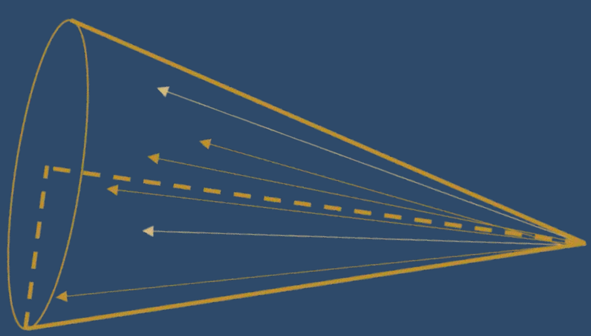
Soft Drop Condition

$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut} \left(\frac{R_g}{R_{Jet}}\right)^{\beta}$$

$$R_g = \sqrt{\Delta\phi^2{}_{1,2} + \Delta y^2{}_{1,2}}$$

- z_g : groomed momentum fraction
- R_g : angular distance between prongs

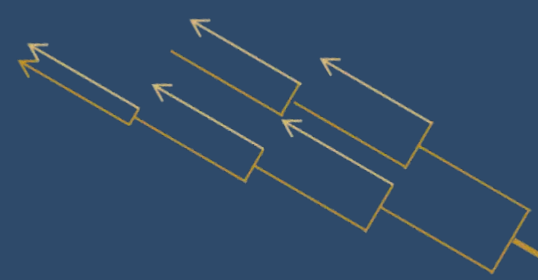
Anti-k_T Jet



Cambridge-Aachen (CA)

Do CA on the constituents

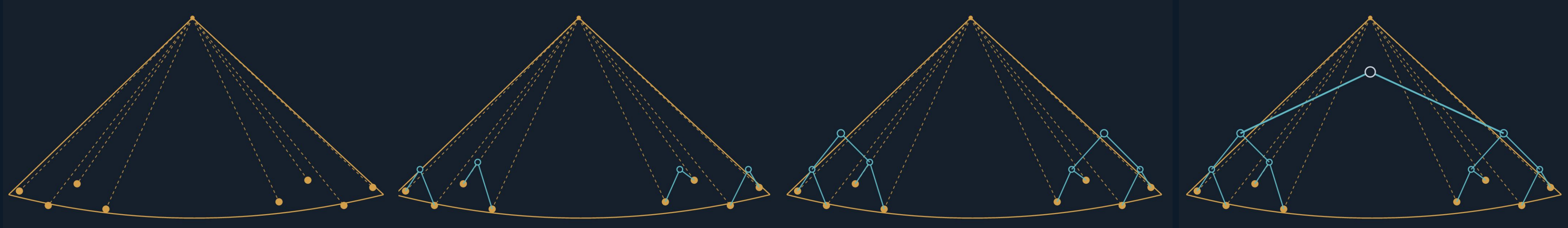
SoftDrop Prong Structure



SoftDrop: JHEP05(2014)146

Walkthrough of Jet Algorithms

Cambridge Aachen then reclusters on found anti-k_T constituents (Cacciari, JHEP04 (2008) 063)



Start from previously found anti-k_T constituents

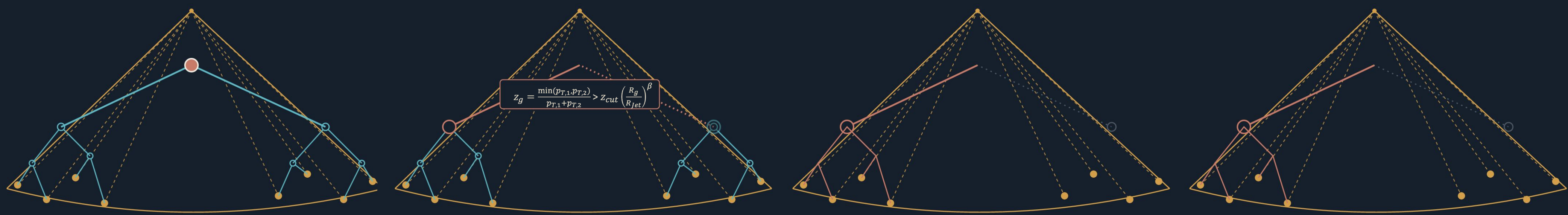
Cluster closest ΔR constituent pairs

Cluster constituent pairs into branch structure

Complete the structured tree of constituents

Walkthrough of Jet Algorithms

SoftDrop evaluates and declusters surviving constituents after CA (Cacciari, JHEP04 (2008) 063)



Start from the "top" of the structure found with CA

Test the SoftDrop conditions on the first branches

$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut} \left(\frac{R_g}{R_{Jet}}\right)^{\beta}$$

Branch failed the condition, begin recursive process

Complete the recursive process when SoftDrop passes. This forms the prong-like structures

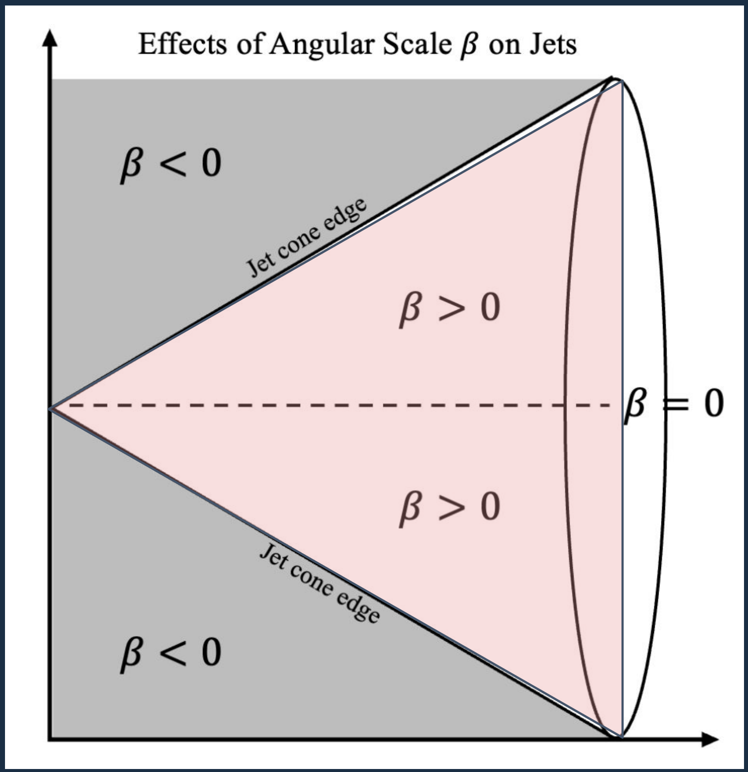
Soft Drop Grooming

$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut} \left(\frac{R_g}{R_{Jet}}\right)^{\beta}$$

Angle-dependent setting ($\beta>0$):

For $\beta>0$ the right-hand side of the equation increases with R_g

For a *wide-angle splitting*, the threshold that z_g has to clear is larger, making it harder to pass. For a *narrow-angle splitting*, the threshold is smaller, making it easier to pass



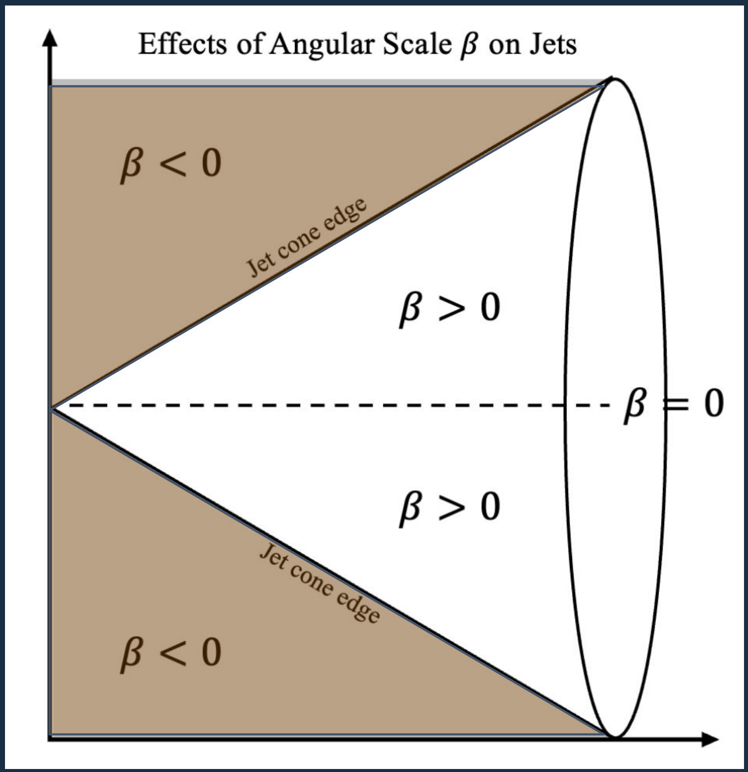
Soft Drop Grooming

$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut} \left(\frac{R_g}{R_{Jet}}\right)^{\beta}$$

Angle-dependent setting ($\beta < 0$):

When $\beta < 0$ SoftDrop can remove both soft and collinear radiation

SoftDrop acts like a "tagger"given it vetoes jets that do not have two well-separated hard prongs



Soft Drop Grooming

$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1}+p_{T,2}} > z_{cut}\left(\frac{R_g}{R_{Jet}}\right)^{\beta}$$

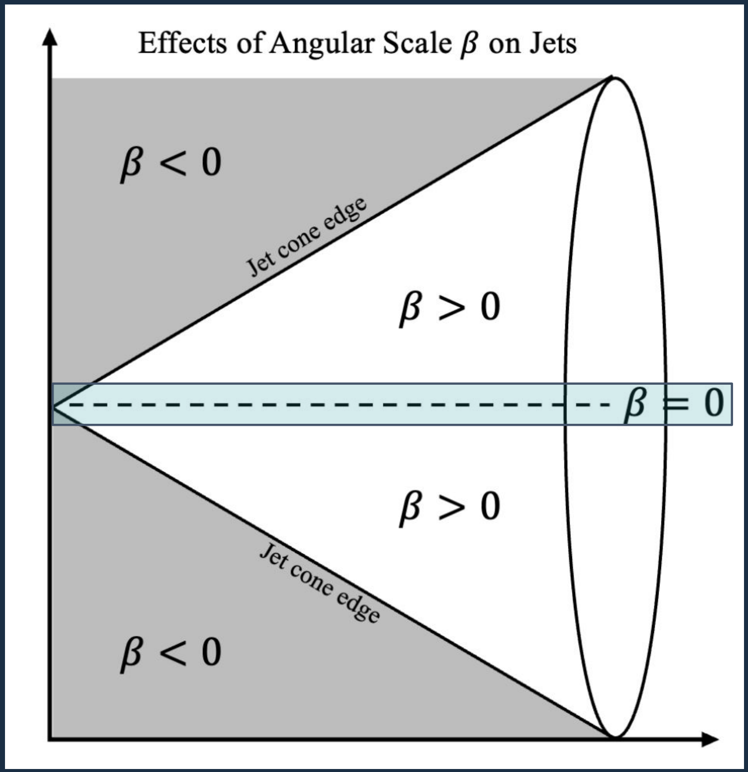
1

sPHENIX

Angle-dependent setting ($\beta = 0$):

When $\beta = 0$ SoftDrop acts as a tag for whether the first splitting satisfies the z_g threshold

So the algorithm doesn't preferentially remove wide-angle soft radiation the way it would for $\beta > 0$



Soft Drop Grooming

sPHENIX

Identifying the hard splitting core of a jet

Soft Drop Condition

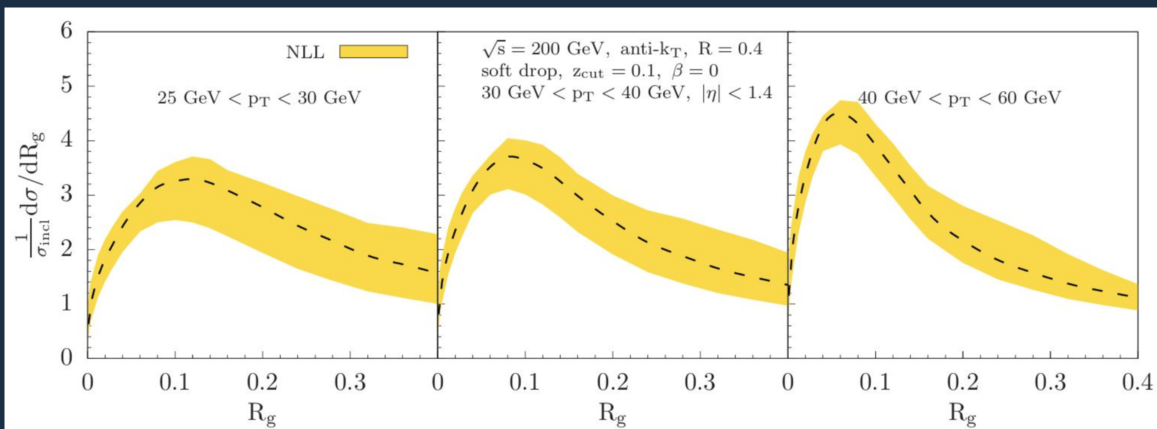
$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut}$$

$$R_g = \sqrt{\Delta\phi^2_{1,2} + \Delta y^2_{1,2}}$$

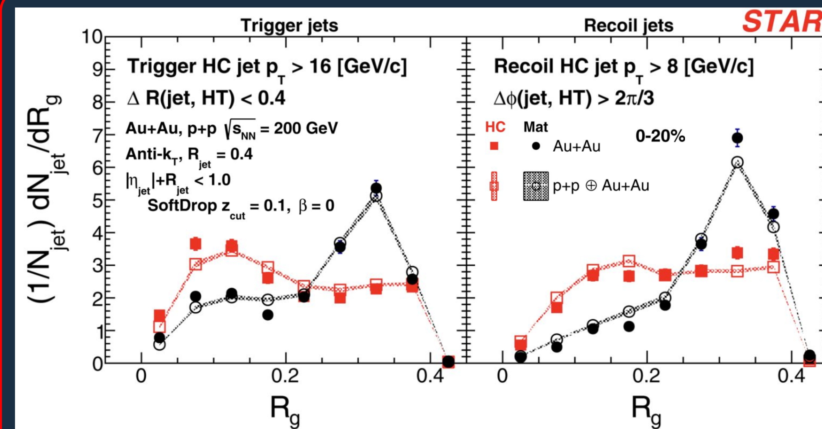
- z_g : groomed momentum fraction
- R_g : angular distance between prongs
- $z_{cut} = 0.1$, $\beta = 0$ ← (used in this analysis)

It's not just the angle or just the momentum balance alone, it's the combination. Wide-angle splittings need progressively harder momentum balance to survive, and β controls how steeply that threshold rises with angle.

Soft Drop Grooming



Kang et al., 10.1007/JHEP02(2020), 054

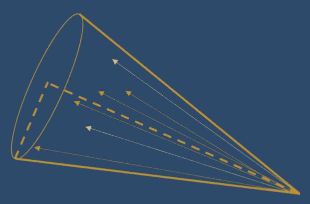


PHYSICAL REVIEW C 105, 044906 (2022)

The soft drop groomed jet radius R_g at NLL in p+p collisions for STAR kinematics at $\sqrt{s}$ = 200 GeV compared to p+p data (red) from STAR.
Peak of R_g shifts as a function of the jet energy towards smaller angles!

Soft Drop Grooming w. Reclustering

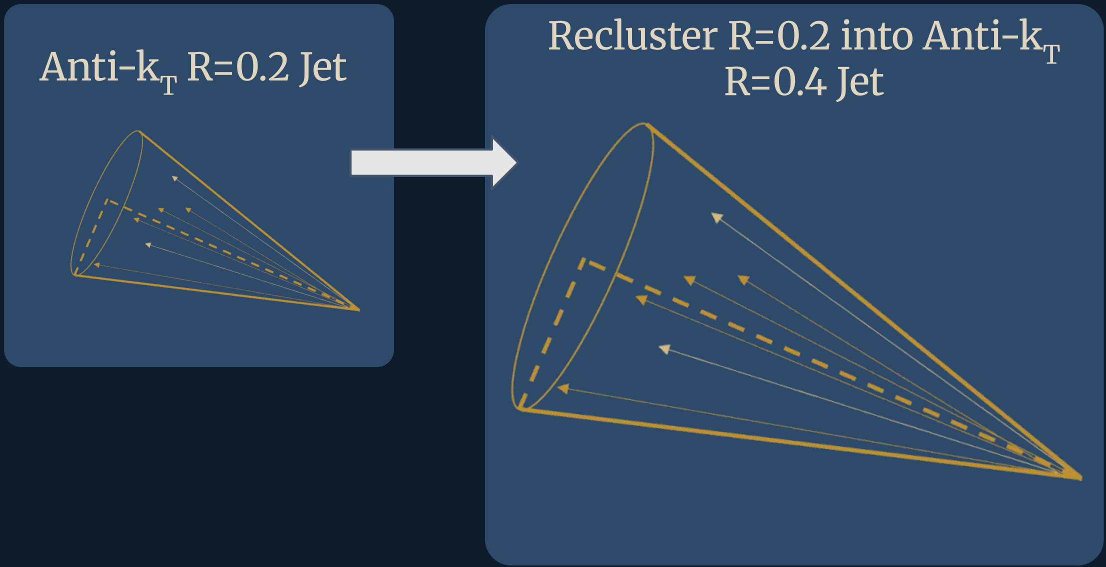
Anti-k_T R=0.2 Jet



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

$$R_g = \sqrt{\Delta\phi^2{}_{1,2} + \Delta y^2{}_{1,2}}$$

Soft Drop Grooming w. Reclustering

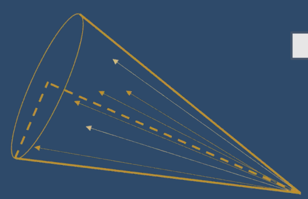


$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut}$$

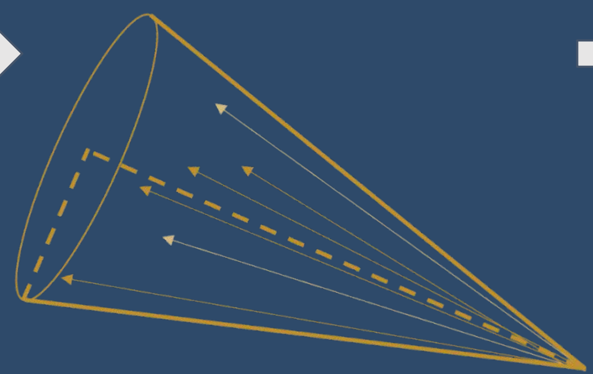
$$R_g = \sqrt{\Delta\phi^2{}_{1,2} + \Delta y^2{}_{1,2}}$$

Soft Drop Grooming w. Reclustering

Anti-k_T R=0.2 Jet



Recluster R=0.2 into Anti-k_T R=0.4 Jet



Cambridge-Aachen (CA)

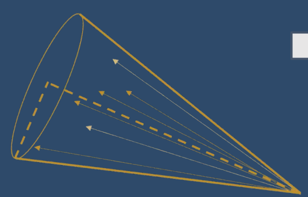
Do CA on the constituents

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

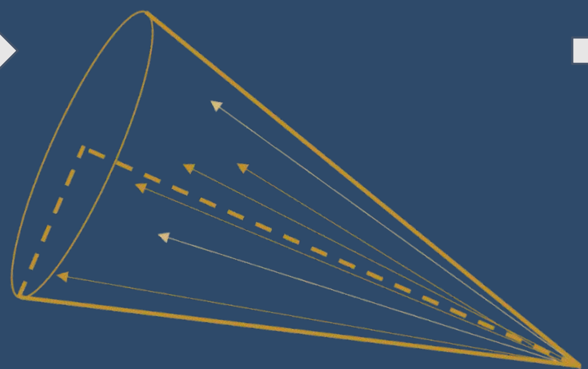
$$R_g = \sqrt{\Delta\phi^2{}_{1,2} + \Delta y^2{}_{1,2}}$$

Soft Drop Grooming w. Reclustering

Anti-k_T R=0.2 Jet



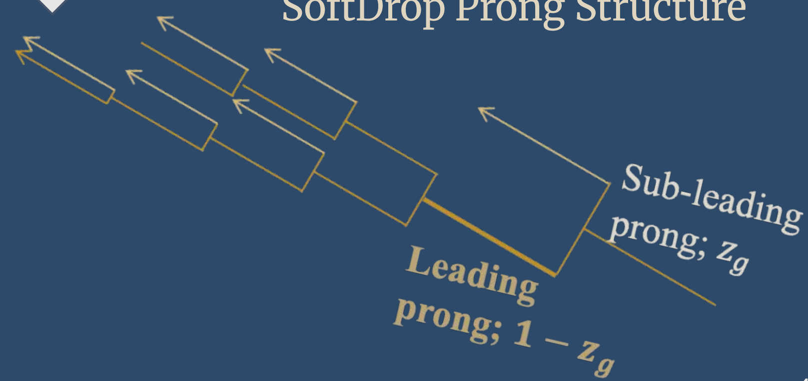
Recluster R=0.2 into Anti-k_T R=0.4 Jet



Cambridge-Aachen (CA)

Do CA on the constituents

SoftDrop Prong Structure



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

$$R_g = \sqrt{\Delta\phi^2{}_{1,2} + \Delta y^2{}_{1,2}}$$

R=0.2 Sub-Jet Reclustering

Applegate, N. - HP2026 Poster

sPHENIX

Step 1

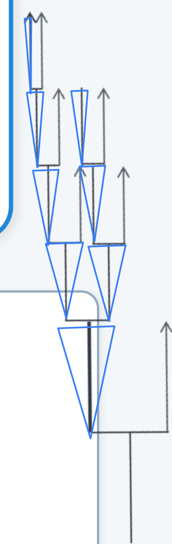
Calo towers → anti-k_T R=0.2 'constituent jets' ($p_T \geq 0.6$ GeV)

Step 2

Constituent jets form anti-k_T R=0.4 'reclustered jets' (matched to reco, $dR \leq 0.1$)

Step 3

Apply Soft Drop to reclustered jet using R=0.2 jets as inputs



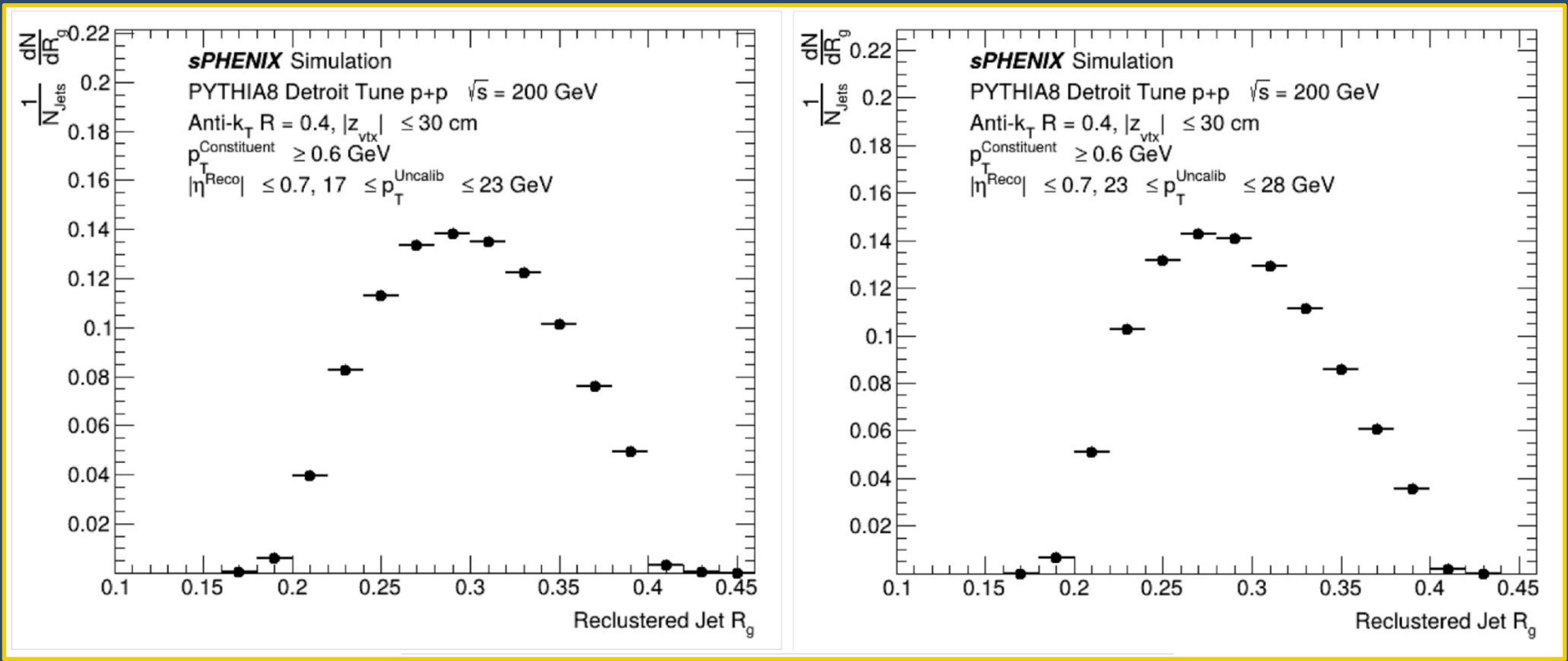
Why reconstructed jets from smaller jets?

- Captures split energy from a particle that deposits across multiple calorimeter towers
- Provides an effective infrared cutoff at the constituent jet resolution (R=0.2)

Reclustered R_g Distributions

Applegate, N. - HP2026 Poster

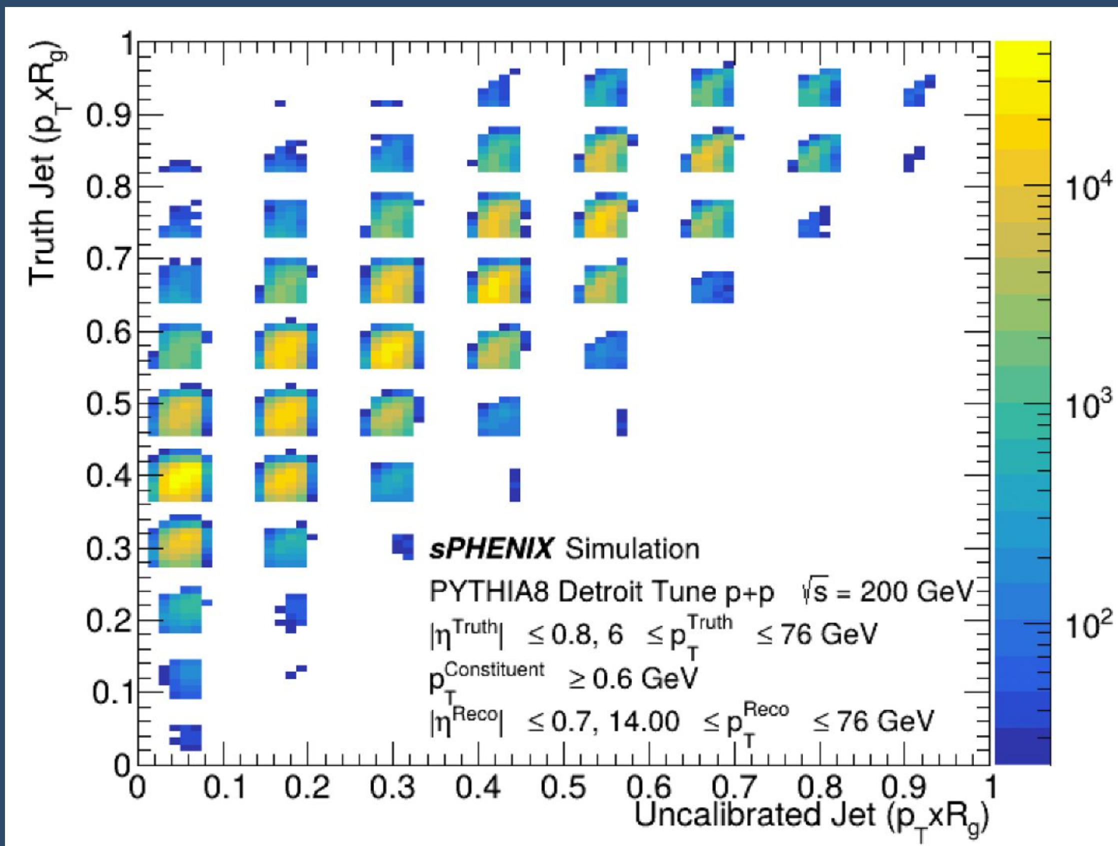
sPHENIX



- Large drop-off in R_g spectrum at $R_g \leq R_{Const.} = 0.2$: splittings within a single constituent jet are merged
- Measurement is sensitive to $R_g > 0.2$ where splittings are resolved between distinct constituent jets
- Motivates unfolding to correct for detector smearing and bin migration

R_g Response Matrix

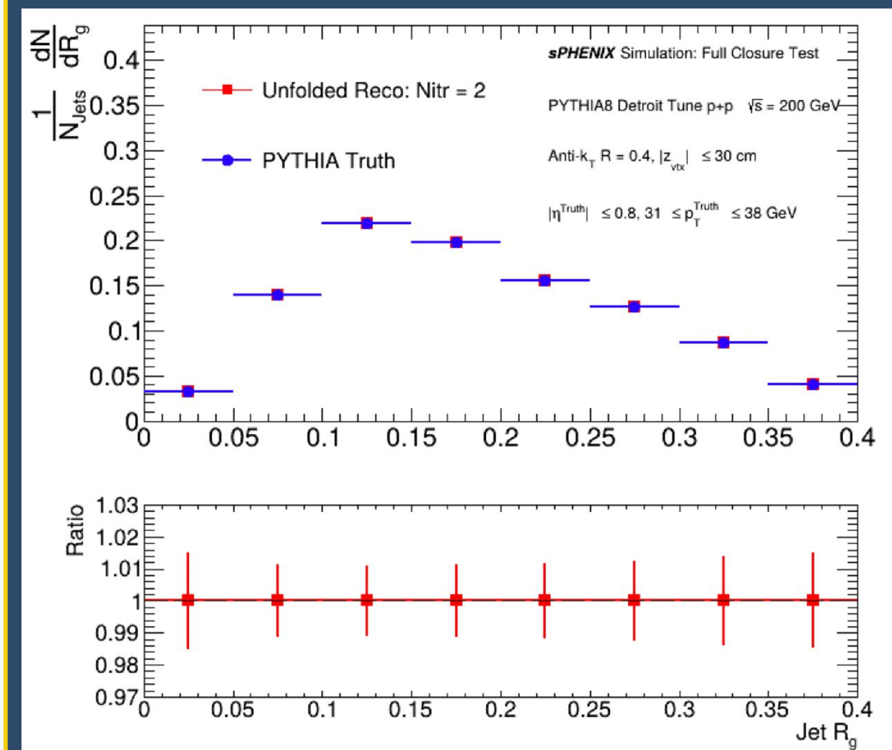
Applegate, N. - HP2026 Poster

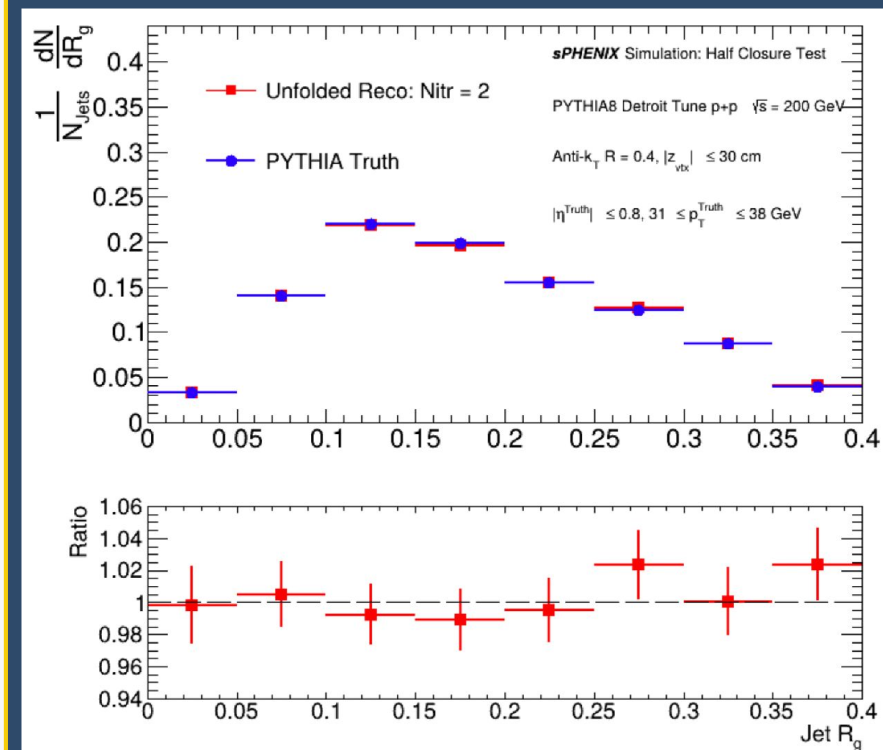


- 2-D response matrix:
  - Jet $p_T \times R_g$ bins of Pythia8 truth jets versus uncalibrated reconstructed jets
  - Using RooUnfold iterative Bayesian
    - $N_{itr} = 2$

R_g Closure Tests

Applegate, N. - HP2026 Poster

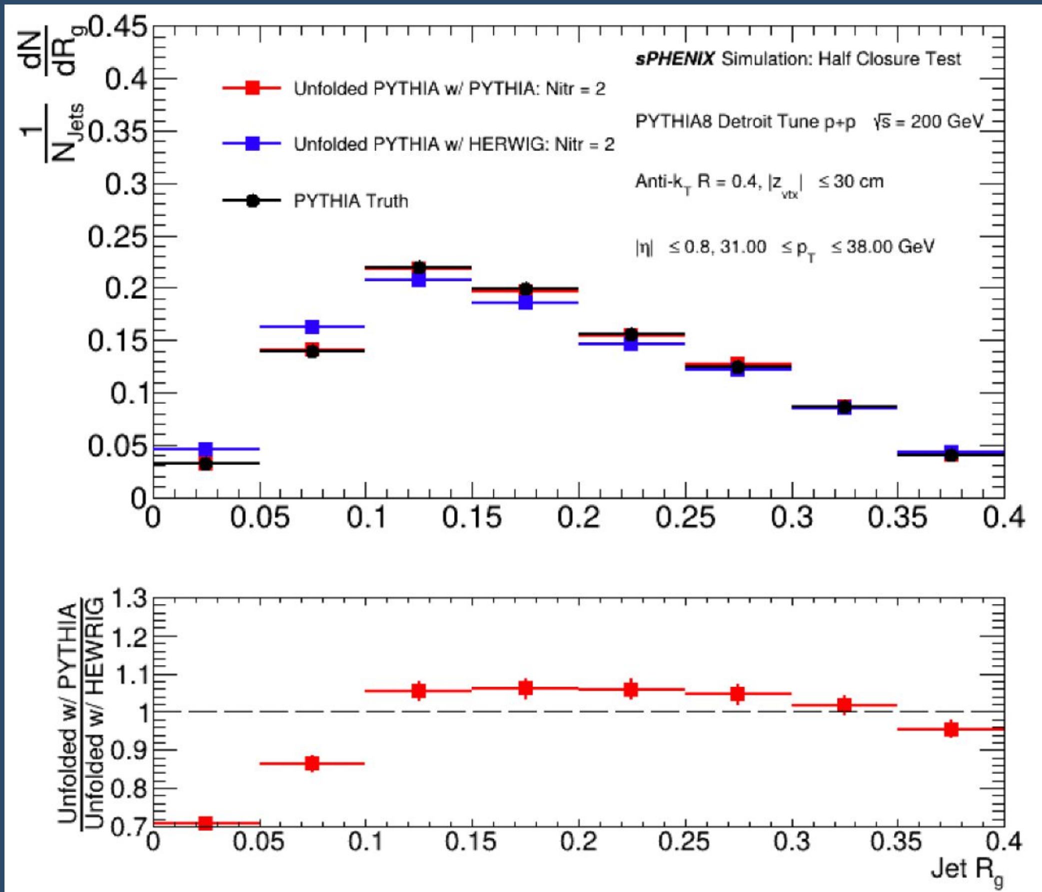




Full closure confirms unfolding mechanics. Half closure tests performance on reconstructed jets (ratio within ±3%)

Generator Dependence: R_g Unfolded

Applegate, N. - HP2026 Poster

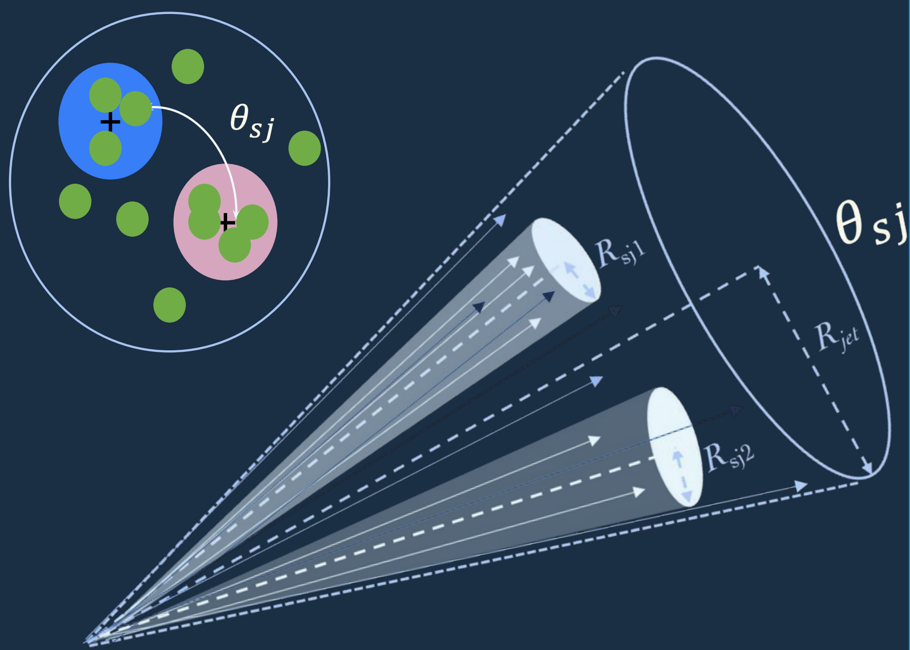


- **Generator bias below R_g = 0.2:** truth prior dominates in that region
  - Expected given constituent jet resolution cutoff
- **Strong PYTHIA/HERWIG agreement** above the constituent jet resolution
- Sets up the same comparison for another substructure observable called θ_{sj}

Subjet Opening Angle θ_{sj}

A complementary probe of jet fragmentation

$$\theta_{sj} = \sqrt{\Delta\phi^2{}_{sj} + \Delta y^2{}_{sj}} \qquad z_{sj} = \frac{\min(p_{T,sj1}, p_{T,sj2})}{p_{T,sj1} + p_{T,sj2}}$$



Example Scenario:
Using anti-k_T R = 0.4 jet constituents,
Cluster into anti-k_T R_{sj}=0.1 subjets

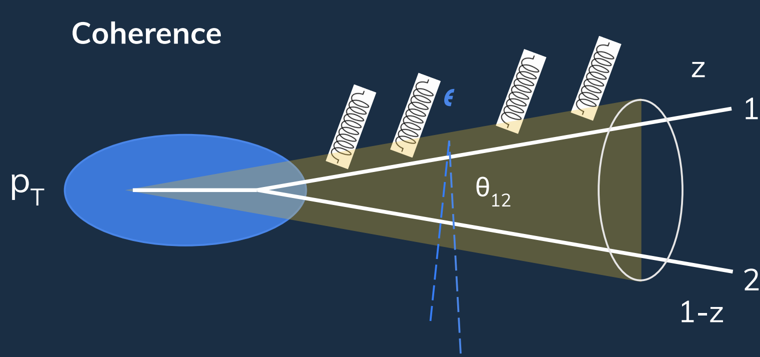
Calculate θ_{sj} directly from these anti-k_T subjets

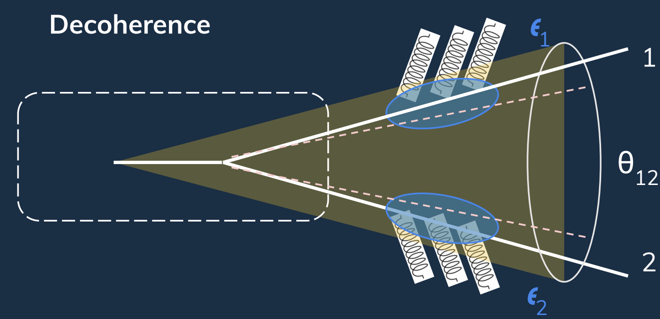
vs. SoftDrop

- R_g finds the hardest splitting via iterative grooming
  - SoftDrop R_g sensitive to background fluctuations
- θ_{sj} captures angular structure at a fixed angular scale
- Together they probe complementary aspects of in-medium modification

Theory Motivation for θ_{sj} & R_g

Mehtar-Tani, Pablos & Tywoniuk $\lambda_{T\equiv} t_d \simeq (\hat{q}\theta_{12}^2)^{-1/3}$





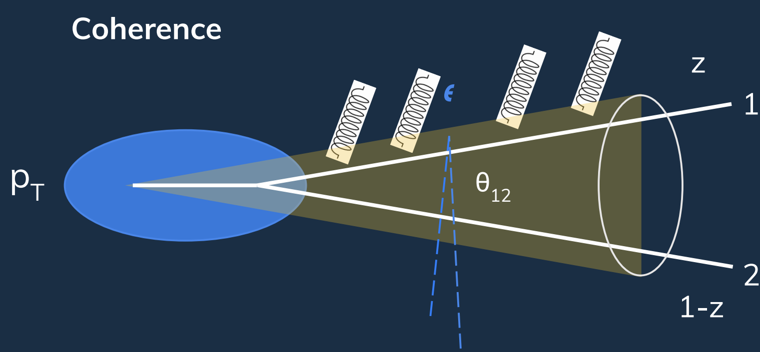
partons close together → medium sees them as one color charge → coherent energy loss, less suppression

partons well separated → medium resolves them individually → incoherent energy loss, more suppression

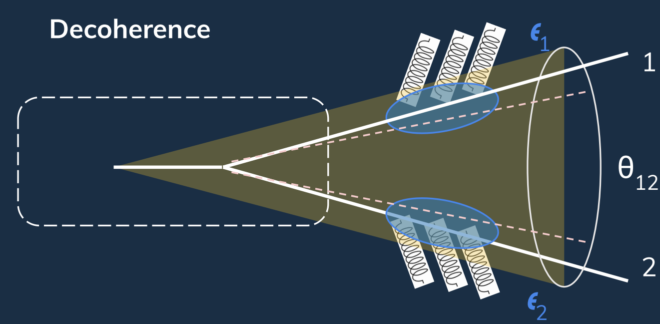
Mehtar-Tani, Pablos, Tywoniuk — PRL 127, 252301 (2021)

Theory Motivation for θ_{sj} & R_g

Mehtar-Tani, Pablos & Tywoniuk $\lambda_{T\equiv}t_d \simeq (\hat{q}\theta_{12}^2)^{-1/3}$



θ_C

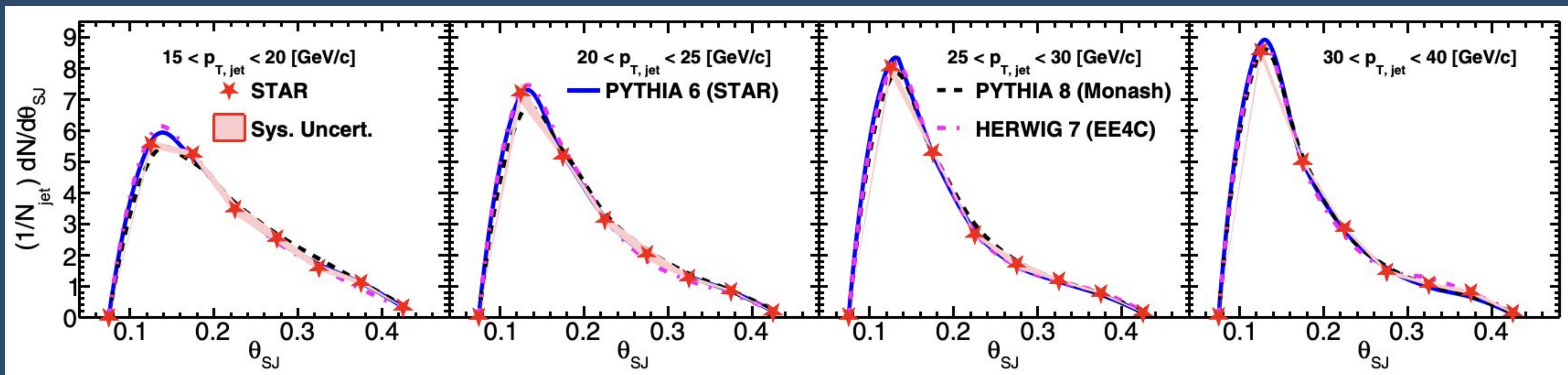


partons close together → medium sees them as one color charge → coherent energy loss, less suppression

partons well separated → medium resolves them individually → incoherent energy loss, more suppression

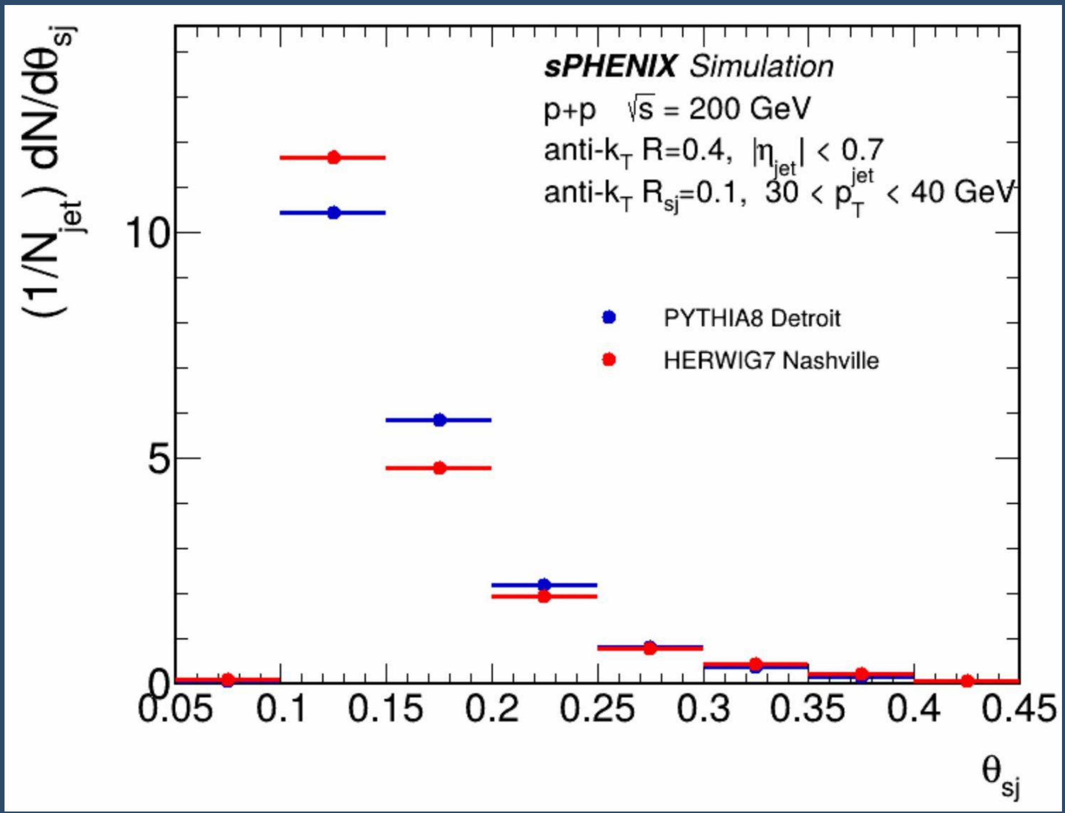
Mehtar-Tani, Pablos, Tywoniuk — PRL 127, 252301 (2021)

Prior Measurement: STAR p+p θ_{sj}



- STAR used EMCal + tracking: sPHENIX currently uses calorimetry only (tracking commissioning)
- Unfolded STAR data consistent with multiple MC generators which validates our simulation approach
- sPHENIX extends this with larger dataset, different systematic profile, and eventual Au+Au comparison

θ_{sj} Truth: PYTHIA vs. HERWIG



- Truth-level generator comparison establishes the theoretical spread before detector effects
- Agreement with STAR unfolded data validates simulation for both generators
- PYTHIA/HERWIG spread quantifies generator systematic for future data unfolding
- Evaluate both substructure observables in Au+Au

Path Forward: Tracking Integration

For R_g

Charged-particle tracks improve angular resolution of the two soft-drop prongs, particularly at small R_g where calorimeter smearing is worst

For θ_{sj}

θ_{sj} is fundamentally an angular observable. A track-based angular resolution extends the fiducial range toward smaller θ_{sj} where the coherence/decoherence physics lives

General

Combined calo+tracking: calorimeter captures neutral energy, tracking provides precision momentum and angular information. sPHENIX TPC commissioning underway serving as a natural next step for both measurements

Au+Au

p+p baseline established → move to Au+Au to observe medium modifications. Tracking also aids background subtraction in the high-multiplicity heavy-ion environment

Summary

1 R_g with R=0.2 sub-jet reclustering

- Two-step calorimeter-based procedure validated in PYTHIA8 simulation
- Closure tests confirm unfolding performance
- Generator dependence (PYTHIA vs. HERWIG) characterized
- An agreement above the constituent jet resolution scale

2 θ_{sj} truth-level generator comparison

- PYTHIA8 Detroit and HERWIG7 Nashville θ_{sj} distributions established and compared to STAR unfolded p+p data
- Validates simulation for both generators; quantifies generator systematic for future data unfolding

3 Outlook

- Data unfolding for both observables in progress
- Tracking integration will extend fiducial reach
- Au+Au analysis will test medium modification against this p+p baseline

Questions

Supported by U.S. DOE Office of Science, Office of Nuclear Physics (Award DE-SC0024660)

Thank you to my:
sPHENIX Collaborators,
Vanderbilt University HENP group,
PI: Rithya Kunnawalkam Elayavalli

For all their support!

BACKUP

Analysis Cuts & Requirements

Level	Cut	Value
Event	z_{vtx}	|z_vtx| ≤ 30 cm
Jet (truth)	p_T range	6 ≤ pT < 76 GeV
Jet (truth)	η acceptance	|η| ≤ 0.8
Jet (reco)	uncalibrated p_T	14 ≤ pT,uncalib < 76 GeV
Jet (reco)	η acceptance	|η| ≤ 1.1 - R (z-vtx corrected)
Jet matching	dR	≤ 0.3 (truth–reco)
Subjet (θ_{sj})	constituents	≥ 2 per subjet
Subjet (R_g)	constituent p_T	pT,constit ≥ 0.5 GeV
Subjet (R_g)	reco–reclus match	dR ≤ 0.1

Analytic Predictions for R_g

Caucal, Soto-Ontoso, Takacs: arXiv:2111.14768v1 [hep-ph] 29 Nov 2021

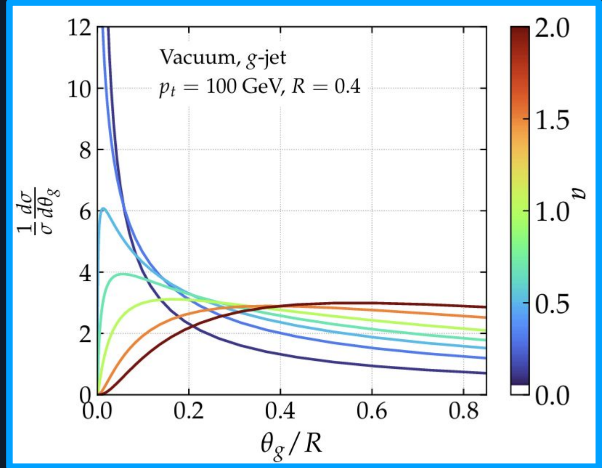
Vacuum (p+p)

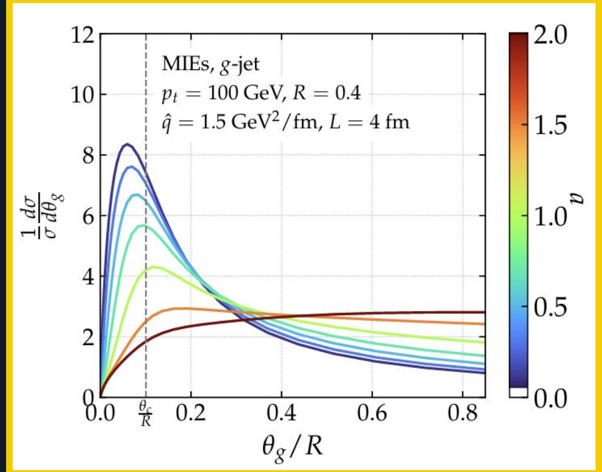
- Analytic R_g distribution calculable in perturbative QCD
- At fixed coupling:

  $dN/dR_g \sim \alpha_s \ln(R/R_g)$
- Rises toward small R_g (collinear enhancement)
- Provides the baseline for our p+p measurement must match

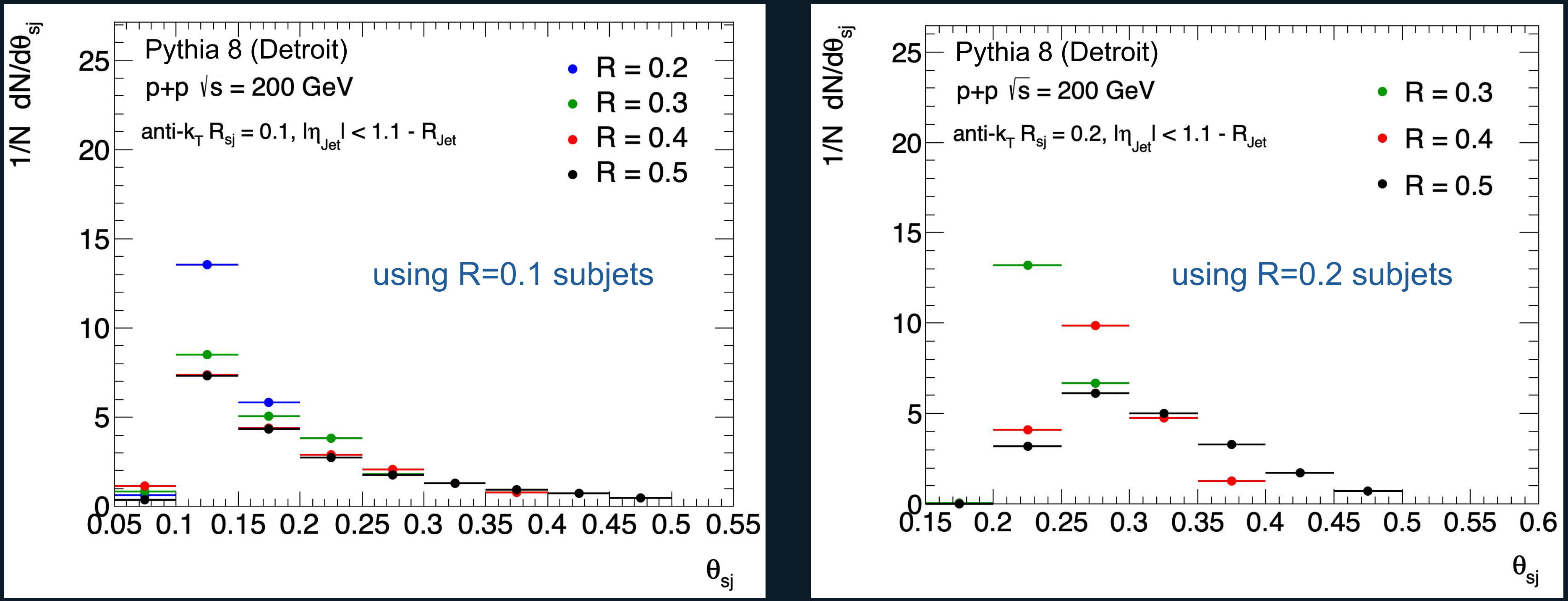
Medium Effects

- QGP suppresses large-angle splittings ($R_g \rightarrow R$) via decoherence
- Enhances small-angle splittings below the medium resolution scale
- Net effect: R_g distribution narrows toward smaller values in A+A





Changing Radius and Subjet Radius for θ_{sj}



With R_{sj} = 0.1, small mother jet radii are geometrically limited to resolving only the jet core, washing out sensitivity to the hard splitting angle. At R_{sj} = 0.2, the subjet radius is large enough that the choice of mother jet radius meaningfully impacts the θ_{sj} distribution, reflecting differences in jet splitting structure. This motivates improving both our calorimeter resolution and study with tracking.

Walkthrough of Jet Algorithms

Building a jet with the anti-k_T algorithm (Cacciari, JHEP04 (2008) 063)

Find hardest particle

Find the nearest constituent via d_{ij}

Continue the process iteratively

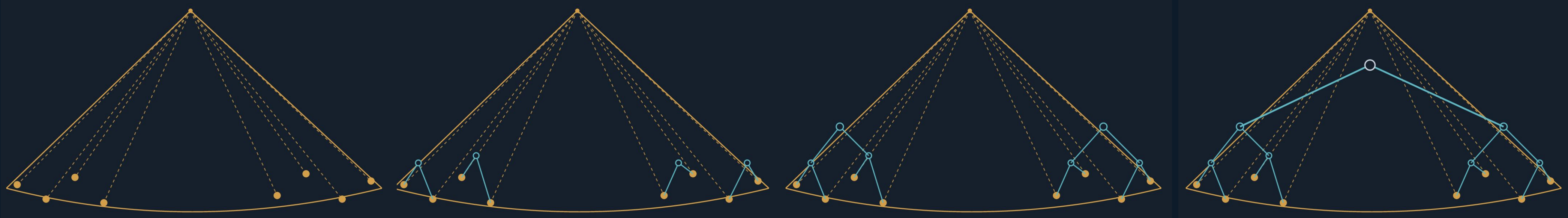
Constituents are then limited to desire jet cone radius

$$d_{ij} = \min(k^{2p}_{ti}, k^{2p}_{tj})*(\Delta^2_{ij}/R^2),$$

anti-k_T $\rightarrow$ $p = -1$

Walkthrough of Jet Algorithms

Cambridge Aachen then used on surviving anti-k_T constituents (Cacciari, JHEP04 (2008) 063)



Start from previously found anti-k_T constituents

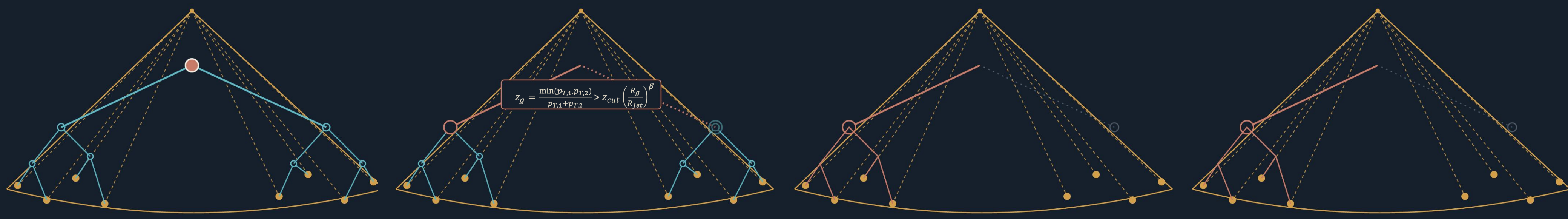
Cluster closest ΔR constituent pairs

Cluster constituent pairs into branch structure

Complete the structured tree of constituents

Walkthrough of Jet Algorithms

SoftDrop evaluates surviving constituents after CA (Cacciari, JHEP04 (2008) 063)



Start from the "top" of the structure found with CA

Test the SoftDrop conditions on the first branches

$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut} \left(\frac{R_g}{R_{Jet}}\right)^{\beta}$$

Branch failed the condition, begin recursive process

Complete the recursive process when SoftDrop passes. This forms the prong-like structures

The k_t algorithm and its siblings

$$d_{ij} = \min(p_{ti}^{2p}, p_{tj}^{2p}) \frac{\Delta y^2 + \Delta\phi^2}{R^2} \qquad d_{iB} = p_{ti}^{2p}$$

p = 1 k_t algorithm

S. Catani, Y. Dokshitzer, M. Seymour and B. Webber, Nucl. Phys. B406 (1993) 187
S.D. Ellis and D.E. Soper, Phys. Rev. D48 (1993) 3160

p = 0 Cambridge/Aachen algorithm

Y. Dokshitzer, G. Leder, S.Moretti and B. Webber, JHEP 08 (1997) 001
M.Wobisch and T.Wengler, hep-ph/9907280

p = -1 **anti-k_t algorithm**

MC, G. Salam and G. Soyez, arXiv:0802.1189

NB: in anti-kt pairs with a **hard** particle will cluster first: if no other hard particles are close by, the algorithm will give **perfect cones**

Quite ironically, a sequential recombination algorithm is the 'perfect' cone algorithm

IRC safe algorithms

k_t	SR $d_{ij} = \min(p_{ti}^2, p_{tj}^2)\Delta R_{ij}^2/R^2$ hierarchical in rel p_t	Catani et al '91 Ellis, Soper '93	NlnN
Cambridge/ Aachen	SR $d_{ij} = \Delta R_{ij}^2/R^2$ hierarchical in angle	Dokshitzer et al '97 Wengler, Wobish '98	NlnN
anti-k_t	SR $d_{ij} = \min(p_{ti}^{-2}, p_{tj}^{-2})\Delta R_{ij}^2/R^2$ gives perfectly conical hard jets	MC, Salam, Soyez '08 (Delsart, Loch)	$N^{3/2}$
SISCone	Seedless iterative cone with split-merge gives 'economical' jets	Salam, Soyez '07	$N^2 \ln N$

'second-generation' algorithms

All are available in FastJet, http://fastjet.fr

(As well as many IRC unsafe ones)